0.1 线性变换的基本概念与矩阵

0.1.1 线性变换

在之前的研究中, 我们考虑了线性映射 $\mathrm{Hom}(V, U)$ 表示从线性空间 V 映射线性空间 U 的映射的全体. 也就是:

那么如果 V 和 U 恰好是同一个东西, 那么我们就可以记这个是**线性变换**. 用花体的字母写出: 也就是

$$
\begin{tikzcd}
\varphi: & V_/ \arrow[r] & U_/ \arrow[r] \arrow[dr] & \varphi\text{是单射} \arrow[r] & V \simeq \varphi(V) \\
 & & & \varphi\text{是满射} \arrow[r] & U \simeq V/\mathrm{Ker}\varphi
\end{tikzcd}
$$

$\mathrm{Hom}(V, V)$是线性变换: $\mathscr{A}: \qquad V \longrightarrow V \qquad$ (从V 到V 自身的线性映射)

如果 V 是 上的线性空间, 称 V 到其自身的线性映射 $\mathscr{A}$ 为 V 上的**线性变换**. 我们不妨把它记作 $\mathrm{End}(V)$.

定义 0.1

设 V 是 上的线性空间, 称 V 到它自身的线性映射 $\mathscr{A}$ 为 V 上的**线性变换**, 即对于任意的 $, \in V, k, l \in$, 都有

$$\mathscr{A}(k + l) = k\mathscr{A} + l\mathscr{A}.$$

下面来举一些例子:

例子 0.1

数乘变换: TODO

例子 0.2

平面上的旋转变换: TODO

例子 0.3

求导与积分: TODO

有了上面的例子, 我们自然要问:End$V$ 是不是也是 上的一个线性空间? 于是我们给出如下的验证:

- 加法: 如果 $\mathscr{A}\mathscr{B} \in \mathrm{End}V, \forall \alpha \in V$, **定义** $(\mathscr{A}+\mathscr{B})\alpha = \mathscr{A}\alpha + \mathscr{B}\alpha$. 那么 $\mathscr{A}+\mathscr{B} \in \mathrm{End}V$ 吗?
  - $(\mathscr{A}+\mathscr{B})(\alpha+\beta) \xlongequal{\text{线性变换}} (\mathscr{A}+\mathscr{B})\alpha + (\mathscr{A}+\mathscr{B})\beta \xlongequal{\text{加法定义}} \mathscr{A}(\alpha+\beta) + \mathscr{B}(\alpha+\beta)$
- 数乘
  - $k \in, \mathscr{A} \in \mathrm{End}V$, $(k\mathscr{A})() = k(\mathscr{A}) \in \mathrm{End}V$.

由此可得, End V 果然是 F 上的一个线性空间.

既然是一个映射, 我们自然要考虑一下它的合成有什么情形. 更一般的, 我们考虑一下如下内容:

- 交换律: 不满足. 实际上很多情形下交换律都不是天然就有的. 即使在考虑映射的换序的时候也不是这样.
- 结合律: 天然满足.
- 零向量: 恒等变换 id, 也就是 $\mathrm{id} \circ \mathscr{A} = \mathscr{A}$. 有时候也成为**单位元**.
- 逆元: 有没有可逆的元素吗? $\mathscr{B} \circ \mathscr{A} = \mathrm{id}$?
- 乘法与加法: $\mathscr{A} \circ (\mathscr{B}+\mathscr{C}) = \mathscr{A}\circ\mathscr{B} + \mathscr{A}\circ\mathscr{C}$, $(\mathscr{B}+\mathscr{C})\circ\mathscr{A} = \mathscr{B}\circ\mathscr{A} + \mathscr{C}\circ\mathscr{A}$.
- 乘法与数乘: $\mathscr{A}\circ(k\mathscr{B}) = k\mathscr{A}\mathscr{B} = (k\mathscr{A})\mathscr{B}$

并且由此定义乘法 $\mathscr{A}, \mathscr{A}^2 = \mathscr{A}\circ\mathscr{A}, \mathscr{A}^3 = \mathscr{A}\circ\mathscr{A}\circ\mathscr{A}, \cdots$ 这样我们就有了多项式. 定义 $f(\mathscr{A}) = a_i^i$.

命题 0

(线性变换的运算律)

0.1.2 与 $^{n\times n}$ 的等价性

以前我们知道, 线性空间实际上就是和 $^{n\times 1}$ 是等价的, 那么线性变换是不是可以和什么等价呢?

如果我们设 $V/$ 是 n 维线性空间, ${}_1..{}_2$ 是 V 的一组基, $V \xrightarrow{\simeq} {}^{n\times 1}$.

$$\mathscr{A}\in \mathrm{End}V: \qquad \alpha\in V \qquad \mathscr{A}\in V$$

$$||$$

$$\sum x_i a_i = ({}_1..{}_n)X, \qquad X=(x_1..x_n)'$$

并且

$$\begin{aligned}\mathscr{A} &= \left(\sum_{i=1}^{n} x_{ii}\right)\\ &=\sum_{i=1}^{n} x_i\mathscr{A}_i\\ &=\mathscr{A}\left(({}_1..{}_n)X\right)\\ &=(\mathscr{A}_1..\mathscr{A}_n)X\end{aligned}$$

也就是 $\mathscr{A}$ 是由 $\mathscr{A}_1..\mathscr{A}_n$ 决定的.

于是我们就得到了如下的内容:

如果 ${}_1..{}_n$ 是一 $V/$ 的一组基, 也就是

$$\mathrm{End}V \ni \qquad \mathscr{A} \xrightarrow{\text{选择基底}_1..{}_n} (\mathscr{A}_1,..,\mathscr{A}_n) \longrightarrow \{{}^{n\times 1}\text{中}n\text{元组}\} \longrightarrow {}^{n\times n}$$

$$\mathscr{A} \xrightarrow{\text{选择基底}_1..{}_n} (\mathscr{A}_1,..,\mathscr{A}_n) \xrightarrow{\text{用基底}_1..{}_n\text{表示}} A$$

也就是完成了 $\mathrm{End}V \to {}^{n\times n}$, 也就是 $\mathscr{A}\to A$ 的映射.

这就意味着我们就把线性变换和矩阵之间搭了一座桥梁, 抽象的线性变换在具体的基的映射下可以变为具体的矩阵. 这样, 在这个线性变换作用在某个向量上的时候, 我们可以这样得到:

定理 0.2

设 $\in V, \mathscr{A} \in \mathrm{End}V$, 且 ${}_{1}..{}_{n}$ 是 V 的一组基, 在这组基下, $\mathscr{A}$ 的基是 A, 的坐标是 X, 那么 $\mathscr{A}$ 的坐标是 AX, 也就是 $\mathscr{A} = ({}_{1}..{}_{n})AX$.

证明. 用形式记号的语言, 有

$\mathscr{A} = \mathscr{A}(({}_{1}..{}_{n})X) = (\mathscr{A}_1..\mathscr{A}_n)X = (({}_{1}..{}_{n})A)X$, 然后验证一下形式记号的结合律是满足的 (第四章已经验证过), 就有原始式子为 $({}_{1}..{}_{n})(AX)$ 的结论. #

这样我们就在线性变换与矩阵之间架了一座桥梁, 方便我们进一步的探讨方阵与线性变换之间的更多性质.

那么我们不妨考虑设 ${}_{1}..{}_{n}$ 是 V 的一组基, 那么从 $\mathrm{End}V$ 到 ${}^{n\times n}$ 的映射 $\mathscr{A} \to (\mathscr{A})$ 有什么性质?

既然是线性空间上面的线性映射, 我们猜测:

- 是单射: 考虑 $\xrightarrow{1..n} A, \xrightarrow{1..n} B$, 要满足单射的条件, 就要考虑从 $A = B$ 推出 $=$. 要证明右侧的式子, 根据定义, $\forall, =$. 两边展开, 就有 $(\sum_{i=1}^{n} x_i{}_i) = (\sum_{i=1}^{n} x_i{}_i)$, 根据定义, 并且把求和记号转化为矩阵乘法, 看出坐标, 就有 $({}_{1}..{}_{n})X = ({}_{1}..{}_{n})X$. 又根据已知的条件 $A = B$, 就可以得到 ${}_i ={}_i$ $(i \in [1..n])$. 就可以推断出这个结论了.

- 是满射: 问题转化为 $\forall A \in^{n\times n}$, 是否存在 $\mathscr{A} \in \mathrm{End}V$, 使得 $() = A$. 把 A 分块形成若干的 $n \times 1$ 的列向量 $A_1..A_n$, 我们自然希望有合理的对应方式. 也就是, 我们定义 $= (\sum_{i=1}^{n} x_i{}_i) = \sum_{i=1}^{n} x_i{}_i \triangleq \sum_{i=1}^{n} x_i{}_i$. 这样就可以方便的验证 $\in \mathrm{End}V$, $\varphi() = A$.

- $(\mathscr{A}+) = () + ()$. 只需要转化成方阵的形式证明就行了.

- $(k) = k()$, 同上.

从上面的内容, 我们知道了这是一个线性同构.

预告: 代数的结构
如同 ${}^{n\times n}$ 满足结合律, 但不满足交换律. $[x]$ 及班组交换律, 又满足结合律. 然而向量的外积既不满足交换律, 又不满足结合律, 但是它满足反交换律和 Jacobi 恒等式.

事实上, 这样的映射 还有一个比较重要的性质, 就是它保持了乘法.

定理 0.3

设 $_1 \cdots _n V$ 的一组基, 则 End$V$ 到 $^{n\times n}$ 的映射 $\to (A)$ 是线性空间 End$V \to^{n\times n}$ 的线性同构, 并且 $\forall, \in \mathrm{End}V, () = ()()$.

证明. 线性同构见上述描述.

下面证明保持乘法:

$$
\begin{aligned}
()(_1 \cdots _n) &= (()_1 \cdots ()_n) \\
&= ((_1) \cdots (_n)) \\
&= \left(\left((_1 \cdots _n)\, B_1\right) \cdots \left((_1 \cdots _n)\, B_n\right)\right) \\
&= ((_1 \cdots _n) B_1 \cdots (_1 \cdots _n) B_n) \\
&= ((_1 \cdots _n) B \\
&= ((_1 \cdots _n) A_n \cdots (_1 \cdots _n) A_n) B \\
&= ((_1 \cdots _n) A) B \\
&= (_1 \cdots _n)(AB) \qquad \text{(需要验证结合律)}
\end{aligned}
$$

#

0.2 线性变换的矩阵

从上一节的内容我们知道 End$V \xrightarrow[_1 \cdots _n]{1-1} {}^{n\times n}$. 这样, 我们就有了 $\mathrm{id} \mapsto I_n, 0 \mapsto 0, k\mathrm{id} \mapsto kI_n$, 可逆变换 $\mapsto$ 可逆矩阵A. 在讲解矩阵的时候, 讲解了矩阵的基本方法之后, 我们阐述了可逆矩阵的概念. 在这里, 我们同样也希望找到可逆的线性变换.

在这之前, 我们还是从维数出发. 因为有限相同维数的情况下, 如果一个映射是单射, 那么就蕴含了它是满射. 也就是有如下的定理:

定理 0.4

若 $\in \mathrm{End}V$, $\dim V = n$, 下列条件等价:

- $\exists$, 使得 $== \mathrm{id}$
- 是双射

那么 是可逆的.

接着保持乘法 (定理 0.1.2) 的叙述, 我们试着举出一些例子:

例子 0.4

$V = (\sqrt{2})/$, 考虑 $\mathrm{End}V$, 如果满足 $\{\in \mathrm{End}V \mid () = ()(), \forall, \in V\}$, 那么 在每一点的取值可以确定吗?

例子 0.5

(推广) 若 $\subseteq \mathbb{E}$, 考虑 $\mathrm{End}\mathbb{E}/$, 如果 $G = \{\in \mathrm{End}\mathbb{E}/ \mid () = ()(), \forall, \in V, (1) \neq 0\}$, 探索 G 有哪些好玩性质.

事实上, 这个集合中的元素满足如下的性质:

- $\mathrm{id} \in G$(存在单位元)
- $^{-1} \in G$(存在逆元)
- 若 $, \psi \in G, \circ\psi \in G$.(具有封闭性)

这样的三条性质在以后会经常遇到. 与群这个概念有关.

例子 0.6

继续考虑上述例子, 如果让 $\mathbb{E} := ({}_{1}..{}_{n})$ 是包含 ${}_{1}..{}_{n}$ 的最小数域, 其中 ${}_{1}..{}_{n}$ 是某多项式的根, 有映射 $\varphi : \mathbb{E} \to \mathbb{E}$, 那么集合中保持着乘法的元素就是多项式根的置换.

这样的技巧可以把 $f(x)$ 求根转化为数域上的问题, 并且抽象成群的问题. 最后, 我们来看几个例子:

- $\{1..n\}$ 的排列到 $i_1..i_n$ 是一一对应的双射

- 有线性空间 $V/$, ${}_{1}\cdot\cdot{}_{n}$ 是 V 的一组基, 只要基的重排唯一, 那么这个线性变换也就唯一了. 一共有 $n!$ 个不同的变换. 这和行列式的奇偶置换有关.

线性变换在不同的基下表示的难度不同, 比如线性变换求导 $\frac{d}{dx}$, 如果我们选取如下的几组基底, 看到有不同的矩阵.

$$
\begin{aligned}
1, x, x^2, \cdots, x^{n-1} &\to \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & n-1 \\ & & & 0 \end{pmatrix} \\
1, x, x^2/2!, \cdots, x^n/n! &\to \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \\
1, (x-a), (x-a)^2, \cdots, (x-a)^n &\to ???
\end{aligned}
$$

注意到这些矩阵的形态和复杂程度有显著的差别. 那么一个自然的问题是有没有办法把这些基变换, 使得矩阵简单一点? 也就是

$$
\begin{tikzcd}
{} \arrow[r, "{}_{1}\cdot\cdot{}_{n}"] \arrow[dr] & A \arrow[d, "{}_{1}\cdot\cdot{}_{n}"', "\text{过渡矩阵}T?"] \\
& B
\end{tikzcd}
$$

0.3 相似与特征值

从本节开始, 我们就要试图选取合适的基 ${}_{1}, \cdot\cdot{}_{n}$, 使得矩阵 B“简单”.

什么样的矩阵简单呢? 上下三角确实比较简单, 但是对角矩阵可能更简单, 因为矩阵乘法是很容易做的. 单位矩阵是更加简单的, 但是它的限制条件可能会比较强. 所以我们的目标是尽量把矩阵化为对角的矩阵.

首先来看不同基底下两组基下矩阵的关系. 我们知道 $({}_{1}\cdot\cdot{}_{n}) = ({}_{1}\cdot\cdot{}_{n})A, B({}_{1}\cdot\cdot{}_{n}) = ({}_{1}\cdot\cdot{}_{n})B$, 并且由过渡矩阵的知识我们有 $({}_{1}\cdot\cdot{}_{n}) = ({}_{1}\cdot\cdot{}_{n})T \quad (*)$, 其中 T 中的每一个元素都是一个列向量, 不妨把 T 展开写作 $(T_1..T_n)$.

对于 (*) 式子两边同时施以线性变换 , 有

$$
\begin{aligned}
((_1 \cdots _n)T) =&((_1 \cdots _n)T)B \\
(_1 \cdots _n)T =&((_1 \cdots _n)T)B \\
((_1 \cdots _n)A)T =&((_1 \cdots _n)T)B \\
AT =&TB \\
B =&T^{-1}AT
\end{aligned}
$$

我们称 B 和 A **相似**.

定义 0.2

(相似)

定理 0.5

如果 $\in \mathrm{End}V/, \dim V/ = n$, 则 在不同基下的矩阵是相似的.

这个命题意味着: 如果把 V 中的所有基在某一个线性变换 下构成的全体看做一个集合的话, 这个集合就是彼此相似的. 也就是我们得到了与 A 相似的矩阵的全体. 并且我们也不难验证相似是一个等价关系.

命题 0

相似是一个等价关系.

在本节的开始, 我们希望让基 "简单一些". 最简单的矩阵是对角形矩阵. 那么我们能不能想办法所有的矩阵化为对角形呢? 根据命题 0.3的等价关系, 意味着 $^{n\times n}$ 可以分解为**不同**的相似等价类的并. 我们就可以从单位矩阵出发, 看看和它相似的矩阵都有哪些, 是不是构成所有的矩阵全体, 从而看一看是不是能够化为所有的矩阵全体.

例子 0.7

如果单位变换 在某一组基 $_1\cdots_n$ 下得到的矩阵是 kI_n, 过渡矩阵使得其基底发生变换, 如果变换成了 $_1\cdots_n$ 时候会发生什么?
事实上, 由于相似关系, 我们有 $T^{-1}(kI_n)T = kI_n$, 因为单位矩阵和任何矩阵都是可换的. 这也就说明与 kI_n 相似的矩阵只有它本身. 因此我们只能退而求其次寻求对角矩阵. 这是稍后要讨论的话题. 不过现在我们可以探讨一下矩阵的相似到底有哪些好玩的性质.

命题 0

如果 A 和 B 两个矩阵相似, 那么

- $r(A) = r(B) = r$, 由此可以定义线性变换的秩 $r() = r$.
- $|A| = |B|$, 同样可以定义线性变换的行列式, 定义作 $|| = |A|$.
- $A = B$, 也同样可以定义线性变换的迹, 定义作 $\mathrm{tr}() = \mathrm{tr}(A)$.
- $B^n = T^{-1}A^nT$, $f(B) = T^{-1}f(A)T$.

下面我们就要通过找到合适的一组基底, 让这个变换看上去简单. 如果现在有一个 A, 我们要找到一个 B 使得 B 简单, 我们采用分步走的形式:

- B 可以是数量矩阵吗? 可以, 当 A 是数量矩阵的时候.
- B 可以是对角矩阵吗?
- B 可以是准对角矩阵吗?

对于上述第二个问题, 我们可以做出如下的推演: 若果 B 是对角矩阵, 就说明

$$(_1\cdots_n) = (_1\cdots_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} = (\lambda_1{}_1\cdots\lambda_n{}_n).$$

这样一来, 就要求

$$\begin{cases} {}_1 = & \lambda_{11} \\ {}_2 = & \lambda_{22} \\ \cdots & \cdots \\ {}_n = & \lambda_{nn} \end{cases}$$

同时成立. 也就是, 它的基元素的像恰好是它本身的常数倍, 这样一来才可以找到一组可以对角化的, 方便的矩阵. 于是我们给出特征值的定义:

定义 0.3

(特征值和特征向量)

如何求解特征值呢? 既然我们要求 $={}_0$, 移项, 有 ${}_0- = ({}_0\mathrm{id}-) = 0$. 这就意味着 $\ker({}_0\mathrm{id}-) \neq 0$, 也就是 ${}_0\mathrm{id}-$ 不是单射, 双射, 满射. 也就是这个是不可逆的变换. 不可逆变换的行列式是 0. 这样的话, 我们若取一下行列式的话, 就有 $|{}_0\mathrm{id}-| = 0$. 用矩阵的语言来书写就是 $|{}_0 I_n - A| = 0$.

既然我们得到了一个变换 ${}_0\mathrm{id}-$, 这时候我们来看一下这个变换核空间. 也就是 $\ker({}_0\mathrm{id}-) = \{\in V| ={}_0\}$, 和上面的有一点点小小的区别是 0 也在这个集合里面. 我们把这个集合称为**属于 ${}_0$ 的特征子空间**.

定义 0.4

(特征子空间)

仔细观察这个属于 ${}_0$ 的特征子空间, 稍微把行列式展开一点, 我们就会发现

$$|{}_0\mathrm{id}-| = 0 \iff |{}_0 I_n - A| = \begin{vmatrix} {}_0 - a_{11} & & -a_{1n} \\ & \ddots & \\ -a_{n1} & & {}_0 - a_{nn} \end{vmatrix} = {}^n_0 + a_{n-1}{}^{n-1}_0 + \cdots + a_0,$$

其中 $a_1..a_{n-1}$ 都是确定的常数. 如果我们用动态的眼光来看这个问题, 即把 λ 视作一个变量的话, 就可以记作以 为变量的 n 次首一的多项式. 这个多项式被称作**特征多项式**. $f(\lambda)$ 的根也称为**特征根**.

定义 0.5

(特征多项式)

定义 0.6

(特征矩阵, 特征根)

例子 0.8

(平面上的旋转变换的特征值)

例子 0.9

(求导操作的特征值)

很多时候计算特征值不是一件简单的工作. 因此就需要一些技巧. 在以前学习行列式的时候有一个例题讲述的是如果 $X \in^{n\times 1}, Y \in^{1\times n}$, 那么有 $|I_n + XY| = 1 + YX$. 在这里我们不妨问一问, 这个命题有什么推广?

问题 0.1

如果有 $A \in^{m\times n}, B \in^{n\times m}, AB \in^{m\times m}, |I_m - AB|$ 与 $|I_n - BA|$ 有什么关系? 是相等吗?

先看左边, 进行化简:

$$|I_m - AB| = |(I_m - \frac{1}{}AB)| =^m |I_m - \frac{1}{}AB| =^m |I_n - \frac{1}{}BA| =^{m\ -n}|I_n - BA|.$$

也就是我们知道了这样的一个命题:

命题 0

如果有 $A \in^{m\times n}, B \in^{n\times m},^n |I_m - AB| =^m |I_n - BA|$.

事实上, 特征多项式在矩阵和线性变换的研究中起到了重要的作用. 将上面的结论推广, 我们有:

定理 0.9

(Hamilton-Cayley) 设 $A \in^{n\times n}$ 的特征多项式为

$$f() = |I_n - A| =^n + a_{n-1}^{n-1} + \cdots + a_0$$

则 $f()$ 是 A 的**零化多项式**, 也就是 $f(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_0 I_n = 0$.

先不证明这个结论同样可以导出很多东西, 不妨留到 0.6节来完成.

0.4 可对角化

回顾在定义 0.3中的内容, 我们这次来看一看什么时候可以对角化. 假设 的不同特征值为 $_{1..k}$, 对应的不同的特征子空间有 $V_{_1}()..V_{_k}()$. 如果我们在上述的 k 个特征子空间里面分别寻找一些基底, 并且我们试图把它们拼凑起来, 试图看一看和全空间的关系. 具体的, 我们需要了解:

- 合并之后是线性无关的吗?
- 向量的个数等于维数吗?

我们经常使用 “直和” 的观念来描述两个线性空间之和看上去像是是 “不相交的”. 这时候我们对 “直和” 的定义 (见定义**??**) 进行扩展:

定义 0.7

(多个元素的直和) 设 $V_1..V_n$ 是 V 的子空间, 那么下列条件等价:

- $\forall \in V_1 + V_2 + \cdots + V_n$, 有唯一的分解式 ${}_1 +_2 + \cdots +_n$ (${}_1 \in V_1, {}_2 \in V_2, \cdots, {}_n \in V_n$).
- 零向量的分解式是唯一的
- $\left(\sum_{j\neq i} V_j\right) \cap V_i = \{0\}$
- $\dim V_1 + \dim V_2 + \cdots + \dim V_n = \dim\left(V_1 + V_2 + \cdots + V_n\right)$
- V_1, V_2 的一组基合并形成 $V_1 + V_2$ 的一组基.

若 $V_1, V_2, \cdots, V_n$ 满足上述五条的任何一条, 则称 $V_1 + V_2 + \cdots + V_n$ 是 $V_1 + V_2 + \cdots + V_n$ 的直和, 通常记作

$$V_1 \oplus V_2 \oplus \cdots \oplus V_n.$$

关于线性无关性, 我们发现如下命题:

命题 0

$V_{_1}() + V_{_2}() + .. + V_{_k}()$ 是直和.

证明. 我们考虑证明零向量的分解式唯一.

假若 $0 =_1 + \cdots +_k$, 能不能推出 ${}_1 = \cdots =_k = 0$?

两边同时作用 上, 有

$$0 =_1 + \cdots +_k =_1 {}_1 +_2 {}_2 + \cdots +_n {}_n = ({}_1 .. {}_k) \begin{pmatrix} {}_1 \\ \vdots \\ {}_k \end{pmatrix}$$

一直作用 $k-1$ 次, 有

$$
\begin{aligned}
0 &= {}^2_1\,{}_1 + .. + {}^2_k\,{}_k &&= ({}_1 .. {}_k)\begin{pmatrix} {}^2_1 \\ \vdots \\ {}^2_k \end{pmatrix} \\
0 &= {}^3_1\,{}_1 + .. + {}^3_k\,{}_k &&= ({}_1 .. {}_k)\begin{pmatrix} {}^3_1 \\ \vdots \\ {}^3_k \end{pmatrix} \\
&\vdots && \vdots \\
0 &= {}^{k-1}_1\,{}_1 + .. + {}^{k-1}_k\,{}_k &&= ({}_1 .. {}_k)\begin{pmatrix} {}^{k-1}_1 \\ \vdots \\ {}^{k-1}_k \end{pmatrix}
\end{aligned}
$$

把上述的列合并起来, 就有

$$
(0, 0, \cdots, 0) = ({}_1, {}_2, \cdots, {}_k)\begin{pmatrix} 1 & {}_1 & \cdots & {}^{k-1}_1 \\ 1 & {}_2 & & {}^{k-1}_2 \\ \vdots & & & \vdots \\ 1 & {}_k & \cdots & {}^{k-1}_k \end{pmatrix}
$$

由于后面的是 Vandermonde 行列式, 一定不为 0, $({}_1, {}_2, \cdots, {}_k)$ 只能为 0. 于是零向量的分解式唯一. #

到底什么时候可以对角化呢?

定理 0.11

(可对角化的判别法则) 如果 可对角化, 那么:

- $\exists$ 一组基, 使得 的矩阵为对角形.
- 有 n 个线性无关特征向量.
- $V_1() \oplus V_2() \oplus .. \oplus V_n() = V$.
- $\dim V_1(A) + \cdots + \dim V_n(A) = \dim V$.

下面来举一个在多项式空间里面求导的一个例子.

例子 0.10

考虑 $\frac{d}{dx} \in \mathrm{End}([x]_n)$,

例子 0.11

考虑度算子 $\mathscr{D} : f(x) \mapsto x\frac{df(x)}{dx}$.

首先考虑一个简单的情况, 也就是若 有 n 个互不相同的特征值 (也就是 $\dim V_{\lambda_i}() = 1$), 那么 可对角化.

例子 0.12

求 $A = \begin{pmatrix} 0 & & & & 1 \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & 0 & \\ & & & 1 & 0 \end{pmatrix}$ 的特征值 (数域在 $\mathbb{C}$ 上).

解答. 求出 $|I_n - A| = \begin{vmatrix} & & & & -1 \\ -1 & & & & \\ & -1 & \ddots & & \\ & & \ddots & & \\ & & & -1 & \end{vmatrix} =^n -1$, 因此它有 n 个不相同的特征根, 分别是 $\omega^k = e^{\sqrt{-1}2\pi/n}, k = 0, 1, .., n-1$. 也就是可以找到一个可逆矩阵 T, 使得 $T^{-1}AT = \begin{pmatrix} 1 & & & \\ & \omega & & \\ & & \ddots & \\ & & & \omega^{n-1} \end{pmatrix}$. #

实际上, 例子 0.4展示了一个有趣的事实. 如果把最后一列换为 $a_0, \cdots, a_{n-1}$, 也就是

$$B = \begin{pmatrix} 0 & & & & a_0 \\ \mathbf{1} & \mathbf{0} & & & a_1 \\ & \mathbf{1} & \ddots & & \vdots \\ & & \ddots & \mathbf{0} & a_{n-2} \\ & & & \mathbf{1} & a_{n-1} \end{pmatrix}$$

B 的特征多项式恰好就是 $|I_n - A| =^n +a_{n-1}^{n-1} + \cdots + a_0 = f()$. 这样就让我们在矩阵和多项式之间搭了一个桥梁. 最早发现这样做的是 Frobenius, 因此这个矩阵也称为 Frobenius 矩阵.

同时发现, 这个矩阵 B 也对应着一个线性方程组. 我们也可以问一问这个 $BX = 0$ 的解空间维数是多少. 注意到 $r(B) \geq n-1$, 因为它的左下角有一块的行列式已经不是 0 了 (加粗的部分).

回到矩阵 A. 我们发现 $A^2, A^3, \cdots, A^{n-1}$ 都是可以对角化的. 并且有 $A^n = I_n$, 这样我们就可以把这些组成一个多项式:

$$a_0 I_n + a_1 A + a_2 A^2 + \cdots + a_{n-1} A^{n-1} = f(A)$$

并且 $f(A)$ 也是可以对角化的. 这就意味着

$$\begin{aligned} T^{-1} f(A) T =& T^{-1}(a_0 I_n + a_1 A + \cdots + a_{n-1} A^{n-1}) T \\ =& \begin{pmatrix} a_0 & & \\ & \ddots & \\ & & a_0 \end{pmatrix} + \begin{pmatrix} a_1\omega & & \\ & \ddots & \\ & & a_1\omega^{n-1} \end{pmatrix} + \cdots + \begin{pmatrix} a_{n-1}\omega & & \\ & \ddots & \\ & & a_{n-1}\omega^{n-1} \end{pmatrix} \\ =& \begin{pmatrix} f(1) & & & \\ & f(\omega) & & \\ & & \ddots & \\ & & & f(\omega^{n-1}) \end{pmatrix}. \end{aligned}$$

这就启示我们可以把一个复杂的内容通过拆分成若干个简单的内容来进行拆解. 所以我们得到的 $f(A)$ 的这个矩阵的特征值是 $f(1)..f(\omega^{n-1})$, $|f(A)| = \prod_{i=1}^{n-1} f(\omega^i)$, 同时 $f(A)$ 特征多项式的矩阵也可以得到是 $f(A) = a_0 I_n + a_1 A + \cdots + a_{n-1} A^{n-1}$. 这样, 我们就可以轻松的求解像这样的**循环矩阵**的行

列式了.

$$C_5 = \begin{pmatrix} a_0 & a_4 & a_3 & a_2 & a_1 \\ a_1 & a_0 & a_4 & a_3 & a_2 \\ a_2 & a_1 & a_0 & a_4 & a_3 \\ a_3 & a_2 & a_1 & a_0 & a_4 \\ a_4 & a_3 & a_2 & a_1 & a_0 \end{pmatrix}$$

现在来总结印证一下前面的内容的框架, 并且从一个方面来说明一下定理 0.3在这种情况下的确是对的.

线性变换和对应同阶方阵之间有一一对应, 有时候我们希望最后的方阵简单一点, 于是我们需要研究方阵之间的变换关系. 有时候可以把最后的方阵化为对角形, 就如下图中的 B 一样

$$\begin{tikzcd} {} \arrow[r, "{}_1\cdots{}_n"] \arrow[dr, "{}_1\cdots{}_n"'] & A \arrow[d, "\text{过渡矩阵}T?"] \\ & B = \mathrm{diag}({}_1\cdots{}_n) = T^{-1}AT \end{tikzcd}$$

这时候这个方阵的特征值就是 ${}_i(1 \leq i \leq n)$, 特征多项式可以表示为

$$(-{}_1)(-{}_2)\cdots(-{}_n).$$

而这就暗示了 id$-$ 可以被表示为

$$({}_1I_n - A)({}_2I_n - A)\cdots({}_nI_n - A) = 0.$$

因为 A, B 相似, 有

$$({}_1I_n - B)({}_2I_n - B)\cdots({}_nI_n - B) = 0$$

$$\begin{pmatrix} 0 & & & \\ & \lambda_2 - {}_1 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix} \cdots \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} = 0$$

从而可以把复杂的问题简单化.
当 A 可对角化的时候, $f() = |I_n - A|$, 则 $f(A) = 0$. 也就是 A 的特征多项式是 A 的**零化多项式**.

接下来来探讨有重根的情形. 假如说有一个矩阵 B 的特征多项式如下:

$$f() = |I_n - B| = (-{}_1)^{i_1}(-{}_2)^{i_2}\cdots(-{}_k)^{i_k}$$

那么它对应的对角形矩阵可以是 $\mathrm{diag}(\overbrace{{}_1, \cdots, {}_1}^{i_1\text{个}}, \overbrace{{}_2\cdots{}_2}^{i_2\text{个}}, \cdots, \overbrace{{}_k\cdots{}_k}^{i_k\text{个}})$.

前面我们看到, 对于矩阵 B 的特征多项式而言, 有 $f() = 0$. 那么还有没有别的多项式, 使得 $f() = 0$ 呢? 其实 $(-{}_1)(-{}_2)\cdots(-{}_k) := m()$ 已经可

以保证 $m() = 0$ 了. 像这样, 首一的次数最小的多项式被称为**最小多项式**, 并且其一定是特征多项式的因子.

定义 0.8

(最小多项式)

从前面开始, 我们一直在强调要找到一个多项式, 使得在把这个多项式的变量赋值为这个线性变换后, 这个多项式等于 0, 还给它起了一个名字**零化多项式**, 这样做有什么意义?

0.5 零化多项式不变子空间

前情提要:
前面的内容我们发现了线性变换和方阵的一一对应关系, 并且随着基的不同, 矩阵表示出来会有所不同. 所以我们希望找到一组好的基底, 使得表示出来的矩阵更加简单. 也就是不同基底下, 线性变换的矩阵满足相似关系. 我们可以用下图的内容展示之. 其中不要忘记线性变换的定义: $({}_1 \cdot\cdot {}_n) = ({}_1 \cdot\cdot {}_n)A$.

$$
\begin{tikzcd}
\mathrm{End}V/ \arrow[r, "{}_1\cdot\cdot{}_n"] & {}^{n\times n} \\
\uparrow\in & \uparrow\in \\
\arrow[r] \arrow[dr, "{}_1\cdot\cdot{}_n"'] & A \\
 & B = T^{-1}AT
\end{tikzcd}
$$

如果我们找的 B 恰好为对角形, 也就是设 $B = \mathrm{diag}({}_1 \cdot\cdot {}_n)$, 那么就必须满足 ${}_i ={}_i {}_i$. 经过变换我们知道 $({}_i\mathrm{id} -)_i = 0$, 意味着 ${}_i \in \ker({}_i\mathrm{id} -)$, 对应到矩阵的表示形式就有 $|{}_i\mathrm{id} - A| = 0$ 这个行列式等于 0 的条件. 此时, ${}_i$ 就叫**特征值**, ${}_i$ 就称为**对应于特征值**的**特征向量**.
为了强调动态的观念, 我们引入**特征多项式**的概念, 也就是把 $({}_i\mathrm{id} -)_i = 0$ 中的 ${}_i$ 这样一个确定的数换做了一个变量 . 也就是定义 $f() = |\mathrm{id} - | = |I_n - A|$ 为这个线性变换的**特征多项式**.
然后我们发现让特征多项式等于 0 的值很不一般, 需要单独分开去考虑. 于是我们定义了**特征子空间**, 也就是如果 $f({}_0) = 0, V_{{}_0}() = \ker({}_0\mathrm{id} -) = \{\in V| ={}_0\}$.

我们发现, 特征多项式有很多神奇的用途. 比如如果记特征多项式是

$$f() = |I_n - A| = \prod_{i=1}^{n}(-{}_i)$$

注意到如果 A 是对角形, 那么根据 0.4节的内容, 我们就知道了 $f(A) = 0$. 并且如果 B 相似与某个对角形, 根据命题 0.3就可以知道 $f(B) = 0$. 实际上, 上面的两个条件都太强了, 对于任意的 A, 如果 $f() = |I_n - A|$, $f(A) = 0$

都是对的 (现在先不给出证明, 因为需要用到定理 0.3, 我们还没有证明). 这样的东西性质非常好, 这样子我们就感到定义**零化多项式**是十分有必要的. 下面, 我们给出一个定义:

定义 0.9

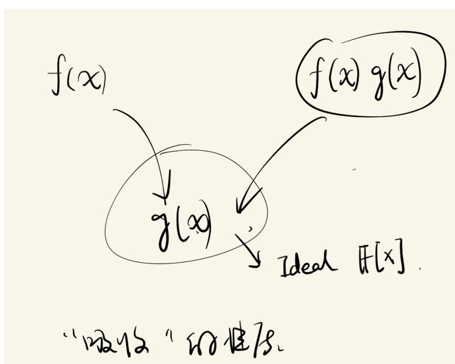
如果 $A \in^{n\times n}, f() \in []$, 称 $f()$ 为**零化多项式**, 如果 $f(A) = 0$.

有了这样一个定义, 我们可以看一看能够得到什么.

- (1) **(存在性)** 存在次数 $\leq n^2$ 的非零多项式 $f()$, 使得 $f(A) = 0$.

  我们有好几种观点来看. 比如 $I_n, A, \cdots, A^{n^2}$ 是线性相关的; $[] \to^{n\times n}$: $f() \to f(A)$ **不是单射**都说明了这个问题.

- (2) 考虑零化多项式的全体 $\ker\varphi = \{g() \in [] | g(A) = 0\}$. 不难发现这样集合里面的元素保持加法和数乘. 更一般的, 我们发现它**保持乘法**: 即 $g(x) \in \ker$, $h(x) \in [], g()h() \in \ker$. 这个集合具有吸收任何东西的感觉. 由于其性质比较理想, 在抽象代数里面, 这个集合叫做 $f[]$ 的**理想**.

- (3) **(存在唯一最小多项式)** 存在唯一的**次数最小的首一多项式** $m_A() \in \ker$, 满足 $m_A() | g()$, 对于 $\forall g() \in \ker$. 也就是说核空间 $\ker = \{m_A()h() | h() \in []\}$. 如果 Hamilton-Cayley 定理 (定理 0.3,A 的特征多项式包含在核里面, 即 $|I_n - A| \in \ker$) 说的是正确的话, 那么上述内容就可以证明了 (至少目前来看可对角化是这样的).

从可对角化出发, 我们来研究一下可对角化情形的最小多项式. 如果我们把对角矩阵的一组基打乱一下, 注意到得到的新的矩阵也是对角矩阵, 因为打乱基的顺序并不影响特征多项式的

$$\begin{tikzcd} \text{可对角化} \arrow[r, "1\cdots n"] \arrow[dr, "{i_1, i_2, \cdots, i_n}", "\text{打乱顺序}"'] & A \\ & B\text{也是对角矩阵} \end{tikzcd} \qquad = \mathrm{diag}(\overbrace{1, .., 1}^{n_1\text{个}}, \overbrace{2..2}^{n_2\text{个}}, \cdots, \overbrace{m..m}^{n_m\text{个}}) = \mathrm{diag}({}_1 I_{n_1}, {}_2 I_{n_2}, \cdots, {}_m I_{n_m})$$

我们知道了 $f() = |I_n - A| = (-_1)^{n_1}(-_2)^{n_2} \cdots (-_k)^{n_k}$, $f(A) = 0, \forall A$. 那么我们考虑零化多项式之一可以是把它们的重数都去掉得到 $(-_1)(-_2)\cdots(-_k)$. 所以我们可以在可对角化的时候得到一个结论: 至少存在一个零化多项式, 使得其没有重根.

$\boxed{\text{中间结论 1: 若 } A \text{ 可对角化, 则存在一个 } A \text{ 的零化多项式 (非零)} g() \text{ 使得 } g() \text{ 无重根.}}$

所以根据最小多项式的定义我们知道 $m_A()|g()$. 那么他们相等吗? 假如说有一个根不在最小多项式里面, 即 $m_A() = (-_{i_1})..(-_{i_l})(\text{当} l < k) \implies m_A() = (A -_{i_1} I_n)..(A -_{i_l} I_n) \neq 0$.

$\boxed{\text{中间结论 2: 若 } A \text{ 可对角化, 那么 } m_A() \text{ 无重根.}}$

那么一个关键问题是如果 $m_A()$ 无重根, 可否推出 A 可对角化 (可对角化)? 如果这个结论成立的话, 就有

定理 0.12

以下条件等价:

- A 可对角化 (可对角化)
- 存在非零的零化多项式 $g()$, 使得 $g()$ **不相同的**一次因式的乘积.
- $m_A()$ 可以分解为**不相同的**一次因式的乘积.

在这里为了保证所有的根都可以去到, 无关数域, 修正为分解为**不相同的**一次因式的乘积. 假设 $m_A() = (-_1)..(-_k)$, 希望得到 $V_1() \oplus .. \oplus V_k() = V$. 这里的一个小问题是我们并不知道 $_i$ 是不是特征值, 但是问题不大, 毕竟直和上 0 也是对结果没有太大影响的.

我们试图把线性变换带入这个零化多项式, 有

$$m_A() = (-_1\mathrm{id})..(-_n\mathrm{id}) = 0$$

然后因为 $V_1 = \ker(-_1\mathrm{id})$, 于是就有:$\ker(-_1\mathrm{id}) \oplus .. \oplus \ker(-_n\mathrm{id})$ 是直和, 我们需要证明这是 $\ker 0$.

如何运用多项式证明这个问题? 假设有 $g()$, 使得 $g() = 0$, 那么先分解太多, 先考虑分解成两项. 假设 $g() = h()k()$, 这时候带入线性变换 就有了 $h()k() = 0$. 我们来关注 $\ker h(), \ker k()$, 注意二者加起来是不是直和, 并且观察相加之后的空间是不是 V. 反映到上面的情形, 就是 $\ker(-_1\mathrm{id}) \oplus \ker((-_2\mathrm{id})..(-_n\mathrm{id})$ 直和加起来是不是等于 V. 接下来一直拆, 可能就有很好的效果.

定理 0.13

如果 $g() \in [], g() = h()k(), \gcd(h(), k()) = 1, \in \mathrm{End}V, \ker g() = \ker h() \oplus \ker k()$.

例子 0.13

(求导的特征多项式是 $^n = \cdot^{n-1} =^k \cdot^{n-k}$)

证明. **是直和.** 如果有 $\in \ker h() \cap \ker k()$, 已知 $h() = 0, k() = 0$. 由于两个多项式**互素**, 即 $\gcd(h(), k()) = 1$, 那么由于 Bezout 定理, 有

$$u()h() + v()k() = 1,$$

带入 , 有

$$u()h() + v()k() = \mathrm{id}, \tag{(*)}$$

作用在 上, 有

$$u()h() + v()k() = \mathrm{id},$$

也就是 $0 + 0 == 0$.

与空间 V 相等. 注意到 $\ker g() \supseteq \ker h() \oplus \ker k()$, 只需要证明反包含关系即可. 也就是 $\ker g() \subseteq \ker h() \oplus \ker k()$. 这时候取 $\in \ker g()$, 由于直和知道 $g() = h()k() = 0$, 与 (*) 式对比, 有 $k()u()h() = 0$, 意味着 $_2 \in \ker k()$. 同理 $_1 \in \ker h()$, 可以证明是相等的. #

这是一个十分强大的定理. 正是因为上面的定理我们才有:

$$
\begin{aligned}
V &= \ker 0 = \ker(-_1\mathrm{id}) \oplus \ker((-_2\mathrm{id})..(-_n\mathrm{id})) \\
&= (-_1\mathrm{id}) \oplus (-_2\mathrm{id}) \oplus .. \oplus (-_n\mathrm{id})
\end{aligned}
$$

也就是 V 是特征子空间的直和, V 可对角化.

例子 0.14

给出 $A^2 = I_n, A^2 = A, A^3 = I_n$ 的条件意味着 $^2 - 1, ^2 -, ^3 -$ 是他们的零化多项式. 但是不一定是最小多项式.

问题 0.2

最小多项式的根是特征值吗? 特征值一定是最小多项式的根吗?

问题 0.3

如果有 $\in \mathrm{End}V, A \in^{n\times n}$, 如果 $f()$ 是特征多项式, $m()$ 是最小多项式, $p_i()$ 是首一不可约多项式, 并且 $f() = p_1()^{k_1} p_2()^{k_2} \cdots p_l()^{k_l}$, 那么 $m()$ 的因式分解形如什么样的? q_i 与 p_i 的关系是什么? k_i 与 k_j 的关系是什么?

那么, 不可对角化的情形是怎么样的呢? 回到定理 0.5, 如果不是一次因式的乘积, 如果 $g() = p_1()^{k_1} p_2()^{k_2} \cdots p_l()^{k_l}$ 是 的零化多项式, p_i 不同的首一不可约, 两两互素多项式, 就可以知道 $V = \ker g() = \ker p_1()^{k_1} \oplus \ker(p_2()^{k_2}..p_n()^{k_n})$, 因为中国剩余定理, 总是可以继续分解下去变成形如

$$
V = \ker p_1()^{k_1} \oplus p_2()^{k_2} \oplus \cdots \oplus p_n()^{k_n}.
$$

那么我们如果有在 $\ker p_1()^{k_1}$ 里面取一组基 $_{11}..{}_{1i_1}$, $\ker p_2()^{k_2}$ 里面取一组基 $_{21}..{}_{2i_2}$,...,$\ker p_n()^{k_n}$ 里面取一组基 $_{n1}..{}_{ni_n}$, 由于是直和, 把它们组合起来就是全空间的一组基.

那么, 一个线性变换在这个基下的矩阵长什么样子呢?

注意到 $_{11} \in \ker p_1()^{k_1}$, $p_1()^{k_1}_{11} = p_1()^{k_1}_{n} = 0$, 也就是它的坐标只需要前面的 i_1 个就行了. 与之类似的, 我们就有

$$\begin{pmatrix} * & * & & & \\ * & * & & & \\ & & * & * & \\ & & * & * & \\ & & & & * \end{pmatrix}$$

这是一个准对角的矩阵, 第一个块长度为 i_1, 第二个块长度为 $i_2, \cdots$, 第 n 个块长度为 i_n. 于是我们知道了, 如果 不可对角化, 那么就把 准对角化. 也就是希望写成

$$\begin{pmatrix} *_{k\times k} & \\ & *_{(n-k)\times(n-k)} \end{pmatrix}$$

的样子. 这就表明了:

$$\begin{cases} {}_1 \in & L({}_1 \cdot\cdot {}_k) := W_1 \\ \vdots & \\ {}_k \in & L({}_1 \cdot\cdot {}_k) := W_1 \\ {}_{k+1} \in & L({}_{k+1} \cdot\cdot {}_n) := W_2 \\ \vdots & \\ {}_n \in & L({}_{k+1} \cdot\cdot {}_n) := W_2 \end{cases}$$

于是我们引入定义:

定义 0.10

设 W 为 V 的子空间, 若对于 $\forall \in W$, 有 $\in W$, 则称 W 为 的**不变子空间**, 记作 $-$ 子空间.

这就意味着取 W 的一组基 ${}_1 \cdot\cdot {}_k \cdot\cdot {}_n$, 那么 在该组基下表示为: $\begin{pmatrix} A_1 & * \\ 0 & * \end{pmatrix}$. 可以把它表示为准上三角的.

0.6 根子空间分解

在上面我们知道了对于空间的第一分解. 对于定义 0.5, 我们还有如下的两种方法来看:

- 限制变换 $|_W, W \to W, \to, |_W \in \mathrm{End}W$.
- $W = L(_1)$ 是 , 其中 $_1 =_1 {}_1$. 将 $_1$ 扩充为 V 的一组基 $_1 .. _n$, 那么 A 的矩阵就形如 $\begin{pmatrix} _1 & \beta \\ \mathbf{0} & \mathbf{A_1} \end{pmatrix}$, 也就是在复数域上, 它可以相似于一个上三角的分块矩阵.

接下来, A_1 又可以经历同样的操作, 让 A_1 更加的小. 也就是 A_1 可以做相似变换使得 $T^{-1}AT = \begin{pmatrix} _2 & * \\ \mathbf{0} & * \end{pmatrix}$. 这样一直做下去, 就可以得到 $T^{-1}AT = D = \begin{pmatrix} _1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & _n \end{pmatrix}$ 的矩阵. 这时候二者的特征多项式是一样的, 也就是

$$f() = |I_n - A| = |I_n - D| = \prod_{i=1}^{n}(-_i).$$

现在分别将 A 与 D 带入, 注意到

$$f(A) = f(TDT^{-1}) = Tf(D)T^{-1}$$
$$f(D) = (D -_1 I_n) \cdots (D -_n I_n)$$
$$= \begin{pmatrix} 0 & * & * & * \\ 0 & _2 -_1 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & _n -_1 \end{pmatrix} \cdot \begin{pmatrix} _1 -_2 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & _n -_2 \end{pmatrix} \cdot \cdots \cdot \begin{pmatrix} _1 -_n & * & * & * \\ 0 & _2 -_n & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

如果它是对角矩阵, 那么就肯定是 0 了. 但是现在是准上三角矩阵, 并且对角线上有一个 0, 那么这个答案是多少? 发现每多一个对角线上就多一个 0. 所以答案是 0. 即

$$f(A) = f(D) = 0.$$

这就是 Hamilton-Cayley 定理 (定理 0.3). 但是刚刚的情况看上去是在复数域上做的, 但是多项式的结果是无关数域的. 在一般数域上不好处理的时候, 都可以放在 $\mathbb{C}$ 上.

当然我们可以在一般数域上证明. 下面给出证明: **证明.** 我们希望把 A 和 $|A|$ 联系起来, 自然想到 A 的伴随矩阵. 由于 $AA^* = |A|I_n$, 于是有

$$(I_n - A)(I_n - A)^* = |I_n - A|I_n = f()I_n$$

并且

$$\begin{pmatrix} -a_{11} & \cdots & \cdots & -a_{1n} \\ \vdots & -a_{22} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ -a_{n1} & \cdots & \cdots & -a_{nn} \end{pmatrix}^* = (A_{ji}) = \begin{pmatrix} b_{11}() & \cdots & b_{1n}() \\ \vdots & & \vdots \\ b_{n1}() & \cdots & b_{nn}() \end{pmatrix},$$

其中 A_{ji} 是一个多项式, 且 $\deg(A_{ji}) \leq n - 1$. 接下来我们把上述式子最右侧按照 的次数拆成若干个小矩阵加起来的和的形式, 也就是:

$$\begin{pmatrix} b_{11}() & \cdots & b_{1n}() \\ \vdots & & \vdots \\ b_{n1}() & \cdots & b_{nn}() \end{pmatrix} =^{n-1} B_{n-1} +^{n-2} B_{n-2} + \cdots + B_1 + B_0.$$

根据条件, $(I_n - A)(I_n - A)^* = f()I_n$, 有

$$(I_n - A)\left(^{n-1}B_{n-1} +^{n-2} B_{n-2} + \cdots + B_1 + TB_0\right) = f()I_n,$$

整理该式子, 展开左右两边并对照系数, 就有

$$\begin{aligned} &^{n}B_{n-1} \quad &&+^{n-1}(B_{n-2} - AB_{n-1}) + \cdots \quad &&+(B_0 - AB_1) - AB_0 \\ &=^{n}I_n &&+a_{n-1}^{n-1}I_n + \cdots &&+a_1I_n + a_0I_n \end{aligned}$$

也就是:

$$\begin{cases} B_{n-1} & = I_n \\ B_{n-2} - AB_{n-1} & = a_{n-1}I_n \\ \vdots & \\ B_1 - AB_2 & = a_2I_n \\ B_0 - AB_1 & = a_1I_n \\ -AB_0 & = a_0I_n \end{cases},$$

逐步回带进去, 我们就有

$$\begin{aligned} B_0 &= a_1 I_n + AB_1 = \cdots \\ &= a_1 I_n + A(a_2 I_n + B_2 I_n) \\ &= \cdots \\ &= \sum_{i=1}^{n-1} A^{i-1} a_i I_n \end{aligned}$$

看最后一个式子就有

$$0 = A^n + \cdots + a_2 A^2 + a_1 A + a_0 I_n = f(A).$$

#

也就是前面的 $m_A()$ 是最小多项式, $f()$ 是特征多项式, 可以知道最小多项式是特征多项式的因子. 即 $m_A() \mid f()$.

问题 0.4

如果 $m_A() = e, f() = \prod_{j=1}^{n} p_j()^{k_j}$, 可以知道 $1 \leq s_i \leq k_i (i \in [1..n])$.

现在我们就得到了一般性结论: 如果 $\in \mathrm{End}V, f() = |\mathrm{id} - | = \prod_{i=1}^{n} p_i()^{s_1}$, 并且 $f() = 0$, 我们就可以推出:

$$f() = p_1\left(\right)^{k_1} \cdots p_l\left(\right)^{k_l}$$

其中 p_i 两两是不同的首一不可约多项式 (**互素**). 这时候使用定理 0.5就可以告诉我们:

$$\ker f() = \ker p_1\left(\right)^{k_1} \oplus \cdots \oplus \ker p_l\left(\right)^{k_l} = V.$$

也就是把 V 分解成了若干个 的直和. 这样选择一组基, 可以让矩阵长成形如 $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_l \end{pmatrix}$ 准对角块的样子. 在复数域上, 这个准对角块可以更方便的研究它. 如令 =, 设 $p_i()$ 都是若干一次因式的乘积. (当然在一般数域上考虑 $f()$ 可以分解成一次因式的乘积的内容), 那么证明内容都是完全

可以平移过去. 这样做的话, 如果记 $W_i = \ker\,(-_i\mathrm{id})^{k_i}$, 就可以做到:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_l.$$

一个线性变换的 ker 和一个线性变换某次方的 ker 的关系是什么? 在定义 0.5下方的探讨中, 我们知道

$$\cdots = \ker^{k+2} = \ker^{k+1} = \ker^{k} \supsetneqq \ker^{k-1} \supsetneqq \cdots \supsetneqq \ker = -\mathrm{id}.$$

也就是 W_i(根子空间) $\supseteq \ker\,(-_i\mathrm{id})$ (特征子空间). 一般记作 $W_i := R_{_i}$. 这样就得到了可对角化的另一个判别条件:

命题 0

V 可分解为 的特征子空间的直和 $\iff$ 可对角化.

并且 始终可以化为根子空间的直和. 如果考虑定义 $_i = |_{R_i()} \in \mathrm{End}(R_i())$. 根据根子空间的定义, $(-_i\mathrm{id})^{k_i}\ |_{R_i()}= 0$. 即, $\left(|_{R_i()} -\mathrm{id}\right)^{k_i} = 0$. 为了方便起见, 不妨先把带有特征值的内容减去, 也就是这样定义: $_i =|_{R_i()} -\mathrm{id}$, 并且 $^{k_i}_i = 0$. 也就是 $B_i \in \mathrm{End}(R_i())$ 也成立. 同时说明了 $_i$ 有一个零化多项式, 是 ^{k_i}. 这时候就可以知道 $_i$ 有一个最小多项式 $^{l_i}, l_i \leq k_i$. $_i$ 的特征多项式是 ^{n_i}. 那么 的所有特征值只能是 $_i$ 吗?

考虑特征值的定义 $_i =,^{k_i}_i =^{k_i}= 0 \implies = 0$, 因此特征值只能是 $_i$. 因此, $_i$ 的特征多项式是 $(-_i)^{n_i}$.

最后, 我们考虑 $|I_n - A| = (-_1)^{k_1}\cdots(-_l)^{k_l}, |\mathrm{id} - | = (-_1)^{n_1}\cdots(-_l)^{n_l}$, 对比系数发现 $n_i = k_i$. 也就是 $n_i = \dim R_i() = k_i$. 于是我们有

命题 0

如果 $f() = |\mathrm{id} - | = (-_1)^{k_1}\cdots(-_l)^{k_l}$, 根子空间的维数是 $\dim R_i() = k_i \geq \dim V_i()$. 其中, $\dim V_i()$ 叫做几何重数, k_i 叫做代数重数, 并且多项式代数重数大于几何重数.

0.7 Jordan 标准形

定义 0.11

(幂零) 设 V 是 上的线性空间, 称 $\in \mathrm{End}V$ 为幂零线性变换, 如果存在正整数 m, 使得 $^m = 0$.

命题 0

设 V 是 上的线性空间, $\in \mathrm{End}V$ 的特征多项式为 $f()$, 那么下列结论等价:

- 是幂零线性变换;
- $f() =^n$;
- $f()$ 在复数域上只有零根.

定理 0.17

设 V 是 上的线性空间, $\in \mathrm{End}V$ 是幂零线性变换, 则 V 可以分解为 的循环子空间的直和. 并且在同构意义下, 分解是唯一的. 也就是如果 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m = W_1 \oplus W_2 \oplus \cdots \oplus W_t$ 为 的循环子空间分解, 则 $m = t$ 且重排 $W_1..W_m$ 的顺序后存在可逆线性变换 $\in \mathrm{End}V$, 使得 $(V_i) = W_i (1 \leq i \leq m)$.

定义 0.12

(Jordan 块和 Jordan 矩阵) 设 $_0 \in$, 称 n 阶方阵

$$J(_0, n) = \begin{pmatrix} _0 & & & \\ 1 & _0 & & \\ & \ddots & \ddots & \\ & & 1 & _0 \end{pmatrix}$$

为 Jordan 块. 设 $_1, _2, \cdots, _m \in$, 称 n 阶方阵

$$J = \mathrm{diag}(J(_1, n_1), J(_2, n_2), \cdots, J(_m, n_m)), \sum_{i=1}^{m} n_i = n$$

为 Jordan 矩阵.

定理 0.18

设 $\in \mathrm{End}V$ 幂零, 则 在某组基下的矩阵为 Jordan 矩阵

$$\mathrm{diag} J(0, n_1), J(0, n_2), \cdots, J(0, n_m), n_1 \geq n_2 \geq \cdots \geq n_m \geq 1.$$

若 $_0 - \mathrm{id}$ 幂零, 则 在某组基下的矩阵为 Jordan 矩阵

$$\mathrm{diag} J(_0, n_1), J(_0, n_2), \cdots, J(_0, n_m), n_1 \geq n_2 \geq \cdots \geq n_m \geq 1.$$

定理 0.19

任意 $\in \mathbb{C}$ 相似于一个 n 阶 Jordan 矩阵, 这个 Jordan 矩阵称为 A 在相似下的标准形, 如果不考虑 Jordan 块的排列次序, A 的 Jordan 标准形唯一.

0.8 多项式矩阵

定义 0.13

设 是数域, 是一个文字, 以 [] 中的多项式为元素的矩阵被称作多项式矩阵或 $-$矩阵, 以 $[]^{m\times n}$ 表示 $m \times n$ 的 $-$矩阵 的全体.

定义 0.14

设 $A() \in []^{n\times n}$, 若存在 $B() \in []^{n\times n}$ 使得 $A()B() = I_n$, 则称 $A()$ 可逆, $B()$ 称为 $A()$ 的逆矩阵.

定理 0.20

$A() \in^{n\times n}$ 可逆当且仅当 $|A()| = d \neq 0$, 此时 $A()$ 有唯一的逆矩阵

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

命题 0

($-$矩阵 的初等变换) 可以对 $-$矩阵 做如下的初等变换:

- 将 $A()$ 的两行 (列) 互换;
- 将 $A()$ 的某两行 (列) 乘以非零常数;
- 将 $A()$ 的某行 (列) 加上另一行 (列) 的 $\varphi()$ 倍, 其中 $() \in []$.

定义 0.15

设 $A(), B() \in []^{m\times n}$, 称 $A()$ 与 $B()$ 相抵 (等价), 如果经过一系列初等行列变换可将 $A()$ 化为 $B()$, 或者说存在 m 阶初等 $-$矩阵:$P_1..P_s$ 以及 $Q_1..Q_t$, 使得 $B() = P_1P_2\cdots P_sA()Q_1Q_2\cdots Q_t$.

引理 0.22

设 $A() \in []^{m\times n}$ 非零, 则存在首一多项式 $b_1(), .., b_r()$ 使得 $A()$ 与 $B() = b_1(), .., b_r(), 0, .., 0 \in^{m\times n}$ 相抵.

引理 0.23

设 $f(), g() \in []$ 为非零多项式, $m() = \gcd(f(), g()), M() = \operatorname{lcm}(f(), g())$, 那么 $\begin{pmatrix} f() & \\ & g() \end{pmatrix}$ 与 $\begin{pmatrix} m() & \\ & M() \end{pmatrix}$ 相抵.

定理 0.24

设 $A() \in []^{m\times n}$ 非零, 则 A 相抵与 $D() = d_1(), .., d_r(), 0, .., 0$, 其中 $d_i()(1 \le i \le r)$ 为首一多项式, 并且 $d_i() \mid d_{i+1}()$.

定义 0.16

称 $D()$ 为 $A()$ 在相抵下的**标准形**, $d_1(), .., d_r()$ 为 $A()$ 的不变因子, 设 $d_i()$ 的标准分解为 $d_i() = p_1()^{k_{i1}} p_2()^{k_{i2}} \cdots p_l()^{k_{il}}, 1 \le i \le n, k_{ij} \ge 0$. 如果 $k_{ij} \ge 1$, 称 $p_j()^{k_{ij}}$ 为 $A()$ 的初等因子.

命题 0

设 $A()$ 与 $B()$ 相抵, 则 $A()$ 与 $B()$ 中的元素分别为彼此的多项式组合.

定义 0.17

设 $A() \in []^{m\times n}$, 记 $D_k(A())$ 为 $A()$ 的所有 k 阶子式的最大公因式, 称为 $A()$ 的 k 阶**行列式因子**, 其中 $1 \le k \le \min(m, n)$, 在不引起混淆的情形简单记作 $D_k()$.

引理 0.26

Jordan 块 $J(_0, n)$ 的特征矩阵的行列式因子为 $1, 1, \cdots, (-_0)^n$.

引理 0.27

初等变换不改变 $-$矩阵 的行列式因子.

定理 0.28

设 $A() \in []^{m\times n}$,$D_1(), D_2() \cdots, D_r()$ 为 $A()$ 的所有非零行列式因子, 则 $A()$ 的不变因子为:

$$D_1(), \frac{D_2()}{D_1()}, \cdots, \frac{D_r()}{D_{r-1}()}.$$

从而标准形是唯一的.

定理 0.29

$A() \in []^{n\times n}$ 可逆当且仅当它是初等 $-$矩阵 的乘积.

定理 0.30

设 $A(), B() \in []^{m\times n}$ 可逆, 那么下列条件等价:

1. $A()$ 与 $B()$ 相抵;
2. 存在可逆的 m 阶方阵 $P()$, n 阶方阵 $Q()$, 使得 $B() = P()A()Q()$;
3. $A()$ 与 $B()$ 有相同的行列式因子;
4. $A()$ 与 $B()$ 有相同的不变因子;
5. $A()$ 与 $B()$ 有相同的初等因子且 $r(A()) = r(B())$.

定义 $A()$ 的不变因子的个数定义为 $A()$ 的秩, 记作 $r(A())$. 注意 $r(A()) = r(B())$ 并不意味 A, B 相抵.

引理 0.31

设 $A \in^{n\times n}, U(), V() \in []^{n\times n}$, 那么存在 $Q(), R() \in []^{n\times n}, U_0, V_0 \in^{n\times n}$ 使得下列等式成立:

$$
\begin{aligned}
U() =&(I_n - A)Q() + U_0;\\
V() =&R()(I_n - A) + V_0.
\end{aligned}
$$

定理 0.32

$A, B \in^{n\times n}$ 相似, 当且仅当 $I_n - A$ 与 $I_n - B$ 相抵.

0.9 线性代数中出现的内容

0.9.1 矩阵的相似

1. 相似矩阵

   (a) Def. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶矩阵, 若存在可逆矩阵 P, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的相似矩阵, 或者说 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 称对 $\boldsymbol{A}$ 进行了相似变换.$\boldsymbol{P}$ 称为把 $\boldsymbol{A}$ 变成 $\boldsymbol{B}$ 的相似变换矩阵.

2. 相似矩阵的性质

   (a) Th. 若 n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值也相同.

   (b) Coll.

      i. 若 n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵 $\lambda_1..\lambda_n$ 相似, 那么 $\lambda_1..\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值

      ii. 若存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\lambda_1..\lambda_n=\boldsymbol{D}$, $\varphi(x)$ 是 m 次多项式, 那么 $\boldsymbol{A}^k=\boldsymbol{P}^{-1}\boldsymbol{D}^k\boldsymbol{P},\varphi(\boldsymbol{A})=\boldsymbol{P}^{-1}\varphi(\boldsymbol{A})\boldsymbol{P}$

3. 矩阵可对角化的条件:

   (a) Th. n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充分必要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

   (b) Coll. 如果 n 阶矩阵的 n 个特征值互不相等, $\boldsymbol{A}$ 与对角矩阵相似.

0.9.2 方阵的特征值与特征向量

1. 特征值. 特征方程. 特征多项式

   (a) Def: 设 $\boldsymbol{A}$ 是 $\boldsymbol{n}$ 阶方阵，如果存在数 λ 和非零 n 维列向量 $\boldsymbol{x}$，使得 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$ 成立，则称 λ 是 $\boldsymbol{A}$ 的一个**特征值**, 列向量 x 称为**属于特征值** λ 的**特征向量**. 称 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=0$ 是 A 的**特征方程**. $f(\lambda)=|\boldsymbol{A}-\lambda\boldsymbol{E}|$ 是 A 的**特征多项式**.

   (b) Note.

      i. 特征值问题是对方阵而言的.

ii. 矩阵 $\boldsymbol{A}$ 的特征值就是 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=0$ 的根. 齐次线性方程组 $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{x}=\mathbf{0}$ 的非零解就是矩阵 $\boldsymbol{A}$ 对于特征值的特征向量.

2. 常见性质

设 n 阶矩阵的特征值为 $\lambda_1\cdots\lambda_n$, 于是有

(a) $\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(\boldsymbol{A})$

(b) $\lambda_1\cdots\lambda_n=|\boldsymbol{A}|$.

3. 特征值与特征向量的性质

(a) 若 x_1,x_2 都是矩阵 $\boldsymbol{A}$ 对应于特征值 λ_0 的特征向量, 则 $k_1x_1+k_2x_2$ 也是 $\boldsymbol{A}$ 的对应特征值 λ_0 的特征向量. 其中 $k_{1,}k_2$ 是任意常数, $k_1x_2+k_2x_2\neq 0$. 所以一个特征值有无数个特征向量与之对应, 一个特征向量只属于一个特征值.

(b) 若 λ 是矩阵 A 的特征值, x 是 $\boldsymbol{A}$ 属于特征值 λ 的特征向量, 那么

i. $k\lambda$ 是 $k\boldsymbol{A}$ 的特征值 (k 是任意常数)

ii. λ^m 是 $\boldsymbol{A}^m$ 的特征值. (m 是任意常数)

iii. 当 $\boldsymbol{A}$ 可逆的时候,λ^{-1} 是 $\boldsymbol{A}^{-1}$ 的特征值.

iv. $\varphi(\lambda)$ 是 $\varphi(\boldsymbol{A})$ 的特征值, 其中 $\varphi(\lambda)=a_0+a_1\lambda+\cdots+a_m\lambda^m,\varphi(\boldsymbol{A})=a_0\boldsymbol{E}+a_1\boldsymbol{A}+\cdots+a_m\boldsymbol{A}^m$ 并且 $\varphi(\boldsymbol{A})$ 是关于矩阵 $\boldsymbol{A}$ 的矩阵多项式.

v. 设 $\lambda_1\cdots\lambda_m$ 是方阵 $\boldsymbol{A}$ 的 m 个特征值,$p_1,p_2,\cdots,p_m$ 是依次与之对应的特征向量, 如果 $\lambda_1,,\lambda_m$ 各不相等, 则 $p_1..p_m$ 线性无关.

0.10 基础习题

问题 0.5

已知 $\boldsymbol{A}=\begin{pmatrix}2&2&1\\2&5&2\\3&6&4\end{pmatrix}$, $\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵, 求 $\boldsymbol{A}^*$ 的特征值和特征向量.

解答. 第一个想法可能是先求出来伴随矩阵, 然后进行计算. 一个自然的问题是能不能找到 $\boldsymbol{A}$ 与 $\boldsymbol{A}^*$ 之间的联系呢?

根据定义 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, 可以导出 $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{A}^*\boldsymbol{x} = |\boldsymbol{A}|\boldsymbol{x} = \lambda\boldsymbol{A}^*\boldsymbol{x}$, 于是知道 $\frac{|\boldsymbol{A}|}{\lambda}\boldsymbol{x} = \boldsymbol{A}^*\boldsymbol{x}$, 也就是 $\boldsymbol{A}^*\boldsymbol{x} = \frac{|\boldsymbol{A}|}{\lambda}\boldsymbol{x}$. 伴随矩阵的若干个特征值是 $|A|/\lambda_i$, 特征向量和原来的矩阵特征向量相同. #

问题 0.6

设 是 n 阶可逆实对称矩阵, 是 n 阶实反对称矩阵, 且 =, 证明 + 是可逆矩阵.

解答. 考虑定义 $+ = ((+^{-1})) = |A||E + A^{-1}B| \neq 0$. #

问题 0.7

如果 n 阶矩阵 的元素全为 k, 那么 的 n 个特征值是.

解答. 可以知道特征多项式是 $^n - nk^{n-1}$, 于是特征值是 $nk, \overbrace{0, 0, \cdots, 0}^{n-1\text{个}}$. #

问题 0.8

若三阶矩阵 的特征值为 $1, -1, 2$, 那么齐次线性方程组 $(+2)\boldsymbol{x} = \boldsymbol{0}$ 的解为.

问题 0.9

设矩阵 $= \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}, P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, = P^{-1*}P$, 求 $+2$ 的特征值和特征向量.

解答. 由问题 0.10可以知道伴随矩阵的特征值是原来矩阵的行列式分别除以对应的原来特征值, 特征向量不变. 即:

$$= \implies {}^* = \frac{||}{}.$$

我们希望由上述式子凑出 $= P^{-1*}P$, 观察特征向量:

$$P^{-1*}PP^{-1} = \frac{||}{}P^{-1} \implies P^{-1} = \frac{||}{}P^{-1}.$$

因此 P^{-1} 是 的特征向量. #