Linear Algebra

Bowen Liu, Chenglong Yu

目录

1 线性方程组 **2**

1.1 引言 2

1.2 列向量空间与线性函数 3

1.2.1 列向量空间以及加法，数乘运算 3

1.3 高斯消元法和最简阶梯型 8

1.4 线性方程组解的结构 12

1.5 选讲: 最简行阶梯形唯一性的证明 14

1.6 选讲: 舒伯特胞腔 15

1.7 作业一 16

1.7.1 基础题 16

1.8 作业二 18

1.8.1 基础题 18

1.8.2 思考题 19

2 矩阵及其运算 **20**

2.1 矩阵乘法 20

2.2 矩阵的转置 24

2.3 分块矩阵 26

2.4 矩阵的行列式 27

2.5 伴随矩阵 31

2.6 矩阵的若干应用 31

2.6.1 快速傅立叶变换 31

1 线性方程组

1.1 引言

PageRank 是 Google 最早的用来计算网页连接的算法，我们考虑一个简单的模型：例如我们有 1, 2, 3, 4 四处网址，相互的超链接情况如下：

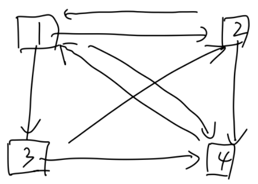

图 1: 简答的网页连接模型

其中 x_i 分别表示第 i 处网页的重要度。对于 1 号网页来说，它的重要度（流量）分成了三份，分别给了 2, 3, 4：

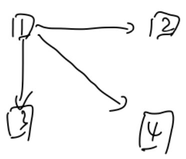

图 2: 1 号网页的流量流出

另一方面，有 2, 4 号网页的流量流入 1 号网页：

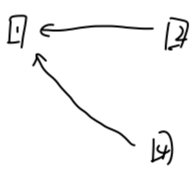

图 3: 1 号网页的流量流入

为了简单起见，我们认为这些流出的流量是被平均分配的，因此根据相互超链

接的关系得到如下的一些等式：

$$\begin{cases} x_1 = \frac{1}{2}x_1 + x_4 \\ x_2 = \frac{1}{3}x_1 + \frac{1}{2}x_3 \\ x_3 = \frac{1}{3}x_1 \\ x_4 = \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3. \end{cases}$$

通过消元法求解，我们可以得到所有的解形如

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = t \cdot \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \\ \frac{4}{3} \\ 1 \end{pmatrix},$$

其中 t 为任意实数。

对于一般的网络连接形式，我们可以问以下问题：

1. 是否一定存在非零解？
2. 解是否都是形如如上的形式？是否各个分量的比例相同，而且可以同为正？
3. 如何求解，以及哪种算法在实际中效率更高？

我们这门课的主要问题：求解线性方程组，了解解的结构。

1.2 列向量空间与线性函数

接下来，我们解释三个方面的问题

1. 在什么集合（空间）上求解？
2. 什么是线性方程？
3. 如何求解？

1.2.1 列向量空间以及加法，数乘运算

我们记 $\mathbb{R}$ 为全体实数的集合。在例子中我们有

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

其中 x_i 都是实数, 这样的 x 被称为一个长度为 4 的实数组. 如果我们考虑长度为 2 的实数组全体构成的集合

$$\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R}\},$$

这个集合记做 $\mathbb{R}^2$, 并且这个集合中的元素一一对应于平面上的点. 同时, $\mathbb{R}^2$ 中的点 P 也可以一一对应与从原点到这个点的向量 $\vec{OP}$, 如下图所示:

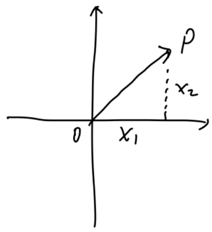

图 4: 点与向量的一一对应

类似的, 我们可以认为 $\mathbb{R}^3$ 一一对应于三维空间中的点.

定义 1.2.1 (列向量空间)**. 列向量空间** *(column vector space)定义为如下的集合*

$$\mathbb{R}^n = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R}\},$$

*其中的元素称为***列向量** *(column vector)*

定义 1.2.2 (加法和数乘)**.** $\mathbb{R}^n$ *上可以定义如下两种运算:*

(1) **加法** *(addition):*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

(2) **数乘** *(scalar product):* 任取实数 $c \in \mathbb{R}$, 定义

$$c\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \\ \vdots \\ cx_n \end{pmatrix}.$$

注 1.2.1. 向量的加法和数乘有其对应的几何意义, 对于 $\mathbb{R}^2$ 来说, 其中向量的加法和数乘的几何意义可以通过如下具体图形象的展示:

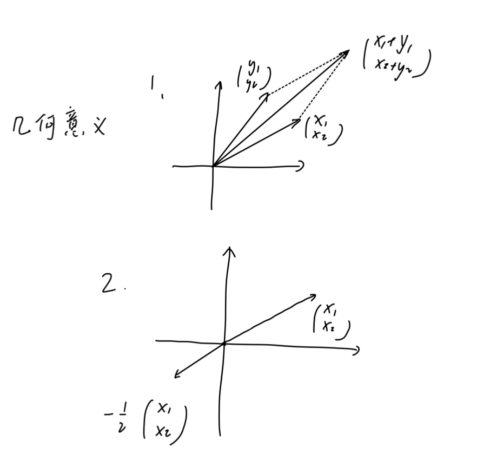

图 5: 向量加法与数乘的几何意义

接下来，我们考虑线性函数:

定义 1.2.3. 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

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

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

例 1.2.1. 如下的 F 是 $\mathbb{R}^2$ 上的线性函数*:*

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

定理 1.2.1 (线性函数的运算性质等价描述)**.** 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,

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

$\square$

命题 1.2.1. 如果 F 是线性函数, 那么 $F(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = 0$.

证明. 任意取 $0 \neq c \in \mathbb{R}$, 根据定义则有

$$F(c\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = cF(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}),$$

因此 $F(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = 0$. □

例 1.2.2. 如下的 F_1, F_2 不是 $\mathbb{R}^2$ 上的线性函数:

$$\begin{aligned} F_1 \colon \mathbb{R}^2 &\to \mathbb{R} \\ \begin{pmatrix}x_1\\x_2\end{pmatrix} &\mapsto x_1 + x_4 + 4, \end{aligned}$$

$$\begin{aligned} F_2 \colon \mathbb{R}^2 &\to \mathbb{R} \\ \begin{pmatrix}x_1\\x_2\end{pmatrix} &\mapsto x_1^2 + x_2^2 \end{aligned}$$

对于函数 F_1, 通过直接的计算得到 $F_1(\begin{pmatrix}0\\0\end{pmatrix}) \neq 0$, 从而利用命题*1.2.1*可知 F_1 不是线性函数. 对于函数 F_2, 我们可以发现

$$F_2(\begin{pmatrix}1\\0\end{pmatrix}) = 1, \quad F_2(2\begin{pmatrix}1\\0\end{pmatrix}) = 4 \neq 2,$$

从而根据线性函数的定义中第二条可知 F_2 不是线性函数.

定义 1.2.4. $\mathbb{R}^n$ 上 m 个线性函数 $F_1, F_2, \ldots, F_m$ 和 m 个实数 $b_1, b_2, \ldots, b_m$ 满足的方程组

$$\begin{cases} F_1(x) = b_1 \\ F_2(x) = b_2 \\ \vdots \\ F_m(x) = b_m \end{cases}$$

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 1.2.2. 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.

1.3 高斯消元法和最简阶梯型

根据注记1.2.2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 1.3.1.

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

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 1.3.1. 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1.3.1, 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 1.3.1. 有限个基础行变换的复合称为**行变换** *(row operations)*.

命题 1.3.1. *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 1.3.1. *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 1.3.2. *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 $(a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{R})$. 当 $m = n$ 时, $A \in M_{n\times n}(\mathbb{R})$ 被称为 n 阶***方阵** *(square matrix).*

例 1.3.2. *$\mathrm{I}_n \in M_{n\times n}(\mathbb{R})$ 是只有 (i, i) 元为 1, 其余分量为零的矩阵, 称为***单位矩阵** *(identity matrix).*

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1.3.1, 我们发现如果我们的增广矩阵有如下的形式, 线性方程组可以直接求解:

(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 1.3.1. *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 $\operatorname{rref} A$.*

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.

最简行阶梯形的唯一性之后再证明，一个好的材料是第一位女菲尔兹奖得主写的短文"A Simple Proof of a Theorem of Schur", https://doi.org/10.1080/00029890.1998.12004879"

□

1.4 线性方程组解的结构

定义 1.4.1. 对于线性方程组的系数矩阵 A, $\operatorname{rref} A$ 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 1.4.1. 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:

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

定理 1.4.1 (线性方程组解的结构定理)**.** 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

(1) 方程有解等价于 $\overline{A}$ 的零行对应的 $\overline{b}_i$ 也是零.

(2) 方程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. 特别地, 方程有唯一解当且仅当没有自由元.

推论 1.4.1. 线性方程组 $Ax = b$

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

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

定义 1.4.2. 方程 $Ax = 0$ 称为**齐次线性方程组** *(system of homogeneous linear equations)*.

定理 1.4.2. 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 1.4.3. 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 1.4.1. 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, 之后我们会用更抽象的观点去描述这种子集, 并称其为一个子空间. 而 $Ax = b$ 的解相当于是将这个子空间做了平移.

定义 1.4.3. 对于矩阵 A, 其主元的个数被定义为 A 的**秩** *(rank)*, 记做 $\operatorname{rank} A$.

定义 1.4.4. 对于线性方程组 $Ax = b$, 我们有如下定义:

(1) 如果方程有解, 我们称这个线性方程组是**相容的** *(consistent)*.

(2) 如果方程无解, 我们称这个线性方程组是**不相容的** *(inconsistent)*.

(3) 如果方程有唯一解, 我们称这个线性方程组是**确定的** *(definite)*.

推论 1.4.2.

(1) 线性方程组 $Ax = b$ 是相容的等价于 $\operatorname{rank} A = \operatorname{rank}(A, b)$.

(2) 线性方程组 $Ax = b$ 是确定的等价于 $Ax = b$ 是相容的, 且 $\operatorname{rank} A$ 等于 A 的列数.

定理 1.4.4. 对于 n 阶方阵 A, 线性方程组 $Ax = b$ 是否有唯一解只取决于 A, 与 b 无关.

证明. 根据推论1.4.2可知, $Ax = b$ 有唯一解等价于 $\operatorname{rank} A = n$. □

例 1.4.2 (Shafarevich-Remizov). 对于互不相同的实数 $c_1,\dots,c_r$, 以及任意实数 $k_1,\dots,k_r$, 存在唯一的次数小于等于 $r-1$ 的多项式 $f(x)=a_0+a_1x+\dots+a_{r-1}x^{r-1}$ 使得对于任意 $i=1,\dots,r$ 使得

$$f(c_i)=k_i \tag{1}$$

证明. 注意到 (1) 是关于 $a_0,\dots,a_{r-1}$ 这 r 个未知元的线性方程组, 不妨记做

$$A\begin{pmatrix}a_0\\a_1\\\vdots\\a_{r-1}\end{pmatrix}=\begin{pmatrix}k_0\\k_1\\\vdots\\k_{r-1}\end{pmatrix}.$$

根据线性方程组的解的结构定理, 上述方程组有唯一解当且仅当下面的线性方程组有唯一解

$$A\begin{pmatrix}a_0\\a_1\\\vdots\\a_{r-1}\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}.$$

这等价于 $c_1,\dots,c_r$ 是 $f(x)$ 的根, 而以 $c_1,\dots,c_r$ 为根的次数不超过 $r-1$ 的多项式是唯一的, 从而说明了 (1) 解的唯一性. □

1.5 选讲: 最简行阶梯形唯一性的证明

令 A 是一个 $m\times n$ 阶的矩阵, 在本节中我们介绍如何证明最简行阶梯形 $\operatorname{rref}A$ 的唯一性.

证明. 我们对 n 做归纳法来证明: 当 $n=1$ 时, $\operatorname{rref}A$ 只有如下两种情况:

$$\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix},\quad\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}.$$

注意到行变换将零矩阵变成零矩阵, 并且由于行变换可逆, 从而行变换将非零矩阵变成非零矩阵. 因此在 $n=1$ 的情况, 最简行阶梯形 $\operatorname{rref}A$ 是否为零完全由 A 是不是零来决定.

现在假设命题对列数是 n 的情况都成立, 对于 $n+1$ 列的矩阵 A, 我们将其写成 $A=(B,y)$ 的形式, 其中 B 是一个 n 列的矩阵, y 是一个列向量. 假设此时 $\operatorname{rref}A$ 有 $A'=(B',y')$ 和 $A''=(B'',y'')$ 两种形式, 则 B' 和 B'' 都是由 B 经过行变换得到, 且都是最简行阶梯形, 从而根据归纳假设 $B'=B''$.

现在考虑线性方程组

$$B\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=y. \tag{2}$$

由于行变换不改变线形方程组解的情况, 从而方程组 (2) 等价于如下两种方程组

$$B\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=y', \quad B\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=y''.$$

对于方程组 (2), 我们有如下两种情况:

(i) (2) 无解, 此时是不相容情形. 假设 B 的秩为 r, 那么

$$y'=\begin{pmatrix}0\\0\\\vdots\\1\\0\\\vdots\\0\end{pmatrix}=y'',$$

其中 1 位于第 $r+1$ 行.

(ii) (2) 有解. 假设 B 的秩为 r, 并且主元分别在 $i_1,\dots,i_r$ 列. 令自由元均取 0, 则得到关于 $x_{i_1},\dots,x_{i_r}$ 这 r 个未知元的线性方程组, 并且有解:

$$x_{i_1}=y'_1,\dots,x_{i_r}=y'_r,$$

或者

$$x_{i_1}=y''_1,\dots,x_{i_r}=y''_r.$$

但由于此时解是唯一的, 从而 $y'=y''$, 从而得到最简行阶梯形 $\operatorname{rref}A$ 是唯一的.

□

1.6 选讲: 舒伯特胞腔

1.7 作业一

这次作业里所有矩阵不加说明都是实数矩阵.

1.7.1 基础题

本部分题必做.

习题 1.7.1. 用消元法解线性方程组

1. 关于两个变元 x_1, x_2 的线性方程组

$$\begin{cases} x_1 - 2x_2 & = 1 \\ 2x_1 - x_2 & = 2 \end{cases}$$

2. 关于四个变元 $x_1, x_2\ x_3, x_4$ 的线性方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 & = 4 \\ x_2 - x_3 + x_4 & = -3 \\ x_1 + 3x_2 + x_4 & = 1 \\ -7x_2 + 3x_3 + x_4 & = -3 \end{cases}$$

3. 关于三个变元 x, y, z 的线性方程组

$$\begin{cases} 2x + y - z = 8 \\ -3x - y + 2z = -11 \\ -2x + y + 2z = -3 \end{cases}$$

习题 1.7.2. 讨论 λ 取何值时，线性方程组

$$\begin{cases} \lambda x_1 + x_2 + x_3 & = 1 \\ x_1 + \lambda x_2 + x_3 & = \lambda \\ x_1 + x_2 + \lambda x_3 & = \lambda^2 \end{cases}$$

有唯一解，无穷多解，无解，并在有解时求其解。

习题 1.7.3. 考虑包含 m 个主元的最简阶梯型的 $m \times n$ 的实矩阵, 如果主元出现的位置相同的这样的矩阵视作一类，求不同的种类数有多少？

习题 1.7.4. 考虑一个连通无向无圈无多重边的有限图 G. 令 V 是顶点的集合, 将其视为网页，构造如下网络. 如果两个顶点之间有边连接，则假设两个网页之间有彼此两个方向之间的超链接连接. 我们利用 *Google* 的 *PageRank* 算法得到关于每个网页重要性 $(x_i)_{i \in V}$ 的线性方程组. 证明此时 x_i 等于经过顶点 i 的边数是这个方程组的一组解.

习题 1.7.5. 构造一个 *3* 阶方阵, 其 *9* 个元素各不相同, 且行简化阶梯形有且只有一个主元.

习题 1.7.6. 设

$$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

证明

1. $Ax = b$ 有解当且仅当 $b_1 + b_2 + b_3 = 0$.
2. $Ax = 0$ 的解集是 $\{kx_1 : k \in \mathbb{R}\}$, 其中 $x_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.
3. 当 $Ax = b$ 有解时, 若 x_0 是一个解, 则解集是 $\{x_0 + kx_1 : k \in \mathbb{R}\}$.

1.8 作业二

这次作业里所有矩阵不加说明都是实数矩阵.

1.8.1 基础题

本部分题必做.

习题 1.8.1. 把下列矩阵化为最简行阶梯型: (默认空格处为 0)

1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{pmatrix}$

4. $\begin{pmatrix} 2 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$

习题 1.8.2. 考虑一个连通无向无圈无多重边的有限图 G. 假设 V 是顶点的集合, 其中只有一条边相连的顶点称为边界点, 有多条边相连的顶点称为内部点. 假设内部点和边界点的集合都非空。对每一个顶点 i 取一个温度 $T_i \in \mathbb{R}$, 称 $T = (T_i)_{i\in V}$ 是一个图上的温度分布. 如果每一个内部点的温度等于与之相连的点的温度的平均值, 则称这一分布称为稳定的的. 证明: 对于每一组边界点的温度值, 存在唯一的内部点的温度取值, 使得这一温度分布是稳定的.

习题 1.8.3. 将下列问题转化为求解线性方程组的问题, 并求解:

1. 设 2×2 矩阵 A 满足 $A\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}4\\8\end{pmatrix}$ 且 A 的第一列元素之和为 2, 求所有可能的 A.

2. 空间中有一个平面经过点 $\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}$, 求所有与该平面垂直的向量.

3. 写出通过 5 点 $M_1(0,1), M_2(2,0), M_3(-2,0), M_4(1,-1), M_5(-1,-1)$ 的二次曲线的方程. 这里二次曲线是 xy-平面上形如 $ax^2+bxy+cy^2+dx+ey+f=0$ 的方程决定的曲线.

习题 1.8.4. 若 A, C 均为 $m\times n$ 矩阵, 如果对任何 b, 线性方程组 $Ax=b$ 与 $Cx=b$ 都有相同的解集, 是否一定有 $A=C$?

1.8.2 思考题

本部分题选做，学期中任何时间都可以交，不计成绩。

习题 1.8.5. （还没讲到，2 月 25 日课会讲）课上我们研究过 $G(m,n)$ 的分解中 $\mathbb{R}^i$ 的个数, 记为 b_i.

1. 求 b_i 的生成函数 $\sum_i b_i t^i$.

2. 验证 $b_i = b_{m(n-m)-i}$.

3. 任取一组正实数 m, n，验证 b_i 是单峰的 (先单调递增后单调递减).

习题 1.8.6. 若 A, A' 均为 $m\times n$ 矩阵, b, b' 为 m 维向量, 方程 $Ax=b$ 与 $A'x=b'$ 的解集相同且非空, 请思考 (A', b') 是否一定可由 (A,b) 经过行变换得到，你能对 $m=n=2$ 写出证明吗？

2 矩阵及其运算

2.1 矩阵乘法

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 $cA:=(ca_{ij})_{m\times n}$.

除了以上两种结构, 还满足额外的乘法结构, 在以前我们已经见到过了类似的例子: 例如 $x\in\mathbb{R}^n$ 和 $A\in M_{m\times n}(\mathbb{R})$, 如果我们将 A 写作是列向量的形式 $A=(v_1,\ldots,v_n)$, 那么

$$Ax:=x_1v_1+\cdots+x_nv_n,$$

也即是以 $x_1,\ldots,x_n$ 为系数的 A 中的列向量 $v_1,\ldots,v_n$ 的线性组合.

定义 2.1.1. 对于向量 $v_1,\ldots,v_n\in\mathbb{R}^m, c_1,\ldots,c_n\in\mathbb{R}$, 则 $c_1v_1+\cdots+c_nv_n$ 称为以 $c_1,\ldots,c_n$ 为系数的 $v_1,v_2,\ldots,v_n$ 的**线性组合** *(linear combination)*.

定义 2.1.2. 对于 $A\in M_{m\times n}(\mathbb{R})$ 以及 $B=(w_1,\ldots,w_\ell)\in M_{n\times\ell}(\mathbb{R})$, 则**矩阵乘法** *(matrix multiplication)*定义为

$$AB:=(Aw_1,\ldots,Aw_\ell).$$

矩阵乘法还可以用如下等价的方式定义:

定义 2.1.3. 对于 $A\in M_{m\times n}(\mathbb{R}), B\in M_{n\times l}(\mathbb{R})$, 则**矩阵乘法** *(matrix multiplication)* 定义为 $AB:=(c_{ij})_{m\times l}$, 其中

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

注 2.1.1. 对于矩阵 A,B, 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.

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

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

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

$\square$

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 2.1.4. 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 2.1.1. 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 2 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 2.1.2. 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 E.

证明. 根据定义验证即可. □

推论 2.1.1. 对 A 做行变换 $O_1 \dots O_k$ 等价于左乘初等矩阵 $E_k \dots E_2 E_1$.

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

(2) 若有 C 使得 $AC = \mathrm{I}_n$, 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 2.1.1. 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) $\operatorname{rref} A = \mathrm{I}_n$.

(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1.4.1, 我们已经证明了 (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

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

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

注 2.1.2. 至此已经证明了如果矩阵 A 存在左逆, 那其一定存在右逆, 即 (2) $\implies$ (3).

(3) $\implies$ (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.

注 2.1.3. 从上述证明可以看出，如果 A 存在左逆，那么其右逆不仅存在，并且还和左逆相同．类似的可以说明如果 A 存在右逆则其左逆不仅存在，也与右逆相同．

□

命题 2.1.3. 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 2.1.4. 上述结论表明，如果 A 可逆，那么 $\operatorname{rref} A = \mathrm{I}_n$，而根据推论*2.1.1*可知行变换等价于左乘初等矩阵，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 2.1.2. 矩阵 A 可逆当且仅当其为初等矩阵的乘积．

例 2.1.2. 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 $A^{-1} = \left(\begin{smallmatrix}-2 & 1\\ \frac{3}{2} & -\frac{1}{2}\end{smallmatrix}\right)$

注 2.1.5. 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 $b_{ij} = a_{ji}$．

定义 2.2.2. 矩阵 $A \in M_{n\times n}(\mathbb{R})$ 被称为**对称矩阵** *(symmetric matrix)*，如果 $A^T = A$．

例 2.2.1. 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 2.2.1. *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 2.2.1. *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2.2.1的 (2) 即可. □

回忆定义1.4.3, 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 2.2.1. *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2.2.1, 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 $x_{i_1}, \dots, x_{i_k}$ 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 $y_{i_{l+1}}, \dots, y_{l_n}$.

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 2.1.1 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 2.2.2. *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 2.2.3. *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 2.2.3. *(实) 矩阵 A, B 之间被称为***相抵** *(equivalent), 如果存在 (实) 可逆矩阵 P, Q 使得 $PAQ = B$.*

定理 2.2.2. *$m \times n$ 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 2.3.1. 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

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

另一方面我们有

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

从而我们有非平凡等式

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

2.4 矩阵的行列式

回忆在注2.1.5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

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

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 2.4.1. *$\mathbb{R}^n$ 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 2.4.2. 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 2.4.1.

例 2.4.1. 当 $n=2$ 时

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

例 2.4.2. 当 $n=3$ 时

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

注 2.4.1. 上述公式又被称为对角线法则.

例 2.4.3. 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) $\det E[ij,c]=1$.

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

命题 2.4.1. 行列式有如下性质:

(1) 如果 A 的某一列为零, 则 $|A|=0$.

(2) $|A|\neq 0$ 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 $A_1,\dots,A_n$, 则 $|A|=|A_1|\dots|A_n|$. 特别地, 如果 A 是上三角矩阵, 并且对角线元素为 $a_1,\dots,a_n$, 则 $|A|=a_1\dots a_n$.

证明. (1). 根据定义2.4.1中 (1) 即可.

(2). 根据例2.4.3可知, 如果 E 是初等矩阵, 则 $|AE|=|A||E|$, 因此因此不妨将 A 写作 $E_k\dots E_1 \operatorname{rref} A$, 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.

[2] 见定义2.1.4

(4). 假设 B 不可逆, 即 $\operatorname{rank} B < n$, 根据 (2) 则有 $|B| = 0$. 而根据命题??有 $\operatorname{rank} AB \le \operatorname{rank} B < n$, 因此 $|AB| = 0$, 即 $|AB| = |A||B|$. 假设 B 可逆, 则根据推论2.1.2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2.2.2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2.2.2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 2.4.2. 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]. □

[3]这个操作称为微扰法, 是矩阵中一个非常经典的技巧.

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

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

推论 2.5.2 (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

$$\begin{aligned} f(x) &= a_0 + a_1 x + \cdots + a_d x^d \\ g(x) &= b_0 + b_1 x + \cdots + b_d x^d \end{aligned}$$

多项式乘积为

$$fg := a_0 b_0 + (a_1 b_0 + b_1 a_0)x + \ldots$$

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 $O(d \log d)$.

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

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

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

习题 2.6.1. $|M| = \prod_{i<j}(x_i - x_j) \neq 0$, 即 M 可逆.

例 2.6.1. 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

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