10 月 13 日线代大作业（2）
第一、二章整理

王凯灵

2021.10.19

1 矩阵概念和运算

1.1 引入

1.1.1 矩阵的定义

设 K 为数域, K 上 m 个 n 元线性型组成的线性型组记为如下形式:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases}$$

其中,

$$a_{ij} \in K, \quad 1 \le i \le m, \quad 1 \le j \le n$$

将以上 mn 个系数安排成如下方形的表:

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

这就是一个矩阵.

定义 1. 设 K 为数域, $a_{ij} \in K, 1 \leq i \leq m, 1 \leq j \leq n$, 令

$$\boldsymbol{A} = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

于是，我们有了矩阵的基本定义.

1.1.2 常用矩阵

下面给出几类常用的矩阵.

(1) 零矩阵

设 $\boldsymbol{A} = (a_{ij})_{m\times n} \in K^{m\times n}$, 若对 $1 \leq i \leq m$ 与 $1 \leq j \leq n$ 都有 $a_{ij} = 0$, 则称 $\boldsymbol{A}$ 为一个 $m \times n$ 零矩阵 (zero matrix), 记作 $\boldsymbol{O}_{m\times n}$ 或 $\boldsymbol{O}$.

(2) 对角矩阵

设 $\boldsymbol{A} = (a_{ij})_{n\times n} \in K^{n\times n}$, 若对 $1 \leq i, j \leq n$, 当 $i \neq j$ 时都有 $a_{ij} = 0$, 即

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$$

则称 $\boldsymbol{A}$ 为 n 阶对角矩阵 (diagonal matrix).

(3) 纯量矩阵

设 $\boldsymbol{A}$ 为 n 阶对角矩阵且 $a_{11} = a_{22} = \cdots = a_{nn} = k$, 即

$$\boldsymbol{A} = \begin{pmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{pmatrix}$$

则称 $\boldsymbol{A}$ 为 n 阶纯量矩阵 (scalar matrix) 或数量矩阵. 特别当 $k = 1$, 即

$$A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

时, A 叫做 n 阶单位矩阵 (identity matrix), 记为 E_n 或 $\boldsymbol{E}$.

(4) 三角形矩阵

设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in K^{n\times n}$. 若当 $1\le j<i\le n$ 时, 都有 $a_{ij}=0$, 即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}$$

则称 $\boldsymbol{A}$ 为 n 阶上三角形矩阵 (upper triangular matrix); 若当 $1\le i<j\le n$ 时, 都有 $a_{ij}=0$, 即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

则称 $\boldsymbol{A}$ 为 n 阶下三角形矩阵 (lower triangular matrix).

(5) 对称矩阵与反对称矩阵

设 $\boldsymbol{A}=(a_{ij})_{m\times n}\in K^{m\times n}$, 令 $\boldsymbol{B}=(b_{ij})_{n\times m}\in K^{n\times m}$, 此处

$$b_{ij}=a_{ji},\quad 1\le i\le n,\quad 1\le j\le m$$

B 称为 A 的转置矩阵 (transposed matrix), 记作 $\boldsymbol{A}^{\mathrm{T}}$ 或 $\boldsymbol{A}'$. 显然 $\left(A^{\mathrm{T}}\right)^{\mathrm{T}}=A$. 设 $\boldsymbol{A}\in K^{n\times n}$, 若 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$, 即

$$a_{ji}=a_{ij},\quad 1\le i,j\le n$$

则称 $\boldsymbol{A}$ 为 n 阶对称矩阵 (symmetric matrix); 若 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$, 即

$$a_{ji}=-a_{ij},\quad 1\le i,j\le n$$

则称 $\boldsymbol{A}$ 为 n 阶反对称矩阵 (anti-symmetric matrix).

(6) 向量

设 $m=1,\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)\in K^{1\times n}$, 则称 $\boldsymbol{\alpha}$ 为 K 上的 n 维行向量 (row vector of dimension n); 设 $n=1$,

$$\boldsymbol{\beta}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}\in K^{m\times 1}$$

则称 $\boldsymbol{\beta}$ 为 K 上 m 维列向量 (column vector).

1.1.3 注

虽然这些东西极其基础, 但是作为线性代数的基本概念, 所以全部列出. 以后的部分才是根据我是否觉得重要删减的.

1.2 加法和乘法

1.2.1 加法

不妨仍然把矩阵看作是方程的系数, 给定两个方程组, 对应系数和常数叠加, 就能得到新的方程组. 实际上, 当我们把矩阵看作一组向量时, 因为每个向量的相加都得到了一个新的向量, 于是我们或许得到了一个新的空间. 当然, 这个新空间维数变化和相加的矩阵本身有关.

很自然的, 我们有:

$$\text{若 } \boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}, \boldsymbol{A}=(a_{ij})_{m\times n}\in K^{m\times n}\boldsymbol{B}=(b_{ij})_{n\times m}\in K^{n\times m}$$

定义 2. 则

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

即 $\boldsymbol{C}=(c_{ij})_{m\times n}$, 其中

$$c_{ij}=a_{ij}+b_{ij}, 1\le i\le m, 1\le j\le n$$

1.2.2 数乘

当所有系数都乘上一个数时,

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

由此很自然地引出纯量乘法的定义.

定义 3.

$$\boldsymbol{A}=(a_{ij})_{m\times n}\in K^{m\times n}$$

令 $D=(d_{ij})_{m\times n}$, 其中

$$d_{ij}=ka_{ij}, 1\le i\le m, 1\le j\le n$$

则称矩阵 $\boldsymbol{D}$ 为数 k 与矩阵 $\boldsymbol{A}$ 的纯量积, 记作 $\boldsymbol{D}=k\boldsymbol{A}$.

以上二者是矩阵的线性变换. 这两者的性质和以前接触的其他加法, 乘法相比, 在结合交换等方面并没有特别的地方. 值得一提的是, 相加再转置, 转置可以视为分配. 数乘再转置, 常数可以提出转置符.

减法, 数量的除法, 都十分简单, 故不赘述.

1.2.3 矩阵的乘法

这个神奇的玩意刚上来实在是令我有点费解. 好好的方程组系数, 有啥子好乘的. 书上的解释是方程的变换. 事实上, 在其他课程的学习中, 我意识到矩阵能够表达诸如伸缩, 旋转的变换, 乘法被我理解为将过程进行封装. 不过, 那是以后研究的内容. 现在先给上定义:

定义 4. 设

$$\boldsymbol{A}=(a_{ij})_{m\times n}\in K^{m\times n},\quad \boldsymbol{B}=(b_{ij})_{n\times s}\in K^{n\times s}$$

令 $\boldsymbol{C}=(c_{ij})_{m\times s}$, 其中

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad 1\le i\le m,\quad 1\le j\le s$$

则称 $m\times s$ 矩阵 C 为矩阵 A 与 B 的乘积 *(product)*, 记作 $C=AB$.

显然能够看出两个矩阵的行列对应关系. 想象成消消乐, 这一点其实很好记.

乘法具有结合律, 分配律. 单位阵相当于 1, 零矩阵相当于 0. 常数可以放在任意一个矩阵前. 这些都很好理解. 比较需要注意的就是不能随便消去. 说真的这一点很烦人.

以下是一些结论:

有些乘法可以交换, 和任意矩阵可交换 $\Leftrightarrow$ 单位阵.

同一个方阵相乘多次, 就是幂运算, 性质平凡.

乘后转置, 由于行列数对应, 不难理解 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$

1.3 迹

定义 5. 迹 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 则称其主对角线上的元素之和为 $\boldsymbol{A}$ 的迹 *(trace)*, 记作 $\operatorname{tr}(\boldsymbol{A})$, 即

$$\operatorname{tr}(\boldsymbol{A})=\sum_{i=1}^{n}a_{ii}$$

设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 阶方阵, k 为数, 以下简单列出性质, 不做说明, 因为我暂时没理解这有啥用:

(1) $\operatorname{tr}\left(\boldsymbol{A}^{\mathrm{T}}\right)=\operatorname{tr}(\boldsymbol{A})$

(2) $\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{tr}(\boldsymbol{A})+\operatorname{tr}(\boldsymbol{B})$.

(3) $\operatorname{tr}(k\boldsymbol{A})=k\operatorname{tr}(\boldsymbol{A})$.

(4) $\operatorname{tr}(\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BA})$.

1.4 行列式

1.4.1 排列

主要了解逆序数, 就是一串数中不按顺序来的个数.

1.4.2 行列式的值

定义 6. 行列式 设 $n \geq 1, \boldsymbol{A} = (a_{ij})_{n\times n} \in K^{n\times n}$. 令

$$|\boldsymbol{A}| = \sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

其中 $\sum_{j_1 j_2 \cdots j_n}$ 表示对所有的 n 级排列求和, 则称 $|\boldsymbol{A}|$ 为 n 阶方阵 $\boldsymbol{A}$ 的行列式或 n 阶行列式, 而把

$$\sum_{j_1 j_2 \cdots j_n} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

叫做 $|\boldsymbol{A}|$ 的展开式. $|\boldsymbol{A}|$ 也常记做

$$|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

上面给出的行列式计算是按行次序来的. 不难发现, 按列来也是以一样的. 所以转置不改变行列式的值:

$$|\boldsymbol{A}| = \left|\boldsymbol{A}^{\mathrm{T}}\right|$$

直接上定义, 记这么几个结论:

当 $n = 1$ 时, 总共只有一个排列: 1. 显然 $\tau(1) = 0$, 因此

$$|\boldsymbol{A}| = |(a_{11})| = (-1)^{\tau(1)} a_{11} = a_{11}$$

当 $n = 2$ 时, 共有 2 个排列: 12,21 . 由于 $\tau(12) = 0, \tau(21) = 1$. 因此

$$\begin{aligned} |\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= \sum_{j_1 j_2} (-1)^{\tau(j_1 j_2)} a_{1j_1} a_{2j_2} \\ &= (-1)^{\tau(12)} a_{11} a_{22} + (-1)^{\tau(21)} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21} \end{aligned}$$

当 $n = 3$ 时, 共有 6 个排列且它们的逆序数为

$$\tau(123) = 0, \tau(231) = \tau(312) = 2, \tau(132) = \tau(213) = 1, \tau(321) = 3$$

因此

$$
\begin{aligned}
|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \sum_{j_1j_2j_3} (-1)^{\tau(j_1j_2j_3)} a_{1j_1}a_{2j_2}a_{3j_3} \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\
&\quad (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33})
\end{aligned}
$$

1.4.3 行列式的计算

超过五阶, 手算就不太现实了. 但是, 我们发现, 当方阵对角线以上或以下都是 0 时, 行列式会很好计算. 这时, 行列式的值显然就是对角线数的乘积. 副对角线也类似, 不过会多一个负号. 这应该是一种很常用的化简计算行列式的思路.

另一种思路就是变换行列式. 有以下三种:

(1) 行列式某行乘以常数 k 时, 显然对于定义 6 的每个式子都乘了 k, 那么行列式也变为 k 倍. 即: 行列式某一行可以提一个常数出来.

(2) 行列式某两行交换, 显然每个逆序数奇偶都改变了, 所以会多一个负号.

(3) 行列式某一行的非零倍加到另一行, 行列式不变. 因为

$$
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_1 & b_2 & \cdots & b_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} & \end{vmatrix}
$$

$$
= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_1 + c_1 & b_2 + c_2 & \cdots & b_n + c_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}
$$

而且如果一个行列式中两行成比例, 交换之, 发现行列式等于负的它自己, 所以是零. 显然有这一结论.

对于列而言, 以上性质都完全一样.

1.4.4 行列式的展开

定义 7. 代数余子式 设 D 为 n 阶行列式, $1 \le k \le n, 1 \le i_1 < i_2 < \cdots < i_k \le n, 1 \le j_1 < j_2 < \cdots < j_k \le n$. 在 D 中取定第 $i_1, i_2, \cdots, i_k$ 行与第 $j_1, j_2, \cdots, j_k$ 列, 将位于这 k 行

k 列交点上的 k^2 个元素依原来的顺序组成一个 k 阶行列式 M, 称为 D 的一个 k 阶子式 *(minor)*. 当 $k < n$ 时, 在 D 中划去这 k 行 k 列后余下的元素按原来顺序组成的 $n-k$ 阶行列式 M' 称为 M 的余子式 *(cofactor)*. 再令

$$A = (-1)^{(i_1+i_2+\cdots+i_k)+(j_1+j_2+\cdots+j_k)} M'$$

则 A 叫做 M 的代数余子式 *(algebraic cofactor)*.

就我个人理解, 代数余子式的产生就是为拉普拉斯展开服务的.

定义 8. *Laplace* 定理 设 D 为 n 阶行列式, $1 \le k \le n-1$. 在 D 中任意取定 k 行, 由这 k 行元素所组成的全体 k 阶子式记作 $M_1, M_2, \cdots, M_t$, 此处 $t = \mathrm{C}_n^k$. 对 $1 \le i \le t$, 令 M_i 的代数余子式为 A_i, 则

$$D = \sum_{i=1}^{t} M_i A_i = M_1A_1 + M_2A_2 + \cdots + M_tA_t$$

按列也能展开, 显然.

紧挨着, 我列出一些行列式的常用技巧:

1.4.5 行列式的技巧

除去之前提过的三角化, 书上的爪形行列式看着又很鸡肋, 故我挑几个好用的:

(1) 数学归纳法永垂不朽.

(2) 递推关系法, 对于一些比较规律的行列式展开递推.

(3) 利用添加一边构造. 常用的有加一列 1 等.

(4) 往范德蒙德行列式靠拢.

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

1.4.6 分块矩阵

分块很好用, 但是性质和普通矩阵差不多. 倒不如说普通矩阵是特殊的分块矩阵.

1.4.7 来自习题的结论

(1) 上 (下) 三角矩阵的乘积仍是上 (下) 三角矩阵.

(2) 若 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{O}$, 则 $\boldsymbol{A} = \boldsymbol{O}$

(3)$(A^*)^{-1}=(A^{-1})^*=\frac{1}{|\boldsymbol{A}|}|\boldsymbol{A}|$, $\quad(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n-2}\boldsymbol{A}$

(4) 若 A 为 n 阶方阵. 对于任意 $n\times 1$ 矩阵 α 都有 $\alpha^{\mathrm{T}}A\alpha=0$ 的充分必要条件为 $\boldsymbol{A}$ 是反对称矩阵.

(5) 若 $\boldsymbol{A}$ 为 n 阶非零的对称矩阵, 存在 $n\times 1$ 矩阵 $\boldsymbol{\alpha}$, 使得 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\neq 0$.

(6) 若 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 证明:

1. $\operatorname{tr}(\boldsymbol{A})=\operatorname{tr}(\boldsymbol{B})$
2. $\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{tr}(\boldsymbol{A})+\operatorname{tr}(\boldsymbol{B})$
3. 若 k 为常数, 则 $\operatorname{tr}(k\boldsymbol{A})=k\operatorname{tr}(\boldsymbol{A})$;
4. 若 $\boldsymbol{A}$ 为可逆矩阵, 则 $\operatorname{tr}(\boldsymbol{A}^*)=|\boldsymbol{A}|\operatorname{tr}\left(\boldsymbol{A}^{-1}\right)$;
5. $\operatorname{tr}\left(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)-\operatorname{tr}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)=0$

(7) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 阶可逆矩阵, 则 $(\boldsymbol{A}\boldsymbol{B})^*=\boldsymbol{B}^*\boldsymbol{A}^*$.

(8) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 阶可逆矩阵, 则 $(-\boldsymbol{A})^*=(-1)^{n-1}\boldsymbol{A}^*$,$\left(\boldsymbol{A}^{\mathrm{T}}\right)^*=(\boldsymbol{A}^*)^{\mathrm{T}}$.

(9) 若 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为上（或下）三角形矩阵. 则:

1. $\boldsymbol{A}$ 可逆的充要条件为 $a_{ii}\neq 0, i=1,2,\cdots,n$;
2. 若 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{A}^{-1}$ 仍为上 (或下) 三角形矩阵;
3. 若 $\boldsymbol{A}$ 可逆, 记 $\boldsymbol{A}^{-1}=(b_{ij})_{n\times n}$, 则 $a_{ii}b_{ii}=1, i=1,2,\cdots,n$.

2 矩阵的秩

感觉好多东西都是再给矩阵的秩做铺垫. 矩阵有了秩的概念, 简直是获得了新生.

2.1 矩阵变换

2.1.1 初等变换

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

再设 $1\leq i,j\leq m, i\neq j$, 对矩阵 $\boldsymbol{A}$ 施行以下三种变换:

(1) 交换 $\boldsymbol{A}$ 的第 i 行和第 j 行, 记作 $r_i\longleftrightarrow r_j$;

(2) 设 $k\in K, k\neq 0$, 用 k 去乘 $\boldsymbol{A}$ 的第 i 行, 记作 $k\cdot r_i$;

(3) 设 $k\in K$, 将 $\boldsymbol{A}$ 的第 j 行的 k 倍加到第 i 行, 记作 $r_i+k\cdot r_j$. 上述三种变换依次叫做矩阵的第一类、第二类和第三类行初等变换 (row elementary transformation).

同样地, 把行改成列, 也有三种变换. 但是列变换不利于我们解方程组.

下面引入阶梯形矩阵的概念.

定义 9. 设 $\boldsymbol{A}$ 为数域 K 上的 $m\times n$ 矩阵, $0\le r\le m, \boldsymbol{A}$ 恰有 r 个非零行 *(即至少有一个个非零元素的行), 其余 $m-r$ 行为全零行 (即所有元素都是零的行). 若 $\boldsymbol{A}$ 满足以下两个条件:*

(1) $\boldsymbol{A}$ 的前 r 行为非零行, 后 $m-r$ 行为全零行;

(2) 对 $1\le i\le r$, 设 $\boldsymbol{A}$ 的第 i 行中最左边的非零元为 a_{ij_i}, 则

$$1\le j_1<j_2<\cdots<j_r\le n$$

则称 $\boldsymbol{A}$ 是一个有 r 级阶梯的行阶梯形矩阵 (row echelon matrix), 各非零行中最左边的非零元叫主元 (pivot).

定义 10. 设 $\boldsymbol{A}$ 为行阶梯形矩阵, 如果 $\boldsymbol{A}$ 还满足以下条件:

(3) 每个主元都是 1;

(4) 各个主元所在列的其余元素都是零.

则称 $\boldsymbol{A}$ 为一个简化行阶梯形矩阵 (reduced row echelon matrix).

2.1.2 用矩阵变换矩阵

定义 11. 设 $\boldsymbol{E}$ 为数域 K 上的 n 阶单位矩阵. 令

$$\boldsymbol{E}(i,j)=\begin{array}{cc}\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&0&\cdots&1&&&\\&&&\vdots&\ddots&\vdots&&&\\&&&1&\cdots&0&&&\\&&&&&&1&&\\&&&&&&&\ddots&\\&&&&&&&&1\end{pmatrix}&\begin{array}{c}\\ \\ \\ (\text{第}i\text{ 行})\\ \\ (\text{第}j\text{ 行})\\ \\ \\ \\ \end{array}\end{array},$$

$$\boldsymbol{E}(i(k))=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&k&&\\&&&\ddots&\\&&&&1\end{pmatrix}(\text{第 } i \text{ 行})$$

$$\boldsymbol{E}(i,j(k))=\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & k & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}\begin{matrix} \\ \\ (\text{第}i\text{ 行}) \\ \\ (\text{第}j\text{ 行}) \\ \\ \\ \end{matrix}.$$

$\boldsymbol{E}(i,j), \boldsymbol{E}(i(k))$ 与 $\boldsymbol{E}(i,j(k))$ 依次称为第一类、第二类和第三类初等矩阵 (*elementary matrix*). 由定义可知, $\boldsymbol{E}(i,j)$ 是由交换单位矩阵 $\boldsymbol{E}$ 的第 i 行与第 j 行所得的矩阵, $\boldsymbol{E}(i(k))$ 是用非零数 k 去乘 $\boldsymbol{E}$ 的第 i 行所得的矩阵, $\boldsymbol{E}(i,j(k))$ 则是将 $\boldsymbol{E}$ 中第 j 行的 k 倍加到第 i 行得到的矩阵. 关于初等矩阵的行列式, 我们有

$$|\boldsymbol{E}(i,j)|=-1, |\boldsymbol{E}(i(k))|=k, |\boldsymbol{E}(i,j(k))|=1$$

把这些初等矩阵乘在矩阵右边, 就相当于对应的初等行变换. 左边就是列变换.

用这些变换, 我们可以将任何矩阵变换成简化阶梯矩阵. 这是很显然的, 其实就是解方程的过程.

并且, 由于变换是可逆的, 我们可以逆向变换从简化阶梯矩阵得到原来的矩阵. 这时, 简化阶梯矩阵右乘的一系列矩阵可以化简为一个矩阵. 利用这个思想, 我们在矩阵右边填上一个方阵, 这样就记录了矩阵的行变换. 当然, 如果是列变换的话只要在下方添置一个矩阵就行了. 我们发现, 这样求逆矩阵比原来方便多了. 所以我甚至跳过了原来的求法.

很自然的, 我们会关心简化阶梯矩阵有几行 (列) 来考察方程解问题, 这玩意可以用秩来考虑.

2.2 秩

定义 12. 设 $\boldsymbol{A}$ 为数域 K 上的 $m \times n$ 矩阵. 若 $\boldsymbol{A}$ 中有一个 r 阶子式不等于零而所有高于 r 阶的子式都等于零, 则定义 r 为矩阵的秩 (rank), 记作 $\operatorname{rank}(\boldsymbol{A})=r$ 或 $r(\boldsymbol{A})=r$. 若 $\boldsymbol{A}$ 为零矩阵, 则定义 $\boldsymbol{A}$ 的秩为零, 记作 $\operatorname{rank}(\boldsymbol{A})=0$ 或 $r(\boldsymbol{A})=0$.

秩这个东西, 放在 n 元一次方程组对应矩阵中, 其实就是说有几个方程是真正有用的. 所以, 秩和方程组解息息相关.

我们发现, 秩就是矩阵向单位矩阵化简结果中 1 的数量. 所以现在我们就这么求秩.

为了更方便解方程组, 我们把常数并入矩阵中一起行变换, 于是有:

定义 13.

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

为矩阵

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

的增广矩阵.

所有的线性方程组, 都可以如此考虑成这样的矩阵.

设线性方程组 $\boldsymbol{Ax} = \boldsymbol{\beta}$, 其中 $\boldsymbol{A}$ 为 $m \times n$ 矩阵, 若 $\boldsymbol{\beta} = 0$, 则

(1) $\boldsymbol{Ax} = \boldsymbol{\beta}$ 只有零解的充分必要条件为 $r(\boldsymbol{A}) = n$, 即 $|\boldsymbol{A}| \neq 0$;

(2) $\boldsymbol{Ax} = \boldsymbol{\beta}$ 有非零解的充分必要条件为 $r(\boldsymbol{A}) < n$, 即 $|\boldsymbol{A}| = 0$.

这其实就是 Crimer 法制. 所以我前面没写, 欸嘿.

若 $\boldsymbol{\beta} \neq 0$, 则

(1) $\boldsymbol{Ax} = \boldsymbol{\beta}$ 有唯一解的充分必要条件为 $r(\boldsymbol{A}) = r(\widetilde{\boldsymbol{A}}) = n$;

(2) $\boldsymbol{Ax} = \boldsymbol{\beta}$ 有无穷多解的充分必要条件为 $r(\boldsymbol{A}) = r(\widetilde{\boldsymbol{A}}) < n$.

2.2.1 解方程组

先将方程组对应增广矩阵通过初等行变换化简成为简化阶梯矩阵. 然后我们得到了一些形式比较简单的方程组.

如果方程组无解或者有唯一解, 很简单就能解出. 但是当方程组有无数解时, 我们也能很简单地给出通解:

首先, 用未知数个数即列数减去秩, 这个差值就是自由变量的个数, 如果忽略自由变量, 我们可以解出其他变量的一组确定解. 再此基础上, 每一个自由变量的值产生的影响可以由前面解出的几个变量消去. 每个自由变量都需要消去, 于是方程解系含有自由变量个数 $+1$ 个向量.

当然, 我讨论的是非齐次线性方程组. 对于齐次线性方程组, 我举得没什么要记的.

2.2.2 向量组分块矩阵变换

向量组这玩意, 我目前没觉得有什么太独特的东西. 无非就是一个向量构成的矩阵, 再在这个基础上考虑. 以平面向量的概念推广, 就目前而言理解毫无难度.

分块矩阵变换也没有什么特别之处. 就是和普通矩阵一样的. 我写在这只是尊重下课本了, 毕竟人家花了大篇幅认认真真地证明.

2.2.3 一些定理和结论

(1)$r(\boldsymbol{AB}) \leq \min(r(\boldsymbol{A}), r(\boldsymbol{B}))$

(2) 矩阵经历初等变换, 其逆矩阵分别经历:

$\boldsymbol{E}(i,j)]$ $\quad$ $\boldsymbol{E}(i(\frac{1}{k}))$ $\quad$ $\boldsymbol{E}(j,i(-k))$

(3)

$$\begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_s \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}_1^{-1} & & & \\ & \boldsymbol{A}_2^{-1} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_s^{-1} \end{pmatrix},$$

$$\begin{pmatrix} & & & \boldsymbol{A}_1 \\ & & \boldsymbol{A}_2 & \\ & \iddots & & \\ \boldsymbol{A}_s & & & \end{pmatrix}^{-1} = \begin{pmatrix} & & & \boldsymbol{A}_s^{-1} \\ & & \boldsymbol{A}_{s-1}^{-1} & \\ & \iddots & & \\ \boldsymbol{A}_1^{-1} & & & \end{pmatrix}.$$

(4) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 阶方阵, 且 $\boldsymbol{E}-\boldsymbol{AB}$ 可逆. 则 $\boldsymbol{E}-\boldsymbol{BA}$ 也可逆.

一般地, 若 $\boldsymbol{A}, \boldsymbol{B}$ 为 n 阶方阵, 数 $\lambda \neq 0$. 则 $|\lambda\boldsymbol{E}-\boldsymbol{AB}| = |\lambda\boldsymbol{E}-\boldsymbol{BA}|$.

(5) 若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 与向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 的秩相等, 且 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$ 可由向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性表示, 则这两个向量组等价.

(6) 若向量组 $\boldsymbol{B}: \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_r$ 能由向量组 $\boldsymbol{A}: \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s$ 线性表示为 $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_r) = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{K}$, 其中 $\boldsymbol{K}$ 为 $s \times r$ 矩阵, 且向量组 $\boldsymbol{A}$ 线性无关. 则向量组 $\boldsymbol{B}$ 线性无关的充分必要条件是矩阵 $\boldsymbol{K}$ 的秩 $r(\boldsymbol{K}) = r$.

(7) 若 $m \geq 2$, 已知 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 为 n 维向量组, 且向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示为

$$\begin{aligned} \boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3 + \cdots + \boldsymbol{\alpha}_m \\ \boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3 + \cdots + \boldsymbol{\alpha}_m \\ &\cdots\ldots\cdots \\ \boldsymbol{\beta}_m &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_{m-1} \end{aligned}$$

则 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 线性无关的充分必要条件是 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关.

(8) 若向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_k$ 的秩为 r, 则其中任意选取 m 个向量所构成的向量组的秩 $\geq r+m-k$.

(9) 若 $\boldsymbol{A},\boldsymbol{B}$ 均为 $m\times n$ 矩阵, $\boldsymbol{C}=(\boldsymbol{A},\boldsymbol{B})$ 为 $m\times 2n$ 矩阵. 则:

$$\max(r(\boldsymbol{A}),r(\boldsymbol{B}))\leq r(\boldsymbol{C})\leq r(\boldsymbol{A})+r(\boldsymbol{B})$$

(10)$r(\boldsymbol{A})=1$ 的充分必要条件是存在非零列向量 $\boldsymbol{\alpha}$ 以及非零行向量 $\boldsymbol{\beta}^{\mathrm{T}}$ 使得 $\boldsymbol{A}=\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$.

(11) 若 $\boldsymbol{A}^{*}$ 为 n 阶方阵 $\boldsymbol{A}$ 的伴随矩阵, $n\geq 2$. 则:

$$r\left(\boldsymbol{A}^{*}\right)=\begin{cases}n, & \text{若} r(\boldsymbol{A})=n,\\ 1, & \text{若} r(\boldsymbol{A})=n-1\\ 0, & \text{若} r(\boldsymbol{A})<n-1\end{cases}$$

(12) 若 A 是 $m\times n$ 实矩阵, 则:

1. $\boldsymbol{A}\boldsymbol{X}=\mathbf{0}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\mathbf{0}$ 是同解方程组;
2. $r(\boldsymbol{A})=r\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)$

(13) 若 A 为 n 阶方阵, 且 $A^2=A$. 则:

$$r(\boldsymbol{A})+r(\boldsymbol{A}-\boldsymbol{E})=n$$

(14) 若 A 为 $m\times n$ 矩阵, B 为 $m\times 1$ 矩阵, 则方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$ 有解的充分必要条件是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\mathbf{0}$ 的任一解向量 $\boldsymbol{Y}_0$ 都是 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}=\mathbf{0}$ 的解向量.,b

2.2.4 课上的一些结论

(1)$r(\boldsymbol{A})=r(-\boldsymbol{A})$. $r(\boldsymbol{A}+\boldsymbol{B})\leq r(\boldsymbol{A})+r(\boldsymbol{B})$

(2) 设A 为$m\times n$ 阵, B 为$n\times p$ 阵, $AB=O$, 则: $r(A)+r(B)\leq n$.

(3) 关于向量组的秩:

1. 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关 $\Leftrightarrow r\left(\alpha_1,\alpha_2,\cdots,\alpha_m\right)=m$
2. 若 $r\left(\alpha_1,\alpha_2,\cdots,\alpha_m\right)=r>0$, 则向量组中任意 $k>r$ 个向量都是线性相关的.
3. 若 $r\left(\alpha_1,\alpha_2,\cdots,\alpha_m\right)=r>0$, 则向量组任意 r 个线性无关的向量都是它的一个极大线性无关组
4. 若向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 可由向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 线性表出, 则 $r\left(\alpha_1,\alpha_2,\cdots,\alpha_m\right)\leq r\left(\beta_1,\beta_2,\cdots,\beta_s\right)$.
5. 等价的向量组有相同的秩.