

矩阵理论复习笔记

前言

本复习笔记是我个人在学习矩阵理论的过程中整理、总结而成，包含课本的内容、上课 PPT 涉及到的内容以及不懂地方的补充知识，希望能够对你有所帮助。

全书排版是利用 \LaTeX 完成的，这也是对我使用 \LaTeX 的一次较大工程的练手，希望我能在撰写完之后对于 \LaTeX 使用有更深层次的理解。

因本人水平有限，故本总结笔记如有不当之处，敬请指出，本人不胜感激！

尹彦江

2024 年 5 月 26 日

目录

第零章 线性代数基本知识	1
0.1 行列式	2
0.1.1 行列式的定义与计算方法	2
0.1.2 行列式的性质	6
0.2 矩阵	8
0.2.1 矩阵的运算性质	8
0.2.2 矩阵的逆	9
0.2.3 矩阵的伴随	10
0.2.4 初等矩阵与初等变换	10
0.2.5 矩阵的秩	12
0.3 矩阵的秩、向量组、线性方程组	13
0.3.1 向量以及向量组	13
0.3.2 方程组	15
0.4 特征值与特征向量、相似理论	17
0.4.1 特征值与特征向量的定义	18
0.4.2 求解特征值与特征向量的步骤	18
0.4.3 特征值、特征向量的性质与重要结论	18
0.4.4 矩阵的相似理论	19
0.5 结语	20
第一章 线性代数基础	21
1.1 线性空间与子空间	22

第零章 线性代数基本知识

本章主要针对矩阵理论中所涉及到的线性代数知识进行复习与回顾,随着课程的进行,当认为有必要需要补充相关的线性代数基本知识时,会补充在本节。

本章会涉及以下内容:

- 行列式
- 矩阵
- 向量组、线性方程组
- 特征值与特征向量, 矩阵相似理论

注意: 本章只作为一个对于基本知识的回顾与复习,并不会涉及到更详细的解释,如果需要查看更详细的解释,请翻阅对应的线性代数教材或上网搜集资料。¹

¹本章的大多数内容是在《2025 张宇考研数学基础 30 讲——线性代数分册》的基础上整理总(抄)结(写)而成

0.1 行列式

行列式作为后续矩阵理论中涉及特征值与特征向量计算的基础，因此有必要复习巩固行列式的定义、行列式的相关性质以及具体行列式的计算。

0.1.1 行列式的定义与计算方法

何为行列式？如果查阅线性代数的教材（这里以同济六版《线性代数》为例），会发现教材是从特例到一般进行推广，首先给出二阶行列式和三阶行列式应如何计算，后面给出全排列的逆序数的概念，最后推广至 n 阶行列式得出 n 阶行列式的定义的，而在介绍二阶行列式的时候，是先从解一个二元线性方程组得出的：

§1 二阶与三阶行列式

一、二元线性方程组与二阶行列式

用消元法解二元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases} \quad (1)$$

为消去未知数 x_2 ，以 a_{22} 与 a_{12} 分别乘上列方程的两端，然后两个方程相减，得

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - a_{12}b_2;$$

类似地，消去 x_1 ，得

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - b_1a_{21}.$$

当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时，求得方程组(1)的解为

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}. \quad (2)$$

(2)式中的分子、分母都是四个数分两对相乘再相减而得，其中分母 $a_{11}a_{22} - a_{12}a_{21}$ 是由方程组(1)的四个系数确定的，把这四个数按它们在方程组(1)中的位置，排成二行二列（横排称行、竖排称列）的数表

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad (3)$$

表达式 $a_{11}a_{22} - a_{12}a_{21}$ 称为数表(3)所确定的二阶行列式，并记作

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}. \quad (4)$$

数 a_{ij} ($i=1,2; j=1,2$) 称为行列式(4)的元素或元。元素 a_{ij} 的第一个下标 i 称为行标，表明该元素位于第 i 行；第二个下标 j 称为列标，表明该元素位于第 j

· 1 ·

二、三阶行列式

定义 1 设有 9 个数排成 3 行 3 列的数表

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad (5)$$

记

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}, \quad (6)$$

(6)式称为数表(5)所确定的三阶行列式。

上述定义表明三阶行列式含 6 项，每项均为不同行不同列的三个元素的乘积再冠以正负号，其规律遵循图 1.2 所示的对角线法则：图中有三条实线看做是平行于主对角线的连线，三条虚线看做是平行于副对角线的连线，实线上三元素的乘积冠正号，虚线上三元素的乘积冠负号。

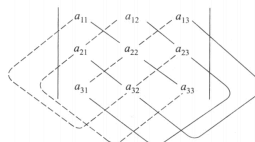


图 1.2

(a) 教材中对于二阶行列式的介绍

(b) 教材中对于三阶行列式的介绍

图 1: 《线性代数》教材中对于二阶、三阶行列式的介绍

这样的定义是从代数的角度出发的，可能对于行列式的理解并不是很清楚，接下来会给出除课本之外的另一种理解行列式的方法。

行列式的第一种定义

假设我们有如下的一个二阶行列式：

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

我们将该行列式的第一行看作一个向量，即 $\mathbf{a}_1 = [a_{11}, a_{12}]$ ，将第二行看作另外一个向量，即 $\mathbf{a}_2 = [a_{21}, a_{22}]$ ，将这两个向量放在平面直角坐标系中，并以这两条边构建成一个平行四边形，如下图所示：

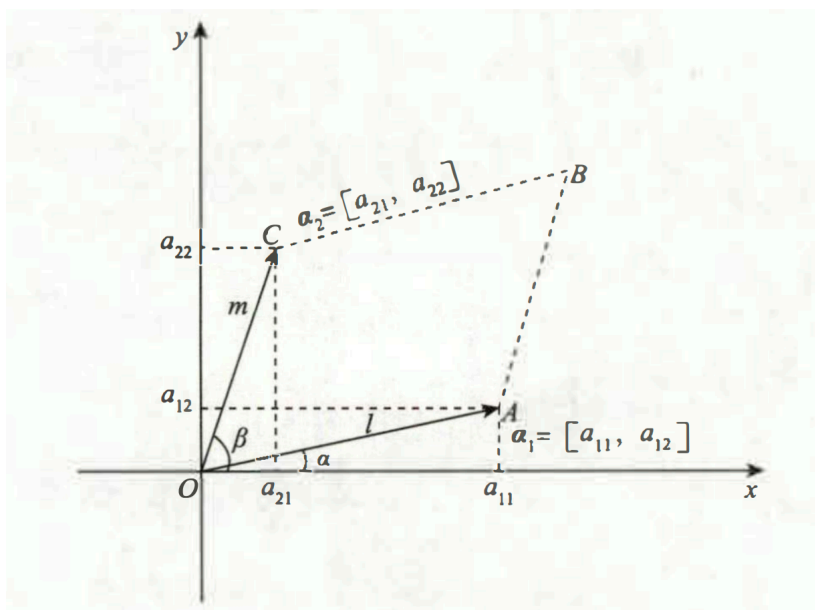


图 2: 将两个向量放在平面直角坐标系中表示的效果图

那么，这个平行四边形的面积是多少呢？我们设 \mathbf{a}_1 与 x 轴的夹角为 α ， \mathbf{a}_2 与 x 轴的夹角为 β ，设 $\|\mathbf{a}_1\| = l$ ， $\|\mathbf{a}_2\| = m$ 。

则，平行四边形的面积为：

$$\begin{aligned} S &= l \cdot m \cdot \sin(\beta - \alpha) \\ &= l \cdot m (\sin \beta \cos \alpha - \cos \beta \sin \alpha) \\ &= l \cos \alpha \cdot m \sin \beta - l \sin \alpha \cdot m \cos \beta \\ &= a_{11}a_{22} - a_{12}a_{21} \end{aligned}$$

这是不是与二阶行列式的结果一样呢？所以我们可以得出一个结论：二阶行列式可以表示为一个平行四边形的面积。

接下来我们讨论三阶行列式，与二阶行列式相似，假设我们有如下三阶行列式：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

其实，三阶行列式的几何意义即为是一个由三个三维向量 $\mathbf{a}_1 = [a_{11}, a_{12}, a_{13}]$, $\mathbf{a}_2 = [a_{21}, a_{22}, a_{23}]$, $\mathbf{a}_3 = [a_{31}, a_{32}, a_{33}]$ 组成的一个平行六面体的体积，如下图：

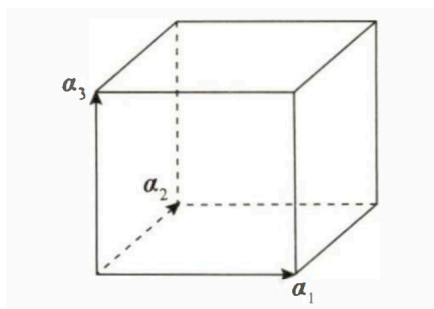


图 3: 三阶行列式的几何意义示意

再进行推广，我们可以得到 n 阶行列式的定义：

定义 0.1.1

一个 n 阶行列式

$$D_n = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

其是由 n 个 n 维向量：

$$\mathbf{a}_1 = [a_{11}, a_{12}, \cdots, a_{1n}]$$

$$\mathbf{a}_2 = [a_{21}, a_{22}, \cdots, a_{2n}]$$

\cdots

$$\mathbf{a}_n = [a_{n1}, a_{n2}, \cdots, a_{nn}]$$

组成的，其结果是以这 n 个向量为边的 n 维图形的“体积”。

这里可以结合后面线性相关和线性无关的性质：如果一个方阵的行列式为 0，就说这个向量组是线性相关的，如果用行列式的几何意义理解将会变得十分简单，因为如果一个向量组线性相关，就说明这个向量组至少有一个向量是与其他向量共线的，如果共线，就说明这个 n 维图形将会少一条边，少一条边“体积”自然为 0。

行列式的第二种定义

后面行列式的第二种与第三种定义与教材中对于行列式的定义相同，第二种行列式引入了“逆序数”这一概念来辅助定义行列式，即一个 n 阶行列式：

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix} = \sum_{j_1 j_2 \cdots j_n} (-1)^{r(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

这里 $r(j_1 j_2 \cdots j_n)$ 表示的是数列 j_1, j_2, \cdots, j_n 的逆序数²。

注. 如何理解上面的公式：计算一个 n 阶行列式只需要找出其所有的行和所有的列的展开乘以对应的系数（-1 的正负），最后将所有的和加起来即为行列式的值。

行列式的第三种定义

这里介绍行列式的第三种定义——代数余子式与余子式形式，这一种定义也是计算行列式的常用方法。

首先，回顾一下什么是余子式，什么是代数余子式：

- 余子式：将一个 n 阶行列式，去掉元素 a_{ij} 所在的行和列（即第 i 行和第 j 列）之后得到的新的行列式，记为 M_{ij} 。
- 代数余子式：余子式乘以 $(-1)^{i+j}$ 记为代数余子式，记为 A_{ij} 。

根据上面的介绍，不难得出下列关系：

$$A_{ij} = (-1)^{i+j} M_{ij}$$

由此，可以得出行列式的第三种定义：

定义 0.1.2

可以按照某一行某一列来展开行列式，即：

$$|A| = \begin{cases} a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^n a_{ij}A_{ij} & (i = 1, 2, \cdots, n) \rightarrow (\text{按行展开}) \\ a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^n a_{ij}A_{ij} & (j = 1, 2, \cdots, n) \rightarrow (\text{按列展开}) \end{cases}$$

²关于逆序数的详细解释，请参阅：<https://zhuanlan.zhihu.com/p/398845457>

注. 行列式的某一行或者某一列乘以其他非对应行或非对应列的代数余子式后再求和, 其结果为 0。

例 0.1.1. 试计算行列式

$$D = \begin{vmatrix} 2 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 2 & -1 \\ -1 & 0 & 0 & 2 \end{vmatrix}$$

解. 计算行列式的值, 按照定义三来解题的话, 尽可能找“0”多的行或者列进行展开, 这样可以简单计算。

将行列式按第一列展开, 得

$$\begin{aligned} D &= 2 \begin{vmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{vmatrix} - 1 \times (-1)^{(4+1)} \begin{vmatrix} -1 & 0 & 0 \\ 2 & -1 & 0 \\ 0 & 2 & -1 \end{vmatrix} \\ &= 16 - 1 = 15 \end{aligned}$$

□

0.1.2 行列式的性质

只知道了行列式的计算还不够, 运用好行列式的性质可以在处理复杂行列式的时候及时发现其规律对其进行处理, 从而达到简便计算的效果。

行列式有下列性质:

- 行列互换, 其值不变, 即 $|\mathbf{A}| = |\mathbf{A}^T|$
- 若行列式中的某一行或者某一列的元素全为 0, 则行列式为 0
- 若行列式中某一行或者某一列有公因子 $k(k \neq 0)$, 则可以将该公因子提到行列式外

面, 即:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{11} & ka_{12} & \cdots & ka_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} ka_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

注.

1. 这叫做行列式的“倍乘”性质
 2. 请注意在行列式中可以提出公因子和矩阵里可以提出公因子的区别, 在矩阵中需要要求所有元素都含有公因子才可以提出来。
- 行列式中某一行或者某一列的元素均是某两个数之和, 则可以将行列式拆成两个行列式之和, 即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \cdots & a_{in} + b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

注. 请仔细留意该性质成立的条件

- 行列式中两行或两列互换, 行列式变号
- 行列式中两行或两列的元素对应成比例, 则行列式结果为 0
- 行列式中的某一行或某一列的 k 倍加到另外一行或另外一列, 行列式不变

注. 这也叫行列式的倍加性质

注. 以上的性质可以结合行列式的几何意义辅助理解

例 0.1.2. 设关于 λ 的方程

$$\begin{vmatrix} \lambda - 1 & -2 & 3 \\ 1 & \lambda - 4 & 3 \\ -1 & a & \lambda - 5 \end{vmatrix} = 0$$

有二重根, 求 a 的值

解. 将行列式的第一行与第二行相减, 得

$$\begin{vmatrix} \lambda - 2 & 2 - \lambda & 0 \\ 1 & \lambda - 4 & 3 \\ -1 & a & \lambda - 5 \end{vmatrix}$$

将行列式的第一列加到第二列, 得

$$\begin{vmatrix} \lambda - 2 & 0 & 0 \\ 1 & \lambda - 3 & 3 \\ -1 & a - 1 & \lambda - 5 \end{vmatrix}$$

按第一行展开, 有

$$\begin{aligned} (\lambda - 2) \begin{vmatrix} \lambda - 3 & 3 \\ a - 1 & \lambda - 5 \end{vmatrix} &= (\lambda - 2)[(\lambda - 3)(\lambda - 5) - 3(a - 1)] \\ &= (\lambda - 2)(\lambda^2 - 8\lambda + 18 - 3a) = 0 \end{aligned}$$

若想让函数有一个二重根, 则需要满足 $\lambda^2 - 8\lambda + 18 - 3a$ 有二重根。

由求根公式得, $\Delta = -\frac{b^2 - 4ac}{2a} = (-8)^2 - 4(18 - 3a) = 0$

解得 $a = \frac{2}{3}$

□

0.2 矩阵

在介绍了行列式之后, 本小节主要介绍矩阵的运算性质、矩阵的逆、矩阵的秩、一些特殊的矩阵, 以及矩阵的初等变换等内容。

0.2.1 矩阵的运算性质

矩阵的运算性质如下:

- 矩阵的加法: 对应项元素相加即可
- 矩阵的乘法: 左边矩阵的每一行乘以右边矩阵的每一列, 将求得和写在对应位置, 请注意矩阵乘法需要满足左侧矩阵的列与右侧矩阵的行相等才能进行

$$\bullet \text{ 数乘矩阵: } k\mathbf{A} = \mathbf{A}k = k \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix}$$

注. 注意这里与前面行列式倍乘性质的区别

- 矩阵的加法满足交换律、结合律, 矩阵的数乘满足分配律、交换律和结合律, 但矩阵之间的乘法只满足结合律和分配律, 一般不满足交换律
- 转置矩阵满足下面的运算律
 1. $(\mathbf{A}^T)^T = \mathbf{A}$
 2. $(k\mathbf{A})^T = k\mathbf{A}^T$
 3. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
 4. $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T \rightarrow \text{“穿脱原则”}$

0.2.2 矩阵的逆

定义 0.2.1: 逆矩阵定义

设 \mathbf{A}, \mathbf{B} 是 n 阶方阵, \mathbf{E} 是 n 阶单位矩阵, 若 $\mathbf{AB} = \mathbf{BA} = \mathbf{E}$, 则称 \mathbf{A} 是可逆矩阵, \mathbf{B} 是 \mathbf{A} 的逆矩阵, 记为 \mathbf{A}^{-1} 。

注. \mathbf{A} 矩阵可逆的充要条件是 $|\mathbf{A}| \neq 0$ 。

逆矩阵的性质与重要公式

设 \mathbf{A}, \mathbf{B} 为同阶可逆矩阵, 则

1. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
2. 若 $k \neq 0$, 则 $(k\mathbf{A})^{-1} = \frac{1}{k}\mathbf{A}^{-1}$
3. \mathbf{AB} 也可逆, 则 $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
4. \mathbf{A}^T 也可逆, 且 $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$
5. $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$

注. $\mathbf{A} + \mathbf{B}$ 不一定可逆, 且 $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$

0.2.3 矩阵的伴随

伴随矩阵的定义

定义 0.2.2

$$A^* = \begin{vmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{vmatrix}$$

其中 A_{ij} 代表代数余子式。

伴随矩阵的性质与重要公式

伴随矩阵具有下列的性质：

- 对任意 n 阶方阵 A ，都有伴随矩阵 A^* ，且有公式

$$AA^* = A^*A = |A|E$$

$$|A|^* = |A|^{n-1}$$

当 $|A| \neq 0$ 的时候，有

- $A^* = |A|A^{-1}$
- $A^{-1} = \frac{1}{|A|}A^*$
- $A = |A|(A^*)^{-1}$
- $(kA)(kA)^* = |kA|E$
- $A^T(A^T)^* = |A^T|E$
- $A^{-1}(A^{-1})^* = |A^{-1}|E$
- $A^*(A^*)^* = |A^*|E$
- $(A^T)^* = (A^*)^T$
- $(A^{-1})^* = (A^*)^{-1}$
- $(AB)^* = B^*A^*$
- $(A^*)^* = |A|^{n-2}A$

注. $(A + B^*) \neq A^* + B^*$

0.2.4 初等矩阵与初等变换

初等变换

有如下三种初等变换：

1. 一个非零常数乘矩阵的某一行或某一列 \rightarrow 倍乘性质
2. 矩阵中的某两行或两列互换位置 \rightarrow 互换性质
3. 将某一行或某一列的 k 倍加到另外一行或另外一列中 \rightarrow 倍加性质

以上三种变换叫做**初等行变换/初等列变换**

初等矩阵

由单位矩阵经过一次初等变换得到的矩阵叫做初等矩阵以三阶矩阵为例，有如下三种初等矩阵：

$$1. E_2(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{E} \text{ 的第 2 行或第 2 列乘 } k \text{ 倍, 称为倍乘初等矩阵}$$

$$2. E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{E} \text{ 的第 1,2 行或第 1,2 列互换, 称为互换初等矩阵}$$

$$3. E_{31}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}, \mathbf{E} \text{ 的第 1 行的 } k \text{ 倍加到第 3 行, 或第 3 列的 } k \text{ 倍加到第 1 列, 称为倍加初等矩阵}$$

行阶梯型矩阵和行最简阶梯型矩阵

具有如下特性的矩阵称为**行阶梯形矩阵**：

- 若有零行，必位于非零行的下方
- 各个非零行左起第一个非零元素的列标从上到下是严格增大的

当每一非零行的第一个元素为 1，并且这些非零元素所在列的其他元素均为 0，则这个矩阵就变成了**最简行阶梯型矩阵**。

用初等变换求逆矩阵

$$\left[\mathbf{A} : \mathbf{E} \right] \xrightarrow{\text{初等行变换}} \left[\mathbf{E} : \mathbf{A}^{-1} \right]$$

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{E} \end{bmatrix} \xrightarrow{\text{初等列变换}} \begin{bmatrix} \mathbf{E} \\ \mathbf{A}^{-1} \end{bmatrix}$$

0.2.5 矩阵的秩

矩阵的秩作为后续解线性方程组、向量组的关键性质，需要认真把握

定义 0.2.3: 矩阵的秩的定义

设 \mathbf{A} 是 $m \times n$ 阶矩阵，若存在 k 阶子式（即，任取 k 行与 k 列构成的行列式）不为零，而任意 $k+1$ 阶子式（如果有的话）全为 0，则 $r(\mathbf{A}) = k$ ，并且若 \mathbf{A} 为 $n \times n$ 阶矩阵，则

$$r(\mathbf{A})_{n \times n} = n \Leftrightarrow |\mathbf{A}| \neq 0 \Leftrightarrow \mathbf{A} \text{ 可逆}$$

如果用定义求矩阵的秩较为麻烦，可以将行列式化成行阶梯型矩阵，其非零行的行数即为 \mathbf{A} 的秩。

矩阵的秩的几个重要式子

设 \mathbf{A} 是 $m \times n$ 阶矩阵， \mathbf{B} 是满足有关矩阵运算要求的矩阵，则：

- $0 \leq r(\mathbf{A}) \leq \min \{m, n\}$
- $r(k\mathbf{A}) = r(\mathbf{A}) (k \neq 0)$
- $r(\mathbf{AB}) \leq \min \{r(\mathbf{A}), r(\mathbf{B})\}$
- $r(\mathbf{A} + \mathbf{B}) \leq r(\mathbf{A}) + r(\mathbf{B})$
- $r(\mathbf{A}^*) = \begin{cases} n, & r(\mathbf{A}) = n \\ 1, & r(\mathbf{A}) = n - 1 \\ 0, & r(\mathbf{A}) < n - 1 \end{cases}$ 其中， \mathbf{A} 为 $n(n \geq 2)$ 阶方阵

- 设 A 为 $m \times n$ 阶矩阵, P, Q 分别是 m, n 阶可逆矩阵, 则

$$r(A) = r(PA) = r(AQ) = r(PAQ)$$

- 若 $A_{m \times n} B_{n \times s} = O$, 则 $r(A) + r(B) \leq n$
- $r(A) = r(A^T) = r(AA^T) = r(A^T A)$

关于矩阵的秩后续更详细介绍, 请看向量组、线性方程组部分。

0.3 矩阵的秩、向量组、线性方程组

向量组与线性方程组的联系十分紧密, 可以说这两个概念分别一个是从向量角度讨论, 一个是在方程组的角度讨论, 其本质是相同的, 而在其中起重要作用的因子就是上一节最后介绍的矩阵的秩。

0.3.1 向量以及向量组

在本小节中, 重点回顾向量组的线性相关、线性无关的概念及其相关性质。

线性相关和线性无关

定义 0.3.1: 线性相关

对于 m 个 n 维向量 a_1, a_2, \dots, a_m , 若存在一组不全为零的数 k_1, k_2, \dots, k_m , 使得

$$k_1 a_1 + k_2 a_2 + \dots + k_m a_m = 0$$

则称向量组 a_1, a_2, \dots, a_m 线性相关。

定义 0.3.2: 线性无关

若不存在不全为零的数 k_1, k_2, \dots, k_m , 使得

$$k_1 a_1 + k_2 a_2 + \dots + k_m a_m = 0$$

则称向量组 a_1, a_2, \dots, a_m 线性无关, 即, 当且仅当 $k_1 = k_2 = \dots = k_n = 0$ 的时候, 才能使等式成立。

判别线性相关性的几大定理

1. 向量组 $a_1, a_2, \dots, a_n (n \geq 2)$ 线性相关的充要条件是向量组中至少有一个向量可以有其他的 $n-1$ 个向量表示

注. 其逆否命题为: 向量组 $a_1, a_2, \dots, a_n (n \geq 2)$ 线性无关的充要条件是向量组中的任何一个向量都不能由其余的 $n-1$ 个向量表示

2. 若向量组 a_1, a_2, \dots, a_n 线性无关, 而 $\beta, a_1, a_2, \dots, a_m$ 线性相关, 则 β 可以由 a_1, a_2, \dots, a_n 线性表示, 且表示法唯一
3. 如果向量组 $\beta_1, \beta_2, \dots, \beta_t$ 可以由向量组 a_1, a_2, \dots, a_s 线性表示, 且 $t > s$, 则 $\beta_1, \beta_2, \dots, \beta_t$ 线性相关 (以少表多, 多的相关)
4. 如果向量组 a_1, a_2, \dots, a_m 有一部分向量线性相关, 则整个向量组也线性相关

注. 其逆否命题为: 如果向量组 a_1, a_2, \dots, a_m 线性无关, 则其任意一部分向量组都线性无关

5. 如果一组 n 维向量 a_1, a_2, \dots, a_s 线性无关, 那么把这些向量各任意添加 m 个分量得到的新向量 ($n+m$ 维) 组 $a_1^*, a_2^*, \dots, a_s^*$ 也是线性无关的

注. 其逆否命题为: 如果向量组 a_1, a_2, \dots, a_s 线性相关, 那么他们各去掉相同的若干分量所得到的新向量组也是线性相关的

极大线性无关组

定义 0.3.3: 极大线性无关组

在向量组 a_1, a_2, \dots, a_s 中, 若存在部分组 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 满足:

1. $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 线性无关
2. 向量组中的任一向量 $a_i (i = 1, 2, \dots, s)$ 均可由 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 线性表示

则称向量组 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 是元向量组的一个极大线性无关组

需要注意的是, 向量组的极大线性无关组一般不唯一, 一个线性无关向量组的极大线性无关组就是该向量组本身

介绍完极大线性无关组后, 结合前面对矩阵的秩的相关定义, 又可以建立起下面的关系:

1. 向量组 a_1, a_2, \dots, a_s 的极大线性无关组 $a_{i_1}, a_{i_2}, \dots, a_{i_r}$ 中所含向量的个数 r 即为向量组的秩
2. $r(\mathbf{A}) = \mathbf{A}$ 的行秩 (\mathbf{A} 的行向量组的秩) = \mathbf{A} 的列秩 (\mathbf{A} 的列向量组的秩)

介绍完了线性相关和线性无关以及极大线性无关组的相关内容之后, 接下来我们引入方程组的内容。

注. 本小节没有涉及到等价向量组的相关内容, 暂时认为对后续学习不重要, 若后续该概念在课程学习中多次提及再来补充

0.3.2 方程组

本小节主要介绍齐次线性方程组和非齐次线性方程组的结构及其解的性质。

齐次线性方程组的结构及其解的性质

齐次方程组的结构如下:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0, \end{cases}$$

这是一个 m 个方程, n 个未知数的齐次线性方程组

如果用向量表示, 则上述方程组可以表示为:

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n = \mathbf{0}$$

其中

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, j = 1, 2, \dots, n$$

用矩阵表示为

$$\mathbf{A}_{m \times n} \mathbf{x} = \mathbf{0}$$

其中

$$\mathbf{A}_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

方程组有解的条件为:

- 当 $r(\mathbf{A}) = n$ (a_1, a_2, \cdots, a_n 线性无关时), 方程组有**唯一零解**
- 当 $r(\mathbf{A}) < n$ (a_1, a_2, \cdots, a_n 线性相关时), 方程组有**非零解** (无穷多解), 且有 $n - r$ 个线性无关解

非齐次线性方程组的结构及其解的性质

非齐次线性方程组的结构如下

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \quad \quad \quad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

这是一个 m 个方程, n 个未知数的非齐次线性方程组, 其向量形式为 $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$, 其中

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, j = 1, 2, \cdots, n, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

用矩阵表示为

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

其中

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

矩阵 $\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right]$ 称为矩阵 \mathbf{A} 的增广矩阵, 简记为 $[\mathbf{A} : \mathbf{b}]$

方程组有解的条件为:

- 若 $r(\mathbf{A}) \neq r([\mathbf{A}, \mathbf{b}])$ (\mathbf{b} 不能由 a_1, a_2, \dots, a_n 线性表示), 方程组无解
- 当 $r(\mathbf{A}) = r([\mathbf{A}, \mathbf{b}]) = n$ (a_1, a_2, \dots, a_n 线性无关, $a_1, a_2, \dots, a_n, \mathbf{b}$ 线性相关), 方程组有唯一解 (无穷多解)
- 若 $r(\mathbf{A}) = r([\mathbf{A}, \mathbf{b}]) = r < n$, 则方程组有无穷多解

非齐次线性方程组与齐次线性方程组的解的关系

非齐次线性方程组的通解 = 非齐次线性方程组的一个特解 + 齐次线性方程组的通解, 即

1. 求出方程组 $\mathbf{Ax} = \mathbf{0}$ 的通解 $k_1\xi_1 + k_2\xi_2 + \cdots + k_{n-r}\xi_{n-r}$
2. 求出 $\mathbf{Ax} = \mathbf{b}$ 的一个特解 $\boldsymbol{\eta}$
3. 则 $\mathbf{Ax} = \mathbf{b}$ 的通解为 $k_1\xi_1 + k_2\xi_2 + \cdots + k_{n-r}\xi_{n-r} + \boldsymbol{\eta}$, 其中 k_1, k_2, \dots, k_{n-r} 为任意常数

注. 本小节没有涉及基础解系、两个方程组的公共解以及同解方程组的相关内容, 其中基础解系概念较为简单故不概述, 后面方程组的公共解以及同解方程组因为课程中暂未涉及, 故也没有介绍, 若后续课程中需要用到这两部分的内容再在这里补充。

0.4 特征值与特征向量、相似理论

前面三节的内容都是为了本节的内容做铺垫的, 因此在回顾本节内容之前请先确保前三节的大题框架没有问题。

0.4.1 特征值与特征向量的定义

定义 0.4.1: 特征值与特征向量

设 A 是 n 阶矩阵, λ 是一个数, 若存在 n 维非零列向量 ξ , 使得

$$A\xi = \lambda\xi \quad (1)$$

则称 λ 是 A 的特征值, ξ 是 A 对应于特征值 λ 的特征向量

0.4.2 求解特征值与特征向量的步骤

由定义中给出的式子, 可以得到

$$(\lambda E - A)\xi = 0$$

由于 $\xi \neq 0$, 故其次方程组 $(\lambda E - A)\xi = 0$ 有非零解, 对应行列式 $|\lambda E - A| = 0$

解该行列式形成的方程, 可以得出特征值, 之后根据接出的特征值回带入其次线性方程组中, 即可求得该特征值对应的特征向量。

0.4.3 特征值、特征向量的性质与重要结论

特征值的性质与重要结论

1. λ_0 是 A 的特征值 $\Leftrightarrow |\lambda_0 E - A| = 0$
2. λ_0 不是 A 的特征值 $\Leftrightarrow |\lambda_0 E - A| \neq 0$
3. 若 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的 n 个特征值, 则

$$\begin{cases} |A| = \lambda_1 \lambda_2 \cdots \lambda_n \\ \text{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad \text{tr}(A) \text{叫做矩阵 } A \text{ 的迹} \end{cases}$$

特征向量的性质与重要结论

1. $\xi (\neq 0)$ 是 A 的属于 λ_0 的特征向量 $\Leftrightarrow \xi$ 是 $(\lambda_0 E - A)x = 0$ 的非零解
2. k 重特征值 λ 至多只有 k 个线性无关的特征向量
3. 若 ξ_1, ξ_2 是 A 的属于不同特征值 λ_1, λ_2 的特征向量, 则 ξ_1, ξ_2 线性无关

4. 若 ξ_1, ξ_2 是 A 的属于同一特征值 λ 的特征向量, 则非零向量 $k_1\xi_1 + k_2\xi_2$ 仍是 A 的属于特征值 λ 的特征向量
5. 若 ξ_1, ξ_2 是 A 的属于不同特征值 λ_1, λ_2 的特征向量, 则当 $k_1 \neq 0, k_2 \neq 0$ 时, $k_1\xi_1 + k_2\xi_2$ 不是 A 的任何特征值的特征向量
6. 设 λ_1, λ_2 是 A 的两个不同的特征值, ξ 是对应于 λ_1 的特征向量, 则 ξ 不是对应于 λ_2 的特征向量 (即一个特征向量不能属于两个不同的特征值)

0.4.4 矩阵的相似理论

矩阵的相似

定义 0.4.2: 矩阵的相似

设 A, B 是两个 n 阶方阵, 若存在 n 阶可逆矩阵 P , 使得 $P^{-1}AP = B$, 则称 A 相似于 B , 记为 $A \sim B$

矩阵相似有如下特点:

1. 反身性: $A \sim A$
2. 对称性: 若 $A \sim B$, 则 $B \sim A$
3. 传递性: 若 $A \sim B, B \sim C$, 则 $A \sim C$

相似矩阵的性质

若 $A \sim B$, 则

- $|A| = |B|$
- $\lambda_A = \lambda_B$, 或 $|\lambda E - A| = |\lambda E - B|$
- $r(A) = r(B)$
- $r(\lambda E - A) = r(\lambda E - B)$
- $\text{tr}(A) = \text{tr}(B)$
- A, B 的各阶主子式之和分别相等

矩阵的相似对角化

定义 0.4.3

设 A 是 n 阶方阵, 若存在 n 阶可逆矩阵 P , 使得 $P^{-1}AP = \Lambda$, 其中 Λ 是对角矩阵, 则称 A 可以相似对角化, 记 $A \sim \Lambda$, 称 Λ 是 A 的相似标准型

判断矩阵能否相似对角化的条件

充要条件：

- n 阶矩阵 \mathbf{A} 可以相似对角化 $\Leftrightarrow \mathbf{A}$ 有 n 个线性无关的特征向量
- n 阶矩阵 \mathbf{A} 可以相似对角化 $\Leftrightarrow \mathbf{A}$ 对应于每个 k_i 重特征值都有 k_i 个线性无关的特征向量

充分条件：

- n 阶矩阵 \mathbf{A} 有 n 个不同的特征值 $\Rightarrow n$ 阶矩阵 \mathbf{A} 可以相似对角化
- n 阶矩阵 \mathbf{A} 为实对称矩阵 $\Rightarrow n$ 阶矩阵 \mathbf{A} 可以相似对角化

0.5 结语

本章主要对于课上出现的内容涉及到线性代数的一部分知识进行了复习，其余没有出现的部分代表暂时在这门课程中没有用到，随着课程的进行，如果后续会有相关知识涉及到了这一章没有出现的内容，会在这里进行补充。

第一章 线性代数基础

本章将会正式进入矩阵理论的知识内容中，首先是第一章，这一章的名字叫做“线性代数基础”，按照课本的话来讲，这一部分不是单纯的对于线性代数知识的简单回顾与复习，而是在已经掌握线性代数的知识的基础上进行深化，同时，这一章也是后面内容的基础。

本章将会涉及到以下内容：

- 线性空间与子空间
- 空间分解与维数定理
- 特征值与特征向量
- 线性变换
- ...

注意：之后的内容会随着课程的进行进行及时更新

1.1 线性空间与子空间

提到“空间”一词，很多人对这个概念应该并不陌生，我们从出生便降临在这个世界中，这个世界便是一个三维空间，我们使用计算机浏览互联网，这也可以称作一个网络空间……类似的例子还有很多很多。

1.1.1 线性空间

在这里我们要讨论的概念叫“线性空间”，是数学上的空间，上面提到的我们生活着的三维空间，抽象出来也属于线性空间。

线性空间的例子，除了上面提到的三维空间，在平面直角坐标系 x, y 组成的一个平面也是一个空间（二维空间），那究竟何为线性空间呢？有什么标准来判断其是否是一个空间？

判断能否组成一个空间的定义如下：

定义 1.1.1: 判断空间的定义

设 V 是一个非空集合， P 是一个数域，在集合 V 的元素之间定义加法 $v = \alpha + \beta$ ，定义数量乘法 $\delta = k\alpha$ ，如果加法与数量乘法满足下列规则：

- $\alpha + \beta = \beta + \alpha$
- $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
- $\exists \mathbf{0} \in V, \forall \alpha \in V, \text{有 } \alpha + \mathbf{0} = \alpha$ （存在零元素）
- $\forall \alpha \in V, \exists \beta \in V, \text{使得 } \alpha + \beta = \mathbf{0}$ （存在负元素）
- $1\alpha = \alpha$
- $k(l\alpha) = (kl)\alpha$
- $(k + l)\alpha = k\alpha + l\alpha$
- $k(\alpha + \beta) = k\alpha + k\beta$

则 V 称为数域 P 上的线性空间

在这里我们需要明确下面一个概念：何为数域？

定义 1.1.2: 数域的概念

如果一个由数字构成的集合（叫做数集） P ，这个数集对于加法、减法、乘法、除法（除数不为 0）封闭，则就把这个数集 P 叫做数域

这里又引出了一个新的词——封闭，何为封闭？封闭的概念较为简单：如果一个集合中的某两个数做某一运算之后的结果仍然在该集合中，那么就称该集合对于该运算是封闭的。

如果还是对于这一概念不理解，希望下面这个例子能够帮你理解：

例 1.1.1. 全体整数组成的集合 \mathbb{Z} 是否是一个数域？

解. 整数集包含两大部分：正整数和负整数，0 是整数，但 0 既不是正数也不是负数

接下来我们来判断整数集是否对于加法封闭：

我们从小学数学的知识就可以得知，两个整数相加依旧是整数，所以整数集对于加法是封闭的

依次类推，两个整数相减，相乘，结果依旧是一个整数，所以很显然，整数集对于四则运算中的加法、减法和乘法都是封闭的

最后，整数集对于除法是否是封闭的呢？

很显然不是，举一个最简单的例子， $a = 1, b = 2$ ， a 除以 b 的结果是 $\frac{1}{2}$ ，它并不是一个整数，而是一个分数，或者说小数，又可以说是一个有理数，所以整数集对于除法并不是封闭的。

综上，可以断定，整数集并不是一个数域。 □

注. 定义 1.1.1 中的八条性质说明了什么？

左侧四条定义了空间对于加法需要满足以下特性：加法的交换律、加法的结合律、存在零元素、存在父元素

右侧的四条定义了空间的数量乘法需要满足以下特性：数量乘法的结合律、数量乘法的分配律（分配律分为两个标量相加的分配律，以及两个向量相加的分配律）。

通常，定义 1.1.1 给出的八个条件即为判断一个集合是否能构成空间的依据，请看下面的例题。

例 1.1.2. 设多项式集合

$$P_n[x] = \{a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \mid a_i \in P, i = 0, 1, \cdots, n-1\}$$

这里 P_n 代表次数不超过 n 的多项式，请问 $P_n[x]$ 是否能构成一个线性空间？

解. 按照定义 1.1.1 的八条规律，依次来判断

首先，多项式的加法一定满足交换律和结合律（由小学数学学过的加法交换律和加法结合律就能得知），同样的，我们可以在这个多项式集合中找到一个元素 0，使得该多项式与 0

相加的结果依旧是该多项式（很明显，这个元素 0 就是数字 0，即 $a_{n-1}, a_{n-2}, \dots, a_1, a_0$ 均为 0 的时候），此外，我们可以构造出下面的一个多项式集合，令其与 $P_n[x]$ 相加的结果为 0：

$$Q_n[x] = -P_n[x]$$

综上，该集合满足空间定义中对于加法的规律，接下来判断乘法

很显然，存在一个元素 1，使得该集合与 1 相乘的结果就是其本身（这个元素 1 就是数字 1，即 $a_{n-1} = a_{n-2} = \dots = a_1 = 0, a_0 = 1$ 的时候）。同时，由小学数学和初中数学的知识可以知道，多项式的乘法满足数乘的结合律与分配律

综上，该多项式集合是一个线性空间。 □

注. 上面的例子大多数情况下运用了一些“显然，由.....的知识可以得知”，没有具体写明如何得出的结果，有以下两点原因，第一点原因是在写这段文字的时候确实懒得打这么长的公式了，二是认为大家应该能明白上面判断的过程，所以就没有写明公式，当然写出公式也是可以的，如果后面确实不理解上面是如何“显然”得来的，会重新更新这部分的例子，用公式表明。

1.1.2 线性空间的维数

在了解了何为线性空间之后，接下来我们再来了解一下如何形容一个线性空间，维数便是形容线性空间的一个度量，我们前面一直所说的“三维空间”中的“三维”，就表明该空间的维数为 3

定义 1.1.3: 线性空间维数的定义

在线性空间 V 中，如过有 n 个向量 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 线性无关，而 V 中任意 $n+1$ 个向量线性相关，则称 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 为 V 的一组**基底**，由于线性空间的所有基底总含有相同数目的向量，则 n 称为线性空间 V 的**维数**，常记为 $\dim V = n$

上面的定义说明了，一个空间中线性无关向量的个数其实就是该空间的维数，同时，这一系列线性无关的向量便可以构成该空间的一组基，回顾第零章线性表示相关的内容，我们同样可以得知：该空间的任意向量都可以由这一组基来线性表示。

注. 请注意, 向量组/线性空间的维数与向量的维数是两个不同的概念, 一定要注意区分。

向量的维数: 向量有几行, 一般就说是向量的维数为几, 如 $\alpha = [1, 2, 3]^T$, 那么这就是一个三维向量

向量组/线性空间的维数: 向量组中线性无关向量的个数, 如 $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$, 虽然 A 是

一个 3×3 的矩阵, 但是经过分析可以看出该矩阵其实只有一个线性无关的向量, 故该向量组/空间是 1 维的, 但是其中按列分块出的三个列向量却是三维的向量。

1.2 空间分解与维数定理

//TODO