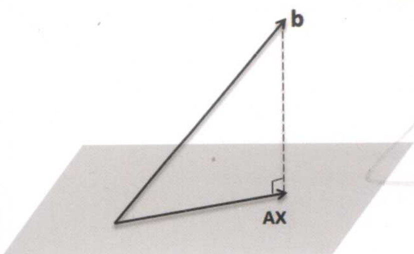
$A^TAx = A^Tb$



线性代数

Linear Algebra

张巍 阚海斌 倪卫明 编著

科学出版社

线性代数

张 巍　阚海斌　倪卫明　编著

科学出版社
北京

内 容 简 介

本书按照高等院校理工科各专业线性代数教学要求而编写，全书共 7 章，包括矩阵、线性方程组、行列式、线性空间与线性变换、特征值与二次型、矩阵分解、矩阵的Jordan标准形. 本书力求深入浅出，在介绍抽象的数学概念时会交代其来龙去脉，突出其实际应用背景. 书中有大量精选例题和习题，较难的印有 *，可以根据情况选用.

本书是高等学校理工科各专业线性代数教材，也可供工程技术类人员、研究生和教师参考.

图书在版编目(CIP)数据

线性代数 / 张巍，阚海斌，倪卫明编著. -- 北京 ：科学出版社，2016.7

ISBN 978-7-03-049409-2

Ⅰ. ①线… Ⅱ. ①张… ②阚… ③倪… Ⅲ. ①线性代数 Ⅳ. ①O151.2

中国版本图书馆 CIP 数据核字(2016)第 165496 号

责任编辑：王艳丽

责任印制：谭宏宇/封面设计：殷 靓

科 学 出 版 社 出版

北京东黄城根北街 16 号

邮政编码：100717

http：//www.sciencep.com

江苏省句容市排印厂 印刷

科学出版社发行 各地新华书店经销

*

2016 年 7 月第 一 版 开本：B5（720×1000）

2016 年 7 月第一次印刷 印张：14 1/4

字数：286 000

定价：20.00 元

（如有印装质量问题，我社负责调换）

前言

线性代数对于自然科学、技术科学及社会科学等各专业学生均是一门非常重要的基础课程,其核心内容包括矩阵理论、线性方程组理论、向量空间理论等,这些概念与理论为许多专业领域中的实际问题提供了表述精确的数学语言,并为分析问题与解决问题提供了强有力的数学工具.

本书几乎涵盖国内外已有线性代数课程的全部内容,比国内大多数教材的内容更加丰富,在很多重要知识点的阐述方面,讲解角度更加新颖、更加体系化.特别针对技术科学类学生,对工程技术专业中常用的线性代数的基本概念、基本理论和基本方法作了较深入的讨论和阐述,力求条理清楚、重点突出,定理证明简洁清晰,例题分析典型直观,使读者在较短时间内系统地掌握线性代数的基本内容.

第 1 章介绍矩阵.主要介绍了矩阵乘法、初等变换,以及秩的概念.从元素、行、列等不同角度讲解矩阵乘法,使读者充分理解为什么这样定义矩阵乘法.特别强调矩阵乘积 $\boldsymbol{AB}$ 的第 j 列是矩阵 $\boldsymbol{A}$ 的所有列的线性组合,组合系数为 $\boldsymbol{B}$ 的第 j 列各元素.从向量线性组合的角度讲解矩阵乘法,将矩阵与向量乘积 $\boldsymbol{Ax}$ 作为关于 $\boldsymbol{A}$ 的列的线性组合,使读者一开始就建立向量的线性组合这一重要概念.

第 2 章介绍线性方程组.首先从行的角度考察线性方程组,求解线性方程组等价于求所有超平面的公共点;其次从列的角度考察线性方程组,求解线性方程组等价于求向量的线性组合系数.从横纵两个角度讲解线性方程组有解的条件,而不是抽象地叙述理论,使读者更直观、更深刻掌握线性方程组理论,也加深了对第 1 章矩阵理论的理解,为后续向量空间的学习打下基础,并将线性方程组理论和向量空间思想联系在一起.在本章最后,基于线性方程组的理论系统地讲解了逆矩阵.

第 3 章介绍行列式.结合矩阵的初等变换讲解行列式的性质使读者温故而知新.在讲解伴随矩阵的基础上,引入行列式按行(列)展开的计算方法以及 Cramer 法则.

第 4 章介绍线性空间与线性变换.主要介绍了基、子空间、内积空间、线性变换等基本概念.从向量投影角度讲解 Schmidt 正交化,使读者深刻理解 Schmidt 正交化过程每一步的含义,避免了硬记公式、学过就忘.结合线性方程组理论,以向量正交投影为基础,引入线性方程组无解情况下求其最小二乘解的意义和方法.

第 5 章介绍特征值与二次型.介绍了特征值、特征向量、二次型、正定矩阵等概念,讨论了矩阵可对角化的条件,揭示矩阵可对角化的内在原因是每个特征值

的几何重数与代数重数相等. 结合二次型引入了 Rayleigh 商, 这个概念对于技术科学类学生后续专业课程的学习是很有用的.

第 6 章介绍矩阵分解. 矩阵分解在解决实际工程问题中起着很大的作用. 将矩阵表示为多个特殊矩阵的乘积, 实际上是对该矩阵中数据的预处理, 将其分解成多个部分. 本章着重介绍了最常用的矩阵 LU 分解、QR 分解、 Cholesky 分解、谱分解和奇异值分解.

第 7 章介绍矩阵的 Jordan 标准形. Jordan 标准形是初等矩阵理论一个核心内容, 我们采用比较简单的 λ-矩阵的方法, 介绍了 λ-矩阵的标准形、不变因子与初等因子、两个矩阵相似当且仅当它们的特征矩阵等价, 从而证明任何一个矩阵都相似于一个 Jordan 形矩阵. 然后, 把矩阵看成一个线性变换, 证明存在一组基, 使得这个线性变换在这组基下的矩阵是 Jordan 形矩阵.

全书编写由阚海斌老师负责统筹安排, 其中第 1, 2, 5, 6 章由张巍老师编写、第 3 章由张巍与阚海斌老师编写、第 4 章由张巍与倪卫明老师编写、第 7 章由阚海斌与倪卫明老师编写. 最后, 阚海斌老师对全书作了修改和审定. 复旦大学 2015 级技术科学试验班的同学们为每章习题的编写提供了不少帮助, 对此表示衷心的感谢.

限于编者的水平与经验, 不妥之处在所难免, 恳请各位专家、学者和读者提出宝贵意见.

编 者

2016年2月29日

目　　录

第 1 章 矩　　阵

1.1 矩阵与向量的概念

1.1.1 矩阵的基本概念

定义1.1.1 由若干个数排成的 m 行 n 列矩形阵列

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为 $m \times n$ **矩阵**(matrix). 习惯上用大写英文字母 $\boldsymbol{A}$ 或者 $[a_{ij}]$ 表示矩阵. 有时为了指明行数与列数, 也可记为 $\boldsymbol{A}_{m\times n}$ 或者 $[a_{ij}]_{m\times n}$. a_{ij} 称为矩阵的第 i 行第 j 列元素(entry), 在本书中有时也采用记号 entry$(\boldsymbol{A}, i, j)$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列元素.

设行数相等、列数相等的两个矩阵 $\boldsymbol{A} = [a_{ij}]_{m\times n}$ 与 $\boldsymbol{B} = [b_{ij}]_{m\times n}$, 若它们的所有元素对应相等, 即 $a_{ij} = b_{ij}$ $(i = 1, \cdots, m; j = 1, \cdots, n)$, 则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相等, 记作 $\boldsymbol{A} = \boldsymbol{B}$.

定义1.1.2 设 $m \times n$ 矩阵

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

若将矩阵 $\boldsymbol{A}$ 的行和列依次互换, 得到的 $n \times m$ 矩阵

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

称为矩阵 $\boldsymbol{A}$ 的**转置矩阵**(transpose matrix), 记作 $\boldsymbol{A}^{\mathrm{T}}$. 显然, $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.

根据定义, 矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列元素与其转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 行第 i 列元素相同, 即 entry$(\boldsymbol{A}, i, j) =$ entry$(\boldsymbol{A}^{\mathrm{T}}, j, i)$.

例1.1.1 矩阵

$$\boldsymbol{A} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} 2 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix}$$

的转置矩阵分别为

$$\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

$$\boldsymbol{B}^{\mathrm{T}} = \begin{bmatrix} 2 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}$$

若矩阵的所有元素均为0, 则该矩阵称为**零矩阵**(zero matrix), 记作 $\boldsymbol{O}$ 或者 $\boldsymbol{O}_{m\times n}$.

当 $m = n$ 时, 称 $\boldsymbol{A}_{m\times n}$ 为 n 阶**方阵**(square matrix); 该方阵从左上角至右下角的元素 $a_{11}, a_{22}, \cdots, a_{nn}$ 称为对角线元素; 方阵 $\boldsymbol{A}$ 的对角线元素之和 $\sum_{i=1}^{n} a_{ii}$ 称为该方阵的**迹**(trace), 记作 tr$(\boldsymbol{A})$.

定义1.1.3 若 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$, 则称 $\boldsymbol{A}$ 为**对称阵**(symmetric matrix).

若 n 阶方阵 $\boldsymbol{A}$ 的非对角线元素均为零, 即

$$\boldsymbol{A} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}$$

则称该方阵为**对角阵**(diagonal matrix), 记为 diag$(a_1, a_2, \cdots, a_n)$. 显然, 对角阵是特殊的对称阵. 特别地, 如果对角阵的对角线元素相等, 则该对角阵又称为**纯量阵**.

若方阵的所有对角线元素均为 1, 所有非对角线元素均为 0, 即

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

则称该方阵为**单位阵**(identity matrix), 记作 $\boldsymbol{I}$ 或 $\boldsymbol{I}_n$ (其中 n 表示阶数). 显然, 单位阵是特殊的对称阵、对角阵、纯量阵.

定义1.1.4 设 n 阶方阵 $\boldsymbol{A} = [a_{ij}]_{n\times n}$, 当 $1 \leqslant j < i \leqslant n$ 时, $a_{ij} = 0$, 称 $\boldsymbol{A}$ 为上三角阵, 如下列第 1 个矩阵所示; 当 $1 \leqslant i < j \leqslant n$ 时, $a_{ij} = 0$, 称 $\boldsymbol{A}$ 为下三角阵, 如下列第 2 个矩阵所示.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

上三角阵和下三角阵统称为**三角阵**(triangular matrix). 如果上(下)三角阵的主对角线元素均为0, 则称为严格上(下)三角阵.

1.1.2 向量的基本概念

定义1.1.5 由 n 个数构成的有序数组 $[a_1, a_2, \cdots, a_n]$ 称为 n 元向量或 n 维向量(vector), 其中 a_i 称为该 n 元向量的第 i 个元素, 习惯上用希腊字母 $\boldsymbol{\alpha}$ 或 $\boldsymbol{\beta}$ 等表示一个向量.

定义1.1.6 设两个 n 元向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$, 若它们的所有对应元素相等时, 称向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 相等, 记作 $\boldsymbol{\alpha} = \boldsymbol{\beta}$.

若向量的所有元素均为0, 则该向量称为**零向量**(zero vector), 记作 $\mathbf{0}$.

一个 n 元向量可写成一行也可写成一列, $\boldsymbol{\alpha} = [a_1, a_2, \cdots, a_n]$ 称为 n 元行向量(row vector),

$$\boldsymbol{\alpha} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

称为n元列向量 (column vector). n元行向量的转置是n元列向量, 反之亦然. 书写时为了节省空间, 通常将n元列向量写成n元行向量的转置$[a_1, a_2, \cdots, a_n]^{\mathrm{T}}$.

本书所讲的n元向量, 若不加声明一般均指n元列向量.

$n \times m$矩阵的每一行可视为一个m元行向量, 每一列可视为一个n元列向量. m元行向量可看作一个$1 \times m$矩阵, n元列向量可看作一个$n \times 1$矩阵.

1.2 矩阵与向量的运算

1.2.1 矩阵(向量)的线性运算

定义1.2.1 设矩阵$\boldsymbol{A} = [a_{ij}]_{m\times n}, \boldsymbol{B} = [b_{ij}]_{m\times n}, \boldsymbol{C} = [c_{ij}]_{m\times n}$, 若满足$c_{ij} = a_{ij} + b_{ij}\,(i = 1, \cdots, m; j = 1, \cdots, n)$, 则称矩阵$\boldsymbol{C}$为$\boldsymbol{A}$与$\boldsymbol{B}$的和, 记作$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$.

上述计算矩阵和的运算称为矩阵的加法.

两个矩阵只有在行数相等且列数相等时才可相加.

可以验证矩阵的加法满足:

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

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

(3) $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.

其中$\boldsymbol{A}, \boldsymbol{B}$为行数相等且列数相等的矩阵.

定义1.2.2 设矩阵$\boldsymbol{A} = [a_{ij}]_{m\times n}, \boldsymbol{B} = [b_{ij}]_{m\times n}, k$为某个数, 若满足$b_{ij} = ka_{ij}\,(i = 1, \cdots, m; j = 1, \cdots, n)$, 则称$\boldsymbol{B}$为数$k$与矩阵$\boldsymbol{A}$的数量乘积, 记作$\boldsymbol{B} = k\boldsymbol{A}$. 上述计算数与矩阵乘积的运算称为矩阵的数乘. 特别地, 若$k = -1$, 则$\boldsymbol{B} = -\boldsymbol{A}$, 称矩阵$\boldsymbol{B}$为$\boldsymbol{A}$的负矩阵.

数与任意一个矩阵相乘总是可行的.

设矩阵$\boldsymbol{A} = [a_{ij}]_{m\times n}, \boldsymbol{B} = [b_{ij}]_{m\times n}$, 矩阵的减法可定义为: $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B})$.

可以验证矩阵的数乘满足:

(1) $k(l\boldsymbol{A}) = (kl)\boldsymbol{A}$(结合律);

(2) $(k+l)\boldsymbol{A} = k\boldsymbol{A} + l\boldsymbol{A}$(分配律);

(3) $k(\boldsymbol{A}+\boldsymbol{B}) = k\boldsymbol{A} + k\boldsymbol{B}$(分配律);

(4) $(k\boldsymbol{A})^{\mathrm{T}} = k\boldsymbol{A}^{\mathrm{T}}$.

其中 k, l 为数, $\boldsymbol{A}, \boldsymbol{B}$ 为行数相等且列数相等的矩阵.

由于 n 元向量可看作是一个特殊的矩阵, 因此 n 元向量的加法、数乘运算与矩阵的加法、数乘运算定义一致. 通常将矩阵(向量)的加法、数乘运算统称为矩阵(向量)的线性运算.

定义1.2.3 设一组 n 元向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$, 若存在一组数 $k_1, k_2, \cdots, k_m$, 使得 $\boldsymbol{\beta} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m$, 则称向量 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的线性组合(linear combination), 或称 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表出, $k_1, k_2, \cdots, k_m$ 称为线性组合系数或线性表出系数.

显然, n 元零向量 $\mathbf{0}$ 可由任意一组 n 元向量线性表出.

1.2.2 向量的内积与矩阵的乘法

定义1.2.4 设两个 n 元向量 $\boldsymbol{\alpha} = [a_1, a_2, \cdots, a_n]$ 与 $\boldsymbol{\beta} = [b_1, b_2, \cdots, b_n]$, 向量 $\boldsymbol{\alpha}$ 的元素与向量 $\boldsymbol{\beta}$ 的对应元素乘积之和 $\sum_{i=1}^{n} a_i b_i$ 称为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积, 记作 $\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle$.

注 上述定义中对向量是行向量还是列向量不作限制, 即 $\boldsymbol{\alpha}$ 可以是行向量也可以是列向量, 同样, $\boldsymbol{\beta}$ 可以是行向量也可以是列向量.

定义1.2.5 设矩阵 $\boldsymbol{A} = [a_{ik}]_{m\times l}, \boldsymbol{B} = [b_{kj}]_{l\times n}, \boldsymbol{C} = [c_{ij}]_{m\times n}$, 若满足

$$c_{ij} = \sum_{k=1}^{l} a_{ik} b_{kj} \quad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$$

则称矩阵 $\boldsymbol{C}$ 为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积, 记作 $\boldsymbol{C} = \boldsymbol{AB}$. 上述计算矩阵乘积的运算称为矩阵的乘法.

仅当矩阵 $\boldsymbol{A}$ 的列数等于矩阵 $\boldsymbol{B}$ 的行数时, 矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 方可相乘; 乘积 $\boldsymbol{AB}$ 也是一个矩阵, 其行数等于矩阵 $\boldsymbol{A}$ 的行数, 其列数等于矩阵 $\boldsymbol{B}$ 的列数; 乘积矩阵 $\boldsymbol{AB}$ 的第 i 行第 j 列元素等于 $\boldsymbol{A}$ 的第 i 行向量与 $\boldsymbol{B}$ 的第 j 列向量的内积.

为了方便起见, 采用记号 $\mathrm{entry}(\boldsymbol{AB}, i, j)$ 表示乘积矩阵 $\boldsymbol{AB}$ 的第 i 行第 j 列元素, $\mathrm{row}(\boldsymbol{A}, i)$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行元素组成的行向量, $\mathrm{col}(\boldsymbol{B}, j)$ 表示矩阵 $\boldsymbol{B}$ 的

第 j 列元素组成的列向量, 则上述矩阵乘法定义可表述为

$$\text{entry}(\boldsymbol{AB}, i, j) = \langle \text{row}(\boldsymbol{A}, i), \text{col}(\boldsymbol{B}, j) \rangle$$

由于 l 元行向量 $\text{row}(\boldsymbol{A}, i)$ 和 l 元列向量 $\text{col}(\boldsymbol{B}, j)$ 分别可以看作是 $1 \times l$ 矩阵和 $l \times 1$ 矩阵, 根据矩阵乘法的定义, $\text{row}(\boldsymbol{A}, i)$ 与 $\text{col}(\boldsymbol{B}, j)$ 的乘积等于它们的内积 $\text{row}(\boldsymbol{A}, i)\text{col}(\boldsymbol{B}, j) = \langle \text{row}(\boldsymbol{A}, i), \text{col}(\boldsymbol{B}, j) \rangle$, 因此上式还可写为

(1) $\text{entry}(\boldsymbol{AB}, i, j) = \text{row}(\boldsymbol{A}, i)\text{col}(\boldsymbol{B}, j)$.

进一步, 不难验证以下有用的等式:

(2) $\text{row}(\boldsymbol{AB}, i) = \text{row}(\boldsymbol{A}, i)\boldsymbol{B} = \sum_{k=1}^{l} a_{ik}\text{row}(\boldsymbol{B}, k)$.

(3) $\text{col}(\boldsymbol{AB}, j) = \boldsymbol{A}\text{col}(\boldsymbol{B}, j) = \sum_{k=1}^{l} b_{kj}\text{col}(\boldsymbol{A}, k)$.

(2) 式表明矩阵乘积 $\boldsymbol{AB}$ 的第 i 行向量是矩阵 $\boldsymbol{B}$ 的所有行向量的线性组合, 组合系数为 $\boldsymbol{A}$ 的第 i 行各元素, 即 $\text{row}(\boldsymbol{AB}, i)$ 只与矩阵 $\boldsymbol{A}$ 的第 i 行有关, 与 $\boldsymbol{A}$ 的其余行无关; (3) 式表明矩阵乘积 $\boldsymbol{AB}$ 的第 j 列向量是矩阵 $\boldsymbol{A}$ 的所有列向量的线性组合, 组合系数为 $\boldsymbol{B}$ 的第 j 列各元素, 即 $\text{col}(\boldsymbol{AB}, j)$ 只与矩阵 $\boldsymbol{B}$ 的第 j 列有关, 与 $\boldsymbol{B}$ 的其余列无关.

可以看出, (1) 式提供了逐个元素求矩阵乘积的算法, 是最细粒度的计算方式; (2) 式和(3) 式分别提供了逐行和逐列求矩阵乘积的算法, 是较粗粒度的计算方式. 在矩阵乘法的实际计算和问题分析中, (2)式和(3)式有时能使我们对矩阵乘积的计算更加便捷、认识更加清晰.

更进一步, 还可验证以下等式:

(4) $\boldsymbol{AB} = \sum_{k=1}^{l} \text{col}(\boldsymbol{A}, k)\text{row}(\boldsymbol{B}, k)$.

从(4) 式看出, 矩阵乘积 $\boldsymbol{AB}$ 等于若干个 $m \times n$ 矩阵 $\text{col}(\boldsymbol{A}, k)\text{row}(\boldsymbol{B}, k)$ 的和.

例1.2.1 设

$$\boldsymbol{A} = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 2 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}$$

求 $\boldsymbol{AB}$ 的第 2 行以及第 1 列.

解 $\boldsymbol{AB}$ 的第 2 行

$$\begin{aligned}\mathrm{row}(\boldsymbol{AB},2) &= \mathrm{row}(\boldsymbol{A},2)\boldsymbol{B}\\ &= 0\times\mathrm{row}(\boldsymbol{B},1)+1\times\mathrm{row}(\boldsymbol{B},2)+0\times\mathrm{row}(\boldsymbol{B},3)\\ &= \mathrm{row}(\boldsymbol{B},2)=[1,3]\end{aligned}$$

$\boldsymbol{AB}$ 的第 1 列

$$\begin{aligned}\mathrm{col}(\boldsymbol{AB},1) &= \boldsymbol{A}\mathrm{col}(\boldsymbol{B},1)\\ &= 2\times\begin{bmatrix}-1\\0\\1\end{bmatrix}+1\times\begin{bmatrix}0\\1\\-1\end{bmatrix}+0\times\begin{bmatrix}2\\0\\1\end{bmatrix}=\begin{bmatrix}-2\\1\\1\end{bmatrix}\end{aligned}$$

例1.2.2 设

$$\boldsymbol{A}=\begin{bmatrix}1&-1\\-1&1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}2&0\\0&-2\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}1&1\\-1&-1\end{bmatrix}$$

求 $\boldsymbol{AB},\boldsymbol{AC},\boldsymbol{CA}$.

解

$$\boldsymbol{AB}=\begin{bmatrix}1&-1\\-1&1\end{bmatrix}\begin{bmatrix}2&0\\0&-2\end{bmatrix}=\begin{bmatrix}2&2\\-2&-2\end{bmatrix}$$

$$\boldsymbol{AC}=\begin{bmatrix}1&-1\\-1&1\end{bmatrix}\begin{bmatrix}1&1\\-1&-1\end{bmatrix}=\begin{bmatrix}2&2\\-2&-2\end{bmatrix}$$

$$\boldsymbol{CA}=\begin{bmatrix}1&1\\-1&-1\end{bmatrix}\begin{bmatrix}1&-1\\-1&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$

从例1.2.2可以看出:

(1) 矩阵乘法的交换律不成立. 首先, 矩阵 $\boldsymbol{A}$ 可以左乘矩阵 $\boldsymbol{C}$ 并不意味着可以右乘 $\boldsymbol{C}$, 可能 $\boldsymbol{AC}$ 与 $\boldsymbol{CA}$ 并非同时有意义; 其次, 即使同时有意义, 也不一定相等. 在矩阵相乘时, 必须考虑是否有意义, 而且说明是左乘还是右乘.

(2) 已知 $\boldsymbol{AC}=\boldsymbol{O}$, 无法推出 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{C}=\boldsymbol{O}$; 已知 $\boldsymbol{AC}=\boldsymbol{O},\boldsymbol{A}\neq\boldsymbol{O}$, 无法推出 $\boldsymbol{C}=\boldsymbol{O}$.

(3) 已知 $\boldsymbol{AB}=\boldsymbol{AC}$, $\boldsymbol{A}\neq\boldsymbol{O}$, 无法推出 $\boldsymbol{B}=\boldsymbol{C}$. 这是因为 $\boldsymbol{AB}=\boldsymbol{AC}$, 即 $\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{C})=\boldsymbol{O}$, 尽管 $\boldsymbol{A}\neq\boldsymbol{O}$, 但由 (2) 知, 不能推出 $\boldsymbol{B}-\boldsymbol{C}=\boldsymbol{O}$.

可以验证矩阵的乘法满足:

(1) $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$ (结合律);

(2) $\boldsymbol{C}(\boldsymbol{A}+\boldsymbol{B})=\boldsymbol{CA}+\boldsymbol{CB}$ (分配律);

(3) $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC}$ (分配律);

(4) $k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$.

其中 k 为数, 矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 的阶数使上述各式有意义.

性质1.2.1 设矩阵 $\boldsymbol{A}_{m\times l},\boldsymbol{B}_{l\times n}$, 则 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

证明

$$\mathrm{entry}((\boldsymbol{AB})^{\mathrm{T}},i,j)=\mathrm{entry}(\boldsymbol{AB},j,i)=\mathrm{row}(\boldsymbol{A},j)\mathrm{col}(\boldsymbol{B},i)=\sum_{k=1}^{l}a_{jk}b_{ki}$$

而

$$\begin{aligned}\mathrm{entry}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}},i,j)&=\mathrm{row}(\boldsymbol{B}^{\mathrm{T}},i)\mathrm{col}(\boldsymbol{A}^{\mathrm{T}},j)=\mathrm{col}(\boldsymbol{B},i)^{\mathrm{T}}\mathrm{row}(\boldsymbol{A},j)^{\mathrm{T}}\\&=\sum_{k=1}^{l}b_{ki}a_{jk}\end{aligned}$$

故 $\mathrm{entry}((\boldsymbol{AB})^{\mathrm{T}},i,j)=\mathrm{entry}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}},i,j)$, 从而 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$. □

性质1.2.2 设矩阵 $\boldsymbol{A}_{m\times n},\boldsymbol{B}_{n\times m}$, 则 $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$.

证明

$$\mathrm{tr}(\boldsymbol{AB})=\sum_{i=1}^{m}\mathrm{entry}(\boldsymbol{AB},i,i)=\sum_{i=1}^{m}\mathrm{row}(\boldsymbol{A},i)\mathrm{col}(\boldsymbol{B},i)=\sum_{i=1}^{m}\sum_{k=1}^{n}a_{ik}b_{ki}$$

$$\mathrm{tr}(\boldsymbol{AB})=\sum_{k=1}^{n}\mathrm{entry}(\boldsymbol{AB},k,k)=\sum_{k=1}^{n}\mathrm{row}(\boldsymbol{B},k)\mathrm{col}(\boldsymbol{A},k)=\sum_{k=1}^{n}\sum_{i=1}^{m}b_{ki}a_{ik}$$

故 $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$. □

例1.2.3 设

$$\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}},\quad \boldsymbol{\beta}=[b_1,b_2,\cdots,b_n]$$

则