Chapter 3

内积空间及其上的线性算子

这一章我们将会回到 $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 上来讨论问题。前面我们已经把一般的域 $\mathbb{K}$ 上的双线性型和线性算子的性质比较完整地讲清楚了，但是如果落回到 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 上，与通常的空间相比 (即我们中学学过的 $\mathbb{R}^3$ 或数学分析中的 $\mathbb{R}^n$)，我们会发现前面的内容里少了一些结构：夹角和距离。这就是我们这一章的主要任务：在带有夹角 (即内积) 或距离的向量空间上进一步讨论双线性型和线性算子的性质。对应到矩阵层面，我们则是讨论一个实 (或复) 矩阵何时能够合同并相似到一个对角矩阵。下面我们正式开始介绍本章的内容。

3.1 欧几里得空间

定义 3.1.1 设 V 是实数域 $\mathbb{R}$ 上的 n 维向量空间，我们在 $V\times V$ 上指定一个满足如下性质的函数 $<\cdot,\cdot>: V\times V\to\mathbb{R}$：

1) $\forall \mathbf{x},\mathbf{y}\in V,\ \ <\mathbf{x},\mathbf{y}>=<\mathbf{y},\mathbf{x}>$；

2) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V,\ \ <\mathbf{x}+\mathbf{y},\mathbf{z}>=<\mathbf{x},\mathbf{z}>+<\mathbf{y},\mathbf{z}>$；

3) $\forall \mathbf{x},\mathbf{y}\in V$ 及 $\lambda\in\mathbb{R},\ \ <\lambda\mathbf{x},\mathbf{y}>=\lambda<\mathbf{x},\mathbf{y}>$；

4) $\forall \mathbf{x}\in V,\ \ <\mathbf{x},\mathbf{x}>\geqslant 0$，并且等号当且仅当 $\mathbf{x}=\mathbf{0}$ 时成立。

也即 $<\cdot,\cdot>$ 是 V 上的一个对称双线性型，其对应的二次型是正定的。则我们称 $(V,<\cdot,\cdot>)$ 是 $\mathbb{R}$ 上的 n 维实内积空间或 n 维欧几里得空间 (Euclid space)，简称欧氏空间[1]。$<\cdot,\cdot>$ 称为 V 上的内积 (inner product)。当内积已经明确时，我们通常省略内积，而直接称 V 是 n 维欧氏空间。

例 3.1.1 (1) 令 $V=\mathbb{R}^n$，任取 $\mathbf{x}=(x_1,\dots,x_n)^t$，$\mathbf{y}=(y_1,\dots,y_n)^t\in V$，定义：

$$<\mathbf{x},\mathbf{y}>=\sum_{i=1}^{n}x_iy_i$$

则容易验证 $<\cdot,\cdot>$ 满足上面的定义，从而是一个内积。我们把这个内积称为 $\mathbb{R}^n$ 上的**标准内积** (standard inner product)，称 $(\mathbb{R}^n,<\cdot,\cdot>)$ 是 n 维标准欧氏空间。

[1] 无穷维的实向量空间上如果有一个满足上述条件的内积，那么我们仍然称其为实内积空间，但只有有限维的内积空间才被称为欧氏空间。

(2) 令 $V = M_n(\mathbb{R})$，任取 $A, B \in V$，定义 $<A, B>= \mathrm{tr}(AB^t)$，则容易验证 $<A, B>$ 是一个内积，从而 $(M_n(\mathbb{R}), <\cdot,\cdot>)$ 是一个 n^2 维的欧氏空间。

(3) 固定 $a, b \in \mathbb{R}$，$a < b$，在 $\mathbb{R}[x]_n$(次数不超过 $n-1$ 的实多项式的构成的向量空间) 上，任取 $f(x), g(x) \in \mathbb{R}[x]_n$，定义：

$$<f, g>= \int_a^b f(t)g(t)\,\mathrm{d}t.$$

则容易验证这是一个内积，即 $(\mathbb{R}[x]_n, <\cdot,\cdot>)$ 是 n 维欧氏空间。

在一个实内积空间上，由定义 3.1.1，任何一个向量和自身的内积都是非负的，从而我们有下面的定义：

定义 3.1.2 设 V 是实内积空间，任取 $\mathbf{x} \in V$，则我们可以定义 $\mathbf{x}$ 的范数 (或称为长度，norm) 为：$\|\mathbf{x}\| = \sqrt{<\mathbf{x}, \mathbf{x}>}$。于是，任取 $\mathbf{x}, \mathbf{y} \in V$，我们称 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{<\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}>}$ 为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 的距离 (distance)。特别地，如果向量 $\mathbf{x}$ 的范数是 1，则称 $\mathbf{x}$ 是一个单位向量 (unit vector)。

例 3.1.2 在 n 维标准欧氏空间 $\mathbb{R}^n$ 上，向量 $\mathbf{x} = (x_1, \ldots, x_n)^t$ 的范数是 $\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$，任何两个向量 $\mathbf{x} = (x_1, \ldots, x_n)^t$，$\mathbf{y} = (y_1, \ldots, y_n)^t$ 的距离是 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1-y_1)^2 + \cdots + (x_n-y_n)^2}$。容易看出当 $n = 2$ 或 3 时上面的范数和距离的定义与我们中学时的定义是一致的。

我们在中学时就已经学过 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 上的的柯西不等式了。下面我们将它推广到一般形式。

定理 3.1.1 (Cauchy-Buniakowsky-Schwarz's inequation, **柯西-布尼亚科夫斯基-施瓦茨不等式**) [1] 设 $(V, <\cdot,\cdot>)$ 是实内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $|<\mathbf{x}, \mathbf{y}>| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\|$，并且等号成立的充分必要条件是 $\mathbf{x} = \mathbf{0}$ 或 $\exists \lambda \in \mathbb{R}$ 使得 $\mathbf{y} = \lambda\mathbf{x}$。

证明: 如果 $\mathbf{x}$ 或 $\mathbf{y}$ 是 $\mathbf{0}$，则不等式两边都是 0，显然成立。下设 $\mathbf{x}$ 和 $\mathbf{y}$ 都不是 $\mathbf{0}$。由内积的定义，对 $\forall \lambda \in \mathbb{R}$ 及 $\mathbf{x}, \mathbf{y} \in V$，我们有 $<\lambda\mathbf{x}-\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}> \geqslant 0$，并且等号当且仅当 $\lambda\mathbf{x}-\mathbf{y} = \mathbf{0}$ 时成立。将 $<\lambda\mathbf{x}-\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}> \geqslant 0$ 利用内积的对称和双线性性质展开可得

$$<\mathbf{x}, \mathbf{x}> \lambda^2 - 2 <\mathbf{x}, \mathbf{y}> \lambda + <\mathbf{y}, \mathbf{y}> \geqslant 0 \text{ 对} \forall \lambda \in \mathbb{R} \text{成立}. \tag{3.1.1}$$

这是一个关于 λ 的一元二次不等式，其恒非负的充要条件是判别式

$$4 <\mathbf{x}, \mathbf{y}>^2 - 4 <\mathbf{x}, \mathbf{x}> \cdot <\mathbf{y}, \mathbf{y}> \leqslant 0,$$

移项并对不等式两边开平方即得到 CBS 不等式。下面讨论等号成立条件。如果 CBS 不等式的等号成立并且 $\mathbf{x}, \mathbf{y}$ 都不是 $\mathbf{0}$，则我们立刻可知式 (3.1.1) 不等号左边的判别式是 0，这说明抛物线 $<\mathbf{x}, \mathbf{x}> \lambda^2 - 2 <\mathbf{x}, \mathbf{y}> \lambda + <\mathbf{y}, \mathbf{y}>$ 与直线 $\lambda = 0$ 相切，即存在 $\lambda \in \mathbb{R}$ 使得 $<\lambda\mathbf{x}-\mathbf{y}, \lambda\mathbf{x}-\mathbf{y}>= 0$，也即 $\lambda\mathbf{x}-\mathbf{y} = 0$。反过来，由 $\exists \lambda \in \mathbb{R}$ 使得 $\mathbf{y} = \lambda\mathbf{x}$ 得到 CBS 不等式取等号是容易的，只需注意到此时

$$|<\mathbf{x}, \mathbf{y}>| = |<\mathbf{x}, \lambda\mathbf{x}>| = |\lambda| \cdot |<\mathbf{x}, \mathbf{x}>|,$$
$$\|\mathbf{x}\| \cdot \|\mathbf{y}\| = \sqrt{<\mathbf{x}, \mathbf{x}>} \cdot \sqrt{<\lambda\mathbf{x}, \lambda\mathbf{x}>} = |\lambda| \cdot |<\mathbf{x}, \mathbf{x}>|.$$

即可。这样我们就完成了证明。 □

[1]简称 CBS 不等式。1821 年柯西最早发现了该不等式，1859 年布尼亚科夫斯基证明了其积分形式，1888 年施瓦茨给出了现代的证明。

注 3.1.1 将 CBS 不等式应用到 n 维标准欧氏空间 $\mathbb{R}^n$ 上可知，任取 $x_1,\ldots,x_n$ 及 $y_1,\ldots,y_n\in\mathbb{R}$，我们有：

$$|x_1y_1+\cdots+x_ny_n|\leqslant\sqrt{x_1^2+\cdots+x_n^2}\cdot\sqrt{y_1^2+\cdots+y_n^2}.$$

即我们熟知的柯西不等式；将 CBS 不等式用到 $(M_n(\mathbb{R}),<A,B>=\mathrm{tr}(AB^t))$ 上可得

$$|\mathrm{tr}(AB^t)|\leqslant\sqrt{\mathrm{tr}(AA^t)}\sqrt{\mathrm{tr}(BB^t)};$$

将 CBS 不等式用到 $(\mathbb{R}[x]_n,<f,g>=\int_a^b f(t)g(t)\,\mathrm{d}t)$ 上可得

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right|\leqslant\sqrt{\int_a^b f^2(t)\,\mathrm{d}t}\cdot\sqrt{\int_a^b g^2(t)\,\mathrm{d}t}.$$

即数学分析中的施瓦茨不等式。(实际上，我们可以将 $\mathbb{R}[x]_n$ 换成区间 $[a,b]$ 上的全体连续函数 $C[a,b]$，这仍然是一个内积空间，但不是有限维的，此时施瓦茨不等式仍然成立。)

推论 3.1.1 (**三角不等式**) 设 $(V,<\cdot,\cdot>)$ 是实内积空间，任取 $\mathbf{x},\mathbf{y}\in V$，则 $\|\mathbf{x}+\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|$，并且等号成立的充分必要条件为 $\exists\alpha,\beta\in\mathbb{R}$ 使得 $\alpha\mathbf{x}+\beta\mathbf{y}=\mathbf{0}$。

将范数全部改写成内积形式，整理后利用 CBS 不等式即可证明。细节留作练习。

有了以上结论，我们很容易验证范数和距离满足以下性质：

命题 3.1.1 设 $(V,<\cdot,\cdot>)$ 是实内积空间，$\|\cdot\|$ 是该内积对应的范数，则

(1) $\forall\mathbf{x}\in V$，$\|\mathbf{x}\|\geqslant 0$，并且 $\|\mathbf{x}\|=0\iff\mathbf{x}=\mathbf{0}$；

(2) $\forall\lambda\in\mathbb{R}$，$\mathbf{x}\in V$，都有 $\|\lambda\mathbf{x}\|=|\lambda|\|\mathbf{x}\|$；

(3) $\forall\mathbf{x},\mathbf{y}\in V$，有 $\|\mathbf{x}+\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|$。

命题 3.1.2 设 $(V,<\cdot,\cdot>)$ 是实内积空间，$\mathrm{d}(\cdot,\cdot)$ 是该内积对应的距离，则

(1) $\forall\mathbf{x},\mathbf{y}\in V$，$\mathrm{d}(\mathbf{x},\mathbf{y})\geqslant 0$，并且等号当且仅当 $\mathbf{y}=\mathbf{x}$ 时成立；

(2) $\forall\mathbf{x},\mathbf{y}\in V$，都有 $\mathrm{d}(\mathbf{x},\mathbf{y})=\mathrm{d}(\mathbf{y},\mathbf{x})$；

(3) $\forall\mathbf{x},\mathbf{y},\mathbf{z}\in V$，有 $\mathrm{d}(\mathbf{x},\mathbf{z})\leqslant\mathrm{d}(\mathbf{x},\mathbf{y})+\mathrm{d}(\mathbf{y},\mathbf{z})$。

实际上，我们也可以绕开内积而直接在向量空间上用这些性质定义范数或距离，当然此时所定义出的空间上不一定还有内积结构，我们会在下一节中继续讨论这种情形。

有了 CBS 不等式，我们就可以定义两个向量的夹角了。

定义 3.1.3 设 $(V,<\cdot,\cdot>)$ 是实内积空间，任取非零向量 $\mathbf{x},\mathbf{y}\in V$，我们定义 $\mathbf{x}$ 和 $\mathbf{y}$ 的夹角是如下的 θ：

$$\cos\theta=\frac{<\mathbf{x},\mathbf{y}>}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|},\ 0\leqslant\theta\leqslant\pi.\qquad(\text{CBS 不等式保证了}\theta\text{的存在性.})$$

其中，若 $\theta=0$，则我们称 $\mathbf{x},\mathbf{y}$ 同向；若 $\theta=\pi$，则我们称 $\mathbf{x},\mathbf{y}$ 反向；若 $\theta=\dfrac{\pi}{2}$，则我们称 $\mathbf{x},\mathbf{y}$ 正交 (orthogonal)，记作 $\mathbf{x}\perp\mathbf{y}$。特别地，我们规定 $\mathbf{0}$ 与任意 V 中的向量都正交，即 $\forall\mathbf{x},\mathbf{y}\in V$, $\mathbf{x}\perp\mathbf{y}\iff<\mathbf{x},\mathbf{y}>=0$。

于是我们立刻有

例 3.1.3 (**勾股定理**) 设 $(V, <\cdot,\cdot>)$ 是实内积空间，$\mathbf{x}, \mathbf{y} \in V$ 且 $\mathbf{x} \perp \mathbf{y}$，则

$$||\mathbf{x}+\mathbf{y}||^2 =< \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} >=< \mathbf{x}, \mathbf{x} > +2 < \mathbf{x}, \mathbf{y} > + < \mathbf{y}, \mathbf{y} >= ||\mathbf{x}||^2 + ||\mathbf{y}||^2.$$

引理 3.1.1 设 $(V, <\cdot,\cdot>)$ 是实内积空间，$\mathbf{x} \in V$，$\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是子空间 U 的一组基。如果 $\mathbf{x} \perp \mathbf{u}_1, \ldots, \mathbf{x} \perp \mathbf{u}_k$ 都成立，则任取 $\mathbf{u} \in U$，都有 $\mathbf{x} \perp \mathbf{u}$。我们把这种情形记作 $\mathbf{x} \perp U$。

证明: 由于 $\mathbf{u}_1, \ldots, \mathbf{u}_k$ 是子空间 U 的一组基, 故任取 $\mathbf{u} \in U$, 存在 $a_1, \ldots, a_k \in \mathbb{R}$ 使得 $\mathbf{u} = \sum_{i=1}^k a_i \mathbf{u}_i$。所以

$$< \mathbf{x}, \mathbf{u} >= \sum_{i=1}^k a_i < \mathbf{x}, \mathbf{u}_i >= 0,$$

即 $\mathbf{x} \perp U$。 □

内积和范数还满足以下性质:

命题 3.1.3 设 $(V, <\cdot,\cdot>)$ 是实内积空间，$\mathbf{x}, \mathbf{y} \in V$，则

(1) (扩展的三角不等式)$\Big| ||\mathbf{x}|| - ||\mathbf{y}|| \Big| \leqslant ||\mathbf{x} \pm \mathbf{y}|| \leqslant ||\mathbf{x}|| + ||\mathbf{y}||$；

(2) (平行四边形法则)$||\mathbf{x}+\mathbf{y}||^2 + ||\mathbf{x}-\mathbf{y}||^2 = 2(||\mathbf{x}||^2 + ||\mathbf{y}||^2)$；

(3) (极化恒等式)$< \mathbf{x}, \mathbf{y} >= \dfrac{1}{4}(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2)$。

其中 (1) 利用三角不等式即可证明，(2) 和 (3) 利用范数的定义和内积的双线性性质展开就可以证明，留作练习。

下面我们考虑 n 维欧氏空间的基底和子空间。由于欧氏空间上有内积结构，因此，我们希望它们在内积结构下还要满足一些好的性质。

定义 3.1.4 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间。如果 V 的一组基 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 满足:

$$\forall i \neq j \in \{1, \ldots, n\}, \quad < \mathbf{v}_i, \mathbf{v}_j >= 0,$$

则我们称 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组正交基 (orthogonal basis)；如果 V 的一组基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是正交基，并且每个 $\mathbf{e}_i$，$i \in \{1, \ldots, n\}$ 的范数都是 1，即

$$< \mathbf{e}_i, \mathbf{e}_j >= \delta_{ij},$$

则我们称 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组标准正交基 (normalized orthogonal basis)。

例 3.1.4 显然 $\mathbb{R}^n$ 的标准基就是一组标准正交基。

引理 3.1.2 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间。如果向量组 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \subset V$ 中的向量两两正交且都不为 $\mathbf{0}$，即 $\forall i \neq j$，$< \mathbf{x}_i, \mathbf{x}_j >= 0$，则 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ 是一个线性无关组。特别地，如果还有 $m = n$，则 $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ 是 V 的一组基。

证明: 设有线性组合 $a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m = \mathbf{0}$，其中 $a_1, \ldots, a_m \in \mathbb{R}$，我们只需证明 $a_1 = \cdots = a_m = 0$ 即可完成证明。注意到对每个 $i \in \{1, \ldots, m\}$，我们有

$$0 =< \mathbf{0}, \mathbf{x}_i >=< a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m, \mathbf{x}_i >= \sum_{k=1}^m a_k < \mathbf{x}_k, \mathbf{x}_i >= a_i < \mathbf{x}_i, \mathbf{x}_i >,$$

而 $\mathbf{x}_i \neq \mathbf{0} \Longrightarrow < \mathbf{x}_i, \mathbf{x}_i > \neq 0$，所以 $a_i = 0$ 对每个 $i \in \{1, \ldots, m\}$ 都成立，即 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 线性无关。特别地，当 $m = n$ 时，由于 V 是 n 维的，故 $\mathbf{x}_1, \ldots, \mathbf{x}_m$ 是 V 的一组基。 □

定义 3.1.5 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的子空间。则我们很容易验证：

$$W = \{\mathbf{x} \in V \mid \forall \mathbf{u} \in U, \quad <\mathbf{x}, \mathbf{u}>= 0\}$$

是 V 的子空间，并且 $V = U \oplus W$ [1] 。我们称 W 是 U 的正交补 (orthogonal complement) 空间，记作 $W = U^{\perp}$。

定理 3.1.2 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，则 V 中一定存在一组标准正交基。

证明: 我们从 V 的任意一组基出发，通过 Gram-Schmidt(**格拉姆-施密特**) **正交化**方法来构造出 V 的一组标准正交基。

任取 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 V 的一组基，令

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 & \mathbf{e}_1 &= \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1\\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{<\mathbf{v}_2, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 & \mathbf{e}_2 &= \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2\\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{<\mathbf{v}_3, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 - \frac{<\mathbf{v}_3, \mathbf{u}_2>}{<\mathbf{u}_2, \mathbf{u}_2>}\mathbf{u}_2 & \mathbf{e}_3 &= \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3\\
\mathbf{u}_4 &= \mathbf{v}_4 - \frac{<\mathbf{v}_4, \mathbf{u}_1>}{<\mathbf{u}_1, \mathbf{u}_1>}\mathbf{u}_1 - \frac{<\mathbf{v}_4, \mathbf{u}_2>}{<\mathbf{u}_2, \mathbf{u}_2>}\mathbf{u}_2 - \frac{<\mathbf{v}_4, \mathbf{u}_3>}{<\mathbf{u}_3, \mathbf{u}_3>}\mathbf{u}_3 & \mathbf{e}_4 &= \frac{1}{\|\mathbf{u}_4\|}\mathbf{u}_4\\
&\vdots & &\vdots\\
\mathbf{u}_n &= \mathbf{v}_n - \sum_{k=1}^{n-1} \frac{<\mathbf{v}_n, \mathbf{u}_k>}{<\mathbf{u}_k, \mathbf{u}_k>}\mathbf{v}_k & \mathbf{e}_n &= \frac{1}{\|\mathbf{u}_n\|}\mathbf{u}_n.
\end{aligned}$$

则容易看出：$\forall k \in \{1, \ldots, n\}$，都有 $\mathbf{u}_k \perp \mathrm{span}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$。这是因为：用数学归纳法，$\mathbf{u}_2 \perp \mathbf{u}_1$ 显然，不妨设 $\mathbf{u}_j \perp \mathrm{span}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}\}$ 对每个 $j = 1, \ldots, k-1$ 成立，下面我们来证明 $\mathbf{u}_k \perp \mathrm{span}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}\}$。任取 $i < k$，则

$$\begin{aligned}
<\mathbf{u}_k, \mathbf{u}_i> &= \left\langle \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{<\mathbf{v}_k, \mathbf{u}_j>}{<\mathbf{u}_j, \mathbf{u}_j>}\mathbf{u}_j \ , \ \mathbf{u}_i \right\rangle\\
&=<\mathbf{v}_k, \mathbf{u}_i> - \sum_{j=1}^{k-1} \frac{<\mathbf{v}_k, \mathbf{u}_j>}{<\mathbf{u}_j, \mathbf{u}_j>} <\mathbf{u}_j, \mathbf{u}_i>\\
&=<\mathbf{v}_k, \mathbf{u}_i> - \frac{<\mathbf{v}_k, \mathbf{u}_i>}{<\mathbf{u}_i, \mathbf{u}_i>} <\mathbf{u}_i, \mathbf{u}_i>= 0
\end{aligned}$$

这就说明 $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 两两正交。而每个 $\mathbf{e}_i$ 与 $\mathbf{u}_i$ 同向，因此 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 也两两正交。又因为每个 $\mathbf{e}_i$ 的范数都是 1，故 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 就是 V 的一组标准正交基。 □

推论 3.1.2 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的 m 维子空间，若 U 有一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_m$，则它可以扩充成 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n$。

类似于上面的定理，用 Gram-Schmidt 正交化方法即可证明。定理 3.1.2的证明过程也是计算标准正交基的方法。我们来看下面的例子。

例 3.1.5 在标准正交基 $\mathbb{R}^3$ 上，用 Gram-Schmidt 正交化方法将 $\mathbf{v}_1 = (3, 0, 4)^t$，$\mathbf{v}_2 = (-1, 0, 7)^t$，$\mathbf{v}_3 = (2, , 9, 11)^t$ 化成一组标准正交基。

[1]这个性质我们将放在稍后证明，当然，我们也会在那里顺便给出一个不依赖于正交基选取的证明。

解: 由 Gram-Schmidt 正交化方法，我们有

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 = (3,0,4)^t & \mathbf{e}_1 &= \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = (\frac{3}{5},0,\frac{4}{5})^t \\
\mathbf{u}_2 &= \mathbf{v}_2 - \frac{<\mathbf{v}_2,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1 = (-4,0,3)^t & \mathbf{e}_2 &= \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = (-\frac{4}{5},0,\frac{3}{5})^t \\
\mathbf{u}_3 &= \mathbf{v}_3 - \frac{<\mathbf{v}_3,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1 - \frac{<\mathbf{v}_3,\mathbf{u}_2>}{<\mathbf{u}_2,\mathbf{u}_2>}\mathbf{u}_2 = (0,9,0)^t & \mathbf{e}_3 &= \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = (0,1,0)^t.
\end{aligned}$$

则 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 就是 $\mathbb{R}^3$ 的一组标准正交基。

从这个例子我们也可以看出，标准正交基的选取并不是唯一的。有了标准正交基以后，我们可以顺便证明正交补空间的以下性质：

定理 3.1.3 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，U 是 V 的子空间，则 $V = U\oplus U^\perp$ 并且 $(U^\perp)^\perp = U$。

证明: 设 $\dim(U) = m$，则 U 在原来的内积下仍然是欧氏空间，取 U 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m$，则对任意的 $\mathbf{x}\in V$，令

$$\mathbf{w} = \mathbf{x} - \sum_{i=1}^{m} <\mathbf{x},\mathbf{e}_i>\mathbf{e}_i,$$

则对任意的 $j\in\{1,\ldots,m\}$，有

$$<\mathbf{w},\mathbf{e}_j> = <\mathbf{x},\mathbf{e}_j> - \sum_{i=1}^{m} <\mathbf{x},\mathbf{e}_i><\mathbf{e}_i,\mathbf{e}_j> = <\mathbf{x},\mathbf{e}_j> - <\mathbf{x},\mathbf{e}_j> = 0.$$

即 $\mathbf{w}$ 与每个 $\mathbf{e}_j$ 都正交，所以 $\mathbf{w}$ 与 U 正交，$\mathbf{w}\in U^\perp$。而 $\sum_{i=1}^m <\mathbf{x},\mathbf{e}_i>\mathbf{e}_i\in U$，于是

$$\mathbf{x} = \sum_{i=1}^{m} <\mathbf{x},\mathbf{e}_i>\mathbf{e}_i + \mathbf{w}\in U + U^\perp,$$

即 $V = U + U^\perp$。下证这个和是直和。任取 $\mathbf{y}\in U\cap U^\perp$，则 $<\mathbf{y},\mathbf{y}> = 0$(前一个 $\mathbf{y}$ 视作 U 中的向量，后一个 $\mathbf{y}$ 视作 $U^\perp$ 中的向量)，于是由内积的正定性可知只能是 $\mathbf{y} = \mathbf{0}$。因此 $V = U\oplus U^\perp$。

最后我们证明 $U = (U^\perp)^\perp$。首先 $U\subset(U^\perp)^\perp$：任取 $\mathbf{x}\in U$，我们欲证 $\mathbf{x}\perp U^\perp$。任取 $\mathbf{y}\in U^\perp$，由 $U^\perp$ 的定义可知 $<\mathbf{x},\mathbf{y}> = \mathbf{0}$，即 $\mathbf{x}\perp\mathbf{y}$，所以 $\mathbf{x}\perp U^\perp$。其次 U 和 $(U^\perp)^\perp$ 的维数相等：利用定理前半部分的结论可知 $V = U\oplus U^\perp = U^\perp\oplus(U^\perp)^\perp$，因此 $\dim(U) = \dim((U^\perp)^\perp)$。由以上两点可知 $U = (U^\perp)^\perp$(包含定理，1.3.2(3))。 □

证明 (**法二**): $U = (U^\perp)^\perp$ 的证法与上一个证明相同。现在我们绕开标准正交基来证明 $V = U\oplus U^\perp$。

设 $\dim(U) = m$，任取 U 的一组基 $\mathbf{u}_1,\ldots,\mathbf{u}_m$，作如下的线性函数：

$$\begin{aligned}
f_1(\mathbf{x}) &= <\mathbf{u}_1,\mathbf{x}>, \\
f_2(\mathbf{x}) &= <\mathbf{u}_2,\mathbf{x}>, \\
&\cdots\cdots \\
f_m(\mathbf{x}) &= <\mathbf{u}_m,\mathbf{x}>,
\end{aligned}$$

则容易验证 $f_i(\mathbf{x}): V\to\mathbb{R}\ \in V^*$，$\forall i\in\{1,\ldots,m\}$，并且这些 f_i 在 V^* 中是线性无关的。这是因为：任取 $\alpha_1,\ldots,\alpha_m$，如果

$$\sum_{i=1}^{m}\alpha_i f_i(\mathbf{x}) = 0\in V^*$$

即 $<\sum_{i=1}^m \alpha_i\mathbf{u}_i, \mathbf{x}>=0$ 对任意的 $\mathbf{x}\in V$ 成立，则只能是 $\sum_{i=1}^m \alpha_i\mathbf{u}_i=\mathbf{0}$ (因为 $<\cdot,\cdot>$ 的正定性)，再由 $\mathbf{u}_1,\ldots,\mathbf{u}_m$ 是 U 的基可知 $\alpha_1=\cdots=\alpha_m=0$，即这些 f_i 在 V^* 中线性无关。

下面我们证明 $U^\perp$ 是抽象齐次线性方程组：

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

的解空间。为此，我们把这个解空间记作 W，则一方面任取 $\mathbf{x}\in U^\perp$，$\mathbf{x}$ 必然与每个 $\mathbf{u}_i$ 正交，所以 $\mathbf{x}\in W$，即 $U^\perp\subset W$；另一方面，任取 $\mathbf{y}\in W$，则 $\mathbf{y}$ 与 U 的每个基底 $\mathbf{u}_i$ 正交，于是由内积的双线性性质，$\mathbf{y}$ 和 U 中的任何元素都正交，即 $\mathbf{y}\in U^\perp$，所以 $W\subset U^\perp$。综上所述，$W=U^\perp$。

因此，由定理 1.5.6，$\dim(U^\perp)=n-\dim(\mathrm{span}\,\{f_1,\ldots,f_m\})=n-m$。再由 $U\cap U^\perp=\{\mathbf{0}\}$ (同上面的证法) 即可得到 $V=U\oplus U^\perp$。

现在我们考虑向量在标准正交基下的坐标。

命题 3.1.4 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是一组标准正交基。任取 $\mathbf{x}\in V$，则

(1) $\mathbf{x}=\sum_{i=1}^n <\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$，我们把 $<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$ 称为 $\mathbf{x}$ 在 $\mathbf{e}_i$ 方向上的**投影** (projection)；

(2) (Parseval 等式)[1] $\|\mathbf{x}\|^2=\sum_{i=1}^n|<\mathbf{x},\mathbf{e}_i>|^2$。

证明: (1) 设 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，我们证明每个 $x_i=<\mathbf{x},\mathbf{e}_i>$ 即可。对每个 $j\in\{1,\ldots,n\}$，作内积

$$<\mathbf{x},\mathbf{e}_j>=\sum_{i=1}^n x_i<\mathbf{e}_i,\mathbf{e}_j>=\sum_{i=1}^n x_i\delta_{ij}=x_j.$$

即得结论。
(2) 由 (1) 的结论，我们有：

$$\begin{aligned}\|\mathbf{x}\|=<\mathbf{x},\mathbf{x}>&=<\sum_{i=1}^n\langle\mathbf{x},\mathbf{e}_i\rangle\mathbf{e}_i,\mathbf{x}>\\ &=\sum_{i=1}^n<\mathbf{x},\mathbf{e}_i><\mathbf{e}_i,\mathbf{x}>\\ &=\sum_{i=1}^n|<\mathbf{x},\mathbf{e}_i>|^2.\end{aligned}$$

即得结论。 □

现在我们来考虑欧氏空间的同构。由于欧氏空间上有内积结构，因此我们还希望同构映射能够保持内积。

定义 3.1.6 设 $(V,<\cdot,\cdot>_V)$ 和 $(W,<\cdot,\cdot>_W)$ 是两个 n 维欧氏空间，如果存在一个线性同构 $f:V\to W$ 使得

$$<\mathbf{x},\mathbf{y}>_V=<f(\mathbf{x}),f(\mathbf{y})>_W,\ \forall\mathbf{x},\mathbf{y}\in V,$$

则称这两个**欧氏**空间 V 和 W 同构。

[1]对于无穷维的内积空间，不等式 $\|\mathbf{x}\|^2\geqslant\sum_{i=1}^\infty|<\mathbf{x},\mathbf{e}_i>|^2$ 总成立，称为 Bessel 不等式。详细的证明可以参考泛函分析的标准教材，如《A Course in Functional Analysis》, Jorn B.Conway, GTM96 的 Chapter I, §4。

定理 3.1.4 设 $(V, <\cdot,\cdot>)$ 是任意一个 n 维欧氏空间，则 V 一定同构于 n 维标准欧氏空间 $(\mathbb{R}^n, <\cdot,\cdot>_{\mathbb{R}^n})$。

证明: 取 V 的一组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$，则任取 $\mathbf{x}\in V$，$\mathbf{x}$ 可以写成线性组合 $\mathbf{x}=x_1\mathbf{e}_1+\cdots+x_n\mathbf{e}_n$，$x_i\in\mathbb{R}$，$i\in\{1,\dots,n\}$，于是我们可以定义：

$$\begin{aligned} f: V &\longrightarrow \mathbb{R}^n \\ \mathbf{x} &\longmapsto (x_1,\dots,x_n)^t \end{aligned}$$

显然 f 是向量空间 V 到 $\mathbb{R}^n$ 的线性同构 (定理 1.3.4)。下面我们只需证明 f 保持内积即可。

任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i, \mathbf{y}=\sum_{j=1}^n y_j\mathbf{e}_j\in V$，我们有

$$\begin{aligned} <\mathbf{x},\mathbf{y}> &= \sum_{i=1}^n\sum_{j=1}^n x_iy_j <\mathbf{e}_i,\mathbf{e}_j> = \sum_{i=1}^n\sum_{j=1}^n x_iy_j\delta_{ij} = \sum_{i=1}^n x_iy_i, \\ <f(\mathbf{x}),f(\mathbf{y})> &= \sum_{i=1}^n x_iy_i. \end{aligned}$$

即 $<f(\mathbf{x}),f(\mathbf{y})>=<\mathbf{x},\mathbf{y}>$，此即这两个欧氏空间同构。 □

现在我们考虑欧氏空间的对偶空间。设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，任取 $\mathbf{v}\in V$，容易验证 $\Phi_{\mathbf{v}}: V\to\mathbb{R}$，$\mathbf{x}\mapsto<\mathbf{v},\mathbf{x}>$ 是一个线性函数，即 $\Phi_{\mathbf{v}}\in V^*$ (在定理 3.1.3的证法二中已经让大家验证过一次了)。实际上，我们有

命题 3.1.5 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，则

$$\Phi: V\longrightarrow V^*,\quad \mathbf{v}\longmapsto\Phi_{\mathbf{v}}\ (\Phi_{\mathbf{v}}\text{的定义如上文所述.})$$

是向量空间 V 到 V^* 的线性同构，并且 Φ 把 V 的任意一组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 映到其对偶基 $\mathbf{e}^1,\dots,\mathbf{e}^n$。

证明: 首先 Φ 是线性映射：任取 $\mathbf{u},\mathbf{v}\in V$ 及 $\alpha,\beta\in\mathbb{R}$，将线性函数作用到任意 $\mathbf{x}\in V$ 上，我们有

$$\begin{aligned} \Phi_{\alpha\mathbf{u}+\beta\mathbf{v}}(\mathbf{x}) &=<\alpha\mathbf{u}+\beta\mathbf{v},\mathbf{x}> \\ &=\alpha<\mathbf{u},\mathbf{x}>+\beta<\mathbf{v},\mathbf{x}> \\ &=\alpha\Phi_{\mathbf{u}}(\mathbf{x})+\beta\Phi_{\mathbf{v}}(\mathbf{x}). \end{aligned}$$

其次 Φ 是单射，这只需验证 $\ker(\Phi)=\{\mathbf{0}\}$：任取 $\mathbf{v}\in\ker(\Phi)$，即 $\Phi_{\mathbf{v}}$ 是零映射，则按 $\Phi_{\mathbf{v}}$ 的定义，对任意 $\mathbf{x}\in V$，都有 $<\mathbf{v},\mathbf{x}>=0$，因此只能是 $\mathbf{v}=\mathbf{0}$[1]。

下面我们说明 Φ 把 V 的标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 映到其对偶基 $\mathbf{e}^1,\dots,\mathbf{e}^n$，从而由维数关系可知 Φ 是满射，即可完成证明。注意到对任意的 $j\in\{1,\dots,n\}$，我们都有

$$\begin{aligned} \Phi_{\mathbf{e}_i}(\mathbf{e}_j) &=<\mathbf{e}_i,\mathbf{e}_j>=\delta_{ij} \\ \mathbf{e}^i(\mathbf{e}_j) &=\delta_{ij} \end{aligned}$$

由定义 1.5.2下方的说明即可得到 $\Phi_{\mathbf{e}_i}=\mathbf{e}^i$ 对每个 $i\in\{1,\dots,n\}$ 都成立。 □

[1]这里利用了内积的正定性，取 $\mathbf{x}=\mathbf{v}$ 即得结论。这是一个简单而又常用的结论，我们就不把它写成定理了。

我们还可以在 V^* 上定义如下的内积：$<\Phi_{\mathbf{u}},\Phi_{\mathbf{v}}>=<\mathbf{u},\mathbf{v}>$，并且很容易证明在此内积下 V^* 也是 n 维欧氏空间，上面命题中的 Φ 是欧氏空间的同构。实际上，这个命题的一个更常用的表述是：任意的 $f\in V^*$，存在唯一的 $\mathbf{v}\in V$ 使得 $f(\mathbf{x})=<\mathbf{v},\mathbf{x}>$，并且 $\|f\|=\|\mathbf{v}\|$。上面的命题是 Riesz 表示定理在欧氏空间上的特殊情形，更一般的情形我们会在泛函分析课程中讨论。

接下来我们考虑两组标准正交基之间的过渡矩阵。设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，由于 $<\cdot,\cdot>$ 是 V 上的对称双线性型，故任取 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$，双线性型 $<\cdot,\cdot>$ 对应的矩阵为

$$G=\begin{pmatrix}<\mathbf{v}_1,\mathbf{v}_1> & <\mathbf{v}_1,\mathbf{v}_2> & \cdots & <\mathbf{v}_1,\mathbf{v}_n>\\ <\mathbf{v}_2,\mathbf{v}_1> & <\mathbf{v}_2,\mathbf{v}_2> & \cdots & <\mathbf{v}_2,\mathbf{v}_n>\\ \vdots & \vdots & & \vdots\\ <\mathbf{v}_n,\mathbf{v}_1> & <\mathbf{v}_n,\mathbf{v}_2> & \cdots & <\mathbf{v}_n,\mathbf{v}_n>\end{pmatrix}$$

我们称上面的 G 为内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的**度量矩阵**或 Gram 矩阵。由第一章的定理 1.6.2可知，如果内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 下的度量矩阵分别为 G 和 G'，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$, $A\in\mathrm{GL}_n(\mathbb{R})$, 则 $G'=A^tGA$。显然 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基 $\Longleftrightarrow$ 其度量矩阵为 E_n。于是，如果 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 都是 V 的标准正交基，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$，则 $E_n=A^tE_nA$，即 $A^tA=E_n$。反之，如果矩阵 A 满足 $A^tA=E_n$，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，令 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$，则显然 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是 V 的基 (A 显然可逆)，并且按上面的讨论，内积在 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的度量矩阵为 $A^tE_nA=E_n$，即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是标准正交基。这就引出了下面的正交矩阵的定义。

定义 3.1.7 设 $A\in\mathrm{GL}_n(\mathbb{R})$，如果 A 满足 $A^tA=E$，即 $A^t=A^{-1}$，则称 A 为正交矩阵 (orthonormal matrix)。容易验证 $\mathbb{R}$ 上所有 n 阶正交矩阵在矩阵乘法运算下构成了一个群 (封闭：$A^t=A^{-1},B^t=B^{-1}$，则 $(AB)^t=B^tA^t=B^{-1}A^{-1}=(AB)^{-1}$，结合律显然，幺元为 E_n，A 的逆元为 A^t。)，我们把这个群称为 (实) 正交群 (orthonormal group)，记作 $O(n)$。

容易看出正交矩阵就是标准正交基之间的过渡矩阵。在上一章中我们已经知道，方阵与线性算子之间是一一对应的。因此，受定义 3.1.7的启发，我们也可以定义正交算子如下：

定义 3.1.8 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的可逆的线性算子。如果 $\mathcal{A}$ 还满足：对任意的 $\mathbf{x},\mathbf{y}\in V$，$<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathcal{A}$ 是 V 上的正交算子 (orthonormal operator) 或正交变换。

我们有以下定理：

定理 3.1.5 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的线性算子，则以下条件等价：

(1) $\mathcal{A}$ 是正交算子；

(2) $\mathcal{A}$ 在标准正交基下对应的矩阵是正交矩阵；

(3) $\mathcal{A}$ 将标准正交基映到标准正交基；

(4) $\forall\mathbf{x}\in V$, $\|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|$。

证明: (1)$\Longrightarrow$(2)：设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，$\mathcal{A}$ 在这组基下的矩阵为 A。首先，令 $\mathbf{v}_i=\mathcal{A}\mathbf{e}_i$, $i=1,\ldots,n$，则

$$<\mathbf{v}_i,\mathbf{v}_j>=<\mathcal{A}\mathbf{e}_i,\mathcal{A}\mathbf{e}_j>=<\mathbf{e}_i,\mathbf{e}_j>=\delta_{ij}$$

即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是 V 的一组标准正交基。显然 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$，由定义 3.1.7前面的那段讨论可知 A 是正交矩阵。

(2)$\Longrightarrow$(3)：即定义 3.1.7前面的讨论（“反之”之后的部分)。

(3)$\Longrightarrow$(4)：设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i\in V$，则 $\mathcal{A}\mathbf{x}=\sum_{i=1}^n x_i\mathcal{A}\mathbf{e}_i$。由于 $\mathcal{A}\mathbf{e}_1,\ldots,\mathcal{A}\mathbf{e}_n$ 也是 V 的一组标准正交基，因此我们有：

$$<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=\sum_{i=1}^n\sum_{j=1}^n x_ix_j<\mathcal{A}\mathbf{e}_i,\mathcal{A}\mathbf{e}_j>=\sum_{i=1}^n\sum_{j=1}^n x_ix_j\delta_{ij}=\sum_{i=1}^n x_i^2;$$
$$<\mathbf{x},\mathbf{x}>=\sum_{i=1}^n\sum_{j=1}^n x_ix_j<\mathbf{e}_i,\mathbf{e}_j>=\sum_{i=1}^n\sum_{j=1}^n x_ix_j\delta_{ij}=\sum_{i=1}^n x_i^2.$$

所以 $\|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|$。

(4)$\Longrightarrow$(1)：由极化恒等式 (命题 3.1.3(3))，任取 $\mathbf{x},\mathbf{y}\in V$，我们有

$$\begin{aligned}<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>&=\frac{1}{4}(\|\mathcal{A}(\mathbf{x}+\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2)\\&=\frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)\\&=<\mathbf{x},\mathbf{y}>.\end{aligned}$$

即 $\mathcal{A}$ 是正交算子。 □

推论 3.1.3 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，则 V 上所有正交算子在算子复合运算下构成群，记作 $O(V)$，则 $O(V)$ 同构于正交群 $O(n)$。

证明: 首先 $O(V)$ 是群：设 $\mathcal{A},\mathcal{B}\in O(V)$，则对任意的 $\mathbf{x},\mathbf{y}\in V$，$<\mathcal{A}\mathcal{B}\mathbf{x},\mathcal{A}\mathcal{B}\mathbf{y}>=<\mathcal{B}\mathbf{x},\mathcal{B}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，即 $\mathcal{A}\mathcal{B}\in O(V)$，封闭性成立；结合律即映射复合的结合律；幺元是恒等算子 $\mathcal{E}$；由于正交算子一定可逆 (将基映到基，满秩)，故 $\mathcal{A}^{-1}$ 存在，并且 $<\mathcal{A}^{-1}\mathbf{x},\mathcal{A}^{-1}\mathbf{y}>=<\mathcal{A}\mathcal{A}^{-1}\mathbf{x},\mathcal{A}\mathcal{A}^{-1}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，即 $\mathcal{A}^{-1}\in O(V)$。因此 $O(V)$ 是群。

任取一组标准正交基，设 $\Phi: O(V)\to O(n)$ 将每个正交算子映到它在这种标准正交基下的矩阵，则由算子复合与矩阵乘法的对应关系可知 Φ 是群同态；由前面的论述显然 Φ 是双射，故 $O(V)\simeq O(n)$。 □

辛空间 *

在本节的最后，我们考虑另一种形式的“度量”结构，它被称为辛结构 (symplectic structure)。辛结构在几何和物理中都有重要的应用。

定义 3.1.9 设 V 是 $\mathbb{R}$ 上的 $n(=2m)$ 维向量空间，$<\cdot,\cdot>$ 是 V 上的非退化的斜对称双线性型[1]，即 $<\cdot,\cdot>$ 满足：

1) $\forall\mathbf{v}\in V$, $\exists\mathbf{x}\in V$ 使得 $<\mathbf{v},\mathbf{x}>\neq 0$;

2) $\forall\mathbf{x},\mathbf{y}\in V$, $<\mathbf{x},\mathbf{y}>=-<\mathbf{y},\mathbf{x}>$;

3) $\forall\mathbf{x},\mathbf{y},\mathbf{z}\in V$, $<\mathbf{x}+\mathbf{y},\mathbf{z}>=<\mathbf{x},\mathbf{z}>+<\mathbf{y},\mathbf{z}>$;

4) $\forall\mathbf{x},\mathbf{y}\in V$ 及 $\lambda\in\mathbb{R}$, $<\lambda\mathbf{x},\mathbf{y}>=\lambda<\mathbf{x},\mathbf{y}>$。

[1]注意，由 1.6.7 小节开头的讨论，只有偶数维的向量空间上才有非退化的斜对称双线性型。

则我们称 $<\cdot,\cdot>$ 是 V 上的一个辛内积或辛结构，称 $(V, \ <\cdot,\cdot>)$ 是 $\mathbb{R}$ 上的 $2m$ 维辛空间 (symplectic space)。设 $\mathcal{A}\in\mathrm{End}(V)$，如果 $\mathcal{A}$ 还保持辛结构，即 $\forall\mathbf{x},\mathbf{y}\in V,\ <\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathcal{A}$ 是一个辛算子 (symplectic operator)。如果 $\mathbf{x},\mathbf{y}\in V$ 满足 $<\mathbf{x},\mathbf{y}>=0$，则我们称 $\mathbf{x}$ 和 $\mathbf{y}$ 是斜正交的，仍记作 $\mathbf{x}\perp\mathbf{y}$。显然 $\forall\mathbf{x}\in V,\ \mathbf{x}\perp\mathbf{x}$。

类似于欧氏内积的度量矩阵，我们也可以定义辛内积在一组基下的“度量”矩阵 (Gram 矩阵)。利用定理 1.6.13，我们很容易得到下面的结论：

定理 3.1.6 设 $(V,<\cdot,\cdot>)$ 是 $n=2m$ 维辛空间，则存在 V 的一组基 $\boldsymbol{\alpha_1},\ldots,\boldsymbol{\alpha_m},\boldsymbol{\beta_1},\ldots,\boldsymbol{\beta_m}$ (称为辛基) 使得辛内积在此基底下的 Gram 矩阵为 $J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}$，即只有 $<\boldsymbol{\alpha_i},\boldsymbol{\beta_i}>=-<\boldsymbol{\beta_i},\boldsymbol{\alpha_i}>=1$，$\forall i\in\{1,\ldots,m\}$，其余基向量之间的辛内积均为 0。

与标准欧氏空间类似，我们也可以定义标准辛空间。

定义 3.1.10 在 $\mathbb{R}^{2m}$ 上定义如下的辛内积：任取 $\mathbf{x}=(x_1,\ldots,x_m,x_{m+1},\ldots,x_{2m})^t$ 和 $\mathbf{y}=(y_1,\ldots,y_m,y_{m+1},\ldots,y_{2m})^t\in\mathbb{R}^{2m}$，定义

$$<\mathbf{x},\mathbf{y}>_s=\sum_{i=1}^{m}(x_{m+i}y_i-x_iy_{m+i})$$

则 $<\cdot,\cdot>_s$ 在 $\mathbb{R}^{2m}$ 的标准基 $\mathbf{e}_1=(1,0,\ldots,0)^t,\ldots\ldots,\mathbf{e}_{2m}=(0,\ldots,0,1)^t$ 下的 Gram 矩阵恰为上面定理中的 J_0。我们称 $(\mathbb{R}^{2m},\ <\cdot,\cdot>_s)$ 为 $2m$ 维的标准辛空间。注意到 $J_0^2=-E_{2m}$，这与虚数单位 i 的性质类似。在稍后的复化与实化一节中，我们还会用到这个空间。

与欧氏空间的情形类似，我们有：

定理 3.1.7 设 $(V,<\cdot,\cdot>)$ 是 $2m$ 维辛空间，则存在线性同构 $f:V\to\mathbb{R}^{2m}$ 使得 $<f(\mathbf{x}),f(\mathbf{y})>_s=<\mathbf{x},\mathbf{y}>$。(我们把这样保持辛内积的线性同构称为辛同构。)

证明方法与定理 3.1.4类似，留作练习。

定义 3.1.11 设 $A\in M_{2m}(\mathbb{R})$，$J_0=\begin{pmatrix}0 & -E_m\\ E_m & 0\end{pmatrix}$，如果 A 满足 $A^tJ_0A=J_0$，则我们称 A 为辛矩阵 (symplectic matrix)。

如果 A 是 $2m$ 阶的辛矩阵，则利用普法夫型的性质 (定理 1.6.14) 可知，$1=\mathrm{Pf}_{2m}(J_0)=\mathrm{Pf}_{2m}(A^tJ_0A)=\det(A)\mathrm{Pf}_{2m}(J_0)=\det(A)$。容易验证所有 $2m$ 阶辛矩阵在矩阵乘法下构成了一个群 (留作练习)，我们把这个群称为辛群，记作 $\mathrm{Sp}_{2m}(\mathbb{R})$。于是显然 $\mathrm{Sp}_{2m}(\mathbb{R})$ 是特殊线性群 $\mathrm{SL}_{2m}(\mathbb{R})$ 的子群。特别地，当 $2m=2$ 时，$\mathrm{Sp}_2(\mathbb{R})=\mathrm{SL}_2(\mathbb{R})$。这是因为任取 $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in\mathrm{SL}_2(\mathbb{R})$，即 $ad-bc=1$，我们有

$$A^tJ_0A=\begin{pmatrix}0 & -(ad-bc)\\ (ad-bc) & 0\end{pmatrix}=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}=J_0$$

即 $A\in\mathrm{Sp}_2(\mathbb{R})$。

下面我们考察辛算子与辛矩阵之间的联系。

定理 3.1.8 设 $(V,<\cdot,\cdot>)$ 是 $2m$ 维辛空间，则所有 V 上所有辛算子在算子复合所定义的乘法下构成一个群，记作 $\mathrm{Sp}(V)$，并且 $\mathrm{Sp}(V)\simeq\mathrm{Sp}_{2m}(\mathbb{R})$。

证明: 由定理 3.1.6，存在辛基 $\mathbf{e}_1, \ldots, \mathbf{e}_m, \ldots \mathbf{e}_{2m}$ 使得

$$(< \mathbf{e}_i, \mathbf{e}_j >)_{2m\times 2m} = J_0 = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$$

令 $\Phi : \mathrm{Sp}(V) \to \mathrm{Sp}_{2m}(\mathbb{R})$, $\mathcal{A} \mapsto A$，其中 A 是算子 $\mathcal{A}$ 在上述辛基下的矩阵。则用类似于定理 3.1.5的证明方法可以得到以下条件等价：

(1) $\mathcal{A}$ 是辛算子；

(2) $\mathcal{A}$ 在辛基下对应的矩阵是辛矩阵；

(3) $\mathcal{A}$ 将辛基映到辛基。

细节留作练习。于是 Φ 是线性双射，并且由于线性算子的复合对应矩阵的乘法，故 Φ 是群同态，因此 $\mathrm{Sp}(V) \simeq \mathrm{Sp}_{2m}(\mathbb{R})$。 □

最后我们关注辛算子的特征多项式的一个有趣的性质。

定理 3.1.9 设 $(V, <\cdot,\cdot>)$ 是 $2m$ 维辛空间，$\mathcal{A}$ 是 V 上的辛算子，则 $\chi_{\mathcal{A}}(t) = t^{2m}\chi_{\mathcal{A}}(\frac{1}{t})$。

证明: 设 $\mathcal{A}$ 在一组辛基下对应矩阵 A，则 A 满足 $A^t J_0 A = J_0$，于是 $A = J_0^{-1}(A^t)^{-1}J_0$。因此我们有：

$$\begin{aligned}
\chi_{\mathcal{A}}(t) &= \det(tE - A) = \det\left(tE - J_0^{-1}\cdot(A^t)^{-1}\cdot J_0\right) = \det\left(J_0^{-1}\left(tE - (A^t)^{-1}\right)J_0\right) \\
&= \det\left(tE - (A^t)^{-1}\right) = \det\left(\left(tE - A^{-1}\right)^t\right) = \det\left(tE - A^{-1}\right) \\
&= \det\left(tE - A^{-1}\right)\det A = \det(tA - E) = \det(E - tA) = t^{2m}\det\left(\frac{1}{t}E - A\right) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right)
\end{aligned}$$

这样我们就完成了证明。 □

3.2 埃尔米特向量空间

上一节我们讨论了实数域上的内积空间，现在我们尝试将内积定义到复向量空间上。为了区分复数的模长和向量的范数这两个记号，我们把复数 α 的模长记作 $|\alpha|$。由于复向量空间上的非退化二次型 $x_1^2+\cdots+x_n^2$ 不再具有正定性 ($i^2=-1$)，因此我们需要寻找新的函数来定义内积，这就是接下来介绍的埃尔米特 (Hermitian) 型。

3.2.1 埃尔米特型

定义 3.2.1 设 V 是 n 维复向量空间，$f: V\times V\to\mathbb{C}$ 满足：

(1) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V,\ \alpha,\beta\in\mathbb{C}$，有 $f(\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{z})=\alpha f(\mathbf{x},\mathbf{z})+\beta f(\mathbf{y},\mathbf{z})$；

(2) $\forall \mathbf{x},\mathbf{y},\mathbf{z}\in V,\ \alpha,\beta\in\mathbb{C}$，有 $f(\mathbf{z},\alpha\mathbf{x}+\beta\mathbf{y})=\overline{\alpha}f(\mathbf{z},\mathbf{x})+\overline{\beta}f(\mathbf{z},\mathbf{y})$，其中 $\overline{\alpha},\overline{\beta}$ 分别是 α 和 β 的共轭复数。

则称 f 是 V 上的半双线性型 (sesquilinear form, 即 f 对第一个变量线性，对第二个变量共轭线性)。如果半双线性型 f 还满足 $\forall \mathbf{x},\mathbf{y}\in V,\ f(\mathbf{y},\mathbf{x})=\overline{f(\mathbf{x},\mathbf{y})}$，则我们称 f 是一个埃尔米特 (Hermitian) 型。对应地，如果 $f(\mathbf{y},\mathbf{x})=-\overline{f(\mathbf{x},\mathbf{y})}$，则称 f 是一个斜埃尔米特型。

任取复向量空间 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$，类似于 1.6 节我们对双线性型的讨论，我们可以定义半双线性型 f 在此基底下的矩阵

$$F=\left(f_{ij}\right)_{n\times n},\ \text{其中} f_{ij}=f(\mathbf{e}_i,\mathbf{e}_j).$$

则任取 $\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i$，$\mathbf{y}=\sum_{j=1}^n y_j\mathbf{e}_j\in V$，利用 f 的半双线性性质可知

$$f(\mathbf{x},\mathbf{y})=\sum_{i=1}^n\sum_{j=1}^n f_{ij}x_i\overline{y_j}=(x_1,\ldots,x_n)\cdot F\cdot\begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}.$$

特别地，如果 f 是 Hermitian 型，则 $f_{ij}=f(\mathbf{e}_i,\mathbf{e}_j)=\overline{f(\mathbf{e}_j,\mathbf{e}_i)}=\overline{f_{ji}}$。我们记矩阵 F 的转置共轭[1]为 F^*(有的书上也记成 F^H 或 $F^\dagger$)，则 f 是 Hermitian 型 $\Longleftrightarrow$ f 在任意一组基下对应的矩阵 F 满足 $F^*=F$。我们称满足 $F^*=F$ 的复矩阵为埃尔米特 (Hermitian) 矩阵。

下面我们考虑基变换下半双线性型对应的矩阵之间的关系。类似于定理 1.6.2的证明方法，设 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$ 是复向量空间 V 的两组基且 $(\mathbf{e}'_1,\ldots,\mathbf{e}'_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$。任取 $\mathbf{x},\mathbf{y}\in V$，设 $\mathbf{x}$ 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$ 下的坐标分别为 $(x_1,\ldots,x_n)^t$ 和 $(x'_1,\ldots,x'_n)^t$，$\mathbf{y}$ 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$ 下的坐标分别为 $(y_1,\ldots,y_n)^t$ 和 $(y'_1,\ldots,y'_n)^t$，f 在基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 和 $\mathbf{e}'_1,\ldots,\mathbf{e}'_n$ 下的矩阵分别为 F 和 F'，则

$$f(\mathbf{x},\mathbf{y})=(x_1,\ldots,x_n)F\begin{pmatrix}\overline{y_1}\\ \cdots\\ \overline{y_n}\end{pmatrix}=(x'_1,\ldots,x'_n)F'\begin{pmatrix}\overline{y'_1}\\ \cdots\\ \overline{y'_n}\end{pmatrix}\tag{3.2.1}$$

由坐标变换公式有

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

[1]很容易验证对复矩阵 F 的每个元素取共轭复数的操作和对 F 取转置的操作是交换的，即 $(\overline{F})^t=\overline{F^t}$。

即

$$(x_1,\ldots,x_n)=(x_1'\ldots,x_n')\cdot A^t,\qquad \begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}=\overline{A}\cdot\begin{pmatrix}\overline{y_1'}\\ \vdots\\ \overline{y_n'}\end{pmatrix}$$

将其代入式 (3.2.1) 得

$$(x_1',\ldots,x_n')A^tF\overline{A}\begin{pmatrix}\overline{y_1'}\\ \ldots\\ \overline{y_n'}\end{pmatrix}=(x_1',\ldots,x_n')F'\begin{pmatrix}\overline{y_1'}\\ \ldots\\ \overline{y_n'}\end{pmatrix}$$

上式对任意 $(x_1',\ldots,x_n')$ 和 $(y_1',\ldots,y_n')$ 都成立，于是我们让 $(x_1',\ldots,x_n')$ 和 $(y_1',\ldots,y_n')$ 分别取遍 $(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,\ldots,0,1)$ 即可知 F 和 F' 每个位置的元素都相等，所以 $F'=A^tF\overline{A}$。

特别地，如果 f 是 Hermitian 型，F' 和 F 是 f 在两组基下的矩阵，则容易看出

$$(F')^*=\overline{(A^tF\overline{A})^t}=\overline{A}^t\cdot\overline{F^t}\cdot\overline{A}=A^tF^*\overline{A}=F',$$

即 V 上的 Hermitian 型与 Hermitian 矩阵是一一对应的。

最后，需要说明的是，如果 f 是复向量空间 V 上的一个 Hermitian 型，则必有 $\forall\mathbf{x}\in V,\ f(\mathbf{x},\mathbf{x})\in\mathbb{R}$ (因为 $f(\mathbf{x},\mathbf{x})=\overline{f(\mathbf{x}.\mathbf{x})}$)。因此，就像我们可以在实对称双线性型对应的二次型上定义正定性那样，如果 $\forall\mathbf{x}\in V,\ f(\mathbf{x},\mathbf{x})\geqslant 0$，则称 f 是半正定的；如果 f 半正定并且 $f(\mathbf{x},\mathbf{x})=0\iff\mathbf{x}=\mathbf{0}$，则称 f 是正定的。

3.2.2 埃尔米特空间

现在我们可以定义复内积空间了。

定义 3.2.2 设 V 是 n 维复向量空间，$<\cdot,\cdot>: V\times V\to\mathbb{C}$ 是 V 上的某个给定的正定 Hermitian 型 (也称 $<\cdot,\cdot>$ 为内积)，则我们称 $(V,<\cdot,\cdot>)$ 是埃尔米特 (Hermitian) 空间或酉空间。或者说 Hermiian 空间上的内积 $<\cdot,\cdot>$ 应该满足以下性质：

1) $\forall\mathbf{x},\mathbf{y}\in V,\quad <\mathbf{x},\mathbf{y}>=\overline{<\mathbf{y},\mathbf{x}>}$;

2) $\forall\mathbf{x},\mathbf{y},\mathbf{z}\in V,\quad <\mathbf{x}+\mathbf{y},\mathbf{z}>=<\mathbf{x},\mathbf{z}>+<\mathbf{y},\mathbf{z}>$;

3) $\forall\mathbf{x},\mathbf{y}\in V$ 及 $\lambda\in\mathbb{C},\quad <\lambda\mathbf{x},\mathbf{y}>=\lambda<\mathbf{x},\mathbf{y}>$;

4) $\forall\mathbf{x}\in V,\quad <\mathbf{x},\mathbf{x}>\geqslant 0$，并且等号当且仅当 $\mathbf{x}=\mathbf{0}$ 时成立。

下面我们来看几个例子。

例 3.2.1 (1) 在 $\mathbb{C}^n$ 上，任取 $\mathbf{x}=(x_1,\ldots,x_n)^t,\mathbf{y}=(y_1,\ldots,y_n)^t\in V$，定义内积 $<\mathbf{x},\mathbf{y}>=\sum_{i=1}^n x_i\overline{y_i}$，则容易验证 $(\mathbb{C}^n,<\cdot,\cdot>)$ 是 Hermitian 空间，它被称为 n 维的标准 Hermitian 空间。

(2) 固定 $a,b\in\mathbb{R},\ a<b$，任取 $f(x),g(x)$ 是闭区间 $[a,b]$ 上的复数值连续函数，定义：

$$<f,g>=\int_a^b f(t)\overline{g(t)}\,\mathrm{d}t.$$

则容易验证这是一个内积，即 $[a,b]$ 上的复数值连续函数在此内积下构成了一个复内积空间 (注意这个空间不是有限维的)。

与实内积空间类似，我们可以定义复内积空间上的范数和距离如下：

定义 3.2.3 设 V 是复内积空间，任取 $\mathbf{x} \in V$，则我们可以定义 $\mathbf{x}$ 的范数 (或称为长度，norm) 为：$\|\mathbf{x}\| = \sqrt{<\mathbf{x}, \mathbf{x}>} \in \mathbb{R}$。于是，任取 $\mathbf{x}, \mathbf{y} \in V$，我们称 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{<\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}>}$ 为向量 $\mathbf{x}$ 和 $\mathbf{y}$ 的距离 (distance)。特别地，如果向量 $\mathbf{x}$ 的范数是 1，则称 $\mathbf{x}$ 是一个单位向量 (unit vector)。

同样我们有复内积空间上的 CBS 不等式：

定理 3.2.1 (**CBS 不等式**) 设 $(V, <\cdot, \cdot>)$ 是复内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $|<\mathbf{x}, \mathbf{y}>| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\|$，并且等号成立的充分必要条件是 $\mathbf{x} = \mathbf{0}$ 或 $\exists \lambda \in \mathbb{C}$ 使得 $\mathbf{y} = \lambda \mathbf{x}$。

证明: 如果 $\mathbf{x}$ 或 $\mathbf{y}$ 是 $\mathbf{0}$，则不等式两边都是 0，显然成立。下设 $\mathbf{x}$ 和 $\mathbf{y}$ 都不是 $\mathbf{0}$。由复数的指数形式，不妨设 $<\mathbf{x}, \mathbf{y}> = |<\mathbf{x}, \mathbf{y}>| \cdot e^{i\theta} \in \mathbb{C}$，其中 $\theta \in \mathbb{R}$。类似于欧氏空间上 CBS 不等式的证明，由于 $\forall t \in \mathbb{R}$，有 $<t\mathbf{x} + e^{i\theta}\mathbf{y}, t\mathbf{x} + e^{i\theta}\mathbf{y}> \ \geqslant 0$，将不等式左边展开得：

$$<\mathbf{x}, \mathbf{x}> t^2 - (<\mathbf{x}, \mathbf{y}> e^{-i\theta} + \overline{<\mathbf{x}, \mathbf{y}>} e^{i\theta})t + <\mathbf{y}, \mathbf{y}> \geqslant 0 \text{ 对} \forall \lambda \in \mathbb{R} \text{成立}.$$

将 $<\mathbf{x}, \mathbf{y}> = |<\mathbf{x}, \mathbf{y}>| \cdot e^{i\theta}$ 代入上式即有：

$$<\mathbf{x}, \mathbf{x}> t^2 - 2|<\mathbf{x}, \mathbf{y}>| \cdot t + <\mathbf{y}, \mathbf{y}> \geqslant 0 \text{ 对} \forall \lambda \in \mathbb{R} \text{成立}.$$

这说明上式左边的判别式非正，即

$$4|<\mathbf{x}, \mathbf{y}>|^2 - 4 <\mathbf{x}, \mathbf{x}> \cdot <\mathbf{y}, \mathbf{y}> \leqslant 0,$$

整理即得 CBS 不等式。等号成立条件的讨论与实内积空间的情形相同，留作练习。 □

于是我们在这里可以直接将实内积空间的一系列定义和性质照搬过来：

推论 3.2.1 (**三角不等式**) 设 $(V, <\cdot, \cdot>)$ 是复内积空间，任取 $\mathbf{x}, \mathbf{y} \in V$，则 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$，并且等号成立的充分必要条件为 $\exists \alpha, \beta \in \mathbb{C}$ 使得 $\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}$。

命题 3.2.1 设 $(V, <\cdot, \cdot>)$ 是复内积空间，$\|\cdot\|$ 是该内积对应的范数，则

(1) $\forall \mathbf{x} \in V$，$\|\mathbf{x}\| \geqslant 0$，并且 $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$；

(2) $\forall \lambda \in \mathbb{R}$, $\mathbf{x} \in V$，都有 $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$；

(3) $\forall \mathbf{x}, \mathbf{y} \in V$，有 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$。

命题 3.2.2 设 $(V, <\cdot, \cdot>)$ 是复内积空间，$\mathrm{d}(\cdot, \cdot)$ 是该内积对应的距离，则

(1) $\forall \mathbf{x}, \mathbf{y} \in V$，$\mathrm{d}(\mathbf{x}, \mathbf{y}) \geqslant 0$，并且等号当且仅当 $\mathbf{y} = \mathbf{x}$ 时成立；

(2) $\forall \mathbf{x}, \mathbf{y} \in V$，都有 $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \mathrm{d}(\mathbf{y}, \mathbf{x})$；

(3) $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$，有 $\mathrm{d}(\mathbf{x}, \mathbf{z}) \leqslant \mathrm{d}(\mathbf{x}, \mathbf{y}) + \mathrm{d}(\mathbf{y}, \mathbf{z})$。

定义 3.2.4 设 $(V, <\cdot, \cdot>)$ 是复内积空间，任取非零向量 $\mathbf{x}, \mathbf{y} \in V$，我们定义 $\mathbf{x}$ 和 $\mathbf{y}$ 的夹角是如下的 θ：

$$\cos\theta = \frac{|<\mathbf{x}, \mathbf{y}>|}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|},\ 0 \leqslant \theta \leqslant \frac{\pi}{2}. \qquad (\text{CBS 不等式保证了}\theta\text{的存在性.})$$

如果 $\theta = \frac{\pi}{2}$，则我们称 $\mathbf{x}, \mathbf{y}$ 正交 (orthogonal)，记作 $\mathbf{x} \perp \mathbf{y}$。特别地，我们规定 $\mathbf{0}$ 与任意 V 中的向量都正交，即 $\forall \mathbf{x}, \mathbf{y} \in V$, $\mathbf{x} \perp \mathbf{y} \iff <\mathbf{x}, \mathbf{y}> = 0$。

命题 3.2.3 (**勾股定理**) 设 $(V,<\cdot,\cdot>)$ 是复内积空间，$\mathbf{x},\mathbf{y}\in V$ 且 $\mathbf{x}\perp\mathbf{y}$，则 $\|\mathbf{x}+\mathbf{y}\|^2=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2$。

引理 3.2.1 设 $(V,<\cdot,\cdot>)$ 是复内积空间，$\mathbf{x}\in V$，$\mathbf{u}_1,\ldots,\mathbf{u}_k$ 是子空间 U 的一组基。如果 $\mathbf{x}\perp\mathbf{u}_1,\ldots,\mathbf{x}\perp\mathbf{u}_k$ 都成立，则任取 $\mathbf{u}\in U$，都有 $\mathbf{x}\perp\mathbf{u}$。我们把这种情形记作 $\mathbf{x}\perp U$。

命题 3.2.4 设 $(V,<\cdot,\cdot>)$ 是复内积空间，$\mathbf{x},\mathbf{y}\in V$，则

(1) (扩展的三角不等式)$\left|\|\mathbf{x}\|-\|\mathbf{y}\|\right|\leqslant\|\mathbf{x}\pm\mathbf{y}\|\leqslant\|\mathbf{x}\|+\|\mathbf{y}\|$；

(2) (平行四边形法则)$\|\mathbf{x}+\mathbf{y}\|^2+\|\mathbf{x}-\mathbf{y}\|^2=2(\|\mathbf{x}\|^2+\|\mathbf{y}\|^2)$；

(3) (极化恒等式)$<\mathbf{x},\mathbf{y}>=\dfrac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)+\dfrac{i}{4}(\|\mathbf{x}+i\mathbf{y}\|^2-\|\mathbf{x}-i\mathbf{y}\|^2)$。

与欧氏空间一样，我们也可以在 Hermitian 空间上定义标准正交基。

定义 3.2.5 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间。如果 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 满足：

$$\forall i\neq j\in\{1,\ldots,n\},\quad <\mathbf{v}_i,\mathbf{v}_j>=0,$$

则我们称 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 是 V 的一组正交基；如果 V 的一组基 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是正交基，并且每个 $\mathbf{e}_i$，$i\in\{1,\ldots,n\}$ 的范数都是 1，即

$$<\mathbf{e}_i,\mathbf{e}_j>=\delta_{ij},$$

则我们称 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基。

显然 $\mathbb{C}^n$ 的标准基 $\mathbf{e}_1=(1,0,\ldots,0)^t,\ldots,\mathbf{e}_n=(0,\ldots,0,1)^t$ 就是一组标准正交基。

定理 3.2.2 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 V 中一定存在一组标准正交基。

用 Gram-Schmidt 正交化方法即可证明，过程与定理 3.1.2的证明完全相同。

推论 3.2.2 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的 m 维子空间，若 U 有一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m$，则它可以扩充成 V 的一组标准正交基 $\mathbf{e}_1,\ldots,\mathbf{e}_m,\mathbf{e}_{m+1},\ldots,\mathbf{e}_n$。

定义 3.2.6 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的子空间。则

$$W=\{\mathbf{x}\in V\mid\forall\mathbf{u}\in U,\quad <\mathbf{x},\mathbf{u}>=0\}$$

是 V 的子空间，称 W 为 U 的正交补空间，记作 $W=U^{\perp}$。

定理 3.2.3 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，U 是 V 的子空间，则 $V=U\oplus U^{\perp}$ 并且 $(U^{\perp})^{\perp}=U$。

证明方法与定理 3.1.3的证明相同，留作练习。

命题 3.2.5 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是一组标准正交基。任取 $\mathbf{x}\in V$，则

(1) $\mathbf{x}=\sum_{i=1}^{n}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$，我们仍把 $<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$ 称为 $\mathbf{x}$ 在 $\mathbf{e}_i$ 方向上的**投影**；

(2) (Parseval 等式)[1] $\|\mathbf{x}\|^2=\sum_{i=1}^{n}|<\mathbf{x},\mathbf{e}_i>|^2$。

与命题 3.1.4的证明方法类似 (注意此时内积对第二个变量是共轭线性的)。

[1]无穷维的复内积空间上 Bessel 不等式仍成立。

定义 3.2.7 设 $(V, <\cdot,\cdot>_V)$ 和 $(W, <\cdot,\cdot>_W)$ 是两个 n 维 Hermitian 空间，如果存在一个线性同构 $f: V \to W$ 使得

$$<\mathbf{x}, \mathbf{y}>_V = <f(\mathbf{x}), f(\mathbf{y})>_W,\ \forall \mathbf{x}, \mathbf{y} \in V,$$

则称这两个 Hermitian 空间 V 和 W 同构。

定理 3.2.4 设 $(V, <\cdot,\cdot>)$ 是任意一个 n 维 Hermitian 空间，则 V 一定同构于 n 维标准 Hermitian 空间 $(\mathbb{C}^n, <\cdot,\cdot>_{\mathbb{C}^n})$。

与定理 3.1.4的证明方法相同，留作练习。

现在我们考虑 Hermitian 空间的对偶空间。类比欧氏空间的情形，我们自然会希望 $\forall \mathbf{v} \in V$, $\Phi_{L,\mathbf{v}}: V \to \mathbb{R}$, $\mathbf{x} \mapsto <\mathbf{v}, \mathbf{x}>$ 或 $\Phi_{R,\mathbf{v}}: V \to \mathbb{R}$, $\mathbf{x} \mapsto <\mathbf{x}, \mathbf{v}>$ (注意 $\Phi_{L,\mathbf{v}}$ 和 $\Phi_{R,\mathbf{v}}$ 是不相等的！) 是线性函数，并且 $\mathbf{v} \mapsto \Phi_{L,\mathbf{v}}$ (或$\Phi_{R,\mathbf{v}}$) 是线性同构。然而，实际情况与我们的期望有一定的差距。

定义 3.2.8 设 V 是一个复向量空间，函数 $f: V \to \mathbb{C}$ 满足：对任意 $\alpha \in \mathbb{C}$ 及 $\mathbf{x}, \mathbf{y} \in V$，有

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

则我们称 f 是 V 到 $\mathbb{C}$ 的半线性函数。我们把 V 上所有半线性映射构成的集合记作 $V^\star$，并且在 $V^\star$ 上定义自然的加法和数乘如下：

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

容易验证 $V^\star$ 在上述运算下构成了 $\mathbb{C}$ 上的向量空间，并且如果 $\dim(V) = n$，则 $\dim(V^\star) = n$。

设 V, W 是两个复向量空间，如果映射 $f: V \to W$ 满足：对 $\forall \alpha, \beta \in \mathbb{C}$ 及 $\mathbf{x}, \mathbf{y} \in V$，都有

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

则我们称 f 是 $V \to W$ 的半线性映射；如果存在半线性映射 $f: V \to W$ 是双射，则我们称 $f: V \to W$ 是 V 到 W 的半线性同构。

我们有下面的定理：

定理 3.2.5 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 $\Phi_{L,\mathbf{v}}: \mathbf{x} \mapsto <\mathbf{v}, \mathbf{x}>$ 是半线性函数，$\Phi_L: V \to V^\star$, $\mathbf{v} \mapsto \Phi_{L,\mathbf{v}}$ 是 V 到 $V^\star$ 的线性同构；$\Phi_{R,\mathbf{v}}: \mathbf{x} \mapsto <\mathbf{x}, \mathbf{v}>$ 是线性函数，$\Phi_L: V \to V^*$, $\mathbf{v} \mapsto \Phi_{R,\mathbf{v}}$ 是 V 到 V^* 的半线性同构。也就是说，V 上的每个线性函数都是 $\mathbf{x} \mapsto <\mathbf{x}, \mathbf{v}>$ 的形式，V 上的每个半线性函数都是 $\mathbf{x} \mapsto <\mathbf{v}, \mathbf{x}>$ 的形式。

与命题 3.1.5的证明方法类似，留作练习。实际上，我们还看到

$$V^* \longrightarrow V^\star,\ (\mathbf{x} \mapsto <\mathbf{x}, \mathbf{v}>) \longmapsto (\mathbf{x} \mapsto <\mathbf{v}, \mathbf{x}>)$$

是 V^* 到 $V^\star$ 的同构，我们称半线性函数 $(\mathbf{x} \mapsto <\mathbf{v}, \mathbf{x}>)$ 是线性函数 $(\mathbf{x} \mapsto <\mathbf{x}, \mathbf{v}>)$ 的共轭。

我们同样可以在 Hermitian 空间上定义内积关于一组基的度量矩阵 (Gram 矩阵)。设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，由于 $<\cdot,\cdot>$ 是 V 上的埃尔米特型，故任取 V 的一组基 $\mathbf{v}_1,\ldots,\mathbf{v}_n$，埃尔米特型 $<\cdot,\cdot>$ 对应的矩阵为

$$G=\begin{pmatrix} <\mathbf{v}_1,\mathbf{v}_1> & <\mathbf{v}_1,\mathbf{v}_2> & \cdots & <\mathbf{v}_1,\mathbf{v}_n> \\ <\mathbf{v}_2,\mathbf{v}_1> & <\mathbf{v}_2,\mathbf{v}_2> & \cdots & <\mathbf{v}_2,\mathbf{v}_n> \\ \vdots & \vdots & & \vdots \\ <\mathbf{v}_n,\mathbf{v}_1> & <\mathbf{v}_n,\mathbf{v}_2> & \cdots & <\mathbf{v}_n,\mathbf{v}_n> \end{pmatrix}$$

我们称上面的 G 为内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的**度量矩阵**或 Gram 矩阵。由 3.2.1 小节的论述可知，如果内积在基底 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 下的度量矩阵分别为 G 和 G'，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$, $A\in \mathrm{GL}_n(\mathbb{C})$, 则 $G'=A^tG\overline{A}$。显然 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基 $\iff$ 其度量矩阵为 E_n。于是，如果 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 和 $\mathbf{v}'_1,\ldots,\mathbf{v}'_n$ 都是 V 的标准正交基，并且 $(\mathbf{v}'_1,\ldots,\mathbf{v}'_n)=(\mathbf{v}_1,\ldots,\mathbf{v}_n)\cdot A$，则 $E_n=A^tE_n\overline{A}$，即 $A^t\overline{A}=E_n$，或者说 $\overline{A}^tA=E_n$，即 $A^*A=E_n$。反之，如果矩阵 A 满足 $A^*A=E_n$，$\mathbf{e}_1,\ldots,\mathbf{e}_n$ 是 V 的一组标准正交基，令 $(\mathbf{v}_1,\ldots,\mathbf{v}_n)=(\mathbf{e}_1,\ldots,\mathbf{e}_n)\cdot A$，则显然 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是 V 的基 (A 显然可逆)，并且按上面的讨论，内积在 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 下的度量矩阵为 $A^*E_nA=E_n$，即 $\mathbf{v}_1,\ldots,\mathbf{v}_n$ 也是标准正交基。这就引出了下面的酉矩阵的定义。

定义 3.2.9 设 $A\in \mathrm{GL}_n(\mathbb{C})$，如果 A 满足 $A^*A=E$，即 $A^*=A^{-1}$，则称 A 为酉矩阵 (unitary matrix)。容易验证 $\mathbb{C}$ 上所有 n 阶酉矩阵在矩阵乘法运算下构成了一个群，我们把这个群称为酉群 (unitary group)，记作 $U(n)$。显然 $O(n)$ 是 $U(n)$ 的子群。

容易看出酉矩阵就是标准正交基之间的过渡矩阵。同样我们也可以定义酉算子如下：

定义 3.2.10 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}$ 是 V 上的可逆的线性算子。如果 $\mathcal{A}$ 还满足：对任意的 $\mathbf{x},\mathbf{y}\in V$，$<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$，则我们称 $\mathcal{A}$ 是 V 上的酉算子 (unitary operator) 或酉变换。

容易看出酉算子一定是 Hermitian 空间的自同构。我们同样有以下定理：

定理 3.2.6 设 $(V,<\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的线性算子，则以下条件等价：

(1) $\mathcal{A}$ 是酉算子；

(2) $\mathcal{A}$ 在标准正交基下对应的矩阵是酉矩阵；

(3) $\mathcal{A}$ 将标准正交基映到标准正交基；

(4) $\forall \mathbf{x}\in V,\ \|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|$。

与定理 3.1.5的证明方法类似，留作练习。

推论 3.2.3 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，则 V 上所有酉算子在算子复合运算下构成群，记作 $U(V)$，则 $U(V)$ 同构于酉群 $U(n)$。

留作练习。

我们看到，Hermitian 空间的性质与欧氏空间的性质是极为类似的，并且我们可以把欧氏空间看成 Hermitian 空间“限制”在实数域上的特例。

3.2.3 赋范向量空间与度量空间

还记得我们在本章开头提到的目标吗？我们希望在向量空间上复原“通常”的距离结构。现在我们已经知道了如何用内积诱导出范数和距离，但我们希望用公理化的方法定义范数和距离。这就是本小节的主要内容。

定义 3.2.11 设 E 是一个非空集合，函数：$\mathrm{d}: E\times E\to[0,+\infty)$ 满足以下性质：

(1) 正定性：$\forall u,v\in E$，$\mathrm{d}(u,v)\geqslant 0$，并且等号当且仅当 $u=v$ 时成立；

(2) 对称性：$\forall u,v\in E$，$\mathrm{d}(u,v)=\mathrm{d}(v,u)$；

(3) 三角不等式：$\forall u,v,w\in E$，$\mathrm{d}(u,w)\leqslant \mathrm{d}(u,v)+\mathrm{d}(v,w)$。

则我们称函数 d 是集合 E 上的一个距离或度量 (distance/metric) 函数，称 (E,d) 是一个度量空间 (metric space)。我们把 E 中的任意一个元素称为一个点 (point)，称 $d(u,v)$ 为点 u 和 v 之间的距离。

下面我们给出一些度量空间的例子。

例 3.2.2 (1) 实 (或) 复的内积空间是度量空间，其度量函数可以由内积诱导出来，即 $\mathrm{d}(\mathbf{x},\mathbf{y})=\sqrt{<\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}>}$。

(2) 在集合 $C[a,b]$ (闭区间 $[a,b]$ 上的所有连续函数所构成的集合) 上可以定义不同的距离函数：$\forall f(x),g(x)\in C[a,b]$，以下函数

$$d_1(f,g)=\sqrt{\int_a^b |f(t)-g(t)|^2\,\mathrm{d}t};$$
$$d_2(f,g)=\int_a^b |f(t)-g(t)|\,\mathrm{d}t;$$
$$d_3(f,g)=\max_{a\leqslant t\leqslant b}|f(t)-g(t)|.$$

都是 $C[a,b]$ 上的度量函数。

(3) 设 p 是大于 1 的实数，定义集合

$$l^p(\mathbb{R})=\left\{\mathbf{x}=\{x_n\}_{n=1}^{\infty}\;\middle|\; x_n\in\mathbb{R},\ \sum_{n=1}^{\infty}|x_n|^p<\infty\right\}$$

容易验证 $l^p(\mathbb{R})$ 在通常的加法和数乘下是 $\mathbb{R}$ 上的无穷维向量空间。设 $\mathbf{x}=\{x_n\}_{n=1}^{\infty}$，$\mathbf{y}=\{y_n\}_{n=1}^{\infty}\in l^p(\mathbb{R})$，我们可以在 $l^p(\mathbb{R})$ 上定义距离函数：

$$\mathrm{d}(\mathbf{x},\mathbf{y})=(\sum_{n=1}^{\infty}|x_n-y_n|^p)^{\frac{1}{p}}$$

则 $l^p(\mathbb{R})$ 在此度量函数下构成了一个度量空间。该度量函数的三角不等式被称为闵可夫斯基 (Minikowski) 不等式，其证明留作习题。我们在以后的泛函分析课程中会经常用到它。

(4) 对任意集合 E，我们可以定义离散度量：$\forall a,b\in E$，$\mathrm{d}(a,b)=\begin{cases}1, & a\neq b\\ 0, & a=b\end{cases}$。验证 (E,d) 是一个度量空间是容易的，留作练习。以后在拓扑学课程中我们会知道，离散度量所诱导的拓扑称为离散拓扑，其性质是十分简单的。

度量空间上有比较好的拓扑和分析结构，因此我们可以进行更复杂的操作。

定义 3.2.12 设 (E, d) 是度量空间，对 $\forall a \in E$ 及 $r > 0$，我们定义

$$
\begin{aligned}
B(a, r) &= \{x \in E \mid \mathrm{d}(a, x) < r\}, \\
\overline{B}(a, r) &= \{x \in E \mid \mathrm{d}(a, x) \leqslant r\}, \\
S(a, r) &= \{x \in E \mid \mathrm{d}(a, x) = r\}.
\end{aligned}
$$

我们把 $B(a, r), \overline{B}(a, r), S(a, r)$ 分别称为以 a 为中心，r 为半径的开球、闭球、和球面。

定义 3.2.13 设 (E, d) 是度量空间，序列 $\{x_n\}_{n=1}^{\infty} \subset E$。

(1) 如果 $\forall \varepsilon$, $\exists N = N(\varepsilon) \in \mathbb{N}$ 使得对 $\forall m, n > N$ 都有 $\mathrm{d}(x_m, x_n) < \varepsilon$，则称 $\{x_n\}$ 是柯西 (Cauchy) 序列。

(2) 如果存在 $x_0 \in E$ 满足 $\forall \varepsilon$, $\exists N = N(\varepsilon) \in \mathbb{N}$ 使得对 $\forall n > N$ 都有 $\mathrm{d}(x_n, x_0) < \varepsilon$ (即 $\lim\limits_{n \to \infty} \mathrm{d}(x_n, x_0) = 0$)，则称序列 $\{x_n\}$ 按距离收敛到 x_0，或称 x_0 是序列 $\{x_n\}$ 的极限。

如果 (E, d) 中每一个柯西序列都能收敛到 E 中的某个元素，则我们称 (E, d) 是一个完备度量空间。

例 3.2.3 $\mathbb{R}$ 和 $\mathbb{C}$ 在通常的度量下显然 (如果大家还记得数学分析课程中实数理论相关的内容的话) 都是完备度量空间。实际上，$\mathbb{R}^n$ 和 $\mathbb{C}^n$ 在下面的度量函数下都是完备度量空间 (设 $\mathbf{x} = (x_1, \ldots, x_n)^t$, $\mathbf{y} = (y_1, \ldots, y_n)^t$ 是 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 中的两个点)：

$$
\begin{aligned}
\mathrm{d}_2(\mathbf{x}, \mathbf{y}) &= \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}; \\
\mathrm{d}_1(\mathbf{x}, \mathbf{y}) &= \sum_{i=1}^{n} |x_i - y_i|; \\
\mathrm{d}_\infty(\mathbf{x}, \mathbf{y}) &= \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}.
\end{aligned}
$$

这是因为这三个度量之间有如下的关系：对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$(或 $\mathbb{C}^n$)，有

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \mathrm{d}_1(\mathbf{x}, \mathbf{y}) \leqslant \mathrm{d}_2(\mathbf{x}, \mathbf{y}) \leqslant \sqrt{n}\, \mathrm{d}_1(\mathbf{x}, \mathbf{y}) \\
\mathrm{d}_\infty(\mathbf{x}, \mathbf{y}) \leqslant \mathrm{d}_2(\mathbf{x}, \mathbf{y}) \leqslant \sqrt{n}\, \mathrm{d}_\infty(\mathbf{x}, \mathbf{y})
\end{aligned}
\tag{3.2.2}
$$

设 $\{\mathbf{x}_k = (x_{k1}, \cdots, x_{kn})^t\}_{k=1}^{\infty}$ 是 d_∞ 距离下的柯西序列，即 $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ 使得 $\forall l, m > N$，都有 $\max\{|x_{l1} - x_{m1}|, |x_{l2} - x_{m2}|, \ldots, |x_{ln} - x_{mn}|\} < \varepsilon$。这说明 $\{x_{k1}\}_{k=1}^{\infty}$, $\{x_{k2}\}_{k=1}^{\infty}, \cdots\cdots, \{x_{kn}\}_{k=1}^{\infty}$ 都是 $\mathbb{R}$(或 $\mathbb{C}$) 中的柯西序列，由 $\mathbb{R}$(或 $\mathbb{C}$) 的完备性可设 $\lim\limits_{n \to \infty} x_{k1} = y_1, \cdots\cdots, \lim\limits_{n \to \infty} x_{kn} = y_n$，令 $\mathbf{y} = (y_1, \ldots, y_n)^t$，则容易验证 $\lim\limits_{k \to \infty} \mathrm{d}_\infty(\mathbf{x}_k, \mathbf{y}) = 0$，即序列 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 按 d_∞ 距离收敛到 $\mathbf{y}$，这就证明了 $\mathbb{R}^n$(或 $\mathbb{C}^n$) 在 d_∞ 距离下是完备度量空间。利用式 (3.2.2) 给出的三个度量函数之间的大小关系即可验证 $\mathbb{R}^n$(或 $\mathbb{C}^n$) 在这三个距离下都是完备的。

例 3.2.4 $C[a, b]$ 在度量函数 $d_3(f, g) = \max\limits_{a \leqslant t \leqslant b} |f(t) - g(t)|$ 下是完备度量空间，但在 $d_2(f, g) = \int_a^b |f(t) - g(t)|\, \mathrm{d}t$ 下就不是完备的。证明完全是分析的技术，可参考《数学分析 · 第二卷》卓里奇，高等教育出版社的第九章 §5 的例 4 和例 5。

我们已经给出了度量的公理化定义，下面我们给出范数的公理化定义。

定义 3.2.14 设 V 是一个实或复的向量空间 (不一定是有限维的)，在 V 上固定一个满足如下性质的函数 $\|\cdot\|: V \to [0,+\infty)$：

(1) $\forall \mathbf{x} \in V$，$\|\mathbf{x}\| \geqslant 0$，并且 $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$；

(2) $\forall \lambda \in \mathbb{R}$，$\mathbf{x} \in V$，都有 $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$；

(3) $\forall \mathbf{x},\mathbf{y} \in V$，有 $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$。

则我们称 $(V,\|\cdot\|)$ 是一个赋范向量空间，$\|\cdot\|$ 被称为 V 上的范数。

显然赋范向量空间一定是度量空间，因为范数诱导了一个自然的距离 $\mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$ (验证这确实是一个距离留作练习)。如果一个赋范向量空间还是完备距离空间，则我们称它为完备赋范空间 (Banach space)。特别地，如果一个实或复的内积空间作为度量空间是完备的，则我们称它为 Hilbert 空间。

本小节介绍的这些空间之间的关系可以用下图表示：

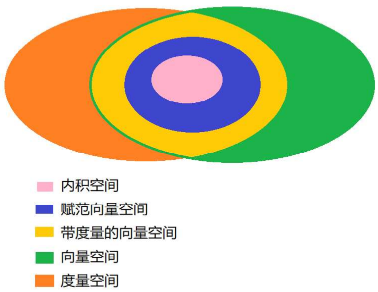

图 3.1: 各种空间的关系 (注意：这里内积空间指实或复内积空间，不包括辛空间。)

思考题 3.2.1 给出以下反例：

(1) 一个赋范向量空间不是内积空间；

(2) 一个带度量的向量空间不是赋范向量空间；

(3) 不完备的内积空间；

(4) 不完备的赋范向量空间；

(5) 不完备的度量空间。

最后，显然欧氏空间和 Hermitian 空间一定是完备的。我们可以在其上讨论收敛性。

定理 3.2.7 设 $(V,<\cdot,\cdot>)$ 是欧氏空间或 Hermitian 空间，则 V 中的序列 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 收敛到 $\mathbf{x} \in V \iff \lim\limits_{k\to\infty} <\mathbf{x}-\mathbf{x}_k,\mathbf{y}>= 0$ 对任意的 $\mathbf{y} \in V$ 成立。

证明: $(\Longrightarrow)$ 如果 $\{\mathbf{x}_k\}_{k=1}^{\infty}$ 收敛到 $\mathbf{x} \in V$，即 $\mathrm{d}(\mathbf{x},\mathbf{x}_k) = \|\mathbf{x}-\mathbf{x}_k\| \xrightarrow{k\to\infty} 0$，故对任意的 $\mathbf{y} \in V$，由 CBS 不等式有

$$<\mathbf{x}-\mathbf{x}_k,\mathbf{y}>\leqslant \|\mathbf{x}-\mathbf{x}_k\|\cdot\|\mathbf{y}\| \xrightarrow{k\to\infty} 0$$

即 $\lim_{k\to\infty} < \mathbf{x}-\mathbf{x}_k, \mathbf{y} >= 0$。
($\Longleftarrow$) 取 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，则有

$$\lim_{k\to\infty} < \mathbf{x}-\mathbf{x}_k, \mathbf{e}_i >= 0,\ \forall i \in \{1, \ldots, n\}.$$

于是 $\|\mathbf{x}-\mathbf{x}_k\|^2 = \sum_{i=1}^n < \mathbf{x}-\mathbf{x}_k, \mathbf{e}_i > \xrightarrow{k\to\infty} 0$ (这里用到了 V 是有限维的)，所以 $\{\mathbf{x}_k\}_{k=1}^\infty$ 收敛到 $\mathbf{x}$。□

本小节的内容只是关于度量空间和赋范向量空间的最简单的介绍，更详细的内容读者可以参考泛函分析的标准教材，如《A Course in Functional Analysis》, Jorn B.Conway, GTM96 或《泛函分析讲义》，张恭庆，北京大学出版社。

3.3 内积空间上的线性算子 I：自伴随算子

本节和接下来两节中我们所说的内积空间特指欧氏空间和 Hermitian 空间 (不包括辛空间)，特别地，我们要求空间都是有限维的。此外，我们提到的域 $\mathbb{K}$ 都是 $\mathbb{R}$ 或 $\mathbb{C}$。现在我们开始讨论内积空间上的线性算子。我们的最终目标是 3.5 节的谱定理，也就是说，我们需要寻找一类内积空间中的算子，它们在某组基 (更进一步，这组基还是标准正交基) 下的矩阵是对角矩阵。用矩阵的语言来说，我们需要寻找一类矩阵，它们能够合同[1]并相似到对角矩阵 (即我们希望 $\exists T \in U(n)$ (或$O(n)$) 使得 T^*AT (或T^tAT) $= T^{-1}AT = \Lambda$，Λ 是对角阵)。

直接讨论上面的问题是困难的。为此，我们从 (半) 双线性型和线性算子的关系入手，先通过型定义伴随算子的概念，然后我们讨论一种特殊的可对角化的算子：自伴随算子。

定义 3.3.1 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A}$ 是 V 上的线性算子，则我们可以定义：

$$f_{\mathcal{A}}: V \times V \longrightarrow \mathbb{K}\ (\mathbb{K} = \mathbb{R}\text{或}\mathbb{C})$$
$$(\mathbf{x}, \mathbf{y}) \longmapsto <\mathcal{A}\mathbf{x}, \mathbf{y}> .$$

容易验证：如果 V 是欧氏空间，则 $f_{\mathcal{A}}$ 是双线性型；如果 V 是 Hermitian 空间，则 $f_{\mathcal{A}}$ 是半双线性型型。由于双线性型和半双线性型型的性质类似，因此我们把欧氏空间上的双线性型和 Hermitian 空间上的半双线性型型统称为 $\theta-$ 线性型 ($\theta = 2$ 表示欧氏空间上的双线性型，$\theta = \frac{3}{2}$ 表示 Hermitian 空间上的半双线性型)。我们采用统一的记号 $L_\theta(V, \mathbb{K})$ 来表示 V 上的 $\theta-$ 线性型 (同样地，$\theta = 2, \mathbb{K} = \mathbb{R}$ 表示欧氏空间上的双线性型，$\theta = \frac{3}{2}, \mathbb{K} = \mathbb{C}$ 表示 Hermitian 空间上的半双线性型)。如果 $\dim(V) = n$，则显然 $L_\theta^+(V, \mathbb{K})$ 是 $\mathbb{R}$ 或 $\mathbb{C}$ 上的 n^2 维向量空间 (定理 1.6.1或 3.2.1 小节的论证)。

定理 3.3.1 设 $(V, <\cdot,\cdot>)$ 是内积空间，定义如下的映射 Φ：

$$\Phi : \mathrm{End}(V) \longrightarrow L_\theta(V, \mathbb{K})$$
$$\mathcal{A} \longmapsto f_{\mathcal{A}}\quad (f_{\mathcal{A}}\text{的定义见于定义 3.3.1}).$$

则 Φ 是 $\mathrm{End}(V)$ 到 $L_\theta(V, \mathbb{K})$ 的线性同构。

证明: 首先 Φ 是线性映射：任取 $\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$ 及 $\alpha, \beta \in \mathbb{K}(= \mathbb{R}\text{或}\mathbb{C})$，则对 $\forall \mathbf{x}, \mathbf{y} \in V$，我们有

$$<(\alpha\mathcal{A} + \beta\mathcal{B})\mathbf{x}, \mathbf{y}> = \alpha <\mathcal{A}\mathbf{x}, \mathbf{y}> + \beta <\mathcal{B}\mathbf{x}, \mathbf{y}>$$

因此 $f_{\alpha\mathcal{A}+\beta\mathcal{B}} = \alpha f_{\mathcal{A}} + \beta f_{\mathcal{B}}$，即 Φ 是线性映射。

其次 Φ 是单射：只需证 $\ker(\Phi) = \{O\}$。如果 $f_{\mathcal{A}} = 0$，即任取 $\mathbf{x}, \mathbf{y} \in V$，$<\mathcal{A}\mathbf{x}, \mathbf{y}> = 0$，则必有 $\mathcal{A}\mathbf{x} = \mathbf{0}$，$\forall \mathbf{x} \in V$，此即 $\mathcal{A} = O$，即 $\ker(\Phi) = \{O\}$。

最后 Φ 是满射：这只需注意到 $\dim(\mathrm{End}(V)) = \dim(L_\theta(V, \mathbb{K})) = n^2$ 即可证明。

综上所述，Φ 是线性同构。 □

下面我们考虑在同一组标准正交基下算子 $\mathcal{A}$ 和 $\theta-$ 线性型 $f_{\mathcal{A}}$ 所对应的矩阵之间的关系。

推论 3.3.1 条件同定理 3.3.1，设 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 V 的一组标准正交基，$\mathcal{A} \in \mathrm{End}(V)$ 在此基底下的矩阵为 A，则 $f_{\mathcal{A}}$ 在此基底下的矩阵为 A^t。

[1]复矩阵则变成 “共轭合同”，即 $A \in M_n(\mathbb{C})$ 与 T^*AT 共轭合同。

证明: 设 $A = (\mathbf{A}^{(1)}, \dots, \mathbf{A}^{(n)}) = (a_{ij})_{n\times n}$，则对每个 $i \in \{1, \dots, n\}$，有

$$\mathcal{A}\mathbf{e}_i = (\mathbf{e}_1, \dots, \mathbf{e}_n) \cdot \mathbf{A}^{(i)} = \sum_{k=1}^{n} a_{ki}\mathbf{e}_k,$$

所以 $< \mathcal{A}\mathbf{e}_i, \mathbf{e}_j >= a_{ji}$, $\forall i, j \in \{1, \dots, n\}$。而 $f_{\mathcal{A}}$ 在 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的矩阵就是 $(< \mathcal{A}\mathbf{e}_i, \mathbf{e}_j >)_{n\times n} = (a_{ji})_{n\times n}$，此即 $f_{\mathcal{A}}$ 在此基底下的矩阵为 A^t。 □

现在我们可以定义一个算子的伴随算子了。

定理 3.3.2 设 $(V, < \cdot, \cdot >)$ 是内积空间，$\mathcal{A}$ 是 V 上的线性算子，则存在唯一的 $\mathcal{B} \in \mathrm{End}(V)$ 使得 $\forall \mathbf{x}, \mathbf{y} \in V$，有 $< \mathbf{x}, \mathcal{B}\mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} >$。我们称 $\mathcal{B}$ 为 $\mathcal{A}$ 的 (Hilbert) 伴随算子 (或共轭算子，adjoint operator)，记作 $\mathcal{B} = \mathcal{A}^*$。

证明: 先证存在性：任取 $\mathbf{y} \in V$，我们只要给出 $\mathcal{B}\mathbf{y} \in V$ 的取法就证明了 $\mathcal{B}$ 的存在性。首先，$\mathbf{y}$ 给出了一个 V 上的线性函数 f：

$$\begin{aligned} f \; : \; V &\longrightarrow \mathbb{K} \\ \mathbf{x} &\longmapsto < \mathcal{A}\mathbf{x}, \mathbf{y} > . \end{aligned}$$

而由定理 3.1.5和定理 3.2.5，我们知道一定存在唯一的 $\mathbf{v} \in V$ 使得 f 是 $\mathbf{x} \mapsto < \mathbf{x}, \mathbf{v} >$ 的形式。我们定义 $\mathcal{B}\mathbf{y} = \mathbf{v}$，则首先 $\mathcal{B}$ 确实是一个映射 ($\mathbf{v}$ 的存在唯一性)；其次 $\mathcal{B}$ 是线性的，这是因为：如果 $\mathcal{B}\mathbf{y}_1 = \mathbf{v}_1$，$\mathcal{B}\mathbf{y}_2 = \mathbf{v}_2$，即 $< \mathcal{A}\mathbf{x}, \mathbf{y}_1 >=< \mathbf{x}, \mathbf{v}_1 >$，$< \mathcal{A}\mathbf{x}, \mathbf{y}_2 >=< \mathbf{x}, \mathbf{v}_2 >$ 对 $\forall \mathbf{x} \in V$ 成立，则对任意的 $\alpha, \beta \in \mathbb{K}$，有

$$\begin{aligned} < \mathcal{A}\mathbf{x}, \alpha\mathbf{y}_1 + \beta\mathbf{y}_2 > &= \overline{\alpha} < \mathcal{A}\mathbf{x}, \mathbf{y}_1 > + \overline{\beta} < \mathcal{A}\mathbf{x}, \mathbf{y}_2 > \\ &= \overline{\alpha} < \mathbf{x}, \mathbf{v}_1 > + \overline{\beta} < \mathbf{x}, \mathbf{v}_2 > \\ &=< \mathbf{x}, \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 > \end{aligned}$$

这就说明 $\mathcal{B}(\alpha\mathbf{y}_1 + \beta\mathbf{y}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$，即 $\mathcal{B}$ 是线性算子。

再证唯一性：设 $\mathcal{B}' \in \mathrm{End}(V)$ 也满足 $\forall \mathbf{x}, \mathbf{y} \in V$, $< \mathbf{x}, \mathcal{B}'\mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} >$，那么，对任意的 $\mathbf{x}, \mathbf{y} \in V$，有 $< \mathbf{x}, (\mathcal{B}' - \mathcal{B})\mathbf{y} >= 0$，这说明 $(\mathcal{B}' - \mathcal{B})\mathbf{y} = \mathbf{0}$ 对任意的 $\mathbf{y} \in V$ 成立，此即 $\mathcal{B}' = \mathcal{B}$，唯一性成立。 □

推论 3.3.2 条件同定理 3.3.2，设 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 V 的一组标准正交基，$\mathcal{A} \in \mathrm{End}(V)$ 在此基底下的矩阵为 A，则伴随算子 $\mathcal{A}^*$ 在此基底下的矩阵为 $\overline{A^t} = A^*$。

证明: 设 $\mathcal{A}^*$ 在标准正交基 $\mathbf{e}_1, \dots, \mathbf{e}_n$ 下的矩阵为 $B = (b_{ij})_{n\times n}$，则对每个 $j \in \{1, \dots, n\}$，有

$$\mathcal{A}^*\mathbf{e}_j = \sum_{k=1}^{n} b_{kj}\mathbf{e}_k,$$

所以

$$< \mathcal{A}\mathbf{e}_i, \mathbf{e}_j >=< \mathbf{e}_i, \mathcal{A}^*\mathbf{e}_j >= \overline{b_{ij}}, \; \forall i, j \in \{1, \dots, n\}.$$

另一方面，由推论 3.3.1，设 $\mathcal{A}$ 在此基底下的矩阵是 $A = (a_{ij})_{n\times n}$，则 $< \mathcal{A}\mathbf{e}_i, \mathbf{e}_j >= a_{ji}$。这说明 $a_{ji} = \overline{b_{ij}}$, $\forall i, j \in \{1, \dots, n\}$，即 $B = A^*$。 □

既然每个线性算子在内积空间上都存在唯一的伴随算子，那么我们可以考察“取伴随”这个运算的性质。

定理 3.3.3 设 $(V, <\cdot,\cdot>)$ 是内积空间，定义取伴随的映射 $*: \mathrm{End}(V) \to \mathrm{End}(V)$，$\mathcal{A} \mapsto \mathcal{A}^*$，则 $*$ 映射满足以下性质：对任意的 $\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$ 及 $\lambda \in \mathbb{K}$，有

(1) $(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$；

(2) $(\lambda\mathcal{A})^* = \overline{\lambda}\mathcal{A}^*$；

(3) $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$；

(4) $(\mathcal{A}^*)^* = \mathcal{A}$。

证明是简单的，留作练习。实际上，我们可以将以上性质抽象出来，作为 $*-$ 代数的定义 (当然，我们还会附加一些范数的要求)。更详细的论述会在算子代数课程中介绍，读者可以参考《A Short Course on Spectral Theory》, William Arveson, GTM209 以及更深入的《An Invitation to C*-Algebra》, William Arveson, GTM39。

下面我们定义几类特殊的线性算子。

定义 3.3.2 设 $(V, <\cdot,\cdot>)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$，则

(1) 如果 $\mathcal{A}^* = \mathcal{A}$，则称 $\mathcal{A}$ 为自伴随 (self-adjoint) 算子或 Hermitian 算子；如果 $\mathcal{A}^* = -\mathcal{A}$，则称 $\mathcal{A}$ 为斜自伴随 (skew self-adjoint) 算子或斜 Hermitian 算子 (也有的书译成反 Hermitian 算子)。欧氏空间上的自伴随算子也称为对称算子。

(2) 如果 $\forall \mathbf{x}, \mathbf{y} \in V$，都有 $<\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{x}> = <\mathbf{x}, \mathbf{x}>$，则称 $\mathcal{A}$ 为保距 (isometric) 算子。保距算子实际上是欧氏空间上的正交算子和 Hermitian 空间上的酉算子的统称，我们会在下一节中证明这一点。

(3) 如果 $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$，则称 $\mathcal{A}$ 为正规 (normal) 算子。

对应地，设 $A \in M_n(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}$或$\mathbb{C}$)，则

(1) 如果 $A^* = A$，则称 A 为自伴随矩阵或 Hermitian 矩阵；如果 $A^* = -A$，则称 $\mathcal{A}$ 为斜自伴随矩阵或斜 Hermitian 矩阵 (实矩阵则分别是对称矩阵和斜对称矩阵)。

(2) 如果 $A^*A = AA^* = E$，则称 A 为酉矩阵 (若 A 是实矩阵则称为正交矩阵)。

(3) 如果 $AA^* = A^*A$，则称 A 为正规矩阵。

显然上面的每一种算子在 V 的任意一组标准正交基下的矩阵就是下面对应类型的矩阵。

本节我们主要考虑自伴随算子 (和矩阵) 的情形，接下来的两节我们会依次考虑保距算子和正规算子的情形。

以下是本节的主要结论：

定理 3.3.4 设 $(V, <\cdot,\cdot>)$ 是 n 维内积空间，$\mathcal{A} \in \mathrm{End}(V)$ 是自伴随算子，则存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}$ 在这组基下的矩阵是实对角矩阵，即

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad \lambda_1, \ldots, \lambda_n \in \mathbb{R}(\lambda_1, \ldots, \lambda_n\text{中可以有相等的元素}).$$

我们也可以用矩阵语言将上述定理重新叙述如下：

定理 3.3.4′ 设 $A \in M_n(\mathbb{C})$ 满足 $A^* = A$(实矩阵则是 $A^t = A$)，则存在酉矩阵 (实情形则是正交矩阵)B 使得

$$B^*AB(\text{或}B^tAB) = B^{-1}AB = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},\ \ \lambda_1,\ldots,\lambda_n \in \mathbb{R}(\lambda_1,\ldots,\lambda_n\text{中可以有相等的元素}).$$

为了证明上述定理，我们先给出下面的引理。为了证明的方便，我们先只考虑 Hermitian 空间上的自伴随算子，欧氏空间上对称算子的情形我们会在稍后给出说明。

引理 3.3.1 设 $(V, <\cdot,\cdot>)$ 是 Hermitian 空间，$\mathcal{A} \in \mathrm{End}(V)$ 是自伴随算子，则

(1) $\forall p(t) \in \mathbb{R}[t]$，线性算子 $p(\mathcal{A})$ 也是自伴随算子。特别地，对任意的 $\lambda \in \mathbb{R}$，有 $\mathcal{A} - \lambda\mathcal{E}$ 是自伴随算子。

(2) 自伴随算子一定有特征值和特征向量，并且其所有特征值都是实数。

(3) 如果 W 是 $\mathcal{A}-$ 子空间，则 $W^\perp$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 都是自伴随算子。

(4) $\ker(\mathcal{A}) = (\mathrm{im}(\mathcal{A}^*))^\perp$，$\mathrm{im}(\mathcal{A}) = (\ker(\mathcal{A}^*))^\perp$。

(5) $\forall k \in \mathbb{Z}^+$，$\ker(\mathcal{A}^k) = \ker(\mathcal{A})$。

(6) 如果 λ, μ 是 $\mathcal{A}$ 的两个不同的特征值，$\mathbf{x}$ 和 $\mathbf{y}$ 分别是 λ 和 μ 对应的特征向量，则 $\mathbf{x} \perp \mathbf{y}$。

证明: (1) 首先，由定理 3.3.3的 (3) 可知 $\forall k \in \mathbb{Z}^+$，有 $(\mathcal{A}^k)^* = (\mathcal{A}^*)^k$。不妨设 $p(t) = a_0 + a_1 t + \cdots + a_n t^n \in \mathbb{R}[t]$，则由于 $*$ 是半线性映射 (定理 3.3.3的 (1) 和 (2))，我们有

$$[p(\mathcal{A})]^* = \overline{a_0}\mathcal{E} + \overline{a_1}\mathcal{A}^* + \cdots + \overline{a_n}(\mathcal{A}^*)^n = a_0\mathcal{E} + a_1\mathcal{A} + \cdots + a_n\mathcal{A}^n = p(\mathcal{A}).$$

特例是显然的。

(2) 设 $\chi_{\mathcal{A}}(t) \in \mathbb{C}[t]$ 是 $\mathcal{A}$ 的特征多项式，则 $\chi_{\mathcal{A}}(t)$ 在复数域 $\mathbb{C}$ 上一定有根，从而一定有特征值和特征向量。下面我们证明 $\chi_{\mathcal{A}}(t)$ 的所有根都是实根。为此，任取 $\lambda \in \mathbb{C}$ 是 $\chi_{\mathcal{A}}(t)$ 的根，$\mathbf{v} \in V$ 是 λ 对应的非零特征向量。则

$$\lambda < \mathbf{v}, \mathbf{v} > = < \lambda\mathbf{v}, \mathbf{v} > = < \mathcal{A}\mathbf{v}, \mathbf{v} > = < \mathbf{v}, \mathcal{A}\mathbf{v} > = < \mathbf{v}, \lambda\mathbf{v} > = \overline{\lambda} < \mathbf{v}, \mathbf{v} > .$$

由于 $< \mathbf{v}, \mathbf{v} > \neq 0$，故 $\lambda = \overline{\lambda}$，即 $\lambda \in \mathbb{R}$。

(3) 先证 $W^\perp$ 是 $\mathcal{A}-$ 子空间。任取 $\mathbf{x} \in W^\perp$，我们需要证明 $\mathcal{A}\mathbf{x} \in W^\perp$，即 $\mathcal{A}\mathbf{x} \perp W$。任取 $\mathbf{y} \in W$，则由 W 是 $\mathcal{A}-$ 子空间可知 $\mathcal{A}\mathbf{y} \in W$，于是 $< \mathcal{A}\mathbf{x}, \mathbf{y} > = < \mathbf{x}, \mathcal{A}\mathbf{y} > = 0$，即 $\mathcal{A}\mathbf{x} \perp W$ 成立。

$\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 都是自伴随算子是显然的。

(4) 先证明 $\ker(\mathcal{A}) = (\mathrm{im}(\mathcal{A}^*))^\perp$。任取 $\mathbf{x} \in \ker(\mathcal{A})$，则对任意的 $\mathcal{A}^*\mathbf{y} \in \mathrm{im}(\mathcal{A}^*)$，我们有

$$< \mathbf{x}, \mathcal{A}^*\mathbf{y} > = < \mathcal{A}\mathbf{x}, \mathbf{y} > = < \mathbf{0}, \mathbf{y} > = 0,$$

即 $\mathbf{x} \perp \mathrm{im}(\mathcal{A}^*)$，所以 $\mathbf{x} \in (\mathrm{im}(\mathcal{A}^*))^\perp$，$\ker(\mathcal{A}) \subset (\mathrm{im}(\mathcal{A}^*))^\perp$。反过来，任取 $\mathbf{z} \in (\mathrm{im}(\mathcal{A}^*))^\perp$，即 $\mathbf{z} \perp \mathcal{A}^*\mathbf{y}$，$\forall \mathbf{y} \in V$。这说明

$$< \mathcal{A}\mathbf{z}, \mathbf{y} > = < \mathbf{z}, \mathcal{A}^*\mathbf{y} > = 0,\ \ \forall \mathbf{y} \in V$$

所以只能是 $\mathcal{A}\mathbf{z} = \mathbf{0}$，即 $\mathbf{z} \in \ker(\mathcal{A})$，$\ker(\mathcal{A}) \supset (\mathrm{im}(\mathcal{A}^*))^\perp$。所以 $\ker(\mathcal{A}) = (\mathrm{im}(\mathcal{A}^*))^\perp$。以 $\mathcal{A}^*$ 替换上式中的 $\mathcal{A}$ 可知 $\ker(\mathcal{A}^*) = (\mathrm{im}(\mathcal{A}^{**}))^\perp = \mathrm{im}(\mathcal{A})^\perp$，即 $\mathrm{im}(\mathcal{A}) = (\ker(\mathcal{A}^*))^\perp$。

(5) 我们先证明：$\forall k \in \mathbb{Z}^+$，$\ker(\mathcal{A}^k) = \ker(\mathcal{A}^{k+1})$。首先 $\ker(\mathcal{A}^k) \subset \ker(\mathcal{A}^{k+1})$ 是显然成立的，因此我们只需要证明 $\ker(\mathcal{A}^k) \supset \ker(\mathcal{A}^{k+1})$ 即可。任取 $\mathbf{x} \in \ker(\mathcal{A}^{k+1})$，我们有 $\mathcal{A}^{k+1}\mathbf{x} = 0$，因此

$$\|\mathcal{A}^k\mathbf{x}\|^2 =< \mathcal{A}^k\mathbf{x}, \mathcal{A}^k\mathbf{x} >=< \mathcal{A}^{k+1}\mathbf{x}, \mathcal{A}^{k-1}\mathbf{x} >= 0,$$

这说明 $\mathcal{A}^k\mathbf{x} = \mathbf{0}$, 即 $\mathbf{x} \in \ker(\mathcal{A}^k)$, $\ker(\mathcal{A}^k) \supset \ker(\mathcal{A}^{k+1})$。特别地，取 $k = 1$ 即有 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2)$，因此 $\ker(\mathcal{A}) = \ker(\mathcal{A}^2) = \cdots = \ker(\mathcal{A}^k)$, $\forall k \in \mathbb{Z}^+$。

(6) 首先由 (2) 可知 λ 和 μ 都是实数。利用特征向量的定义和内积的半双线性性质，我们有

$$\begin{aligned}\lambda < \mathbf{x}, \mathbf{y} > &=< \lambda\mathbf{x}, \mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} > \\ &=< \mathbf{x}, \mathcal{A}\mathbf{y} >=< \mathbf{x}, \mu\mathbf{y} > \\ &= \overline{\mu} < \mathbf{x}, \mathbf{y} >= \mu < \mathbf{x}, \mathbf{y} >\end{aligned}$$

而 $\lambda \neq \mu$，故只能是 $< \mathbf{x}, \mathbf{y} >= 0$，即 $\mathbf{x} \perp \mathbf{y}$。 □

下面我们先简单地讨论一下欧氏空间对应的情形。显然上面引理的 (1)(3)(4)(5)(6) 在欧氏空间上都是成立的，只有 (2) 用到了 $\mathbb{C}$ 的代数闭域性质，因此需要重新讨论。实际上，(2) 在欧氏空间上也成立，理由如下。任取欧氏空间 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$，则对称算子 $\mathcal{A}$ 在标准正交基下的矩阵显然是实对称矩阵。而实对称矩阵 $A \in M_n(\mathbb{R})$ 可以视作标准 Hermitian 空间 $\mathbb{C}^n$ 上的自伴随的线性算子，故由上面的引理，A 的所有特征值都是实数，并且每个特征值都有复的特征向量。设 $\lambda \in \mathbb{R}$ 是 A 的特征值，$\mathbf{x} = (x_1, \ldots, x_n)^t \in \mathbb{C}^n$ 是 λ 对应的复的特征向量，记 $\overline{\mathbf{x}} = (\overline{x_1}, \ldots, \overline{x_n})^t$，则

$$A\mathbf{x} = \lambda\mathbf{x},\ A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}} = \lambda\overline{\mathbf{x}}.$$

即 $\overline{\mathbf{x}}$ 也是 λ 对应的复特征向量。于是 $\mathbf{x} + \overline{\mathbf{x}} \in \mathbb{R}^n$ 就是 λ 对应的 A 的实特征向量，对应地，$(\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot (\mathbf{x} + \overline{\mathbf{x}})$ 就是算子 $\mathcal{A}$ 的实特征值 λ 对应的特征向量。

定理 3.3.4 的证明: 借助 3.3.1的 (2) 和 (3)，我们可以用数学归纳法证明该定理，留作练习。下面我们给出一个更精彩的证明。任取自伴随算子 $\mathcal{A}$ 的一个特征值 $\lambda \in \mathbb{R}$，则其极小多项式应该是下面的形式

$$\mu_{\mathcal{A}}(t) = (t - \lambda)^k \cdot f(t),\ k \in \mathbb{Z}^+.$$

下面我们证明 k 只能等于 1。由极小多项式的定义，任取 $\mathbf{v} \in V$，我们都有 $\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mathbf{0}$，即 $(\mathcal{A} - \lambda\mathcal{E})^k f(\mathcal{A})\mathbf{v} = \mathbf{0}$。所以对任意的 $\mathbf{v} \in V$，都有 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A} - \lambda\mathcal{E})^k)$。由引理 3.3.1的 (1) 可知 $\mathcal{A} - \lambda\mathcal{E}$ 也是自伴随算子，故由引理 3.3.1的 (5)，$\ker((\mathcal{A} - \lambda\mathcal{E})^k) = \ker((\mathcal{A} - \lambda\mathcal{E}))$，所以 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A} - \lambda\mathcal{E}))$，即 $(\mathcal{A} - \lambda\mathcal{E})f(\mathcal{A})\mathbf{v} = \mathbf{0},\ \forall \mathbf{v} \in V$。因此，如果 $k > 1$，则有 $(t - \lambda) \cdot f(t)$ 零化 $\mathcal{A}$ 并且 $\deg[(t - \lambda)f(t)] < \deg(\mu_{\mathcal{A}}(t))$，矛盾！因此只能是 $k = 1$。

由于 $\mathcal{A}$ 的所有特征值都是实数，故不妨设 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t) = (t - \lambda_1)^{p_1} \cdots (t - \lambda_s)^{p_s}$，其中 $\lambda_1, \ldots, \lambda_s \in \mathbb{R}$ 两两不同，$p_1 + \cdots + p_s = n$。则由上面的论述以及推论 2.4.2可知 $\mu_{\mathcal{A}}(t) = (t - \lambda_1) \cdots (t - \lambda_s)$，故由定理 2.5.8的 (vi) 可知 $\mathcal{A}$ 是可对角化的。

最后我们只需要找到一组标准正交基作为 $\mathcal{A}$ 的特征向量即可。由于 $\mathcal{A}$ 可对角化，故 $\mathcal{A}$ 的每个特征值的代数重数都等于几何重数，即 $V = V^{\lambda_1} \oplus V^{\lambda_2} \oplus \cdots \oplus V^{\lambda_s}$，其中 $\dim(V^{\lambda_i}) = p_i$, $\forall i \in \{1, \ldots, s\}$。由定理 3.1.2或定理 3.2.2，我们可以取出每个特征子空间 V^{λ_i} 的一组标准正交基 $\mathbf{e}_{i1}, \ldots, \mathbf{e}_{i,p_i}$，而

由引理 3.3.1的 (6)，$V_{\lambda_i} \perp V_{\lambda_j}$，故将所有 V_{λ_i} 的标准正交基合到一起就得到了 V 的一组标准正交基 $\mathbf{e}_{11}, \ldots, \mathbf{e}_{1,p_1}, \cdots\cdots, \mathbf{e}_{s1}, \ldots, \mathbf{e}_{s,p_s}$。这样我们就完成了定理 3.3.4的证明。 □

我们先来说明矩阵形式的定理 3.3.4′ 的由来。满足 $A^* = A$(或$A' = A$) 的矩阵可以视作 $\mathbb{C}^n$(或$\mathbb{R}^n$) 上的自伴随算子 $\mathcal{A}$，故由定理 3.3.4，可取 $\mathbb{C}^n$(或$\mathbb{R}^n$) 的一组标准正交基 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 使得 $\mathcal{A}$ 可对角化。不妨设 $\mathbb{C}^n$(或$\mathbb{R}^n$) 的标准基到 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 的过渡矩阵为 B，则 B 是酉矩阵 (或正交矩阵，定理 3.2.6/3.1.5)，则算子 $\mathcal{A}$ 在基底 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 下的矩阵为 $B^{-1}AB$，是对角阵。

利用定理 3.3.1和 3.3.4我们也可以顺便讨论一下 Hermitian 型和对称双线性型的对角化。

定理 3.3.5 设 V 是 Hermitian 空间 (或欧氏空间)，$f: V \times V \to \mathbb{C}$(或$\mathbb{R}$) 是 V 上的 Hermitian 型 (或对称双线性型)，由定理 3.3.1，存在唯一的 $\mathcal{A} \in \mathrm{End}(V)$ 使得 $f = f_{\mathcal{A}}$。则 $\mathcal{A}$ 是自伴随算子，并且存在一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 f 有以下形式：任取 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$，$\mathbf{y} = \sum_{j=1}^n y_j \mathbf{e}_j \in V$，有

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

其中 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 是 $\mathcal{A}$ 的所有特征值 (计重数)。

证明: 首先，$\mathcal{A}$ 的存在唯一性是由定理 3.3.1保证的，而 f 的 Hermitian 性 (对称性) 则可以导出：$\forall \mathbf{x}, \mathbf{y} \in V$, $<\mathcal{A}\mathbf{x}, \mathbf{y}> = \overline{<\mathcal{A}\mathbf{y}, \mathbf{x}>} = <\mathbf{x}, \mathcal{A}\mathbf{y}>$，即 $\mathcal{A} = \mathcal{A}^*$。由定理 3.3.4，存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}$ 在这组基下的矩阵是特征值组成的对角阵，即对 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$，有 $\mathcal{A}\mathbf{x} = \sum_{i=1}^n \lambda_i x_i \mathbf{e}_i$，其中 $\lambda_1, \ldots, \lambda_n$ 是 $\mathcal{A}$ 的所有特征值 (计重数，可以相等)。于是

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= <\mathcal{A}\mathbf{x}, \mathbf{y}> = <\sum_{i=1}^n \lambda_i x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j> \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i x_i \overline{y_j} <\mathbf{e}_i, \mathbf{e}_j> = \sum_{i=1}^n \lambda_i x_i \overline{y_i}. \end{aligned}$$

这样我们就完成了证明。 □

我们知道，实对称双线性型与实二次型是一一对应的。同理，我们也可以在 Hermitian 空间 V 上定义 Hermitian 二次型：任取 f 是 $V \times V \to \mathbb{C}$ 的 Hermitian 型，称 $q_f: V \to \mathbb{R}$, $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{x})$ 是 V 上的 Hermitian 二次型。我们同样可以证明 Hermitian 型和 Hermitian 二次型也是一一对应的 (留作练习)。因此，定理 3.3.5有下面的推论：

推论 3.3.3 设 V 是 Hermitian 空间 (或欧氏空间)，$q: V \to \mathbb{R}$ 是 V 上的 Hermitian 二次型 (或实双线性型)，则存在唯一的 $\mathcal{A} \in \mathrm{End}(V)$ 使得 $q(\mathbf{x}) = <\mathcal{A}\mathbf{x}, \mathbf{x}>$，其中 $\mathcal{A}$ 是自伴随算子，并且存在一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 q 有以下形式：任取 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \in V$，有

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

其中 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 是 $\mathcal{A}$ 的所有特征值 (计重数)。

在定理 3.3.5的证明中取 $\mathbf{y} = \mathbf{x}$ 即可证明，细节留作练习。

下面我们先来看一个具体计算的例子。

例 3.3.1 设 $A=\begin{pmatrix}0&1&1&-1\\1&0&-1&1\\1&-1&0&1\\-1&1&1&0\end{pmatrix}\in M_4(\mathbb{R})$，求正交矩阵 P 使得 P^tAP 是对角矩阵。

解: 显然矩阵 A 是线性算子 $\mathcal{A}:\mathbb{R}^4\to\mathbb{R}^4$，$\mathbf{x}=(x_1,x_2,x_3,x_4)^t\mapsto A\mathbf{x}$ 在标准基下的矩阵。我们需要寻找 $\mathbb{R}^4$ 的一组标准正交基使得 $\mathcal{A}$ 在这组标准正交基下的矩阵是对角矩阵。为此，按照定理3.3.4的证明过程，我们需要先将 $\mathbb{R}^4$ 分解成 $\mathcal{A}$ 的特征子空间的直和，再求出每个特征子空间的标准正交基，最后将它们合在一起。

首先，$\chi_{\mathcal{A}}(t)=\det(tE-A)=(t-1)^3(t+3)$，则由定理 3.3.4的证明过程可知，$\mathcal{A}$ 的极小多项式 $\mu_{\mathcal{A}}(t)=(t-1)(t+3)$，所以 $\mathcal{A}$ 有两个特征值 1 和-3，并且 $\dim(V^1)=3$，$\dim(V^{-3})=1$。分别解线性方程组 $(E-A)\mathbf{x}=\mathbf{0}$ 和 $(-3E-A)\mathbf{x}=\mathbf{0}$ 即可得到：

$$V^1=\mathrm{span}\,\{\mathbf{v}_1=(1,1,0,0)^t,\ \mathbf{v}_2=(0,0,1,1)^t,\ \mathbf{v}_3=(-1,0,0,1)^t\},\ V^{-3}=\mathrm{span}\,\{\mathbf{v}_4=(1,-1,-1,1)^t\}.$$

我们利用 Gram-Schmidt 正交化方法将上面的 V^1 和 V^{-3} 的基化成标准正交基：

对V^1：

$$\mathbf{u}_1=\mathbf{v}_1=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}\qquad \mathbf{e}_1=\frac{1}{||\mathbf{u}_1||}\mathbf{u}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\\0\end{pmatrix}$$

$$\mathbf{u}_2=\mathbf{v}_2-\frac{<\mathbf{v}_2,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1=\begin{pmatrix}0\\0\\1\\1\end{pmatrix}\qquad \mathbf{e}_2=\frac{1}{||\mathbf{u}_2||}\mathbf{u}_2=\begin{pmatrix}0\\0\\\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}$$

$$\mathbf{u}_3=\mathbf{v}_3-\frac{<\mathbf{v}_3,\mathbf{u}_1>}{<\mathbf{u}_1,\mathbf{u}_1>}\mathbf{u}_1-\frac{<\mathbf{v}_3,\mathbf{u}_2>}{<\mathbf{u}_2,\mathbf{u}_2>}\mathbf{u}_2=\begin{pmatrix}-\frac12\\\frac12\\-\frac12\\\frac12\end{pmatrix}\quad \mathbf{e}_3=\frac{1}{||\mathbf{u}_3||}\mathbf{u}_3=\begin{pmatrix}-\frac12\\\frac12\\-\frac12\\\frac12\end{pmatrix};$$

对V^{-3}：

$$\mathbf{e}_4=\frac{1}{||\mathbf{v}_4||}\mathbf{v}_4=\begin{pmatrix}\frac12\\-\frac12\\-\frac12\\\frac12\end{pmatrix}.$$

于是 $\mathbf{e}_1,\ldots,\mathbf{e}_4$ 就是使 $\mathcal{A}$ 对角化的标准正交基，由于标准基到 $\mathbf{e}_1,\ldots,\mathbf{e}_4$ 的过渡矩阵就是

$$P=(\mathbf{e}_1,\ldots,\mathbf{e}_4)=\begin{pmatrix}\frac{1}{\sqrt{2}}&0&-\frac12&\frac12\\\frac{1}{\sqrt{2}}&0&\frac12&-\frac12\\0&\frac{1}{\sqrt{2}}&-\frac12&-\frac12\\0&\frac{1}{\sqrt{2}}&\frac12&\frac12\end{pmatrix}.$$

故显然 P 即是我们所要求的正交矩阵，并且 $P^tAP = P^{-1}AP = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 3 \end{pmatrix}$。

需要注意的是，使一个自伴随算子 $\mathcal{A}$ 对角化的标准正交基不是唯一的，这是因为每个 V^{λ_i} 的标准正交基都不只有一种取法。例如，在上面的例子中，取 $P' = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$ 时同样有 $P'^tAP' = P'^{-1}AP' = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 3 \end{pmatrix}$。在这里我们还可以顺手解释一下注 1.6.2。如果实二次型 q 在某组基下 $\mathbf{v}_1, \ldots, \mathbf{v}_s, \ldots, \mathbf{v}_r, \ldots, \mathbf{v}_n$ 的矩阵为 $\begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}$，则我们可以将实向量空间 V 分解成直和 $V = V^+ \oplus V^- \oplus V^0$，其中

$$
\begin{aligned}
V^+ &= \mathrm{span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}, \\
V^- &= \mathrm{span}\,\{\mathbf{v}_{s+1}, \ldots, \mathbf{v}_r\}, \\
V^0 &= \mathrm{span}\,\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}.
\end{aligned}
$$

实际上，设 q 在另外的某组基下对应实对称矩阵 A，其全部两两不同的特征值为 $\lambda_1 > \cdots > \lambda_{k-1} > \lambda_k = 0 > \lambda_{k+1} > \cdots > \lambda_l$，则不难验证 $V^+ = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_{k-1}}$，$V^- = V^{\lambda_{k+1}} \oplus \cdots \oplus V^{\lambda_l}$，具体过程留给读者作为练习。设 $\mathcal{B}, \mathcal{C}$ 分别是 V^+、V^- 上的正交算子，$\mathcal{D}$ 是 V^0 上的任意线性算子，它们在 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 下的矩阵分别是 B, C, D。则我们把 $\mathcal{B}, \mathcal{C}, \mathcal{D}$ 分别作用到 V^+, V^-, V^0 的基上，就可以得到 V 的另一组基 $\mathbf{u}_1, \ldots, \mathbf{u}_s, \ldots, \mathbf{u}_r, \ldots, \mathbf{u}_n$。于是 q 在 $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 下的矩阵为 $\begin{pmatrix} B^tE_sB & & \\ & -C^tE_{r-s}C & \\ & & O \end{pmatrix} = \begin{pmatrix} E_s & & \\ & -E_{r-s} & \\ & & O \end{pmatrix}$。这就是注 1.6.2的构造方法。

在本节的最后，我们介绍一类特殊的自伴随算子：(半) 正定算子。

定义 3.3.3 设 $(V, < \cdot, \cdot >)$ 是内积空间，$\mathcal{A} \in \mathrm{End}(V)$ 是自伴随算子，如果 $\forall \mathbf{x} \in V$，二次型 (或 Hermitian 二次型)$q(\mathbf{x}) =< \mathcal{A}\mathbf{x}, \mathbf{x} >\geqslant 0$，则称 $\mathcal{A}$ 是半正定算子；如果半正定算子 $\mathcal{A}$ 还是非退化 (可逆) 算子，则称 $\mathcal{A}$ 为正定算子。

由推论 3.3.3可知，存在一组标准正交基使得二次型 (或 Hermitian 二次型) $q(\mathbf{x}) =< \mathcal{A}\mathbf{x}, \mathbf{x} >$ 有标准型

$$
q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i |x_i|^2,\ \forall \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \in V.
$$

其中 λ_i, $i = 1, \ldots, n$ 是 $\mathcal{A}$ 的所有特征值 (计重数)。因此我们立刻有：

命题 3.3.1 设 $\mathcal{A}$ 是内积空间 $(V,<\cdot,\cdot>)$ 上的自伴随算子，则 $\mathcal{A}$ 是半正定算子 $\Longleftrightarrow$ 其所有特征值都非负；$\mathcal{A}$ 是正定算子 $\Longleftrightarrow$ 其所有特征值都大于 0。

我们知道，非负实数可以开平方，现在我们希望将开平方的操作推广到半正定算子 (半正定矩阵) 上。为此，不妨设半正定算子 $\mathcal{A}$ 在某组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵为

$$A=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},\quad \lambda_i\geqslant 0,\ \forall i\in\{1,\dots,n\}.$$

则线性算子 $\mathcal{B}:\mathbf{x}=\sum_{i=1}^n x_i\mathbf{e}_i\mapsto\sum_{i=1}^n\sqrt{\lambda_i}x_i\mathbf{e}_i$(即 $\mathcal{B}$ 在此基底下对应的矩阵为 $B=\begin{pmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{pmatrix}$) 也是半正定的自伴随算子，并且 $\mathcal{B}^2=\mathcal{B}^*\mathcal{B}=\mathcal{A}$。特别地，如果 $\mathcal{A}$ 正定，则 $\mathcal{B}$ 也正定。我们把算子 $\mathcal{B}$ 称为算子 $\mathcal{A}$ 的**平方根**，记作 $\mathcal{B}=\sqrt{\mathcal{A}}$。实际上半正定算子的平方根是唯一的，我们会在稍后的 3.5 节中证明它，或者先留作练习[1]。

于是我们有

定理 3.3.6 设 $(V,<\cdot,\cdot>)$ 是内积空间，$\mathcal{A}\in\mathrm{End}(V)$，则以下条件等价：

(1) $\mathcal{A}$ 是半正定算子 (正定算子)；

(2) 存在 $\mathcal{B}\in\mathrm{End}(V)$ 是半正定算子 (正定算子) 使得 $\mathcal{A}=\mathcal{B}^2$ 且 $\mathcal{B}^*=\mathcal{B}$；

(3) 存在线性算子 (可逆算子)C 使得 $\mathcal{A}=C^*C$ 或 $\mathcal{A}=CC^*$。

证明: (1)$\Longrightarrow$(2) 即上面的论述。(2)$\Longrightarrow$(3) 是平凡的。下面我们证明 (3)$\Longrightarrow$(1)。我们只证明 $\mathcal{A}=C^*C$ 的情形，另一个同理。首先 $\mathcal{A}^*=(C^*C)^*=C^*C=\mathcal{A}$，即 $\mathcal{A}$ 是自伴随算子；其次任取 $\mathbf{x}\in V$，有 $<C^*C\mathbf{x},\mathbf{x}>=<C\mathbf{x},C\mathbf{x}>\ \geqslant 0$，即 C 是半正定算子。特别地，如果 C 是可逆算子，则 $\mathcal{A}$ 显然也是可逆的，故 $\mathcal{A}$ 是正定算子。 □

[1]也可参考《高等代数学习指南》，蓝以中，北京大学出版社的第五章 §2 的例 2.9-例 2.11。

3.4 内积空间上的线性算子 II：保距算子

现在我们开始考虑内积空间 V 上另一类特殊的算子：保距算子。之所以我们要把满足 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>$, $\forall\mathbf{x}\in V$ 的算子 $\mathcal{A}$ 称为保距算子，是因为 $\mathcal{A}$ 满足：$\forall\mathbf{x},\mathbf{y}\in V$, $\mathrm{d}(\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y})=\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|=\|\mathbf{x}-\mathbf{y}\|=\mathrm{d}(\mathbf{x},\mathbf{y})$，即算子 $\mathcal{A}$ 保持距离。下面我们先证明保距算子在有限维的内积空间上就是酉算子 (或正交算子)。

定理 3.4.1 设 $(V,<\cdot,\cdot>)$ 是内积空间，则 $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x}\in V$, $\|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|\iff\forall\mathbf{x},\mathbf{y}\in V$, $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$。

证明: 由内积空间中范数的定义，$\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x}\in V$, $\|\mathcal{A}\mathbf{x}\|=\|\mathbf{x}\|$ 是显然的。下面我们只需证明 $\mathcal{A}$ 是保距算子 $\iff\forall\mathbf{x},\mathbf{y}\in V$, $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathbf{y}>$。
($\Longleftarrow$) 取 $\mathbf{y}=\mathbf{x}$ 即得 $\mathcal{A}$ 是保距算子。
($\Longrightarrow$) 如果 V 是欧氏空间，$\mathcal{A}$ 是保距算子，则 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>$, $\forall\mathbf{x}\in V$，于是由极化恒等式 (命题 3.1.3的 (3))，我们有：任取 $\mathbf{x},\mathbf{y}\in V$，有

$$\begin{aligned}<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}> &= \frac{1}{4}(\|\mathcal{A}(\mathbf{x}+\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2)\\ &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)\\ &=<\mathbf{x},\mathbf{y}>.\end{aligned}$$

如果 V 是 Hermitian 空间，$\mathcal{A}$ 是保距算子，则 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathbf{x},\mathbf{x}>$, $\forall\mathbf{x}\in V$，于是由极化恒等式 (命题 3.2.4的 (3))，我们有：任取 $\mathbf{x},\mathbf{y}\in V$，有

$$\begin{aligned}<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}> &= \frac{1}{4}(\|\mathcal{A}(\mathbf{x}+\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-\mathbf{y})\|^2)+\frac{i}{4}(\|\mathcal{A}(\mathbf{x}+i\mathbf{y})\|^2-\|\mathcal{A}(\mathbf{x}-i\mathbf{y})\|^2)\\ &= \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2-\|\mathbf{x}-\mathbf{y}\|^2)+\frac{i}{4}(\|\mathbf{x}+i\mathbf{y}\|^2-\|\mathbf{x}-i\mathbf{y}\|^2)\\ &=<\mathbf{x},\mathbf{y}>.\end{aligned}$$

即 $\mathcal{A}$ 是正交算子或酉算子。这样我们就完成了证明。 □

推论 3.4.1 设 $(V,<\cdot,\cdot>)$ 是内积空间，则 $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\mathcal{A}^*\mathcal{A}=\mathcal{E}$。

证明: $\mathcal{A}\in\mathrm{End}(V)$ 是保距算子 $\iff\forall\mathbf{x},\mathbf{y}\in V$, $<\mathbf{x},\mathbf{y}>=<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathbf{x},\mathcal{A}^*\mathcal{A}\mathbf{y}>\iff\forall\mathbf{x},\mathbf{y}\in V$, $<\mathbf{x},(\mathcal{A}^*\mathcal{A}-\mathcal{E})\mathbf{y}>=0\iff\forall\mathbf{y}\in V$, $(\mathcal{A}^*\mathcal{A}-\mathcal{E})\mathbf{y}=\mathbf{0}\iff\mathcal{A}^*\mathcal{A}=\mathcal{E}$。 □

于是任取 V 的一组标准正交基，保距算子对应的矩阵就是酉矩阵 (或正交矩阵)。

注 3.4.1 在有限维空间中，$\mathcal{A}^*\mathcal{A}=\mathcal{E}$ 显然能够保证 $\mathcal{A}$ 可逆。但在无穷维空间中，保距算子不一定是酉算子，这是因为 $\mathcal{A}^*\mathcal{A}=\mathcal{E}$ 并不能保证 $\mathcal{A}$ 是可逆的 (这只能说明 $\mathcal{A}$ 有左逆，但右逆可以不存在，当然如果右逆存在的话必有左逆等于右逆)，而酉算子必须是可逆算子。反例可参考《A Course in Functional Analysis》，Jorn B.Conway, GTM96 的 Chapter I, §4 的 Example 5.3。

下面我们来考虑保距算子的对角化问题。保距算子的对角化可以作为下一节的结论的特例，但为了循序渐进，我们在这里给出一个比较简单的归纳证明 (用它读者也可以类比地写出定理 3.3.4和下一节正规算子可对角化的证明)。

定理 3.4.2 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}$ 是 V 上的保距算子 (即酉算子)，则存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得：

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

其中 $\lambda_i \in \mathbb{C}$ 且 $|\lambda_i| = 1$, $\forall i \in \{1, \ldots, n\}$(各个 λ_i 之间可以相等)。

为了证明上述定理，我们首先需要以下引理：

引理 3.4.1 设 $\mathcal{A}$ 是 Hermitian 空间 $(V, <\cdot,\cdot>)$ 上的保距算子，则

(1) 若 $\lambda \in \mathbb{C}$ 是 $\mathcal{A}$ 的特征值，则 $|\lambda| = 1$；

(2) 若 W 是 $\mathcal{A}-$ 子空间，则 $W^\perp$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 都是保距算子。

证明: (1) 设 $\mathbf{v} \in V$ 是 λ 对应的一个非零特征向量，则

$$< \mathbf{v}, \mathbf{v} > = < \mathcal{A}\mathbf{v}, \mathcal{A}\mathbf{v} > = < \lambda\mathbf{v}, \lambda\mathbf{v} > = \lambda \cdot \overline{\lambda} < \mathbf{v}, \mathbf{v} > .$$

因为 $< \mathbf{v}, \mathbf{v} > \neq 0$，故只能是 $\lambda\overline{\lambda} = |\lambda|^2 = 1$，即 $|\lambda| = 1$。

(2) 由于 W 是 $\mathcal{A}-$ 子空间，而 $\mathcal{A}$ 是可逆算子，故 $\forall \mathbf{u} \in W$, $\exists \mathbf{w} \in W$ 使得 $\mathcal{A}\mathbf{w} = \mathbf{u}$。下面我们验证 $W^\perp$ 也是 $\mathcal{A}-$ 子空间。任取 $\mathbf{v} \in W^\perp$(即 $\mathbf{v} \perp W$)，我们只要证明 $\mathcal{A}\mathbf{v} \in W^\perp$，即 $\mathcal{A}\mathbf{v} \perp W$ 即可完成证明。任取 $\mathbf{u} \in W$，则 $< \mathcal{A}\mathbf{v}, \mathbf{u} > = < \mathcal{A}\mathbf{v}, \mathcal{A}\mathbf{w} > = < \mathbf{v}, \mathbf{w} > = 0$，此即 $\mathcal{A}\mathbf{v} \perp W$。$\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^\perp}$ 是保距算子是显然的。 □

定理 3.4.2 **的证明**: 对 V 的维数 n 用数学归纳法。$n = 1$ 时定理显然成立。下设定理对所有 $n-1$ 维 Hermitian 空间都成立，我们来证明 $\dim(V) = n$ 的情形。

由于特征多项式 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{C}$ 上至少有 1 个根，故我们总可以取 $\mathcal{A}$ 的某个特征值 λ_1 和它对应的一个特征向量 $\mathbf{e}_1$。由引理 3.4.1，我们有 $|\lambda_1| = 1$，并且 span $\{\mathbf{e}_1\}$ 和 span $\{\mathbf{e}_1\}^\perp$ 都是 $\mathcal{A}-$ 子空间。由于 $\mathcal{A}$ 限制到 span $\{\mathbf{e}_1\}^\perp$ 上也是保距算子，并且 $\dim(\text{span}\,\{\mathbf{e}_1\}^\perp) = n-1$，故由归纳假设，存在 span $\{\mathbf{e}_1\}^\perp$ 的一组标准正交基 $\mathbf{e}_2, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}$ 限制到 span $\{\mathbf{e}_1\}^\perp$ 上对应的矩阵是对角阵 $\begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}_{(n-1)\times(n-1)}$，并且对角元的模长都是 1。因此，结合上述论证与推论 2.3.1，可知 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下的矩阵是 $\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}_{n\times n}$，并且所有对角元的模长都是 1。这样我们就完成了归纳证明。 □

推论 3.4.2 设 $A \in M_n(\mathbb{C})$ 是酉矩阵，则存在 $B \in U(n)$ 使得

$$B^*AB = B^{-1}AB = \begin{pmatrix} e^{i\varphi_1} & & \\ & \ddots & \\ & & e^{i\varphi_n} \end{pmatrix}$$

其中 $\varphi_1, \ldots, \varphi_n \in \mathbb{R}$(可以相等)，$i$ 是虚数单位。

将矩阵视作 $\mathbb{C}^n$ 上的线性算子即可证明，注意得到最后的对角矩阵形式时利用了复数的指数形式 (讲义上册 5.1 节)。

现在我们来考虑欧氏空间上的正交算子的情形。此时情况要复杂一些，因为正交算子的特征值可能是虚数，这样我们就不能将正交算子在实数域上对角化。但退而求其次，我们可以将正交算子在实数域上分块对角化。

引理 3.4.2 设 $\mathcal{A}$ 是 Hermitian 空间 $(V, <\cdot,\cdot>)$ 上的保距算子，则

(1) 若 $\lambda \in \mathbb{R}$ 是 $\mathcal{A}$ 的实特征值，则 $\lambda = \pm 1$；

(2) 若 W 是 $\mathcal{A}-$ 子空间，则 $W^{\perp}$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_W$ 和 $\mathcal{A}\big|_{W^{\perp}}$ 都是保距算子。

证明与引理 3.4.1相同，留作练习。

引理 3.4.3 $\mathbb{R}$ 上的一阶正交矩阵只有 (± 1)；$\mathbb{R}$ 上的二阶正交矩阵一定是 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$，$\theta \in \mathbb{R}$ 的形式。

证明: 一阶的情况是显然的。下面考虑二阶的情形。任取二阶实正交矩阵 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则 $A^t A = E$，即 $\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$，于是

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

不妨取 $a = \cos\theta,\ c = \sin\theta,\ b = \sin\varphi,\ d = \cos\varphi$，则 $ab + cd = 0 \Rightarrow \sin(\theta + \varphi) = 0,\ \varphi = 2k\pi - \theta, k \in \mathbb{Z}$，所以 $b = -c, d = a$，即 $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$。 □

引理 3.4.4 设 V 是 n 维实向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 V 一定有一个 1 维或 2 维的 $\mathcal{A}-$ 子空间。

证明: 设 $\mathcal{A}$ 的特征多项式是 $\chi_{\mathcal{A}}(t)$，则 $\chi_{\mathcal{A}}(t)$ 在 $\mathbb{R}[t]$ 中可以分解成一些 1 次因子和 2 次的因子的乘积 (讲义上册推论 6.5.2)，则由定理 2.4.4的证明方法，归纳即可得到 $\mathcal{A}$ 一定有一个 1 维或 2 维的不变子空间。细节留作思考。 □

定理 3.4.3 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A}$ 是 V 上的保距算子 (即正交算子)，则存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得：

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} A_1 & & & & \\ & \ddots & & & \\ & & A_r & & \\ & & & E_s & \\ & & & & -E_t \end{pmatrix}.$$

其中 $A_i = \begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix} \in O(2),\ \forall i \in \{1, \ldots, r\}$ (各个 φ_i 之间可以相等)，E_s 和 E_t 分别是 $s \times s$ 和 $t \times t$ 阶的单位矩阵，并且 $2r + s + t = n$。

证明: 首先，V 一定有一个 1 维或 2 维的 $\mathcal{A}-$ 子空间 U，则 $U^{\perp}$ 也是 $\mathcal{A}-$ 子空间，并且 $\mathcal{A}\big|_{U^{\perp}}$ 上也是正交算子。因此，类似于定理 3.4.2的证明方法，我们可以归纳地证明：V 是一些两两正交的 1 维或 2 维的 $\mathcal{A}-$ 子空间的直和 (细节留作练习)。不失一般性，我们可以设：

$$V = V_1 \oplus \cdots \oplus V_r \oplus V_{r+1} \oplus \cdots \oplus V_{r+s} \oplus V_{r+s+1} \oplus \cdots \oplus V_{r+s+t}. \tag{3.4.1}$$

其中 $V_1, \ldots, V_r$ 是 2 维的 $\mathcal{A}-$ 子空间，$V_{r+1}, \ldots, V_{r+s}$ 是 $\mathcal{A}$ 的 1 维特征子空间 (对应的特征值是 1)，$V_{r+s+1}, \ldots, V_{r+s+t}$ 是 $\mathcal{A}$ 的 1 维特征子空间 (对应的特征值是-1)。因此，由引理 3.4.2，$\mathcal{A}$ 限制到每个 V_i 上都是正交算子，即每个 V_i 中都存在一组标准正交基使得 $\mathcal{A}\big|_{V_i}$ 对应的矩阵是 1 阶或 2 阶的正交矩阵。故由引理 3.4.3及推论 2.3.1，将每个 V_i 的标准正交基合在一起构成 V 的标准正交基，则 $\mathcal{A}$ 在此标准正交基下的矩阵就是我们所需要的形式。$2r + s + t = n$ 可以由直和分解 (3.4.1) 直接得到。 □

在本节的最后，我们介绍一下内积空间上可逆线性算子的**极化分解** (Polar decomposition)。我们知道任何非零的复数 z 都能写成指数形式 $z = re^{i\theta}$，其中 $r \in \mathbb{R}^+$，$\theta \in \mathbb{R}$。实际上，我们可以对可逆线性算子作类似的分解。

定理 3.4.4 设 $(V, <\cdot, \cdot>)$ 是 n 维内积空间，$\mathcal{A} \in \mathrm{End}(V)$ 可逆，则

(1) 存在正定算子 $\mathcal{P}$ 和保距算子 $\mathcal{Q}$ 使得 $\mathcal{A} = \mathcal{P}\mathcal{Q}$。更进一步，$\mathcal{P}$ 和 $\mathcal{Q}$ 都是唯一的。

(2) 存在正定算子 $\mathcal{P}'$ 和保距算子 $\mathcal{Q}'$ 使得 $\mathcal{A} = \mathcal{Q}'\mathcal{P}'$。更进一步，$\mathcal{P}'$ 和 $\mathcal{Q}'$ 也都是唯一的。

证明: (1) 由定理 3.3.6，$\mathcal{A}\mathcal{A}^*$ 是正定算子，因此 $\mathcal{A}\mathcal{A}^*$ 存在 (唯一) 的平方根 $\mathcal{P} = \sqrt{\mathcal{A}\mathcal{A}^*}$。则 $\mathcal{P}^2 = \mathcal{P}\mathcal{P}^* = \mathcal{A}\mathcal{A}^*$，并且 $\mathcal{P}$ 也正定，故 $\mathcal{P}$ 可逆，于是可令 $\mathcal{Q} = \mathcal{P}^{-1}\mathcal{A}$。下面我们证明 $\mathcal{Q}$ 是保距算子，从而 $\mathcal{A} = \mathcal{P}\mathcal{Q}$ 即我们所需要的分解。注意到 $\mathcal{Q}^* = \mathcal{A}^*(\mathcal{P}^{-1})^* = \mathcal{A}^*(\mathcal{P}^*)^{-1}$ (注意取 $*$ 和求逆可交换是定理 3.3.3的直接推论，留作练习)，故

$$\mathcal{Q}^*\mathcal{Q} = \mathcal{A}^*(\mathcal{P}^*)^{-1}\mathcal{P}^{-1}\mathcal{A} = \mathcal{A}^*(\mathcal{P}\mathcal{P}^*)^{-1}\mathcal{A} = \mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A} = \mathcal{A}^*(\mathcal{A}^*)^{-1}\mathcal{A}^{-1}\mathcal{A} = \mathcal{E}$$

此即 $\mathcal{Q}$ 是保距算子，故分解存在。下证唯一性。设另有正定算子 $\mathcal{P}_1$ 和保距算子 $\mathcal{Q}_1$ 也满足 $\mathcal{A} = \mathcal{P}_1\mathcal{Q}_1$，则 $\mathcal{A} = \mathcal{P}\mathcal{Q} = \mathcal{P}_1\mathcal{Q}_1$，对上式取 $*$ 即得 $\mathcal{Q}^*\mathcal{P} = \mathcal{Q}_1^*\mathcal{P}_1$，故

$$\mathcal{A}\mathcal{A}^* = \mathcal{P}^2 = \mathcal{P}\mathcal{Q}\mathcal{Q}^*\mathcal{P} = \mathcal{P}\mathcal{Q}\mathcal{Q}_1^*\mathcal{P}_1 = \mathcal{P}_1\mathcal{Q}_1\mathcal{Q}_1^*\mathcal{P}_1 = \mathcal{P}_1^2$$

因此，由正定算子平方根的唯一性 (会在下一节证明) 可知 $\mathcal{P} = \mathcal{P}_1$，于是 $\mathcal{Q} = \mathcal{Q}_1$ 也是唯一的。这样我们就完成了证明。

(2) 取 $\mathcal{P}' = \sqrt{\mathcal{A}^*\mathcal{A}}$, $\mathcal{Q} = \mathcal{A}(\mathcal{P}')^{-1}$ 即可，证明过程同上，留作练习。 □

3.5 内积空间上的线性算子 III：正规算子

现在我们可以来探讨内积空间上的线性算子何时可以用标准正交基对角化了。

定义 3.5.1 设 $(V,<\cdot,\cdot>)$ 是内积空间，$\mathcal{A}\in \mathrm{End}(V)$，如果 $\mathcal{A}$ 还满足 $\mathcal{A}^*\mathcal{A}=\mathcal{A}\mathcal{A}^*$，则我们称 $\mathcal{A}$ 是正规算子 (normal operator)。对应到矩阵层面，如果矩阵 A 满足 $A^*A=AA^*$(实数域情形则是 $A^tA=AA^t$)，则我们称 A 是正规矩阵 (normal matrix)。显然 $\mathcal{A}$ 是正规算子 $\iff$ $\mathcal{A}$ 在标准正交基下的矩阵是正规矩阵。

容易看出欧氏空间或 Hermitian 空间上的自伴随算子和保距算子都是正规算子。但正规算子显然不只有这两种类型，例如 $\mathbb{C}^2$ 上的算子 $\mathbf{x}(x_1,x_2)^t\mapsto\begin{pmatrix}2&0\\0&i\end{pmatrix}$ 就是正规算子，但显然它不是自伴随算子或保距算子。

本节中我们的主要任务是研究正规算子的对角化问题，即谱定理。我们主要在复数域上讨论问题，实数域的情形我们会给出对应的结论和参考文献。

定理 3.5.1 设 $(V,<\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}\in\mathrm{End}(V)$，则 $\mathcal{A}$ 是正规算子 $\iff$ 存在 V 的一组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 使得

$$(\mathcal{A}\mathbf{e}_1,\dots,\mathcal{A}\mathbf{e}_n)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix},\quad \lambda_1,\dots,\lambda_n\in\mathbb{C}(\lambda_1,\dots,\lambda_n\text{中可以有相等的元素}).$$

我们同样可以用数学归纳法证明以上定理。但为了更清楚地理解正规算子的结构，我们将仿照定理 3.3.4的证明方法给出上面定理的证明。

引理 3.5.1 设 $(V,<\cdot,\cdot>)$ 是内积空间，$\mathcal{A}\in\mathrm{End}(V)$，则 $\ker(\mathcal{A}^*\mathcal{A})=\ker(\mathcal{A})$。

证明: $\ker(\mathcal{A}^*\mathcal{A})\supset\ker(\mathcal{A})$：任取 $\mathbf{x}\in\ker(\mathcal{A})$，则 $\mathcal{A}\mathbf{x}=\mathbf{0}$，于是 $\mathcal{A}^*\mathcal{A}\mathbf{x}=\mathbf{0}$，即 $\mathbf{x}\in\ker(\mathcal{A}^*\mathcal{A})$。
$\ker(\mathcal{A}^*\mathcal{A})\subset\ker(\mathcal{A})$：任取 $\mathbf{x}\in\ker(\mathcal{A}^*\mathcal{A})$，即 $\mathcal{A}^*\mathcal{A}\mathbf{x}=\mathbf{0}$，于是 $<\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{x}>=<\mathcal{A}^*\mathcal{A}\mathbf{x},\mathbf{x}>=\mathbf{0}$，因此 $\mathcal{A}\mathbf{x}=\mathbf{0}$，即 $\mathbf{x}\in\ker(\mathcal{A})$。这样我们就完成了证明。 □

引理 3.5.2 设 $(V,<\cdot,\cdot>)$ 是 Hermitian 空间，$\mathcal{A}\in\mathrm{End}(V)$ 是正规算子，则

(1) $\mathcal{A}^*$ 也是正规算子。$\forall p(t)\in\mathbb{C}[t]$，线性算子 $p(\mathcal{A})$ 也是正规算子。特别地，对任意的 $\lambda\in\mathbb{C}$，有 $\mathcal{A}-\lambda\mathcal{E}$ 是正规算子。

(2) $\forall\mathbf{x},\mathbf{y}\in V,\ <\mathcal{A}\mathbf{x},\mathcal{A}\mathbf{y}>=<\mathcal{A}^*\mathbf{x},\mathcal{A}^*\mathbf{y}>$。因此，$\forall\mathbf{x}\in V$，$\|\mathcal{A}\mathbf{x}\|=\|\mathcal{A}^*\mathbf{x}\|$，并且 $\ker(\mathcal{A})=\ker(\mathcal{A}^*)$。

(3) $\mathcal{A}\mathbf{x}=\lambda\mathbf{x}\iff\mathcal{A}^*\mathbf{x}=\overline{\lambda}\mathbf{x}$，其中 $\lambda\in\mathbb{C}$ 是 $\mathcal{A}$ 的特征值，$\mathbf{x}\in V$ 是 λ 对应的非零特征向量。

(4) $\forall k\in\mathbb{Z}^+$，$\ker(\mathcal{A}^k)=\ker(\mathcal{A})$。

(5) 如果 λ,μ 是 $\mathcal{A}$ 的两个不同的特征值，$\mathbf{x}$ 和 $\mathbf{y}$ 分别是 λ 和 μ 对应的特征向量，则 $\mathbf{x}\perp\mathbf{y}$。

证明: (1) $\mathcal{A}^*$ 的正规性是显然的。由于 $\mathcal{A}^*\mathcal{A}=\mathcal{A}\mathcal{A}^*$，则容易验证 $\forall i,j\in\mathbb{Z}^+$，$(\mathcal{A}^*)^i\mathcal{A}^j=\mathcal{A}^j(\mathcal{A}^*)^i$。

不妨设 $p(t)=a_0+a_1t+\cdots+a_nt^n\in\mathbb{C}[t]$，则有

$$p(\mathcal{A})=a_0\mathcal{E}+a_1\mathcal{A}^*+\cdots+a_n(\mathcal{A}^*)^n,\ [p(\mathcal{A})]^*=\overline{a_0}\mathcal{E}+\overline{a_1}\mathcal{A}^*+\cdots+\overline{a_n}(\mathcal{A}^*)^n,$$

于是

$$\begin{aligned}[p(\mathcal{A})]^* \cdot p(\mathcal{A}) &= \sum_{i=0}^{n}\sum_{j=0}^{n} \overline{a_i} a_j (\mathcal{A}^*)^i \mathcal{A}^j \\ &= \sum_{i=0}^{n}\sum_{j=0}^{n} a_j \overline{a_i} \mathcal{A}^j (\mathcal{A}^*)^i \\ &= p(\mathcal{A}) \cdot [p(\mathcal{A})]^*\end{aligned}$$

即 $p(\mathcal{A})$ 也是正规算子。取 $p(t) = t - \lambda$ 即可看出特例是显然的。

(2) 利用正规性可得

$$\begin{aligned}< \mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{y} > &=< \mathbf{x}, \mathcal{A}^*\mathcal{A}\mathbf{y} > \\ &=< \mathbf{x}, \mathcal{A}\mathcal{A}^*\mathbf{y} > \\ &=< \mathcal{A}^*\mathbf{x}, \mathcal{A}^*\mathbf{y} >\end{aligned}$$

在上式中取 $\mathbf{y} = \mathbf{x}$ 即得 $\|\mathcal{A}\mathbf{x}\| = \|\mathcal{A}^*\mathbf{x}\|$。所以

$$\mathcal{A}\mathbf{x} = \mathbf{0} \iff \|\mathcal{A}\mathbf{x}\| = 0 \iff \|\mathcal{A}^*\mathbf{x}\| = 0 \iff \mathcal{A}^*\mathbf{x} = \mathbf{0},$$

即 $\ker(\mathcal{A}) = \ker(\mathcal{A}^*)$。

(3) 由 (1) 可知 $\mathcal{A} - \lambda\mathcal{E}$ 是正规算子，并且 $(\mathcal{A} - \lambda\mathcal{E})^* = \mathcal{A}^* - \overline{\lambda}\mathcal{E}$，因此 $\ker(\mathcal{A} - \lambda\mathcal{E}) = \ker(\mathcal{A}^* - \overline{\lambda}\mathcal{E})$。于是任取 $\mathbf{x} \in \ker(\mathcal{A} - \lambda\mathcal{E})$ 即得结论。

(4) 由于 $(\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}$，即 $\mathcal{A}^*\mathcal{A}$ 是自伴随算子，故由引理 3.3.1的 (5)，我们有

$$\forall k \in \mathbb{Z}^+,\ \ker[(\mathcal{A}^*\mathcal{A})^k] = \ker(\mathcal{A}^*\mathcal{A}).$$

而由正规性容易验证

$$(\mathcal{A}^*\mathcal{A})^k = (\mathcal{A}^*)^k\mathcal{A}^k = (\mathcal{A}^k)^*\mathcal{A}^k,$$

因此由引理 3.5.1可得

$$\ker(\mathcal{A}^k) = \ker[(\mathcal{A}^k)^*\mathcal{A}^k] = \ker[(\mathcal{A}^*\mathcal{A})^k] = \ker(\mathcal{A}^*\mathcal{A}) = \ker(\mathcal{A}).$$

即得结论。

(5) 由于 $\mathcal{A}\mathbf{x} = \lambda\mathbf{x}$，$\mathcal{A}\mathbf{y} = \mu\mathbf{y}$，故首先由 (3) 可知 $\mathcal{A}^*\mathbf{y} = \overline{\mu}\mathbf{y}$。于是利用特征向量的定义和内积的半双线性性质，我们有

$$\begin{aligned}\lambda < \mathbf{x}, \mathbf{y} > &=< \lambda\mathbf{x}, \mathbf{y} >=< \mathcal{A}\mathbf{x}, \mathbf{y} > \\ &=< \mathbf{x}, \mathcal{A}^*\mathbf{y} >=< \mathbf{x}, \overline{\mu}\mathbf{y} > \\ &= \mu < \mathbf{x}, \mathbf{y} >\end{aligned}$$

而 $\lambda \neq \mu$，故只能是 $< \mathbf{x}, \mathbf{y} >= 0$，即 $\mathbf{x} \perp \mathbf{y}$。 □

定理 3.5.1的证明: ($\Longleftarrow$) 如果 $\mathcal{A}$ 在某一组标准正交基下的矩阵是 $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$，则 $\mathcal{A}^*$ 在此基底下的矩阵是 $\begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix}$。由于

$$A^*A = AA^* = \begin{pmatrix} |\lambda_1|^2 & & \\ & \ddots & \\ & & |\lambda_n|^2 \end{pmatrix}$$

故 $\mathcal{A}$ 是正规算子。

($\Longrightarrow$) 任取正规算子 $\mathcal{A}$ 的一个特征值 $\lambda \in \mathbb{C}$，则其极小多项式应该是下面的形式

$$\mu_{\mathcal{A}}(t) = (t-\lambda)^k \cdot f(t),\ k \in \mathbb{Z}^+.$$

下面我们证明 k 只能等于 1。由极小多项式的定义，任取 $\mathbf{v} \in V$，我们都有 $\mu_{\mathcal{A}}(\mathcal{A})\mathbf{v} = \mathbf{0}$，即 $(\mathcal{A}-\lambda\mathcal{E})^k f(\mathcal{A})\mathbf{v} = \mathbf{0}$。所以对任意的 $\mathbf{v} \in V$，都有 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A}-\lambda\mathcal{E})^k)$。由引理 3.5.2的 (1) 可知 $\mathcal{A}-\lambda\mathcal{E}$ 也是正规算子，故由引理 3.5.2的 (4)，$\ker((\mathcal{A}-\lambda\mathcal{E})^k) = \ker((\mathcal{A}-\lambda\mathcal{E}))$，所以 $f(\mathcal{A})\mathbf{v} \in \ker((\mathcal{A}-\lambda\mathcal{E}))$，即 $(\mathcal{A}-\lambda\mathcal{E})f(\mathcal{A})\mathbf{v} = \mathbf{0},\ \forall \mathbf{v} \in V$。因此，如果 $k > 1$，则有 $(t-\lambda)\cdot f(t)$ 零化 $\mathcal{A}$ 并且 $\deg[(t-\lambda)f(t)] < \deg(\mu_{\mathcal{A}}(t))$，矛盾！因此只能是 $k = 1$。

设 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t) = (t-\lambda_1)^{p_1}\cdots(t-\lambda_s)^{p_s}$，其中 $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$ 两两不同，$p_1+\cdots+p_s = n$。则由上面的论述以及推论 2.4.2可知 $\mu_{\mathcal{A}}(t) = (t-\lambda_1)\cdots(t-\lambda_s)$，故由定理 2.5.8的 (vi) 可知 $\mathcal{A}$ 是可对角化的。

最后我们只需要找到一组标准正交基作为 $\mathcal{A}$ 的特征向量即可。由于 $\mathcal{A}$ 可对角化，故 $\mathcal{A}$ 的每个特征值的代数重数都等于几何重数，即 $V = V^{\lambda_1} \oplus V^{\lambda_2} \oplus \cdots \oplus V^{\lambda_s}$，其中 $\dim(V^{\lambda_i}) = p_i,\ \forall i \in \{1, \ldots, s\}$。由定理 3.1.2或定理 3.2.2，我们可以取出每个特征子空间 V^{λ_i} 的一组标准正交基 $\mathbf{e}_{i1}, \ldots, \mathbf{e}_{i,p_i}$，而由引理 3.5.2的 (5)，$V_{\lambda_i} \perp V_{\lambda_j}$，故将所有 V_{λ_i} 的标准正交基合到一起就得到了 V 的一组标准正交基 $\mathbf{e}_{11}, \ldots, \mathbf{e}_{1,p_1}, \cdots\cdots, \mathbf{e}_{s1}, \ldots, \mathbf{e}_{s,p_s}$。这样我们就完成了定理 3.5.1的证明。 □

与归纳法证明相比，以上证明看起来更加复杂。但实际上，上面的证明方法同时也给出了正规算子的极小多项式和特征子空间的精细结构，这对我们理解下面的谱定理是有帮助的。在我们引入谱定理之前，我们先直接给出欧氏空间上正规算子的对角化结论。与正交算子的情形类似，在欧氏空间上我们只能使正规算子分块对角化。

命题 3.5.1 设 $(V, <\cdot,\cdot>)$ 是 n 维欧氏空间，$\mathcal{A} \in \mathrm{End}(V)$，则 $\mathcal{A}$ 是正规算子 $\Longleftrightarrow$ 存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得

$$(\mathcal{A}\mathbf{e}_1, \ldots, \mathcal{A}\mathbf{e}_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & A_1 & & \\ & & & & \ddots & \\ & & & & & A_m \end{pmatrix}.$$

其中 $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$，$A_i = \begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix} \in M_2(\mathbb{R})$，$i \in \{1, \ldots, m\}$，并且 $k + 2m = n$。

证明可参考 Advanced Linear Algebra,Steven Roman,GTM135 的 Chapter10,Theorem10.10。

接下来我们引入谱定理。我们的最终目的是将一个正规算子唯一地“拆解”成比较简单的形式，即一些正交投影算子的线性组合。实际上，在定理 3.5.1的证明过程中我们就已经得到：$\mathcal{A} = \sum_{i=1}^{s} \lambda_i \mathcal{P}_i$，其中 $\lambda_i, i = 1, \ldots, s$ 是 $\mathcal{A}$ 的两两不同的特征值，$\mathcal{P}_i$ 在标准正交基 $\mathbf{e}_{11}, \ldots, \mathbf{e}_{s,p_s}$ 下的矩阵是 $\begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & E_{p_i} & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$ (即只有 $\mathbf{e}_{i1}, \ldots, \mathbf{e}_{i,p_i}$ 对应的对角线元素是 1，其余位置都是 0)。显然每个 $\mathcal{P}_i$ 都是投影算子 (例 2.3.2)。这就是我们得到谱定理的思路。

定义 3.5.2 设 V 是**一般**的域 $\mathbb{K}$ 上的 n 维向量空间。如果两个投影算子 $\mathcal{P}_1, \mathcal{P}_2 \in \mathrm{End}(V)$ 满足 $\mathcal{P}_1\mathcal{P}_2 = \mathcal{O}$，则我们称 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 正交。更进一步，如果 $V = V_1 \oplus \cdots \oplus V_r$，则任取 $\mathbf{x} \in V$，在每个 V_i 内存在唯一的 $\mathbf{x}_i$ 使得 $\mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i$。对每个 $i \in \{1, \ldots, r\}$，我们令 $\mathcal{P}_i : V \to V_i$，$\mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i \mapsto \mathbf{x}_i$，则 $\mathcal{P}_1, \ldots, \mathcal{P}_r$ 满足：

$$\mathcal{P}_i\mathcal{P}_j = \delta_{ij}\mathcal{P}_i\ ,\ \ \sum_{i=1}^{r} \mathcal{P}_i = \mathcal{E} \tag{3.5.1}$$

我们称 $\mathcal{P}_1, \ldots, \mathcal{P}_r$ 是**一个完全正交的投影算子组**。实际上，如果一组线性算子满足式 (3.5.1)，则我们就可以找到相对应的直和分解 $V = V_1 \oplus \cdots \oplus V_r$，使得每个 $\mathcal{P}_i$ 就是 V 到 V_i 的投影算子 (证明是第二章的一道习题，留作练习)。

需要说明的是，内积空间上投影算子并不一定都是自伴随的，尽管其特征值只有 0 和 1。这是因为使得投影算子对角化的基底并不一定是标准正交基，换言之，0 对应的特征子空间和 1 对应的特征子空间不一定是正交的。例如，标准欧氏空间 $\mathbb{R}^2$ 上的线性算子 $\mathcal{P} : \mathbf{x} = (x_1, x_2)^t \mapsto A\mathbf{x}$，$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ 就是投影算子 (验证留作练习)，但它显然不是自伴随算子 ($A \neq A^t$)。实际上，投影算子 $\mathcal{P}$ 是自伴随算子 $\iff \ker(\mathcal{P}) \perp \mathrm{im}(\mathcal{P})$，证明是 3.3 节的一道习题，留作练习。

定理 3.5.2 (**谱定理**,the spectral theorem) 设 V 是**一般**的域 $\mathbb{K}$ 上的 n 维向量空间，$\mathcal{A} \in \mathrm{End}(V)$ 是可对角化的线性算子。如果 $\mathcal{A}$ 的全部的两两不同的特征值为 $\lambda_1, \ldots, \lambda_r$，则存在一个完全正交的投影算子组 $\mathcal{P}_1, \ldots, \mathcal{P}_r$ 使得：

(1) $\sum_{i=1}^{r} \mathcal{P}_i = \mathcal{E}$;

(2) $\sum_{i=1}^{r} \lambda_i\mathcal{P}_i = \mathcal{A}$(称为 $\mathcal{A}$ 的谱分解，spectral decomposition)；

(3) 谱分解是唯一的，即如果另有一个完全正交的投影算子组 $Q_1, \ldots, Q_s$ 也满足 $\sum_{i=1}^{s} Q_i = \mathcal{E}$ 和 $\sum_{i=1}^{s} \mu_i Q_i = \mathcal{A}$ ($\mu_1, \ldots, \mu_s$ 两两不同)，那么 $r = s$，并且 $\mu_1, \ldots, \mu_r$ 是 $\lambda_1, \ldots, \lambda_r$ 的一个排列，当 $\mu_i = \lambda_j$ 时，$Q_i = \mathcal{P}_j$。

(4) 存在多项式 $f_1(t), \ldots, f_r(t) \in \mathbb{K}[t]$ 使得 $\forall i, j \in \{1, \ldots, r\}$，$f_i(\lambda_j) = \delta_{ij}$，并且 $f_i(\mathcal{A}) = \mathcal{P}_i$。

证明: (1) 由于 $\mathcal{A}$ 可对角化，故 $V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_r}$，对每个 $i \in \{1, \ldots, r\}$，作线性算子

$$\mathcal{P}_i : V \to V^{\lambda_i},\ \mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i\ (\text{其中}\mathbf{x}_i \in V^{\lambda_i}) \mapsto \mathbf{x}_i$$

则每个 $\mathcal{P}_i$ 都是投影算子，并且 $\forall \mathbf{x} \in V$，有 $(\sum_{i=1}^{r} \mathcal{P}_i)\mathbf{x} = \sum_{i=1}^{r} \mathbf{x}_i = \mathbf{x} = \mathcal{E}\mathbf{x}$，即 $\sum_{i=1}^{r} \mathcal{P}_i = \mathcal{E}$。

(2) 对 $\forall \mathbf{x}=\sum_{i=1}^r \mathbf{x}_i \in V$ (其中$\mathbf{x}_i \in V^{\lambda_i}$)，由于 $\mathbf{x}_i \in V^{\lambda_i}$，故 $\mathcal{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i$，因此

$$\mathcal{A}\mathbf{x}=\mathcal{A}\mathcal{E}\mathbf{x}=\mathcal{A}(\sum_{i=1}^r \mathcal{P}_i)\mathbf{x}=\mathcal{A}(\sum_{i=1}^r \mathbf{x}_i)=\sum_{i=1}^r \mathcal{A}\mathbf{x}_i=\sum_{i=1}^r \lambda_i\mathbf{x}_i.$$

而另一方面，按照 $\mathcal{P}_i$ 的定义我们有

$$(\sum_{i=1}^r \lambda_i\mathcal{P}_i)\mathbf{x}=\sum_{i=1}^r \lambda_i\mathbf{x}_i.$$

因此 $\sum_{i=1}^r \lambda_i\mathcal{P}_i=\mathcal{A}$。

(3) 令 $U_i=\mathrm{im}(Q_i)$, $i=1,\dots,s$，则任取 $Q_i\mathbf{x}\in U_i$(再次注意这个取法)，有

$$\mathcal{A}(Q_i\mathbf{x})=(\sum_{i=1}^s \mu_iQ_i)Q_i\mathbf{x}=\mu_iQ_i^2\mathbf{x}=\mu_iQ_i\mathbf{x}.$$

这说明每个 μ_i 都是 $\mathcal{A}$ 的特征值，由于 μ_i 之间两两不同，故 $s\leqslant r$，并且如果设 $\mu_i=\lambda_j$，则 $U_i\subset V^{\lambda_j}$。于是和空间 $U_1+\cdots+U_s$ 是直和。又因为 $\sum_{i=1}^s Q_i=\mathcal{E}$，所以任取 $\mathbf{x}\in V$，都有 $\mathbf{x}=\mathcal{E}\mathbf{x}=\sum_{i=1}^s(Q_i\mathbf{x})$，即 $V=U_1+\cdots+U_s$。所以 $V=U_1\oplus\cdots\oplus U_s$，结合每个 U_i 都是 V^{λ_j} 的子空间以及 $V=V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_r}$ 可知 $r=s$，并且 $U_1,\dots,U_s$ 就是 $V^{\lambda_1},\dots,V^{\lambda_r}$ 的一个排列，即 $\mu_1,\dots,\mu_r$ 是 $\lambda_1,\dots,\lambda_r$ 的一个排列，当 $\mu_i=\lambda_j$ 时，$Q_i=\mathcal{P}_j$。

(4) 任意固定 $i\in\{1,\dots,r\}$，我们构造出 $f_i(t)$ 即可。由于 $f_i(t)$ 应该满足 $f_i(\lambda_i)=1$ 以及 $\forall j\neq i$, $f_i(\lambda_j)=0$，因此由 Lagrange 插值公式 (讲义上册定理 6.2.4) 可知：

$$f_i(t)=\sum_{k=1}^r\left[f_i(\lambda_k)\cdot\frac{\prod\limits_{m\neq k}(t-\lambda_m)}{\prod\limits_{m\neq k}(\lambda_k-\lambda_m)}\right]=\frac{\prod\limits_{m\neq i}(t-\lambda_m)}{\prod\limits_{m\neq i}(\lambda_i-\lambda_m)}$$

下面我们证明 $f_i(\mathcal{A})=\mathcal{P}_i$。首先我们注意到：

$$\begin{aligned}
\mathcal{A}^0&=\mathcal{E}=\sum_{i=1}^r\mathcal{P}_i\\
\mathcal{A}&=\sum_{i=1}^r\lambda_i\mathcal{P}_i\\
\mathcal{A}^2&=(\sum_{i=1}^r\lambda_i\mathcal{P}_i)(\sum_{i=1}^r\lambda_i\mathcal{P}_i)=\sum_{i=1}^r\lambda_i^2\mathcal{P}_i \quad \text{(展开括号并利用式 (3.5.1) 整理即得.)}\\
\mathcal{A}^3&=(\sum_{i=1}^r\lambda_i\mathcal{P}_i)(\sum_{i=1}^r\lambda_i^2\mathcal{P}_i)=\sum_{i=1}^r\lambda_i^3\mathcal{P}_i\\
&\cdots\cdots\\
\mathcal{A}^k&=(\sum_{i=1}^r\lambda_i\mathcal{P}_i)(\sum_{i=1}^r\lambda_i^{k-1}\mathcal{P}_i)=\sum_{i=1}^r\lambda_i^k\mathcal{P}_i \quad ,\ \forall k\in\mathbb{N}\\
&\cdots\cdots
\end{aligned}$$

于是对任意的 $f(t)\in\mathbb{K}[t]$，都有 $f(\mathcal{A})=\sum_{j=1}^r f(\lambda_j)\mathcal{P}_j$。特别地，取 $f=f_i$，利用 $f_i(\lambda_j)=\delta_{ij}$ 即得 $f_i(\mathcal{A})=\mathcal{P}_i$。这样我们就完成了证明。 □

我们注意到，上面定理的结论对于一般的域都成立。将定理 3.5.1与定理 3.5.2的结论结合起来，我们立刻得到：

命题 3.5.2 (**正规算子谱定理**) 设 $\mathcal{A}$ 是 n 维 Hermitian 空间 V 上的正规算子，则 $\mathcal{A}$ 可对角化，并且 $\mathcal{A}$ 有唯一的谱分解：$\mathcal{A} = \sum_{i=1}^r \lambda_i \mathcal{P}_i$，其中 $\lambda_1, \ldots, \lambda_r$ 是 $\mathcal{A}$ 的全部的两两不同的特征值，$\mathcal{P}_i$ 是如下的投影算子：

$$V \to V^{\lambda_i}, \quad \mathbf{x} = \sum_{i=1}^r \mathbf{x}_i \text{ (其中} \mathbf{x}_i \in V^{\lambda_i}) \mapsto \mathbf{x}_i.$$

需要说明的是，在上面的命题中，每个 $\mathcal{P}_i$ 都是自伴随的投影算子。这来源于正规算子的特殊性质：$V^{\lambda_i} \perp V^{\lambda_j}, \forall i \neq j$，再利用我们在定理 3.5.2之前的那段论述即可得到 $\mathcal{P}_i$ 的自伴随性质。

现在我们考虑如何将两个正规算子用同一组标准正交基对角化。这当然是有条件的。我们有如下定理：

定理 3.5.3 设 $(V, <\cdot,\cdot>)$ 是 n 维 Hermitian 空间，$\mathcal{A}, \mathcal{B} \in \mathrm{End}(V)$ 都是正规算子。则我们有：存在 V 的一组标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使得 $\mathcal{A}, \mathcal{B}$ 在此基底下都是对角矩阵 $\iff \mathcal{AB} = \mathcal{BA}$。

证明: $(\Longrightarrow)$ 方向是显然的，因为两个对角矩阵显然交换。下面证明相反的方向。
$(\Longleftarrow)$ 任取 $\mathcal{A}$ 的某个特征子空间 V^{λ_i}，再任取 $\mathbf{x} \in V^{\lambda_i}$，则

$$\mathcal{AB}\mathbf{x} = \mathcal{BA}\mathbf{x} = \lambda_i \mathcal{A}\mathbf{x} \Longrightarrow \mathcal{B}\mathbf{x} \in V^{\lambda_i}.$$

这说明每个 $\mathcal{A}$ 的特征子空间 V_{λ_i} 都是 $\mathcal{B}$ 的不变子空间。将 $\mathcal{B}$ 限制到 V^{λ_i} 上可知 V^{λ_i} 中一定有 $\mathcal{B}$ 的特征向量 (利用了复数域上的算子一定有特征向量，定理 2.3.2)。于是 $\mathcal{A}$ 和 $\mathcal{B}$ 一定有公共的特征向量，设为 $\mathbf{y}$。取 $\mathbf{e}_1 = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$，则 $\|\mathbf{e}_1\| = 1$ 且 $\mathrm{span}\,\{\mathbf{e}_1\}$ 是 $\mathcal{A}$ 和 $\mathcal{B}$ 公共的特征子空间。

下面我们证明：$W = \mathrm{span}\,\{\mathbf{e}_1\}^\perp$ 同时是 $\mathcal{A}$ 和 $\mathbb{B}$ 的不变子空间，并且 $\mathcal{A}$ 和 $\mathcal{B}$ 限制到 W 上也是正规算子。

任取 $\mathbf{w} \in W$，即 $\mathbf{w} \perp \mathbf{e}_1$，我们有

$$<\mathcal{A}\mathbf{w}, \mathbf{e}_1> = <\mathbf{w}, \mathcal{A}^*\mathbf{e}_1> = <\mathbf{w}, \overline{\lambda_i}\mathbf{e}_1> = \lambda_i <\mathbf{w}, \mathbf{e}_1> = 0 \quad \text{(这里用到了引理 3.5.2的 (3).)}$$

即 $\mathcal{A}\mathbf{w} \in \mathrm{span}\,\{\mathbf{e}_1\}^\perp = W$，所以 W 是 $\mathcal{A}-$ 子空间。对 $\mathcal{B}$ 重复一遍上述过程即可得到 W 也是 $\mathcal{B}-$ 子空间。$\mathcal{A}, \mathcal{B}$ 限制到 W 上是正规算子上显然的。

因此，对 $\mathcal{A}\big|_W$ 和 $\mathcal{B}\big|_W$ 而言，重复我们一开始的操作，我们仍可以在 W 中找到 $\mathcal{A}, \mathcal{B}$ 的公共特征向量 $\mathbf{e}_2$，$\cdots\cdots$ 归纳法即可找到 V 的标准正交基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 使 $\mathcal{A}, \mathcal{B}$ 在此基底下都是对角矩阵。 □

需要注意的是，以上定理在欧氏空间上不对，例如在 $\mathbb{R}^2$ 上取正规矩阵 $A = E_2$，$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，则 $AB = BA$，但 B 不存在实的特征值和特征向量！

在本节的最后。我们来回答 3.3 节遗留下来的一个问题：(半) 正定算子的平方根为什么是唯一的。(由于时间紧迫，以后补充，已经在当时给出了参考文献。)

3.6 复化与实化 *

我们在前面几节中已经看到，实内积空间和复内积空间的性质十分相似，这启发我们考虑实向量空间和复向量空间本身的关系。我们在讲义上册第五章中已经知道，$\mathbb{C}$ 可以视作 $\mathbb{R}$ 上的二维向量空间 $\left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in \mathbb{R}\right\}$，将它推广到 n 维情形就是我们本节的内容。本节内容我们大多数只给出结论，证明可参考代数学引论的相关章节。

首先，设 V 是 $\mathbb{R}$ 上的 n 维向量空间，我们希望通过 V 构造出复向量空间 V'，并且 V' 上的复数数乘可以用 V 上的某种“操作”表示出来。这就是所谓的“复化”(complexification)。

定义 3.6.1 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，$\mathcal{J} \in \mathrm{End}(V)$ 并且 $\mathcal{J}^2 = -\mathcal{E}$，则我们可以在 V 上定义如下的复数数乘：

$$\cdot : \ \mathbb{C} \times V \to V, \ (a + bi \ , \ \mathbf{v}) \mapsto a\mathbf{v} + b\mathcal{J}\mathbf{v} \ (\text{其中} a, b \in \mathbb{R}.) \tag{3.6.1}$$

验证这确实是一个数乘运算 (即验证它满足定义 1.1.1的 (2) 和 (3)) 留作练习。于是 V 在原来的加法和式 (3.6.1) 定义的数乘之下是 $\mathbb{C}$ 上的向量空间，我们把这个向量空间记作 $V_{\mathcal{J}}$，其在 $\mathbb{C}$ 上的维数记作 $\dim_{\mathbb{C}}(V_{\mathcal{J}})$，并称式 (3.6.1) 为 V 上的**复结构**。

我们先来看两个复结构的具体例子。

例 3.6.1 (1) 在 $\mathbb{R}^2$ 上算子 $\mathcal{J} : \mathbf{x} = (x_1, x_2)^t \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathbf{x}$ 可以定义一个复结构：

$$(a + bi) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto a\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + b\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 - bx_2 \\ ax_2 + bx_1 \end{pmatrix}.$$

这实际上就是复数乘法：$(a + bi)(x_1 + x_2 i) = (ax_1 - bx_2) + (ax_2 + bx_1)i$。

(2) 在 $2m$ 维的辛空间 V 上有一个自然的复结构：取辛空间的一组辛基 $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_m}, \boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_m}$，则我们可以定义算子 $\mathcal{J}$ 使得 $\mathcal{J}$ 在这组辛基下的矩阵恰为 $J_0 = \begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$。容易验证 $\mathcal{J}$ 是辛算子，并且 $\mathcal{J}^2 = -\mathcal{E}$，于是可以用 $\mathcal{J}$ 在 V 上定义一个自然的复结构。

我们有以下定理：

定理 3.6.1 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，如果 V 上存在一个用线性算子 $\mathcal{J}$ ($\mathcal{J}^2 = -\mathcal{E}$) 定义的复结构，则 V 的实维数 (记为 $\dim_{\mathbb{R}}(V)$) 是偶数 $n = 2m$。更进一步，我们一定能找到某组基 $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_m}, \boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_m}$ 使得 $\mathcal{J}$ 在这组基下的矩阵形如 $\begin{pmatrix} 0 & -E_m \\ E_m & 0 \end{pmatrix}$，并且 $\dim_{\mathbb{C}}(V_{\mathcal{J}}) = \frac{1}{2}\dim_{\mathbb{R}}(V) = m$。

证明过程实际上就是寻找斜对称双线性型的标准基，留作思考。

于是，我们可以从一般的实向量空间 V 出发，先构造出偶数维的 $V \oplus V$(外直和，下面会严格定义)，再在 $V \oplus V$ 上定义一个特殊的复结构就可以得到一个复维数与 $\dim_{\mathbb{R}}(V)$ 相等的向量空间。这就是下面的复化。

定义 3.6.2 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，则容易看出 $V\times V$ 在下面的加法和实数乘法下是 $\mathbb{R}$ 上的 $2n$ 维的向量空间：

$$+:\ (V\times V)\times(V\times V)\to(V\times V),\ (\mathbf{x}_1,\mathbf{x}_2)+(\mathbf{y}_1,\mathbf{y}_2)=(\mathbf{x}_1+\mathbf{y}_1,\mathbf{x}_2+\mathbf{y}_2);$$
$$\cdot_{\mathbb{R}}:\ \mathbb{R}\times(V\times V)\to(V\times V),\ a\cdot(\mathbf{x}_1,\mathbf{x}_2)=(a\mathbf{x}_1,a\mathbf{x}_2).$$

我们把 $V\times V$ 在上述运算下构成的向量空间称为 V 和 V 的外直和，记作 $V\oplus V$ [1]。设 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，则容易看出 $(\mathbf{e}_1,\mathbf{0}),\dots,(\mathbf{e}_n,\mathbf{0}),(\mathbf{0},\mathbf{e}_1),\dots,(\mathbf{0},\mathbf{e}_n)$ 是 $V\oplus V$ 的一组基。显然在 $V\oplus V$ 上线性算子 $\mathcal{J}:(\mathbf{u},\mathbf{v})\mapsto(-\mathbf{v},\mathbf{u})$ 满足 $\mathcal{J}^2=-\mathcal{E}$ (实际上 $\mathcal{J}$ 在上述基底下的矩阵恰好是 $\begin{pmatrix}0&-E_n\\E_n&0\end{pmatrix}$)，于是 $\mathcal{J}$ 可以定义 $V\oplus V$ 上的一个复结构，称为标准复结构。相应地，我们把复向量空间 $(V\oplus V)_{\mathcal{J}}$ 称为 V 的复化空间，记作 $V^{\mathbb{C}}$。

我们也可以把 $(V\oplus V)_{\mathcal{J}}$ 中的向量 $(\mathbf{u},\mathbf{v})$ 记作 $\mathbf{u}+i\mathbf{v}$。容易验证

$$(a+bi)(\mathbf{u}+i\mathbf{v})=a(\mathbf{u}+i\mathbf{v})+b(-\mathbf{v}+i\mathbf{u})=(a\mathbf{u}-b\mathbf{v})+i(a\mathbf{v}+b\mathbf{u}).$$

即这样的形式记号与复数数乘是相容的。稍后我们会看到，这种记号实际上是先复化，再实化。特别地，我们看到 $\mathbb{C}$ 就是 $\mathbb{R}$ 的复化。由定理 3.6.1立刻有：$\dim_{\mathbb{C}}(V^{\mathbb{C}})=\dim_{\mathbb{R}}(V)$。实际上，我们很容易验证：如果 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 V 的一组基，则 $\mathbf{e}_1+i\mathbf{0},\dots,\mathbf{e}_n+i\mathbf{0}$ 也是 $V^{\mathbb{C}}$ 的一组基。

下面我们考虑线性算子的复化。

定义 3.6.3 设 V 是 $\mathbb{R}$ 上的 n 维向量空间，$\mathcal{A}\in\mathrm{End}(V)$，则我们可以定义 $\mathcal{A}$ 的复化算子 $\mathcal{A}^{\mathbb{C}}:V^{\mathbb{C}}\to V^{\mathbb{C}}$，$\mathbf{u}+i\mathbf{v}\mapsto\mathcal{A}\mathbf{u}+i\mathcal{A}\mathbf{v}$。

利用复化算子我们也可以很容易地证明：实向量空间上的线性算子一定有 1 维或 2 维的不变子空间，留作思考。

下面我们考虑相反方向的问题：给出一个 n 维复向量空间 U，我们如何得到一个实向量空间呢？这个方向是简单的，我们只需要把数乘运算限制到实数域上就可以了。

定义 3.6.4 设 U 是 n 维复向量空间，$\mathbf{e}_1,\dots,\mathbf{e}_n$ 是 U 的一组基，则显然 $U=\{\sum_{k=1}^n a_k\mathbf{e}_k+\sum_{k=1}^n b_k(i\mathbf{e}_k)\mid a_k,b_k\in\mathbb{R}\}$，即我们可以把 U 视作 $\mathbb{R}$ 上的 $2n$ 维实向量空间，称作 U 的实化空间，记为 $U_{\mathbb{R}}$。同样地，U 上的线性算子 $\mathcal{A}$ 也可以视作 $U_{\mathbb{R}}$ 上的线性算子，称为 $\mathcal{A}$ 的实化算子，记作 $\mathcal{A}_{\mathbb{R}}$。

设 $\mathcal{A}\in\mathrm{End}(U)$ 在 U 的基底 $\mathbf{e}_1,\dots,\mathbf{e}_n$ 下的矩阵为 $A=A_1+iA_2$，$A_1,A_2\in M_n(\mathbb{R})$，则

$$\mathcal{A}(i\mathbf{e}_1,\dots,i\mathbf{e}_n)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot i(A_1+iA_2)=(\mathbf{e}_1,\dots,\mathbf{e}_n)\cdot(-A_2+iA_1).$$

因此 $\mathcal{A}_{\mathbb{R}}$ 在 $\mathbf{e}_1,\dots,\mathbf{e}_n,i\mathbf{e}_1,\dots,i\mathbf{e}_n$ 下的矩阵为 $A_{\mathbb{R}}=\begin{pmatrix}A_1&-A_2\\A_2&A_1\end{pmatrix}$。由此可见，一个 $U_{\mathbb{R}}$ 上的算子要想称为 U 中某个算子的实化是需要一定条件的。

定理 3.6.2 设 U 是 n 维复向量空间，$\mathcal{A}\in\mathrm{End}(U)$，则 $\det(\mathcal{A}_{\mathbb{R}})=|\det(\mathcal{A})|^2$。

证明留作思考。

容易验证 (过程留作练习)$U_{\mathbb{R}}$ 上所有能称为某个 $\mathrm{End}(U)$ 中算子的实化算子的线性算子构成了一个代数，我们把这个代数记作 $\mathrm{End}(U)_{\mathbb{R}}$，它显然是 $\mathrm{End}(U_{\mathbb{R}})$ 的一个子代数。实际上，可以证明 $\dim(\mathrm{End}(U)_{\mathbb{R}})=\dfrac{1}{2}\dim(\mathrm{End}(U_{\mathbb{R}}))$，留作思考。那么，$\mathrm{End}(U)_{\mathbb{R}}$ 中的元素是不是都具有 $\begin{pmatrix}A_1&-A_2\\A_2&A_1\end{pmatrix}$ 这样的矩阵形式呢？答案是肯定的，这就是下面的定理。

[1]为了与 1.2 节的直和 (称为内直和) 区分，也有的书将外直和记作 $V\boxplus V$。

定理 3.6.3 设 U 是 n 维复向量空间，$\mathbf{e}_1, \dots, \mathbf{e}_n$ 是 U 的一组基，则 $\mathrm{End}(U)_{\mathbb{R}}$ 由所有与 $\mathcal{J}$ 交换的算子构成，其中 $\mathcal{J} \in \mathrm{End}(U_{\mathbb{R}})$ 在 $U_{\mathbb{R}}$ 的基底 $\mathbf{e}_1, \dots, \mathbf{e}_n, i\mathbf{e}_1, \dots, i\mathbf{e}_n$ 下的矩阵是 $\begin{pmatrix} 0 & -E_n \\ E_n & 0 \end{pmatrix}$。换言之，任取 $\mathcal{A}_{\mathbb{R}} \in \mathrm{End}(U)_{\mathbb{R}}$，一定存在 $\mathcal{U}_{\mathbb{R}}$ 的一组基使得 $\mathcal{A}_{\mathbb{R}}$ 具有 $\begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix}$ 的矩阵形式。

证明留作思考。我们称这种情况为 $\mathcal{A}$ 与实向量空间 V(实际上等于 $U_{\mathbb{R}}$) 上的某个复结构相容。

特别地，我们有：

定理 3.6.4 二维实向量空间上的没有特征向量的线性算子一定与某个复结构相容。即设 V 是 2 维实向量空间，$\mathcal{A} \in \mathrm{End}(V)$，则 V 是某个一维复向量空间 U 的实化，并且存在 $\mathcal{B} \in \mathrm{End}(U)$ 使得 $\mathcal{B}_{\mathbb{R}} = \mathcal{A}$。

证明留作思考。

接下来我们考虑将实向量空间先复化，再实化。设 V 是 n 维实向量空间，令 $W = (V^{\mathbb{C}})_{\mathbb{R}}$，则容易验证 $W = V \oplus iV$，我们称 V 是 W 的实平面，iV 是 W 的虚平面。令 $\mathcal{J} = (i\mathcal{E})_{\mathbb{R}} \in \mathrm{End}(W)$，则 $\mathcal{J}$ 交换实平面与虚平面。

我们知道一个复数可以取其共轭复数，类似地，我们可以定义 W 上的取共轭如下：$\overline{\mathbf{u} + i\mathbf{v}} = \mathbf{u} - i\mathbf{v}$。进一步，我们可以定义一个算子 $\mathcal{A} \in V^{\mathbb{C}}$ 的复共轭算子：$\overline{\mathcal{A}} : \mathbf{u} + i\mathbf{v} \mapsto \overline{\mathcal{A}(\mathbf{u} + i\mathbf{v})}$。容易验证 $\mathcal{A} = \overline{\mathcal{A}} \iff \mathcal{A}$ 限制到实平面上的像集仍在实平面内，即 $\mathcal{A}$ 是某个 $\mathcal{B} \in \mathrm{End}(V)$ 的复化算子。此外容易验证下面的等式：

$$\mathrm{tr}(\mathcal{A}_{\mathbb{R}}) = \mathrm{tr}(\mathcal{A}) + \mathrm{tr}(\overline{\mathcal{A}}).$$

在本节的最后，我们相对应地考虑将复向量空间先实化，再复化。首先，设 V 是 n 维复向量空间，则我们可以自然地定义 V 的复共轭空间 $\overline{V}$：$\overline{V}$ 中的元素和其上的加法与 V 一致，而数乘 $\odot$ 则定义成：$\lambda \odot \mathbf{x} = \overline{\lambda}\mathbf{x},\ \forall \lambda \in \mathbb{C}, \mathbf{x} \in V$。我们可以证明：$(V_{\mathbb{R}})^{\mathbb{C}}$ 同构于 $V \oplus \overline{V}$，过程留作思考。

3.7 正交展开 *

这一节我们简单地介绍一下无穷维内积空间的部分性质。要想研究无穷维空间，一个不可避免的问题就是在无穷维空间中我们能否找到与有限维空间的基底性质相似的数学概念，这就是我们本节的主题。

定义 3.7.1 设 $(V,<\cdot,\cdot>)$ 是无穷维的实 (或复) 内积空间，$\{\mathbf{e}_n\}_{n=1}^{\infty}\subset V$ 是一个所有向量都非零的序列。如果 $\forall i,j\in\mathbb{Z}^+$，都有 $\mathbf{e}_i\perp\mathbf{e}_j$，则称 $\{\mathbf{e}_n\}$ 是一个可数的正交系；如果 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 还满足 $\|\mathbf{e}_i\|=1,\forall i\in\mathbb{Z}^+$，则称 $\{\mathbf{e}_n\}$ 是一个标准的 (也称规范的) 正交系。容易验证一个正交系中任意有限个向量都是线性无关的[1]。如果正交系 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 还满足 $\operatorname{span}(\{\mathbf{e}_n\}_{n=1}^{\infty})^{\perp}=\{\mathbf{0}\}$，即 $(<\mathbf{v},\mathbf{e}_i>=0,\forall i\in\mathbb{Z}^+\Longrightarrow\mathbf{v}=\mathbf{0})$，那么我们称 $\{\mathbf{e}_n\}_{n=1}^{\infty}$ 是完备的正交系 (也有的书称其为完全正交系)。如果一个正交系既是标准的又是完备的，则我们称其为标准完备正交系。

我们希望标准完备正交系在无穷维内积空间中的作用正如标准正交基在有限维内积空间中的作用，也就是说，我们希望 $\forall\mathbf{x}\in V$，有 $\mathbf{x}=\sum_{i=1}^{\infty}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$。这就产生了两个问题：一是 $\sum_{i=1}^{\infty}<\mathbf{x},\mathbf{e}_i>\mathbf{e}_i$ 何时是 (以何种意义) 收敛的，二是如果收敛那么极限何时恰好是 $\mathbf{x}$。下面我们逐步回答这两个问题。

引理 3.7.1 设 $(V,<\cdot,\cdot>)$ 是实 (或复) 内积空间，则内积 $<\mathbf{x},\mathbf{y}>$ 关于坐标 $(\mathbf{x},\mathbf{y})$ 是连续的。换言之，任取序列 $\{\mathbf{x}_n\}_{n=1}^{\infty}\xrightarrow{n\to\infty}\mathbf{x}_0$，$\{\mathbf{y}_n\}_{n=1}^{\infty}\xrightarrow{n\to\infty}\mathbf{y}_0$(按距离收敛)，则 $<\mathbf{x}_n,\mathbf{y}_n>\xrightarrow{n\to\infty}<\mathbf{x},\mathbf{y}>$。

利用 CBS 不等式即可证明，留作思考。

容易验证以下事实：$\forall\mathbf{x}\in V$，如果 $\mathbf{x}=\sum_{i=1}^{\infty}a_i\mathbf{e}_i$，$a_i\in\mathbb{R}$(或$\mathbb{C}$)，那么每个 $a_i=<\mathbf{x},\mathbf{e}_i>$ 都必然是唯一的，留作练习。我们称 $a_i=<\mathbf{x},\mathbf{e}_i>$ 是向量 $\mathbf{x}$ 关于标准正交系 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 的傅里叶 (Fourier) 系数。

引理 3.7.2 (Bessel **不等式**) 设 $(V,<\cdot,\cdot>)$ 是实 (或复) 内积空间，$\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是标准正交系，则 $\forall\mathbf{x}\in V$，都有 $\|\mathbf{x}\|^2\geqslant\sum_{i=1}^{\infty}|<\mathbf{x},\mathbf{e}_i>|^2$。

证明留作思考，我们在 3.1 节的脚注中已经给出了参考文献。

定理 3.7.1 在完备内积空间 (即柯西列都收敛,Hilbert 空间)$(V,<\cdot,\cdot>)$ 中，对于完备的标准正交系 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 而言，任取 $\mathbf{x}\in V$，部分和 $\sum_{i=1}^{n}<\mathbf{x},\mathbf{e}_i>$ 在 $n\to\infty$ 时都按距离收敛到 $\mathbf{x}$ 本身。

证明可参考《A Course in Functional Analysis》, Jorn B.Conway, GTM96 的 Chapter I, §4, Theorem4.13，或者《泛函分析讲义》，张恭庆，北京大学出版社的第一章，定理 1.6.25。

注意，如果内积空间不是完备的，那么部分和 $\sum_{i=1}^{n}<\mathbf{x},\mathbf{e}_i>$ 在 $n\to\infty$ 时甚至可以不收敛，反例留作思考。

定理 3.7.2 设 $(V,<\cdot,\cdot>)$ 是实 (或复) 内积空间，则一定存在完备内积空间 $\hat{V}$ 使得：

1. $\hat{V}$ 中有一个子空间 V_1 同构于 V；

2. $\hat{V}$ 中每个向量都是 V_1 中的某个柯西序列的极限，即 V_1 在 V 中稠密。

我们称 $\hat{V}$ 是 V 的完备化，并且在同构意义下完备化是唯一的。

[1]我们不妨证明 $\mathbf{e}_1,\ldots,\mathbf{e}_n$ 线性无关。设 $\sum_{i=1}^{n}a_i\mathbf{e}_i=\mathbf{0}$，等式两边与每个 $\mathbf{e}_i$ 作内积即得 $a_i\|\mathbf{e}_1\|^2=0,\forall i$，而 $\|\mathbf{e}_i\|\neq0$，故每个 $a_i=0$。此即线性无关。

完备化的过程实际上可以推广到一般的度量空间上，过程可参考《泛函分析讲义》，张恭庆，北京大学出版社的第一章 §2。

定理 3.7.3 设 $(V, <\cdot,\cdot>)$ 是实 (或复) 内积空间，如果 V 的子空间 U 在 V 中稠密 (即 U 的闭包包含 V，详细的定义参考分析学的标准教材)，那么 U 在 $\hat{V}$ 中也稠密。

证明留作思考。

此外，我们很容易证明 Gram-Schmidt 正交化方法在无穷维空间中仍然能够使用 (留作练习)。于是将以上的结论都结合起来，我们有：

定理 3.7.4 设 $(V, <\cdot,\cdot>)$ 是实 (或复) 内积空间，向量序列 $\{\mathbf{v}_i\}_{i=1}^{\infty} \subset V$ 并且其中任意有限个向量线性无关，对 $\{\mathbf{v}_i\}_{i=1}^{\infty}$ 施以 Gram-Schmidt 正交化得到向量序列 $\{\mathbf{e}_i\}_{i=1}^{\infty}$。

(1) 如果 span $\{\{\mathbf{v}_i\}_{i=1}^{\infty}\}$ 在 V 中稠密，则 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是 $\hat{V}$ 中的标准完备正交系，从而 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 也是 V 的标准完备正交系。

(2) 如果 V 是完备的，则 V 中任何标准完备正交系张成的子空间在 V 中稠密。

证明留作练习。

推论 3.7.1 设 $\{\mathbf{e}_i\}_{i=1}^{\infty}$ 是内积空间 V 的标准正交系，如果 span $\{\{\mathbf{e}_i\}_{i=1}^{\infty}\}$ 在 V 中稠密，则 $\forall \mathbf{x} \in V$，都有 $\mathbf{x} = \sum_{i=1}^{\infty} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i$。

最后我们看两个例子。

例 3.7.1 (1) 闭区间 $[a,b]$ 上的连续函数空间 $C[a,b]$ 在内积 $< f, g >= \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$ 下不是完备的，但是，容易验证 span $\{t^n \mid \forall n \in \mathbb{N}\}$ 在 $C[a,b]$ 中是稠密的 (数学分析中的 Weierstrass 逼近定理)，故对 $\{t^n \mid \forall n \in \mathbb{N}\}$ 进行 Gram-Schmidt 正交化即可得到 $C[a,b]$ 的一个标准完备正交系。这是稍后正交多项式的讨论内容。

(2) 在 $L^2[0,2\pi]$(平方可积函数构成的内积空间，相关定义可参考标准的实分析或泛函分析教材) 中，定义内积 $< f, g >= \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$，这是一个完备内积空间，容易验证 $e_n(t) = \dfrac{1}{\sqrt{2\pi}} e^{in\pi}$，$n = 0, \pm 1, \pm 2, \ldots$ 是一个标准完备正交系。这是傅里叶分析 (傅里叶级数) 讨论的内容。

3.8 正交投影与最小二乘法

这一节我们考虑如下问题: 设 $(V, <\cdot,\cdot>)$ 是 n 维内积空间, U 是 V 的 m 维子空间, 我们定义 $\mathbf{x}\in V$ 到 U 的距离 $\mathrm{d}(\mathbf{x},U)=\inf\limits_{\mathbf{y}\in U}\mathrm{d}(\mathbf{x},\mathbf{y})$，那么我们会问：是否存在 $\mathbf{y}_0\in U$ 使得 $\mathrm{d}(\mathbf{x},U)=\mathrm{d}(\mathbf{x},\mathbf{y_0})$? 如果存在，那么 $\mathbf{y}_0$ 是否是唯一的，以及如何计算 $\mathbf{y}_0$？我们接下来会逐步回答这些问题。

定理 3.8.1 条件如上所述，则满足上述要求的 $\mathbf{y}_0$ 是存在唯一的。实际上，由定理 3.1.3或定理 3.2.3，设 $V=U\oplus U^{\perp}$，则 $\mathbf{x}$ 可以唯一地分解成 $\mathbf{x}=\mathbf{y}_0+\mathbf{z}$, $\mathbf{y}_0\in U,\mathbf{z}\in U^{\perp}$，于是 $\mathbf{y}_0$ 即为我们所求。

证明: 我们只需证明 $\forall\mathbf{u}\in U$，都有 $\mathrm{d}(\mathbf{x},\mathbf{u})\geqslant\mathrm{d}(\mathbf{x},\mathbf{y}_0)$，并且等号仅在 $\mathbf{u}=\mathbf{y}_0$ 处成立即可。事实上，注意到 $\mathbf{z}=\mathbf{x}-\mathbf{y}_0\in U^{\perp}$，而 $\mathbf{y}_0-\mathbf{u}\in U$，故 $(\mathbf{x}-\mathbf{y}_0)\perp(\mathbf{y}_0-\mathbf{u})$, $\forall\mathbf{u}\in U$，因此我们有:

$$\begin{aligned}||\mathbf{x}-\mathbf{u}||^2&=||(\mathbf{x}-\mathbf{y}_0)+(\mathbf{y}_0-\mathbf{u})||^2\\&=||\mathbf{x}-\mathbf{y}_0||^2+||\mathbf{y}_0-\mathbf{u}||^2\quad\text{(勾股定理)}\\&\geqslant||\mathbf{x}-\mathbf{y}_0||^2\end{aligned}$$

即 $[\mathrm{d}(\mathbf{x},\mathbf{u})]^2\geqslant[\mathrm{d}(\mathbf{x},\mathbf{y}_0)]^2$，并且等号当且仅当 $||\mathbf{y}_0-\mathbf{u}||=0$，即 $\mathbf{y}_0=\mathbf{u}$ 时成立。这样我们就完成了证明。 □

在上面的条件下，我们定义投影算子 $\mathcal{P}_U: V\to U,\ \mathbf{x}\mapsto\mathbf{y}_0$，则我们称 $\mathcal{P}_U$ 是正交投影算子，$\mathbf{y}_0$ 是向量 $\mathbf{x}$ 到 U 的正交投影。

例 3.8.1 设 $V=C[-\pi,\pi]$ 是闭区间 $[-\pi,\pi]$ 上的实值连续函数空间，内积定义成 $<f,g>=\int_{-\pi}^{\pi}f(t)g(t)\,\mathrm{d}t$，设 $U=\mathrm{span}\,\{1,t,t^2,t^3,t^4,t^5\}$，求 $f(t)=\sin t$ 到 U 的正交投影以及 $\mathrm{d}(f,U)$。

解: 对 U 的基底 $1,\dots,t^5$ 作 Gram-Schmidt 正交化, 可以得到 U 的一组标准正交基 $\mathbf{e}_1(t),\dots,\mathbf{e}_6(t)$。于是 f 到 U 的正交投影 $\mathcal{P}_U(f)$ 为:

$$\mathcal{P}_U(f)=\sum_{i=1}^{6}<f(t),\mathbf{e}_i(t)>\mathbf{e}_i(t)=\sum_{i=1}^{6}\left(\int_{-\pi}^{\pi}[\mathbf{e}_i(s)\sin s]\,\mathrm{d}s\right)\cdot\mathbf{e}_i(t)\approx0.98-0.155t^3+0.0056t^5$$

并且 $\mathrm{d}(f,U)=\sqrt{\int_{-\pi}^{\pi}(f(t)-\mathcal{P}_U(f)(t))^2\,\mathrm{d}t}\approx0.001$。读者可以自行验证在题设距离的意义下这个逼近比 Taylor 展开逼近要精确。

下面我们来考虑一个更常见的计算场景。设 $A\in\mathbb{R}^{m\times n}$, $m>n$, $\mathbf{b}\in\mathbb{R}^m$，未定向量 $\mathbf{x}\in\mathbb{R}^n$，那么我们知道方程组 $A\mathbf{x}=\mathbf{b}$ 不一定有解。但是，按照上面的讨论，我们总可以在 $\mathbb{R}^n$ 中找到 $\mathbf{x}_0$ 使得在 $\mathbb{R}^m$ 的标准内积下 $\mathrm{d}(\mathbf{A}\mathbf{x}_0,\mathbf{b})$ 最小，因为在 $\mathbf{x}_0$ 取遍 $\mathbb{R}^n$ 时 $A\mathbf{x}_0$ 恰好取遍 A 的列空间 $V_c(A)$ (把 A 视作线性算子则是 $\mathrm{im}(A)$)，于是我们只需寻找 $\mathbf{b}$ 到 $V_c(A)$ 的正交投影即可找到 $A\mathbf{x}_0$。下面我们开始计算这个正交投影。

我们不妨设 $\mathrm{rank}(A)=r$ 并且 A 的列空间是 $\mathrm{span}\,\{\mathbf{A}^{(1)},\dots,\mathbf{A}^{(r)}\}$，则对 $\mathbf{A}^{(1)},\dots,\mathbf{A}^{(r)}$ 施以 Gram-Schmidt 正交化可以得到 $V_c(A)$ 的一组标准正交基 $\mathbf{e}_1,\dots,\mathbf{e}_r\in\mathbb{R}^m$，令

$$\mathbf{b}_0=\sum_{k=1}^{r}<\mathbf{b},\mathbf{e}_k>\mathbf{e}_k\in\mathbb{R}^m,$$

其中 $<\cdot,\cdot>$ 是 $\mathbb{R}^m$ 上的标准内积。则 $\mathbf{b}_0$ 就是 $\mathbf{b}$ 到 $V_c(A)$ 的正交投影。之后，我们求解非齐次线性方程组 $A\mathbf{x}_0=\mathbf{b}_0$，则 $\mathbf{x}_0\in\mathbb{R}^n$ 即是我们所求。注意方程组 $A\mathbf{x}_0=\mathbf{b}_0$ 总是存在唯一解，这是因为我们很容易验证 $\mathrm{rank}(A)=\mathrm{rank}(A\mid\mathbf{b}_0)=r$，即系数矩阵的秩等于增广矩阵的秩。

事实上, 如果 $\mathrm{rank}(A) = n$, 那么我们可以更简单地写出 $\mathbf{x}_0$。首先, 此时 $\mathrm{rank}(A^tA) = \mathrm{rank}(A) = n$(引理 3.5.1), 于是 A^tA 可逆。设 $\mathbf{b} = A\mathbf{x}_0 + \mathbf{w}$ 是 $\mathbf{b}$ 沿着 $\mathbb{R}^m = V_c(A) \oplus (V_c(A))^\perp$ 的直和分解, 即 $A\mathbf{x}_0 \perp \mathbf{w} = \mathbf{b} - A\mathbf{x}_0$, 则

$$< A\mathbf{x}_0, \mathbf{b} - A\mathbf{x}_0 >= \mathbf{x}_0^t A^t(\mathbf{b} - A\mathbf{x}_0) = 0, \tag{3.8.1}$$

即 $\mathbf{x}_0^t(A^tA\mathbf{x}_0 - A^t\mathbf{b}) = 0$。故取 $\mathbf{x}_0 = (A^tA)^{-1}A^t\mathbf{b}$ 时, 式 (3.8.1) 成立, 而由正交投影的唯一性可知 $\mathbf{x}_0$ 是唯一的, 即为所求。

我们把上面的求解正交投影的方法称为**最小二乘法** (ordinary least squares)。最小二乘法在实际问题中有广泛的应用, 如统计学中的回归分析即是最小二乘法的应用。有关最小二乘法及回归分析的内容, 读者可以参考《线性统计模型: 线性回归与方差分析》, 王松桂, 陈敏, 陈丽萍, 高等教育出版社。

3.9 正交多项式

在本章的最后，我们考虑多项式空间 $\mathbb{R}[t]$ 在不同的内积下的正交基，它们被称为正交多项式。

首先，我们考虑内积 $<f,g>=\int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$。在此内积下我们对基底 $1,t,t^2,\cdots\cdots$ 施以 Gram-Schmidt 正交化 (不对向量的范数归一，只进行正交化的部分)，则可以得到 $\mathbb{R}[t]$ 的一组正交基：

$$u_0(t)=1,\ u_1(t)=t,\ u_2(t)=t^2-\frac{1}{3},\ u_3(t)=t^3-\frac{3}{5}t,\cdots\cdots$$

这些正交基的一般公式如下：

$$u_n(t)=\frac{2^n(n!)^2}{(2n)!}P_n(t),\ \text{其中}\ P_n(t)=\frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n.$$

可以证明 $P_n(t)$ 之间满足如下的正交性质：

$$\int_{-1}^{1} P_i(t)P_j(t)\,\mathrm{d}t=\frac{2}{2j+1}\delta_{ij}.$$

我们把上面出现的 $P_n(t)$ 的显式表达式称为 Legendre(勒让德) 多项式的 Rodrigue 公式。Rodrigue 公式的证明可以参考《数值逼近》，王仁宏，高等教育出版社的第四章 §7 或《函数逼近论方法》，莫国端，刘开第，科学出版社的 §5.2。

Legendre 多项式也可以由递推式 $P_0(t)=1,P_1(t)=t,P_{n+1}(t)=\dfrac{2n+1}{n+1}tP_n(t)-\dfrac{n}{n+1}P_{n-1}(t)$ 定义，可以验证这个定义与 Rodrigue 公式是等价的。这是因为：在 $\mathbb{R}[t]$ 上定义自伴随算子 (验证之)：

$$\mathcal{S}(f(t))=\frac{\mathrm{d}}{\mathrm{d}t}\left((t^2-1)\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right)$$

验证：$\mathcal{S}P_n(t)=n(n+1)P_n(t),\forall n\in\mathbb{N}$ 即可得到上面的递推式。细节留作练习或者仍参考上面的两篇文献。

Legendre 多项式是 $\mathbb{R}[t]$ 的完备正交系。利用分部积分公式证明任意的多项式 $Q(x)$ 如果与 $P_n(x)$ 正交，则 $Q(x)=0$ 即可。细节留作练习。

下面我们考虑 $\mathbb{R}[t]$ 在其它内积下的正交系，它们被称为加权正交多项式。

1. 第一类 Chebyshev 多项式：内积 $<f,g>=\int_{-1}^{1} f(t)g(t)\dfrac{1}{\sqrt{1-t^2}}\,\mathrm{d}t$。则此基底下的完备正交系是

   $$T_n(t)=\cos(n\arccos t),\ \forall n\in\mathbb{N}.$$

   我们有

   $$\int_{-1}^{1} T_i(t)T_j(t)\,\mathrm{d}t=\begin{cases}\pi, i=j=0\\ \dfrac{\pi}{2}, i=j\neq 0\\ 0, i\neq j\end{cases}$$

   可以验证 $T_n(t)$ 是算子 $\mathcal{T}=(t^2-1)\dfrac{\mathrm{d}^2}{\mathrm{d}t^2}-t\dfrac{\mathrm{d}}{\mathrm{d}t}$ 的特征向量，并且 $\mathcal{T}T_n(t)=n^2T_n(t)$。

2. 第二类 Chebyshev 多项式：内积 $<f,g>=\int_{-1}^{1} f(t)g(t)\sqrt{1-t^2}\,\mathrm{d}t$。则此基底下的完备正交系是

   $$u_n(\cos t)=\frac{\sin(n+1)t}{\sin t},\ \forall n\in\mathbb{N}.$$

我们有

$$\int_{-1}^{1} u_i(t)u_j(t)\,\mathrm{d}t = \frac{\pi}{2}\delta_{ij}.$$

可以验证 $u_n(t)$ 满足 $(1-t^2)\frac{\mathrm{d}^2}{\mathrm{d}t^2}u_n(t) - 3t\frac{\mathrm{d}}{\mathrm{d}t}u_n(t) + n(n+2)u_n(t) = 0$，并且 $T_n(t) = u_n(t) - tu_{n-1}(t)$。

3. Hermite 多项式：内积 $< f, g >= \int_{-\infty}^{+\infty} f(t)g(t)e^{-t^2}\,\mathrm{d}t$。则此基底下的完备正交系 $\{H_n(t)\}_{n=1}^{\infty}$ 满足：$H_0(t) = 1, H_1(t) = 2t, H_{n+1}(t) = 2tH_n(t) - \frac{\mathrm{d}}{\mathrm{d}t}H_n(t)$，并且

$$\int_{-\infty}^{\infty} H_i(t)H_j(t)\,\mathrm{d}t = 2^j j!\,\sqrt{\pi}\delta_{ij}$$

可以验证 $T_n(t)$ 是算子 $\mathcal{H} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} - 2t\frac{\mathrm{d}}{\mathrm{d}t}$ 的特征向量，并且 $\mathcal{H}T_n(t) = -2nH_n(t)$。

有关正交多项式的更多性质，读者可以自行阅读上面的两篇文献。