线性代数

王崇政

2023 年 3 月 23 日

目录

1 导引和预备知识

1.1 导引

待补。

1.2 常用符号

下面是本讲义中会采用的常见数学符号，它们大部分是标准记号。

$\mathbb{Z}$	整数集。
$\mathbb{Z}_{\geq 0}$ 或 $\mathbb{N}$	非负整数 (自然数) 集。
$\mathbb{Q}$	有理数集。
$\mathbb{R}$	实数集。
$\mathbb{C}$	复数集。
大写字母 $A, B, C, \cdots$	（某个）集合。
小写字母 $a, b, c, \cdots$	集合中的元素。
$\{x \mid P(x)\}$	集合中满足某个性质 P 的元素集合。
$\emptyset$	空集。
$a \in A$	a 是集合 A 的一个元素。
$a \notin A$	a 不是集合 A 的一个元素。
$A \subseteq B$	A 是 B 的一个子集。
$A \cap B$	集合 A 、B 的交。
$A \cup B$	集合 A 、B 的并。
$\#A$ or $\lvert A\rvert$	集合 A 的基数（元素个数）。
$f : A \to B$	f 是一个从 A 到 B 的映射。
$g \circ f$ 或 gf	映射 f, g 的复合。
$a \mapsto b$	元素 a 被映射到元素 b。
$f(A)$	映射 $f : A \to B$ 的像。
$f^{-1}(B)$	映射 $f : A \to B$ 的原像。
$A \times B$	集合 A 、B 的笛卡尔积。
(a, b)	笛卡尔积中的元素（有序数组）。
$\exists$	存在。
$\forall$	对所有。

$A\coprod B$	集合 A,B 的不交并。
$:=$ 或 $\stackrel{def}{=}$	被定义为。
$k[x_1,\cdots,x_n]$	系数在 k 中以 $x_1,\cdots,x_n$ 为自变量的多项式。

以上部分记号的确切含义将会在下文中仔细讨论。在课程当中我们还会介绍相关内容中的其他常见符号。

遵循规范的数学符号有利于数学资料的正确理解和写作。

1.3 集合及相关概念；笛卡尔积；关系和映射

这一节我们简要的回顾集合及其相关概念，如集合的概念、集合间映射、笛卡尔积。关于集合论的基础知识可参考 [G13], 第 3，4 章；[R19] 第 2，9 章。

在本课程中我们只使用**朴素集合论**当中最常用的部分。由于内容所限本课程不讨论较为严肃的**公理集合论**，例如 ZFC 公理体系。有兴趣者可参考 [W18] 第二章。

在中学当中我们已经初步有了元素与集合的概念。在朴素集合论当中，一些概念是不做定义的元概念，例如元素、集合、元素相同（不同）。直观的讲，集合是由一群元素构成的群体。在集合论当中我们使用以下基本设定。

1. 由任意一族元素 $a,b,c\cdots$ 可以构成集合
$$A:=\{a,b,c,\cdots\}$$
元素 a,b,c 也称为集合 A 的**成员**，记作
$$a\in A.$$
2. 任意两个元素 a,a' 存在且仅存在
$$a=a'$$
与
$$a\neq a'$$
两种之一的可能。
3. 两个集合 A,B 满足
$$A=B$$
当且仅当它们具有相同的成员。

4. 存在不包含任何元素的集合 $\emptyset$，称为空集。

5. 若 A 的任何元素都是 B 的元素，则称 A 是 B 的子集合，或者等价的，B 包含 A，记作
$$A \subseteq B \text{ 或者} B \supseteq A.$$
我们还规定空集是任意集合的子集合。

6. 集合自身也可以看作元素而构成新集合的成员。

7. 对任意集合 X 的两个集合 A, B，存在交集
$$A \cap B$$
由共同属于集合 A, B 的元素构成；以及并集
$$A \cup B$$
由属于 A, B 其中之一的元素构成；和 A 在 X 中的补集合
$$A^c \text{ 或者} X/A \text{ 或者} X - A,$$
由 X 中不属于 A 的元素构成。

以上是中学中已经熟知的内容，但我们还进一步的需要一些关于集合构造的知识。

在中学组合数学部分我们学过所谓“乘法原理”。更具体的，将“乘法原理”中枚举出的元素罗列出来，即是集合构造中重要的的**笛卡尔积**，它相当于对集合利用“乘法”运算得到了一个新的集合。

定义 1. 给定 m 个集合 $A_1, \cdots, A_m$, 它们的笛卡尔积 $A_1 \times \cdots \times A_m$ 定义为 m 元有序组
$$A_1 \times \cdots \times A_m := \{(a_1, \cdots, a_m) |\, a_i \in A_i\}$$
构成的集合。

直观的看，集合的笛卡尔积相当于 m 个集合 $A_1, \cdots, A_m$ 的一个带有箭头顺序的“串”，每个集合 A_i 恰好被“射中”一个元素。

$$(a_{1,j_1}, \cdots, a_{i,j_i}, \cdots, a_{m,j_m}) = \begin{array}{ccccc} a_{1,s_1} & & a_{i,s_i} & & a_{m,s_m} \\ \vdots & & \vdots & & \vdots \\ (a_{1,j_1}) & \longrightarrow & (a_{i,j_i}) & \longrightarrow & (a_{m,j_m}) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{11} & & a_{i1} & & a_{m1} \\ & & & & \\ S_1 & & A_i & & A_m \end{array}$$

我们需要注意的是

1. 与集合中的元素不同的是，在笛卡尔积的定义中需强调“有序性”，例如若 $A_1 = A_2 = A$，$a, b \in A$ 且 $a \neq b$，则
$$(a,b) \neq (b,a).$$

2. 笛卡尔积不满足“结合律”。一般的，
$$(A_1 \times A_2) \times A_3,\, A_1 \times (A_2 \times A_3),\, A_1 \times A_2 \times A_3$$
是三个不同的集合（为什么？)。

利用笛卡尔积，我们可以严格的重新叙述一些熟知的基本概念。

给定集合 A, B。

定义 2. 笛卡尔积 $A \times B$ 的一个子集合 $M \subseteq A \times B$ 称为 A 到 B 的一个**对应**或 A, B 间的**二元关系**，且若 $(a,b) \in M$，则称 b *(*在 M 下*)* 与 a **对应**。特别的当 $B = A$ 是同一个集合时，笛卡尔积 $A \times A$ 中的一个子集 $R \subseteq A \times A$ 称为 A 上的一个**二元关系**（或 A 上的一个**对应**）。

在数学当中需要研究各种各样的对应与二元关系。这其中重要的是基本概念是**映射**、**（全，偏）序关系**，以及 **等价关系**。我们进行依次介绍。

设 M 是 A 到 B 的一个对应。

定义 3. 称
$$Dom(M) := \{a |\, a \in A,\ \text{存在}\, b \in B\ \text{使得}(a,b) \in M\},$$
为 M 的**定义域**。特别的若
$$Dom(M) = A$$
则称 M 是 **完全定义在** A **上的**。

我们可以将映射（函数）定义为特殊的对应，即进一步要求与任何定义域中的元素对应的元素总是唯一的。

定义 4. A 到 B 间的一个对应 $M \subseteq A \times B$ 称为一个 **(部分定义的) 从** A **到** B **的映射**，或者等价的，**(部分定义在)** A **上的** B**-值函数**，如果对任意 $a \in Dom(M)$，存在唯一的 $b \in B$ 对应 a，也即存在唯一 $b \in B$ 使得
$$(a,b) \in M.$$

通常的我们以一个小写字母 f 来表示一个映射 M 并写作

$$f : A \to B, a \mapsto b,$$

且称与 a 唯一对应的元素 b 为在 f 映射下的**像**，记作

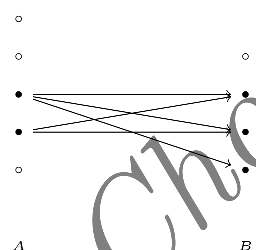
$$b = f(a).$$

我们也用符号

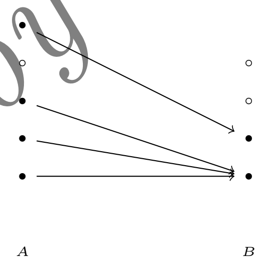
$$Dom(f) := Dom(M)$$

表示 f 的**定义域**。特别的当 $A = B$ 是同一个集合时，A 到自身的映射也称为集合 A 上的一个**变换**。

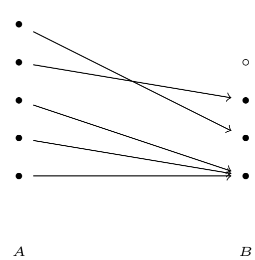
利用图示可以直观表明例子。以下是一个（部分定义的）关系（对应）的图示：



一个部分定义但不完全定义的映射的例子：



一个完全定义的映射（函数）的例子：



由以上的定义我们可以推出

命题 1. 若 f, g 为 A 到 B 的映射，则

$$f = g$$

当且仅当

$$f(a) = g(a)$$

对所有 $a \in A$ 成立。

对子集 $A' \subseteq A$, 称集合

$$f(A') = \{f(a) |\, a \in A'\}$$

为 A' 在映射 f 下的的**像**，特别的对完全定义的 f 我们也简称 $f(A)$ 为 f 的**像**，有时也记作 $Im(f)$。对子集 $B' \subseteq B$, 称集合

$$f^{-1}(B') := \{a |\, a \in A,\, f(a) \in B'\}$$

为 B' 在映射 f 下的**原像**，特别的我们也简称 $f^{-1}(B)$ 为 f 的**原像**。若对任意 $a \neq a'$ 推出

$$f(a) \neq f(a')$$

则称 f 为一个 **单射**。若

$$f(A) = B$$

则称 f 为一个**满射**。既是单射又是满射的映射称为一个**双射**。

需要注意的是

1. 在含义上“函数”与“映射”,“对应”与“(二元)关系”是两组同义词，但在不同场合会有所偏向。例如 A 到 B 的“对应”更强调 A 到 B 的“方向性”，进一步的用以定义函数与映射；而“(二元)关系”则主要强调 A, B 的“联系”，用来衍生出等价关系、偏序关系等概念。

2. 在一些场合中我们确实需要考虑一般的对应（例如代数几何），以及部分定义的映射（函数）（例如泛函分析）。但在本课程中，我们主要考虑映射和函数，并且如无特别声明，映射和函数都是**完全定义**的，即只考虑 $Dom(f) = A$ 的情形。此时我们也不加区分的将定义集合 A 称为 f 的定义域。

3. 在某些文献中对部分定义映射和函数，也将定义集合 A 称为“定义域”，像集合 $Im(f)$ 称作“值域”，取值集合 B 称为 **上域 (codomain)**，需小心区分。

在数学中还有一些常见定义需要将笛卡尔积与映射结合使用。

定义 5. 从笛卡尔积 $A^n := A \times \cdots \times A$ 到集合 B 的一个映射

$$f : A \to B$$

称为一个定义在 A 上到 B 的 n **元映射** ***(n-array maps)***。特别的当 $B = A$ 是同一个集合时，称一个映射

$$f : A^n \to A$$

为 A 上的 n **元运算** ***(n-array operations)***。

在我们熟知的例子中，加减乘除运算都是二元运算；取相反数和取倒都是一元运算。

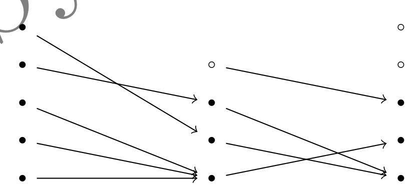
给定集合 A, B, C, 以及映射

$$f : A \to B, g : B \to C.$$

A 到 C 的映射

$$g \circ f : A \to C, a \mapsto g(f(a)) \text{ 对所有} a \in A$$

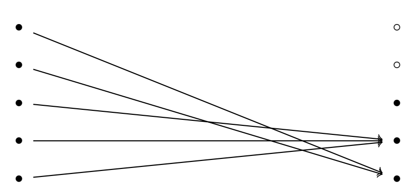
称作映射 f, g 的**复合**或**乘积**。直观的看映射复合即沿箭头方向进行路径的“合成”。例如：



$$A \xrightarrow{f} B \xrightarrow{g} C$$

$$A \xrightarrow{g \circ f} C$$

=



需要注意一般情况下映射 f, g 并不总是可以做复合的，而可以做复合的条件是 f 的值域包含在 g 的定义域中。另外需要注意的是，由于通用的书写习惯，复合 $g \circ f$ 中的 f 总是写在 g 的右侧，但复合 $g \circ f$ 表示的映射是先 f 而后 g。另外今后我们也常用箭头图来表示：

$$A \xrightarrow{f} B \xrightarrow{g} C,$$

在现代数学当中，箭头图具有很强的直观性并导致产生了新的概念。关于映射复合我们可以推出

命题 2. 三个映射 f, g, h 在可以做复合的情况下，相应的复合满足结合律

$$(h \circ g) \circ f = h \circ (g \circ f).$$

1.4 逻辑基础

此部分暂略，可按需要查阅相关参考书 [R19], 第 1 章；[W18] 第二章；[G13] 第 1，2 章。

1.5 偏序关系；等价关系和等价类

定义 6. 集合 S 上的一个 **二元关系** R 称为一个 **偏序关系**,如果对任意 $(x,y) \in R$ 有：

1. (自反性) $(x,x) \in R$ 对所有 $x \in S$。
2. $(x,y) \in R$ 且 $(y,x) \in R$ 推出 $x = y$。
3. （传递性）$(x,y),(y,z) \in R$ 推出 $(x,z) \in R$。

习惯上我们将一个偏序关系

$$x \sim_R y$$

写作

$$x \preccurlyeq y.$$

如果一对元素 $x, y \in S$ 满足 $x \preccurlyeq y$ 或者 $y \preccurlyeq x$，则称 x, y **可比较**。若任意一对元素 $x, y \in S$ 对应偏序关系 $\preccurlyeq$ 都是可比较的，则称此偏序关系为一个 **全序关系**。

偏序中“偏”字的直观含义是，偏序关系不要求对所有的有序对都定义，而一般只需对部分元素定义集合。

例 1. 整数、有理数、实数上的小于等于关系，都是全序关系。

例 2. 正整数的整除关系是一个偏序关系而不是全序关系，因为任意给定的一组非零整数，未必其中一个可以整除另外一个。

等价关系和由等价关系划分的等价类，代数学中最重要的基本构造之一。

定义 7. 集合 S 上的一个二元关系 $R \subseteq S \times S$ 如果满足下列三个条件，则称 R 是一个 **等价关系**：对任意 $(x, y) \in R$ 有：

1. (自反性) $(x, x) \in R$ 对所有 $x \in S$。
2. （对称性）$(x, y) \in R$ 推出 $(y, x) \in R$。
3. （传递性）$(x, y), (y, z) \in R$ 推出 $(x, z) \in R$。

如果 $(x, y) \in R$，我们称 x **等价于** y 并将其写作

$$x \sim_R y.$$

如无歧义可简记为 $x \sim y$ 。

令 S 是一个集合, R 是 S 上的一个二元关系。对任意元素 $x \in S$，用

$$\bar{x} : \{a | \, a \in S, a \sim_R x\}$$

表示集合 S 中和元素 x 等价的元素组成的集合. 我们把 $\bar{x}$ 称作 **一个等价类 (相应于元素 x 的)**. 注意到对于两个等价类 $\bar{x}, \bar{y}$, 它们或者相同：

$$\bar{x} = \bar{y},$$

或者不交：

$$\bar{x} \cap \bar{y} = \emptyset.$$

我们也把由 S 中等价类 $\bar{x}$ 构成的集合称作 **由等价关系 R 划分的等价类**, 并用

$$S/\sim_R := \{\bar{x} | \, x \in S\}, ---$$

或简记为 $S/\sim$ 。由此注意到，相应的存在一个映射

$$\pi : S \to S/\sim, x \mapsto \bar{x},$$

称为 **标准投射**。

我们在中小学数学当中已经接触过许多使用等价关系和等价类的例子

例 3. 在小学数学中，我们将有理数定义为“分数”，即由整数分子和非零整数分母够成熟“数”，用等价关系和等价类的语言予，我们可以对此进行严格叙述。我们用笛卡尔积

$$\mathbb{Z} \times \mathbb{Z}_{\neq 0}$$

中的一个元素

$$(a, b) \in \mathbb{Z} \times \mathbb{Z}_{\neq 0}$$

来表达一个“有理分数”

$$\frac{a}{b},$$

其中 a, b 分别作为“分子”和“分母”。注意到不同的一组分子和分母可能表达了同一个有理数，例如

$$\frac{1}{2} = \frac{2}{4} = \frac{2022}{4044} = \cdots.$$

于是在 $\mathbb{Z} \times \mathbb{Z}_{\neq 0}$ 定义如下二元关系：

$$(a, b) \sim (c, d) \text{ 当且仅当 } ad = bc.$$

可以验证这是一个等价关系，进而

$$\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}_{\neq 0} / \sim.$$

定义 8. 集合 S 的一个 **划分** 是指 S 的一个子集族 $\mathcal{P} = \{P_\lambda | P_\lambda \subseteq S\}$, 满足

$$S = \bigcup_{P_\lambda \in \mathcal{P}} P_\lambda$$

且对任何两个集合 $P_i, P_j \in \mathcal{P}$，$P_i \neq P_j$ 有

$$P_i \cap P_j = \emptyset.$$

其中子集族 $\{P_\lambda\}$ 中的单个集合 P_i 称作划分的一个**部分**。

命题 3. 集合 S 上的一个等价关系自然的决定了 S 上的一个划分，且等价类一一对应于划分中的部分。反之，集合 S 上的任何一个划分决定了 S 上的一个等价关系。

例 4. 集合 S 上的所有划分，也存在一个自然的偏序关系。给定两个划分 $\mathcal{P}$, $\mathcal{P}'$，称 $\mathcal{P}$ 比 $\mathcal{P}'$ **粗** (或 $\mathcal{P}'$ 比 $\mathcal{P}$ **细**)，如果划分 $\mathcal{P}'$ 中任何一个部分都是划分 $\mathcal{P}$ 中某个部分的子集，即对任意 $P' \in \mathcal{P}'$，存在 $P \in \mathcal{P}$ 使得

$$P' \subseteq P.$$

容易验证以上给出了集合划分的一个偏序关系，且若 $\mathcal{P}'$ 比 $\mathcal{P}$ 更细，则 $\mathcal{P}'$ 对 $\mathcal{P}$ 中的每一个部分都诱导了一个划分。

另外注意到，集合上的任意两个划分并不总是能够比较的，故以上偏序关系一般不是全序关系。

与等价关系相关的映射

集合以及等价类本身的结构比较单薄，因而我们经常还需要进一步的考虑和等价类有关的映射。假设存在原像集合 S 上的一个等价关系 $\sim$, 同时存在一个从集合 S 到到集合 T 的映射：

$$f : S \to T.$$

倘若存在唯一的一个映射 $\tilde{f} : S/\sim \to T$ 使得

$$f = \tilde{f} \circ \pi.$$

成立，即有下面的**交换图表**:

$$\begin{tikzcd} & T \\ S \arrow[ur, "f"] \arrow[r, "\pi"'] & S/\sim \arrow[u, dashed, "\tilde{f}"'] \end{tikzcd}$$

那么我们说映射 f **穿过**（或者 **良定义在**, **继承到**）等价类 $S/\sim$。在交换图中，通常以虚线箭头来表示一个条件性存在的映射。下面这个命题是简单显然的，但是它作为一个基本模式，会反复出现在代数学的许多问题和构造当中。

引理 1. *以上映射 f 良定义在 $S/\sim$ 上当且仅当对任意的 $x \sim y$，有*

$$f(x) = f(y),$$

且相应的映射 $\tilde{f}$ 由 f 唯一决定。

类似的当等价类作为映射的像是，也有类似问题。假定有映射

$$g : T \to S/\sim,$$

我们称标准映射 $\pi : S \to S/\sim$ 可被 **提升到** T, 如果存在映射 $\tilde{g}$ 使得

$$g = \pi \circ \tilde{g},$$

即有交换图

$$
\begin{tikzcd}
 & T \arrow[d, "g"] \\
S \arrow[ur, dashed, "\tilde{g}"] \arrow[r, "\pi"'] & S/\sim
\end{tikzcd}
$$

同样类似的

引理 2. 标准映射 $\pi : S \to S/\sim$ 可被提升到 T 当且仅当 g 是满射。

需要注意的是与前一情况不同，若 π 可以被提升，则一般来讲提升不是唯一的。

实数和复数

关于复数、数域、求和符号的预备知识可参考 [L02a] 1.1 小节。

在中学阶段，我们已经熟知**有理数集** $\mathbb{Q}$ 及其子集**整数集** $\mathbb{Z}$，它们都可以进行加减乘除（除数非零）四则运算。由于数学上的需要，我们考虑由有理数“扩展”得到**实数集** $\mathbb{R}$。形式上我们可以将实数等同于无限（十进制）小数。在数学分析课程中，我们将知道，实数是通过有理数做**完备化**得到的，继承了有理数所具有的四则运算，并且实数是**完备的**。

进一步的，由于代数上的需要，我们还需要将实数拓展为**复数**。例如，我们回忆实系数二次方程

$$ax^2 + bx + c = 0$$

在判别式 Δ 满足

$$\Delta < 0$$

的时候，是无实数解的。特别的，

$$x^2 + 1 = 0$$

没有实数解。解决这一障碍的办法，是“强行添加**纯虚数** $\sqrt{-1}$”到实数集合上构成**复数**。

我们可以通过纯粹形式的办法来定义复数。

定义 9. 作为集合，**复数集合** $\mathbb{C}$ 定义为实数二元有序组构成的集合

$$\mathbb{C} := \{(a, b) \,|\, a, b \in \mathbb{R}\}.$$

通常我们以

$$a+bi \text{ 或者 } a+b\sqrt{-1}$$

来表示二元有序组

$$(a,b)$$

对应的复数，并且我们分别称 a,b 为 z 的**实部**和**虚部**，记作

$$a=Re(z), b=Im(z).$$

复数的加法、减法和乘法定义为

$$(a+bi)+(c+di):=(a+c)+(b+d)i$$
$$(a+bi)(c+di):=(ac-bd)+(ad+bc)i.$$

有定义可以看出，任意一个复数唯一对应平面上的一个坐标点。但是相比纯粹的坐标点，复数还具有四则运算。

我们可以直接验证

命题 4. 复数集合在加法和乘法下是封闭的，并且加法、乘法都是交换的：

$$z_1+z_2=z_2+z_1,\ z_1z_2=z_2z_1,\ \text{对所有} z_1,z_2\in\mathbb{C}.$$

对任意非零复数 z，它具有唯一的乘法逆 z^{-1}:

$$zz^{-1}=z^{-1}z=1.$$

对于复数 $z=a+bi\neq 0$，我们容易验证它的逆 z^{-1} 为

$$\frac{a-bi}{a^2+b^2}.$$

复数具有**复共轭**这一特别的运算。

定义 10. 对任意复数 $z=a+bi$，定义其**复共轭**为

$$\bar{z}:=a-bi.$$

复共轭的重要性质在于，它与四则运算“交换”。

命题 5.

$$\overline{z_1+z_2}=\bar{z}_1+\bar{z}_2,\ \overline{z_1z_2}=\bar{z}_1\bar{z}_2,\ \overline{z^{-1}}=\bar{z}^{-1},$$

并且

$$z=\bar{z}$$

当且仅当

$$z\in\mathbb{R}.$$

我们可以证明复共轭运算是唯一满足以上性质的运算。

在实践中我们还常用复数的**辐角表示**，它使得乘法具有较为简单的形式。一个复数 $z = a + bi$ 的**模长**定义为实数

$$|z| := \sqrt{a^2 + b^2}.$$

命题-定义 1. *若 $z \neq 0$，存在唯一的 $\theta \in [0, 2\pi)$ 使得*

$$a = r \cdot cos(\theta), b = r \cdot sin(\theta).$$

当 $r = 0$，即 $z = 0$ 时，我们可以取 θ 为任意实数而仍是的上述命题成立。

在数学分析中我们将证明 **欧拉公式**：

$$cos(\theta) + \sqrt{-1} \cdot sin(\theta) = e^{i\theta}.$$

特别的，

$$e^{2\pi i} = 1.$$

因而在前述命题的约定下，有

$$z = re^{i\theta}.$$

我们称此为复数 z 的**三角表示**。在此表示下复数的乘法运算变得较为简单：

$$(r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) = r_1 r_2 e^{i(\theta_1 + \theta_2)}.$$

此外对相应的复共轭我们有

$$\overline{re^{i\theta}} = re^{-i\theta}$$

单位根是一类比较特殊的复数，并且在代数中频繁出现。给定一个整数 $n \geq 2$, 称一个复数 ω 为一个 n **次单位根**，若

$$\omega^n = 1.$$

显然对任意 n，1 都是 n 次单位根，但除此之外我们并没有一个“标准”的单位根，这是由于共有 n 个单位根。利用三角函数和幅角表示，我们容易得到所有的 n 次单位根为

$$exp(\frac{2l\pi i}{n}), l = 0, \cdots, n-1.$$

另外，n 次单位根的逆和乘积仍是 n 次单位根：

$$exp(\frac{2k\pi i}{n})exp(\frac{2l\pi i}{n}) = exp(\frac{2(k+l)\pi i}{n}), exp(\frac{2l\pi i}{n})^{-1} = exp(\frac{-2l\pi i}{n}).$$

在熟悉整数模运算的情况下，我们注意到 n 次单位根的乘法运算与模 n 加法运算具有相似性。

最后我们再来陈述代数基本定理，它表明了复数的一个重要基本性质。

定理 1. *若 $f(x) \in \mathbb{C}[x]$ 是一个非常数复系数多项式，则*

$$f(x) = 0$$

在 $\mathbb{C}$ 存在根。

在复变函数和抽象代数中会给出代数基本定理比较简洁的证明。

数域和一般域的概念

分析总结有理数 $\mathbb{Q}$, 实数 $\mathbb{R}$ 与复数 $\mathbb{C}$ 这几个集合，我们发现它们都具有加法和乘法，并且具有以下性质：

1. 加法具有"加法单位元"0, 每个元素 a 都满足

$$a + 0 = 0 + a = a,$$

并且存在唯一的元素 $-a$ 使得

$$a + (-a) + = (-a) + a = 0.$$

2. 乘法具有"乘法单位元"1, 每个元素 a 都满足

$$a \cdot 1 = 1 \cdot a = a,$$

并且对每个非零元素 a 都存在唯一的元素 a^{-1} 使得

$$aa^{-1} = a^{-1}a = 1.$$

3. 加法和乘法都满足交换律和结合律：

$$a + b = b + a,\ ab = ba;\ a + (b + c) = (a + b) + c,\ a(bc) = (ab)c.$$

4. 乘法对加法满足分配律

$$a(b + c) = ab + ac.$$

在本课程当中我们会看到，许多情况下，我们仅仅用到了数域及其以上运算的性质，而没有用到一些其他性质，如实数、有理数的全序性、实数的完备性等等。因此我们有必要抽象出更一般的具有以上性质的概念。

定义 11. *我们称一个具有加法运算 $+$、乘法运算 $\times$, 以及有加法单位元 0、乘法单位元 1，满足以上性质 (1)——(4) 的集合为一个***域***。*

如果两个域 k, K 具有包含关系

$$k \subseteq K,$$

且 k 中的加法乘法运算，以及加法乘法单位元与 K 中的相同，则称 k 为 K 的**子域**, K 为 k 的 **扩域**。显然由定义，$\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 都是域, 且后者是前者的扩域。本节中我们暂时都以 k 表示一个域。

可以看到在这里，域的定义比较冗长。在抽象代数课程中，利用阿贝尔群的概念，我们可以比较简洁的重新叙述域的公理。

复数域 $\mathbb{C}$ 的一个子域称作**数域**。需要注意的是数域**不是**数学文献中的标准术语，且数学文献当数域的定义与此定义有略微差异。

我们给出一个的非平凡域的例子，它们和熟知的几个数域 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 都不同。

例 5. 我们知道 $\sqrt{2}$ 是无理数，故对 $a, b \in \mathbb{Q}$,

$$a + b\sqrt{2} = 0$$

当且仅当

$$a = b = 0.$$

令

$$\mathbb{Q}[\sqrt{2}] := \{a + b\sqrt{2} | \, a, b \in \mathbb{Q}\}.$$

则 $\mathbb{Q}[\sqrt{2}]$ 构成一个 $\mathbb{R}$ 的一个子域。为了验证 $\mathbb{Q}[\sqrt{2}]$ 是一个域，仅需验证 $\mathbb{Q}[\sqrt{2}]$ 关于乘法和（非零）乘法逆封闭，而这可以通过类似复数的直接计算得出：

$$\begin{aligned}
&(a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2}) = (a_1a_2 + 2b_1b_2) + (a_1b_2 + a_2b_1)\sqrt{2}, \\
&(a + b\sqrt{2})^{-1} = \frac{a - b\sqrt{2}}{a^2 + 2b^2}.
\end{aligned}$$

此外我们会在初等数论与抽象代数中遇到许多不是数域的域的例子，。

求和与乘积符号

在数学当中我们还经常使用**求和符号** Σ 和**乘积符号** $\prod$，来表达求和与连乘积。这里 Σ 和 Π 的命名规则，分别是求和（summation）与乘积 (product) 西文首字母的对应的大写希腊字母。

以较为简单的情况为例，设 $\{a_i | \, i = 0, 1, \cdots\}$ 是一串由自然数标记的数，其中 i 为**指标**。我们以

$$\sum_{i=0}^{n} a_i$$

表示下标等于 0 到 n 之间数字 a_i 的和:

$$a_0 + \cdots + a_n.$$

其中 0 和 n 分别称作求和的**上限**和**下限**，指标所在的自然数集合 $\mathbb{N}$ 成为 **指标集合**。在一般的使用中我们还可以依据稍做推广，例如

1. 指标集合可以推广到整数集合。

2. 指标求和上下限可以是任意一对整数。

另外我们约定，一切使用求和符号的求和需要预先**指定**或默认好指标集合。

在使用当中我们还需要做一些进一步的推广。例如在某些情况下，给定指标集合 I，我们需要的求和范围是 I 的一个子集合，由一个条件 P 来决定。例如数学分析当中经常使用

$$\sum_{i\geq 0} a_i$$

来表示无限求和

$$a_0 + a_1 + \cdots.$$

此处的求和范围是由条件 $i \geq 0$ 给出的指标集合 $\mathbb{Z}$ 的子集合 $\{i| i \geq 0\}$，它等价于另一种常用的利用上下限的记号

$$\sum_{i=0}^{\infty} a_i$$

最一般的情况下，求和的形式为

$$\sum_{P(i)} a_i,$$

其中 P 为求和指标所满足的约束条件。特别的若指标集合确定但无约束条件，则

$$\sum_{i} a_i$$

表示对指标集合对应所有元素求和。例如若指标集合为正整数集合 $\mathbb{Z}_{>0}$，则

$$\sum_{p\ \text{是素数}} \frac{1}{p^2}$$

表示对所有满足 p 是素数的数 $\frac{1}{p^2}$ 进行求和，而

$$\sum_{i} \frac{1}{i^2}$$

表示对所有形如 $\frac{1}{i^2}$ 的数（i 是正整数）进行求和。以上描述形式大大推广了传统的上下限形式求和。

另一种常用的求和形式是指标集合为一个集合的**笛卡尔积**，即**多重求和**。例如，若数字集合 $\{a_{ij}\}$ 对应的指标集合，为双自然数指标集合 $\mathbb{N}\times\mathbb{N}$，则

$$\sum_{i=0}\sum_{j=0}a_{i,j}$$

表示对集合 $\{a_{ij}\}$ 中所有元素，先对对应指标 j 求和, 再对对应指标 i 求和：

$$\begin{aligned}&(a_{0,0}+a_{0,1}\cdots)\\+&(a_{1,0}+a_{1,1}\cdots)\\+&\ddots.\end{aligned}$$

在不需要考虑求和次序的情况下，多重求和可以仅使用一个求和符号而直接写为

$$\sum_{i,j}a_{ij}.$$

进一步的，我们可以推广到指标集合为多个集合的 (k-重) 笛卡尔积的情况，其中的分量集合也可以不同：

$$\sum_{i_1,\cdots,i_k}a_{i_1,\cdots,i_k},\, i_j\in A_j.$$

多重求和与利用条件限制指标集合求和, 也可以结合使用。例如，取指标集合为正整数集合的二元笛卡尔积

$$\mathbb{Z}_{>0}\times\mathbb{Z}_{>0},$$

则求和

$$\sum_{i<j}\frac{1}{ij}$$

表示所有形如 $\frac{1}{ij}$ 且 $i<j$ 的数字之和。此处求和条件中 $i<j$ 是集合

$$\{(i,j)|\, i<j,\, i,j\in\mathbb{Z}_{>0}\}$$

的简写。

最后，在使用求和符号的过程中，还需要注意求和范围对应的指标是可以依据需要进行替换的，因而也称为。例如取指标集合为整数集合，则

$$\sum_{i>0}\frac{1}{i^2}$$

当中的指标 i 即为哑指标，它可以替换为任何字母，例如

$$\sum_{i>0}\frac{1}{i^2}=\sum_{j>0}\frac{1}{j^2}=\sum_{k>0}\frac{1}{k^2}.$$

根据实际情况进行哑指标符号替换，是数学写作中常用的技巧。另外需要注意的是，在一些表达式当中可能会出现求和（哑）指标与非求和指标共存的情况，此时非求和指标是不能随意做替换的。

对乘积符号我们也有与求和符号完全平行的一套规则，此处略。

自然数的 Peano 公理和数学归纳法

我们在小学就已经学过自然数的概念。在很长一段时间内，人们没有意识到需要从基础层面定义自然数。意大利数学家 Peano 第一个注意到了这个问题，他在 1888 年首次提出了刻画自然数的 Peano 公理。

公理 1. *自然数集合 $\mathbb{N}$ 是满足以下公理的集合*

1. *存在唯一的一个特殊元素称作零，记为 0。*
2. *$\mathbb{N}$ 存在一个称作***后继***的一元运算：*

$$\cdot^+:\mathbb{N}\to\mathbb{N},\, a\mapsto a^+,$$

   即，任何元素存在唯一后继。

3. *任何非零元素都存在唯一的“前缀”，即若*

$$a^+=b^+$$

   则

$$a=b.$$

4. *0 不存在“前缀”，即不存在元素 a 使得 $a^+=0$。*
5. *（归纳公理）一族由自然数标记的性质 $\{P(n)\}, n\in\mathbb{N}$，若满足*
   (a) *$P(0)$ 成立。*
   (b) *对任意 k，$P(k)$ 成立总推出 $P(k^+)$ 也成立。*

   则所有 $P(n)$ 都成立。

若我们将“后继”a^+ 理解为

$$a+1,$$

则以上几条公理是非常容易理解的。

由以上公理出发，我们可以推出所有熟知的整数的性质，这里不做详细探讨。但是在这其中尤为重要的是第五条的**归纳公理**，它是数学中极为常用的**数学归纳法**的公理基础。归纳公理中 (a) 条称为**归纳基础**，(b) 条称为**归纳假设**。在利用归纳法证明的过程中，两者缺一不可。在许多情况下，直接证明一个以上形式的性质会比较困难，但是分别验证归纳公理的两条会比较容易。

在实际使用当中还会遇到以上形式归纳法的各种变形。

2 矩阵运算，线性空间，与线性变换

2.1 矩阵运算

2.1.1 矩阵的基本概念和基本运算

定义 12. 给定域 k, 以及两个正整数 m,n，称以下将 $n\times m$ 个数 $a_{ij}\in k$ 按 n 行 m 列方式依次排布

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

并在左右两侧添加圆括号的形式为 $n\times m$ **矩阵**，有时也简记作

$$(a_{ij})_{n\times m},$$

对任意一对数 $1\leq i\leq n, 1\leq j\leq m$，其中 (i,j) 位置处的数 a_{ij} 也称为 A 在相应位置处的 **矩阵元素**，或**坐标分量**。$n\times m$ 矩阵构成的全体集合也记作 $M_{n\times m}(k)$，且当 $m=n$ 时可简记为 $M_n(k)$。

作为集合，形式上我们可以将一个矩阵看作域 k 上的 nm 重笛卡尔积

$$\underbrace{k\times\cdots\times k}_{nm\text{ 次}},$$

而每个矩阵可以看作是 mn 个数 a_{ij} 组成的有序数组:

$$(a_{11},\cdots,a_{1,n};\cdots;a_{n1},\cdots,a_{nm}).$$

我们会看到，矩阵按 n 行 m 列形式排布，在形式之外还具有进一步的性质和意义。

我们把 $n\times 1$ 矩阵称为 n **维列向量**:

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

$1\times n$ 的矩阵称为 n **维行向量**:

$$(a_1,\cdots,a_n).$$

特别的，k 中的一个元素 $a\in k$ 也可以看做一个 1×1 的矩阵，为和一般矩阵相区分，也常常称 a 为一个**标量**。

注. 一般的，矩阵中的元素可以取为更为一般的**环**中的元素，并且相应的矩阵运算规则仍然适用。但在本讲义中我们主要讨论域的情况，有时会涉及到**交换环**。这其中重要的例子是元素为多项式的情况。

矩阵可以看作是推广了的（标量）数的概念，它也具有两种最基本的运算：加法和乘法。任意给定两个 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},\ B=\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \cdots & \cdots & \cdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix},$$

它们具有相同的行数 n 和相通的列数 m。此时我们称 A，B 是**可相加**的。我们定义它们的加法为如下的一个新的 $n\times m$ 矩阵：

$$A+B=\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{pmatrix},$$

即 $A+B$ 为 A,B 对应元素相加。称 $A+B$ 为 (可相加矩阵) A，B 的**和**。

矩阵的乘法运算比矩阵加法稍加复杂。任意给定一个 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

和一个 $m\times l$ 矩阵

$$B=\begin{pmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{ml} \end{pmatrix},$$

满足 A 的列数 (等于 m) 等于 B 的行数（也等于 m）。此时我们称 A，B 是 **可相乘**的，并且称下面的 $n \times l$ 矩阵

$$AB := \begin{pmatrix} c_{11} & \cdots & c_{1l} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nl} \end{pmatrix}, \, c_{ij} := \sum_k a_{ik} b_{kj}.$$

为 AB 为 A **左乘** B, 或者 B **右乘** A，简称为 A，B 的**乘积**。通过直接计算验证我们可以得到：

命题 6. *若 A, B, C 三个矩阵，满足 A, B 与 B, C 分别可乘，则 A, BC 和 AB, C 也分别可乘，且成立结合律，即：*

$$A(BC) = (AB)C.$$

另外我们需要特别注意以下两点

1. 与通常数字的的加法乘法运算不同，矩阵的加法和乘法运算是**部分定义**的运算，它们只在可相加和可相乘的情况下才可以定义相应操作。

2. 矩阵的可乘性与次序有关。一般的即使 A, B 可乘，B, A 也未必可乘。

3. 矩阵乘法一般不满足交换律。即使 A, B 与 B, A 都可乘，也可能会有

$$AB \neq BA.$$

另外我们注意到，两个 n 阶方阵总是可相加且可乘的。

我们习惯用一些记号来表达一些特殊的矩阵。在矩阵的行和列已经确定的情况下，我们用

$$E_{ij}$$

来表示 i 行 j 列元素为 1，其余位置元素为 0 的矩阵。通过计算容易得到在 n 阶方阵的情况时，

$$E_{ij}E_{kl} = \delta_{jk}E_{il}.$$

对于 $n \times m$ 矩阵 $A = (a_{ij})_{n \times m}$，我们可以将其重新表达为

$$A = \sum_{i,j} a_{ij} E_{ij}.$$

在 n 阶方阵 $M_n(k)$ 中有两个特别的元素。称下面对角元全为 1 非对角元全为 0 的方阵

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

为**单位矩阵**，记作 I_n。对任意常数 c，我们以

$$cI_n$$

表示矩阵 I_n 每个元素乘以 c 得到的矩阵

$$\begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c \end{pmatrix}.$$

我们称之为**标量（倍数）矩阵**

需要注意当 $n > 1$ 时，标量 c 与 I_n 在矩阵乘法意义下不是可乘的，因而 cI_n 仅作为一个形式符号使用。称下面元素全为零的方阵

$$\begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}$$

为**零矩阵**，记作 O。零矩阵也可以理解为一个“零倍”标量矩阵：

$$O = 0I_n.$$

I_n 和 O 分别起到乘法和加法单位元的作用，即

$$A \cdot I_n = A,\ I_n \cdot A = A,\ A + O = O + A = A.$$

容易看到，单位矩阵和零矩阵分别可以类比于特殊的数字 1 和 0。另外我们看到

$$I_n = E_{11} + \cdots + E_{nn}.$$

2.1.2 数乘，转置，和矩阵的逆

对任何一个 $n\times m$ 矩阵 $A=(a_{ij})_{n\times m}$，以及一个标量 c，将每个矩阵元素都乘以 c 得到的新矩阵

$$\begin{pmatrix} ca_{11} & \cdots & ca_{1m} \\ \cdots & \cdots & \cdots \\ ca_{n1} & \cdots & ca_{nm} \end{pmatrix}$$

为 c 与 A 的一个**数乘 (scalar multiplication)**，也简记为

$$cA.$$

我们看大前面定义的标量矩阵 cI_n 可以看作数乘的一个特例。和标量矩阵 cI_n 类似的是，当 $n>1$ 时标量 c 与矩阵 A 在矩阵乘法意义下也不是可乘的，因而 cA 也仅仅是一个形式记号。但是我们看到利用标量矩阵 cI_n 乘法，仍可以实现一般矩阵的数乘，即有

$$cA=(cI_n)A=A(cI_m).$$

进一步的，我们将可加矩阵 A,B 的相减定义为加法和数乘导出的一个运算，即：

$$A-B:=A+(-1)B,$$

因而不再对其多做分析。

特别的我们看到，若将数字标量视作 1×1 矩阵，相应的加法与乘法运算与原先数字的加法和乘法运算一致。

关于矩阵的大部分常见运算都可以通过矩阵加法和乘法来实现，但矩阵还有一些较为特别的操作。对任意给定一个 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

称 $m\times n$ 矩阵

$$A^t:=\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots \\ a_{1m} & \cdots & a_{nm} \end{pmatrix}$$

为 A 的**转置 (transpose)**，即 A^t 的 (j,i) 位置元素等于原矩阵 (i,j) 位置元素。特别的，任意一个 n 维列向量 u，经过转置之后得到一个 n 维行向量 u^t。

我们可以通过定义直接验证以下常用的矩阵乘法与转置之间的关系：

命题 7.

$$(A^t)^t = A, (AB)^t = B^t A^t.$$

例 6. 根据定义，两个 n 维列向量

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

在 $n \geq 2$ 时不是可乘的。但是转置 u, v 得到的行向量 u^t, v^t，使得 u^t, v 与 v, u^t 总是可乘的。特别的由直接计算，他们的乘积由对应分量相乘后相加：

$$u^t v = u_1 v_1 + \cdots + u_n v_n = v^t u$$

进而得到的是一个标量，有时也称作 n 维列向量 u, v 的**标量积** ***(scalar product)***。在无歧义时也简写为

$$(u, v)$$

或者

$$\langle u^t, v \rangle$$

我们已经对矩阵定义了类似数的“四则运算”中的“加减乘”三种运算。对矩阵也存在类似“倒数”的概念，即矩阵**逆**。对于一个标量 a，我们知道它具有乘法逆 a^{-1} 当且仅当 $a \neq 0$，对于矩阵的情况，我们也需要首先分析“可逆性”，但它比标量的情况复杂得多。

首先注意到，由于矩阵乘法的非交换性，一般情况下我们还需要区分左逆和右逆。

定义 13. 对于一个 $n \times m$ 矩阵 A，若存在 $m \times n$ 方阵 B 使得

$$BA = I_m,$$

则称 B 为 A 的左逆 *(*矩阵*)*，否则则称 A 左不可逆；类似的若存在 $n \times m$ 方阵 B' 使得

$$AB = I_n,$$

则称 B 为 A 的右逆 *(*矩阵*)*，否则则称 A 右不可逆。

我们先给出存在单侧逆的必要条件。

定理 2. 若 A 存在左逆，则 $n \geq m$；若 A 存在右逆，则 $m \geq n$。

从以上定理可以看到，A 既存在左逆又存在右逆的必要条件是 A 为**方阵**。从运算可定义性的角度来看，讨论同阶方阵也是的加法和乘法运算总能够进行，因而进一步对“逆”的讨论，将暂时集中在方阵情况。

从概念上看，左逆与右逆的存在性是相互独立的。但是对于方阵这一特殊情况，存在如下的“意外巧合”：

定理 3. *方阵 A 存在左逆当且仅当 A 存在右逆。*

定理 1 和定理 2 的成立并不显然，此处不能通过简单计算得到，因而我们暂时承认它，并且将证明放在 2.6 节和 2.11 节讨论。在证明它们之前，我们也不需要利用它们证明相关结论。

但是我们容易推出如下结论。

定理-定义 1. *若方阵 A 存在左逆或存在右逆，则 A 的左逆和右逆相同，因而在任何一种情况下我们称 A* **可逆** *，并且左逆和右逆统称为 A 的***逆**，*记作*

$$A^{-1}.$$

否则，则称 A **不可逆**。

通过直接计算我们发现，若 B, B' 分别为 A 的左逆和右逆:

$$BA = AB' = I_n$$

则利用结合律和 I_n 作为乘法单位元，

$$B = BI_n = B(AB') = (BA)B' = I_nB' = B',$$

即当左右逆均存在时，它们一定相同。再次由结合律，我们发现

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n,$$

以及类似的

$$(B^{-1}A^{-1})(AB) = I_n,$$

因此得到以下重要结论：

命题 8. *若 A, B 都是可逆矩阵，则 AB 也是可逆矩阵，且*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

n 阶可逆矩阵全体在乘法下构成一个群，称作n **阶一般线性群** ***(general linear group)***，*记作 $GL_n(k)$。*

由此我们得到可逆矩阵集合在乘法下是**封闭**的。

方阵的可逆性，是线性代数基本的问题之一。另外，对于非方阵情形，人们也希望找到类似逆矩阵的推广，这其中就包括了重要的 **矩阵的广义逆**，例如在 $k=\mathbb{C}$ 是复数域情况时有著名的 Moore-Penrose 广义逆，我们将其放在应用部分加以讨论。

2.1.3 矩阵分块，标量积

在考虑矩阵矩阵时，有时它们以 **分块矩阵**的形式出现，有时也称为对一个矩阵的**分块**。通俗的来说，分块矩阵进一步放宽了矩阵中元素的要求，使得形状较小的子矩阵也可称为大矩阵的“矩阵元素”。

设 $n_1,\cdots,n_l,m_1,\cdots,m_k$ 是一组非负整数，$A_{i,j},\ i=1,\cdots,l,\ j=1,\cdots,k$ 是 $l\times k$ 个矩阵，满足：对任意 i, A_{ij} 具有相同的行数 n_i；对任意 $j, A_{i,j}$ 具有相同的列数 m_j，则容易观察到由 A_{ij} 按如下方式依次排布

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \cdots & \cdots & \cdots \\ A_{l1} & \cdots & A_{lk} \end{pmatrix} \tag{1}$$

构成了一个 $n\times m$ 矩阵，其中 $n=\sum_i n_i, m=\sum_j m_j$。我们称 A 是一个**分块矩阵**。

在实际应用当中，我们也经常利用相反方向的描述，即将一个 $n\times m$ 大矩阵 A 根据具体情况进行**分块**。若

$$n=n_1+\cdots+n_l,\ m=m_1+\cdots+m_k$$

是 n,m 的有序整数划分，则这一整数划分自然的决定了形如 (1) 的分块矩阵。

分块矩阵和通常矩阵具有类似的运算规则。我们可以直接验证：

命题 9. 设两个分块矩阵

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \cdots & \cdots & \cdots \\ A_{l1} & \cdots & A_{lk}, \end{pmatrix} B=\begin{pmatrix} B_{11} & \cdots & B_{1k} \\ \cdots & \cdots & \cdots \\ B_{l1} & \cdots & B_{lk} \end{pmatrix}$$

满足对所有 i,j，子矩阵 A_{ij}, B_{ij} 可相加，即 A_{ij}、B_{jt} 的具有相同的行数和列数，则它们的和等于分块矩阵

$$A+B=\begin{pmatrix} C_{11} & \cdots & C_{1k} \\ \cdots & \cdots & \cdots \\ C_{l1} & \cdots & C_{lk} \end{pmatrix}$$

其中

$$C_{ij} = A_{ij} + B_{ij}.$$

类似的，设两个分块矩阵

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \cdots & \cdots & \cdots \\ A_{l1} & \cdots & A_{lk}, \end{pmatrix} B = \begin{pmatrix} B_{11} & \cdots & B_{1s} \\ \cdots & \cdots & \cdots \\ B_{k1} & \cdots & B_{ks} \end{pmatrix}$$

满足对所有 i, j, t，子矩阵 A_{ij}, B_{jt} 可相乘，即 A_{ij} 的列数等于 B_{jt} 的行数，则 A, B 可相乘，且它们的乘积等于分块矩阵

$$AB = \begin{pmatrix} C_{11} & \cdots & C_{1s} \\ \cdots & \cdots & \cdots \\ C_{l1} & \cdots & C_{ls} \end{pmatrix}$$

其中

$$C_{it} = \sum_j A_{ij} B_{jt}.$$

简而言之，分块矩阵的加法和乘法运算与通常矩阵加法和乘法具有相通的运算规则，但是需要特别注意的是

1. 对分块矩阵做加法或乘法时，对应的子矩阵需要满足可加或可乘性。

2. 子矩阵的乘法顺序和大矩阵的乘法顺序应保持一致。

例 7. 我们有时需要将矩阵看作由行向量或者列向量构成的分块矩阵。例如对于 $n \times m$ 矩阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

和 $m \times l$ 矩阵

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{ml} \end{pmatrix},$$

我们可以将他们分别看作由 n 个 m 维行向量：

$$a_1^t, \cdots, a_n^t,\ a_i^t := (a_{i1}, \cdots, a_{im})$$

和 l 个 m 维行向量：

$$b_1, \cdots, b_l,\ v_i := (b_{mi}, \cdots, b_{mi})^t$$

作为子矩阵构成的分块矩阵：

$$A=\begin{pmatrix} a_1^t \\ \vdots \\ a_n^t \end{pmatrix}, B=(b_1,\cdots,b_l).$$

我们看到 A,B 的分块使得它们相应的子矩阵也是对应可乘的，于是 AB 可以重新表达为相对简洁的形式：

$$AB=\begin{pmatrix} a_1^t b_1 & \cdots & a_1^t b_l \\ \cdots & \cdots & \cdots \\ a_n^t b_1 & \cdots & a_n^t b_l \end{pmatrix}$$

即，乘积矩阵 AB 的 i 行 j 列分量等于 m 维列向量 u_i, v_j 的标量积 $u_i^t v_j$ 。

2.1.4 总结

以上我们看到，本节给出的矩阵相关概念，暂时只能看作是一种形式的定义。矩阵运算的数学实质，是通过域上的线性空间及线性变换体现出来的。后面我们会看到，在线性代数中，矩阵可以给出三种意义完全不同对象的表示。它们的具体对应，以及相互之间区别与联系的讨论，将在相应章节进行：

1. 线性空间中一组向量在取定基底之后的坐标表示（[2.3 节]）。

2. 线性变换在取定基底之后的坐标表示（[2.5 节] ）。

3. 双线性型在取定基底之后的坐标表示（[2.9 节] ）。

2.2 线性方程组求解与 Gauss 消元法

2.2.1 线性方程组

方程求解是古典数学的核心问题之一，在一般类型方程当中，具有**线性性**的方程是相对较容易求解的一类。在实际应用当中我们会经常遇到如下形式的求解问题：给定已知量 nm 个已知量 $a_{ij}, i=1,\cdots,n, j=1,\cdots,m$，以及另外 n 个已知量 $b_1,\cdots,b_n$，需要求解关于 m 个未知量 $x_1,\cdots,x_m$ 的方程组：

$$\begin{cases} a_{11}x_1+\cdots+a_{1m}x_m=b_1 \\ \vdots \\ a_{n1}x_1+\cdots+a_{nm}x_m=b_n. \end{cases} \tag{2}$$

我们称方程组 (2) 为**线性（非齐次）方程组**。

若利用上一节的矩阵记号:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \cdots \\ x_m \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \cdots \\ b_n \end{pmatrix},$$

则方程组 (2) 可以用矩阵形式表达为

$$Ax = b. \tag{3}$$

我们还常常利用 A 的行向量分块形式，将方程组 (2) 表达为以下两种形式：第一种方式，是对 A 做行向量分块，将 (3) 写作由以下 n 个标量积形式约束构成的方程组：

$$\begin{cases} a_1^t x = b_1 \\ \cdots \\ a_n^t x = b_n. \end{cases}, A = \begin{pmatrix} a_1^t \\ \vdots \\ a_n^t \end{pmatrix} \tag{4}$$

第二种方式，是对 A 的列向量分块，将 (3) 写作一个关于列向量线性组合约束的形式：

$$x_1 v_1 + \cdots x_m v_m = b, \, A = (v_1, \cdots, v_m) \tag{5}$$

研究方程组 (3) 的求解，是引入线性代数核心概念的主要动机和出发点之一。而等价形式 (4) (5) 还将进一步引申出其他相关问题。

关于 (2) 我们面临两个基本问题：

1. 如何判定方程组 (2) 何时有解？

2. 若方程组 (2) 有解，其解的结构有何形状？

从形式上来看，线性方程组 (2) 中的整数 n, m 分别代表了约束和（未知）自变量的个数。从直观上看，约束越多或者自变量个数越少，方程越“自由”，解集倾向于“变大”，反之则解集倾向于“变小”。然而实际情况经常体现出和此直观相悖的复杂性。我们从最简单的例子做一观察。

例 8. 方程组

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 0 \end{cases}$$

当中，自变量的个数 $m=3$ 大于约束个数 $n=2$，然而容易看出若此方程有解，两式联立将会导致

$$\text{“}1=0\text{”}$$

的矛盾，因此方程组解集为空。

例 9. 方程组

$$\begin{cases} x_1+x_2=1 \\ 2x_1+2x_2=2 \\ 3x_1+3x_2=3 \end{cases}$$

当中，自变量的个数 $m=2$ 小于约束个数 $n=3$。然而此方程组实际上有无穷多解，且其通解可以由参数化

$$x_1=1-t, x_2=t$$

给出。实际上容易看出，三个约束方程中的任何一个约束都可以推出另外两个约束自动成立，因而“真实”约束仅仅有 *1* 个而非 *3* 个。

由以上两个例子看出，在相对较为简单的方程组中，“何时有解”的问题就已经体现出了复杂性。我们看到一般情况下，约数个数 n 和自变量个数 m 之间的大小关系甚至无法准确体现解集的大小，因而感知到解集的情况进一步的与系数矩阵 A 和 b 的数值相关。在后面我们会看到，**维数**，以及由维数引出的 **秩**的概念，才是真正便于体现解集基本性质的内蕴量。

2.2.2 初等（行）变换和初等矩阵

本节将首先从算法角度，给出线性方程组 (2) 求解的一般方法, 即 **Gauss 消元法**。利用它可以直接给出线性方程组求解的完整解答。

Gauss 消元的基本思想，即利用不改变方程组解集的**初等变换**，使得方程组中的系数含有尽量多的零元素，最终转化为形式简单可直接进行求解的方程组。

Gauss 消元法即分为两步骤：

1. 将原方程化为形如 $\binom{A'}{O}x=b'$ 的形式，其中 A' 为每行元素不全为零的(行) 阶梯型矩阵。
2. 判定步骤 1 中的方程 $\binom{A'}{O}x=b'$ 是否有解，若有解则进行求解。

我们先来描述步骤 1，其核心是下面将要介绍的**初等变换**。初等变换是一类较为简单且具有**可逆性**的变换，并且它们不改变方程的解集。

首先注意到以下显然的命题

命题 10. 将方程组 *(2)* 当中任何一行乘以非零常数得到的新方程组，与原方程组具有相同的解集。具体的来讲，对任意整数 $1 \le i \le n$, 以及任意常数 $c \neq 0$, 令

$$A' = \begin{pmatrix} a'_{11} & \cdots & a'_{1m} \\ \cdots & \cdots & \cdots \\ a'_{n1} & \cdots & a'_{nm} \end{pmatrix}, b' = \begin{pmatrix} b'_1 \\ \cdots \\ b'_n \end{pmatrix},$$

其中的元素满足对所有的 $k \neq i, l = 1, \cdots, m$, 有

$$a'_{kl} = a_{kl}, b'_k = b_k,$$

以及

$$a'_{il} = ca_{il}, b'_i = cb_i,$$

则 x 是 $Ax = b$ 的解当且仅当 x 是 $A'x = b'$ 的解。

显然，将方程 $A'x = b'$ 的第 i 行乘以 c^{-1} 便恢复为原方程 $Ax = b$，因而两组方程具有相同解集。

定义 14. 将矩阵一行元素乘以同一个非零数 c 变换称为矩阵的**第一类初等（行）变换**。

类似的

命题 11. 交换方程组 *(2)* 当中任何两行的到的新方程组，与原方程组具有相同的解集。具体的来讲，对任意一对整数 $1 \le i < j \le n$, 令

$$A' = \begin{pmatrix} a'_{11} & \cdots & a'_{1m} \\ \cdots & \cdots & \cdots \\ a'_{n1} & \cdots & a'_{nm} \end{pmatrix}, b' = \begin{pmatrix} b'_1 \\ \cdots \\ b'_n \end{pmatrix},$$

其中的元素满足对所有的 $k \neq i, j, l = 1, \cdots, m$, 有

$$a'_{kl} = a_{kl}, b'_k = b_k,$$

以及

$$a'_{il} = a_{jl}, b'_i = b_j,$$

$$a'_{jl} = a_{il},\, b'_j = b_i.$$

则 x 是 $Ax = b$ 的解当且仅当 x 是 $A'x = b'$ 的解。

直观来看，新的方程 $A'x = b'$ 仅仅是调换了方程组中方程的次序，因而与原方程具有相同解集。

定义 15. 称交换矩阵 $i, j (i \neq j)$ 两行元素的操作为矩阵的**第二类初等（行）变换**。

但是注意到以上两类显然的初等变换还并不能化简一般线性方程组为阶梯型方程组，因为以上两种变换不能使系数矩阵中产生更多的零元素。我们需要引入进一步的**消元**操作。

例 10. 考虑二元方程组

$$\begin{cases} x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{6}$$

其中 $a_{11} = 1$。将第一行两侧乘以 $-a_{21}$ 并加到第二行，使得左下角元素变为 0 而得到阶梯型方程组

$$\begin{cases} x_1 + a_{12}x_2 = b_1 \\ 0x_1 + (a_{22} - a_{12}a_{21})x_2 = b_2 - a_{21}b_1 \end{cases} \tag{7}$$

另外我们注意到从方程组 *(7)* 还可以复原到方程组 *(6)*，即第一行两侧乘以 a_{21} 加到第二行上，因此以上两个方程组是等价的。

以上的特殊情况可以直接推广到多变量情况

命题 12. 将 *(2)* 当中第 j 行乘以一个数 c 后加到第 j 行 $(j \neq i)$ 得到的新方程组，与原方程组具有相同的解集。具体的来讲，对任意一对整数 $1 \leq i, j \leq n$, $i \neq j$, 令

$$A' = \begin{pmatrix} a'_{11} & \cdots & a'_{1m} \\ \cdots & \cdots & \cdots \\ a'_{n1} & \cdots & a'_{nm} \end{pmatrix}, b' = \begin{pmatrix} b'_1 \\ \cdots \\ b'_n \end{pmatrix},$$

其中的元素满足对所有的 $k \neq j, l = 1, \cdots, m$, 有

$$a'_{kl} = a_{kl},\, b'_k = b_k,$$

以及

$$a'_{jl} = a_{il}c + a_{jl},\, b'_j = b_i c + b_j.$$

则 x 是 $Ax = b$ 的解当且仅当 x' 是 $A'x = b'$ 的解。

定义 16. 将矩阵第 j 行元素乘以同一个数 c 后加到第 i 行 $(j \neq i\ )$ 的变换称为矩阵的**第三类初等（行）变换**。所有三类初等行变换统称为**初等（行）变换**。

注. 需要注意的是在第三类初等（行）变换的定义中，需正确理解“将矩阵第 i 行元素乘以同一个数 c 后……”一句的语义，即第三类初等行变换**不改变**第 i 行的元素，而是利用第 i 行的元素改变第 j 行元素的值。

下面我们用矩阵乘法来表示初等行变换，他们在进一步的分析中起到关键作用。

命题 13. 给定一个 $n \times m$ 矩阵 A。则对矩阵第 i 行做乘以非零常数 c 的第一类初等变换得到的新矩阵，等于 n 阶方阵 $M_i(c)$ 左乘 A：

$$M_i(c)A,$$

其中

$$M_i(c) = \sum_{j \neq i, 1 \leq j \leq n} E_{jj} + cE_{ii}.$$

命题 14. 给定一个 $n \times m$ 矩阵 A。则交换矩阵第 $i, j(i \neq j)$ 行的第二类初等变换得到的新矩阵，等于 n 阶方阵 P_{ij} 左乘：

$$P_{ij}A,$$

其中

$$P_{ij} = \sum_{k \neq i, j, 1 \leq k \leq n} E_{kk} + E_{ij} + E_{ji}.$$

命题 15. 给定一个 $n \times m$ 矩阵 A。则通过将第 j 行元素乘以 c 加到第 $i(i \neq j)$ 行上的操作第三类初等变换得到的新矩阵，等于 n 阶方阵 $T_{ij}(c)$ 左乘：

$$T_{ij}(c)A,$$

其中

$$T_{ij}(c) = I_n + cE_{ij}$$

命题 8，9，10 中的矩阵 $M_i(c)(c \neq 0), P_{ij}(i \neq j), T_{ij}(c)(i \neq j)$ 也称为**初等矩阵**。

我们已经看到初等矩阵左乘对应的初等变换都是“可逆”的变换。进一步的，通过直接计算我们得到初等变换的“可逆性”对应初等矩阵的可逆性。

命题 16. 初等矩阵 $M_i(c)(c \neq 0), P_{ij}(i \neq j), T_{ij}(c)(i \neq j)$ 都是可逆矩阵，且

$$M_i(c) = M_i(c^{-1}),\ P_{ij}^{-1} = P_{ij}, T_{ij}(c)^{-1} = T_{ij}(c).$$

容易看到，初等矩阵的逆矩阵左乘，对应的恰好是初等变换的逆变换。

更一般的，我们可以将以上初等行变换的相关结论，推广到可逆矩阵左乘变换的情况

命题 17. *若 T 是一个可逆矩阵且 T, A 可乘，则*

$$Ax = b$$

与

$$TAx = Tb$$

具有相同的解集。特别的，初等矩阵左乘诱导的初等行变换，不改变方程组 (3) 的解集。

2.2.3 Gauss 消元法

为了方便起见，我们可以忽略自变量 x 构成的列向量，而仅仅考虑由系数矩阵 A 和列向量构成的 $n \times (m+1)$ 阶矩阵

$$(A|b)$$

称作关于线性方程组 (3) 的**增广矩阵 (augmented matrix)**，其中为表示示性质上的区别，常将增广矩阵前 m 列与最后一列之间用竖线分开。显然

$$T(A|b) = (TA|Tb),$$

从矩阵 $(A|b)$ 经由初等变换 T 变为新矩阵 $(TA|Tb)$ 的过程，也常简写为

$$(A|b) \sim (TA|Tb).$$

定义 17. *一个矩阵 A 如果满足以下条件：*

1. *存在 r 使得所有满足 $i > r$ 的第 i 行全为零，且满足 $1 \leq j \leq r$ 的第 j 行至少有一个非零元素。*

2. *第 $j(2 \leq j \leq r)$ 行最左位置非零元素所处的列数，小于第 $i-1$ 行最左位置非零元所处的列数。*

*则称 A 为（行）**阶梯型矩阵 (row echelon form)**。其中非零行个数 r 称为 A 的**行秩**。非零行最左位置非零元素，称为这一行的**主元 (pivot)**，其位置称为**主元位置**。*

“阶梯”一词形象的指出了阶梯形矩阵的一般形状：

$$\begin{pmatrix} & \boxed{*} & * & * & * & * & * & * & * \\ & & \boxed{*} & * & * & * & * & * & * \\ & & & & & \boxed{*} & * & * & * \\ & & & & & & & \boxed{*} & * \\ & & & & & & & & \boxed{*} \\ & & & & & & & & \end{pmatrix}$$

其中无“星号”位置元素为 0，方形“盒子”标出的是主元位置。

下面的定理是利用初等行变换化简方程的核心。

定理 4. *所有矩阵 A 可以经过有限次初等行变换之后化为主元均为 1 的阶梯状矩阵。*

定理证明的核心即 **Gauss 消元法**，或者**主元 (pivoting) 消元法**。Gauss 消元法的核心步骤有三方面：

1. 利用第一类初等变换将主元“标准化”为 1。
2. 利用主元和第三类初等变换消元。
3. 遇到“意外”情况，通过第二类初等变换调整。
4. 遇到“退化”情况，进一步调整主元位置。

我们通过如下 3 阶矩阵的几个例子来展示 Gauss 消元法的具体过程中需要处理的情况。

例 11. *考虑一般的 3 元方程组*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_2 \end{cases} \tag{8}$$

并暂时假设 $a_{11} \neq 0$。相应的增广矩阵来表示为

$$\left(\begin{array}{ccc|c} \boxed{a_{11}} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_2 \end{array}\right)$$

其中用方框框出的非零元素 a_{11} 即为第一行的**主元**。首先将第一行乘以 a_{11}^{-1} 将主元化为 1:

$$\left(\begin{array}{ccc|c} \boxed{a_{11}} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_2 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & a'_{12} & a'_{13} & b'_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_2 \end{array}\right)$$

其中第一行 $a'_{1j} = \frac{a_{1j}}{a_{11}}, b'_1 = \frac{b_1}{a_{11}}$，然后，将第一行分别乘以 $-a_{-21}, -a_{31}$ 加到第二行、第三行上，使得主元下方元素全部为零：

$$\left(\begin{array}{ccc|c} \boxed{1} & a'_{12} & a'_{13} & b'_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_2 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & a'_{12} & a'_{13} & b'_1 \\ 0 & a'_{22} & a'_{23} & b'_2 \\ 0 & a'_{32} & a'_{33} & b'_2 \end{array}\right)$$

其中第二行第三行满足

$$a'_{ij} = a_{ij} - a'_{1j}a_{i1}, b'_i = b_i - b'_1 a_{i1}.$$

我们称此为以 $(1,1)$ 位置为主元的一次**消元**。我们发现新的增广矩阵所代表的线性方程组中，后两组方程已经不含第一个变量 x_1，从而可以进一步减少变量数而分析更小的子矩阵

$$\left(\begin{array}{cc|c} \boxed{a'_{22}} & a'_{23} & b'_2 \\ a'_{32} & a'_{33} & b'_2 \end{array}\right)$$

及其对应的方程，即例 5 中的情形。此时第一行对应的主元消元已经完成，进而选取原先主元位置 $(1,1)$ 的右下方临近位置 $(2,2)$ 作为下一步预定的主元位置。

在理想情况下，我们仿照例 6 中 $n = 3$ 的情况，反复利用主元和第一类、第三类初等变换，消去主元下方的非零元素。归纳的来看，若已经完成了第 i 行以 (i,j) 位置为主元的 Gauss 消元得到增广矩阵的分块形式

$$\left(\begin{array}{cc|c} A_{11} & A_{12} & b_1 \\ 0 & A_{22} & b_2 \end{array}\right) \tag{9}$$

其中 A_{11} 为 $i \times j$ 的阶梯矩阵，且 (i,j) 为矩阵的第 i 行主元，则在下一步我们考虑以 $(i+1, j+1)$ 位置为第 $i+1$ 行主元，即等价的针对子矩阵

$$(A_{22}|b_2)$$

以 $(1,1)$ 位置为这一子矩阵第一行的主元位置。不过在一般情况下，需要考虑到选取位置的元素为零的**意外**情况，此时需要利用第二类初等变换进行调整。

例 12. 仍以例 *6* 中的增广矩阵为例进行考虑：

$$\left(\begin{array}{ccc|c} \boxed{a_{11}} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_2 \end{array}\right)$$

我们看到若 $a_{11}=0$，需要进行如下调整：考虑 a_{11} 所在列的所有元素 $a_{i,1}, i=2,\cdots,n$，寻找非零元素 $a_{i1}\neq 0$，并利用第二类初等变换 P_{1i} 调换第一行和第 i 行的位置。例如假定 $a_{31}\neq 0$，则利用第二类初等矩阵 $P_{1,3}$ 得到

$$\left(\begin{array}{ccc|c} 0 & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{a_{31}} & a_{32} & a_{33} & b_3 \\ a_{21} & a_{22} & a_{23} & b_2 \\ 0 & a_{12} & a_{13} & b_1 \end{array}\right),$$

此时 $(1,1)$ 仍然为第一行主元所在位置，并且矩阵化为例 *6* 的情况。

注：以上需要利用第二类初等变换处理意外情况的操作，会体现在一些有用的矩阵分解当中，例如可逆方阵的 **Gauss 分解**。

在最一般的情况中，我们还需处理一些**退化**的情况。

例 13. 以分块增广矩阵 *(9)* 为例，我们须考虑到存在从下一步预定的主元位置 $(i+1,j+1)$ 开始朝下方全部元素为零，或等价的，子矩阵 A_{22} 第一列全部为零的可能性：

$$A_{22}=\left(\begin{array}{cccc|c} \boxed{0} & * & \cdots & & b_{i+1} \\ 0 & * & \cdots & & b_{i+2} \\ \vdots & \vdots & \cdots & & \vdots \\ 0 & * & \cdots & & b_n \end{array}\right)$$

此时交换行的办法已经失效，唯一的办法是**改变** A 第 $i+1$ 行主元的列位置，即列右移一步将 A 的 $(i+1,j+2)$ 位置调整为第 $i+1$ 行新的主元位置*:*

$$\left(\begin{array}{cccc|c} 0 & \boxed{*} & \cdots & & b_{i+1} \\ 0 & * & \cdots & & b_{i+2} \\ \vdots & \vdots & \cdots & & \vdots \\ 0 & * & \cdots & & b_n \end{array}\right)$$

在更极端的退化情况中，右移一次主元位置仍可能会出现全列为零的情况，因而需要继续按上面进行调整，直到找到非零元素或列出界。这一类型的调整既能保证整个矩阵最终呈阶梯形状，又能够递归操作并在有限步内停止。

Algorithm 1: Gauss 消元法

输入: 矩阵系数 $a[i,j]$，行数 n，列数 m。

输出: 矩阵 (A) 的行阶梯标准型

$i \leftarrow 0; j \leftarrow 0;$
while $i \leq n$ *and* $j \leq m$ **do**
  $i \leftarrow i+1; j \leftarrow j+1; i' \leftarrow i;$
  while $a_{ij} = 0$ *and* $j \leq m$ **do**
    $i' \leftarrow i'+1;$
    if $i' > n$ **then**
      $j \leftarrow j+1$;$i' \leftarrow i;$
  if $j \leq m$ **then**
    for $l = j$ *to* m **do**
      $swap(a[i,l], a[i',l])$;a[i,l]← a[i,l]/a[i,j];
    for $k = i+1$ *to* n **do**
      for $l = j$ *to* m **do**
        $a[k,l] \leftarrow a[k,l] - a[k,j]a[i,l];$

命题 18. 增广矩阵 $(A|b)$ 经过有限次初等行变换之后，可化为阶梯状矩阵

$$\left(\begin{array}{c|c} A' & b' \\ 0 & \tilde{b}' \end{array}\right)$$

其中 A' 是每一行均不为零的且主元素全为 1 的阶梯矩阵，因而 *(3)* 的解集等同于阶梯状方程

$$\begin{pmatrix} A' \\ 0 \end{pmatrix} x = \begin{pmatrix} b' \\ \tilde{b}' \end{pmatrix}$$

与定理 3 所陈述的情况稍有不同的是，对利用增广矩阵 $(A|b)$ 的初等变换求解 (2) 这一情况，我们在以上算法中的最右列为 A 的最右列，即第 m 列。当主元所处列 j 满足 $j > m$ 时算法即停止，而不需要继续对第 $m+1$ 列进行消元操作。

2.2.4 齐次线性方程组

命题 19. 阶梯状方程

$$\begin{pmatrix} A' \\ 0 \end{pmatrix} x = \begin{pmatrix} b' \\ \tilde{b}' \end{pmatrix}$$

有解当且仅当

$$\tilde{b}' = 0$$

首先我们注意到，若 $\tilde{b}' \neq 0$，则至少对某个 i 会出现无解的子方程

$$0x_i = \tilde{b}'_i, \tilde{b}'_i \neq 0,$$

因而“仅当”部分是显然的。其次若 $\tilde{b}' = 0$，则从第 $r+1$ 行开始是平凡成立的方程

$$0x_i = 0, i = r+1, \cdots, n$$

于是我们仅需要具体找到前 r 行非平凡方程组

$$A'x = b'$$

的一个解。注意到 A' 的主元全部为 1 且位于不同的列上。若假设第 i 行的主元位于第 l_i 列，则 r 个非平凡方程的形状为

$$x_{l_i} + \text{非主元对应自变量} = b'_i, i = 1, \cdots, r$$

因而方程存在一个显然的解

$$x_j = \begin{cases} b'_i, & \text{若} j = l_i \text{是第} i \text{ 行主元对应的列} \\ 0, & \text{若} j \text{不是某个主元对应的列} \end{cases}$$

由此我们也看到，行秩 r 给出了方程组 (3) 的“真实”约束个数，我们会在后面进一步加以讨论。

至此我们已经完全给出了问题 1 的回答。对于问题 2，通过上面对“主元”的分析，我们能够得到如下结论

命题 20. *若 $r = m$，则线性方程组 (3) 存在唯一解。*

但是对 $r < m$ 且有解的情况，我们还没有给出描述。下面我们将定方程组 (3) 存在解。

我们注意到若 $x, \tilde{x}$ 都是 (3) 的解，则他们的差 $x - \tilde{x}$ 满足方程

$$A(x - \tilde{x}) = 0,$$

将 (3) 中左侧向量组 b 消去。反之，若 y 是方程组

$$Ay = 0 \tag{10}$$

的一个解，则通过 (3) 的任何一个解 x 得到

$$\tilde{x} = x + y$$

也是 (3) 的一个解。因而对解空间的描述，可以进一步转化为对 (10) 解空间的研究。我们称 (10) 为关于 (3) 的**齐次线性方程组**。

对 (3)，通过简单计算我们知道若 $x, \tilde{x}$ 都是解，则它们的和

$$x + \tilde{x}$$

与标量倍数

$$cx$$

一般都不再是 (3) 的解。但是对齐次方程组 (10)，我们发现对任意两个解 y, y'，以及标量 c，由矩阵运算的分配律与标量乘法规则，得到

$$A(y + y') = Ay + Ay' = 0 + 0 = 0, \; A(c \cdot y) = c \cdot Ay = 0.$$

即 (10) 的解集在加法和标量乘法下，都是封闭的，而这一性质是 (3) 所不具备的。将 (10) 解集所具有的这一特性加以抽象化，便引入了重要的**线性空间**的概念。我们从前面的例子看到，一般情况下方程组 (3) 和 (10) 是可能存在无穷多解的，但在后面我们会看到，通过对（有限维）线性空间的讨论，最终可以将问题 2 的答案转化为一个“有限”问题。

在讨论完线性空间和线性变换的基本理论后，我们会在 2.6 节给出问题 2 的完整回答。

2.3 线性空间

2.3.1 线性空间的基本定义的实例

回忆第一节，我们引入了矩阵、矩阵乘法、（行或列）向量的概念。在第二节我们通过解线性方程组的问题，阐明了矩阵运算的一些含义。本节我们进一步引入了线性（向量）空间，和线性空间上的线性变换的概念，最终完成对上节末尾问题的回答。

在第一节当中我们定义了由 k 中的 n 个数 $a_1, \cdots, a_n \in k$ 组成的列向量

$$\begin{pmatrix} a_1 \\ \cdots \\ a_n \end{pmatrix}$$

的 n 维列向量空间。我们注意到对 k^n 可以做两种运算：

1. 两个列向量

$$u = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, v = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

可以做加法：

$$u + v = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix},$$

满足交换律和结合律，且存在一个特殊的元素“零向量”

$$0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

作为“加法单位元”：

$$u + 0 = 0 + u = u,$$

且对每个列向量 u 存在唯一的“加法逆” $-u$ 满足

$$u + (-u) = -u + u = 0.$$

2. 域 k 中的常数 $a \in k$ 可与 k^n 中的元素 u 做数乘：

$$a \cdot u = ua = \begin{pmatrix} au_1 \\ \vdots \\ au_n \end{pmatrix},$$

并且和 k 中自然的乘法，和列向量的自然加法运算相匹配：

$$(ab) \cdot u = a(b \cdot u).$$

需要注意的是如果将数乘看作矩阵乘法，那么常数应当进行右乘。我们进一步的抽象出线性空间的概念。

定义 18. 域 k 上的一个**线性空间**, 不妨记作 V，是一个带有加法二元运算

$$+ : V \times V \to V$$

和数乘的二元运算

$$\cdot : k \times V \to V$$

的集合，满足：

1. 零向量 0 为“加法单位元”，满足

$$u + 0 = 0 + u = u.$$

2. 存在“加法逆”$-u$ 满足

$$u + (-u) = -u + u = 0.$$

3. 加法满足交换律和结合律

$$u + v = v + u, (u + v) + w = u + (v + w).$$

4. k 中的乘法、数乘和向量加法运算“匹配”，即对任何 $a, b \in k, u, v \in V$ 有

$$(ab)u = a(bu), a(u + v) = au + av$$

且乘法单位元 1 的标量乘法为恒等变换：

$$1 \cdot v = v.$$

线性空间中的一个元素，也称作线性空间中的一个**向量**。线性空间的一个子集合，若在原线性空间的运算下也构成一个线性空间，则称这个子集为原线性空间的一个**线性子空间**，也简称为**子空间**。

在实际情况中，为了验证线性空间 V 的一个子集合 W 是一个子空间，我们不需要验证 W 满足定义 7 中线性空间的所有的性质，而仅仅需要验证其中部分条件，这是由于其余的条件继承自 V 的线性空间结构而自动成立。

命题-定义 2. 一个线性空间 V 的子集合 W 是一个子空间当且仅当 W 在加法和标量乘法下封闭，即对任意 $c \in k$, $w_1, w_2 \in W$ 有

$$cw_1, w_1 + w_2 \in W.$$

例 14. 在任何线性空间 V 中都包含零向量 0，零向量 0 自身构成 V 的一个子空间：

$$\{0\}$$

也称为**零空间**。因此任何两个子空间的交都必然包含零空间，因而交总是非空的。

命题 21. 线性空间 V 的两个子空间 V_1, V_2 的交

$$V_1 \cap V_2$$

仍是子空间。

但我们容易找到例子表明，子空间的并一般不是子空间 (为什么？)。

例 15. 由本节开始的讨论，n 维列向量空间 k^n 是 k 上的一个线性空间。特别的域 $k = k^1$ 自身是一维 k-线性空间。

注. 对于标量 $c \in k$ 与列向量 $v \in k^n$ 的标量乘法，为了和通常线性组合表达一致，我们通常都以第一节当中标量乘法

$$cv$$

进行表达，而不写作

$$`vc',$$

尽管后者在矩阵乘法上是恰当的形式。

例 16. 进一步的，$n \times m$ 矩阵集合 $M_{n\times m}(k)$，在数乘和矩阵加法下，也构成一个线性空间。

例 17. 回忆线性方程组 *(3)*

$$Ax = b.$$

当 $b \neq 0$ 时，方程的解集并不构成一个线性空间。但是相应的齐次线性方程组 *(10)*

$$Ay = 0$$

的解集构成一个 m 维列向量空间 k^m 的子空间。

我们需要注意线性空间公理与上面例子之间的关系：n 维列向量空间 k^n 给出的是抽象的公理化线性空间**范畴 (cateogyr)** 的一个具体**对象 (object)**。列向量空间这一“对象”保证了线性空间公理所定义是自洽且满足公理的集合是非空的，并且本节所有关于抽象线性空间的概念和命题，特别的都适用于列向量空间。后面我们还会看到，对任何一个抽象的有限维线性空间，都可以通过取定**基底**的办法来“等同于”一个具体的列向量空间。在讨论范畴论基本概念的部分，它还将作为我们的一个例子。

我们注意到线性空间的结构中，仅规定了标量与向量的乘法，而一般的来讲，并不存在“向量乘法”。这一问题将在关于张量得部分中进一步讨论。

我们观察到域的定义和线性空间的概念在属性上略有不同：域是一个绝对的概念，一个域不需要相对于其他对象来定义，而可以作为一个独立存在的对象；而线性空间是一个相对的概念，它需要相对于一个域 k 来进行讨论。很多情况下为了叙述上的方便，在默认基域 k 无歧义情况下，我们可以直接说“一个线性空间 V”。

2.3.2 线性组合，线性相关（无关）性，和线性空间的基底

在例 9 当中，我们通过其次线性方程约束来间接的定义了一个线性子空间，但更直接且便于分析的“制造”子空间的方式，则是由一些已知向量的**线性组合**来“生成”子空间。

任意给出向量空间 V 的一组向量 $v_1, \cdots, v_m$, 从线性空间的定义除法容易推出，若子空间 W 包含 $v_1, \cdots, v_m$，则

$$c_1v_1 + \cdots + c_mv_m, \, c_i \in k$$

也在 W 中。以上形式的向量称作向量组 $v_1, \cdots, v_m$ 的一个 **k-线性组合**，也简称为**线性组合**。进一步的，由线性空间的定义我们都可以得出

定理-定义 2. 线性空间 V 中的一个向量组 $v_1, \cdots, v_m$ 的所有线性组合构成的是 V 的一个子空间，称作由 $v_1, \cdots, v_m$ **张成的子空间**：

$$span\{v_1, \cdots, v_m\} := \{c_1v_1 + \cdots + c_mv_m | \, c_i \in k\}.$$

若进一步的有

$$V = span\{v_1, \cdots, v_m\},$$

则称向量组 $v_1, \cdots, v_m$ **张成 *(span)*** 线性空间 V

在此需注意，线性组合仅允许有限求和。

我们对单个向量，也有相应的概念。

定义 19. 我们称 $v \in V$ 可以被有限多个向量 $v_1, \cdots, v_m$ **线性表出**，若

$$v \in span\{v_1, \cdots, v_m\},$$

即存在 $c_1, \cdots, c_m \in k$ 使得

$$v = c_1v_1 + \cdots + c_mv_m.$$

c_i 也称作这一线性表示的**分量系数**。

例 18. 回忆线性方程组的形式 *(5)*。用线性表示的语言重新叙述，*(5)* 存在解当且仅当

$$b \in span\{v_1, \cdots, v_m\}.$$

我们也说一组向量（构成的向量组集合 S）可以被另一组向量 (构成的向量组集合 S'）**线性表出**，如果 S 的每个向量都可以被 S' 中的向量线性表出。

由定义可以直接推出“线性表出”具有单向传递性。

命题 22. 若第一组向量 $v_1, \cdots, v_l$ 可以线性表出第二组向量 $v'_1, \cdots, v'_m$，而第二组向量 $v'_1, \cdots, v'_m$ 又可以线性表出第三组向量 $v''_1, \cdots, v''_n$，则第一组向量 $v_1, \cdots, v_l$ 可以线性表出第三组向量 $v''_1, \cdots, v''_n$

定义 20. 若线性空间 V 中存在有限个向量 $v_1, \cdots, v_n$ 张成 V:

$$V = span\{v_1, \cdots, v_n\},$$

则称 V 是一个**有限维线性空间**。

我们会看到，有限维线性空间具有一般线性空间所不具备的特性。在本讲义中，如无特别声明，除去少量有用的例子外，我们仅仅讨论**有限维线性空间**。

从定义可以看到，对有限维线性空间 V，以及张成 V 的一组向量 $v_1, \cdots, v_n$，对任意向量 v 我们总可以找到 n 个数字 $c_1, \cdots, c_n$ 构成的 n 元数组来表示 v，从而向量的描述变成了一个“有限”问题。但仅仅这样做还存在一个明显的缺陷，即一般情况下，n 个数字 $c_1, \cdots, c_n$ 未必是唯一确定的，即表示可能存在一定“冗余”。这一方面的缺陷需要进一步线性相关（无关）的概念来弥补。

我们注意到零向量 0 总是可以被任何一组向量线性表出，即“平凡”的取分量系数全为 0，但我们更关心的是“非平凡”的情况。

定义 21. 称 V 上的有限多个向量 $v_1, \cdots, v_m$ **线性相关**，如果存在一组不全为零的数 $c_1, \cdots, c_m$ 使得 0 被 $v_1, \cdots, v_m$ 关于分量系数 $c_1, \cdots, c_m$ 的线性组合线性表出：

$$c_1 v_1 + \cdots + c_m v_m = 0.$$

反之则称它们线性无关，即若 0 被 $v_1, \cdots, v_m$ 线性表出

$$c_1 v_1 + \cdots + c_m v_m = 0,$$

则它一定是一个“平凡”的表示：

$$c_1 = \cdots = c_m = 0.$$

例 19. 例如我们对线性方程组 *(2)* 解的唯一性的讨论，归结为对齐次线性方程组 *(10)* 是否存在非零解的讨论。显然 *(2)* 的解（若存在）是唯一的，当且仅当 0 关于 A 的列向量，仅有平凡线性组合表示。

由定义可得到线性相关的另一个等价定义。

命题 23. 一组向量 $v_1, \cdots, v_m$ 线性相关，当且仅当其中的某一个向量是其余向量的线性组合，即存在 i 及 $m-1$ 个常数 $c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_m$ 使得

$$v_i = \sum_{j \neq i} c_j v_j.$$

类似的由定义出发，我们容易推出

命题 24. 一组线性无关的向量 $v_1, \cdots, v_m$ 的任何一个非空子集形成的向量组，仍然是线性无关的。

命题 25. v 可以被一组向量 $v_1, \cdots, v_m$ 唯一的线性表出，当且仅当

$$v \in span\{v_1, \cdots, v_m\}$$

且

$$v_1, \cdots, v_m$$

线性无关，即零向量 0 在向量组 $v_1, \cdots, v_m$ 之下仅有平凡表示。

下面的命题给出线性无关性与线性表出的一个关系。

命题 26. 若 $v_1, \cdots, v_m$ 线性无关，则 $v_1, \cdots, v_m, v$ 线性无关，当且仅当

$$v \notin span\{v_1, \cdots, v_m\}.$$

我们希望得到的，是线性空间 V 中，兼具“张成”和“线性无关”两方面特性的向量组。这就引入线性代数中最为重要的概念之一：线性空间的的**基底**。

定义 22. 称 V 的一个子集 $\mathcal{B} \subseteq V$ 为 V 的一组**基底** ***(basis)***，简称 **基**，如果 $\mathcal{B}$ 满足如下两条性质：

1. $\mathcal{B}$ 中任何有限多个元素都线性无关。

2. V 中任何元素都可被 $\mathcal{B}$ 中有限多个元素线性表出。

结合命题 19，我们得到

命题 27. V 的一个子集 $\mathcal{B}\subseteq V$ 为 V 的一组基底，当且仅当 V 中任何向量可以唯一的表示为 $\mathcal{B}$ 中元素的线性组合。

以下一般性定理表明线性空间总是存在一组基。

定理 5. 任何一个线性空间都存在一组基。

但由于一般情况下的证明方法在本课程中很少用到, 故我们仅证明**有限维线性空间**的情况，即：

命题 28. 若 V 是有限维线性空间, 即存在有限多个向量 $v_1,\cdots,v_m$ 张成 V，则存在向量组 $\{v_1,\cdots,v_m\}$ 的一个子集 $\mathcal{B}$ 使得 $\mathcal{B}$ 构成 V 的一组基底。

直观的来看，我么剔除掉 $\{v_1,\cdots,v_m\}$ 当中可以被其余向量线性表出的“冗余”向量即可。我们用如下方法来给出 $\mathcal{B}$ 的一种构造。首先设定

$$\mathcal{B}_0=\emptyset$$

为空集，并对 $k=0,\cdots,k-1$ 依次归纳的定义

$$\mathcal{B}_{k+1}:=\begin{cases}\mathcal{B}_k,\ \text{如果} v_{k+1}\in span\{\mathcal{B}_k\};\\ \mathcal{B}_k\cup\{v_{k+1}\},\ \text{如果} v_{k+1}\notin span\{\mathcal{B}_k\}.\end{cases}$$

由我们的取法，

$$span\{\mathcal{B}_k\}=span\{v_1,\cdots,v_k\}$$

对所有 $k=1,\cdots,m$ 都成立，特别的

$$span\{\mathcal{B}_m\}=span\{v_1,\cdots,v_m\}=V.$$

并且，由命题 20，$\mathcal{B}_m$ 中的元素都是线性无关的，因而

$$\mathcal{B}=\mathcal{B}_m$$

即为所求。

在本课程的大多数内容中我们都假定所讨论的线性空间是有限维的，但在某些场合仍会讨论一些重要的无限维线性空间，例如 k 上的多项式环 $k[x]$。

线性空间基底的存在性，还并不足以有效描绘线性空间的性质。另一方面我们会看到，即使是有限维线性空间，其基底也有不同的选择。但是对有限维线性空间，幸运的是任何一组基底具有**相同**元素个数，因而我们只需考虑一种基底的选择，并利用矩阵对线性空间进行分析和计算。

2.3.3 线性空间的维数与向量组的秩；Steinitz 替换引理及其推论

在许多教材中，是通过上节介绍的 Gauss 消元法，来证明基底具有相同元素个数的. 但此处我们利用所谓“替换引理”来构建这一基本结论，它可以摆脱矩阵运算，给出比较精简的证明。此外替换引理还被 S.Maclane 进一步推广并引入所谓矩阵胚 (matroid) 的概念。

定理 6 (Steinitz 替换引理)**.** *给定一个线性空间 V (不一定要求有限维)，若 $S := \{v_1, \cdots, v_m\}$ 是 V 中一组线性无关的向量，$S' = \{v'_1, \cdots, v'_n\}$ 是 V 中另一组向量，满足 $V = span\{v'_1, \cdots, v'_n\}$，则*

1. *$m \leq n$;*

2. *存在 S' 的一个子集 T 满足 $|T| = n - m$ 且 $S \cup T$ 张成 V。*

此引理中的 (2) 表明，张成 V 中的向量组 $v'_1, \cdots, v'_n$ 中的某 m 个元素可以被 m 个线性无关向量 $v_1, \cdots, v_m$ 替换掉而仍然张成整个空间，故得名。

需要注意到 $v'_1, \cdots, v'_n$ 张成 V 蕴含了 S' 可以线性表出 S。另外此定理包含了 $n = m$ 的情况，在这一情况下允许 (2) 中的 T 为空集。

证明. 我们对 m 进行归纳证明，并且我们约定当求和下限大于上限时，相应的和式为 0。当 $m = 1$ 时显然 (1) 成立，即

$$n \geq m = 1.$$

此时 $v_1 \neq 0$，且存在一组不全为 0 的数 $c'_1, \cdots, c'_n$ 使得

$$v_1 = \sum_{i=1}^{n} c'_i v'_i.$$

不妨设 $c'_1 \neq 0$，于是 v'_1 可以被 $v_1, v'_2, \cdots, v'_n$ 线性表出：

$$v'_1 = \frac{1}{c'_1}(v_1 - \sum_{i=2}^{m} c'_i v'_i).$$

因而取 $T = \{v'_2, \cdots, v'_n\}$ 即满足条件，即

$$v_1, v'_2, \cdots, v'_n$$

张成 V，故证明了 (2)。

假设引理对 $k = m - 1 (m \geq 2)$ 的情况都已经成立，现在证明引理对 $k = m$ 也成立。我们首先证明 (1)，即若 $v_1, \cdots, v_m$ 线性无关，且它们可以被 $v'_1, \cdots, v'_n$ 线性表出，则 $m \leq n$。我们注意到 $v_1, \cdots, v_{m-1}$ 也线性无关，因而 $n \geq m - 1$，

否则 $v'_1, \cdots, v'_n$ 线性表出 $v_1, \cdots, v_{m-1}$ 将与 $k = m-1$ 时的归纳假设矛盾。我们现在证明 $n = m-1$ 也是不能成立的，即一定有 $n \geq m$.

如若不然，则存在 $m-1$ 个向量 $v'_1, \cdots, v'_{m-1}$ 可以线性表出 $v_1, \cdots, v_m$，显然它们也可以线性表出 $v_1, \cdots, v_{m-1}$。由归纳假设关于 $k = m-1$ 情况（2）中的部分，$v'_1, \cdots, v'_{m-1}$ 也可以被 $v_1, \cdots, v_{m-1}$ 线性表出，即这两组向量可以相互线性表出：

$$span\{v_1, \cdots, v_{m-1}\} = span\{v'_1, \cdots, v'_{m-1}\}.$$

此时 $T = \emptyset$ 是空集。由线性表出的单向传递性，$v_1, \cdots, v_{m-1}$ 可以线性表出 v_m，这与 $v_1, \cdots, v_m$ 线性无关矛盾，故（1）部分得到证明。

下面进一步证明 (2),即线性无关的向量组 $v_1, \cdots, v_m$ 总可以通过添加 $n-m$ 个 $v'_1, \cdots, v'_n$ 中的元素而张成 V。由 (1) 我们知道 $n \geq m$。由归纳假设关于 $k = m-1$ 的情况，通过添加 $n-m+1$ 个 $v'_1, \cdots, v'_n$ 中的元素可以使得 $m-1$ 个向量 $v_1, \cdots, v_{m-1}$ 与这 $n-m+1$ 个元素张成 V. 适当调整指标后，我们不妨假设这 $n-m+1$ 个元素是 v'_j, $j = m, \cdots, n$，即成立

$$span\{v_1, \cdots, v_{m-1}, v'_m, \cdots, v'_n\} = span\{v'_1, \cdots, v'_{m-1}, v'_m, \cdots, v'_n\} = V.$$

于是存在 $c_1, \cdots, c_{m-1}, c'_m, \cdots, c'_n \in k$ 使得

$$v_m = c_1 v_1 + \cdots + c_{m-1} v_{m-1} + c'_m v'_m + \cdots + c'_n v'_n.$$

由于 $v_1, \cdots, v_m$ 线性无关，必然存在 $c'_j, j \in \{m, \cdots, n\}$ 使得 $c'_j \neq 0$。通过调整下标我们不妨假设 $c'_m \neq 0$，于是

$$v'_m = \frac{1}{c'_m}(v_m - \sum_{i=1}^{m} c_i v_i - \sum_{i=m+1}^{n} c'_i v'_i).$$

因而我们取 $n-m$ 个向量构成的向量组

$$T = \{v'_{m+1}, \cdots, v'_n\}$$

即为所求。

由 Steinitz 引理我们立即得到：

定理 7. 任何一个有限维线性空间 V 的任何一组基其元素个数都是相同的。这个相同的正整数若为 n, 则称 V 为 n 维线性空间，记作

$$dim(V) = n.$$

需要注意的是以上讨论均要求空间是非零空间。对于退化情况的零空间 $\{0\}$，我们规定

$$dim(\{0\}) = 0.$$

另外两个常用的推论是

命题 29. *若 $m > n$，则 n 维线性空间中的任意 m 个向量都是线性相关的。*

命题 30. *若 m 个向量 $v_1, \cdots, v_m$ 张成 V，则*

$$dim(V) \le m.$$

例 20. 虽然前面我们已经先于“维数”概念引入了“n 维列向量空间” k^n 这一例子，但是我们还需要验证

$$dim(k^n) = n.$$

事实上，我们可以找到 k^n 的一组显然的基底：

$$\{e_i | i = 1, \cdots, n\}, e_i = (0, \cdots, \underbrace{1}_{\text{第}i\text{个位置}} \cdots, 0)^t.$$

容易它们满足定义中的两条性质。有时我们称 $\{e_i\}$ 为 k^n 的一组**标准基底**。

将命题 24 应用于列向量空间得情况，我们得到：

命题 31. *当 $m > n$ 时，任意 m 个 n 维列向量空间中的向量 $v_1, \cdots, v_m$ 都线性相关。等价的，$m > n$ 时，任意以矩阵*

$$A = (v_1, \cdots, v_m)$$

为系数，的 m 元齐次线性方程组

$$Ay = 0$$

都存在非零解。

事实上，我们利用 Gauss 消元法也可以很容易的证明以上命题，由此可以给出定理 6 的另一个证明。

通过以上的分析，我们可以看到抽象的线性空间相比于具体的列向量空间在分析上的一个优点：在很多问题当中，即使是对具体的列向量空间，我们也不必关注其具体元素，而仅仅需要利用它们所满足的的公理性质。

由维数的概念，可以引申出“秩”的概念。

定义 23. 给定一组向量 $v_1, \cdots, v_m$,我们也称他们张成的线性空间 $span\{v_1, \cdots, v_m\}$ 的维数

$$dim(span\{v_1, \cdots, v_m\})$$

为这一向量组的**秩**，记作

$$r(v_1, \cdots, v_m).$$

特别的，若一个 $n \times m$ 矩阵 A 以列向量分块形式表示 m 个 n 维列向量 $v_1, \cdots, v_m$:

$$A = (v_1, \cdots, v_m),$$

则称向量组 $v_1, \cdots, v_m$ 的秩为矩阵 A 的 **列秩**；类似的，若 A 以列向量分块形式表示 n 个 m 维行向量 $u_1^t, \cdots, u_n^t$:

$$A = \begin{pmatrix} u_1^t \\ \vdots \\ u_n^t \end{pmatrix},$$

则称向量组 $u_1^t, \cdots, u_n^t$ 的秩为矩阵 A 的 **行秩**。

在下一节，我们会用另一种方式定义矩阵 A 的秩，并最终证明这同一个矩阵在不同定义下的秩，是同一个数值。我们还可以推出如下结论：

命题 32.

$$r(v_1, \cdots, v_m) \leq m.$$

我们注意到，有限维线性空间的基是没有“标准”选择的。由 Steinitz 引理可以得到另一重要结论：

定理 8. 若 $v_1, \cdots, v_m$ 为 n 维线性空间 V 上一组线性无关向量，$m \leq n$，则可添加 $n - m$ 个向量 $v_{m+1}, \cdots, v_n$ 使得 $v_1, \cdots, v_n$ 为 V 的一组基。

由同样的原理，上面的定理也包含了 $m = n$ 的情况，即“不添加任何元素”。

以上的两个定理是整个线性代数中最为基本的结论，本讲义中的绝大部份结论都可以以此为基础而建立。我们看到它们的证明仅仅利用到了 k 是域，以及线性空间的公理化定义这两样事实。

我们通过下面的一些命题看到，线性空间的包含关系与维数大小关系相匹配，因而维数可以有效刻画线性空间的包含关系。

命题 33. 若 W 是 V 的一个子空间，则

1. $dim(W)=0$ 当且仅当 $W=\{0\}$。

2.
$$dim(W)\le dim(V),$$
并且等号成立，当且仅当 $V=W$。

2.3.4 向量与向量组的矩阵表示

最后我们在描述如何从抽象的线性空间回到具体的由列向量组成的线性空间。假设

$$\mathcal{B}_V:=\{v_1,\cdots,v_m\}$$

是 n 维线性空间 V 的一组基底。由基的性质，对任何一个向量 $v\in V$ 都存在唯一的一组数 $c_1,\cdots,c_m\in k$ 使得

$$v=c_1v_1+\cdots+c_nv_m.$$

因而在取定一组基之后，线性空间中的向量和列向量之间是一一对应的。我们把列向量

$$\begin{pmatrix} c_1\\ \cdots\\ c_m\end{pmatrix}$$

称为**向量 v 在基底 $\mathcal{B}_V$ 之下的坐标表示**或者 **列向量表示**，记作

$$M_{\mathcal{B}}(v):=\begin{pmatrix} c_1\\ \cdots\\ c_m\end{pmatrix}$$

进一步的容易验证，向量加法和标量乘法，与列向量加法和标量乘法是一致的。在应用数学当中，若给出 m 个向量 $v_1',\cdots,v_m'\in V$, 它们在基底 $\mathcal{B}_V$ 之下的坐标表示为

$$M_{\mathcal{B}}(v_i')=\begin{pmatrix} c_{1i}\\ \cdots\\ c_{ni}\end{pmatrix},$$

则也习惯 m 个列向量坐标表示合成为一个 $n\times m$ 矩阵

$$\begin{pmatrix} c_{11} & \cdots & c_{1m}\\ \cdots & \cdots & \cdots\\ c_{n1} & \cdots & c_{nm}\end{pmatrix}.$$

我们也将其记作

$$M_{\mathcal{B}}(v_1', \cdots, v_m') = \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nm} \end{pmatrix}$$

相应的，我们也需要研究坐标列向量在基底变换下的变化，这需要进一步讨论线性映射。

2.4 线性映射与线性变换

2.4.1 线性映射的概念与运算

在建立了线性空间的概念之后，我们自然的需要研究线性空间之间的“线性”关系，即线性映射。

定义 24. 称一个从线性空间 V 到线性空间 W 的映射 $A: V \to W$ 为一个 k-**线性映射**，如果下面的**线性性**

$$A(u+v) = Au + Av, A(cu) = cA(u)$$

对所有 $c \in k$, $u, v \in V$ 成立，并将 V 到 W 的所有 k-线性映射集合记作

$$Hom_k(V, W),$$

或者在基域 k 选取无歧义情况下将其简记为

$$Hom(V, W).$$

集合 $Hom(V, W)$ 在文献中有时也称其为**线性同态 (linear homomorphism)**。如果 $V = W$, 则也称相应的线性映射为 **线性自同态 (linear endomorphism)**，**线性变换 (linear transformations)**，或**线性算子 (linear operators)**，并将 V 上所有记作 k-线性自同态构成的集合记作

$$End_k(V) := Hom_k(V, V),$$

或在无歧义的情况下简记为

$$End(V).$$

可逆的线性同态，也称为**线性同构**。

例 21. 在任何线性空间 V 上，对任意标量 c，相应的数乘

$$v \mapsto cv$$

是一个线性变换。特别的当 $c=1$ 时，数乘给出的是恒等（线性）映射，记作

$$Id_V \text{ 或者} Id$$

类似的我们将用 c 数乘诱导的变换记作

$$cId.$$

例 22. 考虑 m 维列向量空间 $V=k^m$ 和 n 维列向量空间 $V=k^n$。对于给定的 $n\times m$ 矩阵 A:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

根据矩阵乘法的规则，我们可以直接验证由 A 左乘一个 m 维列向量 $v\in k^m$ 得到一个 n 维列向量的运算 $Av\in k^n$，是从列向量空间 k^m 到列向量空间 k^n 的一个 k-线性变换：

$$A: v\mapsto Av.$$

类似的，由 A 左乘一个 n 维行向量 $u\in(k^n)^t$ 得到一个 m 维行向量的运算 $uA\in(k^m)^t$，给出了从行向量空间 $(k^n)^t$ 到行向量空间 $(k^m)^t$ 的一个 k-线性变换：

$$A: u\mapsto uA.$$

由此我们可以说，矩阵通过左乘列向量或右乘行向量的运算，给出了线性映射的一个**具体实现**。

例 23. 给定一个 $n\times m$ 矩阵 A，则用 A 右乘一个 $m\times l$ 矩阵 B:

$$B\mapsto AB$$

诱导了一个从 $m\times l$ 矩阵集合 $M_{m\times l}(k)$ 到 $n\times l$ 矩阵集合 $M_{n\times l}(k)$ 的线性映射。特别的，我们在第二节提到的初等（行）变换，都是（系数矩阵集合上的）线性变换。

例 24.

对两个线性映射 $A,B\in Hom(V,W)$，定义标量乘法

$$c\cdot A:=cId_W\circ A=A\circ cId_V$$

以及加法

$$(A+B)v:=Av+Bv.$$

命题 34. 线性映射 $Hom(V,W)$ 也以上定义的标量乘法和加法运算下也构成一个 k-线性空间。

尽管线性映射也是线性空间，但相比于一般的线性空间，线性变换还具有额外的运算，即线性变换的**合成**。

定义 25. 给定三个线性空间 U,V,W, 以及两个线性映射

$$A \in Hom(V,W),\ B \in Hom(U,V)$$

则称 A,B 是**可合成**的，并且定义 A,B 的**合成**或 **乘法**为 A,B 作为映射的合成

$$AB := A \circ B.$$

容易验证

$$AB \in Hom(U,W).$$

需要注意并不是任意两个线性映射都能做合成。但是若所有线性映射都是同一个空间 V 上的线性变换，则它们总是可以做合成。

命题 35. 可合成的线性变换，其合成满足结合律

$$(AB)C = A(BC),$$

以及分配律

$$A(B+C) = AB + AC.$$

此外还存在**恒等变换** $Id_V \in End(V)$, $Id_W \in End(W)$,以及**零映射** $O \in Hom(V,W)$ 分别作为合成以及加法的单位元：

$$A \cdot Id_V = Id_W \cdot A = A,\ A + O = A = O + A.$$

2.4.2 核空间，像空间，维数基本关系及其推论

对给定的线性变换 $A \in Hom(V,W)$，他可以自然诱导出 V,W 的两个重要子空间。

定理-定义 3. 对任何一个线性变换 $A \in Hom(V,W)$, 它的**核** ***(kernel)***

$$\mathcal{N}(A) := \{Av = 0 | v \in V\}$$

和 **像** ***(image)***

$$\mathcal{R}(A) := \{Av | v \in V\}$$

分别是 V 和 W 的子空间。核空间与像空间，有时也分别称作**零化空间** ***(null space)*** 和**值域空间** ***(range space)***。

从直观上来看，核空间描述了线性空间在线性变换下被“压缩为零”而导致“损失”的部分，像空间描述了线性空间在线性变换下被“保留”的部分。

容易验证

命题 36. $A \in Hom(V,W)$ 是（集合意义下）的单射，当且仅当

$$\mathcal{N}(A) = \{0\}.$$

相应的 A 是满射当且仅当

$$\mathcal{R}(A) = W.$$

对于线性映射我们有如下的基本关系

定理 9.

$$dim(\mathcal{N}(A)) + dim(\mathcal{R}(A)) = dim(V).$$

证明. 假设 $dim(V) = n$，$v_1, \cdots, v_m$ 为 $\mathcal{N}(A)$ 一组基底，由定义

$$m = dim(\mathcal{N}(A)).$$

由定理 5，存在 $n-m$ 个元素 $v_{m+1}, \cdots, v_n$ 使得添加他们之后得到的 n 个向量

$$v_1, \cdots, v_n$$

构成 V 一组基底。现只需证明 $n-m$ 个向量

$$Av_{m+1}, \cdots, Av_n$$

线性无关，即它们构成了 $\mathcal{R}(A)$ 一组基底。如若不然，存在不全为 0 的常数使得

$$c_{m+1}Av_{m+1} + \cdots + c_nAv_n = 0,$$

即

$$A(c_{m+1}v_{m+1} + \cdots + c_nv_n) = 0,$$

因而

$$c_{m+1}v_{m+1} + \cdots + c_nv_n \in \mathcal{N}(A).$$

然而这表明 $\mathcal{N}(A)$ 中的某个向量可由 $v_{m+1}, \cdots, v_n$ 线性表出，与 $v_1, \cdots, v_n$ 线性无关矛盾。

从“基本关系”我们看到，尽管对于不同的线性变换，相应核空间与像空间的维数一般也会不同，但它们之和是“守恒”的，并且“守恒”总量仅仅与原像空间 V 有关而与 W 无关。

基本关系的一个重要应用是用来分析线性变换可逆性，与线性变换（作为集合之间映射）是单射、满射之间的关系。

命题 37.

$$dim(\mathcal{R}(A)) = dim(V) - dim(\mathcal{N}(A)) \leq dim(V)$$

这一命题表明，值域空间的维数在线性映射“箭头”所指的方向具有“单调性”，它是保持不增的。

定义 26. 称

$$dim(\mathcal{R}(A))$$

为线性映射 A 的 **秩**，记作

$$r(A).$$

显然

$$0 \leq r(A) \leq dim(W).$$

当

$$r(A) = dim(W)$$

时，称 A 是**满秩 (full rank)** 的；当

$$r(A) = dim(V)$$

时，称 A 是**保秩 (rank preserving)** 的。由定义立即推出

命题 38. A 是**满秩 *(full rank)*** 的，当且仅当

$$\mathcal{R}(A) = V;$$

A 是**保秩 *(rank preserving)*** 的，当且仅当

$$\mathcal{N}(A) = \{0\}.$$

引理 3. $A \in Hom(V, W)$ 是单射，当且仅当 A 是保秩的；$A \in Hom(V, W)$ 是满射，当且仅当 A 是满秩的。

关于“满射”部分由定义即可得到，下面我们证明“单射”部分。若 A 是单射，则由单射定义，对任意 $v \neq 0$ 都须有

$$Av \neq 0 = A0,$$

因而

$$\mathcal{N}(A) = \{0\}.$$

若 A 不是单射，则存在 $v, v', v \neq v'$ 使得

$$Av = Av'.$$

由线性性，

$$A(v - v') = 0$$

得出

$$0 \neq v - v' \in \mathcal{N}(A)$$

于是

$$\mathcal{N}(A) \neq \{0\}.$$

下面的命题进一步解释了“保秩”的含义，即保秩映射与线性映射的复合过程中是“保（线性映射的）秩”的。

命题 39. 设 $A \in Hom(V, W)$，$B \in Hom(V', V)$，$C \in Hom(W, W')$ 是三个线性映射，则

$$r(AB) \leq r(A), r(CA) \leq r(A).$$

其中第一式等号成立，当且仅当 B 是单射；第二式成立，当且仅当 C 是单射。进一步的若 B, C 都是单射，则

$$r(A) = r(CA) = r(AB) = r(ABC).$$

定义 27. 线性映射 $A \in Hom(V, W)$ 称为**左可逆**,如果存在映射 $B \in Hom(W, V)$ 使得

$$BA = Id_V.$$

B 也称为 A 的一个**左逆**。相应的线性映射 $A \in Hom(V, W)$ 称为**右可逆**，如果存在映射 $B' \in Hom(W, V)$ 使得

$$AB' = Id_W.$$

B' 也称为 A 的一个**右逆** 如果 A 既存在左逆又存在右逆，则称 A **可逆**，或说 A 给出了 V, W 之间的一个**线性同构**。

我们注意到一般的来讲，存在仅有单侧逆但却不可逆的例子。

命题 40. A 存在左（右）逆当且仅当 A 是单（满）射。

需要注意的是，根据定义 A 是单（满）射推出 A 存在集合意义下的左（右）逆映射，但是以上命题进一步要求我们证明存在相应是线性映射的逆映射，因而以上命题并非一个平凡叙述。

命题 41. *线性空间 V 和线性空间 W 线性同构，当且仅当*

$$dim(V) = dim(W).$$

由此可见，线性空间的维数完全标记了线性空间的线性同构等价类。

我们证明若 A 是单的，则 A 存在左逆。设 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$ 是 V 的一组基底，显然

$$\{Av_1, \cdots, Av_m\}$$

张成 $\mathcal{R}(A)$, 又由于 A 是单射，由命题 33 ，A 是保秩的，因而 $\{Av_1, \cdots, Av_m\}$ 实际上构成 $\mathcal{R}(A)$ 的一组基。

我们有如下的关于线性映射的基本定理，它表明，对任意线性映射 $A \in Hom(V, W)$，数量 $dim(V), dim(W), r(A)$ 决定了线性映射的类型。

定理 10. *对任意的线性映射 $A \in Hom(V, W)$, 若*

$$r(A) = r,$$

则存在 V, W 各自的一组基 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$, $\mathcal{B}_W = \{w_1, \cdots, w_n\}$, $m, n \geq r$, 使得

$$Av_i = w_i,\ i = 1, \cdots, r;\ Av_j = 0,\ j = r + 1, \cdots, n.$$

最后我们进一步讨论 $A \in End(V)$ 是线性变换的特殊情况，我们会发现还有更进一步的关系。

命题 42. *若线性变换 $A \in End(V)$ 存在左 (右) 逆，则它也存在右（左）逆，因而在任何一种情况下 A 可逆。*

由维数关系，

$$dim(\mathcal{N}(A)) + dim(\mathcal{R}(A)) = dim(V),$$

显然

$$dim(\mathcal{N}(A)) = 0,$$

当且仅当

$$dim(\mathcal{R}(A)) = dim(V),$$

因而 A 是满的，当且仅当 A 是单的。

因此与一般线性映射不同的是，线性变换单侧逆的存在性自动蕴含了可逆，因而对一个线性变换 A 而言，

$$A \text{ 是单射} = A \text{ 是满射} = A \text{ 可逆} = A \text{ 满秩} = A \text{ 保秩}$$

容易看到这对一般线性映射是不成立的。

我们还容易得出以下结论

命题 43. 设 $A_1, \cdots, A_k \in End(V)$ 是一些 A 上的线性变换，则

$$r(A_1 \cdots A_k) \leq dim(V).$$

其中的等号成立，当且仅当每个 A_i 都是可逆的，且此时

$$A_1 \cdots A_k$$

也可逆。

本节的主要结论，将在 2.6 节当中用来给出一些关于矩阵的定理的便捷证明。

2.5 线性映射和线性空间的矩阵表示

2.5.1 用矩阵表示线性映射

在前面我们已经指出，列向量空间和矩阵分别给出了线性空间和线性映射的具体模型实现，并且抽象的向量也可以通过基底的坐标表示来实现为列向量。现在我们利用线性空间总存在**基底**这一基本事实，来说明任何抽象的线性空间和线性映射，总可以通过基底的选取来分别等同于例 16 和例 17 中的列向量空间和矩阵。由于线性空间总存在基底，所有基底具有相同的元素个数，且线性映射满足“线性性”，因而我们能够通过有限多个标量构成的列向量以及矩阵，来描述向量以及线性映射。这使得线性映射具有一般映射（或函数）所不具备的便利。

本节我们固定用

$$\mathcal{B}_V := \{v_1, \cdots, v_m\}$$

与

$$\mathcal{B}_W := \{w_1, \cdots, w_n\}$$

来表示 V, W 分别选取的一组基。

我们注意到在线性映射 A 的作用下，基向量 $v_1, \cdots, v_m$ 被线性映射 A 映射成为了 W 中的 m 个向量 $Av_1, \cdots, Av_m$。现在我们来证明相反的方向来表明线性映射集合的元素“足够多”，使得基向量像的选择具有“任意性”。

命题 44. 给定 V 一组基底

$$\mathcal{B}_V := \{v_1, \cdots, v_m\},$$

则对线性空间 W 中任何给定的 m 个向量 $\tilde{w}_1, \cdots, \tilde{w}_m \in W$，存在唯一的线性映射 A 使得

$$Av_i = \tilde{w}_i, i = 1, \cdots, m.$$

等价的，取定基底 $\mathcal{B}_V$ 后，线性映射空间

$$Hom(V, W)$$

与 n 次笛卡尔积空间

$$W^n := \underbrace{W \times \cdots \times W}_{n次}$$

是一一对应的，

首先我们证明存在性。设 $v \in V$ 是任意一个向量，由基底的性质，存在唯一一组系数 $c_1, \cdots, c_m$ 使得

$$v = c_1 v_1 + \cdots + c_m v_m.$$

定义映射 $A : V \to W$ 使得对 v 有

$$A : v \mapsto c_1 \tilde{w}_1 + \cdots + c_m \tilde{w}_m.$$

显然由以上定义，

$$Av_i = \tilde{w}_i,$$

并且我们可以直接验证，对任意 $v, v' \in V$，$c \in k$，A 满足线性性:

$$A(v + v') = Av + Av', A(cv) = cAv.$$

下面再证明唯一性：若 A, A' 是两个满足题设条件的线性映射：

$$Av_i = \tilde{w}_i = A'v_i, i = 1, \cdots, m,$$

因而对任意 $v = c_1 v_1 + \cdots + c_m v_m$，由线性性得到

$$(A - A')v = c_1(Av_1 - A'v_1) + \cdots + c_m(Av_m - A'v_m) = 0,$$

即

$$A = A'.$$

以上命题告诉我们，在选定 V 的一组基底后，线性映射空间 $Hom(V,W)$ 与 W 中的 $n = dim(V)$ 元（有序）向量组一一对应。这个对应有助于我们利用线性映射来分析向量组的性质。下面我们给出具体的行或列向量空间的例子。

例 25. 回忆线性方程组 *(2)* 的形式 *(4)* 当中，对任意给定的 $n \times m$ 矩阵 A, 它左乘列向量空间 k^m 也诱导了一个线性映射

$$A : x \mapsto Ax, \, x \in k^m.$$

具体的若 A 以列向量形式分块

$$A = (v_1, \cdots, v_m),$$

则

$$x \mapsto Ax = (v_1, \cdots, v_m) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1 v_1 + \cdots + x_m v_m.$$

特别的，若取 k^m 的标准基底 $\{e_1, \cdots, e_m\}$ 则

$$e_i \mapsto Ae_i = v_i,$$

因而以上给出了列向量组 $v_1, \cdots, v_m$ 与 $Hom(k^m, k^n)$ 的一一对应关系。我们还看到，A 作为线性变换

$$\mathcal{R}(A) = span\{v_1, \cdots, v_m\},$$

因而

$$r(A) = dim(\mathcal{R}(A)) = dim(span\{v_1, \cdots, v_m\}) = r(v_1, \cdots, v_m) = r_{col}(A).$$

平行的，A 诱导的行向量空间之间的线性变换，也具有类似的性质。

命题 45. 矩阵 A 做为列向量空间上线性变换的秩，与 A 的列秩相同；矩阵 A 作为行向量线性空间上线性变换的秩，与 A 的行秩相同。

命题 46. 方阵 A 可逆，当且仅当 A 列满秩。

由此可以看到，决定一个线性映射 $A \in Hom(V, W)$ 不需要烦劳对每个元素给出像，而仅仅需要给出一组基底 $\mathcal{B}_V$ 中元素的像即可。以上的证明中构造 A 的过程，我们也称为由 A **在基底 $\mathcal{B}_V$ 上的定义，通过线性性（唯一的）拓展为整个空间 V 到 W 的线性映射**。在线性代数中，会频繁使用这一说法。

进一步的，若给定 W 的一组基 $\mathcal{B}_W = \{w_1, \cdots, w_n\}$ 若考虑由向量组 $\tilde{w}_1, \cdots, \tilde{w}_m \in W$ 决定的线性映射

$$Av_i = \tilde{w}_i, \, i = 1, \cdots, m.$$

则再次由基底的性质，对每个 $\tilde{w}_i \in W$，存在相应的唯一一组的数 $a_{ji}, j = 1, \cdots, n$ 使得

$$\tilde{w}_i = \sum_j a_{ji} w_j,$$

即 $\tilde{w}_i$ 具有列向量坐标表示

$$M_{\mathcal{B}_W}(\tilde{w}_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

由此向量组

$$(\tilde{w}_1, \cdots, \tilde{w}_m)$$

通过 $\mathcal{B}_W$ 的坐标表示，可以进一步得到一个矩阵。我们称相应的 $n \times m$ 坐标表示矩阵矩阵

$$M_{\mathcal{B}_W}(\tilde{w}_1, \cdots, \tilde{w}_n) = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

为线性映射 A **在 V、W 各自的一组基 $\mathcal{B}_V$、$\mathcal{B}_W$ 下的矩阵表示**，并将其记作

$$M_{\mathcal{B}_V}^{\mathcal{B}_W}(A).$$

特别，若 $W = V$ 是同一个线性空间，则 $A \in End(V)$ 是一个 V 上的线性变换，V 同时作为原像空间和像空间，可选取 V 上的同一组基底 $\mathcal{B}_V$。在无歧义的情况下为方便起见，我们也说 $M_{\mathcal{B}_V}^{\mathcal{B}_V}(A)$ 是线性变换 $A \in End(V)$ **在基底 $\mathcal{B}_V$ 下的矩阵表示**，并简记作

$$M_{\mathcal{B}_V}(A) := M_{\mathcal{B}_V}^{\mathcal{B}_V}(A).$$

反之，若给定了一个 $n \times m$ 矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

则根据命题 39，存在唯一的线性变换 A 使得

$$Av_i = \sum_j a_{ji} w_j.$$

因而我们得到了相反方向的对应。

需要注意的是在一般情况下根据需要，V 同时作为像空间和原像空间，可以独立的取两组基底 $\mathcal{B}_V$、$\mathcal{B}'_V$ 来进行表示，这与上面的用语的含义是有所不同。

综上我们得到

命题 47. 若选取了 V, W 各自的一组基底 $\mathcal{B}_V = \{v_1, \cdots, v_n\}$,$\mathcal{B}_W = \{w_1, \cdots, w_m\}$，则 $n \times m$ 矩阵空间

$$M_{n\times m}(k)$$

与线性映射空间

$$Hom(V, W)$$

是一一对应的。即一方面，若给定了矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

则存在唯一的线性映射 $A \in Hom(V, W)$ 使得

$$Av_i = \sum_j a_{ji} w_j$$

成立；另一方面，若给定了线性映射 A，则上面的规则

$$A \mapsto M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)$$

唯一决定了对应的 $n \times m$ 矩阵。

注 1. 在此处，线性空间“基”的概念比先前的定义略有加强而是指**有序基**。在先前“基”的概念中，并不强调基元素的顺序，但是在这里为了写下矩阵表达式，我们必须强调基的顺序，例如

$$\{v_1, v_2, v_3\}, \{v_2, v_1, v_3\}, \{v_3, v_2, v_1\}$$

是三组不同的有序基。今后凡是涉及矩阵表示的问题，相应的基选取均指有序基。

注 2. 我们注意到在表达式

$$\begin{cases} Av_1 = a_{11}w_1 + \cdots + a_{n1}w_n, \\ \cdots \\ Av_i = a_{1i}w_1 + \cdots + a_{ni}w_n, \\ \cdots \\ Av_m = a_{1m}w_1 + \cdots + a_{nm}w_n, \end{cases}$$

当中，规定求和是对系数 a_{ji} 的第一个下指标分量 j 做的。另外，由系数 a_{ji} 按照"正常排列"得到的 $m \times n$ 矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{nm} \end{pmatrix}$$

是表示矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

的转置，并且无论我们如何安排求和指标都是如此。我们将从下面的矩阵运算看到这一安排的必要性，而这一现象的根源在对偶空间部分会得到自然解释。

2.5.2 线性映射、线性空间的矩阵表示，及其运算相容性

首先我们注意到，线性变换作用在向量上的矩阵表示，与相应的矩阵表示乘法一致。

命题 48. *线性变换在向量上的作用对应矩阵与列向量的乘法，即*

$$M_{\mathcal{B}_U}^{\mathcal{B}_V}(B) M_{\mathcal{B}_U}(u) = M_{\mathcal{B}_V}(Bu),$$

进一步的我们来看矩阵的其他运算。若在 U, V 的基底 $\mathcal{B}_U, \mathcal{B}_V$ 下的情况下，$B, C \in Hom(U, V)$，B, C 分别有矩阵表示

$$M_{\mathcal{B}_U}^{\mathcal{B}_V}(B) = \begin{pmatrix} b_{11} & \cdots & b_{1l} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{ml} \end{pmatrix}, M_{\mathcal{B}_U}^{\mathcal{B}_V}(C) = \begin{pmatrix} c_{11} & \cdots & c_{1l} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{ml} \end{pmatrix}.$$

则经过简单计算，$B + C$ 相应的矩阵表示为

$$M_{\mathcal{B}_U}^{\mathcal{B}_V}(B + C) = \begin{pmatrix} b_{11} + c_{11} & \cdots & b_{1l} + c_{1l} \\ \cdots & \cdots & \cdots \\ b_{m1} + c_{m1} & \cdots & b_{ml} + c_{ml} \end{pmatrix},$$

而 A, B 的复合的表示矩阵是一个 $n\times l$ 矩阵

$$M_{\mathcal{B}_U}^{\mathcal{B}_W}(AB)=\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \cdots & \cdots & \cdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}\begin{pmatrix}b_{11} & \cdots & b_{1l}\\ \cdots & \cdots & \cdots\\ b_{m1} & \cdots & b_{ml}\end{pmatrix},$$

于是我们得到

命题 49. 我们有

$$M_{\mathcal{B}_U}^{\mathcal{B}_V}(B+C)=M_{\mathcal{B}_U}^{\mathcal{B}_V}(B)+M_{\mathcal{B}_U}^{\mathcal{B}_V}(C).$$
$$M_{\mathcal{B}_U}^{\mathcal{B}_W}(AB)=M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)M_{\mathcal{B}_U}^{\mathcal{B}_V}(B),$$

即线性变换复合恰好对应相应矩阵表示的乘积。我们注意到这一对应恰恰是由于前面的规定而成立。

至此我们已经系统的讨论了抽象公理化的线性空间、线性映射与列向量、矩阵的关系。通过这些对象我们可以为近代数学中的**内蕴 (intrinsic)** 和**外蕴 (extrinsic)** 概念提供典型范例。我们注意到，列向量、矩阵给出了抽象线性空间和抽象线性变换的**具体实现**，但是这些具体实现，依赖于抽象线性空间当中基底的**选取**。对于一个固定的向量或线性变换，其列向量或矩阵实现也可能会随着基底的选取不同而不同，因而我们将它们归类为**外蕴**范畴，将线性空间和线性变换归类为**内蕴**范畴。后面我们会遇到一些与线性空间和线性变换相关的**内蕴量**和**内蕴性质**，它们虽然可以通过列向量或矩阵来表达，但与基底选取无关，因而在某些文献中也常常称内蕴性为**坐标无关性 (coordinate free)**。对内蕴对象的研究是数学和理论物理的重要内容之一。

在很多情况下，利用线性空间和线性映射的抽象概念，可以避免复杂的矩阵运算而专注于数学结构的分析。在数学中，还习惯利用箭头图来直观表示线性映射，例如

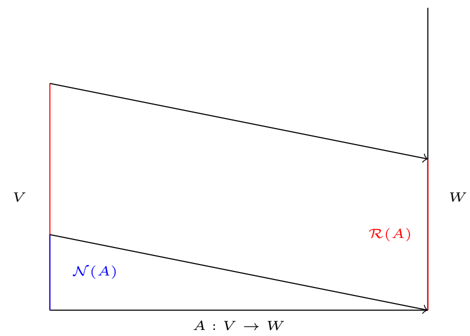
$$U\xrightarrow{B}V$$

表示 B 是一个 U 到 V 的线性映射。类比水的流动我们也把其中 U 称作线性映射 B 的**源**，把 V 称作**汇**。而进一步的

$$U\xrightarrow{B}V\xrightarrow{A}W,$$

表示了线性映射 A 和 B 的复合，而且可以看到顺向箭头表示的线性映射可以自然的做复合运算。对一个线性自同态 $T\in End(V)$，则用一个到自身的箭头来表示：

$$T\circlearrowright V.$$



2.5.3 矩阵表示与基底变换

由于线性变换的表示矩阵表示会因基选取的改变而改变，因此若使用矩阵研究线性变换，我们还需要研究线性变换的矩阵表示是如何随着基的变换而变换的。

给定 V 的两组基 $\mathcal{B}_V = \{v_1, \cdots, v_n\}$，$\mathcal{B}'_V = \{v'_1, \cdots, v'_n\}$, 我们知道存在唯一的线性变换 $T_{\mathcal{B}_V}^{\mathcal{B}'_V}$ 使得

$$v'_i = T_{\mathcal{B}_V}^{\mathcal{B}'_V} v_i, \ i = 1, \cdots, n.$$

我们称 $T_{\mathcal{B}_V}^{\mathcal{B}'_V}$ 为从基底 $\mathcal{B}_V$ 到基底 $\mathcal{B}'_V$ 的 **过渡线性变换**。由基底的性质，存在一组数字 t_{ij} 使得

$$v'_i = \sum_j t_{ji} v_j,$$

因而矩阵

$$(t_{ij})_{n\times n}$$

是 $T_{\mathcal{B}_V}^{\mathcal{B}'_V}$ 在基底 $\mathcal{B}_V$ 下的表示矩阵。我们也称

$$M_{\mathcal{B}_V}^{\mathcal{B}_V}(T_{\mathcal{B}_V}^{\mathcal{B}'_V}) = (t_{ij})_{n\times m}$$

为从基底 $\mathcal{B}_V$ 到基底 $\mathcal{B}'_V$ 的 **过渡矩阵**，简记为

$$M_{\mathcal{B}_V}^{\mathcal{B}'_V}.$$

我们注意到过渡矩阵、过渡线性变换有以下一些性质

命题 50. *1.* $M_{\mathcal{B}_V}^{\mathcal{B}'_V}(T_{\mathcal{B}_V}^{\mathcal{B}'_V}) = I_m$, $M_{\mathcal{B}'_V}^{\mathcal{B}_V}(Id_V) = M_{\mathcal{B}_V}^{\mathcal{B}'_V}$.

2. $M_{\mathcal{B}_V}^{\mathcal{B}'_V} M_{\mathcal{B}'_V}^{\mathcal{B}_V} = M_{\mathcal{B}'_V}^{\mathcal{B}_V} M_{\mathcal{B}_V}^{\mathcal{B}'_V} = I_m$, $T_{\mathcal{B}_V}^{\mathcal{B}'_V} T_{\mathcal{B}'_V}^{\mathcal{B}_V} = T_{\mathcal{B}'_V}^{\mathcal{B}_V} T_{\mathcal{B}_V}^{\mathcal{B}'_V} = Id_V$.

3. 数乘变换的矩阵表示与基底选取无关，即对任意基底 $\mathcal{B}_V=\{v_1,\cdots,v_m\}$，$\mathcal{B}_W=\{w_1,\cdots,w_m\}$ 有

$$M_{\mathcal{B}_V}(cId_V):=M_{\mathcal{B}_V}^{\mathcal{B}_V}(cId_V)=cI_m.$$

在此需要注意 (3) 当中“与基底选取无关”，要求 V 作为像空间和原像空间，需要取相同的基底 $\mathcal{B}_V$，否则根据 (1)，若 V 的原像空间的基底 $\mathcal{B}_V$ 与其作为像空间的基底 $\mathcal{B}'_V$ 不同，则恒等变换的表示矩阵也可以不是单位矩阵。

遵循前面的记号，假设 $\mathcal{B}'_V=\{v'_1,\cdots,v'_m\}$ 为 V 的另一组基底，我们来研究 $M_{\mathcal{B}'_V}^{\mathcal{B}_W}(A)$ 与 $M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)$ 的关系。利用

$$A=A\cdot Id_V$$

以及命题 44，我们得到

$$M_{\mathcal{B}'_V}^{\mathcal{B}_W}(A)=M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)M_{\mathcal{B}'_V}^{\mathcal{B}_V}(Id_V)=M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)M_{\mathcal{B}_V}^{\mathcal{B}'_V}$$

类似的若 $\mathcal{B}'_w=\{w'_1,\cdots,w'_n\}$ 为 W 的另一组基底，则

$$M_{\mathcal{B}_V}^{\mathcal{B}'_W}(A)=M_{\mathcal{B}_W}^{\mathcal{B}'_W}(Id_W)M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)=M_{\mathcal{B}'_W}^{\mathcal{B}_W}M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)$$

以上对 V 和 W 的基变换可同时进行，即

$$M_{\mathcal{B}'_V}^{\mathcal{B}'_W}(A)=M_{\mathcal{B}'_W}^{\mathcal{B}_W}M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)M_{\mathcal{B}_V}^{\mathcal{B}'_V}=(M_{\mathcal{B}_W}^{\mathcal{B}'_W})^{-1}M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)M_{\mathcal{B}_V}^{\mathcal{B}'_V}.$$

由此可见对源 V 做基变换对应的矩阵变换是对变换矩阵右侧乘以过度矩阵，而对汇 W 做基变换对应的矩阵变换是对变换矩阵左侧乘以过度矩阵的逆，因而源和汇在变换下的情况有所不同。

特别的注意到若 $W=V$ 是同一个空间时，V 既是源也是汇，所以将基底 $\mathcal{B}_V$ 换为 $\mathcal{B}'_V$ 时源和汇需要**同时**进行基变换，因此

$$M_{\mathcal{B}'_V}(A)=M_{\mathcal{B}'_V}^{\mathcal{B}_V}M_{\mathcal{B}_V}(A)M_{\mathcal{B}_V}^{\mathcal{B}'_V}=(M_{\mathcal{B}_V}^{\mathcal{B}'_V})^{-1}M_{\mathcal{B}_V}(A)M_{\mathcal{B}_V}^{\mathcal{B}'_V}$$

寻找适当的一组基底，使得线性自同台具有“较好”的形式，相较于上面的问题更加复杂，它将是第三章的主要内容。

2.6 矩阵的秩与 Smith 标准型

2.6.1 初等列变换与高斯消元

本节我们重新回到利用矩阵运算，讨论 Gauss 消元法求解线性方程组的问题，同时利用 2.3，2.4，2.5 节关于抽象线性空间、线性映射的结论，给出一些矩阵运算的解释，并提供一些较为简洁的证明。

我们先回到利用矩阵初等变换求解方程。回忆在 2.2 节当中，在有解的情况下，我们将方程 (2) 转化为了阶梯状方程

$$\begin{pmatrix} A' \\ O \end{pmatrix} x = \begin{pmatrix} b' \\ 0 \end{pmatrix} \tag{11}$$

的求解。我们会看到，在余下的步骤中，需要对阶梯矩阵 A' 进行列变换。

与行变换部分不同的是，对阶梯矩阵进一步做初等列变换所得到的新方程，一般来说与原方程的解集是不同的。但是，利用初等变换的可逆性，新的方程的解集与原有方程通过可逆矩阵，仍然可以建立起一一对应关系。

命题 51. *若 A' 是一个 $r \times m$ 的阶梯矩阵，且没有全为零的一行，则存在 m 阶可逆矩阵 Q 使得*

$$\begin{pmatrix} A' \\ O \end{pmatrix} Q = D.$$

其中

$$D = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

是对角元素全为 1 其余为 0 的 $r \times m$ 矩阵。

证明这一命题的方法，仍然是利用主元进行高斯消元。首先，假设

$$l_1 < l_2 \cdots < l_r$$

分别是 $1, \cdots, r$ 行主元所在列，则我们发现依次交换第 i, l_i, $i = 1, \cdots, r$ 两列，或等价的，依次将矩阵右乘 P_{i,l_i}, $i = 1, \cdots, r$，则原阶梯矩阵化为对角现全部为 1 的上三角形状矩阵，即

$$A' P_{1,l_1} \cdots P_{r,l_r} = \begin{pmatrix} \boxed{1} & * & * & * & * & * \\ & \boxed{1} & * & * & * & * \\ & & \ddots & & & \\ 0 & & & \boxed{1} & * & * \end{pmatrix}$$

这样，第 i 行的主元位置即第 i 列。

然后，我们可以从第一行开始，用位于主元上的 1，从每一行从左向右，依

次消去每一行主元左侧的非零元素。例如，假设矩阵

$$\begin{pmatrix} \boxed{1} & 0 & & & & & & & 0 \\ & \boxed{1} & 0 & & & & & & 0 \\ & & \ddots & & & & & & \\ & & & \boxed{1} & 0 & 0 & a_{ij} & * & * \\ & & & & \ddots & & & & \\ & & & & & \boxed{1} & * & * & * \end{pmatrix}$$

当中，前 $i-1$ 行，以及第 i 行第 j 列之前的非主元位置元素都已经为零，那么我们通过右乘 $T_{ij}(-a_{ij})$ ，即将 (i,j) 位置元素化为零，且不影响先前结果:

$$\begin{pmatrix} \boxed{1} & 0 & & & & & & & 0 \\ & \boxed{1} & 0 & & & & & & 0 \\ & & \ddots & & & & & & \\ & & & \boxed{1} & 0 & 0 & a_{ij} & * & * \\ & & & & \ddots & & & & \\ & & & & & \boxed{1} & * & * & * \end{pmatrix} \xrightarrow{\text{右乘}T_{ij}(-a_{ij})}$$

$$\sim \begin{pmatrix} \boxed{1} & 0 & & & & & & & 0 \\ & \boxed{1} & 0 & & & & & & 0 \\ & & \ddots & & & & & & \\ & & & \boxed{1} & 0 & 0 & 0 & * & * \\ & & & & \ddots & & & & \\ & & & & & \boxed{1} & * & * & * \end{pmatrix}$$

自此我们证明了命题。

事实上我们纯粹借助矩阵运算，证明了命题 46。

我们回到阶梯状方程的求解。我们注意到，系数矩阵转化为对角矩阵的方程

$$Dx' = \begin{pmatrix} b' \\ 0 \end{pmatrix},$$

即

$$\begin{cases} x_1' & & & & & & & = b_1' \\ & \ddots & & & & & & = \vdots \\ & & x_i' & & & & & = b_i' \\ & & & \ddots & & & & = \vdots \\ & & & & x_r' = & & & = b_r' \\ & & & & & 0x_{r+1}' & & = 0 \\ & & & & & & \ddots & \vdots \\ & & & & & & 0x_m' & = 0 \end{cases}$$

已经完全实现了 **解耦合**，因而可以直接读出其一般解为

$$x' = (b_1', \cdots, b_r', x_{r+1}', \cdots, x_m')^t.$$

其中 $x_{r+1}', \cdots, x_m'$ 是可以去任意值的“自由变量”。为了建立这两个方程之间的关系，注意到若 x 是 $A'x = b'$ 一个解。考虑到 Q 是 **可逆矩阵**，并利用矩阵分块运算，得到

$$x = I_m x = QQ^{-1}x$$

因而

$$b' = \begin{pmatrix} A' \\ O \end{pmatrix} QQ^{-1}x = DQ^{-1}x$$

反之若我们设 x' 是

$$Dx' = \begin{pmatrix} b' \\ 0 \end{pmatrix}$$

的一个解，则 $x := Q^{-1}x'$ 是 (11) 的一个解。因而我们证明了

命题 52. $x \in k^m$ 是阶梯状方程

$$\begin{pmatrix} A' \\ O \end{pmatrix} x = b'$$

的一个解，当且仅当 $x' := Q^{-1}x \in k^m$ 是

$$Dx' = b'$$

的一个解。

2.6.2 Smith 标准型及其推论

我们得到

定理 11. 对任意非零的 $n \times m$ 矩阵 A，存在正整数 r 与 n 阶可逆方阵 P，和 m 阶可逆方阵 Q，使得

$$PAQ = D$$

其中

$$D = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

等价的，

$$A = P^{-1}DQ^{-1}.$$

它也称为 A 的 ***Smith* 标准型**。

矩阵的行秩、列秩之间也存在如下"巧合"：

定理 12. 对给定的 $n \times m$ 矩阵 A，它的行秩、列秩，和作为行、列向量空间上线性变换的秩都相同，且等于定理 *10* 中的整数 r。

引理 4. A 的列秩 $r_c(A)$，等于将 A 等同于列空间线性变换的秩。

假设 A 按列分块的形式为

$$A = (v_1, \cdots, v_m).$$

取 k^m 的标准基底 $e_1, \cdots, e_m$，通过直接计算我们得到

$$Ae_i = v_i,$$

因而由定义，

$$\mathcal{R}(A) = span\{v_1, \cdots, v_m\}.$$

定义 28. 若一个 n 阶方阵 A 满足

$$r(A) = n,$$

则称矩阵 A 是**满秩**的。

由 Smith 标准型的结论，若一个矩阵的行数和列数不同，则它不可能同时是行满秩与列满秩的，因而满秩概念仅对方阵才有意义。

我们将 A 视作 $Hom(k^m, k^n)$ 中的一个元素。将定理 10 应用于情形 $V = V' = k^m, W = W' = k^n$，则得到

命题 53. *若 $B \in M_m(k), M_n(k)$ 都是可逆方阵，则*

$$r_{col}(CAB) = r_{col}(A).$$

即，任何一个矩阵左乘或者右乘可逆方阵，不改变矩阵的行秩。

类似的，将以上的“行向量”全部换为“列向量”，则同样得到

命题 54. *若 $B \in M_m(k), M_n(k)$ 都是可逆方阵，则*

$$r_{row}(CAB) = r_{row}(A).$$

即，任何一个矩阵左乘或者右乘可逆方阵，不改变矩阵的列秩。

设 A 的 Smith 标准型为

$$A = P^{-1}DQ^{-1},$$

利用上面的命题，并注意到

$$r_{row}(D) = r_{col}(D) = r,$$

显然得到

$$r_{row}(A) = r_{row}(PAQ) = r_{row}(D) = r_{col}(D) = r_{col}(P^{-1}DQ^{-1}) = r_{col}(A) = r.$$

作为线性空间和线性变换理论的应用，我们证明了

命题 55. *给定一个 $n \times m$ 矩阵 A 。*

1. *A 所诱导的列向量线性变换是满秩的，当且仅当它是列满秩的。*
2. *A 所诱导的列向量线性变换是保秩的，当且仅当它是行满秩的。*
3. *进一步的若 A 为方阵，则 A 是可逆矩阵，当且仅当它是满秩的。*

2.7 线性空间与线性变换的构造

我们知道数域中数的基本运算有加减乘除四种。在线性代数中，我们也需要通过对线性空间进行“运算”。与线性空间自身存在的向量加法和数乘运算不同，这些“运算”得到的是一些新的线性空间。对线性空间的基本“运算”有以下四种：

1. 线性空间“求和”得到直和空间。

2. 线性空间“做商”得到商空间。

3. 线性空间“取对偶”得到对偶空间。

4. 线性空间“做乘积”得到张量空间。

后面我们会看到，这其中的“求和”、“做乘积”与“做商”在某些方面，分别类似正整数的加法、乘法和减法，但含义上又有进一步的丰富和延伸，而“取对偶”则难以找到类比。本节我们主要讨论前两种操作，即直和空间与商空间。

2.7.1 线性空间直和

假设 V_1, V_2 是 V 的子空间，则

$$V_1 + V_2 := \{v_1 + v_2 | \, v_1 \in V_1, \, v_2 \in V_2\}$$

也是 V 的子空间，称作 V_1, V_2 的**和**。进一步的若

$$V_1 \cap V_2 = \{0\},$$

则称 $V_1 + V_2$ 为 V_1, V_2 的**（内）直和**，记作

$$V_1 \oplus V_2.$$

类似的我们可以仿照以上形式定义 V 的子空间 $V_1, \cdots, V_n$ 的和为

$$V_1 + \cdots + V_n := \{\sum v_i | \, v_i \in V_i, \, i = 1, \cdots, n\}.$$

容易验证

命题 56. $V_1 + \cdots + V_n$ 也是 V 的子空间。

进一步的，

定义 29. 设 $V_1, \cdots, V_n$ 是 V 的子空间，若 V_i 满足对任意整数 $i = 1, \cdots, n-1$，有

$$(V_1 + \cdots + V_i) \cap V_{i+1} = \{0\},$$

则称 $V_1 + \cdots + V_n$ 为 $V_1, \cdots, V_n$ 的（内）直和，也记作

$$\oplus_i V_i.$$

一般内直和也可以通过归纳的方式进行定义，即对任意 i, $V_1+\cdots+V_i$ 是 $V_1,\cdots,V_i$ 的直和，且

$$V_1+\cdots+V_{i+1}$$

是 $V_1+\cdots+V_i$ 与 V_{i+1} 的直和。若 $V=V_1\oplus\cdots\oplus V_i$，则也称 $V_1,\cdots,V_n$ 给出了 V 的一个**直和分解**。在此我们注意到，内直和是满足附加条件的线性空间的和。

另外需特别注意的是，$V_1,\cdots,V_n$ 满足

$$V_i\cap V_j=\{0\}$$

一般并不能推出 $V_1+\cdots+V_n$ 是 $V_1,\cdots,V_n$ 的直和（为什么?）。

直和分解的重要之处在于，他保证我们可以对较大的空间分治为较小的空间来进行研究。

命题 57. $V_1+\cdots+V_n$ 是 $V_1,\cdots,V_n$ 的直和，当且仅当对任意 $v\in V_1+\cdots+V_n$，满足 $v_1\in V_1,\cdots,v_n\in V_n$ 且

$$v=v_1+\cdots+v_n$$

的分解是唯一的。

特别的，当以上命题的条件成立时，注意到零向量总有平凡分解

$$0=0+\cdots+0,$$

因而

$$0=w_1+\cdots+w_n,\, w_i\in V_i$$

推出

$$w_1=\cdots=w_n=0.$$

反之若零向量仅仅存在非平凡分解，即

$$0=w_1'+\cdots+w_n',\, w_i\in V_i$$

推出

$$w_1'=\cdots=w_n',$$

则对任意向量 $v\in V$ 的两个分解

$$v=v_1+\cdots+v_n=v_1'+\cdots+v_n',,\, v_iv_i'\in V_i$$

得到

$$0=(v_1-v_1')+\cdots+(v_n-v_n')$$

进而由 $v_i-v_i'\in V_i$ 推出

$$v_i=v_i'$$

即 v 总是存在唯一分解。于是我们得到

命题 58. $V_1+\cdots+V_n$ 是 $V_1,\cdots,V_n$ 的直和，当且仅当 $w_1\in V_1,\cdots,w_n\in V_n$ 且

$$0=w_1+\cdots+w_n$$

推出

$$w_i=0.$$

我们先证明 $n=2$ 的情况。假设有分解

$$0=w_1+w_2,w_i\in V_i,$$

则得到

$$w_1=-w_2\in V_1\cap V_2.$$

若 $V=V_1\oplus V_2$ 是直和分解，则有定义 $V_1\cap V_2=\{0\}$，得到

$$w_1=w_2=0.$$

反之若 $V_1\cap V_2\neq\{0\}$，则存在 $0\neq w\in V_1\cap V_2$，因而 $w\in V_1$, $-w\in V_2$，$w\neq 0$ ，因而

$$0=w+(-w)$$

得到 0 的非平凡分解。

现在对一般的 n 用归纳法来证明。在前面已经证明了 $n=2$ 的情况。假设命题对 $n=k$ 的情况已经成立，现证明 $n=k+1$ 的情况也成立。若 $V=V_1\oplus\cdots\oplus V_{k+1}$ 是 $V_1,\cdots,V_{k+1}$ 的直和,则依据定义,$V_1+\cdots+V_{k+1}$ 是 $V_1+\cdots+V_k$ 与 V_{k+1} 的直和，且 $V_1+\cdots+V_k$ 是 $V_1,\cdots,V_k$ 的直和。若

$$0=w_1+\cdots+w_{k+1},\ w_i\in V_i,$$

注意到

$$w_1+\cdots+w_k\in V_1+\cdots+V_k=V_1\oplus\cdots\oplus V_k,$$

因而由 $n=2$ 的情况，

$$w_{k+1}=0,\ w_1+\cdots+w_k=0.$$

再由归纳假设，得到

$$w_i = 0, i = 1, \cdots, k,$$

得到 0 的分解的唯一性。反之，若 $V = V_1 + \cdots + V_n$ 不是 $V_1, \cdots, V_n$ 的直和，依据定义，存在 i 使得

$$(V_1 + \cdots + V_i) \cap V_{i+1} \neq 0,$$

即 $V = V_1 + \cdots + V_{i+1}$ 不是 $V_1 + \cdots + V_i$ 与 V_{i+1} 的直和。取 $0 \neq w \in (V_1 + \cdots + V_i) \cap V_{i+1}$。特别的

$$w = w_1 + \cdots + w_i.$$

又注意到 $0 \neq w \in V_{i+1}$，因而

$$0 = w_1 + \cdots + w_i + (-w) + 0 + \cdots + 0$$

得出 0 的一个非平凡分解。

我们注意到，以上关于线性空间和与直和的定义，以及直和分解与向量表示分解唯一性的等价的证明这些内容当中，只用到了向量的加法而完全没有用到标量乘法，因而以上所有的定义与结论可以推广到更为一般的**阿贝尔群直和**而无需做任何修改。

定义 30. *若 V_1 是 V 的一个子空间，则称满足*

$$V = V_1 \oplus V_2$$

*的子空间 V_2 为 V_1 在 V 当中的*__补子空间__*。*

定理 13. *若 V_1 是 V 的一个子空间，则补空间总存在。*

我们需要注意的是，补空间通常并不唯一。

在上面的线性直和空间构造中，我们利用了 V_1, V_2 是 V 的子空间这一点。在代数学当中，另一种常用的构造是（外）直和，它从任意线性空间 V_1, V_2 出发直接构造一个新的与 V_1, V_2 相关联的线性空间，而不需要 V_1, V_2 是某个空间的子空间。我们会看到这种模式会继续出现在商空间和张量空间的构造中。

定义 31. *对任意的线性空间 V_1, V_2, 考虑他们的笛卡尔积*

$$V_1 \times V_2.$$

在 $V_1 \times V_2$ 上定义加法和标量乘法运算为

$$(v_1, v_2) + (v_1', v_2') = (v_1 + v_1', v_2 + v_2')$$

$$a(v_1, v_2) = (av_1, av_2),$$

使得 $V_1 \times V_2$ 称为一个新的线性空间，称作 V_1, V_2 的 **(外) 直和**，记作

$$V_1 \oplus V_2.$$

2.7.2 直和分解诱导的线性变换

给定了一个线性空间 V 的直和分解

$$V = V_1 \oplus V_2,$$

对任意向量 $v \in V$，我们有相应的分解

$$v = v_1 + v_2,\ v_i \in V_i,$$

并且 v_1, v_2 是由直和分解唯一确定的。同时，我们还可以得到四个线性映射 p_1, p_2, i_1, i_2

$$V_1 \underset{p_1}{\overset{i_1}{\rightleftarrows}} V_1 \oplus V_2 \underset{i_2}{\overset{p_2}{\rightleftarrows}} V_2$$

$$i_1 : v_1 \mapsto v,\ i_2 : v_2 \mapsto v,$$

$$p_1 : v \mapsto v_1,\ p_2 : v \mapsto v_2$$

容易验证四个线性映射满足

$$p_1 i_1 = I_{V_1},\ p_2 i_2 = I_{V_2},\ i_1 p_1 + p_2 i_2 = Id_V.$$

我们称 i_1, i_2 为子空间 V_1, V_2 分别到 V 的**包含映射**, p_1, p_2 为 V 关于直和分解

$$V = V_1 \oplus V_2$$

分别到 V_1, V_2 的**投影算子**。

进一步的，我们注意到

$$P_1 := i_1 p_1 \in End(V)$$

是 V 上的一个线性变换，它满足

$$P_1^2 = i_1 p_1 i_1 p_1 = i_1 p_1 = P_1.$$

于是我们引入更为一般的概念

定义 32. 若线性变换 $P \in End(V)$ 满足

$$P^2 = P$$

则称 P 是 V 上的一个 **投影 *(projection)* 算子**。有时也称 P 是一个**幂等 *(Idempotent)* 算子**。

投影（幂等）算子是代数学中频繁出现的对象。

命题 59. 给定线性空间 V，则 V 上的投影算子集合，与 V 的直和分解是一一对应的。

在此我们需要注意，仅仅给定线性空间 V 的一个子空间 V_1 并不足以确定一个 V 到子空间 V_2 的投影算子，而需要进一步由补空间 V_2 的选择来决定。

设 $A \in Hom(V, W)$, 并且 V, W 分别有直和分解

$$V = V_1 \oplus V_2, W = W_1 \oplus W_2.$$

对应的嵌入和投影算子分别为

$$V_1 \underset{p_1}{\overset{i_1}{\rightleftarrows}} V_1 \oplus V_2 \underset{i_2}{\overset{p_2}{\rightleftarrows}} V_2$$
$$W_1 \underset{q_1}{\overset{j_1}{\rightleftarrows}} W_1 \oplus W_2 \underset{j_2}{\overset{q_2}{\rightleftarrows}} W_2,$$

则我们相应的得到由直和分解诱导的子空间之间的线性映射

$$A_{11} = q_1 \circ A \circ i_1 \in Hom(V_1, W_1),\ A_{12} = q_2 \circ A \circ i_1 \in Hom(V_1, W_2)$$
$$A_{21} = q_1 \circ A \circ i_2 \in Hom(V_2, W_1),\ A_{22} = q_2 \circ A \circ i_2 \in Hom(V_2, W_2)$$

我们也可以将 A 形式的写为分块形式

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

因而这给出了矩阵分块的一个解释。

2.7.3 商空间及其诱导的线性变换

下面我们引入商空间的概念。它与上面的直和概念具有一定关系但又存在一些本质的不同。

给定一个线性空间 V 和一个子空间 $W \subseteq V$。定义一个 V 上的二元关系

$$\sim_W : u \sim_W v \text{ 当且仅当} u - v \in W.$$

由线性空间性质，$\sim_W$ 是一个等价关系。我们把相应的等价类

$$V/\sim_W$$

称作 **线性空间 V 模掉子空间 W 得到的商空间** ，记作

$$V/W.$$

由定义看出，V 中的一个等价类是形如

$$v+W$$

的一个集合，我们称其为 W 在 V 中的**陪集**。容易验证由陪集构成的等价类，即商空间也是一个线性空间：

定理 14. *线性空间 V 上的加法和数乘良定义在商空间 V/W 上, 故商空间也是一个线性空间。具体的，*

$$(u+W)+(v+W):=(u+v)+W, a(u+W):=au+W.$$

进一步的，映射

$$\pi: V\to V/W$$

*也是一个线性映射，称作从 V 到商空间 V/W 的***标准投射***。*

在前面讨论直和空间的部分我们引入了补空间的概念。我们看到商空间与补空间在概念上是不同的，但补空间和商空间仍存在一定联系。

定理 15. *若 W' 为 W 在 V 中的补，则在 W 关于 V 的任何一个陪集中存在唯一的一个 W' 中的代表元。*

进一步的，我们需要讨论 V 上的线性变换是否可以“诱导”为一个商空间上的线性变换。首先我们需要注意到，商空间分别作为线性映射的“源”和“汇”时，情况是由很大不同的。一方面若由映射

$$B: U\to V,$$

则同自然投射 π 复合

$$\pi\circ B: U\to V/W$$

自然的得到了 U 到 V/W 的线性映射。

另一方面，一个线性映射

$$A: V\to U$$

未必能自然的“诱导”一个从商空间 V/W 到 U 的线性映射 $\tilde{A}$，使得

$$\tilde{A} \circ \pi = A.$$

由于良定义的问题，一般情况下 $\tilde{A}$ 可能会不存在，但是容易得到：

引理 5. 以上诱导映射 $\tilde{A}$ 存在，或者说 A 良定义在 V/W 上当且仅当

$$W \subseteq \mathcal{N}(A),$$

进一步的，若”源“和”汇“都为商空间 V/W, 那么由上面分析得到

命题 60. 一个线性自同态

$$A : V \to V$$

良定义成为一个商空间 V/W 上的线性自同态，当且仅当 W 是 A 的不变子空间，即

$$AW \subseteq W.$$

上述概念在处理矩阵标准型时起着基本作用。

2.8 对偶空间和对偶线性变换

2.8.1 对偶空间及其诱导的线性变换

在前面我们已经讨论了一般的线性变换，并且线性变换自身也天然具有线性空间的结构，因而给出了构造新的线性空间的一种方法。在本节我们讨论以一维线性空间 k 为像空间的线性映射，即**对偶空间**，它也是线性空间的重要基本构造之一。

定义 33. 若 $f : V \to k$ 是一个对任意 $c_1, c_2 \in k$ ，$v_1, v_2 \in V$ 成立

$$f(c_1v_1 + c_2v_2) = c_1f(v_1) + c_2f(v_2)$$

的映射，则称 f 为 V 上的一个 k-**线性函数（或线性泛函）**。

命题 61. 对任意的线性函数 f_1, f_2, 以及 $c \in k, v \in V$, 定义标量乘法

$$(c \cdot f)(v) := cf(v)$$

和加法

$$(f_1 + f_2)(v) := f_1(v) + f_2(v),$$

则全体 k-线性函数在以上自然定义的运算下构成一个 n 维线性空间，称作 V **的对偶空间**, 记作 V^*。

定理-定义 4. 假设 $\mathcal{B}=\{v_1,\cdots,v_n\}$ 是 V 的一组基底，则满足

$$v_i^*(v_j)=\delta_{ij}$$

的 n 个线性泛函 $v_1^*,\cdots,v_n^*\in V^*$ 构成了 V^* 的一组基，称作关于 V 的基底 $\mathcal{B}$ 的**对偶基**。

下面来进一步考虑对偶空间之间的线性变换。设 $A\in Hom(V,W)$, 由于 W^* 由 W 上的（线性）函数构成，由预备知识知道，相应的对偶空间的变换应当是反向的。对任意的 $f\in W^*$, 容易验证 A 与 f 的复

$$A\circ f$$

是 V 上的一个 k-线性函数，即它是 V^* 中的一个元素。更进一步的，映射

$$f\mapsto A\circ f$$

是 W^* 到 V^* 的一个线性变换，称作 A 的 **对偶线性变换**，记作

$$A^\vee.$$

我们看到 $A^\vee\in Hom(W^*,V^*)$. 对偶线性变换的另一种等价描述是，$A^\vee$ 满足对所有 $f\in W^*, v\in V$ 成立

$$\langle A^\vee f,v\rangle=\langle f,Av\rangle.$$

定理-定义 5. 对任意一个线性变换 $A\in End(V)$, 存在唯一的线性变换 $A^\vee\in End(V^*)$ 使得

$$\langle A^\vee f,v\rangle=\langle f,Av\rangle$$

对任意 $f\in V^*, v\in V$ 成立。$A^\vee$ 称作 A 的**对偶变换**。

一个空间 V 的对偶空间 V^* 还可以进一步的取二次对偶 $(V^*)^*$, 也可简单记作

$$V^{**}.$$

我们注意到，每个向量 $v\in V$ 也可自然的视作 V^{**} 中的元素，即

$$\langle v,f\rangle:=\langle f,v\rangle=f(v).$$

容易得到

$$(AB)^\vee=B^\vee A^\vee.$$

2.8.2 对偶空间和线性变换的矩阵表示

利用对偶空间我们可以重新叙述线性变换、线性空间等概念与其矩阵表示之间的关系。回忆当线性空间 V 取定了一组基底 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$ 之后，任何一个向量

$$v = c_1 v_1 + \cdots + \cdots c_m v_m \in V.$$

都可以有列向量坐标表示

$$M_{\mathcal{B}_V}(v) = \begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix}$$

其中 c_i 即表示第 i 个坐标分量的数值。

利用对偶空间，一个更便于分析的观点是，我们应当将系数 c_i 视作 V 上的一个**线性函数**，即一个对偶空间中的元素 v_i^* 在 v 的**取值**:

$$v_i^* : V \to k,\, v \mapsto v_i^*(v) = \langle v_i^*, v \rangle = c_i.$$

因而

$$M_{\mathcal{B}_V}(v) = \begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix} = \begin{pmatrix} \langle v_1^*, v \rangle \\ \cdots \\ \langle v_m^*, v \rangle \end{pmatrix}.$$

与此同时，我们还注意到映射

$$\begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix} \mapsto c_i$$

恰好等于第 i 个标准列向量 e_i 与其做标量积得到的结果，即

$$(0, \cdots, \underbrace{1}_{\text{第}i\text{个位置}} \cdots, 0) \begin{pmatrix} c_1 \\ \cdots \\ c_m \end{pmatrix} = c_i,$$

并且行向量 e_i^t 恰好为 v_i^* 在对偶基 $\mathcal{B}_V^* := \{v_1^*, \cdots, v_n^*\}$ 下列向量坐标表示的**转置**:

$$(0, \cdots, \underbrace{1}_{\text{第}i\text{个位置}} \cdots, 0) = M_{\mathcal{B}_V^*}(v_i^*)^t.$$

因此，若我们取定列向量为向量在一组基底下的坐标表示，则行向量为对偶空间中的向量，在对偶基下的坐标表示，并且标准对和标量积吻合。一般的我们有：

命题 62. 对任意 $f \in V^*$, $v \in V$,

$$\langle f, v\rangle = M_{\mathcal{B}_V^*}(f)^t \cdot M_{\mathcal{B}}(v)$$

我们注意到，上面命题的等式中，左侧与基底选取无关，而右侧与基底选取是有关的。

类似的，对给定的一个线性变换

$$A : V \to W,$$

去对应的取定基底之后的表示系数 $(a_{i,j})$

$$Av_i = \sum_{ij} a_{ij} w_j$$

也可以很方便的用对偶空间来进行描述。具体的来说，用 w_k^* 作用在两侧，我们得到

$$a_{ik} = \langle w_k^*, \sum_{ij} a_{ij} w_j\rangle = \langle w_k^*, Av_i\rangle.$$

更一般的，由 A 可以诱导 $W^* \times V$ 上的一个双线性函数

$$(f, v)_A := \langle f, Av\rangle, \tag{12}$$

而 a_{ik} 即为以上双线性函数在特殊情况 $f = w_k^*, v = v_i$ 时的取值。因此在许多文献中，我们也把 (12) 称作 A (关于元素 f, v) 的**矩阵系数**，它自然的推广了通常意义下的矩阵系数 a_{ik}。

我们在线性映射部分描述的线性映射关系 $w = Av$ 的矩阵表示

$$M_{\mathcal{B}_W}(w) = M_{\mathcal{B}_W}(Av) = M_{\mathcal{B}_V}^{\mathcal{B}_W}(A) M_{\mathcal{B}_V}(v)$$

实际上应当理解为对偶线性变换 $A^\vee \in Hom(W^*, V^*)$ 在 W^* 上的作用。

具体的来说，令

$$M_{\mathcal{B}_w}(w) = \begin{pmatrix} d_1 \\ \cdots \\ d_n \end{pmatrix} = \begin{pmatrix} \langle w_1^*, w\rangle \\ \cdots \\ \langle w_n^*, w\rangle \end{pmatrix}$$

是 w 的坐标列向量表示，则命题 57 当中的矩阵表示关系写为

$$\begin{pmatrix} d_1 \\ \cdots \\ d_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

即

$$d_i = \sum_j a_{ij} c_j.$$

利用对偶空间的记号，上式转换为

$$\langle w_i^*, w\rangle = \sum_j a_{ij} \langle v_j^*, v\rangle.$$

利用

$$w = Av$$

以及对偶变换的性质我们得到

$$\langle A^\vee w_i^*, v\rangle = \langle w_i^*, Av\rangle = \langle w_i^*, w\rangle = \sum_j \langle a_{ij} v_j^*, v\rangle.$$

由 v 的任意性，我们得到

$$A^\vee w_i^* = \sum_j a_{ij} v_j^*$$

是对偶变换 $A^\vee$ 的作用表达式，

因此若假定

$$A^\vee w_i^* = \sum_j a_{ji}^\vee v_j^*,$$

即其中待定系数 $a_{ji}^\vee$ 是表示矩阵 $M_{\mathcal{B}_W^*}^{\mathcal{B}_V^*}(A)$ 的矩阵元，则我们发现

$$a_{ki}^\vee = a_{ik},$$

即表示矩阵 $M_{\mathcal{B}_W^*}^{\mathcal{B}_V^*}(A)$ 和表示矩阵 $M_{\mathcal{B}_V}^{\mathcal{B}_W}(A)$ 恰好互为转置。

命题 63. 假设 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$，$\mathcal{B}_W = \{w_1, \cdots, w_n\}$ 分别为 V, W 的一组基，$\mathcal{B}_V^* = \{v_1^*, \cdots, v_m^*\}$，$\mathcal{B}_W^* = \{w_1^*, \cdots, w_n^*\}$ 为相应的对偶基，则任意的线性变换 $A \in Hom(V, W)$ 及其对偶线性变换 $A^\vee$，他们的矩阵表示满足关系

$$M_{\mathcal{B}_W^*}^{\mathcal{B}_V^*}(A^\vee) = (M_{\mathcal{B}_V}^{\mathcal{B}_W}(A))^t.$$

进一步的，上面的矩阵表示关系与

与之相仿但略有不同的是对偶基在坐标变换下的变化。

命题 64. 假设 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$ 是 V 的一组基，$\mathcal{B}_V^* = \{v_1^*, \cdots, v_m^*\}$ 为其对偶基，$\mathcal{B}'_V = \{v'_1, \cdots, v'_m\}$ 是 V 的另一组基，$\mathcal{B}_V^{*'} = \{v_1^{*'}, \cdots, v_m^{*'}\}$ 是 $\mathcal{B}'_V$ 的对偶基，则

$$M_{\mathcal{B}_V}^{\mathcal{B}'_V} \cdot (M_{\mathcal{B}_V^*}^{\mathcal{B}_V^{*'}})^t = I_m.$$

2.8.3 方阵和线性变换的迹

迹是方阵和线性变换空间上重要的线性函数。

定义 34. 定义方阵 $A=(a_{ij})_{n\times n}$ 的 **迹** 为

$$Tr(A) := \sum_{i=1,\cdots,n} a_{ii}.$$

命题 65. 迹函数

$$A \mapsto Tr(A),\ M_n(k) \to k$$

是线性函数，且对任意的 $A, B \in M_n(k)$,

$$Tr(AB) = Tr(BA).$$

反之若线性函数 $f: M_n(k) \to k$ 满足对任意的 $A, B \in M_n(k)$ 成立

$$Tr(AB) = Tr(BA), Tr(I_n) = n, \tag{13}$$

则

$$f = Tr.$$

命题 66. 设 $\mathcal{B}_V = \{v_1, \cdots, v_m\}$ 是 V 的一组基，$\mathcal{B}_V^* = \{v_1^*, \cdots, v_m^*\}$ 为其对偶基，则

$$Tr(A) = \sum_i \langle v_i^*, Av_i \rangle,$$

且 $Tr(A)$ 的以上定义与基底 $\mathcal{B}_V$ 取法无关。

在张量部分我们会重新讨论以上命题。

2.9 双线性映射与双线性 (二次) 型

在前面我们已经完成了线性代数基本理论的讨论。在线性代数和线性几何当中，尽管有一些重要的映射并非线性映射，但是与线性映射具有紧密的关系，这就是本节要讨论的多重线性映射。这其中尤为重要的是**双线性映射**和**双线性型**。

2.9.1 线性映射

设 V_i, W 都是有限维线性空间。

定义 35. 称一个映射

$$F : V_1 \times \cdots \times V_m \to W$$

为从 $V_1, \cdots, V_m$ 到 W 的 **多重线性映射**，如果对任意 $1 \le i \le m$，在固定元素 $v_j \in V_j, j \neq i$ 后，F 关于第 i 个分量的映射

$$\{v_1\} \times \cdots V_i \cdots \times \{v_m\} \to W, \, v \mapsto F(v_1, \cdots, v, \cdots, v_m)$$

是线性的，即

$$\begin{aligned} &F(v_1, \cdots, c'v_i' + c''v_i'', \cdots, v_m) \\ =&c'F(v_1, \cdots, v_i', \cdots, v_m) + c''F(v_1, \cdots, v_i'', \cdots, v_m) \end{aligned}$$

对任意 $c', c'' \in k$ 和 $v_i', v_i'' \in V_i$ 成立。

当 $W = k$ 是一维线性空间时，称 F 为**多重线性函数**，更进一步的，当 $m = 2$ 时也称 F 为**双线性函数**或**双线性型**。

特别的若 $V_1 = V_2 = V$ 是同一个空间，则也称 V_1, V_2 上的双线性函数为为关于 V 上的**双线性型**。双线性函数和双线性型相较于一般的多重线性映射，与线性映射有着更为紧密的联系。

需要注意的是，若我们将笛卡尔积 $V_1 \times V_2$ 等同于直和空间

$$V_1 \oplus V_2,$$

那么 V_1, V_2 上的双线性映射一般不是 $V_1 \oplus V_2$ 上的线性映射。

例. 从定义出发容易得到存在一个关于 V^*, V 的标准双线性映射 $\langle \cdot, \cdot \rangle$: $V^* \times V \to k$:

$$\langle f, v \rangle := f(v),$$

称作 V 和其对偶空间 V^* 之间的 **标准对**。

例. 给定一个 V 上的双线性映射 B, 它自然的诱导了一个从 V 到 V^* 的线性映射,

$$v \mapsto B(v, \cdot),$$

即元素 v 被映射为一个线性函数 $f_v := B(v, \cdot)$:

$$f_v(v') := B(v, v').$$

我们注意到，若给定了线性空间 V, 则其对偶空间 V^* 以及相应的标准对 $\langle \cdot, \cdot \rangle$ 是仅仅依赖于 V 天然存在而无需其他结构的。与此不同的是，V 上的双线

性函数是具有许多选择而不存在标准选取的，这一情况类似于 V 上一组基底的非唯一性。

在第一节我们曾经提到过，矩阵还可以用来表示一个双线性映射。取 U,V 的一组基底 $\mathcal{B}_U := \{u_1, \cdots, u_l\}$ $\mathcal{B}_V := \{v_1, \cdots, v_m\}$ 以及 U,V 上的双线性函数 $B : U \times V \to k$, 我们得到一个 $l \times m$ 矩阵

$$(B(u_i, v_j))_{l\times m}.$$

反之由双线性性，若取定了同样的一组基底，任意给定一个 $l \times m$ 矩阵

$$B := (b_{ij})_{l\times m}$$

都唯一确定了 V 上的一个双线性函数 B 满足

$$B(u_i, v_j) = b_{ij},$$

它通过双线性性拓展为 V 上的双线性函数。类似的我们记作

$$M_{\mathcal{B}_U,\mathcal{B}_V}(B) := (b_{ij})_{l\times m}$$

也称为 **基底 $\mathcal{B}_U$，$\mathcal{B}_V$** 下 B 的表示矩阵。

在许多讨论当中，我们还常常将一个双线性函数用多元函数的形式表达。延续上文的设定，通过直接计算我们可以得到

命题 67. 对任意 $u \in U, v \in V$，若在基底 $\mathcal{B}_U$, $\mathcal{B}_V$ 之下，它们的坐标表示为

$$u = \sum_i x_i u_i,\ v = \sum_j y_j v_j,\ x_i, y_j \in k,$$

则

$$B(u, v) = \sum_{i=1,\cdots,l;j=1,\cdots,m} b_{ij} x_i y_j.$$

以上命题表明，若我们将系数 $x_1, \cdots, x_l, y_1, \cdots, y_m$ 看作是是 $l + m$ 个自变量，则矩阵

$$B = (b_{ij})_{l\times m}$$

唯一决定了在基底 $\mathcal{B}_U$, $\mathcal{B}_V$ 之下对应的双线性函数，它可以由关于自变量 $x_1, \cdots, x_l; y_1, \cdots, y_m$ 的 $l + m$ 元二次齐次函数

$$B(x, y) := \sum_{i=1,\cdots,l;j=1,\cdots,m} b_{ij} x_i y_j$$

来表达。进一步的，如果我们以 l 维和 m 维列向量形式

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_l \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

表示对应的自变量，则双线性函数还可以用矩阵乘法形式表示

$$B(u,v) = B(x,y) = x^t B y.$$

若我们再次回忆 2.8.2 当中关于向量坐标系数和对偶空间的讨论，则自变量 x_i, y_j 不仅仅是通常的标量数字，而应当进一步视作对偶基 $\mathcal{B}_U^*$, $\mathcal{B}_V^*$ 中的元素，即它们作为 U, V 上的线性函数。具体的我们有

$$x_i(u) = \langle u_i^*, u\rangle,\, y_i(v) = \langle v_i^*, v\rangle.$$

需要注意的是，即使在 $U = V$ 是同一个线性空间的情况下（即 $B(\cdot,\cdot)$ 是 V 上的双线性型），仅用 l 个变量也还不足以刻画 $B(\cdot,\cdot)$。最一般的情况下我们仍然需要将 $B(\cdot,\cdot)$ 视为 $2l$ 个变量 $x_1, \cdots, x_l; y_1, \cdots, y_l$ 的函数。

2.9.2 双线性函数、对偶空间，和线性变换

在下面的讨论中，我们进一步假设 $\mathcal{B}_U^* := \{u_1^*, \cdots, u_l^*\}$ $\mathcal{B}_V^* := \{v_1^*, \cdots, v_m^*\}$ 为相应的对偶基。

回忆在 V 到自身的全体线性变换 $End(V)$ 中，**可逆线性变换**是较为特殊的一类线性变换。对于 V 上的双线性函数，我们同样有类似于可逆线性变换的概念。

定义 36. 给定 V, W 上的双线性函数 B，定义 B 的**左根**

$$Rad_L(B) := \{v|\, v \in V,\, B(v,y) = 0 \text{ 对所有} y \in W \text{ 成立}\}$$

和**右根**

$$Rad_R(B) := \{w|\, w \in W,\, B(x,w) = 0 \text{ 对所有} x \in V \text{ 成立}\}.$$

若

$$Rad_L(B) = \{0\}$$

则称 B **左非退化**，并且类似的若

$$Rad_R(B) = \{0\}$$

则称 B **右非退化**。

我们注意到给定 V, W 上的双线性函数 $B(\cdot,\cdot)$，自然的有相应的线性映射

$$A_B^L : v \mapsto B(v,\cdot),\, V \to W^*,$$

以及

$$A_B^R : w \mapsto B(\cdot, w),\, W \to V^*.$$

事实上我们注意到，利用基底我们可以写出映射 A_B^L, A_B^R 的表达式：

$$A_B^L(v) = \sum_i B(v, w_i) w_i^*,\, A_B^L(w) = \sum_i B(v_i, w) v_i^*.$$

由于上面两式左侧中 A_B^L, A_B^R 的原像基底取法无关，因此上面的表达式右侧实际上与基底取法无关。我们在后面会反复利用这一论证模式。由定义我们得到：

命题 68. *V, W 上的双线性函数 B 左非退化当且仅当线性映射 A_B^L 是单射；类似的，B 左非退化当且仅当线性映射 A_B^R 是单射.*

我们还注意到

$$B \mapsto A_B^L$$

给出了 V, W 上的双线性函数 $B(\cdot,\cdot)$ 到 $Hom(V, W)$ 的线性映射。反之，若给定一个线性映射

$$A : V \to W^*,$$

则

$$B_A^L(v, w) := \langle A(v), w \rangle$$

显然是 V, W 上的双线性函数。进一步的我们还注意到

$$B \mapsto A_B^L$$

与

$$A \mapsto B_A^L$$

是两个互为逆的线性映射。于是我们得到

命题 69. *V, W 上的双线性函数 $B(\cdot,\cdot)$ 与 V 到 W 的线性映射全体 $Hom(V, W)$ 作为线性空间是自然（线性）同构的。*

在这一命题中我们尚未解释何为“自然同构”。回忆在 2.3.3 当中我们已经指出，线性空间的维数完全标记了线性空间的线性同构等价类，而此命题当中的线性空间，还额外具有双线性函数和线性映射的结构，因而一般的线性空间

同构相较于“自然同构”，未反映出额外的结构而显得更为薄弱，我们将在范畴论的部分对此予以进一步解释。

下面进一步讨论 $V=W$ 的情形。作为上面的特例我们得到

命题 70. V 上的双线性函数 $B(\cdot,\cdot)$ 与 V 上线性变换的全体 $End(V)$ 作为线性空间是线性同构的。

定理-定义 6. 若 $V=W$, 则 V 上的双线性型 B 左非退化当且仅当其右非退化，因此在这两种情况下我们都简称 B 非退化。

定理 16. B 是 V 上的非退化双线性型，当且仅当在某一组基底 $\mathcal{B}_V$ 下的矩阵

$$(b_{ij})_{n\times n}$$

可逆，或等价的，对任意 $v\in V$, B 诱导的线性映射

$$v\mapsto B(v,\cdot)(\text{ 或}v\mapsto B(\cdot,v)),\ V\to V^*$$

是可逆映射。

以上的定理即是所谓 Riesz 表示定理在有限维空间的版本，即对任意线性函数 $f\in V^*$，都可以由一个相应的向量 $v^L_{B,f}$（或 $V^R_{B,f}$）表示：

$$f(v')=B(v^L_{B,f},v')(\text{ 或}f(v')=B(v',v^R_{B,f})\text{ 对任意}v'\in V.$$

特别的,关于一组基底 $\mathcal{B}_V=\{v_1,\cdots,v_n\}$ 的对偶基 $\mathcal{B}^*_V=\{v^*_1,\cdots,v^*_n\}$ 中的每个元素 v^*_i,都可以用相应的一个向量 $v^L_{B,v^*_i}\in V$ 来表示,因而为了简便起见,我们也常将(左)表示向量 v^L_{B,v^*_i} 用简单的用 v^*_i 表达,并且也称 $\mathcal{B}^*_V=\{v^*_1,\cdots,v^*_n\in V\}$ 为 $\mathcal{B}_V=\{v_1,\cdots,v_n\}$（关于非退化双线性型 $B(\cdot,\cdot)$ ）的对偶基。由此我们经常说，V 的对偶空间 V^*，通过非退化双线性函数 $B(\cdot,\cdot)$，自然同构于 V。对于右表示向量的情况是类似的，在此不再重复。

2.9.3 共轭变换

定理-定义 7. 设 B_1,B_2 分别为 V, W 上的非退化双线性型, 对任意线性映射 $A\in Hom(V,W)$, 存在唯一的线性映射 $A^*\in Hom(W,V)$ 使得

$$B_2(Av,w)=B_1(v,A^*w)$$

对任意 $v\in V,w\in W$ 成立。称 A^* 为 A 的 (关于非退化双线性型 B_1 和 B_2) **右伴随（共轭）变换**。类似的，对任意线性映射 $A\in Hom(W,V)$, 存在唯一的线性映射 ${}^*A\in Hom(V,W)$ 使得

$$B_2({}^*Av,w)=B_1(v,Aw)$$

对任意 $v \in V, w \in W$ 成立。称 ${}^*A$ 为 A 的 *(关于非退化双线性型 B_1 和 B_2)* **左伴随（共轭）变换**。

事实上由同样的理由，我们注意到 W 到 V 的线性映射

$$A_R(w) := \sum_i B_2(Av_i^*, w)v_i,\ A_L(v) := \sum_i B_1(v, Aw_i^*)w_i$$

在上面的讨论中我们看到，一般情况下我们总是要进行“左”和“右”的区分。在应用当中常见的双线性型通常还具有更多的“对称性”，它们是更为重要的研究对象。

定义 37. 一个 V 上的双线性型 B 若满足

$$B(u,v) = B(v,u)$$

则称其为**对称双线性型**。类似的，B 若满足

$$B(u,v) = -B(v,u)$$

则称其为**反对称双线性型**。

命题 71. 双线型 $B(\cdot,\cdot)$ 对称（或反对称），当且仅当在 V 的某一组基底 $\mathcal{B}_V$ 下，表示矩阵 $M_{\mathcal{B}_V}$ 对称（或反对称）。

对称双线性型是较为特殊的双线性型。

定义 38. 对 V 上给定的对称双线性型 $B(\cdot,\cdot)$，定义 V 上的函数

$$Q_B(v) := B(v,v).$$

我们称 $Q_B(\cdot)$ 为 V 上 ***(和双线性性函数 $B(\cdot,\cdot)$ 相关联的) 二次型 (Quadratic form associated with the symmetric bilinear form $B(\cdot,\cdot)$)***。

显然 V 上的二次型也构成一个线性空间。

命题 72. 若 $char(k) \neq 2$ 则 V 上的对称双线性型与二次型构成的线性空间同构。

在定义中我们已经给出了从对称双线性函数到二次型的对应，因而我们需要给出反方向的对应。若给定了 V 上与对称双线性函数 $B(\cdot,\cdot)$ 相关联的二次型 $Q(\cdot) := Q_B(\cdot)$，定义 $V \times V$ 上的函数

$$B_Q(u,v) := \frac{1}{2}(Q(u+v,u+v) - Q(u,u) - Q(v,v)).$$

我们容易验证

$$B(\cdot,\cdot) = B_Q(\cdot,\cdot),$$

等于原先的对称双线性函数 $B(\cdot,\cdot)$, 因而

$$B \mapsto Q_B,\, Q \mapsto B_Q$$

给出了两个线性空间之间的一个线性同构。

类似的，若取定了基底 $\mathcal{B}_U, \mathcal{B}_V$，并且以关于系数的 $l+m$ 元函数

$$B(x,y) := \sum_{i=1,\cdots,l;j=1,\cdots,m} b_{ij}x_iy_j$$

表达对称双线性函数 $B(\cdot,\cdot)$，则对应二次型为

$$Q(x) := Q_B(x) = B(x,x) = x^tBx,$$

并且

$$B(x,y) = \frac{1}{2}(Q(x+y,x+y) - Q(x,x) - Q(y,y)).$$

由此我们看到与一般双线性型不同的是，由于对称性，V 上的二次型仅仅用 l 元变量 $x_1,\cdots,x_l$ 就可以进行表示。

特别的，当 $V=W$ 时 $A^*, {}^*A \in End(V)$。

定义 39. 若 B 是 V 上的非退化双线性型，$A \in End(V)$ 满足

$$A^* = A$$

则称 A 是 *(关于 B 的)* 一个自伴随（共轭）算子。类似的若 $A \in End(V)$ 满足

$$A^* = -A$$

则称 A 是 *(关于 B 的)* 一个反（斜）自伴随（共轭）算子。

下一章我们会看到，在某些情况下, 还需要对双线性型加上“扭曲”。

2.9.4 半双线性型

定义 40. 假定 k 是一个域，若 $\sigma \in Aut(k)$ 且

$$\sigma^2 = Id_k,$$

则称 σ 为 k 的一个**对合** ***(involution)***。

定义 41. 给定 k 的对合 σ，以及一个 k-线性空间 V。若一个映射 $A: V \to W$ 满足

$$A(v_1 + v_2) = Av_1 + Av_2,\ A(cv) = \sigma(c)Av,$$

则称 A 是（关于 A 和 σ）**共轭线性**的。

定义 42. 一个 V 上**半双线性型** ***(sesquilinear form)*** 是一个满足如下条件的映射 $B: V \times V \to k$

1. $B(c_1v_1 + c_2v_2, v) = c_1B(v_1, v) + c_2B(v_2, v)$。
2. $B(v, c_1v_1 + c_2v_2) = \sigma(c_1)B(v, v_1) + \sigma(c_2)B(v, v_2)$。

也就是说，半双线性型关于第一个分量是线性的，但关于第二个分量是共轭线性的。

例 26. 考虑复数域 $k = \mathbb{C}$，σ 为复共轭：

$$\sigma(c) = \bar{c},$$

以及相应的 n 维列向量空间 $V = \mathbb{C}^n$，则容易验证

$$(x, y) := x^t\bar{y}$$

是 V 上的半双线性函数。我们也称之为 $\mathbb{C}^n$ 上的**标准内积**。

关于 $\mathbb{C}^n$ 上内积的讨论将在内积空间的部分进行。

2.10 行列式与线性相关性

在前面我们已经利用矩阵初等变换，给出了求解一般线性方程组的 Gauss 消元算法。我们看到 Gauss 消元法并没有给出线性方程组的具体公式。由于一般线性方程组可能无解，也可能有无穷多解，因而我们不能期待存在这样的具体公式。但是，对于“可逆方阵”的情况，我们知道此时存在唯一解。在这时我们可以利用矩阵的**行列式**给出解的具体表达式。

历史上人们对行列式的研究远远早于矩阵概念的引入。但我们看到，线性代数中大部分基本结论的建立，并不需要依赖于行列式。不过在更深层次的理论中，行列式扮演者比较重要的角色。有时利用行列式可以给出一些结论较为简洁的证明。

2.10.1 置换和对称群

行列式概念的引入需要引入**对称群**及其相关的一些概念。

定义 43. 给定一个 n 元集合 $\mathcal{S}(n)$，则称 $\mathcal{S}(n)$ 到自身的一个双射 $\sigma : \mathcal{S}(n) \to \mathcal{S}(n)$ 为 $\mathcal{S}(n)$ 上的一个**置换** ***(permutation)***。$\mathcal{S}(n)$ 上的所有置换构成一个群，称作 n **元对称群**，记作 S_n。

由于对称群中的元素都是置换，因而有时也称对称群为**置换群**。通常我们取 n 元集合 $\mathcal{S}(n)$ 为整数 $1, \cdots, n$ 构成的集合：

$$\mathcal{S}(n) = \{1, \cdots, n\}.$$

特别的，对任意两个不同的整数 i, j，定义

$$(ij) : i \mapsto j,\, j \mapsto i,\, k \mapsto k \text{对所有} k \neq i, j,$$

为交换 i, j 并保持其他元素不同的映射，显然它是一个置换。我们把这样的特殊置换称作**对换**（**transposition**）。由归纳法容易证明

命题 73. 任何一个置换 σ 都是对换的乘积，即存在对换 $\tau_1, \cdots, \tau_k$ 使得

$$\sigma = \tau_1 \cdots \tau_k.$$

我们注意到用对换乘积表达置换的方式一般不是唯一的。

定义 44. 定义 ***Vandermonde*** **多项式**

$$V(x) := \prod_{1 \leq i < j \leq n} (x_i - x_j),$$

并定义

$$\Delta(x) := V(x)^2 = \prod_{1 \leq i < j \leq n} (x_i - x_j)^2$$

对一个置换 $\sigma \in S_n$ 和一个多项式 $f(x)$，我们定义多项式

$$\sigma(f)(x_1, \cdots, x_n) := f(x_{\sigma^{-1}(1)}, \cdots, x_{\sigma^{-1}(n)})$$

为 σ 排列下标所得到的新多项式。容易验证

$$\sigma(fg) = \sigma(f)\sigma(g),\, (\sigma\tau)(f) = \sigma(\tau(f)) \tag{14}$$

由于 $\sigma : \mathcal{S}(n) \to \mathcal{S}(n)$ 是双射，因而

$$\sigma : \{(x_i - x_j)^2 | \, 1 \leq i < j \leq n\} \to \{(x_{\sigma(i)} - x_{\sigma(j)})^2 | \, 1 \leq i < j \leq n\}$$

诱导出的也是双射，故

$$\sigma(\Delta)(x) = \Delta(x),$$

即 $\Delta(x)$ 是一个**对称多项式**。因此由

$$(\sigma(V)(x))^2 = V(x)^2 = \Delta(x)$$

得到

$$\sigma(V)(x) = V(x) \text{ 或者 } -V(x).$$

由此我们引出

定理-定义 8. 对任意置换 $\sigma \in S_n$，定义它的**符号**（***sign***） $sgn(\sigma)$ 为使得

$$\sigma(V)(x) = sgn(\sigma)V(x)$$

成立的数，它的取值为 ±1。符号为 1 和 -1 的排列也分别称为是奇（排列）和偶（排列）的。

特别的容易看到

引理 6. 若 τ 是一个对换，则

$$sgn(\tau) = -1.$$

对任意两个置换 σ, σ'，将它们作用在 Vandermonde 多项式上，由 (14)，

$$sgn(\sigma\sigma')(V)(x) = sgn(\sigma)(\sigma'(V))(x) = sgn(\sigma)sgn(\sigma')(V)(x),$$

因而推出

命题 74. 置换的符号具有可乘性，即

$$sgn(\sigma\sigma') = sgn(\sigma)sgn(\sigma') = sgn(\sigma')sgn(\sigma)$$

特别的，利用 $\sigma\sigma^{-1} = 1$ 我们容易得到

$$sgn(\sigma) = sgn(\sigma^{-1}).$$

结合引理我们得到利用对换定义符号的另一种方法，它在后面的证明中起到关键作用。

定理-定义 9. 若一个置换 σ 都是对换的乘积，即存在 k 个对换 $\tau_1, \cdots, \tau_k$ 使得

$$\sigma = \tau_1 \cdots \tau_k$$

则

$$sgn(\sigma) = (-1)^k.$$

由于定理-定义 8 并不依赖于置换表达成对换的乘积的形式，因而事实上我们还证明了以下结论，尽管这一结论我们不需要用到

命题 75. 对一个置换 σ 若存在两组对换 $\tau_1,\cdots,\tau_k$，$\tau_1',\cdots,\tau_l'$ 使得

$$\sigma=\tau_1\cdots\tau_k=\tau_1',\cdots,\tau_l',$$

则

$$sgn(\sigma)=(-1)^k=(-1)^l$$

因而 $k-l$ 总是一个偶数。

我们注意到可以考虑置换直接在 Vandermonde 行列式上的作用，来得到置换的符号。这一方法得到的是在文献中经常出现的利用**逆序数 (inversion)** 定义符号的方法。对任意一个置换 σ 和一对整数 $1\leq i\neq j\leq n$，令

$$k=min(\sigma(i),\sigma(j)),\, l=max(\sigma(i),\sigma(j)).$$

显然对 Vandermonde 多项式中相应的项 x_i-x_j，我们有

$$\sigma^{-1}(x_i-x_j)=x_k-x_l \text{ 或者 } -(x_k-x_l)$$

由于置换是 n 个元素的双射，因而置换 σ 排列了 Vandermonde 多项式乘积中的项，但是若 $\sigma(i)>\sigma(l)$，则相应的项为了与 Vandermonde 多项式比较，会贡献出一个负号，因此符号 $sgn(\sigma)$，由所有整数对 $1\leq i\neq j\leq n$ 在排列 σ 作用下所贡献的负号数量决定。

定理-定义 10. 对任何一个排列 σ，定义它的**逆序数**为满足

$$\sigma(i)>\sigma(j)$$

的有序对 $1\leq i<j\leq n$ 的个数。进一步，若 k 为 σ 的逆序数，则

$$sgn(\sigma)=(-1)^k.$$

逆序数通常便于计算排列的奇偶性，但它并不适合我们做推导。

2.10.2 行列式的定义与性质

定义 45. 对任意矩阵元在 k 中的方阵 $A=(a_{ij})_{n\times n}$，定义它的**行列式**为标量：

$$\sum_{\sigma\in S_n} sgn(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}. \tag{15}$$

我们注意到，以上定义中仅仅用到加减乘运算而未用到除法，因而以上定义可以放宽到矩阵元在任意交换环 R 中的方阵。这一推广在定义特征多项式时会用到。

命题 76. *n 阶行列式满足如下性质：*

1. （归一性）$det(I_n) = 1$

2. （（列）多重线性性）行列式对任何单独一列具有线性性，即对任意 $c \in k$ 有

$$det\begin{pmatrix} a_{11} & \cdots & ca_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & ca_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & ca_{1j} & \cdots & a_{nn} \end{pmatrix} = c \cdot det\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{1j} & \cdots & a_{nn} \end{pmatrix},$$

以及

$$det\begin{pmatrix} a_{11} & \cdots & a'_{1j} + a''_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a'_{ij} + a''_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a'_{nj} + a''_{nj} & \cdots & a_{nn} \end{pmatrix}$$

$$= det\begin{pmatrix} a_{11} & \cdots & a'_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a'_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a'_{nj} & \cdots & a_{nn} \end{pmatrix} + det\begin{pmatrix} a_{11} & \cdots & a''_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a''_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a''_{nj} & \cdots & a_{nn} \end{pmatrix}$$

3. *(列反对称性)* 交换任意不同两列行列式反号，即

$$det\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{1j} & \cdots & a_{1k} & \cdots & a_{nn} \end{pmatrix}$$

$$= - det \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{1k} & \cdots & a_{1j} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{1k} & \cdots & a_{1j} & \cdots & a_{nn} \end{pmatrix}$$

反之，任何一个满足以上三条的从 $M_n(k)$ 到 k 的映射都是 n 阶行列式（映射）。

行列式的公理化定义，体现出行列式的几何意义是由向量组构成高维空间中的平行多面体的**有向体积**。公理 1，2 符合我们对一般“体积”概念的认知，而公理 3 的反对称性则指出，在线性空间上有意义的“体积”，需要额外考虑带有符号的**有向**“体积”，而这自然要求行列式具有反线性性性。

反对称性是近现代数学和物理理论中普遍出现的现象。

我们可以通过行列式的定义直接验证它满足以上几条性质。为了证明满足以上三条性质的映射等于 (15) 所给出的表达式，我们需要做一些准备。

为方便起见假设 A 按列向量分块：

$$A = (v_1, \cdots, v_n).$$

于是多重线性性性可以重新叙述为

$$\begin{aligned} & det(v_1, \cdots, cv_i + c'v_i', \cdots, v_n) \\ =& c \cdot det(v_1, \cdots, v_i, \cdots, v_n) + c' \cdot det(v_1, \cdots, v_i', \cdots, v_n), \end{aligned}$$

而反对称性为

$$det(v_1, \cdots, v_i, \cdots, v_j, \cdots, v_n) = -det(v_1, \cdots, v_j, \cdots, v_i, \cdots, v_n).$$

首先，反复利用列反对称性，我们得到

引理 7. *若 A' 由 A 经过 k 次列变换得到，则*

$$det(A') = (-1)^k det(A).$$

特别的，

引理 8. *若 A 的 i, j 列相同 $(i \neq j)$，则*

$$det(A) = 0.$$

这是由于此时对换 (ij) 保持 A 不变，得到 $det(A) = -det(A)$ 因而推出推论。

引理 9. 设 $\mu : \{1, \cdots, \} \to \{1, \cdots, n\}$ 是 $\mathcal{S}(n)$ 到自身的一个映射。定义矩阵

$$P_\mu := \sum_{i=1,\cdots,n} E_{\mu(i),i}.$$

则

$$det(P_\mu) = \begin{cases} 0, & \text{若 } \mu \text{ 不是一个置换。} \\ sgn(\mu), & \text{若 } \mu \text{ 是一个置换。} \end{cases}$$

首先我们注意到，若 μ 不是一个置换，则存在 $i \neq j$ 使得 $\mu(i) = \mu(j)$，因而 P_μ 的第 i, j 两列完全相同，故由反对称性，

$$det(P_\mu) = 0.$$

若 μ 是一个置换，首先我们考虑 $\mu = (i, j)$ 是一个对换的情况。由于交换 $P_{(ij)}$ 的 i, j 列得到的是单位矩阵 I_n，因而由反对称性我们得到

$$det(P_\mu) = -1$$

又注意到若

$$\mu = \tau_1 \cdots \tau_l$$

则

$$P_\mu = P_{\tau_1} \cdots P_{\tau_l} I_n$$

并结合归一化条件 $det(I_n) = 1$ 得到

$$det(P_\mu) = (-1)^l det(I_n) = sgn(\mu).$$

我们延续假设 $A = (a_{ij})_{n\times n}$ 按列向量分块：

$$A = (v_1, \cdots, v_n),$$

根据多重线性性，我们可以将矩阵系数 a_{ij} 提出：

$$\begin{aligned} det(A) = det((v_1, \cdots, v_n)) =& det(\sum_{k_1} a_{k_1,i} e_{k_1}, \cdots, \sum_{k_n} a_{k_n,i} e_{k_n}) \\ =& \sum_{k_1,\cdots,k_n=1,\cdots,n} a_{k_1,1} \cdots a_{k_n,n} det((e_{k_1}, \cdots, e_{k_n})) \end{aligned}$$

因而 A 的行列式转化为每一列仅有一个元素为 1 的矩阵行列式的线性组合。注意到多重有序数组集合 $k_1, \cdots, k_n$ 与 $\mathcal{S}(n)$ 到自身的映射一一对应，且对给定的多重有序数组 $k_1, \cdots, k_n$ 若它对应映射

$$\mu : i \mapsto k_i,\, i = 1, \cdots, n$$

则

$$(e_{k_1}, \cdots, e_{k_n}) = P_\mu$$

为引理 7 当中定义的矩阵，因而仅当 $k_1, \cdots, k_n$ 互不相同，即 μ 是排列时，对应的单项非零，因此最终得到

$$det(A) = \sum_{\mu} a_{k_1,1} \cdots a_{k_n,n} det(P_\mu) = \sum_{\sigma \in S_n} sgn(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

定理 17. 设 $A \in End(V)$ 是一个线性变换。对 V 的任意两组基底 $\mathcal{B}, \mathcal{B}'$，我们有

$$det(M_{\mathcal{B}}(A)) = det(M_{\mathcal{B}'}(A)),$$

即线性变换 A 在任何一组基底下的表示矩阵的行列式都是同一个数值。我们将这个共同的数值定义为**线性变换 A 的行列式**。

下面是关于行列式的一些重要性质。他们都可以作为行列式公理化定义的推论。

命题 77. 对任意 n 阶矩阵 A 我们有

$$det(A^t) = det(A). \tag{16}$$

并且对两个任意 n 阶矩阵 A, B，

$$det(AB) = det(A)det(B) = det(BA). \tag{17}$$

我们利用定义 29 中给出的行列式的直接定义来证明命题。这个描述不可避免的要利用较为笨重的记号和繁琐的论证，但可以直接推广到矩阵元素为交换环的情形。对第一式，我们仅需注意到

$$sgn(\sigma^{-1}) = sgn(\sigma)$$

且映射

$$\sigma \mapsto \sigma^{-1}$$

给出了 S_n 到自身的一个双射，因而由行列式定义，

$$\begin{aligned} det(A) &= \sum_{\sigma \in S_n} sgn(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma^{-1}) a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} sgn(\sigma^{-1}) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = det(A^t). \end{aligned}$$

对于第二式我们仅需证明

$$det(A)det(B) = det(AB)$$

而另一半自动可以推出。由定义

$$\begin{aligned}
det(A)det(B) =& (\sum_{\sigma\in S_n} sgn(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)})(\sum_{\tau\in S_n} sgn(\tau)b_{1,\tau(1)}\cdots b_{n,\tau(n)}) \\
=& \sum_{\sigma,\tau\in S_n} sgn(\sigma\tau)(a_{1,\sigma(1)}b_{1,\tau(1)})\cdots(a_{n,\sigma(n)}b_{n,\tau(n)}) \\
=& \sum_{\sigma,\tau\in S_n} sgn(\sigma\sigma\tau)(a_{1,\sigma(1)}b_{\sigma(1),\sigma\tau(1)})\cdots(a_{n,\sigma(n)}b_{\sigma(n),\sigma\tau(n)}) \\
=& \sum_{\sigma,\tau\in S_n} sgn(\tau)(a_{1,\sigma(1)}b_{\sigma(1),\tau(1)})\cdots(a_{n,\sigma(n)}b_{\sigma(n),\tau(n)})
\end{aligned}$$

其中我们利用了哑指标替换

$$\tau\mapsto\sigma\tau$$

并注意到对固定的 σ 上面给出了 S_n 到自己的一个双射。另一方面设

$$C=(c_{ij})_{n\times n}=AB,$$

则由矩阵乘法

$$c_{ij}=\sum_k a_{ik}b_{kj},$$

因而仍由行列式定义，得到

$$\begin{aligned}
det(AB) = det(C) =& \sum_{k_1,\cdots,k_n=1,\cdots,n;\tau\in S_n} sgn(\tau)(a_{1,k_1}b_{k_1,\tau(1)})\cdots(a_{n,k_n}b_{k_n,\tau(n)}) \\
=& \sum_{\mu;\tau\in S_n} sgn(\tau)(a_{1,\mu(1)}b_{\mu(1),\tau(1)})\cdots(a_{n,\mu(n)}b_{\mu(n),\tau(n)}) \\
=& \underbrace{\sum_{\mu,\tau\in S_n} sgn(\tau)(a_{1,\mu(1)}b_{\mu(1),\tau(1)})\cdots(a_{n,\mu(n)}b_{\mu(n)\,tau(n)})}_{=det(A)det(B)} \\
&+ \sum_{\mu,\mu\notin S_n;\tau\in S_n} sgn(\tau)(a_{1,\mu(1)}b_{\mu(1),\tau(1)})\cdots(a_{n,\mu(n)}b_{\mu(n),\tau(n)})
\end{aligned}$$

上式第二行中，μ 跑遍所有 $\mathcal{S}(n)$ 到自身的映射，并且我们注意到 $\mathcal{S}(n)$ 到自身的映射通过如下方式一一对应于 n 元数组

$$\mu: i\mapsto k_i.$$

为了完成证明，我们仅需要证明上式最后一行求和为 0:

$$\sum_{\mu,\mu\notin S_n;\tau\in S_n} sgn(\tau)(a_{1,\mu(1)}b_{\mu(1),\tau(1)})\cdots(a_{n,\mu(n)}b_{\mu(n),\tau(n)}) = 0.$$

进一步我们只需我们证明，对固定的 μ，若 $\mu\notin S_n$ ，即 μ 不是一个双射时，

$$\sum_{\tau\in S_n} sgn(\tau)(a_{1\mu(1)}b_{\mu(1)\tau(1)})\cdots(a_{n\mu(n)}b_{\mu(n)\tau(n)}) = 0$$

我们可以再次利用 $\mu: i\mapsto k_i$ 将上式左侧改写为固定多重有序数组 $k_1,\cdots,k_n$ 的形式：

$$\sum_{\tau\in S_n} sgn(\tau)(a_{1k_1}b_{k_1\tau(1)})\cdots(a_{nk_n}b_{k_n\tau(n)})$$

由于 μ 不是一个双射，因此 $k_1,\cdots,k_n$ 中至少由一对不同下标的数字相同，由对称性不妨设 $k_1 = k_2$。

考虑由对换 $(1,2)$ 左乘给出的 S_n 到自身的双射，自然的给出了 S_n 上的一个等价关系。他将所有置换 S_n 划分为 $\frac{n!}{2}$ 个对，其中 σ,τ 组成一个对，当且仅当

$$\sigma = \tau(1,2)$$

在上式跑遍 S_n 的求和中，我们只需证明在以上定义的任意一对置换 σ,τ 对应的两项，其和为 0，即

$$sgn(\tau)(a_{1,k_1}b_{k_1,\sigma(1)})\cdots(a_{n,k_n}b_{k_n,\sigma(n)}) + sgn(\tau)(a_{1,k_1}b_{k_1,\tau(1)})\cdots(a_{n,k_n}b_{k_n,\tau(n)}) = 0.$$

注意到第一项当中，利用

$$(1,2)(1) = 2,\ (1,2)(2) = 1,\ sgn((1,2)) = -1,$$

得到第一项

$$\begin{aligned}
&sgn(\sigma)(a_{1,k_1}b_{k_1,\sigma(1)})\cdots(a_{n,k_n}b_{k_n,\sigma(n)})\\
=&sgn(\tau(12))(a_{1,k_1}b_{k_1,\tau(12)(1)})(a_{1,k_2}b_{k_12,\tau(12)(2)})\cdots(a_{n,k_n}b_{k_n,\tau(12)(n)})\\
=&-sgn(\tau)(a_{1,k_1}b_{k_1,\tau(2)})(a_{1,k_2}b_{k_2,\tau(1)})\cdots(a_{n,k_n}b_{k_n\tau(12)(n)}),
\end{aligned}$$

由于矩阵元素 a_{ij}, b_{ij} 是交换的，我们发现第一项与第二项除去前置符号相反外，乘积部分仅仅是排列了其中的项而最终相同，故两项的和相互抵消为零。

行列式理论的一个重要应用，是可以对线性相关或无关性给出具体的刻画。回忆若 A 可逆当且仅当存在 A^{-1} 使得

$$AA^{-1} = I_n,$$

因而

$$det(A)det(A^{-1}) = 1,$$

故推出

$$det(A) \neq 0.$$

命题 78. *$A \in M_n(k)$ 可逆当且仅当，则*

$$det(A) \neq 0.$$

我们已经证明了命题“仅当”部分。对另一部分的证明有两种处理办法，本节我们采取证明“仅当”命题的否命题的办法，即

命题 79. *若 $A \in M_n(k)$ 不可逆，则*

$$det(A) = 0.$$

证明上述命题的一个技巧，将 A 不可逆，转化为列向量的线性相关性。

命题 80. *若 A 的列向量线性相关，则*

$$det(A) = 0.$$

若 $v_1, \cdots, v_n$ 线性相关，不妨设 v_1 可以由 $v_2, \cdots, v_n$ 线性表出：

$$v_1 = c_2 v_2 + \cdots + c_n v_n.$$

由于第二类初等变换不改变行列式的值，因而将 A 的第 j 列，$j = 2, \cdots, n$ 分别乘以 c_i 加到第一列得到新矩阵第一列为零向量：

$$det(v_1, v_2, \cdots, v_n) = det(0, v_2, \cdots, v_n) = 0.$$

结合前几节的内容，我们得到方阵 A 可逆的几种不同形式的等价刻画

命题 81. *以下性质与方阵 A 可逆等价：*

1. *A 存在左逆。*

2. *A 存在右逆。*

3. *A 列满秩。*

4. *A 行满秩。*

5. *齐次方程组 $Ay = 0$ 仅存在平凡解 $y = 0$。*

6. 齐次方程组 $y^tA=0$ 仅存在平凡解 $y=0$。

7. $det(A)\neq 0$。

我们看到，行列式给出了线性相关性的一个**封闭公式 (closed formula)** 刻画，它和几何上的直观也是符合的，即一组向量构成空间的基底，当且仅当它们构成的平行多面体的“定向体积”为 0。

我们看到当 $det(A)$ 时，A^{-1} 是唯一存在的，但上面的论证办法仅仅从侧面论证了存在唯一性而没有具体给出 A^{-1} 的公式。这将在下一节进行讨论。

2.10.3 矩阵逆和 Cramer 法则

在上一节当中我们证明了方阵 A 可逆当且仅当 $det(A)\neq 0$。由此可以得知当 $det(A)\neq 0$ 时，A^{-1} 是存在且唯一的。为了给出 A^{-1} 的表达式，我们需要进一步引入一些概念。

定义 46. 给定一个 $n\times m$ 矩阵 A，以及两组整数 $1\leq i_1<\cdots<i_k\leq n$，$1\leq j_1<\cdots<j_l\leq n$。称将位于 (i_s,j_t), $s=1,\cdots,k$, $t=1,\cdots,l$ 位置元素取出并依次排布得到的 $k\times l$ 子矩阵

$$(a'_{s,t})_{k\times l},\ a'_{s,t}:=a_{i_s,j_t}$$

为 A 关于 $i_1,\cdots,i_k$ 和 $j_1,\cdots,j_l$ 的 **子式（*minor*）**，记作

$$A\binom{i_1,\cdots,i_k}{j_1,\cdots,j_l},$$

并称将位于第 $i_1,\cdots,i_k$ 行，或者 $j_1,\cdots,j_l\leq n$ 列元素划去得到的 $(n-k)\times(m-l)$ 矩阵称为相应的 **余子式 *(co-minor)***，记作

$$\hat{A}\binom{i_1,\cdots,i_k}{j_1,\cdots,j_l},$$

本节我们主要使用 A 和 A 的子式为方阵的情形。

我们有如下的 Laplace 展开定理：

定理 18. 设 A 为 n 阶方阵，对固定的 k 个数 $1\leq i_1<\cdots<i_k\leq n$，有

$$det(A)=\sum_{1\leq j_1<\cdots<j_k\leq n}(-1)^{\sum i_s+\sum j_t}det(A\binom{i_1,\cdots,i_k}{j_1,\cdots,j_k})det(\hat{A}\binom{i_1,\cdots,i_k}{j_1,\cdots,j_k})$$

这一定理可以通过行列式的定义直接予以证明。

引理 10.

$$\sum_j (-1)^{i+j} a_{ij} det(\hat{A}\binom{k}{j}) = \delta_{ik} det(A).$$

定义 47. 称 N 阶方阵

$$adj(A) = (a'_{i,j})_{n\times n},\ a'_{i,j} = (-1)^{i+j} det(A\binom{j}{i})$$

为 A 的**伴随矩阵**。

命题 82.

$$A \cdot adj(A) = det(A) I_n = adj(A) \cdot A.$$

因此若 $det(A) \neq 0$ 则 A^{-1} 存在且

$$A^{-1} = det(A)^{-1} adj(A).$$

回忆线性方程组 (2)：

$$Ax = b.$$

现考虑 A 为方阵的情况。将两侧同乘以 $adj(A)$ 得到

$$det(A)b = adj(A)b = (\hat{b}_1, \cdots, \hat{b}_n),$$

其中

$$\hat{b}_i = \sum_j b_j (-1)^{i+j} det(\hat{A}\binom{k}{j}) = det(\hat{B}_i),$$

而

$$\hat{B}_i = (v_1, \cdots, \underbrace{b}_{\text{第}i\text{个位置}}, \cdots, v_n).$$

是将 A 第 i 列替换为 b 的方阵。当 $det(A) \neq 0$ 时便得到 Cramer 法则：

命题-定义 3.

$$Ax = b$$

当 $det(A) \neq 0$ 时有唯一解

$$x_i = \frac{det(\hat{B}_i)}{det(A)}.$$

最后我们再指出（非齐次）线性方程组 (2) 解的存在性，与相应的齐次线性方程组 (10) 非平凡（非零）解的（非）存在性的一个重要区别。通过行列式，我们已经知道 (10) 非平凡（非零）非解存在性，由关于系数 a_{ij} 的行列式函

数是否等于零刻画，但是关于线性方程组 (2) 解，无论其存在性还是非存在性，都不可能通过“封闭等式”进行刻画，即不存在关于系数 a_{ij}, b_i 的多项式函数 $f_1, \cdots, f_m$ 使得 (2) 存在（或不存在）解，当且仅当

$$f_1(a,b) = \cdots = f_m(a,b) = 0$$

成立。在习题中我们会看到，非齐次线性方程组解（不）存在性的充要条件需，要由等式和不等式联立进行刻画。在讨论射影空间时，我们还会看到与齐次线性方程组 (10) 非平凡（非零）解的（非）存在性类似的规律。

我们看到，本节当中所介绍的有关行列式的理论与应用，看上去具有一定的巧合性，且行列式的定义本身也具有一定复杂性。后面我们在讨论张量的部分，会通过反对称张量空间，给出行列式及其推广的内蕴刻画。

2.11 习题

待补充。

3 线性变换的约化和标准型

已初步完成，待完善。

4 （半）双线性型的约化及其标准型

待完善。

5 内积空间

待完善。

6 张量积与张量代数

已初步完成，待完善。

7 线性几何与代数簇简介

已初步完成，待完善。

8 数值线性代数和矩阵计算简介

待完善。

参考文献

[W18] 汪芳庭，数学基础. 中国科技大学出版社，2018.

[G13] 耿素云等，离散数学. 清华大学出版社，2013.

[R19] K.H. Rosen， Discrete Mathematics and its appilcations，8-th ed; 中译本：离散数学及其应用（第八版）. 中译本：机械工业出版社,2019.

[L02a] 蓝以中，高等代数简明教程（上）（下）. 北京大学出版社, 2002.

[HK71] Hoffmann, Kenneth, and Ray Alden Kunze. Linear algebra. New Jersey: Prentice-Hall, 1971.

[H43] Halmos, Paul R. Finite-dimensional vector spaces, 1943; Courier Dover Publications reprint, 2017.

[B98] Bourbaki, Nicolas. Algebra I: chapters 1-3. Vol. 1. Springer Science Business Media, 1998.

[S06] Strang, Gilbert. Linear algebra and its applications. Belmont, CA: Thomson, Brooks/Cole, 2006.