线性代数学习笔记

February 23, 2023

写在前面

这是一份考前复习资料, 但是其中囊括了很多超过考试范围的内容, 一般是为了自己学习与记录使用, 因此一定是漏洞百出的. 对于准备考试的同学来讲, 直接查看每一章最后的线性代数中出现的内容一节是比较好的一种行为.

参考资料: 工程高等代数 (第二版) 中国地质大学出版; 高等代数与解析几何朱富海等. 科学出版社; Linear Algebra and its applications. 5th Ed.; The Essence of Linear Algebra. Grant Sanderson.; 工程数学线性代数同步辅导及习题全解中国水利水电出版社;

第一章　多项式

1.1 数域

每次做不封闭的运算的时候, 我们就要拓展数的关系了, 所以给出如下定义.

Definition 1.1. 设 $\mathbb{F}$ 是某些复数组成的集合, 其中至少含一个非零数, 且 $\mathbb{F}$ 对于四则运算封闭, 我们称 $\mathbb{F}$ 是一个数域.

数域上面自然有多项式的运算.

Definition 1.2. 设 $\mathbb{F}$ 是一个数域, x 是一个不定元, 那么 $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ 是以 x 为不定元的一元多项式.

1.2 一元多项式

1.3 辗转相除法

1.4 因式分解

1.5 不可约多项式

1.6 多元多项式

1.7 线性代数中出现的内容

TODO

1.8 基础习题

由于本章不作为线性代数的考试范围, 所以暂时没有总结.

第二章　行列式

2.1 行列式的定义以及完全展开

参考程序: det.py

2.2 行列式的性质

2.3 行列式的典型计算法

2.4 Cramer 法则

2.5 线性代数中出现的内容

一些简单的行列式

1. 问题的提出: 在求解二元一次方程组的时候, 注意到可以使用合适的记号表示整个方程问题. 见 AA 例 2.1.1.

2. 二阶行列式

   - 定义: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$ 应用: 求解方程组 $\begin{cases} a_{11}x_1 + a_{12}x_2 & = b_1 \\ a_{12}x_1 + a_{22}x_2 & = b_2 \end{cases}$ 的解为 (前提是 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 不是 0).

   $$x_1 = \frac{\begin{vmatrix} \boldsymbol{b_1} & a_{12} \\ \boldsymbol{b_2} & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, x_2 = \frac{\begin{vmatrix} a_{11} & \boldsymbol{b_1} \\ a_{21} & \boldsymbol{b_2} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

   - 几何意义: 描述一个空间经历线性变换之后, (面积/体积/更高维度的度量) 相对于原来改变的倍数.(EOLA)

3. 三阶行列式

   - 问题的提出: 我们猜测这个内容可以推广到更高阶解方程的问题上. 因此我们尝试解三解方程. 见 AA 例 2.1.3.
   - 定义及其由来: 注意到: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$,

继续展开就有:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{12}a_{13}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}.$$

- 应用: 解三元一次方程组 $\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3 & = b_1 \\ a_{12}x_1+a_{22}x_2+a_{23}x_3 & = b_2 \\ a_{13}x_1+a_{23}x_2+a_{33}x_3 & = b_3 \end{cases}$ 的解为 (前提是 $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ 不是 0).

$$x_1 = \frac{\begin{vmatrix} \boldsymbol{b_1} & a_{12} & a_{13} \\ \boldsymbol{b_2} & a_{22} & a_{23} \\ \boldsymbol{b_3} & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, x_2 = \frac{\begin{vmatrix} a_{11} & \boldsymbol{b_1} & a_{13} \\ a_{21} & \boldsymbol{b_2} & a_{23} \\ a_{31} & \boldsymbol{b_3} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & \boldsymbol{b_1} \\ a_{21} & a_{22} & \boldsymbol{b_2} \\ a_{31} & a_{32} & \boldsymbol{b_3} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

4. n 阶行列式

   - 定义及其由来: 定义一阶方阵的行列式 $A_1 = (a_{11})$ 为 a_{11}, 如果 $n-1$ 阶行列式已经定义好, 那么 n 阶行列式的定义为:

   $$\sum_{j=1}^{n}(-1)^{1+j}a_{1j}M_{1j}$$

   其中 M_{1j} 是划去 A 的第 1 行第 j 列的新得到的行列式,

   - 常见矩阵的行列式值
     - 上三角矩阵: $|A| = a_{11}a_{22}\cdots a_{nn}$, 其中 n 是上三角矩阵的阶数.
   - 下一步的要求: 需要找到 n 阶行列式的完全展开.

全排列和对换

1. 问题的提出: 在上面套娃展开的时候, 我们注意到所有的项数都有出现, 并且前面的符号只能为 +1 或者-1. 为了探究这个规律是不是成立以及何时前面是 +1, -1, 引入如下的操作.

2. 奇排列和偶排列. 设 $i_1, i_2, \cdots, i_n$ 是 n 个不同自然数的排列, 如果 $k < j$ 的时候 $i_k < i_j$, 那么称 i_k, i_j 构成一个**正序**. 否则称 i_k, i_j 构成一个**逆序**. 称一个排列的逆序数总数为这个排列的**逆序数**. 记为 $\tau(i_1, i_2, \cdots, i_n)$, 如果是奇数 (偶数), 称这个排列是奇 (偶) 排列.

3. 性质. 一个排列中的任意两个元素调换位置, 排列改变奇偶性.

4. 行列式的完全展开:

$$|A| = \sum_{(i_1 i_2 \cdots i_n)} (-1)^{\tau(i_1, i_2, \cdots, i_n)} a_{1i_1} a_{1i_2} \cdots a_{1i_n}$$

- 下一步操作: 需要寻找计算行列式的简单的方法, 因为现有的方法过于复杂.

行列式的性质

单个行列式的一些性质.

- 转置值不变: $|A| = |A^T|$.(AA 定理 2.2.2)
- 对换要变号: 任意对换行列式的两行 (两列), 行列式要变号. (AA 命题 2.2.4)
- 线性性 I: : 若某行 (列) 存在公因式, 可以提到外面来. (AA 命题 2.2.5)
- 线性性 II: 如果行列式某行某个元素都是两个元素的和, 这个行列式可以拆分成两个行列式的和. (AA 命题 2.2.6)
- 某一行 (列) 的 k 倍加到另一行 (列) 去, 行列式值不变.(AA 命题 2.2.8)

"划去" 的一般规律.

- 动机: 由于总是划去第一行有时候不是很方便, 因此希望可以每一次不在第一行划去, 或者挑选多行划去.
- 余子式和代数余子式: 见 AA 定理 2.2.10.
  - 可以自然推导出 **Cramer 法则**, 并且引入 **Kronecker 记号**, 见 AA 引理 2.4.1.
- Laplace 展开: 形式化叙述见 AA 定义 2.2.13. 先选定 m 行 (列) 不动, 然后在列 (行) 上面枚举所有可能的选择 m 列 (行) 的情况 (一共要枚举 $\binom{n}{m}$ 种情形), 并且根据所选两个排列的和的奇偶性判定前面符号, 最后相加.

一些特殊的例子:

$$\begin{vmatrix} A & C \\ \mathbf{0} & B \end{vmatrix} = \begin{vmatrix} A & \mathbf{0} \\ D & B \end{vmatrix} = |A||B| (\text{暗示了矩阵乘法})$$

$$\begin{vmatrix} 0 & C \\ -I_n & B \end{vmatrix} = |C|$$

Vandermonde 行列式.

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1\le j\le i\le n}(x_i - x_j).$$

(与算法竞赛的联系) 这个公式保证了单位根构造多项式是可行的, 因此在 FFT 中间遇见过很多次.

2.6 常见习题及其思想

求行列式的值:

求下列行列式的值:

Problem 2.1. $D_5 = \begin{vmatrix} b & c & & & \\ a & b & c & & \\ & a & b & c & \\ & & a & b & c \\ & & & a & b \end{vmatrix}$

Solution. 考虑递推关系. 不妨先按照第一行第一列展开, 有

$$\begin{aligned} D_5 =& b\begin{vmatrix} b & c & & \\ a & b & c & \\ & a & b & c \\ & & a & b \end{vmatrix} - c\begin{vmatrix} a & c & & \\ & b & c & \\ & a & b & c \\ & & a & b \end{vmatrix} \\ =& bD_4 - c\left(a\begin{vmatrix} b & c & \\ a & b & c \\ & a & b \end{vmatrix}\right) \\ =& bD_4 - acD_3 \end{aligned}$$

然后就可以按照这样的方法得到递推的解答. 不难发现, $D_n = bD_{n-1} + acD_{n-2}$ 是对于一般的 n 成立的.

Problem 2.2. $D = \begin{vmatrix} a_1 + x & a_2 & a_3 & a_4 \\ -x & x & & \\ & -x & x & \\ & & -x & x \end{vmatrix}$

Solution. 不妨把后面三列的所有的元素加到第一列上面去, 这样就可以得到类似 $\begin{vmatrix} x+\sum_{i=1}^4 a_i & & & \\ & x & & \\ & -x & x & \\ & & -x & x \end{vmatrix}$

的矩阵. 展开之后就可以处理剩下的式子 $\begin{vmatrix} x & & \\ -x & x & \\ & -x & x \end{vmatrix}$, 就是比较轻松的. 根据2.1得到的答案是 $x^3\left(x+\sum_{i=1}^4 a_i\right)$.

这个问题给我们另一个启发就是, 如果行与行之间的元素是相似或者可以相互抵消的, 可以把它们加到同一行上面去.

Problem 2.3. $D=\begin{vmatrix} a_1 & 1 & 1 & 1 \\ 1 & a_2 & & \\ 1 & & a_3 & \\ 1 & & & a_4 \end{vmatrix}$

Solution. 不妨提取出 a_2,a_3,a_4, 并且最后可以让上面的 1 来与之相减. 具体的, 有

$$D=a_2a_3a_4\begin{vmatrix} a_1 & 1 & 1 & 1 \\ 1/a_2 & 1 & & \\ 1/a_3 & & 1 & \\ 1/a_4 & & & 1 \end{vmatrix}=a_2a_3a_4\begin{vmatrix} a_1-\sum_{i=2}^4 1/a_i & 0 & 0 & 0 \\ 1/a_2 & 1 & & \\ 1/a_3 & & 1 & \\ 1/a_4 & & & 1 \end{vmatrix}=a_2a_3a_4\left(a_1-\sum_{i=2}^4 1/a_i\right)$$

Problem 2.4. $D=\begin{vmatrix} x & 0 & 0 & \cdots & 0 & a_0 \\ -1 & x & 0 & \cdots & 0 & a_1 \\ 0 & -1 & x & \cdots & 0 & a_2 \\ & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & a_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & x+a_{n-1} \end{vmatrix}$

Solution. 这个矩阵巧妙的把多项式和矩阵联系在了一起. 事实上, 他表示的是 $x^n+a_{n-1}x^{n-1}+\cdots+a_0$. 事实上, 这个想法在后面证明多项式与矩阵的联系 (Jordan 标准型一节, 线性代数不讲) 打下了坚实的基础. 下面我们来证明这件事情.

$$D_{n-1}=\begin{vmatrix} x & 0 & 0 & \cdots & 0 & a_0 \\ -1 & x & 0 & \cdots & 0 & a_1 \\ 0 & -1 & x & \cdots & 0 & a_2 \\ & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & a_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & x+a_{n-1} \end{vmatrix}$$

$$=x\begin{vmatrix} x & 0 & \cdots & 0 & a_1 \\ -1 & x & \cdots & 0 & a_2 \\ & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x & a_{n-2} \\ 0 & 0 & \cdots & -1 & x+a_{n-1} \end{vmatrix}+\begin{vmatrix} 0 & 0 & \cdots & 0 & a_0 \\ -1 & x & \cdots & 0 & a_2 \\ & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x & a_{n-2} \\ 0 & 0 & \cdots & -1 & x+a_{n-1} \end{vmatrix}$$

$$=xD_{n-2}+\left(\begin{vmatrix} 0 & \cdots & 0 & a_0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & x & a_{n-2} \\ 0 & \cdots & -1 & x+a_{n-1} \end{vmatrix}\right)=xD_{n-2}.$$

根据递推关系就可以得到这个结论. 这样的方法在2.1中也曾经用到过.

与之同样有趣的是,$D=\begin{vmatrix} x & 0 & 0 & \cdots & 0 & a_0 \\ -1 & x & 0 & \cdots & 0 & a_1 \\ 0 & -1 & x & \cdots & 0 & a_2 \\ & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & a_{n-1} \\ 0 & 0 & 0 & \cdots & -1 & a_n \end{vmatrix}$一个 $n+1$ 阶的行列式表示的也是一个多项式 $\sum_{i=0}^{n} a_i x^i$, 这次它的系数和指标是相对应的.

第三章　矩阵

3.1 矩阵的基本概念和运算

3.2 可逆矩阵

3.3 矩阵的初等变换和初等矩阵

3.4 矩阵的相抵和秩

矩阵在图形处理的应用

3.5 线性代数中出现的内容

问题的提出: 由上一章的内容, 对于 n 个未知数, n 个变量的方程组, 我们可以使用 Creamer 法则判定是否存在解. 但是对于解的**自由度**有多少还是没有概念. 并且, 如果我们有 n 个未知数, m 个方程, 这样子解的自由度又是多少呢? 于是我们考虑引入矩阵的概念, 同样可以看作线性方程组的系数抽离开来得到的抽象的内容.

矩阵的加法和数乘以及转置

1. 定义 (略). 参见 AA 定义 3.1.1, 3.1.2
   (a) 注意: 矩阵的提取公因式和释放公因式是对于整个矩阵的全体元素的, 因此有 $|\lambda A| = \lambda^n |A|$.

矩阵的乘法

1. 问题的提出: 我们可以运用代换法代换线性方程组的某一些元素, 描述见 AA3.1.3 下方的文本.
2. 定义: 见 AA 定义 3.1.4, 并且注意只有 A 的列数等于 B 的行数的时候才可以完成乘法.
3. 性质: 没有交换律, 但是有结合律. 见 AA 命题 3.1.5.
4. 矩阵的幂: 因为满足结合律因此可以定义, 参见 AA 命题 3.1.5 下方文本.
   (a) 与多项式的联系: 如果把多项式 $f(x)$ 的变量 x 换做矩阵 A, 那么知道 $f(A)+g(A)=(f+g)(A)$, $f(A)g(A)=(fg)(A)$.
5. 下一步: 定义了乘法, 就希望有一个类似于除法的东西. 也就是需要研究可逆矩阵.

可逆矩阵

1. 问题的提出: 定义了乘法, 就希望有一个地位与除法相似的内容, 通常也成为 “逆元”.
2. 行列式与矩阵的联系: $|AB| = |A||B|$.
3. 可逆的充分必要条件: $|A| \neq 0$. 参见 AA3.2.4, 结合此, 就可以得到**伴随矩阵**的定义.
4. 计算可逆矩阵
   (a) 想法 1: 伴随矩阵 (Cramer 法则)
      i. 定义: 见 AA 定义 3.2.5(略)
      ii. 性质: $A^*A = AA^* = |A|I_n$.(与 Cramer 法则等价)
   (b) 想法 2: 矩阵的初等变换 (消元法)
      i. 定义: 三类初等矩阵: $P(i,j)$,$P(i(c))$, $P(i,j(c))$. 见 AA3.3 节.
      ii. 作用: 左 (右) 边乘变换目标矩阵的行 (列), P(i, j) 交换 i, j 两行 (列), P(i(c)), 把第 i 行 (列) 乘上 c 倍; P(i,j(c)), 第 j 行 (列) 乘上 c 倍加到第 i 行 (列).
      iii. 矩阵的秩: 因为矩阵总是可以化为形状如 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ 的. 因此定义 r 是矩阵的秩. (等价关系, 揭示了同样自由度的解的相互关系, 告诉我们了方程的解的自由度到底是多少).
      iv. 性质: 初等行列变换总是不改变矩阵的秩 (因为操作都可逆).

分块矩阵及其的逆

1. 分块矩阵的基础运算
   (a) 分块矩阵的加法 (略). 条件: A 与 B 要有**相同的分块方法**.
   (b) 分块矩阵的数乘.(略)
   (c) 分块矩阵的转置. 不仅分块矩阵的行列互换, 每个分块还要转置.
   (d) 分块矩阵的乘法. 前提: A 的**列分法**与 B 的**行分法**要一致.
2. 特殊的分块矩阵
   (a) 分块对角矩阵: $A_{s\times s} = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_s \end{pmatrix}$. 如果可逆, 其逆是 $(A_{s\times s})^{-1} = \begin{pmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \cdots & \\ & & & A_s^{-1} \end{pmatrix}$.

(b) 按行分块, 得到列向量.

(c) 按列分块, 得到行向量.

3. 分块矩阵的初等变换

(a) 第一类: $\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{\text{交换两行}} \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix}$,

(b) 第二类: $\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{\text{某一行乘上}P} \begin{pmatrix} P & 0 \\ 0 & I_m \end{pmatrix}$ 或 $\begin{pmatrix} I_n & 0 \\ 0 & P \end{pmatrix}$.

(c) 第三类: $\begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \xrightarrow{\text{某一行乘上}P\text{加到另一行}} \begin{pmatrix} I_n & 0 \\ P & I_m \end{pmatrix}$.

4. 应用: 解矩阵方程.

(a) AX=B, 那就 $\begin{pmatrix} A & \vdots B \end{pmatrix} \to \begin{pmatrix} I & \vdots A^{-1}B \end{pmatrix}$

(b) XA=B, 那就 $\begin{pmatrix} A \\ B \end{pmatrix} \to \begin{pmatrix} I \\ BA^{-1} \end{pmatrix}$.

矩阵的秩

1. 秩的等价定义:

(a) 化为 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ 类型之后, r 就是它的秩 (AA 定理 3.4.6)

(b) $r(A) = r$ 当且仅当 A 中存在一个 r 阶子式不为 0, 任意 $r+1$ 阶子式全为 0.(AA 定理 3.4.9)

(c) 后续学习向量组后的行 (列) 秩, 列秩也和矩阵秩的定义是等价的.

2. 常见不等式: 通常情况下无法得到准确的秩

(a) $r(A) + r(B) = r\begin{pmatrix} A & \\ & B \end{pmatrix} \leq r\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$.(AA 命题 3.4.11)

(b) 如果 $A_{m\times n}, B_{n\times s}$, 那么 $r(AB) + n \geq r(A) + r(B)$. (命题 3.4.12, Sylvester 不等式)

i. 思路: 考虑 $r\begin{pmatrix} AB & \\ & -I_n \end{pmatrix} \overset{=}{\text{等价变换}} r\begin{pmatrix} AB & A \\ 0 & -I_n \end{pmatrix} = r\begin{pmatrix} 0 & A \\ B & -I_n \end{pmatrix} = r\begin{pmatrix} A & 0 \\ -I_n & B \end{pmatrix} \geq r\begin{pmatrix} A & \\ & B \end{pmatrix}$

3. 应用: 解方程组

(a) 如果方程组 $AX = 0$ 只有零解, 那么 $|A| \neq 0$.

(b) 如果有方程组 $A_{m\times n}X = \beta_{m\times 1}$, 称 $(A|\beta)$ 为**增广矩阵**, 方程组有解当且仅当 $r(A) = r(A|\beta)$.

(c) 如果有 $A_{m\times n}, B_{m\times 1}, P_{m\times m}$, 且 P 可逆, $AX = \beta$ 与 $(PA)X = P\beta$ 的解相同.

3.6 基础问题

3.6.1 矩阵的相关运算

Problem 3.1. 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶对称矩阵, 则 $\boldsymbol{AB}$ 是对称矩阵的充分必要条件是 ____________.

Solution. $\boldsymbol{AB}$ 对称等价于 $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}' = \boldsymbol{BA}$, 所以要求两个矩阵是可换的.

Problem 3.2. 设 α,β 是三维列向量, 如果 $\alpha\beta' = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$, 求 $\alpha'\beta =$____________.

Solution. 不妨设 $\alpha = [x_1, x_2, x_3]'$,$\beta = [y_1, y_2, y_3]'$, 那么 $\alpha\beta' = \begin{pmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \\ x_3y_1 & x_3y_2 & x_3y_3 \end{pmatrix}$, $\alpha'\beta = x_1y_1 + x_2y_2 + x_3y_3 = a + e + i$, 也就是这个矩阵的**迹 (trace)**. 一般记作 tr(A). 在后面我们还可以看到这个内容再次出现.

Problem 3.3. 已知 $A = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{pmatrix}$, 证明 $A^2 = \mathrm{tr(A)}A$.

Solution. 可以把它分解开成两个矩阵的乘积. 然后用结合律和上一题的结论就可以知道了.

Problem 3.4. 已知 $A = \begin{pmatrix} \lambda & & \\ 1 & \lambda & \\ & 1 & \lambda \end{pmatrix}$, 求 A^n.

Solution. 注意到这是后续的 Jordan 块. 实际上是求导的具体的线性变换. 从这个角度来看, 我们可以先把他拆开, 然后用这样的方法:

$$\lambda = \lambda I_3 + J$$

使用如下的公式:

$$
\begin{aligned}
A^n =&(\lambda I_3 + J)^n \\
=&\lambda^n I_3 + \binom{n}{1}\lambda^{n-1}J + \binom{n}{2}\lambda^{n-2}J^2 \\
=&\begin{pmatrix} \lambda^n & & \\ & \lambda^n & \\ & & \lambda^n \end{pmatrix} + \begin{pmatrix} & \binom{n}{1}\lambda^{n-1} & \\ & & \binom{n}{1}\lambda^{n-1} \\ & & \end{pmatrix} + \begin{pmatrix} & & \binom{n}{2}\lambda^{n-2} \\ & & \\ & & \end{pmatrix}
\end{aligned}
$$

Problem 3.5. 已经知道 $A = \begin{pmatrix} 3 & -1 \\ 9 & 3 \end{pmatrix}$, 求 A^n.

Solution. 对于这样的问题, 我们可以在一个更方便的基底下做这个操作, 也就是让它成为 $P = T^{-1}AT$. 然后进行操作就行了.

Problem 3.6. 已知 A, B 满足 n 阶矩阵,

- 有 $AB = A - 2B$, 求 $(A + 2E)^{-1}$.
- 有 $B = (E + A)^{-1}(E - A)$, 求 $(E + B)^{-1}$.
- 有 $A^2 + 3A - 2E = \mathbf{0}$, 求 $(A + E)^{-1}$.

那么如果有 $AB = A - 2B$, 可以求 $(A + 3E)^{-1}$ 吗?
那么如果有 $A^2 - A - 2E = 0$, 可以求 $(A + 3E)^{-1}$ 吗?

Solution. 注意到前面的三个可以用配方法加上逆矩阵去做, 它的主题思路是现通过配凑的方法干掉平方项, 再干掉一次项, 最后只留下常数项. 类似于中学学习过的分离常数. 只不过在下面的问题里面进行分离常数之后剩下的内容形式已经不是那么好了, 也就是得到了 $(A + 3E)(B - 1) = B - 3E$, 如果 A 和 B 是同一个矩阵的话形式就会简洁一些.

Problem 3.7. 如果 $A, B, A + B, A^{-1} + B^{-1}$ 均为 n 阶可逆矩阵, 那么 $(A^{-1} + B^{-1})^{-1}$ 是 _______.

Solution. 这道题实际上是3.6问题的一个拓展. 关键是如何把这个和写成一个乘积的形式. 实际上是做一个基变换. 于是我们有如下的内容:

$$
\begin{aligned}
(A^{-1} + B^{-1})^{-1} =&(EA^{-1} + B^{-1}E)^{-1} \\
=&(B^{-1}BA^{-1} + B^{-1}AA^{-1}) \\
=&\left(B^{-1}(A + B)A^{-1}\right)^{-1} \\
=&A(A + B)^{-1}B
\end{aligned}
$$

Problem 3.8. 设 $\boldsymbol{A}$ 为 n 阶非奇异矩阵, $\boldsymbol{\alpha}$ 为 n 维列向量, b 为常数, 记分块矩阵

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{E} & 0\\ -\boldsymbol{\alpha}^T\boldsymbol{A}^* & |\boldsymbol{A}|\end{pmatrix},\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}^T & b\end{pmatrix}$$

计算并化简 $\boldsymbol{PQ}$; 证明 $\boldsymbol{Q}$ 可逆的充分必要条件为 $\boldsymbol{\alpha}^T\boldsymbol{A}^{-1}\boldsymbol{\alpha}\neq b$.

Solution. 这个题目中出现的矩阵问题有什么更加深刻的见解吗?

Problem 3.9. 若已知 n 阶行列式 $|A|=\det\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 2 & \cdots & 0\\ \vdots & \vdots & & & \vdots\\ 0 & 0 & 0 & \cdots & n-1\\ n & 0 & 0 & 0 & 0\end{pmatrix}$, 则 |A| 的第 k 行代数余子式的和是 _______.

Solution. 运用分块法和一些基础的操作即可.

3.6.2 矩阵的初等行列变换和矩阵的秩

Problem 3.10. 如果 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{B}$ 是 $n\times s$ 矩阵, 证明 $r(\boldsymbol{AB})\leq\min(\mathrm{rk}(\boldsymbol{A}),\mathrm{rk}(\boldsymbol{B}))$.

Problem 3.11. 证明伴随矩阵秩的公式: $\mathrm{rk}(\boldsymbol{A}^*)=\begin{cases}n & \mathrm{rk}(\boldsymbol{A})=n\\ 1 & \mathrm{rk}(\boldsymbol{A})=n-1\\ 0 & \mathrm{rk}(\boldsymbol{A})<n-1\end{cases}$.

Problem 3.12. 如果 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{B}$ 是 $n\times s$ 矩阵, 证明 $r(\boldsymbol{AB})\leq\mathrm{rk}(\boldsymbol{B})$.

Problem 3.13. 如果 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{B}$ 是 $n\times s$ 矩阵, 如果 $\boldsymbol{AB}=\boldsymbol{0}$, 证明 $\mathrm{rk}(\boldsymbol{A})+\mathrm{rk}(\boldsymbol{B})\leq n$.

第四章　线性空间

4.1 线性空间的基本概念

我们研究问题是从具体到抽象的, 从特殊到一般的, 于是自然就要考虑一下我们以前探讨过的多项式, 矩阵, 和普通的数在操作有什么共通的地方.

我们先来考虑加法的共性. 得到了 $V1 \sim V4$.

然后考虑减法的共性, 这时候就要依赖一个数域 F, 得到了 $V6 \sim V8$.

其中 $V5$ 在这里好像不是必须的, 因为我们可以从其他的几条性质里面推导出来.

但是这八条性质有更加深刻的意义. $V2 + V3 + V4 =$ 群的定义, $V1 +$ 群的定义 $=$ 交换群的定义, $V5 + V6 + V7 + V8 =$ 模的定义. 注意我们在写定义的时候不要用上交换律的性质: 在左边都在左边, 在右边都在右边.

Definition 4.1. (线性空间的八条性质) 设 V 是一个非空集合, $\mathbb{F}$ 是一个数域, 在 V 中定义了一个一种运算, 成为**加法**, 即, 对于任意的 $\alpha,\beta \in V$, 存在唯一的元素 $\gamma \in V$ 与之对应, 记作 $\gamma = \alpha + \beta$. 在 $\mathbb{F}$ 与 V 之间定义了一个运算, 存在唯一的元素 $\gamma \in V$ 与之对应, 记为 $\gamma = \alpha + \beta$; 在 $\mathbb{F}$ 与 V 之间定义了一个运算, 称为**数乘**. 即对于任意的 $k \in \mathbb{F}, \alpha \in V$, 存在唯一的 $\delta \in V$ 与之对应. 记作 $\delta = k\alpha$ 或 $k \cdot \alpha$. 称 V 为 $\mathbb{F}$ 上的**线性空间**或**向量空间**. 如果对于任意的 $\alpha, \beta, \gamma \in V$ 和 $k, l \in \mathbb{F}$, 加法和数乘满足下列八条运算法则:

(1) 加法交换律: $\alpha + \beta = \beta + \alpha$;

(2) 加法结合律: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$;

(3) 存在零元: 存在 $0 \in V$, 使得对于任意的 $\alpha \in V, 0 + \alpha = \alpha$, 这个元素称为 V 的零向量.

(4) 存在负向量: 对于任意的 $\alpha \in V$, 存在 $\beta \in V$, 使得 $\alpha + \beta = 0$. 称 β 是 α 的**负向量**.

(5) 存在单位元: 对 $\mathbb{F}$ 中的 1, 任意的 $\alpha \in V$, 有 $1 \cdot \alpha = \alpha$.

(6) 对于任意的 $k, l \in \mathbb{F}, \alpha \in V$, 有 $(kl)\alpha = k(l\alpha)$.

(7) 对于任意的 $k, l \in \mathbb{F}, \alpha \in V$, 有 $(k + l)\alpha = k\alpha + l\alpha$.

(8) 对于任意的 $k \in \mathbb{F}, \alpha, \beta \in V$, 有 $k(\alpha + \beta) = k\alpha + k\beta$.

此时 V 中的元素也称为向量, $\mathbb{F}$ 为 V 的**基域**.

从线性空间的定义, 我们可以知道一些简单的性质:

(1) 零向量是唯一的.

如果有 $0'$ 使得 $\forall$

4.2 线性空间的基与维数

在之前的内容中, 我们了解了线性空间的定义. 在之前我们做的事情就是做线性组合, 形式化的我们给出定义:

Definition 4.2. (线性组合) TODO

我们来举一些例子:

Example 4.1. (有限维多项式)

Example 4.2. (矩阵多项式) 如果 $A \in \mathbb{F}^{n\times n}$, $f(x) \in \mathbb{F}[x], f(x) = a_o x_0 + a_1 x^1 + \cdots + a_n x^n$, 如果对 $f(x)$ 赋值为一个矩阵 $f(A) = a_0 I_n + a_1 A + a_2 A^2 + \cdots + a_n A^n$, 问是否存在 $f(A) = 0$, 也就是 $I_n, A, \cdots, A^n$ 的线性组合为 0?

回顾上面两个例子, 我们为了刻画这种关系, 来描述**线性相关性**如下:

Definition 4.3. (线性相关性)

Example 4.3. (平面向量, 空间向量的线性相关和线性无关)

我们希望找到一些最少的向量把空间 “撑起来”, 于是我们给出如下的定义:

Definition 4.4. (极大线性无关组)

如何求解极大线性无关组呢? 我们可以使用**筛选法**.

Definition 4.5. (筛选法)

运用筛选法得到的极大线性无关组是不唯一的, 但是筛选得到的元素的个数是唯一的. 我们把它定义为这个向量组的**秩**.

Definition 4.6. (基与维数) 设 $V_{/\mathbb{F}}$, 若 $\alpha_1..\alpha_n \in V$ 且线性无关, 且 $\forall \alpha \in V$ 都可以用 $\alpha_1..\alpha_n$ **的线性组合表出**, 则称 $\alpha_1..\alpha_n$ 为 V 的一组基. 称 n 为 $V/\mathbb{F}$ 的维数.

Example 4.4. 平面是 2 维的, 空间是 3 维的.

Theorem 4.1. *若 $\beta_1..\beta_n$ 是 V 的一组基, 则 $\beta_1..\beta_n$ 与 $\alpha_1..\alpha_n$ 线性等价且可以互相表出 $\Rightarrow n = m$.*

?proofname? 由于 $\beta_1..\beta_m$ 可以与 $\alpha_1..\alpha_n$ 线性等价, 我们可以得到这样的方程组, 其中 $a_{11}..a_{mn}$ 是常数:

$$\begin{cases} \beta_1 = & a_{11}\alpha_1 + .. + a_{1n}\alpha_n \\ \beta_2 = & a_{21}\alpha_1 + .. + a_{2n}\alpha_n \\ \vdots & \vdots \\ \beta_m = & a_{m1}\alpha_1 + .. + a_{mn}\alpha_n \end{cases}$$

由于这个记号看起来非常的难看, 我们不妨考虑简化记号: 比如

$$\sum_{i=1}^{n} a_i b_i = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \end{pmatrix}$$

用上这个表示技巧, 我们就可以把 β_1 表示为如下所示的内容. 注意, $(\alpha_1..\alpha_n)$ 只是形式上的行向量, 每一个 $\alpha_1\cdots\alpha_n$ 其实上是一个列向量. 后面的内容才是真正在 $\mathbb{F}^{n\times 1}$ 中的元素.

$$
\begin{aligned}
\beta_1 &= \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}a_{11}\\ \vdots \\ a_{1n}\end{pmatrix}\\
\beta_2 &= \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}a_{21}\\ \vdots \\ a_{2n}\end{pmatrix}\\
&\cdots\cdots\\
\beta_m &= \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}a_{m1}\\ \vdots \\ a_{mn}\end{pmatrix}
\end{aligned}
$$

于是考虑把上述的内容组合到一起:

$$
(\beta_1,\beta_2,\cdots,\beta_m) = (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ \ldots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}
$$

分类讨论:

Case 1. 若 $n\neq m$,

Case i. 若 $n<m$, $A\in\mathbb{F}^{n\times m}, AX=0$, 说明一定存在非零解 $\eta\in\mathbb{F}^{m\times 1}$ 成立. 证明如下:

$$
(\beta_1..\beta_n) = (\alpha_1..\alpha_n)\,A
$$

两端同时右乘 η, 有 $(\beta_1..\beta_n)\,\eta = ((\alpha_1..\alpha_n)\,A)\,\eta$. 可以证明形式矩阵的乘法具有结合律, 于是有 $(\beta_1..\beta_n)\,\eta = (\alpha_1..\alpha_n)\,(A\eta)$, 而 $A\eta=0$, 说明 $\beta_1..\beta_n$ 线性相关, 矛盾!

Case ii. $n<m$, 同理.

Case 2. $n=m$, 由于自然数的三歧性, 维数存在, 说明 $n=m$ 成立.

□

上例中形式矩阵具有结合律的一个证明:

$$\begin{aligned}
\left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix} A\right)\eta &= \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}a_{11} & .. & a_{1m}\\ \vdots & & \vdots\\ a_{n1} & .. & a_{nm}\end{pmatrix}\right)\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}\\
&= \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix}, .., \begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}a_{1m}\\ \vdots\\ a_{nm}\end{pmatrix}\right)\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}\\
&= \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix}\right)\eta_1 + .. + \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}a_{1m}\\ \vdots\\ a_{nm}\end{pmatrix}\right)\eta_n\\
&= \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}\eta_1 a_{11}\\ \vdots\\ \eta_1 a_{n1}\end{pmatrix}\right) + .. + \left(\begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}\begin{pmatrix}\eta_n a_{1m}\\ \vdots\\ \eta_n a_{nm}\end{pmatrix}\right)\\
&= \begin{pmatrix}\alpha_1 & .. & \alpha_n\end{pmatrix}(A\eta)
\end{aligned}$$

类似于平面几何和立体几何, 我们也希望有一个坐标来更好的研究线性空间, 于是我们给出坐标的定义:

Definition 4.7. (坐标)

可以验证, 在基给定的情形下, 坐标是唯一确定的. 从4.1可以看出, 每组基的元素个数是相等的. 也就是

Theorem 4.2. *若 $\alpha_1..\alpha_n$ 是线性空间 V 上的一组基, $\beta_1..\beta_m$ 也是 V 上的一组基, 则 m=n.*

那么我们自然的会想到有时候基不是很方便, 我们希望得到一个更加方便的基底, 这样我们就可以更好的考察这个线性空间的若干性质. 所以我们定义**过渡矩阵**. 通过右乘一个这样的矩阵, 我们可以顺利的把基进行变换. 那么, 要变成什么样的矩阵比较好呢? 实际上, 这样的矩阵会在下一章讲解线性变换的时候有大用处, 并且还会定义一个方阵的特征值和特征向量.

Definition 4.8. (过渡矩阵) 我们可以使用右乘过渡矩阵方法完成基的变换. 设 $\alpha_1..\alpha_n$ 和 $\beta_1, .., \beta_n$ 是 V 上的两组基, 有

$$(\beta_1..\beta_n) = (\alpha_1..\alpha_n)\begin{pmatrix}a_{11} & .. & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & .. & a_{nn}\end{pmatrix} := (\alpha_1..\alpha_n)(T_1..T_n)$$

称 $\begin{pmatrix} a_{11} & .. & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & .. & a_{nn} \end{pmatrix}$ 是从 $(\alpha_1..\alpha_n)$ 到 $(\beta_1..\beta_n)$ 的**过渡矩阵**. 并且 T_j 是 β_j 在 $\alpha_1..\alpha_n$ 下的坐标.

Theorem 4.3. 上述过渡矩阵一定可逆.

?proofname? 如果过渡矩阵不可逆, 意味着 $AX = 0$ 存在非零解答, 左右两边同时右乘非零解, 矛盾. □

Lemma 4.1. 下列命题等价*:*

- V 为 $n(n < \infty)$ 维线性空间
- $\alpha_1..\alpha_n$ 是一组基底
- $\forall \alpha \in V$ 都可以被 $\alpha_1..\alpha_n$ 线性表出

Problem 4.1. 如果 $A \in \mathbb{F}^{n\times n}$, 是否存在 $f(x) \in \mathbb{F}[x]$, 使得 $f(A) = 0$, 但是 a_i 不全为 0?

线性变换中, 在哪个数域上面看也是不可少的. 有时候在一个数域上面看是线性空间, 在另一个数域上面看就不是了. 下面来举几个例子.

Example 4.5. $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}|a, b \in \mathbb{Q}\}$, $\mathbb{Q}(\sqrt{2})$ 既可以看作有理数域上的线性空间, 又可以看作自身的线性空间, 即 $\mathbb{Q}(\sqrt{2})_{/\mathbb{Q}(\sqrt{2})}$.

Theorem 4.4. $\mathbb{F} \subseteq \mathbb{K} \subseteq \mathbb{E}$ 是数域*,* $dim\mathbb{E}/\mathbb{F} = dim\mathbb{E}/\mathbb{K} \cdot dim\mathbb{K}/\mathbb{F}$*.*

4.3 线性同构和线性映射

回忆我们曾经提到过的线性空间, 有加法和数乘. 线性空间中的一个元素 α 都可以用这个线性空间的一组基以及他们前面搭配上对应的系数 (在给定的数域中) 来表示. 比如 $\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n$. 并且在上一节中我们也定义了坐标的概念. 那么, 是不是空间里面的所有的元素都能用坐标表示呢? 你可能会觉得这是显然可以的. 下面来简单看一下这方面的更深层次的内涵.

我们知道, 如果在 $\mathbb{F}$ 上的 n 维线性空间里面, 如果 $\alpha_1..\alpha_n$ 是一组基, $\forall \alpha \in V, \exists! x_1..x_n \in \mathbb{F}$, 有下图关系的映射:

$$
\begin{array}{ccc}
\boldsymbol{\alpha} & = \boldsymbol{x_1 a_1} + .. & + \boldsymbol{x_n a_n} \\
& = \quad (\alpha_1..\alpha_n) & \begin{pmatrix} \boldsymbol{x_1} \\ \vdots \\ \boldsymbol{x_n} \end{pmatrix} \\
\varphi : \boldsymbol{V} & \longrightarrow & \mathbb{F}^{\boldsymbol{n\times 1}} \\
\uparrow\in & \text{唯一确定} & \\
\alpha & \longrightarrow & X = \varphi(\alpha) \\
\beta & \longrightarrow & Y = \varphi(\beta) \\
\alpha + \beta & \longrightarrow & X + Y = \varphi(\alpha) + \varphi(\beta) \\
k\alpha & \longrightarrow & kX = \varphi(k\alpha)
\end{array}
$$

为了看看我们定义的“坐标”是不是真正可以完成这些内容的转换, 我们自然要问这些问题:

1. φ 是单射吗? 是的. 因为上节课中定理4.1可以看出.

2. φ 是满射吗? 是的, 这是基的定义告诉我们的. 这样就推出了 φ 是双射.

3. φ 保持加法吗? 可以验证 $\varphi(\alpha+\beta) = \varphi(\alpha) + \varphi(\beta)$.

4. φ 保持数乘吗? 可以验证 $\varphi(k\alpha) = k\varphi(\alpha)$.

于是我们说, 在线性空间上, $\mathbb{F}^{1\times n}$ 与 V 没有差别.

顺便提一句, 如果 3, 4, 我们称这个映射为一个**线性映射**, 如果同时满足 1, 2, 3, 4, 我们称之为一个**线性同构**.

Definition 4.9. (线性映射) 设 V, U 都是 $\mathbb{F}$ 上的线性空间, 若对任意 $\alpha, \beta \in V, k \in \mathbb{F}$, 映射 $\mathscr{A} : V \to U$, 满足:

$$\mathscr{A}(\alpha+\beta) = \mathscr{A}(\alpha) + \mathscr{A}(\beta), \mathscr{A}(k\alpha) = k\mathscr{A}(\alpha).$$

Definition 4.10. (线性同构) 记作 $U \simeq V$

Problem 4.2. 如果 $\dim V/\mathbb{F} = n, V \simeq \mathbb{F}^{n\times 1}$, 并且 $\dim U/\mathbb{F} = n, U \simeq \mathbb{F}^{n\times 1}$, 问 $V \simeq U$ 吗?

或许答案是肯定的, 这时候我们只需要从 $\mathbb{F}^{n\times 1}$ 中转一下就可以了. 具体的, 有:

$$
\begin{tikzcd}
V \arrow[r, "\varphi"] \arrow[dr, "\psi^{-1}\circ\varphi"'] & \mathbb{F}^{n\times 1} \arrow[d, shift left, "\psi^{-1}"] \\
& U \arrow[u, shift left, "\psi"]
\end{tikzcd}
$$

上面的图示是我们的设想, 我们还有一些问题需要解答:

1. ψ^{-1} 存在吗? 由于上面是双射, 所以是存在的, 并且也为双射.

2. $\psi^{-1} \circ \varphi$ 也是双射吗? 是的, 因为两个映射都是双射.
3. 这个合成是线性的吗? 只需要证明 ψ^{-1} 是线性的, φ 也是线性的, 两个线性映射的合成也设线性的就可以了.

证明: ψ^{-1} 是线性的

TODO

证明: 两个线性映射的合成是线性的.

TODO

我们发现线性同构也是一个等价关系, 也就是满足自反性, 对称性, 传递性.

Proposition 4.1. *(线性同构是等价关系)*

至此我们可能要问一问为什么要引入线性空间. 毕竟在一个 $\mathbb{F}^{n\times 1}$ 的上面做事情也挺好的. 原因之一是很多时候很多空间难以处理. 例如这个空间:

Example 4.6. 考虑 $\mathbb{R}^+$, 定义 $a \oplus b = ab, k \circ a = a^k, k \in \mathbb{R}, a, b \in \mathbb{R}^+$. 现在求这个空间的维数.

我们立刻发现这个空间的维数不是很好直接观察出来. 因为基就十分的难找. 这时候我们不妨使用 $\ln, \exp$ 两个在这个空间下是线性的映射来看一看, 这样我们就看到了这个空间化为了 $\mathbb{R}^+$, 因此这个空间的维数就是 1.

在无穷维的情况下, 考虑 $\mathbb{F}^{1\times\infty}$ 是无法办到的. 这时候就需要我们直面这个抽象的线性空间了.

下面我们来解决一个同样维数之间 (不是无穷维) 的映射, 单射与满射的问题. 因为之后的问题解决单射和满射可能不是很直接, 但是如果我们能得到两个空间维数相同, 那么单射和满射就等价了. 于是我们给出这个引理:

Lemma 4.2. *考虑 $\varphi : V \to U$, V, U 均为 n 维的, 那么 φ 是单射 $\Leftrightarrow \varphi$ 是满射.*

?proofname? 记 $\mathrm{Im}\ \varphi = \{\varphi(\alpha) | \alpha \in V\} = \varphi(V)$,TODO □

Definition 4.11. 若 $\varphi : V \to U$ 满足:

1. $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$
2. $\varphi(k\alpha) = k\varphi(\alpha)$

就称其为**线性映射**. 记所有线性映射的全体为 $\mathrm{Hom}(V, U)$.

Example 4.7. (线性方程组)

我们知道可以把线性方程组写成如下的形式:

$$AX = \beta, A \in \mathbb{F}^{m\times n}, \beta \in F^{m\times 1}$$

如果**把左乘 A 看做一个映射**的话, 这个映射是把 $\mathbb{F}^{n\times 1}$ 映到了 $\mathbb{F}^{m\times 1}$ 了. 如果这样的映射我们记作 φ_A 的话, 那么 φ_A 是线性映射吗? 更具体的, 我们有如下的图示意:

$$\begin{tikzcd} V^n \arrow[r, "\varphi"] \arrow[d, "\simeq"'] & U^m \arrow[d, "\simeq"'] \\ \mathbb{F}^{n\times 1} \arrow[r, "\simeq", "\hat{\varphi}"'] & \mathbb{F}^{m\times 1} \end{tikzcd}$$

其中 $\hat{\varphi}$ 一定形如某个 φ. 下面我们来问几个问题:

1. $AX=\beta$ 有解吗? 答案与如下的内容等价:

   (a) $r(A)=r(A|\beta)$

   (b) β 可以被 A 的列向量线性表出.

   就像在引理4.2中的那样, 我们可以考虑把映射作用在空间里面所有可能矩阵之后的**像**放在集合里面, 称为**像集**, 也就是 $\mathrm{Im}\varphi_A=\{\varphi_A(x)|x\in\mathbb{F}^{n\times 1}\}$, 类似于在函数中自变量跑遍所有可能的取值, 观察因变量的取值, 称为**值域**一样.

2. 求 $AX=0$ 的解.

   事实上, 这个解集是十分重要的. 也就是求 $\{X\in\mathbb{F}^{n\times 1}|\varphi_A(X)=0\}$, 我们很多时候把它叫做这个映射的**核**. 用记号 $\ker\varphi_A$ 来表示上面的集合.
   为什么核集合如此重要呢? 因为右侧从等于 0 到等于 β 是需要一点点变动, 也就是一个**特解**. 这样我们就可以得到所有的解集了.

3. 求特解 η_0, 也就是 $A\eta_0=\beta$, 这样一来, $Ax=\beta$ 的所有解就都可以写作 $\eta_0+\eta,\eta\in\ker\varphi_A$.

Example 4.8. (一个奇怪的例子, 和 Lie 群有关)

Example 4.9. (Lagrange 插值公式)

Example 4.10. (求导)

4.4 子空间

4.4.1 基本定义

在上一节的例子中, 我们其实从求解线性方程组发现了一个比较一般的内容:

Example 4.11. $A\in\mathbb{F}^{m\times n},\beta\in\mathbb{F}^{m\times 1}$, 求线性方程组 $AX=\beta$ 的步骤是:

1. 探讨解的存在性;

2. 寻找 $AX=0$ 的解空间 W

3. $AX=0$ 的解空间为 $n-r(A)$

4. $AX=\beta$ 的解集是 $\{\eta_0+\alpha|\alpha\in W\}$, 并且有性质 $\{\eta_0+\alpha|\alpha\in W\}\underset{-\eta_0}{\overset{+\eta_0}{\rightleftharpoons}}W$.

我们在使用线性映射的观点去看整个内容的时候发现这件事情是很有价值的. 于是我们给出映射的核与像的定义:

Definition 4.12. (映射的核与像)

一个自然的问题就是, 一个映射的核上面有什么的性质? 更一般的, 线性映射的核与像是不是线性空间?

Lemma 4.3. *如果 $\varphi\in Hom(V,U)$, $\ker\varphi, Im\varphi$ 都是线性空间.*

从上面的证明我们可以感受到, $\ker\varphi,\mathrm{Im}\varphi$ 继承且保留了 V,U 之间的运算及其性质. 并且我们还可以感受到, $\mathrm{Ker}\varphi\subset V,\mathrm{Im}\varphi\subset U$. 于是我们给出如下的定义:

Definition 4.13. 若 $V/\mathbb{F}$, $W\subseteq V$ 是非空子集, 若 W 中对于 V 的加法, 数乘也构成线性空间, 则称 W 为 V 的子空间.

在理解这个定义的时候要注意如下几点:

1. 这个集合要是非空的

2. 这个集合要继承下来 W 中对于加法, 数乘的定义, 并且对加法, 数乘具有封闭性.

Proposition 4.2. *设 W 是 $V/\mathbb{F}$ 的非空子集, 则下列条件等价:*

- *W 为 V 的子空间*
- *W 对 V 中的加法和数乘封闭*
- $\forall\alpha,\beta\in W,k,l\in\mathbb{F},k\alpha+l\beta\in W$

那么, 有没有最小的子空间呢? 其实 $\{0\}$ 永远是 V 平凡的子空间. 下面同样给出几个空间包含的例子.

Example 4.12. (平面和三维空间)

Example 4.13. (多项式空间)

Example 4.14. (连续函数的空间)

那么, 核空间和像空间之间有什么联系呢? 还是看我们的线性方程组的例子开始. 对于一个线性方程组, $\dim\ker\varphi_A=n-r(A)$, $\dim\mathrm{Im}\varphi_A=r(A)$, 而把它们的维数加起来就有和 $\mathbb{F}^{n\times 1}$ 是等价的. 也就是 $\dim\ker\varphi_A+\dim\mathrm{Im}\varphi_A=\dim V$. 下面我们来给出这个内容的一个证明.

?proofname? TODO □

更加广阔的, 我们是不是可以证明

Proposition 4.3. *若 $\varphi : V \to U$ 是线性映射, $\dim \ker \varphi_A + \dim Im\varphi_A = \dim V$, 当 V 是有限维线性空间的时候.*

这个命题我们会在4.4.2和4.5.1两节给出不同角度的证明.

4.4.2 交与和

有了子空间, 我们自然希望寻求一组子空间的基底. 首先我们可能会猜测: 子空间的基也可能由原空间 "继承" 下来.

Problem 4.3. 设 $\dim V = n$,W 为 V 的子空间, W 的一组基是 $\alpha_1..\alpha_k$. 设 $q_1..q_n$ 是 V 的一组基, 则取 $\{q_1..q_n\} \cap W$ 是 W 的一组基吗?

答案当然不是. 在平面上可以找到反例.

Example 4.15. (平面上子空间)

因此这告诉我们应当从小空间往大找, 而不是大空间往小的找. 具体的, 我们有这个定理:

Theorem 4.5. *设 $\dim V = n < \infty$, W 为 V 的子空间, W 的一组基可以扩充为 V 的一组基.*

证明的思路大致是取 V 的一组基,$\beta_1..\beta_n$, 将 $\alpha_1..\alpha_k\beta_1..\beta_n$ 放在一起, 然后做筛选法就可以了.

Problem 4.4. (思考题) 若 V 是无穷维的, 子空间可以基的扩充的方法做吗?

有了这样的想法, 我们就可以来证明命题4.3($\dim \ker \varphi_A + \dim \mathrm{Im}\varphi_A = \dim V$). 证明如下:

?proofname? 取 $\ker \varphi$ 的一组基 $\alpha_1..\alpha_k$, 扩充一下成为 V 的一组基 : $\alpha_1..\alpha_k..\alpha_n$. 其中 $\varphi(\alpha_{k+1})..\varphi(\alpha_n)$ 是 $\mathrm{Im}\varphi$ 的生成元 (由它们线性组合得到 $\mathrm{Im}\varphi$ 的元素, 记作 $L(\varphi(\alpha_{k+1})..\varphi(\alpha_n))$.)

问题转换为需要证明 $\dim \mathrm{Im}\varphi = n - k \Leftrightarrow \varphi(\alpha_{k+1})..\varphi(\alpha_n)$ 是线性无关的. 也就是 $l_{k+1}\varphi(\alpha_{k+1}) + .. + l_n\varphi(\alpha_n) = 0$ 是不是可以说明 $l_{k+1}..l_n$ 全为 0? 于是给出如下的证明:

$l_{k+1}\varphi(\alpha_{k+1}) + .. + l_n\varphi(\alpha_n) = 0$ 意味着 $l_{k+1}\alpha_{k+1} + .. + l_n\alpha_n = 0 \in \ker \varphi$, 也就是 $l_{k+1}\alpha_{k+1} + .. + l_n\alpha_n = l_1\alpha_1 + .. + l_k\alpha_k$, 移动项有: $l_{k+1}\alpha_{k+1} + .. + l_n\alpha_n - (l_1\alpha_1 + .. + l_k\alpha_k) = 0$, 由于 $l_1..l_k$ 的线性无关性, 可以知道 $\varphi(\alpha_{k+1})..\varphi(\alpha_n)$ 是线性无关的. 因此是 $\mathrm{Im}\varphi$ 中的一组基. □

接下来我们来看线性映射的全体之间单射、满射之间的关系. 在这之前, 我们再次引入一下下列记号;

Definition 4.14. (Hom 记号) 设 $\varphi : V \to U, \mathrm{Hom}(V/\mathbb{F}, U/\mathbb{F})$ 是所有从 V 到 U 的线性映射的全体. 可记作 $\varphi \in \mathrm{Hom}(V/\mathbb{F}, U/\mathbb{F})$.

如何让这个线性映射是满射? 其实很简单, 只要把 U 这个集合缩小到 $\mathrm{Im}\varphi$ 就行了.

那如何让这个线性映射是一个单射呢? 我们希望只要缩小 V 这个集合就行了. 具体的, 只要让每一个被映过去的元素只有一个就行了. 于是我们可以找到一个 V 的子集 W. 可以验证这个 W 也是 V 的一个子空间.

$$\begin{tikzcd}
\ker\varphi \arrow[r, "\text{单射}"] & V/\mathbb{F} \arrow[r, "\varphi"] \arrow[dr, "\varphi"] & U/\mathbb{F} \\
 & W \arrow[u, "\text{单射}"] \arrow[r, "\varphi|_W"'] & \mathrm{Im}\varphi
\end{tikzcd}$$

上面这张图里面的下面那一条线就得到了一个新的映射. 一般叫做作**限制映射**. 如果我们考察这个映射的核空间的话, 也就是 $\ker\varphi|_W = W \cap \ker\varphi$, 会发现这个集合里面只有 0, 并且 $\varphi(W) = \mathrm{Im}\varphi = \varphi(V)$ 下面我们来说明这件事情.

Proposition 4.4. 限制映射 $\varphi|_W$ 的核空间与空间 W 的交只有 $\{0\}$, 且 $\varphi(W) = Im\varphi = \varphi(V)$.

?proofname? $\forall \alpha \in V$, $\exists \beta \in W$, 使得 $\varphi(\alpha) = \varphi(\beta)$, $r := \varphi(\alpha - \beta) = 0 \implies \alpha - \beta \in \ker\varphi$. □

从上面的证明和命题中我们可以看出任意的一个 $\alpha \in V$ 可以被分解为 $\beta \in W$ 和 $\gamma \in \ker\varphi$. 这个在后续会用到.

下面来给出子空间的交的定义:

Definition 4.15. (子空间的交) 设 V_1, V_2 为 V 的子空间, 称 $V_1 \cap V_2$ 为 V_1 与 V_2 的交.

Definition 4.16. (子空间的和) 设 $W_1, W_2 \subset V, W_1 + W_2 = \{\alpha_1 + \alpha_2 | \alpha_1 \in W_1, \alpha_2 \in W_2\}$.

我们在比较熟悉的地方来举几个例子:

Example 4.16. (平面上子空间的交)

Example 4.17. (两个子空间的维数和其和维数之间的关系)

4.4.3 直和

有了子空间的和的问题之后, 我们自然要考虑子空间的和和子空间的并是一回事吗? 它们的区别是什么?

Definition 4.17. (直和) 设 V_1, V_2 都是 V 的子空间, 则下列条件等价:

1. $V_1 \cap V_2 = \{0\}$

2. $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$
3. V_1 的一组基与 V_2 的一组基可以合并成 $V_1 + V_2$ 的一组基
4. $\forall \alpha \in V_1 + V_2$, $\alpha_1 \in V_1, \alpha_2 \in V_2$, α 的分解式唯一.
5. 零向量的分解式唯一

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

$$V_1 \oplus V_2.$$

最后我们用矩阵的例子来看一看 "直和".

Example 4.18. (矩阵的分类与直和)

4.5 商空间

4.5.1 准备工作

在开始之前, 请回顾一下直和的定义 (定义4.17) 和维数公式 (命题4.3).

其实命题4.3的维数公式是一个特殊情形, 我们可以给出如下的维数公式:

Theorem 4.6. *(维数公式)* $\dim(V_1 \cap V_2) + \dim(V_1 + V_2) = \dim V_1 + \dim V_2$.

?proofname? TODO □

有时候我们可能需要空间的一个补集, 于是我们需要如下的定义:

Definition 4.18. (补子空间) V_1 是的 V 子空间, 如果存在 V_2, 使得 $V_1 \oplus V_2 = V$, 称 V_2 为 V 的**补子空间**.

下面我们来通过一个实际的例子来看一看以前我们的线性方程组的思想 (例子4.11) 是不是可以被抽象出来.

回顾左乘 A 矩阵的例子实际上是 $\varphi_A : \mathbb{F}^{n\times 1} \to \mathbb{F}^{m\times 1} \ni \beta$, 那么步骤大概是这样的

1. $AX = \beta$ 有解 $\iff$ $\beta \in \mathrm{Im}\varphi_A$
2. 若有解, 先求特解 $\eta_0, A\eta_0 = \beta$, 要求任意解, 必须先求 $AX = 0$ 的解, 也就是 $\ker \varphi_A$.
3. $AX = \beta$ 的解集为 $\eta_0 + \ker \varphi_A$
4. 能否选取特解 η, 使得特解的全体构成 $\mathbb{F}^{n\times 1}$ 的子空间?

不妨设 $AX = \beta_0$ 对应的解空间为 $\eta_0 + W$, $AX = \beta_1$ 对应的解空间为 $\eta_1 + W$. 那么 $AX = \beta_0 + \beta_1$ 对应的解空间为 $\eta_0 + \eta_1 + W$, $AX = k\beta_0$ 的解空间为 $k\eta_0 + W$. 这是一件很好玩的事情.

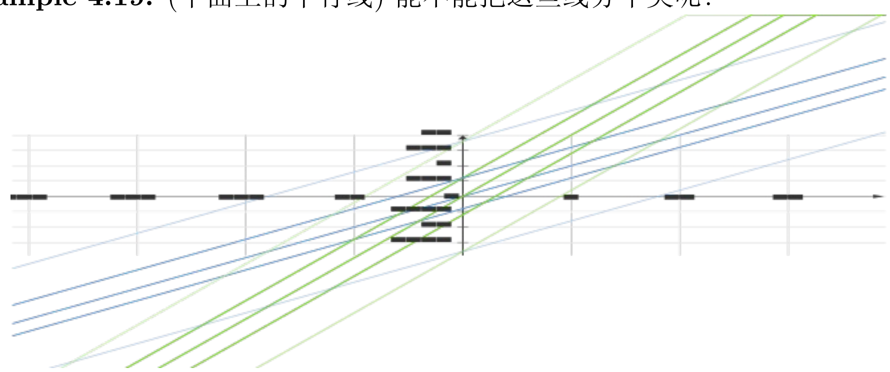
再继续之前, 先来看看我们都玩了些什么.

$$
\begin{tikzcd}
\text{一些例子} \arrow[r] \arrow[d] & \text{线性空间} \arrow[r, "\text{同构}"] \arrow[d] & \mathbb{F}^{n\times 1} & \\
\text{齐次线性方程组}AX=0 \arrow[r] \arrow[d] & \text{基与维数} \arrow[r] & \text{坐标} \arrow[u] & \\
\text{线性方程组}AX=\beta \arrow[r] \arrow[dd] & \text{线性映射} \arrow[r] \arrow[dr] & \text{核} \arrow[dr] & \\
 & & \text{像} \arrow[r] & \text{子空间} \\
\text{特解与通解} & & &
\end{tikzcd}
$$

4.5.2 商空间

商空间是从解线性方程组里面出来的. 加入 $A\eta_0 = \beta$ 成立, 且 η_0 是特解, 那么所有解就会形如 $\eta_0 + W = \{\eta_0 + \alpha | \alpha \in W\}$, $AX = 0$ 的解空间记作 W. 这样的情形就像我们在平面上考虑一些平行线一样.

Example 4.19. (平面上的平行线) 能不能把这些线分个类呢?



还是从映射的角度考虑. 假设考虑 $V(\mathbb{F}^{n\times 1}) \to U(\mathbb{F}^{m\times 1}), (m < n)$ 的映射, 恰当的选取元素, 使得这是一个满射. 因为这样比较好考虑.

$$
\begin{array}{ccc|ccccc}
\varphi : V \to U\text{是满射} & & (\varphi_A : & \mathbb{F}^{n\times 1} & \to & \mathbb{F}^{m\times 1} & (m<n) &) \\
 & & \uparrow\in & \varphi_A^{-1}(\beta) & \to & \beta & & \\
\text{有原象}\varphi^{-1}(\beta_1) & \to & \text{给定}\beta_1 & \text{有原像}\varphi_A^{-1}(\beta) & \to & \text{给定}\beta & & \\
\text{有原象}\varphi^{-1}(\beta_2) & & \text{给定}\beta_2 & (\text{也就是} & Ax=B & \text{的解集}) & & \\
 & & & & & & &
\end{array}
$$

注意到每一个 β 总是对应**唯一的**一个原像. 并且原像构成的集合是 V 的一个子集. 因为不同的常数项在右边, 导致方程组的解集是不一样的. 也就是说

1. 若 $\beta_1 \neq \beta_2$, $\varphi^{-1}(\beta_1) \cap \varphi^{-1}(\beta_2) = \varnothing$.

2. $\forall \alpha \in V, \exists \beta$, 使得 $\alpha \in \varphi^{-1}(\beta)$. 只需要取得 $\beta = \varphi(\alpha)$ 即可 (取成 0).

3. $V = \bigcup_{\beta \in U} \varphi^{-1}(\beta)$, (由 1 是不交并).

现在我们希望这个集合是单射, 因此, 不妨构造这样的一个集合 X, 使得把映成同样元素的内容放在一起看. 这样构造的集合 X 与 U 之间的元素是一一对应的. 也就是

$$X = \left\{\varphi^{-1}(\beta) | \beta \in U\right\} \subseteq \mathsf{PSET}(U)$$

$\mathsf{PSET}(U)$ 是 U 的幂集. 也就是 U 的所有子集构成的集合, 比如 $\mathsf{PSET}(\{1,2,3\}) = \{\{1\},\{2\},\{3\},\{1,2\},\{2,3\},\{1,3\},\{1,2,3\}\}$. 相关内容可以参看离散数学集合论相关的内容.

那么这个集合性质很好, 我们能不能把 U 上的加法和数乘**平移到 X 上面去**?

$$\begin{tikzcd}
\text{不知道怎么加} \rightarrow & \varphi^{-1}(\beta_1) \arrow[d, "\varphi"] & + & \varphi^{-1}(\beta_2) \arrow[d, "\varphi"] & \varphi^{-1}(\beta_1+\beta_2) \\
\text{映射到这里} \rightarrow & \beta_1 & + & \beta_2 & \beta_1+\beta_2 \in U \arrow[u, "\varphi^{-1}", "\text{映射回来}"']
\end{tikzcd}$$

从上, 可以得到

- X 是 $\mathbb{F}$ 上面的线性空间.

这个 X 和 U, V 有什么关系呢?

用 V 中的运算看 X 的运算:
用原象的记号表达可能并不能让我们了解到底有什么联系. 我们具体的写一下:
考虑这个线性空间中的一个特殊的元素零向量, 那么 $\varphi^{-1}(0) = \ker \varphi$.
令 $W = \ker \varphi \subset V$, 若 $\varphi(\alpha) = \beta$, 考虑 $\varphi^{-1}(\beta) = \alpha + \gamma | \gamma \in W$. 也就是可以通过特解和核空间里面的通解加起来得到. 简单记作 $\alpha + W$.
如果 $\varphi(\alpha_1) = \beta_1, \varphi(\alpha_2) = \beta_2$, 由线性空间的性质, $\varphi^{-1}(\beta_1) + \varphi^{-1}(\beta_2) = \varphi^{-1}(\beta_1 + \beta_2)$, 有 $\varphi^{-1}(\beta_1) = \alpha_1 + W, \varphi^{-1}(\beta_2) = \alpha_2 + W, \varphi^{-1}(\beta_1) + \varphi^{-1}(\beta_2) = \alpha_1 + \alpha_2 + W$.

由此我们可以发现:

Theorem 4.7. *若 $\overline{\alpha} \cap \overline{\beta}$ 非空, 则有下列条件:*

1. $\bar{\alpha} = \bar{\beta}$

2. $\alpha - \beta \in W$

3. $\alpha \in \bar{\beta}$

4. $\beta \in \bar{\alpha}$

我们可以重新考虑一下这个子空间.

$W \subset V$ 的子空间:
$\forall \alpha \in V$, 考虑**子集** $\alpha + W = \alpha + \gamma | \gamma \in W := \bar{\alpha}$. 这样给定一个子空间之后, 通过改变 γ 的值, 就可以找到很多的子集. V 也一定是这些子集的并集. 我们来考虑 $\alpha, \beta \in V$ 之间的元素的关系. 也就是

$$\left.\begin{array}{ccccc} V & \ni & \alpha & \to & \bar{\alpha} := \alpha + W \\ V & \ni & \beta & \to & \bar{\beta} := \beta + W \end{array}\right\} \bar{\alpha} \cap \bar{\beta} = ?$$

假设存在 $\delta = \alpha + \gamma_0 = \beta + \gamma_1$, 那么有 $\alpha - \beta = \gamma_1 - \gamma_0 \in W$, $\alpha = \beta + \gamma \in \beta + W$; $\beta = \alpha - \gamma \in \alpha + W$. 得到若 $\bar{\alpha} \cap \bar{\beta}$ 非空, 那么 $\alpha = \beta$. 这样会被称作广义的平行.

在上述的这个子空间里面, 也有如下的一些性质可以满足:

$W \subset V$ 的子空间:
记 $V/W = \{\alpha + W | \alpha \in V\}$, 那么满足

1. 加法: $\alpha + W + \beta + W = \alpha + \beta + W$
2. 数乘: $(\alpha + W)k = k\alpha + W$.

也就是简单记作 $\bar{\alpha} + \bar{\beta} = \overline{\alpha + \beta}, k\overline{\alpha} = \overline{k\alpha}$.
刚刚的内容让我们选取了一个代表元进行操作. 那么, 商空间的性质是不是和代表元的选取有关呢? 若更换 $\alpha \leftarrow \alpha_1, \beta \leftarrow \beta_1$, 那么 $\overline{\alpha_1 + \beta_1} = \overline{\alpha + \beta}$ 成立吗? 如果要让定义合理的话, 这个要成立才可以.

在看到之前的问题的时候, 考虑如下的映射:

$$V \xrightarrow{\pi} V/W$$

$$\alpha \longrightarrow \overline{\alpha} = \alpha + W$$

并且记 π 作为**自然映射**. 那么这个映射有这些的性质:

1. π 是满射
2. $\pi(\alpha + \beta) = \overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta} = \pi(\alpha) + \pi(\beta)$
3. $\pi(k\alpha) = \overline{k\alpha} = k\overline{\alpha} = k\pi(\alpha)$
4. $\ker \varphi = W$
5. 取 $V_1 \subset V$ 是 W 的补子空间, 使得 $\pi|_{V_1} : V_1 \to V/W$ 是线性同构. 因为 $V_1 \cap \ker \pi = 0, V_1 + \ker \pi = V$, 意味着 $W \oplus V_1 = V$.
6. $V/W \simeq V_1 \implies \dim V/W = \dim V - \dim W$. 注意: V/W 不是 V 的子空间.

7. 找到 V/W 的一组基 $\iff$ 找到 W 的子空间的一组基, 记作 $\alpha_1..\alpha_k$. 扩充为 V 的一组基: $\alpha_1..\alpha_k\alpha_{k+1}..\alpha_n$, 则 $L(\alpha_{k+1}..\alpha_n)$ 为 W 的补子空间.

Proposition 4.5. 若 $\pi(\alpha_1)=\overline{\alpha_1},\pi(\alpha_2)=\overline{\alpha_2},...,\pi(\alpha_k)=\overline{\alpha_k},\pi(\alpha_{k+1})=\overline{\alpha_{k+1}}..,\pi(\alpha_n)=\overline{\alpha_n}$, $\overline{\alpha_1}..\overline{\alpha_n}$ 是商空间 V/W 的基.

Proposition 4.6. 若 $\overline{\alpha_1}..\overline{\alpha_k\alpha_{k+1}}..\overline{\alpha_n}$, $\overline{\alpha_1}..\overline{\alpha_k}$ 为 W 的一组基, 若 $\alpha_1..\alpha_k$ 是 W 的一组基, 那么 $\alpha_1..\alpha_k\alpha_{k+1}..\alpha_n$ 是 V 的一组基.

下面来说明基是不依赖于代表元的选取.

(TODO)

Example 4.20. (在数列极限中的应用)

Definition 4.19. (等价关系)

4.6 线性代数中出现的内容

4.6.1 向量组以及其线性组合

1. n 维向量
   (a) 定义: n 个数 $a_1,a_2,\cdots,a_n$ 构成的有序数组 $(a_1,a_2,\cdots,a_n)^T$ 称为一个 n 维向量. 这 n 个数称为该向量的分量, 第 i 个数 a_i 称为这个向量的第 i 个分量.
   (b) 运算:
      i. 加法 $(a_1,a_2,\cdots,a_n)+(b_1,b_2,\cdots,b_n)=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$
      ii. 数乘 $k(a_1,a_2,\cdots,a_n)=(ka_1,ka_2,\cdots,ka_n),k\in\mathbb{F}$.
2. 线性组合: 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\in\mathbb{F}^n$,$k_1,k_2,\cdots,k_n\in\mathbb{F}$, 那么 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n$ 称为向量组 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 的一个线性组合.
3. 线性表示: 设 $\boldsymbol{\beta}\in\mathbb{R}^n$, 若**存在**一组数 $k_1,k_2,\cdots,k_n$, 使得 $\beta=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n$, 那么称向量 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}$ **线性表示** (**线性表出**).
   (a) Th. 向量 β 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示的充分必要条件是矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 和矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta})$ 的秩相等.
4. 向量组的等价: 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是两个向量组, 若向量组 $\boldsymbol{A}$ 的**每个**向量都可以由向量组 $\boldsymbol{B}$ 线性表示, 则称向量组 $\boldsymbol{A}$ 可由向量组 $\boldsymbol{B}$ 线性表示. 若向量组 $\boldsymbol{A}$ 和向量组 $\boldsymbol{B}$ 可以**互相线性表示**, 则称向量组 $\boldsymbol{A}$ 和向量组 $\boldsymbol{B}$ **等价**.
   (a) Prop. 这是一个**等价关系**.
      i. Coll. 向量组 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 等价的充要条件是 $\mathrm{rk}(\boldsymbol{A})=\mathrm{rk}(\boldsymbol{B})=\mathrm{rk}(\boldsymbol{A},\boldsymbol{B})$.

4.6.2 向量组的线性相关性

1. 线性相关与线性无关: 给定向量组 $\boldsymbol{A}:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n(n\geq 1)$, 如果**存在**一组**不全为 0** 的数 $k_1,k_2,\cdots,k_n$, 使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}$, 则称向量组 $\boldsymbol{A}$ 线性相关, 否则称向量组线性无关.
   (a) 注: 这个命题的否定是 $(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0})\Longrightarrow(\forall k_1,k_2,\cdots,k_n=0)$. 可以参看离散数学中的命题逻辑一节学习.
2. 若干结论:
   (a) 单独的一个非零向量线性无关
   (b) 如果一个向量组线性无关, 那么它的任何一个非空部分组也线性无关
   (c) 若 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 组线性无关, 那么在相同的位置都添加一个分量之后得到的 $n+1$ 维的向量组也线性无关.
   (d) 若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关, 但是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}$ 线性相关, 则向量 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 唯一线性表出.
   (e) 如果一个向量组的一部分线性相关, 那么这个向量组线性相关
   (f) 如果向量组 $(\circledast)\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 可由向量组 $(\clubsuit)\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性表示, 且 $\circledast$ 线性无关, 那么 $\circledast$ 所含的向量个数不大于 $\clubsuit$ 含的像两个数. 也就是 $n\leq r$.
   (g) 一个向量组线性相关 $\Leftrightarrow$ 它所构成矩阵的秩**小于**向量个数; 一个向量组线性无关 $\Leftrightarrow$ 它所构成矩阵的秩**等于**向量个数.

4.6.3 向量组的秩

1. 最大线性无关组
   (a) Def. 给定向量组 $\boldsymbol{A}$, 若存在 $\boldsymbol{A}$ 的一个**部分组** $\boldsymbol{B}$:$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$, 满足以下的两个条件:
      - 向量组 $\boldsymbol{B}$ **线性无关**
      - 向量组 $\boldsymbol{A}$ 中的**任意** $s+1$ 个向量**都能**由向量组 $\boldsymbol{B}$ 线性表示

      则称向量组 $\boldsymbol{B}$ 是向量组 $\boldsymbol{A}$ 的一个极大线性无关组.
   (b) Prop.
      i. 向量组与其最大线性无关组等价.
      ii. 一个向量组的最大线性无关组包含的个数是唯一的
2. 向量组的秩
   (a) Def. 向量组 $\boldsymbol{A}$ 的最大线性无关组包含的向量个数称为向量组 $\boldsymbol{A}$ 的秩. 规定只有零向量的向量组的秩为 0.

3. 矩阵的行秩和列秩

   (a) 分块: 对于矩阵 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{sm} \end{pmatrix}$, 我们可以按照行来分块:

$$\begin{matrix} \boldsymbol{\alpha}_1 & \rightarrow & a_{11} & a_{12} & \cdots & a_{1n} \\ \boldsymbol{\alpha}_2 & \rightarrow & a_{21} & a_{22} & \cdots & a_{2n} \\ \boldsymbol{\alpha}_{\cdots} & \rightarrow & \vdots & \vdots & & \vdots \\ \boldsymbol{\alpha}_s & \rightarrow & a_{s1} & a_{s2} & \cdots & a_{sm} \end{matrix}$$

   形成列向量: $\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_s \end{pmatrix}$.

   也可以按照行进行分块:

$$\begin{matrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_n \\ \downarrow & \downarrow & \downarrow & \downarrow \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{sm} \end{matrix}$$

   形成行向量:$\begin{pmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_n \end{pmatrix}$.

   (b) 行 (列) 秩: $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 的秩称为 A 的**行秩**,$\boldsymbol{A}$ 的列向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$ 的秩称为**列秩**.

4.6.4 线性方程组的解的结构

1. n 元线性方程组的矩阵表示:

   记矩阵 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{sm} \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix}, x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, 那么线性方程组 $Ax = b$ 称为 n 元线性方程组.

   (a) 非齐次线性方程组: $\boldsymbol{Ax} = \boldsymbol{b}(\boldsymbol{b} \neq \boldsymbol{0})$, $\boldsymbol{A}$ 称为方程的系数矩阵, $\boldsymbol{b}$ 为常数项矩阵, $\boldsymbol{x}$ 为未知量矩阵.

      i. $\boldsymbol{Ax} = \boldsymbol{b}$ 有解 $\Leftrightarrow$ 系数矩阵 $(\boldsymbol{A})$ 和增广矩阵 $(\boldsymbol{A}|\boldsymbol{b})$ 的秩相等

(b) 齐次线性方程组: $\boldsymbol{Ax}=\mathbf{0}$, 零向量 $\mathbf{0}$ 是齐次线性方程组的一个解, 称为**零解**. 如果存在一个另外的解, 但是这个解不是 $\mathbf{0}$, 那么称为这个解为**非零解**.

i. $\boldsymbol{Ax}=\mathbf{0}$ 有非零解 $\Leftrightarrow$ 系数矩阵的秩小于等于 n

2. 解空间

(a) Def. 设 S 是齐次线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的所有解向量组成的集合, 也就是 $S=\{\boldsymbol{Ax}=\mathbf{0}\mid \boldsymbol{x}\in\mathbb{R}^n\}$, S 被称为齐次线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的解空间.

3. 基础解系

(a) 齐次线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的一组解 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_t$ 称为基础解系, 如果如下的两条同时满足:

- $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_t$ 线性无关
- $\boldsymbol{Ax}=\mathbf{0}$ 的任意一个解都能由 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_t$ 线性表示.

(b) $\boldsymbol{Ax}=\mathbf{0}$ 的基础解系就是其解空间的最大线性无关组

4. 通解

(a) 设齐次线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 有非零解, 其基础解系是 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$, 那么 $\xi=c_1\boldsymbol{\xi}_1+c_2\boldsymbol{\xi}_2+\cdots+c_n\boldsymbol{\xi}_n$, c_i 是任意常数. 表示了 $\boldsymbol{Ax}=\mathbf{0}$ 的全部的解, 称为 $\boldsymbol{Ax}=\mathbf{0}$ 的**通解**.

5. 非齐次线性方程组

(a) 导出组:$\boldsymbol{Ax}=\mathbf{0}$ 称为非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的导出组.

i. 非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的任意两个解是它的导出组的解.

ii. $\boldsymbol{Ax}=\boldsymbol{b}$ 的一解和它的导出组的一个解的和是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解.

(b) 解的结构: 如果 $\boldsymbol{\xi}_0$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个解 (称为**特解**), 那么非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的任一解向量可以表示为 $\xi=\xi_0+k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}$, 其中 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}$ 是导出组 $\boldsymbol{Ax}=\mathbf{0}$ 的基础解系, $k_1,k_2,\cdots,k_{n-r}$ 是任意实数.

4.6.5 向量空间

1. Def. 设 V 为 n 维向量的集合, 如果集合非空, 却集合 V 对于加法和数乘两种运算封闭, 则称集合 V 为向量空间.

2. r 维向量空间: 设 V 为向量空间, 如果 r 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in V$, 且满足

- $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ **线性无关**
- V 中**任意**一个向量**都能**由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表示

那么 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 就是空间 V 的一组基, r 是 V 的维数, V 为 r 维向量空间.

4.7 基础习题

第五章　线性变换

5.1 线性变换的基本概念与矩阵

5.1.1 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

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

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

如果 V 是 $\mathbb{F}$ 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 $\mathrm{End}(V)$.

Definition 5.1. 设 V 是 $\mathbb{F}$ 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 $\alpha,\beta \in V, k,l \in \mathbb{F}$, 都有

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

下面来举一些例子:

Example 5.1. 数乘变换: TODO

Example 5.2. 平面上的旋转变换: TODO

Example 5.3. 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 $\mathbb{F}$ 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
  - $k \in \mathbb{F}, \mathscr{A} \in \mathrm{End}V$, $(k\mathscr{A})(\alpha) = k(\mathscr{A}\alpha) \in \mathrm{End}V$.

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 $f(\mathscr{A})=a_i\alpha^i$.

Proposition 5.1. *(线性变换的运算律)*

5.1.2 与 $\mathbb{F}^{n\times n}$ 的等价性

以前我们知道, 线性空间实际上就是和 $\mathbb{F}^{n\times 1}$ 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 $V/\mathbb{F}$ 是 n 维线性空间, $\alpha_1..\alpha_2$ 是 V 的一组基, $V\xrightarrow{\simeq}\mathbb{F}^{n\times 1}$.

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

$$||$$

$$\sum x_i a_i=(\alpha_1..\alpha_n)X,\qquad X=(x_1..x_n)'$$

并且

$$\begin{aligned}\mathscr{A}\alpha=&\alpha\left(\sum_{i=1}^n x_i\alpha_i\right)\\=&\sum_{i=1}^n x_i\mathscr{A}\alpha_i\\=&\mathscr{A}\left((\alpha_1..\alpha_n)X\right)\\=&(\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)X\end{aligned}$$

也就是 $\mathscr{A}$ 是由 $\mathscr{A}\alpha_1..\mathscr{A}\alpha_n$ 决定的.

于是我们就得到了如下的内容:

如果 $\alpha_1..\alpha_n$ 是一 $V/\mathbb{F}$ 的一组基, 也就是

$$\mathrm{End}V \ni \quad \mathscr{A} \xrightarrow{\text{选择基底}\alpha_1..\alpha_n} (\mathscr{A}\alpha_1,..,\mathscr{A}\alpha_n) \longrightarrow \{\mathbb{F}^{n\times 1}\text{中}n\text{元组}\} \longrightarrow \mathbb{F}^{n\times n}$$

$$\mathscr{A} \xrightarrow{\text{选择基底}\alpha_1..\alpha_n} (\mathscr{A}\alpha_1,..,\mathscr{A}\alpha_n) \xrightarrow{\text{用基底}\alpha_1..\alpha_n\text{表示}} A$$

也就是完成了 $\mathrm{End}V \to \mathbb{F}^{n\times n}$, 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:

Theorem 5.1. *设 $\alpha \in V, \mathscr{A} \in \mathrm{End}V$, 且 $\alpha_1..\alpha_n$ 是 V 的一组基, 在这组基下, $\mathscr{A}$ 的基是 A, α 的坐标是 X, 那么 $\mathscr{A}\alpha$ 的坐标是 AX, 也就是 $\mathscr{A}\alpha = (\alpha_1..\alpha_n)AX$.*

?proofname? 用形式记号的语言, 有

$\mathscr{A}\alpha = \mathscr{A}((\alpha_1..\alpha_n)X) = (\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)X = ((\alpha_1..\alpha_n)A)X$, 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 $(\alpha_1..\alpha_n)(AX)$ 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 $\alpha_1..\alpha_n$ 是 V 的一组基, 那么从 $\mathrm{End}V$ 到 $\mathbb{F}^{n\times n}$ 的映射 $\mathscr{A} \to \varphi(\mathscr{A})$ 有什么性质?

既然是线性空间上面的线性映射, 我们猜测:

- φ 是单射: 考虑 $\mathscr{A} \xrightarrow{\alpha_1..\alpha_n} A, \mathscr{B} \xrightarrow{\beta_1..\beta_n} B$, 要满足单射的条件, 就要考虑从 $A = B$ 推出 $\mathscr{A} = \mathscr{B}$. 要证明右侧的式子, 根据定义, $\forall\alpha, \mathscr{A}\alpha = \mathscr{B}\alpha$. 两边展开, 就有 $\mathscr{A}(\sum_{i=1}^n x_i\alpha_i) = \mathscr{B}(\sum_{i=1}^n x_i\alpha_i)$, 根据定义, 并且把求和记号转化为矩阵乘法, 看出坐标, 就有 $(\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)X = (\mathscr{B}\alpha_1..\mathscr{B}\alpha_n)X$. 又根据已知的条件 $A = B$, 就可以得到 $\mathscr{A}\alpha_i = \mathscr{B}\alpha_i (i \in [1..n])$. 就可以推断出这个结论了.
- φ 是满射: 问题转化为 $\forall A \in \mathbb{F}^{n\times n}$, 是否存在 $\mathscr{A} \in \mathrm{End}V$, 使得 $\varphi(\mathscr{A}) = A$. 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 $\mathscr{A}\alpha = \mathscr{A}(\sum_{i=1}^n x_i\alpha_i) = \sum_{i=1}^n x_i\mathscr{A}\alpha_i \triangleq \sum_{i=1}^n x_i\beta_i$. 这样就可以方便的验证 $\mathscr{A} \in \mathrm{End}V, \varphi(\mathscr{A}) = A$.
- $\varphi(\mathscr{A} + \mathscr{B}) = \varphi(\mathscr{A}) + \varphi(\mathscr{B})$. 只需要转化成方阵的形式证明就行了.
- $\varphi(k\mathscr{A}) = k\varphi(\mathscr{A})$, 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 $\mathbb{F}^{n\times n}$ 满足结合律, 但不满足交换律. $\mathbb{F}[x]$ 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 φ 还有一个比较重要的性质, 就是它保持了乘法.

Theorem 5.2. 设 $\alpha_1..\alpha_n V$ 的一组基, 则 $\mathrm{End}V$ 到 $\mathbb{F}^{n\times n}$ 的映射 $\mathscr{A}\to\varphi(A)$ 是线性空间 $\mathrm{End}V\to\mathbb{F}^{n\times n}$ 的线性同构, 并且 $\forall\mathscr{A},\mathscr{B}\in\mathrm{End}V,\varphi(\mathscr{A}\mathscr{B})=\varphi(\mathscr{A})\varphi(\mathscr{B})$.

?proofname? 线性同构见上述描述.

下面证明保持乘法:

$$\begin{aligned}(\mathscr{A}\mathscr{B})(\alpha_1..\alpha_n)=&((\mathscr{A}\mathscr{B})\alpha_1..(\mathscr{A}\mathscr{B})\alpha_n)\\=&(\mathscr{A}(\mathscr{B}\alpha_1)..\mathscr{A}(\mathscr{B}\alpha_n))\\=&\left(\mathscr{A}\left((\alpha_1..\alpha_n)B_1\right)..\mathscr{A}\left((\alpha_1..\alpha_n)B_n\right)\right)\\=&((\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)B_1..(\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)B_n)\\=&((\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)B\\=&((\alpha_1..\alpha_n)A_n..(\alpha_1..\alpha_n)A_n)B\\=&((\alpha_1..\alpha_n)A)B\\=&(\alpha_1..\alpha_n)(AB)\qquad(\text{需要验证结合律})\end{aligned}$$

□

5.2 线性变换的矩阵

从上一节的内容我们知道 $\mathrm{End}V\xrightarrow[\alpha_1..\alpha_n]{1-1}\mathbb{F}^{n\times n}$. 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 $\mathscr{A}\mapsto$ 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:

Theorem 5.3. 若 $\mathscr{A}\in\mathrm{End}V$, $\dim V=n$, 下列条件等价:

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

那么 $\mathscr{A}$ 是可逆的.

接着保持乘法 (定理5.2) 的叙述, 我们试着举出一些例子:

Example 5.4. $V=\mathbb{Q}(\sqrt{2})/\mathbb{Q}$, 考虑 $\mathrm{End}V$, 如果满足 $\{\varphi\in\mathrm{End}V\mid\varphi(\alpha\beta)=\varphi(\alpha)\varphi(\beta),\forall\alpha,\beta\in V\}$, 那么 φ 在每一点的取值可以确定吗?

Example 5.5. (推广) 若 $\mathbb{F}\subseteq\mathbb{E}$, 考虑 $\mathrm{End}\mathbb{E}/\mathbb{F}$, 如果 $G=\{\varphi\in\mathrm{End}\mathbb{E}/\mathbb{F}\mid\varphi(\alpha\beta)=\varphi(\alpha)\varphi(\beta),\forall\alpha,\beta\in V,\varphi(1)\neq 0\}$, 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)
- $\varphi^{-1} \in G$(存在逆元)
- 若 $\varphi, \psi \in G, \varphi \circ \psi \in G$.(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.

Example 5.6. 继续考虑上述例子, 如果让 $\mathbb{E} := \mathbb{F}(\alpha_1..\alpha_n)$ 是包含 $\alpha_1..\alpha_n$ 的最小数域, 其中 $\alpha_1..\alpha_n$ 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 $V/\mathbb{F}$, $\alpha_1..\alpha_n$ 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.

$$1, x, x^2, \cdots, x^{n-1} \to \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & n-1 \\ & & & 0 \end{pmatrix}$$

$$1, x, x^2/2!, \cdots, x^n/n! \to \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

$$1, (x-a), (x-a)^2, \cdots, (x-a)^n \to ???$$

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

$$\begin{tikzcd} \mathscr{A} \arrow[r, "\alpha_1..\alpha_n"] \arrow[dr, "\beta_1..\beta_n"'] & A \arrow[d, "\text{过渡矩阵}T?"] \\ & B \end{tikzcd}$$

5.3 相似与特征值

从本节开始, 我们就要试图选取合适的基 $\beta_1,..\beta_n$, 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 $\mathscr{A}(\alpha_1..\alpha_n) = (\alpha_1..\alpha_n)A, B(\alpha_1..\alpha_n) = (\alpha_1..\alpha_n)B$, 并且由过渡矩阵的知识我们有 $(\beta_1..\beta_n) = (\alpha_1..\alpha_n)T \quad (*)$, 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 $\mathscr{A}$, 有

$$\begin{aligned}
\mathscr{A}((\alpha_1..\alpha_n)T) =&((\alpha_1..\alpha_n)T)B \\
(\mathscr{A}\alpha_1..\mathscr{A}\alpha_n)T =&((\alpha_1..\alpha_n)T)B \\
((\alpha_1..\alpha_n)A)T =&((\alpha_1..\alpha_n)T)B \\
AT =&TB \\
B =&T^{-1}AT
\end{aligned}$$

我们称 B 和 A **相似**.

Definition 5.2. (相似)

Theorem 5.4. *如果 $\mathscr{A} \in \mathrm{End}V/\mathbb{F}, \dim V/\mathbb{F} = n$, 则 $\mathscr{A}$ 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 $\mathscr{A}$ 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.

Proposition 5.2. *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5.2的等价关系, 意味着 $\mathbb{F}^{n\times n}$ 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.

Example 5.7. 如果单位变换 $\mathscr{A}$ 在某一组基 $\alpha_1..\alpha_n$ 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 $\beta_1..\beta_n$ 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.

Proposition 5.3. *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 $r(\mathscr{A}) = r$.
- $|A| = |B|$, 同样可以定义线性变换的行列式, 定义作 $|\mathscr{A}| = |A|$.
- $\operatorname{tr}(A) = \operatorname{tr}(B)$, 也同样可以定义线性变换的迹, 定义作 $\operatorname{tr}(\mathscr{A}) = \operatorname{tr}(A)$.
- $B^n = T^{-1}A^nT$, $f(B) = T^{-1}f(A)T$.

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

$$\mathscr{A}(\beta_1..\beta_n) = (\beta_1..\beta_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} = (\lambda_1\beta_1..\lambda_n\beta_n).$$

这样一来, 就要求

$$\begin{cases}\mathscr{A}\beta_1 = & \lambda_1\beta_1 \\ \mathscr{A}\beta_2 = & \lambda_2\beta_2 \\ \cdots & \cdots \\ \mathscr{A}\beta_n = & \lambda_n\beta_n\end{cases}$$

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:

Definition 5.3. (特征值和特征向量)

如何求解特征值呢? 既然我们要求 $\mathscr{A}\alpha = \lambda_0\alpha$, 移项, 有 $\lambda_0\alpha - \mathscr{A}\alpha = (\lambda_0\mathrm{id} - \mathscr{A})\alpha = 0$. 这就意味着 $\ker(\lambda_0\mathrm{id} - \mathscr{A}) \neq 0$, 也就是 $\lambda_0\mathrm{id} - \mathscr{A}$ 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 $|\lambda_0\mathrm{id} - \mathscr{A}| = 0$. 用矩阵的语言来书写就是 $|\lambda_0 I_n - A| = 0$.

既然我们得到了一个变换 $\lambda_0\mathrm{id} - \mathscr{A}$, 这时候我们来看一下这个变换核空间. 也就是 $\ker(\lambda_0\mathrm{id} - \mathscr{A}) = \{\alpha \in V | \mathscr{A}\alpha = \lambda_0\alpha\}$, 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 λ_0 的特征子空间**.

Definition 5.4. (特征子空间)

仔细观察这个属于 λ_0 的特征子空间, 稍微把行列式展开一点, 我们就会发现

$$|\lambda_0 \mathrm{id} - \mathscr{A}| = 0 \iff |\lambda_0 I_n - A| = \begin{vmatrix} \lambda_0 - a_{11} & & -a_{1n} \\ & \ddots & \\ -a_{n1} & & \lambda_0 - a_{nn} \end{vmatrix} = \lambda_0^n + a_{n-1}\lambda_0^{n-1} + \cdots + a_0,$$

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 λ 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**.

Definition 5.5. (特征多项式)

Definition 5.6. (特征矩阵, 特征根)

Example 5.8. (平面上的旋转变换的特征值)

Example 5.9. (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 $X \in \mathbb{F}^{n\times 1}, Y \in \mathbb{F}^{1\times n}$, 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?

Problem 5.1. 如果有 $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}, AB \in \mathbb{F}^{m\times m}$, $|\lambda I_m - AB|$ 与 $|\lambda I_n - BA|$ 有什么关系? 是相等吗?

先看左边, 进行化简:

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

也就是我们知道了这样的一个命题:

Proposition 5.4. *如果有 $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}, \lambda^n|\lambda I_m - AB| = \lambda^m |\lambda I_n - BA|$.*

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:

Theorem 5.5. *(Hamilton-Cayley) 设 $A \in \mathbb{F}^{n\times n}$ 的特征多项式为*

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

*则 $f(\lambda)$ 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5.6节来完成.

5.4 可对角化

回顾在定义5.4中的内容, 我们这次来看一看什么时候可以对角化. 假设 $\mathscr{A}$ 的不同特征值为 $\lambda_1..\lambda_k$, 对应的不同的特征子空间有 $V_{\lambda_1}(\mathscr{A})..V_{\lambda_k}(\mathscr{A})$. 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4.17) 进行扩展:

Definition 5.7. (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:

- $\forall\alpha \in V_1+V_2+\cdots+V_n$, 有唯一的分解式 $\alpha_1+\alpha_2+\cdots+\alpha_n(\alpha_1 \in V_1, \alpha_2 \in V_2, \cdots, \alpha_n \in V_n)$.
- 零向量的分解式是唯一的
- $\left(\sum_{j\neq i} V_j\right) \cap V_i = \{0\}$
- $\dim V_1 + \dim V_2 + \cdots + \dim V_n = \dim (V_1 + V_2 + \cdots + V_n)$
- 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:

Proposition 5.5. $V_{\lambda_1}(\mathscr{A}) + V_{\lambda_2}(\mathscr{A}) + .. + V_{\lambda_k}(\mathscr{A})$ *是直和.*

?proofname? 我们考虑证明零向量的分解式唯一.

假若 $0 = \alpha_1 + \cdots + \alpha_k$, 能不能推出 $\alpha_1 = \cdots = \alpha_k = 0$?

两边同时作用 $\mathscr{A}$ 上, 有

$$0 = \mathscr{A}\alpha_1 + \cdots + \mathscr{A}\alpha_k = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n = (\alpha_1..\alpha_k)\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_k\end{pmatrix}$$

一直作用 $k-1$ 次, 有

$$
\begin{aligned}
0 &= \lambda_1^2\alpha_1 + .. + \lambda_k^2\alpha_k && = (\alpha_1..\alpha_k)\begin{pmatrix}\lambda_1^2\\ \vdots\\ \lambda_k^2\end{pmatrix}\\
0 &= \lambda_1^3\alpha_1 + .. + \lambda_k^3\alpha_k && = (\alpha_1..\alpha_k)\begin{pmatrix}\lambda_1^3\\ \vdots\\ \lambda_k^3\end{pmatrix}\\
&\vdots && \vdots\\
0 &= \lambda_1^{k-1}\alpha_1 + .. + \lambda_k^{k-1}\alpha_k && = (\alpha_1..\alpha_k)\begin{pmatrix}\lambda_1^{k-1}\\ \vdots\\ \lambda_k^{k-1}\end{pmatrix}
\end{aligned}
$$

把上述的列合并起来, 就有

$$
(0,0,\cdots,0) = (\alpha_1,\alpha_2,\cdots,\alpha_k)\begin{pmatrix}1 & \lambda_1 & \cdots & \lambda_1^{k-1}\\ 1 & \lambda_2 & & \lambda_2^{k-1}\\ \vdots & & & \vdots\\ 1 & \lambda_k & \cdots & \lambda_k^{k-1}\end{pmatrix}
$$

由于后面的是 Vandermonde 行列式, 一定不为 0, $(\alpha_1,\alpha_2,\cdots,\alpha_k)$ 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?

Theorem 5.6. *(可对角化的判别法则) 如果 $\mathscr{A}$ 可对角化, 那么:*

- *$\exists$ 一组基, 使得 $\mathscr{A}$ 的矩阵为对角形.*
- *$\mathscr{A}$ 有 n 个线性无关特征向量.*
- *$V_{\lambda_1}(\mathscr{A}) \oplus V_{\lambda_2}(\mathscr{A}) \oplus .. \oplus V_{\lambda_n}(\mathscr{A}) = V$.*
- *$\dim V_{\lambda_1}(A) + \cdots + \dim V_{\lambda_n}(A) = \dim V$.*

下面来举一个在多项式空间里面求导的一个例子.

Example 5.10. 考虑 $\frac{d}{dx} \in \mathrm{End}(\mathbb{F}[x]_n)$,

Example 5.11. 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 $\mathscr{A}$ 有 n 个互不相同的特征值 (也就是 $\dim V_{\lambda_i}(\mathscr{A}) = 1$), 那么 $\mathscr{A}$ 可对角化.

Example 5.12. 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).

Solution. 求出 $|\lambda I_n - A| = \begin{vmatrix}\lambda & & & & -1\\ -1 & \lambda & & & \\ & -1 & \ddots & & \\ & & \ddots & \lambda & \\ & & & -1 & \lambda\end{vmatrix} = \lambda^n - 1$, 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5.12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 $|\lambda I_n - A| = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = f(\lambda)$. 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

$$\begin{aligned}
T^{-1}f(A)T =&T^{-1}(a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1})T \\
=& \begin{pmatrix} a_0 & & \\ & \ddots & \\ & & a_0 \end{pmatrix} + \begin{pmatrix} a_1\omega & & \\ & \ddots & \\ & & a_1\omega^{n-1} \end{pmatrix} + \cdots + \begin{pmatrix} a_{n-1}\omega & & \\ & \ddots & \\ & & a_{n-1}\omega^{n-1} \end{pmatrix} \\
=& \begin{pmatrix} f(1) & & & \\ & f(\omega) & & \\ & & \ddots & \\ & & & f(\omega^{n-1}) \end{pmatrix}.
\end{aligned}$$

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 $f(A) = a_0I_n + a_1A + \cdots + a_{n-1}A^{n-1}$. 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5.5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

$$\begin{tikzcd}
\mathscr{A} \arrow[r, "\alpha_1..\alpha_n"] \arrow[dr, "\beta_1..\beta_n"'] & A \arrow[d, "过渡矩阵T?"] \\
& B = \mathrm{diag}(\lambda_1..\lambda_n) = T^{-1}AT
\end{tikzcd}$$

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

$$(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

而这就暗示了 $\lambda\mathrm{id}-\mathscr{A}$ 可以被表示为

$$(\lambda_1 I_n - A)(\lambda_2 I_n - A)\cdots(\lambda_n I_n - A) = 0.$$

因为 A, B 相似, 有

$$(\lambda_1 I_n - B)(\lambda_2 I_n - B)\cdots(\lambda_n I_n - B) = 0$$

$$\begin{pmatrix} 0 & & & \\ & \lambda_2-\lambda_1 & & \\ & & \ddots & \\ & & & \ddots \end{pmatrix} \cdots \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} = 0$$

从而可以把复杂的问题简单化.

当 A 可对角化的时候, $f(\lambda) = |\lambda I_n - A|$, 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:

$$f(\lambda) = |\lambda I_n - B| = (\lambda-\lambda_1)^{i_1}(\lambda-\lambda_2)^{i_2}\cdots(\lambda-\lambda_k)^{i_k}$$

那么它对应的对角形矩阵可以是 $\mathrm{diag}(\overbrace{\lambda_1,..,\lambda_1}^{i_1个}, \overbrace{\lambda_2..\lambda_2}^{i_2个}, \cdots, \overbrace{\lambda_k..\lambda_k}^{i_k个})$.

前面我们看到, 对于矩阵 B 的特征多项式而言, 有 $f(\mathscr{B}) = 0$. 那么还有没有别的多项式, 使得 $f(\mathscr{B}) = 0$ 呢? 其实 $(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k) := m(\lambda)$ 已经可以保证 $m(\mathscr{B}) = 0$ 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.

Definition 5.8. (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?

5.5 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: $\mathscr{A}(\alpha_1..\alpha_n) = (\alpha_1..\alpha_n)A$.

$$\mathrm{End}V/\mathbb{F} \xrightarrow{\alpha_1..\alpha_n} \mathbb{F}^{n\times n}$$

$$\uparrow\in \qquad\qquad \uparrow\in$$

$$\begin{array}{ccc} \mathscr{A} & \longrightarrow & A \\ & \searrow^{\beta_1..\beta_n} & \\ & & B = T^{-1}AT \end{array}$$

如果我们找的 B 恰好为对角形, 也就是设 $B = \mathrm{diag}(\lambda_1..\lambda_n)$, 那么就必须满足 $\mathscr{A}\beta_i = \lambda_i\beta_i$. 经过变换我们知道 $(\lambda_i\mathrm{id} - \mathscr{A})\beta_i = 0$, 意味着 $\beta_i \in \ker(\lambda_i\mathrm{id} - \mathscr{A})$, 对应到矩阵的表示形式就有 $|\lambda_i\mathrm{id} - A| = 0$ 这个行列式等于 0 的条件. 此时, λ_i 就叫**特征值**, β_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 $(\lambda_i\mathrm{id} - \mathscr{A})\beta_i = 0$ 中的 λ_i 这样一个确定的数换做了一个变量 λ. 也就是定义 $f(\lambda) = |\lambda\mathrm{id} - \mathscr{A}| = |\lambda I_n - A|$ 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 $f(\lambda_0) = 0, V_{\lambda_0}(\mathscr{A}) = \ker(\lambda_0\mathrm{id} - \mathscr{A}) = \{\alpha \in V | \mathscr{A}\alpha = \lambda_0\alpha\}$.

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

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

注意到如果 A 是对角形, 那么根据5.4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5.4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 $f(\lambda) = |\lambda I_n - A|$, $f(A) = 0$ 都是对的 (现在先不给出证明, 因为需要用到定理5.5, 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:

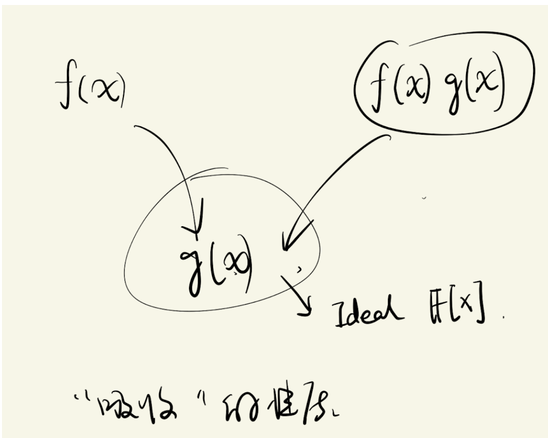
Definition 5.9. 如果 $A \in \mathbb{F}^{n\times n}, f(\lambda) \in \mathbb{F}[\lambda]$, 称 $f(\lambda)$ 为**零化多项式**, 如果 $f(A) = 0$.

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 $\le n^2$ 的非零多项式 $f(\lambda)$, 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; $\mathbb{F}[\lambda]\to\mathbb{F}^{n\times n}: f(\lambda)\to f(A)$ **不是单射**都说明了这个问题.

- (2) 考虑零化多项式的全体 $\ker\varphi=\{g(\lambda)\in\mathbb{F}[\lambda]|g(A)=0\}$. 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 $g(x)\in\ker\varphi$, $h(x)\in\mathbb{F}[\lambda], g(\lambda)h(\lambda)\in\ker\varphi$. 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 $f[\lambda]$ 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** $m_A(\lambda)\in\ker\varphi$, 满足 $m_A(\lambda)|g(\lambda)$, 对于 $\forall g(\lambda)\in\ker\varphi$. 也就是说核空间 $\ker\varphi=\{m_A(\lambda)h(\lambda)|h(\lambda)\in\mathbb{F}[\lambda]\}$. 如果 Hamilton-Cayley 定理 (定理5.5,A 的特征多项式包含在核里面, 即 $|\lambda I_n-A|\in\ker\varphi$) 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

$$\begin{tikzcd}
\mathscr{A}\text{可对角化} \arrow[r, "\alpha_1..\alpha_n"] \arrow[dr, "\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_n}", "\text{打乱顺序}"'] & A \\
& B\text{也是对角矩阵}
\end{tikzcd}
\qquad = \mathrm{diag}(\overbrace{\lambda_1,..,\lambda_1}^{n_1\text{个}},\overbrace{\lambda_2..\lambda_2}^{n_2\text{个}},\cdots,\overbrace{\lambda_m..\lambda_m}^{n_m\text{个}})=\mathrm{diag}(\lambda_1I_{n_1},\lambda_2I_{n_2},\cdots,\lambda_mI_{n_m})$$

我们知道了 $f(\lambda)=|\lambda I_n-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_k)^{n_k}$, $f(A)=0,\forall A$. 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 $(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$. 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.

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

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

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

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

Theorem 5.7. *以下条件等价:*

- *A 可对角化 ($\mathscr{A}$ 可对角化)*
- *存在非零的零化多项式 $g(\lambda)$, 使得 $g(\lambda)$* **不相同的**一次因式的乘积*.*
- *$m_A(\lambda)$ 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 $m_A(\lambda) = (\lambda-\lambda_1)..(\lambda-\lambda_k)$, 希望得到 $V_{\lambda_1}(\mathscr{A})\oplus..\oplus V_{\lambda_k}(\mathscr{A}) = V$. 这里的一个小问题是我们并不知道 λ_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

$$m_A(\mathscr{A}) = (\mathscr{A}-\lambda_1\mathrm{id})..(\mathscr{A}-\lambda_n\mathrm{id}) = 0$$

然后因为 $V_{\lambda_1} = \ker(\mathscr{A}-\lambda_1\mathrm{id})$, 于是就有:$\ker(\mathscr{A}-\lambda_1\mathrm{id})\oplus..\oplus\ker(\mathscr{A}-\lambda_n\mathrm{id})$ 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 $g(\lambda)$, 使得 $g(\mathscr{A}) = 0$, 那么先分解太多, 先考虑分解成两项. 假设 $g(\lambda) = h(\lambda)k(\lambda)$, 这时候带入线性变换 $\mathscr{A}$ 就有了 $h(\mathscr{A})k(\mathscr{A}) = 0$. 我们来关注 $\ker h(\mathscr{A}), \ker k(\mathscr{A})$, 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 $\ker(\mathscr{A}-\lambda_1\mathrm{id})\oplus\ker((\mathscr{A}-\lambda_2\mathrm{id})..(\mathscr{A}-\lambda_n\mathrm{id})$ 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.

Theorem 5.8. *如果 $g(\lambda)\in\mathbb{F}[\lambda], g(\lambda) = h(\lambda)k(\lambda), \gcd(h(\lambda),k(\lambda)) = 1, \mathscr{A}\in\mathrm{End}V, \ker g(\mathscr{A}) = \ker h(\mathscr{A})\oplus\ker k(\mathscr{A})$.*

Example 5.13. (求导的特征多项式是 $\lambda^n = \lambda\cdot\lambda^{n-1} = \lambda^k\cdot\lambda^{n-k}$)

***?proofname?* 是直和.** 如果有 $\alpha\in\ker h(\mathscr{A})\cap\ker k(\mathscr{A})$, 已知 $h(\mathscr{A})\alpha = 0, k(\mathscr{A})\alpha = 0$. 由于两个多项式**互素**, 即 $\gcd(h(\lambda),k(\lambda)) = 1$, 那么由于 Bezout 定理, 有

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

带入 $\mathscr{A}$, 有

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

作用在 α 上, 有

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

也就是 $0 + 0 = \alpha = 0$.

与空间 V 相等. 注意到 $\ker g(\mathscr{A}) \supseteq \ker h(\mathscr{A}) \oplus \ker k(\mathscr{A})$, 只需要证明反包含关系即可. 也就是 $\ker g(\mathscr{A}) \subseteq \ker h(\mathscr{A}) \oplus \ker k(\mathscr{A})$. 这时候取 $\alpha \in \ker g(\mathscr{A})$, 由于直和知道 $g(\mathscr{A}) = h(\mathscr{A})k(\mathscr{A}) = 0$, 与 (*) 式对比, 有 $k(\mathscr{A})u(\mathscr{A})h(\mathscr{A})\alpha = 0$, 意味着 $\alpha_2 \in \ker k(\mathscr{A})$. 同理 $\alpha_1 \in \ker h(\mathscr{A})$, 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:

$$\begin{aligned} V = \ker 0 &= \ker(\mathscr{A} - \lambda_1 \mathrm{id}) \oplus \ker((\mathscr{A} - \lambda_2 \mathrm{id})..(\mathscr{A} - \lambda_n \mathrm{id})) \\ &= (\mathscr{A} - \lambda_1 \mathrm{id}) \oplus (\mathscr{A} - \lambda_2 \mathrm{id}) \oplus .. \oplus (\mathscr{A} - \lambda_n \mathrm{id}) \end{aligned}$$

也就是 V 是特征子空间的直和, V 可对角化.

Example 5.14. 给出 $A^2 = I_n, A^2 = A, A^3 = I_n$ 的条件意味着 $\lambda^2 - 1, \lambda^2 - \lambda, \lambda^3 - \lambda$ 是他们的零化多项式. 但是不一定是最小多项式.

Problem 5.2. 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?

Problem 5.3. 如果有 $\mathscr{A} \in \mathrm{End}V, A \in \mathbb{F}^{n\times n}$, 如果 $f(\lambda)$ 是特征多项式, $m(\lambda)$ 是最小多项式, $p_i(\lambda)$ 是首一不可约多项式, 并且 $f(\lambda) = p_1(\lambda)^{k_1}p_2(\lambda)^{k_2}\cdots p_l(\lambda)^{k_l}$, 那么 $m(\lambda)$ 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5.8, 如果不是一次因式的乘积, 如果 $g(\lambda) = p_1(\lambda)^{k_1}p_2(\lambda)^{k_2}\cdots p_l(\lambda)^{k_l}$ 是 $\mathscr{A}$ 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 $V = \ker g(\mathscr{A}) = \ker p_1(\lambda)^{k_1} \oplus \ker(p_2(\lambda)^{k_2}..p_n(\lambda)^{k_n})$, 因为中国剩余定理, 总是可以继续分解下去变成形如

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

那么我们如果有在 $\ker p_1(\lambda)^{k_1}$ 里面取一组基 $\alpha_{11}..\alpha_{1i_1}$, $\ker p_2(\lambda)^{k_2}$ 里面取一组基 $\alpha_{21}..\alpha_{2i_2}$,...,$\ker p_n(\lambda)^{k_n}$ 里面取一组基 $\alpha_{n1}..\alpha_{ni_n}$, 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 $\mathscr{A}\alpha_{11} \in \ker p_1(\mathscr{A})^{k_1}$, $p_1(\mathscr{A})^{k_1}\mathscr{A}\alpha_{11} = \mathscr{A}p_1(\mathscr{A})^{k_1}\alpha_n = 0$, 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 $\mathscr{A}$ 不可对角化, 那么就把 $\mathscr{A}$ 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:

$$\begin{cases} \mathscr{A}\alpha_1 \in & L(\alpha_1..\alpha_k) := W_1 \\ \vdots & \\ \mathscr{A}\alpha_k \in & L(\alpha_1..\alpha_k) := W_1 \\ \mathscr{A}\alpha_{k+1} \in & L(\alpha_{k+1}..\alpha_n) := W_2 \\ \vdots & \\ \mathscr{A}\alpha_n \in & L(\alpha_{k+1}..\alpha_n) := W_2 \end{cases}$$

于是我们引入定义:

Definition 5.10. 设 W 为 V 的子空间, 若对于 $\forall\alpha \in W$, 有 $\mathscr{A}\alpha \in W$, 则称 W 为 $\mathscr{A}$ 的**不变子空间**, 记作 $\mathscr{A}-$ 子空间.

这就意味着取 W 的一组基 $\alpha_1..\alpha_k..\alpha_n$, 那么 $\mathscr{A}$ 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.

5.6 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5.10, 我们还有如下的两种方法来看:

- 限制变换 $\mathscr{A}|_W, W \to W, \alpha \to \mathscr{A}\alpha, \mathscr{A}|_W \in \mathrm{End}W$.
- $W = L(\alpha_1)$ 是 $\mathscr{A}-$ 子空间, 其中 $\mathscr{A}\alpha_1 = \lambda_1\alpha_1$. 将 α_1 扩充为 V 的一组基 $\alpha_1..\alpha_n$, 那么 A 的矩阵就形如 $\begin{pmatrix} \lambda_1 & \beta \\ \mathbf{0} & \mathbf{A_1} \end{pmatrix}$, 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 $T^{-1}AT = \begin{pmatrix} \lambda_2 & * \\ \mathbf{0} & * \end{pmatrix}$. 这样一直做下去, 就可以得到 $T^{-1}AT = D = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix}$ 的矩阵. 这时候二者的特征多项式是一样的, 也就是

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

现在分别将 A 与 D 带入, 注意到

$$f(A) = f(TDT^{-1}) = Tf(D)T^{-1}$$
$$f(D) = (D-\lambda_1 I_n)\cdots(D-\lambda_n I_n)$$
$$= \begin{pmatrix} 0 & * & * & * \\ 0 & \lambda_2-\lambda_1 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n-\lambda_1 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1-\lambda_2 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \lambda_n-\lambda_2 \end{pmatrix} \cdots\cdots \begin{pmatrix} \lambda_1-\lambda_n & * & * & * \\ 0 & \lambda_2-\lambda_n & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

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5.5). 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:

?proofname? 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

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

并且

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

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

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

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

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

整理该式子, 展开左右两边并对照系数, 就有

$$\begin{aligned}&\lambda^n B_{n-1} \quad &&+\lambda^{n-1}\left(B_{n-2}-AB_{n-1}\right)+\cdots \quad &&+\lambda(B_0-AB_1)-AB_0\\ =&\lambda^n I_n &&+a_{n-1}\lambda^{n-1}I_n+\cdots &&+a_1\lambda I_n+a_0 I_n\end{aligned}$$

也就是:

$$\begin{cases}B_{n-1} & =I_n\\ B_{n-2}-AB_{n-1} & =a_{n-1}I_n\\ \vdots & \\ B_1-AB_2 & =a_2 I_n\\ B_0-AB_1 & =a_1 I_n\\ -AB_0 & =a_0 I_n\end{cases},$$

逐步回带进去, 我们就有

$$\begin{aligned}B_0&=a_1I_n+AB_1=\cdots\\ &=a_1I_n+A(a_2I_n+B_2I_n)\\ &=\cdots\\ &=\sum_{i=1}^{n-1}A^{i-1}a_iI_n\end{aligned}$$

看最后一个式子就有

$$0=A^n+\cdots+a_2A^2+a_1A+a_0I_n=f(A).$$

□

也就是前面的 $m_A(\lambda)$ 是最小多项式, $f(\lambda)$ 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 $m_A(\lambda)\mid f(\lambda)$.

Problem 5.4. 如果 $m_A(\lambda)=e, f(\lambda)=\prod_{j=1}^n p_j(\lambda)^{k_j}$, 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 $\mathscr{A}\in \mathrm{End}V, f(\lambda)=|\lambda\mathrm{id}-\mathscr{A}|=\prod_{i=1}^n p_i(\lambda)^{s_1}$, 并且 $f(\mathscr{A})=0$, 我们就可以推出:

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

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5.8就可以告诉我们:

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

也就是把 V 分解成了若干个 $\mathscr{A}-$ 子空间 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 $\mathbb{F}=\boldsymbol{C}$, 设 $p_i(\lambda)$ 都是若干一次因式的乘积. (当然在一般数域上考虑 $f(\lambda)$ 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 $W_i=\ker(\mathscr{A}-\lambda_i\mathrm{id})^{k_i}$, 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5.9下方的探讨中, 我们知道

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

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

Proposition 5.6. *V 可分解为 $\mathscr{A}$ 的特征子空间的直和 $\iff\mathscr{A}$ 可对角化.*

并且 $\mathscr{A}$ 始终可以化为根子空间的直和. 如果考虑定义 $\mathscr{A}_i=\mathscr{A}|_{R_{\lambda_i(\mathscr{A})}}\in\mathrm{End}(R_{\lambda_i}(\mathscr{A}))$. 根据根子空间的定义, $(\mathscr{A}-\lambda_i\mathrm{id})^{k_i}|_{R_{\lambda_i}(\mathscr{A})}=0$. 即, $\left(\mathscr{A}|_{R_{\lambda_i}(\mathscr{A})}-\lambda\mathrm{id}\right)^{k_i}=0$. 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: $\mathscr{B}_i=\mathscr{A}|_{R_{\lambda_i}(\mathscr{A})}-\lambda\mathrm{id}$, 并且 $\mathscr{B}_i^{k_i}=0$. 也就是 $B_i\in\mathrm{End}(R_{\lambda_i}(\mathscr{A}))$ 也成立. 同时说明了 $\mathscr{B}_i$ 有一个零化多项式, 是 λ^{k_i}. 这时候就可以知道 $\mathscr{B}_i$ 有一个最小多项式 $\lambda^{l_i},l_i\le k_i$. $\mathscr{B}_i$ 的特征多项式是 λ^{n_i}. 那么 $\mathscr{A}$ 的所有特征值只能是 λ_i 吗?

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

最后, 我们考虑 $|\lambda I_n-A|=(\lambda-\lambda_1)^{k_1}\cdots(\lambda-\lambda_l)^{k_l},|\lambda\mathrm{id}-\mathscr{A}|=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_l)^{n_l}$, 对比系数发现 $n_i=k_i$. 也就是 $n_i=\dim R_{\lambda_i}(\mathscr{A})=k_i$. 于是我们有

Proposition 5.7. *如果 $f(\lambda)=|\lambda\mathrm{id}-\mathscr{A}|=(\lambda-\lambda_1)^{k_1}\cdots(\lambda-\lambda_l)^{k_l}$, 根子空间的维数是 $\dim R_{\lambda_i}(\mathscr{A})=k_i\ge\dim V_{\lambda_i}(\mathscr{A})$. 其中, $\dim V_{\lambda_i}(\mathscr{A})$ 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.*

5.7 Jordan 标准形

Definition 5.11. (幂零) 设 V 是 $\mathbb{F}$ 上的线性空间, 称 $\mathscr{B}\in\mathrm{End}V$ 为幂零线性变换, 如果存在正整数 m, 使得 $\mathscr{B}^m=0$.

Proposition 5.8. *设 V 是 $\mathbb{F}$ 上的线性空间, $\mathscr{B}\in\mathrm{End}V$ 的特征多项式为 $f(\lambda)$, 那么下列结论等价:*

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

Theorem 5.9. 设 V 是 $\mathbb{F}$ 上的线性空间, $\mathscr{B}\in\mathrm{End}V$ 是幂零线性变换, 则 V 可以分解为 $\mathscr{B}$ 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 $\mathscr{B}$ 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 $\mathscr{A}\in\mathrm{End}V$, 使得 $\mathscr{A}(V_i)=W_i(1\le i\le m)$.

Definition 5.12. (Jordan 块和 Jordan 矩阵) 设 $\lambda_0\in\mathbb{F}$, 称 n 阶方阵

$$J(\lambda_0,n)=\begin{pmatrix}\lambda_0 & & & \\ 1 & \lambda_0 & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_0\end{pmatrix}$$

为 Jordan 块. 设 $\lambda_1,\lambda_2,\cdots,\lambda_m\in\mathbb{F}$, 称 n 阶方阵

$$J=\mathrm{diag}(J(\lambda_1,n_1),J(\lambda_2,n_2),\cdots,J(\lambda_m,n_m)),\sum_{i=1}^{m}n_i=n$$

为 Jordan 矩阵.

Theorem 5.10. 设 $\mathscr{A}\in\mathrm{End}V$ 幂零, 则 $\mathscr{A}$ 在某组基下的矩阵为 *Jordan* 矩阵

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

若 $\lambda_0-\mathrm{id}$ 幂零, 则 $\mathscr{A}$ 在某组基下的矩阵为 *Jordan* 矩阵

$$\mathrm{diag}J(\lambda_0,n_1),J(\lambda_0,n_2),\cdots,J(\lambda_0,n_m),n_1\ge n_2\ge\cdots\ge n_m\ge 1.$$

Theorem 5.11. 任意 $\mathscr{A}\in\mathbb{C}$ 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.

5.8 多项式矩阵

Definition 5.13. 设 $\mathbb{F}$ 是数域, λ 是一个文字, 以 $\mathbb{F}[\lambda]$ 中的多项式为元素的矩阵被称作多项式矩阵或 $\lambda-$矩阵, 以 $\mathbb{F}[\lambda]^{m\times n}$ 表示 $m\times n$ 的 $\lambda-$矩阵 的全体.

Definition 5.14. 设 $A(\lambda)\in\mathbb{F}[\lambda]^{n\times n}$, 若存在 $B(\lambda)\in\mathbb{F}[\lambda]^{n\times n}$ 使得 $A(\lambda)B(\lambda)=I_n$, 则称 $A(\lambda)$ 可逆, $B(\lambda)$ 称为 $A(\lambda)$ 的逆矩阵.

Theorem 5.12. $A(\lambda)\in\mathbb{F}^{n\times n}$ 可逆当且仅当 $|A(\lambda)|=d\neq 0$, 此时 $A(\lambda)$ 有唯一的逆矩阵

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

Proposition 5.9. *($\lambda-$ 矩阵 的初等变换) 可以对 $\lambda-$ 矩阵 做如下的初等变换:*

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

Definition 5.15. 设 $A(\lambda),B(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$, 称 $A(\lambda)$ 与 $B(\lambda)$ 相抵 (等价), 如果经过一系列初等行列变换可将 $A(\lambda)$ 化为 $B(\lambda)$, 或者说存在 m 阶初等 $\lambda-$ 矩阵:$P_1..P_s$ 以及 $Q_1..Q_t$, 使得 $B(\lambda)=P_1P_2\cdots P_sA(\lambda)Q_1Q_2\cdots Q_t$.

Lemma 5.1. 设 $A(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$ 非零, 则存在首一多项式 $b_1(\lambda),..,b_r(\lambda)$ 使得 $A(\lambda)$ 与 $B(\lambda)=diag(b_1(\lambda),..,b_r(\lambda),0,..,0)\in\mathbb{F}^{m\times n}$ 相抵.

Lemma 5.2. 设 $f(\lambda),g(\lambda)\in\mathbb{F}[\lambda]$ 为非零多项式, $m(\lambda)=\gcd(f(\lambda),g(\lambda)),M(\lambda)=\operatorname{lcm}(f(\lambda),g(\lambda))$, 那么 $\begin{pmatrix} f(\lambda) & \\ & g(\lambda)\end{pmatrix}$ 与 $\begin{pmatrix} m(\lambda) & \\ & M(\lambda)\end{pmatrix}$ 相抵.

Theorem 5.13. 设 $A(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$ 非零, 则 A 相抵与 $D(\lambda)=diag(d_1(\lambda),..,d_r(\lambda),0,..,0)$, 其中 $d_i(\lambda)(1\le i\le r)$ 为首一多项式, 并且 $d_i(\lambda)\mid d_{i+1}(\lambda)$.

Definition 5.16. 称 $D(\lambda)$ 为 $A(\lambda)$ 在相抵下的**标准形**, $d_1(\lambda),..,d_r(\lambda)$ 为 $A(\lambda)$ 的不变因子, 设 $d_i(\lambda)$ 的标准分解为 $d_i(\lambda)=p_1(\lambda)^{k_{i1}}p_2(\lambda)^{k_{i2}}\cdots p_l(\lambda)^{k_{il}},1\le i\le n,k_{ij}\ge 0$. 如果 $k_{ij}\ge 1$, 称 $p_j(\lambda)^{k_{ij}}$ 为 $A(\lambda)$ 的初等因子.

Proposition 5.10. 设 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 则 $A(\lambda)$ 与 $B(\lambda)$ 中的元素分别为彼此的多项式组合.

Definition 5.17. 设 $A(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$, 记 $D_k(A(\lambda))$ 为 $A(\lambda)$ 的所有 k 阶子式的最大公因式, 称为 $A(\lambda)$ 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 $D_k(\lambda)$.

Lemma 5.3. *Jordan* 块 $J(\lambda_0,n)$ 的特征矩阵的行列式因子为 $1,1,\cdots,(\lambda-\lambda_0)^n$.

Lemma 5.4. 初等变换不改变 $\lambda-$ 矩阵 的行列式因子.

Theorem 5.14. 设 $A(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$,$D_1(\lambda),D_2(\lambda)\cdots,D_r(\lambda)$ 为 $A(\lambda)$ 的所有非零行列式因子, 则 $A(\lambda)$ 的不变因子为:

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

从而标准形是唯一的.

Theorem 5.15. $A(\lambda)\in\mathbb{F}[\lambda]^{n\times n}$ 可逆当且仅当它是初等 $\lambda-$ 矩阵 的乘积.

Theorem 5.16. 设 $A(\lambda),B(\lambda)\in\mathbb{F}[\lambda]^{m\times n}$ 可逆, 那么下列条件等价:

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

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

Lemma 5.5. 设 $A \in \mathbb{F}^{n\times n}, U(\lambda), V(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$, 那么存在 $Q(\lambda), R(\lambda) \in \mathbb{F}[\lambda]^{n\times n}, U_0, V_0 \in \mathbb{F}^{n\times n}$ 使得下列等式成立:

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

Theorem 5.17. $A, B \in \mathbb{F}^{n\times n}$ 相似, 当且仅当 $\lambda I_n - A$ 与 $\lambda I_n - B$ 相抵.

5.9 线性代数中出现的内容

5.9.1 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 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}$, 则称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的相似矩阵, 或者说 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 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 $\operatorname{diag}(\lambda_1..\lambda_n)$ 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 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \operatorname{diag}(\lambda_1..\lambda_n) = \boldsymbol{D}$, $\varphi(x)$ 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.

5.9.2 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 $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ 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 $(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$ 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.

5.10 基础习题

Problem 5.5. 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.

Solution. 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.

Problem 5.6. 设 $\boldsymbol{A}$ 是 n 阶可逆实对称矩阵, $\boldsymbol{B}$ 是 n 阶实反对称矩阵, 且 $\boldsymbol{AB} = \boldsymbol{BA}$, 证明 $\boldsymbol{A} + \boldsymbol{B}$ 是可逆矩阵.

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

Problem 5.7. 如果 n 阶矩阵 $\boldsymbol{A}$ 的元素全为 k, 那么 $\boldsymbol{A}$ 的 n 个特征值是.

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

Problem 5.8. 若三阶矩阵 $\boldsymbol{A}$ 的特征值为 $1, -1, 2$, 那么齐次线性方程组 $(\boldsymbol{A} + 2\boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$ 的解为.

Problem 5.9. 设矩阵 $\boldsymbol{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$, $\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}$, 求 $\boldsymbol{B} + 2\boldsymbol{E}$ 的特征值和特征向量.

Solution. 由问题5.5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:

$$\boldsymbol{A}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha} \implies \boldsymbol{A}^*\boldsymbol{\alpha} = \frac{|\boldsymbol{A}|}{\lambda}\boldsymbol{\alpha}.$$

我们希望由上述式子凑出 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}$, 观察特征向量:

$$\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{\alpha} = \frac{|\boldsymbol{A}|}{\lambda}\boldsymbol{P}^{-1}\boldsymbol{\alpha} \implies \boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{\alpha} = \frac{|\boldsymbol{A}|}{\lambda}\boldsymbol{P}^{-1}\boldsymbol{\alpha}.$$

因此 $\boldsymbol{P}^{-1}\boldsymbol{\alpha}$ 是 $\boldsymbol{B}$ 的特征向量.

第六章　线性函数与双线性函数

6.1 线性函数与对偶空间

Definition 6.1. (线性函数) 设 V 是 $\mathbb{F}$ 上的线性空间, 我们称 V 到 $\mathbb{F}$ 的线性映射为**线性函数**. V 上所有的线性函数的全体记为 V^* 或 $\mathrm{Hom}(V,\mathbb{F})$. V^* 也是 $\mathbb{F}$ 上的线性空间, 称为 V 的**对偶空间**.

Proposition 6.1. *对于线性映射而言, 有如下的命题成立:*

- *映射 $f : V \to \mathbb{F}$ 为线性函数当且仅当对任意 $k, l \in \mathbb{F}$, $\alpha, \beta \in V$, 有 $f(k\alpha + l\beta) = kf(\alpha) + lf(\beta)$.*
- *若 $f \in V^*$, 则 $f(0) = 0$, 且对任意 $\alpha \in V, f(-\alpha) = -f(\alpha)$.*
- *若 $f \in V^*, \alpha_i \in V, k_i \in \mathbb{F}, 1 \le i \le n$, 则*

$$f\left(\sum_{i=1}^{n} x_i a_i\right) = \sum_{i=1}^{n} k_i f(\alpha_i).$$

Theorem 6.1. *设 $\alpha_1..\alpha_n$ 为 V 的一组基, 对任意 $(b_1..b_n) \in \mathbb{F}^{1\times n}$, 存在唯一的线性函数 $f \in V^*$, 使得 $f(\alpha_i) = b_i, i = 1, 2, \cdots, n$.*

Definition 6.2. 设 $\alpha_1, .., \alpha_n$ 为 V 的一组基, 称 $(f(\alpha_1), f(\alpha_2), \cdots, f(\alpha_n))$ 为 f **在基 $\alpha_1, .., \alpha_n$ 下的矩阵**.

Theorem 6.2. *映射 φ 是从 V^* 到 $\mathbb{F}^{1\times n}$ 的线性同构, 因此 $\dim V = \dim V^*$.*

Definition 6.3. (对偶基) 如果 $f_i(\alpha_j) = \delta_{ij}, 1 \le i, j \le n$, 那么称 $f_1, .., f_n$ 为 $\alpha_1..\alpha_n$ 的对偶基. 相当于比较简单的一种情形.

Theorem 6.3. *(对偶基之间的变换) 设 $\alpha_1, .., \alpha_n$ 与 $\beta_1, .., \beta_n$ 是线性空间 V 的两组基, 其过渡矩阵为 T, $f_1, .., f_n$ 与 $g_1, .., g_n$ 分别是他们的对偶基, 过渡矩阵为 S, 则 $S'T = I_n$.*

Lemma 6.1. *对任意的 $\alpha \in V, \alpha^{**}$ 是 V^* 上的线性函数.*

Proposition 6.2. *映射 φ 是线性映射, 当 V 是有限维的时候, φ 是线性函数.*

6.2 双线性函数

Definition 6.4. (双线性函数) 设 V 是 $\mathbb{F}$ 上的线性空间, $f(\alpha, \beta)$ 是 V 上的一个二元函数 ($\forall(a, b) \in V$, 唯一确定 $\mathbb{F}$ 中的一个数, 记作 $f(\alpha, \beta)$.) 如果称一个函数是**双线性函数 (或双线性型)**, 如果对于任意的 $k, l \in \mathbb{F}, \alpha, \beta, \gamma \in V$, 有

$$\begin{aligned} f(k\alpha + l\beta, \gamma) =& kf(\alpha, \gamma) + lf(\beta, \gamma); \\ f(\gamma, k\alpha + l\beta) =& kf(\gamma, \alpha) + lf(\gamma, \beta). \end{aligned}$$

Example 6.1. (矩阵的迹与双线性函数)

Example 6.2. 设 $V=\mathbb{F}^n$, $A\in\mathbb{F}^{n\times n}$, 定义

$$f(X,Y)=X'AY,\ \forall X,Y\in V,$$

那么 $f(X,Y)$ 是 V 上的双线性函数.

根据例子6.2可以知道, 线性函数都可以表示成这样的类型. 如果设 $\alpha_1,..,\alpha_n$ 是 V 的一组基, 对任意的 $\alpha=\sum_{i=1}^n x_i\alpha_i,\beta=\sum_{j=1}^n y_j\alpha_j$, 我们有

$$f(\alpha,\beta)=f\left(\sum_{i=1}^n x_i\alpha_i,\sum_{j=1}^n y_j\alpha_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_i f(x_i,y_j)y_j=X'AY.$$

称 A 为 $f(\alpha,\beta)$ 在基 $\alpha_1..\alpha_n$ 下的**度量矩阵**.

Definition 6.5. (度量矩阵) 如果将 n 维线性空间 V 的一组基为 $\alpha_i=\sum_{i=1}^n x_i\alpha_i$, 现在希望把它换为另外一组基 $\beta=\sum_{j=1}^n y_i\alpha_j$, 如果有双线性函数 $f(\alpha,\beta)$, 那么有

$$f(\alpha,\beta)=f\left(\sum_{i=1}^n x_i\alpha_i,\sum_{j=1}^n y_j\alpha_j\right)=\sum_{i=1}^n\sum_{j=1}^n x_i f(\alpha_i,\alpha_j)y_j:=X'AY,$$

其中 $X=(x_1,x_2,\cdots,x_n)',Y=(y_1,y_2,\cdots,y_n)',A=(f(\alpha_i,\alpha_j))\in\mathbb{F}^{n\times n}$, 称 A 为 $f(\alpha,\beta)$ 在基 $\alpha_1..\alpha_n$ 下的**度量矩阵**.

Proposition 6.3. *考虑所有的双线性函数的全体, 映射 $\varphi:\mathbb{B}\to\mathbb{F}^{n\times n}$ 是双射.*

在研究矩阵的时候可以注意到可逆的矩阵可以给我们带来很大的便利, 于是我们定义非退化 (满秩) 双线性函数:

Definition 6.6. (满秩双线性函数) 如果 n 维线性空间上 V 的双线性函数 f 满足:

$$f(\alpha,\beta)=0\iff(\forall\alpha\in V\implies\beta=0),$$

那么 f 是**非退化**或**满秩**的.

Theorem 6.4. *设 f 是 V 上的双线性函数, $\epsilon_1,\epsilon_2,\cdots,\epsilon_n$ 是 V 的一组基, f 在这组基下的度量矩阵为 $A_f=(f(\epsilon_i,\epsilon_j))$, 则下列条件等价:*

- *f 是非退化的;*
- *A_f 可逆;*
- *若 $f(\alpha,\beta)=0,\forall\beta\in V\implies\alpha=0$;*
- *对于任意的非零 $\alpha\in V$, 存在 $\beta\in V$, 使得 $f(\alpha,\beta)\neq0$.*

很多时候, 双线性函数并不具有对称性, 也就是说 $f(\alpha,\beta)\neq f(\beta,\alpha)$. 这样, 等于零的情形就有两个. 分别是左边的等于 0 时和右边等于 0 时. 于是我们给出正交补的概念:

Definition 6.7. (左正交补与右正交补) 对任意 V 的子空间 W, V 中的子集

$$
\begin{aligned}
{}^{\perp}W &= \{\alpha \in V | f(\alpha,\beta) = 0, \forall \beta \in W\};\\
W^{\perp} &= \{\alpha \in V | f(\beta,\alpha) = 0, \forall \beta \in W\},
\end{aligned}
$$

分别定义为 W 的**左正交补**与**右正交补**. 一般情形下, ${}^{\perp}W$ 与 $W^{\perp}$ 并不相等

Lemma 6.2. 设 f 是 n 维线性空间 V 上的非退化双线性函数, 对于任意的 $\beta \in V, g_\beta(\alpha) = f(\alpha,\beta)$ 为 V 上的线性函数, 即我们有映射

$$\varphi : V \to V^*, \beta \mapsto g_\beta,$$

那么 φ 是线性映射且 $\ker\varphi = V^{\perp}$, 从而是线性同构.

如果给定 V 的子空间, W, f 自然可以看做 W 上的双线性函数, 记作 $f|_W$, 那么我们有

Lemma 6.3. 设 f 是 n 维线性空间 V 上的双线性函数, 对任意的 V 的子空间 W, 如果 $f|_W$ 非退化, 那么 $V = W \oplus W^{\perp} = W^{\perp} \oplus W$.

这个引理我们可以考虑使用向量的分解式唯一的方法去证明之.

现在我们可以研究一般的双线性函数与矩阵之间的关系, 也就是双线性函数在不同基下的度量矩阵有什么关系? 也就是如果有两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到 $\beta_1, \beta_2, \cdots, \beta_n$ 之间的过渡矩阵为 $T = (t_{ij})$, 那么

$$
\begin{aligned}
f(\beta_i,\beta_j) &= f\left(\sum_{k=1}^{n} t_{ki}\alpha_k, \sum_{l=1}^{n} t_{lj}\alpha_l\right) = \sum_{k=1}^{n}\sum_{l=1}^{n} t_{ki} f(\alpha_k,\alpha_l) t_{lj}\\
&= (t_{1i}, \cdots, t_{ni}) \begin{pmatrix} f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n)\\ \vdots & & \vdots\\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n) \end{pmatrix} \begin{pmatrix} t_{1j}\\ \vdots\\ t_{nj} \end{pmatrix}
\end{aligned}
$$

由于定理6.3可以知道基变换对矩阵的影响可以表示为

$$
\begin{aligned}
B = (T_i' A T_j) &= \begin{pmatrix} T_1'AT_1 & \cdots & T_1'AT_n\\ \vdots & & \vdots\\ T_n'AT_1 & \cdots & T_n'AT_n \end{pmatrix}\\
&= \begin{pmatrix} T_1'\\ \vdots\\ T_n' \end{pmatrix} (AT_1, \cdots, AT_n)\\
&= \begin{pmatrix} T_1'\\ \vdots\\ T_n' \end{pmatrix} A(T_1, T_2, \cdots, T_n)\\
&= T'AT
\end{aligned}
$$

于是我们得到了不同基下两个矩阵的关系. 我们把它定义为**合同**.

Definition 6.8. (矩阵的合同) 称 $A, B \in \mathbb{F}^{n\times n}$ 合同, 意味着存在可逆矩阵 $T \in \mathbb{F}^{n\times n}$, 使得

$$B = T^{-1}AT.$$

Proposition 6.4. *矩阵的合同关系是等价关系.*

这就意味着

Theorem 6.5. *数域 $\mathbb{F}$ 上的 n 维线性空间 V 上的双线性函数在不同基下的度量矩阵是合同的. 反之, 如果 $A, B \in \mathbb{F}^{n\times n}$ 合同, 则 A, B 是 V 上某个双线性函数在不同基下的矩阵.*

这样, 我们希望能不能对于一般的线性函数, 选择一组基, 使得其可以变换成一个容易研究的矩阵呢? 在一般的线性函数上, 好像没有特别好的性质. 我们来考虑若干特殊的情形.

首先考虑某组基下的度量矩阵是对角矩阵 Λ, 那么它在其他的基下的度量矩阵 $B = T'\Lambda T$, 是一个对称矩阵. 由于之前的情形, 我们知道任意一个矩阵可以写成一个对称矩阵与一个反对称矩阵之和, 于是自然我们应该考虑其对应的双线性函数: 对称双线性函数与反对称双线性函数. 引入如下的定义:

Definition 6.9. (对称双线性函数与反对称双线性函数) 称线性空间 V 上的双线性函数 $f(\alpha,\beta)$ 为**对称双线性函数**, 如果

$$f(\alpha,\beta) = f(\beta,\alpha), \forall \alpha,\beta \in V;$$

称 $f(\alpha,\beta)$ 为**反对称双线性函数**, 如果

$$f(\alpha,\beta) = -f(\beta,\alpha), \forall \alpha,\beta \in V.$$

这样就有一个好处, 就是他们的左正交补与右正交补是相等的. 称作**正交补**.

反对称双线性函数有什么性质?

Theorem 6.6. *(反对称双线性函数的标准形式) 设 $f(\alpha,\beta)$ 是 n 维线性空间 V 上的犯规陈双线性函数, 则存在 V 的一组基 $\beta_1, \beta_2, \cdots, \beta_n$ 使得*

$$(f(\beta_i,\beta_j) = diag(S_2, S_2, \cdots, S_2, 0, 0, \cdots, 0),$$

其中 $S_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

?proofname? 考虑 $f \neq 0$, 即存在 $\beta_1, \beta_2 \in V$ 使得 $f(\beta_1,\beta_2) = c \neq 0$, 用 $\frac{1}{c}\beta_2$ 代替 β_2, 使得 $f(\beta_1,\beta_2) = 1$. 于是 $W = L(\beta_1\beta_2)$ 是 V 的二维子空间, 并且 $f|_W$ 非退化. 自然 β_1, β_2 线性无关, 同时也是 W 的一组基, 并且 $f|_W$ 在 β_1, β_2 下的矩阵为 S_2. 于是 $V = W \oplus W^{\perp}$, $f|_{W^{\perp}}$ 也是反对称双线性函数. 利用归纳法知道存在 $W^{\perp}$ 的基 $\beta_3, \cdots, \beta_n$ 使得 $f|_{W^{\perp}}$ 在该组基下的度量矩阵是

$$\mathrm{diag}(\underbrace{S_2, \cdots, S_2}_{k-1\text{个}}, 0, \cdots, 0).$$

故 $\beta_1, \beta_2, \cdots, \beta_n$ 为 V 的基, 且

$$(f(\beta_i, \beta_j)) = \mathrm{diag}(\underbrace{S_2, \cdots, S_2}_{k\text{个}}, 0, \cdots, 0).$$

□

这就意味着反对称线性函数的矩阵的秩一定是偶数.

6.3 对称双线性函数与二次型

下面研究对称双线性函数的标准情形. 由 $f(\alpha, \beta)$ 对称有: $f(\alpha+\beta, \alpha+\beta) = f(\alpha, \alpha) + f(\beta, \beta) + 2f(\alpha, \beta)$. 如果我们记 $Q_f(\alpha) = (\alpha, \alpha)$, 那么有 $f(\alpha, \beta) = \frac{1}{2}(Q_f(\alpha+\beta) - Q_f(\alpha) - Q_f(\beta))$. 也就是说, f 是有 Q_f 唯一定义, Q_f 也是由 f 唯一确定. 于是我们定义:

Definition 6.10. (二次型函数函数 Q_f) 如果 $\mathbb{F}$ 上的线性空间 V 到 $\mathbb{F}$ 的映射 Q 满足:

- (二次齐次性) 对于任意 $k \in \mathbb{F}, \alpha \in V$, 有 $Q(k\alpha) = k^2 Q(\alpha)$;
- $f(\alpha, \beta) = \frac{1}{2}(Q_f(\alpha+\beta) - Q_f(\alpha) - Q_f(\beta))$ 是 V 上的对称双线性函数.

则称 Q 是 V 的一个**二次型函数.**

并且我们可以发现 V 上的对称双线性函数与二次型函数存在一一对应.

Proposition 6.5. *(双线性函数与二次型函数的一一对应) V 上的双线性函数与二次型函数之间一一对应.*

这就意味着对于对称双线性函数而言, 我们就可以用带着一个变量的字母来考虑整件事情了. 这时候, 如果 V 是 n 维线性空间, 选定 V 的一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$, 设 α, β 的坐标分别是 $(x_1, x_2, \cdots, x_n)', (y_1, y_2, \cdots, y_n)', A = (a_{ij})$, 其中 $a_{ij} = f(\epsilon_i, \epsilon_j)$, 那么 $f(\alpha, \beta) = X'AY, Q_f(\alpha) = X'AX$. 这样 $Q_f(\alpha) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ 是 $\mathbb{F}$ 上的二元齐次多项式. 于是我们定义

Definition 6.11. (二元齐次多项式) 设 $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ 对称, 称二次其次多项式

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

为 $\mathbb{F}$ 上的一个 n **元二次型**, 简称**二次型**. 称 A 为**二次型** $f(x_1, x_2, \cdots, x_n)$ **的矩阵**, 也记为 A_f.

于是我们就有了如下的研究对象:

- V 上的对称双线性函数;
- V 上的二次型函数;

- $\mathbb{F}$ 上的 n 元二次型;
- $\mathbb{F}$ 上的 n 阶对称矩阵.

这四个数学对象的本质是一致的, 只是解答问题的时候需要我们按需选择要用的形式.

选择对称矩阵的一大好处就是我们可以把它对角化.

Theorem 6.7. *设 V 是 $\mathbb{F}$ 上的 n 维线性空间, $f(\alpha,\beta)$ 是 V 上的对称双线性函数, 则存在 V 的一组基, $\beta_1,\beta_2,\cdots,\beta_n$ 使得 $(f(\beta_i,\beta_j))$ 是对角矩阵.*

?proofname? 若对于任意 $\alpha\in V$, 有 $Q_f(\alpha)=f(\alpha,\alpha)=0$, 则这个函数恒等于 0, 结论自然成立.

如果 $f(\alpha,\alpha)$ 不恒等于 0, 取一个 $\beta_1\in V$, 使得 $f(\beta_1,\beta_1)\neq 0$. 于是 f 在 V 的子空间 $W=L(\beta_1)$ 上的限制是非退化的. 故 $V=W\oplus W^{\perp}$. 且 $f|_{W^{\perp}}$ 是 $W^{\perp}$ 的对称双线性函数. 利用归纳法可以知道存在一组基 $\beta_2,\cdots,\beta_n$ 使得 $f|_{W^{\perp}}$ 在这组基下的矩阵是对角矩阵. 由于 $\beta_i\in W^{\perp}, i=2,3,\cdots,n$, 意味着 $f(\beta_1,\beta_i)=0$. 于是 f 在 V 的矩阵 $\beta_1,\beta_2,\cdots,\beta_n$ 为对角矩阵. □

这样一来, 给定一组基, 如果有 $\alpha=\sum_{i=1}^n x_i\beta_i, \beta_i=\sum_{j=1}^n y_j\beta_j$, 我们有 $f(\alpha,\beta)=\sum_{i=1}^n d_ix_iy_i$. 用矩阵的语言可以这样描述:

Corollary 6.1. *设 $A\in\mathbb{F}^{n\times n}$ 对称, 存在可逆矩阵 $T\in\mathbb{F}^{n\times n}$ 使得 $T'AT$ 为对角矩阵.*

我们还可以把上述的内容表述为二次型的情形. 我们思考一下**线性替换**.

Definition 6.12. (线性替换) 设 $X=(x_1,x_2,\cdots,x_n)'$ 与 $Y=(y_1,y_2,\cdots,y_n)'$ 是两组变量, $T=(t_{ij})\in\mathbb{F}^{n\times n}$. 称

$$\begin{cases} x_1=t_{11}y_1+t_{12}y_2+\cdots+t_{1n}y_n & (1)\\ x_2=t_{21}y_1+t_{22}y_2+\cdots+t_{2n}y_n & (2)\\ \cdots & \cdots\\ x_n=t_{n1}y_1+t_{n2}y_2+\cdots+t_{nn}y_n & (n)\end{cases} \text{或} X=TY$$

为 $x_1,x_2,\cdots,x_n$ 到 $y_1,y_2,\cdots,y_n$ 的一个线性替换, 如果 T 可逆, 则称这个线性代换是**可逆**的. 称 $\mathbb{F}$ 上的二次型 $f(x_1,x_2,\cdots,x_n)$ 与 $g(y_1,y_2,\cdots,y_n)$ 等价, 如果存在可逆线性替换 $X=TY$ 使得

$$f(x_1,x_2,\cdots,x_n)=g(y_1,y_2,\cdots,y_n).$$

由于上述的探索, 我们发现

Lemma 6.4. *(合同变换变为标准形) 数域上的二次型 $f(x_1,x_2,\cdots,x_n)$ 与 $g(y_1,y_2,\cdots,y_n)$ 等价当且仅当两个二次型的度量矩阵合同.*

Theorem 6.8. *(二次型的标准形.)* 数域 $\mathbb{F}$ 上的二次型 $f(x_1,x_2,\cdots,x_n)$ 可以经过可逆线性替换可以化为平方和的形式 $X=TY,|T|\neq 0$, 使得

$$f(x_1,x_2,\cdots,x_n)=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2.$$

称等号右手边的式子称为**二次型的标准形.**

上面的定理告诉我们二次型的标准形是一定存在的. 定理6.7告诉了我们从双线性函数的角度求解标准形的方法. 下面介绍两种更加容易操作的方法: 分别是**配方法**和**初等变换法**.

从定理6.7导出一组可以用的解. 设在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下, V 上的双线性函数为 $f(\alpha,\beta)=X'AY$.

(1) 如果存在 i, 使得 $f(\alpha_i,\alpha_i)\neq 0$, 那么重排基的顺序, 不妨设 $i=1,\beta_1=\alpha_1$, 那么 $\beta_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一组基, 要求 $L(\beta_1)^\perp$, 只要求一组基. 也就是线性无关的 $\gamma_2,\cdots,\gamma_n$ 为 V 的一组基, 并且满足 $f(\beta_1,\gamma_i)=0$, 对于 $i=2,3,\cdots,n$. 待定系数, 令 $\gamma_i=\alpha_i+k_i\beta_1$, 由于 $f(\gamma_i,\beta_1)=0$, 根据双线性函数的性质有

$$f(\alpha_i,\beta_1)+k_if(\beta_1,\beta_1)=0,$$

即 $k_i=-\frac{f(\alpha_i,\beta_1)}{f(\beta_1,\beta_1)}$. 于是令 $\gamma_i=\alpha_i-\frac{f(\alpha_i,\beta_1)}{f(\beta_1,\beta_1)}\beta_1$ 即可. 可以验证并知道 $\gamma_2,\cdots,\gamma_n$ 为 $L(\beta_1)^\perp$ 的一组基. 继续对于 $L(\beta_1)^\perp$ 进行类似的操作.

(2) 如果对于任意的 i, 都有 $f(\alpha_i,\alpha_i)=0$, 由于 f 不恒为 0, 存在 $f(\alpha_i,\alpha_j)\neq 0$. 于是可以取 $\alpha_1+\alpha_2,\alpha_1-\alpha_2,\alpha_3,\cdots,\alpha_n$. 这样的好处是 $f(\alpha_1+\alpha_2,\alpha_1-\alpha_2)=0$. 然后令 $\beta=\alpha_1+\alpha_2$ 就可以转化为 (1) 的情形.

用二次型的语言来描述. 设

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j,\quad a_{ij}=a_{ji}.$$

(1) 若 a_{ii} 不全为 0, 不妨设 $a_{11}\neq 0$, 对于 x_1 配方得:

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)&=a_{11}x_1^2+2x_1\sum_{j=2}^n a_{1j}x_j+\sum_{i=2}^n\sum_{j=2}^n a_{ij}x_ix_j\\&=a_{11}\left(x_1+\sum_{j=2}^n\frac{a_{1j}}{a_{11}}x_j\right)^2-a_{11}\left(\sum_{j=2}^n\frac{a_{1j}}{a_{11}}x_j\right)^2+\sum_{i=2}^n\sum_{j=2}^n a_{ij}x_ix_j\end{aligned}$$

于是做可逆线性变换

$$\begin{cases}y_1=x_1+\sum_{j=2}^n\frac{a_{1j}}{a_{11}}x_j, & i=1\\ y_i=x_i, & 2\leq i\leq n\end{cases}$$

则二次型化为了

$$a_{11}y_1^2-\frac{1}{a_{11}}\left(\sum_{j=2}^n a_{1j}y_j\right)^2+\sum_{i=2}^n\sum_{j=2}^n a_{ij}x_ix_j.$$

这样就变成了一个 $n-1$ 个代凑的的二次型, 继续操作即可.

(2) 若 $a_{11}=a_{22}=\cdots=a_{nn}=0$, 总存在 $i,j,i\neq j$ 使得 $a_{ij}\neq 0$. 不妨设 $a_{12}\neq 0$. 做可逆线性替换 $x_1=z_1+z_2, x_2=z_1-z_2, x_i=z_i, 3\leq i\leq n$. 这样的好处是做完替换后 z_1z_2 的系数为 0. 之后就化为了情形 (1) 继续就可以了.

Example 6.3. 把 $f(x,y,z)=x^2+5y^2+5z^2+4xy-2xz+4yz$ 化为标准型.

Solution. 根据上面的内容, 有

$$\begin{aligned}f(x,y,z)&=x^2+5y^2+5z^2+4xy-2xz+4yz\\&=x^2+2x(2y-z)+(2y-z)^2-(2y-z)^2\\&\quad+5y^2+5z^2+4yz\\&=(x+2y-z)^2+y^2+8yz+4z^2\\&=(x+2y-z)^2+(y^2+8yz+16z^2)-16z^2+4z^2\\&=(x+2y-z)^2+(x+4z)^2-12z^2\\&=(x+2y-z)^2+(x+4z)^2-12z^2.\end{aligned}$$

Example 6.4. 用配方法将二次型 $f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3+2x_2x_3$ 化为标准型.

Solution. 先做替换 $x_1=z_1+z_2, x_2=z_1-z_2, x_3=z_3$, 那么有

$$\begin{aligned}f(x_1,x_2,x_3)&=2z_1^2-2z_2^2+4z_1z_3\\&=2(z_1^2+2z_1z_3+z_3^2)-2z_2-2z_3\\&=2(z_1+z_3)^2-2z_2-2z_3,\end{aligned}$$

于是令 $\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}$, 就有 $f(x_1,x_2,x_3)=2y_1^2-2y_2^2-2y_3^2$.

对于二次型做可逆线性变换的过程其实是对度量矩阵做一系列的合同变换. 具体地, 对于任意的对称矩阵 A, 存在可逆矩阵 T 使得 $T'AT$ 为对角矩阵, 由于 T 可逆, 于是可以分解成若干个初等矩阵的乘积 $T=P_1P_2\cdots P_k$. 此时, $T'=P_k'\cdots P_2'P_1'$. 于是

$$T'AT=P_k'(\cdots(P_2'(P_1'AP_1)P_2)\cdots)P_k.$$

所以我们可以把二次型的配方法用初等矩阵记录下来. 这就是初等变换法.

我们来看一看初等变换会对目标带来怎样的效果.

- $P=P(i(c))$. 此时, $P'AP$ 的效果是将 A 的第 i 行乘以 c, 得到 $P'A$, 再将此矩阵的第 i 列乘 c.

- $P = P(i,j)$. 此时, $P'AP$ 的效果是将 A 的第 i 行与第 j 列互换, 然后将 A 的第 i 列与第 j 列互换.
- $P = P(i,j(k))$. 此时, 将 A 的第 j 行加上第 i 行的 k 倍, 再将 A 的第 j 列加上第 i 列的 k 倍.

可以看到, 我们可以通过先做一次行变换, 再做一次列变换的方法来解决这个问题. 知道把左侧的对称矩阵化为单位矩阵. 这样右侧的就记录下了 P'. 就是说, $(A|I_n) \to (P'AP|P')$.

Example 6.5. 用初等变换法把二次型 $f(x_1,x_2,x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$ 化为标准型.

Solution. 该二次型的矩阵为 $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, 那么对 $(A|I_3)$ 进行初等变换:

$$\begin{pmatrix} 0 & 1 & 1 & | & 1 & & \\ 1 & 0 & 1 & | & & 1 & \\ 1 & 1 & 0 & | & & & 1 \end{pmatrix} \xrightarrow{\text{行}} \begin{pmatrix} 0 & 1 & 2 & | & 1 & 1 & \\ 1 & 0 & 1 & | & & 1 & \\ 1 & 1 & 0 & | & & & 1 \end{pmatrix} \xrightarrow{\text{列}} \begin{pmatrix} 2 & 1 & 2 & | & 1 & 1 & \\ 1 & 0 & 1 & | & & 1 & \\ 2 & 1 & 0 & | & & & 1 \end{pmatrix}$$

$$\xrightarrow{\text{行}} \begin{pmatrix} 2 & 1 & 2 & | & 1 & 1 & 0 \\ 0 & -1/2 & 0 & | & -1/2 & 1/2 & 0 \\ 0 & 0 & 2 & | & -1 & -1 & 1 \end{pmatrix} \xrightarrow{\text{列}} \begin{pmatrix} 2 & 0 & 0 & | & 1 & 1 & 0 \\ 0 & -1/2 & 0 & | & -1/2 & 1/2 & 0 \\ 0 & 0 & 2 & | & -1 & -1 & 1 \end{pmatrix}$$

由此可得.

6.4 惯性定理与正定二次型

上一节里面我们了解了 n 元二次型的标准形的产生方法. 例如, $\mathbb{F}$ 上的 n 元二次型有标准型 $f(x_1,x_2,\cdots,x_n) = d_1^2y_1 + d_2^2y_2 + \cdots + d_n^2y_n$. 如果我们取全部为 0 的可逆线性替换, 使得 $y_i = t_iz_i$, 可以发现二次型的标准形并不唯一. 如果在 $\mathbb{C}$ 里面, 我们就可以恰当地选取 t_i 使得其前面的系数全部为 1. 于是有定理

Theorem 6.9. *(复数域上二次型的规范形可以唯一) n 元复二次型 $f(x_1,x_2,\cdots,x_n)$ 经过可逆线性变换可以化为 $z_1^2 + z_2^2 + \cdots + z_n^2$. 即:*

$$f(x_1,x_2,\cdots,x_n) = z_1^2 + z_2^2 + \cdots + z_n^2.$$

*其中右式称为二次型的**规范形**. 并且二次型的规范形是唯一的.*

用矩阵的语言叙述之, 就是任意复对称的 n 阶方阵 A 比对称与 $\mathrm{diag}(I_r, 0)$. 其中 $r = r(A)$. 两个复对称的 n 阶方阵合同当且仅当它们的秩相等.

实数域上的情形好像更加复杂一些. 因为我们不能通过一些变换将负的系数变为正的. 因此, 我们猜测经过线性替换并不能更改掉系数前面正负号的个数. 事实上也确实是这样的, 这个定理被叫做 "**惯性定理**".

Theorem 6.10. *(惯性定理) 设 n 元实二次型 $f(x_1,x_2,\cdots,x_n)=X'AX$ 经过可逆线性替换可以化为如下所示的***规范形**

$$f(x_1,x_2,\cdots,x_n)=z_1^2+z_2^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2.$$

?proofname? 设 $f(x_1,x_2,\cdots,x_n)$ 的标准形为 $d_1y_1^2+\cdots+d_py_p^2+d_{p+1}y_{p+1}^2+\cdots+d_ry_r^2$, 其中 $d_1,\cdots,d_p$ 为证书, $d_{p+1},\cdots,d_r$ 为负数. 做线性替换

$$y_i=\begin{cases}\frac{1}{\sqrt{d_i}}z_i, & 1\le i\le p\\ \frac{1}{\sqrt{-d_i}}z_i, & p+1\le i\le r\\ z_i, & r+1\le i\le n\end{cases}$$

则可得

$$f(x_1,x_2,\cdots,x_n)=z_1^2+z_2^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2.$$

设 $w_1^2+w_2^2+\cdots+w_s^2-w_{s+1}^2-\cdots-w_r^2$ 为 $f(x_1,x_2,\cdots,x_n)$ 的另一个规范形. 利用二次型与双线性函数的一一对应关系知道, 存在 n 维实线性空间 V 上的对称双线性函数 $f(\alpha,\beta)$, 使得其在基 $\eta_1,\eta_2,\cdots,\eta_n,\xi_1,\xi_2,\cdots,\xi_n$ 下的度量矩阵分别为 $\operatorname{diag}(I_p,I_{r-p},0)$ 与 $\operatorname{diag}(I_s,I_{r-s},0)$. 令 $V_1=L(\eta_1,\cdots,\eta_p)$, 对于任意的 $\alpha\in V_1$, $f(\alpha,\alpha)\ge 0$, 并且当且仅当 $\alpha=0$ 的时候 $f(\alpha,\alpha)=0$. 令 $V_2=L(\xi_{s+1},\cdots,\xi_n)$, 则对任意的 $\beta\in V_2, f(\beta,\beta)\le 0$. 容易证明 $V_1\cap V_2=\{0\}$, 也就是 V_1+V_2 是直和. 因此

$$p+n-s=\dim V_1+\dim V_2=\dim(V_1+V_2)\le n,$$

从而 $p\le s$, 同理, $s\le p$, 于是 $s=p$, 从而规范形唯一. □

Definition 6.13. (正负惯性指数) 在实二次型 $f(x_1,x_2,\cdots,x_n)$ 的规范形中, 正平方项的个数 p 称为 $f(x_1,x_2,\cdots,x_n)$ 的**正惯性指数**, 负平方项的个数 $q=r-p$ 称为 $f(x_1,x_2,\cdots,x_n)$ 的**负惯性指数**, 称他们的差 $p-q=2p-r$ 为 $f(x_1,x_2,\cdots,x_n)$ 的**符号差**.

接下来我们来考虑一些比较特殊的二次型. 也就是二次型 $X'AX$ 的正惯性指数为 n. 此时, 二次型的规范形为 $y_1^2+y_2^2+\cdots+y_n^2$. 对于任意的 y, 最后的函数值都是正的. 这样一来, 我们在一些问题里面研究起来就非常方便. 于是我们可以给出定义**正定二次型**.

Definition 6.14. (正定二次型) 称实二次型 $f(x_1,x_2,\cdots,x_n)=X'AX$ 为正定二次型, 如果对任意的不全为 0 的实数 $c_1,c_2,\cdots,c_n$,

$$f(c_1,c_2,\cdots,c_n)>0.$$

于是我们称这样的矩阵 A 为正定矩阵.

于是我们可以由若干个二次型的等价条件:

Theorem 6.11. *(二次型的等价条件) 设 $f(x_1,x_2,\cdots,x_n)=XAX$ 是实二次型, 则下列条件等价:*

- *(1)* $f(x_1, x_2, \cdots, x_n)$ *的正惯性指数为* n*;*
- *(2)* A *合同与* I_n*;*
- *(3)* $A = TT'$*, 其中* T *可逆;*
- *(4)* $f(x_1, x_2, \cdots, x_n)$ *是正定二次型.*

?proofname? 只需要证明 (4) $\implies$ (1), 设正定二次型 $f(x_1, x_2, \cdots, x_n) = X'AX$ 经过可逆变换 $Y = TX$ 之后化为规范形为 $y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - y_r^2$, 如果 $p < n$, 则令 $(c_1, c_2, \cdots, c_n)' = T^{-1}(0, 0, 0, \cdots, 1)'$. 这时候 $f(c_1, c_2, \cdots, c_n) \leq 0$, 与正定性矛盾. 于是 $p = n$. □

由于 $|A| = |T'T| = |T|^2$, 我们有

Proposition 6.6. *正定矩阵的行列式大于 0.*

所以, 我们得到了一个必要的条件. 一个问题是, 这个条件充分吗? 事实上, 这个条件并不充分. 如果我们对于任意的 $1 \leq i_1 \leq i_2 \leq \cdots \leq i_m \leq n$, 当 $j \notin \{i_1, \cdots, i_m\}$ 的时候令 $x_j = 0$, 那么二次型就化为了 $\sum_{k=1}^m \sum_{l=1}^m a_{i_k i_l} x_{i_k} x_{i_l}$. 于是得到一个新的 m 元二次型 $X_1' A_1 X_1$, 其中

$$X_1 = \begin{pmatrix} x_{i_1} \\ \vdots \\ x_{i_m} \end{pmatrix} \quad A_1 = \begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_m} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_m i_1} & a_{i_m i_2} & \cdots & a_{i_m i_m} \end{pmatrix}.$$

如果 A 是正定的, 容易看出 $X_1' A_1 X_1$ 也是正定的. 故由上述的推论可以知道 $|A_1| > 0$, 从而 $|A_1|$ 即为 A 的子式 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_m \\ i_1 & i_2 & \cdots & i_m \end{pmatrix}$. 于是, 我们给出定义:

Definition 6.15. (主子式) 设 $A \in \mathbb{R}^{n\times n}$, 称 A 的子式 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_m \\ i_1 & i_2 & \cdots & i_m \end{pmatrix}$ 为 A 的一个 m 阶**主子式**. 特别地, 称 $A\begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix}, m = 1, 2, \cdots, n$ 为 A 的**顺序主子式**.

于是我们可以了解到矩阵正定 (二次型正定) 的判别法则:

Theorem 6.12. *(二次型正定的判别法)* 设 $A \in \mathbb{R}^{n\times n}$ *对称, 那么下列条件等价:*

- *(1)* A *正定;*
- *(2)* A *的主子式都大于 0;*
- *(3)* A *的顺序主子式都大于 0.*

?proofname? 只需要证明 (3) $\Longrightarrow$ (1). 考虑对于 n 的阶数做归纳法. 当 n 为 1 阶的时候, 结论显然成立, 假设 A 的阶数是 $n-1$ 的时候结论成立, 当 $A=(a_{ij})\in\mathbb{R}^{n\times n}$ 时, 考虑分块矩阵:

$$A=\begin{pmatrix}A_1 & \beta\\ \beta' & a_{nn}\end{pmatrix},\beta=(a_{1n},a_{2n},\cdots,a_{n-1,n})'$$

根据条件知道 $|A_1|=A\begin{pmatrix}1 & 2 & \cdots & n-1\\ 1 & 2 & \cdots & n-1\end{pmatrix}>0$, 则 A_1 可逆. 下面做合同变换:

$$\begin{pmatrix}I_{n-1} & 0\\ -\beta' A_1^{-1} & 1\end{pmatrix}\begin{pmatrix}A & \beta\\ \beta' & a_{nn}\end{pmatrix}\begin{pmatrix}I_{n-1} & -\beta' A_1^{-1}\\ 0 & 1\end{pmatrix}=\begin{pmatrix}A_1 & 0\\ 0 & c\end{pmatrix}.$$

其中 $c=a_{nn}-\beta' A_1^{-1}\beta$. 由于 $|A|=|A_1|c$, 所以 $c=\frac{|A|}{|A_1|}>0$. 利用归纳假设可以知道 A_1 为正定矩阵, 因此存在 $n-1$ 阶实可逆的矩阵 S 使得 $S'A_1S=I_{n-1}$, 于是有

$$\begin{pmatrix}S' & \\ & \frac{1}{\sqrt{c}}\end{pmatrix}\begin{pmatrix}A_1 & \\ & c\end{pmatrix}\begin{pmatrix}S' & \\ & \frac{1}{\sqrt{c}}\end{pmatrix}=I_n,$$

因此 $\begin{pmatrix}A_1 & \\ & c\end{pmatrix}$ 与 I_n 合同, 即为正定矩阵, 从而 A 正定. □

所以我们可以运用顺序主子式的方法来判断一个二次型正定.

6.5 线性代数中出现的内容

6.5.1 二次型及其标准形

1. 二次型

   (a) 含有 n 个变量 $x_1..x_n$ 的二次型多项式 (每一项都是二次的多项式)$f(x_1..x_n)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j$ 称为 n 的二次型, 令 $\boldsymbol{x}=(x_1..x_n)^T,\boldsymbol{A}=(a_{ij})$, 则二次型可以用矩阵乘法表示为 $f(x_1..x_n)=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$, 其中 A 是 n 阶实对称矩阵, 称 A 为二次型 $f(x_1..x_n)$ 的矩阵. 矩阵 $\boldsymbol{A}$ 的秩 $\mathrm{rk}(\boldsymbol{A})$ 称为二次型 f 的秩, 记作 $\mathrm{rk}(f)$.

2. 二次型的标准形

   (a) 如果二次型含有变量的平方项, 所有混合项 (如$x_ix_j,i\neq j$) 的系数都等于 0, 也就是 $f(x_1..x_n)=x^T\boldsymbol{A}x=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$, 其中 d_i 都是实数. 称这样的二次型为标准形.

   (b) Prop. 在标准形中, **正平方项的个数** p 称为二次型的**正惯性指数**, **负平方项**的个数 q 称为二次型的**负惯性指数**, 且 $r(f)+r(A)=p+q$.

   (c) Thm. 任意 n 元二次型 $x^{-1}\boldsymbol{A}x$ 都可以通过正交变换 $x=\boldsymbol{C}y$(C 是可逆矩阵) 化为标准形. 也就是 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\overset{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}}{\Longrightarrow}\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}=d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$, 其中 $\boldsymbol{A}=\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}$, 特别的, 存在正交变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$($\boldsymbol{C}$ 是正交矩阵) 化 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 为标准形.

6.5.2 用配方法化二次型为标准形

Case 1. 二次型中含有某变量 x_i 的平方项和交叉项: 先集中 x_i 的交叉项, 然后将 x_i^2 配方, 化为完全平方, 用新变量代替各个平方项中的变量, 同时立刻写出它的逆变换, 也就是用新的变量表示旧的变量. 一直这样做下去, 知道各个项数完全化为平方项即可.

Case 2. 二次型中没有平方项, 只有交叉项: 先利用平方差公式进行可逆变换的构造, 把二次型化为含平方项的二次型, 如当 x_ix_j 的系数 $a_{ij} \neq 0$ 时, 进行可逆线性变换:

$$x_i = y_i - y_j, x_j = y_i + y_j, x_k = y_k (k \neq i, j)$$

转化为情形 1 处理.

6.5.3 正定二次型

1. 正定二次型

   (a) Def. 设二次型矩阵 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$, 如果对于任意的 $\boldsymbol{x} \neq \boldsymbol{0}$, 恒有 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$, 则称二次型 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 为正定二次型, 正定二次型的矩阵 $\boldsymbol{A}$ 称为**正定矩阵**.

   (b) 二次型正定的充分必要条件

      i. $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 的正惯性指数 $p = n$

      ii. $\boldsymbol{A}$ 与 $\boldsymbol{E}$ 合同, 即有可逆矩阵 $\boldsymbol{C}$, 是的 $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} = \boldsymbol{E}$

      iii. $\boldsymbol{A}$ 的所有是特征值全大于 0

      iv. $\boldsymbol{A}$ 的顺序主子式全大于 0

      v. 存在可逆矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{A} = \boldsymbol{C}^T\boldsymbol{C}$

6.6 基础习题

Problem 6.1. 将二次型 $2x_1^2 - 2x_1x_2 + 2x_1x_3 - 2x_2x_3$ 通过正交变换化为标准形, 并且指出所用的正交变换.

Solution. 二次型矩阵为 $\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$, 它的特征值是 $3, -1, 0$. 对应的特征向量是 $(2, -1, 1)^T, (0, 1, 1)^T, (-1, -1, 1)^T$, 属于不同特征值的向量线性无关, 在这种情形下正交, 因此单位化, 得到正交矩阵 $P = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$, 得到的对角矩阵

$$\begin{pmatrix} 3 & & \\ & -1 & \\ & & 0 \end{pmatrix}.$$

Problem 6.2. 设三阶实对称矩阵 $\boldsymbol{A}$ 的特征值为 $6,3,3$, 与特征值 6 对应的特征向量 $\boldsymbol{p}_1 = (1,1,1)^T$, 求矩阵 $\boldsymbol{A}$.

Solution. 考虑实对称矩阵各个特征向量之间是正交的. 据此可以求出 $\boldsymbol{P}$.

Problem 6.3. 设 $\boldsymbol{A},\boldsymbol{B}$ 为同阶方阵, 证明: 如果 $\boldsymbol{A},\boldsymbol{B}$ 相似, 那么他们的特征多项式相等. 举反例说明逆命题不成立. 并且证明当 $\boldsymbol{A},\boldsymbol{B}$ 为实对称矩阵的时候, 逆命题成立.

Solution. 注意单位矩阵和任意矩阵相乘都是可逆的. 并借此提出两边的内容即可. 第二问举例参考 $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ 即可. 第三问参考二者都相似与对角形.

Problem 6.4. 证明: 上三角矩阵, 下三角矩阵, 对角矩阵的特征值是它对角线上的各个元素.

Solution. 通过特征多项式的计算可以知道.

Problem 6.5. 设二次型 $f(x_1,x_2,x_3) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 的秩为 1, $\boldsymbol{A}$ 中各行元素为 3, 则 f 在正交变换 $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ 下的标准形为.

第七章　Euclid 空间

7.1 Euclid 空间的基本概念

7.2 标准正交基

7.3 正交矩阵与正交变换

7.4 正规变换

7.5 酉空间

7.6 线性代数中出现的内容

7.6.1 向量的内积. 长度. 正交性

1. 内积
   (a) Def. 设 n 维向量 $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, 称 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i$ 为向量的内积.
   用矩阵的记号就是 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^T \boldsymbol{y}$.
   (b) Prop. 设 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ 为 n 维向量,$\lambda \in \mathbb{R}$
      i. 对称性: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$
      ii. 线性性: $\langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{y} \rangle; \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
      iii. 非负性: 当 $\boldsymbol{x} = \mathbf{0}$ 时 $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$, 否则 $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$.
      iv. Swartz 不等式: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \leq \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle$.
2. 长度, 夹角和正交
   (a) 长度. 设 $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 为 n 维向量, 称 $||\boldsymbol{x}|| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$ 是 x 的长度 (范数). 当 $||\boldsymbol{x}|| = 1$ 的时候称其为单位向量, 若 $\boldsymbol{x} \neq 0$, 称 $\boldsymbol{x}/||\boldsymbol{x}||$ 为单位向量, 这种操作称为把 $\boldsymbol{x}$ 单位化.
   (b) 夹角. 当 $\boldsymbol{x} \neq 0, \boldsymbol{y} \neq 0$ 时, 称 $\theta = \arccos \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{||\boldsymbol{x}|| \cdot ||\boldsymbol{y}||} (0 \leq \theta \leq \pi)$ 为向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的夹角
   (c) 正交. 当 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ 的时候, 称 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交. 零向量与任何向量正交.

(d) 规范正交基: 设 n 维向量 $\boldsymbol{e}_1..\boldsymbol{e}_r$ 是向量空间 $V(V \subset \mathbb{R}^n)$ 的一个基, 如果 $\boldsymbol{e}_1..\boldsymbol{e}_r$ 两两正交且都是单位向量, 则称 $\boldsymbol{e}_1..\boldsymbol{e}_r$ 是 V 的一个**规范正交基**.

3. 长度的性质
   (a) 非负性. 当 $\boldsymbol{x} = 0$ 时 $||\boldsymbol{x}|| = 0$, 否则 $||\boldsymbol{x}|| > 0$.
   (b) 齐次性 $||\lambda\boldsymbol{x}|| = |\lambda| \cdot ||\boldsymbol{x}||$.
   (c) 三角不等式 $||\boldsymbol{x}+\boldsymbol{y}|| \le ||\boldsymbol{x}|| + ||\boldsymbol{y}||$.
4. 正交向量组的性质
   (a) Th. 若 n 维向量是 $\boldsymbol{a}_1..\boldsymbol{a}_r$ 是一组两两正交的非零向量, 则 $\boldsymbol{a}_1..\boldsymbol{a}_r$ 线性无关
5. Schmidt 正交化
   (a) 设向量组 $\boldsymbol{a}_1..\boldsymbol{a}_r$ 线性无关, 经过 Schmidt 正交化将其化为两两正交的向量组 $\boldsymbol{b}_1..\boldsymbol{b}_n$, 二者是等价的. Schmidt 正交化的过程如下:

$$\begin{aligned}
\boldsymbol{b}_1 =& \boldsymbol{a}_1 \\
\boldsymbol{b}_2 =& \boldsymbol{a}_2 - \frac{\langle \boldsymbol{b}_1, \boldsymbol{a}_2 \rangle}{\langle \boldsymbol{b}_1, \boldsymbol{a}_1 \rangle}\boldsymbol{b}_1 \\
&\vdots \\
\boldsymbol{b}_r =& \boldsymbol{a}_r - \frac{\langle \boldsymbol{b}_1, \boldsymbol{a}_r \rangle}{\langle \boldsymbol{b}_1, \boldsymbol{a}_1 \rangle}\boldsymbol{b}_1 - \frac{\langle \boldsymbol{b}_2, \boldsymbol{a}_r \rangle}{\langle \boldsymbol{b}_2, \boldsymbol{a}_2 \rangle}\boldsymbol{b}_2 - \cdots - \frac{\langle \boldsymbol{b}_{r-1}, \boldsymbol{a}_r \rangle}{\langle \boldsymbol{b}_{r-1}, \boldsymbol{a}_{r-1} \rangle}\boldsymbol{b}_{r-1}
\end{aligned}$$

6. 正交矩阵
   (a) Def. 如果 $\boldsymbol{A}$ 是 $\boldsymbol{n}$ 阶矩阵, $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{E}$, 也就是 $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$, 称 $\boldsymbol{A}$ 为正交矩阵, $\boldsymbol{P}$ 为正交矩阵, 则线性变换 $\boldsymbol{y} = \boldsymbol{P}\boldsymbol{x}$ 是正交变换.
   (b) Prop
      i. $\boldsymbol{A}$ 正交 $\Leftrightarrow$ $\boldsymbol{A}$ 的向量都是单位向量且两两正交
      ii. $\boldsymbol{A}$ 正交 $\Leftrightarrow$ 行向量都是单位向量且两两正交
      iii. 正交矩阵的行列式是 1 或者 -1.
      iv. $\boldsymbol{A}, \boldsymbol{B}$ 是正交矩阵, 那么 $\boldsymbol{AB}$ 也是正交矩阵
      v. 若 $\boldsymbol{y} = \boldsymbol{P}\boldsymbol{x}$ 是正交变换, $||\boldsymbol{y}|| = ||\boldsymbol{x}||$.

7.7 基本习题

Problem 7.1. TODO