高等代数补充内容

本部分参考资料，网站，小文章，小技巧，高峡老师课件，对个人之前没有整理到的以及觉得值得记录的做一整理.

北京大学李帛阳

1. 如矩阵 $E-AB$ 可逆，可以求出其逆矩阵为：$E+A(E-BA)^{-1}B$（注意到利用上述性质我们可以直接找到 $E-BA$ 的逆矩阵从而证明前者是满秩的（可逆的））.

2. 考察 $m\times n$ 阶矩阵 A 与 $n\times m$ 阶矩阵 B，如其乘积为一个 $m\times m$ 阶满秩方阵，利用秩不等式有：

$$m = rank(AB) \le min\{rank(A), rank(B)\} \le m$$

从而所有的不等号只能取等，这也就证明了 A 和 B 分别是行满秩和列满秩.

3. AB 和 BA 的特征值相同，但是其特征值的几何重数不一定相同（代数重数一定相等），因为 0 特征值及其对应的特征向量空间不一定相同.

设 A 和 B 是 $n\times n$ 阶方阵。我们需要证明矩阵 AB 和 BA 的特征值相同，并且每个特征值的代数重数也相等，这相当于它们的特征多项式相等.

对于任意方阵 M，其特征多项式定义为：

$$p_M(\lambda) = \det(\lambda I - M),$$

其中 λ 是一个标量变量，I 是与 M 同阶的单位矩阵。

我们需要证明：

$$p_{AB}(\lambda) = p_{BA}(\lambda),$$

即：

$$\det(\lambda I - AB) = \det(\lambda I - BA).$$

考虑以下分块矩阵：

$$M = \begin{bmatrix} \lambda I & A \\ B & I \end{bmatrix}.$$

计算 M 的行列式时，可以使用分块矩阵的行列式公式。具体来说，如果矩阵 M 可以写成如下形式：

$$M = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

则有：

$$\det(M) = \det(P)\cdot\det(S - RP^{-1}Q),$$

其中 P 是可逆矩阵。

在我们的例子中：- $P = \lambda I$（显然是可逆的，因为 $\lambda \neq 0$）。- $Q = A$，$R = B$，$S = I$。

因此：

$$\det(M)=\det(\lambda I)\cdot\det(I-B(\lambda I)^{-1}A).$$

注意到 $(\lambda I)^{-1}=\frac{1}{\lambda}I$，所以：

$$\det(M)=\lambda^n\cdot\det\left(I-\frac{1}{\lambda}BA\right)=\lambda^n\cdot\frac{\det(\lambda I-BA)}{\lambda^n}=\det(\lambda I-BA).$$

故有：

$$\det(M)=\det(\lambda I-AB)=\det(\lambda I-BA).$$

这表明 AB 和 BA 的特征多项式相同，从而它们的特征值集合相同，且每个特征值的代数重数也相同。

4. 判断二次型是否可以分解（实数域和复数域上的分析情况类似）

注意到一个二次型可以分解相当于齐次二次多元多项式可以被分解为两个一次齐次式相乘，对应的秩至多为 2. 注意到实数域上的二次型合同于一个同阶的对角阵，对角元素为 1 或-1，于是考察正负惯性指数为

$$(1,0)\qquad(0,1)$$

时显然可以分解，考察正负惯性指数为 $(2,0)\qquad(0,2)\qquad(1,1)$ 的情形，这里不妨选择变量 x_1,x_2 为主元，利用判别式可以证明之.

5. 矩阵的相似和数域扩张无关

设 $A,B\in M_n(\mathbb{Q})$，则如果存在复的可逆矩阵 U 使得 $A=UBU^\top$，则存在可逆矩阵 $P\in M_n(\mathbb{Q})$ 使得 $A=PBP^\top$.

将 $X=(x_{ij})$ 的 n^2 个元素 x_{ij} 看做未知量，则 $AX=XB$ 可以看成 n^2 个方程的有理系数齐次方程组. 不管在 C 上还是在 Q 上按 Gauss 消元法计算，此方程组系数矩阵的行化简阶梯形都是一样的，故不论从 C 上看还是从 Q 上看，方程组的系数矩阵的秩（解空间的维数）都是一样的。

设 $U_1,\cdots,U_r$ 是子空间

$$\{X\in M_n(Q)|AX=XB\}$$

的一组基. 则 $U_1,\cdots,U_r$ 也是将 $AX=XB$ 看成复系数方程组后解空间的基，即 $AX=XB$ 在 C 上的解 X 都可以唯一地表示成

$$k_1U_1+\cdots+k_rU_r,k_i\in C$$

由题意知存在一组 $c_1,\cdots c_r\in C$ 使得

$$c_1U_1+\cdots+c_rU_r=U\qquad(|U|\neq 0)$$

于是有 $r\geq 1$.

记 $f(x_1,\cdots,x_r)=|x_1U_1+\cdots+x_rU_r|$. 展开后 f 是字符 $x_1,\cdots,x_r$ 的有理系数多项式. 由 $f(c_1,\cdots c_r)=|U|\neq 0$ 知 f 在 C^r 上不是零函数，从而 f 不是（$\mathbb{Q}^r$）上的零函数（零多项式），故存在一组解 $b_1,\cdots,b_r$ 满足 $f(b_1,\cdots b_r)\neq 0$，则对应的 $P=b_1U_1+\cdots+b_rU_r$ 满足题设.

射影空间

在向量空间 K^{n+1} 中，由全体过原点的直线（视线）构成的集合称为 n 维射影空间，记为 $P^n(K)$. 严格地，先在 K^{n+1} 中引入等价关系 $\sim$:

$$(x_0,\cdots,x_n)\sim(y_0,\cdots,y_n)\Longleftrightarrow\exists 0\neq k\in\mathbf{K},s.t.$$

$$(y_0,\cdots y_n)=k(x_0,\cdots,x_n)$$

则:

$$P^n(K):=\{K^{n+1}-\{0,\cdots,0\}\}/\sim=\{[x_0:\cdots:x_n]|x_0,\cdots,x_n\in\mathbf{K}\text{不全为}0\}$$

在 K^3 中，齐次多项式 $x^2+z^2-yz=0$ 的零点集是由一条条过原点的直线组成的（圆锥），故可看做 2 维射影空间 $P^2(K)$ 的子集，类似的，射影空间内的齐次曲线都可被称为射影曲线，注意到对任意不等于 0 的 k 都有 $[x:y:z]\sim[kx:ky:kz]$，这也从另一方面表现了射影曲线的齐次性，如此一来仿射曲线可以被看做射影曲线在子空间的投影（或者截面）. 但是仿射曲线可以代表仿射空间上的所有的曲线.

椭圆曲线

3 次光滑射影曲线经有理变换都能写成:

$$E:y^2=x^3+ax+b,\ \ a,b\in\mathbb{C}$$

这里 x^3+ax+b 无重根, 即 $4a^3+27b^2\neq 0$, 这样的曲线 E 被称为椭圆曲线. 例如, 射影曲线 $x^3+y^3=z^3$ 的点 $(x:y:z)\in P^2(K)$ 经过双有理变换:

$$x=\frac{x}{z-y},\qquad\qquad y=\frac{z+y}{z-y}$$

变成椭圆曲线 $y^2=\frac{4}{3}x^3-\frac{1}{3}$ 上的点. 反之，$(x,y)\mapsto(2x:y-1:y+1)$ 给出了这个变换的逆映射.

先给出一些复变函数上的基本概念.

全纯函数: 设 $U\subset\mathbb{C}$ 为开子集，且 $f:U\longrightarrow\mathbb{C}$ 是一个单复变函数，称 f 在 $z_0\in U$（复）可微或全纯: 如果极限

$$f'(z_0):=\lim_{z\to z_0}\frac{f(z)-f(z_0)}{z-z_0}$$

处处存在.

Cauchy-Riemann equation 复变函数 $f(z)=u(x,y)+iv(x,y)$ 对应的 C-R 方程包含两个方程:

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}$$

$$\frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}$$

如果两个方程都满足，则称这个复变函数全纯.

称一个函数为亚纯函数，如果去除若干个孤立点的集合外处处全纯（可解析）. 每个亚纯函数可以表示为两个全纯函数的比，这些孤立点的集合被称为亚纯函数的极点.

格（Lattice） 记作 Λ，是复平面上的一组离散点的集合，可以通过两个线性无关的复数 u_1, u_2 生成，具体而言：

$$\Lambda \overset{def}{=} \{mu_1 + nu_2 | m, n \in \mathbb{Z}\}$$

由此可以得到双周期函数的定义：如果函数 $f(z, \Lambda) = 0$ 满足 Λ 是一个格，则它是一个双周期函数. 特别地，对于一般的椭圆函数（曲线），经过适当的变量代换，可以得到标准椭圆曲线方程，其形式如前述定义. 给出 Weierstrass 椭圆曲线：对于 $\mathbb{C}$ 上的一个确定的格 $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, 形如下式的函数被称为 Weierstrass 函数：

$$\wp(z, \Lambda) = \frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right)$$

, 该函数满足微分方程：

$$(\wp'(z))^2 = 4\wp(z)^3 - g_2\wp(z) - g_3$$

,不难看出这其实几乎是标准椭圆方程. 于是椭圆曲线可以被表示为二维参数方程:$\Gamma(z) = f(\wp(z), \wp'(z))$，根据 $\wp(z)$ 的双周期性，$\Gamma(z)$ 和 $\mathbb{C}/\Lambda$ 有一一映射，注意到复平面关于格的商空间和环面同胚，所以椭圆曲线和环面同胚，因此亏格 $g = 1$.

椭圆曲线的加法 加法定理

$$\det \begin{pmatrix} \wp(z) & \wp'(z) & 1 \\ \wp(y) & \wp'(y) & 1 \\ \wp(z+y) & -\wp'(z+y) & 1 \end{pmatrix} = 0$$

假设 $u + v + w = 0$，上式有一个较对称的版本

$$\det \begin{pmatrix} \wp(u) & \wp'(u) & 1 \\ \wp(v) & \wp'(v) & 1 \\ \wp(w) & \wp'(w) & 1 \end{pmatrix} = 0$$

此外

$$\wp(z+y) = \frac{1}{4} \left\{ \frac{\wp'(z) - \wp'(y)}{\wp(z) - \wp(y)} \right\}^2 - \wp(z) - \wp(y).$$

魏尔斯特拉斯椭圆函数满足复制公式：若 $2z$ 不是周期，则

$$\wp(2z) = \frac{1}{4} \left\{ \frac{\wp''(z)}{\wp'(z)} \right\}^2 - 2\wp(z)$$

几何加法 规定椭圆曲线的几何加法：$\wp(z)$ 上两点 P,Q，规定 P+Q 由下述操作生成：

当 $P \neq Q$ 时，P+Q 被定义为直线 PQ 和 $\wp(z)$ 的第三点（此时要求两点不能关于 x 轴对称，即横坐标不同）关于 x 轴的对称点；当 $P \equiv Q$ 时，$P + Q$ 被定义为 $\wp(z)$ 过 P 的切线和 $\wp(z)$ 的另一个交点关于 x 轴的对称点. 进一步地，对椭圆曲线（**本身包含无穷远点 $\mathcal{O}$, 对应加法单位元**）定义加法逆元：Q 的逆元相当于 Q 关于 x 轴的对称点，于是，椭圆曲线关于几何加法为 Abel 群.

伽罗瓦 (Galois) 理论

有限域 也叫伽罗瓦域. 有限域的特征必为某个质数 p. 若 $\mathbb{F}_p$ 为特征为 p 的有限域，则 $\mathbb{F}_{p,n}$ 的阶为 p^n，其中 n 为某个正整数，则 $\mathbb{F}_{p,n}$ 的乘法群 $< \mathbb{F}_{p,n}, \cdot >$ 为循环群，即存在 $a \in \mathbb{F}_{p,n}$ 使得

$$< \mathbb{F}_{p,n}, \cdot > = < a > = \{a, a^2, \cdots, a^{q-2}, a^{q-1} = 1\}$$

这里所谓的特征就是域中最小的正整数 z 使得 z 个单位元的和等于零元，记作 $z = charF_p$，否则称特征为 0.

对特征 $p > 0$ 的域 F，有 $(u+v)^p = u^p + v^p (u, v \in F)$，若 F 是有限域，称映射 $\sigma_p : a \longrightarrow a^p$ 为 F 的 Frobenius 自同构.

扩域 如果 F 是 K 的子域，则称 K 是 F 的扩域，同时也称 K 为 F 的域扩张.(类比扩环)

定义 设 K 是 F 的扩域,$S \subseteq K$，所谓 F 上由 S 生成的域是指 K 中包含 S 的最小的 F 的扩域，记为 $F(S)$, 若 $K = F(S)$，则称 S 为 K **在 F 上的生成系**，亦称 **K 在 S 上由 F 生成**，可以由有限集合生成的 F 的扩域被称为 F 上**有限生成**的域，否则为**无限生成**的域. 可以由一个元素（组成的集合）生成的扩域被称为**单扩张**.

定义 设 E,F 为 K 的子域，称 K 中包含 E 和 F 的最小子域为 E 和 F 的**合成**，记为 EF.

设 R, R_1 是环, 映射：$\varphi : R \longrightarrow R_1$ 称为由 R到R_1 的一个**环同态**，如果 φ 保持环运算，即对于所有的 $a, b \in R$, 都有 $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$. 又如果 φ 是单（满）射，则称 φ 为**单（满）同态**. 既单又满的同态称为**同构**. 如果存在一个由 R到R_1 的同构，则称 R 同构于 R_1 也说二者是同构的，记为 $R \cong R_1$. 如果在前述定义中规定 $\varphi(ab) = \varphi(b)\varphi(a)$，则这样的映射被称为**反同态**，类似的可以定义**反同构**.

设 $\varphi : R \longrightarrow R_1$ 是环同态，$\varphi(R)$ 称为 φ 的像，记为 $Im(\varphi)$，0 的原像被称为 φ 的核，记为 $Ker\varphi$，即：$Ker\varphi = \{a \in R \,|\, \varphi(a) = 0\}$.

定义 设 $(R, +, \cdot)$ 是环，I 是 R 的加法子群，并且对于任意的 $r \in R$，都有 $rI \subseteq R$，则 I 为 R 的一个**左理想**（类似地可以定义右理想），如果某个 I 可以同时为左、右理想，则称 I 为**双边理想**，也简称**理想**.

定理 若 $\varphi : R \longrightarrow R_1$ 是环同态，则 φ 单射 $\Longleftrightarrow Ker\varphi = \{0\}$，且 $Im\varphi$ 是 R_1 的子环，$Ker\varphi$ 是 R 的理想.

定义 域的同态就是环同态，域的同构就是环同构. 由于同态的核是理想，而域只有两个理想（0 和域本身)，所以域的同态只有单同态（核为零理想）和零同态（核为域本身）两种. 域的单同态也被称为**域嵌入**，如果 K 和 E 都是域 F 的扩域,$\sigma : K \longrightarrow E$ 是一个域嵌入，并且 σ 在 F 上的限制是恒同映射，则称 σ 是一个 **F-嵌入**，进一步的可以定义 **F-同构**.

定义 设 K/F 是域扩张，$t_1, t_2 \cdots t_n \in K$，如果存在系数在 F 中的非零多项式 $f(x_1, \cdots, x_n)$，使得 $f(t_1, \cdots, t_n) = 0$，则称 $t_1, \cdots, t_n$ 在 F 上**代数相关**，否则称他们在 F 上**代数无关**. 特别地，如果 K 中一个元素 t 在 F 上是代数相关的，则称 t 是 F 上的**代数元**，K 中不是 F 上的代数元的元素称为 F 上的**超越元**，如果 K 的所有的元素都是 F 上的代数元，则称 K/F 为**代数扩张**，否则称为**超越扩张**.

命题 设 F 是域，$K = F(\alpha)$，如果 α 是 F 上的超越元，则 $F(\alpha)$ 同构于一个 F 上的一元有理分式 $F(x)$.

极小多项式 设 F, α, K 满足 $K = F(\alpha)$，F 是域，且 α 是 F 上的代数元，称 $F[x]$ 中满足 $f(\alpha) = 0$ 的次数最低的首一多项式 $f(x)$ 为 α 在 F 上的**极小多项式**，记作 $Irr(\alpha, F)$.

命题 $f(x)$ 为 α 在 F 上的极小多项式的充要条件是 $f(x)$是$F[x]$中以α 为零点的首一不可约多项式.

命题 设 F 是域, $K = F(\alpha)$, α 是 F 上的代数元, $f(x)$ 为 α 在 F 的极小多项式, 则 $K \cong F[x]/(f(x))$.

证明. 考虑映射:

$$\sigma \ : \ F[x] \longrightarrow F(\alpha)$$

$$g(x) \mapsto g(\alpha)$$

易见 σ 是同态. 于是可以证明 $F(\alpha) = F[\alpha]$，所以 σ 是满同态. 只要再证明 $ker\sigma = (f(x))$. 对于任意 $g(x) \in (f(x))$，设 $g(x) = q(x)f(x)$，有 $\sigma(g(x)) = g(\alpha) = q(\alpha)f(\alpha) = 0$，故 $g(x) \in ker\sigma$. 这说明 $(f(x)) \subseteq ker\sigma$，反之，设 $g(x) \in ker\sigma$，即 $g(\alpha) = 0$，由于 $f(x)$ 不可约，所以 $(f(x), g(x)) = 1$或$f(x)$. 如果 $(f(x), g(x)) = 1$，则存在 $u(x), v(x) \in F[x]$ 使得 $u(x)f(x)+v(x)g(x) = 1$，两端用 $x = \alpha$ 代入，得 0=1，矛盾. 从而 $(f(x), g(x)) = f(x)$，即 $f(x)|g(x)$，亦即 $g(x) \in (f(x))$，这就证明了 $ker\sigma \subseteq (f(x))$，从而命题得证. □

定义 设 K/F 是域扩张，则 K 作为 F 上的线性空间维数 $dim_F K$ 称为 K/F 的**次数**，记作 $[K : F]$，如果前者有限，则称 K/F 为**有限扩张**，否则称为**无限扩张**.

命题 设 K/E 和 E/F 都是域扩张，则 K/F 是有限扩张当且仅当 K/E 和 E/F 都是有限扩张，此时有:

$$[K : F] = [K : E][E : F]$$

(分别取出后两个扩域上的一组基即可，第一个扩域的基是后两者基两两乘积的组合)

命题 设 L/F 是域扩张，$\alpha \in L$，则 α 为 F 上的代数元的充要条件是 $[F(\alpha) : F] = degIrr(\alpha, F)$.

证明. 令 $K = F(\alpha), f(x) = Irr(\alpha, F)$，设 $degf(x) = n$.

必要性 只要证明 $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ 是 K 的 F-基。首先不难看出这 n 个元素 F-线性无关。否则存在不全为零的 $a_i \in F$ $(i = 0, 1, \ldots, n-1)$，使得 $\sum_{i=0}^{n-1} c_i\alpha^i = 0$，即 $\sum_{i=0}^{n-1} c_i x^i$ 是以 α 为零点的非零多项式，其次数小于 n。这矛盾于 $f(x)$ 是 α 在 F 上的最小多项式。其次，如命题 1.8 的证明中所述，我们有 $K = F(\alpha) = F[\alpha]$。所以对于 $\beta \in K$，可设 $\beta = g(\alpha)$，其中 $g(x) \in F[x]$。由带余除法，存在 $q(x), r(x) \in F[x]$，使得 $g(x) = q(x)f(x) + r(x)$，$r(x) = 0$ 或 $\deg r(x) < n$。于是 $g(\alpha) = q(\alpha)\cdot 0+r(\alpha) = r(\alpha)$。而 $r(x)$ 是 $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ 的 F-线性组合。这就证明了 $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ 是 K 的 F-基。

充分性 因为 $\dim_F F(\alpha) = n$，所以 $1, \alpha, \alpha^2, \ldots, \alpha^n$ 在 F 上线性相关，即存在不全为零的 $a_i \in F$ $(0 \leq i \leq n)$ 使得 $\sum_{i=0}^{n} a_i\alpha^i = 0$。故 α 是 F 上的代数元。 □

不可约元 设 p 是整环 R 上的非零元素，且 p 不是可逆元，若从 p 的任意分解 $p = ab(a, b \in R)$ 总能推出 a 或 b 之一是可逆元，则称 p 是 R 的不可约元 (只有平凡的分解).

定义 如果有理数域 $\mathbb{Q}$ 上的代数元的极小多项式的次数为 n，则称该代数元为 **n 次代数数**.

命题 K/F 是有限扩张当且仅当 K/F 是有限生成的代数扩张.

推论 设 K/F 是域扩张，则 K 中在 F 上的代数元的全体构成 K 的子域，称为 F 在 K 中的**代数闭包**.

7. **结式** 对于多 2 个多项式

$$f(x) = a_0x^m + a_1x^{m-1} + \cdots + a_m$$

$$g(x) = b_0x^n + b_1x^{n-1} + \cdots + b_n$$

则其结式 $res(f, g)$ 定义为:

$$res(f, g) = a_0^n b_0^m \prod_{\substack{1 \le i \le m \\ 1 \le j \le n}} (\alpha_i - \beta_j)$$

于是可以得到性质:(利用范德蒙行列式和多项式性质)

$$res(f, g) = a_0^n \prod_{1 \le i \le m} g(\alpha_i), \qquad \text{其中 } \alpha_i \text{ 是 f 的根}$$

$$res(g, f) = (-1)^{mn} res(f, g)$$

$$res(fg, h) = res(f, h)res(g, h)$$

若 $h \in \mathbb{C}[x], deg(f - hg) = r > 0$, 则

$$res(g, f) = b_0^{m-r} res(g, f - hg)$$

8. 多项式 $f = a_0x^m + a_1x^{m-1} + \cdots + a_m = a_0 \prod_{1 \le i \le m}(x - \alpha_i)$ 的判别式定义为

$$disc(f) = a_0^{2m-2} \prod_{1 \le i \le j \le m} (\alpha_1 - \alpha_j)^2$$

利用 $f(x)$和$f'(x)$ 的关系，我们还有:

$$a_0 disc(f) = (-1)^{\frac{m(m-1)}{2}} res(f, f')$$

8. **有限域上的子空间计数问题** 在有限域 $F_{q=p^n}$ 上，n 维线性空间 V 的 k 维子空间的个数可以用高斯系数表示:

$$N_q(k, n) = \frac{(q^n)(q^n - q) \cdots (q^n - q^{k-1})}{(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1})}$$

计数过程类似施密特正交化方法.

9. **定理** 设 $V_1, V_2, \cdots, V_s$ 是方阵 A 的特征子空间，则 $V_1 + V_2 \cdots + V_s$ 是直和.(利用 A 的 s 个特征向量分别被变换 A 作用，整理出范德蒙行列式).

10. **定理** 设 $V_i = Ker(A - \lambda_i I)^{r_i}$ 是方阵 A 的属于不同特征值 $\lambda_i (1 \le i \le s)$ 的广义特征子空间，则 $\sum V_i$ 是直和.

证明.（反证法）若有 $\alpha_i \in V_i$ 不全为 0，使得

$$\alpha_1 + \cdots + \alpha_s = 0.$$

不妨设 $\alpha_1 \neq 0$，则存在 $0 \leq k \leq r_1$，使得

$$\beta = (A - \lambda_1 I)^k \alpha_1 \neq 0, \qquad (A - \lambda_1 I)^{k+1} \alpha_1 = 0$$

于是 $\beta \neq 0$, $(A - \lambda_1 I)\beta = 0 \Longrightarrow A\beta = \lambda_1 \beta$

将算子 $(A - \lambda_1 I)^k (A - \lambda_2 I)^{r_2} \cdots (A - \lambda_s I)^{r_s}$ 作用在

$$\alpha_1 + \alpha_2 + \cdots + \alpha_s = 0$$

上，得:

$$(A - \lambda_2 I)^{r_2} \cdots (A - \lambda_s I)^{r_s} \beta \tag{1}$$

$$= (\lambda_1 - \lambda_2)^{r_2} \cdots (\lambda_1 - \lambda_s)^{r_s} \beta \neq 0 \tag{2}$$

这是一个矛盾! □

11. 矩阵可以对角化的充要条件是它可以表示成特征子空间的直和（可以直观看出来），相当于矩阵有一个零化多项式，它可以被分解成互异的一次因式的乘积.

12. **定义 外直和**：设 $V_1, V_2 \cdots, V_s$ 是 K-线性空间，则 $V_1, V_2, \cdots, V_s$ 的外直和定义为：

$$V_1 \oplus \cdots \oplus V_s = \{(\alpha_1, \cdots, \alpha_s) | \alpha_i \in V_i\}$$

同样的符号表示，进一步的，有，令 $V_i' = \{(0, \cdots, \alpha_i, \cdots, 0) \in V_i' \mapsto \alpha_i \in V_i\}$，则它是 V 的子空间且同构于 $V_i (1 \leq i \leq s)$，$((0, \cdots, \alpha_i, \cdots, 0) \in V_i' \mapsto \alpha_i \in V_i)$，不难发现 $V = V_1' \oplus \cdots \oplus V_s'$(外直和和内直和的关系).

13. 将 $m \times n$ 矩阵 A 的元素自左向右一列一列地写在一个 mn 维列向量上，所得的列向量称为矩阵 A 的列展开，记作 $cs(A)$. 类似地，还有行展开.

14. 定义矩阵 $A = [a_{ij}] \in M_{m,n}(K)$和矩阵$B \in M_{p,q}(K)$ 的 Kronecker 积（张量积），是一个 $mp \times nq$ 矩阵，它将每个 A 中的元素 a_{ij} 替换成小矩阵 $a_{ij}B$ 后得到的大分块矩阵，记为 $A \otimes B$. 显然张量积的运算几乎不满足交换律，但是满足结合律. 以下是一些简单的计算性质：

- $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
- $(A \otimes B)(C \otimes D) = AC \otimes BD$
- $(A \otimes B)^\top = A^\top \otimes B^\top$

- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
- 当矩阵 A,B 都是正交矩阵时，$A \otimes B$ 也是正交矩阵.
- $A \otimes B$ 的秩等于 A 和 B 的秩之积.
- 设 A,B 分别是 m 阶与 n 阶方阵，则有

$$|A \otimes B| = |A|^n |B|^m$$

- 设 A,B 的复特征值分别为 $\lambda_1, \cdots, \lambda_m$ 与 $\eta_1, \cdots, \eta_n$, 则 $A \otimes B$ 的特征值为 $\lambda_i \eta_j$, $A \otimes I_n + I_m \otimes B$ 的特征值为 $\lambda_i + \eta_j$.

设 $B \in M_{m,n}(K)$, 则

$$\begin{aligned}
cs(ABC) &= cs((AB)C) && (3)\\
&= (C^\top \otimes I_m) cs(AB) && (4)\\
&= (C^\top \otimes I_m)(I_n \otimes A) cs(B) && (5)\\
&= (C^\top \otimes A) cs(B) && (6)
\end{aligned}$$

15. 一个有趣的问题. 将实数域 $\mathbb{R}$ 看成 $\mathbb{Q}-$ 线性空间，证明：$1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ $\mathbb{Q}-$ 线性无关.

证明. 有问题有关的空间是 $\mathbb{Q}(\sqrt{2}, \sqrt{3})$，关键在于找到他的一组基.

记 $\theta = \sqrt{2} + \sqrt{3}$，首先证明向量组 $1, \theta, \theta^2, \theta^3$ $\mathbb{Q}-$ 线性无关.

$$\begin{aligned}
f(x) &= (x - \sqrt{2} - \sqrt{3})(x - \sqrt{2} + \sqrt{3})(x + \sqrt{2} - \sqrt{3})(x + \sqrt{2} + \sqrt{3}) && (7)\\
&= (x^2 - 2\sqrt{2}x - 1)(x^2 + 2\sqrt{2}x - 1) && (8)\\
&= x^4 - 10x^2 + 1 && (9)
\end{aligned}$$

以 θ 为根，并且 $f(x)$ 在 $\mathbb{Q}$ 上不可约，于是可以推出前述四个元在 $\mathbb{Q}$ 上线性无关.

利用同构的思想，注意到:

$$(1, \sqrt{2}, \sqrt{3}, \sqrt{6}) = (1, \theta, \theta^2, \theta^3) \begin{bmatrix} 1 & 0 & 0 & -\frac{5}{2} \\ 0 & -\frac{9}{2} & \frac{11}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{bmatrix}$$

（显然过过渡矩阵是满秩的），于是 $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ 线性无关. □

16. SVD 的几何理解：一个矩阵作用在一个单位高维球上（单位矩阵）相当于对一个高维球，在奇异值矩阵确定的主方向轴上作相应倍数的放缩，变成一个椭球，左奇异矩阵和右奇异矩阵（正交矩阵）相当于只旋转了椭球的方向（仿射变换），值得注意的是，对于高维球实际没有定义的维度无需进行仿射变换，对于不足维度的未定义伸缩也不用变换.

17. **根数**：某个整数或多项式的所有素因子的一次乘积. 记作 $rad(X)$.

整数环上的 ABC 猜想 存在绝对常数 $C>0$，使得对于每组满足条件的 $a+b=c$ 且 $(a,b,c)=1$ 的正整数 a,b,c，都有：

$$c < Crad(abc)^2$$

ABC 定理 设 $f,g,h\in K[x]$ 不全为常数，且 $f+g=h$, $(f,g,h)=1$. 则有：

$$deg\,f, deg\,g, deg\,h < deg\,(r = rad(fgh))$$

证明. 不妨设 $deg\,h \geq deg\,f, deg\,g$，容易看出 f,g,h 两两互素，我们有：

$$r = fgh/(f,f')(g,g')(h,h')$$

又由 $f+g=h$ 得 $f'+g'=h'$

于是 $f'g-fg'=f'h-fh'$

从而 $(f,f')(g,g')(h,h')|f'g-fg'$;　　fgh|(f'g-fg')r

若 $f'g-fg'\neq 0$，则：

$$\begin{aligned} deg(fgh) &\leq deg(f'g-fg') + \deg r & (10)\\ &\leq \deg(fg) - 1 + \deg r & (11)\end{aligned}$$

故 $deg\,h \leq \deg r - 1$

若 $f'g-fg'=0$，则 f 和 g 相伴，于是只能是 f,g,h 都是常数，和题设矛盾. □

例题 设 n 是大于 2 的整数，证明：

不存在 $f,g,h\in K[x], deg\,fgh>0$，使得

$$f^n+g^n=h^n, \qquad (f,g,h)=1$$

18. **多项式的可约性**.

- 多项式在复数域的可约性.
  显然 n 次多项式在复数域上一定有 n 次根，因此复数域上的不可约多项式一定是一次式.
  Lemma 复多项式 $f(x)$ 的实值函数 $|f(x)|$ 一定能够在复平面 $\mathbb{C}$ 上取到最小值.
  设 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n = a_0x^n\left(1+\frac{a_1}{a_0x}+\cdots+\frac{a_n}{a_0x^n}\right) = a_0x^n(1+b(x))$
  取 R 足够大，可使对 $\forall|x|>R$, 有$|b(x)|<\frac{1}{2}$，进而 $|f(x)|\geq\frac{1}{2}|a_0|R^n(\geq|f(0)|)$. 故在有界闭集 $|x|\leq R$ 上，连续函数 $|f(x)|$ 可取到 $\mathbb{C}$ 上的最小值.
  代数基本定理（复数域上，次数大于等于 1 的多项式都能被唯一分解成一次因式的乘积）的证明：(反证法)，假设 $f(x)$ 次数 ≥ 1 且没有复根，则 $|f(x)|$ 必在某点 x_0 取到复平面上的最小值

$|f(x_0)| > 0$. 可乘适当复数，不妨设 $f(x_0) = 1$，记 $y = x - x_0$，则

$$f(x) = 1 + c_k y^k + \cdots + c_n y^n \qquad (c_k \neq 0, k \geq 1) \tag{12}$$
$$= 1 + c_k y^k (1 + d(y)) \tag{13}$$

取 $\varepsilon > 0$ 足够小，可使当 $|y| < \varepsilon$ 时，$|d(y)| < \frac{1}{3}$
令 $x = x_0 + y = x_0 + \varepsilon e^{iw}$，此时 $f(x)$ 相当于绕 $f(x_0) = 1$ 转了 k 圈，记 $c_k = \rho e^{i\theta}, \rho > 0, 0 \leq \theta < 2\pi$，取 $y = re^{i(\pi-\theta)/k}$ 使得 $0 < r < \varepsilon$ 且 $r^k\rho < 1$. 则有：

$$|f(x)| = |f(x_0 + y)| \tag{14}$$
$$= |1 - \rho r^k (1 + d(y))| \tag{15}$$
$$\leq 1 - \rho r^k + \rho r^k/3 < 1 \tag{16}$$

这与 1 是 $|f(x)|$ 在复平面上的最小值矛盾!

- 在实数域上，非零多项式都能被唯一的写成一次因式与判别式 <0 的二次因式的乘积.
- 称各项系数最大公因数为 1 的整系数多项式为**本原多项式**.
  Gauss 引理 本原多项式的乘积仍是本原多项式.(考察某个素数 p 整除多项式 f，g 的积式的每一系数，分别考察 f 和 g 的不能被 p 整除的最高次项系数，然而它们应该被 p 整除，和 f,g 本原矛盾！)
  Eisenstein 判别法 某个素数不整除多项式 $f(x)$ 的最高次项，它的平方不整除常数项，它可以整除所有其他项系数，则该多项式在 $\mathbb{Q}[x]$ 不可约.
  例题 证明分圆多项式 $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ 不可约，这里 p 是一个素数.
  (考察 $f(x+1)$ 的 Eisenstein 判别法)
- **定理** 若 R 是唯一分解整环（UFD），则 $R[x_1, \cdots, x_n]$ 也是唯一分解整环.

19. **皮亚诺公理系统** 若集合 $\mathbb{N}$ 包含一个特殊元素 0，并存在一个 $\ltimes$ 到自身的映射 $f : n \longrightarrow n^+$（后继映射），满足：

1 f 是单射;
2 0 不在 f 的像集内;
3 若 A 是 $\mathbb{N}$ 的一个包含 0 的子集，且满足条件：若 n 属于 A，则 n^+ 亦属于 A，则必有 $A = \mathbb{N}$

则称 $\mathbb{N}$ 为**自然数集**，$1 = 0^+, 2 = 1^+, \cdots$，其中公理 3 被称为**数学归纳法**.

20. 快速判断一个数 n 是不是素数：取 $(a, n) = 1$，看是否有 $a^{n-1} \equiv 1 \mod n$. 是则肯定否则不是.

21. 定理：若 $V = V_1 \oplus V_2 \cdots \oplus V_s$，则在 V 上存在投影变换 $P_1, P_2, \cdots P_s$，s.t.:

- $P_i P_j = 0,\ \forall i \neq j$;
- $P_1 + \cdots + P_s = I$;
- $Im\, P_i = V_i,\ \forall i$

反之亦成立 (可以利用直和表示的 0 向量表示法唯一).

22. （复合映射的像空间维数公式） 若 $\mathcal{B}: U \longrightarrow V$ 与 $\mathcal{A}: V \longrightarrow W$ 是线性映射，则有：

$$\dim Im\mathcal{B} = \dim Im(\mathcal{A}\mathcal{B}) + \dim(Im\mathcal{B} \cap Ker\mathcal{A})$$

推论：设矩阵 A,B 满足：A 的列数 =B 的行数 =n，则有：

$$rank(A) + rank(B) - n \le rank(AB)$$

等号成立当且仅当 A 的解空间包含于 B 的列空间.

23. 定理：设 $\mathcal{A} \in Hom(U, V)$ 在 U 的基 $\{\alpha_i\}$ 与 V 的基 $\{\beta_j\}$ 下的矩阵是 A，即：

$$\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\beta_1, \beta_2, \cdots, \beta_m)A$$

设 U 基 $\{\alpha_i\}$ 到 $\{\alpha_i'\}$ 的过度矩阵为 P，V 基 $\{\beta_j\}$ 到 $\{\beta_j'\}$ 过度矩阵是 Q，则 $\mathcal{A}$ 在新基 $\{\alpha_i'\}$ 和 $\{\beta_j'\}$ 下的矩阵为 $Q^{-1}AP$.

设 $\mathcal{A}: U \longrightarrow V$ 是线性映射，则存在一组 U 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 V 的一组基 $\beta_1, \cdots, \beta_m$，使得 $\mathcal{A}$ 在这两组基下的矩阵为

$$A = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

上述定理的两个向量空间的基底的构造方法：

取 $Im\mathcal{A}$ 的一组基 $\beta_1, \beta_2, \cdots, \beta_r$，任取（原像）$\alpha_i \in U$，使得：$\mathcal{A}\alpha_1 = \beta_1, \cdots, \mathcal{A}\alpha_r = \beta_r$，再任取 $Ker\mathcal{A}$ 的一组基 $\alpha_{r+1}, \cdots, \alpha_n$. 由线性映射基本定理，$U/Ker\mathcal{A} \cong Im\mathcal{A}$，$\alpha + Ker\mathcal{A} \mapsto \alpha$.

由于 $\mathcal{A}\alpha_1, \cdots, \mathcal{A}\alpha_r$ 是 $Im\mathcal{A}$ 的一组基，对应的，$\alpha_1 + Ker\mathcal{A}, \cdots, \alpha_r + Ker\mathcal{A}$ 构成 $U/Ker\mathcal{A}$ 的基，再由商空间基的定理，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成 U 的基，将 $\beta_1, \cdots \beta_r$ 扩成 V 的基 $\beta_1, \cdots, \beta_m$，则有：

$$\mathcal{A}(\alpha_1, \cdots, \alpha_n) = (\beta_1, \cdots, \beta_m)\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

推论： 线性空间 V 上的任意线性变换 $\mathcal{A}$ 都能写成 $\mathcal{A} = PU$，其中 P 是投影变换，U 是可逆变换.(考察 V 上的基和 A 作用后的自空间的基底，可以现将子空间的基底扩成 V 的基，对应的原基组到这组新基的矩阵就是满秩变换，新基到子空间的基是投影变换)

如果线性映射的作用对象是基向量的话对应的线性变换的矩阵是右乘，反之如果对应的是向量坐标则是左乘.

线性变换的秩, 迹, 行列式及特征多项式就是它在任意一个基底下的矩阵的秩, 迹, 行列式及特征多项式.

24. 设 $\mathcal{A} \in \mathrm{Hom}(V)$，$\lambda \in K$. 若 $\mathrm{Ker}(\mathcal{A} - \lambda I) = \{0\}$，则 λ 不是 $\mathcal{A}$ 的特征值；若 $\mathrm{Ker}(\mathcal{A} - \lambda I) \neq \{0\}$，则 λ 是 $\mathcal{A}$ 的特征值，$\mathrm{Ker}(\mathcal{A} - \lambda I)$ 称为 $\mathcal{A}$ 的特征子空间.$\mathrm{Ker}(\mathcal{A} - \lambda I) \subseteq \mathrm{Ker}((\mathcal{A} - \lambda I)^2) \subseteq \cdots$ 称为 $\mathcal{A}$ 的广义特征子空间。

25. 矩阵（线性变换）可以对角化的充要条件：

$$
\begin{aligned}
V &= \ker(A-\lambda_1 I)\oplus\ker(A-\lambda_2 I)\oplus\cdots\oplus\ker(A-\lambda_s I)\quad V\ \text{是}\ A\ \text{的特征子空间},\\
&= \ker((A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_s I)),\\
&\iff A\ \text{有}\ n\ \text{个线性无关的特征向量},\\
&\iff A\ \text{的不同特征子空间的维数之和等于}\ n,\\
&\iff A\ \text{有一个零化多项式（最小多项式），能在}\ \mathbb{K}\ \text{上分解成互异的一次因式的乘积},\\
&\iff A\ \text{的每个特征值的代数重数等于几何重数}.
\end{aligned}
$$

26. 设 $A\in Hom(U,V)$, $B\in Hom(V,U)$，则有：

$$AB\in Hom(U)\ ,\ BA\in Hom(V)$$

有相同的非零特征值（且有相同的几何重数和代数重数）.

证明. 设 λ 是 AB 的一个特征值，则存在 $u\in U\neq 0$ 使得：$ABu=\lambda u$，令 $v=Bu\in V$，则有 $v\neq 0$，否则和 $u\neq 0$ 矛盾！于是 $BAv=BA(Bu)=B(\lambda u)=\lambda v\in V$，从而 λ 也是 BA 的特征值. 同理到另一侧.

另一方面,为了证明 AB 和 BA 关于每个特征值的代数重数一样,只需 $dim\,Ker(AB-\lambda I_U)=\dim Ker(AB-\lambda I_V)\qquad u\mapsto Bv$. 为此，构造两个核空间上的线性映射 Φ，往证 Φ 是双射，先证明它是单射，只需证明 $Ker\Phi=0$，这是显然的；再证明它是满射，对任意 $v\in\mathrm{Ker}(\mathcal{B}\mathcal{A}-\lambda I_V)$，存在 $u=\frac{1}{\lambda}\mathcal{A}v\in U$，使得：

$$\Phi(u)=\mathcal{B}\left(\frac{1}{\lambda}\mathcal{A}v\right)=\frac{1}{\lambda}\mathcal{B}\mathcal{A}v=\frac{1}{\lambda}\cdot\lambda v=v.$$

进一步，$u\in\mathrm{Ker}(\mathcal{A}\mathcal{B}-\lambda I_U)$：

$$\mathcal{A}\mathcal{B}u=\mathcal{A}\left(\mathcal{B}\cdot\frac{1}{\lambda}\mathcal{A}v\right)=\mathcal{A}\left(\frac{1}{\lambda}\mathcal{B}\mathcal{A}v\right).$$

代入 $\mathcal{B}\mathcal{A}v=\lambda v$，得：

$$\mathcal{A}\mathcal{B}u=\mathcal{A}\left(\frac{1}{\lambda}\cdot\lambda v\right)=\mathcal{A}v=\lambda\cdot\frac{1}{\lambda}\mathcal{A}v=\lambda u.$$

因此 $\mathcal{A}\mathcal{B}u=\lambda u$，即 $u\in\mathrm{Ker}(\mathcal{A}\mathcal{B}-\lambda I_U)$，从而 Φ 是满射.

对于代数重数相同,只需广义特征子空间对应维数相同: $\dim Ker(AB-\lambda I_U)^k=\dim Ker(BA-\lambda I_V)^k$，这里 K 足够大（$(AB-\lambda I_U)^k\neq 0$且$(BA-\lambda I_V)^{k+1}=0$）. □

定理 设 A 是 K-线性空间 V 上的线性变换，若在 $K[x]$ 中有分解：

$$f(x)=f_1(x)\cdots f_s(x)\ ,\ f_i(x)\text{两两互素}$$

则有：$Ker\,f(A)=Ker\,f_1(A)\oplus\cdots\oplus Ker\,f_s(A)$.

定理 相似矩阵具有相同的最小多项式.(零化多项式的集合).

27. **交换矩阵的性质**

1. 如 A,B 是可交换矩阵，它们必有公共特征向量.(这是因为 A 的所有特征子空间必然都是 B 的不变空间，对 B 同理到另一侧)

2.(2 的推广)，见 Schur 定理第一条.

3. 交换矩阵一般和根子空间上作用的不变性有关, 如 $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$, 对角块 $A_i = \begin{pmatrix} \lambda_i & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_i \end{pmatrix}$ 都是上三角矩阵，且 λ_i 互不相同. 则若矩阵 B 和 A 可交换，则 B 和 A 有相同的分块形式.(这是因为可以把 A 作用在根子空间的直和上，每个限制就是一个分块，易见 B 在每个分块上是线性变换，于是 B 在这组基下也是分块矩阵. 也可以认为是 B 和 A 是线性变换，它们在这组基下可以同时准对角化.)

28. **Sylvester 变换** 设 A, 分别是 m 阶与 n 阶复方阵，则 $S:\ X \longrightarrow AX - XB$ 是矩阵空间 $M_{m,n}(C)$ 上的 Sylvester 变换.

- S 可逆当且仅当方阵 A,B 没有公共的复特征值；

  证明. 若 A, B 有公共复特征值 λ，设

  $$\begin{aligned} A\alpha &= \lambda\alpha, \\ B^T\beta &= \lambda\beta, \quad \alpha, \beta \neq 0. \end{aligned}$$

  则 $X = \alpha\beta^T \neq 0$，而

  $$SX = A\alpha\beta^T - \alpha\beta^T B = 0.$$

  由此知 S 不是单射，故不可逆。
  反之，设 A, B 无公共复特征值。若有矩阵 X 使得 $AX = XB$，则

  $$A^2X = A(AX) = A(XB) = (AX)B = XB^2.$$

  类似地，有 $A^kX = XB^k$. 更一般地，对任意多项式 $f(x)$，有 $f(A)X = Xf(B)$.
  取 $f(x) = (x-\lambda_1)\cdots(x-\lambda_m)$ 为 A 的特征多项式. 则 $f(A) = 0$ 而 $f(B) = (B-\lambda_1 I)\cdots(B-\lambda_m I)$ 可逆.

  $$0 = f(A)X = Xf(B) \quad \Rightarrow \quad X = 0.$$

  由以上推导知线性变换 S 是单射，故也是满射。 □

- **Sylvester 方程** 方程最初的面貌是 $AX - XB = 0$，其中矩阵 A，B 分别是 $\mathbb{R}^{m\times m}$和$\mathbb{R}^{n\times n}$ 上的方阵，更一般的 Sylvester 方程是 $AX + XB = C$，其中 C 是 $\mathbb{R}^{m\times n}$ 上的矩阵。求解：

  $$\begin{aligned} & AX + XB = C && (17) \\ \iff & (I_n \otimes A + B^\top \otimes I_m)\, cs(X) = cs(C) \qquad (cs\text{表示矩阵的列展开}) && (18) \\ & && (19) \end{aligned}$$

- 若 A,B 可以对角化，则 S 也可以对角化.

  证明. 设 $\{a_i\}$ 是 A 的特征向量集合，$\{\lambda_i\}$ 是每个特征值对应的特征值，同理对 B 定义 $\{b_j\}$ 和 $\{\mu_j\}$，考察矩阵 $X_{ij} = a_i b_j^\top$，有：$AX_{ij} = A(a_i b_j^\top) = (Aa_i)b_j^\top = \lambda_i a_i b_j^\top$，$X_{ij}B = (a_i b_j^\top)B = a_i(\mu_j b_j^\top) = \mu_j a_i b_j^\top$，于是，$SX_{ij} = \lambda_i a_i b_j^\top - \mu_j a_i b_j^\top = (\lambda_i - \mu_j)X_{ij}$，这表明 X_{ij} 是 S 的特征向量，对应的特征值为 $\lambda_i - \mu_j$.
  注意到 A,B 都可对角化，因此他们的特征值的外积可以构成 $\mathbb{R}^{m\times n}$ 的基底 $\{X_{ij}\}$，恰好是 S 的特征值，因此 S 可以对角化. □

- **Lyapunov 方程** $\mathbb{R}^{n\times n}$ 上的矩阵 A,W,X 满足：

$$AX + XA^\top = W$$

  Bartels-Stewart 算法指出，将 ASchur 分解为 $A = Q\Lambda Q^\top$，其中 Λ 是上三角矩阵，则方程等价于 $Q^\top WQ = \Lambda Y + Y\Lambda^\top$，后续可以将矩阵分块简化计算. 事实上，这种算法还可以被用于求解 Sylvester 方程.

-

29. **Frobenius 矩阵（友矩阵）** 设多项式 $f(t) = t^n + a_1 t^{n-1} + \cdots + a_{n-1}t + a_n$，则 n 阶矩阵

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix}$$

称为多项式 $f(t)$ 的友矩阵 (伴侣矩阵)，用友矩阵构造的分块对角矩阵被称为矩阵的有理标准型.

性质：

- 友矩阵的特征多项式就是它对应的伴侣多项式.
- 友矩阵的最小多项式等于特征多项式.

  证明. 考察标准正交基 $\{\varepsilon_i\}$,从而有 $A\varepsilon_1 = \varepsilon_2$, $A\varepsilon_2 = \varepsilon_3$, $\cdots$, $A\varepsilon_{n-1} = \varepsilon_n$,于是 $\{\varepsilon_1, A\varepsilon_1, \cdots, A\varepsilon_{n-1}\}$ 线性无关，因此 $deg(m(\lambda)) > n-1$，这里 $m(\lambda)$ 是 A 的最小多项式（零化条件推出），于是必有 $m(\lambda) = f(\lambda)$，这里 $f(\lambda)$ 是 A 的特征多项式. □

- 对一个线性变换 A，若其最小多项式是 $m(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$，则可以找到向量 v 使得 $V = span\{v, Av, A^2v, \cdots, A^{n-1}v\}$ 线性无关（否则与最小多项式的次数违背），从而 $A|_V$ 的矩阵为 $m(x)$ 对应的友矩阵.(有理标准块)
  $\Longrightarrow$（**推论**）：对 V 上的一个线性变换 A，记 A 的最小多项式为 $m(x)$，则存在一个 A 在 V 上的不变子空间 W 使得 $A|_W$ 的矩阵为 $m(x)$ 的友矩阵.

称 v 维子空间 $<\beta, A\beta, \cdots, A^{v-1}\beta> \stackrel{def}{=} Z(\beta, A)$ 为 β 生成的 A-循环子空间，这个循环子空间由 A 关于 β 的最小多项式决定. 若 $A \in M_n(K)$，则存在 $y_1, \cdots, y_s \in K^n$ 使得 $Z(y_1, A) \oplus \cdots \oplus Z(y_s, A) = K^n$，且 A 在这些循环子空间 $Z(y_i, A)$ 上的最小多项式 $d_i(x)$ 满足：$d_1(x)|d_2(x)|\cdots|d_s(x)$. 它们由 A 唯一

决定，称为 A 的**不变因子**. 把 A 限制在每个循环子空间内，得到对角分块矩阵 $B=\begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_s \end{bmatrix}$.

称 B 为 A 的**有理标准型**. 这样的 B 在不考虑对角块顺序的情况下是唯一的.

证明. 利用 Smith 标准型,知存在可逆 x-矩阵 $P(x),Q(x)$,使得:$P(x)(xI-A)Q(x)=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & d_1(x) & & \\ & & & \ddots & \\ & & & & d_s(x) \end{pmatrix}$

我们记 x-矩阵 $Q(x)$ 第 i 个列向量为 Q_i，对 A 的每个不变因子 $d_i(x)$，我们将 Q(x) 中与之对应的列拆写成：$Q_{n-s+i}\equiv\beta_1^{(i)}+x\beta_2^{(i)}+\cdots+x^{v_i-1}\beta_{v_i}^{(i)}\mod d_i(x)$，其中 $v_i=\deg d_i(x)$，并用 γ_i 表示最后一个列向量 $\beta_{v_i}^{(i)}$. 下面只需证明 $d_i(A)\gamma_i=0$且$\gamma_1,A\gamma_1,\cdots\gamma_s,A\gamma_s,\cdots,A^{v_s-1}\gamma_s$K-线性无关. 证明不难，直接计算即可. □

从上述可以推出有理标准型的存在性和唯一性.

有理标准型的每个对角块都是对应位置的不变因子的友阵.

30. **根子空间分解（主分解定理）** 若 A 的特征多项式在域 K 上有分解

$$f_A(x)=(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_s)^{n_s}$$

则 $V=V_1\oplus V_2\cdots\oplus V_S$. 其中 $V_i=Ker(A-\lambda_i I)^{n_i}$ 称为 λ_i 的**根子空间**,$\dim V_i=$ 特征值 λ_i 的代数重数 n_i.

考察**广义特征子空间升链**：$\{0\}\subsetneq Ker(A-\lambda I)\subsetneq(A-\lambda I)^2\subsetneq\cdots\subsetneq Ker(A-\lambda I)^{r_\lambda}=Ker(A-\lambda I)^{r_\lambda+1}=\cdots$，称开始稳定的广义特征子空间为根子空间. 广义特征子空间的升链的增大幅度是不断减小的，证明如下：利用 Frobenius 不等式，

$$rank(AB)+rank(BC)\le rank(ABC)+rank(B)$$

在上式取 $A=C$ 为方阵且 $B=A^{k-1}$，有：$rank(A^{k-1})-rank(A^k)\ge rank(A^k)-rank(A^{k+1})$，再利用解空间性质：$rank(X)=\dim V-\dim Ker(X)$，即可.

又： 考察 A 在根子空间 $V_\lambda=Ker(A-\lambda I)^{r_\lambda}$ 上的限制，$A|V_\lambda$ 的最小多项式为 $(x-\lambda)^{r_\lambda}$，$A|V_\lambda$ 的特征多项式为 $(x-\lambda)^{\dim V_\lambda}$

又： 对于前述 V 的子空间分解，不难证明对于所有的 V 上的 A-不变子空间 W，都有 $W=(W\cap V_1)\oplus(W\cap V_2)\oplus\cdots\oplus(W\cap V_S)$，即 W 也可以表示成子空间的直和.(可以利用投影变换证明)

根子空间的维数为代数重数，特征子空间的维数为几何重数$\Longrightarrow$ 应用：利用广义子空间升链的维数性质推导最小多项式.

约当标准型的存在性：设 V 是有限维 K-线性空间, $\mathcal{A}\ HomV$. 若 $\mathcal{A}$ 的特征多项式在域 K 上能分解成一次因式的乘积, 则存在 V 的一组基, 使得 $\mathcal{A}$ 在此基下的矩阵为若当形矩阵.（对 $\mathcal{A}$ 主分解，每个广

义特征子空间都可以找出 A 限制在该子空间上的幂零矩阵，将所以的根子空间分解成幂零变换的直和 => 它们显然在自己的循环基下是 Jordan 标准型）. 循环基降幂排列得到的就是标准型的上三角形，反之则反. 特征值为 λ 的约当块，$\geq t$ 阶块数可以表示为 $Ker(A-\lambda I)^t$.

【例】 对一般多项式 $g(x) \in K[x]$, $Ker\, g(A)$ 是 A 的不同特征值的广义特征子空间的直和.(将 g(x) 写作 $(x-\lambda_1)^{e_1}\cdots(x-\lambda_S)^{e_S}h(x)$，其中 h(x) 和 A 的最小多项式 $m(x)$ 互质，利用裴蜀定理有 $\exists u(x), s.t. u(A)h(A)=1$，从而 h(A) 可逆，于是 $Ker\, h(A) = Ker\,(A-\lambda_1)^{e_1} \oplus \cdots \oplus Ker\,(A-\lambda_S)^{e_S} \oplus Ker\, h(A)$，且 $Ker\, h(A)=\{0\}$).

31.最小多项式确定 $K[A]$ 的结构:

为说明这点我们先给出 2 个引理:

1. 在交换环 R 中，理想 I 是极大理想当且仅当商环 R/I 是一个域;

(充分性) 对任意非零元 $a+I \in R/I$，若其无逆元，考察 I 的扩张理想 $J = I+(a)$，即 $J=\{i+ra|\ i\in I, r\in R\}$，利用 I 是极大理想，有 $J=R$，利用域含单位元，有：$\exists i \in I$ 使得 $i+ra=1$，因此每个 I 中的元素都有逆元，$\Longrightarrow R/I$ 是域.

(必要性) 利用 R/I 是域，它的理想只有 0 和自身可证.

2. 若 $p(x)\in F[x]$ 是不可约多项式，则商环 $F[x]/(p(x))$ 作为 F-向量空间的维数是 $\deg(p)$.(这是显然的)

结合 1、2，注意到不可约多项式的生成理想是极大理想，因此 $K[A]$ 是域等价于 $m(x)$ 即 A 的最小多项式在 K 上不可约. 同时有：

$$K[A] \cong K[x]/(m(x))$$

32. 分块矩阵的最小多项式：

$$g(A) = \begin{bmatrix} g(A_1) & & \\ & \ddots & \\ & & g(A_S) \end{bmatrix} = 0$$

当且仅当每个分块为 0；其最小多项式为对每个分块矩阵的最小多项式的不同因式，取所有分块中的次数最高者，并相乘. 应用：Jordan 型的最小多项式；幂零矩阵 A 的最小多项式（Jordan 块），$\Longrightarrow A^{rank(A)}=0$.

33. **一些线性变换与约当标准型**:

投影变换

投影变换一定是幂等变换.$\Longrightarrow$ 正交投影.

【例】 若实矩阵 A 列满秩，则映射 $X \mapsto A(A^\top A)^{-1}A^\top X$ 是欧氏空间到 A 列空间的正交投影. 若 A 是实对称矩阵且幂等, 则映射 $X\mapsto AX$ 是欧氏空间到 A 列空间的正交投影.

幂等变换

定理: 设 $\mathcal{P}_1,\mathcal{P}_2,\cdots,\mathcal{P}_S$ 是 V 上线性变换，$\mathcal{P}=\mathcal{P}_1+\cdots+\mathcal{P}_S$，则 $\mathcal{P}_1,\cdots,\mathcal{P}_S$ 是两两正交的幂等变换当且仅当 $\mathcal{P}$ 是幂等变换，且

$$rank(\mathcal{P})=\sum_{i=1}^{S} rank(\mathcal{P}_i)$$

幂等变换和投影变换：假设 A 是幂等变换，我们可以把 A 看成 A 沿 $Ker\,A$ 在 $Im\,A$ 上的投影变换.

命题 设 K 是数域，V 是有限维 K-线性空间，则 V 上的线性变换 A 都能写成有限个 V 上的投影变换的 K-线性组合.(直接拆分成单位基矩阵即可)

定理 n 级矩阵 A 可以写成 $n+1$ 个幂等矩阵的线性组合.

幂等变换和投影变换 若 P 是线性空间 V 上的幂等变换，即 $P=P^2$，则 $V=Ker\,(P-I)\oplus Ker\,P$. 此时 P 可以看做 $Ker\,P$ 向 $Ker\,(P-I)=Im\,P$ 的投影变换.

幂零变换

非零向量利用幂零变换生成的循环子空间被称为强循环子空间，次数最大者被称为强循环子空间的尾项. 有性质：

- 强循环子空间间是直和当且仅当循环基的尾项线性无关.
- 考察 k 维循环基 $\alpha,B\alpha,\cdots,B^{k-1}\alpha\quad(B^k\alpha=0)$ 的尾项 $B^{k-1}\alpha\in Ker\,B\cap ImB^{k-1}=W_k$，则有 $\dim W_k=\dim Im\,B^{k-1}-\dim Im\,B^k$(考虑线性映射 φ：$Im\,B^{k-1}\longrightarrow Im\,B^k$；$\alpha\mapsto B\alpha$，得到线性空间的同构：$Ker\,\varphi=Im\,B^{k-1}\cap Ker\,B=W_k$，于是 $Im\,B^{k-1}/W_k\cong Im\,B^k$).

定理 设 B 是线性空间 V 上的幂零变换，$s=\dim Ker\,B$，则存在 $\alpha_1,\alpha_2,\cdots,\alpha_S\in V$ 使得：$V=Z(\alpha_1,B)\oplus\cdots\oplus Z(\alpha_1^{(t-1)},B)$.(证明只需要考察 1.V 中的所有向量都能用列出的子空间表示；2. 子空间是直和即其维数和等于 V 的维数.)

【例】 设 A,B 是 V 上的线性变换，且 A 与 $C=AB-BA$ 可交换，则 C 是幂零变换.（找到一组基使得 C 在这组基下化为若当标准型，利用准对角矩阵的可交换矩阵的充要条件是自身也是准对角矩阵可以 A 可以在这组基下化为和 C 同分块的准对角阵. 将 B 表示为类似地分块矩阵，则 $C=AB-BA$ 的对角分块矩阵的 trace 都是 0，结合 C 是约当标准型知 C 是幂零矩阵.）

幂零矩阵有如下性质：

- 幂零矩阵 A 的秩为 r，则 $A^r=0$.
- $|I_n-A|=1$，利用特征多项式代入 $\lambda=1$.
- 幂零矩阵的迹一定为 0. 幂零矩阵 Jordan 标准型和原矩阵同迹，因此为 0.

正交变换和酉变换 内积空间上的变换 P 是正交变换，当它满足如下性质：$(P\alpha,P\beta)=(\alpha,\beta)$. 可见正交变化不改变模长. 正交变换的充要条件是它在标准正交基下的矩阵是正交矩阵. 有性质：

1. 正交矩阵的特征值模长为 1（不妨正交矩阵为 P，它的一个特征向量为 α，且 $P\alpha=\lambda\alpha$，于是两边取范数，由正交变换不改变模长即证）.

2. 正交变换在某组标准正交基下可以写成 $diag\left(\begin{bmatrix}cos\theta_1 & -sin\theta_1\\ sin\theta_1 & cos\theta_1\end{bmatrix},\cdots,\begin{bmatrix}cos\theta_2 & -sin\theta_2\\ sin\theta_2 & cos\theta_2\end{bmatrix},\pm1,\cdots,\pm1\right)$

3. 行列式为 1 的正交变换为第一类正交变换，表示旋转 θ 角；行列式为-1 的正交变换称为第二类正交变换，表示关于 $\frac{\theta}{2}$ 的反射.

4. **三维旋转正交变换** 设 $\alpha = [a_1, a_2, a_3]^|top$ 是单位列向量，设 A 是绕 α 旋转 θ 角的正交变换，求 A 在标准基下的矩阵 A. 一种方法是用 α 的外积矩阵表示，$C = \begin{bmatrix} 0 & a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_3 & 0 \end{bmatrix}$，且 $A = I + sin\theta C + (1 - cos\theta)C^2$. 此时有 $C\beta = \alpha \times \beta$, $\forall \beta \in \mathbb{R}^3$，证明只需考察空间中任意向量在 α 的生成空间和它的正交补上的投影，不难发现此时有 $A - A^\top = 2\sin\theta C$. 利用 θ 的正负和矩阵幂级数可以算出 $e^{\theta C} = I - sin\theta C + (1 - cos\theta)C^2$

5. 实方阵 A 是第一类正交矩阵当且仅当 A 可以写作 $e^C = I + C + \frac{1}{2!}C^2 + \cdots$，其中 C 是实反对称矩阵.(利用转置和原始矩阵的乘积为单位矩阵可知它是正交矩阵，另一方面，考察相似矩阵的行列式不变，将反对称矩阵 C 变成准对角矩阵，这里可以直接考察子块 $K = \begin{bmatrix} 0 & -\theta \\ \theta & 0 \end{bmatrix}$，不难得到 $e^K = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ 这里 θ 确实直接表示角度，且行列式为 1，即证.)

5.**Cartan-Dieudonne 定理** n 维正交变换的乘积是不超过 n 个反射变换的乘积.

定义了复内积的空间被称为**酉空间**，称一个变换 A 是**酉变换**，当 $\forall \alpha, \beta \in U$, $(A\alpha, A\beta) = (\alpha, \beta)$. 酉变换的特征值模长为 1（选择酉变换的特征向量 α 考内积 (α, α)）. 酉变换的充要条件是，存在一组标准正交基，使得它在这组基下是一个对角元模长均为 1 的对角矩阵.

与自身共轭转置乘积为单位阵的复方阵称为酉矩阵，酉变换的充要条件是它在一组标准正交基下的矩阵为酉矩阵.

正规变换 若 A 是正规变换，当 $AA^H = A^H A$；称一个矩阵 A 为正规矩阵，当 $AA^H = A^H A$. 不难发现正规变换的充要条件是它在某组标准正交基下的矩阵为正规矩阵.

谱定理 A 是正规变换，当且仅当存在一组标准正交基, 使得 A 的矩阵是复对角矩阵.（等价命题：A 是正规矩阵，当且仅当存在酉矩阵 U 和复对角矩阵 D，使得 $A = UDU^H$）（归纳证明. 现在找到一个复特征值 k 和对应的一个复特征向量 a，把 a 扩成全空间的一组基后，使用类似 Schmidt 正交化的方法. 此时在这组基下，A 变成 $\begin{bmatrix} k & C \\ 0 & B \end{bmatrix}$，这里 B 是 $(n-1) \times (n-1)$ 阶方阵. 利用正规变换的定义可以证明 C=0）

矩阵 A 是正规矩阵等价于 A 酉相似于一个对角阵(后推前容易，前推后：由 Schur 引理知存在酉矩阵 U 使得 $U^H AU = B$ 是上三角矩阵. 注意到 $B^H B = BB^H$，故 B 是正规矩阵，设 $B = \begin{bmatrix} b_{11} & \alpha \\ 0 & C \end{bmatrix}$，这里 C 也是上三角正规阵，代入 $B^H B = BB^H$ 不难看出 $\alpha = 0$，依此类推，B 是对角阵，即 A 可以被酉对角化.)

实正规矩阵的对称部分和反对称部分可交换.

Hermite 变换/自伴随变换 酉空间上满足下述的变换：$(A\alpha, \beta) = (\alpha, A\beta)$, $\forall \alpha, \beta \in V$. 当 A 是 Hermite 变换，当且仅当它在一组标准正交基下的矩阵是 Hermite 矩阵（自伴随矩阵，即等于自身的共轭

转置）. 有性质：

1.A 是 Hermite 矩阵，等价于 A 复正定（$\forall X \neq 0, X^\top A\overline{X} > 0$）；等价于 A 的特征值都大于 0；等价于 A 可以分解成 $P^\top P$（**Cholesky 分解**），P 是上三角可逆复矩阵；等价于 A 的顺序主子式都大于 0.($1 \to 2$，设出 Hermite 矩阵的特征值，代入定义式易见特征值一定是实数；另一方面，假设特征向量是 X，又有 $X^\top AX = \lambda||X||^2 > 0$，从而 $\lambda > 0$；$Hermite \to 3$;Cholesky 分解：先利用 LU 分解得到 $A = LU$，这里 L 是单位下三角阵，进一步，可以把 U 分解为 $U = DU_0$，这里 U_0 为单位上三角阵，只需 $U_0 = L^*$. 这一点利用 LU 分解的唯一性和 A 等于自己的共轭转置可得;$4 \to 1$，利用正定矩阵的充要条件是各阶顺序主子式的行列式都大于 0 可证等价性. 也可以对 A 作共轭成对的行列变换，从而相邻顺序主子式行列式的比值 (大比小) 为化出的对角元，一定为正，利用对角化的第一个元 $a_{11} > 0$ 即可，利用这种对角化的方法也可得出变形（过渡）矩阵是单位上三角矩阵，又有 Cholesky 分解)

2.Hermite 矩阵是正规矩阵，因此可以酉对角化. 假设 A 的对角化结果为 $U^H\Sigma U$，由于 Σ 的对角线元素全为正，开方得到 $\Sigma = D^2$，于是 $A = (DU)^H DU$，这里 U 是一个酉矩阵，可见 A 可以分解为一个矩阵和其共轭转置的乘积.

共轭变换 设 A 是酉空间 V 上的变换,若存在 V 上的线性变换 A^* 满足 $(A\alpha, \beta) = (\alpha, A^*\beta)$, $\forall\alpha, \beta \in V$，则称 A^* 是 A 的共轭变换. 共轭变换一定存在，它的充要条件是在同一组基下，A 的共轭变换的矩阵是 A 的矩阵的共轭转置.

正交空间与辛空间 设 K 是特征不为 2 的域, V 是 K-线性空间，在 V 上指定一个 (非退化) 双线性函数 f . 若 f 是对称双线性函数, 则称 (V, f) 是 **正交空间**; 若 f 是反对称双线性函数, 则称 (V, f) 是**辛空间**.

Witt 消去定理 设 V 是有限维正交 (辛) 空间, V , V” 是两个同构的子空间, 则此同构可扩充为 V 上的正交变换.

反射变换和镜面反射

半单变换

对称变换 若线性变换 A 是对称变换，则 $(A\alpha, \beta) = (\alpha, A\beta)$, $\forall\alpha, \beta$. 对称变换等价于在标准正交基下的矩阵是实对称矩阵. 设 f 是特征不为 2 的域 F 上的 n 维线性空间上的对称双线性函数，则 V 中存在一组基使得 f 在这组基下的度量矩阵为对角矩阵（合同标准型）.（归纳法 + 施密特正交化构造）

一个重要的性质是对称变换的特征值都是实数（设出特征值和特征向量，利用反对称变换中的取共轭方法）.

Cauchy Interlacing 原理 设 n 级实对称矩阵 A 的特征值为 $\lambda_1 \geq, \cdots \geq \lambda_n$，设 B 是 A 的一个 n-1 级主子阵，特征值为 $\mu_1 \geq \cdots \geq \mu_{n-1}$，则有：$\lambda_k \geq \mu_k \geq \lambda_{k+1}$, $\forall 1 \leq k \leq n$.

Weyl 定理 设 n 级实对称矩阵 A , B 的特征值分别为 $\lambda_1 \geq \cdots \geq \lambda_n$, $\mu_1 \geq \cdots \geq \mu_n$，设 A+B 的特征值为 $\tau_1 \geq \cdots \geq \tau_n$，则有：$\lambda_k + \mu_1 \geq \tau_k \geq \lambda_k + \mu_n$, $\forall 1 \leq k \leq n$.

更多实对称的特殊性质看“正定矩阵”.

反对称变换 设 f 是特征不为 2 的域 F 上的 n 维线性空间中的反对称双线性函数，则存在一组基使得

f 在这组基下的度量矩阵为（也叫反对称矩阵的合同标准型）

$$diag\left\{\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},0,\cdots,0\right\}$$

(利用数学归纳法，找内积非零的元素分组并递归构造)

实系数反对称矩阵的特征值只能是纯虚数和 0.(假设反对称矩阵 A 的特征值是 a，a 的一个特征向量是 b，则 Ab=ab，两边分别左乘 b 的共轭转置，再取结果的共轭转置和原式比较，利用 $A^*=A^\top$，可得特征值只能为纯虚数.)

反对称矩阵的相似准对角化:反对称矩阵在某组标准正交基下可以准对角化为 $diag\{R_1,R_2,\cdots,R_s,0,\cdots\}$，其中 R_i 是形如 $\begin{bmatrix}0 & q_i\\ -q_i & 0\end{bmatrix}$ 的分块矩阵.(考察归纳法：取出反对称矩阵 A 的一个特征向量 λ，考察 λ 的正交补，易见正交补空间中也是 A 的不变子空间，归纳即证. 反之若只有虚特征值，由于**实反对称矩阵的复特征值只能是纯虚数**，因此利用共轭特征值的特征向量即证 A 在这组向量张成的空间中可以化成 R_i 的形式).

实反对称矩阵一定是正规矩阵，但是复反对称矩阵不一定. 故只有实反对称矩阵才能保证酉对角化.

35. **同时对角化问题**

1. 若矩阵 A,B 都是 n 阶正定矩阵，则 A,B 可以同时合同对角化.

36. **几个反例**

1.$KerT$ 和 ImT 的维数和一定是全空间的维数, 但是 $KerT$ 和 ImT 不一定是直和. 如: 考察投影映射 $T:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ 定义为 $(x,y)\mapsto(x,0)$. 可以分别求得 $ImT=\{(x,0)|x\in\mathbb{R}\}$，$KerT=\{(0,y)|y\in\mathbb{R}\}$ 显然二者不是直和.（ImT 和 KerT 的交集不为空）

2. 任何一个线性空间都能写成它的一个子空间 V 和 V 的正交补的直和.(正交补在内积下定义，如果 V 是内积空间则命题才是真命题)

3. 设 A,B 是两个同阶矩阵，AB 可交换时不一定存在矩阵 C 使得 A 和 B 都能被 C 表征. 如 $A=\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix}$，$B=\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}$，则 AB 不能被某个 C 表征. 这是因为 C 的生成空间一定可以和 A,B 可交换，但是和 A,B 可交换的矩阵应该形如 $V=span<I,A,B>$，验证方程的有解性即可.

4.3 的特殊化：不存在一个矩阵 A，使得存在 $n*n$ 个实系数多项式 $p_{ij}(x)$ 使得 $p_{ij}(A)=E_{ij}$，这里 E_{ij} 是满足仅有 $a_{ij}=1$，其他元素均为 0 的 n 阶方阵的标准基.(利用 $p_{12}(A)p_{11}(A)=p_{11}(A)p_{12}(A)=E_{11}E_{12}=E_{12}E_{11}$，这显然矛盾.)

5. 若 V 是无限维线性空间，则 V 和 V 的对偶空间不一定同构. 如设 $V=\mathbb{Q}[x]$，则:

$$V\cong\{(a_0,a_1,a_2,\dots)\mid a_i\in\mathbb{Q},\text{仅有限个}a_i\neq0\}.$$

定义 V^* 为所有 $\mathbb{Q}$-线性泛函的集合, 即 $V^*=\{b^*=(b_0,b_1,b_2,\dots)\mid b_i\in\mathbb{Q}\}$, 其中 $b^*:(a_0,a_1,a_2,\dots)\mapsto$

$\sum_{i=0}^{\infty} a_i b_i$，并且这是一个有限和. 而 $V^* \cong \mathbb{Q}[[x]]$，它的基底是不可数的，但是 $V = \mathbb{Q}[x]$ 的基是可数的，即证.

6. 满秩矩阵不一定可对角化.(如特征值为 1 的约当矩阵)

37. **矩阵的分解**

1.

38. 一个有趣的问题： 设 $\mathcal{A}$ 是线性空间 V 上的线性变换，$k \geq 1$，则：

$$\dim Ker\mathcal{A}^k - \dim Ker\mathcal{A}^{k-1} \geq \dim Ker\mathcal{A}^{k+1} - \dim Ker\mathcal{A}^k$$

（利用映射 $S : \alpha \mapsto \mathcal{A}\alpha + Ker\,\mathcal{A}^{k-1}$ 是 $Ker\mathcal{A}^{k+1}$ 到 $Ker\mathcal{A}^k / Ker\mathcal{A}^{k-1}$ 的线性映射,且 $KerS = Ker\,\mathcal{A}^k$）

39. **特殊的矩阵**

1.Fourier 矩阵和 Fourier 变换

2.Hilbert 矩阵

40.**Schur 定理** 设 $A_1, A_2, \cdots, A_S \in Hom(V)$ 两两可交换，则：

- $A_1, \cdots, A_S$ 存在公共的特征向量.（先对空间的维数归纳. 显然空间维数均为 1 时可以，现考察 A_1 的某个特征值 λ 和相应的特征空间 $V_\lambda = Ker\,(A_1 - \lambda I)$. 如 V_λ 是小于 n 维的，我们可以把 A_j $(j \neq 1)$ 限制在 V_λ 上，现在它们的维数都小于 n，因此可以在不变子空间 V_λ 中找到一个公共的特征向量. 如 $\dim V_\lambda = n$，即 $A_1 = P^{-1} diag\{\lambda, \cdots, \lambda\} P = diag\{\lambda, \cdots, \lambda\}$，这里 A_1 经过相似对角化变成 λ 倍的单位向量，它的特征向量是平凡的，于是可以不考虑 A_1，次时对矩阵的数量作归纳即可.）
- 存在 V 的一组基，使得 $A_1, \cdots, A_S$ 在这组基下的矩阵是上三角矩阵.（对空间的维数归纳：注意到 A_i 有相同的特征向量，考察它们在该生成空间的商空间中的可交换性，由于可以同时上三角化，且扩展回原来的维度后仍然线性无关，于是在这组基下可以同时上三角化.）
- 若矩阵 A,B 可交换且 B 幂零，则有 $|A + B| = |A|$.（利用 A,B 可以同时上三角化，且 B 的上三角化的结果是对角元素全为 0 的矩阵）

41. 小技巧：

1. 如有矩阵 A,B 满足：$AB = 0$，则，$Im(A) \subseteq Ker(A)$，且 $Row(B^\top) = Im(B)$，$Row(B^\top) = Col(B)$，对称的，$Ker(A) = Col(A)^\perp$，于是 $Im(B) \subseteq Ker(A) \Longleftrightarrow Col(B) \subseteq Col(A)^\perp$. 直接讨论 $Im(A)$ 中的元素可知 $Im(A) \subseteq Ker\,(B)$.

2. 矩阵解空间是列空间的正交补.

42.**x-矩阵** 设 K 是一个数域，x 是一个字符，若一个矩阵的元素都是 K[x] 里的多项式，则称该矩阵为 x-矩阵. 和符号数矩阵同理可以定义 x-矩阵的逆矩阵. 有性质：n 级 x-矩阵 A(x) 可逆当且仅当 $|A(x)|$ 是 K 中的非零数.

Smith 矩阵 若 $A \in M_n(K)$，则存在可逆的 x-矩阵 $P(x), Q(x)$ 使得 Smith 标准型：

$$P(x)(xI-A)Q(x) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & d_1(x) & & \\ & & & \ddots & \\ & & & & d_s(x) \end{bmatrix}$$

这里 $d_i(x)$ 是次数 ≥ 1 的首一多项式，满足：$d_1(x)|d_2(x)|\cdots|d_s(x)$，我们称 $d_1(x), \cdots, d_s(x)$ 为 A 的**不变因子（组）**. 在 $d_1(x), \cdots, d_n(x)$ 中出现的 x 的一次因式方幂 (相同的需要重复记录) 称为 A 的**初等因子**. 上述 | 表示整除关系，更直白的说，Smith 标准型就是主理想整环上的矩阵的一种标准形式，通过初等变换可以把矩阵转化为对角形. 计算方法：初等变换 + 辗转相除法，初始矩阵写作：$\begin{bmatrix} xI-A \\ I \end{bmatrix}$，$xI-A$ 对角化后，过度矩阵 Q(x) 就会被记录在 I 的位置. 利用 Q(x) 的列向量，可以求出 $A-\lambda I$ 的强循环子空间. 这里 λ 为初等因子的根：(在 Q(x) 的第 i 个列向量中分出该列对应的初等因子的因式 $1, \cdots, (x-\lambda)^r, r$是该因式在该初等因子中的次数-1).

行列式因子 设 x-矩阵 $A(x)$ 的秩为 r，对于正整数 k,$1 \leq k \leq r$，若 $A(x)$ 必有非零 k 阶子式，$A(x)$ 的全部 k 阶子式的首项系数为 1 的最大公因式 $D_k(x)$ 称为 $A(x)$ 的 k 阶行列式因子.

推论：

- 方阵的初等因子组在相差一个初等因子排列的意义下唯一.(初等因子组和若当标准型一一对应)
- 两个矩阵相似当且仅当他们的初等因子组相同，或所有 k 阶行列式因子相同，或不变因子相同.
- 次数最高的不变因子是矩阵的最小多项式.
- x-矩阵 A 的不变因子依次记作 $d_1(x), \cdots$，其行列式因子依次记作 $D_1(x), \cdots$，则有关系：

$$d_1(x) = D_1(x),\ d_i(x) = \frac{D_i(x)}{D_{i-1}(x)},\ \cdots$$

43. **范数**

向量范数满足正定性（除了零元一定大于 0）、齐次性（数域上的常量的模可以提出去）、三角不等式（$||\alpha+\beta|| \leq ||\alpha|| + ||\beta||$）；带范数的线性空间被称为赋范空间.

范数不等式/Minkovski 不等式 对任意 $\alpha = [a_1 \cdots a_n]^\top$，$\beta = [b_1 \cdots b_n]^\top \in \mathbb{C}^n$ 有：

$$\left(\sum_{i=1}^n |a_i+b_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p\right)^{\frac{1}{p}}$$

即，$||\alpha+\beta||_p \leq ||\alpha||_p + ||\beta||_p$.

定理 设 V 是数域 K 上的 n 维向量空间，$||*||$与$||*||'$ 是 V 上的两个范数，则存在常数 $C_1, C_2 > 0$ 使得 $C_1||\alpha|| \leq ||\alpha||' \leq C_2||\alpha|| \qquad \forall \alpha \in V$. 这是因为单位化 α 后，可以看出 $f(\alpha) = ||\alpha||'$ 是一个连续函数，限制 $||\alpha||$ 是一个有界闭集，因此可以找到前面连续函数的最大值和最小值 $C_2 \geq C_1 \geq 0$.

对于矩阵的范数还需要额外满足相容性：$||AB|| \leq ||A|| + ||B||$.

矩阵 $A \in M_n(K)$ 的复特征值 $\lambda_1, \cdots \lambda_n$ 称为 A 的**谱**，$\rho(A) := \underset{1\leq i\leq n}{max}|\lambda_i|$ 称为 A 的**谱半径**. 有性质，矩阵的谱半径一定不大于 A 的任何范数.

定义 给定向量空间 K^n 上的一个向量范数 $||*||_v$，则 $||A|| := \underset{||\alpha||_v=1}{max}||A\alpha||_v = \underset{0\neq\alpha\in K^n}{max}\frac{||A\alpha||}{||\alpha_v||}, \qquad \forall A \in M_n(K)$ 是一个矩阵范数，被称为从属于向量范数 $||*||_v$ 的从属范数. 从属于 l_2 范数的矩阵范数被称为**谱范数**，记作 $||*||_2$.v=1 的从属范数也叫 (欧氏) 诱导范数，它的值是 $\sqrt{\lambda_1}$，这里λ_1 是矩阵的最大特征值. 诱导范数有矩阵范数的一切性质，另外还满足转置相等性.

二次型的取值 对任意列向量 $X \in \mathbb{R}^n$，有 $\lambda_n||X||^2 \leq X^\top AX \leq \lambda_1||X||^2$，这里 λ_1, λ_n 分别是实对称矩阵 A 的最大、最小的特征值.（把 A 对角相似分解，利用均值不等式即可）

定理: 对任意 $A = [a_{ij}]_{1\leq i,j\leq n} \in M_n(K)$，有

$$\|A\|_2 := \max_{\|\alpha\|_2=1}\|A\alpha\|_2 = \max_{\|\alpha\|_2=1}\sqrt{\alpha^T\overline{A}^TA\alpha} = \sqrt{\rho(\overline{A}^TA)}$$

(首先 $\overline{A}^\top A$ 一定是非负定矩阵，相似对角化后得到 $P\Sigma P^\top$，取 P 的列向量表示 α，直接代入计算，利用加权平均不等式即可.)

设赋范空间 $(V, \|\cdot\|)$ 完备。若范数级数 $\sum_{n=1}^\infty \|\alpha_n\|$ 收敛，则 V 中的向量级数 $\sum_{n=1}^\infty \alpha_n$ 收敛。此时我们称向量级数绝对收敛。绝对收敛的向量级数可任意交换求和项而和不变，即对任意自然数集到自身的双射 σ，都有

$$\sum_{n=1}^\infty \alpha_{\sigma(n)} = \sum_{n=1}^\infty \alpha_n.$$

(对任意幂级数 $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$，存在唯一的 r（$0 \leq r \leq +\infty$），使得当 $|x| < r$ 时幂级数绝对收敛，当 $|x| > r$ 时幂级数一定发散。这样的 r 称为幂级数的收敛半径。)

Frobenius 范数 记作 $||A||_F = \sqrt{\sum_{i=1}^m\sum_{j=1}^n|a_{ij}|^2} = tr(A^HA)$，它不是诱导范数.

定理: 设矩阵 $A \in M_n(\mathbb{C})$ 的谱半径 $\rho(A)$ 小于幂级数 $f(x) = \sum_{k=0}^\infty a_kx^k$ 的收敛半径 R，则矩阵级数 $f(A) = \sum_{k=0}^\infty a_kA^k$ 绝对收敛，且存在次数小于 d 的多项式 $g(x)$ 使得：

$$f(A) = g(A) \in R_d[A],$$

其中 $R_d[A]$ 是 A 的最小多项式 $m(x)$ 生成的商空间.

44. **正方阵** 元素都是正数的矩阵称为正矩阵, 元素都非负的矩阵称为非负矩阵.

Perron-Frobenius 定理 对于正方阵 A, 以下命题成立:

- A 有正特征值. 设 λ 是 A 的最大正特征值 (**Frobenius 根**，谱半径), 则 λ 的代数重数为 1 , 且 λ 有正特征向量 α ;

- 除 λ 以外, A 其它复特征值的绝对值 $< \lambda$;
- 正方阵 $A^\top$ 与 A 有相同的最大正特征值.

45. **正定矩阵** (显然满足对称矩阵的所有性质)

1. 实对称矩阵 A 是正定矩阵等价于它的特征值全是正数.(直接设出特征值和特征向量用二次型.)(等价于 A 的所有主子式都是正的，也等价于 A 的各级顺序主子式都是正数.)

2.（Schur） 若 $A \in C^{n\times n}$，则存在酉矩阵 $U \in C^{n\times n}$，及上三角矩阵 $T \in C^{n\times n}$，使得 $U^H AU = T$. 这里 T 的主对角线上都是 A 的特征值. 将过度矩阵用对角矩阵作用，还能得到额外的限制.

3.(Schur 不等式) 若 n 阶复方阵 $A = (a_{ij}) \in C^{n\times n}$，A 的特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则 $\sum_{i=1}^n |\lambda_i|^2 \le \sum_{i,j=1}^n |a_{ij}|^2$. 等号成立当且仅当 A 是正规矩阵.(利用 Schur 引理，存在酉矩阵 U 和上三角矩阵 T 满足前述，使得 $U^H AU = T$，于是 $\sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n |t_{ii}|^2 \le \sum_{i=1}^n |t_{ii}|^2 + \sum_{i<j} |t_{ij}|^2 = tr(T^H T) = tr(A^H A) = \sum_{i,j=1}^n |a_{ij}|^2$，等号成立当且仅当 $\sum_{i<j} |t_{ij}|^2 = 0$，即 A 酉相似于对角阵，即 A 是正规矩阵.)

4. 存在唯一的正定矩阵 S 使得 $A = S^2$.

5. 若 A,B 都是正定矩阵，则 AB 的特征值都是正数.（利用 4，设正定矩阵 C 满足 $A = C^2$，注意到 $C^{-1}ABC = (C^{-1}ABC)^\top$，于是 $C^{-1}ABC$ 实对称；另一方面，分解 $B = D^2$，任取 $\alpha \in \mathbb{R}^n$，于是 $\alpha^\top C^{-1}ABC\alpha = \alpha^\top CD^2C\alpha = (DC\alpha, DC\alpha) \ge 0$，由于 DC 满秩，显然上述等号不取等，并且该矩阵和 AB 相似，从而 AB 的特征值也全是正数.）

6. 如果矩阵 A 是正定矩阵（n 阶），B 半正定实对称矩阵且 $|B| \ne 0$，则 $|A + B| > max\{|A|, |B|\}$.

证明. 由于 A 正定，于是存在可逆矩阵 P 使得 $P^T AP = I$. 注意到 $P^T BP$ 也是实对称矩阵，于是它可以对角化，即存在正交矩阵 Q 使得 $Q^T P^T BPQ = D = \mathrm{diag}\{\mu_1, \mu_2, \cdots, \mu_n\}$.

由于 B 是半正定矩阵，于是 $\mu_i \ge 0$ 且不全为零.

令 $C = PQ$ 得 $C^T AC = I$ 且 $C^T BC = D$. 于是

$$|A + B| = |C^{-1}I(C^T)^{-1} + C^{-1}D(C^T)^{-1}| = |C^{-1}|^2|I + D| = |C^{-1}|^2 \prod_{i=1}^n (1 + \mu_i)$$

另一方面，

$$|A| = |C^{-1}I(C^T)^{-1}| = |C^{-1}|^2, \quad |B| = |C^{-1}D(C^T)^{-1}|^2 = |C^{-1}|^2 \mu_1\mu_2 \cdots \mu_n,$$

进而有

$$|A + B| > \max\{|A|, |B|\}$$

□

7. 如果 $A = (a_{ij})$ 是一个 n 级正定矩阵，则 $|A| \le a_{11}a_{22} \cdots a_{nn}$.

Step 1: 先证明：若

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$$

是 n 级正定矩阵，其中 A 是 r 级矩阵，则 $A, D, D-B^TA^{-1}B$ 都是正定矩阵。

显然 $\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$ 的所有主子式全是正数，因此 A 与 D 的所有顺序主子式是正数，进而是正定矩阵。由于

$$\begin{pmatrix} I_r & 0 \\ -B^TA^{-1} & I_{n-r} \end{pmatrix}\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}\begin{pmatrix} I_r & -A^{-1}B \\ 0 & I_{n-r} \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D-B^TA^{-1}B \end{pmatrix}$$

于是 $\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$ 合同于矩阵 $\begin{pmatrix} A & 0 \\ 0 & D-B^TA^{-1}B \end{pmatrix}$，于是后者也是正定矩阵，进而 $D-B^TA^{-1}B$ 的各级主子式也都是正数，所以 $D-B^TA^{-1}B$ 也是正定矩阵。

Step 2: 证明在 Step 1 的条件下有 $|M| \le |A||D|$ 且等号成立当且仅当 $B=0$。

由前面的步骤知 $|M| = |A||D-B^TA^{-1}B|$，记 $H = D-B^TA^{-1}B$，由于 $B^TA^{-1}B$ 半正定，因此由性质 6 知 $|D| = |H+B^TA^{-1}B| \ge |H|$，等号成立当且仅当 $B=0$. 于是

$$|M| = |A||H| \le |A||D|$$

Step 3: 使用数学归纳法，假设性质 3.3 对于 $n-1$ 的情形是成立的，于是将 A 写成

$$A = \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{pmatrix}$$

由于 A 是正定矩阵，于是 A_{n-1} 也是正定矩阵，由归纳假设知

$$|A_{n-1}| \le a_{11}a_{22}\cdots a_{n-1,n-1}$$

结合 Step 2 知

$$|A| \le |A_{n-1}|a_{nn} \le a_{11}a_{22}\cdots a_{nn}$$

【例】 设 $A=[a_{ij}]$, $B=[b_{ij}]$ 是 n 阶半正定矩阵, 则 $C=[a_{ij}b_{ij}]$ 也是半正定矩阵.(设 $A=PDP^\top$, $P=[\alpha_1\alpha_2,\cdots\alpha_n]=[p_{ij}]$ 是正交矩阵, 且 $D=diag\{\lambda_1,\cdots,\lambda_n\}, \lambda_i \ge 0$, 则 $A=\lambda_1\alpha_1\alpha_1^\top+\cdots+\lambda_n\alpha_n\alpha_n^\top$, 于是有 $a_{ij}=\lambda_1p_{i1}p_{j1}+\cdots+\lambda_np_{in}p_{jn}$，代入 $X^\top CX=\sum_{i,j=1}^n a_{ij}b_{ij}x_ix_j=\sum_k\lambda_k\sum_{i,j}b_{ij}(p_{ik}x_i)(p_{jk}x_j)$, 对固定的 k 令 $y_i=p_{ik}x_i$ 有：$\sum_{i,j}b_{ij}(p_{ik}x_i)(p_{jk}x_j)=\sum_{i,j}b_{ij}y_iy_j \ge 0$，于是 $X^\top CX \ge 0, \ \forall X \in \mathbb{R}^n$)

46. 若 A 是复矩阵，则存在唯一复矩阵 X，满足

$$AXA = A, \quad XAX = X,$$

$$(AX)^H = AX, \quad (XA)^H = XA.$$

X 称为 A 的 **Penrose 广义逆**，记为 A^+，这里 $X^H = \overline{X}^T$。

若实矩阵 A 列满秩，则 A 的 Penrose 广义逆

$$A^+ = (A^TA)^{-1}A^T$$

复矩阵的广义逆：若 $A = PSQ^H$，则

$$A^+ = Q \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix}_{n\times m} P^H$$

47. **张量积**