代数 0 课程讲义

Instructor: 余成龙

Notes Taker: 刘博文

Qiuzhen College, Tsinghua University

2022 Spring

目录

第一章　线性方程组

1.1　线性函数与线性方程

定义 1.1.1. 对于 $\mathbb{R}^n$ 上的函数 $F\colon \mathbb{R}^n \to \mathbb{R}$, 如果存在 $a_1, \ldots, a_n \in \mathbb{R}$ 是常数, 使得 F 有如下表达式

$$F(x) = a_1 x_1 + \cdots + a_n x_n$$

那么称 F 是 $\mathbb{R}^n$ 上的**线性函数** (linear function).

例子. *如下的 F 是 $\mathbb{R}^2$ 上的线性函数:*

$$\begin{aligned} &F\colon \mathbb{R}^2 \to \mathbb{R} \\ &\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto (x_1+1)^2 - (x_1-1)^2 + (x_2-1)^2 - (x_1+1)^2 \end{aligned}$$

例子. *如下的 F 不是 $\mathbb{R}^2$ 上的线性函数*

$$\begin{aligned} &F\colon \mathbb{R}^2 \to \mathbb{R} \\ &\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2^2 \end{aligned}$$

定理 1.1.2. *函数 $F\colon \mathbb{R}^n \to \mathbb{R}$ 是线性函数当且仅当 F 满足*

(1) *对任意 $x, y \in \mathbb{R}^n$, 有*

$$F(x+y) = F(x) + F(y)$$

(2) *对任意 $c \in \mathbb{R}, x \in \mathbb{R}^n$, 有*

$$F(cx) = cF(x)$$

证明: 如果 F 是线性函数, 可以直接验证 F 满足 (1), (2) 两条性质；另一方面, 如果 F 满足 (1),(2), 我们记

$$a_1 = F(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}), \quad a_2 = F(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}), \quad \ldots$$

则

$$F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = a_1 x_1, \quad F(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}) = a_2 x_2, \quad \ldots$$

那么任取 $v_1, \ldots, v_m \in \mathbb{R}^n$, 则

$$\begin{aligned} F(v_1 + v_2 + \cdots + v_m) &= F((v_1 + \cdots + v_{m-1}) + v_m) \\ &= F(v_1 + \cdots + v_{m-1}) + F(v_m) \\ &= F(v_1) + \cdots + F(v_m) \end{aligned}$$

因此,

$$F(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}) = F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}) = F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) + \cdots + F(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}) = a_1x_1 + \cdots + a_nx_n$$

□

定义 1.1.3. $\mathbb{R}^n$ 上 m 个线性函数 $F_1, \ldots, F_m$ 和 m 个实数 $b_1, \ldots, b_m$ 满足的方程组

$$\begin{cases} F_1(x) = b_1 \\ \vdots \\ F_m(x) = b_m \end{cases}$$

称为 n 个变元的**线性方程组** (system of linear equations), 带入方程组使得其成立的 x 称为**线性方程组的解** (solution of system of linear equations)

注记. 我们可以给线性方程组如下的一些几何解释:

(1) 在 $\mathbb{R}^2$ 中, 单个线性函数 $F_1(x) = a_1x_1 + a_2x_2$ 以及实数 b_1 给出的线性方程组 $a_1x_1 + a_2x_2 = b_1$ 的解是 $\mathbb{R}^2$ 中的一条直线.

(2) 在 $\mathbb{R}^2$ 中, 根据 (1) 的几何解释不难理解如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

的解是 $\mathbb{R}^2$ 中两条直线的交点. 注意, 在 $\mathbb{R}^2$ 中两条直线不一定相交, 即如上线性方程组不一定有有解, 但是如果有解一定只有唯一解.

(3) 在 $\mathbb{R}^3$ 中, 如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases}$$

的解可以看成是 $\mathbb{R}^3$ 中两个平面的交线. 注意: 在 $\mathbb{R}^3$ 中两个平面不一定相交, 即如上线性方程组不一定有有解, 并且如果相交, 也是交出一条线, 即此时解不唯一.

(4) 在更高维中也有同样的解释: 由一个线性函数给出的线性方程的解可以看成是一个低一维的超平面, 而多个线性函数给出的线性方程组的解则是这些超平面的交.

1.2 高斯消元法

根据注记1.1可知对于一个线性方程组其可能没有解, 并且即使有解也不一定只有唯一解, 那么该如何求解线性方程组呢? 在本节中我们将利用高斯消元法, 来求解一般的线性方程组. 我们先来看下面的一个简单的例子.

例子.

$$\begin{cases} 4x_2 - x_3 = 7 & (r_1) \\ x_1 + 2x_2 = 5 & (r_2) \\ 2x_1 + x_3 = 3 & (r_3) \end{cases}$$

显然我们交换 r_1, r_2 不改变上述方程组的解, 因此我们得到:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1') \\ 4x_2 - x_3 = 7 & (r_2') \\ 2x_1 + x_3 = 3 & (r_3') \end{cases}$$

我们考虑如下操作: 保持 r_1', r_2' 不变, 用 r_3' 减去 $2r_1'$, 得到如下的方程组:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1'') \\ 4x_2 - x_3 = 7 & (r_2'') \\ -4x_2 + x_3 = -7 & (r_3'') \end{cases}$$

上述操作并不改变方程组的解, 因为可由 r_1'', r_2'', r_3'' 恢复出 r_1', r_2', r_3'. 类似的最后再保持 r_1'', r_2'' 不变, 用 r_3'' 加上 r_2'', 得到

$$\begin{cases} x_1 + 2x_2 = 5 \\ 4x_2 - x_3 = 7 \\ 0 = 0 \end{cases}$$

对于上述方程组我们可以用 x_3 来如下的表示 x_1, x_2, 其中 x_3 可以取任意的实数

$$x_1 = -\frac{1}{2}x_3 + \frac{3}{2}$$
$$x_2 = \frac{1}{4}x_3 + \frac{7}{4}$$

因此我们可以将方程组的解写作

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} \\ \frac{7}{4} \\ 0 \end{pmatrix}$$

注记. 根据上述结果可以发现该线性方程组有无穷组解, 这对应于几何解释中 $\mathbb{R}^3$ 中三个平面相交出一条线.

回顾例1.2, 在解方程中我们主要用到了如下三种操作:

(E1) 交换方程组的某两行.
(E2) 将某一行乘以非零常数 c.
(E3) 将某一行的非零常数 c 倍加到另一行上.

我们称如上的三种操作为**基础行变换** (elementary row operations). 不难发现基础行变换均可逆, 并且其逆也是基础行变换.

定义 1.2.1. 有限个基础行变换的复合称为**行变换** (row operations).

命题 1.2.2. *行变换均可逆, 并且其逆也为行变换.*

证明: 因为基础行变换可逆, 且其逆也为基础行变换, 并且操作 O_1O_2 的逆为 $O_2^{-1}O_1^{-1}$. □

推论 1.2.3. *行变换不改变线性方程组的解.*

由于作行变换只关注方程的系数以及右侧常数项, 因此对于如下的线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

我们将其系数及常数项提出来记作

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

并将这个线性方程组记做 $Ax = b$, 这也引出了矩阵的概念.

定义 1.2.4. 由 $m \times n$ 个数 a_{ij} 排成的 m 行 n 列的 (实) 数表称为 m 行 n 列**矩阵** (matrix), 记做 $(a_{ij}) \in M_{m\times n}(\mathbb{R})$.

对于线性方程组 $Ax = b$, A 称为**系数矩阵** (coefficient matrix), (A, b) 称为**增广矩阵** (augmented matrix), 并将上述方程记作 $Ax = b$. 现在我们即可以通过行变换来操作我们的增广矩阵, 使其最终的形式方便于我们求解, 那么究竟该操作到什么样子为止呢?

根据例子1.2, 我们可以观察出行变换的目标为:

1. 所有非零行在零行的上面.
2. 对某一非零行, 称最左边的非零元为**主元** (pivot), 第 i 行的主元严格比第 $i+1$ 行的主元靠左.

满足上述条件的增广矩阵称为**行阶梯型** (row echelon form), 并且如果主元所在列的其他元素均为零, 主元本身为 1, 则称此时为**最简行阶梯型** (reduced row echelon form).

定理 1.2.5. *矩阵 A 可通过行变换变成最简行阶梯型, 并且该最简行阶梯型不依赖于行变换的选取, 记作* $\mathrm{rref}(A)$.

证明: 对 $m \times n$ 矩阵的列作归纳: 假设 $n = 1$ 时, 对于 $m \times 1$ 矩阵

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

如果 $a_{11} = \cdots = a_{m1} = 0$, 则此时已经是最简行阶梯型. 若 $a_{11} = a_{21} = \cdots = a_{(k-1)1} = 0, a_{k1} \neq 0$, 那么通过 (E1) 将 a_{k1} 换到第一行, 用 (E2) 将第一行乘以 $(a_{k1})^{-1}$ 使得主元变为 1, 再用 (E3) 将第一行以下变为零, 因此此时最简行阶梯型为

$$\operatorname{rref}(A) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

并且不依赖于行列变换的选取. 假设对列数为 n 的时候成立, 对于 $m \times (n+1)$ 的矩阵 A, 将其写作 $A = (B, y)$, 其中 B 是 $m \times n$ 矩阵. 根据归纳假设 B 可由行变换得到最简行阶梯型, 记作 B', 将同样的变换作用在 A 上得到 $A' = (B', y')$. 如果 B' 没有非零行, 则次数 A' 已经是最简行阶梯型. 如果 B' 从 $k+1$ 行开始是零行, 则对

$$\begin{pmatrix} y'_{k+1} \\ \vdots \\ y'_m \end{pmatrix}$$

应用 $n = 1$ 时的结论, 可做行变换得到最简行阶梯型, 同时也对 B' 作. 但由于行变换不改变零矩阵, 因此不改变 B', 得到的矩阵记作 A''. 考虑如下两种情况:

(1) 如果

$$\operatorname{rref}(\begin{pmatrix} y'_{k+1} \\ \vdots \\ y'_m \end{pmatrix}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

则此时 A'' 已经是最简行阶梯型.

(2) 如果

$$\operatorname{rref}(\begin{pmatrix} y'_{k+1} \\ \vdots \\ y'_m \end{pmatrix}) = \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$$

则作 (E3) 将第 $k+1$ 行加到第 $1, 2, \ldots, k$ 行, 将 $y'_1, \ldots, y'_k$ 变成零, 此时得到的矩阵也是最简行阶梯型. 由于上述操作只依赖于 B 和 y, 即 A 的最简行阶梯型只依赖于 A 本身.

□

定义 1.2.6. 对于矩阵 A, $\operatorname{rref}(A)$ 的主元个数称为 A 的秩 (rank), 记作 $\operatorname{rank} A$.

注记. 根据定义, 对于 $m \times n$ 矩阵, $\operatorname{rank} A \leq m$.

命题 1.2.7. *对于矩阵 A, B, 有 $\operatorname{rank} AB \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}$*

1.3 线性方程组解的结构

定义 1.3.1. 对于线性方程组的系数矩阵 A, $\mathrm{rref}(A)$ 中主元对应的未知元称为**主元** (principal unknowns), 其余未知元称为**自由元** (free unknowns).

例子. 例如线性方程组 $Ax = b$, 其中

$$A = \begin{pmatrix} 1 & 0 & 3 & 5 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix}$$

则 x_1, x_2, x_5 是主元, x_3, x_4, x_6 是自由元. 并且根据上述最简行阶梯型, 我们可以直接分析出方程组的解的情况:

1. 如果 b_4 或者 b_5 不是零, 则方程组 $Ax = b$ 无解.
2. 如果 $b_4 = b_5 = 0$, 则 x_3, x_4, x_6 取定任意实数后, 主元由方程组唯一确定:

$$\begin{aligned} x_1 &= b_1 - 3x_3 - 5x_4 - x_6 \\ x_2 &= b_2 - 2x_3 - x_4 - 2x_6 \\ x_5 &= b_3 \end{aligned}$$

定理 1.3.2. 对于方程 $Ax = b$, 用行变换将 (A, b) 化作最简行阶梯型 $(\overline{A}, \overline{b})$, 则方程有解等价于 $\overline{A}$ 的零行对应的 $\overline{b}_i$ 也是零. 并且有解时自由元可以任意取值, 且自由元的每一组取值都唯一决定了一组解.

推论 1.3.3. 线性方程组 $Ax = b$

(1) 有解当且仅当 $\mathrm{rank}\, A = \mathrm{rank}(A, b)$.

(2) 有唯一解当且仅当 $\mathrm{rank}\, A$ 等于 A 的列数相同.

定义 1.3.4. 方程 $Ax = 0$ 称为**齐次线性方程组** (system of homogeneous linear equations).

定理 1.3.5. 齐次线性方程组 $Ax = 0$ 的解在加法和数乘下封闭.

证明: 注意到

$$\begin{aligned} A(x + y) &= Ax + Ay \\ A(cx) &= cAx \end{aligned}$$

□

定理 1.3.6. 对于线性方程组 $Ax = b$, 如果 $\widetilde{x}$ 是其某一解 (特解), 则 $Ax = b$ 的所有解均可唯一的表达为 $x = y + \widetilde{x}$, 其中 y 是 $Ax = 0$ 的解.

证明: 只需验证如下两点:

(1) 验证 $y + \widetilde{x}$ 是解.

(2) 验证当 x 是解时, $x = (x - \widetilde{x}) + \widetilde{x}$, 其中 $x - \widetilde{x}$ 满足 $Ax = 0$.

□

注记. 从几何上来看, 齐次线性方程组 $Ax = 0$ 的解构成了 $\mathbb{R}^n$ 中的一个对加法数乘封闭的子集, 称为一个子空间. 而 $Ax = b$ 的解相当于是将这个子空间做了平移.

第二章　矩阵及其运算

2.1　矩阵乘法

在求解线性方程组的时候, 我们已经接触到了 $m \times n$ 阶矩阵的概念, 即定义1.2.4. 当 $m = n$ 时, $n \times n$ 阶矩阵被称为 n 阶**方阵** (square matrix). 在 $M_{m\times n}(\mathbb{R})$ 上有如下的运算

(1) 加法: $A = (a_{ij})_{m\times n}, B = (b_{ij})_{m\times n}$, 则 $A + B := (a_{ij} + b_{ij})_{m\times n}$.

(2) 数乘: $A = (a_{ij})_{m\times n}, c \in \mathbb{R}$, 则 $cA := (ca_{ij})_{m\times n}$.

并且对于 $A \in M_{m\times n}(\mathbb{R}), B \in M_{n\times l}(\mathbb{R})$, 可以定义矩阵的乘法为 $AB := (c_{ij})_{m\times l}$, 其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

即用 A 的第 i 行去乘 B 的第 j 列然后作求和. 并且注意只有 A 的列数与 B 的行数相同时, 其才能做矩阵乘法, 否则无意义. 因此在 $M_{n\times n}(\mathbb{R})$ 上还有第三种运算

(3) 乘法: $A = (a_{ij})_{n\times n}, B = (b_{ij})_{n\times n}$, 则 $AB := (\sum_{k=1}^{n} a_{ik}b_{kj})_{n\times n}$.

注记. 对于矩阵 $A = (a_{ij})_{m\times n}, x = (x_i)_{n\times 1}$ 以及 $b = (b_i)_{n\times 1}$, 矩阵乘法的等式 $Ax = b$ 给出了一个线性方程组, 这也是为什么我们之前如此记线性方程组的原因.

命题 2.1.1. *矩阵乘法具有结合律.*

证明: 假设 $A, B, C \in M_{n\times n}(\mathbb{R})$, 不妨记 $C = (c_1, \ldots, c_n)$, 其中 c_i 是列向量. 那么

$$\begin{aligned}(AB)C &= ((AB)c_1, \ldots, (AB)c_n)\\ A(BC) &= A(Bc_1, \ldots, Bc_n) = (A(Bc_1), \ldots, A(Bc_n))\end{aligned}$$

因此只需对每个 c_i 验证 $(AB)c_i = A(Bc_i)$ 即可, 因此我们不妨假设 C 是 $n \times 1$ 的矩阵. 将 A 写作

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

其中 a_i 是行向量, 那么

$$(AB)C = \begin{pmatrix} a_1B \\ a_2B \\ \vdots \\ a_nB \end{pmatrix} C = \begin{pmatrix} a_1BC \\ a_2BC \\ \vdots \\ a_nBC \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} a_1(BC) \\ a_2(BC) \\ \vdots \\ a_n(BC) \end{pmatrix}$$

因此只需要对每一个 a_i 验证即可, 因此我们不妨假设 A 是 $1 \times n$ 的矩阵, 那么

$$((a_1, \ldots, a_n)B) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \sum_{j=1}^{n} (\sum_{i=1}^{n} a_i b_{ij}) c_j = \sum_{i,j} a_i b_{ij} c_j$$

$$(a_1, \ldots, a_n)(B \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n) \end{pmatrix} = \sum_{j=1}^{n} a_i (\sum_{i=1}^{n} b_{ij} c)_j = \sum_{i,j} a_i b_{ij} c_j$$

□

有了矩阵乘法, 我们可以将之前对系数矩阵做的初等行变换用矩阵的语言来再说一遍.

定义 2.1.2. 如下的三类矩阵被称为**初等矩阵** (elementary matrix)

(E1)

$$E[ij] = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

即交换 I_n 的第 i 行与第 j 行得到的矩阵.

(E2)

$$E[i, c] = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

即将 I_n 的第 i 行乘以 c 得到的矩阵, 其中 $c \neq 0$.

(E3)

$$E[ij,c]=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&\vdots&\ddots&&&\\&&c&\cdots&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}$$

即将 I_n 的第 i 行乘以 c 加到第 j 行得到的矩阵, 其中 $c\neq 0$.

例子. 假设系数矩阵

$$A=\begin{pmatrix}r_1\\r_2\\r_3\end{pmatrix}$$

那么如果我们对其作用 *(E3)* 将第三行的二倍加到第一行上去, 得到的新的系数矩阵记做 A', 那么

$$A'=\begin{pmatrix}r_1+2r_3\\r_2\\r_3\end{pmatrix}=\begin{pmatrix}1&0&2\\0&1&0\\0&0&1\end{pmatrix}A$$

因此我们可以看出初等行变换 *(E3)* 可以看作是初等矩阵 *(E3)* 左乘.

命题 2.1.3. *对 $A\in M_{m\times n}(\mathbb{R})$ 做初等行变换 O 等价于对左乘相应的初等矩阵 B.*

证明: 根据定义验证即可. □

推论 2.1.4. *对 A 做行变换 $O_1\dots O_k$ 等价于左乘初等矩阵 $B_k\dots B_2B_1$.*

注意到我们的初等行变换是可逆的, 用矩阵的语言来说, 对于初等矩阵 $B\in M_{n\times n}(\mathbb{R})$, 总存在另一个初等矩阵 B' 使得 $BB'A=\mathrm{I}_n$, 其中 I_n 是只有对角线为 1, 其余地方全为零的 $n\times n$ 矩阵.

定义 2.1.5. 对于矩阵 $A\in M_{n\times n}(\mathbb{R})$

(1) 若有 B 使得 $BA=\mathrm{I}_n$, 则成 B 为 A 的**左逆** (left inverse).

(2) 若有 C 使得 $AC=\mathrm{I}_n$, 则成 C 为 A 的**右逆** (right inverse).

(3) 如果左逆右逆均存在, 则称 A **可逆** (invertible).

定理 2.1.6. *对于矩阵 $A\in M_{n\times n}(\mathbb{R})$, 如下叙述等价:*

(1) *A 可逆.*

(2) *A 存在左逆.*

(3) *A 存在右逆.*

(4) $\mathrm{rref}(A)=\mathrm{I}_n$.

(5) *$Ax=b$ 有唯一解.*

(6) *$Ax=0$ 有唯一解.*

(7) $\operatorname{rank} A = n$.

证明: 根据线性方程组解的结构定理, 即定理1.3.2, 我们已经见到了 (4),(5),(6),(7) 的等价性.

(2) $\Longrightarrow$ (6): 如果 A 存在左逆, 那么 $A^{-1}Ax = 0$ 意味着 $x = 0$, 即 $Ax = 0$ 只有唯一的解.

(5) $\Longrightarrow$ (3): 如果 $Ax = b$ 存在唯一解, 那么我们不妨取

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots$$

那么不妨记 $Ax = b_1, Ax = b_2, \dots, Ax = b_n$ 唯一的解分别是 $w_1, \dots, w_n$, 则 $A(w_1, \dots, w_n) = \mathrm{I}_n$, 即 $(w_1, \dots, w_n)$ 是 A 的右逆.

(3) $\Longrightarrow$ (2): 假设 A 有右逆, 存在 C 使得即 $AC = \mathrm{I}_n$, 从而 $CAC = C$. 另一方面, 由于 C 存在左逆, 从而 C 存在右逆, 不妨记为 D, 因此

$$CA = CACD = CD = \mathrm{I}_n$$

即 C 也是 A 的左逆. 从这个证明可以看出, 如果 A 存在左逆, 那么其右逆不仅存在, 并且还和左逆相同, 从而 A 可逆, 这也证明了 (3) $\Longrightarrow$ (1), 类似的我们可以证明 (2) $\Longrightarrow$ (1), 而 (1) $\Longrightarrow$ (2) 和 (1) $\Longrightarrow$ (3) 根据定义即可. □

命题 2.1.7. *若 A 可逆, 则左逆与右逆均唯一存在且相同, 记做 A^{-1}.*

证明: 我们只需要证明如果 A 可逆, 那么其左逆右逆都唯一. 假设 C 是 A 的一个左逆, D 是 A 的一个右逆, 那么

$$C = C\,\mathrm{I}_n = CAD = \mathrm{I}_n\,D = D$$

即 A 的任何左逆与右逆都相同. 那么假设 C_1, C_2 是 A 的两个左逆, 由于 C_1 也是 A 的右逆, 从而 $C_1 = C_2$, 即 A 的左逆唯一, 类似的, 我们也可以说明 A 的右逆唯一. □

注记. 上述结论表明, 如果 A 可逆, 那么 $\operatorname{rref}(A) = \mathrm{I}_n$, 因此将其化为最简行阶梯型的初等矩阵的乘积就是 A^{-1}. 那么我们该如何将这些初等矩阵的乘积记录下来呢? 考虑矩阵 (A, I_n), 对其进行操作使得 A 化为最简行阶梯型则有 (I, A^{-1}), 这也给出了我们求逆的办法, 并且我们也有如下简单的推论.

推论 2.1.8. *矩阵 A 可逆当且仅当其为初等矩阵的乘积.*

例子. *考虑 $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, 则*

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

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

注记. 更一般的, 如果 2×2 矩阵可逆, 那么

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

2.2 矩阵的转置

定义 2.2.1. 对于矩阵 $A = (a_{ij})_{m\times n}$, 其**转置矩阵** (transpose matrix), 记做 A^T, 是一个 $n \times m$ 阶矩阵 $(b_{ij})_{n\times n}$, 其中 $b_{ij} = a_{ji}$.

例子. *对于列向量来说, 其转置为行向量; 对于行向量来说, 其转置为列向量.*

命题 2.2.2. *对于矩阵转置来说, 我们有如下简单的性质:*

(1) $(A^T)^T = A$.

(2) $(AB)^T = B^T A^T$.

(3) $AA^T = 0$ *当且仅当* $A = 0$.

证明: 直接验证即可. □

推论 2.2.3. *对* $A \in M_{m\times n}(\mathbb{R})$ *做列变换等价于右乘可逆矩阵* $B \in M_{n\times n}(\mathbb{R})$.

证明: 利用转置矩阵的观点, 对矩阵 A 进行列变换, 等价于对 A^T 进行行变换再转置, 而列变换等价于左乘可逆矩阵, 因此根据命题2.2.2的 (2) 即可. □

回忆定义1.2.6, 我们定义矩阵 A 的秩为其最简行阶梯型的主元个数. 一个自然的问题就是 A^T 的秩与 A 的秩有什么关系呢?[1] 我们可以证明 $\operatorname{rank} A = \operatorname{rank}(A^T)$, 这主要依赖于下面的定理.

定理 2.2.4. *列变换不改变矩阵* A *的秩.*

证明: 假设 A 是 $m \times n$ 阶矩阵, 根据推论2.2.3, 我们只需要对可逆矩阵 $B \in M_{n\times n}(\mathbb{R})$ 证明 $\operatorname{rank} A = \operatorname{rank}(AB)$. 我们不妨记 $\operatorname{rank} A = k, \operatorname{rank}(AB) = l$. 根据线性方程组解的理论可知

1. $Ax = 0$ 有主元 $x_{i_1}, \ldots, x_{i_k}$ 以及自由元 $x_{i_{k+1}}, \ldots, x_{i_n}$.
2. $ABy = 0$ 有主元 $y_{i_1}, \ldots, y_{i_l}$ 以及自由元 $y_{i_{l+1}}, \ldots, y_{l_n}$.

由于 $Ax = 0$ 的解与 $ABy = 0$ 的解之间满足 $x = By$, 由于 B 是可逆矩阵, 根据定理 2.1.6 可知两者解之间存在一一对应, 从而主元与自由元的情况是相同的, 从而 $k = l$. □

推论 2.2.5. *对于矩阵* A *来说,* $\operatorname{rank} A = \operatorname{rank}(A^T)$.

证明: 我们对 A 的最简行阶梯型 $\operatorname{rref}(A)$ 作列变换, 将其化作如下形式

$$\begin{pmatrix} \mathrm{I}_k & \mathrm{O}_{k\times n-k} \\ \mathrm{O}_{m-k\times k} & \mathrm{O}_{m-k\times n-k} \end{pmatrix}$$

其中 O 代表分量全为零的矩阵. 此时 A 与 A^T 都为最简行阶梯型, 从而 $\operatorname{rank} A = \operatorname{rank}(A^T) = k$. □

从上述证明过程中, 根据可逆矩阵与行列变换的关系, 我们还能看出:

[1]在一些教材中我们这里定义的矩阵的秩又被称为行秩, A^T 的秩被称为 A 的列秩, 即我们要证明矩阵的行秩与列秩相同.

推论 2.2.6. 对于矩阵 $A \in M_{m\times n}$, 存在可逆矩阵 P, Q 使得

$$PAQ = \begin{pmatrix} \mathrm{I}_k & \mathrm{O}_{k\times n-k} \\ \mathrm{O}_{m-k\times k} & \mathrm{O}_{m-k\times n-k} \end{pmatrix}$$

其中 $k = \operatorname{rank} A$, 这被称为 A 的**相抵标准型** *(canonical form)*.

定义 2.2.7. 矩阵 A, B 之间被称为**相抵** (equivalent), 如果存在可逆矩阵 P, Q 使得 $PAQ = B$.

定理 2.2.8. $m \times n$ 阶矩阵 A, B 之间相抵当且仅当 $\operatorname{rank} A = \operatorname{rank}(B)$, 即相抵关系完全由矩阵的秩分类.

证明: 由于行列变换不改变矩阵的秩, 从而相抵矩阵有相同的秩; 另一方面, 如果 A, B 有相同的秩, 它们的相抵标准型相同, 从而相抵. □

2.3 分块矩阵

一般来说, 当 n 较大时, 求解 $n \times n$ 矩阵的逆对人工操作来说是相对较麻烦的, 但如果矩阵有相对较好的形式, 此时的求解也可以化简. 下面将介绍分块矩阵的想法, 给定矩阵 A, 我们可以做适当的划分, 将其看作矩阵元素是矩阵的矩阵. 例如

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

我们可以将其看成 2×2 的矩阵 $(A_{ij})_{2\times 2}$, 其中

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad A_{12} = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} \quad A_{21} = \begin{pmatrix} a_{31} & a_{32} \end{pmatrix} \quad A_{22} = \begin{pmatrix} a_{33} \end{pmatrix}$$

如果可逆矩阵 A 可以写成分块对角的形式, 即

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & A_3 & \\ & & & A_4 \end{pmatrix}$$

那么则有

$$A^{-1} = \begin{pmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & A_3^{-1} & \\ & & & A_4^{-1} \end{pmatrix}$$

同样的, 我们可以对分块矩矩阵进行分块行列变换, 得到相对较好的形式. 例如对于分块矩阵

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

其中 A,B,C,D 都是方阵. 如果 A 可逆, 那么

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \mathrm{I} & -A^{-1}B \\ 0 & \mathrm{I} \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D-CA^{-1}B \end{pmatrix}$$

即通过行列变换将其下三角化, 对于 B,C,D 可逆的时候我们也可以做类似的事情. 特别地是, 如果我们采取不同的变换得到相同的等式, 这有时候可以给我们带来一些非平凡的结果.

例子. 对于列向量 α,β, 考虑

$$\begin{pmatrix} \mathrm{I} & \alpha \\ \beta^T & 1 \end{pmatrix}^{-1}$$

一方面我们考虑

$$\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ \beta^T & 1 & \mathrm{O} & 1 \end{pmatrix} \to \begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ 0 & 1-\beta^T\alpha & -\beta^T & 1 \end{pmatrix} \to \begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ 0 & 1 & -\beta^T(1-\beta^T\alpha)^{-1} & (1-\beta^T\alpha)^{-1} \end{pmatrix}$$
$$\to \begin{pmatrix} \mathrm{I} & 0 & \mathrm{I}+\alpha(1-\beta^T\alpha)^{-1}\beta^T & -(1-\beta^T\alpha)^{-1}\alpha \\ 0 & 1 & -(1-\beta^T\alpha)^{-1}\beta^T & (1-\beta^T\alpha)^{-1} \end{pmatrix}$$

另一方面我们有

$$\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ \beta^T & 1 & \mathrm{O} & 1 \end{pmatrix} \to \begin{pmatrix} \mathrm{I}-\alpha\beta^T & 0 & \mathrm{I} & -\alpha \\ \beta^T & 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} \mathrm{I} & 0 & (\mathrm{I}-\alpha\beta^T)^{-1} & -(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \\ \beta^T & 1 & 0 & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} \mathrm{I} & 0 & (\mathrm{I}-\alpha\beta^T)^{-1} & -(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \\ 0 & 1 & -\beta^T(\mathrm{I}-\alpha\beta^T)^{-1} & 1+\beta^T(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \end{pmatrix}$$

从而我们有非平凡等式

$$(\mathrm{I}-\alpha\beta^T)^{-1} = \mathrm{I}+\alpha(1-\beta^T\alpha)^{-1}\beta^T$$
$$(1-\beta^T\alpha)^{-1} = 1+\beta^T(\mathrm{I}-\alpha\beta^T)^{-1}\alpha$$

2.4 矩阵的行列式

回忆在注2.1中, 我们有

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

因此 2×2 矩阵可逆当且仅当 $ad-bc\neq 0$. 实际上, $ad-bc$ 有如下的几何含义: 考虑在 $\mathbb{R}^2$ 中由列向量 $v=\binom{a}{c}, w=\binom{b}{d}$ 围成的平行四边形, 向量 v,w 的夹角 θ 满足

$$\cos\theta = \frac{ab+cd}{\sqrt{a^2+c^2}\sqrt{b^2+d^2}}$$

从而 v,w 围成的平行四边形面积为

$$\begin{aligned} S &= \sqrt{1-\cos\theta^2}\sqrt{a^2+c^2}\sqrt{b^2+d^2} \\ &= \sqrt{(ad-bc)^2} \\ &= |ad-bc| \end{aligned}$$

注意 $ad-bc>0$ 与 <0 两种情况分别对应了 v 在 w 左侧或右侧两种情况, 因此 $ad-bc$ 可以看作是 v,w 围成的平行四边形的 "有向面积", 我们将要定义的行列式, 就是这种有向面积的高维推广.

定义 2.4.1. $\mathbb{R}^n$ 上的 n 次**多重反对称线性函数** (multiple skew symmetric linear function)(多重反对称线性函数, multiple skew symmetric linear function) 是函数

$$f:\underbrace{\mathbb{R}^n\times\cdots\times\mathbb{R}^n}_{n\text{个}}\to\mathbb{R}$$

满足条件

(1) $f(v_1,\dots,cv_i,\dots,v_n)=cf(v_1,\dots,v_n)$, 其中 $c\in\mathbb{R}$.

(2) $f(v_1,v_2,\dots,v_i'+v_i,\dots,v_n)=f(v_1,\dots,v_i',\dots,v_n)+f(v_1,\dots,v_i,\dots,v_n)$.

(3) $f(v_1,\dots,v_i,\dots,v_j,\dots,v_n)=-f(v_1,\dots,v_j,\dots,v_i,\dots,v_n)$.

(4) $f(e_1,\dots,e_n)=1$.

定理 2.4.2. *$\mathbb{R}^n$ 上的 n 次多重反对称线性函数存在且唯一.*

证明: 对 n 进行归纳: 当 $n=1$ 时, 由于 $f(e_1)=1$ 以及任取 $v\in\mathbb{R}$ 我们有 $v=ce_1$, 从而 $f(v)=c$, 即此时被唯一确定. 假设当 $n<k$ 时假设成立, 考虑 $n=k$ 的情形. 任取 $v=(v_1,\dots,v_k)\in\mathbb{R}^k$, □

定义 2.4.3. 假设 f 是 $\mathbb{R}^n$ 上 n 次多重反对称线性函数, 对于 $A\in M_{n\times n}(\mathbb{R})$, 记 $A=(v_1,\dots,v_n)$, 则 A 的**行列式** (determinant)定义为

$$\det A=|A|:=f(v_1,\dots,v_n)$$

推论 2.4.4.

例子. 当 $n=2$ 时

$$|\begin{pmatrix}a&b\\c&d\end{pmatrix}|=ad-bc$$

例子. 当 $n=3$ 时

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

注记. 上述公式又被称为对角线法则.

例子. 对于三种初等矩阵[2], 其行列式分别为

(E1) $\det E[ij]=1$.

(E2) $\det E[i,c]=c$.

(E3) $\det E[ij,c]=1$.

[2]见定义2.1.2

特别地，对于矩阵 A 以及初等矩阵 E，有 $|AE| = |A||E|$.

命题 2.4.5. 行列式有如下性质：

(1) 如果 A 的某一列为零，则 $|A| = 0$.

(2) $\det A \neq 0$ 当且仅当 $\operatorname{rank} A = n$.

(3) $|\,\mathrm{I}_n\,| = 1$.

(4) $|AB| = |A||B|$.

(5) $|A^{-1}| = |A|^{-1}, |PAP^{-1}| = |A|$.

(6) $|A| = |A^T|$.

(7) 假设 A 是分块上三角矩阵，并且对角线分块矩阵为 $A_1, \ldots, A_n$，则 $|A| = |A_1| \ldots |A_n|$. 特别地，如果 A 是上三角矩阵，并且对角线元素为 $a_1, \ldots, a_n$，则 $|A| = a_1 \ldots a_n$.

证明： (1). 根据定义2.4.1中 (1) 即可.

(2). 根据例2.4可知，如果 E 是初等矩阵，则 $|AE| = |A||E|$，因此因此不妨将 A 写作 $E_k \ldots E_1 \operatorname{rref}(A)$，其中 E_i 是初等矩阵. 而 $\operatorname{rank} A = n$ 当且仅当 $\operatorname{rref}(A)$ 没有零列，并且由于 $|E_i| \neq 0$，从而根据 (1) 可知 $|A| \neq 0$ 当且仅当 $\operatorname{rank} A = n$.

(3). 根据定义2.4.1中 (4) 即可.

(4). 假设 B 不可逆，即 $\operatorname{rank} B < n$，根据 (2) 则有 $|B| = 0$. 而根据命题1.2.7有 $\operatorname{rank} AB \leq \operatorname{rank} B < n$，因此 $|AB| = 0$，即 $|AB| = |A||B|$. 假设 B 可逆，则根据推论2.1.8不妨将其写成初等矩阵的乘积，再根据初等矩阵的性质即可.

(5). 由 (4) 即得.

(6). 根据推论2.2.5可知 $\operatorname{rank} A = \operatorname{rank} A^T$. 因此如果 $\operatorname{rank} A < n$，则 $\operatorname{rank} A^T < n$，即根据 (2) 可知 $|A| = |A^T| = 0$. 假设 $\operatorname{rank} A = n$，再根据推论2.2.5将 A 写成初等矩阵的乘积，并注意到对于初等矩阵 E 有 $|E| = |E^T|$.

(7). 不妨假设

$$A = \begin{pmatrix} A_1 & A_2 \\ \mathrm{O} & A_3 \end{pmatrix}$$

假设 A_1, A_3 中有一个不可逆，则此时 A 也不可逆，因此 $|A| = |A_1||A_2| = 0$. 假设 A_1, A_2 都可逆，则此时 $\operatorname{rref}(A) = \mathrm{I}_n$. 将 A_1 化作最简行阶梯型的初等矩阵记做 $E_1, \ldots, E_k$，将 A_3 化作最简行阶梯型的初等矩阵记做 $E'_1, \ldots, E'_l$，则考虑

$$\widetilde{E}_i = \begin{pmatrix} E_i & \mathrm{O} \\ \mathrm{O} & \mathrm{I} \end{pmatrix}, \quad \widetilde{E}'_j = \begin{pmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & E'_j \end{pmatrix}$$

则 $\widetilde{E}_k \ldots \widetilde{E}_1 \widetilde{E}'_l \ldots \widetilde{E}'_1 A = \mathrm{I}_n$，从而 $|A|^{-1} = |\widetilde{E}_k| \ldots |\widetilde{E}_1||\widetilde{E}'_l| \ldots |\widetilde{E}'_1|$. 注意到 $|\widetilde{E}_i| = |E_i|, |\widetilde{E}'_j| = |E'_j|$ 以及 $|A_1| = \prod_{i=1}^{k} |E_i|, |A_3| = \prod_{j=1}^{l} |E'_j|$ 即可. □

命题 2.4.6. 对于 $2n \times 2n$ 阶矩阵

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

其中 A, B, C, D 是 $n \times n$ 阶矩阵，并且 $AC = CA$，则 $|M| = |AD - CB|$

证明: 首先我们假设 A 可逆, 则

$$\begin{aligned}|M|| &= |\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}| \\ &= |A||D-CA^{-1}B| \\ &= |A(D-CA^{-1}B)| \\ &= |AD-CB|\end{aligned}$$

而当 A 不可逆时, 我们不妨考虑 $A_\lambda = A+\lambda I$. 由于 $|A_\lambda|$ 是 λ 的 n 次多项式, 从而对有限多个 λ 之外 A_λ 总可逆, 因此我们不妨取足够小的 λ 使得 A_λ 总可逆, 根据可逆时的情形我们有

$$|\begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}| = |A_\lambda D - CB|$$

从而我们令 $\lambda \to 0$ 即有我们期待的结果[3]. □

2.5 伴随矩阵

命题 2.5.1. 对于 $A \in M_{n\times n}(\mathbb{R})$, 有

$$AA^* = A^*A = \det A\, \mathrm{I}_n$$

推论 2.5.2. 如果 $A \in M_{n\times n}(\mathbb{R})$ 可逆, 则

$$A^{-1} = \frac{1}{\det A}A^*$$

推论 2.5.3 (克拉姆法则 (cramer's rule))**.** 对于线性方程组 $Ax=b$, 如果 A 可逆, 则有唯一解

$$x = \frac{1}{\det A}A^*b$$

2.6 矩阵的若干应用

2.6.1 快速傅立叶变换

给定两个 d 次多项式

$$\begin{aligned}f(x) &= a_0 + a_1x + \cdots + a_dx^d \\ g(x) &= b_0 + b_1x + \cdots + b_dx^d\end{aligned}$$

多项式乘积为

$$fg := a_0b_0 + (a_1b_0 + b_1a_0)x + \dots$$

在实际计算中, 计算两个 d 次多项式的乘积需要进行 d^2 次运算, 即复杂度为 $O(d^2)$. 在本节中我们将介绍快速傅立叶变换 (fast fourier transformation), 将复杂度降到 $O(d\log d)$.

[3]这个操作称为微扰法, 是矩阵中一个非常经典的技巧.

注意到对于 d 次多项式 $f(x)$, 其可以由在 $d+1$ 个不同的 x 处取值决定, 即我们有如下的矩阵表达式

$$\begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_d) \end{pmatrix} = M \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix} = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^d \\ 1 & x_1 & x_1^2 & \dots & x_1^d \\ \vdots & & & & \vdots \\ 1 & x_d & x_d^2 & \dots & x_d^d \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix}$$

其中 M 被称为**范德蒙德矩阵** (Vandermond matrix).

练习. $|M| = \prod_{i<j}(x_i - x_j) \neq 0$, 即 M 可逆.

例子. 例如 $f(x) = x^4 + 3x^3 + 2x^2 + x + 1$, 我们将其写作

$$\begin{aligned} f(x) &= (x^4 + 2x^2 + 1) + x(3x^2 + 1) \\ &= f_e(x^2) + x f_o(x^2) \end{aligned}$$

从而

$$\begin{aligned} f(x_i) &= f_e(x_i^2) + x_i f_o(x_i^2) \\ f(-x_i) &= f_e(x_i^2) - x_i f_o(x_i^2) \end{aligned}$$

第三章　线性空间与线性变换

3.1　$\mathbb{R}^n$ 情形

3.1.1　$\mathbb{R}^n$ 的子空间

回忆对于线性方程组 $Ax=0$ 的解构成的集合, 其满足

(1) 加法封闭.

(2) 数乘封闭.

定义 3.1.1. 将满足 (1) 和 (2) 的 $\mathbb{R}^n$ 的非空子集称为 $\mathbb{R}^n$ 的**子空间** (subspace).

例子. *$\{0\}$ 是子空间*[1].

命题 3.1.2. *$\mathbb{R}^n$ 的任何一个子空间 W 均包含 0.*

证明: 任取 $w \in W$, 由于 W 对于数乘封闭, 那么 $0w = 0 \in W$, 即零向量在 W 中. □

注记. $\{0\}$ 是在包含关系下最小的子空间.

定义 3.1.3. 齐次线性方程组 $Ax=0$ 的解是 $\mathbb{R}^n$ 的子空间, 被称为解空间, 也被称为 A 的**核** (kernel), 记为 $\ker A$.

例子. *如果 A 是 3×3 阶矩阵, 根据线性方程组解的理论, 即定理1.3.2, 我们可以发现 $\ker A$ 在 $\mathbb{R}^3$ 中的形式与 $\operatorname{rank} A$ 关系密切.*

(1) *$\operatorname{rank} A = 0$, 此时 A 是零矩阵, 从而 $\ker A = \mathbb{R}^3$.*

(1) *$\operatorname{rank} A = 1$, 此时 $\ker A$ 是通过原点的平面.*

(1) *$\operatorname{rank} A = 2$, 此时 $\ker A$ 是通过原点的直线.*

(1) *$\operatorname{rank} A = 3$, 此时 A 可逆, 从而 $\ker A = \{0\}$.*

定义 3.1.4. 给定 $v_1,\dots,v_n \in \mathbb{R}^m$, 则

$$\operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_n\} := \{x_1v_1+\cdots+x_nv_n \mid x_i \in \mathbb{R}\}$$

称为 $v_1,\dots,v_n$ 的**线性张成** (linearly span), 可以直接验证 $\operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_n\}$ 是子空间.

注记. 现在我们来看一下线性方程组与线性张成的关系: 给定 $A \in M_{m\times n}(\mathbb{R})$, 将其写作 $A = (v_1,\dots,v_n)$, 那么 $Ax=b$ 有解当且仅当

$$b \in \operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_n\}$$

[1]这里的 0 指代零向量, 之后可能会用 0 即代指实数零又代指零向量, 请读者注意自己区分.

这个时候也称 $\mathrm{span}_{\mathbb{R}}\{v_1,\ldots,v_n\}$ 是 A 的**列空间** (column space), 记做 $\mathrm{im}\,A$. 类似的, 我们也可以定义其**行空间** (row space), 记做 $\mathrm{im}\,A^T$.

命题 3.1.5. 列变换不改变 A 的列空间.

证明: 不妨将 A 写作 $A=(v_1,\ldots,v_n)$, 根据推论2.2.3, 做列变换等价于右乘可逆矩阵 B, 因此不妨记 A 做列变换得到的矩阵为 $AB=(\overline{v}_1,\ldots,\overline{v}_n)$. 根据矩阵乘法的定义, 我们有 $\overline{v}_1,\ldots,\overline{v}_n$ 都是 $v_1,\ldots,v_n$ 的线性组合, 从而

$$\mathrm{span}_{\mathbb{R}}\{\overline{v}_1,\ldots,\overline{v}_n\}\subseteq \mathrm{span}_{\mathbb{R}}\{v_1,\ldots,v_n\}$$

并且由于列变换是可逆的, 即 A 也可以由 AB 做列变换得到, 从而

$$\mathrm{span}_{\mathbb{R}}\{v_1,\ldots,v_n\}\subseteq \mathrm{span}_{\mathbb{R}}\{\overline{v}_1,\ldots,\overline{v}_n\}$$

即列变换不改变 A 的列空间. □

推论 3.1.6. 行变换不改变 A 的行空间.

证明: 证明同上. □

3.1.2 线性相关性

定义 3.1.7. $v_1,\ldots,v_n\in\mathbb{R}^m$ 被称为**线性无关** (linearly independent), 如果 0 表示为 $v_1,\ldots,v_n$ 的线性组合的方式唯一, 即只有

$$0=0v_1+\cdots+0v_n$$

否则 $v_1,\ldots,v_n$ 被称为**线性相关** (linearly dependent).

命题 3.1.8. 若 $v_1,\ldots,v_n$ 线性无关, 则对于 $v\in\mathrm{span}_{\mathbb{R}}\{v_1,\ldots,v_n\}$, 其被写成 $v_1,\ldots,v_n$ 线性组合式子的系数是唯一的.

证明: 不妨假设

$$\begin{aligned} v&=a_1v_1+\cdots+a_nv_n\\ &=b_1v_1+\cdots+b_nv_n\end{aligned}$$

则

$$0=(a_1-b_1)v_1+\cdots+(a_n-b_n)v_n$$

根据线性无关的定义可知 $a_i=b_i$ 对任意的 $1\le i\le n$ 成立. □

定理 3.1.9. $v_1,\ldots,v_n\in\mathbb{R}^m$ 线性无关当且仅当 $\mathrm{rank}\,A=n$, 其中 $A=(v_1,\ldots,v_n)$ 是 $m\times n$ 阶矩阵, 此时也称 A 是列满秩的.

证明: 注意到 $v_1,\ldots,v_n$ 线性无关当且仅当 $Ax=0$ 只有零解, 根据定理1.3.2可知这当且仅当 $\mathrm{rank}\,A=n$. □

推论 3.1.10. $\mathbb{R}^m$ 中 k 个列向量当 $k>m$ 时一定线性相关.

证明: $\mathbb{R}^m$ 中 k 个列向量组成的矩阵 A 的秩在 $k > m$ 时最大为 m. □

推论 3.1.11. 列变换不改变矩阵列向量的线性相关性.

证明: 根据定理2.2.4, 列变换不改变矩阵的秩, 从而不改变列向量的线性相关性. □

定理 3.1.12. 假设 $v_1, \ldots, v_n$ 线性无关, 则 $v_1, \ldots, v_n, v_{n+1}$ 线性相关等价于 $v_{n+1} \in \mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\}$.

证明: 如果 $v_{n+1} \in \mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\}$, 则显然 $v_1, \ldots, v_{n+1}$ 线性相关. 另一方面, 假设

$$a_1 v_1 + \cdots + a_{n+1} v_{n+1} = 0$$

的系数 $a_1, \ldots, a_{n+1}$ 不全为零, 那么一定有 $a_{n+1} \neq 0$, 否则有

$$a_1 v_1 + \cdots + a_n v_n = 0$$

并且 $a_1, \ldots, a_n$ 不全为零, 这与线性无关相矛盾, 从而

$$v_{n+1} = -a_{n+1}^{-1}(a_1 v_1 + \cdots + a_n v_n) \in \mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\}$$

□

定义 3.1.13. 对于 $v_1, \ldots, v_n \in \mathbb{R}^m$, 称 $v_{i_1}, \ldots, v_{i_k}$ 是**极大线性无关组** (maximal linearly independent set), 如果

(1) $v_{i_1}, \ldots, v_{i_k}$ 线性无关.

(2) 任何包含 $v_{i_1}, \ldots, v_{i_k}$ 的 $\{v_1, \ldots, v_n\}$ 的子集中的向量都线性相关.

注记. 若 $v_1, \ldots, v_n$ 不全为零, 则其一定存在极大线性无关组: 假设 $v_1 \neq 0$, 考虑 v_1, v_2 是否线性相关, 如线性相关则剔除 v_2, 线性无关则保留 v_2. 再依次考虑 $v_3, v_4, \ldots$ 即可.

定理 3.1.14. 对于 $v_1, \ldots, v_n \in \mathbb{R}^m$, $v_{i_1}, \ldots, v_{i_k}$ 是其极大线性无关组, 则

$$\mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\} = \mathrm{span}_{\mathbb{R}}\{v_{i_1}, \ldots, v_{i_k}\}$$

证明: 显然 $\mathrm{span}_{\mathbb{R}}\{v_{i_1}, \ldots, v_{i_k}\} \subseteq \mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\}$, 并且根据定理以及极大线性无关组的定义可知任取 $i \in \{1, \ldots, n\} \backslash \{i_1, \ldots, i_k\}$ 有

$$v_i \in \mathrm{span}_{\mathbb{R}}\{v_{i_1}, \ldots, v_{i_k}\}$$

从而 $\mathrm{span}_{\mathbb{R}}\{v_1, \ldots, v_n\} = \mathrm{span}_{\mathbb{R}}\{v_{i_1}, \ldots, v_{i_k}\}$. □

注意在定义极大线性无关组时的极大意味着在包含意义下极大, 但是其实际上也是绝对数目的极大.

定理 3.1.15. 假设 $v_1, \ldots, v_n$ 的某个极大线性无关组中有 k 个向量, 则对于任意 $v_{j_1}, \ldots, v_{j_l}$, 其中 $l > k$, 其线性相关.

证明: 假设 $v_{i_1},\dots,v_{i_k}$ 是 $v_1,\dots,v_n$ 的一个极大线性无关组, 任取 $l>k$ 以及 $v_{j_1},\dots,v_{j_l}$, 根据极大线性无关组的定义有

$$(v_{j_1},\dots,v_{j_l})=(v_{i_1},\dots,v_{i_k})A$$

其中 $A\in M_{k\times l}(\mathbb{R})$, 从而

$$x_1v_{j_1}+\cdots+x_lv_{j_l}=(v_{j_1},\dots,v_{j_l})\begin{pmatrix}x_1\\x_2\\\vdots\\x_l\end{pmatrix}=(v_{i_1},\dots,v_{i_k})A\begin{pmatrix}x_1\\x_2\\\vdots\\x_l\end{pmatrix}$$

由于 $l>k$, 从而根据线性方程组解的结构定理1.3.2可知 $Ax=0$ 有非零解, 从而 $v_{j_1},\dots,v_{j_l}$ 线性相关. $\square$

推论 3.1.16. *$v_1,\dots,v_n$ 的极大线性无关组中向量的数目是确定的.*

3.1.3 $\mathbb{R}^n$ 子空间的维数

定义 3.1.17. $\mathbb{R}^n$ 的子空间 W 中的一个极大线性无关组被称作**基** (basis).

例子. $W=\mathbb{R}^n$, 则

$$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\quad e_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}\quad\dots\quad e_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

是 $\mathbb{R}^n$ 的一组基.

定义 3.1.18. $\mathbb{R}^n$ 的子空间 W 的极大线性无关组中向量的个数被称为 W 的**维数** (dimension).

注记. 根据推论3.1.16可知 W 的维数是良定义的, 并且如下三条中任意满足两条即可说明 $v_1,\dots,v_k$ 是 W 的基:

(1) $W=\mathrm{span}_{\mathbb{R}}\{v_1,\dots,v_k\}$.

(2) $v_1,\dots,v_k$ 线性无关.

(3) $\dim W=k$.

命题 3.1.19. *若 $W_1\subseteq W_2$ 都是 $\mathbb{R}^n$ 的子空间, 则*

(1) $\dim W_1\le\dim W_2$, *并且等号成立当且仅当 $W_1=W_2$.*

(2) *W_1 的基可以扩充为 W_2 的基.*

证明: (1). 由于 W_1 中的线性无关组一定是 W_2 中的线性无关组, 从而$\dim W_1\le\dim W_2$, 等号取得是显然的.

(2). 假设 W_1 的基是 $v_1,\dots,v_k$, W_2 的基是 $w_1,\dots,w_l$. 我们在取 $v_1,\dots,v_k,w_1,\dots,w_l$ 的极大线性无关组的时候仔细一些: 即前 k 个向量取 $v_1,\dots,v_k$, 这是可以做到的, 因为 $v_1,\dots,v_k$ 本身线性无关. 这样取出的极大线性无关组就是由 $v_1,\dots,v_k$ 扩充得到的 W_2 的基. $\square$

定理 3.1.20. 对于矩阵 A 来说, $\operatorname{rank} A$ 是行空间维数, $\operatorname{rank} A^T$ 是列空间维数.

证明: 根据推论3.1.6可知行变换不改变行空间, 因此我们通过行变换将其化作最简行阶梯型, 此时主元所在的行向量构成了行空间的一组基, 因此 $\operatorname{rank} A$ 是行空间维数. 类似的可以证明 $\operatorname{rank} A^T$ 是列空间维数. □

推论 3.1.21. 对于矩阵 A 来说, 行空间维数等于列空间维数.

证明: 根据推论2.2.5即可. □

3.2 域上的线性空间

定义 3.2.1. 集合 $\mathbb{F}$ 被称为**域** (field), 如果其上带有运算

(1) $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, 记做 $(a, b) \mapsto a + b$.

(2) $\times : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, 记做 $(a, b) \mapsto ab$.

满足

(3) 交换律, 即 $a + b = b + a, ab = ba$.

(4) 结合律, 即 $(a + b) + c = a + (b + c), (ab)c = a(bc)$.

(5) 分配律, 即 $a(b + c) = ab + ac$.

(6) 单位元, 即存在 $1, 0 \in \mathbb{F}$, 使得 $a + 0 = a, 1a = a$.

(7) 逆元, 即存在 $b \in \mathbb{F}$ 是的 $a + b = 0$; 对 $a \neq 0$, 存在 c 使得 $ac = 1$.

例子. 常见的 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 都是域.

例子. $\mathbb{R} \times \mathbb{R}$ 上定义

$$(a, b) + (c, d) := (a + c, b + d)$$
$$(a, b)(c, d) := (ac - bd, ad + bc)$$

可以直接验证如上运算使得 $\mathbb{R} \times \mathbb{R}$ 构成了域.

注记. 在给定域 $\mathbb{F}$ 之后, 我们之前对实数域 $\mathbb{R}$ 所做的事情在一般的域 $\mathbb{F}$ 上均成立, 即线性函数, 线性方程组, 线性方程组解的结构, 秩, 子空间以及维数等等.

定义 3.2.2. 给定集合 V, 如果其上有如下结构

(1) 加法 $V \times V \to V$, 记做 $(v_1, v_2) \mapsto v_1 + v_2$.

(2) 数乘 $\mathbb{F} \times V \to V$, 记做 $(c, v) \mapsto cv$.

满足

(3) 加法满足交换律, 结合律, 0 向量以及逆元.

(4) 数乘满足结合律以及单位元.

(5) 加法和数乘满足分配律.

则称 V 构成了一个 $\mathbb{F}$-**线性空间** (vector space).

例子. $\mathbb{F}^n$ 构成了一个 $\mathbb{F}$-线性空间. 特别地, $\mathbb{R}^n$ 是一个 $\mathbb{R}$-线性空间.

例子. 假设有域之间的包含关系 $\mathbb{F} \subseteq \mathbb{E}$, 则任何 $\mathbb{E}$-线性空间都可以视作 $\mathbb{F}$-线性空间. 特别地, 任何 $\mathbb{C}$-线性空间都可以视作 $\mathbb{R}$-线性空间.

例子. $M_{m\times n}(\mathbb{F})$ 构成了 $\mathbb{F}$-线性空间, 但全体可逆矩阵不构成线性空间.

例子. 全体 $\mathbb{R}$-系数多项式 $\mathbb{R}[x]$ 构成了 $\mathbb{R}$-线性空间.

例子. 全体次数小于等于 n 的 $\mathbb{R}$-系数多项式, 记做 $\mathbb{R}[x]_{\leq n}$ 构成了 $\mathbb{R}$-线性空间.

例子. $V=\{f(x)\mid f(x)\text{是定义在 }[a,b]\text{ 上的光滑函数}\}$, 这常常记做 $C^\infty([a,b])$.

我们之前对于 $\mathbb{R}^n$ 所发展的理论都可以对一般的 $\mathbb{F}$-线性空间发展.

定义 3.2.3. 给定 $\mathbb{F}$-线性空间 V, $W\subseteq V$ 被称为 V 的**子空间** (subspace), 如果 W 对 V 上的数乘与加法都封闭.

定义 3.2.4. $\mathbb{F}$-线性空间 V 中的向量 $v_1,\dots,v_n$ 称为 **$\mathbb{F}$-线性无关** (linearly independent)(线性无关, linearly independent), 如果 0 表示为 $v_1,\dots,v_n$ 的线性组合方式唯一, 即只有

$$0=0v_1+\cdots+0v_n$$

否则 $v_1,\dots,v_n$ 称为**线性相关** (linearly dependent).

定义 3.2.5. $\mathbb{F}$-线性空间 V 的一个 $\mathbb{F}$-极大线性无关组被称作 V 的 **$\mathbb{F}$-基** (basis).

定义 3.2.6. $\mathbb{F}$-线性空间 V 的一个 $\mathbb{F}$-极大线性无关组中向量的个数被称作 V 的 **$\mathbb{F}$-维数** (dimension).

注记. 注意, 在这里我们强调它作为某个域上线性空间的维数, 因为同一个集合作为不同线性空间的维数可能不同, 例如 $\dim_{\mathbb{C}}\mathbb{C}=1$ 但是 $\dim_{\mathbb{R}}\mathbb{C}=2$.

注记. 对于一般的 $\mathbb{F}$-线性空间 V, 其维数不一定有限, 比如下面的例子.

例子. 对于 $\mathbb{R}$-线性空间 $\mathbb{R}[x]$, $1,x,x^2,\dots$ 构成了一个极大线性无关组, 因此 $\dim_{\mathbb{R}}\mathbb{R}[x]=\infty$.

但是在这门课程中, 我们主要关心有限维的线性空间. 在之后, 如果不加特殊说明, 我们总假设线性空间是有限维的.

例子. $\mathbb{F}$-线性空间 $\mathbb{F}^n$ 有一组自然的基 $\{e_1,\dots,e_n\}$, 其中 e_i 是只有第 i 行为 1, 其余行为 0 的列向量.

例子. 基本矩阵 E_{ij}, 即只有 (i,j) 元是 1, 其余地方为零的矩阵构成了 $M_{m\times n}(\mathbb{R})$ 的一组基, 因此 $\dim_{\mathbb{R}}M_{m\times n}(\mathbb{R})=mn$.

例子. $1,x,x^2,\dots,x^n$ 构成了 $\mathbb{R}[x]_{\leq n}$ 的一组基, 因此 $\dim_{\mathbb{R}}\mathbb{R}[x]_{\leq n}=n+1$.

例子. $V=\{\frac{ax^2+bx+c}{x^3-x}\mid a,b,c\in\mathbb{R}\}$ 构成了一个 $\mathbb{R}$-线性空间, 其有如下两组不同的基:

$$B=\{\frac{x^2}{x^3-x},\frac{x}{x^3-x},\frac{1}{x^3-x}\}$$
$$C=\{\frac{1}{x},\frac{1}{x-1},\frac{1}{x+1}\}$$

定义 3.2.7. V 是 $\mathbb{F}$-线性空间, $B=\{v_1,\dots,v_n\}$ 是其一组 $\mathbb{F}$-基, 则对于 $v\in V$, 有

$$v=a_1v_1+\cdots+a_nv_n$$

其中 $a_i\in\mathbb{F}$. 列向量

$$[v]_B=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\in\mathbb{F}^n$$

被称为 v 在基 B 下的**坐标** (coordinate).

对于 $\mathbb{F}$-线性空间 V, 给定其一组基 $B=\{v_1,\dots,v_n\}$, 我们可以将其视作矩阵 $B=(v_1,\dots,v_n)$, 此时我们有矩阵乘法的等式

$$v=B[v]_B$$

并且对于另一个基 $C=\{w_1,\dots,w_n\}$, 任取 $1\le i\le j$, 有

$$w_i=p_{i1}v_1+\cdots+p_{in}v_n$$

即如果我们记矩阵 $P=(p_{ij})_{n\times n}$, 则有

$$C^T=PB^T$$

对于 $v\in V$, 有

$$\begin{aligned}v&=a_1w_1+\cdots+a_nw_n\\&=a_1(\sum_{i=1}^n p_{1i}v_i)+\cdots+a_n(\sum_{i=1}^n p_{ni}v_i)\\&=\sum_{i=1}^n a_ip_{1i}v_1+\cdots+\sum_{i=1}^n a_ip_{in}v_n\end{aligned}$$

即有

$$[v]_C=P^T[v]_B$$

我们通常用 P_{CB} 记上述 P^T, 称为基 B 到基 C 的**转移矩阵** (transition matrix).

命题 3.2.8. 对于转移矩阵, 我们有如下简单的性质.

(1) $P_{B_3B_1}=P_{B_3B_2}P_{B_2B_1}$.

(2) $P_{BB}=\mathrm{I}_n$.

(3) P_{BC} 是可逆矩阵, 并且 $P_{BC}^{-1}=P_{CB}$.

基的转移矩阵在实用中往往非常有用, 比如在例子3.2中, 考虑

$$v=\frac{3x^2+x+2}{x^3-x}$$

则有

$$[v]_B=\begin{pmatrix}2\\1\\3\end{pmatrix}$$

然而很难直接写出 $[v]_C$. 但是对基 C 来说, 其有很多好处, 例如便于积分.

3.3 线性空间的构造

3.3.1 直和

定义 3.3.1. 给定 $\mathbb{F}$-线性空间 V, W, 定义线性空间的**外直和** (external direct sum)为集合 $V \times W$, 其上带有结构

(1) $(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2)$.

(2) $c(v, w) := (cv, cw)$.

使得其成为一个 $\mathbb{F}$-线性空间, 记做 $V \oplus W$.

例子. $\mathbb{R}^n$ 可以视作 n 个 $\mathbb{R}$ 的外直和.

命题 3.3.2. 给定 $\mathbb{F}$-线性空间 V, W, 有

$$\dim V \oplus W = \dim V + \dim W$$

证明: 假设 $\{v_1, \ldots, v_n\}$ 是 V 的一组基, $\{w_1, \ldots, w_m\}$ 是 W 的一组基, 直接验证 $\{(v_1, 0), \ldots, (v_n, 0), (0, w_1), \ldots,$ 构成了 $V \oplus W$ 的一组基. □

定义 3.3.3. 给定 $\mathbb{F}$-线性空间 V 以及其子空间 V_1, V_2, 如果

(1) $V = V_1 + V_2$, 即任何 V 中的向量可以表示成 V_1, V_2 中向量的组合.

(2) $V_1 \cap V_2 = \{0\}$.

则称 V 是 V_1 和 V_2 的**内直和** (internal direct sum).

命题 3.3.4. 给定 $\mathbb{F}$-线性空间 V 以及其子空间 V_1, V_2, 如果 V 是 V_1 和 V_2 的内直和, 那么

$$T : V_1 \oplus V_2 \to V_1 + V_2$$
$$(v_1, v_2) \mapsto v_1 + v_2$$

是线性同构, 即 $V \cong V_1 \oplus V_2$.

证明: 线性映射 T 显然是满射, 并且根据内直和的定义有 $\ker T = V_1 \cap V_2 = \{0\}$. □

注记. 这意味着内直和与外直和是一体两面, 因此我们之后并不再区分内外直和, 而统称为直和.

定义 3.3.5. 给定 $\mathbb{F}$-线性空间 V 以及子空间 V_1, 如果存在子空间 V_2 满足 $V = V_1 \oplus V_2$, 则称 V_2 是 V_1 的补空间 (complement space).

注记. 根据命题3.1.19, 我们总可以将子空间的一组基延拓成全空间的一组基, 因此补空间总是存在的.

命题 3.3.6. $\mathbb{F}$ 线性空间 $V = V_1 \oplus \cdots \oplus V_k$ 当且仅当

(1) $V = V_1 + \cdots + V_k$.

(2) 对任意 $1 \le i \le k$, 有 $V_i \cap \sum_{i \neq j} V_j = \{0\}$.

3.3.2 商空间

定义 3.3.7. 给定 $\mathbb{F}$-线性空间 V 以及其子空间 W, **商空间** (quotient space)定义为集合 V/W, 其上带有结构

(1) $(v_1 + W) + (v_2 + W) := (v_1 + v_2) + W$.

(2) $c(v + W) := cv + W$.

注记. 注意我们在处理商集的时, 定义的良好性是我们始终要考虑的问题, 即我们的定义不依赖于代表元的选取.

命题 3.3.8. *给定 $\mathbb{F}$-线性空间以及商空间 V/W, 自然投射*

$$\begin{aligned}\pi : V &\to V/W\\ v &\mapsto v + W\end{aligned}$$

是满的线性映射, 并且 $\ker \pi = W$.

命题 3.3.9. *给定 $\mathbb{F}$-线性空间以及商空间 V/W, 有*

$$\dim V = \dim V/W + \dim W$$

证明: 假设 $v_1+W, \ldots, v_n+W, w_1, \ldots, w_m$ 分别构成了 V/W 和 W 的基, 直接验证 $v_1, \ldots, v_n, w_1, \ldots, w_m$ 构成了 V 的一组基. □

3.4 线性映射

3.4.1 定义与例子

定义 3.4.1. 对于 $\mathbb{F}$-线性空间 V, W, 映射 $T : V \to W$ 被称为 **$\mathbb{F}$-线性映射** (linear map), 如果

(1) 对任意 $v_1, v_2 \in V$ 有 $T(v_1 + v_2) = Tv_1 + Tv_2$.

(2) 对于任意 $c \in \mathbb{F}, v \in V$ 有 $T(cv) = cv$.

定义 3.4.2. $\mathbb{F}$-线性映射 T 是**线性同构** (linear isomorphism), 如果其既是双射也是满射.

定义 3.4.3. $\mathbb{F}$-线性空间 V, W 称为**线性同构** (linear isomorphism), 如果存在线性同构 $T : V \to W$.

例子. *逆时针旋转角度 θ 是 $\mathbb{R}^2 \to \mathbb{R}^2$ 的 $\mathbb{R}$-线性同构.*

例子. *$\mathbb{R}^n$ 到子空间 W 的投影映射是 $\mathbb{R}$-线性映射.*

命题 3.4.4. *给定 n 维 $\mathbb{F}$-线性空间 V 以及其一组基 $B = \{v_1, \ldots, v_n\}$, 如下映射是 $\mathbb{F}$-线性同构*

$$\begin{aligned}T : V &\to \mathbb{F}^n\\ \sum_{i=1}^n a_i v_i &\mapsto \sum_{i=1}^n a_i e_i\end{aligned}$$

证明: 映射 T 显然是 $\mathbb{F}$-线性的, 并且满射也是显然的. 假设

$$T(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i e_i = 0$$

根据 $\{e_i\}$ 的线性无关性可知 $a_i = 0$, 从而 $\sum_{i=1}^{n} a_i v_i = 0$. □

推论 3.4.5. $\mathbb{F}$-线性空间 V, W 之间同构当且仅当 $\dim_{\mathbb{F}} V = \dim_{\mathbb{F}} W$, 即维数是线性空间的完全不变量.

证明: 显然同构的线性空间有相同的维数; 另一方面, 假设 $\dim_{\mathbb{F}} V = \dim_{\mathbb{F}} W = n$, 则 $V \cong \mathbb{F}^n \cong W$. □

例子. 对于矩阵 $A \in M_{m\times n}(\mathbb{F})$, 对于 n 维 $\mathbb{F}$-线性空间 V 以及 m 维 $\mathbb{F}$-线性空间 W

$$\begin{aligned} T_A : V &\to W \\ v &\mapsto CA[v]_B \end{aligned}$$

是 $\mathbb{F}$-线性映射, 其中 B, C 分别是 V 和 W 的基. T_A 为线性同构当且仅当 A 可逆.

定义 3.4.6. 给定 $\mathbb{F}$-线性空间 V, W, 全体 V 到 W 的 $\mathbb{F}$ 线性映射组成的集合记做 $\mathrm{Hom}_{\mathbb{F}}(V, W)$.

注记. $\mathrm{Hom}_{\mathbb{F}}(V, W)$ 上有自然的加法与 $\mathbb{F}$-数乘结构如下:

(1) $(T_1 + T_2)v = T_1 v + T_2 v$.

(2) $(cT)v = cTv$.

并且可以直接验证上述结构使得 $\mathrm{Hom}_{\mathbb{F}}(V, W)$ 是 $\mathbb{F}$-线性空间.

命题 3.4.7. 给定 $\mathbb{F}$-线性映射 $T_1 : V_1 \to V_2, T_2 : V_2 \to V_3$, 则 $T_2 \circ T_1 \in \mathrm{Hom}_{\mathbb{F}}(V_1, V_3)$.

3.4.2 线性映射与矩阵

定理 3.4.8. 对于 n 维 $\mathbb{F}$-线性空间 V 以及 m 维 $\mathbb{F}$-线性空间 W, 有如下 $\mathbb{F}$-线性空间的同构

$$M_{m\times n}(\mathbb{F}) \longleftrightarrow \mathrm{Hom}_{\mathbb{F}}(V, W)$$

证明: 考虑映射

$$\begin{aligned} M_{m\times n}(\mathbb{F}) &\to \mathrm{Hom}_{\mathbb{R}}(V, W) \\ A &\mapsto T_A \end{aligned}$$

首先线性性是显然的. 我们固定 V 的一组基 $\{v_1, \dots, v_n\}$ 以及 W 的基 $C = \{w_1, \dots, w_m\}$. 对于矩阵 $A_1, A_2 \in M_{m\times n}(\mathbb{F})$, 注意到 $[T_{A_1} v_i]_C$ 是 A_1 的第 i 列, $[T_{A_2} v_i]_C$ 是 A_2 的第 i 列, 从而如果 $T_{A_1} = T_{A_2}$, 从而 $A_1 = A_2$; 另一方面, 对于 $T \in \mathrm{Hom}_{\mathbb{F}}(V, W)$, 我们可以得到矩阵[2] $[T]_B^C = ([Tv_1]_C, \dots, [Tv_n]_C) \in M_{m\times n}(\mathbb{F})$, 直接验证有

$$\begin{aligned} Tv &= T((v_1, \dots, v_n)[v]_B) \\ &= (Tv_1, \dots, Tv_n)[v]_B \\ &= C([Tv_1]_C, \dots, [Tv_n]_C)[v]_B \\ &= T_{[T]_B^C} v \end{aligned}$$

□

[2]我们称矩阵 $[T]_B^C$ 为线性映射 T 在基 B, C 下的矩阵.

注记. 这是一个非常重要的观点, 即当我们取定线性空间 V 的一组基之后, 根据命题3.4.4我们可以将 V 看作 $\mathbb{F}^n$, 并且根据定理3.4.8我们可以将线性映射看作矩阵, 反之亦然.

注意到定理3.4.8中的对应是依赖于基的选取的, 那么一个自然的问题是给定一个 $\mathbb{F}$-线性映射, 选取不同的基得到的矩阵之间有什么关系?

命题 3.4.9. 假设 B_1, B_2 都是 $\mathbb{F}$-线性空间 V 的基, C_1, C_2 都是 W 的基, 则对于 $T \in \mathrm{Hom}_{\mathbb{F}}(V, W)$, 有

$$[T]_{B_1}^{C_1} = P_{C_1C_2}[T]_{B_2}^{C_2}P_{B_2B_1}$$

定义 3.4.10. 矩阵 $A, B \in M_{n\times n}(\mathbb{F})$ 称为**相似** (similar), 如果存在可逆矩阵 $P \in M_{n\times n}(\mathbb{F})$ 使得 $PAP^{-1} = B$.

因此, 对于 $\mathbb{F}$-线性映射 $T : V \to V$, 其在不同基下的矩阵是相似的, 这也是相似矩阵的几何解释. 如果对于矩阵来说一些量是相似不变量[3], 则我们可以对线性映射定义.

定义 3.4.11. 对于 $\mathbb{F}$-线性映射 $T : V \to W$, 其**行列式** (determinant)定义为 $\det[T]_B^C$, 其中 B, C 分别是 V, W 的基, 记做 $\det T$.

注记. 根据命题2.4.5的 (5) 即可.

定义 3.4.12. 对于 $\mathbb{F}$-线性映射 $T : V \to W$, 其**秩** (rank)定义为 $\mathrm{rank}[T]_B^C$, 其中 B, C 分别是 V, W 的基, 记做 $\mathrm{rank}\, T$.

注记. 由于相似矩阵只相差可逆矩阵左乘右乘, 并且左乘右乘可逆矩阵不改变矩阵的秩, 从而线性映射的秩定义是良好的.

例子. 考虑如下线性映射

$$\begin{aligned} T : \mathbb{R}[x]_{\le n} &\to \mathbb{R}[x]_{\le n} \\ f(x) &\mapsto \frac{\mathrm{d}f}{\mathrm{d}x} \end{aligned}$$

考虑基 $B = \{1, x, \dots, x^n\}$, 则有

$$[T]_B^B = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & n & 0 \end{pmatrix}_{(n+1)\times(n+1)}$$

注记. 在上面的例子中, $T^{n+1} = 0$, 并且 $\mathrm{rank}\, T = n$.

[3]即指如果两个矩阵相似, 它们的这个量相同.

3.5 维数公式

定义 3.5.1. 对于 $\mathbb{F}$-线性映射 $T: V \to W$, 其**核** (kernel)定义为

$$\ker T := \{v \in V \mid Tv = 0\}$$

其**像** (image)定义为

$$\operatorname{im} T := \{Tv \in W \mid v \in V\}$$

命题 3.5.2. 对于 $\mathbb{F}$-线性映射 $T: V \to W$, $\ker T$ 是 V 的子空间, $\operatorname{im} T$ 是 W 的子空间.

定理 3.5.3 (维数公式)**.** 对于 $\mathbb{F}$-线性映射 $T: V \to W$, 有

$$\dim V = \dim \ker T + \dim \operatorname{im} T$$

推论 3.5.4. 对于 $A \in M_{m\times n}(\mathbb{F})$, 有如下维数公式成立

$$\dim \ker A + \operatorname{rank} A = n$$

推论 3.5.5. 对于 $A \in M_{m\times n}(\mathbb{F})$, 线性方程组 $Ax = 0$ 解空间的维数为 $n - \operatorname{rank} A$.

第四章　对角化

对于 $\mathbb{F}$-线性映射 $T: V \to V$, 能否找到 V 的一组基 $\{v_1, \ldots, v_n\}$ 满足对任意的 $1 \leq i \leq n$ 有 $Tv_i = \lambda_i v_i$, 其中 $\lambda_i \in \mathbb{F}$, 称为线性映射的可对角化问题. 用矩阵的语言来说[1], 给定矩阵 $A \in M_{n\times n}(\mathbb{F})$, 是否存在可逆矩阵 $P \in M_{n\times n}(\mathbb{F})$ 使得 PAP^{-1} 是对角矩阵.

当 $\dim_{\mathbb{F}} V = 1$ 时是显然的, 因为任何 1×1 阶矩阵都是对角矩阵. 但是对一般的维数来说, 并不是所有的线性映射都是可对角化的. 考虑 $\dim_{\mathbb{F}} = 2$ 的情形: 固定 V 的一组基 B 将 T 看作矩阵

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

那么 $0 \neq v \in V$ 满足 $Tv = \lambda v$, 其中 $\lambda \in \mathbb{F}$, 等价于 $(\lambda \mathrm{I}_2 - A)v = 0$ 有非零解, 根据定理2.1.6这等价于 λ 满足 $|\lambda \mathrm{I}_2 - A| = 0$, 即

$$\lambda^2 - (a+d)\lambda + ad - bc = 0$$

我们考虑如下的情况

(1) 如果 $\lambda^2 - (a+d)\lambda + ad - bc = 0$ 不存在根, 那么一定不存在 V 的一组基 $\{v_1, v_2\}$ 使得 $Tv_i = \lambda_i v_i$, 其中 $i = 1, 2$.

(2) 如果 $\lambda^2 - (a+d)\lambda + ad - bc = 0$ 有两个不同的根 λ_1, λ_2, 那么考虑对于每一个 λ_i 考虑 $(\lambda_i \mathrm{I}_2 - A)v = 0$ 的非零解 v_i, 那么我们断言 $\{v_1, v_2\}$ 构成了一组基, 因此满足我们的要求: 假设 $\{v_1, v_2\}$ 线性相关, 不妨假设 $v_1 = \mu v_2$, 从而

$$\lambda_1 v_1 = Tv_1 = T(\mu v_2) = \mu \lambda_2 v_2$$

这意味着 $\mu(\lambda_1 - \lambda_2) = 0$, 从而有 $\mu = 0$, 从而证明了 v_1, v_2 线性无关.

(3) 如果 $\lambda^2 - (a+d)\lambda + ad - bc = 0$ 有重根 λ, 我们要考虑 $\lambda \mathrm{I}_2 - A$ 的秩:

　(a) $\mathrm{rank}(\lambda \mathrm{I}_2 - A) = 0$, 此时根据推论3.5.5解空间的维数是 2, 因此可以找到两个线性无关的向量 $\{v_1, v_2\}$ 满足 $Tv_i = \lambda v_i$, 其中 $i = 1, 2$.

　(b) $\mathrm{rank}(\lambda \mathrm{I}_2 - A) = 1$, 此时根据推论3.5.5解空间的维数是 1, 则此时找不到两个线性无关的向量.

注记. 实际上, 我们对 2 维情况做的分析, 已经暗示了对一般的维数该如何去分析.

[1] 在之后会根据需要选择用不同的语言来叙述问题, 但本质都是同一件事.

4.1 特征值与特征向量

4.1.1 线性映射语言

定义 4.1.1. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, $\lambda \in \mathbb{F}$ 被称为 T 的**特征值** (eigenvalue), 如果存在 $0 \neq v \in V$ 使得

$$Tv = \lambda v$$

此时 v 称为特征值 λ 对应的**特征向量** (eigenvector).

定义 4.1.2. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, 其中 $\dim_{\mathbb{F}} V = n$. 关于 λ 的 n 次多项式 $|\lambda \mathrm{I}_n - T|$ 称为 T 的**特征多项式** (characteristic polynomial).

注记. 显然, $\mathbb{F}$-线性映射 T 拥有的特征值是特征多项式在域 $\mathbb{F}$ 中的根, 因此 n 维线性空间上的线性变换至多有 n 个不同的特征值.

命题 4.1.3. *假设 V 是 n 维 $\mathbb{C}$-线性空间, 给定 $\mathbb{C}$-线性映射 $T: V \to V$, 存在 V 的一组基 B 使得 $[T]_B^B$ 是上三角矩阵.*

证明: 我们对维数 n 做归纳法, 当 $n = 1$ 的时候是显然的. 假设命题对 $n < k$ 都成立, 当 $n = k$ 时, 由于 $\mathbb{C}$ 是代数闭域, 因此 T 的特征多项式总存在根, 即 T 总有特征值. 不妨假设 $v_1 \in V$ 是 T 对于特征值 λ_1 的特征向量. 将 v_1 扩充为 V 的一组基 $B = \{v_1, \ldots, v_n\}$, 则

$$[T]_B^B = \begin{pmatrix} \lambda_1 & * \\ \mathrm{O} & A \end{pmatrix}$$

其中 A 是 $(k-1) \times (k-1)$ 阶矩阵. 根据归纳法存在可逆矩阵 P 使得 PAP^{-1} 是上三角矩阵, 从而考虑基 $B' = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} B$, 则此时 T 在这组基下的矩阵为上三角矩阵. □

推论 4.1.4. *对于 $\mathbb{C}$-线性映射 $T: V \to V$, 其特征值恰为 $[T]_B^B$ 对角线上的元素, 其中 B 是使得 $[T]_B^B$ 为上三角矩阵的基.*

定义 4.1.5. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, 全体特征值 λ 对应的特征向量构成了一个 $\mathbb{F}$-线性空间, 称为特征值 λ 的**特征子空间** (eigenspace), 记做 V_λ.

命题 4.1.6. *对于 $\mathbb{F}$-线性映射 $T: V \to V$ 以及其特征值 λ, 有*

$$V_\lambda = \ker(\lambda \mathrm{I} - T)$$

证明: 根据定义即可. □

定理 4.1.7. *$\mathbb{F}$-线性映射 $T: V \to V$ 可对角化当且仅当 V 有特征子空间分解, 即 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 其中 $\lambda_1, \ldots, \lambda_s$ 是 T 的全体不同特征值.*

证明: 假设 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 则任取 $V_{\lambda_1}, \ldots, V_{\lambda_s}$ 的基, 将其并起来得到 V 的一组基, 则在这组基下 T 对应的矩阵是对角矩阵; 另一方面, 假设 T 可对角化, 则可以找到 V 的一组由特征向量构成的基, 这组基给出了 V 的特征子空间分解. □

定义 4.1.8. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, 特征值 λ 的特征子空间 V_λ 的维数称为 λ 的**几何重数** (geometric multiplicity).

定义 4.1.9. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, 特征值 λ 在特征多项式中的重数称为 λ 的**代数重数** (algebraic multiplicity).

命题 4.1.10. 对于 $\mathbb{F}$-线性映射 $T: V \to V$, 其代数重数为

$$\dim\{v \in V \mid 存在\ n \in \mathbb{N}_{\geq 0}\ 使得\ (\lambda \mathrm{I} - T)^n v = 0\}$$

推论 4.1.11. 对于 $\mathbb{F}$-线性映射 $T: V \to V$ 的特征值 λ, 其几何重数小于等于代数重数.

定理 4.1.12. $\mathbb{F}$-线性映射 $T: V \to V$ 可对角化当且仅当对于每一个特征值 λ 其几何重数等于代数重数.

4.1.2 矩阵语言

定义 4.1.13. 对于矩阵 $A \in M_{m\times n}(\mathbb{F})$, $\lambda \in \mathbb{F}$ 被称为 A 的**特征值** (eigenvalue), 如果存在非零列向量使得

$$Av = \lambda v$$

此时 v 称为特征值 λ 对应的**特征向量** (eigenvector).

定义 4.1.14. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 关于 λ 的 n 次多项式 $|\lambda \mathrm{I}_n - A|$ 称为 A 的**特征多项式** (characteristic polynomial).

定义 4.1.15. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 全体特征值 λ 对应的特征向量构成了一个 $\mathbb{F}$-线性空间, 称为特征值 λ 的**特征子空间** (eigenspace), 记做 V_λ.

命题 4.1.16. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$ 以及其特征值 λ, 有

$$V_\lambda = \ker(\lambda \mathrm{I}_n - T)$$

定理 4.1.17. 矩阵 $A \in M_{n\times n}(\mathbb{F})$ 可对角化当且仅当 $\mathbb{F}^n$ 有特征子空间分解, 即 $\mathbb{F}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 其中 $\lambda_1, \ldots, \lambda_s$ 是 T 的全体不同特征值.

定义 4.1.18. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 特征值 λ 在特征多项式中的重数称为 λ 的**代数重数** (algebraic multiplicity).

定义 4.1.19. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 特征值 λ 的特征子空间 V_λ 的维数称为 λ 的**几何重数** (geometric multiplicity).

命题 4.1.20. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其代数重数为

$$\dim\{v \in \mathbb{F}^n \mid 存在\ k \in \mathbb{N}_{\geq 0}\ 使得\ (\lambda \mathrm{I} - T)^k v = 0\}$$

推论 4.1.21. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$ 的特征值 λ, 其几何重数小于等于代数重数.

定理 4.1.22. 矩阵 $A \in M_{n\times n}(\mathbb{F})$ 可对角化当且仅当对于每一个特征值 λ 其几何重数等于代数重数.

注记. 在往后的章节中, 我们将只对矩阵定义相应的概念, 读者可以自行将线性映射部分平行的补全.

4.2 极小多项式

定义 4.2.1. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 多项式 $f(\lambda)$ 称为其**零化多项式** (annihilation polynomial), 如果 $f(A) = 0$.

定义 4.2.2. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其次数最低的首一零化多项式被称为**极小多项式** (minimal polynomial).

注记. 相似的矩阵有相同的极小多项式.

命题 4.2.3. *对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其极小多项式整除其任何一个零化多项式.*

证明: 假设 $m(\lambda)$ 是 A 的极小多项式, $f(\lambda)$ 是 A 的某个零化多项式, 作带余除法则有

$$f(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$

其中 $r(\lambda)$ 显然也是 A 的零化多项式. 如果 $r(\lambda)$ 不为零, 根据带余除法的结果我们有 $\deg r(\lambda) < m(\lambda)$, 这与极小多项式的定义矛盾. □

推论 4.2.4. *对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其极小多项式是唯一的.*

证明: 假设 $m_1(\lambda), m_2(\lambda)$ 都是 A 的极小多项式, 从而有 $m_1(\lambda) \mid m_2(\lambda)$ 以及 $m_2(\lambda) \mid m_1(\lambda)$, 因此 $m_1(\lambda), m_2(\lambda)$ 之间相差一个非零常数 c, 再利用首一性可知 $m_1(\lambda) = m_2(\lambda)$. □

定理 4.2.5 (Cayley-Hamilton)**.** *对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其特征多项式是其零化多项式.*

证明: □

推论 4.2.6. *对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 其极小多项式整除其特征多项式.*

命题 4.2.7. *对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 如果 $f(\lambda)$ 是其零化多项式, 则其特征值 λ 也是该多项式的根. 特别地, 是其极小多项式的根.*

注记. 在我们不知道矩阵特征多项式, 只知道其满足某个多项式的时候, 这是一个用来控制矩阵特征值的好办法.

定义 4.2.8. 矩阵 $A \in M_{n\times n}(\mathbb{F})$ 被称为**幂零矩阵** (nilpotent matrix), 如果存在 $k \in \mathbb{N}_{\geq 0}$ 使得 $A^k = 0$.

命题 4.2.9. *幂零矩阵的特征值都为零.*

证明: 如果 A 是幂零矩阵, 则某个单项式 λ^k 是 A 的零化多项式, 特别地, 根据命题4.2.7有幂零矩阵的特征值都为零. □

例子. 对于矩阵

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

其特征多项式 $f(\lambda)=\lambda^2-5\lambda-2$, 其有两个实根记为 λ_1,λ_2. 我们断言其极小多项式 $m(\lambda)=f(\lambda)$, 不然根据推论*4.2.6*有 $m(\lambda)=\lambda-\lambda_1$ 或者 $\lambda-\lambda_2$. 不妨假设 $m(\lambda)=\lambda-\lambda_1$, 假设 v 是 λ_2 对应的特征向量, 则

$$(A-\lambda_1\mathrm{I}_2)v=(\lambda_2-\lambda_1)v\neq 0$$

相矛盾, 从而 $m(\lambda)=f(\lambda)$.

定理 4.2.10. *对于矩阵 $A\in M_{n\times n}(\mathbb{F})$, 其在 $\mathbb{F}$ 上可对角化当且仅当极小多项式 $m(\lambda)$ 在 $\mathbb{F}[\lambda]$ 中可分解为一次因式的乘积, 且没有重根.*

证明: 假设 A 是可对角矩阵, 我们不妨假设 A 就是对角阵, 对角线元素分别是 $\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1\text{个}},\ldots,\underbrace{\lambda_k,\ldots,\lambda_s}_{n_s\text{个}}$, 其中 $\lambda_1,\ldots,\lambda_s$ 互不相同. 此时 A 的极小多项式 $m(\lambda)=(\lambda-\lambda_1)\ldots(\lambda-\lambda_s)$; 另一方面, 假设 A 的极小多项式为 $m(\lambda)=(\lambda-\lambda_1)\ldots(\lambda-\lambda_s)$, 其中 $\lambda_1,\ldots,\lambda_s$ 互不相同, 考虑

$$h_i(\lambda)=\frac{m(\lambda)}{(\lambda-\lambda_i)}$$

则 $\gcd(h_1,\ldots,h_s)=1$, 从而根据裴蜀定理存在 $k_1(\lambda),\ldots,k_s(\lambda)\in\mathbb{F}[\lambda]$ 使得

$$k_1(\lambda)h_1(\lambda)+\cdots+k_s(\lambda)h_s(\lambda)=1$$

即

$$k_1(A)h_1(A)+\cdots+k_s(A)h_s(A)=\mathrm{I}_n$$

从而任取列向量 $v\in\mathbb{F}^n$, 我们可以将其分解为 $v=v_1+\cdots+v_s$, 其中

$$v_i=k_i(A)h_i(A)v$$

如果我们记 $V_i=\ker(\lambda_i\mathrm{I}_n-A)$, 那么

(1) $v_i\in V_i$, 这是因为 $(A-\lambda_i\mathrm{I}_n)v_i=k_i(A)m(A)v=0$.

(2) 任取 $1\leq i\leq s$ 以及 $v\in V_i\cap\sum_{j\neq i}V_j$, 则有

$$0=(k_1(A)h_1(A)+\cdots+k_s(A)h_s(A))v=v$$

从而根据命题3.3.6, 我们将 $\mathbb{F}^n$ 分解成了

$$\mathbb{F}^n=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}$$

根据定理4.1.17可知 A 可对角化. □

推论 4.2.11. *$\mathbb{F}$-线性映射 $T:V\to V$ 可对角化, 假设 W 是 T-不变子空间, 则 $T|_W:W\to W$ 也可对角化.*

证明: 假设 $m_V(\lambda)$ 是 $T:V\to V$ 的极小多项式, 那么其也是 $T|_W:W\to W$ 的零化多项式, 从而 $m_W(\lambda)\mid m_V(\lambda)$. 从而如果 $m_V(\lambda)$ 可以分解为没有重根的一次因式乘积, $m_W(\lambda)$ 也可以. □

定义 4.2.12. 矩阵 $A\in M_{n\times n}(\mathbb{F})$ 称为**幂等矩阵** (idempotent matrix), 如果 $A^2=\mathrm{I}_n$.

命题 4.2.13. *假设 $\mathbb{F}$ 的特征不为 2, 则幂等矩阵 A 在 $\mathbb{F}$ 上可对角化.*

证明: 注意到 λ^2-1 是 A 的零化多项式, 从而其极小多项式整除 λ^2-1. 如果 $\mathbb{F}$ 的特征不为 2, λ^2-1 在 $\mathbb{F}$ 可以分解为 $(\lambda-1)(\lambda+1)$, 从而根据定理4.2.10可知 A 可对角化. □

索引