线性代数教案 (II)

支丽红 (主讲)　禹天石　梁昊 (助教)

2022 年 4 月 21 日

目录

Chapter 1

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 $\lambda\mathbf{x} = \mathbf{0} \nRightarrow \lambda = 0$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 $\mathrm{char}(\mathbb{K})=0$，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 $a=\underbrace{1_{\mathbb{K}}+\cdots+1_{\mathbb{K}}}_{m\text{个}}\in\mathbb{K}$，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

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

另一方面

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

这与数乘的结合律矛盾！

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

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

另一方面

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

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 $\mathrm{span}\,\{\mathbf{x}_1,\ldots\mathbf{x}_n\}$。

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

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

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 $\bigcap_{i\in I} V_i$ 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 $\mathbf{x}, \mathbf{y} \in \bigcap_{i\in I} V_i$，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 $\lambda \mathbf{x} + \mu \mathbf{y} \in \bigcap_{i\in I} V_i$，即 $\bigcap_{i\in I} V_i$ 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 $U_1,U_2,\ldots,U_n$ 的直和，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，$\mathrm{char}(\mathbb{K})\neq 2$。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 $\mathrm{char}(\mathbb{K})\neq 2$，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 $U=U_1\oplus U_n$ 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。

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。

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

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 $\forall x\in\mathbb{R}$, $\lambda_1e^{\alpha_1x}+\cdots+\lambda_ne^{\alpha_nx}=\mathbf{0}(x)\Longrightarrow\lambda_1=\cdots=\lambda_n=0$ 即可。

在 $\lambda_1e^{\alpha_1x}+\cdots+\lambda_ne^{\alpha_nx}=\mathbf{0}(x)$ 中依次取 $x=0,1,\ldots,n-1$ 即有

$$\begin{cases}\lambda_1e^0+\lambda_2e^0+\cdots+\lambda_ne^0=0\\ \lambda_1e^{\alpha_1}+\lambda_2e^{\alpha_2}+\cdots+\lambda_ne^{\alpha_n}=0\\ \qquad\vdots\\ \lambda_1e^{(n-1)\alpha_1}+\lambda_2e^{(n-1)\alpha_2}+\cdots+\lambda_ne^{(n-1)\alpha_n}=0\end{cases}$$

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 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ 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，$\mathbf{e}_1, \ldots, \mathbf{e}_s, \mathbf{f}_1, \ldots, \mathbf{f}_t \in V$，其中 $\mathbf{e}_1, \ldots, \mathbf{e}_s$ 线性无关且每个 $\mathbf{e}_i$ 都是 $\mathbf{f}_1, \ldots, \mathbf{f}_t$ 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 $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$，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 $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$ 的方程组

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

有非零解，即存在不全为 0 的 $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$ 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 $\mathbf{e}_1, \ldots, \mathbf{e}_s$ 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 $\mathbb{K}[x_1, \dots, x_n]$ 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 $\{x_1^{s_1} \cdots x_n^{s_n} \mid s_1, \dots, s_n \in \mathbb{N}且 s_1 + \cdots + s_n = m\}$，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 $\{\mathbf{e}_1 = (1, 0, 0, \dots, 0)^t,\ \mathbf{e}_2 = (0, 1, 0, \dots, 0)^t, \cdots\cdots,\ \mathbf{e}_n = (0, 0, \dots, 0, 1)^t\}$，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 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 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 $\mathbf{f}_1, \dots, \mathbf{f}_s \in V$ 线性无关，则存在 $\mathbf{f}_{s+1}, \dots, \mathbf{f}_n \in V$ 使得 $\mathbf{f}_1, \dots, \mathbf{f}_s, \mathbf{f}_{s+1}, \cdots, \mathbf{f}_n$ 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 $U \subsetneq W \iff \dim(U) < \dim(W)$。

证明: (1) 由于 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组基，故对任意 $\mathbf{x} \in V$，我们有 $\mathbf{x}, \mathbf{e}_1, \dots, \mathbf{e}_n$ 线性相关，即存在不全为 0 的 $\lambda, \lambda_1, \dots, \lambda_n \in \mathbb{K}$ 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 $\lambda_1, \dots, \lambda_n$ 不全为 0 且 $\lambda_1 \mathbf{e}_1 + \cdots + \lambda_n \mathbf{e}_n = \mathbf{0}$，这说明 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 线性相关，与 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 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 $\beta_1, \dots, \beta_n \in \mathbb{K}$ 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 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组基可知 $\forall i = 1, \dots, n,\ \alpha_i - \beta_i = 0$，即 $\alpha_i = \beta_i$，这就证明了唯一性。
(2) 作 $V_1 = \operatorname{span}\{\mathbf{f}_1, \dots, \mathbf{f}_s\}$ 是 V 的子空间，则 $\dim(V_1) = s$。如果 $s = n$，则 $\mathbf{f}_1, \dots, \mathbf{f}_s$ 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 $\mathbf{f}_1, \dots, \mathbf{f}_s, \mathbf{f}_{s+1}$ 线性无关 (否则 $\mathbf{f}_{s+1} \in V_1$)。令 $V_2 = \operatorname{span}\{\mathbf{f}_1, \dots, \mathbf{f}_s, \mathbf{f}_{s+1}\}$，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 $\mathbf{f}_1, \dots, \mathbf{f}_s, \mathbf{f}_{s+1}, \mathbf{f}_{s+2}$ 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，$U_1, \ldots, U_m$ 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 $\mathrm{codim}(U)=\dim(W)$。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 $f(x)=a_0+a_1x+\cdots+a_nx^n$ 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 $(f(a),f'(a),\ldots,\dfrac{f^{(n)}(a)}{n!},0,0,\ldots)^t$。

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。

容易看到过渡矩阵一定是可逆矩阵 (思考之)。

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

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

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 $\mathrm{im}(f) = \{\mathbf{y}\in W \mid \exists\mathbf{x}\in V$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 $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$，那么对 $\forall\mathbf{v}\in V$，$\mathbf{v}$ 有一组坐标 $(\lambda_1,\cdots,\lambda_n)^t$，即

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

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

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

[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。

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，$V = \bigcup_{\mathbf{x} \in V} (\mathbf{x} + U)$，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。 □

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 定理 1.5.2对无穷维空间不成立。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，因此它们之间不可能有同构。

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

首先，如果 $\mathbf{x}_1,\dots\mathbf{x}_r$ 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，$i=1,\dots,n$，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 $\mathbf{A}^{(1)},\dots,\mathbf{A}^{(r)}$ 线性相关，矛盾！故 $\mathbf{x}_1,\dots\mathbf{x}_r$ 线性无关。

其次，由于 $\mathrm{rank}(A)=r$，则对任意的 $j=r+1,\dots,m$，都有 $\mathbf{A}^{(j)}\in\mathrm{span}\,\{\mathbf{A}^{(1)},\dots,\mathbf{A}^{(r)}\}$，即存在 $\alpha_1,\dots,\alpha_r\in\mathbb{K}$ 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，$i=1,\dots,n$，都有

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

注意到 $f_1,\dots,f_n$ 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 $\forall j=r+1,\dots,m$，$\mathbf{x}_j$ 都能够写成 $\mathbf{x}_1,\dots\mathbf{x}_r$ 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 $\varepsilon_1,\dots,\varepsilon_n$，作其对偶基 $\varepsilon^1,\dots,\varepsilon^n$，则 $\varepsilon^1,\dots,\varepsilon^n$ 与 $f_1,\dots,f_n$ 之间有基变换 $(f_1,\dots,f_n)=(\varepsilon^1,\dots,\varepsilon^n)\cdot A$，$A$ 是可逆矩阵，作 V 上的基变换 $(\mathbf{e}_1,\dots,\mathbf{e}_n)=(\varepsilon_1,\dots,\varepsilon_n)\cdot(A^t)^{-1}$，容易验证 $f_1,\dots,f_n$ 是 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 的对偶基 (细节留作练习)。

对于 (2)，我们利用 (1) 的结论，取 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得 $f_1,\dots,f_n$ 是 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 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 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是域 $\mathbb{K}$ 上的 n 维向量空间 V 的一组基，$f_1,\dots,f_m\in V^*$，则 $\dim_{\mathbb{K}}(\mathrm{span}\,\{f_1,\dots,f_m\})=\mathrm{rank}(B)$，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 $\dim(V) = n$，$\dim(U) = m$，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 $\mathbf{e}_j \in U$，$j \in \{1,\ldots,k\}$，都有 $f(\mathbf{e}_j) = \alpha_j = 0$，即 $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$，所以 $f \in \mathrm{span}\,\{\mathbf{e}^{k+1},\ldots,\mathbf{e}^n\}$，即 $\mathrm{span}\,\{\mathbf{e}^{k+1},\ldots,\mathbf{e}^n\} \supset U^0$。综上所述，$\mathrm{span}\,\{\mathbf{e}^{k+1},\ldots,\mathbf{e}^n\} = U^0$。这样我们就完成了证明。 □

例 1.5.5 设 $V = \mathbb{R}^2$，$U = \mathrm{span}\,\{(2,3)^t\}$，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 $f(\mathbf{x}) = \alpha_1x_1 + \alpha_2x_2$，$\alpha_1,\alpha_2 \in \mathbb{R}$ 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 $U^0 = \mathrm{span}\,\{-3x_1 + 2x_2\}$。

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，$f_1,\ldots,f_m \in V^*$，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 $U = \{\mathbf{x} \in V \mid \forall i = 1, \ldots, m,\ f_i(\mathbf{x}) = 0\}$(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 $r = \dim(\mathrm{span}\,\{f_1, \ldots, f_m\})$，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 $f_1, \ldots, f_r$ 在 V^* 中线性无关，则由基扩充定理，存在一组 $\varphi_1, \ldots, \varphi_{n-r} \in V^*$ 使得 $f_1, \ldots, f_r, \varphi_1, \ldots, \varphi_{n-r}$ 是 V^* 的一组基。由定理 1.5.3(1)，存在 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 的对偶基恰为 $f_1, \ldots, f_r, \varphi_1, \ldots, \varphi_{n-r}$。那么，显然有 $\mathrm{span}\,\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ 是方程组 (H) 的解空间，即 $\dim(U) = n - r$。
(2) 取 U 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_r$，由基扩充定理，存在 $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n \in V$ 使得 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组基。取这组基的对偶基 $\mathbf{e}^1, \ldots, \mathbf{e}^n \in V^*$，则显然方程组

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

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 $V_1,\ldots,V_m,W$ 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 $i\in\{1,\ldots,m\}$，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

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

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 $f: V_1\times\cdots\times V_m\to\mathbb{K}$, $(x_1,\ldots,x_m)\mapsto x_1x_2\cdots x_m$，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 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组基，$\mathbf{e}^1,\ldots,\mathbf{e}^n$ 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

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

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 $\Phi(f) = (f(\mathbf{e}_i,\mathbf{e}_j))_{n\times n}$ 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

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

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，$\Phi(f)$ 就是这个函数的“参数”(或“系数”)。

现在我们可以证明 Φ 是同构了。首先 Φ 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

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

上式对任意 $(x'_1,\dots,x'_n)$ 和 $(y'_1,\dots,y'_n)$ 都成立，于是我们让 $(x'_1,\dots,x'_n)$ 和 $(y'_1,\dots,y'_n)$ 分别取遍 $(1,0,\dots,0),(0,1,\dots,0),\dots,(0,\dots,0,1)$ 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

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

即

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

此外，容易求出 $\mathrm{rank}(f)=\mathrm{rank}(A)=3$，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 $\Longleftrightarrow$ $\Phi(f)$ 是对称矩阵；f 是斜对称双线性型 $\Longleftrightarrow$ $\Phi(f)$ 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 $\Longleftrightarrow$ $\forall\mathbf{e}_i,\mathbf{e}_j$，有 $f(\mathbf{e}_i,\mathbf{e}_j)=f(\mathbf{e}_j,\mathbf{e}_i)$ $\Longleftrightarrow$ $\Phi(f)=\Phi(f)^t$；f 是斜对称双线性型 $\Longleftrightarrow$ $\forall\mathbf{e}_i,\mathbf{e}_j$，有 $f(\mathbf{e}_i,\mathbf{e}_j)=-f(\mathbf{e}_j,\mathbf{e}_i)$ $\Longleftrightarrow$ $\Phi(f)=-\Phi(f)^t$。□

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 $\mathrm{char}(\mathbb{K})\neq 2$，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 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

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

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 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 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 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下一定有

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

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 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 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

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

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，$\dim(U \oplus \mathrm{span}\,\{\mathbf{e}_1\}) = n = \dim(V)$，即 $V = U \oplus \mathrm{span}\,\{\mathbf{e}_1\}$。

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

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

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 $f(\mathbf{v}_1,\mathbf{v}_1)=2\neq0$。

(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 $\mathbf{x}=(x_1,x_2,x_3)^t\in\mathbb{R}^3$，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 $\mathbf{u}_1=(1,-1,0)^t$，$\mathbf{u}_2=(2,0,-1)^t$。

[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。

(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

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

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

- 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：

$$A \xrightarrow{r_1+r_2} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow{c_1+c_2} \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow{r_2-\frac{1}{2}r_1} \begin{pmatrix} 2 & 1 & 2 \\ 0 & -\frac{1}{2} & 0 \\ 2 & 1 & 0 \end{pmatrix} \xrightarrow{c_2-\frac{1}{2}c_1} \begin{pmatrix} 2 & 0 & 2 \\ 0 & -\frac{1}{2} & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{r_3-r_1} \begin{pmatrix} 2 & 0 & 2 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -2 \end{pmatrix} \xrightarrow{c_3-c_1} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

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)：

$$(A|E) = \left(\begin{array}{ccc:ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{左乘}F_{1,2}(1)} \left(\begin{array}{ccc:ccc} 1 & 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{右乘}F_{2,1}(1)} \left(\begin{array}{ccc:ccc} 2 & 1 & 2 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{\text{左乘}F_{2,1}(-\frac{1}{2})} \left(\begin{array}{ccc:ccc} 2 & 1 & 2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{右乘}F_{1,2}(-\frac{1}{2})} \left(\begin{array}{ccc:ccc} 2 & 0 & 2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

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

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，我们将在下一节中进行介绍。

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 $f(\mathbf{0},\mathbf{0})=f(\mathbf{0},-\mathbf{0})\implies f(\mathbf{0},\mathbf{0})=0$(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

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

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 $A \in M_n(\mathbb{C})$, $\mathrm{rank}(A) = r$ 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 $A \in M_n(\mathbb{R})$, $\mathrm{rank}(A) = r$ 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

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

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 $\text{rank}(q) = n$，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，$\text{rank}(q) = r$，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 $\mathbf{e}_1'=\frac{1}{a_{11}}\mathbf{e}_1$，则 f 在基 $\mathbf{e}_1'$ 下的矩阵为 $(\frac{1}{a_{11}})\cdot(a_{11})\cdot(\frac{1}{a_{11}})=(\frac{1}{a_{11}})$。由于 $\Delta_0=1$，$\Delta_1=a_{11}$，故 $\frac{1}{a_{11}}=\frac{\Delta_0}{\Delta_1}$ 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 $V_1=\operatorname{span}\{\mathbf{e}_1,\ldots,\mathbf{e}_{n-1}\}$，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

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

则由于 $\dim(\operatorname{span}\{f_1,\ldots,f_{n-1}\})\leqslant n-1$，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 $\mathbf{e}_1',\ldots,\mathbf{e}_{n-1}'$ 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 $\alpha_1,\ldots,\alpha_{n-1}$ 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 $\mathbf{e}_1',\ldots,\mathbf{e}_{n-1}',\lambda\mathbf{x}_0$ 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 $\mathbf{e}'_1,\ldots,\mathbf{e}'_{n-1},\lambda\mathbf{x}_0$ 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 $\mathbf{e}_1,\ldots,\mathbf{e}_{n-1}$ 到 $\mathbf{e}'_1,\ldots,\mathbf{e}'_{n-1}$ 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

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

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

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

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 $A'=T_\lambda^tAT_\lambda$ 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 $\mathbf{e}'_1,\ldots,\mathbf{e}'_{n-1},\mathbf{e}'_n$ 就满足 ($*$) 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 的正惯性指数等于序列 $1,\Delta_1,\ldots,\Delta_n$ 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 $\iff$ 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 $\iff$ $\forall 1\leqslant i_1<i_2<\cdots<i_m\leqslant n$，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 $\mathbf{x}^tA_m\mathbf{x}\geqslant 0$。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 $\widetilde{\mathbf{x}}^tA\widetilde{\mathbf{x}}\geqslant 0$，而 $\mathbf{x}^tA_m\mathbf{x}=\widetilde{\mathbf{x}}^tA\widetilde{\mathbf{x}}$，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 $A_m=C^tC$，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 $\Delta_m=t^m+p_{m-1}t^{m-1}+\cdots+p_1t+p_0$。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。

[1]这个结论我们在下一章还会用到。

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 $f\big|_{V_1}$ 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 $\exists \beta \in V$ 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 $\mathrm{span}\,\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$ 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

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

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 $\mathbb{K}=\mathbb{Q}(t_{12},\ldots,t_{1n},\ldots\ldots,t_{(n-1)n})$。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 $\det(T)\in\mathbb{Q}[t_{12},\ldots,t_{1n},\ldots\ldots,t_{(n-1)n}]$。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

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

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 $\mathrm{Pf}(t_{12},\ldots,t_{(n-1)n})$ 或 $\mathrm{Pf}(\mathbf{t})$。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：

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

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 $\mathrm{Pf}_4(\mathbf{t})=t_{12}t_{34}-t_{13}t_{24}+t_{14}t_{23}$ 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 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。