高等代数 II
Advanced Linear Algebra

龚舒凯
中国人民大学
Renmin University of China
School of Applied Economics/ School of Statistics
shukai_gong@ruc.edu.cn

目录

1 多项式

1.1 最大公因式

1.1.1 最大公因式表示定理

存在 $u(x), v(x)$ 使得 $d(x) = u(x)f(x) + v(x)g(x)$

- **（题）** $f(x), g(x) \neq 0$，若存在 $u(x), v(x)$ 使得 $(f(x),\, g(x)) = f(x)u(x) + g(x)v(x)$，则：$(u(x), v(x)) = 1$
- **（题）** 设 $(f(x), g(x)) = 1$，$\partial f(x)$,$\partial g(x)$ 均大于 0，可取 $u(x), v(x)$ 满足 $\partial u < \partial g$，$\partial v < \partial f$ 使得

$$f(x)u(x) + g(x)v(x) = 1$$

且 $u(x), v(x)$ 是唯一的一对多项式。

1.1.2 最大公因式的判定定理

法一：$d(x)$ 是公因式 $+d(x)$ 是最大公因式（用 $\varphi(x)$ 做）
法二：$d(x)$ 是公因式 + 存在 $u(x), v(x)$ 使得 $d(x) = u(x)f(x) + v(x)g(x)$

- **（题）** 对任意 $n \in \mathbf{Z}^+$，有 $(f(x), g(x))^n = (f^n(x), g^n(x))$

1.1.3 互素判定

法一：存在 $u(x), v(x)$ 使得 $1 = u(x)f(x) + v(x)g(x)$（充要条件）
法二：设 $(f(x), g(x)) = d(x)$ 证明 $d(x) = 1$

- **（题）** 非零多项式 $f(x), g(x) \in \mathbf{P}[\mathbf{x}]$ 不互素的充分必要条件是：存在 $u(x), v(x) \neq 0$ 满足 $f(x)u(x) = g(x)v(x)$ 且 $\partial u(x) < \partial g(x)$，$\partial v(x) < \partial f(x)$

1.1.4 互素性质

1. 若 $(f(x), g(x)) = 1$，且 $f(x)|g(x)h(x)$，则 $f(x)|h(x)$
2. 若 $f(x)|h(x)$，$g(x)|h(x)$ 且 $(f(x), g(x)) = 1$，则 $f(x)g(x)|h(x)$

- **（题）** 若 $(f(x), g(x)) = 1$，则对任意 $m \in \mathbf{Z}^+$，有 $(f(x^m), g(x^m)) = 1$
- **（题）** 若 $(f(x), g(x)) = 1$，则对任意 $m, n \in \mathbf{Z}^+$，有 $(f^m(x), g^n(x)) = 1$

1.1.5 带余除法、辗转相除、综合除法

1. 带余除法的应用：（1）求商式和余式（2）判断整除
2. 辗转相除的应用：（1）求最大公因式 **（最大公因式是除尽前的上一个余式** $r_s(x)$**）**（2）判断互素（辗转相除到最后，最大公因式是不是 1？是的话就互素）

- **（题）** $x^m - 1$ 和 $x^n - 1$ 的最大公因式

3. 综合除法的应用：（1）用一次多项式去除多项式的商式与余式（2）多项式按方幂展开（3）判断根（用 $x-a$ 综合除法最后余数是 0，则 a 是根）（4）

1.2 因式分解定理

1.2.1 因式分解唯一性定理

1. 数域 $\mathbf{P}$ 上次数大于 0 的多项式 $f(x)$ 都可以**唯一分解为 $\mathbf{P}$** 上的一些不可约多项式的乘积。
2. 数域 $\mathbf{P}$ 上次数 ≥ 1 的任一多项式 $f(x)$ 都能有**标准分解式**：

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)...p_s^{r_s}(x),\ (r_i\geq 1,\ i=1,2,...,s)$$

- **（题）** 如果 $f(x),g(x)\in\mathbf{P}[\mathbf{x}]$，则
$$\frac{f(x)g(x)}{(f(x),g(x))}=[f(x),g(x)]$$
- **（题）** $g(x)|f(x)$ 的充分必要条件是 $g^2(x)|f^2(x)$

1.2.2 不可约多项式

1. **不可约多项式** $p(x)$ **与任意多项式的** $f(x)$ **的关系只有两种**：$p(x)|f(x)$，或者 $p(x)$ 与 $f(x)$ 互素。
2. 数域 $\mathbf{P}$ 上，若对于任意两个多项式 $f(x),g(x)$ 有 $p(x)|f(x)g(x)$，则一定有 $p(x)|f(x)$ 或 $p(x)|g(x)$ **[注]：一种常见的方法是：通过 n 次多项式至多有 n 个根，和导出来的多项式有超过 n 个根来推矛盾。**

- **（题）** 对任意 $f(x),p(x)\in\mathbf{Q}[\mathbf{x}]$，$p(x)$ 在 $\mathbf{Q}$ 上不可约，且 $f(x)$ 与 $p(x)$ 有一个公共复根，则 $p(x)|f(x)$
- **（题）** 数域 $\mathbf{P}$ 上次数大于 0 的多项式 $f(x)$ 是某个不可约多项式 $p(x)$ 的方幂的充分必要条件是：对于任意 $g(x)\in\mathbf{P}[\mathbf{x}]$，或者 $(f(x),g(x))=1$，或者存在正整数使得 $f(x)|g^m(x)$
- **（题）** 设 $a_1,...a_n$ 为互异的整数，则 $f(x)=\prod_{i=1}^{n}(x-a_i)-1$ 在有理数域上不可约.

1.3 重因式

1.$p(x)$ 是 $f(x)$ 的重因式 $\iff$ $p(x)$ 是 $(f(x),f'(x))$ 的公因式

- 推论：找出所有 $f(x)$ 的重因式的方法：$\dfrac{f(x)}{(f(x),f'(x))}=p_1(x)...p_k(x)$
- 如果成立 $f(x)=\dfrac{f(x)}{(f(x),f'(x))}\cdot g(x)$，则 $g(x)$ 无重因式

2.$f(x)$ 没有重因式 $\iff$ $1=(f(x),f'(x))$
3.a 是 $f(x)$ 的 k 重根 $\longrightarrow$ a 是 $f'(x)$ 的 $k-1$ 重根

- **（题）** 首项系数为 1 的 n 次多项式 $f(x)$ 可以写成 $f(x)=(x-a)^n$ 的充分必要条件是 $f'(x)|f(x)$

1.4 多项式函数的根与重根

- **（作业）**设 $a_1,...,a_n$ 是 n 个不同的数，而 $F(x)=(x-a_1)...(x-a_n)$，证明：（1）$\sum\limits_{i=1}^{n}\dfrac{F(x)}{(x-a_i)F'(a_i)}=1$（2）任意多项式 $f(x)$ 用 $F(x)$ 所除的余式为 $\sum\limits_{i=1}^{n}\dfrac{f(a_i)F(x)}{(x-a_i)F'(a_i)}$
- **（作业）**如果 $f(x)|f(x^n)$，那么 $f(x)$ 的根只能是 0 或者单位根
- **（作业）**$1+x+\dfrac{x^2}{2!}+...+\dfrac{x^n}{n!}$ 没有重根
- **（作业）**如果 $(x^2+x+1)|f_1(x^3)+xf_2(x^3)$，那么 $(x-1)|f_1(x),(x-1)|f_2(x)$

1.5 多项式根与系数的关系

设 $x_1,...,x_n$ 是 n 次多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+...+a_1x+a_0$ 的 n 个根，则根与系数的关系为：

$$\begin{aligned}
&\sum_i x_i=(-1)^1\frac{a_{n-1}}{a_n}\\
&\sum_{i<j} x_ix_j=(-1)^2\frac{a_{n-2}}{a_n}\\
&...\\
&\sum_{i_1<i_2<...<i_{n-1}} x_{i_1}x_{i_2}...x_{i_{n-1}}=(-1)^{n-1}\frac{a_1}{a_n}\\
&x_1...x_n=(-1)^n\frac{a_0}{a_n}
\end{aligned}$$

1.6 有理系数多项式

1.6.1 有理系数多项式因式分解

1. 本原多项式：$f(x)$ 的系数互素
2.**Gauss 引理：**两个本原多项式的乘积还是本原多项式
3. 如果整系数多项式能分解为两个次数较低的**有理系数多项式**的乘积，那么它可以分解为两个次数较低的**整系数多项式**的乘积
4. 已知整系数多项式 $f(x),g(x)$ 为整系数多项式，且 $g(x)$ 本原，如果 $f(x)=g(x)h(x)$，则 $h(x)$ 为整系数多项式.

- **（题）**设 $f(x)$ 是整系数多项式，如果 $f(1),f(0)$ 都是奇数，证明 $f(x)$ 没有整数根

1.6.2 有理根问题

1.**Eisenstein 判别法：**$f(x)=a_nx^n+a_{n-1}x^{n-1}+...+a_1x+a_0$ 是整系数多项式，如果存在素数 p 满足：

- $p\nmid a_n$
- $p|a_i\ (i=\mathbf{0},1,...,n-1))$
- $p^2\nmid a_0$

则 $f(x)$ 不可约。

1'.**Eisenstein 判别法的变形**：设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+...+a_1x+a_0$，若 $g(x)=x^nf(\frac{1}{x})$ 不可约，则 $f(x)$ 不可约，**只需对 $g(x)$（即系数恰好颠倒的多项式）使用 Eisenstein 判别法即可。**

2. 设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+...+a_1x+a_0$ 是整系数多项式，而 $\frac{r}{s}$ 是其一个有理根且 r,s 互素，**那么必有 $s|a_n,r|a_0$**.

- **（题）**$p_1,...,p_s$ 是 s 个互不相同的素数，$n\in\mathbf{Z}^+$，则 $\sqrt[n]{p_1p_2...p_s}$ 为无理数。

2 线性空间

2.1 常用的基与空间

- 无限维线性空间：多项式空间 $V=\mathbf{P}[\mathbf{x}]$
- **无限维线性空间：连续函数空间** $V=C[a,b]$：例如

  （题）$1,\sin x,\sin 2x,...,\sin nx$ 线性无关（证明：求两次导）

  （题）$1,\cos x,\cos 2x,...,\cos nx$ 线性无关（证明：求两次导）

  （题）$1,\sin x,\cos x,\sin 2x,\cos 2x,...,\sin nx,\cos nx$ 线性无关（证明：求两次导）

  （21-22 复旦线代期中）$e^{a_1}x,e^{a_2}x,...,e^{a_n}x$ 线性无关（证明：Vandermonde 行列式）
- 有限维线性空间：$\mathbb{C}_{\mathbb{C}}$、$\mathbb{C}_{\mathbb{R}}$、$\mathbf{P}^n$

2.2 基变换与坐标变换

2.2.1 过渡矩阵的求法

n 维向量空间 $\mathbf{P}^n$ 中，设方阵 $B_{n\times n}=\left[\xi_1,\xi_2,...,\xi_n\right]$，$C_{n\times n}=\left[\eta_1,\eta_2,...,\eta_n\right]$，则基 $\xi_1,...,\xi_n$ 到 $\eta_1,...,\eta_n$ 的过渡矩阵是 $A=B^{-1}C$

具体实践方法是，排一个矩阵 $\left[\xi_1,\xi_2,...,\xi_n,\eta_1,\eta_2,...,\eta_n\right]\Longrightarrow$(行变换)$\Longrightarrow$ 消成 $\left[\varepsilon_1,\varepsilon_2,...,\varepsilon_n,\sigma_1,...,\sigma_n\right]=\left[\boldsymbol{E},\boldsymbol{B^{-1}C}\right]$

2.2.2 坐标变换公式

设 A 为过渡矩阵，如果老基$\left[\varepsilon_1,...,\varepsilon_n\right]$ 到新基$\left[\varepsilon_1',...,\varepsilon_n'\right]$ 的基变换是

$$\left[\varepsilon_1',...,\varepsilon_n'\right]=\left[\varepsilon_1,...,\varepsilon_n\right]\cdot A$$

那么老坐标$\left[x_1,...,x_n\right]^T$ 和新坐标$\left[y_1,...,y_n\right]^T$ 的关系是：

$$\left[x_1,...,x_n\right]^T=A\cdot\left[y_1,...,y_n\right]^T$$
$$\left[y_1,...,y_n\right]^T=A^{-1}\cdot\left[x_1,...,x_n\right]^T$$

2.3 子空间

2.3.1 子空间判定准则

1. 线性空间 V 要求满足八条法则：交换律、结合律、零元素、负元素、一元素、乘法分配 *2、乘法结合
2. 子空间只要求 V 中的加法和数乘运算封闭

2.3.2 一些子空间的标识

设 A 是数域 $\mathbf{P}$ 上的 $n\times m$ 矩阵，其行向量 $\alpha_1,...\alpha_n$，列向量 $\beta_1,...,\beta_m$，则：
（1）**行空间**：$\mathcal{R}(A^T)=L(\alpha_1,...\alpha_n)$
（2）**列空间**：$\mathcal{R}(A)=L(\beta_1,...\beta_m)$
（3）**零化空间**：$\mathcal{N}(A)=\{x\in\mathbf{P^m}|Ax=0\}$

2.3.3 子空间的性质

1.$W\subset V\Rightarrow \dim W\le \dim V$.

2.$L(\alpha_1,...,\alpha_s)=L(\beta_1,...,\alpha_t)\iff \beta_1,...,\beta_t$ 与 $\alpha_1,...,\alpha_s$ 等价.

3.$\dim L(\alpha_1,...,\alpha_s)=r(\alpha_1,...,\alpha_s)$

4. 基扩张定理：
$W\subset V$，$\alpha_1,...,\alpha_m$ 是 W 的一组基，又设 $e_1,...,e_n$ 是 V 的一组基，则可以从 $e_1,...,e_n$ 中选 $\boldsymbol{n-m}$ **个向量**，扩充到 $\alpha_1,...,\alpha_m$ 从而构成 V 的一组基.

- **（推论 1）** V 的任意线性无关组都可以扩充为 V 的一组基
- **（推论 2）** W 的子空间的基可以扩充为 W 的基

5. **（题）** 若 $V_1\subset V_2$，$\dim V_1=\dim V_2\ne+\infty$，则 $V_1=V_2$

6. **（题）** $L(\alpha_1,...,\alpha_s)+L(\beta_1,...,\beta_t)=L(\alpha_1,...,\alpha_s,\beta_1,...,\beta_t)$

- 若 $\alpha_1,...\alpha_s$ 是 V 的一组基，则 $V=L(\alpha_1)+...+L(\alpha_s)$

7. 设 $\alpha_1,...,\alpha_n$ 为 n 维线性空间 V 的一组基，A 为一个 $n\times s$ 矩阵，若 $(\beta_1,...,\beta_s)=(\alpha_1,...,\alpha_n)A$，则 $\dim L(\beta_1,...,\beta_s)=\mathrm{rank}A$

2.4 子空间的运算

2.4.1 子空间的交与和

1. 若 $V_i\subset V$，则：
（1）子空间的交：$\bigcap\limits_{i=1}^{n}V_i$ 也是子空间

（2）子空间的和：$\sum\limits_{i=1}^{n} V_i$ 也是子空间
（3）V_1+V_2 是包含 $V_1\cup V_2$ 的最小子空间

2.$V_1\cup V_2$ 不一定是子空间
（反例：$V=\mathbf{R}, V_1=\{(x,y)|x=0\}, V_2=\{(x,y)|y=0\}$，$(1,1)\in V_1\cup V_2$ 但 $(1,1)\notin V$）
3. 子空间只有**交换律**和**结合律**，没有分配律

4. 设 V_1,V_2,W 都是 V 的子空间，则：
（1）若 $W\subset V_1$，$W\subset V_2$，则 $W\subset V_1\cap V_2$
（2）若 $W\supset V_1$，$W\supset V_2$，则 $W\supset V_1+V_2$
[注]：怎么验证 $W_1+W_2=V$？，通常而言，都是先 $W_1+W_2\subset V$，再 $W_1+W_2\supset V$

- **（题）**数域 $\mathbb{F}$ 上 n 维线性空间 $\mathbb{F}^n$ 的任一子空间 U 都是数域 $\mathbb{F}$ 上你某个齐次线性方程组的解空间
- **（题）**设 $V_1,V_2,...,V_k$ 是线性空间 V 上的真子空间，则 $\bigcup\limits_{i=1}^{k} V_i\neq V$
- **（题）**设 V 为数域 $\mathbb{F}$ 上 n 维线性空间，则 V 的 r 维子空间有无数个（$1<r<n$）
- **（题）**V_1,V_2 是 V 的两个子空间，则 $V_1+V_2=V_1\cup V_2 \iff V_1\subset V_2$ 或 $V_1\supset V_2$

2.4.2 子空间维数计算

1. 维数公式：$\dim V_1+\dim V_2=\dim(V_1+V_2)+\dim(V_1\cap V_2)$

- **（推论）**在 n 维线性空间 V 中，$V_1,V_2\subset V$，$\dim V_1+\dim V_2>n$，则 $V_1\cap V_2$ 有非零公共向量

2. $\boxed{\textbf{求子空间 } V_1\cap V_2 \textbf{ 的维数}}$：设 V_1 的基为 $\alpha_1,...,\alpha_s$，V_2 的基为 $\beta_1,...,\beta_t$，则 $V_1\cap V_2$ 的基就是

$$\alpha_1x_1+...+\alpha_sx_s+\beta_1x_{s+1}+...+\beta_tx_{s+t}=0$$

的解空间的基（基础解系）
3. $\boxed{\textbf{求子空间 } V_1+V_2 \textbf{ 的维数}}$：设 V_1 的基为 $\alpha_1,...,\alpha_s$, V_2 的基为 $\beta_1,...,\beta_t$, 则 V_1+V_2 的基就是 $L(\alpha_1,...,\alpha_s,\beta_1,...,\beta_t)$ 的基。**把** $\alpha_1,...,\alpha_s,\beta_1,...,\beta_t$ **竖过来排成矩阵作初等行变换，他们的极大无关组就是基。**

2.5 直和

1. 直和：和空间 V_1+V_2 中每一个向量的分解方式**唯一**：一个来自 V_1，一个来自 V_2。
直和的例子有：

- **（题）**V 为全体 n 阶矩阵，V_1 为全体 n 阶对称矩阵，V_2 为全体 n 阶反称矩阵，则 $V=V_1\oplus V_2$
- **（题）**$A^2=A$，则 $P^{n\times n}=\mathcal{N}(A)\oplus\mathcal{R}(A)$
- **（题）**$V=\mathbf{P}[x]$，$W_1=\mathbf{P}$，$W_2=\{xf(x)|f(x)\in V\}$，则 $V=W_1\oplus W_2$

- **(题)** 设 $A \in \mathbf{M_n(P)}$, $f(x), g(x) \in \mathbf{P[x]}$, 如果 $(f_1(x), g(x)) = 1$, $A = f(M)$, $B = g(M)$, 则 $\mathcal{N}(AB) = \mathcal{N}(A) \oplus \mathcal{N}(B)$

2. **直和补**：W 是 n 维线性空间 V 的子空间，那么一定存在 V 的子空间 U 使得 $V = W \oplus U$

[注]：取 V 的一组基 $\alpha_1, ..., \alpha_s$，把它扩成 V 的一组基 $\alpha_1, ..., \alpha_s, ..., \alpha_n$，那么取 $U = L(\alpha_{m+1}, ..., \alpha_n)$

[注 2]：通常选取 $\varepsilon_1, ..., \varepsilon_n$ 来扩基。检验扩充后的基是否线性无关的方法是：竖过来排成一个方阵 A，看 $\det A$ 是不是 $\neq 0$

2.5.1 直和的判定定理

1.$V_1 + V_2$ 是直和 $\iff$ 零元唯一分解，即 $\alpha_1 + \alpha_2 = 0$，$\alpha_1 \in V_1$，$\alpha_2 \in V_2$ 时必有 $\alpha_1 = 0, \alpha_2 = 0$

2.$V_1 + V_2$ 是直和 $\iff$ $V_1 \cap V_2 = 0$

3. （当子空间是有限维的时候）$V_1 + V_2$ 是直和 $\iff$ $\dim V_1 + \dim V_2 = \dim(V_1 + V_2)$

- 要证 $W = V_1 \oplus V_2$，需证（1）$W = V_1 + V_2$（2）$V_1 + V_2$ 是直和

- 要证 $V = V_1 \oplus V_2$，需证（1）$V = V_1 + V_2$（其实只需证 $V \subset V_1 + V_2$，$V_1 + V_2$ 必是子空间）（2）$V_1 + V_2$ 是直和

2.5.2 多个子空间的直和

1. **多个子空间的直和**：和空间 $V_1 + ... + V_s$ 中每一个向量 α 的分解方式**唯一**：$\alpha = \alpha_1 + ... + \alpha_s$，其中 $\alpha_i \in V_i$

2.$V_1 \oplus ... \oplus V_s$ 的充要条件：

- 零元唯一分解

- $\boldsymbol{V_i \cap \sum\limits_{j \neq i} V_j = 0}$（注：这个条件比两两相交等于 0 强）

- $\dim(V_1 + ... + V_s) = \dim V_1 + ... + \dim V_s$

- **（复旦高代）** $V_1, ..., V_s$ 的一组基可以拼成 V 的一组基

  设 $\varepsilon_1, ...\varepsilon_n$ 和 $\eta_1, ..., \eta_n$ 分别是 V_1, V_2 的一组基，则 $V_1 \oplus V_2 \Longleftrightarrow \varepsilon_1, ...\varepsilon_n, \eta_1, ..., \eta_n$ 线性无关

[注]：第二条充要条件，不能两两相交的反例：设 $\mathbf{R}^2$ 上 $V_1 = L[(1,0)]$，$V_2 = L[(0,1)]$，$V_3 = L[(1,1)]$，则 $V_i \cap V_j = 0$ $(i \neq j, i, j = 1, 2, 3)$，但 $V_1 + V_2 + V_3$ 不是直和，因为分解不唯一：

$$(x, x) = x(1,0) + x(0,1) + 0(1,1) = 0(1,0) + 0(0,1) + x(1,1)$$

- **（题）** 如果 $V = V_1 \oplus V_2$，$V_1 = V_{11} \oplus V_{22}$，则 $V = V_{11} \oplus V_{22} \oplus V_2$

3. 设 $\alpha_1, ..., \alpha_n$ 是 n 维线性空间 V 的一组基，那么 $V = L(\alpha_1) \oplus L(\alpha_2) \oplus ... \oplus V(\alpha_n)$

2.6 同构

线性同构的概念与构造 1. 若 V 到 V' 是同构：一共四件事

- $\sigma : V \longmapsto V'$ 是一一对应（分别验证单射和满射）
- $\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta)$
- $\sigma(k\alpha) = k\sigma(\alpha)$

2. **重要的同构映射：线性空间到坐标的映射**
设 V 是数域 $\mathbb{F}$ 上的一个线性空间，则 $V \cong \mathbb{F}^n$，即存在同构线性映射使得

$$\begin{aligned}\sigma : &V \longmapsto \mathbb{F}^n \\ &\alpha \longmapsto (a_1, ..., a_n)^T\end{aligned}$$

其他常见的同构线性映射有：

- 对称方阵与上三角方阵同构
- $\mathbb{C}_{\mathbb{R}}$ 与 $\mathbb{R}^2$ 同构（复数到复平面）
- n 维多项式空间 $\mathbf{P}[\mathbf{x}]_n$ 与 $\mathbb{P}^n$ 同构

3. 构造线性空间到坐标映射的一般方法：任取 V 的一组基 $\varepsilon_1, ..., \varepsilon_n$，则可以找到一个同构线性映射 σ 使得：

$$\begin{gathered}\sigma : V \longmapsto \mathbb{F}^n \\ \alpha = a_1\varepsilon_1 + ... + a_n\varepsilon_n \longmapsto (a_1, ..., a_n)^T\end{gathered}$$

2.6.1 同构映射的性质

1.$\sigma(0) = 0, \sigma(-\alpha) = -\sigma(\alpha)$
2.$\sigma(k_1\alpha_1 + ... + k_n\alpha_n) = k_1\sigma(\alpha_1) + ... + k_n\sigma(\alpha_n)$
3.V 中 $\alpha_1, ...\alpha_n$ 线性相关（无关） $\iff \sigma(\alpha_1), ..., \sigma(\alpha_n)$ 线性相关（无关）
4.V_1 是 V 的子空间，则 $\sigma(V_1)$ 是 $\sigma(V)$ 的子空间，且 $\dim V_1 = \dim \sigma(V_1)$
5. **同构映射的乘积** $\sigma\tau$ 和**同构映射的逆映射** σ^{-1} 都是同构映射
6. 线性空间的同构是等价关系

2.6.2 线性同构判定定理

数域 $\mathbb{F}$ 上**有限维**线性空间 V_1, V_2 同构 $\iff \dim V_1 = \dim V_2$

- 因此**有限维**线性空间不可能和其真子空间同构
- 但如果条件改为**无限维**线性空间 V，则 V 可能和其真子空间 V' 同构。例如：$V = \mathbf{P}[x]$，$V' = \{xf(x)|f(x) \in V\} \subsetneqq V$，但存在线性映射 $\sigma(f(x)) = xf(x)$ 使得 $V' = \sigma(V)$

2.6.3 同构映射的应用

将抽象向量转化为具体坐标解决

（1）设 V 中 $\alpha_1, ..., \alpha_s$ 在某个基下的坐标为 $\xi_1, ..., \xi_s$（其中 ξ_i 为 n 维向量），则

$\alpha_1, ..., \alpha_s$ 线性相关

$\Longleftrightarrow$ 存在不全为 0 的 k_i 使得 $k_1\alpha_1 + ... + k_s\alpha_s = 0$

$\Longleftrightarrow$ 齐次线性方程组 $x_1\xi_1 + ... + x_s\xi_s = 0$ 有非零解

换言之，向量与其坐标有着相同的线性相关性，我们关心向量的线性相关性，只需要把他们的坐标竖着排成一个方阵 A，然后关心这个方阵是否满秩即可。

3 线性变换

3.1 线性变换的概念

1. 线性变换：

$$\mathcal{A}: V \to V$$

满足对任意 $\alpha, \beta \in V$

$$\mathcal{A}(\alpha+\beta) = \mathcal{A}(\alpha) + \mathcal{A}(\beta)$$
$$\mathcal{A}(k\alpha) = k\mathcal{A}(\alpha)$$

几种常见的线性变换有：

- **零变换**：$\mathcal{O}(\alpha) = 0$
- **恒等变换**：$\mathcal{E}(\alpha) = \alpha$
- **数乘变换**：$\mathcal{K}(\alpha) = k\alpha$
- **投影变换**：设子空间 W 的一组基为 $\varepsilon_1, ..., \varepsilon_m$，将其扩充为 V 的一组基 $\varepsilon_1, ..., \varepsilon_m, ..., \varepsilon_n$，记投影变换为

$$\mathcal{A}(\varepsilon_i) = \begin{cases} \varepsilon_i & 1 \le i \le m \\ 0 & m+1 \le i \le n \end{cases}$$

  - **投影**表现为：对任意 $\alpha \in V$，$\alpha = k_1\varepsilon_1 + ... + k_n\varepsilon_n$，那么

$$\mathcal{A}(\alpha) = \mathcal{A}(k_1\varepsilon_1 + ... + k_n\varepsilon_n) = k_1\mathcal{A}(\varepsilon_1) + ... + k_n\mathcal{A}(\varepsilon_n) = k_1\varepsilon_1 + ... + k_m\varepsilon_m$$

    相当于把 V 中的向量 α 投影到子空间 W 中

2. 线性变换的性质

- $\mathcal{A}(0) = 0, \mathcal{A}(-\alpha) = -\mathcal{A}(\alpha)$
- $\mathcal{A}(k_1\varepsilon_1 + ... + k_n\varepsilon_n) = k_1\mathcal{A}(\varepsilon_1) + ... + k_n\mathcal{A}(\varepsilon_n)$
- **保线性相关性**：若 $\varepsilon_1, ..., \varepsilon_n$ 线性相关，则 $\mathcal{A}(\varepsilon_1), ..., \mathcal{A}(\varepsilon_n)$ 线性相关
- **不保线性无关**：若 $\varepsilon_1, ..., \varepsilon_n$ 线性无关，则 $\mathcal{A}(\varepsilon_1), ..., \mathcal{A}(\varepsilon_n)$ 不一定线性无关
  - 反例：取 $\mathcal{A} = \mathcal{O}$ 即可
  - 根本原因是：无法保证 $\mathcal{A}$ 是双射
- 设 $\varepsilon_1, ..., \varepsilon_n$ 是一组基，则 $\mathcal{A} = \mathcal{B} \iff \mathcal{A}(\varepsilon_i) = \mathcal{B}(\varepsilon_i)$
- 设 $\varepsilon_1, ..., \varepsilon_n$ 是一组基，对**任意一组向量 $\boldsymbol{\alpha_1, ..., \alpha_n}$，总存在线性变换** $\mathcal{A}$ 使得 $\mathcal{A}(\varepsilon_i) = \alpha_i$

3.2 线性变换的矩阵

3.2.1 对比：线性变换后的坐标与基变换后的坐标

设 $\boldsymbol{V}$ 是线性空间，$\boldsymbol{\varepsilon_1,\varepsilon_2,...,\varepsilon_n}$ 是 $\boldsymbol{V}$ 的一组基，$\boldsymbol{V}$ 的线性变换 $\mathcal{A}$ 在这组基下对应的矩阵为 $\boldsymbol{A}$,$\boldsymbol{V}$ 中向量 $\boldsymbol{\alpha}$ 在这组基下的坐标为 $\boldsymbol{X}=(x_1,x_2,...,x_n)^T$，$\mathcal{A}(\boldsymbol{\alpha})$ 在这组基下的坐标为 $\boldsymbol{X'}=(x_1,x_2,...,x_n)^T$

	基变换	线性变换
变换后基	$(\boldsymbol{\eta_1,...,\eta_n})=(\boldsymbol{\varepsilon_1,...,\varepsilon_n})\boldsymbol{P}$	仍为 $(\boldsymbol{\varepsilon_1,...,\varepsilon_n})$
变换后向量	$\boldsymbol{X'}=(\boldsymbol{\eta_1,...,\eta_n})\begin{pmatrix}\boldsymbol{x'_1}\\ \boldsymbol{...}\\ \boldsymbol{x'_n}\end{pmatrix}$	$\boldsymbol{X'}=\sum\limits_{i=1}^{n}x_i\mathcal{A}(\boldsymbol{\varepsilon_i})$
变换后坐标	$\boldsymbol{X'}=\boldsymbol{P}^{-1}\boldsymbol{X}$	$\boldsymbol{X'}=\boldsymbol{AX}$

3.2.2 相似矩阵

相似矩阵是矩阵等价的一个特例：矩阵相似是 $\boldsymbol{A}\sim\boldsymbol{B}\iff\exists\boldsymbol{P},\ \boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$，而矩阵等价是 $\boldsymbol{A}\cong\boldsymbol{B}$ $\iff\exists\boldsymbol{P},\boldsymbol{Q},\ \boldsymbol{B}=\boldsymbol{PAQ}$，显然，矩阵相似也是等价关系。

1. 过渡矩阵与相似的关系： $\boldsymbol{V}$ 的线性变换 $\boldsymbol{\sigma}$ 在 $\{\boldsymbol{e_1,...,e_n}\}$ 和 $\{\boldsymbol{f_1,...,f_n}\}$ 下的矩阵分别为 $\boldsymbol{A},\boldsymbol{B}$，则 $\boldsymbol{A}\sim\boldsymbol{B}$ 且 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$，其中 $\boldsymbol{P}$ 是 $\{\boldsymbol{e_1,...,e_n}\}$ 到 $\{\boldsymbol{f_1,...,f_n}\}$ 的过渡矩阵.

2. 相似矩阵的性质：

- 若 $\boldsymbol{A}\sim\boldsymbol{B}$，则 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{B})$，$\boldsymbol{A^T}\sim\boldsymbol{B^T}$，$\boldsymbol{A}^{-1}\sim\boldsymbol{B}^{-1}$，$|\boldsymbol{A}|=|\boldsymbol{B}|$，$\boldsymbol{A}+k\boldsymbol{E}\sim\boldsymbol{B}+k\boldsymbol{E}$
- $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}\Rightarrow\boldsymbol{B}^m=\boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}$
- $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}\Rightarrow f(\boldsymbol{B})=\boldsymbol{P}^{-1}f(\boldsymbol{A})\boldsymbol{P}$

3.3 特征值与特征向量

3.3.1 特征值、特征向量与特征子空间

1. 特征值与特征向量： 设 $\lambda\in\mathbb{P}$，若存在向量 $\xi\in V$，使得

$$\mathcal{A}(\xi)=\lambda\xi$$

则称 λ 为特征值，ξ 为**线性变换 $\mathcal{A}$ 的属于 λ 的**特征向量

由于线性变换 $\mathcal{A}$ 在给定基下的矩阵 $\boldsymbol{A}$ 唯一，定义矩阵 $\boldsymbol{A}$ 特征值 λ 和特征向量 xi 满足：

$$\boldsymbol{A}(\xi) = \lambda\xi$$

- **几何意义**：特征向量 ξ 经过线性变换 $\mathcal{A}$ 后仍在同一直线上
- **特征子空间**：对任意特征值 λ，属于 λ 的特征向量及零向量构成的集合 $V_\lambda = \{\xi : \mathcal{A}(\xi) = \lambda\xi\}$ 是 V 的一个子空间

3.3.2 特征值、特征向量的求法

1. 计算特征值：只需计算令**特征多项式**

$$\left|\lambda\boldsymbol{E} - \boldsymbol{A}\right| = 0$$

的**所有** λ 即可（可以有重根），其中 $\boldsymbol{A}$ 是线性变换 $\mathcal{A}$ 在某组基下的矩阵

2. 计算 $\mathcal{A}$ 的特征向量

- 解出所有特征值 $\lambda_i(i = 1, 2, ..., n)$
- 代回方程组得到 $(\lambda_i\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = 0$，解出其基础解系 $\boldsymbol{\eta_{i_1}, ..., \eta_{i_{r_i}}}$
  - 注 1：$\boldsymbol{\eta_{i_1}, ..., \eta_{i_{r_i}}}$ 也就是 V_{λ_i} 的一组基
  - 注 2：在复数域 $\mathbb{C}$ 上讨论时，全部特征根都是存在的
- 得到 V_{λ_i} 的一组基 $\xi_{i_j} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)\eta_{i_j}$ $(j = 1, 2, ..., r_i)$
- 线性组合 $k_1\xi_{i_1} + ... + k_{i_{r_i}}\xi_{i_{r_i}}$ 就是 $\mathcal{A}$ 的属于 λ_i 的特征向量

3.3.3 特征值、特征向量、特征多项式的性质

1. 计算后的特征值：设 λ 是 $\mathcal{A}$ 的特征值，ξ 是 $\mathcal{A}$ 的特征向量（对矩阵 $\boldsymbol{A}$ 完全同理）

- $k\mathcal{A}$ 的特征值为 $k\alpha$，$\mathcal{A} + k\mathcal{E}$ 的特征值为 $\lambda + k$，$\mathcal{A}^m$ 的特征值为 λ^m
- $f(\mathcal{A})$ 的特征值为 $f(\lambda)$，$\mathcal{A}^{-1}$ 的特征值为 λ^{-1}，$\mathcal{A}^*$ 的特征值为 $\dfrac{|A|}{\lambda}$，$\mathcal{A}^{\mathcal{T}}$ 的特征值为 λ

2. 特征多项式的性质

- $\boldsymbol{A}$ 与 $\boldsymbol{A^T}$ 有相同的特征多项式，即 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = |\lambda\boldsymbol{E} - \boldsymbol{A^T}|$
- $\boldsymbol{A} \sim \boldsymbol{B}$，则 $\boldsymbol{A, B}$ 具有相同的特征多项式，进而有相同的特征值
  - 这说明有限维空间上线性变换的特征值 λ，与所取的基没有关系（因为 $\mathcal{A}$ 在不同基下的矩阵式相似的），这种特征多项式可以记为 $f_{\mathcal{A}}(x)$

– **（逆命题）**有相同特征值的矩阵不一定是相似的, 反例：两者特征值均为 1，但不相似

$$A=\begin{bmatrix}1&1\\0&1\end{bmatrix},E$$

- 矩阵多项式的部分展开：

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda^n-(a_{11}+...+a_{nn})\lambda^{n-1}+...+(-1)^n|A|$$

  – $\mathrm{tr}(\boldsymbol{A})=\lambda_1+...+\lambda_n$
  – $|A|=\lambda_1...\lambda_n$（**$\boldsymbol{A}$ 可逆 $\iff$ 特征值全不为 0**）

***3. 迹 (Trace)**

- $\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{tr}(\boldsymbol{A})+\mathrm{tr}(\boldsymbol{B})$，$\mathrm{tr}(\boldsymbol{kA})=k\mathrm{tr}(\boldsymbol{A})$，$\mathrm{tr}(\boldsymbol{A^T})=\mathrm{tr}(\boldsymbol{A})$
- $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$，$\mathrm{tr}(\boldsymbol{ABC})=\mathrm{tr}(\boldsymbol{BAC})=\mathrm{tr}(\boldsymbol{CAB})$
- 若 $A\sim B$，则 $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{B})$

3.3.4 Hamilton-Cayley 定理及其应用

1.Hamilton-Cayley 定理

记 $f(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$，则 $f(\boldsymbol{A})=\boldsymbol{O}$

- $f(\mathcal{A})=\mathcal{O}$

2.Hamilton-Cayley 定理的应用

- **求逆矩阵**：若 $f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+...+a_1\lambda+a_0$，由 Hamilton-Cayley 定理，$f(\boldsymbol{A})=\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+...+a_1\boldsymbol{A}+a_0\boldsymbol{E}=\boldsymbol{O}$，那么

$$\boldsymbol{A}^{-1}=\frac{1}{a_0}(A^{n-1}+a_{n-1}\boldsymbol{A}^{n-2}+...+a_1\boldsymbol{E})$$

  也可以用此方法判断 $\boldsymbol{A}$ 是否可逆（特征多项式是否有常数项）

- **求矩阵方幂**：设特征多项式 $f(x)=x^n+a_{n-1}x^{n-1}+...+a_1x+a_0$，要求 x^m，根据带余除法易知

$$x^m=f(x)q(x)+(r_{n-1}x^{n-1}+...+r_1x+r_0)\tag{1}$$

  代入 $x=\boldsymbol{A}$，再由 $f(\boldsymbol{A})=\boldsymbol{O}$ 知

$$A^m=r_{n-1}\boldsymbol{A}^{n-1}+...+r_1\boldsymbol{A}+r_0E\tag{2}$$

  情况 1: 若特征值无重根，则把全部的 $\lambda_1,...,\lambda_n$ 带入 (1)，从而得到线性方程组

$$\begin{cases}\lambda_1^m=r_{n-1}\lambda_1^{n-1}+...+r_1\lambda_1+r_0\\ ...\\ \lambda_n^m=r_{n-1}\lambda_n^{n-1}+...+r_1\lambda_n+r_0\end{cases}$$

把 $(r_0, ..., r_{n-1})$ 全部代入回（2）即可得到 $\boldsymbol{A}^m$

情况 2: 若特征值 λ_i 有 k_i 根，则对

$$\lambda_i^m = r_{n-1}\lambda_i^{n-1} + ... + r_1\lambda_i + r_0$$

求 k_i 次导，得到 k_i 个方程，由于 $k_1 + ... + k_n = n$，这样还是能得到 n 个 n 元方程组，解出 $(r_0, ..., r_{n-1})$ 全部代入回（2）即可得到 $\boldsymbol{A}^m$

3.4 对角化

本节的目标是把复杂的 $\boldsymbol{A}$ 对角化为一个形式更简单的 $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, ..., \lambda_n\}$，这样有利于我们做很多操作，比如：求矩阵方幂、二次型标准化、主成分分析（PCA）、解耦等等。

1. 对角化定义:

- 矩阵可对角化：若 $A \sim \mathrm{diag}$，则 A 可对角化
- 线性变换可对角化：若 $\mathcal{A}$ 在某组基下的矩阵为 diag，则 $\mathcal{A}$ 可对角化

2. 特征值与特征子空间

设 $\lambda_1, ..., \lambda_k$ 是 V 上线性变换 $\mathcal{T}$ 的不同特征值，V_i 是 λ_i 的特征子空间，则

$$V_1 + ... + V_k = V_1 \oplus ... \oplus V_k$$

- **属于不同特征值的特征向量**线性无关（因为直和）

3.4.1 代数重数与几何重数

1. 定义

- λ_i 的代数重数：λ_i 的重根数；
- λ_i 的几何重数：$\dim(V_{\lambda_i})$

2. 几何重数不超过代数重数：设 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = (\lambda - \lambda_1)^{r_1} ... (\lambda - \lambda_s)^{r_s}$，则 $\dim \boldsymbol{V_{\lambda_i}} \le r_i, i = 1, 2, ..., s$

- 直观上，这告诉我们如果特征值 λ 有 r 重根，则属于 λ 的线性无关的特征向量不会超过 r 个

3.4.2 对角化的条件

设 $\mathcal{T}$ 有 k 个特征根，每个特征根的重根数为 s_i

（1）$\mathcal{T}$ 可对角化 $\iff \mathcal{T}$ 有 n 个线性无关的特征向量 $\iff V = V_{\lambda_1} \oplus ... \oplus V_{\lambda_k}$

$\iff \sum\limits_{i=1}^{k} \dim(V_{\lambda_i}) = n \iff \mathrm{rank}(\lambda_i\boldsymbol{E} - \boldsymbol{A}) = n - s_i \iff \dim(V_{\lambda_i}) = s_i$（代数重数 = 几何重数）

（2）$\mathcal{T}$ 有 n 个不同特征值 $\Rightarrow$ $\mathcal{T}$ 可对角化

（3）$\mathcal{T}$ 在 $V_{\mathbb{C}}$ 中无重根 $\Rightarrow \mathcal{T}$ 可对角化

4. 对角化的一般方法

- 求出 $\boldsymbol{A}$ 的全部特征值 $\lambda_1, ..., \lambda_k$
- 求出每个特征值对应的特征向量 $\{\boldsymbol{v_{11}, ..., v_{1n_1}}\}, \{\boldsymbol{v_{21}, ..., v_{2n_2}}\}, ..., \{\boldsymbol{v_{k1}, ..., v_{kn_k}}\}$
  - 注 $_1$：**若特征值无重根，则必可对角化**（n 阶矩阵正好解出来 n 个特征值，正好对应 n 个特征向量）
  - 注 $_2$：**若特征值有重根，则若第 r 重特征值有 r 个线性无关特征向量，**（使得 k 个特征值能正好对应 n 个特征向量）**则也可对角化**
  - 注 $_3$：**否则不可对角化**
- 将 λ_i 的特征向量 $\{\boldsymbol{v_{i1}, ..., v_{in_i}}\}$ 按 i 的顺序排成矩阵 $\boldsymbol{P} = (\boldsymbol{v_{11}, ..., v_{1n_1}, v_{21}, ..., v_{2n_2}, ..., v_{k1}, ..., v_{kn_k}})$，则 $\boldsymbol{P^{-1}AP} = \mathbf{diag}(\boldsymbol{\lambda_1, \lambda_1, ..., \lambda_k, \lambda_k})$（特征根重复几次就排几次）
  - 这里的矩阵 $\boldsymbol{P}, \boldsymbol{P^{-1}AP}$ 都是 n 阶矩阵

3.5 像空间与核空间

3.5.1 概念

1. 像空间： $\mathrm{Im}\mathcal{A} = \mathcal{A}(\boldsymbol{V}) = \{\mathcal{A}\boldsymbol{\alpha} | \boldsymbol{\alpha} \in \boldsymbol{V}\}$

- $\mathcal{A}$ 的秩：$\dim \mathrm{Im}\mathcal{A}$

2. 核空间： $\ker \mathcal{A} = \mathcal{A}^{-1}(0) = \{\boldsymbol{\alpha} | \boldsymbol{\alpha} \in \boldsymbol{V}, \mathcal{A}(\boldsymbol{\alpha}) = 0\}$

- $\mathcal{A}$ 的零度：$\dim \ker \mathcal{A}$

3.5.2 像空间与核空间的性质

设 $\boldsymbol{\varphi}$ 是 n 维线性空间的线性变换，$\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 是一组基，$\boldsymbol{\varphi}$ 在这组基下的矩阵为 $\boldsymbol{A}$

1. $\mathrm{Im}\boldsymbol{\varphi} = L(\boldsymbol{\varphi}(\boldsymbol{\varepsilon_1}), ..., \boldsymbol{\varphi}(\boldsymbol{\varepsilon_n}))$
2. $\dim \mathrm{Im}\boldsymbol{\varphi} = \mathrm{rank}(\boldsymbol{A})$（$\boldsymbol{\varphi}$ 的秩是 $\boldsymbol{A}$ 的秩）
3. $\dim \mathrm{Im}\boldsymbol{\varphi} + \dim \ker \boldsymbol{\varphi} = \dim \boldsymbol{V} = n$
   - 这还不够说明 $\mathrm{Im}\boldsymbol{\varphi} + \ker \boldsymbol{\varphi} = \boldsymbol{V}$
   - $\mathrm{Im}\boldsymbol{\varphi} \oplus \ker \boldsymbol{\varphi} = \boldsymbol{V} \iff \mathrm{Im}\boldsymbol{\varphi} \cap \ker \boldsymbol{\varphi} = \mathbf{0}$
4. $\boldsymbol{\varphi}$ 是满射 $\iff \mathrm{Im}\boldsymbol{\varphi} = \boldsymbol{V}$，$\boldsymbol{\varphi}$ 是单射 $\iff \ker \boldsymbol{\varphi} = \{0\}$，$\boldsymbol{\varphi}$ 是满射 $\iff \boldsymbol{\varphi}$ 是单射

- **（题）** 设 $\boldsymbol{\sigma}$ 为数域 $\mathbb{P}$ 上 n 维线性空间 $\boldsymbol{V}$ 的线性变换，$f_1(x), f_2(x), ..., f_s(x) \in \mathbf{P}[\mathbf{x}]$，$s \geq 2$，且他们两两互素，令 $f(x) = f_1(x)f_2(x)...f_n(x)$，则

$$\ker(f(\boldsymbol{\sigma})) = \ker(f_1(\boldsymbol{\sigma})) \oplus \ker(f_2(\boldsymbol{\sigma})) \oplus ... \oplus \ker(f_s(\boldsymbol{\sigma}))$$

3.6 不变子空间：目的是为了准对角化

1.$\sigma-$ **子空间/**σ **的不变子空间：**$\boldsymbol{\sigma}$ 是 $\boldsymbol{V}$ 上的线性变换，$\boldsymbol{W}$ 是 $\boldsymbol{V}$ 的子空间且 $\sigma(\boldsymbol{W})\subset \boldsymbol{W}$

2. 不变子空间的性质：

- 设 $\boldsymbol{W}=L(\varepsilon_1,...,\varepsilon_n)$，则 $\boldsymbol{W}$ 是 $\sigma-$ 子空间 $\iff$ $\boldsymbol{\sigma(\varepsilon_1),\sigma(\varepsilon_2),...,\sigma(\varepsilon_n)\in W}$
- 两个 $\sigma-$ 子空间的交与和仍是 $\sigma-$ 子空间.

3. 重要的不变子空间

- $\boldsymbol{V}$ 的平凡子空间对任意 $\boldsymbol{\sigma}$ 都是 $\sigma-$ 子空间
- 任何子空间都是数乘变换 $\mathcal{K}$ 的不变子空间
- $\mathrm{Im}\boldsymbol{\sigma}$ 与 $\ker\boldsymbol{\sigma}$ 都是 $\boldsymbol{\sigma}$ 的不变子空间.
- 若 $\boldsymbol{\sigma\tau=\tau\sigma}$，则 $\mathrm{Im}\boldsymbol{\tau}$ 和 $\ker\boldsymbol{\tau}$ 都是 $\boldsymbol{\sigma}$ 的不变子空间，$\mathrm{Im}\boldsymbol{\sigma}$ 和 $\ker\boldsymbol{\sigma}$ 都是 $\boldsymbol{\tau}$ 的不变子空间
- $f(x)\in \mathbf{P}[\mathbf{x}]$，则 $\mathrm{Im}f(\boldsymbol{\sigma})$ 和 $\ker f(\boldsymbol{\sigma})$ 都是 $\boldsymbol{\sigma}$ 的不变子空间（$*$）
- 特征子空间 $\boldsymbol{V_{\lambda_0}}$ 是 $\boldsymbol{\sigma}$ 的不变子空间
- 由 $\boldsymbol{\sigma}$ 的特征向量生成的子空间是 $\boldsymbol{\sigma}$ 的不变子空间
  - 特别地，$\boldsymbol{\sigma}$ 的一个特征向量 ξ 生成的子空间 $L(\boldsymbol{\xi})$ 是一个一维 $\boldsymbol{\sigma}-$ 子空间；
    而一个一维 $\boldsymbol{\sigma}-$ 子空间必可看成 $\boldsymbol{\sigma}$ 的一个特征向量 $\boldsymbol{\xi'}$ 生成的子空间（充要条件）
- **根子空间：**设 $\boldsymbol{\sigma}$ 的特征多项式 $f(\lambda)$ 具有分解式

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}...(\lambda-\lambda_s)^{r_s}$$

  则 $V^{\lambda_i}=\{\boldsymbol{\xi}\in \boldsymbol{V}|(\boldsymbol{\sigma}-\boldsymbol{\lambda_i}\boldsymbol{\mathcal{E}})^{r_i}\boldsymbol{\xi}=0\}$ 为 $\boldsymbol{\sigma}$ 的属于特征值 λ_i 的**根子空间**

  - **根子空间为 $\boldsymbol{\sigma}$ 的不变子空间**
  - **记 $f_i(\lambda)=\dfrac{f(\lambda)}{(\lambda-\lambda_i)}$，则 $\mathrm{Im}f_i(\sigma)$ 为 $\boldsymbol{\sigma}$ 的不变子空间（来自 $*$）**

4. 不变子空间 W 上的线性变换：把线性空间 $\boldsymbol{V}$ 上的线性变换 $\boldsymbol{\sigma}$ 看作不变子空间 $\boldsymbol{W}$ 上的线性变换 $\boldsymbol{\sigma}|_{\boldsymbol{W}}$

- $\boldsymbol{\xi}\in \boldsymbol{W}$ 时，$\boldsymbol{\sigma}|_{\boldsymbol{W}}(\boldsymbol{\xi})=\boldsymbol{\sigma}(\boldsymbol{\xi})$
- $\boldsymbol{\xi}\notin \boldsymbol{W}$ 时，$\boldsymbol{\sigma}|_{\boldsymbol{W}}$ 无意义（不变子空间上的线性变换必须在不变子空间上起作用）
- $\boldsymbol{\sigma}|_{\boldsymbol{W}}(\boldsymbol{W})\subset \boldsymbol{W}$
- $\boldsymbol{\sigma}|_{\ker\boldsymbol{\sigma}}=\boldsymbol{\mathcal{O}}$（任意线性变换在其核上为零变换）
- $\boldsymbol{\sigma}|_{V_{\lambda_0}}=\boldsymbol{\lambda\mathcal{E}}$（$\boldsymbol{\sigma}$ 在其特征子空间上是数乘变换）

3.6.1 不变子空间：化简线性变换的矩阵

1. 设 $\boldsymbol{\sigma}$ 是 n 维线性空间 $\boldsymbol{V}$ 的线性变换，$\boldsymbol{W}$ 是 $\boldsymbol{V}$ 的 $\sigma-$ 子空间，$\boldsymbol{e_1, e_2, ..., e_k}$ 为 $\boldsymbol{W}$ 的一组基，把它扩充为 $\boldsymbol{V}$ 的一组基 $\boldsymbol{e_1, ..., e_k, e_{k+1}, .., e_n}$；若 $\boldsymbol{\sigma}|_{\boldsymbol{W}}$ 在基 $\boldsymbol{e_1, e_2, ..., e_k}$ 下的矩阵为 $\boldsymbol{A_1} \in \boldsymbol{P}^{k\times k}$，则 $\boldsymbol{\sigma}$ 在基 $\boldsymbol{e_1, ..., e_k, e_{k+1}, .., e_n}$ 下的矩阵具有下列形状：

$$\begin{pmatrix} \boldsymbol{A_1} & \boldsymbol{A_2} \\ \boldsymbol{O} & \boldsymbol{A_3} \end{pmatrix}$$

注：这是因为 $\boldsymbol{W}$ 是 $\boldsymbol{V}$ 的不变子空间，所以 $\boldsymbol{\sigma(e_1), \sigma(e_2), ..., \sigma(e_k)} \in \boldsymbol{W}$，这说明 $\boldsymbol{\sigma(e_1), \sigma(e_2), ..., \sigma(e_k)} \in \boldsymbol{W}$ 均可以被 $\boldsymbol{e_1, e_2, ..., e_k}$ 线性表出：

$$\begin{cases} \boldsymbol{\sigma(e_1)} = a_{11}\boldsymbol{e_1} + ... + a_{k1}\boldsymbol{e_k} \\ \boldsymbol{\sigma(e_2)} = a_{12}\boldsymbol{e_1} + ... + a_{k2}\boldsymbol{e_k} \\ ... \\ \boldsymbol{\sigma(e_k)} = a_{1k}\boldsymbol{e_1} + ... + a_{kk}\boldsymbol{e_k} \end{cases}$$

那么，

$$\boldsymbol{\sigma(e_1, e_2, ..., e_n)} = \boldsymbol{(e_1, e_2, ..., e_n)} \begin{pmatrix} a_{11} & a_{12} & ... & a_{1k} & a_{1,k+1} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2k} & a_{2,k+1} & ... & a_{2n} \\ ... & ... & ... & ... & ... & ... & ... \\ a_{k1} & a_{k2} & ... & a_{kk} & a_{k,k+1} & ... & a_{kn} \\ 0 & 0 & ... & 0 & a_{k+1,k+1} & ... & a_{kn} \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & a_{n,k+1} & ... & a_{nn} \end{pmatrix}$$

2.（1 的推广）$\boldsymbol{V}$ 的线性变换 $\boldsymbol{\sigma}$ 在某组基下的矩阵为准对角形 $\Longleftrightarrow$ $\boldsymbol{V}$ 可分解为一些 $\boldsymbol{\sigma}$ 的不变子空间的直和.

- （$\Rightarrow$）设 $\boldsymbol{\sigma}$ 是 n 维线性空间 $\boldsymbol{V}$ 的线性变换，$\boldsymbol{W_i}$ 都是 $\boldsymbol{\sigma}$ 的不变子空间，$\boldsymbol{e_{i1}, e_{i2}, ..., e_{in_i}}$ 是 $\boldsymbol{W_i}$ 的一组基，且 $\boldsymbol{\sigma}|_{\boldsymbol{W_i}}$ 在这组基下的矩阵为 $\boldsymbol{A_i}$，$\boldsymbol{A_i} \in \boldsymbol{P}^{n_i\times n_i}, i = 1, 2, ..., s$，若 $\boldsymbol{V} = \boldsymbol{W_1} \oplus \boldsymbol{W_2} \oplus ... \oplus \boldsymbol{W_s}$，则 $\boldsymbol{e_{11}, ..., e_{1n_1}, ..., e_{s1}, ..., e_{sn_s}}$ 为 $\boldsymbol{V}$ 的一组基，且在这组基下 $\boldsymbol{\sigma}$ 的矩阵为准对角阵

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{A_1} & & & \\ & \boldsymbol{A_2} & & \\ & & ... & \\ & & & \boldsymbol{A_s} \end{pmatrix}$$

- （$\Leftarrow$）反之，若 $\boldsymbol{\sigma}$ 在基 $\boldsymbol{e_{11}, e_{12}, ..., e_{1n_1}, ..., e_{s1}, e_{s2}, ..., e_{sn_s}}$ 下的矩阵为准对角矩阵 $\boldsymbol{\Lambda}$，则由 $\boldsymbol{e_{i1}, e_{i2}, ..., e_{in_i}}$ 生成的子空间 $\boldsymbol{W_i}$ 为 $\boldsymbol{\sigma}$ 的不变子空间，且 $\boldsymbol{V}$ 具有直和分解：

$$\boldsymbol{V} = \boldsymbol{W_1} \oplus \boldsymbol{W_2} \oplus ... \oplus \boldsymbol{W_s}$$

3.6.2 线性空间的分解

$\boldsymbol{\sigma}$ 为 $\boldsymbol{V}$ 上线性变换，则 $\boldsymbol{V}$ 可以分解为其根子空间 $\boldsymbol{V}^{\boldsymbol{\lambda}_i} = \{\boldsymbol{\xi} \mid (\mathcal{A}-\lambda_i\mathcal{E})^{r_i}\boldsymbol{\xi}=0\}$ 的直和

$$\boldsymbol{V} = \boldsymbol{V}^{\boldsymbol{\lambda}_1} \oplus \boldsymbol{V}^{\boldsymbol{\lambda}_2} \oplus \ldots \oplus \boldsymbol{V}^{\boldsymbol{\lambda}_s}$$

- **推论 1**：$\boldsymbol{V}_{\mathbb{C}}$ 均可以分解为根子空间 $\boldsymbol{V}^{\boldsymbol{\lambda}_i}$ 的直和
- **推论 2**：$\boldsymbol{V}_{\mathbb{C}}$ 上的线性变换 $\mathcal{T}$ 均可以准对角化
- **推论 2'**：$\mathbb{C}$ 上的 n 阶方阵都可以准对角化
- **推论 3**：数域 $\mathbb{F}$ 上的 n 维线性空间上的线性变换 $\mathcal{T}$ 可准对角化 $\Longleftrightarrow$ 其特征多项式可以分解为一次多项式的乘积

$$f_{\mathcal{T}}(x) = (x-\lambda_1)^{r_1}\ldots(x-\lambda_s)^{r_s}，\ r_1+\ldots+r_s=n$$

3.7 Jordan 标准型

引言：上节说到，不是所有的 n 阶方阵都能对角化（可能凑不出 n 个特征向量），但是一定可以**准对角化**（根子空间直和分解，每个对角块都是根子空间 $\boldsymbol{V}^{\boldsymbol{\lambda}_i}$ 的矩阵），本节中 Jordan 标准型给出了一个准对角化的方案：让每个准对角块都是 Jordan 块。Jordan 标准型不是中国人民大学高等代数考试的重点，对其结论与性质了解即可。

1.Jordan 块：其中 $\lambda_0 \in \mathbb{C}$

$$\boldsymbol{J}(\boldsymbol{\lambda_0}, \boldsymbol{k}) = \begin{pmatrix} \lambda_0 & & & & \\ 1 & \lambda_0 & & & \\ & & \ddots & & \\ & & & \lambda_0 & \\ & & & 1 & \lambda_0 \end{pmatrix}_{k\times k}$$

2.Jordan 型矩阵：由若干个 Jordan 块组成的准对角矩阵:

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}(\boldsymbol{\lambda_1}, \boldsymbol{k_1}) & & & \\ & \boldsymbol{J}(\boldsymbol{\lambda_2}, \boldsymbol{k_2}) & & \\ & & \ddots & \\ & & & \boldsymbol{J}(\boldsymbol{\lambda_s}, \boldsymbol{k_s}) \end{pmatrix}$$

其中 $\lambda_i \in \mathbb{C}$ 且可以部分相同

- 设 $\boldsymbol{A}$ 的 Jordan 标准型为 $\boldsymbol{J}$，则如果 $\boldsymbol{J}$ 中 Jordan 块下的 1 没有出现，说明 $\boldsymbol{A}$ 能对角化，否则不行

3.Jordan 标准型结论:

- **结论 1**：设 $\mathcal{A}$ 是 $\mathbb{C}$ 上 n 维线性空间 $\boldsymbol{V}$ 的线性变换，则在 $\boldsymbol{V}$ 中存在一组基，使得 $\mathcal{A}$ 在该基下的矩阵为 Jordan 标准型。
  而且这个 Jordan 标准型除了 Jordan 块的次序之外由 $\mathcal{A}$ 唯一决定，称之为 $\mathcal{A}$ 的 Jordan 标准型。

- **结论 2**：$\mathbb{C}$ 上 n 阶矩阵都相似与一个 Jordan 标准型，其中对角线上的元素是 $\boldsymbol{A}$ 的全部特征值。

$$\boldsymbol{J}=\begin{pmatrix}\lambda_1 & & & & & \\ 1 & \lambda_1 & & & & \\ & 1 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \lambda_s & \\ & & & & 1 & \lambda_s\end{pmatrix}$$

  – 即可以找到一个矩阵 $\boldsymbol{P}=(\boldsymbol{\varepsilon_{11},\varepsilon_{12},...,\varepsilon_{1k_1},\varepsilon_{21},\varepsilon_{22},...,\varepsilon_{2k_2},...,\varepsilon_{s1},\varepsilon_{s2},...,\varepsilon_{sk_s}})$，使得 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{AP}$

- **结论 3**：Jordan 块的块数 = 特征向量的个数

- **结论 4**：特征值 λ_i 的代数重数 = Jordan 标准型中 λ_i 出现的次数，λ_i 的几何重数 = 含 λ_i 的块数

- *** 结论 5**：k 阶 Jordan 块的个数为：

$$\operatorname{rank}(\boldsymbol{A}-\boldsymbol{\lambda I})^{k-1}+\operatorname{rank}(\boldsymbol{A}-\boldsymbol{\lambda I})^{k+1}-2\operatorname{rank}(\boldsymbol{A}-\boldsymbol{\lambda I})^{k}$$

3.8 最小多项式

（引言：Hamilton-Cayley 告诉我们，能找到 $\boldsymbol{A}$ 的零化多项式（比如特征多项式）使得 $f(\boldsymbol{A})=\boldsymbol{O}$，我们关心的是：次数最低的零化多项式和 $\boldsymbol{A}$ **的对角化间的关系**。事实上，最小多项式与特征多项式有许多相似的性质。）

1. 最小多项式：$f(\boldsymbol{A})=\boldsymbol{O}$，次数最低，首 1

3.8.1 最小多项式的基本性质

- $\boldsymbol{A}$ 的最小多项式是唯一的

- 设 $g(x)$ 为 $\boldsymbol{A}$ 的最小多项式，则：$h(x)$ 以 $\boldsymbol{A}$ 为根 $\iff g(x)|h(x)$

- $\boldsymbol{A}$ 的最小多项式是 $\boldsymbol{A}$ 的特征多项式的一个因式

  – $\boldsymbol{A}$ 是数量矩阵 $k\boldsymbol{E} \iff \boldsymbol{A}$ 的最小多项式为一次多项式

- 相似矩阵具有相同的最小多项式，但具有最小多项式的矩阵未必相似

- 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A_1} & & & \\ & \boldsymbol{A_2} & & \\ & & \ddots & \\ & & & \boldsymbol{A_s}\end{pmatrix}$，$g_1(x),g_2(x),...,g_s(x)$ 分别为 $\boldsymbol{A_1,A_2,...,A_s}$ 的最小多项式，则 $\boldsymbol{A}$ 的最小多项式为 $[g_1(x),g_2(x),...,g_s(x)]$

- k 阶 Jordan 块

$$\boldsymbol{J}=\begin{pmatrix}\lambda & & & & & \\ 1 & \lambda & & & & \\ & 1 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & 1 & \lambda & \\ & & & & 1 & \lambda\end{pmatrix}$$

  的最小多项式为 $(x-\lambda)^k$

- **（重要定理）$\boldsymbol{A}$ 与对角阵 $\boldsymbol{\Lambda}$ 相似 $\iff$ $\boldsymbol{A}$ 的最小多项式是数域 $\mathbb{F}$ 上互素的一次因式的乘积**
  - $\boldsymbol{A}\in\mathbb{C}^{\boldsymbol{n}\times\boldsymbol{n}}\sim\mathrm{diag}\iff\boldsymbol{A}$ 的最小多项式没有重根

3.8.2 最小多项式与 Jordan 标准型的关系

1.Jordan 块最高阶数的确定：设矩阵 $\boldsymbol{A}$ 的最小多项式是 $m(x)=\prod_{i=1}^{s}(x-\lambda_i)^{r_i}$，则 $\boldsymbol{A}$ 的 Jordan 标准型中以 λ_i 为主对角元的 Jordan 块的最高阶数为 r_i

通过 **3.7 Jordan 标准型**的结论 2，3，4，5 和 **3.8.2 最小多项式与 Jordan 标准型的关系**，可以算出低阶矩阵的 Jordan 标准型。

4 Euclid 空间

4.1 Euclid 空间的定义

1. 内积：设 $\boldsymbol{V}$ 是 $\mathbb{R}$ 上的线性空间，$\forall \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in \boldsymbol{V}$，$\forall k \in \mathbb{R}$，定义一个**二元实函数** $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 满足：

- 对称性：$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\boldsymbol{\beta}, \boldsymbol{\alpha})$
- 数乘：$(k\boldsymbol{\alpha}, \boldsymbol{\beta}) = k(\boldsymbol{\alpha}, \boldsymbol{\beta})$
- 可加性：$(\boldsymbol{\alpha} + \boldsymbol{\gamma}, \boldsymbol{\beta}) = (\boldsymbol{\alpha}, \boldsymbol{\beta}) + (\boldsymbol{\gamma}, \boldsymbol{\beta})$
- 正定性：$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \geq 0$，当且仅当 $\boldsymbol{\alpha} = \boldsymbol{0}$ 时 $(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \boldsymbol{0}$

则称 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 为 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 的内积.
内积的简单性质有：

- $(\sum\limits_{i=1}^{s} k_i \boldsymbol{\alpha_i}, \sum\limits_{j=1}^{s} l_j \boldsymbol{\beta_j}) = \sum\limits_{i=1}^{s} \sum\limits_{j=1}^{t} k_i l_j (\boldsymbol{\alpha_i}, \boldsymbol{\beta_j})$
- $(\boldsymbol{0}, \boldsymbol{\beta}) = \boldsymbol{0}$

2.Euclid 空间：定义了内积的 $\mathbb{R}$ 上的线性空间 $\boldsymbol{V}$ 为 Euclid 空间.

- Euclid 空间 $\boldsymbol{V}$ 是特殊的线性空间，除向量的线性运算外还有“内积”运算
- $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}$
- “内积”不仅限于点积，可以是
  - $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{\beta}$
  - $(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})) = \displaystyle\int_a^b f(x) g(x) \mathrm{d}x$

4.1.1 Euclid 空间中向量的长度

1. 模：$\|\boldsymbol{\alpha}\| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}$

- $\boldsymbol{\alpha}$ 的单位化：$\dfrac{\boldsymbol{1}}{\|\boldsymbol{\alpha}\|}\boldsymbol{\alpha}$

4.1.2 Euclid 空间中向量的夹角

1.Cauchy-Buniakowsky 不等式：对 Euclid 空间 $\boldsymbol{V}$ 中任意两个向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$，成立

$$|(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leq \|\boldsymbol{\alpha}\| \, \|\boldsymbol{\beta}\|$$

等号成立条件：$\boldsymbol{\alpha}, \boldsymbol{\beta}$ 线性相关.

- 三角不等式：

$$\|\boldsymbol{\alpha_1} + \boldsymbol{...} + \boldsymbol{\alpha_s}\| \leq \|\boldsymbol{\alpha_1}\| + ... + \|\boldsymbol{\alpha_m}\|$$

2. 非零向量间的夹角：定义 Euclid 空间 $\boldsymbol{V}$ 中任意两非零向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$ 的夹角为

$$\langle\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle = \arccos\frac{(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\,\|\boldsymbol{\beta}\|}$$

- 正交：$\boldsymbol{\alpha}\perp\boldsymbol{\beta} \iff (\alpha, \beta) = 0 \iff \langle\alpha, \beta\rangle = \frac{\pi}{2} \iff \cos\langle\alpha, \beta\rangle = 0$
- $\mathbf{0}$ 与任意向量正交
- $\boldsymbol{\alpha}\perp\boldsymbol{\alpha} \iff \boldsymbol{\alpha} = \mathbf{0}$

4.1.3 正交向量的性质

- 正交投影：$\boldsymbol{\alpha} \in \boldsymbol{V}$ 为一固定非零向量，把 $\boldsymbol{V}$ 中每个向量 $\boldsymbol{\xi}$ 变成它在 $\boldsymbol{\alpha}$ 上的**内射影**是 $\boldsymbol{V}$ 上的一个线性变换

$$\begin{aligned}\mathcal{P}: \boldsymbol{V} &\longmapsto \boldsymbol{V}\\ \boldsymbol{\xi} &\longmapsto \frac{(\boldsymbol{\alpha}, \boldsymbol{\xi})}{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}\boldsymbol{\alpha}\end{aligned}$$

  - $\boldsymbol{\xi}$ 在 $\boldsymbol{\alpha}$ 上的正交投影为 $\frac{(\boldsymbol{\alpha}, \boldsymbol{\xi})}{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}\boldsymbol{\alpha}$
  - $(\boldsymbol{\xi} - \frac{(\boldsymbol{\alpha}, \boldsymbol{\xi})}{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}\boldsymbol{\alpha})\perp\boldsymbol{\alpha}$

- 勾股定理：若 Euclid 空间 $\boldsymbol{V}$ 中向量 $\boldsymbol{\alpha_1}, \ldots, \boldsymbol{\alpha_s}$ 两两正交，则

$$\|\boldsymbol{\alpha_1} + \ldots + \boldsymbol{\alpha_s}\|^2 = \|\boldsymbol{\alpha_1}\|^2 + \ldots + \|\boldsymbol{\alpha_m}\|^2$$

4.1.4 Euclid 空间中的距离

1. 距离：$d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|$

- $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = d(\boldsymbol{\beta}, \boldsymbol{\alpha})$，$d(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \leq d(\boldsymbol{\alpha}, \boldsymbol{\beta}) + d(\boldsymbol{\beta}, \boldsymbol{\gamma})$

4.1.5 Euclid 空间中内积的矩阵表示

本质上说的是：两个向量用基展开时，两个向量的内积可以用二次型来简化表达

设 $\boldsymbol{V}$ 为 Euclid 空间，$\boldsymbol{\varepsilon_1}, \ldots, \boldsymbol{\varepsilon_n}$ 为 $\boldsymbol{V}$ 的一组基，对 $\boldsymbol{V}$ 中任意两个向量

$$\alpha = \sum_{i=1}^{n} x_i\boldsymbol{\varepsilon_i}, \boldsymbol{\beta} = \sum_{j=1}^{n} y_j\boldsymbol{\varepsilon_j}$$

他们的内积表达为

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\sum_{i=1}^{n} x_i\boldsymbol{\varepsilon_i}\sum_{j=1}^{n} y_j\boldsymbol{\varepsilon_j}) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j(\boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_j}) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j a_{ij} = \boldsymbol{X^T A Y}$$

其中 $\boldsymbol{A}$ 为度量矩阵 $\boldsymbol{A} = \begin{pmatrix} (\boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_1}) & (\boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_2}) & \ldots & (\boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_n}) \\ (\boldsymbol{\varepsilon_2}, \boldsymbol{\varepsilon_1}) & (\boldsymbol{\varepsilon_2}, \boldsymbol{\varepsilon_2}) & \ldots & (\boldsymbol{\varepsilon_2}, \boldsymbol{\varepsilon_n}) \\ \ldots & \ldots & \ldots & \ldots \\ (\boldsymbol{\varepsilon_n}, \boldsymbol{\varepsilon_1}) & (\boldsymbol{\varepsilon_n}, \boldsymbol{\varepsilon_2}) & \ldots & (\boldsymbol{\varepsilon_n}, \boldsymbol{\varepsilon_n}) \end{pmatrix}$，$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$，$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}$

- $\boldsymbol{A} = \boldsymbol{A}^T$
- Euclid 空间在任意一组基下的度量矩阵 $\boldsymbol{A}$ 是正定矩阵，**反之**，给定正定矩阵 $\boldsymbol{A}$ 和线性空间 $\boldsymbol{V}$ 的一组基，都可以定义 V 上的内积，使得 $\boldsymbol{V}$ 是 Euclid 空间且在给定的这组基下的度量矩阵为 $\boldsymbol{A}$
- 基确定下来则度量矩阵确定，**同一内积在不同基下的度量矩阵是合同的**
- 我们可以选择基为**正交向量组 $\boldsymbol{\epsilon_1, ..., \epsilon_n}$**，这样度量矩阵 $\boldsymbol{A}$ 为正对角阵，于是：
  实对称矩阵 $\boldsymbol{A}$ 正定 $\iff$ $\boldsymbol{A}$ 与任一正对角阵合同 $\iff$ $\boldsymbol{A}$ 与单位矩阵 $\boldsymbol{E}$ 合同

4.1.6 Euclid 子空间

Euclid 空间的非空子集合，只要对加法和数乘封闭，则为 Euclid 子空间，维数就是线性子空间的维数。

4.2 标准正交基

4.2.1 标准正交基的定义

1. 正交向量组：$\boldsymbol{\alpha_1, ..., \alpha_m}$ 两两正交

- 一个向量 $\boldsymbol{\alpha} \neq \boldsymbol{0}$ 也是正交向量组
- 正交向量组必线性无关，但线性无关组未必为正交向量组（**向量正交是比线性无关更强的条件**）
- n 维 Euclid 空间中正交向量组所含向量个数 $\leq n$

2. 正交基：n 维 Euclid 空间中由 n 个向量构成的正交向量组
3. 标准正交组：都是单位向量的正交向量组
4. 标准正交基：都是单位向量的正交基

- Euclid 空间中标准正交基总是存在的
- $\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 是标准正交基 $\iff$ 度量矩阵 $\boldsymbol{A} = \boldsymbol{E_n}$
- $\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 是标准正交基，将 $\boldsymbol{V}$ 中向量表为 $\boldsymbol{\alpha} = x_1\boldsymbol{\varepsilon_1} + ... + x_n\boldsymbol{\varepsilon_n}$，则 $x_i = (\boldsymbol{\alpha}, \boldsymbol{\varepsilon_i})$
- $\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 是标准正交基，$\boldsymbol{\alpha} = x_1\boldsymbol{\varepsilon_1} + ... + x_n\boldsymbol{\varepsilon_n}$，$\boldsymbol{\beta} = y_1\boldsymbol{\varepsilon_1} + ... + y_n\boldsymbol{\varepsilon_n}$，则 $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = x_1y_1 + ... + x_ny_n$

4.2.2 构造标准正交基

1.n 维 Euclid 空间中任一个正交向量组都能扩充为一组正交基
2. 通过度量矩阵正交化

1. 取 $\boldsymbol{V}$ 中一组基 $\boldsymbol{\varepsilon_1, ...\varepsilon_n}$，则**它的度量矩阵 $\boldsymbol{A}$ 是正定的**
2. 正定矩阵必合同于 $\boldsymbol{E}$，那么存在**可逆矩阵 $\boldsymbol{C}$** 使得 $\boldsymbol{C^T A C = E}$

3. 令 $\begin{pmatrix}\boldsymbol{\eta_1} & \boldsymbol{\eta_2} & \boldsymbol{\dots} & \boldsymbol{\eta_n}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\varepsilon_1} & \boldsymbol{\varepsilon_2} & \boldsymbol{\dots} & \boldsymbol{\varepsilon_n}\end{pmatrix}\boldsymbol{C}$，则 $\begin{pmatrix}\boldsymbol{\eta_1} & \boldsymbol{\eta_2} & \boldsymbol{\dots} & \boldsymbol{\eta_n}\end{pmatrix}$ 是标准正交基
（因为这组基的度量矩阵就是 $\boldsymbol{C^TAC=E} \iff$ 是标准正交基）

3.Schmidt 正交化过程

1. **正交化**：将线性无关的向量组 $\boldsymbol{\alpha_1,\dots,\alpha_m}$ 化成正交向量组 $\boldsymbol{\beta_1,\dots,\beta_m}$：

$$\boldsymbol{\beta_1=\alpha_1},\qquad \boldsymbol{\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1},\qquad \boldsymbol{\beta_j=\alpha_j-\sum_{i=1}^{j-1}\frac{(\alpha_j,\beta_i)}{(\beta_i,\beta_i)}\beta_i},\ j=2,3,...,m$$

2. **单位化**：将正交向量组单位化得到标准正交组 $\boldsymbol{\eta_1,\dots,\eta_m}$

$$\boldsymbol{\eta_i}=\frac{1}{\|\boldsymbol{\beta_i}\|}\boldsymbol{\beta_i},\ i=1,2,...,m$$

4. 若对 n 维 Euclid 空间中任一组基 $\boldsymbol{\varepsilon_1,\dots,\varepsilon_n}$ Schmidt 正交化对应得到 $\boldsymbol{\eta_1,\dots,\eta_n}$，则

- $L(\boldsymbol{\varepsilon_1,\dots,\varepsilon_i})=L(\boldsymbol{\eta_1,\dots,\eta_i})$，$i=1,...,n$
- 若 $(\boldsymbol{\eta_1,\dots,\eta_n})=(\boldsymbol{\varepsilon_1,\dots,\varepsilon_n})\boldsymbol{T}$，则 $\boldsymbol{T}$ 为上三角形

4.2.3 正交矩阵

1. 正交矩阵 $\boldsymbol{A} \iff \boldsymbol{A^{-1}=A^T}$
$\iff \boldsymbol{A}$ 的列向量组是 Euclid 空间 $\mathbb{R}^n$ 的标准正交基 $\iff \boldsymbol{A^TA=E}$
$\iff \boldsymbol{A}$ 的行向量组是 Euclid 空间 $\mathbb{R}^n$ 的标准正交基 $\iff \boldsymbol{AA^T=E}$

- $\boldsymbol{A}$ 为正交矩阵 $\Rightarrow |\boldsymbol{A}|=\pm1$
- 若 $\boldsymbol{\varepsilon_1,\dots,\varepsilon_n}$ 为标准正交基，$\boldsymbol{A}$ 为正交矩阵，若 $(\boldsymbol{\eta_1,\dots,\eta_n})=(\boldsymbol{\varepsilon_1,\dots,\varepsilon_n})\boldsymbol{A}$，则 $\boldsymbol{\eta_1,\dots,\eta_n}$ 也是标准正交基
- $\boldsymbol{A,B}$ 为正交阵 $\Rightarrow\boldsymbol{AB}$ 为正交阵，$\boldsymbol{A^T}$ 为正交阵，$\boldsymbol{A^*}$ 为正交阵
- $\forall\boldsymbol{\alpha,\beta}\in\mathbb{R}^n$，$\boldsymbol{(A\alpha,A\beta)=(\alpha,\beta)}$，从而 $\|\boldsymbol{A\alpha}\|=\|\boldsymbol{\alpha}\|$，夹角 $\langle\boldsymbol{A\alpha,A\beta}\rangle=\langle\boldsymbol{\alpha,\beta}\rangle$
- **标准正交基之间的过渡矩阵 $\boldsymbol{A}$ 为正交矩阵**
- 正交矩阵的实特征值为 1 或 -1，复特征值的模为 1

2. 寻找一个正交矩阵的方法

- 将 n 个标准正交基竖着/横着排成一个矩阵 $\boldsymbol{A}$
- 求算两个标准正交基之间的过渡矩阵

3. 正交矩阵的 $\boldsymbol{QR}$ 分解：设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 列满秩，则存在一个列正交矩阵 $\boldsymbol{Q_{m\times n}}$ 和一个主对角线上元素都是正的上三角矩阵 $\boldsymbol{T_{n\times n}}$，使得 $\boldsymbol{A=QR}$，且这样的分解唯一

- 先对 $\boldsymbol{A}$ 的列向量 Schmidt 正交化，得到标准正交基 $\boldsymbol{\eta_1, ...\eta_n}$，竖着排成矩阵 $\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{\eta_1} & \boldsymbol{\eta_2} & \boldsymbol{...} & \boldsymbol{\eta_n} \end{pmatrix}$

- 将 $\boldsymbol{A}$ 的列向量表示为 $\boldsymbol{\eta_1, ..., \eta_n}$ 的线性组合：$\boldsymbol{A_i = k_{i1}\eta_1 + ... + k_{in}\eta_n}$，那么 $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{k_{11}} & \boldsymbol{k_{21}} & \boldsymbol{...} & \boldsymbol{k_{n1}} \\ \boldsymbol{k_{12}} & \boldsymbol{k_{22}} & \boldsymbol{...} & \boldsymbol{k_{n2}} \\ \boldsymbol{...} & \boldsymbol{...} & \boldsymbol{...} & \\ \boldsymbol{k_{1n}} & \boldsymbol{k_{2n}} & \boldsymbol{...} & \boldsymbol{k_{nn}} \end{pmatrix}$

4.3 Euclid 空间的同构

1.Euclid 空间的同构： 存在一个一一对应的 $\boldsymbol{\sigma}$：

$$\boldsymbol{V}_{\mathbb{R}} \longmapsto \mathbb{R}^n$$
$$\alpha \longmapsto (x_1, x_2, ..., x_n)^T$$

满足

- $\boldsymbol{\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)}$
- $\boldsymbol{\sigma(k\alpha)} = k\boldsymbol{\sigma(\alpha)}$
- $\boldsymbol{(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)}$

2. 线性空间的同构与 Euclid 空间的同构

- 线性空间中 $\boldsymbol{V}$ 到自身的同构：**所有可逆线性变换**
- Euclid 空间中 $\boldsymbol{V}$ 到自身的同构：不仅要保证可逆线性变换，还要"保内积"$\Rightarrow$ **所有正交变换**

3.Euclid 空间的重要性质：

- 两个有限维 Euclid 空间同构 $\iff$ 维数相同
  - 无论内积的定义如何，n 维 Euclid 空间都与 $\mathbb{R}^n$ 同构
- 两个同维 Euclid 空间同构的建立：分别找出两个空间的标准正交基，使得基向量一一对应

4.4 正交变换

1. 正交变换： Euclid 空间 $\boldsymbol{V}$ 的线性变换 $\boldsymbol{\sigma}$ 保持向量的内积不变

$$\boldsymbol{(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)}$$

2. 正交变换的性质：

- $\boldsymbol{\sigma}$ 是正交变换
  $\iff \|\boldsymbol{\sigma(\alpha)}\| = \|\boldsymbol{\alpha}\|$
  $\iff$ 若 $\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 是 $\boldsymbol{V}$ 的标准正交基，则 $\boldsymbol{\sigma(\varepsilon_1), ..., \sigma(\varepsilon_n)}$ 也是 $\boldsymbol{V}$ 的标准正交基
  $\iff \boldsymbol{\sigma}$ 在任一组标准正交基下的矩阵是正交矩阵
  $\iff \boldsymbol{\sigma}$ 保持向量间距离不变：$\|\boldsymbol{\sigma(\xi) - \sigma(\eta)}\| = \|\boldsymbol{\xi - \eta}\|$

- 正交变换 $\boldsymbol{\sigma}$ 保持向量夹角不变

- 正交变换 $\boldsymbol{\sigma}$ 保持正交性不变

- 正交变换的特征值只能是 ± 1

- Euclid 空间 $\boldsymbol{V}$ 的正交变换是 $\boldsymbol{V}$ 到自身的同构映射
  - 正交变换的逆变换是正交变换（反身性）
  - 正交变换的乘积是正交变换（传递性）

3. 正交变换的分类

- **第一类正交变换**：$|A|=1$（旋转）

- **第二类正交变换**：$|A|=-1$（反射）

4. 镜面反射：设 Euclid 空间中的一组标准正交基：$\boldsymbol{\varepsilon_1,...,\varepsilon_n}$，称变换 $\boldsymbol{\sigma(\varepsilon_i)}=\begin{cases}-\boldsymbol{\varepsilon_i} & i=1\\ \boldsymbol{\varepsilon_i} & i=2,...,n\end{cases}$ 为**镜面反射**

- $\boldsymbol{\sigma}$ 在基 $\boldsymbol{\varepsilon_1,...,\varepsilon_n}$ 下的矩阵为 $\boldsymbol{M}=\begin{pmatrix}-1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$

- 镜面反射是第二类正交变换，$\boldsymbol{\sigma^2=\mathcal{E},\sigma^{-1}=\sigma}$

- 相当于一个向量关于一个**超平面**反射

4.5 Euclid 子空间

4.5.1 Euclid 空间的正交子空间

1. 子空间的正交：$\forall\boldsymbol{\alpha}\in\boldsymbol{V_1},\boldsymbol{\beta}\in\boldsymbol{V_2}$，恒有 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0\iff\boldsymbol{V_1}\perp\boldsymbol{V_2}$

2. 向量与子空间的正交：给定向量 $\boldsymbol{\alpha}\in\boldsymbol{V}$，对 $\forall\boldsymbol{\beta}\in\boldsymbol{V_1}$，恒有 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0\iff\boldsymbol{\alpha}\perp\boldsymbol{V_1}$

- 特别地，若 $\boldsymbol{\beta}\perp\boldsymbol{\alpha_i},(i=1,...,s)$，则 $\boldsymbol{\beta}\perp\sum\limits_{i=1}^{s}\boldsymbol{k_i\alpha_i}\Rightarrow\boldsymbol{\beta}\perp\boldsymbol{L(\alpha_1,...,\alpha_s)}$

- $\boldsymbol{V_1}\perp\boldsymbol{V_2}\iff\boldsymbol{V_1}$ 中的每个向量都与 $\boldsymbol{V_2}$ 正交

- $\boldsymbol{V_1}\perp\boldsymbol{V_2}\Rightarrow\boldsymbol{V_1}\cap\boldsymbol{V_2}=\{\boldsymbol{0}\}$

- $\boldsymbol{\alpha}\perp\boldsymbol{V_1}$ 且 $\boldsymbol{\alpha}\in\boldsymbol{V_1}\Rightarrow\boldsymbol{\alpha}=\boldsymbol{0}$

3. 正交子空间与直和：两两正交的子空间的和必是直和，即若 $\boldsymbol{V_1,...,V_s}$ 两两正交，则 $\boldsymbol{V_1}\oplus...\oplus\boldsymbol{V_s}$

4.5.2 子空间的正交补

1. 子空间的正交补：若 Euclid 子空间 $\boldsymbol{V_1}, \boldsymbol{V_2}$ 满足 $\boldsymbol{V_1} \perp \boldsymbol{V_2}$ 且 $\boldsymbol{V_1} + \boldsymbol{V_2} = \boldsymbol{V}$，则称 $\boldsymbol{V_2}$ 为 $\boldsymbol{V_1}$ 的正交补，记作 $\boldsymbol{V_2} = \boldsymbol{V_1^\perp}$. 显然 $\boldsymbol{V_1^\perp} = \{\boldsymbol{\alpha} \in \boldsymbol{V} | \ \boldsymbol{\alpha} \perp \boldsymbol{V_1}\}$

2. 正交补的性质：

- n 维 Euclid 空间 $\boldsymbol{V}$ 的每个子空间 $\boldsymbol{W}$ 都有**唯一的正交补** $\boldsymbol{W}^\perp$
- $\boldsymbol{W}$ 的正交补一定是 $\boldsymbol{W}$ 的直和补（但 $\boldsymbol{W}$ 的直和补不一定是 $\boldsymbol{W}$ 的正交补）

3. 正交补的计算法则：

- $(\boldsymbol{W}^\perp)^\perp = \boldsymbol{W}$，$\dim \boldsymbol{W}^\perp + \dim \boldsymbol{W} = n$，$\boldsymbol{W}^\perp \oplus \boldsymbol{W} = \boldsymbol{V}$
- $(\boldsymbol{V_1} + \boldsymbol{V_2})^\perp = \boldsymbol{V_1^\perp} \cap \boldsymbol{V_2^\perp}$，$(\boldsymbol{V_1} \cap \boldsymbol{V_2})^\perp = \boldsymbol{V_1^\perp} + \boldsymbol{V_2^\perp}$

4. 重要子空间的正交补：对于 $\boldsymbol{A^{s\times n}}$

- 零空间是行空间的正交补：$\boldsymbol{\mathcal{N}(A)^\perp = \mathcal{R}(A^T)}$
- 左零空间是列空间的正交补：$\boldsymbol{\mathcal{N}(A^T)^\perp = \mathcal{R}(A)}$
- **（题）**域 $\mathbb{F}$ 上 n 维线性空间 $\mathbb{F}^n$ 的任一子空间 $\boldsymbol{U}$ 是某个齐次线性方程组的解空间（**本质上就是求 $\boldsymbol{U}^\perp$**）

4.6 内射影/正交投影

1. 内射影：设 $\boldsymbol{W}$ 维 Euclid 空间 $\boldsymbol{V}$ 的子空间，对任意 $\boldsymbol{\alpha} \in \boldsymbol{V}$，可唯一分解为 $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$，其中 $\boldsymbol{\alpha_1} \in \boldsymbol{W}, \boldsymbol{\alpha_2} \in \boldsymbol{W}^\perp$，

- 正交投影：$\boldsymbol{\alpha_1}$ 为 $\boldsymbol{\alpha}$ 在 $\boldsymbol{W}$ 中的正交投影，记为 $\boldsymbol{\alpha_1 = \mathcal{P}_W(\alpha) = \hat{\alpha}}$
- 正交分量：$\boldsymbol{\alpha_2}$ 为 $\boldsymbol{\alpha}$ 在 $\boldsymbol{W}$ 的正交分量，换句话说 $\boldsymbol{\alpha_2}$ 为 $\boldsymbol{\alpha}$ 在 $\boldsymbol{W}^\perp$ 的正交投影

2. 内射影变换 $\mathcal{P}_{\boldsymbol{W}}$（即 3.1 线性变换的概念中的投影变换）

$$\mathcal{P}_{\boldsymbol{W}} : \boldsymbol{V} \longmapsto \boldsymbol{W}$$

$$\boldsymbol{\alpha} \longmapsto \boldsymbol{\alpha_1},\ \boldsymbol{\alpha_1} \in \boldsymbol{W}$$

3. 正交分解定理：设 $\boldsymbol{V}$ 是 n 为 Euclid 空间，非零子空间 $\boldsymbol{W}$ 的一组正交基为 $\boldsymbol{\eta_1, ..., \eta_m}$，则 $\forall \boldsymbol{\alpha} \in V$，$\boldsymbol{\alpha}$ 在 $\boldsymbol{W}$ 中的正交投影为：

$$\boldsymbol{\mathcal{P}_W(\alpha) = \frac{(\alpha, \eta_1)}{(\eta_1, \eta_1)}\eta_1 + \frac{(\alpha, \eta_2)}{(\eta_2, \eta_2)}\eta_2 + ... + \frac{(\alpha, \eta_m)}{(\eta_m, \eta_m)}\eta_m}$$

- 如果 $\boldsymbol{\eta_1, ..., \eta_m}$ 为标准正交基，则 $\boldsymbol{\mathcal{P}_W(\alpha) = (\alpha, \eta_1)\eta_1 + (\alpha, \eta_2)\eta_2 + ... + (\alpha, \eta_m)\eta_m}$
- Schmidt 正交化的本质是：求 $\boldsymbol{\alpha_{i+1}}$ 在 $\boldsymbol{W = L(\eta_1^*, ..., \eta_i^*)}$ 上的正交分量：

$$\boldsymbol{\eta_{i+1}^* = \alpha_{i+1} - \frac{(\alpha_{i+1}, \eta_1^*)}{(\eta_1^*, \eta_1^*)}\eta_1^* - ... - \frac{(\alpha_{i+1}, \eta_i^*)}{(\eta_i^*, \eta_i^*)}\eta_i^*}$$

4. 向量到子空间的距离：$\|\boldsymbol{\delta}\| = \|\boldsymbol{\alpha - \mathcal{P}_W(\alpha)}\|$（$\boldsymbol{\alpha}$ 到子空间 $\boldsymbol{W}$ 的最佳逼近）

- 弱化地，$\boldsymbol{\alpha}$ 到 $\boldsymbol{\eta}$ 的距离 $\iff$ $\boldsymbol{\alpha}$ 到 $\boldsymbol{L(\eta)}$：$\|\boldsymbol{\delta}\| = \mathcal{P}_{\boldsymbol{W}}(\boldsymbol{\alpha}) = \boldsymbol{\frac{(\alpha, \eta)}{(\eta, \eta)}\eta}$

4.7 实对称矩阵的标准型

4.7.1 实对称矩阵与对称变换

1. 实对称矩阵 $\boldsymbol{A}$ 的特征值均为实数

2. 对称变换：$\boldsymbol{\sigma}$ 为 Euclid 空间中的线性变换，且满足对 $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \boldsymbol{V}$，$(\boldsymbol{\sigma}(\boldsymbol{\alpha}), \boldsymbol{\beta}) = (\boldsymbol{\alpha}, \boldsymbol{\sigma}(\boldsymbol{\beta}))$

- n 维 Euclid 空间 $\boldsymbol{V}$ 的对称变换与 n 阶实对称矩阵在标准正交基下是相互确定的：
  - **实对称矩阵在标准正交基下可确定一个正交变换**：设**实对称矩阵** $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$，$\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 为 $\boldsymbol{V}$ 的一组**标准正交基**，定义 $\boldsymbol{V}$ 的线性变换 $\boldsymbol{\sigma}$：

    $$\boldsymbol{\sigma}(\boldsymbol{\varepsilon_1, ..., \varepsilon_n}) = (\boldsymbol{\varepsilon_1, ..., \varepsilon_n})\boldsymbol{A}$$

    则 $\boldsymbol{\sigma}$ 即为 $\boldsymbol{V}$ 的**对称变换**

  - **对称变换在标准正交基下的矩阵为实对称矩阵**：设 $\boldsymbol{\sigma}$ 为 n 维 Euclid 空间上的**对称变换**，$\boldsymbol{\varepsilon_1, ..., \varepsilon_n}$ 为一组**标准正交基**，$\boldsymbol{A} = (\boldsymbol{a_{ij}}) \in \mathbb{R}^{n \times n}$ 为 $\boldsymbol{\sigma}$ 在这组基下的矩阵，则 $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$

3.$\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$，定义 n 维 Euclid 空间 $\mathbb{R}^n$ 上的线性变换 $\boldsymbol{\sigma}$：

$$\boldsymbol{\sigma}(\boldsymbol{X}) = \boldsymbol{A}\boldsymbol{X}$$

则对任意 $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{\boldsymbol{n}}$，$(\boldsymbol{\sigma}(\boldsymbol{X}), \boldsymbol{Y}) = (\boldsymbol{X}, \boldsymbol{\sigma}(\boldsymbol{Y}))$

4. 记 $\boldsymbol{\sigma}$ 为 Euclid 空间的对称变换，则若 $\boldsymbol{V_1}$ 是 $\boldsymbol{\sigma}-$ 子空间，那么 $\boldsymbol{V_1^{\perp}}$ 也是 $\boldsymbol{\sigma}-$ 子空间

5. 实对称矩阵属于不同特征值的特征向量是正交的：设 λ, μ 是实对称矩阵 $\boldsymbol{A}$ 的两个不同特征值，$\boldsymbol{\xi}, \boldsymbol{\zeta}$ 分别是 $\boldsymbol{A}$ 的属于 λ, μ 的特征向量（$\boldsymbol{A\xi} = \boldsymbol{\lambda\xi}$，$\boldsymbol{A\zeta} = \boldsymbol{\mu\zeta}$），则 $\boldsymbol{\xi}^{\boldsymbol{T}}\boldsymbol{\zeta} = 0$

4.7.2 实对称矩阵的正交相似对角化

回顾二次型：二次型的矩阵 $\boldsymbol{A}$ 是对称的。之前将二次型转换为规范形的方法是配方，但通过实对称矩阵的正交相似对角化，我们能更方便的将二次型转换为规范形

1. 对 $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$，总有**正交矩阵** $\boldsymbol{Q}$，使得 $\boldsymbol{Q}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}\{\lambda_1, ..., \lambda_n\}$

2. 正交相似对角化的步骤：设 $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{T}}$

- 求出 $\boldsymbol{A}$ 所有的特征值：$\lambda_1, ..., \lambda_s$（特征值重数之和为 n）
- 对每个 λ_i 解齐次线性方程组 $(\lambda_i \boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$，求出其一个基础解系 $\boldsymbol{\xi_{i1}, ..., \xi_{in_i}}$（$n_i = \dim V_{\lambda_i}$）
- 对 $\boldsymbol{\xi_{i1}, ..., \xi_{in_i}}$ Schmidt 正交化得到 $\boldsymbol{\eta_{i1}, ..., \eta_{in_i}} \Rightarrow$ 这是 $\boldsymbol{V_{\lambda_i}}$ 的一组标准正交基
- $\boldsymbol{\eta_{11}, ..., \eta_{1n_1}, ..., \eta_{s1}, ..., \eta_{sn_s}}$ 就是 $\mathbb{R}^n$ 的一组标准正交基，竖着排成 $\boldsymbol{Q}$ 即得到正交矩阵 $\boldsymbol{Q}$

注意事项

1. 使得 $\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}\{\lambda_1,...,\lambda_n\}$ 的正交矩阵 $\boldsymbol{Q}$ 不唯一

   - 可进一步要求 $|\boldsymbol{Q}|=1$：如果 $|\boldsymbol{Q}|=-1$，取正交矩阵 $\boldsymbol{S}=\mathrm{diag}\{-1,1,...,1\}$，则 $\boldsymbol{Q}'=\boldsymbol{Q}\boldsymbol{S}$ 也为正交矩阵，且 $|\boldsymbol{Q}'|=|\boldsymbol{Q}||\boldsymbol{S}|=1$，且 $(\boldsymbol{Q}')^T\boldsymbol{A}(\boldsymbol{Q}')=\boldsymbol{S}^T(\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q})\boldsymbol{S}=\mathrm{diag}\{\lambda_1,...,\lambda_n\}$

2. 可以通过 $\boldsymbol{A}=\boldsymbol{A}^T$ 的特征值刻画其正定性 **(因为正交相似的矩阵是合同的)**：不妨设 $\lambda_1\geq...\geq\lambda_n$

   - $\boldsymbol{A}$（半）正定 $\iff\lambda_n>0(\geq 0)$
   - $\boldsymbol{A}$（半）负定 $\iff\lambda_1<0(\leq 0)$
   - $\boldsymbol{A}$ 不定 $\iff\lambda_1>0,\lambda_n<0$

3. $\boldsymbol{A}=\boldsymbol{A}^T$ 的正、负惯性指数 = 正、负特征值的个数（重根按重数计算），$n-\mathrm{rank}(\boldsymbol{A})$ 是 0 为 $\boldsymbol{A}$ 特征值的重数（回顾：实对称矩阵都合同于 $\begin{pmatrix}\boldsymbol{E}_p & & \\ & -\boldsymbol{E}_{r-p} & \\ & & \boldsymbol{O}\end{pmatrix}$，其中 p 为正惯性系数，$r=\mathrm{rank}(\boldsymbol{A})$，$r-p$ 为负惯性系数）

4.7.3 实二次型的正交替换

从平面/空间解析几何的视角，$\mathbb{R}^2$ 上有心二次曲线或 $\mathbb{R}^3$ 上有心二次曲面通过坐标变换可以旋转为标准型（例如椭圆 $C:a^2x^2+b^2y^2=1$），而且坐标变换的矩阵是正交矩阵（例如 $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$），推广到 $\mathbb{R}^n$ 上，任意 n 元实二次型可通过正交线性替换替换为标准型。

1. 正定的相关性质

- 对任意 n 维向量 $\boldsymbol{X}$，$\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}>0\iff\boldsymbol{A}$ 正定 $\iff\boldsymbol{A}$ 的正惯性指数为 $n\iff\boldsymbol{A}$ 合同于正定矩阵 $\iff\boldsymbol{A}$ 合同于 $\boldsymbol{E}\iff$ 存在可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A}=\boldsymbol{C}^T\boldsymbol{C}\iff\boldsymbol{A}$ 的顺序主子式 $|\boldsymbol{P}_k|$ 全大于 $0\iff$ **$\boldsymbol{A}$ 的主子式全大于 0.**

  (注：最后一个充要条件的充分性，设 $\boldsymbol{A}_k$ 为正定矩阵 $\boldsymbol{A}$ 的主子式，只需要取 $\boldsymbol{X}=(a_{i_1i_1},0,...,a_{i_2i_2},...,a_{i_ki_k})\neq 0$，那么 $\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}=\boldsymbol{X}_k^T\boldsymbol{A}\boldsymbol{X}_k>0$，因此 $\boldsymbol{A}_k$ 正定，进而 $|\boldsymbol{A}_k|>0$。必要性是显然的，因为主子式大于 0 必然顺序主子式也都大于 0)

- 对任意 n 维向量 $\boldsymbol{X}$，$\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}<0\iff\boldsymbol{A}$ 负定 $\iff\boldsymbol{A}$ 的负惯性指数为 $n\iff\boldsymbol{A}$ 合同于负定矩阵 $\iff\boldsymbol{A}$ 合同于 $-\boldsymbol{E}\iff$ 存在可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A}=-\boldsymbol{C}^T\boldsymbol{C}\iff\boldsymbol{A}$ 的顺序主子式 $|-\boldsymbol{P}_k|=(-1)^k|\boldsymbol{P}_k|$ 全大于 $0\iff$ **$-\boldsymbol{A}$ 的主子式全大于 0.**

- 对任意 n 维向量 $\boldsymbol{X}$，$\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}\geq 0\iff\boldsymbol{A}$ 半正定 $\iff\boldsymbol{A}$ 的正惯性指数 $p=r\iff\boldsymbol{A}$ 合同于半正定矩阵 $\iff\boldsymbol{A}$ 合同于 $\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\iff$ 存在矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A}=\boldsymbol{C}^T\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{C}\iff$ **$\boldsymbol{A}$ 的主子式全大于等于 0.**

（注 1：这里“$\boldsymbol{A}$ 的顺序主子式 $|\boldsymbol{P_k}|$ 全大于等于 0 是不对的!）

（注 2：这里不要求 $\boldsymbol{C}$ 是可逆矩阵，取 $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{E_r} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{C_1}$，则 $\boldsymbol{A}=\boldsymbol{C^T C}$）

- 对任意 n 维向量 $\boldsymbol{X}$，$\boldsymbol{X^T A X}\le 0 \Longleftrightarrow$ $\boldsymbol{A}$ 半负定 $\Longleftrightarrow \boldsymbol{A}$ 的负惯性指数 $r-p=r \Longleftrightarrow \boldsymbol{A}$ 的正惯性指数 $p=0 \Longleftrightarrow \boldsymbol{A}$ 合同于半负定矩阵 $\Longleftrightarrow \boldsymbol{A}$ 合同于 $\begin{bmatrix}-\boldsymbol{E_r} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix} \Longleftrightarrow$ 存在矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{A}=\boldsymbol{C^T}\begin{bmatrix}-\boldsymbol{E_r} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{C} \Longleftrightarrow$ **$-\boldsymbol{A}$ 的主子式全大于等于 0.**

- 若 $\boldsymbol{A}$ 正定，则 $k\boldsymbol{A},\boldsymbol{A^{-1}},\boldsymbol{A^*}$ 均正定
  $\boldsymbol{A}$ 为任意 n 阶实可逆矩阵，则 $\boldsymbol{A^T A}$ 必正定
  $\boldsymbol{A}$ 为任意 $m\times n$ 阶实矩阵，则 $\boldsymbol{A^T A}$ 正定 $\Longleftrightarrow$ $\boldsymbol{A}$ 列满秩（$r(\boldsymbol{A})=n$）
  $\boldsymbol{A}$ 为任意 $m\times n$ 阶实矩阵，则 $\boldsymbol{A A^T}$ 正定 $\Longleftrightarrow$ $\boldsymbol{A}$ 行满秩（$r(\boldsymbol{A})=m$）

- **正定矩阵的介值性：**
  设 $\boldsymbol{X^T A X}$ 为一实二次型. 已知有 n 维实向量 $\boldsymbol{X_1},\boldsymbol{X_2}$，使得 $\boldsymbol{X_1^T A X_1}>0,\boldsymbol{X_2^T A X_2}<0$，则必存在 n 维实向量 $\boldsymbol{X_0}\ne\boldsymbol{0}$ 使得 $\boldsymbol{X_0^T A X_0}=0$
  （提示：通过非退化线性变化 $\boldsymbol{X}=\boldsymbol{CY}$ 化为规范形后，由条件知正惯性系数 $p>0$, 负惯性系数 $r-p>0$，取向量 $\boldsymbol{Y}=(y_1,...,y_{p+1},...,y_n)=(1,...,1,...0)$, 那么取 $\boldsymbol{X_0}=\boldsymbol{C}Y$）

- 设 $\boldsymbol{A}$ 是实对称矩阵，则当 t 充分大时，$t\boldsymbol{E}+\boldsymbol{A}$ 是正定矩阵.
  (注：因为 $t\boldsymbol{E}+\boldsymbol{A}$ 的第 k 个顺序主子式为 $\begin{vmatrix}t+a_{11} & a_{12} & ... & a_{1n}\\ a_{21} & t+a_{22} & ... & a_{2n}\\ ... & ... & ... & ...\\ a_{n1} & a_{n2} & ... & t+a_{nn}\end{vmatrix}$, 行列式最高次数为 t^n，显然当 t 充分大时 t^n 占主导，那么顺序主子式大于 0，那么 $t\boldsymbol{E}+\boldsymbol{A}$ 正定)

- （上一结论的推论）设 $\boldsymbol{A}$ 是实对称矩阵，则存在一正实数 c 使得对任意向量 X 都有 $|\boldsymbol{X^T A X}|\le \boldsymbol{c X^T X}$.
  （注：当 c 充分大时，$c\boldsymbol{E}+\boldsymbol{A}$ 和 $c\boldsymbol{E}-\boldsymbol{A}$ 都是正定矩阵）

- $\boldsymbol{A},\boldsymbol{B}$ 正定，则 $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 也正定

2. 正交线性替换：线性替换 $\boldsymbol{X}=\boldsymbol{CY}$ 中，$\boldsymbol{C}$ 为正交矩阵

3. 任一 n 元实二次型 $f(x_1,...,x_n)=\boldsymbol{X^T A X}=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}x_i x_j$，都能通过正交线性替换 $\boldsymbol{X}=\boldsymbol{CY}$ 变成标准形 $\lambda_1 y_1^2+...+\lambda_n y_n^2$，其中 $\lambda_1,...,\lambda_n$ 为 $\boldsymbol{A}$ 的全部特征值

4.8 最小二乘法

4.8.1 最小二乘法

1. 最小二乘问题：求 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n}\|\boldsymbol{Ax}-\boldsymbol{b}\|_2$，其中 $\boldsymbol{b}\in\mathbb{R}^{\boldsymbol{n}}\boldsymbol{,b}\notin\boldsymbol{\mathcal{R}(A)}$

- 写成求和形式就是：记 $\boldsymbol{A}=(a_{ij})$，$\boldsymbol{b}=(b_1,...b_n)^T$，求使得

$$\min\sum_{i=1}^{n}(a_{i1}x_1+...+a_{in}x_n-b_i)^2$$

  的最小二乘解 $\boldsymbol{x}=(\hat{x}_1,\hat{x}_2,...,\hat{x}_n)$

- 最小二乘解的推导见附录 4.7.2 最小二乘法

2. 正规方程组：$\boldsymbol{A^TAX=A^Tb}$，正规方程组的解就是最小二乘解

5 附录：详细证明

2.3.3 子空间的性质

7. 设 $\alpha_1,...,\alpha_n$ 为 n 维线性空间 V 的一组基,A 为一个 $n\times s$ 矩阵,若 $(\beta_1,...,\beta_s)=(\alpha_1,...,\alpha_n)A$,则 $\dim L(\beta_1,...,\beta_s)=$ $\mathrm{rank}A$

证：设 $A=(\eta_1,...,\eta_s)$，则 $(\beta_1,...,\beta_s)=(\alpha_1,...,\alpha_n)(\eta_1,...,\eta_s)$，设 $\mathrm{rank}A$，且 A 的列向量的极大无关组为 $\eta_{i_1},...,\eta_{i_r}$，则若

$$\begin{aligned}k_1\beta_{i_1}+...+k_r\beta_{i_r}=\begin{bmatrix}\beta_{i_1},...,\beta_{i_r}\end{bmatrix}\begin{bmatrix}k_1\\ ...\\ k_r\end{bmatrix}&=\begin{bmatrix}\alpha_1,...,\alpha_n\end{bmatrix}\begin{bmatrix}\eta_{i_1},...,\eta_{i_r}\end{bmatrix}\begin{bmatrix}k_1\\ ...\\ k_r\end{bmatrix}\\ &=\begin{bmatrix}\alpha_1,...,\alpha_n\end{bmatrix}\begin{bmatrix}k_1\eta_{i_1}+...+k_r\eta_{i_r}\end{bmatrix}=0\end{aligned}$$

由于 $\alpha_1,...\alpha_n$ 为一组基，因此只能

$$k_1\eta_{i_1}+...+k_r\eta_{i_r}=0$$

又由 A 的列向量的极大无关组为 $\eta_{i_1},...,\eta_{i_r}$，因此

$$k_1=...=k_r=0$$

这说明 $\beta_{i_1},...,\beta_{i_r}$ 线性无关，类似地，从 $\beta_1,...\beta_s$ 中**任取** β_j，那么由（其中 $k_1,...k_{r+1}$ 不全为 0），

$$k_1\eta_{i_1}+...+k_r\eta_{i_r}+k_{r+1}\eta_j=0$$

可以推出

$$k_1\beta_{i_1}+...+k_r\beta_{i_r}+k_{r+1}\beta_j=0$$

这说明 $\beta_{i_1},...,\beta_{i_r},\beta_j$ 线性相关，进而 $\dim L(\beta_1,...,\beta_s)=\mathrm{rank}(\beta_1,...,\beta_s)=r$

3.4.1 代数重数和几何重数

2. 几何重数不超过代数重数：设 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)^{r_1}...(\lambda-\lambda_s)^{r_s}$，则 $\dim \boldsymbol{V_{\lambda_i}}\le r_i, i=1,2,...,s$

证：设 $\boldsymbol{A}$ 的特征值 λ_i 的特征子空间 $\boldsymbol{W_i}$ 维数为 r，其一组基为 $\boldsymbol{e_1,...,e_r}$，将其扩充为 $\boldsymbol{V}$ 的一组基 $\boldsymbol{e_1,...,e_r}$ $\boldsymbol{f_1,...,f_n}$，令

$$\boldsymbol{P=(e_1,...,e_r,f_1,...,f_{n-r})}$$

显然 $\boldsymbol{P^{-1}}$ 可逆，且

$$\begin{aligned}\boldsymbol{P^{-1}AP}&=\boldsymbol{P^{-1}(Ae_1,...,Ae_r,Af_1,...,Af_{n-r})}\\ &=(\lambda_1\boldsymbol{P^{-1}e_1},...,\lambda_r\boldsymbol{P^{-1}e_r},\boldsymbol{P^{-1}Af_1},...,\boldsymbol{P^{-1}Af_{n-r}})\end{aligned}$$

又因为

$$\boldsymbol{E} = \boldsymbol{P}^{-1}\boldsymbol{P} = (\boldsymbol{P}^{-1}\boldsymbol{e}_1, ..., \boldsymbol{P}^{-1}\boldsymbol{e}_r)$$

因此

$$\boldsymbol{\varepsilon}_1 = \boldsymbol{P}^{-1}\boldsymbol{e}_1, \ \dots \ , \boldsymbol{\varepsilon}_r = \boldsymbol{P}^{-1}\boldsymbol{e}_r$$

进而

$$\begin{aligned}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} &= (\lambda_1\boldsymbol{\varepsilon}_1, ..., \lambda_r\boldsymbol{\varepsilon}_r, \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{f}_1, ..., \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{f}_{n-r}) \\ &= \begin{pmatrix} \boldsymbol{\lambda}_i\boldsymbol{E}_r & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}\end{aligned}$$

又因为相似矩阵有相同的矩阵多项式，因此

$$\begin{aligned}|\lambda\boldsymbol{E} - \boldsymbol{A}| &= |\lambda\boldsymbol{E} - \begin{pmatrix} \boldsymbol{\lambda}_i\boldsymbol{E}_r & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}| \\ &= \begin{vmatrix} \lambda\boldsymbol{E}_r - \lambda_i\boldsymbol{E}_r & -\boldsymbol{B} \\ \boldsymbol{O} & \lambda\boldsymbol{E}_{n-r} - \boldsymbol{C} \end{vmatrix} = |\lambda - \lambda_i|^r|\lambda\boldsymbol{E}_{n-r} - C|\end{aligned}$$

令 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$，解得的 λ 的重根数至少有 r 个（除了来自于 $|\lambda - \lambda_i|$，可能 $|\lambda\boldsymbol{E}_{n-r} - C|$ 里面也有根），因此 $r \leq r_i$，即 $\dim V_{\lambda_i} \leq r_i$

3.6.2 线性空间的分解

$\mathcal{A}$ 为 $\boldsymbol{V}$ 上线性变换，$\mathcal{A}$ 的特征多项式 $f(\lambda)$ 具有分解式

$$f(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2}...(\lambda - \lambda_s)^{r_s}$$

证明：$\boldsymbol{V}$ 可以分解为其根子空间 $\boldsymbol{V}^{\boldsymbol{\lambda}_i} = \{\boldsymbol{\xi} \mid (\mathcal{A} - \lambda_i\mathcal{E})^{r_i}\boldsymbol{\xi} = 0\}$ 的直和
证：令

$$f_i(\lambda) = \frac{f(\lambda)}{(\lambda - \lambda_i)^{r_i}} = (\lambda - \lambda_1)^{r_1}...(\lambda - \lambda_{i-1})^{r_{i-1}}(\lambda - \lambda_{i+1})^{r_{i+1}}...(\lambda - \lambda_s)^{r_s}$$

记 $\boldsymbol{W}_i = \mathrm{Im} f_i(\mathcal{A})$，则由之前结论可知 $\boldsymbol{W}_i$ 是 $\mathcal{A}-$ 子空间. 下面证明三件事：

1. $\boldsymbol{V} = \boldsymbol{W}_1 + \boldsymbol{W}_2 + ... + \boldsymbol{W}_s$
2. $\boldsymbol{V}^{\boldsymbol{\lambda}_1} \oplus \boldsymbol{V}^{\boldsymbol{\lambda}_2} \oplus ... \oplus \boldsymbol{V}^{\boldsymbol{\lambda}_s}$
3. $\boldsymbol{V}^{\boldsymbol{\lambda}_i} = \boldsymbol{W}_i$

对于 1，因为

$$(f_1(\lambda), f_2(\lambda), ..., f_s(\lambda)) = 1$$

那么存在多项式 $u_1(\lambda), u_2(\lambda), ..., u_s(\lambda)$ 使得

$$u_1(\lambda)f_1(\lambda) + u_2(\lambda)f_2(\lambda) + ... + u_s(\lambda)f_s(\lambda) = 1$$

那么

$$u_1(\mathcal{A})f_1(\mathcal{A})+u_2(\mathcal{A})f_2(\mathcal{A})+...+u_s(\mathcal{A})f_s(\mathcal{A})=\mathcal{E}$$

对任意 $\boldsymbol{\alpha}\in\boldsymbol{V}$，有

$$\begin{aligned}\boldsymbol{\alpha}&=\mathcal{E}(\boldsymbol{\alpha})\\&=(u_1(\mathcal{A})f_1(\mathcal{A})+u_2(\mathcal{A})f_2(\mathcal{A})+...+u_s(\mathcal{A})f_s(\mathcal{A}))(\boldsymbol{\alpha})\\&=u_1(\mathcal{A})f_1(\mathcal{A})(\boldsymbol{\alpha})+u_2(\mathcal{A})f_2(\mathcal{A})(\boldsymbol{\alpha})+...+u_s(\mathcal{A})f_s(\mathcal{A})(\boldsymbol{\alpha})\\&=f_1(\mathcal{A})(u_1(\mathcal{A})(\boldsymbol{\alpha}))+f_2(\mathcal{A})(u_2(\mathcal{A})(\boldsymbol{\alpha}))+...+f_s(\mathcal{A})(u_s(\mathcal{A})(\boldsymbol{\alpha}))\end{aligned}$$

由于 $f_i(\mathcal{A})(u_i(\mathcal{A})(\boldsymbol{\alpha}))\in\mathrm{Im}f_i(\mathcal{A})=\boldsymbol{W_i}$，因此 $\boldsymbol{V}\subset\boldsymbol{W_1}+...+\boldsymbol{W_s}$，进而 $\boldsymbol{V}=\boldsymbol{W_1}+...+\boldsymbol{W_s}$

对于 2，通过零元唯一分解来证明：
设

$$\boldsymbol{\beta_1}+...+\boldsymbol{\beta_s}=\mathbf{0} \tag{3}$$

其中 $\boldsymbol{\beta_i}$ 满足 $(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}(\boldsymbol{\beta_i})=0$，要证明 $\boldsymbol{\beta_i}=\mathbf{0}$
等式（3）成立时，显然有 $(\lambda-\lambda_j)^{r_j}|f_i(\lambda)$, $i\neq j$，因此存在 $h(\lambda)$ 使得 $f_i(\lambda)=h(\lambda)(\lambda-\lambda_j)^{r_j}$，于是

$$f_i(\mathcal{A})=h(\mathcal{A})(\lambda-\lambda_j)^{r_j}$$

那么

$$f_i(\mathcal{A})(\boldsymbol{\beta_j})=h(\mathcal{A})(\lambda-\lambda_j)^{r_j}(\boldsymbol{\beta_j})=h(\mathcal{A})(\mathbf{0})=\mathbf{0},\ i\neq j$$

再对（3）两边取 $\mathcal{A}$，再利用以上结果可以得到

$$f_i(\mathcal{A})(\boldsymbol{\beta_1})+...+f_i(\mathcal{A})(\boldsymbol{\beta_i})+...+f_i(\mathcal{A})(\boldsymbol{\beta_s})=f_i(\mathcal{A})(\boldsymbol{\beta_i})=\mathbf{0}$$

又因为 $(f_i(\lambda),(\lambda-\lambda_i)^{r_i})=1$，因此存在多项式 $u(\lambda),v(\lambda)$ 使得

$$\begin{aligned}&u(\lambda)f_i(\lambda)+v(\lambda)(\lambda-\lambda_i)^{r_i}=1\\\Rightarrow&u(\mathcal{A})f_i(\mathcal{A})+v(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}=\mathcal{E}\end{aligned}$$

由此，

$$\begin{aligned}\boldsymbol{\beta}=\mathcal{E}(\boldsymbol{\beta})&=(u(\mathcal{A})f_i(\mathcal{A})+v(\mathcal{A})(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}(\boldsymbol{\beta_i})\\&=u(\mathcal{A})[f_i(\mathcal{A})\boldsymbol{\beta_i}]+v(\mathcal{A})[(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}\boldsymbol{\beta_i}]\\&=u(\mathcal{A})(\mathbf{0})+v(\mathcal{A})(\mathbf{0})=\mathbf{0}\end{aligned}$$

对于 3，首先 $\boldsymbol{\alpha}\in\boldsymbol{W_i}$，存在 $\boldsymbol{\beta}\in\boldsymbol{V}$，使得 $f_i(\mathcal{A})(\boldsymbol{\beta})=\boldsymbol{\alpha}$，那么 $(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}(\boldsymbol{\alpha})=(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}f_i(\mathcal{A})(\boldsymbol{\beta})=f(\mathcal{A})\boldsymbol{\beta}$,

由 Hamiliton-Cayley 定理，$f(\mathcal{A})=\mathbf{0}$，因此 $(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}(\boldsymbol{\alpha})=\mathbf{0}$，进而 $\boldsymbol{\alpha}\in \boldsymbol{V}^{\boldsymbol{\lambda}_i}$，得到 $\boldsymbol{W}_i\subset \boldsymbol{V}^{\boldsymbol{\lambda}_i}$
再任取 $\boldsymbol{\alpha}\in \boldsymbol{V}^{\boldsymbol{\lambda}_i}$，设

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+...+\boldsymbol{\alpha}_s,\ \boldsymbol{\alpha}_i\in \boldsymbol{W}_i \tag{4}$$

即 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+...+(\boldsymbol{\alpha}_i-\boldsymbol{\alpha})+...+\boldsymbol{\alpha}_s=\mathbf{0}$，令 $\boldsymbol{\beta}_j=\boldsymbol{\alpha}_j(j\neq i)$，$\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i-\boldsymbol{\alpha}$,，那么（4）化为

$$\boldsymbol{\beta}_1+...+\boldsymbol{\beta}_i+...+\boldsymbol{\beta}_s=\mathbf{0} \tag{5}$$

由于 $\boldsymbol{\alpha}_j\in \boldsymbol{W}_j\subset \boldsymbol{V}^{\boldsymbol{\lambda}_j}$ $(j=1,2,...,s)$，所以 $\boldsymbol{\beta}_j\in \boldsymbol{V}^{\boldsymbol{\lambda}_j}$ 且 $\boldsymbol{\beta}_i=\boldsymbol{\alpha}_i-\boldsymbol{\alpha}\in \boldsymbol{V}^{\boldsymbol{\lambda}_i}$
由**第二部分证明**已知 $\boldsymbol{V}^{\boldsymbol{\lambda}_1}\oplus \boldsymbol{V}^{\boldsymbol{\lambda}_2}\oplus ...\oplus \boldsymbol{V}^{\boldsymbol{\lambda}_s}$，因此（5）的零元分解式唯一，进而 $\boldsymbol{\beta}_j=\mathbf{0}$，$j=1,2,...,s$，进而 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_i\in \boldsymbol{W}_i$，因此 $\boldsymbol{V}^{\boldsymbol{\lambda}_i}\subset \boldsymbol{W}_i$，综合两点知 $\boldsymbol{V}^{\boldsymbol{\lambda}_i}=\boldsymbol{W}_i$

最后，综合三部分证明，$\boldsymbol{V}=\boldsymbol{V}^{\boldsymbol{\lambda}_1}\oplus ...\oplus \boldsymbol{V}^{\boldsymbol{\lambda}_s}$ 得证

4.7.2 最小二乘法

求 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2$，其中 $\boldsymbol{b}\in\mathbb{R}^n,\boldsymbol{b}\notin\boldsymbol{\mathcal{R}}(\boldsymbol{A})$
解：等价于最小化

$$\begin{aligned}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2&=(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})\\&=\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{A}\boldsymbol{x}-\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{b}+\boldsymbol{b}^T\boldsymbol{b}\end{aligned}$$

求导得到

$$\frac{\partial\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2}{\partial\boldsymbol{x}}=\mathbf{2}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}-\mathbf{2}\boldsymbol{A}^T\boldsymbol{b}=0$$

即得到正规方程组

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^T\boldsymbol{b}$$

若 $\boldsymbol{A}^T\boldsymbol{A}$ 满秩，则立刻可解出最小二乘解 $\boldsymbol{x}=(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$