

MIT 18.06 线性代数

第三讲：矩阵乘法与逆矩阵

整理自 MIT 18.06 课程及相关公开笔记

引言

在线性代数中，矩阵乘法是连接线性变换、向量空间与方程组的核心运算。本讲将从四个不同的视角来理解矩阵乘法 $AB = C$ ，这不仅有助于加深对矩阵本质的理解，也为后续学习矩阵分解打下基础。此外，本讲还将重点讨论方阵逆矩阵的性质，以及一种系统化求逆矩阵的方法——高斯-若尔当消元法。

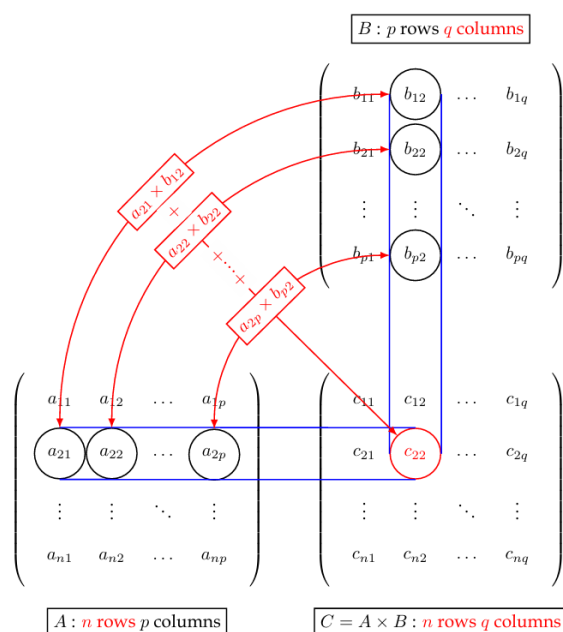
1 矩阵乘法的四种视角

设 A 为 $m \times n$ 的矩阵， B 为 $n \times p$ 的矩阵，则它们的乘积 $C = AB$ 为一个 $m \times p$ 的矩阵。

1.1 标准方法（行乘以列）

这是矩阵乘法最常规的定义方式。矩阵 C 中第 i 行第 j 列的元素 c_{ij} ，等于矩阵 A 的第 i 行行向量与矩阵 B 的第 j 列列向量的点积（内积）：

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$



1.2 列组合视角（列视角）

该视角将矩阵 B 视为 p 个列向量的组合，即 $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_p]$ 。由此，矩阵乘法 AB 可以写为：

$$AB = A [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_p] = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \cdots \ A\mathbf{b}_p]$$

其中， $A\mathbf{b}_j$ 是矩阵 A 的所有列向量的线性组合，组合系数正是列向量 $\mathbf{b}_j$ 的各个分量。**重要推论：**矩阵 C 的每一列，都是矩阵 A 的所有列向量的线性组合。这个视角清晰地解释了“用矩阵 B 的列去线性组合矩阵 A 的列”。

1.3 行组合视角（行视角）

该视角将矩阵 A 视为 m 个行向量的组合：

$$AB = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{a}_1 B \\ \mathbf{a}_2 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix}$$

其中， $\mathbf{a}_i B$ 是矩阵 B 的所有行向量的线性组合，组合系数正是行向量 $\mathbf{a}_i$ 的各个分量。**重要推论：**矩阵 C 的每一行，都是矩阵 B 的所有行向量的线性组合。这个视角清晰地说明了“用矩阵 A 的行去线性组合矩阵 B 的行”。

1.4 列乘以行（外积求和）

任何矩阵乘法都可以表示为一组“列向量乘以行向量”的矩阵之和。设 A 的第 k 列是一个 $m \times 1$ 的列向量， B 的第 k 行是一个 $1 \times p$ 的行向量，两者的乘积将得到一个 $m \times p$ 的矩阵：

$$AB = \sum_{k=1}^n \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} \begin{bmatrix} b_{k1} & b_{k2} & \cdots & b_{kp} \end{bmatrix}$$

这是一个非常重要的视角：它表明矩阵乘法可以分解为 n 个秩为 1 的矩阵之和。此类“列乘行”产生的矩阵，其所有列向量共线，所有行向量也共线。

1.5 分块乘法

当矩阵的维度较大时，我们可以将它们划分为尺寸匹配的分块矩阵，然后按块进行乘法运算。只要分块满足矩阵乘法的维数要求，分块乘法的规则与常规乘法完全一致。

例如，若将 A 和 B 分块为：

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

则它们的分块乘积为：

$$AB = \begin{bmatrix} A_1 B_1 + A_2 B_3 & A_1 B_2 + A_2 B_4 \\ A_3 B_1 + A_4 B_3 & A_3 B_2 + A_4 B_4 \end{bmatrix}$$

类似地，可以通过 $C_1 = A_1 B_1 + A_2 B_3$ 分块计算出矩阵块的元素。这一技巧是计算机实现高效矩阵运算（如 Strassen 算法）的重要基础。

2 逆矩阵

对于方阵 A ，如果存在一个矩阵 A^{-1} ，使得 $AA^{-1} = A^{-1}A = I$ (I 为单位矩阵)，则称 A 为可逆矩阵（或非奇异矩阵）， A^{-1} 为其逆矩阵。

2.1 奇异矩阵（不可逆矩阵）

如果矩阵 A 不可逆，则称之为奇异矩阵。奇异矩阵有一个关键性质：存在非零向量 $\mathbf{x}$ ，使得 $A\mathbf{x} = \mathbf{0}$ 。

理由：如果 A 可逆，那么 $A\mathbf{x} = \mathbf{0}$ 两边左乘 A^{-1} 得 $\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$ ，这与 $\mathbf{x}$ 是非零向量矛盾。

几何意义：矩阵 A 的列向量线性相关（例如成倍数关系），无法张满整个空间。

举例： $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ 的两列成比例，为奇异矩阵（行列式为零）。对于非零向量 $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ ，有：

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

2.2 高斯-若尔当方法求逆矩阵

求逆矩阵本质上是解方程组 $A\mathbf{x} = \mathbf{b}$ 的特殊形式：对应右端项为单位阵。

解法步骤：

1. 将单位矩阵放在 A 的右侧，建立增广矩阵 $[A | I]$ ；
2. 对增广矩阵进行高斯消元与回代，将左侧的 A 化为单位矩阵 I ；
3. 左侧变为 I 时，右侧得到的矩阵即为逆矩阵 A^{-1} 。

$$[A | I] \xrightarrow{\text{行变换}} [I | A^{-1}]$$

原理证明：我们使用消元矩阵 E 来表示对矩阵 $[A | I]$ 进行的行变换。变换后的结果是：

$$E[A | I] = [EA | EI]$$

由于最终左侧变成了单位矩阵，即 $EA = I$ ，故根据逆矩阵的定义，可知 $E = A^{-1}$ 。同时右侧 $EI = E = A^{-1}$ ，与预期一致。

示例：求 $A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ 的逆矩阵。

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \xrightarrow{r_2 - 2r_1} \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \xrightarrow{r_1 - 3r_2} \begin{bmatrix} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

因此， $A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}$ 。

2.3 逆矩阵的运算性质

与实数乘法类似，矩阵乘法和转置的逆运算遵循以下规则：

- 乘积的逆： $(AB)^{-1} = B^{-1}A^{-1}$ （注意运算顺序反转）
- 转置的逆： $(A^T)^{-1} = (A^{-1})^T$

总结

- 矩阵乘法有四种理解视角：逐元素点积（行乘以列）、列向量的线性组合、行向量的线性组合，以及列向量乘行向量外积求和。后三种视角强调了“线性组合”的核心思想，是贯穿整个课程的主线。
- 分块乘法是处理大型矩阵的重要技巧，对于实现快速矩阵算法（如 Strassen 算法）至关重要。
- 方阵可逆的充要条件是方程 $A\mathbf{x} = \mathbf{0}$ 仅存在零解（列向量线性无关）。
- 高斯-若尔当法通过构建增广矩阵 $[A \mid I]$ ，利用行变换同时处理多个方程组，以此系统化地求解逆矩阵，其本质是用初等矩阵 E 左乘 A 来将 A 化为单位矩阵。

补充：Kronecker 积

设 $A = (a_{ij})$ 和 $B = (b_{ij})$ 分别是 $m \times n$ 阶和 $k \times l$ 阶矩阵，定义他们的 Kronecker 积为一个 $mk \times nl$ 阶矩阵：

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

Kronecker 积满足如下九条性质：

1. $(A + B) \otimes C = A \otimes C + B \otimes C$
2. $A \otimes (B + C) = A \otimes B + A \otimes C$
3. $(kA) \otimes B = k(A \otimes B) = A \otimes (kB)$
4. $I_m \otimes I_n = I_{mn}$
5. $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$
6. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
7. 若 A 和 B 是可逆矩阵，则 $A \otimes B$ 也是可逆矩阵，且 $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
8. 设 A 是 m 阶方阵， B 是 n 阶方阵，则 $|A \otimes B| = |A|^n |B|^m$
9. 设 A 是 m 阶方阵， B 是 n 阶方阵，则 $r(A \otimes B) = r(A) \cdot r(B)$