1.3 初等变换与初等矩阵

本节以观看视频为主，对下面几个方面再做些强调和补充。

1.3.1 初等变换

1. 初等变换在线性代数中非常重要，希望同学们重视。线性代数中很多问题的计算都是先用初等变换对矩阵进行化简，然后通过化简以后的矩阵来计算原来的问题。

2. 矩阵的初等变换起源于解线性方程组的消元法,将消元法所用的变换加以推广给出了矩阵初等行变换的概念，将行变换再做进一步的推广给出了矩阵初等列变换的概念。

3. 注：（1）用 r 表示行(row)，用 c 表示列(column).

（2）$r_j + r_i$ 与 $r_i + r_j$ 的含义不同，$r_j + r_i$ 表示将矩阵的第 i 行加到第 j 行，

$r_i + r_j$ 表示将矩阵的第 j 行加到第 i 行.

（3）形如 $r_i \times \frac{1}{2}$ 和 $r_j + (-2)r_i$ 的记号也可写成 $r_i \div 2$ 和 $r_j - 2r_i$ 的形式.

4. 注意：当 **A** 经过初等变换变成 **B** 时,**A** 与 **B** 一般是不相等的,所以不能写成 **A=B**

5.当 $n \times n$ 型线性方程组 $\mathbf{Ax} = \mathbf{b}$ 有唯一解时,其解法是:用**初等行变换**将增广矩阵 $[\mathbf{A}, \mathbf{b}]$ 化为 $[\mathbf{E}, \mathbf{c}]$,则其解为 $\mathbf{x} = \mathbf{c}$.

注：（1）解方程组**只允许对矩阵做初等行变换，不允许对矩阵做初等列变换**。

（2）虽然解方程组时不允许对矩阵做初等列变换，但是初等列变换可用于处理矩阵的其他问题，所以也要学习。

（3）线性方程组什么时候有唯一解？这个问题将在后面给出答案。

6. 在后面我们经常会遇到**将一个方阵化成对角矩阵**的问题，下面以三阶方阵为例讲述用**初等行变换将其化为对角矩阵**的做法。

第一步：化简第一列，重点考虑 a_{11} . 若 $a_{11} = 0$,但 a_{21}, a_{31} 不全为 0，则需先**通过对调变换或者倍加变换做个调整**，然后再用下面行减去第一行的倍数，将第一列下方全化为 0. 记化完第一列以后的矩阵为 $\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{bmatrix}$.

例如：$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{r_2 - 2r_1} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$

$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{r_1 + r_3} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow[r_3 - r_1]{r_2 - 2r_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -3 \\ 0 & -1 & -1 \end{bmatrix}$

（注：对调变换比倍加变换简单）

注意，若 $a_{11} \neq 0$, 但 a_{21}, a_{31} 不是 a_{11} 的整数倍，为了计算简单一般也要先通过对调变换或者倍加变换做个调整，然后再用下面行减去第一行的倍数，将第一列下方全化为 0.

例如：$\begin{bmatrix} 3 & 4 & 6 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 3 \\ 3 & 4 & 6 \end{bmatrix} \xrightarrow[r_3-3r_1]{r_2-2r_1} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$

$$\begin{bmatrix} 3 & 4 & 6 \\ 2 & 2 & 3 \\ 4 & 7 & 2 \end{bmatrix} \xrightarrow{r_1 - r_2} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 4 & 7 & 2 \end{bmatrix} \xrightarrow[r_3-5r_1]{r_2-2r_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -3 \\ 0 & -1 & -10 \end{bmatrix}$$

第二步:化简第二列，重点考虑 b_{22}，若 $b_{22} = 0$, 或者 b_{12}, b_{22}, b_{32} 之间的倍数不好，则需先**通过对调变换或者倍加变换做个调整（注意：这时只能让第二行与第三行做调整,不能让第二行与第一行做调整）**，调整好以后，分别用第一行和第三行减去第二行的倍数，将矩阵化成 $\begin{bmatrix} c_{11} & 0 & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix}$

第三步：化简第三列。如果能化成对角矩阵，要么 c_{13}, c_{23}, c_{33} 全为 0，要么 $c_{33} \neq 0$。这时，只需用第一行和第二行分别减去第三行的倍数，就可化成对角矩阵 $\begin{bmatrix} c_{11} & 0 & 0 \\ 0 & c_{22} & 0 \\ 0 & 0 & c_{33} \end{bmatrix}$。

注：（1）若 c_{11}, c_{22}, c_{33} 都不为 0，则进一步做行变换能化成单位矩阵。这时，一般还需做倍乘行变换。

（2）如果想用行变换把一个方阵化成上三角矩阵，则只需将对角线下方化成 0 就可以，想法和上面讲的差不多。

1.3.2 初等矩阵

1. **定义** 由**单位矩阵 E** 经过**一次初等变换**所得到的矩阵叫做**初等矩阵**.与三种初等变换相对应有以下三种初等矩阵：

(1)对调 $\mathbf{E}$ 的 i 和 j 两行(列)所得到的方阵叫做对调矩阵,记作 $\mathbf{E}_{i,j}$.

(2)将 $\mathbf{E}$ 的第 i 行(列)乘以非零数 k 所得到的方阵叫做倍乘矩阵,记作 $\mathbf{E}_i(k)$.

(3)将 $\mathbf{E}$ 的第 j 行乘以数 k 加到第 i 行（或将 $\mathbf{E}$ 的第 i 列乘以数 k 加到第 j 列）所得到的方阵叫做倍加阵,记作 $\mathbf{E}_{i,j}(k)$.　　**（注意：倍加矩阵对于行和列表达的含义不一样）**

2. 初等矩阵的特点：

$$\mathbf{E}_{i,j}=\begin{bmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{bmatrix},\qquad \mathbf{E}_i(k)=\begin{bmatrix}1 & & & & \\ & \ddots & & & \\ & & k & & \\ & & & \ddots & \\ & & & & 1\end{bmatrix},$$

$$\mathbf{E}_{i,j}(k)=\begin{bmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & \cdots & k & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{bmatrix},\quad i<j$$

$$\mathbf{E}_{i,j}(k)=\begin{bmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & k & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{bmatrix},\quad i>j$$

3. 初等矩阵具有下列性质：

性质 1-1 $\mathbf{E}_{i,j}^T=\mathbf{E}_{i,j}$, $\quad \mathbf{E}_i^T(k)=\mathbf{E}_i(k)$, $\quad \mathbf{E}_{i,j}^T(k)=\mathbf{E}_{j,i}(k)$.

性质 1-2 （这个性质非常重要）

（1）$\mathbf{A}\xrightarrow{r_i\leftrightarrow r_j}\mathbf{B}$ 等同于 $\mathbf{E}_{i,j}\mathbf{A}=\mathbf{B}$，$\mathbf{A}\xrightarrow{c_i\leftrightarrow c_j}\mathbf{B}$ 等同于 $\mathbf{A}\mathbf{E}_{i,j}=\mathbf{B}$；

（2）$\mathbf{A}\xrightarrow{r_i\times k}\mathbf{B}$ 等同于 $\mathbf{E}_i(k)\mathbf{A}=\mathbf{B}$，$\mathbf{A}\xrightarrow{c_i\times k}\mathbf{B}$ 等同于 $\mathbf{A}\mathbf{E}_i(k)=\mathbf{B}$；

（3）$\mathbf{A}\xrightarrow{r_j+kr_i}\mathbf{B}$ 等同于 $\mathbf{E}_{j,i}(k)\mathbf{A}=\mathbf{B}$，

$\mathbf{A}\xrightarrow{c_j+kc_i}\mathbf{B}$ 等同于 $\mathbf{A}\mathbf{E}_{i,j}(k)=\mathbf{B}$.

注意 (1)“行变换”对应于“左乘初等矩阵”，“列变换”对应于“右乘初等矩阵”.

(2) 对矩阵 A 做有限次初等行(列)变换相当于用有限个相应的初等矩阵左乘(右乘) A.

性质 1-3 $\mathbf{E}_{i,j}\mathbf{E}_{i,j}=\mathbf{E}$, $\quad \mathbf{E}_i(k)\mathbf{E}_i(k^{-1})=\mathbf{E}$, $\quad \mathbf{E}_{i,j}(k)\mathbf{E}_{i,j}(-k)=\mathbf{E}$

证明 (只证明第一个式子.)

设 $\mathbf{E}=\left[\mathbf{e}_1,\cdots,\mathbf{e}_i,\cdots,\mathbf{e}_j,\cdots,\mathbf{e}_n\right]$.

证法 1 根据性质 1-2，由 $\mathbf{E}_{i,j}=\left[\mathbf{e}_1,\cdots,\mathbf{e}_j,\cdots,\mathbf{e}_i,\cdots,\mathbf{e}_n\right]\xrightarrow{c_i\leftrightarrow c_j}\mathbf{E}$，

得 $\mathbf{E}_{i,j}\mathbf{E}_{i,j}=\mathbf{E}$.

证法 2　可转化为证明 $\mathbf{E}\mathbf{E}_{i,j}\mathbf{E}_{i,j}=\mathbf{E}$.

根据性质 1-2, $\mathbf{E}\mathbf{E}_{i,j}\mathbf{E}_{i,j}$ 的含义是对 $\mathbf{E}$ 连着做两次对调第 i 列和第 j 列的对调列变换，显然,得到的矩阵还是 $\mathbf{E}$，所以结论正确.

例　设 $\mathbf{A}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$，计算 $\mathbf{E}_{2,1}(-2)\mathbf{A}$.

解　根据性质 1-2, 用 $\mathbf{E}_{2,1}(-2)$ 左乘 $\mathbf{A}$ 所得矩阵和对 $\mathbf{A}$ 做相应的倍加行变换 $r_2+(-2)r_1$

所得矩阵是相等的.于是,可通过初等行变换求出 $\mathbf{E}_{2,1}(-2)\mathbf{A}$.

$$\mathbf{E}_{2,1}(-2)\mathbf{A}=\begin{bmatrix}1&2&3\\2&1&0\\7&8&9\end{bmatrix}.$$

1.3.1　矩阵的等价标准形

定理 1-1　对于任何方阵 **A,** 只用有限次倍加行变换(或有限次倍加列变换)都能将 **A** 化为上三角形矩阵,即一定有倍加矩阵 $\mathbf{P}_i(i=1,2,\cdots,k)$ [或 $\mathbf{Q}_j(j=1,2,\cdots,l)$]，使得 $\mathbf{P}_m\cdots\mathbf{P}_2\mathbf{P}_1\mathbf{A}$ （或 $\mathbf{A}\mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_l$ ）为上三角形矩阵。

定理 1-2　对于任何 $m\times n$ 非零矩阵 **A,** 必能用初等变换把它化为形如 $\mathbf{F}=\begin{bmatrix}\mathbf{E}_s&\mathbf{O}\\\mathbf{O}&\mathbf{O}\end{bmatrix}$ 的矩阵,即存在 m 阶初等矩阵 $\mathbf{P}_1,\mathbf{P}_2,\cdots,\mathbf{P}_k$ 和 n 阶初等矩阵 $\mathbf{Q}_1,\mathbf{Q}_2,\cdots,\mathbf{Q}_l$,使得

$$\mathbf{P}_k\cdots\mathbf{P}_2\mathbf{P}_1\mathbf{A}\ \mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_l=\mathbf{F}.$$

$\mathbf{F}=\begin{bmatrix}\mathbf{E}_s&\mathbf{O}\\\mathbf{O}&\mathbf{O}\end{bmatrix}$ 叫做矩阵 **A** 的**等价(相抵)标准形**.

注意　**F** 包括 $[\mathbf{E}_m\ ,\ \mathbf{O}]$、$\begin{bmatrix}\mathbf{E}_n\\\mathbf{O}\end{bmatrix}$、**E** 三种特殊情况,它们分别对应于 s=m<n,s=n<m, s=m=n.

(在 5.2 节将会看到 s 就是矩阵 **A** 的秩,它是由 **A** 唯一确定的,因此一个矩阵的等价标准形是唯一的.)