高等代数笔记

副标题

Zhang Liang

2025 年 9 月 20 日

前言标题

前言内容

2025 年 9 月 20 日

目录

符号表

符号	含义
$(V, F, +, \cdot)$	域 F 上的线性空间 V
$\mathrm{span}(S)$	至多可数集 S 张成的线性子空间；不可数集 S 的全体有限子集张成向量的集合
$\mathrm{rank}(S)$	至多可数集 S 的秩
$\dim_F V$	线性空间 V 在域 F 上的维数
$V \cong W$	线性空间 V 和 W 同构
$V \oplus W$	线性空间 V 和 W 的直和
V/W	线性空间 V 关于子空间 W 的商空间
V^*	线性空间 V 的对偶空间
$V_1 \otimes V_2$	V_1 和 V_2 的张量积

第一章　线性空间

作为第一章，我们研究高等代数的核心研究对象：线性空间

正如之前研究的群、环、域等代数结构一样，我们如下构建线性空间的理论。

在第一节中，我们先提出线性空间的定义；第二至四节中，我们从线性空间中的基本运算——线性组合出发，逐渐寻找表示一个向量集乃至整个空间的最佳方式

从第五节开始，我们开始探索不同空间的相互联系。第五节中，我们将研究一种极为良好的关系：同构；第六节中，我们将研究如何将线性空间作唯一分解；第七节中，我们将研究如何从空间和它的一个子空间导出新的结构——商空间；

在第八节中，作为第一章的最后一节，我们将为第二章的核心研究对象——线性映射作铺垫，研究一类特殊的线性映射的线性空间性质

1.1 线性空间的定义

1.1.1 线性空间的定义

定义 1.1.1: 线性空间

设 F 是一个域，V 是一个集合，存在两个运算 $+:V\times V\to V$ 和 $\cdot:F\times V\to V$，分别称为加法和乘法，使得：

① $\exists\mathbf{0}\in V,\forall\alpha\in V,\mathbf{0}+\alpha=\alpha$

② $\forall\alpha\in V,\exists-\alpha\in V,\text{s.t. }\alpha+(-\alpha)=\mathbf{0}$

③ $\forall\alpha,\beta\in V,\alpha+\beta=\beta+\alpha$

④ $\forall\alpha,\beta,\gamma\in V,(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$

⑤ $\forall\alpha\in V,1\cdot\alpha=\alpha$

⑥ $\forall k,l\in F,\alpha\in V,(k\cdot l)\cdot\alpha=k\cdot(l\cdot\alpha)$

⑦ $\forall k,l\in F,\alpha\in V,(k+l)\cdot\alpha=k\cdot\alpha+l\cdot\alpha$

⑧ $\forall k\in F,\alpha\in V,k\cdot(\beta+\gamma)=k\cdot\beta+k\cdot\gamma$

那么我们称 $(V,F,+,\cdot)$ 是一个线性空间。

此时我们也说 V 是一个 F 上的线性空间（或向量空间）

1.1.2 线性子空间

定义 1.1.2: 线性子空间

设 V 是一个线性空间，集合 $W\subseteq V$，如果 W 在 V 的运算构成一个线性空间，那么我们称 W 是 V 的一个线性子空间

1.1.3 线性空间的性质

1.

命题 1.1.1. $\mathbf{0}$ 是唯一的

证明: 不妨假设命题不成立，$\mathbf{0}_1,\mathbf{0}_2$ 均是零元，并且 $\mathbf{0}_1\neq\mathbf{0}_2$

我们注意到：$\mathbf{0}_1=\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_2$，与假设矛盾，于是命题得证 □

2.

命题 1.1.2. $\forall\alpha\in V$，$-\alpha$ 是唯一的

证明: 不妨假设命题不成立，α 有两个逆元 β_1, β_2，并且 $\beta_1 \neq \beta_2$

我们注意到：$\beta_1 = \mathbf{0} + \beta_1 = (\beta_2 + \alpha) + \beta_1 = \beta_2 + (\alpha + \beta_1) = \beta_2 + \mathbf{0} = \beta_2$，与假设矛盾，于是命题得证 □

3.

命题 1.1.3. $\forall \alpha \in V, 0 \cdot \alpha = \mathbf{0}$

证明: $0 \cdot \alpha = (0+0) \cdot \alpha = 0 \cdot \alpha + 0 \cdot \alpha$

$\Rightarrow 0 \cdot \alpha + (-0 \cdot \alpha) = 0 \cdot \alpha + 0 \cdot \alpha + (-0 \cdot \alpha)$

$\Rightarrow \mathbf{0} = 0 \cdot \alpha$ □

4.

命题 1.1.4. $\forall k \in F, k \cdot \mathbf{0} = \mathbf{0}$

证明: $k \cdot \mathbf{0} = k \cdot (0 \cdot \alpha) = (k \cdot 0) \cdot \alpha = 0 \cdot \alpha = \mathbf{0}$ □

5.

命题 1.1.5. $\forall \alpha \in V, (-1) \cdot \alpha = -\alpha$

证明: $\mathbf{0} = 0 \cdot \alpha = (1 + (-1)) \cdot \alpha = 1 \cdot \alpha + (-1) \cdot \alpha$

$\Rightarrow \mathbf{0} + (-\alpha) = (-\alpha) + \alpha + (-1) \cdot \alpha \Rightarrow (-1) \cdot \alpha = -\alpha$ □

6. 事实上，验证一个子集是否是线性子空间，只需要验证封闭性即可，其他的条件都是不必要的。

命题 1.1.6. *$W \subseteq V$ 是 V 的线性子空间，当且仅当：*

$\forall k \in F, \alpha, \beta \in V, \alpha + \beta \in V, k \cdot \alpha \in V$

证明: 充分性是显然的，对于必要性，我们依次验证：

首先，$0 \in F$，因此 $0 \cdot \alpha = \mathbf{0} \in W$

其次，因为 $-1 \in F$，所以 $(-1) \cdot \alpha = -\alpha \in W$

剩余的六条运算律，因为 W 上的运算即是 V 上的运算在 W 上的限制，所以显然成立。

那么命题得证。 □

1.2 线性组合

本节中，我们着手研究一系列向量以相加，相乘的形式组合而成的新向量，称之为线性组合。

其最一般的形式是 $\sum\limits_{i=1}^{n}\alpha_i$，但是我们希望能够进一步拓展，考虑无穷多个向量的线性组合。

因此，我们需要先定义向量列的极限，其最好方式就是先定义出线性空间中的度量。

1.2.1 度量线性空间

定义 1.2.1: 度量线性空间、线性组合的定义

设 V 是一个域 F 上的线性空间，映射 $d: V\times V\to F$ 如果满足：

① $\forall\alpha,\beta\in V, d(\alpha,\beta)=0\Leftrightarrow\alpha=\beta$

② $\forall\alpha,\beta\in V, d(\alpha,\beta)=d(\beta,\alpha)$

③ $\forall\alpha,\beta,\gamma\in V, d(\alpha,\beta)+d(\beta,\gamma)\geqslant d(\alpha,\gamma)$

那么我们称 (V,d) 是一个度量线性空间（或 Banach 空间），d 称为 V 中的度量

这样，我们就可以定义出线性组合

定义 1.2.2: 线性组合

设 V 是一个域 F 上的线性空间，$\{\alpha_i\in V\}$ 是一个至多可数的向量组

那么如果向量 $\beta\in V$ 满足：$\forall\varepsilon>0,\exists,n\in\mathbb{N}_+,\text{s.t. } d(\sum\limits_{i=1}^{n}k_i\alpha_i,\beta)<\varepsilon, k_i\in F$

那么我们称 β 是一个由向量组 $\{\alpha_i\in V\}$ 张成的的线性组合，并称 β 可以由 $\{\alpha_i\in V\}$ 线性表出

如果 $\{\alpha_i\in V\}$ 可数，记作：$\beta=\sum\limits_{i=1}^{\infty}\alpha_i$，否则记作 $\beta=\sum\limits_{i=1}^{n}\alpha_i$，其中 n 是向量组的元素数

由向量组 S 张成的全体线性组合形成的子空间记作：$span_F(S)$（或 $L(S)$，$\langle S\rangle$）。特别地，如果 S 不可数，我们记 $span_F(S)$ 为由 S 的有限子集张成的全体向量的集合

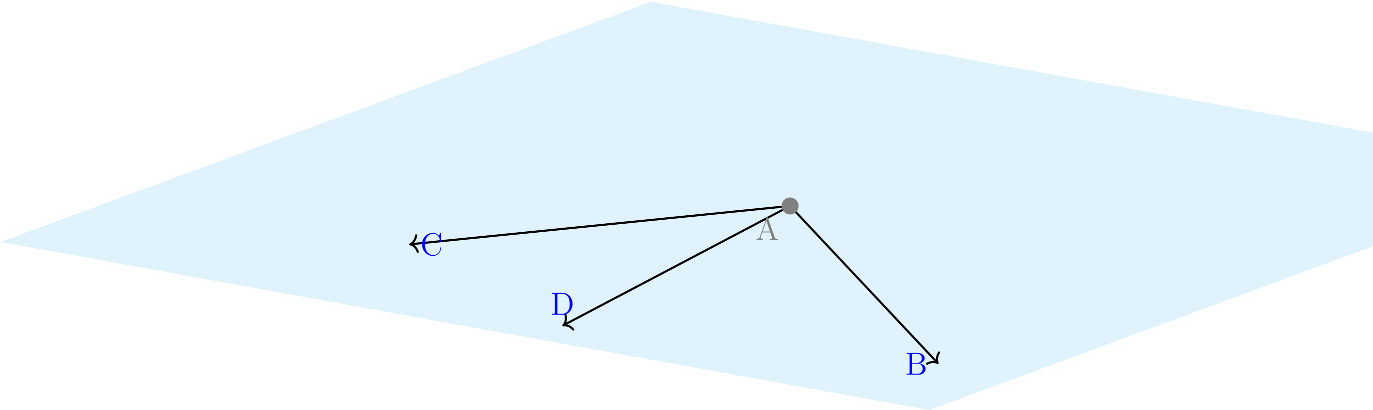
我们不考虑不可数的向量组张成的线性组合，因为不可数的向量的求和的定义通常是不良好的。

此后除非特别说明，我们默认我们所说的向量组指的是至多可数的向量组。

1.2.2 向量组的线性相关性和线性无关性

我们观察到一个现象：向量组张成线性子空间时，并非每一个向量都有贡献，有些向量是不必要的。就像下图中，尽管三个向量张成了整个平面，但是其实任取其中之二，就足以

张成整个平面。



我们会自然地认为：如果每一个向量都和其他向量的“方向”不一样，就不会有不必要的向量，或者说，不应该可以在赋予非零系数时，张成零向量

此时，向量组被称为是“线性无关的”。

定义 1.2.3: 线性无关的向量组

向量组 $\{\alpha_x \in V|x \in A\}$ 如果满足：

$\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0} \Rightarrow \forall x \in A, k_x = 0$

那么我们称 $\{\alpha_x \in V|x \in A\}$ 是一个线性无关的向量组

如果一个向量组不是线性无关的，那么我们称它是线性相关的，我们可以写出以下定义

定义 1.2.4: 线性相关的向量组

向量组 $\{\alpha_x \in V|x \in A\}$ 如果满足：

存在不全为零的一组系数 $k_x, x \in A$，使得 $\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0}$

那么我们称 $\{\alpha_x \in V|x \in A\}$ 是一个线性相关关的向量组

1.2.3 线性相关和线性无关的性质

1. 单个向量当且仅当它是零向量是线性相关

   命题 1.2.1. α 线性相关 $\Leftrightarrow \alpha = \mathbf{0}$

   证明： 先证充分性，如果 α 线性相关，那么 $\exists k \in F, k \neq 0, k\alpha = \mathbf{0} \Rightarrow \alpha = \mathbf{0}$

   必要性是显然的 □

2. 如果一个部分组线性相关，整个向量组也是线性相关的

命题 1.2.2. 如果 $\{\alpha_x \in V|x \in A\}$ 是一个向量组,它的一个部分组 $\{\alpha_x \in V|x \in B\}, B \subseteq A$ 线性相关，那么 $\{\alpha_x \in V|x \in A\}$ 也是线性相关的

证明： 因为 $\{\alpha_x \in V|x \in B\}$ 线性相关，所以存在不全为零的系数 $k_x, x \in B$，使得 $\sum\limits_{x\in B} k_x\alpha_x = \mathbf{0}$

我们可以将 k_x 扩展到 A 上，令 $k_x = 0, x \in A - B$，那么 $\sum\limits_{x\in A} k_x\alpha_x = \sum\limits_{x\in B} k_x\alpha_x = \mathbf{0}$，所以 $\{\alpha_x \in V|x \in A\}$ 线性相关 □

反过来，我们可以推论出，线性无关向量组的任意部分组也是线性无关的

推论 1.2.1: 线性无关向量组的任意部分组都线性无关

若向量组 $\{\alpha_x \in V|x \in A\}$ 线性无关，那么它的任意一个部分组 $\{\alpha_x \in V|x \in B\}, B \subseteq A$ 也是线性无关的

证明： 如果并非如此，那么至少有一个部分组线性相关，那么按照命题，整个向量组线性相关，这与假设矛盾，所以命题得证 □

因此，一个显然的事实是，一个包含零向量的向量组是一定线性相关的

命题 1.2.3. 如果 $\mathbf{0} \in \{\alpha_x \in V|x \in A\}$，那么 $\{\alpha_x \in V|x \in A\}$ 线性相关

证明： 显然 □

3. 向量组线性相关的充要条件是：存在一个向量可以由其他向量线性表出

命题 1.2.4. 向量组线性相关 $\Leftrightarrow$ 向量组中存在一个向量可以由其他向量线性表出

证明： 先证充分性。如果 $\{\alpha_x \in V|x \in A\}$ 线性相关,那么存在不全为零的系数 $k_x, x \in A$，使得 $\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0}$

取其中任意一个不为零的系数 $k_a \neq 0$，那么有 $\alpha_a = -\frac{1}{k_a}\sum\limits_{x\in A-\{a\}} k_x\alpha_x$，充分性得证

接下来证明必要性，如果 $\alpha_a = \sum\limits_{x\in A-\{a\}} k_x\alpha_x$，那么我们只需令 $k_a = -1$

那么就有 $\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0}$ □

4. 能唯一地表出一个向量的向量组线性无关

命题 1.2.5. $\beta \in V$ 可以被 $\{\alpha_x \in V|x \in A\}$ 唯一地线性表出 $\Leftrightarrow \{\alpha_x \in V|x \in A\}$ 线性无关

证明: 先证充分性。

这其实是显然的,因为如果 $\{\alpha_x \in V | x \in A\}$ 线性相关,那么有一个线性组合 $\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0}$ 并且 k_x 不全为零

假设 $\beta = \sum\limits_{x\in A} l_x\alpha_x$,那么也有 $\sum\limits_{x\in A}(k_x + l_x)\alpha_x$,这与假设矛盾

再证必要性。

不妨假设 $\beta = \sum\limits_{x\in A} l_x\alpha_x = \sum\limits_{x\in A} k_x\alpha_x \exists x \in A, k_x \neq l_x$

那么 $\sum\limits_{x\in A}(l_x - k_x)\alpha_x = \mathbf{0}$,又因为 $\{\alpha_x \in V | x \in A\}$ 线性无关,所以 $l_x - k_x = 0$。因此表出方式是唯一的,命题得证 □

5. 向线性无关向量组中加入一个向量,如果变得线性相关,那么这个向量可以由其他向量线性表出

命题 1.2.6. 如果 $\{\alpha_x \in V | x \in A\}$ 线性无关,$\{\beta, \alpha_x \in V | x \in A\}$ 线性相关

那么 β 可由 $\{\alpha_x \in V | x \in A\}$ 线性表出

证明: 因为 $\{\beta, \alpha_x \in V | x \in A\}$ 线性相关,所以存在不全为零的系数 $k_x, l \in F, x \in A$,使得 $\sum\limits_{x\in A} k_x\alpha_x + l\beta = \mathbf{0}$

我们只需证 $l \neq 0$。我们不妨假设 $l = 0$

那么 $\sum\limits_{x\in A} k_x\alpha_x = \mathbf{0}$。但是 $\{\alpha_x \in V | x \in A\}$ 线性无关,所以 $\forall x \in A, k_x = 0$,这与系数不全为零的假设矛盾

那么有 $l \neq 0$,于是 $\beta = -\frac{1}{l}\sum\limits_{x\in A} k_x\alpha_x$,命题得证 □

1.3 极大线性无关向量组、向量组的秩

前面我们讨论了线性无关向量组,它可以说是“所有元素都有贡献”的,接下来我们讨论,如果把一个向量集剔除至恰好包含所有“有贡献”的向量,这种向量组有什么性质

我们的一个想法是,如果剔除至向量组本身依旧线性无关,但是加入任意向量就变得线性相关,那么这个向量组就是我们想要的“极大线性无关向量组”

1.3.1 极大线性无关向量组

定义 1.3.1: 极大线性无关向量组

设 $\{\alpha_x \in V | x \in A\}$ 是一个向量组

如果它的一个部分组 $\{\alpha_x \in V | x \in B\}, B \subseteq A$ 线性无关，并且 $\forall \beta \in V, \{\beta, \alpha_x | x \in B\}$ 线性相关

那么我们称 $\{\alpha_x \in V | x \in B\}$ 是 $\{\alpha_x \in V | x \in A\}$ 的一个极大线性无关向量组

我们直观地认为，每一个极大线性无关向量组应该具备相同的元素数（或势），就像在三维空间中，我们可以选取不同的坐标系，但是每个坐标系都恰好三个坐标轴

1.3.2 向量组的等价

我们在向量组上定义以下等价关系：

定义 1.3.2: 向量组的线性表出

如果向量组 $\{\alpha_x \in V | x \in A\}$ 中的任意向量可由 $\{\beta_x \in V | x \in B\}$ 线性表出

那么我们称 $\{\alpha_x \in V | x \in A\}$ 可由 $\{\beta_x \in V | x \in B\}$ 线性表出

定义 1.3.3: 向量组的等价

如果向量组 $\{\alpha_x \in V | x \in A\}$ 可由 $\{\beta_x \in V | x \in B\}$ 线性表出，并且 $\{\beta_x \in V | x \in B\}$ 可由 $\{\alpha_x \in V | x \in A\}$ 线性表出

那么我们称 $\{\alpha_x \in V | x \in A\}$ 和 $\{\beta_x \in V | x \in B\}$ 是等价的，记作：$\{\alpha_x \in V | x \in A\} \cong \{\beta_x \in V | x \in B\}$

我们首先验证它的确是一个等价关系

命题 1.3.1. 向量组的等价是一个等价关系

证明: 自反性是显然的，因为向量组中的任意向量 α_x 可以由它自己线性表出；

对称性由定义是显然的；

最后证明传递性，假设 $\{\alpha_x \in V | x \in A\} \cong \{\beta_x \in V | x \in B\}, \{\beta_x \in V | x \in B\} \cong \{\gamma_x \in V | x \in C\}$

设 $\alpha_x = \sum\limits_{y \in B} k_{xy} \beta_y, \beta_y = \sum\limits_{z \in C} l_{yz} \gamma_z$

那么 $\alpha_x = \sum\limits_{y\in B} k_{xy}\left(\sum\limits_{z\in C} l_{yz}\gamma_z\right) = \sum\limits_{y\in B,z\in C} k_{xy}l_{yz}\gamma_z$

因此 $\{\alpha_x \in V|x \in A\}$ 可由 $\{\alpha_x \in V|x \in A\}$ 线性表出，同理可证 $\{\gamma_x \in V|x \in C\}$ 可由 $\{\alpha_x \in V|x \in A\}$ 线性表出

所以 $\{\alpha_x \in V|x \in A\} \cong \{\gamma_x \in V|x \in C\}$，传递性得证。于是命题得证 □

接下来讨论一些向量组等价的性质：

1. 任何向量组和它的极大线性无关向量组等价

命题 1.3.2. 任何向量组和它的极大线性无关向量组等价

证明： 极大线性无关向量组显然可以由原向量组线性表出；

而按照极大线性无关向量组的定义，向其中加入任意一个向量，向量组会变得线性相关，于是，原向量组中的任意一个向量都可以由极大线性无关向量组线性表出。命题得证 □

2. 向量组的两个极大线性无关向量组等价

命题 1.3.3. 向量组的两个极大线性无关向量组等价

证明： 只需利用上面的性质和等价的传递性即可完成证明。 □

3. 一个向量组如果可以由一个势比它小的向量组线性表出，那么它线性相关 在给出命题及其证明前，我们先证明一个显然的引理

引理 1.3.1

向量组 $\{\alpha_1, \cdots, \alpha_n\}$ 如果线性相关，

那么一定 $\exists k, \text{s.t. } \alpha_k$ 可由 $\alpha_1, \cdots, \alpha_{k-1}$ 线性表出，并且 $\text{span}(\alpha_1, \cdots, \alpha_{i-1}, \alpha_{i+1}, \cdots, \alpha_n) = \text{span}(\alpha_1, \cdots, \alpha_n)$

接下来给出命题

证明： 因为 $\{\alpha_1, \cdots, \alpha_n\}$ 线性相关，所以一定 $\exists k_1, \cdots, k_n, k_1\alpha_1 + \cdots + k_n\alpha_n = \mathbf{0}$

只需选取满足 $k_i \neq 0$ 的最大的下标 p，那么有 $\alpha_1 + \cdots + k_p\alpha_p = \mathbf{0}$

于是 $\alpha_p = -\frac{1}{k_p}\sum\limits_{i=1}^{p-1}\alpha_i$

此时因为 α_p 可由前面的向量线性表出，因此去除其不会对生成的子空间有任何影响，命题得证。 □

命题 1.3.4. 如果向量组 $\{\alpha_x \in V|x \in A\}$ 可由 $\{\beta_x \in V|x \in B\}$ 线性表出，并且 $Card\ A > Card\ B$

那么 $\{\alpha_x \in V|x \in A\}$ 线性相关

证明: 先考虑两者都是有限向量组的情况

我们可以证明：如果 $\{\alpha_1, \cdots, \alpha_m\}$ 可由 $\{\beta_1, \cdots, \beta_n\}$ 线性表出，并且 $m > n$，那么 $\{\alpha_1, \cdots, \alpha_m\}$ 线性相关

记 $W = \mathrm{span}(\beta_1, \ldots, \beta_n)$。

我们采取以下方法证明：用 α_i 逐步替换 β_j，但是保持生成的子空间不变。最后，我们一定会发现，β_j 已经被全部替换，但是此时依旧有 α_i 未曾被加入向量组过。这就说明了 $\{\alpha_1, \cdots, \alpha_m\}$ 中有向量可以被其它向量线性表出，即 $\{\alpha_1, \cdots, \alpha_m\}$ 线性相关

也就是说，我们希望构造一系列向量组 $B_0, B_1, \ldots, B_n$，使得对于 $j = 0, 1, \ldots, n$，向量组 B_j 满足：

(a) B_j 的长度为 n。

(b) $\mathrm{span}(B_j) = W$。

(c) B_j 包含向量 $\alpha_1, \ldots, \alpha_j$ 以及从原始向量组 $\{\beta_1, \ldots, \beta_n\}$ 中剩下的 $n - j$ 个向量。

我们首先取 $B_0 = \{\beta_1, \ldots, \beta_n\}$

我们如此由 B_{j-1} 迭代出向量组 B_j:

考虑向量组 $\alpha_j \cup B_{j-1}$（并且不失一般性地将 α_j 置于向量组的首位）。由于 $\alpha_j \in W$，所以一定有 $\mathrm{span}(\alpha_j \cup B_{j-1}) = W$

由于 $\mathrm{span}(B_{j-1}) = W$，向量 α_j 在 B_{j-1} 生成的子空间中，因此向量组 $\alpha_j \cup B_{j-1}$ 是线性相关的。

根据前面的引理，线性相关的向量组 $\alpha_j \cup B_{j-1}$ 中存在一个向量，它是其前面向量的线性组合。设这个向量为 w。

如果 $w = \alpha_j$，则 α_j 是空向量组的线性组合，意味着 $\alpha_j = \mathbf{0}$，于是该向量组是线性相关的，证明结束。

接下来考虑 w 是 B_{j-1} 中的一个向量。如果 $w \in \{\alpha_1, \cdots, \alpha_{j-1}\}$，那么它可以被 $\{\alpha_1, \cdots, \alpha_m\}$ 中的向量线性表出，那么向量组线性相关，证明结束。

因此只需考虑这个向量 w 是 B_{j-1} 中那些来自 $\{\beta_1, \ldots, \beta_n\}$ 的向量

由于 w 可以由向量组中前面的向量线性表出，我们可以从 $\alpha_j \cup B_{j-1}$ 中移除 w，得到新的向量组 $B_j = (\alpha_j \cup B_{j-1}) - w$。这个新的向量组 B_j 仍然生成子空间 W

这样，我们就完成了向量组序列 $B_0, B_1, \ldots, B_n$ 的构造。

此时，我们只需注意到，$\text{span}(B_n) = \text{span}(\alpha_1, \cdots, \alpha_n) = W, \alpha_{n+1} \in W$。于是 α_{n+1} 可由 $\alpha_1, \cdots, \alpha_n$ 线性表出，于是 $\{\alpha_1, \cdots, \alpha_m\}$ 线性相关，证明完毕。

最后，我们考虑 $\text{Card}A = \mathbb{N}, \text{Card}B < \mathbb{N}$ 的情况。此时，一定能找到 $\{\alpha_x \in V | x \in A\}$ 的一个有限部分组线性相关，于是整个向量组就线性相关了。

于是命题得证。 □

4. 向量组的两个极大线性无关向量组等势

命题 1.3.5. 向量组的两个极大线性无关向量组等势

证明: 设有两个极大线性无关向量组 $\{\alpha_x \in V | x \in A\}, \{\beta_x \in V | x \in B\}$

根据前面的性质,他们相互等价,因此可以相互线性表出。所以一定有 Card $A \leqslant$ Card B, Card BCar

那么一定有 Card $A =$ Card B，所以命题得证 □

1.3.3 向量组的秩

在完成了前面的准备工作后，我们可以定义向量组的秩了

定义 1.3.4: 向量组的秩

向量组 $G = \{\alpha_x \in V | x \in A\}$ 的极大线性无关向量组的势称为向量组的秩，记作：$rank(G)$

我们前面的命题已经保证了，无论如何选取极大线性无关组，它的势都相同，因而秩也是相同的。

接下来讨论它的性质

1.

命题 1.3.6. 向量组 $\{\alpha_1, \cdots, \alpha_n\}$ 线性无关 $\Leftrightarrow rank(\alpha_1, \cdots, \alpha_n) = n$

证明: 先证充分性，如果向量组线性无关，那么它自身即是自己的一个极大线性无关向量组，所以 $rank(\alpha_1, \cdots, \alpha_n) = n$

再证必要性。如果 $rank(\alpha_1, \cdots, \alpha_n) = n$，那么它仅有一个极大线性无关向量组，即向量组本身，所以它线性无关，命题得证 □

值得注意的是，这个性质只对有限向量组成立，因为对于可数的向量组，可以存在一个可数的极大线性无关组，但是他的秩依旧和整个向量组一致。

2.

命题 1.3.7. 如果向量组 $A = \{\alpha_x \in V | x \in A\}$ 可由 $B = \{\beta_x \in V | x \in B\}$ 线性表出

那么 $rank(A) \leqslant rank(B)$

证明: 选取两者的任意一个极大线性无关向量组 $C \subseteq A, D \subseteq B$

因为 $A \cong C, B \cong D$，所以 C 可由 D 线性表出

由 C 的线性无关性可知，$rank(A) = \text{Card } C \leqslant \text{Card } D = rank(B)$ □

1.4 基、线性空间的维数

上一节中，我们提出了极大线性无关向量组的概念。

从线性子空间的角度看，极大线性无关组是张成子空间的最小的向量组；

本节中，我们尝试进一步地，让极大线性无关组张成整个线性空间。这样的极大线性无关组的存在性并不是平凡的，因为我们甚至不知道是否存在一个向量组可以张成整个空间

1.4.1 基的定义

我们定义两种不同的基，后续我们会看到为何我们需要两种基

定义 1.4.1: Harmel 基

设 V 是一个线性空间，它的一个子集 $B \subseteq V$ 如果满足:

1. B 的任意有限子集线性无关
2. V 中的任意向量可由 B 的一个有限子集线性表出

那么我们称 B 是 V 的一个 Harmel 基

我们特别规定: $\{\mathbf{0}\}$ 的基是 $\emptyset$

定义 1.4.2: Schauder 基

设 V 是一个线性空间，它的一个子集 $B \subseteq V$ 如果满足：

1. B 线性无关
2. V 中的任意向量可由 B 线性表出

那么我们称 B 是 V 的一个 Schauder 基

我们特别规定：$\{\mathbf{0}\}$ 的基是 $\emptyset$

可以看出来，这两种基的根本区别是：Schauder 基允许无限个向量来逼近一个向量，而 Harmel 基要求向量必须由有限个向量表出

我们接下来将看到这个区别的重要影响：即两种基的存在性是不同的

1.4.2 Harmel 基的存在性

我们将证明：Harmel 基总是存在的

定理 1.4.1: Harmel 基的存在性

设 V 是一个线性空间，那么 V 中一定存在一个 Harmel 基

证明： 记 H 为 V 中所有满足全部有限子集都线性无关的子集的集合。

我们定义 H 上的一个序关系：$A, B \in H, A \leqslant B \Leftrightarrow A \subseteq B$，显然它的确是一个序，因而 H 成为一个偏序集

接下来证明，对于任意一个 H 中的链 C，$\bigcup\limits_{P \in C} P$ 正是 C 的一个上界

先证明 $\bigcup\limits_{P \in C} P$ 是 H 中的一个元素。

取 $\bigcup\limits_{P \in C} P$ 的一个有限子集 $\alpha_1, \cdots, \alpha_s$

事实上，$\alpha_1, \cdots, \alpha_s$ 一定属于某个 $K \in C$（因为链是按照相互包含的关系形成的），因此 $\alpha_1, \cdots, \alpha_s$ 线性无关。

而 $\forall K \in C, K \leqslant \bigcup_{P \in C} P$ 是显然的。于是 H 中的任意一个链有上界

那么，依据 Zorn 引理，H 中一定存在一个极大元素 B。我们接下来证明，B 是 V 的一个 Harmel 基

首先，由 B 的构造方式，它的任意有限子集线性无关是显然的

我们证明任意一个向量 $\alpha \in V$ 都可以被 B 的有限子集线性表出

如果不然，那么任意的 $K \cup \{\alpha\} \subseteq B \cup \{\alpha\}$，$K \cup \{\alpha\}$ 都是线性无关的（利用命题 1.2.6）所以 $B \cup \{\alpha\} \in H$，于是，$B \leqslant B \cup \{\alpha\}$，这与 B 是 H 的极大元素矛盾。

于是命题得证。 □

在这个证明中，我们使用了 Zorn 引理，如果我们不承认选择公理，Harmel 基的存在性就不再成立。

事实上，线性空间的 Schauder 基是不一定存在的，但是它的反例非常复杂，此处我们简单讨论一下为何我们不能按照证明 Harmel 基存在性的方式来证明 Schauder 基的存在性

前面的证明中，有一个容易被忽视的地方：为何任意一个链都有上界。在 Harmel 基的定义下，这是可以成立的，因为 Harmel 基要求的是有限子集，从而可以被链上的一个元素包含

但是当考虑 Schauder 基，它要求全体向量一起作用，此时，链的并就不一定有这样的性质了，比如考虑以下例子：

例 1.4.1. 考虑平方可和空间 $l^2 = \{(a_1, \cdots, a_n, \cdots) | \sum\limits_{i=1}^{\infty} |a_i|^2 < +\infty\}$

并定义其中的度量为：$d\left((a_1, \cdots, a_n, \cdots), (b_1, \cdots, b_n, \cdots)\right) = \sqrt{\sum\limits_{i=1}^{\infty} (a_i - b_i)^2}$

取它上面的一个向量组 $\{e_i\}, e_n = (0, \cdots, 1, \cdots)$，其中 1 处于第 n 个位置

我们构造以下链 $\{S_k\}$：

$S_1 = \{e_1\}, S_{n+1} = S_n \cup \{e_{n+1}\}$

证明：$\forall k, S_k$ 线性无关，但是 $\bigcup_{i=1}^{\infty} S_i$ 线性相关

证明： 首先，$\forall k, S_k$ 线性无关。这是显然的，因为其实 S_k 全部是有限集

设 $c_n = \begin{cases} \frac{1}{2k+1}, n = 2k+1, k \in \mathbb{N} \\ -\frac{1}{2k-1}, n = 2k, k \in \mathbb{N} \end{cases}$

注意到：$\sum\limits_{i=1}^{\infty} c_i e_i = \mathbf{0}$

因为：$\forall \varepsilon > 0, \exists n \in N_+, \text{s.t.}$

$d\left(\sum\limits_{i=1}^{n} c_i e_i, \mathbf{0}\right) = \sum\limits_{i=1}^{\infty} c_n < \varepsilon$

此时系数显然不全为零，所以 $\bigcup_{i=1}^{\infty} S_i$ 线性相关 □

所以，论证至此就失效了，我们并不能同样地证明 Schauder 基的存在性。

此后除非特别说明，我们默认我们所说的基指的是 Hamel 基

1.4.3 线性空间的维数

我们常常将基的势称为维数。自然地，因为 Schauder 基并非一直存在，我们一般选取 Hamel 基来定义维数。

但是，目前还有一个问题：我们如何保证选取不同基时，维数依旧不变？也就是说，为何两个基一定等势？

对于有限基，这个问题是容易解决的，但是对于无限基，特别是不可数基，这个问题非常难以解决。接下来，我们将花费大量篇幅论证这一点。

引理 1.4.2: Steinitz 替换引理

设 V 是一个线性空间，$S,T \subseteq V$，并且 S 的任意有限子集线性无关，$V = \mathrm{span}(T)$

那么我们断言：$|S| \leqslant |T|$

证明: 我们希望构造一个映射 $S \ni S \mapsto C_s \subseteq T$，但是此时我们发现我们无从下手。

我们退而求其次，构造 S 的子集和 T 的子集之间的映射，同时用一个偏序集逼近 T

为了保证偏序集的每一个链都有上界（后面我们会发现为何我们采取这样的奇怪构造），我们如下构造

设 $\Omega := \{\,(C,f) \mid C \subset T,\ f : C \to S$ 单射且$(S - f(C)) \cup C$ 线性无关$\,\}$.

并在它上面定义序关系：$(C_1, f_1) \preceq (C_2, f_2) \iff C_1 \subset C_2$ 且$f_2|_{C_1} = f_1$

我们先验证 Ω 中的每个链的确有上界

设 $\mathscr{C} = \{(C_\alpha, f_\alpha)\}_{\alpha\in\Lambda}$ 是 Ω 中的全序链

设 $C := \bigcup_{\alpha\in\Lambda} C_\alpha, f := \bigcup_{\alpha\in\Lambda} f_\alpha$

此处映射的并指的是将全部元素的映射关系一同合并（由 Ω 的定义，此次操作的定义显然是良好的）

那么，$f : C \to S$ 的单射性是显然的；

而因为 $\forall\alpha, (S - f(C)) \cup C_\alpha \subseteq (S - f(C_\alpha) \cup C_\alpha)$ 线性无关，因此 $(S - f(C)) \cup C$ 线性无关。

于是 (C, f) 是链的上界。

故 Ω 满足 Zorn 引理的条件，我们记其极大元为 (C, f)。

我们接下来证明 $S \subset f(C) \cup C$。如果这个成立，我们可以立即用集合的势的性质推出结论。

若不然，那一定可以取 $s \in S - (f(C) \cup C)$

记 $S' := (S - f(C)) \cup C$。因为 (C, f) 是极大元，因此 S' 线性无关，所以 $S' - \{s\}$ 也是线性无关的。

我们说，一定存在一个 $t \in T - C, t \notin \operatorname{span}(S' - \{s\})$

如果不是这样，即 $T - C \subset \operatorname{span}(S - \{s\})$

那么 $s \in \operatorname{span}(T) = \operatorname{span}(T - C) + \operatorname{span}(C) \subset \operatorname{span}(S' - \{s\})$,

那么 s 可由 $S' - s$ 线性表出，那么 S' 就是线性相关的，与 S' 的线性无关性矛盾。

因此可取 $t \in T - C, t \notin \operatorname{span}(S' - \{s\})$

此时，我们定义：$g : C \cup \{t\} \longrightarrow S, \quad g|_C = f,\ g(t) = s$

而 $(S - g(C \cup \{t\})) \cup (C \cup \{t\}) = (S' - \{s\}) \cup \{t\}$ 仍线性无关（因为 $S' - s$ 线性无关且 t 不能被 $S' - s$ 线性表出），

故 $(C \cup \{t\}, g) \in \Omega$，与 (C, f) 极大矛盾。于是 $S \subset f(C) \cup C$

至此，我们可以推出最终的结论了，由 f 单射可知；

$|S| \leqslant |f(C)| + |C| = 2|C| \leqslant 2|T|$. □

进而我们知道，线性空间的两个基是等势的。

引理 1.4.3: 线性空间的两个基等势

设 V 是一个线性空间，B_1, B_2 是 V 的两个基。

那么：$|B_1| = |B_2|$

证明： 这是显然的，因为 B_1, B_2 均满足有限子集线性无关和张成 V

所以一定有 $|B_1| \leqslant |B_2|, |B_2| \leqslant |B_1|$，于是有 $|B_1| = |B_2|$，命题得证 □

证明了这个引理后，我们就可以给出维数的定义了

定义 1.4.3: 线性空间的维数

设 V 是一个 F 上的线性空间，B 是 V 的一个 Hamel 基

我们定义：

$$dim_F V = \text{Card } B \tag{1.1}$$

称为 V 的维数，简记作 $dimV$

如果 $dim_F V < +\infty$，我们称 V 是一个有限维线性空间，反之称为无限维线性空间

可以看出来，线性空间的维数并非只有有限和无限两类，比如以下例子分别给出了 $\aleph_0$ 和

$\aleph_1$ 维的线性空间

例 1.4.2. $\dim_K K[x] = \aleph_0$

证明： 显然，$\{x^n | n \in \mathbb{N}\}$ 是 $K[x]$ 的一个基

因此，$\dim_K K[x] = |\{x^n | n \in \mathbb{N}\}| = \aleph_0$ □

例 1.4.3. $\dim_{\mathbb{R}} \{f : \mathbb{R} \to \mathbb{R} | \mathrm{Card} supp(f) < \aleph_0\} = \aleph_1$

证明： 我们指出，$\left\{ f_x | x \in \mathbb{R}, f_x(y) = \begin{cases} 1, y = x \\ 0, y \neq x \end{cases} \right\}$ 是 $\{f : \mathbb{R} \to \mathbb{R} | \,\mathrm{Card}\, supp(f) < \aleph_0\}$ 的一个基

因为：首先这个集合的任意有限子集线性无关，其次空间的任意向量可以被这个集合的一个有限子集线性表出（因为空间内的函数的支撑集都有限）

因此 $\dim_{\mathbb{R}} \{f : \mathbb{R} \to \mathbb{R} | \mathrm{Card} supp(f) < \aleph_0\} = \mathrm{Card} \left\{ f_x | x \in \mathbb{R}, f_x(y) = \begin{cases} 1, y = x \\ 0, y \neq x \end{cases} \right\} = \aleph_1$ □

1.4.4 线性空间的维数的性质

现在讨论一些维数的性质

1.

命题 1.4.1. 设 V 是一个线性空间，集合 $S \subseteq V$ 如果 $|S| > dimV$ 那么 S 至少有一个有限子集线性相关。

证明： 这其实是前面的 Steinitz 替换引理的直接结论

任取 V 的一个基 B，由基的定义和引理可知，如果 S 线性无关，那么一定有 $|S| \leqslant |B|$，与前提矛盾。

于是命题得证。 □

2.

命题 1.4.2. 线性空间 V 的一个任意一个子集 B，如果它的任意有限子集都线性无关，那么它可以扩充成一个 *Harmel* 基

证明： 我们只需要仿照证明 Harmel 基存在性的方式，构造一个偏序集 $H = \{S \supseteq B | S\text{的任意有限子集线性无关}\}$

同样地，H 一定存在一个极大元 $B' \supseteq B$，并且容易验证它的确是一个基。于是命题得证 □

3.

命题 1.4.3. 设 V 是一个线性空间, $W \subseteq V$ 是一个线性子空间, 那么有: $dimW \leqslant dimV$

证明: 取 W 的一个基 B。依定义，B 的任意子集在 W 中线性无关，于是它的任意子集在 V 中也线性无关

那么按照我们前面证明的命题，B 可以扩充成 V 的一个基 B'

那么有 $|B| \leqslant |B'|$，于是 $dimW \leqslant dimV$，命题得证 □

4.

命题 1.4.4. 设 V 是一个线性空间，它的一个子集 $S \subseteq V$，那么:

如果 S 是 V 的一个基，那么 S 的任意有限子集线性无关，并且 $\forall \alpha, S \cup \{\alpha\}$ 存在一个有限线性相关子集;

如果 $V \neq \{\mathbf{0}\}$,S 的任意有限子集线性无关，并且 $\forall \alpha, S \cup \{\alpha\}$ 存在一个有限线性相关子集，那么 S 是 V 的一个基

证明: 这是显然的，因为:

对于第一条结论，$\forall \alpha \in V, \exists B \subseteq S, |B| < \aleph_0, \alpha \in \operatorname{span}(B)$，此时一定有 $B \cup \{\alpha\}$ 线性相关和线性无关;

对于第二条结论，$\forall \alpha \in V, \exists B \subseteq S, |B| < \aleph_0, S \cup \{\alpha\}$ 线性相关，而又 S 线性无关，所以 α 可由 B 线性表出，因此 S 是一个基

于是命题得证。 □

如果 S 的任意有限子集线性无关，并且 $\forall \alpha, S \cup \{\alpha\}$ 存在一个有限线性相关子集，此时我们也说 S 是代数意义下 V 的一个极大线性无关集

接下来考虑几个只适用于有限维的结论

5.

命题 1.4.5. 设 V 是一个线性空间，$dimV = n$，那么 V 中的任意 n 个线性无关向量都可以构成 V 的一个基

证明: 这是显然的，因为这 n 个向量可以扩充成一个基，但是因为维数就是 n，如果添加任意向量都会导致线性相关，因此只有可能它本来就是一个基 □

6.

命题 1.4.6. 设 V 是一个线性空间，$dimV = n$，如果 $\forall \alpha \in V, \alpha \in span(\gamma_, \cdots, \gamma_n)$，那么 $\{\gamma_1, \cdots, \gamma_n\}$ 是 V 的一个基

证明: 任取 V 的一个基 $\{\beta_1, \cdots, \beta_n\}$

于是依命题条件，$\{\beta_1, \cdots, \beta_n\}$ 可由 $\{\gamma_1, \cdots, \gamma_n\}$ 线性表出

那么有 $rank(\beta_1, \cdots, \beta_n) \leqslant rank(\gamma_1, \cdots, \gamma_n) \leqslant n$

于是 $rank(\gamma_1, \cdots, \gamma_n) = n$，运用前面证明的命题后命题得证。 □

7.

命题 1.4.7. 设 $S = \{\alpha_x \in V | x \in A, |A| \leqslant \aleph_0\}$ 是一个至多可数的向量组

那么有：$dim\ span(S) = rank(S)$

命题 1.4.8. 只需取 S 的一个极大线性无关组 $B \subseteq S$。由极大线性无关组的性质，$span(S) = span(B)$

于是有 $dim\ span\ (S) = dim\ span\ (B) = |B| = rank(B) = rank(S)$，命题得证。

8.

命题 1.4.9. 设 $S = \{\alpha_x \in V | x \in A, |A| \leqslant \aleph_0\}, G = \{\alpha_x \in V | x \in B, |B| \leqslant \aleph_0\}$

那么：$span(S) = span(G) \Leftrightarrow S \cong G$

证明: 充分性是显然的，因为如果生成的子空间相同，两个向量组一定可以相互线性表出

必要性也是显然的，因为如果 S 可由 G 线性表出，一定有 $\mathrm{span}(S) \subseteq \mathrm{span}(G)$，同理有 $\mathrm{span}(S) \supseteq \mathrm{span}(G)$，于是命题得证。 □

1.5 线性空间的同构

在前面的章节中，我们讨论了线性组合、线性空间的基和维数。

本节开始，我们着手研究不同的空间之间的关系。本节我们考虑一个问题：不同的空间是否可以拥有相同的结构呢？我们应该如何刻画这个“相同”呢？

这个问题的答案是肯定的。这种关系称为同构，它具有一系列极为良好的性质。

1.5.1 同构的定义

定义 1.5.1: 线性空间的同构

V_1, V_2 是两个 F 上的线性空间，如果存在双射 $f : V_1 \to V_2$

满足：$\forall k \in F, \alpha, \beta \in V_1$

$f(k \cdot \alpha) = k \cdot f(\alpha), f(\alpha + \beta) = f(\alpha) + f(\beta)$

那么我们称 V_1 和 V_2 同构，记作 $V_1 \cong V_2$，并称 f 是 V_1 到 V_2 的同构映射

1.5.2 同构的性质

接下来讨论同构的一些性质

1.

命题 1.5.1. 同构是一种等价关系，即：

(a) $V_1 \cong V_1$

(b) $V_1 \cong V_2 \Rightarrow V_2 \cong V_1$

(c) $V_1 \cong V_2, V_2 \cong V_3 \Rightarrow V_1 \cong V_3$

证明： 首先证明自反性，即 $V_1 \cong V_1$。

考虑恒等映射 $I : V_1 \to V_1$，定义为 $I(\alpha) = \alpha$ 对于所有 $\alpha \in V_1$。

显然，I 是一个双射

$\forall \alpha_1, \alpha_2 \in V_1, k \in F$

$I(\alpha_1 + \alpha_2) = \alpha_1 + \alpha_2 = I(\alpha_1) + I(\alpha_2), I(k\alpha_1) = k\alpha_1 = kI(\alpha_1)$

于是自反性成立。

接下来证明对称性，即如果 $V_1 \cong V_2$，那么 $V_2 \cong V_1$。

假设 $V_1 \cong V_2$。这意味着存在一个线性同构 $T : V_1 \to V_2$。

由于 T 是同构映射，它一定是双射的，因此它的逆映射 $T^{-1} : V_2 \to V_1$ 存在，并且也是一个双射

接下来证明 T^{-1} 也是同构映射

设 $\beta_1, \beta_2 \in V$。因为 T 是满射，存在唯一的 $\alpha_1, \alpha_2 \in U$ 使得 $T(\alpha_1) = \beta_1$ 且 $T(\alpha_2) = \beta_2$。

因此 $\alpha_1 = T^{-1}(\beta_1)$ 且 $\alpha_2 = T^{-1}(\beta_2)$。

$T^{-1}(\beta_1 + \beta_2) = T^{-1}(T(\alpha_1) + T(\alpha_2))$

$= T^{-1}(T(\alpha_1+\alpha_2)) = \alpha_1+\alpha_2 = T^{-1}(\beta_1)+T^{-1}(\beta_2)$。

设 $\beta a \in V_2$ 和纯量 $k \in F$。存在唯一的 $\alpha \in U$ 使得 $T(\alpha)=\beta$，因此 $T^{-1}(\beta)=\alpha$。

$T^{-1}(k\beta) = T^{-1}(kT(\alpha))$

$= T^{-1}(T(k\alpha)) = k\alpha = kT^{-1}(\beta)$。

于是对称性是成立的

最后，我们证明传递性，即如果 $V_1 \cong V_2$ 且 $V_2 \cong V_3$，那么 $V_1 \cong V_3$。

假设 $V_1 \cong V_2$ 且 $V_2 \cong V_3$。

这意味着存在一个同构映射 $T_1 : V_1 \to V_2$ 和一个同构映射 $T_2 : V_2 \to V_3$。

考虑复合映射 $T_2 \circ T_1 : V_1 \to V_3$。因为 T_1 和 T_2 都是双射，因此他也是双射

设 $\alpha_1, \alpha_2 \in V_1$ 和纯量 $k \in F$。

$(T_2 \circ T_1)(\alpha_1+\alpha_2) = T_2(T_1(\alpha_1+\alpha_2))$

$= T_2(T_1(\alpha_1)+T_1(\alpha_2))$

$= (T_2 \circ T_1)(\alpha_1) + (T_2 \circ T_1)(\alpha_2)$。

$(T_2 \circ T_1)(k\alpha_1) = T_2(T_1(k\alpha_1))$

$= T_2(kT_1(\alpha_1))$

$= k(T_2 \circ T_1)(\alpha_1)$。

于是传递性成立，命题得证。 □

2.

命题 1.5.2. 如果 $f : V_1 \to V_2$ 是一个同构映射，那么 $f(\mathbf{0}_{V_1}) = \mathbf{0}_{V_2}$

证明： $f(\mathbf{0}_{V_1}) = f(0 \cdot \mathbf{0}_{V_1}) = 0 \cdot f(\mathbf{0}_{V_1}) = \mathbf{0}_{V_2}$ □

3.

命题 1.5.3. 如果 $f : V_1 \to V_2$ 是一个同构映射，那么 $f(-\alpha) = -f(\alpha)$

证明： $f((-1)\cdot\alpha) = (-1)\cdot f(\alpha) = -f(\alpha)$ □

4.

命题 1.5.4. 如果 $V_1 \cong V_2$，$f : V_1 \to V_2$ 是一个同构映射，$\alpha_1, \cdots, \alpha_n \in V_1$ 在 V_1 中线性相关

那么 $f(\alpha_1), \cdots, f(\alpha_n) \in V_2$ 在 V_2 中线性相关

证明： $k_1\alpha_1 + \cdots + k_n\alpha_n = \mathbf{0}_{V_1}$

$\Leftrightarrow f(k_1\alpha_1 + \cdots + k_n\alpha_n) = f(\mathbf{0}_{V_1}) = \mathbf{0}_{V_2}$

$\Leftrightarrow k_1 f(\alpha_1)+\cdots+k_n f(\alpha_n)=\mathbf{0}_{V_2}$

因此，$\alpha_1,\cdots,\alpha_n$ 线性相关

$\Leftrightarrow \exists k_1,\cdots,k_n,k_1\alpha_1+\cdots+k_n\alpha_n=\mathbf{0}_{V_1}$，其中 k_i 不全为零

$\Leftrightarrow \exists k_1,\cdots,k_n,k_1 f(\alpha_1)+\cdots+k_n f(\alpha_n)=\mathbf{0}_{V_2}$，其中 k_i 不全为零

$\Leftrightarrow f(\alpha_1),\cdots,f(\alpha_n)$ 线性相关 □

其实，这个命题也指出了线性无关的向量组在同构映射作用下也是线性无关的，也就是说同构线性相关和线性无关性

另外一个有趣的点是，其实在这个命题的证明中，我们并没有用到同构映射的满射性，只需要它是单射即可。

5.

命题 1.5.5. 设 V_1,V_2 是两个线性空间，$\dim V_1=\dim V_2=n<\aleph_0$

映射 $T:V_1\to V_2$ 如果是单射，并且满足：

$T(\alpha+\beta)=T(\alpha)+T(\beta),T(k\alpha)=kT(\alpha)$

那么 T 是一个同构映射

证明: 取 V_1 的一组基 $\{\alpha_1,\cdots,\alpha_n\}$，按基的定义，它一定线性无关

运用前面命题的证明方法可得，$T(\alpha_1),\cdots,T(\alpha_n)$ 线性无关。

而 $\dim V_2=n$，于是 $\{T(\alpha_1),\cdots,T(\alpha_n)\}$ 是 V_2 的一个基

那么，我们说，T 一定也是满射。因为:

$\forall\beta=k_1T(\alpha_1)+\cdots+k_nT(\alpha_n)\in V_2$，注意到:

$T(k_1\alpha_1+\cdots+k_n\alpha_n)=k_1T(\alpha_1)+\cdots+k_nT(\alpha_n)=\beta$，于是命题得证。 □

这个命题指出，其实同构映射满射的条件也可以用维数有限且相等代替。值得注意的是，维数有限是必须的。

6.

命题 1.5.6. 如果 $V_1\cong V_2$，$f:V_1\to V_2$ 是一个同构映射，$B_1\subseteq V_1$ 是 V_1 的一个基

那么 $B_2=f(B_1)\subseteq V_2$ 是 V_2 的一个基

证明: 由前面的命题可知 $B_2=f(B_1)$ 的线性无关性

现在取 $\forall\beta\in V_2$

一定存在 $\alpha\in V_1,f(\alpha)=\beta$

因为 B_1 是 V_1 的基，所以一定能找到它的一个有限子集 $\{\eta_1,\cdots,\eta_n\}$

$\exists k_1, \cdots, k_n, k_1\eta_1 + \cdots + k_n\eta_n = \alpha$

$\Rightarrow k_1 f(\eta_1) + \cdots, k_n f(\eta_n) = \beta$

而 $\{f(\eta_i)\} \subseteq B_2 = f(B_2)$

因此 B_2 的确是 V_2 的基，命题得证 □

7.

命题 1.5.7. $V_1 \cong V_2 \Leftrightarrow \dim V_1 = \dim V_2$

证明： 充分性是显然的，因为任取 V_1 的一个基 B_1 和一个同构映射 f，依照前面的命题，$f(B_1)$ 一定是 V_2 的一个基

接下来证明必要性。分别取 V_1, V_2 的一个基 B_1, B_2

因为 $\dim V_1 = \dim V_2$，一定有一个双射 $g : B_1 \to B_2$

我们构造以下映射：

$$f : V_1 \ni \alpha = \sum_{i=1}^{n} k_i\alpha_i \mapsto \beta = \sum_{i=1}^{n} k_i g(\alpha_i) \in V_2, \alpha_i \in B_1 \tag{1.2}$$

我们验证它的确是一个同构映射。

首先，它是一个满射，因为 k_i 是可以在 F 中任意选取的，而 $B_2 = g(B_1)$ 是一个基。

其次，它是一个单射，因为基的线性无关性，k_i 一定是唯一的。

我们最后验证可加性和齐次性

我们不失一般性的认为接下来的 α, β 可由同一有限子集线性表出，因为如果表出它们的子集不同，我们只需要取并，随后在没有分量的位置赋予零的系数

设 $\alpha = \sum\limits_{i=1}^{n} k_i\eta_i, \beta = \sum\limits_{i=1}^{n} l_i\eta_i, p \in F$

$f(\alpha + \beta) = f(\sum\limits_{i=1}^{n} k_i\eta_i + \sum\limits_{i=1}^{n} l_i\eta_i)$

$= f\left(\sum\limits_{i=1}^{n}(k_i + l_i)\eta_i\right) = \sum\limits_{i=1}^{n}(k_i + l_i)g(\eta_i)$

$= \sum\limits_{i=1}^{n} k_i g(\eta_i) + \sum\limits_{i=1}^{n} l_i g(\eta_i) = f(\alpha) + f(\beta)$

$f(p\alpha) = f(p\sum\limits_{i=1}^{n} k_i\eta_i) = f(\sum\limits_{i=1}^{n} pk_i\eta_i)$

$= \sum\limits_{i=1}^{n} pk_i g(\eta_i) = p\sum\limits_{i=1}^{n} k_i g(\eta_i) = pf(\alpha)$

于是命题得证 □

它的一个常用推论是：

推论 1.5.1

如果 $\dim_F V = n < \aleph_0$，那么 $V \cong F^n$

证明： 依照上面的命题，这是显然的 $\square$

8.

命题 1.5.8. 设 $V_1 \cong V_2$ 是同构的两个线性空间，$f : V_1 \to V_2$ 是一个同构映射

那么，如果 W 是 V_1 的一个子空间，那么 $f(W)$ 是 V_2 的一个子空间，并且 $\dim W = \dim f(W)$

证明： 首先，因为 $\mathbf{0}_{V_1} \in W$，所以 $f(\mathbf{0}_{V_1}) \in f(W)$，因此 $f(W) \neq \emptyset$

取 $\forall f(\alpha), f(\beta) \in f(W), k \in F$

$f(\alpha) + f(\beta) = f(\alpha + \beta) \in f(W), kf(\alpha) = f(k\alpha) \in f(W)$，因此 $f(W)$ 是 V_2 的子空间

接下来证明 $\dim W = \dim f(W)$。设 $g = f|_W : W \to f(W)$

首先，g 显然是一个双射，因为 f 就是一个双射

注意到：$g(\alpha + \beta) = g(\alpha) + g(\beta), g(k\alpha) = kg(\alpha)$

因此 g 是一个 W 到 $f(W)$ 的同构映射，于是 $W \cong f(W)$，那么 $\dim W = \dim f(W)$，命题得证 $\square$

1.6 线性子空间的直和

在研究了同构后，我们转而研究另一种关系：直和。与之间我们研究的是结构相同的空间不同，我们此时寻求一个空间可以分裂为数个空间，而且最好可以对每一个向量唯一分解。

先提出线性空间和的概念。

1.6.1 子空间的和

定义 1.6.1: 子空间的和

设 V 是一个线性空间，$V_1, V_2 \subseteq V$ 是 V 的子空间

我们定义：$V_1 + V_2 = \{\alpha_1 + \alpha_2 | \alpha_1 \in V_1, \alpha_2 \in V_2\}$

称为 V_1 和 V_2 的和

接下来讨论子空间和的性质

1.

命题 1.6.1. 设 V 是一个线性空间，$V_1, V_2 \subseteq V$ 是 V 的子空间

$V_1 + V_2$ 也是 V 的一个子空间

证明： 只需验证 $V_1 + V_2$ 非空且对加法和纯量乘法封闭

首先，$\mathbf{0} \in V_1, \mathbf{0} \in V_2$，所以一定有 $\mathbf{0} = \mathbf{0} + \mathbf{0} \in V_1 + V_2$

取 $\forall k \in F, \alpha_1 + \beta_1, \alpha_2 + \beta_2 \in V_1 + V_2, \alpha_1, \alpha_2 \in V_1, \beta_1, \beta_2 \in V_2$

于是 $(\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \in V_1 + V_2$

$k \cdot (\alpha_1 + \beta_1) = k \cdot \alpha_1 + k \cdot \beta_1 \in V_1 + V_2$，于是命题得证 □

2.

命题 1.6.2. $V_1 + V_2 = V_2 + V_1$

证明： 这是显然的，因为 $\forall \alpha \in V_1, \beta \in V_2, \alpha + \beta = \beta + \alpha$ □

3.

命题 1.6.3. $V_1 + (V_2 + V_3) = (V_1 + V_2) + V_3$

证明： 这是显然的，因为 $\forall \alpha \in V_1, \beta \in V_2, \gamma \in V_3, \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ □

4.

命题 1.6.4. $\operatorname{span}(S) + \operatorname{span} span(B) = \operatorname{span}(S \cup B), |S|, |B| \leqslant \aleph_0$

证明： 设 $S = \{\alpha_x \in V | x \in A\}, B = \{\beta_y \in V | y \in B\}$

$\forall \sum\limits_{x \in A} k_x \alpha_x \in \operatorname{span}(S), \sum\limits_{y \in B} l_y \beta_y \in \operatorname{span}(B)$

$\sum\limits_{x \in A} k_x \alpha_x + \sum\limits_{y \in B} l_y \beta_y \in \operatorname{span}(S \cup B) \Rightarrow \operatorname{span}(S) + \operatorname{span}(B) \subseteq \operatorname{span}(S \cup B)$

$$\forall \sum_{z \in A \cup B} b_z \gamma_z, \gamma_z = \begin{cases} \alpha_z, z \in A \\ \\ \beta_z, z \in B - A \end{cases}$$

$\sum\limits_{z \in A \cup B} b_z \gamma_z = \sum\limits_{x \in A} b_x \alpha_x + \sum\limits_{y \in B - A} b_y \beta_y \in \operatorname{span}(S) + \operatorname{span}(B)$

于是命题得证 □

引理 1.6.1

设 V 是一个线性空间，V_1, V_2 是 V 的一个子空间

那么 $V_1 \cap V_2$ 也是 V 的一个子空间

5.

证明： 因为 $\mathbf{0} \in V_1, V_2, \mathbf{0} \in V_1 \cap V_2$

现在取 $\forall k \in F, \alpha, \beta \in V_1 \cap V_2$

$\because \alpha \in V_1, \beta \in V_1$

$\therefore \alpha + \beta \in V_1$

同理有 $\alpha + \beta \in V_2$。因此，$\alpha + \beta \in V_1 \cap V_2$

$\because \alpha \in V_1, \therefore k\alpha \in V_1$

同理有 $k\alpha \in V_2$，于是有 $k\alpha \in V_1 \cap V_2$

于是命题得证

$\square$

命题 1.6.5. $\dim(V_1 + V_2) + \dim(V_1 \cap V_2) = \dim(V_1) + \dim(V_2)$

证明： 取 $V_1 \cap V_2$ 的一个基 B_0。按照前面已经证明的结论，我们可以将其分别补充为 $B_0 \cup B_1, B_0 \cup B_2$ 分别作为 V_1, V_2 的基

我们只需证明：$B_0 \cup B_1 \cup B_2$ 是 $V_1 + V_2$ 的一个基

我们首先证明它线性无关。我们只需取它的一个有限子集 $S_0 \cup S_1 \cup S_2 \subseteq B_0 \cup B_1 \cup B_2, S_0 \subseteq B_0, S_1 \subseteq B_1, S_2 \subseteq B_2$

设 $B_0 = \{\alpha_1, \cdots, \alpha_m\}, B_1 = \{\beta_1, \cdots, \beta_n\}, B_2 = \{\gamma_1, \cdots, \gamma_l\}$

考虑线性组合 $\sum\limits_{i=1}^{m} a_i\alpha_i + \sum\limits_{j=1}^{n} b_j\beta_j + \sum\limits_{k=1}^{l} c_k\gamma_k = \mathbf{0}$

那么，$\sum\limits_{i=1}^{m} a_i\alpha_i + \sum\limits_{j=1}^{n} b_j\beta_j = -\sum\limits_{k=1}^{l} c_k\gamma_k$

但是，等式左边的向量是由 $S_0 \cup S_1$ 张成的，因此是 V_1 的元素；右边的向量由 S_2 张成，因此是 V_2 的元素

再由线性空间的性质可得 $\sum\limits_{i=1}^{m} a_i\alpha_i + \sum\limits_{j=1}^{n} b_j\beta_j \in V_1 \cap V_2$

事实上，一定有 $\forall b_j = 0$，因为：

$\sum\limits_{i=1}^{m} a_i\alpha_i + \sum\limits_{j=1}^{n} b_j\beta_j \in V_1 \cap V_2$，那么它一定也可以在 S_0 上表示，即

$\exists\{d_i\}, \sum\limits_{i=1}^{m} a_i\alpha_i + \sum\limits_{j=1}^{n} b_j\beta_j = \sum\limits_{i=1}^{m} d_i\alpha_i$

此时如果 $\exists b_j \neq 0$，那么就给出了一个向量在线性无关向量组 $S_0 \cap S_1$ 上的两种不同表示（注意等式右侧可以视为 S_1 部分缺省的一种表出），而这是不可能的。

于是有 $\sum\limits_{i=1}^{m} a_i\alpha_i = -\sum\limits_{k=1}^{l} c_k\gamma_k$，同理可推出 $\forall a_i, c_k = 0$。

于是我们证明了 $B_0 \cup B_1 \cup B_2$ 线性无关。而由构造的方式可知，它一定可以表出 $V_1 + V_2$ 中的任意向量。

此时有 $\mathrm{Card}(B_0 \cup B_1 \cup B_2) + \mathrm{Card}\, B_0 = \mathrm{Card}(B_0 \cup B_1) + \mathrm{Card}(B_0 \cup B_2)$

$\Rightarrow \dim(V_1+V_2)+\dim(V_1\cap V_2)=\dim V_1+\dim V_2$，于是命题得证 □

可以立即得到下面的推论：

推论 1.6.2

$\dim(V_1+V_2)=\dim V_1+\dim V_2 \Leftrightarrow V_1\cap V_2=\{\mathbf{0}\}$

证明： 这是显然的，因为维数等于零的线性空间有且只有 $\{\mathbf{0}\}$ □

1.6.2 子空间的直和

我们接下来转入一种特殊的和：使和空间中任意向量表出都唯一的空间。这样的空间的好处是：未来我们对向量做分解时，只会“投影”出唯一的分量

定义 1.6.2: 子空间的直和

设 V 是一个线性空间，$V_1,V_2\subseteq V$ 是 V 的子空间

$W=V_1+V_2$ 如果满足：$\forall\alpha\in W,\alpha=\alpha_1+\alpha_2,\alpha_1\in V_1,\alpha_2\in V_2$ 的表出方式唯一

那么我们称 W 是 V_1 和 V_2 的直和，记作 $W=V_1\oplus V_2$。

如果 $V=V_1\oplus V_2$，我们称 V_2 是 V_1 的补空间

我们首先探究一些更加简洁易用的直和充要条件

定理 1.6.3: 子空间直和的等价命题

设 V 是一个线性空间，$V_1,V_2\subseteq V$ 是 V 的子空间

以下四个命题等价：

1. V_1+V_2 是直和
2. V_1+V_2 中 $\mathbf{0}$ 的表法唯一
3. $V_1\cap V_2=\{\mathbf{0}\}$
4. 分别任取 V_1,V_2 的一个基 B_1,B_2，$B_1\cup B_2$ 是 V_1+V_2 的一个基

证明： 我们首先证明前三个命题等价。

第一个命题显然可以推出第二个命题，因为这就是直和的定义；

接下来证明第二个命题可以推出第三个命题。取 $\alpha\in V_1\cap V_2$，这是一定可以取到的，因为 V_1,V_2 至少包含零向量

那么，因为 $V_1 \cap V_2$ 是一个线性空间，一定 $\exists -\alpha \in V_1 \cap V_2, \alpha + (-\alpha) = \mathbf{0}$

此时，$\alpha \in V_1, -\alpha \in V_2$，但是 $\mathbf{0}$ 仅有一种表出方式，所以一定有 $\alpha = \mathbf{0}$，得证

接下来证明第三个命题可推出第一个命题。取 $\gamma \in V_1 + V_2$，假设它有两种表出方式 $\gamma = \alpha_1 + \beta_1 = \alpha_2 + \beta_2$

那么有 $(\alpha_1 - \alpha_2) + (\beta_1 - \beta_2) = \mathbf{0}$

此时，$\beta_1 - \beta_2 = -(\alpha_1 - \alpha_2) \in V_1$，于是 $\beta_1 - \beta_2 \in V_1 \cap V_2$

所以 $\beta_1 - \beta_2 = \mathbf{0} \Rightarrow \alpha_1 - \alpha_2 = \mathbf{0}$，这说明表出方式的确是唯一的。

接下来证明第三个命题和第四个命题可以互推。只需要仿照前面证明和空间维数性质的方式证明即可。

于是命题得证。 □

我们可以得出一个后面很常用的推论

推论 1.6.4

$V_1 + V_2$ 是直和 $\Leftrightarrow \dim(V_1 + V_2) = \dim V_1 + \dim V_2$

证明: $V_1 + V_2$ 是直和 $\Leftrightarrow V_1 \cap V_2 = \{\mathbf{0}\} \Leftrightarrow \dim(V_1 + V_2) = \dim V_1 + \dim V_2$ □

这个定理的另一个推论是：补空间总是存在的。

推论 1.6.5

线性空间 V 的任意子空间 W 的补空间总是存在的。

证明: 取 W 的一个基 B_0，将其补充至其成为 V 的一个基 $B_0 \cap B_1$

设 $E = \{\alpha | \alpha \in \operatorname{span}(B), B \subseteq B_1, |B| \leqslant \aleph_0\}$，它的一个基是

事实上，显然一定有 $W + E = V$，而它们的基的并是和的基

因此有 $W \oplus E = V$，命题得证。 □

1.6.3 多个子空间的直和

我们仿照前面的方式定义有限个子空间的直和

定义 1.6.3: 有限个子空间的直和

设 V 是一个线性空间，$V_1, \cdots, V_n \subseteq V$ 是 V 的子空间

$W = V_1 + \cdots + V_n$ 如果满足：$\forall \alpha \in W, \alpha = \alpha_1 + \cdots + \alpha_n, \alpha_1 \in V_1, \cdots, \alpha_n \in V_n$ 的表出方式唯一

那么我们称 W 是 $V_1, \cdots, V_2$ 的直和，记作 $W = \bigoplus_{i=1}^n V_i = V_1 \oplus \cdots \oplus V_2$。

但是，如果希望拓展到可数乃至不可数个子空间的直和，就无法这样定义了，因为此时向量求和逐渐变得难以定义

我们转而定义直和为一系列映射的集合，我们后面将看到，对于有限情形，它和前面的定义是同构的。

定义 1.6.4: 子空间直和的一般定义

V 是一个线性空间，$\{V_x \subseteq V | x \in I\}$ 是 V 的一族子空间

我们定义：

$$\bigoplus_{x \in I} V_x = \{f : I \to \bigcup_{x \in I} V_x | f(x) \in V_x, |supp(f)| < \aleph_0\} \tag{1.3}$$

称为这一族子空间的直和

接下来讨论直和的性质

1\.

命题 1.6.6. $V_1 \oplus \cdots \oplus V_n \cong \{f : I \to \bigcup_{i=1}^n V_i | f(i) \in V_i, |supp(f)| < \aleph_0\}, I = \{1, \cdots, n\}$

证明： 考虑映射 $T : V_1 \oplus \cdots \oplus V_n \ni \alpha = \alpha_1 + \cdots + \alpha_n \mapsto \left(f_\alpha : I \ni i \mapsto \alpha_i \in \bigcup_{i=1}^n V_i\right) \in$ $\{f : I \to \bigcup_{i=1}^n V_i | f(i) \in V_i, |supp(f)| < \aleph_0\}$

这个映射的定义是良好的，因为直和中向量的标出方式唯一

我们验证它是一个同构映射。

首先，它显然是一个单射；其次，它也是一个满射，因为直和中每一个子空间中的向量的选取没有任何限制

接下来验证可加性和齐次性

$\forall \alpha = \alpha_1 + \cdots + \alpha_n, \beta = \beta_1 + \cdots + \beta_n \in V_1 \oplus \cdots \oplus V_n, k \in F$

$T(\alpha + \beta) = f_{\alpha+\beta}, f_{\alpha+\beta}(i) = \alpha_i + \beta_i = f_\alpha(i) + f_\beta(i)$

$\Rightarrow T(\alpha+\beta)=f_{\alpha+\beta}=f_\alpha+f_\beta=T(\alpha)+T(\beta)$

$T(k\alpha)=f_{k\alpha}, f_{k\alpha}(i)=k\alpha_i=kf_\alpha$

$\Rightarrow T(k\alpha)=f_{k\alpha}=kf_\alpha=kT(\alpha)$

于是命题得证。 □

因此第二种定义的确可以视为第一种定义的推广。

值得注意的是，我们尽管定义了无限个子空间的直和，但是并没有定义和，所以之前讨论的性质很多只能受限在有限范围

2.

命题 1.6.7. 设 V 是一个线性空间，$V_1,\cdots,V_n\subseteq V$ 是 V 的子空间

以下四个命题等价：

(a) $\bigcup\limits_{i=1}^{n} V_i$ 是直和

(b) $\bigcup\limits_{i=1}^{n} V_i$ 中 $\mathbf{0}$ 的表法唯一

(c) $\forall i, V_i\cap(\bigcup\limits_{j\neq i} V_j)=\{\mathbf{0}\}$

(d) 分别任取 $V_1,\cdots,V_n$ 的一个基 $B_1,\cdots,B_n$，$\bigcup\limits_{i=1}^{n} B_i$ 是 $\bigcup\limits_{i=1}^{n} V_i$ 的一个基

证明： 仿照两个子空间的情形即可。 □

也可以得出类似的推论：

推论 1.6.6

$V_1+\cdots+V_n$ 是直和 $\Leftrightarrow \dim(V_1+\cdots+V_n)=\dim V_1+\cdots+\dim V_n$

证明： 这是前面命题的直接推论。因为从证明过程我们知道，直和当且仅当子空间的基互不相交地组成和空间的基 □

1.7 商空间

本节中，我们考虑一种新的线性空间的构造方式：商空间。商空间是将一个线性空间中的向量划分为等价类，进而构造出一个新的线性空间。

比如说，在三维空间 $\mathbb{R}^3$ 中，我们可以把所有的水平面看作一个子空间，然后将每个平面上的向量看作一个等价类。这样，我们就可以得到一个新的线性空间，其中的元素是平面，

而不是具体的向量。

同时，我们也可以看到，于是同时每一个巨大的平面变得可以由与 z 轴的交点来唯一标识，从而我们做到了“消去二维平面的影响”。

商空间比较抽象，但是在后续线性映射的研究中有巨大作用。

1.7.1 商空间的定义

在定义商空间之前，我们先定义一个关于子空间的等价关系

定义 1.7.1: 关于子空间的等价

设 V 是一个线性空间，W 是 V 的子空间，我们定义以下等价关系 $\sim$:

$\alpha, \beta \in V, \alpha \sim \beta \Leftrightarrow \alpha - \beta \in W$

我们首先验证它的确是一个等价关系

命题 1.7.1. $\sim$ *是一个等价关系*

证明: 首先验证它满足自反性，$\forall \alpha, \alpha - \alpha = \mathbf{0} \in W$

接下来验证它满足对称性，$\alpha \sim \beta \Leftrightarrow \alpha - \beta \in W \Leftrightarrow \beta - \alpha = -(\alpha - \beta) \in W \Leftrightarrow \beta \sim \alpha$

最后验证它满足传递性,$\alpha \sim \beta, \beta \sim \gamma \Leftrightarrow \alpha - \beta \in W, \beta - \gamma \in W \Rightarrow \alpha - \gamma \in W \Rightarrow \alpha \sim \gamma$ □

在这个等价关系下 $\alpha \in V$ 的等价类 $\bar{\alpha}$ 也常常记作 $\alpha + W$，这个表示方式是自然的，因为:

$\bar{\alpha} = \{\beta | \beta - \alpha \in W\} = \{\alpha + \beta | \beta \in W\}$ 现在我们来定义商空间

定义 1.7.2: 商空间

设 V 是一个线性空间，W 是 V 的一个子空间，我们定义:

$$V/W := \{\alpha + W | \alpha \in V\} \tag{1.4}$$

称为 V 关于 W 的商空间，其中的加法和纯量乘法定义为:

$(\alpha + W) + (\beta + W) := (\alpha + \beta) + W$

$k \cdot (\alpha + W) = k\alpha + W$

我们首先验证 V/W 的加法和乘法的定义是良好的

命题 1.7.2. 如果 $\alpha_1 \sim \alpha_2, \beta_1 \sim \beta_2, k \in F$，那么

$(\alpha_1 + W) + (\beta_1 + W) = (\alpha_2 + W) + (\beta_2 + W)$

$k(\alpha_1 + W) = k(\alpha_2 + W)$

证明: 首先，$\alpha_1 \sim \alpha_2 \Leftrightarrow \alpha_1 - \alpha_2 \in W$

$\beta_1 \sim \beta_2 \Leftrightarrow \beta_1 - \beta_2 \in W$

于是有 $(\alpha_1 + \beta_1) - (\alpha_2 + \beta_2) = (\alpha_1 - \alpha_2) + (\beta_1 - \beta_2) \in W$

因此 $(\alpha_1 + W) + (\beta_1 + W) = (\alpha_2 + W) + (\beta_2 + W)$

同理可得 $k(\alpha_1 + W) = k(\alpha_2 + W)$，命题得证 □

于是它的定义的确是良好的。V/W 中的加法和乘法符合公理是显然的。

这样，我们就完成了商空间的定义，并验证了它的确是一个线性空间

1.7.2 商空间的性质

1. **定理 1.7.1**

   设 V 是一个线性空间，$V = W \oplus U$

   那么，$V/W \cong U$

证明: 设 $T : U \ni \alpha \mapsto \alpha + W \in V/W$

我们验证它的确是一个同构映射

首先验证它是一个单射。设 $T(\alpha) = T(\beta), \alpha, \beta \in U$

$\Rightarrow \alpha + W = \beta + W \Rightarrow \alpha - \beta = W$

记 $\alpha - \beta = \gamma$，事实上，只能有 $\gamma = \mathbf{0}$，否则 $\alpha = \beta + \gamma$ 就给出了 α 的第二种表示方式，但这在直和中是不可能的。

因此 $\alpha = \beta$，T 的确是单射

接下来证明它是一个满射，取 $\forall \alpha \in V, \alpha = \alpha_1 + \alpha_2, \alpha_1 \in W, \alpha_2 \in U$

注意到，$T(\alpha_2) = \alpha_2 + W = \alpha + W$，因为 $\alpha - \alpha_2 = \alpha_1 \in W$，于是 T 的确是一个满射

最后我们验证可加性和齐次性。

$T(\alpha + \beta) = (\alpha + \beta) + W = (\alpha + W) + (\beta + W) = T(\alpha) + T(\beta)$

$T(k\alpha) = k\alpha + W = k(\alpha + W) = kT(\alpha)$

于是命题得证

□

值得注意的是，这个定理的逆命题并不成立（尽管逆命题的形式优美到让人难以想象它是错误的），比如说：

考虑 $\mathbb{R}^3/(\mathbb{R},\mathbb{R},0) \cong (\mathbb{R},0,0)$，但是 $(\mathbb{R},\mathbb{R},0)+(\mathbb{R},0,0)=(\mathbb{R},\mathbb{R},0)\neq\mathbb{R}^3$

后面我们看到，如果追加一些基的限制，我们可以让逆命题成立

2. 商空间的维数

利用上面的定理可以立即得出以下结论：

推论 1.7.2: 商空间的维数

$\dim V/W = \dim V - \dim W$

证明: 我们之前已经证明，任意子空间都有补空间

所以只需取 W 的补空间 U，此时有 $V = W \oplus U$

$\Rightarrow V/W \cong U \Rightarrow \dim V/W = \dim U = \dim V - \dim W$，命题得证 □

3.

命题 1.7.3. 设 V 是一个线性空间，$V = W + U, U = \{\alpha|\alpha\in \operatorname{span}(S), S\subseteq B, |S| < \aleph_0\}$ 那么，V/W 的一个基是 $\{\alpha + W|\alpha\in B\} \Leftrightarrow V = W\oplus U, B$ 是 U 的一个基

证明: 先证明充分性。

首先，B 一定是线性无关的，因为：$\sum\limits_{i=1}^{n} k_i(\alpha_i + W) = \mathbf{0}_{V/W} = W$

$\Leftrightarrow \sum\limits_{i=1}^{n} k_i\alpha_i + W = W \Leftrightarrow \sum\limits_{i=1}^{n} k_i\alpha_i = \mathbf{0}$

那么因为 $\{\alpha + W|\alpha\in B\}$ 是基，B 也一定线性无关

于是 B 是 U 的一个基，接下来证明 $V = W\oplus U$

取 $\forall\alpha\in W\cap U$，一定有 $\alpha = \sum\limits_{i=1}^{n} k_i\alpha_i$，其中 $\alpha_i\in B$

又因为 $\alpha\in W$，那么有 $\alpha + W = W \Rightarrow \sum\limits_{i=1}^{n} k_i(\alpha_i + W) = W \Rightarrow k_i = 0$

于是 $\alpha = \mathbf{0} \Rightarrow W\cap U = \{\mathbf{0}\}$，因此 $V = W\oplus U$，充分性得证。

必要性利用前面的定理是显然的，因为 $V/W\cong U$，其同构映射为 $V/W \ni \alpha + W \mapsto \alpha\in U$

那么按照基在同构映射下映射为一个基的性质，必要性得证。

于是命题得证。 □

1.8 对偶空间

本节中，我们研究一种比较具体的线性空间：对偶空间。对偶空间是一个映射集，作为下一章的铺垫。我们不会聚焦于线性映射的具体结构，而是关注它的线性空间性质

1.8.1 对偶空间的定义

定义 1.8.1: 对偶空间

设 V 是一个 F 上的线性空间，我们定义：

$$V^* = \{f : V \to F | \forall \alpha, \beta \in V, k \in F, f(\alpha+\beta) = f(\alpha) + f(\beta), f(k\alpha) = kf(\alpha)\} \tag{1.5}$$

其中 V^* 的基域是 F，它的加法和纯量乘法是即是函数的加法和纯量乘法

称为 V 的对偶空间

1.8.2 有限维对偶空间的对偶基

我们希望知道，对偶空间的维数到底是多少。在本节中，我们仅仅探讨有限维线性空间的对偶空间，因为进一步的讨论需要线性映射的进一步知识

我们指出，此时有一类与原空间的基密切相关的基，我们称之为对偶基

定义 1.8.2: 对偶基

设 V 是一个线性空间，$\dim V = n < \aleph_0$，$B = \{\alpha_1, \cdots, \alpha_n\}$ 是它的一个基

集合 $B^* = \{f_i \in V^* | f_i(\alpha_j) = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}, 1 \leqslant i \leqslant n\}$

称为 B 的对偶基

对偶基的定义的确是良好的，因为 $B = \{\alpha_1, \cdots, \alpha_n\}$ 是一个 V 的基，又函数是线性的，所以其在基上的映射的确给出了函数的完整定义

我们接下来验证 B^* 是 V^* 的一个基。

命题 1.8.1. *B^* 是 V^* 的一个基*

证明: 首先证明 B^* 是线性无关的。

取 B^* 的任意有限子集 $\{f_1, \cdots, f_n\}$

考虑线性组合 $k_1f_1+\cdots+k_nf_n=\mathbf{0}_{V^*}$

代入 α_i，得：$k_1f_1(\alpha_i)+\cdots+k_nf_n(\alpha_i)=0\Rightarrow k_i=0$。因此 B^* 线性无关。

接下来证明它的确可以生成整个对偶空间

$\forall A\in V^*$，考虑线性组合 $T=\sum\limits_{i=1}^{n}A(\alpha_i)f_i\in V^*$

我们来证明，$A=T$。

$\forall\gamma=\sum\limits_{i=1}^{n}k_i\alpha_i\in V$

$T(\gamma)=\sum\limits_{i=1}^{n}A(\alpha_i)f_i(\gamma)=\sum\limits_{i=1}^{n}A(\alpha_i)f_i(\sum\limits_{i=1}^{n}k_i\alpha_i)$

$=\sum\limits_{i=1}^{n}A(\alpha_i)k_i=A(\sum\limits_{i=1}^{n}k_i\alpha_i)=A(\gamma)$。b 于是命题得证。 □

它的直接推论是我们之前提出的问题的答案，即：

推论 1.8.1: 有限维线性空间的对偶空间的维数

如果 $\dim V=n<\aleph_0$，那么 $\dim V^*=n$

证明过程中还使用了一个后续很常用的性质：

命题 1.8.2. 设 $\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一个基，$\{f_1,\cdots,f_n\}$ 是它的对偶基，那么 $\forall A\in V^*,A=\sum\limits_{i=1}^{n}A(\alpha_i)f_i$

类似地，我们也可以对向量作一个展开

命题 1.8.3. 设 $\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一个基，$\{f_1,\cdots,f_n\}$ 是它的对偶基，那么 $\forall\gamma\in V,\gamma=\sum\limits_{i=1}^{n}f_i(\gamma)\alpha_i$

1.8.3 双重对偶空间

定义 1.8.3: 双重对偶空间

V^* 的对偶空间称为 V 的双重对偶空间，记作 V^{**}

双重对偶空间的最大意义是，它体现了对偶并非是一个“越来越远”的操作，反而每两次操作在同构意义下可以抵消

接下来的命题将逐步指出为什么我们这么说

定理 1.8.2: 线性空间与其双重对偶空间的自然同构

设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$。

那么 $V \cong V^{**}$，并且存在不依赖于基的自然同构

$$\psi : V \ni \alpha \mapsto (\alpha^{**} : V^* \ni f \mapsto f(\alpha) \in F) \in V^{**} \tag{1.6}$$

证明: 首先验证 ψ 是一个单射。

如果不成立，那么 $\exists \alpha_1 \neq \alpha_2, \psi(\alpha_1) = \psi(\alpha_2) = \alpha^{**}$

依照 ψ 的定义，此时有：$\forall f \in V^*, f(\alpha_1) = f(\alpha_2)$。

我们只需举出一个 V^* 中的映射即可推出矛盾。比如说：

记 $\gamma = \alpha_1 - \alpha_2$，将其补充为 V 的一个基 $\{\gamma, \beta_1, \cdots, \beta_{n-1}\}$

考虑映射 $f \in V^*, f(\gamma) = 1, f(\beta_i) = 0$。但是本来理应 $f(\gamma) = 0$，与假设矛盾。

所以 ψ 是一个单射。

我们接下来验证可加性和齐次性。

$\forall \alpha, \beta \in V, k \in F$

$\psi(\alpha+\beta)(f) = f(\alpha+\beta) = f(\alpha) + f(\beta) = \psi(\alpha)(f) + \psi(\beta)(f) \Rightarrow \psi(\alpha+\beta) = \psi(\alpha) + \psi(\beta)$

$\psi(k\alpha)(f) = f(k\alpha) = kf(\alpha) = k\psi(\alpha)(f) \Rightarrow \psi(k\alpha) = k\psi(\alpha)$

我们其实不需要验证满射性，因为 $\dim V = \dim V^{**} = n < \aleph_0$，结合单射性和可加性和齐次性，我们可以直接得出 ψ 是一个同构映射。

于是命题得证。 □

之所以我们说“两次对偶相互抵消”，是因为在一般的线性空间之间，两次同构会使基的作用堆叠，其同构依旧是依赖于基的。但是自然同构却不依赖于任何基

接下来的命题让我们看到，自然同构不仅仅是消除了基的影响的同构，而且还恰好是由两次对偶基导出的同构映射的复合

命题 1.8.4. 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$。

设 $\{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一个基，$\{f_1, \cdots, f_n\}$ 是它的对偶基，$\{\alpha_1^{**}, \cdots, \alpha_n^{**}\}$ 是 $\{f_1, \cdots, f_n\}$ 的对偶基

设 $\sigma : V \ni \sum\limits_{i=1}^{n} k_i\alpha_i \mapsto \sum\limits_{i=1}^{n} k_i f_i \in V^*$

$\tau : V^* \ni \sum\limits_{i=1}^{n} l_i f_i \mapsto \sum\limits_{i=1}^{n} l_i \alpha_i^{**} \in V^{**}$

那么，$\psi = \tau \circ \sigma$

证明： $\forall \gamma = \sum_{i=1}^{n} k_i \alpha_i \in V, f \in V^*$

$$(\tau \circ \sigma)(\gamma)(f) = \tau\left(\sigma(\sum_{i=1}^{n} k_i \alpha_i)\right)(f)$$

$$= \tau(\sum_{i=1}^{n} k_i f_i)(f) = \sum_{i=1}^{n} k_i \alpha_i^{**}(f)$$

$$= \sum_{i=1}^{n} k_i \alpha_i^{**}\left(\sum_{j=1}^{n} f(\alpha_j) f_j\right)$$

$$= \sum_{i=1}^{n} k_i f(\alpha_i) = f\left(\sum_{i=1}^{n} k_i \alpha_i\right)$$

$$= f(\gamma) = \psi(\gamma)(f) \Rightarrow \psi = \tau \circ \sigma$$

于是命题得证。 $\square$

第二章　线性映射及其矩阵

本章开始，我们转向线性映射的研究。

我们将用三章完成线性映射的研究。本章我们将从映射最基本的研究方式：向量的作用开始

线性映射的独特之处在于：一方面它能在常规的映射加法和纯量乘法下构成一个线性空间；另一方面，如果我们将复合视为乘法，它可以构成一个幺环。

第一节中，我们将给出线性映射的定义及运算，并研究基本性质；

第二节中，我们将研究两种由线性映射导出的子空间，核和像，并借此提出一个概念：秩。它和我们之前的秩也有很强的联系；

第三节到第五节中，我们将研究矩阵，它是将线性映射在基下的作用写成的一张数表，非常便于在数值上研究矩阵；

第六节中，我们将研究行列式，它是一个反对称多线性函数，我们以此为工具，为后续我们对线性映射分解的研究铺垫。

2.1　线性映射的定义和运算

2.1.1　线性映射的定义

我们首先给出线性映射的定义

定义 2.1.1: 线性映射

设 V_1, V_2 是一个 F 上的两个线性空间,，映射 $A : V_1 \to V_2$ 如果满足：

$\forall \alpha, \beta \in V_1, k \in F$

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

$A(k\alpha) = kA(\alpha)$

那么我们称 A 是一个从 V_1 到 V_2 的线性映射

全体 V_1 到 V_2 的线性映射的集合记作 $\hom_F(V_1, V_2)$，或简记作 $\hom(V_1, V_2)$

特别地，如果 $V_1 = V_2$，我们称 A 是一个 V_1 上的线性变换

有一些常用的线性映射，我们在这里列出来：

定义 2.1.2: 一些常用的线性映射

1. 恒等变换：$I : V \ni \alpha \mapsto \alpha \in V$
2. 数乘变换：$k : V \ni \alpha \mapsto k\alpha \in V$
3. 零变换：$0 : V_1 \ni \alpha \mapsto \mathbf{0}_{V_2} \in V_2$

2.1.2 线性映射的运算

前面我们定义了线性映射，现在我们开始赋予 $\hom(V_1, V_2)$ 线性空间和环的性质。

我们会定义三种运算：加法、纯量乘法、乘法

定义 2.1.3: 线性映射的运算

我们定义：

映射 $+: \hom(V_1, V_2) \times \hom(V_1, V_2) \to \hom(V_1, V_2)$，称为加法，如果满足：

$$\forall A, B \in \hom(V_1, V_2), \alpha \in V_1, (A+B)(\alpha) = A(\alpha) + B(\alpha) \tag{2.1}$$

映射 $\cdot: F \times \hom(V_1, V_2) \to \hom(V_1, V_2)$，称为纯量乘法，如果满足：

$$\forall k \in F, A \in \hom(V_1, V_2), \alpha \in F, (k \cdot A)(\alpha) = kA(\alpha) \tag{2.2}$$

映射 $\circ: \hom(V_1, V_2) \times \hom(V_1, V_2) \to \hom(V_1, V_2)$，称为乘法，如果满足：

$$\forall A, B \in \hom(V_1, V_2), \alpha \in V_1, (A \circ B)(\alpha) = A(B(\alpha)) \tag{2.3}$$

我们也常常把 $k \cdot A$ 简记为 kA，将 $A \circ B$ 简记为 AB

显然，$(\hom(V_1, V_2), F, +, \cdot)$ 是一个线性空间，0 是它的零向量；

$(\hom(V_1, V_2), +, \circ)$ 是一个幺环，0 是它的加法单位元，I 是它的乘法单位元

除此之外，还有一些运算，但是它们是针对特殊的线性映射的，比如说：

定义 2.1.4: 线性变换的幂

$\forall A \in \hom(V, V)$

我们定义：$A^m := \begin{cases} A \circ A^{m-1}, m \geqslant 1 \\ I, m = 0 \end{cases} \quad, m \geqslant 0$

如果一个线性变换的幂不会使其本身变化，我们称它是一个幂等变换

定义 2.1.5: 幂等映射

$A \in \hom(V, V)$ 如果有：

$A = A^2$

我们称它是一个幂等变换

如果一个线性变换，存在一个自然数，使其在这个自然数次幂下为零变换，那么我们称它是一个幂零变换

定义 2.1.6: 幂零变换

$A \in \hom(V, V)$ 如果有：

$\exists l \in \mathbb{N}_+, A^l = 0$，且 $\forall n \in N_+, n < l, A^n \neq 0$

那么我们称 A 是一个幂零指数为 l 的幂零变换

我们不再讨论其他的运算，我们接下来转入线性映射一般性质的研究

2.1.3 线性映射的性质

1.

命题 2.1.1. $\forall A \in \hom(V_1, V_2), A(\mathbf{0}_{V_1}) = \mathbf{0}_{V_2}$

证明： $A(\mathbf{0}_{V_1}) = A(0 \cdot \mathbf{0}_{V_1}) = 0 \cdot A(\mathbf{0}_{V_1}) = \mathbf{0}_{V_2}$ □

2.

命题 2.1.2. $\forall A \in \hom(V_1, V_2), A(-\alpha) = -A(\alpha)$

证明： $A(-\alpha) = A((-1) \cdot \alpha) = (-1) \cdot A(\alpha) = -A(\alpha)$ □

3.

命题 2.1.3. $\forall A \in \hom(V_1, V_2), A(\sum\limits_{i=1}^{n} k_i\alpha_i) = \sum\limits_{i=1}^{n} k_i A(\alpha_i)$

证明： 对 n 使用数学归纳法易证。 □

值得注意，这个定理并不能随意地推广到 $\aleph_0$，因为此时依赖于度量线性空间或线性映射的进一步性质。

4.

命题 2.1.4. $\forall A \in \hom(V, V), m \geqslant 1, A^m = A^{m-1} \circ A$

证明： 对 m 作数学归纳法。

首先，当 $m = 2$ 时，$A^2 = A \circ A$，命题成立

现在假设 m 时成立，我们来证明 $m + 1$ 时命题也成立：

$A^m = A \circ A^{m-1} = A \circ A^{m-2} \circ A = A^{m-1} \circ A$，于是命题得证。 □

这个命题看似显然，但是是必要的，因为线性映射环不交换。这个命题指出：递归式地推导幂时，从左右方向都是等价的。进一步，在递归中不断变换方向也不会影响结果。

5.

命题 2.1.5. $\forall A \in \hom(V_1, V_2)$

如果 $\alpha_1, \cdots, \alpha_s$ 线性相关，那么 $A(\alpha_1), \cdots, A(\alpha_s)$ 也线性相关

证明： $\alpha_1, \cdots, \alpha_s$ 线性相关

$\Rightarrow \exists k_1, \cdots, k_s, k_1\alpha_1 + \cdots + k_n\alpha_s = \mathbf{0}$，其中 $k_1, \cdots, k_s$ 不全为零

$\Rightarrow A(k_1\alpha_1 + \cdots + k_n\alpha_s) = k_1A(\alpha_1) + \cdots + k_nA(\alpha_s) = \mathbf{0}$

$\Rightarrow A(\alpha_1), \cdots, A(\alpha_n)$ 线性无关 □

值得注意的是，不同于同构映射，在这个命题中把线性相关改为线性无关会使命题变得不成立。比如说，零映射会把任何线性无关的向量组变得线性相关

6.

命题 2.1.6. $\forall T, Y \in \hom(V_1, V_2)$，$B$ 是 V_1 的一个基

如果 $\forall \alpha \in B, T(\alpha) = Y(\beta)$，那么 $T = Y$

证明： 只需证明：$\forall \gamma \in V_1, T(\gamma) = Y(\gamma)$

因为 B 是 V_1 的基，所以一定有 $\gamma = k_1\alpha_1 + \cdots + k_s\alpha_s, \alpha_1, \cdots, \alpha_s \in B$

此时，$T(\gamma) = k_1T(\alpha_1) + \cdots + k_sT(\alpha_s) = k_1Y(\alpha_1) + \cdots + k_sY(\alpha_s) = Y(\gamma)$，于是命题得证。 □

这个命题指出，线性映射完全由其在基上的作用决定，因为我们其实只需要指定基上的像就指定了线性映射本身。

最后，还有一个问题未被解决：$\hom(V, W)$ 的维数和 $\dim V, \dim W$ 有什么关系。这个问题比较复杂，因为涉及无限维时，维数的公式会变得和有限情形完全不同。在讲解行列式后，我们将在附录一中解决这个问题。

2.2 线性映射的核和像

本节中，我们将借助核与像继续研究线性映射。所谓核，即是线性映射映到零的那一部分；像则是值域。

同时，我们会引入对偶映射，借助本节中核与像的工具，我们将看到对偶和本身之间的联系。

2.2.1 核与像的定义

定义 2.2.1: 线性映射的核

设 $A \in \hom(V_1, V_2)$，我们定义：

$$\ker A := \{\alpha \in V_1 | A(\alpha) = \mathbf{0}_{V_2}\} \tag{2.4}$$

称为线性映射 A 的核

定义 2.2.2: 线性映射的像

设 $A \in \hom(V_1, V_2)$，我们定义：

$$\operatorname{Im} A := A(V_1) := \{A(\alpha) | \alpha \in V_1\} \tag{2.5}$$

.

称为线性映射 A 的像

特别地，核与像的维数我们分别称为零化度和秩：

定义 2.2.3: 线性映射的零化度

设 $A \in \hom(V_1, V_2)$，我们定义：

$$\operatorname{nullity}(A) := \dim(\ker A) \tag{2.6}$$

称为线性映射 A 的零化度

定义 2.2.4: 线性映射的秩

设 $A \in \hom(V_1, V_2)$，我们定义：

$$\operatorname{rank}(A) := \dim(\operatorname{Im} A) \tag{2.7}$$

称为线性映射 A 的秩

事实上，线性映射的秩和之前我们曾提及的向量组的秩有着很大的联系，我们将在后续看到这一点。

为了方便后续性质的研究，接下来我们给出对偶映射的概念

定义 2.2.5: 对偶映射

设 $T \in \hom(V_1, V_2)$，我们定义对偶映射 $T^* \in \hom(V_2^*, V_1^*)$

如果满足：$\forall f \in V_2^*, \alpha \in V_1$

$$(T^*(f))(\alpha) = (f \circ T)(\alpha) \tag{2.8}$$

2.2.2 核与像的性质

接下来研究核与像的性质

1.

命题 2.2.1. $\forall A \in \hom(V, W), \ker A, \operatorname{Im} A$ 都是线性空间

证明： 先证明 $\ker A$ 是一个线性空间

注意到，$\mathbf{0}_V \in \ker A$，因为 $A(\mathbf{0}_V) = \mathbf{0}_W$，因此 $\ker A$ 非空

那么，只需注意到 $\forall \alpha, \beta \in \ker A, A(\alpha + \beta) = A(\alpha) + A(\beta) = \mathbf{0}_W \Rightarrow \alpha + \beta \in \ker A$

$\forall \alpha \in \ker A, k \in F, A(k\alpha) = kA(\alpha) = \mathbf{0}_W \Rightarrow k\alpha \in \ker A$

再证明 $\operatorname{Im} A$ 是一个线性空间

注意到，$\mathbf{0}_W = A(\mathbf{0}_V) \in \operatorname{Im} A \Rightarrow \operatorname{Im} A \neq \emptyset$

那么，只需注意到 $\forall A(\alpha), A(\beta) \in \operatorname{Im} A, A(\alpha) + A(\beta) = A(\alpha + \beta) \in \operatorname{Im} A$

$\forall A(\alpha) \in \operatorname{Im} A, k \in F, kA(\alpha) = A(k\alpha) \in \operatorname{Im} A$

于是命题得证 □

2.

命题 2.2.2. $A \in \hom(V, W)$，那么：

A 是单射 $\Leftrightarrow \ker A = \{\mathbf{0}_V\}$

证明： 先证明充分性。注意到，$A(\mathbf{0}_V) = \mathbf{0}_W$，而 A 为单射，于是一定有 $\ker A = \{\mathbf{0}_V\}$

再证明必要性。设 $A(\alpha) = A(\beta)$，那么 $A(\alpha - \beta) = \mathbf{0}_W \Rightarrow \alpha - \beta = \mathbf{0}_V \Rightarrow \alpha = \beta \Rightarrow A$ 是单射 □

3.

命题 2.2.3. $A \in \hom(V, W)$，那么：

A 是单射 $\Leftrightarrow \operatorname{Im} A = W$

证明： 这是显然的。 □

4. 秩-零化度定理 在给出定理前，我们先给出一个引理。它是一个线性空间的“同态基本定理”，证明方法也很相似

引理 2.2.1

$\forall A \in \hom(V, W)$

$$V / \ker A \cong \operatorname{Im} A \tag{2.9}$$

证明： 设 $\phi : V / \ker A \ni \alpha + \ker A \mapsto A(\alpha) \in \operatorname{Im} A$

首先验证它的确是一个映射

假设 $\alpha + \ker A = \beta + \ker A$

那么 $\alpha - \beta \in \ker A \Rightarrow A(\alpha - \beta) = \mathbf{0}_W \Rightarrow A(\alpha) = A(\beta)$

所以 ϕ 的确是一个映射

接下来验证它是一个单射

设 $\phi(\alpha + \ker A) = \phi(\beta + \ker A)$

$\Rightarrow A(\alpha) = A(\beta) \Rightarrow A(\alpha - \beta) = \mathbf{0}_W$

$\Rightarrow \alpha - \beta \in \ker A \Rightarrow \alpha + \ker A = \beta + \ker A$

于是 ϕ 是单射。ϕ 显然是满射

接下来验证线性性：

$\forall \alpha + \ker A, \beta + \ker A \in V / \ker A, k \in F$

$\phi\left((\alpha + \ker A) + (\beta + \ker A)\right) = \phi\left((\alpha + \beta) + \ker A\right) = A(\alpha + \beta) = A(\alpha) + A(\beta) = \phi(\alpha + \ker A) + \phi(\beta + \ker A)$

$\phi\left(k \cdot (\alpha + \ker A)\right) = \phi(k\alpha + \ker A) = A(k\alpha) = kA(\alpha) = k\phi(\alpha + \ker A)$

于是命题得证。 □

它的直接推论是被称为秩-零化度定理的结论，它揭示了秩和零化度的联系

定理 2.2.2: 秩-零化度定理

$\forall A \in \hom(V, W)$

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = \dim V \tag{2.10}$$

证明： $V / \ker A \cong \operatorname{Im} A$

$\Rightarrow \dim(V / \ker A) = \dim \operatorname{Im} A$

$\Rightarrow \dim V - \dim \ker A = \dim \operatorname{Im} A$

$\Rightarrow \operatorname{rank}(A) + \operatorname{nullity}(A) = \dim V$ □

5. 有限维映射和其对偶的秩相同 我们先证明一个引理：

引理 2.2.3: 子空间及其零化子维数和为原空间维数

设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$，U 是 V 的一个线性子空间

我们定义 $U' = \{A \in V^* | \forall \alpha \in U, A(\alpha) = 0\}$

那么有：$\dim U + \dim U' = \dim V$

证明： 设 $\dim U = m, 0 \leqslant m \leqslant n$

取 U 的一个基 $\{\alpha_1, \cdots, \alpha_m\}$，并补全为 V 的一个基 $\{\alpha_1, \cdots, \alpha_n\}$

设 $\{\alpha_1, \cdots, \alpha_n\}$ 的对偶基为 $\{f_1, \cdots, f_n\}$

我们来证明：$\{f_{m+1}, \cdots, f_n\}$ 是 U' 的一个基

取 $\forall A \in U'$。因为 $U' \subseteq V^*$，那么 A 一定可以被对偶基线性表出

不妨设 $A = \sum\limits_{i=1}^{n} k_i f_i$

注意到，当 $1 \leqslant j \leqslant m, A(\alpha_j) = \sum\limits_{i=1}^{n} k_i f_i(\alpha_j) = k_j$

但因为 $\alpha_j \in U$，因此 $k_j = A(\alpha_j) = 0$

因此有 $A = \sum\limits_{i=m+1}^{n} k_i f_i$

这说明 U' 可由 $\{f_{m+1}, \cdot, f_n\}$ 线性表出。又因为 $\{f_1, \cdots, f_n\}$ 线性无关，因此 $\{f_{m+1}, \cdot, f_n\}$ 也线性无关

因此 $\{f_{m+1}, \cdot, f_n\}$ 是 U' 的一个基，那么 $\dim U' = n - m$，命题得证。 □

接下来给出命题

命题 2.2.4. 设 $T \in \hom(V, W), \dim V = m < \aleph_0, \dim W = n < \aleph_0$

那么 $\operatorname{rank}(T) = \operatorname{rank}(T^*)$

证明： 只需注意到，$\ker T^* = \{f \in W^* | T^*(f) = \mathbf{0}_{V^*}\}$

$= \{f \in W^* | f \circ T = \mathbf{0}_{V^*}\}$

$= \{f \in W^* | \forall \alpha \in V, f(T(\alpha)) = 0\}$

$= \{f \in W^* | \forall \beta \in \operatorname{Im} T, f(\beta) = 0\}$

而按照引理，有：$\dim \operatorname{Im} T + \dim\{f \in W^* | \forall \beta \in \operatorname{Im} T, f(\beta) = 0\} = \dim W$

$\Rightarrow \dim \operatorname{Im} T + \dim \ker T^* = \dim W$

再运用秩-零化度定理，有：$\dim \operatorname{Im} T^* + \dim \ker T^* = \dim W^* = \dim W$

因此有 $\dim \operatorname{Im} T = \dim \operatorname{Im} T^* \Rightarrow \operatorname{rank}(T) = \operatorname{rank}(T^*)$，命题得证 □

6.

命题 2.2.5. 设 $A \in \hom(V, W)$，如果 $\dim V = \dim W$，那么：

A 是满射 $\Leftrightarrow$ A 是单射

证明： 因为 $V/\ker A \cong \operatorname{Im} A$

A 是满射 $\Leftrightarrow \operatorname{Im} A = W \Leftrightarrow \operatorname{Im} A \cong W \Leftrightarrow \operatorname{Im} A \cong W \cong V$

$\Leftrightarrow V \cong V/\ker A \Leftrightarrow \dim \ker A = 0 \Leftrightarrow \ker A = \{\mathbf{0}_V\} \Leftrightarrow A$ 是单射 □

7.

命题 2.2.6. $A \in \hom(V, V), A$ 可逆 $\Leftrightarrow \operatorname{rank}(A) = \dim V$

证明： A 可逆 $\Leftrightarrow A$ 是双射 $\Leftrightarrow \operatorname{Im} A = V \Leftrightarrow \operatorname{rank}(A) = \dim V$ □

在这里，我们利用了因为线性变换的定义域和陪域一定同维，所以其实双射就是满射的性质

事实上，容易发现线性映射可逆的必要条件是它的秩等于配域的维数

2.3 矩阵

为了方便后续的研究，我们引入一种“数表”，即矩阵。它和有限维空间上的线性映射完全同构

2.3.1 矩阵的定义

定义 2.3.1: 矩阵

形如以下的矩形阵列称为一个域 F 上的矩阵 $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, a_{ij} \in F$

简记为 $(a_{ij})_{m\times n}$ 或 (a_{ij})。m 称为矩阵的行数，n 称为矩阵的列数。

特别地，如果 $m = n$，我们称它是一个 m 阶方阵。

F 上的全体 $m \times n$ 矩阵的集合记作 $M_{m\times n}(F)$，特别地如果 $m = n$，记作 $M_n(F)$。

我们也将矩阵 $\boldsymbol{A}$ 在 m 行 n 列处的元素记作 $\boldsymbol{A}_{ij}$

我们也常常将 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 记作 $(\alpha_1,\cdots,\alpha_n)$ 或 $\begin{pmatrix}\beta_1\\ \vdots\\ \beta_m\end{pmatrix}$，其中 $\alpha_i=\begin{pmatrix}a_{1i}\\ \vdots\\ a_{mi}\end{pmatrix}, \beta_j=(a_{j1},\cdots,a_{jn})$ 可以看出来，矩阵也是可以视为一个向量组，因此，我们也可以对其定义秩：

定义 2.3.2: 矩阵的秩

设 $\boldsymbol{A}=(\alpha_1,\cdots,\alpha_n)\in M_{m\times n}(F)$

我们定义：

$$\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\alpha_1,\cdots,\alpha_n) \tag{2.11}$$

称为矩阵的秩

后续我们会看到，矩阵秩、线性映射秩、向量组秩其实是统一的

2.3.2 矩阵的运算

1. 相等

定义 2.3.3: 矩阵的相等

设 $\boldsymbol{A}\in M_{m\times n}(F),\boldsymbol{B}\in M_{m\times n}(F)$，如果 $\forall i,j,\boldsymbol{A}(i;j)=\boldsymbol{B}(i;j)$，则称 $\boldsymbol{A}=\boldsymbol{B}$。

2. 转置

定义 2.3.4: 矩阵的转置

设 $\boldsymbol{A}\in M_{m\times n}(F)$，

我们定义矩阵 $\boldsymbol{A}^T\in M_{n\times m}(F)$ 为满足 $\boldsymbol{A}^T(i;j)=\boldsymbol{A}(j;i)$ 的矩阵，称为 $\boldsymbol{A}$ 的转置。

3. 加法

定义 2.3.5: 矩阵的加法

设 $\boldsymbol{A}\in M_{m\times n}(F),\boldsymbol{B}\in M_{m\times n}(F)$

我们定义：$(\boldsymbol{A}+\boldsymbol{B})(i;j)=\boldsymbol{A}(i;j)+\boldsymbol{B}(i;j)$。

4. 纯量乘法

定义 2.3.6: 矩阵的纯量乘法

设 $\boldsymbol{A} \in M_{m\times n}(F), k \in F$，

我们定义矩阵 $k\cdot\boldsymbol{A} \in M_{m\times n}(F)$ 为满足 $(k\cdot\boldsymbol{A})(i;j) = k\cdot\boldsymbol{A}(i;j)$ 的矩阵。

5. 乘法

定义 2.3.7: 矩阵的乘法

设 $\boldsymbol{A} \in M_{m\times n}(F), \boldsymbol{B} \in M_{n\times p}(F)$，

我们定义矩阵 $\boldsymbol{A}\cdot\boldsymbol{B} \in M_{m\times p}(F)$ 为满足 $(\boldsymbol{A}\cdot\boldsymbol{B})(i;j) = \sum_{k=1}^{n}\boldsymbol{A}(i;k)\boldsymbol{B}(k;j)$ 的矩阵

6. 幂

定义 2.3.8: 方阵的幂

$\forall \boldsymbol{A} \in M_n(F)$

我们定义：$\boldsymbol{A}^m := \begin{cases} \boldsymbol{A}\cdot\boldsymbol{A}^{m-1}, m \geqslant 1 \\ \boldsymbol{I}, m = 0 \end{cases}, m \geqslant 0$

7. 和行向量的乘法

最后，为了后续研究线性映射，我们再定义一种新运算

定义 2.3.9: V^n 中的行向量和矩阵的乘积

设 $(\alpha_1, \cdots, \alpha_m) \in V^m, \boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m\times n}(F)$，

我们定义：

$(\alpha_1, \cdots, \alpha_m)\cdot\boldsymbol{A} = (\sum_{i=1}^{m} a_{i1}\alpha_i, \cdots, \sum_{i=1}^{m} a_{im}\alpha_i)$

值得注意的是，这个运算并不是矩阵乘法的退化，因为此处乘法的左元中的向量是一个一般的向量，而不一定是列向量

2.3.3 线性映射的矩阵

给出矩阵及其运算的基本定义后，我们可以开始讨论如何将线性映射写成矩阵了。

定义 2.3.10: 线性映射的矩阵

设 $A \in \hom(V,W), \dim V = m < \aleph_0, \dim W = n < \aleph_0$

取 V 的一个基 $\{\alpha_1,\cdots,\alpha_m\}$，$W$ 的一个基 $\{\beta_1,\cdots,\beta_n\}$

如果矩阵 $\boldsymbol{A} = (a_{ij}) \in M_{m\times n}(F)$ 满足：

$$(A(\alpha_1),\cdots,A(\alpha_m)) = (\beta_1,\cdots,\beta_n)\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \tag{2.12}$$

那么我们称 $\boldsymbol{A}$ 是线性映射 A 在 $\{\alpha_1,\cdots,\alpha_m\}$ 和 $\{\beta_1,\cdots,\beta_n\}$ 下的矩阵

特别地，如果 $V = W$ 且 $\alpha_i = \beta_i$，我们称 $\boldsymbol{A}$ 是线性变换 A 在 $\{\alpha_1,\cdots,\alpha_m\}$ 下的矩阵

我们也常常简记 $(A(\alpha_1),\cdots,A(\alpha_m))$ 为 $A(\alpha_1,\cdots,\alpha_m)$

我们习惯上用大写字母表示线性映射，而用加粗的大写字母表示它在某个基下的矩阵

我们接下来会看到，矩阵的运算其实就是线性映射的运算

2.3.4 矩阵的性质

1.

命题 2.3.1. $(\alpha_1,\cdots,\alpha_m)(\boldsymbol{A}+\boldsymbol{B}) = (\alpha_1,\cdots,\alpha_m)\boldsymbol{A} + (\alpha_1,\cdots,\alpha_m)\boldsymbol{B}$

$(\alpha_1,\cdots,\alpha_m)(k\boldsymbol{A}) = k\left((\alpha_1,\cdots,\alpha_m)\boldsymbol{A}\right)$

$(\alpha_1,\cdots,\alpha_m)(\boldsymbol{AB}) = \left((\alpha_1,\cdots,\alpha_m)\boldsymbol{A}\right)\boldsymbol{B}$

证明： $(\alpha_1,\cdots,\alpha_m)(\boldsymbol{A}+\boldsymbol{B}) = (\sum_{i=1}^m(\boldsymbol{A}+\boldsymbol{B})(i;1)\alpha_i,\cdots,\sum_{i=1}^m(\boldsymbol{A}+\boldsymbol{B})(i;m)\alpha_i)$

$= (\sum_{i=1}^m \boldsymbol{A}(i;1)\alpha_i,\cdots,\sum_{i=1}^m \boldsymbol{A}(i;m)\alpha_i) + (\sum_{i=1}^m \boldsymbol{B}(i;1)\alpha_i,\cdots,\sum_{i=1}^m \boldsymbol{B}(i;m)\alpha_i)$

$= (\alpha_1,\cdots,\alpha_m)\boldsymbol{A} + (\alpha_1,\cdots,\alpha_m)\boldsymbol{B}$

$(\alpha_1,\cdots,\alpha_m)(k\boldsymbol{A}) = (\sum_{i=1}^m(k\boldsymbol{A})(i;1)\alpha_i,\cdots,\sum_{i=1}^m(k\boldsymbol{A})(i;m)\alpha_i)$

$= k(\sum_{i=1}^m \boldsymbol{A}(i;1)\alpha_i,\cdots,\sum_{i=1}^m \boldsymbol{A}(i;m)\alpha_i)$

$= k\left((\alpha_1,\cdots,\alpha_m)\boldsymbol{A}\right)$

设 $\boldsymbol{A} \in M_{m\times n}(F), \boldsymbol{B} \in M_{n\times s}(F)$

$(\alpha_1,\cdots,\alpha_m)(\boldsymbol{AB})=(\sum_{i=1}^m(\boldsymbol{AB})(i;1)\alpha_i,\cdots,\sum_{i=1}^m(\boldsymbol{AB})(i;m)\alpha_i)$

$=(\sum_{i=1}^m(\sum\limits_{j=1}^n\boldsymbol{A}(i;j)\boldsymbol{B}(j;1))\alpha_i,\cdots,\sum_{i=1}^m(\sum\limits_{j=1}^n\boldsymbol{A}(i;j)\boldsymbol{B}(j;m))\alpha_i)$

$=(\sum\limits_{j=1}^n\boldsymbol{B}(j;1)(\sum_{i=1}^m\boldsymbol{A}(i;j)\alpha_i),\cdots,\sum\limits_{j=1}^n\boldsymbol{B}(j;m)(\sum_{i=1}^m\boldsymbol{A}(i;j)\alpha_i))$

$=(\sum_{i=1}^m\boldsymbol{A}(i;1)\alpha_i,\cdots,\sum_{i=1}^m\boldsymbol{A}(i;m)\alpha_i)\boldsymbol{B}$

$=((\alpha_1,\cdots,\alpha_m)\boldsymbol{A})\,\boldsymbol{B}$ □

2.

命题 2.3.2. 设 $A,B\in\hom(V,W)$，其中 V,W 是 F 上的有限维线性空间

分别取 V,W 的一个基 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$

设 A,B 在 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$ 下的矩阵分别为 $\boldsymbol{A},\boldsymbol{B}$

那么：$A+B$ 在 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$ 下的矩阵为 $\boldsymbol{A}+\boldsymbol{B}$

证明： 只需证明：$(A+B)(\alpha_1,\cdots,\alpha_m)=(\beta_1,\cdots,\beta_n)(\boldsymbol{A}+\boldsymbol{B})$

注意到，$(A+B)(\alpha_1,\cdots,\alpha_m)=A(\alpha_1,\cdots,\alpha_m)+B(\alpha_1,\cdots,\alpha_m)$

$=(\beta_1,\cdots,\beta_n)\boldsymbol{A}+(\beta_1,\cdots,\beta_n)\boldsymbol{B}=(\beta_1,\cdots,\beta_n)(\boldsymbol{A}+\boldsymbol{B})$

命题得证。 □

3.

命题 2.3.3. 设 $A\in\hom(V,W),k\in F$，其中 V,W 是 F 上的有限维线性空间

分别取 V,W 的一个基 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$

设 A 在 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$ 下的矩阵为 $\boldsymbol{A}$

那么：kA 在 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$ 下的矩阵为 $k\boldsymbol{A}$

证明： 只需证明：$(kA)(\alpha_1,\cdots,\alpha_m)=(\beta_1,\cdots,\beta_n)(k\boldsymbol{A})$

注意到，$(kA)(\alpha_1,\cdots,\alpha_m)=k\,(A(\alpha_1,\cdots,\alpha_m))$

$=k\,((\beta_1,\cdots,\beta_n)\boldsymbol{A})=(\beta_1,\cdots,\beta_n)(k\boldsymbol{A})$

命题得证。 □

4.

命题 2.3.4. 设 $A\in\hom(W,U),B\in\hom(V,W)$，其中 V,W 是 F 上的有限维线性空间

分别取 V,W,U 的一个基 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\},\{\gamma_1,\cdots,\gamma_k\}$

设 A 在 $\{\beta_1,\cdots,\beta_n\},\{\gamma_1,\cdots,\gamma_k\}$ 下的矩阵为 $\boldsymbol{A}$

B 在 $\{\alpha_1,\cdots,\alpha_m\},\{\beta_1,\cdots,\beta_n\}$ 下的矩阵为 $\boldsymbol{B}$

那么：AB 在 $\{\alpha_1,\cdots,\alpha_m\},\{\gamma_1,\cdots,\gamma_k\}$ 下的矩阵为 $\boldsymbol{AB}$

证明： 只需证明：$(AB)(\alpha_1,\cdots,\alpha_m)=(\gamma_1,\cdots,\gamma_k)(\boldsymbol{AB})$

注意到，$(AB)(\alpha_1,\cdots,\alpha_m)=A\left(B(\alpha_1,\cdots,\alpha_m)\right)$

$=A((\beta_1,\cdots,\beta_n)\boldsymbol{B})=A(\sum\limits_{i=1}^{n}\boldsymbol{B}(i,1)\beta_i,\cdots,\sum\limits_{i=1}^{n}\boldsymbol{B}(i,m)\beta_i)$

$=(\sum\limits_{i=1}^{n}\boldsymbol{B}(i,1)A(\beta_i),\cdots,\sum\limits_{i=1}^{n}\boldsymbol{B}(i,m)A(\beta_i))=(A(\beta_1),\cdots,A(\beta_n))\,\boldsymbol{B}$

$((\gamma_1,\cdots,\gamma_k)\boldsymbol{A})\,\boldsymbol{B}=((\gamma_1,\cdots,\gamma_k))\,(\boldsymbol{AB})$

命题得证。 □

至此，我们证明了有限维到有限维的线性映射的运算和矩阵的运算别无二致

5.

命题 2.3.5. $\dim_F M_{m\times n}(F)=mn$

证明： 设 $\boldsymbol{E_{ij}}\in M_{m\times n}(F)$ 满足：

$\boldsymbol{E_{ij}}(k;l)=\delta_{ik}\delta_{jk}$

显然 $\{\boldsymbol{E}_{ij}\}$ 线性无关

那么，只需注意到：

$\forall\boldsymbol{A}\in M_{m\times n}(F),\boldsymbol{A}=\sum\limits_{i=1}^{m}\left(\sum\limits_{j=1}^{n}A(i;j)\boldsymbol{J_{ij}}\right)$

于是 $\dim_F M_{m\times n}(F)=\mathrm{Card}\{\boldsymbol{J}_{ij}\}=mn$，命题得证。 □

6.

命题 2.3.6. 设 V,W 是 F 上的线性空间，$\dim V=n<\aleph_0,\dim W=m<\aleph_0$

那么：$\hom(V,W)\cong M_{m\times n}(F)$

证明： 注意到：$\dim hom(V,W)=mn$

依照前面的命题和同构的性质，命题得证。 □

此命题中我们使用了有关线性映射空间的维数的结论，证明参见附录。

7. 矩阵乘法的交换律

命题 2.3.7. 设 $\boldsymbol{A}\in M_{m\times n}(F),\boldsymbol{B}\in M_{n\times k}(F),\boldsymbol{C}\in M_{k\times s}(F)$

那么有：$\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}$

命题 2.3.8. 分别任取 F 上的线性空间 V,W,U,H，使得 $\dim V=n,\dim W=m,\dim U=k,\dim H=s$

依照前面的命题，一定能找到 $A \in \hom(V, W), B \in \hom(U, V), C \in \hom(H, U)$，使得 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 分别是它们的矩阵

注意到，$\boldsymbol{A}(\boldsymbol{BC}), (\boldsymbol{AB})\boldsymbol{C}$ 分别是 $A(BC), (AB)C$ 的矩阵。

而因为 $A(BC) = (AB)C$，所以一定有 $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$，命题得证。

8. 矩阵乘法的分配律

命题 2.3.9. 如果 $\boldsymbol{A} \in M_{m\times n}(F), \boldsymbol{B} \in M_{n\times k}(F), \boldsymbol{C} \in M_{n\times k}(F)$

那么有：$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$；

如果 $\boldsymbol{A} \in M_{n\times k}(F), \boldsymbol{B} \in M_{n\times k}(F), \boldsymbol{C} \in M_{k\times s}(F)$

那么有：$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$

命题 2.3.10. 分别任取 F 上的线性空间 V, W, U, H，使得 $\dim V = n, \dim W = m, \dim U = k, \dim H = s$

先证左分配律：

依照前面的命题，一定能找到 $A \in \hom(V, W), B \in \hom(U, V), C \in \hom(U, V)$，使得 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 分别是它们的矩阵

注意到，$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}), \boldsymbol{AB} + \boldsymbol{AC}$ 分别是 $A(B + C), AB + AC$ 的矩阵。

而因为 $A(B + C) = AB + AC$，所以一定有 $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$

再证右分配律：

依照前面的命题，一定能找到 $A \in \hom(U, V), B \in \hom(U, V), C \in \hom(H, U)$，使得 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 分别是它们的矩阵

注意到，$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C}, \boldsymbol{AC} + \boldsymbol{BC}$ 分别是 $(A + B)C, AC + BC$ 的矩阵。

而因为 $(A + B)C = AC + BC$，所以一定有 $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$

命题得证。

于是，我们证明了 $(M_n(F), +, \cdot)$ 是一个环。事实上，它是一个非交换幺环。

非交换是显然的，我们接下来讨论它的幺元。

9. 单位矩阵的性质

命题 2.3.11. 设 $\boldsymbol{I_n} \in M_n(F)$ 满足：$\boldsymbol{I_n}(i; j) = \delta_{ij}$

那么有：$\forall \boldsymbol{A} \in M_{n\times m}(F), \boldsymbol{I_n}\boldsymbol{A} = \boldsymbol{A}$

$\forall \boldsymbol{A} \in M_{m\times n}(F), \boldsymbol{A}\boldsymbol{I_n} = \boldsymbol{A}$

$\forall \boldsymbol{A} \in M_n(F), \boldsymbol{A}\boldsymbol{I_n} = \boldsymbol{I_n}\boldsymbol{A} = \boldsymbol{A}$

证明: 事实上，以上都是显然的，因为只需注意到：

任取一个 n 维线性空间 V 和它的一个基 $\{\alpha_1,\cdots,\alpha_n\}$

$I(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)\boldsymbol{I_n}$

因此，$\boldsymbol{A}$ 是 I 在任意基上的矩阵，因此上述性质是显然的。 □

上述结论说明：$\boldsymbol{I_n}$ 是 $M_n(F)$ 的乘法单位元，它一般也被称为单位矩阵

在证明过程中，我们还发现：不同于绝大多数线性映射的矩阵和其基的选择有关，恒等变换在任意基下的矩阵都是单位矩阵

我们将在后续用一节专门研究可逆矩阵

10.

命题 2.3.12. $\forall k\in F,\boldsymbol{A}\in M_{n\times m}(F),\boldsymbol{B}\in M_{m\times k}(F)$

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

命题 2.3.13. 随意取三个 F 上的线性空间 V,W,U,使得 $\dim V=m,\dim W=n,\dim U=k$

依照前面的命题，一定能找到 $A\in\hom(V,W),B\in\hom(U,V)$，使得 $\boldsymbol{A},\boldsymbol{B}$ 分别是它们的矩阵

注意到，$k(\boldsymbol{AB}),(k\boldsymbol{A})\boldsymbol{B},\boldsymbol{A}(k\boldsymbol{B})$ 分别是 $k(AB),(kA)B,A(kB)$ 的矩阵。

而因为 $k(AB)=(kA)B=A(kB)$，所以一定有 $k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$

命题得证。

11.

命题 2.3.14. $\boldsymbol{A}^m=\boldsymbol{A}\cdot\boldsymbol{A}^{m-1}=\boldsymbol{A}^{m-1}\cdot\boldsymbol{A},m\geqslant 1$

证明: 只需证明：$\forall A\in\hom(V,V)$，任取 V 的一个基 $\{\alpha_1,\cdots,\alpha_n\}$，如果 $\boldsymbol{A}$ 是 A 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵，那么 $\boldsymbol{A}^m$ 是 A^m 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵

使用数学归纳法。首先，当 $m=1$ 时，结论就假设，显然成立

现在假设 m 时命题成立，于是 $\boldsymbol{A}^{m+1}=A\circ A^m$ 是在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为 $\boldsymbol{A}\cdot\boldsymbol{A}^m=\boldsymbol{A}^{m+1}$

因此结论成立，依照之前我们证明的线性映射性质，命题显然成立。 □

12.

命题 2.3.15. $(\boldsymbol{A}+\boldsymbol{B})^T=\boldsymbol{A}^T+\boldsymbol{B}^T$

证明: $(\boldsymbol{A}+\boldsymbol{B})^T(j;i)=(\boldsymbol{A}+\boldsymbol{B})(i;j)=\boldsymbol{A}(i;j)+\boldsymbol{B}(i;j)=\boldsymbol{A}^T(j;i)+\boldsymbol{B}^T(j;i)$

命题得证 □

13.

命题 2.3.16. $(k\boldsymbol{A})^T=k\boldsymbol{A}^T$

证明: $(k\boldsymbol{A})^T(j;i)=(k\boldsymbol{A})(i;j)=k\boldsymbol{A}(i;j)=k\boldsymbol{A}^T(j;i)$

命题得证 □

14.

命题 2.3.17. 设 V,W 是 F 上的两个线性空间，$\dim V=n<\aleph_0,\dim W=m<\aleph_0$

分别取 V,W 的一个基 $\{\alpha_1,\cdots,\alpha_n\},\{\beta_1,\cdots,\beta_m\}$，设其对偶基是 $\{f_1,\cdots,f_n\},\{g_1,\cdots,g_m\}$

$\forall A\in\hom(V,W)$，如果它在 $\{\alpha_1,\cdots,\alpha_n\},\{\beta_1,\cdots,\beta_m\}$ 下的矩阵是 $\boldsymbol{A}$

那么 $A^*\in\hom(W^*,V^*)$ 在 $\{g_1,\cdots,g_m\},\{f_1,\cdots,f_n\}$ 下的矩阵是 $\boldsymbol{A}^T$

证明: 设 $A^*\in\hom(W^*,V^*)$ 在 $\{g_1,\cdots,g_m\},\{f_1,\cdots,f_n\}$ 下的矩阵是 $\boldsymbol{A}'$

$A^*(g_j)=\sum\limits_{i=1}^{n}\boldsymbol{A}'(i;j)f_i$

$\Rightarrow A^*(g_j)(\alpha_k)=\sum\limits_{i=1}^{n}\boldsymbol{A}'(i;j)f_i(\alpha_k)$

$\Rightarrow (g_jA)(\alpha_k)=\sum\limits_{i=1}^{n}\boldsymbol{A}'(i;j)\delta_{ik}$

$\Rightarrow g_j(\sum\limits_{l=1}^{m}\boldsymbol{A}(l;k)\beta_l)=\boldsymbol{A}'(k;j)$

$\Rightarrow \boldsymbol{A}(j;k)=\boldsymbol{A}'(k;j)$

$\Rightarrow \boldsymbol{A}'=\boldsymbol{A}^T$

命题得证 □

15. 我们先证明一个引理

引理 2.3.1

设 V,W,U 是 F 上的两个线性空间，$\dim V=n<\aleph_0,\dim W=m<\aleph_0,\dim U=k<\aleph_0$

$\forall A\in\hom(V,W),B\in\hom(U,V)$

那么 $(AB)^*=B^*A^*$

证明: $\forall g\in W^*,\gamma\in U$

$(AB)^*(g)(\gamma)=(gAB)(\gamma)=B^*(A^*(g))(\gamma)=(B^*A^*)(g)(\gamma)$

$\Rightarrow (AB)^* = B^*A^*$ □

命题 2.3.18. $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

证明： 任取 F 上的三个线性空间 V, W, U，使得 $\dim V = m < \aleph_0, \dim W = n < \aleph_0, \dim U = k < \aleph_0$

分别取 V, W, U 的一个基 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}, \{\gamma_1, \cdots, \gamma_k\}$

并记其对偶基为 $\{f_1, \cdots, f_m\}, \{g_1, \cdots, g_n\}, \{u_1, \cdots, u_k\}$

显然存在 $A \in \hom(V, W)$，使得它在 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}$ 下的矩阵为 $\boldsymbol{A}$

存在 $B \in \hom(U, V)$，使得它在 $\{\gamma_1, \cdots, \gamma_k\}, \{\alpha_1, \cdots, \alpha_m\}$ 下的矩阵为 $\boldsymbol{B}$

注意到：$(\boldsymbol{AB})^T$ 是 $(AB)^*$ 在 $\{g_1, \cdots, g_n\}, \{u_1, \cdots, u_k\}$ 下的矩阵

$\boldsymbol{B}^T\boldsymbol{A}^T$ 是 B^*A^* 在 $\{g_1, \cdots, g_n\}, \{u_1, \cdots, u_k\}$ 下的矩阵

那么，因为 $(AB)^* = B^*A^*$，所以一定有 $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$ □

16.

命题 2.3.19. 设 V, W 是 F 上的两个线性空间，$\dim V = m < \aleph_0, \dim W = n < \aleph_0$

任取 V, W 的一个基 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}$

$\forall A \in \hom(V, W)$，设 $\boldsymbol{A}$ 是 A 在 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}$ 下的矩阵

那么有：$\operatorname{rank}(A) = \operatorname{rank}(\boldsymbol{A})$

证明： 设 $\boldsymbol{A} = (\gamma_1, \cdots, \gamma_m)$

我们只需证明:$\dim \operatorname{Im}(A) = \dim \operatorname{span}(\gamma_1, \cdots, \gamma_m)$,因为 $\operatorname{rank}(\gamma_1, \cdots, \gamma_m) = \dim \operatorname{span}(\gamma_1, \cdots, \gamma_m)$（参阅命题 1.4.7）

显然，$\operatorname{Im}(A) = \operatorname{span}(A(\alpha_1), \cdots, A(\alpha_m))$

考虑自然同构：$\phi : W \ni \sum\limits_{i=1}^{n} a_i\beta_i \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n$

显然 ϕ 是一个同构映射。

那么，只需注意到：$\forall k_1A(\alpha_1) + \cdots, k_mA(\alpha_m) \in \operatorname{Im}(A)$

$\phi(k_1A(\alpha_1) + \cdots, k_mA(\alpha_m)) = k_1\gamma_1 + \cdots, k_m\gamma_m \in \operatorname{span}(\alpha_1, \cdots, \alpha_m)$

$\forall k_1\gamma_1 + \cdots, k_m\gamma_m \in \operatorname{span}(\alpha_1, \cdots, \alpha_m)$

$\phi^{-1}(k_1\gamma_1 + \cdots, k_m\gamma_m) = k_1A(\alpha_1) + \cdots, k_mA(\alpha_m) \in \operatorname{Im}(A)$

因此，$\operatorname{Im}(A) \cong \operatorname{span}(\gamma_1, \cdots, \gamma_m)$，于是命题得证 □

至此，我们证明了，矩阵、向量组、线性映射的秩是统一的。

17. 矩阵行秩等于列秩

命题 2.3.20. 设 $\boldsymbol{A} = (\alpha_1, \cdots, \alpha_n) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$ 那么，$\operatorname{rank}(\alpha_1, \cdots, \alpha_n) = \operatorname{rank}(\beta_1, \cdots, \beta_m)$

证明: 设 V, W 是域 F 上的两个线性空间，$\dim V = n < \aleph_0, \dim W = m < \aleph_0$

分别任取 V, W 的一个基 $\{\gamma_1, \cdots, \gamma_n\}, \{\delta_1, \cdots, \delta_m\}$

显然存在 $A \in \hom(V, W)$，使得它在 $\{\gamma_1, \cdots, \gamma_n\}, \{\delta_1, \cdots, \delta_m\}$ 下的矩阵是 $\boldsymbol{A}$

于是有:$\operatorname{rank}(\alpha_1, \cdots, \alpha_n) = \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(A) = \operatorname{rank}(A^*) = \operatorname{rank}(\boldsymbol{A}^T) = \operatorname{rank}(\beta_1, \cdots, \beta_m)$

于是命题得证。 □

18.

命题 2.3.21. 设 $\boldsymbol{A} \in M_{m\times n}(F), \boldsymbol{B} \in M_{n\times k}(F)$

$\operatorname{rank}(\boldsymbol{AB}) \leqslant \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}$

证明: 先证明 $\operatorname{rank}(\boldsymbol{AB}) \leqslant \operatorname{rank}(\boldsymbol{A})$

设 V, W, U 是 F 上的三个线性空间,$\dim V = m < \aleph_0, \dim W = n < \aleph_0, \dim U = k < \aleph_0$

任取 V, W, U 的一个基 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}, \{\gamma_1, \cdots, \gamma_k\}$

显然存在 $A \in \hom(V, W), B \in \hom(U, V)$，使得 A 在 $\{\alpha_1, \cdots, \alpha_m\}, \{\beta_1, \cdots, \beta_n\}$ 下的矩阵是 $\boldsymbol{A}$，B 在 $\{\gamma_1, \cdots, \gamma_k\}, \{\alpha_1, \cdots, \alpha_m\}$ 下的矩阵为 $\boldsymbol{B}$

于是，$\operatorname{rank}(\boldsymbol{AB}) = \operatorname{rank}(AB) = \dim \operatorname{Im}(AB)$

注意到：$\operatorname{Im}(AB) = \{A(B(\alpha)) | \alpha \in U\} \subseteq \{A(\beta) | \beta \in V\} = \operatorname{Im}(A)$

$\Rightarrow \dim \operatorname{Im}(AB) \leqslant \dim \operatorname{Im}(A) \Rightarrow \operatorname{rank}(AB) \leqslant \operatorname{rank}(A) \Rightarrow \operatorname{rank}(\boldsymbol{AB}) \leqslant \operatorname{rank}(\boldsymbol{A})$

又注意到：$\operatorname{rank}(\boldsymbol{AB}) = \operatorname{rank}\left((\boldsymbol{AB})^T\right) = \operatorname{rank}(\boldsymbol{B}^T\boldsymbol{A}^T) \leqslant \operatorname{rank}(\boldsymbol{B}^T) = \operatorname{rank}(\boldsymbol{B})$

于是命题得证。 □

19.

命题 2.3.22. 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$

取 V 的一个基 $\{\alpha_1, \cdots, \alpha_n\}$，设 $A \in \hom(V, V)$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为 $\boldsymbol{A}$

A 是幂等变换 $\Leftrightarrow \boldsymbol{A}$ 是幂等矩阵

证明: A 是幂等变换 $\Leftrightarrow A^2 = A \Leftrightarrow \boldsymbol{A}^2 = \boldsymbol{A} \Leftrightarrow \boldsymbol{A}$ 是幂等矩阵 □

20.

命题 2.3.23. 设 V, W 是域 F 上的两个线性空间，$\dim V = n < \aleph_0, \dim W = m < \aleph_0$

设 V, W 的一个基分别是 $\{\alpha_1, \cdots, \alpha_n\}, \{\beta_1, \cdots, \beta_m\}$

$\forall \mathcal{A} \in \hom(V, W)$，如果它在 $\{\alpha_1, \cdots, \alpha_n\}, \{\beta_1, \cdots, \beta_m\}$ 下的矩阵是 $\boldsymbol{A}$

$\gamma \in V$ 满足 $\gamma = (\alpha_1, \cdots, \alpha_n)\mathbf{x}$，其中 $\mathbf{x} \in F^n$

那么，$\mathcal{A}(\gamma)$ 在 $\{\beta_1, \cdots, \beta_m\}$ 下的坐标是 $\boldsymbol{A}\mathbf{x}$。

证明： 设 $\mathbf{x} = (x_1, \cdots, x_n)^T$

那么，$\mathcal{A}(\gamma) = \mathcal{A}\left((\alpha 1, \cdots, \alpha_n)\mathbf{x}\right)$

$= \mathcal{A}(x_1\alpha_1 + \cdots + x_n\alpha_n) = x_1\mathcal{A}(\alpha_1) + \cdots + x_n\mathcal{A}(\alpha_n) = (\mathcal{A}(\alpha_1, \cdots, \alpha_n))\,\mathbf{x}$

$((\beta_1, \cdots, \beta_m)\boldsymbol{A})\,\mathbf{x} = (\beta_1, \cdots, \beta_m)\,(\boldsymbol{A}\mathbf{x})$

于是命题得证 □

2.4 特殊矩阵

本节中我们介绍一些后续需要用到的特殊矩阵

1. 单位矩阵

定义 2.4.1: 单位矩阵

n 阶矩阵环 $M_n(F)$ 的乘法单位元称为单位矩阵，记作 $\boldsymbol{I_n}$，或简记为 $\boldsymbol{I}$

在前面一节，我们已经知道，$\boldsymbol{I}$ 的具体形式，其实就是满足 $\boldsymbol{I}(i;j) = \delta_{ij}$ 的矩阵

也就是 $\boldsymbol{I_n} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$

2. 数量矩阵

定义 2.4.2: 数量矩阵

我们称单位矩阵与一纯量的乘法的结果为数量矩阵：

$$k\boldsymbol{I} = \begin{pmatrix} k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & k \end{pmatrix} \in M_n(F), k \in F \tag{2.13}$$

记作 $k\boldsymbol{I}$

全体数量矩阵显然构成一个线性子空间，我们一般记为 $F\boldsymbol{I}$

3. 基本矩阵

定义 2.4.3: 基本矩阵

我们定义基本矩阵 $\boldsymbol{E_{ij}} \in M_{m\times n}(F)$ 为满足以下条件的矩阵：

$\boldsymbol{E_{ij}}(k;l) = \delta_{ik}\delta_{jl}$

我们之前已经指出，全体基本矩阵是矩阵空间的一个基

容易注意到以下结果：

命题 2.4.1. 设 $\mathbf{e}_i$ 为 F^n 中仅在第 i 个元素上为 1，其他元素均为 0 的向量

那么有：$\boldsymbol{E}_{ij} = \mathbf{e}_i\mathbf{e}_j^T$

4. 对角矩阵

定义 2.4.4: 对角矩阵

我们定义：

$$diag\{d_1, \cdots, d_n\} = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix} \in M_n(F), d_1, \cdots, d_n \in F \tag{2.14}$$

5. 初等矩阵

定义 2.4.5: 初等矩阵

设 $\mathbf{e}_i$ 为 F^n 中仅在第 i 个元素上为 1，其他元素均为 0 的向量

我们定义：

$$\boldsymbol{P}(i,j) = \boldsymbol{I} - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \in M_n(F) \tag{2.15}$$

$$\boldsymbol{P}(i(c)) = \boldsymbol{I} + (c-1)\mathbf{e}_i\mathbf{e}_i^T \in M_n(F), c \in F, c \neq 0 \tag{2.16}$$

$$\boldsymbol{P}(j, i(c)) = \boldsymbol{I} + c\mathbf{e}_j\mathbf{e}_i^T \in M_n(F), i \neq j, c \in F \tag{2.17}$$

这三种矩阵统称为初等矩阵

容易注意到它们的具体形式：$\boldsymbol{P}(i,j)=\begin{pmatrix}1&&&&&&0\\&\ddots&&&&&\\&&0&\cdots&1&&\\&&\vdots&&\vdots&&\\&&1&\cdots&0&&\\&&&&&\ddots&\\0&&&&&&1\end{pmatrix}$

$$\boldsymbol{P}(i(c))=\begin{pmatrix}1&&&&&&0\\&\ddots&&&&&\\&&k&&&&\\&&&\ddots&&&\\&&&&1&&\\&&&&&\ddots&\\0&&&&&&1\end{pmatrix}\boldsymbol{P}(j,i(c))=\begin{pmatrix}1&&&&&&0\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\0&&&&&&1\end{pmatrix}$$

6. 对称矩阵、反对称矩阵

定义 2.4.6: 对称矩阵与反对称矩阵

我们定义：

如果 $\boldsymbol{A}\in M_n(F)$ 满足：

$\boldsymbol{A}^T=\boldsymbol{A}$，那么我们称 $\boldsymbol{A}$ 是一个对称矩阵；

$\boldsymbol{A}^T=-\boldsymbol{A}$，那么我们称 $\boldsymbol{A}$ 是一个反对称矩阵。

7. 上三角矩阵、下三角矩阵

定义 2.4.7: 我

定义：

如果 $\boldsymbol{A}\in M_n(F)$ 满足：

$\forall i>j,\boldsymbol{A}(i;j)=0$，那么我们称 $\boldsymbol{A}$ 是一个上三角矩阵；

$\forall i<j,\boldsymbol{A}(i;j)=0$，那么我们称 $\boldsymbol{A}$ 是一个下三角矩阵。

2.5 可逆矩阵

可逆矩阵其实就是矩阵环 $M_n(F)$ 中的可逆元。本节中我们将研究可逆矩阵的一些基本性质，以及矩阵环中的特殊性质。最后，我们会看到它的一个简单应用：过渡矩阵和相似

2.5.1 可逆矩阵的定义

定义 2.5.1: 可逆矩阵

设 $\boldsymbol{A} \in M_n(F)$，如果 $\exists \boldsymbol{A}^{-1} \in M_n(F)$，使得：

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} \tag{2.18}$$

那么我们称 $\boldsymbol{A}$ 是一个可逆矩阵，并称 $\boldsymbol{A}^{-1}$ 是 $\boldsymbol{A}$ 的逆矩阵

我们记全体可逆矩阵的集合为 $GL_n(F)$，称为 F 上的 n 阶可逆矩阵群

2.5.2 可逆矩阵的性质

1.

命题 2.5.1. $\boldsymbol{A} \in GL_n(F) \Rightarrow \left(\boldsymbol{A}^{-1}\right)^{-1} = \boldsymbol{A}$

证明： 依照定义显然。 □

2.

命题 2.5.2. 设 V 是一个 F 上的线性空间，且它们一个基是 $\{\alpha_1, \cdots, \alpha_n\}$

则 $\forall A \in \hom(V,V)$，若 A 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$，那么：

A 可逆 $\Leftrightarrow \boldsymbol{A} \in GL_n(F)$

且此时 A^{-1} 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 $\boldsymbol{A}^{-1}$

证明： A 可逆 $\Leftrightarrow \exists B \in \hom(V,V), AB = BA = I$

记 B 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为 $\boldsymbol{B}$，则：

A 可逆 $\Leftrightarrow \boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I} \Leftrightarrow \boldsymbol{A} \in GL_n(F), \boldsymbol{A}^{-1} = \boldsymbol{B}$

于是命题得证。 □

3.

命题 2.5.3. $\boldsymbol{A} \in GL_n(F) \Leftrightarrow \operatorname{rank}(\boldsymbol{A}) = n$

证明： 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$

任取 V 的一个基 $\{\alpha_1, \cdots, \alpha_n\}$

显然存在 $\mathcal{A} \in \hom(V, V)$，使得 $\mathcal{A}$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 上的矩阵是 $\boldsymbol{A}$

$\boldsymbol{A} \in GL_n(F) \Leftrightarrow \mathcal{A}$ 可逆 $\Leftrightarrow \mathrm{rank}(\mathcal{A}) = n \Leftrightarrow \mathrm{rank}(\boldsymbol{A}) = n$ □

4.

命题 2.5.4. $\boldsymbol{A}, \boldsymbol{B} \in M_n(F)$

那么：$\boldsymbol{AB} = \boldsymbol{I} \Rightarrow \boldsymbol{A}, \boldsymbol{B} \in GL_n(F), \boldsymbol{A}^{-1} = \boldsymbol{B}, \boldsymbol{B}^{-1} = \boldsymbol{A}$

证明： $n = \mathrm{rank}(\boldsymbol{I}) = \mathrm{rank}(\boldsymbol{AB}) \leqslant \mathrm{rank}(\boldsymbol{A}) \Rightarrow \mathrm{rank}(\boldsymbol{A}) = n \Rightarrow \boldsymbol{A} \in GL_n(F)$

同理可证 $\boldsymbol{B} \in GL_n(F)$，$\boldsymbol{A}^{-1} = \boldsymbol{B}, \boldsymbol{B}^{-1} = \boldsymbol{A}$ 依照定义是显然的。

于是命题得证 □

5.

命题 2.5.5. $(\boldsymbol{P}(i, j))^{-1} = \boldsymbol{P}(j, i)$

$(\boldsymbol{P}(i(c)))^{-1} = \boldsymbol{P}(i(\frac{1}{c}))$

$(\boldsymbol{P}(j, i(c)))^{-1} = \boldsymbol{P}(j, i(-c))$

证明： 注意到：

$\boldsymbol{P}(i, j)\boldsymbol{P}(j; i) = \left(\boldsymbol{I} - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T\right)\left(\boldsymbol{I} - (\mathbf{e}_j - \mathbf{e}_i)(\mathbf{e}_j - \mathbf{e}_i)^T\right)$

$= (\boldsymbol{I} - (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T)^2 = \boldsymbol{I} - 2(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T + (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T$

当 $i \neq j$，有 $(\mathbf{e}_i - \mathbf{e}_j)^T(\mathbf{e}_i - \mathbf{e}_j) = \delta_{ii} + \delta_{jj} - 2\delta_{ij} = 2$

于是有 $\boldsymbol{P}(i, j)\boldsymbol{P}(j; i) = \boldsymbol{I}$

注意到：

$\boldsymbol{P}(i(c))\boldsymbol{P}(i(\frac{1}{c})) = \left(\boldsymbol{I} + (c - 1)\mathbf{e}_i\mathbf{e}_i^T\right)\left(\boldsymbol{I} + (\frac{1}{c} - 1)\mathbf{e}_i\mathbf{e}_i^T\right)$

$= \boldsymbol{I} + (c + \frac{1}{c} - 2)\mathbf{e}_i\mathbf{e}_i^T + (1 - c - \frac{1}{c} + 1)\mathbf{e}_i\mathbf{e}_i^T\mathbf{e}_i\mathbf{e}_i^T = \boldsymbol{I}$

注意到：

$\boldsymbol{P}(j, i(c))\boldsymbol{P}(j, i(-c)) = \left(\boldsymbol{I} + c\mathbf{e}_i\mathbf{e}_j^T\right)\left(\boldsymbol{I} - c\mathbf{e}_i\mathbf{e}_j^T\right)$

$= \boldsymbol{I} - c^2\mathbf{e}_i\mathbf{e}_j^T\mathbf{e}_i\mathbf{e}_j^T = \boldsymbol{I}$

于是命题得证。 □

2.5.3 过渡矩阵

我们研究一种矩阵，称为过渡矩阵。

定义 2.5.2: 过渡矩阵

设 V 是一个 F 上的有限维线性空间，$\{\alpha_1,\cdots,\alpha_n\},\{\beta_1,\cdots,\beta_n\}$ 是它的两个基

如果 $\boldsymbol{S}\in M_n(F)$ 满足：

$$(\alpha_1,\cdots,\alpha_n)=(\beta_1,\cdots,\beta_n)\boldsymbol{S} \tag{2.19}$$

那么我们称 $\boldsymbol{S}$ 是 $\{\beta_1,\cdots,\beta_n\}$ 到 $\{\alpha_1,\cdots,\alpha_n\}$ 的过渡矩阵

容易看出，其实过渡矩阵就是恒等变换 I 在定义域和值域上选取不同基后的矩阵

我们可以看出以下性质：

1.

命题 2.5.6. 如果 $\boldsymbol{S}$ 是 $\{\beta_1,\cdots,\beta_n\}$ 到 $\{\alpha_1,\cdots,\alpha_n\}$ 的过渡矩阵，那么 $\boldsymbol{S}\in GL_n(F)$

证明： 注意到：I 在 $\{\beta_1,\cdots,\beta_n\},\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵 $\boldsymbol{P}$ 满足 $\boldsymbol{PS}=\boldsymbol{SP}=\boldsymbol{I}$

于是命题得证 □

2. 过渡矩阵最大的应用，就是给出了切换线性变换的矩阵所依赖的基后，矩阵究竟会如何变化。

命题 2.5.7. 设 V 是 F 上的线性空间，$\dim V=n<\aleph_0$，$\{\beta_1,\cdots,\beta_n\},\{\alpha_1,\cdots,\alpha_n\}$ 是它的两个基

设 $A\in\hom(V,V)$ 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$，$\boldsymbol{S}$ 是 $\{\alpha_1,\cdots,\alpha_n\}$ 到 $\{\beta_1,\cdots,\beta_n\}$ 的过渡矩阵

那么，A 在 $\{\beta_1,\cdots,\beta_n\}$ 下的矩阵 $\boldsymbol{B}$ 有：

$\boldsymbol{B}=\boldsymbol{S}^{-1}\boldsymbol{AS}$

证明： 这是显然的，因为 $A=IAI$，而 I 在 $\{\beta_1,\cdots,\beta_n\},\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为 $\boldsymbol{S}$，在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵为 $\boldsymbol{S}^{-1}$ □

如果两个矩阵具备以上所述的关系——即同一变换在不同基下，我们称这种关系为相似。接下来我们正式阐述相似矩阵

2.5.4 矩阵的相似

定义 2.5.3: 相似矩阵

设 $\boldsymbol{A},\boldsymbol{B}\in M_n(F)$

如果 $\exists \boldsymbol{P}\in GL_n(F)$，使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$

那么我们称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，记作 $\boldsymbol{A}\sim\boldsymbol{B}$

接下来我们研究相似矩阵的性质

1.

命题 2.5.8. 设 $\boldsymbol{A},\boldsymbol{B}\in M_n(F)$

那么：$\boldsymbol{A}\sim\boldsymbol{B}\Leftrightarrow\exists T\in\hom(V,V),\dim_F V=n$，$\boldsymbol{A},\boldsymbol{B}$ 是 T 分别在基 B_1,B_2 下的矩阵

证明： 根据前面同一变换在不同基下矩阵的关系，这是显然的 □

2. 相似是一种等价关系

这是前面命题的直接推论：

推论 2.5.1: 矩

的相似是一种等价关系

证明： 这是显然的，因为我们已经指出，相似矩阵只是同一映射在不同基下的表示 □

3. 相似矩阵具有相同的秩

这事实上也只是前面命题的推论

推论 2.5.2: 如

$\boldsymbol{A},\boldsymbol{B}\in M_n(F),\boldsymbol{A}\sim\boldsymbol{B}$

那么有 $\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{B})$

证明： 这是显然的，因为我们之前已经证明了，无论选取哪个基，映射的矩阵一定和映射具有相同的秩。 □

2.6 迹和行列式

2.6.1 张量积

在研究迹和行列式前，我们先研究张量积。张量积一般应用于多重线性映射的研究中，后面我们会看到：行列式是一种特殊的多重线性映射。

张量积的本意是将两个线性空间组合在一起形成新的空间，并且希望可以将原来的两个空间和新的空间对应。

在双线性映射，即对每一分量都线性的映射的研究中，我们将两个空间组合在一起，视为同一空间研究

我们希望：新的空间和原来的两个空间应该别无二致。因此，如果确立了两个空间 V,W 到张量积空间 U 的映射，我们希望每一个 V,W 到 Z 的双线性映射，都只能确定唯一的 U 到 Z 的映射。

以上想法给出了以下定义：

定义 2.6.1: 双线性映射

设 V,W,Z 是三个线性空间，

映射 $f:V\times W\to Z$ 如果满足：

1. $\forall\alpha,\beta\in V,\gamma\in W,f(\alpha+\beta,\gamma)=f(\alpha,\gamma)+f(\beta+\gamma)$
2. $\forall\alpha\in V,\beta,\gamma\in W,f(\alpha,\beta+\gamma)=f(\alpha,\beta)+f(\alpha,\gamma)$
3. $\forall k\in F,\alpha\in V,\beta\in W,f(k\cdot\alpha,\beta)=k\cdot f(\alpha,\beta),f(\alpha,k\cdot\beta)=k\cdot f(\alpha,\beta)$

那么我们称 f 是一个双线性映射

定义 2.6.2: 线性空间的张量积

设 V,W,Z 是三个 F 上的线性空间，如果线性空间 $V\otimes W$ 满足：

存在映射 $f:V\times W\to V\otimes W$，使得：

任意的 $f:V\times W\to Z$ 的双线性映射 g，存在唯一的线性映射 $h:V\otimes W\to Z$，使得 $g=h\circ f$

那么，我们称 $V\otimes W$ 是 V 和 W 的张量积

此时，我们也记 $\alpha\otimes\beta:=f(\alpha,\beta)$

在继续研究之前，我们需要先证明张量积的确是存在的：

命题 2.6.1. 任意两个 F 上的线性空间 V,W 的张量积 $V\otimes W$ 是存在的，并且在同构意义下唯一

我们也容易证明，张量积在同构意义下满足结合律

命题 2.6.2. $(V\otimes W)\otimes Z\cong V\otimes(W\otimes Z)$

2.6.2　张量和外积

我们首先介绍张量的概念

定义 2.6.3: 张量

设 V 是 F 上的一个线性空间，我们定义：

$$T_s^r := T^rV\otimes T^sV^* \tag{2.20}$$

其中

$$T^pW := \begin{cases} F, p=0 \\ W\otimes T^{p-1}W, p\geqslant 1 \end{cases} \tag{2.21}$$

我们称 T_s^r 中的元素为一个 (r,s) 型张量

定义 2.6.4: 张量代数

我们定义：

$$TV=\bigoplus_{r=0}^{\infty}T^rV \tag{2.22}$$

称为 V 上的张量代数

其中，TV 的加法和纯量乘法为映射的加法和纯量乘法

同时，我们定义 TV 中的乘法为：$\forall f,g\in TV$

$(f\cdot g)(k)=\sum\limits_{i+j=k}f(i)\otimes g(j)$

定义 2.6.5: 外代数

记 $S=\{\alpha\otimes\alpha|\alpha\in V\}, I=(S)$

我们定义：

$$\bigwedge V=TV/I \tag{2.23}$$

称为 V 上的外代数

特别地，我们记：$I_p=I\cap T^pV, \bigwedge^p V=T^pV/I_p$

并定义：$\alpha_1\wedge\cdots\wedge\alpha_m:=\alpha_1\otimes\cdots\otimes\alpha_m+I$

称为 $\alpha_1,\cdots,\alpha_m$ 的外积

2.6.3 迹

在完成了预备工作后，我们准备开始定义矩阵的迹

迹是体现线性映射的不变量之一，其核心在于对映射复合的交换下不变

定义 2.6.6: 迹

设 F 是一个域，V 是 F 上的一个线性空间，并且 $dim_FV=n$

满足以下条件的映射 $\mathrm{tr}:\hom(V,V)\to F$ 称为迹：

1. $\forall A,B\in\hom(V,V),\mathrm{tr}(AB)=\mathrm{tr}(BA)$
2. $\mathrm{tr}(I)=\dim V=n$

我们将看到，迹是存在且唯一的。

我们也可以定义矩阵的迹：

定义 2.6.7: 矩阵的迹

设 $\boldsymbol{A}\in M_n(F)$，V 是一个 F 上的 n 维线性空间

如果 $A\in\hom(V,V)$ 在基 B 上的矩阵为 $\boldsymbol{A}$

那么我们定义：

$$\mathrm{tr}(\boldsymbol{A}):=\mathrm{tr}(A) \tag{2.24}$$

迹有以下性质：

1.

命题 2.6.3. tr 存在且唯一

证明: 取 V 的一个基 $\{\alpha_1, \cdots, \alpha_n\}$ 和它的对偶基 $\{f_1, \cdots, f_n\}$

我们定义：$f : \hom(V, V) \ni T \mapsto \sum_{i=1}^{n} f_i\left(T(\alpha_i)\right) \in F$

$tr(TS) = \sum_{i=1}^{n} f_i\left(T \sum_{j=1}^{n} f_j(S(\alpha_i))\alpha_j\right)$

□

2.6.4 行列式

定义 2.6.8: 行列式

设 F 是一个域，V 是 F 上的一个线性空间，并且 $dim_F V = n$

对于 $T \in \hom(V, V)$，我们定义映射 $\bigwedge^n T : \bigwedge^n V \to \bigwedge^n V$ 为：

$$\left(\bigwedge^n T\right)(\alpha_1 \wedge \cdots \wedge \alpha_n) := T(\alpha_1) \wedge \cdots \wedge T(\alpha_n) \tag{2.25}$$

由下式导出的映射 $\det : \hom(V, V) \to F$ 称为行列式：

$$\left(\bigwedge^n T\right)(\omega) = \det(T)\omega \tag{2.26}$$

我们自然地把矩阵的行列式定义为其映射的行列式：

定义 2.6.9: 矩阵的行列式

设 $\boldsymbol{A} \in M_n(F)$，$V$ 是一个 F 上的 n 维线性空间

如果 $A \in \hom(V, V)$ 在基 B 上的矩阵为 $\boldsymbol{A}$

那么我们定义：

$$\det(\boldsymbol{A}) := \det(A) \tag{2.27}$$

A 的行列式的另外一种常用的符号是记作 $|A|$；特别地，我们也常常记 $\boldsymbol{A} = (\alpha_1, \cdots, \alpha_n)$ 的行列式为 $\det(\alpha_1, \cdots, \alpha_n)$

我们也给出子式的定义，它是矩阵的一个子矩阵的行列式，常常用于对行列式进行降阶

定义 2.6.10: 子式、余子式

设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in M_n(F), I=\{i_1,\cdots,i_k\}, J=\{j_1,\cdots,j_k\}$，其中 $i_1<\cdots<i_k, j_1<\cdots<j_k$

我们定义：

$\boldsymbol{A}_J^I:=\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k}:=(a_{ij})_{k\times k}, i\in I, j\in J$

行列式 $\det(\boldsymbol{A}_J^I)$ 称为 $\boldsymbol{A}$ 的一个 k 阶子式

特别地，如果 $I'=\{i_1',\cdots,i_{n-k}'\}, J'=\{j_1',\cdots,j_{n-k}'\}, i_1'<\cdots<i_{n-k}', j_1'<\cdots<j_{n-k}', I\cup I'=\mathbb{Z}_n, J\cup J'=\mathbb{Z}_n$

那么我们 $\det(\boldsymbol{A}_{J'}^{I'})$ 是 $\det(\boldsymbol{A}_J^I)$ 的余子式

映射行列式的讨论比较简单，我们接下来讨论中，主要是矩阵的行列式

1. 存在性和唯一性

   事实上，在定义中，我们甚至不能看出行列式是否存在。我们首先证明行列式是存在且唯一的

   命题 2.6.4. *det 是存在且唯一的*

   证明: 设 V 是一个 F 上的线性空间，$\dim V=n$，并取它的一个基 $\{\gamma_1,\cdots,\gamma_n\}$

   设 $\omega=\gamma_1\wedge\cdots\wedge\gamma_n$，$T(\gamma_i)=\sum\limits_{j=1}^n a_{ji}\gamma_j$

   注意到：$\left(\bigwedge^n T\right)(\omega)=T(\gamma_1)\wedge\cdots\wedge T(\gamma_n)=\sum\limits_{i_1=1}^n a_{i_1,1}\gamma_{i_1}\wedge\cdots\wedge\sum\limits_{i_n=1}^n a_{i_n,n}\gamma_{i_n}$

   $=\sum\limits_{i_1=1}^n\cdots\sum\limits_{i_n=1}^n\prod\limits_{k=1}^n a_{i_k,i}\gamma_{i_1}\wedge\cdots\wedge\gamma_{i_n}$

   $=\sum\limits_{\sigma\in S_n}\left(\prod\limits_{k=1}^n \operatorname{sgn}(\sigma)a_{\sigma(k),k}\right)\omega$

   $\Rightarrow \det(T)=\sum\limits_{\sigma\in S_n}\operatorname{sgn}(\sigma)\left(\prod\limits_{k=1}^n a_{\sigma(k),k}\right)$

   至此，我们推导出了行列式的具体形式，这说明行列式是存在且唯一的。 □

   从证明过程中可见以下推论，这恰恰是矩阵行列式的计算方式

   推论 2.6.1

   $\boldsymbol{A}=(a_{ij})\in M_n(F)$

   $$|\boldsymbol{A}|=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^n a_{\sigma(i),i} \tag{2.28}$$

2. 行列式的交错线性性

事实上，如果考察矩阵的行列式，它呈现出交错多重线性函数的特征，这是其外积定义带来的，正如以下

定理 2.6.2: 矩阵行列式的多重线性性

对于 $\alpha_1, \cdots, \alpha_n, \beta_i \in F^n, k \in F$

(a) $\det(\alpha_1, \cdots, \alpha_i + \beta_i, \cdots, \alpha_n) = \det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_n) + \det(\alpha_1, \cdots, \beta_i, \cdots, \alpha_n)$

(b) $\forall k \in F, \det(\alpha_1, \cdots, k \cdot \alpha_i, \cdots, \alpha_n) = k \cdot \det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_n)$

(c) $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = -\det(\alpha_1, \cdots, \alpha_j, \cdots, \alpha_i, \cdots, \alpha_n)$

证明:

□

这也是矩阵行列式的基本性质。进一步，我们可以推出以下推论：

推论 2.6.3

量组 $\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n$ 如果有 $\alpha_i = \alpha_j$

那么 $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = 0$

证明: $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = -\det(\alpha_1, \cdots, \alpha_j, \cdots, \alpha_i, \cdots, \alpha_n)$

但因为 $\alpha_i = \alpha_j$，所以必有 $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = 0$ □

进一步我们可以推出，如果两个变量成系数关系，那么行列式也为零

推论 2.6.4

量组 $\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n$ 如果有 $\alpha_i = k\alpha_j, k \in F$

那么 $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = 0$

证明: $\det(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_j, \cdots, \alpha_n) = k \cdot \det(\alpha_1, \cdots, \alpha_j, \cdots, \alpha_j, \cdots, \alpha_n) = 0$ □

3. 行列式的置换展开

事实上，的展开的下标，只需要选取两个置换即可，其顺序和选取都是任意的

推论 2.6.5

设 $\boldsymbol{A}=(a_{ij})\in M_n(F)$

那么，有：

$$\det(\boldsymbol{A})=\operatorname{sgn}(\rho)\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}a_{\rho(i),\sigma(i)} \tag{2.29}$$

其中 $\rho\in S_n$

证明： $\det(\boldsymbol{A})=\sum_{\tau\in S_n}\operatorname{sgn}(\tau)\prod_{i=1}^{n}a_{\tau(i),i}$

对指标作置换 ρ，累乘的结果不会变化，所以有：

$\det(\boldsymbol{A})=\sum_{\tau\in S_n}\operatorname{sgn}(\tau)\prod_{i=1}^{n}a_{(\rho\circ\tau)(i),\rho(i)}$

记 $\sigma=\rho\circ\tau$，那么 $\det(\alpha_1,\cdots,\alpha_n)=\sum_{\rho^{-1}\circ\sigma\in S_n}\operatorname{sgn}(\rho^{-1}\circ\sigma)\prod_{i=1}^{n}a_{\sigma(i),\rho(i)}$

因为 $\operatorname{sgn}(\rho^{-1}\circ\sigma)=\operatorname{sgn}(\rho)\operatorname{sgn}(\sigma)$

所以 $\det(\alpha_1,\cdots,\alpha_n)=\operatorname{sgn}(\rho)\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\prod_{i=1}^{n}a_{\sigma(i),\rho(i)}$ □

4. 行列式关于矩阵转置和可逆性的性质

命题 2.6.5. $\boldsymbol{A}\in GL_n(F)\Leftrightarrow|\boldsymbol{A}|\neq 0$

证明： $\boldsymbol{A}\notin GL_n(F)\Leftrightarrow\operatorname{rank}(\boldsymbol{A})<n$

那么，$\boldsymbol{A}$ 一定有一个列向量，可被其他列线性表出

依照之前证明的交错线性性质，一定有 $|\boldsymbol{A}|=0$

于是命题得证。 □

5. 行列式的余子式展开

引理 2.6.6

设 V 是一个 F 上的线性空间，$\dim V=n<\aleph_0$，$\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一个基

对于 $\boldsymbol{A}\in M_n(F),i_1<\cdots<i_k,j_1<\cdots<j_k$，如果 $A\in\hom(V,V)$ 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$

那么有：

$$\left(\bigwedge^k A\right)(\alpha_{i_1}\wedge\cdots\wedge\alpha_{i_k})=\sum_{j_1<\cdots<j_k}\det(\boldsymbol{A}^{i_1,\cdots,i_k}_{j_1,\cdots,j_k})\cdot\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_k}$$

证明： 设 $\boldsymbol{A}=(a_{ij})$

$\left(\bigwedge^k A\right)(\alpha_{i_1}\wedge\cdots\wedge\alpha_{i_k})$

$$= \left(\sum_{s_1=1}^{n} a_{s_1 i_1}\alpha_{s_1}\right)\wedge\cdots\wedge\left(\sum_{s_k=1}^{n} a_{s_k i_k}\alpha_{s_k}\right)$$
$$= \sum_{\rho\in S_{\{j_1,\cdots,j_k\}}} \prod_{l=1}^{k} a_{\rho(l)i_l}\alpha_{\rho(1)}\wedge\cdots\wedge\alpha_{\rho(l)}$$
$$= \sum_{j_1<\cdots<j_k}\left(\sum_{\rho\in S_k}\mathrm{sgn}(\rho)\prod_{l=1}^{k} a_{j_{\rho(l)}i_l}\right)\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_k}$$
$$= \sum_{j_1<\cdots<j_k}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\cdot\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_k}$$ □

下面的定理指出了行列式关于 k 列展开的方法：

定理 2.6.7: Laplace 定理

设 $\boldsymbol{A}\in M_n(F)$，那么：

$$\det(\boldsymbol{A}) = \sum_{j_1<\cdots<j_k}(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\det(\boldsymbol{A}_{j'_1,\cdots,j'_{n-k}}^{i'_1,\cdots,i'_{n-k}}) \tag{2.30}$$

其中 $\{i_1,\cdots,i_k\}\cup\{i'_1,\cdots,i'_{n-k}\} = \{j_1,\cdots,j_k\}\cup\{j'_1,\cdots,j'_{n-k}\} = \mathbb{Z}_n, i'_1<\cdots<i'_{n-k}, j'_1<\cdots<j'_{n-k}$

证明：任取一个 F 上的 n 维线性空间 V，并取它的一个基 $\{\alpha_1,\cdots,\alpha_n\}$

那么，一定存在一个 $A\in\hom(V,V)$，并且它在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$

于是：$\det(\boldsymbol{A})\cdot\alpha_1\wedge\cdots\wedge\alpha_n$

$$= A(\alpha_1)\wedge\cdots\wedge A(\alpha_n)$$
$$= \left(\bigwedge^n A\right)(\alpha_1\wedge\cdots\wedge\alpha_n)$$
$$= (-1)^{(i_1+\cdots+i_k)-k}\left(A(\alpha_{i_1})\wedge\cdots\wedge A(\alpha_{i_k})\right)\wedge\left(A(\alpha_{i'_1})\wedge\cdots\wedge A(\alpha_{i'_{n-k}})\right)$$
$$= (-1)^{(i_1+\cdots+i_k)-k}\left(\sum_{j_1<\cdots<j_k}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\cdot\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_k}\right)\wedge$$
$$\left(\sum_{j'_1<\cdots<j'_{n-k}}\det(\boldsymbol{A}_{j'_1,\cdots,j'_{n-k}}^{i'_1,\cdots,i'_{n-k}})\cdot\alpha_{j'_1}\wedge\cdots\wedge\alpha_{j'_{n-k}}\right)$$
$$= (-1)^{(i_1+\cdots+i_k)-k}\sum_{j_1<\cdots<j_k}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\det(\boldsymbol{A}_{j'_1,\cdots,j'_{n-k}}^{i'_1,\cdots,i'_{n-k}})\left(\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_k}\wedge\alpha_{j'_1}\wedge\cdots\wedge\alpha_{j'_{n-k}}\right)$$
$$= (-1)^{(i_1+\cdots+i_k)-k}\sum_{j_1<\cdots<j_k}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\det(\boldsymbol{A}_{j'_1,\cdots,j'_{n-k}}^{i'_1,\cdots,i'_{n-k}})(-1)^{(j_1+\cdots+j_k)-k}(\alpha_1\wedge\cdots\wedge\alpha_n)$$
$$= \left(\sum_{j_1<\cdots<j_k}(-1)^{(i_1+\cdots+i_k)+(j_1+\cdots+j_k)}\det(\boldsymbol{A}_{j_1,\cdots,j_k}^{i_1,\cdots,i_k})\det(\boldsymbol{A}_{j'_1,\cdots,j'_{n-k}}^{i'_1,\cdots,i'_{n-k}})\right)(\alpha_1\wedge\cdots\wedge\alpha_n)$$

于是命题得证。 □

6. 矩阵乘积的行列式

第三章　线性变换的表示与分解

在上一章中，我们研究了线性映射的基础理论。本章中，我们进一步，寻求对线性映射的表示和分解。

我们已经看到：选定一个映射和两个基，就可以写出矩阵。但是，其实这些矩阵中有些是容易使用的，而有些则是难以处理的。如果有一个基下映射的矩阵我们觉得是良好的，我们常常会称这类矩阵为标准型

本章中我们不研究所有线性映射，而是聚焦一种特殊的线性映射——线性变换。我们会看到，如果域是代数闭的，在特定条件下我们可以将其写成对角矩阵；而更一般地，它一定可以写成 Jordan 块组成的分块对角矩阵。

本章的第一节我们讲述对角标准型的基础工作——特征值、特征向量；第二节讲解对角标准型的定义及判定是否可对角化的方式；

第三至第七节，我们逐渐搭建出 Jordan 标准型的理论。

3.1　线性变换的特征值和特征向量

在本节和下一节中，我们将解决对角标准型的问题。

对角标准型的想法是：尽管并非所有线性映射呈现出 $A(\alpha)=\lambda\alpha$ 的形式，但是很多线性映射都呈现为将基向量仅仅进行伸缩的形式。这样的线性映射会在指定基上呈现为对角矩阵，其对角线元素我们一般称为特征值。

我们首先构建特征值和其基向量——特征向量的理论。

3.1.1 特征值和特征向量的定义

定义 3.1.1: 线性变换的特征值、特征向量、特征子空间

设 V 是一个 F 上的线性空间，$A \in \hom(V,V)$

那么如果 $\exists \lambda \in F, \alpha \in V, \alpha \neq \mathbf{0}$ 使得

$$A(\alpha) = \lambda\alpha \tag{3.1}$$

成立

那么我们称 λ 是 A 的一个特征值，α 是 A 的隶属于 λ 的一个特征向量

同时我们定义：

$$V_\lambda := \{\alpha | A(\alpha) = \lambda\alpha\} \tag{3.2}$$

称为 A 的属于特征值 λ 的特征子空间

特征值代表着一个线性映射对一个向量的伸缩程度，这里我们要求至少能找到一个向量不为零是因为：如果一个特征值只对于零向量成立，那么它过于平凡，而且会扰乱后续我们一些关于特征值数量的命题

类似地，我们自然可以定义矩阵的特征值

定义 3.1.2: 矩阵的特征值、特征向量

设 $\boldsymbol{A} \in M_n(F)$

那么如果 $\exists \lambda \in F, \alpha \in F^n$，使得

$$\boldsymbol{A}\alpha = \lambda\alpha \tag{3.3}$$

成立。

那么我们称 λ 是 $\boldsymbol{A}$ 的一个特征值，α 是 $\boldsymbol{A}$ 的隶属于 λ 的一个特征向量

3.1.2 特征值和特征向量的性质

我们探讨一些相关的基本性质

1. 我们首先验证，特征子空间的确是子空间

   命题 3.1.1. $\forall A \in \hom(V,V)$，V_λ 是 V 的一个线性子空间

证明： 取 $\forall\alpha,\beta\in V_\lambda, k\in F$

$A(\alpha+\beta)=\lambda\alpha+\lambda\beta=\lambda(\alpha+\beta)\Rightarrow\alpha+\beta\in V_\lambda$

$A(k\alpha)=k\lambda\alpha=\lambda(k\alpha)\Rightarrow k\alpha\in V_\lambda$

于是命题得证 □

请注意：特征子空间并非全体特征向量的集合，而是全体特征向量和零向量的集合

2. 从特征值的定义可以看出，特征值就像是把线性变换的一部分转换为纯量乘法。我们猜想：每一个特征值体现了映射的不同的“方向”，不同特征子空间的线性无关向量组的并也应该线性无关

以下命题证实了这个猜想

命题 3.1.2. 设 $A\in\hom(V,V)$，λ_1,λ_2 是 A 的两个特征值，$\lambda_1\neq\lambda_2$

如果 $\alpha_1,\cdots,\alpha_m\in V_{\lambda_1}$ 线性无关，$\beta_1,\cdots,\beta_n\in V_{\lambda_2}$ 线性无关

那么 $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_n$ 也线性无关

证明： 取线性组合并设其为零:

$k_1\alpha_1+\cdots+k_m\alpha_m+l_1\beta_1+\cdots+l_n\beta_n=\mathbf{0}$

将 A 在其上进行变换得：$k_1A(\alpha_1)+\cdots+k_mA(\alpha_m)+l_1A(\beta_1)+\cdots+l_nA(\beta_n)=\mathbf{0}$

$\Rightarrow k_1\lambda_1\alpha_1+\cdots+k_m\lambda_1\alpha_m+l_1\lambda_2\beta_1+\cdots+l_n\lambda_2\beta_n=\mathbf{0}$

但是，如果我们把最初的线性组合乘以 λ_1，得：

$k_1\lambda_1\alpha_1+\cdots+k_m\lambda_1\alpha_m+l_1\lambda_1\beta_1+\cdots+l_n\lambda_1\beta_n=\mathbf{0}$

将两式相减，得：

$l_1(\lambda_2-\lambda_1)\beta_1+\cdots+l_n(\lambda_2-\lambda_1)\beta_n=\mathbf{0}$

$\Rightarrow\forall i, l_i(\lambda_2-\lambda_1)=0\Rightarrow l_i=0$

$\Rightarrow k_1\alpha_1+\cdots+k_m\alpha_m=\mathbf{0}\Rightarrow\forall i, k_i=0$

于是命题得证 □

一个显然的推论是此结论的 n 个子空间的版本：

推论 3.1.1

$A\in\hom(V,V)$，$\lambda_1,\cdots,\lambda_n$ 是 A 的 n 个互不相同的特征值

如果 $\forall i,\alpha_{ir_1},\cdots,\alpha_{ir_m}\in V_{\lambda_i}$ 线性无关

那么向量组 $\{\alpha_{jr_k}\}$ 也线性无关

证明： 由数学归纳法易证 □

3. 容易注意到，矩阵的特征值和线性映射的特征值其实是一样的，正如下面的命题：

命题 3.1.3. 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$，$\{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一个基
$A \in \hom(V, V)$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$
那么，λ 是 A 的一个特征值，α 是 A 的一个隶属于 λ 的一个特征向量 $\Leftrightarrow$
λ 是 $\boldsymbol{A}$ 的一个特征值，并且 α 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的坐标 $\mathbf{x}$ 是 $\boldsymbol{A}$ 的一个隶属于 λ 的特征向量

证明： $A(\alpha) = \lambda\alpha$
$\Leftrightarrow \boldsymbol{A}\mathbf{x} = \lambda\mathbf{x}$（参见命题 2.3.23）
于是命题得证。 □

3.1.3 特征矩阵与特征多项式

前面我们研究了特征值的性质，接下来我们想知道：是否可以直接去寻找计算特征值的直接方法？事实上，是可以的，我们指出：特征值是特征多项式的一个根，而特征向量是对应映射的核的一个元素

先给出定义：

定义 3.1.3: 线性变换的特征多项式

设 $A \in \hom(V, V)$，我们称
$\chi_A(\lambda) = \det(\lambda I - A)$
为 A 的特征多项式

类似地，我们可以定义矩阵的特征矩阵和特征多项式：

定义 3.1.4: 矩阵的特征矩阵和特征多项式

设 $\boldsymbol{A} \in M_n(F)$，我们称 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 是 $\boldsymbol{A}$ 的特征矩阵
并称 $\chi_{\boldsymbol{A}}(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A})$ 是 $\boldsymbol{A}$ 的特征多项式

下面的性质指出了我们想要的结果：

1.

命题 3.1.4. 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$，$\{\alpha_1, \cdots, \alpha_n\}$

如果 $A \in \hom(V, V)$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$

那么我们断言：$\chi_A(\lambda) = \chi_{\boldsymbol{A}}(\lambda)$

证明： 这是显然的，因为矩阵的行列式的定义就是其矩阵的行列式 □

2.

命题 3.1.5. $\forall \boldsymbol{A} \in M_n(F), \chi_{\boldsymbol{A}} \in F[\lambda]$

证明： 设 $\boldsymbol{A} = (a_{ij}) \in M_n(F)$

那么，$\lambda \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}$

由行列式的置换展开可知，$\chi_{\boldsymbol{A}}(\lambda) \in F[\lambda]$ □

推论 3.1.2

$\forall A \in \hom(V, V), \dim V = n < \aleph_0, \chi_A(\lambda) \in F[\lambda]$

证明： 这是显然的，因为我们只需要任取一个基 $\{\alpha_1, \cdots, \alpha_n\}$，并利用 A 在此基下的矩阵 $\boldsymbol{A}$

利用前面线性映射与其矩阵的特征多项式相同的命题，即可得证。 □

定理 3.1.3: 线性变换的特征值即是特征多项式在域内的根

设 V 是一个 F 上的 n 维线性空间，$A \in \hom(V, V)$

3. 那么：λ 是 A 的一个特征值，$\alpha \in V_\lambda \Leftrightarrow \chi_A(\lambda) = 0, \alpha \in \ker(\lambda I - A)$

证明： λ 是 A 的一个特征值，$\alpha \in V_\lambda$

$\Leftrightarrow A(\alpha) = \lambda\alpha, V_\lambda \neq \{\mathbf{0}\}$

$\Leftrightarrow (\lambda I - A)(\alpha) = \mathbf{0}, \ker(\lambda I - A) \neq \{\mathbf{0}\}$

$\Leftrightarrow \operatorname{rank}(\lambda I - A) < n, \alpha \in \ker(\lambda I - A)$

$\Leftrightarrow \chi_A(\lambda) = 0, \alpha \in \ker(\lambda I - A)$ □

其自然推论是其矩阵版本：

推论 3.1.4

设 $\boldsymbol{A} \in M_n(F)$

那么：λ 是 $\boldsymbol{A}$ 的一个特征值，$\mathbf{x} \in F^n$ 是 $\boldsymbol{A}$ 从属于 λ 的特征向量 $\Leftrightarrow \chi_{\boldsymbol{A}}(\lambda) = 0, (\lambda \boldsymbol{I} - \boldsymbol{A})\mathbf{x} = \mathbf{0}, \mathbf{x} \neq \mathbf{0}$

这个推论也指出了一下结果：

推论 3.1.5

设 $\boldsymbol{A}, \boldsymbol{B} \in M_n(F), \boldsymbol{A} \sim \boldsymbol{B}$

那么 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的特征值，并且 $\chi_{\boldsymbol{A}}(\lambda) = \chi_{\boldsymbol{B}}(\lambda)$

我们之后不再完全讨论矩阵版本的特征值相关定理，因为上面的结果已经说明了：相似矩阵，以及相互对应的矩阵和映射在特征值理论中并无区别

4.

命题 3.1.6. 设 V 是一个 F 上的 n 维线性空间，$A \in \hom(V, V)$

那么：

$$\chi_A(\lambda) = \lambda^n - \operatorname{tr}(A)\lambda^{n-1} + \cdots + (-1)^{n-k}\left(\sum_{j_1<\cdots<j_{n-k}} A^{j_1,\cdots,j_{n-k}}_{j_1,\cdots,j_{n-k}}\right)\lambda^k + \cdots + (-1)^n \det(A) \tag{3.4}$$

证明： 事实上，λ^k 项即是下面的行列式求和：$\displaystyle\sum_{j'_1<\cdots<j'_k} \begin{vmatrix} -a_{11} & \cdots & 0 & \cdots & 0 & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ -1 & \cdots & \lambda & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ -1 & \cdots & 0 & \cdots & \lambda & \cdots & 1 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & 0 & \cdots & 0 & \cdots & -a_{nn} \end{vmatrix}$ 其中行列式的 $j'_1, \cdots, j'_k$ 列为仅在第 j_i 个元素为 λ，其他位置都是 0 的列。

这是因为，如果想要产生 λ^k，那么求和中的每一项必须选取对角线上的 k 个元素；同时，如果在进一步展开中选取了 $-a_{ii}$ 而不是 λ，那么会导致次数降低

注意到：$\begin{vmatrix} -a_{11} & \cdots & 0 & \cdots & 0 & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ -1 & \cdots & \lambda & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ -1 & \cdots & 0 & \cdots & \lambda & \cdots & 1 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & 0 & \cdots & 0 & \cdots & -a_{nn} \end{vmatrix}$

$= (-1)^{(j'_1+\cdots+j'_k)+(j'_1+\cdots+j'_k)} \det(diag\{\lambda,\cdots,\lambda\})(-1)^{n-k} A^{j_1,\cdots,j_{n-k}}_{j_1,\cdots,j_{n-k}}$

这个结果只需要对 $j'_1,\cdots,j'_k$ 列展开即可得出，其中 $(-1)^{n-k}$ 是将 $-a_{ij}$ 前面的负号提出去得到的，而 $j_1,\cdots,j_{n-k}$ 是与 $j'_1,\cdots,j'_k$ 互补的有限序列

于是命题得证。 □

其推论是以下结论：

推论 3.1.6

设 F 是一个代数闭域，V 是一个 F 上的 n 维线性空间，$A \in \hom(V,V)$

设 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征值（$\chi_A(\lambda)$ 中的重根按重数计算）

那么有：

$$\prod_{i=1}^{n} \lambda_i = \det(A) \tag{3.5}$$

$$\sum_{i=1}^{n} \lambda_i = \mathrm{tr}(A) \tag{3.6}$$

证明： 对 $\chi_A(\lambda)$ 作唯一分解：

$\chi_A(\lambda) = \prod\limits_{i=1}^{n}(\lambda-\lambda_i)$

$= \lambda^n - (\lambda_1+\cdots+\lambda_n)\lambda^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i$

结合前面的命题既得。 □

请注意：这个命题必须要求 F 代数闭，否则可能 $\chi_A(\lambda)$ 的根并非全体特征值，而导致论证失效。

3.2 线性变换的对角标准型

本节中我们构建对角标准型的理论。对角标准型的理论是完全自然的，我们无需更多讨论，只需给出定义并直接推导充要条件即可。

3.2.1 对角标准型的定义

定义 3.2.1: 线性映射的对角标准型

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

如果存在一个基 $\{\alpha_1, \cdots, \alpha_n\}$，使 $\mathcal{A}$ 在其上的矩阵 $\boldsymbol{A}$ 是一个对角矩阵

那么我们称 $\boldsymbol{A}$ 为 $\mathcal{A}$ 的对角标准型，此时称 $\mathcal{A}$ 可对角化。

类似地，我们可以定义矩阵的对角标准型

定义 3.2.2: 矩阵的对角标准型

设 $\boldsymbol{A} \in M_n(F)$，如果存在 $\boldsymbol{P} \in GL_n(F)$ 和对角矩阵 $\boldsymbol{D} \in M_n(F)$，使得

$\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{D}\boldsymbol{P}$

那么我们称 $\boldsymbol{D}$ 为 $\boldsymbol{A}$ 的对角标准型，此时称 $\boldsymbol{A}$ 可对角化。

容易看出，线性映射的对角标准型与矩阵的对角标准型是等价的。特别地，如果认为 $\boldsymbol{A}$ 是 $\mathcal{A} \in \hom(F^n, F^n)$ 在 $\{\mathbf{e}_i\}$ 下的矩阵，那么容易发现：其实 $\boldsymbol{P}^{-1}$ 的列向量就是 $\mathcal{A}$ 的特征向量。

3.2.2 线性变换可对角化的条件

观察对角矩阵的结构，既得以下显然的条件：

定理 3.2.1: 线性映射可对角化的条件 (1)

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

那么：$\mathcal{A} \in \hom(V, V)$ 可对角化 $\Leftrightarrow$ 存在一个由 $\mathcal{A}$ 的特征向量组成的 V 的基

证明： $\mathcal{A} \in \hom(V, V)$ 可对角化

$\Leftrightarrow$ 存在一个基 $\{\alpha_1, \cdots, \alpha_n\}$，使得 $\mathcal{A}$ 在其上的矩阵为 $diag\{\lambda_1, \cdots, \lambda_n\}$

$\Leftrightarrow \mathcal{A}(\alpha_1, \cdots, \alpha_n) = diag\{\lambda_1, \cdots, \lambda_n\}(\alpha_1, \cdots, \alpha_n)$

$\Leftrightarrow A(\alpha_1,\cdots,\alpha_n)=(\lambda_1\alpha_1,\cdots,\lambda_n\alpha_n)$

$\Leftrightarrow \lambda_i$ 是 A 的特征值，$\alpha_i\in V_{\lambda_i}$

$\Leftrightarrow$ 存在一个由 A 的特征向量组成的 V 的基 □

推论 3.2.2: 线性映射可对角化的条件 (2)

设 V 是一个 F 上的一个线性空间，$\dim V=n<\aleph_0$

那么：$A\in\hom(V,V)$ 可对角化 $\Leftrightarrow$ 存在 n 个线性无关的特征向量

证明: 这是显然的，因为 V 中的 n 个线性无关的向量一定组成一个基 □

定理 3.2.3: 线性映射可对角化的条件 (3)

设 V 是一个 F 上的一个线性空间，$\dim V=n<\aleph_0$

那么：$A\in\hom(V,V)$ 可对角化 $\Leftrightarrow$ A 的全体不同的特征值 $\lambda_1,\cdots,\lambda_s$，满足 $V_{\lambda_1},\cdots,V_{\lambda_s}$ 的基的并集是 V 的基

证明: 必要性由之前的命题是显然的。

接下来证明充分性。

如果 A 可对角化，那么，它一定能选取一个由特征向量组成的基 $\{\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s}\}$，其中 $\alpha_{ij}\in V_{\lambda_i},r_1+\cdots+r_s=n$

我们只需证明，其实 $\{\alpha_{i1},\cdots,\alpha_{ir_i}\}$ 就是 V_{λ_i} 的一个基

首先，它一定线性无关，因为它是 V 的一个基的一个子集。不妨假设它不是 V_{λ_i} 的基，那么，一定 $\exists\beta\in V_{\lambda_i}$，使得它不能被 $\{\alpha_{i1},\cdots,\alpha_{ir_i}\}$ 线性表出

那么，一定有 $\{\alpha_{i1},\cdots,\alpha_{ir_i},\beta\}$ 线性无关。因为如果它线性相关，那么 $\{\alpha_{i1},\cdots,\alpha_{ir_i}\}$ 一定是它的一个极大线性无关组，这使得 β 可以被 $\{\alpha_{i1},\cdots,\alpha_{ir_i}\}$ 线性表出，矛盾。

但是，我们之前已经指出，若干个特征子空间的线性无关向量的并集必定也是线性无关的，这说明 $\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s},\beta$ 线性无关。但是 $\dim V=n$，其中的 $n+1$ 个向量不可能线性无关，矛盾。

于是命题得证 □

其推论是接下来的两个条件：

推论 3.2.4: 线性映射可对角化的条件 (4)

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

那么：$A \in \hom(V, V)$ 可对角化 $\Leftrightarrow$ 如果 $\lambda_1, \cdots, \lambda_s$ 是 A 的全体不同的特征值，那么

$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s} = V$

证明： 这是显然的，因为前面的命题指出了 V 的一个基正是 V_{λ_i} 的基组成的，而这正是直和的充要条件。 □

推论 3.2.5: 线性映射可对角化的条件 (5)

V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

那么：$A \in \hom(V, V)$ 可对角化 $\Leftrightarrow$ 如果 $\lambda_1, \cdots, \lambda_n$ 是 A 的全体不同的特征值，那么

$\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_s} = n$

证明： 充分性由上面的命题是显然的。

必要性只需注意到前面命题的结果，即存在一个 V 的基由全体特征子空间的基组成即可。 □

至此，我们基本上得到了全部结果，最后我们将看到一个最终结果，它也是最常用的充要条件

其实，我们已经注意到：如果特征子空间的维数之和不足，那么它就不可对角化。我们的一个天然的猜想是：特征子空间的维数是不是受限于特征值的重数？如果是的话，其实只需要填满这个重数就可以对角化。

现在给出一个引理

引理 3.2.6: 几何重数小于代数重数

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

设 $\lambda_i \in F$ 是 $A \in \hom(V, V)$ 的一个特征值，并且其特征多项式有分解：

$\chi_A(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{r_i}$

那么我们断言：$\dim V_{\lambda_i} \leqslant r_i$

其中 $\dim V_{\lambda_i}$ 也常常称为 λ_i 的几何重数，r_i 称为 λ_i 的代数重数

证明： 不妨设 $\dim V_{\lambda_i} = d$

取 V_{λ_i} 的一个基 $\{\alpha_1, \cdots, \alpha_d\}$，并补全为 V 的一个基 $\{\alpha_1, \cdots, \alpha_d, \cdots, \alpha_n\}$

考虑 A 在此基上的矩阵 $\boldsymbol{A}$，它一定具备以下形式：

$\boldsymbol{A} = \begin{pmatrix} \lambda_i \boldsymbol{I}_d & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}$

因为 A 在 $\{\alpha_1, \cdots, \alpha_d\}$ 上的作用即是按照相同的特征值 λ_i 进行纯量乘法

那么，特征矩阵 $\lambda \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} (\lambda - \lambda_i)\boldsymbol{I}_d & -\boldsymbol{B} \\ \boldsymbol{O} & \lambda \boldsymbol{I}_{n-d} - \boldsymbol{C} \end{pmatrix}$

于是有：$\chi_A(\lambda) = \det((\lambda - \lambda_i)\boldsymbol{I}_d)\det(\boldsymbol{I}_{n-d} - \boldsymbol{C})$

$= (\lambda - \lambda_i)^d \det(\boldsymbol{I}_d)\det(\boldsymbol{I}_{n-d} - \boldsymbol{C})$

因此一定有 $d \leqslant r_i$，命题得证 □

那么有以下命题：

定理 3.2.7: 线性映射可对角化的条件 (6)

设 F 是一个代数闭域，V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

那么 $A \in \hom(V, V)$ 可对角化 $\Leftrightarrow$ 如果 $\chi_A(\lambda) = \prod_{i=1}^{s}(\lambda - \lambda_i)^{r_i}$，那么 $\dim V_{\lambda_i} = r_i$

证明: 必要性是显然的，因为这正是前面的命题给出的结果

充分性只需注意到，如果可对角化必定有 $\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_s}$，但是引理已经指出 $\dim V_{\lambda_i} \leqslant r_i$，又因为 $r_1 + \cdots + r_s = n$，那么一定有 $\dim V_{\lambda_i} = r_i$

于是命题得证。 □

对角标准型的性质就是对角矩阵的性质，它们都很显然故无需阐述。可以看出来，对角标准型的条件是极其严苛的——这也是它优良性质的缘由。

3.3 线性变换的根子空间、不变子空间

我们转向不可对角化的线性映射的表示。

我们观察可对角化的条件 (4)，它的另一种表述是：

$V = \ker(\lambda_1 I - A) \oplus \cdots \oplus \ker(\lambda_s I - A)$

但是，我们知道，直和并不总是成立。我们的一个想法是：如果我们调整 $\lambda_i I - A$ 的幂，是否能让他构成直和呢？如果能构成直和，这个幂有办法快速求出来吗？

一个相当胡扯的想法是：特征多项式的分解 $\chi_A(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_s)^{r_s}$ 似乎很符合我们的结构

我们从特征多项式在映射上的作用，及其分解入手。

3.3.1 线性变换的多项式

此处我们希望讨论的，其实是一个具体的映射的多项式，也就是以下定义：

定义 3.3.1: 线性映射的行列式

设 F 是一个域，$A \in \hom(V,V)$

我们定义：

$$F[A] := \{\sum_{k=0}^{n} a_k A^k | a_k \in F, n \in \mathbb{N}\} \tag{3.7}$$

称为 A 的多项式环，其元素称为 A 的多项式

一个显然的事实是：$F[A]$ 是 $\hom(V,V)$ 的子环

接下来探讨线性变换的多项式的性质。

1. Hamilton-Caylay 定理

定理 3.3.1: Hamilton-Caylay 定理

设 V 是一个 F 上的 n 维线性空间，$A \in \hom(V,V)$

那么有：

$$\chi_A(A) = 0 \tag{3.8}$$

证明： 设 M 是一个 $F[A]$ 上的一个模，其元素和 V 一致，保持 V 中的加法，而其纯量乘法则为：

$f(A) \cdot \alpha := f(A)(\alpha), f(A) \in F[A], \alpha \in V$

那么容易注意到，此时 $A \in \hom(M,M)$，即成为 M 的一个自同态

接下来我们尝试给出 $\det(\lambda I - A)$ 在自同态下的一个表示。此处，我们尝试借助矩阵来完成这一步

由于 M 其实和 V 有相同的元素，并且 F 本就是 $F[A]$ 的子环，因此只需取 V 的一个基 $\{e_1, \cdots, e_n\}$，它既是 M 的生成元

那么，一定能写出 A 的在这一组生成元下的矩阵 $\boldsymbol{A} \in M_n(F[A])$

那么，$\det(\lambda I - A) = \det(\lambda \boldsymbol{I} - \boldsymbol{A})$（请注意，此处的 $\boldsymbol{I}, \boldsymbol{A}$ 是环上的矩阵，而非域上的矩阵）

此时只需考察 $\chi_A(A)=\det(A\boldsymbol{I}-\boldsymbol{A})$（由于此时矩阵已经是建立在环 $F[A]$ 上的了，A 成为了纯量，可以直接代入）

记 $\boldsymbol{P}=A\boldsymbol{I}-\boldsymbol{A}, adj(\boldsymbol{P})=(\frac{1}{|\boldsymbol{P}|}\boldsymbol{P}^{\mathbb{N}-\{j\}}_{\mathbb{N}-\{i\}})$，并注意到：

$\chi_A(A)(e_1,\cdots,e_n)=(e_1,\cdots,e_n)\det(\boldsymbol{P})=(e_1,\cdots,e_n)\boldsymbol{P}adj(\boldsymbol{P})$

$(A(e_1,\cdots,e_n)-\boldsymbol{A}(e_1,\cdots,e_n))adj(\boldsymbol{P})=0$

于是命题得证。 □

这个定理看起来非常令人惊讶：特征多项式竟然可以直接把映射本身化为零。

这个定理的一个常见的伪证是：直接把 A 代入，得到 $\det(AI-A)=\det(0)=0$。但是，其实 λ 是一个纯量，并不能将映射代入。

我们此处的证明正是基于这种想法进行了一个修改：将 F 视为环 $F[A]$ 的子环，从而实现了正确的代入。

2.

命题 3.3.1. 设 $A\in\hom(V,V)$,$f(x),f_1(x),f_2(x)\in F[x]$,$f(x)=f_1(x)f_2(x)$,$(f_1(x),f_2(x))=1$

那么有：$\ker f(A)=\ker f_1(A)\oplus\ker f_2(A)$

证明： 我们首先证明，$\ker f_1(A)$ 和 $\ker f_2(A)$ 的确是 $\ker f(A)$ 的子空间

$\forall\alpha\in\ker f_1(A)$

$\Rightarrow f_1(A)(\alpha)=\mathbf{0}\Rightarrow f(A)(\alpha)=\left(f_2(A)f_1(A)\right)(\alpha)=\mathbf{0}\Rightarrow\alpha\in\ker f(A)\Rightarrow\ker f_1(A)\subseteq\ker f(A)$

同理可证 $\ker f_2(A)\subseteq\ker f(A)$

于是我们证明了 $\ker f_1(A)+\ker f_2(A)\subseteq\ker f(A)$。现在我们证明 $\ker f(A)\subseteq\ker f_1(A)+\ker f_2(A)$

$\forall\alpha\in\ker f(A)$

由于 $(f_1(x),f_2(x))=1$，一定存在 $u(x),v(x)\in F[x],u(x)f_1(x)+v(x)f_2(x)=1$

于是有：$\alpha=I(\alpha)=\left(u(A)f_1(A)+v(A)f_2(A)\right)(\alpha)=\left(u(A)f_1(A)\right)(\alpha)+\left(v(A)f_2(A)\right)(\alpha)$

记 $\alpha_1=\left(u(A)f_1(A)\right)(\alpha),\alpha_2=\left(v(A)f_2(A)\right)(\alpha)$

注意到：$f_2(A)(\alpha_1)=\left(f_2(A)u(A)f_1(A)\right)(\alpha)=\left(u(A)f(A)\right)(\alpha)=u(A)(\mathbf{0})=\mathbf{0}\Rightarrow\alpha_1\in\ker f_2(A)$

同理可证 $\alpha_2\in\ker f_1(A)$。于是我们证明了 $\ker f(A)=\ker f_1(A)+\ker f_2(A)$

我们最后证明它们的确构成直和。

$\forall\beta \in \ker f_1(A) \cap \ker f_2(A)$

注意到:$\beta = I(\beta) = (u(A)f_1(A) + v(A)f_2(A))\,(\beta) = (u(A)f_1(A))\,(\beta)+(v(A)f_2(A))\,(\beta) = \mathbf{0}$

于是命题得证。 □

这个命题的直接推论是：

推论 3.3.2: –

设 $A \in \hom(V,V)$，$f(x), f_1(x), \cdots, f_s(x) \in F[x]$，$f(x) = \prod_{i=1}^{s} f_i(x)$，$\forall i,j,(f_i(x), f_j(x)) = 1$

那么有：$\ker f(A) = \bigoplus_{i=1}^{s} \ker f_i(A)$

至此，其实我们已经可以看出来一些方向了。我们已经证明了多项式如果分解为多个互素多项式，那么他们的核也可以构成直和。我们又注意到，任意映射在特诊多项式下为零。

那么，其实特征多项式的唯一分解一定可以直和为零映射的核——即线性空间本身，这正是我们所期望的结果

3.3.2 线性变换的根子空间

我们将这种理想的子空间为根子空间，因为它的常数是从特征多项式求出来的

定义 3.3.2: 根子空间

设 F 是一个代数闭域，V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$

$\chi_A(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{r_i}$

那么我们称 $\ker(A - \lambda_i I)^{r_i}$ 为 A 的一个根子空间（又称广义特征子空间）

我们现在正式叙述我们观察到的结果

定理 3.3.3: 子空间分解定理

设 F 是一个代数闭域，V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$

$\chi_A(\lambda) = \prod_{i=1}^{s}(\lambda-\lambda_i)^{r_i}$

那么有：

$$V = \bigoplus_{i=1}^{s} \ker(A-\lambda_i I)^{r_i} \tag{3.9}$$

至此，我们证实了，代数闭域上的任意有限维线性映射，都可以形成一个根子空间分解。

接下来我们转向矩阵表示。在之前的对角标准型中，从特征值的定义可以看出来，每一个特征子空间表示为一个对角线上的元素。那么，现在我们希望：A 在根子空间上的限制，都可以写成一个矩阵；特别地，如果每一个限制所属的子空间不同，此时我们便有可能写成分块对角矩阵

3.3.3 线性变换的不变子空间

一个首要的问题是：$A|_{\ker(A-\lambda_i I)^{r_i}}$ 到底能不能写成一个方阵的形式。如果它可以，那么，我们将不仅可以把映射写成分块对角矩阵，而且每一个块的具体形式可以借助线性变换来求得。

总之，我们希望验证：是否任意代数闭域上的线性变换在根子空间上的限制是一个线性变换，而且这些线性变换在指定基下的矩阵，恰好可以分块地组成线性变换的矩阵

我们首先给出不变子空间的定义，它其实就是使得变换的限制依旧为变换的子空间

定义 3.3.3: 不变子空间

设 V 是 F 上的一个线性空间，$A \in \hom(V,V)$

如果 V 的子空间 $W \subseteq V$ 满足：

$\forall \alpha \in W, A(\alpha) \in W$

那么我们称 W 是 A 的一个不变子空间

显然，V 和 $\{0\}$ 都是 A 的一个不变子空间，它们一般被称为是平凡的不变子空间

接下来我们给出一些性质

1.

命题 3.3.2. 设 V 是 F 上的一个线性空间，$A \in \hom(V,V)$，λ 是 A 的一个特征值

那么 $\ker A, \operatorname{Im} A, V_\lambda$ 是 A 的不变子空间

证明： $\forall \alpha \in \ker A, A(\alpha) = \mathbf{0} \in \ker A$

$\forall \alpha \in \operatorname{Im} A, A(\alpha) \subseteq \operatorname{Im} A$

$\forall \alpha \in V_\lambda, A(\alpha) = \lambda\alpha \in V_\lambda$ □

2.

命题 3.3.3. 设 V 是 F 上的一个线性空间，$A, B \in \hom(V, V), AB = BA$，$\lambda$ 是 A 的一个特征值

那么 $\ker A, \operatorname{Im} A, V_\lambda$ 是 B 的不变子空间

证明： $\forall \alpha \in \ker A, A(B(\alpha)) = B(A(\alpha)) = B(\mathbf{0}) = \mathbf{0} \Rightarrow B(\alpha) \in \ker A$

$\forall A(\alpha) \in \operatorname{Im} A, B(A(\alpha)) = A(B(\alpha)) \in \operatorname{Im} A0$

$\forall \alpha \in V_\lambda, A(B(\alpha)) = B(A(\alpha)) = B(\lambda\alpha) = \lambda B(\alpha) \Rightarrow B(\alpha) \in V_\lambda$

于是命题得证。 □

3.

命题 3.3.4. 一个线性变换的两个不变子空间的和与交，也是不变子空间

证明： 设 $A \in \hom(V, V)$，W, U 是 A 的不变子空间

注意到：$\forall \alpha + \beta \in W + U, \alpha \in W, \beta \in U, A(\alpha + \beta) = A(\alpha) + A(\beta) \in W + U$

$\forall \gamma \in W \cap U, A(\gamma) \in W, A(\gamma) \in U \Rightarrow A(\gamma) \in W \cap U$

于是命题得证 □

4. 最后一条性质，是我们所希望的结果：即如果把映射限制在一系列构成直和的不变子空间，那么矩阵也可以自然构成分块

命题 3.3.5. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0, A \in \hom(V, V)$，那么：$V = W_1 \oplus \cdots \oplus W_s$，其中 W_i 是 A 的不变子空间，且 $A|_{W_i}$ 在基 $B_i = \{\alpha_{i1}, \cdots, \alpha_{ir_i}\}$ 下的矩阵为 $\boldsymbol{A}_i$

$\Leftrightarrow A$ 在 $\{\alpha_{11}, \cdots, \alpha_{1r_1}, \cdots, \alpha_{s1}, \cdots, \alpha_{sr_s}\}$ 下的矩阵为 $\begin{pmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_s \end{pmatrix}$

证明： 先证充分性。

首先，由于 $W_1, \cdots, W_s$ 构成直和，因此 $B_1 \cup \cdots \cup B_s$ 的确是 V 的一个基

因为 $A|_{W_i}$ 在 B_i 下的矩阵为 $\boldsymbol{A}_i$，因此有：$A|_{W_i}(\alpha_{ij}) = \sum\limits_{k=1}^{r_i} \boldsymbol{A}_i(k;j)\alpha_{ik}$

考察 A 在同样的向量上的作用：

$A(\alpha_{ij}) = A|_{W_i}(\alpha_{ij}) = \sum\limits_{k=1}^{r_i} \boldsymbol{A}_i(k;j)\alpha_{ik}$

由基的线性无关性可知，以上是 $A(\alpha_{ij})$ 的唯一表示。于是，A 在基 $B_1 \cup \cdots \cup B_s$ 下的矩阵 $\boldsymbol{A}$ 分块后的第 i 列，当且仅当在 i 行为 $\boldsymbol{A}_i$，其他均为 $\boldsymbol{O}$

于是一定有 $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_s \end{pmatrix}$，充分性得证。

接下来证明必要性。

取 $W_i = \mathrm{span}(\alpha_{i1}, \cdots, \alpha_{ir_i})$

由直和的充要条件可知，的确有 $V = W_1 \oplus \cdots \oplus W_s$

且由分块矩阵的结构可见：

$(A|_{W_i})(\alpha_{i1}, \cdots, \alpha_{ir_i}) = (\alpha_{i1}, \cdots, \alpha_{ir_i})\boldsymbol{A}$

于是命题得证。 □

5. 我们本节提出的最后一个命题，是指出根子空间也是不变子空间，自然可以应用上面关于分块矩阵的命题

命题 3.3.6. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，

如果 $A \in \hom(V, V)$ 的特征多项式在 F 内有以下分解：

$\chi_A(\lambda) = \prod\limits_{i=1}^{s} (\lambda - \lambda_i)^{r_i}$

那么 A 的根子空间 $\ker(A - \lambda_i I)^{r_i}$ 是 A 的不变子空间

证明： 这其实是显然的，我们只需要注意到：$A(A - \lambda_i I) = A^2 - \lambda_i A = (A - \lambda_i I)A$

即 A 和 $A - \lambda_i I$ 是可交换的

那么，按照数学归纳法，容易证明 $A(A - \lambda_i I)^{r_i} = (A - \lambda_i I)^{r_i} A$

于是有：

$\forall \alpha \in \ker(A - \lambda_i I)^{r_i}, (A - \lambda_i I)^{r_i}(\alpha) = \mathbf{0}$

$\Rightarrow A(A - \lambda_i I)^{r_i}(\alpha) = \mathbf{0} \Rightarrow (A - \lambda_i I)^{r_i}(A(\alpha)) = \mathbf{0}$

$\Rightarrow A(\alpha) \in \ker(A - \lambda_i I)^{r_i}$

于是命题得证。 □

3.4 线性变换的最小多项式

前面我们指出了：只要特征多项式可分解为一次因式的乘积，那么线性空间就是根子空间的直和；进一步，我们只需求出线性变换在根子空间上的限制的矩阵，就可以对角地写出线性变换的矩阵

本节中我们依旧选取多项式作为一个出发点，试图找到这些分块的具体形式。

3.4.1 最小多项式的定义

定义 3.4.1: 最小多项式

设 V 是 F 上的一个线性空间，$A \in \hom(V, V)$

如果 $f(x) \in F[x], f \neq 0$ 满足：

1. $f(A) = 0$
2. $\forall g \in F[x] \wedge g(A) = 0, \deg f \leqslant \deg g$
3. f 是首一多项式

那么我们称 f 是 A 的最小多项式。

3.4.2 最小多项式的性质

1. 最小多项式是唯一的

命题 3.4.1. 任意线性变换的最小多项式是唯一的

证明: 不妨假设 $m_1(x), m_2(x)$ 均是 $A \in \hom(V, V)$ 的最小多项式，且 $m_1(x), m_2(x)$ 不妨假设 $m_1(x) = \sum\limits_{k=0}^{n} a_k x^k, m_2(x) = \sum\limits_{k=0}^{n} b_k x^k$（由最小多项式的定义可知，它们一定同次）

由最小多项式的定义可知，一定有 $a_n = b_n = 1$

容易注意到，$m_1(x) - m_2(x) = \sum\limits_{i=0}^{n-1} (a_k - b_k) x^k$ 是零化的，因为 $m_1(A) - m_2(A) = 0 - 0 = 0$

那么这是不可能的，因为此时，只需找出 $m_1(x) - m_2(x)$ 相伴的首一多项式 $m(x)$，它正是一个最小多项式，并且它的次数比 $m_1(x), m_2(x)$ 低。 □

2.

命题 3.4.2. 设 $m(x)$ 是 $A \in \hom(V, V)$ 的最小多项式，那么：

$p(x) \in F[x], p(A) = 0 \Leftrightarrow m(x) \mid p(x)$

证明: 先证充分性。

对 $p(x), m(x)$ 作带余除法:

$\exists q(x), r(x) \in F[x], \deg r(x) < \deg m(x), p(x) = q(x)m(x) + r(x)$

$\Rightarrow p(A) = q(A)m(A) + r(A) \Rightarrow 0 = 0 + r(A) \Rightarrow r(A) = 0$

事实上，此时只可能有 $r(A) = 0$，因为我们已经指出，最小多项式是次数最低的零化多项式。

那么此时有 $p(x) = q(x)m(x)$，充分性得证。

再证必要性。

这其实是显然的，因为如果有 $p(x) = q(x)m(x)$，此时 $p(A) = q(A)m(A) = 0$，于是必要性得证。

于是命题得证。 □

这个命题指出了一个事实：零化多项式环其实不仅是一个子环，而且是个理想；特别的，它是由最小多项式生成的主理想 $(m(x))$

3. 接下来我们探讨有限维空间上的线性映射的最小多项式。这样的线性映射的最小多项式一定是存在的，因为我们知道特征多项式正好就是一个零化多项式

命题 3.4.3. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$，$m(x)$ 是 A 的最小多项式

那么 $m(x)$ 和 $\chi_A(x)$ 在 F 中有相同的根（但是重数不一定一致）

证明: 由之前的命题可知，$m(x) \mid \chi_A(x)$，于是 $m(x)$ 的任意根都是 $\chi_A(x)$ 的根。

于是我们只需证明 $\chi_A(x)$ 在 F 中的任意根都是 $m(x)$ 的根。

设 $m(x) = \sum\limits_{k=0}^{n} a_k x^k$

我们之前已经指出，$\chi_A(x)$ 在 F 的根其实就是 A 的特征值。于是不妨设 $\chi_A(\lambda) = 0, \lambda \in F$

$\Rightarrow \forall \alpha \in V_\lambda, A(\alpha) = \lambda\alpha$

$\Rightarrow \mathbf{0} = 0(\alpha) = m(A)(\alpha) = (a_0 I + a_1 A + \cdots + a_n A^n)(\alpha)$

$= a_0\alpha + a_1 A(\alpha) + \cdots + a_n A^n(\alpha) = a_0\alpha + a_1\lambda\alpha + \cdots + a_n\lambda^n\alpha$

$= (a_0 + a_1\lambda + \cdots + a_n\lambda^n)\alpha = m(\lambda)\alpha$

又 $\lambda \neq \mathbf{0}$，于是一定有 $m(\lambda) = 0$。

于是命题得证。 □

4.

命题 3.4.4. 设 $A \in \hom(V,V), V = W_1 \oplus \cdots \oplus W_s$，其中 W_i 是 A 的不变子空间那么如果 $m_i(x)$ 是 $A|_{W_i}$ 的最小多项式,那么 A 的最小多项式 $m(x)$ 就是 $m_1(x),\cdots,m_s(x)$ 的首一最小公倍式。

证明: 设 $q(x)$ 是 $m_1(x),\cdots,m_s(x)$ 的首一最小公倍式。

我们其实只需证明 $q(x) \sim m(x)$

先证明 $q(x) \mid m(x)$

考虑 $\forall \alpha_i \in W_i, m(A|_{W_i})(\alpha_i) = m(A)(\alpha_i) = \mathbf{0}$

由于 α_i 的选取是任意的，因此一定有 $m(A|_{W_i}) = 0$，即 $m(x)$ 也是 $A|_{W_i}$ 的零化多项式。

那么，由之前的命题一定有 $m_i(x) \mid m(x)$。

于是 $m(x)$ 也是 $m_1(x),\cdots,m_s(x)$ 的公倍式。由于 $q(x)$ 是他们的最小公倍式，一定有 $q(x) \mid m(x)$

接下来证明 $m(x) \mid q(x)$。

只需证明 $q(A) = 0$ 即可。设 $\forall \alpha \in V, \alpha = \alpha_1 + \cdots, \alpha_s, \alpha_i \in W_i$

又因为 $q(x)$ 是最小公倍式，因此一定有 $q(x) = h_i(x)m_i(x)$

注意到：$q(A)(\alpha) = q(A)\left(\sum\limits_{i=1}^{s} \alpha_i\right) = \sum\limits_{i=1}^{s} q(A)\left(\alpha_i\right)$

$= \sum\limits_{i=1}^{s} q(A|_{W_i})\left(\alpha_i\right) = \sum\limits_{i=1}^{s} h_i(A|_{W_i})m_i(A|_{W_i})\left(\alpha_i\right) = \mathbf{0}$

因为 α 是随意选取的，因此必须有 $q(A) = 0$。那么按照前面的命题，有 $m(x) \mid q(x)$

于是我们证明了 $m(x) \sim q(x)$。又两者都是首一的，因此必须有 $m(x) = q(x)$

于是命题得证。 □

5.

命题 3.4.5. $A \in \hom(V,V)$ 是一个幂零指数为 l 的幂零多项式 $\Leftrightarrow$ A 的最小多项式 $m(x) = x^l$

证明: 先证明充分性，注意到：$A^l = 0$，因此 x^l 是 A 的零化多项式。

那么一定有 $m(x) \mid x^l$。但是我们知道，$\forall n < l, A^n \neq 0$，因此只可能 $m(x) = x^l$

必要性是显然的，因为如果 $m(x) = x^l$，那么一定有 $A^l = 0$，且 $\forall n < l, A^n \neq 0$

于是命题得证。 □

6.

命题 3.4.6. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V,V)$，$m(x)$ 是 A 的最小多项式

那么 A 可对角化 $\Leftrightarrow m(x)$ 可在 F 中分解为互不相同的一次因式的乘积

证明： 设 A 的全体互不相同的特征值为 $\lambda_1, \cdots, \lambda_s$

A 可对角化 $\Leftrightarrow V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$

注意到，$\forall \alpha \in \ker(A-\lambda_i I), (A-\lambda_i I)|_{\ker(A-\lambda_i I)}(\alpha) = A(\alpha) - \lambda_i\alpha = \lambda_i\alpha - \lambda_i\alpha = \mathbf{0}$

$\Rightarrow x - \lambda_i$ 是 $A|_{\ker(A-\lambda_i I)}$ 的最小多项式（因为 $x-\lambda_i$ 是零化多项式，而其唯一非零低阶因子 a 不是零化多项式）

$\Leftrightarrow A$ 的零化多项式为 $(x-\lambda_1)\cdots(x-\lambda_s)$

因此 A 可对角化 $\Leftrightarrow m(x)$ 可在 F 中分解为互不相同的一次因式的乘积 □

其直接推论是：

推论 3.4.1: 幂零指数大于 1 的幂零变换不可对角化

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V,V)$ 是一个幂零指数为 l 的幂零变换

那么，如果 $l > 1$，则 A 不可对角化

证明： 这是显然的，因为此时 A 的最小多项式为 x^l，无法分解为不同的一次因式的乘积。 □

推论 3.4.2: 幂等变换可对角化

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V,V)$ 是一个幂等变换

那么 A 一定可对角化

证明： 这是显然的，因为只需注意到 $A^2 = A \Rightarrow A(A-I) = 0 \Rightarrow x(x-1)$ 是 A 的零化多项式

那么，A 的最小多项式作为 $x(x-1)$ 的因子，一定可以分解为不同的一次因式的乘积

于是命题得证 □

7. 最后，我们给出我们所需要的最终成果：

命题 3.4.7. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V,V)$，$\lambda_1, \cdots, \lambda_s$ 是 A 的全体互不相同的特征值

$m(x) = (x-\lambda_1)^{t_1} \cdots (x-\lambda_s)^{t_s}$ 是 A 的最小多项式

那么，$A|_{\ker(A-\lambda_i I)^{t_i}}$ 的最小多项式为 $(x-\lambda_i)^{t_i}$，

且 $(A-\lambda_i I)|_{\ker(A-\lambda_i I)^{t_i}}$ 是一个幂零指数为 t_i 的幂零变换

证明： 因为最小多项式和特征多项式同根，此时一定有分解 $\chi_A(x) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_s)^{r_s}$

我们之前已经证明了：$V = \ker(A-\lambda_1 I)^{r_1} \oplus \cdots \oplus \ker(A-\lambda_s I)^{r_s}$

注意到：$\forall \alpha \in \ker(A-\lambda_i I)^{r_i}, (A-\lambda_i I)^{r_i}|_{\ker(A-\lambda_i I)^{r_i}}(\alpha) = \mathbf{0}$

因此，$(x-\lambda_i)^{r_i}$ 是 $A|_{\ker(A-\lambda_i I)^{r_i}}$ 的零化多项式。不妨假设 $A|_{\ker(A-\lambda_i I)^{r_i}}$ 的最小多项式 $m_i(x) = (x-\lambda_i)^{l_i}, l_i \leqslant r_i$

那么，按照之前的命题，A 的最小多项式一定是 $m_1(x), \cdots, m_s(x)$ 的首一最小公倍式，又因为 $m_1(x), \cdots, m_s(x)$ 两两互素，因此一定有 $m(x) = (x-\lambda_1)^{l_1} \cdots (x-\lambda_s)^{l_s}$

但是，我们已经假设了 $m(x) = (x-\lambda_1)^{t_1} \cdots (x-\lambda_s)^{t_s}$，因此一定有 $l_i = t_i$

于是我们证明了 $A|_{\ker(A-\lambda_i I)^{r_i}}$ 的最小多项式为 $(x-\lambda_i)^{t_i}$

此时只需注意到：因为 $t_i < r_i$，所以一定有 $\forall \alpha \in \ker(A-\lambda_i)^{t_i}, (A-\lambda_i)^{r_i}(\alpha) = \mathbf{0}$，即 $\ker(A-\lambda_i I)^{t_i} \subseteq \ker(A-\lambda_i I)^{r_i}$

同时，由于 $(x-\lambda_i)^{t_i}$ 是 $A|_{\ker(A-\lambda_i I)^{r_i}}$ 的最小多项式，那么 $\forall \alpha \in \ker(A-\lambda_i I)^{r_i}, (A-\lambda_i I)^{t_i}|_{\ker(A-\lambda_i I)^{r_i}}(\alpha) = 0(\alpha) = \mathbf{0}$

那么，一定有 $\ker(A-\lambda_i I)^{r_i} = \ker(A-\lambda_i I)^{t_i}$，于是命题的前半得证

对于 $(A-\lambda_i I)|_{\ker(A-\lambda_i)^{t_i}}$ 是幂零映射的证明，只需注意到：

因为 $A|_{\ker(A-\lambda_i I)^{t_i}}$ 的最小多项式为 $(x-\lambda_i)^{t_i}$，所以 $(A-\lambda_i I)^{t_i}|_{\ker(A-\lambda_i I)^{t_i}} = 0$，且 $\forall k < t_i, (A-\lambda_i I)^k|_{\ker(A-\lambda_i I)^{t_i}} \neq 0$

因此 x^l 是 $(A-\lambda_i I)|_{\ker(A-\lambda_i)^{t_i}}$ 次数最小的首一零化多项式，即最小多项式

于是命题得证 □

在这个结论的证明过程中，我们还指出了一个特殊的结果：

推论 3.4.3

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V,V)$，$\lambda_1, \cdots, \lambda_s$ 是 A 的全体互不相同的特征值

$m(x) = (x-\lambda_1)^{t_1} \cdots (x-\lambda_s)^{t_s}$ 是 A 的最小多项式

那么，$V = \ker(A-\lambda_i I)^{t_1} \oplus \cdots \oplus \ker(A-\lambda_i I)^{t_s}$

前面的命题指出了两个事实：首先，线性变换可以分解为一些幂零变换；其次，特征多项式其实是不必要的，根子空间中的 r_i 其实也可以用 t_i 代替。

3.5 线性变换的 Jordan 标准型

3.5.1 Jordan 矩阵与 Jordan 标准型

我们首先给出 Jordan 矩阵的定义

定义 3.5.1: Jordan 块与 Jordan 矩阵

设 F 是一个域，$\lambda \in F, t \in \mathbb{N}^+$

那么我们称矩阵

$$\boldsymbol{J}_t(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_t(F) \tag{3.10}$$

为一个主对角元为 a 的 t 阶的 Jordan 块

特别地，我们称形如 $diag\{\boldsymbol{J}_{t_1}(\lambda_1), \cdots, \boldsymbol{J}_{t_s}(\lambda_s)\}$ 的分块对角矩阵为 Jordan 矩阵

接下来我们给出 Jordan 标准型的定义

定义 3.5.2: 线性变换的 Jordan 标准型

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$

如果存在 V 的一个基，使得 A 在该基下的矩阵是一个 Jordan 矩阵，那么我们称该矩阵为 A 的 Jordan 标准型

我们之前已经指出，只要特征多项式的根都在域 F 中，那么线性变换就可以分解为一些幂零变换。因此，本节中我们暂时先讨论幂零变换的 Jordan 标准型

为了给出幂零变换的 Jordan 标准型，我们先给出循环子空间的定义：

定义 3.5.3: 循环子空间

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$

如果 $\exists \alpha \in V$，使得 $\alpha, A(\alpha), \cdots, A^{t-1}(\alpha)$ 线性无关

且 $\forall r \in \mathbb{N}_+, A^r(\alpha) \in \operatorname{span}(\alpha, A(\alpha), \cdots, A^{t-1}(\alpha))$

那么我们称 $\operatorname{span}\{\alpha, A(\alpha), A^2(\alpha), \cdots, A^{t-1}(\alpha)\}$ 是一个 A-循环子空间

特别地，如果还有 $A^t(\alpha) = \mathbf{0}$，我们称 $\operatorname{span}\{\alpha, A(\alpha), A^2(\alpha), \cdots, A^{t-1}(\alpha)\}$ 是一个 A-强循环子空间

容易注意到，$A-$ 强循环子空间是 A 的一个不变子空间

3.5.2 幂零变换的 Jordan 标准型

引理 3.5.1

V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$

如果 $\alpha \in V$ 满足：$A^l(\alpha) = \mathbf{0}, A^{l-1}(\alpha) \neq \mathbf{0}$

那么，对于 $B = \{A^{l-1}(\alpha), \cdots, A(\alpha), \alpha\}$，$\operatorname{span}(B)$ 是一个 A-强循环子空间，且 $A|_{\operatorname{span}(B)}$ 在基 B 下的矩阵为 $J_l(0)$

证明： 首先证明 $A^{l-1}(\alpha), \cdots, A(\alpha), \alpha$ 线性无关。

考察线性组合 $k_{l-1}A^{l-1}(\alpha) + \cdots + k_1 A(\alpha) + k_0\alpha$

对 k_i 作数学归纳法：

首先，$k_0 = 0$，因为：$A^{l-1}(k_{l-1}A^{l-1}(\alpha) + \cdots + k_1 A(\alpha) + k_0\alpha) = \mathbf{0}$

$\Rightarrow k_0 A^{l-1}(\alpha) = \mathbf{0} \Rightarrow k_0 = 0$

现在假设 $\forall t \leqslant n-1, k_t = 0$，我们来证明 $k_n = 0$

注意到，此时 $A^{l-n-1}(k_{l-1}A^{l-1}(\alpha) + \cdots + k_n A^n(\alpha)) = k_n A^{l-1}(\alpha) = \mathbf{0} \Rightarrow k_n = 0$

由强归纳原理，一定有 $\forall i, k_i = 0$，于是线性无关性得证。

$\forall r \in \mathbb{N}_+, A^r(\alpha) \in \operatorname{span}(B)$ 是显然的，因为当 $r \geqslant t, A^r(\alpha) = \mathbf{0}$

于是我们证明了 $\operatorname{span}(B)$ 是一个 A-强循环子空间。

接下来我们只需注意到，$A|_{\operatorname{span}(B)}(A^{l-i}(\alpha)) = A^{l-i+1}(\alpha) = \sum\limits_{k=1}^{l} \delta_{i-1,k} A^{l-k}(\alpha)$

因此，$A|_{\operatorname{span}(B)}$ 在 B 下的矩阵 $\boldsymbol{A}$ 一定满足：$\boldsymbol{A}(i-1; i) = 1$，而其他元素均为 0

那么，这其实就是 $\boldsymbol{A} = J_l(0)$

于是命题得证。 □

在证明过程中，我们其实也证明了：

推论 3.5.2

V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$ 是一个幂零指数为 l 的幂零变换

那么一定有 $l \leqslant n$

证明: 这是显然的，因为我们已经证明了：对于 $\alpha \in V, A^l(\alpha) = \mathbf{0}, A^{l-1}(\alpha) \neq \mathbf{0}$

此时 $A^{l-1}(\alpha), \cdots, A(\alpha), \alpha$ 线性无关。

那么，其实一定有 $l \leqslant n$，因为 n 维线性空间中不可能有超过 n 个线性无关的向量。

于是命题得证。 □

引理 3.5.3

V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$ 是一个幂零指数为 l 的幂零变换

那么，V 可分解为 $\dim V_0 = \dim\ker A$ 个强循环子空间的直和。

证明: 对 n 作数学归纳法：

当 $n = 1$，由于我们已经指出幂零指数无法超出维数，因此一定有 $l = 1$（$A^0 = I \neq 0$，因此 l 不是零）

此时，$A = 0$，那么，其实只需要取 V 的一个基 $\{\alpha\}$，此时显然 $V = \mathrm{span}(\alpha)$ 是一个 A-强循环子空间，因此命题显然成立

现在假设 $r < n$ 时命题成立

首先，我们已经证明了 $A = 0$ 时命题成立，因此我们只需考虑 $A \neq 0$ 的情况

此时，$1 \leqslant \dim(V/\ker A) = \dim V - \dim\ker A < \dim V = n$

于是，只需找到 $V/\ker A$ 上的一个幂零变换，就可以在 $V/\ker A$ 上应用假设。

考虑映射：$\phi : V/\ker A \ni \alpha + \ker A \mapsto A(\alpha) + \ker A$

我们在证明秩-零化度定理时已经验证了这个映射的良定义性，此处不再讨论

注意到：$\forall \alpha \in V/\ker A, \phi^l(\alpha + \ker A) = \phi^l(\alpha) + \ker A = \ker A$。

因此 ϕ 是一个幂零指数小于等于 l 的幂零变换

于是，依照假设，$V/\ker A$ 一定存在以下分解：

$V/\ker A=\operatorname{span}(\alpha_1+\ker A,\cdots,\phi^{t_1-1}(\alpha_1+\ker A))\oplus\cdots\oplus\operatorname{span}(\alpha_k+\ker A,\cdots,\phi^{t_k-1}(\alpha_k+\ker A))$

且其中 $\operatorname{span}(\alpha_i+\ker A,\cdots,\phi^{t_i-1}(\alpha_i+\ker A))$ 是一个 ϕ-强循环子空间，即有 $\phi^{t_i}(\alpha_i+\ker A)=\ker A$

那么，依照直和的等价条件，$\alpha_1+\ker A,\cdots,\phi^{t_1-1}(\alpha_1+\ker A),\cdots,\alpha_k+\ker A,\cdots,\phi^{t_k-1}(\alpha_k+\ker A)$ 是 $V/\ker A$ 的一个基

我们记 $U=\operatorname{span}(\alpha_1,\cdots,A^{t_1-1}(\alpha_1),\cdots,\alpha_k,\cdots,A^{t_k-1}(\alpha_k))$

那么此时有 $V=U\oplus\ker A$

此时注意之前强循环子空间的关键条件：$\phi^{t_i}(\alpha_i+\ker A)=\ker A\Rightarrow A^{t_i}(\alpha_i)\in\ker A$

考察线性组合 $p_1A^{t_1}(\alpha_1)+\cdots+p_kA^{t_k}(\alpha_k)=\mathbf{0}$

$\Rightarrow A(p_1A^{t_1-1}(\alpha_1)+\cdots+p_kA^{t_k-1}(\alpha_k))=\mathbf{0}$

$\Rightarrow p_1A^{t_1-1}(\alpha_1)+\cdots+p_kA^{t_k-1}(\alpha_k)\in\ker A$

$\Rightarrow p_1(A^{t_1-1}(\alpha_1)+\ker A)+\cdots+p_k(A^{t_k-1}(\alpha_k)+\ker A)\in\ker A$

$\Rightarrow p_1,\cdots,p_k=0$（因为之前证明了 $\alpha_1+\ker A,\cdots,\phi^{t_1-1}(\alpha_1+\ker A),\cdots,\alpha_k+\ker A,\cdots,\phi^{t_k-1}(\alpha_k+\ker A)$ 是 $V/\ker A$ 的一个基，即线性无关）

于是我们证明了 $A^{t_1}(\alpha_1),\cdots,A^{t_k}(\alpha_k)\in\ker A$ 线性无关

于是它一定能补全成 $\ker A$ 的一个基 $\{A^{t_1}(\alpha_1),\cdots,A^{t_k}(\alpha_k),\beta_1,\cdots,\beta_s\}$

那么此时：$V=U\oplus\ker A$

$=\operatorname{span}(\alpha_1,\cdots,A^{t_1-1}(\alpha_1),\cdots,\alpha_k,\cdots,A^{t_k-1}(\alpha_k),\beta_1,\cdots,\beta_s)\oplus\operatorname{span}(A^{t_1}(\alpha_1),\cdots,A^{t_k}(\alpha_k),\beta_1,\cdots,\beta_s)$

$=\operatorname{span}(\alpha_1,\cdots,A^{t_1-1}(\alpha_1),A^{t_1}(\alpha_1))\oplus\cdots\oplus\operatorname{span}(\alpha_k,\cdots,A^{t_k-1}(\alpha_k),A^{t_k}(\alpha_k))\oplus\operatorname{span}(\beta_1)\oplus\cdots\oplus\operatorname{span}(\beta_s)$

我们来验证这些子空间的确是 A-强循环子空间。其实，依照之前的命题，我们只需要验证最后一个元素是否是 A 的核即可，但是，这些元素正是 $\ker A$ 的基，这是必定成立的。

最后，我们注意到：$\dim\ker A=\operatorname{Card}\{A^{t_1}(\alpha_1),\cdots,A^{t_k}(\alpha_k),\beta_1,\cdots,\beta_s\}=t+s$，而分解的数量也恰好是 $t+s$，因此分解的数量正好是 $\dim\ker A$ □

现在，我们可以给出结论：幂零变换的 Jordan 标准型一定是存在的。

定理 3.5.4: 幂零变换的 Jordan 标准型

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$ 是一个幂零指数为 l 的幂零变换

那么一定存在一个基，使得它在此基下的矩阵是一个 Jordan 矩阵 $diag\{J_{t_1}(0), \cdots, J_{t_k}(0)\}$，且其中 $t_i \leqslant l$

特别地，Jordan 块的个数是 $\dim\ker A$，其中 t 级 Jordan 块的个数 $N(t) = \operatorname{rank}(B^{t-1}) - 2\operatorname{rank}(B^t) + \operatorname{rank}(B^{t+1})$

证明: 首先，Jordan 标准型的存在性是显然的，因为我们只需找到它的一个强循环子空间分解 $V = W_1 \oplus \cdots \oplus W_s$，而按照我们之前证明的引理，

我们首先证明：$\operatorname{rank}(J_l(0)^t) = l - t, 0 \leqslant t \leqslant l$

我们来证明，$J_l(0)^t$ 是满足 $J_l(0)(i; i+t) = 1$，而其他元素均为 0 的矩阵。

对 t 作数学归纳法。首先，当 $t = 1$，这就是 Jordan 块的定义，自然成立

现在假设 t 时成立，我们来证明 $t+1$ 时也成立：

$J_l(0)^{t+1}(i; j) = \sum\limits_{k=1}^{l} J_l(0)^t(i; k) J_l(0)(k; j)$

显然，只有当 $k = i + t = j - 1$ 时，$J_l(0)^t(i; k) J_l(0)(k; j) = 1$，其他情况均为 0，即当且仅有 $J_l(0)^{t+1}(i; i+t) = 1$

于是由归纳法得证。那么，显然有：$\operatorname{rank}(J_l(0)^t) = l - t, 0 \leqslant t \leqslant l$

接下来我们只需注意到：$\operatorname{rank}(A^t) = \sum\limits_{k=t+1}^{l} N(k)(k-t)$（因为 $diag\{\boldsymbol{J}_{t_1}(\lambda_1), \cdots, \boldsymbol{J}_{t_s}(\lambda_s)\}^t = diag\{\boldsymbol{J}_{t_1}(\lambda_1)^t, \cdots, \boldsymbol{J}_{t_s}(\lambda_s)^t\}$）

于是有：$\operatorname{rank}(A^{t-1}) = \sum\limits_{k=t}^{l} N(k)(k - t + 1)$

$\operatorname{rank}(A^{t+1}) = \sum\limits_{k=t+2}^{l} N(k)(k - t - 1)$

于是可以注意到：$\operatorname{rank}(A^{t-1}) + \operatorname{rank}(A^{t+1}) - 2\operatorname{rank}(A^t)$

$= \sum\limits_{k=t+2}^{l} 2N(k)(k-t) + N(t) + 2N(t+1) - 2\sum\limits_{k=t+2}^{l} N(k)(k-t) - 2N(t+1) = N(t)$

于是命题得证。 □

3.5.3 线性变换的 Jordan 标准型

最后，我们应用我们给出的全部结论，给出我们的最终结果：线性变换的 Jordan 标准型

我们回顾之前给出的结果，我们指出了可以分解为若干个不变子空间，每个子空间上是一个幂零变换和数量变换的和；我们又指出了幂零变换可以写出 Jordan 标准型

结果已经很显然了：如果能给出根子空间分解，我们就可以写出 Jordan 标准型

命题 3.5.1. 设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$

那么，A 的 *Jordan* 标准型存在 $\Leftrightarrow$ A 的最小多项式在 F 中存在分解 $m(x) = (x - \lambda_1)^{t_1} \cdots (x - \lambda_s)^{t_s}$

证明: 先证明必要性。

由之前的命题,如果此形式的分解存在,那么一定有 $V = \ker(A-\lambda_1 I)^{t_1} \oplus \cdots \oplus \ker(A-\lambda_s I)^{t_s}$

但是,我们已经证明了,幂零变换 $(A-\lambda_i I)|_{\ker(A-\lambda_i I)^{t_i}}$ 的 Jordan 标准型 $diag\{J_{t_{i1}}(0), \cdots, J_{t_{ik_i}}(0)\}$

那么 $A|_{\ker(A-\lambda_i I)^{t_i}}$ 的 Jordan 标准型就是 $diag\{J_{t_{i1}}(\lambda_i), \cdots, J_{t_{ik_i}}(\lambda_i)\}$

于是 A 一定有标准型 $diag\{J_{t_{11}}(\lambda_1), \cdots, J_{t_{1k_1}}(\lambda_1), \cdots, J_{t_{s1}}(\lambda_s), \cdots, J_{t_{sk_s}}(\lambda_s)\}$，于是必要性得证。

我们接下来证明充分性

充分性是显然的，因为对于 Jordan 标准型 $diag\{J_{l_1}(\lambda_1), \cdots, J_{l_s}(\lambda_s)\}$，$J_{l_i}(\lambda_i)$ 所对应的线性变换的最小多项式是 $(x - \lambda_i)^{l_i}$

因此线性变换的最小多项式一定可以在 F 中分解为互若干一次因式的乘积，于是命题得证。 □

有了以上命题，接下来我们阐述 Jordan 标准型的具体形式

定理 3.5.5: 线性变换的 Jordan 标准型

设 V 是 F 上的一个线性空间，$\dim V = n < \aleph_0$，$A \in \hom(V, V)$，

$m(x) = (x - \lambda_1)^{t_1} \cdots (x - \lambda_s)^{t_s}$ 是 A 的最小多项式

那么，A 的 Jordan 标准型一定存在，且对角线元素为 λ_i 的 Jordan 块个数 N_i，及 $J_t(\lambda_i)$ 的个数 $N_i(T)$ 分别为：

$$N_i = \text{nullity}(A - \lambda_i I) \tag{3.11}$$

$$N_i(t) = \text{rank}((A - \lambda_i I)^{t-1}) - 2\,\text{rank}((A - \lambda_i I)^t) + \text{rank}((A - \lambda_i I)^{t+1}), t \leqslant t_i \tag{3.12}$$

证明: 首先，我们已经证明了此时 Jordan 标准型存在。我们来计算块的个数

首先，我们已经指出，$V = \ker(A - \lambda_1 I)^{t_1} \oplus \cdots \oplus \ker(A - \lambda_s I)^{t_s}$

此时，对于每一个子空间，一定能找到一个基 B_i，使 $(A - \lambda_i I)|_{\ker(A-\lambda_i I)^{t_i}}$ 在 B_i 下的矩阵为 $diag\{J_{t_{i1}}(0), \cdots, J_{t_{ik_i}}(0)\}$

于是，此时 $A|_{\ker(A-\lambda_i I)^{t_i}}$ 在 B 下的矩阵为 $diag\{J_{t_{i1}}(\lambda_i),\cdots,J_{t_{ik_i}}(\lambda_i)\}$。

由于 λ_i 是互不相同的，其实此时一定有 $N_i=t_i=\dim\ker((A-\lambda_i I)|_{\ker(A-\lambda_i I)^{t_i}})$

其实，此时 $\ker(A-\lambda_i I)^{t_i}$ 的限制是完全没有必要的，因为注意到：$\forall\alpha\in\ker(A-\lambda_i I)\Rightarrow\alpha\in\ker(A-\lambda_i I)^{t_i}$

因此，此时有 $N_i=\text{nullity}(A-\lambda_i I)$

而 $N_i(t)=\text{rank}((A-\lambda_i I)^{t-1}|_{\ker(A-\lambda_i I)^{t_i}})-2\,\text{rank}((A-\lambda_i I)^{t}|_{\ker(A-\lambda_i I)^{t_i}})+\text{rank}((A-\lambda_i I)^{t+1}|_{\ker(A-\lambda_i I)^{t_i}})$

$=(\dim\ker(A-\lambda_i I)^{t_i}-\dim\ker((A-\lambda_i I)^{t-1}|_{\ker(A-\lambda_i I)^{t_i+1}})-2(\dim\ker(A-\lambda_i I)^{t_i}-\dim\ker((A-\lambda_i I)^{t}|_{\ker(A-\lambda_i I)^{t_i}}))+(\dim\ker(A-\lambda_i I)^{t_i}-\dim\ker((A-\lambda_i I)^{t+1}|_{\ker(A-\lambda_i I)^{t_i}}))$

$=-\dim\ker((A-\lambda_i I)^{t-1}|_{\ker(A-\lambda_i I)^{t_i+1}})+2\dim\ker((A-\lambda_i I)^{t}|_{\ker(A-\lambda_i I)^{t_i}}))-\dim\ker((A-\lambda_i I)^{t+1}|_{\ker(A-\lambda_i I)^{t_i}})$

事实上，有：$\forall t,\ker((A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}})=\mathbf{0},i\neq j$

这是因为：如果 $i\neq j$，那么一定有 $((x-\lambda_i)^t,(x-\lambda_j)^{t_j})=1$

$\Rightarrow\exists u(x),v(x)\in F[x],u(x)(x-\lambda_i)^t+v(x)(x-\lambda_j)^{t_j}=1$

代入 $A|_{\ker(A-\lambda_j I)^{t_j}}$：

$\Rightarrow u(A|_{\ker(A-\lambda_j I)^{t_j}})(A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}}+v(A|_{\ker(A-\lambda_j I)^{t_j}})(A-\lambda_j I)^{t_j}|_{\ker(A-\lambda_j I)^{t_j}}=1$

$\Rightarrow u(A|_{\ker(A-\lambda_j I)^{t_j}})(A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}}=1$

由 $F[x]$ 可交换可知，此时 $((A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}})^{-1}=u(A)|_{\ker(A-\lambda_j I)^{t_j}}$

既然 $(A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}}$ 可逆，那么一定有 $\ker((A-\lambda_i I)^t|_{\ker(A-\lambda_j I)^{t_j}})=\mathbf{0}$

那么，其实之前 $N_i(t)$ 的计算公式中的子空间限制都是没有必要的（因为在其他子空间上的核都是零），因此有：

$N_i(t)=-\dim\ker((A-\lambda_i I)^{t-1}|_{\ker(A-\lambda_i I)^{t_i+1}})+2\dim\ker((A-\lambda_i I)^{t}|_{\ker(A-\lambda_i I)^{t_i}}))-\dim\ker((A-\lambda_i I)^{t+1}|_{\ker(A-\lambda_i I)^{t_i}})$

$-\dim((A-\lambda_i I)^{t-1})+2\dim\ker((A-\lambda_i I)^{t})-\dim\ker((A-\lambda_i I)^{t+1})$

$=(\dim V-\dim\ker((A-\lambda_i I)^{t-1}))-2(\dim V-\dim\ker((A-\lambda_i I)^{t}))+(\dim V-\dim\ker((A-\lambda_i I)^{t+1}))$

$=\text{rank}((A-\lambda_i I)^{t-1})-2\,\text{rank}((A-\lambda_i I)^{t})+\text{rank}((A-\lambda_i I)^{t+1})$

于是命题得证。 □

第四章　多重线性代数

4.1　多重线性映射

4.2　张量积

4.3　张量代数

4.4　外代数

第五章　多重线性代数

5.1　多重线性映射

5.2　张量积

5.3　张量代数

5.4　外代数

第六章　附录

这一部分中，对于正文中因为逻辑结构无法提及的部分，进行补充。

6.1　$\hom(V,W)$ 的维数

这篇附录中，我们解决一个问题：$\hom(V,W)$ 的维数是多少。

引理 6.1.1

$\hom(V,W) \cong \bigoplus_{i=1}^{n} \hom(F,W)$

证明： 任取 V 的一个基 $\{\alpha_1,\cdots,\alpha_n\}$，我们定义：

$\phi : \hom(V,W) \ni f \mapsto (\{1,\cdots,n\} \ni i \mapsto (F \ni k \mapsto kf(\alpha_i) \in W) \in \hom(F,W)) \in \bigoplus_{i=1}^{n} \hom(F,W)$

（即满足 $\phi(f)(i)(k) = kf(\alpha_i)$ 的唯一映射）

接下来验证 ϕ 是同构映射

首先，容易注意到 $\phi(f) = \phi(g) \Rightarrow f = g$，因此 ϕ 是单射

其次，$\forall\, (\{1,\cdots,n\} \ni i \mapsto g_i \in \hom(F,W)) \in \bigoplus_{i=1}^{n} \hom(F,W)$

注意到，如果设 $f : V \ni \sum\limits_{i=1}^{n} k_i\alpha_i \mapsto \sum\limits_{i=1}^{n} k_i g_i(1_F)$，那么 $\phi(f)(i) = g_i$，于是 ϕ 是满射

接下来验证它是一个线性映射

$\forall f,g \in \hom(V,W), \phi(f+g)(i)(k) = k(f+g)(\alpha_i) = kf(\alpha_i) + kg(\alpha_i) = \phi(f)(i)(k) + \phi(g)(i)(k) \Rightarrow \phi(f+g) = \phi(f) + \phi(g)$

$\forall f \in \hom(V,W), l \in F, \phi(lf)(i)(k) = k(lf)(\alpha_i) = lkf(\alpha_i) = l\phi(f)(i)(k) \Rightarrow \phi(lf) = l\phi(f)$

因此它是一个同构映射，命题得证。 □

引理 6.1.2

$\hom(F,W)\cong W$

证明： 任取 V 的一个基 $\{\alpha_1,\cdots,\alpha_n\}$，我们定义：

$\psi:\alpha\in W\mapsto(F\ni k\mapsto k\alpha\in W)\in\hom(F,W)$

注意到：$\psi(\alpha)=\psi(\beta)\Rightarrow\psi(\alpha)(k)=\psi(\beta)(k)\Rightarrow k\alpha=k\beta\Rightarrow\alpha=\beta$，于是 ψ 是单射

$\forall f\in\hom(F,W),\psi(f(1_F))(k)=kf(1_F)=f(k)\Rightarrow\psi(f(1_F))=f$，于是 ψ 是满射

$\forall\alpha,\beta\in W,\psi(\alpha+\beta)(k)=k(\alpha+\beta)=k\alpha+k\beta=\psi(\alpha)(k)+\psi(\beta)(k)\Rightarrow\psi(\alpha+\beta)=\psi(\alpha)+\psi(\beta)$

$\forall\alpha\in W,l\in F,\psi(l\alpha)(k)=lk\alpha=l\psi(\alpha)(k)\Rightarrow\psi(l\alpha)=l\psi(\alpha)$

于是 ψ 是一个同构映射。 □

引理 6.1.3

$|\hom(V,W)|=|W|^{\dim V}$

证明： 事实上，线性映射只由基上的作用决定，因此如果设 V 的一个基为 B

$|\hom(V,W)|=\mathrm{Card}\{f:B\to W\}=|W|^{|B|}=|W|^{\dim V}$ □

引理 6.1.4

$\dim\hom_F(V,W)\geqslant|F|$

证明： 取 V 的一个基 B，任取 B 的一个可数子集 $\{\alpha_1,\cdots,\alpha_n,\cdots\}$，又任取 $\gamma\in W,\gamma\neq\mathbf{0}_W$

我们定义：$k\in F,f_k(\alpha_i)=k^{i-1}\gamma$，并且 $\forall\beta\in B-\{\alpha_1,\cdots,\alpha_n\},f_k(\beta)=\mathbf{0}_W$

显然，如此定义的映射集 $\{f_k|k\in F\}$ 中的任意映射都是线性映射（因为我们给出且只给出了在基上的像）

我们来验证，$\{f_k|k\in F\}$ 的任意有限子集线性无关

取它的一个有限子集 $\{f_{k_1},\cdots,f_{k_s}\}$

如果 $\sum\limits_{i=1}^{s}l_if_{k_i}=0$

$\Rightarrow\forall\alpha_j,\sum\limits_{i=1}^{s}l_if_{k_i}(\alpha_j)=\mathbf{0}_W$

$\Rightarrow\sum\limits_{i=1}^{s}l_ik_i^{j-1}\gamma=\mathbf{0}_W$

$\Rightarrow\sum\limits_{i=1}^{s}l_ik_i^{j-1}=0$

因此，$\forall j\in\mathbb{N}_+,\sum\limits_{i=1}^{s} l_i k_i^{j-1}=0$，我们取前 s 项，并写成方程组的形式：

$\boldsymbol{A}\mathbf{x}=\mathbf{0}_{F^s}$

其中 $\boldsymbol{A}=\begin{pmatrix}1 & \cdots & 1\\ \vdots & \ddots & \vdots\\ k_1^{s-1} & \cdots & k_s^{s-1}\end{pmatrix},\mathbf{x}=\begin{pmatrix}l_1\\ \ddots\\ l_s\end{pmatrix}$

注意到，$|\boldsymbol{A}|=\prod_{1\leqslant i<j\leqslant s}(k_j-k_i)\neq 0$

因此方程组 $\boldsymbol{A}\mathbf{x}=\mathbf{0}_{F^s}$ 仅有平凡解 $\mathbf{x}=\mathbf{0}_{F^s}$

又因为这个有限子集是随意选取的，因此 $\{f_k|k\in F\}$ 线性无关。

那么，就必须有 $\dim\hom(V,W)\geqslant|F|$，命题得证 □

引理 6.1.5

设 U 是 F 上的一个线性空间，那么 $|U|=|F|\cdot\dim_F U$

证明：

□

引理 6.1.6

$\hom(V,W)\cong\{f:B_1\to G|G\subseteq F^{B_2},\forall g\in G,supp(g)<\aleph_0\}$

其中 B_1,B_2 分别是是 V,W 的基

证明： 记 $\mathcal{P}=\{f:B_1\to G|G\subseteq F^{B_2},\forall g\in G,supp(g)<\aleph_0\}$

我们定义 $\phi:\hom(V,W)\ni T\mapsto\big(B_1\ni\alpha\mapsto\big(B_2\ni\beta\mapsto T_{\alpha\beta}\in F\big)\in G\big)\in\mathcal{P}$

其中 $T(\alpha)=\sum\limits_{\beta\in B_2}T_{\alpha\beta}\beta$

这个定义是合理的，因为我们知道，任意一个向量可由有限个基向量线性组合，因此 $B_2\ni\beta\mapsto T_{\alpha\beta}\in F$ 是有限支撑的（即使 B_2 无限）

接下来验证 ϕ 是同构映射

先验证它是线性的：

注意到：$\forall A,B\in\hom(V,W),\phi(A+B)(\alpha)(\beta)=(A+B)_{\alpha\beta}=A_{\alpha\beta}+B_{\alpha\beta}=\phi(A)(\alpha)(\beta)+\phi(B)(\alpha)(\beta)\Rightarrow\phi(A+B)=\phi(A)+\phi(B)$

$\forall k\in F,A\in\hom(V,W),\phi(kA)(\alpha)(\beta)=(kA)_{\alpha\beta}=k\cdot A_{\alpha\beta}=k\phi(A)(\alpha)(\beta)\Rightarrow\phi(kA)=k\phi(A)$

接下来验证它是一个双射：

注意到：$\forall g\in\mathscr{P}$，令 $T(\alpha)=\sum\limits_{\beta\in B_2}g(\alpha)(\beta)\beta,\alpha\in B_1$。这个定义是良好的，因为它仅仅给出了在基上的作用，而且由于 $g(\alpha)$ 是有限支撑的，所以求和是良好的。

$\phi(T)(\alpha)(\beta)=g(\alpha)(\beta)\Rightarrow\phi(T)=g$

又注意到: $\phi(A)=\phi(B)\Rightarrow\forall\alpha\in B_1,\beta\in B_2,\phi(A)(\alpha)(\beta)=\phi(B)(\alpha)(\beta)\Rightarrow A_{\alpha\beta}=B_{\alpha\beta}\Rightarrow A=B$

于是 ϕ 的确是一个同构映射，命题得证。 □

定理 6.1.7: $\hom_F(V_1,V_2)$ 的维数

$$\dim\hom_F(V,W)=\begin{cases}0,\dim V=0\vee\dim W=0\\(\dim V)\cdot(\dim W),\dim V=n<\aleph_0\\|\hom(V,W)|=|W|^{\dim V}=(|F|\cdot\dim W)^{\dim V},\dim V\geqslant\aleph_0\end{cases}$$

证明: 首先考虑 V,W 中存在零空间的情形。

如果 $\dim V=0$，即 $V=\{\mathbf{0}_V\}$，而我们知道线性映射仅能把零向量映射到零向量，此时 $\hom(V,W)$ 中仅存在零映射，维数为零

如果 $\dim W=0$。即 $W=\{\mathbf{0}_W\}$，此时映射只有零映射（因为 W 中仅存在零向量），维数为零

接下来考虑 V,W 中不存在零空间，且 $\dim V=n<\aleph_0$ 的情形。

此时，$\dim\hom(V,W)=\dim\bigoplus_{i=1}^n\hom(F,W)=n\cdot\dim\hom(F,W)=n\cdot\dim W=(\dim V)\cdot(\dim W)$

最后我们证明 $\dim V,\dim W$ 均是无限的情形。

□