



高等代数笔记

作者: Reichtum

组织: ZheJiang University

时间: September 20, 2022

版本: 0.1



There was no royal road to geometry. —Euclid

目录

第 1 章 多项式	1
1.1 一元多项式	1
1.1.1 整除理论	1
1.1.2 最大公因式	2
1.1.3 互素	3
1.1.4 不可约多项式与因式分解	4
1.1.5 余数定理及其应用	5
1.1.6 重因式与重根	6
1.1.7 整数与多项式	7
1.2 特殊数域上的一元多项式理论	8
1.2.1 复数域上的一元多项式	8
1.2.2 实数域上的一元多项式	8
1.2.3 有理数域上一元多项式因式分解	8
1.2.4 一元有理多项式的根	10
1.3 多元多项式	10
1.3.1 基本概念	10
1.3.2 对称多项式	11
1.3.3 判别式	12
1.3.4 结式	12
第 2 章 矩阵	14
2.1 矩阵基础	14
2.1.1 矩阵概念与运算	14
2.1.2 矩阵的逆	14
2.1.3 伴随矩阵的性质	15
2.1.4 广义逆矩阵	16
2.1.5 Moore-Penrose 广义逆	17
2.2 分块矩阵	18
2.2.1 常见矩阵拆分技巧	18
2.2.2 分块初等变换	19
2.2.3 分块打洞技巧总结	19
2.2.4 分块矩阵行列式	20
2.2.5 分块矩阵求逆	20
2.3 矩阵的秩	21
2.3.1 秩的概念	21
2.3.2 初等变换标准型与矩阵分解	21
2.3.3 利用初等变换标准型进行矩阵分解	22
2.3.4 秩不等式	22
2.3.5 空间分解法研究矩阵秩	24
2.4 矩阵的迹	25
2.5 几种技巧矩阵	25

2.5.1	初等矩阵	25
2.5.2	基本矩阵: 仅一个元素为 1 的矩阵	25
2.5.3	置换矩阵	26
2.6	特殊矩阵性质总结	27
2.6.1	对角矩阵	27
2.6.2	上、下三角矩阵	27
2.6.3	实对称矩阵	28
2.6.4	反对称矩阵	28
2.6.5	Hermite 矩阵	30
2.6.6	正交矩阵	30
2.6.7	Unitary 矩阵	31
2.6.8	幂零矩阵	31
2.6.9	幂等矩阵	31
2.6.10	对角占优矩阵	33
第 3 章	行列式	34
3.1	行列式基本概念与性质	34
3.2	行列式展开	34
3.3	特殊的行列式	34
3.3.1	爪型行列式	34
3.3.2	大对角型 (三斜线)	35
3.3.3	行和相同型	35
3.3.4	循环型	35
3.3.5	行列式加边法	36
3.4	Cramer 法则	36
第 4 章	线性方程组	37
4.1	线性方程组解的判定定理	37
4.1.1	解存在判定定理	37
4.1.2	线性方程组同解	37
4.1.3	解个数判定定理	37
4.2	线性方程组的求解	37
4.2.1	$Ax = 0$ 求解	37
4.2.2	$Ax = b$ 求解	38
4.2.3	线性方程组公共解的计算	39
4.3	最小二乘解	39
4.3.1	$A^T A$ 的研究	39
4.3.2	正规方程组	39
第 5 章	线性空间与线性变换	41
5.1	线性空间	41
5.1.1	线性空间的基本概念	41
5.1.2	线性相关、线性无关与向量组的秩	41
5.1.3	线性空间同构	42
5.1.4	基、坐标	43
5.1.5	坐标变换	44

5.1.6	和与交、维数定理	45
5.1.7	直和理论	47
5.1.8	商空间	48
5.2	线性映射	48
5.2.1	线性映射的概念	48
5.2.2	核与像	48
5.2.3	幂等变换的核与像	50
5.2.4	一般线性变换的核像分解	50
5.3	线性变换	51
5.3.1	线性变换基本概念	51
5.3.2	线性变换在基下的矩阵、变换后的坐标	51
5.3.3	线性变换在不同基下的矩阵：相似关系	52
5.3.4	相似矩阵的性质	53
5.4	特征值、相似对角化	53
5.4.1	特征值与特征向量	53
5.4.2	特征值与特征向量进一步结论	54
5.4.3	相似的进一步研究	55
5.4.4	相似对角化	56
5.4.5	相似上三角型	57
5.5	最小多项式	58
5.5.1	最小多项式与基本性质	58
5.5.2	Hamilton-Cayley 定理	58
5.5.3	最小多项式与相似对角化	59
第 6 章	Jordan 标准型	60
6.1	Jordan 标准形：空间分解法	60
6.1.1	空间第二分解定理：循环不变子空间分解	60
6.1.2	核像升链与根子空间	60
6.1.3	空间第一分解定理：根子空间分解	61
6.1.4	Jordan 标准形：空间分解法	62
6.2	空间分解法：零化多项式角度	62
6.2.1	根子空间分解：矩阵角度	62
6.2.2	根子空间分解：线性变换角度	63
6.2.3	Jordan 阵与循环不变子空间	64
6.3	Jordan 标准形： λ 矩阵法	65
6.3.1	λ 矩阵与相抵	65
6.3.2	行列式因子、不变因子	66
6.3.3	初等因子与 Jordan 标准型	67
6.4	λ 矩阵的进一步研究	68
6.4.1	Jordan 标准型的直接计算	68
6.4.2	k 阶 Jordan 块的个数	70
6.4.3	分块矩阵的 Jordan 阵	71
6.4.4	过渡矩阵的计算	71
第 7 章	双线性函数与二次型	73

7.1	双线性函数	73
7.1.1	双线性函数概念	73
7.1.2	双线性函数度量矩阵	73
7.1.3	双线性函数不同基下的矩阵: 合同	73
7.2	二次型	74
7.2.1	二次型概念及其矩阵	74
7.2.2	二次型的标准型存在与配方法	75
7.2.3	二次型的标准型: 初等变换法	77
7.2.4	二次型的标准型: 正交变换法	77
7.2.5	$\mathbb{R}, \mathbb{C}$ 上的规范型	79
7.3	(半) 正定二次型的研究	79
7.3.1	正定二次型的概念与判别	79
7.3.2	半正定二次型的概念与判别	81
7.3.3	$C^T C$ 的 (半) 正定性	82
7.3.4	正定矩阵同时对角化	83
7.3.5	正定矩阵的四则运算性质	84
7.3.6	正定矩阵分解进阶	84
7.3.7	负定二次型	85
第 8 章	内积空间	86
8.1	Euclid 空间	86
8.1.1	Euclid 空间与正交	86
8.1.2	标准正交基	87
8.1.3	正交和与正交补	88
8.1.4	正交变换	89
8.1.5	镜面反射	90
8.2	Unitary 空间	90
8.2.1	Unitary 空间概念	90
第 9 章	对偶空间	91
9.1	对偶空间与对偶基	91

第1章 多项式

1.1 一元多项式

1.1.1 整除理论

定义 1.1 (整除)

$g(x), f(x)$ 是 $\mathbb{P}$ 上的多项式, 若 $\exists h(x)$ 使得 $f(x) = g(x)h(x)$, 则称 $g(x)$ 整除 $f(x)$, 记为 $g(x)|f(x)$ 。称 $g(x)$ 是 $f(x)$ 的因式, $f(x)$ 是 $g(x)$ 的倍式。



定理 1.1 (整除的数域不变性)

$\mathbb{P}, \mathbb{F}$ 为两个数域, $\mathbb{P} \subseteq \mathbb{F}$, $f(x), g(x) \in \mathbb{P}[x]$ (在 $\mathbb{P}$ 中), 则 $\mathbb{P}$ 中 $g(x)|f(x)$ 当且仅当 $\mathbb{F}$ 中 $g(x)|f(x)$



笔记 该定理的意义在于可以在最大的数域 (方便因式分解) 中证明整除, 即可证明小数域中整除。

定理 1.2 (带余除法)

对于 $\forall f(x), g(x) \in \mathbb{P}[x]$, 存在唯一 $q(x), r(x) \in \mathbb{P}[x]$ 使得 $f(x) = q(x)g(x) + r(x)$ 。这里 $r(x) = 0$ 或者 $\partial(r(x)) < \partial(g(x))$



练习 1.1 带余除法的计算 (1) $f(x) = x^4 + 2x^3 - 5x + 7, g(x) = x^2 - 3x + 1$, 求 $g(x)$ 除 $f(x)$ 的商式与余式
解 (1) 商为 $x^2 + 5x + 14$, 余式为 $32x - 7$

练习 1.2 带余除法的应用 (1) 设 $g(x) = ax + b, f(x) \in \mathbb{P}[x]$, 证明: $g(x)|f^2(x)$ 当且仅当 $g(x)|f(x)$
(2) $f(x), h(x), g(x)$ 为三个非零多项式, 证明: $h(x)|(f(x) - g(x))$ 当且仅当 $f(x), g(x)$ 除以 $h(x)$ 的余式相同
(3) $f(x), g(x)$ 为多项式, 证明: $g^m(x)|f^m(x)$ 当且仅当 $g(x)|f(x)$

证明 (1) 右推左显然。左推右: $f^2(x) = g(x)h(x)$, 若 $f(x) = q(x)g(x) + r(x)$, 则 $f^2(x) = q^2(x)g^2(x) + 2q(x)g(x)r(x) + r^2(x)$, 根据 $g(x)|f(x)$ 得到 $g(x)|r^2(x)$, 显然 $r(x) = 0$

(2) 设 $f(x) = q_1(x)h(x) + r_1(x), g(x) = q_2(x)h(x) + r_2(x)$, 相减得到

$$f(x) - g(x) = h(x)[q_1(x) - q_2(x)] + r_1(x) - r_2(x)$$

显然 $\partial(r_1(x) - r_2(x)) < \partial(h(x))$, 而 $h(x)|(f(x) - g(x))$, 因此当且仅当 $r_1(x) = r_2(x)$

(3) 右推左: 用 $f(x) = h(x)g(x)$, 两侧取 m 次方即可。

左推右: 设 $f(x) = c_1 p_1^{r_1}(x) \cdots p_s^{r_s}(x), g(x) = c_2 p_1^{t_1}(x) \cdots p_s^{t_s}(x)$, 其中 $p_1(x) \sim p_s(x)$ 两两不可约且互素, r_i 可以为 0 或非负整数。根据条件可知 $mt_i \leq mr_i$, 因此 $t_i \leq r_i$, 因此 $g(x)|f(x)$

注意本题用带余除法可能不好说, 因此稳妥起见还是用因式分解更好。

1.1.2 最大公因式

定义 1.2 (最大公因式)

$f(x), g(x), d(x) \in \mathbb{P}[x]$, 若 $d(x)$ 同时整除 $f(x), g(x)$, 则称其为公因式。若没有可以整除 $d(x)$ 的公因式存在, 则称其为最大公因式, 记为

$$d(x) = (f(x), g(x))$$



定理 1.3 (最大公因式表示定理)

$\forall f(x), g(x) \in \mathbb{P}[x]$, 均存在 $d(x) = (f(x), g(x)) \in \mathbb{P}[x]$, 且 $\exists u(x), v(x) \in \mathbb{P}[x]$ 满足 $d(x) = u(x)f(x) + v(x)g(x)$ (反之不可, 互素时才可)



证明 假设通过辗转相除法得到:

$$\begin{cases} f(x) = q_1(x)g(x) + r_1(x) \\ g(x) = q_2(x)r_1(x) + r_2(x) \\ r_1(x) = q_3(x)r_2(x) + r_3(x) \\ \vdots \\ r_{s-3}(x) = q_{s-1}(x)r_{s-2}(x) + r_{s-1}(x) \\ r_{s-2}(x) = q_s(x)r_{s-1}(x) + r_s(x) \\ r_{s-1}(x) = q_{s+1}(x)r_s(x) + 0 \end{cases}$$

则此时可以得到 $r_s(x)$ 是 $f(x), g(x)$ 的最大公因式, 并且从倒数第二个等式开始往回推, 得到

$$\begin{aligned} r_s(x) &= r_{s-2}(x) - q_s(x)r_{s-1}(x) \\ &= r_{s-2}(x) - q_s(x)(r_{s-3}(x) - q_{s-1}(x)r_{s-2}(x)) \\ &= \dots \\ &= u(x)f(x) + v(x)g(x) \end{aligned}$$

即逐个将前面的等式代入方程, 逐个消去 $r_{s-2}(x), \dots, r_2(x), r_1(x)$, 最终只剩下 $g(x), f(x)$

笔记 最大公因式 $d(x) = u(x)f(x) + v(x)g(x)$ 的 $u(x)$ 和 $v(x)$ 不是唯一的, 如果要唯一则需要加上额外的限制条件。

练习 1.3 辗转相除法 (1) $f(x) = x^4 + 3x^3 - x^2 - 4x - 3, g(x) = 3x^3 + 10x^2 + 2x - 3$, 求 $(f(x), g(x))$, 并求 $u(x), v(x)$ 使得 $u(x)f(x) + v(x)g(x) = (f(x), g(x))$

(2) 已知 $f(x) = x^4 + x^3 - 3x^2 - 4x - 1, g(x) = x^3 + x^2 - x - 1$, 求满足 $u(x)f(x) + v(x)g(x) = x + 1$ 的 $u(x), v(x)$

解 (1) 第一题过程如下图所示。(2) 可先求出最大公因式, 再提取

定理 1.4 (保最大公因式)

$f(x), g(x) \in P[x]$, 若 $a, b, c, d \in P$ 满足 $ad - bc \neq 0$, 则

$$(f(x), g(x)) = (af(x) + bg(x), cf(x) + dg(x))$$



证明 假设左侧为 $d(x)$, 右侧为 $d_1(x)$, 下面要证 $d(x) = d_1(x)$, 不妨证明 $d(x)|d_1(x), d_1(x)|d(x)$ 。

(1) $d(x)|d_1(x)$: $d(x)$ 是左侧最大公因式, 其是右侧每个部分的最大公因式, 因此 $d(x)|d_1(x)$

(2) $d_1(x)|d(x)$: 由于 $ad - bc \neq 0$, 从方程组角度 $f(x), g(x)$ 也可以用 $af(x) + bg(x), cf(x) + dg(x)$ 线性表出,

Handwritten work showing the Euclidean algorithm for finding the GCD of two polynomials $f(x)$ and $g(x)$.

① $f(x) = x^4 + 3x^3 - x^2 - 4x - 3$
 $g_1(x) = x - \frac{1}{3}$

② $\frac{1}{3}g(x) = x^3 + \frac{10}{3}x^2 + \frac{2}{3}x - 1$
 $g_2(x) = x - \frac{5}{3}$

Division steps:
 $f(x) \div \frac{1}{3}g(x) = 1 \text{ remainder } -\frac{1}{3}x^3 - \frac{5}{3}x^2 - 3x - 3$
 $-\frac{1}{3}x^3 - \frac{5}{3}x^2 - 3x - 3$
 $-\frac{1}{3}x^3 - \frac{10}{9}x^2 - \frac{2}{9}x + \frac{1}{3}$
 $r_1(x) = -\frac{5}{9}x^2 - \frac{25}{9}x - \frac{10}{3}$
 $-\frac{5}{9}r_1(x) = x^2 + 5x + 6$

③ $r_2(x) = 3x + 9$
 $\frac{1}{3}r_2(x) = x + 3$

因由: ① $f(x) = \frac{1}{3}g_1(x)g(x) + r_1(x)$ ② $g(x) = g_2(x)(-\frac{5}{3}r_1(x)) + r_2(x)$
 ③ $-\frac{9}{5}r_1(x) = (x+3)\frac{1}{3}r_2(x)$ $r_2(x)$ 与 $x+3$ 互素
 $\Rightarrow$ ①: $r_1(x) = f(x) - \frac{1}{3}g_1(x)g(x)$ ②: $r_2(x) = g(x) + \frac{5}{3}g_2(x)r_1(x)$
 将 $r_1(x)$ 代入 ② 中得

具体用 cramer 法则求出即可, 如 $f(x) = \frac{df_1(x) - bg_1(x)}{ad - bc}$, $g(x) = \frac{ag_1(x) - cf_1(x)}{ad - bc}$, 因此与 (1) 同理

推论 1.1 (加减保最大公因式)

$$(f(x), f(x) - g(x)) = (f(x), f(x) + g(x)) = (f(x), g(x))$$

定理 1.5 (最小公倍式的表示)

$$f(x), g(x) \text{ 的首 1 最小公倍式 } [f(x), g(x)] = \frac{f(x)g(x)}{(f(x), g(x))}$$

证明 设 $(f(x), g(x)) = d(x)$, $f(x) = d(x)f_1(x)$, $g(x) = d(x)g_1(x)$, 且 $(f_1(x), g_1(x)) = 1$, 记 $s(x) = \frac{f(x)g(x)}{d(x)} = d(x)f_1(x)g_1(x) = f(x)g_1(x) = f_1(x)g(x)$, 下面说明 $s(x)$ 是最小公倍式: (1) 是公倍式: 显然 (2) 最小: 设 $h(x)$ 也是公倍式, 设 $h(x) = f(x)m(x)$, 根据 $g(x)|h(x)$ 得到 $d(x)g_1(x)|d(x)f_1(x)m(x)$, 根据 $g_1(x)$ 与 $f_1(x)$ 互素, 得到 $g_1(x)|m(x)$, 因此 $g_1(x)f(x)|f(x)m(x)$, 得到 $s(x)|h(x)$, 因此 $s(x)$ 是最小的。

1.1.3 互素

定义 1.3 (互素)

若 $f(x), g(x) \in \mathbb{P}[x]$ 的最大公因式为 1, 则称它们互素。

定理 1.6 (互素充要条件)

$$f(x), g(x) \in \mathbb{P}[x] \text{ 互素当且仅当 } \exists u(x), v(x) \in \mathbb{P}[x] \text{ 使得 } u(x)f(x) + v(x)g(x) = 1$$

证明 左推右显然 (根据最大公因式定理)。右推左: 设 $(f(x), g(x)) = d(x)$, 则 $d(x)|f(x), d(x)|g(x)$, 设 $f(x) = d(x)f_1(x), g(x) = d(x)g_1(x)$, 从而

$$d(x)[u(x)f_1(x) + v(x)g_1(x)] = 1$$

因此 $d(x), u(x)f_1(x) + v(x)g_1(x)$ 均为常数, 故 $f(x), g(x)$ 互素

 **笔记** 上述 $u(x), v(x)$ 在无限制下也是不一定唯一的。

 **练习 1.4 互素相关应用** (1) 若 $f(x), g(x)$ 互素, 证明: $f(x^m), g(x^m)$ 互素

(2) 重点: 若 $(f(x), g(x)) = d(x)$, 证明: $(f(x^m), g(x^m)) = d(x^m)$

(3) 设 $(f(x), g(x)) = 1, (f(x), h(x)) = 1$, 证明: $(f(x), g(x)h(x)) = 1$

证明 (1) 由于 $f(x), g(x)$ 互素, 得到 $u(x)f(x) + v(x)g(x) = 1$, 从而 $u(x^m)f(x^m) + v(x^m)g(x^m) = 1$, 由于是互素充要条件, 因此 $f(x^m), g(x^m)$ 互素

(2) 不能直接用最大公因式表示, 因为不是充要条件, 要转到互素上。设 $f(x) = d(x)m(x), g(x) = d(x)n(x)$, 则 $(m(x), n(x)) = 1$, 因此

$$u(x^k)m(x^k) + v(x^k)n(x^k) = 1$$

得到 $m(x^k), n(x^k)$ 互素, 而 $f(x^k) = d(x^k)m(x^k), g(x^k) = d(x^k)n(x^k)$, 因此 $(f(x^k), g(x^k)) = d(x^k)$

(3) 根据互素, 有 $u_1(x)f(x) + v_1(x)g(x) = 1, u_2(x)f(x) + v_2(x)h(x) = 1$, 两式相乘得到

$$[u_1(x)u_2(x)f(x) + v_1(x)g(x) + v_2(x)h(x)]f(x) + v_1(x)v_2(x)g(x)h(x) = 1$$

因此根据充要条件得到 $(f(x), g(x)h(x)) = 1$

 **练习 1.5 互素充要条件推广** $f(x), g(x)$ 是次数大于 0 的互素多项式, 则存在唯一的 $u(x), v(x)$ 使得 $u(x)f(x) + v(x)g(x) = 1$, 这里 $\partial(u(x)) < \partial(g(x)), \partial(v(x)) < \partial(f(x))$

证明 (1) 存在性: 由于 $f(x), g(x)$ 互素, 故存在 $a(x), b(x)$ 使得 $a(x)f(x) + b(x)g(x) = 1$, 显然 $a(x), g(x)$ 也互素, 作带余除法 $a(x) = q(x)g(x) + u(x)$, 代入公式得到

$$(q(x)g(x) + u(x))f(x) + b(x)g(x) = 1 \Rightarrow u(x)f(x) + (b(x) + q(x)f(x))g(x) = 1$$

记 $v(x) = b(x) + q(x)f(x)$, 下面证明 $\partial(v(x)) < \partial(f(x))$ 。反设 $\partial(v(x)) \geq \partial(f(x))$, 根据 $\partial(u(x)) < \partial(g(x))$ 可知 $\partial(v(x)g(x)) > \partial(u(x)f(x))$, 显然它们的和不等于 1, 矛盾。

(2) 唯一性: 假设还有 $u_1(x)f(x) + v_1(x)g(x) = 1$, 那么有 $(u(x) - u_1(x))f(x) + (v(x) - v_1(x))g(x) = 0$, 即

$$(u(x) - u_1(x))f(x) = (v_1(x) - v(x))g(x)$$

由于 $f(x), g(x)$ 互素, 得到 $f(x) | (v_1(x) - v(x))$, 而后者次数低于前者, 只能 $v_1(x) = v(x)$

1.1.4 不可约多项式与因式分解

定义 1.4 (不可约多项式)

设 $p(x) \in \mathbb{P}[x]$ 且 $\partial(p(x)) \geq 1$, 若 $p(x)$ 不能表示成 $\mathbb{P}[x]$ 中两个次数小于 $p(x)$ 的多项式的乘积, 则称 $p(x)$ 为 $\mathbb{P}[x]$ 中的不可约多项式。



定理 1.7 (不可约多项式的性质)

若 $p(x) \in \mathbb{P}[x]$ 是不可约多项式, 则

- $p(x)$ 与 $\forall f(x) \in \mathbb{P}[x]$ 要么互素, 要么 $p(x) | f(x)$
- 若 $p(x) | f(x)g(x)$, 则要么 $p(x) | f(x)$ 要么 $p(x) | g(x)$



 **笔记** 不可约多项式经常用于证明乘积形式 $(h(x), f(x)g(x)) = 1$ 的命题。例如反设 $(h(x), f(x)g(x)) = d(x)$, 不可约多项式 $p(x) | d(x)$, 则 $p(x)$ 要么整除 $f(x)$, 要么整除 $g(x)$

 **练习 1.6 利用不可约多项式性质 + 反证法** (1) 重点: 已知 $(f(x), g(x)) = 1$, 证明 $(f(x)g(x), f(x) + g(x)) = 1$

(2) 已知 $f_1(x), f_2(x), \dots, f_s(x)$ 两两互素, 记 $F_i(x) = f_1(x) \cdots f_{i-1}(x) f_{i+1}(x) \cdots f_s(x)$, 证明 $F_1(x), F_2(x), \dots, F_s(x)$ 互素

(3) 稍加变化: 复数域上的多项式 $f(x)$ 没有重因式, 证明 $(f(x) + 2f'(x), 3f^2(x)) = 1$

证明 (1) 反设有不可约多项式 $q(x)$ 作为公因式, 则 $q(x)$ 同时是 $f(x)g(x), f(x) + g(x)$ 因式。根据前者得到 $q(x)$ 至少整除 $f(x), g(x)$ 之一, 不妨设 $q(x) | f(x)$, 根据第二个式子得到 $q(x)$ 也一定整除 $g(x)$, 否则无法得到多项式。因此矛盾。

(2) 同种方法, 随便挑 $F_i(x), F_j(x)$ 说一说就行

(3) 同样设不可约 $q(x)$, 根据 $3f^2(x)$ 可知 $q(x) | f(x)$, 再根据 $q(x) | f(x) + 2f'(x)$ 可知 $q(x) | f'(x)$ (否则 $\frac{f(x) + 2f'(x)}{q(x)}$ 不是多项式), 根据重因式充要条件可知。

练习 1.7 证明不可约多项式 (1) $p(x)$ 是次数大于 0 的多项式, 若对任意多项式 $f(x), g(x)$, 根据 $p(x) | f(x)g(x)$ 可以得到 $p(x) | f(x)$ 或者 $p(x) | g(x)$, 则 $p(x)$ 是不可约多项式。

证明 (1) 反设 $p(x)$ 不是不可约多项式, 则存在 $f(x), g(x)$ 使得 $p(x) = f(x)g(x)$, 此时 $p(x) | f(x)g(x)$, 但是 $p(x) \nmid f(x), p(x) \nmid g(x)$

定理 1.8 (因式分解唯一性)

$f(x)$ 是 $\mathbb{P}$ 上次数大于等于 1 的多项式, 则 $f(x)$ 可唯一分解为 $\mathbb{P}$ 上有限个不可约多项式乘积:

$$f(x) = cp_1^{r_1}(x)p_2^{r_2}(x) \cdots p_s^{r_s}(x)$$



1.1.5 余数定理及其应用

定理 1.9 (余数定理)

$f(x) \in \mathbb{P}[x]$, 对 $\forall \alpha \in \mathbb{P}$, 总有

$$f(x) = (x - \alpha)q(x) + f(\alpha)$$



证明 用带余除法得到 $f(x) = (x - \alpha)q(x) + c$, 而代入 $x = \alpha$ 得到 $f(\alpha) = 0 + c$, 即 $c = f(\alpha)$, 因此

$$f(x) = (x - \alpha)q(x) + f(\alpha)$$

推论 1.2

$\alpha \in \mathbb{P}$ 是多项式 $f(x)$ 的根当且仅当 $(x - \alpha) | f(x)$



笔记 余数定理的主要价值在于 $\mathbb{P}$ 上的因式等价于 $\mathbb{P}$ 上的根, 两者可以灵活互换。判断重根可以转换为判断重因式, 判断整除可以转换为根。

练习 1.8 余数定理的应用 (1) 若 $(x - 1) | f(x^m)$, 证明 $(x^m - 1) | f(x^m)$

(2) 证明 $x^2 + x + 1 | x^{3m} + x^{3n+1} + x^{3p+2}$

(3) 已知 $x^2 + x + 1 | f_1(x^3) + xf_2(x^3)$, 证明 $(x - 1) | f_1(x)$ 且 $(x - 1) | f_2(x)$

(4) 已知 $x^2 + 1 | f_1(x^4) + xf_2(x^4)$, 证明 1 是 $f_1(x), f_2(x)$ 的公共根。

证明 (1) 由于 $(x - 1) | f(x^m)$, 因此 $x = 1$ 是 $f(x^m)$ 的根, 故 $x = 1$ 是 $f(x)$ 的根, 因此 $(x - 1) | f(x)$, 故 $(x^m - 1) | f(x^m)$

(2) 放到 $\mathbb{C}$ 上考虑: $\omega = e^{i\frac{2\pi}{3}}$, 则 $x^2 + x + 1 = (x - \omega)(x - \omega^2)$, 代入 ω 则 $\omega^{3m} + \omega^{3n+1} + \omega^{3p+2} = 1 + \omega + \omega^2 = 0$, 代入 ω^2 则 $\omega^{6m} + \omega^{6n+2} + \omega^{6p+4} = 1 + \omega^2 + \omega^4 = 1 + \omega^2 + \omega = 0$, 根据余数定理可知成立

(4) $x^2 + 1$ 复数根为 $x = i$, 根据 $(x^2 + 1) | f_1(x^4) + xf_2(x^4)$ 可知 $f_1(1) + if_2(1) = 0$, 从而为公共根。

1.1.6 重因式与重根

定理 1.10 (重因式性质汇总)

$f(x)$ 是次数大于 0 的多项式, $p(x)$ 是不可约多项式, 则

- 核心: 若 $p(x)$ 是 $f(x)$ 的 k 重因式, 则其是 $f'(x)$ 的 $k-1$ 重因式。
- 无重因式充要: $f(x)$ 没有重因式当且仅当 $f(x), f'(x)$ 互素。
- 不可约因式: $\frac{f(x)}{(f(x), f'(x))}$ 是 $f(x)$ 所有不可约因式 (重数均为 1) 乘积的常数倍



笔记 重因式与重根的关系: 重因式不随数域而改变, 而重根会随着数域而改变。例如 $(x^2+1)^2$, 其在 $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ 上都有重因式, 但是只在 $\mathbb{C}$ 上有重根。因此有重根一定有重因式, 有重因式不一定有重根。

练习 1.9 用根判断因式 (1) 重点: 设 $h(x), f(x), g(x)$ 为实系数多项式满足以下等式, 证明 $x^2+1|f(x), x^2+1|g(x)$

$$\begin{cases} (x^2+1)h(x) + (x+1)f(x) + (x-2)g(x) = 0 \\ (x^2+1)k(x) + (x-1)f(x) + (x+2)g(x) = 0 \end{cases}$$

证明 (1) 要证明 $\mathbb{R}$ 上 $(x^2+1)|f(x)$, 只需要证明 $\mathbb{C}$ 上 $(x-i)|f(x), (x+i)|f(x)$, 等价于 $x=\pm i$ 在 $\mathbb{C}$ 上为 $f(x)$ 的根。代入 $x=i$ 得到:

$$\begin{cases} (i+1)f(i) + (i-2)g(i) = 0 \\ (i-1)f(i) + (i+2)g(i) = 0 \end{cases}$$

解得 $f(i) = g(i) = 0$ 。同理可证 $f(-i) = g(-i) = 0$

定理 1.11 (重根性质汇总)

$f(x)$ 是 $\mathbb{P}[x]$ 上次数大于 0 的多项式, 则

- 重根: x_0 是 $f(x)$ 的 k 重根当且仅当 $f(x_0) = f'(x_0) = \dots = f^{(k-1)}(x_0) = 0$, 而 $f^{(k)}(x_0) \neq 0$
- 重根判定定理: $p(x)$ 是 $f(x)$ 重因式当且仅当它是 $f(x), f'(x)$ 公因式。



笔记 $\mathbb{C}$ 上的重根问题等价于 $\mathbb{C}$ 上的重因式问题, 因此判断多项式是否有重根时, 可判断其在 $\mathbb{C}$ 上是否有重因式。重因式的好处是可以用多项式理论, 重根的好处是可以带进去算, 要相互补充。

练习 1.10 判定重根存在性 (1) 证明 $f(x) = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}$ 没有重根
(2) 证明 $f(x) = x^n + nx^{n-1} + n(n-1)x^{n-2} + \dots + n!$ 无重根

证明 (1) 直接证重根不好证, 转换为重因式, 若在 $\mathbb{C}$ 上无重因式则肯定无重根。根据 $f'(x) = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^{n-1}}{(n-1)!}$ 得到

$$(f(x), f'(x)) = (f(x), f(x) - f'(x)) = (f(x), \frac{x^n}{n!}) = (f(x), x^n)$$

由于 x^n 项的原因, $(f(x), f'(x)) = x^k, k \geq 0$ 。而 $f(0) = 1$, 这说明 x 不是 $f(x)$ 的因式, 从而 $k = 0$, 因此 $f(x), f'(x)$ 互素, 无重根。

(2) 也用无重因式证明

练习 1.11 重根、重因式相互转化 (1) 求 $x^5 + x^4 - 6x^3 - 14x^2 - 11x - 3$ 的重根和重数

(2) 重点: 求 A, B 使得 $(x-1)^2 | Ax^4 + Bx^2 + 1$

(3) 重点: 求 t 使得 $f(x) = x^3 - 3x^2 + tx - 1$ 有重根

(4) 重点 ZJU2018.2: 证明 $x^3 + px + q$ 有重根的充要条件 $4p^3 + 27q^2 = 0$

(5)ZJU2021: 求 t 为何值时, $f(x) = x^3 + 6x^2 + tx + 8$ 有重根, 并求重根

解 (1) 可以尝试猜。一般用辗转相除法算 $f(x), f'(x)$ 的最大公因式, 最终答案为 $(x+1)^4(x-3)$

(2) 根据重根与重因式关系, 在 $\mathbb{C}$ 上必有 $f(1) = 0, f'(1) = 0$, 因此得出 $A+B+1=0, 4A+2B=0$, 得到 $A=-2, B=1$

(3) 同理考虑 $\mathbb{C}$ 上的因式分解, $f'(x) = 3x^2 - 6x + t$, 因此考虑 $f(x)$ 与 $\frac{1}{3}f'(x) = x^2 - 2x + \frac{t}{3}$ 辗转相除法, 第一步得出 $f(x) = \frac{1}{3}f'(x)(x-1) + (\frac{2}{3}t-2)x + (\frac{t}{3}-1)$, 如果 $t=3$ 则直接整除有重根, 如果 $t \neq 3$, 则接着算, $\frac{1}{3}f' = (\frac{1}{\frac{2}{3}t-2})r_1(x)(x-\frac{5}{2}) + \frac{t}{3} + \frac{5}{4}$, 则需要 $\frac{t}{3} + \frac{5}{4} = 0$, 得到 $t = -\frac{15}{4}$

(4) 在 $\mathbb{C}$ 上考虑, 有重根即有重因式, $f'(x) = 3x^2 + p$, 先有 $f(x) = \frac{1}{3}f'(x)x + \frac{2}{3}px + q$, 若 $p=q=0$, 则成立, 下面考虑 $p \neq 0$ 。得到 $r_1(x) = \frac{3}{2p}r_1(x)(x - \frac{3q}{2p}) + \frac{p}{3} + \frac{9q^2}{4p^2}$, 因此得到 $4p^3 + 27q^2 = 0$

(5) $t=12$ 时有 3 重 $x=-2$ 根, $t=-15$ 时有 2 重 $x=1$ 根。

练习 1.12 导数重根反推函数重根? x_0 是 $f'(x)$ 的 k 重根, 举例说明其不能推出 x_0 是 $f(x)$ 的 $k+1$ 重根, 该命题需要 x_0 是 $f(x)$ 的根。

解 例如 $f(x) = x^{k+1} + 1, x_0 = 0$

笔记 上面的结论非常重要, 在利用导数验证根的重数时, 一定要验证其是 $f(x)$ 本身的根。

1.1.7 整数与多项式

定理 1.12 (整数与多项式)

m, n 为正整数, $(m, n) = d$ 的充要条件是 $(x^m - 1, x^n - 1) = x^d - 1$



证明 (1) 左推右是公因式: 设 $m = kd$, 则 $x^m - 1 = (x^d)^k - 1 = (x^d - 1)(1 + x^d + \cdots + x^{(k-1)d})$, 这说明 $x^d - 1 | x^m - 1$, 另一侧同理, 因此是公因式

(2) 最大: 设 $p(x)$ 是 $x^n - 1, x^m - 1$ 的公因式, 下面证明 $p(x) | x^d - 1$ 。由于 $x^n - 1$ 在 $\mathbb{C}$ 上无重根 (因为和导数互素), 因此 $p(x)$ 也没有重根, 从而 $p(x) = a(x - x_1) \cdots (x - x_s)$, 这里 $x_1 \sim x_s$ 是互异复数, 由于 $x_i^n = x_i^m = 1$, 得到 $x_i^d = x_i^{um+vn} = (x_i^m)^u (x_i^n)^v = 1$, 因此 $x - x_i | x^d - 1, (x - x_1) \cdots (x - x_s) | x^d - 1$, 因此 $p(x) | x^d - 1$

(3) 右推左: 若 $(m, n) = d_1$, 则根据前面的结论 $(x^m - 1, x^n - 1) = x^{d_1} - 1 = x^d - 1$, 因此 $d_1 = d$

练习 1.13 整除与多项式 (1) $d, n \in \mathbb{N}^+$, 证明 $x^d - 1 | x^n - 1$ 的充要条件为 $d | n$

(2) m, n 是大于 1 的自然数, 证明: m, n 互素的充要条件是 $1 + x + x^2 + \cdots + x^{m-1}$ 与 $1 + x + x^2 + \cdots + x^{n-1}$ 互素

(3) $g(x) = 1 + x + x^2 + \cdots + x^n, f(x) = g(x^k), (k, n+1) = 1$, 证明: $g(x) | f(x)$

证明 (1) 本题即前面题目的特例。

右推左: 设 $x^d - 1 = (x - x_1) \cdots (x - x_d) (\mathbb{C} \text{ 上})$, 因此 $x_i^d = 1$, 由于 $n = dk$, 因此 $x_i^n = (x_i^d)^k = 1$, 因此 $x^d - 1 | x^n - 1$ 。

左推右: 设 $x^d - 1 = (x - x_1) \cdots (x - x_d)$, 由于 $x^d - 1 | x^n - 1, x_i^d = x_i^n = 1$, 假设 $n = kd + r, r \neq 0$, 则 $x_i^n = x_i^{kd+r} = x_i^r$, 由于 $x_i = 1$ 只有一重, 其他 $x_i \neq 1$, 对于这些 x_i 有 $x_i^r \neq 1$, 因此矛盾。

(2) m, n 互素当且仅当 $(x^m - 1, x^n - 1) = x - 1$, 等价于 $(1 + x + x^2 + \cdots + x^{m-1}, 1 + x + \cdots + x^{n-1}) = 1$

(3) 根据 $g(x) = \frac{x^{n+1} - 1}{x - 1}, f(x) = \frac{x^{k(n+1)} - 1}{x^k - 1}$, 此时 $g(x) | f(x)$ 等价于 $(x^k - 1)(x^{n+1} - 1) | (x - 1)(x^{k(n+1)} - 1)$, 由于 $(k, n+1) = 1$, 因此 $x^k - 1$ 与 x^{n+1} 的最大公因为 $(x - 1)$, 而 $(x^{k(n+1)} - 1)$ 是 $(x^k - 1)(x^{n+1} - 1)$, 因此根据最小公倍式的性质显然可知。

1.2 特殊数域上的一元多项式理论

1.2.1 复数域上的一元多项式

定理 1.13 (代数学基本定理)

任意次数大于等于 1 的多项式在 $\mathbb{C}$ 上至少有一个根。



定理 1.14 (复数域因式分解)

$\mathbb{C}$ 上任意次数大于等于 1 的多项式可唯一地分解为一次因式的乘积, 因此 n 次多项式在 $\mathbb{C}$ 上恰有 n 个复根。



1.2.2 实数域上的一元多项式

定理 1.15 (实数域因式分解)

$\mathbb{R}$ 上的多项式总能唯一地分解为一次因式和不可约二次因式的乘积。



1.2.3 有理数域上一元多项式因式分解

复不可约多项式是一次的, 实不可约多项式是一次或二次的, 下面将说明有理不可约多项式的次数可以是任意的。

定义 1.5 (本原多项式)

若整系数多项式 $g(x) = b_n x^n + \cdots + b_1 x + b_0$ 的系数 $b_0 \sim b_n$ 是互素的 (整体互素而非两两互素), 即

$$(b_1, \cdots, b_n) = 1$$

则称 $g(x)$ 是一个本原多项式



定理 1.16 (Gauss 引理)

两个本原多项式的乘积仍为本原多项式



证明 设 $f(x) = \sum_{k=0}^n a_k x^k, g(x) = \sum_{k=0}^m b_k x^k$, 设 $h(x) = f(x)g(x) = \sum_{k=0}^{n+m} d_k x^k$ 。假设 $h(x)$ 不是本原多项式, 则存在素数 p 整除所有 d_k , 由于 $f(x), g(x)$ 本原, 因此从常数项开始搜索, 存在 i, j , 使得

$$\begin{cases} p \mid a_0, \cdots, p \mid a_{i-1}, p \nmid a_i \\ p \mid b_0, \cdots, p \mid b_{j-1}, p \nmid b_j \end{cases}$$

而 $d_{i+j} = a_i b_j + \sum_{k=0, k \neq i}^{i+j} a_k b_{i+j-k}$, 除了第一项外均至少有一项被 p 整除, 第一项不被 p 整除, 因此 $p \nmid d_{i+j}$, 这与假设矛盾。

定理 1.17 (有理多项式变本原多项式)

设 $f(x)$ 是有理多项式, $g(x)$ 为本原多项式, 则 $f(x)$ 可表示为有理数 r 与本原多项式 $g(x)$ 的乘积



笔记 有理多项式分解说明了有理多项式的因式分解问题完全可以转换成本原多项式的因式分解, 而下面的定理说明了本原多项式能否分解为两个有理多项式之积与能否分解为整系数多项式之积是等价的。

 **练习 1.14 本原多项式** (1) $f(x), g(x)$ 是本原多项式, $g(x) \mid f(x)$, 证明: $\frac{f(x)}{g(x)}$ 也是本原多项式

(2) $f(x)$ 是本原多项式, 证明: $f(x+1)$ 也是本原多项式。

证明 (1) 设 $f(x) = g(x)q(x)$, 由于 $f(x), g(x)$ 为本原多项式, $q(x)$ 必然是有理系数的, 设 $q(x) = r q_1(x)$, 这里 $q_1(x)$ 为本原多项式, 因此

$$f(x) = r g(x) q_1(x)$$

根据 Gauss 引理可知 $g(x)q_1(x)$ 仍为本原多项式, 因此 $r = \pm 1$, 因此 $q(x) = \pm q_1(x)$ 是本原多项式。

(2) 直接展开得到 $f(x+1) = b_n x^n + \cdots + b_1 x + b_0$, 其中 $b_i = \sum_{k=i}^n \binom{k}{i} a_k$ 。下面证明互素, 假设 $(b_0, b_1, \cdots, b_n) = d$, 则

$$\begin{cases} d \mid b_n \Rightarrow d \mid a_n \\ d \mid b_{n-1} \Rightarrow d \mid \binom{n}{1} a_n + a_{n-1} \Rightarrow d \mid a_{n-1} \end{cases}$$

同理可以得到 $d \mid a_{n-2}, \cdots, d \mid a_1, d \mid a_0$, 这与 $f(x)$ 是本原多项式矛盾。

定理 1.18 (整系数多项式因式分解)

整系数多项式若能分解为两个较低次有理多项式乘积, 则一定也能分解为两个较低次整系数多项式乘积 

证明 设 $f(x) = g(x)h(x)$, 这里 $g(x), h(x) \in \mathbb{Q}[x]$, 则存在本原多项式 $f_1(x), h_1(x), g_1(x)$ 使得

$$f(x) = a f_1(x), g(x) = r g_1(x), h(x) = s h_1(x), a \in \mathbb{Z}, r, s \in \mathbb{Q}$$

从而 $a f_1(x) = r s g_1(x) h_1(x)$, 根据 Gauss 引理可知 $g_1(x) h_1(x)$ 是本原多项式, 不妨设最高项系数均正, 则 $f_1(x) = g_1(x) h_1(x), a = r s$, 得到 $r s \in \mathbb{Z}$, 于是

$$f(x) = (r s g_1(x)) h_1(x)$$

其中 $r s g_1(x), h_1(x)$ 都是整系数多项式。

 **练习 1.15 整系数多项式整数根** (1) 若 $\alpha \neq \pm 1$ 为整系数多项式 $f(x)$ 的整数根, 证明: $\frac{f(1)}{\alpha-1}, \frac{f(-1)}{\alpha+1}$ 都是整数

(2) ZJU2017.1: $f(x)$ 为整系数多项式, $f(0), f(1)$ 为奇数, 证明: $f(x)$ 无整数根

(3) ZJU2019.4: $a_1, \cdots, a_n$ 为互不相同的整数, $a_1 \cdots a_n + 1$ 不是某个整数的平方, 证明 $f(x) = (x+a_1) \cdots (x+a_n) + 1$ 不能表示为 $\mathbb{Q}$ 上两个次数大于等于 1 的多项式的乘积。

证明 (1) 由于 α 是 $f(x)$ 的根, 因此 $f(x) = (x-\alpha)(b_{n-1}x^{n-1} + \cdots + b_1x + b_0)$, 这里 $b_0, \cdots, b_{n-1}$ 也都是整数, 因此

$$f(1) = (1-\alpha)(b_{n-1} + \cdots + b_1 + b_0) \Rightarrow \frac{f(1)}{\alpha-1} = -b_0 - \cdots - b_{n-1} \in \mathbb{Z}$$

(2) 若 f 有整数根, 则 $f(x) = (x-a)f_0(x)$, 这里 $f_0(x)$ 为整系数多项式, 而 $f(0) = (-a)f_0(0), f(1) = (1-a)f_0(1)$, 而根据 $f(0), f(1)$ 为奇数得到 $-a, 1-a$ 同时为奇数, 矛盾。

(3) 若 $f(x)$ 在 $\mathbb{Q}[x]$ 上可约, 则 $f(x)$ 在 $\mathbb{Z}[x]$ 也可约, 不妨设 $f(x) = g(x)h(x), g(x), h(x) \in \mathbb{Z}[x]$, 则 $f(-a_i) = g(-a_i)h(-a_i) = 1$, 这说明 $g(-a_i) = h(-a_i) = \pm 1$, 因此 $g(x) - h(x)$ 有 n 个不同的根, 而 $\partial(g(x) - h(x)) < n$, 因此 $g(x) \equiv h(x)$, 根据条件 $f(x) = g^2(x)$, 常数项为整数的平方, 这与条件不是平方矛盾。

1.2.4 一元有理多项式的根

定理 1.19 (有理多项式的根)

$f(x) = a_n x^n + \cdots + a_1 x + a_0$ 是整系数多项式, 若既约分数 $\alpha = \frac{q}{p}$ 是 $f(x)$ 的有理根, 则必有 $q \mid a_0, p \mid a_n$

证明 假设 $\frac{q}{p}$ 是解, 则代入多项式得到:

$$a_n \frac{q^n}{p^n} + \cdots + a_1 \frac{q}{p} + a_0 = 0$$

两边乘上 p^n 并移项得到:

$$\begin{cases} a_n q^n = p(-a_{n-1} q^{n-1} - a_{n-2} p q^{n-2} - \cdots - a_0 p^{n-1}) \\ a_0 p^n = q(-a_n q^{n-1} - a_{n-1} p q^{n-2} - \cdots - a_{n-1} p^{n-1}) \end{cases}$$

因为 $(p, q) = 1$, 因此 $(p, q^n) = 1, (q, p^n) = 1$, 因此根据上面两个式子得到

$$p \mid a_0, q \mid a_n$$

练习 1.16 计算有理根 (1) 证明: $f(x) = x^3 + 2x^2 + x + 1$ 在有理数域上不可约

(2) 求 $f(x) = 3x^4 + 5x^3 + x^2 + 5x - 2$ 的有理根

证明 (1) 有理数域不可约即无有理根。而有理根只可能是 ± 1 , $f(\pm 1) \neq 0$, 因此不可约。

(2) 全部列出来验证, 结果为 $\frac{1}{3}, -2$

定理 1.20 (Eisenstein 判别法)

$f(x) = a_n x^n + \cdots + a_1 x + a_0$ 是整系数多项式, 若存在素数 p 使得: (1) $p \nmid a_n$ (2) $p \mid a_{n-1}, \cdots, a_0$ (3) $p^2 \nmid a_0$, 则 $f(x)$ 在有理数域上是不可约多项式。

练习 1.17 使用 Eisenstein 判别法 证明在 $\mathbb{Q}$ 不可约

(1) $f(x) = 2x^4 + 3x^3 - 9x^2 - 3x + 6$

(2) $f(x) = x^n + 2x + 2$

证明 (1) 取 $p = 3$ 即可

(2) 取 $p = 2$

1.3 多元多项式

1.3.1 基本概念

定义 1.6 (字典排序)

对于多元多项式 $p(x_1, x_2, \cdots, x_n)$ 的乘积项, 先按 x_1 次数由高到低排序, 再根据 $x_2, x_3, \cdots$ 次数排序, 将字典排序得到的第一项称为首项。

定理 1.21 (乘积保首项)

$f(x_1, \cdots, x_n), g(x_1, \cdots, x_n)$ 为多元多项式, 则它们乘积的首项等于首项的乘积。

1.3.2 对称多项式

定义 1.7 (对称多项式)

若多元多项式 $f(x_1, \dots, x_n)$ 互换任意 x_i, x_j 后多项式保持不变, 则称其为对称多项式

定理 1.22 (对称多项式的性质)

对称多项式的和、差、乘积、多项式复合仍为对称多项式。

定义 1.8 (初等对称多项式)

将形如 $\sigma_i = \sum_{k_1 < k_2 < \dots < k_i} x_{k_1} x_{k_2} \dots x_{k_i}$ 的对称多项式称为初等对称多项式:

$$\begin{cases} \sigma_1 = x_1 + x_2 + \dots + x_n \\ \sigma_2 = x_1 x_2 + x_1 x_3 + \dots + x_{n-1} x_n \\ \vdots \\ \sigma_n = x_1 x_2 \dots x_n \end{cases}$$

定理 1.23 (Vieta 定理)

对于 $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$, 假设其解为 $x_1, \dots, x_n$, 则

$$\begin{cases} a_{n-1} = -\sigma_1 \\ a_{n-2} = \sigma_2 \\ \vdots \\ a_{n-i} = (-1)^i \sigma_i \\ \vdots \\ a_0 = (-1)^n \sigma_n \end{cases}$$

定理 1.24 (复合定理)

任意对称多项式 $f(x_1, \dots, x_n)$ 可唯一地写成初等多项式的多项式 $\varphi(\sigma_1, \dots, \sigma_n)$

证明 设 $f(x_1, \dots, x_n)$ 的首项为 $ax_1^{l_1}x_2^{l_2}\dots x_n^{l_n}$, 且 $l_1 \geq l_2 \geq \dots \geq l_n$ (根据对称多项式可交换性可知该性质), 做多项式 $\varphi_1 = a\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\dots\sigma_n^{l_n}$, 得到 φ_1 的首项为 $ax_1^{l_1-l_2}(x_1x_2)^{l_2-l_3}\dots(x_1\dots x_n)^{l_n} = ax_1^{l_1}\dots x_n^{l_n}$, 取 $f_1(x) = f(x) - \varphi_1$, 以此类推可以得到结果。

练习 1.18 (1) 将 $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$ 表达为 $\sigma_1, \sigma_2, \sigma_3$ 的多项式

解 (1) 根据首项可知 $l_1 = 3, l_2 = l_3 = 0$, $\varphi_1 = \sigma_1^{l_1}\sigma_2^{l_2}\sigma_3^{l_3} = (x_1 + x_2 + x_3)^3$, 得到

$$f_1(x) = f(x) - \varphi_1 = -3(x_1^2x_2 + x_1^2x_3 + \dots) - 6x_1x_2x_3$$

$f_1(x)$ 首项为 $-3x_1^2x_2$, 因此 $\varphi_2 = -3\sigma_1^{2-1}\sigma_2 = -3(x_1^2x_2 + x_2^2x_1 + \dots) - 9x_1x_2x_3$, 因此 $f_2(x_1, x_2, x_3) = 3\sigma_3$ 。综上 $f(x_1, x_2, x_3) = 3\sigma_3 - 3\sigma_1\sigma_2 + \sigma_1^3$

1.3.3 判别式

对称多项式理论可应用于判断一元高次多项式重根的存在性, 这就导出了判别式的概念。

定义 1.9 (判别式)

给定多项式 $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{C}[x]$, 设 $x_1, \cdots, x_n$ 是 $f(x)$ 的根, 则定义判别式为

$$D(f) := \prod_{1 \leq i < j \leq n} (x_i - x_j)^2$$

显然 $f(x)$ 有重根当且仅当 $D(f) = 0$



定理 1.25 (判别式的计算)

由于 $D(f)$ 是 $x_1, \cdots, x_n$ 的对称多项式, 因此可以表示为初等对称多项式的多项式, 再根据 Vieta 定理即可表达为 $a_0, \cdots, a_{n-1}$ 的多项式



练习 1.19 判别式的计算 (1) 计算 $f(x) = x^2 + px + q$ 的判别式

解 (1) 设 $g(x_1, x_2) = (x_1 - x_2)^2$, 可以将 $g(x_1, x_2)$ 表达为初等对称多项式的多项式:

$$g(x_1, x_2) = \sigma_1^2 - 4\sigma_2$$

而根据 Vieta 定理, $p = -\sigma_1, q = \sigma_2$, 因此 $D(f) = p^2 - 4q$

1.3.4 结式

给定多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, g(x) = b_m x^m + \cdots + b_0$, 虽然可以用辗转相除法计算它们的最大公因式, 但判断它们是否互素有更好的解决办法, 这就是结式

定义 1.10 (结式)

考虑两个多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_0$, 则定义两个矩阵:

$$A = \begin{bmatrix} a_n & a_{n-1} & a_{n-2} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & a_{n-2} & \cdots & a_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_n & a_{n-1} & a_{n-2} & \cdots & a_0 \end{bmatrix}_{m \times (n+m)}$$

$$B = \begin{bmatrix} b_m & b_{m-1} & b_{m-2} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & b_{m-1} & b_{m-2} & \cdots & b_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & 0 & \cdots & b_m & b_{m-1} & b_{m-2} & \cdots & b_0 \end{bmatrix}_{n \times (n+m)}$$

则定义它们的结式 $R(f, g) = \begin{vmatrix} A \\ B \end{vmatrix}$



定理 1.26 (结式判定定理)

给定 $f(x) = a_n x^n + \cdots + a_0, g(x) = b_m x^m + \cdots + b_0, a_n, b_m \neq 0$, 则 f, g 不互素的充要条件为 $R(f, g) = 0$



证明 设 $\partial(f) = n, \partial(g) = m$, 首先 $f(x), g(x)$ 不互素当且仅当 $\exists u(x), v(x)$ 满足 $\partial(u) < \partial(g), \partial(v) < \partial(f)$, 使得

$u(x)f(x) = v(x)g(x)$, 不妨设

$$\begin{cases} u(x) = u_{m-1}x^{m-1} + \cdots + u_0 \\ v(x) = v_{n-1}x^{n-1} + \cdots + v_0 \end{cases}$$

将 $u(x)f(x) = v(x)g(x)$ 展开对比系数得到

$$\begin{cases} a_n u_{m-1} = b_m v_{n-1} \\ a_{n-1} u_{m-1} + a_n u_{m-2} = b_{m-1} v_{n-1} + b_m v_{n-2} \\ \vdots \\ a_0 u_0 = b_0 v_0 \end{cases}$$

将上述线性方程组看成 $u_0, \dots, u_{m-1}, v_0, \dots, v_{n-1}$ 的线性方程组, 则系数矩阵的转置 C^T 为结式定义中的矩阵:

$$C^T = \begin{pmatrix} A \\ -B \end{pmatrix}$$

因此有解当且仅当 $|C^T| = 0$, 也等价于 $R(f, g) = 0$

定理 1.27 (结式计算公式)

给定前面的 $f(x), g(x)$, $\alpha_1, \dots, \alpha_n$ 和 $\beta_1, \dots, \beta_m$ 分别是 $f(x), g(x)$ 的复根, 则

$$R(f, g) = a_n^m b_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j)$$



定理 1.28 (判别式计算公式)

$\mathbb{C}[x]$ 中多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$, $a_n \neq 0$, 则 $f(x)$ 的判别式

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-(2n-1)} R(f, f')$$



第2章 矩阵

2.1 矩阵基础

2.1.1 矩阵概念与运算

略

2.1.2 矩阵的逆

定义 2.1 (矩阵的逆)

若 $AB = BA = E$, 则称 B 为 A 的逆, 记 A^{-1}

定理 2.1 (矩阵逆的基本性质)

(1) $(kA)^{-1} = k^{-1}A^{-1}$ (2) $(AB)^{-1} = B^{-1}A^{-1}$ (3) $(A^T)^{-1} = (A^{-1})^T$ (4) $|A^{-1}| = |A|^{-1}$

证明 (2) $(AB)^{-1}AB = E$, 因此 $(AB)^{-1} = B^{-1}A^{-1}$

(3) $(A^{-1})^T A^T = (AA^{-1})^T = E$

(4) $|A^{-1}A| = |A^{-1}| \cdot |A| = 1$, 因此 $|A^{-1}| = |A|^{-1}$

定理 2.2 (可逆充要条件)

矩阵 A 可逆的充要条件如下:

- $|A| \neq 0$
- A 可表示为初等矩阵的积

定理 2.3 (利用初等行变换求逆)

若 A 可逆, 则利用初等行变换可以得到 $(A, E) \rightarrow (E, A^{-1})$ 。特别地, $(A, B) \rightarrow (E, A^{-1}B)$

 **笔记** 同样可以通过初等列变换得到: $\begin{pmatrix} A \\ B \end{pmatrix} \rightarrow \begin{pmatrix} E \\ BA^{-1} \end{pmatrix}$

 **练习 2.1 利用初等变换求逆** (1) 计算下面矩阵的逆:

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

解 结果为 $\frac{1}{3} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -2 \\ -3 & 2 & -1 \end{bmatrix}$

 **练习 2.2 逆矩阵理论练习** (1) A 为方阵, 若 $A^k = O$, 证明 $E - A$ 可逆, 且 $(E - A)^{-1} = E + A + A^2 + \cdots + A^{k-1}$

(2) A, B 为 n 阶方阵, 若 $A + B = AB$, 证明: $A - E$ 可逆, 并求逆矩阵

(3) 若 $A + B$ 可逆, 证明 $A^{-1} + B^{-1}$ 也可逆, 并求其逆

解 (1) 直接做乘法验证:

$$(E - A)(E + A + A^2 + \cdots + A^{k-1}) = E + A + A^2 + \cdots + A^{k-1} - A - A^2 - \cdots - A^k = E - A^k = E$$

(2) 由于 $A + B = AB$, 则 $(A - E)(B - E) = AB - (A + B) + E = E$, 因此 $(A - E)^{-1} = B - E$

(3) 根据 $B(A^{-1} + B^{-1})A = (A + B)$ 得到

$$(A^{-1} + B^{-1})B = A^{-1}(A + B) \Rightarrow (A^{-1} + B^{-1})B(A + B)^{-1} = A^{-1}$$

因此 $(A^{-1} + B^{-1})^{-1} = B(A + B)^{-1}A$

定义 2.2 ((代数) 余子式)

$A = (a_{ij})$ 为 $n \times n$ 矩阵, 划去 i 行 j 列的行列式称为余子式 M_{ij} , 代数余子式为 $(-1)^{i+j}M_{ij}$

定义 2.3 (伴随矩阵)

$A = (a_{ij})_n$, A_{ij} 是 a_{ij} 的代数余子式, 则其伴随为 (注意行列颠倒!!):

$$A^* = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

定理 2.4 (利用伴随求逆)

A 及其伴随矩阵 A^* 满足 $AA^* = |A|E$, 若 A 可逆, 则 $A^{-1} = \frac{1}{|A|}A^*$

证明 考虑 $B = AA^*$, 则 $B_{ij} = \sum_{k=1}^n a_{ik}A_{kj}$, 根据行列式的展开式可知

$$B_{ij} = \begin{cases} |A|, & i = j \\ 0, & i \neq j \end{cases}$$

因此 $AA^* = |A|E$

2.1.3 伴随矩阵的性质

定理 2.5 (伴随运算的性质)

A 为 n 阶方阵, 则

1. $(A^*)^T = (A^T)^*$
2. $(A^*)^{-1} = (A^{-1})^*$
3. $(AB)^* = B^*A^*$

证明 (1) 设 $A = (a_{ij})$, 则

$$(A^*)^T = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

同理设 $(A^T)^* = (B_{ij})$, 则 $B_{ij} = A_{ji}^T$, 因此行列式大小相同。

(2) $A^* = |A|A^{-1}$, 则 $(A^*)^{-1} = \frac{1}{|A|}A$ 因此

$$(A^{-1})^* = |A^{-1}|A = \frac{1}{|A|}A$$

因此结论成立

$$(3)(AB)^* = |AB|(AB)^{-1} = |B|(B)^{-1}|A|A^{-1} = B^*A^*$$

定理 2.6 (伴随矩阵的行列式)

$$|A^*| = |A|^{n-1}$$



证明 $|A^*| = ||A|A^{-1}| = |A|^n \cdot |A^{-1}| = |A|^{n-1}$

定理 2.7 (伴随矩阵的秩)

A^* 为 A 的伴随矩阵, 则

$$r(A^*) = \begin{cases} n, & r(A) = n \\ 1, & r(A) = n-1 \\ 0, & r(A) < n-1 \end{cases}$$



证明 $r(A) = n$ 时, $|A^*| = |A|^{n-1} \neq 0$, 因此满秩。 $r(A) = n-1$ 时, 由于 $AA^* = |A|E = O$, 根据秩不等式:

$$r(AA^*) \geq r(A) + r(A^*) - n = r(A^*) - 1 \Rightarrow r(A^*) \leq 1$$

而由于 $r(A) = n-1$, 其至少有一个 $n-1$ 阶代数余子式不为 0, 因此 $r(A^*) \geq 1$, 综上 $r(A^*) = 1$ 。其余显然

定理 2.8 (伴随矩阵的特征值)

矩阵 A 的特征值为 $\lambda_1, \dots, \lambda_n$, 则 A^* 的特征值 $\lambda_j^* = \prod_{i \neq j} \lambda_i$



证明 考虑 $A = P^{-1}JP$, 由于 J 是上三角阵, 根据初等变换法, J^{-1} 是上三角阵, $J^* = |J|J^{-1}$ 也是上三角阵, 而 J 对角元素的代数余子式为 $\prod_{i \neq j} \lambda_i$, 因此 J^* 的对角元为 $\prod_{i \neq j} \lambda_i$, 代数余子式在初等变换下不变 (位置可能变化, 但值整体相同), 因此结论成立。

2.1.4 广义逆矩阵

定义 2.4 (广义逆)

$A_{m \times n}$ 是数域 $\mathbb{P}$ 上的矩阵, 若存在 $G_{n \times m}$ 使得

$$AGA = A$$

则称 G 是 A 的广义逆矩阵。



笔记 广义逆矩阵一般不唯一。

定理 2.9 (广义逆的表达式)

设 $A_{m \times n}$ 的秩为 r , 设 $A = P \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q$, 其中 $P_{m \times m}, Q_{n \times n}$ 为可逆矩阵, 则 A 全部的广义逆矩阵为

$$G = Q^{-1} \begin{pmatrix} E_r & C \\ D & F \end{pmatrix} P^{-1}$$

这里 C, D, F 为任意 $r \times (m-r), (n-r) \times r, (m-r) \times (n-r)$ 阶矩阵。



证明 (1) 验证 G 是广义逆: 直接乘起来就行, 得到 $AGA = A$

(2) 验证广义逆都是 G : 设 $G = Q^{-1} \begin{pmatrix} B & C \\ D & F \end{pmatrix} P^{-1}$, 乘法得到

$$AGA = P \begin{pmatrix} B & O \\ O & O \end{pmatrix} Q$$

因此得到 $B = E_r$

定理 2.10 (非方阵线性方程组解的表达式)

$A_{m \times n}$, G 是 A 的广义逆, 则 $AX = b$ 的通解可表示为

$$X = Gb + (E_n - GA)Y$$

这里 Y 是任一 n 维向量



证明 (1) 证明是解: 设 Z 是 $AZ = b$ 的解, 则此时有 $b = AZ = AGAZ = AGb$. 将 X 表达式代入 AX 验证:

$$AX = AGb + (A - AGA)Y = b + O = b$$

(2) 证明解都有该形式: 设 X_0 是 $AX = b$ 的解, 则

$$X_0 = Gb + X_0 - Gb = Gb + X_0 - GAX_0 = Gb + (E_n - GA)X_0$$

2.1.5 Moore-Penrose 广义逆

定义 2.5 (共轭转置)

A 是 $\mathbb{C}$ 上 $n \times m$ 矩阵, 则定义 $A^H = (\overline{A})^T$



定义 2.6 (Moore-Penrose 广义逆)

A 是 $\mathbb{C}$ 上 $m \times n$ 阶矩阵, 若存在 $\mathbb{C}$ 上 $n \times m$ 阶矩阵 G 满足:

1. $AGA = A$
2. $GAG = G$
3. $(AG)^H = AG$
4. $(GA)^H = GA$

则称 G 为 A 的 M-P 广义逆



定理 2.11 (M-P 广义逆的存在唯一性)

$A_{m \times n}$ 的 M-P 广义逆总是存在且唯一的, 且若 $\beta_1, \dots, \beta_r$ 是 A 列向量的极大线性无关组, 记 $B = [\beta_1, \dots, \beta_r]$, 分解 $A = BC$, 则 M-P 广义逆的具体表达式为:

$$A^+ = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$$



证明 存在性: 设 A 的秩为 r , 取 $\beta_1, \dots, \beta_r$ 为 A 列向量的极大线性无关组, 此时

$$A = (\alpha_1, \dots, \alpha_n) = (\beta_1, \dots, \beta_r) \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{r1} & \cdots & c_{rn} \end{pmatrix} = BC$$

此时 $r(BC) = r \leq r(C) \leq \min\{r, n\} \leq r$, 得到 $r(C) = r$ 行满秩, B 显然列满秩, 因此 B^HB, CC^H 是 r 阶可逆矩阵 (根据 $r(A^HA), r(AA^H)$ 的性质). 令

$$G = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$$

可验证 G 满足上述所有条件

唯一性：设 X 是另一个广义逆，则进行下面步骤，本质是用 $AXA = A$ 消去 X

$$\begin{aligned} X &= XAX = X(AX)^H = XX^H A^H = XX^H(AGA)^H \\ &= XX^H A^H(AG)^H = X(AX)^H(AG)^H = XAXAG \\ &= XAG = X(AGA)G = (XA)^H(GA)^H G \\ &= (GAXA)^H G = (GA)^H G = (GA)G \\ &= G. \end{aligned}$$

 **练习 2.3 计算 M-P 逆** (1) $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 0 \\ 0 & 2 & 1 \end{pmatrix}$ 的 M-P 逆

解 (1) 先进行分解，得到

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix}$$

最终结果得到

$$A^+ = \frac{1}{9} \begin{pmatrix} 3 & 2 & -1 \\ 0 & 2 & 2 \\ -3 & -1 & 2 \end{pmatrix}$$

定理 2.12 (最小二乘解的表达式)

$AX = b$ 最小二乘解为 $X = A^+b + (E_n - A^+A)Y$ ，其中 A^+b 是唯一的极小最小二乘解。



2.2 分块矩阵

2.2.1 常见矩阵拆分技巧

定理 2.13 (点态分解)

A 是 $n \times n$ 矩阵，则

$$A = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix} \Rightarrow A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}$$



定理 2.14 ($A^T A$ 的和式表达)

A 为 $n \times m$ 矩阵，将 A 做下面的行划分，则 $A^T A = \sum_{i=1}^n \alpha_i^T \alpha_i$

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{pmatrix}$$



2.2.2 分块初等变换

定义 2.7 (分块矩阵初等变换)

分块矩阵初等变换分为行变换和列变换:

- 行变换: (1) 左乘矩阵 P , 加到另一行 (2) 互换两行位置 (3) 将可逆矩阵左乘至一行
- 列变换: (1) 右乘矩阵 P , 加到另一列 (2) 互换两列位置 (3) 将可逆矩阵右乘至一列



定义 2.8 (分块初等矩阵)

对应三种分块初等变换有三种分块初等矩阵, 常用的即加至另一行/列的初等分块矩阵:

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



定理 2.15 (初等变换的分块初等矩阵表达)

分块初等矩阵左乘一个矩阵即做行变换, 右乘即做列变换:

$$\begin{bmatrix} E & 0 \\ P & E \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ PA_1 + A_3 & PA_2 + A_4 \end{bmatrix}$$



2.2.3 分块打洞技巧总结

定理 2.16 (分块打洞)

对于分块矩阵 $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ 可通过分块初等变换转换为:

$$\begin{bmatrix} A & B \\ O & D - CA^{-1}B \end{bmatrix} \quad \begin{bmatrix} A - BD^{-1}C & O \\ O & D \end{bmatrix}$$

记忆方式: 除去被减的位置, 减去的内容顺序都是顺时针转的。



推论 2.1 (两种特殊形式)

对称形式和归纳形式:

- 对称形式: $\begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \rightarrow \begin{bmatrix} A & O \\ O & D - B^T A^{-1} B \end{bmatrix}$
- 归纳形式: $\begin{bmatrix} A & \alpha \\ \beta^T & a_{nn} \end{bmatrix} \rightarrow \begin{bmatrix} A & \alpha \\ 0 & a_{nn} - \beta^T A^{-1} \alpha \end{bmatrix}$



练习 2.4 分块矩阵与数学归纳法 (1) n 阶方阵 A 的顺序主子式非零, 证明: 存在 n 阶下三角矩阵 B 使得 BA 是上三角阵

(2) A 是 n 阶实方阵, A 所有顺序主子式大于 0, 且非主对角元素为负, 证明: A^{-1} 的元素均为正数

证明 (1) $n = 1$ 显然成立。设 $n - 1$ 时成立, 考虑如下分块, 由于 A 顺序主子式非零, 因此 A_1 顺序主子式非零, 存在 $n - 1$ 阶下三角矩阵 B_1 满足 $B_1 A_1$ 为上三角阵, 因此做对应分块初等变换

$$A = \begin{bmatrix} A_1 & \alpha \\ \beta^T & a_{nn} \end{bmatrix} \Rightarrow \begin{bmatrix} B_1 & \\ & 1 \end{bmatrix} \cdot \begin{bmatrix} E_{n-1} & O \\ -\beta^T A_1^{-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} A_1 & \alpha \\ \beta^T & a_{nn} \end{bmatrix} = \begin{bmatrix} B_1 A_1 & B_1 \alpha \\ O & a_{nn} - \beta^T A_1^{-1} \alpha \end{bmatrix}$$

右侧即一个上三角矩阵, 把两个初等矩阵的乘积视作 B 即可。

(2) $n = 1$ 显然。设 $n - 1$ 时成立, 对 A 做下述分块, 则 A_1^{-1} 元素都是正的

$$A = \begin{bmatrix} A_1 & \alpha \\ \beta^T & a_{nn} \end{bmatrix} \Rightarrow \begin{bmatrix} E_{n-1} & O \\ -\beta^T A_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} A_1 & \alpha \\ \beta^T & a_{nn} \end{bmatrix} \begin{bmatrix} E_{n-1} & -A_1^{-1}\alpha \\ O & 1 \end{bmatrix} = \begin{bmatrix} A_1 & O \\ O & a_{nn} - \beta^T A_1^{-1}\alpha \end{bmatrix}$$

根据 $|A| > 0, |A_1| > 0$ 得到 $a_{nn} - \beta^T A_1^{-1}\alpha > 0$ 。记上述等式为 $BAC = D$, 则 $A^{-1} = CD^{-1}B$, 根据矩阵乘法乘开即可获得结论。

2.2.4 分块矩阵行列式

定理 2.17 (分块对角矩阵的行列式)

根据 Laplace 多行展开定理可以得到:

$$\begin{vmatrix} A & 0 \\ B & C \end{vmatrix} = |A||B|, \quad \begin{vmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix} = |A_{11}||A_{22}|\cdots|A_{nn}|$$

定理 2.18 (一般分块矩阵的行列式)

设 A, D 分别为 r, s 阶矩阵, A, D 可逆, 则

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D||A - BD^{-1}C| = |A||D - CA^{-1}B|$$

证明 (1) 如果 D 可逆, 则把 B 消掉 (左乘 BD^{-1}), 取行列式即可

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \Rightarrow \begin{bmatrix} A - BD^{-1}C & 0 \\ C & D \end{bmatrix}$$

定理 2.19 (行列式计算技巧 1: 配凑分块矩阵)

矩阵 A 可分解为 $A = J - u^T u$, 要计算 $|A|$, 则可以构造分块矩阵:

$$\begin{vmatrix} J & u \\ u^T & 1 \end{vmatrix} = |J| \cdot |J - u^T u| = |J| \cdot |1 - uJ^{-1}u^T| \Rightarrow |A| = |J| \cdot |1 - uJ^{-1}u^T|$$

注意这里 u 是列向量

2.2.5 分块矩阵求逆

一般分块矩阵求逆都先转换为分块对角矩阵, 具体如下。

定理 2.20 (上三角分块矩阵求逆)

当下面的 A_1, A_2 都是方阵, 则 A 可逆当且仅当 A_1, A_2 都可逆, 且逆矩阵如下:

$$A = \begin{bmatrix} A_1 & A_3 \\ O & A_2 \end{bmatrix} \Rightarrow A^{-1} = \begin{bmatrix} A_1^{-1} & -A_1^{-1}A_3A_2^{-1} \\ 0 & A_2^{-1} \end{bmatrix}$$

证明 根据分块矩阵初等变换可知

$$\begin{bmatrix} E & -A_2^{-1}A_3 \\ O & E \end{bmatrix} \begin{bmatrix} A_1 & A_3 \\ O & A_2 \end{bmatrix} = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix} \Rightarrow \begin{bmatrix} E & -A_2^{-1}A_3 \\ O & E \end{bmatrix}^{-1} \begin{bmatrix} A_1 & A_3 \\ O & A_2 \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{-1} & O \\ O & A_2^{-1} \end{bmatrix}$$

因此可以得到结论。

定理 2.21 (一般分块矩阵可逆条件)

分块矩阵 $H = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 可逆当且仅当 $A - BD^{-1}C$ 和 $D - CA^{-1}B$ 都可逆



练习 2.5 分块矩阵求逆 A, B 为 n, m 阶可逆方阵, 求 $\begin{bmatrix} O & A \\ B & O \end{bmatrix}$ 的逆矩阵

解 先做交换行的变换: $\begin{bmatrix} O & A \\ B & O \end{bmatrix} \begin{bmatrix} O & E \\ E & O \end{bmatrix} = \begin{bmatrix} A & \\ & B \end{bmatrix}$, 两侧求逆即可得到:

$$\begin{bmatrix} O & A \\ B & O \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & O \\ O & B^{-1} \end{bmatrix} \begin{bmatrix} O & E \\ E & O \end{bmatrix} = \begin{bmatrix} O & B^{-1} \\ A^{-1} & O \end{bmatrix}$$

2.3 矩阵的秩

2.3.1 秩的概念

定义 2.9 (矩阵的秩)

矩阵的秩等于:

- 矩阵行列向量的秩
- 最大非零子式的阶数



2.3.2 初等变换标准型与矩阵分解

定理 2.22 (初等变换标准型)

设 A 是秩为 r 的 $s \times n$ 矩阵, 则存在 s 阶可逆矩阵 P 和 n 阶可逆矩阵 Q 使得

$$A = P \begin{bmatrix} E_r & O \\ O & O \end{bmatrix} Q$$



笔记 如果一个题目中没给出条件, 或者只给出了矩阵的秩, 则往往需要用到初等变换标准型

定理 2.23 (行列满秩的标准型)

A 是 $m \times n$ 矩阵, A 列满秩当且仅当存在 m 阶可逆矩阵 P 使得 $A = P \begin{bmatrix} E_n \\ O \end{bmatrix}$ 。 A 行满秩当且仅当存在 n 阶可逆矩阵 Q , 使得 $A = \begin{bmatrix} E_m & O \end{bmatrix} Q$



证明 由于列满秩, 可以用行初等变换消成阶梯型, 再从下到上消回去即可 (若列不满秩, 则阶梯型可能消不干净, 因此需要列变换)

2.3.3 利用初等变换标准型进行矩阵分解

定理 2.24 (矩阵分解)

$s \times n$ 矩阵 $\begin{bmatrix} E_r & O \\ O & O \end{bmatrix}$ 可分解为

$$\begin{bmatrix} E_r & O \\ O & O \end{bmatrix} = \begin{bmatrix} E_r \\ O \end{bmatrix} \begin{bmatrix} E_r & O \end{bmatrix}$$

若为方阵时, 可以分解为:

$$\begin{bmatrix} E_r & O \\ O & O \end{bmatrix} = \begin{bmatrix} E_r & O \\ O & O \end{bmatrix} \begin{bmatrix} E_r & O \\ O & O \end{bmatrix}$$



练习 2.6 行列满秩分解 (1) 证明: 任意非零矩阵都可以分解为列满秩和行满秩矩阵的乘积

(2) 设 B_1, B_2 为数域 P 上 $s \times n$ 列满秩矩阵, 证明: 存在 P 上 s 阶可逆矩阵 C , 使得 $B_2 = CB_1$ (即 B_1, B_2 只需要行变换就可以相互转换)

(3) C_1, C_2 为 P 上 $s \times n$ 行满秩矩阵, 证明: C_1, C_2 可以仅通过列变换相互转换。

证明 (1) 使用等价标准型:

$$A = P \begin{bmatrix} E_r & O \\ O & O \end{bmatrix} Q = P \begin{bmatrix} E_r \\ O \end{bmatrix} \begin{bmatrix} E_r & O \end{bmatrix} Q$$

(2) 根据列满秩初等标准型, $\exists P$ 使得 $PB_1 = \begin{bmatrix} E_n \\ O \end{bmatrix}$, 同理存在 Q 使得 $QB_2 = \begin{bmatrix} E_n \\ O \end{bmatrix}$, 结论成立。

(3) 和 (2) 一样

练习 2.7 幂等矩阵分解 (1) 证明: 任意方阵可以分解为可逆矩阵与幂等矩阵乘积

证明 (1) 利用等价标准型得到:

$$A = P \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q = PQQ^{-1} \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q$$

PQ 为可逆矩阵, 后面为幂等

2.3.4 秩不等式

引理 2.1 (乘可逆矩阵不改变秩)

P, Q 为可逆矩阵, 则 $r(A) = r(PAQ)$



证明 相当于做了一系列初等变换, 因此不改变秩

练习 2.8 非方阵情况 A, B 分别为 $m \times n, n \times s$ 矩阵, 证明:

(1) 若 $r(A) = n$, 则 $r(AB) = r(B)$

(2) 若 $r(B) = n$, 则 $r(AB) = r(A)$

证明 (1) 存在可逆矩阵 P 使得 $PA = \begin{bmatrix} A_1 \\ O \end{bmatrix}$, 这里 A_1 为可逆矩阵, 因此

$$r(AB) = r\left(\begin{bmatrix} A_1 \\ O \end{bmatrix} B\right) = r\left(\begin{pmatrix} A_1 B \\ O \end{pmatrix}\right) = r(A_1 B) = r(B)$$

(2) 同理 (转置即可)

引理 2.2 (分块矩阵的秩)

- $r \begin{pmatrix} A & O \\ O & B \end{pmatrix} = r(A) + r(B)$
- $r \begin{pmatrix} A & O \\ C & B \end{pmatrix} \geq r(A) + r(B)$



证明 (1) 根据初等变换等价标准型有 $P_1 A Q_1 = \begin{pmatrix} E_{r_1} & O \\ O & O \end{pmatrix}$, $P_2 B Q_2 = \begin{pmatrix} E_{r_2} & O \\ O & O \end{pmatrix}$, 因此

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix} \begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} = \begin{pmatrix} E_{r_1} & O & O & O \\ O & O & O & O \\ O & O & E_{r_2} & O \\ O & O & O & O \end{pmatrix}$$

(2) 根据初等变换等价标准型有 $P_1 A Q_1 = \begin{pmatrix} E_{r_1} & O \\ O & O \end{pmatrix}$, $P_2 B Q_2 = \begin{pmatrix} E_{r_2} & O \\ O & O \end{pmatrix}$, 因此

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix} \begin{pmatrix} A & O \\ C & B \end{pmatrix} \begin{pmatrix} Q_1 & O \\ O & Q_2 \end{pmatrix} = \begin{pmatrix} E_{r_1} & O & O & O \\ O & O & O & O \\ C_{11} & C_{12} & E_{r_2} & O \\ C_{21} & C_{22} & O & O \end{pmatrix} \rightarrow \begin{pmatrix} E_{r_1} & O & O & O \\ O & O & O & O \\ O & O & E_{r_2} & O \\ O & C_{22} & O & O \end{pmatrix}$$

定理 2.25 (矩阵秩基本不等式)

- $r(AB) \leq \min\{r(A), r(B)\}$
- $r(A+B) \leq r(A) + r(B)$
- Sylvester 不等式: A, B 是 n 阶方阵, 则 $r(A) + r(B) \leq r(AB) + n$
- Frobenius 不等式: A, B, C 是 n 阶方阵, 则 $r(ABC) \geq r(AB) + r(BC) - r(B)$
- A, B 为 $s \times n, n \times m$ 阶矩阵, $AB = O$, 则 $r(A) + r(B) \leq n$
- A 是 $m \times n$ 的实矩阵, 则 $r(A^T A) = r(A)$



证明 (1) 将 B 写为行向量形式: $B = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$, 则

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} a_{11}\beta_1 + \cdots + a_{1n}\beta_n \\ \vdots \\ a_{n1}\beta_1 + \cdots + a_{nn}\beta_n \end{pmatrix}$$

因此 AB 每个行向量都可被 B 行向量线性表出, 从而 $r(AB) \leq r(B)$, 同理可知 $r(AB) \leq r(A)$

(2) 用列向量角度, $A+B$ 的列向量显然可以被 A, B 的列向量线性表出, 从而结论成立。

(3) 是 (4) 的推论

(4) Frobenius: 即构造 $r(ABC) + r(B) \geq r(AB) + r(BC)$

$$\begin{bmatrix} ABC & \\ & B \end{bmatrix} \rightarrow \begin{bmatrix} ABC & O \\ BC & B \end{bmatrix} \rightarrow \begin{bmatrix} O & -AB \\ BC & B \end{bmatrix}$$

根据分块矩阵秩不等式可知

练习 2.9 秩不等式的推广 (1) A, B 为 n 阶方阵, 则 $r(A - ABA) = r(A) + r(E - BA) - n$

(2) A, B 是 n 阶方阵, $AB = BA = O, r(A^2) = r(A)$, 证明: $r(A+B) = r(A) + r(B)$

证明 (1) 等号需要构造单位矩阵, 即等式中的 n 用 E_n 代替:

$$\begin{bmatrix} A & \\ & E_n - BA \end{bmatrix} \rightarrow \begin{bmatrix} A & O \\ BA & E_n - BA \end{bmatrix} \rightarrow \begin{bmatrix} A & A \\ BA & E_n \end{bmatrix} \rightarrow \begin{bmatrix} A - ABA & O \\ O & E_n \end{bmatrix}$$

从右往左推非常难, 因此一般从无 E_n 的一边推另一边 (后面的练习也会体现这一点)

(2) 注意从复杂的一侧 $A+B$ 开始, 否则 A, B 不知道怎么搞,

$$\begin{pmatrix} A+B & O \\ O & O \end{pmatrix} \rightarrow \begin{pmatrix} A+B & O \\ A^2+AB & O \end{pmatrix} = \begin{pmatrix} A+B & O \\ A^2 & O \end{pmatrix} \rightarrow \begin{pmatrix} A+B & A^2 \\ A^2 & A^3 \end{pmatrix}$$

由于 $r(A) = r(A^2)$, 因此矩阵方程 $A^2X = A$ 有解, 故 $\exists P$ 使得 $A^2P = A$, 故第二列可右乘 P 消去第一列:

$$\begin{pmatrix} A+B & A^2 \\ A^2 & A^3 \end{pmatrix} \rightarrow \begin{pmatrix} B & A^2 \\ O & A^3 \end{pmatrix}$$

因此根据分块矩阵的秩得到 $r(A+B) \geq r(A^3) + r(B) = r(A) + r(B)$, 而不等式另一侧显然.

练习 2.10 用秩不等式研究矩阵性质 (1) 证明零矩阵: $A_{m \times n}, B_{n \times s}, r(B) = n$, 证明: 若 $AB = O$, 则 $A = O$

(2) 判定可逆: n 为奇数, A, B 为 n 阶矩阵, $A^2 = O$, 证明: $AB - BA$ 不可逆

证明 (1) 由于 $r(A) + r(B) \leq r(AB) + n$, 而 $r(B) = n$, 得到 $r(A) = 0$, 即 $A = O$

(2) $n = 2k + 1$, 根据 $A^2 = O$ 得到 $2r(A) \leq 2k + 1 + 0$, 从而 $r(A) \leq k$, 此时

$$r(AB - BA) \leq r(AB) + r(BA) \leq r(A) + r(A) \leq 2k$$

因此不可逆

2.3.5 空间分解法研究矩阵秩

定理 2.26 (多项式矩阵的秩关系)

$f(x) = f_1(x)f_2(x)$ 是多项式, $(f_1(x), f_2(x)) = 1$, 则 $f(A) = O$ 当且仅当 $r(f_1(A)) + r(f_2(A)) = n$. 常见的推论如下:

- $A^2 = A$ 当且仅当 $r(A - E) + r(A) = n$
- $A^2 = E$ 当且仅当 $r(A - E) + r(A + E) = n$



证明 根据 $(f_1(x), f_2(x)) = 1$, 有 $u(x)f_1(x) + v(x)f_2(x) = 1$, 因此 $u(A)f_1(A) + v(A)f_2(A) = E_n$, 根据分块矩阵初等变换:

$$\begin{bmatrix} f_1(A) & \\ & f_2(A) \end{bmatrix} \rightarrow \begin{bmatrix} f_1(A) & f_1(A)u(A) + v(A)f_2(A) \\ O & f_2(A) \end{bmatrix} = \begin{bmatrix} f_1(A) & E_n \\ O & f_2(A) \end{bmatrix}$$

再做初等变换:

$$\begin{bmatrix} f_1(A) & E_n \\ O & f_2(A) \end{bmatrix} \rightarrow \begin{bmatrix} O & E_n \\ -f_2(A)f_1(A) & f_2(A) \end{bmatrix} \rightarrow \begin{bmatrix} O & E_n \\ -f(A) & O \end{bmatrix}$$

因此得到 $r(f_1(A)) + r(f_2(A)) = n + r(f(A))$

笔记 分块矩阵、多项式互素、空间直和本质上是一体的。

练习 2.11 从线性空间角度解决难题 A, B 是 P 上 n 阶矩阵, $AB = BA$, 证明: $r(A) + r(B) \geq r(AB) + r(A+B)$

证明 设 $AX = 0, BX = 0$ 的解空间分别为 V_1, V_2 , 设 $ABX = BAX = 0$ 的解空间为 W_1 , $(A+B)X = 0$ 的解空间为 W_2 . 显然 $V_1 \subseteq W_1, V_2 \subseteq W_1$, 从而 $V_1 + V_2 \subseteq W_1$, 显然 $V_1 \cap V_2 \subseteq W_2$, 根据维数公式:

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2) \leq \dim W_1 + \dim W_2$$

因此推出：

$$(n - r(A)) + (n - r(B)) \leq (n - r(AB)) + (n - r(A + B))$$

化简得到结论。

 **笔记** 这种 $r(A), r(B), r(A + B), r(AB)$ 相关问题绝大多数从线性空间角度都能迎刃而解

2.4 矩阵的迹

定理 2.27 (矩阵迹的性质)

A, B 为 n 阶方阵，则

- $\text{tr}(A) = \text{tr}(kA)$
- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
- $\text{tr}(AB) = \text{tr}(BA)$
- 若 A 为实方阵，则 $A = O$ 当且仅当 $\text{tr}(A^T A) = 0$



证明 显然

2.5 几种技巧矩阵

2.5.1 初等矩阵

定义 2.10 (初等矩阵)

单位矩阵通过一次初等变换得到的矩阵称为初等矩阵。记作 $P(i, j), P(i(c)), P(i, j(k))$ 表示 i, j 行互换, i 行乘 c , $P_{ij} = k$ 的矩阵



定理 2.28 (初等矩阵乘法)

AP 相当于做一次列初等变换, PA 相当于做一次行初等变换,



2.5.2 基本矩阵：仅一个元素为 1 的矩阵

定义 2.11 (基本矩阵)

基本矩阵 E_{ij} 表示 $n \times n$ 矩阵，但是仅有 (i, j) 元素为 1，其余元素为 0 的矩阵。



定理 2.29 (基本矩阵乘法)

基本矩阵乘法有以下性质：

- AE_{ij} 将 A 的第 i 列移动到第 j 列，
- $E_{ij}A$ 将 A 的第 j 行移动到第 i 行。
- 基础矩阵分解： $E_{ij} = e_i e_j^T$
- 基本矩阵相乘： $E_{ik} E_{kj} = E_{ij}$, $k \neq l$ 时 $E_{ik} E_{lj} = 0$



 **练习 2.12 用基础矩阵研究幂次** (1) 设 $A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$, 证明: $A^k = \begin{pmatrix} O & E_{n-k} \\ E_k & O \end{pmatrix}$.

(2) A 是 n 阶上三角阵且主对角上元素均为 0, 证明 $A^n = O$

证明 (1) 写为 $A = (e_n, e_1, \cdots, e_{n-1})$, 则 $A^2 = (Ae_n, Ae_1, \cdots, Ae_{n-1})$. 显然 Ae_i 为 A 的第 i 列, 因此

$$A^2 = (e_{n-1}, e_n, e_1, \cdots, e_{n-2})$$

以此类推容易得到结论

(2) $A = \sum_{i < j} a_{ij} E_{ij}$, 而 $j \neq k$ 时 $E_{ij} E_{kl} = O$, 因此 A^n 展开式中非零项只能是 $E_{i_1 j_1} E_{j_1 j_2} \cdots E_{j_{n-1} j_n}$ 这样的项, 但满足 $1 \leq i < j_1 < \cdots < j_n \leq n$ 的项不存在, 因此 $A^n = O$

 **练习 2.13 用二次型研究矩阵性质** (1) 证明: n 阶对称矩阵 A 为零矩阵当且仅当 $\forall \alpha \in \mathbb{R}^n$ 有 $\alpha^T A \alpha = 0$

(2) 证明: n 阶方阵 A 是反对称矩阵当且仅当 $\forall \alpha \in \mathbb{R}^n$ 有 $\alpha^T A \alpha = 0$

证明 (1) 左推右显然. 先设 $\alpha = e_i$, 则 $e_i^T A e_i = a_{ii} = 0$, 再取 $\alpha = e_i + e_j$, 则

$$0 = (e_i + e_j)^T A (e_i + e_j) = a_{ii} + a_{jj} + a_{ij} + a_{ji}$$

而 $a_{ij} = a_{ji}$, 因此得到 $a_{ij} = 0, \forall i, j$

(2) 左推右用 $\alpha^T A \alpha = (\alpha^T A \alpha)^T = -\alpha^T A \alpha$ 即可. 右推左: 根据条件得到 $\alpha^T A^T \alpha = 0$, 因此 $\alpha^T (A + A^T) \alpha = 0$, 由于 $A + A^T$ 为对称矩阵, 根据上题结论得到 $A + A^T = O$, 因此 $A^T = -A$

2.5.3 置换矩阵

定义 2.12 (置换矩阵)

每行只有一个元素为 1, 每列也只有一个元素为 1, 则称该矩阵为置换矩阵. 一般写为 $P = [\epsilon_{i_1}, \cdots, \epsilon_{i_n}] = [\epsilon_{j_1}^T, \cdots, \epsilon_{j_n}^T]^T$, 其中 ϵ_i 表示仅第 i 元素为 1 的列向量, i_k, j_k 分别表示列、行的重排顺序.

定理 2.30 (置换矩阵是正交矩阵)

置换矩阵是正交矩阵。

证明 证明 $P^T P = I$

定理 2.31 (置换矩阵的作用)

左乘置换矩阵相当于把行原本的 $j_1, j_2, \cdots, j_n$ 行换到 $1, 2, \cdots, n$ 行, 右乘相当于把列原本的 $i_1, \cdots, i_n$ 列换到 $1, 2, \cdots, n$ 列。

定理 2.32 (置换矩阵正交相似)

$i_1 \cdots i_n$ 是 $1 \sim n$ 的排列, $A = (a_{ij})$, 取 $P = [\epsilon_{i_1}, \cdots, \epsilon_{i_n}]$, 则:

$$P^T A P = P^{-1} A P = \begin{bmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_n} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_n} \\ \vdots & \vdots & & \vdots \\ a_{i_n i_1} & a_{i_n i_2} & \cdots & a_{i_n i_n} \end{bmatrix}$$

相当于将行列均按新顺序排列了一遍 (先把行放到对应位置, 再把列放到对应位置), 且是一种正交变换。

 **练习 2.14 置换矩阵正交相似例子** $i_1 \cdots i_n$ 是 $1, 2, \cdots, n$ 的排列, $A = \text{diag}\{\lambda_1, \cdots, \lambda_n\}$, $B = \text{diag}\{\lambda_{i_1}, \cdots, \lambda_{i_n}\}$, 证明 A, B 是正交相似的

证明 取 $P = [\epsilon_{i_1}, \cdots, \epsilon_{i_n}]$ 即可。

定义 2.13 (特殊的置换矩阵)

特殊置换矩阵 $K = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \end{bmatrix}$, $K^T A K$ 的功能是绕中心转 180 度 (行列全部倒排)



2.6 特殊矩阵性质总结

2.6.1 对角矩阵

定理 2.33 (对角矩阵乘法)

矩阵 A 右乘一个对角阵, 则每列乘上对角相应元素。 A 左乘对角阵, 则每行乘上对角相应元素。



定理 2.34 (对角矩阵求逆)

$\text{diag}\{d_1, \cdots, d_n\}^{-1} = \text{diag}\{d_1^{-1}, \cdots, d_n^{-1}\}$



2.6.2 上、下三角矩阵

定理 2.35 (上下三角乘法、逆、伴随)

A, B 是上三角矩阵, 则

- AB 仍然是上三角矩阵, 且 AB 对角元为 $a_{11}b_{11}, \cdots, a_{nn}b_{nn}$
- A^{-1} 也是上三角阵, 且 A^{-1} 对角元为 $a_{11}^{-1}, \cdots, a_{nn}^{-1}$
- A^* 也是上三角阵, 且 A^* 的对角元为 $a_{22}a_{33} \cdots a_{nn}, a_{11}a_{33} \cdots a_{nn}, \cdots, a_{11}a_{22} \cdots a_{n-1,n-1}$ (也是特征值)



证明 (1) 令 $C = AB$, 若 $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = \sum_{k=1}^{i-1} a_{ik}b_{kj} + \sum_{k=i}^n a_{ik}b_{kj}$, 当 $i > j$ 时, 显然 $c_{ij} = 0$ 。而 $c_{ii} =$

$$\sum_{k=1}^{i-1} a_{ik}b_{ki} + a_{ii}b_{ii} + \sum_{k=i+1}^n a_{ik}b_{ki} = a_{ii}b_{ii}$$

(2) 根据行变换可以将 $(A, I) \rightarrow (I, A^{-1})$, 当 A 为上三角时, 先用最后一行消去 A 最后一列其余行的元素, 以此类推可以发现 A^{-1} 为上三角, 根据矩阵乘法显然有 A^{-1} 对角为 A 对角的逆。

(3) 根据 $A^* = |A|A^{-1}$ 即可

定理 2.36 (n 个上三角矩阵的乘积)

$A_1, \cdots, A_n$ 是 n 个对角为 0 的 n 阶上三角矩阵, 证明 $A_1 A_2 \cdots A_n = O$



证明 用数学归纳法, 设 $n-1$ 时成立, 取 $A_n = \begin{bmatrix} B_{n-1} & * \\ 0 & 0 \end{bmatrix}$, 此时

$$A_1 \cdots A_n = \begin{bmatrix} B_1 \cdots B_{n-1} & * \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} B_n & * \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & * \\ 0 & 0 \end{bmatrix} \begin{bmatrix} B_n & * \\ 0 & 0 \end{bmatrix} = O$$

2.6.3 实对称矩阵

定义 2.14 (实对称矩阵)

每个元素都是实数, 且满足 $A^T = A$ 的矩阵 A 称为实对称矩阵

定理 2.37 (实对称矩阵的矩阵性质)

实对称矩阵 A 具有以下性质:

- A^{-1} 也是实对称矩阵

证明 (1) 由于 $(AA^{-1})^T = E_n$, 得到 $(A^{-1})^T A^T = E_n$, 而 A 实对称, 因此 $(A^{-1})^T A = E_n$, 这说明 $(A^{-1})^T = A^{-1}$

定理 2.38 (实对称矩阵的特征值性质)

实对称矩阵有如下性质:

- 每个特征值都是实数
- 不同特征值的特征向量正交
- 一定正交相似 (相似且合同) 于对角矩阵

证明 (1) 共轭转置法: 设 $\lambda \in \mathbb{C}$ 是特征值, $\xi \in \mathbb{C}^n$ 是特征向量, 一方面根据 $A\xi = \lambda\xi$ 两侧同时左乘 $\bar{\xi}^T$ 得到

$$\bar{\xi}^T A\xi = \lambda \bar{\xi}^T \xi$$

另一方面直接对 $A\xi = \lambda\xi$ 取共轭转置, 再同时右乘 ξ 得到

$$\bar{\xi}^T A = \bar{\xi}^T \bar{\lambda} \Rightarrow \bar{\xi}^T A\xi = \bar{\lambda} \bar{\xi}^T \xi$$

因此比较得到 $\lambda = \bar{\lambda}$

(2) 不同特征值的特征向量正交: 设 λ_1, λ_2 是两个互异特征值, ξ_1, ξ_2 分别属于 λ_1, λ_2 的实特征向量, 则

$$\lambda_1(\xi_1, \xi_2) = (\lambda_1 \xi_1)^T \xi_2 = (A\xi_1)^T \xi_2 = \xi_1^T (A\xi_2) = \lambda_2(\xi_1, \xi_2)$$

从而得到 $(\lambda_2 - \lambda_1)(\xi_1, \xi_2) = 0$, 而 $\lambda_1 \neq \lambda_2$, 从而 $(\xi_1, \xi_2) = 0$

(3) 做数学归纳法, $n=1$ 显然, 假设 $n-1$ 成立, 任取特征值 λ , 对应单位特征向量 α_1 , 将 α_1 扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\alpha_1, \dots, \alpha_n$, 令 $T = (\alpha_1, \dots, \alpha_n)$, 则

$$AT = T \begin{pmatrix} \lambda & \alpha^T \\ 0 & A_{n-1} \end{pmatrix} \Rightarrow T^T AT = \begin{pmatrix} \lambda & \alpha^T \\ 0 & A_{n-1} \end{pmatrix}$$

由于 $T^T AT$ 对称, 因此 $\alpha^T = 0$, 再根据归纳假设, 存在正交矩阵 P_1 使得 $P_1^T A_{n-1} P_1$ 为对角阵, 因此令 $P = \text{diag}(1, P_1)$, 此时

$$P^T T^T AT P = \begin{pmatrix} \lambda & 0 \\ 0 & P^T A_{n-1} P \end{pmatrix}$$

显然 $TP = Q$, 则 Q 也是正交阵。

2.6.4 反对称矩阵

定理 2.39 (反对称矩阵的行列式)

奇数阶反对称矩阵的行列式必为 0。

证明 根据 $A^T = -A$ 得到 $|A| = (-1)^n |A|$, 当 n 为奇数时 $|A| = -|A|$, 因此 $|A| = 0$

定理 2.40 (反对称矩阵的矩阵性质)

A 为反对称矩阵, 则

- A^{-1} 为反对称矩阵



证明 (1) 类似实对称矩阵的

定理 2.41 (反对称矩阵的二次型)

A 为反对称矩阵当且仅当任意 $X \in \mathbb{R}^n$, 有 $X^T A X = 0$



证明 在基础矩阵技巧一节中给出过证明

定理 2.42 (反对称矩阵的特征值、特征向量)

反对称矩阵 A 的特征值是 0 或者纯虚数。且纯虚数对应特征向量的实部和虚部的实向量等长、正交。



证明 (1) 设 A 的特征值 $\lambda = a + bi$, 对应特征向量 $x = u + vi$, $A(u + vi) = (a + bi)(u + vi)$, 因此 $Au + iAv = (au - bv) + (bu + av)i$, 实虚部分离得到:

$$Au = au - bv, Av = bu + av \Rightarrow u^T Au = au^T u - bu^T v, v^T Av = bv^T u + av^T v$$

因此得到 $u^T Au + v^T Av = a(|u|^2 + |v|^2)$, 由于 $u^T Au \in \mathbb{C}$ 是一个数, 因此

$$u^T Au = (u^T Au)^T = u^T A^T u = -u^T Au \Rightarrow u^T Au = 0$$

同理有 $v^T Av = 0$, 因此 $a(|u|^2 + |v|^2) = 0$, 由于 $u + vi \neq 0$, 得到 $a = 0$, 即 $\lambda = a + bi = bi$

(2) 设 $\alpha + \beta i$ 是特征值 ki 的特征向量, 即 $A(\alpha + i\beta) = ki(\alpha + i\beta)$, 对比实部虚部, 得到

$$A\alpha = -k\beta, A\beta = k\alpha$$

由于 $\alpha^T A\alpha = 0$, 得到 $\alpha^T A\alpha = \alpha^T (-k\beta) = -k\alpha^T \beta = 0$, 这说明 α, β 正交, 再由于

$$\alpha^T A\beta + \beta^T A\alpha = \begin{cases} \alpha^T A\beta - (\alpha^T A\beta)^T = 0 \\ k(\alpha^T \alpha - \beta^T \beta) \end{cases}$$

得到 $\alpha^T \alpha - \beta^T \beta = 0$, 即模长相等

定理 2.43 (反对称矩阵的标准型)

A 为 n 阶反对称矩阵, 则

- A 正交相似于如下准对角阵

$$\text{diag} \left\{ \begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, 0, \dots, 0 \right\}$$

- A 合同于标准型

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

**定理 2.44 (将矩阵拆为对称 + 反对称)**

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$



2.6.5 Hermite 矩阵

定义 2.15 (Hermite 矩阵)

$\overline{A}^T = A$ 的矩阵称为 Hermite 矩阵



定理 2.45 (Hermite 矩阵的性质)

Hermite 矩阵的特征值都是实数



2.6.6 正交矩阵

定义 2.16 (正交矩阵)

若矩阵 T 满足 $TT^T = I$, 则称 T 为正交矩阵。



定理 2.46 (正交矩阵的性质)

A, B 为 n 阶正交矩阵, 则

- 求逆: $A^{-1} = A^T$
- A^{-1}, A^T, AB 都是正交矩阵
- $S = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ 也是正交矩阵
- A 正交当且仅当 A 的列向量是 $\mathbb{R}^n$ 的一组标准正交基。



证明 (1) 显然

(2) $(AB)^T AB = I$

(3) 直接用 $S^T S$ 验证即可

(4) $A = [\alpha_1, \dots, \alpha_n]$, 计算 $A^T A = I$ 可得出结论。

练习 2.15 正交矩阵练习 b 是单位列向量, 证明存在 n 阶对称正交矩阵 A , 使得 b 为 A 的第一列

证明

定理 2.47 (分块上三角正交矩阵)

若 $A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix}$ 是正交矩阵, 则 $A_2 = 0$, A_1, A_4 是正交矩阵



证明 根据分块矩阵转置以及乘法可以得到结论

$$A^T A = \begin{bmatrix} A_1^T & 0 \\ A_2^T & A_4^T \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix}$$

定理 2.48 (正交矩阵的行列式与特征值)

A 是 n 阶正交矩阵, 则 A 的特征值模长为 1, 复特征值成对出现, $|A| = \pm 1$, 进一步地

- 若 $|A| = -1$: -1 一定是 A 的特征值
- 若 A 是奇数阶的, 且 $|A| = 1$, 则 1 一定是 A 的特征值
- 复特征值的特征向量实部和虚部模长相等且正交



证明 (1) 主命题: $A\alpha = \lambda\alpha$, 两侧取共轭得到 $A\bar{\alpha} = \bar{\lambda}\bar{\alpha}$, 两侧取转置得到 $\bar{\alpha}^T A^T = \bar{\lambda}\bar{\alpha}^T$, 两侧右乘 α 得到 $\bar{\alpha}^T A^{-1}\alpha = \bar{\lambda}\bar{\alpha}^T\alpha$, 因此 $\frac{1}{\lambda}\bar{\alpha}^T\alpha = \bar{\lambda}\bar{\alpha}^T\alpha$, 因此 $\lambda\bar{\lambda} = 1$, 即模长为 1。根据 $|A^T A| = 1$ 可知行列式为 ± 1

(2) $|A| = -1$ 情况: 根据 $|A|$ 是特征值乘积可知至少有 1 个 -1

(3) $|A| = 1$ 情况: 同理

 **笔记** 在计算 A 的特征值时, 如果 A 的特殊性需要 A^T 来体现, 则经常两侧取共轭转置 (例如实对称、反对称、正交矩阵)。

定理 2.49 (n 阶正交矩阵的标准型)

A 是 n 阶正交矩阵, 则 A 总是正交相似于

$$\text{diag} \left\{ \lambda_1, \dots, \lambda_r, \begin{pmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_m & -\sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{pmatrix} \right\}$$



2.6.7 Unitary 矩阵

定理 2.50 (Unitary 矩阵的性质)

Unitary 矩阵的特征值模长均为 1



2.6.8 幂零矩阵

定义 2.17 (幂零矩阵)

若存在整数 l , 使得 $A^l = O$, 则称 A 为幂零矩阵。



定理 2.51 (幂零矩阵充要条件)

A 是幂零矩阵当且仅当

1. 特征值: A 的特征值只有 0
2. 特征多项式: A 的特征多项式为 x^n
3. 最小多项式: A 的最小多项式为 $x^k, k \leq n$
4. 迹 (重点): 对 $\forall k > 0, \text{tr}(A^k) = 0$



2.6.9 幂等矩阵

定义 2.18 (幂等矩阵)

若 n 阶矩阵 A 满足 $A^2 = A$, 则称 A 为幂等矩阵



定理 2.52 (幂等矩阵充要条件)

n 阶方阵 A 是幂等矩阵当且仅当 $r(A) + r(A - E) = n$



证明 由于 $A^2 = A$, 即 $A(A - E) = O$, 根据空间分解研究矩阵秩定理 (具体见线性空间-直和与空间分解)

$$n = n - r(A) + n - r(A - E)$$

因此 $r(A) + r(A - E) = n$

定理 2.53 (幂等矩阵的性质)

A 为 n 阶幂等矩阵, 则

- A 的特征值只有 1 和 0
- A 一定可以对角化, 且标准型为 $\text{diag}\{E_r, O\}$
- $r(A) = \text{tr}(A)$



证明 (1) 设 λ 是 A 的特征值, 故 $A\alpha = \lambda\alpha$, $A^2\alpha = \lambda^2\alpha$, 由于 $A = A^2$, 得到 $\lambda^2 = \lambda$, 因此 $\lambda = 0, 1$
 (2) 根据最小多项式理论, $f(A) = A(A - E) = O$, 因此最小多项式只有一次项乘积, 故可对角化。
 (3) 根据秩、迹在相似下的不变性可知

定理 2.54 (加减乘法的幂等性)

A, B 为两个 n 阶幂等矩阵, 则

1. $A + B$ 为幂等矩阵当且仅当 $AB = BA = O$
2. $A - B$ 为幂等矩阵当且仅当 $AB = BA = B$
3. 若 $AB = BA$, 则 AB 为幂等矩阵。反之不一定
4. 如 $E - A - B$ 可逆, 则 $r(A) = r(B)$



证明 (1) 右推左: 由于 $AB = BA = O$, 因此 $(A + B)^2 = A^2 + AB + BA + B^2 = A + B$ 。左推右: $(A + B)^2 = A + AB + BA + B = A + B$, 因此 $AB + BA = O$, 得到

$$AB = AB^2 = (-BA)B = -BAB = B(BA) = BA$$

练习 2.16 幂等矩阵分解 (1) 证明: 任意方阵都可以分解为可逆矩阵与幂等矩阵的乘积

证明 (1) 利用等价标准型得到:

$$A = P \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q = PQQ^{-1} \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q$$

PQ 为可逆矩阵, 后面为幂等

定理 2.55 (幂等矩阵列的性质)

$A_1, \dots, A_m$ 为 n 阶方阵, 且 $A_1 + A_2 + \dots + A_m = E_n$, 则下面三个命题等价:

- $A_1, \dots, A_m$ 为幂等矩阵
- $r(A_1) + \dots + r(A_m) = n$
- $\forall i \neq j$ 有 $A_i A_j = O$



证明 (1) $\Rightarrow$ (2): 由于 $r(A_i) = \text{tr}(A_i)$, 因此

$$r(A_1) + \dots + r(A_m) = \text{tr}(A_1) + \dots + \text{tr}(A_m) = \text{tr}(A_1 + \dots + A_m) = \text{tr}(E_n) = n$$

(2) $\Leftarrow$ (3): 比较经典, 先通过分块矩阵的初等变换凑出 E :

$$\begin{pmatrix} A_1 & & & & \\ & A_2 & & & \\ & & A_3 & & \\ & & & \ddots & \\ & & & & A_m \end{pmatrix} \begin{pmatrix} A_1 & A_2 & A_3 & \cdots & A_m \\ A_2 & & & & \\ & A_3 & & & \\ & & \ddots & & \\ & & & A_m & \end{pmatrix} \rightarrow \begin{pmatrix} E & A_2 & A_3 & \cdots & A_m \\ A_2 & A_2 & & & \\ A_3 & & A_3 & & \\ \vdots & & & \ddots & \\ A_m & & & & A_m \end{pmatrix}$$

再通过 E 消去其他部分：

$$\rightarrow \begin{pmatrix} E & O & O & \cdots & O \\ A_2 & A_2 - A_2^2 & -A_2 A_3 & \cdots & -A_2 A_m \\ A_3 & -A_3 A_2 & A_3 - A_3^2 & \cdots & -A_3 A_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_m & -A_m A_2 & -A_m A_3 & \cdots & A_m - A_m^2 \end{pmatrix} \rightarrow \begin{pmatrix} E & O & O & \cdots & O \\ O & A_2 - A_2^2 & -A_2 A_3 & \cdots & -A_2 A_m \\ O & -A_3 A_2 & A_3 - A_3^2 & \cdots & -A_3 A_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O & -A_m A_2 & -A_m A_3 & \cdots & A_m - A_m^2 \end{pmatrix}$$

根据 $r(A_1) + r(A_2) + \cdots + r(A_m) = n$ ，得到

$$\begin{pmatrix} A_2 - A_2^2 & -A_2 A_3 & \cdots & -A_2 A_m \\ -A_3 A_2 & A_3 - A_3^2 & \cdots & -A_3 A_m \\ \vdots & \vdots & \ddots & \vdots \\ -A_m A_2 & -A_m A_3 & \cdots & A_m - A_m^2 \end{pmatrix} = O$$

这说明 $\forall i \neq j$ 有 $A_i^2 = A_i$ ，且 $A_i A_j = O (i, j = 2, 3, \cdots, m)$ ，再对 $A_1 + \cdots + A_m = E_n$ 两边同乘 A_i 得到

$$A_1 A_i + A_i = A_i$$

因此 $A_1 A_i = O$ ，结论成立

(3) $\Rightarrow$ (1)：直接对 $A_1 + \cdots + A_m = E_n$ 两边右乘 A_i 得到

$$A_1 A_i + A_2 A_i + \cdots + A_m A_i = A_i^2 = A_i$$

因此都是幂等矩阵

 **笔记** 此定理在 ZJU2020 中考过

2.6.10 对角占优矩阵

第3章 行列式

3.1 行列式基本概念与性质

定理 3.1 (行列式初等变换)

矩阵初等变换相对行列式的变换为：

- 互换两行/列：取反号
- 同行/列可同时提出公因式
- 倍加：不改变行列式大小



定理 3.2 (行列式分拆)

行列式可以关于一行或者一列进行分拆，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{s1} + c_{s1} & b_{s2} + c_{s2} & \cdots & b_{sn} + c_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{s1} & b_{s2} & \cdots & b_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{s1} & c_{s2} & \cdots & c_{sn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$



3.2 行列式展开

定义 3.1 (余子式与代数余子式)

将矩阵 A 去除第 i 行，第 j 列后矩阵的行列式称为 a_{ij} 的余子式 M_{ij} ，代数余子式 $A_{ij} := (-1)^{i+j} M_{ij}$



定理 3.3 (行列式展开)

固定列 j ，则按列展开为 $|A| = \sum_{i=1}^n a_{ij} A_{ij}$ 。固定行 i ，则按行展开为 $|A| = \sum_{j=1}^n a_{ij} A_{ij}$



定理 3.4 (不同行相乘为 0)

A 的第 i 行元素与 k 行的代数余子式乘积和为 0： $\sum_{j=1}^n a_{ij} A_{kj} = 0$



3.3 特殊的行列式

3.3.1 爪型行列式

定理 3.5 (正爪型行列式)

正爪型行列式通过对角消去一条直边，再按三角行列式展开



定理 3.6 (异爪型行列式)

异爪型行列式 (一横两斜) 分以下几种情况:

- 行和为 0 型: 全部加到第一行, 按第一列展开
- 两对角各自相等型: 按斜直夹边展开, 获得递推公式

**练习 3.1 异爪型行列式**

(1) 计算 $|A| =$

$$\begin{vmatrix} 0 & -1 & -1 & \cdots & -1 & -1 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{vmatrix}$$

(2) 设 $s(x) = \begin{cases} \frac{x}{|x|}, & x \neq 0 \\ 0, & x = 0 \end{cases}$, 已知 $A = (a_{ij})$, $a_{ij} = s(i-j)$, 求 $|A|$

解 (1) 按最后一列展开, 得到 $D_n = (-1)^{n+2} + D_{n-1}$, 因此 n 为奇数时 $D_n = 0$, n 为偶数时 $D_n = 1$

(2) 上下相减后变为 (1)

3.3.2 大对角型 (三斜线)**3.3.3 行和相同型****定理 3.7 (行和相同型)**

行和相同型行列式分以下两种情况:

- 内部大量相同: 加到第一列, 提出公因式, 消去内部大量元素
- 循环型: 上下相减 (例题见循环型一节)

**练习 3.2 内部大量相同型**

计算行列式

$$\begin{vmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & & & \vdots \\ b & b & b & \cdots & a \end{vmatrix}$$

解 加到第一列, 提出 $[a + (n-1)b]$, 用第一列消去其他元素, 最终得到答案 $(a-b)^{n-1}[a + (n-1)b]$

3.3.4 循环型

循环型一般无论形态都是上下相减, 一般只有行和相同型和反对称型两种。

练习 3.3 循环型

计算 $|A|, |B|$

$$|A| = \begin{vmatrix} 0 & 1 & 2 & \cdots & n-2 & n-1 \\ 1 & 0 & 1 & \cdots & n-3 & n-2 \\ 2 & 1 & 0 & \cdots & n-4 & n-3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-2 & n-3 & n-4 & \cdots & 0 & 1 \\ n-1 & n-2 & n-3 & \cdots & 1 & 0 \end{vmatrix} \quad |B| = \begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ n & 1 & 2 & \cdots & n-2 & n-1 \\ n-1 & n & 1 & \cdots & n-3 & n-2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 3 & 4 & 5 & \cdots & 1 & 2 \\ 2 & 3 & 4 & \cdots & n & 1 \end{vmatrix}$$

解 (1) 依次上面一行减下面一行, 变成如下行列式, 再每一列加上第一列即可

$$\begin{vmatrix} -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & 1 \\ n-1 & n-2 & n-3 & \cdots & 1 & 0 \end{vmatrix}$$

(2) 依次上面一行减下面一行, 然后将最后一列的 -1 倍加到每一列。

3.3.5 行列式加边法

定理 3.8 (行列式加边法)

计算矩阵 A 的行列式 $|A|$ 时, 可加上一边:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$



笔记 行列式加边法往往在内部元素大量相同, 但行和不同的情况。

3.4 Cramer 法则

定理 3.9 (Cramer 法则)

n 个方程的 n 元线性方程组 $AX = B$, 系数矩阵行列式 $|A| \neq 0$, 则其唯一解为

$$\left(\frac{|B_1|}{|A|}, \frac{|B_2|}{|A|}, \dots, \frac{|B_n|}{|A|} \right)$$

其中 B_i 为 A 将第 i 列换为 B 组成的矩阵。



第4章 线性方程组

4.1 线性方程组解的判定定理

4.1.1 解存在判定定理

定理 4.1 (解存在充要条件)

K 上线性方程组 $Ax = b$ 有解的充要条件为 $r(A) = r(\bar{A})$, 这里 $\bar{A} := [A, b]$



4.1.2 线性方程组同解

定理 4.2 (同解定理)

$A_{s \times n}, B_{t \times n}$ 为两个矩阵, 则 $AX = 0, BX = 0$ 同解当且仅当 A, B 行向量等价



练习 4.1 具体应用 (1) 已知下面两个线性方程组同解, 求 a, b, c 的值, 且求通解

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 0 \\ 2x_1 + 3x_2 + 5x_3 = 0 \\ x_1 + x_2 + ax_3 = 0 \end{cases} \quad \begin{cases} x_1 + bx_2 + cx_3 = 0 \\ 2x_1 + b^2x_2 + (c+1)x_3 = 0 \end{cases}$$

(2) $A_{s \times n}, B_{t \times n}$ 为两个矩阵, $AX = 0, BX = 0$ 的解空间分别为 V_1, V_2 , 证明: $V_1 \subseteq V_2$ 当且仅当存在矩阵 $C_{t \times s}$ 使得 $B = CA$

证明 (1) 设方程组为 $AX = 0, BX = 0$, 即 $AX = 0$ 与 $\begin{pmatrix} A \\ B \end{pmatrix} X = 0$ 同解, 做初等行变换即可。

(2) $V_1 \subseteq V_2$ 当且仅当 $AX = 0$ 与 $\begin{pmatrix} A \\ B \end{pmatrix} X = 0$ 同解。

4.1.3 解个数判定定理

定理 4.3 (解个数判定定理)

若 K 上 $Ax = b$ 满足 $r(A) = r(\bar{A})$:

- 若 $r(A) = n$: 则 $Ax = b$ 有唯一解
- 若 $r(A) < n$: 则 $Ax = b$ 有无穷解



4.2 线性方程组的求解

4.2.1 $Ax = 0$ 求解

定义 4.1 (自由未知量)

消元后多余的 $0 = 0$ 位置对应的变量 $x_{r+1}, \dots, x_n$ 称为方程组自由未知量



定理 4.4 (自由变量的个数)自由变量的个数为 $n - r(A)$ **定义 4.2 (基础解系)**让自由未知量 x_{r-1+i} 取 1, 其余取 0, 构造解向量 η_i 称为基础解系:

$$\begin{cases} \eta_1 = (*, \cdots, *, 1, 0, \cdots, 0) \\ \eta_2 = (*, \cdots, *, 0, 1, \cdots, 0) \\ \vdots \\ \eta_{n-r} = (*, \cdots, *, 0, 0, \cdots, 1) \end{cases}$$

 **笔记** 如果只有一个自由未知数, 则一般取 1 而非 0, 否则无法表示出基础解系, 只是要给特解。**定理 4.5 (齐次线性方程组的解)**齐次线性方程组 $Ax = 0$ 的任意解是其基础解系的线性组合**4.2.2 $Ax = b$ 求解****定理 4.6 (非齐次线性方程组的解)**非齐次线性方程组 $Ax = b$ 的任意解是其特解 γ_0 加上齐次方程组 $Ax = 0$ 基础解系线性组合, 即 $x = \gamma_0 + L\{\eta_1, \cdots, \eta_{n-r}\}$  **练习 4.2 求解非齐次线性方程组** (1) 讨论以下线性方程组解的情况

$$\begin{cases} x_1 + x_2 - x_3 = 1 \\ 2x_1 + 3x_2 + kx_3 = 3 \\ x_1 + kx_2 + 3x_3 = 2 \end{cases}$$

(2) 已知 $A = \begin{pmatrix} 1 & -1 & -1 \\ 4 & t & 1 \\ -2 & 2 & t \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \\ -t \end{pmatrix}$, 当 t 为何值时, $AX = b$ 无解, 有唯一解, 有无穷多解, 并

给出有解时对应的解。

解 (1) 对增广矩阵做初等变换得到

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & k+2 & 1 \\ 0 & 0 & (k+3)(k-2) & k-2 \end{bmatrix}$$

先讨论解的情况: 有解当且仅当 $r(A) = r(\bar{A})$, $k = -3$ 无解, $k = 2$ 有无穷多解。其余情况有唯一解。 $k = 2$ 时特解以及齐次基础解系:

$$\gamma = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, y = \begin{pmatrix} 5 \\ -4 \\ 1 \end{pmatrix} \Rightarrow x = \gamma + k_1 y$$

唯一解情况直接解出结果即可。

(2) 做增广矩阵初等变换得到

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 0 & t+4 & 5 & -2 \\ 0 & 0 & t-2 & 2-t \end{bmatrix}$$

$t = -4$ 时无解。 $t = 2$ 无穷多解，其他情况有唯一解。具体求解略去。

4.2.3 线性方程组公共解的计算

 **练习 4.3 公共解的情况** (1) 线性方程组 I 的通解为 $(1, 1, 0, 0)^T + k_1(1, 0, -1, 0)^T + k_2(2, 3, 0, 1)^T$ ，线性方程组 II 如下，若两个方程组有无穷组公共解，求 a, b 的值并表示出所有公共解

$$\begin{cases} 7x_1 - 6x_2 + 3x_3 = b \\ 8x_1 - 9x_2 + ax_4 = 7 \end{cases}$$

解 由于有无穷多组解，因此有无穷组 k_1, k_2 使得 $(1 + k_1 + 2k_2, 1 + 3k_2, -k_1, k_2)$ 为方程组 II 的解。代入方程化简得到

$$\begin{cases} 4k_1 - 4k_2 = b - 1 \\ 8k_1 + (a - 11)k_2 = 8 \end{cases} \Rightarrow \begin{cases} 4k_1 - 4k_2 = b - 1 \\ 0k_1 + (a - 3)k_2 = 10 - 2b \end{cases}$$

需要有无穷多解，得有自由位置量，因此 $a = 3, b = 5, k_1 = k_2 + 1$ ，解系为 $(2, 1, -1, 0)^T + k_2(3, 3, -1, 1)^T$

 **笔记** 若给两个完整的线性方程组，则先求出一个的基础解系，再往另一个方程中代。

4.3 最小二乘解

4.3.1 $A^T A$ 的研究

定理 4.7 ($A^T A$ 与 A 的秩)

设 A 是 $m \times n$ 的实矩阵，则 $r(A^T A) = r(A)$ 。复数域有 $r(A \bar{A}^T) = r(\bar{A}^T A) = r(A)$ ，但不一定有 $r(A^T A) = r(A)$ 。

证明 (1) 只需要考虑 $AX = 0$ 和 $A^T AX = 0$ 解的情况。首先根据 $AX = 0$ 可以推出 $A^T AX = 0$ 。其次 $A^T AX = 0$ 可推出 $X^T A^T AX = (AX)^T (AX) = 0$ ，而 AX 是一个实向量， $\alpha^T \alpha = 0$ 当且仅当 $\alpha = 0$ ，因此 $AX = 0$ 。故它们同解， $n - r(A) = n - r(A^T A)$ 得到 $r(A) = r(A^T A)$

(2) 若 $AX = 0$ ，则 $\bar{A}^T AX = 0$ 。若 $\bar{A}^T AX = 0$ ，则 $\bar{X}^T \bar{A}^T AX = (\bar{A} \bar{X})^T AX = 0$ ，得到 $AX = 0$

 **笔记** 该定理最重要的结论是最小二乘解一定存在。

 **练习 4.4 $A^T A$ 与 A 性质** (1) A 是 $\mathbb{R}$ 上任意 $s \times n$ 矩阵，证明： $r(AA^T A) = r(A)$

证明 (1) $r(AA^T A) \leq r(A)$ ，而 $r(AA^T A) \geq r(A^T AA^T A) = r(A^T A) = r(A)$

4.3.2 正规方程组

定义 4.3 (最小二乘解)

对于一个无解的线性方程组 (矛盾方程组) $AX = b$ ，若向量 $(x_1, \dots, x_s) \in \mathbb{R}^s$ 使得距离

$$d = \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{is}x_s - b_i)^2$$

最小，则称该向量为最小二乘解。

定理 4.8 (正规方程组)

对于任一方程组 $AX = b$, 最小二乘解一定存在, 且 X_0 是最小二乘解当且仅当 $A^T AX = A^T b$



证明 设 $\alpha_1, \dots, \alpha_s$ 是 A 的列向量, 定义 $W = L(\alpha_1, \dots, \alpha_s)$, X_0 是最小二乘解当且仅当 AX_0 是 b 在 W 上的投影, 即 $(b - AX_0) \perp \alpha_i$, 因此

$$\alpha_i^T (b - AX_0) \Rightarrow A^T (b - AX_0) = 0 \Rightarrow A^T AX_0 = A^T b$$

定理 4.9 (正规方程组解的存在性)

方程组 $A^T AX = A^T b$ 一定有解。



证明 因为 $r(A^T A) = r(A)$, 先将增广矩阵拆开得到 $r([A^T A, A^T b]) = r(A^T [A, b])$, 用秩不等式得到 $r(A^T [A, b]) \leq r(A^T) = r(A) = r(A^T A)$, 而显然 $r([A^T A, A^T b]) \geq r(A^T A)$, 因此增广矩阵的秩等于系数矩阵的秩, 综上得到最小二乘法解存在。

定理 4.10 (凸闭集最佳逼近元的唯一性)

$\mathcal{H}$ 是 Hilbert 空间, $M \subset \mathcal{H}$ 是凸闭集, 则 $\forall x \in \mathcal{H}$ 在 M 中存在唯一最佳逼近。



笔记 一般而言最小二乘解不唯一, 除非是凸闭集。

第5章 线性空间与线性变换

5.1 线性空间

5.1.1 线性空间的基本概念

定义 5.1 (线性空间)

$\mathbb{P}$ 是一个数域, V 是非空集合, V 上定义了二元运算 $+$ 和数乘, 若这两种运算满足:

- 二元运算: (1) 交换 (2) 结合 (3) 零元 (4) 负元
- 数乘: (1) 一元 (2) 结合 $(c_1 c_2) \alpha = c_1 (c_2 \alpha)$ (3) 向量分配 $(c_1 + c_2) \alpha = c_1 \alpha + c_2 \alpha$ (4) 数分配 $c(\alpha + \beta) = c\alpha + c\beta$

则称 V 是 $\mathbb{P}$ 上的线性空间



5.1.2 线性相关、线性无关与向量组的秩

定义 5.2 (线性相关和线性无关)

$\alpha_1, \dots, \alpha_n$ 为向量组, 若存在不全为 0 的 $k_1, \dots, k_n$ 使得

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n = 0$$

则称它们线性相关。若该等式成立当且仅当 $k_1 = k_2 = \dots = k_n = 0$, 则称线性无关。



练习 5.1 证明线性无关 (1) 幂零基: $\mathcal{A}$ 为 V 上的线性变换, $\xi \in V$, 若 $\mathcal{A}^{k-1} \xi \neq 0, \mathcal{A}^k \xi = 0$, 证明 $\xi, \mathcal{A}\xi, \dots, \mathcal{A}^{k-1} \xi$ 线性无关

证明 设 $k_0 \xi + k_1 \mathcal{A}\xi + \dots + k_{n-1} \mathcal{A}^{n-1} \xi = 0$, 两侧同时作用 $\mathcal{A}^{k-1}$ 得到 $k_0 \mathcal{A}^{k-1} \xi = 0$, 因此 $k_0 = 0$, 以此类推得到 $k_0, k_1, \dots, k_{n-1} = 0$

定理 5.1 (极大线性无关组的计算)

给定一系列向量 $\alpha_1, \dots, \alpha_n$, 将它们作为矩阵的列向量拼成一个矩阵, 对矩阵做初等行变换化为阶梯型矩阵, 阶梯头所在列对应的向量为极大线性无关组。(其实按行也可以, 按列更不容易出错, 因为换行不影响结果)



练习 5.2 极大线性无关组的计算 计算下列向量组的极大线性无关组

$$\begin{cases} \alpha_1 = (1, -1, 2, 4) \\ \alpha_2 = (0, 3, 1, 2) \\ \alpha_3 = (3, 0, 7, 14) \\ \alpha_4 = (1, -1, 2, 0) \\ \alpha_5 = (2, 1, 5, 6) \end{cases}$$

解 拼成一个矩阵, 并做初等行变换:

$$\begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ -1 & 3 & 0 & -1 & 1 \\ 2 & 1 & 7 & 2 & 5 \\ 4 & 2 & 14 & 0 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 3 & 1 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

因此极大线性无关组为 $\alpha_1, \alpha_2, \alpha_4$

定义 5.3 (向量组的秩)

$\alpha_1, \dots, \alpha_n$ 为向量组, 称其中极大线性无关组的向量个数为它们的秩。

定理 5.2 (源泉定理及其推论)

设 $\alpha_1, \dots, \alpha_r$ 可以被 $\beta_1, \dots, \beta_s$ 线性表出, $r > s$, 则 $\alpha_1, \dots, \alpha_r$ 线性相关。

5.1.3 线性空间同构

定义 5.4 (同构)

若两个代数系统之间存在双射, 且该双射保持代数系统的运算, 则称两个代数系统同构。

定理 5.3 (有限维线性空间的同构映射)

V 为 $\mathbb{P}$ 上线性空间, $\alpha_1, \dots, \alpha_n$ 为 V 的一组基, 构造线性映射 f , 将 $\forall \alpha \in V$ 映射到其坐标, 则 f 为 $V \rightarrow \mathbb{P}^n$ 的同构映射

定理 5.4 (用坐标研究向量组)

$\alpha_1, \dots, \alpha_n$ 是 V 的一组基, $A_{n \times s}$, 向量组 $\beta_1, \dots, \beta_s$ 满足

$$(\beta_1, \dots, \beta_s) = (\alpha_1, \dots, \alpha_n)A$$

则 $\dim(\beta_1, \dots, \beta_s) = r(A)$

证明 $A = [X_1, \dots, X_s]$ 即 $\beta_1, \dots, \beta_s$ 的坐标, 因此 A 的列向量与 $\beta_1, \dots, \beta_s$ 等价, 从而 $r(A) = \dim(\beta_1, \dots, \beta_s)$

推论 5.1 (线性无关情况)

$\alpha_1, \dots, \alpha_n$ 线性无关, $(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n)A$, 则 β_i 线性无关当且仅当 A 可逆。

练习 5.3 同构的应用 (1) $\beta_1 = \alpha_1 + \alpha_2, \beta_2 = \alpha_1 + \alpha_3, \beta_3 = \alpha_1 + \alpha_2$, 证明: $\alpha_1, \alpha_2, \alpha_3$ 和 $\beta_1, \beta_2, \beta_3$ 等价

(2) 证明 n 为奇数时, $\alpha_1, \dots, \alpha_n$ 线性无关当且仅当 $\alpha_1 + \alpha_2, \dots, \alpha_{n-1} + \alpha_n, \alpha_n + \alpha_1$ 线性无关

(3) 设 $\beta_1 = \alpha_2 + \dots + \alpha_n, \beta_2 = \alpha_1 + \alpha_3 + \dots + \alpha_n, \beta_n = \alpha_1 + \dots + \alpha_{n-1}$, 证明 $\alpha_1, \dots, \alpha_n$ 和 $\beta_1, \dots, \beta_n$ 有相同的秩

证明 (1) 如下同构映射至坐标, 显然 A 可逆, 因此等价

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3) \begin{bmatrix} 1 & 1 & 1 \\ 1 & & 1 \\ 0 & 1 & \end{bmatrix} := (\alpha_1, \alpha_2, \alpha_3)A$$

(2) 将后者设为 $\beta_1, \dots, \beta_n$, 如下线性表出

$$(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n) \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & \\ & & 1 & \ddots & \\ & & & \ddots & 1 \\ & & & & 1 & 1 \end{bmatrix}$$

用 Gauss 消元法可看出 n 为奇数时矩阵满秩, 因此等价。

(3) 同理列矩阵, 用加边法可算出行列式非零, 因此 A 可逆, 即等价。

5.1.4 基、坐标

定义 5.5 (基与坐标)

V 是线性空间, $\epsilon_1, \dots, \epsilon_n$ 是 V 中线性无关的向量组, 若 $\forall \alpha \in V$ 可被该向量组线性表出, 则称该向量组为 V 的一组基。且若

$$\alpha = (\epsilon_1, \dots, \epsilon_n)X$$

则称 X 为 α 在 $\epsilon_1, \dots, \epsilon_n$ 基下的坐标



 **笔记** 坐标的计算: 直接根据 $\alpha = (\epsilon_1, \dots, \epsilon_n)X$ 解线性方程组 (或者用初等变换求逆)

$$X = (\epsilon_1, \dots, \epsilon_n)^{-1} \alpha$$

 **练习 5.4 证明基** 给定 $\alpha_1 = (1, 1, \dots, 1, 1), \alpha_2 = (1, 1, \dots, 1, 0), \dots, \alpha_n = (1, 0, \dots, 0)$, 证明其是 $\mathbb{P}^n$ 的一组基
证明 当成列向量排成一个矩阵, 满秩即可

 **练习 5.5 坐标的计算** (1) 已知 $\epsilon_1, \epsilon_2, \epsilon_3$ 是 $\mathbb{P}^3$ 的一组基, 计算 $\alpha = (1, 2, 1)^T$ 在这组基下的坐标

$$\epsilon_1 = (1, 1, 1)^T, \epsilon_2 = (1, 1, -1)^T, \epsilon_3 = (1, -1, -1)^T$$

解 (1) 只需要解方程组 $\alpha = k_1\epsilon_1 + k_2\epsilon_2 + k_3\epsilon_3$, 即 $[\epsilon_1, \epsilon_2, \epsilon_3]X = \alpha$:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{bmatrix} X = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

解得 $k_1 = 1, k_2 = \frac{1}{2}, k_3 = -\frac{1}{2}$

 **练习 5.6 难度进阶** (1) 求非零向量 $\alpha \in \mathbb{P}^4$, 使得其在下面两组基中有相同的坐标

$$\begin{cases} \epsilon_1 = (1, 0, 0, 0)^T \\ \epsilon_2 = (0, 1, 0, 0)^T \\ \epsilon_3 = (0, 0, 1, 0)^T \\ \epsilon_4 = (0, 0, 0, 1)^T \end{cases}, \begin{cases} \eta_1 = (2, 1, -1, 1)^T \\ \eta_2 = (0, 3, 1, 0)^T \\ \eta_3 = (5, 3, 2, 1)^T \\ \eta_4 = (6, 6, 1, 3)^T \end{cases}$$

解 (1) 设 α 在第一组基下的坐标为 $(x_1, x_2, x_3, x_4)^T$, 则由于同坐标

$$\alpha = x_1\eta_1 + x_2\eta_2 + x_3\eta_3 + x_4\eta_4$$

解线性方程组 $[\eta_1, \eta_2, \eta_3, \eta_4]X = \alpha$ 即可 (注意将 α 减到左边):

$$\begin{bmatrix} 1 & 0 & 5 & 6 \\ 1 & 2 & 3 & 6 \\ -1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix} X = 0$$

解得 $\alpha = k(1, 1, 1, -1)$

5.1.5 坐标变换

定义 5.6 (过渡矩阵)

给定 $(\epsilon_1, \dots, \epsilon_n), (\eta_1, \dots, \eta_n)$ 两组基, 若存在矩阵 T 满足:

$$(\epsilon_1, \dots, \epsilon_n) = (\eta_1, \dots, \eta_n)T$$

则称 T 为 $\eta_1, \dots, \eta_n$ 到 $\epsilon_1, \dots, \epsilon_n$ 的过渡矩阵 (不要搞反)

 **笔记** 过渡矩阵的计算: 直接根据 $(\epsilon_1, \dots, \epsilon_n) = (\eta_1, \dots, \eta_n)T$ 得到

$$T = (\eta_1, \dots, \eta_n)^{-1}(\epsilon_1, \dots, \epsilon_n)$$

用初等变换求 $A^{-1}B$ 的方式直接求 T 即可。

定理 5.5 (坐标变换)

V 是 n 维线性空间, $\eta_1, \dots, \eta_n$ 和 $\epsilon_1, \dots, \epsilon_n$ 是空间两组基, 且 $(\eta_1, \dots, \eta_n) = (\epsilon_1, \dots, \epsilon_n)T$, 向量 $x \in V$ 在 ϵ_i, η_i 下的坐标分别为 X, Y , 则

$$X = TY$$

证明 $x = (\epsilon_1, \dots, \epsilon_n)X = (\eta_1, \dots, \eta_n)Y = (\epsilon_1, \dots, \epsilon_n)TY$ 即可。

 **笔记** 坐标变换用荡秋千记忆, 即 $(\eta_1, \dots, \eta_n) = (\epsilon_1, \dots, \epsilon_n)T$, 分别对应 Y, X , 则把 T 荡到另一边, 即 $TY = X$ 。这与线性变换在不同基下的矩阵一样都需要记住。

 **练习 5.7 过渡矩阵的计算** 给定两组基, 求它们之间的过渡矩阵, 并求 $\xi = (1, 0, 0, -1)$ 在 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 下的坐标

$$\alpha_1 = (1, 1, 1, 1) \quad \alpha_2 = (1, 1, -1, -1) \quad \alpha_3 = (1, -1, 1, -1) \quad \alpha_4 = (1, -1, -1, 1)$$

$$\beta_1 = (1, 1, 0, 1) \quad \beta_2 = (2, 1, 3, 1) \quad \beta_3 = (1, 1, 0, 0) \quad \beta_4 = (0, 1, -1, -1)$$

解 (1) 先取标准基 $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$, 设

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)A, (\beta_1, \beta_2, \beta_3, \beta_4) = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)B$$

则

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & -1 \\ 1 & 1 & 0 & -1 \end{bmatrix}$$

由此得到

$$(\beta_1, \beta_2, \beta_3, \beta_4) = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)B = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)A^{-1}B$$

综上得到过渡矩阵为 (具体计算时用行初等变换: $(A|B) \rightarrow (E|A^{-1}B)$)

$$A^{-1}B = \frac{1}{4} \begin{bmatrix} 3 & 7 & 2 & -1 \\ 1 & -1 & 2 & 3 \\ -1 & 3 & 0 & -1 \\ 1 & -1 & 0 & -1 \end{bmatrix}$$

(2) ξ 在 $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ 下的坐标即本身, 因此在 $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ 下坐标为 $A^{-1}\xi$ (具体计算时也 and $(A|B)$ 摆在一起), 结果为 $(0, \frac{1}{2}, \frac{1}{2}, 0)^T$

5.1.6 和与交、维数定理

定义 5.7 (线性空间和与交)

V_1, V_2 为线性空间, 则

$$V_1 \cap V_2 := \{\alpha : \alpha \in V_1, \alpha \in V_2\}, V_1 + V_2 := \{\alpha = \alpha_1 + \alpha_2 : \alpha_1 \in V_1, \alpha_2 \in V_2\}$$

且显然 $V_1 \cap V_2, V_1 + V_2$ 都是子空间。



 **笔记** $V_1 \cup V_2$ 由于运算不封闭, 一般而言不是子空间。注意分别 $V_1 \cup V_2$ 和 $V_1 + V_2$

 **练习 5.8 和与交基本计算** (1) 已知 $V_1 = L(\alpha_1, \alpha_2, \alpha_3), V_2 = L(\beta_1, \beta_2)$, 其中

$$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 0 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \beta_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \beta_2 = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 7 \end{pmatrix}$$

求 $V_1 + V_2, V_1 \cap V_2$ 的基与维数

(2) 已知 $W_1 = L(\alpha_1, \alpha_2, \alpha_3), W_2 = L(\beta_1, \beta_2)$, 求 $W_1 \cap W_2, W_1 + W_2$ 的基与维数。

$$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -3 \end{pmatrix}, \alpha_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \alpha_3 = \begin{pmatrix} -1 \\ -3 \\ 0 \\ 5 \end{pmatrix}, \beta_1 = \begin{pmatrix} -1 \\ 0 \\ 4 \\ 2 \end{pmatrix}, \beta_2 = \begin{pmatrix} 0 \\ 5 \\ 9 \\ -14 \end{pmatrix}$$

解 (1) $V_1 + V_2 = L(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2)$, 因此只需要排成列, 做初等变换:

$$\begin{pmatrix} 1 & -1 & 0 & 2 & 1 \\ 2 & 1 & 3 & -1 & -1 \\ 1 & 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 1 & 7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} := A$$

因此维数为 3, 基为 $\alpha_1, \alpha_2, \beta_1$

任取 $\alpha \in V_1 \cap V_2$, 设 $\alpha = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = y_1\beta_1 + y_2\beta_2$, 因此系数需要满足方程组:

$$x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 - y_1\beta_1 - y_2\beta_2 = 0$$

设 $x_4 = -y_1, x_5 = y_2$, 即求解 $AX = [\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2]X = 0$, 解得基础解系为

$$\eta_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \eta_2 = \begin{pmatrix} 1 \\ -4 \\ 0 \\ -3 \\ 1 \end{pmatrix}$$

对应 $V_1 \cap V_2$ 的基础解系为:

$$-\alpha_1 - \alpha_2 + \alpha_3 = (0, 0, 0, 0)^T, \alpha_1 - 4\alpha_2 = (5, -2, -3, -4)^T$$

只取非零的, 故只能为 $k(5, -2, -3, -4)$, 维数为 1, 基即 $\alpha_1 - 4\alpha_2$ 。

- 练习 5.9 和与并的区别** (1) 设 V_1, V_2 是 $\mathbb{P}$ 上线性空间 V 的两个非平凡子空间, 证明: $\exists \alpha \in V$ 使得 $\alpha \notin V_1, \alpha \notin V_2$
 (2) $V_1, \dots, V_s$ 是 V 的 s 个非平凡子空间, 证明: V 中至少有一个向量不属于 $V_1, \dots, V_s$ 中的任意一个。
 (3) V_1, V_2 是两个子空间, 证明: $V_1 + V_2 = V_1 \cup V_2$ 当且仅当 $V_1 \subseteq V_2$ 或者 $V_2 \subseteq V_1$

证明 (1) 由于 V_1 非平凡, $\exists \alpha \notin V_1$, 若 $\alpha \notin V_2$, 则结论成立; 否则 $\alpha \in V_2$, 由于 V_2 非平凡, 因此 $\exists \beta \notin V_2$, 若 $\beta \notin V_1$, 则结论成立; 否则 $\beta \in V_1$, 此时有

$$\alpha \notin V_1, \beta \in V_1, \alpha \in V_2, \beta \notin V_2$$

因此 $\gamma = \alpha + \beta \notin V_1, \gamma \notin V_2$

(2) 用归纳法

(3) 右推左显然。左推右: 设 $V_1 + V_2 = V_1 \cup V_2$, 但 $V_1 \not\subseteq V_2$, 则 $\exists \alpha_1 \in V_1, \alpha_1 \notin V_2$, 但

$$\forall \alpha_2 \in V_2, \alpha = \alpha_1 + \alpha_2 \in V_1 + V_2$$

由于 $V_1 + V_2 = V_1 \cup V_2$, 因此 α 至少属于 V_1, V_2 中的一个, 显然 $\alpha \notin V_2$ (根据加法封闭), 因此 $\alpha \in V_1$, 故

$$\alpha_2 = \alpha - \alpha_1 \in V_1$$

综上得到 $V_2 \subseteq V_1$

定理 5.6 (维数公式)

设 V_1, V_2 是有限维线性空间 V 的两个子空间, 则

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$$



证明 设小扩大。设 $V_1 \cap V_2$ 的基为 $\alpha_1, \dots, \alpha_s$, V_1, V_2 的基分别为:

$$\begin{cases} \alpha_1, \dots, \alpha_s, \beta_{s+1}, \dots, \beta_m \\ \alpha_1, \dots, \alpha_s, \gamma_{s+1}, \dots, \gamma_n \end{cases}$$

则 $V_1 + V_2 = L(\alpha_1, \dots, \alpha_s, \beta_{s+1}, \dots, \beta_m, \gamma_{s+1}, \dots, \gamma_n)$, 下面证明上面的向量全部线性无关。不妨设

$$k_1\alpha_1 + \dots + k_s\alpha_s + k_{s+1}\beta_{s+1} + \dots + k_n\beta_n + l_{s+1}\gamma_{s+1} + \dots + l_n\gamma_n = 0$$

将 $l_i\gamma_i$ 的部分移到右侧, 此时左侧线性组合属于 V_1 , 右侧属于 V_2 , 因此都属于 $V_1 \cap V_2$ 。根据左侧得到:

$$k_1\alpha_1 + \dots + k_s\alpha_s + k_{s+1}\beta_{s+1} + \dots + k_n\beta_n \in V_1 \cap V_2 \Rightarrow k_{s+1} = k_{s+2} = \dots = k_n = 0 \Rightarrow k_1 = \dots = k_s = 0$$

因此左侧系数全为 0, 同理右侧系数也全为 0

- 练习 5.10 维数公式相关** (1) 经典: V_1, V_2 是有限维线性空间 V 的子空间, $\dim(V_1 + V_2) = \dim(V_1 \cap V_2) + 1$, 证明: 要么 $V_1 \subseteq V_2$, 要么 $V_2 \subseteq V_1$

证明 (1) 根据维数公式得到

$$\dim(V_1) + \dim(V_2) = \dim(V_1 + V_2) + \dim(V_1 \cap V_2) = 2\dim(V_1 \cap V_2) + 1$$

等价于

$$(\dim V_1 - \dim(V_1 \cap V_2)) + (\dim V_2 - \dim(V_1 \cap V_2)) = 1$$

由于都是整数, 只能一个 0, 一个 1, 因此结论成立。

5.1.7 直和理论

定义 5.8 (直和)

V_1, V_2 是线性空间 V 的子空间, 若 $V_1 \cap V_2 = \{0\}$, 则记 $V_1 \oplus V_2 := V_1 + V_2$

定义 5.9 (补空间)

若 $V = V_1 \oplus V_2$, 则称 V_2 是 V_1 的补空间。

 **笔记** 补空间一般是不唯一的, 因为 $\alpha_1, \dots, \alpha_m$ 扩充为 $\alpha_1, \dots, \alpha_n$ 的方式不唯一。

定理 5.7 (直和等价条件)

$W = V_1 \oplus V_2$ 当且仅当

- W 中任意向量 $\alpha = \alpha_1 + \alpha_2, \alpha_1 \in V_1, \alpha_2 \in V_2$ 唯一表出
- W 中零向量被 V_1, V_2 中向量唯一表出: $0 = 0 + 0$
- $\dim W = \dim V_1 + \dim V_2$

 **练习 5.11 判断直和** (1) 数域 P 上所有 n 阶矩阵组成线性空间 $V = M_n(P)$, V_1 表示所有对称矩阵组成的集合, V_2 表示所有反对称矩阵组成的集合, 证明 V_1, V_2 是子空间且 $V = V_1 \oplus V_2$

证明 (1) 显然关于加法和数乘封闭, 因此是子空间。首先证明 $V = V_1 + V_2$, 由于矩阵 A 都可以写为:

$$A = \frac{A + A^T}{2} + \frac{A^T - A}{2}$$

从而 $V = V_1 + V_2$ 。

证明 $V = V_1 \oplus V_2$: 可用 $V_1 \cap V_2 = \{0\}$, 此时取 $A \in V_1 \cap V_2$, 则 $A = A^T = -A^T$, 得到 $A^T = A = O$, 因此为直和。也可以用 O 分解唯一, 设 $O = B + C = B^T - C^T$, 而 $O^T = B^T + C^T$, 因此 $B = C = O$ 。

定理 5.8 (多空间直和等价条件)

$V_1, \dots, V_s$ 是 W 的子空间, 且 $W = V_1 + V_2 + \dots + V_s$, 则 $W = V_1 \oplus V_2 \oplus \dots \oplus V_s$ 当且仅当

- W 中零向量分解唯一
- W 中任意向量 α 分解唯一
- $V_i \cap \sum_{j \neq i} V_j = \{0\}$
- $\dim W = \dim V_1 + \dots + \dim V_s$

证明 (3) 只证明第三条。第三条推直和: 若

$$0 = \alpha_1 + \dots + \alpha_n \Rightarrow \alpha_i = \alpha_1 + \dots + \alpha_{i-1} + \alpha_{i+1} + \dots + \alpha_n$$

由于 $V_i \cap \sum_{j \neq i} V_j = \{0\}$, 因此 $\alpha_k = 0, \forall k = 1, 2, \dots, s$, 从而是直和

直和推第三条: 若 $\alpha \in V_i \cap \sum_{j \neq i} V_j$, 则 $\alpha \in V_i$, 且

$$\alpha = \alpha_1 + \dots + \alpha_{i-1} + \alpha_{i+1} + \dots + \alpha_n \Rightarrow \alpha_1 + \dots + \alpha_{i-1} + (-\alpha) + \alpha_{i+1} + \dots + \alpha_n = 0$$

由于 $\sum_{i=1}^n V_i$ 为直和, 因此 $\alpha = 0$, 即第三条结论成立。

定理 5.9 (直和合并定理)

若 $V = V_1 \oplus V_2$, $V_1 = V_{11} \oplus V_{12} \oplus \cdots \oplus V_{1s}$, $V_2 = V_{21} \oplus \cdots \oplus V_{2t}$, 则

$$V = V_{11} \oplus \cdots \oplus V_{1s} \oplus V_{21} \oplus \cdots \oplus V_{2t}$$



证明 只证明 $V = V_1 \oplus V_2$, $V_1 = V_{11} \oplus V_{12}$ 时, $V = V_{11} \oplus V_{12} \oplus V_2$ 。

首先显然有 $V = V_{11} + V_{12} + V_2$ 。设

$$0 = \alpha_{11} + \alpha_{12} + \alpha_2 = (\alpha_{11} + \alpha_{12}) + \alpha_2$$

由于 $V = V_1 \oplus V_2$, 因此得到 $\alpha_{11} + \alpha_{12} = \alpha_2 = 0$, 同理根据 $V_1 = V_{11} \oplus V_{12}$ 得到 $\alpha_{11} = \alpha_{12} = 0$, 综上 0 元素分解唯一, 为直和。

5.1.8 商空间

定义 5.10 (商空间的元素)

W 是线性空间 V 的子空间, $\alpha \in V$, 定义

$$\alpha + W := \{\alpha + \omega : \omega \in W\}$$

**定义 5.11 (商空间)**

定义 V/W 为 $\{\alpha + W : \alpha \in V\}$, 即线性空间子集组成的集合。可以证明其是一个线性空间

**定义 5.12 (自然映射)**

定义 $\eta : V \rightarrow V/W$ 为自然映射



5.2 线性映射

5.2.1 线性映射的概念

定义 5.13 (线性映射)

V, W 是 $\mathbb{P}$ 上的线性空间, $\varphi : V \rightarrow W$ 若满足:

- $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$
- $\varphi(c\alpha) = c\varphi(\alpha)$

则称 φ 为 V 到 W 的线性映射, 将 V 到 W 的线性映射全体记为集合 $\text{Hom}_{\mathbb{P}}(V, W)$



5.2.2 核与像

核与像是最常见的不变子空间, 因此在进一步研究线性映射之前, 要先研究这两个特殊的集合。

定义 5.14 (核与像)

U, V 是 $\mathbb{P}$ 上的线性空间, $f : U \rightarrow V$, 定义 $\text{Ker}(f) = \{\alpha \in U : f(\alpha) = 0\}$, $\text{Im}(f) = \{f(\alpha) : \alpha \in U\}$



笔记 像的维度对应着 $(\alpha_1, \cdots, \alpha_n)A$ 的秩, 核的维度对应着线性方程组 $AX = 0$ 的解空间维数。

定理 5.10 (核像维数公式)

$f \in \text{Hom}(U, V)$ 其核与像满足:

$$\dim \text{Ker}(f) + \dim \text{Im}(f) = \dim(U)$$



证明 设 $\alpha_1, \dots, \alpha_r$ 是 $\text{Ker} f$ 的基, $\alpha_1, \dots, \alpha_n$ 是 U 的基, 而 $\text{Im} f = L(f(\alpha_1), \dots, f(\alpha_n)) = L(f(\alpha_{r+1}), \dots, f(\alpha_n))$, 下面证明线性无关。假设 $k_{r+1}f(\alpha_{r+1}) + \dots + k_nf(\alpha_n) = 0$, 根据线性性可知 $f(k_{r+1}\alpha_{r+1} + \dots + k_n\alpha_n) = 0$, 即 $k_{r+1}\alpha_{r+1} + \dots + k_n\alpha_n \in \text{Ker}(f)$, 根据基的线性无关性, 显然 $k_{r+1}, \dots, k_n$ 均为 0



笔记 该定理的证明会在题目中出现, 主要方式是先设原空间的子空间的一组基, 扩展成原空间的一组基, 再作用 f 后与像空间进行联系。

练习 5.12 核像维数公式推广证明 证明以下核像命题

(1) 定义域为子空间: $\mathcal{A}$ 是有限维线性空间 V 的线性变换, W 是 V 的子空间, 证明:

$$\dim(\mathcal{A}W) + \dim(\text{Ker} \mathcal{A} \cap W) = \dim(W)$$

(2) 复合映射: σ, τ 是有限维空间 $U \rightarrow V, V \rightarrow W$ 的线性映射, 证明:

$$\dim(\text{Ker}(\sigma)) + \dim(\text{Im}(\sigma) \cap \text{Ker}(\tau)) = \dim(\text{Ker}(\tau\sigma))$$

证明 (1) 设 $\alpha_1, \dots, \alpha_r$ 是 $\text{Ker} \mathcal{A} \cap W$ 的一组基, $\alpha_1, \dots, \alpha_n$ 是 W 的一组基, 则 $\mathcal{A}W = L(\mathcal{A}\alpha_{r+1}, \dots, \mathcal{A}\alpha_n)$ 。考虑 $\mathcal{A}W$ 的维数, 设 $k_{r+1}\mathcal{A}\alpha_{r+1} + \dots + k_n\mathcal{A}\alpha_n = 0$, 则有 $k_{r+1}\alpha_{r+1} + \dots + k_n\alpha_n \in \text{Ker} \mathcal{A} \cap W$, 因此得到 $k_{r+1} = \dots = k_n = 0$ 。

(2) 先得找原空间最小的空间, 可知 $\text{Ker}(\sigma) \subseteq \text{Ker}(\tau\sigma)$, 因此设 $\alpha_1, \dots, \alpha_r$ 是 $\text{Ker}(\sigma)$ 的基, $\alpha_1, \dots, \alpha_n$ 是 $\text{Ker}(\tau\sigma)$ 的基。设 $V_1 = L(\sigma\alpha_{r+1}, \dots, \sigma\alpha_n)$ 。显然 $V_1 \subseteq \text{Im}(\sigma) \cap \text{Ker}(\tau)$ 。

另一侧 $\forall \alpha \in \text{Im}(\sigma) \cap \text{Ker}(\tau)$, $\exists \beta$ 使得 $\sigma(\beta) = \alpha, \tau(\alpha) = 0$, 这说明 $\tau(\sigma(\beta)) = 0$, 因此 $\beta \in \text{Ker}(\tau\sigma)$, 即 $\alpha \in V_1$, 最终得出 $V_1 = \text{Im}(\sigma) \cap \text{Ker}(\tau)$ 。

定理 5.11 (核像与单射、满射)

σ 是 $U \rightarrow W$ 的线性映射, 则:

- σ 是满射当且仅当 $\text{Im}(\sigma) = W$
- σ 是单射当且仅当 $\text{Ker}(\sigma) = \{0\}$
- 推论: 有限维线性空间上线性变换 (非映射) 是单射当且仅当其是满射 (一般单射更好证)

**定理 5.12 (核、像任意性)**

V 是 P 上 n 维线性空间, V_1, V_2 是其子空间, $\dim V_1 + \dim V_2 = n$, 则存在 V 上的线性变换 $\mathcal{A}$ 使得 $\text{Ker} \mathcal{A} = V_1, \text{Im} \mathcal{A} = V_2$



证明 设 V_1 的一组基为 $\alpha_1, \dots, \alpha_r$, 扩充为 V 的一组基 $\alpha_1, \dots, \alpha_n$, 设 $\beta_{r+1}, \dots, \beta_n$ 是 V_2 的一组基, 构造映射 $\mathcal{A}(k_1\alpha_1 + \dots + k_n\alpha_n) = k_{r+1}\beta_{r+1} + \dots + k_n\beta_n$ 。可证明 $\mathcal{A}$ 是线性变换, 且核、像满足条件。

练习 5.13 计算核与像 (1) $\mathcal{A}: \mathbb{P}^4 \rightarrow \mathbb{P}^3$, 求 $\text{Ker} \mathcal{A}, \text{Im} \mathcal{A}$

$$\mathcal{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 - 3x_2 + x_3 - 2x_4 \\ -x_1 - 11x_2 + 2x_3 - 5x_4 \\ 3x_1 + 5x_2 + x_4 \end{pmatrix}$$

解 (1) $\text{Ker} \mathcal{A}$ 相当于 $AX = 0$ 的解空间, $\text{Im} \mathcal{A}$ 相当于 A 的列空间

5.2.3 幂等变换的核与像

一般而言核与像虽然有维数关系，但不具备直和关系(可能相交也可能和不为全空间)，但是幂等变换是一个特例。

定义 5.15 (幂等变换)

若线性变换 $\mathcal{A}$ 满足 $\mathcal{A}^2 = \mathcal{A}$ ，则称其为幂等变换。

定理 5.13 (幂等变换核像空间分解)

$\mathcal{A}$ 是 V 上的幂等变换，则

$$V = \text{Ker}\mathcal{A} \oplus \text{Im}\mathcal{A}$$

证明 (1) 先证明 V 是核与像的和: $\forall \alpha \in V$, 有 $\alpha = \alpha - \mathcal{A}\alpha + \mathcal{A}\alpha$, 其中 $\mathcal{A}\alpha \in \text{Im}\mathcal{A}$ 。而 $\mathcal{A}(\alpha - \mathcal{A}\alpha) = \mathcal{A}\alpha - \mathcal{A}^2\alpha = 0$, 因此 $\alpha - \mathcal{A}\alpha \in \text{Ker}\mathcal{A}$

(2) 证明直和: $\forall \alpha \in \text{Ker}\mathcal{A} \cap \text{Im}\mathcal{A}$, 则 $\mathcal{A}\alpha = 0$ 且 $\exists \beta, \mathcal{A}\beta = \alpha$, 两侧同时作用 $\mathcal{A}$, 左侧得到 $\mathcal{A}^2\beta = \mathcal{A}\beta = \alpha$, 右侧 $\mathcal{A}\alpha = 0$, 因此 $\alpha = 0$

练习 5.14 经典幂等变换核像题 $\mathcal{A}, \mathcal{B}$ 为 $\mathbb{P}$ 上线性空间中的幂等变换，证明：

(1) $\mathcal{A}, \mathcal{B}$ 有相同的像当且仅当 $\mathcal{A}\mathcal{B} = \mathcal{B}, \mathcal{B}\mathcal{A} = \mathcal{A}$

(2) $\mathcal{A}, \mathcal{B}$ 有相同的核当且仅当 $\mathcal{A}\mathcal{B} = \mathcal{A}, \mathcal{B}\mathcal{A} = \mathcal{B}$

证明 (1) 左推右: $\forall \alpha \in V$ 有 $\mathcal{B}\alpha \in \text{Im}(\mathcal{B})$, 根据条件即 $\mathcal{B}\alpha \in \text{Im}\mathcal{A}$, 因此根据幂等变换投影的性质 ($\mathcal{A}$ 不改变 $\text{Im}\mathcal{A}$ 中的向量) 得到

$$\mathcal{A}(\mathcal{B}\alpha) = \mathcal{B}\alpha \Rightarrow \mathcal{A}\mathcal{B} = \mathcal{B}$$

同理得到 $\mathcal{B}\mathcal{A} = \mathcal{A}$

右推左: $\forall \gamma \in \text{Im}(\mathcal{A})$, 有 $\mathcal{A}\gamma = \gamma$, 而 $\mathcal{B}\mathcal{A}\gamma = \mathcal{A}\gamma = \gamma$, 因此 $\mathcal{B}\gamma = \gamma$, 即 $\gamma \in \text{Im}(\mathcal{B})$, 另一侧同理

(2) 左推右: 由于 $V = \text{Im}(\mathcal{A}) \oplus \text{Ker}(\mathcal{A})$, 因此 $\forall \alpha \in V$ 有

$$\alpha = \alpha_1 + \alpha_2, \alpha_1 \in \text{Im}(\mathcal{A}), \alpha_2 \in \text{Ker}\mathcal{A}$$

$\alpha - \alpha_1 \in \text{Ker}\mathcal{A}$, 因此 $\alpha - \alpha_1 \in \text{Ker}\mathcal{B}$, 故 $\mathcal{B}(\alpha - \alpha_1) = 0$, 即 $\mathcal{B}\alpha = \mathcal{B}\alpha_1$ 。而 $\mathcal{A}\alpha = \alpha_1$, 得到

$$(\mathcal{B}\mathcal{A})\alpha = \mathcal{B}\alpha_1 = \mathcal{B}\alpha$$

因此得出 $\mathcal{B}\mathcal{A} = \mathcal{B}$

右推左

5.2.4 一般线性变换的核像分解

定理 5.14 (线性变换核像空间分解)

假设 $\mathcal{A}$ 为线性空间 V 中的线性变换, $\dim V = n$, 则 V 可进行如下分解:

$$V = \text{Ker}(\mathcal{A}^n) \oplus \text{Im}(\mathcal{A}^n)$$

笔记 该结论在后续证明空间分解法中会用到

5.3 线性变换

5.3.1 线性变换基本概念

定义 5.16 (线性变换)

V 是数域 $\mathbb{P}$ 上的线性空间, $\mathcal{A}$ 是 V 到自身的线性映射, 则称 $\mathcal{A}$ 是 V 内的线性变换. V 内全体线性变换组成集合 $\text{End}(V)$.



练习 5.15 线性变换基本练习 (1) 若 $T_1, \dots, T_m$ 是 m 个不同的线性变换, 证明存在 $\alpha \in V$, 使得 $T_1\alpha, \dots, T_m\alpha$ 互不相同

证明 (1) 令 $V_{ij} = \{x \in V : T_i x = T_j x\}$, 则显然 V_{ij} 是 V 的子空间. 因为 $T_1, \dots, T_m$ 两两不同, 故 $\forall T_i, T_j, \exists \beta$ 使得 $T_i \beta \neq T_j \beta$, 从而 $V_{ij} \neq V$. 设 V_{ij} 中有 s 个是零空间, 则其余均为非平凡子空间, 而根据线性空间习题 (和与交一节) 可知存在 α 不属于任一非平凡子空间, 故 $\exists \alpha$ 不属于 $V_{ij}, \forall i, j$, 因此 $T_1\alpha, \dots, T_m\alpha$ 两两不同.

定理 5.15 (线性变换的可逆性)

线性变换 T 可逆当且仅当

- $T\epsilon_1, \dots, T\epsilon_n$ 线性无关
- T 是单射
- T 是满射



证明 (1) 从矩阵角度看即可

(2) 左推右: 若 $T\alpha_1 = T\alpha_2$, 则两侧作用 T^{-1} 得到 $\alpha_1 = \alpha_2$

右推左: 下证 $T\epsilon_1, \dots, T\epsilon_n$ 线性无关, 若 $k_1 T\epsilon_1 + \dots + k_n T\epsilon_n = 0$, 则

$$T(k_1\epsilon_1 + \dots + k_n\epsilon_n) = 0$$

由于 T 是单射, 得到 $k_1\epsilon_1 + \dots + k_n\epsilon_n = 0$, 因此 $k_1 = \dots = k_n = 0$, 根据命题 (1) 知可逆

(3) 左推右: 设 T 可逆, 则 $\forall \alpha \in V$, 存在 $T^{-1}\alpha$, 而 $\alpha = T(T^{-1}\alpha)$, 因此是满射

右推左: 由于满射, 对任一组基 $\eta_1, \dots, \eta_n$, 总存在 $\epsilon_1, \dots, \epsilon_n$, 使得

$$T\epsilon_1 = \eta_1, \dots, T\epsilon_n = \eta_n$$

下面证明 $\epsilon_1, \dots, \epsilon_n$ 线性无关, 设 $k_1\epsilon_1 + \dots + k_n\epsilon_n = 0$, 则

$$T(k_1\epsilon_1 + \dots + k_n\epsilon_n) = k_1\eta_1 + \dots + k_n\eta_n = 0$$

这说明 $k_1 = \dots = k_n = 0$, 因此可逆。

5.3.2 线性变换在基下的矩阵、变换后的坐标

定义 5.17 (线性变换的矩阵)

$\mathcal{A}$ 是 V 内的线性变换, $V = L(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$, 若 $(\mathcal{A}\epsilon_1, \dots, \mathcal{A}\epsilon_n) = (\epsilon_1, \dots, \epsilon_n)A$, 则 A 称为 $\mathcal{A}$ 在 $\epsilon_1, \dots, \epsilon_n$ 基下的矩阵.



笔记 计算线性变换在基下的矩阵: 根据 $(\mathcal{A}\epsilon_1, \dots, \mathcal{A}\epsilon_n) = (\epsilon_1, \dots, \epsilon_n)A$ 推出

$$A = (\epsilon_1, \dots, \epsilon_n)^{-1}(\mathcal{A}\epsilon_1, \dots, \mathcal{A}\epsilon_n)$$

可以用矩阵求逆方式求解。

 **练习 5.16 线性变换在基下的矩阵** 已知 $A = (a_{ij}), i, j = 1, 2, 3$ 是 $\mathcal{A}$ 在 $\epsilon_1, \epsilon_2, \epsilon_3$ 下的矩阵, 计算

- (1) $\mathcal{A}$ 在 $\epsilon_3, \epsilon_2, \epsilon_1$ 下的矩阵
- (2) $\mathcal{A}$ 在 $\epsilon_1, k\epsilon_2, \epsilon_3$ 下的矩阵
- (3) $\mathcal{A}$ 在 $\epsilon_1 + \epsilon_2, \epsilon_2, \epsilon_3$ 下的矩阵

解 根据 $T(\epsilon_1, \epsilon_2, \epsilon_3) = (\epsilon_1, \epsilon_2, \epsilon_3)A$, 得到

$$\begin{cases} T\epsilon_1 = a_{11}\epsilon_1 + a_{21}\epsilon_2 + a_{31}\epsilon_3 \\ T\epsilon_2 = a_{12}\epsilon_1 + a_{22}\epsilon_2 + a_{32}\epsilon_3 \\ T\epsilon_3 = a_{13}\epsilon_1 + a_{23}\epsilon_2 + a_{33}\epsilon_3 \end{cases}$$

(1) 重新排列得到:

$$\begin{cases} T\epsilon_3 = a_{33}\epsilon_3 + a_{23}\epsilon_2 + a_{13}\epsilon_1 \\ T\epsilon_2 = a_{32}\epsilon_3 + a_{22}\epsilon_2 + a_{12}\epsilon_1 \\ T\epsilon_1 = a_{31}\epsilon_3 + a_{21}\epsilon_2 + a_{11}\epsilon_1 \end{cases}$$

得到矩阵为

$$\begin{bmatrix} a_{33} & a_{32} & a_{31} \\ a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{bmatrix}$$

(2)(3) 同理列开即可

 **练习 5.17 幂零变换的矩阵** T 是 n 维线性空间 V 上的线性变换, $T^{n-1} \neq \mathcal{O}, T^n = \mathcal{O}$, 证明: T 在某组基下的矩阵为

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

证明 首先 $\alpha, T\alpha, \dots, T^{n-1}\alpha$ 均线性无关, 因此它们是 V 的一组基, T 在这组基下的矩阵即目标。

定理 5.16 (线性变换后的坐标)

若 α 在 $(\epsilon_1, \dots, \epsilon_n)$ 基下的坐标为 X , 则 $\mathcal{A}\alpha$ 在这组基下的坐标为 AX 。



证明 $\mathcal{A}\alpha = \mathcal{A}(x_1\epsilon_1 + \dots + x_n\epsilon_n) = \mathcal{A}(\epsilon_1, \dots, \epsilon_n)X = (\epsilon_1, \dots, \epsilon_n)AX$,

5.3.3 线性变换在不同基下的矩阵: 相似关系

定理 5.17 (线性变换在不同基下的矩阵)

$\epsilon_i, \eta_i (i = 1, \dots, n)$ 是 V 两组基, 且 $(\eta_1, \dots, \eta_n) = (\epsilon_1, \dots, \epsilon_n)T$ 。若 $\mathcal{A}$ 在 ϵ_i, η_i 基下的矩阵分别为 A, B , 则有 $B = T^{-1}AT$ 。



证明 由线性变换矩阵定义, $\mathcal{A}(\epsilon_1, \dots, \epsilon_n) = (\epsilon_1, \dots, \epsilon_n)A$, $\mathcal{A}(\eta_1, \dots, \eta_n) = (\eta_1, \dots, \eta_n)B$ 。而 $\mathcal{A}(\eta_1, \dots, \eta_n) = \mathcal{A}(\epsilon_1, \dots, \epsilon_n)T = (\epsilon_1, \dots, \epsilon_n)AT$, 且 $\mathcal{A}(\eta_1, \dots, \eta_n) = (\epsilon_1, \dots, \epsilon_n)TB$ 上面两式取等号即可。

 **笔记** 记忆时根据 $(\eta_1, \dots, \eta_n) = (\epsilon_1, \dots, \epsilon_n)T$, 将 η, ϵ 基直接换成对应的矩阵 B, A , 有 T 的一边添上 T^{-1} 即可, 即 $B = T^{-1}AT$

 **练习 5.18 线性变换在不同基下的矩阵** 给定下述两组基, 线性变换 T 满足 $T\epsilon_i = \eta_i$,

$$\begin{cases} \epsilon_1 = (1, 0, 1)^T, \epsilon_2 = (2, 1, 0)^T, \epsilon_3 = (1, 1, 1)^T \\ \eta_1 = (1, 2, -1)^T, \eta_2 = (2, 2, -1)^T, \eta_3 = (2, -1, -1)^T \end{cases}$$

(1) 写出 $\epsilon_1, \epsilon_2, \epsilon_3$ 到 η_1, η_2, η_3 的过渡矩阵

(2) 写出 T 在 $\epsilon_1, \epsilon_2, \epsilon_3$ 下的矩阵

(3) 写出 T 在 η_1, η_2, η_3 下的矩阵

解 (1) e_1, e_2, e_3 下的矩阵为

$$(\eta_1, \eta_2, \eta_3) =$$

定理 5.18 (线性变换构造的等价类)

矩阵 A, B 相似的充要条件为它们是一个线性变换在两组基下的矩阵。



5.3.4 相似矩阵的性质

定理 5.19 (相似矩阵的基本性质)

K 上矩阵 A, B 相似, 且满足 $B = P^{-1}AP$, 则

- $r(A) = r(B)$
- $A^m \sim B^m$
- $f(x)$ 为多项式, 则 $f(A) \sim f(B)$
- 若 $A_1 \sim B_1, A_2 \sim B_2$, 不一定有 $A_1 + A_2 \sim B_1 + B_2$, 要求过渡矩阵相同才行



证明 (1) $r(A) = r(B)$: 因为 $B = P^{-1}AP$, 而 P 满秩, 等价于一系列初等变换, 故不改变秩。

5.4 特征值、相似对角化

5.4.1 特征值与特征向量

定义 5.18 (特征值与特征向量)

$\mathcal{A}$ 是数域 $\mathbb{P}$ 上线性空间 V 内的一个线性变换, 若对于 $\lambda \in \mathbb{P}, \exists \xi \neq 0$, 使得 $\mathcal{A}\xi = \lambda\xi$, 则称 λ 是 $\mathcal{A}$ 的一个特征值, ξ 是 $\mathcal{A}$ 属于 λ 的特征向量。



定义 5.19 (特征子空间)

对于 $\lambda \in \mathbb{P}$, 定义其特征子空间为所有满足 $\mathcal{A}\xi = \lambda\xi$ 的特征向量 ξ 所组成的集合 V_λ 。



引理 5.1 (特征子空间的性质)

特征子空间 V_λ 有如下性质:

- V_λ 是 V 的子空间。
- 对于 $\lambda \in \mathbb{P}, V_\lambda \neq \{0\}$ 的充要条件为 $|\lambda E - A| = 0$, 其中 A 是 $\mathcal{A}$ 在任意基下的矩阵。



证明 (1) 只需要证明 V_λ 对 V 内的数乘和加法封闭即可。

(2) 对于 $\lambda\alpha = \mathcal{A}\alpha$, 可以取 V 一组基, 若 α 在这组基下的坐标为 X 。则等价于求解 $\lambda X = AX$, 其中 A 是 $\mathcal{A}$ 在这组基下的矩阵。等价于求解 $(\lambda E - A)X = 0$, 而线性方程组有非零解的充要条件为系数矩阵行列式非零。

定义 5.20 (特征多项式)

对于 $\mathbb{P}$ 上的矩阵 A , 特征多项式定义为 $f(\lambda) = |\lambda E - A|$ 。

**定理 5.20 (相似矩阵的特征多项式)**

相似矩阵有相同特征多项式。即对于同一个线性变换, 特征多项式是不变的。



证明 若 $B = T^{-1}AT$, 则 $|\lambda E - B| = |\lambda E - T^{-1}AT| = |T^{-1}||\lambda E - A||T| = |\lambda E - A|$ 。

定理 5.21 (特征值基本性质)

$\lambda_1, \dots, \lambda_n$ 是矩阵 A 的特征值。则 $\sum \lambda_i = \text{tr}(A)$, $\prod \lambda_i = |A|$



证明 用行列式展开与韦达定理证明。

练习 5.19 特征值与特征向量的基本练习 (1) λ_1, λ_2 是 T 两个不同的特征值, α_1, α_2 分别是对应的特征向量, 证明: $\alpha_1 + \alpha_2$ 不是 T 的特征向量

(2) T 是可逆线性变换, 若 λ 是 T 的特征值, 则 λ^{-1} 是 T^{-1} 的特征值

证明 (1) $T\alpha_1 = \lambda_1\alpha_1, T\alpha_2 = \lambda_2\alpha_2$, 因此

$$T(\alpha_1 + \alpha_2) = \lambda_1\alpha_1 + \lambda_2\alpha_2$$

若 $\alpha_1 + \alpha_2$ 是特征向量, λ 是特征值, 则 $T(\alpha_1 + \alpha_2) = \lambda(\alpha_1 + \alpha_2)$, 得到

$$(\lambda - \lambda_1)\alpha_1 + (\lambda - \lambda_2)\alpha_2 = 0$$

因此 $\lambda - \lambda_1 = \lambda - \lambda_2 = 0$, 这与假设矛盾。

(2) 由于 $T\alpha = \lambda\alpha$, 因此 $\alpha = \lambda T^{-1}(\alpha)$, 即 $T^{-1}(\alpha) = \lambda^{-1}\alpha$, 因此 λ^{-1} 是 T^{-1} 的特征值

5.4.2 特征值与特征向量进一步结论**定理 5.22 (AB 与 BA 的特征值)**

V, V' 是 K 上的线性空间, $A \in \text{Hom}(V, V'), B \in \text{Hom}(V', V)$ (矩阵同理), 则

- AB 与 BA 有相同非零特征值
- 若 ξ 是 AB 属于 λ_0 的特征向量, 则 $B\xi$ 是 BA 属于 λ_0 的一个特征向量



证明 设 λ_0 是 AB 非零特征值, $\exists \xi \in V'$ 使得 $(AB)\xi = \lambda_0\xi$, 两侧同时作用 B 得到:

$$B(AB)\xi = B(\lambda_0\xi) \Rightarrow (BA)(B\xi) = \lambda_0(B\xi)$$

定理 5.23 (Sylvester 公式)

$A_{m \times n}, B_{n \times m}$, $f_{AB}(\lambda), f_{BA}(\lambda)$ 是 AB, BA 的特征多项式, 则

$$\lambda^n f_{AB}(\lambda) = \lambda^m f_{BA}(\lambda)$$



证明 证明: 设 $\lambda \neq 0$, 根据

$$\begin{bmatrix} E_n & 0 \\ -A & E_m \end{bmatrix} \begin{bmatrix} E_n & \frac{1}{\lambda}B \\ A & E_m \end{bmatrix} = \begin{bmatrix} E_n & \frac{1}{\lambda}B \\ O & E_m - \frac{1}{\lambda}AB \end{bmatrix} \Rightarrow |H| = \begin{vmatrix} E_n & \frac{1}{\lambda}B \\ O & E_m - \frac{1}{\lambda}AB \end{vmatrix} = \left(\frac{1}{\lambda}\right)^m |\lambda E_m - AB|$$

同理根据

$$\begin{bmatrix} E_n & \frac{1}{\lambda}B \\ A & E_m \end{bmatrix} \cdot \begin{bmatrix} E_n & 0 \\ -A & E_m \end{bmatrix} = \begin{bmatrix} E_n - \frac{1}{\lambda}AB & \frac{1}{\lambda}B \\ 0 & E_m \end{bmatrix} \Rightarrow |H| = \left(\frac{1}{\lambda}\right)^n |\lambda E_n - BA|$$

根据 $|H|$ 的值即可推出结论。

 **练习 5.20 相关结论** (1) A, B 是两个 n 阶方阵, 证明: $AB + A$ 和 $BA + A$ 有相同的特征多项式

(2) 计算下面矩阵 C 的特征值

$$C = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix}$$

(3) 特例: 已知 $A = (a_1, \cdots, a_n)$, 证明 $A^T A$ 的特征值为 0 和 $a_1^2 + \cdots + a_n^2$, 这里 0 是 $n-1$ 重根

证明 (1) 由于 $AB + A = A(B + E)$, $BA + A = (B + E)A$, 因此根据前面定理可知。

(2) 令 $A = (a_1, \cdots, a_n)^T, B = (b_1, \cdots, b_n)$, 则 $C = AB$, 而 $BA = \sum_{i=1}^n a_i b_i$ 为一阶矩阵, 其特征值为 $\sum_{i=1}^n a_i b_i$, AB 和 BA 特征多项式相差 λ^{n-1} , 因此 $C = AB$ 的特征多项式为

$$\lambda^{n-1}(\lambda - \sum_{i=1}^n a_i b_i)$$

(3) 即 (2) 的特例

定理 5.24 (可逆的特征值表达)

n 阶方阵 A 可逆当且仅当 A 的特征值全不为 0

证明 用 $|A| = \lambda_1 \cdots \lambda_n$ 即可

定理 5.25 (逆矩阵的特征值)

n 阶可逆矩阵 A 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则 A^{-1} 的全部特征值为 $\lambda_1^{-1}, \cdots, \lambda_n^{-1}$

证明 用 A 相似于 Jordan 阵 J , J^{-1} 为上三角且对角为 $\lambda_1^{-1}, \cdots, \lambda_n^{-1}$, 因此 A^{-1} 的特征值也为 $\lambda_1^{-1}, \cdots, \lambda_n^{-1}$

5.4.3 相似的进一步研究

定理 5.26 (相似的矩阵分解)

A, B 相似当且仅当存在方阵 P, Q , 使得

$$A = PQ, B = QP$$

这里 P, Q 至少有一个可逆

证明 由于 A, B 相似, 故存在可逆 P 使得 $P^{-1}AP = B$, 令 $Q = P^{-1}A$, 则得到 $A = PQ, B = QP$

定理 5.27 (一些相似结论)

A, B 为 n 阶方阵, 则

- AB 与 BA : 若方阵 A 可逆, 则 AB 与 BA 相似
- A^{-1} 与 B^{-1} : 若 A, B 可逆, 则 A^{-1} 与 B^{-1} 相似
- 多项式: 任一多项式 $f(x)$, 有 $f(A), f(B)$ 相似

• 伴随矩阵：若 A, B 相似，则 A^* 与 B^* 相似



证明 (1) $BA = (A^{-1}A)(BA) = A^{-1}(AB)A$

(2) $(P^{-1}AP)^{-1} = B^{-1}$ ，得到 $P^{-1}A^{-1}P = B^{-1}$

(3) 显然

(4) $B^* = (P^{-1}AP)^* = P^*A^*(P^{-1})^* = P^*A^*(P^*)^{-1}$

推论 5.2 (伴随矩阵的特征值)

若矩阵 A 的特征值为 $\lambda_1, \dots, \lambda_m$ ，则 A^* 的特征值为 $\prod_{i \neq j} \lambda_i$



证明 考虑 A 在 $\mathbb{C}$ 上相似于 Jordan 阵 J ，而 J^* 的特征值易算，相似时特征值不变，因此可算出 A^* 的特征值。

5.4.4 相似对角化

定义 5.21 (相似对角化)

$\mathcal{A}$ 是 V 内的线性变换，如果存在 V 的一组基使得 $\mathcal{A}$ 在该基下的矩阵为对角阵，则称 $\mathcal{A}$ 可对角化。



定义 5.22 (代数重数、几何重数)

代数重数指特征值 λ_i 在特征多项式 $f(\lambda)$ 中的重数。几何重数指 λ_i 对应特征子空间的维数。



引理 5.2 (代数重数与几何重数关系)

对每个特征值 λ_i ，其代数重数大于等于几何重数。



引理 5.3 (不同特征值的特征向量)

$\mathcal{A}$ 属于不同特征值的特征向量线性无关。



证明 设 $\lambda_1, \dots, \lambda_s$ 是 A 的 s 个不同的特征值，对应特征向量 $\xi_1, \dots, \xi_s$ 。用归纳法，假设 $\xi_1, \dots, \xi_{k-1}$ 线性无关，若

$$c_1\xi_1 + \dots + c_k\xi_k = 0 \quad (5.1)$$

两侧作用 $\mathcal{A}$ 得到 $c_1\lambda_1\xi_1 + \dots + c_k\lambda_k\xi_k = 0$ ，用该式子减去 λ_k 倍的 (5.1) 得到

$$c_1(\lambda_1 - \lambda_k)\xi_1 + \dots + c_{k-1}(\lambda_{k-1} - \lambda_k)\xi_{k-1} = 0$$

由于 λ_i 两两不相等，因此 $c_1 = \dots = c_{k-1} = 0$ ，即特征向量彼此线性无关。

定理 5.28 (可对角化条件)

若 V 是 n 维线性空间，则 $\mathcal{A}$ 可对角化的条件按照递进关系如下排列：

- 充要条件 1: $\mathcal{A}$ 有 n 个线性无关特征向量。
- 充分条件：若 $\mathcal{A}$ 有 n 个不同特征值，则 $\mathcal{A}$ 可对角化
- 充要条件 2: V 是所有特征子空间的直和 $V = \bigoplus_{i=1}^m V_i$ (特征子空间维数可能没有 λ_i 重数高)
- 充要条件 3: 所有特征值的代数重数等于几何重数
- 充要条件 4: $\dim V = \sum_{i=1}^m \dim(V_i)$
- 充要条件 5(最小多项式): 最小多项式都是一次幂的乘积



证明 (1) 左推右: 由对角化形式, 右侧乘开 n 个基向量均是特征向量, 显然基向量都线性无关。

$$(A\eta_1, A\eta_2, \dots, A\eta_n) = (\eta_1, \eta_2, \dots, \eta_n) \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix}$$

右推左: n 个线性无关特征向量当一组基即可。

(2) 根据引理5.3可知

后面三个都根据第一个可以推出

 **练习 5.21 对角化计算** (1) 计算 $A = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix}$ 的特征值, 并计算 Q 使得 $Q^{-1}AQ$ 是对角矩阵

定理 5.29 (相关矩阵的可对角化性质)

若矩阵 A 可对角化, 则

1. A^* 可对角化
2. $f(x)$ 为多项式, 则 $f(A)$ 可对角化
3. 若 A 可逆, 则 A^{-1} 可对角化



证明 (1) 设 $P^{-1}AP = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, 两侧取伴随得到 $P^*A^*(P^*)^{-1} = P^{-1}A^*P = \text{diag}\{\prod_{i \neq 1} \lambda_i, \prod_{i \neq 2} \lambda_i, \dots, \prod_{i \neq n} \lambda_i\}$ 。

其中伴随转逆用了 $A^{-1} = \frac{1}{|A|}A^*$

(2,3) 显然

5.4.5 相似上三角型

定理 5.30 (相似上三角型)

任意 n 阶矩阵 A 都相似于一个上三角矩阵, 且上三角对角为 A 的特征值。



证明 $n=1$ 时显然成立。假设 $n-1$ 时成立, 则取 x_1 为 A 对应 λ_1 的特征向量, 构造矩阵 $P_1 = [x_1, \dots, x_n]$ ($x_2 \sim x_n$ 无要求), 可以得到:

$$AP_1 = (x_1, \dots, x_n) \begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n1} & \cdots & b_{nn} \end{bmatrix} \Rightarrow P_1^{-1}AP_1 = \begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n1} & \cdots & b_{nn} \end{bmatrix} = \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix}$$

根据归纳假设可知对 A_1 有可逆矩阵 Q 使得 $Q^{-1}A_1Q$ 为上三角且对角为特征值, 取 $P_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & Q \end{bmatrix}$, 取 $S = P_1P_2$, $S^{-1}AS$ 即所求上三角

5.5 最小多项式

5.5.1 最小多项式与基本性质

定义 5.23 (零化多项式与最小多项式)

A 是数域 K 上的 n 阶方阵, $g(x)$ 是 K 上的多项式,

- 零化多项式: 若 $g(A) = 0$, 则称 $g(x)$ 是 A 的一个零化多项式。
- 最小多项式: 将 A 的次数最低首一零化多项式称为 A 的最小多项式



定理 5.31 (最小多项式的性质)

最小多项式有以下性质:

- 唯一性: 最小多项式是唯一的
- 整除性: 最小多项式整除任意零化多项式
- 分块对角: $A = \text{diag}\{A_1, \dots, A_s\}$, 则 $m_A(\lambda) = [m_1(\lambda), \dots, m_s(\lambda)]$ (最小公倍数)
- 相似: 相似矩阵有相同最小多项式



证明 (1) 假设有两个不同最小多项式 $m_1(\lambda), m_2(\lambda)$, 则令 $h(\lambda) = m_1(\lambda) - m_2(\lambda)$, 有 $h(A) = m_1(A) - m_2(A) = 0$, 显然 $h(\lambda)$ 是最小多项式, 矛盾。

(2) 对任意零化多项式 $f(x)$, 其可以表示为 $f(x) = q(x)m(x) + r(x)$, 代入 $x = A$ 得到 $r(A) = 0$, 其只能为 $r(x) = 0$, 否则与最小多项式矛盾, 因此整除。

(3) 由于 $m(A) = \text{diag}\{m(A_1), \dots, m(A_s)\}$, 最大公因式即可满足所有分块均为 0。

(4) 若 $B = P^{-1}AP$, 则 $f(B) = f(P^{-1}AP) = P^{-1}f(A)P$, 因此可知最小多项式相同。

5.5.2 Hamilton-Cayley 定理

定理 5.32 (Hamilton-Cayley 定理)

A 是 K 上的 n 阶矩阵, $f(\lambda) = |\lambda E - A|$ 即 A 的特征多项式, 则 $f(x)$ 是 A 的零化多项式。



证明 设 $B(\lambda)$ 是 $\lambda E - A$ 的伴随矩阵, 则 $B(\lambda)(\lambda E - A) = f(\lambda)E$, 由于 $B(\lambda)$ 元素含 $1 \sim \lambda^{n-1}$, 因此可以写为 $B(\lambda) = \lambda^{n-1}B_0 + \dots + B_{n-1}$, 从而得到:

$$B(\lambda)(\lambda E - A) = \lambda^n B_0 + \lambda^{n-1}(B_1 - B_0 A) + \dots + \lambda(B_{n-1} - B_{n-2} A) - B_{n-1} A$$

根据 $B(\lambda)(\lambda E - A) = f(\lambda)E = \lambda^n E + a_1 \lambda^{n-1} E + \dots + a_{n-1} \lambda E + a_n E$ 进行比较得到:

$$\begin{cases} B_0 = E \\ B_1 - B_0 A = a_1 E \\ \vdots \\ B_{n-1} - B_{n-2} A = a_{n-1} E \\ -B_{n-1} A = a_n E \end{cases} \Rightarrow \begin{cases} B_0 A^n = A^n \\ B_1 A^{n-1} - B_0 A^n = a_1 A^{n-1} \\ \vdots \\ B_{n-1} A - B_{n-2} A^2 = a_{n-2} A \\ -B_{n-1} A = a_n E \end{cases}$$

两端分别相加, 左侧为 0 , 因此右侧为 0

证明 使用 Jordan 阵证明: 分块看, 显然每个 λ_i 对应的块的阶数小于 λ_i 的代数重数, 每个分块都是零, 因此是零化多项式。

推论 5.3 (Hamilton-Cayley 定理推论)

Hamilton-Cayley 定理有以下推论：

- 最小多项式与特征多项式同根 (重数可以不同)。

**5.5.3 最小多项式与相似对角化****定理 5.33 (最小多项式与相似对角化)**

$\mathbb{P}$ 上的 n 阶矩阵 A 可对角化当且仅当 A 的最小多项式在 $\mathbb{P}$ 上可分解为互素一次因子乘积。



练习 5.22 最小多项式判断可对角化 (1) 幂等矩阵一定可对角化

(2) 对合矩阵 $A^2 = E$ 一定可对角化

(3) 幂零矩阵 $A^l = O$ 不一定可对角化

证明 (1) $A(A - E) = O$ ，一定为一次因子

(2) $(A + E)(A - E) = O$ ，一定可对角化

(3) 例如 Jordan 块

第 6 章 Jordan 标准型

6.1 Jordan 标准形：空间分解法

6.1.1 空间第二分解定理：循环不变子空间分解

定义 6.1 (幂零线性变换)

$\mathcal{A}$ 是 V 内的线性变换, 若 $\exists m$ 使得 $\mathcal{A}^m = 0$, 则 $\mathcal{A}$ 是幂零线性变换。

引理 6.1 (幂零基)

$\mathcal{A}$ 是幂零变换, $\forall \alpha \in V$ 若 $\mathcal{A}^{k-1}\alpha \neq 0, \mathcal{A}^k\alpha = 0$, 则 $\alpha, \mathcal{A}\alpha, \dots, \mathcal{A}^{k-1}\alpha$ 线性无关。

证明 设 $a_0\alpha + a_1\mathcal{A}\alpha + \dots + a_{k-1}\mathcal{A}^{k-1}\alpha = 0$, 若从左至右第一个非零的是 a_i , 即 $a_i\mathcal{A}^i\alpha + \dots + a_k\mathcal{A}^{k-1}\alpha = 0$, 两侧同时作用 $\mathcal{A}^{k-1-i}$, 上式变为 $a_i\mathcal{A}^{k-1}\alpha = 0$, 而 $\mathcal{A}^{k-1}\alpha \neq 0$, 故 $a_i = 0$, 与假设矛盾。

定义 6.2 (循环不变子空间)

对 $\forall \alpha \in V$, $\exists k$ 使得 $\mathcal{A}^k = 0, \mathcal{A}^{k-1} \neq 0$, 则 $I(\alpha) := L(\alpha, \mathcal{A}\alpha, \dots, \mathcal{A}^{k-1}\alpha)$ 称为 α 生成的 $\mathcal{A}$ 的循环不变子空间。

定理 6.1 (循环不变子空间分解)

对于 V 内的幂零变换 $\mathcal{A}$, $\exists v_1, \dots, v_n \in V, m_1, \dots, m_n \in \mathbb{N}$ 使得 $V = I(v_1) \oplus I(v_2) \oplus \dots \oplus I(v_n)$ 。其中 $I(v_i) = L(v_i, \mathcal{A}v_i, \dots, \mathcal{A}^{m_i}v_i)$, 且 $\mathcal{A}^{m_i+1}v_i = 0$ 。

证明 采用数归与核像分解。(1) 若 $\dim V = 1$, 则显然可分解。(2) 若 $\dim V > 1$, 则用归纳法。由于 $\mathcal{A}$ 不是单射 (因为将不同的向量均映射到 0), 因此 $\mathcal{A}$ 不是满射 (因为线性映射是满射当且仅当是双射), 因此 $\mathcal{A}$ 的像维数低于 $\dim V$ 。对 $\text{Im } \mathcal{A}$ 用归纳假设, 得到 $\text{Im } \mathcal{A}$ 的一组基 $\mathcal{A}^{m_1}v_1, \dots, \mathcal{A}v_1, v_1, \dots, \mathcal{A}^{m_n}v_n, \dots, v_n$, 由于上述向量组是 $\text{Im } \mathcal{A}$ 的一组基, 对 $\forall j$ 存在 V 中向量 u_j 使得 $v_j = \mathcal{A}u_j$ 。

下面说明下面的向量组线性无关:

$$\mathcal{A}^{m_1+1}u_1, \dots, \mathcal{A}u_1, u_1, \dots, \mathcal{A}^{m_n+1}u_n, \dots, u_n \quad (6.1)$$

原因: 可设上述向量的一个线性组合为 0, 两侧同时作用 $\mathcal{A}$, 则仅有 $\mathcal{A}^{m_1+1}u_1, \dots, \mathcal{A}^{m_n+1}u_n$ 的系数可以非零 (因为它们被作用后为 0)。去除其他项后, 相当于设上面这组向量的某个线性组合等于 0, 由于它们也是 $\text{Im } \mathcal{A}$ 基的一部分, 彼此线性无关, 所以它们的系数也必须为零。

于是将 (6.1) 扩成 V 的一组基 $\mathcal{A}^{m_1+1}u_1, \dots, u_1, \dots, \mathcal{A}^{m_n+1}u_n, \dots, u_n, w_1, \dots, w_p$, 由于 $\mathcal{A}w_j \in \text{Im } \mathcal{A}$, 因此存在 (6.1) 张成空间中元素 x_j , 使得 $\mathcal{A}w_j = \mathcal{A}x_j$ 。令 $u_{n+j} = w_j - x_j$, 则有 $\mathcal{A}u_{n+j} = 0$, 从而 (6.1) 配上 $u_{n+1}, \dots, u_{n+p}$ 是所需 V 的一组基。

6.1.2 核像升链与根子空间

设 $\mathcal{A}$ 是 V 上的线性变换, 定义 $M_0 = \{0\}, M_i = \text{Ker } \mathcal{A}^i, N_0 = V, N_i = \text{Im } (\mathcal{A}^i)$, 下面来研究它们的性质。

引理 6.2 (核像空间升链)

上述空间有: $\{0\} = M_0 \subseteq M_1 \subseteq M_2 \subseteq \dots, V = N_0 \supseteq N_1 \supseteq N_2 \supseteq \dots$

证明 (1) 对 $\forall \alpha \in M_i$, 由定义有 $\mathcal{A}^i\alpha = 0$, 显然 $\mathcal{A}^{i+1}\alpha = 0$, 因此 $M_i \subseteq M_{i+1}$ 。

(2) 对于 $\alpha \in N_i$, 由定义 $\exists \beta \in V$ 使得 $\alpha = \mathcal{A}^i \beta$, 从而 $\alpha = \mathcal{A}^{i-1}(\mathcal{A}\beta) \in N_{i-1}$ 。由 α 任意性可知 $N_i \subseteq N_{i-1}$ 。

引理 6.3 (核空间增长)

核空间 M_i 的增长满足以下性质:

- 若 $M_k = M_{k+1}$, 则 $M_k = M_{k+1} = M_{k+2} = \cdots$
- 若 $\dim V = n$, 则 $M_n = M_{n+1} = M_{n+2} = \cdots$ 。



证明 (1) 显然有 $M_{k+1} \subseteq M_{k+2}$, 只需要证明 $M_{k+2} \subseteq M_{k+1}$ 。对于 $\forall \alpha \in M_{k+2}$, 有 $\mathcal{A}^{k+2}\alpha = 0$, 即 $\mathcal{A}^{k+1}(\mathcal{A}\alpha) = 0$, 因此 $\mathcal{A}\alpha \in M_{k+1} = M_k$, 即 $\mathcal{A}^{k+1}\alpha = 0 \rightarrow \alpha \in M_{k+1}$, 由任意性可知 $M_{k+2} \subseteq M_{k+1}$, 故 $M_{k+1} = M_{k+2}$, 后续类似。

(2) 只需要证明 $M_n = M_{n+1}$, 假设不然, 则 $\{0\} = M_0 \subset M_1 \subset \cdots \subset M_n \subset M_{n+1}$, $\dim M_{n+1} > n$, 与 $M_{n+1} \subseteq V$ 矛盾。

定理 6.2 (幂次变换核像分解)

$V = M_n \oplus N_n$, 其中 $n = \dim V$ 。



证明 显然由核像性质有 $\dim M_n + \dim N_n = n$, 只需证 $M_n \cap N_n = \{0\}$ 。 $\forall v \in M_n \cap N_n$, 有 $\mathcal{A}^n v = 0$, 且 $\exists u \in V$ 使得 $v = \mathcal{A}^n u$, 从而 $\mathcal{A}^{2n} u = 0$, 而 $M_{2n} = M_{2n-1} = \cdots = M_n$, 故 $\mathcal{A}^n u = 0$, 从而 $v = 0$ 。

定义 6.3 (广义特征向量与根子空间 (广义特征子空间))

- 广义特征向量: 线性变换 $\mathcal{A}$ 关于特征值 λ 的广义特征向量定义为满足 $\exists j \in \mathbb{N}^+, (\mathcal{A} - \lambda E)^j v = 0$ 的所有向量。
- 根子空间: 线性变换 $\mathcal{A}$ 关于 λ 的所有广义特征向量加上零向量组成的空间称为 $\mathcal{A}$ 关于特征值 λ 的根子空间, 记作 $G(\lambda, \mathcal{A})$ 。



定理 6.3 (根子空间的具体表达)

根子空间 $G(\lambda, \mathcal{A})$ 满足 $G(\lambda, \mathcal{A}) = \text{Ker}(\mathcal{A} - \lambda E)^{\dim V}$ 。



证明 右推左: 若 $v \in \text{Ker}(\mathcal{A} - \lambda E)^{\dim V}$, 则由 $G(\lambda, \mathcal{A})$ 定义可知 $v \in G(\lambda, \mathcal{A})$ 。左推右: $v \in G(\lambda, \mathcal{A})$ 即 $\exists j \in \mathbb{N}$ s.t. $(\mathcal{A} - \lambda E)^j v = 0$, 而前面已证明这样的 v 必然满足 $v \in \text{Ker}(\mathcal{A} - \lambda E)^{\dim V}$ 。



笔记 为了知识体系的完整, 这里补充一下: λ_i 根子空间的次数最低压缩为 λ_i 对应 Jordan 块的最高阶数, 即最小多项式中的 λ_i 的阶数。其他书也常用特征多项式的代数重数表示 (题目中会遇到), 上面引理的 $\dim V$ 明显放缩过大。

6.1.3 空间第一分解定理: 根子空间分解

定理 6.4 (根子空间的不变性)

根子空间 $G(\lambda, \mathcal{A})$ 是 $\mathcal{A}$ 不变的。



证明 由定理 6.3 可知 $G(\lambda, \mathcal{A}) = \text{Ker}(\mathcal{A} - \lambda E)^n$, $\forall \alpha \in G(\lambda, \mathcal{A})$ 有 $(\mathcal{A} - \lambda E)^n \alpha = 0$, 显然 $(\mathcal{A} - \lambda E)^n \mathcal{A}\alpha = \mathcal{A}((\mathcal{A} - \lambda E)^n \alpha) = 0$, 故 $\mathcal{A}\alpha \in G(\lambda, \mathcal{A})$ 。

定理 6.5 (空间第一分解定理: 根子空间分解)

若 V 是复线性空间, $\mathcal{A}$ 是线性变换, $\lambda_1, \dots, \lambda_m$ 是其特征值, 则 $V = G(\lambda_1, \mathcal{A}) \oplus \cdots \oplus G(\lambda_m, \mathcal{A})$ 。



证明 用归纳法, (1) $n = 1$ 时显然成立 (2) 假设对所有小于 n 维空间均成立, 由定理 6.2 可知 $V = G(\lambda_1, \mathcal{A}) \oplus \text{Im}(\mathcal{A} - \lambda_1 E)^n$, 若 $U = \text{Im}(\mathcal{A} - \lambda_1 E)^n$ 为空, 则显然成立。其他情况首先显然 $\mathcal{A}$ 在 U 上是不变的 ($\mathcal{A}$ 与 $(\mathcal{A} - \lambda_1 E)$ 可交换),

去除 λ_1 所有对应特征向量后, $\mathcal{A}|_U$ 在 U 有 $\lambda_2, \dots, \lambda_m$ 特征值, 对 U 用归纳法, $U = G(\lambda_2, \mathcal{A}|_U) \oplus \dots \oplus G(\lambda_m, \mathcal{A}|_U)$, 由于 $\forall j \geq 2, G(\lambda_j, \mathcal{A}|_U) = G(\lambda_j, \mathcal{A})$, 从而得证。

6.1.4 Jordan 标准形：空间分解法

定理 6.6 (幂零变换的 Jordan 标准型)

幂零变换 $\mathcal{A}$ 在循环不变子空间分解基下的矩阵如下。称为幂零变换的 Jordan 标准形。

$$J = \begin{bmatrix} J_1 & & 0 \\ & J_2 & \\ 0 & & \ddots \\ & & & J_s \end{bmatrix}, \quad J_i = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}_{n_i \times n_i}$$

证明 代入计算显然可得。

定理 6.7 (核限制变换是幂零变换)

对于 $\forall$ 线性变换 $\mathcal{A}$, 令 $\mathcal{B} = \mathcal{A} - \lambda E$, 则 $\mathcal{B}$ 在 $M_k = \text{Ker} \mathcal{B}^k$ 中的限制变换 $\mathcal{B}|_{M_k}$ 是幂零变换。

证明 由于 $M_k = \text{Ker} \mathcal{B}^k$, 因此 $\forall \alpha \in M_k$ 有 $\mathcal{B}^k \alpha = 0$, 故 $(\mathcal{B}|_{M_k})^k = 0$ 。

定理 6.8 (Jordan 标准型：空间分解法)

对于 $\mathbb{C}$ 上线性变换 $\mathcal{A}$ 而言, 其有在先根子空间分解, 根子空间再循环不变子空间分解的基下有以下形式。称为线性变换的 Jordan 标准形。

$$J = \begin{bmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_s \end{bmatrix}, \quad J_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}_{n_i \times n_i}$$

证明 先由空间分解第一定理将 V 分解为根子空间直和, 由于 $\mathcal{A} - \lambda E$ 在 $G(\lambda, \mathcal{A})$ 中的限制变换是幂零变换, 因此再对 $G(\lambda, \mathcal{A})$ 进行循环不变子空间分解, 再由幂零变换的 Jordan 标准形可证。

6.2 空间分解法：零化多项式角度

6.2.1 根子空间分解：矩阵角度

定理 6.9 (解空间分解)

A 是 n 阶方阵, $f(x), f_1(x), f_2(x) \in P[x]$, 且 $f(x) = f_1(x)f_2(x), (f_1(x), f_2(x)) = 1$, 设 $f(A)X = 0, f_1(A)X = 0, f_2(A)X = 0$ 的解空间分别为 V, V_1, V_2 , 则

$$V = V_1 \oplus V_2$$

证明 由于互素, 得到 $\exists u(x), v(x)$, 使得 $E = u(A)f_1(A) + v(A)f_2(A)$, 从而 $\forall \alpha \in V$ 有

$$\alpha = u(A)f_1(A)\alpha + v(A)f_2(A)\alpha := \alpha_1 + \alpha_2$$

且满足 $f(A)\alpha = f_1(A)f_2(A)\alpha = 0$ 。故考察 $f_2(A)\alpha_1$ ：

$$f_2(A)\alpha_1 = f_2(A)u(A)f_1(A)\alpha = u(A)f_1(A)f_2(A)\alpha = 0 \Rightarrow \alpha_1 \in V_2$$

同理得到 $f_1(A)\alpha_2 = 0, \alpha_2 \in V_1$ ，从而得到 $V = V_1 + V_2$ 。

下证直和， $\forall \alpha \in V_1 \cap V_2$ ，则 $f_1(A)\alpha = f_2(A)\alpha = 0$ ，因此

$$\alpha = u(A)f_1(A)\alpha + v(A)f_2(A)\alpha = 0$$

因此结论成立。

定理 6.10 (多空间分解)

A 是数域 P 上 n 阶方阵， $f(x), f_1(x), \dots, f_s(x) \in P[x]$ ， $f(x) = f_1(x)f_2(x) \cdots f_s(x)$ ，且 $f_1(x), \dots, f_s(x)$ 两两互素，设 $f(A)X = 0, f_1(A)X = 0, \dots, f_s(A)X = 0$ 的解空间为 $V, V_1, \dots, V_s$ ，则

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$



证明 用数学归纳法即可。

定理 6.11 (空间分解研究矩阵秩)

考虑 $f(A) = f_1(A) \cdots f_s(A)$ ，则

$$n - r(f(A)) = n - r(f_1(A)) + \cdots + n - r(f_s(A))$$



证明 设 $f(A)X = 0, f_1(A)X = 0, \dots, f_s(A)X = 0$ 的解空间为 $V, V_1, \dots, V_s$ ，根据解空间分解得到 $V = V_1 \oplus \cdots \oplus V_s$ ，因此根据解空间的维数可得到结论。

练习 6.1 空间分解研究矩阵的秩 证明以下命题

(1) $A^2 = A$ 当且仅当 $r(A) + r(A - E) = n$

(2) $A^2 = E$ 当且仅当 $r(A + E) + r(A - E) = n$

(3) $A^3 = E$ 当且仅当 $r(A - E) + r(A^2 + A + E) = n$

证明 (1) $f(A) = A^2 - A = A(A - E)$ ，因此 $n - r(A^2 - A) = n - r(A) + n - r(A - E)$ ，左侧等于 n ，右侧等于 $2n - r(A) - r(A - E)$ ，因此 $r(A) + r(A - E) = n$

6.2.2 根子空间分解：线性变换角度

引理 6.4 (零变换的核空间)

V 是 $\mathbb{P}$ 上的线性空间， $\mathcal{O}$ 是零变换，则其核空间为整个空间：

$$\text{Ker}(\mathcal{O}) = V$$



定理 6.12 (核空间分解)

$\mathcal{A}$ 为线性变换， $f(x) = f_1(x)f_2(x)$ 为多项式，且 $(f_1(x), f_2(x)) = 1$ ，则有如下核空间分解：

$$\text{Ker}(f(\mathcal{A})) = \text{Ker}(f_1(\mathcal{A})) \oplus \text{Ker}(f_2(\mathcal{A}))$$



证明 根据 $(f_1(x), f_2(x)) = 1$ 得到 $\exists u(x), v(x)$ 使得 $u(x)f_1(x) + v(x)f_2(x) = 1$ ，因此

$$\mathcal{E} = u(\mathcal{A})f_1(\mathcal{A}) + v(\mathcal{A})f_2(\mathcal{A})$$

对 $\forall \alpha \in \text{Ker}(f(\mathcal{A}))$

$$\alpha = \alpha_1 + \alpha_2 := u(\mathcal{A})f_1(\mathcal{A})\alpha + v(\mathcal{A})f_2(\mathcal{A})\alpha$$

显然 $f_2(\mathcal{A})\alpha_1 = u(\mathcal{A})f(\mathcal{A})\alpha = 0$, 得到 $\alpha_1 \in \text{Ker}(f_1(\mathcal{A}))$, 同理有 $\alpha_2 \in \text{Ker}(f_2(\mathcal{A}))$, 综上得到

$$\text{Ker}(f(\mathcal{A})) = \text{Ker}(f_1(\mathcal{A})) + \text{Ker}(f_2(\mathcal{A}))$$

下证直和：对 $\forall \alpha \in \text{Ker}(f_1(\mathcal{A})) \cap \text{Ker}(f_2(\mathcal{A}))$, 有

$$\alpha = u(\mathcal{A})f_1(\mathcal{A})\alpha + v(\mathcal{A})f_2(\mathcal{A})\alpha = 0$$

因此为直和

定理 6.13 (全空间分解：零化多项式角度)

V 是 $\mathbb{P}$ 上的线性空间, $\mathcal{A}$ 是 V 上的线性变换, $f(x) = f_1(x)f_2(x)\cdots f_s(x)$ 为零化多项式, 且 $f_1(x), \dots, f_s(x)$ 两两互素, 则

$$V = \text{Ker}(f_1(\mathcal{A})) \oplus \cdots \oplus \text{Ker}(f_s(\mathcal{A}))$$



证明 由于 $f(\mathcal{A}) = \mathcal{O}$, 根据零变换的核空间以及核空间分解定理可知。

引理 6.5 (Hamilton-Cayley)

若 $f(\lambda)$ 是线性变换 $\mathcal{A}$ 的特征多项式, 则 $f(\lambda)$ 是 $\mathcal{A}$ 的零化多项式



证明 该定理在最小多项式一章已经证明过了

定理 6.14 (根子空间分解：零化多项式角度)

若 $\mathbb{C}$ 上线性变换 $\mathcal{A}$ 的特征多项式为 $f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_s)^{k_s}$, 最小多项式为 $m(\lambda) = (\lambda - \lambda_1)^{l_1} \cdots (\lambda - \lambda_s)^{l_s}$, 则有以下根子空间分解

$$V = \text{Ker}(\mathcal{A} - \lambda_1 \mathcal{E})^{k_1} \oplus \cdots \oplus \text{Ker}(\mathcal{A} - \lambda_s \mathcal{E})^{k_s} = \text{Ker}(\mathcal{A} - \lambda_1 \mathcal{E})^{l_1} \oplus \cdots \oplus \text{Ker}(\mathcal{A} - \lambda_s \mathcal{E})^{l_s}$$



证明 根据 Hamilton-Cayley 定理以及零化多项式全空间分解可知。

练习 6.2 根子空间分解：零化多项式角度 (1)ZJU2021.8: V 是线性空间, φ 是 V 上线性变换, 特征多项式 $f(\lambda) = (\lambda - 2)^6(\lambda + 2)^4$, 将 V 分解为两个非平凡不变子空间的直和, 并证明结论。

定理 6.15 (根子空间分解与 Jordan 标准型)

根子空间分解将 Jordan 阵按照特征值 λ_i 分为若干份, 每份由 λ_i 为对角的若干个 Jordan 块组成, 对角 λ_i 在 Jordan 阵中的出现次数等于其代数重数。



6.2.3 Jordan 阵与循环不变子空间

引理 6.6 (核限制变换是幂零变换)

λ_i 为 $\mathcal{A}$ 的特征值, 则 $(\mathcal{A} - \lambda_i \mathcal{E})|_{\text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k}$ 在 $\text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k$ 中为幂零变换



证明 由于 $\forall \alpha \in \text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k$ 满足 $(\mathcal{A} - \lambda_i \mathcal{E})^k \alpha = 0$, 因此限制变换满足 $(\mathcal{A} - \lambda_i \mathcal{E})^k|_{\text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k} = \mathcal{O}$

定理 6.16 (循环不变子空间分解)

对于 V 内的幂零变换 $\mathcal{A}$, $\exists v_1, \dots, v_n \in V$, $m_1, \dots, m_n \in \mathbb{N}$ 使得 $V = I(v_1) \oplus I(v_2) \oplus \cdots \oplus I(v_n)$ 。其中 $I(v_i) = L(v_i, \mathcal{A}v_i, \dots, \mathcal{A}^{m_i-1}v_i)$, 且 $\mathcal{A}^{m_i+1}v_i = 0$ 。



证明 在空间第二分解定理中已经证明过了。

定理 6.17 (Jordan 阵与循环不变子空间)

线性变换 $\mathcal{A}$ 在根子空间 $\text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k$ 下的限制变换 $\mathcal{A}|_{\text{Ker}(\mathcal{A} - \lambda_i \mathcal{E})^k}$ 在一组循环不变基 $((\mathcal{A} - \lambda_i \mathcal{E})^s \alpha, (\mathcal{A} - \lambda_i \mathcal{E})^{s-1} \alpha, \dots, \alpha)$ 下的矩阵为一个 Jordan 块:

$$J_s(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}$$



证明 直接验证即可。

定理 6.18 (Jordan 阵与几何重数)

特征值 λ_i 的代数重数是特征多项式 $f(\lambda)$ 的 $\lambda = \lambda_i$ 根的重数, 几何重数 ($\lambda_i E - A = 0$ 解空间维数) 是 λ_i 对应以 λ_i 为对角的 Jordan 块个数。



证明 根据 $\lambda_i E - A = 0$ 的解空间维数等于 $\lambda_i E - J = 0$ 解空间维数, 而每有一个以 λ_i 为对角的 Jordan 块, $\lambda_i E - J$ 上就有一行全为 0, 因此秩减 1, 从而块数等于几何重数。

笔记 即以 λ_i 为对角的 Jordan 块个数为 $n - r(\lambda E - A)$, 反过来写也要分得清。

6.3 Jordan 标准形: λ 矩阵法

6.3.1 λ 矩阵与相抵

定义 6.4 (λ 矩阵)

$A(\lambda)$ 是元素为 λ 多项式的矩阵, 则将 $A(\lambda)$ 称为 λ 矩阵。

**定义 6.5 (λ 矩阵初等变换)**

λ 矩阵有三种初等变换:

- 互换两行/列
- 用常数 c 乘一行/列 (不能乘 λ 多项式!)
- 将一行/列乘以一个 λ 多项式加到另一行/列上

**定义 6.6 (相抵)**

若可以用初等变换将一个 λ 矩阵转换为另一个 λ 矩阵, 则称两个 λ 矩阵相抵。

**定理 6.19 (相抵标准形: 法式)**

对于数值矩阵 A , $\lambda E - A$ 总是相抵于 $\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_m(\lambda)\}$, 其中 $d_i(\lambda)$ 是首一多项式, 且 $d_i(\lambda) | d_{i+1}(\lambda)$ 。

**引理 6.7 (对角整除化)**

$A(\lambda)$ 为非零 λ 矩阵, 则 $A(\lambda)$ 一定相抵于 $B(\lambda)$, 满足 $b_{11}(\lambda) \neq 0$ 且 $b_{11}(\lambda)$ 可整除 $B(\lambda)$ 中的所有元素



证明 先将 $A(\lambda)$ 中次数最低的元素放置于对角 $a_{11}(\lambda)$

(1) 消去第一行、第一列: 不妨设 $a_{i1}(\lambda)$ 不被 $a_{11}(\lambda)$ 整除, 则做带余除法 $a_{i1}(\lambda) = q(\lambda)a_{11}(\lambda) + r(\lambda)$, 再将 $r(\lambda)$ 放置于 a_{11} , 直至全部整除, 可消去第一列和第一行。

(2) 考虑第一列、第一行外的其他元素, 若存在 $a_{ij}(\lambda)$ 不被 a_{11} 整除, 则把第 j 列加到第一列上, 做带余除法, 把余项放到 a_{11} , 以此类推。

定理 6.20 (法式的计算: 初等变换法)

每次把最低次移动到角上, 尝试消除, 消除过程按照以下步骤

- 通过初等变换将当前对角元素变为所有非零元素中次数最低的 λ 多项式
- 将矩阵对角整除化: 具体见引理的过程
- 消去一列, 再消去一行, 重复对角整除化和消去

定理 6.21 (相抵与相似)

A 与 B 相似当且仅当 $\lambda E - A$ 与 $\lambda E - B$ 相抵。

证明 (1) 左推右: 若 $B = P^{-1}AP$, 则 $\lambda E - B = P^{-1}(\lambda E - A)P$, 显然右侧两个 λ 矩阵相抵。

(2) 右推左: 若 $\lambda E - B = P(\lambda)(\lambda E - A)Q(\lambda)$, 设 $Q(\lambda) = Q_k\lambda^k + \cdots + Q_1\lambda + Q_0$, $Q(\lambda)^{-1} = R_m\lambda^m + \cdots + R_1\lambda + R_0$, 下记 $W = Q(B)$ 。由 $Q(\lambda)Q^{-1}(\lambda) = E$ 得 $R_mQ(\lambda)\lambda^m + \cdots + R_1Q(\lambda)\lambda + R_0Q(\lambda) = E$, 代入 B 得到 $R_mWB^m + \cdots + R_1WB + R_0W = E$ 。

由于 $P(\lambda)^{-1}(\lambda E - B) = (\lambda E - A)Q(\lambda) = Q(\lambda)\lambda - AQ(\lambda)$, 代入 $\lambda = B$ 得到 $Q(B)B = AQ(B)$ (前面一式子左 $=0$), 即 $WB = AW$, 从而 $WB^l = A^lW$ (多作用一个 A 相当于多右乘一个 B)。

再由前面的式子得到 $(R_mA^m + \cdots + R_0)W = E$, 则表明 W 可逆, 从而 $B = W^{-1}AW$ 。

6.3.2 行列式因子、不变因子

定义 6.7 (行列式 k 阶子式)

$|A|$ 的 k 阶子式是其任选 k 行、 k 列组成的矩阵的行列式。

定义 6.8 (行列式因子)

$A(\lambda)$ 是 λ 矩阵, 若其所有 k 阶子式 (行列式) 均为 0, 则定义其 k 阶行列式因子 $D_k(\lambda) = 0$, 否则定义为所有非零 k 阶子式的首一最大公因式。

笔记 一般行列式因子比初等变换方便, 不过一定注意这里的子式是所有子式, 而非主子式, 甚至可以不连着, 0 的情况也得分开考虑。

引理 6.8 (行列式因子的相除性)

设 $D_1(\lambda), \dots, D_r(\lambda)$ 是 $A(\lambda)$ 非零行列式因子, 则 $D_i(\lambda) | D_{i+1}(\lambda)$ 。

证明 设 A_{i+1} 是 $A(\lambda)$ 任意 $i+1$ 阶子式, 这个行列式按某一行展开, 则其每一展开项都是一个多项式与一个 i 阶子式乘积, 而 $D_i(\lambda)$ 是所有 i 阶子式公因子, 因此 $D_i(\lambda) | A_{i+1}(\forall i)$, 而 $D_{i+1}(\lambda)$ 是所有 A_{i+1} 的最大公因式, 因此证毕。

定义 6.9 (不变因子)

定义 $d_1(\lambda) = D_1(\lambda)$, $d_i(\lambda) = \frac{D_i(\lambda)}{D_{i-1}(\lambda)} (\forall i \neq 1)$ 为 $A(\lambda)$ 的不变因子。

定理 6.22 (行列式因子、不变因子的性质)

- 相抵矩阵的因子：相抵的 λ 矩阵有相同的行列式因子、不变因子 (初等变换不改变行列式、不变因子)
- 法式与不变因子： $\lambda E - A$ 法式的所有对角元素即 $\lambda E - A$ 的不变因子。



证明 (1) 只需证明行列式因子在三种初等变换不变即可。而不变因子由行列式因子唯一确定。

(2) 由于不变因子在初等变化中不变，求出法式的不变因子 (即对角)，因此 $\lambda E - A$ 的不变因子即其法式对角元素。

定理 6.23 (不变因子与最小多项式)

A 的最小多项式即 $\lambda E - A$ 的最后一个不变因子



证明 设 A 的互异特征值为 $\lambda_1, \dots, \lambda_s$ ，因此可以列出所有的初等因子 (具体见下一节)，最小多项式即初等因子的最小公倍式，最后一个不变因子也是，因此结论成立

6.3.3 初等因子与 Jordan 标准型**定义 6.10 (初等因子)**

将所有不变因子分解为不可约因子之积，这些不可约因子全体称为初等因子组。

**定理 6.24 (对角化计算初等因子)**

$\lambda E - A$ 经过初等变化变为对角阵 (不一定要法式)

$$\begin{pmatrix} f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda) \end{pmatrix}$$

其中 $f_i(\lambda)$ 为非零首一多项式，则 A 的初等因子组是所有 $f_i(\lambda)$ 做不可约分解得到的因式组合。

**引理 6.9 (Jordan 块、Jordan 阵的初等因子)**

r 阶 λ_0 对角的 Jordan 块的初等因子组为 $(\lambda - \lambda_0)^r$ 。对角为 $J_1, J_2, \dots, J_k$ 的 Jordan 阵的初等因子组为 $(\lambda - \lambda_1)^{r_1}, \dots, (\lambda - \lambda_k)^{r_k}$ 。



证明 (1) Jordan 块的初等因子组直接用行列式因子计算即可 (只有 r 阶非 1，因此不变因子好算)

(2) Jordan 阵的初等因子进行分块初等变换，将各个 Jordan 块分块变成法式，再用引理 6.24 即可。

定理 6.25 (使用初等因子计算 Jordan 标准型)

若 K 上的矩阵 A 有初等因子组 $(\lambda - \lambda_1)^{r_1}, \dots, (\lambda - \lambda_k)^{r_k}$ ，则 A 相似于 Jordan 阵 $J = \text{diag}\{J_1, \dots, J_k\}$ ，其中 J_i 是以 λ_i 为对角的 r_i 阶 Jordan 块。

**定理 6.26 (Jordan 标准型存在充要条件)**

K 上矩阵 A 有 Jordan 标准型当且仅当其初等因子都是一次项的幂次乘积。



6.4 λ 矩阵的进一步研究

6.4.1 Jordan 标准型的直接计算

一、初等变换方法：只需要将 $\lambda E - A$ 转换为对角矩阵，对角所有分解即初等因子组，直接构造 Jordan 标准型

 **练习 6.3 无法直接整除的情况** (1) 计算下面矩阵 A 的法式：

$$A = \begin{bmatrix} -1 & 1 & 3 \\ 3 & 0 & -4 \\ -2 & 1 & 4 \end{bmatrix}$$

解 (1) 进行到下面步骤的时候无法整除，即发现不整除列上的多项式，要进行对角整除化

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & -\lambda + 1 \\ 0 & \lambda^2 + \lambda - 3 & -3\lambda + 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & -\lambda + 1 \\ 0 & (\lambda^2 + \lambda - 3) - (\lambda - 1)(\lambda + 2) & (-3\lambda + 4) - (\lambda + 2)(-\lambda + 1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & -\lambda + 1 \\ 0 & -1 & \lambda^2 - \lambda - 2 \end{bmatrix}$$

因此最终可以化简为：

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & (\lambda - 1)^3 \end{bmatrix}$$

 **练习 6.4 初等变换方法计算 Jordan 标准型** 计算 $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ 2 & -2 & 4 \end{bmatrix}$ 的 Jordan 标准型

解 通过初等变换得到 $\lambda E - A$ 的法式为 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & -\lambda^2 + 2\lambda \end{bmatrix}$ ，因此初等因子为 $\lambda, \lambda, (\lambda - 2)$ ，对应 Jordan 阵为：

$$J = \begin{bmatrix} 0 & & \\ & 0 & \\ & & 2 \end{bmatrix}$$

二、行列式因子方法：求一般矩阵的 Jordan 阵用行列式因子可能比较麻烦，但是对于抽象的 n 阶矩阵却有很大的作用，原因在于只要有两个子式互素，则行列式因子就为 1。

 **练习 6.5 经典而重要的题目** $a, b \in \mathbb{C}$ ，求下面矩阵 A 的 Jordan 标准型

$$A = \begin{bmatrix} a & b & \cdots & b & b \\ & a & \cdots & b & b \\ & & \ddots & \vdots & \vdots \\ & & & a & b \\ & & & & a \end{bmatrix}$$

解 若 $b = 0$, 则 Jordan 标准型即 aE . 若 $b \neq 0$, 则考虑 $\lambda E - A$:

$$\lambda E - A = \begin{bmatrix} \lambda - a & -b & \cdots & -b & -b \\ & \lambda - a & \cdots & -b & -b \\ & & \ddots & \vdots & \vdots \\ & & & \lambda - a & -b \\ & & & & \lambda - a \end{bmatrix}$$

考虑 $n-1$ 阶子式, 右下子式行列式为 $(\lambda - a)^{n-1}$, 右上子式在 $\lambda = a$ 时显然不是 0 (根据余数定理), 因此互素, 即初等因子为 $(\lambda - a)^n$, 因此 Jordan 阵为一个 n 阶的以 a 为对角的 Jordan 块

 **练习 6.6 行列式因子应用相关练习** 求不变因子: (1) $A = \begin{bmatrix} & & 1 & \lambda + 2 \\ & 1 & \lambda + 2 \\ 1 & \lambda + 2 \\ \lambda + 2 \end{bmatrix}$ (2) $A = \begin{bmatrix} \lambda & -1 & & \\ & \lambda & -1 & \\ & & \lambda & -1 \\ 5 & 4 & 3 & \lambda + 2 \end{bmatrix}$

解 (1) 左上行列式因子为 -1 , 因此不变因子为 $1, 1, 1, (\lambda + 2)^4$

(2) 右上行列式因子为 -1 , 因此不变因子 $1, 1, 1, |A|$

定理 6.27 (重要 λ 矩阵)

下面 λ 矩阵 A 的不变因子为 $1, 1, \dots, \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n$

$$A = \begin{bmatrix} \lambda & & & & a_n \\ -1 & \lambda & & & a_{n-1} \\ & -1 & \lambda & & a_{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & -1 & \lambda & a_2 \\ & & & & -1 & \lambda + a_1 \end{bmatrix}$$



证明 考虑左下方行列式因子, 其值为 1, 因此不变因子为 $1, 1, \dots, 1, |A|$, 而计算出 $|A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n$, 因此结论成立

推论 6.1

下列矩阵 A 在 $\mathbb{C}$ 可对角化当且仅当 $\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n$ 在 $\mathbb{C}$ 上无重根

$$A = \begin{bmatrix} 0 & & & & a_n \\ -1 & 0 & & & a_{n-1} \\ & -1 & 0 & & a_{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & -1 & 0 & a_2 \\ & & & & -1 & a_1 \end{bmatrix}$$



 **练习 6.7 一些用行列式因子计算方便的特殊矩阵** 计算 Jordan 标准型 (1) $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ & 1 & 2 & 3 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}$ (2) $A = \begin{bmatrix} a & & & \\ a & a & & \\ \vdots & \vdots & \ddots & \\ a & a & \cdots & a \end{bmatrix}$

$$(3) A = \begin{bmatrix} a & & & & \\ b_1 & a & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & b_{n-1} & a \end{bmatrix} \quad (4) A = \begin{bmatrix} a & a_{12} & a_{13} & \cdots & a_{1n} \\ & a & a_{23} & \cdots & a_{2n} \\ & & \ddots & & \vdots \\ & & & a & a_{n-1,n} \\ & & & & a \end{bmatrix}, \quad a_{12}a_{23}\cdots a_{n-1,n} \neq 0$$

解 (1) $\lambda E - A$ 左上行列式为 $(\lambda - 1)^3$, 右上在 $\lambda = 1$ 是显然非零, 因此不变因子 $1, 1, 1, |\lambda E - A|$

(2) 左上行列式 $(\lambda - a)^{n-1}$, 左下在 $\lambda = a$ 显然非零, 因此不变因子 $1, 1, \cdots, 1, |\lambda E - A|$

(3)(4) 显然

 **笔记** 如果对角相等, 则用行列式因子计算可能会比较方便

6.4.2 k 阶 Jordan 块的个数

定理 6.28 (k 阶 Jordan 块个数)

n 阶复矩阵 A 的 Jordan 标准型为 J , 则 J 中对角元 λ 的 k 阶 Jordan 块 $J_k(\lambda)$ 的个数如下。这里约定 $r((A - \lambda E_n)^0) = n$

$$r((A - \lambda E_n)^{k-1}) - 2r((A - \lambda E_n)^k) + r((A - \lambda E_n)^{k+1})$$



证明 考虑零对角 Jordan 块的幂次:

$$r(J_r(0)^k) = \begin{cases} r-k, & 0 \leq k < r \\ 0, & k \geq r \end{cases} \quad r(J_r(0)^k) - r(J_r(0)^{k+1}) = \begin{cases} 1, & 0 \leq k < r \\ 0, & k \geq r \end{cases}$$

而非零对角 Jordan 块的幂次满足 $r(J_r(a)^k) - r(J_r(a)^{k+1}) = 0$ 。

设 A 的 Jordan 标准型为 $J = \text{diag}\{J_{r_1}(\lambda_1), \cdots, J_{r_s}(\lambda_s)\}$, $(A - \lambda E)^k, (J - \lambda E)^k$ 相似。根据 J 是分块对角的性质可知 $r((A - \lambda E)^k) = r((J - \lambda E)^k) = \sum_{i=1}^s r(J_{r_i}(\lambda_i - \lambda)^k)$, 因此得到:

$$r((A - \lambda E)^k) - r((A - \lambda E)^{k+1}) = \sum_{i=1}^s [r(J_{r_i}(\lambda_i - \lambda)^k) - r(J_{r_i}(\lambda_i - \lambda)^{k+1})]$$

根据前面的分析, $r((A - \lambda E)^k) - r((A - \lambda E)^{k+1})$ 等于 $\lambda_i = \lambda$ 的且阶数大于 k 那些块个数, 同理 $r((A - \lambda E)^{k-1}) - r((A - \lambda E)^k)$ 等于对角为 λ 的且阶数大于 $k-1$ 的块个数, 相减即可得到阶数为 k 的块个数。

 **练习 6.8 上述定理的应用** 计算 A 的 Jordan 标准型

$$A = \begin{bmatrix} a & 0 & 1 & & \\ & a & 0 & \ddots & \\ & & a & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & a \end{bmatrix}$$

 **练习 6.9** A, B 是两个 n 阶矩阵, 证明 A, B 相似的充要条件是 $\forall \lambda \in \mathbb{C}, k = 1, 2, \cdots, n$, 均有 $r((A - \lambda E)^k) = r((B - \lambda E)^k)$

证明 直接用结论即可

6.4.3 分块矩阵的 Jordan 阵

定理 6.29 (分块矩阵的 Jordan 标准型)

P 上 n 阶矩阵 A, B 无公共复特征值, 且 A, B 的 Jordan 标准型分别为 J_1, J_2 , 则对任意 $C \in P^{m \times n}$, 则:

$$\begin{bmatrix} A & C \\ O & B \end{bmatrix} \sim \text{diag}\{J_1, J_2\}$$



证明 当 A, B 无公共特征值时, 根据 $AX - XB$ 的矩阵方程结论可知, 对任意矩阵 C , $AX - XB = C$ 都有唯一解 X_0 使得 $AX_0 - X_0B = C$, 因此:

$$\begin{bmatrix} E_m & X_0 \\ O & E_n \end{bmatrix} \begin{bmatrix} A & C \\ O & B \end{bmatrix} \begin{bmatrix} E_m & -X_0 \\ O & E_n \end{bmatrix} = \begin{bmatrix} A & C - (AX_0 - X_0B) \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

因此得到:

$$\begin{bmatrix} A & C \\ O & B \end{bmatrix} \sim \begin{bmatrix} A & O \\ O & B \end{bmatrix} \sim \begin{bmatrix} J_1 & O \\ O & J_2 \end{bmatrix}$$

证明 第二种证明: 直接看 Jordan 标准型, λ_i 为对角的 k 阶 Jordan 块结论直接往上套即可。

推论 6.2

对于分块上三角矩阵 A , $A_{11}, A_{22}, \dots, A_{ss}$ 两两无公共特征值, A_{ii} 的 Jordan 标准型为 J_i , 则 A 的 Jordan 标准型为 $\text{diag}\{J_1, \dots, J_s\}$

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ & A_{22} & \cdots & A_{2s} \\ & & \ddots & \vdots \\ & & & A_{ss} \end{bmatrix}$$



练习 6.10 分块矩阵应用 求 A 的 Jordan 标准型:

$$A = \begin{bmatrix} 1 & 5 & -1 & 2 \\ 0 & 2 & 4 & 3 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 4 & -3 \end{bmatrix}$$

解 $\text{diag}\{1, 2, J_2(-3)\}$

6.4.4 过渡矩阵的计算

定理 6.30 (过渡矩阵计算总结)

若已知 A 的 Jordan 标准型为 J , 要计算 P 使得 $P^{-1}AP = J$, 则可以按照以下步骤:

1. 对于非一阶, 以 λ 为对角的 Jordan 块, 设对应的 P 位置为 $[X_1, \dots, X_k]$, 则满足

$$AX_1 = \lambda X_1, AX_2 = \lambda X_2 + X_1, \dots, AX_k = \lambda X_k + X_{k-1}$$

设右端项为 $b = [b_1, \dots, b_n]^T$ (未知量), 统一对增广矩阵做初等变换 $[A - \lambda E, b] \rightarrow [A', b']$, 选取顺序为 $X_1, \dots, X_k$ 。 X_1 要保证不能选 $[0, \dots, 0]^T$, 且选的 X_k 要保证 X_{k+1} 有解。

2. 对于一阶: 即 $AX = \lambda X$ 特征向量, 先求完非一阶再求, 不要选 $[0, \dots, 0]^T$, 保证与其他列向量线性无关。



 **练习 6.11** 已知矩阵 A 及其 Jordan 标准型 J , 求过渡矩阵 P 使得 $P^{-1}AP = J$

$$A = \begin{bmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{bmatrix} \quad J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

解 设 $P = (X_1, X_2, X_3)$, 根据 $P^{-1}AP = J$ 得到 $(AX_1, AX_2, AX_3) = (X_1, X_2, X_3)J$, 因此 $AX_1 = X_1, AX_2 = 3X_2, AX_3 = X_2 + 3X_3$, 即 $(A - E)X_1 = 0, (A - 3E)X_2 = 0, (A - 3E)X_3 = X_2$ 。先算后面非一阶, 做初等变换得到:

$$[A - 3E|b] \rightarrow \begin{bmatrix} -1 & 3 & 2 & b_1 \\ 0 & 8 & 4 & b_1 + b_2 \\ 0 & 8 & 4 & \frac{1}{2}b_1 + \frac{5}{2}b_2 + b_3 \end{bmatrix}$$

有解必须满足 $\frac{1}{2}b_1 + \frac{5}{2}b_2 + b_3 = 0$, 因此 $AX_2 = 0$ 等价于:

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & 8 & 4 \\ \frac{1}{2} & \frac{5}{2} & 1 \end{bmatrix} X_2 = 0$$

而最后一个条件显然可以被前两行消掉, 因此可以无视, 随便取 $X_2 = [1, -1, 2]^T$, 因此 X_3 满足

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & 8 & 4 \\ 0 & 0 & 0 \end{bmatrix} X_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

X_3 取法无所谓, 因此取 $X_3 = [-1, 0, 0]^T$ 。而 X_1 取线性无关即可, 如 $X_1 = [-2, 0, 1]^T$

 **练习 6.12** 已知矩阵 A 及其 Jordan 标准型 J , 求过渡矩阵 P 使得 $P^{-1}AP = J$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}, J = \begin{bmatrix} 2 & & \\ & 2 & 1 \\ & & 2 \end{bmatrix}$$

解 设 $P = [X_1, X_2, X_3]$, 则 $(AX_1, AX_2, AX_3) = (2X_1, 2X_2, X_2 + 2X_3)$, 得到方程组 $(A - 2E)X_1 = 0, (A - 2E)X_2 = 0, (A - 2E)X_3 = X_2$ 。做初等变换后:

$$[A|b] \rightarrow \begin{bmatrix} -2 & 1 & 0 & b_1 \\ 0 & 0 & 0 & b_2 - 2b_1 \\ 0 & 0 & 0 & b_3 - b_1 \end{bmatrix}$$

因此要使 $AX_3 = X_2$ 有解, X_2 必须满足 $b_3 = b_1$, 先选取 $X_2 = [1, 2, 1]^T$, 解出 $X_3 = [-\frac{1}{2}, 0, 0]^T$, X_1 要线性无关, 取 $[0, 0, 1]^T$ 即可。

第 7 章 双线性函数与二次型

7.1 双线性函数

7.1.1 双线性函数概念

定义 7.1 (线性函数)

V 是 K 上的线性空间, $f: V \rightarrow K$ 且满足 $f(\alpha + \beta) = f(\alpha) + f(\beta)$, $f(k\alpha) = kf(\alpha)$, 则称 f 是 V 上的线性函数。

定义 7.2 (双线性函数)

V 是 K 上的线性空间, $f: V \times V \rightarrow K$, 且满足 $f(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1f(\alpha_1, \beta) + k_2f(\alpha_2, \beta)$, $f(\alpha, l_1\beta_1 + l_2\beta_2) = l_1f(\alpha, \beta_1) + l_2f(\alpha, \beta_2)$, 则称 f 是 V 上的双线性函数

7.1.2 双线性函数度量矩阵

定义 7.3 (度量矩阵)

V 是 K 上的 n 维线性空间, $\epsilon_1, \dots, \epsilon_n$ 是 V 的一组基, f 为 V 上的双线性函数, 则 f 的度量矩阵 G 定义如下:

$$G = \begin{bmatrix} f(\epsilon_1, \epsilon_1) & f(\epsilon_1, \epsilon_2) & \cdots & f(\epsilon_1, \epsilon_n) \\ f(\epsilon_2, \epsilon_1) & f(\epsilon_2, \epsilon_2) & \cdots & f(\epsilon_2, \epsilon_n) \\ \vdots & \vdots & & \vdots \\ f(\epsilon_n, \epsilon_1) & f(\epsilon_n, \epsilon_2) & \cdots & f(\epsilon_n, \epsilon_n) \end{bmatrix}$$

定理 7.1 (双线性函数度量矩阵唯一确定)

若 $x, y \in V$ 在 $\epsilon_1, \dots, \epsilon_n$ 下的坐标为 X, Y , 则 $f(x, y) = X^T G Y$

7.1.3 双线性函数不同基下的矩阵: 合同

定义 7.4 (合同)

A, B 是 K 上的 n 阶方阵, 若存在 K 上的可逆 n 阶方阵 T 使得 $B = T^T A T$, 则称 A, B 合同。



笔记 从矩阵角度来看, 合同相当于 A 通过一系列相同的行与列变换变为 B

定理 7.2 (不同基下度量矩阵关系)

矩阵 A, B 合同当且仅当它们是双线性函数两组基下的矩阵。

定理 7.3 (分块矩阵合同)

这里只考虑 A 为 n 阶实对称矩阵, 希望对 A 整体或者一部分做合同变换:

• $A = \begin{pmatrix} A_1 & \alpha \\ \alpha^T & a_{nn} \end{pmatrix}$, 若 A_1 可逆, 则取 $P = \begin{pmatrix} E_{n-1} & -A_1^{-1}\alpha \\ 0 & 1 \end{pmatrix}$ (用列变换记忆), 得到

$$P^T A P = \begin{pmatrix} E_{n-1} & 0 \\ 0 & a_{nn} - \alpha^T A_1^{-1} \alpha \end{pmatrix}$$

• $A = \begin{pmatrix} a_{11} & \alpha^T \\ \alpha & A_{n-1} \end{pmatrix}$, 若 $a_{11} \neq 0$, 取 $P = \begin{pmatrix} 1 & -\frac{1}{a_{11}}\alpha^T \\ 0 & E_{n-1} \end{pmatrix}$, 则

$$P^T A P = \begin{pmatrix} a_{11} & 0 \\ 0 & A_{n-1} - \frac{1}{a_{11}}\alpha\alpha^T \end{pmatrix}$$

• 若想对上面的右下角分块 $A_{n-1} - \frac{1}{a_{11}}\alpha\alpha^T$ 做合同变换, 变为 $P_1^T \left(A_{n-1} - \frac{1}{a_{11}}\alpha\alpha^T \right) P_1$, 则 $Q =$

$P \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix}$ (用列变换记忆), 得到

$$Q^T A Q = \begin{pmatrix} a_{11} & 0 \\ 0 & P_1^T \left(A_{n-1} - \frac{1}{a_{11}}\alpha\alpha^T \right) P_1 \end{pmatrix}$$



7.2 二次型

7.2.1 二次型概念及其矩阵

定义 7.5 (二次型)

定义多元二次多项式 $f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij}x_i x_j$ 为二次型



定义 7.6 (二次型的矩阵)

对于任意二次型 $f(x_1, x_2, \dots, x_n)$, 总能找到一个对称矩阵 A , 使得 $f(x_1, \dots, x_n) = x^T A x$, 称矩阵 A 为二次型的矩阵



笔记 对于一个二次型 $f(x_1, \dots, x_n)$, 存在不止一个 A 使得 $f(x_1, \dots, x_n) = x^T A x$, 但是二次型的矩阵特指那个对称矩阵。

练习 7.1 二次型的矩阵 (1) 给定二次型 $f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$, 计算二次型的矩阵
(2) 给定二次型 $f(X)$, 计算二次型的矩阵:

$$f(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(3) 重点: 求 $f(x_1, \dots, x_n) = \sum_{i=1}^m (a_{i1}x_1 + \dots + a_{in}x_n)^2$ 的矩阵

解 (1) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix}$

(2) 展开得到 $f(x_1, x_2) = 2x_1^2 + 4x_1x_2 + x_2^2$, 因此矩阵为

$$\begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}$$

(3) 令 $A = (a_{ij})$, 则

$$\begin{aligned} f &= \sum_{i=1}^m (a_{i1}x_1 + \cdots + a_{in}x_n)^2 = \sum_{i=1}^m \left[\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \right]^2 \\ &= \sum_{i=1}^m \left[\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right] = X^T (A^T A) X \end{aligned}$$

其中计算出 $A^T A$ 的过程为

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \Rightarrow A^T A = \begin{bmatrix} A_1^T & A_2^T & \cdots & A_m^T \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} = \sum_{i=1}^m A_i^T A_i$$

定义 7.7 (二次型等价)

给定二次型 $f = x^T A x, g = y^T B y$, 若存在可逆矩阵 T , 使得 $x = Ty$, 且 $f = y^T (T^T A T) y = y^T B y = g$, 则称二次型 f, g 等价 (即 A, B 合同)



定理 7.4 (合同的性质)

若 $A = C^T B C$, 则

- 秩不变: $r(A) = r(B)$
- C 不一定唯一
- 分块合同: 若 A_1, A_2 分别与 B_1, B_2 合同, 则 $\text{diag}\{A_1, A_2\}$ 与 $\text{diag}\{B_1, B_2\}$ 合同



证明 (1) 由于 C 是满秩矩阵, 相当于一系列初等变换, 因此不改变秩

(2) 例如 $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 在 $\mathbb{C}$ 上合同, 但是

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & i \end{pmatrix}$$

练习 7.2 反对称矩阵的二次型 证明: A 是反对称矩阵当且仅当 $\forall X, X^T A X = 0$

证明 左推右显然。右推左根据 $X^T A X = X^T A^T X = -X^T A X = 0$ 可知

7.2.2 二次型的标准型存在与配方法

定理 7.5 (二次型标准型的存在性)

对数域 P 上任意对称矩阵 A , 其都合同于一个对角矩阵, 且对角矩阵不唯一。



证明 $n = 1$ 则显然。设 $n - 1$ 时成立, 则考虑 n 阶矩阵 $A = \begin{pmatrix} a_{11} & \alpha^T \\ \alpha & A_1 \end{pmatrix}$, 这里 A_1 对称。若 $a_{11} \neq 0$, 则根据

前面分块矩阵合同结论, 取 $P = \begin{pmatrix} 1 & -\frac{1}{a_{11}}\alpha^T \\ 0 & E_{n-1} \end{pmatrix}$, 得到

$$P^T A P = \begin{pmatrix} a_{11} & 0 \\ 0 & A_1 - \frac{1}{a_{11}}\alpha\alpha^T \end{pmatrix} := \begin{pmatrix} a_{11} & 0 \\ 0 & B_1 \end{pmatrix}$$

根据归纳假设, 存在 Q_1 使得 $Q_1^T B_1 Q_1$ 为对角阵, 则取 $Q = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$, 则

$$Q^T P^T A P Q = \begin{pmatrix} a_{11} & \\ & Q_1^T B_1 Q_1 \end{pmatrix}$$

若 $a_{11} = 0$, 但 $a_{kk} \neq 0$, 则取 $P(1, k)$ 为 E_n 第一列与第 k 列互换的矩阵, 则 $P^T(1, k) A P(1, k)$ 的 $a'_{11} \neq 0$, 且两个矩阵合同, 转换为第一种情况。

若对角全为 0, 讨论 A 第一行, 若 A 第一行全为 0, 则显然。否则不妨设 $a_{12} \neq 0$, 则 A 如下, 做相同行列变换变为:

$$A = \begin{pmatrix} 0 & a_{12} & * \\ a_{12} & 0 & * \\ * & * & * \end{pmatrix} \rightarrow \begin{pmatrix} 2a_{12} & a_{12} & * \\ a_{12} & a_{22} & * \\ * & * & * \end{pmatrix}$$

转换为第一种类型

 **笔记** 要特别注意, 如果对角均为零, 则将同行的非零元素加到对角上!

定义 7.8 (二次型标准型)

对于二次型 $f = x^T A x$, 若 A 与对角阵 D 合同, 则称 $g = z^T D z$ 为 f 的标准型



定理 7.6 (配方法)

配方法按照以下步骤进行:

1. 若无 x_1 的平方项, 但有 $x_1 x_2$ 交叉项, 则用 $x_1 = y_1 + y_2, x_2 = y_1 - y_2$ 换元。
2. 把 x_1 的项放在一起, 全部配方。
3. 把凑成的 $(x_1 + \dots)^2$ 设为新变量
4. 以此类推



 **笔记** 配方法一般算过渡矩阵不太方便, 不过算标准型还行

 **练习 7.3 配方法** (1) 用配方法计算 $f(x_1, x_2, x_3) = 2x_1 x_2 + 2x_1 x_3 - 6x_2 x_3$ 的标准型

解 (1) 无 x_1^2 项, 因此令 $x_1 = y_1 + y_2, x_2 = y_1 - y_2, x_3 = y_3$, 得到

$$f = 2(y_1 - y_3)^2 - 2y_2^2 - 2y_3^2 + 8y_2 y_3$$

令 $x_1 = y_1 - y_3, x_2 = y_2, x_3 = y_3$ 得到 $f = 2x_1^2 - 2(x_2^2 - 4x_2 x_3 + 4x_3^2) + 6x_3^2$, 再做换元 $y_1 = x_1, y_2 = x_2 - 2x_3, y_3 = x_3$, 得到

$$f = 2y_1^2 - 2y_2^2 + 6y_3^2$$

7.2.3 二次型的标准型：初等变换法

定理 7.7 (标准型的计算：初等变换法)

对于对称矩阵 A 使用以下格式计算初等变换 (注意这不一定得出正交矩阵):

$$\begin{pmatrix} A \\ E \end{pmatrix} \rightarrow \begin{pmatrix} C^T AC \\ C \end{pmatrix} \quad \text{对 } A \text{ 行变换和对应列变换, 对 } I \text{ 只列变换}$$

具体的步骤就是证明标准型存在性中的过程: 若 A 首对角为零, 存在非零对角, 则互换位置。若对角均为零, 则找到第一行非零的元素加到首对角位置 (下面的题目马上会用)



笔记 做同样的行列变换一定要小心, 两种变换并非“同时”做的, 而是先后做的, 即做行变换时的矩阵已经是列变换后的, 而非原本的矩阵!

练习 7.4 标准型计算：初等变换法 (1) 简单: 计算 $f = x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_2x_3 + 4x_3^2$ 的标准型
(2) 对角均零: 计算 $f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$, 计算其标准型, 并求出过渡矩阵 C

解 (1) 结果为 $f = y_1^2 + y_2^2$, 千万小心别行列变换错了

(2) 首先显然 $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix}$, 列出 $\begin{pmatrix} A \\ E \end{pmatrix}$ 并做行列初等变换, 由于 A 对角均为 0, 因此将第二列

和第二行加到首对角位置:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & -3 \\ -2 & -3 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & -2 \\ 0 & -2 & -2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \dots$$

最终变为结果为:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 6 \\ 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} C^T AC \\ C \end{pmatrix}$$

7.2.4 二次型的标准型：正交变换法

定理 7.8 (正交对角化的可行性)

任何实对称矩阵 A 一定正交相似于对角阵。



证明 做数学归纳法, $n=1$ 显然, 假设 $n-1$ 成立, 任取特征值 λ , 对应单位特征向量 α_1 , 将 α_1 扩充为 $\mathbb{R}^n$ 的一组标准正交基 $\alpha_1, \dots, \alpha_n$, 令 $T = (\alpha_1, \dots, \alpha_n)$, 则

$$AT = T \begin{pmatrix} \lambda & \alpha^T \\ 0 & A_{n-1} \end{pmatrix} \Rightarrow T^T AT = \begin{pmatrix} \lambda & \alpha^T \\ 0 & A_{n-1} \end{pmatrix}$$

由于 $T^T AT$ 对称, 因此 $\alpha^T = 0$, 再根据归纳假设, 存在正交矩阵 P_1 使得 $P_1^T A_{n-1} P_1$ 为对角阵, 因此令 $P =$

$\text{diag}(1, P_1)$, 此时

$$P^T T^T A T P = \begin{pmatrix} \lambda & 0 \\ 0 & P^T A_{n-1} P \end{pmatrix}$$

显然 $TP = Q$, 则 Q 也是正交阵。

定理 7.9 (标准型计算: 正交变换)

正交变换计算标准型过程如下:

- 计算 A 的特征值 $\lambda_1, \dots, \lambda_n$
- 计算 λ_i 对应的特征向量 $\mathbf{v}_i$, 不同特征值的特征向量必正交, 同一特征值的特征向量需正交化
- 将特征向量单位化, 组成 $T = [\mathbf{v}_1, \dots, \mathbf{v}_n]$



练习 7.5 正交变换法 (1) 已知矩阵 A 如下, (a) 若 $tE + A$ 为正定矩阵, 求 t 的取值范围 (b) 求正交矩阵 T 使得 $T^T A T$ 为对角矩阵

$$A = \begin{bmatrix} 2 & 2 & -1 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$$

解 (1) $|\lambda E - A| = (\lambda - 1)^2(\lambda - 10)$, 特征值为 1, 10, 正定要求 $t > -1$ 。

先求 $(E - A)X = 0$, 得到:

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} X = 0 \Rightarrow X = k_1 \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

单位化得到 $\frac{1}{\sqrt{5}}(-2, 1, 0)^T, \frac{1}{\sqrt{5}}(2, 0, 1)^T$

再求 $(10E - A)X = 0$:

$$\begin{bmatrix} 2 & 4 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} X = 0 \Rightarrow X = k_1 \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

单位化得到 $\frac{1}{3}(1, 2, -2)^T$ 。

最终结果为:

$$T = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & \frac{2}{3} \\ -\frac{1}{\sqrt{5}} & \frac{5}{3\sqrt{5}} & -\frac{2}{3} \\ 0 & \frac{1}{3\sqrt{5}} & -\frac{2}{3} \end{bmatrix}, T^T A T = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 10 \end{bmatrix}$$

定理 7.10 (二次型与特征值)

设 n 阶实对称矩阵 A 的全部特征值排序为: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, 则

$$\max_{X \in \mathbb{R}^n} \frac{X^T A X}{X^T X} = \lambda_1, \min_{X \in \mathbb{R}^n} \frac{X^T A X}{X^T X} = \lambda_n$$



证明 由于 $\lambda_1 E - A$ 的特征值均非负, 故 $\lambda_1 E - A$ 半正定, 因此 $\forall X \in \mathbb{R}^n$ 有 $X^T(\lambda_1 E - A)X \geq 0$, 得到 $X^T A X \leq \lambda_1 X^T X$, 且 λ_1 是 $\lambda_1 E - A$ 的特征值, 故 $\exists X_0 \neq 0$ 使得 $X_0^T(\lambda_1 E - A)X_0 = 0$

练习 7.6 二次型的特征值研究 (1) 重点: A, B 是 n 阶实对称矩阵, A 的特征值都大于 a , B 的特征值都大于 b ,

证明 $A+B$ 的特征值都大于 $a+b$

(2) B 是 n 阶实矩阵, $B^T B$ 的全部特征值按照大小顺序排为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, 证明: B 的任意实特征值 μ 满足 $\sqrt{\lambda_n} \leq |\mu| \leq \sqrt{\lambda_1}$

证明 (1) 根据前面定理可知 $X^T A X > a X^T X$, $X^T B X > b X^T X$, 从而 $X^T (A+B) X > (a+b) X^T X$, 因此 $A+B$ 的特征值大于 $a+b$

(2) 由于 $B\alpha = \mu\alpha$, 因此 $\alpha^T B^T = \mu\alpha^T$, 得到

$$\alpha^T B^T B \alpha = \mu^2 \alpha^T \alpha \in [\lambda_n \alpha^T \alpha, \lambda_1 \alpha^T \alpha] \Rightarrow \lambda_n \leq \mu^2 \leq \lambda_1$$

7.2.5 $\mathbb{R}, \mathbb{C}$ 上的规范型

前面根据标准型对二次型进行分类, 但是可以发现二次型的标准型是不唯一的, 因此在 $\mathbb{R}, \mathbb{C}$ 上可以进一步划分, 即每个规范型表示一个类。

定理 7.11 ($\mathbb{C}$ 上的规范型)

$\mathbb{C}$ 上任一二次型 f 都等价于

$$g = u_1^2 + \cdots + u_r^2$$

推论 7.1 ($\mathbb{C}$ 上合同充要条件)

$\mathbb{C}$ 上两个对称矩阵合同的充要条件为它们的秩相同。

定理 7.12 ($\mathbb{R}$ 上的规范型)

$\mathbb{R}$ 上每个二次型 f 都等价于一个如下二次型。将 p 称为正惯性指数, $r-p$ 称为负惯性指数

$$g = u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_r^2$$

推论 7.2 ($\mathbb{R}$ 上合同充要条件)

$\mathbb{R}$ 上两个对称矩阵充要条件是它们有相同的正负惯性指数。

练习 7.7 规范型的应用 (1) 设 A 是 n 阶实对称矩阵, $|A| < 0$, 证明: $\exists X \in \mathbb{R}^n$ 满足 $X^T A X < 0$

(2) $f(X) = X^T A X$ 为实二次型, 若存在 X_1, X_2 使得 $X_1^T A X_1 > 0, X_2^T A X_2 < 0$, 证明: $\exists X_0 \neq 0$ 使得 $X_0^T A X_0 = 0$

证明 (1) 由于 $|A| < 0$, 可知 A 满秩且负惯性指数个数不为零, 做非退化线性替换 $X = CY$ 即可得到 $Y^T C^T A C Y$, 这里

$$C^T A C = \text{diag}\{1, 1, \cdots, 1, -1, -1, \cdots, -1\}$$

取 $Y = e_n$ 时 $Y^T C^T A C Y = -1$

(2) 正负惯性指数均非零, 因此可转换为规范型 $f(X) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_n^2$, 取 $Y = e_1 + e_{p+1}$ 即可

7.3 (半) 正定二次型的研究

7.3.1 正定二次型的概念与判别

正定二次型是 $\mathbb{R}$ 中二次型中的特殊等价类, 其代表为 $y^T E y$

定义 7.9 (正定二次型)

给定实数域上的二次型 $f(x_1, x_2, \dots, x_n)$, 若对任意不全为 0 的 $x_1, \dots, x_n$ 有 $f(x_1, \dots, x_n) > 0$, 则称 $f(x_1, \dots, x_n)$ 为正定矩阵。

**定理 7.13 (正定二次型的规范型)**

正定二次型的规范型是 $y^T E y = y_1^2 + y_2^2 + \dots + y_n^2$

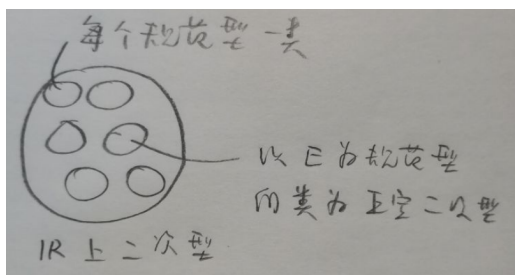


图 7.1: 二次型分类示意图

定理 7.14 (正定充要条件)

$f = x^T A x$ 是正定二次型当且仅当:

- A 合同于矩阵 E
- 存在可逆矩阵 C , 使得 $A = C^T C$, 也可以分解为非方阵 $A = \begin{pmatrix} C^T & O^T \end{pmatrix} \begin{pmatrix} C \\ O \end{pmatrix}$
- 所有特征值都大于 0
- A 的各阶顺序主子式/所有主子式行列式大于 0



证明 等价于顺序主子式大于 0 需要证明: 对 A 做如下分块:

$$A = \begin{bmatrix} A_1 & \alpha \\ \alpha^T & a_{nn} \end{bmatrix}$$

正定推顺序主子式: 若 A 正定, 则 A_1 正定, 因此 A_1 特征值均为正, $|A_1| > 0, |A| > 0$, 根据归纳法成立。

顺序主子式推正定: $n = 1$ 时显然成立, A 是 n 阶矩阵时, A_1 是 $n - 1$ 阶矩阵且顺序主子式均正, 根据归纳假设 A_1 正定, 故存在 P_1 使得 $P_1^T A_1 P_1 = E_{n-1}$, 因此令:

$$P_2 = \begin{bmatrix} P_1 & \\ & 1 \end{bmatrix}, P_3 = \begin{bmatrix} E_{n-1} & -P_1^T \alpha \\ & 1 \end{bmatrix}$$

得到:

$$P_3^T P_2^T A P_2 P_3 = \begin{bmatrix} E_{n-1} & \\ & a_{nn} - \alpha^T P_1 P_1^T \alpha \end{bmatrix}$$

而根据乘积矩阵的行列式性质得到

$$a_{nn} - \alpha^T P_1 P_1^T \alpha = |A| |P_2|^2 |P_3|^2 > 0$$

令 $P = P_2 P_3 \text{diag}\{E_{n-1}, \frac{1}{\sqrt{a_{nn} - \alpha^T P_1 P_1^T \alpha}}\}$ 得到

$$P^T A P = E$$

因此正定

- 练习 7.8 正定二次型的判定** (1) t 为何值时, $f = x_1^2 + x_2^2 + 5x_3^2 + 2tx_1x_2 - 2x_1x_3 + 4x_2x_3$ 为正定的?
 (2) A 为实对称矩阵, 则 $\exists M, \forall t > M$ 有 $tE + A$ 为正定矩阵

解 (1) 算顺序主子式即可, 得到 $1 - t^2 > 0, -5t^2 - 4t > 0$, 因此 $-\frac{4}{5} < t < 0$
 (2) 看特征值即可

 **练习 7.9 正定矩阵性质的简单应用** (1) A 为 n 阶正定矩阵, 证明: 任意 n 阶实矩阵 B 有 $r(B^T AB) = r(B)$

(2) A 为 n 阶正定矩阵, B 为 $n \times m$ 实矩阵, 证明: 若 $r(B) = m$, 则 $C = B^T AB$ 为正定矩阵

证明 (1) 由于 A 正定, 存在满秩矩阵 C 使得 $A = C^T C$, 因此 $B^T AB = (CB)^T (CB)$, 而根据 $A^T A$ 的结论可知 $r((CB)^T (CB)) = r(CB) = r(B)$, 因此 $r(B^T AB) = r(B)$

(2) 由于 A 正定, 因此 $B^T AB$ 实对称, $\forall X, X^T CX = (BX)^T ABX$, 根据 A 的正定性可知其大于 0, 因此 $B^T AB$ 为正定矩阵

定理 7.15 (正定矩阵对角性质)

A 为 n 阶正定矩阵, 则

1. $a_{ii} > 0, \forall i = 1, 2, \dots, n$
2. $2|a_{ij}| < a_{ii} + a_{jj} (i \neq j)$
3. A 的所有元素中, 绝对值最大的元素一定是在主对角线上, 且一定是正数



证明 (1) 不妨取 $X = e_i$, 则 $X^T AX = a_{ii} > 0$

(2) 取 $X = e_i + e_j$, 则 $X^T AX = a_{ii} + a_{jj} + 2a_{ij} > 0$, 取 $X = e_i - e_j$, 则 $X^T AX = a_{ii} + a_{jj} - 2a_{ij} > 0$, 综上得到结论成立。

(3) 根据 (2) 可知 $\forall i, j$, 有 $|a_{ij}| < \max\{a_{ii}, a_{jj}\}$, 因此结论成立

7.3.2 半正定二次型的概念与判别

定义 7.10 (半正定二次型)

若对任意 $x_1, \dots, x_n$ 均有 $f(x_1, \dots, x_n) \geq 0$, 则称 $f(x_1, \dots, x_n)$ 为半正定二次型, 对应矩阵 A 为半正定矩阵。



定理 7.16 (半正定二次型的基础充要条件)

$f(x_1, \dots, x_n) = X^T AX$ 为半正定二次型当且仅当:

- A 和 $\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$ 合同, A 的正惯性指数与 A 的秩相同
- A 的特征值都大于等于 0
- A 的所有主子式 (不等价于顺序主子式) 大于等于 0



 **笔记** 半正定不等价于顺序主子式大于等于 0, 例如 $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$, 顺序主子式都为 0, 但是半负定。

定理 7.17 (半正定矩阵的 $C^T C$ 分解)

$f(x_1, \dots, x_n) = X^T AX$ 为半正定二次型等价于下面两个命题

- 存在不列满秩的 $n \times n$ 矩阵 C , 使得 $A = C^T C$
- 存在行满秩 (列不满秩) 的 $r \times n (r < n)$ 矩阵 C , 使得 $A = C^T C$



证明 根据 $A = C^T \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} C$ 以及

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} E_r \\ O \end{pmatrix} \begin{pmatrix} E_r & O \end{pmatrix}$$

可以得到两种分解:

$$\begin{cases} C_1 = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} C \Rightarrow A = C_1^T C_1 \\ C_2 = \begin{pmatrix} E_r & O \end{pmatrix} C \Rightarrow A = C_2^T C_2 \end{cases}$$

无论哪种分解列均不满秩。

定理 7.18 (半正定矩阵的 CC^T 分解)

$f(x_1, \dots, x_n) = X^T A X$ 为半正定二次型等价于下面两个命题

- 存在不行满秩的 $n \times n$ 矩阵 C , 使得 $A = CC^T$
- 存在列满秩 (行不满秩) 的 $n \times r (r < n)$ 矩阵 C , 使得 $A = CC^T$



证明 右推左: 直接 $X^T CC^T X = (C^T X)^T (C^T X)$, 该多项式是 $C^T X$ 列向量中元素的平方和, 因此大于等于 0。

左推右: 合同的 $A = B^T \text{diag}\{E_r, O\} B$ 也可以写为 $A = C \text{diag}\{E_r, O\} C^T$, 因此把中间的分解了即可。

笔记 特别的可以取 $A = \alpha^T \alpha$, α 为列向量 (列满秩), 此时 A 是半正定的。

练习 7.10 半正定矩阵 CC^T 分解 (1) $\alpha \in \mathbb{R}^n$, 证明: $\alpha \alpha^T$ 一定是半正定矩阵

(2) 重点: A 的元素 $a_{ij} = (i \times j)$, B 的元素 $b_{ij} = \left(\frac{1}{i \times j}\right)$, 则 A, B 都是半正定矩阵

证明 (1) 显然 $\alpha \alpha^T$ 是实对称矩阵, 从而 $X^T \alpha \alpha^T X = (\alpha^T X)^T (\alpha^T X) = (\alpha^T X)^2 \geq 0$

(2) A, B 可进行如下分解, 根据 (1) 的结论可证是半正定矩阵

$$A = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix} \begin{bmatrix} 1 & 2 & \cdots & n \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{n} \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \end{bmatrix}$$

定理 7.19 (半正定矩阵对角性质)

半正定矩阵 A 有以下性质:

1. 对角元素非负
2. $2|a_{ij}| \leq a_{ik} + a_{ki}$
3. 若 $a_{ii} = 0$, 则 $\forall k, a_{ik} = a_{ki} = 0$



证明 (1) 只需要取 $x = e_i$, $x^T A x \geq 0$ 即可

(2) 取 $x = e_i + e_j$, 得到 $a_{ii} + a_{jj} + 2a_{ij} \geq 0$, 同理取 $x = e_i - e_j$, 得到 $a_{ii} + a_{jj} - 2a_{ij} \geq 0$, 因此 $2|a_{ij}| \leq a_{ik} + a_{ki}$

(3) 根据 (2) 可以得到

7.3.3 $C^T C$ 的 (半) 正定性

$C^T C$ 对应着特殊的二次型 $f(X) = \sum_{i=1}^n (a_{i1}x_1 + \cdots + a_{in}x_n)^2$, 且有着特殊的正定性, 是正定二次型中特殊的一类。

引理 7.1 ($r(A^T A)$ 与 $r(A)$)

A 为实数域上的 $m \times n$ 的矩阵, 则

$$r(A^T A) = r(A)$$

证明 在最小二乘法一节中已经给出了证明

定理 7.20 ($A^T A$ 的正定性)

考虑 A 为实数域上的 $m \times n$ 的矩阵, 则

1. $A^T A, AA^T$ 是半正定矩阵

2. $A^T A$ 是正定矩阵当且仅当 A 列满秩, AA^T 正定当且仅当 A 行满秩

且 $A^T A$ 对应的二次型为 $f(X) = \sum_{i=1}^n (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n)^2$

证明 (半) 正定性: 显然 $A^T A$ 实对称, 因此 $X^T A^T A X = (AX)^2 \geq 0$, 显然 $A^T A$ 半正定。而 $A^T A$ 正定当且仅当 $r(A^T A) = n$, 根据 (1) 当且仅当 $r(A) = n$, 这里 n 等于 A 的列数, 因此等价于 A 列满秩

二次型表达式: 在二次型矩阵一节已经作为练习给出。

7.3.4 正定矩阵同时对角化**定理 7.21** (同时合同对角化)

若 A, B 是 n 阶半正定矩阵, 则存在一个 n 阶可逆矩阵 P , 使得 $P^T A P, P^T B P$ 同时为对角矩阵

证明 由于 A, B 半正定, 因此 $A + B$ 半正定, 存在可逆矩阵 M 使得

$$M^T (A + B) M = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} = M^T A M + M^T B M$$

由于 $M^T A M, M^T B M$ 半正定, 因此根据半正定矩阵的对角性质, 当 $r+1 \leq i \leq n$, 其第 i 个对角元都非负, 因此 $r+1 \sim n$ 的对角元都必须为 0。再根据对角性质可知 $r+1 \sim n$ 行、列元素均为 0, 因此

$$M^T A M = \begin{pmatrix} H_r & O \\ O & O \end{pmatrix}, M^T B M = \begin{pmatrix} E_r - H_r & O \\ O & O \end{pmatrix}$$

存在正交矩阵 T 使得 $T^T H_r T = \Lambda = \text{diag}\{\lambda_1, \dots, \lambda_r\}$, 令 $P = M \text{diag}\{T, E_{n-r}\}$, 此时有

$$P^T A P = \begin{pmatrix} \Lambda & O \\ O & O \end{pmatrix}, P^T B P = \begin{pmatrix} E_r - \Lambda & O \\ O & O \end{pmatrix}$$

同时为对角矩阵

定理 7.22 ($AB = BA$ 同时相似对角化)

若 A, B 可相似对角化, 且 $AB = BA$, 则存在可逆矩阵 P , 使得

$$P^{-1} A P, P^{-1} B P$$

同时为对角矩阵

证明 由于 A 可相似对角化, 因此存在 M , 使得 $M^{-1} A M = \text{diag}\{\lambda_1 E, \lambda_2 E, \dots, \lambda_r E\}$, 而

$$M^{-1} A B M = (M^{-1} A M)(M^{-1} B M) = M^{-1} B A M = (M^{-1} B M)(M^{-1} A M)$$

将 $M^{-1} B M$ 以同样的形式进行分块, 得到 $M^{-1} B M = (B_{ij})_{r \times r}$, 对于每个 B_{ii} , $\exists Q_i$ 使得 $Q_i^{-1} B_{ii} Q_i = \Lambda_i$, 因此令

$$T = M \text{diag}\{Q_1, \dots, Q_r\}$$

可以得到

$$T^{-1}AT = \text{diag}\{\lambda_1 E, \dots, \lambda_r E\}, T^{-1}BT = \text{diag}\{\Lambda_1, \dots, \Lambda_r\}$$

同时为对角矩阵。

7.3.5 正定矩阵的四则运算性质

定理 7.23 (正定的四则运算)

若 A, B 是正定矩阵, 则 $A^{-1}, A^k, A+B, A^*$ 都是正定矩阵, 但 AB 不一定正定 (原因在于不一定对称), 若 AB 实对称, 则 AB 是正定矩阵



证明 (1) A^{-1}, A^k, A^* : 由于 A 正定, 因此 A 的特征值 λ_i 均大于 0, A^{-1} 的特征值 $\lambda_i^{-1} > 0$, A^* 的特征值 $\prod_{j \neq i} \lambda_j > 0$, A^k 的特征值 $\lambda_i^k > 0$, 因此均正定。

(2) $A+B$ 不能用特征值, 只能用合同, 显然 $A+B$ 实对称, 而任意非零 $x \in \mathbb{R}^n$, 有 $X^T(A+B)X = X^TAX + X^TBX > 0$, 因此 $A+B$ 实对称。

(3) AB 正定: 由于 AB 实对称, 此时 $(AB)^T = B^T A^T = BA$ 。由于 A, B 正定, 故存在可逆 P, Q 使得 $A = P^T P, B = Q^T Q, AB = P^T P Q^T Q$, 而

$$QABQ^{-1} = QP^T P Q^T (QQ^{-1}) = (PQ^T)^T (PQ^T)$$

因此 AB 与 $(PQ^T)^T (PQ^T)$ 相似, 而后者正定 (因为列满秩), 因此 AB 为正定矩阵。

7.3.6 正定矩阵分解进阶

定理 7.24 (正定矩阵平方分解)

A 是 n 阶正定矩阵, 则存在实对称矩阵 C_1 , 使得 $A = C_1^2$



证明 由于 $\exists P$ 使得 $P^T A P = \Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, 根据正定, 所有特征值都为正, 因此取

$$C_1 = P^T \sqrt{\Lambda} P$$

即可得到 $A = C_1^2$

笔记 上述 $A = C_1^2$ 也是正定矩阵分解! 且是唯一的。

定理 7.25 (正定矩阵三角分解)

A 是 n 阶正定矩阵, 则存在可逆上三角阵 R , 使得 $A = R^T R$



证明 首先 $A = C^T C$, C 为可逆矩阵, 再对 C 做 QR 分解得到 $C = QR$, 因此 $A = R^T Q^T Q R = R^T R$, 即所求。

定理 7.26 (极分解定理)

对任一可逆矩阵 A , 一定存在一个正交矩阵 T 和两个正定矩阵 S_1, S_2 使得

$$A = T S_1 = S_2 T$$

且这种分解是唯一的。



证明 (1) 存在性: 由于 A 可逆, 因此 $A^T A$ 正定, 根据前面结论存在正定矩阵 S_1 , 满足 $A^T A = S_1^2$, 设 $T = (A^T)^{-1} S_1$, 则

$$A = (A^T)^{-1} S_1^2 = T S_1$$

经过验证可发现 $T T^T = (A^T)^{-1} S_1 S_1^T A^{-1} = (A^T)^{-1} S_1^2 A^{-1} = (A^T)^{-1} A^T A A^{-1} = E$, 因此为正交矩阵。

另一方面, $A = TS_1T^{-1}T$, 记 $S_2 = TS_1T^{-1}$, 则 $A = S_2T$, S_2 正定。

(2) 唯一性: 设 $A = TS_1 = \tilde{T}\tilde{S}_1$, 则

$$\begin{cases} A^T A = S_1^2 \\ A^T A = \tilde{S}_1^2 \end{cases}$$

根据前面正定矩阵分解唯一性可知唯一。

7.3.7 负定二次型

定义 7.11 (负定二次型)

若对于任意不全为零的 $x_1, \dots, x_n$, 满足 $f(x_1, \dots, x_n) < 0$, 则称 $f(x_1, \dots, x_n)$ 为负定二次型



定理 7.27 (负定二次型的充要条件)

A 负定当且仅当 $-A$ 正定, 因此充要条件为: A 的所有奇数阶顺序主子式小于 0, 偶数阶顺序主子式大于 0。



证明 用特征值判断即可。

第8章 内积空间

8.1 Euclid 空间

8.1.1 Euclid 空间与正交

定义 8.1 (欧式空间)

V 是 $\mathbb{R}$ 上线性空间, 定义内积 (α, β) 满足:

- 线性性: $(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1(\alpha_1, \beta) + k_2(\alpha_2, \beta)$
- 对称性: $(\alpha, \beta) = (\beta, \alpha)$
- 正定性: $(\alpha, \alpha) \geq 0$ 当且仅当 $\alpha = 0$



定义 8.2 (度量矩阵)

给定 Euclid 空间 V 的一组基 $\epsilon_1, \dots, \epsilon_n$, 则定义这组基对应的度量矩阵为

$$A = \begin{bmatrix} (\epsilon_1, \epsilon_1) & (\epsilon_1, \epsilon_2) & \cdots & (\epsilon_1, \epsilon_n) \\ (\epsilon_2, \epsilon_1) & (\epsilon_2, \epsilon_2) & \cdots & (\epsilon_2, \epsilon_n) \\ \vdots & \vdots & & \vdots \\ (\epsilon_n, \epsilon_1) & (\epsilon_n, \epsilon_2) & \cdots & (\epsilon_n, \epsilon_n) \end{bmatrix}$$



定理 8.1 (内积的度量矩阵表示)

若 Euclid 空间中的向量 α, β 在基 $(\epsilon_1, \dots, \epsilon_n)$ 中的坐标分别为 X, Y , $(\epsilon_1, \dots, \epsilon_n)$ 对应的度量矩阵为 A , 则

$$(\alpha, \beta) = X^T A Y$$



练习 8.1 度量矩阵 已知 Euclid 空间 V 的一组基为:

$$\epsilon_1 = e_1 + e_2, \epsilon_2 = e_1 + e_3, \epsilon_3 = e_4 - e_1, \epsilon_4 = e_1 - e_2 - e_3 + e_4$$

向量 α, β 在 $(\epsilon_1, \dots, \epsilon_4)$ 下的坐标分别为 $(1, 2, 3, 4)^T, (2, 0, 1, 0)^T$, 计算 $(\epsilon_1, \dots, \epsilon_4)$ 的度量矩阵 A , 以及 (α, β)

解 把 $\epsilon_1, \dots, \epsilon_4$ 看成 $\mathbb{R}^n$ 中的坐标:

$$\begin{cases} \epsilon_1 = (1, 1, 0, 0) \\ \epsilon_2 = (1, 0, 1, 0) \\ \epsilon_3 = (-1, 0, 0, 1) \\ \epsilon_4 = (1, -1, -1, 1) \end{cases}$$

因此根据向量的坐标内积公式可以得到度量矩阵为:

$$A = \begin{bmatrix} 2 & 1 & -1 & 0 \\ 1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, (\alpha, \beta) = 5$$

定理 8.2 (正交向量线性无关)

若 Euclid 空间 V 中的向量 $\alpha_1, \dots, \alpha_s$ 彼此正交, 则它们线性无关。



证明 不妨设 $c_1\alpha_1 + \cdots + c_s\alpha_s = 0$, 则两侧与 α_i 做内积得到:

$$(c_1\alpha_1 + \cdots + c_s\alpha_s, \alpha_i) = (0, \alpha_i) = 0$$

根据正交定义可知左侧等于 $c_i(\alpha_i, \alpha_i)$, 而 $(\alpha_i, \alpha_i) \neq 0$, 因此 $c_i = 0$, 根据 i 的任意性可知线性无关。

8.1.2 标准正交基

一、标准正交基的概念与存在性

定义 8.3 (标准正交基)

将 n 维 Euclid 空间中 n 个两两正交的单位向量称为标准正交基



二、标准正交基之间的过渡矩阵

定理 8.3 (标准正交基的过渡矩阵)

$\epsilon_1, \cdots, \epsilon_n$ 是 n 维 Euclid 空间 V 的一组基, 若 $\eta_1, \cdots, \eta_n$ 下述等式, 则 $\eta_1, \cdots, \eta_n$ 是标准正交基当且仅当 T 是正交矩阵。

$$(\eta_1, \cdots, \eta_n) = (\epsilon_1, \cdots, \epsilon_n)T$$



三、标准正交基的计算: Schmidt 正交化

定理 8.4 (Schmidt 正交化)

给定一组线性无关的向量 $\alpha_1, \cdots, \alpha_s$, 要求做出一组相互正交的向量 (单位化单独进行) $\epsilon_1, \cdots, \epsilon_s$, 可以按照以下步骤:

$$\begin{aligned}\epsilon_1 &= \alpha_1 \\ \epsilon_2 &= \alpha_2 - \frac{(\alpha_2, \epsilon_1)}{(\epsilon_1, \epsilon_1)}\epsilon_1 \\ &\vdots \\ \epsilon_s &= \alpha_s - \sum_{i=1}^{s-1} \frac{(\alpha_s, \epsilon_i)}{(\epsilon_i, \epsilon_i)}\epsilon_i\end{aligned}$$



练习 8.2 Schmidt 正交化 已知 $\epsilon_1 = (1, 2, -1)^T, \epsilon_2 = (-1, 3, 1)^T, \epsilon_3 = (4, -1, 0)^T$, 用 Schmidt 正交化将其化为标准正交基

解 最终结果为 $\eta_1 = \frac{1}{\sqrt{6}}(1, 2, -1)^T, \eta_2 = \frac{1}{\sqrt{3}}(-1, 1, 1)^T, \eta_3 = \frac{1}{\sqrt{2}}(1, 0, 1)^T$

定理 8.5 (QR 分解)

任意实可逆矩阵可分解为正交矩阵以及主对角元为正的上三角矩阵的乘积, 且这种分解是唯一的



证明 假设原本的矩阵 A 的列向量为 $\alpha_1, \cdots, \alpha_n$, 经过 Schmidt 正交化后得到 $\beta_1, \cdots, \beta_n$, 则满足:

$$\begin{cases} \alpha_1 = \beta_1 \\ \alpha_2 = a_{12}\beta_1 + \beta_2 \\ \vdots \\ \alpha_n = a_{1n}\beta_1 + a_{2n}\beta_2 + \cdots + \beta_n \end{cases}$$

再将 β 进行单位化得到 $\eta_i = \frac{\beta_i}{|\beta_i|}$, 此时写为

$$(\alpha_1, \dots, \alpha_n) = (\eta_1, \dots, \eta_n) \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{pmatrix}$$

此时对角 $b_{11} = |\beta_1|, \dots, b_{nn} = |\beta_n|$, 而 $(\eta_1, \dots, \eta_n)$ 是标准正交基。因此是正交矩阵乘上三角阵

 **练习 8.3QR 分解** 将下面的矩阵 A 分解为正交矩阵 T 与对角都是正数的上三角矩阵 B 的乘积

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

解 先进行 $\alpha_1, \alpha_2, \alpha_3$ 正交化到 $\beta_1, \beta_2, \beta_3$:

$$\begin{cases} \beta_1 = \alpha_1 = (1, 1, 1)^T, |\beta_1| = \sqrt{3} \\ \beta_2 = \alpha_2 - \frac{3}{3}(1, 1, 1)^T = (0, -1, 1)^T, |\beta_2| = \sqrt{2} \\ \beta_3 = \alpha_3 + \frac{1}{2}\beta_2 - \frac{2}{3}\beta_1 = (\frac{1}{3}, -\frac{1}{6}, -\frac{1}{6})^T, |\beta_3| = \sqrt{\frac{1}{6}} \end{cases}$$

进行单位化得到

$$\begin{cases} \eta_1 = (\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}) \\ \eta_2 = (0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \\ \eta_3 = (\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}) \end{cases}$$

反向得到

$$\begin{cases} \alpha_1 = \beta_1 \\ \alpha_2 = \beta_1 + \beta_2 \\ \alpha_3 = \frac{2}{3}\beta_1 - \frac{1}{2}\beta_2 + \beta_3 \end{cases}$$

因此

$$(\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2, \beta_3) \begin{pmatrix} 1 & 1 & \frac{2}{3} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} = (\eta_1, \eta_2, \eta_3) \begin{pmatrix} |\beta_1| & & \\ & |\beta_2| & \\ & & |\beta_3| \end{pmatrix} \begin{pmatrix} 1 & 1 & \frac{2}{3} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

8.1.3 正交和与正交补

定义 8.4 (正交和、正交补)

$V_1, \dots, V_s$ 为一列子空间, 若 $\forall i, j$ 有 $V_i \perp V_j$, 则称 $V_1 + \dots + V_s$ 为正交和。若 $V = V_1 + V_2$ 且为正交和, 则 V_1, V_2 互为正交补。



定理 8.6 (正交和必为直和)

若 $V_1 + \dots + V_s$ 为正交和, 则其必为直和。



证明 设 $0 = \alpha_1 + \dots + \alpha_s$, 两侧与 α_i 做内积可知

$$0 = (\alpha_i, \alpha_i)$$

因此 $\alpha_i = 0$, 故零元素分解唯一, 为直和。

定理 8.7 (正交补的唯一性)

Euclid 空间 V 的任一子空间 V_1 有唯一的正交补 $W^\perp$



证明 $V_1 = \{0\}$ 时, 即 $V_1^\perp = V$ 。否则设 $V_1 = L(\epsilon_1, \dots, \epsilon_s)$ 为标准正交基, 将其扩充为 V 的标准正交基:

$$V = L(\epsilon_1, \dots, \epsilon_n)$$

则 $V_1^\perp = L(\epsilon_{s+1}, \dots, \epsilon_n)$ 为正交补

唯一性: 若 V_1 有正交补 V_2, V_3 , 则 $\forall \alpha \in V_2 \subset V_1 \oplus V_3, \exists \alpha_1 \in V_1, \alpha_3 \in V_3$ 使得 $\alpha = \alpha_1 + \alpha_3$, 而 $\alpha \perp \alpha_1, \alpha_1 \perp \alpha_3$, 因此

$$0 = (\alpha, \alpha_1) = (\alpha_1, \alpha_1) + (\alpha_3, \alpha_1) = (\alpha_1, \alpha_1)$$

即 $\alpha_1 = 0, \alpha = \alpha_3 \in V_3$, 得到 $V_2 \subseteq V_3, V_3 \subseteq V_2$, 因此 $V_2 = V_3$

定理 8.8 (正交补的具体表达式)

V_1 是线性空间 V 的子空间, 则 $V_1^\perp$ 恰由所有与 V_1 正交的向量组成, 即

$$V_1^\perp = \{\alpha \in V, \alpha \perp V_1\}$$



证明 令 $W = \{\alpha \in V : \alpha \perp V_1\}$, 下面证明 $V_1^\perp = W$ 。首先 $V_1 \perp V_1^\perp$, 因此 $V_1^\perp \subseteq W$ 。

反之, $\forall \alpha \in W$, 根据 $V = V_1 \oplus V_1^\perp$, 因此可分解 $\alpha = \alpha_1 + \alpha_2$, 得到

$$(\alpha_1, \alpha_1) = (\alpha - \alpha_2, \alpha_1) = (\alpha, \alpha_1) - (\alpha_2, \alpha_1) = 0$$

因此 $\alpha_1 = 0$, 即 $\alpha \in V_1^\perp$

定理 8.9 (正交补的性质)

$$(1)(U^\perp)^\perp = U \quad (2)(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp \quad (3)(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$$



证明 (1) 根据正交补的具体表达式一下得出

(2) 显然 $(U_1 + U_2)^\perp \subseteq U_1^\perp, (U_1 + U_2)^\perp \subseteq U_2^\perp$, 因此 $(U_1 + U_2)^\perp \subseteq U_1^\perp \cap U_2^\perp$ 。

另一方面, $\forall \alpha \in U_1^\perp \cap U_2^\perp, \beta \in U_1 + U_2$, 有 $\beta = \beta_1 + \beta_2$, 因此

$$(\alpha, \beta) = (\alpha, \beta_1 + \beta_2) = (\alpha, \beta_1) + (\alpha, \beta_2) = 0$$

因此 $U_1^\perp \cap U_2^\perp \subseteq (U_1 + U_2)^\perp$

(3) 可转换为 (2)

8.1.4 正交变换

定义 8.5 (正交变换)

V 是 n 维 Euclid 空间, $\mathcal{A}$ 是 V 上线性变换, 若 $\forall \alpha, \beta \in V$ 有 $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta)$, 则称 $\mathcal{A}$ 为正交变换



定理 8.10 (正交变换充要条件)

$\mathcal{A}$ 是正交变换与下面几个命题等价:

- $\mathcal{A}$ 在标准正交基下矩阵为正交矩阵
- 标准正交基在 $\mathcal{A}$ 的作用下仍为标准正交基
- $\mathcal{A}$ 保持向量长度不变
- 若 $\epsilon_1, \dots, \epsilon_n$ 是 V 的标准正交基, 则 $\mathcal{A}(\epsilon_1), \dots, \mathcal{A}(\epsilon_n)$ 也是 V 的标准正交基



8.1.5 镜面反射

定义 8.6 (镜面反射)

若对 $\forall \alpha \in V$, $\mathcal{A}: V \rightarrow V$ 满足

$$\mathcal{A}\alpha = \alpha - 2(\eta, \alpha)\eta$$

其中 η 是单位向量



引理 8.1 (镜面反射是正交变换)

镜面反射是正交变换



证明

定理 8.11 (镜面反射的特征)



8.2 Unitary 空间

8.2.1 Unitary 空间概念

定义 8.7 (Unitary 空间)

V 是 $\mathbb{C}$ 上内积空间, 其上定义内积 (α, β) 满足 $\forall k_1, k_2 \in \mathbb{C}$

- $(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1(\alpha_1, \beta) + k_2(\alpha_2, \beta)$
- $(\alpha, \beta) = \overline{(\beta, \alpha)}$
- $(\alpha, \alpha) \geq 0$ 且 $(\alpha, \alpha) = 0$ 当且仅当 $\alpha = 0$



定义 8.8 (酉矩阵)

$\mathbb{C}$ 上矩阵 $U^{-1} = \overline{U}^T$, 则称 U 为酉矩阵。



定义 8.9 (酉变换)

$\mathcal{U}$ 满足 $(\mathcal{U}\alpha, \mathcal{U}\beta) = (\alpha, \beta)$, 则称 $\mathcal{U}$ 为酉变换



定义 8.10 (共轭变换)

$\mathcal{A}$ 是酉空间 V 的线性变换, 若 $\mathcal{A}^*$ 满足 $(\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}^*\beta)$, 则称 $\mathcal{A}^*$ 是 $\mathcal{A}$ 的共轭变换



定义 8.11 (正规变换)

若酉空间 V 的线性变换 $\mathcal{A}$ 与其共轭变换 $\mathcal{A}^*$ 可交换, 则称 $\mathcal{A}$ 为正规变换。



定义 8.12 (Hermite 变换)

若酉空间 V 的线性变换 $\mathcal{A}$ 满足 $\mathcal{A} = \mathcal{A}^*$, 则称 $\mathcal{A}$ 为 Hermite 变换



第9章 对偶空间

9.1 对偶空间与对偶基

定义 9.1 (对偶空间)

V 是 K 上的线性空间, 将定义于 V 上的所有有界线性泛函组成的空间称为对偶空间 V^* 。

定义 9.2 (对偶基)

给定 V 的一组基 $\epsilon_1, \dots, \epsilon_n$, 定义线性泛函 f_i 满足:

$$f_i(\epsilon_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

则 $\{f_1, \dots, f_n\}$ 是 V^* 的一组基, 称其为 $\epsilon_1, \dots, \epsilon_n$ 的对偶基。

 **笔记** 计算对偶基时, 将 f_i 视为一个向量, 即 $f_i(\alpha) = f_{i1}x_1 + \dots + f_{in}x_n = f_i^T \alpha$ 。计算 $\epsilon_1, \dots, \epsilon_n$ 的对偶基只需要满足:

$$\begin{bmatrix} f_1^T \\ f_2^T \\ \vdots \\ f_n^T \end{bmatrix} \begin{bmatrix} \epsilon_1 & \epsilon_2 & \cdots & \epsilon_n \end{bmatrix} = E$$

 **练习 9.1** 给定 $\epsilon_1 = [1, 1, 0]^T, \epsilon_2 = [0, 1, 1]^T, \epsilon_3 = [0, 1, -1]^T$, 求其对偶基

解 显然 f_1, f_2, f_3 都是三个系数组成的线性函数, 不妨记 $A = [f_1, f_2, f_3]^T, P = [\epsilon_1, \epsilon_2, \epsilon_3]$, 则

$$AP = \begin{bmatrix} f_1^T \\ f_2^T \\ f_3^T \end{bmatrix} [\epsilon_1, \epsilon_2, \epsilon_3] = E$$

因此 $A = P^{-1}$, 只需求出 P^{-1} 即可。最终求出:

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \Rightarrow \begin{cases} f_1^T = (1, 0, 0) \\ f_2^T = (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \\ f_3^T = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) \end{cases}$$