

代数 0 课程讲义



Instructor: 余成龙
Notes Taker: 刘博文

Qiuuzhen College, Tsinghua University
2022 Spring



目录

第一章 线性方程组	2
1.1 线性函数与线性方程	2
1.2 高斯消元法	4
1.3 线性方程组解的结构	6
第二章 矩阵及其运算	9
2.1 矩阵乘法	9
2.2 矩阵的转置	12
2.3 分块矩阵	13
第三章 线性空间	15
3.1 \mathbb{R}^n 的子空间	15
3.2 线性相关性	16
3.3 维数	18

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 $F: \mathbb{R}^n \rightarrow \mathbb{R}$, 如果存在 $a_1, \dots, a_n \in \mathbb{R}$ 是常数, 使得 F 有如下表达式

$$F(x) = a_1x_1 + \dots + a_nx_n$$

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:

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto (x_1 + 1)^2 - (x_1 - 1)^2 + (x_2 - 1)^2 - (x_1 + 1)^2$$

例子. 如下的 F 不是 \mathbb{R}^2 上的线性函数

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2^2$$

定理 1.1.2. 函数 $F: \mathbb{R}^n \rightarrow \mathbb{R}$ 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

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

则

$$F\left(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right) = a_1x_1, \quad F\left(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}\right) = a_2x_2, \quad \dots$$

那么任取 $X_1, \dots, X_m \in \mathbb{R}^n$, 则

$$\begin{aligned} F(X_1 + X_2 + \dots + X_m) &= F((X_1 + \dots + X_{m-1}) + X_m) \\ &= F(X_1 + \dots + X_{m-1}) + F(X_m) \\ &= F(X_1) + \dots + F(X_m) \end{aligned}$$

因此,

$$F\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = F\left(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right) + \dots + F\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}\right) = F\left(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right) + \dots + F\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}\right) = a_1 x_1 + \dots + a_n x_n$$

□

定义 1.1.3. \mathbb{R}^n 上 m 个线性函数 F_1, \dots, F_m 和 m 个实数 b_1, \dots, 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 $F_1(x) = a_1 x_1 + a_2 x_2$ 以及实数 b_1 给出的线性方程组 $a_1 x_1 + a_2 x_2 = b_1$ 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解, 但是如果解一定只有唯一解.

- (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

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

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 $O_1 O_2$ 的逆为 $O_2^{-1} O_1^{-1}$. \square

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

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

我们将其系数及常数项提出来记作

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

其中 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? 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 1.2.4. 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 $\text{rref}(A)$.

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

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

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

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

则此时 A'' 已经是最简行阶梯型.

2. 如果

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

则作 (E3) 将第 $k+1$ 行加到第 $1, 2, \dots, k$ 行, 将 y'_1, \dots, 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.

□

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 , $\text{rref}(A)$ 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

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

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:

$$x_1 = b_1 - 3x_3 - 5x_4 - x_6$$

$$x_2 = b_2 - 2x_3 - x_4 - 2x_6$$

$$x_5 = b_3$$

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 (\bar{A}, \bar{b}) , 则方程有解等价于 \bar{A} 的零行对应的 \bar{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确定了一组解.

定义 1.3.3. 方程 $Ax = 0$ 称为齐次线性方程组 (system of homogeneous linear equations).

定理 1.3.4. 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

$$A(x + y) = Ax + Ay$$

$$A(cx) = cAx$$

□

定理 1.3.5. 对于线性方程组 $Ax = b$, 如果 \tilde{x} 是其某一解 (特解), 则 $Ax = b$ 的所有解均可唯一的表达为 $x = y + \tilde{x}$, 其中 y 是 $Ax = 0$ 的解.

证明: 只需验证如下两点:

- (1) 验证 $y + \tilde{x}$ 是解.
- (2) 验证当 x 是解时, $x = (x - \tilde{x}) + \tilde{x}$, 其中 $x - \tilde{x}$ 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.

定义 1.3.6. 线性方程有解称其为相容的 (consistent); 无解称为不相容的 (inconsistent); 方程有唯一解称为确定的 (definite).

定义 1.3.7. 对于矩阵 A , $\text{rref}(A)$ 的主元个数称为 A 的秩 (rank), 记作 $\text{rank } A$.

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

推论 1.3.8. 方程 $Ax = b$ 是

- (1) 相容的当且仅当 $\text{rank } A = \text{rank}(A, b)$.
- (2) 确定的当且仅当其为相容的, 并且 $\text{rank } A$ 等于 A 的列数相同.

练习. 对于线性方程组, 如果其增广矩阵有相同的最简行阶梯型, 则其有一样的解集, 反之是否成立?



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, 我们已经接触到了矩阵这个概念. 一个 $m \times n$ 的 (实) 矩阵 (matrix) 定义为

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

也被记作 $(a_{ij})_{m \times n}$. 全体 $m \times n$ 的矩阵的集合记作 $M_{m \times n}(\mathbb{R})$, 并且当 $m = n$ 时, 被称为方阵 (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 $C = (c_1, \dots, c_n)$, 其中 c_i 是列向量. 那么

$$(AB)C = ((AB)c_1, \dots, (AB)c_n)$$

$$A(BC) = A(Bc_1, \dots, Bc_n) = (A(Bc_1), \dots, A(Bc_n))$$

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

$$(AB)C = \begin{pmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_n B \end{pmatrix} C = \begin{pmatrix} a_1 BC \\ a_2 BC \\ \vdots \\ a_n BC \end{pmatrix}$$

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

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

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

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

□

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. 我们看下面的例子.

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 $(E3)$ 可以看作是矩阵左乘.

命题 2.1.2. 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 .

推论 2.1.3. 对 A 做行变换 $O_1 \dots O_k$ 等价于左乘初等矩阵 $B_k \dots B_2 B_1$.

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 $BB'A = I_n$, 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 2.1.4. 对于矩阵 $A \in M_{n \times n}(\mathbb{R})$

- (1) 若有 B 使得 $BA = I_n$, 则称 B 为 A 的**左逆** (left inverse).
- (2) 若有 C 使得 $AC = I_n$, 则称 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 2.1.5. 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) $\text{rref}(A) = I_n$.
- (5) $Ax = b$ 有唯一解.
- (6) $Ax = 0$ 有唯一解.
- (7) $\text{rank } A = n$.

证明: 在求解线性方程组的时候, 我们已经见到了 (4),(5),(6),(7) 的等价性.

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

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 2 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots$$

那么不妨记 $Ax = b_1, Ax = b_2, \dots$ 的唯一解分别是 w_1, \dots, w_n , 则 $A(w_1, \dots, w_n) = I_n$, 即 (w_1, \dots, w_n) 是 A 的右逆.

(3) \implies (2): 假设 A 有右逆, 存在 C 使得即 $AC = I_n$, 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

$$CA = CACD = CD = I_n$$

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) \implies (1), 类似的我们可以证明 (2) \implies (1), 而 (1) \implies (2) 和 (1) \implies (3) 根据定义即可. \square

命题 2.1.6. 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

$$C = CI_n = CAD = I_n D = D$$

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. \square

注记. 上述结论表明, 如果 A 可逆, 那么 $\text{rref}(A) = I_n$, 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 2.1.7. 矩阵 A 可逆当且仅当其为初等矩阵的乘积.

例子. 考虑 $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, 则

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

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, 并且矩阵元 $b_{ij} = a_{ji}$.

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 1.3.7, 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 $\text{rank}(A) = \text{rank}(A^T)$, 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 $\text{rank}(A) = \text{rank}(AB)$. 我们不妨记 $\text{rank}(A) = k, \text{rank}(AB) = l$. 根据线性方程组解的理论可知

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.

1. $Ax = 0$ 有主元 x_{i_1}, \dots, x_{i_k} 以及自由元 $x_{i_{k+1}}, \dots, x_{i_n}$.

2. $ABy = 0$ 有主元 y_{i_1}, \dots, y_{i_l} 以及自由元 $y_{i_{l+1}}, \dots, y_{i_n}$.

由于 $Ax = 0$ 的解与 $ABy = 0$ 的解之间满足 $x = By$, 而根据定理 2.1.5, 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 $k = l$. \square

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

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

$$\begin{pmatrix} I_k & O_{k \times n-k} \\ O_{m-k \times k} & O_{m-k \times n-k} \end{pmatrix}$$

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 $\text{rank}(A) = \text{rank}(A^T) = k$. \square

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:

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

$$PAQ = \begin{pmatrix} I_k & O_{k \times n-k} \\ O_{m-k \times k} & O_{m-k \times n-k} \end{pmatrix}$$

其中 $k = \text{rank}(A)$, 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 $\text{rank}(A) = \text{rank}(B)$.

证明: 显然. \square

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

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

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

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

一方面我们考虑

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

另一方面我们有

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

从而我们有非平凡等式

$$\begin{aligned} (I - \alpha \beta^T)^{-1} &= I + \alpha(1 - \beta^T \alpha)^{-1} \beta^T \\ (1 - \beta^T \alpha)^{-1} &= 1 + \beta^T (I - \alpha \beta^T)^{-1} \alpha \end{aligned}$$

第三章 线性空间

3.1 \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. 线性方程组 $Ax = 0$ 的解是子空间, 也被称为解空间, 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 $\text{rank } A$ 关系密切.

- (1) $\text{rank}(A) = 0$, 此时 A 是零矩阵, 从而 $\ker A = \mathbb{R}^3$.
- (1) $\text{rank}(A) = 1$, 此时 $\ker A$ 是通过原点的平面.
- (1) $\text{rank}(A) = 2$, 此时 $\ker A$ 是通过原点的直线.
- (1) $\text{rank}(A) = 3$, 此时 $Ax = 0$ 只有零解, 即 $\ker A = \{0\}$.

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

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

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

注记. 现在我们来了解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

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

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

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.

命题 3.1.5. 列变换不改变 A 的列空间.

证明: 不妨将 A 写作 $A = (v_1, \dots, v_n)$, A 做列变换得到的矩阵记做 $AB = (\bar{v}_1, \dots, \bar{v}_n)$, 其中 B 是可逆矩阵, 因此 $\bar{v}_1, \dots, \bar{v}_n$ 都是 v_1, \dots, v_n 的线性组合, 从而

$$\text{span}_{\mathbb{R}}\{\bar{v}_1, \dots, \bar{v}_n\} \subseteq \text{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$$

并且由于 A 也可以由 AB 做列变换得到, 从而

$$\text{span}_{\mathbb{R}}\{v_1, \dots, v_n\} \subseteq \text{span}_{\mathbb{R}}\{\bar{v}_1, \dots, \bar{v}_n\}$$

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. □

3.2 线性相关性

定义 3.2.1. $v_1, \dots, v_n \in \mathbb{R}^m$ 被称为**线性无关** (linearly independent), 如果 0 表示为 v_1, \dots, v_n 的线性组合的方式唯一, 即只有

$$0 = 0v_1 + \dots + 0v_n$$

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

命题 3.2.2. 若 v_1, \dots, v_n 线性无关, 则对于 $v \in \text{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$, 其被写成 v_1, \dots, v_n 线性组合式子的系数是唯一的.

证明: 不妨假设

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

则

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

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

定理 3.2.3. $v_1, \dots, v_n \in \mathbb{R}^m$ 线性无关当且仅当 $\text{rank}(A) = n$, 其中 $A = (v_1, \dots, v_n)$. 此时也称 A 是列满秩的.

证明: 注意到 v_1, \dots, 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 $\text{rank}(A) = n$. □

推论 3.2.4. \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 3.2.5. 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. \square

定理 3.2.6. 假设 v_1, \dots, v_n 线性无关, 则 v_1, \dots, v_n, v_{n+1} 线性相关等价于 $v_{n+1} \in \text{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$.

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

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

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

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

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

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

\square

定义 3.2.7. 对于 $v_1, \dots, v_n \in \mathbb{R}^m$, 称 v_{i_1}, \dots, v_{i_k} 是极大线性无关组 (maximal linearly independent set), 如果

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

注记. 若 v_1, \dots, 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, \dots 即可.

定理 3.2.8. 对于 $v_1, \dots, v_n \in \mathbb{R}^m$, v_{i_1}, \dots, v_{i_k} 是其极大线性无关组, 则

$$\text{span}_{\mathbb{R}}\{v_1, \dots, v_n\} = \text{span}_{\mathbb{R}}\{v_{i_1}, \dots, v_{i_k}\}$$

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

$$v_i \in \text{span}_{\mathbb{R}}\{v_{i_1}, \dots, v_{i_k}\}$$

从而 $\text{span}_{\mathbb{R}}\{v_1, \dots, v_n\} = \text{span}_{\mathbb{R}}\{v_{i_1}, \dots, v_{i_k}\}$. \square

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 3.2.9. 假设 v_1, \dots, 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}, \dots, 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

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

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. □

推论 3.2.10. v_1, \dots, v_n 的极大线性无关组中向量的数目是确定的.

证明: 显然. □

3.3 维数





索引

- 不容的, inconsistent, 7
主元, pivot, 5
主元, principal unknowns, 6
列空间, column space, 15
可逆, invertible, 11
右逆, right inverse, 11
基础行变换, elementary row operations, 5
增广矩阵, augmented matrix, 5
子空间, subspace, 15
左逆, left inverse, 11
方阵, square matrix, 9
最简行阶梯型, reduced row echelon form, 5
极大线性无关组, maximal linearly independent set, 17
核, kernel, 15
相容的, consistent, 7
相抵, equivalent, 13
相抵标准型, canonical form, 13
矩阵, matrix, 9
确定的, definite, 7
秩, rank, 7
系数矩阵, coefficient matrix, 5
线性函数, linear function, 2
线性张成, linearly span, 15
线性方程组, solution of system of linear equations, 3
线性方程组, system of linear equations, 3
线性无关, linearly independent, 16
线性相关, linearly dependent, 16
自由元, free unknowns, 6
行变换, row operations, 5
行空间, row space, 15
行阶梯型, row echelon form, 5
转置矩阵, transpose matrix, 12
齐次线性方程组, system of homogeneous linear equations, 7