MIT-18.06：线性代数

课程视频 和 配套教材（提取码：whs9）在这里！

如果我大二的时候学的是这套线代教程，我的人生轨迹就会发生变化：可能绩点会有大的改观，可能拿到保研资格，可能不会因为考研不考数学而跟医工院深度绑定，可能离开天津前往上海，可能硕士毕业就参加工作，可能不会在一个不喜欢的行业继续读博苟延残喘，之后的人生理应处处发生改变吧。

这份文档并非是面面俱到的笔记，而是我在学习相关课程中发现的重要结论与新的收获。

若无特殊说明，本文中大写字母表示二维矩阵，小写字母表示行向量。

---

1. 方程组的几何解释

1.1 矩阵的“列视图”

对于多元一次线性方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{1-1}$$

式 (1-1) 可以改写为矩阵 & 向量的乘法形式：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{1-2}$$

$$\boldsymbol{A}\boldsymbol{x}^T = \boldsymbol{b}^T \tag{1-3}$$

在面对式 (1-2) 这种类型的矩阵乘法时，我们通常习惯以“**元素的加权和**”这一视角来描述一种“数群”至“数字”的运算过程，即存在这样的思维习惯：$b_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3$。此时我们接受系数矩阵 $\boldsymbol{A}$ 的方式是逐行进行的，是一种略显复杂的动态过程。

列视图是一种我以前未曾发现的全新视角，其本质为矩阵列的线性组合，基于列视图，我们可以把式 (1-2) 改写为：

$$
x_1\begin{bmatrix}a_{11}\\a_{21}\\a_{31}\end{bmatrix}+x_2\begin{bmatrix}a_{12}\\a_{22}\\a_{32}\end{bmatrix}+x_3\begin{bmatrix}a_{13}\\a_{23}\\a_{33}\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix} \tag{1-4}
$$

显然，在面对形如“矩阵右乘向量”之类的运算时，使用**列视图**能够更好地简化思路，在实际应用中帮助我们明确信息流传递的真实意义。

---

2. 矩阵消元

2.1 矩阵的“行视图”

对于向量 & 矩阵的乘法 $\boldsymbol{aX}=\boldsymbol{b}$:

$$
\begin{bmatrix}a_1 & a_2 & a_3\end{bmatrix}\begin{bmatrix}x_{11} & x_{12} & x_{13}\\x_{21} & x_{22} & x_{23}\\x_{31} & x_{32} & x_{33}\end{bmatrix}=\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix} \tag{2-1}
$$

与式 (1-4) 不同，这是“矩阵左乘向量”的运算形式，因此可以使用**行视图**对其进行观察与分析：

$$
a_1\begin{bmatrix}x_{11} & x_{12} & x_{13}\end{bmatrix}+a_2\begin{bmatrix}x_{21} & x_{22} & x_{23}\end{bmatrix}+a_3\begin{bmatrix}x_{31} & x_{32} & x_{33}\end{bmatrix}=\begin{bmatrix}b_1 & b_2 & b_3\end{bmatrix} \tag{2-2}
$$

掌握矩阵的**行视图**与**列视图**有助于我们进一步理解矩阵初等变换以及逆矩阵的本质。

2.2 矩阵乘法的新视角

由 **1.1** 以及 **2.1** 的内容可知，矩阵之间的乘法可视为**左矩阵**与多个**列向量**的乘法，抑或为多个**行向量**与**右矩阵**的乘法。例如 $\boldsymbol{AB}=\boldsymbol{C}$ 除了最基础、最抽象的**运算型** (*regular way*) 表示：

> The regular way: a row of $\boldsymbol{A}$ times a column of $\boldsymbol{B}$ to get a number of $\boldsymbol{C}$

$$
\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\\a_{31} & a_{32}\end{bmatrix}\begin{bmatrix}b_{11} & b_{12} & b_{13}\\b_{21} & b_{22} & b_{23}\end{bmatrix}=\begin{bmatrix}\sum_i a_{1i}b_{i1} & \sum_i a_{1i}b_{i2} & \sum_i a_{1i}b_{i3}\\\sum_i a_{2i}b_{i1} & \sum_i a_{2i}b_{i2} & \sum_i a_{2i}b_{i3}\\\sum_i a_{3i}b_{i1} & \sum_i a_{3i}b_{i2} & \sum_i a_{3i}b_{i3}\end{bmatrix} \tag{2-3}
$$

还可以表示为**列型** (*column way*) 和**行型** (*row way*) :

> The column way: matrix A times a column of B to get a column of C
>
> The row way: a row of A times matrix B to get a row of C

$$\boldsymbol{AB}=\boldsymbol{A}\begin{bmatrix}b_{11}\\b_{21}\end{bmatrix}\bigoplus\boldsymbol{A}\begin{bmatrix}b_{21}\\b_{22}\end{bmatrix}\bigoplus\boldsymbol{A}\begin{bmatrix}b_{31}\\b_{32}\end{bmatrix}\tag{2-4}$$

$$\boldsymbol{AB}=\begin{bmatrix}a_{11}&a_{12}\end{bmatrix}\boldsymbol{B}\bigoplus\begin{bmatrix}a_{21}&a_{22}\end{bmatrix}\boldsymbol{B}\bigoplus\begin{bmatrix}a_{31}&a_{32}\end{bmatrix}\boldsymbol{B}\tag{2-5}$$

此时 $\bigoplus$ 视情况分别表示**列向量在横向**或**行向量在纵向**的拼接。除了上述三种运算角度，其实还有**行视图**、**列视图**混用的第四种形式：

> The 4th way: a column of A times a row of B to get a partial matrix of C

$$\boldsymbol{AB}=\begin{bmatrix}a_{11}\\a_{21}\\a_{31}\end{bmatrix}\begin{bmatrix}b_{11}&b_{12}&b_{13}\end{bmatrix}+\begin{bmatrix}a_{12}\\a_{22}\\a_{32}\end{bmatrix}\begin{bmatrix}b_{21}&b_{22}&b_{23}\end{bmatrix}=\sum_i\boldsymbol{A}(:,i)\boldsymbol{B}(i,:)\tag{2-6}$$

这种混合视图隐藏的信息非常重要：

（1） $\boldsymbol{AB}$ 结果中的**每一列**，都是左矩阵 $\boldsymbol{A}$ 中**各列**的**线性组合**

（2） $\boldsymbol{AB}$ 结果中的**每一行**，都是右矩阵 $\boldsymbol{B}$ 中**各行**的**线性组合**

上述结论通常与行（列）空间结合，帮助我们判断矩阵是否可逆（满秩）。

2.3 矩阵初等变换的意义

对矩阵 $\boldsymbol{X}$ 进行单次初等行变换，相当于 $\boldsymbol{X}$ 左乘一个操作矩阵 $\boldsymbol{H}$，例如矩阵**第一行不变**、**第二行减去第一行的两倍**：

$$\begin{bmatrix}x_{11}&x_{12}\\x_{21}-2x_{11}&x_{22}-2x_{12}\end{bmatrix}\Leftrightarrow\begin{bmatrix}1&0\\-2&1\end{bmatrix}\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}\tag{2-7}$$

上述过程太抽象？回过头去看看式 (2-1) 和 (2-5)，想想矩阵的**行视图**、矩阵乘法的**行型**：

$$\begin{bmatrix}1&0\\-2&1\end{bmatrix}\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}=\begin{bmatrix}1&0\end{bmatrix}\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}+\begin{bmatrix}-2&1\end{bmatrix}\begin{bmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{bmatrix}\tag{2-8}$$

$$\begin{cases} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = 1 \times \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} + 0 \times \begin{bmatrix} x_{21} & x_{22} \end{bmatrix}, \ (row\ 1) \\ \\ \begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = -2 \times \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} + 1 \times \begin{bmatrix} x_{21} & x_{22} \end{bmatrix}, \ (row\ 2) \end{cases} \tag{2-9}$$

对应地，单次初等列变换等价于 $\boldsymbol{X}$ 右乘一个操作矩阵 $\boldsymbol{V}$，例如矩阵**第一列除以2、第二列乘3以后加上第一列变换后的5倍**：

$$\begin{bmatrix} \dfrac{x_{11}}{2} & 3x_{12} + \dfrac{5x_{11}}{2} \\ \\ \dfrac{x_{21}}{2} & 3x_{22} + \dfrac{5x_{21}}{2} \end{bmatrix} \Leftrightarrow \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} \dfrac{1}{2} & \dfrac{5}{2} \\ \\ 0 & 3 \end{bmatrix} \tag{2-10}$$

朝闻道，夕死可矣。

3 逆矩阵

3.1 判断矩阵是否可逆

书接上回，例如病态矩阵 $\boldsymbol{A}$：

$$\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} \tag{2-7}$$

要判断 $\boldsymbol{A}$ 是否可逆，即寻找一个矩阵 $\boldsymbol{A}^{-1}$ 使得 $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$。根据上一节末尾的结论，$\boldsymbol{A}\boldsymbol{A}^{-1}$ 的每一列都是 $\boldsymbol{A}$ 中各列的线性组合。但是 $\boldsymbol{A}$ 的各列是**共线**的，而 $[1,0]^T$ 所在直线并不在方向向量为 $[1,2]^T$ 的线簇中，因此在本例中我们永远无法组合出单位向量，换句话说 $\boldsymbol{A}$ 不可逆。

更一般地，如果我们能找到一个非零列向量 $\boldsymbol{x} \neq \boldsymbol{0}$ 使得以下方程成立，则 $\boldsymbol{A}$ 不可逆：

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \tag{2-8}$$

这一结论可以从两方面理解：

（1）$\boldsymbol{A}$ 中存在某一列（行）可以由其它列（行）的线性组合表示，即该列（行）对于填充矩阵必要信息没有起到任何作用，因此这个矩阵就是“病态”的，是不完整的，也就不可逆

（2）若 $\boldsymbol{A}$ 可逆，则有 $\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$，从而 $\boldsymbol{x} = \boldsymbol{0}$，这与非零列向量的前提是矛盾的，所以 $\boldsymbol{A}$ 不可逆

3.2 求解方阵的逆矩阵

简单起见，我们先具体到一个非奇异的二阶方阵 $\boldsymbol{B}$：

$$\boldsymbol{B}=\begin{bmatrix}1&3\\2&7\end{bmatrix} \tag{2-9}$$

关于如何求解逆矩阵 $\boldsymbol{B}^{-1}$，有两种主要思路，**第一种**是 *Gauss's way*，即依次求解二元一次线性方程组：

$$\boldsymbol{B}\boldsymbol{x}_1=\begin{bmatrix}1\\0\end{bmatrix},\ \boldsymbol{B}\boldsymbol{x}_2=\begin{bmatrix}0\\1\end{bmatrix} \tag{2-10}$$

$$\boldsymbol{B}^{-1}=\begin{bmatrix}\boldsymbol{x}_1&\boldsymbol{x}_2\end{bmatrix} \tag{2-11}$$

需要注意的是，*Gauss* 通过增广矩阵 (*augmented matrix*) 与初等行变换来实现方程组求解，以 $\boldsymbol{B}\boldsymbol{x}_1$ 为例：

$$\hat{\boldsymbol{B}}=\begin{bmatrix}1&3|&1\\2&7|&0\end{bmatrix}\to\begin{bmatrix}1&3|&1\\0&1|&-3\end{bmatrix}\to\begin{bmatrix}1&0|&7\\0&1|&-2\end{bmatrix} \tag{2-12}$$

显然我们已经得到了 $\boldsymbol{x}_1$ ，以此类推可解出 $\boldsymbol{x}_2$。

第二种是 *Gauss-Jordan's way*，即一次性求解多个多元一次线性方程组，具体实施方法是在 *Gauss* 的基础上进一步拓展增广矩阵 $\hat{\boldsymbol{B}}$，并通过彻底的初等行变换将增广矩阵的左半部分化简为单位阵，则增广部分即为逆矩阵：

$$\hat{\boldsymbol{B}}=\begin{bmatrix}1&3|&1&0\\2&7|&0&1\end{bmatrix}\to\begin{bmatrix}1&3|&1&0\\0&1|&-2&1\end{bmatrix}\to\begin{bmatrix}1&0|&7&-3\\0&1|&-2&1\end{bmatrix} \tag{2-13}$$

$$\boldsymbol{B}^{-1}=\begin{bmatrix}7&-3\\-2&1\end{bmatrix} \tag{2-14}$$
