



高代

作者: Huang

时间: May 31, 2023

这可能是你见过最全的一个高代资料

目录

第 1 章 线性映射	1
1.1 导学	1
1.2 前置知识	1
1.3 线性变换中的重点题型	8
1.3.1 利用矩阵进行证明	8
1.3.2 利用基或定义进行证明	10
1.3.3 乘积可交换的变换	11
1.3.4 幂等变换	13
1.3.5 幂零变换	15
第 2 章 多项式	17
2.1 导学	17
2.2 前置知识	17
2.3 复系数与实系数多项式	22
2.4 有理系数和整系数多项式	25
2.5 重点题型	26
2.5.1 最大公因式总结	26
2.5.2 中国剩余定理	28
2.5.3 重因式	28
2.5.4 有理根和有理数域上的可约性	29
第 3 章 特征值	31
3.1 导学	31
3.2 特征值和特征多项式	31
3.2.1 前置知识	31
3.2.2 特征向量的结构	32
3.2.3 相似矩阵	32
3.2.4 秩一矩阵	33
3.2.5 伴随矩阵	34
3.2.6 AB 与 BA	34
3.3 可对角化	34
3.4 极小多项式	35
第 4 章 相似标准形	36
4.1 导学	36
4.2 前置知识	36
4.3 Frobenius 标准形和 Jordan 标准形	37
第 5 章 欧氏空间	40
5.1 导学	40
5.2 前置知识	40
5.3 正交矩阵性质总结	44
5.4 实对称矩阵性质总结	46

第 6 章 二次型	47
6.1 引言	47
6.2 前置知识	47
6.3 重点题型归纳	52
6.3.1 求二次型对应的矩阵	52
6.3.1.1 矩阵可直接求	53
6.3.1.2 矩阵不可直接求	55
6.3.2 合同变换法	61
6.3.2.1 对角矩阵过渡	61
6.3.3 正定二次型	63
6.3.3.1 判定正定二次型或者利用正定二次型的性质	63

第 1 章 线性映射

1.1 导学

线性映射是从一个线性空间再到另一个线性空间的保持加法和数量乘法的映射，这一章我们主要围绕线性映射的运算，核与像以及线性映射和线性变换的矩阵表示展开，这个就是我们研究线性映射的主线。

1.2 前置知识

首先，我们来了解一下映射的概念

定义 1.1

设 S, T 为非空集合，映射 $\varphi: S \rightarrow T$ 是指一个对应法则， $\forall \alpha \in S, \exists$ 唯一的 $\beta \in T$ 与之对应，记为 $\varphi(\alpha) = \beta$ 或 $\varphi: \alpha \mapsto \beta$ 。称 β 为 α 在 φ 下的像， α 称为 β 在 φ 下的原像。

在这样的对应法则下，我们就会遇到很多种对应情况，其中有两种比较特殊的情况。

定义 1.2 (单射)

设 $\varphi: S \rightarrow T$ 。若 $\forall \alpha_1, \alpha_2 \in S$ 且 $\alpha_1 \neq \alpha_2$ ，有 $\varphi(\alpha_1) \neq \varphi(\alpha_2)$ ，则称 φ 是单映射，简称单射。等价说法是，对 $\forall \alpha_1, \alpha_2 \in S$ ，若 $\varphi(\alpha_1) = \varphi(\alpha_2)$ ，必有 $\alpha_1 = \alpha_2$

当然，若一个线性映射的像中 0 元素的原项只有 0 元素，那么我们也可以推出这个映射是单射 (这是我们后面的知识)

定义 1.3 (满射)

设 $\varphi: S \rightarrow T$ 。若 $\forall \beta \in T, \exists \alpha \in S$ ，使得 $\varphi(\alpha) = \beta$ ，则称 φ 是满映射，简称满射。等价说法是， $Im\varphi = T$

如果一个映射即使单射，又是满射，那么我们称之为双射。

接着我们定义恒等映射

定义 1.4 (恒等映射)

$id_S: S \rightarrow S, a \mapsto a$ 是双射，称为 S 上的恒等映射

接下来我们来定义映射的合成

定义 1.5 (映射合成)

映射 $\varphi: S \rightarrow T$ 与 $\psi: T \rightarrow U$ 的合成定义为

$$\psi\varphi: S \rightarrow U, \alpha \mapsto \psi(\varphi(\alpha))$$

注: $\psi\varphi$ 不可写成 $\varphi\psi$ 。

注: 设 $\varphi: S \rightarrow T, \psi: T \rightarrow U, \rho: U \rightarrow V$ ，则有

(1) $(\rho\psi)\varphi = \rho(\psi\varphi)$;

(2) $\varphi = id_T\varphi = \varphi id_S$

然后我们继续定义逆映射 (知识点太多了，没办法)

定义 1.6 (逆映射)

设 $\varphi: S \rightarrow T$ 。若存在 $\psi: T \rightarrow S$, s.t. $\psi\varphi = id_S, \varphi\psi = id_T$, 则称 φ 是可逆映射, 并称 ψ 是 φ 的逆映射



如果一个映射是可逆映射的话, 按照它的定义, 则其像和原像必须是一一对应的, 即必须是单射, 然后我们要求这个映射的像都要能找到原像, 即必须是满射, 故其必须是双射, 由此我们得到以下定理

定理 1.1

$\varphi: S \rightarrow T$ 可逆的充要条件是 φ 是双射



注: 可逆映射的逆映射是唯一的。可逆映射 φ 的逆映射记为 φ^{-1}

注: $\varphi: S \rightarrow T$ 和 $\psi: T \rightarrow U$ 均可逆, 则 $\psi\varphi$ 也可逆, 且 $(\psi\varphi)^{-1} = \varphi^{-1}\psi^{-1}$ 。

然后我们还要注意一个结论

结论 设 $\varphi: S \rightarrow T$ 和 $\psi: T \rightarrow U$, 则有

(1) 若 $\psi\varphi$ 单, 则 φ 单; (2) 若 $\psi\varphi$ 满, 则 ψ 满; (3) 若 $\psi\varphi$ 双, 则 φ 单且 ψ 满

如何理解这一段话, 我们可以想象一下输入信号, 有一个输入器, 一个中转站, 一个显示屏, 满射代表显示屏所有的元素都亮了, 单射代表着我们输入一个数, 显示屏上都唯一对应一个数, 我们要想显示屏上元素全亮, 无论我们输入什么, 其都只是传输到中转站上, 最后的结果还是要看中转站和显示屏的关系, 此时必须要中转站的输出结果为显示屏上所有元素。类似的, 我们要想按一个数, 对应一个结果, 就必须保证输入器与中转站的元素是一一对应的, 这样就能保证我们输出的结果是唯一的。

下面一道例题, 特别的重要, 当成结论记下来, 对做题很有用

例题 1.1 设 $A \in F^{m \times n}, \varphi_A: F^n \rightarrow F^m, X \mapsto AX$ 。证明:

(1) $r(A) = n \Leftrightarrow \varphi_A$ 单; (2) $r(A) = m \Leftrightarrow \varphi_A$ 满

解:

(1) 要证明单射, 只要证明零元素原像为只0

即 $AX = 0$ 只有零解

又 $r(A) = n$, 故 $AX = 0$ 只有零解

故单射的充要条件为 $r(A) = n$

(2) $\varphi_A(F^n) = \{AX | X \in F^n\} = \langle A_1, A_2, \dots, A_n \rangle$

其中 A_1, A_2, \dots, A_n 为 A 的列向量, 若为满射, 则

$\dim \varphi_A(F^n) = m = \dim(A_1, A_2, \dots, A_n) = r(A)$

则 φ_A 为满射的充要条件是 $r(A) = m$

到此, 我们就结束了映射的基础定义。现在, 我们就开始我们对与线性映射研究的主线, 首先我们来了解一下何为线性映射

定义 1.7 (线性映射)

设 V, U 是 F 上线空, 映射 $\varphi: V \rightarrow U$ 若满足:

(1) $\forall \alpha, \beta \in V$ 有 $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$;

(2) $\forall \alpha \in V, k \in F$ 有 $\varphi(k\alpha) = k\varphi(\alpha)$, 则称 φ 是从 V 到 U 的线性映射



注: 设 V, U 是 F 上线空, 从 V 到 U 的线性映射全体构成的集合记为 $L(V, U)$ 。 $L(V, V)$ 中的元素称为 V 上的线性变换。 $L(V, V)$ 也记为 $L(V)$ 。

注: 设 V, U 是 F 上线空, $\varphi: V \rightarrow U$, 则 $\varphi \in L(V, U) \Leftrightarrow$ 对 $\forall \alpha_1, \alpha_2 \in V, k_1, k_2 \in F$, 有 $\varphi(k_1\alpha_1 + k_2\alpha_2) =$

注：设 V, U 是 F 上 n, m 维线空， $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的基， $\beta_1, \beta_2, \dots, \beta_m$ 是 U 的基， $\alpha \in V, \varphi \in L(V, U)$ 且 φ 在 $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_m$ 下矩阵为 A 。若 α 在 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下坐标为 X ，则 $\varphi(\alpha)$ 在 $\beta_1, \beta_2, \dots, \beta_m$ 下坐标为 AX 。

这个定理说明了，在确定一组基之后，一个线性映射可以用一个矩阵来表示且表示法唯一，这个结论就给我们提供了一个思路：线性映射就是在玩矩阵

我们知道，基和基之间是存在着过渡矩阵的，由上面的知识，我们知道，同一个线性映射在不同基下的矩阵是不同的，那么这几个不同的矩阵之间又有什么关系呢？接着就有以下定理

定理 1.2

设 V, U 是 F 上 n, m 维线空， $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\alpha'_1, \alpha'_2, \dots, \alpha'_n$ 是 V 的两个基且 $(\alpha'_1, \alpha'_2, \dots, \alpha'_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P$,

$\beta_1, \beta_2, \dots, \beta_m$ 和 $\beta'_1, \beta'_2, \dots, \beta'_m$ 是 U 的两个基且

$(\beta'_1, \beta'_2, \dots, \beta'_m) = (\beta_1, \beta_2, \dots, \beta_m)Q$

$\varphi \in L(V, U)$ 且

$$\varphi(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_m)A$$

$$\varphi(\alpha'_1, \alpha'_2, \dots, \alpha'_n) = (\beta'_1, \beta'_2, \dots, \beta'_m)B$$

则 $B = Q^{-1}AP$ ，即 A 相抵于 B 。反之，若 A 相抵于 B ，则 A, B 可看成同一个线性映射在两组不同基下的矩阵。



这下好了，在选定一组基下，线性映射与矩阵一一对应，在不同基下的矩阵是相抵的。矩阵这一个玩意，无疑就成了我们研究线性映射的神兵利器。

上文中，我们提到了，线性映射和矩阵是一个一一对应的关系，那么这种关系是什么关系呢？这就要来到我们接下来要讨论的一种关系——同构，首先来到它的定义

定义 1.11 (同构映射)

设 V, U 是 F 上线空， $\varphi \in L(V, U)$ ，若 φ 是双射，则称 φ 是一个（线性）同构映射，并称 V 和 U 是同构的线空，记为 $\varphi: V \cong U$



所以，我们就能得到下面的定理

定理 1.3

数域 F 上的 $n(n>0)$ 维线空同构于 F^n



这个定理表明，我们可以通过研究 F^n 来研究任何与其同构的线性空间，这属实方便了不少。那么，问题来了，该怎么样判断同构呢？下面的一个定理给了我们启发

定理 1.4

F 上两个有限维线空同构的充要条件是它们维数相同



这无疑就是给了我们判断同构的方法——只要维数相同，就同构。

让我们回到一开始的问题，线性映射和矩阵是一种什么关系呢？我们有以下定理

定理 1.5

设 V, U 是 F 上的 n, m 维线空，则 $L(V, U) \cong F^{m \times n}$



这个定理告诉我们，线性映射实际上是和矩阵是一个同构关系，研究了矩阵的性质，就是研究了这个线性映射的性质。于是乎，凭借着这一同构的关系，我们能求出一些平时求不出来的线性映射的维数

结论 设 V, U 是 F 上的 n, m 维线空, 则 $\dim L(V, U) = mn$

设 V 是 F 上的 n 维线空, 则 $\dim L(V) = n^2$

接下来我们要讨论的, 是线性映射中比较重要的两个子空间, 也就是我们的另一条主线, 线性空间的像与核, 其中衍生出了大量的结论, 希望大家要掌握, 我们首先来到定义

定义 1.12 (像与核)

设 V, U 是 F 上线空, $\varphi \in L(V, U)$, 则 $\text{Im}\varphi = \varphi(V)$ 是 U 的子空间, 称为 φ 的像, 并将 $\dim \text{Im}\varphi$ 称为 φ 的秩; $\text{Ker}\varphi = \varphi^{-1}(0)$ 是 V 的子空间, 称为 φ 的核, 并将 $\dim \text{Ker}\varphi$ 称为 φ 的零度



接着, 我们就有以下定理

定理 1.6

设 V, U 是 F 上的 n, m 维线空, $\alpha_1, \alpha_2, \dots, \alpha_n; \beta_1, \beta_2, \dots, \beta_m$ 分别是 V, U 的基, $\varphi \in L(V, U)$ 且 $\varphi(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_m)A$,

$$\sigma_1: V \rightarrow F^n, \alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \mapsto \sigma_1(\alpha) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

$$\sigma_2: U \rightarrow F^m, \beta = (\beta_1, \beta_2, \dots, \beta_m) \mapsto \sigma_2(\beta) = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

$$\varphi_A: F^n \rightarrow F^m, \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \text{ 则:}$$

- (1) $\sigma_2\varphi = \varphi_A\sigma_1$;
- (2) $\sigma_2(\text{Im}\varphi) = \text{Im}\varphi_A, \sigma_1(\text{Ker}\varphi) = \text{Ker}\varphi_A$;
- (3) $\dim \text{Im}\varphi = r(A), \dim \text{Ker}\varphi = n - r(A)$



由此, 我们就能得到重要的维数公式

结论 设 V, U 是 F 上的 n, m 维线空, $\varphi \in L(V, U)$,

则 $\dim \text{Im}\varphi + \dim \text{Ker}\varphi = n$

设 V 是 F 上的 n 维线空, $\varphi \in L(V)$,

则 $\dim \text{Im}\varphi + \dim \text{Ker}\varphi = n$

同时, 我们还能得到以下重要结论

结论 设 V, U 是 F 上的 n, m 维线空, $\varphi \in L(V, U)$, 则:

- (1) φ 是单射 $\Leftrightarrow \text{Ker}\varphi = 0 \Leftrightarrow r(A) = n$;
- (2) φ 是满射 $\Leftrightarrow \text{Im}\varphi = U \Leftrightarrow r(A) = m$ 。

以上的内容, 大致就为线性映射的主体框架, 现在, 我们来研究其中较为特殊的线性映射——线性变换。为了更好的研究, 我们需要引入代数的概念, 这个在我们近世代数会更为系统的学习。

定义 1.13 (代数)

设 V 是 F 上线空。如果在 V 上定义乘法 \circ , 满足:

- (1) $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma), \forall \alpha, \beta, \gamma \in V$;
- (2) $\alpha \circ (\beta + \gamma) = \alpha \circ \beta + \alpha \circ \gamma, (\alpha + \beta) \circ \gamma = \alpha \circ \gamma + \beta \circ \gamma, \forall \alpha, \beta, \gamma \in V$;
- (3) $k(\alpha \circ \beta) = (k\alpha) \circ \beta = \alpha \circ (k\beta), \forall \alpha, \beta \in V, k \in F$,

则称 V 是 F 上的代数。若还满足:

(4) $\exists e \in V$ s.t. 对 $\forall \alpha \in V$, 有 $e \circ \alpha = \alpha = \alpha \circ e$,

则称 V 是 F 上带单位元 e 的代数。

若满足 (1) – (3), 还满足:

(5) $\alpha \circ \beta = \beta \circ \alpha, \forall \alpha, \beta \in V$, 则称 V 是 F – 交换代数。

不满足 (5) 的代数称为非交换代数。



注:

带单位元的代数, 单位元唯一。

然后我们要引入代数同构的概念

定义 1.14 (代数同构)

设 V, U 是 F 上的两个代数。若存在线性空间同构映射 $\varphi: V \rightarrow U$ 满足:

$$\varphi(\alpha \circ \beta) = \varphi(\alpha) \circ \varphi(\beta), \forall \alpha, \beta \in V,$$

则称 φ 是 F 上代数同构 (映射), 称 V 与 U 代数同构。



在定义完一个东西之后, 我们最想了解其的运算法则, 在这里的话我们研究的是乘法。

定义 1.15 (乘法)

设 V 是 F 上线空, $\varphi, \psi \in L(V)$, $n \in \mathbb{N}^+$, , 定义 φ 与 ψ 的乘积为 φ 与 ψ 的合成,

$$\text{即: } \psi\varphi: V \rightarrow V, \alpha \mapsto \psi(\varphi(\alpha))$$

定义 φ^n 为 n 个 φ 的乘积。

定义 $\varphi^0 = id_V$



注:

设 V 是 F 上线空, $\varphi \in L(V)$, $n, m \in \mathbb{N}$, 则:

$$(1) \varphi^n \varphi^m = \varphi^{n+m};$$

$$(2) (\varphi^n)^m = \varphi^{nm}$$

然后我们就有以下定理

定理 1.7

设 V 是 F 上 n 维线空, $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的基, 如下定义从 $L(V)$ 到 $F^{n \times n}$ 的映射 T 设 $\varphi \in L(V)$ 且 $\varphi(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A$, 定义 $T(\varphi) = A$, 则 T 是代数同构



我们继续回到主线, 既然线性变换是一种特殊的线性映射, 那么它有什么特殊的性质呢?

结论 设 V 是 F 上 n 维线空, $\varphi \in L(V)$, 则下列条件等价:

(1) φ 可逆;

(2) φ 是同构映射;

(3) φ 单;

(4) φ 满;

(5) φ 在任意基下的矩阵是可逆阵

所以说, 在线性变换下, 单射就是满射, 这是别的线性映射没有的性质。这个在证明题的时候很好用。

此外, 我们想起, 线性变换在确定一组基的情况下, 是与矩阵同构的。同一个线性变换, 在不同基下的的矩阵是不同的, 那么这些矩阵之间又有什么关系呢?

我们有以下定理

定理 1.8

设 V 是 F 上 n 维线空, $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\beta_1, \beta_2, \dots, \beta_n$ 是 V 的两个基且 $(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P, \varphi \in L(V)$
 且 $\varphi(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A, \varphi(\beta_1, \beta_2, \dots, \beta_n) = (\beta_1, \beta_2, \dots, \beta_n)B$,
 则 $B = P^{-1}AP$ 。



我们发现, 同一个线性变换下, 不同基的矩阵满足上述关系, 我们称满足上述关系的矩阵是**相似的**, 于是我们就有下述定义

定义 1.16 (相似)

设 $A, B \in F^{n \times n}$ 。若存在可逆矩阵 P 使得 $B = P^{-1}AP$, 则称 A 相似于 B 。



有了这个, 我们立马能得到一个结论:

同一个线性变换在不同基下的矩阵是**相似的**

那么相似矩阵之间, 又有什么关系呢?

结论 设 $A, B \in F^{n \times n}$ 。若 A 相似于 B , 则有以下结论

1. A, B 有相同的秩

2. A, B 有相同的行列式

3. A, B 有相同的迹

注:

与 kE 相似的矩阵只能是 kE

注:

$$B = P^{-1}APB^n = P^{-1}A^nP.$$

注:

$$B = P^{-1}AP \Rightarrow f(B) = P^{-1}f(A)P, f(x) \in F[x].$$

注:

$$A \text{ 相似于 } B \Rightarrow f(A) \text{ 相似于 } f(B), f(x) \in F[x]$$

就此, 我们本章的主线也到此结束, 我们最后来研究一下在这个主线背景下的一条支线任务——不变子空间, 我们先来到它的定义

定义 1.17 (不变子空间)

设 V 是 F 上线空, U 是 V 的子空间, $\varphi \in L(V)$ 。若 $\varphi(U) \subset U$, 则称 U 是 φ -不变子空间或 φ -子空间。将 φ 限制在 U 上, 导出 U 的线性变换, 称为由 φ 导出变换 (或称为 φ 在 U 上的限制变换), 记为 $\varphi|_U$ 。



注:

φ 与 $\varphi|_U$ 的相同点是在 U 上的对应法则一样;

不同点是 $\varphi \in L(V)$, 而 $\varphi|_U \in L(U)$ 。

注:

定义中 U 是子空间这个条件不可少, 否则无法导出 U 上的线性变换 $\varphi|_U$ 。

按这个定义出发, 显然有

设 φ 是线空 V 的线性变换, 则 $\text{Ker}\varphi, \text{Im}\varphi$ 是 φ -子空间。

我们还能得到一个结论

结论 设 V 是 F 上 n 维线空 $\varphi \in L(V)$, U 是 V 的 φ -子空间。设 $\alpha_1, \alpha_2, \dots, \alpha_r$ 是 U 的一个基, 将其扩充为 V

的一个基 $\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n$, 则 φ 在此基下的矩阵为

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{pmatrix}.$$

反之, 若 φ 在基 $\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n$ 下的矩阵为上述形式矩阵, 则 $\langle \alpha_1, \alpha_2, \dots, \alpha_r \rangle$ 是一个 φ -子空间

到此为止, 线性映射的大体知识框架便搭建完成, 如果需要更为细致地填充, 大概还需看老师的讲义中的思考题。接下来我们就来到题型部分的总结。

1.3 线性变换中的重点题型

1.3.1 利用矩阵进行证明

在这一个部分, 常常是遇到了一些用线性变换知识无法解决或很难解决的问题, 我们通常考虑利用在确立一组基后, 一个线性变换同构于一个矩阵来进行解题, 此外在这里, 你还会看到大量的计算题。

例题 1.2 设 V, U 是 F 上 n, m 维线空, $\varphi \in L(V, U)$, 证: 存在 V 的一个基 $\xi_1, \xi_2, \dots, \xi_n$ 和 U 的一个基 $\eta_1, \eta_2, \dots, \eta_m$, 使得 $\varphi(\xi_1, \xi_2, \dots, \xi_n) = (\eta_1, \eta_2, \dots, \eta_m) \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ 。

解

不妨设 $\varphi(\xi'_1, \xi'_2, \dots, \xi'_n) = (\beta_1, \beta_2, \dots, \beta_m) B_{m \times n}$, 我们再观察结论的形式, 我们得想办法向这个结论靠拢。

遇事不决, 相抵标准形, 不妨设 $\text{rank}(B) = r$ 即存在 m 阶可逆阵 P, n 阶可逆阵 Q , 使得

$$PB_{m \times n}Q = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

于是我们里结论越来越近了, 不妨假设

$$(\xi_1, \xi_2, \dots, \xi_n)X = (\xi'_1, \xi'_2, \dots, \xi'_n), (\eta_1, \eta_2, \dots, \eta_m)Y = (\beta_1, \beta_2, \dots, \beta_m)$$

代入, 可得

$$YB_{m \times n}X = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

比较系数, 可得

$$Y^{-1} = P, X = Q$$

故存在 V 的一个基 $\xi_1, \xi_2, \dots, \xi_n$ 和 U 的一个基 $\eta_1, \eta_2, \dots, \eta_m$, 满足

$$(\xi_1, \xi_2, \dots, \xi_n)Q = (\xi'_1, \xi'_2, \dots, \xi'_n), (\eta_1, \eta_2, \dots, \eta_m)P^{-1} = (\beta_1, \beta_2, \dots, \beta_m)$$

使得

$$\varphi(\xi_1, \xi_2, \dots, \xi_n) = (\eta_1, \eta_2, \dots, \eta_m) \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

例题 1.3 设 V, U 是 F 上的 n, m 维线空 $\varphi \in L(V, U)$, 求证: $\exists \psi \in L(U, V)$, s.t. $\varphi\psi\varphi = \varphi, \psi\varphi\psi = \psi$

解

在确定一组基后, 设 φ 对应的矩阵为 $A_{m \times n}, \psi$ 对应的矩阵为 $B_{n \times m}$

即证明 $ABA = A, BAB = B$,

无从下手, 遇事不决, 相抵标准形, 存在 n 阶可逆阵 P, m 阶可逆阵 Q , 使得

$$QA_{m \times n}P = \begin{pmatrix} E_{r(A)} & 0 \\ 0 & 0 \end{pmatrix}$$

令

$$P \begin{pmatrix} E_{r(A)} & 0 \\ 0 & 0 \end{pmatrix} Q$$

则有 $ABA = A, BAB = B$

由线性映射的同构关系, 结论得证。

例题 1.4 设 V, U 是 F 上的 n, m 维线空 $\varphi, \psi \in L(V, U)$, $\text{Ker} \varphi \subset \text{Ker} \psi$, 证明: $\exists \sigma \in L(U, U)$, s.t. $\psi = \sigma \varphi$

解

在确定一组基后, 设 φ 对应的矩阵为 A, ψ 对应的矩阵为 B

由题, 即

$$AX = 0, \text{ 和 } \begin{pmatrix} A \\ B \end{pmatrix} X = 0 \text{ 同解}$$

于是有

$$\text{rank}(A) = \text{rank} \begin{pmatrix} A \\ B \end{pmatrix}$$

于是存在 m 阶可逆阵 C 使得

$$B = CA$$

由线性映射的同构, 令

$$C \cong \sigma$$

故存在

例题 1.5 设 V 是 3 维线空, $\varphi \in L(V)$ 且 φ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵为 $\begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$, 求证: $U = \langle \alpha_3, \alpha_1 + \alpha_2 + 2\alpha_3 \rangle$ 是 φ -子空间

解

因为

$$\varphi(\alpha_3) = -\alpha_1 - \alpha_2 = 2\alpha_3 - (\alpha_1 + \alpha_2 + 2\alpha_3) \in U$$

$$\varphi(\alpha_1 + \alpha_2 + 2\alpha_3) = (3\alpha_1 + 2\alpha_2 + 2\alpha_3) + (\alpha_1 + 2\alpha_2 + 2\alpha_3) - 2(\alpha_1 + \alpha_2) = 2(\alpha_1 + \alpha_2 + 2\alpha_3) \in U$$

所以 $\varphi(U) \subseteq U$, 即 U 是 φ -子空间

例题 1.6 设 V, U 是 F 上的 5, 4 维线空 $\alpha_1, \dots, \alpha_5; \beta_1, \dots, \beta_4$ 分别是 V, U 的基, $\varphi \in L(V, U)$ 且 $\varphi(\alpha_1, \dots, \alpha_5) =$

$$(\beta_1, \dots, \beta_4) \begin{pmatrix} 1 & 2 & 1 & -3 & 2 \\ 2 & 1 & 1 & 1 & -3 \\ 1 & 1 & 2 & 2 & -2 \\ 2 & 3 & -5 & -17 & 10 \end{pmatrix}, \text{ 求 } \text{Im} \varphi, \text{Ker} \varphi$$

解

令

$$A = \begin{pmatrix} 1 & 2 & 1 & -3 & 2 \\ 2 & 1 & 1 & 1 & -3 \\ 1 & 1 & 2 & 2 & -2 \\ 2 & 3 & -5 & -17 & 10 \end{pmatrix}$$

对矩阵作行初等变换, 有

$$B = \begin{pmatrix} 1 & 0 & 0 & 1 & -\frac{9}{4} \\ 0 & 1 & 0 & -3 & \frac{11}{4} \\ 0 & 0 & 1 & 2 & -\frac{5}{4} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

我们可以知道 $r(A) = 3$, 由于行初等变换不改变列向量的线性相关性, 于是乎, A 的前三列线性无关, 于是

$$\text{Im}\varphi = \langle \beta_1 + 2\beta_2 + \beta_3 + 2\beta_4, 2\beta_1 + \beta_2 + \beta_3 + 3\beta_4, \beta_1 + \beta_2 + 2\beta_3 - 5\beta_4 \rangle$$

解 $AX = 0$, 得到基础解系

$$\alpha_1 = \begin{pmatrix} -1 \\ 3 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 9 \\ -11 \\ 5 \\ 0 \\ 4 \end{pmatrix}$$

所以

$$\text{Ker}\varphi = \langle -\beta_1 + 3\beta_2 - 2\beta_3 + \beta_4, 9\beta_1 + 11\beta_2 + 5\beta_3 + 4\beta_4 \rangle$$

例题 1.7 设 V 是 n 维线性空间, U 是 m 维线性空间, $\varphi, \psi: V \rightarrow U$ 是线性映射。求证:

$\text{ker}\varphi = \text{ker}\psi$ 的充要条件是存在可逆变化 $\theta \in \mathcal{L}(U)$, 使得 $\varphi = \theta\psi$

解

在确定一组基后, 设 φ 对应的矩阵为 A , ψ 对应的矩阵为 B

有题有, $AX = 0$ 与 $BX = 0$ 同解 \iff

$$r(A) = r \begin{pmatrix} A \\ B \end{pmatrix} = r(B)$$

\iff 存在可逆矩阵 C 使得 $A = CB$

令

$$C \cong \theta$$

由线性映射的同构关系, 结论得证。

1.3.2 利用基或定义进行证明

这种类型的题目, 通常就按定义来就行了, 当然, 有的时候得注意一下, 我们在用基进行证明的时候, 常常要用到扩基定理。

例题 1.8 设 φ 是 n 维线性空 V 的线性变换, U 是 V 的 φ -子空间。若 φ 可逆, 求证:

(1) $\varphi|_U$ 可逆;

(2) U 是 φ^{-1} -子空间, 且 $(\varphi|_U)^{-1} = \varphi^{-1}|_U$ 。

解

(1) φ 可逆, 则有

$$\text{Ker}\varphi = 0$$

$\forall \alpha \in U$, 当 $\varphi|_U(\alpha) = 0$ 时, 有

$$\alpha = 0$$

故 $\varphi|_U$ 可逆

(2) $\forall \alpha \in U, \exists \beta \in U$ 使得

$$\varphi|_U(\alpha) = \beta = \varphi(\alpha)$$

由 $\varphi|_U, \varphi$ 可逆, 可得

$$\varphi^{-1}(\beta) = (\varphi|_U)^{-1}(\beta) = \alpha \in U$$

于是 U 是 φ^{-1} -子空间, 且 $(\varphi|_U)^{-1} = \varphi^{-1}|_U$

例题 1.9 设 φ 是 n 维线性空间 V 的线性变换, 则

1. $\text{Ker}\varphi \subset \text{Ker}\varphi^2 \subset \text{Ker}\varphi^3 \subset \cdots \subset \text{Ker}\varphi^m \subset \cdots$;
2. $\text{Im}\varphi \supset \text{Im}\varphi^2 \supset \text{Im}\varphi^3 \supset \cdots \supset \text{Im}\varphi^m \supset \cdots$;
3. 存在 $s \in \mathbb{N}^+$, 使得 $\text{Ker}\varphi^s = \text{Ker}\varphi^{s+1}$;
4. 存在 $t \in \mathbb{N}^+$, 使得 $\text{Im}\varphi^t = \text{Im}\varphi^{t+1}$;
5. 若 $\text{Ker}\varphi^s = \text{Ker}\varphi^{s+1}$, 则对 $i \in \mathbb{N}$, 有 $\text{Ker}\varphi^s = \text{Ker}\varphi^{s+i}$;
6. 若 $\text{Im}\varphi^t = \text{Im}\varphi^{t+1}$, 则对 $i \in \mathbb{N}$, 有 $\text{Im}\varphi^t = \text{Im}\varphi^{t+i}$;
7. $\text{Ker}\varphi^s = \text{Ker}\varphi^{s+1} \Leftrightarrow \text{Im}\varphi^s = \text{Im}\varphi^{s+1}$;
8. 若 $\text{Ker}\varphi^s = \text{Ker}\varphi^{s+1}$, 则 $V = \text{Ker}\varphi^s \oplus \text{Im}\varphi^s$.

解

看看就完事, 我懒得写了。

1.3.3 乘积可交换的变换

在做题的时候, 我们通常会遇到这种等式 $AB = BA$ 我们称之为乘积可交换, 那么乘积可交换到底有何性质呢? 我们分别从线性映射和矩阵的角度进行讨论。这一部分难度可能会较大, 因为涉及到的内容较多。

我们先从线性空间的角度来看, 首先以这一道例题开始

例题 1.10 设 V 是 F 上线性空间, $\varphi, \psi \in \mathcal{L}(V)$ 且 $\varphi\psi = \psi\varphi$, W 是 V 的 φ 子空间, 则 $\psi(W)$ 也是 V 的 φ 子空间。

从这里, 我们能得到, 如果乘积可交换, 那么其可保持**不变子空间的传递性**。我们可以从下面一题得到我们的结论

例题 1.11 设 V 是 F 上线性空间, $\varphi, \psi \in \mathcal{L}(V)$ 且 $\varphi\psi = \psi\varphi$, 求证: $\text{Ker}\varphi$ 是 ψ 子空间, $\text{Im}\varphi$ 是 ψ 子空间

这道题正面验证了上一题我们所得到的结论, 同时, 这一题也是很重要的一个结论。我们知道, 特征子空间也是线性空间中的一个很重要的不变子空间, 接着一道例题是关于它的

例题 1.12 设 V 是 F 上线性空间, $\varphi, \psi \in \mathcal{L}(V)$ 且 $\varphi\psi = \psi\varphi$, 设 V_{λ_0} 是属于 φ 的属于特征值 λ_0 的特征子空间, 则 V_{λ_0} 也是 ψ 子空间

我们用接下来的一题来宣告线性映射部分的结束

例题 1.13 设 V 是 n 维复线性空间, $\varphi, \psi \in \mathcal{L}(V)$ 且 $\varphi\psi = \psi\varphi$, 若 φ, ψ 均可对角化证明: φ, ψ 可在同一组基下对角化

解: 我们知道, 如果 φ 可对角化, 则 V 可表示为:

$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_p}$, 其中 $\lambda_1, \cdots, \lambda_p$ 为 φ 的所有特征值。由上面的结论, 我们可以知道, φ 的特征子空间也是 ψ 的特征子空间, 而对于任意特征子空间 V_{λ_i} ($i = 1, 2, \cdots, p$) 我们可选取其对应的一组基 $(e_{i1}, e_{i2}, \cdots, e_{im})$, 其中 $\dim V_i = m$ 在这组基下, 属于特征子空间 V_{λ_i} 的均可表示为对角阵

我们将 i 推广至 p , 于是有在 $(e_{11}, \cdots, e_{1m_1}, e_{21}, \cdots, e_{pm_p})$ 下, φ, ψ 可对角化, 其中 $m_1 + \cdots + m_p = n$

接着我们来到矩阵的部分, 在这一部分的话, 其和特征值与特征向量是密切相关的。我们知道, 在确定一组基之后, 线性映射是与矩阵实际上是同构的关系, 在这样一种同构的关系下, 矩阵的性质就是线性映射的性质。那么, 在上述性质之外, 我们还能挖掘出上面别的性质呢? 我们先以下面一道例题为开场。

例题 1.14 设 \mathbb{C} 上方阵 A, B 满足 $AB = BA$, 则 A, B 至少有一个公共特征向量

解

定义在选定一组基后, φ 对应的矩阵为 A , ψ 对应的矩阵为 B , 且满足 $\varphi\psi = \psi\varphi$

由线性映射的乘积可交换性质, 我们可以知道, φ 的特征子空间也是 ψ 的不变子空间。

设 λ_0 为 φ 的一个特征值, 于是就有 φ 的特征子空间 V_{λ_0} 也是 ψ 的不变子空间

于是乎 $\forall \xi \in V_{\lambda_0}$ 我们有, $\varphi(\xi) = \lambda_0 \xi$

由于 V_{λ_0} 也是 ψ 的不变子空间

则 $\psi|_{V_{\lambda_0}}$ 作用于 ξ 必有特征值。

因此, ψ, φ 至少有一个公共特征向量, 由于线性映射的同构关系, A, B 至少有一个公共特征向量

由于属于不同特征值的特征向量必线性无关, 于是乎我们可以加强一下命题

例题 1.15 设 \mathbb{C} 上方阵 A, B 满足 $AB = BA$, 若 A 至少有 s 个两两不同的特征值, 则 A, B 至少有 s 个公共特征向量, 且它们线性无关。

这道题的证明思路与上题类似, 不多加以赘述。我们甚至还能加强到多个方阵可交换的情景, 在这不多加讨论。

下面我们来到一个很有意思的例题

例题 1.16 设 F 上 n 阶矩阵, A, B 满足 $AB = BA$, 且 $f_A(\lambda), f_B(\lambda)$ 的根全在 F 上, 则存在 n 阶可逆阵 P , 使得 $P^{-1}AP, P^{-1}BP$ 同时为上三角矩阵。

解

对阶数做数学归纳法

当 $n=1$ 时, 显然结论成立

设命题对 $n-1$ 阶矩阵成立, 下证对 n 阶矩阵仍成立

我们设 λ_1 为 A 的一个特征值, 由于 $AB = BA$, 我们可以得到 A 的特征子空间 V_{λ_1} 也是 B 的不变子空间, 于是对于 $X_1 \in V_{\lambda_1}$, 有

$$AX_1 = \lambda_1 X_1$$

$$BX_1 = \mu_1 X_1$$

定义在选定一组基后, φ 对应的矩阵为 A , ψ 对应的矩阵为 B , 且满足 $\varphi\psi = \psi\varphi$

我们把 X_1 扩充为 F^n 的一组基: X_1, X_2, \dots, X_n 令 $P_1 = (X_1, X_2, \dots, X_n)$, 则 P_1 可逆, 并且有

$$\varphi(X_1, X_2, \dots, X_n) = (X_1, X_2, \dots, X_n) \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix}$$

即 $P_1^{-1}AP_1 = \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix}$, 同理, 我们可以得到

$$\psi(X_1, X_2, \dots, X_n) = (X_1, X_2, \dots, X_n) \begin{pmatrix} \mu_1 & \beta \\ 0 & B_1 \end{pmatrix}$$

即 $P_1^{-1}BP_1 = \begin{pmatrix} \mu_1 & \beta \\ 0 & B_1 \end{pmatrix}$, 又 $\varphi\psi = \psi\varphi$, 于是有

$$\begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} \mu_1 & \beta \\ 0 & B_1 \end{pmatrix} = \begin{pmatrix} \mu_1 & \beta \\ 0 & B_1 \end{pmatrix} \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix}$$

于是, 我们有 $A_1B_1 = B_1A_1$, 由归纳假设, 存在 $n-1$ 阶矩阵 P_2 使得 $P_2^{-1}AP_2, P_2^{-1}BP_2$ 同时为上三角矩阵, 令

$$P = P_1 \begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix}$$

则 P 为 n 阶可逆阵, 且满足

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} P_1^{-1}AP_1 \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \alpha P_2 \\ 0 & P_2^{-1}A_1P_2 \end{pmatrix} \\ P^{-1}BP &= \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} P_1^{-1}BP_1 \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P_2^{-1} \end{pmatrix} \begin{pmatrix} \mu_1 & \beta \\ 0 & B_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} \mu_1 & \beta P_2 \\ 0 & P_2^{-1}B_1P_2 \end{pmatrix} \end{aligned}$$

则 $P^{-1}AP, P^{-1}BP$ 同时为上三角矩阵。由数学归纳法原理, 对一切正整数 n 都成立, 命题为真。

从这一题里我们可以推出一个很好玩的性质, 若乘积可交换, 只要有 n 个特征值, 就可以相似于一个上三

角矩阵，这个就蛮有意思的。由于考试不怎么考像这样有趣的，咱们这个部分就到此结束。

1.3.4 幂等变换

幂等变换，是一个特殊的线性变换，由于在确定一组基基后，线性映射同构于矩阵，此时幂等变换对应的矩阵我们称作幂等矩阵。接下来我们来研究一下这类的性质。

首先我们给出幂等矩阵的定义

定义 1.18 (幂等矩阵)

若 A 为方阵，且 $A^2 = A$ 我们便称之为幂等矩阵



其有以下性质

例题 1.17 幂等矩阵的特征值只有 0 或 1

证明：

设 λ 为 A 的特征值，则存在非零向量 α 使得

$$A\alpha = \lambda\alpha$$

同时乘 A , 有

$$\lambda\alpha = A\alpha = A^2\alpha = \lambda A\alpha = \lambda^2\alpha$$

移项，有

$$(\lambda^2 - \lambda)\alpha = 0$$

解得

$$\lambda = 1 \text{ 或 } 0$$

例题 1.18 数域 F 的幂等矩阵一定可对角化

证明：

因为

$$A^2 - A = 0$$

于是 A 的零化多项式为

$$x(x - 1) = 0$$

则其为极小多项式，有因为极小多项式没有重因式，故可对角化。

例题 1.19 若 A 为幂等矩阵，则 $\text{rank}(A) = \text{tr}(A)$

证明： A 可对角化，则有

$$P^{-1}AP = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

又相似矩阵有相同的迹，得证。接下来，我们来看看线性变换又有什么性质。

例题 1.20 设 V 是 F 上 n 维线性空间 $\varphi \in L(V)$ ，则 $\exists \psi, \sigma \in L(V)$ s.t. $\varphi = \psi\sigma$ ，其中 $\psi^2 = \psi$ ， σ 是可逆变换

解
选定一组基后，设 φ 对应的矩阵为 A ， ψ 所对应的基为 B ， σ 所对应的矩阵为 C
遇事不决，相抵标准形，存在 n 阶可逆阵 P, Q 使得

$$A = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} P^{-1} PQ$$

令

$$B = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} P^{-1}, C = PQ$$

显然满足题目要求, 由线性映射的同构, 得证。

例题 1.21 设 V 是 F 上 n 维线空, $\varphi \in L(V)$ 满足 $\varphi^2 = \varphi$ 且 $\dim \operatorname{Im} \varphi = r < n$ 。证明: 存在 V 的一个基 $\xi_1, \xi_2, \dots, \xi_n$ 使得

$$\varphi(\xi_1, \xi_2, \dots, \xi_n) = (\xi_1, \xi_2, \dots, \xi_n) \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

。

解

设在一组基下, φ 的矩阵为 A , 题设条件变为 $A^2 = A$

现在我们想要证明这个结论, 要知道, 在不同基下, 同一个线性映射在不同基下的矩阵是相似的, 所以要证明的就是 A 相似于 $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$, 即存在可逆阵 P 使得 $P^{-1}AP = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$

遇事不决, 相抵标准形, 即存在 n 阶可逆阵 P_1, Q_1 , 使得

$$A = P_1 \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_1$$

化简, 可得

$$P_1^{-1}AQ_1^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

即

$$P_1^{-1}AP_1^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_1P_1$$

令

$$Q_1P_1 = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

其中 X 为 r 阶方阵, 代入有

$$P_1^{-1}AP_1^{-1} = \begin{pmatrix} X & Y \\ 0 & 0 \end{pmatrix}$$

又 $A^2 = A$, 则有

$$\begin{pmatrix} X & Y \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} X & Y \\ 0 & 0 \end{pmatrix}$$

解得

$$X = E_r, Y = 0$$

于是有

$$P_1^{-1}AP_1^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

得证。

例题 1.22 设 V 是 F 上 n 维线空 $\varphi \in L(V)$ 满足 $\varphi^2 = \varphi$ 。证明:

$$V = \operatorname{Im} \varphi \oplus \operatorname{Ker} \varphi$$

解

因为 $\forall \alpha \in V$ 都有

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

显然有 $\varphi(\alpha) \in \text{Im}\varphi$, 又因为 $\varphi(\alpha - \varphi(\alpha)) = \varphi(\alpha) - \varphi^2(\alpha) = 0$

于是有

$$V = \text{Im}\varphi + \text{Ker}\varphi$$

下证 $\text{Im}\varphi \cap \text{Ker}\varphi = 0$

设 $\text{Im}\varphi \cap \text{Ker}\varphi = \alpha$, 因为 $\alpha \in \text{Ker}\varphi$ 于是有 $\varphi(\alpha) = 0$ 。又因为 $\alpha \in \text{Im}\varphi$, 故 $\exists \beta \in V$, 使得 $\varphi(\beta) = \alpha$, 即 $\varphi(\beta) = \varphi^2(\beta) = \alpha = \varphi(\alpha) = 0$

结论得证。

1.3.5 幂零变换

我们还有一个特殊的变换, 其为幂零变化, 同样的, 其也与矩阵同构, 相应的矩阵称为幂零矩阵, 我们来到定义

定义 1.19 (幂零矩阵)

对于一个 n 阶矩阵, 如果存在正整数 l , 使得 $A^l = 0$, 我们便称其为幂零矩阵



性质如下

例题 1.23 n 阶矩阵是幂零矩阵的充要条件是特征值全为 0

证明:

\Rightarrow :

设 A 为幂零矩阵, 设 A 的特征值为 λ , 则存在非零向量 α , 使得

$$A^l \alpha = \lambda^l \alpha = 0$$

解得

$$\lambda = 0$$

\Leftarrow :

设 A 的特征值全为零, 对于 A 的特征多项式, 我们有

$$|\lambda E - A| = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_0$$

由 Vieta 定理, 有

$$a_1 = \cdots = a_0 = 0$$

于是有

$$|\lambda E - A| = \lambda^n = 0$$

由 Hamilton-Cayley Theorem, 得证

例题 1.24 设 A 为数域 F 上的矩阵, 则 A 是幂零矩阵的充要条件是对于任意正整数 k , 都有 $\text{tr}(A^k) = 0$

证明:

\Rightarrow :

A 为幂零矩阵, 则 A 的特征值全为 0, A^k 的特征值也为 0, 由由于迹为所有特征值的和, 得证

\Leftarrow :

假设有 n 个特征值, 记为 $\lambda_1, \lambda_2, \dots, \lambda_n$ 由题, 有

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n = 0$$

$$\operatorname{tr}(A^2) = \lambda_1^2 + \lambda_2^2 + \dots + \lambda_n^2 = 0$$

...

$$\operatorname{tr}(A^n) = \lambda_1^n + \lambda_2^n + \dots + \lambda_n^n = 0$$

设 A 的所有非零互异的特征值是 $\lambda_1, \lambda_2, \dots, \lambda_r$, 其代数重数为 s_1, s_2, \dots, s_r

对于上述式子, 我们可以化简为

$$s_1 \lambda_1 + \dots + s_r \lambda_r = 0$$

$$s_1 \lambda_1^2 + \dots + s_r \lambda_r^2 = 0$$

...

$$s_1 \lambda_1^n + \dots + s_r \lambda_r^n = 0$$

将其看成关于 s 的方程, 由范德蒙行列式的性质, 此时只有 0 解, 故解得 A 的所有特征值为 0, 于是 A 为幂零矩阵。

例题 1.25 幂零矩阵不可对角化

证明:

$A^k = 0$, 故 A 的极小多项式次数至少为 2, 于是其有重根, 故不可对角化。

接下来, 我们再来看看线性映射的性质, 这个就很少了, 就一个题目, 很有趣的

例题 1.26 设 n 维度线性空间 V 上的线性变换 φ 满足 $\varphi^n = 0$, $\varphi^{n-1} \neq 0$. 求证:

存在 $\alpha \in V$, 使得 $\alpha, \varphi(\alpha), \dots, \varphi^{n-1}(\alpha)$ 线性无关。

解

自证

这个东西, 就是 Frobenius 标准型的循环基, 有意思的很。

第2章 多项式

2.1 导学

任给一个一元多项式，在给定数域情况下，我们该如何分解，又能分解成什么样？这个便是我们要研究的主线，由此，我们便在不同数域下对多项式进行讨论。

2.2 前置知识

首先，我们要了解的，肯定是一元多项式的定义。

定义 2.1 (一元多项式)

形如 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ ，我们便称其为称为 F 上关于 x 的一元多项式，其中 F : 数域； $a_i \in F, i = 0, 1, \cdots, n; n \in \mathbb{Z}_{\geq 0}$ ； x : 未定元
同时，我们记 F 上一元多项式全体为 $F[x]$

注

1. 当 $a_n \neq 0$ 时，次数记为 $\deg f(x) = n$ (定义 0 多项式次数为负无穷)
2. 两个多项式称为相等当且仅当它们的次数相同且各次项的系数相等，即若

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$$

则 $f(x) = g(x)$ 当且仅当 $m = n, a_i = b_i, i = 0, 1, \cdots, n$ 。

接着我们来定义一下其的运算法则

定义 2.2 (加法和数乘)

定义加法： $f(x) + g(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0)$

定义数乘： $cf(x) = ca_n x^n + ca_{n-1} x^{n-1} + \cdots + ca_1 x + ca_0$ 。

然后就这两个运算，我们能推出其满足以下性质：

结论 (1) 结合律： $(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$

(2) 交换律： $f(x) + g(x) = g(x) + f(x)$

(3) 存在零元： $f(x) + 0 = f(x)$ ，定义 $-f(x) = \sum_{i=0}^n (-a_i)x^i$ ，同样存在

(4) 存在负元： $f(x) + (-f(x)) = 0$

(5) $c(f(x) + g(x)) = cf(x) + cg(x)$

(6) $(c + d)f(x) = cf(x) + df(x)$

(7) $(cd)f(x) = c(df(x))$

(8) $1 \cdot f(x) = f(x)$

我们立即就能得到， $F[x]$ 关于多项式的加法与数乘构成 F 上的无限维线性空间，因为这个恰好就是那验证线性空间的那八条性质，这个全满足。

既然是线性空间，那我们就肯定要有基，于是，我们就有以下结论

结论 $1, x, x^2, \cdots, x^n, \cdots$ 是 $F[x]$ 的一个基

$1, x, \frac{x^2}{2!}, \cdots, \frac{x^n}{n!}, \cdots$ 也是 $F[x]$ 的一个基

然后我们要继续定义多项式之间的乘法

定义 2.3 (乘法)

$f(x)g(x) = h(x)$, 其中 $h(x) = c_{n+m}x^{n+m} + c_{n+m-1}x^{n+m-1} + \cdots + c_1x + c_0 \in F[x]$

$$c_{n+m} = a_nb_m$$

$$c_{n+m-1} = a_nb_{m-1} + a_{n-1}b_m$$

.....

$$c_k = \sum_{i=0}^n \sum_{j=0}^m \sum_{i+j=k} a_ib_j$$

.....

$$c_0 = a_0b_0$$



满足以下性质

结论 (9) $(f(x)g(x))h(x) = f(x)(g(x)h(x))$

$$(10) f(x)g(x) = g(x)f(x)$$

$$(11) (f(x) + g(x))h(x) = f(x)h(x) + g(x)h(x)$$

$$(12) c(f(x)g(x)) = (cf(x))g(x) = f(x)(cg(x)), c \in F$$

$$(13) 1 \cdot f(x) = f(x)$$

再定义完这些的加法和乘法后, 我们想看看在这样的运算法则下, 其次数又满足怎样的性质, 如下

结论 设 $f(x), g(x) \in F[x]$, 则

$$(1) \deg(f(x) + g(x)) \leq \max\{\deg f(x), \deg g(x)\}$$

$$(2) \deg(cf(x)) = \deg f(x), 0 \neq c \in F$$

$$(3) \deg(f(x)g(x)) = \deg f(x) + \deg g(x)$$

接下来的内容, 是课上关于这一部分的重点。

结论 若 $f(x) \neq 0$ 且 $f(x)g(x) = f(x)h(x)$, 则 $g(x) = h(x)$

这个告诉我们, 多项式的乘法满足消去律, 前提是消去的是不为零。

结论 若 $f(x), g(x) \in R[x]$ 且 $f^2(x) + g^2(x) = 0$, 则 $f(x) = g(x) = 0$ 。

我们要注意的, 这个结论在复数域上不成立。

在定义完这些基本的运算之后, 我们就可以来到我们的主线了, 首先, 我们先来到整除的概念。

定义 2.4 (整除)

设 $f(x), g(x) \in F[x]$ 。若存在 $h(x) \in F[x]$, 使得

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

则称 $g(x)$ 是 $f(x)$ 的因式, $f(x)$ 是 $g(x)$ 的倍式, 或称 $f(x)$ 能被 $g(x)$ 整除, 或称 $g(x)$ 整除 $f(x)$, 记为 $g(x)|f(x)$ 。否则称 $f(x)$ 不能被 $g(x)$ 整除,



注:

零多项式可以被任意多项式整除。

零多项式只能整除零多项式。

零次多项式可以整除任意多项式。

零次多项式只能被零次多项式整除。

整除关系, 是良性的关系, 其满足

结论 (1) 反身性: $f(x)|f(x)$

(2) 传递性: 若 $f(x)|g(x)$, $g(x)|h(x)$, 则 $f(x)|h(x)$ 。

(3) 互伴性: 若 $f(x)|g(x)$, $g(x)|f(x)$, 则存在 $0 \neq c \in F$ 使得 $f(x) = cg(x)$, 此时称 $f(x), g(x)$ 为相伴多项式, 记作 $f(x) \sim g(x)$ 。

(4) 若 $f(x)|g(x)$, $f(x)|h(x)$, 则对任意 $u(x), v(x) \in F[x]$,

有 $f(x)|u(x)g(x) + v(x)h(x)$

通常来讲，我们遇到的式子是不能整除的，于是我们就来到了一个重要的定理

定义 2.5 (带余除法)

设 $f(x), g(x) \in F[x]$ 且 $g(x) \neq 0$ ，则存在唯一的 $q(x), r(x) \in F[x]$ 使得

$$f(x) = q(x)g(x) + r(x)$$

其中 $\deg r(x) < \deg g(x)$ ，我们称 $q(x), r(x)$ 分别称为 $g(x)$ 除 $f(x)$ （或 $f(x)$ 除以 $g(x)$ ）的商式和余式。

注：

1. 带余除法与数域的扩张无关。

2. 整除关系与数域的扩张无关。

由整除关系，我们就自然过渡到了下一个版块——公因式，定义如下

定义 2.6 (公因式)

设 $f(x), g(x), d(x) \in F[x]$ ，若 $d(x)|f(x), d(x)|g(x)$ ，则称 $d(x)$ 是 $f(x), g(x)$ 的一个公因式

注：

1. 零次多项式是任意两个多项式的公因式

2. 若 $d(x)$ 是 $f(x), g(x)$ 的一个公因式，则对 $\forall 0 \neq k \in F$ ， $kd(x)$ 也是 $f(x), g(x)$ 的公因式

紧接着，我们便引出最大公因式的概念

定义 2.7 (最大公因式)

设 $f(x), g(x) \in F[x]$ ， $d(x)$ 是 $f(x), g(x)$ 的一个公因式。若对 $f(x), g(x)$ 的任意一个公因式 $h(x)$ 都有 $h(x)|d(x)$ ，则称 $d(x)$ 是 $f(x), g(x)$ 的一个最大公因式

注

1. 最大公因式不唯一，最多差一个非零常数倍，但两个不全为零的多项式 $f(x), g(x)$ 的首项系数为 1 的最大公因式是唯一确定的，记为 $d(x) = (f(x), g(x))$ 特别的， $(0, 0) = 0$

2. 最大公因式与数域的扩张有关

3. 首一最大公因式与数域的扩张无关。

那么问题来了，有什么好的方法来求最大公因式么？于是乎，就来到了一个优秀的算法，辗转相除法

定义 2.8 (辗转相除法)

设 $f(x), g(x) \in F[x]$ ，则 $f(x), g(x)$ 存在最大公因式 $d(x)$ ，且存在 $u(x), v(x) \in F[x]$ 使得 $d(x) = u(x)f(x) + v(x)g(x)$

注

1. 设 $f(x), g(x), d(x) \in F[x]$ 且 $d(x)$ 是 $f(x), g(x)$ 的公因式。若存在 $u(x), v(x) \in F[x]$ 使得 $d(x) = u(x)f(x) + v(x)g(x)$ ，则 $d(x)$ 是 $f(x), g(x)$ 的最大公因式

2. 定理中的 $u(x), v(x)$ 不唯一

接着我们将其推广到多个多项式的最大公因式

定义 2.9 (最大公因式)

设 $f_1(x), \dots, f_s(x) \in F[x]$, $d(x)$ 是 $f_1(x), \dots, f_s(x)$ 的一个公因式。若对 $f_1(x), \dots, f_s(x)$ 的任意一个公因式 $h(x)$ 都有 $h(x)|d(x)$, 则称 $d(x)$ 是 $f_1(x), \dots, f_s(x)$ 的一个最大公因式

注

1. 设 $f_1(x), \dots, f_s(x) \in F[x]$ 且 $f_1(x), \dots, f_s(x)$ 不全为零, 则 $f_1(x), \dots, f_s(x)$ 的首一最大公因式记为 $(f_1(x), \dots, f_s(x))$

2. $((f(x), g(x)), h(x)) = (f(x), g(x), h(x)) = (f(x), (g(x), h(x)))$

然后, 我们就到了一种特殊情况, 即首一最大公因式为 1 的时候, 在这个时候我们称之为互素, 就有以下定义

定义 2.10 (互素)

设 $f(x), g(x) \in F[x]$, 若 $(f(x), g(x)) = 1$, 则称 $f(x)$ 与 $g(x)$ 互素或互质

注

1. 互素与数域的扩大无关

2. 若 $f(x), g(x), h(x)$ 两两互素, 则 $(f(x), g(x), h(x)) = 1$, 反之不成立

最后, 我们仿照公因式定义公倍式

定义 2.11 (公倍式)

设 $f_1(x), \dots, f_s(x), c(x) \in F[x]$, 若 $f_i(x)|c(x) (i = 1, \dots, s)$, 则称 $c(x)$ 是 $f_1(x), \dots, f_s(x)$ 的一个公倍式

同理, 我们有最小公倍式

定义 2.12 (最小公倍式)

设 $f_1(x), \dots, f_s(x) \in F[x]$, $c(x)f_1(x), \dots, f_s(x)$ 的一个公倍式。若对 $f_1(x), \dots, f_s(x)$ 的任意一个公倍式 $h(x)$ 都有 $c(x)|h(x)$, 则称 $c(x)$ 是 $f_1(x), \dots, f_s(x)$ 的一个最小公倍式

注

1. 设 $f_1(x), \dots, f_s(x) \in F[x]$ 且 $f_1(x), \dots, f_s(x)$ 全不为零, 则 $f_1(x), \dots, f_s(x)$ 的首一最小公倍式, 记为 $[f_1(x), \dots, f_s(x)]$

接着, 我们来到最后一个工具——标准分解式, 这个与定义, 带余除法, 辗转相除法共称为证明的四大金刚。为了解释何为标准分解式, 我们首先要引入可约和不可约的概念。

定义 2.13 (可约性)

设 $f(x) \in F[x]$ 且 $\deg f(x) \geq 1$ 。若 $f(x)$ 能表为两个次数更低的多项式之积, 则称 $f(x)$ 是 F 上的可约多项式, 否则称为 F 上的不可约多项式。

注

1. 多项式可约不可约与数域有关。

2. F 上不可约多项式 $f(x)$ 的因式只能是 F 上非零常数及 $f(x)$ 的相伴多项式

3. 一次多项式必不可约。

接下来, 我们来到一个很重要的结论

结论 设 $p(x) \in F[x]$ 且 $\deg p(x) > 0$, 则 $p(x)$ 不可约 \Leftrightarrow 对任意 $f(x) \in F[x]$ 有 $(p(x), f(x)) = 1$ 或 $p(x)|f(x)$

这个结论告诉我们, 不可约多项式和别的多项式是一种非黑即白的关系, 要么整除, 要么就互素, 这个就很有趣。

接下来我们来到多项式唯一分解定理, 这个是标准分解式的基石。

定理 2.1 (唯一分解定理)

设 $f(x) \in F[x]$ 且 $\deg f(x) \geq 1$, 则

- (1) $f(x) = p_1(x)p_2(x) \cdots p_s(x)$, 其中 $p_i(x)$ 是 F 上不可约多项式 ($i = 1, \dots, s$);
- (2) 若 $f(x) = p_1(x)p_2(x) \cdots p_s(x) = q_1(x)q_2(x) \cdots q_t(x)$, 其中 $p_i(x), q_j(x)$ 是 F 上不可约多项式 ($i = 1, \dots, s, j = 1, \dots, t$), 则必有 $s = t$ 且经过适当调换因式顺序后, $p_i(x)$ 与 $q_i(x)$ 相伴 ($i = 1, \dots, s$)



由此, 我们能得到任意次数大于零的多项式标准分解式

定理 2.2

设 $f(x) \in F[x]$ 且 $\deg f(x) \geq 1$, 则

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

其中 $p_i(x) (i = 1, \dots, s)$ 是首一的两两互素的不可约多项式, $r_i \geq 1 (i = 1, \dots, s)$ 。



我们还能得到以下结论

结论 设 $f(x) = cp_1^{a_1}(x)p_2^{a_2}(x) \cdots p_s^{a_s}(x), g(x) = dp_1^{b_1}(x)p_2^{b_2}(x) \cdots p_s^{b_s}(x)$, 其中 $p_i(x) (i = 1, \dots, s)$ 是首一的两两互素的不可约多项式, $a_i \geq 0, b_i \geq 0, a_i + b_i > 0 (i = 1, \dots, s)$, 则

- (1) $f(x)g(x) = cd p_1^{a_1+b_1}(x)p_2^{a_2+b_2}(x) \cdots p_s^{a_s+b_s}(x)$;
- (2) $(f(x), g(x)) = p_1^{c_1}(x)p_2^{c_2}(x) \cdots p_s^{c_s}(x), c_i = \min\{a_i, b_i\} (i = 1, \dots, s)$;
- (3) $[f(x), g(x)] = p_1^{d_1}(x)p_2^{d_2}(x) \cdots p_s^{d_s}(x), d_i = \max\{a_i, b_i\} (i = 1, \dots, s)$;
- (4) $(f(x), g(x))[f(x), g(x)] = c^{-1}d^{-1}f(x)g(x)$;
- (5) $f(x)|g(x) \Leftrightarrow a_i \leq b_i (i = 1, \dots, s)$

事实上, 这玩意在我们实际计算的时候, 还真的不好用, 最多在证明题里使用。

由不可约多项式, 我们又可以衍生出一个新的概念——重因式。

定义 2.14 (重因式)

设 $f(x), p(x) \in F[x]$ 且 $p(x)$ 不可约。若存在 $k \geq 0$ 使得 $p^k(x)|f(x), p^{k+1}(x) \nmid f(x)$, 则称 $p(x)$ 是 $f(x)$ 的 k 重因式。当 $k = 0$ 时, $p(x)$ 不是 $f(x)$ 的因式; 当 $k = 1$ 时, $p(x)$ 称为 $f(x)$ 的单因式; 当 $k > 1$ 时, $p(x)$ 称为 $f(x)$ 的重因式



注

1. $p(x)$ 是否是 $f(x)$ 的重因式与数域的扩张有关;

2. 但 $f(x)$ 是否有重因式与数域的扩张无关

有几个结论需要我们记住

结论 不可约多项式 $p(x)$ 是 $f(x)$ 的 k 重因式 \Leftrightarrow 存在 $h(x)$ 使得 $f(x) = p^k(x)h(x)$ 且 $(p(x), h(x)) = 1$

结论 不可约多项式 $p(x)$ 是 $f(x)$ 的重因式 $\Leftrightarrow p(x)$ 是 $f(x)$ 和 $f'(x)$ 的公因式

非零多项式 $f(x)$ 无重因式 $\Leftrightarrow (f(x), f'(x)) = 1$

结论 设 $d(x) = (f(x), f'(x))$, 则 $f_1(x)$ 是一个无重因式的多项式, 且此多项式的每一个不可约因式与 $f(x)$ 的不可约因式相同

到此为止, 我们研究多项式的工具就结束了, 接下来的内容, 便是在已有工具下的补充。

定理 2.3 (余数定理)

设 $f(x) \in F[x], b \in F$, 则存在唯一的 $g(x) \in F[x]$ 使得 $f(x) = (x - b)g(x) + f(b)$



由此, 我们能得到一个算法——综合除法

定义 2.15 (综合除法)

设 $f(x) = a_n x^n + \cdots + a_1 x + a_0 = (x - b)(b_{n-1} x^{n-1} + \cdots + b_1 x + b_0) + f(b)$, 比对系数, 有
 $b_{n-1} = a_n b_{n-2} = a_{n-1} + b b_{n-1}, b_{n-3} = a_{n-2} + b b_{n-2}, \cdots, b_1 = a_2 + b b_2, b_0 = a_1 + b b_1, f(b) = a_0 + b b_0$

可表示为下图

	a_n	a_{n-1}	a_{n-2}	\cdots	a_1	a_0
b		$b_{n-1}b$	$b_{n-2}b$	\cdots	b_1b	b_0b
	b_{n-1}	b_{n-2}	b_{n-3}	\cdots	b_0	$f(b)$

定理 2.4

设 $f(x) \in F[x]$ 且 $\deg f(x) = n > 0$, 则 $f(x)$ 在 F 内至多有 n 个根 (重根按重数计算)

定义 2.16 (Lagrange 插值公式)

设 a_1, a_2, \cdots, a_m 是数域 F 上 m 个不同的数, 则对任意 m 个数 b_1, b_2, \cdots, b_m , 存在唯一次数小于 m 的多项式

$$L(x) = \sum_{i=1}^m b_i \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

满足对于任意的 $i (1 \leq i \leq m)$, 都有 $L(a_i) = b_i$

这个看上去有点抽象, 我们举个例子吧

如 $f(1) = \sqrt{3}, f(2) = 114514, f(3) = 1919810$ 这个肯定是不能硬来的, 所以我们要用拉格朗日插值公式

$$L(x) = 1919810 \frac{(x-1)(x-2)}{(3-1)(3-2)} + 114514 \frac{(x-1)(x-3)}{(2-1)(2-3)} + \sqrt{3} \frac{(x-2)(x-3)}{(1-2)(1-3)}$$

这样, 这个函数, 完美的满足题目所给的条件, 很不错。

定义 2.17 (重根)

设 $b \in F$ 。若 $(x-b)^k | f(x)$, 但 $(x-b)^{k+1} \nmid f(x)$, 则称 b 是 $f(x)$ 的一个 k 重根。若 $k=1$, 则称 b 为单根

注

$f(x)$ 在 F 上有重根, 则在 F 上必有重因式。反之未必, 不能保证重因式有根。

此外, 还有几个结论, 蛮有意思的

结论 $f(x) \in R[x]$, 若 $a + bi, a, b \in R, b \neq 0$ 是 $f(x)$ 的根, $a - bi$ 也是 $f(x)$ 的根

这个结论说明了, 如果有复根, 一定是成对存在的。

结论 $p(x) \in F$ 上不可约, 则必在复数域上无重根

结论 $f(x), p(x)$ 是 F 上多项式, $p(x)$ 在 F 上不可约, 且 $f(x)p(x)$ 在 C 上有公共根, 则 $p(x) | f(x)$

到此, 前置知识就结束了, 我们开启主线任务。

2.3 复系数与实系数多项式

把实数域扩充为复数域比较容易实现, 因此找出实数域上的不可约多项式, 可以利用复数域多项式的信息, 故我们先从复系数多项式开始研究。

定理 2.5 (代数学基本定理)

每个次数大于零的复系数多项式在复数域上至少有一个根。

由定理 (2.5) 我们可以得到：每一个次数大于零的复系数多项式在复数域上至少有一个根，从而次数大于 1 的复系数多项式都是可约的，于是得到

推论 2.1

复数域上的不可约多项式都是一次的。

**定理 2.6 (复系数多项式唯一因式分解定理)**

每一个次数大于 0 的复系数多项式在复数域上都可以唯一地分解为一次因式的乘积。



由此，我们可以得到 $n(n > 0)$ 次复系数多项式 $f(x)$ 在复数域上的标准分解式：

$$f(x) = c(x - a_1)^{r_1}(x - a_2)^{r_2} \cdots (x - a_s)^{r_s} \quad (2.1)$$

其中 a_1, a_2, \dots, a_s 两两不同， c 为 $f(x)$ 首项系数 $r_1, r_2, \dots, r_s \in N^+$ 且 $r_1 + r_2 + \cdots + r_s = n$

于是我们能立即推出：

推论 2.2

$n(n \geq 0)$ 次复系数多项式在复数域上恰有 n 个根（重根按重数计）。



至此我们完全解决了复数域上的不可约多项式，从而细化了复系数多项式的唯一因式分解定理。

利用复系数多项式唯一因式分解定理，我们可以得到数域 F 上的 $n(n > 0)$ 次复系数多项式 $f(x)$ 的复根与它的系数之间的关系：

定理 2.7

设 $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in F[x]$ 在 F 中有 n 个根 c_1, c_2, \dots, c_n ，则有

$$\begin{aligned} \sum_{1 \leq i \leq n} c_i &= -a_{n-1} \\ \sum_{1 \leq i < j \leq n} c_i c_j &= a_{n-2} \\ &\vdots \\ c_1 c_2 \cdots c_n &= (-1)^n a_0 \end{aligned} \quad (2.2)$$



接下来我们来讨论实系数多项式

定理 2.8

设 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in R[x]$ 。若 $a + bi (a, b \in R, b \neq 0)$ 是 $f(x)$ 的复根，则 $a - bi$ 也是 $f(x)$ 的复根



这个说明了虚根是成对出现的。

由此我们立即得到

推论 2.3

实系数的奇次多项式至少有一个实根。

**定理 2.9**

实数域上的不可约多项式或为一次多项式或为二次多项式 $ax^2 + bx + c$ ，其中 $b^2 - 4ac < 0$



由标准分解式和定理 (2.9)，我们可以立即推出：

2.4 有理系数和整系数多项式

有理数域上的不可约多项式有哪些？如何判别一个有理数多项式是否不可约？这就是我们要讨论的问题

设 $f(x) \in Q[x]$, 由于 $f(x)$ 和它的相伴元只差一个非零有理数因子, 因此 $f(x)$ 和它的相伴元在有理数域上有相同的因式, 从而 $f(x)$ 在 Q 上不可约当且仅当它的相伴元在 Q 上不可约。这样我们就可以从 $f(x)$ 的相伴元中选择一个最简单的多项式作为代表研究它的不可约性。这个代表可以很自然地如下选取: 例如, $f(x) = \frac{1}{2}x^3 + \frac{1}{3}x^2 - 2x + 1 = \frac{1}{6}(3x^3 + 2x^2 - 12x + 6)$, 显然 $3x^3 + 2x^2 - 12x + 6$ 是与 $f(x)$ 相伴的最简单的多项式。一般的, 设 $f(x)$ 的各项系数的分母的最小公倍数为 m , 则 $f(x) = m^{-1}mf(x)$, 其中 $mf(x)$ 的各项系数都为整数。设 $mf(x)$ 的各项系数的最大公因数为 d , 则 $mf(x) = d\frac{m}{d}f(x)$, 其中 $\frac{m}{d}f(x)$ 的各项系数的最大公因数为 ± 1 , 于是 $\frac{m}{d}f(x)$ 就是与 $f(x)$ 相伴的最简单的多项式, 由此抽象出本原多项式的概念:

定义 2.18 (本原多项式)

一个非零的整系数多项式 $f(x)$, 如果它的各项系数的最大公因式只有 ± 1 , 那么称 $f(x)$ 是一个本原多项式。

从前文知道, 任何一个非零的有理系数多项式 $f(x)$ 都与一个本原多项式相伴。我们进一步可以得到: 与 $f(x)$ 相伴的本原多项式在相差一个正负号的情况下是唯一的, 即

定理 2.11

两个本原多项式 $f(x), g(x)$ 在 $Q[x]$ 中相伴当且仅当 $f(x) = \pm g(x)$

任何一个次数大于 0 的有理系数多项式都与一个本原多项式相伴, 因此我们只需要去研究本原多项式是否不可约。由于因式分解涉及乘法, 因此自然要问: 两个本原多项式的乘积是否还是本原多项式? 下面的定理就回答了这一问题:

定理 2.12

两个本原多项式的乘积还是本原多项式

要寻找本原多项式不可约的充分条件, 不是很好找。我们可以反过来思考: 从一个本原多项式可约能推出怎样的结论? 从不可约多项式的等价条件得出, 如果一个次数大于 0 的本原多项式可约, 那么它可以分解为两个次数较低的有理系数多项式的乘积。从高斯引理我们可以判断它可以分解为两个次数较低的本原多项式的乘积, 于是我们就有:

定理 2.13

一个次数大于 0 的本原多项式 $f(x)$ 在 Q 上可约当且仅当 $f(x)$ 可以分解为两个较低的本原多项式的乘积。

由这个定理, 我们可以推出

推论 2.4

一个次数大于 0 的整系数多项式 $f(x)$ 在 Q 上可约当且仅当 $f(x)$ 能分解为两个次数较低的整系数多项式的乘积

接下来, 有一个定理, 很有意思:

定理 2.14

整系数多项式 $f(x)$ 在 Q 上可约 $\Leftrightarrow f(x)$ 在 Z 上可约

有这个定理, 我们可以把有理系数多项式在 Q 上的可约问题可以转化为整系数多项式在 Z 上的可约问题。下述定理给出了本原多项式组成的集合的结构

定理 2.15

每一个次数大于 0 的本原多项式 $f(x)$ 都可以唯一地分解为 Q 上地不可约多项式的乘积。唯一性是指, 假如 $f(x)$ 有两个这样的分解式:

$$f(x) = p_1(x)p_2(x)\cdots p_s(x), f(x) = q_1(x)q_2(x)\cdots q_t(x)$$

则 $s = t$ 且经过适当排列因式的次序后, 有

$$p_i(x) = \pm q_i(x), i = 1, 2, \cdots, s$$



一个次数大于 1 的整系数多项式 $f(x)$ 如果有一次因式, 那么 $f(x)$ 可约, 因此次数大于 1 的整系数多项式 $f(x)$ 在 $Q[x]$ 不可约的必要条件是 $f(x)$ 没有一次因式, 而 $f(x)$ 有一次因式当且仅当 $f(x)$ 在 Q 中有根。接下来我们来研究一下整系数多项式在 Q 中有根的必要条件。

定理 2.16

设 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in Z[x]$, 则有理数 $\frac{q}{p}$ 是 $f(x)$ 的根的必要条件是 p 是 a_n 的因数, q 是 a_0 的因数, 其中 p, q 是互素的整数。



当 ± 1 不是 $f(x)$ 的根时, 我们可以推出一个很有意思的结论

结论 设 $c \in Z, f(x) \in Z[x]$ 且 $f(c) = 0$, 则 $\frac{f(-1)}{c+1} \in Z (c \neq -1), \frac{f(1)}{c-1} \in Z (c \neq 1)$

故如果计算出 $\frac{f(-1)}{c+1} \notin Z (c \neq -1), \frac{f(1)}{c-1} \notin Z (c \neq 1)$, 那么 c 就不是 $f(x)$ 的根, 这个方法在判断多项式的有理根的时候很有用。

用定理 (2.16) 可以判断二次或三次整系数多项式是否在 Q 上不可约: 二次或三次整系数多项式在 Q 上不可约当且仅当它没有有理根。

注

对于四次以及四次以上的整系数多项式 $f(x)$, 如果它没有有理根, 那么只能说明 $f(x)$ 没有一次因式, 并不能说明 $f(x)$ 在 Q 上不可约, 因为 $f(x)$ 可能有二次因式或者次数大于二的因式。这表明, 对于四次以及四次以上的整系数多项式 $f(x)$, 没有有理根只是 $f(x)$ 在 Q 上不可约的必要条件, 而不是充分条件。

那么有什么办法来判断 $f(x)$ 在 Q 上不可约呢? 下面的定理给了我们答案

定理 2.17 (Eisenstein 判别法)

设多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in Z[x], a_n \neq 0, n \geq 1, p$ 是一个素数, 若 $p|a_i (i = 0, 1, 2, \cdots, n-1), p \nmid a_n, p^2 \nmid a_0$, 则 $f(x)$ 在 Q 上不可约



由这个可以推出

定理 2.18

在 Q 上存在 $n(n \geq 1)$ 次的不可约多项式



2.5 重点题型

2.5.1 最大公因式总结

结论

若 $d(x)$ 是 $f(x), g(x)$ 的公因式, 且能写成

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

那么 $d(x)$ 为最大公因式。

例题 2.2 设 $t(x)$ 首一, $(f(x), g(x)) = d(x)$, 则 $(f(x)t(x), g(x)t(x)) = d(x)t(x)$

解

要证明 $d(x)t(x)$ 是首一最大公因式, 首先得证明是公因式, 显然有

$$d(x)t(x) | f(x)t(x), d(x)t(x) | g(x)t(x)$$

又因为, $(f(x), g(x)) = d(x)$ 于是有

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

同乘 $t(x)$, 有

$$u(x)f(x)t(x) + v(x)g(x)t(x) = d(x)t(x)$$

得证

例题 2.3 设 $f(x), g(x)$ 不全为 0, 证明 $\left(\frac{f(x)}{(f(x), g(x))}, \frac{g(x)}{(f(x), g(x))}\right) = 1$

解

要证明 $\left(\frac{f(x)}{(f(x), g(x))}, \frac{g(x)}{(f(x), g(x))}\right) = 1$, 即证明

$$u(x)\frac{f(x)}{(f(x), g(x))} + v(x)\frac{g(x)}{(f(x), g(x))} = 1$$

又 $f(x), g(x)$ 不全为 0, 则 $(f(x), g(x)) \neq 0$, 即证明

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

显然成立。

例题 2.4 若 $(f_1(x), g(x)) = 1, (f_2(x), g(x)) = 1$, 则 $(f_1(x)f_2(x), g(x)) = 1$

解

由 $(f_1(x), g(x)) = 1, (f_2(x), g(x)) = 1$, 则有

$$u_1(x)f_1(x) + v_1(x)g(x) = 1, u_2(x)f_2(x) + v_2(x)g(x) = 1$$

两式相乘, 则有

$$u_1(x)f_1(x)u_2(x)f_2(x) + u_1(x)f_1(x)v_2(x)g(x) + v_1(x)g(x)u_2(x)f_2(x) + v_1(x)g(x)v_2(x)g(x) = 1$$

即

$$u_1(x)f_1(x)u_2(x)f_2(x) + g(x)(u_1(x)f_1(x)v_2(x) + v_1(x)u_2(x)f_2(x) + v_1(x)g(x)v_2(x))) = 1$$

则有 $(f_1(x)f_2(x), g(x)) = 1$

例题 2.5 如果 $(f(x), g(x)) = 1$, 那么 $(f(x)g(x), f(x) + g(x)) = 1$

解 由 $(f(x), g(x)) = 1$, 则有

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

对照形式, 则有

$$u(x)[f(x) + g(x)] + [v(x) - u(x)]g(x) = 1$$

$$v(x)[f(x) + g(x)] + [v(x) - u(x)]f(x) = 1$$

于是有

$$(f(x), f(x) + g(x)) = 1, (g(x), f(x) + g(x)) = 1$$

由上一题, 结论得证。

2.5.2 中国剩余定理

为了增强文化自信,弘扬中华民族传统美德,这个定理我们就需要掌握,具体内容是啥呢?如下

定理 2.19 (中国剩余定理)

设 $p_1(x), \dots, p_m(x) \in F[x]$ 且两两互素, $g_1(x), \dots, g_m(x) \in F[x]$ 且 $\deg g_i(x) < \deg p_i(x) (i = 1, \dots, m)$, 则存在唯一的 $g(x), q_i(x) \in F[x] (i = 1, \dots, m)$ 使得 $\deg g(x) < \sum_{i=1}^m \deg p_i(x)$ 且 $g(x) = p_i(x)q_i(x) + g_i(x) (i = 1, \dots, m)$ 。



虽然说,看上去有点复杂,实际上,就是有点复杂,事实上,我们只要会做题就行了。按我给你的套路走,不了解原理也能做题,拿分,以练习为例题。

例题 2.6 设 $f(x)$ 除以 $x^2 + 1, x^2 + 2$ 的余式分别为 $4x + 4, 4x + 8$, 求 $f(x)$ 除以 $(x^2 + 1)(x^2 + 2)$ 的余式

解:

三步走, 第一步, 找 $q(x), p(x)$

这里 $q_1(x) = x^2 + 1, q_2(x) = x^2 + 2, p_1(x) = 4x + 4, p_2(x) = 4x + 8$

第二步, 求 $f_i(x)$, 就把除了 $q_i(x)$ 以外的所有 $q(x)$ 相乘, 凑成 1,

这里就两项, 有 $(-1)(x^2 + 1) + 1(x^2 + 2) = 1$ 于是 $f_1(x) = x^2 + 2, f_2(x) = (-1)(x^2 + 1)$

第三步, $p(x)f(x)$ 累加再加上 $t(x)$ 乘所有 $q(x)$

这里有 $f(x) = t(x)(x^2 + 1)(x^2 + 2) + (4x + 4)(x^2 + 2) + (4x + 8)(-1)(x^2 + 1) = t(x)(x^2 + 1)(x^2 + 2) - 4x^2 + 4x$

例题 2.7 (韩信点兵) 有一次战斗后, 韩信要清点士兵的人数。让士兵三人一组, 就有两人没法编组; 五人一组, 就有三人无法编组; 七人一组, 就有两人无法编组。那么请问这些士兵至少有几?

解

三步走, 第一步, 找 $q(x), p(x)$

这里 $q_1(x) = 3, q_2(x) = 5, q_3(x) = 7, p_1(x) = 2, p_2(x) = 3, p_3 = 2$

第二步, 求 $f_i(x)$, 就把除了 $q_i(x)$ 以外的所有 $q(x)$ 相乘, 凑成 1

这里有三项, 我们挨个求

$q_1(x) = 3, q_2(x)q_3(x) = 35, 12 \cdot 3 + (-1) \cdot 35 = 1$ 于是 $f_1(x) = -35$

$q_2(x) = 5, q_1(x)q_3(x) = 21, (-4) \cdot 5 + 21 = 1$ 于是 $f_2(x) = 21$

$q_3(x) = 7, q_1(x)q_2(x) = 15, (-2) \cdot 7 + 15 = 1$ 于是 $f_3(x) = 15$

第三步, $p(x)f(x)$ 累加再加上 $t(x)$ 乘所有 $q(x)$

这里有 $f(x) = 105t(x) + (-35) \cdot 2 + 21 \cdot 3 + 15 \cdot 2 = 105t(x) + 23$

至少 23 人

2.5.3 重因式

结论

1. 若不可约多项式 $p(x)$ 是 $f(x)$ 的 $k(k \geq 1)$ 重因式, 则 $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式

2. 已知不可约多项式 $p(x)$ 是 $f(x)$ 的因式。若 $p(x)$ 是 $f'(x)$ 的 $k(k > 0)$ 重因式, 则 $p(x)$ 是 $f(x)$ 的 $k+1$ 重因式

3. 若不可约多项式 $p(x)$ 是 $f(x)$ 的 $k(k \geq 1)$ 重因式, 则 $p(x)$ 是 $f'(x), f''(x), \dots, f^{(k-1)}(x)$ 的因式, 但不是 $f^{(k)}(x)$ 的因式

例题 2.8 如果 $(x-1)^2 | (Ax^4 + Bx^2 + 1)$, 求 A, B

解

由题, 1 是多项式的二重根, 由此, 有

$$\begin{cases} f(1) = 0 \\ f'(1) = 0 \end{cases}$$

解得

$$\begin{cases} A = 1 \\ B = -2 \end{cases}$$

例题 2.9 如果 a 是 $f'''(x)$ 的一个 k 重根, 证明: a 是

$$g(x) = \frac{(x-a)}{2}[f'(x) + f'(a)] - f(x) + f(a)$$

的一个 $k+3$ 重根

解

要证明是 $k+3$ 重根, 只需要证明是二阶导的 $k+1$ 重根, 有

$$g''(x) = \frac{(x-a)}{2}f'''(x)$$

又有

$$(x-a)^{k+1}|g''(x), (x-a)^{k+2} \nmid g''(x)$$

得证

2.5.4 有理根和有理数域上的可约性

结论 判断是否有有理根

第一步: 写出 a_n, a_0 所有因子 s_i, r_j , 组合成有理根的所有候选 $\frac{r_j}{s_i}$

第二步: 用代入法检验 1 和 -1 是不是根

1.1 和 -1 都是, 用综合除法或者代入验证

2.1 和 -1 至少有一个不是, 用 $\frac{f(1)}{\frac{r_j}{s_i}-1}$ 和 $\frac{f(-1)}{\frac{r_j}{s_i}+1}$ 排除不是整数的候选, 其它的代入或者用综合除法验证

3. 求出根的, 用综合除法验证重数

例题 2.10 证明: $f(x) = x^5 - 12x^3 + 36x + 12$ 没有有理根。

解

显然有有理根仅有 $\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12$ 这几种情况

经检验, 上述情况都不成立, 于是没有有理根

例题 2.11 求多项式 $f(x) = x^5 + x^4 - 6x^3 - 14x^2 - 11x - 3$ 的有理根

解

显然有有理根仅有 $\pm 1, \pm 3$ 这几种情况

$$f(1) = -32, f(-1) = 0, \frac{-32}{4} = -8, \frac{-32}{2} = 16$$

于是需要综合除法, 经检验 3 为单根, -3 不是根, 1 是四重根

结论 Eisenstein 判别法

设多项式 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in Z[x], a_n \neq 0, n \geq 1, p$ 是一个素数, 若 $p|a_i (i = 0, 1, 2, \cdots, n-1), p \nmid a_n, p^2 \nmid a_0$, 则 $f(x)$ 在 Q 上不可约

例题 2.12 判断 $x^4 - 8x^3 + 12x^2 + 2$ 在有理数域上是否可约

解

$$2|2, 2 \nmid 1, 4 \nmid 2$$

由 Eisenstein 判别法, Q 上不可约

例题 2.13 证明: 当 n 为素数时, $f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$ 在 Q 上不可约。

解

化简, 有

$$n!f(x) = x^n + \cdots + n!$$

$$n!|n!, n! \nmid 1, (n!)^2 \nmid n!$$

由 Eisenstein 判别法, $n!f(x)$ 在 Q 上不可约, 进而 $f(x)$ 在 Q 上不可约

例题 2.14 若 p 为素数, 证明: $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ 在 Q 上不可约。

解

注意到

$$(x-1)f(x) = x^p - 1$$

令 $y = x - 1$, 有

$$f(y+1) = \frac{(y+1)^p - 1}{y}$$

又因为

$$\begin{aligned} (a+b)^n &= C_n^0 a^n + C_n^1 a^{n-1}b + C_n^2 a^{n-2}b^2 + \cdots + C_n^{n-1} a b^{n-1} + C_n^n b^n \\ &= \sum_{k=0}^n C_n^k a^{n-k} b^k \end{aligned}$$

所以原式等于

$$f(y+1) = \frac{\sum_{k=0}^p C_n^k y^{p-k} - 1}{y} = \frac{\sum_{k=0}^{p-1} C_n^k y^{p-k}}{y} = \sum_{k=0}^{p-1} C_p^k y^{p-k-1}$$

又因为

$$C_p^k = \frac{p!}{i!(p-i)!} \in Z$$

故有

$$p | C_p^k$$

又因为 $p \nmid 1, p^2 \nmid C_p^k$, 由 Eisenstein 判别法, $f(y)$ 在 Q 上不可约, 进而 $f(x)$ 在 Q 上不可约

例题 2.15 设 $f(x) = (x-a_1)(x-a_2)\cdots(x-a_n) - 1$, 其中 a_1, a_2, \cdots, a_n 为两两不同的整数。求证: $f(x)$ 在 Q 上不可约

解

设 $f(x)$ 可约, 则有

$$f(x) = g(x)h(x), 0 < \deg h(x), \deg g(x) < n$$

于是有

$$\begin{cases} f(a_1) = g(a_1)h(a_1) = -1 \\ f(a_2) = g(a_2)h(a_2) = -1 \\ \cdots \\ f(a_n) = g(a_n)h(a_n) = -1 \end{cases}$$

又因为 $g(x), h(x)$ 为整系数多项式, 于是 $g(x) = \pm 1, h(x) = \mp 1$ 即

$$h(x) + g(x) = 0$$

对 a_1, \cdots, a_n 恒成立。又因为该等式次数小于 n , 又有 n 个根, 故该等式为恒等式。于是有

$$f(x) = -g^2(x)$$

这与 $f(x)$ 首一矛盾, 故 $f(x)$ 在 Q 上不可约。

第3章 特征值

3.1 导学

这一章节是围绕着特征值和特征多项式进行展开,进而衍生出相应的可对角化概念。然后再提出了零化多项式和极小多项式的概念。**Hamilton-Cayley Theorem** 提供出了一个较为简便的找出一个矩阵的零化多项式的方法,然后我们根据极小多项式可以整除零化多项式,可以缩小寻找极小多项式的范围。事实上这样的方法还是不够简便,如何寻找极小多项式会在下一章提到。

可以从目录看出来,本章的重点还是在特征值和特征多项式上,事实也是如此,可对角化的考点较少,思维较为固定,更别说是稍微一提的极小多项式了。本章的知识点主要是从矩阵的角度出发,鲜有线性空间。主要是对高代老师课上内容的补充和整理,再加上一些个人的思考。

3.2 特征值和特征多项式

3.2.1 前置知识

Define: 设 $A \in F^{n \times n}$ 。若存在 $\lambda \in F$, 使得 $A\alpha = \lambda\alpha$, 则称 λ 是 A 的一个特征值, α 是 A 的属于特征值 λ 的一个特征量向量 (当然, 这里 α 不为零向量)。

Attention: 设 $A \in F^{n \times n}$ 。 α 是 A 的属于特征值 λ 的一个特征量向量, $i \in N^+$, 则 α 是 A^i 属于特征值 λ^i 的一个特征量向量。

Attention: 一个特征向量只能属于一个特征值。

Attention: A 的属于特征值 λ 的特征量向量的非零线性组合 仍是 A 的属于特征值 λ 的特征量向量。

Define: 设 $A \in F^{n \times n}$, λ 是 A 的一个特征值, 称 V_λ 是的属于特征值 λ 的特征子空间 (特征向量本身就包含对加法和纯量乘法封闭的性质, 此时只要包含零向量即可构成线性空间。)

Theorem: 设 $A \in F^{n \times n}$, 则 A 的属于不同特征值的特征向量线性无关。

Define: 设 $A \in F^{n \times n}$, $f(A) = \sum_{i=0}^n a_i A^i$, α 是 A 的属于特征值 λ 的一个特征量向量, 则 α 是 $f(A)$ 属于特征值 $f(\lambda)$ 的特征向量。

Define: 设 $A \in F^{n \times n}$, 矩阵 $\lambda E - A$, 称为 A 的特征矩阵, $\det(\lambda E - A)$ 称为 A 的特征多项式, 记作 $f_A(\lambda)$ 。 $f_A(\lambda)=0$, 称为 A 的特征方程, $f_A(\lambda)$ 的根称为 A 的特征根。

Attention: 设 $A \in F^{n \times n}$, α 是 A 的属于特征值 λ_0 的一个特征向量, 则 λ_0 是特征多项式 $f_A(\lambda)$ 的根, α 是齐次线性方程组 $(\lambda_0 E - A)x=0$ 的非零解。

Attention: 设 $A \in F^{n \times n}$, 则 A 的所有特征值之和为 A 的迹, A 的所有特征值之积为 $\det(A)$ 。

Attention: 相似矩阵有相同的特征多项式, 从而有相同的特征值。转置矩阵之间也有相同的特征多项式, 同理也有相同的特征值。

线性空间有关部分，本资料不做更多归纳

3.2.2 特征向量的结构

Conclusion: A 是 n 阶矩阵, $\lambda_1, \lambda_2, \lambda_3$ 分别是 A 的 k_1, k_2, k_3 重特征值, 且 $k_1 + k_2 + k_3 = n$ 。则有 $\xi_1, \xi_2, \dots, \xi_s$ 是 $(\lambda_1 E - A)x=0$ 的一个基础解系 $\eta_1, \eta_2, \dots, \eta_t$ 是 $(\lambda_2 E - A)x=0$ 的一个基础解系 $\delta_1, \delta_2, \dots, \delta_r$ 是 $(\lambda_3 E - A)x=0$ 的一个基础解系此时 $\xi_1, \xi_2, \dots, \xi_s, \eta_1, \eta_2, \dots, \eta_t, \delta_1, \delta_2, \dots, \delta_r$ 线性无关。即 λ_1 所对应的特征向量不可由 λ_2, λ_3 所对应的特征向量表出。

Theorem: 若 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的全部特征值, 则 $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$ 是 $f(A)$ 的全部特征值。

Infer: 现在讨论特殊情况

$f(x)=bx(b \neq 0)$ 则此时 $\det(\lambda E - bA) = (\lambda - b\lambda_1) \cdots (\lambda - b\lambda_n)$

$f(x)=x+b$ 则此时 $\det(\lambda E - (A + bE)) = (\lambda - (\lambda_1 + b)) \cdots (\lambda - (\lambda_n + b))$

$f(x)=x^{-1}$ 则此时 $\det(\lambda E - A^{-1}) = (\lambda - \lambda_1^{-1}) \cdots (\lambda - \lambda_n^{-1})$

$f(x)=x^m(m \text{ 为正整数})$ 则此时 $\det(\lambda E - A^m) = (\lambda - \lambda_1^m) \cdots (\lambda - \lambda_n^m)$

Application:

例 1. A 是三阶矩阵, 若 $A^2 - 6A + 11E = 6A^{-1}$, 则 A 的特征值可能是多少?

解: 原式等价于 $\lambda^2 - 6\lambda + 11 = 6\lambda^{-1}$, 其中 λ 为任意一个特征值。令其等于 0, 解得 $\lambda = 1, 2, 3$ 。故 A 的特征值只能为这三个中的某一个。

例 2. A 是 n 阶矩阵, 证明: $A^2 + E$ 可逆

证明: 不妨设 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则 $A^2 + E$ 的特征值为 $\lambda_1^2 + 1, \lambda_2^2 + 1, \dots, \lambda_n^2 + 1$ 则 $\det(A^2 + E) = (\lambda_1^2 + 1)(\lambda_2^2 + 1) \cdots (\lambda_n^2 + 1) \neq 0$ 故可逆。

3.2.3 相似矩阵

Attention: 相似矩阵有着相同的特征多项式, 故有着相同的特征值和重数。

Attention: 设 $A, B \in F^{n \times n}$ 且 $A \sim B$, 则存在 n 阶可逆阵 $P \in F^{n \times n}$ 使得 $B = P^{-1}AP$ 。若 α 是 A 的属于特征值 λ_0 的一个特征向量, 则 $P^{-1}\alpha$ 是 B 的属于特征值 λ_0 的一个特征向量

Application: 设 $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$, $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $B = P^{-1}A^*P$, 求 $B + 2E$ 的特征值和特征向量。

解: 由 $A^* = \det(A)A^{-1}$ 可知 $B = P^{-1}\det(A)A^{-1}P$ 。即求 A^{-1} 的特征值和特征向量, 事实上, A, A^{-1} 有着相同的特征向量和特征值的重数。即求 A 的特征值和特征向量。

$$\lambda E - A = \begin{bmatrix} \lambda-3 & -2 & -2 \\ -2 & \lambda-3 & -2 \\ -2 & -2 & \lambda-3 \end{bmatrix} = \lambda^3 - 9\lambda^2 + 15\lambda - 7 = 0$$

解得 $\lambda = 1, 1, 7$

$$\lambda=1 \text{ 时, } E-A=\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \alpha_1=\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \alpha_2=\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},$$

$$\lambda=7 \text{ 时, } 7E-A=\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \alpha_3=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

要求 $B+2E$ 的特征值和特征向量, 即求 $A^{-1} \det(A) + 2E$ 的特征值和特征向量。 $\lambda=1, 1, 7$ 则 A^{-1} 的特征值为 $1, 1, \frac{1}{7}$ 特征向量都为 $\alpha_1, \alpha_2, \alpha_3$ $B+2E$ 的特征值为 $9, 9, 3$ 。此时对应的特征向量为 $P^{-1}\alpha_1, P^{-1}\alpha_2, P^{-1}\alpha_3$

3.2.4 秩一矩阵

Conclusion: A 是 n 阶矩阵, $r(A)=1$, 则 A 的所有特征值为 $\underbrace{0, \dots, 0}_{n-1}, \operatorname{tr}(A)$ 。当 $\operatorname{tr}(A)$ 大于 0 时, A 可对角化, 当 $\operatorname{tr}(A)$ 等于 0 时, A 不可对角化,

证明: $Ax=0=0x$. 故 0 为 A 的特征值, $r(A)=1$, 所以 0 几何重数为 $n-1$, 又代数重数大于几何重数, 所以 0 的重数大于等于 $n-1$ 又有 A 的特征值的和为 $\operatorname{tr}(A)$, 并且前 $n-1$ 个特征值全为 0, 故最后一个特征值为 $\operatorname{tr}(A)$ 。

Conclusion: $\operatorname{tr}(A)=1$ 时, A 可表示为 $\alpha\beta^T$, 其中 α, β 为非零列向量, 此时 $\operatorname{tr}(A)=\alpha^T\beta=\beta^T\alpha$

Infer:

当 $\operatorname{tr}(A)=0$ 时, A 的所有特征向量为 $\beta^T x=0$ 的所有非零解

当 $\operatorname{tr}(A) \neq 0$ 时, A 的所有特征向量为 $\beta^T x=0$ 的所有非零解 $+k\alpha (k \neq 0)$

0 对应的特征向量为 $\beta^T x=0$ 的所有非零解, $\operatorname{tr}(A)$ 对应的所有特征向量为 $k\alpha (k \neq 0)$

证明:

$Ax=\alpha\beta^T x=0x$, 又 α 非零, 则 $\beta^T x=0=0x$, 故 0 对应的特征向量为 $\beta^T x=0$ 的所有非零解

$A\alpha=\alpha\beta^T \alpha=\operatorname{tr}(A)\alpha$ 。故 $\operatorname{tr}(A)$ 对应的所有特征向量为 $k\alpha (k \neq 0)$

Application:

例 1. n 阶方阵 A 所有元素为 1, 求 A 的所有特征值和特征向量。

解: 由题, $\operatorname{tr}(A)=1$, 则阵 A 的所有特征值为 $\underbrace{0, \dots, 0}_{n-1}, n$ 。且 $A=\alpha\beta^T=\begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}$, 0 对应的特征向量为 $\beta^T x=0$ 的所有非零解, $\operatorname{tr}(A)$ 对应的所有特征向量为 $k\alpha (k \neq 0)$ 解 $\beta^T x=0$ 即 $\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} x=0$; 解得 0 的特征

向量为 $\begin{bmatrix} 1 \\ -1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ \dots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ -1 \end{bmatrix}, n$ 对应的特征向量为 $\begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix}$ 。

例 2. $A=E+\alpha\beta^T$ 其中 α, β 为列向量且 $\beta^T \alpha=1$, 求 $\det(A+E)$

解: $\alpha\beta^T$ 的特征值为 $\underbrace{0, \dots, 0}_{n-1}, 1$ 则 $A+E$ 的特征值为 $\underbrace{2, \dots, 2}_{n-1}, 3$, 则 $\det(A+E)=3 \times 2^{n-1}$

例 3. β 是 n 维列向量, 证明: 若 $\beta^T \beta = 1$, 则 $E - \beta \beta^T$ 不可逆, 若 $\beta^T \beta \neq 1$, 则 $E - \beta \beta^T$ 可逆.

证明: $E - \beta \beta^T$ 的所有特征值为 $\underbrace{1, \dots, 1}_{n-1}, 1 - \beta^T \beta$. 若 $\beta^T \beta = 1$, 则 $\det(E - \beta \beta^T) = 0$, 不可逆. 同理.

3.2.5 伴随矩阵

Conclusion: 若 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶方阵 A 的全部特征值

Case1. $r(A) = n, A^* = \det(A) A^{-1}$

A^* 的特征值为 $\frac{|A|}{\lambda_1}, \dots, \frac{|A|}{\lambda_n}$.

$$\text{tr}(A^*) = \sum_{j=1}^n \prod_{\substack{i=1 \\ i \neq j}}^n \lambda_i$$

Case2. $r(A) = n-1, r(A^*) = 1$

A^* 的特征值为 $\underbrace{0, \dots, 0}_{n-1}, \text{tr}(A^*)$.

$$\text{tr}(A^*) = \sum_{j=1}^n \prod_{\substack{i=1 \\ i \neq j}}^n \lambda_i$$

A^* 的特征多项式 $|\lambda E - A^*| = \lambda^{n-1}(\lambda - \text{tr}(A^*))$

Case3. $r(A) < n-1, r(A^*) = 0$

A^* 的特征值为 $\underbrace{0, \dots, 0}_n$

A^* 的特征多项式 $|\lambda E - A^*| = \lambda^n$

Application:

例 1. $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix}$, 求 A^* 的特征值

解: $|\lambda E - A| = (\lambda - 1)(\lambda - 2)(\lambda - 3)(\lambda - 4)$. 故 A 的特征值为 1, 2, 3, 4, A^* 的特征值为 6, 8, 12, 24.

3.2.6 AB 与 BA

Conclusion: A, B 分别是 $m \times n, n \times m$ 矩阵, 则 $|\lambda E_m - AB| = \lambda^{m-n} |E_n - BA|$ (结论不多, 但很重要)

Infer: AB 与 BA 的 0 特征值重数差 $m-n$ 其他特征值重数一样并且 $|AB| = |BA|$ 要么说明 0 是 AB 与 BA 的特征值, 要么 0 都不是 AB 与 BA 的特征值。

3.3 可对角化

Define: 设 $A \in F^{n \times n}$. 存在 n 阶可逆阵 $P \in F^{n \times n}$ 使得 $P^{-1}AP$ 为对角阵, 则称 A 是可对角化的。(此时的对角线上的元素恰好为 A 的特征值, P 的列向量为 A 的特征向量。)

Theorem: A 可对角化的充要条件为 A 的特征多项式的根全在 F 上, 且每个特征值的代数重数 (特征多项式中对应特征值的重数) 等于几何重数 (特征值对应特征向量个数)。

Application:

例 1. 已知 $A = \begin{bmatrix} 1 & -1 & 1 \\ x & 4 & y \\ -3 & -3 & 5 \end{bmatrix}$ 有三个线性无关的特征向量, $\lambda=2$ 是二重特征值。试求可逆矩阵 P , 使得 $P^{-1}AP$ 为对角阵。

解: 由 $\lambda=2$ 是二重特征值, 可知 $r(2E - A)=3-2=1$, 对 $2E - A$ 作初等变换, 可得 $\begin{bmatrix} 1 & -1 & 1 \\ 2-x & 0 & -2-y \\ 0 & 0 & 0 \end{bmatrix}$ 即

$$x=2, y=-2. \text{ 此时 } A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 4 & -2 \\ -3 & -3 & 5 \end{bmatrix}$$

由 A 的所有特征值之和为 10, 则可得到 A 的另一个特征值为 6, 则 A 的特征值为 2, 2, 6。

解 $(2E - A)x=0$, 得到以下基础解系 $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ 。

解 $(6E - A)x=0$, 得到以下基础解系 $\begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix}$,

$$\text{令 } P = \begin{bmatrix} -1 & 1 & \frac{1}{3} \\ 1 & 0 & -\frac{1}{3} \\ 0 & 1 & 1 \end{bmatrix}, \text{ 则 } P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

3.4 极小多项式

Define: 设 $A \in F^{n \times n}$ 。若 $f(A)=0$, 则称 $f(\lambda)$ 是 A 的零化多项式。

Hamilton - Cayley Theorem: 一个方阵的特征多项式为这个方阵的零化多项式。

Define: 设 $A \in F^{n \times n}$ 。 A 的 **次数最低且首一的零化多项式** 称作 A 的极小多项式。记作 $m_A(\lambda)$ 。

Theorem:

1. 方阵 A 的极小多项式整除 A 的所有零化多项式。
2. 方阵的极小多项式唯一。
3. 相似矩阵具有相同的极小多项式。

Attention: 在不计重数的情况下, $m_A(\lambda)$ 与 $f_A(\lambda)$ 有完全相同的根。

Theorem: 设 $A, B \in F^{n \times n}$, 则 $(m_A(\lambda), m_B(\lambda))=1 \Leftrightarrow (f_A(\lambda), f_B(\lambda))=1$

Infer: 设 $A, B \in F^{n \times n}$, 若 $(f_A(\lambda), f_B(\lambda))=1$, 则 $f_A(B), f_B(A)$ 可逆。

Application :

$$\text{例 1. 设 } A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 2 & 0 \end{bmatrix}, \text{ 求 } A^{500}.$$

$$\text{解: 由 Hamilton-Cayley Theorem, 可得 } A^3 - A^2 = 0 \text{ 则 } A^{500} = A^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}$$

第4章 相似标准形

4.1 导学

在这一章节，我们主要解决的是之前遗留下来的问题：矩阵相似的充要条件是什么？有没有找极小多项式的简便方法？另外从可对角化我们得到了启发：能不能找到一个与原矩阵相似的形式较为简便的矩阵，来供我们来研究原矩阵的部分性质，因此衍生出来了相似标准形。其中就有 Frobenius 标准形和 Jordan 标准形。前者由矩阵不变因子所决定，后者由矩阵初等因子决定，各有各的好处。本章内容安排与书本上不太一致，主要理解为主，并没有安排习题，想要做题请参考书本。

4.2 前置知识

Define： 类比与数量矩阵，我们将数量矩阵里的元素替换为关于 λ 的多项式，即可得到 λ -矩阵，其拥有数量矩阵的初等变换、相抵关系等关系。（当然，一部分满秩的 λ -矩阵还具有可逆的性质，这里要特别注意的是，特征矩阵满秩，但其一定不可逆）

数量矩阵具有相抵标准形，即通过一系列初等变换可变为单位矩阵，类似的 λ -矩阵也有类似的性质，只不过变为的矩阵为法式

Define： 设 $A(\lambda)$ 是一个 $m \times n$ 的 λ -矩阵且 $r(A(\lambda))=r$, 则

$$A(\lambda) \simeq \begin{pmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

其中 $d_i(\lambda) (i = 1, 2, \dots, r)$ 为首一多项式，且 $d_i(\lambda) | d_{i+1}(\lambda) (i = 1, 2, \dots, r-1)$ 。

我们知道，特征矩阵肯定是满秩的，这样我们就得到了特征矩阵的法式。

Infer： 设 A 是数域 F 上的 n 阶方阵，则 A 的特征矩阵 $\lambda E - A$ 的法式为

$$\text{diag}(d_1(\lambda), \dots, d_n(\lambda))$$

其中 $d_i(\lambda) (i = 1, 2, \dots, n)$ 为首一多项式，且 $d_i(\lambda) | d_{i+1}(\lambda) (i = 1, 2, \dots, n-1)$ 。

我们知道，就算一个矩阵不为方阵，但其仍可取出 k 阶子式 (k 小于等于行数和列数中的最小值) 就有如下定义

Define： 设 $A(\lambda)$ 是一个 $m \times n$ 的 λ -矩阵且 $r(A(\lambda))=r$, 则其所有 k 阶子式 (k 小于等于行数和列数中的最小值) 中的首一最大公因式为 $A(\lambda)$ 的 k 阶行列式因子。记作 $D_k(\lambda)$ ，其个数与 $A(\lambda)$ 的秩个数相同。

接下来我们再回到一般 λ -矩阵的法式

$$A(\lambda) \simeq \begin{pmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

其中 $d_i(\lambda) (i = 1, 2, \dots, r)$ 为首一多项式, 且 $d_i(\lambda) | d_{i+1}(\lambda) (i = 1, 2, \dots, r-1)$ 。根据 $d_i(\lambda) | d_{i+1}(\lambda) (i = 1, 2, \dots, r-1)$, 其行列式因子为

$$D_1 = d_1(\lambda)$$

$$D_2 = d_1(\lambda)d_2(\lambda)$$

...

$$D_r = d_1(\lambda)d_2(\lambda) \cdots d_r(\lambda)$$

我们发现 $D_i(\lambda) | D_{i+1}(\lambda) (i = 1, 2, \dots, r-1)$, 并由此引出一个新的概念

Define: 设 $D_i(\lambda) (i = 1, 2, \dots, r)$ 为 $A(\lambda)$ 的行列式因子, 称

$$d_1(\lambda) = D_1(\lambda), d_i(\lambda) = \frac{D_{i+1}(\lambda)}{D_i(\lambda)} (i = 2, \dots, r-1)$$

为 $A(\lambda)$ 的不变因子, 显然, 其和的 λ -矩阵的行列式因子相互确定。

在此时, 我们发现不变因子已经是形式比较简单的了, 我们是否还能找到比不变因子的形式还简单的一个式子来与不变因子相互确定呢? 答案是肯定的, 因为在求不变因子的时候, 我们并没有要求不变因子在数域 F 上一定可约所以。由因式分解定理, 任何一个多项式在 F 都可分解为不可约因式的方幂之积, 因此我们对 $A(\lambda)$ 的不变因子 $d_i(\lambda) (i = 1, \dots, r-1)$ 进行分解

$$d_1(\lambda) = p_1^{e_{11}}(\lambda)p_2^{e_{12}}(\lambda) \cdots p_t^{e_{1t}}(\lambda)$$

$$d_2(\lambda) = p_1^{e_{21}}(\lambda)p_2^{e_{22}}(\lambda) \cdots p_t^{e_{2t}}(\lambda)$$

.....

$$d_r(\lambda) = p_1^{e_{r1}}(\lambda)p_2^{e_{r2}}(\lambda) \cdots p_t^{e_{rt}}(\lambda)$$

其中 p_j 是首一的两两互素的不可约多项式, e_{ij} 是非负整数, 且 $e_{1j} \leq e_{2j} \leq \cdots \leq e_{rj} (j = 1, 2, \dots, t)$ 由此可见, 我们找到了比不变因子的形式还简单的一个式子来与不变因子相互确定, 然后再引出一个新的概念。

Define: 我们将上面分解式中满足 $e_{ij} > 0$ 的 $p_j^{e_{ij}}(\lambda)$ 叫做 $A(\lambda)$ 的一个初等因子, $A(\lambda)$ 的全体初等因子称为 $A(\lambda)$ 的初等因子组。 $A(\lambda)$ 的不变因子由其初等因子组和秩唯一确定。

接下来, 我们回到之前的一个问题: 矩阵相似的充要条件是什么?

Theorem: 对于两个 $m \times n$ 的 λ -矩阵 $A(\lambda), B(\lambda)$, 下列叙述等价

1. $A(\lambda), B(\lambda)$ 相抵
2. $A(\lambda), B(\lambda)$ 具有相同的法式
3. $A(\lambda), B(\lambda)$ 具有相同的行列式因子
4. $A(\lambda), B(\lambda)$ 具有相同的不变因子
5. $A(\lambda), B(\lambda)$ 具有相同的初等因子组

因此, 我们就解决了目前的一个主要问题, 剩下的问题交给下一节。

4.3 Frobenius 标准形和 Jordan 标准形

本节我们主要是围绕 Frobenius 标准形和 Jordan 标准形进行展开, 来比较二者的特点, 顺便解决一下上一章遗留下来的问题: 如何较为简便地寻找极小多项式?

首先我们来研究 Frobenius 标准形, 为此, 我们需要引进 Frobenius 块的概念。

Define : r 阶矩阵

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{r-1} \end{pmatrix}$$

的行列式因子和不变因子均为 $\underbrace{1, \dots, 1}_{r-1}, f(\lambda)$, 其中

$$f(\lambda) = \lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_1\lambda + a_0$$

我们称之为关于 $f(\lambda)$ 的 **Frobenius 块**, 记作 $F(f(\lambda))$ 。由于其行列式因子和不变因子的特殊性, 我们可以推出这一 **Frobenius 块** 所对应 a 的特征多项式和极小多项式全为 $f(\lambda)$ 。(记住这个结论即可)

根据这个形式, 我们发现, 这一形式的极小多项式恰好为最后一个不变因子。能不能推出更一般的结论, 使得我们能够快速的找到一个矩阵的极小多项式呢?

Theorem: 设数域 F 上 n 阶矩阵 A 的不变因子为

$$\underbrace{1, \dots, 1}_{n-k}, d_1(\lambda), d_1(\lambda), \dots, d_k(\lambda)$$

则

$$m_A(\lambda) = d_k(\lambda)$$

因此的话, 我们又解决了一个问题, 要求一个矩阵的极小多项式, 我们只要求出它最后一个不变因子即可, 这无疑是快了不少。接着我们继续回到 **Frobenius 标准形**。

Define : 设数域 F 上 n 阶矩阵 A 的不变因子为

$$\underbrace{1, \dots, 1}_{n-k}, d_1(\lambda), d_2(\lambda), \dots, d_k(\lambda)$$

其中 $\deg d_i(\lambda) \geq 1, d_i(\lambda) \mid d_{i+1}(\lambda) (i = 1, 2, \dots, k-1)$ 则 A 相似与分块对角矩阵

$$F = \text{diag}(\underbrace{1, \dots, 1}_{n-k}, F(d_1(\lambda)), F(d_2(\lambda)), \dots, F(d_k(\lambda)))$$

我们便称 F 为 **Frobenius 标准形** 由于相似矩阵有着相同的法式, 则 $\lambda E - F$ 的法式为

$$\text{diag}(\underbrace{1, \dots, 1}_{n-k}, d_1(\lambda), d_2(\lambda), \dots, d_k(\lambda))$$

我们发现, 一个方阵的 **Frobenius 标准形** 是由这个方阵的不变因子唯一确定的, 对于每一个 **Frobenius 块** 的排列有着严格的限制, 必须要求所对应的 $d_i(\lambda)$ 满足 $d_i(\lambda) \mid d_{i+1}(\lambda)$ 并且有的时候我们要研究的对象是对角矩阵, 这个时候再从行列式因子来得到不变因子就显得繁琐了, 那有没有什么较为简便的呢? 那就是接下来我们要研究的 **Jordan 标准形**。

首先我们先引入一个广义 **Jordan 块** 的概念。

Define: 设 $p(\lambda)$ 是数域 F 上 m 次首一不可约多项式, $F(p(\lambda))$ 是关于 $p(\lambda)$ 的 **Frobenius 块**。令

$$C = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

并且其为 m 阶方阵, 则 em 阶方阵

$$J = \begin{pmatrix} F(p(\lambda)) & & & \\ C & F(p(\lambda)) & & \\ & \ddots & \ddots & \\ & & C & F(p(\lambda)) \end{pmatrix}$$

的行列式因子和不变因子为 $\underbrace{1, \dots, 1}_{e-1}, p^e(\lambda)$, 初等因子为 $p^e(\lambda)$. 我们称 F 为关于 $p^e(\lambda)$ 的广义 Jordan 块, 记作 $J(p^e(\lambda))$.

Define: 设 A 的初等因子组为 $p_1^{e_1}(\lambda), p_2^{e_2}(\lambda), \dots, p_k^{e_k}(\lambda)$, 则 A 相似于分块对角矩阵

$$J = \text{diag}(J(p_1^{e_1}(\lambda)), J(p_2^{e_2}(\lambda)), \dots, J(p_k^{e_k}(\lambda)))$$

我们称这个对角矩阵为 A 的广义 Jordan 标准形 (因为不保证初等因子组中每一个不可约多项式都为一次多项式)

由于初等因子组并没有硬性的顺序要求, 又因为 A 的初等因子组是唯一确定的, 故可以说 A 的广义 Jordan 标准形和初等因子组相互唯一确定。

现在我们将初等因子组放在复数域上进行讨论, 则其不变因子必定可以分解为多个一次多项式的次幂的乘积, 我们类比原来的广义 Jordan 块, 由于此时所有的不变因子都形如 $(\lambda - \lambda_0)^e$, 其中 λ_0 是 A 的特征值则其对应的广义 Jordan 块为

$$J((\lambda - \lambda_0)^e) = \begin{pmatrix} \lambda_0 & & & \\ 1 & \lambda_0 & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_0 \end{pmatrix}$$

其行列式因子和不变因子为 $\underbrace{1, \dots, 1}_{e-1}, (\lambda - \lambda_0)^e$, 初等因子为 $(\lambda - \lambda_0)^e$.

我们称 $J((\lambda - \lambda_0)^e)$ 为属于 λ_0 的 e 阶 Jordan 块, 简记为 $J(\lambda_0, e)$ 。

Define: 设 A 的初等因子组为 $(\lambda - \lambda_1)^{e_1}, (\lambda - \lambda_2)^{e_2}, \dots, (\lambda - \lambda_m)^{e_m}$ 则 A 相似于分块对角矩阵

$$J = \text{diag}(J(\lambda_1, e_1), \dots, J(\lambda_m, e_m))$$

我们称 J 为 A 的 Jordan 标准形。由于初等因子组并没有硬性的顺序要求, 又因为 A 的初等因子组是唯一确定的, 故可以说 A 的 Jordan 标准形和初等因子组相互唯一确定。关于 Jordan 标准形, 还有一些好玩的推论。

Infer: A 是复数域上的方阵下列命题等价:

1. A 可对角化
2. A 的初等因子全是一次的 (一次的初等因子对应的 Jordan 块是一个数, 则所有的初等因子所对应的 Jordan 标准形为对角阵)
3. $m_A(\lambda)$ 无重根 (极小多项式对应的是最后一个不变因子, 若无重根代表前面的不变因子也无重根)

这个推论能够帮我们快速的判断一些特定矩阵是否可对角化

如 $A^2 - E = 0$ 就可轻松地判断出其可对角化, 因为其对应的特征多项式为 $(\lambda - 1)(\lambda + 1) = 0$, 很容易的判断出这是极小多项式, 又没有重根, 就可对角化。

Infer: A 是复数域上的方阵下列命题等价:

1. A 相似于 cE_n
2. A 的初等因子全是一次且相同 (一次的初等因子对应的 Jordan 块是一个数, 则所有的初等因子所对应的 Jordan 标准形为对角阵, 且对角元素相同)

3. $m_A(\lambda)$ 为一次多项式 (极小多项式具有多项式所有的根, 若极小多项式为一次, 在不考虑重数的情况下, 代表多项式仅有一个根, 则其不变因子次数全为一次, 且又前者整除后者, 则 A 的极小多项式全为一次且相等)

第5章 欧氏空间

5.1 导学

在前面的章节，我们研究的是线性空间与线性空间之间元素的数量关系. 用向量来解释的话就是不同向量之间的线性关系。现在抛出一个问题？一个向量在空间中到底该如何描述？无非就是长度，方向。抽象到线性空间，我们要更深一步地研究元素之间的度量关系

5.2 前置知识

由三维的的向量，我们描述一个向量的度量关系往往借助向量的内积来表示，所以我们来引进线性空间内积的概念

定义 5.1

设 V 是实数域 \mathbb{R} 上的线性空间，我们称映射 $(-, -) : V \times V \mapsto \mathbb{R}$ 为内积。如果对于任意的 $\alpha, \beta, \gamma \in V, c \in \mathbb{R}$ 都有

$$(1) (\alpha, \beta) = (\beta, \alpha) \text{ (对称性)}$$

$$(2) (\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$$

$$(3) (c\alpha, \beta) = c(\alpha, \beta)$$

$$(4) (\alpha, \alpha) \geq 0 \text{ 且等号成立的条件为 } \alpha = 0 \text{ (正定性)}$$

我们就称 V 为关于内积的欧几里得空间



在此定义下，我们就自然就会有以下的内积

定义 5.2 (标准内积)

在实数域上的 n 维列向量空间 \mathbb{R}^n 中，对于 $X = (a_1, a_2, \dots, a_n), Y = (b_1, b_2, \dots, b_n)$ ，定义

$$(X, Y) = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

为标准内积。



定义完内积之后，我们就可以研究线性空间的度量关系，首先是长度

定义 5.3

设 V 是欧氏空间， $\alpha \in V$ 我们定义 α 的长度为 $\sqrt{(\alpha, \alpha)}$ ，记作 $|\alpha|$ ，设 $X = (a_1, a_2, \dots, a_n)$ ，其长度用坐标表示就是

$$|X| = \sqrt{(X, X)} = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

对 $\frac{\alpha}{|\alpha|}$ ，我们称做将 α 化为单位向量。



定义完了长度，接下来就来到了向量之间的角度，为了保证定义的合理性，我们需要引用一个不等式

定理 5.1 (Cauchy-Schwarz 不等式)

设 V 是欧氏空间，对于任意的 $\alpha, \beta \in V$ ，总有

$$(\alpha, \beta)^2 \leq (\alpha, \alpha)(\beta, \beta)$$

当且仅当 α, β 线性相关时，等号成立。



由此我们可以得到

$$-1 \leq \frac{(\alpha, \beta)}{|\alpha||\beta|} \leq 1$$

因此，我们可以定义向量之间的夹角

定义 5.4

设 V 是欧氏空间，对于任意的非零向量 $\alpha, \beta \in V$ ，它们的夹角 θ 由

$$\cos \theta = \frac{(\alpha, \beta)}{|\alpha||\beta|}, (0 \leq \theta \leq \pi)$$

其中，当 $\theta = \frac{\pi}{2}$ 时，两个向量内积为 0，我们就得到了一种特殊情况，定义如下

定义 5.5

设 V 是欧氏空间，对于任意的非零向量 $\alpha, \beta \in V$ ，如果满足

$$(\alpha, \beta) = 0$$

我们就称 α, β 正交，记为 $\alpha \perp \beta$ 。特别要注意的是，零向量和任何向量都正交。

然后我们就有一个结论

结论 欧氏空间 V 中不含零向量的正交向量组必线性无关。

由于在线性空间中，个数和线性空间相等的一组线性无关的向量，可以构成这个线性空间的一组基，我们就有以下定义

定义 5.6

欧氏空间 V 中的正交向量组构成的基称为正交基，标准正交向量组构成的基称为标准正交基。

注意的是，在标准正交基下，向量的内积形式如下

$$\text{设 } \alpha = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n, \beta = b_1\xi_1 + b_2\xi_2 + \cdots + b_n\xi_n$$

$$(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

其中 $(\xi_i, \xi_j) = \delta_{ij}$

既然标准正交基有这么好的性质，那么我们该如何得到一组标准正交基呢？于是乎，就有了下面的定理

定理 5.2 (Schmidt 正交化)

取欧氏空间 V 中的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ ，令

$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1$$

$$\vdots$$

$$\beta_k = \alpha_k - \frac{\langle \alpha_k, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle} \beta_1 - \frac{\langle \alpha_k, \beta_2 \rangle}{\langle \beta_2, \beta_2 \rangle} \beta_2 - \cdots - \frac{\langle \alpha_k, \beta_{k-1} \rangle}{\langle \beta_{k-1}, \beta_{k-1} \rangle} \beta_{k-1}$$

则 $\beta_1, \beta_2, \cdots, \beta_n$ ，为 V 中的一组正交基再令

$$\gamma_i = \frac{\beta_i}{|\beta_i|}, (i = 1, 2, \cdots, n)$$

则 $\gamma_1, \gamma_2, \cdots, \gamma_n$ ，为 V 中的一组标准正交基。

其中第一步是正交化，第二步是单位化

这个算法，就是我们求标准正交向量组的方法。

现在，我们来定义正交矩阵

定义 5.7

如果 n 阶实方阵 Q 满足 $Q^T Q = E$, 则称 Q 为正交矩阵。



由这个定义出发, 我们能推出以下结论

结论 1. 正交矩阵是可逆矩阵, 并且其逆矩阵也为正交矩阵

2. 两个正交矩阵的乘积还是正交矩阵

3. n 阶实方阵 Q 是正交矩阵的充要条件是 $Q^T = Q^{-1}$

4. Q 是正交矩阵的充要条件是 Q 的行 (列) 向量是欧氏空间 \mathbb{R}^n 的标准正交基

结论 设 $\gamma_1, \gamma_2, \dots, \gamma_n$, 为 V 中的一组标准正交基

$$(\beta_1, \beta_2, \dots, \beta_n) = (\gamma_1, \gamma_2, \dots, \gamma_n)Q$$

则 $\beta_1, \beta_2, \dots, \beta_n$, 为 V 中的一组标准正交基当且仅当 Q 是正交矩阵

效仿向量组之间的正交, 我们可以将其推广到子空间的正交。

定义 5.8

设 U 为欧氏空间 V 的欧氏子空间, 如果向量 β 与 U 中所有向量正交, 则称向量 β 与 U 正交, 记为 (β, U) 或 $\beta \perp U$



由这个定义, 我们可以立即推出一个结论

结论 向量 β 与 U 正交的充要条件是向量 β 与 U 的基向量正交

接下来, 我们有这样一个定义

定义 5.9

设 U 为欧氏空间 V 的欧氏子空间, 若存在 V 的子空间 W , 使得 $V = U \oplus W$ 且 $U \perp W$, 则称 W 是 U 在 V 中的补空间



那么问题来了, 这样的一个补空间是唯一的么? 答案是补空间是唯一的, 我们将其记为 U^\perp

我们知道, 在确定一种内积之后, 其所对应的欧氏空间也就确定, 这就造成了欧氏空间的多样性。难道我们没有一种统一的标准, 来研究不同的欧氏空间么? 带着这个问题, 我们先提出一个定义

定义 5.10

设 U, V 是两个欧氏空间, $\varphi: V \rightarrow U$ 是线性映射, 如果 φ 是同构映射且满足: 对于任意的 $\alpha, \beta \in V$, 有

$$(\varphi(\alpha), \varphi(\beta)) = (\alpha, \beta)$$

我们则称 φ 是欧氏空间的同构映射, 也称 U, V 是同构的欧氏空间, 记为 $U \cong V$



在这样一个定义下, 我们能得到一个强有力的结论, 这个结论帮助我们统一了欧氏空间

结论 任意 n 维欧氏空间都同构于欧氏空间 \mathbb{R}^n

所以说, 研究任意 n 维的欧氏空间, 本质上就是在研究欧氏空间 \mathbb{R}^n , 由此, 我们又可以得到一个推论。

推论 5.1

两个有限维欧氏空间同构的充要条件是它们的维数相等



到此为止, 欧氏空间宏观的性质已经结束, 说到底, 我们都是在研究欧氏空间 \mathbb{R}^n , 接下来我们来探寻在以这样一个欧氏空间下的一些线性映射的性质。

我们知道, 在同一线性变换下, 不同基所对应的矩阵是相似的。由之前推出的结论, 如果这个基为标准正交基, 那么就能推出其所对应的矩阵是正交矩阵这一个强力的结论, 如果是在那么同一线性变换下的不同标准正交基所对应的矩阵又有什么关系呢?

定义 5.11

设 $\xi_1, \xi_2, \dots, \xi_n, \eta_1, \eta_2, \dots, \eta_n$ 为 n 维欧式空间 V 的标准正交基, 从 $\xi_1, \xi_2, \dots, \xi_n$ 到 $\eta_1, \eta_2, \dots, \eta_n$ 的过渡矩阵为正交矩阵 Q , 即

$$(\eta_1, \eta_2, \dots, \eta_n) = (\xi_1, \xi_2, \dots, \xi_n)Q$$

设 φ 是 V 的线性变换, φ 在两个标准正交基下的矩阵分别是 A, B , 即

$$\varphi(\xi_1, \xi_2, \dots, \xi_n) = (\xi_1, \xi_2, \dots, \xi_n)A$$

$$\varphi(\eta_1, \eta_2, \dots, \eta_n) = (\eta_1, \eta_2, \dots, \eta_n)B$$

则有 $B = Q^{-1}AQ = Q^T AQ$

我们称这一种关系为 A 正交相似于 B



我们知道, 在平面上向量绕定点旋转保持长度不变, 向量关于一条直线的反射保持向量的长度, 两个向量的夹角。由此, 我们来看一种线性变换

定义 5.12

设 φ 是 V 的线性变换, φ 保内积, 即 且满足: 对于任意的 $\alpha, \beta \in V$, 有

$$(\varphi(\alpha), \varphi(\beta)) = (\alpha, \beta)$$

我们则称 φ 是正交变换



由此, 我们能得到以下结论

结论 设 φ 是 V 的线性变换, 则下列命题等价:

1. φ 是正交变换
2. φ 保持长度不变, 即对于任意的 $\alpha \in V$, 有 $|\varphi(\alpha)| = |\alpha|$
3. φ 将 V 的标准正交基变为标准正交基
4. φ 在 V 下的矩阵是正交矩阵

我们考虑这样一个正交矩阵:

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

我们发现, 这个矩阵在实数域上没有特征值, 在复数域上有特征值 i 和 $-i$, 由此我们能得到以下结论:

结论 设 n 阶实矩阵 A 是正交矩阵, 则

1. $|A| = \pm 1$
2. A 在 \mathbb{C} 上的特征值的长度为 1

由于正交投影具有以下性质

$$(\varphi(\alpha), \beta) = (\alpha, \varphi(\beta)), \forall \alpha, \beta \in V$$

由此我们抽象这样一个定义

定义 5.13

设 φ 是 V 的线性变换, φ 保内积, 即 且满足: 对于任意的 $\alpha, \beta \in V$, 有

$$(\varphi(\alpha), \beta) = (\alpha, \varphi(\beta))$$

我们则称 φ 是对称变换



由此, 我们能得到以下结论

结论 设 φ 是 V 的线性变换, 则下列命题等价:

1. φ 是对称变换

2. 存在 V 的一个标准正交基 $\xi_1, \xi_2, \dots, \xi_n$, 使得 $(\varphi(\xi_i), \xi_j) = (\xi_i, \varphi(\xi_j)) (i, j = 1, 2, \dots, n)$ (当然, 对任意的标准正交基也成立)

3. φ 在 V 中任意一个正交基下的矩阵实对称矩阵

下面有两个定理, 十分的重要

定理 5.3

设 A 是 n 阶实对称矩阵, 则 A 的特征值全为实数, 并且属于不同特征值的特征向量在 \mathbb{R}^n 中相互正交

定理 5.4

设 A 是 n 阶实对称矩阵, 则存在正交矩阵 Q , 使得 $Q^{-1}AQ = Q^T A Q$ 为对角矩阵, 且对角元为 A 的特征值。

我们发现, 仅仅是多了实对称的一个条件, 就直接蹦出了如定理 1 的特征值全为实数的强力条件呢? 有, 还牛的不得了!

结论 设 A, B , 为实对称矩阵, 则 A 正交相似于 B 的充要条件是 A, B 有相同的特征值

这个条件可就厉害得很了, 之前我们说有相同的特征值知识矩阵相似的必要条件, 并不能推出这些矩阵相似, 可是对于是实对称矩阵的话, 这个就是一个充要条件, 这就好的不行。对于实对称矩阵, 我们就没必要要求它们的不变因子什么的了, 直接算它们的特征值我们就能判断其是否相似。

5.3 正交矩阵性质总结

这个部分我们将列举出正交矩阵常考的性质, 加以梳理, 并在此给出证明

结论 A 是正交矩阵, $|A| = \pm 1$

证明:

由 A 是正交矩阵, 则有 $AA^T = E$, 等号两边同取行列式, 可得 $|A|^2 = 1$, 解得 $|A| = \pm 1$

结论 $A^{-1} = A^T$

证明:

由 $AA^T = E$ 可以直接得到

结论 两个正交矩阵的乘积依旧是正交矩阵

证明:

设 A, B 为正交矩阵, 则

$$AB(AB)^T = ABB^T A^T = E$$

结论 A 是正交向量 $\iff A$ 的行 (列) 向量是标准内积空间 \mathbb{R}^n 的标准正交基。

证明:

不妨证明列向量是标准正交基

设

$$A = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

又 $AA^T = E$

则

$$\alpha_i^T \alpha_j = \delta_{ij}, i, j = 1, 2, \dots$$

则在标准内积空间 \mathbb{R}^n 中

$$(\alpha_i^T, \alpha_j) = \delta_{ij}, i, j = 1, 2, \dots$$

则 A 的列向量是标准内积空间 \mathbb{R}^n 的标准正交基。行向量同理。

结论 正交变换不改变向量之间的内积，向量的长度

证明：

由正交变换的保内积性显然

结论 A 是正交矩阵， A 如果有实数特征值，只能为 ± 1

注：正交矩阵不一定有实数特征值，比如 $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

证明：

设 λ 为 A 的特征值，则有

$$A\alpha = \lambda\alpha (\alpha \neq 0)$$

同时取转置

$$\alpha^T A^T = \lambda \alpha^T$$

$$\text{则 } \alpha^T A^T A \alpha = \alpha^T \alpha = \lambda^2 \alpha^T \alpha$$

$$\text{故 } (\lambda^2 - 1) \alpha^T \alpha = 0$$

$$\text{又 } \alpha^T \alpha \neq 0, \text{ 则 } \lambda = \pm 1$$

结论 关于 $|A \pm E|$ 的性质

A 是 n 阶正交矩阵，则有

$$(|A| - 1)|A + E| = 0$$

$$(|A| - (-1)^n)|A - E| = 0$$

证明：

$$|A^T||A + E| = |A^T A + A^T E| = |E + A^T| = |A + E|$$

$$\Rightarrow (|A| - 1)|A + E| = 0$$

$$|A^T||A - E| = |A^T A - A^T E| = |E - A^T| = (-1)^n |A + E|$$

$$\Rightarrow (|A| - (-1)^n)|A - E| = 0$$

结论 设 A 是正交矩阵，若 A 有特征值 ± 1 ，则 -1 和 1 对应的特征向量正交

证明：

设 1 对应的特征向量为 α ， -1 对应的特征向量为 β

$$\text{则有 } A\alpha = \alpha, A\beta = -\beta (\alpha, \beta \neq 0)$$

$$\text{则 } \alpha^T = \alpha^T A^T$$

$$\alpha^T (-\beta) = \alpha^T A^T A \beta = \alpha^T \beta$$

$$\Rightarrow \alpha^T \beta = 0$$

结论 设 A 是三阶非零正交矩阵，则

$$|A| = 1 \iff A_{ij} = a_{ij}$$

$$|A| = -1 \iff A_{ij} = -a_{ij}$$

证明：

$$A^* = |A|A^{-1} = \pm A^T$$

$$\text{又 } A^* = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}, A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

$$\text{则 } A_{ij} = \pm a_{ij}$$

5.4 实对称矩阵性质总结

- 结论**
1. 实对称矩阵一定与对角矩阵相似
 2. 实对称矩阵的不同特征值对应的不同特征向量相互正交
 3. 实对称矩阵可用正交矩阵相似对角化

定理 5.5 (谱分解定理)

以三阶实对称矩阵 A 为例, 若 $\lambda_1, \lambda_2, \lambda_3$ 为 A 的三个特征值, 且 e_1, e_2, e_3 是属于特征值 $\lambda_1, \lambda_2, \lambda_3$ 的两两正交的单位正交向量, 则 $A = \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \lambda_3 e_3 e_3^T$

证明:

A 为实对称矩阵, 则存在正交矩阵 $Q = (e_1, e_2, e_3)$

$$\text{使得 } Q^T A Q = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

$$\text{反解可得 } A = (e_1, e_2, e_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} (e_1, e_2, e_3)^T$$

$$= \lambda_1 e_1 e_1^T + \lambda_2 e_2 e_2^T + \lambda_3 e_3 e_3^T$$

这个结论有多好用? 来一题

A 特征值为 2, 1, 1, $\xi_1 = (1, 1, 1)^T$ 是属于特征值 2 的特征向量, 求实对称矩阵 A 解:

$A - E$ 的特征值为 1, 0, 0, $\xi_1 = (1, 1, 1)^T$ 是属于特征值 1 的特征向量

将 ξ_1 单位化, 得到 $\frac{1}{\sqrt{3}}(1, 1, 1)^T$

$$\text{则 } A - E = 1 \times \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

直接秒杀了, 快的很

第6章 二次型

6.1 引言

为了把二次曲面 S 化简, 需要做直角坐标变换, 使得 S 的新方程不含交叉项, 从而抽象出: 把一个二次齐次多项式化成只含平方项的形式, 这个就是我们要研究的问题中心。

6.2 前置知识

定义 6.1

数域 F 上的一个 n 元二次型是系数在 F 中的 n 个变量的二元齐次多项式, 它的一般形式是

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = & a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{1n}x_1x_n \\ & + a_{22}x_2^2 + \dots + 2a_{2n}x_2x_n \\ & \dots \\ & + a_{nn}x_n^2 \end{aligned} \quad (6.1)$$

其也可写成

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j \quad (6.2)$$

其中 $a_{ij} = a_{ji}, 1 \leq i, j \leq n$



我们可以把 (6.2) 式中的系数按照原来的顺序排列成一个矩阵 A 其中

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix} \quad (6.3)$$

我们则称 A 是二次型 $f(x_1, x_2, \dots, x_n)$ 的矩阵, 它是对称矩阵。显然二次型 $f(x_1, x_2, \dots, x_n)$ 的矩阵是唯一的: 它的主对角元是 x_1^2, \dots, x_n^2 的系数, 它的 (i, j) 元是 $x_i x_j$ 系数的一半, 其中 $i \neq j$ 。令

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (6.4)$$

则二次型 (6.1) 可以写成

$$f(x_1, x_2, \dots, x_n) = X^T A X \quad (6.5)$$

其中 A 是二次型 $f(x_1, x_2, \dots, x_n)$ 的矩阵。

为了讨论方便, 允许 (6.1) 系数全为 0, 即 $A = 0$ 。

令 $Y = (y_1, y_2, \dots, y_n)^T$, 设 C 是数域 F 上的 n 阶可逆矩阵, 则关系式

$$X = CY \quad (6.6)$$

称为 x 到 y 的一个非退化线性替换

n 元二次型 $X^T A X$ 经过非退化线性替换 $X = CY$ 变成

$$(CY)^T A (CY) = Y^T C^T A C Y \quad (6.7)$$

记 $B = C^T AC$, 则 (6.7) 可写为 $Y^T BY$, 这是变量 y_1, y_2, \dots, y_n 的一个二次型, 由于

$$B^T = C^T AC = B \quad (6.8)$$

于是 B 是对称矩阵, 则 B 是二次型 $Y^T BY$ 的矩阵。由此受到启发, 我们引出下述两个概念

定义 6.2

数域 F 上的两个 n 元二次型 $X^T AX, Y^T BY$, 如果存在一个非退化线性替换 $X = CY$, 把 $X^T AX$ 变为 $Y^T BY$, 那么称 $X^T AX$ 与 $Y^T BY$ 等价, 记作 $X^T AX \cong Y^T BY$.



定义 6.3

数域 F 上的两个 n 阶矩阵, A, B 如果存在数域 F 上的 n 阶可逆矩阵 C 使得

$$C^T AC = B \quad (6.9)$$

那么称 A 与 B 合同, 记作 $A \simeq B$



从 (6.7) 容易看出:

命题 6.1

数域 F 上的两个 n 元二次型 $X^T AX, Y^T BY$ 等价当且仅当 n 阶对称矩阵 A 与 B 合同



本章研究的基本问题是: 数域 F 上 n 元二次型能不能等价于一个只含平方项的二次型? 容易看出, 二次型只含平方项的充要条件是它的矩阵是对角矩阵。因此用矩阵的术语, 研究的就是: 数域 F 上 n 阶对称矩阵能不能合同与一个对角矩阵?

如果一个二次型 $X^T AX$ 等价于一个只含平方项的二次型, 那么这个只含平方项的二次型称为 $X^T AX$ 的一个**标准型**。

如果对称矩阵 A 合同与一个对角矩阵, 那么这个对角矩阵称为 A 的一个**合同标准型**

命题 6.2

实数域上的 n 元二次型 $X^T AX$ 有一个标准型为

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 \quad (6.10)$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的全部特征值。



如果 C 为正交矩阵, 我们便称变量的替换 $X = CY$ 为**正交替换**

引理 6.1

设 A, B 是数域 F 上的 n 阶矩阵, 则 A 合同于 B 当且仅当 A 经过一系列成对的初等行、列变换可以变成 B , 此时对 E 只作其中的初等列变换得到的可逆矩阵 C , 就使得 $C^T AC = B$



定理 6.1

数域 F 上任一对称矩阵都合同于一个对角矩阵



定理 6.2

数域 F 上任一 n 元二次型都等价于一个只含平方项的二次型



由引理 6.1, 定理 6.1 定理 6.2 我们就可以得到求二次型的标准型的有一种方法: 对于数域 F 上 n 元二次型 $X^T AX$,

$$\left(\begin{array}{c} A \\ E \end{array} \right) \xrightarrow[\text{对 } E \text{ 只作其中的初等列变换}]{\text{对 } A \text{ 作成对的初等行、列变换}} \left(\begin{array}{c} D \\ C \end{array} \right) \quad (6.11)$$

其中 D 是对角矩阵 $\text{diag}\{d_1, d_2, \dots, d_n\}$, 则

$$C^T AC = D \quad (6.12)$$

令 $X = CY$, 则得到 $X^T AX$ 的一个标准形

$$d_1 y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2 \quad (6.13)$$

上述方法我们称之为矩阵的合同变换。

命题 6.3

数域 F 上任一 n 元二次型 $X^T AX$ 的任一标准形中, 系数不为 0 的平方项个数等于它的矩阵 A 的秩。

于是, 我们便称二次型 $X^T AX$ 的矩阵 A 的秩为二次型 $X^T AX$ 的秩

我们知道, 一个二次型的标准形是不唯一的, 这个取决于你所采用的非退化线性替换。现在我们来讨论二次型的那些量于所作的非退化线性替换有关。我们已经得到了数域 F 上任一 n 元二次型 $X^T AX$ 的任一标准形中, 系数不为 0 的平方项个数等于它的矩阵 A 的秩, 这个与它所作的线性替换无关。二次型的标准形中还有那些量具有这种性质? 在一个二次型的等价类里最简单的二次型形式是怎么样的, 其是否唯一?

我们首先对实数域上的二次型进行讨论。实数域上的二次型简称实二次型。

n 元实二次型 $X^T AX$ 经过一个适当的非退化线性替换 $X = CY$ 可以化为下述形式的标准形:

$$d_1 y_1^2 + \dots + d_p y_p^2 - d_{p+1} y_{p+1}^2 - \dots - d_r y_r^2 \quad (6.14)$$

其中 $d_i > 0, i = 1, 2, \dots, r$ 由命题 (6.3) 可知, r 为这个二次型的秩, 由此我们再作一个非退化线性替换:

$$y_i = \frac{1}{\sqrt{d_i}} z_i, i = 1, 2, \dots, r$$

$$y_j = z_j, j = r+1, \dots, n$$

则二次型 (6.14) 可变为

$$z_1^2 + \dots + z_p^2 - z_{p+1}^2 - \dots - z_r^2 \quad (6.15)$$

因此实二次型 $X^T AX$ 有形如 (6.15) 的一个标准形, 我们称之为 $X^T AX$ 的规范形, 其特征是: 只含平方项, 且平方项的系数为 1, -1, 0; 系数为 1 的平方项都在前面。实二次型 $X^T AX$ 的规范形 (6.15) 被两个自然数 p 和 r 决定。

定理 6.3 (惯性定理)

n 元实二次型 $X^T AX$ 的规范形是唯一的。

定义 6.4

在实二次型 $X^T AX$ 的规范形中, 系数为 +1 的平方项个数 p 称为 $X^T AX$ 的正惯性指数, 系数为 -1 的平方项个数 $r - p$ 称为 $X^T AX$ 的负惯性指数; 正惯性指数减去负惯性指数所得的差 $2p - r$ 称为 $X^T AX$ 的符号差。

由此可知, 实二次型 $X^T AX$ 规范形被它的秩和正惯性指数决定。由此, 我们可以得到。

命题 6.4

两个 n 元二次型等价 \iff 它们的规范形相同 \iff 它们的秩相等, 并且正惯性指数也相等。

从实二次型 $X^T AX$ 经过非退化线性替换化成规范形的过程中我们可以看到, $X^T AX$ 的任意标准形中系数为正的平方项个数等于 $X^T AX$ 的正惯性指数; 系数为负的平方项个数等于 $X^T AX$ 的负惯性指数。从而虽然 $X^T AX$ 的标准形不唯一, 但是标准形中系数为正的平方项个数是唯一的, 系数为负的平方项个数也是唯一的。

从惯性定理我们能得到:

推论 6.1

任一 n 级实对称矩阵 A 合同于对角矩阵 $\text{diag}\{1, \dots, 1, -1, \dots, -1, 0, \dots, 0\}$, 其中 1 的个数等于 X^TAX 的正惯性指数, -1 的个数等于 X^TAX 的负惯性指数, 这个对角矩阵称为 A 的合同规范形



从上面的结论容易得出, n 阶实对称矩阵 A 的合同标准形中, 主对角元为正 (负) 指数的个数等于 X^TAX 的正 (负) 惯性指数。

从命题 (6.4) 我们能立即得到:

推论 6.2

两个 n 阶实对称矩阵合同 \iff 它们的秩相等, 并且正惯性指数也相等。



推论 (6.2) 表明, 秩和正惯性指数恰好完全决定 n 阶实对称的合同类, 因此由所有 n 阶实对称矩阵组成的集合中, 秩和正惯性指数是合同关系下的一组完全不变量。

现在讨论复数域上的二次型, 简称为复二次型。

n 元复二次型 X^TAX 经过一个适当的非退化线性替换 $X = CY$ 可以化为下述形式的标准形:

$$d_1 y_1^2 + \dots + d_r y_r^2 \quad (6.16)$$

其中 $d_i > 0, i = 1, 2, \dots, r; r$ 为这个二次型的秩。

设 $d_j = r_j (\cos \theta_j + i \sin \theta_j)$ 。其中 $0 \leq \theta_j \leq 2\pi$, 显然有

$$\left[\pm \sqrt{r_j} \left(\cos \frac{\theta_j}{2} + i \sin \frac{\theta_j}{2} \right) \right]^2 = d_j$$

把 $\left[\sqrt{r_j} \left(\cos \frac{\theta_j}{2} + i \sin \frac{\theta_j}{2} \right) \right]$ 记作 $\sqrt{d_j}$, 再作一个非退化线性替换:

$$y_i = \frac{1}{\sqrt{d_i}} z_i, i = 1, 2, \dots, r$$

$$y_j = z_j, j = r+1, \dots, n$$

则得到 X^TAX 的下述形式的标准形

$$z_1^2 + \dots + z_r^2 \quad (6.17)$$

我们把这个称为复二次型 X^TAX 的**规范形**。其特征是: 只含平方项, 且平方项的系数为 0 或 1。显然, 复二次型 X^TAX 的规范性完全由它的秩所决定。于是有

定理 6.4

复二次型 X^TAX 的规范形是唯一的。



类比可得

命题 6.5

两个 n 元复二次型等价 \iff 它们的规范形相同 \iff 它们的秩相等

**推论 6.3**

任一 n 阶复对称矩阵 A 合同于对角阵

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

其中 $r = \text{rank}(A)$



推论 6.4

两个 n 阶复对称矩阵合同 \iff 它们的秩相等



由推论 (6.4) 可得出, 秩是 n 阶复对称矩阵组合的集合在合同关系下的完全不变量。

定义 6.5

如果对于 R^n 中任意非零列向量 α 都有 $\alpha^T A \alpha > 0$, 我们称 n 元实二次型 $X^T A X$ 是正定的

**定理 6.5**

n 元实二次型 $X^T A X$ 是正定的当且仅当它的正惯性指数为 n



从定理 (6.5) 我们立即能推出:

推论 6.5

n 元实数二次型 $X^T A X$ 是正定的 \iff 它的规范形为 $y_1^2 + \dots + y_n^2 \iff$ 它的标准形中 n 个系数全大于零。

**定义 6.6**

如果实二次型 $X^T A X$ 是正定的, 那么实对称矩阵 A 称为正定的, 即对于 R^n 中任意非零列向量 α 都有 $\alpha^T A \alpha > 0$



正定的实对称矩阵称为正定矩阵。

从上面的结论, 我们可以推出:

定理 6.6

n 阶实对称矩阵 A 是正定的 $\iff A$ 的正惯性指数为 $n \iff A \simeq E \iff A$ 的合同标准形中主对角元全大于 0 $\iff A$ 的特征值全大于 0

**推论 6.6**

与正定矩阵合同的也是正定矩阵。

**推论 6.7**

与正定二次型等价的实二次型也是正定的, 从而非退化线性替换不改变实二次型的正定性。

**推论 6.8**

正定矩阵的行列式大于 0

**定理 6.7**

实对称矩阵 A 是正定的充要条件是 A 的所有顺序主子式全大于 0



由该定理, 我们能立即得到

推论 6.9

实二次型 $X^T A X$ 是正定的充要条件是 A 的所有顺序主子式全大于 0

**定义 6.7**

n 元二次型 $X^T A X$ 称为半正定 (负定, 半负定的), 如果对于 R^n 中任意非零列向 α , 都有

$$\alpha^T A \alpha \geq 0 (\alpha^T A \alpha < 0, \alpha^T A \alpha \leq 0)$$

如果 X^TAX 既不是半正定的, 又不是半负定的, 那么称它为不定的。

定义 6.8

如果实二次型 X^TAX 是半正定(负定, 半负定的), 那么实对称矩阵 A 是半正定(负定, 半负定的)

定理 6.8

n 元实二次型 X^TAX 是半正定的 \iff 它的正惯性指数等于它的秩 \iff 它的规范形为 $y_1^2 + \cdots + y_r^2 (0 \leq r \leq n) \iff$ 它的标准形中 n 个系数全非负

由定理 (6.8) 我们能立即得到

推论 6.10

n 阶实对称矩阵是半正定的 $\iff A$ 的正惯性指数等于它的秩 $\iff A \simeq \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ 其中 $r = \text{rank}(A) \iff A$ 的合同标准形中 n 个主对角线元全非负 $\iff A$ 的特征值全非负

定理 6.9

实对称矩阵 A 是半正定的充要条件是 A 的所有主子式全非负

定理 6.10

实对称矩阵 A 负定的充要条件是: 它的奇数阶顺序主子式全小于零, 偶数阶顺序主子式全大于 0.

6.3 重点题型归纳

6.3.1 求二次型对应的矩阵

在开启这一部分的内容前, 我们先要学会读题, 将题目中所给出的条件挖掘出来。

例题 6.1 $A = [\alpha_1, \alpha_2, \alpha_3], \alpha = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, f(x) = X^TAX$ 在正交变换 $X = CY$ 化成 $ay_1^2 + by_2^2 + cy_3^2$

这个是前置的条件, 现在我们加条件进去, 然后翻译翻译。

笔记

1. 若给出 $A\alpha = 0 \iff \alpha$ 是 $AX = 0$ 的非零解 $\iff \alpha_1 + 3\alpha_3 = 0$

2. 若给出 $A\alpha = 3\alpha \iff \alpha$ 是方程 $AX = 3X$ 的非零解 $\iff \alpha$ 是方程 $(3E - AX) = 0$ 的非零解 $\iff \alpha_1 + 3\alpha_3 = \alpha$

3. $(A^{-1} - 3E)\alpha = 0 \Rightarrow A^{-1}\alpha = 3\alpha \Rightarrow A\alpha = \frac{1}{3}\alpha$

4. $(A^* - 3E)\alpha = 0 \Rightarrow A^*\alpha = 3\alpha \Rightarrow A\alpha = \frac{|A|}{3}\alpha$

5. A 的各行各列元素之和全为三 $\Rightarrow A \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

5. $r(A) = 1$ 或 $r(A) = 2 \Rightarrow 0$ 为 A 的二重或一重特征值

证明:

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

相似矩阵有相同的特征值, 右边式子的秩取决于非零特征值的个数。

6. $r(A - 2E) = 1$ 或 $r(A - 2E) = 2 \Rightarrow 2$ 为 A 的二重或一重特征值

证明

$$P^{-1}(A - 2E)P = \begin{pmatrix} \lambda_1 - 2 & 0 & 0 \\ 0 & \lambda_2 - 2 & 0 \\ 0 & 0 & \lambda_3 - 2 \end{pmatrix}$$

相似矩阵有相同的特征值, 右边式子的秩取决于非 2 特征值的个数。

7. $A^2(A - 2E) = 0 \iff \lambda = 0$ 或 2

8. C 的第一列是 $c_1 \Rightarrow Ac_1 = ac_1$

证明:

$$C^T AC = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

$$AC = C \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} = (c_1, c_2, c_3) \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \Rightarrow A(c_1, c_2, c_3) = (ac_1, bc_2, cc_3)$$

6.3.1.1 矩阵可直接求

直接求出 A 的所有元素, 一般从特征多项式、迹、 $|\lambda E - A|$ 、 $r(\lambda E - A)$ 入手

例题 6.2 三元二次型 $X^T AX = cx_1^2 + 2ax_1x_2 + 2bx_1x_3 + 2x_2x_3 (a > b)$ 经过正交变换 $X = PY$ 变为 $y_1^2 + y_2^2 - 2y_3^2$, 求 A

解

先拿基础分, 求二次型的矩阵

$$A = \begin{pmatrix} c & a & b \\ a & 0 & 1 \\ b & 1 & 0 \end{pmatrix}$$

因为是通过正交变换, 所以有相同的特征值, 于是矩阵 A 的特征值为 $1, 1, -2$

由 $\text{tr}(A) = 0$, 可得 $c = 0$

$$|\lambda E - A| = \lambda^3 - (a^2 + b^2 + 1)\lambda - 2ab = (\lambda - 1)^2(\lambda + 2) = \lambda^3 - 3\lambda + 2$$

比较系数, 有

$$\begin{cases} ab = -1 \\ a^2 + b^2 = 2 \end{cases} \Rightarrow \begin{cases} a = 1 \\ b = -1 \end{cases} \quad \text{或} \quad \begin{cases} a = -1 \\ b = 1 \end{cases}$$

例题 6.3 三元二次型 $X^T AX (A = A^T)$ 的平方项系数全为 1, $\alpha = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ 是 $AX = 0$ 的唯一解, 求 A

解

先拿低保

$$A = \begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix}$$

由题, 有

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+a+2b \\ a+1+2c \\ b+c+2 \end{pmatrix} = 0$$

解得

$$\begin{cases} a = 1 \\ b = -1 \\ c = -1 \end{cases}$$

例题 6.4 三元二次型 $X^TAX = ax_1^2 + x_2^2 - x_3^2 + 2x_1x_2 + 2bx_1x_3 + 2cx_2x_3$, A 满足 $AB = 0$, $B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

解

先拿低保

$$A = \begin{bmatrix} a & 1 & b \\ 1 & 1 & c \\ b & c & -1 \end{bmatrix}$$

由题, 有

$$\begin{bmatrix} a & 1 & b \\ 1 & 1 & c \\ b & c & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a-b & 0 & b-a \\ 1-c & 0 & c-1 \\ b+1 & 0 & -1-b \end{bmatrix} = 0$$

解得

$$\begin{cases} a = 1 \\ b = -1 \\ c = 1 \end{cases}$$

例题 6.5 三元二次型 $X^TAX = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2bx_1x_3 + 2cx_2x_3$, $r(A) = 1$, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 是 A 的一个特征向量。

解

先拿低保

$$A = \begin{bmatrix} 1 & 1 & b \\ 1 & 1 & c \\ b & c & 1 \end{bmatrix}$$

由 $r(A) = 1$, 则 A 的特征值为 $0, 0, 3$ 故对其讨论

$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 对应的特征值为 0 时, 有

$$\begin{bmatrix} 1 & 1 & b \\ 1 & 1 & c \\ b & c & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0$$

显然无解。

$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 对应的特征值为 3 时, 有

$$\begin{bmatrix} 1 & 1 & b \\ 1 & 1 & c \\ b & c & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

解得

$$\begin{cases} b = 1 \\ c = 1 \end{cases}$$

6.3.1.2 矩阵不可直接求

像这类题, 我们通常是通过求 A 的特征值 $\lambda_1, \lambda_2, \lambda_3$ 以及对应的特征向量 $\alpha_1, \alpha_2, \alpha_3$ 然后用

$$A = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} (\alpha_1, \alpha_2, \alpha_3)^T$$

求出 A

主要分为以下四种情况

$$\text{三阶实对称矩阵 } A \text{ 的特征值为 } \begin{cases} a, b, b \text{ 且 } a : \alpha \\ a, b, b \text{ 且 } b : \beta_1, \beta_2 \\ a, b, c \text{ 且 } b : \beta, c : \gamma \\ a, a, a \end{cases}$$

有以下结论, 以下出现的特征向量均为我随便写的, 实际上对任意满足下述条件的特征向量均成立。

$$\begin{cases} a, b, b \text{ 且 } a : \alpha = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \text{ 则 } \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \text{ 的非零解都是 } b \text{ 对应的特征向量} \\ a, b, b \text{ 且 } b : \beta_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \beta_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ 则 } \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \text{ 非零解都是 } a \text{ 对应的特征向量} \\ a, b, c \text{ 且 } b : \beta = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, c : \gamma = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ 则 } \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \text{ 非零解都是 } a \text{ 对应的特征向量} \\ a, a, a \text{ 则任意非零向量都是 } a \text{ 对应的特征向量} \end{cases}$$

例题 6.6 已知 3 阶实对称矩阵 A 的各行元素之和均为 1, $r(A - 6E) = 1$

(1) 求正交矩阵 P , 使得 $P^{-1}AP$ 均为对角阵

(2) 求 A

解

(1) 由题意, 有 1 为 A 的特征值且 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 为属于 1 的特征向量, 6 为 A 的二重特征值

由先前结论可知, 属于 0 的特征向量为

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} X = 0$$

的非零解

解得

$$\beta_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \beta_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

正交化 + 单位化, 有 (这里只需要对 β_1, β_2 进行施密特正交化就行了)

$$\gamma_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \gamma_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \gamma_3 = \begin{bmatrix} \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

取 $P = (\gamma_1, \gamma_2, \gamma_3)$, 即

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

此时 P 为正交矩阵, 有

$$P^T A P = P^{-1} A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

(2) 由题, 有

$$A = (\gamma_1, \gamma_2, \gamma_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} (\gamma_1, \gamma_2, \gamma_3)^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{13}{3} & -\frac{5}{3} & -\frac{5}{3} \\ -\frac{5}{3} & \frac{13}{3} & -\frac{5}{3} \\ -\frac{5}{3} & -\frac{5}{3} & \frac{13}{3} \end{pmatrix}$$

由于求 A 过于套路化, 后面的求 A 我都不写出, 有兴趣的自行去做

例题 6.7 已知 3 阶实对称矩阵 A , $Ax = 0$ 的通解为 $k_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $0 < r(A - E) < 3$ 求正交矩阵 P 使得

$P^{-1}AP$ 为对角矩阵

解

由题 $r(A) = 1$, A 的特征值为 $0, 0, 1$

由先前的结论, 属于 1 的特征向量为

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

的非零解。

解得

$$\alpha = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

正交化 + 单位化, 有 (这里只需要对 β_1, β_2 进行施密特正交化就行了)

$$\gamma_1 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \gamma_2 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \gamma_3 = \begin{bmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

剩下过程同上。

例题 6.8 3 阶实对称矩阵 A 满足: $\text{tr}(A) = 1, |A| = -1, \alpha = [1, 1, 1]^T$ 是方程 $(A^* - E)x = 0$ 的解, 求正交矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

因为 $(A^* - E)x = 0$, 由前面结论, 解得 $Ax = -x$, 故 -1 为 A 的一个特征值。

则 A 的特征值为 $0, \lambda_1, \lambda_2$, 由 $\text{tr}(A) = 1, |A| = -1$, 解得特征值为 $1, 1, -1$

由先前结论可知, 属于 1 的特征向量为

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T X = 0$$

的非零解

解得

$$\beta_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \beta_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

正交化 + 单位化, 有 (这里只需要对 β_1, β_2 进行施密特正交化就行了)

$$\gamma_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \gamma_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \gamma_3 = \begin{bmatrix} \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

例题 6.9 3 阶实对称矩阵 $A = [\alpha_1, \alpha_2, \alpha_3], \alpha_1 = 2\alpha_2 + \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix}, r(A) = 1$, 求正交矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

$$\alpha_1 = 2\alpha_2 + \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix} \Rightarrow \alpha_1 - 2\alpha_2 = \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = A \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

于是 2 为 A 的一个特征值, 又 $r(A) = 1$, 于是 $0, 0, 1$ 为 A 的所有特征值。

属于 0 的所有特征向量为方程

$$\begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

的非零解。

剩下的过程自己写一下, 都是一样的。

例题 6.10 3 阶实对称矩阵 $A = [\alpha_1, \alpha_2, \alpha_3], \alpha_1 = 2\alpha_2 + \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix} = -\frac{1}{2}\alpha_2 - \alpha_3 + \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}, |\alpha_1, \alpha_2, \alpha_3| = 0$, 求正交

矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

$$\alpha_1 = 2\alpha_2 + \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix} \Rightarrow \alpha_1 - 2\alpha_2 = \begin{bmatrix} 2 \\ -4 \\ 0 \end{bmatrix} =$$

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = A \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

于是 2 为 A 的一个特征值

$$\alpha_1 = -\frac{1}{2}\alpha_2 - \alpha_3 + \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \Rightarrow \alpha_1 + \frac{1}{2}\alpha_2 + \alpha_3 = [\alpha_1, \alpha_2, \alpha_3] \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} \Rightarrow A \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

于是 4 为 A 的一个特征值

又 $|A| = 0$, 0 也为 A 的一个特征值。

于是属于特征值 0 的特征向量为

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

的非零解, 剩下同上

例题 6.11 3 阶实对称矩阵 $A = [\alpha_1, \alpha_2, \alpha_3]$, $A \begin{bmatrix} 4 & 0 \\ -3 & 5 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 4 \\ -4 & 12 \\ 8 & 8 \end{bmatrix}$, $|\alpha_1, \alpha_2, \alpha_3| = 0$ 求正交矩阵 P 使得 $P^{-1}AP$

为对角矩阵

解

很明显, 这对应的行和列是不成比例的, 所以我们无法直观地看出特征值和特征向量, 所以我们不妨待定系数, 即

$$A \begin{bmatrix} 4+0k \\ -3+5k \\ 2+2k \end{bmatrix} = \begin{bmatrix} 12+4k \\ -4+12k \\ 8+8k \end{bmatrix} \Rightarrow \begin{cases} 12+4k = 4\lambda \\ -4+12k = (-3+5k)\lambda \\ 8+8k = (2+2k)\lambda \end{cases} \Rightarrow \begin{cases} \lambda = 4 \\ k = 1 \end{cases} \quad \text{或} \quad \begin{cases} \lambda = 2 \\ k = -1 \end{cases}$$

故 A 的属于特征值 4 的特征向量为 $\begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$ A 的属于特征值 2 的特征向量为 $\begin{bmatrix} 4 \\ -8 \\ 0 \end{bmatrix}$

又 $|A| = 0$, A 的特征值为 0, 4, 2

于是属于特征值 0 的特征向量为

$$\begin{bmatrix} 4 & 2 & 4 \\ 4 & -8 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

剩下的过程就是算算算。

例题 6.12 A 是 3 阶实对称矩阵, α, β 线性无关且满足 $A\alpha = 2\beta$, $A\beta = 4\alpha$, $\gamma \neq 0$ 且满足 $A\gamma = 0$. 求 A 的所有特征值以及对应的特征向量, 可逆矩阵 P 使得 $P^{-1}AP$ 为对角矩阵. (P 要用 α, γ, β 表示)

解

容易证明 α, γ, β 线性无关, 由题则有

$$A(\alpha, \beta, \gamma) = (\alpha, \beta, \gamma) \begin{pmatrix} 0 & 4 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow (\alpha, \beta, \gamma)^{-1} A(\alpha, \beta, \gamma) = \begin{pmatrix} 0 & 4 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = B$$

$|\lambda E - B| = \lambda(\lambda^2 - 8)$ 解得 B 的特征值为 $0, 2\sqrt{2}, -2\sqrt{2}$ 解得, 分别对应的特征向量为

$$\gamma_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \gamma_2 = \begin{bmatrix} \sqrt{2} \\ 1 \\ 0 \end{bmatrix}, \gamma_3 = \begin{bmatrix} -\sqrt{2} \\ 1 \\ 0 \end{bmatrix}$$

于是存在一个可逆阵 $Q = (\gamma_1, \gamma_2, \gamma_3)$, 使得 $Q^{-1}BQ = Q^{-1}(\alpha, \beta, \gamma)^{-1}A(\alpha, \beta, \gamma)Q = \begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & -2\sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$$\text{此时 } P = (\alpha, \beta, \gamma)Q = (\alpha, \beta, \gamma) \begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

例题 6.13 A 是 3 阶实对称矩阵, $\alpha_1, \alpha_2, \alpha_3$ 线性无关且满足 $A(\alpha_1 - 2\alpha_2 - \alpha_3) = -2\alpha_2, A\alpha_2 = \alpha_3, A\alpha_3 = \alpha_1$ 。求 A 的所有特征值以及对应的特征向量, 可逆矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

法一:

由题, 有

$$A(\alpha_1 - 2\alpha_2 - \alpha_3, \alpha_2, \alpha_3) = A(\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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

$$\text{于是有 } (\alpha_1, \alpha_2, \alpha_3)^{-1} A(\alpha_1, \alpha_2, \alpha_3) = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = B$$

$$|\lambda E - B| = (\lambda - 1)(\lambda^2 - 2)$$

解 $(\sqrt{2}E - B)X = 0$, 得

$$\begin{bmatrix} \frac{1}{\sqrt{2}-1} \\ -\frac{\sqrt{2}}{\sqrt{2}-1} \\ 1 \end{bmatrix}$$

解 $(-\sqrt{2}E - B)X = 0$, 得

$$\begin{bmatrix} -\frac{1}{\sqrt{2}+1} \\ -\frac{\sqrt{2}}{\sqrt{2}+1} \\ 1 \end{bmatrix}$$

解 $(E - B)X = 0$ 得

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

于是, $P = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} \frac{1}{\sqrt{2}-1} & -\frac{1}{\sqrt{2}+1} & -\frac{1}{2} \\ -\frac{\sqrt{2}}{\sqrt{2}-1} & -\frac{\sqrt{2}}{\sqrt{2}+1} & 1 \\ 1 & 1 & 0 \end{pmatrix}$

法二: 直接待定系数

$$A(\alpha_1 - 2\alpha_2 - \alpha_3 + m\alpha_2 + n\alpha_3) = -2\alpha_2 + m\alpha_3 + n\alpha_1 = \lambda(\alpha_1 - 2\alpha_2 - \alpha_3 + m\alpha_2 + n\alpha_3)$$

比较系数, 得

$$\begin{cases} n = \lambda \\ -2 = -2\lambda(m-2) \\ m = \lambda(n-1) \end{cases} \implies \begin{cases} m = 0 \\ n = 1 \\ \lambda = 1 \end{cases} \quad \text{或} \quad \begin{cases} m = 2 - \sqrt{2} \\ n = \sqrt{2} \\ \lambda = \sqrt{2} \end{cases} \quad \text{或} \quad \begin{cases} m = 2 + \sqrt{2} \\ n = -\sqrt{2} \\ \lambda = -\sqrt{2} \end{cases}$$

例题 6.14 A 是 3 阶实对称矩阵, 满足 $A^2(A - 2E) = 0$, $r(A) = 2$, $A = [\alpha_1, \alpha_2, \alpha_3]$, $\alpha_1 = 2\alpha_2$. 求 A 的所有特征值以及对应的特征向量, 可逆矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

由题 $A^2(A - 2E) = 0 \Rightarrow \lambda^2(\lambda - 2) = 0$, 解得 A 的特征值为 0 或 2

又 $r(A) = 2$, 于是 A 的特征值为 0, 2, 2

又因为 $(\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = A \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$

于是 $\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ 是属于特征值 0 的特征向量, 剩下的就跟前面一样。

例题 6.15 二次型 $f(x) = X^TAX$ 在正交变换 $X = PY$ 化成 $y_1^2 + 2y_2^2 + 2y_3^2$, P 的第一列为 $\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix}$ 求 A 的所有

特征值以及对应的特征向量使得可逆矩阵 P 使得 $P^{-1}AP$ 为对角矩阵

解

由题, 经过 $X = PY$ 变换之后 $X^TAX = (PY)^TPY = Y^TP^TAPY$

于是有 $(P_1, P_2, P_3)^T A(P_1, P_2, P_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

即 $A(P_1, P_2, P_3) = (P_1, P_2, P_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = (P_1, 2P_2, 2P_3)$

于是 P_1 是属于特征值 1 的特征向量, P_2, P_3 为方程

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

的非零解, 剩下过程同上。

6.3.2 合同变换法

例题 6.16 求 $f(x_1, x_2, \dots, x_n) = 2x_1x_2 + x_1x_3$ 的规范形。

解

先求矩阵拿低保

$$A = \begin{bmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}$$

对矩阵进行合同变换

$$\begin{bmatrix} 0 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{4} \end{bmatrix}$$

即可看出规范形。

6.3.2.1 对角矩阵过渡

例题 6.17 $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 求一个正交矩阵 Q , 使得 $Q^{-1}AQ = B$

解

由题, 存在正交矩阵 Q_1 , 使得 $Q_1^{-1}AQ_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$ 存在正交矩阵 Q_2 , 使得 $Q_2^{-1}BQ_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

于是有

$$Q_1^{-1}AQ_1 = Q_2^{-1}BQ_2$$

则 $Q_2Q_1^{-1}AQ_1Q_2^{-1} = B$ 此时 $Q = Q_1Q_2^{-1}$, 计算过程就省略

例题 6.18 $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

- (1) 求一个可逆矩阵 C , 使得 $C^3 = A$
- (2) 求一个正定矩阵 C , 使得 $C^2 = A + 2E$
- (3) 求一个正定矩阵 C , 使得 $C^n = A + 2E$

解

(1)

解 $|\lambda E - A|$ 得 A 的特征值为 $1, 1, -1$

由题, 存在正交矩阵 Q_1 , 使得 $Q_1^{-1}AQ_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

则 $A = Q_1 \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} Q_1^{-1}$

令 $C = Q_1 \begin{pmatrix} \sqrt[3]{\lambda_1} & 0 & 0 \\ 0 & \sqrt[3]{\lambda_2} & 0 \\ 0 & 0 & \sqrt[3]{\lambda_3} \end{pmatrix} Q_1^{-1}$ 则 C 可逆且满足 $C^3 = A$

代入数据 $C = Q_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} Q_1^{-1}$

(2) 由题, $A + 2E$ 的特征值为 $3, 3, 1$

由题, 存在正交矩阵 Q_2 , 使得 $Q_2^{-1} A Q_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

则 $A = Q_2 \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} Q_2^{-1}$

令 $C = Q_2 \begin{pmatrix} \sqrt{\lambda_1 + 2} & 0 & 0 \\ 0 & \sqrt{\lambda_2 + 2} & 0 \\ 0 & 0 & \sqrt{\lambda_3 + 2} \end{pmatrix} Q_2^{-1}$ 则 C 正交且满足 $C^2 = A + 2E$

代入数据 $C = Q_2 \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} Q_2^{-1}$

剩下自己算, 正交矩阵实际上一个样。

(3) 同上

例题 6.19 若可逆线性替换 $X = PY$ 可将二次型 $f(x_1, x_2) = x_1^2 + 2x_2^2 + 2x_1x_2$ 化为规范形 $y_1^2 + y_2^2$ 同时将二次型 $g(x_1, x_2) = -x_1^2 + 2x_2^2 + 2x_1x_2$ 化为标准形 $k_1y_1^2 + k_2y_2^2$, 求可逆矩阵 P 以及 k_1k_2

解

先求矩阵拿保底

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

由题, 有

$$P^T A P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P^T B P = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \quad (6.18)$$

(6.18) 左右两边同取行列式, 有

$$|P|^2 = 1, -3|P|^2 = k_1k_2$$

(6.18) 左右两边相减后同取行列式, 有

$$(k_1 - 1)(k_2 - 1) = 0$$

解得

$$k_1 = 1 \text{ 或 } k_2 = 1$$

分别对 $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ 和 $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ 作合同变换, 有

$$P_1 = \begin{bmatrix} 0 & \sqrt{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, P_2 = \begin{bmatrix} \sqrt{2} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

6.3.3 正定二次型

定义 6.9

如果对于 R^n 中任意非零列向量 α 都有 $\alpha^T A \alpha > 0$, 我们称 n 元实二次型 $X^T A X$ 是正定的



注

A 不是对称矩阵的时候, A 不是二次型 $X^T A X$ 所对应的矩阵, 此时 $\frac{1}{2}(A + A^T)$ 才是二次型 $X^T A X$ 所对应的矩阵

6.3.3.1 判定正定二次型或者利用正定二次型的性质

在这个地方, 我们需要知道二次型 $X^T A X$ 正定的充要条件, 如下

结论

1. 对于任意非零向量 $X, X^T A X > 0$
2. A 是正定矩阵
3. A 的特征值均为正数
4. A 的正惯性指数为 n
5. 存在可逆阵 C , 使得 $A = C^T C$
6. A 和单位阵合同
7. A 的所有顺序主子式大于 0

例题 6.20 设 A 是 n 阶矩阵, B 是 $m \times n$ 实矩阵, 证明: $B^T A B$ 是正定矩阵的充要条件是 $r(B) = n$

解

\Rightarrow , 对于任意 $X \neq 0$ 都有 $X^T B^T A B X > 0$ 即此时 $(BX) \neq 0$

令 $BX = 0$, 若无非零解, 表明 $r(B) = n$

\Leftarrow 此时 $BX = 0$ 只有零解, 于是对于任意 $X \neq 0$ 都有 $X^T B^T A B X > 0$ 故为正定矩阵

例题 6.21 设 n 元实二次型 $f(x_1, x_2, \dots, x_n) = (x_1 + a_1 x_2)^2 + (x_2 + a_2 x_3)^2 + \dots + (x_{n-1} + a_{n-1} x_n)^2 + (x_n + a_n x_1)^2$, 其中 a_1, a_2, \dots, a_n 为实数, 试问 a_1, a_2, \dots, a_n 满足什么条件时, 二次型 $f(x_1, x_2, \dots, x_n)$ 是正定二次型

解

$$\begin{aligned} \text{由题 } f &= \begin{bmatrix} x_1 + a_1 x_2 \\ x_2 + a_2 x_3 \\ \vdots \\ x_n + a_n x_1 \end{bmatrix}^T \begin{bmatrix} x_1 + a_1 x_2 \\ x_2 + a_2 x_3 \\ \vdots \\ x_n + a_n x_1 \end{bmatrix} = \\ &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}^T \begin{bmatrix} 1 & a_1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & a_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{n-1} \\ a_n & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = X^T A^T A X \end{aligned}$$

若正定, 则 $|A^T A| > 0$, 即 $|A|^2 > 0 \Rightarrow |A| \neq 0$

剩下计算就跳过, 这个行列式直接按行列展开即可

例题 6.22 设矩阵 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, 矩阵 $B = (kE + A)^2$, 求对角矩阵 P 使得 B 与 P 相似, 并求 k 为何值时, B

为正定矩阵

解

解 $|\lambda E - A| = 0$, 得 A 的特征值为 0, 2, 2

B 的特征值为 $(2+k)^2, (2+k)^2, k^2$, 正定矩阵要求特征值全大于 0

则 $k \neq 0$ 且 $k \neq -1$

例题 6.23 设 A 为三阶实对称矩阵, 且满足 $A^2 + A = 0$, 且 $r(A) = 2$, 问 k 为何值时, 矩阵 $A + kE$ 为正定矩阵

解

由题, A 的特征值为 0 或 -1 , 又 $r(A) = 2$, 则 A 的特征值为 $0, -1, -1$, $A + kE$ 的特征值为 $k, k-1, k-1$
又正定矩阵要求特征值全大于 0, 于是有 $k > 1$

例题 6.24 考虑二次型 $x_1^2 + 4x_2^2 + 4x_3^2 + 2\lambda x_1x_2 - 2x_1x_3 + 4x_2x_3$, 当 λ 取得何值时, 该二次型为正定二次型

解

先求矩阵拿低保

$$A = \begin{bmatrix} 1 & \lambda & -1 \\ \lambda & 4 & 2 \\ -1 & 2 & 4 \end{bmatrix}$$

正定矩阵要求所有的顺序主子式 > 0 解得

$$-2 < \lambda < 1$$

例题 6.25 若二次型 $2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + tx_2x_3$, 当 λ 为正定二次型, t 为何值

解

先求矩阵拿低保

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & \frac{t}{2} \\ 0 & \frac{t}{2} & 1 \end{bmatrix}$$

正定矩阵要求所有的顺序主子式 > 0 解得

$$-\sqrt{2} < \lambda < \sqrt{2}$$

n 元二次型 $X^TAX (A = A^T)$ 正定的必要条件如下

结论

1. $|A| > 0$
2. A 主对角线所有元素大于 0
3. A 最大的元素只能位于主对角线上