中国科学院大学课程讲义

线性代数 II

中国科学院大学　　数学科学学院
作者：支丽红
编者：禹天石
鸣谢：梁昊 郑涛 代梓灏
2024 年 12 月 30 日

目录

第一章　向量空间与二次型

在上册我们已经学习了实数域上的 n 维向量空间 $\mathbb{R}^n$，那么一般域上的向量空间应该有什么样的结构和性质呢？这就是我们接下来所要讨论的问题。

本章和接下来的两章将始终沿着下面的思路进行讨论：我们为抽象的空间选取一组合适的基，在这组基下，向量可以用坐标表示，二次型或算子可以用矩阵表示，那么就产生了以下两个问题：

- 基变换与坐标变换：不同的基下同一个向量 (二次型或算子) 的坐标 (矩阵) 之间有什么关系？
- 矩阵的简化与空间分解：我们希望使二次型 (或算子) 在某一组基下的矩阵尽可能简单 (如对角形或分块对角形)，那么，应该如何选取合适的基呢？基的选取也对应着将空间进行合适的分解。

我们要在抽象的空间、二次型、算子和具体的坐标向量、矩阵之间建立深刻的联系，这也是线性代数这门课程最为核心的思想。

1.1　抽象向量空间

回忆讲义上册 2.1.1 小节中 $\mathbb{R}^n$ 的性质，我们将 $\mathbb{R}^n$ 所满足的性质抽象出来，并推广到一般的域 $\mathbb{K}$ 上，就有如下的定义：

定义 1.1.1. 设 $\mathbb{K}$ 是一个一般的域，我们有一个与 $\mathbb{K}$ 相关联的集合 V 及 V 上的两个运算：

$$\begin{aligned} &\text{加法：}V\times V\to V &&\text{数乘：}\mathbb{K}\times V\to V\\ &(\mathbf{x},\mathbf{y})\mapsto \mathbf{x}+\mathbf{y} &&(\lambda,\mathbf{x})\mapsto \lambda\cdot\mathbf{x}(\text{简记为}\lambda\mathbf{x}) \end{aligned}$$

V 满足以下条件：

(1) V 关于加法成交换群，即加法满足封闭性、交换律、结合律，V 中存在加法单位元 $\mathbf{0}\in V(\forall\mathbf{x}\in V,\ \mathbf{x}+\mathbf{0}=\mathbf{x})$，并且 $\forall\mathbf{x}\in V$，存在加法负元 $\mathbf{y}$ 满足 $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}=\mathbf{0}$(记为 $\mathbf{y}=-\mathbf{x}$)。

(2) V 关于数乘运算满足结合律和酉性：即任意 $\alpha,\beta\in\mathbb{K}$ 及 $\mathbf{x}\in V$，我们有 $(\alpha\beta)\mathbf{x}=\alpha(\beta\mathbf{x})$，并且域 $\mathbb{K}$ 中的乘法单位元 $1_{\mathbb{K}}$ 满足 $\forall\mathbf{x}\in V,\ 1_{\mathbb{K}}\cdot\mathbf{x}=\mathbf{x}$。

(3) V 上的加法与数乘满足分配律：$\forall\alpha,\beta\in\mathbb{K}$ 及 $\mathbf{x},\mathbf{y}\in V$，我们有 $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$, $\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$。

则我们称 V 是域 $\mathbb{K}$ 上的抽象**向量空间** (abstract vector space) 或者线性空间 (linear space)。

我们先来看向量空间的一些简单性质。

命题 1.1.1. 设 $(\mathbb{K},+,0,\cdot,1)$ 是域，V 是 $\mathbb{K}$ 上的向量空间，则

(1) $\mathbf{0}\in V$ 是唯一的，并且 $\forall\mathbf{x}\in V$，有 $0\cdot\mathbf{x}=\mathbf{0}$；

(2) $\forall\mathbf{x}\in V,\lambda\in\mathbb{K}$，我们有 $\lambda\mathbf{x}=\mathbf{0}\iff\lambda=0$ 或 $\mathbf{x}=\mathbf{0}$；

(3) $\forall \mathbf{x} \in V$，$-\mathbf{x}$ 是唯一的，并且 $-\mathbf{x} = (-1) \cdot \mathbf{x}$；

(4) 我们记 $\underbrace{\mathbf{x} + \cdots + \mathbf{x}}_{n\text{个}}$ 为 $n\mathbf{x}$，则 $n\mathbf{x} = (n \cdot 1)\mathbf{x}$($n \cdot 1$ 表示 n 个 1 相加)。特别地，如果 $\mathrm{char}(\mathbb{K}) = p$，则 $p\mathbf{x} = \mathbf{0}$。

证明. (1) $\mathbf{0}$ 的唯一性可由群 $(V, +, \mathbf{0})$ 中单位元的唯一性得到。又因为 $0 \cdot \mathbf{x} = (0 + 0) \cdot \mathbf{x} = 0 \cdot \mathbf{x} + 0 \cdot \mathbf{x}$，即 $0 \cdot \mathbf{x} = \mathbf{0}$。

(2) $(\Longleftarrow) \lambda = 0$ 推出 $\lambda\mathbf{x} = \mathbf{0}$ 已经在 (1) 中证明，而由 $\lambda\mathbf{0} = \lambda(\mathbf{0} + \mathbf{0}) = \lambda\mathbf{0} + \lambda\mathbf{0}$ 即得 $\lambda\mathbf{0} = 0$；$(\Longrightarrow)$ 我们已知 $\lambda\mathbf{x} = \mathbf{0}$，如果 $\lambda = 0$，则证明结束，否则由 $\mathbb{K}$ 是域可知存在 $\lambda^{-1} \in \mathbb{K}$ 使得 $\mathbf{0} = \lambda^{-1} \cdot \mathbf{0} = \lambda^{-1} \cdot (\lambda\mathbf{x}) = (\lambda^{-1} \cdot \lambda)\mathbf{x} = 1 \cdot \mathbf{x}$，即 $\mathbf{x} = 0$。

(3) $-\mathbf{x}$ 的唯一性可由群 $(V, +, \mathbf{0})$ 中加法逆元的唯一性得到。注意到 $\mathbf{x} + (-1) \cdot \mathbf{x} = 1 \cdot \mathbf{x} + (-1) \cdot \mathbf{x} = (1 + (-1))\mathbf{x} = 0 \cdot \mathbf{x} = \mathbf{0}$，即 $(-1) \cdot \mathbf{x}$ 是 $\mathbf{x}$ 的加法逆元，再由加法逆元的唯一性即有 $-\mathbf{x} = (-1) \cdot \mathbf{x}$。

(4) 由分配律立刻可证。

□

注 1.1.1. 如果将域 $\mathbb{K}$ 换成任意的交换环 R，则称 V 是一个左 $R-$ 模。与域不同的是，在左 $R-$ 模 V 中 $\lambda\mathbf{x} = \mathbf{0} \nRightarrow \lambda = 0 \text{或} \mathbf{x} = \mathbf{0}$。

我们先来看一些向量空间的具体例子。

例 1.1.1. 1. 容易验证上册中我们详细讨论的 $\mathbb{R}^n$ 是 $\mathbb{R}$ 上的向量空间。特别地，$n = 1$ 时 $\mathbb{R}$ 是 $\mathbb{R}$ 自己上的向量空间。类似地，若 $\mathbb{F}$ 是域，则

$$\mathbb{F}^n = \{(x_1, \ldots, x_n)^t \mid x_i \in \mathbb{F},\ i = 1, \ldots, n\}$$

在通常的加法和数乘下也是 $\mathbb{F}$ 上的向量空间。

2. 设 $\mathbb{F}$ 是域，令 $V = \mathbb{F}^{n \times m} = \{\mathbb{F}\text{上所有的}n \times m\text{矩阵}\}$，则 V 在通常的矩阵加法和数乘下是 $\mathbb{F}$ 上的向量空间。

3. 设 $\mathbb{F}$ 是域，则 $\mathbb{F}[x_1, \ldots, x_n]$ 在通常的多项式加法和数乘下是 $\mathbb{F}$ 上的向量空间。

4. 设 $\mathbb{K}$ 是域，X 是任意一个非空集合，我们令

$$\mathbb{K}^X = \{\text{所有的映射} f : X \to \mathbb{K}\}$$

在 $\mathbb{K}^X$ 上逐点定义如下的加法和数乘：任取 $f, g : X \to \mathbb{K}$，$\lambda \in \mathbb{K}$，则

$$\begin{aligned} f + g : & X \to \mathbb{K} & \qquad \lambda \cdot f : & X \to \mathbb{K} \\ & x \mapsto f(x) + g(x) & & x \mapsto \lambda f(x) \end{aligned}$$

首先上面定义的加法和数乘是良定义的 (即加法和数乘得到的仍是 $\mathbb{K}^X$ 中的映射)；其次，容易验证 $\mathbb{K}^X$ 在上面定义的加法和数乘之下是 $\mathbb{K}$ 上的向量空间 (细节留作练习)。

例 1.1.2. $\mathbb{Z}$ 不是任何域 $\mathbb{K}$ 上的向量空间。

证明. 用反证法。

(1) 如果 $\operatorname{char}(\mathbb{K}) = 0$，则 $\mathbb{K}$ 中存在一个子域与 $\mathbb{Q}$ 同构。取 $\mathbb{K}$ 中的乘法幺元 $1_{\mathbb{K}}$，作 $a = \underbrace{1_{\mathbb{K}} + \cdots + 1_{\mathbb{K}}}_{m个} \in \mathbb{K}$，则由 $\mathbb{K}$ 是域可知存在 $a^{-1} \in \mathbb{K}$。由于 $\mathbb{Z}$ 是 $\mathbb{K}$ 上的向量空间，故取 $1_{\mathbb{Z}} \in \mathbb{Z}$，$1_{\mathbb{Z}}$ 可以视作向量，故可以用 $\mathbb{K}$ 中的元素数乘 $1_{\mathbb{Z}}$，得到的结果仍在 $\mathbb{Z}$ 中。那么就有

$$a \cdot 1_{\mathbb{Z}} = m_{\mathbb{Z}} \in \mathbb{Z},\ a^{-1} \cdot 1_{\mathbb{Z}} = n_{\mathbb{Z}} \in \mathbb{Z}.$$

于是，一方面我们有

$$a^{-1}\cdot(a\cdot 1_{\mathbb{Z}}) = a^{-1}\cdot m_{\mathbb{Z}} = a^{-1}\cdot(\underbrace{1_{\mathbb{Z}} + \cdots + 1_{\mathbb{Z}}}_{m个}) = \underbrace{a^{-1} \cdot 1_{\mathbb{Z}} + \cdots + a^{-1} \cdot 1_{\mathbb{Z}}}_{m个} = \underbrace{n_{\mathbb{Z}} + \cdots + n_{\mathbb{Z}}}_{m个} = (m\cdot n)_{\mathbb{Z}}$$

另一方面

$$(a^{-1}a) \cdot 1_{\mathbb{Z}} = 1_{\mathbb{K}} \cdot 1_{\mathbb{Z}} = 1_{\mathbb{Z}}\ (1_{\mathbb{K}}的酉性)$$

这与数乘的结合律矛盾!

(2) 如果 $\operatorname{char}(\mathbb{K}) = p$ 为素数，则取 $1_{\mathbb{K}} \in \mathbb{K}$，一方面

$$(\underbrace{1_{\mathbb{K}} + \cdots + 1_{\mathbb{K}}}_{p个}) \cdot 1_{\mathbb{Z}} = 0_{\mathbb{K}} \cdot 1_{\mathbb{Z}} = 0_{\mathbb{Z}}$$

另一方面

$$(\underbrace{1_{\mathbb{K}} + \cdots + 1_{\mathbb{K}}}_{p个}) \cdot 1_{\mathbb{Z}} = \underbrace{1_{\mathbb{K}} \cdot 1_{\mathbb{Z}} + \cdots + 1_{\mathbb{K}} \cdot 1_{\mathbb{Z}}}_{p个} = \underbrace{1_{\mathbb{Z}} + \cdots + 1_{\mathbb{Z}}}_{p个} = p_{\mathbb{Z}}$$

两者显然矛盾!

综上所述，$\mathbb{Z}$ 不是任何域 $\mathbb{K}$ 上的向量空间。 □

与 $\mathbb{R}^n$ 类似，我们仍然有线性组合、线性包和子空间的概念。

定义 1.1.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\lambda_1, \ldots, \lambda_n \in \mathbb{K}$，向量 $\mathbf{x}_1, \ldots \mathbf{x}_n \in V$，则我们称 $\lambda_1\mathbf{x}_1 + \cdots + \lambda_n\mathbf{x}_n = \sum_{i=1}^n \lambda_i\mathbf{x}_i$ 是 $\mathbf{x}_1, \ldots \mathbf{x}_n$ 的一个线性组合 (linear combination)。我们称 $\mathbf{x}_1, \ldots \mathbf{x}_n$ 的所有线性组合所构成的集合为 $\mathbf{x}_1, \ldots \mathbf{x}_n$ 的线性包 (linear span)，记为 span $\{\mathbf{x}_1, \ldots \mathbf{x}_n\}$。

定义 1.1.3. 设 V 是域 $\mathbb{K}$ 上的向量空间，$U \subset V$ 非空，如果 $\forall \mathbf{x}, \mathbf{y} \in U$，$\lambda, \mu \in \mathbb{K}$，有 $\lambda\mathbf{x} + \mu\mathbf{y} \in U$，则称 U 是 V 的一个子空间 (subspace)。U 中的运算与 V 中的运算相同。显然 $\mathbf{x}_1, \ldots \mathbf{x}_n$ 的线性包是子空间，称为 $\mathbf{x}_1, \ldots \mathbf{x}_n$ 生成的子空间。此外，显然我们有 $U \subset V$ 是子空间 $\iff$ U 中任意有限个向量的线性组合仍在 U 中。

显然任何向量空间至少都有两个子空间：$\{\mathbf{0}\}$ 和 V 本身，它们称为 V 的平凡子空间 (trivial subspace)，其余的子空间称为非平凡子空间。

下面我们看一些例子。

例 1.1.3. 设 V 是域 $\mathbb{K}$ 上的向量空间，$M = \{\mathbf{x}_i \in V \mid i \in I\} \subset V$，容易验证集合

$$\{\sum_{i \in I} \lambda_i\mathbf{x}_i \mid \lambda_i \in \mathbb{K}且仅有有限多个\lambda_i非零\}$$

是 V 的子空间，我们称其为 M 的线性包或者由 M 生成的子空间，记为 $\operatorname{span} M$。需要注意的是，这里的线性组合由于有仅有有限多个组合系数非零的要求，所以 $\operatorname{span} M$ 中的元素仍然是有限和的形式。

例 1.1.4. $SM_n(\mathbb{K})$(域 $\mathbb{K}$ 上的所有 n 阶对称矩阵的集合) 是 $M_n(\mathbb{K})$ 的子空间。这是因为 $\forall A, B \in SM_n(\mathbb{K})$，任取 $\lambda, \mu \in \mathbb{K}$，我们有

$$(\lambda A + \mu B)^t = \lambda A^t + \mu B^t = \lambda A + \mu B$$

即 $\lambda A + \mu B \in SM_n(\mathbb{K})$，所以 $SM_n(\mathbb{K})$ 是 $M_n(\mathbb{K})$ 的子空间。

例 1.1.5. 设 $\mathbb{K}$ 是域，$m \in \mathbb{Z}^+$，则 $\mathbb{K}[x_1, \ldots, x_n]^m = \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid \deg(f) < m\}$ 是 $\mathbb{K}[x_1, \ldots, x_n]$ 的子空间。

例 1.1.6. 容易验证 $\mathbb{K}^n$ 中的向量都可以写成 $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)^t$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)^t, \cdots\cdots, \mathbf{e}_n = (0, 0, \ldots, 0, 1)^t$ 的线性组合。

与 $\mathbb{R}^n$ 的情形类似，我们可以定义子空间的交与和，并证明它们都是子空间。

命题 1.1.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，则

(1) 设 $V_1, \ldots, V_k$ 是 V 的子空间，定义 V 的子集

$$V_1 + \cdots + V_k = \{\mathbf{x}_1 + \cdots \mathbf{x}_k \mid \forall i = 1, \ldots, k,\ \mathbf{x}_i \in V_i\}$$

则 $V_1 + \cdots + V_k$ 是 V 的子空间，称为 $V_1, \ldots, V_k$ 的和空间。

(2) 设 $\{V_i \mid i \in I\}$ 是 V 的一族子空间 (可以是无穷个甚至不可数个)，则它们的交 $\bigcap\limits_{i\in I} V_i$ 也是 V 的一个子空间。

证明. (1) 设 $\mathbf{x}, \mathbf{y} \in V_1 + \cdots + V_k$，则 $\mathbf{x} = \mathbf{x}_1 + \cdots \mathbf{x}_k$，$\mathbf{y} = \mathbf{y}_1 + \cdots \mathbf{y}_k$，其中 $\forall i = 1, \ldots, k$ 有 $\mathbf{x}_i, \mathbf{y}_i \in V_i$，于是对 $\forall \lambda, \mu \in \mathbb{K}$，我们有

$$\lambda \mathbf{x} + \mu \mathbf{y} = (\lambda \mathbf{x}_1 + \mu \mathbf{y}_1) + \cdots + (\lambda \mathbf{x}_k + \mu \mathbf{y}_k)$$

由于 $V_1, \ldots, V_k$ 都是 V 的子空间，因此 $\forall i = 1, \ldots, k$，有 $\lambda \mathbf{x}_i + \mu \mathbf{y}_i \in V_i$，因此 $\lambda \mathbf{x} + \mu \mathbf{y} \in V_1 + \cdots + V_k$。此即 $V_1 + \cdots + V_k$ 是 V 的子空间。

(2) 设 $\mathbf{x}, \mathbf{y} \in \bigcap\limits_{i\in I} V_i$，任取 $\lambda, \mu \in \mathbb{K}$，则对任意 $i \in I$，由 V_i 是子空间可得 $\lambda \mathbf{x} + \mu \mathbf{y} \in V_i$，所以 $\lambda \mathbf{x} + \mu \mathbf{y} \in \bigcap\limits_{i\in I} V_i$，即 $\bigcap\limits_{i\in I} V_i$ 是 V 的子空间。

□

1.2 子空间的直和

与 $\mathbb{R}^n$ 中子空间的直和类似，我们也可以定义一般向量空间中什么样的子空间的和是直和。

定义 1.2.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，$U_1,\ldots,U_n\subset V$ 是子空间，令 $U=U_1+\cdots+U_n$。如果对 $\forall\mathbf{x}\in U$，存在唯一一组 $\mathbf{x}_1\in U_1,\mathbf{x}_2\in U_2,\ldots,\mathbf{x}_n\in U_n$ 使得 $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_n$，则称 U 是 $U_1,U_2,\ldots,U_n$ 的（内）直和 (internal direct sum)，记作 $U=U_1\oplus U_2\oplus\cdots\oplus U_n$ （n 可以取无穷）。

例 1.2.1. 取 $V=\mathbb{R}^3$，$U_1=\{(\lambda_1,\lambda_2,0)^t\mid\lambda_1,\lambda_2\in\mathbb{R}\}$，$U_2=\{(0,\lambda_1,\lambda_2)^t\mid\lambda_1,\lambda_2\in\mathbb{R}\}$，则容易验证 $V=U_1+U_2$，但不是直和。例如，

$$(0,1,0)^t=\underbrace{(0,1,0)^t}_{\in U_1}+\underbrace{(0,0,0)^t}_{\in U_2}=\underbrace{(0,0,0)^t}_{\in U_1}+\underbrace{(0,1,0)^t}_{\in U_2}.$$

例 1.2.2. 设 $\mathbb{K}$ 是域，$\operatorname{char}(\mathbb{K})\neq 2$。令 $SM_n(\mathbb{K})=\{A\in M_n(\mathbb{K})\mid A=A^t\}$, $AM_n(\mathbb{K})=\{A\in M_n(\mathbb{K})\mid A=-A^t\}$，则 $M_n(\mathbb{K})=SM_n(\mathbb{K})\oplus AM_n(\mathbb{K})$。这是因为 $\operatorname{char}(\mathbb{K})\neq 2$，每个矩阵 $A\in M_n(\mathbb{K})$ 都可以表示成对称矩阵 $B=\dfrac{A+A^t}{2}$ 与斜对称矩阵 $C=\dfrac{A-A^t}{2}$ 的和，并且如果另有 $B'\in SM_n(\mathbb{K}),C'\in AM_n(\mathbb{K})$ 使得 $A=B'+C'$，则有 $A^t=(B')^t+(C')^t=B'-C'$，因此 $B'=B$, $C'=C$，这说明这个表达式是唯一的。

命题 1.2.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，$U_1,\ldots U_n$ 是 V 的子空间，令 $U=U_1+\cdots+U_n$，则以下条件等价：

1) $U=U_1\oplus\cdots\oplus U_n$;

2) 若 $\mathbf{x}_1+\cdots+\mathbf{x}_n=\mathbf{0}$，其中 $\forall i\in\{1,\ldots,n\}$ 有 $\mathbf{x}_i\in U_i$，则 $\mathbf{x}_1=\cdots=\mathbf{x}_n=\mathbf{0}$;

3) $\forall i\in\{1,\ldots,n\}$，有 $U_i\cap(U_1+\cdots+U_{i-1}+U_{i+1}+\cdots+U_n)=\{\mathbf{0}\}$ [1]。

证明. 1)⇒2): 由 $\mathbf{0}=\mathbf{x}_1+\cdots+\mathbf{x}_n=\mathbf{0}+\cdots+\mathbf{0}$，按直和的定义即有 $\mathbf{x}_1=\cdots=\mathbf{x}_n=\mathbf{0}$。
2)⇒3): 不妨设 $i=1$(其余同理)，设 $\mathbf{x}\in U_1\cap(U_2+\cdots+U_n)$，则存在 $\mathbf{x_2}\in U_2,\ldots,\mathbf{x}_n\in U_n$ 使得 $\mathbf{x}=\mathbf{x}_2+\cdots+\mathbf{x}_n$，即 $\mathbf{0}=\underbrace{(-\mathbf{x})}_{\in U_1}+\underbrace{\mathbf{x}_2}_{\in U_2}+\cdots+\underbrace{\mathbf{x}_n}_{\in U_n}$，由 2) 的条件即有 $\mathbf{x}=\mathbf{x}_2=\cdots=\mathbf{x}_n=\mathbf{0}$。
3)⇒1): 设 $\mathbf{x}=\mathbf{x}_1+\cdots+\mathbf{x}_n=\mathbf{x}'_1+\cdots+\mathbf{x}'_n$，其中 $\forall i\in\{1,\ldots,n\}$, $\mathbf{x}_i,\mathbf{x}'_i\in U_i$，我们需要证明 $\mathbf{x}_i=\mathbf{x}'_i$，从而说明唯一性。注意到

$$(\mathbf{x}_1-\mathbf{x}'_1)+(\mathbf{x}_2-\mathbf{x}'_2)+\cdots+(\mathbf{x}_n-\mathbf{x}'_n)=\mathbf{0},$$

即

$$\begin{array}{ccccc}(\mathbf{x}_1-\mathbf{x}'_1) & = & -(\mathbf{x}_2-\mathbf{x}'_2) & -\cdots & -(\mathbf{x}_n-\mathbf{x}'_n)\\ \in U_1 & & \in U_2 & & \in U_n\end{array}$$

也即 $(\mathbf{x}_1-\mathbf{x}'_1)\in U_1\cap(U_2+\cdots+U_n)$，故由 3) 的条件可知 $\mathbf{x}_1-\mathbf{x}'_1=\mathbf{0}$，即 $\mathbf{x}_1=\mathbf{x}'_1$。同理可以证明其余的 $\mathbf{x}_i-\mathbf{x}'_i$。这样我们就完成了证明。 □

注 1.2.1. 设 $U=U_1\oplus\cdots\oplus U_n$ 是直和，$i_1,\ldots,i_s\in\{1,\ldots,n\}$ 两两不同，则 $U_{i_1}+U_{i_2}+\cdots+U_{i_s}$ 也是直和。

---

[1] $U_1+\cdots+U_{i-1}+U_{i+1}+\cdots+U_n$ 也可以简记作 $U_1+\cdots+\widehat{U_i}+\cdots+U_n$。

例 1.2.3. 沿用例1.2.2的记号。显然当 $\mathrm{char}(\mathbb{K}) \neq 2$ 时，若 $A \in SM_n(\mathbb{K}) \cap AM_n(\mathbb{K})$，则 $A = A^t = -A^t$，于是 $A^t = O$，$A = O$。所以由命题1.2.1即有 $SM_n(\mathbb{K}) + AM_n(\mathbb{K})$ 是直和。注意当 $\mathrm{char}(\mathbb{K}) = 2$ 时，$E_n \in SM_n(\mathbb{K}) \cap AM_n(\mathbb{K})$ 且 $E_n \neq O$，故此时 $SM_n(\mathbb{K}) + AM_n(\mathbb{K})$ 不是直和。

利用内直和我们也可以定义一组向量空间的 (外) 直和如下。

定义 1.2.2. 设 $V_1, \ldots, V_n$ 是域 $\mathbb{K}$ 上的向量空间，则我们可以定义向量空间 V:

$$
\begin{aligned}
&V = \{(\mathbf{v}_1, \ldots, \mathbf{v}_n) \mid \mathbf{v}_i \in V_i,\ i \in 1, \ldots, n\} \\
&+ :\ (\mathbf{v}_1, \ldots, \mathbf{v}_n) + (\mathbf{v}'_1, \ldots, \mathbf{v}'_n) = (\mathbf{v}_1 + \mathbf{v}'_1, \ldots, \mathbf{v}_n + \mathbf{v}'_n) \\
&\ \cdot :\ \lambda \cdot (\mathbf{v}_1, \ldots, \mathbf{v}_n) = (\lambda \mathbf{v}_1, \ldots, \lambda \mathbf{v}_n)
\end{aligned}
$$

显然每个 V_i 可以视作 V 的一个子空间 (令其它分量为零向量即可)，并且利用命题1.2.1很容易看出 $V = V_1 \oplus \cdots \oplus V_n$。我们把 V 称为 $V_1, \ldots, V_n$ 的 (外) 直和 (external direct sum)。由于外直和与内直和的性质是一样的，因此今后我们不再区分外直和与内直和，而是统称直和。

1.3 线性相关性，维数与基底

这一节的内容仍是 $\mathbb{R}^n$ 中相对应内容的简单推广。

定义 1.3.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{v}_1, \cdots, \mathbf{v}_n \in V$，如果存在不全为 0 的 $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ 使得

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}$$

成立，则称 $\mathbf{v}_1, \cdots, \mathbf{v}_n$ 在域 $\mathbb{K}$ 上线性相关 (linearly dependent)；反之，若

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0} \Longrightarrow \lambda_1 = \cdots = \lambda_n = 0,$$

则称 $\mathbf{v}_1, \cdots, \mathbf{v}_n$ 在域 $\mathbb{K}$ 上线性无关 (linearly independent)。

例 1.3.1. 将 $\mathbb{C}$ 视作 $\mathbb{R}$ 上的向量空间，则 $1, \sqrt{2}, i$ 线性相关，因为 $\sqrt{2}\cdot 1+(-1)\cdot\sqrt{2}+0\cdot i = 0$；将 $\mathbb{C}$ 视作 $\mathbb{Q}$ 上的向量空间，则 $1, \sqrt{2}, i$ 线性无关，这是因为若 $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{Q}$ 满足 $\lambda_1\cdot 1+\lambda_2\cdot\sqrt{2}+\lambda_3\cdot i = 0$，则 $\mathrm{Re}(\lambda_1 \cdot 1 + \lambda_2 \cdot \sqrt{2} + \lambda_3 \cdot i) = \lambda_1 + \sqrt{2}\lambda_2 = 0$，$\mathrm{Im}(\lambda_1 \cdot 1 + \lambda_2 \cdot \sqrt{2} + \lambda_3 \cdot i) = \lambda_3 = 0$，再由 $\sqrt{2}$ 是无理数可知 $\lambda_1 = \lambda_2 = 0$。

例 1.3.2. 设 $\mathbb{K}$ 是域，一元多项式环 $\mathbb{K}[x]$ 是 $\mathbb{K}$ 上的向量空间，则对 $\forall d \in \mathbb{Z}^+$，$\{1, x, \ldots, x^{d-1}\}$ 线性无关，这是因为零多项式的系数必须全为 0。

我们很容易证明，上册的命题 2.1.1 可以直接推广到一般的向量空间上，即：设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{v}_1, \cdots, \mathbf{v}_k \in V$，$1 \leqslant i \leqslant k$，则

(i) 如果 $\mathbf{v}_1, \cdots, \mathbf{v}_i$ 线性相关，则 $\mathbf{v}_1, \cdots, \mathbf{v}_k \in \mathbb{R}^n$ 线性相关；

(ii) 如果 $\mathbf{v}_1, \cdots, \mathbf{v}_i, \cdots, \mathbf{v}_k$ 线性无关，则 $\mathbf{v}_1, \cdots, \mathbf{v}_i$ 线性无关；

(iii) $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性相关 $\Longleftrightarrow$ $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 中某个向量是其它向量的线性组合；

(iv) 设 $\mathbf{v} \in \mathbb{R}^n$ 且 $\mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性无关，则 $\mathbf{v}, \mathbf{v}_1, \cdots, \mathbf{v}_k$ 线性相关 $\Longleftrightarrow$ $\exists!$一组$\alpha_1, \cdots, \alpha_k \in \mathbb{R}$ 使得 $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_k \mathbf{v}_k$。

证明留作练习。

例 1.3.3. 易证所有 $f(x) : \mathbb{R} \to \mathbb{R}$ 的函数的集合 $\mathbb{R}^{\mathbb{R}}$ 是 $\mathbb{R}$ 上的向量空间，$\cos x$，$\sin x \in \mathbb{R}^{\mathbb{R}}$，则 $\cos x$，$\sin x$ 在 $\mathbb{R}$ 上线性无关。这是因为如果有 $\alpha, \beta \in \mathbb{R}$ 使得对任意 $x \in \mathbb{R}$ 都有 $\alpha\cos x + \beta\sin x = \mathbf{0}(x)$，则令 $x = 0$ 可得 $\alpha = 0$，令 $x = \dfrac{\pi}{2}$ 可得 $\beta = 0$。

例 1.3.4. 设 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ 两两不同，求证：$e^{\alpha_1 x}, \ldots, e^{\alpha_n x} \in \mathbb{R}^{\mathbb{R}}$ 在 $\mathbb{R}$ 上线性无关。

证明. 只需证 $\forall x \in \mathbb{R}$，$\lambda_1 e^{\alpha_1 x} + \cdots + \lambda_n e^{\alpha_n x} = \mathbf{0}(x) \Longrightarrow \lambda_1 = \cdots = \lambda_n = 0$ 即可。

在 $\lambda_1 e^{\alpha_1 x} + \cdots + \lambda_n e^{\alpha_n x} = \mathbf{0}(x)$ 中依次取 $x = 0, 1, \ldots, n-1$ 即有

$$\begin{cases} \lambda_1 e^0 + \lambda_2 e^0 + \cdots + \lambda_n e^0 = 0 \\ \lambda_1 e^{\alpha_1} + \lambda_2 e^{\alpha_2} + \cdots + \lambda_n e^{\alpha_n} = 0 \\ \vdots \\ \lambda_1 e^{(n-1)\alpha_1} + \lambda_2 e^{(n-1)\alpha_2} + \cdots + \lambda_n e^{(n-1)\alpha_n} = 0 \end{cases}$$

该方程组的系数矩阵$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{\alpha_1} & e^{\alpha_2} & \cdots & e^{\alpha_n} \\ \vdots & \vdots & & \vdots \\ (e^{\alpha_1})^{n-1} & (e^{\alpha_2})^{n-1} & \cdots & (e^{\alpha_n})^{n-1} \end{pmatrix}$是 Vandermonde 矩阵，由于 $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ 两两不同，故其行列式不为 0，所以该齐次方程组有唯一解 $\lambda_1 = \cdots = \lambda_n = 0$。 □

类似于 $\mathbb{R}^n$ 的情形，我们有

定理 1.3.1 (线性组合引理). 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{e}_1, \dots, \mathbf{e}_s, \mathbf{f}_1, \dots, \mathbf{f}_t \in V$，其中 $\mathbf{e}_1, \dots, \mathbf{e}_s$ 线性无关且每个 $\mathbf{e}_i$ 都是 $\mathbf{f}_1, \dots, \mathbf{f}_t$ 的线性组合，则 $s \leqslant t$。

证明. 用反证法。不妨设

$$\begin{aligned} \mathbf{e}_1 &= a_{11}\mathbf{f}_1 + \cdots + a_{1t}\mathbf{f}_t \\ \mathbf{e}_2 &= a_{21}\mathbf{f}_1 + \cdots + a_{2t}\mathbf{f}_t \\ &\vdots \\ \mathbf{e}_s &= a_{s1}\mathbf{f}_1 + \cdots + a_{st}\mathbf{f}_t \end{aligned}$$

其中 $a_{ij} \in \mathbb{K}$ 都是已知的。任取 $\lambda_1, \dots, \lambda_s \in \mathbb{K}$，考虑线性组合

$$\lambda_1\mathbf{e}_1 + \cdots + \lambda_s\mathbf{e}_s = \sum_{i=1}^{t}(\sum_{j=1}^{s}\lambda_j a_{ji})\mathbf{f}_i$$

那么，如果 $s > t$，由于关于 $\lambda_1, \dots, \lambda_s \in \mathbb{K}$ 的方程组

$$\begin{cases} a_{11}\lambda_1 + \cdots + a_{s1}\lambda_s = 0 \\ a_{12}\lambda_1 + \cdots + a_{s2}\lambda_s = 0 \\ \qquad\qquad\qquad\quad \vdots \\ a_{1t}\lambda_1 + \cdots + a_{st}\lambda_s = 0 \end{cases}$$

有非零解，即存在不全为 0 的 $\lambda_1, \dots, \lambda_s \in \mathbb{K}$ 使得 $\lambda_1\mathbf{e}_1 + \cdots + \lambda_s\mathbf{e}_s = \sum_{i=1}^{t} 0 \cdot \mathbf{f}_i = \mathbf{0}$，这与 $\mathbf{e}_1, \dots, \mathbf{e}_s$ 线性无关矛盾！ □

下面的内容是上册 2.1.3 小节的自然推广。

定义 1.3.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，$M \subset V$ 是 V 的**子集**，如果 M 中任意有限个向量都线性无关，则称 M 是 V 的一个线性无关组 (linearly independent set)。此外，如果 V 的两个子集 M_1, M_2 满足 $M_1 \subset \operatorname{span} M_2$，$M_2 \subset \operatorname{span} M_1$，则我们称 M_1 与 M_2 等价。

由线性组合引理可知，两个等价的线性无关组或者同时包含无穷多个向量，或者向量个数都有限且相等。

定义 1.3.3. 设 V 是域 $\mathbb{K}$ 上的向量空间，$M \subset V$ 是 V 的子集，$N \subset M$。如果 N 是线性无关组且 $\operatorname{span} N \supset M$，则称 N 是 M 的极大线性无关组，N 中的元素个数 (或者说 N 的势或基数)$|N|$ 称为向量集 M 的秩。

定义 1.3.4. 设 V 是域 $\mathbb{K}$ 上的向量空间，$M \subset V$ 是 V 的一个极大线性无关组，如果 M 是无穷集，则称 V 是无穷维向量空间；如果 M 是有限集，则称 M 是有限维向量空间，并称 M

中的元素为 V 的一组基底或基 (basis)，称 M 中的元素个数 $|M|$ 为 V 的维数 (dimension)，记为 $\dim_{\mathbb{K}}(V) = |M|$ 或简记为 $\dim(V) = |M|$。一个向量空间显然可以选取不同的基底。

例 1.3.5. (1) 设 $\mathbb{K}$ 是域，则 $\mathbb{K}[x]$ 是 $\mathbb{K}$ 上的无穷维向量空间，$M = \{1, x, x^2, \cdots\}$ 是 $\mathbb{K}[x]$ 的一组基。

(2) 设 $\mathbb{K}$ 是域，则 $\mathbb{K}[x_1, \ldots, x_n]$ 中所有的 m 次齐次多项式 (包括 0 多项式) 构成的集合 V 是 $\mathbb{K}$ 上的向量空间，它的一组基为 $\{x_1^{s_1} \cdots x_n^{s_n} \mid s_1, \ldots, s_n \in \mathbb{N} 且 s_1 + \cdots + s_n = m\}$，其维数 $\dim(V) = \binom{n+m-1}{m}$(挡板法，留作练习)。

(3) $\mathbb{K}^n$ 是 n 维向量空间，其一组基为 $\{\mathbf{e}_1 = (1, 0, 0, \ldots, 0)^t,\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0)^t, \cdots\cdots,\ \mathbf{e}_n = (0, 0, \ldots, 0, 1)^t\}$，称之为 $\mathbb{K}^n$ 的标准基。

以后如果不作特别的声明，我们所研究的向量空间都是有限维的。无穷维的向量空间将在后续课程泛函分析中继续学习。与 $\mathbb{R}^n$ 类似，我们同样有基扩充定理、包含定理和维数公式。

定理 1.3.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\dim(V) = n$。则

(1) 若 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基，则对 $\forall \mathbf{v} \in V$，存在唯一一组 $\alpha_1, \cdots, \alpha_n \in \mathbb{K}$，使得 $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$。

(2)(基扩充定理) 设 $s \leqslant n$ 且 $\mathbf{f}_1, \ldots, \mathbf{f}_s \in V$ 线性无关，则存在 $\mathbf{f}_{s+1}, \ldots, \mathbf{f}_n \in V$ 使得 $\mathbf{f}_1, \ldots, \mathbf{f}_s, \mathbf{f}_{s+1}, \cdots, \mathbf{f}_n$ 是 V 的一组基。

(3)(包含定理) 设 U, W 都是 V 的子空间且 $U \subset W$，则 $U \subsetneqq W \iff \dim(U) < \dim(W)$。

证明. (1) 由于 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基，故对任意 $\mathbf{x} \in V$，我们有 $\mathbf{x}, \mathbf{e}_1, \ldots, \mathbf{e}_n$ 线性相关，即存在不全为 0 的 $\lambda, \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ 使得

$$\lambda \mathbf{x} + \lambda_1 \mathbf{e}_1 + \cdots + \lambda_n \mathbf{e}_n = \mathbf{0}$$

下面说明 $\lambda \neq 0$。若不然，则 $\lambda_1, \ldots, \lambda_n$ 不全为 0 且 $\lambda_1 \mathbf{e}_1 + \cdots + \lambda_n \mathbf{e}_n = \mathbf{0}$，这说明 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 线性相关，与 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基相矛盾！

因此，我们可以将上式移项后两边同时乘以 λ^{-1}，得到

$$\mathbf{x} = -\lambda^{-1} \lambda_1 \mathbf{e}_1 - \cdots - \lambda^{-1} \lambda_n \mathbf{e}_n,$$

取 $\alpha_i = -\lambda^{-1} \lambda_i$ 即我们所求。

下证上式中的系数是唯一的。如果另有 $\beta_1, \ldots, \beta_n \in \mathbb{K}$ 使得 $\mathbf{v} = \sum_{i=1}^n \beta_i \mathbf{e}_i$，那么，将上式与 $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$ 作差可得 $\sum_{i=1}^n (\alpha_i - \beta_i) \mathbf{e}_i = \mathbf{0}$，由 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基可知 $\forall i = 1, \ldots, n$，$\alpha_i - \beta_i = 0$，即 $\alpha_i = \beta_i$，这就证明了唯一性。

(2) 作 $V_1 = \mathrm{span}\,\{\mathbf{f}_1, \ldots, \mathbf{f}_s\}$ 是 V 的子空间，则 $\dim(V_1) = s$。如果 $s = n$，则 $\mathbf{f}_1, \ldots, \mathbf{f}_s$ 已经是 V 的一组基，命题证毕；否则，V_1 是 V 的真子空间，于是 $\exists \mathbf{f}_{s+1} \in V \setminus V_1$，那么必有 $\mathbf{f}_1, \ldots, \mathbf{f}_s, \mathbf{f}_{s+1}$ 线性无关 (否则 $\mathbf{f}_{s+1} \in V_1$)。令 $V_2 = \mathrm{span}\,\{\mathbf{f}_1, \ldots, \mathbf{f}_s, \mathbf{f}_{s+1}\}$，则 $\dim(V_2) = s + 1$，如果 $s + 1 = n$ 则命题证毕，否则可以找到 $\mathbf{f}_{s+2} \in V \setminus V_2$ 使得 $\mathbf{f}_1, \ldots, \mathbf{f}_s, \mathbf{f}_{s+1}, \mathbf{f}_{s+2}$ 线性无关。重复以上步骤，则以上操作在有限步内一定终止 (因为 $\dim(V) = n$ 有限)，这样我们就完成了证明。[1]

(3) 显然 (或者参考上册定理 2.1.2 的证明)。 □

---

[1] 在无穷维空间中，我们仍然可以通过类似的操作得到一组基，称为 Hamel 基。证明细节见习题课讲义。

定理 1.3.3 (维数公式). 设 V 是域 $\mathbb{K}$ 上的向量空间，$U_1, U_2 \subset V$ 是子空间，则

$$\dim(U_1+U_2)+\dim(U_1\cap U_2)=\dim U_1+\dim U_2.$$

证明. 若 $U_1=\{\mathbf{0}\}$ 或 $U_2=\{\mathbf{0}\}$，则结论显然成立。

下设 $U_1\cap U_2\neq\{\mathbf{0}\}$，$\mathbf{w}_1,\cdots,\mathbf{w}_m$ 是 $U_1\cap U_2$ 的一组基。则由基扩充定理，$\exists\mathbf{u}_1,\cdots,\mathbf{u}_s\in U_1$及$\mathbf{v}_1,\cdots,\mathbf{v}_t\in U_2$ 使得：

(A)$\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s$ 是 U_1 的一组基；

(B)$\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 是 U_2 的一组基。

下面证明 $\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 是线性无关的。

如果存在 $\alpha_1,\cdots,\alpha_m,\lambda_1,\cdots,\lambda_s,\mu_1,\cdots,\mu_t\in\mathbb{R}$ 满足

$$\sum_{i=1}^{m}\alpha_i\mathbf{w}_i+\sum_{i=1}^{s}\lambda_i\mathbf{u}_i+\sum_{i=1}^{t}\mu_i\mathbf{v}_i=\mathbf{0}, \tag{$*$}$$

则令 $\mathbf{w}=\sum_{i=1}^{m}\alpha_i\mathbf{w}_i$，$\mathbf{u}=\sum_{i=1}^{s}\lambda_i\mathbf{u}_i$，$\mathbf{v}=\sum_{i=1}^{t}\mu_i\mathbf{v}_i$。显然 $\mathbf{w}\in U_1\cap U_2$，$\mathbf{u}\in U_1$，$\mathbf{v}\in U_2$。又因为 $\mathbf{w}+\mathbf{u}=-\mathbf{v}$，故也有 $\mathbf{v}\in U_1$，于是 $\mathbf{v}\in U_1\cap U_2$。则存在$\beta_1,\cdots,\beta_m\in\mathbb{R}$ 使得 $\mathbf{v}=\beta_1\mathbf{w}_1+\cdots+\beta_m\mathbf{w}_m$。即

$$(\alpha_1+\beta_1)\mathbf{w}_1+\cdots+(\alpha_m+\beta_m)\mathbf{w}_m+\lambda_1\mathbf{u}_1+\cdots+\lambda_s\mathbf{u}_s=\mathbf{0}.$$

由 (A) 知 $\lambda_1=\cdots=\lambda_s=0$，代回 ($*$) 式得到

$$\sum_{i=1}^{m}\alpha_i\mathbf{w}_i+\sum_{i=1}^{t}\mu_i\mathbf{v}_i=\mathbf{0}$$

再由 (B) 可知 $\alpha_1=\cdots=\alpha_m=\mu_1=\cdots=\mu_t=0$。

即 $\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 是线性无关的。

再证 $\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 的线性包络是 U_1+U_2。

设 $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2\in U_1+U_2$，其中 $\mathbf{x}_1\in U_1$，$\mathbf{x}_2\in U_2$，则有：

$$\mathbf{x}_1=a_1\mathbf{w}_1+\cdots+a_m\mathbf{w}_m+c_1\mathbf{u}_1+\cdots+c_s\mathbf{u}_s;$$
$$\mathbf{x}_2=b_1\mathbf{w}_1+\cdots+b_m\mathbf{w}_m+d_1\mathbf{v}_1+\cdots+d_t\mathbf{v}_t.$$

于是

$$\mathbf{x}=(a_1+b_1)\mathbf{w}_1+\cdots+(a_m+b_m)\mathbf{w}_m+c_1\mathbf{u}_1+\cdots+c_s\mathbf{u}_s+d_1\mathbf{v}_1+\cdots+d_t\mathbf{v}_t\in<\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t>$$

即 $\mathbf{w}_1,\cdots,\mathbf{w}_m,\mathbf{u}_1,\cdots,\mathbf{u}_s,\mathbf{v}_1,\cdots,\mathbf{v}_t$ 是 U_1+U_2 的基。

因此，$\dim(U_1+U_2)=m+s+t=(m+s)+(m+t)-m=\dim U_1+\dim U_2-dim(U_1\cap U_2)$。特别的。当 $U_1\cap U_2=\{\mathbf{0}\}$ 时，令 $m=0$ 即可得到结论。这样我们就完成了证明。

(与上册证明完全相同，为防止大家忘记，直接粘贴到了这里。) □

推论 1.3.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，$U_1,\dots,U_m$ 是 V 的子空间，则 $U=U_1+\cdots+U_m$ 是直和 $\iff\dim(U)=\sum_{i=1}^{m}\dim(U_i)$。

对 m 归纳即可，留作练习。

推论 1.3.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\dim(V)=n$，则对 V 的任意一个 m 维子空间 U，都

存在一个 V 的 $n-m$ 维子空间 W 使得 $V = U \oplus W$。

证明也是容易的，留作练习。此时我们称 W 是 U 在 V 中的补空间，并称 W 的维数为 U 的余维数，记作 $\text{codim}(U) = \dim(W)$。需要注意的是，补空间并不是唯一的，但所有的补空间都是同构的，我们会在后面学习商空间时证明这一点。此外，我们称余维数是 1 的子空间为超平面 (hyperplane)。

定义 1.3.5. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基。则任取 $\mathbf{v} \in V$，由定理1.3.2(1)，存在唯一一组 $(\alpha_1, \ldots, \alpha_n)^t \in \mathbb{K}^n$ 使得 $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$，我们称这组 $(\alpha_1, \ldots, \alpha_n)^t$ 为 $\mathbf{v}$ 在基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标 (coordinate)。

定理1.3.2(1) 表明向量与坐标是一一对应的。此外，同一个向量在不同的基下其坐标也是不同的，我们看下面的例子。

例 1.3.6. 我们已经知道一元多项式环 $\mathbb{R}[x]$ 是 $\mathbb{R}$ 上的向量空间，$\{1, x, x^2, \ldots\}$ 是 $\mathbb{R}[x]$ 的一组基，则多项式 $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ 在这组基下的坐标为 $(a_0, a_1, \ldots, a_n, 0, 0, \ldots)^t$。

另一方面，我们也可以固定 $a \in \mathbb{R}$，容易证明 $\{1, x-a, (x-a)^2, \ldots\}$ 也是 $\mathbb{R}[x]$ 的一组基，下面我们求 $f(x)$ 在这组基下的坐标。由泰勒公式有

$$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

于是 $f(x)$ 在基 $\{1, x-a, (x-a)^2, \ldots\}$ 下的坐标为 $(f(a), f'(a), \ldots, \frac{f^{(n)}(a)}{n!}, 0, 0, \ldots)^t$。

定义 1.3.6. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ 是 V 的两组基，则由基的定义，存在 a_{ij}, $i, j \in \{1, 2, \ldots, n\}$ 满足

$$\begin{aligned}
\mathbf{e}'_1 &= a_{11}\mathbf{e}_1 + \cdots + a_{n1}\mathbf{e}_n \\
\mathbf{e}'_2 &= a_{12}\mathbf{e}_1 + \cdots + a_{n2}\mathbf{e}_n \\
&\cdots \\
\mathbf{e}'_n &= a_{1n}\mathbf{e}_1 + \cdots + a_{nn}\mathbf{e}_n
\end{aligned}$$

这些方程可以写成矩阵的形式如下：

$$(\mathbf{e}'_1, \mathbf{e}'_2, \cdots, \mathbf{e}'_n) = (\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

我们记上面的矩阵为 A，称之为基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 到 $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ 的转换矩阵或者过渡矩阵 (transition matrix)，称向量空间从一组基换到另一组基的过程为基变换。

过渡矩阵就是将新基在旧基下的坐标按列排起来。由坐标的唯一性 (定理1.3.2(1)) 可知，一组基到另一组基的过渡矩阵一定是唯一的。容易看出一组基到自身的过渡矩阵一定是单位矩阵。

过渡矩阵一定是可逆矩阵。事实上，如果 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n$ 分别是 V 的两组基，

那么按照基的定义，我们知道：存在矩阵 A 使得

$$(\mathbf{e}_1', \mathbf{e}_2', \cdots, \mathbf{e}_n') = (\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n) \cdot A,$$

以及存在矩阵 A' 使得

$$(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n) = (\mathbf{e}_1', \mathbf{e}_2', \cdots, \mathbf{e}_n') \cdot A'$$

于是

$$(\mathbf{e}_1', \mathbf{e}_2', \cdots, \mathbf{e}_n') = (\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n) \cdot A = (\mathbf{e}_1', \mathbf{e}_2', \cdots, \mathbf{e}_n') \cdot A' \cdot A,$$

因此 $A' \cdot A = E_n$，即过渡矩阵可逆。

由基变换我们可以导出同一向量在不同基底下的坐标之间的关系，即坐标变换。设 $\mathbf{v} \in V$ 在基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标为 $(\lambda_1, \cdots, \lambda_n)^t$，在基 $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$ 下的坐标为 $(\lambda_1', \cdots, \lambda_n')^t$，则

$$\mathbf{v} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix} = (\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n')\begin{pmatrix}\lambda_1'\\ \vdots\\ \lambda_n'\end{pmatrix} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot A \cdot \begin{pmatrix}\lambda_1'\\ \vdots\\ \lambda_n'\end{pmatrix}$$

通过对比上式，我们立刻有

$$\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix} = A \cdot \begin{pmatrix}\lambda_1'\\ \vdots\\ \lambda_n'\end{pmatrix} \qquad \text{或者} \qquad \begin{pmatrix}\lambda_1'\\ \vdots\\ \lambda_n'\end{pmatrix} = A^{-1} \cdot \begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}. \tag{1.3.1}$$

此即**坐标变换公式**[1]。

最后我们来考虑向量空间的同构。

定义 1.3.7. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，如果映射 $f : V \to W$ 满足：对 $\forall \alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in V$，都有

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

则我们称 f 是 V 到 W 的线性映射 (linear map)。于是我们也可以定义线性映射的核 $\ker(f) = \{\mathbf{x} \in V \mid f(\mathbf{x}) = \mathbf{0}_W\}$，像 $\mathrm{im}(f) = \{\mathbf{y} \in W \mid \exists \mathbf{x} \in V \text{使得} \mathbf{y} = f(\mathbf{x})\}$。容易验证 $\ker(f)$ 是 V 的子空间，$\mathrm{im}(f)$ 是 W 的子空间，并且 f 是单射 $\iff \ker(f) = \{\mathbf{0}\}$，$f$ 是满射 $\iff \mathrm{im}(f) = W$(留作练习)。

定义 1.3.8. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，如果存线性映射 $f : V \to W$ 是双射，则我们称向量空间 V 和 W 同构，记作 $V \simeq W$。

显然同构是一个等价关系。我们有以下定理：

定理 1.3.4. 设 $\mathbb{K}$ 是域，$n \in \mathbb{Z}^+$，则 $\mathbb{K}$ 上所有 n 维向量空间 V 都同构于 $\mathbb{K}^n$。

证明. 取 V 的一组基 $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$，那么对 $\forall \mathbf{v} \in V$，$\mathbf{v}$ 有一组坐标 $(\lambda_1, \cdots, \lambda_n)^t$，即

$$\mathbf{v} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}$$

---

[1] 可以这样记：新基 = 旧基 × 过渡矩阵，新坐标 = 过渡矩阵的逆 × 旧坐标。

我们作如下映射:

$$\begin{aligned} f : V &\longrightarrow \mathbb{K}^n \\ \mathbf{v} &\longmapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \end{aligned}$$

显然 f 是双射 (向量和坐标的一一对应)，下面证明 f 是线性映射。任取 $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$，$\mathbf{y} = \sum_{i=1}^n \beta_i \mathbf{e}_i \in V$，我们有

$$f(\mathbf{y}) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad f(\mathbf{y}) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

另一方面，任取 $\alpha, \beta \in \mathbb{K}$，我们有

$$\alpha \mathbf{x} + \beta \mathbf{y} = \alpha(\sum_{i=1}^n \alpha_i \mathbf{e}_i) + \beta(\sum_{i=1}^n \beta_i \mathbf{e}_i) = \sum_{i=1}^n (\alpha\alpha_i + \beta\beta_i)\mathbf{e}_i$$

于是

$$f(\alpha \mathbf{x} + \beta \mathbf{y}) = \begin{pmatrix} \alpha\alpha_1 + \beta\beta_1 \\ \vdots \\ \alpha\alpha_n + \beta\beta_n \end{pmatrix}$$

所以我们有

$$\alpha f(\mathbf{x}) + \beta f(\mathbf{y}) = \alpha \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \beta \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 + \beta\beta_1 \\ \vdots \\ \alpha\alpha_n + \beta\beta_n \end{pmatrix} = f(\alpha \mathbf{x} + \beta \mathbf{y}).$$

即 f 保持线性运算。这样我们就完成了证明。 □

这个定理告诉我们，有限维向量空间在同构意义下每个维数只有一种结构，即 $\mathbb{K}^n$。所以我们可以把一切向量空间的结论用坐标的形式表达出来，也可以将一切 $\mathbb{K}^n$ 的性质推广到一般的 n 维向量空间上。

1.4 商空间

我们已经知道，给定域 $\mathbb{K}$ 上的向量空间 V 及子空间 W，其补空间不唯一。我们就从这里开始讨论。

例 1.4.1. 取向量空间 $V = \mathbb{R}^2$ 及子空间 $W = \operatorname{span}\{(1,0)^t\}$。令子空间 $U_1 = \operatorname{span}\{(0,1)^t\}$，$U_2 = \operatorname{span}\{(1,1)^t\}$，则容易验证 $U_1 \oplus W = U_2 \oplus W = V$，但显然 U_1, U_2 是两个不同的子空间。

我们注意到，虽然补空间不唯一，但余维数是固定的，即两个补空间之间一定存在同构映射。那么，是否可以构造一个“标准”的子空间，它同构于所有的补空间呢？这就是我们下面定义的商空间 (quotient space)。

首先，设 V 是域 $\mathbb{K}$ 上的向量空间，子空间 $U \subset V$。任取 $\mathbf{x} \in V$，我们定义 U 的陪集 (coset)：$\mathbf{x} + U = \{\mathbf{x} + \mathbf{u} \mid \mathbf{u} \in U\}$，称这里的 $\mathbf{x}$ 为陪集的代表元，则：

(1) 取 $\mathbf{x}_1$，$\mathbf{x}_2 \in V$，则 $\mathbf{x}_1 + U = \mathbf{x}_2 + U \iff \mathbf{x}_1 - \mathbf{x}_2 \in U$。这是因为：

$(\Longrightarrow)\mathbf{x}_1 + U = \mathbf{x}_2 + U$ 即任取 $\mathbf{y} = \mathbf{x}_1 + \mathbf{u}_1 \in \mathbf{x}_1 + U$，存在唯一的 $\mathbf{u}_2 \in U$ 使得 $\mathbf{y} = \mathbf{x}_2 + \mathbf{u}_2$，所以 $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{u}_2 - \mathbf{u}_1 \in U$；

$(\Longleftarrow)$ 设 $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{u} \in U$，则若 $\mathbf{y} = \mathbf{x}_1 + \mathbf{u}_1 \in \mathbf{x}_1 + U$，则 $\mathbf{y} = \mathbf{x}_2 + (\mathbf{x}_1 - \mathbf{x}_2) + \mathbf{u}_1 = \mathbf{x}_2 + \mathbf{u} + \mathbf{u}_1$，显然 $\mathbf{u} + \mathbf{u}_1 \in U$，所以我们有 $\mathbf{y} \in \mathbf{x}_2 + U$，即 $\mathbf{x}_1 + U \subset \mathbf{x}_2 + U$。同理可证 $\mathbf{x}_1 + U \supset \mathbf{x}_2 + U$，所以 $\mathbf{x}_1 + U = \mathbf{x}_2 + U$。

由以上结论可得，若 $\mathbf{x}_1 + U \cap \mathbf{x}_2 + U \neq \emptyset$，就有 $\mathbf{x}_1 + U = \mathbf{x}_2 + U$。这是因为：若 $\mathbf{y} = \mathbf{x}_1 + \mathbf{u}_1 = \mathbf{x}_2 + \mathbf{u}_2 \in \mathbf{x}_1 + U \cap \mathbf{x}_2 + U$，则 $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{u}_2 - \mathbf{u}_1 \in U$，即有 $\mathbf{x}_1 + U = \mathbf{x}_2 + U$。

(2) 固定子空间 U，任取 $\mathbf{x} \in V$，则存在唯一的陪集 $\mathbf{x} + U$ 使得 $\mathbf{x} \in \mathbf{x} + U$。存在性是显然的，下面我们说明唯一性。若 $\mathbf{x}$ 也在陪集 $\mathbf{y} + U$ 中，则 $\mathbf{x} \in \mathbf{x} + U \cap \mathbf{y} + U$，于是由上面的结论就有 $\mathbf{x} + U = \mathbf{y} + U$，这就证明了唯一性。

由以上两点可知，$V = \bigcup\limits_{\mathbf{x} \in V} (\mathbf{x} + U)$，即所有陪集构成了向量空间 V 的一个分割 (上册 1.4.5 小节)，也即“$\mathbf{x}_1 + U = \mathbf{x}_2 + U \iff \mathbf{x}_1 - \mathbf{x}_2 \in U$”是 V 上的一个等价关系。于是我们可以作商集合 $V/U = \{\mathbf{x} + U \mid \forall \mathbf{x} \in V\}$。要使商集合也成为向量空间，我们需要在其上定义合适的加法和数乘运算。我们给出如下的自然定义：

$$\begin{aligned}
+ \ : \ & V/U \times V/U \longrightarrow V/U \\
& (\mathbf{x} + U, \mathbf{y} + U) \longmapsto (\mathbf{x} + \mathbf{y}) + U \\
\cdot \ : \ & \mathbb{K} \times V/U \longrightarrow V/U \\
& (\lambda, \mathbf{x} + U) \longmapsto (\lambda \mathbf{x}) + U
\end{aligned}$$

首先我们需要验证在商集合上定义的这两个运算是良定义的，即我们需要运算的结果不依赖于代表元的选取。

首先验证加法：若 $\mathbf{x}_1 + U = \mathbf{x}_2 + U$，$\mathbf{y}_1 + U = \mathbf{y}_2 + U$，则 $\mathbf{x}_1 - \mathbf{x}_2 \in U$，$\mathbf{y}_1 - \mathbf{y}_2 \in U$，所以 $(\mathbf{x}_1 + \mathbf{y}_1) - (\mathbf{x}_2 + \mathbf{y}_2) \in U$，即 $(\mathbf{x}_1 + \mathbf{y}_1) + U = (\mathbf{x}_2 + \mathbf{y}_2) + U$，即加法是良定义的。

其次验证数乘：若 $\mathbf{x}_1 + U = \mathbf{x}_2 + U$，即 $\mathbf{x}_1 - \mathbf{x}_2 \in U$，所以 $\lambda(\mathbf{x}_1 - \mathbf{x}_2) \in U$，即 $\lambda\mathbf{x}_1 + U = \lambda\mathbf{x}_2 + U$，即数乘是良定义的。

最后，V/U 在上面定义的加法和数乘下是向量空间。V/U 关于加法成交换群：封闭性显然，交换律和结合律来自 V 上加法的交换律和结合律，零元是陪集 $\mathbf{0} + U$，$\mathbf{x} + U$ 的负元

是 $-\mathbf{x}+U$。V/U 关于数乘的结合律来自 V 的数乘结合律，$1_{\mathbb{K}}$ 仍满足酉性。分配律也来自 V 上的分配律。由此，V/U 是向量空间，我们称之为**商空间**。

显然 V 到 V/U 有一个自然的映射 $\pi : V \to V/U$, $\mathbf{x} \mapsto \mathbf{x}+U$，容易验证 π 是满的线性映射，留作练习。我们称 π 是 V 到 V/U 的自然投射。

我们有如下定理：

定理 1.4.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，U, M 是 V 的子空间并且 $V = U \oplus M$。则映射

$$f : M \longrightarrow V/U,\ \mathbf{x} \longmapsto \mathbf{x}+U$$

是向量空间之间的同构[1]。

证明. 首先，f 是线性映射：任取 $\alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in M$，我们有

$$\begin{aligned} \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) &= \alpha(\mathbf{x}+U) + \beta(\mathbf{y}+U) \\ &= (\alpha\mathbf{x}+\beta\mathbf{y}) + U \\ &= f(\alpha\mathbf{x}+\beta\mathbf{y}) \end{aligned}$$

其次，f 是单射，这只需证 $\ker(f) = \{\mathbf{0}\}$。若 $f(\mathbf{x}) = \mathbf{0}+U$，即 $\mathbf{x}+U = \mathbf{0}+U$，所以 $\mathbf{x} \in U$，然而 $\mathbf{x} \in M$，即 $\mathbf{x} \in U \cap M$，由 $U+M$ 是直和可知 $U \cap M = \{\mathbf{0}\}$，所以 $\mathbf{x} = \mathbf{0}$，此即 $\ker(f) = \{\mathbf{0}\}$。

最后，f 是满射。任取陪集 $\mathbf{v}+U \in V/U$，由 $V = U \oplus M$ 可知存在唯一的 $\mathbf{u} \in U$ 及 $\mathbf{x} \in M$ 使得 $\mathbf{v} = \mathbf{u}+\mathbf{x}$，则 $\mathbf{x}+U = \mathbf{v}+U$，所以 $f(\mathbf{x}) = \mathbf{x}+U = \mathbf{v}+U$。这就说明 f 是满射。

综上所述，f 是同构。 □

我们注意到定理1.4.1的证明中并不涉及到基的选取，因此可以推广到无穷维向量空间。

推论 1.4.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，U 是 V 的子空间，则 $\dim(V/U) = \operatorname{codim}(U) = \dim(V) - \dim(U)$。

我们也可以用上面的定理证明无穷维向量空间中补空间的存在性，并将余维数定义成商空间的维数，细节留作思考。

类比讲义上册的定理 1.4.1(以及抽象代数中的群同态基本定理)，我们有以下定理：

定理 1.4.2. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，线性映射 $f : V \to W$ 是满射，则 f 诱导了同构 $\widetilde{f} : V/\ker(f) \to W$，并且 $f = \widetilde{f} \circ \pi$。

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/\ker(f) \arrow[ur, "\widetilde{f}"'] & \end{tikzcd}$$

证明. 我们定义

$$\begin{aligned} \widetilde{f} : V/\ker(f) &\longrightarrow W \\ \mathbf{x}+\ker(f) &\longmapsto f(\mathbf{x}). \end{aligned}$$

首先 $\widetilde{f}$ 是良定义的：设 $\mathbf{x}_1+\ker(f) = \mathbf{x}_2+\ker(f)$ ，则 $\mathbf{x}_1-\mathbf{x}_2 \in \ker(f)$，即 $f(\mathbf{x}_1-\mathbf{x}_2) = \mathbf{0}$，由 f 是线性映射即得 $f(\mathbf{x}_1) = f(\mathbf{x}_2)$，也即 $\widetilde{f}(\mathbf{x}_1+\ker(f)) = \widetilde{f}(\mathbf{x}_2+\ker(f))$，即 $\widetilde{f}$ 的值与代表元的选取无关，这满足映射的定义。

---

[1] 这是一个自然的同构，与基底的选取无关。

其次 $\widetilde{f}$ 是线性映射：任取 $\alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x} + \ker(f), \mathbf{y} + \ker(f) \in V/\ker(f)$，我们有

$$\begin{aligned}
&\alpha\widetilde{f}(\mathbf{x} + \ker(f)) + \beta\widetilde{f}(\mathbf{y} + \ker(f)) \\
={}& \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \\
={}& f(\alpha\mathbf{x} + \beta\mathbf{y}) \\
={}& \widetilde{f}((\alpha\mathbf{x} + \beta\mathbf{y}) + \ker(f)) \\
={}& \widetilde{f}(\alpha(\mathbf{x} + \ker(f)) + \beta(\mathbf{y} + \ker(f)))
\end{aligned}$$

再次 $\widetilde{f}$ 是双射：首先由于 f 是满射，即任取 $\mathbf{w} \in W$，存在 $\mathbf{x} \in V$ 使得 $f(\mathbf{x}) = \mathbf{w}$。那么，存在 $\mathbf{x} + \ker(f) \in V/\ker(f)$ 使得 $\widetilde{f}(\mathbf{x} + \ker(f)) = f(\mathbf{x}) = \mathbf{w}$。此即 $\widetilde{f}$ 是满射。再设 $\widetilde{f}(\mathbf{x} + \ker(f)) = \widetilde{f}(\mathbf{y} + \ker(f))$，即 $f(\mathbf{x}) = f(\mathbf{y})$，所以 $f(\mathbf{x} - \mathbf{y}) = \mathbf{0}$，即 $\mathbf{x} - \mathbf{y} \in \ker(f)$，所以 $\mathbf{x} + \ker(f) = \mathbf{y} + \ker(f)$，即 $\widetilde{f}$ 是单射。

最后验证 $f = \widetilde{f} \circ \pi$：对 $\forall \mathbf{x} \in V$, $\widetilde{f} \circ \pi(\mathbf{x}) = \widetilde{f}(\mathbf{x} + \ker(f)) = f(\mathbf{x})$，即 $f = \widetilde{f} \circ \pi$。 □

1.5 线性函数与对偶空间

定义 1.5.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，函数 $f: V \to \mathbb{K}$ 满足：对任意 $\alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in V$，有

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

(或者等价地，$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$，$f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$。) 则我们称 f 是 V 到 $\mathbb{K}$ 的线性函数。

例 1.5.1. 我们知道 $M_n(\mathbb{K})$ 是域 $\mathbb{K}$ 上的向量空间，我们定义矩阵的迹 (trace) 如下：

$$\mathrm{tr}: \ M_n(\mathbb{K}) \longrightarrow \mathbb{K}, \quad A = (a_{ij})_{n\times n} \longmapsto \sum_{i=1}^{n} a_{ii}$$

则迹函数是线性函数。验证如下：任取 $\alpha, \beta \in \mathbb{K}$ 及 $A = (a_{ij})$，$B = (b_{ij}) \in M_n(\mathbb{K})$，我们有

$$\begin{aligned}\mathrm{tr}(\alpha A + \beta B) &= \sum_{i=1}^{n} (\alpha a_{ii} + \beta b_{ii}) \\ &= \alpha \sum_{i=1}^{n} a_{ii} + \beta \sum_{i=1}^{n} b_{ii} \\ &= \alpha \mathrm{tr}(A) + \beta \mathrm{tr}(B)\end{aligned}$$

即 tr 是线性函数。

定义 1.5.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，令 $V^* = \{V \to \mathbb{K}\text{的所有线性函数}\}$，在 V^* 上定义如下的加法和数乘运算：设 $f, g \in V^*$ 是线性函数，$\lambda \in \mathbb{K}$，定义

$$\begin{aligned} + \ : \ & V^* \times V^* \longrightarrow V^* \\ & (f, g) \longmapsto (f+g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}), \ \forall \mathbf{x} \in V \\ \cdot \ : \ & \mathbb{K} \times V^* \longrightarrow V^* \\ & (\lambda, f) \longmapsto (\lambda f)(\mathbf{x}) = \lambda f(\mathbf{x}), \ \forall \mathbf{x} \in V \end{aligned}$$

容易验证 V^* 在上述运算下构成了 $\mathbb{K}$ 上的向量空间，称为 V 的对偶空间 (dual space)。

通常而言，验证一个函数具有某种性质总是比较困难的，因为我们需要将函数作用到定义域的每个点上。但对于线性函数，我们可以比较简单地处理。设 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是向量空间 V 的一组基，$f \in V^*$，则 f 由 $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$ 的值唯一确定。这是因为任取 $\mathbf{x} \in V$，存在唯一一组 $(\lambda_1, \ldots, \lambda_n)^t \in \mathbb{K}^n$ 使得 $\mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{e}_i$，于是 $f(\mathbf{x}) = \sum_{i=1}^n \lambda_i f(\mathbf{e}_i)$ 就确定了。

定义 1.5.3. 我们定义一些特殊的线性函数：设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是一组基，定义线性函数 $\mathbf{e}^i \in V^*$，$i \in \{1, \ldots, n\}$ 如下：

$$\begin{aligned} \mathbf{e}^i \ : \ & V \longrightarrow \mathbb{K} \\ & \mathbf{x} = \sum_{j=1}^{n} \lambda_j \mathbf{e}_j \longmapsto \lambda_i \end{aligned}$$

也即 $\mathbf{e}^i$ 是由 $\mathbf{e}^i(\mathbf{e}_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ (Kronecker 符号，见讲义上册定义 3.2.1) 唯一确定的

线性函数。有时我们也用 $\mathbf{e}_i^*$ 来记 $\mathbf{e}^i$。此外，我们把零函数记作 $\mathbf{0}^*$。

定理 1.5.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是一组基，则 $\mathbf{e}^1, \ldots, \mathbf{e}^n$ 是 V^* 的一组基，称为 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 的对偶基。

证明. 首先，$\mathbf{e}^1, \ldots, \mathbf{e}^n$ 线性无关，这是因为如果线性组合 $\lambda_1\mathbf{e}^1 + \cdots + \lambda_n\mathbf{e}^n = \mathbf{0}^*$，则对 $\forall i \in \{1, \ldots, n\}$，有

$$0 = \mathbf{0}^*(\mathbf{e}_i) = (\lambda_1\mathbf{e}^1 + \cdots + \lambda_n\mathbf{e}^n)(\mathbf{e}_i) = \lambda_i.$$

此即 $\mathbf{e}^1, \ldots, \mathbf{e}^n$ 线性无关。

其次，任取 $f \in V^*$，都有 $f \in \operatorname{span}\{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$. 这是因为：令 $\lambda_i = f(\mathbf{e}_i)$，则对 $\forall \mathbf{x} = \sum_{i=1}^n \alpha_i\mathbf{e}_i \in V$，都有

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i f(\mathbf{e}_i) = \sum_{i=1}^n \alpha_i\lambda_i$$

另一方面，

$$(\sum_{i=1}^n \lambda_i\mathbf{e}^i)(\mathbf{x}) = \sum_{i=1}^n \lambda_i(\sum_{i=j}^n \alpha_j\mathbf{e}^i(\mathbf{e}_j)) = \sum_{i=1}^n \lambda_i\alpha_i$$

即 $f = \sum_{i=1}^n \lambda_i\mathbf{e}^i$。此即 $f \in \operatorname{span}\{\mathbf{e}^1, \ldots, \mathbf{e}^n\}$。

综上所述，$\mathbf{e}^1, \ldots, \mathbf{e}^n$ 是 V^* 的一组基。 □

从该定理的结论也可以看出，如果取定 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是有限维空间 V 的一组基，$\mathbf{e}^1, \ldots, \mathbf{e}^n$ 是其对偶基。那么，任取 $\mathbf{x} = \sum_{i=1}^n \alpha_i\mathbf{e}_i \in V$ 及 $f = \sum_{i=1}^n \lambda_i\mathbf{e}^i \in V^*$，我们有

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i\lambda_i = \begin{pmatrix}\lambda_1 & \cdots & \lambda_n\end{pmatrix} \cdot \begin{pmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{pmatrix}.$$

也就是说，如果我们把原空间中的向量视作列向量 (坐标)，那么对偶空间中的元素可以视作行向量，其作用到原空间上的方式就是行向量乘以列向量。我们会在第三章更加深入地讨论对偶空间与原空间的联系。

推论 1.5.1. 对于有限维向量空间 V，有 $\dim(V) = \dim(V^*)$。

这个推论告诉我们，对有限维向量空间 V 而言，有 $V \simeq V^*$，并且如果 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基，则映射 $\mathbf{e}_i \mapsto \mathbf{e}^i$ 是同构映射。我们可以看到，这个同构映射依赖于基的选取。我们接下来会看到，V 和 V 的对偶的对偶的同构是不依赖于基底的选择的。

例 1.5.2. $\mathbb{R}^n$ 的标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 有对偶基 $\mathbf{e}^1, \ldots, \mathbf{e}^n$，其中 $\mathbf{e}^i$ 是满足

$$\mathbf{e}^i(\mathbf{e}_j) = \delta_{ij} = \begin{cases}1, & i = j \\ 0, & i \neq j\end{cases}$$

的线性函数。

例 1.5.3. 我们知道 $V = \{f \in \mathbb{K}[x] \mid \deg(f) \leqslant n-1\}$ 是域 $\mathbb{K}$ 上的 n 维向量空间，$\{1, x, \ldots, x^{n-1}\}$ 是 V 的一组基，则其对偶基为

$$1^* = \mathrm{id},\ x^* = \frac{\mathrm{d}}{\mathrm{d}x}\Big|_{x=0}, \ldots,\ (x^{n-1})^* = \frac{1}{(n-1)!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big|_{x=0}$$

这是因为 $\frac{1}{i!}\frac{\mathrm{d}^i}{\mathrm{d}x^i}\Big|_{x=0}(x^j) = \delta_{ij}$ 。

我们也可以考虑对偶空间的对偶 V^{**}，在有限维的情形下，V^{**} 的结构是简单的。

首先我们说明 V^{**} 中的元素是什么。设 $\theta \in V^{**}$，则 θ 是 $V^* \to \mathbb{K}$ 的函数，即任意 $f \in V^*$，$\theta(f)$ 是 $\mathbb{K}$ 中的一个数。我们首先考虑 V^{**} 中一些特殊的元素，它们是由 V 中的元素确定的。固定 $\mathbf{x} \in V$，我们定义 $\varepsilon_{\mathbf{x}} \in V^{**}$ 如下：对任意 $f \in V^*$，$\varepsilon_{\mathbf{x}}(f) = f(\mathbf{x})$。可以验证

$$\varepsilon_{\mathbf{x}}(\alpha f + \beta g) = \alpha f(\mathbf{x}) + \beta g(\mathbf{x}) = \alpha\varepsilon_{\mathbf{x}}(f) + \beta\varepsilon_{\mathbf{x}}(g).$$

因此，$\varepsilon_{\mathbf{x}} \in V^{**}$ 是 $V^* \to \mathbb{K}$ 的线性函数。取遍 $\mathbf{x} \in V$，我们将所有的 $\varepsilon_{\mathbf{x}}$ 放在一起做成一个集合，容易证明这个集合在 V^{**} 的运算下是 V^{**} 的子空间。稍后我们会证明，在有限维空间的条件下，上述子空间就是 V^{**} 本身。我们称 $V \to V^{**}$ 的映射：ε : $\mathbf{x} \to \varepsilon_{\mathbf{x}}$ 为自然映射或典型映射。

定理 1.5.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，则 V 与 V^{**} 同构，并且同构映射可以取自然映射。

证明. 取自然映射

$$\begin{aligned} \varepsilon: \ V &\longrightarrow V^{**} \\ \mathbf{x} &\longmapsto \varepsilon_{\mathbf{x}}\ ,\ \text{其中 } \varepsilon_{\mathbf{x}}(f) = f(\mathbf{x}) \text{ 对 } \forall f \in V^* \text{ 成立.} \end{aligned}$$

首先，ε 是线性映射，理由如下。任取 $\alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in V$，我们有 $\varepsilon(\alpha\mathbf{x} + \beta\mathbf{y}) = \varepsilon_{\alpha\mathbf{x}+\beta\mathbf{y}}$，注意到对任意的 $f \in V^*$，我们有

$$\varepsilon_{\alpha\mathbf{x}+\beta\mathbf{y}}(f) = f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) = \alpha\varepsilon_{\mathbf{x}}(f) + \beta\varepsilon_{\mathbf{y}}(f) = (\alpha\varepsilon_{\mathbf{x}} + \beta\varepsilon_{\mathbf{y}})(f)$$

即 $\varepsilon(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\varepsilon(\mathbf{x}) + \beta\varepsilon(\mathbf{y})$。此即 ε 是线性映射。

其次，ε 是单射。我们只需证明 $\ker(\varepsilon) = \{\mathbf{0}\}$。设 $\mathbf{x} \in \ker(\varepsilon)$，即 $\varepsilon_{\mathbf{x}} = \mathbf{0}^{**}$，也即对 $\forall f \in V^*$ 有

$$f(\mathbf{x}) = \varepsilon_{\mathbf{x}}(f) = \mathbf{0}^{**}(f) = 0$$

如果 $\mathbf{x} \neq \mathbf{0}$，则由基扩充定理，可以将 $\mathbf{x}$ 扩充成 V 的一组基 $\mathbf{e}_1 = \mathbf{x}, \mathbf{e}_2 \ldots, \mathbf{e}_n$，取这组基的对偶基 $\mathbf{e}^1, \ldots, \mathbf{e}^n$，则 $\mathbf{e}^1(\mathbf{x}) = 1$，这与 $\forall f \in V^*$，$f(\mathbf{x}) = 0$ 矛盾！故由 $\mathbf{x} \in \ker(\varepsilon)$ 即可得到 $\mathbf{x} = \mathbf{0}$，即 ε 是单射。

最后，ε 是满射。这是因为由推论1.5.1立刻有 $\dim(V^{**}) = \dim(V^*) = \dim(V) = n$，如果 ε 不是满射，则 $\dim(V^{**}) > \dim(V) = n$ 矛盾！

综上，n 维空间 $V \simeq V^{**}$。 □

例 1.5.4. 推论1.5.1和定理1.5.2对无穷维空间都不成立。例如取 $\mathbb{Q}[x]$ 是 $\mathbb{Q}$ 上的无穷维向量空间，它有一组可数的基底 $\{1, x, \ldots, x^n, \ldots\}$，现在我们考虑 $\mathbb{Q}[x]$ 的对偶空间。$\mathbb{Q}[x] \to \mathbb{Q}$ 的线性函数应该把每个基底中的元素映到一个有理数，从而将任意一个多项式映到这些有理数数的线性组合。因此，$\mathbb{Q}[x]$ 的对偶空间中的元素可以视作有理数的无穷序列 $(a_0, a_1, \ldots)$。我们知道(或者以后会知道) 所有有理数的无穷序列构成的集合是不可数的，然而 $\mathbb{Q}[x]$ 是可数的，因此 $\mathbb{Q}[x]$ 和它的对偶空间之间不可能有同构。实际上，在无穷维空间上，我们总有 $\dim(V^*)$(定义为 V^*Hamel 基的势) 等于 V^* 本身的势，而 V^* 本身的势大于或等于 $\dim(V)$ 幂集的势，从而

此时对偶空间的维数总是严格大于原空间的维数（细节留作思考，或参考 Advanced Linear Algebra, Steve Roman, GTM135 的 Theorem 3.12）。因此，$\dim(V^{**}) > \dim(V^*) > \dim(V)$，从而 V 和 V^{**} 不同构。

定理 1.5.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$f_1, \ldots, f_n$ 是 V^* 的一组基，则

(1) 存在唯一的 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得其对偶基恰为 $f_1, \ldots, f_n$。

(2) 若 $\mathbf{v} \in V$ 满足对任意 f_i，$i \in \{1, \ldots, n\}$ 都有 $f_i(\mathbf{v}) = 0$，则必有 $\mathbf{v} = \mathbf{0}$。

(3) 任取 V 的子空间 $U = \operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$，则 $\dim(U) = \operatorname{rank}(A)$，其中矩阵 $A = \begin{pmatrix} f_1(\mathbf{x}_1) & \cdots & f_1(\mathbf{x}_m) \\ f_2(\mathbf{x}_1) & \cdots & f_2(\mathbf{x}_m) \\ \vdots & & \vdots \\ f_n(\mathbf{x}_1) & \cdots & f_n(\mathbf{x}_m) \end{pmatrix}$。

证明. (1) 由于 $f_1, \ldots, f_n$ 是 V^* 的基，故存在唯一一组对偶基：$f_1^*, \ldots, f_n^* \in V^{**}$ 使得 $f_i^*(f_j) = \delta_{ij}$。由定理1.5.2可知对每个 $f_i^* \in V^{**}$，存在唯一的 $\mathbf{e}_i \in V$ 使得对任意 $f \in V^*$，都有 $f_i^*(f) = f(\mathbf{e}_i)$。特别地，$f_j(\mathbf{e}_i) = f_i^*(f_j) = \delta_{ij}$，即对偶基。

(2) 由 $f_1, \ldots, f_n$ 是 V^* 的基可知任意 $f \in V^*$，都有 $f(\mathbf{v}) = 0$，那么由定理1.5.2，我们立刻有 $\varepsilon_{\mathbf{v}}(f) = 0$ 对任意 $f \in V^*$ 都成立，所以 $\varepsilon_{\mathbf{v}} = \mathbf{0}^{**}$，故再由定理1.5.2可知 $\mathbf{v} = \mathbf{0}$。

(3) 设 $r = \operatorname{rank}(A)$，不失一般性，不妨设 A 的前 r 列 $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(r)}$ 线性无关，我们只需证 $\mathbf{x}_1, \ldots \mathbf{x}_r$ 是子空间 U 的一组基。为此，只需证 $\mathbf{x}_1, \ldots \mathbf{x}_r$ 线性无关并且 $\mathbf{x}_{r+1}, \ldots \mathbf{x}_m$ 能够写成 $\mathbf{x}_1, \ldots \mathbf{x}_r$ 的线性组合即可。

首先，如果 $\mathbf{x}_1, \ldots \mathbf{x}_r$ 的有一个系数不全为 0 的线性组合 $\lambda_1\mathbf{x}_1 + \cdots + \lambda_r\mathbf{x}_r = \mathbf{0}$，那么，对每个 f_i，$i = 1, \ldots, n$，都有 $f_i(\lambda_1\mathbf{x}_1 + \cdots + \lambda_r\mathbf{x}_r) = 0$，即

$$\begin{aligned} &\lambda_1 f_1(\mathbf{x}_1) + \cdots + \lambda_r f_1(\mathbf{x}_r) = 0 \\ &\lambda_1 f_2(\mathbf{x}_1) + \cdots + \lambda_r f_2(\mathbf{x}_r) = 0 \\ &\cdots \\ &\lambda_1 f_n(\mathbf{x}_1) + \cdots + \lambda_r f_n(\mathbf{x}_r) = 0 \end{aligned}$$

即 $\lambda_1\mathbf{A}^{(1)} + \cdots + \lambda_r\mathbf{A}^{(r)} = \mathbf{0}$，所以 $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(r)}$ 线性相关，矛盾！故 $\mathbf{x}_1, \ldots \mathbf{x}_r$ 线性无关。

其次，由于 $\operatorname{rank}(A) = r$，则对任意的 $j = r+1, \ldots, m$，都有 $\mathbf{A}^{(j)} \in \operatorname{span}\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(r)}\}$，即存在 $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$ 使得 $\mathbf{A}^{(j)} = \alpha_1\mathbf{A}^{(1)} + \cdots + \alpha_r\mathbf{A}^{(r)}$，具体写出来就是

$$\begin{aligned} f_1(\mathbf{x}_j) &= \alpha_1 f_1(\mathbf{x}_1) + \cdots + \alpha_r f_1(\mathbf{x}_r) \\ f_2(\mathbf{x}_j) &= \alpha_1 f_2(\mathbf{x}_1) + \cdots + \alpha_r f_2(\mathbf{x}_r) \\ &\cdots\cdots \\ f_n(\mathbf{x}_j) &= \alpha_1 f_n(\mathbf{x}_1) + \cdots + \alpha_r f_n(\mathbf{x}_r) \end{aligned}$$

即对任意的 f_i，$i = 1, \ldots, n$，都有

$$f_i(\mathbf{x}_j - \alpha_1\mathbf{x}_1 - \cdots - \alpha_r\mathbf{x}_r) = 0$$

注意到 $f_1,\ldots,f_n$ 是 V^* 的一组基，由 (2) 的结论即有 $\mathbf{x}_j = \alpha_1\mathbf{x}_1 + \cdots + \alpha_r\mathbf{x}_r$，也即 $\forall j = r+1,\ldots,m$，$\mathbf{x}_j$ 都能够写成 $\mathbf{x}_1,\ldots\mathbf{x}_r$ 的线性组合。这样我们就完成了证明。

□

注 1.5.1. 上面的定理证明过程中，(1) 和 (2) 的证明也可以绕开定理1.5.2进行。

对于 (1)，我们可以任取 V 的一组基 $\varepsilon_1,\ldots,\varepsilon_n$，作其对偶基 $\varepsilon^1,\ldots,\varepsilon^n$，则 $\varepsilon^1,\ldots,\varepsilon^n$ 与 $f_1,\ldots,f_n$ 之间有基变换 $(f_1,\ldots,f_n) = (\varepsilon^1,\ldots,\varepsilon^n)\cdot A$，$A$ 是可逆矩阵，作 V 上的基变换 $(\mathbf{e}_1,\ldots,\mathbf{e}_n) = (\varepsilon_1,\ldots,\varepsilon_n)\cdot(A^t)^{-1}$，容易验证 $f_1,\ldots,f_n$ 是 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 的对偶基 (细节留作练习)。

对于 (2)，我们利用 (1) 的结论，取 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得 $f_1,\ldots,f_n$ 是 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 的对偶基，设 $\mathbf{v} = \sum_{i=1}^n \lambda_i\mathbf{e}_i$，则由 $f_i(\mathbf{v}) = \lambda_i = 0$ 可知 $\mathbf{v} = \mathbf{0}$。

这两种证明方法体现了线性代数的两条路径，一条路径是从抽象空间本身的性质出发，最后落回具体问题；另一条路径则是通过基的选取和变换将问题转换成具体的向量、矩阵、方程组的问题从而加以解决，有时还会回到抽象的问题。

推论 1.5.2. 若 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是域 $\mathbb{K}$ 上的 n 维向量空间 V 的一组基，$f_1,\ldots,f_m \in V^*$，则 $\dim_{\mathbb{K}}(\mathrm{span}\,\{f_1,\ldots,f_m\}) = \mathrm{rank}(B)$，其中矩阵 $B = \begin{pmatrix} f_1(\mathbf{e}_1) & \cdots & f_1(\mathbf{e}_n) \\ f_2(\mathbf{e}_1) & \cdots & f_2(\mathbf{e}_n) \\ \vdots & & \vdots \\ f_m(\mathbf{e}_1) & \cdots & f_m(\mathbf{e}_n) \end{pmatrix}$。

证明. 由 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的基可知 $\varepsilon_{\mathbf{e}_1},\ldots,\varepsilon_{\mathbf{e}_n}$ 是 V^{**} 的基，则由定理1.5.3(3) 可知

$$\begin{aligned}\dim_{\mathbb{K}}(\mathrm{span}\,\{f_1,\ldots,f_m\}) &= \mathrm{rank}((\varepsilon_{\mathbf{e}_i}(f_j))_{n\times m}) \\ &= \mathrm{rank}((f_j(\mathbf{e}_i))_{n\times m}) \\ &= \mathrm{rank}((f_i(\mathbf{e}_j))_{m\times n})\end{aligned}$$

即得结论。 □

下面我们考虑子空间的对偶。

定义 1.5.4. 设 U 是域 $\mathbb{K}$ 上向量空间 V 的子空间，我们定义 U 的零化子集为

$$U^0 = \{f \in V^* \mid \forall \mathbf{x} \in U,\ f(\mathbf{x}) = 0\}.$$

定理 1.5.4. U^0 是 V^* 的子空间，并且 $\dim(U^0) = \dim(V) - \dim(U)$。

证明. 首先我们证明 U^0 是子空间。任取 $\alpha,\beta \in \mathbb{K}$ 及 $f,g \in U^0$，我们需要验证 $\alpha f + \beta g \in U^0$，即我们只需证任取 $\mathbf{x} \in U$ 都有 $(\alpha f + \beta g)(\mathbf{x}) = 0$ 即可。注意到

$$(\alpha f + \beta g)(\mathbf{x}) = \alpha f(\mathbf{x}) + \beta g(\mathbf{x}) = 0 + 0 = 0.$$

即得结论。

下面我们计算 $\dim(U^0)$。设 $\dim(V) = n$，$\dim(U) = k$，并且 $\mathbf{e}_1,\ldots,\mathbf{e}_k$ 是 U 的一组基。则由基扩充定理，可以得到 $\mathbf{e}_{k+1},\ldots,\mathbf{e}_n \in V$ 使得 $\mathbf{e}_1,\ldots,\mathbf{e}_k,\mathbf{e}_{k+1},\ldots,\mathbf{e}_n$ 是 V 的基。取其对偶基 $\mathbf{e}^1,\ldots,\mathbf{e}^k,\mathbf{e}^{k+1},\ldots,\mathbf{e}^n$，我们验证 $\mathrm{span}\,\{\mathbf{e}^{k+1},\ldots,\mathbf{e}^n\} = U^0$，从而 $\dim(U^0) = n - k$。

首先，按对偶基的定义，$\mathbf{e}^{k+1},\ldots,\mathbf{e}^n$ 作用到 $\mathbf{e}_1,\ldots,\mathbf{e}_k$ 上是 0，即 $\mathbf{e}^{k+1},\ldots,\mathbf{e}^n \in U^0$，所以 $\mathrm{span}\,\{\mathbf{e}^{k+1},\ldots,\mathbf{e}^n\} \subset U^0$。其次，任取 $f \in U^0$，由 $\mathbf{e}^1,\ldots,\mathbf{e}^n$ 是 V^* 的基可知 $f = \sum_{i=1}^n \alpha_i\mathbf{e}^i$。

由于对任意的 $\mathbf{e}_j \in U$, $j \in \{1,\dots,k\}$，都有 $f(\mathbf{e}_j) = \alpha_j = 0$，即 $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$，所以 $f \in \operatorname{span}\{\mathbf{e}^{k+1},\dots,\mathbf{e}^n\}$，即 $\operatorname{span}\{\mathbf{e}^{k+1},\dots,\mathbf{e}^n\} \supset U^0$。综上所述，$\operatorname{span}\{\mathbf{e}^{k+1},\dots,\mathbf{e}^n\} = U^0$。这样我们就完成了证明。 □

例 1.5.5. 设 $V = \mathbb{R}^2$，$U = \operatorname{span}\{(2,3)^t\}$，求 U^0 的一组基。

解. 取 $\mathbb{R}^2$ 的标准基 $\mathbf{e}_1 = (1,0)^t$，$\mathbf{e}_2 = (0,1)^t$，则 $\mathbb{R}^2$ 中的元素都可以写成坐标形式 $\mathbf{x} = (\mathbf{e}_1,\mathbf{e}_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$，于是 V^* 中任意元素 f 都能写成 $f(\mathbf{x}) = \alpha_1 x_1 + \alpha_2 x_2$，$\alpha_1, \alpha_2 \in \mathbb{R}$ 的形式。注意到 $\dim(U) = 1$, 故 $\dim(U^0) = 1$, 故只需找到 U^0 中的一个非零线性函数即可。由于 $f((2,3)^t) = 0$，即 $2\alpha_1 + 3\alpha_2 = 0$，故可取 $f(\mathbf{x}) = -3x_1 + 2x_2 \in U^0$，即 $U^0 = \operatorname{span}\{-3x_1 + 2x_2\}$。 □

定理 1.5.5. *设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，U 是 V 的子空间，则 $(U^0)^0 \simeq U$，同构映射是自然映射。*

利用定理1.5.2及定理1.5.4立刻可证。

在本节的最后，我们讨论一下抽象的齐次线性方程组。设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$f_1,\dots,f_m \in V^*$，则我们称方程组

$$\begin{cases} f_1(\mathbf{x}) = 0 \\ f_2(\mathbf{x}) = 0 \\ \quad\vdots \\ f_m(\mathbf{x}) = 0 \end{cases} \text{，其中}\mathbf{x} \in V. \tag{H}$$

是抽象的齐次线性方程组，称子空间 $U = \{\mathbf{x} \in V \mid \forall i = 1,\dots,m,\ f_i(\mathbf{x}) = 0\}$(验证这是 V 的一个子空间，留作练习) 是该方程组的解空间。

显然，如果我们取 $V = \mathbb{R}^n$，取基为 $\mathbb{R}^n$ 的标准基，将抽象的线性函数写成坐标形式，就可以看到这里的方程组和解空间的定义回到了上册中我们熟知的形式。对抽象的情形，我们有下面的定理：

定理 1.5.6. *(1) 令 $r = \dim(\operatorname{span}\{f_1,\dots,f_m\})$，则方程组 (H) 的解空间 U 的维数为 $n - r$。*
(2)V 的任意子空间 U 都是某个齐次线性方程组的解空间。

证明. (1) 不妨设 $f_1,\dots,f_r$ 在 V^* 中线性无关，则由基扩充定理，存在一组 $\varphi_1,\dots,\varphi_{n-r} \in V^*$ 使得 $f_1,\dots,f_r,\varphi_1,\dots,\varphi_{n-r}$ 是 V^* 的一组基。由定理1.5.3(1)，存在 V 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 的对偶基恰为 $f_1,\dots,f_r,\varphi_1,\dots,\varphi_{n-r}$。那么，显然有 $\operatorname{span}\{\mathbf{e}_{r+1},\dots,\mathbf{e}_n\}$ 是方程组 (H) 的解空间，即 $\dim(U) = n - r$。
(2) 取 U 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_r$，由基扩充定理，存在 $\mathbf{e}_{r+1},\dots,\mathbf{e}_n \in V$ 使得 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基。取这组基的对偶基 $\mathbf{e}^1,\dots,\mathbf{e}^n \in V^*$，则显然方程组

$$\begin{cases} \mathbf{e}^{r+1}(\mathbf{x}) = 0 \\ \qquad\vdots \\ \mathbf{e}^{n}(\mathbf{x}) = 0 \end{cases} ,\quad \mathbf{x} \in V.$$

的解空间恰为 U。 □

1.6 双线性型和二次型

1.6.1 多重线性映射简介

我们在 1.3 节的最后 (定义1.3.7) 已经定义了线性映射，下面我们给出多重线性映射的定义。

定义 1.6.1. 设 $V_1,\dots,V_m,W$ 是域 $\mathbb{K}$ 上的向量空间，映射 $f:V_1\times\cdots\times V_m\to W$，如果对任意的 $i\in\{1,\dots,m\}$，固定除了 $\mathbf{x}_i$ 之外其余的变元，任取 $\alpha,\beta\in\mathbb{K}$ 及 $\mathbf{u},\mathbf{v}\in V_i$，都有

$$f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)=\alpha f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{u},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)+\beta f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{v},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)$$

则称 f 是 m 重线性映射 (m-multilinear map)。即如果 f 在任意固定 $m-1$ 个变元后是关于剩下那个变元的线性映射，那么 f 就是 m 重线性映射。

例 1.6.1. 设 $V_1=\cdots=V_m=\mathbb{K}$ 是域 $\mathbb{K}$ 上的向量空间，定义 $f:V_1\times\cdots\times V_m\to\mathbb{K}$, $(x_1,\dots,x_m)\mapsto x_1x_2\cdots x_m$，则容易看出 f 是 m 重线性映射。

例 1.6.2. 在上册第三章中，我们已经知道 $\mathbb{R}$ 上的 n 阶行列式是 n 重线性映射。我们也可以将行列式的定义推广到一般的域上。设 $\mathbb{K}$ 是任意的一个域，则容易验证行列式

$$\det: M_n(\mathbb{K})\to\mathbb{K},\ A=(a_{ij})_{n\times n}\mapsto\sum_{\sigma\in S_n}\varepsilon_\sigma a_{\sigma(1)1}\cdots a_{\sigma(n)n}$$

仍是关于矩阵列 (或行) 的 n 重线性函数。

例 1.6.3. 设 V 是域 $\mathbb{K}$ 上的向量空间，令 $\mathcal{A}:V^*\times V\to\mathbb{K}$, $(f,\mathbf{x})\mapsto f(\mathbf{x})$，则 $\mathcal{A}$ 是 2 重线性函数。验证如下：

- 任取 $\alpha,\beta\in\mathbb{K}$ 及 $f,g\in V^*$，有

$$\mathcal{A}(\alpha f+\beta g,\mathbf{x})=\alpha f(\mathbf{x})+\beta g(\mathbf{x})=\alpha\mathcal{A}(f,\mathbf{x})+\beta\mathcal{A}(g,\mathbf{x}).$$

- 任取 $\alpha,\beta\in\mathbb{K}$ 及 $\mathbf{x},\mathbf{y}\in V$，则

$$\mathcal{A}(f,\alpha\mathbf{x}+\beta\mathbf{y})=\alpha f(\mathbf{x})+\beta f(\mathbf{y})=\alpha\mathcal{A}(f,\mathbf{x})+\beta\mathcal{A}(f,\mathbf{y}).$$

故 $\mathcal{A}$ 是 2 重线性映射。

特别地，令 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，$\mathbf{e}^1,\dots,\mathbf{e}^n$ 是其对偶基，则有

$$f=\sum_{i=1}^n\alpha_i\mathbf{e}^i,\ \mathbf{x}=\sum_{i=1}^n\beta_i\mathbf{e}_i.$$

则 $\mathcal{A}(f,\mathbf{x})=f(\mathbf{x})=\sum_{i=1}^n\alpha_i\beta_i$。

在讲义的最后一章中我们会看到，每个 2 重线性函数 $\mathcal{A}:V^*\times V\to\mathbb{K}$ 与 V 上的一个线性算子是一一对应的，具体的分析我们留到最后一章。

关于多重线性映射的更多性质，我们会在最后一章张量中统一地讨论。张量是一种灵巧的结构，它统一了我们所学的向量、矩阵、双线性型和线性算子等代数结构并加以拓展，在理论和应用中都有重要的作用。

1.6.2 双线性型

定义 1.6.2. 设 V 是域 $\mathbb{K}$ 上的向量空间，$f: V\times V\to\mathbb{K}$ 是 2 重线性映射，则我们称 f 是 V 上的一个双线性型 (bilinear form)。

换言之，双线性型 $f(\mathbf{x},\mathbf{y})$ 是一个 $V\times V\to\mathbb{K}$ 的函数，它满足 $f(\alpha\mathbf{x}_1+\beta\mathbf{x}_2,\mathbf{y})=\alpha f(\mathbf{x}_1,\mathbf{y})+\beta f(\mathbf{x}_2,\mathbf{y})$，$f(\mathbf{x},\alpha\mathbf{y}_1+\beta\mathbf{y}_2)=\alpha f(\mathbf{x},\mathbf{y}_1)+\beta f(\mathbf{x},\mathbf{y}_2)$。

我们可以把 V 上所有的双线性型放在一起做成一个集合 $L_2(V,\mathbb{K})$，在 $L_2(V,\mathbb{K})$ 上定义如下的加法和数乘：

$$\begin{aligned}
+ &: L_2(V,\mathbb{K})\times L_2(V,\mathbb{K})\to L_2(V,\mathbb{K}),\quad (f,g)\mapsto(f+g):\ (f+g)(\mathbf{x},\mathbf{y})=f(\mathbf{x},\mathbf{y})+g(\mathbf{x},\mathbf{y})\\
\cdot &: \mathbb{K}\times L_2(V,\mathbb{K})\to L_2(V,\mathbb{K}),\quad (\lambda,f)\mapsto(\lambda f):\ (\lambda f)(\mathbf{x},\mathbf{y})=\lambda f(\mathbf{x},\mathbf{y})
\end{aligned}$$

则 $L_2(V,\mathbb{K})$ 在上述加法和数乘下构成了 $\mathbb{K}$ 上的向量空间。下面我们考虑这个向量空间的结构。

定理 1.6.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基，则映射

$$\begin{aligned}
\varPhi: L_2(V,\mathbb{K}) &\longmapsto M_n(\mathbb{K})\\
f &\longmapsto (f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}
\end{aligned}$$

是线性双射，即 $L_2(V,\mathbb{K})$ 与 $M_n(\mathbb{K})$ 同构。

证明. 首先，任取 $\mathbf{x},\mathbf{y}\in V$，我们考虑 $f(\mathbf{x},\mathbf{y})$ 的值和矩阵 $\varPhi(f)=(f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}$ 的关系。不妨设 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{j=1}^n y_j\mathbf{e}_j$，则

$$\begin{aligned}
f(\mathbf{x},\mathbf{y}) &= f(\sum_{i=1}^n x_i\mathbf{e}_i,\sum_{j=1}^n y_j\mathbf{e}_j)\\
&= \sum_{i=1}^n x_i f(\mathbf{e}_i,\sum_{j=1}^n y_j\mathbf{e}_j)\\
&= \sum_{i=1}^n x_i\sum_{j=1}^n y_j f(\mathbf{e}_i,\mathbf{e}_j)\\
&= \sum_{i,j=1}^n x_i y_j f(\mathbf{e}_i,\mathbf{e}_j) \qquad (1.6.1)\\
&= (x_1,\ldots,x_n)\begin{pmatrix} f(\mathbf{e}_1,\mathbf{e}_1) & \cdots & f(\mathbf{e}_1,\mathbf{e}_n)\\ \vdots & & \vdots\\ f(\mathbf{e}_n,\mathbf{e}_1) & \cdots & f(\mathbf{e}_n,\mathbf{e}_n)\end{pmatrix}\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}\\
&= (x_1,\ldots,x_n)\varPhi(f)\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}
\end{aligned}$$

由于向量在给定基底后与坐标是一一对应的，故由上面的推导可知，双线性型 f 可以视作坐标的函数，$\varPhi(f)$ 就是这个函数的“参数”(或“系数”)。

现在我们可以证明 $\varPhi$ 是同构了。首先 $\varPhi$ 是线性映射。任取 $f,g\in L_2(V,\mathbb{K})$ 及 $\lambda,\mu\in\mathbb{K}$，

则对矩阵的每个位置，都有

$$(\lambda f+\mu g)(\mathbf{e}_i,\mathbf{e}_j)=\lambda f(\mathbf{e}_i,\mathbf{e}_j)+\mu g(\mathbf{e}_i,\mathbf{e}_j)$$

故 $\Phi(\lambda f+\mu g)=\lambda\Phi(f)+\mu\Phi(g)$。

其次，Φ 是单射。只需证 $\ker(\Phi)=\{0\}$。如果 $\Phi(f)=O_{n\times n}$，那么任取 $\mathbf{x},\mathbf{y}\in V$，由推导 (1.6.1) 可得 $f(\mathbf{x},\mathbf{y})=0$，即 f 是零映射。

最后，f 是满射。任取矩阵 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{K})$，则可作双线性型

$$f(\sum_{i=1}^n x_i\mathbf{e}_i,\sum_{j=1}^n y_j\mathbf{e}_j)=(x_1,\dots,x_n)\cdot A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

满足 $\Phi(f)=A$。

综上所述，Φ 是同构。□

我们称 $\Phi(f)$ 为双线性型 f 在基底 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵。这个定理告诉我们，选定一组基后双线性型与矩阵一一对应，并且，设 $\mathbf{x},\mathbf{y}$ 在给定基底下的坐标分别为 $(x_1,\dots,x_n)^t$ 和 $(y_1,\dots,y_n)^t$，$\Phi(f)=(a_{ij})_{n\times n}$，则 f 可以写成关于坐标的函数 $f(\mathbf{x},\mathbf{y})=\sum_{i,j=1}^n a_{ij}x_iy_j$。那么，接下来一个自然的问题就是，在基变换下双线性型对应的矩阵如何变换呢？这就是下面的定理。

定理 1.6.2. *设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\dots,\mathbf{e}'_n$ 是 V 的两组基，并且这两组基之间有基变换 $(\mathbf{e}'_1,\dots,\mathbf{e}'_n)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot A$。若 f 在这两组基下的矩阵分别为 $\Phi(f)$ 和 $\Phi'(f)$，则 $\Phi'(f)=A^t\Phi(f)A$。*

证明. 任取 $\mathbf{x},\mathbf{y}\in V$，设 $\mathbf{x}$ 在基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\dots,\mathbf{e}'_n$ 下的坐标分别为 $(x_1,\dots,x_n)^t$ 和 $(x'_1,\dots,x'_n)^t$，$\mathbf{y}$ 在基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\dots,\mathbf{e}'_n$ 下的坐标分别为 $(y_1,\dots,y_n)^t$ 和 $(y'_1,\dots,y'_n)^t$，则由推导 (1.6.1)，有

$$f(\mathbf{x},\mathbf{y})=(x_1,\dots,x_n)\Phi(f)\begin{pmatrix}y_1\\ \dots\\ y_n\end{pmatrix}=(x'_1,\dots,x'_n)\Phi'(f)\begin{pmatrix}y'_1\\ \dots\\ y'_n\end{pmatrix}\tag{1.6.2}$$

由坐标变换公式有

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=A\cdot\begin{pmatrix}x'_1\\ \vdots\\ x'_n\end{pmatrix},\qquad\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=A\cdot\begin{pmatrix}y'_1\\ \vdots\\ y'_n\end{pmatrix}$$

即

$$(x_1,\dots,x_n)=(x'_1\dots,x'_n)\cdot A^t,\qquad\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=A\cdot\begin{pmatrix}y'_1\\ \vdots\\ y'_n\end{pmatrix}$$

将其代入式 (1.6.2) 得

$$(x'_1,\ldots,x'_n)A^t\Phi(f)A\begin{pmatrix}y'_1\\ \ldots\\ y'_n\end{pmatrix} = (x'_1,\ldots,x'_n)\Phi'(f)\begin{pmatrix}y'_1\\ \ldots\\ y'_n\end{pmatrix}$$

上式对任意 $(x'_1,\ldots,x'_n)$ 和 $(y'_1,\ldots,y'_n)$ 都成立，于是我们让 $(x'_1,\ldots,x'_n)$ 和 $(y'_1,\ldots,y'_n)$ 分别取遍 $(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,\ldots,0,1)$ 即可知 $\Phi'(f)$ 和 $A^t\Phi(f)A$ 每个位置的元素都相等，所以 $\Phi'(f)=A^t\Phi(f)A$。 □

同一个双线性型在两组不同的基底下对应了两个矩阵，我们称这两个矩阵是合同等价的，即下面的定义。

定义 1.6.3. 设 $\mathbb{K}$ 是域，$A,B\in M_n(\mathbb{K})$ 并且存在可逆矩阵 $P\in M_n(\mathbb{K})$ 使得 $B=P^tAP$，则我们称 A 与 B 是合同等价 (congruent) 的，简称 A 和 B 合同。容易验证合同等价是一个等价关系 (留作练习)。此外，设 $P\in M_n(\mathbb{K})$ 是可逆矩阵，则我们称 $M_n(\mathbb{K})\to M_n(\mathbb{K})$, $A\mapsto P^tAP$ 为由 P 定义的合同变换。

显然若两个矩阵合同等价，则它们的秩相等。于是我们有以下定义。

定义 1.6.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基，f 是 V 上的双线性型，且 f 在这组基下的矩阵为 A。如果 $\mathrm{rank}(A)=r$，则我们称 f 的秩 (rank) 是 r，记为 $\mathrm{rank}(f)=r$。如果 $r=n$，则称 f 是非退化的双线性型，否则称为退化的双线性型。由矩阵合同的定义可知，双线性型 f 的秩是 f 的固有性质 (不变量)，与基底的选择无关。

既然双线性型是函数，那我们当然要研究其“零点”。我们有以下定义。

定义 1.6.5. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，f 是 V 上的双线性型。我们称

$$L_f=\{\mathbf{x}\in V\mid \forall\mathbf{y}\in V,\ f(\mathbf{x},\mathbf{y})=0\}$$

为 f 的左核 (或左根, left kernel)。容易验证 L_f 是 V 的子空间, 并且 $\dim(L_f)=n-\mathrm{rank}(f)$(需利用上学期的对偶定理)。设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基，f 在这组基下的矩阵为 A，由定义我们也可以将左核写成

$$L_f=\{\mathbf{x}\in V\mid f(\mathbf{x},\mathbf{e}_1)=\cdots=f(\mathbf{x},\mathbf{e}_n)=0\}=\left\{(\mathbf{e}_1,\ldots,\mathbf{e}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\mid (x_1,\ldots,x_n)^t\in\mathbb{K}^n\text{且}(x_1,\ldots,x_n)\cdot A=\mathbf{0}\right\}.$$

此外，我们也可以对称地定义 f 的右核 $R_f=\{\mathbf{y}\in V\mid\forall\mathbf{x}\in V,\ f(\mathbf{x},\mathbf{y})=0\}$，很容易看到 R_f 也是 V 的子空间并且 $\dim(R_f)=\dim(L_f)=n-\mathrm{rank}(f)$。

在本小节的最后，我们来看一个具体计算的例子。

例 1.6.4. 在 $\mathbb{R}^3$ 上取标准基 $\mathbf{e}_1=(1,0,0)^t$，$\mathbf{e}_2=(0,1,0)^t,\mathbf{e}_3=(0,0,1)^t$，任取 $\mathbf{x}=\sum_{i=1}^3 x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{j=1}^3 y_j\mathbf{e}_j$，定义 $f(\mathbf{x},\mathbf{y})=x_1y_2+x_1y_3+x_2y_1+x_2y_3+x_3y_1+x_3y_2$，显然 f 是 $\mathbb{R}^3$ 上的双线性型，如果我们另取一组基 $\mathbf{e}'_1=(1,1,0)^t$，$\mathbf{e}'_2=(1,-1,0)^t$，$\mathbf{e}'_3=(-1,-1,1)^t$，则两组基之间有基变换

$(\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3') = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \cdot P$ ，其中 $P = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$。$f$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下可以写成

$$f(\mathbf{x}, \mathbf{y}) = (x_1, x_2, x_3) \underbrace{\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{A} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

在基底 $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$ 下，设 $\mathbf{x} = \sum_{i=1}^{3} x_i' \mathbf{e}_i'$，$\mathbf{y} = \sum_{j=1}^{3} y_j' \mathbf{e}_j'$，$f$ 在基底 $\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'$ 下的矩阵为 B，则由定理1.6.2，有

$$B = P^t A P = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}^t \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

即

$$f(\mathbf{x}, \mathbf{y}) = 2x_1'y_1' - 2x_2'y_2' - 2x_3'y_3'$$

此外，容易求出 $\operatorname{rank}(f) = \operatorname{rank}(A) = 3$，因此 $\dim(L_f) = 0$，即 f 的左核为 $\{\mathbf{0}\}$。

1.6.3 对称双线性型

接下来我们的问题是，如何选取一组合适的基底使得双线性型在这组基下的矩阵较为简单，最好是对角形。这个问题对于一般的双线性型是不成立的，因为与对角形矩阵合同等价的矩阵必然是对称矩阵。因此，我们首先要考虑对称矩阵所对应的双线性型。

定义 1.6.6. 设 V 是域 $\mathbb{K}$ 上的向量空间，$f \in L_2(V, \mathbb{K})$，若任取 $\mathbf{x}, \mathbf{y} \in V$，都有 $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$，则称 f 是对称双线性型 (symmtric bilinear form)；如果对任意 $\mathbf{x}, \mathbf{y} \in V$，都有 $f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$，则称 f 是斜对称双线性型 (antisymmtric bilinear form)。

斜对称双线性型的性质我们会在下下节讨论。首先我们有以下命题。

命题 1.6.1. f 是对称双线性型 $\iff$ $\Phi(f)$ 是对称矩阵；f 是斜对称双线性型 $\iff$ $\Phi(f)$ 是斜对称矩阵。

证明. 任取 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，则 f 是对称双线性型 $\iff$ $\forall \mathbf{e}_i, \mathbf{e}_j$，有 $f(\mathbf{e}_i, \mathbf{e}_j) = f(\mathbf{e}_j, \mathbf{e}_i)$ $\iff$ $\Phi(f) = \Phi(f)^t$；f 是斜对称双线性型 $\iff$ $\forall \mathbf{e}_i, \mathbf{e}_j$，有 $f(\mathbf{e}_i, \mathbf{e}_j) = -f(\mathbf{e}_j, \mathbf{e}_i)$ $\iff$ $\Phi(f) = -\Phi(f)^t$。 □

定义 1.6.7. 设 V 是域 $\mathbb{K}$ 上的向量空间，我们记 V 上所有对称双线性型的集合为 $L_2^+(V, \mathbb{K})$，所有斜对称双线性型的集合为 $L_2^-(V, \mathbb{K})$，容易验证 $L_2^+(V, \mathbb{K})$ 和 $L_2^-(V, \mathbb{K})$ 都是 $L_2(V, \mathbb{K})$ 的子空间。

定理 1.6.3. 如果域 $\mathbb{K}$ 满足 $\operatorname{char}(\mathbb{K}) \neq 2$，则 $L_2(V, \mathbb{K}) = L_2^+(V, \mathbb{K}) \oplus L_2^-(V, \mathbb{K})$。

证明. 我们可以利用例1.2.2直接得到结论，下面给出一个抽象层面的证明。首先，任取 $f \in L_2(V, \mathbb{K})$，令

$$g(\mathbf{x}, \mathbf{y}) = \frac{f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x})}{2}, \quad h(\mathbf{x}, \mathbf{y}) = \frac{f(\mathbf{x}, \mathbf{y}) - f(\mathbf{y}, \mathbf{x})}{2}.$$

则容易验证 $g \in L_2^+(V,\mathbb{K})$，$h \in L_2^-(V,\mathbb{K})$，即 $L_2(V,\mathbb{K}) = L_2^+(V,\mathbb{K}) + L_2^-(V,\mathbb{K})$。

下证 $L_2^+(V,\mathbb{K}) \cap L_2^-(V,\mathbb{K}) = \{0\}$。如果 $f \in L_2^+(V,\mathbb{K}) \cap L_2^-(V,\mathbb{K})$，则任取 $\mathbf{x},\mathbf{y} \in V$，我们有 $f(\mathbf{y},\mathbf{x}) = f(\mathbf{x},\mathbf{y}) = -f(\mathbf{x},\mathbf{y})$，故 $2f(\mathbf{x},\mathbf{y}) = 0$，由于 $\mathrm{char}(\mathbb{K}) \neq 2$，故只能是 $f(\mathbf{x},\mathbf{y}) = 0$ 对 $\forall \mathbf{x},\mathbf{y} \in V$ 成立，即 $f = 0$。

综上所述，$L_2(V,\mathbb{K}) = L_2^+(V,\mathbb{K}) \oplus L_2^-(V,\mathbb{K})$。 □

下面我们考虑对称双线性型的标准型 (典范型)。

定义 1.6.8. 设 V 是域 $\mathbb{K}$ 上的向量空间，$f \in L_2^+(V,\mathbb{K})$，如果 f 在某组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下的矩阵是对角形，则称 f 在这组基下的坐标表达式为 f 的标准型或典范型，相应的这组基称为典范基。即 f 在典范基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下一定有

$$f(\mathbf{x},\mathbf{y}) = (x_1,\ldots,x_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots \\ y_n\end{pmatrix}, \quad \text{其中 } \mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i,\ \mathbf{y} = \sum_{i=1}^n y_i\mathbf{e}_i.$$

的形式。

下面的定理告诉我们对称双线性型一定可以化成标准型。

定理 1.6.4. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathrm{char}(\mathbb{K}) \neq 2$，$f \in L_2^+(V,\mathbb{K})$，则存在 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得 f 在这组基下为标准型，即任取 $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y} = \sum_{i=1}^n y_i\mathbf{e}_i \in V$，有

$$f(\mathbf{x},\mathbf{y}) = (x_1,\ldots,x_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots \\ y_n\end{pmatrix} = \sum_{i=1}^n \lambda_i x_i y_i.$$

证明. 以下证明是 Lagrange 给出的。

当 $f = 0$ 时定理显然成立。下设 $f \neq 0$，则存在 $\mathbf{x}_0,\mathbf{y}_0 \in V$ 使得 $f(\mathbf{x}_0,\mathbf{y}_0) \neq 0$，利用 2 重线性性质容易验证

$$f(\mathbf{x}_0,\mathbf{y}_0) = \frac{1}{2}(f(\mathbf{x}_0+\mathbf{y}_0,\mathbf{x}_0+\mathbf{y}_0) - f(\mathbf{x}_0,\mathbf{x}_0) - f(\mathbf{y}_0,\mathbf{y}_0))$$

于是 $f(\mathbf{x}_0+\mathbf{y}_0,\mathbf{x}_0+\mathbf{y}_0)$，$f(\mathbf{x}_0,\mathbf{x}_0)$，$f(\mathbf{y}_0,\mathbf{y}_0)$ 三者中至少有一个不为 0(否则与 $f(\mathbf{x}_0,\mathbf{y}_0) \neq 0$ 矛盾！)。即 $f \neq 0$ 可以推出 $\exists \mathbf{e}_1 \in V$ 使得 $f(\mathbf{e}_1,\mathbf{e}_1) \neq 0$。

对 $\dim(V)$ 归纳。$\dim(V) = 1$ 时任取 $\mathbf{e}_1 \neq 0 \in V$，都有 $\mathbf{e}_1$ 是 V 的一组基，此时若 $\mathbf{x} = x_1\mathbf{e}_1$，$\mathbf{y} = y_1\mathbf{e}_1$，则必有 $f(\mathbf{x},\mathbf{y}) = f(\mathbf{e}_1,\mathbf{e}_1)\cdot \mathbf{x}_1\mathbf{y}_1$ 是标准型。

下设定理对所有维数小于 n 的向量空间都成立，我们来证明 $\dim(V) = n$ 的情形。

由证明第一段的论述可知，$\exists \mathbf{e}_1 \in V$ 使得 $f(\mathbf{e}_1,\mathbf{e}_1) = \lambda_1 \neq 0$，我们令

$$U = \{\mathbf{u} \in V \mid f(\mathbf{u},\mathbf{e}_1) = 0\}$$

则容易验证：U 是线性函数 $\varphi : V \to \mathbb{K}$, $\mathbf{x} \mapsto f(\mathbf{x},\mathbf{e}_1)$ 的核，即 $U = \ker(\varphi)$。我们首先证明：$\dim(U) = n-1$。显然 φ 是满射，故由定理1.4.2，我们有 $V/\ker(\varphi) \simeq \mathbb{K}$，故 $\dim(V)-\dim(U) = 1$，即 $\dim(U) = n-1$。

其次，$V = U \oplus \mathrm{span}\,\{\mathbf{e}_1\}$。任取 $\mathbf{v} \in U \cap \mathrm{span}\,\{\mathbf{e}_1\}$，则一方面 $\mathbf{v} = \lambda\mathbf{e}_1$，$\lambda \in \mathbb{K}$，另一方面 $\varphi(\mathbf{v}) = \lambda f(\mathbf{e}_1,\mathbf{e}_1) = 0$，而 $f(\mathbf{e}_1,\mathbf{e}_1) \neq 0$，故只有 $\lambda = 0$，即 $\mathbf{v} = \mathbf{0}$，即 $U + \mathrm{span}\,\{\mathbf{e}_1\}$ 是直和。于

是由推论 1.3.1，$\dim(U\oplus \operatorname{span}\{\mathbf{e}_1\}) = n = \dim(V)$，即 $V = U\oplus \operatorname{span}\{\mathbf{e}_1\}$。

现在我们可以使用归纳假设了。令 $g = f\big|_U$，即

$$g\ :\ U\times U \longrightarrow \mathbb{K},\ (\mathbf{x}',\mathbf{y}')\longmapsto f(\mathbf{x}',\mathbf{y}').$$

显然 g 是 U 上的对称双线性型，由于 $\dim(U) = n-1$，按归纳假设，存在 U 的一组基 $\mathbf{e}_2,\dots,\mathbf{e}_n$，任取 $\mathbf{x}' = \sum_{i=2}^n x_i\mathbf{e}_i$，$\mathbf{y}' = \sum_{i=2}^n y_i\mathbf{e}_i$，有

$$g(\mathbf{x}',\mathbf{y}') = \sum_{i=2}^n \lambda_i x_i y_i$$

即 g 在基底 $\mathbf{e}_2,\dots,\mathbf{e}_n$ 下的矩阵是对角形。由于 $V = U\oplus \operatorname{span}\{\mathbf{e}_1\}$，故 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，注意到 $U = \ker(\varphi)$，所以任取 $i\in\{2,\dots,n\}$，都有 $f(\mathbf{e}_1,\mathbf{e}_i) = f(\mathbf{e}_i,\mathbf{e}_1) = 0$。于是利用 2 重线性性质及刚刚得到的结论，任取 $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y} = \sum_{i=1}^n y_i\mathbf{e}_i\in V$，有

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &= x_1y_1f(\mathbf{e}_1,\mathbf{e}_1) + x_1\sum_{i=2}^n y_i f(\mathbf{e}_1,\mathbf{e}_i) + y_1\sum_{i=2}^n x_i f(\mathbf{e}_i,\mathbf{e}_1) + f(\sum_{i=2}^n x_i\mathbf{e}_i,\sum_{i=2}^n y_i\mathbf{e}_i)\\ &= \lambda_1x_1y_1 + 0 + 0 + \sum_{i=2}^n x_iy_i\\ &= \sum_{i=1}^n \lambda_i x_i y_i. \end{aligned}$$

即定理对 n 维向量空间也成立。由数学归纳法可知定理总成立。 □

由于双线性型空间与矩阵空间同构，因此我们也可以将定理1.6.4写成矩阵形式如下：

推论 1.6.1. 域 $\mathbb{K}$ 上的每个 n 阶对称矩阵都合同等价于一个 n 阶对角矩阵。

证明是显然的。这个推论也可以绕开双线性型而直接从矩阵的初等变换角度证明，我们留给读者思考[1]。

1.6.4 对称双线性型标准型的计算

现在我们考虑如何将一个具体的对称双线性型化成标准型。

归纳法

实际上，定理1.6.4的证明过程已经给我们提供了计算方法，即我们只要归纳地选出一组典范基即可。下面我们通过具体的例子来展示计算过程。

例 1.6.5. 我们仍然使用例1.6.4的双线性型。在 $\mathbb{R}^3$ 上取标准基 $\mathbf{e}_1 = (1,0,0)^t$，$\mathbf{e}_2 = (0,1,0)^t, \mathbf{e}_3 = (0,0,1)^t$，任取 $\mathbf{x} = \sum_{i=1}^3 x_i\mathbf{e}_i$，$\mathbf{y} = \sum_{j=1}^3 y_j\mathbf{e}_j$，定义 $f(\mathbf{x},\mathbf{y}) = x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2$，容易验证 f 是 $\mathbb{R}^3$ 上的对称双线性型，其在标准基下的矩阵为 $A = \begin{pmatrix} 0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$。我们下面将 f 化成标准型。

(i) 取 $\mathbf{v}_1 = (1,1,0)^t$，容易验证 $f(\mathbf{v}_1,\mathbf{v}_1) = 2\neq 0$。

(ii) 求出子空间 $U_1 = \{\mathbf{x}\in\mathbb{R}^3 \mid f(\mathbf{x},\mathbf{v}_1) = 0\}$ 的一组基。过程如下：任取 $\mathbf{x} = (x_1,x_2,x_3)^t\in\mathbb{R}^3$，

---

[1] 或参考《高等代数 · 上》，丘维声，清华大学出版社的 §6.1，定理 1。

要使 $f(\mathbf{x}, \mathbf{v}_1) = 0$，即

$$f(\mathbf{x}, \mathbf{v}_1) = (x_1, x_2, x_3) \cdot A \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = x_1 + x_2 + 2x_3 = 0$$

这是一个齐次线性方程 (组)，其解空间的一组基为 $\mathbf{u}_1 = (1, -1, 0)^t$，$\mathbf{u}_2 = (2, 0, -1)^t$。

(iii) 降维。我们令 $f_1 = f\big|_{U_1}$，即 f_1 应满足：任取 $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$，$\mathbf{y} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 \in U_1$，应该有 $f_1(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y})$。那么，按照定理1.6.1，$f_1$ 在基底 $\mathbf{u}_1, \mathbf{u}_2$ 下的矩阵为

$$\Phi(f_1) = \begin{pmatrix} f_1(\mathbf{u}_1, \mathbf{u}_1) & f_1(\mathbf{u}_1, \mathbf{u}_2) \\ f_1(\mathbf{u}_2, \mathbf{u}_1) & f_1(\mathbf{u}_2, \mathbf{u}_2) \end{pmatrix} = \begin{pmatrix} f(\mathbf{u}_1, \mathbf{u}_1) & f(\mathbf{u}_1, \mathbf{u}_2) \\ f(\mathbf{u}_2, \mathbf{u}_1) & f(\mathbf{u}_2, \mathbf{u}_2) \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ -2 & -4 \end{pmatrix}.$$

(iv) 类似步骤 (i)，取 $\mathbf{v}_2 = \mathbf{u}_1 = (1, -1, 0)^t$ 使得 $f_1(\mathbf{v}_2, \mathbf{v}_2) = f(\mathbf{u}_1, \mathbf{u}_1) = -2 \neq 0$。

(v) 求子空间 $U_2 = \{\mathbf{x} \in U_1 \mid f_1(\mathbf{x}, \mathbf{v}_2) = 0\}$ 的一组基。我们可以在标准基下计算，也可以在基 $\mathbf{u}_1, \mathbf{u}_2$ 下计算，分别展示如下：

- 在标准基下：取 $\mathbf{x} = (x_1, x_2, x_3)^t \in U_2$，则 $\mathbf{x}$ 应该满足

  $$\begin{cases} f(\mathbf{x}, \mathbf{v}_1) = x_1 + x_2 + 2x_3 = 0 \\ f(\mathbf{x}, \mathbf{v}_2) = -x_1 + x_2 = 0 \end{cases}$$

  该齐次方程组的解空间为 $\operatorname{span}\{(-1, -1, 1)^t\}$。

- 在基 $\mathbf{u}_1, \mathbf{u}_2$ 下：设 $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 \in U_1$，则

  $$f_1(\mathbf{x}, \mathbf{v}_2) = (\alpha_1, \alpha_2)\begin{pmatrix} -2 & -2 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (\text{注意} \mathbf{v}_2 = (\mathbf{u}_1, \mathbf{u}_2)\begin{pmatrix} 1 \\ 0 \end{pmatrix}.)$$

  即 $-2\alpha_1 - 2\alpha_2 = 0$，$\alpha_2 = -\alpha_1$，所以 U_2 的一组基为 $\mathbf{u}_1 - \mathbf{u}_2 = (-1, -1, 1)^t$。

可以看到这两种方法的计算结果是一致的。

(vi) 取 $\mathbf{v}_3 = (-1, -1, 1)^t \in U_2$，令 $f_2 = f\big|_{U_2}$，容易验证 $f(\mathbf{v}_3, \mathbf{v}_3) = -2 \neq 0$，但我们看到 $U_3 = \{\mathbf{x} \in U_2 \mid f_2(\mathbf{x}, \mathbf{v}_3) = 0\} = \{\mathbf{0}\}$，故算法终止。

于是在典范基 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 下，f 对应的矩阵为

$$\begin{pmatrix} f(\mathbf{v}_1, \mathbf{v}_1) & & \\ & f(\mathbf{v}_2, \mathbf{v}_2) & \\ & & f(\mathbf{v}_3, \mathbf{v}_3) \end{pmatrix} = \begin{pmatrix} 2 & & \\ & -2 & \\ & & -2 \end{pmatrix}$$

即 f 有标准型：任取 $\mathbf{x} = \sum_{i=1}^{3} x_i'\mathbf{v}_i$，$\mathbf{y} = \sum_{i=1}^{3} y_i'\mathbf{v}_i \in V$，都有

$$f(\mathbf{x}, \mathbf{y}) = (x_1', x_2', x_3') \cdot \begin{pmatrix} 2 & & \\ & -2 & \\ & & -2 \end{pmatrix} \cdot \begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}.$$

这个例子可以改写成一般的算法 (算法的终止条件是某一步算出 f_k 是零映射或 $U_k = \{0\}$)，留作练习。当然，归纳法的计算比较繁琐，我们更常用的是下面介绍的方法。

合同变换法

我们已经证明了双线性型空间和矩阵空间是同构的，那么，将双线性型 f 化成标准型等价于寻找可逆矩阵 P 使得 $P^t\Phi(f)P$ 是对角形矩阵。我们知道，可逆矩阵是有限个初等矩阵的乘积。而对一个矩阵左乘 (或右乘) 初等矩阵等价于对这个矩阵作初等行 (或列) 变换。于是，我们只要对 f 对应的矩阵作对称的初等行 (列) 变换 (即合同变换)，即可将对称矩阵化成对角形。下面我们看一个具体的例子。

例 1.6.6. 继续讨论例1.6.4，这一次我们采用合同变换的方法。f 在标准基下所对应的矩阵为 $A = \begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$，我们对 A 作如下的对称的初等变换：

$$
\begin{aligned}
A &\xrightarrow{r_1+r_2} \begin{pmatrix}1&1&2\\1&0&1\\1&1&0\end{pmatrix} \xrightarrow{c_1+c_2} \begin{pmatrix}2&1&2\\1&0&1\\2&1&0\end{pmatrix}\\
&\xrightarrow{r_2-\frac{1}{2}r_1} \begin{pmatrix}2&1&2\\0&-\frac{1}{2}&0\\2&1&0\end{pmatrix} \xrightarrow{c_2-\frac{1}{2}c_1} \begin{pmatrix}2&0&2\\0&-\frac{1}{2}&0\\2&0&0\end{pmatrix}\\
&\xrightarrow{r_3-r_1} \begin{pmatrix}2&0&2\\0&-\frac{1}{2}&0\\0&0&-2\end{pmatrix} \xrightarrow{c_3-c_1} \begin{pmatrix}2&0&0\\0&-\frac{1}{2}&0\\0&0&-2\end{pmatrix}
\end{aligned}
$$

即 f 在某组基下对应对角矩阵 $\begin{pmatrix}2&0&0\\0&-\frac{1}{2}&0\\0&0&-2\end{pmatrix}$。实际上我们可以把这组基写出来，因为对 A 作合同变换相当于坐标变换，我们把初等变换对应的矩阵乘起来就可以得到基变换的过渡矩阵。我们通常采用如下的格式求解过渡矩阵 (对比求逆矩阵的格式，讲义上册 2.6-2.7 节)：

$$
\begin{aligned}
(A|E) = \left(\begin{array}{ccc|ccc}0&1&1&1&0&0\\1&0&1&0&1&0\\1&1&0&0&0&1\end{array}\right) &\xrightarrow{\text{左乘}F_{1,2}(1)} \left(\begin{array}{ccc|ccc}1&1&2&1&1&0\\1&0&1&0&1&0\\1&1&0&0&0&1\end{array}\right) \xrightarrow{\text{右乘}F_{2,1}(1)} \left(\begin{array}{ccc|ccc}2&1&2&1&1&0\\1&0&1&0&1&0\\2&1&0&0&0&1\end{array}\right)\\
&\xrightarrow{\text{左乘}F_{2,1}(-\frac{1}{2})} \left(\begin{array}{ccc|ccc}2&1&2&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\2&1&0&0&0&1\end{array}\right) \xrightarrow{\text{右乘}F_{1,2}(-\frac{1}{2})} \left(\begin{array}{ccc|ccc}2&0&2&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\2&0&0&0&0&1\end{array}\right)\\
&\xrightarrow{\text{左乘}F_{3,1}(-1)} \left(\begin{array}{ccc|ccc}2&0&2&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\0&0&-2&-1&-1&1\end{array}\right) \xrightarrow{\text{右乘}F_{1,3}(-1)} \left(\underbrace{\begin{matrix}2&0&0\\0&-\frac{1}{2}&0\\0&0&-2\end{matrix}}_{B} \;\vdots\; \underbrace{\begin{matrix}1&1&0\\-\frac{1}{2}&\frac{1}{2}&0\\-1&-1&1\end{matrix}}_{P^t}\right)
\end{aligned}
$$

即

$$
(P^tAP \mid P^t) = \left(\begin{array}{ccc|ccc}2&0&0&1&1&0\\0&-\frac{1}{2}&0&-\frac{1}{2}&\frac{1}{2}&0\\0&0&-2&-1&-1&1\end{array}\right) = (B \mid P^t),
$$

所以

$$f(\mathbf{x},\mathbf{y}) = (x_1, x_2, x_3) \cdot A \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = (x_1, x_2, x_3)(P^{-1})^t \cdot B \cdot P^{-1} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

即 $\mathbf{y}$ 在新基底 $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$ 下的新坐标 $(y'_1, y'_2, y'_3)^t$ 满足

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix} = P^{-1} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

于是由坐标变换和基变换的关系，我们有

$$(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \cdot P$$

即 $\mathbf{e}'_1 = (1,1,0)^t$，$\mathbf{e}'_2 = (-\frac{1}{2}, \frac{1}{2}, 0)^t$，$\mathbf{e}'_3 = (-1,-1,1)^t$，这就是我们所需要的典范基。

我们可以看到，典范基和标准型是不唯一的，有关于标准型的不变量，我们将在稍后进行介绍。

变量替换法

我们看到，矩阵合同变换法的本质是坐标变换，而坐标变换可以视作变量替换，那么我们可以直接对二次型进行变量替换将二次型化成标准型。

例 1.6.7. 继续讨论例1.6.4。在 $f(\mathbf{x},\mathbf{y}) = x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2$ 中直接令

$$\begin{cases} x_1 = x'_1 + x'_2 - x'_3 \\ x'_2 = x'_1 - x'_2 - x'_3 \\ x'_3 = x_3 \end{cases}, \quad \begin{cases} y_1 = y'_1 + y'_2 - y'_3 \\ y'_2 = y'_1 - y'_2 - y'_3 \\ y'_3 = y_3 \end{cases}$$

则 $f(\mathbf{x},\mathbf{y}) = 2x'_1y'_1 - 2x'_2y'_2 - 2x'_3y'_3$ 即为标准型。

上面的变量替换很难直接看出，我们会在下一小节中给出找到这组变量替换的方法。

雅可比方法

我们有下面的定理。

定义 1.6.9. 设 $\mathbb{K}$ 是域，$A = (a_{ij})_{n\times n} \in M_n(\mathbb{K})$，$m \in \{1, 2, \ldots, n\}$。我们称行列式

$$\Delta_m = \det \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$

为 A 的 m 阶顺序主子式。特别地，我们令 $\Delta_0 = 1$。

定理 1.6.5 (Jacobian). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基，$f \in L_2^+(V, \mathbb{K})$，$A$ 是 f 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下对应的矩阵，如果 A 的各阶顺序主子式都不为 0，则存在 V 的一

组基 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 使得 f 在这组基下对应的矩阵为

$$\begin{pmatrix} \frac{\Delta_0}{\Delta_1} & & & \\ & \frac{\Delta_1}{\Delta_2} & & \\ & & \ddots & \\ & & & \frac{\Delta_{n-1}}{\Delta_n} \end{pmatrix}.$$

或者用矩阵的语言来说，设 A 是可逆的对称矩阵，则 A 一定合同等价于上面的对角矩阵。

这个定理我们将在稍后证明。需要注意的是，雅可比方法并不能给出典范基的取法，因此有些“鸡肋”。然而，它在下面的正定二次型判定中有重要作用。

1.6.5 二次型的定义和标准型

定义 1.6.10. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})\neq 2)$ 上的 n 维向量空间，函数 $q: V\to\mathbb{K}$ 满足如下两条性质：

(1) 对 $\forall \mathbf{v}\in V$，都有 $q(-\mathbf{v})=q(\mathbf{v})$；

(2) 由公式

$$f_q(\mathbf{x},\mathbf{y})=\frac{1}{2}(q(\mathbf{x}+\mathbf{y})-q(\mathbf{x})-q(\mathbf{y})),\quad (\text{极化恒等式}) \tag{1.6.3}$$

决定的的映射 $f_q: V\times V\to\mathbb{K}$ 是对称双线性型。

则我们称 q 是 V 上的二次型 (quadratic form)，并称 f_q 是与 q 配极 (polarization, 极化) 的双线性型。此外，我们将二次型 q 的秩定义成 $\mathrm{rank}(q)=\mathrm{rank}(f_q)$，将 f_q 的左核 (由 f_q 是对称双线性型可知其左核和右核相等) 称为 q 的迷向子空间 (isotropic subspace)，记为 L_q。

注 1.6.1. 二次型 q 的迷向子空间也可以直接定义成 $L_q=\{\mathbf{u}\in V\mid \forall\mathbf{v}\in V,\ q(\mathbf{u}+\mathbf{v})=q(\mathbf{u})+q(\mathbf{v})\}$。很容易验证这两个定义是等价的，留作练习。

我们看到，二次型和对称双线性型之间有紧密的联系。我们也可以从另一个角度定义二次型。

任给 $f\in L_2^+(V,\mathbb{K})$，我们可以定义函数 $q_f: V\to\mathbb{K},\ \mathbf{x}\mapsto f(\mathbf{x},\mathbf{x})$。下面我们验证 q_f 满足定义1.6.10。首先，$q(-\mathbf{x})=f(-\mathbf{x},-\mathbf{x})=f(\mathbf{x},\mathbf{x})=q(\mathbf{x})$。其次，$f_{q_f}(\mathbf{x},\mathbf{y})=\frac{1}{2}(q_f(\mathbf{x}+\mathbf{y})-q_f(\mathbf{x})-q_f(\mathbf{y}))=\frac{1}{2}(f(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})-f(\mathbf{x},\mathbf{x})-f(\mathbf{y},\mathbf{y}))=f(\mathbf{x},\mathbf{y})$ 是对称双线性型。于是这里定义的 q_f 是定义1.6.10里的二次型。

那么我们会问，是不是定义1.6.10里的每个二次型都可以用上面的方式定义呢？答案是肯定的。实际上，设 q 是定义1.6.10里的二次型，则 $q=q_{f_q}$。这是因为在式 (1.6.3) 中令 $\mathbf{y}=-\mathbf{x}$，则

$$f_q(\mathbf{x},-\mathbf{x})=\frac{1}{2}(q(\mathbf{0})-2q(\mathbf{x}))$$

即 $q(\mathbf{x})=f_q(\mathbf{x},\mathbf{x})+\frac{1}{2}q(\mathbf{0})$，又因为 $f(\mathbf{0},\mathbf{0})=f(\mathbf{0},-\mathbf{0})\Longrightarrow f(\mathbf{0},\mathbf{0})=0$(用到了 $\mathrm{char}(\mathbb{K})\neq 2$)，即 $q(\mathbf{0})=0$，所以 $q(\mathbf{x})=f_q(\mathbf{x},\mathbf{x})=q_{f_q}(\mathbf{x})$。

综合以上论述，我们有：

定理 1.6.6. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})\neq 2)$ 上的 n 维向量空间，则 V 上的每个二次型 q 都唯一确定一个对称双线性型 f_q，反之每个对称双线性型 f 都唯一确定二次型 q_f。

二次型		对称双线性型
$q(\mathbf{x})$	$\rightharpoonup$	$f_q(\mathbf{x},\mathbf{y})=\frac{1}{2}(q(\mathbf{x}+\mathbf{y})-q(\mathbf{x})-q(\mathbf{y}))$
$q_f(\mathbf{x})=f(\mathbf{x},\mathbf{x})$	$\leftharpoondown$	$f(\mathbf{x},\mathbf{y})$

我们知道，选定基底后 n 维向量空间 V 上的对称双线性型与 n 阶对称矩阵是一一对应的，因此，设 q 是 V 上的二次型，取 $\mathbf{e}_1,\dots\mathbf{e}_n$ 是 V 的一组基，则我们称 f_q 在这组基下对应的矩阵 $A=(a_{ij})_{n\times n}$ 为二次型 q 在这组基下对应的矩阵。显然，任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i\in V$，我们有

$$q(\mathbf{x})=(x_1,\dots,x_n)\cdot A\cdot\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\sum_{i,j=1}^n a_{ij}x_ix_j=\sum_{i=1}^n a_{ii}x_i^2+2\sum_{1\leqslant i<j\leqslant n}a_{ij}x_ix_j. \tag{1.6.4}$$

即二次型在给定基底下可以视作关于坐标的二次齐次多项式。

在上一节中我们已经详细讨论了对称双线性型的标准型，现在我们将这些内容直接类比到二次型中。

定义 1.6.11. 设 V 是域 $\mathbb{K}$ 上的向量空间，q 是 V 上的二次型，如果 q 在某组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵是对角形，则称 q 在这组基下的坐标表达式为 q 的标准型或典范型，相应的这组基称为典范基。即 q 在典范基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下一定有

$$q(\mathbf{x})=(x_1,\dots,x_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad \text{其中 } \mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i.$$

的形式。

定理 1.6.7. 设 V 是域 $\mathbb{K}$ 上的向量空间，$\mathrm{char}(\mathbb{K})\neq 2$，$q$ 是 V 上的二次型，则存在 V 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得 q 在这组基下为标准型，即任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，有

$$q(\mathbf{x})=(x_1,\dots,x_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\sum_{i=1}^n \lambda_i x_i^2.$$

利用 $q(\mathbf{x})=f_q(\mathbf{x},\mathbf{x})$ 及定理1.6.4即可证明。

下面我们考虑如何将二次型化成标准型。

例 1.6.8. 在 $\mathbb{R}^3$ 上取标准基 $\mathbf{e}_1=(1,0,0)^t$，$\mathbf{e}_2=(0,1,0)^t,\mathbf{e}_3=(0,0,1)^t$，任取 $\mathbf{x}=\sum_{i=1}^3 x_i\mathbf{e}_i$，定义二次型 $q(\mathbf{x})=2x_1x_2+2x_1x_3+2x_2x_3$，则由式 (1.6.4)，易得 $q(\mathbf{x})$ 对应的矩阵为 $A=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$(即例1.6.4的矩阵)，我们欲将 q 化成标准型，即我们需要寻找一组典范基或者将矩阵 A 合同变换到一个对角形矩阵，显然我们可以采用例1.6.5的归纳法或例 1.6.6的合同变换法来解决此问题 (只需在最后将双线性型转到二次型即可)。现在我们讨论变量替换法。由于二次型是关于坐标的二次齐次多项式，故我们可以通过配平方的方法来寻找合适的变量替换。

对于本例来说，由于表达式中没有平方项，因此我们需要通过旋转变换 (我们会在仿射空间这一章中解释这个变换的几何意义) 作出一个平方项。令

$$\begin{cases}x_1=y_1+y_2\\ x_2=y_1-y_2\ ,\\ x_3=y_3\end{cases}$$

则

$$q(\mathbf{x})=2(y_1^2-y_2^2)+2(y_1+y_2)y_3+2(y_1-y_2)y_3=2y_1^2+4y_1y_3-2y_2^2,$$

对 y_1^2 配平方得

$$q(\mathbf{x})=2(y_1+y_3)^2-2y_2^2-2y_3^2,$$

再令

$$\begin{cases} z_1 = y_1 + y_3 \\ z_2 = y_2 \\ z_3 = y_3 \end{cases},$$

即有

$$q(\mathbf{x}) = 2z_1^2 - 2z_2^2 - 2z_3^2.$$

将两步变量替换合在一起，写成矩阵形式就是

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

按坐标变换公式，上式中的矩阵就是基变换的过渡矩阵，这与我们在例1.6.4中所得的的结果一致。

对于一般的二次型，我们也可以用**配方法**求其标准型，具体过程我们不再赘述。此外我们看到，由对称双线性型和二次型的一一对应关系，我们可以先将对称双线性型化成二次型，通过配方将二次型化成标准型后再化回对称双线性型，这就是例1.6.7中变量替换的由来。

1.6.6 复二次型和实二次型

回到我们常用的复数域和实数域上进行讨论。我们已经知道，二次型的标准型不是唯一的，那么，不同的标准型之间有什么联系呢？换言之，能否找到合同等价的矩阵的不变量呢？这就是我们接下来讨论的内容。

首先，由合同变换的定义可知，二次型的秩在不同基底下是相等的，即合同等价的矩阵秩相同。更具体地说，如果二次型 q 的秩是 r，则其标准型必形如 $q(\mathbf{x}) = \sum_{i=1}^{r} \lambda_i x_i^2$，或者说 q 对应的矩阵必定合同于 $\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$，$\lambda_i \neq 0, i \in \{1, \ldots, r\}$。

我们先考虑复数域的情形。设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，q 是 V 上的二次型，其秩为 r。任取 $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$，由定理 1.6.7，存在一组典范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathbb{C}$ 上的二次型在这组基下可表示为 $q(\mathbf{x}) = \sum_{i=1}^{r} \lambda_i x_i^2$，$\lambda_i \in \mathbb{C}$。由于复数域上每个数都可以开平方，故可以作基变换 $\mathbf{e}_i' = \frac{1}{\sqrt{\lambda_i}} \mathbf{e}_i$，$\forall i \in \{1, \ldots, n\}$，则 $x_i = \frac{1}{\sqrt{\lambda_i}} x_i'$，所以 $q(\mathbf{x}) = \sum_{i=1}^{r} \lambda_i (\frac{x_i'}{\sqrt{\lambda_i}})^2 = \sum_{i=1}^{r} (x_i')^2$。用矩阵的语言表述，即每个 $A \in M_n(\mathbb{C})$，$\operatorname{rank}(A) = r$ 都合同等价于 $\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{n\times n}$。这样我们对复二次型的标准型和复矩阵的合同等价矩阵就彻底讨论清楚了。

实数域上的情形要复杂一些。我们同样可设秩为 r 的二次型 $q(\mathbf{x})$ 在典范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下有表达式 $q(\mathbf{x}) = \sum_{i=1}^{s} \lambda_i x_i^2 - \sum_{j=s+1}^{r} \lambda_j x_j^2$，其中 $\lambda_1, \ldots, \lambda_r$ 都是正实数 (即 q 在这组基下对应的矩阵对角元上有 s 个正数和 $r-s$ 个负数)。那么，我们同样可以作基变换 $\mathbf{e}_i' = \frac{1}{\sqrt{\lambda_i}} \mathbf{e}_i$，$\forall i \in \{1, \ldots, n\}$，则 $q(\mathbf{x}) = \sum_{i=1}^{s} (x_i')^2 - \sum_{j=s+1}^{r} (x_j')^2$。用矩阵的语言表述，即每个 $A \in M_n(\mathbb{R})$，$\operatorname{rank}(A) = r$ 都合同等价于 $\begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}_{n\times n}$。然而，这里有一个问题没有讨论清楚，即 s 在不同的典

范基下是否是确定的呢？答案是肯定的，即下面的定理。

定理 1.6.8 (惯性定理). 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，q 是 V 上秩为 r 的二次型，如果 q 在两组典范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ 下分别对应矩阵 $\begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}_{n\times n}$ 和 $\begin{pmatrix} E_t & & \\ & -E_{r-t} & \\ & & O \end{pmatrix}_{n\times n}$，则 $s = t$。

证明. 用反证法，如果 $s \neq t$，不妨设 $t < s$，我们令

$$L = \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_s\},\ L' = \operatorname{span}\{\mathbf{e}'_{t+1}, \ldots, \mathbf{e}'_n\}.$$

则 L 和 L' 都是 V 的子空间，由维数公式我们有

$$\dim(L + L') + \dim(L \cap L') = \dim(L) + \dim(L') = s + n - t > n,$$

而 $\dim(L + L') \leqslant n$，故 $\dim(L \cap L') > 0$，所以 $\exists \mathbf{x} \in L \cap L'$，$\mathbf{x} \neq \mathbf{0}$ 使得

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_s\mathbf{e}_s = x'_{t+1}\mathbf{e}'_{t+1} + \cdots + x'_n\mathbf{e}'_n.$$

分别在这两组基下计算 $q(\mathbf{x})$ 可得

$$q(\mathbf{x}) = \sum_{i=1}^{s} x_i^2 > 0$$

$$q(\mathbf{x}) = \sum_{j=t+1}^{n} (x'_j)^2 \leqslant 0$$

(注意 L 的取法保证了 $q(\mathbf{x}) \neq 0$) 这两者显然矛盾！故 $s = t$。 □

注 1.6.2. 需要注意的是，一个二次型在两组基下的矩阵相同不能推出这两组基相同。例如，在 $\mathbb{R}^4$ 上，取基底为标准基，二次型 $q(\mathbf{x}) = x_1^2 + x_2^2 - x_3^2$ 的矩阵为 $A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \end{pmatrix}$。如果取基底 $\mathbf{e}'_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0, 0)^t$，$\mathbf{e}'_2 = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0, 0)^t$，$\mathbf{e}'_3 = (0, 0, 1, 0)^t$，$\mathbf{e}'_4 = (0, 0, 0, 2)^t$，则 q 在新的基底下矩阵仍为 A(验证留作练习)。学完欧氏空间后，我们会对此有更深入的理解。

我们称上面的 s 为二次型 q(或其矩阵) 的**正惯性指数**，$r - s$ 为**负惯性指数**，$s - (r - s)$ 为**符号差**。惯性定理告诉我们，实对称矩阵的正 (负) 惯性指数和符号差在合同变化下不变，或者说实二次型的正 (负) 惯性指数和符号差是其本身的固有性质。

接下来我们讨论一类更特殊的实二次型。

定义 1.6.12. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，q 是 V 上的非零二次型。如果 $\operatorname{rank}(q) = n$，则我们称 q 是非退化 (non-degenerate) 的二次型。如果 q 满足：$\forall \mathbf{x} \in V$，都有 $q(\mathbf{x}) \geqslant 0$，则称 q 是半正定 (semi-positive) 的二次型；如果半正定的二次型 q 还是非退化的，则称 q 是正定 (positive) 的二次型。相反地，如果 $-q$ 是 (半) 正定的二次型，则我们称 q 是 (半) 负定 ((semi-)negative) 的二次型。如果存在 $\mathbf{x}, \mathbf{y} \in V$ 使得 $q(\mathbf{x}) > 0$, $q(\mathbf{y}) < 0$，则称 q 是不定的二次型。相应地，取定一组基后我们把 q 对应的矩阵称为 (半) 正定/负定/不定的矩阵。

显然我们只需要搞清楚 (半) 正定二次型的等价条件就可以直接推出 (半) 负定的情形，因此我们以下只讨论 (半) 正定的情形。我们将实二次型化成标准型，就很容易得到下面的定理。

定理 1.6.9. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，q 是 V 上的二次型，$\mathrm{rank}(q) = r$，q 的正惯性指数为 s，则

(1) q 是半正定二次型 $\iff r = s$。

(2) q 是正定二次型 $\iff r = s = n$。

对应到矩阵情形，设 $A \in SM_n(\mathbb{R})$，$\mathrm{rank}(A) = r$，A 的正惯性指数为 s，则

(1) A 是半正定矩阵 $\iff r = s. \iff \exists$ 可逆矩阵 P 使得 $P^tAP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$，其中 $\forall i \in \{1, \ldots, n\}$，$\lambda_i \geqslant 0. \iff \exists C \in M_n(\mathbb{R})$ 使得 $A = C^t \cdot C$。

(2) A 是正定矩阵 $\iff r = s = n. \iff \exists$ 可逆矩阵 P 使得 $P^tAP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$，其中 $\forall i \in \{1, \ldots, n\}$，$\lambda_i > 0. \iff \exists$ 可逆矩阵 $C \in M_n(\mathbb{R})$ 使得 $A = C^t \cdot C$。

证明. 我们只证明二次型情况的 (1)，其余留作练习。
($\Longrightarrow$) 由于 $\mathrm{rank}(q) = r$，q 的正惯性指数为 s，故可以取一组典范基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得在这组基下 $q(\mathbf{x}) = \sum_{i=1}^{s} x_i^2 - \sum_{j=s+1}^{r} x_j^2$。用反证法，如果 $s < r$，那么取 $\mathbf{x} = \mathbf{e}_{s+1}$ 即有 $q(\mathbf{x}) = -1 < 0$，这与 q 半正定矛盾！
($\Longleftarrow$) 由于 $r = s$，故在典范基下，任取 $\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i \in V$，都有 $q(\mathbf{x}) = \sum_{i=1}^{r} x_i^2 \geqslant 0$，即 q 半正定。 □

二次型 (或对称矩阵) 的正定性有什么用呢？我们看下面的例子。

例 1.6.9. 在数学分析中，我们可以用多元函数在极值点处的二阶全微分的系数矩阵 (Hesse 矩阵) 是正定 (或负定) 矩阵来判断极值点是极小值点 (或极大值点)。例如，设 $\varphi(x, y)$ 是 $\mathbb{R}$ 上的二元函数，在点 $(0, 0)$ 的一个邻域内二阶偏导数都连续，且 $\varphi'_x(0, 0) = \varphi'_y(0, 0) = 0$，那么 $(0, 0)$ 是 $\varphi(x, y)$ 的临界点。由多元泰勒公式，在 $(0, 0)$ 附近我们有

$$\varphi(x, y) = \varphi(0, 0) + \varphi'_x(0, 0)x + \varphi'_y(0, 0)y + \frac{1}{2}\left(\varphi''_{xx}(0, 0)x^2 + 2\varphi''_{xy}(0, 0)xy + \varphi''_{yy}(0, 0)y^2\right) + o(x^2, y^2)$$

令二次型 $q(x, y) = \varphi''_{xx}(0, 0)x^2 + 2\varphi''_{xy}(0, 0)xy + \varphi''_{yy}(0, 0)y^2$，则由数学分析的知识很容易得到：当 $q(x, y)$ 正定时，$(0, 0)$ 是 $\varphi(x, y)$ 的极小值点；当 $q(x, y)$ 负定时，$(0, 0)$ 是 $\varphi(x, y)$ 的极大值点；当 $q(x, y)$ 不定时，$(0, 0)$ 是 $\varphi(x, y)$ 的鞍点。更复杂的情形读者可以参考数学分析的标准教材。

在本节的最后，我们来考虑如何利用行列式判断一个二次型 (或对称矩阵) 是否正定。在上一节中我们提及了雅可比方法，它是行列式和二次型 (对称双线性型) 联系的桥梁。显然定理1.6.5与下面的定理是等价的。

定理 1.6.10 (Jacobian). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基，q 是 V 上的二次型，A 是 q 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下对应的矩阵，如果 A 的各阶顺序主子式都不为 0，

则存在 V 的一组基 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 使得 q 在这组基下对应的矩阵为

$$\begin{pmatrix} \frac{\Delta_0}{\Delta_1} & & & \\ & \frac{\Delta_1}{\Delta_2} & & \\ & & \ddots & \\ & & & \frac{\Delta_{n-1}}{\Delta_n} \end{pmatrix}.$$

我们还是对定理1.6.5的形式进行证明。

证明. 对 V 的维数 n 用数学归纳法。

当 $n = 1$ 时，任取 $\mathbf{e}_1 \neq 0$，则它是 V 的一组基。设此时 f 对应的矩阵为 $A = (a_{11})$，则显然 $a_{11} = f(\mathbf{e}_1, \mathbf{e}_1) \neq 0$。令 $\mathbf{e}_1' = \dfrac{1}{a_{11}}\mathbf{e}_1$，则 f 在基 $\mathbf{e}_1'$ 下的矩阵为 $(\dfrac{1}{a_{11}}) \cdot (a_{11}) \cdot (\dfrac{1}{a_{11}}) = (\dfrac{1}{a_{11}})$。由于 $\Delta_0 = 1$，$\Delta_1 = a_{11}$，故 $\dfrac{1}{a_{11}} = \dfrac{\Delta_0}{\Delta_1}$ 显然成立。

现在我们假设定理对任意的 $n-1$ 维空间都成立，下面考虑 $\dim(V) = n$ 的情形。令 $V_1 = \text{span}\,\{\mathbf{e}_1,\ldots,\mathbf{e}_{n-1}\}$，则 $f\big|_{V_1}$ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_{n-1}$ 下的矩阵为 $B = \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$。于是 B 的各阶顺序主子式恰为 $\Delta_1,\ldots,\Delta_{n-1}$。由归纳假设，存在 V_1 的一组基 $\mathbf{e}_1',\ldots,\mathbf{e}_{n-1}'$ 使得 $f\big|_{V_1}$ 在这组基下的矩阵为

$$\begin{pmatrix} \frac{\Delta_0}{\Delta_1} & & & \\ & \frac{\Delta_1}{\Delta_2} & & \\ & & \ddots & \\ & & & \frac{\Delta_{n-2}}{\Delta_{n-1}} \end{pmatrix}_{(n-1)\times(n-1)}.$$

下面我们需要寻找一个合适的 $\mathbf{e}_n'$ 使得 $\mathbf{e}_1',\ldots,\mathbf{e}_{n-1}',\mathbf{e}_n'$ 是 V 的一组基，并且 f 在这组基下的矩阵恰好是定理所要求的矩阵。 (※)

为此，记 $f_i(\mathbf{x}) = f(\mathbf{x},\mathbf{e}_i')$，$\forall i \in \{1,\ldots,n\}$，则 $f_i \in V^*$。作抽象的齐次线性方程组

$$\begin{cases} f_1(\mathbf{x}) = 0 \\ f_2(\mathbf{x}) = 0 \\ \qquad\vdots \\ f_{n-1}(\mathbf{x}) = 0 \end{cases},\quad \mathbf{x} \in V. \tag{1.6.5}$$

则由于 $\dim(\text{span}\,\{f_1,\ldots,f_{n-1}\}) \leqslant n-1$，由定理1.5.6可知方程组 (1.6.5) 的解空间维数大于等于 1，即该方程组有非零解。

在方程组 (1.6.5) 的解空间任取一个非零解 $\mathbf{x}_0$，则对任意的 $\lambda \in \mathbb{K}\backslash\{0\}$，我们有 $\mathbf{e}_1',\ldots,\mathbf{e}_{n-1}',\lambda\mathbf{x}_0$ 是 V 的一组基。理由如下。用反证法，如果 $\alpha_1,\ldots,\alpha_n \in \mathbb{K}$ 不全为 0，而线性组合

$$\alpha_1\mathbf{e}_1' + \cdots + \alpha_{n-1}\mathbf{e}_{n-1}' + \alpha_n\lambda\mathbf{x}_0 = \mathbf{0}.$$

那么，首先有 $\alpha_n \neq 0$ (否则与 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}$ 是 V_1 的一组基矛盾！)，那么我们就有

$$\mathbf{x}_0 = -(\alpha_n \lambda)^{-1}(\alpha_1 \mathbf{e}'_1 + \cdots + \alpha_{n-1}\mathbf{e}'_{n-1})$$

显然此时 $\alpha_1, \dots, \alpha_{n-1}$ 也不全为 0 (否则 $\mathbf{x}_0 = \mathbf{0}$ 矛盾！)，不妨设 $\alpha_i \neq 0$，则

$$f(\mathbf{x}_0, \mathbf{e}'_i) = -(\alpha_n \lambda \alpha_i) f(\mathbf{e}'_i, \mathbf{e}'_i) = -(\alpha_n \lambda \alpha_i)\frac{\Delta_{i-1}}{\Delta_i} \neq 0,$$

这与 $\mathbf{x}_0$ 是方程组 (1.6.5) 的解矛盾！故 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}, \lambda\mathbf{x}_0$ 线性无关，从而是 V 的一组基。

到此时，我们已经得到 f 在基底 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}, \lambda\mathbf{x}_0$ 下的矩阵为

$$A' = \begin{pmatrix} \frac{\Delta_0}{\Delta_1} & & & & \\ & \frac{\Delta_1}{\Delta_2} & & & \\ & & \ddots & & \\ & & & \frac{\Delta_{n-2}}{\Delta_{n-1}} & \\ & & & & f(\lambda\mathbf{x}_0, \lambda\mathbf{x}_0) \end{pmatrix}_{n\times n}.$$

最后我们来确定 λ 的值使得 $f(\lambda\mathbf{x}_0, \lambda\mathbf{x}_0) = \dfrac{\Delta_{n-1}}{\Delta_n}$ 即可。不妨设 V_1 的基 $\mathbf{e}_1, \dots, \mathbf{e}_{n-1}$ 到 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}$ 的过渡矩阵为 $T = (t_{ij})_{(n-1)\times(n-1)}$，并且 $\mathbf{x}_0 = \sum_{i=1}^n a_i \mathbf{e}_i$，则

$$(\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}, \lambda\mathbf{x}_0) = (\mathbf{e}_1, \dots, \mathbf{e}_n) \cdot \begin{pmatrix} t_{11} & \cdots & t_{1,n-1} & \lambda a_1 \\ \vdots & & \vdots & \vdots \\ t_{n-1,1} & \cdots & t_{n-1,n-1} & \lambda a_{n-1} \\ 0 & \cdots & 0 & \lambda a_n \end{pmatrix}.$$

于是 V 的基 $\mathbf{e}_1, \dots, \mathbf{e}_{n-1}, \mathbf{e}_n$ 到 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}, \lambda\mathbf{x}_0$ 的过渡矩阵为

$$T_\lambda = \begin{pmatrix} T & \mathbf{a} \\ \mathbf{0} & \lambda a_n \end{pmatrix}, \quad 其中 \mathbf{a} = (a_1, \dots, a_{n-1})^t.$$

所以 $\det(T_\lambda) = \lambda a_n \det(T)$ 是 λ 的线性函数。我们令 $\lambda = (a_n \det(T)\Delta_n)^{-1}$，即 $\det(T_\lambda) = \Delta_n^{-1}$，则由 $A' = T_\lambda^t A T_\lambda$ 可得

$$\frac{1}{\Delta_{n-1}} f(\lambda\mathbf{x}_0, \lambda\mathbf{x}_0) = \det(A') = [\det(T_\lambda)]^2 \det(A) = \frac{1}{\Delta_n^2}\Delta_n = \frac{1}{\Delta_n}$$

即此时 $f(\lambda\mathbf{x}_0, \lambda\mathbf{x}_0) = \dfrac{\Delta_{n-1}}{\Delta_n}$。

综上所述，令 $\mathbf{e}'_n = \lambda\mathbf{x}_0$，$\lambda = (a_n \det(T)\Delta_n)^{-1}$，则 $\mathbf{e}'_1, \dots, \mathbf{e}'_{n-1}, \mathbf{e}'_n$ 就满足 $(\ast)$ 的要求。这样我们就完成了归纳证明。 □

由雅可比方法我们很容易得到下面的推论。

推论 1.6.2. 设 $A \in SM_n(\mathbb{R})$，则 A 的负惯性指数等于序列 $1, \Delta_1, \dots, \Delta_n$ 的变号数。

同时我们得到了以下的判定实对称矩阵正定的方法。

定理 1.6.11 (Sylvester 准则). 设 $A \in SM_n(\mathbb{R})$，则 A 是正定矩阵 $\Longleftrightarrow$ A 的各阶顺序主子式都大于 0。

证明都是显然的。

我们也可以用行列式工具来判断一个实对称矩阵是否是半正定的，这就是下面的定理。

定理 1.6.12. 设 $A = (a_{ij})_{n\times n} \in SM_n(\mathbb{R})$，则 A 是半正定矩阵 $\Longleftrightarrow$ $\forall 1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant n$，都有 $M_A\binom{i_1,\cdots,i_m}{i_1,\cdots,i_m} \geqslant 0$ ($M_A\binom{i_1,\cdots,i_m}{i_1,\cdots,i_m}$ 称为 A 的主子式，这个记号的含义参见讲义上册 §3.3 节)。

证明. ($\Longrightarrow$) 我们首先证明：如果实对称矩阵 $A = (a_{ij})_{n\times n}$ 是半正定矩阵，则 A 的任意一个子矩阵

$$A_m = \begin{pmatrix} a_{i_1,i_1} & a_{i_1,i_2} & \cdots & a_{i_1,i_m} \\ a_{i_2,i_1} & a_{i_2,i_2} & \cdots & a_{i_2,i_m} \\ \vdots & \vdots & & \vdots \\ a_{i_m,i_1} & a_{i_m,i_2} & \cdots & a_{i_m,i_m} \end{pmatrix}$$

也是半正定的。为此，只需证 $\forall \mathbf{x} = (x_{i_1,\ldots,x_{i_m}})^t \in \mathbb{R}^m$，都有 $\mathbf{x}^t A_m \mathbf{x} \geqslant 0$。作延长向量 $\widetilde{\mathbf{x}} = (x'_1, \ldots, x'_n)$，其中

$$x'_k = \begin{cases} x_{i_j}, & 若 k = i_j, j \in \{1, \ldots, m\}; \\ 0, & 其它. \end{cases}.$$

那么，由 A 半正定可知 $\widetilde{\mathbf{x}}^t A \widetilde{\mathbf{x}} \geqslant 0$，而 $\mathbf{x}^t A_m \mathbf{x} = \widetilde{\mathbf{x}}^t A \widetilde{\mathbf{x}}$，故 A_m 半正定。

于是，由定理1.6.9可知，存在 $C \in M_m(\mathbb{R})$ 使得 $A_m = C^t C$，故 $\det(A_m) = (\det(C))^2 \geqslant 0$，即 $M_A\binom{i_1,\cdots,i_m}{i_1,\cdots,i_m} \geqslant 0$。

($\Longleftarrow$) 我们先证明：若 A 的所有主子式都非负，则 $\forall t > 0$，都有 $tE + A$ 是正定矩阵。为此，我们考虑 $tE + A$ 的各阶顺序主子式

$$\Delta_0 = 1,\ \Delta_1 = t + a_{11},\ \Delta_2 = \begin{vmatrix} t + a_{11} & a_{12} \\ a_{21} & t + a_{22} \end{vmatrix}, \cdots\cdots, \Delta_n = \begin{vmatrix} t + a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & t + a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & t + a_{nn} \end{vmatrix}.$$

我们需要证明它们在 $t > 0$ 时全为正数。实际上，任取 $m \in \{1, \ldots, n\}$ 并记 $A_{(m)}$ 是 Δ_m 所对应的矩阵，则容易看出 Δ_m 是关于 t 的 m 次多项式。下面我们具体地求解 Δ_m 关于 t 的各项系数。

显然 t^m 的系数是 1，故我们可设 $\Delta_m = t^m + p_{m-1}t^{m-1} + \cdots + p_1 t + p_0$。令 $t = 0$ 可知 $p_0 = M_A\binom{1,\cdots,m}{1,\cdots,m}$；对 t 求一阶导数后令 $t = 0$ 可得 (利用行列式的求导法则，参考《数学分析(第三版)》，华东师范大学出版社，第五章总练习题 9)：

$$p_1 = \begin{vmatrix} 1 & a_{12} & \cdots & a_{1m} \\ 0 & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2} & \cdots & a_{mm} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & \cdots & a_{1m} \\ a_{21} & 1 & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & 0 & \cdots & a_{mm} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & 1 \end{vmatrix} = A_{(m)}\text{的所有 } m-1 \text{ 阶主子式之和}.$$

类似地，对 t 求 k 阶导数后令 $t = 0$，可知 p_k 是 $A_{(m)}$ 的所有 $m - k$ 阶主子式之和[1](细节留作

---

[1] 这个结论我们在下一章还会用到。

练习)。由于 A 的所有主子式都非负，而 $t>0$，故 $\Delta_m \geqslant t^m > 0$。因此，由 Sylvester 准则，$tE+A$ 是正定矩阵。

下面我们可以证明 A 是半正定矩阵了。用反证法，如果 A 不是半正定矩阵，那么一定存在 $\mathbf{y}=(y_1,\ldots,y_n)^t \in \mathbb{R}^n$ 使得 $\mathbf{y}^tA\mathbf{y}=-c<0$。令 $t=\dfrac{c}{2\mathbf{y}^t\mathbf{y}}>0$，则

$$\mathbf{y}^t(tE+A)\mathbf{y}=t\mathbf{y}^t\mathbf{y}+\mathbf{y}^tA\mathbf{y}=\frac{1}{2}c-c=-\frac{1}{2}c<0.$$

这与任取 $t>0$，$tE+A$ 正定相矛盾！故 A 是半正定矩阵。 □

最后我们看一个计算的例子。

例 1.6.10. 设 $A=\begin{pmatrix}\lambda & -1 & -1\\ -1 & \lambda & 1\\ -1 & 1 & \lambda\end{pmatrix}\in M_3(\mathbb{R})$ 是正定矩阵，求 λ 的取值范围。

解. 由 Sylvester 准则，我们有

$$\Delta_1=\lambda>0,\ \Delta_2=\begin{vmatrix}\lambda & -1\\ -1 & \lambda\end{vmatrix}=\lambda^2-1>0,\ \Delta_3=\det(A)=\lambda^3-3\lambda+2>0.$$

解得 $\lambda\in(1,\infty)$。 □

1.6.7 斜对称双线性型与普法夫型

前面我们已经比较深入地讨论了对称双线性型及其标准型，那么，与之相对地，我们会问：斜对称双线性型具有什么样的性质呢？这就是我们这一节的主要内容。

设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})\neq 2)$ 上的 n 维向量空间，$f\in L_2^-(V,\mathbb{K})$，也即 $\forall \mathbf{x},\mathbf{y}\in V$，$f(\mathbf{x},\mathbf{y})=-f(\mathbf{y},\mathbf{x})$。我们同样可以考虑 f 的左核 (等于右核，思考之)$L_f=\{\mathbf{v}\in V\mid \forall\mathbf{y}\in V,\ f(\mathbf{v},\mathbf{y})=0\}$，显然 L_f 是 V 的子空间。设 V_1 是 L_f 在 V 中的补空间，则容易验证 $f|_{V_1}$ 是 V_1 上的非退化双线性型 (留作练习)。

下面我们考虑斜对称双线性型在合同变换下的标准型。首先，设 $f\in L_2^-(V,\mathbb{K})$，$V=V_1\oplus L_f$，$\mathbf{e}_1,\ldots,\mathbf{e}_m$ 是 V_1 的一组基，$\mathbf{e}_{m+1},\ldots,\mathbf{e}_n$ 是 L_f 的一组基。设 f 在 V 的基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下对应的矩阵为 A，则容易证明，A 的左上角 $m\times m$ 的子块是满秩的斜对称矩阵，而其余部分都为 0(留作练习)。

因此，我们只需要考虑 f 的 “非退化部分”。换言之，我们首先要考虑 f 何时是非退化的。取定 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 后，设 f 对应的矩阵为 $A\in M_n(\mathbb{K})$，则由命题1.6.1，$A^t=-A$，所以 $\det(A)=\det(A^t)=\det(-A)=(-1)^n\det(A)$。故如果 V 的维数 n 是奇数，必有 $\det(A)=0$，即 f 退化。因此，以后我们只需要讨论偶数维空间上的非退化斜对称双线性型即可。

我们有下面的定理：

定理 1.6.13. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})\neq 2)$ 上的 $n\ (=2m,\ m\in\mathbb{Z}^+)$ 维向量空间，$f\in L_2^-(V,\mathbb{K})$ 非退化，则存在 V 的一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2,\ldots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_m$ 使得 f 在这组基下对应的矩阵为分块对角矩阵

$$\begin{pmatrix}D & & & \\ & D & & \\ & & \ddots & \\ & & & D\end{pmatrix}_{2m\times 2m},$$

其中 $D=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\in M_2(\mathbb{K})$。我们称上面的这组基 $\boldsymbol{\alpha_1},\boldsymbol{\beta_1},\boldsymbol{\alpha_2},\boldsymbol{\beta_2},\ldots,\boldsymbol{\alpha_m},\boldsymbol{\beta_m}$ 为辛基 (symplectic basis)。

证明. 对 m 做数学归纳法。$m=1$ 时，任取 $\boldsymbol{\alpha_1}\in V$，由 f 非退化可知 $\exists\boldsymbol{\beta}\in V$ 使得 $f(\boldsymbol{\alpha_1},\beta)\neq 0$。令 $\boldsymbol{\beta_1}=(f(\boldsymbol{\alpha_1},\beta))^{-1}\boldsymbol{\beta}$，则

$$f(\boldsymbol{\alpha_1},\boldsymbol{\beta_1})=-f(\boldsymbol{\beta_1},\boldsymbol{\alpha_1})=1,$$

显然 $f(\boldsymbol{\alpha_1},\boldsymbol{\alpha_1})=f(\boldsymbol{\beta_1},\boldsymbol{\beta_1})=0$，即 f 在基 $\boldsymbol{\alpha_1},\boldsymbol{\beta_1}$ 下的矩阵为 $D=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$。

下设 $\dim(V)=2m$ 且定理对 $\dim(V)=2(m-1)$ 的情形成立。由于 f 非零，故存在 $\boldsymbol{\alpha_1},\boldsymbol{\beta}\in V$ 使得 $f(\boldsymbol{\alpha_1},\boldsymbol{\beta})\neq 0$，取 $\boldsymbol{\beta_1}=(f(\boldsymbol{\alpha_1},\beta))^{-1}\boldsymbol{\beta}$ 即得 $f(\boldsymbol{\alpha_1},\boldsymbol{\beta_1})=-f(\boldsymbol{\beta_1},\boldsymbol{\alpha_1})=1$。容易证明此时 $\boldsymbol{\alpha_1},\boldsymbol{\beta_1}$ 线性无关 (否则 $f(\boldsymbol{\alpha_1},\boldsymbol{\beta_1})=0$)，于是 V^* 中的线性函数 $f_1:\mathbf{x}\mapsto f(\mathbf{x},\boldsymbol{\alpha_1})$ 和 $f_2:\mathbf{x}\mapsto f(\mathbf{x},\boldsymbol{\beta_1})$ 也线性无关 (向 f_1,f_2 的线性组合中分别代入 $\boldsymbol{\alpha_1},\boldsymbol{\beta_1}$ 即可，留作练习)。考虑抽象齐次线性方程组

$$f_1(\mathbf{x})=0,\ f_2(\mathbf{x})=0.$$

则由定理1.5.6，其解空间 U 是 V 的 $2m-2$ 维子空间。容易证明 $f\big|_U$ 仍是非退化的斜对称双线性型，故按照归纳假设，存在 U 的一组基 $\boldsymbol{\alpha_2},\boldsymbol{\beta_2},\ldots,\boldsymbol{\alpha_m},\boldsymbol{\beta_m}$ 使得 $f\big|_U$ 在这组基下的矩阵为

$$\begin{pmatrix}D & & \\ & \ddots & \\ & & D\end{pmatrix}_{(2m-2)\times(2m-2)},\quad D=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}.$$

再由 U 的选取就知道 f 在 V 的基底 $\boldsymbol{\alpha_1},\boldsymbol{\beta_1},\boldsymbol{\alpha_2},\boldsymbol{\beta_2},\ldots,\boldsymbol{\alpha_m},\boldsymbol{\beta_m}$ 下的矩阵就是定理所求 (细节留作练习)。这样我们就完成了证明。 □

在上面的证明中，我们称满足 $f(\boldsymbol{\alpha},\boldsymbol{\beta})=-f(\boldsymbol{\beta},\boldsymbol{\alpha})=1$ 的 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$ 所张成的子空间 span $\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ 为辛平面 (symplectic plane)。上面的定理也可以表述成如下的矩阵形式：

推论 1.6.3. 任何域 $\mathbb{K}(\mathrm{char}(\mathbb{K})\neq 2)$ 上的偶数阶非退化斜对称矩阵 $A\in AM_{2m}(\mathbb{K})$ 都可以合同变换到

$$J=\begin{pmatrix}D & & & \\ & D & & \\ & & \ddots & \\ & & & D\end{pmatrix}_{2m\times 2m},\ D=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\ \text{或者}\ J_0=\begin{pmatrix}0 & E_m\\ -E_m & 0\end{pmatrix}.$$

证明. A 可以合同变换到 J 是定理1.6.13的直接推论。将基的顺序调整成 $\boldsymbol{\alpha_1},\ldots,\boldsymbol{\alpha_m},\boldsymbol{\beta_1},\ldots,\boldsymbol{\beta_m}$ 即可证明 A 可以合同变换到 J_0。 □

在本节的最后，我们利用斜对称双线性型构造一类特殊的多项式，它们被称为 Pfaffian(普法夫) 多项式。设 $1\leqslant i<j\leqslant n=2m$，我们取 $\dfrac{n(n-1)}{2}$ 个符号 (未定元)t_{ij}，作 $\mathbb{Z}$ 上的多元多项式环 $\mathbb{Z}[t_{12},\ldots,t_{1n},\ldots\ldots,t_{(n-1)n}]$。令 $\mathbb{K}$ 是上面多项式环的分式域，即

$\mathbb{K}=\mathbb{Q}(t_{12},\dots,t_{1n},\dots\dots,t_{(n-1)n})$。作斜对称矩阵

$$T=\begin{pmatrix} 0 & t_{12} & t_{13} & \cdots & t_{1n}\\ -t_{12} & 0 & t_{23} & \cdots & t_{2n}\\ -t_{13} & -t_{23} & 0 & \cdots & t_{3n}\\ \vdots & \vdots & \vdots & & \vdots\\ -t_{1n} & -t_{2n} & -t_{3n} & \cdots & 0\end{pmatrix}\in AM_n(\mathbb{K})$$

则显然 $\det(T)\in\mathbb{Q}[t_{12},\dots,t_{1n},\dots\dots,t_{(n-1)n}]$。首先，$\det(T)\neq 0$ (这是因为当 t_{ij} 赋具体值时 $\det(T)$ 不全为 0)，即 T 在 $\mathbb{K}$ 上非退化 (可逆)，故由推论1.6.3我们知道，存在可逆矩阵 $A\in M_n(\mathbb{K})$ 使得

$$A^tTA=\begin{pmatrix} 0 & E_m\\ -E_m & 0\end{pmatrix}.$$

两边取行列式即得 $[\det(A)]^2\det(T)=1$，即 $\det(T)=\left(\dfrac{1}{\det(A)}\right)^2$。这说明 $\det(T)$ 是域 $\mathbb{K}$ 上的平方元。但我们注意到 $\det(T)$ 实际上是整系数多项式，因此它一定是某个多项式

$$P_n(t_{12},\dots,t_{1n},\dots\dots,t_{n-1,n})\in\mathbb{Z}[t_{12},\dots,t_{1n},\dots\dots,t_{n-1,n}]$$

的平方。这是因为：设 $\det(T)=P_n^2$，如果 $P_n=\dfrac{f}{g}$，f,g 是整系数多项式且 $\gcd(f,g)=1$，那么必有 $f^2=g^2\cdot\det(T)$，由于整系数多项式环是唯一因子分解整环，f 中不含有 g 的因子，故 $g=1$，P_n 是整系数多项式。显然也有 $(-P_n)^2=\det(T)$，因此，为了保证我们定义出来的东西是一个确定的多项式，我们还需要规定 P_n 的符号。一般我们称在

$$T=\begin{pmatrix} 0 & t_{12} & \cdots & t_{1n}\\ -t_{12} & 0 & \cdots & t_{2n}\\ \vdots & \vdots & & \vdots\\ -t_{1n} & -t_{2n} & \cdots & 0\end{pmatrix}=\begin{pmatrix} 0 & E_m\\ -E_m & 0\end{pmatrix}$$

处取值 $P_n=(-1)^{m(m-1)/2},n=2m$ 的 $\det(T)$ 的平方根为 $n(=2m)$ 维**普法夫多项式**，记作 $\mathrm{Pf}(t_{12},\dots,t_{(n-1)n})$ 或 $\mathrm{Pf}(\mathbf{t})$[1]。规定所有奇数维普法夫多项式为零多项式。此外，设 $\mathbb{K}$ 是特征不为 2 的域，我们称由普法夫多项式确定的在 n 阶斜对称矩阵上的赋值同态：

$$\begin{aligned}\mathrm{Pf}_n\ :\ AM_n(\mathbb{K})&\longrightarrow\mathbb{K}\\ (f_{ij})_{n\times n}&\longmapsto\mathrm{Pf}_n(f_{12},\dots,f_{(n-1)n}).\end{aligned}$$

为 n 阶普法夫型 (我们用记号 $\mathrm{Pf}_n(F)$ 代表将 t_{ij} 替换为 f_{ij} 赋值同态)。

例 1.6.11. 2 维普法夫多项式 $\mathrm{Pf}_2(\mathbf{t})=t_{12}$ 是 $\begin{vmatrix} 0 & t_{12}\\ -t_{12} & 0\end{vmatrix}$ 的平方根；4 维普法夫多项式 $\mathrm{Pf}_4(\mathbf{t})=$

---

[1] 需要修改!!! 或者把普法夫多项式定义成在 $T=\begin{pmatrix} 0 & -E_m\\ E_m & 0\end{pmatrix}$ 处取 1 的 $\det(T)$ 的平方根，读者可以自行验证这两个定义是一致的。

$t_{12}t_{34} - t_{13}t_{24} + t_{14}t_{23}$ 是

$$\begin{vmatrix} 0 & t_{12} & t_{13} & t_{14} \\ -t_{12} & 0 & t_{23} & t_{24} \\ -t_{13} & -t_{23} & 0 & t_{34} \\ -t_{14} & -t_{24} & -t_{34} & 0 \end{vmatrix}$$

的平方根。

我们有以下定理。

定理 1.6.14. [1] 设 $\mathbb{K}$ 是特征不为 2 的域，$F \in AM_{2m}(\mathbb{K})$，则 $\det(F) = [\mathrm{Pf}_{2m}(F)]^2$，并且任取 $A \in M_{2m}(\mathbb{K})$，有 $\mathrm{Pf}_{2m}(A^tFA) = \det(A) \cdot \mathrm{Pf}_{2m}(F)$。

证明. 利用 $\det(T) = [\mathrm{Pf}(\mathbf{t})]^2$ 及赋值同态即得 $\det(F) = [\mathrm{Pf}_{2m}(F)]^2$。于是

$$\begin{aligned}[\mathrm{Pf}_{2m}(A^tFA)]^2 &= \det(A^tFA) \\ &= [\det(A)]^2 \det(F) \\ &= [\det(A) \cdot \mathrm{Pf}_{2m}(F)]^2\end{aligned}$$

即 $\mathrm{Pf}_{2m}(A^tFA) = \pm\det(A) \cdot \mathrm{Pf}_{2m}(F)$。令 $A = E_{2m}$，$F = \begin{pmatrix} 0 & E_m \\ -E_m & 0 \end{pmatrix}$，可以推出得我们应取 $\mathrm{Pf}_{2m}(A^tFA) = \det(A) \cdot \mathrm{Pf}_{2m}(F)$。 □

我们这一章中研究双线性型和二次型的思路在接下来的两章中会继续使用，希望读者认真体会。

---

[1] 这里有问题，记得提醒编者修改。

1.7 习题

抽象向量空间, 直和, 线性相关性, 维数与基底

1. 下列集合是否构成域 $\mathbb{R}$ 上的向量空间:
   (1) $M_n(\mathbb{R})$ 中秩为一固定数 r 的所有矩阵;
   (2) $M_n(\mathbb{R})$ 中所有对称矩阵 (${}^tA = A$);
   (3) $M_n(\mathbb{R})$ 中所有斜对称矩阵 (${}^tA = -A$);
   (4) $M_n(\mathbb{R})$ 中所有行列式为零的矩阵;
   (5) $M_n(\mathbb{R})$ 中所有迹为零 $\operatorname{tr} A = 0$ 的矩阵 (矩阵 $A = \left(a_{ij}\right)$ 的迹用关系式 $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$ 定义);
   (6) $M_n(\mathbb{R})$ 中所有迹为正数的矩阵;
   (7) 对固定的 n, 所有形如 $f(t) = (a_1 + a_2 + \cdots + a_n) - a_1 t - \cdots - a_n t^n$ 的多项式 ($a_i \in \mathbb{R}$).

2. 在有限的 p 元域 $\mathbb{F}_p$ 上的长度为 n 的行向量 $(x_1, \cdots, x_n)$ 的坐标空间 $\mathbb{F}_p^n$ 中共有多少个元素? 在 $\mathbb{F}_p^n$ 中方程 $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \left(\alpha_i \in \mathbb{F}_p \text{ 不全等于零 }\right)$ 有多少个解?

3. 对复数域上的坐标空间 $\mathbb{C}^n$, 定义新的纯量乘为 $b \circ (a_1, \cdots, a_n) = (b\overline{a}_1, \cdots, b\overline{a}_n)$. 在运算 $+$ 和 $\circ$ 下 $\mathbb{C}^n$ 是否为向量空间?

4. 设 M 是有限集, 它的所有的子集形成的集合记作 2^M. 取域为 $\mathbb{Z}_2$. 定义 2^M 中的加法和 $\mathbb{Z}_2$ 与 2^M 中元素的乘法如下:

$$A + B = (A \cup B) \backslash (A \cap B), \quad 1A = A, \quad 0A = \varnothing.$$

   证明: 在这两个运算下, 2^M 成为域 $\mathbb{Z}_2$ 上的向量空间.

5. 交换群 A 能成为域 $\mathbb{Z}_p$ 上的向量空间当且仅当对任意的 $x \in A$ 有 $px = 0$.

6. 交换群 A 能成为域 $\mathbb{Q}$ 上的向量空间当且仅当 A 中的非零元都是无限阶的, 且对任意的正整数 n 和 $a \in A$, 方程 $nx = a$ 在 A 中有解.

7. 在单变量实函数空间中, 向量 $f_1, \cdots, f_n$ 线性无关当且仅当存在实数 $a_1, \cdots, a_n$ 使得 $\det\left(f_i\left(a_j\right)\right) \neq 0$.

8. 在 q 元域 $\mathbb{F}_q$ 上的 n 维向量空间 V 中, 对于 $1 \leqslant k \leqslant n$, 有多少个 k 维的子空间?

9. 试指出由下列各种 n 阶实方阵构成的空间的维数: (1) 对称矩阵; (2) 斜对称矩阵; (3) 迹为零的矩阵.

10. 由所有的一个变元 t 的次数 $\leqslant n$ 的满足条件 $f(1) = 0$ 的多项式组成的空间的维数是多少? 找出这个空间的一个基底来.

11. 和 $U = U_1 + U_2 + \cdots + U_m$ 是直和的充要条件是

$$(U_1 + \cdots + U_{i-1}) \cap U_i = 0, \quad 1 < i \leqslant m.$$

12. 找出空间 P_n 的基底 $\left(1, t, \cdots, t^{n-1}\right)$ 向同一空间的基底 $\left(1, (t - \alpha), \cdots, (t - \alpha)^{n-1}\right)$ 的过渡矩阵.

13. 设 θ 是 $\mathbb{Q}$ 上不可约多项式 $f \in \mathbb{Q}[t]$ 的一个复根. 求出空间 $\mathbb{Q}[\theta] = \left\langle 1, \theta, \cdots, \theta^k, \cdots \right\rangle_{\mathbb{Q}}$ 在 $\mathbb{Q}$ 上的维数.

14. 试举出反例: 直和不满足消去律, 也就是说, 由等式 $U \oplus W_1 = U \oplus W_2$, 一般说来, 不能得到 $W_1 = W_2$, 尽管等式有一相同的被加项.

15. 设 $V_1, \ldots, V_k$ 是有限维向量空间 V 的子空间, 求证:
    (1) $\dim(\sum_{i=1}^k V_i) = \sum_{i=1}^k \dim(V_i) - \sum_{i=2}^k \dim(V_i \cap (\sum_{j=1}^{i-1} V_j))$;
    (2) $\dim(\bigcap\limits_{i=1}^{k} V_i) = \dim(\sum_{i=1}^k V_i) - \sum_{j=2}^k \dim((\bigcap\limits_{i=1}^{j-1} V_i) + V_j)$;
    (3) 如果 $\sum_{i=1}^k \dim(V_i) > (k-1) \cdot \dim(V)$, 那么 $\bigcap\limits_{i=1}^{k} V_i \neq \{\mathbf{0}\}$.

16. 证明: 向量空间 V 是无限维的当且仅当 V 中有一个无限的向量序列 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \cdots$ 使得对每个正整数 k, 向量 $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ 线性无关.

商空间, 对偶空间

1. 设 U 是 V 的子空间, 证明: V 与 $U \oplus V/U$ 同构.

2. 对下列子空间 L 说明商空间 $\mathbb{K}[t]/L$ 是否为有限维的:
    (1) L 是所有次数不超过 $n-1$ 的多项式形成的子空间;
    (2) L 是所有可以被 t^n 整除的多项式形成的子空间;
    (3) L 是所有 t^2 的多项式形成的子空间.

3. 设 V 是任意一个向量空间 (不一定是有限维的), 证明 V 的任何一个子空间 U 都有补空间, 并且所有补空间都同构于商空间 V/U.

4. 如果商空间 V/U 是有限维的, 则称向量空间 V 的子空间 U 是余有限维的. 如果 U 和 W 都是 V 的余有限维子空间, 证明:
    (1) $U + W$ 和 $U \cap W$ 也是余有限维子空间;
    (2) $\operatorname{codim}(U + W) + \operatorname{codim}(U \cap W) = \operatorname{codim} U + \operatorname{codim} W$.

5. 设 $\mathbb{K}$ 是域, $V = M_n(\mathbb{K})$. 证明, V 上的每个线性函数 f 必形如 $f(X) = \operatorname{tr} AX$, 其中矩阵 $A = A_f$ 是唯一确定的.

6. 设 $a(t)$ 是 $\mathbb{R}[t]$ 中的一个固定的多项式, P_n 是所有次数 $\leqslant n-1$ 的实的多项式子空间, 判断下列函数是否是 P_n 上的线性函数:
    (1) $f(u) = \int_{-1}^{1} a(t)u(t)dt; \quad u(t) \in P_n$;
    (2) $f(u) = \int_0^1 a(t)u\left(t^2\right)dt$;
    (3) $f(u) = \int_0^1 a(t)[u(t)]^2 dt$;
    (4) $f(u) = \frac{d^3}{dt^3} a(t)u(t)\Big|_{t=-1}$.

7. 设 V 是向量空间, 再设 $f, g \in V^*$ 且 $\operatorname{Ker} f = \operatorname{Ker} g$. 证明, 必有某个纯量 λ 使得 $g = \lambda f$.

8. 设 $\mathbf{x}$ 是向量空间 V 的一个非零向量. 条件 $f(\mathbf{x}) = 1$ 能将函数 $f \in V^*$ 唯一决定吗?

9. 证明, 在域 $\mathbb{K}$ 上的 n 维空间 V 中, 对任意非零的线性函数 f 都能找到空间 V 的一个基底 $(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ 使得

$$f(\alpha_1 \mathbf{e}_1 + \cdots + \alpha_n \mathbf{e}_n) = \alpha_i, \quad \forall \alpha_i \in \mathbb{K}.$$

10. 设 V 是域 $\mathbb{K}$ 上的向量空间, $f, g \in V^*$. 证明: 如果对所有的向量 $\mathbf{x} \in V$ 都有 $f(\mathbf{x})g(\mathbf{x}) = 0$, 那么这两个线性函数中有一个是零函数.

11. 设 V 是域 $\mathbb{K}$ 上的向量空间, $f_0, f_1, \ldots, f_n \in V^*$, 并且 $f_1, \ldots, f_n$ 线性无关. 若 $\ker(f_0) \supseteq \bigcap\limits_{k=1}^{n} \ker(f_k)$, 证明: f_0 能够写成 $f_1, \ldots, f_n$ 的线性组合.

双线性型和二次型

1. 设 $\Delta_1, \cdots, \Delta_n = F$ 是矩阵 F 对应的实二次型 q 的顺序主子式. 证明, 只有当 $(-1)^k\Delta_k > 0$, 对所有 $k = 1, 2, \cdots, n$ 都成立, q 和 F 才能是负定的.

2. 举出反例:
   (1) 正定矩阵 $A = \left(a_{ij}\right)$, 但有某个对 (i, j) 使得 $a_{ij} < 0$;
   (2) 矩阵 $A = \left(a_{ij}\right)$, 对所有的指标都有 $a_{ij} > 0$, 但 A 却不是正定的.

3. 找出所有 $\lambda, \mu \in \mathbb{R}$, 使得矩阵

$$\begin{pmatrix} 1 & \lambda & \lambda \\ \lambda & 1 & \lambda \\ \lambda & \lambda & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & \mu \\ 1 & \mu & 1 \\ \mu & 1 & 1 \end{pmatrix}$$

   是正定的.

4. 说明下面的函数是否是某个空间的双线性型, 其中 $\mathbb{K}$ 是域:
   (1) $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t \cdot \mathbf{y}$ ($x, y \in \mathbb{K}^n$ 是列向量);
   (2) $f(A, B) = \operatorname{tr}(A^tB)\,(A, B \in M_{m,n}(K))$;
   (3) $f(A, B) = \det(AB)\left(A, B \in M_{\mathrm{p}}(K)\right)$;
   (4) $f(A, B)$ 是矩阵 AB 在 (i, j) 处的值;
   (5) $f(u, v) = \int_a^b uvt^2dt \quad (u, v$ 是闭区间 $[a, b]$ 上的连续函数);
   (6) $f(u, v) = \int_a^b (u + v)^2dt$;
   (7) $f(u, v) = (uv)(a) \quad (u, v \in K[x], a \in K)$;
   (8) $f(u, v) = \frac{d}{dt}(uv)(a)$.

5. 对列向量空间 $\mathbb{R}^2$ 上的双线性型 $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}\mathbf{y}$ 求典范基.

6. 求出以下双线性型的左核，并分别通过合同变换法、配方法和雅可比法将其化成典范式:

   (a) $f(\mathbf{x}, \mathbf{y}) = x_1y_1 - x_1y_2 + x_1y_3 - x_2y_1 + x_2y_2 + 2x_2y_3 + x_3y_1 + 2x_3y_2 + x_3y_3$.

   (b) $f(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + 2x_2y_3 + 2x_3y_2 + 4x_3y_3$ (不需雅可比法).

7. 判断下面两个矩阵是否合同:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & -1 \\ 3 & -1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 1 \\ 0 & 1 & 5 \end{pmatrix}.$$

8. 已知实二次型 $f(\mathbf{x}) = ax_1^2 + x_2^2 + 3x_3^2 - 2x_1x_2$ 的秩为 2，
(1) 写出二次型所对应的矩阵 A, 并求参数 a;
(2) 把二次型化成标准形.

9. 设 $A, B \in M_n(\mathbb{R})$ 是正定矩阵, 求证 A^{-1} 和 $A + B$ 也正定.

10. 设 $A \in M_n(\mathbb{R})$, 求证 $\operatorname{rank}(A^tA) = \operatorname{rank}(A)$.

11. 实二次型

$$\lambda x_1^2 - 2x_2^2 - 3x_3^2 + 2x_1x_2 - 2x_1x_3 + 2x_2x_3$$

在 λ 取什么值时是负定的.

12. 设 $A = \left(a_{ij}\right) \in M_n(\mathbb{R})$ 是对称矩阵,

$$L(f) = \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

是多项式环 $\mathbb{R}[x_1, \cdots, x_n]$ 上的微分算子, 证明 (1) 如果 $C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, C \in GL_n(\mathbb{R})$, 是变量替换, 那么

$$L(f) = \sum_{i,j=1}^{n} b_{ij} \frac{\partial^2 f}{\partial y_i \partial y_j},$$

其中 $\left(b_{ij}\right) = CA^tC$;
(2) 可以找到某个 $C \in GL_n(\mathbb{R})$ 使得

$$L(f) = \frac{\partial^2 f}{\partial y_1^2} + \cdots + \frac{\partial^2 f}{\partial y_s^2} - \frac{\partial^2 f}{\partial y_{s+1}^2} - \cdots - \frac{\partial^2 f}{\partial y_t^2},$$

其中 $0 \leqslant s \leqslant t \leqslant n$.

13. 设 $x = (x_1, x_2, x_3)^t \in \mathbb{C}^3, q(x) = x_1^3 + x_2^3 + x_3^3 - 3x_1x_2x_3, \varepsilon$ 是 1 的三次本原根. 利用等式

$$q(x) = (x_1 + x_2 + x_3)\left(x_1 + \varepsilon x_2 + \varepsilon^2 x_3\right)\left(x_1 + \varepsilon^2 x_2 + \varepsilon x_3\right)$$

证明: (1) 存在对称双线性型 $z_i = z_i(\mathbf{x}, \mathbf{y}) = \sum_{j,k} a_{jk}^{(i)} x_j y_k$ 使得 $q(\mathbf{x})q(\mathbf{y}) = q(\mathbf{z})$, 其中 $\mathbf{z} = (z_1, z_2, z_3)^t$;
(2) 求出这些双线性型.

14. (Hadamard 不等式) 设 $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$, 求证:

$$\det A \leqslant \sqrt{\sum_{i=1}^{n} a_{i1}^2} \cdot \sqrt{\sum_{i=1}^{n} a_{i2}^2} \cdots \sqrt{\sum_{i=1}^{n} a_{in}^2}.$$

15. (Schur 乘积定理) 设 $A = (a_{ij})_{n\times n}, B = (b_{ij})_{n\times n} \in M_n(\mathbb{R})$ 是半正定矩阵, 定义 $A \odot B = (a_{ij} \cdot b_{ij})_{n\times n}$, 求证 $A \odot B$ 也是半正定的.

16. 设 A 是域 $\mathbb{K}$ 上的 n 阶方阵, $\mathrm{char}(\mathbb{K}) \neq 2$. 证明 A 是斜对称的当且仅当对所有的 $\mathbf{x} = (x_1, \cdots, x_n)^t$ 有 $\mathbf{x}^t A\mathbf{x} = 0$.

17. 把斜对称双线性型

$$x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1 + x_2y_3 - x_3y_2 + x_2y_4 - x_4y_2$$

化成典范式.

18. 设 A 是 n 阶斜对称实方阵, E 是 n 阶单位矩阵, λ 是复数. 如果 $\det(\lambda E - A) = 0$, 那么 λ 的实部为 0 .

19. 设 A 是实对称矩阵, E 是同阶单位矩阵, 证明: 存在充分小的实数 $\varepsilon > 0$, 使得方阵 $E + \varepsilon A$ 是正定的.

第二章　线性算子代数与矩阵的若尔当标准型

这一章我们仍然关注两个问题：一是取定向量空间的基后，线性算子如何用矩阵表示，以及基变换下线性算子对应的矩阵如何变换；二是如何选取一组基使得线性算子在这组基下对应的矩阵有比较简单的形式，对应于向量空间的分解。我们将在这一章比较完整地回答这两个问题。

2.1　向量空间上的线性映射

实际上，我们在上一章已经定义过线性映射了。这里我们再重复一下这个定义。

定义 2.1.1. 设 V, W 分别是域 $\mathbb{K}$ 上的 n 和 m 维向量空间，如果映射 $f: V \to W$ 满足：对 $\forall \alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in V$，都有

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

则我们称 f 是 V 到 W 的线性映射 (linear map)。我们把 V 到 W 的所有线性映射所构成的集合记作 $L(V, W)$ 或 $\mathrm{Hom}(V, W)$。显然我们可以在 $\mathrm{Hom}(V, W)$ 上自然地定义加法和数乘如下：任取 $\mathbf{x} \in V$ 及 $\alpha \in \mathbb{K}$，定义

$$(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}),\ \ (\alpha f)(\mathbf{x}) = \alpha \cdot f(\mathbf{x})$$

容易验证在此定义下 $\mathrm{Hom}(V, W)$ 是 $\mathbb{K}$ 上的向量空间。

定义 2.1.2. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，$f \in \mathrm{Hom}(V, W)$。我们也可以定义线性映射的核 $\ker(f) = \{\mathbf{x} \in V \mid f(\mathbf{x}) = \mathbf{0}_W\}$，像 $\mathrm{im}(f) = \{\mathbf{y} \in W \mid \exists \mathbf{x} \in V \text{使得} \mathbf{y} = f(\mathbf{x})\}$。容易验证 $\ker(f)$ 是 V 的子空间，$\mathrm{im}(f)$ 是 W 的子空间，并且 f 是单射 $\Longleftrightarrow$ $\ker(f) = \{\mathbf{0}\}$，f 是满射 $\Longleftrightarrow$ $\mathrm{im}(f) = W$(留作练习)。

下面我们来看一些例子。

例 2.1.1. (1) 设 V 是域 $\mathbb{K}$ 上的向量空间，则零映射 $0: V \to V$, $\mathbf{x} \mapsto \mathbf{0}$ 和恒同映射 $\mathrm{id}: V \to V$, $\mathbf{x} \mapsto \mathbf{x}$ 都是线性映射。

(2) 设 $A \in \mathbb{R}^{m\times n}$，由讲义上册第二章的内容我们知道 $f: \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \mapsto A\mathbf{x}$ 是线性映射，并且 $\ker(f)$ 是方程组 $A\mathbf{x} = \mathbf{0}$ 的解空间，$\mathrm{im}(f)$ 是矩阵 A 的列空间 $V_c(A)$，f 是单射 $\Longleftrightarrow$ $\mathrm{rank}(A) = n$，f 是满射 $\Longleftrightarrow$ $\mathrm{rank}(A) = m$。

类比于上册中 $\mathbb{R}^n$ 上的对偶定理，我们有：

定理 2.1.1 (对偶定理 (抽象线性映射版)). 设 V, W 是域 $\mathbb{K}$ 上的向量空间，$f \in \mathrm{Hom}(V, W)$，则

$$\dim(\ker(f)) + \dim(\mathrm{im}(f)) = \dim(V).$$

证明. 类似于讲义上册定理 2.3.2 的证明，我们可以利用用基扩充定理来证明上面的结论。这里我们给出另一个证明。

由于 $\ker(f)$ 是 V 的子空间，故可以作商空间 $V/\ker(f)$。显然 $f: V \to \mathrm{im}(f)$ 是满射，故由定理1.4.2，我们有 $V/\ker(f) \simeq \mathrm{im}(f)$，所以 $\dim(V/\ker(f)) = \dim(\mathrm{im}(f))$。而由推论1.4.1，$\dim(V/\ker(f)) = \dim(V) - \dim(\ker(f))$，即 $\dim(\ker(f)) + \dim(\mathrm{im}(f)) = \dim(V)$。 □

推论 2.1.1. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，$f \in \mathrm{Hom}(V, W)$，则 $\dim(\mathrm{im} f) \leqslant \dim(V)$，等号当且仅当 f 是单射时成立。

由对偶定理可以直接证明该命题。

线性映射与矩阵有密切的联系。

定义 2.1.3. 设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维向量空间，$f \in \mathrm{Hom}(V, W)$，分别取 V 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 和 W 的一组基 $\mathbf{w}_1, \dots, \mathbf{w}_m$，显然 $\forall i \in \{1, \dots, n\}$，$f(\mathbf{v}_i)$ 都是 $\mathbf{w}_1, \dots, \mathbf{w}_m$ 的线性组合，即我们有

$$\begin{aligned} f(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + \cdots + a_{m1}\mathbf{w}_m \\ f(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + \cdots + a_{m2}\mathbf{w}_m \\ &\vdots \\ f(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + \cdots + a_{mn}\mathbf{w}_m \end{aligned}$$

写成矩阵形式就是

$$(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)) = (\mathbf{w}_1, \dots, \mathbf{w}_m)\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{M_f}$$

我们把上面的系数矩阵 $M_f \in \mathbb{K}^{m \times n}$ 称为线性映射 f 在基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 和 $\mathbf{w}_1, \dots, \mathbf{w}_m$ 下的矩阵。

例 2.1.2. 在 $\mathbb{R}^2$ 上取标准基 $\mathbf{e}_1 = (1, 0)^t$，$\mathbf{e}_2 = (0, 1)^t$，设 $\alpha \in \mathbb{R}$，线性映射 $f : \mathbb{R}^2 \to \mathbb{R}^2$ 满足：

$$f(\mathbf{e}_1) = \cos\alpha\ \mathbf{e}_1 + \sin\alpha\ \mathbf{e}_2, \quad f(\mathbf{e}_2) = -\sin\alpha\ \mathbf{e}_1 + \cos\alpha\ \mathbf{e}_2.$$

则 f 在标准基下的矩阵为 $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$。

定义 2.1.4. 设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维向量空间，$f \in \mathrm{Hom}(V, W)$。我们定义 f 的秩 $\mathrm{rank}(f) = \dim(\mathrm{im}(f))$。显然 f 的秩与基底的选取无关。

我们已经知道，$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{m \times n}$(讲义上册推论 2.4.1)。类似地，我们同样可以证明如下定理：

定理 2.1.2. 设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维向量空间，$f \in \mathrm{Hom}(V, W)$，分别取 V 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 和 W 的一组基 $\mathbf{w}_1, \dots, \mathbf{w}_m$，记 f 在此基底下的矩阵为 M_f，则

(1) $\Phi : \mathrm{Hom}(V, W) \to \mathbb{K}^{m \times n}$，$f \mapsto M_f$ 是线性双射，即 $\mathrm{Hom}(V, W) \simeq \mathbb{K}^{m \times n}$；

(2) $\mathrm{rank}(f) = \mathrm{rank}(M_f)$；

(3) (线性映射基本定理) 在 W 中任取 n 个向量 $\mathbf{u}_1, \dots, \mathbf{u}_n$，都存在唯一的 $f \in \mathrm{Hom}(V, W)$ 使得 $\forall i \in \{1, \dots, n\}$，$f(\mathbf{v}_i) = \mathbf{u}_i$。

证明. (1) 和 (3) 的证明可以参考讲义上册定理 2.4.1，推论 2.4.1 和定理 2.3.1，只需将 $\mathbb{R}$ 换成任意的域 $\mathbb{K}$ 即可。下面证明 (2)。

注意到映射 $\varphi : V_c(M_f) \to \mathrm{im}(f)$，$(b_1, \ldots, b_m)^t = \sum_{i=1}^n \alpha_i (a_{1i}, \ldots, a_{mi})^t \mapsto \sum_{j=1}^m b_j \mathbf{w}_j$ 是线性双射 (验证留作练习)，故

$$\mathrm{rank}(M_f) = \dim(V_c(M_f)) = \dim(\mathrm{im}(f)) = \mathrm{rank}(f).$$

□

推论 2.1.2. 设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维向量空间，则 $\dim(\mathrm{Hom}(V, W)) = m \cdot n$。

下面我们讨论线性映射的复合。我们知道，在 $\mathbb{R}^n$ 中矩阵乘法等价于线性映射的复合，这一点对一般的线性映射仍然成立。

定理 2.1.3. 设 U, V, W 分别是域 $\mathbb{K}$ 上的 n, s, m 维向量空间，$f \in \mathrm{Hom}(U, V)$，$g \in \mathrm{Hom}(V, W)$，则 $g \circ f \in \mathrm{Hom}(U, W)$。并且，分别取定 U, V, W 的某组基底后 f，g，$g \circ f$ 在此基底下的矩阵分别为 $M_f \in \mathbb{K}^{s \times n}$，$M_g \in \mathbb{K}^{m \times s}$ 和 $M_{g \circ f} \in \mathbb{K}^{m \times n}$，则 $M_{g \circ f} = M_g \cdot M_f$。

证明. 首先，验证 $g \circ f \in \mathrm{Hom}(U, V)$ 是容易的，留作练习。(只需验证 $\forall \alpha, \beta \in \mathbb{K}$ 及 $\mathbf{x}, \mathbf{y} \in U$，都有 $g \circ f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha \ g \circ f(\mathbf{x}) + \beta \ g \circ f(\mathbf{y})$)。类似于 $\mathbb{R}^n$ 的情形，我们可以证明矩阵乘法 $M_g \cdot M_f$ 等同于 $\mathbb{K}^n \xrightarrow{\varphi_f} \mathbb{K}^s \xrightarrow{\varphi_g} \mathbb{K}^m$ 的线性映射的复合 (参考讲义上册命题 2.4.4 到定义 2.4.3 之间的部分)，再利用如下交换图即可证明原命题。

$$\begin{tikzcd}
U \arrow[r, leftrightarrow] \arrow[d, "f"] & \mathbb{K}^n \arrow[d, "\varphi_f"] \\
V \arrow[r, leftrightarrow] \arrow[d, "g"] & \mathbb{K}^s \arrow[d, "\varphi_g"] \\
W \arrow[r, leftrightarrow] & \mathbb{K}^m
\end{tikzcd}$$

细节比较繁琐但不难，留给读者补充。 □

关于线性映射的复合，很容易得到以下推论：

推论 2.1.3. 设 U, V, W 分别是域 $\mathbb{K}$ 上的 n, s, m 维向量空间，$f \in \mathrm{Hom}(U, V)$，$g \in \mathrm{Hom}(V, W)$，则 $\dim(\mathrm{im}(g \circ f)) \leqslant \dim(\mathrm{im}(g))$，$\dim(\mathrm{im}(g \circ f)) \leqslant \dim(\mathrm{im}(f))$。

证明. 借助上面的定理以及秩不等式 $\mathrm{rank}(M_g \cdot M_f) \leqslant \min\{\mathrm{rank}(M_g), \mathrm{rank}(M_f)\}$(讲义上册定理 2.4.2) 即可得到结论。或者也可以通过子空间的包含关系及对偶定理证明结论 (留作练习)。 □

在本节的最后，我们来讨论坐标与线性映射的矩阵之间的关系。

设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维线性空间，V 的一组基为 $\mathbf{v}_1, \ldots, \mathbf{v}_n$，$W$ 的一组基为 $\mathbf{w}_1, \ldots, \mathbf{w}_m$，$f \in \mathrm{Hom}(V, W)$。任取 $\mathbf{x} \in V$，设 $\mathbf{x}$ 在基底 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 下的坐标为 $(x_1, \ldots, x_n)^t$，$f(\mathbf{x})$ 在基底 $\mathbf{w}_1, \ldots, \mathbf{w}_m$ 下的坐标为 $(y_1, \ldots, y_m)^t$，并且 f 在基底 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 和 $\mathbf{w}_1, \ldots, \mathbf{w}_m$ 下的矩阵为 M_f，则

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i$$

于是

$$f(\mathbf{x}) = \sum_{i=1}^{n} x_i f(\mathbf{v}_i) = (f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\mathbf{w}_1, \dots, \mathbf{w}_m) \cdot M_f \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

另一方面，

$$f(\mathbf{x}) = \sum_{i=1}^{m} y_i \mathbf{w}_i = (\mathbf{w}_1, \dots, \mathbf{w}_m) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

对比即有

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = M_f \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{2.1.1}$$

也就是说，取定基底后，一个向量在线性映射的下的像的坐标等于线性映射的矩阵乘以原坐标。

我们也可以考虑基变换下线性映射的矩阵的变换。

定理 2.1.4. 设 V, W 分别是域 $\mathbb{K}$ 上的 n, m 维线性空间，$f \in \mathrm{Hom}(V, W)$，$(\mathbf{v}_1, \dots, \mathbf{v}_n)$ 和 $\left(\mathbf{v}'_1, \dots, \mathbf{v}'_n\right)$ 是 V 的两组基，$(\mathbf{w}_1, \dots, \mathbf{w}_m)$ 和 $\left(\mathbf{w}'_1, \dots, \mathbf{w}'_m\right)$ 是 W 的两组基。f 关于 $(\mathbf{v}_i)$ 和 $\left(\mathbf{w}_j\right)$ 的矩阵是 M_f，关于 $\left(\mathbf{v}'_i\right)$ 和 $\left(\mathbf{w}'_j\right)$ 的矩阵是 M'_f。又设 $\left(\mathbf{v}'_1, \dots, \mathbf{v}'_n\right) = (\mathbf{v}_1, \dots, \mathbf{v}_n) \cdot P$，$\left(\mathbf{w}'_1, \dots, \mathbf{w}'_m\right) = (\mathbf{w}_1, \dots, \mathbf{w}_m) \cdot Q$，则

$$M'_f = Q^{-1} M_f P.$$

证明. 任取 $\mathbf{x} \in V$，设 $\mathbf{x}$ 在基底 $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ 和 $\left(\mathbf{v}'_1, \dots, \mathbf{v}'_n\right)$ 下的坐标分别为 $(x_1, \dots, x_n)^t$ 和 $(x'_1, \dots, x'_n)^t$，$f(\mathbf{x})$ 在基底 $(\mathbf{w}_1, \dots, \mathbf{w}_m)$ 和 $\left(\mathbf{w}'_1, \dots, \mathbf{w}'_w\right)$ 下的坐标分别为 $(y_1, \dots, y_m)^t$ 和 $(y'_1, \dots, y'_m)^t$，则由式 (2.1.1)，我们有

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = M_f \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{2.1.2}$$

$$\begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = M'_f \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}. \tag{2.1.3}$$

由坐标变换公式，我们有

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = Q \cdot \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix}$$

将上式代入式 (2.1.2) 得

$$Q \cdot \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = M'_f \cdot P \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

即

$$\begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix} = Q^{-1} M'_f \cdot P \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

将上式与式 (2.1.3) 对比即得 $M'_f = Q^{-1} M_f P$。 □

2.2 线性算子代数

本节的内容以具体例子为主，起着承上启下的作用，希望读者认真体会。

定义 2.2.1 (代数,algebra). 设 $\mathbb{K}$ 是域，如果环 A(乘法记为为 $\circ$) 满足：

(1) A 是域 $\mathbb{K}$ 上的向量空间；

(2) 任取 $a, b \in A$ 及 $k \in \mathbb{K}$，有 $k(a \circ b) = (ka) \circ b = a \circ (kb)$ (即数乘关于 A 的环乘法满足结合律)。

则我们称环 A 是 $\mathbb{K}$ 上的代数 (algebra)。如果 A 是交换环，则称 A 是一个交换的代数，否则称为非交换的代数。

实际上，代数的定义可以将条件放宽到 $\mathbb{K}$ 是交换幺环，A 是 $\mathbb{K}$ 上的左模。我们会在抽象代数课程中继续深入学习。

例 2.2.1. 设 $\mathbb{K}$ 是域，显然 $M_n(\mathbb{K})$ 和 $\mathbb{K}[x_1, \ldots, x_n]$ 在通常的加法、数乘和乘法下构成代数，且前者是非交换的代数，后者是交换的代数。

下面我们考虑一类特殊的线性映射。设 V 是域 $\mathbb{K}$ 上的向量空间，$f, g \in \operatorname{Hom}(V, V)$，则在 $\operatorname{Hom}(V, V)$ 上有一个自然的乘法，即映射复合。容易验证 $\operatorname{Hom}(V, V)$(也记作 $\operatorname{End}(V)$，即 endomorphism，自同态，或记作 $\mathcal{L}(V)$) 在线性映射的加法、数乘和乘法下构成了 $\mathbb{K}$ 上的一个代数。我们称 $\operatorname{End}(V)$ 为 V 上的线性算子代数，其中的元素也称为**线性算子** (linear operator) 或**线性变换** (linear transformation)。我们把 $\operatorname{End}(V)$ 作为向量空间的维数称为代数 $\operatorname{End}(V)$ 的维数。设 $\dim(V) = n$，则由推论2.1.2，显然 $\operatorname{End}(V)$ 的维数是 n^2。我们通常用花体的大写字母 $\mathcal{A}, \mathcal{B}, \ldots$ 来记线性算子，并且将算子 $\mathcal{A}$ 作用到 $\mathbf{x}$ 上得到的向量记作 $\mathcal{A}\mathbf{x}$。

定义 2.2.2. 设 A, B 是域 $\mathbb{K}$ 上的两个代数 (乘法分别记为 $\cdot$ 和 $\circ$)，如果存在 $\varphi : A \to B$ 是 A, B 作为向量空间的同构，并且任取 $a, b \in A$，都满足 $\varphi(a \cdot b) = \varphi(a) \circ \varphi(b)$，则我们称 A 和 B 作为代数是同构的 (isomorphic)，φ 是同构映射 (isomorphism)。

设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，则取定 V 的基底 $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 后 V 上的线性算子 $\mathcal{A} \in \operatorname{End}(V)$ 唯一对应了矩阵 $A \in M_n(\mathbb{K})$(定义2.1.3)，我们称 A 为线性算子在此基底下的矩阵。

定理 2.2.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，则 $\operatorname{End}(V)$ 与 $M_n(\mathbb{K})$ 代数同构，并且定理2.1.2(1) 中的映射 Φ 是同构映射。

利用定理2.1.2和定理2.1.3立刻可证。特别地，容易验证 $\operatorname{End}(V)$ 所有可逆线性算子在 $\operatorname{End}(V)$ 的乘法 (即复合) 下构成一个群，记作 $\operatorname{GL}(V)$。我们很容易验证，$\operatorname{GL}(V)$ 与一般线性群 $\operatorname{GL}_n(\mathbb{K})$ 是同构的，同构映射仍然是上面的 Φ。

下面我们开始研究线性算子在不同基底下的矩阵之间的关系。

定理 2.2.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 和 $\left(\mathbf{v}_1', \ldots, \mathbf{v}_n'\right)$ 是 V 的两组基，$\mathcal{A} \in \operatorname{End}(V)$ 且 $\mathcal{A}$ 在基底 $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ 和 $\left(\mathbf{v}_1', \ldots, \mathbf{v}_n'\right)$ 下的矩阵分别为 A 和 A'，$\left(\mathbf{v}_1', \ldots, \mathbf{v}_n'\right) = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \cdot P$，则 $A' = P^{-1}AP$。

这是定理2.1.4的直接推论。

于是我们有以下定义：

定义 2.2.3. 设 $\mathbb{K}$ 是域，$A, A' \in M_n(\mathbb{K})$，如果存在可逆矩阵 $P \in M_n(\mathbb{K})$ 使得 $A' = P^{-1}AP$，则我们称 A' 和 A 相似 (similar)。容易验证矩阵的相似是一个等价关系 (留作练习)。此外，我们称 $M_n(\mathbb{K})$ 上的变换 $A \mapsto P^{-1}AP$ 为由可逆矩阵 P 确定的相似变换。

定理2.2.2也即：同一个线性算子在不同基底下的矩阵是相似的。

矩阵的相似具有非常好的性质。例如，首先我们有：

命题 2.2.1. 设 $A, A' \in M_n(\mathbb{K})$ 并且存在可逆矩阵 $P \in M_n(\mathbb{K})$ 使得 $A' = P^{-1}AP$，则 $\det(A) = \det(A')$，$\operatorname{tr}(A) = \operatorname{tr}(A')$(tr 的定义见于例1.5.1)。

证明. $\det(A') = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \det(P^{-1}P)\det(A) = \det(E_n)\det(A) = \det(A)$。

我们先证明 tr 满足如下性质：$\forall A, B \in M_n(\mathbb{K})$，$\operatorname{tr}(AB) = \operatorname{tr}(BA)$。设 $A = (a_{ij})_{n\times n}$，$B = (b_{ij})_{n\times n}$，则容易计算出：

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}a_{ji}\\ &= \operatorname{tr}(BA).\end{aligned}$$

现在可以证明 $\operatorname{tr}(A) = \operatorname{tr}(A')$ 了。利用上面的性质即得到 $\operatorname{tr}(A') = \operatorname{tr}(P^{-1}AP) = \operatorname{tr}(PP^{-1}A) = \operatorname{tr}(A)$。这样我们就完成了证明。 □

于是我们可以定义线性算子 $\mathcal{A} \in \operatorname{End}(V)$ 的**行列式**和**迹**就是它们在任意一组基下的矩阵的行列式和迹，仍记作 $\det(\mathcal{A})$ 和 $\operatorname{tr}(\mathcal{A})$。由上面的定理可知算子的行列式和迹与基底的选取无关。

利用矩阵的相似我们也可以方便地计算矩阵的方幂。设 $A' = P^{-1}AP$，则

$$(A')^k = \underbrace{(P^{-1}AP)\cdots(P^{-1}AP)}_{k\text{个}} = P^{-1}A^kP.$$

因此，为了计算 A'，我们希望将 A' 相似变换到一个比较简单的形式 (如对角形或分块对角形)。

下面我们先看一些线性算子的例子。

例 2.2.2. (1) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，则零算子 $O : V \to V, \mathbf{x} \mapsto \mathbf{0}$ 和恒等算子 $\mathrm{id} : V \to V, \mathbf{x} \mapsto \mathbf{x}$(id 也记为 $\mathcal{E}$) 都在 $\operatorname{End}(V)$ 中，取定一组基后，它们对应的矩阵分别为 $O_{n\times n}$ 和 E_n。

(2) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\lambda \in \mathbb{K}$，则伸缩算子 $\mathcal{A} : V \to V, \mathbf{x} \mapsto \lambda\mathbf{x}$ 在 $\operatorname{End}(V)$ 中，取定一组基后，它对应的矩阵为 λE_n。

(3) 设 $\mathbb{K}$ 是域，记 $\mathbb{K}[t]_n = \{f \in \mathbb{K}[t] \mid \deg(f) \leqslant n-1\}$，显然 $\mathbb{K}[t]_n$ 在通常的加法和数乘下是 $\mathbb{K}$ 上的 n 维向量空间，则形式微分算子 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 是 $\mathbb{K}[t]_n$ 上的线性算子，并且在基底

$(1, t, \ldots, t^{n-1})$ 下 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 对应矩阵

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \\ & & & & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

如果 $\mathrm{char}(\mathbb{K}) = 0$，则可取 $\mathbb{K}[t]_n$ 的另一组基底 $(1, t, \dfrac{t^2}{2}, \ldots, \dfrac{t^{n-1}}{(n-1)!})$，在此基底下 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 对应矩阵

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

(4) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间并且 $V = U \oplus W$，则对任意的 $\mathbf{x} \in V$，可做分解 $\mathbf{x} = \mathbf{x}_U + \mathbf{x}_W$，其中 $\mathbf{x}_U \in U$，$\mathbf{x}_W \in W$，并且该分解是唯一的。于是我们可以作线性算子：

$$\begin{aligned} \mathcal{P} : & V \longrightarrow V \\ & \mathbf{x} \longmapsto \mathbf{x}_U. \end{aligned}$$

(验证 $\mathcal{P} \in \mathrm{End}(V)$ 是容易的，留作练习)。容易验证 $\mathcal{P}$ 满足性质：$\mathcal{P}^2 = \mathcal{P}$ ($\forall \mathbf{x} \in V$, $\mathcal{P}^2(\mathbf{x}) = \mathcal{P}(\mathbf{x}_U) = \mathbf{x}_U = \mathcal{P}(\mathbf{x})$)。

实际上，设 $\mathcal{A} \in \mathrm{End}(V)$ 且 $\mathcal{A}^2 = \mathcal{A}$(我们称这样的算子为**幂等算子** (idempotent operator))，则必有 $\mathcal{A}$ 是投影算子。证明只需利用直和分解 $V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A})$ 即可，后者的证明留作练习或者参考稍后的定理2.2.3。

显然取定一组基后，投影算子对应的矩阵是幂等矩阵 (讲义上册定义 2.5.4)。

(5) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A}(\neq O) \in \mathrm{End}(V)$ 且存在 $k \in \mathbb{Z}^+$ 使得 $\mathcal{A}^k = O$，则我们称 $\mathcal{A}$ 是**幂零算子** (nilpotent operator)。例如：$\mathcal{A} : \mathbb{R}^2 \to \mathbb{R}^2$，$\mathbf{x} \mapsto A\mathbf{x}$，$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 就是一个幂零算子。显然取定一组基后，幂零算子对应的矩阵是幂零矩阵。

(6) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，我们称算子 $\mathcal{A}(\neq O) \in \mathrm{End}(V)$ 是可逆算子，若 $\exists \mathcal{B} \in \mathrm{End}(V)$ 使得 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} = \mathcal{E}$。显然可逆算子在任意一组基底下对应的矩阵都是可逆矩阵。

线性算子作为特殊的线性映射，我们显然会考虑其核和像。设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则由对偶定理，$n = \dim(\ker(\mathcal{A})) + \dim(\mathrm{im}(\mathcal{A}))$。这看起来很像是直和分解：$V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A})$。但这个直和分解实际上并不总成立，因为和空间 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 不总是

等于 V。例如，取 $\mathcal{A} \in \mathrm{End}(\mathbb{R}^2) : (x_1, x_2)^t \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot (x_1, x_2)^t$，则 $\ker(\mathcal{A}) = \mathrm{span}\ \{(1,0)^t\}$，$\mathrm{im}(\mathcal{A}) = \mathrm{span}\ \{(1,0)^t\}$，显然 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 不是直和并且 $\mathbb{R}^2 \neq \ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$。事实上，我们有下面的定理：

定理 2.2.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A}) \iff \mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{A}^2)$(线性算子的秩即线性映射的秩，定义2.1.4)。

证明. ($\Longleftarrow$) 不妨设 $\mathrm{rank}(\mathcal{A}) = n - d$，则由对偶定理，$\dim(\ker(\mathcal{A})) = d$。于是可设 $\ker(\mathcal{A})$ 有一组基 $\mathbf{x}_1, \ldots, \mathbf{x}_d$，$\mathrm{im}(\mathcal{A})$ 有一组基 $\mathcal{A}\mathbf{x}_{d+1}, \ldots, \mathcal{A}\mathbf{x}_n$ 是 $\mathrm{im}(\mathcal{A})$ 的一组基 (思考 $\mathrm{im}(\mathcal{A})$ 的基为什么可以这样设？)，我们来证明 $\mathbf{x}_1, \ldots, \mathbf{x}_d, \mathcal{A}\mathbf{x}_{d+1}, \ldots, \mathcal{A}\mathbf{x}_n$ 是 V 的一组基。注意到它们是 V 中的 n 个向量，由于 $\dim(V) = n$，故我们只需证明它们线性无关即可证明它们是一组基。为此，设 $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$，作线性组合

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_d\mathbf{x}_d + \alpha_{d+1}\mathcal{A}\mathbf{x}_{d+1} + \cdots + \alpha_n\mathcal{A}\mathbf{x}_n = \mathbf{0}. \tag{2.2.1}$$

我们只需证明所有的组合系数 $\alpha_1, \ldots, \alpha_n$ 都是 0。上式两边同时用算子 $\mathcal{A}$ 作用得

$$\mathcal{A}^2(\alpha_{d+1}\mathbf{x}_{d+1} + \cdots + \alpha_n\mathbf{x}_n) = \mathbf{0},$$

(利用了线性性质及 $\mathcal{A}\mathbf{x}_1, \ldots, \mathcal{A}\mathbf{x}_d$ 都为 0)，即 $\alpha_{d+1}\mathbf{x}_{d+1} + \cdots + \alpha_n\mathbf{x}_n \in \ker(\mathcal{A}^2)$。而注意到 $\mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{A}^2)$，由对偶定理，$\dim(\ker(\mathcal{A})) = \dim(\ker(\mathcal{A}^2))$，然而 $\ker(\mathcal{A}) \subset \ker(\mathcal{A}^2)$(因为 $\mathcal{A}\mathbf{x} = \mathbf{0} \Rightarrow \mathcal{A}^2\mathbf{x} = \mathbf{0}$)，故由包含定理 (定理1.3.2(3)) 即得 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$。所以 $\alpha_{d+1}\mathbf{x}_{d+1} + \cdots + \alpha_n\mathbf{x}_n \in \ker(\mathcal{A})$，即

$$\alpha_{d+1}\mathcal{A}\mathbf{x}_{d+1} + \cdots + \alpha_n\mathcal{A}\mathbf{x}_n = \mathbf{0}.$$

而 $\mathcal{A}\mathbf{x}_{d+1}, \ldots, \mathcal{A}\mathbf{x}_n$ 是 $\mathrm{im}(\mathcal{A})$ 的一组基，故 $\alpha_{d+1}, \ldots, \alpha_n$ 全为 0。将 $\alpha_{d+1} = \cdots = \alpha_n = 0$ 代回式 (2.2.1) 即得

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_d\mathbf{x}_d = \mathbf{0}.$$

由 $\mathbf{x}_1, \ldots, \mathbf{x}_d$ 是 $\ker(\mathcal{A})$ 的一组基即得 $\alpha_1 = \cdots = \alpha_d = 0$。综上所述，$\mathbf{x}_1, \ldots, \mathbf{x}_d, \mathcal{A}\mathbf{x}_{d+1}, \ldots, \mathcal{A}\mathbf{x}_n$ 线性无关，构成了 V 的一组基，故 $V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A})$。

($\Longrightarrow$) 由于 $\mathrm{im}(\mathcal{A}^2) \subset \mathrm{im}(\mathcal{A})$，故 $\mathrm{rank}(\mathcal{A}^2) \leqslant \mathrm{rank}(\mathcal{A})$。用反证法。如果 $\mathrm{rank}(\mathcal{A}^2) \neq \mathrm{rank}(\mathcal{A})$，即 $\mathrm{rank}(\mathcal{A}^2) < \mathrm{rank}(\mathcal{A})$，则由对偶定理，$\dim(\ker(\mathcal{A}^2)) > \dim(\ker(\mathcal{A}))$。于是按照包含定理，$\ker(\mathcal{A}) \subsetneq \ker(\mathcal{A}^2)$，即存在 $\mathbf{x} \in \ker(\mathcal{A}^2) \setminus \ker(\mathcal{A})$，也即 $\mathcal{A}\mathbf{x} \neq \mathbf{0}$ 而 $\mathcal{A}^2\mathbf{x} = \mathbf{0}$。令 $\mathbf{y} = \mathcal{A}\mathbf{x}$，则 $\mathbf{y} \neq \mathbf{0}$ 而 $\mathcal{A}\mathbf{y} = \mathbf{0}$，即 $\mathbf{0} \neq \mathbf{y} \in \ker(\mathcal{A}) \cap \mathrm{im}(\mathcal{A})$，这说明 $\ker(\mathcal{A}) + \mathrm{im}(\mathcal{A})$ 不是直和，与条件矛盾！

这样我们就完成了证明。 □

利用上面的定理很容易证明：幂等算子一定是投影算子。

定义 2.2.4. 设 A 是域 $\mathbb{K}$ 上的代数，$B \subset A$。如果 B 在 A 的运算下也成为一个 $\mathbb{K}$ 上的代数，则称 B 是 A 的子代数 (subalgebra)。

除了平凡的子代数 (0 和本身) 之外，最简单的子代数是由一个元素生成的子代数。回到 $\mathrm{End}(V)$ 上讨论，设非零算子 $\mathcal{A} \in \mathrm{End}(V)$，令

$$\mathbb{K}[\mathcal{A}] = \{\sum_{i=0}^{s} \alpha_i\mathcal{A}^i \mid \forall s \in \mathbb{N},\ \alpha_0, \ldots, \alpha_s \in \mathbb{K}\}. \text{(注：以后始终约定}\mathcal{A}^0 = \mathcal{E}\text{)}$$

则容易验证 $\mathbb{K}[\mathscr{A}]$ 是 $\mathrm{End}(V)$ 的子代数，并且是包含算子 $\mathscr{A}$ 的最小的子代数。此外，显然 $\mathbb{K}[\mathscr{A}]$ 是交换的代数 (注意 $\mathrm{End}(V)$ 一般不是交换的代数)。我们可以看出，$\mathbb{K}[\mathscr{A}]$ 实际上是把算子“代入”到多项式中得到的，即我们有赋值同态：

$$\varphi_{\mathscr{A}}:\mathbb{K}[t]\to\mathbb{K}[\mathscr{A}],\ f(t)=\sum_{i=0}^{s}\alpha_i t^i\mapsto\sum_{i=0}^{s}\alpha_i\mathscr{A}^i.$$

我们把 $\varphi_{\mathscr{A}}(f)$ 简记作 $f(\mathscr{A})$。

下面我们考虑 $\mathbb{K}[\mathscr{A}]$ 的维数。首先，由于 $\mathrm{End}(V)$ 是 n^2 维的，故 $\dim(\mathbb{K}[\mathscr{A}])\leqslant n^2$。也就是说，一定存在非零多项式 $f(t)\in\mathbb{K}[t]$，$\deg(f)\leqslant n^2$ 使得 $f(\mathscr{A})=O$。我们称这样的 f 为 $\mathscr{A}$ 的**零化多项式**，其中次数最低的首一多项式称为 $\mathscr{A}$ 的极小多项式 (minimal polynomial)。我们有下面的定理：

定理 2.2.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A}\in\mathrm{End}(V)$，则存在唯一的极小多项式 $\mu_{\mathscr{A}}(t)=t^m+u_{m-1}t^{m-1}+\cdots+u_1t+u_0\in\mathbb{K}[t]$，使得 $\mu_{\mathscr{A}}(\mathscr{A})=O$，并且任取 $\mathscr{A}$ 的零化多项式 $q(t)$，都有 $\mu_{\mathscr{A}}(t)\mid q(t)$。

证明. 令

$$S=\{f\in\mathbb{K}[t]\mid f(A)=O\text{且}f\text{首一}\}.$$

由上面的讨论，显然 S 非空，故极小多项式 $\mu_{\mathscr{A}}(t)$ 的存在性是显然的。下证 $\mu_{\mathscr{A}}(t)$ 一定整除任意零化多项式 $q(t)$。任取 $q(t)$ 满足 $q(\mathscr{A})=O$，则由 $\mu_{\mathscr{A}}(t)$ 的取法可知一定有 $\deg(\mu_{\mathscr{A}}(t))\leqslant\deg(q(t))$，故可以作带余除法：存在唯一的 $a(t),r(t)$ 使得 $q(t)=a(t)\mu_{\mathscr{A}}(t)+r(t)$，其中 $r(t)=0$ 或 $\deg(r(t))<\deg(\mu_{\mathscr{A}}(t))$。将 $\mathscr{A}$ 代入 $q(t)$ 中可知 $r(\mathscr{A})=O$，这说明只能取 $r(t)=0$(否则 r 是次数比极小多项式更低的零化多项式，矛盾！)，即 $q(t)=a(t)\mu_{\mathscr{A}}(t)$，$\mu_{\mathscr{A}}(t)\mid q(t)$。

由此我们很容易证明 $\mu_{\mathscr{A}}(t)$ 的唯一性。设 $\mu(t),\mu'(t)$ 都是极小多项式，则 $\mu(t)\mid\mu'(t)$，$\mu'(t)\mid\mu(t)$，于是 $\mu(t)=\mu'(t)$(注意首一)。这样我们就完成了证明。 □

推论 2.2.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，则 $\mathscr{A}\in\mathrm{End}(V)$ 可逆 $\Longleftrightarrow$ $\mu_{\mathscr{A}}(t)$ 的常数项不为 0。

证明是显然的，留作练习。

例 2.2.3. 幂零算子的极小多项式一定形如 t^k，$k\in\mathbb{Z}^+$(我们把幂零算子极小多项式的次数称为**幂零指数**)；幂等算子的极小多项式一定是 t^2-t。证明是容易的，留作练习。

2.3 不变子空间与特征问题

从本节开始，我们开始将向量空间按照某一个线性算子的特征进行分解，对应地我们研究如何将矩阵进行对角化。

定义 2.3.1 (不变子空间). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，如果一个子空间 U 满足 $\mathcal{A}(U) \subset U$，则我们称 U 是算子 $\mathcal{A}$ 的一个不变子空间 (invariant subspace)，也简记为 U 是 $\mathcal{A}-$ 子空间。

例 2.3.1. 显然 $\{\mathbf{0}\}$ 和 V 是 $\mathcal{A}$ 的平凡的不变子空间；容易验证 $\ker(\mathcal{A})$ 和 $\mathrm{im}(\mathcal{A})$ 都是 $\mathcal{A}$ 的不变子空间，留作练习。

定理 2.3.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，$U$ 是 $\mathcal{A}-$ 子空间，$\mathbf{e}_1, \dots, \mathbf{e}_d$ 是 U 的一组基，将其扩充成 V 的一组基 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$，则算子 $\mathcal{A}$ 在基底 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 下的矩阵 A 一定形如

$$\begin{pmatrix} B_{d\times d} & C_{d\times(n-d)} \\ O_{(n-d)\times d} & D_{(n-d)\times(n-d)} \end{pmatrix}.$$

证明. 对任意的 $j \in \{1, \dots, d\}$，由于 U 是 $\mathcal{A}-$ 子空间，故 $\mathcal{A}\mathbf{e}_j \in U$，即存在 $b_{11}, \dots, b_{d1}, \cdots\cdots, b_{1d}, \dots, b_{dd} \in \mathbb{K}$ 使得

$$\mathcal{A}\mathbf{e}_j = \sum_{i=1}^{d} b_{ij}\mathbf{e}_j.$$

而对 $\forall k \in \{d+1, \dots, n\}$，由 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的基可知存在 $\alpha_{1,d+1}, \dots, \alpha_{n,d+1}, \cdots\cdots, \alpha_{1n}, \dots, \alpha_{nn} \in \mathbb{K}$ 使得

$$\mathcal{A}\mathbf{e}_k = \sum_{i=1}^{n} \alpha_{ik}\mathbf{e}_k.$$

将上面的式子改写成矩阵形式即是

$$(\mathcal{A}\mathbf{e}_1, \dots, \mathcal{A}\mathbf{e}_d, \mathcal{A}\mathbf{e}_{d+1}, \dots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n)\begin{pmatrix} b_{11} & \cdots & b_{1d} & \alpha_{1,d+1} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{d1} & \cdots & b_{dd} & \alpha_{d,d+1} & \cdots & \alpha_{dn} \\ 0 & \cdots & 0 & \alpha_{d+1,d+1} & \cdots & \alpha_{d+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \alpha_{n,d+1} & \cdots & \alpha_{nn} \end{pmatrix}$$

这样我们就完成了证明。 □

我们把 $\mathcal{A}$ 限制到不变子空间 U 上得到的算子记为 $\mathcal{A}\big|_U$。若 $\mathcal{A}$ 在上面定理中的基底下的矩阵为 A，则 $\mathcal{A}\big|_U$ 在此基底下的矩阵是定理中的 B，我们记 $B = A_U$。

推论 2.3.1. 条件同定理2.3.1。如果我们还有 $W = \mathrm{span}\,\{\mathbf{e}_{d+1}, \dots, \mathbf{e}_n\}$ 也是 $\mathcal{A}-$ 子空间，则 $\mathcal{A}$ 在基底 $\mathbf{e}_1, \dots, \mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 下的矩阵 A 一定是分块对角形：

$$\begin{pmatrix} A_U & O_{d\times(n-d)} \\ O_{(n-d)\times d} & A_W \end{pmatrix}.$$

更一般的，如果 V 是 m 个 $\mathcal{A}-$ 子空间 $U_1, \dots, U_m$ 的直和，将每个 U_i 的基底合在一起构成

V 的一组基，则算子 $\mathcal{A}$ 在此基底下的矩阵 A 有分块对角形式

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix},$$

其中 A_i 是 $\mathcal{A}\big|_{U_i}$ 的矩阵。特别地，如果每个 U_i 都是一维的，则矩阵 A 是对角阵。

证明是容易的，留作练习。

例 2.3.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$ 是投影算子，$\mathrm{rank}(\mathcal{A}) = d$，则由定理2.2.3，我们有直和分解 $V = \ker(\mathcal{A}) \oplus \mathrm{im}(\mathcal{A})$，由于 $\ker(\mathcal{A})$ 和 $\mathrm{im}(\mathcal{A})$ 是 $\mathcal{A}-$ 子空间，$\mathcal{A}\big|_{\ker(\mathcal{A})}$ 是零算子，而如果取 $\mathrm{im}(\mathcal{A})$ 的一组基为 $\mathcal{A}\mathbf{e}_1, \dots, \mathcal{A}\mathbf{e}_d$(思考这样取的理由，我们已经在定理2.2.3的证明中遇到过一次了)，则由 $\mathcal{A}^2 = \mathcal{A}$ 可知 $\mathcal{A}\big|_{\mathrm{im}(\mathcal{A})}$ 在基底 $\mathcal{A}\mathbf{e}_1, \dots, \mathcal{A}\mathbf{e}_d$ 下的矩阵为 E_d。于是由推论2.3.1，$\mathcal{A}$ 在基底 $\mathcal{A}\mathbf{e}_1, \dots, \mathcal{A}\mathbf{e}_d, \mathbf{e}_{d+1}, \dots, \mathbf{e}_n$($\mathbf{e}_{d+1}, \dots, \mathbf{e}_n$ 是 $\ker(\mathcal{A})$ 的一组基) 下的矩阵为

$$\begin{pmatrix} E_d & O \\ O & O \end{pmatrix}$$

即 $\mathrm{tr}(\mathcal{A}) = d = \mathrm{rank}(\mathcal{A})$。这是投影矩阵的重要性质。

一个线性算子不一定存在非平凡的不变子空间 (存在性与域本身有关)。例如，我们取 $\mathrm{End}(\mathbb{R}^2)$ 中的线性算子 $\mathcal{A}$ 如下：

$$\mathcal{A} : \mathbb{R}^2 \to \mathbb{R}^2,\ (x_1, x_2)^t \mapsto A \cdot (x_1, x_2)^t,\ \text{其中} A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},\ \alpha \in \mathbb{R} \text{且} 0 < \alpha < \pi.$$

则如果 U 是非平凡的 $\mathcal{A}-$ 子空间，则必有 $\dim(U) = 1$，于是取 U 中的一个非零元 $\mathbf{x}$，必有 $U = \mathrm{span}\,\{\mathbf{x}\}$。于是 $\exists \lambda \in \mathbb{R}$ 使得 $\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$，这说明方程组 $(A - \lambda E)(x_1, x_2)^t = \mathbf{0}$ 有非零解，即 $\det(A - \lambda E) = 0$。然而，$\det(A - \lambda E) = \lambda^2 - 2\cos\alpha \cdot \lambda + 1 = 0$ 在 $0 < \alpha < \pi$ 时没有实数根，矛盾！这说明算子 $\mathcal{A}$ 没有非平凡的不变子空间。

我们看到，上面的例子里多项式方程在 $\mathbb{R}$ 中无根是造成非平凡不变子空间不存在的原因。为此，我们需要代数闭域的概念。

定义 2.3.2. 设 $\mathbb{K}$ 是域，如果任取 $f(x) \in \mathbb{K}[x] \setminus \mathbb{K}$，$f(x)$ 在 $\mathbb{K}$ 上都至少有一个根，则称 $\mathbb{K}$ 是代数闭的 (algebraic closed)。

例如，复数域 $\mathbb{C}$ 是代数闭的，但实数域 $\mathbb{R}$ 不是。利用代数闭域我们可以得到如下结论：

定理 2.3.2. 设域 $\mathbb{K}$ 是代数闭的，V 是 $\mathbb{K}$ 上的 n 维向量空间，则任取非零线性算子 $\mathcal{A} \in \mathrm{End}(V)$，都有一个 1 维的 $\mathcal{A}-$ 子空间。

证明. 设 $\mu_{\mathcal{A}}(x)$ 是 $\mathcal{A}$ 的极小多项式，则显然 $\deg(\mu_{\mathcal{A}}(x)) \geqslant 1$，于是由 $\mathbb{K}$ 是代数闭域可知 $\mu_{\mathcal{A}}(x)$ 在 $\mathbb{K}$ 上至少有一个根，设其为 $\lambda \in \mathbb{K}$，则 $\mu_{\mathcal{A}}(x) = (x - \lambda)g(x)$，其中 $g(x) \in \mathbb{K}[x]$ 且 $\deg(g(x)) < \deg(\mu_{\mathcal{A}}(x))$。由 $\mu_{\mathcal{A}}(x)$ 是极小多项式可知 $g(\mathcal{A})$ 一定不是零算子，故存在 $\mathbf{v} \in V$ 使得 $g(\mathcal{A})(\mathbf{v}) \neq \mathbf{0}$。令 $\mathbf{u} = g(\mathcal{A})(\mathbf{v})$，下面我们证明 $\mathrm{span}\,\{\mathbf{u}\}$ 是 $\mathcal{A}-$ 子空间。注意到

$$\begin{aligned} \mathcal{A}\mathbf{u} &= (\mathcal{A} - \lambda\mathcal{E} + \lambda\mathcal{E})\mathbf{u} = (\mathcal{A} - \lambda\mathcal{E})g(\mathcal{A})(\mathbf{v}) + \lambda\mathbf{u} \\ &= \mu_{\mathcal{A}}(\mathcal{A})(\mathbf{v}) + \lambda\mathbf{u} = \mathbf{0} + \lambda\mathbf{u} = \lambda\mathbf{u}. \end{aligned}$$

于是对任意的 $\mathbf{y} = \alpha\mathbf{u} \in \operatorname{span}\{\mathbf{u}\}$，都有 $\mathscr{A}\mathbf{y} = \alpha\lambda\mathbf{u} \in \operatorname{span}\{\mathbf{u}\}$，此即 $\operatorname{span}\{\mathbf{u}\}$ 是 $\mathscr{A}-$ 子空间。 □

上面的证明过程启发我们考虑这样一个问题：对线性算子 $\mathscr{A} \in \operatorname{End}(V)$ 而言，是否存在 $\lambda \in \mathbb{K}$ 及 $\mathbf{v} \in V$ 使得 $\mathscr{A}\mathbf{v} = \lambda\mathbf{v}$ 成立呢？这就是我们下面的定义。

定义 2.3.3 (特征问题). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \operatorname{End}(V)$，如果存在 $\lambda \in \mathbb{K}$ 及 $\mathbf{v} \in V$ 使得 $\mathscr{A}\mathbf{v} = \lambda\mathbf{v}$，则称 λ 是算子 $\mathscr{A}$ 的特征值 (eigenvalue or characteristic value or proper value)，$\mathbf{v}$ 是特征值 λ 对应的一个特征向量 (eigenvector)。我们把算子 $\mathscr{A}$ 的所有特征值所构成的的集合称为 $\mathscr{A}$ 的谱 (spectrum)，记作 $\operatorname{spec}(\mathscr{A})$。即

$$\operatorname{spec}(\mathscr{A}) = \{\lambda \in \mathbb{K} \mid \exists \mathbf{v} \in V\text{使得}\mathscr{A}\mathbf{v} = \lambda\mathbf{v}\}.$$

显然如果 $\mathbf{v}$ 是特征值 λ 对应的一个特征向量，则 $\operatorname{span}\{\mathbf{v}\}$ 是 $\mathscr{A}$ 的一维不变子空间。但是，$\operatorname{span}\{\mathbf{v}\}$ 并不一定是全部的 λ 对应的特征向量[1]。容易证明 (验证细节留作练习)：对算子 $\mathscr{A}$ 的每个固定特征值 λ，

$$V^{\lambda} = \{\mathbf{v} \in V \mid \mathscr{A}\mathbf{v} = \lambda\mathbf{v}\} \subset V$$

是 V 的子空间，我们称 V^{λ} 是 V 的特征子空间 (eigenspace)。容易验证 V^{λ} 是 $\mathscr{A}-$ 子空间 (练习)。此外，我们称 $\dim(V^{\lambda})$ 为特征值 λ 的几何重数 (geometric multiplicity)。

将 $M_n(\mathbb{K})$ 中的矩阵视作 $\operatorname{End}(\mathbb{K}^n)$ 中的算子，我们可以同样地定义方阵的特征值、特征向量与特征多项式，完整的叙述留给读者补充。

定理 2.3.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \operatorname{End}(V)$。则 $\lambda \in \mathbb{K}$ 是 $\mathscr{A}$ 的特征值 $\iff \det(\lambda\mathcal{E} - \mathscr{A}) = 0$。

证明. $\lambda \in \mathbb{K}$ 是 $\mathscr{A}$ 的特征值 $\iff \exists \mathbf{v} \in V$ 使得 $(\lambda\mathcal{E} - \mathscr{A})\mathbf{v} = \mathbf{0} \iff \ker(\lambda\mathcal{E} - \mathscr{A}) \neq \{\mathbf{0}\} \iff \operatorname{rank}(\lambda\mathcal{E} - \mathscr{A}) < n \iff \det(\lambda\mathcal{E} - \mathscr{A}) = 0$。 □

定义 2.3.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \operatorname{End}(V)$，$\mathscr{A}$ 在取定的一组基下对应矩阵 A，则我们称多项式 $\chi_{\mathscr{A}}(t) = \det(tE - A) \in \mathbb{K}[t]$ 为算子 $\mathscr{A}$ 的**特征多项式** (characteristic polynomial)。

命题 2.3.1. 相似的矩阵具有相同的特征多项式。

由于不同基底下同一个算子对应的矩阵是相似的，因此特征多项式与基底的选取无关。

显然 $\deg_t(\chi_{\mathscr{A}}(t)) = n$。有了特征多项式的定义以后，定理2.3.3也可以表述为：$\lambda \in \mathbb{K}$ 是 $\mathscr{A}$ 的特征值 $\iff$ λ 是特征多项式 $\chi_{\mathscr{A}}(t)$ 的一个根。下面我们考虑特征多项式的系数。设 $\mathscr{A} \in \operatorname{End}(V)$ 在某组基下的矩阵为 $A = (a_{ij})_{n\times n} \in M_n(\mathbb{K})$，则利用行列式函数的求导法则 (或者直接展开) 可以得到：

$$\chi_{\mathscr{A}}(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

其中 $c_k, 1 \leqslant k \leqslant n$ 是 A 的所有 $n-k$ 阶主子式之和乘以 $(-1)^{n-k}$，其证明方法我们在定理1.6.12中已经用过一次了。特别地，$c_0 = (-1)^n \det(A)$，$c_{n-1} = -\operatorname{tr}(A)$。

---

[1] 例如在 $\mathbb{R}^n$ 中 1 是 id 的特征值，$\mathbb{R}^n$ 中的每个向量都是 1 对应的特征向量。

例 2.3.3. 设 $\mathbb{K}$ 是域，则 $\mathbb{K}[t]_n$ 上的形式微分算子 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 在基底 $(1, t, \dots, t^{n-1})$ 下对应的矩阵为

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \\ & & & & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

于是其特征多项式 $\chi(t) = \det(tE - A) = t^n$。于是算子 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 有且仅有特征值 0，并且容易验证 0 所对应的特征子空间为 span $\{1\}$(即 $\mathbb{K}$)。

由定理2.3.3，我们也可以给出下面的定义：

定义 2.3.5. 条件同定义2.3.3。设 $\chi_{\mathscr{A}}(t)$ 是算子 $\mathscr{A}$ 的特征多项式，如果 $\lambda \in \mathbb{K}$ 是 $\chi_{\mathscr{A}}(t)$ 的 k 重根，则我们称特征值 λ 的**代数重数** (algebraic multiplicity) 是 k。

命题 2.3.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \mathrm{End}(V)$，$\lambda$ 是 $\mathscr{A}$ 的特征值，则 λ 的几何重数小于等于其代数重数。

证明. 设 $m = \dim(V^\lambda)$ 是 λ 的几何重数，取 V^λ 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_m$，则每个 $\mathbf{v}_i$ 都满足 $\mathscr{A}\mathbf{v}_i = \lambda \mathbf{v}_i$。由基扩充定理，我们可以找到 V 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \dots, \mathbf{v}_n$，在这组基下算子 $\mathscr{A}$ 对应的矩阵一定有

$$A = \begin{pmatrix} \lambda E_m & B \\ O & D \end{pmatrix}$$

的形式 (定理2.3.1)。因此，算子 $\mathscr{A}$ 的特征多项式

$$\chi_{\mathscr{A}}(t) = \det(tE_n - A) = (t - \lambda)^m \det(tE_{n-m} - D)$$

即 λ 作为特征多项式的根的重数至少是 m。这样我们就完成了证明。 □

命题 2.3.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \mathrm{End}(V)$ 且 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$，其中 $V_1, \dots, V_m$ 是 $\mathscr{A}$– 子空间，令 $\mathscr{A}_i$ 是算子 $\mathscr{A}$ 在 V_i 上的限制，$\chi_i(t) \in \mathbb{K}[t]$, $(i = 1, \dots, m)$ 是 $\mathscr{A}_i$ 的特征多项式，则 $\chi_{\mathscr{A}}(t) = \chi_1(t) \cdots \chi_m(t)$。

证明. 不妨设 $\dim(V_i) = k_i$，$i \in \{1, \dots, m\}$。由推论2.3.1，我们可取 $V_1, \dots, V_m$ 的基放到一起，得到 V 的一组基，此时 $\mathscr{A}$ 在这组基底下的矩阵是分块对角阵

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix},$$

其中 A_1 是 $\mathscr{A}_i$ 对应的矩阵。那么

$$\chi_A(t) = \det(tE_n - A) = \det \begin{pmatrix} tE_{k_1} - A_1 & & \\ & \ddots & \\ & & tE_{k_m} - A_m \end{pmatrix} = \det(tE_{k_1} - A_1) \cdots \det(tE_{k_m} - A_m) = \chi_1(t) \cdots \chi_m(t).$$

这样我们就完成了证明。 □

命题 2.3.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，$\lambda_1, \ldots, \lambda_m$ 是 $\mathcal{A}$ 的互不相同的 m 个特征值，则任取 $\mathbf{v}_i \in V^{\lambda_i}$，$i = 1, \ldots, m$，则 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 线性无关。换言之，和空间 $V^{\lambda_1} + \cdots + V^{\lambda_m}$ 是直和 (但这个和空间不一定等于 V，稍后的定理会更清楚地表明这一点)。

证明. 对 m 用数学归纳法。$m = 1$ 时命题显然成立。下设 $m \geqslant 2$ 且命题对 $m - 1$ 个特征值的情形成立。则对 m 个特征值的情形，设线性组合

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m = \mathbf{0},\ a_i \in \mathbb{K}.$$

我们来证明组合系数 $a_1 = \cdots = a_m = 0$。用算子 $\mathcal{A}$ 作用到上式两边，得到

$$\mathcal{A}(a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m) = a_1\lambda_1\mathbf{v}_1 + a_2\lambda_2\mathbf{v}_2 + \cdots + a_m\lambda_m\mathbf{v}_m = \mathbf{0}$$

用第二个式子减去第一个式子的 λ_1 倍即得

$$a_2(\lambda_2 - \lambda_1)\mathbf{v}_2 + \cdots + a_m(\lambda_m - \lambda_1)\mathbf{v}_m = \mathbf{0}$$

由归纳假设可知 $a_2(\lambda_2 - \lambda_1) = \cdots = a_m(\lambda_m - \lambda_1) = 0$，而 $\lambda_2 - \lambda_1, \ldots, \lambda_m - \lambda_1$ 都不为 0，所以只能是 $a_2 = \cdots = a_m = 0$。再由第一个式子可知 $a_1\mathbf{v}_1 = \mathbf{0}$，即 $a_1 = 0$。这样我们就完成了归纳证明。

特别地，命题结论也表明 $\mathbf{0}$ 在和空间 $V^{\lambda_1} + \cdots + V^{\lambda_m}$ 中的表法是唯一的，于是由命题1.2.1可知 $V^{\lambda_1} + \cdots + V^{\lambda_m}$ 是直和。 □

下面我们讨论线性算子的矩阵何时是对角阵，这可以视作若尔当标准型理论的前置知识。

定义 2.3.6. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，如果存在 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

即算子 $\mathcal{A}$ 在此基底下的矩阵 A 为对角阵 (记作 $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$)，则称算子 $\mathcal{A}$ 是可对角化的 (diagonalizable)。

定理 2.3.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则算子 $\mathcal{A}$ 是可对角化的 $\Longleftrightarrow$ $\chi_{\mathcal{A}}(t)$ 的所有根都在 $\mathbb{K}$ 中并且 $\mathcal{A}$ 的每个特征值的几何重数都等于代数重数。

证明. ($\Longrightarrow$) 假设 $\mathscr{A}$ 可对角化，即存在一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得

$$(\mathscr{A}\mathbf{e}_1,\dots,\mathscr{A}\mathbf{e}_n)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\begin{pmatrix}\lambda_1&&&&&&\\&\ddots&&&&&\\&&\lambda_1&&&&\\&&&\ddots&&&\\&&&&\lambda_m&&\\&&&&&\ddots&\\&&&&&&\lambda_m\end{pmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\}n_1\\ \vdots\\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\}n_m\end{matrix}\tag{2.3.1}$$

其中 $n_1+\cdots+n_m=n$。则 $\chi_{\mathscr{A}}(t)=(t-\lambda_1)^{n_1}\cdots(t-\lambda_m)^{n_m}$，其中 $\lambda_1,\dots,\lambda_m\in\mathbb{K}$ 是 $\mathscr{A}$ 的全部互不相同的特征值。设 $\lambda_1,\dots,\lambda_m$ 所对应的特征子空间分别是 $V^{\lambda_1},\dots,V^{\lambda_m}$，则由命题2.3.2可知 $\dim(V^{\lambda_i})\leqslant n_i$，$\forall i\in\{1,\dots,m\}$。下面我们只需证明 $V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_m}$ ($\star$)，即可由 $n=\sum_{i=1}^m\dim(V^{\lambda_i})\leqslant\sum_{i=1}^m n_i=n$ 得到 $\dim(V^{\lambda_i})=n_i$，$\forall i$。等式 ($\star$) 右边是直和已经在命题2.3.4中证明过了，因此只需证明：$V=V^{\lambda_1}+\cdots+V^{\lambda_n}$。

注意到由式 (2.3.1) 可知 $\operatorname{span}\{\mathbf{e}_1,\dots,\mathbf{e}_{n_1}\}\subset V^{\lambda_1}$，……，$\operatorname{span}\{\mathbf{e}_{n_1+\cdots+n_{m-1}+1},\dots,\mathbf{e}_n\}\subset V^{\lambda_m}$。由于 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i\in V$，都有：

$$\mathbf{x}=\underbrace{\sum_{i=1}^{n_1}x_i\mathbf{e}_i}_{\in V^{\lambda_1}}+\cdots\cdots+\underbrace{\sum_{i=n_1+\cdots+n_{m-1}+1}^{n}\mathbf{e}_i}_{\in V^{\lambda_m}}$$

即 $V=V^{\lambda_1}+\cdots+V^{\lambda_n}$ 成立。这样我们就完成了 ($\Longrightarrow$) 方向的证明。

($\Longleftarrow$) 由条件不妨设 $\lambda_1,\dots,\lambda_m\in\mathbb{K}$ 是特征多项式 $\chi_{\mathscr{A}}(t)$ 的全部互不相同的根，并且 $\chi_{\mathscr{A}}(t)=(t-\lambda_1)^{l_1}\cdots(t-\lambda_m)^{l_m}$，其中 $l_1,\dots,l_m$ 分别是特征值 $\lambda_1,\dots,\lambda_m$ 的代数重数，并且 $\sum_{i=1}^m l_i=n$。由条件，每个 λ_i 的几何重数都等于代数重数，即 $l_i=\dim(V^{\lambda_i})$，即 $\sum_{i=1}^m\dim(V^{\lambda_i})=n$。由于和 $V^{\lambda_1}+\cdots+V^{\lambda_n}$ 是直和 (命题2.3.4)，故 $V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_m}$。于是取每个 V^{λ_i}，$i=1,\dots,m$ 的一组基合在一起构成 V 的一组基，即可证明在此基底下 $\mathscr{A}$ 对应对角矩阵 (细节留给读者补充)。这样 ($\Longleftarrow$) 方向的证明也完成了。 □

推论 2.3.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A}\in\mathrm{End}(V)$，如果特征多项式 $\chi_{\mathscr{A}}(t)$ 在 $\mathbb{K}$ 中有 n 个互不相同的根，则 $\mathscr{A}$ 是可对角化的。

这是定理2.3.4的直接推论。

例 2.3.4. 例2.3.3中的形式微分算子有且仅有特征值 0，其代数重数是 n，几何重数是 1，故不可对角化。

在学完若尔当标准型理论之后，我们会重新总结线性算子 (或方阵) 可以对角化 (相似于对角阵) 的充分必要条件。

2.4　商算子和对偶算子

2.4.1　商算子

我们在上一章已经学过了商空间，也知道商集合可以诱导商映射 (讲义上册定义 1.4.9)，那么，线性算子是否可以诱导出商空间上的商算子呢？实际上，这在一般情况下是做不到的，因为对于一般的子空间 $U \subset V$ 而言，$\mathbf{x} \in U \nRightarrow \mathcal{A}\mathbf{x} \in U$，即 $\overline{\mathcal{A}}(\mathbf{x}+U) \neq \overline{\mathcal{A}}(\mathbf{0}+U)$，这不是良定义的！因此，我们需要对子空间 U 加以限制。

定义 2.4.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，$U$ 是一个 $\mathcal{A}-$ 子空间，即 $\mathcal{A}(U) \subset U$，则在 V/U 上可以定义商算子 $\overline{\mathcal{A}}$ 如下：

$$\begin{aligned}\overline{\mathcal{A}}: V/U &\longrightarrow V/U\\ \mathbf{x}+U &\longmapsto (\mathcal{A}\mathbf{x})+U\end{aligned}$$

容易验证这是一个良定义 $(\mathbf{x}+U=\mathbf{y}+U \Longrightarrow \overline{\mathcal{A}}(\mathbf{x}+U)=\overline{\mathcal{A}}(\mathbf{y}+U))$ 的**线性**算子，验证过程留作练习。

定理 2.4.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，$U$ 是 $\mathcal{A}-$ 子空间，$\mathbf{e}_1,\ldots,\mathbf{e}_d$ 是 U 的一组基，将这组基扩充成 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$。再设 $\mathcal{A}\big|_U$ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_d$ 下的矩阵为 A_U，商算子 $\overline{\mathcal{A}}$ 在商空间的基底 $\mathbf{e}_{d+1}+U,\ldots,\mathbf{e}_n+U$ 下的矩阵为 B，则算子 $\mathcal{A}$ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下的矩阵 A 一定具有如下形式：

$$A=\begin{pmatrix}A_U & C\\ O & B\end{pmatrix}.$$

证明. 由定理2.3.1可知算子 $\mathcal{A}$ 在此基底下的矩阵一定形如

$$A=\begin{pmatrix}A_U & A_2\\ O & A_3\end{pmatrix}.$$

下面只需验证 $A_3=B$ 即可。任取 $j \in \{d+1,\ldots,n\}$，我们有

$$\begin{aligned}\mathcal{A}\mathbf{e}_j &= (\mathbf{e}_1,\ldots,\mathbf{e}_n)\begin{pmatrix}\mathbf{A}_2^{(j-d)}\\ \mathbf{A}_3^{(j-d)}\end{pmatrix}\\ &= (\mathbf{e}_1,\ldots,\mathbf{e}_d)\mathbf{A}_2^{(j-d)}+(\mathbf{e}_{d+1},\ldots,\mathbf{e}_n)\mathbf{A}_3^{(j-d)}\end{aligned}$$

显然 $(\mathbf{e}_1,\ldots,\mathbf{e}_d)\mathbf{A}_2^{(j-d)} \in U$，故

$$(\mathcal{A}\mathbf{e}_j)+U=(\mathbf{e}_{d+1}+U,\ldots,\mathbf{e}_n+U)\mathbf{A}_3^{(j-d)}$$

即 $\mathbf{B}^{(j-d)}=\mathbf{A}_3^{(j-d)}$，$\forall j \in \{d+1,\ldots,n\}$，所以 $A_3=B$。 □

推论 2.4.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，$U$ 是 $\mathcal{A}-$ 子空间，记 $\mathcal{A}\big|_U=\mathcal{B}$，商算子$\overline{\mathcal{A}}=C$，则它们的特征多项式满足 $\chi_{\mathcal{A}}(t)=\chi_{\mathcal{B}}(t)\cdot\chi_{C}(t)$。

证明是容易的，留作练习。

定理 2.4.2. 设 V 是**代数闭域** $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则存在一组 V 的基使得

$\mathcal{A}$ 在这组基下的矩阵是上三角矩阵。

证明. 对 n 用数学归纳法。$n = 1$ 时定理显然成立。下设 $n > 1$ 并且定理对所有 $n - 1$ 维向量空间都成立，我们需要证明定理对 n 维空间也成立。

由定理2.3.2，存在 $\mathcal{A}$ 的 1 维不变子空间 $U = \operatorname{span}\{\mathbf{e}_1\}$，则 V/U 是域 $\mathbb{K}$ 上的 $n - 1$ 维向量空间。由归纳假设，存在 V/U 的一组基 $\mathbf{e}_2 + U, \dots, \mathbf{e}_n + U$ 使得 V/U 上的商算子 $\overline{\mathcal{A}}$ 在此基底下对应上三角矩阵 B。那么，容易证明 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 是 V 的一组基，则由定理2.4.1，$\mathcal{A}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ 下的矩阵 A 形如

$$A = \begin{pmatrix} \lambda & C \\ O_{(n-1)\times 1} & B \end{pmatrix}$$

由于 B 是上三角的，故 A 也是上三角的。这样我们就完成了归纳证明。 □

商算子的一个重要应用就是证明下面的哈密顿-凯莱 (Hamilton-Cayley) 定理。当然，这个定理有许多证明方法，我们会在习题课讲义中给出一个纯矩阵版本的证明，以后还会有更高观点下的证明。

定理 2.4.3 (Hamilton-Cayley 定理). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \operatorname{End}(V)$，则特征多项式 $\chi_{\mathcal{A}}(t)$ 是 $\mathcal{A}$ 的零化多项式。

证明. 对 n 用数学归纳法。当 $n = 1$ 时，不妨设 $\mathbf{v}$ 是 V 的一组基，且 $\mathcal{A}\mathbf{v} = \lambda\mathbf{v}$, $\lambda \in \mathbb{K}$，则 $\mathcal{A}$ 在基底 $\mathbf{v}$ 下的矩阵为 (λ)。所以特征多项式 $\chi_{\mathcal{A}}(t) = t - \lambda$，于是代入算子 $\mathcal{A}$ 得 $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{A} - \lambda\mathcal{E}$。下面我们说明 $\mathcal{A} - \lambda\mathcal{E}$ 是零算子。任取 $\mathbf{x} = \alpha\mathbf{v} \in V$，则 $(\mathcal{A} - \lambda\mathcal{E})\mathbf{x} = \mathcal{A}\mathbf{x} - \lambda\mathbf{x} = \alpha\mathcal{A}\mathbf{v} - \alpha\lambda\mathbf{v} = \mathbf{0}$，此即 $n = 1$ 时定理成立。

下设 $n > 1$ 且定理对小于 n 维的向量空间都成立，我们来证明定理对 n 维空间成立。我们分如下两种情形讨论：

(1) 如果 V 有非平凡的 $\mathcal{A}-$ 子空间 U，则 $\dim(U) < n$, $\dim(V/U) < n$。记 $\mathcal{B} = \mathcal{A}\big|_U$, $\mathcal{C} = \overline{\mathcal{A}}$，于是由归纳假设，$\chi_{\mathcal{B}}(\mathcal{B}) = O$, $\chi_{\mathcal{C}}(\mathcal{C}) = O$。前者表明 $\forall \mathbf{u} \in U$, $\chi_{\mathcal{B}}(\mathcal{A})\mathbf{u} = \chi_{\mathcal{B}}(\mathcal{B})\mathbf{u} = \mathbf{0}$(第一个等号来源于 $\mathcal{B}$ 的定义)；后者表明任取 $\mathbf{x} \in V$，都有 $\chi_{\mathcal{C}}(\mathcal{C})(\mathbf{x} + U) = \mathbf{0} + U$，即 $\chi_{\mathcal{C}}(\mathcal{A})\mathbf{x} \in U$(从商空间回到原空间)。

因此，由推论2.4.1，$\chi_{\mathcal{A}}(t) = \chi_{\mathcal{B}}(t)\chi_{\mathcal{C}}(t)$，代入算子 $\mathcal{A}$ 得

$$\chi_{\mathcal{A}}(\mathcal{A}) = \chi_{\mathcal{B}}(\mathcal{A})\chi_{\mathcal{C}}(\mathcal{A}).$$

于是任取 $\mathbf{x} \in V$，都有

$$\chi_{\mathcal{A}}(\mathcal{A})\mathbf{x} = \chi_{\mathcal{B}}(\mathcal{A})\underbrace{\chi_{\mathcal{C}}(\mathcal{A})\mathbf{x}}_{\in U} = \mathbf{0} \qquad \text{(利用了归纳假设)}.$$

即此时 $\chi_{\mathcal{A}}(\mathcal{A})$ 是 V 上的零算子。

(2) 如果 $\mathcal{A}$ 没有非平凡的不变子空间，则任取非零向量 $\mathbf{v} \in V$，以下 n 个向量

$$\mathbf{v}, \mathcal{A}\mathbf{v}, \dots, \mathcal{A}^{n-1}\mathbf{v}$$

线性无关。事实上，若不然，则存在不全为 0 的 $a_0, \ldots, a_{n-1} \in \mathbb{K}$ 使得

$$a_0\mathbf{v} + a_1\mathcal{A}\mathbf{v} \cdots + a_{n-1}\mathcal{A}^{n-1}\mathbf{v} = \mathbf{0}.$$

那么，如果 $a_{n-1} \neq 0$，那么 $\mathcal{A}^{n-1}\mathbf{v} \in \operatorname{span}\{\mathbf{v}, \ldots, \mathcal{A}^{n-2}\mathbf{v}\}$；如果 $a_{n-1} = 0$ 而 $a_{n-2} \neq 0$，那么 $\mathcal{A}^{n-2}\mathbf{v} \in \operatorname{span}\{\mathbf{v}, \ldots, \mathcal{A}^{n-3}\mathbf{v}\}$；…… 这样进行下去，由于 $a_0, \ldots, a_{n-1} \in \mathbb{K}$ 不全为 0，我们一定能找到某个 $i \in \{1, \ldots, n-1\}$ 使得 $\mathcal{A}^i\mathbf{v} \in \operatorname{span}\{\mathbf{v}, \ldots, \mathcal{A}^{i-1}\mathbf{v}\}$，即 $\operatorname{span}\{\mathbf{v}, \ldots, \mathcal{A}^{i-1}\mathbf{v}\}$ 是 $\mathcal{A}$– 子空间，这与 $\mathcal{A}$ 没有非平凡的不变子空间矛盾！

因此，$\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{n-1}\mathbf{v}$ 是 V 的一组基。于是 $\mathcal{A}^n\mathbf{v}$ 可以用它们的线性组合表示。不妨设

$$\mathcal{A}^n\mathbf{v} = \sum_{k=0}^{n-1} a_k\mathcal{A}^k\mathbf{v},\ a_k \in \mathbb{K} \tag{2.4.1}$$

则容易看出 $\mathcal{A}$ 在这组基下的矩阵为

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{n-1} \end{pmatrix}$$

于是 $\chi_{\mathcal{A}}(t) = \det(tE - A) = t^n - a_{n-1}t^{n-1} - \cdots - a_1t - a_0$，于是 $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{A}^n - a_{n-1}\mathcal{A}^{n-1} - \cdots - a_1\mathcal{A} - a_0\mathcal{E}$，所以由式 (2.4.1) 可得

$$\chi_{\mathcal{A}}(\mathcal{A})\mathbf{v} = (\mathcal{A}^n - a_{n-1}\mathcal{A}^{n-1} - \cdots - a_1\mathcal{A} - a_0\mathcal{E})\mathbf{v} = \mathbf{0}.$$

于是对任意 $k \in \mathbb{N}$，我们有

$$\chi_{\mathcal{A}}(\mathcal{A})(\mathcal{A}^k\mathbf{v}) = \mathcal{A}^k\chi_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mathbf{0}.$$

这里利用了算子 $\mathcal{A}^k$ 与 $\chi_{\mathcal{A}}(\mathcal{A})$ 都在 $\mathbb{K}[\mathcal{A}]$ 中，从而可交换。所以，任取 $\mathbf{x} = \sum_{k=0}^{n-1} x_k\mathcal{A}^k\mathbf{v}$，都有

$$\begin{aligned} \chi_{\mathcal{A}}(\mathcal{A})\mathbf{x} &= \chi_{\mathcal{A}}(\mathcal{A})(\sum_{k=0}^{n-1} x_k\mathcal{A}^k\mathbf{v}) \\ &= \sum_{k=0}^{n-1} x_k\chi_{\mathcal{A}}(\mathcal{A})(\mathcal{A}^k\mathbf{v}) = \mathbf{0}. \end{aligned}$$

此即 $\chi_{\mathcal{A}}(\mathcal{A})$ 是零算子。

综合 (1),(2) 可知定理对 n 维空间成立。这样我们就完成了归纳证明。 □

推论 2.4.2. (1) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \operatorname{End}(V)$，则 $\mu_{\mathcal{A}}(t) \mid \chi_{\mathcal{A}}(t)$。
(2) 若 $\lambda \in \mathbb{K}$ 是 $\mathcal{A}$ 的特征值，则 $t - \lambda \mid \mu_{\mathcal{A}}(t)$。

证明. (1) 由哈密顿-凯莱定理和定理2.2.4可以直接得到 $\mu_{\mathcal{A}}(t) \mid \chi_{\mathcal{A}}(t)$。

(2) 若 $\lambda\in\mathbb{K}$ 是算子 $\mathcal{A}$ 的特征值，则 $\exists\mathbf{v}(\neq\mathbf{0})\in V$ 使得 $\mathcal{A}\mathbf{v}=\lambda\mathbf{v}$。做带余除法

$$\mu_{\mathcal{A}}(t)=q(t)(t-\lambda)+r,\ \ r\in\mathbb{K}.$$

将 $\mathcal{A}$ 代入上式并作用到向量 $\mathbf{v}$ 上可得

$$\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v}=q(\mathcal{A})(\mathcal{A}-\lambda\mathcal{E})\mathbf{v}+r\mathbf{v}=r\mathbf{v}.$$

另一方面，按极小多项式的定义，$\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v}=\mathbf{0}$，因此 $r=0$，即 $t-\lambda\mid\mu_{\mathcal{A}}(t)$。

□

由上面的推论我们看到，特征多项式和极小多项式具有相同的根，但每个根的重数可能不同。在学完若尔当标准型理论以后，我们会对特征多项式和极小多项式的本质有更深刻的理解。

定理 2.4.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$，如果 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{K}(t)$ 中可以分解成两个正次数多项式的乘积，则 V 有非平凡的 $\mathcal{A}-$ 子空间。

证明. 不妨设 $\chi_{\mathcal{A}}(t)=\xi(t)\eta(t)$，其中 $0<p=\deg(\xi(t))<n$，$0<q=\deg(\eta(t))<n$。任取 $\mathbf{v}(\neq\mathbf{0})\in V$，如果 $\eta(\mathcal{A})\mathbf{v}=\mathbf{0}$，则类似于哈密顿-凯莱定理的证明过程中的 (2)，容易验证 $\mathrm{span}\,\{\mathbf{v},\mathcal{A}\mathbf{v}\ldots,\mathcal{A}^{q-1}\mathbf{v}\}$ 是 $\mathcal{A}-$ 子空间；否则令 $\mathbf{u}=\eta(\mathcal{A})\mathbf{v}\neq\mathbf{0}$，同样我们可以验证 $\mathrm{span}\,\{\mathbf{u},\mathcal{A}\mathbf{u}\ldots,\mathcal{A}^{p-1}\mathbf{u}\}$ 是 $\mathcal{A}-$ 子空间。细节留给读者补充。 □

2.4.2 对偶算子

我们已经学过了对偶空间，那么，对偶空间上的线性算子与原空间上的线性算子又有什么关系呢？这就是我们本小节要讨论的问题。

定义 2.4.2 (对偶算子). 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$，定义如下映射：

$$\begin{aligned}\mathcal{A}^*\ :\ V^*&\longrightarrow V^*\\ f&\longmapsto f\circ\mathcal{A}\end{aligned}$$

首先，容易验证 $f\circ\mathcal{A}:\mathbf{x}\mapsto f(\mathcal{A}\mathbf{x})$ 是 V^* 中的线性函数；其次，容易验证 $\mathcal{A}^*$ 满足线性性质，即 $\mathcal{A}^*\in\mathrm{End}(V^*)$(验证留作练习)。我们称 $\mathcal{A}^*$ 是 $\mathcal{A}$ 的对偶算子或共轭算子 (dual operator)。

下面我们考察映射 $*:\mathrm{End}(V)\to\mathrm{End}(V^*)$，$\mathcal{A}\mapsto\mathcal{A}^*$。首先，容易验证 $*$ 是一个双射；其次，$*$ 保持线性运算：$\forall\mathcal{A},\mathcal{B}\in\mathrm{End}(V)$，$f\in V^*,\mathbf{x}\in V$，我们有

$$((\alpha\mathcal{A}+\beta\mathcal{B})^*(f))(\mathbf{x})=f((\alpha\mathcal{A}+\beta\mathcal{B})\mathbf{x})=\alpha f(\mathcal{A}\mathbf{x})+\beta f(\mathcal{B}\mathbf{x})=\alpha(\mathcal{A}^*f)(\mathbf{x})+\beta(\mathcal{B}^*f)\mathbf{x}$$

即 $(\alpha\mathcal{A}+\beta\mathcal{B})^*=\alpha\mathcal{A}^*+\beta\mathcal{B}^*$；最后，$*$ 使算子乘法反序：$\forall\mathcal{A},\mathcal{B}\in\mathrm{End}(V)$，$f\in V^*$，都有 $(\mathcal{A}\mathcal{B})^*f=f\circ(\mathcal{A}\circ\mathcal{B})=(f\circ\mathcal{A})\circ\mathcal{B}=\mathcal{B}^*(f\circ\mathcal{A})=(\mathcal{B}^*\mathcal{A}^*)f$，即 $(\mathcal{A}\mathcal{B})^*=\mathcal{B}^*\mathcal{A}^*$。因此，我们称上面的 $*$ 映射为 $\mathrm{End}(V)$ 到 $\mathrm{End}(V^*)$ 的**反同构** (anti-isomorphism)。

下面我们考察对偶算子在对偶基下的矩阵。

定理 2.4.5. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$，$\mathcal{A}$ 在基底 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 下的矩阵为 A，则对偶算子 $\mathcal{A}^*$ 在对偶基 $\mathbf{e}^1,\ldots,\mathbf{e}^n$ 下的矩阵为 A^t。

证明. 设 $A = (a_{ij})_{n\times n}$，显然 $\mathcal{A}\mathbf{e}_j = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \mathbf{A}^{(j)} = \sum_{k=1}^{n} a_{kj}\mathbf{e}_j$，故

$$\mathbf{e}^i(\mathcal{A}\mathbf{e}_j) = \sum_{k=1}^{n} a_{kj}\delta_{ij} = a_{ij}.$$

另一方面，设 $\mathcal{A}^*$ 在对偶基下的矩阵为 $B = (b_{ij})_{n\times n}$，则同理有 $\mathcal{A}^*\mathbf{e}^i = (\mathbf{e}^1, \ldots, \mathbf{e}^n) \cdot \mathbf{B}^{(i)} = \sum_{k=1}^{n} b_{ki}\mathbf{e}^i$，因此

$$(\mathcal{A}^*\mathbf{e}^i)\mathbf{e}_j = \sum_{k=1}^{n} b_{ki}\delta_{ij} = b_{ji}.$$

由对偶算子的定义可知 $\mathbf{e}^i(\mathcal{A}\mathbf{e}_j) = (\mathcal{A}^*\mathbf{e}^i)\mathbf{e}_j$，因此 $a_{ij} = b_{ji}$, $\forall i, j \in \{1, \ldots, n\}$。此即 $B = A^t$。 □

推论 2.4.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $(\mathcal{A}^*)^* = \mathcal{A}$。

证明是显然的，当然也可以由 V 与 V^{**} 的自然同构得到结论。

现在我们已经做好了所有的准备工作，我们接下来将要介绍线性代数最深刻的知识点：算子的循环子空间分解与矩阵的若尔当标准型理论。

2.5　若尔当标准型理论介绍

我们已经知道，一个矩阵并不总是可以相似对角化的 (例2.3.4)，那么，我们是否可以减弱一些要求，让一个矩阵能够相似于某种分块对角矩阵呢？对应地，我们能否将向量空间 V 按照某个线性算子 $\mathcal{A}$ 分解成一些简单的 $\mathcal{A}-$ 子空间的直和呢？这就是若尔当标准型理论的出发点。在本节中，我们始终假定**域 $\mathbb{K}$ 是代数闭的**。

本节我们的主要任务有两个，一是证明任何一个代数闭域上的 $n\times n$ 阶矩阵 A 都相似于一个若尔当标准型 J，并证明 J 在某种意义下的唯一性；二是任给算子 $\mathcal{A}\in\mathrm{End}(V)$，求出一组若尔当基 (也即求出可逆矩阵 P 使得 $P^{-1}AP=J$)。

2.5.1　若尔当标准型的存在唯一性

定义 2.5.1. 设 $\mathbb{K}$ 是代数闭域，V 是 $\mathbb{K}$ 上的 n 维向量空间。

(1) 称 $\mathbb{K}$ 上的 m 阶方阵

$$J_m(\lambda)=\begin{pmatrix}\lambda&1&0&\cdots&0&0\\0&\lambda&1&\cdots&0&0\\0&0&\lambda&\cdots&0&0\\\vdots&\vdots&\vdots&\ddots&\ddots&\vdots\\0&0&0&\cdots&\lambda&1\\0&0&0&\cdots&0&\lambda\end{pmatrix}_{m\times m}$$

为特征值为 λ 的 m 阶 (上) 若尔当块 (Jordan block)[1]。

(2) 如果一个矩阵 $J\in M_n(\mathbb{K})$ 是分块对角矩阵，并且对角块都是若尔当块 (非对角块全为零矩阵)，即 J 形如

$$J=\begin{pmatrix}J_{m_1}(\lambda_1)&&&\\&J_{m_2}(\lambda_2)&&\\&&\ddots&\\&&&J_{m_s}(\lambda_s)\end{pmatrix},\ (\text{其中}\lambda_1,\dots,\lambda_s\text{和}m_1,\dots,m_s\text{可以相同也可以不同})$$

则称 J 是一个若尔当矩阵或若尔当标准型 (Jordan canonical form)。

(3) 若 $\mathcal{A}\in\mathrm{End}(V)$ 在 V 的某组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵是若尔当矩阵，则称这组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是关于算子 $\mathcal{A}$ 的若尔当基 (Jordan basis)。

例 2.5.1. 设 $\mathrm{char}(\mathbb{K})=0$，则 $\mathbb{K}[t]_n$ 上的形式微分算子 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 在基底 $(1,t,\dfrac{t}{2},\dots,\dfrac{t^{n-1}}{(n-1)!})$ 下对应的矩阵

$$\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\0&0&0&\cdots&0&0\\&&&\ddots&\ddots&\\&&&&0&1\\0&0&0&\cdots&0&0\end{pmatrix}=J_n(0).$$

就是若尔当标准型。

---

[1] 很容易验证 $J_m(\lambda)$ 的特征值有且仅有 λ，并且 λ 的几何重数为 1，代数重数为 m。

下面是本小节的主要定理，我们会用整个小节来证明它。

定理 2.5.1. 设 $A \in M_n(\mathbb{K})$，则存在可逆矩阵 $P \in M_n(\mathbb{K})$ 使得 $P^{-1}AP = J$，其中 $J = \begin{pmatrix} J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_s}(\lambda_s) \end{pmatrix}$ 是若尔当标准型。并且，在不计若尔当块的次序的条件下，J 是唯一的。

这就是本节的主要结论，我们将分步骤给出其证明。由线性算子代数 $\mathrm{End}(V)$ 与矩阵代数 $M_n(\mathbb{K})$ 之间的同构关系，我们先按线性算子逐步分解空间 V，然后再对应到矩阵上 (利用推论2.3.1)。

定义 2.5.2 (根子空间). 设 V 是代数闭域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \mathrm{End}(V)$，$\lambda \in \mathbb{K}$ 是 $\mathscr{A}$ 的一个特征值，则我们称子空间 $V(\lambda) = \{\mathbf{v} \in V \mid \exists k \in \mathbb{Z}^+ 使得 (\mathscr{A} - \lambda\mathscr{E})^k\mathbf{v} = \mathbf{0}\}$(验证这是一个子空间留作练习) 是特征值 λ 对应的**根子空间** (root subspace)。

引理 2.5.1. 设 V 是代数闭域 $\mathbb{K}$ 上的 n 维向量空间，$\mathscr{A} \in \mathrm{End}(V)$ 的特征多项式 $\chi_{\mathscr{A}}(t)$ 在 $\mathbb{K}$ 上可以分解成如下一次因式的乘积：

$$\chi_{\mathscr{A}}(t) = (t-\lambda_1)^{r_1} \cdot (t-\lambda_2)^{r_2} \cdots (t-\lambda_p)^{r_p}，其中 r_1 + \cdots + r_p = n$$

则 V 可以分解成如下的 $\mathscr{A}-$ 子空间的直和：

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_p$$

其中 $V_i = \{\mathbf{v} \in V \mid (\mathscr{A} - \lambda_i\mathscr{E})^{r_i}\mathbf{v} = \mathbf{0}\}$，$\forall i \in \{1, \ldots, p\}$。

证明. 我们令

$$f_i(t) = \frac{\chi_{\mathscr{A}}(t)}{(t-\lambda_i)^{r_i}} \in \mathbb{K}[t],\ i \in \{1, \ldots, p\}.$$

及

$$U_i = f_i(\mathscr{A})V = \{f_i(\mathscr{A})\mathbf{v} \mid \forall \mathbf{v} \in V\}$$

则显然 U_i 是子空间，并且由于任取 $f_i(\mathscr{A})\mathbf{v} \in U_i$，我们有

$$\mathscr{A}(f_i(\mathscr{A})\mathbf{v}) = f_i(\mathscr{A})(\mathscr{A}\mathbf{v}) \in f_i(\mathscr{A})V, \text{ (这里利用了 } \mathscr{A} \text{ 与 } f_i(\mathscr{A}) \text{ 可交换)}$$

即 U_i 是 $\mathscr{A}-$ 子空间。我们先来证明 $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$，之后再来证明 $U_i = V_i$。

(I) 首先，$V = U_1 + U_2 + \cdots + U_n$。实际上，首先容易看出

$$\gcd(f_1(t), \ldots, f_p(t)) = 1.$$

于是我们将讲义上册的定理 6.1.7 稍加推广，即有：存在 $u_1(t), \ldots, u_p(t) \in \mathbb{K}[t]$ 使得

$$\sum_{i=1}^{p} f_i(t)u_i(t) = 1. \text{ (Bezout 关系)}$$

以算子 $\mathscr{A}$ 代入上式得

$$\sum_{i=1}^{p} f_i(\mathscr{A})u_i(\mathscr{A}) = \mathscr{E}$$

于是任取 $\mathbf{v} \in V$，都有

$$\mathbf{v} = \mathcal{E}\mathbf{v} = \sum_{i=1}^{p} f_i(\mathcal{A})(u_i(\mathcal{A})\mathbf{v})$$

显然每个 $f_i(\mathcal{A})(u_i(\mathcal{A})\mathbf{v}) \in U_i$，这就证明了 $V = U_1 + U_2 + \cdots + U_n$。

(II) 为了证明 (I) 中的和是直和，我们需要证明：如果 $\mathbf{v}_1 + \cdots + \mathbf{v}_p = \mathbf{0}$，其中每个 $\mathbf{v}_i \in U_i$，则一定有 $\mathbf{v}_1 = \cdots = \mathbf{v}_p = \mathbf{0}$。不妨设每个 $\mathbf{v}_i = f_i(\mathcal{A})\mathbf{w}_i$，则对任意固定的 $i \in \{1, \ldots, p\}$，将 $f_i(\mathcal{A})$ 作用到 $\mathbf{v}_1 + \cdots + \mathbf{v}_p = \mathbf{0}$ 两边得

$$f_i(\mathcal{A})(f_1(\mathcal{A})\mathbf{w}_1 + \cdots + f_p(\mathcal{A})\mathbf{w}_p) = \mathbf{0} \tag{2.5.1}$$

注意到对任意的 $i \neq j$，$\chi_{\mathcal{A}}(t) \mid f_i(t)f_j(t)$(利用了 f_i 的构造)，即 $\exists g_j(t) \in \mathbb{K}[t]$ 使得 $g_j(t)\chi_{\mathcal{A}}(t) = f_i(t)f_j(t)$，这表明 $f_i(\mathcal{A})f_j(\mathcal{A})\mathbf{w}_j = g_j(\mathcal{A})\chi_{\mathcal{A}}(\mathcal{A})\mathbf{w}_j = \mathbf{0}$(利用了哈密顿-凯莱定理)。于是，式 (2.5.1) 即

$$f_i(\mathcal{A})f_i(\mathcal{A})\mathbf{w}_i = f_i(\mathcal{A})\mathbf{v}_i = \mathbf{0} \tag{2.5.2}$$

下面我们说明 $\mathbf{v}_i = \mathbf{0}$, $\forall i$。首先，由 f_i 的构造可知 $\gcd(f_i, (t-\lambda_i)^{r_i}) = 1$，于是存在 $a_i(t), b_i(t) \in \mathbb{K}[t]$ 使得

$$a_i(t)f_i(t) + b_i(t)(t-\lambda_i)^{r_i} = 1.$$

以算子 $\mathcal{A}$ 代入上式得

$$a_i(\mathcal{A})f_i(\mathcal{A}) + b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i} = \mathcal{E},$$

再作用到 $\mathbf{v}_i$ 上得

$$\begin{aligned}
\mathbf{v}_i &= \mathcal{E}\mathbf{v}_i \\
&= a_i(\mathcal{A})f_i(\mathcal{A})\mathbf{v}_i + b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i}\mathbf{v}_i \\
&= \mathbf{0} + b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i} f_i(\mathcal{A})\mathbf{w}_i \\
&= b_i(\mathcal{A})\chi_{\mathcal{A}}(\mathcal{A})\mathbf{w}_i \\
&= \mathbf{0}
\end{aligned}$$

这里用到了式 (2.5.2) 和哈密顿-凯莱定理。

由上面的 (I)、(II) 两点及命题1.2.1即可知 $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$。最后我们来说明每个 $U_i = V_i = \ker((\mathcal{A} - \lambda_i\mathcal{E})^{r_i})$。一方面，$(\mathcal{A} - \lambda_i\mathcal{E})^{r_i}U_i = \{\mathbf{0}\}$(哈密顿-凯莱定理)，即 $U_i \subset V_i$；另一方面，任取 $\mathbf{x} \in V_i$，设 $\mathbf{x}$ 可以唯一分解成如下向量之和：

$$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_p,\ \mathbf{x}_j \in U_j,\ j \in \{1, \ldots, p\}. \tag{2.5.3}$$

我们只要证明 $\mathbf{x} = \mathbf{x}_i \in V_i$ 即可。为此，将式 (2.5.3) 改写成

$$\mathbf{0} = \mathbf{x}_1 + \cdots + (\mathbf{x}_i - \mathbf{x}) + \cdots + \mathbf{x}_p$$

则类似于上面 (II) 的证明过程，将 $f_i(\mathcal{A})$ 作用到上式两边可得

$$f_i(\mathcal{A})(\mathbf{x}_i - \mathbf{x}) = 0.$$

再以 $a_i(\mathcal{A})f_i(\mathcal{A}) + b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i} = \mathcal{E}$ 作用到上式两边得

$$\mathbf{x}_i - \mathbf{x} = \mathcal{E}(\mathbf{x}_i - \mathbf{x}) = a_i(\mathcal{A})f_i(\mathcal{A})(\mathbf{x}_i - \mathbf{x}) + b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i}\mathbf{x}_i - b_i(\mathcal{A})(\mathcal{A} - \lambda_i\mathcal{E})^{r_i}\mathbf{x} = \mathbf{0}$$

即 $\mathbf{x} = \mathbf{x}_i \in V_i$。这样我们就完成了证明。 □

实际上，上面引理中的 V_i 就是根子空间 $V(\lambda_i)$。这是因为：一方面，由定义立刻有 $V_i \subset V(\lambda_i)$；另一方面，任取 $\mathbf{x} \in V(\lambda_i)$，不妨设 $(\mathcal{A} - \lambda_i\mathcal{E})^k\mathbf{x} = \mathbf{0}$，我们只需证明 $k \leqslant r_i$ 即可证明 $V_i \supset V(\lambda_i)$。用反证法，如果存在 $\mathbf{x} \in V(\lambda_i)$ 使得 $(\mathcal{A}-\lambda_i\mathcal{E})^k\mathbf{x} = \mathbf{0}$ 而 $(\mathcal{A}-\lambda_i\mathcal{E})^{k-1}\mathbf{x} \neq \mathbf{0}$, $k \geqslant r_i+1$，那么容易验证 $U = \operatorname{span}\{\mathbf{x}, (\mathcal{A}-\lambda_i\mathcal{E})\mathbf{x}, \ldots, (\mathcal{A}-\lambda_i\mathcal{E})^{k-1}\mathbf{x}\}$ 是 $\mathcal{A}-$ 子空间 (练习)，将线性无关 [1] 的向量组 $\mathbf{x}, (\mathcal{A} - \lambda_i\mathcal{E})\mathbf{x}, \ldots, (\mathcal{A} - \lambda_i\mathcal{E})^{k-1}\mathbf{x}$ 扩充成 V 的一组基，则在此基底下 $\mathcal{A}$ 对应的矩阵具有如下形式 (利用定理2.3.1)：

$$A = \begin{pmatrix} A_1 & A_2 \\ O & A_3 \end{pmatrix}$$

其中 $A_1 \in M_k(\mathbb{K})$ 是 $\mathcal{A}\big|_U$ 在此基底下的矩阵，且满足 $(A - \lambda_i E_k)^k = O_{k\times k}$。于是，$\mathcal{A}$ 的特征多项式

$$\chi_{\mathcal{A}}(t) = \det(tE - A) = \det(tE_k - A_1) \cdot \det(tE_{n-k} - A_3).$$

其中 $\det(tE_k - A_1) = \chi_{\mathcal{A}|_U}(t)$。容易验证 $\mathcal{A}\big|_U$ 的极小多项式是 $(t - \lambda_i)^k$(思考之，或参考下面关于循环不变子空间的论述)，故 $(t - \lambda_i)^k \mid \det(tE_k - A_1)$(实际上两者是相等的，利用次数就可以看出这一点)。因此，$(t - \lambda_i)^k \mid \chi_{\mathcal{A}}(t)$ $(k > r_i)$。但是，由 $\chi_{\mathcal{A}}$ 的因式分解唯一性，这是不可能的！这样我们就证明了 $V_i \supset V(\lambda_i)$，于是 $V_i = V(\lambda_i)$。

我们还有：$\mathcal{A} - \lambda_i\mathcal{E}$ 作用在 $V_1 \oplus \cdots V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_p$ 上是非退化 (可逆) 的。这是显然的，因为 $\ker(\mathcal{A} - \lambda_i\mathcal{E})$ 包含在 V_i 中，而 V_i 与其余空间的直和的交是零子空间。

结合引理2.5.1及以上论述，我们有：

定理 2.5.2. 设 V 是代数闭域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \operatorname{End}(V)$ 的特征多项式 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{K}$ 上可以分解成如下一次因式的乘积：

$$\chi_{\mathcal{A}}(t) = (t - \lambda_1)^{r_1} \cdot (t - \lambda_2)^{r_2} \cdots (t - \lambda_p)^{r_p}, \text{其中} r_1 + \cdots + r_p = n$$

则 $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$，其中每个根子空间 $V(\lambda_i)$ 都是 $\mathcal{A}-$ 子空间，其维数是 r_i。$\mathcal{A} - \lambda_i\mathcal{E}$ 限制在 $V(\lambda_i)$ 上是幂零算子，限制到 $V(\lambda_1) \oplus \cdots V(\lambda_{i-1}) \oplus V(\lambda_{i+1}) \oplus \cdots \oplus V(\lambda_p)$ 上则是可逆算子。最后，λ_i 是 $\mathcal{A}\big|_{V(\lambda_i)}$ 的唯一一个特征值。

上面的根子空间分解用矩阵的语言叙述出来就是：任何代数闭域 $\mathbb{K}$ 上的 n 阶方阵 A 都一定相似于分块对角矩阵

$$\begin{pmatrix} B_1 + \lambda_1 E_{r_1} & & \\ & \ddots & \\ & & B_p + \lambda_p E_{r_p} \end{pmatrix} \tag{2.5.4}$$

---

[1] 这个向量组确实是线性无关的。事实上，如果 $\lambda_1\mathbf{x} + \cdots + \lambda_k(\mathcal{A} - \lambda_i\mathcal{E})^{k-1}\mathbf{x} = \mathbf{0}$，$\lambda_1, \ldots, \lambda_k \in \mathbb{K}$，那么以 $(\mathcal{A} - \lambda_i\mathcal{E})^{k-1}$ 作用于上式两边即得 $\lambda_1 = 0$，以之代入上式后再以 $(\mathcal{A} - \lambda_i\mathcal{E})^{k-2}$ 作用于等式两边可得 $\lambda_2 = 0$，…… 递归地可以证明 $\lambda_1 = \cdots = \lambda_k = 0$，即线性无关。

其中 $B_1, \ldots, B_p$ 分别是 $r_1, \ldots, r_p$ 阶的幂零矩阵，$\lambda_1, \ldots, \lambda_p$ 是 A 的全部特征值，其代数重数分别为 $r_1, \ldots, r_p$。因此，为了进一步化简上面的矩阵 (即将根子空间作进一步的分解)，我们只需要对幂零算子进行讨论。在此之前，我们先介绍循环不变子空间的概念。

定义 2.5.3. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，如果有向量 $\mathbf{v} \in V$ 满足

$$\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{k-1}\mathbf{v}$$

线性无关，而 $\mathcal{A}^k\mathbf{v} \in \mathrm{span}\,\{\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{k-1}\mathbf{v}\}$，则称 $U = \mathrm{span}\,\{\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{k-1}\mathbf{v}\}$ 是由算子 $\mathcal{A}$ 和向量 $\mathbf{v}$ 生成的循环子空间 (cyclic subspace)。

显然循环子空间一定是算子 $\mathcal{A}$ 的不变子空间，并且我们有：

定理 2.5.3. 条件同定义2.5.3。令 $\mathcal{B} = \mathcal{A}\big|_U$，则 $\mu_{\mathcal{B}}(t) = \chi_{\mathcal{B}}(t)$。

证明. 不妨设 $\mathcal{A}^k\mathbf{v} = \sum_{i=0}^{k-1} a_i\mathcal{A}^i\mathbf{v}$，则显然在 U 的基底 $(\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{k-1}\mathbf{v})$ 下算子 $\mathcal{B}$ 对应的矩阵 B 为

$$B = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{k-1} \end{pmatrix}$$

所以 $\chi_{\mathcal{B}}(t) = t^k - a_{k-1}t^{k-1} - \cdots - a_1t - a_0$。用反证法，如果 $\mu_{\mathcal{B}}(t)$ 是 $\chi_{\mathcal{B}}(t)$ 的真因子，即

$$\mu_{\mathcal{B}}(t) = t^d + b_{d-1}t^{d-1} + \cdots + b_1t + b_0 \mid \chi_{\mathcal{B}}(t),\ d < k,\ b_0, \ldots, b_{d-1} \in \mathbb{K}.$$

由于 $\mu_{\mathcal{B}}(\mathcal{B}) = \mathcal{O}$ 可知 $\mathcal{B}^d\mathbf{v} + b_{d-1}\mathcal{B}^{d-1}\mathbf{v} + \cdots + b_1\mathcal{B}\mathbf{v} + b_0\mathcal{E}\mathbf{v} = \mathbf{0}$，这说明向量组 $\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^d\mathbf{v}$ $(d < k)$ 线性相关，这与 $\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{k-1}\mathbf{v}$ 是基矛盾！这样我们就完成了证明。 □

将上面的结论用到幂零算子上，我们立刻得到：

推论 2.5.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$ 是幂零指数为 l 的幂零算子，则存在 $\mathbf{v} \in V$ 使得 $\mathrm{span}\,\{\mathbf{v}, \mathcal{A}\mathbf{v}, \ldots, \mathcal{A}^{l-1}\mathbf{v}\}$ 是循环不变子空间。

证明是显然的。作为一个例子，我们看到 $\mathbb{K}[t]_n$ $(\mathrm{char}(\mathbb{K}) = 0)$ 上的形式微分算子就是一个幂零算子，其极小多项式为 $\chi(x) = x^n$，并且由 $\dfrac{\mathrm{d}}{\mathrm{d}t}$ 和 $\frac{1}{(n-1)!}t^{n-1}$ 生成的循环子空间 $\mathrm{span}\,\{\frac{1}{(n-1)!}t^{n-1}, \ldots, t, 1\}$ 就是 $\mathbb{K}[t]_n$ 本身。

下面我们讨论一下若尔当块的特征多项式和极小多项式。直接计算可知若尔当块 $J_m(\lambda)$ 的特征多项式为 $(t-\lambda)^m$。注意到 $J_m(\lambda) - \lambda E_m$ 是幂零矩阵 $\begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$，幂零指数是 m，于是 $J_m(\lambda) - \lambda E_m$ 的极小多项式为 t^m。所以 $J_m(\lambda)$ 的极小多项式也是 $(t-\lambda)^m$。(首先，$(t-\lambda)^m$ 一定是 $J_m(\lambda)$ 的零化多项式，于是 $J_m(\lambda)$ 的极小多项式一定形如 $(t-\lambda)^k$，$k \leqslant m$。若 $k < m$，则 $J_m(\lambda) - \lambda E_m$ 有零化多项式 t^k，$k < m$，矛盾！故 $k = m$)。

现在我们已经做好了分解根子空间 (按幂零算子将根子空间分解成循环不变子空间) 的所有准备工作了。

定理 2.5.4. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{B} \in \mathrm{End}(V)$ 是幂零算子，则存在如下形式的 V 的一组基：

$$\begin{aligned}
&\mathcal{B}^{k_1-1}\mathbf{e}_1, \ldots, \mathcal{B}\mathbf{e}_1, \mathbf{e}_1,\\
&\mathcal{B}^{k_2-1}\mathbf{e}_2, \ldots\ldots \mathcal{B}\mathbf{e}_2, \mathbf{e}_2,\\
&\ldots\ldots\\
&\mathcal{B}^{k_s-1}\mathbf{e}_s, \ldots\ldots\ldots, \mathcal{B}\mathbf{e}_s, \mathbf{e}_s.
\end{aligned} \tag{2.5.5}$$

其中 $\mathcal{B}^{k_1}\mathbf{e}_1 = \mathcal{B}^{k_2}\mathbf{e}_2 = \cdots = \mathcal{B}^{k_s}\mathbf{e}_s = \mathbf{0}$，$k_1, \ldots, k_s \in \mathbb{N}$ 可以相同也可以不同。在此基底下，算子 $\mathcal{B}$ 的矩阵 B 为

$$B = \begin{pmatrix} J_{k_1}(0) & & & \\ & J_{k_2}(0) & & \\ & & \ddots & \\ & & & J_{k_s}(0) \end{pmatrix}.$$

证明. 对维数 n 用数学归纳法。$n = 1$ 时设 $V = \mathrm{span}\,\{\mathbf{e}_1\}$，则由 $\mathcal{B}$ 是幂零算子可知 $\mathcal{B}$ 的特征值有且只有 0，故此时 $\mathcal{B}$ 对应的矩阵就是零矩阵 $J_1(0)$(这里把零算子暂且当作特殊的幂零算子)，此时定理成立。

下设定理对一切 $\mathbb{K}$ 上的维数小于 n 的向量空间都成立，我们来证明定理对 n 维空间成立。考虑 $\mathbb{B}$ 的像空间 $\mathrm{im}(\mathcal{B})$，在例2.3.1中我们已经知道 $\mathrm{im}(\mathcal{B})$ 是 $\mathcal{B}-$ 子空间，如果 $\dim(\mathrm{im}(\mathcal{B})) = n$，则说明算子 $\mathcal{B}$ 可逆 (满秩)，这与 $\mathcal{B}$ 是幂零算子矛盾！故 $\dim(\mathrm{im}(\mathcal{B})) < n$。将 $\mathcal{B}$ 限制到 $\mathrm{im}(\mathcal{B})$ 上，则限制算子仍然是幂零的，由归纳假设，$\mathrm{im}(\mathcal{B})$ 有一组基

$$\begin{aligned}
&\mathcal{B}^{k_1-1}\mathbf{v}_1, \ldots, \mathcal{B}\mathbf{v}_1, \mathbf{v}_1\\
&\mathcal{B}^{k_2-1}\mathbf{v}_2, \ldots\ldots \mathcal{B}\mathbf{v}_2, \mathbf{v}_2\\
&\ldots\ldots\\
&\mathcal{B}^{k_t-1}\mathbf{v}_t, \ldots\ldots\ldots, \mathcal{B}\mathbf{v}_t, \mathbf{v}_t.
\end{aligned} \tag{2.5.6}$$

其中 $\mathcal{B}^{k_1}\mathbf{v}_1 = \mathcal{B}^{k_2}\mathbf{v}_2 = \cdots = \mathcal{B}^{k_t}\mathbf{v}_t = \mathbf{0}$。注意到 $\mathbf{v}_1, \ldots, \mathbf{v}_t \in \mathrm{im}(\mathcal{B})$，即存在 $\mathbf{e}_1, \ldots, \mathbf{e}_t \in V$ 使得 $\mathcal{B}\mathbf{e}_i = \mathbf{v}_i, \forall i \in \{1, \ldots, t\}$，于是 V 中的以下向量是线性无关的：

$$\begin{aligned}
&\mathcal{B}^{k_1-1}\mathbf{e}_1, \ldots, \mathcal{B}\mathbf{e}_1, \mathbf{e}_1,\\
&\mathcal{B}^{k_2-1}\mathbf{e}_2, \ldots\ldots \mathcal{B}\mathbf{e}_2, \mathbf{e}_2,\\
&\ldots\ldots\\
&\mathcal{B}^{k_t-1}\mathbf{e}_t, \ldots\ldots\ldots, \mathcal{B}\mathbf{e}_t, \mathbf{e}_t.
\end{aligned} \tag{2.5.7}$$

这是因为 $\mathcal{B}$ 作用到向量组 (2.5.7) 的每个向量上正是向量组 (2.5.6) 中对应位置的那个向量，如果向量组 (2.5.7) 线性相关，则向量组 (2.5.6) 也线性相关，这与向量组 (2.5.6) 是 $\mathrm{im}(\mathcal{B})$ 的一组基矛盾！在 $\ker(\mathcal{B})$ 中取向量 $\mathcal{B}^{k_1}\mathbf{e}_1, \ldots, \mathcal{B}^{k_t}\mathbf{e}_t$，容易看出它们线性无关，于是可以将它们扩充成 $\ker(\mathcal{B})$ 的一组基 $\mathcal{B}^{k_1}\mathbf{e}_1, \ldots, \mathcal{B}^{k_t}\mathbf{e}_t, \mathbf{e}_{t+1}, \ldots, \mathbf{e}_s$，则将 $\ker(\mathcal{B})$ 的一组基与向量组 (2.5.7) 合在一起就是 V 的基 (首先合在一起的向量组线性无关，这是因为如果它们的一个线性组合是 $\mathbf{0}$，则以算子 $\mathcal{B}$ 作用到这个线性组合上，可知向量组 (2.5.7) 的组合系数全是 0，再由 $\mathbf{e}_{t+1}, \ldots, \mathbf{e}_s$ 是 $\ker(\mathcal{B})$ 的基可知 $\mathbf{e}_{t+1}, \ldots, \mathbf{e}_s$ 前面的组合系数也是 0；利用对偶定理可知合在一起的向量组含有 n 个向量)。

显然向量组 (2.5.7) 和 $\mathcal{B}^{k_1}\mathbf{e}_1, \ldots, \mathcal{B}^{k_t}\mathbf{e}_t, \mathbf{e}_{t+1}, \ldots, \mathbf{e}_s$ 合在一起之后仍然具有 (2.5.5) 所要求

的基的形式 (其中 $k_{t+1} = \cdots = k_s = 0$)。这样我们就完成了归纳证明。

最后，验证在基底 (2.5.5) 下 $\mathcal{B}$ 的矩阵是 B 是容易的。(这里空白太小，真的写不下排出的矩阵，只好给出 $\mathcal{B}$ 作用到 (2.5.5) 上的结果) 这是因为 $\mathcal{B}$ 作用到 (2.5.5) 上得到的结果为：

$$\begin{aligned}&\mathbf{0}, \mathcal{B}^{k_1-1}\mathbf{e}_1, \ldots, \mathcal{B}\mathbf{e}_1,\\ &\mathbf{0}, \mathcal{B}^{k_2-1}\mathbf{e}_2, \ldots\ldots \mathcal{B}\mathbf{e}_2,\\ &\ldots\ldots\\ &\mathbf{0}, \mathcal{B}^{k_s-1}\mathbf{e}_s, \ldots\ldots\ldots, \mathcal{B}\mathbf{e}_s.\end{aligned}$$

□

回过头来考察基底 (2.5.5)。我们看到，基底 (2.5.5) 的第 i 行中的向量恰好是由 $\mathcal{B}$ 和 $\mathbf{e}_i$ 生成的 k_i 维循环不变子空间 (推论2.5.1)。我们记

$$\mathbb{K}[\mathcal{B}]\mathbf{e}_i = \operatorname{span}\{\mathcal{B}^{k_i-1}\mathbf{e}_i, \ldots, \mathcal{B}\mathbf{e}_i, \mathbf{e}_i\},\ i \in \{1, \ldots, s\}$$

则定理2.5.4也可以表述为：$\mathbb{K}$ 上的 n 维向量空间 V 可以按幂零算子 $\mathcal{B}$ 分解成循环不变子空间的直和 $V = \mathbb{K}[\mathcal{B}]\mathbf{e}_1 \oplus \cdots \oplus \mathbb{K}[\mathcal{B}]\mathbf{e}_s$。

定理2.5.4告诉我们，分块对角矩阵 (2.5.4) 中的每个分块 $B_i + \lambda_i E_{r_i}$ 都一定相似于若尔当标准型

$$\begin{pmatrix} J_{k_{i,1}}(\lambda_1) & & & \\ & J_{k_{i,2}}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{k_{i,s_i}}(\lambda_i) \end{pmatrix}$$

其中 $k_{i,1} + \cdots + k_{i,s_i} = r_i$。将 $i = 1, \ldots, p$ 对应的若尔当标准型合到一起就证明了任何代数闭域上的 n 阶方阵都相似于若尔当标准型

$$J = \begin{pmatrix} \underbrace{\begin{matrix} J_{k_{1,1}}(\lambda_1) & & \\ & \ddots & \\ & & J_{k_{1,s_1}}(\lambda_1) \end{matrix}}_{r_1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \underbrace{\begin{matrix} J_{k_{p,1}}(\lambda_p) & & \\ & \ddots & \\ & & J_{k_{p,s_p}}(\lambda_p) \end{matrix}}_{r_p} \end{pmatrix} \tag{2.5.8}$$

也即任给 $\mathcal{A} \in \operatorname{End}(V)$，$V$ 可以分解成一些 $\mathcal{A} - \lambda_i\mathcal{E}$, $i = 1, \ldots, p$ (每个 λ_i 都是 $\mathcal{A}$ 的特征值) 的循环不变子空间的直和。到这里，若尔当标准型的存在性就已经证完了。

现在我们距离完全证明定理2.5.1还差证明唯一性。由于根子空间分解是唯一的，因此我们只需要说明：在不计若尔当块的次序的意义下，每个 $B_i + \lambda_i E_{r_i}$ 的若尔当标准型是唯一的。实际上，我们只需要说明：对 $\mathcal{A}$ 的每个特征值 λ_i 而言，m 阶若尔当块 $J_m(\lambda_i)$ 的个数 $N(m, \lambda_i)$ 是确定的。下面的定理给出了计算 $N(m, \lambda_i)$ 的方法。

定理 2.5.5. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，其在一组若尔当基下的矩阵形如 (2.5.8) 中的 J。记 J 中特征值 λ_i 对应的 m 阶若尔当块 $J_m(\lambda_i)$ 的个数为 $N(m, \lambda_i)$ 及 $\mathrm{rk}_t = \mathrm{rank}(\mathcal{A} - \lambda_i \mathcal{E})^t$，$t \in \mathbb{N}$。则

$$N(m, \lambda_i) = \mathrm{rk}_{m-1} - 2\mathrm{rk}_m + \mathrm{rk}_{m+1}.$$

证明. 不妨设 $V = V(\lambda_i) \oplus V'$，其中 $V' = \overset{p}{\underset{j=1, j\neq i}{\oplus}} V(\lambda_j)$。由定理2.5.4，$V(\lambda_i)$ 有循环不变子空间的直和分解：

$$V(\lambda_i) = \overset{s_i}{\underset{j=1}{\oplus}} \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j,$$

其中

$$\mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j = \mathrm{span}\,\{\mathbf{e}_j, (\mathcal{A} - \lambda_i \mathcal{E})\mathbf{e}_j, \ldots, (\mathcal{A} - \lambda_i \mathcal{E})^{k_{i,j}-1}\mathbf{e}_j\} \tag{2.5.9}$$

不妨设上面的 $k_{i,1} \leqslant k_{i,2} \leqslant \cdots \leqslant k_{i,s_i}$，则我们要计数的 $N(m, \lambda_i)$ 实际上就是集合 $\left\{j \in \{1, \ldots, s_i\} \mid k_{i,j} = m\right\}$ 中的元素个数。

下面我们考虑 rk_t。首先

$$\mathrm{rk}_t = \mathrm{rank}(\mathcal{A} - \lambda_i \mathcal{E})^t = \dim[\mathrm{im}((\mathcal{A} - \lambda_i \mathcal{E})^t)]$$

而按照定理2.5.2和定理2.5.4，我们知道

$$\begin{aligned}\mathrm{im}((\mathcal{A} - \lambda_i \mathcal{E})^t) &= [(\mathcal{A} - \lambda_i \mathcal{E})^t V(\lambda_i)] \oplus [(\mathcal{A} - \lambda_i \mathcal{E})^t V'] \\ &= \left[\overset{s_i}{\underset{j=1}{\oplus}} (\mathcal{A} - \lambda_i \mathcal{E})^t \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j\right] \oplus (\mathcal{A} - \lambda_i \mathcal{E})^t V'\end{aligned}$$

定理2.5.2告诉我们，$(\mathcal{A} - \lambda_i \mathcal{E})^t$ 限制到 V' 上是可逆算子，即 $\dim((\mathcal{A} - \lambda_i \mathcal{E})^t V') = n - \mathrm{rk}_i$。因此我们只需要考虑 $\overset{s_i}{\underset{j=1}{\oplus}} (\mathcal{A} - \lambda_i \mathcal{E})^t \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j$ 的维数即可。注意到

$$\dim\left[(\mathcal{A} - \lambda_i \mathcal{E})^t \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j\right] = \begin{cases} 0, & k_{i,j} \leqslant t; \\ k_{i,j} - t, & k_{i,j} > t. \end{cases} \tag{2.5.10}$$

式 (2.5.10) 成立是因为：当 $k_{i,j} \leqslant t$ 时 $(\mathcal{A} - \lambda_i \mathcal{E})^t \mathbf{e}_j = \mathbf{0}$ （$(\mathcal{A} - \lambda_i \mathcal{E})^{k_{i,j}} \mathbf{e}_j$ 就已经是 $\mathbf{0}$ 了），于是 $(\mathcal{A} - \lambda_i \mathcal{E})^t$ 作用到 $\mathbf{e}_j$ 生成的循环子空间上必然为零。当 $k_{i,j} > t$ 时，将 $(\mathcal{A} - \lambda_i \mathcal{E})^t$ 作用到循环子空间 (2.5.9) 上可得

$$(\mathcal{A} - \lambda_i \mathcal{E})^t \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j = \mathrm{span}\,\{(\mathcal{A} - \lambda_i \mathcal{E})^t \mathbf{e}_j, \ldots, (\mathcal{A} - \lambda_i \mathcal{E})^{k_{i,j}-1}\mathbf{e}_j\}.$$

即 $\dim\left[(\mathcal{A} - \lambda_i \mathcal{E})^t \mathbb{K}[\mathcal{A} - \lambda_i \mathcal{E}]\mathbf{e}_j\right] = k_{i,j} - t$。

现在我们可以计算 rk_t 了。让 j 取遍 $1, \ldots, s_i$，将式 (2.5.10) 累加起来，再加上 $\dim((\mathcal{A} - \lambda_i \mathcal{E})^t V')$ 就可以得到 rk_t。写出来就是：

$$\mathrm{rk}_t = \sum_{j=1,\ k_{i,j} > t}^{s_i} (k_{i,j} - t) + (n - \mathrm{rk}_i)$$

因此

$$\begin{aligned} \mathrm{rk}_t - \mathrm{rk}_{t+1} &= \sum_{j=1,\ k_{i,j}>t}^{s_i} (k_{i,j}-t) - \sum_{j=1,\ k_{i,j}>t+1}^{s_i} [k_{i,j}-(t+1)] \\ &= \sum_{j=1,\ k_{i,j}=t+1}^{s_i} 1 + \sum_{j=1,\ k_{i,j}>t+1}^{s_i} 1 \\ &= \sum_{j=1,\ k_{i,j}\geqslant t+1}^{s_i} 1 \\ &= N(t+1,\lambda_i) + N(t+2,\lambda_i) + \cdots + N(r_i,\lambda_i) \end{aligned} \tag{2.5.11}$$

(注意若尔当块最大只能是 r_i 阶的)，于是

$$\begin{aligned} &\ \mathrm{rk}_{m-1} - 2\mathrm{rk}_m + \mathrm{rk}_{m+1} \\ =&\ (\mathrm{rk}_{m-1} - \mathrm{rk}_m) - (\mathrm{rk}_m - \mathrm{rk}_{m+1}) \\ =&\ (N(m,\lambda_i) + N(m+1,\lambda_i) + \cdots + N(r_i,\lambda_i)) - (N(m+1,\lambda_i) + N(m+2,\lambda_i) + \cdots + N(r_i,\lambda_i)) \\ =&\ N(m,\lambda_i) \end{aligned}$$

注意我们在证明中虽然使用了循环不变子空间的基底，但得到的结论却与基底的选取无关。也就是说，无论若尔当基如何选取，以 λ_i 为特征值的 m 阶若尔当块的个数都是定值 (只与算子 $\mathcal{A}$ 本身有关)。这样我们就完成了定理的证明，同时也完成了定理2.5.1中唯一性的证明。 □

至此，结合定理2.5.2，定理2.5.4和定理2.5.5的结论，我们就彻底完成了定理2.5.1的证明。

2.5.2 若尔当基的计算与若尔当标准型的应用

现在我们进入第二个任务，计算若尔当基。这是一件比较繁琐的事情。实际上定理2.5.1的证明过程已经给出了计算若尔当基的思路。

定理 2.5.6. *条件同定理2.5.5。则以 λ_i 为特征值的若尔当块的个数为 $n - \mathrm{rk}_1 = \dim(\ker(\mathcal{A} - \lambda_i\mathcal{E}))$。*

证明. 由定理2.5.5的证明过程中的式 (2.5.11)，以 λ_i 为特征值的若尔当块的个数为：

$$\begin{aligned} &N(1,\lambda_i) + N(2,\lambda_i) + \cdots + N(r_i,\lambda_i) \\ =&\ \mathrm{rk}_0 - \mathrm{rk}_1 = n - \mathrm{rk}_1 \end{aligned}$$

由对偶定理即得 $n - \mathrm{rk}_1 = \dim(\ker(\mathcal{A} - \lambda_i\mathcal{E}))$。 □

下面我们开始寻找算子 $\mathcal{A}$ 的若尔当基。首先，设 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t) = (t-\lambda_1)^{r_1}\cdot(t-\lambda_2)^{r_2}\cdots(t-\lambda_k)^{r_k}$，则按照定理2.5.2，我们只需要求出每个 $\mathcal{A}\big|_{V(\lambda_i)}$ 的若尔当基，再合并在一起就得到了 $\mathcal{A}$ 的一组若尔当基。将任意固定的一个 λ_i 简记作 λ，并设 $\mathcal{B} = (\mathcal{A} - \lambda\mathcal{E})\big|_{V(\lambda)}$ 的幂零指数是 p。按照定理2.5.4，我们需要找到一组符合定理要求的 $\mathbf{e}_1,\ldots,\mathbf{e}_s$，再将 $\mathcal{B}$ 反复作用到这组向量上直到得到的向量全为 $\mathbf{0}$，将所有得到的非零向量合起来就是根子空间 $V(\lambda)$ 的基。然而，直接寻找 $\mathbf{e}_1,\ldots,\mathbf{e}_s$ 是不容易的，而 $\ker(\mathcal{B})$ 的基 (即 λ 所对应的特征向量) 是容易计算的，因此我们需要从 $\ker(\mathcal{B})$ 出发，通过基扩充的方式对 $V(\lambda)$ 进行进一步的 “拆解”。

注意到

$$\begin{array}{ccccccccccc} & & \ker(\mathscr{B}) & \subset & \ker(\mathscr{B}^2) & \subset & \cdots & \subset & \ker(\mathscr{B}^{p-1}) & \subset & \ker(\mathscr{B}^p) = V(\lambda) \\ & & \cup & & \cup & & & & \cup & & \\ \{\mathbf{0}\} = \operatorname{im}(\mathscr{B}^p) & \subset & \operatorname{im}(\mathscr{B}^{p-1}) & \subset & \operatorname{im}(\mathscr{B}^{p-2}) & \subset & \cdots & \subset & \operatorname{im}(\mathscr{B}) & & \end{array}$$

因此，对正整数 i，令

$$V(\lambda)_i = \operatorname{im}(\mathscr{B}^{i-1}) \cap \ker(\mathscr{B})$$

那么

$$\ker(\mathscr{B}) = V(\lambda)_1 \supset V(\lambda)_2 \supset \cdots \supset V(\lambda)_p \neq \{\mathbf{0}\}, \quad V(\lambda)_{p+1} = \{\mathbf{0}\}.$$

在 $V(\lambda)$ 中选取线性无关的向量 $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_{s_p}$ 使得 $\mathscr{B}^{p-1}\mathbf{e}_1, \cdots, \mathscr{B}^{p-1}\mathbf{e}_{s_p}$ 是 $V(\lambda)_p$ 的基。则此时

$$\begin{aligned} & \mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_{s_p} \in \ker(\mathscr{B}^p), \\ & \mathscr{B}\mathbf{e}_1, \mathscr{B}\mathbf{e}_2, \cdots, \mathscr{B}\mathbf{e}_{s_p} \in \ker(\mathscr{B}^{p-1}), \\ & \cdots\cdots \\ & \mathscr{B}^{p-1}\mathbf{e}_1, \mathscr{B}^{p-1}\mathbf{e}_2, \cdots, \mathscr{B}^{p-1}\mathbf{e}_{s_p} \in \ker(\mathscr{B}). \end{aligned}$$

接下来在 $V(\lambda)$ 中取线性无关的向量 $\mathbf{e}_{s_p+1}, \mathbf{e}_{s_p+2}, \cdots, \mathbf{e}_{s_{p-1}}$ 使得向量组

$$\mathscr{B}^{p-1}\mathbf{e}_1, \cdots, \mathscr{B}^{p-1}\mathbf{e}_{s_p}, \mathscr{B}^{p-2}\mathbf{e}_{s_p+1}, \cdots, \mathscr{B}^{p-2}\mathbf{e}_{s_{p-1}}$$

是 $V(\lambda)_{p-1}$ 的基。如此下去，最后得到向量组

$$\mathbf{e}_1, \cdots, \mathbf{e}_{s_p}, \mathbf{e}_{s_p+1}, \cdots, \mathbf{e}_{s_{p-1}}, \cdots, \mathbf{e}_{s_2+1}, \cdots, \mathbf{e}_{s_1}$$

使得对于不超过 p 的正整数 k，向量组

$$\mathscr{B}^{p-1}\mathbf{e}_1, \cdots, \mathscr{B}^{p-1}\mathbf{e}_{s_p}, \cdots, \mathscr{B}^{k-1}\mathbf{e}_{s_{k+1}+1}, \cdots, \mathscr{B}^{k-1}\mathbf{e}_{s_k}$$

是 $V(\lambda)_k$ 的基。把算子 $\mathscr{B} = \mathscr{A} - \lambda\mathscr{E}\big|_{V(\lambda)}$ 的幂作用在向量组 $(\mathbf{e}_i)$ 上，得到如下的向量组：

$$
\begin{array}{rl}
\text{s_p个p阶若尔当块}\left\{\begin{array}{ccccc}
\mathbf{e}_1 & \mathscr{B}\mathbf{e}_1 & \cdots\cdots & \mathscr{B}^{p-2}\mathbf{e}_1 & \mathscr{B}^{p-1}\mathbf{e}_1\\
\vdots & \vdots & & \vdots & \vdots\\
\mathbf{e}_{s_p} & \mathscr{B}\mathbf{e}_{s_p} & \cdots\cdots & \mathscr{B}^{p-2}\mathbf{e}_{s_p} & \mathscr{B}^{p-1}\mathbf{e}_{s_p}
\end{array}\right\} & p\times s_p\\
\text{$s_{p-1}-s_p$个$p-1$阶若尔当块}\left\{\begin{array}{ccccc}
\mathbf{e}_{s_p+1} & \mathscr{B}\mathbf{e}_{s_p+1} & \cdots\cdots & \mathscr{B}^{p-2}\mathbf{e}_{s_p+1} & \\
\vdots & \vdots & & \vdots & \\
\mathbf{e}_{s_{p-1}} & \mathscr{B}\mathbf{e}_{s_{p-1}} & \cdots\cdots & \mathscr{B}^{p-2}\mathbf{e}_{s_{p-1}} &
\end{array}\right\} & (p-1)\times(s_{p-1}-s_p)\\
\vdots\quad\vdots\quad\vdots\qquad\qquad\qquad\qquad\qquad\qquad\vdots & \\
\vdots\quad\vdots\quad\vdots\qquad\qquad\qquad\qquad\qquad\qquad\vdots & \\
\text{s_2-s_3个2阶若尔当块}\left\{\begin{array}{ccccc}
\mathbf{e}_{s_3+1} & \mathscr{B}\mathbf{e}_{s_3+1} & & & \\
\vdots & \vdots & & & \\
\mathbf{e}_{s_2} & \mathscr{B}\mathbf{e}_{s_2} & & &
\end{array}\right\} & 2(s_2-s_3)\\
\text{s_1-s_2个1阶若尔当块}\left\{\begin{array}{ccccc}
\mathbf{e}_{s_2+1} & & & & \\
\vdots & & & & \\
\mathbf{e}_{s_1} & & & &
\end{array}\right\} & s_1-s_2
\end{array}
\tag{2.5.12}
$$

这些向量是线性无关的。事实上，设有线性组合

$$
\sum_{i=1}^{s_1}\alpha_{i,1}\mathbf{e}_i + \sum_{i=1}^{s_2}\alpha_{i,2}\mathscr{B}\mathbf{e}_i + \cdots + \sum_{i=1}^{s_p}\alpha_{i,p}\mathscr{B}^{p-1}\mathbf{e}_i = \mathbf{0} \tag{2.5.13}
$$

把算子 $\mathscr{B}^{p-1}$ 作用到上式，得

$$
\sum_{i=1}^{s_p}\alpha_{i,1}\mathscr{B}^{p-1}\mathbf{e}_i = \mathbf{0}
$$

从而

$$
\alpha_{i,1} = 0, \quad 1 \leqslant i \leqslant s_p.
$$

把算子 $\mathscr{B}^{p-2}$ 作用到式 (2.5.13)，得

$$
\sum_{i=s_{p+1}}^{s_{p-1}}\alpha_{i,1}\mathscr{B}^{p-2}\mathbf{e}_i + \sum_{i=1}^{s_p}\alpha_{i,2}\mathscr{B}^{p-1}\mathbf{e}_i = \mathbf{0}
$$

根据向量 $\mathbf{e}_i$ 的选择，得

$$
\alpha_{i,1} = \alpha_{j,2} = 0, \quad s_p + 1 \leqslant i \leqslant s_{p-1}, \quad 1 \leqslant j \leqslant s_p.
$$

如此下去，可知式 (2.5.13) 中的系数全为 0。这个向量组所含的向量的个数是

$$
\begin{aligned}
&ps_p + (p-1)\left(s_{p-1} - s_p\right) + \cdots + 2\left(s_2 - s_3\right) + \left(s_1 - s_2\right) \\
=&s_1 + s_2 + \cdots + s_p \\
=&\dim V(\lambda)_1 + \dim V(\lambda)_2 + \cdots + \dim V(\lambda)_p \\
=&\sum_{i=1}^{p} \dim\left(\mathrm{im}\mathscr{B}^{i-1} \cap \ker\mathscr{B}\right) = \sum_{i=1}^{p}\left(\dim\ker\mathscr{B}^{i} - \dim\ker\mathscr{B}^{i-1}\right) \\
=&\dim\ker\mathscr{B}^{p} = \dim V(\lambda)
\end{aligned}
$$

其中上面倒数第二行的等式是因为：容易验证映射 $\ker(\mathscr{B}^i) \to \mathrm{im}(\mathscr{B}^{i-1}) \cap \ker(\mathscr{B})$, $\mathbf{x} \mapsto \mathscr{B}^{i-1}\mathbf{x}$ 是满射，其核是 $\ker(\mathscr{B}^{i-1})$，于是 $\ker(\mathscr{B}^i)/\ker(\mathscr{B}^{i-1}) \simeq \mathrm{im}(\mathscr{B}^{i-1}) \cap \ker(\mathscr{B})$。由上面的计算可知向量组 (2.5.12) 是 $V(\lambda)$ 的基，其构造方式表明这是 $\mathscr{B}$ 的若尔当基。向量组 (2.5.12) 中的每一行向量给出了 $\mathscr{B}$(等价地，$\mathscr{A}$) 的一个若尔当块。

到此为止，我们已经完成了本节一开始就提到的两个任务。下面我们通过具体的例子来进一步展示若尔当标准型和若尔当基的计算方法，以及一些实际计算中可以减少计算量的小技巧。

例 2.5.2. 设 $A = \begin{pmatrix} 6 & -9 & 5 & 4 \\ 7 & -13 & 8 & 7 \\ 8 & -17 & 11 & 8 \\ 1 & -2 & 1 & 3 \end{pmatrix} \in M_4(\mathbb{C})$，求 A 的若尔当标准型 J 及满足 $P^{-1}AP = J$ 的可逆矩阵 P。

解. 容易计算出：矩阵 A 的特征多项式 $\chi_A(t) = (t-1)(t-2)^3$。于是根子空间 $V(1)$ 是一维的，即特征值 1 对应的若尔当块只有一个，即 $J_1(1)$。下面考虑 $V(2)$，它是 3 维的，即各个若尔当块的阶数之和应该是 3，同时若尔当块的最大阶数不超过 3。经计算可得：

$$
A - 2E = \begin{pmatrix} 4 & -9 & 5 & 4 \\ 7 & -15 & 8 & 7 \\ 8 & -17 & 9 & 8 \\ 1 & -2 & 1 & 1 \end{pmatrix},\ (A-2E)^2 = \begin{pmatrix} -3 & 6 & -3 & -3 \\ -6 & 12 & -6 & -6 \\ -7 & 14 & -7 & -7 \\ -1 & 2 & -1 & -1 \end{pmatrix},
$$

$$
(A-2E)^3 = \begin{pmatrix} 3 & -6 & 3 & 3 \\ 6 & -12 & 6 & 6 \\ 7 & -14 & 7 & 7 \\ 1 & -2 & 1 & 1 \end{pmatrix},\ (A-2E)^4 = \begin{pmatrix} -3 & 6 & -3 & -3 \\ -6 & 12 & -6 & -6 \\ -7 & 14 & -7 & -7 \\ -1 & 2 & -1 & -1 \end{pmatrix}
$$

分别计算上述矩阵的秩，可得

$$
\mathrm{rank}(A-2E) = 2,\ \mathrm{rank}((A-2E)^2) = \mathrm{rank}((A-2E)^3) = \mathrm{rank}((A-2E)^4) = 1.
$$

于是按照定理2.5.5，以 2 为特征值的若尔当块个数如下：

$$
N(1,2) = 4 - 2 \times 2 + 1 = 1,\ N(2,2) = 2 - 2 \times 1 + 1 = 1,\ N(3,2) = 1 - 2 \times 1 + 1 = 0
$$

即特征值 2 对应了两个若尔当块：$J_1(2)$ 和 $J_2(2)$。综上所述，A 的若尔当标准型为

$$J = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}.$$

下面我们来求满足要求的矩阵 P。设 $\mathscr{A} \in \mathrm{End}(\mathbb{C}^4)$ 在标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_4$ 下对应的矩阵为 A，在基 $\mathbf{v}_1, \ldots, \mathbf{v}_4$ 下对应的矩阵为 J，则 P 是标准基到基底 $\mathbf{v}_1, \ldots, \mathbf{v}_4$ 的过渡矩阵，即

$$(\mathbf{v}_1, \ldots, \mathbf{v}_4) = (\mathbf{e}_1, \ldots, \mathbf{e}_4) \cdot P.$$

因此，我们只需求出 $\mathscr{A}$ 的一组若尔当基即可得到矩阵 P，而这组若尔当基应该满足

$$\begin{cases} (\mathscr{A} - \mathscr{E})\mathbf{v}_1 = \mathbf{0} \\ (\mathscr{A} - 2\mathscr{E})\mathbf{v}_2 = \mathbf{0} \\ (\mathscr{A} - 2\mathscr{E})\mathbf{v}_3 = \mathbf{0} \\ (\mathscr{A} - 2\mathscr{E})\mathbf{v}_4 = \mathbf{v}_3 \end{cases}$$

首先，解方程组 $(A - E)\mathbf{x} = \mathbf{0}$ 即可得到根子空间 $V(1)$ 的基为 $\mathbf{v}_1 = (3, 6, 7, 1)^t$。下面求解 $V(2)$ 的基。首先，解方程组 $(A - 2E)\mathbf{x} = \mathbf{0}$ 可知特征值 2 所对应的特征子空间 V^2 有一组基

$$\mathbf{v}_2 = (1, 0, 0, -1)^t,\ \mathbf{v}_3 = (0, 1, 1, 1)^t.$$

不妨设 $\mathbf{v}_2$ 对应的若尔当块是 1 阶的，则 $\mathbf{v}_3$ 对应的若尔当块是 2 阶的，于是我们需要 $\mathbf{v}_3$ 是 $(\mathscr{A} - 2\mathscr{E})\mathbf{v}_4$ 的形式，即 $\mathbf{v}_3$ 应在 $\ker(\mathscr{A} - 2\mathscr{E}) \cap \mathrm{im}(\mathscr{A} - 2\mathscr{E})$ 中。然而，很容易验证

$$(A - 2E)\mathbf{x} = \begin{pmatrix} 4 & -9 & 5 & 4 \\ 7 & -15 & 8 & 7 \\ 8 & -17 & 9 & 8 \\ 1 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

无解。这说明我们选取的 $\mathbf{v}_3$ 不符合要求。重新选取 $\mathbf{v}_3' = \mathbf{v}_2 + \lambda \mathbf{v}_3$，显然此时 $\mathbf{v}_3' \in \ker(\mathscr{A} - 2\mathscr{E})$，我们需要 $\mathbf{v}_3'$ 也在 $\mathrm{im}(\mathscr{A} - 2\mathscr{E})$ 中，即我们需要非齐次线性方程组

$$(A - 2E)\mathbf{x} = \begin{pmatrix} 4 & -9 & 5 & 4 \\ 7 & -15 & 8 & 7 \\ 8 & -17 & 9 & 8 \\ 1 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \mathbf{v}_3' = \begin{pmatrix} 1 \\ \lambda \\ \lambda \\ \lambda - 1 \end{pmatrix}$$

有解。简单计算可得：取 $\lambda = 1$ 时方程组有解 $\mathbf{v}_4 = (0, 0, 1, -1)^t$，此时 $\mathbf{v}_3' = (1, 1, 1, 0)^t$。故 $\mathbf{v}_1 = (3, 6, 7, 1)^t$，$\mathbf{v}_2 = (1, 0, 0, -1)^t$，$\mathbf{v}_3' = (1, 1, 1, 0)^t$，$\mathbf{v}_4 = (0, 0, 1, -1)^t$ 是算子 $\mathscr{A}$ 对应的一组若

尔当基，即

$$P = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3', \mathbf{v}_4) = \begin{pmatrix} 3 & 1 & 1 & 0 \\ 6 & 0 & 1 & 0 \\ 7 & 0 & 1 & 1 \\ 1 & -1 & 0 & -1 \end{pmatrix} \in M_4(\mathbb{C})$$

是从标准基到基底 $(\mathbf{v}_i)$ 的过渡矩阵，所以 $P^{-1}AP = J$。 □

注 2.5.1. 注意若尔当基的选取不是唯一的。例如，在上面的例子中 $\mathcal{A}$ 的若尔当基也可以取 $\mathbf{u}_1 = (-3,-6,-7,-1)^t$，$\mathbf{u}_2 = (-1,0,0,1)^t$，$\mathbf{u}_3 = (1,1,1,0)^t$，$\mathbf{u}_4 = (3,6,7,2)^t$，读者可以自行验证。

在上面的计算过程中，我们可以看到，在矩阵规模比较小时，我们不一定非要像前面的论述过程那样，直接找出 $V(\lambda)_p, V(\lambda)_{p-1}, \ldots$ 的基，而也可以从特征向量出发，通过调整逐步得到 $V(\lambda)_p, V(\lambda)_{p-1}, \ldots$ 的基。

思考题 2.5.1. 计算 $B = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{pmatrix} \in M_3(\mathbb{C})$ 的若尔当标准型 J 以及求矩阵 P 使得 $P^{-1}BP = J$。

在本节的最后，我们来考虑若尔当标准型理论的应用。首先，有了若尔当标准型理论，我们可以将线性算子 (或矩阵) 的特征多项式和极小多项式的构成讲清楚了。

定理 2.5.7. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$ 在一组基下的矩阵是式 (2.5.8) 中的若尔当标准型 J，则

$$\chi_{\mathcal{A}}(t) = (t-\lambda_1)^{r_1} \cdot (t-\lambda_2)^{r_2} \cdots (t-\lambda_p)^{r_p}, \quad \mu_{\mathcal{A}}(t) = (t-\lambda_1)^{q_1} \cdot (t-\lambda_2)^{q_2} \cdots (t-\lambda_p)^{q_p}$$

其中 $q_i = \max\{k_{i,1}, k_{i,2}, \ldots, k_{i,s_i}\}$，即 q_i 是特征值 λ_i 所对应的若尔当块的最大的阶数。

证明. 定理前半部分可以直接计算得到；而 $\mu_{\mathcal{A}}(t)$ 和 $\chi_{\mathcal{A}}(t)$ 具有相同的根 (推论2.4.2)，故我们只需考虑每个 $t-\lambda_i$ 的次数。将 V 分解成循环不变子空间的直和，我们就可以看到：在每个循环不变子空间上，$\mathcal{A}-\lambda_i\mathcal{E}$ 都幂零，并且幂零指数就是循环不变子空间的维数。因此，$\mathcal{A}-\lambda_i\mathcal{E}$ 在根子空间上的幂零指数就是其分解出的循环不变子空间的最大维数，也即 λ_i 对应的若尔当块的最大阶数。这样我们就完成了证明。 □

由此，我们可以总结线性算子 (或矩阵) 可对角化的等价条件如下：

定理 2.5.8. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$，其全部的两两不同的特征值为 $\lambda_1, \ldots, \lambda_p$。则以下条件等价：

(i) $\mathcal{A}$ 是可对角化的 (定义2.3.6)。

(ii) $\mathcal{A}$ 的若尔当标准型中若尔当块都是 1 阶的。

(iii) $\mathcal{A}$ 有 n 个线性无关的特征向量。

(iv) $V = V^{\lambda_1} \oplus V^{\lambda_2} \oplus \cdots \oplus V^{\lambda_p}$。

(v) $n = \sum_{i=1}^{p} \dim(V^{\lambda_i})$。

(vi) $\mu_{\mathcal{A}}(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_p)$。

其中 (i) $\Longleftrightarrow$ (ii) 是显然的，而 (ii) 与其余各条件等价都可以用定理2.5.2，定理2.5.4或定理2.5.7推出。细节留作练习。

利用若尔当标准型，我们还可以方便地计算矩阵的方幂。容易验证若尔当块 $J_m(\lambda)$ 的方幂是：

$$J_m(\lambda)^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \cdots & \binom{k}{m-2}\lambda^{k-m+2} & \binom{k}{m-1}\lambda^{k-m+1} \\ & \lambda^k & \cdots & \binom{k}{m-3}\lambda^{k-m+3} & \binom{k}{m-2}\lambda^{k-m+2} \\ & & \ddots & \vdots & \vdots \\ & & & \lambda^k & k\lambda^{k-1} \\ & & & & \lambda^k \end{pmatrix}$$

注意：我们规定当 $l > k$ 时二项系数 $\binom{k}{l} = 0$。验证上面的等式是简单的，只需要用到二项式定理即可，留作练习。因此，我们能够方便地计算一个若尔当标准型的方幂。对于一般的 $A \in M_n(\mathbb{C})$，设 $P^{-1}AP = J$，J 是若尔当标准型，则 $A^k = PJ^kP^{-1}$ 即完成了计算。

2.6 λ 矩阵理论简介与矩阵的有理标准型

我们已经比较完整地介绍了若尔当标准型理论，但是我们看到，用前面介绍的方法来计算一个矩阵的若尔当标准型是比较繁琐的。那么，是否有更简单的方法呢？这就是我们下面介绍的由主理想整环上有限生成模结构定理得到的 λ 矩阵理论。

定义 2.6.1. 设 $\mathbb{K}$ 是域，$\mathbb{K}[\lambda]$ 是 $\mathbb{K}$ 上以 λ 为未定元的一元多项式环。我们称

$$A(\lambda) = (a_{ij}(\lambda))_{n\times n} \in M_n(\mathbb{K}[\lambda]) \qquad \Big(\text{即} a_{ij}(\lambda) \in \mathbb{K}[\lambda]\Big)$$

为 $\mathbb{K}$ 上的 λ 矩阵。

与 $M_n(\mathbb{K})$ 中的矩阵的初等变换类似，我们将以下变换称为 λ 矩阵的**初等行**（**或列**）**变换**：

1. 将 $A(\lambda)$ 的两行（或列）互换，称为第 (I) 类初等变换；
2. 将 $A(\lambda)$ 的某一行（或列）乘以某个 $f(\lambda) \in \mathbb{K}[\lambda]$ 后加到另一行（或列）上，称为第 (II) 类初等变换；
3. 将 $A(\lambda)$ 的某一行（或列）乘以某个常数 $c \in \mathbb{K} \setminus \{0\}$，称为第 (III) 类初等变换。

定义 2.6.2. 如果 λ 矩阵 $A(\lambda)$ 经过初等变换可以得到 $B(\lambda)$，则称 $A(\lambda)$ 与 $B(\lambda)$ 等价。

例 2.6.1. 容易验证 $\mathbb{C}$ 上的两个 λ 矩阵

$$\begin{pmatrix} \lambda & -1 & 0 \\ 4 & \lambda-4 & 0 \\ 2 & -1 & \lambda-2 \end{pmatrix} \quad \text{和} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -(\lambda-2) & 0 \\ 0 & 0 & (\lambda-2)^2 \end{pmatrix}$$

是等价的。细节留作练习。

定理 2.6.1. 任意的 $A(\lambda) \in M_n(\mathbb{K}[\lambda])$ 都等价于以下标准型（称为 Smith 标准型）

$$S(\lambda) = \begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

其中 $r \geqslant 1$，$d_i(\lambda)(i = 1, 2, \ldots, r)$ 是首项系数为 1 的 $\mathbb{K}[\lambda]$ 中的多项式，且 $d_i(\lambda) \mid d_{i+1}(\lambda) \quad (i = 1, 2, \ldots, r-1)$。而且，$A(\lambda)$ 的 Smith 标准型是唯一的，即 $S(\lambda)$ 与所做的初等变换无关。

证明从略，感兴趣的读者可以参考《高等代数（第四版）》，王萼芳，石生明修订，高等教育出版社的第 8 章 §2。

我们称定理2.6.1中出现的 $d_1(\lambda), \ldots, d_r(\lambda)$ 为 $A(\lambda)$ 的**不变因子**，r 称为 $A(\lambda)$ 的**秩**。

定义 2.6.3. 设 λ-矩阵 $A(\lambda)$ 的秩为 r，对于每个正整数 $k, 1 \leqslant k \leqslant r$，则 $A(\lambda)$ 中一定存在 $k \times k$ 阶子矩阵的行列式不为 0。我们把 $A(\lambda)$ 中全部 $k \times k$ 阶子矩阵的行列式的首项系数为 1 的最大公因式 $D_k(\lambda)$ 称为 $A(\lambda)$ 的 k 阶**行列式因子**。

显然 $A(\lambda)$ 的不变因子与行列式因子之间有如下关系：

$$d_1(\lambda) = D_1(\lambda), d_2(\lambda) = \frac{D_2(\lambda)}{D_1(\lambda)}, \cdots\cdots, d_r(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}$$

现在我们回到域上的矩阵上来。我们把 λ 矩阵 $\lambda E - A$(称为 A 的**特征矩阵**) 的不变因子称为 $A \in M_n(\mathbb{K})$ 的不变因子 [1]。需要注意的是，特征矩阵总是满秩的。于是我们可以证明如下推论：

推论 2.6.1. 任取 $A \in M_n(\mathbb{K})$，则存在可逆[2]的 λ 矩阵 $P(\lambda)$，$Q(\lambda)$ 使得

$$P(\lambda)(\lambda E_n - A)Q(\lambda) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(\lambda) & & \\ & & & & \ddots & \\ & & & & & d_s(\lambda) \end{pmatrix}.$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda) \quad (i = 1, 2, \ldots, s-1)$ 是首一多项式。

证明从略，可参考《高等代数 (第四版)》，王萼芳，石生明修订，高等教育出版社的第 8 章 §3 ~ 4。

定义 2.6.4. 条件同推论2.6.1。将 $d_1(\lambda), \ldots, d_s(\lambda)$ 在域 $\mathbb{K}$ 上进行完全的因式分解，即我们可设

$$\begin{aligned} d_1(\lambda) &= p_1(\lambda)^{k_{11}} p_2(\lambda)^{k_{12}} \cdots p_t(\lambda)^{k_{1t}}, \\ d_2(\lambda) &= p_1(\lambda)^{k_{21}} p_2(\lambda)^{k_{22}} \cdots p_t(\lambda)^{k_{2t}}, \\ \cdots\cdots & \\ d_s(\lambda) &= p_1(\lambda)^{k_{s1}} p_2(\lambda)^{k_{s2}} \cdots p_t(\lambda)^{k_{st}}. \end{aligned}$$

其中 $k_{ij} \geqslant 0$，并且每个 $p_j(\lambda)$ 都是 $\mathbb{K}$ 上的不可约多项式。我们称所有的 $\left\{p_j(\lambda)^{k_{ij}} \mid k_{ij} > 0\right\}$ 为矩阵 A 的**初等因子组**，其中的每个元素称为 A 的一个初等因子。

例 2.6.2. 可以验证 $\mathbb{C}$ 上的 m 阶若尔当块 $J_m(\lambda_0)$ 的初等因子组为 $\{(\lambda - \lambda_0)^m\}$；因此，$\mathbb{C}$ 上的若尔当标准型 $J = \begin{pmatrix} J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_s}(\lambda_s) \end{pmatrix}$ 的初等因子组为 $\{(\lambda - \lambda_1)^{m_1}, (\lambda - \lambda_2)^{m_2}, \ldots, (\lambda - \lambda_s)^{m_s}\}$。细节留作思考。

现在我们不加证明地给出以下结论：

---

[1] 为了方便，我们通常在 $\lambda E - A$ 中把平凡的不变因子 1 去掉，不说成 A 的不变因子。

[2] λ 矩阵 $A(\lambda)$ 可逆 $\iff \exists\, \lambda$ 矩阵 $B(\lambda)$ 使得 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = E$。

定理 2.6.2. 两个 n 阶的非零 λ 矩阵等价的充分必要条件是它们有相同的不变因子。

定理 2.6.3. 任取 $A, B \in M_n(\mathbb{K})$，则 A, B 相似 $\iff$ A, B 的不变因子相同。

推论 2.6.2. 任取 $A, B \in M_n(\mathbb{K})$，则 A, B 相似 $\iff$ A, B 的初等因子组因子相同。

我们可以利用以上结论，结合若尔当标准型的初等因子组来证明若尔当标准型的存在唯一性，并求出一般的复矩阵的若尔当标准型，有关细节可参考《高等代数 (第四版)》，王萼芳，石生明修订，高等教育出版社的第 8 章 §6。

对于非代数闭域，我们也可以有类似于若尔当标准型的结论，称为**有理标准型**。我们不加说明地给出以下定义：

定义 2.6.5. 对域 $\mathbb{K}$ 上的一个多项式

$$d(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n,$$

我们称矩阵

$$A = \begin{pmatrix} 0 & 0 & & 0 & -a_n \\ 1 & 0 & & 0 & -a_{n,-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 1 & -a_n \end{pmatrix}$$

为多项式 $d(\lambda)$ 的友矩阵 (companion matrix)。

定义 2.6.6. 下列分块对角矩阵

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix},$$

其中 A_i 分别是域 $\mathbb{K}$ 上某些多项式 $d_i(\lambda)(i = 1, 2, \cdots, s)$ 的友矩阵，且满足 $d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_s(\lambda)$，则我们称 A 为 $\mathbb{K}$ 上的一个**有理标准型**。

利用 λ 矩阵理论，我们可以证明：

定理 2.6.4. 设 $\mathbb{K}$ 是域 (不一定代数闭)，$B \in M_n(\mathbb{K})$，则在 $\mathbb{K}$ 上 B 相似于唯一的一个有理标准型，称为矩阵 B 的有理标准型。

证明细节可参考《高等代数 (第四版)》，王萼芳，石生明修订，高等教育出版社的第 8 章 §7。

以上我们关于 λ 矩阵理论和有理标准型的介绍是非常简略的，感兴趣的读者可以自行阅读其他的高等代数教材。有关更一般的主理想整环上的有限生成模的结构，读者可以参考《代数导引》，万哲先，科学出版社的第 6 章；《Algebra》，Thomas W. Hungerford，GTM73 的 §4.6 或者《Advanced Linear Algebra》，Steven Roman，GTM135 的 Chapter6。

2.7 习题

向量空间上的线性映射

1. 把坐标列 $\mathbf{x} = (x_1, x_2, x_3, x_4)^t$ 记成矩阵 $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \in M_2(\mathbb{K})$ 形式, 再取一个固定的矩阵 $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in M_2(\mathbb{K})$, 定义两个线性映射

$$f_L : X \mapsto A\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} = \begin{pmatrix} x_1' & x_2' \\ x_3' & x_4' \end{pmatrix} = X',$$

$$f_R : X \mapsto \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}A = \begin{pmatrix} x_1'' & x_2'' \\ x_3'' & x_4'' \end{pmatrix} = X'',$$

它们分别对应矩阵 M_{f_L} 和 M_{f_R}. 试验证

$$M_{f_L} = \begin{pmatrix} a_1 & 0 & a_2 & 0 \\ 0 & a_1 & 0 & a_2 \\ a_3 & 0 & a_4 & 0 \\ 0 & a_3 & 0 & a_4 \end{pmatrix}, M_{f_R} = \begin{pmatrix} a_1 & a_3 & 0 & 0 \\ a_2 & a_4 & 0 & 0 \\ 0 & 0 & a_1 & a_3 \\ 0 & 0 & a_2 & a_4 \end{pmatrix}.$$

2. 设 $f : V \to W$ 是线性映射. 证明:
   (1) V 的任意子空间 U 的像 $f(U) = \{f(u) \mid u \in U\}$ 是 W 的子空间;
   (2) W 的任意子空间 Z 的逆像 $f^{-1}(Z) = \{v \in V \mid f(v) \in Z\}$ 是 V 的子空间.

3. 设 $f : V \to W$ 是线性映射. 证明存在 V 的基和 W 的基, 使得 f 关于这两个基的矩阵具有如下形式

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

其中 I_r 是 r 阶单位矩阵, $r = \operatorname{rank} f$.

4. 给出一个非线性函数 $f : \mathbb{R}^2 \to \mathbb{R}$ 使得对任意的 $a \in \mathbf{R}$ 和 $\mathbf{v} \in \mathbb{R}^2$ 有

$$f(a\mathbf{v}) = af(\mathbf{v}).$$

5. 设 U 是有限维向量空间 V 的子空间. 证明:
   (1) 任何线性映射 $g : U \to W$ 都可以延拓成 V 到 W 的线性映射, 即存在线性映射 $f : V \to W$ 使得对任何 $\mathbf{u} \in U$ 有 $f(\mathbf{u}) = g(\mathbf{u})$.
   (2) 自然映射 $V \to V/U, \mathbf{x} \to \mathbf{x} + U$ 是线性映射.
   (3) 存在线性映射 $h : V \to U$ 使得 h 限制在 U 上是恒等映射.

6. 设 F 是 q 元域. 找出
   (1) 从 F^n 到 F^k 的线性映射的个数;
   (2) 从 F^n 到 F^k 的单线性映射的个数;
   (3) 从 F^n 到 F^k 的满线性映射的个数.

7. 设 $f : V \to W$ 是线性映射, 证明: 如果 V 是无限维向量空间, 则 $\ker f$ 和 $\operatorname{Im} f$ 中至少有一个是无限维的.

8. 设 P_n 是域 $\mathbb{K}$ 上次数不超过 $n-1$ 的多项式全体形成的空间, 定义 $\varphi : P_n \to P_n$, $u(t) \mapsto tu'(t) - u(t)$, 证明 φ 是线性映射, 并求出其核与秩.

9. 设 $A \in M_n(\mathbb{K})$ 是非退化矩阵, 证明: 映射 $f_C : M_n(\mathbb{K}) \to M_n(\mathbb{K}), X \to C^{-1}XC$ 是线性映射且 $f_C(XY) = f_C(X)f_C(Y)$.

10. 设 V 是区间 $[-\pi, \pi]$ 上所有的连续实函数形成的向量空间. 命 U 为 V 中满足如下三个条件的所有函数 f 形成的集合:

$$\int_{-\pi}^{\pi} f(t)\mathrm{d}t = 0, \quad \int_{-\pi}^{\pi} \mathrm{f(t)\cos tdt} = 0, \quad \int_{-\pi}^{\pi} \mathrm{f(t)\sin tdt} = 0.$$

证明:

(1) U 是 V 的子空间;

(2) U 包含所有的函数 $\cos nx$ 和 $\sin nx, n = 2, 3, \cdots$;

(3) U 是无限维空间.

定义线性映射 $T : V \to V$ 如下: 对 $f \in V, g = T(f)$, 有

$$g(x) = \int_{-\pi}^{\pi} [1 + \cos(x - t)] f(t)\mathrm{d}t.$$

(4) 证明 $\operatorname{Im} T$ 是有限维向量空间并求出它的一个基;

(5) 求出 T 的核;

(6) 求所有的实数 $c \neq 0$ 和 V 中所有的非零元素 f 使得 $T(f) = cf$. (注意这样的 f 在 T 的像中，)

线性算子代数

1. 对次数不超过 n 的多项式形成的空间, 验证映射 $f(x) \to f(x+1) - f(x)$ 是线性算子.

2. (1) 验证幂零算子和幂等算子的特征多项式;

(2) 证明第 1 题中的算子是幂零算子. 特别地，如果基域的特征是 0, 那么其幂零指数是 $n+1$.

3. 证明微分算子

(1) 在次数不超过 n 的多项式形成的空间上是退化的;

(2) 在函数 $\cos t, \sin t, \cdots, \cos nt, \sin nt$ 张成的空间上是非退化的.

4. 证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}, \mathcal{B}$, 有等式

$$\operatorname{rank} \mathcal{A} = \operatorname{rank} \mathcal{B}\mathcal{A} + \dim(\operatorname{Im} \mathcal{A} \cap \ker \mathcal{B}).$$

5. 利用上一题的等式证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}, \mathcal{B}, \mathcal{C}$, 有

$$\operatorname{rank} \mathcal{B}\mathcal{A} + \operatorname{rank} \mathcal{A}\mathcal{C} \leqslant \operatorname{rank} \mathcal{A} + \operatorname{rank} \mathcal{B}\mathcal{A}\mathcal{C}.$$

6. 证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}$ 和任意正整数 i 都有等式

$$\dim\left(\operatorname{Im}\mathcal{A}^{i-1}\cap\ker\mathcal{A}\right)=\dim\ker\mathcal{A}^{i}-\dim\ker\mathcal{A}^{i-1}.$$

(由于总是约定 $\mathcal{A}^0=\mathcal{E}$, 所以对 $i=1$, 公式是显然的.)

7. 证明: 域 K 上的两个同阶方阵 A, B 中如果有一个是非退化的, 那么 AB 和 BA 相似. 举例说明如果 A, B 都是退化的, 则 AB 和 BA 可以不相似.

8. 命 $\mathrm{gl}_n(\mathbb{K})$ 为代数闭域 $\mathbb{K}$ 上的 n 阶方阵全体形成的向量空间. 这个空间中的元素 A 称为半单的如果它与对角矩阵相似. 设 $A\in\mathrm{gl}_n(\mathbb{K})$. 证明:
(1) 映射 $f_A: X\to AX-XA$ 是 $\mathrm{gl}_n(\mathbb{K})$ 上的线性算子;
(2) f_A 是退化的;
(3) 如果 A 是半单的, 我么 f_A 在某个基下的矩阵是对角矩阵.

9. 证明: 两个 n 阶实矩阵如果在复数域上相似, 则一定在实数域上相似.

10. 设 $\mathbb{K}$ 是 q 元域, V 是 $\mathbb{K}$ 上的 n 维向量空间. 求出 V 上自同构的个数. 由此知道 $|\mathrm{GL}_n(\mathbb{K})|$. 进而求出 $\mathrm{GL}_n(\mathbb{K})$ 中行列式为 1 的矩阵全体形成的群 $\mathrm{SL}_n(\mathbb{K})$ 的阶数.

11. 设 $\mathcal{A}$ 是二维向量空间上的线性算子. 证明多项式 $t^2-(\operatorname{tr}\mathcal{A})t+\det\mathcal{A}$ 零化 $\mathcal{A}$, 即 $\mathcal{A}^2-(\operatorname{tr}\mathcal{A})\mathcal{A}+(\det\mathcal{A})\mathcal{E}=0$.

12. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子. 称多项式 $f(t)$ 是算子 $\mathcal{A}$ 对 $\mathbf{v}\in V$ 的零化多项式如果 $f(\mathcal{A})\mathbf{v}=0$. 在算子 $\mathcal{A}$ 对 $\mathbf{v}$ 的零化多项式中次数最小且首项系数为 1 者称为 $\mathcal{A}$ 对向量 v 的极小多项式, 记为 $\mu_{\mathcal{A},\mathbf{v}}(t)$. 假设基域 $\mathbb{K}$ 是无限域. 证明:
(1) $\mu_{\mathcal{A},\mathbf{v}}(t)\mid\mu_{\mathcal{A}}(t)$;
(2) 存在向量 $\mathbf{u}\in V$ 使得 $\mu_{\mathcal{A},\mathbf{u}}(t)=\mu_{\mathcal{A}}(t)$.

13. 设 A 是特征为 0 的域 $\mathbb{K}$ 上的 n 阶方阵. 证明: 如果 A 的迹为零, 那么 A 相似于某个主对角线均取零值的矩阵 B ($B=\left(b_{ij}\right), b_{11}=b_{22}=\cdots=b_{nn}=0$).

14. 上一题中域的特征为 0 是必要条件么?

15. 设 V 是个向量空间, U, W 是它的两个子空间, 而且

$$V=V_1\oplus V_2,\quad W=W_1\oplus W_2$$

是直和分解, 其中 $W_i\subseteq V_i, i=1,2$. 其次, 再设 $\mathcal{P}_i: V\to V_i$, $\mathbf{x}=\mathbf{x}_i+\mathbf{x}_2\mapsto\mathbf{x}_i$, $i=1,2$ (称 $\mathcal{P}_i$ 是 V_i 的平行于 $V_j, j\neq i$ 的投影). 证明:
(1) 如果

$$V_1=W_1+U\cap V_1,\quad V_2=W_2+\mathcal{P}_2(U),\tag{$*$}$$

那么, $V=W+U$;
(2) 如果 $V=W+U$ 且 $\mathcal{P}_2(U)\cap W_2=\mathbf{0}$, 那么, 对于 V_1 和 V_2 的分解式 $(*)$ 总成立, 而且 $W\cap U=W_1\cap U$.

不变子空间和特征问题

1. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 证明: 如果 $V_1, \cdots, V_m$ 是 $\mathcal{A}$ 的不变子空间, 那么这些子空间的和与交都是 $\mathcal{A}$ 的不变子空间.

2. 证明: 如果 V 的子空间 U 在 V 的每个线性算子下都是不变的, 则 U 是 V 的平凡子空间, 即 $U = \{0\}$ 或 $U = V$.

3. 设 $\mathcal{A}, \mathcal{B}$ 是向量空间 V 上的线性算子. 证明: 如果 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, 那么对任意的 $\lambda \in K$, $\ker(\lambda\mathcal{E} - \mathcal{A})$ 是 $\mathcal{B}$ 的不变子空间.

4. 设 $\mathcal{A}, \mathcal{B}$ 是向量空间 V 上的线性算子 (基域 $\mathbb{K}$ 是代数闭域). 证明: 如果 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$, 那么 $\mathcal{A}, \mathcal{B}$ 有共同的特征向量.

5. 证明: 如果 $\mathcal{A}$ 是可逆线性算子, 那么 $\mathcal{A}$ 的不变子空间也是 $\mathcal{A}^{-1}$ 的不变子空间. 特别, $\mathcal{A}$ 的特征向量也是 $\mathcal{A}^{-1}$ 的特征向量.

6. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 其平方是恒等算子. 证明: 对任意的向量 $\mathbf{v} \in V$, 向量 $\mathbf{v} - \mathcal{A}\mathbf{v}$ 是零向量或以 -1 为特征值的特征向量. 如果 $\mathrm{char}\mathbb{K} \neq 2$, 证明 V 是特征空间 V^1 和 V^{-1} 的直和.

7. 设 $\mathcal{A}$ 是向量空间 V 上的幂等线性算子 (即 $\mathcal{A}^2 = \mathcal{A}$). 证明:
   (1) $\mathcal{B} = \mathcal{E} - \mathcal{A}$ 也是幂等线性算子, 而且 $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}, \mathcal{E} = \mathcal{A} + \mathcal{B}$;
   (2) $\mathrm{Im}\,\mathcal{A}$ 是 $\mathcal{A}$ 的以 1 为特征值的特征空间 $V^1, \ker\mathcal{A}$ 是 $\mathcal{A}$ 的以 0 为特征值的特征空间;
   (3) $\ker\mathcal{A} = \mathrm{Im}\,\mathcal{B}$;
   (4) $V = \mathrm{Im}\,\mathcal{A} \oplus \ker\mathcal{A} = \mathrm{Im}\,\mathcal{A} \oplus \mathrm{Im}(\mathcal{E} - \mathcal{A})$.

8. 更一般地, 设 $\mathcal{A}_1, \cdots, \mathcal{A}_{m-1}$ 是向量空间 V 上的幂等线性算子, 且相互正交 (即 $\mathcal{A}_i\mathcal{A}_j = \mathcal{A}_j\mathcal{A}_i = \mathcal{O}$ 如果 $i \neq j$). 证明:
   (1) $\mathcal{A} = \mathcal{A}_1 + \cdots + \mathcal{A}_{m-1}$ 是幂等线性算子, 且 $\mathcal{A}\mathcal{A}_i = \mathcal{A}_i\mathcal{A} = \mathcal{A}_i, 1 \leqslant i \leqslant m-1$;
   (2) 令 $A_m = \mathcal{E} - \mathcal{A}$, 则有 $\mathcal{A}_m\mathcal{A}_i = \mathcal{A}_i\mathcal{A}_m = \mathcal{O}, \quad 1 \leqslant i \leqslant m-1$;
   (3) $V = \mathrm{Im}\,\mathcal{A}_1 \oplus \mathrm{Im}\,\mathcal{A}_2 \oplus \cdots \oplus \mathrm{Im}\,\mathcal{A}_m$.

9. 设 $\mathcal{D} : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ 是域 $\mathbb{K}$ 上的 n 阶方阵空间上的非零线性算子. 证明: 如果它保持乘法, 即对所有 $A, B \in M_n(\mathbb{K})$ 有

$$\mathcal{D}(AB) = \mathcal{D}(A)\mathcal{D}(B),$$

   那么, 存在非退化矩阵 $C \in M_n(\mathbb{K})$ 使得 $\mathcal{D} = f_C$ (见习题” 向量空间上的线性映射” 的第 9 题). 找出 f_C 的以 1 为特征值的特征空间。

10. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 对某个正整数 k 有 $\mathrm{Im}\,\mathcal{A}^k = \mathrm{Im}\,\mathcal{A}^{k+1}$, 证明: 在这种情形有 $V = \ker\mathcal{A}^k \oplus \mathrm{Im}\,\mathcal{A}^k$ 是 $\mathcal{A}$ 的两个不变子空间的直和.

11. 找出下列线性算子的特征值和特征空间:
   (1) 次数不超过 n 的实多项式空间上的微分算子;
   (2) 次数不超过 n 的实多项式空间上的算子 $t\frac{\mathrm{d}}{\mathrm{d}t}$;
   (3) 次数不超过 n 的实多项式空间上的算子 $\frac{1}{t}\int_0^t f(x)\mathrm{d}x$;

(4) 方阵空间 $M_n(K)$ 上的算子 XX^t; (5) 函数空间 $\langle \cos t, \sin t, \cdots, \cos mt, \sin mt\rangle$ 上的算子 $f \to \frac{\mathrm{d}^n f}{\mathrm{d}t^n}$;
(6) 函数空间 $\langle \cos t, \sin t, \cdots, \cos mt, \sin mt\rangle$ 上的算子 $f \to \int_0^t f(x)\mathrm{d}x$.

12. 证明: 次数不超过 n 的实多项式空间上的算子 $f(t) \to f(at+b)$ 的特征值是 $1, a, \ldots, a^n$.

13. 证明: 如果 λ^2 是线性算子 $\mathcal{A}^2$ 的特征值, 那么 λ 和 $-\lambda$ 中有一个是 $\mathcal{A}$ 的特征值.

14. 设 $\mathbb{K}$ 是代数闭域, $A, B \in M_n(\mathbb{K})$ 且 $AB = BA$, 求证: 存在可逆矩阵 $P \in M_n(\mathbb{K})$ 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵 (这是 Lie 定理在线性代数中的特殊情形).

15. 设 A, B 是域 $\mathbb{K}$ 上的 n 阶方阵, 且 $AB = BA, B$ 幂零. 证明: $\det(tE-(A+B)) = \det(tE-A)$.

16. 证明: 如果 n 维向量空间上的线性算子 $\mathcal{E}, \mathcal{A}, \cdots, \mathcal{A}^{n-1}$ 线性无关, 那么, 存在向量 $v \in V$ 使得
$$V = \left\langle v, \mathcal{A}v, \cdots, \mathcal{A}^{n-1}v \right\rangle.$$
(此时说 V 是 $\mathcal{A}$ 的循环空间。) 而且, $\mathcal{A}$ 的特征多项式就是 $\mathcal{A}$ 的极小多项式.

17. 设 A 是 n 阶实方阵, 没有实的特征根, 特别, n 是偶数且 A 可逆. 证明: 存在实矩阵 B 使得 $AB = BA$ 且 $B^2 = -E$, 其中 E 是单位矩阵.

18. 证明: 对任意的方阵 $A, B \in M_n(\mathbb{K})$, 方阵 AB 和 BA 的特征多项式相同. 举例说明 AB 和 BA 的极小多项式可以不同.

19. 利用简单的关系式
$$A = a_0E + a_1B + a_2B^2, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$
求出复数域上的循环矩阵
$$A = \begin{pmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{pmatrix}$$
的特征根.

20. 设 $A, B \in M_n(\mathbb{K})$, $\mathrm{char}K \neq 2$. 若有 $D = \mathrm{diag}\,(\theta_1, \cdots, \theta_n)\,, \theta_i = \pm 1$, 使得 $B = DA$, 则记为 $B \sim A$. 显然 $\sim$ 是等价关系. 如果 $S(A)$ 表含 A 的等价类, 那么 $\mathrm{Card}\,S(A) \leqslant 2^n$. 证明: 每个等价类中都含有这样的矩阵, 1 不是其特征根.

21. 计算特征多项式并求出复特征值和特征向量:
(1) $\begin{pmatrix} -2 & 2 \\ -2 & 3 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$; (3) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$;
(4) $\begin{pmatrix} -2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$; (5) $\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$.

22. 方阵 A 和 A^t 是否有相同的特征向量? 是否有相同的特征值?

23. 设 $\mathcal{A},\mathcal{B}$ 是有限维复向量空间上的线性算子. 证明: 如果 $\mathcal{AB}-\mathcal{BA}$ 的秩不超过 1 , 那么这两个算子有共同的特征向量.

商算子和对偶算子

1. 利用推论2.4.2求出下列矩阵的极小多项式 (其定义和算子的极小多项式的定义类似):
(1) $\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 4 & -2 & 2 \\ -5 & 7 & -5 \\ -6 & 7 & -4 \end{pmatrix}$.

2. Hamilton-Cayley 定理 (定理2.4.3) 可以用于计算线性算子和方阵代入多项式得到的结果. 设 $\mathcal{A}$ 是向量空间上 V 上的线性算子, $f(t)$ 是多项式. 带余除法给出

$$f(t) = q(t)\chi_{\mathcal{A}}(t) + r(t).$$

于是

$$f(\mathcal{A}) = r(\mathcal{A}).$$

如果 λ 是 $\chi_{\mathcal{A}}(t)$ 的 k 重根, 那么, 对 $i = 0, 1, \cdots, k-1$, 多项式 f 的 i 阶导数在 λ 处的值 $f^{(i)}(\lambda)$ 等于 r 的 i 阶导数在 λ 处的值 $r^{(i)}(\lambda)$. (约定 $f^{(0)} = f$.) 这个性质对计算 r 是有用的. 计算:
(1) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{100}$;
(2) $\begin{pmatrix} 7 & -4 \\ 14 & -8 \end{pmatrix}^{66}$;
(3) $\begin{pmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{pmatrix}^{50}$.

3. 假设基域是代数闭域. 证明: 线性算子 $\mathcal{A}$ 是幂零的当且仅当其特征多项式没有非零根.

若尔当标准型理论介绍

1. 非零的 4 阶幂零方阵的若尔当标准形只有下面四个:

$$\begin{aligned} A_1 &= J_2(0) + J_1(0) + J_1(0), & A_2 &= J_2(0) + J_2(0), \\ A_3 &= J_3(0) + J_1(0), & A_4 &= J_4(0). \end{aligned}$$

下列矩阵都是幂零的, 它们各相似于哪个 A_i :

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 7 & 0 & 0 \end{pmatrix}.$$

2. (1) 如果矩阵的特征多项式为 $(t+3)^2(t-2)^3$, 它的若尔当标准形有哪些可能?
(2) 如果矩阵的特征多项式为 $(t+3)^2(t-2)^3$, 以 -3 为特征值的特征空间的维数是 1 , 以 2 为特征值的特征空间的维数是 2 , 这个矩阵的若尔当标准形有哪些可能?

3. 求出下列矩阵的若尔当标准形: (1) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$,

(2) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$, (3) $\begin{pmatrix} -5 & -3 & -2 & 4 \\ 2 & 0 & 1 & -1 \\ 10 & 7 & -4 & 9 \\ 2 & 0 & 1 & 0 \end{pmatrix}$.

4. (1) 验证矩阵

$$A = \begin{pmatrix} 6 & 2 & -2 \\ -2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

有相同的特征多项式;

(2) 求出它们的极小多项式;

(3) 求出它们的若尔当标准形.

5. 设域 $\mathbb{K}$ 的特征为 0, A 是 $\mathbb{K}$ 上的 n 阶方阵. 证明: A 是幂零的当且仅当 $\mathrm{tr}\left(A^k\right) = 0, \ \forall 1 \leqslant k \leqslant n$.

6. 如果线性算子 $\mathcal{A}$ 的若尔当标准形只有一个若尔当块, 确定 $\mathcal{A}$ 的所有不变子空间.

7. 证明: 方阵 $A \in M_n(\mathbb{C})$ 和它的转置 A^t 相似.

8. 设 $A \in M_n(\mathbb{C})$ 的特征根都等于 1 . 证明: 对任意非雲整数 k, A 和 A^k 相似.

9. 证明: 对于矩阵 $A \in M_n(\mathbb{C})$, 等式 $A^m = E$ 成立当且仅当 A 可对角化且它的特征值都是 m 次单位根. 举例说明这个结论对 $\mathbb{Z}_p$ 上的方阵不成立, 即存在 $\mathbb{Z}_p$ 上的方阵, 其某个幂为单位矩阵, 但它不能对角化 (即不相似于对角矩阵).

10. (Jordan-Chevalley 分解) 设 $\mathcal{A}$ 是 n 维复向量空间 V 上的线性算子, 证明: 存在唯一的可对角化算子 $\mathcal{S}$ 和幂零算子 $\mathcal{N}$ 使得

$$\mathcal{A} = \mathcal{S} + \mathcal{N}, \quad \mathcal{S}\mathcal{N} = \mathcal{N}\mathcal{S}.$$

而且, $\mathcal{S}$ 和 $\mathcal{N}$ 都可以表达成 $\mathcal{A}$ 的多项式. (可对角化算子也称为半单算子.)

11. 对任意正整数 k, 计算 $(J_n(\lambda))^k$.

12. 如果线性算子 $\mathcal{A}$ 在基 $e_1, \cdots, e_n$ 下的矩阵是 $\begin{pmatrix} 3 & 2 & 3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$. $\begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 2 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}$. 求出 $\mathcal{A}$ 的一个若尔当基和 $\mathcal{A}$ 在若尔当基下的矩阵.

13. 证明 n 维复向量空间上与给定线性变换 $\mathcal{A}$ 可交换的线性变换组成的集合是维数 $\geqslant n$ 的一个向量空间.

14. (双中心化定理) 证明如果复向量空间上的一个线性变换 $\mathcal{B}$ 与跟线性变换 $\mathcal{A}$ 可交换的任一线性变换可交换, 那么 $\mathcal{B}$ 是 $\mathcal{A}$ 的一个多项式.

15. 证明: 如果复向量空间 V 上的线性算子 $\mathcal{A}$ 和 $\mathcal{B}$ 满足关系 $\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A}=\mathcal{B}$, 那么 $\mathcal{B}$ 是幂零的.

16. 算子 $\mathcal{A}=\frac{\partial}{\partial x}+\frac{\partial}{\partial y}$ 作用在次数不超过 n 的二元复多项式形成的空间上. 找出 $\mathcal{A}$ 的若尔当标准形.

17. 解矩阵方程: (1) $X^2=\begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}$,

(2) $X^2=\begin{pmatrix} 6 & 2 \\ 3 & 7 \end{pmatrix}$.

18. 命 $\mathfrak{N}$ 为 $M_n(\mathbb{C})$ 中的幂零矩阵全体, $\mathfrak{U}$ 为 $M_n(\mathbb{C})$ 中特征值都等于 1 的矩阵全体. 证明映射

$$\exp : \mathfrak{N}\to\mathfrak{U},\quad A\to\exp A=E+A+\frac{A^2}{2!}+\cdots+\frac{A^{n-1}}{(n-1)!}$$

是双射并求出这个双射的逆映射.

第三章　内积空间及其上的线性算子

这一章我们将会回到 $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 上来讨论问题。前面我们已经把一般的域 $\mathbb{K}$ 上的双线性型和线性算子的性质比较完整地讲清楚了，但是如果落回到 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 上，与通常的空间相比 (即我们中学学过的 $\mathbb{R}^3$ 或数学分析中的 $\mathbb{R}^n$)，我们会发现前面的内容里少了一些结构：夹角和距离。这就是我们这一章的主要任务：在带有夹角 (即内积) 或距离的向量空间上进一步讨论双线性型和线性算子的性质。对应到矩阵层面，我们则是讨论一个实 (或复) 矩阵何时能够合同并相似到一个对角矩阵。下面我们正式开始介绍本章的内容。

3.1　欧几里得空间

定义 3.1.1. 设 V 是实数域 $\mathbb{R}$ 上的 n 维向量空间，我们在 $V \times V$ 上指定一个满足如下性质的函数 $<\cdot,\cdot>: V\times V \to \mathbb{R}$:

1) $\forall \mathbf{x},\mathbf{y} \in V,\quad <\mathbf{x},\mathbf{y}>=<\mathbf{y},\mathbf{x}>$;

2) $\forall \mathbf{x},\mathbf{y},\mathbf{z} \in V,\quad <\mathbf{x}+\mathbf{y},\mathbf{z}>=<\mathbf{x},\mathbf{z}>+<\mathbf{y},\mathbf{z}>$;

3) $\forall \mathbf{x},\mathbf{y} \in V$ 及 $\lambda \in \mathbb{R},\quad <\lambda\mathbf{x},\mathbf{y}>=\lambda<\mathbf{x},\mathbf{y}>$;

4) $\forall \mathbf{x} \in V,\quad <\mathbf{x},\mathbf{x}>\geqslant 0$，并且等号当且仅当 $\mathbf{x}=\mathbf{0}$ 时成立。

也即 $<\cdot,\cdot>$ 是 V 上的一个对称双线性型，其对应的二次型是正定的。则我们称 $(V,<\cdot,\cdot>)$ 是 $\mathbb{R}$ 上的 n 维实内积空间或 n 维欧几里得空间 (Euclid space)，简称欧氏空间[1]。$<\cdot,\cdot>$ 称为 V 上的内积 (inner product)。当内积已经明确时，我们通常省略内积，而直接称 V 是 n 维欧氏空间。

例 3.1.1.　(1) 令 $V=\mathbb{R}^n$，任取 $\mathbf{x}=(x_1,\ldots,x_n)^t$，$\mathbf{y}=(y_1,\ldots,y_n)^t \in V$，定义：

$$<\mathbf{x},\mathbf{y}>=\sum_{i=1}^{n} x_i y_i$$

则容易验证 $<\cdot,\cdot>$ 满足上面的定义，从而是一个内积。我们把这个内积称为 $\mathbb{R}^n$ 上的**标准内积** (standard inner product)，称 $(\mathbb{R}^n,<\cdot,\cdot>)$ 是 n 维标准欧氏空间。

(2) 令 $V=M_n(\mathbb{R})$，任取 $A,B\in V$，定义 $<A,B>=\mathrm{tr}(AB^t)$，则容易验证 $<A,B>$ 是一个内积，从而 $(M_n(\mathbb{R}),<\cdot,\cdot>)$ 是一个 n^2 维的欧氏空间。

(3) 固定 $a,b\in\mathbb{R}$，$a<b$，在 $\mathbb{R}[x]_n$(次数不超过 $n-1$ 的实多项式的构成的向量空间) 上，任取 $f(x),g(x)\in\mathbb{R}[x]_n$，定义：

$$<f,g>=\int_a^b f(t)g(t)\,\mathrm{d}t.$$

则容易验证这是一个内积，即 $(\mathbb{R}[x]_n,<\cdot,\cdot>)$ 是 n 维欧氏空间。

---

[1] 无穷维的实向量空间上如果有一个满足上述条件的内积，那么我们仍然称其为实内积空间，但只有有限维的内积空间才被称为欧氏空间。

在一个实内积空间上，由定义3.1.1，任何一个向量和自身的内积都是非负的，从而我们有下面的定义：

定义 3.1.2. 设 V 是实内积空间，任取 $\mathbf{x} \in V$，则我们可以定义 $\mathbf{x}$ 的范数 (或称为长度，norm) 为：$\|\mathbf{x}\| = \sqrt{<\mathbf{x},\mathbf{x}>}$。于是，任取 $\mathbf{x},\mathbf{y} \in V$，我们称 $\mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{<\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}>}$ 为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 的距离 (distance)。特别地，如果向量 $\mathbf{x}$ 的范数是 1，则称 $\mathbf{x}$ 是一个单位向量 (unit vector)。

例 3.1.2. 在 n 维标准欧氏空间 $\mathbb{R}^n$ 上，向量 $\mathbf{x} = (x_1,\ldots,x_n)^t$ 的范数是 $\|\mathbf{x}\| = \sqrt{x_1^2+\cdots+x_n^2}$，任何两个向量 $\mathbf{x} = (x_1,\ldots,x_n)^t$，$\mathbf{y} = (y_1,\ldots,y_n)^t$ 的距离是 $\mathrm{d}(\mathbf{x},\mathbf{y}) = \sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2}$。容易看出当 $n=2$ 或 3 时上面的范数和距离的定义与我们中学时的定义是一致的。

我们在中学时就已经学过 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 上的的柯西不等式了。下面我们将它推广到一般形式。

定理 3.1.1 (Cauchy-Buniakowsky-Schwarz's inequation, 柯西-布尼亚科夫斯基-施瓦茨不等式). [1]
设 $(V,<\cdot,\cdot>)$ 是实内积空间，任取 $\mathbf{x},\mathbf{y} \in V$，则 $|<\mathbf{x},\mathbf{y}>| \leqslant \|\mathbf{x}\|\cdot\|\mathbf{y}\|$，并且等号成立的充分必要条件是 $\mathbf{x} = \mathbf{0}$ 或 $\exists\lambda \in \mathbb{R}$ 使得 $\mathbf{y} = \lambda\mathbf{x}$。

证明. 如果 $\mathbf{x}$ 或 $\mathbf{y}$ 是 $\mathbf{0}$，则不等式两边都是 0，显然成立。下设 $\mathbf{x}$ 和 $\mathbf{y}$ 都不是 $\mathbf{0}$。由内积的定义，对 $\forall\lambda \in \mathbb{R}$ 及 $\mathbf{x},\mathbf{y} \in V$，我们有 $<\lambda\mathbf{x}-\mathbf{y},\lambda\mathbf{x}-\mathbf{y}> \geqslant 0$，并且等号当且仅当 $\lambda\mathbf{x}-\mathbf{y} = \mathbf{0}$ 时成立。将 $<\lambda\mathbf{x}-\mathbf{y},\lambda\mathbf{x}-\mathbf{y}> \geqslant 0$ 利用内积的对称和双线性性质展开可得

$$<\mathbf{x},\mathbf{x}>\lambda^2 - 2<\mathbf{x},\mathbf{y}>\lambda + <\mathbf{y},\mathbf{y}> \geqslant 0 \text{ 对}\forall\lambda \in \mathbb{R}\text{成立}. \tag{3.1.1}$$

这是一个关于 λ 的一元二次不等式，其恒非负的充要条件是判别式

$$4<\mathbf{x},\mathbf{y}>^2 - 4<\mathbf{x},\mathbf{x}>\cdot<\mathbf{y},\mathbf{y}>\leqslant 0,$$

移项并对不等式两边开平方即得到 CBS 不等式。下面讨论等号成立条件。如果 CBS 不等式的等号成立并且 $\mathbf{x},\mathbf{y}$ 都不是 $\mathbf{0}$，则我们立刻可知式 (3.1.1) 不等号左边的判别式是 0，这说明抛物线 $<\mathbf{x},\mathbf{x}>\lambda^2 - 2<\mathbf{x},\mathbf{y}>\lambda + <\mathbf{y},\mathbf{y}>$ 与直线 $\lambda = 0$ 相切，即存在 $\lambda \in \mathbb{R}$ 使得 $<\lambda\mathbf{x}-\mathbf{y},\lambda\mathbf{x}-\mathbf{y}>= 0$，也即 $\lambda\mathbf{x}-\mathbf{y} = 0$。反过来，由 $\exists\lambda \in \mathbb{R}$ 使得 $\mathbf{y} = \lambda\mathbf{x}$ 得到 CBS 不等式取等号是容易的，只需注意到此时

$$\begin{aligned} |<\mathbf{x},\mathbf{y}>| &= |<\mathbf{x},\lambda\mathbf{x}>| = |\lambda|\cdot|<\mathbf{x},\mathbf{x}>|, \\ \|\mathbf{x}\|\cdot\|\mathbf{y}\| &= \sqrt{<\mathbf{x},\mathbf{x}>}\cdot\sqrt{<\lambda\mathbf{x},\lambda\mathbf{x}>} = |\lambda|\cdot|<\mathbf{x},\mathbf{x}>|. \end{aligned}$$

即可。这样我们就完成了证明。 □

注 3.1.1. 将 CBS 不等式应用到 n 维标准欧氏空间 $\mathbb{R}^n$ 上可知，任取 $x_1,\ldots,x_n$ 及 $y_1,\ldots,y_n \in \mathbb{R}$，我们有：

$$|x_1y_1+\cdots+x_ny_n| \leqslant \sqrt{x_1^2+\cdots+x_n^2}\cdot\sqrt{y_1^2+\cdots+y_n^2}.$$

---

[1] 简称 CBS 不等式。1821 年柯西最早发现了该不等式，1859 年布尼亚科夫斯基证明了其积分形式，1888 年施瓦茨给出了现代的证明。

即我们熟知的柯西不等式；将 CBS 不等式用到 $(M_n(\mathbb{R}), < A, B >= \mathrm{tr}(AB^t))$ 上可得

$$|\mathrm{tr}(AB^t)| \leqslant \sqrt{\mathrm{tr}(AA^t)}\sqrt{\mathrm{tr}(BB^t)};$$

将 CBS 不等式用到 $(\mathbb{R}[x]_n, < f, g >= \int_a^b f(t)g(t)\,\mathrm{d}t)$ 上可得

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leqslant \sqrt{\int_a^b f^2(t)\,\mathrm{d}t} \cdot \sqrt{\int_a^b g^2(t)\,\mathrm{d}t}.$$

即数学分析中的施瓦茨不等式。(实际上，我们可以将 $\mathbb{R}[x]_n$ 换成区间 $[a, b]$ 上的全体连续函数 $C[a, b]$，这仍然是一个内积空间，但不是有限维的，此时施瓦茨不等式仍然成立。)

推论 3.1.1 (三角不等式). 设 $(V, < \cdot, \cdot >)$ 是实内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$，并且等号成立的充分必要条件为 $\exists \alpha, \beta \in \mathbb{R}$ 使得 $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$。

将范数全部改写成内积形式，整理后利用 CBS 不等式即可证明。细节留作练习。

有了以上结论，我们很容易验证范数和距离满足以下性质：

命题 3.1.1. 设 $(V, < \cdot, \cdot >)$ 是实内积空间，$\|\cdot\|$ 是该内积对应的范数，则

(1) $\forall \mathbf{x} \in V$，$\|\mathbf{x}\| \geqslant 0$，并且 $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$；

(2) $\forall \lambda \in \mathbb{R}$, $\mathbf{x} \in V$，都有 $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$；

(3) $\forall \mathbf{x}, \mathbf{y} \in V$，有 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$。

命题 3.1.2. 设 $(V, < \cdot, \cdot >)$ 是实内积空间，$\mathrm{d}(\cdot, \cdot)$ 是该内积对应的距离，则

(1) $\forall \mathbf{x}, \mathbf{y} \in V$，$\mathrm{d}(\mathbf{x}, \mathbf{y}) \geqslant 0$，并且等号当且仅当 $\mathbf{y} = \mathbf{x}$ 时成立；

(2) $\forall \mathbf{x}, \mathbf{y} \in V$，都有 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \mathrm{d}(\mathbf{y}, \mathbf{x})$；

(3) $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$，有 $\mathrm{d}(\mathbf{x}, \mathbf{z}) \leqslant \mathrm{d}(\mathbf{x}, \mathbf{y}) + \mathrm{d}(\mathbf{y}, \mathbf{z})$。

实际上，我们也可以绕开内积而直接在向量空间上用这些性质定义范数或距离，当然此时所定义出的空间上不一定还有内积结构，我们会在下一节中继续讨论这种情形。

有了 CBS 不等式，我们就可以定义两个向量的夹角了。

定义 3.1.3. 设 $(V, < \cdot, \cdot >)$ 是实内积空间，任取非零向量 $\mathbf{x}, \mathbf{y} \in V$，我们定义 $\mathbf{x}$ 和 $\mathbf{y}$ 的夹角是如下的 θ：

$$\cos\theta = \frac{< \mathbf{x}, \mathbf{y} >}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|},\ 0 \leqslant \theta \leqslant \pi. \qquad (\text{CBS 不等式保证了}\theta\text{的存在性.})$$

其中，若 $\theta = 0$，则我们称 $\mathbf{x}, \mathbf{y}$ 同向；若 $\theta = \pi$，则我们称 $\mathbf{x}, \mathbf{y}$ 反向；若 $\theta = \dfrac{\pi}{2}$，则我们称 $\mathbf{x}, \mathbf{y}$ 正交 (orthogonal)，记作 $\mathbf{x} \perp \mathbf{y}$。特别地，我们规定 $\mathbf{0}$ 与任意 V 中的向量都正交，即 $\forall \mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} \perp \mathbf{y} \iff < \mathbf{x}, \mathbf{y} >= 0$。

于是我们立刻有

例 3.1.3 (勾股定理). 设 $(V, < \cdot, \cdot >)$ 是实内积空间，$\mathbf{x}, \mathbf{y} \in V$ 且 $\mathbf{x} \perp \mathbf{y}$，则

$$\|\mathbf{x} + \mathbf{y}\|^2 =< \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} >=< \mathbf{x}, \mathbf{x} > +2 < \mathbf{x}, \mathbf{y} > + < \mathbf{y}, \mathbf{y} >= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

引理 3.1.1. 设 $(V, <\cdot,\cdot>)$ 是实内积空间，$\mathbf{x} \in V$，$\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是子空间 U 的一组基。如果 $\mathbf{x} \perp \mathbf{u}_1, \ldots, \mathbf{x} \perp \mathbf{u}_k$ 都成立，则任取 $\mathbf{u} \in U$，都有 $\mathbf{x} \perp \mathbf{u}$。我们把这种情形记作 $\mathbf{x} \perp U$。

证明. 由于 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是子空间 U 的一组基，故任取 $\mathbf{u} \in U$，存在 $a_1, \ldots, a_k \in \mathbb{R}$ 使得 $\mathbf{u} = \sum_{i=1}^{k} a_i \mathbf{u}_i$。所以

$$< \mathbf{x}, \mathbf{u} > = \sum_{i=1}^{k} a_i < \mathbf{x}, \mathbf{u}_i > = 0,$$

即 $\mathbf{x} \perp U$。 □

如果 U_1, U_2 是 V 的两个子空间，并且任取 $\mathbf{x} \in U_1$，都有 $\mathbf{x} \perp U_2$，则我们称子空间 U_1 和 U_2 正交，记作 $U_1 \perp U_2$。显然 $U_1 \perp U_2 \Longrightarrow U_1 \cap U_2 = \{\mathbf{0}\}$。

内积和范数还满足以下性质：

命题 3.1.3. 设 $(V, <\cdot,\cdot>)$ 是实内积空间，$\mathbf{x}, \mathbf{y} \in V$，则

(1) (扩展的三角不等式)$\Big| \|\mathbf{x}\| - \|\mathbf{y}\| \Big| \leqslant \|\mathbf{x} \pm \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$；

(2) (平行四边形法则)$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$；

(3) (极化恒等式)$< \mathbf{x}, \mathbf{y} > = \dfrac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2)$。

其中 (1) 利用三角不等式即可证明，(2) 和 (3) 利用范数的定义和内积的双线性性质展开就可以证明，留作练习。

下面我们考虑 n 维欧氏空间的基底和子空间。由于欧氏空间上有内积结构，因此，我们希望它们在内积结构下还要满足一些好的性质。

定义 3.1.4. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间。如果 V 的一组基 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 满足：

$$\forall i \neq j \in \{1, \ldots, n\}, \quad < \mathbf{v}_i, \mathbf{v}_j > = 0,$$

则我们称 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组正交基 (orthogonal basis)；如果 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是正交基，并且每个 $\mathbf{e}_i$, $i \in \{1, \ldots, n\}$ 的范数都是 1，即

$$< \mathbf{e}_i, \mathbf{e}_j > = \delta_{ij},$$

则我们称 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组标准正交基 (normalized orthogonal basis)。

例 3.1.4. 显然 $\mathbb{R}^n$ 的标准基就是一组标准正交基。

引理 3.1.2. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间。如果向量组 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \subset V$ 中的向量两两正交且都不为 $\mathbf{0}$，即 $\forall i \neq j$, $< \mathbf{x}_i, \mathbf{x}_j > = 0$，则 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ 是一个线性无关组。特别地，如果还有 $m = n$，则 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ 是 V 的一组基。

证明. 设有线性组合 $a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m = \mathbf{0}$，其中 $a_1, \ldots, a_m \in \mathbb{R}$，我们只需证明 $a_1 = \cdots = a_m = 0$ 即可完成证明。注意到对每个 $i \in \{1, \ldots, m\}$，我们有

$$0 = < \mathbf{0}, \mathbf{x}_i > = < a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m, \mathbf{x}_i > = \sum_{k=1}^{m} a_k < \mathbf{x}_k, \mathbf{x}_i > = a_i < \mathbf{x}_i, \mathbf{x}_i >,$$

而 $\mathbf{x}_i \neq \mathbf{0} \Longrightarrow < \mathbf{x}_i, \mathbf{x}_i > \neq 0$，所以 $a_i = 0$ 对每个 $i \in \{1, \ldots, m\}$ 都成立，即 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 线性无关。特别地，当 $m = n$ 时，由于 V 是 n 维的，故 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 是 V 的一组基。 □

定义 3.1.5. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的子空间。则我们很容易验证：

$$W = \{\mathbf{x} \in V \mid \forall \mathbf{u} \in U, \quad <\mathbf{x}, \mathbf{u}>= 0\}$$

是 V 的子空间，并且 $V = U \oplus W$ [1] 。我们称 W 是 U 的正交补 (orthogonal complement) 空间，记作 $W = U^{\perp}$。

定理 3.1.2. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，则 V 中一定存在一组标准正交基。

证明. 我们从 V 的任意一组基出发，通过 Gram-Schmidt(**格拉姆-施密特**) **正交化**方法来构造出 V 的一组标准正交基。

任取 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组基，令

$$\begin{array}{ll}
\mathbf{u}_1 = \mathbf{v}_1 & \mathbf{e}_1 = \dfrac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 \\
\mathbf{u}_2 = \mathbf{v}_2 - \dfrac{<\mathbf{v}_2, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 & \mathbf{e}_2 = \dfrac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 \\
\mathbf{u}_3 = \mathbf{v}_3 - \dfrac{<\mathbf{v}_3, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 - \dfrac{<\mathbf{v}_3, \mathbf{u}_2>}{<\mathbf{u}_2, \mathbf{u}_2>}\mathbf{u}_2 & \mathbf{e}_3 = \dfrac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 \\
\mathbf{u}_4 = \mathbf{v}_4 - \dfrac{<\mathbf{v}_4, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 - \dfrac{<\mathbf{v}_4, \mathbf{u}_2>}{<\mathbf{u}_2, \mathbf{u}_2>}\mathbf{u}_2 - \dfrac{<\mathbf{v}_4, \mathbf{u}_3>}{<\mathbf{u}_3, \mathbf{u}_3>}\mathbf{u}_3 & \mathbf{e}_4 = \dfrac{1}{\|\mathbf{u}_4\|}\mathbf{u}_4 \\
\quad\vdots & \quad\vdots \\
\mathbf{u}_n = \mathbf{v}_n - \displaystyle\sum_{k=1}^{n-1} \dfrac{<\mathbf{v}_n, \mathbf{u}_k>}{<\mathbf{u}_k, \mathbf{u}_k>}\mathbf{u}_k & \mathbf{e}_n = \dfrac{1}{\|\mathbf{u}_n\|}\mathbf{u}_n.
\end{array}$$

则容易看出：$\forall k \in \{1, \ldots, n\}$，都有 $\mathbf{u}_k \perp \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$。这是因为：用数学归纳法，$\mathbf{u}_2 \perp \mathbf{u}_1$ 显然，不妨设 $\mathbf{u}_j \perp \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}\}$ 对每个 $j = 1, \ldots, k-1$ 成立，下面我们来证明 $\mathbf{u}_k \perp \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$。任取 $i < k$，则

$$\begin{aligned}
<\mathbf{u}_k, \mathbf{u}_i> &= \left\langle \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{<\mathbf{v}_k, \mathbf{u}_j>}{<\mathbf{u}_j, \mathbf{u}_j>}\mathbf{u}_j \ , \ \mathbf{u}_i \right\rangle \\
&=<\mathbf{v}_k, \mathbf{u}_i> - \sum_{j=1}^{k-1} \frac{<\mathbf{v}_k, \mathbf{u}_j>}{<\mathbf{u}_j, \mathbf{u}_j>} <\mathbf{u}_j, \mathbf{u}_i> \\
&=<\mathbf{v}_k, \mathbf{u}_i> - \frac{<\mathbf{v}_k, \mathbf{u}_i>}{<\mathbf{u}_i, \mathbf{u}_i>} <\mathbf{u}_i, \mathbf{u}_i>= 0
\end{aligned}$$

这就说明 $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 两两正交。而每个 $\mathbf{e}_i$ 与 $\mathbf{u}_i$ 同向，因此 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 也两两正交。又因为每个 $\mathbf{e}_i$ 的范数都是 1，故 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 就是 V 的一组标准正交基。 □

推论 3.1.2. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的 m 维子空间，若 U 有一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_m$，则它可以扩充成 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n$。

类似于上面的定理，用 Gram-Schmidt 正交化方法即可证明。定理3.1.2的证明过程也是计算标准正交基的方法。我们来看下面的例子。

例 3.1.5. 在标准正交基 $\mathbb{R}^3$ 上，用 Gram-Schmidt 正交化方法将 $\mathbf{v}_1 = (3, 0, 4)^t$，$\mathbf{v}_2 = (-1, 0, 7)^t$，$\mathbf{v}_3 = (2, , 9, 11)^t$ 化成一组标准正交基。

---

[1] 这个性质我们将放在稍后证明，当然，我们也会在那里顺便给出一个不依赖于正交基选取的证明。

解. 由 Gram-Schmidt 正交化方法，我们有

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 = (3,0,4)^t & \mathbf{e}_1 &= \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = (\frac{3}{5},0,\frac{4}{5})^t \\ \mathbf{u}_2 &= \mathbf{v}_2 - \frac{<\mathbf{v}_2,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1 = (-4,0,3)^t & \mathbf{e}_2 &= \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = (-\frac{4}{5},0,\frac{3}{5})^t \\ \mathbf{u}_3 &= \mathbf{v}_3 - \frac{<\mathbf{v}_3,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1 - \frac{<\mathbf{v}_3,\mathbf{u}_2>}{<\mathbf{u}_2,\mathbf{u}_2>}\mathbf{u}_2 = (0,9,0)^t & \mathbf{e}_3 &= \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = (0,1,0)^t. \end{aligned}$$

则 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 就是 $\mathbb{R}^3$ 的一组标准正交基。 □

从这个例子我们也可以看出，标准正交基的选取并不是唯一的。有了标准正交基以后，我们可以顺便证明正交补空间的以下性质：

定理 3.1.3. 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的子空间，则 $V=U\oplus U^{\perp}$ 并且 $(U^{\perp})^{\perp}=U$。

证明. 设 $\dim(U)=m$，则 U 在原来的内积下仍然是欧氏空间，取 U 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m$，则对任意的 $\mathbf{x}\in V$，令

$$\mathbf{w}=\mathbf{x}-\sum_{i=1}^{m}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i,$$

则对任意的 $j\in\{1,\ldots,m\}$，有

$$<\mathbf{w},\mathbf{e}_j>=<\mathbf{x},\mathbf{e}_j>-\sum_{i=1}^{m}<\mathbf{x},\mathbf{e}_i><\mathbf{e}_i,\mathbf{e}_j>=<\mathbf{x},\mathbf{e}_j>-<\mathbf{x},\mathbf{e}_j>=0.$$

即 $\mathbf{w}$ 与每个 $\mathbf{e}_j$ 都正交，所以 $\mathbf{w}$ 与 U 正交，$\mathbf{w}\in U^{\perp}$。而 $\sum_{i=1}^{m}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i\in U$，于是

$$\mathbf{x}=\sum_{i=1}^{m}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i+\mathbf{w}\in U+U^{\perp},$$

即 $V=U+U^{\perp}$。下证这个和是直和。任取 $\mathbf{y}\in U\cap U^{\perp}$，则 $<\mathbf{y},\mathbf{y}>=0$(前一个 $\mathbf{y}$ 视作 U 中的向量，后一个 $\mathbf{y}$ 视作 $U^{\perp}$ 中的向量)，于是由内积的正定性可知只能是 $\mathbf{y}=\mathbf{0}$。因此 $V=U\oplus U^{\perp}$。

最后我们证明 $U=(U^{\perp})^{\perp}$。首先 $U\subset(U^{\perp})^{\perp}$：任取 $\mathbf{x}\in U$，我们欲证 $\mathbf{x}\perp U^{\perp}$。任取 $\mathbf{y}\in U^{\perp}$，由 $U^{\perp}$ 的定义可知 $<\mathbf{x},\mathbf{y}>=\mathbf{0}$，即 $\mathbf{x}\perp\mathbf{y}$，所以 $\mathbf{x}\perp U^{\perp}$。其次 U 和 $(U^{\perp})^{\perp}$ 的维数相等：利用定理前半部分的结论可知 $V=U\oplus U^{\perp}=U^{\perp}\oplus(U^{\perp})^{\perp}$，因此 $\dim(U)=\dim((U^{\perp})^{\perp})$。由以上两点可知 $U=(U^{\perp})^{\perp}$(包含定理，1.3.2(3))。 □

证明. [法二] $U=(U^{\perp})^{\perp}$ 的证法与上一个证明相同。现在我们绕开标准正交基来证明 $V=U\oplus U^{\perp}$。

设 $\dim(U)=m$，任取 U 的一组基 $\mathbf{u}_1,\ldots,\mathbf{u}_m$，作如下的线性函数：

$$f_1(\mathbf{x})=<\mathbf{u}_1,\mathbf{x}>,$$

$$f_2(\mathbf{x})=<\mathbf{u}_2,\mathbf{x}>,$$

$$\cdots\cdots$$

$$f_m(\mathbf{x})=<\mathbf{u}_m,\mathbf{x}>,$$

则容易验证 $f_i(\mathbf{x}) : V \to \mathbb{R} \ \in V^*$，$\forall i \in \{1, \ldots, m\}$，并且这些 f_i 在 V^* 中是线性无关的。这是因为：任取 $\alpha_1, \ldots, \alpha_m$，如果

$$\sum_{i=1}^{m} \alpha_i f_i(\mathbf{x}) = 0 \in V^*$$

即 $< \sum_{i=1}^{m} \alpha_i \mathbf{u}_i, \mathbf{x} >= 0$ 对任意的 $\mathbf{x} \in V$ 成立，则只能是 $\sum_{i=1}^{m} \alpha_i \mathbf{u}_i = \mathbf{0}$ (因为 $< \cdot, \cdot >$ 的正定性)，再由 $\mathbf{u}_1, \ldots, \mathbf{u}_m$ 是 U 的基可知 $\alpha_1 = \cdots = \alpha_m = 0$，即这些 f_i 在 V^* 中线性无关。

下面我们证明 $U^{\perp}$ 是抽象齐次线性方程组：

$$\begin{cases} f_1(\mathbf{x}) = 0 \\ f_2(\mathbf{x}) = 0 \\ \quad \vdots \\ f_m(\mathbf{x}) = 0 \end{cases}, \quad \text{其中} \mathbf{x} \in V. \tag{H}$$

的解空间。为此，我们把这个解空间记作 W，则一方面任取 $\mathbf{x} \in U^{\perp}$，$\mathbf{x}$ 必然与每个 $\mathbf{u}_i$ 正交，所以 $\mathbf{x} \in W$，即 $U^{\perp} \subset W$；另一方面，任取 $\mathbf{y} \in W$，则 $\mathbf{y}$ 与 U 的每个基底 $\mathbf{u}_i$ 正交，于是由内积的双线性性质，$\mathbf{y}$ 和 U 中的任何元素都正交，即 $\mathbf{y} \in U^{\perp}$，所以 $W \subset U^{\perp}$。综上所述，$W = U^{\perp}$。

因此，由定理1.5.6，$\dim(U^{\perp}) = n - \dim(\text{span}\,\{f_1, \ldots, f_m\}) = n - m$。再由 $U \cap U^{\perp} = \{\mathbf{0}\}$ (同上面的证法) 即可得到 $V = U \oplus U^{\perp}$。 □

现在我们考虑向量在标准正交基下的坐标。

命题 3.1.4. 设 $(V, < \cdot, \cdot >)$ 是 n 维欧氏空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是一组标准正交基。任取 $\mathbf{x} \in V$，则

(1) $\mathbf{x} = \sum_{i=1}^{n} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$，我们把 $< \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$ 称为 $\mathbf{x}$ 在 $\mathbf{e}_i$ 方向上的**投影** (projection)；

(2) (Parseval 等式)[1] $\|\mathbf{x}\|^2 = \sum_{i=1}^{n} | < \mathbf{x}, \mathbf{e}_i > |^2$。

证明. (1) 设 $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$，我们证明每个 $x_i =< \mathbf{x}, \mathbf{e}_i >$ 即可。对每个 $j \in \{1, \ldots, n\}$，作内积

$$< \mathbf{x}, \mathbf{e}_j >= \sum_{i=1}^{n} x_i < \mathbf{e}_i, \mathbf{e}_j >= \sum_{i=1}^{n} x_i \delta_{ij} = x_j.$$

即得结论。

(2) 由 (1) 的结论，我们有：

$$\begin{aligned} \|\mathbf{x}\| =< \mathbf{x}, \mathbf{x} > &=< \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i, \mathbf{x} > \\ &= \sum_{i=1}^{n} < \mathbf{x}, \mathbf{e}_i >< \mathbf{e}_i, \mathbf{x} > \\ &= \sum_{i=1}^{n} | < \mathbf{x}, \mathbf{e}_i > |^2. \end{aligned}$$

即得结论。 □

---

[1] 对于无穷维的内积空间，不等式 $\|\mathbf{x}\|^2 \geqslant \sum_{i=1}^{\infty} | < \mathbf{x}, \mathbf{e}_i > |^2$ 总成立，称为 Bessel 不等式。详细的证明可以参考泛函分析的标准教材，如《A Course in Functional Analysis》，John B.Conway, GTM96 的 Chapter I, §4。

现在我们来考虑欧氏空间的同构。由于欧氏空间上有内积结构，因此我们还希望同构映射能够保持内积。

定义 3.1.6. 设 $(V, <\cdot,\cdot>_V)$ 和 $(W, <\cdot,\cdot>_W)$ 是两个 n 维欧氏空间，如果存在一个线性同构 $f : V \to W$ 使得

$$< \mathbf{x}, \mathbf{y} >_V = < f(\mathbf{x}), f(\mathbf{y}) >_W, \ \forall \mathbf{x}, \mathbf{y} \in V,$$

则称这两个**欧氏**空间 V 和 W 同构。

定理 3.1.4. 设 $(V, <\cdot,\cdot>)$ 是任意一个 n 维欧氏空间，则 V 一定同构于 n 维标准欧氏空间 $(\mathbb{R}^n, <\cdot,\cdot>_{\mathbb{R}^n})$。

证明. 取 V 的一组标准正交基 $\mathbf{e}_1, \dots, \mathbf{e}_n$，则任取 $\mathbf{x} \in V$，$\mathbf{x}$ 可以写成线性组合 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$，$x_i \in \mathbb{R}$，$i \in \{1, \dots, n\}$，于是我们可以定义：

$$\begin{aligned} f : & V \longrightarrow \mathbb{R}^n \\ & \mathbf{x} \longmapsto (x_1, \dots, x_n)^t \end{aligned}$$

显然 f 是向量空间 V 到 $\mathbb{R}^n$ 的线性同构 (定理1.3.4)。下面我们只需证明 f 保持内积即可。

任取 $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i, \mathbf{y} = \sum_{j=1}^n y_j\mathbf{e}_j \in V$，我们有

$$< \mathbf{x}, \mathbf{y} > = \sum_{i=1}^n \sum_{j=1}^n x_i y_j < \mathbf{e}_i, \mathbf{e}_j > = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{ij} = \sum_{i=1}^n x_i y_i,$$

$$< f(\mathbf{x}), f(\mathbf{y}) > = \sum_{i=1}^n x_i y_i.$$

即 $< f(\mathbf{x}), f(\mathbf{y}) > = < \mathbf{x}, \mathbf{y} >$，此即这两个欧氏空间同构。 □

现在我们考虑欧氏空间的对偶空间。设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，任取 $\mathbf{v} \in V$，容易验证 $\Phi_{\mathbf{v}} : V \to \mathbb{R}$，$\mathbf{x} \mapsto < \mathbf{v}, \mathbf{x} >$ 是一个线性函数，即 $\Phi_{\mathbf{v}} \in V^*$ (在定理3.1.3的证法二中已经让大家验证过一次了)。实际上，我们有

命题 3.1.5. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，则

$$\Phi : V \longrightarrow V^*, \quad \mathbf{v} \longmapsto \Phi_{\mathbf{v}} \ (\Phi_{\mathbf{v}}\text{的定义如上文所述.})$$

是向量空间 V 到 V^* 的线性同构，并且 Φ 把 V 的任意一组标准正交基 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 映到其对偶基 $\mathbf{e}^1, \dots, \mathbf{e}^n$。

证明. 首先 Φ 是线性映射：任取 $\mathbf{u}, \mathbf{v} \in V$ 及 $\alpha, \beta \in \mathbb{R}$，将线性函数作用到任意 $\mathbf{x} \in V$ 上，我们有

$$\begin{aligned} \Phi_{\alpha\mathbf{u}+\beta\mathbf{v}}(\mathbf{x}) &= < \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{x} > \\ &= \alpha < \mathbf{u}, \mathbf{x} > + \beta < \mathbf{v}, \mathbf{x} > \\ &= \alpha\Phi_{\mathbf{u}}(\mathbf{x}) + \beta\Phi_{\mathbf{v}}(\mathbf{x}). \end{aligned}$$

其次 Φ 是单射，这只需验证 $\ker(\Phi) = \{\mathbf{0}\}$：任取 $\mathbf{v} \in \ker(\Phi)$，即 $\Phi_{\mathbf{v}}$ 是零映射，则按 $\Phi_{\mathbf{v}}$ 的定义，对任意 $\mathbf{x} \in V$，都有 $< \mathbf{v}, \mathbf{x} > = 0$，因此只能是 $\mathbf{v} = \mathbf{0}$[1]。

---

[1] 这里利用了内积的正定性，取 $\mathbf{x} = \mathbf{v}$ 即得结论。这是一个简单而又常用的结论，我们就不把它写成定理了。

下面我们说明 Φ 把 V 的标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 映到其对偶基 $\mathbf{e}^1,\ldots,\mathbf{e}^n$，从而由维数关系可知 Φ 是满射，即可完成证明。注意到对任意的 $j\in\{1,\ldots,n\}$，我们都有

$$\Phi_{\mathbf{e}_i}(\mathbf{e}_j) =< \mathbf{e}_i,\mathbf{e}_j >= \delta_{ij}$$
$$\mathbf{e}^i(\mathbf{e}_j) = \delta_{ij}$$

由定义1.5.2下方的说明即可得到 $\Phi_{\mathbf{e}_i}=\mathbf{e}^i$ 对每个 $i\in\{1,\ldots,n\}$ 都成立。

□

我们还可以在 V^* 上定义如下的内积：$<\Phi_{\mathbf{u}},\Phi_{\mathbf{v}}>=<\mathbf{u},\mathbf{v}>$，并且很容易证明在此内积下 V^* 也是 n 维欧氏空间，上面命题中的 Φ 是欧氏空间的同构。实际上，这个命题的一个更常用的表述是：任意的 $f\in V^*$，存在唯一的 $\mathbf{v}\in V$ 使得 $f(\mathbf{x}) =< \mathbf{v},\mathbf{x} >$，并且 $\|f\|=\|\mathbf{v}\|$。上面的命题是 Riesz 表示定理在欧氏空间上的特殊情形，更一般的情形我们会在泛函分析课程中讨论。

接下来我们考虑两组标准正交基之间的过渡矩阵。设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，由于 $<\cdot,\cdot>$ 是 V 上的对称双线性型，故任取 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$，双线性型 $<\cdot,\cdot>$ 对应的矩阵为

$$G=\begin{pmatrix} <\mathbf{v}_1,\mathbf{v}_1> & <\mathbf{v}_1,\mathbf{v}_2> & \cdots & <\mathbf{v}_1,\mathbf{v}_n> \\ <\mathbf{v}_2,\mathbf{v}_1> & <\mathbf{v}_2,\mathbf{v}_2> & \cdots & <\mathbf{v}_2,\mathbf{v}_n> \\ \vdots & \vdots & & \vdots \\ <\mathbf{v}_n,\mathbf{v}_1> & <\mathbf{v}_n,\mathbf{v}_2> & \cdots & <\mathbf{v}_n,\mathbf{v}_n> \end{pmatrix}$$

我们称上面的 G 为内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的**度量矩阵**或 Gram 矩阵。由第一章的定理1.6.2可知，如果内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 下的度量矩阵分别为 G 和 G'，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$, $A\in \mathrm{GL}_n(\mathbb{R})$，则 $G'=A^tGA$。显然 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基 $\Longleftrightarrow$ 其度量矩阵为 E_n。于是，如果 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 都是 V 的标准正交基，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$，则 $E_n=A^tE_nA$，即 $A^tA=E_n$。反之，如果矩阵 A 满足 $A^tA=E_n$，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，令 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$，则显然 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是 V 的基 (A 显然可逆)，并且按上面的讨论，内积在 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的度量矩阵为 $A^tE_nA=E_n$，即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是标准正交基。这就引出了下面的正交矩阵的定义。

定义 3.1.7. 设 $A\in\mathrm{GL}_n(\mathbb{R})$，如果 A 满足 $A^tA=E$，即 $A^t=A^{-1}$，则称 A 为正交矩阵 (orthogonal matrix)。容易验证 $\mathbb{R}$ 上所有 n 阶正交矩阵在矩阵乘法运算下构成了一个群 (封闭：$A^t=A^{-1},B^t=B^{-1}$，则 $(AB)^t=B^tA^t=B^{-1}A^{-1}=(AB)^{-1}$，结合律显然，幺元为 E_n，A 的逆元为 A^t。)，我们把这个群称为 (实) 正交群 (orthogonal group)，记作 $O(n)$。

容易看出正交矩阵就是标准正交基之间的过渡矩阵。在上一章中我们已经知道，方阵与线性算子之间是一一对应的。因此，受定义3.1.7的启发，我们也可以定义正交算子如下：

定义 3.1.8. 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathscr{A}$ 是 V 上的可逆的线性算子。如果 $\mathscr{A}$ 还满足：对任意的 $\mathbf{x},\mathbf{y}\in V$，$<\mathscr{A}\mathbf{x},\mathscr{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathscr{A}$ 是 V 上的正交算子 (orthogonal operator) 或正交变换。

我们有以下定理：

定理 3.1.5. 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathscr{A}$ 是 V 上的线性算子，则以下条件等价：

(1) $\mathscr{A}$ 是正交算子；

(2) $\mathscr{A}$ 在标准正交基下对应的矩阵是正交矩阵；

(3) $\mathscr{A}$ 将标准正交基映到标准正交基；

(4) $\forall \mathbf{x} \in V,\ \|\mathscr{A}\mathbf{x}\| = \|\mathbf{x}\|$。

证明. (1)$\Longrightarrow$(2)：设 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组标准正交基，$\mathscr{A}$ 在这组基下的矩阵为 A。首先，令 $\mathbf{v}_i = \mathscr{A}\mathbf{e}_i$，$i = 1, \dots, n$，则

$$< \mathbf{v}_i, \mathbf{v}_j >=< \mathscr{A}\mathbf{e}_i, \mathscr{A}\mathbf{e}_j >=< \mathbf{e}_i, \mathbf{e}_j >= \delta_{ij}$$

即 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 也是 V 的一组标准正交基。显然 $(\mathbf{v}_1, \dots, \mathbf{v}_n) = (\mathbf{e}_1, \dots, \mathbf{e}_n) \cdot A$，由定义3.1.7前面的那段讨论可知 A 是正交矩阵。
(2)$\Longrightarrow$(3)：即定义3.1.7前面的讨论（“反之”之后的部分)。
(3)$\Longrightarrow$(4)：设 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组标准正交基，任取 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \in V$，则 $\mathscr{A}\mathbf{x} = \sum_{i=1}^n x_i \mathscr{A}\mathbf{e}_i$。由于 $\mathscr{A}\mathbf{e}_1, \dots, \mathscr{A}\mathbf{e}_n$ 也是 V 的一组标准正交基，因此我们有：

$$\begin{aligned} < \mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{x} > &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j < \mathscr{A}\mathbf{e}_i, \mathscr{A}\mathbf{e}_j >= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \delta_{ij} = \sum_{i=1}^n x_i^2; \\ < \mathbf{x}, \mathbf{x} > &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j < \mathbf{e}_i, \mathbf{e}_j >= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \delta_{ij} = \sum_{i=1}^n x_i^2. \end{aligned}$$

所以 $\|\mathscr{A}\mathbf{x}\| = \|\mathbf{x}\|$。
(4)$\Longrightarrow$(1)：由极化恒等式 (命题3.1.3(3))，任取 $\mathbf{x}, \mathbf{y} \in V$，我们有

$$\begin{aligned} < \mathscr{A}\mathbf{x}, \mathscr{A}\mathbf{y} > &= \frac{1}{4}(\|\mathscr{A}(\mathbf{x}+\mathbf{y})\|^2 - \|\mathscr{A}(\mathbf{x}-\mathbf{y})\|^2) \\ &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) \\ &=< \mathbf{x}, \mathbf{y} > . \end{aligned}$$

即 $\mathscr{A}$ 是正交算子。 □

推论 3.1.3. 设 $(V, < \cdot, \cdot >)$ 是 n 维欧氏空间，则 V 上所有正交算子在算子复合运算下构成群，记作 $O(V)$，则 $O(V)$ 同构于正交群 $O(n)$。

证明. 首先 $O(V)$ 是群：设 $\mathscr{A}, \mathscr{B} \in O(V)$，则对任意的 $\mathbf{x}, \mathbf{y} \in V$，$< \mathscr{A}\mathscr{B}\mathbf{x}, \mathscr{A}\mathscr{B}\mathbf{y} >=< \mathscr{B}\mathbf{x}, \mathscr{B}\mathbf{y} >=< \mathbf{x}, \mathbf{y} >$，即 $\mathscr{A}\mathscr{B} \in O(V)$，封闭性成立；结合律即映射复合的结合律；幺元是恒等算子 $\mathscr{E}$；由于正交算子一定可逆 (将基映到基，满秩)，故 $\mathscr{A}^{-1}$ 存在，并且 $< \mathscr{A}^{-1}\mathbf{x}, \mathscr{A}^{-1}\mathbf{y} >=< \mathscr{A}\mathscr{A}^{-1}\mathbf{x}, \mathscr{A}\mathscr{A}^{-1}\mathbf{y} >=< \mathbf{x}, \mathbf{y} >$，即 $\mathscr{A}^{-1} \in O(V)$。因此 $O(V)$ 是群。

任取一组标准正交基，设 $\Phi : O(V) \to O(n)$ 将每个正交算子映到它在这种标准正交基下的矩阵，则由算子复合与矩阵乘法的对应关系可知 Φ 是群同态；由前面的论述显然 Φ 是双射，故 $O(V) \simeq O(n)$。 □

辛空间 *

在本节的最后，我们考虑另一种形式的“度量”结构，它被称为辛结构 (symplectic structure)。辛结构在几何和物理中都有重要的应用。

定义 3.1.9. 设 V 是 $\mathbb{R}$ 上的 $n(=2m)$ 维向量空间，$<\cdot,\cdot>$ 是 V 上的非退化的斜对称双线性型[1]，即 $<\cdot,\cdot>$ 满足：

1) $\forall \mathbf{v}\in V,\ \exists \mathbf{x}\in V$ 使得 $<\mathbf{v},\mathbf{x}>\neq 0$；

2) $\forall \mathbf{x},\mathbf{y}\in V,\quad <\mathbf{x},\mathbf{y}>=-<\mathbf{y},\mathbf{x}>$；

3) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V,\quad <\mathbf{x}+\mathbf{y},\mathbf{z}>=<\mathbf{x},\mathbf{z}>+<\mathbf{y},\mathbf{z}>$；

4) $\forall \mathbf{x},\mathbf{y}\in V$ 及 $\lambda\in\mathbb{R},\quad <\lambda\mathbf{x},\mathbf{y}>=\lambda<\mathbf{x},\mathbf{y}>$。

则我们称 $<\cdot,\cdot>$ 是 V 上的一个辛内积或辛结构，称 $(V,\ <\cdot,\cdot>)$ 是 $\mathbb{R}$ 上的 $2m$ 维辛空间 (symplectic space)。设 $\mathcal{A}\in\mathrm{End}(V)$，如果 $\mathcal{A}$ 还保持辛结构，即 $\forall\mathbf{x},\mathbf{y}\in V,\ <\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathcal{A}$ 是一个辛算子 (symplectic operator)。如果 $\mathbf{x},\mathbf{y}\in V$ 满足 $<\mathbf{x},\mathbf{y}>=0$，则我们称 $\mathbf{x}$ 和 $\mathbf{y}$ 是斜正交的，仍记作 $\mathbf{x}\perp\mathbf{y}$。显然 $\forall\mathbf{x}\in V,\ \mathbf{x}\perp\mathbf{x}$。

类似于欧氏内积的度量矩阵，我们也可以定义辛内积在一组基下的“度量”矩阵 (Gram 矩阵)。利用定理1.6.13，我们很容易得到下面的结论：

定理 3.1.6. 设 $(V,<\cdot,\cdot>)$ 是 $n=2m$ 维辛空间，则存在 V 的一组基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m$ (称为辛基) 使得辛内积在此基底下的 Gram 矩阵为 $J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}$，即只有 $<\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i>=-<\boldsymbol{\beta}_i,\boldsymbol{\alpha}_i>=1,\ \forall i\in\{1,\ldots,m\}$，其余基向量之间的辛内积均为 0。

与标准欧氏空间类似，我们也可以定义标准辛空间。

定义 3.1.10. 在 $\mathbb{R}^{2m}$ 上定义如下的辛内积：任取 $\mathbf{x}=(x_1,\ldots,x_m,x_{m+1},\ldots,x_{2m})^t$ 和 $\mathbf{y}=(y_1,\ldots,y_m,y_{m+1},\ldots,y_{2m})^t\in\mathbb{R}^{2m}$，定义

$$<\mathbf{x},\mathbf{y}>_s=\sum_{i=1}^{m}(x_{m+i}y_i-x_iy_{m+i})$$

则 $<\cdot,\cdot>_s$ 在 $\mathbb{R}^{2m}$ 的标准基 $\mathbf{e}_1=(1,0,\ldots,0)^t,\ldots\ldots,\mathbf{e}_{2m}=(0,\ldots,0,1)^t$ 下的 Gram 矩阵恰为上面定理中的 J_0。我们称 $(\mathbb{R}^{2m},\ <\cdot,\cdot>_s)$ 为 $2m$ 维的标准辛空间。注意到 $J_0^2=-E_{2m}$，这与虚数单位 i 的性质类似。在稍后的复化与实化一节中，我们还会用到这个空间。

与欧氏空间的情形类似，我们有：

定理 3.1.7. 设 $(V,<\cdot,\cdot>)$ 是 $2m$ 维辛空间，则存在线性同构 $f:V\to\mathbb{R}^{2m}$ 使得 $<f(\mathbf{x}),f(\mathbf{y})>_s=<\mathbf{x},\mathbf{y}>$。(我们把这样保持辛内积的线性同构称为辛同构。)

证明方法与定理3.1.4类似，留作练习。

定义 3.1.11. 设 $A\in M_{2m}(\mathbb{R})$，$J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}$，如果 A 满足 $A^tJ_0A=J_0$，则我们称 A 为辛矩阵 (symplectic matrix)。

如果 A 是 $2m$ 阶的辛矩阵，则利用普法夫型的性质 (定理1.6.14) 可知，$1=\mathrm{Pf}_{2m}(J_0)=\mathrm{Pf}_{2m}(A^tJ_0A)=\det(A)\mathrm{Pf}_{2m}(J_0)=\det(A)$。容易验证所有 $2m$ 阶辛矩阵在矩阵乘法下构成了一个群 (留作练习)，我们把这个群称为辛群，记作 $\mathrm{Sp}_{2m}(\mathbb{R})$。于是显然 $\mathrm{Sp}_{2m}(\mathbb{R})$ 是特殊线性群

---

[1] 注意，由 1.6.7 小节开头的讨论，只有偶数维的向量空间上才有非退化的斜对称双线性型。

$\mathrm{SL}_{2m}(\mathbb{R})$ 的子群。特别地，当 $2m = 2$ 时，$\mathrm{Sp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$。这是因为任取 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$，即 $ad - bc = 1$，我们有

$$A^t J_0 A = \begin{pmatrix} 0 & -(ad-bc) \\ (ad-bc) & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = J_0$$

即 $A \in \mathrm{Sp}_2(\mathbb{R})$。

下面我们考察辛算子与辛矩阵之间的联系。

定理 3.1.8. 设 $(V, <\cdot,\cdot>)$ 是 $2m$ 维辛空间，则所有 V 上所有辛算子在算子复合所定义的乘法下构成一个群，记作 $\mathrm{Sp}(V)$，并且 $\mathrm{Sp}(V) \simeq \mathrm{Sp}_{2m}(\mathbb{R})$。

证明. 由定理3.1.6，存在辛基 $\mathbf{e}_1, \ldots, \mathbf{e}_m, \ldots \mathbf{e}_{2m}$ 使得

$$(<\mathbf{e}_i, \mathbf{e}_j>)_{2m\times 2m} = J_0 = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$$

令 $\Phi : \mathrm{Sp}(V) \to \mathrm{Sp}_{2m}(\mathbb{R})$，$\mathcal{A} \mapsto A$，其中 A 是算子 $\mathcal{A}$ 在上述辛基下的矩阵。则用类似于定理3.1.5的证明方法可以得到以下条件等价：

(1) $\mathcal{A}$ 是辛算子；

(2) $\mathcal{A}$ 在辛基下对应的矩阵是辛矩阵；

(3) $\mathcal{A}$ 将辛基映到辛基。

细节留作练习。于是 Φ 是线性双射，并且由于线性算子的复合对应矩阵的乘法，故 Φ 是群同态，因此 $\mathrm{Sp}(V) \simeq \mathrm{Sp}_{2m}(\mathbb{R})$。 □

最后我们关注辛算子的特征多项式的一个有趣的性质。

定理 3.1.9. 设 $(V, <\cdot,\cdot>)$ 是 $2m$ 维辛空间，$\mathcal{A}$ 是 V 上的辛算子，则 $\chi_{\mathcal{A}}(t) = t^{2m}\chi_{\mathcal{A}}(\frac{1}{t})$。

证明. 设 $\mathcal{A}$ 在一组辛基下对应矩阵 A，则 A 满足 $A^t J_0 A = J_0$，于是 $A = J_0^{-1}(A^t)^{-1}J_0$。因此我们有：

$$\begin{aligned}
\chi_{\mathcal{A}}(t) &= \det(tE - A) = \det\left(tE - J_0^{-1}\cdot(A^t)^{-1}\cdot J_0\right) = \det\left(J_0^{-1}\left(tE - (A^t)^{-1}\right)J_0\right) \\
&= \det\left(tE - (A^t)^{-1}\right) = \det\left(\left(tE - A^{-1}\right)^t\right) = \det\left(tE - A^{-1}\right) \\
&= \det\left(tE - A^{-1}\right)\det A = \det(tA - E) = \det(E - tA) = t^{2m}\det\left(\frac{1}{t}E - A\right) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right)
\end{aligned}$$

这样我们就完成了证明。 □

3.2 埃尔米特向量空间

上一节我们讨论了实数域上的内积空间，现在我们尝试将内积定义到复向量空间上。为了区分复数的模长和向量的范数这两个记号，我们把复数 α 的模长记作 $|\alpha|$。由于复向量空间上的非退化二次型 $x_1^2+\cdots+x_n^2$ 不再具有正定性 ($i^2=-1$)，因此我们需要寻找新的函数来定义内积，这就是接下来介绍的埃尔米特 (Hermitian) 型。

3.2.1 埃尔米特型

定义 3.2.1. 设 V 是 n 维复向量空间，$f: V\times V\to\mathbb{C}$ 满足：

(1) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V$, $\alpha,\beta\in\mathbb{C}$，有 $f(\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{z})=\alpha f(\mathbf{x},\mathbf{z})+\beta f(\mathbf{y},\mathbf{z})$；

(2) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V$, $\alpha,\beta\in\mathbb{C}$，有 $f(\mathbf{z},\alpha\mathbf{x}+\beta\mathbf{y})=\overline{\alpha}f(\mathbf{z},\mathbf{x})+\overline{\beta}f(\mathbf{z},\mathbf{y})$，其中 $\overline{\alpha},\overline{\beta}$ 分别是 α 和 β 的共轭复数。

则称 f 是 V 上的半双线性型 (sesquilinear form, 即 f 对第一个变量线性，对第二个变量共轭线性)。如果半双线性型 f 还满足 $\forall\mathbf{x},\mathbf{y}\in V$, $f(\mathbf{y},\mathbf{x})=\overline{f(\mathbf{x},\mathbf{y})}$，则我们称 f 是一个埃尔米特 (Hermitian) 型。对应地，如果 $f(\mathbf{y},\mathbf{x})=-\overline{f(\mathbf{x},\mathbf{y})}$，则称 f 是一个斜埃尔米特型。

任取复向量空间 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$，类似于 1.6 节我们对双线性型的讨论，我们可以定义半双线性型 f 在此基底下的矩阵

$$F=\left(f_{ij}\right)_{n\times n},\text{ 其中}f_{ij}=f(\mathbf{e}_i,\mathbf{e}_j).$$

则任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{j=1}^n y_j\mathbf{e}_j\in V$，利用 f 的半双线性性质可知

$$f(\mathbf{x},\mathbf{y})=\sum_{i=1}^n\sum_{j=1}^n f_{ij}x_i\overline{y_j}=(x_1,\ldots,x_n)\cdot F\cdot\begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}.$$

特别地，如果 f 是 Hermitian 型，则 $f_{ij}=f(\mathbf{e}_i,\mathbf{e}_j)=\overline{f(\mathbf{e}_j,\mathbf{e}_i)}=\overline{f_{ji}}$。我们记矩阵 F 的转置共轭[1]为 F^*(有的书上也记成 F^H 或 $F^\dagger$)，则 f 是 Hermitian 型 $\Longleftrightarrow$ f 在任意一组基下对应的矩阵 F 满足 $F^*=F$。我们称满足 $F^*=F$ 的复矩阵为埃尔米特 (Hermitian) 矩阵。

下面我们考虑基变换下半双线性型对应的矩阵之间的关系。类似于定理1.6.2的证明方法，设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 是复向量空间 V 的两组基且 $(\mathbf{e}_1',\ldots,\mathbf{e}_n')=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$。任取 $\mathbf{x},\mathbf{y}\in V$，设 $\mathbf{x}$ 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 下的坐标分别为 $(x_1,\ldots,x_n)^t$ 和 $(x_1',\ldots,x_n')^t$，$\mathbf{y}$ 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 下的坐标分别为 $(y_1,\ldots,y_n)^t$ 和 $(y_1',\ldots,y_n')^t$，f 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}_1',\ldots,\mathbf{e}_n'$ 下的矩阵分别为 F 和 F'，则

$$f(\mathbf{x},\mathbf{y})=(x_1,\ldots,x_n)F\begin{pmatrix}\overline{y_1}\\ \ldots\\ \overline{y_n}\end{pmatrix}=(x_1',\ldots,x_n')F'\begin{pmatrix}\overline{y_1'}\\ \ldots\\ \overline{y_n'}\end{pmatrix}\tag{3.2.1}$$

---

[1] 很容易验证对复矩阵 F 的每个元素取共轭复数的操作和对 F 取转置的操作是交换的，即 $(\overline{F})^t=\overline{F^t}$。

由坐标变换公式有

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \cdot \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix}$$

即

$$(x_1, \ldots, x_n) = (x'_1 \ldots, x'_n) \cdot A^t, \qquad \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} = \overline{A} \cdot \begin{pmatrix} \overline{y'_1} \\ \vdots \\ \overline{y'_n} \end{pmatrix}$$

将其代入式 (3.2.1) 得

$$(x'_1, \ldots, x'_n) A^t F \overline{A} \begin{pmatrix} \overline{y'_1} \\ \ldots \\ \overline{y'_n} \end{pmatrix} = (x'_1, \ldots, x'_n) F' \begin{pmatrix} \overline{y'_1} \\ \ldots \\ \overline{y'_n} \end{pmatrix}$$

上式对任意 $(x'_1, \ldots, x'_n)$ 和 $(y'_1, \ldots, y'_n)$ 都成立，于是我们让 $(x'_1, \ldots, x'_n)$ 和 $(y'_1, \ldots, y'_n)$ 分别取遍 $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, \ldots, 0, 1)$ 即可知 F 和 F' 每个位置的元素都相等，所以 $F' = A^t F \overline{A}$。

特别地，如果 f 是 Hermitian 型，F' 和 F 是 f 在两组基下的矩阵，则容易看出

$$(F')^* = \overline{(A^t F \overline{A})^t} = \overline{\overline{A}^t} \cdot \overline{F^t} \cdot \overline{A} = A^t F^* \overline{A} = F',$$

即 V 上的 Hermitian 型与 Hermitian 矩阵是一一对应的。

最后，需要说明的是，如果 f 是复向量空间 V 上的一个 Hermitian 型，则必有 $\forall \mathbf{x} \in V$, $f(\mathbf{x}, \mathbf{x}) \in \mathbb{R}$ (因为 $f(\mathbf{x}, \mathbf{x}) = \overline{f(\mathbf{x}.\mathbf{x})}$)。因此，就像我们可以在实对称双线性型对应的二次型上定义正定性那样，如果 $\forall \mathbf{x} \in V$, $f(\mathbf{x}, \mathbf{x}) \geqslant 0$，则称 f 是半正定的；如果 f 半正定并且 $f(\mathbf{x}, \mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$，则称 f 是正定的。

3.2.2 埃尔米特空间

现在我们可以定义复内积空间了。

定义 3.2.2. 设 V 是 n 维复向量空间, $< \cdot, \cdot >: V \times V \to \mathbb{C}$ 是 V 上的某个给定的正定 Hermitian 型 (也称 $< \cdot, \cdot >$ 为内积)，则我们称 $(V, < \cdot, \cdot >)$ 是埃尔米特 (Hermitian) 空间或酉空间。或者说 Hermiian 空间上的内积 $< \cdot, \cdot >$ 应该满足以下性质：

1) $\forall \mathbf{x}, \mathbf{y} \in V, \quad < \mathbf{x}, \mathbf{y} > = \overline{< \mathbf{y}, \mathbf{x} >}$;

2) $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, \quad < \mathbf{x} + \mathbf{y}, \mathbf{z} > = < \mathbf{x}, \mathbf{z} > + < \mathbf{y}, \mathbf{z} >$;

3) $\forall \mathbf{x}, \mathbf{y} \in V$ 及 $\lambda \in \mathbb{C}, \quad < \lambda \mathbf{x}, \mathbf{y} > = \lambda < \mathbf{x}, \mathbf{y} >$;

4) $\forall \mathbf{x} \in V, \quad < \mathbf{x}, \mathbf{x} > \geqslant 0$，并且等号当且仅当 $\mathbf{x} = \mathbf{0}$ 时成立。

下面我们来看几个例子。

例 3.2.1. (1) 在 $\mathbb{C}^n$ 上，任取 $\mathbf{x} = (x_1, \ldots, x_n)^t, \mathbf{y} = (y_1, \ldots, y_n)^t \in V$，定义内积 $< \mathbf{x}, \mathbf{y} > =$

$\sum_{i=1}^{n} x_i\overline{y_i}$，则容易验证 $(\mathbb{C}^n, <\cdot,\cdot>)$ 是 Hermitian 空间，它被称为 n 维的标准 Hermitian 空间。

(2) 固定 $a, b \in \mathbb{R}$, $a < b$，任取 $f(x), g(x)$ 是闭区间 $[a, b]$ 上的复数值连续函数，定义：

$$< f, g >= \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t.$$

则容易验证这是一个内积，即 $[a, b]$ 上的复数值连续函数在此内积下构成了一个复内积空间 (注意这个空间不是有限维的)。

与实内积空间类似，我们可以定义复内积空间上的范数和距离如下：

定义 3.2.3. 设 V 是复内积空间，任取 $\mathbf{x} \in V$，则我们可以定义 $\mathbf{x}$ 的范数 (或称为长度，norm) 为：$\|\mathbf{x}\| = \sqrt{<\mathbf{x},\mathbf{x}>} \in \mathbb{R}$。于是，任取 $\mathbf{x}, \mathbf{y} \in V$，我们称 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{<\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}>}$ 为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 的距离 (distance)。特别地，如果向量 $\mathbf{x}$ 的范数是 1，则称 $\mathbf{x}$ 是一个单位向量 (unit vector)。

同样我们有复内积空间上的 CBS 不等式：

定理 3.2.1 (CBS 不等式). 设 $(V, <\cdot,\cdot>)$ 是复内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $| <\mathbf{x},\mathbf{y}> | \leqslant \|\mathbf{x}\|\cdot\|\mathbf{y}\|$，并且等号成立的充分必要条件是 $\mathbf{x} = \mathbf{0}$ 或 $\exists\lambda \in \mathbb{C}$ 使得 $\mathbf{y} = \lambda\mathbf{x}$。

证明. 如果 $\mathbf{x}$ 或 $\mathbf{y}$ 是 $\mathbf{0}$，则不等式两边都是 0，显然成立。下设 $\mathbf{x}$ 和 $\mathbf{y}$ 都不是 $\mathbf{0}$。由复数的指数形式，不妨设 $<\mathbf{x},\mathbf{y}>= | <\mathbf{x},\mathbf{y}> | \cdot e^{i\theta} \in \mathbb{C}$，其中 $\theta \in \mathbb{R}$。类似于欧氏空间上 CBS 不等式的证明，由于 $\forall t \in \mathbb{R}$，有 $< t\mathbf{x} + e^{i\theta}\mathbf{y}, t\mathbf{x} + e^{i\theta}\mathbf{y} > \ \geqslant 0$，将不等式左边展开得：

$$<\mathbf{x},\mathbf{x}> t^2 - (<\mathbf{x},\mathbf{y}> e^{-i\theta} + \overline{<\mathbf{x},\mathbf{y}>}e^{i\theta})t + <\mathbf{y},\mathbf{y}> \geqslant 0 \text{ 对}\forall\lambda \in \mathbb{R}\text{成立}.$$

将 $<\mathbf{x},\mathbf{y}>= | <\mathbf{x},\mathbf{y}> | \cdot e^{i\theta}$ 代入上式即有：

$$<\mathbf{x},\mathbf{x}> t^2 - 2| <\mathbf{x},\mathbf{y}> | \cdot t + <\mathbf{y},\mathbf{y}> \geqslant 0 \text{ 对}\forall\lambda \in \mathbb{R}\text{成立}.$$

这说明上式左边的判别式非正，即

$$4| <\mathbf{x},\mathbf{y}> |^2 - 4 <\mathbf{x},\mathbf{x}> \cdot <\mathbf{y},\mathbf{y}> \leqslant 0,$$

整理即得 CBS 不等式。等号成立条件的讨论与实内积空间的情形相同，留作练习。 □

于是我们在这里可以直接将实内积空间的一系列定义和性质照搬过来：

推论 3.2.1 (三角不等式). 设 $(V, <\cdot,\cdot>)$ 是复内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$，并且等号成立的充分必要条件为 $\exists\alpha, \beta \in \mathbb{C}$ 使得 $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$。

命题 3.2.1. 设 $(V, <\cdot,\cdot>)$ 是复内积空间，$\|\cdot\|$ 是该内积对应的范数，则

(1) $\forall\mathbf{x} \in V$，$\|\mathbf{x}\| \geqslant 0$，并且 $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$；

(2) $\forall\lambda \in \mathbb{R}$, $\mathbf{x} \in V$，都有 $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$；

(3) $\forall\mathbf{x}, \mathbf{y} \in V$，有 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$。

命题 3.2.2. 设 $(V, <\cdot,\cdot>)$ 是复内积空间，$\mathrm{d}(\cdot,\cdot)$ 是该内积对应的距离，则

(1) $\forall \mathbf{x},\mathbf{y} \in V$，$\mathrm{d}(\mathbf{x},\mathbf{y}) \geqslant 0$，并且等号当且仅当 $\mathbf{y} = \mathbf{x}$ 时成立；

(2) $\forall \mathbf{x},\mathbf{y} \in V$，都有 $\mathrm{d}(\mathbf{x},\mathbf{y}) = \mathrm{d}(\mathbf{y},\mathbf{x})$；

(3) $\forall \mathbf{x},\mathbf{y},\mathbf{z} \in V$，有 $\mathrm{d}(\mathbf{x},\mathbf{z}) \leqslant \mathrm{d}(\mathbf{x},\mathbf{y}) + \mathrm{d}(\mathbf{y},\mathbf{z})$。

定义 3.2.4. 设 $(V, <\cdot,\cdot>)$ 是复内积空间，任取非零向量 $\mathbf{x},\mathbf{y} \in V$，我们定义 $\mathbf{x}$ 和 $\mathbf{y}$ 的夹角是如下的 θ：

$$\cos\theta = \frac{|<\mathbf{x},\mathbf{y}>|}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|},\ 0 \leqslant \theta \leqslant \frac{\pi}{2}. \qquad (\text{CBS 不等式保证了}\theta\text{的存在性.})$$

如果 $\theta = \frac{\pi}{2}$，则我们称 $\mathbf{x},\mathbf{y}$ 正交 (orthogonal)，记作 $\mathbf{x} \perp \mathbf{y}$。特别地，我们规定 $\mathbf{0}$ 与任意 V 中的向量都正交，即 $\forall \mathbf{x},\mathbf{y} \in V,\ \mathbf{x} \perp \mathbf{y} \iff <\mathbf{x},\mathbf{y}>= 0$。

命题 3.2.3 (勾股定理). 设 $(V, <\cdot,\cdot>)$ 是复内积空间，$\mathbf{x},\mathbf{y} \in V$ 且 $\mathbf{x} \perp \mathbf{y}$，则 $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2+\|\mathbf{y}\|^2$。

引理 3.2.1. 设 $(V, <\cdot,\cdot>)$ 是复内积空间，$\mathbf{x} \in V$，$\mathbf{u}_1,\ldots,\mathbf{u}_k$ 是子空间 U 的一组基。如果 $\mathbf{x} \perp \mathbf{u}_1,\ldots,\mathbf{x} \perp \mathbf{u}_k$ 都成立，则任取 $\mathbf{u} \in U$，都有 $\mathbf{x} \perp \mathbf{u}$。我们把这种情形记作 $\mathbf{x} \perp U$。

如果 U_1, U_2 是 V 的两个子空间，并且任取 $\mathbf{x} \in U_1$，都有 $\mathbf{x} \perp U_2$，则我们称子空间 U_1 和 U_2 正交，记作 $U_1 \perp U_2$。显然 $U_1 \perp U_2 \Longrightarrow U_1 \cap U_2 = \{\mathbf{0}\}$。

命题 3.2.4. 设 $(V, <\cdot,\cdot>)$ 是复内积空间，$\mathbf{x},\mathbf{y} \in V$，则

(1) (扩展的三角不等式)$\Big|\|\mathbf{x}\| - \|\mathbf{y}\|\Big| \leqslant \|\mathbf{x} \pm \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$；

(2) (平行四边形法则)$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$；

(3) (极化恒等式)$<\mathbf{x},\mathbf{y}>= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) + \frac{i}{4}(\|\mathbf{x}+i\mathbf{y}\|^2 - \|\mathbf{x}-i\mathbf{y}\|^2)$。

与欧氏空间一样，我们也可以在 Hermitian 空间上定义标准正交基。

定义 3.2.5. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间。如果 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 满足：

$$\forall i \neq j \in \{1,\ldots,n\},\quad <\mathbf{v}_i,\mathbf{v}_j>= 0,$$

则我们称 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 V 的一组正交基；如果 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是正交基，并且每个 $\mathbf{e}_i$，$i \in \{1,\ldots,n\}$ 的范数都是 1，即

$$<\mathbf{e}_i,\mathbf{e}_j>= \delta_{ij},$$

则我们称 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基。

显然 $\mathbb{C}^n$ 的标准基 $\mathbf{e}_1 = (1,0,\ldots,0)^t,\ldots,\mathbf{e}_n = (0,\ldots,0,1)^t$ 就是一组标准正交基。

定理 3.2.2. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 V 中一定存在一组标准正交基。

用 Gram-Schmidt 正交化方法即可证明，过程与定理3.1.2的证明完全相同。

推论 3.2.2. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的 m 维子空间，若 U 有一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m$，则它可以扩充成 V 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m,\mathbf{e}_{m+1},\ldots,\mathbf{e}_n$。

定义 3.2.6. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的子空间。则

$$W = \{\mathbf{x} \in V \mid \forall \mathbf{u} \in U, \quad <\mathbf{x},\mathbf{u}>= 0\}$$

是 V 的子空间，称 W 为 U 的正交补空间，记作 $W = U^{\perp}$。

定理 3.2.3. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的子空间，则 $V = U \oplus U^{\perp}$ 并且 $(U^{\perp})^{\perp} = U$。

证明方法与定理3.1.3的证明相同，留作练习。

命题 3.2.5. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是一组标准正交基。任取 $\mathbf{x} \in V$，则

(1) $\mathbf{x} = \sum_{i=1}^{n} <\mathbf{x},\mathbf{e}_i> \mathbf{e}_i$，我们仍把 $<\mathbf{x},\mathbf{e}_i> \mathbf{e}_i$ 称为 $\mathbf{x}$ 在 $\mathbf{e}_i$ 方向上的**投影**；

(2) (Parseval 等式)[1] $\|\mathbf{x}\|^2 = \sum_{i=1}^{n} |<\mathbf{x},\mathbf{e}_i>|^2$。

与命题3.1.4的证明方法类似 (注意此时内积对第二个变量是共轭线性的)。

定义 3.2.7. 设 $(V, <\cdot,\cdot>_V)$ 和 $(W, <\cdot,\cdot>_W)$ 是两个 n 维 Hermitian 空间，如果存在一个线性同构 $f : V \to W$ 使得

$$<\mathbf{x},\mathbf{y}>_V=< f(\mathbf{x}), f(\mathbf{y})>_W, \ \forall \mathbf{x},\mathbf{y} \in V,$$

则称这两个 Hermitian 空间 V 和 W 同构。

定理 3.2.4. 设 $(V, <\cdot,\cdot>)$ 是任意一个 n 维 Hermitian 空间，则 V 一定同构于 n 维标准 Hermitian 空间 $(\mathbb{C}^n, <\cdot,\cdot>_{\mathbb{C}^n})$。

与定理3.1.4的证明方法相同，留作练习。

现在我们考虑 Hermitian 空间的对偶空间。类比欧氏空间的情形，我们自然会希望 $\forall \mathbf{v} \in V$，$\Phi_{L,\mathbf{v}} : V \to \mathbb{R}$，$\mathbf{x} \mapsto <\mathbf{v},\mathbf{x}>$ 或 $\Phi_{R,\mathbf{v}} : V \to \mathbb{R}$，$\mathbf{x} \mapsto <\mathbf{x},\mathbf{v}>$ (注意 $\Phi_{L,\mathbf{v}}$ 和 $\Phi_{R,\mathbf{v}}$ 是不相等的！) 是线性函数，并且 $\mathbf{v} \mapsto \Phi_{L,\mathbf{v}}$ (或$\Phi_{R,\mathbf{v}}$) 是线性同构。然而，实际情况与我们的期望有一定的差距。

定义 3.2.8. 设 V 是一个复向量空间，函数 $f : V \to \mathbb{C}$ 满足：对任意 $\alpha \in \mathbb{C}$ 及 $\mathbf{x},\mathbf{y} \in V$，有

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}), \ f(\alpha\mathbf{x}) = \overline{\alpha} f(\mathbf{x}),$$

则我们称 f 是 V 到 $\mathbb{C}$ 的半线性函数。我们把 V 上所有半线性映射构成的集合记作 $V^\star$，并且在 $V^\star$ 上定义自然的加法和数乘如下：

$$\begin{aligned}
+ \ : \ & V^\star \times V^\star \longrightarrow V^\star \\
& (f, g) \longmapsto (f+g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}), \ \forall \mathbf{x} \in V \\
\cdot \ : \ & \mathbb{C} \times V^\star \longrightarrow V^\star \\
& (\lambda, f) \longmapsto (\lambda f)(\mathbf{x}) = \lambda f(\mathbf{x}), \ \forall \mathbf{x} \in V
\end{aligned}$$

容易验证 $V^\star$ 在上述运算下构成了 $\mathbb{C}$ 上的向量空间，并且如果 $\dim(V) = n$，则 $\dim(V^\star) = n$。

---

[1] 无穷维的复内积空间上 Bessel 不等式仍成立。

设 V,W 是两个复向量空间，如果映射 $f:V\to W$ 满足：对 $\forall\alpha,\beta\in\mathbb{C}$ 及 $\mathbf{x},\mathbf{y}\in V$，都有

$$f(\alpha\mathbf{x}+\beta\mathbf{y})=\overline{\alpha}f(\mathbf{x})+\overline{\beta}f(\mathbf{y})$$

则我们称 f 是 $V\to W$ 的半线性映射；如果存在半线性映射 $f:V\to W$ 是双射，则我们称 $f:V\to W$ 是 V 到 W 的半线性同构。

我们有下面的定理：

定理 3.2.5. 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 $\Phi_{L,\mathbf{v}}:\mathbf{x}\mapsto<\mathbf{v},\mathbf{x}>$ 是半线性函数，$\Phi_L:V\to V^\star$，$\mathbf{v}\mapsto\Phi_{L,\mathbf{v}}$ 是 V 到 $V^\star$ 的线性同构；$\Phi_{R,\mathbf{v}}:\mathbf{x}\mapsto<\mathbf{x},\mathbf{v}>$ 是线性函数，$\Phi_L:V\to V^*$，$\mathbf{v}\mapsto\Phi_{R,\mathbf{v}}$ 是 V 到 V^* 的半线性同构。也就是说，V 上的每个线性函数都是 $\mathbf{x}\mapsto<\mathbf{x},\mathbf{v}>$ 的形式，V 上的每个半线性函数都是 $\mathbf{x}\mapsto<\mathbf{v},\mathbf{x}>$ 的形式。

与命题3.1.5的证明方法类似，留作练习。实际上，我们还看到

$$V^*\longrightarrow V^\star,\ (\mathbf{x}\mapsto<\mathbf{x},\mathbf{v}>)\longmapsto(\mathbf{x}\mapsto<\mathbf{v},\mathbf{x}>)$$

是 V^* 到 $V^\star$ 的同构，我们称半线性函数 $(\mathbf{x}\mapsto<\mathbf{v},\mathbf{x}>)$ 是线性函数 $(\mathbf{x}\mapsto<\mathbf{x},\mathbf{v}>)$ 的共轭。

我们同样可以在 Hermitian 空间上定义内积关于一组基的度量矩阵 (Gram 矩阵)。设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，由于 $<\cdot,\cdot>$ 是 V 上的埃尔米特型，故任取 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$，埃尔米特型 $<\cdot,\cdot>$ 对应的矩阵为

$$G=\begin{pmatrix}<\mathbf{v}_1,\mathbf{v}_1> & <\mathbf{v}_1,\mathbf{v}_2> & \cdots & <\mathbf{v}_1,\mathbf{v}_n>\\ <\mathbf{v}_2,\mathbf{v}_1> & <\mathbf{v}_2,\mathbf{v}_2> & \cdots & <\mathbf{v}_2,\mathbf{v}_n>\\ \vdots & \vdots & & \vdots\\ <\mathbf{v}_n,\mathbf{v}_1> & <\mathbf{v}_n,\mathbf{v}_2> & \cdots & <\mathbf{v}_n,\mathbf{v}_n>\end{pmatrix}$$

我们称上面的 G 为内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的**度量矩阵**或 Gram 矩阵。由 3.2.1 小节的论述可知，如果内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 下的度量矩阵分别为 G 和 G'，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$，$A\in\mathrm{GL}_n(\mathbb{C})$，则 $G'=A^tG\overline{A}$。显然 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基 $\iff$ 其度量矩阵为 E_n。于是，如果 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 都是 V 的标准正交基，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$，则 $E_n=A^tE_n\overline{A}$，即 $A^t\overline{A}=E_n$，或者说 $\overline{A}^tA=E_n$，即 $A^*A=E_n$。反之，如果矩阵 A 满足 $A^*A=E_n$，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，令 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$，则显然 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是 V 的基 (A 显然可逆)，并且按上面的讨论，内积在 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的度量矩阵为 $A^*E_nA=E_n$，即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是标准正交基。这就引出了下面的酉矩阵的定义。

定义 3.2.9. 设 $A\in\mathrm{GL}_n(\mathbb{C})$，如果 A 满足 $A^*A=E$，即 $A^*=A^{-1}$，则称 A 为酉矩阵 (unitary matrix)。容易验证 $\mathbb{C}$ 上所有 n 阶酉矩阵在矩阵乘法运算下构成了一个群，我们把这个群称为酉群 (unitary group)，记作 $U(n)$。显然 $O(n)$ 是 $U(n)$ 的子群。

容易看出酉矩阵就是标准正交基之间的过渡矩阵。同样我们也可以定义酉算子如下：

定义 3.2.10. 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}$ 是 V 上的可逆的线性算子。如果 $\mathcal{A}$ 还满足：对任意的 $\mathbf{x},\mathbf{y}\in V$，$<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathcal{A}$ 是 V 上的酉算子 (unitary operator) 或酉变换。

容易看出酉算子一定是 Hermitian 空间的自同构。我们同样有以下定理：

定理 3.2.6. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的线性算子，则以下条件等价：

(1) $\mathcal{A}$ 是酉算子；

(2) $\mathcal{A}$ 在标准正交基下对应的矩阵是酉矩阵；

(3) $\mathcal{A}$ 将标准正交基映到标准正交基；

(4) $\forall \mathbf{x} \in V$, $\|\mathcal{A}\mathbf{x}\| = \|\mathbf{x}\|$。

与定理3.1.5的证明方法类似，留作练习。

推论 3.2.3. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 V 上所有酉算子在算子复合运算下构成群，记作 $U(V)$，则 $U(V)$ 同构于酉群 $U(n)$。

留作练习。

我们看到，Hermitian 空间的性质与欧氏空间的性质是极为类似的，并且我们可以把欧氏空间看成 Hermitian 空间“限制”在实数域上的特例。

3.2.3 赋范向量空间与度量空间

还记得我们在本章开头提到的目标吗？我们希望在向量空间上复原“通常”的距离结构。现在我们已经知道了如何用内积诱导出范数和距离，但我们希望用公理化的方法定义范数和距离。这就是本小节的主要内容。

定义 3.2.11. 设 E 是一个非空集合，函数：$\mathrm{d}: E \times E \to [0, +\infty)$ 满足以下性质：

(1) 正定性：$\forall u, v \in E$，$\mathrm{d}(u, v) \geqslant 0$，并且等号当且仅当 $u = v$ 时成立；

(2) 对称性：$\forall u, v \in E$，$\mathrm{d}(u, v) = \mathrm{d}(v, u)$；

(3) 三角不等式：$\forall u, v, w \in E$，$\mathrm{d}(u, w) \leqslant \mathrm{d}(u, v) + \mathrm{d}(v, w)$。

则我们称函数 d 是集合 E 上的一个距离或度量 (distance/metric) 函数，称 (E, d) 是一个度量空间 (metric space)。我们把 E 中的任意一个元素称为一个点 (point)，称 $d(u, v)$ 为点 u 和 v 之间的距离。

下面我们给出一些度量空间的例子。

例 3.2.2. (1) 实（或）复的内积空间是度量空间，其度量函数可以由内积诱导出来，即 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \sqrt{<\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}>}$。

(2) 在集合 $C[a, b]$ (闭区间 $[a, b]$ 上的所有连续函数所构成的集合) 上可以定义不同的距离函数：$\forall f(x), g(x) \in C[a, b]$，以下函数

$$d_1(f, g) = \sqrt{\int_a^b |f(t) - g(t)|^2 \, \mathrm{d}t};$$

$$d_2(f, g) = \int_a^b |f(t) - g(t)| \, \mathrm{d}t;$$

$$d_3(f, g) = \max_{a \leqslant t \leqslant b} |f(t) - g(t)|.$$

都是 $C[a, b]$ 上的度量函数。

(3) 设 p 是大于 1 的实数，定义集合

$$l^p(\mathbb{R}) = \left\{ \mathbf{x} = \{x_n\}_{n=1}^{\infty} \;\middle|\; x_n \in \mathbb{R},\ \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$$

容易验证 $l^p(\mathbb{R})$ 在通常的加法和数乘下是 $\mathbb{R}$ 上的无穷维向量空间。设 $\mathbf{x} = \{x_n\}_{n=1}^{\infty}$，$\mathbf{y} = \{y_n\}_{n=1}^{\infty} \in l^p(\mathbb{R})$，我们可以在 $l^p(\mathbb{R})$ 上定义距离函数：

$$\mathrm{d}(\mathbf{x},\mathbf{y}) = (\sum_{n=1}^{\infty} |x_n - y_n|^p)^{\frac{1}{p}}$$

则 $l^p(\mathbb{R})$ 在此度量函数下构成了一个度量空间。该度量函数的三角不等式被称为闵可夫斯基 (Minikowski) 不等式，其证明留作习题。我们在以后的泛函分析课程中会经常用到它。

(4) 对任意集合 E，我们可以定义离散度量：$\forall a, b \in E$，$\mathrm{d}(a,b) = \begin{cases} 1, & a \neq b \\ 0, & a = b \end{cases}$。验证 (E, d) 是一个度量空间是容易的，留作练习。以后在拓扑学课程中我们会知道，离散度量所诱导的拓扑称为离散拓扑，其性质是十分简单的。

度量空间上有比较好的拓扑和分析结构，因此我们可以进行更复杂的操作。

定义 3.2.12. 设 (E, d) 是度量空间，对 $\forall a \in E$ 及 $r > 0$，我们定义

$$\begin{aligned} B(a,r) &= \{x \in E \mid \mathrm{d}(a,x) < r\}, \\ \overline{B}(a,r) &= \{x \in E \mid \mathrm{d}(a,x) \leqslant r\}, \\ S(a,r) &= \{x \in E \mid \mathrm{d}(a,x) = r\}. \end{aligned}$$

我们把 $B(a,r), \overline{B}(a,r), S(a,r)$ 分别称为以 a 为中心，r 为半径的开球、闭球、和球面。

定义 3.2.13. 设 (E, d) 是度量空间，序列 $\{x_n\}_{n=1}^{\infty} \subset E$。

(1) 如果 $\forall \varepsilon$，$\exists N = N(\varepsilon) \in \mathbb{N}$ 使得对 $\forall m, n > N$ 都有 $\mathrm{d}(x_m, x_n) < \varepsilon$，则称 $\{x_n\}$ 是柯西 (Cauchy) 序列。

(2) 如果存在 $x_0 \in E$ 满足 $\forall \varepsilon, \exists N = N(\varepsilon) \in \mathbb{N}$ 使得对 $\forall n > N$ 都有 $\mathrm{d}(x_n, x_0) < \varepsilon$ (即 $\lim_{n\to\infty} \mathrm{d}(x_n, x_0) = 0$)，则称序列 $\{x_n\}$ 按距离收敛到 x_0，或称 x_0 是序列 $\{x_n\}$ 的极限。

如果 (E, d) 中每一个柯西序列都能收敛到 E 中的某个元素，则我们称 (E, d) 是一个完备度量空间。

例 3.2.3. $\mathbb{R}$ 和 $\mathbb{C}$ 在通常的度量下显然 (如果大家还记得数学分析课程中实数理论相关的内容的话) 都是完备度量空间。实际上，$\mathbb{R}^n$ 和 $\mathbb{C}^n$ 在下面的度量函数下都是完备度量空间 (设 $\mathbf{x} = (x_1, \ldots, x_n)^t$，$\mathbf{y} = (y_1, \ldots, y_n)^t$ 是 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 中的两个点)：

$$\begin{aligned} \mathrm{d}_2(\mathbf{x},\mathbf{y}) &= \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}; \\ \mathrm{d}_1(\mathbf{x},\mathbf{y}) &= \sum_{i=1}^{n} |x_i - y_i|; \\ \mathrm{d}_\infty(\mathbf{x},\mathbf{y}) &= \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}. \end{aligned}$$

这是因为这三个度量之间有如下的关系：对任意 $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$(或 $\mathbb{C}^n$)，有

$$\begin{gathered}\frac{1}{\sqrt{n}}\,\mathrm{d}_1(\mathbf{x},\mathbf{y})\leqslant\mathrm{d}_2(\mathbf{x},\mathbf{y})\leqslant\sqrt{n}\,\mathrm{d}_1(\mathbf{x},\mathbf{y})\\ \mathrm{d}_\infty(\mathbf{x},\mathbf{y})\leqslant\mathrm{d}_2(\mathbf{x},\mathbf{y})\leqslant\sqrt{n}\,\mathrm{d}_\infty(\mathbf{x},\mathbf{y})\end{gathered}\tag{3.2.2}$$

设 $\{\mathbf{x}_k=(x_{k1},\cdots,x_{kn})^t\}_{k=1}^{\infty}$ 是 d_∞ 距离下的柯西序列，即 $\forall\varepsilon>0$，$\exists N\in\mathbb{N}$ 使得 $\forall l,m>N$，都有 $\max\{|x_{l1}-x_{m1}|,|x_{l2}-x_{m2}|,\ldots,|x_{ln}-x_{mn}|\}<\varepsilon$。这说明 $\{x_{k1}\}_{k=1}^{\infty}$，$\{x_{k2}\}_{k=1}^{\infty}$，$\cdots\cdots,\{x_{kn}\}_{k=1}^{\infty}$ 都是 $\mathbb{R}$(或 $\mathbb{C}$) 中的柯西序列，由 $\mathbb{R}$(或 $\mathbb{C}$) 的完备性可设 $\lim\limits_{n\to\infty}x_{k1}=y_1,\cdots\cdots,\lim\limits_{n\to\infty}x_{kn}=y_n$，令 $\mathbf{y}=(y_1,\ldots,y_n)^t$，则容易验证 $\lim\limits_{k\to\infty}\mathrm{d}_\infty(\mathbf{x}_k,\mathbf{y})=0$，即序列 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 按 d_∞ 距离收敛到 $\mathbf{y}$，这就证明了 $\mathbb{R}^n$(或 $\mathbb{C}^n$) 在 d_∞ 距离下是完备度量空间。利用式 (3.2.2) 给出的三个度量函数之间的大小关系即可验证 $\mathbb{R}^n$(或 $\mathbb{C}^n$) 在这三个距离下都是完备的。

例 3.2.4. $C[a,b]$ 在度量函数 $d_3(f,g)=\max\limits_{a\leqslant t\leqslant b}|f(t)-g(t)|$ 下是完备度量空间，但在 $d_2(f,g)=\int_a^b|f(t)-g(t)|\,\mathrm{d}t$ 下就不是完备的。证明完全是分析的技术，可参考《数学分析 · 第二卷》卓里奇，高等教育出版社的第九章 §5 的例 4 和例 5。

我们已经给出了度量的公理化定义，下面我们给出范数的公理化定义。

定义 3.2.14. 设 V 是一个实或复的向量空间 (不一定是有限维的)，在 V 上固定一个满足如下性质的函数 $\|\cdot\|:V\to[0,+\infty)$：

(1) $\forall\mathbf{x}\in V$，$\|\mathbf{x}\|\geqslant 0$，并且 $\|\mathbf{x}\|=0\iff\mathbf{x}=\mathbf{0}$；

(2) $\forall\lambda\in\mathbb{R}$，$\mathbf{x}\in V$，都有 $\|\lambda\mathbf{x}\|=|\lambda|\|\mathbf{x}\|$；

(3) $\forall\mathbf{x},\mathbf{y}\in V$，有 $\|\mathbf{x}+\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|$。

则我们称 $(V,\|\cdot\|)$ 是一个赋范向量空间，$\|\cdot\|$ 被称为 V 上的范数。

显然赋范向量空间一定是度量空间，因为范数诱导了一个自然的距离 $\mathrm{d}(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|$ (验证这确实是一个距离留作练习)。如果一个赋范向量空间还是完备距离空间，则我们称它为完备赋范空间 (Banach space)。特别地，如果一个实或复的内积空间作为度量空间是完备的，则我们称它为 Hilbert 空间。

本小节介绍的这些空间之间的关系可以用下图表示：

图 3.2-1 各种空间的关系 (注意：这里内积空间指实或复内积空间，不包括辛空间。)

思考题 3.2.1. 给出以下反例：

(1) 一个赋范向量空间不是内积空间；

(2) 一个带度量的向量空间不是赋范向量空间；

(3) 不完备的内积空间；

(4) 不完备的赋范向量空间；

(5) 不完备的度量空间。

最后，显然欧氏空间和 Hermitian 空间一定是完备的。我们可以在其上讨论收敛性。

定理 3.2.7. 设 $(V, <\cdot,\cdot>)$ 是欧氏空间或 Hermitian 空间，则 V 中的序列 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 收敛到 $\mathbf{x} \in V \iff \lim\limits_{k\to\infty} <\mathbf{x}-\mathbf{x}_k, \mathbf{y}>= 0$ 对任意的 $\mathbf{y} \in V$ 成立。

证明. $(\Longrightarrow)$ 如果 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 收敛到 $\mathbf{x} \in V$，即 $\mathrm{d}(\mathbf{x}, \mathbf{x}_k) = \|\mathbf{x}-\mathbf{x}_k\| \xrightarrow{k\to\infty} 0$，故对任意的 $\mathbf{y} \in V$，由 CBS 不等式有

$$<\mathbf{x}-\mathbf{x}_k, \mathbf{y}>\leqslant \|\mathbf{x}-\mathbf{x}_k\| \cdot \|\mathbf{y}\| \xrightarrow{k\to\infty} 0$$

即 $\lim\limits_{k\to\infty} <\mathbf{x}-\mathbf{x}_k, \mathbf{y}>= 0$。
$(\Longleftarrow)$ 取 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，则有

$$\lim_{k\to\infty} <\mathbf{x}-\mathbf{x}_k, \mathbf{e}_i>= 0,\ \forall i \in \{1, \ldots, n\}.$$

于是 $\|\mathbf{x}-\mathbf{x}_k\|^2 = \sum_{i=1}^{n} <\mathbf{x}-\mathbf{x}_k, \mathbf{e}_i> \xrightarrow{k\to\infty} 0$ (这里用到了 V 是有限维的)，所以 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 收敛到 $\mathbf{x}$。 □

本小节的内容只是关于度量空间和赋范向量空间的最简单的介绍，更详细的内容读者可以参考泛函分析的标准教材，如《A Course in Functional Analysis》, John B.Conway, GTM96 或《泛函分析讲义》，张恭庆，北京大学出版社。

3.3 内积空间上的线性算子 I：自伴随算子

本节和接下来两节中我们所说的内积空间特指欧氏空间和 Hermitian 空间 (不包括辛空间)，特别地，我们要求空间都是有限维的。此外，我们提到的域 $\mathbb{K}$ 都是 $\mathbb{R}$ 或 $\mathbb{C}$。现在我们开始讨论内积空间上的线性算子。我们的最终目标是 3.5 节的谱定理，也就是说，我们需要寻找一类内积空间中的算子，它们在某组基 (更进一步，这组基还是标准正交基) 下的矩阵是对角矩阵。用矩阵的语言来说，我们需要寻找一类矩阵，它们能够合同[1]并相似到对角矩阵 (即我们希望 $\exists T\in U(n)$ (或$O(n)$) 使得 T^*AT (或$T'AT$) $=T^{-1}AT=\Lambda$，Λ 是对角阵)。

直接讨论上面的问题是困难的。为此，我们从 (半) 双线性型和线性算子的关系入手，先通过型定义伴随算子的概念，然后我们讨论一种特殊的可对角化的算子：自伴随算子。

定义 3.3.1. 设 $(V,<\cdot,\cdot>)$ 是内积空间，$\mathcal{A}$ 是 V 上的线性算子，则我们可以定义：

$$f_{\mathcal{A}}:V\times V\longrightarrow\mathbb{K}\ (\mathbb{K}=\mathbb{R}或\mathbb{C})$$
$$(\mathbf{x},\mathbf{y})\longmapsto<\mathcal{A}\mathbf{x},\mathbf{y}>.$$

容易验证：如果 V 是欧氏空间，则 $f_{\mathcal{A}}$ 是双线性型；如果 V 是 Hermitian 空间，则 $f_{\mathcal{A}}$ 是半双线性型型。由于双线性型和半双线性型型的性质类似，因此我们把欧氏空间上的双线性型和 Hermitian 空间上的半双线性型型统称为 $\theta-$ 线性型 ($\theta=2$ 表示欧氏空间上的双线性型，$\theta=\frac{3}{2}$ 表示 Hermitian 空间上的半双线性型)。我们采用统一的记号 $L_\theta(V,\mathbb{K})$ 来表示 V 上的 $\theta-$ 线性型 (同样地，$\theta=2,\mathbb{K}=\mathbb{R}$ 表示欧氏空间上的双线性型，$\theta=\frac{3}{2},\mathbb{K}=\mathbb{C}$ 表示 Hermitian 空间上的半双线性型)。如果 $\dim(V)=n$，则显然 $L_\theta^+(V,\mathbb{K})$ 是 $\mathbb{R}$ 或 $\mathbb{C}$ 上的 n^2 维向量空间 (定理1.6.1或 3.2.1 小节的论证)。

定理 3.3.1. 设 $(V,<\cdot,\cdot>)$ 是内积空间，定义如下的映射 $\varPhi$：

$$\varPhi:\mathrm{End}(V)\longrightarrow L_\theta(V,\mathbb{K})$$
$$\mathcal{A}\longmapsto f_{\mathcal{A}}\quad (f_{\mathcal{A}}的定义见于定义3.3.1).$$

则 $\varPhi$ 是 $\mathrm{End}(V)$ 到 $L_\theta(V,\mathbb{K})$ 的线性同构。

证明. 首先 $\varPhi$ 是线性映射：任取 $\mathcal{A},\mathcal{B}\in\mathrm{End}(V)$ 及 $\alpha,\beta\in\mathbb{K}(=\mathbb{R}或\mathbb{C})$，则对 $\forall\mathbf{x},\mathbf{y}\in V$，我们有

$$<(\alpha\mathcal{A}+\beta\mathcal{B})\mathbf{x},\mathbf{y}>=\alpha<\mathcal{A}\mathbf{x},\mathbf{y}>+\beta<\mathcal{B}\mathbf{x},\mathbf{y}>$$

因此 $f_{\alpha\mathcal{A}+\beta\mathcal{B}}=\alpha f_{\mathcal{A}}+\beta f_{\mathcal{B}}$，即 $\varPhi$ 是线性映射。

其次 $\varPhi$ 是单射：只需证 $\ker(\varPhi)=\{\mathcal{O}\}$。如果 $f_{\mathcal{A}}=0$，即任取 $\mathbf{x},\mathbf{y}\in V$，$<\mathcal{A}\mathbf{x},\mathbf{y}>=0$，则必有 $\mathcal{A}\mathbf{x}=\mathbf{0}$, $\forall\mathbf{x}\in V$，此即 $\mathcal{A}=\mathcal{O}$，即 $\ker(\varPhi)=\{\mathcal{O}\}$。

最后 $\varPhi$ 是满射：这只需注意到 $\dim(\mathrm{End}(V))=\dim(L_\theta(V,\mathbb{K}))=n^2$ 即可证明。

综上所述，$\varPhi$ 是线性同构。 □

下面我们考虑在同一组标准正交基下算子 $\mathcal{A}$ 和 $\theta-$ 线性型 $f_{\mathcal{A}}$ 所对应的矩阵之间的关系。

推论 3.3.1. 条件同定理3.3.1，设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，$\mathcal{A}\in\mathrm{End}(V)$ 在此基底下的矩阵为 A，则 $f_{\mathcal{A}}$ 在此基底下的矩阵为 A'。

---

[1] 复矩阵则变成"共轭合同"，即 $A\in M_n(\mathbb{C})$ 与 T^*AT 共轭合同。

证明. 设 $A = (\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(n)}) = (a_{ij})_{n\times n}$，则对每个 $i \in \{1, \ldots, n\}$，有

$$\mathscr{A}\mathbf{e}_i = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \mathbf{A}^{(i)} = \sum_{k=1}^{n} a_{ki}\mathbf{e}_k,$$

所以 $< \mathscr{A}\mathbf{e}_i, \mathbf{e}_j >= a_{ji}$, $\forall i, j \in \{1, \ldots, n\}$。而 $f_{\mathscr{A}}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵就是 $(< \mathscr{A}\mathbf{e}_i, \mathbf{e}_j >)_{n\times n} = (a_{ji})_{n\times n}$，此即 $f_{\mathscr{A}}$ 在此基底下的矩阵为 A^t。 □

现在我们可以定义一个算子的伴随算子了。

定理 3.3.2. 设 $(V, < \cdot, \cdot >)$ 是内积空间，$\mathscr{A}$ 是 V 上的线性算子，则存在唯一的 $\mathscr{B} \in \mathrm{End}(V)$ 使得 $\forall \mathbf{x}, \mathbf{y} \in V$，有 $< \mathbf{x}, \mathscr{B}\mathbf{y} >=< \mathscr{A}\mathbf{x}, \mathbf{y} >$。我们称 $\mathscr{B}$ 为 $\mathscr{A}$ 的 (Hilbert) 伴随算子 (或共轭算子，adjoint operator)，记作 $\mathscr{B} = \mathscr{A}^*$。

证明. 先证存在性：任取 $\mathbf{y} \in V$，我们只要给出 $\mathscr{B}\mathbf{y} \in V$ 的取法就证明了 $\mathscr{B}$ 的存在性。首先，$\mathbf{y}$ 给出了一个 V 上的线性函数 f:

$$\begin{aligned} f\ :\ V &\longrightarrow \mathbb{K} \\ \mathbf{x} &\longmapsto < \mathscr{A}\mathbf{x}, \mathbf{y} > . \end{aligned}$$

而由定理3.1.5和定理3.2.5，我们知道一定存在唯一的 $\mathbf{v} \in V$ 使得 f 是 $\mathbf{x} \mapsto < \mathbf{x}, \mathbf{v} >$ 的形式。我们定义 $\mathscr{B}\mathbf{y} = \mathbf{v}$，则首先 $\mathscr{B}$ 确实是一个映射 ($\mathbf{v}$ 的存在唯一性)；其次 $\mathscr{B}$ 是线性的，这是因为：如果 $\mathscr{B}\mathbf{y}_1 = \mathbf{v}_1$, $\mathscr{B}\mathbf{y}_2 = \mathbf{v}_2$，即 $< \mathscr{A}\mathbf{x}, \mathbf{y}_1 >=< \mathbf{x}, \mathbf{v}_1 >$, $< \mathscr{A}\mathbf{x}, \mathbf{y}_2 >=< \mathbf{x}, \mathbf{v}_2 >$ 对 $\forall \mathbf{x} \in V$ 成立，则对任意的 $\alpha, \beta \in \mathbb{K}$，有

$$\begin{aligned} < \mathscr{A}\mathbf{x}, \alpha\mathbf{y}_1 + \beta\mathbf{y}_2 > &= \overline{\alpha} < \mathscr{A}\mathbf{x}, \mathbf{y}_1 > + \overline{\beta} < \mathscr{A}\mathbf{x}, \mathbf{y}_2 > \\ &= \overline{\alpha} < \mathbf{x}, \mathbf{v}_1 > + \overline{\beta} < \mathbf{x}, \mathbf{v}_2 > \\ &=< \mathbf{x}, \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 > \end{aligned}$$

这就说明 $\mathscr{B}(\alpha\mathbf{y}_1 + \beta\mathbf{y}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$，即 $\mathscr{B}$ 是线性算子。

再证唯一性：设 $\mathscr{B}' \in \mathrm{End}(V)$ 也满足 $\forall \mathbf{x}, \mathbf{y} \in V$, $< \mathbf{x}, \mathscr{B}'\mathbf{y} >=< \mathscr{A}\mathbf{x}, \mathbf{y} >$，那么，对任意的 $\mathbf{x}, \mathbf{y} \in V$，有 $< \mathbf{x}, (\mathscr{B}' - \mathscr{B})\mathbf{y} >= 0$，这说明 $(\mathscr{B}' - \mathscr{B})\mathbf{y} = \mathbf{0}$ 对任意的 $\mathbf{y} \in V$ 成立，此即 $\mathscr{B}' = \mathscr{B}$，唯一性成立。 □

推论 3.3.2. 条件同定理3.3.2，设 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组标准正交基，$\mathscr{A} \in \mathrm{End}(V)$ 在此基底下的矩阵为 A，则伴随算子 $\mathscr{A}^*$ 在此基底下的矩阵为 $\overline{A^t} = A^*$。

证明. 设 $\mathscr{A}^*$ 在标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵为 $B = (b_{ij})_{n\times n}$，则对每个 $j \in \{1, \ldots, n\}$，有

$$\mathscr{A}^*\mathbf{e}_j = \sum_{k=1}^{n} b_{kj}\mathbf{e}_k,$$

所以

$$< \mathscr{A}\mathbf{e}_i, \mathbf{e}_j >=< \mathbf{e}_i, \mathscr{A}^*\mathbf{e}_j >= \overline{b_{ij}}, \ \forall i, j \in \{1, \ldots, n\}.$$

另一方面，由推论3.3.1，设 $\mathscr{A}$ 在此基底下的矩阵是 $A = (a_{ij})_{n\times n}$，则 $< \mathscr{A}\mathbf{e}_i, \mathbf{e}_j >= a_{ji}$。这说明 $a_{ji} = \overline{b_{ij}}$, $\forall i, j \in \{1, \ldots, n\}$，即 $B = A^*$。 □

既然每个线性算子在内积空间上都存在唯一的伴随算子，那么我们可以考察“取伴随”这个运算的性质。

定理 3.3.3. 设 $(V, <\cdot,\cdot>)$ 是内积空间，定义取伴随的映射 $*: \mathrm{End}(V) \to \mathrm{End}(V)$, $\mathcal{A} \mapsto \mathcal{A}^*$，则 $*$ 映射满足以下性质：对任意的 $\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$ 及 $\lambda \in \mathbb{K}$，有

(1) $(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$；

(2) $(\lambda\mathcal{A})^* = \overline{\lambda}\mathcal{A}^*$；

(3) $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$；

(4) $(\mathcal{A}^*)^* = \mathcal{A}$。

证明是简单的，留作练习。实际上，我们可以将以上性质抽象出来，作为 $*-$ 代数的定义 (当然，我们还会附加一些范数的要求)。更详细的论述会在算子代数课程中介绍，读者可以参考《A Short Course on Spectral Theory》, William Arveson, GTM209 以及更深入的《An Invitation to C*-Algebra》, William Arveson, GTM39。

下面我们定义几类特殊的线性算子。

定义 3.3.2. 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$，则

(1) 如果 $\mathcal{A}^* = \mathcal{A}$，则称 $\mathcal{A}$ 为自伴随 (self-adjoint) 算子或 Hermitian 算子；如果 $\mathcal{A}^* = -\mathcal{A}$，则称 $\mathcal{A}$ 为斜自伴随 (skew self-adjoint) 算子或斜 Hermitian 算子 (也有的书译成反 Hermitian 算子)。欧氏空间上的自伴随算子也称为对称算子。

(2) 如果 $\forall \mathbf{x}, \mathbf{y} \in V$，都有 $<\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{x}> = <\mathbf{x}, \mathbf{x}>$，则称 $\mathcal{A}$ 为保距 (isometric) 算子。保距算子实际上是欧氏空间上的正交算子和 Hermitian 空间上的酉算子的统称，我们会在下一节中证明这一点。

(3) 如果 $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$，则称 $\mathcal{A}$ 为正规 (normal) 算子。

对应地，设 $A \in M_n(\mathbb{K})$ $(\mathbb{K} = \mathbb{R}或\mathbb{C})$，则

(1) 如果 $A^* = A$，则称 A 为自伴随矩阵或 Hermitian 矩阵；如果 $A^* = -A$，则称 $\mathcal{A}$ 为斜自伴随矩阵或斜 Hermitian 矩阵 (实矩阵则分别是对称矩阵和斜对称矩阵)。

(2) 如果 $A^*A = AA^* = E$，则称 A 为酉矩阵 (若 A 是实矩阵则称为正交矩阵)。

(3) 如果 $AA^* = A^*A$，则称 A 为正规矩阵。

显然上面的每一种算子在 V 的任意一组标准正交基下的矩阵就是下面对应类型的矩阵。

本节我们主要考虑自伴随算子 (和矩阵) 的情形，接下来的两节我们会依次考虑保距算子和正规算子的情形。

以下是本节的主要结论：

定理 3.3.4. 设 $(V, <\cdot,\cdot>)$ 是 n 维内积空间，$\mathcal{A} \in \mathrm{End}(V)$ 是自伴随算子，则存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}$ 在这组基下的矩阵是实对角矩阵，即

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad \lambda_1, \ldots, \lambda_n \in \mathbb{R}(\lambda_1, \ldots, \lambda_n\text{中可以有相等的元素}).$$

我们也可以用矩阵语言将上述定理重新叙述如下：

定理 3.3.4′ 设 $A \in M_n(\mathbb{C})$ 满足 $A^* = A$(实矩阵则是 $A^t = A$)，则存在酉矩阵 (实情形则是正交矩阵)B 使得

$$B^*AB(\text{或}B^tAB) = B^{-1}AB = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}(\lambda_1, \dots, \lambda_n\text{中可以有相等的元素}).$$

为了证明上述定理，我们先给出下面的引理。为了证明的方便，我们先只考虑 Hermitian 空间上的自伴随算子，欧氏空间上对称算子的情形我们会在稍后给出说明。

引理 3.3.1. 设 $(V, <\cdot,\cdot>)$ 是 Hermitian 空间，$\mathcal{A} \in \mathrm{End}(V)$ 是自伴随算子，则

(1) $\forall p(t) \in \mathbb{R}[t]$，线性算子 $p(\mathcal{A})$ 也是自伴随算子。特别地，对任意的 $\lambda \in \mathbb{R}$，有 $\mathcal{A} - \lambda\mathcal{E}$ 是自伴随算子。

(2) 自伴随算子一定有特征值和特征向量，并且其所有特征值都是实数。

(3) 如果 W 是 $\mathcal{A}-$ 子空间，则 $W^\perp$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 都是自伴随算子。

(4) $\ker(\mathcal{A}) = (\mathrm{im}(\mathcal{A}^*))^\perp$，$\mathrm{im}(\mathcal{A}) = (\ker(\mathcal{A}^*))^\perp$。

(5) $\forall k \in \mathbb{Z}^+$，$\ker(\mathcal{A}^k) = \ker(\mathcal{A})$。

(6) 如果 λ, μ 是 $\mathcal{A}$ 的两个不同的特征值，$\mathbf{x}$ 和 $\mathbf{y}$ 分别是 λ 和 μ 对应的特征向量，则 $\mathbf{x} \perp \mathbf{y}$。

证明. (1) 首先，由定理3.3.3的 (3) 可知 $\forall k \in \mathbb{Z}^+$，有 $(\mathcal{A}^k)^* = (\mathcal{A}^*)^k$。不妨设 $p(t) = a_0 + a_1 t + \cdots + a_n t^n \in \mathbb{R}[t]$，则由于 $*$ 是半线性映射 (定理3.3.3的 (1) 和 (2))，我们有

$$[p(\mathcal{A})]^* = \overline{a_0}\mathcal{E} + \overline{a_1}\mathcal{A}^* + \cdots + \overline{a_n}(\mathcal{A}^*)^n = a_0\mathcal{E} + a_1\mathcal{A} + \cdots + a_n\mathcal{A}^n = p(\mathcal{A}).$$

特例是显然的。

(2) 设 $\chi_{\mathcal{A}}(t) \in \mathbb{C}[t]$ 是 $\mathcal{A}$ 的特征多项式，则 $\chi_{\mathcal{A}}(t)$ 在复数域 $\mathbb{C}$ 上一定有根，从而一定有特征值和特征向量。下面我们证明 $\chi_{\mathcal{A}}(t)$ 的所有根都是实根。为此，任取 $\lambda \in \mathbb{C}$ 是 $\chi_{\mathcal{A}}(t)$ 的根，$\mathbf{v} \in V$ 是 λ 对应的非零特征向量。则

$$\lambda < \mathbf{v}, \mathbf{v} >=< \lambda\mathbf{v}, \mathbf{v} >=< \mathcal{A}\mathbf{v}, \mathbf{v} >=< \mathbf{v}, \mathcal{A}\mathbf{v} >=< \mathbf{v}, \lambda\mathbf{v} >= \overline{\lambda} < \mathbf{v}, \mathbf{v} > .$$

由于 $< \mathbf{v}, \mathbf{v} > \neq 0$，故 $\lambda = \overline{\lambda}$，即 $\lambda \in \mathbb{R}$。

(3) 先证 $W^\perp$ 是 $\mathcal{A}-$ 子空间。任取 $\mathbf{x} \in W^\perp$，我们需要证明 $\mathcal{A}\mathbf{x} \in W^\perp$，即 $\mathcal{A}\mathbf{x} \perp W$。任取 $\mathbf{y} \in W$，则由 W 是 $\mathcal{A}-$ 子空间可知 $\mathcal{A}\mathbf{y} \in W$，于是 $< \mathcal{A}\mathbf{x}, \mathbf{y} >=< \mathbf{x}, \mathcal{A}\mathbf{y} >= 0$，即 $\mathcal{A}\mathbf{x} \perp W$ 成立。

$\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 都是自伴随算子是显然的。

(4) 先证明 $\ker(\mathcal{A}) = (\mathrm{im}(\mathcal{A}^*))^\perp$。任取 $\mathbf{x} \in \ker(\mathcal{A})$，则对任意的 $\mathcal{A}^*\mathbf{y} \in \mathrm{im}(\mathcal{A}^*)$，我们有

$$< \mathbf{x}, \mathcal{A}^*\mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} >=< \mathbf{0}, \mathbf{y} >= 0,$$

即 $\mathbf{x} \perp \operatorname{im}(\mathcal{A}^*)$，所以 $\mathbf{x} \in (\operatorname{im}(\mathcal{A}^*))^\perp$，$\ker(\mathcal{A}) \subset (\operatorname{im}(\mathcal{A}^*))^\perp$。反过来，任取 $\mathbf{z} \in (\operatorname{im}(\mathcal{A}^*))^\perp$，即 $\mathbf{z} \perp \mathcal{A}^*\mathbf{y}$，$\forall \mathbf{y} \in V$。这说明

$$< \mathcal{A}\mathbf{z}, \mathbf{y} >=< \mathbf{z}, \mathcal{A}^*\mathbf{y} >= 0,\ \forall \mathbf{y} \in V$$

所以只能是 $\mathcal{A}\mathbf{z} = \mathbf{0}$，即 $\mathbf{z} \in \ker(\mathcal{A})$，$\ker(\mathcal{A}) \supset (\operatorname{im}(\mathcal{A}^*))^\perp$。所以 $\ker(\mathcal{A}) = (\operatorname{im}(\mathcal{A}^*))^\perp$。以 $\mathcal{A}^*$ 替换上式中的 $\mathcal{A}$ 可知 $\ker(\mathcal{A}^*) = (\operatorname{im}(\mathcal{A}^{**}))^\perp = \operatorname{im}(\mathcal{A})^\perp$，即 $\operatorname{im}(\mathcal{A}) = (\ker(\mathcal{A}^*))^\perp$。

(5) 我们先证明：$\forall k \in \mathbb{Z}^+$，$\ker(\mathcal{A}^k) = \ker(\mathcal{A}^{k+1})$。首先 $\ker(\mathcal{A}^k) \subset \ker(\mathcal{A}^{k+1})$ 是显然成立的，因此我们只需要证明 $\ker(\mathcal{A}^k) \supset \ker(\mathcal{A}^{k+1})$ 即可。任取 $\mathbf{x} \in \ker(\mathcal{A}^{k+1})$，我们有 $\mathcal{A}^{k+1}\mathbf{x} = 0$，因此

$$\|\mathcal{A}^k\mathbf{x}\|^2 =< \mathcal{A}^k\mathbf{x}, \mathcal{A}^k\mathbf{x} >=< \mathcal{A}^{k+1}\mathbf{x}, \mathcal{A}^{k-1}\mathbf{x} >= 0,$$

这说明 $\mathcal{A}^k\mathbf{x} = \mathbf{0}$，即 $\mathbf{x} \in \ker(\mathcal{A}^k)$，$\ker(\mathcal{A}^k) \supset \ker(\mathcal{A}^{k+1})$。特别地，取 $k = 1$ 即有 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$，因此 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2) = \cdots = \ker(\mathcal{A}^k)$，$\forall k \in \mathbb{Z}^+$。

(6) 首先由 (2) 可知 λ 和 μ 都是实数。利用特征向量的定义和内积的半双线性性质，我们有

$$\begin{aligned}\lambda < \mathbf{x}, \mathbf{y} > &=< \lambda\mathbf{x}, \mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} >\\ &=< \mathbf{x}, \mathcal{A}\mathbf{y} >=< \mathbf{x}, \mu\mathbf{y} >\\ &= \overline{\mu} < \mathbf{x}, \mathbf{y} >= \mu < \mathbf{x}, \mathbf{y} >\end{aligned}$$

而 $\lambda \neq \mu$，故只能是 $< \mathbf{x}, \mathbf{y} >= 0$，即 $\mathbf{x} \perp \mathbf{y}$。

□

下面我们先简单地讨论一下欧氏空间对应的情形。显然上面引理的 (1)(3)(4)(5)(6) 在欧氏空间上都是成立的，只有 (2) 用到了 $\mathbb{C}$ 的代数闭域性质，因此需要重新讨论。实际上，(2) 在欧氏空间上也成立，理由如下。任取欧氏空间 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，则对称算子 $\mathcal{A}$ 在标准正交基下的矩阵显然是实对称矩阵。而实对称矩阵 $A \in M_n(\mathbb{R})$ 可以视作标准 Hermitian 空间 $\mathbb{C}^n$ 上的自伴随的线性算子，故由上面的引理，A 的所有特征值都是实数，并且每个特征值都有复的特征向量。设 $\lambda \in \mathbb{R}$ 是 A 的特征值，$\mathbf{x} = (x_1, \ldots, x_n)^t \in \mathbb{C}^n$ 是 λ 对应的复的特征向量，记 $\overline{\mathbf{x}} = (\overline{x_1}, \ldots, \overline{x_n})^t$，则

$$A\mathbf{x} = \lambda\mathbf{x},\ A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}} = \lambda\overline{\mathbf{x}}.$$

即 $\overline{\mathbf{x}}$ 也是 λ 对应的复特征向量。于是 $\mathbf{x} + \overline{\mathbf{x}} \in \mathbb{R}^n$ 就是 λ 对应的 A 的实特征向量，对应地，$(\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot (\mathbf{x} + \overline{\mathbf{x}})$ 就是算子 $\mathcal{A}$ 的实特征值 λ 对应的特征向量。

定理 3.3.4 的证明：借助3.3.1的 (2) 和 (3)，我们可以用数学归纳法证明该定理，留作练习。下面我们给出一个更精彩的证明。任取自伴随算子 $\mathcal{A}$ 的一个特征值 $\lambda \in \mathbb{R}$，则其极小多项式应该是下面的形式

$$\mu_{\mathcal{A}}(t) = (t - \lambda)^k \cdot f(t),\ k \in \mathbb{Z}^+.$$

下面我们证明 k 只能等于 1。由极小多项式的定义，任取 $\mathbf{v} \in V$，我们都有 $\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mathbf{0}$，即 $(\mathcal{A} - \lambda\mathcal{E})^k f(\mathcal{A})\mathbf{v} = \mathbf{0}$。所以对任意的 $\mathbf{v} \in V$，都有 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A} - \lambda\mathcal{E})^k)$。由引理3.3.1的 (1)

可知 $\mathscr{A}-\lambda\mathcal{E}$ 也是自伴随算子，故由引理3.3.1的 (5)，$\ker((\mathscr{A}-\lambda\mathcal{E})^k)=\ker((\mathscr{A}-\lambda\mathcal{E}))$，所以 $f(\mathscr{A})\mathbf{v}\in\ker((\mathscr{A}-\lambda\mathcal{E}))$，即 $(\mathscr{A}-\lambda\mathcal{E})f(\mathscr{A})\mathbf{v}=\mathbf{0}$，$\forall\mathbf{v}\in V$。因此，如果 $k>1$，则有 $(t-\lambda)\cdot f(t)$ 零化 $\mathscr{A}$ 并且 $\deg[(t-\lambda)f(t)]<\deg(\mu_{\mathscr{A}}(t))$，矛盾！因此只能是 $k=1$。

由于 $\mathscr{A}$ 的所有特征值都是实数，故不妨设 $\mathscr{A}$ 的特征多项式 $\chi_{\mathscr{A}}(t)=(t-\lambda_1)^{p_1}\cdots(t-\lambda_s)^{p_s}$，其中 $\lambda_1,\ldots,\lambda_s\in\mathbb{R}$ 两两不同，$p_1+\cdots+p_s=n$。则由上面的论述以及推论2.4.2可知 $\mu_{\mathscr{A}}(t)=(t-\lambda_1)\cdots(t-\lambda_s)$，故由定理2.5.8的 (vi) 可知 $\mathscr{A}$ 是可对角化的。

最后我们只需要找到一组标准正交基作为 $\mathscr{A}$ 的特征向量即可。由于 $\mathscr{A}$ 可对角化，故 $\mathscr{A}$ 的每个特征值的代数重数都等于几何重数，即 $V=V^{\lambda_1}\oplus V^{\lambda_2}\oplus\cdots\oplus V^{\lambda_s}$，其中 $\dim(V^{\lambda_i})=p_i$，$\forall i\in\{1,\ldots,s\}$。由定理3.1.2或定理3.2.2，我们可以取出每个特征子空间 V^{λ_i} 的一组标准正交基 $\mathbf{e}_{i1},\ldots,\mathbf{e}_{i,p_i}$，而由引理3.3.1的 (6)，$V_{\lambda_i}\perp V_{\lambda_j}$，故将所有 V_{λ_i} 的标准正交基合到一起就得到了 V 的一组标准正交基 $\mathbf{e}_{11},\ldots,\mathbf{e}_{1,p_1},\cdots\cdots,\mathbf{e}_{s1},\ldots,\mathbf{e}_{s,p_s}$。这样我们就完成了定理3.3.4的证明。 □

我们先来说明矩阵形式的定理3.3.4′ 的由来。满足 $A^*=A(\text{或}A^t=A)$ 的矩阵可以视作 $\mathbb{C}^n(\text{或}\mathbb{R}^n)$ 上的自伴随算子 $\mathscr{A}$，故由定理3.3.4，可取 $\mathbb{C}^n(\text{或}\mathbb{R}^n)$ 的一组标准正交基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 使得 $\mathscr{A}$ 可对角化。不妨设 $\mathbb{C}^n(\text{或}\mathbb{R}^n)$ 的标准基到 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 的过渡矩阵为 B，则 B 是酉矩阵 (或正交矩阵，定理3.2.6/3.1.5)，则算子 $\mathscr{A}$ 在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的矩阵为 $B^{-1}AB$，是对角阵。

利用定理3.3.1和3.3.4我们也可以顺便讨论一下 Hermitian 型和对称双线性型的对角化。

定理 3.3.5. 设 V 是 Hermitian 空间 (或欧氏空间)，$f:V\times V\to\mathbb{C}(\text{或}\mathbb{R})$ 是 V 上的 Hermitian 型 (或对称双线性型)，由定理3.3.1，存在唯一的 $\mathscr{A}\in\mathrm{End}(V)$ 使得 $f=f_{\mathscr{A}}$。则 $\mathscr{A}$ 是自伴随算子，并且存在一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得 f 有以下形式：任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{j=1}^n y_j\mathbf{e}_j\in V$，有

$$f(\mathbf{x},\mathbf{y})=\sum_{i=1}^n\lambda_i x_i\overline{y_i}=(x_1,\ldots,x_n)\cdot\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\cdot\begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}.$$

其中 $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$ 是 $\mathscr{A}$ 的所有特征值 (计重数)。

证明. 首先，$\mathscr{A}$ 的存在唯一性是由定理3.3.1保证的，而 f 的 Hermitian 性 (对称性) 则可以导出：$\forall\mathbf{x},\mathbf{y}\in V$，$<\mathscr{A}\mathbf{x},\mathbf{y}>=\overline{<\mathscr{A}\mathbf{y},\mathbf{x}>}=<\mathbf{x},\mathscr{A}\mathbf{y}>$，即 $\mathscr{A}=\mathscr{A}^*$。由定理3.3.4，存在 V 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得 $\mathscr{A}$ 在这组基下的矩阵是特征值组成的对角阵，即对 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，有 $\mathscr{A}\mathbf{x}=\sum_{i=1}^n\lambda_i x_i\mathbf{e}_i$，其中 $\lambda_1,\ldots,\lambda_n$ 是 $\mathscr{A}$ 的所有特征值 (计重数，可以相等)。于是

$$\begin{aligned}f(\mathbf{x},\mathbf{y})&=<\mathscr{A}\mathbf{x},\mathbf{y}>=<\sum_{i=1}^n\lambda_i x_i\mathbf{e}_i,\sum_{j=1}^n y_j\mathbf{e}_j>\\&=\sum_{i=1}^n\sum_{j=1}^n\lambda_i x_i\overline{y_j}<\mathbf{e}_i,\mathbf{e}_j>=\sum_{i=1}^n\lambda_i x_i\overline{y_i}.\end{aligned}$$

这样我们就完成了证明。 □

我们知道，实对称双线性型与实二次型是一一对应的。同理，我们也可以在 Hermitian 空间 V 上定义 Hermitian 二次型：任取 f 是 $V\times V\to\mathbb{C}$ 的 Hermitian 型，称 $q_f:V\to\mathbb{R}$，$\mathbf{x}\mapsto f(\mathbf{x},\mathbf{x})$ 是 V 上的 Hermitian 二次型。我们同样可以证明 Hermitian 型和 Hermitian 二次型也是一一对应的 (留作练习)。因此，定理3.3.5有下面的推论：

推论 3.3.3. 设 V 是 Hermitian 空间 (或欧氏空间)，$q : V \to \mathbb{R}$ 是 V 上的 Hermitian 二次型 (或实双线性型)，则存在唯一的 $\mathcal{A} \in \mathrm{End}(V)$ 使得 $q(\mathbf{x}) =< \mathcal{A}\mathbf{x}, \mathbf{x} >$，其中 $\mathcal{A}$ 是自伴随算子，并且存在一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 q 有以下形式：任取 $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i \in V$，有

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i |x_i|^2 = (x_1, \ldots, x_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix}.$$

其中 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 是 $\mathcal{A}$ 的所有特征值 (计重数)。

在定理3.3.5的证明中取 $\mathbf{y} = \mathbf{x}$ 即可证明，细节留作练习。

下面我们先来看一个具体计算的例子。

例 3.3.1. 设 $A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R})$，求正交矩阵 P 使得 P^tAP 是对角矩阵。

解. 显然矩阵 A 是线性算子 $\mathcal{A} : \mathbb{R}^4 \to \mathbb{R}^4$，$\mathbf{x} = (x_1, x_2, x_3, x_4)^t \mapsto A\mathbf{x}$ 在标准基下的矩阵。我们需要寻找 $\mathbb{R}^4$ 的一组标准正交基使得 $\mathcal{A}$ 在这组标准正交基下的矩阵是对角矩阵。为此，按照定理3.3.4的证明过程，我们需要先将 $\mathbb{R}^4$ 分解成 $\mathcal{A}$ 的特征子空间的直和，再求出每个特征子空间的标准正交基，最后将它们合在一起。

首先，$\chi_{\mathcal{A}}(t) = \det(tE - A) = (t-1)^3(t+3)$，则由定理3.3.4的证明过程可知，$\mathcal{A}$ 的极小多项式 $\mu_{\mathcal{A}}(t) = (t-1)(t+3)$，所以 $\mathcal{A}$ 有两个特征值 1 和-3，并且 $\dim(V^1) = 3$，$\dim(V^{-3}) = 1$。分别解线性方程组 $(E - A)\mathbf{x} = \mathbf{0}$ 和 $(-3E - A)\mathbf{x} = \mathbf{0}$ 即可得到：

$$V^1 = \mathrm{span}\,\{\mathbf{v}_1 = (1, 1, 0, 0)^t,\ \mathbf{v}_2 = (0, 0, 1, 1)^t,\ \mathbf{v}_3 = (-1, 0, 0, 1)^t\},\ V^{-3} = \mathrm{span}\,\{\mathbf{v}_4 = (1, -1, -1, 1)^t\}.$$

我们利用 Gram-Schmidt 正交化方法将上面的 V^1 和 V^{-3} 的基化成标准正交基:

对V^1：

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{e}_1 = \frac{1}{||\mathbf{u}_1||}\mathbf{u}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{<\mathbf{v}_2, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{e}_2 = \frac{1}{||\mathbf{u}_2||}\mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{<\mathbf{v}_3, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 - \frac{<\mathbf{v}_3, \mathbf{u}_2>}{<\mathbf{u}_2, \mathbf{u}_2>}\mathbf{u}_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad \mathbf{e}_3 = \frac{1}{||\mathbf{u}_3||}\mathbf{u}_3 = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix};$$

对V^{-3}：

$$\mathbf{e}_4 = \frac{1}{||\mathbf{v}_4||}\mathbf{v}_4 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

于是 $\mathbf{e}_1, \ldots, \mathbf{e}_4$ 就是使 $\mathcal{A}$ 对角化的标准正交基，由于标准基到 $\mathbf{e}_1, \ldots, \mathbf{e}_4$ 的过渡矩阵就是

$$P = (\mathbf{e}_1, \ldots, \mathbf{e}_4) = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

故显然 P 即是我们所要求的正交矩阵，并且 $P^tAP = P^{-1}AP = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -3 \end{pmatrix}$。 □

需要注意的是，使一个自伴随算子 $\mathcal{A}$ 对角化的标准正交基不是唯一的，这是因为每个 V^{λ_i} 的标准正交基都不只有一种取法。例如，在上面的例子中，取 $P' = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$ 时同样有 $P'^tAP' = P'^{-1}AP' = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -3 \end{pmatrix}$。在这里我们还可以顺手解释一下注1.6.2。如果

实二次型 q 在某组基下 $\mathbf{v}_1,\dots,\mathbf{v}_s,\dots,\mathbf{v}_r,\dots,\mathbf{v}_n$ 的矩阵为 $\begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}$，则我们可以将实向量空间 V 分解成直和 $V = V^+ \oplus V^- \oplus V^0$，其中

$$V^+ = \text{span}\,\{\mathbf{v}_1,\dots,\mathbf{v}_s\},$$
$$V^- = \text{span}\,\{\mathbf{v}_{s+1},\dots,\mathbf{v}_r\},$$
$$V^0 = \text{span}\,\{\mathbf{v}_{r+1},\dots,\mathbf{v}_n\}.$$

实际上，设 q 在另外的某组基下对应实对称矩阵 A，其全部两两不同的特征值为 $\lambda_1 > \cdots > \lambda_{k-1} > \lambda_k = 0 > \lambda_{k+1} > \cdots > \lambda_l$，则不难验证 $V^+ = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_{k-1}}$，$V^- = V^{\lambda_{k+1}} \oplus \cdots \oplus V^{\lambda_l}$，具体过程留给读者作为练习。设 $\mathcal{B},\mathcal{C}$ 分别是 V^+、V^- 上的正交算子，$\mathcal{D}$ 是 V^0 上的任意线性算子，它们在 $\mathbf{v}_1,\dots,\mathbf{v}_n$ 下的矩阵分别是 B,C,D。则我们把 $\mathcal{B},\mathcal{C},\mathcal{D}$ 分别作用到 V^+,V^-,V^0 的基上，就可以得到 V 的另一组基 $\mathbf{u}_1,\dots,\mathbf{u}_s,\dots,\mathbf{u}_r,\dots,\mathbf{u}_n$。于是 q 在 $\mathbf{u}_1,\dots,\mathbf{u}_n$ 下的矩阵为 $\begin{pmatrix} B^tE_sB & & \\ & -C^tE_{r-s}C & \\ & & O \end{pmatrix} = \begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}$。这就是注1.6.2的构造方法。

在本节的最后，我们介绍一类特殊的自伴随算子：(半) 正定算子。

定义 3.3.3. 设 $(V,<\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \text{End}(V)$ 是自伴随算子，如果 $\forall \mathbf{x} \in V$，二次型 (或 Hermitian 二次型)$q(\mathbf{x}) =< \mathcal{A}\mathbf{x},\mathbf{x} >\geqslant 0$，则称 $\mathcal{A}$ 是半正定算子；如果半正定算子 $\mathcal{A}$ 还是非退化 (可逆) 算子，则称 $\mathcal{A}$ 为正定算子。

由推论3.3.3可知，存在一组标准正交基使得二次型 (或 Hermitian 二次型) $q(\mathbf{x}) =< \mathcal{A}\mathbf{x},\mathbf{x} >$ 有标准型

$$q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i |x_i|^2,\ \forall \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \in V.$$

其中 λ_i，$i = 1,\dots,n$ 是 $\mathcal{A}$ 的所有特征值 (计重数)。因此我们立刻有：

命题 3.3.1. 设 $\mathcal{A}$ 是内积空间 $(V,<\cdot,\cdot>)$ 上的自伴随算子，则 $\mathcal{A}$ 是半正定算子 $\iff$ 其所有特征值都非负；$\mathcal{A}$ 是正定算子 $\iff$ 其所有特征值都大于 0。

我们知道，非负实数可以开平方，现在我们希望将开平方的操作推广到半正定算子 (半正定矩阵) 上。为此，不妨设半正定算子 $\mathcal{A}$ 在某组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵为

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},\quad \lambda_i \geqslant 0,\ \forall i \in \{1,\dots,n\}.$$

则线性算子 $\mathcal{B} : \mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i \mapsto \sum_{i=1}^{n} \sqrt{\lambda_i} x_i \mathbf{e}_i$(即 $\mathcal{B}$ 在此基底下对应的矩阵为 $B = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}$) 也是半正定的自伴随算子，并且 $\mathcal{B}^2 = \mathcal{B}^*\mathcal{B} = \mathcal{A}$。特别地，如果 $\mathcal{A}$ 正定，则 $\mathcal{B}$ 也正定。我们把算子 $\mathcal{B}$ 称为算子 $\mathcal{A}$ 的**平方根**，记作 $\mathcal{B} = \sqrt{\mathcal{A}}$。实际上半正定算子的平方根是唯一的，我们会在稍后的 3.5 节中证明它，或者先留作练习[1]。

---

[1] 也可参考《高等代数学习指南》，蓝以中，北京大学出版社的第五章 §2 的例 2.9-例 2.11。

于是我们有

定理 3.3.6. 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$，则以下条件等价：

(1) $\mathcal{A}$ 是半正定算子（正定算子）；

(2) 存在 $\mathcal{B} \in \mathrm{End}(V)$ 是半正定算子（正定算子）使得 $\mathcal{A} = \mathcal{B}^2$ 且 $\mathcal{B}^* = \mathcal{B}$；

(3) 存在线性算子（可逆算子）C 使得 $\mathcal{A} = C^*C$ 或 $\mathcal{A} = CC^*$。

证明. (1)$\Longrightarrow$(2) 即上面的论述。(2)$\Longrightarrow$(3) 是平凡的。下面我们证明 (3)$\Longrightarrow$(1)。我们只证明 $\mathcal{A} = C^*C$ 的情形，另一个同理。首先 $\mathcal{A}^* = (C^*C)^* = C^*C = \mathcal{A}$，即 $\mathcal{A}$ 是自伴随算子；其次任取 $\mathbf{x} \in V$，有 $< C^*C\mathbf{x}, \mathbf{x} >=< C\mathbf{x}, C\mathbf{x} > \ \geqslant 0$，即 C 是半正定算子。特别地，如果 C 是可逆算子，则 $\mathcal{A}$ 显然也是可逆的，故 $\mathcal{A}$ 是正定算子。 □

3.4　内积空间上的线性算子 II：保距算子

现在我们开始考虑内积空间 V 上另一类特殊的算子：保距算子。之所以我们要把满足 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>,\ \forall\mathbf{x}\in V$ 的算子 $\mathcal{A}$ 称为保距算子，是因为 $\mathcal{A}$ 满足：$\forall\mathbf{x},\mathbf{y}\in V,\ \mathrm{d}(\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y})=\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|=\|\mathbf{x}-\mathbf{y}\|=\mathrm{d}(\mathbf{x},\mathbf{y})$，即算子 $\mathcal{A}$ 保持距离。下面我们先证明保距算子在有限维的内积空间上就是酉算子 (或正交算子)。

定理 3.4.1. 设 $(V,<\cdot,\cdot>)$ 是内积空间，则 $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x}\in V,\ \|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|\iff\forall\mathbf{x},\mathbf{y}\in V,\ <\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$。

证明. 由内积空间中范数的定义，$\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x}\in V,\ \|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|$ 是显然的。下面我们只需证明 $\mathcal{A}$ 是保距算子 $\iff\forall\mathbf{x},\mathbf{y}\in V,\ <\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$。

($\Longleftarrow$) 取 $\mathbf{y}=\mathbf{x}$ 即得 $\mathcal{A}$ 是保距算子。

($\Longrightarrow$) 如果 V 是欧氏空间，$\mathcal{A}$ 是保距算子，则 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>,\ \forall\mathbf{x}\in V$，于是由极化恒等式 (命题3.1.3的 (3))，我们有：任取 $\mathbf{x},\mathbf{y}\in V$，有

$$\begin{aligned}<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}> &= \frac{1}{4}(\|\mathcal{A}(\mathbf{x}+\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2)\\ &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)\\ &=<\mathbf{x},\mathbf{y}>.\end{aligned}$$

如果 V 是 Hermitian 空间，$\mathcal{A}$ 是保距算子，则 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>,\ \forall\mathbf{x}\in V$，于是由极化恒等式 (命题3.2.4的 (3))，我们有：任取 $\mathbf{x},\mathbf{y}\in V$，有

$$\begin{aligned}<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}> &= \frac{1}{4}(\|\mathcal{A}(\mathbf{x}+\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2)+\frac{i}{4}(\|\mathcal{A}(\mathbf{x}+i\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-i\mathbf{y})\|^2)\\ &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)+\frac{i}{4}(\|\mathbf{x}+i\mathbf{y}\|^2-\|\mathbf{x}-i\mathbf{y}\|^2)\\ &=<\mathbf{x},\mathbf{y}>.\end{aligned}$$

即 $\mathcal{A}$ 是正交算子或酉算子。这样我们就完成了证明。□

推论 3.4.1. 设 $(V,<\cdot,\cdot>)$ 是内积空间，则 $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\mathcal{A}^*\mathcal{A}=\mathcal{E}$。

证明. $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x},\mathbf{y}\in V,\ <\mathbf{x},\mathbf{y}>=<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathcal{A}^*\mathcal{A}\mathbf{y}>\iff\forall\mathbf{x},\mathbf{y}\in V,\ <\mathbf{x},(\mathcal{A}^*\mathcal{A}-\mathcal{E})\mathbf{y}>=0\iff\forall\mathbf{y}\in V,\ (\mathcal{A}^*\mathcal{A}-\mathcal{E})\mathbf{y}=\mathbf{0}\iff\mathcal{A}^*\mathcal{A}=\mathcal{E}$。□

于是任取 V 的一组标准正交基，保距算子对应的矩阵就是酉矩阵 (或正交矩阵)。

注 3.4.1. 在有限维空间中，$\mathcal{A}^*\mathcal{A}=\mathcal{E}$ 显然能够保证 $\mathcal{A}$ 可逆。但在无穷维空间中，保距算子不一定是酉算子，这是因为 $\mathcal{A}^*\mathcal{A}=\mathcal{E}$ 并不能保证 $\mathcal{A}$ 是可逆的 (这只能说明 $\mathcal{A}$ 有左逆，但右逆可以不存在，当然如果右逆存在的话必有左逆等于右逆)，而酉算子必须是可逆算子。反例可参考《A Course in Functional Analysis》，John B.Conway, GTM96 的 Chapter I, §4 的 Example 5.3。

下面我们来考虑保距算子的对角化问题。保距算子的对角化可以作为下一节的结论的特例，但为了循序渐进，我们在这里给出一个比较简单的归纳证明 (用它读者也可以类比地写出定理3.3.4和下一节正规算子可对角化的证明)。

定理 3.4.2. 设 $(V, <\cdot, \cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}$ 是 V 上的保距算子 (即酉算子)，则存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得：

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

其中 $\lambda_i \in \mathbb{C}$ 且 $|\lambda_i| = 1$，$\forall i \in \{1, \ldots, n\}$(各个 λ_i 之间可以相等)。

为了证明上述定理，我们首先需要以下引理：

引理 3.4.1. 设 $\mathcal{A}$ 是 Hermitian 空间 $(V, <\cdot, \cdot>)$ 上的保距算子，则

(1) 若 $\lambda \in \mathbb{C}$ 是 $\mathcal{A}$ 的特征值，则 $|\lambda| = 1$；

(2) 若 W 是 $\mathcal{A}-$ 子空间，则 $W^{\perp}$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^{\perp}}$ 都是保距算子。

证明. (1) 设 $\mathbf{v} \in V$ 是 λ 对应的一个非零特征向量，则

$$< \mathbf{v}, \mathbf{v} >=< \mathcal{A}\mathbf{v}, \mathcal{A}\mathbf{v} >=< \lambda\mathbf{v}, \lambda\mathbf{v} >= \lambda \cdot \overline{\lambda} < \mathbf{v}, \mathbf{v} > .$$

因为 $< \mathbf{v}, \mathbf{v} >\neq 0$，故只能是 $\lambda\overline{\lambda} = |\lambda|^2 = 1$，即 $|\lambda| = 1$。

(2) 由于 W 是 $\mathcal{A}-$ 子空间，而 $\mathcal{A}$ 是可逆算子，故 $\forall \mathbf{u} \in W$, $\exists \mathbf{w} \in W$ 使得 $\mathcal{A}\mathbf{w} = \mathbf{u}$。下面我们验证 $W^{\perp}$ 也是 $\mathcal{A}-$ 子空间。任取 $\mathbf{v} \in W^{\perp}$(即 $\mathbf{v} \perp W$)，我们只要证明 $\mathcal{A}\mathbf{v} \in W^{\perp}$，即 $\mathcal{A}\mathbf{v} \perp W$ 即可完成证明。任取 $\mathbf{u} \in W$，则 $< \mathcal{A}\mathbf{v}, \mathbf{u} >=< \mathcal{A}\mathbf{v}, \mathcal{A}\mathbf{w} >=< \mathbf{v}, \mathbf{w} >= 0$，此即 $\mathcal{A}\mathbf{v} \perp W$。$\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^{\perp}}$ 是保距算子是显然的。

□

定理 3.4.2 **的证明**: 对 V 的维数 n 用数学归纳法。$n = 1$ 时定理显然成立。下设定理对所有 $n - 1$ 维 Hermitian 空间都成立，我们来证明 $\dim(V) = n$ 的情形。

由于特征多项式 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{C}$ 上至少有 1 个根，故我们总可以取 $\mathcal{A}$ 的某个特征值 λ_1 和它对应的一个特征向量 $\mathbf{e}_1$。由引理3.4.1，我们有 $|\lambda_1| = 1$，并且 $\operatorname{span}\{\mathbf{e}_1\}$ 和 $\operatorname{span}\{\mathbf{e}_1\}^{\perp}$ 都是 $\mathcal{A}-$ 子空间。由于 $\mathcal{A}$ 限制到 $\operatorname{span}\{\mathbf{e}_1\}^{\perp}$ 上也是保距算子，并且 $\dim(\operatorname{span}\{\mathbf{e}_1\}^{\perp}) = n - 1$，故由归纳假设，存在 $\operatorname{span}\{\mathbf{e}_1\}^{\perp}$ 的一组标准正交基 $\mathbf{e}_2, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}$ 限制到 $\operatorname{span}\{\mathbf{e}_1\}^{\perp}$ 上对应的矩阵是对角阵 $\begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}_{(n-1)\times(n-1)}$，并且对角元的模长都是 1。因此，结合上述论证与推论2.3.1，可知 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵是 $\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}_{n\times n}$，并且所有对角元的模长都是 1。这样我们就完成了归纳证明。

推论 3.4.2. 设 $A \in M_n(\mathbb{C})$ 是酉矩阵，则存在 $B \in U(n)$ 使得

$$B^*AB = B^{-1}AB = \begin{pmatrix} e^{i\varphi_1} & & \\ & \ddots & \\ & & e^{i\varphi_n} \end{pmatrix}$$

其中 $\varphi_1, \ldots, \varphi_n \in \mathbb{R}$(可以相等)，$i$ 是虚数单位。

将矩阵视作 $\mathbb{C}^n$ 上的线性算子即可证明，注意得到最后的对角矩阵形式时利用了复数的指数形式 (讲义上册 5.1 节)。

现在我们来考虑欧氏空间上的正交算子的情形。此时情况要复杂一些，因为正交算子的特征值可能是虚数，这样我们就不能将正交算子在实数域上对角化。但退而求其次，我们可以将正交算子在实数域上分块对角化。

引理 3.4.2. 设 $\mathcal{A}$ 是 Hermitian 空间 $(V, <\cdot,\cdot>)$ 上的保距算子，则

(1) 若 $\lambda \in \mathbb{R}$ 是 $\mathcal{A}$ 的实特征值，则 $\lambda = \pm 1$；

(2) 若 W 是 $\mathcal{A}-$ 子空间，则 $W^{\perp}$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^{\perp}}$ 都是保距算子。

证明与引理3.4.1相同，留作练习。

引理 3.4.3. $\mathbb{R}$ 上的一阶正交矩阵只有 (± 1)；在 $\mathbb{R}$ 上不能对角化的二阶正交矩阵一定是 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$，其中 $\theta \in \mathbb{R} \setminus \{2k\pi \mid k \in \mathbb{Z}\}$ 的形式。

证明. 一阶的情况是显然的。下面考虑二阶的情形。任取二阶实正交矩阵 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，首先 $A \neq \pm E$；接下来由于 $A^t A = E$，即 $\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$，于是

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

不妨取 $a = \cos\theta$, $c = \sin\theta$, $b = \sin\varphi$, $d = \cos\varphi$，则 $ab + cd = 0 \Rightarrow \sin(\theta + \varphi) = 0$, $\varphi = k\pi - \theta, k \in \mathbb{Z}$，所以 $b = -c, d = a$(k是偶数) 或者 $b = c, d = -a$(k是奇数)，然而后者会导致 $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$, $\chi_A(t) = t^2 - 1$, 此时 A 有两个实特征值 ± 1, A 可以在实数域上对角化, 与我们的要求不符。因此, 我们只能取 k 是偶数的情形, 即 $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $\theta \in \mathbb{R} \setminus \{2k\pi \mid k \in \mathbb{Z}\}$。 □

引理 3.4.4. 设 V 是 n 维实向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 V 一定有一个 1 维或 2 维的 $\mathcal{A}-$ 子空间。

证明. 设 $\mathcal{A}$ 的特征多项式是 $\chi_{\mathcal{A}}(t)$，则 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{R}[t]$ 中可以分解成一些 1 次因子和 2 次的因子的乘积 (讲义上册推论 6.5.2)，则由定理2.4.4的证明方法，归纳即可得到 $\mathcal{A}$ 一定有一个 1 维或 2 维的不变子空间。细节留作思考。 □

定理 3.4.3. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的保距算子 (即正交算子)，则存在 V

的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得:

$$(\mathcal{A}\mathbf{e}_1,\ldots,\mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot\begin{pmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_r & & \\ & & & E_s & \\ & & & & -E_t \end{pmatrix}.$$

其中 $A_i = \begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix} \in O(2)$, $\forall i \in \{1,\ldots,r\}$ (各个 φ_i 之间可以相等), E_s 和 E_t 分别是 $s\times s$ 和 $t\times t$ 阶的单位矩阵, 并且 $2r+s+t=n$。

证明. 首先, V 一定有一个 1 维或 2 维的 $\mathcal{A}-$ 子空间 U, 则 $U^\perp$ 也是 $\mathcal{A}-$ 子空间, 并且 $\mathcal{A}\big|_{U^\perp}$ 上也是正交算子。因此, 类似于定理3.4.2的证明方法, 我们可以归纳地证明: V 是一些两两正交的 1 维或 2 维的 $\mathcal{A}-$ 子空间的直和 (细节留作练习)。不失一般性, 我们可以设:

$$V = V_1\oplus\cdots\oplus V_r\oplus V_{r+1}\oplus\cdots\oplus V_{r+s}\oplus V_{r+s+1}\oplus\cdots\oplus V_{r+s+t}. \tag{3.4.1}$$

其中 $V_1,\ldots,V_r$ 是 2 维的 $\mathcal{A}-$ 子空间, $V_{r+1},\ldots,V_{r+s}$ 是 $\mathcal{A}$ 的 1 维特征子空间 (对应的特征值是 1), $V_{r+s+1},\ldots,V_{r+s+t}$ 是 $\mathcal{A}$ 的 1 维特征子空间 (对应的特征值是-1)。因此, 由引理3.4.2, $\mathcal{A}$ 限制到每个 V_i 上都是正交算子, 即每个 V_i 中都存在一组标准正交基使得 $\mathcal{A}\big|_{V_i}$ 对应的矩阵是 1 阶或 2 阶的正交矩阵。故由引理3.4.3及推论2.3.1, 将每个 V_i 的标准正交基合在一起构成 V 的标准正交基, 则 $\mathcal{A}$ 在此标准正交基下的矩阵就是我们所需要的形式。$2r+s+t=n$ 可以由直和分解 (3.4.1) 直接得到。 □

下面我们介绍一下内积空间上线性算子的**极化分解** (Polar decomposition)。我们知道任何非零的复数 z 都能写成指数形式 $z=re^{i\theta}$, 其中 $r\in\mathbb{R}^+$, $\theta\in\mathbb{R}$。实际上, 我们可以对线性算子作类似的分解。

定理 3.4.4. 设 $(V,<\cdot,\cdot>)$ 是 n 维内积空间, $\mathcal{A}\in\mathrm{End}(V)$, 则

(1) 存在半正定算子 $\mathcal{P}'$ 和保距算子 $\mathcal{Q}'$ 使得 $\mathcal{A}=\mathcal{Q}'\mathcal{P}'$。更进一步, 如果 $\mathcal{A}$ 可逆, 则 $\mathcal{P}'$ 是正定的, $\mathcal{P}'$ 和 $\mathcal{Q}'$ 也都是唯一的。

(2) 存在半正定算子 $\mathcal{P}$ 和保距算子 $\mathcal{Q}$ 使得 $\mathcal{A}=\mathcal{P}\mathcal{Q}$。更进一步, 如果 $\mathcal{A}$ 可逆, 则 $\mathcal{P}$ 是正定的, 且 $\mathcal{P}$ 和 $\mathcal{Q}$ 都是唯一的。

证明. (1) 由定理3.3.6, $\mathcal{A}^*\mathcal{A}$ 是正定算子, 因此 $\mathcal{A}\mathcal{A}^*$ 存在 (唯一) 的平方根 $\mathcal{P}'=\sqrt{\mathcal{A}^*\mathcal{A}}$。在 $\mathrm{im}(\mathcal{P}')$ 上定义线性算子 $\mathcal{Q}':\mathrm{im}(\mathcal{P}')\to V$ 如下:

$$\forall \mathcal{P}'\mathbf{x}\in\mathrm{im}(\mathcal{P}'),\ \mathcal{Q}'(\mathcal{P}'\mathbf{x}) = \mathcal{A}\mathbf{x}.$$

我们需要验证这样定义的 $\mathcal{Q}'$ 是良定义的。事实上, 如果 $\mathcal{P}'\mathbf{x}=\mathcal{P}'\mathbf{y}$, 即 $\mathbf{x}-\mathbf{y}\in\ker(\mathcal{P}')$, 那么

$$0 = \|\mathcal{P}'(\mathbf{x}-\mathbf{y})\|^2 =< \mathbf{x}-\mathbf{y}, (\mathcal{P}')^2(\mathbf{x}-\mathbf{y}) >=< \mathbf{x}-\mathbf{y}, (\mathcal{A}^*\mathcal{A})(\mathbf{x}-\mathbf{y}) >= \|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2,$$

即 $\mathcal{A}(\mathbf{x}-\mathbf{y})=0$, $\mathcal{A}\mathbf{x}=\mathcal{A}\mathbf{y}$ 成立。这说明 $\mathcal{Q}'$ 确实是良定义的。

下面验证 $\mathcal{Q}'$ 是一个保距映射。任取 $\mathcal{P}'\mathbf{x} \in \mathrm{im}(\mathcal{P}')$，我们有

$$\|\mathcal{Q}'(\mathcal{P}'\mathbf{x})\|^2 = \|\mathcal{A}\mathbf{x}\|^2 =< \mathbf{x}, \mathcal{A}^*\mathcal{A}\mathbf{x} >=< \mathcal{P}\mathbf{x}, \mathcal{P}\mathbf{x} >= \|\mathcal{P}\mathbf{x}\|^2,$$

这说明 $\mathcal{Q}'$ 确实是一个保距映射。

接下来将 $\mathcal{Q}'$ 扩张到 $V \to V$ 上。取 $\mathrm{im}(\mathcal{P}')$ 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_k$，由于 $\mathcal{Q}'$ 保持向量的范数不变，因此 $\mathcal{Q}'$ 是单射，于是 $\mathcal{Q}'\mathbf{e}_1, \ldots, \mathcal{Q}'\mathbf{e}_k$ 也是 V 中的一组标准正交的向量。将 $\mathbf{e}_1, \ldots, \mathbf{e}_k$ 扩充成 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，$\mathcal{Q}'\mathbf{e}_1, \ldots, \mathcal{Q}'\mathbf{e}_k$ 扩充成 V 的另一组标准正交基 $\mathcal{Q}'\mathbf{e}_1, \ldots, \mathcal{Q}'\mathbf{e}_k, \mathbf{e}'_{k+1}, \ldots, \mathbf{e}'_n$ (推论3.2.2)，并定义

$$\begin{aligned} \mathcal{Q}' : V &\to V \\ \mathbf{e}_i &\mapsto \mathcal{Q}'\mathbf{e}_i, \ i \in \{1, \ldots, k\}; \\ \mathbf{e}_i &\mapsto \mathbf{e}'_i, \ i \in \{k+1, \ldots, n\}. \end{aligned}$$

则 $\mathcal{Q}'$ 是 V 上的保距算子 (定理3.2.6) 且按定义 $\mathcal{Q}'\mathcal{P}' = \mathcal{A}$。

最后，如果 $\mathcal{A}$ 可逆，我们来证明极化分解的唯一性。此时，注意到 $\mathcal{P}', \mathcal{Q}'$ 满足：

$$\mathcal{P}'^* = \mathcal{P}', \ \mathcal{Q}'^*\mathcal{Q}' = \mathcal{E},$$

因此我们有

$$\mathcal{A}^*\mathcal{A} = \mathcal{P}'^*\mathcal{Q}'^*\mathcal{Q}\mathcal{P} = \mathcal{P}'^2$$

即 $\mathcal{P}'$ 一定是 $\mathcal{A}^*\mathcal{A}$ 的平方根，因此它是唯一的，并且此时 $\mathcal{A}^*\mathcal{A}$ 正定 $\Longrightarrow \mathcal{P}$ 也正定，于是 $\mathcal{Q}' = \mathcal{A}\mathcal{P}^{-1}$ 也是唯一的。

(2) 由 (1)，我们知道 $\mathcal{A}^*$ 有极化分解

$$\mathcal{A}^* = \mathcal{Q}''\mathcal{P}'',$$

则 $\mathcal{A} = \mathcal{P}''^*\mathcal{Q}''^*$。令

$$\mathcal{P} = \mathcal{P}''^*, \mathcal{Q} = \mathcal{Q}''^*,$$

则 $\mathcal{P}$ 是半正定算子，$\mathcal{Q}$ 是保距算子。$\mathcal{A}$ 可逆时分解的唯一性同上，留作练习。 □

事实上，极化分解可以被推广到 Hilbert 空间上的有界线性算子。此时，保距算子会被弱化成部分等距算子。感兴趣的同学可以参考 A Course in Functional Analysis, John B.Conway, GTM96 的 Chapter VIII, 3.10。

在本节的最后，我们来介绍实或复矩阵的奇异值分解 (singular value decomposition)。

定理 3.4.5. 设 V 是 n 维内积空间，$\mathcal{A} \in \mathrm{End}(V)$。则存在 V 的两组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\mathbf{f}_1, \ldots \mathbf{f}_n$ 使得映射 $\mathcal{A} : V \to V$ 在定义域的基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和值域的基 $\mathbf{f}_1, \ldots, \mathbf{f}_n$ 下的矩阵为

$$\mathcal{A}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = (\mathbf{f}_1, \ldots, \mathbf{f}_n)\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}.$$

其中 $s_1, \ldots, s_n$ 是非负实数，被称为算子 $\mathcal{A}$ 的奇异值 (singular value)。对应到矩阵形式，设

A 是 n 阶复矩阵 (或实矩阵)，则存在酉矩阵 (或正交矩阵)U, V 使得

$$A = U\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}V.$$

当 A 可逆时，A 的奇异值分解在奇异值的次序固定时是唯一的，并且奇异值全为正数。

证明. 取 $\mathcal{A}$ 的极化分解 $\mathcal{A} = \mathcal{Q}\mathcal{P}$，其中 $\mathcal{Q}$ 是保距算子，P 是 $\mathcal{A}^*\mathcal{A}$ 的平方根。由于 $\mathcal{P}$ 是半正定算子，因此由定理3.3.4，存在一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得

$$\mathcal{P}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}.$$

以 $\mathcal{Q}$ 作用到上式两边得

$$\mathcal{A}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \mathcal{Q}\mathcal{P}(\mathbf{e}_1, \ldots, \mathbf{e}_n) = (\mathcal{Q}\mathbf{e}_1, \ldots, \mathcal{Q}\mathbf{e}_n)\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}.$$

由于 $\mathcal{Q}$ 是保距算子，因此 $\mathcal{Q}\mathbf{e}_1, \ldots, \mathcal{Q}\mathbf{e}_n$ 也是一组标准正交基。令 $\mathbf{f}_1 = \mathcal{Q}\mathbf{e}_1, \ldots, \mathbf{f}_n = \mathcal{Q}\mathbf{e}_n$ 就得到了算子的奇异值分解。矩阵的奇异值分解则可由算子的奇异值分解和定理2.1.4得到，留作练习。当 A 可逆时，由极化分解的唯一性可以得到奇异值分解的唯一性，并且此时 $\mathcal{P}$ 正定，而奇异值是 $\mathcal{P}$ 的特征值，因此都是正数。这样我们就完成了证明。 □

例 3.4.1. 矩阵

$$A = \begin{pmatrix} \frac{11\sqrt{2}}{2} & -\frac{5\sqrt{2}}{2} \\ -\sqrt{2} & 5\sqrt{2} \end{pmatrix} \in M_n(\mathbb{C}).$$

(1) 求 A 的极化分解 (只需写出定理3.3.4(1))；

(2) 求 A 的奇异值分解。

解. (1) 容易计算出 A^*A 的特征值为 25 和 100，其对应的单位特征向量分别为 $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})^t$ 和 $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})^t$，因此

$$P' = \sqrt{A^*A} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{15}{2} & -\frac{5}{2} \\ -\frac{5}{2} & \frac{15}{2} \end{pmatrix},$$

而

$$Q' = A \cdot P'^{-1} = \begin{pmatrix} \frac{7\sqrt{2}}{10} & -\frac{\sqrt{2}}{10} \\ \frac{\sqrt{2}}{10} & \frac{7\sqrt{2}}{10} \end{pmatrix},$$

(2) 令

$$U = Q' \cdot \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix},\ V = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

则 $A = Q'P' = Q' \cdot V \cdot \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \cdot V^* = U \cdot \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \cdot V^*$ 即为奇异值分解。 □

3.5 内积空间上的线性算子 III：正规算子

现在我们可以来探讨内积空间上的线性算子何时可以用标准正交基对角化了。

定义 3.5.1. 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$，如果 $\mathcal{A}$ 还满足 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$，则我们称 $\mathcal{A}$ 是正规算子 (normal operator)。对应到矩阵层面，如果矩阵 A 满足 $A^*A = AA^*$(实数域情形则是 $A^tA = AA^t$)，则我们称 A 是正规矩阵 (normal matrix)。显然 $\mathcal{A}$ 是正规算子 $\Longleftrightarrow$ $\mathcal{A}$ 在标准正交基下的矩阵是正规矩阵。

容易看出欧氏空间或 Hermitian 空间上的自伴随算子和保距算子都是正规算子。但正规算子显然不只有这两种类型，例如 $\mathbb{C}^2$ 上的算子 $\mathbf{x}(x_1, x_2)^t \mapsto \begin{pmatrix} 2 & 0 \\ 0 & i \end{pmatrix}$ 就是正规算子，但显然它不是自伴随算子或保距算子。

本节中我们的主要任务是研究正规算子的对角化问题，即谱定理。我们主要在复数域上讨论问题，实数域的情形我们会给出对应的结论和参考文献。

定理 3.5.1. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $\mathcal{A}$ 是正规算子 $\Longleftrightarrow$ 存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad \lambda_1, \ldots, \lambda_n \in \mathbb{C}(\lambda_1, \ldots, \lambda_n \text{中可以有相等的元素}).$$

我们同样可以用数学归纳法证明以上定理。但为了更清楚地理解正规算子的结构，我们将仿照定理3.3.4的证明方法给出上面定理的证明。

引理 3.5.1. 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $\ker(\mathcal{A}^*\mathcal{A}) = \ker(\mathcal{A})$。

证明. $\ker(\mathcal{A}^*\mathcal{A}) \supset \ker(\mathcal{A})$：任取 $\mathbf{x} \in \ker(\mathcal{A})$，则 $\mathcal{A}\mathbf{x} = \mathbf{0}$，于是 $\mathcal{A}^*\mathcal{A}\mathbf{x} = \mathbf{0}$，即 $\mathbf{x} \in \ker(\mathcal{A}^*\mathcal{A})$。$\ker(\mathcal{A}^*\mathcal{A}) \subset \ker(\mathcal{A})$：任取 $\mathbf{x} \in \ker(\mathcal{A}^*\mathcal{A})$，即 $\mathcal{A}^*\mathcal{A}\mathbf{x} = \mathbf{0}$，于是 $<\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{x}> = <\mathcal{A}^*\mathcal{A}\mathbf{x}, \mathbf{x}> = \mathbf{0}$，因此 $\mathcal{A}\mathbf{x} = \mathbf{0}$，即 $\mathbf{x} \in \ker(\mathcal{A})$。这样我们就完成了证明。 □

引理 3.5.2. 设 $(V, <\cdot,\cdot>)$ 是 Hermitian 空间，$\mathcal{A} \in \mathrm{End}(V)$ 是正规算子，则

(1) $\mathcal{A}^*$ 也是正规算子。$\forall p(t) \in \mathbb{C}[t]$，线性算子 $p(\mathcal{A})$ 也是正规算子。特别地，对任意的 $\lambda \in \mathbb{C}$，有 $\mathcal{A} - \lambda\mathcal{E}$ 是正规算子。

(2) $\forall \mathbf{x}, \mathbf{y} \in V$，$<\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y}> = <\mathcal{A}^*\mathbf{x}, \mathcal{A}^*\mathbf{y}>$。因此，$\forall \mathbf{x} \in V$，$\|\mathcal{A}\mathbf{x}\| = \|\mathcal{A}^*\mathbf{x}\|$，并且 $\ker(\mathcal{A}) = \ker(\mathcal{A}^*)$。

(3) $\mathcal{A}\mathbf{x} = \lambda\mathbf{x} \Longleftrightarrow \mathcal{A}^*\mathbf{x} = \overline{\lambda}\mathbf{x}$，其中 $\lambda \in \mathbb{C}$ 是 $\mathcal{A}$ 的特征值，$\mathbf{x} \in V$ 是 λ 对应的非零特征向量。

(4) $\forall k \in \mathbb{Z}^+$，$\ker(\mathcal{A}^k) = \ker(\mathcal{A})$。

(5) 如果 λ, μ 是 $\mathcal{A}$ 的两个不同的特征值，$\mathbf{x}$ 和 $\mathbf{y}$ 分别是 λ 和 μ 对应的特征向量，则 $\mathbf{x} \perp \mathbf{y}$。

证明. (1) $\mathcal{A}^*$ 的正规性是显然的。由于 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$，则容易验证 $\forall i, j \in \mathbb{Z}^+$，$(\mathcal{A}^*)^i\mathcal{A}^j = \mathcal{A}^j(\mathcal{A}^*)^i$。不妨设 $p(t) = a_0 + a_1t + \cdots + a_nt^n \in \mathbb{C}[t]$，则有

$$p(\mathcal{A}) = a_0\mathcal{E} + a_1\mathcal{A}^* + \cdots + a_n(\mathcal{A}^*)^n,\ [p(\mathcal{A})]^* = \overline{a_0}\mathcal{E} + \overline{a_1}\mathcal{A}^* + \cdots + \overline{a_n}(\mathcal{A}^*)^n,$$

于是

$$\begin{aligned}[p(\mathcal{A})]^* \cdot p(\mathcal{A}) &= \sum_{i=0}^{n}\sum_{j=0}^{n} \overline{a_i} a_j (\mathcal{A}^*)^i \mathcal{A}^j \\ &= \sum_{i=0}^{n}\sum_{j=0}^{n} a_j \overline{a_i} \mathcal{A}^j (\mathcal{A}^*)^i \\ &= p(\mathcal{A}) \cdot [p(\mathcal{A})]^*\end{aligned}$$

即 $p(\mathcal{A})$ 也是正规算子。取 $p(t) = t - \lambda$ 即可看出特例是显然的。

(2) 利用正规性可得

$$\begin{aligned}< \mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y} > &=< \mathbf{x}, \mathcal{A}^*\mathcal{A}\mathbf{y} > \\ &=< \mathbf{x}, \mathcal{A}\mathcal{A}^*\mathbf{y} > \\ &=< \mathcal{A}^*\mathbf{x}, \mathcal{A}^*\mathbf{y} >\end{aligned}$$

在上式中取 $\mathbf{y} = \mathbf{x}$ 即得 $\|\mathcal{A}\mathbf{x}\| = \|\mathcal{A}^*\mathbf{x}\|$。所以

$$\mathcal{A}\mathbf{x} = \mathbf{0} \iff \|\mathcal{A}\mathbf{x}\| = 0 \iff \|\mathcal{A}^*\mathbf{x}\| = 0 \iff \mathcal{A}^*\mathbf{x} = \mathbf{0},$$

即 $\ker(\mathcal{A}) = \ker(\mathcal{A}^*)$。

(3) 由 (1) 可知 $\mathcal{A} - \lambda\mathcal{E}$ 是正规算子，并且 $(\mathcal{A} - \lambda\mathcal{E})^* = \mathcal{A}^* - \overline{\lambda}\mathcal{E}$，因此 $\ker(\mathcal{A} - \lambda\mathcal{E}) = \ker(\mathcal{A}^* - \overline{\lambda}\mathcal{E})$。于是任取 $\mathbf{x} \in \ker(\mathcal{A} - \lambda\mathcal{E})$ 即得结论。

(4) 由于 $(\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}$，即 $\mathcal{A}^*\mathcal{A}$ 是自伴随算子，故由引理3.3.1的 (5)，我们有

$$\forall k \in \mathbb{Z}^+, \ \ker[(\mathcal{A}^*\mathcal{A})^k] = \ker(\mathcal{A}^*\mathcal{A}).$$

而由正规性容易验证

$$(\mathcal{A}^*\mathcal{A})^k = (\mathcal{A}^*)^k\mathcal{A}^k = (\mathcal{A}^k)^*\mathcal{A}^k,$$

因此由引理3.5.1可得

$$\ker(\mathcal{A}^k) = \ker[(\mathcal{A}^k)^*\mathcal{A}^k] = \ker[(\mathcal{A}^*\mathcal{A})^k] = \ker(\mathcal{A}^*\mathcal{A}) = \ker(\mathcal{A}).$$

即得结论。

(5) 由于 $\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$, $\mathcal{A}\mathbf{y} = \mu\mathbf{y}$，故首先由 (3) 可知 $\mathcal{A}^*\mathbf{y} = \overline{\mu}\mathbf{y}$。于是利用特征向量的定义和内积的半双线性性质，我们有

$$\begin{aligned}\lambda < \mathbf{x}, \mathbf{y} > &=< \lambda\mathbf{x}, \mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} > \\ &=< \mathbf{x}, \mathcal{A}^*\mathbf{y} >=< \mathbf{x}, \overline{\mu}\mathbf{y} > \\ &= \mu < \mathbf{x}, \mathbf{y} >\end{aligned}$$

而 $\lambda \neq \mu$，故只能是 $< \mathbf{x}, \mathbf{y} >= 0$，即 $\mathbf{x} \perp \mathbf{y}$。

□

定理3.5.1的证明：($\Longleftarrow$) 如果 $\mathcal{A}$ 在某一组标准正交基下的矩阵是 $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$，则 $\mathcal{A}^*$ 在此基底下的矩阵是 $\begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix}$。由于

$$A^*A = AA^* = \begin{pmatrix} |\lambda_1|^2 & & \\ & \ddots & \\ & & |\lambda_n|^2 \end{pmatrix}$$

故 $\mathcal{A}$ 是正规算子。

($\Longrightarrow$) 任取正规算子 $\mathcal{A}$ 的一个特征值 $\lambda \in \mathbb{C}$，则其极小多项式应该是下面的形式

$$\mu_{\mathcal{A}}(t) = (t-\lambda)^k \cdot f(t),\ k \in \mathbb{Z}^+.$$

下面我们证明 k 只能等于 1。由极小多项式的定义，任取 $\mathbf{v} \in V$，我们都有 $\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mathbf{0}$，即 $(\mathcal{A}-\lambda\mathcal{E})^k f(\mathcal{A})\mathbf{v} = \mathbf{0}$。所以对任意的 $\mathbf{v} \in V$，都有 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A}-\lambda\mathcal{E})^k)$。由引理3.5.2的 (1) 可知 $\mathcal{A}-\lambda\mathcal{E}$ 也是正规算子，故由引理3.5.2的 (4)，$\ker((\mathcal{A}-\lambda\mathcal{E})^k) = \ker((\mathcal{A}-\lambda\mathcal{E}))$，所以 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A}-\lambda\mathcal{E}))$，即 $(\mathcal{A}-\lambda\mathcal{E})f(\mathcal{A})\mathbf{v} = \mathbf{0},\ \forall \mathbf{v} \in V$。因此，如果 $k > 1$，则有 $(t-\lambda)\cdot f(t)$ 零化 $\mathcal{A}$ 并且 $\deg[(t-\lambda)f(t)] < \deg(\mu_{\mathcal{A}}(t))$，矛盾！因此只能是 $k = 1$。

设 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t) = (t-\lambda_1)^{p_1}\cdots(t-\lambda_s)^{p_s}$，其中 $\lambda_1,\ldots,\lambda_s \in \mathbb{C}$ 两两不同，$p_1 + \cdots + p_s = n$。则由上面的论述以及推论2.4.2可知 $\mu_{\mathcal{A}}(t) = (t-\lambda_1)\cdots(t-\lambda_s)$，故由定理2.5.8的 (vi) 可知 $\mathcal{A}$ 是可对角化的。

最后我们只需要找到一组标准正交基作为 $\mathcal{A}$ 的特征向量即可。由于 $\mathcal{A}$ 可对角化，故 $\mathcal{A}$ 的每个特征值的代数重数都等于几何重数，即 $V = V^{\lambda_1} \oplus V^{\lambda_2} \oplus \cdots \oplus V^{\lambda_s}$，其中 $\dim(V^{\lambda_i}) = p_i$，$\forall i \in \{1,\ldots,s\}$。由定理3.1.2或定理3.2.2，我们可以取出每个特征子空间 V^{λ_i} 的一组标准正交基 $\mathbf{e}_{i1},\ldots,\mathbf{e}_{i,p_i}$，而由引理3.5.2的 (5)，$V_{\lambda_i} \perp V_{\lambda_j}$，故将所有 V_{λ_i} 的标准正交基合到一起就得到了 V 的一组标准正交基 $\mathbf{e}_{11},\ldots,\mathbf{e}_{1,p_1},\cdots\cdots,\mathbf{e}_{s1},\ldots,\mathbf{e}_{s,p_s}$。这样我们就完成了定理3.5.1的证明。 □

与归纳法证明相比，以上证明看起来更加复杂。但实际上，上面的证明方法同时也给出了正规算子的极小多项式和特征子空间的精细结构，这对我们理解下面的谱定理是有帮助的。在我们引入谱定理之前，我们先直接给出欧氏空间上正规算子的对角化结论。与正交算子的情形类似，在欧氏空间上我们只能使正规算子分块对角化。

命题 3.5.1. 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $\mathcal{A}$ 是正规算子 $\Longleftrightarrow$ 存在 V 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 使得

$$(\mathcal{A}\mathbf{e}_1,\ldots,\mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & A_1 & & \\ & & & & \ddots & \\ & & & & & A_m \end{pmatrix}.$$

其中 $\lambda_1,\ldots,\lambda_k\in\mathbb{R}$，$A_i=\begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}\in M_2(\mathbb{R})$，$i\in\{1,\ldots,m\}$，并且 $k+2m=n$。

证明可参考 Advanced Linear Algebra,Steven Roman,GTM135 的 Chapter10,Theorem10.10。

接下来我们引入谱定理。我们的最终目的是将一个正规算子唯一地“拆解”成比较简单的形式，即一些正交投影算子的线性组合。实际上，在定理3.5.1的证明过程中我们就已经得到：$\mathcal{A}=\sum_{i=1}^{s}\lambda_i\mathcal{P}_i$，其中 $\lambda_i, i=1,\ldots,s$ 是 $\mathcal{A}$ 的两两不同的特征值，$\mathcal{P}_i$ 在标准正交基 $\mathbf{e}_{11},\ldots,\mathbf{e}_{s,p_s}$ 下的矩阵是 $\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & E_{p_i} & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$ (即只有 $\mathbf{e}_{i1},\ldots,\mathbf{e}_{i,p_i}$ 对应的对角线元素是 1，其余位置都是 0)。显然每个 $\mathcal{P}_i$ 都是投影算子 (例2.3.2)。这就是我们得到谱定理的思路。

定义 3.5.2. 设 V 是**一般**的域 $\mathbb{K}$ 上的 n 维向量空间。如果两个投影算子 $\mathcal{P}_1,\mathcal{P}_2\in\mathrm{End}(V)$ 满足 $\mathcal{P}_1\mathcal{P}_2=\mathcal{O}$，则我们称 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 正交。更进一步，如果 $V=V_1\oplus\cdots\oplus V_r$，则任取 $\mathbf{x}\in V$，在每个 V_i 内存在唯一的 $\mathbf{x}_i$ 使得 $\mathbf{x}=\sum_{i=1}^{r}\mathbf{x}_i$。对每个 $i\in\{1,\ldots,r\}$，我们令 $\mathcal{P}_i: V\to V_i$，$\mathbf{x}=\sum_{i=1}^{r}\mathbf{x}_i\mapsto\mathbf{x}_i$，则 $\mathcal{P}_1,\ldots,\mathcal{P}_r$ 满足：

$$\mathcal{P}_i\mathcal{P}_j=\delta_{ij}\mathcal{P}_i\ ,\quad \sum_{i=1}^{r}\mathcal{P}_i=\mathcal{E} \tag{3.5.1}$$

我们称 $\mathcal{P}_1,\ldots,\mathcal{P}_r$ 是一个**完全正交的投影算子组**。实际上，如果一组线性算子满足式 (3.5.1)，则我们就可以找到相对应的直和分解 $V=V_1\oplus\cdots\oplus V_r$，使得每个 $\mathcal{P}_i$ 就是 V 到 V_i 的投影算子 (证明是第二章的一道习题，留作练习)。

需要说明的是，内积空间上投影算子并不一定都是自伴随的，尽管其特征值只有 0 和 1。这是因为使得投影算子对角化的基底并不一定是标准正交基，换言之，0 对应的特征子空间和 1 对应的特征子空间不一定是正交的。例如，标准欧氏空间 $\mathbb{R}^2$ 上的线性算子 $\mathcal{P}:\mathbf{x}=(x_1,x_2)^t\mapsto A\mathbf{x}$，$A=\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ 就是投影算子 (验证留作练习)，但它显然不是自伴随算子 ($A\neq A^t$)。实际上，投影算子 $\mathcal{P}$ 是自伴随算子 $\iff\ker(\mathcal{P})\perp\mathrm{im}(\mathcal{P})$，证明是 3.3 节的一道习题，留作练习。

定理 3.5.2 (谱定理,the spectral theorem). 设 V 是**一般**的域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$ 是可对角化的线性算子。如果 $\mathcal{A}$ 的全部的两两不同的特征值为 $\lambda_1,\ldots,\lambda_r$，则存在一个完全正交的投影算子组 $\mathcal{P}_1,\ldots,\mathcal{P}_r$ 使得：

(1) $\sum_{i=1}^{r}\mathcal{P}_i=\mathcal{E}$;

(2) $\sum_{i=1}^{r}\lambda_i\mathcal{P}_i=\mathcal{A}$(称为 $\mathcal{A}$ 的谱分解，spectral decomposition)；

(3) 谱分解是唯一的，即如果另有一个完全正交的投影算子组 $\mathcal{Q}_1,\ldots,\mathcal{Q}_s$ 也满足 $\sum_{i=1}^{s}\mathcal{Q}_i=\mathcal{E}$ 和 $\sum_{i=1}^{s}\mu_i\mathcal{Q}_i=\mathcal{A}$ ($\mu_1,\ldots,\mu_s$ 两两不同)，那么 $r=s$，并且 $\mu_1,\ldots,\mu_r$ 是 $\lambda_1,\ldots,\lambda_r$ 的一个排列，当 $\mu_i=\lambda_j$ 时，$\mathcal{Q}_i=\mathcal{P}_j$。

(4) 存在多项式 $f_1(t),\ldots,f_r(t)\in\mathbb{K}[t]$ 使得 $\forall i,j\in\{1,\ldots,r\}$，$f_i(\lambda_j)=\delta_{ij}$，并且 $f_i(\mathcal{A})=\mathcal{P}_i$。

证明. (1) 由于 $\mathcal{A}$ 可对角化，故 $V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_r}$，对每个 $i\in\{1,\ldots,r\}$，作线性算子

$$\mathcal{P}_i: V\to V^{\lambda_i},\ \mathbf{x}=\sum_{i=1}^{r}\mathbf{x}_i\ (\text{其中}\mathbf{x}_i\in V^{\lambda_i})\mapsto\mathbf{x}_i$$

则每个 $\mathcal{P}_i$ 都是投影算子，并且 $\forall\mathbf{x}\in V$，有 $(\sum_{i=1}^{r}\mathcal{P}_i)\mathbf{x}=\sum_{i=1}^{r}\mathbf{x}_i=\mathbf{x}=\mathcal{E}\mathbf{x}$，即 $\sum_{i=1}^{r}\mathcal{P}_i=\mathcal{E}$。

(2) 对 $\forall \mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i \in V$ (其中$\mathbf{x}_i \in V^{\lambda_i}$)，由于 $\mathbf{x}_i \in V^{\lambda_i}$，故 $\mathscr{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i$，因此

$$\mathscr{A}\mathbf{x} = \mathscr{A}\mathcal{E}\mathbf{x} = \mathscr{A}(\sum_{i=1}^{r} \mathcal{P}_i)\mathbf{x} = \mathscr{A}(\sum_{i=1}^{r} \mathbf{x}_i) = \sum_{i=1}^{r} \mathscr{A}\mathbf{x}_i = \sum_{i=1}^{r} \lambda_i \mathbf{x}_i.$$

而另一方面，按照 $\mathcal{P}_i$ 的定义我们有

$$(\sum_{i=1}^{r} \lambda_i \mathcal{P}_i)\mathbf{x} = \sum_{i=1}^{r} \lambda_i \mathbf{x}_i.$$

因此 $\sum_{i=1}^{r} \lambda_i \mathcal{P}_i = \mathscr{A}$。

(3) 令 $U_i = \mathrm{im}(\mathcal{Q}_i)$，$i = 1, \ldots, s$，则任取 $\mathcal{Q}_i\mathbf{x} \in U_i$(再次注意这个取法)，有

$$\mathscr{A}(\mathcal{Q}_i\mathbf{x}) = (\sum_{i=1}^{s} \mu_i \mathcal{Q}_i)\mathcal{Q}_i\mathbf{x} = \mu_i \mathcal{Q}_i^2\mathbf{x} = \mu_i \mathcal{Q}_i\mathbf{x}.$$

这说明每个 μ_i 都是 $\mathscr{A}$ 的特征值，由于 μ_i 之间两两不同，故 $s \leqslant r$，并且如果设 $\mu_i = \lambda_j$，则 $U_i \subset V^{\lambda_j}$。于是和空间 $U_1 + \cdots + U_s$ 是直和。又因为 $\sum_{i=1}^{s} \mathcal{Q}_i = \mathcal{E}$，所以任取 $\mathbf{x} \in V$，都有 $\mathbf{x} = \mathcal{E}\mathbf{x} = \sum_{i=1}^{s}(\mathcal{Q}_i\mathbf{x})$，即 $V = U_1 + \cdots + U_s$。所以 $V = U_1 \oplus \cdots \oplus U_s$，结合每个 U_i 都是 V^{λ_j} 的子空间以及 $V = V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_r}$ 可知 $r = s$，并且 $U_1, \ldots, U_s$ 就是 $V^{\lambda_1}, \ldots, V^{\lambda_r}$ 的一个排列，即 $\mu_1, \ldots, \mu_r$ 是 $\lambda_1, \ldots, \lambda_r$ 的一个排列，当 $\mu_i = \lambda_j$ 时，$\mathcal{Q}_i = \mathcal{P}_j$。

(4) 任意固定 $i \in \{1, \ldots, r\}$，我们构造出 $f_i(t)$ 即可。由于 $f_i(t)$ 应该满足 $f_i(\lambda_i) = 1$ 以及 $\forall j \neq i$，$f_i(\lambda_j) = 0$，因此由 Lagrange 插值公式 (讲义上册定理 6.2.4) 可知：

$$f_i(t) = \sum_{k=1}^{r} \left[f_i(\lambda_k) \cdot \frac{\prod\limits_{m \neq k}(t - \lambda_m)}{\prod\limits_{m \neq k}(\lambda_k - \lambda_m)} \right] = \frac{\prod\limits_{m \neq i}(t - \lambda_m)}{\prod\limits_{m \neq i}(\lambda_i - \lambda_m)}$$

下面我们证明 $f_i(\mathscr{A}) = \mathcal{P}_i$。首先我们注意到：

$$\mathscr{A}^0 = \mathcal{E} = \sum_{i=1}^{r} \mathcal{P}_i$$

$$\mathscr{A} = \sum_{i=1}^{r} \lambda_i \mathcal{P}_i$$

$$\mathscr{A}^2 = (\sum_{i=1}^{r} \lambda_i \mathcal{P}_i)(\sum_{i=1}^{r} \lambda_i \mathcal{P}_i) = \sum_{i=1}^{r} \lambda_i^2 \mathcal{P}_i \quad \text{(展开括号并利用式 (3.5.1) 整理即得.)}$$

$$\mathscr{A}^3 = (\sum_{i=1}^{r} \lambda_i \mathcal{P}_i)(\sum_{i=1}^{r} \lambda_i^2 \mathcal{P}_i) = \sum_{i=1}^{r} \lambda_i^3 \mathcal{P}_i$$

……

$$\mathscr{A}^k = (\sum_{i=1}^{r} \lambda_i \mathcal{P}_i)(\sum_{i=1}^{r} \lambda_i^{k-1} \mathcal{P}_i) = \sum_{i=1}^{r} \lambda_i^k \mathcal{P}_i \ \ , \ \forall k \in \mathbb{N}$$

……

于是对任意的 $f(t) \in \mathbb{K}[t]$，都有 $f(\mathscr{A}) = \sum_{j=1}^{r} f(\lambda_j)\mathcal{P}_j$。特别地，取 $f = f_i$，利用 $f_i(\lambda_j) = \delta_{ij}$

即得 $f_i(\mathcal{A}) = \mathcal{P}_i$。这样我们就完成了证明。

□

我们注意到，上面定理的结论对于一般的域都成立。将定理3.5.1与定理3.5.2的结论结合起来，我们立刻得到：

命题 3.5.2 (正规算子谱定理). 设 $\mathcal{A}$ 是 n 维 Hermitian 空间 V 上的正规算子，则 $\mathcal{A}$ 可对角化，并且 $\mathcal{A}$ 有唯一的谱分解：$\mathcal{A} = \sum_{i=1}^{r} \lambda_i \mathcal{P}_i$，其中 $\lambda_1, \ldots, \lambda_r$ 是 $\mathcal{A}$ 的全部的两两不同的特征值，$\mathcal{P}_i$ 是如下的投影算子：

$$V \to V^{\lambda_i}, \quad \mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i \text{ (其中} \mathbf{x}_i \in V^{\lambda_i}\text{)} \mapsto \mathbf{x}_i.$$

需要说明的是，在上面的命题中，每个 $\mathcal{P}_i$ 都是自伴随的投影算子。这来源于正规算子的特殊性质：$V^{\lambda_i} \perp V^{\lambda_j}, \forall i \neq j$，再利用我们在定理3.5.2之前的那段论述即可得到 $\mathcal{P}_i$ 的自伴随性质。

现在我们考虑如何将两个正规算子用同一组标准正交基对角化。这当然是有条件的。我们有如下定理：

定理 3.5.3. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$ 都是正规算子。则我们有：存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}, \mathcal{B}$ 在此基底下都是对角矩阵 $\iff$ $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A}$。

证明. ($\Longrightarrow$) 方向是显然的，因为两个对角矩阵显然交换。下面证明相反的方向。
($\Longleftarrow$) 任取 $\mathcal{A}$ 的某个特征子空间 V^{λ_i}，再任取 $\mathbf{x} \in V^{\lambda_i}$，则

$$\mathcal{A}\mathcal{B}\mathbf{x} = \mathcal{B}\mathcal{A}\mathbf{x} = \lambda_i \mathcal{A}\mathbf{x} \Longrightarrow \mathcal{B}\mathbf{x} \in V^{\lambda_i}.$$

这说明每个 $\mathcal{A}$ 的特征子空间 V_{λ_i} 都是 $\mathcal{B}$ 的不变子空间。将 $\mathcal{B}$ 限制到 V^{λ_i} 上可知 V^{λ_i} 中一定有 $\mathcal{B}$ 的特征向量 (利用了复数域上的算子一定有特征向量，定理2.3.2)。于是 $\mathcal{A}$ 和 $\mathcal{B}$ 一定有公共的特征向量，设为 $\mathbf{y}$。取 $\mathbf{e}_1 = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$，则 $\|\mathbf{e}_1\| = 1$ 且 $\mathrm{span}\ \{\mathbf{e}_1\}$ 是 $\mathcal{A}$ 和 $\mathcal{B}$ 公共的特征子空间。

下面我们证明：$W = \mathrm{span}\ \{\mathbf{e}_1\}^{\perp}$ 同时是 $\mathcal{A}$ 和 $\mathbb{B}$ 的不变子空间，并且 $\mathcal{A}$ 和 $\mathcal{B}$ 限制到 W 上也是正规算子。

任取 $\mathbf{w} \in W$，即 $\mathbf{w} \perp \mathbf{e}_1$，我们有

$$<\mathcal{A}\mathbf{w}, \mathbf{e}_1> = <\mathbf{w}, \mathcal{A}^*\mathbf{e}_1> = <\mathbf{w}, \overline{\lambda_i}\mathbf{e}_1> = \lambda_i <\mathbf{w}, \mathbf{e}_1> = 0 \quad \text{(这里用到了引理3.5.2的 (3).)}$$

即 $\mathcal{A}\mathbf{w} \in \mathrm{span}\ \{\mathbf{e}_1\}^{\perp} = W$，所以 W 是 $\mathcal{A}-$ 子空间。对 $\mathcal{B}$ 重复一遍上述过程即可得到 W 也是 $\mathcal{B}-$ 子空间。$\mathcal{A}, \mathcal{B}$ 限制到 W 上是正规算子上显然的。

因此，对 $\mathcal{A}\big|_W$ 和 $\mathcal{B}\big|_W$ 而言，重复我们一开始的操作，我们仍可以在 W 中找到 $\mathcal{A}, \mathcal{B}$ 的公共特征向量 $\mathbf{e}_2$，$\cdots\cdots$ 归纳法即可找到 V 的标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使 $\mathcal{A}, \mathcal{B}$ 在此基底下都是对角矩阵。 □

需要注意的是，以上定理在欧氏空间上不对，例如在 $\mathbb{R}^2$ 上取正规矩阵 $A = E_2$，$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，则 $AB = BA$，但 B 不存在实的特征值和特征向量！

在本节的最后。我们来回答 3.3 节遗留下来的一个问题：(半) 正定算子的平方根为什么是唯一的。

定理 3.5.4. 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}$ 是 V 上的半正定线性算子，则存在唯一的半正定线性算子 $\mathcal{B}$，使得 $\mathcal{B}^2 = \mathcal{A}$。

证明. 存在性已经在定理3.3.6中证明过了，下证唯一性。设另有半正定算子 $\mathcal{B}' \in \mathrm{End}(V)$ 也满足 $\mathcal{B}'^2 = \mathcal{A}$，由谱定理（定理3.5.2），我们知道 $\mathcal{A}, \mathcal{B}, \mathcal{B}'$ 分别有谱分解

$$\begin{aligned}
\mathcal{A} &= \sum_{i=1}^{r} \lambda_i \mathcal{P}_i,\\
\mathcal{B} &= \sum_{j=1}^{s} \mu_i \mathcal{Q}_j,\\
\mathcal{B}' &= \sum_{k=1}^{s} \mu'_k \mathcal{Q}'_k.
\end{aligned}$$

其中每个 $\lambda_i \geqslant 0$，$i \in \{1, \ldots, r\}$。于是由谱定理中投影的正交关系可得

$$\begin{aligned}
\mathcal{A} &= \sum_{i=1}^{r} \lambda_i \mathcal{P}_i\\
=\mathcal{B}^2 &= \sum_{j=1}^{s} \mu_j^2 \mathcal{Q}_j\\
=\mathcal{B}' &= \sum_{k=1}^{s} \mu_k'^2 \mathcal{Q}'_k.
\end{aligned}$$

这给出了 $\mathcal{A}$ 的三个谱分解。由谱分解的唯一性可知 $r = s = t$ 且经过适当的次序调整我们有

$$\forall i \in \{1, \ldots, r\},\ \lambda_i = \mu_i^2 = \mu_i'^2 \text{ 且 } \mathcal{P}_i = \mathcal{Q}_i = \mathcal{Q}'_i.$$

由于 $\mathcal{B}$ 和 $\mathcal{B}'$ 都是半正定的，即每个 μ_i 和 μ'_i 都非负，因此我们有

$$\mu_i = \mu'_i,\ \forall i \in \{1, \ldots, r\}.$$

这说明 $\mathcal{B}$ 和 $\mathcal{B}'$ 具有相同的谱分解，即 $\mathcal{B} = \mathcal{B}'$。这样我们就证明了平方根的唯一性。 □

3.6 复化与实化 *

我们在前面几节中已经看到，实内积空间和复内积空间的性质十分相似，这启发我们考虑实向量空间和复向量空间本身的关系。我们在讲义上册第五章中已经知道，$\mathbb{C}$ 可以视作 $\mathbb{R}$ 上的二维向量空间 $\left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \,\middle|\, a, b \in \mathbb{R}\right\}$，将它推广到 n 维情形就是我们本节的内容。本节内容我们大多数只给出结论，证明可参考代数学引论的相关章节。

首先，设 V 是 $\mathbb{R}$ 上的 n 维向量空间，我们希望通过 V 构造出复向量空间 V'，并且 V' 上的复数数乘可以用 V 上的某种 “操作” 表示出来。这就是所谓的 “复化” (complexification)。

定义 3.6.1. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，$\mathcal{J} \in \mathrm{End}(V)$ 并且 $\mathcal{J}^2 = -\mathcal{E}$，则我们可以在 V 上定义如下的复数数乘：

$$\cdot : \ \mathbb{C} \times V \to V, \ (a + bi \ , \ \mathbf{v}) \mapsto a\mathbf{v} + b\mathcal{J}\mathbf{v} \ (\text{其中} a, b \in \mathbb{R}.) \tag{3.6.1}$$

验证这确实是一个数乘运算 (即验证它满足定义1.1.1的 (2) 和 (3)) 留作练习。于是 V 在原来的加法和式 (3.6.1) 定义的数乘之下是 $\mathbb{C}$ 上的向量空间，我们把这个向量空间记作 $V_{\mathcal{J}}$，其在 $\mathbb{C}$ 上的维数记作 $\dim_{\mathbb{C}}(V_{\mathcal{J}})$，并称式 (3.6.1) 为 V 上的**复结构**。

我们先来看两个复结构的具体例子。

例 3.6.1. (1) 在 $\mathbb{R}^2$ 上算子 $\mathcal{J} : \mathbf{x} = (x_1, x_2)^t \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\mathbf{x}$ 可以定义一个复结构：

$$(a + bi) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto a\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 - bx_2 \\ ax_2 + bx_1 \end{pmatrix}.$$

这实际上就是复数乘法：$(a + bi)(x_1 + x_2 i) = (ax_1 - bx_2) + (ax_2 + bx_1)i$。

(2) 在 $2m$ 维的辛空间 V 上有一个自然的复结构：取辛空间的一组辛基 $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_m}, \boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_m}$，则我们可以定义算子 $\mathcal{J}$ 使得 $\mathcal{J}$ 在这组辛基下的矩阵恰为 $J_0 = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$。容易验证 $\mathcal{J}$ 是辛算子，并且 $\mathcal{J}^2 = -\mathcal{E}$，于是可以用 $\mathcal{J}$ 在 V 上定义一个自然的复结构。

我们有以下定理：

定理 3.6.1. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，如果 V 上存在一个用线性算子 $\mathcal{J}$ ($\mathcal{J}^2 = -\mathcal{E}$) 定义的复结构，则 V 的实维数 (记为 $\dim_{\mathbb{R}}(V)$) 是偶数 $n = 2m$。更进一步，我们一定能找到某组基 $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_m}, \boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_m}$ 使得 $\mathcal{J}$ 在这组基下的矩阵形如 $\begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$，并且 $\dim_{\mathbb{C}}(V_{\mathcal{J}}) = \frac{1}{2}\dim_{\mathbb{R}}(V) = m$。

证明过程实际上就是寻找斜对称双线性型的标准基，留作思考。

于是，我们可以从一般的实向量空间 V 出发，先构造出偶数维的 $V \oplus V$(外直和，下面会严格定义)，再在 $V \oplus V$ 上定义一个特殊的复结构就可以得到一个复维数与 $\dim_{\mathbb{R}}(V)$ 相等的向量空间。这就是下面的复化。

定义 3.6.2. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，则容易看出 $V \times V$ 在下面的加法和实数乘法下

是 $\mathbb{R}$ 上的 $2n$ 维的向量空间：

$$+:\ (V\times V)\times(V\times V)\to(V\times V),\ (\mathbf{x}_1,\mathbf{x}_2)+(\mathbf{y}_1,\mathbf{y}_2)=(\mathbf{x}_1+\mathbf{y}_1,\mathbf{x}_2+\mathbf{y}_2);$$
$$\cdot_{\mathbb{R}}:\ \mathbb{R}\times(V\times V)\to(V\times V),\ a\cdot(\mathbf{x}_1,\mathbf{x}_2)=(a\mathbf{x}_1,a\mathbf{x}_2).$$

我们把 $V\times V$ 在上述运算下构成的向量空间称为 V 和 V 的外直和，记作 $V\oplus V$[1]。设 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，则容易看出 $(\mathbf{e}_1,\mathbf{0}),\dots,(\mathbf{e}_n,\mathbf{0}),(\mathbf{0},\mathbf{e}_1),\dots,(\mathbf{0},\mathbf{e}_n)$ 是 $V\oplus V$ 的一组基。显然在 $V\oplus V$ 上线性算子 $\mathcal{J}:(\mathbf{u},\mathbf{v})\mapsto(-\mathbf{v},\mathbf{u})$ 满足 $\mathcal{J}^2=-\mathcal{E}$ (实际上 $\mathcal{J}$ 在上述基底下的矩阵恰好是 $\begin{pmatrix}0 & -E_n\\ E_n & 0\end{pmatrix}$)，于是 $\mathcal{J}$ 可以定义 $V\oplus V$ 上的一个复结构，称为标准复结构。相应地，我们把复向量空间 $(V\oplus V)_{\mathcal{J}}$ 称为 V 的复化空间，记作 $V^{\mathbb{C}}$。

我们也可以把 $(V\oplus V)_{\mathcal{J}}$ 中的向量 $(\mathbf{u},\mathbf{v})$ 记作 $\mathbf{u}+i\mathbf{v}$。容易验证

$$(a+bi)(\mathbf{u}+i\mathbf{v})=a(\mathbf{u}+i\mathbf{v})+b(-\mathbf{v}+i\mathbf{u})=(a\mathbf{u}-b\mathbf{v})+i(a\mathbf{v}+b\mathbf{u}).$$

即这样的形式记号与复数数乘是相容的。稍后我们会看到，这种记号实际上是先复化，再实化。特别地，我们看到 $\mathbb{C}$ 就是 $\mathbb{R}$ 的复化。由定理3.6.1立刻有：$\dim_{\mathbb{C}}(V^{\mathbb{C}})=\dim_{\mathbb{R}}(V)$。实际上，我们很容易验证：如果 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，则 $\mathbf{e}_1+i\mathbf{0},\dots,\mathbf{e}_n+i\mathbf{0}$ 也是 $V^{\mathbb{C}}$ 的一组基。

下面我们考虑线性算子的复化。

定义 3.6.3. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$，则我们可以定义 $\mathcal{A}$ 的复化算子 $\mathcal{A}^{\mathbb{C}}:V^{\mathbb{C}}\to V^{\mathbb{C}}$，$\mathbf{u}+i\mathbf{v}\mapsto\mathcal{A}\mathbf{u}+i\mathcal{A}\mathbf{v}$。

利用复化算子我们也可以很容易地证明：实向量空间上的线性算子一定有 1 维或 2 维的不变子空间，留作思考。

下面我们考虑相反方向的问题：给出一个 n 维复向量空间 U，我们如何得到一个实向量空间呢？这个方向是简单的，我们只需要把数乘运算限制到实数域上就可以了。

定义 3.6.4. 设 U 是 n 维复向量空间，$\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 U 的一组基，则显然 $U=\{\sum_{k=1}^n a_k\mathbf{e}_k+\sum_{k=1}^n b_k(i\mathbf{e}_k)\mid a_k,b_k\in\mathbb{R}\}$，即我们可以把 U 视作 $\mathbb{R}$ 上的 $2n$ 维实向量空间，称作 U 的实化空间，记为 $U_{\mathbb{R}}$。同样地，U 上的线性算子 $\mathcal{A}$ 也可以视作 $U_{\mathbb{R}}$ 上的线性算子，称为 $\mathcal{A}$ 的实化算子，记作 $\mathcal{A}_{\mathbb{R}}$。

设 $\mathcal{A}\in\mathrm{End}(U)$ 在 U 的基底 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵为 $A=A_1+iA_2$，$A_1,A_2\in M_n(\mathbb{R})$，则

$$\mathcal{A}(i\mathbf{e}_1,\dots,i\mathbf{e}_n)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot i(A_1+iA_2)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot(-A_2+iA_1).$$

因此 $\mathcal{A}_{\mathbb{R}}$ 在 $\mathbf{e}_1,\dots,\mathbf{e}_n,i\mathbf{e}_1,\dots,i\mathbf{e}_n$ 下的矩阵为 $A_{\mathbb{R}}=\begin{pmatrix}A_1 & -A_2\\ A_2 & A_1\end{pmatrix}$。由此可见，一个 $U_{\mathbb{R}}$ 上的算子要想称为 U 中某个算子的实化是需要一定条件的。

定理 3.6.2. 设 U 是 n 维复向量空间，$\mathcal{A}\in\mathrm{End}(U)$，则 $\det(\mathcal{A}_{\mathbb{R}})=|\det(\mathcal{A})|^2$。

证明留作思考。

容易验证 (过程留作练习)$U_{\mathbb{R}}$ 上所有能称为某个 $\mathrm{End}(U)$ 中算子的实化算子的线性算子构成了一个代数，我们把这个代数记作 $\mathrm{End}(U)_{\mathbb{R}}$，它显然是 $\mathrm{End}(U_{\mathbb{R}})$ 的一个子代数。实际

---

[1] 为了与 1.2 节的直和 (称为内直和) 区分，也有的书将外直和记作 $V\boxplus V$。

上，可以证明 $\dim(\mathrm{End}(U)_{\mathbb{R}}) = \dfrac{1}{2}\dim(\mathrm{End}(U_{\mathbb{R}}))$，留作思考。那么，$\mathrm{End}(U)_{\mathbb{R}}$ 中的元素是不是都具有 $\begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix}$ 这样的矩阵形式呢？答案是肯定的，这就是下面的定理。

定理 3.6.3. 设 U 是 n 维复向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 U 的一组基，则 $\mathrm{End}(U)_{\mathbb{R}}$ 由所有与 $\mathcal{J}$ 交换的算子构成，其中 $\mathcal{J} \in \mathrm{End}(U_{\mathbb{R}})$ 在 $U_{\mathbb{R}}$ 的基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n, i\mathbf{e}_1, \ldots, i\mathbf{e}_n$ 下的矩阵是 $\begin{pmatrix} 0 & -E_n \\ E_n & 0 \end{pmatrix}$。换言之，任取 $\mathcal{A}_{\mathbb{R}} \in \mathrm{End}(U)_{\mathbb{R}}$，一定存在 $\mathcal{U}_{\mathbb{R}}$ 的一组基使得 $\mathcal{A}_{\mathbb{R}}$ 具有 $\begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix}$ 的矩阵形式。

证明留作思考。我们称这种情况为 $\mathcal{A}$ 与实向量空间 V(实际上等于 $U_{\mathbb{R}}$) 上的某个复结构相容。

特别地，我们有：

定理 3.6.4. 二维实向量空间上的没有特征向量的线性算子一定与某个复结构相容。即设 V 是 2 维实向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 V 是某个一维复向量空间 U 的实化，并且存在 $\mathcal{B} \in \mathrm{End}(U)$ 使得 $\mathcal{B}_{\mathbb{R}} = \mathcal{A}$。

证明留作思考。

接下来我们考虑将实向量空间先复化，再实化。设 V 是 n 维实向量空间，令 $W = (V^{\mathbb{C}})_{\mathbb{R}}$，则容易验证 $W = V \oplus iV$，我们称 V 是 W 的实平面，iV 是 W 的虚平面。令 $\mathcal{J} = (i\mathcal{E})_{\mathbb{R}} \in \mathrm{End}(W)$，则 $\mathcal{J}$ 交换实平面与虚平面。

我们知道一个复数可以取其共轭复数，类似地，我们可以定义 W 上的取共轭如下：$\overline{\mathbf{u} + i\mathbf{v}} = \mathbf{u} - i\mathbf{v}$。进一步，我们可以定义一个算子 $\mathcal{A} \in V^{\mathbb{C}}$ 的复共轭算子：$\overline{\mathcal{A}} : \mathbf{u} + i\mathbf{v} \mapsto \overline{\mathcal{A}(\mathbf{u} + i\mathbf{v})}$。容易验证 $\mathcal{A} = \overline{\mathcal{A}} \iff \mathcal{A}$ 限制到实平面上的像集仍在实平面内，即 $\mathcal{A}$ 是某个 $\mathcal{B} \in \mathrm{End}(V)$ 的复化算子。此外容易验证下面的等式：

$$\mathrm{tr}(\mathcal{A}_{\mathbb{R}}) = \mathrm{tr}(\mathcal{A}) + \mathrm{tr}(\overline{\mathcal{A}}).$$

在本节的最后，我们相对应地考虑将复向量空间先实化，再复化。首先，设 V 是 n 维复向量空间，则我们可以自然地定义 V 的复共轭空间 $\overline{V}$：$\overline{V}$ 中的元素和其上的加法与 V 一致，而数乘 $\odot$ 则定义成：$\lambda \odot \mathbf{x} = \overline{\lambda}\mathbf{x},\ \forall \lambda \in \mathbb{C}, \mathbf{x} \in V$。我们可以证明：$(V_{\mathbb{R}})^{\mathbb{C}}$ 同构于 $V \oplus \overline{V}$，过程留作思考。

3.7 正交展开 *

这一节我们简单地介绍一下无穷维内积空间的部分性质。要想研究无穷维空间，一个不可避免的问题就是在无穷维空间中我们能否找到与有限维空间的基底性质相似的数学概念，这就是我们本节的主题。

定义 3.7.1. 设 $(V, <\cdot,\cdot>)$ 是无穷维的实 (或复) 内积空间，$\{\mathbf{e}_n\}_{n=1}^{\infty} \subset V$ 是一个所有向量都非零的序列。如果 $\forall i, j \in \mathbb{Z}^+$，都有 $\mathbf{e}_i \perp \mathbf{e}_j$，则称 $\{\mathbf{e}_n\}$ 是一个可数的正交系；如果 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 还满足 $\|\mathbf{e}_i\| = 1, \forall i \in \mathbb{Z}^+$，则称 $\{\mathbf{e}_n\}$ 是一个标准的 (也称规范的) 正交系。容易验证一个正交系中任意有限个向量都是线性无关的[1]。如果正交系 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 还满足 $\text{span}\,\{\{\mathbf{e}_n\}_{n=1}^{\infty}\}^{\perp} = \{\mathbf{0}\}$，即 $(<\mathbf{v}, \mathbf{e}_i> = 0, \forall i \in \mathbb{Z}^+ \Longrightarrow \mathbf{v} = \mathbf{0})$，那么我们称 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 是完备的正交系 (也有的书称其为完全正交系)。如果一个正交系既是标准的又是完备的，则我们称其为标准完备正交系。

我们希望标准完备正交系在无穷维内积空间中的作用正如标准正交基在有限维内积空间中的作用，也就是说，我们希望 $\forall \mathbf{x} \in V$，有 $\mathbf{x} = \sum_{i=1}^{\infty} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$。这就产生了两个问题：一是 $\sum_{i=1}^{\infty} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$ 何时是 (以何种意义) 收敛的，二是如果收敛那么极限何时恰好是 $\mathbf{x}$。下面我们逐步回答这两个问题。

引理 3.7.1. 设 $(V, <\cdot,\cdot>)$ 是实 (或复) 内积空间，则内积 $<\mathbf{x}, \mathbf{y}>$ 关于坐标 $(\mathbf{x}, \mathbf{y})$ 是连续的。换言之，任取序列 $\{\mathbf{x}_n\}_{n=1}^{\infty} \xrightarrow{n\to\infty} \mathbf{x}_0$，$\{\mathbf{y}_n\}_{n=1}^{\infty} \xrightarrow{n\to\infty} \mathbf{y}_0$(按距离收敛)，则 $<\mathbf{x}_n, \mathbf{y}_n> \xrightarrow{n\to\infty} <\mathbf{x}, \mathbf{y}>$。

利用 CBS 不等式即可证明，留作思考。

容易验证以下事实：$\forall \mathbf{x} \in V$，如果 $\mathbf{x} = \sum_{i=1}^{\infty} a_i\mathbf{e}_i$，$a_i \in \mathbb{R}$(或$\mathbb{C}$)，那么每个 $a_i = <\mathbf{x}, \mathbf{e}_i>$ 都必然是唯一的，留作练习。我们称 $a_i = <\mathbf{x}, \mathbf{e}_i>$ 是向量 $\mathbf{x}$ 关于标准正交系 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 的傅里叶 (Fourier) 系数。

引理 3.7.2 (Bessel 不等式). 设 $(V, <\cdot,\cdot>)$ 是实 (或复) 内积空间，$\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是标准正交系，则 $\forall \mathbf{x} \in V$，都有 $\|\mathbf{x}\|^2 \geqslant \sum_{i=1}^{\infty} |<\mathbf{x}, \mathbf{e}_i>|^2$。

证明留作思考，我们在 3.1 节的脚注中已经给出了参考文献。

定理 3.7.1. 在完备内积空间 (即柯西列都收敛,Hilbert 空间)$(V, <\cdot,\cdot>)$ 中，对于完备的标准正交系 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 而言，任取 $\mathbf{x} \in V$，部分和 $\sum_{i=1}^{n} < \mathbf{x}, \mathbf{e}_i >$ 在 $n \to \infty$ 时都按距离收敛到 $\mathbf{x}$ 本身。

证明可参考《A Course in Functional Analysis》, John B.Conway, GTM96 的 Chapter I, §4, Theorem4.13, 或者《泛函分析讲义》, 张恭庆, 北京大学出版社的第一章, 定理 1.6.25。

注意，如果内积空间不是完备的，那么部分和 $\sum_{i=1}^{n} < \mathbf{x}, \mathbf{e}_i >$ 在 $n \to \infty$ 时甚至可以不收敛，反例留作思考。

定理 3.7.2. 设 $(V, <\cdot,\cdot>)$ 是实 (或复) 内积空间，则一定存在完备内积空间 $\hat{V}$ 使得：

1. $\hat{V}$ 中有一个子空间 V_1 同构于 V；
2. $\hat{V}$ 中每个向量都是 V_1 中的某个柯西序列的极限，即 V_1 在 V 中稠密。

我们称 $\hat{V}$ 是 V 的完备化，并且在同构意义下完备化是唯一的。

完备化的过程实际上可以推广到一般的度量空间上，过程可参考《泛函分析讲义》，张恭庆，北京大学出版社的第一章 §2。

---

[1] 我们不妨证明 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 线性无关。设 $\sum_{i=1}^{n} a_i\mathbf{e}_i = \mathbf{0}$，等式两边与每个 $\mathbf{e}_i$ 作内积即得 $a_i\|\mathbf{e}_i\|^2 = 0, \forall i$，而 $\|\mathbf{e}_i\| \neq 0$，故每个 a_i 都是 0。此即线性无关。

定理 3.7.3. 设 $(V, < \cdot, \cdot >)$ 是实 (或复) 内积空间，如果 V 的子空间 U 在 V 中稠密 (即 U 的闭包包含 V，详细的定义参考分析学的标准教材)，那么 U 在 $\hat{V}$ 中也稠密。

证明留作思考。

此外，我们很容易证明 Gram-Schmidt 正交化方法在无穷维空间中仍然能够使用 (留作练习)。于是将以上的结论都结合起来，我们有：

定理 3.7.4. 设 $(V, < \cdot, \cdot >)$ 是实 (或复) 内积空间，向量序列 $\{\mathbf{v}_i\}_{i=1}^{\infty} \subset V$ 并且其中任意有限个向量线性无关，对 $\{\mathbf{v}_i\}_{i=1}^{\infty}$ 施以 Gram-Schmidt 正交化得到向量序列 $\{\mathbf{e}_i\}_{i=1}^{\infty}$。

(1) 如果 span $\{\{\mathbf{v}_i\}_{i=1}^{\infty}\}$ 在 V 中稠密，则 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是 $\hat{V}$ 中的标准完备正交系，从而 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 也是 V 的标准完备正交系。

(2) 如果 V 是完备的，则 V 中任何标准完备正交系张成的子空间在 V 中稠密。

证明留作练习。

推论 3.7.1. 设 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是内积空间 V 的标准正交系，如果 span $\{\{\mathbf{e}_i\}_{i=1}^{\infty}\}$ 在 V 中稠密，则 $\forall \mathbf{x} \in V$，都有 $\mathbf{x} = \sum_{i=1}^{\infty} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$。

最后我们看两个例子。

例 3.7.1. (1) 闭区间 $[a, b]$ 上的连续函数空间 $C[a, b]$ 在内积 $< f, g >= \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$ 下不是完备的，但是，容易验证 span $\{t^n \mid \forall n \in \mathbb{N}\}$ 在 $C[a, b]$ 中是稠密的 (数学分析中的 Weierstrass 逼近定理)，故对 $\{t^n \mid \forall n \in \mathbb{N}\}$ 进行 Gram-Schmidt 正交化即可得到 $C[a, b]$ 的一个标准完备正交系。这是稍后正交多项式的讨论内容。

(2) 在 $L^2[0, 2\pi]$(平方可积函数构成的内积空间，相关定义可参考标准的实分析或泛函分析教材) 中，定义内积 $< f, g >= \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$，这是一个完备内积空间，容易验证 $e_n(t) = \dfrac{1}{\sqrt{2\pi}} e^{in\pi}$, $n = 0, \pm 1, \pm 2, \ldots$ 是一个标准完备正交系。这是傅里叶分析 (傅里叶级数) 讨论的内容。

3.8 正交投影与最小二乘法

这一节我们考虑如下问题：设 $(V, <\cdot,\cdot>)$ 是 n 维内积空间，U 是 V 的 m 维子空间，我们定义 $\mathbf{x} \in V$ 到 U 的距离 $\mathrm{d}(\mathbf{x}, U) = \inf\limits_{\mathbf{y}\in U} \mathrm{d}(\mathbf{x}, \mathbf{y})$，那么我们会问：是否存在 $\mathbf{y}_0 \in U$ 使得 $\mathrm{d}(\mathbf{x}, U) = \mathrm{d}(\mathbf{x}, \mathbf{y}_0)$？如果存在，那么 $\mathbf{y}_0$ 是否是唯一的，以及如何计算 $\mathbf{y}_0$？我们接下来会逐步回答这些问题。

定理 3.8.1. *条件如上所述，则满足上述要求的 $\mathbf{y}_0$ 是存在唯一的。实际上，由定理3.1.3或定理3.2.3，设 $V = U \oplus U^{\perp}$，则 $\mathbf{x}$ 可以唯一地分解成 $\mathbf{x} = \mathbf{y}_0 + \mathbf{z}$，$\mathbf{y}_0 \in U, \mathbf{z} \in U^{\perp}$，于是 $\mathbf{y}_0$ 即为我们所求。*

证明. 我们只需证明 $\forall \mathbf{u} \in U$，都有 $\mathrm{d}(\mathbf{x}, \mathbf{u}) \geqslant \mathrm{d}(\mathbf{x}, \mathbf{y}_0)$，并且等号仅在 $\mathbf{u} = \mathbf{y}_0$ 处成立即可。事实上，注意到 $\mathbf{z} = \mathbf{x} - \mathbf{y}_0 \in U^{\perp}$，而 $\mathbf{y}_0 - \mathbf{u} \in U$，故 $(\mathbf{x} - \mathbf{y}_0) \perp (\mathbf{y}_0 - \mathbf{u})$, $\forall \mathbf{u} \in U$，因此我们有：

$$\begin{aligned} \|\mathbf{x} - \mathbf{u}\|^2 &= \|(\mathbf{x} - \mathbf{y}_0) + (\mathbf{y}_0 - \mathbf{u})\|^2 \\ &= \|\mathbf{x} - \mathbf{y}_0\|^2 + \|\mathbf{y}_0 - \mathbf{u}\|^2 \text{（勾股定理）} \\ &\geqslant \|\mathbf{x} - \mathbf{y}_0\|^2 \end{aligned}$$

即 $[\mathrm{d}(\mathbf{x}, \mathbf{u})]^2 \geqslant [\mathrm{d}(\mathbf{x}, \mathbf{y}_0)]^2$，并且等号当且仅当 $\|\mathbf{y}_0 - \mathbf{u}\| = 0$，即 $\mathbf{y}_0 = \mathbf{u}$ 时成立。这样我们就完成了证明。 □

在上面的条件下，我们定义投影算子 $\mathcal{P}_U : V \to U$, $\mathbf{x} \mapsto \mathbf{y}_0$，则我们称 $\mathcal{P}_U$ 是正交投影算子，$\mathbf{y}_0$ 是向量 $\mathbf{x}$ 到 U 的正交投影。

例 3.8.1. *设 $V = C[-\pi, \pi]$ 是闭区间 $[-\pi, \pi]$ 上的实值连续函数空间，内积定义成 $< f, g >= \int_{-\pi}^{\pi} f(t)g(t)\,\mathrm{d}t$，设 $U = \mathrm{span}\,\{1, t, t^2, t^3, t^4, t^5\}$，求 $f(t) = \sin t$ 到 U 的正交投影以及 $\mathrm{d}(f, U)$。*

解. 对 U 的基底 $1, \ldots, t^5$ 作 Gram-Schmidt 正交化，可以得到 U 的一组标准正交基 $\mathbf{e}_1(t), \ldots, \mathbf{e}_6(t)$。于是 f 到 U 的正交投影 $\mathcal{P}_U(f)$ 为：

$$\mathcal{P}_U(f) = \sum_{i=1}^{6} < f(t), \mathbf{e}_i(t) > \mathbf{e}_i(t) = \sum_{i=1}^{6} \left(\int_{-\pi}^{\pi} [\mathbf{e}_i(s) \sin s]\,\mathrm{d}s \right) \cdot \mathbf{e}_i(t) \approx 0.98 - 0.155t^3 + 0.0056t^5$$

并且 $\mathrm{d}(f, U) = \sqrt{\int_{-\pi}^{\pi} (f(t) - \mathcal{P}_U(f)(t))^2\,\mathrm{d}t} \approx 0.001$。读者可以自行验证在题设距离的意义下这个逼近比 Taylor 展开逼近要精确。 □

下面我们来考虑一个更常见的计算场景。设 $A \in \mathbb{R}^{m\times n}$, $m > n$, $\mathbf{b} \in \mathbb{R}^m$，未定向量 $\mathbf{x} \in \mathbb{R}^n$，那么我们知道方程组 $A\mathbf{x} = \mathbf{b}$ 不一定有解。但是，按照上面的讨论，我们总可以在 $\mathbb{R}^n$ 中找到 $\mathbf{x}_0$ 使得在 $\mathbb{R}^m$ 的标准内积下 $\mathrm{d}(A\mathbf{x}_0, \mathbf{b})$ 最小，因为在 $\mathbf{x}_0$ 取遍 $\mathbb{R}^n$ 时 $A\mathbf{x}_0$ 恰好取遍 A 的列空间 $V_c(A)$ (把 A 视作线性算子则是 $\mathrm{im}(A)$)，于是我们只需寻找 $\mathbf{b}$ 到 $V_c(A)$ 的正交投影即可找到 $A\mathbf{x}_0$。下面我们开始计算这个正交投影。

我们不妨设 $\mathrm{rank}(A) = r$ 并且 A 的列空间是 $\mathrm{span}\,\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(r)}\}$，则对 $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(r)}$ 施以 Gram-Schmidt 正交化可以得到 $V_c(A)$ 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_r \in \mathbb{R}^m$，令

$$\mathbf{b}_0 = \sum_{k=1}^{r} < \mathbf{b}, \mathbf{e}_k > \mathbf{e}_k \in \mathbb{R}^m,$$

其中 $<\cdot,\cdot>$ 是 $\mathbb{R}^m$ 上的标准内积。则 $\mathbf{b}_0$ 就是 $\mathbf{b}$ 到 $V_c(A)$ 的正交投影。之后，我们求解非齐次线性方程组 $A\mathbf{x}_0=\mathbf{b}_0$，则 $\mathbf{x}_0\in\mathbb{R}^n$ 即是我们所求。注意方程组 $A\mathbf{x}_0=\mathbf{b}_0$ 总是存在唯一解，这是因为我们很容易验证 $\mathrm{rank}(A)=\mathrm{rank}(A\mid\mathbf{b}_0)=r$，即系数矩阵的秩等于增广矩阵的秩。

事实上，如果 $\mathrm{rank}(A)=n$，那么我们可以更简单地写出 $\mathbf{x}_0$。首先，此时 $\mathrm{rank}(A^tA)=\mathrm{rank}(A)=n$(引理3.5.1)，于是 A^tA 可逆。设 $\mathbf{b}=A\mathbf{x}_0+\mathbf{w}$ 是 $\mathbf{b}$ 沿着 $\mathbb{R}^m=V_c(A)\oplus(V_c(A))^\perp$ 的直和分解，即 $A\mathbf{x}_0\perp\mathbf{w}=\mathbf{b}-A\mathbf{x}_0$，则

$$<A\mathbf{x}_0,\mathbf{b}-A\mathbf{x}_0>=\mathbf{x}_0^tA^t(\mathbf{b}-A\mathbf{x}_0)=0, \tag{3.8.1}$$

即 $\mathbf{x}_0^t(A^tA\mathbf{x}_0-A^t\mathbf{b})=0$。故取 $\mathbf{x}_0=(A^tA)^{-1}A^t\mathbf{b}$ 时，式 (3.8.1) 成立，而由正交投影的唯一性可知 $\mathbf{x}_0$ 是唯一的，即为所求。

我们把上面的求解正交投影的方法称为**最小二乘法** (ordinary least squares)。最小二乘法在实际问题中有广泛的应用，如统计学中的回归分析即是最小二乘法的应用。有关最小二乘法及回归分析的内容，读者可以参考《线性统计模型：线性回归与方差分析》，王松桂，陈敏，陈丽萍，高等教育出版社。

3.9　正交多项式

在本章的最后，我们考虑多项式空间 $\mathbb{R}[t]$ 在不同的内积下的正交基，它们被称为正交多项式。

首先，我们考虑内积 $< f, g >= \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$。在此内积下我们对基底 $1, t, t^2, \cdots\cdots$ 施以 Gram-Schmidt 正交化 (不对向量的范数归一，只进行正交化的部分)，则可以得到 $\mathbb{R}[t]$ 的一组正交基：

$$u_0(t) = 1,\ u_1(t) = t,\ u_2(t) = t^2 - \frac{1}{3},\ u_3(t) = t^3 - \frac{3}{5}t, \cdots\cdots$$

这些正交基的一般公式如下：

$$u_n(t) = \frac{2^n(n!)^2}{(2n)!}P_n(t),\ \text{其中}\ P_n(t) = \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n.$$

可以证明 $P_n(t)$ 之间满足如下的正交性质：

$$\int_{-1}^{1} P_i(t)P_j(t)\,\mathrm{d}t = \frac{2}{2j+1}\delta_{ij}.$$

我们把上面出现的 $P_n(t)$ 的显式表达式称为 Legendre(勒让德) 多项式的 Rodrigue 公式。Rodrigue 公式的证明可以参考《数值逼近》，王仁宏，高等教育出版社的第四章 §7 或《函数逼近论方法》，莫国端，刘开第，科学出版社的 §5.2。

Legendre 多项式也可以由递推式 $P_0(t) = 1, P_1(t) = t, P_{n+1}(t) = \dfrac{2n+1}{n+1}tP_n(t) - \dfrac{n}{n+1}P_{n-1}(t)$ 定义，可以验证这个定义与 Rodrigue 公式是等价的。这是因为：在 $\mathbb{R}[t]$ 上定义自伴随算子 (验证之)：

$$\mathcal{S}(f(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\left((t^2-1)\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right)$$

验证：$\mathcal{S}P_n(t) = n(n+1)P_n(t), \forall n \in \mathbb{N}$ 即可得到上面的递推式。细节留作练习或者仍参考上面的两篇文献。

Legendre 多项式是 $\mathbb{R}[t]$ 的完备正交系。利用分部积分公式证明任意的多项式 $Q(x)$ 如果与 $P_n(x)$ 正交，则 $Q(x) = 0$ 即可。细节留作练习。

下面我们考虑 $\mathbb{R}[t]$ 在其它内积下的正交系，它们被称为加权正交多项式。

1. 第一类 Chebyshev 多项式：内积 $< f, g >= \int_{-1}^{1} f(t)g(t)\dfrac{1}{\sqrt{1-t^2}}\,\mathrm{d}t$。则此基底下的完备正交系是

$$T_n(t) = \cos(n \arccos t),\ \forall n \in \mathbb{N}.$$

   我们有

$$\int_{-1}^{1} T_i(t)T_j(t)\frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t = \begin{cases} \pi, i = j = 0 \\ \dfrac{\pi}{2}, i = j \neq 0 \\ 0, i \neq j \end{cases}$$

   可以验证 $T_n(t)$ 是算子 $\mathcal{T} = (t^2-1)\dfrac{\mathrm{d}^2}{\mathrm{d}t^2} - t\dfrac{\mathrm{d}}{\mathrm{d}t}$ 的特征向量，并且 $\mathcal{T}T_n(t) = n^2T_n(t)$。

2. 第二类 Chebyshev 多项式：内积 $< f, g >= \int_{-1}^{1} f(t)g(t)\sqrt{1-t^2}\,\mathrm{d}t$。则此基底下的完备正交系是

$$u_n(\cos t) = \frac{\sin(n+1)t}{\sin t},\ \forall n \in \mathbb{N}.$$

我们有

$$\int_{-1}^{1} u_i(t)u_j(t)\sqrt{1-t^2}\,\mathrm{d}t = \frac{\pi}{2}\delta_{ij}.$$

可以验证 $u_n(t)$ 满足 $(1-t^2)\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}u_n(t)-3t\dfrac{\mathrm{d}}{\mathrm{d}t}u_n(t)+n(n+2)u_n(t)=0$，并且 $T_n(t)=u_n(t)-tu_{n-1}(t)$。

3. Hermite 多项式：内积 $< f,g >= \int_{-\infty}^{+\infty} f(t)g(t)e^{-t^2}\,\mathrm{d}t$。则此基底下的完备正交系 $\{H_n(t)\}_{n=1}^{\infty}$ 为

$$H_n(t)=(-1)^n e^{t^2}\frac{\mathrm{d}^n}{\mathrm{d}t^n}e^{-t^2},$$

它满足：$H_0(t)=1, H_1(t)=2t, H_{n+1}(t)=2tH_n(t)-\dfrac{\mathrm{d}}{\mathrm{d}t}H_n(t)$，并且

$$\int_{-\infty}^{\infty} H_i(t)H_j(t)e^{-t^2}\,\mathrm{d}t = 2^j j!\sqrt{\pi}\delta_{ij}$$

可以验证 $H_n(t)$ 是算子 $\mathcal{S}=\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}-2t\dfrac{\mathrm{d}}{\mathrm{d}t}$ 的特征向量，并且 $\mathcal{S}H_n(t)=-2nH_n(t)$，即 $H_n(t)$ 满足微分方程 $u''-2tu'+2nu=0$。

有关正交多项式的更多性质，读者可以自行阅读上面的两篇文献。

到此为止，我们已经完成了关于内积空间上以及其上的线性型和线性算子的讨论。在本讲义的最后一章，我们将学习张量这一数学语言，从而将线性型和线性算子的理论整合到一起。在此之前，我们将先应用前面所学的知识，代数化地“构建”我们熟悉的几何空间，这就是我们接下来讨论的解析几何的内容。

3.10　习题

欧几里得空间

1. 在所有次数 < 3 的多项式空间 P_3 中, 对于由公式 $< f, g >= \int_{-1}^{1} f(t)g(t)\mathrm{d}t$ 给定的纯量乘积, 向量 1 和向量 t 是正交的. 找出: (1) 子空间 $\langle 1, t\rangle^{\perp}$; (2) P_3 的一个标准正交基底.

2. 设 $(V, < *, * >)$ 是 3 维欧几里得向量空间, 对任意 $\mathbf{x} \in V, \|\mathbf{x}\|^2 =< \mathbf{x}, \mathbf{x} >= 3x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2 - 2x_1x_3 + 2x_2x_3$ (验证, 这个二次型是正定的), 试找出:
   (1) 向量 $\mathbf{x} = (1, 1, 1)^t, \mathbf{y} = (2, 2, 1)^t$ 之间的夹角 α;
   (2) 所有与 $\mathbf{x}$ 正交的向量.

3. 运用 Gram-Schmidt 正交化过程证明, 任意一个可逆矩阵 $A = \left(a_{ij}\right) \in M_n(\mathbb{R})$ 都可以写成乘积 $A = BC$ 的形式, 其中 B 是正交矩阵, C 是上三角形矩阵, 且 $\pm \det A = \det C$.

4. 命 $\mathfrak{o}(n)$ 为 n 阶斜对称实方阵全体形成的空间. Cayley 变换

$$K = (E - A)^{-1}(E + A), \quad A = (E - K)^{-1}(E + K), \tag{3.10.1}$$

   建立了正交群 $O(n)$ 和向量空间 $o(n)$ 之间一个很好的联系. 证明: 如果 $A \in O(n), \det(E - A) \neq 0$, 那么 $K \in \mathfrak{o}(n)$. 反之, 如果 $K \in \mathfrak{o}(n)$, 那么上式定义的矩阵 A 是正交矩阵, 且 $1 \notin \operatorname{Spec} A$.

   容易看出正交群 $O(n)$ 是 n^2 维向量空间 $M_n(\mathbb{R})$ 中的子集, 由 $n(n+1)/2$ 个方程定义. 所以, $O(n)$ 应该是具有维数 $n^2 - n(n+1)/2 = n(n-1)/2 = \dim o(n)$ 的一个 “代数流形” (代数簇). 在 $O(n)$ 中满足方程 $\det(E - A) = 0$ 的矩阵 A 形成 $O(n)$ 的一个 “超曲面” S. Cayley 变换建立了 $O(n) \setminus S$ 和 $\mathfrak{o}(n)$ 之间的联系.

   可以提出另一个 Cayley 变换:

$$K = (E + A)^{-1}(E - A), \quad A = (E + K)^{-1}(E - K). \tag{3.10.2}$$

   这时, 需要从 $O(n)$ 中挖去方程 $\det(E + A) = 0$ 确定的 “超曲面” T, 即, 挖去特征值中含 -1 的那些正交矩阵。证明: 变换 (3.10.2) 建立了 $O(n) \setminus T$ 和 $\mathfrak{o}(n)$ 之间的联系.

5. 证明: 在凯莱变换 (3.10.1) 或 (3.10.2) 中得到的正交矩阵的行列式均为 1 (用 SO(n) 表示所有这样的矩阵的集合).

6. 设 K 是 n 阶斜对称实方阵, λ 是其非零的纯虚特征根, $\mathbf{x}$ 是 λ 的特征向量: $K\mathbf{x} = \lambda\mathbf{x}$. 记 $\mathbf{x} = \mathbf{y} + i\mathbf{z}$, 其中 $\mathbf{y}, \mathbf{z}$ 都是列向量欧氏空间 $\mathbb{R}^n$ 中的元素. 证明: $\mathbf{y}$ 和 $\mathbf{z}$ 正交且长度相等.

7. 证明, $n \times n$ 阶正交矩阵 A 的特征多项式 $\chi_A(t)$ 具有性质:

$$t^n\chi_A(1/t) = \pm\chi_A(t).$$

8. 设 $A = (\mathbf{A}_1, \cdots, \mathbf{A}_n)^t$ 是一个由相互正交的行排成的矩阵, 证明

$$|\det A| = \|\mathbf{A}_1\| \cdot \|\mathbf{A}_2\| \cdots \|\mathbf{A}_n\|.$$

   (向量在 $\mathbb{R}^n$ 的标准范数).

9. 用 Gram-Schmidt 正交化再次证明 Hadamard 不等式：设 $X = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^t$ 是 $M_n(\mathbb{R})$ 中的任意一个矩阵, 则

$$|\det X| \leqslant \|\mathbf{X}_1\| \cdot \|\mathbf{X}_2\| \cdots \|\mathbf{X}_n\|.$$

埃尔米特向量空间

1. 在向量空间 $M_n(\mathbb{C})$ 上定义半双线性型

$$f(X, Y) = \mathrm{tr}\left(X \cdot \overline{Y}^t\right).$$

(1) 证明：它是正定的埃尔米特型, 从而是 $M_n(\mathbb{C})$ 的一个内积. 记这个内积为 $< *, * >$. 找出如下子空间的正交补;

(2) 迹为零的矩阵全体;

(3) 所有的上三角矩阵.

2. 把下面的向量组扩充为埃尔米特空间的正交基:

(1) $(1, 1-i, 2), (2, -1+3i, 1-i)$;

(2) $(-i, 2, -2-i), (4-i, -i, i)$.

3. 运用正交化方法找出埃氏空间 $\mathbf{C}^4$ 的子空间的一个标准正交基:

(1) $\langle(2, 1, -i, 1), (1, -i, 2, 0), (-i, 0, 1, -i)\rangle$;

(2) $\langle(0, 1-i, 2, 0), (1, 0, 2, i), (1-i, -1, 0, -i)\rangle$.

4. 在 $\mathbb{C}^3$ 中找出子空间 $\mathrm{span}\,\{(0, 1+2i, -i), (1, -1, 2-i)\}$ 的正交补.

5. 复正交群定义为 $\mathbb{C}O(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^t \cdot A = E\}$. 显然, $O(n) \subset \mathbb{C}O(n)$ 且 $\mathrm{SO(n)} \subset \mathrm{S\mathbb{C}O(n)}$ (后二者分别是正交群和复正交群中行列式为 1 的矩阵形成的子群). 问题：是否有 $\mathbb{C}O(n) \subset U(n)$ 和 $S\mathbb{C}O(n) \subset SU(n)$.

6. 对任意实数 $p \geqslant 1$, 在空间 $\mathbb{R}^n$ 上定义

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

证明:

(1) $\| * \|_p$ 是 $\mathbb{R}^n$ 的一个范数;

(2) $\lim_{p\to\infty} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} = \max_{1\leqslant i\leqslant n} |x_i|$.

于是, 范数 $\|x\|_P$ 包括了前面提到的三个范数: $\|x\|_\infty = d_1(x, 0), \|x\|_2 = d(x, 0), \|x\|_1 = d_2(x, 0)$.

在闭区间 $[a, b]$ 上的连续函数全体 $C(a, b)$ 上可以类似定义范数:

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt\right)^{1/p}.$$

同样有

$$\|f\|_\infty = \max_{a\leqslant t\leqslant b} |f(t)|, \quad \|f\|_1 = \int_a^b |f(t)| dt.$$

7. 设 p 和 q 是正数, 满足等式 $\dfrac{1}{p}+\dfrac{1}{q}=1$. 证明: 对标准欧氏空间 $\mathbb{R}^n$ 中的向量 $\mathbf{x},\mathbf{y}$, 有 Hölder 不等式
$$|<\mathbf{x},\mathbf{y}>|\leqslant\|\mathbf{x}\|_p\cdot\|\mathbf{y}\|_q.$$
(当 $p=q=2$ 时, Hölder 不等式就是 CBS 不等式.)

8. 有限维实或复向量空间 V 上的两个范数 $\|\cdot\|_*$ 和 $\|\cdot\|_\circ$ 称为等价的如果存在常数 $c_1,c_2\in\mathbb{R}^+$ $(c_1\leqslant c_2)$ 使得对所有的 $\mathbf{x}\in V$ 有
$$c_1\|\mathbf{x}\|_*\leqslant\|\mathbf{x}\|_\circ\leqslant c_2\|\mathbf{x}\|_*.$$
证明: 有限维实或复向量空间 V 上的任何两个范数都是等价的.

自伴算子

1. 设 V 是内积空间, $\mathcal{A}$ 是 V 上的线性算子. 证明: 如果 $U\subset V$ 是 $\mathcal{A}$ 的不变子空间, 那么 U 的正交补是 $\mathcal{A}^*$ 的不变子空间.

2. 试举反例说明复数域上的**对称**矩阵不一定相似于对角矩阵.

3. 设 $\mu(t)=a_0+a_1t+\cdots+a_mt^m$ 是线性算子 $\mathcal{A}$ 的极小多项式. 证明: $\overline{\mu}(t)=\overline{a}_0+\overline{a}_1t+\cdots+\overline{a}_mt^m$ (小横杆表示复共轭) 是线性算子 $\mathcal{A}^*$ 的极小多项式.

4. 证明: 如果 $\mathcal{A}$ 和 $\mathcal{B}$ 都是向量空间 V 上的自伴随线性算子, 则线性变换
$$\mathcal{A}+\mathcal{B},\quad i(\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A})$$
也是自伴随的.

5. 设 $\mathcal{P}$ 是内积空间 V 上的投影算子. 证明: $\mathcal{A}$ 是自伴随的当且仅当 $\ker(\mathcal{P})\perp\operatorname{im}(\mathcal{P})$.

6. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是域 $\mathbb{K}$ 上的向量空间的算子, 证明: 如果它们交换且可对角化, 那么, 它们可同时对角化, 也就是说, 存在 V 的一个基底, 基底中的向量既是 $\mathcal{A}$ 的特征向量, 也是 $\mathcal{B}$ 的特征向量.

7. 证明: 如果 $\mathcal{A}$ 和 $\mathcal{B}$ 是正定算子且 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$, 那么 $\mathcal{A}\mathcal{B}$ 也是正定的.

8. 假设 A 是可逆斜对称实矩阵. 证明: $-A^2$ 是对称的正定矩阵. 特别, 斜对称实矩阵的非零特征值必为纯虚数.

9. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是两个半正定的自伴随算子, 其中一个是可逆的. 证明: $\mathcal{A}\mathcal{B}$ 的特征值都是非负实数。

10. 假设内积空间上线性算子在某个标准正交基下的矩阵如下:
(1) $\begin{pmatrix}2&1\\1&2\end{pmatrix}$,
(2) $\begin{pmatrix}11&2&-8\\2&2&10\\-8&10&5\end{pmatrix}$,

(3) $\begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}$,

(4) $\begin{pmatrix} 3 & 2+2i \\ 2-2i & 1 \end{pmatrix}$,

(5) $\begin{pmatrix} 3 & -i \\ i & 3 \end{pmatrix}$.

求一个标准正交基使得线性算子在这个基下的矩阵是对角的.

11. 三维欧氏空间 V 上的 $\mathcal{A}$ 在某个标准正交基下的矩阵是

$$\begin{pmatrix} 13 & 14 & 4 \\ 14 & 24 & 18 \\ 4 & 18 & 29 \end{pmatrix}.$$

求 $\mathcal{A}$ 的平方根 $\mathcal{B}$ (即 $\mathcal{B}$ 正定且 $\mathcal{B}^2 = \mathcal{A}$) 在这个基下的矩阵.

12. 求出下列二次型的符合推论3.3.3的典范形式，以及对应的一组标准正交基:
(1) $6x_1^2 + 5x_2^2 + 7x_3^2 - 4x_1x_2 + 4x_1x_3$;
(2) $9x_1^2 + 5x_2^2 + 5x_3^2 + 8x_4^2 + 8x_2x_3 - 4x_2x_4 + 4x_3x_4$;
(3) $5\,|x_1|^2 + i\sqrt{3}x_1\overline{x}_2 - i\sqrt{3}x_2\overline{x}_1 + 6\,|x_2|^2$;
(4) $2\,|x_1|^2 + 2\,|x_2|^2 + 2\,|x_3|^2 - 2ix_1\overline{x}_2 + 2i3x_2\overline{x}_1 + 2ix_2\overline{x}_3 + 2ix_3\overline{x}_2$.

13. 给定欧氏空间上的二次型 $f(x) = \sum_{i=1}^{n} \lambda_i x_i^2$. 证明

$$\max\{|\lambda_1|, \cdots, |\lambda_r|) = \max_{|x|=1} |f(x)|.$$

设 $q(\mathbf{x})$ 是欧氏空间 V 上的二次型. 问: 在单位球面 $S = \{\mathbf{x} \in V \mid \|\mathbf{x}\| = 1\}$ 上哪些点处二次型 q 能达到极大和极小. 更一般地, 单位球面上哪些点是 q 的静止点 (stationary point), 即 q 在该点的各个方向的导数都是 0? 证明下述论断正确:

二次型 $q(\mathbf{x})$ 在单位球面上的静止点恰是二次型 $q(\mathbf{x}) =< \mathcal{F}\mathbf{x}, \mathbf{x} >$ 确定的对称算子 $\mathcal{F}$ 的单位特征向量. 特别, 二次型 $q(\mathbf{x})$ 在单位球面上的极大值就是 $\mathcal{F}$ 的特征值中的最大者, 极小值就是 $\mathcal{F}$ 的特征值中的最小者. 也就是说, 通过保距变换把 $q(\mathbf{x})$ 化成典范式, 那么典范式的系数中的最大者 (最小者) 就是二次型 $q(x)$ 在单位球面上的极大值 (极小值).

14. 如往常, E 是 n 阶单位矩阵, E_{ij} 是在 (i,j) 处的值为 1 , 其他地方的值为 0 的 n 阶方阵. 验证: 矩阵族

$$\mathfrak{S} = \left\{E_{ij} \mid 1 \leqslant i \leqslant [n/2], [n/2] + 2 \leqslant j \leqslant n\right\} \cup \{E\}$$

含有 $\left[n^2/4\right] + 1$ 个元素, 它们线性无关. 而且, 对 $E_{ij}, E_{kl} \in \mathfrak{S}$, 有 $E_{ij}E_{kl} = E_{kl}E_{ij} = 0$. (对实数 a, 记号 $[a]$ 表示比 a 小的整数中的最大者, 即 $[a]$ 是整数, 且 $0 \leqslant a - [a] < 1$. 如 $[-1.2] = -2, [3.1] = 1, [\pm 5] = \pm 5$.)

15. (I. Schur, 1905) 在复代数 $M_n(\mathbb{C})$ 中交换子代数的最大维数是 $\left[n^2/4\right] + 1$. 实际上, 可以

证明更强的结论: 在 $M_n(\mathbb{C})$ 中如果一个向量组中的向量两两交换, 那么这个向量组的秩不超过 $\left[n^2/4\right]+1$. 容易看出, $\mathbb{C}$ 可以用任何域代替.

16. 设 V 是偶数 $2m$ 维欧氏空间, $f(x,y)$ 是 V 上的非退化斜对称双线性型. 证明: 可将 V 分解成两个 m 维子空间的直和 $V = V_1 \oplus V_2$, 并找到一个对称的非退化线性算子 $\mathcal{A}: V \to V$ 使得
$$f(\mathbf{x},\mathbf{y}) =< \mathbf{x}_1, \mathcal{A}\mathbf{y}_2 > - < \mathbf{x}_2, \mathcal{A}\mathbf{y}_1 >,$$
此处 $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2, \quad \mathbf{x}_i, \mathbf{y}_i \in V_i, \quad i = 1, 2.$

酉算子

1. 设 $\mathbf{x}_1, \ldots, \mathbf{x}_k$ 和 $\mathbf{y}_1, \ldots, \mathbf{y}_k$ 是内积空间 V 中的两组向量. 证明: 存在保距算子把 $\mathbf{x}_i$ 映到 $\mathbf{y}_i (i = 1, 2, \cdots, k)$ 的充要条件是 $< \mathbf{x}_i, \mathbf{x}_j >=< y_i, y_j >$, $\forall 1 \leqslant i, j \leqslant k$. 特别地, 如果向量 $\mathbf{x}, \mathbf{y}$ 的范数相等, 则存在保距算子把 $\mathbf{x}$ 映射到 $\mathbf{y}$.

2. 设 $\mathbf{w}$ 是内积空间 V 中的非零向量. 对任何向量 $\mathbf{x} \in V$, 命
$$\mathcal{R}_{\mathbf{w}}(\mathbf{x}) = \mathbf{x} - 2\frac{< \mathbf{x}, \mathbf{w} >}{< \mathbf{w}, \mathbf{w} >}\mathbf{w}.$$
证明:
(1) $\mathcal{R}_{\mathbf{w}}(\mathbf{w}) = -\mathbf{w}$;
(2) $\mathcal{R}_{\mathbf{w}}(\mathbf{y}) = \mathbf{y}$, 如果 $\mathbf{y} \in \operatorname{span}\{\mathbf{w}\}^{\perp}$;
(3) $\mathcal{R}_{\mathbf{w}}$ 是保距算子;
(4) 对 V 中线性无关的向量 $\mathbf{x}, \mathbf{y}$, 存在向量 $\mathbf{v}$ 使得 $\mathcal{R}_{\mathbf{v}}(\mathbf{x}) = \frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}y$ 当且仅当 $< \mathbf{x}, \mathbf{y} >$ 是实数.

3. 证明: 如果内积空间 V 上的线性算子 $\mathcal{A}$ 具有下列三个性质中的任何两个:
(a) 自伴随; (b) 保距; (c) 对合变换 (即 $\mathcal{A}^2 = \mathcal{E}$),
则它也具有第三个性质. 求出具有这三个性质的变换全体.

4. 求证: 酉矩阵 A 能表成换位子 (commutator) $A = XYX^{-1}Y^{-1}$ (其中 X, Y 也是酉矩阵) 的形式, 当且仅当 $\det(A) = 1$.

5. 假设保距算子在某个标准正交基下的矩阵如下, 找出这个算子的一个典范基和算子在这个基下的矩阵:
(1) $\frac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix}$, (2) $\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$,
(3) $\frac{1}{\sqrt{3}}\begin{pmatrix} 1+i & 1 \\ -1 & 1-i \end{pmatrix}$, (4) $\frac{1}{9}\begin{pmatrix} 4+3i & 4i & -6-2i \\ -4i & 4-3i & -2-6i \\ 6+2i & -2-6i & 1 \end{pmatrix}$.

6. (1) 设 $\mathcal{A}$ 是酉算子. 证明: 如果 $\mathcal{A} - \mathcal{E}$ 是可逆的, 那么 $i(\mathcal{A} - \mathcal{E})^{-1}(\mathcal{A} + \mathcal{E})$ 是自伴随的.
(2) 如果酉空间上的线性算子 $\mathcal{A}$ 是自伴随的, 那么 $(\mathcal{A} - i\mathcal{E})^{-1}(\mathcal{A} + i\mathcal{E})$ 是酉算子.

7. 已经知道, 带着内积 $< X, Y >= \operatorname{tr} < X \cdot \overline{Y}^t >$ 下, $M_n(\mathbb{C})$ 成为埃尔米特空间. 证明:
(1) $M_n(\mathbb{C})$ 中的酉矩阵的长度是 $\sqrt{n}$;

(2) $M_n(\mathbb{C})$ 上的算子 $X \to AX$ 的伴随算子是 $X \to \overline{A}^t X$;

(3) 算子 $X \mapsto AX$ 是酉算子当且仅当 A 是酉矩阵.

8. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是正定自伴随算子, $\mathcal{C}$ 是保距算子. 证明: 如果 $\mathcal{A} = \mathcal{B}\mathcal{C}$, 那么 $\mathcal{C} = \mathcal{E}$.

9. 假设算子 $\mathcal{A}$ 在某个标准正交基下的矩阵如下, 找出 $\mathcal{A}$ 的极化分解:

(1) $\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$, (2) $\begin{pmatrix} 4 & -2 & 2 \\ 4 & 4 & -1 \\ -2 & 4 & 2 \end{pmatrix}$.

正规算子

1. 证明: 埃尔米特空间上的算子 $\mathcal{A}$ 是正规的当且仅当存在某个多项式 $p(t)$ 使得 $\mathcal{A}^* = p(\mathcal{A})$.

2. 设 $\mathcal{A}$ 是欧氏空间 V 上的正规算子. 证明: 如果 $\mathcal{A}^2 = -\mathcal{E}$, 那么 $\mathcal{A}^* = -\mathcal{A}$.

3. 设 A 和 B 是交换的 n 阶复矩阵. 如果它们都与对角矩阵相似, 那么存在可逆矩阵 C 使 CAC^{-1} 和 CBC^{-1} 都是对角矩阵.

4. 设 $p(t) = t^2 + at + b$ 是实不可约多项式, $\mathcal{A}$ 是欧氏空间上的正规算子. 证明: 如果 $p(\mathcal{A}) = 0$, 那么 $\mathcal{A}^* = -\mathcal{A} - a\mathcal{E}$.

5. 证明: 埃氏空间上的正规算子 $\mathcal{A}$ 如果与算子 $\mathcal{B}$ 交换, 那么 $\mathcal{A}$ 与 $\mathcal{B}^*$ 交换.

6. 设埃氏空间 V 上的正规算子 $\mathcal{A}$ 和 $\mathcal{B}$ 有相同的特征多项式. 证明: 这两个算子在任何基下的矩阵都是相似的.

7. 证明: 有限维复向量空间上任意一族交换的线性算子都有共同的特征向量.

8. 设 A 是一个 n 阶实雅可比矩阵, 即具有如下形式的矩阵

$$\begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & c_{n-1} & a_n \end{pmatrix}, \quad b_i c_i > 0, \quad i = 1, \cdots, n-1.$$

证明: A 的特征根都是实数且重数都是 1 . (在分析数学中, m 个 n 元可微函数 $f_1(x_1, \cdots, x_n), \cdots, f_n(x_1, \cdots, x_n)$ 的雅可比矩阵是 $m \times n$ 矩阵 $\left(\frac{\partial f_i}{\partial x_j}\right)$.)

最小二乘法, 正交多项式

1. 对下面给定的 A 和 B, 求线性方程组 $AX = B$ 的最小二乘解.

(1) $A = \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$.

(2) $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 5 \\ 6 \\ 6 \end{pmatrix}$.

2. 证明勒让德多项式满足递归公式:

$$P_{n+1}(t) = \frac{2n+1}{n+1} t P_n(t) - \frac{n}{n+1} P_{n-1}(t).$$

3. 证明如下对勒让德多项式的积分很有用的公式:

$$(2n+1)P_n(t) = \frac{d}{dt}\left[P_{n+1}(t) - P_{n-1}(t)\right].$$

4. 利用归纳法和递归公式证明

$$P_n(t) = \sum_{k=0}^{n} (-1)^n \begin{pmatrix} n \\ k \end{pmatrix}^2 \left(\frac{1+t}{2}\right)^{n-k} \left(\frac{1-t}{2}\right)^k.$$

5. 证明: $f(x,t) = \dfrac{1}{\sqrt{1-2xt+x^2}}$ 是勒让德多项式的生成函数:

$$f(x,t) = \sum_{k=0}^{\infty} P_k(t) x^k.$$

6. 设 $\{f_n(t)\}_{n=1}^{\infty}$ 是闭区间 $[a,b]$ 上的正交多项式系, 求证 $f_n(t)$ 的所有根都是落在开区间 (a,b) 内的实根, 并且重数都是 1.

7. 证明 Hermite 多项式满足递归公式

$$H_{n+1}(t) = 2t \cdot H_n(t) - 2n \cdot H_{n-1}(t).$$

第四章　解析几何 (I)：仿射空间

现在我们开始应用前三章的知识讨论解析几何的内容。我们很容易看到，向量空间与不带距离的“真实空间”类似，而欧氏空间则与带距离的“真实空间”类似。然而，向量空间中或欧氏空间中的向量或子空间都要求经过原点 $\mathbf{0}$，而我们在现实中需要不经过原点的向量和平面。因此，我们需要利用向量空间或欧氏空间构造出一个更贴近现实的空间。

4.1　仿射空间

仿射空间的实质与我们中学中学过的建坐标系法是相同的，即以“点”为元素，两点确定一个向量，所有向量构成一个向量空间，取定原点后向量空间的基底就是我们熟悉的坐标系。在仿射空间中，我们主要关注点、线、面之间的位置关系。

定义 4.1.1. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbb{A}$ 是一个非空集合，$\mathbb{A}$ 中的元素称为点 (point)，如果存在满足以下条件的加法映射 $+ : \mathbb{A} \times V \to \mathbb{A}$，$(\dot{p}, \mathbf{x}) \mapsto \dot{p} + \mathbf{x} \in \mathbb{A}$：

(1) $\forall \mathbf{x}, \mathbf{y} \in V$ 及 $\dot{p} \in \mathbb{A}$，有 $(\dot{p} + \mathbf{x}) + \mathbf{y} = \dot{p} + (\mathbf{x} + \mathbf{y})$；

(2) $\forall \dot{p} \in \mathbb{A}$，$\dot{p} + \mathbf{0} = \dot{p}$；

(3) $\forall \dot{p}, \dot{q} \in \mathbb{A}$，存在唯一的 $\mathbf{x} \in V$ 使得 $\dot{p} + \mathbf{x} = \dot{q}$(我们记为 $\mathbf{x} = \overline{pq}$ 或 $\mathbf{x} = \dot{q} - \dot{p}$)。

则称 $\mathbb{A}$ 是一个与 V 相伴的仿射空间 (affine space)。我们定义 $\mathbb{A}$ 的维数 $\dim(\mathbb{A}) = \dim(V) = n$。如果 $\mathbb{K}$ 是 $\mathbb{R}$(或 $\mathbb{C}$)，则我们称 $\mathbb{A}$ 是实 (或复) 仿射空间。

例 4.1.1. (1) 向量空间 V 本身可以视作与 V 相伴的仿射空间，其加法映射就是 V 上的加法。因此，向量空间中的元素也可以视作点。

(2) 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，U 是 V 的子空间。取定 $\mathbf{v}_0 \in V$，以 V 上的加法作为加法映射，则容易验证陪集 $\mathbb{A} = \mathbf{v}_0 + U$ 是与 U 相伴的仿射空间。我们称 $\mathbb{A}$ 是**仿射空间** V 中的一个仿射线性流形或者仿射子空间，U 是这个仿射子空间的方向。

定义 4.1.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基。令 $\mathbb{A}$ 是与 V 相伴的仿射空间，取定点 $\dot{o} \in \mathbb{A}$，则我们称 $(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$ 构成了 $\mathbb{A}$ 的一个仿射坐标系 (或称仿射标架，affine coordinate frame)，其中 $\dot{o}$ 称为仿射坐标系的原点。

我们有：

命题 4.1.1. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$ 是 $\mathbb{A}$ 的一个仿射坐标系，则 $\forall \dot{p} \in \mathbb{A}$，存在唯一一组 $(x_1, \ldots, x_n)^t \in \mathbb{K}^n$ 使得

$$\dot{p} = \dot{o} + \sum_{i=1}^{n} x_i \mathbf{e}_i.$$

我们称 $(x_1, \ldots, x_n)^t$ 是 $\dot{p}$ 在此仿射坐标系下的坐标。

证明. 由仿射空间的定义，我们知道 $\overline{op} \in V$，设向量 $\overline{op}$ 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标是 $(x_1, \ldots, x_n)^t$，则 $\overline{op} = \sum_{i=1}^{n} x_i \mathbf{e}_i$，再用一次仿射空间的定义 (定义4.1.1的 (3)) 即得 $\dot{p} = \dot{o} + \sum_{i=1}^{n} x_i \mathbf{e}_i$。 □

有了坐标系和坐标，我们自然要考虑坐标变换的问题。

命题 4.1.2. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$ 和 $(\dot{o}', \mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ 是 $\mathbb{A}$ 的两个仿射坐标系。设 $\dot{p} \in \mathbb{A}$ 在 $(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$ 下的坐标是 $(x_1, \ldots, x_n)^t$，在 $(\dot{o}', \mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ 下的坐标是 $(x'_1, \ldots, x'_n)^t$，点 $\dot{o}'$ 在 $(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$ 下的坐标是 $(b_1, \ldots, b_n)^t$，而且 $(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot A$，则

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{4.1.1}$$

证明. 利用定义4.1.1的 (1) 和 (3) 我们有：

$$\dot{o} + \overline{op} = \dot{p}$$
$$\dot{o} + \overline{oo'} + \overline{o'p} = \dot{p}$$

于是由两点间向量的唯一性，$\overline{op} = \overline{oo'} + \overline{o'p}$。由坐标的定义我们知道：

$$\overline{op} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\overline{oo'} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

$$\overline{o'p} = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n) \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot A \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

将以上各式代入 $\overline{op} = \overline{oo'} + \overline{o'p}$ 即得

$$(\mathbf{e}_1, \ldots, \mathbf{e}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot A \cdot \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

由于 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的基，一个向量在一组基底下的坐标是唯一的，故式 (4.1.1) 成立。 □

接下来我们研究的思路和研究向量空间类似，我们将依次考虑仿射空间的子空间、仿射组合和仿射包络、仿射相关 (或无关)，以及仿射映射和仿射变换。

定义 4.1.3. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$U \subset V$ 是 V 的 m 维子空间。任意固定 $\dot{p} \in \mathbb{A}$，则集合

$$\Pi = \dot{p} + U = \{\dot{p} + \mathbf{u} \mid \forall \mathbf{u} \in U\}$$

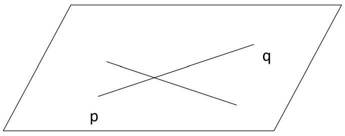
被称为 $\mathbb{A}$ 的一个 m 维仿射子空间或者平面 (flat)，其中 U 称为 Π 的方向子空间。Π 也称为过点 $\dot{p}$ 的以 U 为方向的平面。自然，0 维平面就是点，我们把 1 维的平面也称为直线 (line)，$n-1$ 维的平面称为超平面 (hyperplane)。

我们很容易验证：(1) 如果 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，则 Π 是以 U 为方向的平面 $\iff$ Π 本身也是与 U 相伴的仿射空间 (留作练习)。
(2) 仿射空间中两个平面的交集或者是 $\emptyset$，或者仍是平面 (因为向量空间的子空间之交仍是子空间，细节留作练习)。
(3) 如果 Π 是仿射空间 $\mathbb{A}$ 中的平面，其方向子空间为 U，那么 $\forall \dot{q} \in \Pi$，都有 $\Pi = \dot{q} + U$。这是因为 $\dot{q} + U$ 也是平面，之后验证互相包含即可。

容易看出，仿射空间中的直线与现实中的直线已经很类似了：任取两个不同的点 $\dot{p}, \dot{q} \in \mathbb{A}$，我们可以唯一地确定一条直线：$L_{pq} = \{\dot{r} \in \mathbb{A} \mid \dot{r} = \dot{p} + \lambda\overline{pq}, \forall\lambda \in \mathbb{K}\}$。实际上我们有以下结论：

命题 4.1.3. 设 $\mathbb{A}$ 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) \neq 2)$ 上与 V 相伴的 n 维仿射空间，则子集 $\Pi \subset \mathbb{A}$ 是平面 $\iff$ 任取 $\dot{p}, \dot{q} \in \Pi$，直线 $L_{pq} \subset \Pi$。

证明留作思考。这个命题的几何直观是显然的。



下面我们定义仿射组合。首先给出一个辅助定义：设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，向量组 $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset V$，$\lambda_1, \ldots, \lambda_m \in \mathbb{K}$。如果 $\sum_{i=1}^m \lambda_i = 1$，则我们称线性组合 $\sum_{i=1}^m \lambda_i \mathbf{v}_i$ 是向量 $\mathbf{v}_1, \ldots, \mathbf{v}_m$ 的仿射组合。现在我们需要把仿射组合推广到仿射空间上，为此我们需要以下定理：

定理 4.1.1. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) \neq 2)$ 上的 n 维向量空间，X 是 V 的非空子集，则以下条件等价：

(1) 任取 $\mathbf{x}, \mathbf{y} \in X$，则 $\mathbf{x}, \mathbf{y}$ 的任意仿射组合也在 X 中；

(2) 任取 $\mathbf{x}_1, \ldots, \mathbf{x}_m \in X$，则 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 的任意仿射组合也在 X 中；

(3) 将 V 视作与 V 本身相伴的仿射空间，X 是 V 中的平面。

证明. (1)$\Longrightarrow$(2)：对 m 用数学归纳法。$m = 2$ 显然成立，如果结论对 $m-1$ 成立，则对 m 的情形，不妨设 $\lambda_1\mathbf{x}_1 + \cdots + \lambda_m\mathbf{x}_m$ 是 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 的一个仿射组合 (即 $\sum_{i=1}^m \lambda_i = 1$) 并且 $\lambda_1 \neq 0$，由于 $m-1$ 的情形成立 (归纳假设)，即 $\mathbf{x}_2, \ldots, \mathbf{x}_m$ 的一个仿射组合

$$\frac{\lambda_2}{1-\lambda_1}\mathbf{x}_2 + \cdots + \frac{\lambda_m}{1-\lambda_1}\mathbf{x}_m \in X$$

因此由 (1) 可知

$$\sum_{i=1}^m \lambda_i\mathbf{x}_i = \lambda_1 \underbrace{\mathbf{x}_1}_{\in X} + (1-\lambda_1)\underbrace{\left(\frac{\lambda_2}{1-\lambda_1}\mathbf{x}_2 + \cdots + \frac{\lambda_m}{1-\lambda_1}\mathbf{x}_m\right)}_{\in X} \in X$$

因此 m 时 (1)$\Longrightarrow$(2) 也成立。于是 (1)$\Longrightarrow$(2) 成立。
(2)$\Longrightarrow$(3)：由于 X 非空，因此一定存在 $\mathbf{v} \in X$，要证 X 是仿射空间 V 中的平面 (即仿射子空间)，我们只需证明存在向量空间 V 的线性子空间 U 使得 $X = \mathbf{v} + U$ 即可。

实际上，令 $U = \{\mathbf{x}-\mathbf{v} \mid \forall \mathbf{x} \in X\}$，则 $X = \mathbf{v}+U$ 一定成立，下面我们只需说明这样构造出的 U 是 V 的线性子空间。首先，容易看出 $\mathbf{0} = \mathbf{v}-\mathbf{v} \in U$；其次，任取 $\mathbf{u}_1 = \mathbf{x}_1-\mathbf{v}, \mathbf{u}_2 = \mathbf{x}_2-\mathbf{v} \in U$，即 $\mathbf{x}_1, \mathbf{x}_2 \in X$ 以及 $a_1, a_2 \in \mathbb{K}$，我们有

$$\begin{aligned} a_1\mathbf{u}_1 + a_2\mathbf{u}_2 &= a_1(\mathbf{x}_1 - \mathbf{v}) + a_2(\mathbf{x}_2 - \mathbf{v}) \\ &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{v} - \mathbf{v} \end{aligned}$$

由于 $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{v}$ 是 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}$ 的仿射组合，而由 (2)，X 对仿射组合封闭，，所以 $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{v} \in X$，即 $a_1\mathbf{u}_1 + a_2\mathbf{u}_2 = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + (1 - a_1 - a_2)\mathbf{v} - \mathbf{v}$ 是 $\mathbf{x} - \mathbf{v}, \mathbf{x} \in X$ 的形式，$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 \in U$。此即 U 是 V 的线性子空间。因此 X 是仿射空间 V 中的平面。
(3)$\Longrightarrow$(1)：如果 X 是仿射空间 V 中的平面，则 X 一定可以写成 $\mathbf{v} + U, \mathbf{v} \in X$ 的形式，其中 U 是向量空间 V 的线性子空间。那么，任取 $\mathbf{x}, \mathbf{y} \in X$ 及 $\lambda \in \mathbb{K}$，我们有

$$\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} = \mathbf{v} + \lambda(\mathbf{x} - \mathbf{v}) + (1 - \lambda)(\mathbf{y} - \mathbf{v})$$

由 U 是线性子空间以及 $\mathbf{x} - \mathbf{v}, \mathbf{y} - \mathbf{v} \in U$ 可知 $\lambda(\mathbf{x} - \mathbf{v}) + (1 - \lambda)(\mathbf{y} - \mathbf{v}) \in U$，因此 $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ 也具有 $\mathbf{v} + \mathbf{u}, \mathbf{u} \in U$ 的形式，即 $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in X$。这样我们就完成了证明。

□

下面转入点的仿射组合。

命题 4.1.4. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，任意固定 $k+1$ 个点 $\dot{p}_0, \ldots, \dot{p}_k \in \mathbb{A}$，另外固定 $k+1$ 个数 $\alpha_0, \ldots, \alpha_k \in \mathbb{K}$，$\sum_{i=0}^{k} \alpha_i = 1$，则任取点 $\dot{p} \in \mathbb{A}$，点 $\dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i}$ 的位置都与“基点” $\dot{p}$ 无关。

证明. 我们只需证明：任意另取 $\dot{q} \in \mathbb{A}$，有 $\dot{q} + \sum_{i=0}^{k} \alpha_i \overline{qp_i} = \dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i}$ 即可。设 $\dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i} + \mathbf{v} = \dot{q} + \sum_{i=0}^{k} \alpha_i \overline{qp_i}$，我们来证明 $\mathbf{v} = \mathbf{0}$。注意到

$$\begin{aligned} \mathbf{v} &= \overline{pq} + \sum_{i=0}^{k} \alpha_i \overline{qp_i} - \sum_{i=0}^{k} \alpha_i \overline{pp_i} \\ &= \overline{pq} + \sum_{i=0}^{k} \alpha_i (\overline{qp_i} - \overline{pp_i}) \\ &= \overline{pq} - \sum_{i=0}^{k} \alpha_i \overline{pq} \\ &= \mathbf{0} \end{aligned}$$

此即 $\dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i}$ 的位置与 $\dot{p}$ 无关。 □

于是我们可以定义仿射空间中点的仿射组合如下：

定义 4.1.4. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$k+1$ 个点 $\dot{p}_0, \ldots, \dot{p}_k \in \mathbb{A}$ 的仿射组合 (affine combination) 是指点 $\dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i}$，其中 $\sum_{i=0}^{k} \alpha_i = 1$，$\dot{p} \in \mathbb{A}$ 是任意的点。由上面的推论可知仿射组合与 $\dot{p}$ 的选取无关，因此我们把点 $\dot{p} + \sum_{i=0}^{k} \alpha_i \overline{pp_i}$ 记作 $\alpha_0\dot{p}_0 + \cdots + \alpha_k\dot{p}_k$ 或 $\sum_{i=0}^{k} \alpha_i \dot{p}_i$。

我们有：

推论 4.1.1. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，X 是 $\mathbb{A}$ 的非空子集，则以下条件等价：

(1) 任取 $\dot{p}, \dot{q} \in X$，则 $\dot{p}, \dot{q}$ 的任意仿射组合也在 X 中；

(2) 任取 $\dot{p}_1, \ldots, \dot{p}_m \in X$，则 $\dot{p}_1, \ldots, \dot{p}_m$ 的任意仿射组合也在 X 中；

(3) X 是 $\mathbb{A}$ 中的平面。

证明. (1)$\Longrightarrow$(2)：我们已经在命题4.1.4中说明了点的仿射组合是良定义的，因此此处的证明与定理4.1.1的证明相同。

(2)$\Longrightarrow$(3)：由于 X 非空，因此一定存在 $\dot{p} \in X$，要证 X 是仿射空间 V 中的平面 (即仿射子空间)，我们只需证明存在向量空间 V 的线性子空间 U 使得 $X = \dot{p} + U$ 即可。实际上，令 $U = \{\mathbf{u} = \overline{pq} \in V \mid \forall \dot{q} \in X\}$，则 $X = \dot{p} + U$ 一定成立，于是我们只需说明这样构造出的 U 是 V 的线性子空间。而这个证明与定理4.1.1的证明相同。(3)$\Longrightarrow$(1)：如果 X 是仿射空间 $\mathbb{A}$ 中的平面，则 X 一定可以写成 $\dot{p} + U, \dot{p} \in X$ 的形式，其中 U 是向量空间 V 的线性子空间。那么，任取 $\dot{p}_1, \dot{p}_2 \in X$ 及 $\lambda \in \mathbb{K}$，我们有

$$\lambda \dot{p}_1 + (1-\lambda)\dot{p}_2 = \dot{p} + \lambda\overline{pp_1} + (1-\lambda)\overline{pp_2}$$

由 U 是线性子空间以及 $\overline{pp_1}, \overline{pp_2} \in U$ 可知 $\lambda\overline{pp_1} + (1-\lambda)\overline{pp_2} \in U$，因此 $\lambda\dot{p}_1 + (1-\lambda)\dot{p}_2$ 也具有 $\dot{p} + \mathbf{u}, \mathbf{u} \in U$ 的形式，即 $\lambda\dot{p}_1 + (1-\lambda)\dot{p}_2 \in X$。可以看到这个证明与定理4.1.1的对应部分的证明类似。 □

于是我们可以定义点集的仿射包络如下：

定义 4.1.5. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，M 是 $\mathbb{A}$ 的非空子集，我们称

$$\mathbb{A}(M) = \{\sum_{i=0}^{k} \alpha_i \dot{p}_i \in \mathbb{A} \mid \forall k \in \mathbb{N},\ \dot{p}_0, \ldots, \dot{p}_k \in M,\ \alpha_0, \ldots, \alpha_k \in \mathbb{K}\text{满足} \sum_{i=0}^{k} \alpha_i = 1\}.$$

是 M 的仿射包络 (affine span)，也记作 aff M。

例 4.1.2. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间.

(i) 很容易看出一个点 $\dot{p} \in \mathbb{A}$ 的仿射包络 $\mathbb{A}(\{\dot{p}\}) = \{\dot{p}\}$。、

(ii) 两个点 $\dot{p}, \dot{q} \in \mathbb{A}$ 的仿射包络是过这两点的直线 L_{pq}：任取 $\dot{r} \in \mathbb{A}(\{\dot{p}, \dot{q}\})$，则存在 $\alpha \in \mathbb{K}$ 使得 $\dot{r} = \alpha\dot{p} + (1-\alpha)\dot{q}$，按仿射组合的定义我们有

$$\begin{aligned} \alpha\dot{p} + (1-\alpha)\dot{q} &= \dot{p} + \alpha\overline{pp} + (1-\alpha)\overline{pq} \qquad (\text{把“基点”就取成}\dot{p}) \\ &= \dot{p} + (1-\alpha)\overline{pq} \in L_{pq} \end{aligned}$$

即 $\mathbb{A}(\{\dot{p}, \dot{q}\}) \subset L_{pq}$。反之，很容易验证 $\dot{r} = \dot{p} + \lambda\overline{pq} \in L_{pq}$ 一定可以写成仿射组合 $(1-\lambda)\dot{p} + \lambda\dot{q}$ 的形式，即 $\mathbb{A}(\{\dot{p}, \dot{q}\}) \supset L_{pq}$。综上所述，$\mathbb{A}(\{\dot{p}, \dot{q}\}) = L_{pq}$。

定理 4.1.2. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，M 是 $\mathbb{A}$ 的非空子集，则仿射包络 $\mathbb{A}(M)$ 一定是 $\mathbb{A}$ 中的平面。

证明. 我们利用推论4.1.1的 (1)$\Longleftrightarrow$(3) 来证明这个定理。任取 $\dot{p} = \sum_{i=0}^{l} \alpha_i \dot{p}_i$，$\dot{q} = \sum_{j=0}^{m} \beta_j \dot{q}_j \in \mathbb{A}(M)$，即 $\sum_{i=0}^{l} \alpha_i = \sum_{j=0}^{m} \beta_j = 1$ 并且所有的 $\dot{p}_i$ 和 $\dot{q}_j$ 都是 M 中的点，则任取 $\lambda \in \mathbb{K}$，我们有

$$\lambda\dot{p} + (1-\lambda)\dot{q} = \sum_{i=0}^{l} \lambda\alpha_i \dot{p}_i + \sum_{j=0}^{m} (1-\lambda)\beta_j \dot{q}_j$$

容易计算出 $\sum_{i=0}^{l} \lambda + \sum_{j=0}^{m} (1-\lambda)\beta_j = 1$，因此 $\lambda\dot{p} + (1-\lambda)\dot{q}$ 仍然是 M 中的点的仿射组合，故 $\lambda\dot{p} + (1-\lambda)\dot{q} \in \mathbb{A}(M)$。由推论4.1.1的 (1)$\Longleftrightarrow$(3) 即可得到 $\mathbb{A}(M)$ 是 $\mathbb{A}$ 中的平面。 □

我们很容易证明: 如果 M 是有限集并且 M 中有 $k+1$ 个不同的点, 那么 $\dim(\mathbb{A}(M)) \leqslant k$。将仿射组合改写成点加上向量的线性组合后利用线性包络的维数不超过基向量的个数即可证明，细节留作练习。

有了仿射包络，下面我们就可以定义仿射空间中的仿射相关和仿射无关了。

定义 4.1.6. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$n+1$ 个点 $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n \in \mathbb{A}$。如果 $\dim(\mathbb{A}(\{\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n\})) = n$，则我们称点 $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n$ 仿射无关 (affine independent)(或称这些点处于一般位置)；如果 $\dim(\mathbb{A}(\{\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n\})) < n$，则我们称点 $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n$ 仿射相关 (affine dependent)。

例 4.1.3. 在 $\mathbb{R}$ 上与 $\mathbb{R}^2$ 相伴的 2 维仿射空间 $\mathbb{A}$ 中，容易看出 $\dot{p}_0, \dot{p}_1, \dot{p}_2 \in \mathbb{A}$ 仿射相关 $\Longleftrightarrow$ 这三点位于同一条直线上；换言之，在此仿射空间中任意不共线的三点都仿射无关。

于是我们也可以用点的仿射组合来表示点，这就是下面的重心坐标。

定义 4.1.7. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$n+1$ 个点 $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n \in \mathbb{A}$ 仿射无关。则任取 $\dot{p} \in \mathbb{A}$，存在唯一一组 $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$，$\sum_{i=0}^{n} \lambda_i = 1$ 满足 $\dot{p} = \sum_{i=0}^{n} \lambda_i \dot{p}_i$。我们称有序组 $(\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n)$ 是 $\mathbb{A}$ 的一个重心坐标系，$(\lambda_0, \ldots, \lambda_n)^t \in \mathbb{K}^{n+1}$ 是 $\dot{p}$ 在 $(\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n)$ 下的重心坐标 (barycentric coordinate)。

我们很容易验证：$\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n \in \mathbb{A}$ 仿射无关 $\Longleftrightarrow$ 向量组 $\overline{p_0p_1}, \ldots, \overline{p_0p_n}$ 线性无关，留作练习。

例 4.1.4. 将 $\mathbb{R}^2$ 视作与其本身相伴的仿射空间，任取其内的一个非退化的三角形，将这个三角形的三个顶点作为一个重心坐标系，则三角形重心的重心坐标是 $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^t$。

接下来我们考虑仿射空间之间保持“比例”(或者说保持仿射组合) 的映射，即仿射映射。

定义 4.1.8. 设 $\mathbb{A}$ 和 $\mathbb{A}'$ 分别是是域 $\mathbb{K}$ 上与 V 和 V' 相伴的 n, m 维仿射空间，$f : \mathbb{A} \to \mathbb{A}'$ 满足：任取点 $\dot{p}_0, \ldots, \dot{p}_k \in \mathbb{A}$ 的一个仿射组合 $\sum_{i=0}^{k} \lambda_i \dot{p}_i$(即 $\sum_{i=0}^{k} \lambda_i$)，有 $f(\sum_{i=0}^{k} \lambda_i \dot{p}_i) = \sum_{i=0}^{k} \lambda_i f(\dot{p}_i)$ 成立，则我们称 f 是一个仿射映射 (affine map)。

按定义我们很容易验证: 如果 $f : \mathbb{A} \to \mathbb{A}'$, $g : \mathbb{A}' \to \mathbb{A}''$ 都是仿射映射, 则 $g \circ f : \mathbb{A} \to \mathbb{A}''$ 也是仿射映射，过程留作练习。我们之所以说仿射映射是保持“比例”的映射，是因为我们有以下定理:

定理 4.1.3. 设 $\mathbb{A}$ 和 $\mathbb{A}'$ 分别是是域 $\mathbb{K}$ 上与 V 和 V' 相伴的 n, m 维仿射空间，则以下条件等价:

(1) $f : \mathbb{A} \to \mathbb{A}'$ 是仿射映射。

(2) 任取 $\mathbb{A},\mathbb{A}'$ 的仿射坐标系 $(\dot{o},\mathbf{e}_1,\ldots,\mathbf{e}_n)$ 和 $(\dot{o}',\mathbf{e}'_1,\ldots,\mathbf{e}'_m)$，对任意 $\dot{p}\in\mathbb{A}$，$\dot{p}$ 的坐标是 $(x_1,\ldots,x_n)^t$，则 $f(\dot{p})$ 的坐标是

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} + A\cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

其中 $(v_1,\ldots,v_m)^t$ 是 $f(\dot{o})$ 在 $(\dot{o}',\mathbf{e}'_1,\ldots,\mathbf{e}'_m)$ 下的坐标，$A\in\mathbb{K}^{m\times n}$。换言之，固定 $\dot{o}\in\mathbb{A}$，则存在线性映射 $\mathcal{A}\in\mathrm{Hom}(V,V')$ 使得对 $\forall\dot{p}\in\mathbb{A}$，$f(\dot{p})=f(\dot{o})+\mathcal{A}\overline{op}$。我们称 $\mathcal{A}$ 是 f 的线性部分或微分。

(3) 任取 $\dot{p},\dot{q},\dot{r},\dot{s}\in\mathbb{A}$ 满足 $\overline{rs}=\lambda\overline{pq}$，$\lambda\in\mathbb{K}$，则 $\overline{f(\dot{r})f(\dot{s})}=\lambda\cdot\overline{f(\dot{p})f(\dot{q})}$。

证明. (1)$\Longrightarrow$(2)：我们来证明第二个叙述。注意到利用 f 我们可以定义 $V\to V'$ 的映射

$$\mathcal{A}:V\longrightarrow V',\ \overline{op}\longmapsto\overline{f(\dot{o})f(\dot{p})}$$

则 $f(\dot{p})=f(\dot{o})+\overline{f(\dot{o})f(\dot{p})}=f(\dot{o})+\mathcal{A}\overline{op}$ 一定成立。下面我们需要证明 $\mathcal{A}$ 是线性映射：由于 f 保持仿射组合，故 $\forall x_1,x_2\in\mathbb{K}$ 及 $\overline{op_1},\overline{op_2}\in V$，我们有

$$\begin{aligned} f(\dot{o}+x_1\overline{op_1}+x_2\overline{op_2}) &= f((1-x_1-x_2)\dot{o}+x_1\dot{p}_1+x_2\dot{p}_2) \\ &= (1-x_1-x_2)f(\dot{o})+x_1f(\dot{p}_1)+x_2f(\dot{p}_2) \\ &= f(\dot{o})+x_1\overline{f(\dot{o})f(\dot{p}_1)}+x_2\overline{f(\dot{o})f(\dot{p}_2)} \end{aligned}$$

令 $\dot{q}=\dot{o}+x_1\overline{op_1}+x_2\overline{op_2}$，利用上式及 $\mathcal{A}$ 的定义我们有

$$f(\dot{o})+\mathcal{A}\overline{oq}=f(\dot{o})+x_1\overline{f(\dot{o})f(\dot{p}_1)}+x_2\overline{f(\dot{o})f(\dot{p}_2)}=f(\dot{o})+x_1\mathcal{A}\overline{op_1}+x_2\mathcal{A}\overline{op_2}$$

即 $\mathcal{A}(x_1\overline{op_1}+x_2\overline{op_2})=x_1\mathcal{A}\overline{op_1}+x_2\mathcal{A}\overline{op_2}$，$\mathcal{A}$ 是线性映射得证。之后，设 $\mathcal{A}$ 在 V 和 V' 的基底 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_m$ 的矩阵是 $A\in\mathbb{K}^{m\times n}$，则 $\overline{f(\dot{o})f(\dot{p})}$ 在 $\mathbf{e}'_1,\ldots,\mathbf{e}'_m$ 下的坐标正是 $A\cdot(x_1,\ldots,x_n)^t$，即在仿射坐标系 $(f(\dot{o}),\mathbf{e}'_1,\ldots,\mathbf{e}'_m)$ 下 $f(\dot{p})$ 的坐标是 $A\cdot(x_1,\ldots,x_n)^t$，再利用仿射坐标系的坐标变换 (命题4.1.2) 我们就可以完成该方向的证明 (细节是容易的，留作练习)。

(2)$\Longrightarrow$(3)：利用 (2) 的条件我们立刻有：

$$\begin{aligned} \overline{f(\dot{r})f(\dot{s})} &= \mathcal{A}\overline{rs} \\ &= \lambda\mathcal{A}\overline{pq} \\ &= \lambda\overline{f(\dot{p})f(\dot{q})}. \end{aligned}$$

该方向证毕。

(3)$\Longrightarrow$(2)：令 $\mathcal{A}:V\to V'$，$\overline{pq}\mapsto\overline{f(\dot{p})f(\dot{q})}$，则固定 $\dot{o}\in\mathbb{A}$，对任意的 $\dot{p}\in\mathbb{A}$，$\mathcal{A}$ 一定满足：

$$f(\dot{p})=f(\dot{o})+\mathcal{A}\overline{op}.$$

下面我们只需说明 $\mathcal{A}$ 是线性映射即可证明该方向。首先 $\mathcal{A}$ 是良定义的 ($\mathcal{A}$ 不会出现一对多的情形)：如果 $\overline{pq},\overline{rs}\in V$ 并且 $\overline{pq}=\overline{rs}$，那么按照 (3) 的条件，一定有 $\overline{f(\dot{p})f(\dot{q})}=\overline{f(\dot{r})f(\dot{s})}$，即

$\mathcal{A}\overline{pq} = \mathcal{A}\overline{rs}$，此即良定义性。下面分别验证 $\mathcal{A}$ 保持加法和数乘。对 $\forall \overline{pq}, \overline{qr} \in V$(容易看出 V 中的任意两个向量一定可以写成这样首尾相接的形式)，一方面 $\mathcal{A}(\overline{pq}+\overline{qr}) = \mathcal{A}\overline{pr} = \overline{f(\dot{p})f(\dot{r})}$；另一方面 $\mathcal{A}\overline{pq} + \mathcal{A}\overline{qr} = \overline{f(\dot{p})f(\dot{q})} + \overline{f(\dot{q})f(\dot{r})} = \overline{f(\dot{p})f(\dot{r})}$，这说明 $\mathcal{A}(\overline{pq} + \overline{qr}) = \mathcal{A}\overline{pq} + \mathcal{A}\overline{qr}$，即 $\mathcal{A}$ 保持向量的加法。$\mathcal{A}$ 保持数乘是 (3) 的直接结论，因此 $\mathcal{A} \in \mathrm{Hom}(V, V')$。该方向证毕。

(2)$\Longrightarrow$(1)：固定 $\dot{o} \in \mathbb{A}$，设 f 满足：存在线性映射 $\mathcal{A} \in \mathrm{Hom}(V, V')$ 使得对 $\forall \dot{p} \in \mathbb{A}$，$f(\dot{p}) = f(\dot{o}) + \mathcal{A}\overline{op}$，我们来证明 f 保持仿射组合。任取仿射组合 $\sum_{i=0}^{k} x_i \dot{p}_i \in \mathbb{A}$，其中 $\sum_{i=0}^{k} x_i = 1$，则

$$\begin{aligned}
f(\sum_{i=0}^{k} x_i \dot{p}_i) &= f(\dot{o} + \sum_{i=0}^{k} x_i \overline{op_i}) \\
&= f(\dot{o}) + \mathcal{A}(\sum_{i=0}^{k} x_i \overline{op_i}) \\
&= \sum_{i=0}^{k} x_i (f(\dot{o}) + \mathcal{A}\overline{op_i}) \\
&= \sum_{i=0}^{k} x_i f(\dot{p}_i)
\end{aligned}$$

此即 f 保持仿射组合。

综上所述，定理成立。 □

特别地，我们关注 $\mathbb{A} \to \mathbb{A}$ 本身的仿射映射，我们称之为仿射变换 (affine transformation)。利用上面的定理我们立刻可以得到仿射变换的结构。

定义 4.1.9. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$f : \mathbb{A} \to \mathbb{A}$ 是 $\mathbb{A}$ 上的仿射变换，如果存在 $\mathbf{v} \in V$ 使得 $\forall \dot{p} \in \mathbb{A}$，$f(\dot{p}) = \dot{p} + \mathbf{v}$，则我们称 f 是 $\mathbb{A}$ 上的一个平移 (translation)，并记作 $f = T_{\mathbf{v}}$。

推论 4.1.2. 设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，将 V 本身视作与 V 相伴的仿射空间，则 $f : V \to V$ 是 V 上的仿射变换 $\iff$ 存在 $\mathbf{v} \in V$ 使得 $f = T_{\mathbf{v}} \circ \mathcal{A}$，其中 $\mathcal{A} \in \mathrm{End}(V)$。

证明即定理4.1.3的 (1) $\iff$ (2)。

设 $\mathbb{A}$ 和 $\mathbb{A}'$ 分别是是域 $\mathbb{K}$ 上与 V 和 V' 相伴的仿射空间，如果存在双射 $f : \mathbb{A} \to \mathbb{A}'$ 是仿射映射，则我们称 $\mathbb{A}$ 和 $\mathbb{A}'$ 仿射同构，f 是仿射同构映射。很容易证明 $\mathbb{A}$ 和 $\mathbb{A}'$ 仿射同构 $\iff$ $\dim(\mathbb{A}) = \dim(\mathbb{A}')$，留作练习。特别地，任何与 V 相伴的仿射空间 $\mathbb{A}$ 都一定与 V 本身(视作仿射空间) 仿射同构。

推论 4.1.3. (1) 设 $\mathbb{A}$ 和 $\mathbb{A}'$ 分别是域 $\mathbb{K}$ 上与 V 和 V' 相伴的仿射空间，$f : \mathbb{A} \to \mathbb{A}'$ 是仿射映射并且固定 $\dot{o} \in \mathbb{A}$，对 $\forall \dot{p} \in \mathbb{A}$ 有 $f(\dot{p}) = f(\dot{o}) + \mathcal{A}\overline{op}$，$\mathcal{A} \in \mathrm{Hom}(V, V')$。则 f 是仿射同构映射 $\iff$ $\mathcal{A}$ 是 V 到 V' 的线性同构。

(2) 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，则 $\mathbb{A}$ 的所有自同构 (即 $\mathbb{A} \to \mathbb{A}$ 的同构映射) 在映射复合运算下构成一个群。我们把这个群称为仿射变换群，记作 $\mathrm{Aff}(\mathbb{A})$。

证明. (1) 是显然的，因为容易看出 f 是双射 $\iff$ $\mathcal{A}$ 是双射。下面我们证明 (2)。设 f, g 都是 $\mathbb{A} \to \mathbb{A}$ 的自同构，则显然 $f \circ g$ 也是 $\mathbb{A} \to \mathbb{A}$ 的自同构，封闭性成立；结合律即映射复合的结合律；乘法幺元是恒等映射；固定原点 $\dot{o} \in \mathbb{A}$，$f : \mathbb{A} \to \mathbb{A}$，$\dot{p} \mapsto f(\dot{o}) + \mathcal{A}\overline{op}, \mathcal{A} \in \mathrm{End}(V)$ 的逆元是 $g : \mathbb{A} \to \mathbb{A}$，$\dot{q} \mapsto \dot{o} + \mathcal{A}^{-1}\overline{f(\dot{o})q}$(容易验证 $f \circ g = g \circ f = \mathrm{id}$，留作练习)。这样我们就证明了 $\mathbb{A}$ 的所有自同构在映射复合下构成群。 □

下面考虑仿射群的粗略结构。

命题 4.1.5. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，任意固定原点 $\dot{o} \in \mathbb{A}$，则 $\mathrm{Aff}(\mathbb{A}) = \{T_{\mathbf{v}} \circ \mathbf{A} \mid \mathbf{v} \in V, \mathbf{A} \in \mathrm{Aff}(\mathbb{A})\text{并且}\mathcal{A}\dot{o} = \dot{o}\}$，即 $\mathbf{A}$ 可视作 V 上的可逆线性算子 $\mathcal{A}$，$\mathbf{A}\dot{p} = \dot{o} + \mathcal{A}\overline{op}$。考虑 V 上所有可逆线性算子在映射复合下构成的群 $\mathrm{GL}(V)$，令

$$\Phi : \mathrm{Aff}(\mathbb{A}) \to \mathrm{GL}(V),\ f = T_{\mathbf{v}} \circ \mathbf{A} \mapsto \mathcal{A}$$

则 Φ 是满的群同态，并且 $\ker(\Phi) = \{T_{\mathbf{v}} \mid \forall \mathbf{v} \in V\}$(称为平移群) 是 $\mathrm{Aff}(\mathbb{A})$ 的子群 (实际上是正规子群)。

证明. 推论4.1.2以及 $\mathbb{A}$ 与 V 仿射同构保证了 $\mathrm{Aff}(\mathbb{A})$ 中的元素都可以写成 $T_{\mathbf{v}} \circ \mathbf{A}$ 的形式。下面说明 Φ 是满同态。

首先 Φ 是满射，因为任取 $\mathcal{A} \in \mathrm{GL}(V)$ 都可以定义 $\mathbf{A} : \mathbb{A} \to \mathbb{A}$，$\dot{p} \mapsto \dot{o} + \mathcal{A}\overline{op}$，则 $\mathbf{A} \in \mathrm{Aff}(\mathbb{A})$ 并且 $\Phi(\mathbf{A}) = \mathcal{A}$。

其次，设 $f, g \in \mathrm{Aff}(\mathbb{A})$ 并且 $f = T_{\mathbf{v}} \circ \mathbf{A}$，$g = T_{\mathbf{u}} \circ \mathbf{B}$，$\Phi(f) = \mathcal{A}$，$\Phi(g) = \mathcal{B}$，那么：

$$\begin{aligned} f \circ g &= T_{\mathbf{v}} \circ \mathbf{A} \circ T_{\mathbf{u}} \circ \mathbf{B} \\ &= T_{\mathbf{v}} \circ T_{\mathcal{A}\mathbf{u}} \circ \mathbf{A} \circ \mathbf{B} \\ &= T_{\mathbf{v}+\mathcal{A}\mathbf{u}} \circ \mathbf{A} \circ \mathbf{B} \end{aligned}$$

其中第一行到第二行利用了 $\mathbf{A} \circ T_{\mathbf{u}} = T_{\mathcal{A}\mathbf{u}} \circ \mathbf{A}$，这是因为 $\forall \dot{p} \in \mathbb{A}$，都有 $\mathbf{A} \circ T_{\mathbf{u}}(\dot{p}) = \mathbf{A}(\dot{p} + \mathbf{u}) = \dot{o} + \mathcal{A}(\overline{op} + \mathbf{u}) = \dot{o} + \mathcal{A}\overline{op} + \mathcal{A}\mathbf{u} = T_{\mathcal{A}\mathbf{u}} \circ \mathbf{A}(\dot{p})$。于是 $\Phi(f \circ g) = \mathcal{A} \circ \mathcal{B} = \Phi(f) \circ \Phi(g)$，即 Φ 是同态。

最后我们来计算 $\ker(\Phi)$。如果 $\Phi(f) = \mathcal{E}$，那么一定有 $f = T_{\mathbf{v}} \circ \mathrm{id} = T_{\mathbf{v}}$，即 f 是平移。此即 $\ker(\Phi) = \{T_{\mathbf{v}} \mid \forall \mathbf{v} \in V\}$. 由讲义上册定义 4.2.11 下方的说明立刻可知 $\ker(\Phi)$ 是 $\mathrm{Aff}(\mathbb{A})$ 的正规子群。 □

接下来我们考虑另一类特殊的仿射映射：仿射线性函数，并讨论它和非齐次的线性方程组之间的关系。

定义 4.1.10. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$f : \mathbb{A} \to \mathbb{K}$ 是仿射映射 (将 $\mathbb{K}$ 视作与其本身相伴的 1 维仿射空间)，则我们称 f 是一个仿射线性函数。

命题 4.1.6. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$(\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_n)$ 是 $\mathbb{A}$ 的一个重心坐标系，则每个点映到其坐标分量的函数 $f_k : \mathbb{A} \to \mathbb{K}$，$\sum_{i=0}^{n} x_i \dot{p}_i \mapsto x_k$，$k \in \{1, \ldots, n\}$ 是仿射线性函数。

证明. 我们验证每个 f_k 都保持仿射组合。任取 $\dot{q}_j = \sum_{i=0}^{n} x_{ij} \dot{p}_i \in \mathbb{A}$，$j \in \{0, \ldots, m\}$ 及 $\lambda_0, \ldots, \lambda_m \in \mathbb{K}$，$\sum_{j=0}^{m} \lambda_j = 1$，则

$$\sum_{j=0}^{m} \lambda_j \dot{q}_j = \sum_{j=0}^{m} \lambda_j (\sum_{i=0}^{n} x_{ij} \dot{p}_i) = \sum_{i=0}^{n} (\sum_{j=0}^{m} \lambda_j x_{ij}) \dot{p}_i$$

因此 $f_k(\sum_{j=0}^{m} \lambda_j \dot{q}_j) = \sum_{j=0}^{m} \lambda_j x_{kj}$。另一方面，$\sum_{j=0}^{m} \lambda_j f_k(\dot{p}_j) = \sum_{j=0}^{m} \lambda_j x_{kj}$，故 $f_k(\sum_{j=0}^{m} \lambda_j \dot{q}_j) = \sum_{j=0}^{m} \lambda_j f_k(\dot{p}_j)$。这样我们就完成了证明。 □

我们在 1.5 节中研究过抽象的齐次线性方程组，类似地，我们也可以考虑在仿射空间中求解一些仿射线性函数的核的交集，这就是抽象的非齐次线性方程组。

定义 4.1.11. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$f_1,\dots,f_m$ 是 $\mathbb{A}$ 上的仿射线性函数，则我们称方程组

$$\begin{cases} f_1(\dot{x})=0 \\ f_2(\dot{x})=0 \\ \quad\vdots \\ f_m(\dot{x})=0 \end{cases}，\text{其中}\,\dot{x}\in\mathbb{A}. \tag{L}$$

是抽象的非齐次线性方程组，称 $\Pi=\{\dot{x}\in\mathbb{A}\mid\forall i=1,\dots,m,\ f_i(\dot{x})=0\}$ 是该方程组的解流形。

任取 $\mathbb{A}$ 的一组仿射坐标系 $(\dot{o},\mathbf{e}_1,\dots,\mathbf{e}_n)$，设 $\dot{x}$ 的坐标是 $(x_1,\dots,x_n)^t$，任取仿射线性函数 f，则 $f(\dot{x})=f(\dot{o})+\mathcal{A}\overline{op}$，其中 $\mathcal{A}$ 是 V 上的线性函数。不妨设 $f(\dot{o})=-b\in\mathbb{K}$，并且在 V 的基底 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下 $\mathcal{A}\overline{op}=\sum_{j=1}^n a_jx_j$，则 $f(\dot{x})=\sum_{j=1}^n a_jx_j-b$。也就是说，仿射线性函数在固定仿射坐标系之后是坐标的一次函数。

于是，对于方程组 (L) 而言，固定 $\mathbb{A}$ 的一组仿射坐标系后，设 $f_i(\dot{x})=\sum_{j=1}^n a_{ij}x_j-b_i$, $\forall i\in\{1,\dots,m\}$，则 (L) 等价于下面的非齐次线性方程组：

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \quad (\text{其中所有的}a_{ij},b_i\in\mathbb{K}.) \tag{L'}$$

结合讲义上册的定理 2.9.1 我们就可以得到：

命题 4.1.7. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，则 $\Pi\subset\mathbb{A}$ 是某个抽象的非齐次线性方程组 (L) 的解流形 $\iff$ Π 是 $\mathbb{A}$ 中的平面。

证明细节留作练习。利用线性方程组的知识我们还可以得到：如果方程组 (L') 的系数矩阵和增广矩阵的秩都是 r，则 (L) 的解流形 Π 是 $n-r$ 维的平面；反过来，$\mathbb{A}$ 中的任何一个 $n-r$ 维平面都都是某个含有 r 个方程的抽象非齐次线性方程组的解流形。特别地，超平面等价于一个仿射线性函数的解流形。证明留作练习。

在本节的最后，我们来讨论仿射空间中平面的位置关系。

定义 4.1.12. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$\Pi_1=\dot{x}+U_1$, $\Pi_2=\dot{y}+U_2$ 是 $\mathbb{A}$ 中的两个平面，如果 $U_1\subset U_2$ 或 $U_2\subset U_1$，则我们称 Π_1 和 Π_2 平行，记作 $\Pi_1//\Pi_2$。

需要注意的是，如果限制到真实的三维空间，那么我们这里定义的平行包含了重合、线在面内、线面平行和面面平行。而对 n 维空间，我们有：

定理 4.1.4. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，Π_1, Π_2 是 $\mathbb{A}$ 中的两个平面并且 $\Pi_1//\Pi_2$，则 $\Pi_1\cap\Pi_2$ 一定是以下三者之一：Π_1, Π_2 或 $\emptyset$。

证明. 如果 $\Pi_1\cap\Pi_2=\emptyset$，则证明已经完成；下设 $\Pi_1\cap\Pi_2\neq\emptyset$，即 $\exists\dot{p}\in\Pi_1\cap\Pi_2$，于是可设 $\Pi_1=\dot{p}+U_1$, $\Pi_2=\dot{p}+U_2$。由平行的定义，不妨设 $U_1\subset U_2$，我们来证明此时 $\Pi_1\subset\Pi_2$：任取 $\dot{q}=\dot{p}+\mathbf{x}\in\Pi_1$，则 $\mathbf{x}\in U_1\subset U_2$，即 $\dot{q}\in\dot{p}+U_2=\Pi_2$；如果设 $U_2\subset U_1$，我们同理可证 $\Pi_2\subset\Pi_1$。这样我们就完成了证明。 □

定理 4.1.5 (平行公理). 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，Π_1, Π_2 是 $\mathbb{A}$ 中的两个平面，

其方向子空间分别是 U_1, U_2，则 $\Pi_1 // \Pi_2 \iff \exists \dot{p} \in \Pi_1$ 和 $\dot{q} \in \Pi_2$ 使得 $\overline{pq} + \Pi_1$(即$\dot{p} + \overline{pq} + U_1$) $\subset \Pi_2$ 或者 $\overline{qp} + \Pi_2 \subset \Pi_1$。

证明. ($\Longrightarrow$) 不妨设 $U_1 \subset U_2$，则任取 $\dot{p} \in \Pi_1$ 和 $\dot{q} \in \Pi_2$，都一定有

$$\overline{pq} + \Pi_1 = \dot{p} + \overline{pq} + U_1 = \dot{q} + U_1 \subset \dot{q} + U_2 = \Pi_2;$$

如果设 $U_2 \subset U_1$ 则可以得到另一半结论。
($\Longleftarrow$) 如果存在 $\dot{p} \in \Pi_1$ 和 $\dot{q} \in \Pi_2$ 使得 $\overline{pq} + \Pi_1 \subset \Pi_2$，则 $\dot{p} + \overline{pq} + U_1 \subset \dot{q} + U_2$，即 $\dot{q} + U_1 \subset \dot{q} + U_2$，因此 $U_1 \subset U_2$，由平行的定义即得 $\Pi_1 // \Pi_2$。另一半结论同理。 □

我们知道，三维空间中两条直线可以既不平行也不相交，这种情形可以推广到一般的仿射空间。

定义 4.1.13. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，Π_1, Π_2 是 $\mathbb{A}$ 中的两个平面，如果 $\Pi_1 \cap \Pi_2 = \emptyset$ 并且 Π_1 和 Π_2 不平行，则我们称 Π_1 和 Π_2 是偏斜的 (或异面的)。

现在我们可以对两个平面之间的位置关系进行刻画了。

定理 4.1.6. 设 $\mathbb{A}$ 是域 $\mathbb{K}$ 上与 V 相伴的 n 维仿射空间，$\Pi_1 = \dot{p} + U_1$, $\Pi_2 = \dot{q} + U_2$ 是 $\mathbb{A}$ 中的两个平面并设 $\dim(U_1) = k \leqslant \dim(U_2) = l$。设 $i = \dim(U_1 \cap U_2)$，$m = \dim(\mathbb{A}(\Pi_1 \cup \Pi_2))$，则 Π_1 和 Π_2 的位置关系有且只有以下三种情况：

(1) $\Pi_1 // \Pi_2 \iff i = k$;

(2) Π_1 和 Π_2 相交但没有包含关系 $\iff i < k$ 且 $m = k + l - i$，此时 $\Pi_1 \cap \Pi_2$ 也是平面并且 $\dim(\Pi_1 \cap \Pi_2) = i$;

(3) Π_1 和 Π_2 偏斜 $\iff i < k$ 且 $m = k + l - i + 1$。

证明. 首先，由于 $i \leqslant k$，故只有 $i = k$ 和 $i < k$ 两种可能。
(1) 如果 $i = k$，则说明 $U_1 \cap U_2 = U_1$，即 $U_1 \subset U_2$，此即 $\Pi_1 // \Pi_2$。
(2) 如果 $i < k$，则说明 $U_1 \cap U_2 \subsetneq U_1, U_2$，即 Π_1 和 Π_2 不可能平行。那么我们考虑如下两种情形：

(i) 如果 $\Pi_1 \cap \Pi_2 \neq \emptyset$，即 Π_1 和 Π_2 相交但没有包含关系，那么由定义4.1.3下方的说明 (2) 可知 $\Pi_1 \cap \Pi_2$ 仍是平面并且 $\dim(\Pi_1 \cap \Pi_2) = \dim(U_1 \cap U_2) = i$。不妨设此时 $\dot{p} \in \Pi_1 \cap \Pi_2$ 且 $\mathbb{A}(\Pi_1 \cup \Pi_2)$ 的方向子空间是 W，即 $\mathbb{A}(\Pi_1 \cup \Pi_2) = \dot{p} + W$。则由 $\mathbb{A}(\Pi_1 \cup \Pi_2) \supset \Pi_1$, $\mathbb{A}(\Pi_1 \cup \Pi_2) \supset \Pi_2$ 可知 $W \supset U_1$，$W \supset U_2$，因此 $W \supset U_1 + U_2$。另一方面，容易看出 $\Pi_1 \cup \Pi_2$ 中的点一定都具有 $\dot{p} + \mathbf{w}$, $\mathbf{w} \in U_1 + U_2$ 的形式，因此其仿射组合也具有这样的形式，这说明 $W \subset U_1 + U_2$。综上，$W = U_1 + U_2$，因此由维数公式即得

$$m = \dim(\mathbb{A}(\Pi_1 \cup \Pi_2)) = \dim(W) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2) = k + l - i.$$

(ii) 如果 $\Pi_1 \cap \Pi_2 = \emptyset$，即 Π_1 和 Π_2 偏斜，下面我们来计算 $m = \dim(\mathbb{A}(\Pi_1 \cup \Pi_2))$。任取 $\dot{p} \in \Pi_1$ 和 $\dot{q} \in \Pi_2$，则 $\overline{pq} \notin U_1 + U_2$，这是因为如果 $\overline{pq} \in U_1 + U_2$，即 $\overline{pq} = \mathbf{x} + \mathbf{y}$, $\mathbf{x} \in U_1$, $\mathbf{y} \in U_2$，那么存在 $\dot{r} \in \Pi_1$ 使得 $\overline{pr} = \mathbf{x}$，于是 $\overline{rq} = \overline{pq} - \overline{pr} = \mathbf{y} \in U_2$，而 $\dot{q} \in \Pi_2$，这说明 $\dot{r} \in \Pi_2$，所以 $\dot{r} \in \Pi_1 \cap \Pi_2$，这与 $\Pi_1 \cap \Pi_2 = \emptyset$ 矛盾！

不妨设此时 $\mathbb{A}(\Pi_1\cup\Pi_2)$ 的方向子空间是 W, 则一方面 $W\supset U_1$，$W\supset U_2$ 并且 $\overline{pq}\in W$；另一方面，容易验证 (留作练习)$\mathbb{A}(\Pi_1\cup\Pi_2)$ 中的点一定都可以写成 $\dot{p}+\lambda\overline{pq}+\mathbf{u}_1+\mathbf{u}_2$，$\mathbf{u}\in U_1,\mathbf{u}_2\in U_2,\lambda\in\mathbb{K}$ 的形式。因此，$W=\operatorname{span}\{\overline{pq},U_1+U_2\}$，因此 $m=\dim(W)=\dim(U_1)+\dim(U_2)-\dim(U_1\cap U_2)+1=k+l-i+1$。

综上所述，$\mathbb{A}$ 中两个平面的位置关系有且只有以上情形，证毕。 □

4.2 欧几里得仿射空间

现在我们希望在仿射空间上引入垂直和距离的概念，这就得到了下面的欧氏仿射空间。

定义 4.2.1. 设 $\mathbb{E}$ 是实向量空间 V 相伴的 n 维仿射空间，如果 V 还是欧氏空间，其上的内积记为 $<\cdot,\cdot>$，则我们称 $\mathbb{E}$ 是欧几里得仿射空间，简称为欧氏仿射空间。

显然欧氏仿射空间是一种特殊的仿射空间，因此，我们上一节所讨论的仿射空间的所有性质都可以照搬到欧氏仿射空间上。而我们的目的是引入垂直和距离，这显然需要借助于内积结构来实现。

定义 4.2.2. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，$\dot{p},\dot{q}\in\mathbb{E}$，则 $\overline{pq}\in V$。我们称向量 $\overline{pq}$ 的范数 $\|\overline{pq}\|=\sqrt{<\overline{pq},\overline{pq}>}$ 为 $\dot{p},\dot{q}$ 两点间的距离。

欧氏空间中存在标准正交基，于是我们可以将仿射坐标系换成直角坐标系。

定义 4.2.3. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，任取 $\dot{o}\in\mathbb{E}$ 作为原点以及 V 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$，则任取 $\dot{p}\in\mathbb{E}$，$\dot{p}$ 都可以表示成 $\dot{p}=\dot{o}+\overline{op}=\dot{o}+\sum_{i=1}^n x_i\mathbf{e}_i$，$x_1,\ldots,x_n\in\mathbb{R}$。我们称 $(\dot{o},\mathbf{e}_1,\ldots,\mathbf{e}_n)$ 是 $\mathbb{E}$ 的一组直角坐标系或直角标架，$\dot{p}$ 在此直角坐标系下的坐标为 $(x_1,\ldots,x_n)^t$。

命题 4.2.1 (两点间距离公式). 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，点 $\dot{p},\dot{q}\in\mathbb{E}$ 在直角坐标系 $(\dot{o},\mathbf{e}_1,\ldots,\mathbf{e}_n)$ 下的坐标分别是 $(x_1,\ldots,x_n)^t$ 和 $(y_1,\ldots,y_n)^t$，则

$$\|\overline{pq}\|=\sqrt{\sum_{i=1}^n(y_i-x_i)^2}.$$

利用距离的定义以及内积的性质即可证明，留作练习。

定义 4.2.4. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，点 $\dot{p},\dot{q}\in\mathbb{E}$。则我们称 $\dot{p}\dot{q}=\{(1-\lambda)\dot{p}+\lambda\dot{q}\mid\forall 0\leqslant\lambda\leqslant 1\}\subset\mathbb{E}$ 为连接 $\dot{p},\dot{q}$ 两点的线段 (line segment)，$\|\overline{pq}\|$ 就是线段的长度[1]。

我们把 $\mathbb{R}^n$ 本身视作与 n 维标准欧氏空间相伴的欧氏仿射空间，称为 n 维标准欧氏仿射空间。

定义 4.2.5. 设 $\mathbb{E}_1,\mathbb{E}_2$ 是两个欧氏仿射空间，如果存在仿射同构 $f:\mathbb{E}_1\to\mathbb{E}_2$ 使得 f 保持任意两点间的距离，即 $\forall\dot{p}_1,\dot{p}_2\in\mathbb{E}$，都有 $\|\overline{f(\dot{p}_1)f(\dot{p}_2)}\|=\|\overline{p_1p_2}\|$，则我们称 $\mathbb{E}_1$ 和 $\mathbb{E}_2$ 是等距同构的，f 是等距同构映射。

我们很容易证明：$\mathbb{E}_1$ 和 $\mathbb{E}_2$ 等距同构 $\iff\dim(\mathbb{E}_1)=\dim(\mathbb{E}_2)$。特别地，任何 n 维欧氏仿射空间都与 n 维标准欧氏仿射空间 $\mathbb{R}^n$ 等距同构，证明留作练习。

定理 4.2.1. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，Π 是 $\mathbb{E}$ 中的平面，其方向子空间为 $U\subsetneq V$，点 $\dot{p}\notin\Pi$。则存在唯一的 $\dot{q}\in\Pi$ 使得 $\overline{pq}\perp U$，并且任取 $\dot{r}\in\Pi,\dot{r}\neq\dot{q}$，都有 $\|\overline{pq}\|<\|\overline{pr}\|$。

证明. 由于 $V=U\oplus U^{\perp}$，因此对任意的 $\dot{r}\in\Pi$ 而言，$\overline{pr}$ 存在唯一一组正交分解：$\overline{pr}=\mathbf{x}+\mathbf{y}$，其中 $\mathbf{x}\in U$，$\mathbf{y}\in U^{\perp}$。令 $\dot{q}=\dot{r}-\mathbf{x}\in\Pi$(因为 $\mathbf{x}\in U$)，则 $\overline{qr}=\mathbf{x}$，$\overline{pq}=\overline{pr}-\overline{qr}=\mathbf{y}\in U^{\perp}$，即 $\dot{q}=\dot{p}+\mathbf{y}$ 并且 $\overline{pq}\perp U$。

---

[1] 如果在 $\mathbb{E}$ 上定义曲线的长度，则可以证明两点之间线段最短。不过这已经是变分法 (或黎曼几何) 的内容了。

下面我们证明 $\dot{q}$ 是唯一的，即 $\dot{q}$ 不受 $\dot{r}$ 选取的影响。为此，另取 $\dot{r}_1 \in \Pi$，作正交分解 $\overline{pr_1} = \mathbf{x}_1 + \mathbf{y}_1$，$\mathbf{x}_1 \in U$，$\mathbf{y}_1 \in U^{\perp}$。令 $\dot{q}_1 = \dot{r}_1 - \mathbf{x}_1 = \dot{p} + \mathbf{y}_1$。则一方面，$\mathbf{y} - \mathbf{y}_1 \in U^{\perp}$；另一方面，$\mathbf{y} - \mathbf{y}_1 = \overline{pr} - \overline{pr_1} + \overline{x}_1 - \mathbf{x} = \overline{r_1 r} + \mathbf{x}_1 - \mathbf{x} \in U$。而 $U \cap U^{\perp} = \{\mathbf{0}\}$，故 $\mathbf{y} = \mathbf{y}_1$，所以 $\dot{q}_1 = \dot{p} + \mathbf{y}_1 = \dot{p} + \mathbf{y} = \dot{q}$。这就证明了 $\dot{q}$ 的唯一性。

最后我们来说明任取 $\dot{r} \in \Pi, \dot{r} \neq \dot{q}$，都有 $\|\overline{pq}\| < \|\overline{pr}\|$。由正交分解及勾股定理立刻可得：$\|\overline{pr}\|^2 = \|\overline{pq}\|^2 + \|\overline{qr}\|^2 \geqslant \|\overline{pq}\|^2$，并且等号成立 $\iff$ $\overline{qr} = \mathbf{0}$，即 $\dot{r} = \dot{q}$。这样我们就完成了证明。 □

由此我们可以定义点到平面的距离如下：

定义 4.2.6. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot, \cdot>)$ 相伴的 n 维欧氏仿射空间，Π 是 $\mathbb{E}$ 中的平面，其方向子空间为 $U \subsetneq V$，点 $\dot{p} \notin \Pi$。由定理4.2.1，存在唯一的 $\dot{q} \in \Pi$ 使得 $\overline{pq} \perp U$，我们称直线 L_{pq} 垂直于 Π，垂足是 $\dot{q}$，并称线段 $\dot{p}\dot{q}$ 的长度 $\|\overline{pq}\|$ 为点 $\dot{p}$ 到平面 Π 的距离。特别地，如果 $\dot{p} \in \Pi$，则我们称 $\dot{p}$ 到 Π 的距离是 0。

定理4.2.1也提供了点到平面距离的计算方法。取定 $\mathbb{E}$ 的一个直角坐标系 $(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$，设 $\dot{p}$ 的坐标是 $(x_1, \ldots, x_n)^t$，平面 $\Pi = \dot{r} + U$，其中 $\dot{r}$ 的坐标是 $(y_1, \ldots, y_n)^t$，U 的一组基是 $\mathbf{v}_1, \ldots, \mathbf{v}_m$。则我们对向量 $\overline{pr}$ 沿着 $U \oplus U^{\perp}$ 作正交分解 $\overline{pr} = \mathbf{x} + \mathbf{y}$，$\mathbf{x} \in U$，$\mathbf{y} \in U^{\perp}$，我们知道 $\|\mathbf{y}\| = \|\overline{pq}\|$ 就是 $\dot{p}$ 到 Π 的距离。于是，我们只需要求出 $\overline{pr}$ 在 $U^{\perp}$ 方向上的投影即可。利用基扩充定理可得 V 的一组基 $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n$，对其进行 Gram-Schmidt 正交化即可得到 V 的一组标准正交基 $\mathbf{u}_1, \ldots, \mathbf{u}_n$。则 $\mathbf{u}_{m+1}, \ldots, \mathbf{u}_n$ 就是 $U^{\perp}$ 的一组标准正交基。设 $\mathbf{u}_{m+1}, \ldots, \mathbf{u}_n$ 在基底 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的坐标分别为 $(a_{m+1,1}, \ldots, a_{m+1,n})^t, \cdots\cdots, (a_{n1}, \ldots, a_{nn})^t$，则 $\overline{pr}$ 在 $U^{\perp}$ 方向上的投影为 $\mathbf{y} =< \overline{pr}, \mathbf{u}_{m+1} > u_{m+1} + \cdots + < \overline{pr}, \mathbf{u}_n > u_n$，其长度为

$$\|\mathbf{y}\| = \sqrt{\sum_{i=m+1}^{n} (< \overline{pr}, \mathbf{u}_i >^2)} = \sqrt{\sum_{i=m+1}^{n} \left([\sum_{j=1}^{n} (y_j - x_j) a_{ij}]^2 \right)}.$$

如果我们还需要求出垂足 $\dot{q}$，则只需利用 $\dot{q} = \dot{p} + \mathbf{y}$ 即可。

例 4.2.1. 求标准欧氏仿射空间 $\mathbb{R}^3$ 中点 $\dot{p} = (1, 0, 0)$ 到平面 $\Pi : x + y + z = 2$ 的距离及垂足的坐标。

解. 取平面 Π 上的一个点 $\dot{r} = (2, 0, 0)^t$，则 $\overline{pr} = (1, 0, 0)^t$。平面 Π 的方向子空间

$$U = \text{span}\,\{(-1, 1, 0)^t, (-1, 0, 1)^t\},$$

于是

$$U^{\perp} = \text{span}\,\{\mathbf{v}_3 = (1, 1, 1)^t\},$$

(事实上，如果平面由一个方程组给出，那么方程组的系数向量 (即系数矩阵的行向量) 恰好张成了平面的方向子空间的正交补，这可以由方向子空间和解流形的定义验证，留作练习。)

因此，$\overline{pr}$ 在 $U^{\perp}$ 方向上的投影为

$$\mathbf{y} =< \overline{pr}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} > \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^t,$$

因此 $\dot{p}$ 到平面 π 的距离为 $\|\mathbf{y}\| = \frac{\sqrt{3}}{3}$，垂足为 $\dot{q} = \dot{p} + \mathbf{y} = (\frac{4}{3}, \frac{1}{3}, \frac{1}{3})^t$。 □

下面我们开始考虑欧氏仿射空间中两个平面的位置关系。显然两个平面之间仍然只存在平行、相交 (不包含)、偏斜这三种情形。然而在后两种情形中，我们可以考虑更特殊的位置关系：垂直。我们这里的垂直包括相交垂直和偏斜 (异面) 垂直。

定义 4.2.7. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，$\Pi_1 = \dot{p}+U_1$, $\Pi_2 = \dot{q}+U_2$ 是两个 $\mathbb{E}$ 中的不平行的平面。设 $U_1 \cap U_2 = W$, 并且 $U_1 = W \oplus W_1$, $U_2 = W \oplus W_2$, $W \perp W_1$, $W \perp W_2$。如果我们有 $W_1 \perp W_2$，则我们称 Π_1 和 Π_2 垂直，记作 $\Pi_1 \perp \Pi_2$。

例 4.2.2. 在 $\mathbb{R}^3$ 中平面 $x = 0$ 和 $y = 0$ 垂直。

特别地，我们考虑直线和平面垂直 (此时称直线是平面的垂线)，这种情形是比较简单的。设 n 维欧氏仿射空间 $\mathbb{E}$ 中直线 L_{pq} 和平面 $\Pi = \dot{r} + U$ 垂直，由于 L_{pq} 和 Π 不平行，因此 $\operatorname{span}\{\overline{pq}\} \cap U$ 只能是 0 维的，即 $\operatorname{span}\{\overline{pq}\} \cap U = \{\mathbf{0}\}$。于是按照上面的定义，$W_1 = \operatorname{span}\{\overline{pq}\}$, $W_2 = U$，因此 $L_{pq} \perp \Pi \iff \overline{pq} \perp U$。此外，如果线段所在的直线与某个平面垂直，则我们称该线段是这个平面的垂线段。

由此我们看到，求点到直线的距离实际上就是过平面外一点引平面的垂线段的过程。

下面我们考虑两个平面间的距离。设 $\Pi_1 = \dot{p} + U_1$, $\Pi_2 = \dot{q} + U_2$ 是两个 $\mathbb{E}$ 中的平面，首先，按照几何直观，我们可以定义距离为 $\mathrm{d}(\Pi_1, \Pi_2) = \inf\limits_{\dot{p}\in\Pi_1, \dot{q}\in\Pi_2} \|\overline{pq}\|$，那么，当 $\Pi_1 \cap \Pi_2 \neq \emptyset$ 时显然有 $\mathrm{d}(\Pi_1, \Pi_2) = 0$。那么，当 $\Pi_1 \cap \Pi_2 = \emptyset$ 时我们又该如何求这个距离呢？这就需要下面一系列的结论。

引理 4.2.1. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，$\Pi_1 = \dot{p} + U_1$, $\Pi_2 = \dot{q} + U_2$ 是两个 $\mathbb{E}$ 中的平面，$\Pi_1 \cap \Pi_2 = \emptyset$。那么，存在 $\dot{r} \in \Pi_1$ 和 $\dot{s} \in \Pi_2$ 使得 $\overline{rs} \perp \Pi_1$, $\overline{rs} \perp \Pi_2$。此时我们称直线 L_{rs} 是 Π_1 和 Π_2 的**公垂线**。特别地，公垂线唯一 $\iff$ $U_1 \cap U_2 = \{\mathbf{0}\}$。

证明. 首先，由 $\Pi_1 \cap \Pi_2 = \emptyset$ 可知 $\overline{pq} \notin U_1 + U_2$，其证明方法已经见于定理4.1.6的 (2)(ii)。于是 $\dim(U_1 + U_2) < n$，即 $(U_1 + U_2)^\perp \neq \{\mathbf{0}\}$。由于 $V = (U_1 + U_2) \oplus (U_1 + U_2)^\perp$，故我们可以对 $\overline{pq}$ 作正交分解：

$$\overline{pq} = \mathbf{u} + \mathbf{v},\ \text{其中}\mathbf{u} \in (U_1 + U_2),\ \mathbf{v} \in (U_1 + U_2)^\perp,\ \text{即}\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2,\ \mathbf{u}_1 \in U_1,\ \mathbf{u}_2 \in U_2.$$

于是，令 $\dot{r} = \dot{p} + \mathbf{u}_1 \in \Pi_1$, $\dot{s} = \dot{q} - \mathbf{u}_2 \in \Pi_2$，则 $\overline{rs} = \overline{pq} - (\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{v} \in (U_1 + U_2)^\perp$，即此时 $\overline{rs} \perp U_1$, $\overline{rs} \perp U_2$，所以 $\overline{rs} \perp \Pi_1$, $\overline{rs} \perp \Pi_2$。一般地，由于 $U_1 + U_2$ 不一定是直和，因此 $\mathbf{u}_1, \mathbf{u}_2$ 的选取不唯一，即 $\dot{r}, \dot{s}$ 不唯一。特别地，公垂线 L_{rs} 唯一 $\iff$ $\mathbf{u}_1, \mathbf{u}_2$ 的选取唯一 $\iff$ $U_1 \cap U_2 = \{\mathbf{0}\}$。这样我们就完成了证明。 □

定理 4.2.2. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，$\Pi_1 = \dot{p} + U_1$, $\Pi_2 = \dot{q} + U_2$ 是两个 $\mathbb{E}$ 中的平面，$\Pi_1 \cap \Pi_2 = \emptyset$。设 Π_1 和 Π_2 的公垂线是 L_{rs}, $\dot{r} \in \Pi_1$, $\dot{s} \in \Pi_2$，则对任意的 $\dot{p}_0 \in \Pi_1$ 和 $\dot{q}_0 \in \Pi_2$，都有 $\|\overline{rs}\| \leqslant \|\overline{p_0q_0}\|$，其中等号当且仅当 $\overline{p_0q_0} = \overline{rs}$ 时成立。

证明. 显然 $\overline{p_0q_0} = \overline{p_0r} + \overline{rs} + \overline{sq_0}$，并且由公垂线的定义有 $\overline{rs} \perp \overline{p_0r} + \overline{sq_0}$，于是由勾股定理立刻可得：

$$\|\overline{p_0q_0}\|^2 = \|\overline{p_0r} + \overline{sq_0}\|^2 + \|\overline{rs}\|^2 \geqslant \|\overline{rs}\|^2$$

等号成立当且仅当 $\overline{p_0r} + \overline{sq_0} = \mathbf{0}$，即 $\overline{p_0q_0} = \overline{rs}$。 □

于是我们可以定义两个平面的距离如下：

定义 4.2.8. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，$\Pi_1 = \dot{p} + U_1$，$\Pi_2 = \dot{q} + U_2$ 是两个 $\mathbb{E}$ 中的平面。如果 $\Pi_1 \cap \Pi_2 = \emptyset$，设 Π_1 和 Π_2 的公垂线是 L_{rs}，$\dot{r} \in \Pi_1$，$\dot{s} \in \Pi_2$，则我们称公垂线段的长度 $\|\overline{rs}\|$ 为 Π_1 和 Π_2 之间的距离；如果 $\Pi_1 \cap \Pi_2 \neq \emptyset$，则称 Π_1 和 Π_2 之间的距离是 0。

例 4.2.3. 在标准欧氏仿射空间 $\mathbb{R}^3$ 中给出两条直线的方程

$$l_1 : \begin{cases} 2x - y + z + 2 = 0, \\ x + 2y + 4z - 4 = 0, \end{cases} \qquad l_2 : \begin{cases} x + 2y - 1 = 0, \\ y - z + 2 = 0, \end{cases}$$

求它们的距离和公垂线的方程。

解. 容易计算出点 $\dot{p} = (0,2,0)^t \in l_1$，且 l_1 的方向子空间为 $U_1 = \mathrm{span}\,\{(6,7,-5)^t\}$；同样地，$\dot{q} = (1,0,2)^t \in l_2$，且 l_2 的方向子空间 $U_2 = \mathrm{span}\,\{(-2,1,1)^t\}$。于是，公垂线 l 的方向向量 $\mathbf{x} \perp U_1 + U_2$，于是可以计算出 $\mathbf{x} = (3,1,5)^t$。所以公垂线段的长度 d 是 $\overline{pq}$ 到方向 $\mathbf{x}$ 的投影的长度：

$$d = \frac{|< \overline{pq}, \mathbf{x} >|}{\|\mathbf{x}\|} = \frac{11\sqrt{35}}{35}.$$

下面我们来求公垂线的方程。我们写出 l_1 和 l_2 的参数方程：

$$l_1 : \begin{cases} x_1 = 6t, \\ y_1 = 2 + 7t, \\ z_1 = -5t. \end{cases} \quad l_2 : \begin{cases} x_2 = 1 - 2u, \\ y_2 = u, \\ z_2 = 2 + u. \end{cases}$$

则垂足 $\dot{r} = (x_1, y_1, z_1) \in l_1$，$\dot{s} = (x_2, y_2, z_2) \in l_1$ 应满足 $\overline{rs} // \mathbf{x}$，即

$$\begin{cases} 1 - 2u - 6t = 3\lambda, \\ u - 2 - 7t = \lambda, \\ 2 + u + 5t = 5\lambda. \end{cases}$$

解得

$$\begin{cases} t = -\dfrac{8}{35}, \\ u = \dfrac{5}{7}, \\ \lambda = \dfrac{11}{35}. \end{cases}$$

于是 $\dot{s} = (-\frac{3}{7}, \frac{5}{7}, \frac{19}{7})^t$，公垂线方程为：

$$\frac{x + \frac{3}{7}}{3} = \frac{y - \frac{5}{7}}{1} = \frac{z - \frac{19}{7}}{5}.$$

□

下面我们开始讨论欧氏仿射空间中一类特殊的仿射变换：保距变换。

定义 4.2.9. 设 $\mathbb{E}_1, \mathbb{E}_2$ 分别是 $\mathbb{R}$ 上与 V_1, V_2 相伴的欧氏仿射空间，$f : \mathbb{E}_1 \to \mathbb{E}_2$ 是仿射映射。如果 f 还满足：$\forall \dot{p}, \dot{q} \in \mathbb{E}_1$，都有 $\|\overline{f(\dot{p})f(\dot{q})}\| = \|\overline{pq}\|$，则我们称 f 是一个保距仿射映射，简

称保距映射。特别地，我们主要关注 $\mathbb{E}_1 = \mathbb{E}_2$ 的情形，此时我们称 f 是一个**保距变换**或者**运动**。

定理 4.2.3. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V, <\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，固定其一组直角坐标系 $(\dot{o}, \mathbf{e}_1, \ldots, \mathbf{e}_n)$，则 f 是 $\mathbb{E}$ 上的一个运动 $\iff$ $f = T_{\mathbf{v}} \circ \mathbf{A}$，其中 $\mathbf{v} \in V$, $\mathbf{A} : \mathbb{E} \to \mathbb{E}$, $\dot{p} \mapsto \dot{o} + \mathcal{A}\overline{op}$, $\mathcal{A}$ 是正交算子。换言之，如果 $\dot{p}$ 在这组直角坐标系下的坐标是 $(x_1, \ldots, x_n)^t$，则 $f(\dot{p})$ 的坐标是 $(v_1, \ldots, v_n)^t + A \cdot (x_1, \ldots, x_n)^t$，其中 $(v_1, \ldots, v_n)^t$ 是向量 $\mathbf{v} = \overline{\dot{o}f(\dot{o})}$ 的坐标，A 是 $\mathcal{A}$ 对应的正交矩阵。

证明. ($\Longleftarrow$) 如果 $f = T_{\mathbf{v}} \circ \mathbf{A}$，则由推论4.1.2，$f$ 首先是仿射变换，下面我们验证 f 保持距离。任取 $\dot{p}, \dot{q} \in \mathbb{E}$，我们有

$$\begin{aligned}
\|\overline{f(\dot{p})f(\dot{q})}\| &= \|\overline{(\mathbf{A}\dot{p} + \mathbf{v})(\mathbf{A}\dot{q} + \mathbf{v})}\| = \|\overline{(\mathbf{A}\dot{p})(\mathbf{A}\dot{q})}\| \\
&= \|\overline{(\dot{o} + \mathcal{A}\overline{op})(\dot{o} + \mathcal{A}\overline{oq})}\| = \|\mathcal{A}\overline{oq} - \mathcal{A}\overline{op}\| \\
&= \|\mathcal{A}(\overline{oq} - \overline{op})\| = \|\mathcal{A}\overline{pq}\| \\
&= \|\overline{pq}\| \quad (\text{利用了}\mathcal{A}\text{是正交算子, 定理3.1.5.})
\end{aligned}$$

此即 f 是保距变换 (运动)。

($\Longrightarrow$) 我们来证明一个更强的结论：如果 $f : \mathbb{E} \to \mathbb{E}$ 满足 $\|\overline{f(\dot{p})f(\dot{q})}\| = \|\overline{pq}\|$，则 f 一定是运动并且 f 可以写成 $f = T_{\mathbf{v}} \circ \mathbf{A}$，$\mathbf{A}$ 对应正交算子 $\mathcal{A}$ 的形式。

令 $\mathbf{v} = \overline{\dot{o}f(\dot{o})}$，$\mathbf{A} = T_{-\mathbf{v}} \circ f$，则容易验证 $\mathbf{A}(\dot{o}) = \dot{o}$。首先，由 ($\Longleftarrow$) 方向的证明容易看出平移变换 $T_{-\mathbf{v}}$ 一定是运动。下面我们令 $\mathcal{A} : V \to V$, $\overline{op} \mapsto \overline{\dot{o}\mathbf{A}(\dot{p})}$，则我们有：

(i) $\mathcal{A}$ 保持向量的范数不变。这是因为：任取 $\mathbf{x} \in V$，存在唯一的 $\dot{p} \in \mathbb{E}$ 使得 $\overline{op} = \mathbf{x}$，于是

$$\begin{aligned}
&\|\mathcal{A}\mathbf{x}\| = \|\mathcal{A}\overline{op}\| = \|\overline{\dot{o}\mathbf{A}(\dot{p})}\| \\
&= \|\overline{T_{-\mathbf{v}} \circ f(\dot{o}) \quad T_{-\mathbf{v}} \circ f(\dot{p})}\| \\
&= \|\overline{f(\dot{o})f(\dot{p})}\| \qquad (\text{利用了平移是运动}) \\
&= \|\overline{op}\| \qquad (\text{利用了 } f \text{ 保持距离}) \\
&= \|\mathbf{x}\|
\end{aligned}$$

此即 $\mathcal{A}$ 保持向量的范数不变。

(ii) $\mathcal{A}$ 是线性算子。这是因为：首先，任取 $\mathbf{x}, \mathbf{y} \in V$，利用极化恒等式我们很容易验证 $\mathcal{A}$ 保持内积，即 $< \mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y} > = < \mathbf{x}, \mathbf{y} >$(即定理3.4.1($\Longleftarrow$) 的过程)。于是我们有：

$$\begin{aligned}
&\|\mathcal{A}(\mathbf{x} + \mathbf{y}) - \mathcal{A}\mathbf{x} - \mathcal{A}\mathbf{y}\|^2 \\
&= \|\mathcal{A}(\mathbf{x} + \mathbf{y})\|^2 + \|\mathcal{A}\mathbf{x}\|^2 + \|\mathcal{A}\mathbf{y}\|^2 - 2 < \mathcal{A}(\mathbf{x} + \mathbf{y}), \mathcal{A}\mathbf{x} > -2 < \mathcal{A}(\mathbf{x} + \mathbf{y}), \mathcal{A}\mathbf{y} > +2 < \mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y} > \ (\text{内积双线性性质}) \\
&= \|(\mathbf{x} + \mathbf{y})\|^2 + \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2 < \mathbf{x} + \mathbf{y}, \mathbf{x} > -2 < \mathbf{x} + \mathbf{y}, \mathbf{y} > +2 < \mathbf{x}, \mathbf{y} > \quad (\text{利用 } \mathcal{A} \text{ 保持范数和内积}) \\
&= \|(\mathbf{x} + \mathbf{y}) - \mathbf{x} - \mathbf{y}\|^2 \\
&= 0.
\end{aligned}$$

所以 $\mathcal{A}(\mathbf{x}+\mathbf{y})-\mathcal{A}\mathbf{x}-\mathcal{A}\mathbf{y}=\mathbf{0}$。任取 $\mathbf{x}\in V$ 及 $\lambda\in\mathbb{R}$，则有

$$\begin{aligned}||\mathcal{A}(\lambda\mathbf{x})-\lambda\mathcal{A}\mathbf{x}||^2&=||\mathcal{A}(\lambda\mathbf{x})||^2+\lambda^2||\mathcal{A}\mathbf{x}||^2-2\lambda<\mathcal{A}(\lambda\mathbf{x},\mathcal{A}\mathbf{x})> &&\text{(内积双线性性质)}\\&=||\lambda\mathbf{x}||^2+\lambda^2||\mathbf{x}||^2-2\lambda<\lambda\mathbf{x},\mathbf{x}> &&\text{(利用 }\mathcal{A}\text{ 保持范数和内积)}\\&=0.\end{aligned}$$

所以 $\mathcal{A}(\lambda\mathbf{x})-\lambda\mathcal{A}\mathbf{x}=\mathbf{0}$。综上所述 $\mathcal{A}$ 是线性算子，从而 f 确实是仿射变换。

综合 (i)、(ii) 可知 $\mathcal{A}$ 是正交算子，从而 f 是运动。“换言之”后面的部分即定理的坐标形式，是显然的。这样我们就完成了证明。 □

上面的证明方法同样可以用于弱化定理3.4.1的条件，也就是说，内积空间 (甚至不一定是有限维的) 上保持范数的算子一定是线性算子，并且一定保持内积。证明细节留作练习。

利用所有正交算子构成正交群 $O(V)$(推论3.1.3) 以及命题4.1.5的证明方法，我们很容易验证：欧氏仿射空间上的所有运动构成一个群 (细节留作练习)，因此我们有以下的定义：

定义 4.2.10. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，则 $\mathbb{E}$ 上的所有运动在映射复合的运算下构成一个群，称为 $\mathbb{E}$ 的运动群 (或保距群，Isometry group)，记为 $\mathrm{Iso}(\mathbb{E})$。容易验证 $\mathrm{Iso}(\mathbb{E})$ 是 $\mathrm{Aff}(\mathbb{E})$ 的一个子群，留作练习。

接下来，我们自然要知道运动群中的元素的具体结构，换言之，我们需要对运动群中的元素进行分类。为此，我们首先引入运动的轴的概念。

定理 4.2.4. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧氏仿射空间，原点是 $\dot{o}$，$f=T_{\mathbf{v}}\circ\mathbf{A}$ 是 $\mathbb{E}$ 上的一个运动，其中 $\mathbf{v}\in V$，$\mathbf{A}:\mathbb{E}\to\mathbb{E}$，$\dot{p}\mapsto\dot{o}+\mathcal{A}\overline{op}$，$\mathcal{A}$ 是正交算子，则存在唯一的平面 $\Pi=\dot{p}_0+U$ 使得以下两个条件都成立：

(1) $f(\Pi)=\Pi$ 并且 $f\big|_{\Pi}$ 是平移；

(2) 记 $\mathcal{E}$ 是 V 上的恒等算子，则 $\mathcal{A}-\mathcal{E}$ 限制在 $U^{\perp}$ 上是非退化 (可逆) 的线性算子。

我们称平面 Π 为运动 f 的轴。

证明. 由于 $\mathcal{A}$ 是正交算子，故 $\mathcal{A}$ 的特征值的模长一定是 1。令 U 是特征值 1 对应的特征子空间 (如果 $\mathcal{A}$ 没有特征值 1 则 $U=\{\mathbf{0}\}$)，则 $U\cap U^{\perp}=\{\mathbf{0}\}$，而且 $\mathcal{A}-\mathcal{E}$ 限制到 $U^{\perp}$ 上的特征值一定不为 0，即 $\mathcal{A}-\mathcal{E}$ 限制在 $U^{\perp}$ 上是非退化线性算子。

下面我们需要做的是：找到一个以 U 为方向子空间的平面 Π 使得 f 限制到 Π 上是平移。为此，我们对 $(f=T_{\mathbf{v}}\circ A$ 中的)$\mathbf{v}$ 沿着 U 和 $U^{\perp}$ 的方向进行正交分解，即存在唯一一组 $\mathbf{u}_0\in U$ 和 $\mathbf{w}_0\in U^{\perp}$ 使得 $\mathbf{v}=\mathbf{u}_0+\mathbf{w}_0$。由于 $\mathcal{A}-\mathcal{E}$ 在 $U^{\perp}$ 上非退化，故存在唯一的 $\mathbf{x}$ 使得 $(\mathcal{A}-\mathcal{E})\mathbf{x}=-\mathbf{w}_0$。令点 $\dot{p}=\dot{o}+\mathbf{x}$，取 $\Pi=\dot{p}+U$，则我们注意到

$$\begin{aligned}f(\dot{p})&=T_{\mathbf{v}}\circ\mathbf{A}(\dot{p})\\&=T_{\mathbf{v}}(\dot{o}+\mathcal{A}\overline{op})\\&=T_{\mathbf{v}}(\dot{o}+\mathcal{A}\mathbf{x})\\&=\dot{o}+\mathbf{x}-\mathbf{w}_0+\mathbf{v}\\&=\dot{o}+\mathbf{x}+\mathbf{u}_0\\&=\dot{p}+\mathbf{u}_0\in\Pi.\end{aligned}$$

于是由 f 的线性性质，我们有：对 $\forall \mathbf{q} = \dot{p} + \mathbf{u} \in \Pi$, $\mathbf{u} \in U$，有

$$f(\dot{q}) = f(\dot{p}) + \mathcal{A}\overline{pq} = \dot{p} + \mathbf{u}_0 + \mathcal{A}\mathbf{u}$$

注意到 U 是 $\mathcal{A}$ 的特征值 1 对应的特征子空间，所以 $\mathcal{A}\mathbf{u} = \mathbf{u}$，即

$$f(\dot{q}) = \dot{p} + \mathbf{u}_0 + \mathbf{u} = \dot{q} + \mathbf{u}_0$$

即 $f\big|_{\Pi} = T_{\mathbf{u}_0}$ 是平移，从而自然有 $f(\Pi) = \Pi$。

最后我们还要说明 Π 是唯一的。首先，要使 $f\big|_{\Pi}$ 是平移，则 $f\big|_{\Pi}$ 的线性部分是恒等算子，即 $\mathcal{A}\big|_{U} = \mathcal{E}$，或者说 U 包含在 1 对应的特征子空间中；而 $\mathcal{A} - \mathcal{E}$ 在 $U^{\perp}$ 上非退化则说明 $U^{\perp}$ 中不含特征值 1 的特征向量，因此，U 只能等于 1 对应的特征子空间。因此，如果另有平面 Π_1 也满足条件，则 Π_1 的方向子空间也是 U，并且 $f\big|_{\Pi_1}$ 这个平移与 $f\big|_{\Pi}$ 的平移幅度是相等的 (即都是 $\mathbf{v}$ 沿着 U 的正交投影 $\mathbf{u}_0$，这一点读者很容易验证)。那么，如果 $\Pi_1 \neq \Pi$，则必存在 $\dot{p}_1 \in \Pi_1$ 且 $\dot{p}_1 \notin \Pi$，此时 $\overline{pp_1} \notin U$。因此，$\mathbb{A}(\Pi \cup \Pi_1)$ 的方向子空间真包含 U。然而我们容易验证，如果条件 (1) 和 (2) 在 Π 和 Π_1 上都成立，则条件 (1) 和 (2) 在 $\mathbb{A}(\Pi \cup \Pi_1)$ 上也成立，而 $\mathbb{A}(\Pi \cup \Pi_1)$ 的方向子空间不是 U，这就产生了矛盾！因此必有 $\Pi_1 = \Pi$。这样我们就完成了证明。 □

这个定理的证明过程也给出了运动的轴的计算方法：先求 $\mathcal{A}$ 的特征值 1 的特征子空间 U，再通过正交分解求出 Π 上的一个点 $\dot{p}$ 即可。

现在我们开始对运动分类。限于篇幅，我们只介绍一维和二维空间中运动的分类，更高维空间中运动的分类留给读者思考。

- $\dim(\mathbb{E}) = 1$ 的情形。

  取 $\mathbb{E}$ 的一组直角坐标系 $(\dot{o}, \mathbf{e}_1)$，此时若 $\dot{p}$ 的坐标是 $x(\in \mathbb{R})$，则由定理4.2.3，$f(\dot{p})$ 的坐标是 $y = ax + b$，其中 a 是 1 阶正交矩阵，即 $a^2 = 1$。

  (1) 如果 $a = 1$，即 $y = x + b$，则 f 就是 $\mathbb{E}$ 上的平移，即 f 的轴是整个 $\mathbb{E}$。

  (2) 如果 $a = -1$，即 $y = -x + b$，则 $\mathcal{A} - \mathcal{E}$ 在整个 $\mathbb{E}$ 上非退化，轴是点 $\dfrac{b}{2}$，我们以 $\dfrac{b}{2}$ 作为新原点 $\dot{o}'$，则新坐标满足：$x = x' + \dfrac{b}{2}$，$y = y' + \dfrac{b}{2}$，于是由 $y = -x + b$ 即可得到 $y' = -x'$，这是一个关于新原点 (即旧坐标 $\dfrac{b}{2}$) 的反射。

  综上所述，1 维的运动有且只有两种情形：平移和反射。

- $\dim(\mathbb{E}) = 2$ 的情形。

  取 $\mathbb{E}$ 的一组直角坐标系 $(\dot{o}, \mathbf{e}_1, \mathbf{e}_2)$，设 $\dot{p}$ 的坐标为 $(x_1, x_2)^t$，$f(\dot{p})$ 的坐标为 $(y_1, y_2)^t$，$f(\dot{o})$ 的坐标是 $(v_1, v_2)^t$，则

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \text{其中} AA^t = A^tA = E. \tag{4.2.1}$$

  下面对正交矩阵 A 进行讨论。首先，A 如果有实特征值，则一定是 ± 1；而由引理3.4.3，我们知道在 $\mathbb{R}$ 上不能对角化的正交矩阵一定形如 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$，$\theta \in \mathbb{R} \setminus \{2k\pi \mid k \in \mathbb{Z}\}$。

此外，我们注意到形如 $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$，$a^2+b^2=1$ 的正交矩阵可以进行对角化：令 $a=\cos\theta, b=\sin\theta$，则

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$$

于是我们只需讨论如下三种基本情形，其余情形均可视作通过坐标变换回到以下的基本情形。

(1) $A=E$。此时运动的轴是全空间 $\mathbb{E}$，坐标关系 (4.2.1) 变成

$$\begin{cases} y_1 = x_1 + b_1 \\ y_2 = x_2 + b_2 \end{cases}$$

容易看出这是一个平移。

(2) $A=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$。此时运动的轴是 $\dot{o}+\operatorname{span}\{\mathbf{e}_1\}$，坐标关系 (4.2.1) 变成

$$\begin{cases} y_1 = x_1 + b_1 \\ y_2 = -x_2 + b_2 \end{cases}$$

以 $\dot{o}'=\dot{o}+\dfrac{b_2}{2}$ 为新原点，则在新坐标系 $(\dot{o}',\mathbf{e}_1,\mathbf{e}_2)$ 下，$f(\dot{p})$ 与 $\dot{p}$ 的坐标关系为：

$$\begin{cases} x_1' = x_1 \\ x_2' = x_2 - \dfrac{b_2}{2} \\ y_1' = y_1 \\ y_2' = y_2 - \dfrac{b_2}{2} \end{cases} \implies \begin{cases} y_1' = x_1' + b_1 \\ y_2' = -x_2 \end{cases}$$

这是一个 $\mathbf{e}_1$ 方向上的平移和 $\mathbf{e}_2$ 方向上的反射。

(3) $A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$，$\theta\in\mathbb{R}\setminus\{2k\pi \mid k\in\mathbb{Z}\}$，则 A 没有特征值 1，因此运动的轴是 0 维的 (一个点)，即 f 有唯一的不动点。令 $\dot{o}'=\dot{o}+\xi_1\mathbf{e}_1+\xi_2\mathbf{e}_2$ 是 f 的不动点，换言之，ξ_1,ξ_2 满足

$$\begin{cases} \xi_1 = \cos\theta\xi_1 - \sin\theta\xi_2 + b_1 \\ \xi_2 = \sin\theta\xi_1 + \cos\theta\xi_2 + b_2 \end{cases}$$

则在新坐标系 $(\dot{o}',\mathbf{e}_1,\mathbf{e}_2)$，容易验证

$$\begin{cases} x_1' = x_1 - \xi_1 \\ x_2' = x_2 - \xi_2 \\ y_1' = y_1 - \xi_1 \\ y_2' = y_2 - \xi_2 \end{cases} \implies \begin{cases} y_1' = \cos\theta x_1' - \sin\theta x_2' \\ y_2' = \sin\theta x_1' + \cos\theta x_2' \end{cases}$$

这是一个绕点 $\dot{o}'$ 旋转 θ 角的运动。

三维空间中的运动按照轴 Π 的维数可分为：平移 $(\dim(\Pi)=3)$；沿直线反射或滑动反射 $(\dim(\Pi)=2)$；绕直线旋转或螺旋前进 $(\dim(\Pi)=1)$；镜像旋转（即绕一条直线旋转与沿平面反射合成，直线与平面垂直，此时 $\dim(\Pi)=0$）。

设运动对应的正交算子是 $\mathcal{A}$，我们把 $\det(\mathcal{A})=1$ 的运动称为刚体运动，$\det(\mathcal{A})=-1$ 的运动称为非刚体运动。可以验证，在小于等于 3 维的空间中，平移、旋转和螺旋是刚体运动，而所有含有反射的运动都是非刚体运动，验证细节留作练习。

利用群的观点我们也可以重新定义图形的全等。

定义 4.2.11. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 V 相伴的 n 维欧式空间，Φ_1 和 Φ_2 是 $\mathbb{E}$ 中的两个图形（子集），G 是 $\mathrm{Aff}(\mathbb{E})$ 的一个子群，如果存在变换 $g\in G$ 使得 $\Phi_2=g(\Phi_1)$，则称 Φ_1 和 Φ_2 是 $G-$ 全等（或称为 $G-$ 等价或 $G-$ 叠合，$G-$congruence）。我们把这种情形记作 $\Phi_1 \overset{G}{\simeq} \Phi_2$。

例如，如果取 $G=\mathrm{Aff}(\mathbb{E})$，则 $G-$ 全等是我们中学熟悉的相似关系；如果取 $G=\mathrm{Iso}(\mathbb{E})$，则 $G-$ 全等是我们中学所学的全等关系。

在本节的最后，我们给出欧氏空间中仿射变换的粗略分类。

定理 4.2.5. 设 $\mathbb{E}$ 是 $\mathbb{R}$ 上与 $(V,<\cdot,\cdot>)$ 相伴的 n 维欧式空间，$f=T_{\mathbf{v}}\circ\mathbf{A}\in\mathrm{Aff}(\mathbb{A})$ 是 $\mathbb{E}$ 上的一个仿射变换，如果 $\det(\mathbf{A})>0$，则 f 一定可以写成以下三种变换的复合：

(1) 平移；

(2) 绕原点 $\dot{o}$ 的旋转；

(3) 取定以 $\dot{o}$ 为原点的某个直角坐标系后，沿各个坐标轴方向同时做伸缩变换。

如果 $\det(\mathcal{A})<0$，则在以上变换的基础上还需要加上一个反射，这个反射的轴是过原点 $\dot{o}$ 的某个平面。

证明. 首先由推论4.1.2，我们已经得到了平移 $T_{\mathbf{v}}$ 和原点 $\dot{o}$，使得 $\mathbf{A}\dot{o}=\dot{o}$ 且 $\mathbf{A}$ 是线性变换。对 $\mathbf{A}$ 作极化分解（定理3.4.4）可得 $\mathbf{A}=\mathcal{P}\mathcal{Q}$，其中 $\mathcal{P}$ 是正定算子，$\mathcal{Q}$ 是正交算子。于是由定理3.3.5，可以取定 V 的一组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得

$$\mathcal{P}\mathbf{e}_i=\lambda_i\mathbf{e}_i,\ \lambda_i>0,\ i=1,\dots,n.$$

即 $\mathcal{P}$ 是我们所需的伸缩。另一方面，按照定理3.4.3，我们知道 $\mathcal{Q}$ 是一个绕原点 $\dot{o}$ 的旋转，因为它可以分解成一些绕原点的二维平面旋转 $A_1,\dots,A_r$ 以及反射 $-E_t$，而定理3.4.3中 $-E_t$ 的部分 t 必须是偶数 $(\det(\mathbf{A})>0)$，故 $-E_t$ 这部分可以视作中心对称（旋转角 π）。

如果 $\det(\mathbf{A})<0$，则上面的一些 $\lambda_i<0$，即相当于我们添加了一些沿坐标轴 $\mathbf{e}_i$ 的反射。 □

4.3 凸集

凸集在几何和分析中具有良好的性质和十分重要的意义，并且在优化等领域也有广泛的应用。本节我们将介绍凸集的定义和一些简单性质。

定义 4.3.1. 设 $\mathbb{A}$ 是域 $\mathbb{R}$ 上与 V 相伴的 n 维仿射空间，$M \subset \mathbb{A}$。如果对 $\forall \dot{p}, \dot{q} \in M$，及 $0 \leqslant \lambda \leqslant 1$ 都有

$$(1-\lambda)\dot{p} + \lambda\dot{q} \in M,$$

则我们称 M 是 $\mathbb{A}$ 中的一个凸集 (convex set)。

例 4.3.1. $\mathbb{R}^2$ 中的线段、三角形，$\mathbb{R}^3$ 中的三棱锥、正方体等都是凸集。

下面我们考虑一类简单的凸集：单纯形。

定义 4.3.2. 设 $\dot{p}_0, \dot{p}_1, \ldots, \dot{p}_m$ 是实仿射空间 $\mathbb{A}$ 中 $m+1$ 个仿射无关的点，我们很容易验证

$$\{\sum_{i=0}^{m} \lambda_i \dot{p}_i \mid \sum_{i=0}^{m} \lambda_i = 1 \text{且} \lambda_i \geqslant 0,\ i = 1, \ldots, m\}$$

是凸集，称为 m 维闭单纯形 (closed simplex)。相应的，如果要求改成 $\lambda_i > 0$，则称为 m 维开单纯形 (open simplex)。

由于仿射映射保持仿射组合 (定义4.1.8)，我们知道仿射无关的点在仿射变换下仍然是仿射无关的，因此单纯形在任意仿射变换下还是单纯形。有关单纯形的详细性质是拓扑学中单纯同调论的内容，可以参考《基础拓扑学》，尤承业，北京大学出版社或《Algebraic Topology》，Allen Hatcher, Cambridge University Press。

下面我们给出一些构造凸集的简单方式。

命题 4.3.1. 任意多个凸集的交集仍是凸集。

利用凸集的定义即可验证。

定义 4.3.3. 对 $\mathbb{A}$ 任意子集 S，记

$$\mathrm{C(S)} = \bigcap_{\mathrm{T \supset S,\ T是凸集}} \mathrm{T},$$

则由上面的命题可知 C(S) 也是凸集，我们称 C(S) 为 S 的凸包 (convex hull)。也就是说，S 的凸包是包含 S 的最小凸集。

下面的定理给出了凸包中具体含有哪些点。

定理 4.3.1. 设 $M \subset \mathbb{A}$ 是凸集，$\dot{p} \in \mathbb{A}$，则

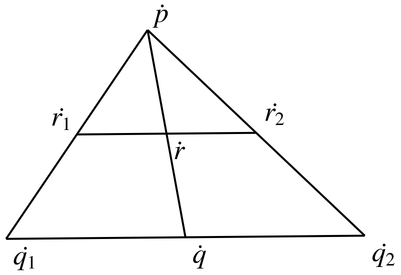
$$\mathrm{C(M \cup \{\dot{p}\})} = \bigcup_{\mathrm{\dot{q} \in M}} \mathrm{\dot{p}\dot{q}},$$

其中 $\dot{p}\dot{q} = \{(1-\lambda)\dot{p} + \lambda\dot{q} \mid \forall 0 \leqslant \lambda \leqslant 1\} \subset \mathbb{A}$ 为连接 $\dot{p}, \dot{q}$ 两点的线段。

证明. $\dot{p} \in M$ 的平凡情形是十分显然的。下面证明 $\dot{p} \notin M$ 的情形。

首先，由于 $\mathrm{C(M \cup \{\dot{p}\})}$ 是凸集，$\dot{p}, \dot{q} \in \mathrm{C(M \cup \{\dot{p}\})}$，所以线段 $\dot{p}\dot{q} \subset \mathrm{C(M \cup \{\dot{p}\})}$，$\forall \mathrm{\dot{q} \in M}$。因此 $\mathrm{C(M \cup \{\dot{p}\})} \supset \bigcup_{\mathrm{\dot{q} \in M}} \mathrm{\dot{p}\dot{q}}$。下面我们只需证明 $\mathrm{C(M \cup \{\dot{p}\})} \subset \bigcup_{\mathrm{\dot{q} \in M}} \mathrm{\dot{p}\dot{q}}$ 即可。而这只需要说明 $\bigcup_{\dot{q} \in M} \dot{p}\dot{q}$ 是凸集。

实际上，$\bigcup\limits_{\dot q\in M}\dot p\dot q$ 是凸集可以从下图中得到：



要证 $\bigcup\limits_{\dot q\in M}\dot p\dot q$ 是凸集，只需证任取 $\dot r_1\in\dot p\dot q_1$，$\dot r_2\in\dot p\dot q_2$，$\dot q_1,\dot q_2\in M$，我们都有 $\dot r_1\dot r_2\subset\bigcup\limits_{\dot q\in M}\dot p\dot q$。不妨设

$$\begin{cases}\dot r_1=\lambda\dot p+(1-\lambda)\dot q_1\\ \dot r_2=\mu\dot p+(1-\mu)\dot q_2\end{cases}\text{，其中}\lambda,\mu\in[0,1].$$

任取 $\dot r\in\dot r_1\dot r_2$，即 $\dot r=\beta\dot r_1+(1-\beta)\dot r_2$，$\beta\in[0,1]$，则

$$\begin{aligned}\dot r&=\beta(\lambda\dot p+(1-\lambda)\dot q_1)+(1-\beta)(\mu\dot p+(1-\mu)\dot q_2)\\&=(\beta\lambda+(1-\beta)\mu)\dot p+(1-\beta\lambda-\mu+\beta\mu)\left(\frac{\beta(1-\lambda)}{1-\beta\lambda-\mu+\beta\mu}\dot q_1+\frac{(1-\beta)(1-\mu)}{1-\beta\lambda-\mu+\beta\mu}\dot q_2\right).\end{aligned}\tag{4.3.1}$$

注意到

$$\frac{\beta(1-\lambda)}{1-\beta\lambda-\mu+\beta\mu}+\frac{(1-\beta)(1-\mu)}{1-\beta\lambda-\mu+\beta\mu}=1,$$

由于 M 是凸集，$\dot q_1,\dot q_2\in M$，因此

$$\frac{\beta(1-\lambda)}{1-\beta\lambda-\mu+\beta\mu}\dot q_1+\frac{(1-\beta)(1-\mu)}{1-\beta\lambda-\mu+\beta\mu}\dot q_1\in\dot q_1\dot q_2\subset M,$$

记 $\frac{\beta(1-\lambda)}{1-\beta\lambda-\mu+\beta\mu}\dot q_1+\frac{(1-\beta)(1-\mu)}{1-\beta\lambda-\mu+\beta\mu}\dot q_1$ 为点 $\dot q$，则由 (4.3.1) 即得 $\dot r\in\dot p\dot q\subset\bigcup\limits_{\dot q\in M}\dot p\dot q$。此即 $\bigcup\limits_{\dot q\in M}\dot p\dot q$ 是凸集，这样我们就完成了证明。 □

更进一步地，我们有

定理 4.3.2. 设 $S=\{\dot p_1,\dots\dot p_k\}$ 是实仿射空间 $\mathbb{A}$ 的有限子集，则

$$\mathrm{C(S)}=\{\lambda_1\dot{\mathrm{p}}_1+\cdots+\lambda_{\mathrm{k}}\dot{\mathrm{p}}_{\mathrm{k}}\mid\sum_{\mathrm{i}=1}^{\mathrm{k}}\lambda_{\mathrm{i}}=1,\ \lambda_{\mathrm{i}}\geqslant 0,\ \mathrm{i}=1,\dots,\mathrm{k}\}.$$

我们把满足 $\sum_{i=1}^k\lambda_i=1$，$\lambda_i\geqslant 0$，$i=1,\dots,k$ 的仿射组合 $\lambda_1\dot p_1+\cdots+\lambda_k\dot p_k$ 称为**凸组合** (convex combination)。

证明. 对 k 做数学归纳法。$k=2$ 时即凸集定义；若定理对 k 成立，则对 $k+1$ 情形，注意到 $\mathrm{C}(\{\dot{\mathrm{p}}_1,\dots,\dot{\mathrm{p}}_{\mathrm{k}+1}\})=\mathrm{C}(\mathrm{C}(\{\dot{\mathrm{p}}_1,\dots,\dot{\mathrm{p}}_{\mathrm{k}}\}\cup\{\dot{\mathrm{p}}_{\mathrm{k}+1}\}))$ (利用凸包的定义验证互相包含)，再利用上面的定理，整理形式后即可得到结论。 □

下面我们考虑一些基本的凸集：超平面、半空间和凸多面体。

定义 4.3.4. 设 $f:\mathbb{A}\to\mathbb{R}$ 是实仿射空间 $\mathbb{A}$ 上的仿射线性函数，记

$$\begin{aligned}
H_f &= \{\dot{p}\in\mathbb{A}\mid f(\dot{p})=0\},\\
H_f^+ &= \{\dot{p}\in\mathbb{A}\mid f(\dot{p})\geqslant 0\},\\
H_f^- &= \{\dot{p}\in\mathbb{A}\mid f(\dot{p})\leqslant 0\}.
\end{aligned}$$

则容易验证 H_f 是凸的超平面，且 H_f^+, H_f^- 都是闭的凸集。我们把 H_f^+ 和 H_f^- 称为以 H_f 为边界的 (闭) **半空间** (half-space)。

有限多个半空间的交集被称为凸多面体 (polyhedron)。

定理 4.3.3 (Minkowski-Weyl). $P\subset\mathbb{A}$ 是有界的凸多面体当且仅当 P 是有限多个点的凸包。

一个比较简单的证明可以参考《Lectures on Polytopes》, Günter M. Ziegler, GTM152 的第 1.1-1.3 节，我们不在这里给出证明细节。

下面的定理是线性规划的理论基础。

定理 4.3.4. 设 f 是实仿射空间 $\mathbb{A}$ 上的仿射线性函数，$P=\mathrm{C}(\{\dot{\mathrm{p}}_0,\dots,\dot{\mathrm{p}}_\mathrm{m}\})$，则 f 在 P 上的极大值一定能在某个 $\dot{p}_i$, $i\in\{0,\dots,m\}$ 处取到。

证明. 对 m 做数学归纳法。$m=0$ 时定理显然成立；下设 $m\geqslant 1$ 且定理对 $m-1$ 的情形成立，我们来证明定理对 m 的情形也成立。事实上，不妨记 $\mathrm{C}(\{\dot{\mathrm{p}}_0,\dots,\dot{\mathrm{p}}_{\mathrm{m}-1}\})=\mathrm{M}$ 且设 f 在 M 上的最大值在 $\dot{p}_0$ 处取到，即

$$\max_{\dot{p}\in M}f(\dot{p})=f(\dot{p}_0).$$

由定理4.3.1可知 $P=\bigcup\limits_{\dot{p}\in M}\dot{p}\dot{p}_m$，于是任取一点 $\dot{q}\in P$，我们总有：存在 $\dot{p}\in M$ 使得

$$\dot{q}=\lambda\dot{p}+(1-\lambda)\dot{p}_m,\ \lambda\in[0,1].$$

于是

$$\begin{aligned}
f(\dot{q}) &= \lambda f(\dot{p})+(1-\lambda)f(\dot{p}_m)\\
&\leqslant \max\{f(\dot{p}),\ f(\dot{p}_m)\}\\
&\leqslant \max\{f(\dot{p}_0),\ f(\dot{p}_m)\}.
\end{aligned}$$

此即定理对 m 也成立。这样我们就完成了证明。 □

有关线性规划的更多理论及其在实际生产生活中的应用，可以参考《运筹学教程》，胡运权，清华大学出版社。

4.4　伪欧几里得空间简介

欧氏空间本质上就是带有正定的对称双线性型的实向量空间，如果我们去掉其中正定的要求，则得到了伪欧氏空间。伪欧氏空间在物理中有重要的应用。

定义 4.4.1. 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，$<\cdot,\cdot>$ 是 V 上非退化的对称双线性型，其正惯性指数为 $s<n$。则由定理1.6.8可知，存在 V 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得对任意的 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{i=1}^n y_i\mathbf{e}_i\in V$，都有

$$<\mathbf{x},\mathbf{y}>=\sum_{i=1}^{s}x_iy_i-\sum_{j=s+1}^{n}x_jy_j.$$

我们把 $(V,<\cdot,\cdot>)$ 称为符号差为 $(s,n-s)$ 的伪欧几里得向量空间，与 $(V,<\cdot,\cdot>)$ 相伴的仿射空间称为伪欧几里得仿射空间。

设 $\mathbb{E}$ 是与 $(V,<\cdot,\cdot>)$ 相伴的 n 维伪欧几里得仿射空间，取定坐标系 $(\dot{o},\mathbf{e}_1,\dots,\mathbf{e}_n)$ 后点 $\dot{p},\dot{q}$ 的坐标分别是 $(x_1,\dots,x_n)^t$ 和 $(y_1,\dots,y_n)^t$，则我们仍可定义点 $\dot{p},\dot{q}$ 的“距离”的平方为

$$\rho(\dot{p},\dot{q})^2=<\overline{pq},\overline{pq}>=\sum_{i=1}^{s}(y_i-x_i)^2-\sum_{j=s+1}^{n}(y_j-x_j)^2.$$

设 $(V,<\cdot,\cdot>)$ 是符号差为 $(s,n-s)$ 的伪欧几里得向量空间，考虑 V 上的保持伪“内积”(即对称双线性型) 的线性算子 $\mathcal{F}$(称为伪正交算子)，设 $\mathcal{F}$ 在 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵为 F，容易验证 $\mathcal{F}\mathbf{e}_1,\dots,\mathcal{F}\mathbf{e}_n$ 也是 V 的一组基，并且双线性型 $<\cdot,\cdot>$ 在 $\mathcal{F}\mathbf{e}_1,\dots,\mathcal{F}\mathbf{e}_n$ 下的矩阵仍为 $\begin{pmatrix}E_s & 0\\ 0 & E_{n-s}\end{pmatrix}$(细节留作练习)，因此由定理1.6.2，我们有

$$F^t\begin{pmatrix}E_s & 0\\ 0 & E_{n-s}\end{pmatrix}F=\begin{pmatrix}E_s & 0\\ 0 & E_{n-s}\end{pmatrix}.$$

于是类似于推论3.1.3，容易验证 (留作练习)V 上所有伪正交算子在映射复合下构成一个群，称之为伪正交群，记作 $O(s,n-s)$。

设 $\mathbb{E}$ 是与 $(V,<\cdot,\cdot>)$ 相伴的 n 维伪欧几里得仿射空间，f 是 $\mathbb{E}$ 上的仿射变换。如果 f 保持伪欧氏距离不变，即 $\forall\dot{p},\dot{q}\in\mathbb{E}$，都有

$$\rho(f(\dot{p}),f(\dot{q}))^2=\rho(\dot{p},\dot{q})^2,$$

则我们称 f 是一个伪正交变换或者运动。容易验证 $\mathbb{E}$ 上的所有伪正交变换在映射复合的运算下构成一个群，我们仍把它记作 $\mathrm{Iso}(\mathbb{E})$。类似于定理4.2.3，容易验证一个伪欧氏仿射空间中的运动一定可以分解成一个平移与一个伪正交算子的复合。

特别地，我们把正惯性指数 $s=1$ 的伪欧几里得向量空间称为闵可夫斯基 (Minkowski) 空间。四维的 Minkowski 空间在狭义相对论中非常重要。我们把四维 Minkowski 空间的伪正交群 $O(1,3)$ 称为洛伦兹 (Lorentz) 群。有关狭义相对论的内容请参考物理学的标准教材。

4.5　习题

仿射空间

1. 证明: 平面 $\Pi' = \dot{p} + U'$ 和 $\Pi'' = \dot{q} + U''$ 的交集非空等价于 $\overline{pq} \in U' + U''$.

2. 设 $\mathbb{A}(\Pi_1, \cdots, \Pi_m)$ 是 n 维仿射空间 $\mathbb{A}$ 中直线 $\Pi_1, \cdots, \Pi_m$ 的仿射包络. 如果 $\mathbb{A}(\Pi_1, \cdots, \Pi_m) = \mathbb{A}$, 那么 m 的极小值是多少?

3. 设 $f : \mathbb{A} \to \mathbb{A}'$ 是仿射映射, Π 和 Π' 分别是 $\mathbb{A}$ 和 $\mathbb{A}'$ 中的平面. 证明:
   (1) $f(\Pi)$ 是 $\mathbb{A}'$ 中的平面;
   (2) 如果 $f^{-1}(\Pi')$ 非空, 那么 $f^{-1}(\Pi')$ 是 $\mathbb{A}$ 中的平面.

4. 设 $\dot{p}_0, \ldots, \dot{p}_m$ 是仿射空间中的点. 证明: (1) 这些点仿射无关当且仅当

$$\dot{p}_i \notin \mathbb{A}(\{\dot{p}_0, \ldots, \dot{p}_m\} \setminus \{\dot{p}_i\}), \quad i = 0, 1, \ldots, m;$$

   (2) 这些点仿射无关当且仅当 $\dot{p}_i \notin A(\dot{p}_0, \ldots, \dot{p}_{i-1}), i = 1, \ldots, m$.

5. 如果仿射空间中有两个二维平面是偏斜的, 这个仿射空间的最小维数是多少?

欧几里得仿射空间

1. 证明下面关于平面对 Π, Π' 的两个性质是等价的:
   (1) 一个平面上的任何直线与另一个平面上的任何直线垂直;
   (2) 平面 Π, Π' 互相垂直, 且至多有一个交点.

2. 设 $\Pi \subset \Gamma$ 是欧几里得空间 $\mathbb{E}$ 的两个平面. 证明: 如果平面 $\Gamma' \subset \mathbb{E}$ 与 Γ 垂直, 且 $\Gamma' \cap \Gamma = \Pi$, 那么 $\dim \Gamma' \leqslant \dim \mathbb{E} - \dim \Gamma + \dim \Pi$. 存在唯一的平面, 它具有这个性质, 且维数等于 $\dim \mathbb{E} - \dim \Gamma + \dim \Pi$.

3. 求出点 $p = (2, 1, -3, 4)$ 到平面

$$\Pi : 2x - 4y - 8z + 13w + 19 = 0, \quad x + y - z + 2w - 1 = 0$$

   的距离.

4. 求出平面

$$\Pi_1 : x_1 + x_2 + x_4 - 2x_5 = 2, \quad x_2 + x_3 - x_4 - x_5 = 3, \quad x_1 - x_2 + 2x_3 - x_5 = 3$$

   和平面

$$\Pi_2 : (1, -2, 5, 8, 2) + \text{span}\,\{(0, 1, 2, 1, 2), (2, 1, 2, -1, 1)\}$$

   之间的距离.

5. 设 $\Pi_1, \Pi_2, \Pi_1', \Pi_2' \subset A$ 是平面, 方向子空间分别是 U_1, U_2, U_1', U_2'. 证明: 如果 $\dim \Pi_1 = \dim \Pi_1', \dim \Pi_2 = \dim \Pi_2', \dim U_1 \cap U_2 = \dim U_1' \cap U_2'$, 且交集 $\Pi_1 \cap \Pi_2$ 和 $\Pi_1' \cap \Pi_2'$ 均为空集或均为非空集, 那么存在仿射变换 $f \in \text{Aff}(\mathbb{A})$, 它把 Π_1 映到 Π_1', Π_2 映到 Π_2'.

6. 描述欧几里得平面上绕不同的点的旋转 f 和 g 的合成是怎样的运动.

7. 证明: 如果欧几里得空间的保距变换保持两个偏斜的平面不变, 且这两个平面的方向子空间的交集是零子空间, 那么它有不动点.

凸集

1. 证明: 凸集在仿射映射下的像是凸集.

2. 证明: 凸集在仿射映射下的逆像是凸集.

3. 欧几里得向量空间 V 同时看作仿射空间, 设 $C \subset V$ 是凸集, 它的对偶 $C^\vee$ 定义为

$$C^\vee = \{\mathbf{u} \in V \mid < \mathbf{u}, \mathbf{v} > \leqslant 1, \forall \mathbf{v} \in C\}.$$

   证明:
   (1) $C^\vee$ 是凸集;
   (2) 如果 C 的内部含有 $\mathbf{0}$, 那么 $C^\vee$ 是有界的; (称集合 S 中的点 p 在 S 中的内部如果 S 包含某个含 p 的开球.)
   (3) 如果 C 有界, 那么 $C^\vee$ 的内部含有 $\mathbf{0}$;
   (4) 如果 C 是有界闭集且 $\mathbf{0}$ 在 C 的内部, 那么 C 是 $C^\vee$ 的对偶; (称集合 $S \subset V$ 是闭集如果 S 中的任何柯西序列的极限都在 S 中.)
   (5) 如果 P 是凸多面体且 $\mathbf{0}$ 在其内部, 那么 $P^\vee$ 也是凸多面体.

第五章　解析几何 (II)：二次曲面与射影空间

二次曲面是数学研究和实际生活中的常见对象。我们现在已经熟悉了二次型和仿射空间，于是可以用这些工具来对 (n 维) 仿射空间上的二次曲面进行分类和研究了。而在对二次曲面的研究中，我们会发现，在投影的观点下二次曲面的分类会简单许多，由此引出了射影空间的概念。相比于前面几章而言，本章的内容更加具体而直观，读者也可以从此看出代数工具在几何研究中的重要意义。

5.1　二次曲面的仿射分类

设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbb{A}$ 是与 V 相伴的仿射空间，则 $\mathbb{A}$ 与 V 作为自身上的仿射空间同构，而 V 与 $\mathbb{K}^n$ 同构，因此我们只需要研究仿射空间 $\mathbb{K}^n$ 中的二次曲面就足够了。

定义 5.1.1. 设对称矩阵 $A=(a_{ij})_{n\times n}\in SM_n(\mathbb{K})$，向量 $\mathbf{b}=(b_1,\dots,b_n)^t\in\mathbb{K}^n$ 常数及 $c\in\mathbb{K}$ 已知，未定向量 $\mathbf{x}=(x_1,\dots,x_n)^t\in\mathbb{K}^n$。我们把满足下面的二次方程

$$\sum_{i,j=1}^{n}a_{ij}x_ix_j+2\sum_{i=1}^{n}b_ix_i+c=0,$$

即

$$\mathbf{x}^tA\mathbf{x}+2\mathbf{b}^t\mathbf{x}+c=0 \tag{5.1.1}$$

的所有点 $\dot{o}+\mathbf{x}$ 组成的几何称为二次曲面 (quadrics)。如果一个二次曲面之间可以通过仿射变换得到另一个二次曲面，则我们称这两个二次曲面仿射等价 (验证这是一个等价关系留作练习)。

为了便于计算，我们只考虑实数域和复数域上的二次曲面。我们希望通过合适的仿射变换将方程 (5.1.1) 化成比较简单的形式，从而得到二次曲面的仿射分类。

首先，方程 (5.1.1) 可以写成

$$\begin{pmatrix}\mathbf{x}^t & 1\end{pmatrix}\begin{pmatrix}A & \mathbf{b}\\ \mathbf{b}^t & c\end{pmatrix}\begin{pmatrix}\mathbf{x}\\ 1\end{pmatrix}=0. \tag{5.1.2}$$

如果我们做适当的适当的仿射变换使得 $\mathbf{x}=P\mathbf{y}+\mathbf{v}$，$P\in\mathrm{GL}_n(\mathbb{K})$，则方程 (5.1.2) 变为

$$\begin{pmatrix}\mathbf{y}^tP^t+\mathbf{v} & 1\end{pmatrix}\begin{pmatrix}A & \mathbf{b}\\ \mathbf{b}^t & c\end{pmatrix}\begin{pmatrix}P\mathbf{y}+\mathbf{v}\\ 1\end{pmatrix}=0.$$

整理后即

$$\begin{pmatrix}\mathbf{y}^t & 1\end{pmatrix}\begin{pmatrix}P^tAP & P^tA\mathbf{v}+P^t\mathbf{b}\\ \mathbf{v}^tAP+\mathbf{b}^tP & \mathbf{v}^t\mathbf{v}+c\end{pmatrix}\begin{pmatrix}\mathbf{y}\\ 1\end{pmatrix}=0.$$

我们知道对称矩阵可以合同等价到对角矩阵，于是如果我们选取合适的 P 和 $\mathbf{v}$ 使得 P^tAP 是对角阵而 $P^tA\mathbf{v}+P^t\mathbf{b}$ 中的 0 尽可能多，那么方程 (5.1.2) 就得到了化简。这个过程具体写出来就是以下定理：

定理 5.1.1. 设 Φ 是 $\mathbb{C}^n$ 中由方程 $\mathbf{x}^tA\mathbf{x}+2\mathbf{b}^t\mathbf{x}+c=0$ 定义的二次曲面。记 $r=\operatorname{rank}(A)$，$r_1=\operatorname{rank}\begin{pmatrix}A & \mathbf{b}\\ \mathbf{b}^t & c\end{pmatrix}$，则 Φ 一定仿射等价于以下二次曲面之一：

(1) 若 $r_1=r$，则 Φ 仿射等价于由方程 $y_1^2+\cdots+y_r^2=0$ 定义的二次曲面；

(2) 若 $r_1=r+1$，则 Φ 仿射等价于由方程 $y_1^2+\cdots+y_r^2+1=0$ 定义的二次曲面；

(3) 若 $r_1=r+2$，则 Φ 仿射等价于由方程 $y_1^2+\cdots+y_r^2+2y_{r+1}=0$ 定义的二次曲面。

证明. 由 1.6.6 小节开头的讨论可知，我们可以找到 $P\in\mathrm{GL}_n(\mathbb{C})$ 使得 $P^tAP=\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}_{n\times n}$，下面我们考虑平移 $\mathbf{v}$ 的选取。令 $\mathbf{x}=P\mathbf{z}$，则 $\mathbf{z}$ 满足方程

$$\begin{pmatrix}\mathbf{z}^t & 1\end{pmatrix}\begin{pmatrix}\begin{pmatrix}E_r & O\\ O & O\end{pmatrix} & P^t\mathbf{b}\\ \mathbf{b}^tP & c\end{pmatrix}\begin{pmatrix}\mathbf{z}\\ 1\end{pmatrix}=0.$$

记 $P^t\mathbf{b}=(d_1,\ldots,d_n)^t$，则上式也可以写成

$$z_1^2+\cdots+z_r^2+2\sum_{i=1}^n d_iz_i+c=0.$$

对其配方得

$$(z_1+d_1)^2+\cdots+(z_r+d_r)^2+2\sum_{j=r+1}^n d_jz_j+(c-\sum_{i=1}^r d_r^2)=0 \tag{5.1.3}$$

令 $\mathbf{y}=(z_1+d_1,\ldots,z_r+d_r,z_{r+1},\ldots,z_n)^t$，即 $\mathbf{y}=\mathbf{z}+\mathbf{f}$，其中 $\mathbf{f}=(d_1,\ldots,d_r,0,\ldots,0)^t$，则方程变为

$$\begin{pmatrix}\mathbf{y}^t & 1\end{pmatrix}\underbrace{\begin{pmatrix}1 & & & & & & 0\\ & \ddots & & & & & \vdots\\ & & 1 & & & & 0\\ & & & 0 & & & d_{r+1}\\ & & & & \ddots & & \vdots\\ & & & & & 0 & d_n\\ 0 & \cdots & 0 & d_{r+1} & \cdots & d_n & c-\sum_{i=1}^r d_r^2\end{pmatrix}}_{\text{记为}A'}\begin{pmatrix}\mathbf{y}\\ 1\end{pmatrix}=0.$$

由于 $r_1\geqslant\operatorname{rank}(A)$ 且

$$r_1=\operatorname{rank}\begin{pmatrix}A & \mathbf{b}\\ \mathbf{b}^t & c\end{pmatrix}\leqslant\operatorname{rank}\begin{pmatrix}A & \mathbf{b}\end{pmatrix}+1\leqslant\operatorname{rank}(A)+2,$$

因此我们有且仅有以下三种情形：

(1) 如果 $r_1=r$，则必有 $d_{r+1}=\cdots=d_n=0$ 且 $c-\sum_{i=1}^r d_r^2=0$，否则 $\operatorname{rank}\begin{pmatrix}A & \mathbf{b}\\ \mathbf{b}^t & c\end{pmatrix}=\operatorname{rank}(A')\geqslant r+1$，矛盾！因此，此时原二次曲面通过仿射变换 $\mathbf{x}=P\mathbf{z}=P\mathbf{y}-P\mathbf{f}$ 得到了

$$y_1^2+\cdots+y_r^2=0$$

即为所需。

(2) 如果 $r_1 = r+1$，则只能是 $d_{r+1} = \cdots = d_n = 0$ 且 $c - \sum_{i=1}^{r} d_r^2 \neq 0$，因为如果某个 d_j，$j \in \{r+1, \ldots, n\}$ 不为 0，则 $\operatorname{rank}(A')$ 至少比 $\operatorname{rank}(A)$ 多 2。此时我们可以继续作伸缩变换

$$\mathbf{y} = \overset{\overbrace{\hphantom{\sqrt{c-\sum_{i=1}^{r} d_r^2}\qquad\ddots\qquad\sqrt{c-\sum_{i=1}^{r} d_r^2}}}^{r\text{个}}}{\begin{pmatrix} \sqrt{c-\sum_{i=1}^{r} d_r^2} & & & & & \\ & \ddots & & & & \\ & & \sqrt{c-\sum_{i=1}^{r} d_r^2} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}} \cdot \mathbf{w}$$

则此时原方程变为

$$(c - \sum_{i=1}^{r} d_r^2)(\mathbf{w}^t\mathbf{w} + 1) = 0, \text{ 即}\mathbf{w}^t\mathbf{w} + 1 = 0,$$

即为所需。

(3) 如果 $r_1 = r+2$，即 d_j，$j \in \{r+1, \ldots, n\}$ 不全为 0，则我们令

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \overset{\overbrace{\hphantom{1\quad\ddots\quad}}^{r\text{个}1}}{\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ 0 & \cdots & 0 & d_{r+1} & d_{r+2} & \cdots & d_n \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}} \cdot \begin{pmatrix} z_1 \\ \vdots \\ z_r \\ z_{r+1} \\ z_{r+2} \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} d_1 \\ \vdots \\ d_r \\ \frac{1}{2}(c - \sum_{i=1}^{r} d_r^2) \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

代入方程 (5.1.3)，则此时原方程变为

$$u_1^2 + \cdots + u_r^2 + 2u_{r+1} = 0,$$

即为所需。这样我们就完成了证明。

□

类似地，对 $\mathbb{R}^n$ 上的二次曲面，我们有以下的仿射分类：

定理 5.1.2. 设 Φ 是 $\mathbb{R}^n$ 中由方程 $\mathbf{x}^tA\mathbf{x} + 2\mathbf{b}^t\mathbf{x} + c = 0$ 定义的二次曲面。设 A 的正惯性指数为 s，记 $r = \operatorname{rank}(A)$，$r_1 = \operatorname{rank}\begin{pmatrix} A & \mathbf{b} \\ \mathbf{b}^t & c \end{pmatrix}$，则 Φ 一定仿射等价于以下二次曲面之一：

(1) 若 $r_1 = r$，则 Φ 仿射等价于由方程 $y_1^2 + \cdots + y_s^2 - y_{s+1}^2 - \cdots - y_r^2 = 0$ 定义的二次曲面；

(2) 若 $r_1 = r+1$，则 Φ 仿射等价于由方程 $y_1^2 + \cdots + y_s^2 - y_{s+1}^2 - \cdots - y_r^2 = \pm 1$ 定义的二次曲面；

(3) 若 $r_1 = r+2$，则 Φ 仿射等价于由方程 $y_1^2+\cdots+y_s^2-y_{s+1}^2-\cdots-y_r^2+2y_{r+1}=0$ 定义的二次曲面。

证明与定理5.1.1是类似的，只需注意到 A 合同等价于 $\begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}$ 即可 (定理1.6.8)，细节留作练习。

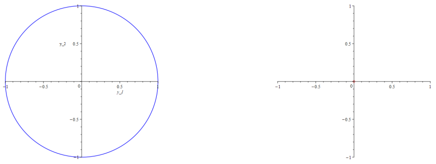
例 5.1.1. 下面我们考察 $\mathbb{R}^2$ 上二次曲面 (即二次曲线) 的仿射分类。$\mathbb{R}^2$ 上的二次曲线可以写成

$$\sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}x_ix_j + 2\sum_{i=1}^{2} b_ix_i + c = 0$$

的形式。记 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$，则

(1) 当 A 合同等价于 $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ 时，则曲线仿射等价于以下三种情形之一：

- $y_1^2+y_2^2=1$，单位圆；
- $y_1^2+y_2^2=0$，原点；
- $y_1^2+y_2^2=-1$，空集；



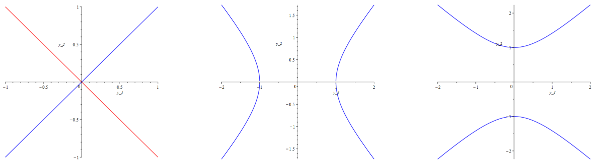
(a) $y_1^2+y_2^2=1$，单位圆

(b) $y_1^2+y_2^2=0$，原点

图 5.1-1

(2) 当 A 合同等价于 $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ 时，则曲线仿射等价于以下两种情形之一：

- $y_1^2-y_2^2=0$，两条相交直线；
- $y_1^2-y_2^2=\pm1$，双曲线；



(a) $y_1^2-y_2^2=0$，两条相交直线

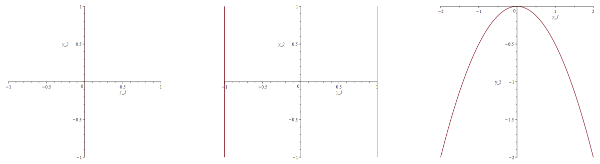
(b) $y_1^2-y_2^2=1$，双曲线

(c) $y_1^2-y_2^2=-1$，双曲线

图 5.1-2

(3) 当 A 合同等价于 $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$ 或 $\begin{pmatrix} -1 & \\ & 0 \end{pmatrix}$ 时，则曲线仿射等价于以下四种情形之一：

- $y_1^2=0$，一条直线；
- $y_1^2=1$，两条平行直线；
- $y_1^2=-1$，空集；
- $y_1^2+2y_2=0$，抛物线。



(a) $y_1^2=0$，一条直线

(b) $y_1^2=1$，两条平行直线

(c) $y_1^2+2y_2=0$，抛物线

图 5.1-3

至此我们就完成了二次曲线的分类。

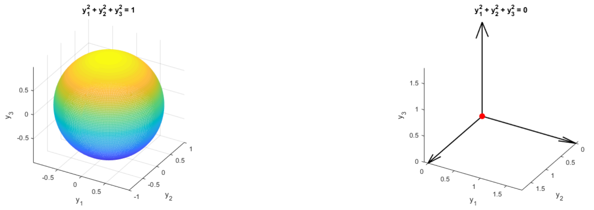
例 5.1.2. 在本节的最后，我们考察 $\mathbb{R}^3$ 上二次曲面的仿射分类。$\mathbb{R}^2$ 上二次曲面的仿射分类。$\mathbb{R}^3$ 上的二次曲线可以写成

$$\sum_{i=1}^{3}\sum_{j=1}^{2}a_{ij}x_ix_j+2\sum_{i=1}^{3}b_ix_i+c=0$$

的形式。记 $A=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}$，则

(1) 当 A 合同等价于 $\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix}$ 时，则曲线仿射等价于以下三种情形之一：

- $y_1^2+y_2^2+y_3^2=1$，单位球面 (sphere surface)；
- $y_1^2+y_2^2+y_3^2=0$，原点 (origin)；
- $y_1^2+y_2^2+y_3^2=-1$，空集；



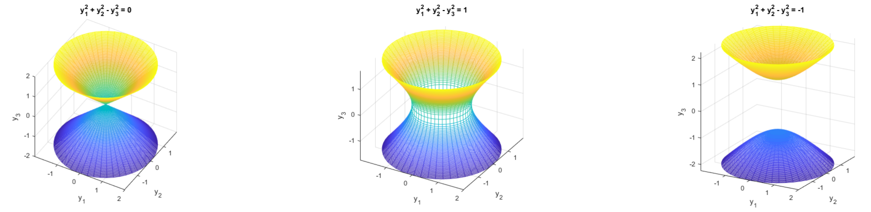
(a) $y_1^2+y_2^2+y_3^2=1$，单位球面

(b) $y_1^2+y_2^2+y_3^2=0$，原点

图 5.1-4

(2) 当 A 合同等价于 $\begin{pmatrix}1&&\\&1&\\&&-1\end{pmatrix}$ 或 $\begin{pmatrix}1&&\\&-1&\\&&-1\end{pmatrix}$ 时，则曲线仿射等价于以下三种情形之一：

- $y_1^2 + y_2^2 - y_3^2 = 0$，圆锥面 (conical surface)；
- $y_1^2 + y_2^2 - y_3^2 = 1$，单叶双曲面 (hyperboloid of one sheet)；
- $y_1^2 + y_2^2 - y_3^2 = -1$，双叶双曲面 (hyperboloid of two sheets)；

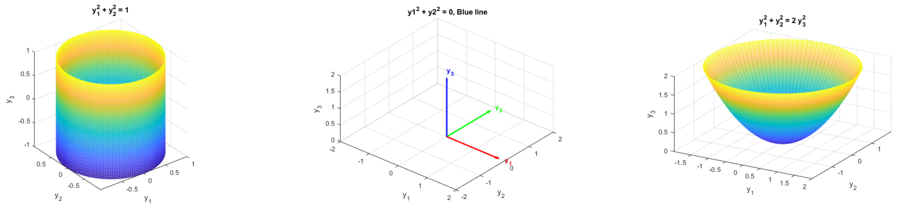


(a) $y_1^2 + y_2^2 - y_3^2 = 0$，圆锥面　(b) $y_1^2 + y_2^2 - y_3^2 = 1$，单叶双曲面　(c) $y_1^2 + y_2^2 - y_3^2 = -1$，双叶双曲面

图 5.1-5

(3) 当 A 合同等价于 $\begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}$ 或 $\begin{pmatrix} -1 & & \\ & -1 & \\ & & 0 \end{pmatrix}$ 时，则曲线仿射等价于以下四种情形之一：

- $y_1^2 + y_2^2 = 1$，圆柱面 (elliptic cylinder)；
- $y_1^2 + y_2^2 = 0$，一条直线 (y_3 坐标轴)；
- $y_1^2 + y_2^2 = -1$，空集；
- $y_1^2 + y_2^2 = 2y_3$，椭圆抛物面 (elliptic paraboloid)；



(a) $y_1^2 + y_2^2 = 1$，圆柱面　(b) $y_1^2 + y_2^2 = 0$，一条直线 (y_3 坐标轴)　(c) $y_1^2 + y_2^2 = 2y_3$，椭圆抛物面

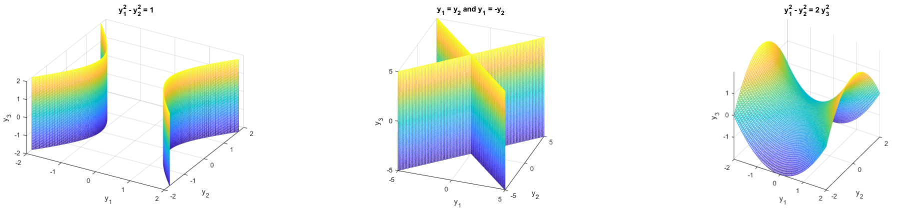
图 5.1-6

(4) 当 A 合同等价于 $\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}$ 时，则曲线仿射等价于以下三种情形之一：

- $y_1^2 - y_2^2 = \pm 1$，双曲柱面 (hyperbolic cylinder)；
- $y_1^2 - y_2^2 = 0$，两个相交平面；
- $y_1^2 - y_2^2 = 2y_3$，双曲抛物面 (马鞍面，hyperbolic paraboloid)；

(5) 当 A 合同等价于 $\begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}$ 或 $\begin{pmatrix} -1 & & \\ & 0 & \\ & & 0 \end{pmatrix}$ 时，则曲线仿射等价于以下四种情形之一：

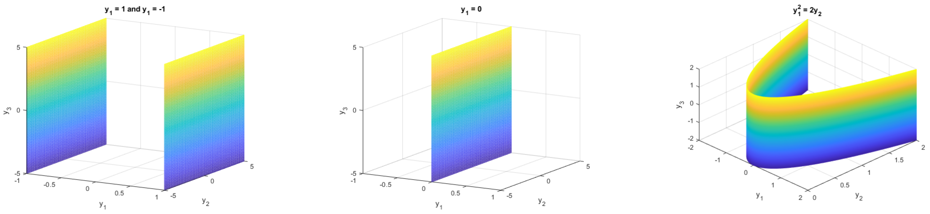
- $y_1 = \pm 1$，两个平行平面；



(a) $y_1^2 - y_2^2 = 1$，双曲柱面 (b) $y_1^2 - y_2^2 = 0$，两个相交平面 (c) $y_1^2 - y_2^2 = 2y_3$，双曲抛物面 (马鞍面)

图 5.1-7

- $y_1^2 = 0$，两个重合平面 (即 y_2Oy_3 平面)；
- $y_1^2 = -1$，两个虚平面 (在实数域上看就是空集)；
- $y_1^2 = 2y_2$，抛物柱面 (parabolic cylinder)。



(a) $y_1 = \pm 1$，两个平行平面 (b) $y_1^2 = 0$，两个重合平面 (c) $y_1^2 = 2y_2$，抛物柱面

图 5.1-8

这就是 $\mathbb{R}^3$ 中二次曲面的所有分类 (共 17 种)。

5.2 二次曲面的正交分类

现在我们来考虑标准欧氏空间 $\mathbb{R}^n$(重点是 $\mathbb{R}^3$) 中二次曲面在保距变换下的分类。也就是说，本节中我们仅允许对曲面做正交变换和平移。我们有以下的定理：

定理 5.2.1. 设 Φ 是 $\mathbb{R}^n$ 中由方程 $\mathbf{x}^t A\mathbf{x}+2\mathbf{b}^t\mathbf{y}+c=0$ 定义的二次曲面，同样记 $r=\operatorname{rank}(A)$, $r_1=\operatorname{rank}\begin{pmatrix} A & \mathbf{b} \\ \mathbf{b}^t & c\end{pmatrix}$, A 的正惯性指数为 s，则 Φ 一定可以通过正交变换和平移得到以下二次曲面之一：

(1) 若 $r_1=r$，则 Φ 可以变换到由方程 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}=0$ 得到的二次曲面；

(2) 若 $r_1=r+1$，则 Φ 可以变换到由方程 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}=\pm1$ 得到的二次曲面；

(3) 若 $r_1=r+2$，则 Φ 可以变换到由方程 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}+2x_{r+1}=0$ 得到的二次曲面。

其中 $a_1,\ldots,a_r$ 是正实数。我们称上面的方程为二次曲面的标准方程。

证明. 由于 A 是对称矩阵，利用定理3.3.4我们可以得到：存在正交矩阵 P 使得

$$P^tAP=P^{-1}AP=\begin{pmatrix} \lambda_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && \lambda_s &&&&&& \\ &&& -\lambda_{s+1} &&&&& \\ &&&& \ddots &&&& \\ &&&&& -\lambda_r &&& \\ &&&&&& 0 && \\ &&&&&&& \ddots & \\ &&&&&&&& 0 \end{pmatrix}.$$

令 $\mathbf{x}=P\mathbf{y}$，类似于定理5.1.1的证明过程，此时方程变为

$$\lambda_1y_1^2+\cdots+\lambda_sy_s^2-\lambda_{s+1}y_{s+1}^2-\lambda_ry_r^2+2\sum_{i=1}^n d_iy_i+c=0.$$

对其配方可得

$$\lambda_1(y_1+\frac{d_1}{\lambda_1})^2+\cdots+\lambda_s(y_s+\frac{d_s}{\lambda_s})^2-\lambda_1(y_{s+1}+\frac{d_{s+1}}{\lambda_{s+1}})^2-\cdots-\lambda_r(y_r+\frac{d_r}{\lambda_r})^2+2\sum_{i=r+1}^n d_iy_i+\left(c-\sum_{i=1}^s\frac{d_i^2}{\lambda_i}+\sum_{i=s+1}^r\frac{d_i^2}{\lambda_i}\right)=0.$$

然后通过平移可以得到：

$$\lambda_1z_1^2+\cdots+\lambda_sz_s^2-\lambda_{s+1}z_{s+1}^2-\lambda_rz_r^2+2\sum_{i=r+1}^n d_iz_i+c'=0.$$

于是类似于定理5.1.1的证明过程，我们分以下情况讨论 (细节不再赘述，留作练习)：

(1) 如果 $r_1 = r$，则 $d_{r+1} = \cdots = d_n = 0$ 且 $c' = 0$，令 $a_i = \dfrac{1}{\sqrt{\lambda_i}}$，$i = 1, \ldots, r$ 即得

$$\frac{z_1^2}{a_1^2} + \cdots + \frac{z_s^2}{a_s^2} - \frac{z_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{z_r^2}{a_r^2} = 0.$$

(2) 如果 $r_1 = r + 1$，则 $d_{r+1} = \cdots = d_n = 0$ 且 $c' \neq 0$，令 $a_i = \sqrt{\dfrac{|c'|}{\lambda_i}}$，$i = 1, \ldots, r$ 即得

$$\frac{z_1^2}{a_1^2} + \cdots + \frac{z_s^2}{a_s^2} - \frac{z_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{z_r^2}{a_r^2} = \pm 1.$$

(3) 如果 $r_1 = r + 2$，即 d_j，$j \in \{r+1, \ldots, n\}$ 不全为 0，则我们再作如下变换：记 $\mu = \sqrt{b_{r+1}^2 + \cdots + b_n^2}$，令 $w_1 = z_1, \cdots, w_r = z_r$，并且

$$\begin{pmatrix} w_{r+1} \\ w_{r+2} \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} \frac{d_{r+1}}{\mu} & \cdots & \frac{d_n}{\mu} \\ & & \\ & * & \\ & & \end{pmatrix} \cdot \begin{pmatrix} z_{r+1} \\ z_{r+2} \\ \vdots \\ z_n \end{pmatrix} + \begin{pmatrix} \frac{c'}{2\mu} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.2.1}$$

其中我们需要使 $\begin{pmatrix} \frac{b_{r+1}}{\mu} & \cdots & \frac{b_n}{\mu} \\ & & \\ & * & \\ & & \end{pmatrix}$ 为正交矩阵（打 $*$ 的部分可以先任意选取，然后用 Gram-Schmidt 正交化得到正交矩阵)。式 (5.2.1) 表明：

$$\sum_{i=r+1}^{n} d_i z_i + c' = 2\mu w_{r+1}.$$

将上式代回曲面方程，并记 $a_i = \sqrt{\dfrac{\mu}{\lambda_i}}$，$i = 1, \ldots, r$，则二次曲面的方程变为：

$$\frac{w_1^2}{a_1^2} + \cdots + \frac{w_s^2}{a_s^2} - \frac{w_{s+1}^2}{a_{s+1}^2} - \cdots - \frac{w_r^2}{a_r^2} + 2w_{r+1} = 0.$$

至此我们就得到了二次曲面在保距变换下的所有分类。 □

由上面的定理我们很容易看出，当 $\operatorname{rank}\begin{pmatrix} A & \mathbf{b} \\ \mathbf{b}^t & c \end{pmatrix} \geqslant n$ 时，二次曲面的标准方程中含有全部 n 个变元，此时我们称二次曲面是非退化的。对于非退化的二次曲面 (不考虑单点集或空集)，我们可以进一步定义：

定义 5.2.1. 条件与记号同定理5.2.1。

(1) 如果 $s = r = n$，$r_1 = n + 1$，即二次曲面的标准方程为 $\dfrac{x_1^2}{a_1^2} + \cdots + \dfrac{x_n^2}{a_n^2} = 1$，则我们称该二次曲面为椭球面；

(2) 如果 $0 < s < r = n\ r_1 = n+1$，即二次曲面的标准方程为 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=\pm 1$，则我们称该二次曲面为双曲面；

(3) 如果 $0 < s < r = n\ r_1 = n$，即二次曲面的标准方程为 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=0$，则我们称该二次曲面为二次锥面；

(4) 如果 $s = r = n-1$，即二次曲面的标准方程为 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_{n-1}^2}{a_{n-1}^2}+2x_n=0$，则我们称该二次曲面为椭圆抛物面；

(5) 如果 $0 < s < r = n-1$，即二次曲面的标准方程为 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_{n-1}^2}{a_{n-1}^2}+2x_n=0$，则我们称该二次曲面为双曲抛物面。

在本节的最后，我们简单介绍一些二次曲面的几何性质。我们主要考虑 $\mathbb{R}^3$ 中的二次曲面，更一般的情形可以类比得到。

首先，比较定理5.2.1和定理5.1.2可知，$\mathbb{R}^3$ 中二次曲面的仿射分类与正交分类是一致的 (17 种情形，例5.1.2)，只有仿射变换会将椭圆伸缩成圆的区别。

其次，我们考虑二次曲面的对称中心。设一个二次曲面的方程为 $\mathbf{x}^tA\mathbf{x}+2\mathbf{b}^t\mathbf{x}+c=0$，如果它有对称中心，设对称中心为点 $\dot{p}=\dot{o}+\mathbf{p}$，其中 $\mathbf{p}=(p_1,\ldots,p_n)\in\mathbb{R}^n$，那么由对称中心的定义，$\mathbf{p}$ 应该满足：对任意的 $\mathbf{x}\in\mathbb{R}^n$，都有

$$(\mathbf{x}+\mathbf{p})^tA(\mathbf{x}+\mathbf{p})+2\mathbf{b}^t(\mathbf{x}+\mathbf{p})+c=(\mathbf{p}-\mathbf{x})^tA(\mathbf{p}-\mathbf{x})+2\mathbf{b}^t(\mathbf{p}-\mathbf{x})+c=0.$$

展开后整理上式即得：$\mathbf{p}$ 应满足 $A\mathbf{p}+\mathbf{b}=\mathbf{0}$。因此我们可以根据该线性方程组解的情况来计算二次曲面的对称中心坐标。具体到 $\mathbb{R}^3$ 的情形，我们得到：椭球面、椭圆锥面、单叶和双叶双曲面，以及退化情形中的椭圆柱面、双曲柱面、直线和平面都是有对称中心的，而椭圆抛物面、双曲抛物面和抛物柱面没有对称中心。

利用我们在中学中学过的方法，我们可以将直线方程与二次曲面的方程联立，来求直线与二次曲面的交点。如果联立消元后得到的方程有重根，则重根处是直线与二次曲面的切点。研究曲线与曲面的切线 (切面，切方向) 是微分几何的内容之一，读者可以参考相关方向的标准教材。

再来考虑双曲面。容易验证双曲面 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=\pm 1$ 在无穷远处会无限趋近于二次锥面 $\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_n^2}{a_n^2}=0$，我们把这个锥面称为双曲面的渐近锥面。可以借助二维和三维的情形直观地理解这一性质。

最后，我们注意到 $\mathbb{R}^3$ 中的一些二次曲面上含有直线，我们把这样的二次曲面称为直纹面 (ruled surface)。除去平凡的情形，容易柱面和椭圆锥面都是直纹面，下面我们验证单叶双曲面和双曲抛物面也是直纹面。对于单叶双曲面，显然我们只需验证 $x^2+y^2-z^2=1$ 中含有直线即可，非标准的单叶双曲面可以通过仿射变换化成标准情形。注意到原方程可以改写成

$$(x+z)(x-z)=(1+y)(1-y)\quad\text{或}\quad(y+z)(y-z)=(1+x)(1-x),$$

因此以下两组直线

$$\begin{cases} a(x+z) = b(1+y) \\ b(x-z) = a(1-y) \end{cases}, \ \forall a, b \in \mathbb{R} \quad \text{和} \quad \begin{cases} c(y+z) = d(1+x) \\ d(y-z) = c(1-x) \end{cases}, \ \forall c, d \in \mathbb{R}$$

都满足曲面方程，即这两组直线都包含于单叶双曲面。我们把这两组直线称为这个单叶双曲面的**直母线** (generatrix)。可以验证，过单叶双曲面上的每个点都可以作两条直母线，它们分别在上面的两组中。

类似地，对于双曲抛物面 $x^2 - y^2 = 2z$，我们可以改写为：

$$(x+y)(x-y) = 2 \cdot z,$$

因此以下两组直母线

$$\begin{cases} a(x+y) = 2b \\ b(x-z) = az \end{cases}, \ \forall a, b \in \mathbb{R} \quad \text{和} \quad \begin{cases} c(x+y) = dz \\ d(x-y) = 2c \end{cases}, \ \forall c, d \in \mathbb{R}$$

包含于曲面。

关于二次曲面的更详细的讨论，可以参考标准的解析几何教材。

5.3 射影空间

我们知道，非退化的二次曲线也称为圆锥曲线，这是因为它们都可以由平面截圆锥面得到。这里就体现了射影空间的思想：我们可以把曲线或曲面放在高一维的仿射空间中，然后利用中心投影考察其性质。下面我们给出射影空间的一般定义。

定义 5.3.1. 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，将 V 的每个一维子空间 L 视作一个点，这样得到的集合 $\mathbb{P}(V) = \{L \subset V \mid \dim(L) = 1\}$ 称为域 $\mathbb{K}$ 上的 n 维射影空间 (projective space)，也记作 $\mathbb{KP}^n$，n 称为该射影空间的维数，仍记作 $\dim(\mathbb{P}(V)) = n$。

接下来，与仿射空间类似，我们同样可以定义射影子空间 (射影平面)、射影直线和射影超平面。如果 U 是向量空间 V 的 $m+1$ $(0 \leqslant m \leqslant n)$ 维子空间，则所有 U 中的一维子空间构成的集合 $\mathbb{P}(U)$ 是射影空间 $\mathbb{P}(V)$ 的一个子集，我们称 $\mathbb{P}(U)$ 是 $\mathbb{P}(V)$ 中的 m 维射影子空间 (或者射影平面)。特别地，1 维和 $n-1$ 维的射影平面分别被称为射影直线和射影超平面。

我们也可以利用非零向量之间的等价关系来定义射影空间：在 $V \setminus \{\mathbf{0}\}$ 上定义等价关系

$$\sim: \ \mathbf{x} \sim \mathbf{y} \iff \exists \lambda \in \mathbb{K}\text{且}\lambda \neq 0, \text{ 使得}\mathbf{x} = \lambda \mathbf{y}.$$

容易验证这确实是一个等价关系，因此这个等价关系诱导了商集合 $V/\sim$(讲义上册定义 1.4.9)，并且如果我们定义

$$\pi: \ V \setminus \{\mathbf{0}\} \to \mathbb{P}(V), \ \mathbf{x} \to \widetilde{x} = \operatorname{span}\{\mathbf{x}\},$$

则容易看出 π 诱导的等价关系就是我们上面定义的 " $\sim$ "(我们称 π 是 $V \setminus \{\mathbf{0}\}$ 到 $\mathbb{P}(V)$ 的**标准映射**)，因此由讲义上册的定理 1.4.1，我们知道 $V/\sim$ 与 $\mathbb{P}(V)$ 之间是一一对应的。

需要强调的是，射影空间中的点之间没有加法和数乘运算，因为 $\operatorname{span}\{\mathbf{x}\} + \operatorname{span}\{\mathbf{y}\}$ 在 V 中可能是一个二维空间，也就是说它不再是射影空间中的一个点；而 $\operatorname{span}\{\lambda\mathbf{x}\} = \operatorname{span}\{\mathbf{x}\}$ $(\lambda \neq 0)$，也就是说数乘得到的点还是本身，没有意义。

下面我们将射影空间坐标化。

定义 5.3.2 (齐次坐标). 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，$\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基。则对 $\mathbb{P}(V)$ 中的任意一个点 $\widetilde{x}$，我们可以找到非零向量 $\mathbf{x}$ 使得 $\widetilde{x} = \operatorname{span}\{\mathbf{x}\}$。如果 $\mathbf{x} = \sum_{i=0}^{n} a_i \mathbf{e}_i$，$a_i \in \mathbb{K}$，$i \in \{0, 1, \ldots, n\}$，那么我们把坐标 $(a_0 : a_1 : \ldots : a_n)$ 称为 $\widetilde{x}$ 的齐次坐标 (homogeneous coordinates, or projective coordinates)。

由于射影空间中两个点 $\operatorname{span}\{\mathbf{x}\} = \operatorname{span}\{\mathbf{x}'\} \iff \exists \lambda \in \mathbb{K} \setminus \{0\}$ 使得 $\mathbf{x} = \lambda \mathbf{x}'$，因此齐次坐标在成比例的意义下是存在且唯一的：两个齐次坐标 $(a_0 : a_1 : \ldots : a_n)$ 和 $(a_0' : a_1' : \ldots : a_n')$ 相等当且仅当 $\exists \lambda \in \mathbb{K} \setminus \{0\}$ 使得 $a_i = \lambda a_i'$，$\forall i \in \{0, 1, \ldots, n\}$。

射影空间 $\mathbb{P}(V)$ 中的 m 维子空间对应于 V 中的 $m+1$ 维子空间，而由定理1.5.6可知，这个子空间是一个齐次线性方程组的解空间，因此我们可以用 V 上的 (抽象) 齐次线性方程组来表示一个射影平面。

射影空间可以用仿射空间直观地表达，这就是下面的仿射图的定义：

定义 5.3.3. 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，$\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基。考虑 V 的子空间 $V_0 = \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$，将 V 视作与自身相伴的仿射空间，则

$$\mathbb{A}_0 = \mathbf{e}_0 + V_0 = \{\mathbf{e}_0 + \mathbf{x} \mid \mathbf{x} \in V_0\}$$

是 V 中的超平面。于是，任取某个 $\widetilde{x}=\operatorname{span}\{\mathbf{x}\}\in\mathbb{P}(V)$，有且只有以下两种情形：

(a) $\widetilde{x}\in\mathbb{P}(V_0)$，即 $\operatorname{span}\{\mathbf{x}\}\subset V_0$，则此时 $\operatorname{span}\{\mathbf{x}\}$ 作为仿射空间中的直线与 $\mathbb{A}_0$ 平行且没有交点（定义4.1.12），此时我们称 $\widetilde{x}$ 是相对于 $\mathbb{A}_0$ 的**无穷远点**；

(b) $\widetilde{x}\in\mathbb{P}(V)\setminus\mathbb{P}(V_0)$，即 $\operatorname{span}\{\mathbf{x}\}\not\subseteq V_0$，也即 $\mathbf{x}\notin V_0$，则直线 $\operatorname{span}\{\mathbf{x}\}$ 与超平面 $\mathbb{A}_0$ 有且只有一个交点。事实上，不妨设 $\mathbf{x}=a_0\mathbf{e}_0+a_1\mathbf{e}_1+\cdots+a_n\mathbf{e}_n$，由 $\mathbf{x}\notin V_0$ 可知 $a_0\neq 0$，于是容易看出 $\mathbf{y}=\mathbf{e}_0+a_0^{-1}a_1\mathbf{e}_1+\cdots+a_0^{-1}a_n\mathbf{e}_n\in\mathbb{A}_0\cap\operatorname{span}\{\mathbf{x}\}$。另一方面，$\mathbb{A}_0$ 中的任意一点（即一个非零向量）都能张成一个 V 中的一维子空间，这个一维子空间显然不在 V_0 中。因此，我们可以构建双射

$$\Phi_0:\ \mathbb{P}(V)\setminus\mathbb{P}(V_0)\to\mathbb{A}_0,\ \widetilde{x}\mapsto\operatorname{span}\{\mathbf{x}\}\cap\mathbb{A}_0.$$

我们称 $(\mathbb{A}_0,\Phi_0)$ 是 $\mathbb{P}(V)$ 的一个仿射图，$\mathbb{P}(V_0)$ 是相对于该仿射图的无穷远射影超平面。

由于基可以任意选取，因此 V 中任何一个不经过原点的超平面都可以做成 $\mathbb{P}(V)$ 的一个仿射图。此外，在上面的定义中，我们也可以用其余的 $\mathbf{e}_i$ 来代替 $\mathbf{e}_0$，这样我们可以得到 $n+1$ 个仿射图 $(\mathbb{A}_i,\Phi_i)$，$i\in\{0,1,\ldots,n\}$。容易验证，对 $\mathbb{P}(V)$ 中任何一个点 $\widetilde{x}$，至少存在一个仿射图 $(\mathbb{A}_i,\Phi_i)$ 使得 $\widetilde{x}$ 不是无穷远点。因此，我们可以用这 $n+1$ 个仿射图来完全描述这个射影空间。

下面我们考虑射影平面的相交。我们知道，向量空间中子空间的交还是子空间，并且有维数公式成立。因此，两个射影平面的交还是射影平面，并且我们有：

定理 5.3.1. 设 $\mathbb{P}(V)$ 是域 $\mathbb{K}$ 上的 n 维射影空间，Π_1 和 Π_2 是 $\mathbb{P}(V)$ 中的射影平面，如果 $\dim(\Pi_1)+\dim(\Pi_2)\geqslant n$，则 $\Pi_1\cap\Pi_2\neq\emptyset$，且

$$\dim(\Pi_1\cap\Pi_2)\geqslant\dim(\Pi_1)+\dim(\Pi_2)-n.$$

证明. 不妨设 $\Pi_1=\mathbb{P}(U_1)$，$\Pi_2=\mathbb{P}(U_2)$，则 $\dim(U_1)=\dim(\Pi_1)+1$，$\dim(U_2)=\dim(\Pi_2)+1$，$\dim(V)=n+1$，于是

$$\begin{aligned}\dim(U_1)+dim(U_2)&=\dim(\Pi_1)+\dim(\Pi_2)+2\\&\geqslant n+2=\dim(V)+1,\end{aligned}$$

因此由维数公式，我们有

$$\dim(U_1\cap U_2)=\dim(U_1)+dim(U_2)-\dim(U_1+U_2)\geqslant\dim(U_1)+dim(U_2)-\dim(V)\geqslant 1.\quad(5.3.1)$$

这说明 $\mathbb{P}(U_1\cap U_2)\neq\emptyset$，并且将式 (5.3.1) 两边替换成射影空间即得

$$\begin{aligned}\dim(\Pi_1\cap\Pi_2)&=\dim(U_1\cap U_2)-1\\&\geqslant(\dim(\Pi_1)+1)+(\dim(\Pi_2)+1)-(\dim(\mathbb{P}(V))+1)-1\\&=\dim(\Pi_1)+\dim(\Pi_2)-n.\end{aligned}$$

这样我们就完成了证明。 □

特别地，如果 $\dim(\mathbb{P}(V))=2$，那么由上面的定理可知，$\mathbb{P}(V)$ 中的两条射影直线一定相交。这条性质通常作为公理化的射影几何的公理之一。

接下来我们介绍射影变换和射影点的一般位置 (类比于线性变换和线性无关性，或仿射变换和仿射无关性)。

定义 5.3.4. 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，$\mathscr{A}$ 是 V 上的可逆线性算子，则容易看出用 $\mathscr{A}$ 可以定义一个 $\mathbb{P}(V)$ 到自身的双射如下：

$$\begin{aligned} \widetilde{\mathscr{A}}\colon\ \mathbb{P}(V) &\longrightarrow \mathbb{P}(V) \\ \operatorname{span}\{\mathbf{x}\}\ (\mathbf{x}\neq\mathbf{0}) &\longmapsto \operatorname{span}\{\mathscr{A}\mathbf{x}\} \end{aligned}$$

首先 $\widetilde{\mathscr{A}}$ 是良定义的，这可以从 $\mathscr{A}(\lambda\mathbf{x})=\lambda\mathscr{A}\mathbf{x},\ \forall\lambda\in\mathbb{K}$ 中看出；其次 $\widetilde{\mathscr{A}}$ 是双射可以从 $\mathscr{A}$ 可逆得到。我们把 $\widetilde{\mathscr{A}}$ 称为 $\mathbb{P}(V)$ 上的一个射影变换 (projective transformation)。

定理 5.3.2. 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，$\mathscr{A},\mathscr{B}$ 是 V 上的可逆线性算子。则 $\widetilde{\mathscr{A}}=\widetilde{\mathscr{B}} \iff \exists\lambda\in\mathbb{K}\setminus\{0\}$ 使得 $\mathscr{B}=\lambda\mathscr{A}$。

证明. ($\Longleftarrow$) 方向是显然的，下面我们来证明 ($\Longrightarrow$) 方向。

如果 $\widetilde{\mathscr{A}}=\widetilde{\mathscr{B}}$，即 $\forall\mathbf{x}\in V\setminus\{\mathbf{0}\}$，都有 $\operatorname{span}\{\mathscr{A}\mathbf{x}\}=\operatorname{span}\{\mathscr{B}\mathbf{x}\}$，那么对于这个已经取定的 $\mathbf{x}$ 来说，我们有：$\exists\lambda_{\mathbf{x}}\in\mathbb{K}\setminus\{0\}$ 使得 $\mathscr{B}\mathbf{x}=\lambda_{\mathbf{x}}\mathscr{A}\mathbf{x}$。下面我们需要证明 $\lambda_{\mathbf{x}}$ 的值与 $\mathbf{x}$ 的选取无关。

另取 $\mathbf{y}\in V\setminus\{\mathbf{0}\}$，则按照上面的讨论，同样存在 $\lambda_{\mathbf{y}}\in\mathbb{K}\setminus\{0\}$ 使得 $\mathscr{B}\mathbf{y}=\lambda_{\mathbf{y}}\mathscr{A}\mathbf{y}$。我们只需证明 $\lambda_{\mathbf{x}}=\lambda_{\mathbf{y}}$ 恒成立即可。$\mathbf{y}$ 的选取只有以下两种情形：

(1) $\mathbf{y}$ 与 $\mathbf{x}$ 线性相关，由于它们都非零，故存在 $\mu\in\mathbb{K}\setminus\{0\}$ 使得 $\mathbf{y}=\mu\mathbf{x}$，因此

$$\mathscr{B}\mathbf{y}=\mathscr{B}(\mu\mathbf{x})=\mu\mathscr{B}\mathbf{x}=\mu\lambda_{\mathbf{x}}\mathscr{A}\mathbf{x}=\lambda_{\mathbf{x}}\mathscr{A}\mathbf{y}.$$

即此时 $\lambda_{\mathbf{y}}=\lambda_{\mathbf{x}}$ 成立。

(2) $\mathbf{y}$ 与 $\mathbf{x}$ 线性无关，考虑 $\mathscr{B}(\mathbf{x}+\mathbf{y})$，一方面我们有：存在 $\lambda_{\mathbf{x}+\mathbf{y}}\in\mathbb{K}\setminus\{0\}$ 使得

$$\mathscr{B}(\mathbf{x}+\mathbf{y})=\lambda_{\mathbf{x}+\mathbf{y}}\mathscr{A}(\mathbf{x}+\mathbf{y})=\lambda_{\mathbf{x}+\mathbf{y}}\mathscr{A}\mathbf{x}+\lambda_{\mathbf{x}+\mathbf{y}}\mathscr{A}\mathbf{y};$$

另一方面，如果我们先将 $\mathscr{B}(\mathbf{x}+\mathbf{y})$ 展开，则：

$$\mathscr{B}(\mathbf{x}+\mathbf{y})=\mathscr{B}\mathbf{x}+\mathscr{B}\mathbf{y}=\lambda_{\mathbf{x}}\mathscr{A}\mathbf{x}+\lambda_{\mathbf{y}}\mathscr{A}\mathbf{y}.$$

因此我们有

$$\lambda_{\mathbf{x}+\mathbf{y}}\mathscr{A}\mathbf{x}+\lambda_{\mathbf{x}+\mathbf{y}}\mathscr{A}\mathbf{y}=\lambda_{\mathbf{x}}\mathscr{A}\mathbf{x}+\lambda_{\mathbf{y}}\mathscr{A}\mathbf{y},$$

即

$$(\lambda_{\mathbf{x}+\mathbf{y}}-\lambda_{\mathbf{x}})\mathscr{A}\mathbf{x}+(\lambda_{\mathbf{x}+\mathbf{y}}-\lambda_{\mathbf{y}})\mathscr{A}\mathbf{y}=\mathbf{0}.$$

由于 $\mathscr{A}$ 可逆，$\ker(\mathscr{A})=\{\mathbf{0}\}$，故上式表明

$$(\lambda_{\mathbf{x}+\mathbf{y}}-\lambda_{\mathbf{x}})\mathbf{x}+(\lambda_{\mathbf{x}+\mathbf{y}}-\lambda_{\mathbf{y}})\mathbf{y}=\mathbf{0}.$$

而 $\mathbf{x}$ 和 $\mathbf{y}$ 线性无关，因此只能是

$$\lambda_{\mathbf{x}+\mathbf{y}}=\lambda_{\mathbf{x}}=\lambda_{\mathbf{y}}.$$

综上所述，$\mathcal{B}\mathbf{x} = \lambda_{\mathbf{x}}\mathcal{A}\mathbf{x}$ 中 $\lambda_{\mathbf{x}}$ 的值与 $\mathbf{x}$ 的选取无关，也即：存在 $\lambda \in \mathbb{K} \setminus \{0\}$，使得对任意的 $\mathbf{x} \in V \setminus \{\mathbf{0}\}$ 都有 $\mathcal{B}\mathbf{x} = \lambda\mathcal{A}\mathbf{x}$，此即 $\mathcal{B} = \lambda\mathcal{A}$。这样我们就完成了证明。 □

定义 5.3.5 (射影线性群). 设 $\mathbb{P}(V)$ 是域 $\mathbb{K}$ 上的 n 维射影空间，则容易验证 $\mathbb{P}(V)$ 上的所有射影变换在映射复合定义的乘法下构成一个群 (留作练习), 我们称之为射影线性群 (projective gdneral linear group)，记作 $\mathrm{PGL}(V)$。

定理5.3.2启发我们考虑 $\mathrm{PGL}(V)$ 与一般线性群 $\mathrm{GL}(V)$ 之间的关系。

定理 5.3.3. 映射 $\sigma : \mathrm{GL}(V) \to \mathrm{PGL}(V)$, $\mathcal{A} \mapsto \widetilde{\mathcal{A}}$ 是满的群同态, 并且 $\ker(\sigma) = \{\lambda\mathcal{E} \mid \lambda \in \mathbb{K}\backslash\{0\}\}$ [1]。

证明. 由射影变换的定义方式立刻可知 σ 是满射。任取 $\mathcal{A}, \mathcal{B} \in \mathrm{GL}(V)$ 以及 $\widetilde{x} = \mathrm{span}\,\{\mathbf{x}\} \in \mathbb{P}(V)$，我们有：

$$\sigma(\mathcal{A}\mathbb{B})\widetilde{x} = \widetilde{\mathcal{A}\mathcal{B}}(\mathrm{span}\,\{\mathbf{x}\}) = \mathrm{span}\,\{\mathcal{A}\mathcal{B}\mathbf{x}\},$$

和

$$\sigma(\mathcal{A})\sigma(\mathcal{B})\widetilde{x} = \widetilde{\mathcal{A}}\widetilde{\mathcal{B}}(\mathrm{span}\,\{\mathbf{x}\}) = \widetilde{\mathcal{A}}(\mathrm{span}\,\{\mathcal{B}\mathbf{x}\}) = \mathrm{span}\,\{\mathcal{A}\mathcal{B}\mathbf{x}\}.$$

因此 $\sigma(\mathcal{A}\mathcal{B}) = \sigma(\mathcal{A})\sigma(\mathcal{B})$，即 σ 是群同态。由定理5.3.2立刻得到

$$\ker(\sigma) = \{\widetilde{\mathcal{A}} \in \mathrm{PGL}(V) \mid \widetilde{\mathcal{A}} = \widetilde{\mathcal{E}}\} = \{\lambda\mathcal{E} \mid \lambda \in \mathbb{K} \setminus \{0\}\}.$$

这样我们就完成了证明。 □

射影线性群可以确定射影几何。射影空间中图形在射影变换下保持不变的那些性质被称为射影性质。射影点处于**一般位置**就是一种射影性质。

定义 5.3.6. 设 $\mathbb{P}(V)$ 是域 $\mathbb{K}$ 上的 n 维射影空间，$n+2$ 个射影点 $\widetilde{x_i} = \mathrm{span}\,\{\mathbf{x}_i\} \in \mathbb{P}(V)$，$i \in \{0, 1, \ldots, n+1\}$。如果其中任何 $n+1$ 个向量 $\mathbf{x}_0, \ldots, \mathbf{x}_{j-1}, \mathbf{x}_{j+1}, \ldots, \mathbf{x}_{n+1}$ 都线性无关，则我们称这 $n+2$ 个射影点处于一般位置。

定理 5.3.4. 设 $\mathbb{P}(V)$ 是域 $\mathbb{K}$ 上的 n 维射影空间，$n+2$ 个射影点 $\widetilde{x_i} = \mathrm{span}\,\{\mathbf{x}_i\} \in \mathbb{P}(V)$，$i \in \{0, 1, \ldots, n+1\}$ 处于一般位置。则有：

(1) 对 $\forall \widetilde{\mathcal{A}} \in \mathrm{PGL}(V)$，$\widetilde{\mathcal{A}}\widetilde{x_i}$，$i \in \{0, 1, \ldots, n+1\}$ 也处于一般位置；

(2) 如果 $\widetilde{y_i} = \mathrm{span}\,\{\mathbf{y}_i\} \in \mathbb{P}(V)$，$i \in \{0, 1, \ldots, n+1\}$ 也处于一般位置，那么存在唯一的 $\widetilde{\mathcal{A}} \in \mathrm{PGL}(V)$ 使得

$$\widetilde{\mathcal{A}}\widetilde{x_i} = \widetilde{y_i},\ \forall i \in \{0, 1, \ldots, n+1\}.$$

证明. (1) 注意到 $\mathcal{A}$ 可逆，它将线性无关的向量仍然映射到线性无关的向量，按定义即得结论。

(2) 首先, 由于 $n+1$ 个向量 $\mathbf{x}_0, \ldots, \mathbf{x}_n$ 线性无关而 $\dim(V) = n+1$, 因此 $\mathbf{x}_{n+1}$ 可以由 $\mathbf{x}_0, \ldots, \mathbf{x}_n$ 线性表出，即存在不全为 0 的 $a_0, \ldots a_n \in \mathbb{K}$ 使得

$$\mathbf{x}_{n+1} = a_0\mathbf{x}_0 + \cdots + a_n\mathbf{x}_n.$$

---

[1] 实际上这表明 $\mathrm{PGL}(V)$ 同构于商群 $\mathrm{GL}(V)/\mathrm{Z}(V)$，其中 $\mathrm{Z}(V)$ 表示 V 上所有非零数乘变换构成的子群，$\mathrm{Z}(V)$ 是一般线性群的中心。

如果某个 $a_j = 0$，那么 $\mathbf{x}_0, \dots, \mathbf{x}_{j-1}, \mathbf{x}_{j+1}, \dots, \mathbf{x}_{n+1}$ 线性相关，这与一般位置的定义矛盾！因此 $a_0, \dots, a_n$ 全都不为 0。

同理，存在另一组全不为 0 的 $b_0, \dots b_n \in \mathbb{K}$ 使得

$$\mathbf{y}_{n+1} = b_0\mathbf{y}_0 + \cdots + b_n\mathbf{y}_n.$$

注意到 $a_0\mathbf{x}_0, \dots, a_n\mathbf{x}_n$ 和 $b_0\mathbf{y}_0 \dots, b_n\mathbf{y}_n$ 也都是 V 的基，因此我们可以定义 $\mathcal{A} \in \mathrm{GL}(V)$ 如下：

$$\mathcal{A} : V \to V,\ a_i\mathbf{x}_i \mapsto b_i\mathbf{y}_i,\ \forall i \in \{0, \dots, n\}.$$

则 $\mathcal{A}$ 显然也满足 $\mathcal{A}\mathbf{x}_{n+1} = \mathbf{y}_{n+1}$，两边取定理5.3.3中的 σ 即得 $\widetilde{\mathcal{A}}\widetilde{x}_i = \widetilde{y}_i$ 对 $\forall i \in \{0, 1, \dots, n+1\}$ 都成立。

下证 $\widetilde{\mathcal{A}}$ 是唯一的。如果另有 $\widetilde{\mathcal{B}}$ 也满足定理要求，即 $\widetilde{\mathcal{A}}\widetilde{x}_i = \widetilde{\mathcal{B}}\widetilde{x}_i = \widetilde{y}_i$, $\forall i \in \{0, 1, \dots, n+1\}$。那么我们有

$$(\widetilde{\mathcal{B}})^{-1}\widetilde{\mathcal{A}}\widetilde{x}_i = \widetilde{x}_i,\ \forall i \in \{0, 1, \dots, n+1\}.$$

由于 $\sigma : \mathcal{B} \mapsto \widetilde{\mathcal{B}}$ 是满同态，所以 $(\widetilde{\mathcal{B}})^{-1}\widetilde{\mathcal{A}} = (\widetilde{\mathcal{B}^{-1}})\widetilde{\mathcal{A}} = \widetilde{\mathcal{B}^{-1}\mathcal{A}}$，即

$$\widetilde{\mathcal{B}^{-1}\mathcal{A}}\widetilde{x}_i = \widetilde{x}_i,\ \forall i \in \{0, 1, \dots, n+1\},$$

上式对应到向量空间即得存在 $\lambda_0, \dots, \lambda_{n+1} \in \mathbb{K} \setminus \{0\}$ 使得

$$\mathcal{B}^{-1}\mathcal{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i,\ \forall i \in \{0, 1, \dots, n+1\}. \tag{5.3.2}$$

注意到 $\mathbf{x}_{n+1} = a_0\mathbf{x}_0 + \cdots + a_n\mathbf{x}_n$，带入到式 (5.3.2) 即得

$$\mathcal{B}^{-1}\mathcal{A}(a_0\mathbf{x}_0 + \cdots + a_n\mathbf{x}_n) = \lambda_{n+1}(a_0\mathbf{x}_0 + \cdots + a_n\mathbf{x}_n),$$

即

$$a_0\lambda_0\mathbf{x}_0 + \cdots + a_n\lambda_n\mathbf{x}_n = \lambda_{n+1}(a_0\mathbf{x}_0 + \cdots + a_n\mathbf{x}_n),$$

由于 $\mathbf{x}_0, \dots, \mathbf{x}_n$ 线性无关，因此只能是 $a_i\lambda_i = a_i\lambda_{n+1}$, $\forall i \in \{0, \dots, n\}$。由于 a_i 全部不为 0，因此只能是 $\lambda_0 = \cdots = \lambda_n = \lambda_{n+1}$，此即说明 $\mathcal{B}^{-1}\mathcal{A} = \lambda_{n+1}\mathcal{E}$，所以 $\widetilde{\mathcal{B}} = \widetilde{\mathcal{A}}$，唯一性证毕。这样我们就完成了整个定理的证明。

□

推论 5.3.1. (1) 对于一维射影空间 $\mathbb{KP}^1$ 而言，设 $\widetilde{x}_i$, $i = 0, 1, 2$ 和 $\widetilde{y}_i$, $i = 0, 1, 2$ 是两个三点组，其内部的三个射影点均两两不同，则存在唯一的 $\widetilde{\mathcal{A}} \in \mathrm{PGL}(\mathbb{KP}^1)$ 使得 $\widetilde{\mathcal{A}}\widetilde{x}_i = \widetilde{y}_i$, $i = 0, 1, 2$；

(2) 设 $\mathbb{P}(U), \mathbb{P}(W)$ 是 $\mathbb{P}(V)$ 中的两个 m 维射影平面，则存在某个射影变换将 $\mathbb{P}(U)$ 变成 $\mathbb{P}(W)$；

(3) 设 $\mathbb{P}(U)$ 是 $\mathbb{P}(V)$ 中的射影平面，则 $\mathbb{P}(U)$ 上的射影变换可以扩张到 $\mathbb{P}(V)$ 上。

(1) 是显然的，这也说明射影变换不再保持比例；利用定理5.3.4和基扩充定理 (定理1.3.2(2)) 即可证明推论 (2) 和 (3)，细节留作练习。

下面我们介绍另一种射影性质：**交比** (cross ratio)。虽然射影变换不保持线段的比例，但它却具有下面的不变性：设在射影空间 V 中，四个射影点 $\widehat{u}_i = \operatorname{span}\{\mathbf{u}_i\}$，$i \in \{1,2,3,4\}$ 在同一条射影直线 $\mathbb{KP}^1 = \mathbb{P}(U)$ (即 $\dim(U) = 2$) 上，并且 $\widehat{u}_1 \neq \widehat{u}_3$，$\widehat{u}_1 \neq \widehat{u}_4$，$\widehat{u}_2 \neq \widehat{u}_3$，$\widehat{u}_2 \neq \widehat{u}_4$。也就是说，$(\mathbf{u}_1,\mathbf{u}_3),(\mathbf{u}_1,\mathbf{u}_4)$, $(\mathbf{u}_2,\mathbf{u}_3),(\mathbf{u}_2,\mathbf{u}_4)$ 都是 U 的基底。我们把 $(\mathbf{u}_1,\mathbf{u}_4)$ 到 $(\mathbf{u}_1,\mathbf{u}_3)$ 的过渡矩阵 (定义1.3.6) 的行列式记为 $\left(\dfrac{\mathbf{u}_1,\mathbf{u}_3}{\mathbf{u}_1,\mathbf{u}_4}\right)$，同理 $(\mathbf{u}_2,\mathbf{u}_4)$ 到 $(\mathbf{u}_2,\mathbf{u}_3)$ 的过渡矩阵的行列式记为 $\left(\dfrac{\mathbf{u}_2,\mathbf{u}_3}{\mathbf{u}_2,\mathbf{u}_4}\right)$，则我们定义如下的数值

$$[\widehat{u}_1,\widehat{u}_2,\widehat{u}_3,\widehat{u}_4] = \left(\frac{\mathbf{u}_1,\mathbf{u}_3}{\mathbf{u}_1,\mathbf{u}_4}\right)\cdot\left(\frac{\mathbf{u}_2,\mathbf{u}_3}{\mathbf{u}_2,\mathbf{u}_4}\right)^{-1} \qquad \text{(注意这是域中的一个数除以另一个数！)}$$

为这四个射影点 (带顺序) $\widehat{u}_1,\widehat{u}_2,\widehat{u}_3,\widehat{u}_4$ 的交比。

下面我们需要说明两件事：一是这样定义的交比是良定义的，即交比的值和 $\mathbf{u}_i$ 的选取无关；二是交比确实是一个射影性质，即它在射影变换下保持不变。

(1) 交比是良定义的：我们不妨设

$$\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}\cdot\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_4\end{pmatrix}, \qquad \begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_2 & \beta_2\\ \gamma_2 & \delta_2\end{pmatrix}\cdot\begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_4\end{pmatrix}. \tag{5.3.3}$$

如果我们以 $\lambda_i\mathbf{u}_i$ $(\lambda_i \neq 0)$ 代替 $\mathbf{u}_i$，则

$$\begin{pmatrix}\lambda_1\mathbf{u}_1\\ \lambda_3\mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_1 & \lambda_1\lambda_4^{-1}\beta_1\\ \lambda_3\lambda_1^{-1}\gamma_1 & \lambda_3\lambda_4^{-1}\delta_1\end{pmatrix}\cdot\begin{pmatrix}\lambda_1\mathbf{u}_1\\ \lambda_4\mathbf{u}_4\end{pmatrix}, \qquad \begin{pmatrix}\lambda_2\mathbf{u}_2\\ \lambda_3\mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_2 & \lambda_2\lambda_4^{-1}\beta_2\\ \lambda_3\lambda_2^{-1}\gamma_2 & \lambda_3\lambda_4^{-1}\delta_2\end{pmatrix}\cdot\begin{pmatrix}\lambda_2\mathbf{u}_2\\ \lambda_4\mathbf{u}_4\end{pmatrix}.$$

此时按定义计算出交比的值应为

$$\begin{aligned}
&\det\begin{pmatrix}\alpha_1 & \lambda_1\lambda_4^{-1}\beta_1\\ \lambda_3\lambda_1^{-1}\gamma_1 & \lambda_3\lambda_4^{-1}\delta_1\end{pmatrix}\cdot\left(\det\begin{pmatrix}\alpha_2 & \lambda_2\lambda_4^{-1}\beta_2\\ \lambda_3\lambda_2^{-1}\gamma_2 & \lambda_3\lambda_4^{-1}\delta_2\end{pmatrix}\right)^{-1}\\
&= \lambda_3\lambda_4^{-1}\det\begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}\cdot\lambda_4\lambda_3^{-1}\left(\det\begin{pmatrix}\alpha_2 & \beta_2\\ \gamma_2 & \delta_2\end{pmatrix}\right)^{-1}\\
&= \det\begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}\cdot\left(\det\begin{pmatrix}\alpha_2 & \beta_2\\ \gamma_2 & \delta_2\end{pmatrix}\right)^{-1}.
\end{aligned}$$

此即说明交比的值与 $\mathbf{u}_i$ 的选取无关。

(2) 交比是射影性质：设 $\widetilde{\mathcal{A}} \in \mathrm{PGL}(V)$，并且两个过渡矩阵仍如式 (5.3.3) 所设，则显然有

$$\begin{pmatrix}\mathcal{A}\mathbf{u}_1\\ \mathcal{A}\mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}\cdot\begin{pmatrix}\mathcal{A}\mathbf{u}_1\\ \mathcal{A}\mathbf{u}_4\end{pmatrix}, \qquad \begin{pmatrix}\mathcal{A}\mathbf{u}_2\\ \mathcal{A}\mathbf{u}_3\end{pmatrix} = \begin{pmatrix}\alpha_2 & \beta_2\\ \gamma_2 & \delta_2\end{pmatrix}\cdot\begin{pmatrix}\mathcal{A}\mathbf{u}_2\\ \mathcal{A}\mathbf{u}_4\end{pmatrix}.$$

即

$$\left(\frac{\mathcal{A}\mathbf{u}_1,\mathcal{A}\mathbf{u}_3}{\mathcal{A}\mathbf{u}_1,\mathcal{A}\mathbf{u}_4}\right)\cdot\left(\frac{\mathcal{A}\mathbf{u}_2,\mathcal{A}\mathbf{u}_3}{\mathcal{A}\mathbf{u}_2,\mathcal{A}\mathbf{u}_4}\right)^{-1} = \left(\frac{\mathbf{u}_1,\mathbf{u}_3}{\mathbf{u}_1,\mathbf{u}_4}\right)\cdot\left(\frac{\mathbf{u}_2,\mathbf{u}_3}{\mathbf{u}_2,\mathbf{u}_4}\right)^{-1} = \det\begin{pmatrix}\alpha_1 & \beta_1\\ \gamma_1 & \delta_1\end{pmatrix}\cdot\left(\det\begin{pmatrix}\alpha_2 & \beta_2\\ \gamma_2 & \delta_2\end{pmatrix}\right)^{-1}.$$

这说明 $[\widetilde{\mathcal{A}}\widehat{u}_1,\widetilde{\mathcal{A}}\widehat{u}_2,\widetilde{\mathcal{A}}\widehat{u}_3,\widetilde{\mathcal{A}}\widehat{u}_4] = [\widehat{u}_1,\widehat{u}_2,\widehat{u}_3,\widehat{u}_4]$，即交比在射影变换下保持不变。

接下来我们将在一组标准基下计算交比。取 U 的一组基 $\mathbf{e}_1, \mathbf{e_2}$，设 $\widehat{u_i}$ 在这组基下的齐次坐标分别为 $(x_i : y_i)$，$i \in \{1, 2, 3, 4\}$，即

$$\widehat{u_i} = \text{span}\,\{\mathbf{u}_i\},\ \mathbf{u}_i = (x_i, y_i)\begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix},\ i \in \{1, 2, 3, 4\}.$$

那么我们立刻有：

$$\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}x_1 & y_1\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix},\quad \begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_4\end{pmatrix} = \begin{pmatrix}x_1 & y_1\\ x_4 & y_4\end{pmatrix}\cdot\begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix},\quad \begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}x_2 & y_2\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix},\quad \begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_4\end{pmatrix} = \begin{pmatrix}x_2 & y_2\\ x_4 & y_4\end{pmatrix}\cdot\begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix}.$$

于是

$$\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}x_1 & y_1\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}x_1 & y_1\\ x_4 & y_4\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_4\end{pmatrix},\qquad \begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_3\end{pmatrix} = \begin{pmatrix}x_2 & y_2\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}x_2 & y_2\\ x_4 & y_4\end{pmatrix}^{-1}\begin{pmatrix}\mathbf{u}_2\\ \mathbf{u}_4\end{pmatrix}.$$

所以交比

$$\begin{aligned}[\widehat{u}_1, \widehat{u}_2, \widehat{u}_3, \widehat{u}_4] &= \det\left(\begin{pmatrix}x_1 & y_1\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}x_1 & y_1\\ x_4 & y_4\end{pmatrix}^{-1}\right)\cdot\det\left(\begin{pmatrix}x_2 & y_2\\ x_3 & y_3\end{pmatrix}\cdot\begin{pmatrix}x_2 & y_2\\ x_4 & y_4\end{pmatrix}^{-1}\right)^{-1}\\ &= \frac{\det\begin{pmatrix}x_1 & y_1\\ x_3 & y_3\end{pmatrix}\cdot\det\begin{pmatrix}x_2 & y_2\\ x_4 & y_4\end{pmatrix}}{\det\begin{pmatrix}x_1 & y_1\\ x_4 & y_4\end{pmatrix}\cdot\det\begin{pmatrix}x_2 & y_2\\ x_3 & y_3\end{pmatrix}}.\end{aligned}\tag{5.3.4}$$

特别地，如果每个 $\widehat{u_i}$ 的齐次坐标都满足 $x_i \neq 0$，那么 $(x_i : y_i) = (1 : \frac{y_i}{x_i})$，记 $\frac{y_i}{x_i} = a_i$，则公式 (5.3.4) 可以化简为

$$\begin{aligned}[\widehat{u}_1, \widehat{u}_2, \widehat{u}_3, \widehat{u}_4] &= \frac{\det\begin{pmatrix}1 & a_1\\ 1 & a_3\end{pmatrix}\cdot\det\begin{pmatrix}1 & a_2\\ 1 & a_4\end{pmatrix}}{\det\begin{pmatrix}1 & a_1\\ 1 & a_4\end{pmatrix}\cdot\det\begin{pmatrix}1 & a_2\\ 1 & a_3\end{pmatrix}}\\ &= \frac{(a_3 - a_1)(a_4 - a_2)}{(a_4 - a_1)(a_3 - a_2)}.\end{aligned}$$

更进一步地，我们可以验证（留作练习）：分别共射影直线的两个两两不同的四点组 $\widehat{u}_1, \widehat{u}_2, \widehat{u}_3, \widehat{u}_4$ 和 $\widehat{v}_1, \widehat{v}_2, \widehat{v}_3, \widehat{v}_4$ 满足 $[\widehat{u}_1, \widehat{u}_2, \widehat{u}_3, \widehat{u}_4] = [\widehat{v}_1, \widehat{v}_2, \widehat{v}_3, \widehat{v}_4] \iff$ 存在 $\widetilde{\mathscr{A}} \in \mathrm{PGL}(V)$ 使得 $\widetilde{\mathscr{A}}\widehat{u_i} = \widehat{v_i}$，$\forall i \in \{1, 2, 3, 4\}$(这是一个等价关系，验证留作练习)。

在本节的最后，我们将介绍射影空间中的二次曲面，并把其与仿射空间中的二次曲面联系起来。为此，我们首先介绍射影代数簇 (projective algebraic variety) 的概念。

定义 5.3.7. 设 V 是域 $\mathbb{K}$ 上的 $n+1$ 维向量空间，则固定 V 的一组基后 $\mathbb{P}(V)$ 中的每个射影点对应一个齐次坐标 $(x_0 : x_1 : \cdots : x_n)$。注意到对于任意的齐次多元多项式 $f \in \mathbb{K}[t_0, t_1, \ldots, t_n]$，如果 $f(x_0, x_1, \ldots, x_n) = 0$，则 $\forall \lambda \in \mathbb{K} \setminus \{0\}$，都有 $f(\lambda x_0, \lambda x_1, \ldots, \lambda x_n) = 0$。因此我们可以把相应的射影点（或者说齐次坐标）视作齐次多项式的零点。

于是，设 $I \subset \mathbb{K}[t_0, \ldots, t_n]$ 且 I 中的多项式都是齐次多项式，则我们称下面的集合

$$X = \{\widehat{x} = (x_0 : x_1 : \cdots : x_n) \mid \forall f \in I,\ f(x_0, x_1, \ldots, x_n) = 0\} \subset \mathbb{P}(V)\ (\text{或者说}\mathbb{KP}^n)$$

为射影代数簇。容易验证任意多个射影代数簇的交以及有限多个射影代数簇的并仍然是射影代数簇[1]，留作练习。

特别地，如果 I 中只有一个多项式，设这个多项式的全次数是 d，即 $X = \{\widehat{x} \mid f(\widehat{x}) = 0,\ \deg(f) = d\}$，则我们称 X 是一个 d 次的射影超曲面。

在本书中我们只讨论二次的射影曲面，也即二次齐次多项式的零点集。我们知道，二次齐次多项式就是二次型，因此我们可以用第一章的知识来处理。我们只讨论复数域和实数域上的情形。

考虑域 $\mathbb{K}(\mathbb{C}$ 或 $\mathbb{R})$ 上的二次齐次多项式 $f(t_0, \ldots, t_n) = \sum_{i=0}^{n} a_{ii}t_i^2 + 2\sum_{0\leqslant i<j\leqslant n} a_{ij}t_it_j$，记它确定的射影二次曲面为 S，我们知道

$$f(t_0,\ldots,t_n) = (t_0,\ldots,t_n)\underbrace{\begin{pmatrix} a_{00} & \cdots & a_{0n} \\ \vdots & \ddots & \vdots \\ a_{n0} & \cdots & a_{nn} \end{pmatrix}}_{\text{记为}A}\begin{pmatrix} t_0 \\ \vdots \\ t_n \end{pmatrix}$$

由 1.6.6 小节的讨论我们知道，当 $\mathbb{K} = \mathbb{C}$ 时，存在 $P \in \mathrm{GL}_{n+1}(\mathbb{C})$ 使得 $P^tAP = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{(n+1)\times(n+1)}$，即作射影变换 $\widetilde{P}$ 后曲面 S 会变为 $t_0^2 + \cdots + t_{r-1}^2 = 0$ 确定的射影二次曲面。也就是说，复射影二次曲面由其二次型的秩完全决定。

而当 $\mathbb{K} = \mathbb{R}$ 时，我们知道：存在 $P \in \mathrm{GL}_{n+1}(\mathbb{R})$ 使得 $P^tAP = \begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}_{(n+1)\times(n+1)}$，也即作射影变换 $\widetilde{P}$ 后曲面 S 会变为 $t_0^2 + \cdots + t_{s-1}^2 - t_s^2 - \cdots - t_{r-1}^2 = 0$ 确定的射影二次曲面。此外，我们可以要求 $s \geqslant r - s$，否则我们可以在方程两端乘以 -1。特别地，当 $s = r < n+1$ 时，S 退化成射影平面；当 $s = r = n+1$ 时，S 是空集。也就是说，实射影二次曲面由其二次型的正惯性指数和秩完全决定。

下面的定理表明，实射影二次曲面中可能含有射影平面。显然我们只需要考虑非退化的情形，即曲面方程中所有变元都是“有效”的，或者说 $r = n+1$。

定理 5.3.5. 设实射影二次曲面 S 由方程 $t_0^2+\cdots+t_{s-1}^2-t_s^2-\cdots-t_n^2 = 0$ 确定，其中 $\dfrac{n+1}{2} \leqslant s < n$。则 S 中含有的射影平面的最大维数是 $n - s$。

证明. 我们先证明以下结论：方程组 $t_s = t_{s+1} = \cdots = t_n = 0$ 确定的射影平面 Π(定义5.3.2下方的讨论告诉我们这确实是一个射影平面) 与 S 的交集是空集。

这是因为：如果 S 上的射影点的齐次坐标满足 $t_s = t_{s+1} = \cdots = t_n = 0$，那么必有 $t_0^2 + \cdots + t_{s-1}^2 = 0$，而在 $\mathbb{R}$ 上这说明 $t_0 = \cdots = t_{s-1} = 0$，即齐次坐标中的数全为 0，这是不可能的！这样我们的结论就一定成立。

利用对偶定理容易验证 Π 作为射影空间的维数是 $s-1$。由定理5.3.1，我们知道 $\mathbb{RP}^n$ 中的任何一个 $n-s+1$ 维射影平面都一定与 Π 有公共的射影点，而这个公共点一定不在 S 中（否则 S 与 Π 有交点，矛盾！），即任何一个 $n-s+1$ 维射影平面一定含有不在 S 中的射影点。这说明 S 中含有的射影平面的最大维数小于或等于 $n-s$。

---

[1] 这构成了一个 Zariski 拓扑，我们会在交换代数课程中学习。

另一方面，方程组

$$\begin{cases} t_0 = t_s \\ \vdots \\ t_{n-s} = t_n \\ t_{n-s+1} = 0 \\ \vdots \\ t_{s-1} = 0 \end{cases}$$

确定了一个 $\mathbb{RP}^n$ 中的 $n-s$ 维平面 Π'（这个方程组中含有 $(n-s+1)+(s-1-(n-s+1)+1) = s$ 个方程，并且满秩，所以它确定了一个 $\mathbb{R}^{n+1}$ 中的 $n+1-s$ 维线性子空间，即 $n-s$ 维射影平面)。容易看出 $\Pi' \subset S$，即 S 中含有 $n-s$ 维射影平面。这样我们就完成了证明。 □

由上面的定理我们立刻知道，对于非退化情形，只有 $s = n$ 时，即曲面方程经射影变换可化为 $t_0^2 + \cdots + t_{n-1}^2 - -t_n^2 = 0$ 时，实射影二次曲面中不含有平面，这时我们称其为 (射影) 椭球面；其余情形我们称为 (射影) 直纹面。

最后，我们利用仿射图将射影二次曲面与仿射二次曲面联系起来。

- $\mathbb{RP}^2$ 情形：此时非退化的射影二次曲面方程只有 $t_0^2 + t_1^2 - t_2^2 = 0$ 一种情形。

  (1) 用 $\mathbb{A}_0 = \mathbf{e}_0 + \mathrm{span}\,\{\mathbf{e}_1, \mathbf{e}_2\}$ 和 Φ_0 去截射影二次曲面，则 $t_0 = 0$ 的射影点是相对于 $\mathbb{A}_0$ 的无穷远点，而其余点则被送到了 $\mathbb{A}_0$ 上：

  $$\Phi_0\left((t_0 : t_1 : t_2)\right) = (1, \frac{t_1}{t_0}, \frac{t_2}{t_0}).$$

  $\mathbb{A}_0$ 上的像点满足 $1 + (\frac{t_1}{t_0})^2 - (\frac{t_2}{t_0})^2 = 0$，这是仿射空间 $\mathbb{R}^2$ 中的双曲线；

  (2) 用 $\mathbb{A}_1 = \mathbf{e}_1 + \mathrm{span}\,\{\mathbf{e}_0, \mathbf{e}_2\}$ 和 Φ_1 去截射影二次曲面，则 $t_1 = 0$ 的射影点是相对于 $\mathbb{A}_1$ 的无穷远点，而其余点则被送到了 $\mathbb{A}_1$ 上：

  $$\Phi_1\left((t_0 : t_1 : t_2)\right) = (\frac{t_0}{t_1}, 1, \frac{t_2}{t_1}).$$

  $\mathbb{A}_1$ 上的像点满足 $(\frac{t_0}{t_1})^2 + 1 - (\frac{t_2}{t_1})^2 = 0$，这仍是 $\mathbb{R}^2$ 中的双曲线；

  (3) 用 $\mathbb{A}_2 = \mathbf{e}_2 + \mathrm{span}\,\{\mathbf{e}_0, \mathbf{e}_1\}$ 和 Φ_2 去截射影二次曲面，则 $t_2 = 0$ 的射影点是相对于 $\mathbb{A}_2$ 的无穷远点，而其余点则被送到了 $\mathbb{A}_2$ 上：

  $$\Phi_2\left((t_0 : t_1 : t_2)\right) = (\frac{t_0}{t_2}, \frac{t_1}{t_2}, 1).$$

  $\mathbb{A}_2$ 上的像点满足 $(\frac{t_0}{t_2})^2 + (\frac{t_1}{t_2})^2 - 1 = 0$，这是 $\mathbb{R}^2$ 中的圆 (不允许伸缩则变换为椭圆)；

  (4) 作射影变换：$u_0 = t_0$，$u_1 = t_1 + t_2, u_2 = t_1 - t_2$，则射影二次曲面的方程变为 $u_0^2 + u_1 u_2 = 0$。在新的基底下用 $\mathbb{A}_1' = \mathbf{e}_1' + \mathrm{span}\,\{\mathbf{e}_0', \mathbf{e}_2'\}$ 和相对应的 Φ_1' 去截射影二次曲面，则除无穷远点以外

  $$\Phi_1'\left((u_0 : u_1 : u_2)\right) = (\frac{u_0}{u_1}, 1, \frac{u_2}{u_1}).$$

$\mathbb{A}_1'$ 上的像点满足 $(\frac{u_0}{u_1})^2 + (\frac{u_2}{u_1}) = 0$，这是 $\mathbb{R}^2$ 中的抛物线。

- $\mathbb{RP}^3$ 情形：此时非退化的射影二次曲面方程有 $t_0^2 + t_1^2 + t_2^2 - t_3^2 = 0$ 和 $t_0^2 + t_1^2 - t_2^2 - t_3^2 = 0$ 两种情形。类似于 $\mathbb{RP}^2$ 上的做法，我们可以得到以下结果，细节留作练习：
  - ⋆ 对 $t_0^2 + t_1^2 + t_2^2 - t_3^2 = 0$：
    - (1) 用 $\mathbb{A}_0$, $\mathbb{A}_1$，或 $\mathbb{A}_2$ 及相对应的 Φ 去截射影二次曲面，在仿射图上得到的是 $\mathbb{R}^3$ 中的双叶双曲面；
    - (2) 用 $\mathbb{A}_3$ 及相对应的 Φ_3 去截射影二次曲面，在仿射图上得到的是 $\mathbb{R}^3$ 中的 (椭) 球面；
    - (3) 作射影变换：$u_0 = t_0$, $u_1 = t_1$, $u_2 = t_2 + t_3$, $u_3 = t_2 - t_3$，则射影二次曲面的方程变为 $u_0^2 + u_1^2 + u_2u_3 = 0$。在新的基底下用 $\mathbb{A}_2'$ 及相对应的 Φ_2' 去截射影二次曲面，在仿射图上得到的得到的是椭圆抛物面；
  - ⋆ 对 $t_0^2 + t_1^2 - t_2^2 - t_3^2 = 0$：
    - (1) 用 $\mathbb{A}_0$, $\mathbb{A}_1$, $\mathbb{A}_2$ 或 $\mathbb{A}_3$ 及相对应的 Φ 去截射影二次曲面，在仿射图上得到的是 $\mathbb{R}^3$ 中的单叶双曲面；
    - (2) 作射影变换：$u_0 = t_0$, $u_1 = t_1 + t_2$, $u_2 = t_1 - t_2$, $u_3 = t_3$，则射影二次曲面的方程变为 $u_0^2 + u_1u_2 - u_3^2 = 0$。在新的基底下用 $\mathbb{A}_1'$ 及相对应的 Φ_1' 去截射影二次曲面，在仿射图上得到的得到的是双曲抛物面；
  - ⋆ 对 $\mathbb{RP}^3$ 中的退化情形，也就是射影二次曲面的方程可化为 $t_0^2+t_1^2-t_2^2 = 0$ 或 $t_0^2-t_1^2 = 0$ 的情形，我们有：
    - (1) 用 $\mathbb{A}_0$ 或 $\mathbb{A}_1$ 及相对应的 Φ 去截 $t_0^2 + t_1^2 - t_2^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的双曲柱面；
    - (2) 用 $\mathbb{A}_2$ 及相对应的 Φ_2 去截 $t_0^2 + t_1^2 - t_2^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的 (椭) 圆柱面；
    - (3) 用 $\mathbb{A}_3$ 及相对应的 Φ_3 去截 $t_0^2 + t_1^2 - t_2^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的 (椭) 圆锥面；
    - (4) 作射影变换：$u_0 = t_0$, $u_1 = t_1 + t_2$, $u_2 = t_1 - t_2$，则方程 $t_0^2 + t_1^2 - t_2^2 = 0$ 变为 $u_0^2 + u_1u_2 = 0$。在新的基底下用 $\mathbb{A}_1'$ 及相对应的 Φ_1' 去截 $u_0^2 + u_1u_2 = 0$，在仿射图上得到的得到的是 $\mathbb{R}^3$ 中的抛物柱面；
    - (5) 用 $\mathbb{A}_0$ 或 $\mathbb{A}_1$ 及相对应的 Φ 去截 $t_0^2 - t_1^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的平行平面；
    - (6) 用 $\mathbb{A}_2$ 或 $\mathbb{A}_3$ 及相对应的 Φ 去截 $t_0^2 - t_1^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的相交平面。
  - ⋆ 如果射影二次曲面的方程可化为 $t_0^2 + t_1^2 + t_2^2 = 0$ 或 $t_0^2 + t_1^2 = 0$ 或 $t_0^2 = 0$，则这些射影二次曲面本质上是射影平面，我们有：
    - (1) 用 $\mathbb{A}_3$ 及相对应的 Φ_3 去截 $t_0^2 + t_1^2 + t_2^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的原点；
    - (2) 用 $\mathbb{A}_2$ 或 $\mathbb{A}_3$ 及相对应的 Φ 去截 $t_0^2 + t_1^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的一条直线；
    - (3) 用 $\mathbb{A}_1$, $\mathbb{A}_2$ 或 $\mathbb{A}_3$ 及相对应的 Φ 去截 $t_0^2 = 0$，在仿射图上得到的是 $\mathbb{R}^3$ 中的一个平面 (两个重合平面)。

这样我们就通过 $\mathbb{RP}^3$ 中的射影二次曲面和仿射图得到了 $\mathbb{R}^3$ 中除了空集以外的所有仿射二次曲面。我们也可以将上述过程反过来，将仿射二次曲面的方程齐次化从而得到射影二次曲面，过程留作练习。利用仿射空间与射影空间的联系我们可以对很多问题的处理进行简化。

有关二次曲面和射影空间的更多内容，读者可以参考《解析几何》，尤承业，北京大学出版社或者？(帮忙提供一本英文参考书，相关的英文教材编者了解不多)。

5.4 习题

二次曲面的仿射分类和正交分类

1. 在 3 维欧几里得空间中通过保距变换把二次曲面化成标准形并判定曲面的类型; 并判断其是否有对称中心, 如果有，请求出:

   (1) $2xy + 2xz + 2yz + 2x + 2y + 2z + 1 = 0$;
   (2) $2x^2 + y^2 - 3z^2 + 12xy + 4xz + 8yz + 18 = 0$;
   (3) $5x^2 + 2y^2 + 2z^2 - 2xy - 4yz + 2xz - 4y - 4z + 4 = 0$;
   (4) $4x^2 + 4y^2 - 8z^2 - 10xy + 4xz + 4yz - 16x - 16y + 10z - 2 = 0$;
   (5) $6x^2 + 5y^2 + 7z^2 + 4xy - 4xz - 8x - 10y + 14z - 6 = 0$.

2. 变量 t 取什么值时, 二次曲面

   $$x^2 + y^2 + z^2 + 2txy + 2txz + 2tyz - 4t = 0$$

   是椭球面?

3. 找出二次曲面与平面的相交曲线的仿射类型: (1) $3y^2 + 4z^2 + 24x + 12y - 72z + 360 = 0, \quad x - y + z = 1$.
   (2) $x^2 + 5y^2 + z^2 + 2xy + 6xz + 2yz - 2x + 6y + 2z = 0, \quad 2x - y + z = 0$.

射影空间

1. 求出 q 元域上 n 维射影空间的点的个数.

2. 设 $\mathbb{K} = \mathbb{F}_q$ 是 q 元域, 求出 $\mathrm{PGL}\left(\mathbb{K}^{n+1}\right)$.

3. 射影平面到自身的双射如果把直线映到直线并且保持每条直线上的交比不变, 那么它是射影变换.

4. 证明: 复射影空间上的射影变换有不动点.

5. 证明: 对复射影空间上的每个射影变换, 都存在一个仿射图使得该射影变换在仿射图上的作用是仿射变换的作用.

6. 根据推论5.3.1的 (1), 对射影直线

   $$\mathbb{P}^1 = \{(\alpha : \beta) \mid \alpha, \beta \in \mathbb{K} \text{ 且不全为零}\} = \mathbb{K} \cup \{\infty\}, \qquad \text{其中} \infty = (0; 1)$$

   上三个给定的互不相同的点 $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3$, 存在唯一的射影变换把 $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3$ 分别映到 $\infty, 0, 1$ (齐次坐标分别是 $(0:1), (1:0), (1:1)$). 证明:
   (1) 如果 $\tilde{u}_4$ 是 $\mathbb{P}^1$ 上的点，那么它在这个射影变换下的像 $\lambda \in \mathbb{K} \cup \{\infty\}$ 就是四点

$\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4$ 的交比;
(2) 对任意的 $\lambda \in \mathbb{K} \cup \{\infty\}$, 存在唯一的点 $\tilde{u}_4$ 使得 $[\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4] = \lambda$;
(3) 四个点的交比值在 $\mathbb{K} \setminus \{0, 1\}$ 中当且仅当四个点是互不相同的.

第一个结论可以作为交比的另一个定义, 并以此为基础建立交比的理论.

7. 假设 $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}$ 是射影直线上四个不同的点. 证明

$$[\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}] = [\tilde{y}, \tilde{x}, \tilde{z}, \tilde{w}]^{-1} = [\tilde{x}, \tilde{y}, \tilde{w}, \tilde{z}]^{-1}$$

$$[\overline{x}, \overline{y}, \overline{z}, \tilde{w}] + [\tilde{x}, \overline{z}, \overline{y}, \tilde{w}] = 1.$$

8. 设 f 是 $n+1$ 维向量空间 V 上的非奇异的双线性函数. 与每一个 $(k+1)$ 维子空间 $U \subset V$, 指定一个 $(n-k)$ 维子空间

$$U^{\perp} = \{y \in V \mid f(x, y) = 0 \text{ 对每个} x \in U\}.$$

这个对应在射影空间 $\mathbb{P}(V)$ 中诱导了一个映射 K_f, 它对于每个 k 维平面指定了一个 $(n-k-1)$ 平面 (关于函数 f 的对射变换). 证明:
a) 对射变换保持关联性:

$$U_1 \subset U_2 \leftrightarrow K_f\,(U_1) \supset K_f\,(U_2)\,;$$

b) 如果函数 f 是对称或斜对称的, 那么对射变换 K_f 是对合:

$$K_f\left(K_f(U)\right) = U;$$

c) 一个对射变换与一个射影变换的乘积是对射变换;
d) 任一对射变换是一个固定的对射变换和某个射影变换的乘积.

9. 求被包含在下述二次曲面里的平面的最大维数:
a) $x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 = 1$;
b) $x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2 = x_n$.

第六章　张量

现在我们来到了这本书的最后一章。在这一章中，我们将把双线性型和线性算子统一起来并加以推广，以更高的观点来回看之前学习的内容。张量的概念起源于微分几何，并且在物理和工程上都有广泛的应用。例如，广义相对论中，引力就是通过张量来描述的；量子力学中，复合系统也是通过张量来描述。张量具有非常灵巧的性质，学习过程中我们会时刻体会到这套语言在不同的代数对象之间起到的联系作用。这一章中，我们将大量使用交换图和泛性质这这类抽象工具，大家需要注意的是，所有的抽象工具都是有具体的构造实现和例子的，我们也将详细地说明每个抽象工具的具体意义，希望大家能够熟练地在抽象观点和具体构造之间进行转换。

6.1　张量计算初步

首先，我们将从抽象和具体两个角度分别构造张量的定义，并证明它们的等价性。我们从多重线性映射开始讲起。实际上，在讲义上册第三章的开头和本册的定义1.6.1中我们已经接触过多重线性映射了，这里再次重复它的定义。

定义 6.1.1. 设 $V_1,\dots,V_m,W$ 是域 $\mathbb{K}$ 上的向量空间，映射 $f:V_1\times\cdots\times V_m\to W$，如果对任意的 $i\in\{1,\dots,m\}$，固定除了 $\mathbf{x}_i$ 之外其余的变元，任取 $\alpha,\beta\in\mathbb{K}$ 及 $\mathbf{u},\mathbf{v}\in V_i$，都有

$$f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)=\alpha f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{u},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)+\beta f(\mathbf{x}_1,\dots,\mathbf{x}_{i-1},\mathbf{v},\mathbf{x}_{i+1},\dots,\mathbf{x}_m)$$

则称 f 是 m 重线性映射 (m-multilinear map)。即如果 f 在任意固定 $m-1$ 个变元后是关于剩下那个变元的线性映射，那么 f 就是 m 重线性映射。

下面两个多重线性映射的例子都是我们已经熟悉的。

例 6.1.1. (1) 设 A 是域 $\mathbb{K}$ 上的一个代数 (定义2.2.1)，则按照代数的定义，立刻有：映射

$$\begin{aligned} f:\ \underbrace{A\times\cdots\times A}_{n\text{个}} &\longrightarrow A\\ (a_1,a_2,\dots,a_n) &\longmapsto a_1\cdot a_2\cdots a_n\ (\text{代数上的乘法运算}) \end{aligned}$$

是多重线性映射。一个更具体的例子就是，线性算子的复合 (或者说矩阵的乘法) 关于每个参与的算子 (或矩阵) 就是一个多重线性映射。

(2) 在讲义上册的第三章中，我们已经知道，行列式函数：

$$\begin{aligned} \det:\ M_n(\mathbb{K})\simeq\underbrace{\mathbb{K}^n\times\cdots\times\mathbb{K}^n}_{n\text{个}} &\longrightarrow \mathbb{K}\\ B=(\mathbf{b}_1,\dots,\mathbf{b}_n) &\longmapsto \det(B) \end{aligned}$$

是一个关于列向量的多重线性映射。

下面的定理保证了向量空间张量积的存在唯一性。

定理 6.1.1. 设 $V_1, \ldots, V_n$ 是域 $\mathbb{K}$ 上的有限维向量空间，则存在向量空间 $T^n(V_1, \ldots, V_n)$ 和多重线性映射

$$\tau: \ V_1 \times \cdots \times V_n \longrightarrow T^n(V_1, \ldots, V_n)$$

使得以下结论 ($\star$) 成立：

($\star$) 任取域 $\mathbb{K}$ 上的向量空间 W 和 $V_1 \times \cdots \times V_n$ 映到 W 的多重线性映射 f，都存在唯一的线性映射 $\varphi_f: \ T^n(V_1, \ldots, V_n) \to W$ 使得 $f = \varphi_f \circ \tau$。也就是说，这个结论表明下面的图表中走红色箭头和蓝色箭头得到的映射是一样的，我们称之为这个图表是**交换**的 (commutative)：

$$\begin{tikzcd}
V_1 \times \cdots \times V_n \arrow[r, "\tau"] \arrow[d, "f"'] & T^n(V_1, \ldots, V_n) \arrow[dl, "\varphi_f"] \\
W &
\end{tikzcd}$$

更进一步地，我们有：如果存在 $\mathbb{K}$– 向量空间 U 和多重线性映射 $\tau' : V_1 \times \cdots \times V_n \to U$ 也满足上述结论 ($\star$)，那么这样的 U 一定都同构于 $T^n(V_1, \ldots, V_n)$，并且存在唯一的同构映射 $\psi : T^n(V_1, \ldots, V_n) \to U$ 使得 $\tau' = \psi \circ \tau$。换句话说，$T^n(V_1, \ldots, V_n)$ 在向量空间的同构这个意义下是唯一的。

这种交换图以后的学习中会经常遇到。在范畴论中，我们称使上述图表交换的 (同构意义下) 存在唯一的 $T^n(V_1, \ldots, V_n)$ 满足**万有性质** (也翻译成**泛性质**, universal property)。

证明. 为使叙述简单，我们只证明 $n = 2$ 的情形，$n > 2$ 情形的证明是很类似的，留作练习。

不妨设 V_1 的一组基是 $\mathbf{e}_1, \ldots, \mathbf{e}_s$，$V_2$ 的一组基是 $\mathbf{e}'_1, \ldots, \mathbf{e}'_t$，我们构造如下的向量空间 $T^2(V_1, V_2)$:

首先，我们将这两组基的集合作笛卡尔乘积，并且把乘积集合中的元素记作 $\mathbf{e}_i \otimes \mathbf{e}'_j$；也就是说，我们得到了下面的集合

$$\Gamma = \{\mathbf{e}_i \otimes \mathbf{e}'_j \mid i = 1, \ldots, s, \ \ j = 1, \ldots, t\} = \{\mathbf{e}_1, \ldots, \mathbf{e}_s\} \times \{\mathbf{e}'_1, \ldots, \mathbf{e}'_t\}.$$

我们以 Γ 中的元素为基底，在域 $\mathbb{K}$ 上张成一个 st 维的向量空间，这个向量空间就是满足要求的 $T^2(V_1, V_2)$，即：

$$T^2(V_1, V_2) = \operatorname{span} \{\mathbf{e}_i \otimes \mathbf{e}'_j \mid i = 1, \ldots, s, \ \ j = 1, \ldots, t\}.$$

下面证明上面的构造确实满足我们的要求。事实上，我们只需构造出 τ 即可。类似于 1.6 节双线性型中式 (1.6.1) 的推导，由于 τ 对每一个变元都是线性的，我们知道: $\forall \mathbf{x} = \sum_{i=1}^{s} x_i \mathbf{e}_i \in V_1$

和 $\mathbf{y} = \sum_{j=1}^{t} y_j\mathbf{e}'_j \in V_2$，都有

$$\begin{aligned}\tau(\mathbf{x},\mathbf{y}) &= \tau(\sum_{i=1}^{s} x_i\mathbf{e}_i, \sum_{j=1}^{t} y_j\mathbf{e}'_j) \\ &= \sum_{i=1}^{s} x_i \cdot \tau(\mathbf{e}_i, \sum_{j=1}^{t} y_j\mathbf{e}'_j) \\ &= \sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot \tau(\mathbf{e}_i, \mathbf{e}'_j)\end{aligned} \tag{6.1.1}$$

因此我们只需要定义出每一个 τ 作用在基底上的像 $\tau(\mathbf{e}_i, \mathbf{e}'_j)$ 就可以得到完全定义出 τ。于是，我们定义 2 重线性映射 τ 如下：

$$\begin{aligned}\tau:\ V_1 \times V_2 &\longrightarrow T^2(V_1, V_2) \\ (\mathbf{e}_i, \mathbf{e}'_j) &\longmapsto \mathbf{e}_i \otimes \mathbf{e}'_j\end{aligned}$$

接下来我们验证定理中的交换图。由于 f 也是 2 重线性映射，与式 (6.1.1) 的推导一样，我们有：

$$\begin{aligned}f(\mathbf{x},\mathbf{y}) &= f(\sum_{i=1}^{s} x_i\mathbf{e}_i, \sum_{j=1}^{t} y_j\mathbf{e}'_j) \\ &= \sum_{i=1}^{s} x_i \cdot f(\mathbf{e}_i, \sum_{j=1}^{t} y_j\mathbf{e}'_j) \\ &= \sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot f(\mathbf{e}_i, \mathbf{e}'_j)\end{aligned}$$

构造如下的 φ_f:

$$\varphi_f:\ T^2(V_1, V_2) \to W,\ \ \mathbf{e}_i \otimes \mathbf{e}'_j \mapsto f(\mathbf{e}_i, \mathbf{e}'_j)$$

这是 φ_f 在基底上的定义，由线性映射基本定理 (定理2.1.2(3)) 可知 φ_f 是良定义的，下面我们验证交换图：$\forall \mathbf{x} = \sum_{i=1}^{s} x_i\mathbf{e}_i \in V_1$ 和 $\mathbf{y} = \sum_{j=1}^{t} y_j\mathbf{e}'_j \in V_2$，都有：

$$\begin{aligned}\varphi_f \circ \tau(\mathbf{x},\mathbf{y}) &= \varphi_f(\sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot \tau(\mathbf{e}_i, \mathbf{e}'_j)) \\ &= \sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot \varphi_f(\tau(\mathbf{e}_i, \mathbf{e}'_j)) \\ &= \sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot \varphi_f(\mathbf{e}_i \otimes \mathbf{e}'_j) \\ &= \sum_{i=1}^{s}\sum_{j=1}^{t} x_i y_j \cdot f(\mathbf{e}_i, \mathbf{e}'_j) \\ &= f(\sum_{i=1}^{s} x_i\mathbf{e}_i, \sum_{j=1}^{t} y_j\mathbf{e}'_j) \\ &= f(\mathbf{x},\mathbf{y}).\end{aligned}$$

此即 $\varphi_f \circ \tau = f$。

此外，φ_f 的选择是唯一的。因为如果另有映射 σ_f 满足和 φ_f 一样的条件，那么必有

$\sigma_f\circ\tau(\mathbf{e}_i,\mathbf{e}'_j)=f(\mathbf{e}_i,\mathbf{e}'_j)$，即

$$\sigma_f(\mathbf{e}_i\otimes\mathbf{e}'_j)=f(\mathbf{e}_i,\mathbf{e}'_j)=\varphi_f(\mathbf{e}_i\otimes\mathbf{e}'_j).$$

即 σ_f 和 φ_f 在基底上的像相同。由线性映射基本定理即得 $\sigma_f=\varphi_f$，即 φ_f 的选择是唯一的。

最后，我们来证明 $T^2(V_1,V_2)$ 在同构意义下是唯一的。这可以从如下的交换图中得到：

$$\begin{tikzcd}
V_1\times V_2 \arrow[r, "\tau"] \arrow[d, "\tau'"'] & T^2(V_1,V_2) \arrow[dl, shift right, "\varphi_\tau"'] \\
U \arrow[ur, shift right, "\varphi_{\tau'}"'] &
\end{tikzcd}$$

也就是说，如果将 U 视作结论 $(\star)$ 中的 W，那么存在唯一的线性映射 φ_τ 使得

$$\tau'=\varphi_\tau\circ\tau, \tag{6.1.2}$$

反过来，如果将 U 放在原来 $T^2(V_1,V_2)$ 的位置，而将 $T^2(V_1,V_2)$ 视作结论 $(\star)$ 中的 W 的话，则存在唯一的线性映射 $\varphi_{\tau'}$ 使得

$$\tau=\varphi_{\tau'}\circ\tau'. \tag{6.1.3}$$

将式 (6.1.2) 和式 (6.1.3) 互相代入彼此就得到：

$$\tau'=\varphi_\tau\circ\varphi_{\tau'}\circ\tau',\quad \tau=\varphi_{\tau'}\circ\varphi_\tau\circ\tau.$$

即如下交换图成立：

$$\begin{tikzcd}
V_1\times V_2 \arrow[r, "\tau'"] \arrow[d, "\tau'"'] & U \arrow[dl, shift right, "\mathrm{id}"'] \arrow[dl, shift left, "\varphi_\tau\circ\varphi_{\tau'}"] \\
U &
\end{tikzcd}
\qquad
\begin{tikzcd}
V_1\times V_2 \arrow[r, "\tau"] \arrow[d, "\tau"'] & T^2(V_1,V_2) \arrow[dl, shift right, "\mathrm{id}"'] \arrow[dl, shift left, "\varphi_{\tau'}\circ\varphi_\tau"] \\
T^2(V_1,V_2) &
\end{tikzcd}$$

由泛性质 $(\star)$ 中的唯一性，上图说明 $\varphi_\tau\circ\varphi_{\tau'}=\mathrm{id}_U$，$\varphi_{\tau'}\circ\varphi_\tau=\mathrm{id}_{T^2(V_1,V_2)}$，即 φ_τ 和 $\varphi_{\tau'}$ 是互逆的同构映射。也就是说，$T^2(V_1,V_2)$ 在同构意义下是唯一的，并且 φ_τ 就是定理中所求的 ψ。这样我们就完成了证明。

□

现在我们可以定义张量了。

定义 6.1.2 (张量). 条件如定理6.1.1。则我们称向量空间 $T^n(V_1,\dots,V_n)$ 为 $V_1,\dots,V_n$ 的张量积空间 (tensor product)，记为 $V_1\otimes_{\mathbb{K}}\cdots\otimes_{\mathbb{K}}V_n$(在基域 $\mathbb{K}$ 明确的情况下可以省略角标 $\mathbb{K}$)；映射 τ 称为张量映射或张量运算 (tensor map)；$V_1\otimes\cdots\otimes V_n$ 中的元素称为 (n 阶反变) 张量 (tensors)，n 称为张量的阶 (contravariant order)。

由于张量运算 τ 是多重线性映射，我们立刻有以下的运算律：

$$\mathbf{x}^{(1)}\otimes\cdots\otimes(\sum_{j=1}^{k}a_j\mathbf{x}_j^{(i)})\otimes\cdots\otimes\mathbf{x}^{(n)}=\sum_{j=1}^{k}a_j\left(\mathbf{x}^{(1)}\otimes\cdots\otimes\mathbf{x}_j^{(i)}\otimes\cdots\otimes\mathbf{x}^{(n)}\right),\quad \forall\mathbf{x}^{(i)}\in V_i,\ a_j\in\mathbb{K}.$$

现在我们将张量积空间中的元素明确地写下来。设对每个 $j\in\{1,\dots,n\}$，V_j 的一组基为 $\mathbf{e}_1^{(j)},\dots,\mathbf{e}_{m_j}^{(j)}$，则由定理6.1.1的证明过程，我们知道 $V_1\otimes\cdots\otimes V_n$ 的基是

$$\{\mathbf{e}_{i_1}^{(1)} \otimes \mathbf{e}_{i_2}^{(2)} \otimes \cdots \otimes \mathbf{e}_{i_n}^{(n)} \mid i_j \in \{1, \ldots, m_j\}, \quad j \in \{1, \ldots, n\}\}. \tag{6.1.4}$$

也就是说，$V_1 \otimes \cdots \otimes V_n$ 中的一个张量 $\mathbf{T}$ 一定可以写成下面的形式：

$$\mathbf{T} = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \cdots \sum_{i_n=1}^{m_n} T_{i_1,i_2,\ldots,i_n} \cdot \mathbf{e}_{i_1}^{(1)} \otimes \mathbf{e}_{i_2}^{(2)} \otimes \cdots \otimes \mathbf{e}_{i_n}^{(n)}, \quad T_{i_1,i_2,\ldots,i_n} \in \mathbb{K}.$$

我们把系数 $T_{i_1,i_2,\ldots,i_n}$ 称为张量在上述这组基下的**坐标**。此外，容易看出张量积空间 $V_1 \otimes \cdots \otimes V_n$ 的维数是 $\prod_{j=1}^{n} m_j$，其中 $m_j = \dim(V_j)$。

从定理6.1.1的证明过程来看，我们在构造张量积空间 $T^n(V_1, \ldots, V_n)$ 时是依赖于基底的选取的，而后再利用泛性质的同构抹去基底的影响。那么，我们是否可以不依赖取基的方法而直接构造张量积空间呢？这就是下面的证明。

定理6.1.1另证：仍然只证明 $n = 2$ 的情形。将 $V_1 \times V_2$ 视作集合，以其中的所有有序对 $(\mathbf{x}, \mathbf{y})$（视作一个符号）为基底，在 $\mathbb{K}$ 上张成一个向量空间 S，即

$$S = \text{span}\ \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in V_1,\ \mathbf{y} \in V_2\} = \{\sum_{i=1}^{s} a_i(\mathbf{x}_i, \mathbf{y}_i),\ \forall a_i \in \mathbb{K},\ s \in \mathbb{N}\}.$$

注意 S 可能是无穷维的。用所有以下形式的元素张成一个 S 中的子空间 U：

$$(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) - (\mathbf{x}_1, \mathbf{y}) - (\mathbf{x}_2, \mathbf{y}),\ (\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) - (\mathbf{x}, \mathbf{y}_1) - (\mathbf{x}, \mathbf{y}_2),\ (a\mathbf{x}, \mathbf{y}) - a(\mathbf{x}, \mathbf{y}),\ (\mathbf{x}, b\mathbf{y}) - b(\mathbf{x}, \mathbf{y}),$$

其中 $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2 \in V_1$，$\mathbf{y}, \mathbf{y}_1, \mathbf{y}_2 \in V_2$ 和 $a, b \in \mathbb{K}$ 都是任意的。

那么，令 $T^2(V_1, V_2) = S/U$ 即为所求的满足结论 $(\star)$ 的向量空间，并且 $\tau: (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, \mathbf{y}) + U$。我们把 S/U 中的元素 $\sum_{i=1}^{s} a_i(\mathbf{x}_i, \mathbf{y}_i) + U$ 也记作 $\sum_{i=1}^{s} a_i \mathbf{x}_i \otimes \mathbf{y}_i$。验证结论 $(\star)$ 是容易的，只需注意到对任意的双线性映射 $f: V_1 \times V_2 \to W$，取

$$\varphi_f: S/U \to W,\ \sum_{i=1}^{s} a_i \mathbf{x}_i \otimes \mathbf{y}_i = \sum_{i=1}^{s} a_i(\mathbf{x}_i, \mathbf{y}_i) + U \mapsto \sum_{i=1}^{s} a_i \cdot f(\mathbf{x}_i, \mathbf{y}_i)$$

即可得到泛性质所需的交换图。验证 φ_f 的良定性和唯一性留作练习。而 $T^2(V_1, V_2)$ 同构意义下的唯一性已经在上一个证法中完成了。至此证毕。

第二个证明可以推广到 $V_1, \ldots, V_n$ 中有无穷维空间的情形。比较这两个证明可以看到，第二个证明的构造更为抽象。那么，如果我们采用这个证明，式 (6.1.4) 给出的张量基空间的基底就不再是一个直接的推论，而是需要如下的证明。

定理 6.1.2. *张量积空间的一组基是式* (6.1.4)，*与* $T^n(V_1, \ldots, V_n)$ *的构造方式无关。*

证明. 为使叙述简单，仍然只证明 $n = 2$ 的情形。设 V_1 的一组基是 $\mathbf{e}_1, \ldots, \mathbf{e}_s$，$V_2$ 的一组基是 $\mathbf{e}'_1, \ldots, \mathbf{e}'_t$，显然任何一个张量 $\sum_{m=1}^{k} a_m \mathbf{x}_m \otimes \mathbf{y}_m$ 都可以拆写成 $\sum_{i,j} b_{ij} \mathbf{e}_i \otimes \mathbf{e}'_j$ 的形式（利用张量运算是多重线性映射，将 $\mathbf{x}_m, \mathbf{y}_m$ 拆开即可）。因此，我们只需证明所有的

$$\{\mathbf{e}_i \otimes \mathbf{e}'_j \mid i = 1, \ldots, s,\ j = 1, \ldots, t\}$$

线性无关。即我们只需证：如果

$$\sum_{i=1}^{s}\sum_{j=1}^{t}\lambda_{ij}\mathbf{e}_i\otimes\mathbf{e}'_j=\mathbf{0}_{V_1}\otimes\mathbf{0}_{V_2},$$

那么必有 $\lambda_{ij}=0,\ \forall i,j$。注意到

$$\sum_{i=1}^{s}\sum_{j=1}^{t}\lambda_{ij}\mathbf{e}_i\otimes\mathbf{e}'_j=\sum_{i=1}^{s}\mathbf{e}_i\otimes(\sum_{j=1}^{t}\lambda_{ij}\mathbf{e}'_j),$$

因此我们只需证明每个 $\sum_{j=1}^{t}\lambda_{ij}\mathbf{e}'_j$ 都是 $\mathbf{0}_{V_2}$，就可以利用 $\mathbf{e}'_j$ 的线性无关性证明每个 λ_{ij} 都是 0。记 $\sum_{j=1}^{t}\lambda_{ij}\mathbf{e}'_j=\mathbf{y}_i$，我们现在来证每个 $\mathbf{y}_i=\mathbf{0}_{V_2}$。

在对偶空间中任取线性函数 $\alpha\in V_1^*$ 和 $\beta\in V_2^*$，容易验证下面定义的

$$g:V_1\times V_2\to\mathbb{K},\ (\mathbf{x},\mathbf{y})\mapsto\alpha(\mathbf{x})\cdot\beta(\mathbf{y})$$

是双线性映射，于是由定理6.1.1，存在 $\varphi_g:V_1\otimes V_2$ 使得

$$\varphi_g(\mathbf{x}\otimes\mathbf{y})=g(\mathbf{x},\mathbf{y})=\alpha(\mathbf{x})\cdot\beta(\mathbf{y}),\ \forall\mathbf{x}\in V_1,\ \mathbf{y}\in V_2.$$

于是，将 φ_g 作用到 $\sum_{i=1}^{s}\mathbf{e}_i\otimes\mathbf{y}_i=\mathbf{0}_{V_1}\otimes\mathbf{0}_{V_2}$ 两边，我们立刻得到：

$$\mathbf{0}_{V_1}\otimes\mathbf{0}_{V_2}=\varphi_g(\sum_{i=1}^{s}\mathbf{e}_i\otimes\mathbf{y}_i)=\sum_{i=1}^{s}\varphi_g(\mathbf{e}_i\otimes\mathbf{y}_i)=\sum_{i=1}^{s}g(\mathbf{e}_i,\mathbf{y}_i)=\sum_{i=1}^{s}\alpha(\mathbf{e}_i)\cdot\beta(\mathbf{y}_i).$$

不妨使 α 取遍对偶基中的每个元素 $\mathbf{e}^k$，$k\in\{1,\dots,s\}$，则此时 $\mathbf{e}^k(\mathbf{e}_i)=\delta_{ki}$，于是对每个 $\mathbf{e}^k$ 有：

$$\mathbf{0}_{V_1}\otimes\mathbf{0}_{V_2}=\sum_{i=1}^{s}\mathbf{e}^k(\mathbf{e}_i)\cdot\beta(\mathbf{y}_i)=\beta(\mathbf{y}_k).$$

注意到 β 是任意选取的，因此只能是每一个 $\mathbf{y}_k$ 都是零向量，这样我们就完成了证明。 □

这个证明过程也告诉我们以下命题：

命题 6.1.1. 设 V_1,V_2 是域 $\mathbb{K}$ 上的向量空间，$\mathbf{u}_1,\dots,\mathbf{u}_s\in V_1$，$\mathbf{v}_1,\dots,\mathbf{v}_s\in V_2$。如果 $\sum_{i=1}^{s}\mathbf{u}_i\otimes\mathbf{v}_i=\mathbf{0}$，而 $\mathbf{u}_1,\dots,\mathbf{u}_s$ 线性无关，则一定是 $\mathbf{v}_1=\cdots=\mathbf{v}_s=\mathbf{0}$。

下面我们先来看两个例子。

例 6.1.2. 设 $V_1=V_2=\mathbb{R}^2$，考虑 $V=V_1\otimes_{\mathbb{R}}V_2$。首先 $\dim(V)=4$，记 V_1 的基为 $\mathbf{e}_1=(1,0)^t$，$\mathbf{e}_2=(0,1)^t$，V_2 的基为 $\mathbf{e}'_1=(1,0)^t$，$\mathbf{e}'_2=(0,1)^t$，则 $V_1\otimes_{\mathbb{R}}V_2$ 的一组基为 $\mathbf{e}_1\otimes\mathbf{e}'_1$，$\mathbf{e}_1\otimes\mathbf{e}'_2$，$\mathbf{e}_2\otimes\mathbf{e}'_1$，$\mathbf{e}_2\otimes\mathbf{e}'_2$。我们可以用如下方式实现这个张量积空间：

(1) 作映射 $V\to M_2(\mathbb{R})$，$\mathbf{e}_i\otimes\mathbf{e}'_j\mapsto\mathbf{e}_i\cdot(\mathbf{e}'_j)^t$，容易验证这是一个同构，也就是说，

$$V\simeq\operatorname{span}\left\{\begin{pmatrix}1\\0\end{pmatrix}(1,0),\ \begin{pmatrix}1\\0\end{pmatrix}(0,1),\ \begin{pmatrix}0\\1\end{pmatrix}(1,0),\ \begin{pmatrix}0\\1\end{pmatrix}(0,1)\right\}=\operatorname{span}\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$$

需要注意的是，此时 V 中的每个张量都可以同构到秩为 1 的矩阵的线性组合，但却不一定能同构到 $\mathbf{x}\cdot\mathbf{y}^t$，$\mathbf{x}\in V_1,\mathbf{y}\in V_2$ 的形式，因为显然存在秩为 2 的矩阵。这种现象可

以用于证明量子力学中纠缠态的存在。

(2) 我们也可以将 V 直接同构到 $\mathbb{R}^4$。令

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \mapsto \begin{pmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{pmatrix},$$

我们把这个运算称为矩阵的 Kronecker 乘积 (也有的书称作直积)，则 Kronecker 乘积就给出了另一个同构：

$$V \simeq \operatorname{span} \{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \}.$$

我们会在稍后正式定义 Kronecker 乘积，并讨论它和张量之间的联系。

需要注意的是，一个张量可能有不同的表达式，或者说，两个不同的表达式

$$\mathbf{T}_1 = \sum_{i=1}^{s} \mathbf{x}_i \otimes \mathbf{y}_i \quad \text{和} \quad \mathbf{T}_2 = \sum_{j=1}^{t} \mathbf{x}'_j \otimes \mathbf{y}'_j, \quad \mathbf{x}_i, \mathbf{x}'_j \in V_1, \ \mathbf{y}_i, \mathbf{y}'_j \in V_2$$

可能表示同一个张量。这可以从下面的例子中看出来。

例 6.1.3. 仍然考虑例6.1.2中的张量积空间 V。令

$$\mathbf{T}_1 = \mathbf{e}_1 \otimes \mathbf{e}'_1 + \mathbf{e}_2 \otimes \mathbf{e}'_1 + \mathbf{e}_2 \otimes \mathbf{e}'_1 + \mathbf{e}_2 \otimes \mathbf{e}'_2, \quad \mathbf{T}_2 = (\mathbf{e}_1 + \mathbf{e}_2) \otimes (\mathbf{e}'_1 + \mathbf{e}'_2).$$

则很容易验证这两个张量相等，放在 $\mathbb{R}^4$ 中它们都等于 $(1, 1, 1, 1)^t$。

现在我们先给出一些张量运算的运算律。

定理 6.1.3. 设 V_1, V_2, V_3 都是域 $\mathbb{K}$ 上的向量空间，则

(1) 结合律：$(V_1 \otimes V_2) \otimes V_3 \simeq V_1 \otimes (V_2 \otimes V_3) \simeq V_1 \otimes V_2 \otimes V_3$;

(2) 交换律：$V_1 \otimes V_2 \simeq V_2 \otimes V_1$;

(3) 分配律：$(V_1 \oplus V_2) \otimes V_3 \simeq (V_1 \otimes V_3) \oplus (V_2 \otimes V_3)$;

(4) $V_1 \otimes_{\mathbb{K}} \mathbb{K} \simeq \mathbb{K} \otimes_{\mathbb{K}} V_1 \simeq V_1$.

证明. 不妨设 $\mathbf{x}_i \in V_i$，$i = 1, 2, 3$，$\lambda \in \mathbb{K}$。只需注意到以下的映射是同构即可，细节留作练习 (注意需要验证这些映射的良定义性)：

(1) $(\mathbf{x}_1 \otimes \mathbf{x}_2) \otimes \mathbf{x}_3 \mapsto \mathbf{x}_1 \otimes (\mathbf{x}_2 \otimes \mathbf{x}_3) \mapsto \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \mathbf{x}_3$;

(2) $\mathbf{x}_1 \otimes \mathbf{x}_2 \mapsto \mathbf{x}_2 \otimes \mathbf{x}_1$;

(3) $(\mathbf{x}_1, \mathbf{x}_2) \otimes \mathbf{x}_3 \mapsto (\mathbf{x}_1 \otimes \mathbf{x}_3, \mathbf{x}_2 \otimes \mathbf{x}_3)$;

(4) $\mathbf{x}_1 \otimes_{\mathbb{K}} \lambda \mapsto \lambda \otimes_{\mathbb{K}} \mathbf{x}_1 \mapsto \lambda \mathbf{x}_1$.

我们把以上的同构映射称为标准同构 (注意这些同构不能用等号代替，因为这些张量积空间背后代表了不一样的泛性质)。 □

上面的定理很容易推广到多个空间的情形，其形式可以参考《Advanced Linear Algebra》，Steven Roman，GTM135 的定理 14.12。证明与上面类似，留作练习。

利用张量可以很容易地理解 3.6 节中实向量空间的复化。我们先讨论更一般的情形：向量空间基域的扩张。设域 $\mathbb{K}$ 是域 $\mathbb{F}$ 的子域，那么，显然 $\mathbb{F}$ 可以视作 $\mathbb{K}$ 上的向量空间 (验证它满足定义1.1.1的所有条件即可)，于是，如果 V 是域 $\mathbb{K}$ 上的向量空间，那么我们可以作向量空间

$$V^{\mathbb{F}} = V \otimes_{\mathbb{K}} \mathbb{F},$$

注意到 $V^{\mathbb{F}}$ 仍然是 $\mathbb{K}$ 上的向量空间，但我们也可以将它视作 $\mathbb{F}$ 上的向量空间，这只需要保持原来的加法运算不变，而数乘 $\mathbb{F} \times V^{\mathbb{F}} \to V^{\mathbb{F}}$ 用以下方式定义：

首先，因为对任意的 $\lambda \in \mathbb{F}$，容易看出映射

$$V \times \mathbb{F} \to V \otimes_{\mathbb{K}} \mathbb{F},\ (\mathbf{x}, a) \mapsto \mathbf{x} \otimes_{\mathbb{K}} (\lambda a) \tag{6.1.5}$$

是域 $\mathbb{K}$ 上的双线性映射，因此按照泛性质，映射 (6.1.5) 唯一地确定了如下的关于域 $\mathbb{K}$ 的线性映射：

$$\varphi_\lambda :\ V \otimes_{\mathbb{K}} \mathbb{F} \to V \otimes_{\mathbb{K}} \mathbb{F},\ \mathbf{x} \otimes_{\mathbb{K}} a \mapsto \mathbf{x} \otimes_{\mathbb{K}} (\lambda a)$$

即以下交换图成立：

$$\begin{tikzcd}
V \times \mathbb{F} \arrow[r] \arrow[dd] & V \otimes_{\mathbb{K}} \mathbb{F} \arrow[ddl] \\
(\mathbf{x}, a) & \mathbf{x} \otimes_{\mathbb{K}} a \\
V \otimes_{\mathbb{K}} \mathbb{F} & \\
\mathbf{x} \otimes_{\mathbb{K}} (\lambda a) &
\end{tikzcd}$$

于是，对任意的 $\mathbf{y} = \sum_{i=1}^{s} \mathbf{x}_i \otimes_{\mathbb{K}} a_i \in V^{\mathbb{F}}$ 和 $\lambda \in \mathbb{F}$，我们定义数乘运算如下：

$$\cdot :\ \mathbb{F} \times V^{\mathbb{F}} \to V^{\mathbb{F}},\ (\lambda, \mathbf{y}) \mapsto \varphi_\lambda(\mathbf{y}) = \sum_{i=1}^{s} \mathbf{x}_i \otimes_{\mathbb{K}} (\lambda a_i)$$

此即我们所需的的数乘映射。由于泛性质给出的 φ_λ 是唯一的，因此这样定义 $\mathbb{F}$ 上的数乘是良定义的。

然而，φ_λ 只是域 $\mathbb{K}$ 上的线性映射，要说明它是 $\mathbb{F}$ 上的的数乘，我们还要验证它满足定义1.1.1中 (2) 和 (3)，其过程如下：

- $1 \cdot \mathbf{y} = \mathbf{y}$ 显然，因为 1 在 $\mathbb{K}$ 中。

- 验证 $\lambda(\mu\mathbf{y}) = (\lambda\mu)\mathbf{y}$，$\forall \lambda, \mu \in \mathbb{F}$，$\mathbf{y} \in V^{\mathbb{F}}$：按照泛性质，这只需验证上式对形如 $\mathbf{y} =$

$\mathbf{x}\otimes_{\mathbb{K}} a,\ \mathbf{x}\in V,\ a\in\mathbb{F}$ 的 $\mathbf{y}$ 成立即可。

$$\begin{aligned}
\lambda\cdot(\mu\cdot(\mathbf{x}\otimes_{\mathbb{K}} a)) &= \lambda\cdot(\mathbf{x}\otimes_{\mathbb{K}}(\mu a)) && (\mathbb{F}\text{上数乘的定义})\\
&= \mathbf{x}\otimes_{\mathbb{K}}(\lambda(\mu a)) && (\mathbb{F}\text{上数乘的定义})\\
&= \mathbf{x}\otimes_{\mathbb{K}}((\lambda\mu)a) && (\mathbb{F}\text{的乘法结合律})\\
&= (\lambda\mu)\cdot(\mathbf{x}\otimes_{\mathbb{K}} a) && (\mathbb{F}\text{上数乘的定义})
\end{aligned}$$

此即数乘结合律成立。

- 验证分配律 $(\lambda+\mu)\mathbf{y}=\lambda\mathbf{y}+\mu\mathbf{y},\ \lambda(\mathbf{y}+\mathbf{z})=\lambda\mathbf{y}+\lambda\mathbf{z},\ \forall\lambda,\mu\in\mathbb{F},\ \mathbf{y},\mathbf{z}\in V^{\mathbb{F}}$。仍然只需验证形如 $\mathbf{y}=\mathbf{x}\otimes_{\mathbb{K}} a,\ \mathbf{z}=\mathbf{w}\otimes_{\mathbb{K}} b,\ \mathbf{x},\mathbf{w}\in V,\ a,b\in\mathbb{F}$ 的情形。

$$\begin{aligned}
(\lambda+\mu)\cdot(\mathbf{x}\otimes_{\mathbb{K}} a) &= \mathbf{x}\otimes_{\mathbb{K}}((\lambda+\mu)a) && (\mathbb{F}\text{上数乘的定义})\\
&= \mathbf{x}\otimes_{\mathbb{K}}(\lambda a+\mu a) && (\mathbb{F}\text{的乘法分配律})\\
&= \mathbf{x}\otimes_{\mathbb{K}}(\lambda a)+\mathbf{x}\otimes_{\mathbb{K}}(\mu a) && (\text{张量运算的多重线性性})\\
&= \lambda\cdot(\mathbf{x}\otimes_{\mathbb{K}} a)+\mu(\mathbf{x}\otimes_{\mathbb{K}} a) && (\mathbb{F}\text{上数乘的定义})
\end{aligned}$$

$$\begin{aligned}
\lambda(\mathbf{x}\otimes_{\mathbb{K}} a+\mathbf{w}\otimes_{\mathbb{K}} b) &= \varphi_\lambda(\mathbf{x}\otimes_{\mathbb{K}} a+\mathbf{w}\otimes_{\mathbb{K}} b) && (\text{定义})\\
&= \varphi_\lambda(\mathbf{x}\otimes_{\mathbb{K}} a)+\varphi_\lambda(\mathbf{w}\otimes_{\mathbb{K}} b) && (\varphi_\lambda\text{是线性映射})\\
&= \lambda\cdot(\mathbf{x}\otimes_{\mathbb{K}} a)+\lambda\cdot(\mathbf{w}\otimes_{\mathbb{K}} b) && (\text{定义})
\end{aligned}$$

这样我们就完成了所有验证，$V^{\mathbb{F}}$ 也是确实是 $\mathbb{F}$ 上的向量空间。下面我们考虑 $V^{\mathbb{F}}$ 作为 $\mathbb{F}$ 上的向量空间的维数。我们有以下命题：

命题 6.1.2. *设 $\mathbb{K}$ 是 $\mathbb{F}$ 的子域，V 是 $\mathbb{K}$ 上的 n 维向量空间，其一组基为 $\mathbf{e}_1,\dots,\mathbf{e}_n$。则 $V^{\mathbb{F}}$ 作为 $\mathbb{F}$ 上的向量空间的维数也是 n，其一组基为 $\mathbf{e}_1\otimes_{\mathbb{K}}1,\dots,\mathbf{e}_n\otimes_{\mathbb{K}}1$。*

证明. 先证明 $\mathbf{e}_1\otimes_{\mathbb{K}}1,\dots,\mathbf{e}_n\otimes_{\mathbb{K}}1$ 在 $\mathbb{F}$ 上线性无关。设有 $\lambda_1,\dots,\lambda_n\in\mathbb{F}$ 使得

$$\lambda_1(\mathbf{e}_1\otimes_{\mathbb{K}}1)+\cdots+\lambda_n(\mathbf{e}_n\otimes_{\mathbb{K}}1)=\mathbf{0}_{V^{\mathbb{F}}},$$

则按照 $V^{\mathbb{F}}$ 中数乘的定义，我们有

$$\mathbf{e}_1\otimes_{\mathbb{K}}\lambda_1+\cdots+\mathbf{e}_n\otimes_{\mathbb{K}}\lambda_n=\lambda_1(\mathbf{e}_1\otimes_{\mathbb{K}}1)+\cdots+\lambda_n(\mathbf{e}_n\otimes_{\mathbb{K}}1)=\mathbf{0}_{V^{\mathbb{F}}},$$

由于 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 线性无关，由命题6.1.1立刻有 $\lambda_1=\cdots=\lambda_n=0_{\mathbb{F}}$，此即线性无关性。

再证明任何一个 $V^{\mathbb{F}}$ 中的元素可以由 $\mathbf{e}_1\otimes_{\mathbb{K}}1,\dots,\mathbf{e}_n\otimes_{\mathbb{K}}1$ 在 $\mathbb{F}$ 上线性表出。任取 $\mathbf{y}=\sum_{i=1}^s\mathbf{x}_i\otimes_{\mathbb{K}}a_i,\ \mathbf{x}_i\in V,\ a_i\in\mathbb{F}$，我们知道每个 $\mathbf{x}_i$ 都可以由 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 线性表出，即每个 $\mathbf{x}_i$ 都可以写成

$$\mathbf{x}_i=\sum_{j=1}^n x_{ij}\mathbf{e}_j,\ x_{ij}\in\mathbb{K}.$$

那么我们有

$$\begin{aligned}\mathbf{y} &= \sum_{i=1}^{s} \mathbf{x}_i \otimes_{\mathbb{K}} a_i \\ &= \sum_{i=1}^{s}(\sum_{j=1}^{n} x_{ij}\mathbf{e}_j) \otimes_{\mathbb{K}} a_i \\ &= \sum_{i=1}^{s} a_i \cdot (\sum_{j=1}^{n} x_{ij}(\mathbf{e}_j \otimes_{\mathbb{K}} 1)) \qquad & \text{(用到了数乘在 } \mathbb{K} \text{ 上是线性映射)} \\ &= \sum_{j=1}^{n}(\sum_{i=1}^{s} a_i x_{ij}) \cdot (\mathbf{e}_j \otimes_{\mathbb{K}} 1)) \qquad & \text{(交换求和顺序)}\end{aligned}$$

显然每个 $\sum_{i=1}^{s} a_i x_{ij} \in \mathbb{F}$，此即任何一个 $V^{\mathbb{F}}$ 中的元素可以由 $\mathbf{e}_1 \otimes_{\mathbb{K}} 1, \dots, \mathbf{e}_n \otimes_{\mathbb{K}} 1$ 在 $\mathbb{F}$ 上线性表出。这样我们就完成了证明。 □

利用基域的扩张，我们可以将线性映射也扩张到更大的域的向量空间上：

命题 6.1.3. 设 $\mathbb{K}$ 是 $\mathbb{F}$ 的子域，V 是 $\mathbb{K}$ 上向量空间，W 是 $\mathbb{F}$ 上的向量空间 (按定义 W 自然也是 $\mathbb{K}$ 上的向量空间，验证留作练习)。那么，对于任意一个 $\mathbb{K}$ 上的线性映射 $f : V \to W$，都存在唯一的 $\mathbb{F}$ 上的线性映射 $\varphi_f : V^{\mathbb{F}} \to W$，使得

$$f = \varphi_f \circ \tau.$$

其中，τ 的定义是 $V \to V^{\mathbb{F}}(= V \otimes_{\mathbb{K}} \mathbb{F})$，$\mathbf{v} \mapsto \mathbf{v} \otimes_{\mathbb{K}} 1_{\mathbb{F}}$。也即下面的交换图成立：

$$\begin{tikzcd} V \arrow[d, "\tau"'] \arrow[dr, "f"] & \\ V^{\mathbb{F}} \arrow[r, "\varphi_f"'] & W \end{tikzcd} \tag{6.1.6}$$

证明. 注意到 $f : V \to W$ 诱导了 $\mathbb{K}$ 上的双线性映射 (验证留作练习)：

$$f' : V \times \mathbb{F} \to W, \ (\mathbf{v}, \lambda) \mapsto \lambda \cdot f(\mathbf{v}),$$

于是由泛性质，存在唯一的 $\mathbb{K}$− 线性映射 $\varphi_{f'} : V \otimes_{\mathbb{K}} \mathbb{F} \to W$ 使得 $f' = \varphi_{f'} \circ \tau'$，其中 $\tau' : V \times \mathbb{F} \to V \otimes_{\mathbb{K}} \mathbb{F}$, $(\mathbf{v}, \lambda) \mapsto \mathbf{v} \otimes_{\mathbb{K}} \lambda$ 是张量映射。即下面的图表交换：

$$\begin{tikzcd} V \times \mathbb{F} \arrow[d, "\tau'"'] \arrow[dr, "f'"] & \\ V \otimes_{\mathbb{K}} \mathbb{F} \arrow[r, "\varphi_{f'}"'] & W \end{tikzcd}$$

在上面的交换图中我们取 $\varphi_f = \varphi_{f'}$ 以及 $\tau(\mathbf{v}) = \tau'(\mathbf{v}, 1_{\mathbb{F}})$ 就得到了交换图 (6.1.6)。然而，这里得到的 φ_f 只是对 $\mathbb{K}$ 上的数乘具有线性性质，我们还需要验证 φ_f 是 $\mathbb{F}$ 线性映射。

事实上，由于加法在 $\mathbb{K}$ 和 $\mathbb{F}$ 下不受影响，因此我们只需验证用 $\mathbb{F}$ 中的元素作数乘与 φ_f

的作用交换即可。任取 $\mathbf{v} \otimes_{\mathbb{K}} \lambda \in V^{\mathbb{F}}$ 及 $\mu \in \mathbb{F}$，我们有

$$\begin{aligned}
\varphi_f(\mu \cdot (\mathbf{v} \otimes_{\mathbb{K}} \lambda)) &= \varphi_f(\mathbf{v} \otimes (\mu\lambda)) && \text{(数乘规则)} \\
&= \varphi_{f'}(\mathbf{v} \otimes \mu\lambda) && (\varphi_f\text{的定义}) \\
&= f'(\mathbf{v}, \mu\lambda) && (\varphi_{f'}\text{的定义}) \\
&= \mu\lambda \cdot f(\mathbf{v}) && (f'\text{的定义}) \\
&= \mu \cdot f'(\mathbf{v}, \lambda) \\
&= \mu \cdot \varphi_{f'}(\mathbf{v} \otimes_{\mathbb{K}} \lambda) \\
&= \mu \cdot \varphi_f(\mathbf{v} \otimes_{\mathbb{K}} \lambda).
\end{aligned}$$

再利用 φ_f 与加法可以交换顺序即证明了 φ_f 是 $\mathbb{F}$ 线性映射。这样我们就完成了证明。 □

推论 6.1.1. 设 $\mathbb{K}$ 是 $\mathbb{F}$ 的子域，V, W 是 $\mathbb{K}$ 上的向量空间。那么，任意一个 $\mathbb{K}$– 线性映射 $f : V \to W$ 都可以唯一地扩张成 $\mathbb{F}$– 线性映射 $\varphi_f : V^{\mathbb{F}} \to W^{\mathbb{F}}$，使得下面的交换图成立：

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "\tau_V"'] & W \arrow[d, "\tau_W"] \\
V^{\mathbb{F}} \arrow[r, "\varphi_f"'] & W^{\mathbb{F}}
\end{tikzcd}$$

其中，$\tau_V : \mathbf{v} \mapsto \mathbf{v} \otimes_{\mathbb{K}} 1_{\mathbb{F}}$，$\tau_W : \mathbf{w} \mapsto \mathbf{w} \otimes_{\mathbb{K}} 1_{\mathbb{F}}$。

只需注意到上面的交换图中红线的部分是一个 $\mathbb{K}$– 线性映射，再利用命题6.1.3即可证明。细节留作练习。

在以上的基域扩张过程中，取 $\mathbb{K} = \mathbb{R}$，$\mathbb{F} = \mathbb{C}$，则很容易验证 $V^{\mathbb{C}}$ 就是 3.6 节中我们定义的实向量空间的复化。

我们已经讨论了向量空间的张量积，那么，一个自然的问题就是考虑线性算子之之间的张量积，我们有下面的定义：

定义 6.1.3. 设 V, W, V', W' 是域 $\mathbb{K}$ 上的向量空间，$\mathcal{A} : V \to V'$ 和 $\mathcal{B} : W \to W'$ 是线性映射，则存在唯一的线性映射 $f : V \otimes W \to V' \otimes W'$ 使得对于任意的 $V \otimes W$ 中的元素 $\mathbf{x} = \sum_{i=1}^{s} \mathbf{v}_i \otimes \mathbf{w}_i$，$\mathbf{v}_i \in V$，$\mathbf{w}_i \in W$，都有 $f(\mathbf{x}) = \sum_{i=1}^{s} (\mathcal{A}\mathbf{v}_i) \otimes (\mathcal{B}\mathbf{w}_i)$。这个存在唯一性是由泛性质保证的，因为下面的交换图成立：

$$\begin{tikzcd}
V \times W \arrow[rr] \arrow[dd] & & V \otimes W \arrow[ddll, "f"] \\
(\mathbf{v}, \mathbf{w}) & & \mathbf{v} \otimes \mathbf{w} \\
V' \otimes W' & & \\
(\mathcal{A}\mathbf{v}) \otimes (\mathcal{B}\mathbf{w}) & &
\end{tikzcd}$$

很容易验证蓝色箭头所定义的映射是一个双线性映射（源于张量的多重线性性质），于是泛性质（定理6.1.1）保证了 f 的存在唯一性。我们把上面的 f 称为 $\mathcal{A}$ 和 $\mathcal{B}$ 的张量积映射，记作 $f = \mathcal{A} \otimes \mathcal{B}$。

同样地，我们可以定义多个线性映射 $\mathcal{A}_k : V_k \to W_k$，$k = 1, \ldots, m$ 的张量积 $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_k$，这里不再赘述。

一个常用的情形是 $V = V'$，$W = W'$，即 $\mathcal{A} \in \mathrm{End}(V)$ 和 $\mathcal{B} \in \mathrm{End}(W)$ 是线性算子的情形。此时，线性算子的张量积有以下的运算律：

定理 6.1.4. 设 V, W 是域 $\mathbb{K}$ 上的向量空间，$\mathcal{A}, \mathcal{C} \in \mathrm{End}(V)$，$\mathcal{B}, \mathcal{D} \in \mathrm{End}(W)$。则

(1) $(\mathcal{A} \otimes \mathcal{B})(\mathcal{C} \otimes \mathcal{D}) = (\mathcal{A}\mathcal{C}) \otimes (\mathcal{B}\mathcal{D})$;

(2) $(\mathcal{A} + \mathcal{C}) \otimes \mathcal{B} = \mathcal{A} \otimes \mathcal{B} + \mathcal{C} \otimes \mathcal{B}$，$\mathcal{A} \otimes (\mathcal{B} + \mathcal{D}) = \mathcal{A} \otimes \mathcal{B} + \mathcal{A} \otimes \mathcal{D}$。

(3) $\forall \lambda \in \mathbb{K}$，$\mathcal{A} \otimes (\lambda\mathcal{B}) = (\lambda\mathcal{A}) \otimes \mathcal{B} = \lambda(\mathcal{A} \otimes \mathcal{B})$;

证明. 只需验证等式的左右两边作用到形如 $\mathbf{v} \otimes \mathbf{w}$，$\mathbf{v} \in V, \mathbf{w} \in W$ 的元素上相等即可，之后由多重线性性质即可得到结论。细节留作练习。 □

这些运算律也可以推广到一般的线性映射的情形上，其形式和证明是类似的，我们不再赘述。

下面我们考虑线性算子张量积的矩阵表达是什么样的。设 V 的一组基是 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，$W$ 的一组基是 $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$，$\mathcal{A} \in \mathrm{End}(V)$ 和 $\mathcal{B} \in \mathrm{End}(W)$ 在各自的基底下对应的矩阵分别是 $A = (a_{ij})_{n\times n}$ 和 $B = (b_{kl})_{m\times m}$。我们在式 (6.1.4) 时已经知道，

$$\mathbf{e}_1 \otimes \mathbf{e}'_1, \ldots, \mathbf{e}_1 \otimes \mathbf{e}'_m, \ldots\ldots, \mathbf{e}_n \otimes \mathbf{e}'_1, \ldots, \mathbf{e}_n \otimes \mathbf{e}'_m$$

是 $V \otimes W$ 的一组基。按线性算子张量积的定义显然我们有

$$\begin{aligned}
(\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_j \otimes \mathbf{e}'_l) &= (\mathcal{A}\mathbf{e}_j) \otimes (\mathcal{B}\mathbf{e}'_l) \\
&= (\sum_{i=1}^{n} a_{ij}\mathbf{e}_i) \otimes (\sum_{k=1}^{m} b_{kl}\mathbf{e}'_k) \\
&= \sum_{i=1}^{n}\sum_{k=1}^{m} a_{ij}b_{kl} \cdot (\mathbf{e}_i \otimes \mathbf{e}'_k), \quad \forall j = 1, \ldots, n,\ l = 1, \ldots, m.
\end{aligned}$$

将上式按照 $(j, l) = (1, 1), \ldots, (1, m), \ldots\ldots, (n, 1), \ldots, (n, m)$ 的顺序排列成矩阵，我们就得到了以下式子：

$$\begin{aligned}
&((\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_1 \otimes \mathbf{e}'_1), \ldots, (\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_1 \otimes \mathbf{e}'_m), \ldots\ldots, (\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_n \otimes \mathbf{e}'_1), \ldots, (\mathcal{A} \otimes \mathcal{B})(\mathbf{e}_n \otimes \mathbf{e}'_m)) \\
&= (\mathbf{e}_1 \otimes \mathbf{e}'_1, \ldots, \mathbf{e}_1 \otimes \mathbf{e}'_m, \ldots\ldots, \mathbf{e}_n \otimes \mathbf{e}'_1, \ldots, \mathbf{e}_n \otimes \mathbf{e}'_m) \cdot \begin{pmatrix}
a_{11}b_{11} & \cdots & a_{11}b_{1m} & & & a_{1n}b_{11} & \cdots & a_{1n}b_{1m} \\
\vdots & & \vdots & \cdots & \cdots & \vdots & & \vdots \\
a_{11}b_{m1} & \cdots & a_{11}b_{mm} & & & a_{1n}b_{m1} & \cdots & a_{1n}b_{mm} \\
 & \vdots & & \ddots & & & \vdots & \\
 & \vdots & & & \ddots & & \vdots & \\
a_{n1}b_{11} & \cdots & a_{n1}b_{1m} & & & a_{nn}b_{11} & \cdots & a_{nn}b_{1m} \\
\vdots & & \vdots & \cdots & \cdots & \vdots & & \vdots \\
a_{n1}b_{m1} & \cdots & a_{n1}b_{mm} & & & a_{nn}b_{m1} & \cdots & a_{nn}b_{mm}
\end{pmatrix}
\end{aligned} \tag{6.1.7}$$

读者一定要亲手验证上式！把上式的矩阵写成分块矩阵的形式，就是：

$$\begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{pmatrix} \tag{$\star\star$}$$

定义 6.1.4. 上面的 $nm \times nm$ 规模的矩阵 $(\star\star)$ 定义成矩阵 A 和 B 的 Kronecker **乘积** (也称为矩阵 A 和 B 的直积或张量积)，记作 $A \otimes B$。我们也可以把 Kronecker 乘积定义到非方阵上，形式与 $(\star\star)$ 相同，可以视作一般的线性映射的张量积对应的矩阵。

有了这个定义，我们很容易看出例6.1.2(2) 中的定义就是 Kronecker 乘积的特殊情形。矩阵的 Kronecker 乘积满足以下性质：

定理 6.1.5. 设 $\mathbb{K}$ 是域，$A, C \in M_n(\mathbb{K})$，$B, D \in M_m(\mathbb{K})$，$F \in M_l(\mathbb{K})$，则

(1) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$;

(2) $(A + C) \otimes B = A \otimes B + C \otimes B,\ A \otimes (B + D) = A \otimes B + A \otimes D$;

(3) $\forall \lambda \in \mathbb{K},\ A \otimes (\lambda B) = (\lambda A) \otimes B = \lambda(A \otimes B)$;

(4) $(A \otimes B) \otimes F = A \otimes (B \otimes F)$，因此可以记作 $A \otimes B \otimes F$;

(5) $(A \otimes B)^t = A^t \otimes B^t$;

(6) 如果 A, B 都可逆，则 $A \otimes B$ 也可逆且 $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$;

(7) 非零矩阵 A, C 同阶, B, D 同阶, 则 $A \otimes B = C \otimes D \iff \exists a \in \mathbb{K} \backslash \{0\}$ 使得 $A = aC,\ B = a^{-1}D$;

(8) $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A) \cdot \operatorname{tr}(B)$;

(9) $\det(A \otimes B) = (\det(A))^m \cdot (\det(B))^n$。

证明. (1),(2),(3) 就是定理6.1.4的矩阵形式。(4),(5) 只需验证矩阵对应位置的元素相等即可，细节留作练习。

对于 (6)，注意到由 (1) 可得 $(A^{-1} \otimes B^{-1}) \cdot (A \otimes B) = (A^{-1}A) \otimes (B^{-1}B) = E_n \otimes E_m = E_{nm}$，同理 $(A \otimes B) \otimes (A^{-1} \otimes B^{-1}) = E_{nm}$，因此 (6) 成立。

对于 (7)，设 $A = (a_{ij})_{n \times n}$，$C = (c_{ij})_{n \times n}$，则

$$A \otimes B = C \otimes D \iff a_{ij}B = c_{ij}D,\ \forall i, j = 1, \ldots, n \iff \exists a \in \mathbb{K} \setminus \{0\} \text{使得} a_{ij} = ac_{ij},\ B = a^{-1}D$$

即得结论。

对于 (8)，设 $A = (a_{ij})_{n \times n}$，由 Kronecker 乘积的定义式 $(\star\star)$ 可知

$$\operatorname{tr}(A \otimes B) = \sum_{i=1}^{n} (a_{ii}\operatorname{tr}(B)) = \operatorname{tr}(A) \cdot \operatorname{tr}(B).$$

对于 (9)，设 $A = (a_{ij})_{n \times n}$，首先我们注意到

$$A \otimes E_m = \begin{pmatrix} a_{11}E_m & \cdots & a_{1n}E_m \\ \vdots & & \vdots \\ a_{n1}E_m & \cdots & a_{nn}E_m \end{pmatrix},$$

于是通过交换矩阵的行和列，我们可知

$$\det(A\otimes E_m) = \det\begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix} = (\det(A))^m;$$

另一方面，我们显然有

$$\det(E_n\otimes B) = \det\begin{pmatrix} B & & \\ & \ddots & \\ & & B \end{pmatrix} = (\det(B))^n.$$

因此 $\det(A\otimes B) = \det((A\otimes E_m)\cdot(B\otimes E_n)) = \det(A\otimes E_m)\cdot\det(B\otimes E_n) = (\det(A))^m\cdot(\det(B))^n$。这样我们就完成了证明。 □

有了 Kronecker 乘积之后我们可以推导张量积空间的坐标变换了。为了推导的简便，我们只写出两个空间的张量积的情形，多个空间的张量积是类似的。

设 V_1, V_2 分别是域 $\mathbb{K}$ 上的 n 维和 m 维向量空间，V_1 有两组基分别是 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$，$V_2$ 有两组基分别是 $\mathbf{w}_1,\ldots,\mathbf{w}_m$ 和 $\mathbf{w}'_1,\ldots,\mathbf{w}'_m$，并且

$$(\mathbf{e}'_1,\ldots,\mathbf{e}'_n) = (\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A,\quad (\mathbf{w}'_1,\ldots,\mathbf{w}'_m) = (\mathbf{w}_1,\ldots,\mathbf{w}_m)\cdot B,\quad A\in \mathrm{GL}_n(\mathbb{K}),\ B\in\mathrm{GL}_m(\mathbb{K}).$$

那么，$V_1\otimes V_2$ 有两组基 $\mathbf{e}_1\otimes\mathbf{w}_1,\ldots,\mathbf{e}_1\otimes\mathbf{w}_m,\ldots\ldots,\mathbf{e}_n\otimes\mathbf{w}_1,\ldots,\mathbf{e}_n\otimes\mathbf{w}_m$ 和 $\mathbf{e}'_1\otimes\mathbf{w}'_1,\ldots,\mathbf{e}'_1\otimes\mathbf{w}'_m,\ldots\ldots,\mathbf{e}'_n\otimes\mathbf{w}'_1,\ldots,\mathbf{e}'_n\otimes\mathbf{w}'_m$。类似于 (6.1.7) 式的推导我们就可以得到：

$$\begin{aligned}&(\mathbf{e}'_1\otimes\mathbf{w}'_1,\ldots,\mathbf{e}'_1\otimes\mathbf{w}'_m,\ldots\ldots,\mathbf{e}'_n\otimes\mathbf{w}'_1,\ldots,\mathbf{e}'_n\otimes\mathbf{w}'_m)\\ =\ &(\mathbf{e}_1\otimes\mathbf{w}_1,\ldots,\mathbf{e}_1\otimes\mathbf{w}_m,\ldots\ldots,\mathbf{e}_n\otimes\mathbf{w}_1,\ldots,\mathbf{e}_n\otimes\mathbf{w}_m)\cdot(A\otimes B).\end{aligned}$$

于是，如果 $\mathbf{T} = \sum_{i=1}^n\sum_{j=1}^m T_{ij}\mathbf{e}_i\otimes\mathbf{w}_j = \sum_{i=1}^n\sum_{j=1}^m T'_{ij}\mathbf{e}'_i\otimes\mathbf{w}'_j$，那么由坐标变换公式（式 (1.3.1)）就有

$$(T'_{11},\ldots,T'_{1m},\ldots\ldots,T'_{n1},\ldots,T'_{mn}) = (T_{11},\ldots,T_{1m},\ldots\ldots,T_{n1},\ldots,T_{mn})\cdot\left((A^{-1})^t\otimes(B^{-1})^t\right)$$

这就是张量积空间的坐标变换。更一般的情形是类似的，留作练习。

下面我们来定义张量的秩。我们很快会看到，矩阵作为一种特殊的张量，其秩的定义与一般的张量秩的定义是一致的。

定义 6.1.5. 设 $V_1,\ldots,V_m$ 是域 $\mathbb{K}$ 上的向量空间，则按照张量的定义，$V_1\otimes\cdots\otimes V_m$ 中的任意一个元素 T 都可以写成下面的表达式：

$$T = \sum_{i=1}^r\left(\mathbf{v}_i^{(1)}\otimes\mathbf{v}_i^{(2)}\otimes\cdots\otimes\mathbf{v}_i^{(m)}\right),\qquad 其中\mathbf{v}_i^{(j)}\in V_j,\ j\in\{1,\ldots,m\}.\tag{6.1.8}$$

这种表达式可能不是唯一的，于是我们把 T 的所有这样的表达式收集起来，将其中最小的求和项数 r 称为张量 T 的 CP-秩 (canonical polyadic rank)，把求和项数最少的 T 的表达式称为 T 的张量秩分解 (tensor rank decomposition 或 canonical polyadic decomposition)。

特别地，如果一个张量的 CP-秩是 1，那么我们称这个张量是可分解 (decomposable) 张量或纯态 (pure state)。

如果 $m = 2$，张量秩分解实际上就是求矩阵的秩标准型的过程。事实上，仿照例6.1.2(1) 的过程，对于 2 阶张量

$$\mathbf{T} = \sum_{i=1}^{s}\sum_{j=1}^{t} a_{ij}\mathbf{e}_i \otimes \mathbf{e}'_j,$$

我们可以把它同构到一个 $s \times t$ 阶的矩阵，即令 $\mathbf{e}_i \otimes \mathbf{e}'_j \mapsto E_{ij}$，则 $\mathbf{T} \mapsto \begin{pmatrix} a_{11} & \cdots & a_{1t} \\ \vdots & & \vdots \\ a_{s1} & \cdots & a_{st} \end{pmatrix}$，我们把这个矩阵记作 A，设 A 的秩是 r，则 A 有秩标准型 $A = P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q$，其中 $P \in \mathrm{GL}_s(\mathbb{K})$, $Q \in \mathrm{GL}_t(\mathbb{K})$。令 $A_i = PE_{ii}Q$, $i \in \{1, \dots, r\}$，则每个 A_i 都是秩为 1 的矩阵且 $A = \sum_{k=1}^{r} A_i$。之后我们将每个 A_i 还原回张量的形式，则它们都是可分解张量 (注意到 $E_{ii} = \mathbf{e}_i \cdot \mathbf{e}'_i$ 即可)，这样我们就得到了张量 $\mathbf{T}$ 的张量秩分解 (最小性是显然的，因为如果有 CP-秩更小的张量秩分解，重复以上过程我们会得到矩阵 A 的秩小于 r，矛盾！)。

但对于 $m \geqslant 3$ 的情形，张量秩分解一般是十分困难的 [1]。此时张量的 CP-秩甚至与基域的选择有关，我们来看下面的例子。

例 6.1.4. [2] 我们考虑如下的张量

$$\mathbf{T} = -\mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1.$$

它可以视作 $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ 中的张量，其 CP-秩为 3，其分解为

$$\mathbf{T} = \frac{1}{2}\left((\mathbf{e}_1 + \mathbf{e}_2) \otimes (\mathbf{e}_1 + \mathbf{e}_2) \otimes (\mathbf{e}_1 + \mathbf{e}_2)\right) + \frac{1}{2}\left((\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}_1 - \mathbf{e}_2)\right) - 2\mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1,$$

但如果将它视作 $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ 中的张量，则它的 CP-秩为 2，且其分解为

$$\mathbf{T} = \frac{i}{2}\left((-i\mathbf{e}_1 + \mathbf{e}_2) \otimes (-i\mathbf{e}_1 + \mathbf{e}_2) \otimes (-i\mathbf{e}_1 + \mathbf{e}_2)\right) - \frac{i}{2}\left((i\mathbf{e}_1 + \mathbf{e}_2) \otimes (i\mathbf{e}_1 + \mathbf{e}_2) \otimes (i\mathbf{e}_1 + \mathbf{e}_2)\right).$$

有关张量在量子计算中的应用，我们会在附录中简单介绍。

---

[1] 一般情况下这是一个 NP 难的问题，参考：Christopher J. Hillar and Lek-Heng Lim. 2013. Most Tensor Problems Are NP-Hard. J. ACM 60, 6, Article 45 (November 2013), 39 pages. https://doi.org/10.1145/2512329.

[2] 这个例子的详细的证明过程可以参考 Tensor Rank and the Ill-Posedness of the Best Low-Rank Approximation Problem Vin de Silva and Lek-Heng Lim SIAM Journal on Matrix Analysis and Applications 2008 30:3, 1084-1127.

6.2 混合张量

上一节中我们已经看到，我们可以将双线性映射转换到张量积空间上的线性映射。那么，我们如何搭建线性算子与张量的联系呢？这就需要我们将对偶空间引入到张量积中。

定义 6.2.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，V^* 是 V 的对偶空间，则我们把张量积空间

$$\underbrace{V \otimes \cdots \otimes V}_{p\text{个}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{个}}$$

记为 $\mathbb{T}_q^p(V)$，把其中的元素称为 V 上的 (q,p) 型混合张量 (mixed tensor)，其中 p 称为张量的反变阶数 (或逆变阶数，contravariant order)，q 称为共变阶数 (或协变阶数，covariant order)。特别地，如果 $p=0$，则我们称这个张量是一个 q 阶共变张量；如果 $q=0$，则我们称这个张量是一个 p 阶反变张量。共变与反变这两个名称来源于基变换时张量坐标的的变化。[1]

定理 6.2.1. 设 $V_1,\ldots,V_n,W$ 都是域 $\mathbb{K}$ 上的向量空间，记 $\mathrm{mul}(V_1,\ldots,V_n;W)=\{f:V_1\times\cdots\times V_n\to W \mid f\text{是多重线性映射}\}$，则 $\mathrm{mul}(V_1,\ldots,V_n;W)$ 在逐点定义的加法和数乘下构成了一个 $\mathbb{K}$ 上的向量空间 (这与所有双线性型构成的向量空间 $L_2(V,\mathbb{K})$ 类似，见定理1.6.1上方)。我们有

$$\Phi:\mathrm{mul}(V_1,\ldots,V_n;W)\to \mathrm{Hom}(V_1\otimes\cdots\otimes V_n,W),\ f\mapsto \varphi_f$$

是向量空间之间的同构，其中 φ_f 是由上一节的泛性质 $(\star)$ 唯一确定的 (定理6.1.1)。

$$\begin{tikzcd}
V_1\times\cdots\times V_n \arrow[r, "\tau"] \arrow[d, "f"'] & V_1\otimes\cdots\otimes V_n \arrow[dl, "\varphi_f"] \\
W &
\end{tikzcd}$$

证明. 首先，Φ 的良定义性已经由泛性质保证。

其次，Φ 是线性映射。事实上，任取 $f,g\in \mathrm{mul}(V_1,\ldots,V_n;W)$ 及 $\lambda,\mu\in\mathbb{K}$，我们只需验证 $\Phi(\lambda f+\mu g)=\lambda\Phi(f)+\mu\Phi(g)$ 即可。为此，我们只需要验证 $\Phi(\lambda f+\mu g)$ 和 $\lambda\Phi(f)+\mu\Phi(g)$ 作用到任意的可分解张量 $\mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_n$ 上都相等，即说明它们作用到任意 CP-秩的张量上都相等。利用泛性质以及 $\mathrm{mul}(V_1,\ldots,V_n;W)$ 中加法和数乘的定义，我们有

$$\begin{aligned}
\Phi(\lambda f+\mu g)(\mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_n) &= (\lambda f+\mu g)(\mathbf{v}_1,\ldots,\mathbf{v}_n)\\
&= \lambda f(\mathbf{v}_1,\ldots,\mathbf{v}_n)+\mu g(\mathbf{v}_1,\ldots,\mathbf{v}_n)\\
&= (\lambda\Phi(f)+\mu\Phi(g))\,(\mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_n).
\end{aligned}$$

此即 Φ 是线性映射。

再次，Φ 是单射，这只需证 $\ker(\Phi)=\{0\}$。事实上，如果 $\Phi(f)$ 是零映射，那么 $\Phi(f)$ 作用到所有可分解张量上都等于 $\mathbf{0}$，由泛性质即得 f 作用到任意的变元组上都得到 $\mathbf{0}$，即 $f=0$。此即 Φ 是单射。

最后，Φ 是满射，因为任取一个 $\mathrm{Hom}(V_1\otimes\cdots\otimes V_n,W)$ 中的线性映射，我们只要考察它在可分解张量上的作用结果，就可以得到一个 $\mathrm{mul}(V_1,\ldots,V_n;W)$ 中的多重线性映射。这样我们就完成了证明。 □

---

[1] 感兴趣的读者可以参考https://en.wikipedia.org/wiki/Covariance_and_contravariance_of_vectors.

我们把上面的 Φ 称为调解态射 (mediating morphism map)。容易看出，这个定理实际上就是定理1.6.1的直接推广。特别地，如果我们取 $W=\mathbb{K}$ 就得到了下面的推论：

推论 6.2.1. 设 $V_1,\dots,V_n$ 都是域 $\mathbb{K}$ 上的向量空间，则

$$\mathrm{mul}(V_1,\dots,V_n;\mathbb{K})\simeq(V_1\otimes\cdots\otimes V_n)^*.$$

实际上，$\mathrm{mul}(V_1,\dots,V_n;W)$ 就与混合张量有关，这就是下面的定理。

定理 6.2.2. 设 $V_1,\dots,V_n,W$ 都是域 $\mathbb{K}$ 上的**有限维**向量空间，则

$$V_1^*\otimes\cdots\otimes V_n^*\otimes W\simeq\mathrm{mul}(V_1,\dots,V_n;W).$$

证明. 首先我们很容易验证下面的映射 τ 是多重线性映射：

$$\begin{aligned}\tau:\ V_1^*\times\cdots\times V_n^*\times W &\longrightarrow \mathrm{mul}(V_1,\dots,V_n;W)\\ (f_1,\dots,f_n,\mathbf{w}) &\longmapsto \begin{array}{rll}\tau_{(f_1,\dots,f_n,\mathbf{w})}: & V_1\times\cdots\times V_n & \to W\\ & (\mathbf{v}_1,\dots,\mathbf{v}_n) & \mapsto f_1(\mathbf{v}_1)\cdots f_n(\mathbf{v}_n)\cdot\mathbf{w}\end{array}\end{aligned}$$

因此由泛性质 (定理6.1.1)，τ 唯一地诱导了张量积空间 $V_1^*\otimes\cdots\otimes V_n^*\otimes W$ 到 $\mathrm{mul}(V_1,\dots,V_n;W)$ 的线性映射 Φ_τ：

$$\begin{aligned}\Phi_\tau:\ V_1^*\otimes\cdots\otimes V_n^*\times W &\longrightarrow \mathrm{mul}(V_1,\dots,V_n;W)\\ \sum_{i=1}^{s}f_1^{(i)}\otimes\cdots\otimes f_n^{(i)}\otimes\mathbf{w}^{(i)} &\longmapsto \begin{array}{rll}\tau_{(\sum_{i=1}^{s}f_1^{(i)}\otimes\cdots\otimes f_n^{(i)}\otimes\mathbf{w}^{(i)})}: & V_1\times\cdots\times V_n & \to W\\ & (\mathbf{v}_1,\dots,\mathbf{v}_n) & \mapsto \sum_{i=1}^{s}f_1^{(i)}(\mathbf{v}_1)\cdots f_n^{(i)}(\mathbf{v}_n)\cdot\mathbf{w}^{(i)}\end{array}\end{aligned}$$

即下面的交换图成立：

$$\begin{tikzcd}
V_1^*\times\cdots\times V_n^*\times W \arrow[r] \arrow[d, "\tau"'] & V_1^*\otimes\cdots\otimes V_n^*\otimes W \arrow[dl, "\Phi_\tau"] \\
\mathrm{mul}(V_1,\dots,V_n;W) &
\end{tikzcd}$$

Φ_τ 的良定义性和线性性已经由泛性质保证。下面我们来说明 Φ_τ 是双射。

- Φ_τ 是满射：对每个 $i\in\{1,\dots,n\}$，不妨设 V_i 的一组基是 $\mathbf{e}_1^{(i)},\dots,\mathbf{e}_{s_i}^{(i)}$，以及 W 的一组基是 $\mathbf{w}_1,\dots,\mathbf{w}_t$，那么，利用多重线性性质我们很容易证明，每个多重线性映射 $f\in\mathrm{mul}(V_1,\dots,V_n;W)$ 都由 f 作用在基底上得到的像所决定 (类似于线性映射基本定理)，于是，我们设

  $$f(\mathbf{e}_{k_1}^{(1)},\dots,\mathbf{e}_{k_n}^{(n)})=\sum_{j=1}^{t}\lambda_j^{(k_1,\dots,k_n)}\mathbf{w}_j,\quad 其中\lambda_j^{(k_1,\dots,k_n)}\in\mathbb{K},\ k_l\in\{1,\dots,s_k\},\ l=1,\dots,n.$$

  如果我们把 $\mathbf{e}_{k_l}^{(m)}$ 的对偶基记作 $\mathbf{e}_{k_l}^{(m)*}$，那么按照 Φ_τ 的定义我们立刻可以验证：

  $$\Phi_\tau\left(\sum_{j=1}^{t}\sum_{k_1=1}^{s_1}\cdots\sum_{k_n=1}^{s_n}\lambda_j^{(k_1,\dots,k_n)}\mathbf{e}_{k_1}^{(1)*}\otimes\cdots\otimes\mathbf{e}_{k_n}^{(n)*}\otimes\mathbf{w}_j\right)=f.$$

  此即 Φ_τ 是满射。

- $\varPhi_\tau$ 是单射：基底的设法同上，则如果 $\varPhi_\tau\left(\sum_{j=1}^{t}\sum_{k_1=1}^{s_1}\cdots\sum_{k_n=1}^{s_n}\lambda_j^{(k_1,\dots,k_n)}\mathbf{e}_{k_1}^{(1)*}\otimes\cdots\otimes\mathbf{e}_{k_n}^{(n)*}\otimes\mathbf{w}_j\right)$ 是零映射，那么只能是每个 $\lambda_j^{(k_1,\dots,k_n)}$ 都是 0，即括号里的张量是零，因此 $\ker(\varPhi_\tau)=\{\mathbf{0}\}$，$\varPhi_\tau$ 是单射。

这样我们就完成了证明。 □

在定理6.2.2中分别取 $W=\mathbb{K}$ 或者 $n=1$，我们就得到以下两个推论：

推论 6.2.2. 设 $V_1,\dots,V_n$ 都是域 $\mathbb{K}$ 上的有限维向量空间，则

$$\mathrm{mul}(V_1,\dots,V_n;\mathbb{K})\simeq(V_1\otimes\cdots\otimes V_n)^*\simeq V_1^*\otimes\cdots\otimes V_n^*.$$

推论 6.2.3. 设 V,W 是域 $\mathbb{K}$ 上的有限维向量空间，则 $\mathrm{Hom}(V,W)\simeq V^*\otimes W$。特别地，$\mathrm{End}(V)\simeq V^*\otimes V$，即线性算子是一个 $(1,1)$ 型的混合张量。

结合推论6.2.1和推论6.2.2，我们有

推论 6.2.4. 设 $V_1,\dots,V_n$ 都是域 $\mathbb{K}$ 上的有限维向量空间，则

$$V_1\otimes\cdots\otimes V_n\simeq\mathrm{mul}(V_1^*,\dots,V_n^*;\mathbb{K}).$$

证明. 由推论6.2.1，$\mathrm{mul}(V_1^*,\dots,V_n^*;\mathbb{K})\simeq(V_1^*\otimes\cdots\otimes V_n^*)^*$；由推论6.2.2，$(V_1^*\otimes\cdots\otimes V_n^*)^*\simeq V_1^{**}\otimes\cdots\otimes V_n^{**}$；再由定理1.5.2，$V_1^{**}\otimes\cdots\otimes V_n^{**}\simeq V_1\otimes\cdots\otimes V_n$。 □

注意到推论6.2.4证明中的每一个同构映射都是自然的，即不依赖于基底的选取，因此我们也可以把推论6.2.4作为张量积空间的另一种定义[1]。

推论6.2.4还告诉我们，混合张量空间 $\mathbb{T}_q^p(V)$ 实际上同构于

$$\mathrm{mul}(\underbrace{V^*,\dots,V^*}_{p\text{个}},\underbrace{V,\dots,V}_{q\text{个}};\mathbb{K}),$$

也就是说，张量本质上就是多重线性映射。这就是《代数学引论》上的定义。特别地，$\mathbb{T}_0^0(V)$ 是域 $\mathbb{K}$ 中的数，$\mathbb{T}_1^0(V)$ 就是对偶空间 V^*V，$\mathbb{T}_0^1(V)$ 是 $V^{**}\simeq V$，$\mathbb{T}_1^1(V)$ 是线性算子空间 $\mathrm{End}(V)$。

接下来我们考虑混合张量的坐标变换。设 V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathbf{e}_1,\dots,\mathbf{e}_n$ 和 $\mathbf{e}_1',\dots,\mathbf{e}_n'$ 分别是 V 的两组基，且

$$(\mathbf{e}_1',\dots,\mathbf{e}_n')=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot A,\quad A\in\mathrm{GL}_n(\mathbb{K}).$$

为了形式的美观，我们设 $A=(a_j^i)_{n\times n}$，其中上指标 i 表示第 i 行，下指标 j 表示第 j 列（注意这里上指标不是幂次！）。也就是说：

$$(\mathbf{e}_1',\dots,\mathbf{e}_n')=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot\begin{pmatrix}a_1^1&\cdots&a_n^1\\ \vdots&&\vdots\\ a_n^1&\cdots&a_n^n\end{pmatrix}\quad\text{或者说}\quad\mathbf{e}_k'=\sum_{i=1}^{n}a_k^i\mathbf{e}_i,\ \forall k=1,\dots,n.$$

---

[1] 参考《微分几何讲义》，陈省身，陈维桓，北京大学出版社的第二章，或者《高等线性代数学》，黎景辉，白正简，周国晖，高等教育出版社的第 2.4 节。

记 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 和 $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ 的对偶基分别为 $\mathbf{e}^1, \ldots, \mathbf{e}^n$ 和 $\mathbf{e}^{1\prime}, \ldots, \mathbf{e}^{n\prime}$，那么，

$$\mathbf{e}^j(\mathbf{e}'_k) = \sum_{i=1}^{n} a_k^i \delta_{ij} = a_k^j.$$

也就是说，

$$\begin{pmatrix} \mathbf{e}^1(\mathbf{e}'_1) & \cdots & \mathbf{e}^1(\mathbf{e}'_n) \\ \vdots & & \vdots \\ \mathbf{e}^n(\mathbf{e}'_1) & \cdots & \mathbf{e}^n(\mathbf{e}'_n) \end{pmatrix} = \begin{pmatrix} a_1^1 & \cdots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \cdots & a_n^n \end{pmatrix} = A. \tag{6.2.1}$$

我们不妨设 $(\mathbf{e}^{1\prime}, \ldots, \mathbf{e}^{n\prime}) = (\mathbf{e}^1, \ldots, \mathbf{e}^n) \cdot B$，那么，

$$\begin{pmatrix} \mathbf{e}^1 \\ \vdots \\ \mathbf{e}^n \end{pmatrix} = (B^t)^{-1} \cdot \begin{pmatrix} \mathbf{e}^{1\prime} \\ \vdots \\ \mathbf{e}^{n\prime} \end{pmatrix},$$

因此应该有

$$\begin{pmatrix} \mathbf{e}^1(\mathbf{e}'_1) & \cdots & \mathbf{e}^1(\mathbf{e}'_n) \\ \vdots & & \vdots \\ \mathbf{e}^n(\mathbf{e}'_1) & \cdots & \mathbf{e}^n(\mathbf{e}'_n) \end{pmatrix} = (B^t)^{-1} \cdot \begin{pmatrix} \mathbf{e}^{1\prime}(\mathbf{e}'_1) & \cdots & \mathbf{e}^{1\prime}(\mathbf{e}'_n) \\ \vdots & & \vdots \\ \mathbf{e}^{n\prime}(\mathbf{e}'_1) & \cdots & \mathbf{e}^{n\prime}(\mathbf{e}'_n) \end{pmatrix} = (B^t)^{-1} \cdot E_n = (B^t)^{-1}. \tag{6.2.2}$$

对比式 (6.2.1) 和式 (6.2.2) 可知 $B = (A^{-1})^t$，即

$$(\mathbf{e}^{1\prime}, \ldots, \mathbf{e}^{n\prime}) = (\mathbf{e}^1, \ldots, \mathbf{e}^n) \cdot (A^{-1})^t, \tag{6.2.3}$$

此即基变换下对偶基的变换公式 (我们也顺手解决了注1.5.1留下的练习)。记 $(A^{-1})^t = (b_j^i)_{n\times n}$，其中 $b_j^i \in \mathbb{K}$ 仍表示矩阵第 i 行第 j 列的元素。有了这个变换公式以后，设

$$\mathbf{T} = \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} \left(T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p} \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \mathbf{e}^{j_1} \otimes \cdots \otimes \mathbf{e}^{j_q} \right) \in \mathbb{T}_q^p(V), \tag{6.2.4}$$

如果改用基底 $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ 及其对偶基后

$$\mathbf{T} = \sum_{i'_1=1}^{n} \cdots \sum_{i'_p=1}^{n} \sum_{j'_1=1}^{n} \cdots \sum_{j'_q=1}^{n} \left(T'^{i'_1,\ldots,i'_p}_{j'_1,\ldots,j'_q} \mathbf{e}'_{i_1} \otimes \cdots \otimes \mathbf{e}'_{i_p} \otimes \mathbf{e}^{j_1\prime} \otimes \cdots \otimes \mathbf{e}^{j_q\prime} \right),$$

那么我们有

$$\begin{aligned} \mathbf{T} &= \sum_{i'_1=1}^{n} \cdots \sum_{i'_p=1}^{n} \sum_{j'_1=1}^{n} \cdots \sum_{j'_q=1}^{n} \left(T'^{i'_1,\ldots,i'_p}_{j'_1,\ldots,j'_q} \mathbf{e}'_{i_1} \otimes \cdots \otimes \mathbf{e}'_{i_p} \otimes \mathbf{e}^{j_1\prime} \otimes \cdots \otimes \mathbf{e}^{j_q\prime} \right) \\ &= \sum_{i'_1=1}^{n} \cdots \sum_{i'_p=1}^{n} \sum_{j'_1=1}^{n} \cdots \sum_{j'_q=1}^{n} \left(T'^{i'_1,\ldots,i'_p}_{j'_1,\ldots,j'_q} (\sum_{k=1}^{n} a_{i'_1}^k \mathbf{e}_k) \otimes \cdots \otimes (\sum_{k=1}^{n} a_{i'_p}^k \mathbf{e}_k) \otimes (\sum_{l=1}^{n} b_{j'_1}^l \mathbf{e}^l) \otimes \cdots \otimes (\sum_{l=1}^{n} b_{j'_q}^l \mathbf{e}^l) \right) \end{aligned} \tag{6.2.5}$$

展开式 (6.2.5) 并对比式 (6.2.4) 的坐标系数可知：

$$T^{i_1,\dots,i_p}_{j_1,\dots,j_q} = \sum_{i'_1,\dots i'_p} \sum_{j'_1,\dots j'_q} a^{i_1}_{i'_1} \cdot a^{i_2}_{i'_2} \cdots a^{i_p}_{i'_p} \cdot T'^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q} \cdot b^{j'_1}_{j_1} \cdot b^{j'_2}_{j_2} \cdots b^{i'_p}_{j_q}. \tag{6.2.6}$$

这就是混合张量的坐标变换公式。这个公式用通常的求和记号来书写是十分复杂的 (再次强调，读者一定要亲手验证 (6.2.6) 式！)，而在微分几何或者物理中，更常用的是爱因斯坦求和约定 (Einstein summation convention)，我们在这里不打算介绍这套记号，感兴趣的读者可以参考《微分几何讲义》，陈省身，陈维桓，北京大学出版社的第二章或者《微分几何入门与广义相对论》上册，梁灿彬，周彬，科学出版社的第二章。

矩阵乘法的 Strassen 算法

在本节的最后，我们来解决一个讲义上册 1.2.5 小节的历史遗留问题：如何降低高斯消元法的乘法复杂度。我们知道，高斯消元法实际上就是不断地左乘一些初等矩阵，因此这个问题和降低矩阵乘法的乘法复杂度是一回事。注意到矩阵乘法是一个双线性映射，因此我们可以用张量来处理矩阵乘法。

考虑 $m \times n$ 的复矩阵与 $n \times p$ 的复矩阵相乘，我们将会得到一个 $m \times p$ 的复矩阵，按照矩阵乘法的定义，我们有：

$$\cdot: \ \mathbb{C}_{m\times n} \times \mathbb{C}_{n\times p} \longrightarrow \mathbb{C}_{m\times p}, \quad (E_{ij}, E_{jk}) \longmapsto E_{ik}, \quad i \in \{1,\dots,m\},\ j \in \{1,\dots,n\},\ k \in \{1,\dots,p\}.$$

这是一个双线性映射，由定理6.2.2，我们知道矩阵乘法这个映射在同构 Φ_τ 下对应于 $(\mathbb{C}_{m\times n})^* \otimes (\mathbb{C}_{n\times p})^* \otimes \mathbb{C}_{m\times p}$ 中的一个张量 $\mathbf{T}$，并且如果我们把 E_{ij} 的对偶基记作 E^*_{ij}，则

$$\mathbf{T} = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} E^*_{ij} \otimes E^*_{jk} \otimes E_{ik}.$$

这个张量是 $m \cdot n \cdot p$ 个可分解张量的和，那么它的 CP-秩是 $m \cdot n \cdot p$ 吗？一个反直觉的答案：不是！以 $m = n = p = 2$ 的情形为例，此时 $\mathbf{T}$ 的 CP-秩是 7 而不是 8！这个结果是 Strassen 在 1969 年证明的，由此开启了矩阵乘法的复杂度理论研究。我们在这里给出 $\mathbf{T}$ 的张量秩分解如下：

$$\begin{aligned}
\mathbf{T} = {} & (E^*_{11} + E^*_{22}) \otimes (E^*_{11} + E^*_{22}) \otimes (E_{11} + E_{22}) \\
& + (E^*_{21} + E^*_{22}) \otimes E^*_{11} \otimes (E_{21} - E_{22}) \\
& + E^*_{11} \otimes (E^*_{12} - E^*_{22}) \otimes (E_{12} + E_{22}) \\
& + E^*_{22} \otimes (E^*_{21} - E^*_{11}) \otimes (E_{11} + E_{21}) \\
& + (E^*_{11} + E^*_{12}) \otimes E^*_{22} \otimes (E_{12} - E_{11}) \\
& + (E^*_{21} - E^*_{11}) \otimes (E^*_{11} + E^*_{12}) \otimes E_{22} \\
& + (E^*_{12} - E^*_{22}) \otimes (E^*_{21} + E^*_{22}) \otimes E_{11}.
\end{aligned}$$

读者可以自行验证 $\mathbf{T}$ 的这两种表达式相等。这个分解的求和项数确实是最少的，证明可以参考《Algebraic Complexity Theory With the Collaboration of Thomas Lickteig》，Peter Bürgisser, Michael Clausen, Mohammad Amin Shokrollahi，Springer-Verlag Berlin Heidelberg GmbH 的 Corollary 17.10 至 Theorem 17.12。

将以上张量秩分解还原成双线性映射的分解，我们就得到了 2×2 矩阵乘法的算法如下：

(Strassen'69) 输入：$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$，$B=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\in M_2(\mathbb{C})$.

- $m_1=(a_{11}+a_{22})(b_{11}+b_{22})$;
- $m_2=(a_{21}+a_{22})b_{11}$;
- $m_3=a_{11}(b_{12}-b_{22})$;
- $m_4=a_{22}(b_{21}-b_{11})$;
- $m_5=(a_{11}+a_{12})b_{22}$;
- $m_6=(a_{21}-a_{11})(b_{11}+b_{12})$;
- $m_7=(a_{12}-a_{22})(b_{21}+b_{22})$;
- $c_{11}=m_1+m_4-m_5+m_7$;
- $c_{12}=m_3+m_5$;
- $c_{21}=m_2+m_4$;
- $c_{22}=m_1-m_2+m_3+m_6$.

输出：$A\cdot B=\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$.

从上表中我们看到，这个算法只使用了 7 次乘法 (即计算 m_1 到 m_7) 就完成了 2×2 矩阵乘法的计算，而按照定义计算则需要计算 8 次乘法。对于规模更大的矩阵，我们只需要将其补充一些全为 0 的行和列，使其成为 $2^n\times 2^n$ 的矩阵，再对其 2×2 分块使用 Strassen 算法，即可降低矩阵乘法中使用乘法的次数。按照 Strassen 算法，$n\times n$ 矩阵乘法的乘法复杂度是 $n^{\log_2 7}\approx n^{2.81}$。这个结果可以进一步改进，最新的结果是 $n^{2.37286}$ 左右 (2020 年)。一个猜想是 $n\times n$ 矩阵乘法的乘法复杂度低于 $n^{2+\varepsilon}$，$\forall\varepsilon>0$，不过目前还不确定猜想是否正确。还有一些专家试图通过深度学习等人工智能方法来寻找乘法次数的更少的算法，以及探索这个复杂度的下界，例如谷歌的 Deepmind 团队就有相关的工作。

6.3 张量的收缩、对称化与交错化

本节中如无特殊说明，始终设 V 是域 $\mathbb{K}$ 上的 n 维向量空间。

6.3.1 张量的乘积与收缩

现在我们考虑张量的运算。首先，张量映射 $\otimes$ 给出了一个 $\mathbb{T}^p_q(V)\times\mathbb{T}^r_s(V)\to\mathbb{T}^{p+r}_{q+s}(V)$ 的双线性映射：

$$\begin{aligned} \otimes:\ & \mathbb{T}^p_q(V)\times\mathbb{T}^r_s(V)\to\mathbb{T}^{p+r}_{q+s}(V)\\ & \left(\mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_p\otimes f_1\otimes\cdots\otimes f_q,\ \mathbf{v}'_1\otimes\cdots\otimes\mathbf{v}'_r\otimes f'_1\otimes\cdots\otimes f'_s\right)\\ \mapsto\ & \mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_p\otimes\mathbf{v}'_1\otimes\cdots\otimes\mathbf{v}'_r\otimes f_1\otimes\cdots\otimes f_q\otimes f'_1\otimes\cdots\otimes f'_s. \end{aligned} \tag{$\star\star\star$}$$

这个运算被称为张量的乘积 (实际上就是张量积)，它的良定义性由泛性质保证。特别地，我们知道 $\mathbb{T}^1_1(V)$ 与 $\mathrm{End}(V)$ 是同构的，而我们现在定义的张量的乘积 $\mathbb{T}^1_1(V)\times\mathbb{T}^1_1(V)\to\mathbb{T}^2_2(V)$ 实际上就是定义6.1.3中所定义的线性算子的张量积。事实上，双线性性意味着我们只需要对可分解张量验证这个一致性就足够了。设 $\mathbf{v}_1\otimes f_1$，$\mathbf{v}_2\otimes f_2\in\mathbb{T}^1_1(V)$，它们分别对应线性算子 $\mathcal{A}:\mathbf{w}\mapsto f_1(\mathbf{w})\cdot\mathbf{v}_1$ 和 $\mathcal{B}:\mathbf{w}'\mapsto f_2(\mathbf{w}')\cdot\mathbf{v}_2$，那么，按照张量乘积的定义，这两个 $(1,1)$ 型张量的乘积是 $\mathbf{v}_1\otimes\mathbf{v}_2\otimes f_1\otimes f_2$，它对应线性算子 $\mathbf{w}\otimes\mathbf{w}'\mapsto f_1(\mathbf{w})f_2(\mathbf{w}')\cdot\mathbf{v}_1\otimes\mathbf{v}_2$，而这个线性算子正是 $\mathcal{A}\otimes\mathcal{B}$。

现在我们来看另一种张量的运算。由于我们可以自然地将 V^* 中的元素作用到 V 上，因此我们可以定义一个线性映射 $\mathbb{T}^p_q(V)\to\mathbb{T}^{p-1}_{q-1}(V)$ 如下：

$$\begin{aligned} \mathrm{tr}^i_j:\ & \mathbb{T}^p_q(V)\to\mathbb{T}^{p-1}_{q-1}(V)\\ & \mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_p\otimes f_1\otimes\cdots\otimes f_q\\ \mapsto\ & (f_j(\mathbf{v}_i))\cdot\mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_{i-1}\otimes\mathbf{v}_{i+1}\otimes\cdots\otimes\mathbf{v}_p\otimes f_1\otimes\cdots\otimes f_{j-1}\otimes f_{j+1}\otimes\cdots\otimes f_q. \end{aligned} \tag{6.3.1}$$

也就是说，我们把第 j 个 V^* 作用到第 i 个 V 上的运算就是 tr^i_j。对可分解张量这样定义以后，由线性性质我们可以把这个映射定义到整个 $\mathbb{T}^p_q(V)$ 上，并且泛性质保证了 tr^i_j 是良定义的。

定义 6.3.1. 我们把式 (6.3.1) 定义的 tr^i_j 称为 (q,p) 型张量按照第 j 个共变指标和第 i 个反变指标的**收缩**或**缩并** (contraction)，也叫做**卷积** (convolution)。

下面我们来计算张量收缩的坐标形式。设取定 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 及其对偶基 $\mathbf{e}^1,\ldots,\mathbf{e}^n$ 后，张量 $\mathbf{T}\in\mathbb{T}^p_q(V)$ 可以表示为

$$\mathbf{T}=\sum_{i_1=1}^{n}\cdots\sum_{i_p=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_q=1}^{n}\left(T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_p}\otimes\mathbf{e}^{j_1}\otimes\cdots\otimes\mathbf{e}^{j_q}\right), \tag{6.3.2}$$

那么，tr_s^r 作用到 $\mathbf{T}$ 上就会得到：

$$
\begin{aligned}
\mathrm{tr}_r^s(\mathbf{T}) &= \sum_{i_1=1}^{n}\cdots\sum_{i_p=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_q=1}^{n}\left(T_{j_1,\dots,j_q}^{i_1,\dots,i_p}\mathrm{tr}_s^r\left(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_p}\otimes\mathbf{e}^{j_1}\otimes\cdots\otimes\mathbf{e}^{j_q}\right)\right)\\
&= \sum_{i_1=1}^{n}\cdots\sum_{i_p=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_q=1}^{n}\left(T_{j_1,\dots,j_q}^{i_1,\dots,i_p}\delta_{i_r,j_s}\cdot\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_{r-1}}\otimes\mathbf{e}_{i_{r+1}}\otimes\cdots\otimes\mathbf{e}_{i_p}\otimes\mathbf{e}^{j_1}\otimes\cdots\otimes\mathbf{e}^{j_{s-1}}\otimes\mathbf{e}^{j_{s+1}}\otimes\cdots\otimes\mathbf{e}^{j_q}\right)\\
&= \sum_{i_1=1}^{n}\cdots\sum_{i_{r-1}}^{n}\sum_{i_{r+1}}^{n}\cdots\sum_{i_p=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_{s-1}}^{n}\sum_{i_{s+1}}^{n}\cdots\sum_{j_q=1}^{n}\\
&\quad\left(\overline{T_{j_1,\dots,j_{s-1},j_{s+1},\dots,j_q}^{i_1,\dots,i_{r-1},i_{r+1},\dots,i_p}}\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_{r-1}}\otimes\mathbf{e}_{i_{r+1}}\otimes\cdots\otimes\mathbf{e}_{i_p}\otimes\mathbf{e}^{j_1}\otimes\cdots\otimes\mathbf{e}^{j_{s-1}}\otimes\mathbf{e}^{j_{s+1}}\otimes\cdots\otimes\mathbf{e}^{j_q}\right),
\end{aligned}
$$

其中

$$
\begin{aligned}
\overline{T_{j_1,\dots,j_{s-1},j_{s+1},\dots,j_q}^{i_1,\dots,i_{r-1},i_{r+1},\dots,i_p}} &= \sum_{i_r=1}^{n}\sum_{j_s=1}^{n}T_{j_1,\dots,j_q}^{i_1,\dots,i_p}\delta_{i_r,j_s}\\
&= \sum_{k=1}^{n}T_{j_1,\dots,j_{s-1},k,j_{s+1},\dots,j_q}^{i_1,\dots,i_{r-1},k,i_{r+1},\dots,i_p}.
\end{aligned}
\tag{6.3.3}
$$

注意到全部的

$$
\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_{r-1}}\otimes\mathbf{e}_{i_{r+1}}\otimes\cdots\otimes\mathbf{e}_{i_p}\otimes\mathbf{e}^{j_1}\otimes\cdots\otimes\mathbf{e}^{j_{s-1}}\otimes\mathbf{e}^{j_{s+1}}\otimes\cdots\otimes\mathbf{e}^{j_q}
$$

是 $\mathbb{T}_q^p(V)$ 的一组基，因此式 (6.3.3) 就是张量 $\mathbf{T}$ 在按共变指标 s 和反变指标 r 收缩后的坐标展开式。

对于一个 (q,p) 型张量，我们可以依次做 $\min\{p,q\}$ 次收缩，这样我们会得到一个 $(0,p-q)$ 的反变张量 $(p>q)$，或者一个纯量 $(p=q)$，或者一个 $(q-p,0)$ 的共变张量，我们把最后得到的张量称为原来张量的完全收缩 (complete contraction)。

下面的例子告诉我们，线性算子的迹和复合都是张量收缩的特殊情形。

例 6.3.1. 取定 V 的一组基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 及其对偶基 $\mathbf{e}^1,\dots,\mathbf{e}^n$，设 $\mathscr{A}\in\mathrm{End}(V)$ 在这组基下的矩阵为 $A=(a_{ij})_{n\times n}$，则按照推论6.2.3，$\mathscr{A}$ 也可以视作一个 $(1,1)$ 型混合张量 $\mathbf{T}_{\mathscr{A}}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{e}_i\otimes\mathbf{e}^j$(验证是容易的，留作练习)。同理我们可设 $\mathscr{B}\in\mathrm{End}(V)$ 的矩阵是为 $B=(b_{ij})_{n\times n}$，对应张量 $\mathbf{T}_{\mathscr{B}}$

(1) 用 tr_1^1 作用到 $\mathbf{T}_{\mathscr{A}}$ 上，则应该得到一个域 $\mathbb{K}$ 中的纯量，按照 (6.3.3) 式我们得到

$$
\mathrm{tr}_1^1(\mathbf{T}_{\mathscr{A}})=\sum_{k=1}^{n}a_{kk}=\mathrm{tr}(A)=\mathrm{tr}(\mathscr{A}),
$$

也就是说，$(1,1)$ 型混合张量的收缩与线性算子 (或者说矩阵) 的迹是一致的，这也是张量收缩这个记号的由来。

(2) 考虑 $\mathbf{T}_{\mathscr{A}}\otimes\mathbf{T}_{\mathscr{B}}$，我们知道它应该等于

$$
\mathbf{T}_{\mathscr{A}}\otimes\mathbf{T}_{\mathscr{B}}=\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}a_{ij}b_{kl}\cdot\left(\mathbf{e}_i\otimes\mathbf{e}_k\otimes\mathbf{e}^j\otimes\mathbf{e}^l\right).
$$

对其沿着第 1 个共变指标和第 2 个反变指标进行收缩，我们得到：

$$\mathrm{tr}_1^2(\mathbf{T}_{\mathcal{A}} \otimes \mathbf{T}_{\mathcal{B}}) = \sum_{i=1}^{n} \sum_{l=1}^{n} (\sum_{m=1}^{n} a_{im} b_{ml}) \cdot (\mathbf{e}_i \otimes \mathbf{e}^l),$$

这是是一个 $(1,1)$ 型混合张量，它对应着线性算子 C，其矩阵 $C = (c_{il})_{n\times n}$，$c_{il} = \sum_{m=1}^{n} a_{im} b_{ml}$，显然 $C = AB$，或者说 $\mathrm{tr}_1^2(\mathbf{T}_{\mathcal{A}} \otimes \mathbf{T}_{\mathcal{B}})$ 对应线性算子 $\mathcal{AB}$，也即两个 $(1,1)$ 型混合张量作乘积后收缩一次，得到的 $(1,1)$ 型张量与线性算子的复合是一致的。

更进一步，我们也可以对 $\mathbf{T}_{\mathcal{A}} \otimes \mathbf{T}_{\mathcal{B}}$ 沿着第 2 个共变指标和第 1 个反变指标收缩，这样得到的 $(1,1)$ 型张量对应线性算子 $\mathcal{BA}$。如果我们对 $\mathrm{tr}_1^2(\mathbf{T}_{\mathcal{A}} \otimes \mathbf{T}_{\mathcal{B}})$ 和 $\mathrm{tr}_2^1(\mathbf{T}_{\mathcal{A}} \otimes \mathbf{T}_{\mathcal{B}})$ 再收缩一次，我们会发现它们是相等的，这也就是矩阵的迹满足 $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ 的原因。细节留作练习。

6.3.2 张量的对称化与交错化

我们在第一章已经研究过两种特殊的双线性映射：对称双线性型与斜对称双线性型。现在我们把它们推广到更一般的多重线性映射情形。本小节中如无特殊说明，我们始终设域 $\mathbb{K}$ 的特征是 0。

定义 6.3.2. 设反变张量 $\mathbf{T} \in \mathbb{T}_0^m(V)$，则由推论6.2.4，$\mathbf{T}$ 可以视作一个 $\underbrace{V^* \times \cdots \times V^*}_{m个} \to \mathbb{K}$ 的多重线性函数 $\Phi_{\mathbf{T}}$，$\Phi_{\mathbf{T}}$ 的定义如下：

如果 $\mathbf{T} = \sum_{i=1}^{s} \mathbf{v}_1^{(i)} \otimes \cdots \otimes \mathbf{v}_m^{(i)}$，则对 $\forall f_1, \ldots, f_m \in V^*$，有

$$\Phi_{\mathbf{T}}(f_1, \ldots, f_m) = \sum_{i=1}^{s} f_1(\mathbf{v}_1^{(i)}) \cdots f_m(\mathbf{v}_m^{(i)}).$$

泛性质保证了 $\Phi_{\mathbf{T}}$ 是良定义的。

现在我们开始定义对称张量和斜对称张量。如果张量 $\mathbf{T}$ 对应的多重线性函数 $\Phi_{\mathbf{T}}$ 是一个对称函数 (讲义上册定义 1.5.6)，即 $\forall \pi \in S_m$(置换群)，都有

$$\Phi_{\mathbf{T}}(f_{\pi(1)}, \ldots, f_{\pi(m)}) = \Phi_{\mathbf{T}}(f_1, \ldots, f_m),$$

那么我们称 $\mathbf{T}$ 是一个对称张量 (symmetric tensor)；如果

$$\Phi_{\mathbf{T}}(f_{\pi(1)}, \ldots, f_{\pi(m)}) = \varepsilon_\pi \Phi_{\mathbf{T}}(f_1, \ldots, f_m), \quad 其中\varepsilon_\pi是置换的符号,$$

那么我们称 $\mathbf{T}$ 是一个斜对称张量 (或反对称张量，antisymmetric tensor)。

类似地，对于共变张量 $\mathbf{T} \in \mathbb{T}_m^0(V)$，它可以视作 $\mathrm{mul}(V \times \cdots \times V; \mathbb{K})$ 中的多重线性函数，如果这个多重线性函数是对称的 (或斜对称的)，那么我们称这个共变张量是对称的 (相对应地，斜对称的)。细节与反变情形相同，不再罗列。注意到有限维空间中 $V \simeq V^*$，因此在取定一组基后我们研究 $\mathbb{T}_m^0(V)$ 中的对称张量和斜对称张量与研究 $\mathbb{T}_0^m(V)$ 中的对称和斜对称张量是类似的，因此本小节之后都只关注反变张量积空间中的情形。

现在我们考虑对称张量和斜对称张量的坐标形式。设 V 的一组基是 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，其对偶基是 $\mathbf{e}^1, \ldots, \mathbf{e}^n$，则 $\mathbf{T}$ 是对称张量 (或斜对称张量) 当且仅当

$$\Phi_{\mathbf{T}}(\mathbf{e}^{k_{\pi(1)}}, \ldots, \mathbf{e}^{k_{\pi(m)}}) = \Phi_{\mathbf{T}}(\mathbf{e}^{k_1}, \ldots, \mathbf{e}^{k_m}) \quad (或者\Phi_{\mathbf{T}}(\mathbf{e}^{k_{\pi(1)}}, \ldots, \mathbf{e}^{k_{\pi(m)}}) = \varepsilon_\pi \Phi_{\mathbf{T}}(\mathbf{e}^{k_1}, \ldots, \mathbf{e}^{k_m}))$$

对 $\forall\pi\in S_m$ 及 $\forall\mathbf{e}^{k_1},\ldots,\mathbf{e}^{k_m}$ 都成立，这只需要注意到 $\Phi_{\mathbf{T}}$ 的多重线性性质，再将 $f_1,\ldots,f_m$ 写成对偶基的线性组合即可证明。如果我们设

$$\mathbf{T}=\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n}T^{i_1,\ldots,i_m}\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_m},$$

那么 $\Phi_{\mathbf{T}}(\mathbf{e}^{k_1},\ldots,\mathbf{e}^{k_m})=T^{k_1,\ldots,k_m}$，因此 $\mathbf{T}$ 是对称张量 (或斜对称张量) 等价于

$$T^{i_{\pi(1)},\ldots,i_{\pi(m)}}=T^{i_1,\ldots,i_m}\quad(\text{或者}T^{i_{\pi(1)},\ldots,i_{\pi(m)}}=\varepsilon_\pi T^{i_1,\ldots,i_m})\tag{6.3.4}$$

对 $\forall\pi\in S_m$ 及 $i_1,\ldots,i_m\in\{1,\ldots,n\}$ 都成立即可。这就是对称张量 (或斜对称张量) 的坐标形式。特别地，如果我们把矩阵 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{K})$ 视作 $(2,0)$ 型的张量 (即 $V*\times V^*\to\mathbb{K}$ 的双线性型)，那么式 (6.3.4) 实际上就是 $a_{ij}=a_{ji}$ (或者$a_{ij}=-a_{ji}$)，$\forall i,j\in\{1,\ldots,n\}$，即 $A=A^t$(或者 $A=-A^t$)，这与第一章中的对称 (或斜对称) 双线性型的情形是一致的。

对于矩阵的情形，我们知道 $\forall A\in M_n(\mathbb{K})$，$\dfrac{1}{2}(A+A^t)$ 是对称矩阵，而 $\dfrac{1}{2}(A-A^t)$ 是斜对称矩阵。我们可以把这个方式推广到张量上，即下面的定义：

定义 6.3.3. *首先，任给 $\pi\in S_m$，π 按如下方式诱导了一个 $\mathbb{T}^m_0(V)\to\mathbb{T}^m_0(V)$ 的线性映射 f_π：*

$$\begin{aligned}f_\pi:\ \mathbb{T}^m_0(V)&\longrightarrow\mathbb{T}^m_0(V)\\ \mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_m&\longmapsto\mathbf{v}_{\pi(1)}\otimes\cdots\otimes\mathbf{v}_{\pi(m)}\end{aligned}$$

f_π 的良定义性由下面的交换图所示的泛性质保证：

$$\begin{tikzcd}
V\times\cdots\times V \arrow[r] \arrow[dd] & \mathbb{T}^m_0(V) \arrow[ddl, "f_\pi"] \\
(\mathbf{v}_1,\ldots,\mathbf{v}_m) & \mathbf{v}_1\otimes\cdots\otimes\mathbf{v}_m \\
\mathbb{T}^m_0(V) & \\
\mathbf{v}_{\pi(1)}\otimes\cdots\otimes\mathbf{v}_{\pi(m)} &
\end{tikzcd}$$

于是我们可以定义 $\mathbb{T}^m_0(V)\to\mathbb{T}^m_0(V)$ 的两个线性算子如下：

$$\mathfrak{S}:\mathbb{T}^m_0(V)\to\mathbb{T}^m_0(V),\quad\mathbf{T}\mapsto\frac{1}{m!}\sum_{\pi\in S_m}f_\pi(\mathbf{T}),$$

$$\mathfrak{A}:\mathbb{T}^m_0(V)\to\mathbb{T}^m_0(V),\quad\mathbf{T}\mapsto\frac{1}{m!}\sum_{\pi\in S_m}\varepsilon_\pi f_\pi(\mathbf{T})$$

我们把 $\mathfrak{S}$ 和 $\mathfrak{A}$ 分别称为对称化算子和斜对称化 (交错化) 算子[1]。

首先，借助于对称 (或斜对称) 张量的坐标形式我们很容易看出，$\mathbf{T}\in\mathbb{T}^m_0(V)$ 是对称张量 (或斜对称张量) $\Longleftrightarrow$ $f_\pi(\mathbf{T})=\mathbf{T}$(或者 $f_\pi(\mathbf{T})=\varepsilon_\pi\mathbf{T}$) 对 $\forall\pi\in S_m$ 成立。其次，我们很容易验证：$\mathbb{T}^m_0(V)$ 中的所有对称张量 (或所有斜对称张量) 在张量的加法和数乘下构成了一个子空间，我们把这个子空间记为 $\mathbb{T}^m_0(V)^+$(或 $\mathbb{T}^m_0(V)^-$)(这个记号与 $L_2^+(V,\mathbb{K})$ 和 $L_2^-(V,\mathbb{K})$ 是一脉相承的)。此外，$\mathbb{T}^m_0(V)^-$ 也可以记作 $\Lambda^m(V)$，我们在下一节会更多地使用这个记号。对称化算子和交错化算子还具有以下性质：

---

[1] 这两个记号分别是德文尖角体的 S 和 A。

定理 6.3.1. 条件同定义6.3.3。则

(1) $\mathrm{im}(\mathfrak{S}) = \mathbb{T}^m_0(V)^+$, $\mathrm{im}(\mathfrak{A}) = \mathbb{T}^m_0(V)^-$.

(2) $\mathfrak{S}^2 = \mathfrak{S}$, $\mathfrak{A}^2 = \mathfrak{A}$.

(3) $\forall \sigma \in S_m$及$\mathbf{T} \in \mathbb{T}^m_0$, $\mathfrak{A}(f_\sigma(\mathbf{T})) = \varepsilon_\sigma \mathfrak{A}(\mathbf{T})$.

证明. 首先，按照 f_π 的定义我们很容易验证：$\forall \pi, \sigma \in S_m$，$f_{\sigma\circ\pi} = f_\sigma \circ f_\pi$。

(1) 先证 $\mathrm{im}(\mathfrak{S}) \subset \mathbb{T}^m_0(V)^+$。任取 $\frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T}) \in \mathrm{im}(\mathfrak{S})$，我们来验证 $\forall \sigma \in S_m$，$f_\sigma(\frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T})) = \frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T})$ 即可说明 $\frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T}) \in \mathbb{T}^m_0(V)^+$。事实上，由于 f_σ 是线性映射，我们有

$$\begin{aligned} f_\sigma(\frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T})) &= \frac{1}{m!}\sum_{\pi\in S_m} f_\sigma \circ f_\pi(\mathbf{T}) \\ &= \frac{1}{m!}\sum_{\pi\in S_m} f_{\sigma\circ\pi}(\mathbf{T}) \\ &= \frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T}). \qquad \text{(注意到求和项是完全一样的)} \end{aligned}$$

再证 $\mathbb{T}^m_0(V)^+ \subset \mathrm{im}(\mathfrak{S})$，事实上，任取对称张量 $\mathbf{T}$，$\mathbf{T}$ 本身就是 $\mathfrak{S}$ 的一个原像。这是因为，此时 $\forall \pi \in S_m$，$f_\pi(\mathbf{T}) = \mathbf{T}$，因此

$$\frac{1}{m!}\sum_{\pi\in S_m} f_\pi(\mathbf{T}) = \frac{1}{m!}\sum_{\pi\in S_m} \mathbf{T} = \mathbf{T}.$$

即 $\mathfrak{S}(\mathbf{T}) = \mathbf{T}$ 成立。

结合以上两方面即得 $\mathrm{im}(\mathfrak{S}) = \mathbb{T}^m_0(V)^+$。同理可证 $\mathrm{im}(\mathfrak{A}) = \mathbb{T}^m_0(V)^-$，留作练习。

(2) 我们在 (1) 中已经证明了 $\mathfrak{S}$ 的像空间是全体对称张量以及 $\mathfrak{S}|_{\mathbb{T}^{m+}_0} = \mathrm{id}$，因此 $\mathfrak{S}^2 = \mathfrak{S}$。同理可证 $\mathfrak{A}^2 = \mathfrak{A}$。

(3) 任取 $\sigma \in S_m$ 及 $\mathbf{T} \in \mathbb{T}^m_0(V)$，我们有

$$\begin{aligned} \mathfrak{A}(f_\sigma(\mathbf{T})) &= \sum_{\pi\in S_m} \varepsilon_\pi \cdot f_\pi \circ f_\sigma(\mathbf{T}) \\ &= \varepsilon_\sigma \sum_{\pi\in S_m} \varepsilon_\pi\varepsilon_\sigma \cdot f_\pi \circ f_\sigma(\mathbf{T}) && (\varepsilon_\sigma^2 = 1) \\ &= \varepsilon_\sigma \sum_{\pi\in S_m} \varepsilon_{\pi\circ\sigma} \cdot f_\pi \circ f_\sigma(\mathbf{T}) && (\varepsilon_\pi\varepsilon_\sigma = \varepsilon_{\pi\circ\sigma}，\text{讲义上册命题 1.5.2}) \\ &== \varepsilon_\sigma \sum_{\pi\circ\sigma\in S_m} \varepsilon_{\pi\circ\sigma} \cdot f_{\pi\circ\sigma}(\mathbf{T}) && (\text{求和项完全一样}) \end{aligned}$$

即得结论。这样我们就完成了证明。

□

需要注意的是，一般情形下 $\mathbf{T} = \mathfrak{S}(\mathbf{T}) + \mathfrak{A}(\mathbf{T})$ 并不成立，例如下面的例子。

例 6.3.2. 在 $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ 中取 $\mathbf{T} = \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1$，则容易计算出

$$\begin{aligned}\mathfrak{S}(\mathbf{T}) &= \frac{1}{3}(\mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \mathbf{e}_1 + \mathbf{e}_1 \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1 \otimes \mathbf{e}_1),\\ \mathfrak{A}(\mathbf{T}) &= \mathbf{0}.\end{aligned}$$

在第一章中我们知道，二次型可以找到与之相配极的对称双线性型，下面我们把这一结论也推广到高次的齐次函数上。

定义 6.3.4. 如果向量空间 V 上的函数 $Q : V \to \mathbb{K}$ 满足：存在 p 重线性函数 $F : \underbrace{V \times \cdots \times V}_{p\text{个}} \to \mathbb{K}$ 使得 $Q(\mathbf{x}) = F(\mathbf{x}, \ldots, \mathbf{x})$，那么我们称 Q 是 V 上的 p 次齐次函数。

由推论6.2.2可知，我们可以把 F 视作 $\mathbb{T}^p_0(V^*)$ 中的一个张量，那我们自然可以把对称化算子作用到 F 上去，再把得到的张量还原成 p 重线性函数，简便起见我们把这个得到的函数直接记作 $\mathfrak{S}(F)$。由定义很容易看出 $\mathfrak{S}(F)$ 是对称函数，并且 $\mathfrak{S}(F)(\mathbf{x}, \ldots, \mathbf{x}) = F(\mathbf{x}, \ldots, \mathbf{x})$, $\forall \mathbf{x} \in V$，即 $\mathfrak{S}(F)$ 与 F 诱导的 p 次齐次函数是相同的。

定义 6.3.5. 对于 p 次齐次函数 $Q : V \to \mathbb{K}$，如果**对称**的 p 重线性函数 F 满足 $Q(\mathbf{x}) = F(\mathbf{x}, \ldots, \mathbf{x})$，那么我们称 F 与 Q 配极 (polarization)，或者说 F 是 Q 的一个**极化**。

定理 6.3.2. p 次齐次函数 $Q : V \to \mathbb{K}$ 的极化是存在且唯一的。

证明. 由定义6.3.4下方的说明可知，$\mathfrak{S}(F)$ 就是 Q 的一个极化，因此存在性显然成立。下证唯一性。

取 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，设 Q 由 p 重线性函数 F 定义，且 F 视作 $\mathbb{T}^p_0(V^*)$ 中的张量时其坐标形式为：

$$F = \sum_{i_1,\ldots,i_p} F_{i_1,\ldots,i_p} \cdot \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_p},$$

由 $\mathfrak{S}$ 的定义可以直接计算出 $\mathfrak{S}(F)$ 的第 $i_1, \ldots, i_p$ 坐标为

$$\mathfrak{S}(F)_{i_1,\ldots,i_p} = \frac{1}{p!} \sum_{\pi \in S_p} F^{i_{\pi(1)},\ldots,i_{\pi(p)}}.$$

另一方面，对于 $\forall \mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n \in V$，我们有

$$Q(\mathbf{x}) = F(\mathbf{x}, \ldots, \mathbf{x}) = \sum_{i_1,\ldots,i_p} F_{i_1,\ldots,i_p} \cdot x_{i_1} \cdots x_{i_p}. \tag{6.3.5}$$

设对称的 p 重线性函数 G 是 Q 的一个极化，且 G 的坐标形式为

$$G = \sum_{i_1,\ldots,i_p} G_{i_1,\ldots,i_p} \cdot \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_p},$$

那么 G 应该满足

$$G_{i_{\pi(1)},\ldots,i_{\pi(p)}} = G_{i_1,\ldots,i_p}, \quad \forall \pi \in S_p,\ i_1, \ldots, i_p \in \{1, \ldots, n\}, \tag{6.3.6}$$

并且

$$\sum_{i_1,\ldots,i_p} G_{i_1,\ldots,i_p} \cdot x_{i_1} \cdots x_{i_p} = \sum_{i_1,\ldots,i_p} F_{i_1,\ldots,i_p} \cdot x_{i_1} \cdots x_{i_p}, \quad \forall x_{i_1}, \ldots, x_{i_p} \in \mathbb{K}. \tag{6.3.7}$$

式 (6.3.7) 实际上就是两个 p 次齐次多项式相等，因此只需要它们对应的系数相等即可。注

意到 $x_{i_{\pi(1)}}\cdots x_{i_{\pi(p)}}=x_{i_1}\cdots x_{i_p}$，因此

$$\sum_{\pi\in S_p}G_{i_{\pi(1)},\dots,i_{\pi(p)}}=\sum_{\pi\in S_p}F_{i_{\pi(1)},\dots,i_{\pi(p)}}.$$

再结合 (6.3.6) 式可知

$$G_{i_1,\dots,i_p}=\frac{1}{p!}\sum_{\pi\in S_p}F_{i_{\pi(1)},\dots,i_{\pi(p)}}.$$

因此 $G=\mathfrak{S}(F)$，唯一性得证。这样我们就完成了证明。 □

从定理6.3.2的证明中我们可以看到，V 上所有的 p 次齐次函数全体实际上与域 $\mathbb{K}$ 上的 n 变元 p 次齐次多项式构成的向量空间 $\mathbb{K}[x_1,\dots,x_n]_p$ 同构，并且同构于 $\mathbb{T}^p_0(V^*)^+$。特别地，把定理6.3.2限制到 $p=2$ 的情形就是我们所熟悉的二次型和对称双线性型的配极。

例 6.3.3. $\mathbb{R}^3$ 上的 3 次齐次函数 $Q:(x_1,x_2,x_3)^t\mapsto x_1x_2^2+x_3^3$ 是由 $\mathbb{T}^3_0(\mathbb{R}^{3^*})$ 中的张量：

$$\mathbf{F}=\mathbf{e}^1\otimes\mathbf{e}^2\otimes\mathbf{e}^2+\mathbf{e}^3\otimes\mathbf{e}^3\otimes\mathbf{e}^3$$

定义的，我们可以计算出 $\mathbf{F}$ 的对称化：

$$\mathfrak{S}(\mathbf{F})=\frac{1}{3}(\mathbf{e}^1\otimes\mathbf{e}^2\otimes\mathbf{e}^2+\mathbf{e}^2\otimes\mathbf{e}^1\otimes\mathbf{e}^2+\mathbf{e}^2\otimes\mathbf{e}^2\otimes\mathbf{e}^1)+\mathbf{e}^3\otimes\mathbf{e}^3\otimes\mathbf{e}^3,$$

因此 Q 的极化 $\mathfrak{S}(F)$ 就是把 $\mathfrak{S}(\mathbf{F})$ 还原成对称的多重线性函数的形式，即

$$\begin{aligned}\mathfrak{S}(F):\mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3&\to\mathbb{R}\\((x_1,x_2,x_3)^t,(y_1,y_2,y_3)^t,(z_1,z_2,z_3)^t)&\mapsto\frac{1}{3}(x_1y_2z_2+x_2y_1z_2+x_2y_2z_1)+x_3y_3z_3.\end{aligned}$$

6.4 张量代数、对称代数与外代数

把向量空间上不同阶的反变 (或共变) 张量放在一起考虑，我们会得到更丰富的代数结构。

定义 6.4.1. 设 V 是域 $\mathbb{K}$ 上的向量空间，考虑反变张量空间 $\mathbb{T}_0^p(V)$，其中 $p \in \mathbb{N}$。特别地，$p = 0$ 时取 $\mathbb{T}_0^0 = \mathbb{K}$。我们考虑以下的无穷外直和：令

$$\mathbb{T}(V) = \bigoplus_{p=0}^{\infty} \mathbb{T}_0^p(V),$$

则 $\mathbb{T}(V)$ 仍然是 $\mathbb{K}$ 上的向量空间，并且我们可以在 $\mathbb{T}(V)$ 上以张量的乘积 $\otimes$ 定义两个不同阶张量的乘法 (6.3.1 小节 $(\star\star\star)$ 式)。$\mathbb{T}(V)$ 上有加法、数乘、和乘法三种运算，我们很容易验证这些运算满足定义2.2.1的条件，即 $\mathbb{T}(V)$ 是一个代数，我们称其为 V 上的张量代数 (tensor algebra)。

类似地，我们也可以定义

$$\mathbb{T}(V^*) = \bigoplus_{q=0}^{\infty} \mathbb{T}_q^0(V),$$

它是 V^* 上的张量代数。

很容易看出张量代数具有以下性质：若 $\mathbf{a} \in \mathbb{T}_0^p(V)$，$\mathbf{b} \in \mathbb{T}_0^q(V)$，则 $\mathbf{a} \otimes \mathbf{b} \in \mathbb{T}_0^{p+q}(V)$。这种性质表明张量代数是一个**分次代数**，下面我们给出分次代数的形式化定义。

定义 6.4.2. 设 A 是域 $\mathbb{K}$ 上的代数，如果 A 能够写成以下形式：

$$A = \bigoplus_{i=0}^{\infty} A_i,$$

其中每个 A_i 都是 A(作为向量空间) 的子空间，并且乘法 $\circ$ 满足：若 $a \in A_i$，$b \in A_j$，则 $a \circ b \in A_{i+j}$，则我们称 A 是一个分次代数 (graded algebra)，并且 A_i 中的元素被称为 i 次齐次的 (homogeneous of degree i)。如果 $a \in A$ 能够分解成：

$$a = a_{i_1} + \cdots + a_{i_n}，\ 其中 a_{i_k} \in A_{i_k}，\ k = 1, \ldots, n，\ i_k \neq i_j,$$

那么每个 a_{i_k} 被称为 a 的 i_k 次齐次部分 (homogeneous component of degree i_k)。

很容易看出，我们在讲义上册第六章中讲过的一元多项式和多元多项式都是分次代数的典型例子 (一元多项式的分次指标是次数，多元多项式的分次指标是全次数)。

下面我们将目光投向对称张量和斜对称张量。设域 $\mathbb{K}$ 的特征是 0，V 是域 $\mathbb{K}$ 上的 n 维向量空间。很容易看出，两个对称张量直接做张量积不一定得到对称张量，一个简单的例子是 $\mathbb{R}^2$ 中 $\mathbf{e}_1$ 和 $\mathbf{e}_2$ 都是 1 阶的对称张量，但 $\mathbf{e}_1 \otimes \mathbf{e}_2$ 却不是对称张量。类似地，两个斜对称的张量直接做张量积也不一定得到斜对称张量。因此，我们需要借助于对称化 (或交错化) 算子来定义对称张量 (或斜对称张量) 之间的乘法。在此之前，我们先要证明对称化 (或交错化) 算子具有以下性质：

引理 6.4.1. 设域 $\mathbb{K}$ 的特征是 0，V 是域 $\mathbb{K}$ 上的 n 维向量空间，$\mathfrak{S}$ 和 $\mathfrak{A}$ 分别是定义6.3.3中的对称化算子和交错化算子。则对任意的 $\mathbf{a} \in \mathbb{T}_0^p(V)$ 和 $\mathbf{b} \in \mathbb{T}_0^q(V)$，我们有

$$\mathfrak{S}(\mathfrak{S}(\mathbf{a}) \otimes \mathbf{b}) = \mathfrak{S}(\mathbf{a} \otimes \mathfrak{S}(\mathbf{b})) = \mathfrak{S}(\mathbf{a} \otimes \mathbf{b}),$$

$$\mathfrak{A}(\mathfrak{A}(\mathbf{a}) \otimes \mathbf{b}) = \mathfrak{A}(\mathbf{a} \otimes \mathfrak{A}(\mathbf{b})) = \mathfrak{A}(\mathbf{a} \otimes \mathbf{b}).$$

证明. 我们知道

$$\mathfrak{S}(\mathbf{a}) = \frac{1}{p!}\sum_{\pi\in S_p} f_\pi(\mathbf{a}),$$

于是

$$\mathfrak{S}(\mathbf{a})\otimes\mathbf{b} = \frac{1}{p!}\sum_{\pi\in S_p}\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right),$$

即

$$\mathfrak{S}(\mathfrak{S}(\mathbf{a})\otimes\mathbf{b}) = \frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}} f_\sigma\left(\frac{1}{p!}\sum_{\pi\in S_p}(f_\pi(\mathbf{a})\otimes\mathbf{b})\right).$$

利用线性性质，我们通过交换求和顺序可得

$$\begin{aligned}\mathfrak{S}(\mathfrak{S}(\mathbf{a})\otimes\mathbf{b}) &= \frac{1}{p!}\sum_{\pi\in S_p}\left(\frac{1}{(p+q)!}\sum_{\sigma\in S_{p+q}} f_\sigma\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right)\right)\\ &= \frac{1}{p!}\sum_{\pi\in S_p}\left(\mathfrak{S}\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right)\right).\end{aligned}$$

下面我们说明：$\forall\pi\in S_p$，$\mathfrak{S}\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right) = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b})$，从而 $\dfrac{1}{p!}\sum_{\pi\in S_p}\left(\mathfrak{S}\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right)\right) = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b})$，结论 $\mathfrak{S}(\mathfrak{S}(\mathbf{a})\otimes\mathbf{b}) = \mathfrak{S}(\mathbf{a}\otimes\mathfrak{S}(\mathbf{b}))$ 得证。为此，我们注意到：

> $\forall\pi\in S_p$，$\exists\widetilde{\pi}\in S_{p+q}$ 使得 $f_\pi(\mathbf{a})\otimes\mathbf{b} = f_{\widetilde{\pi}}(\mathbf{a}\otimes\mathbf{b})$，其中 $\widetilde{\pi}$ 作用在 $\{1,\dots,p\}$ 上与 π 相等，而作用在 $\{p+1,\dots,p+q\}$ 上是恒等映射。

因此

$$\forall\pi\in S_p,\ \mathfrak{S}\left(f_\pi(\mathbf{a})\otimes\mathbf{b}\right) = \mathfrak{S}(f_{\widetilde{\pi}}(\mathbf{a}\otimes\mathbf{b})) = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b}),$$

其中后一步是利用 $\mathfrak{S}$ 的定义，里边的求和项只是换了顺序。这样我们就证明了 $\mathfrak{S}(\mathfrak{S}(\mathbf{a})\otimes\mathbf{b}) = \mathfrak{S}(\mathbf{a}\otimes\mathfrak{S}(\mathbf{b}))$。$\mathfrak{S}(\mathbf{a}\otimes\mathfrak{S}(\mathbf{b})) = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b})$ 同理。

类似地，我们把上面证明中的对称化算子换成交错化算子就可以证明另一个恒等式，这里需要注意置换的符号在运算过程中被消掉了，细节留作练习。 □

定理 6.4.1. *条件同引理*6.4.1。*则* $\forall\mathbf{a}\in\mathbb{T}_0^p(V)$, $\mathbf{b}\in\mathbb{T}_0^q(V)$, $\mathbf{c}\in\mathbb{T}_0^r(V)$，*我们有*

$$\begin{aligned}\mathfrak{S}(\mathfrak{S}(\mathbf{a}\otimes\mathbf{b})\otimes\mathbf{c}) = \mathfrak{S}(\mathbf{a}\otimes\mathfrak{S}(\mathbf{b}\otimes\mathbf{c})) = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}),\\ \mathfrak{A}(\mathfrak{A}(\mathbf{a}\otimes\mathbf{b})\otimes\mathbf{c}) = \mathfrak{A}(\mathbf{a}\otimes\mathfrak{A}(\mathbf{b}\otimes\mathbf{c})) = \mathfrak{A}(\mathbf{a}\otimes\mathbf{b}\otimes\mathbf{c}).\end{aligned}$$

这是引理6.4.1的直接推论。

定义 6.4.3. *设域* $\mathbb{K}$ *的特征是* 0，V *是域* $\mathbb{K}$ *上的* n *维向量空间。考虑以下的无穷外直和：令*

$$\mathsf{S}(V) = \bigoplus_{p=0}^{\infty}\mathbb{T}_0^p(V)^+ \quad (\text{仍令}\mathbb{T}_0^0(V)^+ = \mathbb{K}),$$

$\mathsf{S}(V)$ *仍然是* $\mathbb{K}$ *上的向量空间，并且我们可以在* $\mathsf{S}(V)$ *上定义如下的乘法* ”$\vee$”：

$$\mathbf{a}\vee\mathbf{b} = \mathfrak{S}(\mathbf{a}\otimes\mathbf{b}),\ \forall\mathbf{a}\in\mathbb{T}_0^p(V)^+, \mathbf{b}\in\mathbb{T}_0^q(V)^+,$$

则由定理6.4.1，乘法 ”∨” 满足结合律，因此容易验证 $\mathsf{S}(V)$ 在乘法 ”∨” 下构成了一个 $\mathbb{K}$ 上的分次代数 (留作练习)，称为**对称张量代数** (symmetric tensor algebra)。

同理，令

$$\mathsf{A}(V) = \bigoplus_{p=0}^{\infty} \mathbb{T}_0^p(V)^{-} \quad (\mathbb{T}_0^0(V)^{-} = \mathbb{K}),$$

并在 $\mathsf{A}(V)$ 上定义如下的乘法 ”∧”:

$$\mathbf{a} \wedge \mathbf{b} = \mathfrak{A}(\mathbf{a} \otimes \mathbf{b}),\ \forall \mathbf{a} \in \mathbb{T}_0^p(V)^{-}, \mathbf{b} \in \mathbb{T}_0^q(V)^{-}.$$

则由定理6.4.1，”∧” 也满足结合律，因此容易验证 $\mathsf{A}(V)$ 在乘法 ”∧” 下构成了一个 $\mathbb{K}$ 上的分次代数 (留作练习)，称为**斜对称张量代数** (antisymmetric tensor algebra) 或**外代数** (exterior algebra) 或 Grassmann 代数。我们也把外代数上的乘法运算 ”∧” 称为**外积** (exterior product)，$\mathsf{A}(V)$ 也记作 $\Lambda(V)$。

实际上，对称张量代数和外代数可以不依赖于基域的特征来定义，细节可以参考《Advanced Linear Algebra》，Steven Roman，GTM135 的定理 14.17。

上一节的最后部分已经暗示了我们，对称张量代数实际上同构于多元多项式代数，其齐次部分同构于相应的齐次多项式 (细节留作练习)。以下我们主要考虑外代数。首先需要说明的是，外代数看起来是一个无穷维向量空间，但实际上是有限维的，这是因为：当 $p > n$ 时，$\mathbb{T}_0^p(V)^{-}$(即$\Lambda^p(V)$) $= \{\mathbf{0}\}$。为了说明这一点，我们需要进行以下的推导。

引理 6.4.2. 外积运算 ”∧” 具有以下性质：

(1) $\wedge$ 是双线性映射；

(2) $\forall \mathbf{v} \in V,\ \mathbf{v} \wedge \mathbf{v} = \mathbf{0}$;

(3) $\forall \mathbf{v}, \mathbf{w} \in V,\ \mathbf{v} \wedge \mathbf{w} = -\mathbf{w} \wedge \mathbf{v}$;

(4) $\forall \mathbf{v}_1, \ldots, \mathbf{v}_p \in V,\ \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p = \mathfrak{A}(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_p)$;

(5) $\forall \mathbf{v}_1, \ldots, \mathbf{v}_p \in V$ 及 $\pi \in S_p$，$\mathbf{v}_{\pi(1)} \wedge \cdots \wedge \mathbf{v}_{\pi(p)} = \varepsilon_\pi \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p$;

(6) $\forall \mathbf{v}_1, \ldots, \mathbf{v}_p \in V$，如果存在两个不同的指标 i, j 使得 $\mathbf{v}_i = \mathbf{v}_j$，那么 $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p = \mathbf{0}$;

(7) $\forall \mathbf{v}_1, \ldots, \mathbf{v}_p, \mathbf{w}_1, \ldots, \mathbf{w}_q \in V,\ \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p \wedge \mathbf{w}_1 \wedge \cdots \wedge \mathbf{w}_q = (-1)^{pq} \mathbf{w}_1 \wedge \cdots \wedge \mathbf{w}_q \wedge \mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_p$。更进一步，如果 $\mathbf{T}_1 \in \Lambda^p(V)$，$\mathbf{T}_2 \in \Lambda^q(V)$，则 $\mathbf{T}_1 \wedge \mathbf{T}_2 = (-1)^{pq} \mathbf{T}_2 \wedge \mathbf{T}_1$，这被称为外积的超交换性 (supercommutative)。

证明. (1) 由张量积的双线性性和交错化算子的线性性即证；

(2) $\mathbf{v} \wedge \mathbf{v} = \mathfrak{A}(\mathbf{v} \otimes \mathbf{v}) = \mathbf{v} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{v} = \mathbf{0}$;

(3) 注意到由外积的双线性性质，$\mathbf{0} = (\mathbf{v} + \mathbf{w}) \wedge (\mathbf{v} + \mathbf{w}) = \mathbf{v} \wedge \mathbf{v} + \mathbf{v} \wedge \mathbf{w} + \mathbf{w} \wedge \mathbf{v} + \mathbf{w} \wedge \mathbf{w}$，而 $\mathbf{v} \wedge \mathbf{v} = \mathbf{w} \wedge \mathbf{w} = \mathbf{0}$，因此 $\mathbf{v} \wedge \mathbf{w} = -\mathbf{w} \wedge \mathbf{v}$;

(4) 由外积运算的定义和结合律 (定理6.4.1) 即证；

(5) 由 (3) 归纳即证；

(6) 结合 (2) 和 (5) 即证；

(7) 由 (5) 即证。

□

定理 6.4.2. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})=0)$ 上的 n 维向量空间，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基。则

$$\left\{\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p} \mid 1\leqslant i_1<i_2<\cdots<i_p\leqslant n\right\}$$

是 $\Lambda^p(V)$ 的一组基，从而 $\Lambda^p(V)$ 的维数是组合数 $\binom{n}{p}=\dfrac{n!}{p!(n-p)!}$。

证明. 先证 $\Lambda^p(V)$ 中的任何一个向量都可以由 $\left\{\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p} \mid 1\leqslant i_1<i_2<\cdots<i_p\leqslant n\right\}$ 线性表出。任取 $\mathbf{T}\in\Lambda^p(V)$，不妨设

$$\mathbf{T}=\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}T^{j_1,\ldots,j_p}\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_p},$$

由于 $\mathbf{T}$ 是斜对称张量，因此 $\mathbf{T}=\mathfrak{A}(\mathbf{T})$，即

$$\begin{aligned}\mathbf{T}&=\mathfrak{A}(\mathbf{T})\\&=\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}T^{j_1,\ldots,j_p}\mathfrak{A}(\mathbf{e}_{j_1}\otimes\cdots\otimes\mathbf{e}_{j_p})\\&=\sum_{j_1=1}^{n}\cdots\sum_{j_p=1}^{n}T^{j_1,\ldots,j_p}\mathbf{e}_{j_1}\wedge\cdots\wedge\mathbf{e}_{j_p}.\qquad(\text{引理6.4.2(5)})\end{aligned}$$

而任意的 $\mathbf{e}_{j_1}\wedge\cdots\wedge\mathbf{e}_{j_p}$ 都可以通过交换外积的顺序 (需要改变符号) 变成 $\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p}$，$1\leqslant i_1<i_2<\cdots<i_p\leqslant n$ 的形式，因此所有的斜对称张量 $\mathbf{T}$ 都可以由 $\left\{\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p} \mid 1\leqslant i_1<i_2<\cdots<i_p\leqslant n\right\}$ 线性表出。

再证 $\left\{\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p} \mid 1\leqslant i_1<i_2<\cdots<i_p\leqslant n\right\}$ 线性无关。为此，设有线性组合

$$\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda_{i_1,\ldots,i_p}\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p}=\mathbf{0},\ \lambda_{i_1,\ldots,i_p}\in\mathbb{K},$$

我们只需要证明每个 $\lambda_{i_1,\ldots,i_p}$ 都是 0 即可得到线性无关性。为此，我们把外积用定义展开，得到

$$\begin{aligned}\mathbf{0}&=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda_{i_1,\ldots,i_p}\mathbf{e}_{i_1}\wedge\cdots\wedge\mathbf{e}_{i_p}\\&=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda_{i_1,\ldots,i_p}\mathfrak{A}(\mathbf{e}_{i_1}\otimes\cdots\otimes\mathbf{e}_{i_p})\\&=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda_{i_1,\ldots,i_p}\sum_{\pi\in S_p}\varepsilon_\pi\mathbf{e}_{i_{\pi(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\pi(p)}}\\&=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\sum_{\pi\in S_p}(\varepsilon_\pi\lambda_{i_1,\ldots,i_p})\cdot(\mathbf{e}_{i_{\pi(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\pi(p)}}).\end{aligned}$$

注意到最后得到的这个求和中所有的 $\mathbf{e}_{i_{\pi(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\pi(p)}}$，$\forall\pi\in S_p, 1\leqslant i_1<\cdots<i_p\leqslant n$ 实际上就是式 (6.1.4) 中 $V\otimes\cdots\otimes V(p$ 个$)$ 的基中的所有无重复的那些元素，因此

$$\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\sum_{\pi\in S_p}(\varepsilon_\pi\lambda_{i_1,\ldots,i_p})\cdot(\mathbf{e}_{i_{\pi(1)}}\otimes\cdots\otimes\mathbf{e}_{i_{\pi(p)}})=\mathbf{0}$$

会推出每个 $\varepsilon_\pi \lambda_{i_1,\dots,i_p}$ 都是 0，即 $\lambda_{i_1,\dots,i_p} = 0$。这样我们就完成了证明。 □

于是我们知道，当 $p = n$ 时，$\Lambda^n(V)$ 是一维向量空间，它由 $\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_n$ 张成；当 $p > n$ 时，由于 $\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}$ 中一定会有重复的向量，因此所有的 $\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}$ 都是 $\mathbf{0}$，即 $p > n$ 时 $\Lambda^p(V) = \{\mathbf{0}\}$。我们还得到了 $\Lambda(V)$ 的维数：

$$\dim(\Lambda(V)) = \sum_{p=0}^{n} \binom{n}{p} = 2^n.$$

由定理6.4.2我们还可以得到如下结论：

定理 6.4.3. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) = 0)$ 上的 n 维向量空间，$\mathbf{v}_1, \dots, \mathbf{v}_k \in V$。则 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 线性相关 $\iff$ $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k = \mathbf{0}$。

证明. ($\Longrightarrow$) 如果 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 线性相关，那么不妨设 $\mathbf{v}_1$ 能被其它向量线性表出，即

$$\mathbf{v}_1 = \sum_{j=2}^{k} \lambda_j \mathbf{v}_j,\ \lambda_j \in \mathbb{K}.$$

于是

$$\begin{aligned}
\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k &= (\sum_{j=2}^{k} \lambda_j \mathbf{v}_j) \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_k \\
&= \sum_{j=2}^{k} \lambda_j (\mathbf{v}_j \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_k) \\
&= \underbrace{\mathbf{0} + \cdots + \mathbf{0}}_{k-1\text{个}} \qquad\qquad = \mathbf{0}.
\end{aligned}$$

($\Longleftarrow$) 反证法，如果 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 线性无关，那么可以将其扩充成 V 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$，由定理6.4.2可知 $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k \wedge \mathbf{v}_{k+1} \wedge \cdots \wedge \mathbf{v}_n$ 是 $\Lambda^n(V)$ 的基，因此它不为 $\mathbf{0}$，所以它的一部分 $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k$ 也不为零，这与 $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k = \mathbf{0}$ 矛盾！这样我们就完成了证明。 □

关于外积我们还有以下结论。

定理 6.4.4 (Cartan 引理). 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) = 0)$ 上的 n 维向量空间，$\mathbf{v}_1, \dots, \mathbf{v}_k$ 和 $\mathbf{w}_1, \dots, \mathbf{w}_k$ 是 V 中的两组向量，并且 $\sum_{i=1}^{k} (\mathbf{v}_i \wedge \mathbf{w}_i) = 0$。如果我们还有 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 线性无关，那么每个 $\mathbf{w}_i$，$i = 1, \dots, k$ 都可以写成线性组合

$$\mathbf{w}_i = \sum_{j=1}^{k} a_{ij} \mathbf{v}_j,\ a_{ij} \in \mathbb{K}$$

的形式，并且其中 $a_{ij} = a_{ji}$。

证明. 由于 $\mathbf{v}_1, \dots, \mathbf{v}_k$ 线性无关，因此可以将其扩充成 V 的一组基 $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$。那么，我们可以将每个 $\mathbf{w}_i$ 表示成 $\mathbf{v}_1, \dots, \mathbf{v}_n$ 的线性组合，即

$$\mathbf{w}_i = \sum_{j=1}^{n} a_{ij} \mathbf{v}_j,\ a_{ij} \in \mathbb{K},\ i = 1, \dots, k.$$

将其代入 $\sum_{i=1}^{k}(\mathbf{v}_i \wedge \mathbf{w}_i) = 0$ 中可得

$$\sum_{i=1}^{k}\left(\mathbf{v}_i \wedge \left(\sum_{j=1}^{n} a_{ij}\mathbf{v}_j\right)\right) = \mathbf{0},$$

整理即得

$$\sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}\mathbf{v}_i \wedge \mathbf{v}_j + \sum_{i=1}^{k}\sum_{j=k+1}^{n} a_{ij}\mathbf{v}_i \wedge \mathbf{v}_j = \mathbf{0}.$$

注意到 i, j 都在 $\{1, \ldots, k\}$ 里时，$\mathbf{v}_i \wedge \mathbf{v}_j = -\mathbf{v}_j \wedge \mathbf{v}_i$，所以上式可化为

$$\sum_{1\leqslant i<j\leqslant k}(a_{ij} - a_{ji})\mathbf{v}_i \wedge \mathbf{v}_j + \sum_{i=1}^{k}\sum_{j=k+1}^{n} a_{ij}\mathbf{v}_i \wedge \mathbf{v}_j = \mathbf{0}.$$

由于 $\mathbf{v}_i \wedge \mathbf{v}_j,\ 1 \leqslant i < j \leqslant n$ 是 $\Lambda^2(V)$ 的一组基，即上式中的所有 $\mathbf{v}_i \wedge \mathbf{v}_j$ 线性无关，因此上式中的所有系数都等于 0，即 $a_{ij} = a_{ji},\ \forall i, j \in \{1, \ldots, k\}$ 且 $a_{ij} = 0,\ \forall j \in \{k+1, \ldots, n\}$。所以 $\mathbf{w}_i = \sum_{j=1}^{k} a_{ij}\mathbf{v}_j$，这样我们就完成了证明。 □

下面我们考虑向量做外积的坐标形式。

定理 6.4.5. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) = 0)$ 上的 n 维向量空间，$\mathbf{e}_1, \ldots \mathbf{e}_n$ 是 V 的一组基。则以下 m 个向量

$$\mathbf{x}_i = \sum_{j=1}^{n} a_{ij}\mathbf{e}_j \in V,\ i = 1, \ldots, m,\ a_{ij} \in \mathbb{K}$$

满足

$$\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m = \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \mathit{\Delta}_{i_1,\ldots,i_m}\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_m},$$

其中

$$\mathit{\Delta}_{i_1,\ldots,i_m} = \det\begin{pmatrix} a_{1,i_1} & \cdots & a_{1,i_m} \\ \vdots & & \vdots \\ a_{m,i_1} & \cdots & a_{m,i_m} \end{pmatrix}.$$

证明. 利用外积的斜对称性和多重线性性展开 $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m$，我们有：

$$\begin{aligned}
\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m &= \sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} a_{1,i_1}a_{2,i_2}\cdots a_{m,i_m}\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \\
&= \sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\ \sum_{\pi\in S_m} a_{1,i_{\pi(1)}}a_{2,i_{\pi(2)}}\cdots a_{m,i_{\pi(m)}}\mathbf{e}_{i_{\pi(1)}} \wedge \mathbf{e}_{i_{\pi(2)}} \wedge \cdots \wedge \mathbf{e}_{i_{\pi(m)}} \\
&= \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \underbrace{(\sum_{\pi\in S_m} \varepsilon_\pi a_{1,i_{\pi(1)}}a_{2,i_{\pi(2)}}\cdots a_{m,i_{\pi(m)}})}_{=\mathit{\Delta}_{i_1,\ldots,i_m}}\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \\
&= \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \mathit{\Delta}_{i_1,\ldots,i_m}\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_m}.
\end{aligned}$$

□

下面我们简单介绍一下外代数中的可分解向量。

定义 6.4.4. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}) = 0)$ 上的 n 维向量空间，$\mathbf{T} \in \Lambda^m(V)$。如果存在 $\mathbf{x}_1, \ldots, \mathbf{x}_m \in V$ 使得 $\mathbf{T} = \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m$，则称 $\mathbf{T}$ 是可分解的。

定义 6.4.5. 条件同定义6.4.4。对 $\mathbf{T}$ 而言，容易验证

$$\mathrm{Ann}(\mathbf{T})=\{\mathbf{x}\in V\mid \mathbf{T}\wedge\mathbf{x}=\mathbf{0}\}$$

是 V 的一个子空间，称为 $\mathbf{T}$ 的零化子 (annihilator)。

我们有以下结论：

定理 6.4.6. 条件同定义6.4.4。记 $t=\dim(\mathrm{Ann}(\mathbf{T}))$，则 $t\leqslant m$，并且存在 $\mathbf{x}_1,\dots,\mathbf{x}_t\in\mathrm{Ann}(\mathbf{T})$ 及 $\mathbf{x}\in\Lambda^{m-t}(V)$ 使得 $\mathbf{T}=\mathbf{x}_1\wedge\cdots\wedge\mathbf{x}_t\wedge\mathbf{x}$。特别地，$t=m\iff\mathbf{T}$ 可分解。

定理 6.4.7. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})=0)$ 上的 n 维向量空间，则 $\Lambda^{n-1}(V)$ 中的向量是可分解的。

定理 6.4.8. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K})=0)$ 上的 n 维向量空间，$\mathbf{T}\in\Lambda^2(V)$，则 $\mathbf{T}$ 可分解 $\iff$ $\mathbf{T}\wedge\mathbf{T}=\mathbf{0}$。

这些定理的证明可以在代数学引论中找到，由于它们与我们的主线关系不大，故略去。

我们把 $\Lambda^2(V)$ 中所有的可分解向量的全体称为 Grassmann 锥，其射影化被称为 Grassmann 簇。它们是微分几何的重要研究对象。

在本节的最后，我们来说明对称张量代数和外代数所具有的泛性质。

定理 6.4.9. 设 V 是域 $\mathbb{K}(\mathrm{char}(\mathbb{K}=0))$ 上的 n 维向量空间，则存在多重线性映射

$$\tau:\ \underbrace{V\times\cdots\times V}_{p\text{个}}\longrightarrow\mathbb{T}_0^p(V)^+$$

使得以下结论成立：

任取域 $\mathbb{K}$ 上的向量空间 W 和 $\underbrace{V\times\cdots\times V}_{p\text{个}}$ 映到 W 的**对称的**多重线性映射 f，都存在唯一的线性映射 $\varphi_f:\ \mathbb{T}_0^p(V)^+\to W$ 使得 $f=\varphi_f\circ\tau$。也就是说，下面的交换图表成立：

$$\begin{tikzcd}
V\times\cdots\times V \arrow[r, "\tau"] \arrow[d, "f"'] & \mathbb{T}_0^p(V)^+ \arrow[dl, "\varphi_f"] \\
W &
\end{tikzcd}$$

同理，存在多重线性映射：

$$\tau':\ \underbrace{V\times\cdots\times V}_{p\text{个}}\longrightarrow\Lambda^p(V)$$

使得以下结论成立：

任取域 $\mathbb{K}$ 上的向量空间 W 和 $\underbrace{V\times\cdots\times V}_{p\text{个}}$ 映到 W 的**斜对称的**多重线性映射 f，都存在唯一的斜对称线性映射 $\varphi_f:\ \Lambda^p(V)\to W$ 使得 $f=\varphi_f\circ\tau'$。也就是说，下面的交换

图表成立：

$$\begin{tikzcd}
V\times\cdots\times V \arrow[r, "\tau'"] \arrow[d, "f"'] & \Lambda^p(V) \arrow[dl, "\varphi_f"] \\
W &
\end{tikzcd}$$

证明. 只需构造出：

$$\tau:\ (\mathbf{v}_1,\dots,\mathbf{v}_p)\mapsto \mathbf{v}_1\vee\cdots\vee\mathbf{v}_p$$

和

$$\tau':\ (\mathbf{v}_1,\dots,\mathbf{v}_p)\mapsto \mathbf{v}_1\wedge\cdots\wedge\mathbf{v}_p$$

即可，其余步骤与定理6.1.1的证明类似。细节留作练习。 □

到此为止，张量的基础内容我们已经介绍完了，有关张量的更深入的内容，可以参考抽象代数或微分几何教材中的相关部分，或者专门的张量分析的教材。

6.5 习题

张量计算初步, 混合张量

1. 假设 $\mathbf{T}\in\mathbb{T}^2_3(V)$ 的坐标都是 2，求 $\mathbf{T}(\mathbf{v},\mathbf{v},\mathbf{v},\alpha,\alpha)$, 其中 $\mathbf{v}=\mathbf{e}_1+2\mathbf{e}_2-3\mathbf{e}_3+\mathbf{e}_4,\alpha=\mathbf{e}^1-\mathbf{e}^4$.

2. 假设张量 $\mathbf{T}\in\mathbb{T}^2_3(V)$ 在基 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 下的坐标都是 3，求它在基

$$(\widetilde{\mathbf{e}_1},\widetilde{\mathbf{e}_2},\widetilde{\mathbf{e}_3})=(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}$$

下的坐标 $\tilde{T}^{12}_{123}$.

3. 求张量的坐标: (1) $(\mathbf{e}_1+\mathbf{e}_2)\otimes(\mathbf{e}_1-2\mathbf{e}_2)$;
   (2) $(\mathbf{e}_1+2\mathbf{e}_2)\otimes(\mathbf{e}_3-2\mathbf{e}_4)-(\mathbf{e}_1-\mathbf{e}_2)\otimes(\mathbf{e}_3+\mathbf{e}_4)$.

4. 假设 $\dim V=4,\mathbf{T}=\mathbf{e}^1\otimes\mathbf{e}_2+\mathbf{e}^2\otimes\mathbf{e}_3+\mathbf{e}^3\otimes\mathbf{e}_4\in\mathbb{T}^1_1(V)$. 求出所有的
   (1) $\alpha\in V^*$ 使得 $T(\mathbf{v},\alpha)=0,\forall\mathbf{v}\in V$;
   (2) $\mathbf{v}\in V$ 使得 $T(\mathbf{v},\alpha)=0,\forall\alpha\in V^*$.

5. 证明: 如果 $\mathscr{A}$ 是可对角化的, 那么 $\mathscr{A}\otimes\mathscr{A}$ 也是可对角化的.

6. 假设空间 V 是有限维的. 证明 $\left(T^q_p(V)\right)^*$ 与 $\mathbb{T}^p_q(V)$ 同构.

张量的收缩, 对称化, 交错化; 张量代数, 对称代数, 外代数

1. 假设 $\dim V>1$. 证明向量空间 $\mathbb{T}^q_0(V)$ 上的对称化算子和斜对称化算子有如下性质:
   (1) 它们的核的交等于零如果 $q=2$, 不等于零如果 $q>2$;
   (2) $\mathfrak{S}\cdot\mathfrak{A}=\mathfrak{A}\cdot\mathfrak{S}=0$;
   (3) 算子 $\mathcal{P}=(\mathcal{E}-\mathfrak{S})(\mathcal{E}-\mathfrak{A})$ 是投影算子;
   (4) 如果 $q=3$ 且 V 是有限维空间, 求 $\mathcal{P}$ 的秩.

2. 设 $\mathcal{A}$ 是 n 维向量空间 V 上的线性算子. 证明:

(1) $\mathbb{T}_0^2(V)^+ \subset V \otimes V$ 是 $\mathcal{A} \otimes \mathcal{A}$ 的不变子空间. 算子 $\mathcal{A} \otimes \mathcal{A}$ 在 $\mathbb{T}_0^2(V)^+$ 上的限制记作 $\mathfrak{S}^2\mathcal{A}$.

(2) 如果 $\mathcal{A}$ 的特征多项式的根是 $\lambda_1, \cdots, \lambda_n$ (按重数计算), 那么 $\mathfrak{S}^2\mathcal{A}$ 的特征多项式的根是 $\lambda_i\lambda_j, 1 \leqslant i \leqslant j \leqslant n$.

(3) $\operatorname{tr} \mathfrak{S}^2\mathcal{A} = \dfrac{1}{2}\left[(\operatorname{tr}\mathcal{A})^2 + \operatorname{tr}\mathcal{A}^2\right]$.

3. 设 V 和 W 是向量空间. 证明 $\mathbb{T}_0^q(V \oplus W)^+$ 与 $\bigoplus_{i=0}^q \left(\mathbb{T}_0^i(V)^+ \otimes \mathbb{T}_0^{q-i}(W)^+\right)$ 同构.

4. 假设 $\dim V > 1$. 证明: $\mathbb{T}(V)$ 是无零因子的非交换环, 其中的可逆元只有非零纯量.

5. 设 $V = \mathbb{K}^n$ 是列坐标空间, $\mathbf{A}_1, \cdots, \mathbf{A}_n$ 是 V 的基, $\mathbf{B}$ 是 V 中任意的元素. 证明: 向量方程式

$$\sum_k \lambda_k \mathbf{A}_k = \mathbf{B}$$

的解参数 λ_k 可由关系式

$$\lambda_k (\mathbf{A}_1 \wedge \cdots \wedge \mathbf{A}_n) = \mathbf{A}_1 \wedge \cdots \wedge \mathbf{A}_{k-1} \wedge \mathbf{B} \wedge \mathbf{A}_{k+1} \wedge \cdots \wedge \mathbf{A}_n$$

给出. 试由此导出 Cramer 法则 (讲义上册定理 3.3.2).

后记

这段话会在讲义出版时重写。

时隔两年，编者不负所托，终于完成了这套线性代数讲义的初稿编写 (在此撒花庆祝！)，希望对大家有所帮助。之后编者会补出所有之前为了赶进度而省略的部分，以及对上下册讲义进行彻底的校对与修订，以及以后有机会的话在编者的手稿基础上另写一套习题课讲义，在此过程中欢迎大家随时提出建议。

下面是关于讲义一些细节处的说明。这套讲义是作者和编者的心血之作，它以作者的手稿整理而成，在此过程中，编者对部分内容进行了调整，对符号则进行了统一化，与大部分教材的通用符号保持了一致。讲义中省略的细节，如果标注了“显然”或“容易看出”，那么它一定是十分显然的，并且在以后也不会补充 (因为真的不需要)；如果标注了“容易验证”或“留作练习”，则这部分细节通常需要按照定义进行一定的计算或推导，是编者认为可以独立完成，并且不给出具体过程也可以继续阅读的，这部分细节以后会视具体情况补充细化一部分；如果标注了“留作思考”或者给出了参考文献，则这部分内容通常较难，我们会在完稿后补充出留作思考的部分，至于引用参考文献的内容，我们会补充一些在本科低年级阶段能够接受的内容，更高级的内容感兴趣的同学可以直接阅读相关文献 (毕竟我们只是线性代数的讲义)。

这套讲义的风格是偏“硬”、偏具体的，但并没有回避抽象的内容，而是以直观暴力的推导将抽象的结构拆开，展示给大家看，比如若尔当标准型的证明、谱定理的证明、二次曲面和射影二次曲面的分类、张量的泛性质与坐标变换等内容，讲义中都是不避繁琐，直面问题，将精细的推导过程和数学结构展示给大家。希望大家在学习数学时也要有这种精神，既要学会四两拨千斤的技巧，又要有迎难而上、勇攀高峰的精神 (“所谓重剑无锋，大巧不工”，有时直接算就是最好用的技巧)。同时，大家在阅读讲义时要重视对结构的思考，能够熟练地进行抽象到具体的转换，能够掌握理论并且熟练计算。

这套讲义的手稿以柯斯特利金《代数学引论》和席南华《基础代数》为蓝本，作者和编者在讲义的编写中参考了诸多线性代数和抽象代数的教材，其中主要有《高等代数》丘维声、《高等代数简明教程》蓝以中、《Advanced Linear Algebra》GTM135、《Algebra》GTM73、《Algebra》GTM211、《Linear Algebra and Geometry》沙法列维奇 (英译本)、《抽象代数》邓少强，朱富海、《高等代数学》姚慕生，吴泉水，谢启鸿、《Linear Algebra Done Right》Axler, Sheldon Jay、《高等线性代数学》黎景辉，白正简，周国晖、《高等代数学》张贤科、《高等代数》王萼芳，等等。我们不敢称自己博采众长，但至少避开了一些大家吐槽比较大的点，在一些细节上进行了精心的设计。当然每套教材都有其独特之处，但知识内容总体上是一样的，希望大家可以读一而知百。

祝愿读者学业有成！

编者禹天石，2023.11.06.