线性代数 II 笔记

The Notes of Linear Algebra II

丁毅

中国科学院大学，北京 100049

Yi Ding

University of Chinese Academy of Sciences, Beijing 100049, China

2024.2.4 ～

序言

本文为笔者本科时的线性代数 II 笔记。用灰色字体或灰色方框等表示对主干内容的补充、对晦涩概念的理解、定理的具体证明过程等，采用红色字体对重点部分进行强调，同时适当配有插图。这样的颜色和结构安排既突出了知识的主要框架，也保持了笔记的深度和广度，并且不会因为颜色过多而导致难以锁定文本内容，乃是尝试了多种安排后挑选出的最佳方案。如果读者有更佳的颜色和排版方案，可以将建议发送到笔者邮箱，在此感谢。

由于个人精力及知识水平有限，书中难免有不妥、错误之处，望不吝指正，在此感谢。

目录

第1章　空间与形式

1.1　线性空间

下面两小节是上学期未讲完的内容。

Theorem. 1 (牛顿公式):
在 $\mathbb{F}[x_1,...,x_n]$ 中，设 $k\in\mathbb{N}_+$，规定下面的常见记号：

$$\mathrm{sym}(x_1^{k_1}\cdots x_r^{k_r})=\sum_{1\le i_1<\cdots<i_r\le n}x_{i_1}^{k_1}\cdots x_{i_r}^{k_r}\ ,\ \ s_k=\mathrm{sym}(x_1^k)=\sum_{i=1}^n x_i^k$$

$$\sigma_k=\begin{cases}1 & ,k=0\\ \mathrm{sym}(x_1\cdots x_k)=\sum\limits_{1\le i_1<\cdots<i_r\le n}x_{i_1}x_{i_2}\cdots x_{i_k} & ,k\in[1,n]\\ 0 & ,k>n\end{cases}$$

则 $\forall\, k\in\mathbb{N}_+$，有牛顿公式：

$$\sum_{i=0}^k(-1)^i\sigma_i s_{k-i}=0\Longleftrightarrow s_k=(-1)^k\sum_{i=1}^k\sigma_i s_{k-i}$$

推论：

$$s_k=\begin{vmatrix}\sigma_1 & 2\sigma_2 & 3\sigma_3 & \cdots & (k-1)\sigma_{k-1} & k\sigma_k\\ 1 & \sigma_1 & \sigma_2 & \cdots & \sigma_{k-2} & \sigma_{k-1}\\ 0 & 1 & \sigma_1 & \cdots & \sigma_{k-3} & \sigma_{k-2}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & \sigma_1 & \sigma_2\\ 0 & 0 & 0 & \cdots & 1 & \sigma_1\end{vmatrix}\ ,\ \ \forall\, k\in\mathbb{N}_+$$

用初等对称多项式表示对称多项式：

给定任意对称多项式 $f(x_1,...,x_n)$，将其表示为 $\varphi(\sigma_1,...,\sigma_n)=f(x_1,...,x_n)$ 的步骤如下：
① 确定支配项 (无需确定系数): $f(x_1,...,x_n)=\sum\mathrm{sym}(a_{\vec{\gamma}}x^{\gamma})$
② 确定其对应的初等对称多项式：

$$x_1^{k_1}\cdots x_r^{k_r}\longmapsto\sigma_1^{(k_1-k_2)}\sigma_2^{(k_2-k_3)}\cdots\sigma_{r-1}^{(k_{r-1}-k_r)}\sigma_r^{(k_r-0)}$$

③ 待定系数法求解系数：设 f 为上面初等式的线性组合 (其中第一项系数为 1，这是由首 1 决定的)，取特殊元求解方程组。

例如 homework 1.2 : $f(x_1,x_2,x_3)=(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$

① 支配项为：$x_1^4x_2^2$, $x_1^4x_2x_3$, $x_1^3x_2^3$, $x_1^3x_2^2x_3$, $x_1^2x_2^2x_3^2$ 共五项。
② 对应的初等多项式分别为：$\sigma_1^2\sigma_2^2$, $\sigma_1^3\sigma_2$, σ_2^3, $\sigma_1\sigma_2\sigma_3$, σ_3^2。
③ 设 $f=\sigma_1^2\sigma_2^2+a\sigma_1^3\sigma_2+b\sigma_2^3+c\sigma_1\sigma_2\sigma_3+d\sigma_3^2$,解四阶矩阵得到 $f(x)=\sigma_1^2\sigma_2^2-4\sigma_1^3\sigma_2-4\sigma_2^3+18\sigma_1\sigma_2\sigma_3-27\sigma_3^2$

线性空间定义：

设 $\mathbb{F}$ 为一个域，若集合 V 具有加法和（$\mathbb{F}$ 上的）数乘两种运算，且 V 关于加法构成交换群，关于数乘满足封闭性、结合律、分配律、有数乘幺 $1_{\mathbb{F}}$，则称集合 V 构成 $\mathbb{F}$ 上的**线性空间**。将上述两种运算统称为**线性运算**，线性空间 V 中的元素称为**向量**。

> **“线性空间”与“向量空间”的区别：**
> “线性空间”与“向量空间”有时被看做是同义词，但是也有时对线性空间取广义的含义，对向量空间取狭义的含义（即向量空间 V 是 $\mathbb{R}^n$ 的子空间）。在徐晓平讲义中采用的是前者 (认为两者等同)，在本笔记中采用的是后者 (认为两者不同)。

例如：域 $\mathbb{F}$ 上的全体矩阵 $M_{m\times n}(\mathbb{R})$ 关于矩阵加法、$\mathbb{F}$ 上的数乘（纯量乘积）构成一个线性空间；多项式环 $\mathbb{F}[x]$ 关于多项式加法、$\mathbb{F}$ 上的数乘（纯量乘积）构成一个线性空间。实数域 $\mathbb{R}$ 是 $\mathbb{Q}$ 上的向量空间，复数域 $\mathbb{C}$ 是 $\mathbb{R}$ 上的向量空间。

线性空间基本性质：

$$
\begin{aligned}
&① \ \exists\, 0_{\mathbb{F}} \in \mathbb{F},\ \forall\, u \in V,\ a \in \mathbb{F},\ \text{有} 0\cdot u = 0 = 0\cdot a \\
&② \ a\cdot u = 0_{\mathbb{F}} \Longrightarrow a = 0_{\mathbb{F}} \text{ or } u = 0_V \\
&③ \ \forall\, n \in \mathbb{N},\ n\cdot u = u + u + \cdots u\text{（n 个 }u\text{）} \\
&④ \ (-1)\cdot u = -u
\end{aligned}
$$

线性空间的子空间：

设 U 为 $\mathbb{F}$ 上的线性空间 V 的非空子集，若 U 关于加法和数乘封闭，则称 U 为 V 的一个子空间。

一些常见的线性空间：

1. 矩阵线性空间：域 $\mathbb{F}$ 上的全体矩阵 $M_{m\times n}(\mathbb{R})$ 关于矩阵加法、$\mathbb{F}$ 上的数乘（纯量乘积）构成一个线性空间。
2. 多项式线性空间：多项式环 $\mathbb{F}[x]$ 关于多项式加法、$\mathbb{F}$ 上的数乘（纯量乘积）构成一个线性空间。
3. 映射（函数）线性空间：设 X 是一个非空集合，记 $\mathcal{F}_{\mathbb{F}}(X)$ 为从 X 到 $\mathbb{F}$ 的全体映射（函数），定义 $(af+bg)(x)=af(x)+bg(x)$, $a,b\in\mathbb{F}$, $f,g\in\mathcal{F}_{\mathbb{F}}(X)$，则 $\mathcal{F}_{\mathbb{F}}(X)$ 构成 $\mathbb{F}$ 上的线性空间。

4. 自由线性空间：定义**支撑集** supp $f=\{x\in X|f\in\mathcal{F}_{\mathbb{F}}(X),\ f(x)\neq 0\}$，定义集合 $V_{\mathbb{F}}(X)=\{f\in\mathcal{F}_{\mathbb{F}}(X)\mid |\text{supp } f|<\infty\}$（只在 X 的有限个元素处取非零值的全体函数），则 $V_{\mathbb{F}}(X)$ 构成 $\mathcal{F}_{\mathbb{F}}(X)$ 的一个子空间，我们称 $V_{\mathbb{F}}(X)$ 为由 X 生成的 $\mathbb{F}$ 上的自由线性空间。

5. 数域线性空间：设 $\mathbb{F}$ 是特征为 0 的域，则从 $\mathbb{Z}$ 到 $\mathbb{F}$ 的映射 $m\longmapsto m1_{\mathbb{F}}$ 构成环单同态，从而映射 $\frac{m}{n}\longmapsto(m1_{\mathbb{F}})(n1_{\mathbb{F}})^{-1}$ 构成从 $\mathbb{Q}$ 到 $\mathbb{F}$ 的域单同态，再定义 $\frac{m}{n}\cdot a=[(m1_{\mathbb{F}})(n1_{\mathbb{F}})^{-1}]\cdot a,\ \frac{m}{n}\in\mathbb{Q},\ a\in\mathbb{F}$，则域 $\mathbb{F}$ 关于域的加法和上述数乘构成 $\mathbb{Q}$ 上的线性空间。例如 $\mathbb{R}$ 是 $\mathbb{Q}$ 上的向量空间，$\mathbb{C}$ 是 $\mathbb{R}$ 上的向量空间。特别地，当 $\mathbb{F}$ 的数乘建立在域 $\mathbb{F}_p$ 上时（$\mathbb{F}_p$ 也即 $\mathbb{Z}_p$，它既构成环，也构成域），定义 $\bar{r}\cdot a=(r1_{\mathbb{F}})\cdot a,\ \bar{r}\in\mathbb{F}_p$，则 $\mathbb{F}$ 关于域的加法和上述数乘构成一个 $\mathbb{F}_p$ 上的线性空间。

1.2 基与维数

线性相关/无关：

设 V 是 $\mathbb{F}$ 上的一个线性空间，对于 V 的有限个元素 $\{v_1,v_2,...,v_k\}$，若存在不全为零的纯量 $\alpha_1,\alpha_2,...,\alpha_k\in\mathbb{F}$ 使得：

$$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_kv_k=0_V\in V$$

则称 $\{v_1,v_2,...,v_k\}$ 线性相关，若这样的纯量 α_i 不存在，则称其线性无关。

特别地，对于含有无限个元素的 V 的子集 $S=\{v_1,v_2,...,v_k,...\}$，若 S 中任意有限个向量都是线性无关的，则称 S 线性无关，否则称其线性相关。

关于 n 维实坐标空间 $\mathbb{R}^n$ 的、不涉及无限的结论，对线性空间也依然成立。例如：含零向量的向量组始终线性相关；向量组线性相关的充要条件为其中一个向量是其他向量的线性组合；一个线性无关向量组的任意部分组也线性无关。

基：

线性空间 V 的一个线性无关子集 S 称为它的基如果 $V=\text{Span } S$

例如：$\{1,x,x^2,...,x^n,...\}$ 是 $\mathbb{F}[x]$ 的一组基；记 E_{ij} 为 $M_{m\times n}(\mathbb{F})$ 中 (i,j) 位置元素为 1 而其他位置元素为 0 的矩阵，则 $\{E_{ij}\mid i=1,2,...,m;\ j=1,2,...,n\}$ 是 $M_{m\times n}$ 的一组基；在自由线性空间 $V_{\mathbb{F}}(X)$ 中定义映射 $\delta_y:\delta_y(x)=\begin{cases}1 & \text{if } x=y\\ 0 & \text{if } x\neq y\end{cases}$，则 $\{\delta_y\mid y\in X\}$ 是 $V_{\mathbb{F}}(X)$ 的一组基。

等价向量组：

两个向量组 $\{v_1,v_2,...,v_k\}$，$\{u_1,u_2,...,u_k\}$ 称为等价的如果：

$$\text{Span}\{v_1,v_2,...,v_k\}=\text{Span}\{u_1,u_2,...,u_k\}$$

等价的定义还有：u_i 可由 $\{v_1,v_2,...,v_k\}$ 线性表示，且 v_i 可由 $\{u_1,u_2,...,u_k\}$ 线性表示。

Theorem. 2 （Steinitz 替换定理）:
设 V 是 $\mathbb{F}$ 上的线性空间，若 $u_1, u_2, ..., u_m \in V$ 线性无关且可由 $v_1, v_2, ..., v_n$ 线性表示，则 $m \leq n$，且用 $u_1, u_2, ..., u_m$ 替换 $\{v_1, v_2, ..., v_n\}$ 中的任意 m 个向量得到的新向量组都等价于原向量组 $\{v_1, v_2, ..., v_n\}$。

由定理2可以知道，若两个分别线性无关的向量组 $u_1, u_2, ..., u_m$ 和向量组 $v_1, v_2, ..., v_n$ 等价，则 $m = n$。

维数:

若 $\mathbb{F}$ 上的线性空间 V 存在一组有限基 S，则 S 元素的个数称为 V 的维数，记作 $\dim_{\mathbb{F}} V = |S|$，简记为 $\dim V = |S|$。特别地，我们称零空间 $\{0\}$ 是零维的，称不存在有限基的线性空间是无限维的，记作 $\dim V = \infty$。

例如：$\dim_{\mathbb{Q}} \mathbb{R} = \infty$；$\{1, \sqrt{-1} = i\}$ 是 $\mathbb{R}$ 上的线性空间 $\mathbb{C}$ 的一组基，故 $\dim_{\mathbb{R}} \mathbb{C} = 2$

Theorem. 3 （基扩充定理）:
任意线性无关组 $\{v_1, v_2, ..., v_k\} (k < n)$ 可扩充为 n 维线性空间 V 的一组基 $\{v_1, v_2, ..., v_k, ..., v_n\}$。

转换矩阵:

由 V 的基 $\{u_1, u_2, ..., u_n\}$ 向另一组基 $\{v_1, v_2, ..., v_n\}$ 转换时，有:

$$\begin{cases} v_1 & = a_{11}u_1 + a_{12}u_2 + \cdots a_{1n}u_n \\ v_2 & = a_{21}u_1 + a_{22}u_2 + \cdots a_{2n}u_n \\ \vdots & \qquad\qquad \vdots \\ v_n & = a_{n1}u_1 + a_{n2}u_2 + \cdots a_{nn}u_n \end{cases}$$

记右侧的矩阵为 A，并称 A 为转换矩阵 (transfer matrix)。记 $\vec{u} = [u_1, u_2, ..., u_n]^T$，$\vec{v} = [v_1, v_2, ..., v_n]^T$，则有基转换公式:

$$\vec{v} = A\vec{u}, \text{ 也即 } \vec{u} \longmapsto A\vec{u}$$

设 $\vec{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_n]$ 是向量 $x \in V$ 在基 $\{u_1, u_2, ..., u_n\}$ 下的坐标，$\vec{\beta}$ 为新坐标，则有坐标转换公式:

$$\vec{\beta} = \vec{\alpha}A^{-1}, \text{ 也即 } \vec{\alpha} \longmapsto \vec{\alpha}A^{-1}$$

注意：基向量为列向量，坐标向量为行向量。坐标 [1,3,7] 与常规表示 (1,3,7) 同构。

容易证明，基的转换矩阵都是可逆矩阵。$\vec{\beta} = \vec{\alpha}A^{-1}$ 也可以写作 $A^T\vec{\beta}^T = \vec{\alpha}^T$，求 $\vec{\beta}$ 等价于解方程 $A^T\vec{x} = \vec{\alpha}^T$（推荐用高斯消元法），解得的 $\vec{x}$ 即为 $\vec{\beta}^T$。

Theorem. 4 （同维线性空间）:
$\mathbb{F}$ 上维数相同的两个线性空间必同构（都同构于 $\mathbb{F}^n$，n 为线性空间的维数）。

Theorem. 5 (子空间维数关系):
子空间的交 $U\cap V$、和 $U+V$ 都构成新子空间，但子空间的并 $U\cup V$ 不一定。且有：

$$\dim V+\dim U=\dim(U+V)+\dim(U\cap V)$$

(内外) 直和:

内直和：设 U_1，U_2 是 $\mathbb{F}$ 上线性空间 V 的两个子空间，记 $U=U_1+U_2$，若 $\forall\, u\in U$，$\exists!\, u_1\in U_1$，$u_2\in U_2$ 使 $u=u_1+u_2$，则称 U 是 U_1 和 U_2 的内直和，记作 $U=U_1\oplus U_2$。

外直和：设 U_1, U_2 是 $\mathbb{F}$ 上的任意两个线性空间，记 $U=U_1\times U_2=\{(u_1,u_2)\mid u_1\in U_1,\ u_2\in U_2\}$，并定义数乘 $a\cdot(u_1,u_2)=(a\cdot u_1, a\cdot u_2)$，则 U 构成一个线性空间，称为 U_1, U_2 的外直和，同样记作 $U=U_1\oplus U_2$。

从同构意义下，内外直和没有差别，统称为直和。

类似地，我们可以定义多个子空间的直和 $V=U_1\oplus U_2\oplus\cdots\oplus U_k$，其等价定义见定理7。

Theorem. 6 (二维直和等价定义):
设 U_1，U_2 是 $\mathbb{F}$ 上线性空间 V 的两个子空间，则下面的几个命题等价：

① U_1+U_2 是直和

② 若 $\exists\, u_1\in U_1$, $u_2\in U_2$ 使 $u_1+u_2=0_V$，则 $u_1=u_2=0_V$

③ $U_1\cap U_2=0_V$

④ $\dim(U_1+U_2)=\dim U_1+\dim U_2$

Theorem. 7 (多维直和的等价定义):
设 $U_1, U_2, ..., U_k$ 是 $\mathbb{F}$ 上线性空间 V 的 k 个子空间，则下面的几个命题等价：

① $U_1+U_2+\cdots+U_k$ 是直和

② 若 $\exists\, u_1\in U_1$, $u_2\in U_2, ..., u_k\in U_k$ 使 $u_1+u_2+\cdots u_k=0_V$，则 $u_1=u_2=\cdots=u_k=0_V$

③ $U_s\cap(\sum\limits_{i\neq s}U_i)=0_V$

④ $\dim(U_1+U_2+\cdots+U_k)=\dim U_1+\dim U_2+\cdots+\dim U_k$

一些常见的直和:

1. 矩阵空间的两种直和分解：在矩阵线性空间 $M_{n\times n}(\mathbb{R})$ 中，记 $\mathcal{S}_{n\times n}(\mathbb{R})$ 为全体对称矩阵（简记为 $\mathcal{S}$），记 $\mathcal{O}_{n\times n}(\mathbb{R})$ 为全体斜对称矩阵（简记为 $\mathcal{O}$），记 $\mathcal{T}_{n\times n}(\mathbb{R})$ 为全体上三角矩阵（简记为 $\mathcal{T}$），则有：
$$M_{n\times n}(\mathbb{R})=\mathcal{S}\oplus\mathcal{O},\ \ M_{n\times n}(\mathbb{R})=\mathcal{T}\oplus\mathcal{O}$$

2. 半幻方矩阵空间分解：设 $\mathbb{Q}$ 上的全体半幻方矩阵 (行、列和相同) 为：
$$\mathrm{Smag}_n(\mathbb{Q})=\left\{A=(a_{ij})_{n\times n}\in M_{n\times n}(\mathbb{Q})\mid \sum_{i=1}^{n}a_{ir}=\sum_{i=1}^{n}a_{ri}=\sigma(A),\ r=1,...,n\right\}$$

相应的全体幻方矩阵 (行、列、主副对角线和相同) 为：

$$\mathrm{Mag}_n(\mathbb{Q}) = \left\{ A \in \mathrm{Smag}_n(\mathbb{Q}) \mid \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} a_{i(n+1-i)} = \sigma(A),\ r = 1, ..., n \right\}$$

并记矩阵：

$$I_n = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}_{n\times n}, \qquad D_n = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}_{n\times n}$$

则有结论：

$$\mathrm{Smag}_n(\mathbb{Q}) = \mathrm{Mag}_n(\mathbb{Q}) \oplus \mathbb{Q}I_n \oplus \mathbb{Q}D_n$$

矩阵空间的两种直和分解证明：

对于前者：考虑 $M_{n\times n}(\mathbb{R})$ 的标准基 $\{E_{ij} \mid i = 1, 2, ..., n;\ j = 1, 2, ..., n\}$，注意到

$$E_{ij} = \frac{1}{2}(E_{ij} + E_{ji}) + \frac{1}{2}(E_{ij} - E_{ji}) \in (\mathcal{S} + \mathcal{O})$$

因此

$$\mathrm{Span}\,\{E_{ij}\} = M_{n\times n}(\mathbb{R}) \subseteq \mathcal{S} + \mathcal{O} \Longrightarrow M_{n\times n}(\mathbb{R}) = \mathcal{S} + \mathcal{O}$$

另一方面，设 $A \in \mathcal{S} \cap \mathcal{O}$，则 $A^T = A = -A \Longrightarrow A = 0_{n\times n}$，故 $\mathcal{S} + \mathcal{O}$ 构成直和，也即 $M_{n\times n}(\mathbb{R}) = \mathcal{S} \oplus \mathcal{O}$。（也可从 $\dim \mathcal{S} = \dim \mathcal{T} = \frac{n(n+1)}{2}$，$\dim \mathcal{O} = \frac{(n-1)n}{2}$ 的角度说明构成直和）。

对于后者，设 $1 \le i \le j \le n$，则 $E_{ij} \in \mathcal{T}$，且有：

$$E_{ji} = E_{ij} + (E_{ji} - E_{ij}) \in (\mathcal{T} + \mathcal{O})$$

同理得 $M_{n\times n}(\mathbb{R}) = \mathcal{T} + \mathcal{O}$，进而得 $M_{n\times n}(\mathbb{R}) = \mathcal{S} \oplus \mathcal{O}$。

证必。

余维数：

设线性空间 $V = U \oplus \overline{U}$，则称 $\overline{U}$ 为 U 在 V 中的补空间，称 $\overline{U}$ 的维数为 U 在 V 中的余维数，且有：

$$\mathrm{codim}\, U = \dim \overline{U} = \dim V - \dim U$$

特别地，我们规定，$\{0\}$ 是 V 在 V 中的补空间。

无限维线性空间可以存在有限维子空间，因此子空间的余维数可以是有限的。例如：考虑无限维线性空间 $V = \mathbb{F}[x]$，我们有 $V = \mathcal{P}_n[x] \oplus \mathrm{Span}\{x^i \mid n \le i,\ i \in \mathbb{N}\}$，且 $\dim \mathrm{Span}\{x^i \mid n \le i,\ i \in \mathbb{N}\} = \infty$，$\mathrm{codim}\ \mathrm{Span}\{x^i \mid n \le i,\ i \in \mathbb{N}\} = n$。

超平面：

一个线性空间 V 中，余维数为 1 的子空间称为 V 中的超平面。

Theorem. 8（补空间存在定理）：
设 U 是线性空间 V 中的一个子空间，则 U 在 V 中一定存在补空间 $\overline{U}$。上述结论对有限维、无限维线性空间都成立，证明参考：

https://www.zhihu.com/question/68641016/answer/265785313

商空间：

设 W 是 $\mathbb{F}$ 上线性空间 V 的子空间，定义 V 模 W 的商空间为：

$$V/W = \{\overline{v} \mid v \in V\} = \{v + W \mid v \in V\}$$

定义商空间中的线性运算：$a\overline{v_1} + b\overline{v_2} = \overline{av_1 + bv_2}$, $a, b \in \mathbb{F}$, $v_1, v_2 \in V$，则 V/W 构成一个 $\mathbb{F}$ 上的线性空间。

考虑到 V 可以看作一个加法群，则子空间 W 构成 V 的子群，又 V 交换，因此 $W \trianglelefteq V$，从这个角度来看，商空间本质上还是群模正规子群得到的代数结构。

构造由补空间 $\overline{W}$ 到商空间 V/W 的映射 $\varphi : \varphi(u) = \overline{u}$, $u \in \overline{W}$，则 φ 构成一个同态 (保持线性运算)，且为双射，故 $\overline{W} \cong V/W$，于是 $\dim V/W = \dim \overline{W} = \operatorname{codim} W$。

要证明 φ 是满射，我们首先需要说明一个陪集 $\overline{v} = v + W$ 的代表元不唯一，例如 $\forall w \in W$，显然 $\overline{v} = \overline{v + w}$，这表明我们可以向陪集的代表元"添加" W 中的任意元素。又 V 中的元素可以做直和分解，也即 $\forall\, v \in V, \exists\, w' \in \overline{W}, w \in W$ 使 $v = w' + w,$，于是对于任意的陪集 $\overline{v}$，设 $v = w' + w$，则 $\exists\, w' \in \overline{W}$ 使 $\varphi(w') = \overline{w'} = \overline{w' + w} = \overline{v}$，因此 φ 构成满射。又容易验证它是单射，所以 φ 构成双射。

于是，若 $\{v_1, ..., v_m\}$ 是 $\overline{W}$ 的一组基，则 $\{\overline{v_1}, ..., \overline{v_2}\}$ 构成 V/W 的一组基。

1.3 对偶空间

对偶空间：

记 V^* 为从 V 到 $\mathbb{F}$ 的全体线性映射（函数），则 V^* 构成 $\mathcal{F}_{\mathbb{F}}(V)$ 的子空间，称为 V 的对偶空间，且有 $\dim V = \dim V^*$。

记 V 的一个基向量为 $\vec{v}_0 = [v_1, v_2, ..., v_n]^T$，系数向量 (坐标向量) 为 $\vec{a} = [a_1, a_2, ..., a_n]$，则：对任意的 $f \in V^*$, $v = a_1v_1 + \cdots + a_nv_n = \vec{a} \cdot \vec{v}_0 \in V$，我们有 $f(v) = f(\vec{a} \cdot \vec{v}_0) = \vec{a}f(\vec{v}_0)$，类似地，假设 A 是由基 $\vec{v}_0$ 向基 $\vec{u}_0$ 的转换矩阵，也即 $\vec{u}_0 = A\vec{v}_0$ 我们有 $f(\vec{u}_0) = f(A\vec{v}_0) = Af(\vec{v}_0)$。这就是为什么称向量 $f(\vec{v}_0)$ 是同变的。

特别地，我们指出，在无限维线性空间中，$V^* \neq \operatorname{Span}\{v^1, v^2, ..., v^n\}$。事实上，在无限维的情况下 $\operatorname{Span}\{v^1, v^2, ..., v^n\} \subsetneqq V^*$，这是因为 Span 只能是有限和，如果定义 V 到 $\mathbb{F}$ 的映射 $f : f(v_i) = 1, \forall\, i \in \mathbb{N}$，则 $f \in V^*$ 且 $f \notin \operatorname{Span}\{v^1, v^2, ..., v^n\}$。

Theorem. 9 **(对偶空间重要结论):**
在线性空间及其对偶空间中，我们有一个重要而有用的结论：

$$\{v \in V \mid \forall f \in V^*,\ f(v) = 0\} = \{0_V\}$$

对偶基及其转换:

设 $\{v_1, v_2, ..., v_n\}$ 是 V 的一组基，定义映射 $v^i : v = a_1v_1 + \cdots + a_nv_n \longmapsto a_i$, $i \in \{1, 2, ..., n\}$, $a_i \in \mathbb{F}$，也即 $v^i(v_j) = \delta_{ij}$，则 $\{v^1, v^2, ..., v^n\}$ 构成 V^* 的一组基，称为对偶基。特别地，$\forall f \in V*$，我们有：

$$f = f(v_1)v^1 + f(v_2)v^2 + \cdots f(v_n)v^n$$

也即映射 f 在基 $\{v^1, v^2, ..., v^n\}$ 下的坐标为 $\vec{x} = [f(v_1), ..., f(v_n)]$。

假设 $A \in M_{n\times n}(\mathbb{F})$ 是 V 的基 $\vec{v}_0$ 向基 $\vec{u}_0$ 的转换矩阵，也即 $\vec{u}_0 = A\vec{v}_0$，则 $(A^T)^{-1}$ 是 V^* 的基 $\vec{v}^0$ 向基 $\vec{u}^0$ 的转换矩阵：

$$\vec{u}^0 = (A^T)^{-1}\vec{v}^0,\ \text{也即}\ \vec{v}^0 \longmapsto (A^T)^{-1}\vec{v}^0$$

对偶基转换式的证明:
基 $\vec{v}_0 = [v_1, ..., v_n]$ 对应的对偶基为 $\vec{v}^0 = [v^1, ..., v^n]$，且 $\forall\, f = b_1v^1 + \cdots + b_nv^n \in V^*$，有：

$$f(v_i) = (b_1v^1 + \cdots + b_nv^n)(v_i) = b_i \Longrightarrow f = f(v_1)v^1 + \cdots + f(v_n)v^n$$

类似地，设基 $\vec{u}_0 = [u_1, ..., u_n]$ 的对偶基为 $\vec{u}^0 = [u^1, ..., u^n]$，则有 $f = f(u_1)u^1 + \cdots + f(u_n)u^n$，于是：

$$\begin{aligned}
v^i(u_r) &= v^i(a_{r1}v_1 + \cdots + a_{rn}v_n) = a_{ri} \\
\Longrightarrow v^i &= v^i(u_1)u^1 + \cdots + v^i(u_n)u^n \\
&= a_{1i}u^1 + \cdots + a_{ni}u^n \\
\Longrightarrow \vec{v}^0 &= A^T\vec{u}^0 \\
\Longrightarrow \vec{u}^0 &= (A^T)^{-1}\vec{v}^0
\end{aligned}$$

证毕。

Theorem. 10 **(线性空间向量组的秩):**
设 $\{f_1, f_2, ..., f_n\}$ 是对偶空间 V^* 的一组基，则对 V 中的任意向量组 $\{u_1, u_2, ..., u_k\}$，有：

$$\text{rank}\{u_1, u_2, ..., u_k\} = \text{rank} \begin{bmatrix} f_1(u_1) & f_1(u_2) & \cdots & f_1(u_k) \\ f_2(u_1) & f_2(u_2) & \cdots & f_2(u_k) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(u_1) & f_n(u_2) & \cdots & f_n(u_k) \end{bmatrix} = \text{rank}\ (f_i(u_j))_{n\times k}$$

定理10的证明：

记列向量

$$\vec{a}^i = \begin{bmatrix} f_1(u_i) \\ f_2(u_i) \\ \vdots \\ f_n(u_i) \end{bmatrix}, \ i = 1, 2, ..., k$$

我们只需证明对任意的 $1 \le i_1 < i_2 < \cdots < i_r \le k$，向量组 $\{u_{i_1}, u_{i_2}, ..., u_{i_r}\}$ 线性相关 $\Longleftrightarrow$ 向量组 $\{\vec{a}^{i_1}, \vec{a}^{i_2}, ..., \vec{a}^{i_k}\}$ 线性相关。

1. $\Longrightarrow$:

假设 $\{u_{i_1}, u_{i_2}, ..., u_{i_r}\}$ 线性相关，即 $\exists\, a_1, a_2, ..., a_r \in \mathbb{F}$ 使得：

$$a_1 u_{i_1} + a_2 u_{i_2} + \cdots + a_r u_{i_r} = 0_V$$

依次吧映射 $f_1, f_2, ..., f_n$ 作用在方程两边，得到由 n 个方程构成的方程组：

$$a_1 f_i(u_{i_1}) + a_2 f_i(u_{i_2}) + \cdots + a_r f_i(u_{i_r}) = f_i(0_V) \xlongequal{\text{定理9}} 0$$

这等价于：

$$a_1 \vec{a}^{i_1} + a_2 \vec{a}^{i_2} + \cdots + a_r \vec{a}^{i_r} = \vec{0}$$

因此 $\{\vec{a}^{i_1}, \vec{a}^{i_2}, ..., \vec{a}^{i_k}\}$ 线性相关。

2. $\Longleftarrow$:

将上述过程逆过来即可，略。

有限维线性空间的自反性：

考虑有限维对偶空间的对偶空间 $(V^*)^* = V^{**}$，对任意 $u \in V$，定义从 V^* 到 $\mathbb{F}$ 的映射 $\varepsilon_u \in V^{**}$ 为：$\varepsilon_u(f) = f(u)$, $f \in V^*$，则映射 $u \longmapsto \varepsilon_u$ 给出了 V 到 V^{**} 的同构，也即 $V \cong V^{**}$。要注意，上述结论在无限维不成立，无限维线性空间不具有自反性。在讲义中写作“$V = V^{**}$”，并且称 V 是自反的，其详细含义是什么？自反的定义不应该是 $V \sim V$ 吗？

Theorem. 11 (线性空间的子空间):
设 $\{f_1, f_2, ..., f_k\}$ 是 n 维对偶空间 V^* 中的一个秩为 r 的向量组，则 $\{v \in V \mid f_1(v) = f_2(v) = \cdots = f_k(v) = 0\}$ 构成 V 的 $(n-r)$ 维子空间。定理的几何意义是 V 中 k 个超平面的交。

1.4 双线性型和二次型

双线性型：

V 上的一个双线性型（双线性函数）是一个映射 $f: V\times V\longmapsto \mathbb{F}$ 满足：

$$
\begin{aligned}
&\forall\, v_1,v_2,u_1,u_2\in V,\ a_1,a_2,b_1,b_2\in\mathbb{F},\ \text{有：}\\
&f(a_1v_1+a_2v_2,u)=f(a_1v_1,u)+f(a_2v_2,u)=a_1f(v_1,u)+a_2f(v_2,u)\\
&f(v,b_1u_1+b_2u_2)=f(v,b_1u_1)+f(v,b_2u_2)=b_1f(v,u_1)+b_2f(v,u_2)
\end{aligned}
$$

也即双线性型 f 是一个二重线性映射。特别地，我们指出，双线性型原空间可以是两个不同集合的内积，也即双线性型 $f: X\times Y\longmapsto\mathbb{F}$，后文的相关概念也可据此作出适当延伸。

双线性型线性空间：

记 $\mathcal{L}_2(V,\mathbb{F})$ 为 V 上的全体双线性型，容易验证其关于映射的加法、数乘构成一个线性空间（$\mathcal{F}_{\mathbb{F}(V\times V)}$ 的子空间）。

定义双线性型 f 在 V 的基 $\vec{v}_0=[v_1,v_2,...,v_n]^T$ 下的度量矩阵 (表示矩阵)：

$$
F=(f(v_i,v_j))_{n\times n}=(f_{ij})_{n\times n}=\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n}\\ f_{21} & f_{22} & \cdots & f_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ f_{n1} & f_{n2} & \cdots & f_{nn}\end{bmatrix}=\begin{bmatrix} f(v_1,v_1) & f(v_1,v_2) & \cdots & f(v_1,v_n)\\ f(v_2,v_1) & f(v_2,v_2) & \cdots & f(v_2,v_n)\\ \vdots & \vdots & \ddots & \vdots\\ f(v_n,v_1) & f(v_n,v_2) & \cdots & f(v_n,v_n)\end{bmatrix}
$$

其中 f_{ij} 定义为 $f_{ij}=f(v_i,v_j)$。

并设 $\vec{x},\vec{y}$ 分别是 u,v 在基 $\vec{v}_0$ 下的坐标，则有：

$$f(u,v)=\vec{x}F\vec{y}^T$$

且映射 $f\longmapsto F$ 给出了 $\mathcal{L}_2(V,\mathbb{F})$ 到 $M_{n\times n}(\mathbb{F})$ 的同构。

度量矩阵转换：

设 $A=(a_{ij})_{n\times n}$ 为基 $\vec{v}_0$ 向 $\vec{u}_0$ 的转换矩阵，也即：

$$u_i=a_{i1}v_1+a_{i2}v_2+\cdots a_{in}v_n,\ i=1,2,...,n$$

设 F 为 f 在基 $\vec{v}_0$ 下的度量矩阵，则 f 在 $\vec{u}_0$ 下的新度量矩阵为：

$$F'=AFA^T,\ \text{也即}\ F\longmapsto AFA^T$$

考虑到 A 为可逆矩阵，有 $\operatorname{rank} F'=\operatorname{rank} F$，故度量矩阵 F 的秩不随基而变化，称为 f 的秩。

合同矩阵：

两个 $n \times n$ 矩阵 G, F 称为合同的如果存在可逆矩阵 A 使得 AFA^T。

由此可得，一个双线性型在不同基下的度量矩阵是合同的。

左 (右) 根空间：

给定 $f \in \mathcal{L}_2(V,\mathbb{F})$，定义 f 的左根为 $L_f = \{u \in V \mid \forall\, v \in V,\ f(u,v) = 0\}$，右根为 $R_f = \{v \in V \mid \forall\, u \in V,\ f(u,v) = 0\}$，且有：

$$\dim L_f = \dim R_f = \dim \mathcal{L}_2(V,\mathbb{F}) - \operatorname{rank} f$$

双线性型 f 称为非退化的 (nondegenerate) 如果 $L_f = \{0\} \iff R_f = \{0\} \iff f$满秩，这里的 0 指零映射。

一般 $L_f \neq R_f$，且两者都构成 $\mathcal{L}_2(V,\mathbb{F})$ 的子空间。

(斜) 对称双线性型：

双线性型 $f \in \mathcal{L}_2(V,\mathbb{F})$ 称为对称的如果 $\forall\, u,v \in V,\ f(u,v) = f(v,u)$，称为斜对称的如果 $\forall\, u,v \in V,\ f(u,v) = -f(v,u)$。

记 $\mathcal{L}_2^+(V,\mathbb{F})$ 为全体对称双线性型，记 $\mathcal{L}_2^-(V,\mathbb{F})$ 为全体斜对称双线性型，则有：

$$\mathcal{L}_2(V,\mathbb{F}) = \mathcal{L}_2^+(V,\mathbb{F}) \oplus \mathcal{L}_2^-(V,\mathbb{F})$$

特别地，我们有：双线性型 f 为对称的 $\iff$ f 的度量矩阵 F 是对称矩阵

对称双线性型，或者说实对称矩阵的性质很特殊：

1. 特征值为实数：实对称矩阵的特征多项式在复数域中的每一个根都是实数，因此其特征值都是实数。
2. 特征向量为实向量：实对称矩阵的特征值对应的特征向量都是实向量。
3. 可相似对角化：n 阶实对称矩阵必可对角化，而对角矩阵上的对角元素即为原矩阵的特征值。
4. 秩与非零特征值个数相等：实对称矩阵的秩等于其非零特征值的个数。

二次型：

假设 $\mathbb{F}$ 的特征不是 2(推出 2 在 $\mathbb{F}$ 中可逆)，V 是 $\mathbb{F}$ 上的线性空间，一个映射 $q \in \mathcal{F}_{\mathbb{F}}(V)$ 称为二次型如果 $\forall\, v \in V,\ q(v) = q(-v)$。

定义 q 对应的对称双线性型 f_q，并将 f_q 的秩称为 q 的秩：

$$f_q(u,v) = 2^{-1}(q(u+v) - q(u) - q(v)),\ u,v \in V$$

容易验证 $f_q \in \mathcal{L}_2^+(V,\mathbb{F})$。考虑到我们的 $\mathbb{F}$ 不一定是实数域或复数域，f_q 定义式中的 2 其实是一个抽象含义，例如：令 $\mathbb{F}$ 为模 p 剩余类 $\mathbb{F}_p$，并定义映射 q 从 $\mathbb{F}_p[x]$ 到 $\mathbb{C}$，定义式中的 2 则对应 $\mathbb{F}_p$ 中的 $\overline{2}$，2^{-1} 则对应 $\mathbb{F}_p$ 中的 $\overline{\frac{p+1}{2}}$，比如令 $p = 5$，则 2 对应 $\overline{2}$，2^{-1} 对应 $\overline{3}$。

类似地，给定对称双线性型 $f \in \mathcal{L}_2^+(V, \mathbb{F})$，定义其对应的二次型：

$$q_f(v) = f(v, v),\ u \in V$$

而 $f_q(0,0) = 0$(因为 f 是线性的)，由此我们得到 $q(0) = f_q(0,0) = 0$，且 $f = f_{q_f}$, $q = q_{f_q}$，两者一一对应。我们只在对称双线性型空间中同时讨论 f_q 和 q 及其对应关系。

称对称双线性型 f 在基 $\vec{v}_0$ 下的度量矩阵 F 为 q_f 在基 $\vec{v}_0$ 下的度量矩阵，可以简记为 $F_{f_q} = F_q$。

特别地，对所有坐标空间 (即 $\mathbb{F}^n$) 中的二次型 q，设 q 在基 $\vec{v}_0$ 下的度量矩阵为 F，v 的坐标为 $\vec{x}$，我们有：

$$q(v) = \vec{x}F\vec{x}^T$$

推论：在坐标空间中，二次型始终是二次齐次多元多项式。

迷向空间：

对任意的 q，由于 $f_q(u, v)$ 是对称的，因此 $L_{f_q} = R_{f_q}$，即左右根空间相等，统称为 q 的一个迷向空间 (isotropic space)。并且事实上：

$$L_{f_q} = R_{f_q} = \{v \in V \mid q(u+v) = q(u) + q(v)\}$$

规范型：

V 的一组基 $\vec{v}_0 = [v_1, ...v_n]^T$ 称为二次型 q 的规范基如果对任意 $v = x_1v_1 + \cdots + x_nv_n \in V$，有：

$$q(v) = \vec{x}F\vec{x}^T = f_{11}x_1^2 + f_{22}x_2^2 + \cdots + f_{nn}x_n^2$$

我们称上式为 q 在基 $\vec{v}_0$ 下的规范型，对应生成的 f_q 也称为规范型。且此时，q 的度量矩阵 (也是 f_q 的) 为：

$$F_q = \begin{bmatrix} f(v_1, v_1) & 0 & \cdots & 0 \\ 0 & f(v_2, v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(v_n, v_n) \end{bmatrix}$$

Theorem. 12（对称二次型必有规范基）:
有限维线性空间 V 上的每个二次型 q，都存在规范基。

推论：任意对称矩阵都合同于一个对角阵，也即 $\forall\ A \in \left\{M_{n\times n} \mid A = A^T\right\}, \exists\ D = \text{diag}(a_{11}, a_{22}, ..., a_{nn}),\ B \in M_{n\times n}$ 使得：

$$A = BDB^T$$

标准型：

对于 q 在基 $\vec{v}_0$ 下规范型 $q(v)=\vec{x}F\vec{x}^T=f(v_1,v_1)x_1^2+\cdots+f(v_n,v_n)x_n^2$，由于 $f(v_i,v_i)$ 的取值在 $\mathbb{R}$ 中不定，我们调换 $\{v_1,...,v_n\}$ 中元素的顺序，使得：

$$\begin{cases} f_{ii}>0 & i\in[1,r] \\ f_{ii}<0 & i\in[r+1,r+s] \\ f_{ii}=0 & i\in[r+s+1,n] \end{cases}$$

我们进行基变换：

$$\begin{cases} u_i=\sqrt{f_{ii}}v_i & i\in[1,r] \\ u_i=\sqrt{-f_{ii}}v_i & i\in[r+1,r+s] \\ u_i=v_i & i\in[r+s+1,n] \end{cases}$$

得到基 $\vec{u}_0$，相应地，坐标也由 $\vec{x}$ 变为 $\vec{y}$，则 q 在基 $\vec{u}_0$ 下的式子化为：

$$q(v)=y_1^2+\cdots+y_r^2-y_{r+1}^2-\cdots-y_{r+s}^2,\ \forall\, v=y_1u_1+\cdots+y_nu_n\in V$$

我们称上式为 q 的标准型，也即：

$$\begin{cases} f(u_i,u_i)=1 & i\in[1,r] \\ f(u_i,u_i)=-1 & i\in[r+1,r+s] \\ f(u_i,u_i)=0 & i\in[r+s+1,n] \end{cases}$$

标准基常指的是 $\{\vec{e}_1,\ldots,\vec{e}_n\}$，我们也可以把上面标准型对应的基称为标准基，但要注意辨别，不要引起歧义。

Theorem. 13 （r 与 s 取决于二次型 q）：
对于任意的二次型 q。其标准型中的整数 r 与 s 仅由 q，与相应的规范基无关。

快速确定二次型/双线性型在默认基下的矩阵：

步骤：(二次型 $q\xrightarrow{①}$) 双线性型 $f\xrightarrow{②}$ 度量矩阵 F

① q 转 f：设二次型为 q，对每一项分别作左 y 变换和右 y 变换并相加，最后勿忘乘 $\frac{1}{2}$，即可得到 q 对应的双线性型 f_q。

例如 $q(u)=2x_1x_2-6x_1x_3+2x_1x_3+x_3^2$ 经变换得到：

$$\begin{aligned} f(u,v)&=\frac{1}{2}\left[(2x_1y_2+2y_1x_2)+(-6x_1y_3-6y_1x_3)+(2x_2y_3+2y_2x_3)+(x_3y_3+y_3x_3)\right] \\ &=x_1y_2+y_1x_2-3x_1y_3-3y_1x_3+x_2y_3+y_2x_3+x_3y_3 \end{aligned}$$

② f 转 F：设双线性型 $f(u,v)$，其中 $u=\vec{x}\cdot\vec{e}$，$v=\vec{y}\cdot\vec{e}$，则每一项中 x 的角标表示行，y 的角标表示列，系数代表 F 此处 entry 的值。

继续上面的例子，对应的度量矩阵为 $F=\begin{bmatrix}0&1&-3\\1&0&1\\-3&1&1\end{bmatrix}$，即为在默认基下的矩阵 (一般认为是标准正交基)。

将二次型化为标准型的三种方法：

初等变换法、配方法、偏导数配方法、正交变换法。

https://www.zhihu.com/question/465317828/answer/1943470027

https://www.zhihu.com/question/67528139/answer/1416472234 (偏导数)

https://zhaokaifeng.com/16920/ (偏导数)

偏导数法将二次型化为标准型：

设 q (或双线性型 f 化为 q_f) 为二次型, 步骤如下：

① 平方项的配方：令 $f_1=\frac{\partial f}{\partial x_1}$，求解 $g=f-\frac{1}{4a_{11}}f_1^2$，则 g 中不再含有 x_1；再令 $g_2=\frac{\partial g}{\partial x_2}$，求解 $h=g-\frac{1}{4a_{22}}\cdot g_2^2=f-\frac{1}{4a_{11}}\cdot f_1^2-\frac{1}{4a_{22}}\cdot g_2^2$......(重复上面操作)

② 非平方项的配方：令 $f_1=\frac{\partial f}{\partial x_1}$, $f_2=\frac{\partial f}{\partial x_2}$，求解 $g=f-\frac{1}{4a_{12}}\left[(f_1+f_2)^2-(f_1-f_2)^2\right]$，则 g 中不再含有 x_1,x_2；再令 g_3,g_4，求解 $h=f-\frac{1}{4a_{34}}\left[(g_3+g_4)^2-(g_3-g_4)^2\right]$......(重复上面操作)

③ 最后将结果汇总，即可得到 $f=$ 平方项之和。

特别地，将对称双线性型化为标准型的方法和二次型是一致的 (将双线性转为二次型，标准化后再转回)。

惯性指数：

我们称 $r+s$ 为 V 上的二次型 q 的指数，并称 r 为正惯性指数，s 为负惯性指数。另外，设 q 是 $\mathbb{R}$ 上的有限维向量空间 V 的一个二次型，则：

$$① \ q \text{ 是正定的：} r=\dim V \iff q(v)>0,\ \forall\, 0_V\neq v\in V$$

$$② \ q \text{ 是负定的：} s=\dim V \iff q(v)<0,\ \forall\, 0_V\neq v\in V$$

$$③ \ q \text{ 是半正定的：} s=0 \iff q(v)\geq 0,\ \forall\, 0_V\neq v\in V$$

$$④ \ q \text{ 是半负定的：} r=0 \iff q(v)\leq 0,\ \forall\, 0_V\neq v\in V$$

$$⑤ \ q \text{ 是不定的：} r\neq 0,\ s\neq 0 \iff \exists\, v,u\in V,\ q(v)>0,\ q(u)<0$$

Theorem. 14 (矩阵正定的等价条件)：
设对称矩阵 $A\in M_{n\times n}(\mathbb{R})$，则：

$$\text{矩阵 } A \text{ 正定} \iff \text{存在可逆矩阵 } S\in M_{n\times n}(\mathbb{R}) \text{ s.t. } A=SS^T$$

推论：

$$\text{矩阵 } A \text{ 正定} \Longrightarrow A \text{ 的对角元都} > 0$$
$$\text{矩阵 } A \text{ 半正定} \Longrightarrow A \text{ 的对角元都} \geq 0$$
$$\text{矩阵 } A \text{ 是不定的} \Longrightarrow A \text{ 的对角元既有} > 0 \text{ 也有} < 0$$

设对称矩阵 $A \in \left\{A \in M_{n\times n}(\mathbb{R}) \mid A^T = A\right\}$，定义 A 所对应的对称双线性型为 $f_A(u,v) = uAv^T$ (这表明 A 即为 f_A 的度量矩阵)，从而也有对应的二次型 q_{f_A}，并将 q_{f_A} 的正定性称为对称矩阵 A 的正定性。

主子式：

设 F 是 V 上二次型 q 在基 $\vec{v}_0 = [v_1, ..., v_n]^T$ 下的度量矩阵，则 F 的主子式定义为：

$$\Delta_0 := 1,\ \ \Delta_1 = f_{11},\ \ \Delta_2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix},\ \ldots,\ \ \Delta_n = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{vmatrix} = |F|$$

Theorem. 15 （Jacobi Theorem）：
设二次型 q 的度量矩阵 F 的主子式全都不为 0，则存在一组基 $\{u_1, u_2, ..., u_n\}$ 使得 q 在此基下化成：

$$q(u) = \frac{\Delta_0}{\Delta_1}y_1^2 + \frac{\Delta_1}{\Delta_2}y_2^2 + \cdots + \frac{\Delta_{n-1}}{\Delta_n}y_n^2,\ \ \forall\, u = y_1u_1 + \cdots y_nu_n \in V$$

且上面对应的转换矩阵 T 是下三角的，也即存在下三角可逆矩阵 T 使得：

$$TFT^{-1} = \text{diag}(\frac{\Delta_0}{\Delta_1}, ..., \frac{\Delta_{n-1}}{\Delta_n})$$

推论：二次型的负惯性指数个数是序列 $\Delta_0, \Delta_1, \ldots, \Delta_n$ 的变号个数。特别地，我们有：

$$q \text{ 是正定的} \Longleftrightarrow \Delta_i > 0,\ \forall\, i \in \{0, 1, ..., n-1\}$$

$$q \text{ 是负定的} \Longrightarrow \Delta_i\Delta_{i+1} < 0,\ \forall\, i \in \{0, 1, ..., n-1\}$$

斜对称双线性型：

给定一个双线性型 f，对应的度量矩阵是 F，类似地，我们有：

$$f(u,v) = \vec{x}F\vec{y}^T \ \ ,\ \ \ f \text{ 是斜对称的} \Longleftrightarrow F \text{ 是斜对称的}$$

斜对称双线性型的秩一定是偶数。

Theorem. 16（斜对称双线性型必有规范基）:
设 f 是 V 上的斜对称双线性型，则存在基 $\{v_1,...,v_n\}$ 使得 f 在此基下化成：

$$f(u,v)=(x_1y_2-x_2y_1)+\cdots+(x_{2r-1}y_{2r}-x_{2r}y_{2r-1})=\sum_{k=1}^{r}(-1)^{k+1}(x_{2k-1}y_{2k}-x_{2k}y_{2k-1})$$
$$\forall\, u=x_1v_1+\cdots x_nv_n,\ \ v=y_1v_1+\cdots+y_nv_n$$

并称上式为 f 的标准型。

推论：对任意的斜对称矩阵 $A\in\left\{A\in M_{n\times n}(\mathbb{R})\mid A^T=-A\right\}$，存在可逆矩阵 S 使得：

$$SAS^T=\begin{bmatrix} O_{r\times r} & I & \\ -I & O_{r\times r} & \\ & & O_{(n-2r)\times(n-2r)} \end{bmatrix}$$

且：　A 可逆 $\Longleftrightarrow n=2r\Longrightarrow \det A=(\det S)^{-2}$

这是因为由上面的度量矩阵转换，可以得到：对任意的斜对称矩阵 $A\in\left\{A\in M_{n\times n}(\mathbb{R})\mid A^T=-A\right\}$，记 $H=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$，存在可逆矩阵 S 使得：

$$SAS^T=\begin{bmatrix} H & & & \\ & \ddots & & \\ & & H & \\ & & & O \end{bmatrix}_{n\times n}$$

对 SAS^T 进一步同时做的行变换和列变换，即可得到推论中的形式。特别地，当 A 可逆时，有 $n=2r$，且 $\det A=(\det S)^{-2}$。

将斜对称双线性型 f 化为规范型的方法：

方法一：利用定理16。设 $\{\vec{e}_1,...,\vec{e}_n\}$ 是原基，$\{v_1,...,v_n\}$ 是标准基 (规范基)，步骤如下：

① 找到 v_1,v_2 使得 $f(v_1,v_2)=1$；

② 诱导出 $W'=\{v\in V\mid \forall\, u=\alpha_1v_1+\alpha_2v_2\in W,\ f(u,v)=0\}$，下面找到 W' 的一组基；

③ 任取 $w\in V$，并构造 $v_3=w+f(v_2,w)v_1-f(v_1,w)v_2$，则 $v_3\in W'$；

④ 再取 $w\in V$，依次构造 $v_4,v_5,...$，并分析是否线性无关，直至得到 W' 的一组基。

⑤ 依次验证 $f(v_3,v_4)$，...，$f(v_{n-1},v_n)==1$，若非 1，添加系数得到 v_i'(如 v_3') 使其变为 1。

⑥合并即得标准基 (规范基) $\{v_1,v_2,v_3',v_4,...,v_n\}$

方法二：配凑法 (不好用)。https://zhuanlan.zhihu.com/p/99513090

方法三：矩阵合同初等变换 (在忘记方法一时使用)，注意 A 要做合同变换但是 I_n 仅做行变换 (或仅做列变换)。

第 2 章 线性算子

2.1 向量空间的线性映射

线性映射的秩：

记 $\mathscr{L}_{\mathbb{F}}(U,V)$ 为从 U 到 V 的全体线性映射，对于 $f\in\mathscr{L}_{\mathbb{F}}(U,V)$，定义 f 的“像”和“核”：

$$\operatorname{Im} f=\{f(u)\in V\mid u\in U\}\ ,\ \ \ker f=\{u\in U\mid f(u)=0\}$$

容易验证，$\operatorname{Im} f$ 是 V 的子空间，$\ker f$ 是 U 的子空间。我们称 f 的秩为：

$$\operatorname{rank} f=\dim\operatorname{Im} f$$

线性映射诱导的同构：

设 $\{u_1,...,u_m\}$ 是 U 的基，设 $f\in\mathscr{L}_{\mathbb{F}}(U,V)$，定义映射 $\varphi: f\longmapsto(f(u_1),f(u_2),...,f(u_m))$，则 φ 给出 $\mathscr{L}_{\mathbb{F}}(U,V)$ 到 V^m 的同构。依次验证单射、满射 ($\ker=\{0\}$) 即可证明同构，由此可得 $\dim\mathscr{L}_{\mathbb{F}}(U,V)=\dim U\cdot\dim V$。

Theorem. 17 (线性空间维数分解)：
设 U 是有限维向量空间，则对任意 $f\in\mathscr{L}_{\mathbb{F}}(U,V)$，我们有：

$$\operatorname{rank} f+\dim\ker f=\dim U$$

由此可推导出常用结论：$\operatorname{rank} A+\dim\ker\varphi_A=n$

2.2 线性算子

线性变换 (线性算子)：

设 V 是 $\mathbb{F}$ 上的一个线性空间，从 V 到 V 的线性映射称为线性变换 (linear transformation)，也称为线性算子，相应的集合为 $\mathscr{L}_{\mathbb{F}}(V,V)$，简记为 $\mathscr{L}(V)$。线性算子不一定要求是满的，也即映射的像 (值域) 可以是 V 的子空间。

常见的线性算子：

1. $\mathbb{F}[x]$ 上的求导算子 $\frac{\mathrm{d}}{\mathrm{d}x}$：$f\longmapsto\frac{\mathrm{d}f}{\mathrm{d}x}$
2. $C[a,b]$ 上的积分算子：$f\longmapsto\int_a^x f(t)\mathrm{d}t$

3. 投影算子：设 $V=U\oplus W$，则对于任意 $\xi\in V$，有 $\xi=\varepsilon_U+\xi_W$。定义映射 $\mathscr{P}:\xi\longmapsto\xi_U$，则 $\mathscr{P}$ 构成一个线性映射，称为 V 到 U(与 W 平行) 的投影算子。并且可以证明：

$$f\in\mathscr{L}(V),\ f^2=f\Longleftrightarrow f\text{ 是投影算子}$$

其中映射的积定义为映射的复合，也即 $f^2:x\longmapsto f(f(x))$

代数：

带有双线性乘积的线性空间称为代数。对任意的 $f,g\in\mathscr{L}(V)$，有：

$$(fg)(av+bu)=a(fg)(v)+b(fg)(u)\ ,\ \forall\,a,b\in\mathbb{F},\ v,u\in V$$
$$(af_1+bf_2)g=a(f_1g)+b(f_2g)\ ,\ \forall\,a,b\in\mathbb{F},\ f_1,f_2\in\mathscr{L}(V)$$

故 $\mathscr{L}(V)$ 上映射的乘积 (也即复合) 满足双线性性，称为代数。特别的，由于上述运算也是结合的，称 $\mathscr{L}(V)$ 为结合代数。

线性算子的极小多项式：

定义线性算子的幂：$\varphi^0=e_V,\ \varphi^k=\varphi\varphi\cdots\varphi$ (k个)，由 $\dim\mathscr{L}(V)=n^2<+\infty$ 可知存在正整数 $N\le n^2$ 使得 $\left\{e_V,\varphi,...,\varphi^k\right\}$ 线性相关，也即：

$$\exists\,0\neq f\in\mathbb{F}[x]\ ,\ \ f(\varphi)=a_0\varphi^0+a_1\varphi^1+\cdots+a_N\varphi^N=0$$

此时称多项式 f 零化线性算子 φ，在零化 φ 的多项式中，首一且次数最低的称为 φ 的极小多项式，记为 $\mu_\varphi(x)$。

由多项式的理论：所有零化 φ 的多项式构成由 $\mu_\varphi(x)$ 生成的主理想：

$$\{f\in\mathbb{F}[x]\mid f(\varphi)=0\}=\mu_\varphi(x)\cdot\mathbb{F}[x]$$

例如：零算子 0 满足 $\mu_0(x)=x$，幂零指数为 m 的算子 φ 满足 $\mu_\varphi(x)=x^m$，恒等算子 e_V 满足 $\mu_{e_V}(x)=x-1$，投影算子 φ 满足 $\mu_\varphi(x)=x^2-x$。

求矩阵的最小多项式可参考：

https://www.zhihu.com/question/402082188/answer/1289182806

https://www.zhihu.com/question/605777999/answer/3067899521

Theorem. 18 (线性算子可逆的等价条件)：
设 $\varphi\in\mathscr{L}(V)$，则：

$$\varphi\text{可逆}\Longleftrightarrow\mu_\varphi(0)\neq0$$

定理18的证明：

① φ 可逆 $\Longrightarrow \mu_\varphi(0) \neq 0$：

反证法，假设 $\mu_\varphi(0)=0$，则 $a_0=0$，于是存在 $h(x)$ 使得 $\mu_\varphi(x)=xh(x)$，由 $\deg h < \deg \mu_\varphi$ 且 μ_φ 是极小的知道 $h(\varphi)\neq 0$。另外，$\mu_\varphi(0)=\varphi h(\varphi)=0 \Longrightarrow \varphi^{-1}(\varphi h(\varphi))=h(\varphi)=0$，矛盾。

② φ 可逆 $\Longleftarrow \mu_\varphi(0) \neq 0$：

设 $\mu_\varphi(x)=a_0+a_1x+\cdots+a_mx^m$，其中 $a_0\neq 0$，$\deg \mu_\varphi = m$。记 $g(x)=a_1+a_2x+\cdots a_mx^{m-1}$，则 $\mu_\varphi(x)=a_0+xg(x) \Longrightarrow (-a_0^{-1}g(\varphi))\varphi=\varphi(-a_0^{-1}g(\varphi))=e_V$，因此 $\varphi^1=-a_0^{-1}g(\varphi)$。

Theorem. 19 （由算子生成空间基）:
设 V 是 $\mathbb{F}$ 上的 n 维线性空间，$\varphi\in\mathscr{L}(V)$ 满足 $\{e_V,\varphi,...,\varphi^{n-1}\}$ 线性无关，则：

$$\exists\, v\in V \ \text{ s.t. } \ \left\{v,\varphi(v),...,\varphi^{n-}(v)\right\} \text{ 构成 } V \text{ 的基}$$

Homework 8.1，其中两种证明在:
https://zhuanlan.zhihu.com/p/368560846 https://zhuanlan.zhihu.com/p/499412875

线性算子的矩阵：

设 $\{u_1,...,u_n\}$ 是 V 的一组基且 $\varphi\in\mathscr{L}(V)$，定义线性算子 φ 在基 $\vec{u}_0=[u_1,...,u_n]^T$ 下的矩阵 $M_{\varphi,\vec{u}_0}$ 满足：

$$\varphi(\vec{u}_0)=M_{\varphi,\vec{u}_0}\cdot\vec{u}_0 \Longleftrightarrow \varphi(\vec{u}_0)=\begin{bmatrix}\varphi(u_1)\\ \varphi(u_2)\\ \vdots\\ \varphi(u_n)\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&\ddots&\vdots\\ a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}u_1\\ u_2\\ \vdots\\ u_n\end{bmatrix}$$

容易验证推论：

$$\forall\,\varphi,\psi\in\mathscr{L}(V)\,,\quad M_{\psi\varphi,\vec{u}_0}=M_{\psi,\vec{u}_0}\cdot M_{\varphi,\vec{u}_0}$$

相似矩阵：

两个矩阵 A，B 称为相似的如果存在可逆矩阵 S 使得：

$$B=SAS^{-1}$$

线性算子度量矩阵的转换：

设 A 是基 $\vec{u}$ 向基 $\vec{v}$ 的转换矩阵，即 $\vec{v}=A\vec{u}$，则：

$$M_{\varphi,\vec{v}}=AM_{\varphi,\vec{u}}A^{-1}\,,\quad \text{也即 } M_{\varphi,\vec{u}}\longmapsto AM_{\varphi,\vec{u}}A^{-1}$$

注意这里是相似而不是合同，与之前度量矩阵的转换是不同的！

线性算子的行列式和迹：

设 $\varphi \in \mathscr{L}(V)$，定义 φ 的在基 $\vec{u}_0$ 下的行列式为 $\det M_{\varphi,\vec{u}_0} = |M_{\varphi,\vec{u}_0}|$，容易验证 φ 在任何基下的行列式都相等，称为 φ 的行列式。行列式不同的线性算子必不同，但行列式相同不代表线性算子相同。

类似地，定义 φ 在基 $\vec{u}_0$ 下的迹为 $\operatorname{Tr} M_{\varphi,\vec{u}_0}$，容易验证 φ 在任何基下的迹都相等，称为 φ 的迹。迹不同的线性算子必不同。

2.3 特征值与特征向量

Theorem. 20 (线性空间的投影分解)：
设 $\mathscr{P}_1, \mathscr{P}_2, ..., \mathscr{P}_m \in \mathscr{L}(V)$ 满足：

$$\begin{gathered} \text{① 恒等：} \mathscr{P}_1 + \cdots + \mathscr{P}_m = e_V \\ \text{② 投影：} \mathscr{P}_i^2 = \mathscr{P}_i \\ \text{③ 正交：} \mathscr{P}_i \mathscr{P}_j = 0\ ,\ \ i \neq j \end{gathered}$$

则有结论：

$$V = \mathscr{P}_1(V) \oplus \mathscr{P}_2(V) \oplus \cdots \oplus \mathscr{P}_m(V)$$

例如，已知 $V = W_1 \oplus W_2 \oplus \cdots \oplus W_m$，对任意的 $v = v_1 + \cdots + v_n$, $v_i \in W_i$，定义算子 $\mathscr{P}_i \in \mathscr{L}(V)$，则此算子满足定理条件。

不变子空间：

设 $\varphi \in \mathscr{L}(V)$，$U$ 是 V 的子空间。在 φ 下，U 称为不变的如果 $\varphi(U)$ 嵌入 U，也即 $\varphi(U) \subseteq U$，或者说 $\forall\, u \in U, \varphi(u) \in U$。

例如：在上面的定理中，$\mathscr{P}_i(V)$ 是关于所有 $\mathscr{P}_1, ..., \mathscr{P}_m$ 的不变子空间；在 $\mathbb{F}[x]$ 中，$\mathcal{P}_n[x]$ 是关于求导算子 $\frac{\mathrm{d}}{\mathrm{d}x}$ 的不变子空间；$\{0_V\}$ 和 V 是关于任意算子的不变子空间。

商算子：

设 U 是 V 关于 φ 的不变子空间，定义 φ 关于 U 的商算子 $\overline{\varphi}$：

$$\overline{\varphi}(\overline{v}) = \varphi(v) + U\ ,\ \ \forall\, \overline{v} = v + U \in V/U$$

容易验证 $\overline{\varphi} \in \mathscr{L}(V/U)$。

Theorem. 21 (不变空间补空间的不变性)：
在线性算子 φ 下，设 U 是 V 的一个非零不变真子空间，$\overline{U}$ 为 U 的补空间，则：

$$\overline{U}\ \text{不变} \iff \exists\, 0, e_V \neq \mathscr{P} \in \mathscr{L}(V, U),\ \mathscr{P}^2 = \mathscr{P}\ ,\ \ \text{使得}\ \varphi\mathscr{P} = \mathscr{P}\varphi$$

定理证明的关键在于右推左时，令 $W=(e_V-\mathscr{P})(V)$，则 $V=U\oplus W$ 且 W 是不变的。推论：V 是在 φ 下不变的两个子空间的直和的等价条件是 $\exists\, 0,e_V\neq\mathscr{P}\in\mathscr{L}(V,U)$, $\mathscr{P}^2=\mathscr{P}$，使得$\varphi\mathscr{P}=\mathscr{P}\varphi$。

特征子空间：

当不变子空间 $U=\mathbb{F}u$ 为一维时，若 $\exists\,\lambda\in\mathbb{F},0\neq u\in U$, 使得 $\varphi(u)=\lambda u$，则称 λ 为 φ 的特征值，称 u 为此特征值对应 φ 的特征向量。

考虑到 φ 的线性性，$\varphi^i(u)=\varphi(\varphi\cdots\varphi(u)\cdots)=\lambda^i u$，可推出 $0=\mu_\varphi(\varphi)(u)=\mu_\varphi(\lambda)u\implies\mu_\varphi(\lambda)=0$，$\lambda$ 是极小多项式的根 (反之也成立)。多维时也是类似的，下面我们会讨论。

定义关于 φ 的、特征值为 λ 的特征空间：

$$V_\lambda=\{v\in V\mid\varphi(v)=\lambda v\}$$

则 V_λ 构成关于 φ 的不变子空间，并称 $\dim V_\lambda$ 为 λ 的几何重数 (geometric multiplicity)。

Theorem. 22 (不同特征空间的向量线性无关)：
设 U 是 V 的关于 φ 的不变子空间 (可以是零空间)，设 $\lambda_1,...,\lambda_m$ 是 φ 的 m 个不同特征值，且 $u_i\in V_{\lambda_i}$，有结论：若 $u_1+\cdots+u_m\in U$，则 $u_1,...,u_m\in U$。
推论：设 $\lambda_1,...,\lambda_m$ 为 φ 的不同特征值，v_i 为 λ_i 对应的特征向量，则 $\{v_1,...,v_m\}$ 线性无关。

特征值、特征向量、特征多项式：

① 特征值：

$$\lambda\text{ 是 }\varphi\text{ 的特征值}\iff|\lambda I_n-M_{\varphi,\vec{u}_0}|=0\iff|\lambda I_n-M_{\varphi,\vec{u}_0}|=0\iff\chi_\varphi(\lambda)=0$$

其中 $\vec{u}_0$ 是 V 的一组基，$\vec{e}_0$ 是标准正交基。

② 特征向量：设 λ 对应的特征向量为 $u=\vec{x}\cdot\vec{u}_0$，$\vec{x}\neq\vec{0}$，则：

$$u\text{为特征向量}\iff\varphi(u)=\vec{x}A\vec{u}_0=\lambda u\iff\vec{x}\cdot(M_{\varphi,\vec{u}_0}-\lambda I_n)=\vec{0}_{1\times n}$$

$\vec{x}$ 要么只有零解，要么有无限个解，这里需要用到基础解系的知识。

③ 特征多项式：定义关于 φ 的特征多项式为：

$$\chi_\varphi(x)=|xI_n-M_{\varphi,\vec{u}_0}|=\begin{bmatrix}x-a_{11}&-a_{12}&\cdots&-a_{1n}\\a_{21}&x-a_{22}&\cdots&-a_{2n}\\\vdots&\vdots&\ddots&\vdots\\-a_{n1}&-a_{n2}&\cdots&x-a_{nn}\end{bmatrix}$$

对任意的基 $\vec{u}_0$，$\vec{v}_0$，可以推得 $|xI_n-M_{\varphi,\vec{u}_0}|=|xI_n-M_{\varphi,\vec{v}_0}|$，故特征多项式 $\chi_\varphi(x)$ 与基的选取无关。称 λ 作为 $\chi_\varphi(x)$ 的根的重数 (如 $(x-1)^3$ 重数为 3) 为 λ 的代数重数，可以证明几何重数 $\leq$ 代数重数。

线性算子对角化：

线性算子 $\varphi \in \mathscr{L}(V)$ 称为可对角化的如果 φ 在 V 某组基下的矩阵是对角矩阵。

Theorem. 23 (算子对角化)：
设线性算子 φ 的特征多项式 $\chi_\varphi(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$，则：

$$\varphi \text{ 可对角化} \iff \forall\, \lambda_i,\ \dim V_{\lambda_i} = \dim \ker_{\psi_{\lambda_i}} = m_i$$

定理23的证明：
(1) φ 对角化 $\Longrightarrow \chi_\varphi(x)$ 且 $\dim V_{\lambda_i} = m_i$：
设 φ 在基 $\vec{v}_0 = [v_1, ..., v_n]^T$ 下对角化，则

$$M_{\varphi,\vec{v}_0} = \mathrm{diag}(a_{11}, ..., a_{nn}) \iff \varphi(v_i) = a_{ii}v_i$$

故 a_{ii} 是特征值，v_i 是对应的特征向量。设 $\{\lambda_1, ..., \lambda_r\}$ 是 $\{a_{11}, ..., a_{nn}\}$ 中的所有不同元素 $(r \le n)$，记 $V^j = \mathrm{Span}\{v_i \mid a_{ii} = \lambda_j\}$，$j = 1, 2, ..., r$，记 $m_j = \dim V^j$，则：

$$\chi_\varphi(x) = (x-a_{11})(x-a_{22})\cdots(x-a_{nn}) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$$

再说明 $V^j = V_{\lambda_j}$：
① $V_{\lambda_j} \subseteq V^j$：
$V = \mathrm{Span}\{v_1, ..., v_n\} = V^1 \oplus \cdots \oplus V^r$，设 $v \in V_{\lambda_i} = \{v \in V \mid \varphi(v) = \lambda_i v\} \subseteq V$，设 $v = u_1 + \cdots u_r$ 且 $u_j \in V^j$，由 $\varphi(v) = \lambda_i v$ 得：

$$\varphi(v) = \varphi(u_1 + \cdots u_r) = \varphi(u_1) + \cdots \varphi(u_1) \overset{V^j \subseteq V_{\lambda_j}}{=} \lambda_1 u_1 + \cdots \lambda_r u_r = \lambda_i v = \lambda_i(v_1 + \cdots v_r)$$
$$\iff (\lambda_1 - \lambda_j)u_1 + \cdots (\lambda_r - \lambda_j)u_r = 0$$

根据定理22，推出 $\forall\, k \in \{1, ..., r\} \setminus \{j\}$，$u_k = 0$，于是 $v = u_j \in V^j \Longrightarrow V_{\lambda_j} \subseteq V^j$
② $V_{\lambda_j} \subseteq V^j$：验证定义即知成立，略。
(2) $\chi_\varphi(x)$ 且 $\dim V_{\lambda_i} = m_i \Longrightarrow \varphi$ 对角化：
设：

$$\chi_\varphi(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$$

根据定理22，$V_{\lambda_1} + \cdots V_{\lambda_r}$ 构成直和，且 $\dim(V_{\lambda_1} + \cdots + V_{\lambda_r}) = \lambda_1 + \cdots + \lambda_r = m_1 + \cdots m_r = \deg \chi_\varphi = \dim V$，故 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$，取每个 V_{λ_i} 的一组基 $\mathcal{B}_i$，则 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ 构成 V 的矩阵，由于 V_{λ_i} 是不变的 (有嵌入)，故 φ 在此基下的矩阵：

$$M_\varphi = \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r I \end{bmatrix}$$

也即 φ 在此基下对角化，证毕。

线性算子对角化的方法：

把一个线性算子 $\varphi \in \mathscr{L}$ 对角化，就是要找由其特征向量构成的一组基，因此需要解 φ 的特征值，并根据矩阵方程 (作列初等变换)，求取其对应的特征向量 (找到线性无关的特解即可)，最后由定理得到对角化后的矩阵为：

$$M_\varphi = \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r I \end{bmatrix}$$

借助线性算子对角化，我们可以得到原矩阵的另一种表达式，进一步还可以方便地计算原矩阵的幂 (比如用于数列通项的求解)。

Theorem. 24 （不变子空间与对角化）:
设 $\varphi \in \mathscr{L}(V)$ 且可对角化，W 是 V 关于 φ 不变的子空间。记 φ 在 W 上的限制为 $\varphi|_W$，则:

$$\varphi|_W \in \mathscr{L}(W) \text{ 且 } \varphi|_W \text{ 可对角化}$$

定理24的证明：
W 关于 φ 不变，取 W 的一组基扩充为 V 的一组基使得 φ 的矩阵形如：

$$A = \begin{bmatrix} A_1 & O \\ A_2 & A_3 \end{bmatrix}$$

其中 A_1 是 $\varphi|_W$ 的矩阵，于是

$$\mu_\varphi(A) = \begin{bmatrix} \mu_\varphi(A_1) & O \\ * & \mu_\varphi(A_3) \end{bmatrix} = 0$$
$$\Longrightarrow \mu_\varphi(A_1) = 0 \Longrightarrow \mu_{A_1} \mid \mu_\varphi \text{，而 } \varphi \text{ 可对角化等价于 } \mu_\varphi \text{ 无重因式}$$
$$\Longrightarrow \mu_{A_1} \text{ 无重因式} \Longleftrightarrow \varphi|_W \text{ 可对角化，证毕。}$$

Theorem. 25 （Skolem-Noether）:
设 $0 \neq \varphi \in \mathscr{L}(V)$ 满足 $\varphi(AB) = \varphi(A)\varphi(B)$，$\forall\, A, B \in M_{n\times n}(\mathbb{F})$，也即 φ 是同态线性算子，则：

$$\exists\, T \in M_{n\times n}(\mathbb{F}) \text{ s.t. } \varphi(A) = TAT^{-1}\,,\ \forall\, A \in M_{n\times n}(\mathbb{F})$$

Homework 7.6，证明详见“习题课 7.pdf”。

Theorem. 26 （算子积的特征多项式）:
$\forall\, \varphi, \psi \in \mathscr{L}(V)$，有：

$$\chi_{\varphi\psi}(x) = \chi_{\psi\varphi}(x)$$

证明 $\forall A, B \in M_{n\times n}(\mathbb{F})$，$|AB - xI_n| = |BA - xI_n|$ 即可证明此定理。

对偶算子：

设 $\varphi \in \mathscr{L}(V)$，定义 φ 对应的对偶算子 $\varphi^* : V^* \longmapsto V^*$ 为：

$$\varphi^*(f) = f\varphi\,,\ \forall f \in V^*$$

也即 $\forall\, v \in V,\ (\varphi^*(f))(v) = f\varphi(v) = f(\varphi(v))$，且 $\varphi^* \in \mathscr{L}(V^*)$。

对偶算子的一个应用是证明具有不变超平面的充分条件：设 $\varphi^* \in \mathscr{L}(V^*)$ 且 φ^* 有非零特征值 λ (易证 φ^* 在基下的矩阵是 φ 矩阵的转置)，对应的特征向量为 f。令 $U = \{v \in V \mid f(v) = 0\}$，则 $\forall\, u \in U,\ f(\varphi(u)) = (\varphi^*(f))(u) = \lambda f(u) = 0 \Longrightarrow \varphi(u) \in U$，因此 U 是 V 余维数为 1 的、不变的子空间 (即不变的超平面)

另外，我们有：

$$(\varphi\psi)^* = \psi^*\varphi^*$$

于是映射 $\varphi \longmapsto \varphi^*$ 构成一个代数反同态。

对偶算子的矩阵：

设 $\vec{v}_0 = [v_1, ..., v_n]^T$ 是 V 的一组基，对应的 $\vec{v}^0 = [v^1, ..., v^n]^T$ 是 V^* 的一组基，则有：

$$M_{\varphi^*, \vec{v}^0} = M_{\varphi, \vec{v}_0}^{T}$$

由此可说明映射 $\varphi \longmapsto \varphi^*$ 是一个代数同构。

2.4 Jordan 标准型

本节我们总假设 $\mathbb{F} = \mathbb{C}$。

Theorem. 27 (Hamilton-Cayley Theorem):
设 V 是 $\mathbb{C}$ 上的向量空间且 $\varphi \in \mathscr{L}(V)$，则：

$$\chi_\varphi(\varphi) = |xI_n - M_{\varphi, \vec{u}_0}|_{x=\varphi} = 0$$

广义特征子空间 (根子空间)：

设 φ 的特征多项式 $\chi_\varphi(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$，$\lambda_i$ 为 φ 的特征值，φ 的特征值为 λ_i 的广义特征子空间：

$$V(\lambda_i) = \left\{v \in V \mid \exists\, k \in \mathbb{N} \text{ 使 } (\varphi - \lambda_i)^k(v) = 0\right\}$$

容易验证 $V(\lambda_i)$ 是关于 φ 不变的。

Theorem. 28 (线性空间的广义特征子空间分解):
设 $\varphi \in \mathscr{L}(V)$，$\lambda_1, ..., \lambda_r$ 为 φ 所有不同特征值，则：

$$V = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_r)\,,\ \text{且 } \dim V(\lambda_i) = m_i$$

Jordan 块:

设 φ 满足 $r=1$ 且 $(\varphi-\lambda)$ 的幂零指数是 n，则 $\chi_\varphi(x)=(x-\lambda)^m$。由定理27、定理28，$m=n=\dim V$，存在 v_1 使得 $(\varphi-\lambda)^{n-1}(v_1)\neq 0$，且 φ 在基 $\{v_1,(\varphi-\lambda)(v_1),...,(\varphi-\lambda)^{n-1}(v_1)\}$ 下的矩阵为：

$$J_n(\lambda)=\begin{bmatrix}\lambda & 1 & 0 & \cdots & 0\\ 0 & \lambda & 1 & \cdots & 0\\ 0 & 0 & \lambda & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & 1\\ 0 & 0 & 0 & \cdots & \lambda\end{bmatrix}_{n\times n}$$

称 $J_n(\lambda)$ 是特征值为 λ 的 n 阶 Jordan 块。

循环子空间：

设 V 是 $\mathbb{F}$ 上的线性空间且 $\varphi\in\mathscr{L}(V)$，定义由 v 生成的、关于 φ 的循环子空间：

$$\mathbb{F}[\varphi]v=\text{Span}\left\{\varphi^i(v)\mid i\in\mathbb{N}\right\}=\text{Span}\left\{\varphi^0(v),\varphi(v),\varphi^2(v),...\right\}$$

容易验证 $\mathbb{F}[\varphi]v$ 关于 φ 不变。

Theorem. 29 (幂零算子可诱导循环分解):
设 ψ 是 V 上的幂零算子，令 $t=\dim\ker_\psi$，则存在线性无关的 $v_1,...,v_t\in V$ 使得：

$$V=\mathbb{C}[\psi]v_1\oplus\cdots\oplus\mathbb{C}[\psi]v_t\ ,\ \text{且记}\ k_j=\min\left\{k\mid\psi^k(v_j)=0\right\},\ \text{则有}\ \dim V=\sum_{j=1}^t k_j$$

其中 $\mathbb{C}[\psi]v_j=\text{Span}\left\{\psi^0(v_j),\psi(v_j),...,\psi^{k_j-1}(v_j)\right\}$ 是由 v_j 生成的关于 ψ 的循环子空间。

定理29的证明：
对 $\dim V=n$ 归纳，当 $n=0,1$ 时显然成立，假设结论对 $<n$ 成立：记 $W=ker_\psi=V_\lambda=\{v\in V\mid\psi(v)=0\}$，考虑 V/W 到 V/W 的映射：

$$\tilde{\psi}:\ v+W\longmapsto\psi(v)+W$$

由于 $\dim V/W=\dim V-\dim W=\dim\text{Im}(\psi)$，$\psi$ 幂零因此 $\tilde{\psi}$ 也幂零 $\Longrightarrow\dim\text{Im}(\psi)<\dim V=n$。根据假设，存在 $\overline{v}_1,...,\overline{v}_{\tilde{t}}$，$\tilde{k}_1,...,\tilde{k}_{\tilde{t}}$ 使得：

$$V/W=\mathbb{C}[\tilde{\psi}](\overline{v}_1)\oplus\cdots\oplus\mathbb{C}[\tilde{\psi}](\overline{v}_{\tilde{t}})=\bigoplus_{j=1}^{\tilde{t}}\text{Span}\left\{\overline{\psi^0(v_j)},...,\overline{\psi^{\tilde{k}_j-1}(v_j)}\right\}\ ,\ \text{且}\ \sum_{j=1}^{\tilde{t}}\tilde{k}_j=\dim V/W$$

又 $\tilde{k}_j=\min\{k\mid(\tilde{\psi})^k(\overline{v}_j)=\overline{\psi^k(v)}=\overline{0}\sim\psi^k(v)\in W\}\Longrightarrow k_j=\tilde{k}_j+1$，于是：

$$V=W\oplus\bigoplus_{j=1}^{\tilde{t}}\text{Span}\left\{\psi^0(v_j),...,\psi^{k_j-2}(v_j)\right\}\ ,\ \text{且}\ \sum_{j=1}^{\tilde{t}}k_j+(\dim W-\tilde{t})=n$$

考虑 $\{\psi^{\tilde{k}_1}(v_1),...,\psi^{\tilde{k}_{\tilde{t}}}(v_{\tilde{t}})\}\subset W$ 是否线性相关，假设 $\exists\, a_1,...,a_{\tilde{t}}$ 使：

$$\begin{gathered}
0=a_1\psi^{\tilde{k}_1}(v_1)+\cdots a_{\tilde{t}}\psi^{\tilde{k}_{\tilde{t}}}(v_{\tilde{t}})=\psi\left(a_1\psi^{k_1-2}(v_1)+\cdots a_{\tilde{t}}\psi^{k_{\tilde{t}-2}}(v_{\tilde{t}})\right)\\
\Longrightarrow a_1\psi^{k_1-2}(v_1)+\cdots a_{\tilde{t}}\psi^{k_{\tilde{t}-2}}(v_{\tilde{t}})\in \mathrm{ker}_\psi=W\\
\Longleftrightarrow \exists\, w\in W,\ s.t.\ \ w=a_1\psi^{k_1-2}(v_1)+\cdots a_{\tilde{t}}\psi^{k_{\tilde{t}-2}}(v_{\tilde{t}})\\
\Longrightarrow w\in\bigoplus_{j=1}^{\tilde{t}}\mathrm{Span}\left\{\psi^0(v_j),...,\psi^{k_j-2}(v_j)\right\}\Longrightarrow a_1=\cdots a_{\tilde{t}}=0
\end{gathered}$$

故为线性无关组，将其扩充为 W 的一组基 $\{\psi^{\tilde{k}_1}(v_1),...,\psi^{\tilde{k}_{\tilde{t}}}(v_{\tilde{t}}),w_{\tilde{t}+1},...,w_{\dim W}\}$，此时 $\forall\, j\in\{\tilde{t}+1,...,\dim W\},\ k_j=1$，且：

$$V=\bigoplus_{j=1}^{\tilde{t}}\mathrm{Span}\left\{\psi^0(v_j),...,\psi^{k_j-1}(v_j)\right\}\oplus\bigoplus_{j=\tilde{t}+1}^{t=\dim W}\mathrm{Span}\left\{w_j\right\}\ ,\ \text{且}\ \sum_{j=1}^{t}k_j=n$$

也即：

$$V=\bigoplus_{j=1}^{\tilde{t}}\mathbb{C}[\psi]v_j\oplus\bigoplus_{j=\tilde{t}+1}^{t=\dim W}\mathbb{C}[\psi]v_j=\bigoplus_{j=1}^{t}\mathbb{C}[\psi]v_j\ ,\ \text{且}\ \sum_{j=1}^{t}k_j=n$$

证毕。

Theorem. 30 (算子必有 Jordan 标准型):
设 $\varphi\in\mathscr{L}(V)$，则存在 V 的一组基使得 φ 在其下的矩阵为：

$$J_\varphi=\bigoplus_{i=1}^{r}\bigoplus_{j=1}^{t_i}J_{k_{ij}}(\lambda_i)=\begin{bmatrix}J_1&O&\cdots&O\\O&J_2&\cdots&O\\\vdots&\vdots&\ddots&\vdots\\O&O&\cdots&J_r\end{bmatrix}_{n\times n}\ ,\ J_i=\begin{bmatrix}J_{k_{i1}}(\lambda_i)&O&\cdots&O\\O&J_{k_{i2}}(\lambda_i)&\cdots&O\\\vdots&\vdots&\ddots&\vdots\\O&O&\cdots&J_{k_{it_i}}(\lambda_i)\end{bmatrix}_{m_i\times m_i}$$

称为 φ 的 Jordan 标准型，相应的基称为 Jordan 基。

并且，设以 λ_i 为特征值的 Jordan 块中的最大阶是 k_i，则 φ 的极小多项式为：

$$\mu_\varphi(x)=(x-\lambda_1)^{k_1}\cdots(x-\lambda_r)^{k_r}$$

其中 $\chi_\varphi(x)=(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$，$t_i=\dim\ker_{\psi_i}=\dim V_{\lambda_i}$，$k_{ij}=\min\{k\mid\psi_i^k(v_j)=0\}$，$\sum_{j=1}^{t_i}k_{ij}=m_i=\dim V(\lambda_i)$，$\sum_{i=1}^{r}m_i=n=\dim V$。

对于 J_i，可以简记其分为“维数个部分”(由 $\dim V_{\lambda_i}$ 个 Jordan 块构成)，每个 Jordan 块的大小是“ij 幂零指数”(即 k_{ij}，是 $\psi_i=\varphi-\lambda_i e$ 对 v_j 的幂零指数)。

另外，需要特别注意：

$$V \text{ 的基 } \vec{v}_0 = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in {}^nV\text{，} V \text{ 的元素 } v = \vec{x} \cdot \vec{v}_0 \in V\text{，元素的坐标 } \vec{x} = [x_1, ..., x_n] \in \mathbb{F}^n$$

$$\text{矩阵 } A \text{ 对应的算子 } \varphi_A \text{ 定义为：} \varphi_A(v) = \vec{x}A\vec{v}_0$$

定理30的证明：

(1) 根子空间分解：

设 $\chi_\varphi(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$，其中 $\lambda_1, ..., \lambda_r$ 为 φ 的所有不同特征值。由定理28，我们有：

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_r)$$

设 $\mathcal{B}_i$ 为 $V(\lambda_i)$ 的一组基，则 $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ 构成 V 的一组基，φ 在该基下的矩阵 $A = A_1 \dotplus \cdots \dotplus A_r$(因为广义特征子空间是不变的)。因此只需要证明，限制在 $V(\lambda_i)$ 上的映射 $\varphi|_{V(\lambda_i)}$ 有 Jordan 标准型。

(2) 每个根子空间上有标准型：

对根子空间 $V(\lambda_i)$，令 $\psi_i = \varphi|_{V(\lambda_i)} - \lambda_i e$，由定理27，$\chi_{\varphi|_{V(\lambda_i)}}(\varphi|_{V(\lambda_i)}) = \psi_i^{\dim V(\lambda_i)} = \psi_i^{m_i} = 0$，故 ψ 是 $V(\lambda_i)$ 上的幂零算子。由定理29，记 $t_i = \dim \ker_{\psi_i}$，则存在 $\{v_{i1}, ..., v_{it_i}\} \subset V(\lambda_i)$ 使得：

$$V(\lambda_i) = \bigoplus_{j=1}^{t_i} \mathbb{C}[\psi_i]v_{ij}\ ,\quad \text{且} \sum_{j=1}^{t_i} k_{ij} = \dim V(\lambda_i) = m_i$$

又 $\mathbb{C}[\varphi|_{V(\lambda_i)}]v = \mathbb{C}[\psi_i]v$,因此 $V(\lambda_i) = \bigoplus_{j=1}^{t_i} \mathbb{C}[\psi_i]v_{ij}$,且 $\varphi|_{V(\lambda_i)}$ 在基 $\left\{\psi^0(v_1), ..., \psi^{k_{i1}-1}(v_1)\right\} \cup \cdots \cup \left\{\psi^0(v_{t_i}), ..., \psi^{k_{it_i}-1}(v_{t_i})\right\}$ 下的矩阵是：

$$J_i = J_{k_{i1}}(\lambda_i) \dotplus \cdots \dotplus J_{k_{it_i}}(\lambda_i)$$

(3) 综合：

综合 (1)(2) 得到：

$$J = \bigoplus_{i=1}^{r} J_i = \bigoplus_{i=1}^{r}\bigoplus_{j=1}^{t_i} J_{k_{ij}}(\lambda_i)$$

其中 $t_i = \dim \ker_{\psi_i} = \dim V_{\lambda_i}$，$k_{ij} = \min\{k \mid \psi_i^k(v_j) = 0\}$，$\sum_{j=1}^{t_i} k_{ij} = m_i = \dim V(\lambda_i)$，$\sum_{i=1}^{r} m_i = n = \dim V$。证毕。

https://www.zybuluo.com/ybtang21c/note/1827223 (求 Jordan 标准型的方法及例子)

https://zhuanlan.zhihu.com/p/553660985 (Jordan 标准型理论概要)

https://zhuanlan.zhihu.com/p/75745789 (Jordan 标准型的循环子空间证明)

Theorem. 31 (算子的 Jordan 标准型唯一):
设 $\varphi \in \mathscr{L}_{\mathbb{C}}(V)$，则:

除小 Jordan 块 $J_{k_{ij}}(\lambda_i)$ 的次序外，φ 的 Jordan 标准型是唯一的。

算子 Jordan 化并求 Jordan 基:

依据定理29，定理30，我们给出将线性算子 Jordan 化的系统方法: 设算子 $\varphi \in \mathscr{L}(V)$，$A$ 是 φ 在某组基下的矩阵 (一般认为是标准正交基)，则 Jordan 化步骤如下:

① 求特征值: $\chi_\varphi(x) = |xI_n - M_{\varphi,\vec{u}_0}| = (x-\lambda_1)^{m_1}\cdots(x-\lambda_r)^{m_r} \Longrightarrow V = V(\lambda_1)\oplus\cdots\oplus V(\lambda_r)$ 对于每个 λ_i，令 $\psi = \varphi - \lambda e_V$。

② 确定 $V(\lambda_i)$ 分为几部分: 求出 $\mathrm{rank}\,\psi = \mathrm{rank}\,(A-\lambda I)$，则 "份数" $= \dim\ker_\psi = n - \mathrm{rank}\,\psi$。
"份数" = 几何重数 = 特征子空间维度

③ 确定 $V(\lambda_i)$ 每部分的维数：先根据 m_i 和 $\dim\ker_\psi$ 判断是否能确定维数，若不能，进一步计算 $\psi^2, \psi^3, ...$，直至确定各部分维数。

④ 确定 $V(\lambda_i)$ 的所有小 Jordan 块: 设某份维数是 k，找到 $v \in V$ 使得:

$$\begin{cases} v \notin \ker_{\psi^{k-1}} \iff \vec{x}(A-\lambda I)^{k-1} \neq 0 \\ v \in \ker_{\psi^k} \iff \vec{x}(A-\lambda I)^k = 0 \end{cases}$$

即得到基 $\{v, \psi(v), ..., \psi^{k-1}(v)\}$ 下的一个小 Jordan 块。改变 k 为下一份的值并重复此步骤，得到 $V(\lambda_i)$ 的所有小 Jordan 块。

⑤ 将所有根子空间的基合并，得到最终结果。

Theorem. 32 (极小多项式):
设矩阵 $A = A_1 \dotplus A_2 \dotplus \cdots \dotplus A_m$，则:

$$\mu_A = l.c.m(\mu_{A_1}, ..., \mu_{A_m})$$

对于线性算子 φ，考虑到算子在不同基下的特征多项式不变，可借助 Jordan 标准型求此算子的最小多项式。特别地，如果算子在基下的矩阵就是矩阵直和，则省去了 Jordan 分解的步骤。

Theorem. 33 (特征多项式的性质):
设 $\varphi \in \mathscr{L}(V)$ 的特征多项式为

$$\chi_\varphi(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r} = x^n + a_{n-1}x^{n-1} + \cdots + a_1x^1 + a_0$$

由韦达定理，我们有:

$$a_{n-1} = (-1)^1(m_1\lambda_1 + \cdots + m_r\lambda_r)$$
$$\vdots$$
$$a_0 = (-1)^n\lambda_1^{m_1}\cdots\lambda_r^{m_r} \Longrightarrow \det(M_\varphi) = \lambda_1^{m_1}\cdots\lambda_r^{m_r}$$

Theorem. 34 (幂零矩阵等价于仅有零特征值):
设 $A \in M_{n\times n}(\mathbb{C})$，则：

$$A \text{ 为幂零矩阵} \iff A \text{ 有且仅有零特征值}$$

Homework 10.1

定理34的证明:
(1) 幂零 $\Longrightarrow$ 零特征值:
$\exists\, m \in \mathbb{N}_+$ 使得 $A^m = 0 \Longrightarrow |A^m| = |A|^m = 0 \Longrightarrow |A| = 0 \Longrightarrow \chi_{\varphi_A}(0) = |0I - A| = |A| = 0 \Longrightarrow 0$ 为 A 的特征值。设 λ 为 A 的任一特征值，$0 \neq v_\lambda \in V(\lambda)$ 为一特征向量，则 $varphi^m(v) = \lambda^m v = 0 \Longrightarrow \lambda = 0$，因此 A 有且仅有零特征值。
(2) 幂零 $\Longleftarrow$ 零特征值：A 有且仅有零特征值，因此特征多项式 $\chi_\varphi(x) = (x-0)^n = x^n$，由 Hamilton-Cayley Theorem，$\chi_\varphi(\varphi) = 0 \Longrightarrow A^n = 0 \Longrightarrow A$ 为幂零矩阵。

第 3 章　带有数乘的线性空间：

3.1　欧几里得空间 (Euclidean Space)

欧几里得空间：

一个$\mathbb{R}$ 上的线性空间 V 称为欧式空间如果它带有正定的双线性型 $f : V \times V \longrightarrow \mathbb{R}$

$$f : (u, v) \longmapsto (u \mid v)\ ,\quad u, v \in V$$

称为上面的映射为欧内积，并且有相关概念：

① 模/长度：$\|u\| = \sqrt{(u \mid u)}$

② 距离：$d_{uv} = \|u - v\|$

③ 正交：$u \perp v \Longleftrightarrow (u \mid v) = 0$

④ 夹角：$\theta = \frac{(u|v)}{\|u\| \cdot \|v\|}$

⑤ 单位：$\|u\| = \sqrt{(u \mid u)} = 1$

⑥ 标准正交：一组正交向量 $\{v_1, ...v_r\}$ 称为标准的如果 v_i 是单位的，$i = 1, ..., r$。

欧内积是一个正定的对称双线性型，有其对应的二次型。

例如：通常的 n 维坐标空间 $\mathbb{R}^n$ 中，我们定义的内积是 $f(u, v) = \vec{x} I_n \vec{y}^T = \vec{x} \cdot \vec{y}^T = \sum_{i=1}^{n} x_i y_i$，也就是 $(\vec{x} \mid \vec{y}) = \sum_{i=1}^{n} x_i y_i$, $\vec{x} = [x_1 ... x_n]$, $\vec{y} = [y_1 ..., y_n] \in \mathbb{R}^n$；$[a, b]$ 上的实连续函数空间 $C([a, b])$ 内积定义为 $(u \mid v) = \int_a^b u(x) v(x) \mathrm{d}x$, $u(x), v(x) \in C([a, b])$。

Theorem. 35 (Cauchy-Schwarz Inequality)：
设 V 是欧式空间，$u, v \in V$，则有：

$$|(u \mid v)| \leq \|u\| \cdot \|v\|\ ,\ \text{当且仅当 } u, v \text{ 线性相关时取等}$$

推论：

$$\|u \pm v\| \leq \|u\| + \|v\|$$

Theorem. 36 (欧式空间必有标准正交基)：
设 V 为有限维欧式空间，$\dim V = n$，则：

$$V \text{ 存在正交标准基 } \vec{u}_0 = [u_1, ..., u_n]^T$$

> **定理36的证明：**
> **(1)** 引理 (施密特正交化)：
> 设 $0 \neq v_0, v_1, v_2, ..., v_r \in V$，令 $u = v_0 - \frac{(u|v_1)}{(v_1|v_1)} \cdot v_1 - \cdots - \frac{(u|v_r)}{(v_r|v_r)} \cdot v_r$，则:
>
> $$u \perp v_i, i = 1, ..., r$$

(2) 构造标准正交基：
设 $\vec{v}_0 = [v_1, ..., v_n]^T$ 是 V 的一组基，考虑施密特正交化。令：

$$u_1 = v_1\ ,\ \ u_2 = v_2 - \frac{(v_2 \mid u_1)}{\|u_1\|^2} \cdot u_1\ ,\ \ u_n = v_r - \sum_{i=1}^{n-1} \frac{(v_n \mid u_i)}{\|u_i\|^2} \cdot u_i$$

则 $\vec{u}_0 = [u_1, ..., u_n]^T$ 构成一组正交基，再做标准化：

$$w_i = \frac{u_i}{\|u_i\|}\ ,\ \ i = 1, ..., n$$

即可得到一组标准正交基 $\vec{w}_0 = [w_1, ..., w_n]^T$。证毕。

Theorem. 37 (欧式子空间与其补正交)：
设 V 为有限维欧式空间，$\dim V = n$，U 为 V 的子空间，$\overline{U}$ 是 U 的补空间，则：

$$\overline{U} = U^{\perp} = \{v \in V \mid \forall\, u \in U,\ (v \mid u) = 0\}\ ,\ \ \text{也即 } V = U \oplus U^{\perp}$$

推论①(任意标准正交组可扩充)：

任意一组标准正交向量 $\{v_1, ..., v_r\}$ 可扩充为 V 的标准正交基$\{v_1, ..., v_r, v_{r+1}, ..., v_n\}$

推论②(向量的基表示):

设 $\vec{w}_0$ 为 V 的标准正交基，则: $v = \sum_{i=1}^{n} \langle v \mid w_i \rangle w_i$

推论③(帕塞瓦尔恒等式)：设 $\{w_1, ... w_n\}$ 是 V 的标准正交基，则

$$\sum_{i=1}^{n} (v \mid w_i)(w_i \mid u) = (v \mid u)$$

对偶欧式空间：

设 V 为欧式空间，$u \in V$，定义 $\Phi_u(v) \in V^*$ 为：

$$\Phi_u(v) = (u \mid v)\ ,\ \ \forall\, v \in V$$

定义 V^* 上的内积为：

$$(\Phi_u \mid \Phi_v)^* = (u \mid v)$$

容易验证它构成一个正定的对称双线性型。

另外，映射 $\varphi : u \longmapsto \Phi_u$ 给出了 V 到 V^* 的线性同构，进一步地，φ 构成欧几里得同构。
线性 + 同构 + 保持内积运算

伴随算子：

设 $\varphi \in \mathscr{L}(V)$，定义 φ 的伴随算子 (adjoint operator) $\varphi^* \in \mathscr{L}(V)$ 为：

$$(\varphi^*(u) \mid v) = (u \mid \varphi(v)) \ , \ \ \forall\, u, v \in V$$

设 $\vec{w}_0$ 是任意一组标准正交基，则有：

$$M_{\varphi^*, \vec{w}_0} = M_{\varphi, \vec{w}_0}^T$$

一般情形：$M_{\varphi^*, \vec{v}_0} = AA^T M_{\varphi, \vec{w}_0}^T (A^{-1})^T A^{-1}$

Theorem. 38 (自伴随算子):
φ 为 V 的自伴随线性算子如果

$$\varphi^* = \varphi \iff V = \mathrm{Im}\,\varphi \oplus \ker\varphi \iff \varphi \text{ 在某组标准正交基下的矩阵是对称矩阵}$$

欧算子 (欧自同构):

在本笔记中，我们将"欧几里得自同构"称为"欧算子"，这是为了突出其与酉空间中"酉算子"的对应关系，将酉空间中的"Hermitian 算子"也称为"自伴随算子"，这是为了突出其和欧空间中"自伴随算子"的对应关系。

设 V 是欧式空间，线性算子 φ 称为欧的如果：

$$(\varphi(u) | \varphi(v)) = (u \mid v) \ , \ \ \forall\, u, v \in V$$

更常见的名字为欧几里得自同构。设欧算子的矩阵为 A，有推论：

$$A \text{ 为欧矩阵} \iff AA^T = I_n \iff A^{-1} = A^T \iff A \text{ 为正交矩阵}$$
$$\varphi \text{ 为欧算子} \iff \varphi\varphi^* = e_V \iff \varphi^{-1} = \varphi^* \iff \varphi \text{ 为正交变换}$$

例如，设 $\{v_1, ..., v_n\}$，$\{u_1, ..., u_n\}$ 分别是 V 和 U 的一组标准正交基，定义线性映射 $\varphi : v_i = u_i,\ i = 1, ..., n$，则 φ 构成一个欧几里得同构。

欧式空间 V 上的全体自同构 $\mathrm{Aut}_e(V)$ 关于映射的乘积 (复合) 构成群，且映射 $\varphi \longmapsto M_{\varphi, \vec{w}_0}$ 构成 $\mathrm{Aut}_e(V)$ 到正交群 $O_n(\mathbb{R})$ 的群同构。

这是因为在有限维欧式空间 V，$\forall\, \varphi \in \mathscr{L}(V)$，有：

$$(\mathrm{Im}\,\varphi)^\perp = \ker\varphi^* \Longrightarrow V = \mathrm{Im}\,\varphi \oplus (\mathrm{Im}\,\varphi)^\perp = \mathrm{Im}\,\varphi \oplus \ker\varphi^*$$

3.2 辛空间 (Symplectic Space)

辛空间、辛算子、辛矩阵：

一个线性空间空间 V 称为辛空间如果它带有非退化的斜对称双线性型 (称为辛内积)：

$$(u, v) \longmapsto [u \mid v]$$

线性算子 $\varphi \in \mathscr{L}(V)$ 称为辛算子如果：

$$[\varphi(u) \mid \varphi(v)] = [u \mid v]$$

由第一章内容，$\dim V = 2m$ 为偶数，且辛内积 $[\cdot \mid \cdot]$ 在某组基 $\vec{u}_0$ 下的度量矩阵为：

$$J_0 = \begin{bmatrix} [u_1 \mid u_1] & \cdots & [u_1 \mid u_{2m}] \\ \vdots & \ddots & \vdots \\ [u_{2m} \mid u_1] & \cdots & [u_{2m} \mid u_{2m}] \end{bmatrix} = \begin{bmatrix} O & I_m \\ -I_m & O \end{bmatrix}$$

设辛算子在辛标准基 $\vec{u}_0$ 下的矩阵为 A，则有：$AJ_0A^T = J_0$，并称 A 为辛矩阵。

辛群：

记全体 $2m \times 2m$ 辛矩阵为 $Sp_{2m}(\mathbb{R})$，则 $Sp_{2m}(\mathbb{R})$ 构成 $GL_{2m\times 2m}(\mathbb{R})$ 的一个子群，称为辛群。并且，设 $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$，则有：

$$AJ_0A^T = J_0 \iff \begin{cases} A_{11}A_{22}^T - A_{12}A_{21}^T = I_{2m} \\ (A_{11}A_{12}^T)^T = A_{11}A_{12}^T \\ (A_{21}A_{22}^T)^T = A_{21}A_{22}^T \end{cases}$$

全体辛算子关于映射的乘积构成一个群，记为 $Sp(V)$，且映射 $\varphi \longmapsto M_{\varphi,\vec{u}_0}$ 是 $Sp(V)$ 到 $Sp_{2m}(\mathbb{R})$ 的同构。

构成辛群是因为对于矩阵 A，B，我们有结论：

$$AJ_0A^T = J_0 \iff (A^T)J_0(A^T)^T = J_0 \iff (A^{-1})J_0(A^{-1})^T = J_0$$
$$A,\ B \in Sp_{2m}(\mathbb{R}) \Longrightarrow AB \in Sp_{2m}(\mathbb{R})$$

由等价定义，我们可以构造一些辛矩阵，如下：

$$\begin{bmatrix} A & O \\ O & (A^T)^{-1} \end{bmatrix} \in Sp_{2m}(\mathbb{R}) \text{ , } \forall\, A \in GL_n(\mathbb{R})$$
$$\begin{bmatrix} I_m & A \\ O & I_m \end{bmatrix}, \begin{bmatrix} I_m & A \\ O & I_m \end{bmatrix} \in Sp_{2m}(\mathbb{R}) \text{ , } \forall\, A \in GL_n(\mathbb{R})$$

可以证明辛群是由上述矩阵生成的。特别地，辛矩阵行列式为 1 (注意不是 -1)，也即 $Sp_{2m}(\mathbb{F}) \leq SL_{2m}(\mathbb{F})$。

Theorem. 39（辛算子的特征多项式）：
设 V 是有限维辛空间，$\varphi \in Sp_{2m}(\mathbb{R})$，则：

$$\chi_\varphi(x) = x^{2m}\chi_\varphi(x^{-1})$$

此定理可以导出一些与辛算子特征值有关的结论。在欧空间中，我们类似地有：$\chi_\varphi(x) = x^n\chi_\varphi(x^{-1})$，详见 Homework 11.6

辛空间与欧空间的联系：

设 $V=\mathbb{R}^{2m}$，$\vec{u}_0$ 是 V 的一组基，定义辛内积和欧内积，定义算子 $\mathcal{J}$：

$$[u \mid v]=\sum_{i=1}^{m}(x_iy_{m+i}-x_{m+i}y_i)\ ,\ \ (u \mid v)=\sum_{i=1}^{2m}x_iy_i\ ,\ \forall\, u=\vec{x}\cdot\vec{u}_0,\ v=\vec{y}\cdot\vec{u}_0$$

$$\mathcal{J}(u)=\mathcal{J}(\vec{x})\cdot\vec{u}_0=[x_{m+1},...,x_{2m},-x_1,...,-x_m]\cdot\vec{u}_0\ ,\ \forall\, u=\vec{x}\cdot\vec{u}_0$$

则 $(V,[\cdot\mid\cdot])$ 构成辛空间，$(V,(\cdot\mid\cdot))$ 构成欧空间，$\mathcal{J}\in Sp_{2m}(\mathbb{R})$。且 $[\cdot\mid\cdot]$ 在基 $\vec{u}_0$ 下的度量矩阵为 J_0，$\mathcal{J}$ 在基 $\vec{u}_0$ 下的矩阵为 $-J_0$，$\mathcal{J}$ 构成一个辛算子。

另外，容易验证 $\mathcal{J}^2=-e_V$，$[u\mid v]=(u\mid\mathcal{J}(v))$。

https://zhuanlan.zhihu.com/p/606731586

辛子空间的正交空间：

设 V 是有限维辛空间，W 是 V 的子空间，则：

$$\dim W+\dim W^{\perp}=\dim V\ ,\ \ (W^{\perp})^{\perp}=W$$

$$V=W\oplus W^{\perp}\iff W\cap W^{\perp}=0\iff W\text{ 构成辛空间}\iff W^{\perp}\text{ 构成辛空间}$$

与欧空间类似，其中 $W^{\perp}=\{v\in V\mid [v\mid w]=0,\ \forall\, w\in W\}$。注意：欧空间中一定有 $V=W\oplus W^{\perp}$ 但辛空间不一定。W 称为辛子空间如果 $W\oplus W^{\perp}$，称为迷向子空间如果 $W\subseteq W^{\perp}$，称为 Larange 子空间如果 $W=W^{\perp}$。

且辛迷向的维数 $\leq\frac{\dim V}{2}$。这是因为 $W\perp J(W)\Longrightarrow J(W)\subseteq W^{\perp}\Longrightarrow 2\dim W=\dim W+\dim\mathcal{J}(W)\leq\dim W+\dim W^{\perp}=\dim V$。

这个例子表明，子空间 (满足封闭性) 并不一定能继承原空间的内积成为新的内积空间。

3.3 酉空间 (Unitary Space)

欧空间 V 是 $\mathbb{R}$ 上的，带有正定双线性型 $f:V\times V\longrightarrow\mathbb{R}$ 的线性空间。而我们希望将其拓展到 $\mathbb{C}$ 上，由此产生了 $\mathbb{C}$ 上的，带有正定 Hermitian 型 $f:V\times V\longrightarrow\mathbb{C}$ 的线性空间，称为酉空间。

Hermitian 型:

映射 $f:V\times V\longrightarrow\mathbb{C}$ 称为 Hermitian 型如果：

① 共轭对称 (Hermitian 对称)： $f(u,v)=\overline{f(v,u)}\ ,\forall u,v\in V$

② 左线性： $f(au_1+bu_2,v)=af(u_1,v)+bf(u_2,v)\ ,\forall a,b\in\mathbb{C},\ u_1,u_2,v\in V$

③ 右共轭线性： $f(u,av_1+bv_2)=\overline{a}f(u,v_1)+\overline{b}f(u,v_2)\ ,\forall a,b\in\mathbb{C},\ u,v_1,v_2\in V$

对任意 Hermitian 型 f，其在基 $\vec{v}_0$ 下的矩阵 F 满足 $F=F^H\iff G=G^T\text{且}H=-H^T$，称为 Her 矩阵。

与欧空间中的实双线性型类似，在酉空间下的坐标空间 $\mathbb{C}^n$ 中，Hermitian 型可表示为:

$$f(u,v)=\vec{x}F\vec{y}^H=\vec{x}F\vec{y}^*\ ,\ \forall\, u=\vec{x}\cdot\vec{v}_0,\ v=\vec{y}\cdot\vec{v}_0\in\mathbb{C}$$

$F=M_{f,\vec{v}_0}$ 是 f 在基 $\vec{v}_0$ 下的矩阵。双线性型 (包括对称和斜对称) 的伴随是转置 T，Hermitian 型的伴随是共轭转置 H，常统一用 $*$ 表示伴随。

相应地，可以建立 Hermitian 二次型的概念: $q(u)=f(u,u)=\vec{x}F\vec{x}^H\ ,\ \forall\, u=\vec{x}\cdot\vec{v}_0\in V$

https://www.zhihu.com/question/533224060/answer/3345977116

Hermitian 矩阵空间与实对称矩阵空间同构，并且很多实对称矩阵 (双线性型) 的性质、结论都可以直接推广到 Hermitian 矩阵 (Hermitian 型)。下面是一些基本的性质、结论：

Hermitian 矩阵的性质：

① 对角元素为实数：Hermitian 矩阵对角元素都是实数，因为它们与自身的共轭相等。

② 实特征值：Hermitian 矩阵的特征值都是实数。

③ 正定性：Hermitian 矩阵正定等价于特征值都大于零

④ 可对角化：Hermitian 矩阵可以酉对角化 ($GFG^{-1}=GFG^H=D$)

⑤ 度量矩阵：$F\longmapsto AFA^*$

Theorem. 40 (Hermitian 型分解):

设 $f:V\times V\longrightarrow\mathbb{C}$ 为 Hermitain 型，则存在唯一的实对称双线性 g 和唯一的实斜对称双线性型 h，使得：

$$① \ f(u,v)=g(u,v)+ih(u,v)\iff F=G+iH$$

$$② \ f(u,v)=g(u,v)+ig(u,iv)$$

G 为 g 对应的实对称矩阵，H 为 h 对应的实斜对称矩阵。①②中的 g 是同一个，且反之也成立，即两者一一对应（有待考察）。由 Hermitian 型关于实对称/实斜对称的分解易证，详略。

推论：记全体 Hermitian 矩阵为 $M=\{F\in M_{n\times n}(\mathbb{C})\mid F=F^H\}$，全体实对称矩阵为 $R=\{G\in M_{n\times n}(\mathbb{R})\mid G=G^T\}$，则：

$$M\cong R\iff \text{Hermitian 型 } f \text{ 与实对称双线性型 } g \text{ 一一对应}$$

Theorem. 41 (Hermitian 型正定等价条件):

设 Hermitian 型 $f=g+ih$ 在基 $\vec{u}_0$ 下的矩阵为 $F=G+iH$，则：

$$f,F\text{ 正定}\iff g,G\text{ 正定}\iff \hat{G}=\begin{bmatrix} G & H\\ H^T & G\end{bmatrix}\text{ 正定}$$

Hermitian 型 f 称为正定的如果 $f(u,u)>0,\ \forall\,0\neq u\in V$，其中 $f(u,u)$ 必属于 $\mathbb{R}$。

酉空间 (Unitary Space)：

一个$\mathbb{C}$ 上的线性空间 V 称为酉空间如果它带有正定的 Hermitian 型$f: V\times V \longrightarrow \mathbb{C}$

$$f:(u,v)\longmapsto \langle\cdot \mid \cdot\rangle$$

称上面的内积为酉内积，并且有相关概念：

① 模/长度：$\|u\|=\sqrt{\langle u\mid u\rangle}$

② 复绝对值：$|a|=\sqrt{a\cdot \bar{a}}$，$a\in\mathbb{C}$

③ 正交：$u\perp v \Longleftrightarrow \langle u\mid u\rangle=0$

④ 夹角：$\theta=\frac{\langle u|u\rangle}{\|u\|\cdot\|v\|}$

⑤ 单位：$\|u\|=1$

⑥ 标准正交：一组正交向量 $\{v_1,...v_r\}$ 称为标准的如果 v_i 是单位的，$i=1,...,r$。

Theorem. 42 (酉空间中的定理)：
回想欧空间中出现的定理，很多在有限维酉空间中仍成立，如下：

① Cauchy-Schwarz Inequality:

$$|(u\mid v)|\le\|u\|\cdot\|v\|\ ,\ \text{仅线性相关时取等}\Longrightarrow \|u\pm v\|\le\|u\|+\|v\|$$

② 施密特正交化：

$$\text{设 } 0\ne v_0, v\in V,\ u=v_0-\frac{\langle v_0\mid v\rangle}{\|v\|}v,\ \text{则 } u\perp v\Longrightarrow \text{任意酉空间存在标准正交基}$$

③ 任意子空间与其正交补构成直和：

$$V=U\oplus U^{\perp}\Longrightarrow\ ④⑤⑥$$

④ 标准基扩充：任意标准正交组可扩充为 V 的一组标准正交基。

⑤ 向量的基表示:

$$\text{设 } \vec{w}_0 \text{ 为 } V \text{ 的一组标准正交基，则: } v=\sum_{i=1}^{n}\langle v\mid w_i\rangle w_i$$

⑥ 帕塞瓦尔恒等式 (Parseval’s Identity):

$$\sum_{i=1}^{n}\langle u\mid w_i\rangle\langle w_i\mid v\rangle=\langle u\mid v\rangle$$

⑦ 自伴随算子 (Hermitian 算子): φ 为 V 的自伴随线性算子 (Hermitian 算子) 如果

$$\varphi^*=\varphi\Longleftrightarrow V=\operatorname{Im}\varphi\oplus\ker\varphi\Longleftrightarrow \varphi\ \text{在某组标准正交基下的矩阵是 Hermitian 矩阵}$$

与欧类似，酉空间中的伴随算子 φ^* 定义为: $\langle\varphi^*(u)\mid v\rangle=\langle u\mid\varphi(v)\rangle$，这里的 * 表示 ^{H}，容易验证无论欧 or 酉，任意标准正交基 $\vec{w}_0$ 下的矩阵：$M_{\varphi^*,\vec{w}_0}=M^*_{\varphi,\vec{w}_0}$。

酉算子、酉矩阵：

线性算子 φ 称为酉的如果：

$$\langle\varphi(u)\mid\varphi(v)\rangle=\langle u\mid v\rangle\ ,\ \ \forall\, u,v\in V$$

酉算子 φ 在标准正交基下的矩阵 $M_{\varphi,\vec{w}_0}$ 称为酉矩阵，构成酉群：

$$U_n(\mathbb{C})=\left\{A\mid AA^H=I_n\right\}$$
$$A\text{ 为酉矩阵}\iff AA^*=I_n\iff A^{-1}=A^*$$
$$\varphi\text{ 为酉算子}\iff \varphi\varphi^*=e_V\iff\varphi^{-1}=\varphi^*$$

$\iff A$ 是标准正交基之间的转换矩阵。

全体酉算子记为 $U_o(V)\cong U_n(\mathbb{C})$。

(可逆) 矩阵群 $GL_n(\mathbb{C})$，特殊线性群 $SL_n(\mathbb{R})=\{A\mid\det A=1\}$，正交群 $O_n(\mathbb{R})=\left\{A\mid AA^T=I_n\right\}$，正常正交群 $SO_n(\mathbb{R})=\left\{A\mid AA^T=I_n,\ \det A=1\right\}$。且容易验证 $SO_n(\mathbb{R})\subseteq O_n(\mathbb{R})\cong U_n(\mathbb{R})\subseteq U_n(\mathbb{C})$。

Theorem. 43 (欧酉矩阵、转换矩阵、标准正交基的等价性):
无论是欧空间还是酉空间，设算子 φ 在某组标准正交基下的矩阵为 A，则有：

$$A\text{ 是欧 (酉) 矩阵}\iff A\text{ 是某两组标准正交基间的转换矩阵}\iff A\text{ 是一组标准正交基}$$

3.4 内积空间上的线性算子

正规算子：

$\mathbb{C}$ 上的算子 $\varphi\in\mathscr{L}(V)$ 称为正规算子如果：

$$\varphi\varphi^*=\varphi^*\varphi\iff AA^*=A^*A$$

正规矩阵的性质：

① 可对角化 (等价): A 是正规矩阵 $\iff\exists$ 酉矩阵 $G\ \ s.t.\ GAG^H=GAG^{-1}=D$

② 特征子空间：A 有 n 个不同的、两两正交的特征子空间。

③ A^H 的特征值：若 λ 是 A 的特征值，$u=\vec{x}\cdot\vec{v}_0$ 是对应的特征向量，则 $\overline{\lambda}$ 是 A^H 的特征值，$u=\vec{x}\cdot\vec{v}_0??$ 是对应的特征向量

酉对角化也就是在某组标准正交基下为对角阵，这是因为标准正交基、标准正交基之间的转换矩阵都是酉矩阵 (内积空间中的标准正交矩阵)。

当正规矩阵的特征值全部为实数时，是 Hermitian 矩阵 (对应自伴随算子、Hermitian 型)

当正规矩阵的特征值全部模为 1 时，是酉矩阵 (即复正交矩阵，对应酉算子)

自伴随算子与内积的关系：

设 $f: V \times V \longmapsto \mathbb{F}$ 是一个内积 (对称双线性型或 Hermitian 型)，记它在基下的度量矩阵为 F，则 $F = F^*$。定义 $\varphi \in \mathscr{L}(V)$ 使得 φ 在基下的矩阵为 F，则 φ 是自伴随算子。反之也成立。

因此，对同一个 Hermitian 矩阵 F，既可以把它看作一个自伴随算子 (Hermitian 算子)，也可以看作某个 Hermitian 型的度量矩阵 (但并不构成同构关系)。由此可以知道，Hermitian 矩阵所具有的性质也就是自伴随算子所具有的性质 (如特征值、可对角化等)。

Theorem. 44 (谱定理)：
设 φ 是酉空间 V 上的正规算子，及 φ 的所有不同特征值为 $\lambda_1, ..., \lambda_m$，则：

$$\text{存在唯一正交投影算子组 } \{\mathscr{P}_1, ..., \mathscr{P}_m\} \ s.t. \ \varphi = \sum_{i=1}^{m} \lambda_i \mathscr{P}_i$$

$$\text{进一步，存在 } f_1(x), ..., f_m(x) \in \mathbb{C}[x] \ s.t. \ \mathscr{P}_i = f_i(\varphi) \text{ 且 } f_i(\lambda_j) = \delta_{ij}$$

后一个推论表明：任意可由 $\mathscr{P}_i$ 线性表出的算子 ψ 都可表示为 $\psi = g(\varphi)$。

Theorem. 45 (内积同时对角化)：
为了体现对称双线性型和 Hermitian 型的统一性，我们将两者统称为内积空间上的准内积，其称为内积当且仅当它是正定的。

设 f_1, f_2 是内积空间上的两个准内积 (欧-对称双，酉-Hermi 型)，则：

$$f_1, f_2 \text{ 可同时对角化} \iff \text{其中有一个是正定的}$$

Theorem. 46 (半正定自伴随算子可开根)：
设 φ 是内积空间上一个半正定的自伴随算子，则：

$$\text{存在唯一的半正定自伴随算子 } \varphi_1 \text{ 使得 } \varphi_1^2 = \varphi\text{，记作 } \varphi_1 = \sqrt{\varphi}$$

用矩阵的语言：

$$\text{存在唯一的半正定自伴随矩阵 } A_1 \text{ 使得 } A_1^2 = A\text{，记作 } A_1 = \sqrt{A}$$

注：在以下的内容中，我们将酉空间、欧式空间统称为“内积空间”，Hermitian 型、对称双线性型统称为“准内积”，正定的 Hermitian 型、正定的对称双线性型统称为“内积”，Hermitian 算子、欧空间自伴随算子统称为自伴随算子，酉算子、欧空间自同构 (欧空间正交算子) 统称为正交算子，欧空间对称矩阵、酉空间 Hermitian 矩阵统称为自伴随矩阵。

Theorem. 47 (极化定理)：
设 $\varphi \in \mathscr{L}(V)$ 是内积空间上的线性算子，φ 在某组 (标准正交) 基下的矩阵为 A，则：
矩阵语言：

① 存在唯一的 (半) 正定矩阵 G 和正交矩阵 T 使得 $A = GT$

② A 是正规矩阵 $\iff GT = TG$

③ A 非退化 $\iff T$ 唯一

算子语言：

① 存在唯一的 (半) 正定算子 ζ 和正交算子 ψ 使得 $\varphi=\zeta\psi$

② φ 是正规算子 $\iff \zeta\psi=\psi\zeta$

③ φ 非退化 $\iff \psi$ 唯一

Theorem. 48 (实正规算子半对角化)：
设 V 是欧几里得内积空间且 $\varphi\in\mathscr{L}(V)$，记矩阵 $J[a_1,b_1]=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$，则：

$$\varphi \text{ 是正规算子 } \iff \text{在某组基下 } \varphi \text{ 的矩阵 } J_\varphi=\mathrm{diag}(c_1,...,c_r)\dotplus J[a_1,b_1]\dotplus\cdots\dotplus J[a_s,b_s]$$

$$\varphi\text{是正交算子 } \iff |c_i|=1,\ a_j^2+b_j^2=1,\ i=1,...,r,\ j=1,...,s$$

第 4 章　仿射空间与欧几里得点空间

4.1　仿射空间

放射空间基本概念：

一个 $\mathbb{F}$ 上的非空集合 $\mathbb{A}$ 称为仿射空间如果它与一个 $\mathbb{F}$ 上的向量空间 V 相伴，且存在从 $\mathbb{A}\times V$ 到 $\mathbb{A}$ 的映射 $f:(a,v)\longmapsto a+v,\ \forall\, a\in\mathbb{A},\ v\in V$ 满足：

① 右幺性：$a+0_V=a,\ \forall\, a\in\mathbb{A}$

② 加法结合律：$(a+v_1)+v_2=a+(v_1+v_2),\ \forall\, a\in\mathbb{A},\ v_1,v_2\in V$

③ 唯一性：$\forall\, a,b\in\mathbb{A},\ \exists\,!\, w\in V\ \ s.t.\ \ a+w=b$，记作 $w=\overrightarrow{ab}$

t_u 称为沿 u 对 $\mathbb{A}$ 的平移，且 $\mathbb{A}^\sharp=\{t_u\mid u\in V\}$ 构成群，同构于加法群 V，即 $\mathbb{A}^\sharp\cong V$。

此外，$\overrightarrow{ab}=-\overrightarrow{ba},\ \overrightarrow{aa}=0,\ \overrightarrow{ab}+\overrightarrow{bc}=\overrightarrow{ac}$

Theorem. 49（仿射空间同构）:
设 $\mathbb{A}$ 和 $\mathbb{A}'$ 是 $\mathbb{F}$ 上的相伴向量空间分别为 V 和 V' 的仿射空间，则 $\mathbb{A}\cong\mathbb{A}'\Longleftrightarrow V\cong V'$。

仿射空间的坐标系：

① 定义：给定点 $\dot{o}\in\mathbb{A}$ 和 V 的一组基 $\{v_1,...,v_n\}$，称 $\{\dot{o};v_1,...,v_n\}$ 为仿射空间 $(\mathbb{A},V)$ 的一个坐标系。

② 坐标系变换：对于两个仿射空间坐标系 $O=\{\dot{o};\vec{u}_0\}$ 和 $O'=\{\dot{o}';\vec{v}_0\}$，设点 $\dot{p},\ \dot{o}'$ 在 O 下的坐标分别为 $\vec{x},\ \vec{o}$，且 $\vec{v}=A\vec{u}$，则 $\dot{p}$ 在新坐标系 O' 下的坐标：

$$\vec{y}=(\vec{x}-\vec{o})A^{-1}$$

等价条件：

仿射子空间：

设 $(\mathbb{A},V)$ 是仿射空间而 U 是 V 的子空间，定义 $\mathbb{A}$ 的仿射子空间 $\Pi(\dot{p},U)=\dot{p}+U$，则它是以 U 为伴随空间的仿射空间。若 $\dim U=m<+\infty$，称 $\Pi(\dot{p},U)$ 为 $\mathbb{A}$ 的 m 维平面。

特别地，若 $\dim U=1$，则 $\forall\, 0\neq\overrightarrow{pq}\in\mathbb{A}$，有 $U=\mathbb{F}\overrightarrow{pq}$，也即 $\Pi(\dot{p},U)=\{\dot{p}+t\overrightarrow{pq}\mid t\in\mathbb{F}\}$

等价条件：设 $\mathbb{F}$ 的特征不为 2。Π 为仿射空间 $\mathbb{A}$ 的一个子集，则：

$$\Pi\text{ 是仿射子空间}\Longleftrightarrow\Pi\text{上任意两点的直线都在 }\Pi\text{ 内}$$

推论：仿射子空间的交仍是子空间，且 $\Pi(U_1)\cap\Pi(U_2)=\Pi(U_1\cap U_2)$。

仿射包络：

设 X 为 $\mathbb{A}$ 的一个子空间 (不一定是仿射子空间!)，定义 $\mathbb{A}$ 关于 X 的仿射包络：

$$A(X)=\left\{\dot{p}+\text{Span}\{\overrightarrow{\dot{p}\dot{q}}\mid\forall\,\dot{q}\in X\}\mid\forall\dot{p}\in X\right\}$$

容易验证它和 $\dot{p}$ 或 $\dot{q}$ 的选取无关，且构成 $\mathbb{A}$ 的仿射子空间。

如果 X 只有一个点，则 $A(X)=X$ 是 0 维的仿射子空间。X 有两个点时，$A(X)$ 是过他们的直线。X 有三个点时，$A(X)$ 是由该直线和点确定的平面。

① 仿射无关：若 $X=\{\dot{p}_0,\dot{p}_1,...,\dot{p}_m\}$ 且 $\dim\mathbb{A}(X)=m$，则称 X 仿射无关。这等价于 $\{\overrightarrow{\dot{p}_0\dot{p}_1},...,\overrightarrow{\dot{p}_0\dot{p}_m}\}$ 线性无关

② 重心组合：设 $\dot{p}_0,...,\dot{p}_m\in\mathbb{A}$ 是任意 $m+1$ 个点 (可以相同)，称 $\sum_{i=0}^m a_i\dot{p}_i\ \ s.t.\ \ \sum_{i=0}^m a_i=1$ 为 $\{\dot{p}_0,...,\dot{p}_m\}$ 的重心组合。(上面写法省略了原点 $o'\dot{o}$，因为重心组合的结果与原点选取无关)

③ 仿射映射：映射 $\varphi:\ \mathbb{A}\longrightarrow\mathbb{A}'$ 为仿射映射等价于它保持重心组合，也即：

$$\varphi\left(\sum_{i=0}^m a_i\dot{p}_i\right)=\sum_{i=0}^m a_i\varphi(\dot{p}_i)$$

仿射映射：

线性映射 $\psi:\mathbb{A}\longmapsto\mathbb{A}'$ 称为仿射映射如果满足下面任意一条命题 (第一条为定义，其它为等价条件)

① 定义：

$$\psi(\dot{p}+v)=\psi(\dot{p})+(D\psi)(v)\ \ ,\forall\,\dot{p}\in\mathbb{A},\ v\in V$$

② 保持向量：

$$(D\psi)(\overrightarrow{\dot{p}\dot{q}})=\overrightarrow{\psi(\dot{p})\psi(\dot{q})}$$

③ 保持重心：

$$\psi\left(\sum_{i=0}^m a_i\dot{p}_i\right)=\sum_{i=0}^m a_i\psi(\dot{p}_i)$$

其它性质：

①

②

③

仿射线性映射：

从 $\mathbb{A}$ 到 $\mathbb{F}$ 的仿射映射 φ 称为仿射线性映射。全体

性质：取 $\mathbb{A}$ 仿射无关的 $n+1$ 个点 $\{\dot{p}_1,\dot{p}_2,\}$，它可生成 V 的一组基，且集合

$$\left\{\dot{p}=\dot{p}_0+a_1\overrightarrow{\dot{p}_0\dot{p}_1}+\cdots+a_n\overrightarrow{\dot{p}_0\dot{p}_n}\in\mathbb{A}\mid\varphi(\sum_{i=0}^m a_i\dot{p}_i)=0\right\}$$

构成 $\mathbb{A}$ 的超平面 ($n-1$ 维)，也即 $\varphi^{-1}(0)=ker_{\varphi}$ 构成 $\mathbb{A}$ 的超平面 ($n-1$ 维)。

此处有一个 $\Pi=\bigcap_{i=1}^{n-r}\varphi_i^{-1}(0)$ 的结论，详见讲义 P104。

平行、相交、交错：

① 平行：设 (Π_1,U_1)，(Π_2,U_2) 是 $\mathbb{A}$ 的两个仿射子空间，我们称 Π_1 与 Π_2 平行如果 $U_1\subseteq U_2$。

此时，若 $\Pi_1\cap\Pi_2\neq\emptyset$，则 $\Pi_1\subseteq\Pi_2$；若 $\Pi_1\cap\Pi_2=\emptyset$，考虑它们并集的仿射包络，有 $V(\Pi_1\cup\Pi_2)=\mathbb{F}\overrightarrow{\dot{p}\dot{q}}+U'\Longrightarrow\dim A(\Pi_1\cup\Pi_2)=\dim\Pi_2+1$。

② 相交：$\Pi_1\cap\Pi_2\neq\emptyset$ 且不平行。

③ 交错：既不平行也不相交

4.2 欧几里得点空间

欧几里得点空间：

仿射空间 $(\mathbb{A},V)$ 称为欧几里得点空间如果 $(V,(\cdot\mid\cdot))$ 构成欧内积空间。点空间中有相关定义：

① 距离：$\rho(\dot{p},\dot{q})=\|\overrightarrow{\dot{p}\dot{q}}\|$

② 线段：$\dot{p}\dot{q}=\{\dot{p}+t\overrightarrow{\dot{p}\dot{q}}\mid t\in[0,1]\}$

③ 线段长度：$|\dot{p}\dot{q}|=\|\overrightarrow{\dot{p}\dot{q}}\|=\rho(\dot{p},\dot{q})$

④ 夹角：$\cos\theta=\frac{(\overrightarrow{\dot{p}\dot{q}}|\overrightarrow{\dot{r}\dot{s}})}{\|\overrightarrow{\dot{p}\dot{q}}\|\|\overrightarrow{\dot{r}\dot{s}}\|}$

其中 $\Pi_1=\dot{p}+\mathbb{R}\overrightarrow{\dot{p}\dot{q}}$，$\Pi_1=\dot{r}+\mathbb{R}\overrightarrow{\dot{r}\dot{s}}$ 是 $\mathbb{A}$ 中的两条直线

Theorem. 50 (仿射空间的距离)：
设 $\Pi_1=\Pi_1(\dot{o}_1,U),\Pi_2=\Pi_2(\dot{o}_2,U)$ 是 $\mathbb{A}$ 中的两个不相交的仿射子空间，$\{u_1,...,u_m\}$ 是 $U+V$ 的一组正交基，则有：

$$\rho(\Pi_1,\Pi_2)=\left\|\overrightarrow{\dot{p}\dot{q}}-\sum_{i=1}^{m}\frac{(\overrightarrow{\dot{p}\dot{q}}\mid u_i)}{\|u_i\|^2}u_i\right\|=\left\|\overrightarrow{\dot{p}\dot{q}}-\frac{(\overrightarrow{\dot{p}\dot{q}}\mid u_1)}{(u_1\mid u_1)}u_1-\cdots-\frac{(\overrightarrow{\dot{p}\dot{q}}\mid u_1)}{(u_m\mid u_m)}u_m\right\|$$

法向量：

Theorem. 51 (点到超平面的距离)：
设 $\dot{p}=[\alpha_1,...,\alpha_n]$ 是 $\dot{\mathbb{R}}^n$ 中一点，$\Pi=\{[x_1,...,x_n]\in\mathbb{A}\mid a_1x_1+\cdots+a_nx_n+b=0\}$ 是 $\mathbb{A}$ 的超平面，则：

$$\rho(\dot{p},\Pi)=\frac{|a_1\alpha_1+\cdots+a_n\alpha_n+b|}{\sqrt{a_1^2+\cdots+a_n^2}}$$

受课时所限，课程跳过了下一节“群与几何”，第五章“常见曲面”，直接进入第六章“张量”

第 5 章　常见曲面

第 6 章　张量

6.1