作者	**生姜 DrGinger**
脚本	**生姜 DrGinger**
视频	**崔崔 CuiCui**
开源学习资源	https://github.com/Visualize-ML
平台	https://www.youtube.com/@DrGinger_Jiang https://space.bilibili.com/3546865719052873 https://space.bilibili.com/513194466

3.2 矩阵乘法的第二视角

本节你将掌握的核心技能：

- 外积展开视角：矩阵乘法可以表示为多个外积矩阵之和。
- 每个外积是秩一矩阵。
- 秩一矩阵具有行、列倍数关系：体现线性相关性。
- 零向量外积结果为零矩阵。

需要我用代码或图例帮你复现某一部分内容吗？

矩阵乘法的第二视角是一种将乘法左侧矩阵写成一组列向量、右侧矩阵写成行向量的分解方式。这种视角揭示了矩阵乘法的另一种本质——列向量与行向量的外积之和。通过这种分解，我们可以更深入地理解矩阵乘法的结构和意义。

叠加

对于矩阵乘法 $\boldsymbol{A} @ \boldsymbol{B}$，如图 1 所示，把左侧矩阵 $\boldsymbol{A}$ 写成一组列向量

$$\boldsymbol{A}_{m\times p} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p \end{bmatrix} \tag{1}$$

其中，列向量 $\boldsymbol{a}_k$ $(k = 1, 2, \ldots, p)$ 的形状为 $m \times 1$。

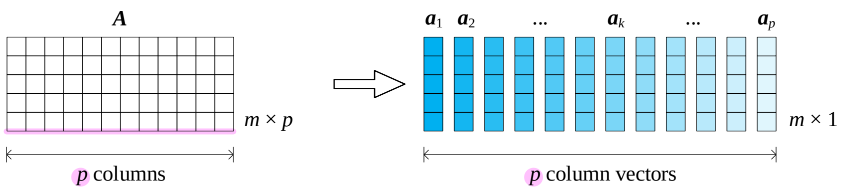

图 1. 把左侧矩阵 $\boldsymbol{A}$ 写成一组列向量

如图 2 所示，把右侧矩阵 $\boldsymbol{B}$ 写成一组行向量

$$\boldsymbol{B}_{p\times n}=\begin{bmatrix}\boldsymbol{b}^{(1)}\\ \boldsymbol{b}^{(2)}\\ \vdots\\ \boldsymbol{b}^{(p)}\end{bmatrix} \tag{2}$$

其中，行向量 $\boldsymbol{b}^{(k)}$的形状为 $1\times n$。

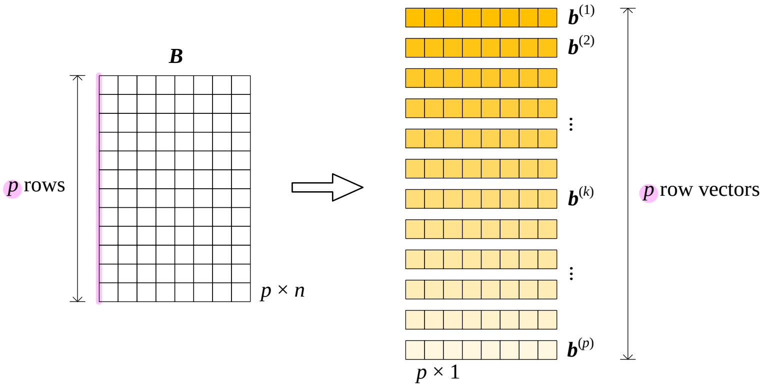

图 2. 把右侧矩阵 $\boldsymbol{B}$ 写成一组行向量

这样矩阵乘法 $\boldsymbol{C} = \boldsymbol{A} @ \boldsymbol{B}$ 可以写成

$$\boldsymbol{A}_{m\times p} @ \boldsymbol{B}_{p\times n}=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_p\end{bmatrix} @ \begin{bmatrix}\boldsymbol{b}^{(1)}\\ \boldsymbol{b}^{(2)}\\ \vdots\\ \boldsymbol{b}^{(p)}\end{bmatrix}=\boldsymbol{a}_1 @ \boldsymbol{b}^{(1)}+\boldsymbol{a}_2 @ \boldsymbol{b}^{(2)}+\cdots+\boldsymbol{a}_p @ \boldsymbol{b}^{(p)}=\sum_{k=1}^{p}\boldsymbol{a}_k @ \boldsymbol{b}^{(k)} \tag{3}$$

大家是否惊奇地发现，我们把矩阵乘法写成一组相同形状矩阵的求和！具体如图 3 所示。

也就是说矩阵乘法 $\boldsymbol{A} @ \boldsymbol{B}$ 转化成了 p 个矩阵叠加。而 p 对应 $\boldsymbol{A}$ 的列数、$\boldsymbol{B}$ 的行数，而 p 是那个“消去”的维度。

令

$$\boldsymbol{C}_k=\boldsymbol{a}_k @ \boldsymbol{b}^{(k)} \tag{4}$$

$\boldsymbol{C}_k$ 就是我们上一节提到的外积。此外，我们注意到矩阵 $\boldsymbol{C}$ 和 $\boldsymbol{C}_k$ 形状完全相同，都是 $m\times n$。

从矩阵形状来看，$(m \times 1) @ (1 \times n)$ 夹在中间的 (1) 被“消去”，矩阵乘法结果的形状为 $(m \times n)$。

这便是，矩阵乘法规则的第二视角——**外积展开** (outer product expansion)。

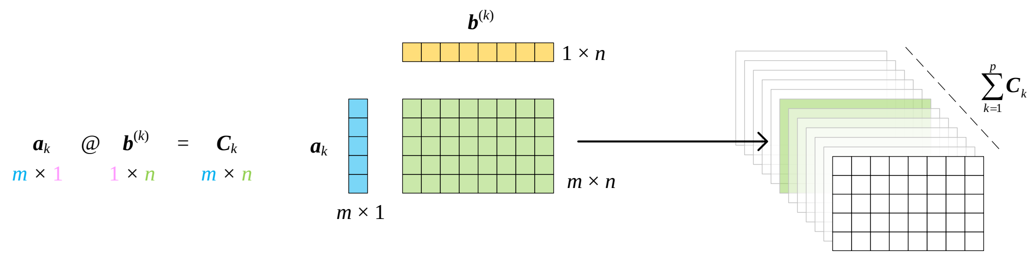

图 3. $\boldsymbol{A}$ @ $\boldsymbol{B}$ 可以写成一组相同形状矩阵的求和

第一个例子

回到本章第一节矩阵乘法的例子。

$$\boldsymbol{A} @ \boldsymbol{B}=\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} @ \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}=\begin{bmatrix}14 & 32\\32 & 77\end{bmatrix} \tag{5}$$

左侧矩阵 $\boldsymbol{A}$ 写成列向量，右侧矩阵 $\boldsymbol{B}$ 写成行向量后，$\boldsymbol{A}$ @ $\boldsymbol{B}$ 可以写成

$$\boldsymbol{A} @ \boldsymbol{B}=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3\end{bmatrix}\begin{bmatrix}\boldsymbol{b}^{(1)}\\\boldsymbol{b}^{(2)}\\\boldsymbol{b}^{(3)}\end{bmatrix}=\boldsymbol{a}_1 @ \boldsymbol{b}^{(1)}+\boldsymbol{a}_2 @ \boldsymbol{b}^{(2)}+\boldsymbol{a}_3 @ \boldsymbol{b}^{(3)} \tag{6}$$

代入具体值

$$\begin{aligned}
\boldsymbol{a}_1 @ \boldsymbol{b}^{(1)}&=\begin{bmatrix}1\\4\end{bmatrix} @ [1,4]=\begin{bmatrix}1\times 1 & 1\times 4\\4\times 1 & 4\times 4\end{bmatrix}=\begin{bmatrix}1 & 4\\4 & 16\end{bmatrix}\\
\boldsymbol{a}_2 @ \boldsymbol{b}^{(2)}&=\begin{bmatrix}2\\5\end{bmatrix} @ [2,5]=\begin{bmatrix}2\times 2 & 2\times 5\\5\times 2 & 5\times 5\end{bmatrix}=\begin{bmatrix}4 & 10\\10 & 25\end{bmatrix}\\
\boldsymbol{a}_3 @ \boldsymbol{b}^{(3)}&=\begin{bmatrix}3\\6\end{bmatrix} @ [3,6]=\begin{bmatrix}3\times 3 & 3\times 6\\6\times 3 & 6\times 6\end{bmatrix}=\begin{bmatrix}9 & 18\\18 & 36\end{bmatrix}
\end{aligned} \tag{7}$$

图 4 所示分别计算 $\boldsymbol{C}_1$、$\boldsymbol{C}_2$、$\boldsymbol{C}_3$。

从矩阵形状来看，(2×1) @ (1×2) 夹在中间的 (1) 被“消去”，矩阵乘法结果的形状为 (2×2)。

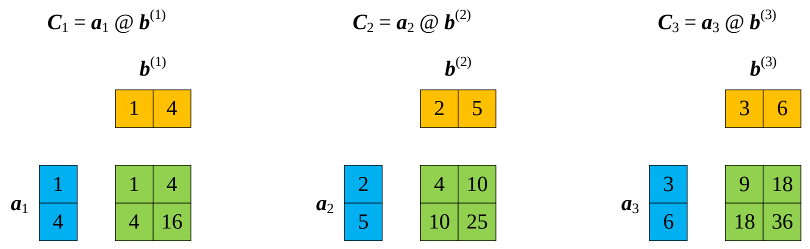

图 4. 分别计算 $\boldsymbol{C}_1$、$\boldsymbol{C}_2$、$\boldsymbol{C}_3$

如图 5 所示，矩阵乘法 $\boldsymbol{C}=\boldsymbol{A}$ @ $\boldsymbol{B}$ 可以写成三个矩阵的叠加

$$\boldsymbol{C} = \begin{bmatrix} 1 & 4 \\ 4 & 16 \end{bmatrix} + \begin{bmatrix} 4 & 10 \\ 10 & 25 \end{bmatrix} + \begin{bmatrix} 9 & 18 \\ 18 & 36 \end{bmatrix} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix} \tag{8}$$

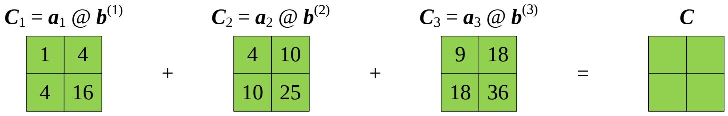

图 5. $\boldsymbol{C}_1$、$\boldsymbol{C}_2$、$\boldsymbol{C}_3$三个矩阵叠加

第二个例子

左侧矩阵 $\boldsymbol{B}$ 写成列向量，右侧矩阵 $\boldsymbol{A}$ 写成行向量后，$\boldsymbol{B}$ @ $\boldsymbol{A}$ 可以写成

$$\boldsymbol{B} @ \boldsymbol{A} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}^{(1)} \\ \boldsymbol{a}^{(2)} \end{bmatrix} = \boldsymbol{b}_1 @ \boldsymbol{a}^{(1)} + \boldsymbol{b}_2 @ \boldsymbol{a}^{(2)} \tag{9}$$

代入具体值

$$\begin{aligned} \boldsymbol{b}_1 @ \boldsymbol{a}^{(1)} &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} @ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1\times1 & 1\times2 & 1\times3 \\ 2\times1 & 2\times2 & 2\times3 \\ 3\times1 & 3\times2 & 3\times3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \\ \boldsymbol{b}_2 @ \boldsymbol{a}^{(2)} &= \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} @ \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 4\times4 & 4\times5 & 4\times6 \\ 5\times4 & 5\times5 & 5\times6 \\ 6\times4 & 6\times5 & 6\times6 \end{bmatrix} = \begin{bmatrix} 16 & 20 & 24 \\ 20 & 25 & 30 \\ 24 & 30 & 36 \end{bmatrix} \end{aligned} \tag{10}$$

图 6 所示分别计算 $\boldsymbol{D}_1$、$\boldsymbol{D}_2$。

从矩阵形状来看，(3 × 1) @ (1 × 3) 夹在中间的 (1) 被“消去”，矩阵乘法结果的形状为 (3 × 3)。

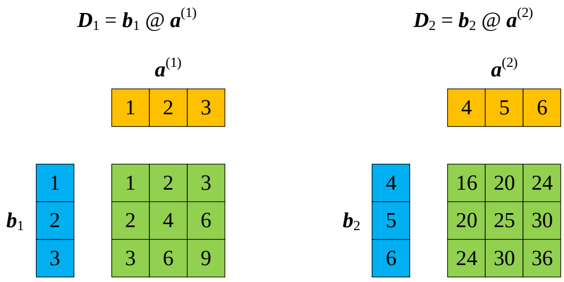

图 6. 分别计算 $\boldsymbol{D}_1$、$\boldsymbol{D}_2$

如图 7 所示，矩阵乘法 $\boldsymbol{D} = \boldsymbol{B}$ @ $\boldsymbol{A}$ 可以写成两个矩阵的叠加

$$\boldsymbol{D} = \boldsymbol{b}_1 @ \boldsymbol{a}^{(1)} + \boldsymbol{b}_2 @ \boldsymbol{a}^{(2)} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} + \begin{bmatrix} 16 & 20 & 24 \\ 20 & 25 & 30 \\ 24 & 30 & 36 \end{bmatrix} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix} \tag{11}$$

$$
\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}}_{\boldsymbol{D}_1 = \boldsymbol{b}_1 @ \boldsymbol{a}^{(1)}} + \underbrace{\begin{bmatrix} 16 & 20 & 24 \\ 20 & 25 & 30 \\ 24 & 30 & 36 \end{bmatrix}}_{\boldsymbol{D}_2 = \boldsymbol{b}_2 @ \boldsymbol{a}^{(2)}} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix}
$$

图 7. $\boldsymbol{D}_1$、$\boldsymbol{D}_2$ 两个矩阵叠加

LA_03_02_01.ipynb 以矩阵乘法第二视角完成以上两个矩阵乘法运算，请大家自学。

热图展示矩阵乘法第二视角

下面让我们用热图来可视化矩阵乘法第二视角。

给定如图 8 所示的矩阵乘法。

图 9 用热图展示如何计算 $\boldsymbol{a}_1 @ \boldsymbol{b}^{(1)}$。

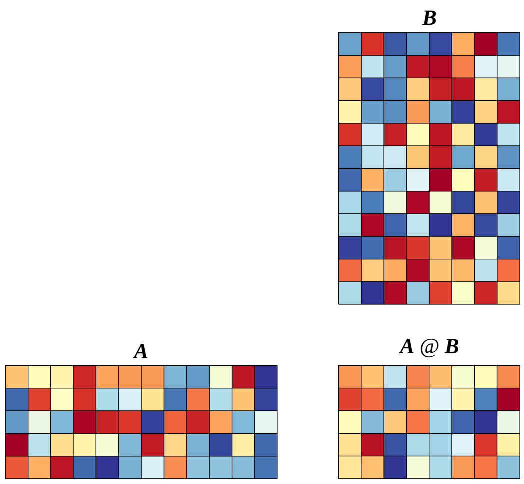

图 8. 矩阵乘法 $\boldsymbol{A} @ \boldsymbol{B}$ 热图，第一视角

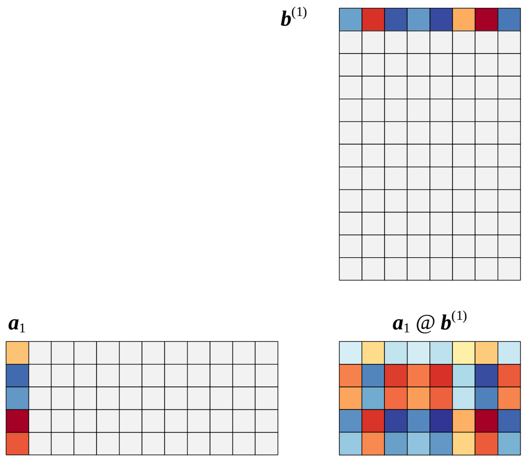

图 9. 热图展示 $\boldsymbol{a}_1$ @ $\boldsymbol{b}^{(1)}$

由于 $\boldsymbol{A}$ 有 12 列 ($\boldsymbol{B}$ 有 12 行)，类似 $\boldsymbol{a}_1$ @ $\boldsymbol{b}^{(1)}$矩阵乘法一共有 12 个，如图 10 所示。

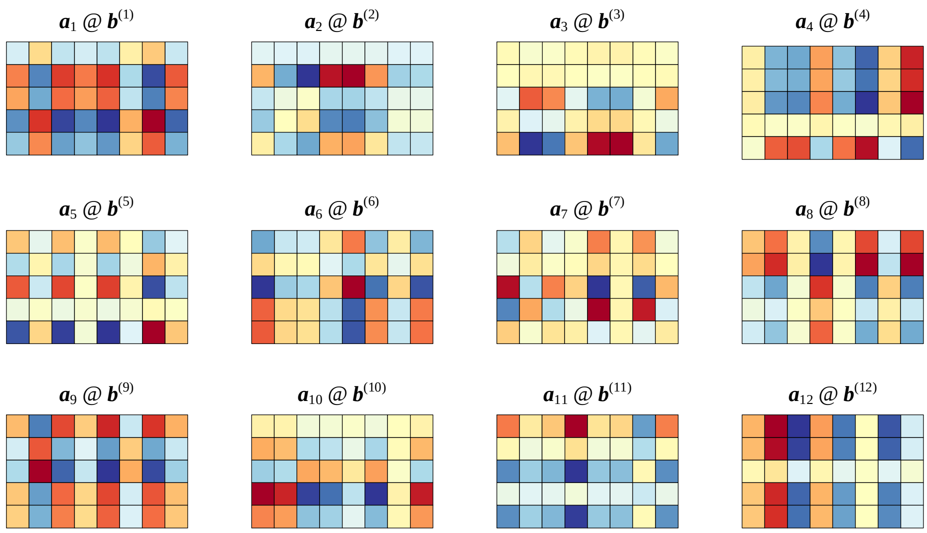

图 10. 12 个 $\boldsymbol{a}_k$ @ $\boldsymbol{b}^{(k)}$

图 10 中 12 个矩阵相加的结果为 $\boldsymbol{A}$ @ $\boldsymbol{B}$。

LA_03_02_02.ipynb 绘制如上热图，请大家自学。

倍数关系

如图 11 所示，给定 $\boldsymbol{a}$、$\boldsymbol{b}$ 两个列向量，矩阵乘法 $\boldsymbol{a}$ @ $\boldsymbol{b}^{\mathrm{T}}$ 的行、列存在倍数关系

$$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} @ \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m b_1 & a_m b_2 & \cdots & a_m b_n \end{bmatrix} \tag{12}$$

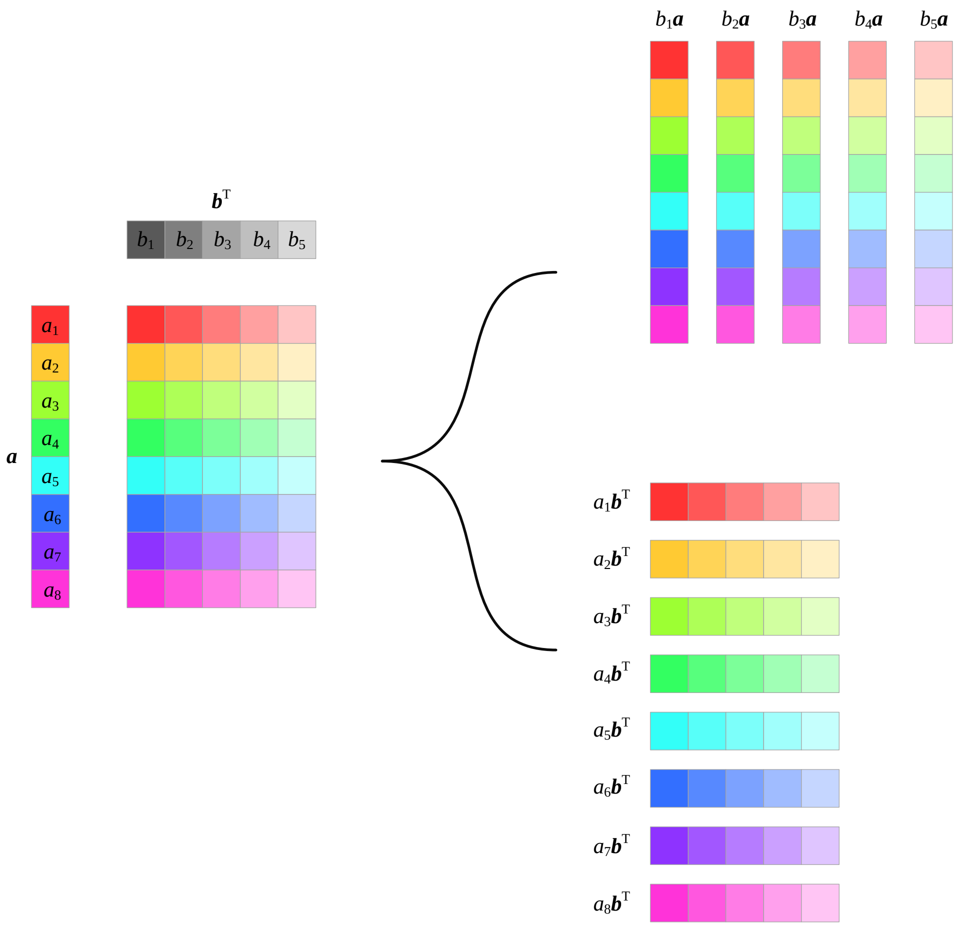

图 11. 矩阵乘法 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 的行、列存在倍数关系

让我们先看 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 的列，即

$$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}} = \boldsymbol{a} @ \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} b_1\boldsymbol{a} & b_2\boldsymbol{a} & \cdots & b_n\boldsymbol{a} \end{bmatrix} \tag{13}$$

展开之后得到

$$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} b_1 \times \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} & b_2 \times \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} & \cdots & b_n \times \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \end{bmatrix} \tag{14}$$

让我们再看 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 的行

$$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \boldsymbol{b}^{\mathrm{T}} = \begin{bmatrix} a_1 \boldsymbol{b}^{\mathrm{T}} \\ a_2 \boldsymbol{b}^{\mathrm{T}} \\ \vdots \\ a_m \boldsymbol{b}^{\mathrm{T}} \end{bmatrix} \tag{15}$$

展开之后得到

$$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} a_1 \times \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \\ a_2 \times \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \\ \vdots \\ a_m \times \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \end{bmatrix} \tag{16}$$

正因为这种行列倍数关系，我们管 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 这种外积叫做**秩一矩阵** (rank-one matrix)。

秩一矩阵的每一行都是同一个行向量的倍数；秩一矩阵的每一列都是同一个列向量的倍数。

这说明秩一矩阵的行和列之间具有线性相关性，即矩阵的秩 (rank) 为 1。简单来说，矩阵的秩是其行或列向量组的最大线性无关个数，等价于该矩阵中不为零的最大线性无关行、列向量的个数。这是本书后续要深入讲解的话题。

⚠ 注意，$\boldsymbol{a}$ 或 $\boldsymbol{b}$ 若为零向量 $\boldsymbol{0}$，$\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 的结果为零矩阵 $\boldsymbol{O}$；此时，矩阵 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 的秩为 0。

回到前文的例子，大家已经能够看到列向量之间的倍数关系

$$\boldsymbol{b}_1 @ \boldsymbol{a}^{(1)} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} @ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} & 2 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} & 3 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \end{bmatrix} \tag{17}$$

再看行向量之间的倍数关系

$$\boldsymbol{b}_1 @ \boldsymbol{a}^{(1)} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} @ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 \times \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \\ 2 \times \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \\ 3 \times \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \end{bmatrix} \tag{18}$$

自定义 Python 函数计算矩阵乘法：外积视角

代码 1 通过自定义函数，利用外积视角完成矩阵乘法。下面聊聊其中关键语句。

ⓐ 定义一个函数，函数名叫 matrix_multiplication_outer，意思是“用外积的方式做矩阵乘法”；函数的意思就是“把某一段代码打包成一个工具”，需要输入矩阵 A 和 B。

ⓑ 开启一个循环，从 k = 0 一直到 p，也就是说我们会把 A 的每一列都拿出来一遍，同时也会取出 B 的每一行；这是准备开始一列一列地做“外积”。

a_k = A[:, [k]] 取出矩阵 A 的第 k 列，变成一个竖着的列向量；b_k = B[[k], :] 取出矩阵 B 的第 k 行。 C += (a_k @ b_k) 计算内积，把这个矩阵加到结果矩阵 C 上，表示把这一步的贡献累加进去。

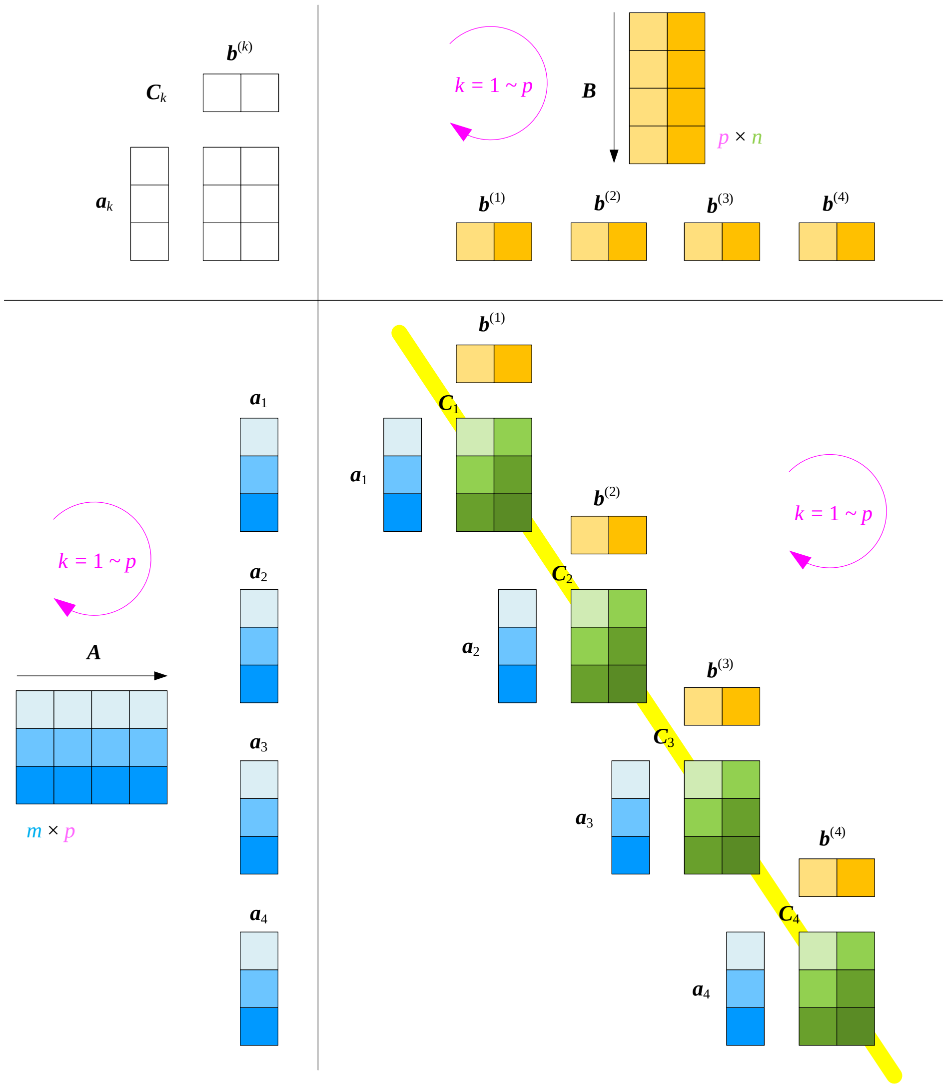

图 12. 矩阵乘法 AB 规则，外积视角

代码 1. 自定义 Python 函数计算矩阵乘法，外积视角 | LA_03_02_03.ipynb

```python
## 初始化
import numpy as np

## 自定义函数，矩阵乘法第二视角
def matrix_multiplication_outer(A, B):

    # 获取矩阵 A 和 B 的形状
    m, p_A = A.shape
    p_B, n = B.shape

    # 检测矩阵形状是否符合矩阵乘法规则
    if p_A != p_B:
        raise ValueError('Dimensions do not match')

    # 初始化结果矩阵 C，形状 (m, n)，初始值设为 0
    C = np.zeros((m, n))

    for k in range(p_A):
        a_k = A[:, [k]]      # A 的第 k 列，二维数组
        b_k = B[[k], :]      # B 的第 k 行，二维数组
        C += (a_k @ b_k)     # 每次计算外积并加到结果中

    return C

## 矩阵乘法
A = np.array([[1, 2, 3],
              [4, 5, 6]])
B = A.T

## 矩阵乘法
matrix_multiplication_outer(A, B)
matrix_multiplication_outer(B, A)
```

请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 给出如下成对矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，分别用矩阵第二视角计算 $\boldsymbol{A}$ @ $\boldsymbol{B}$ 和 $\boldsymbol{B}$ @ $\boldsymbol{A}$

- $\boldsymbol{A}=\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}$
- $\boldsymbol{A}=\begin{bmatrix}1 & 0\\3 & 1\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}1 & 0\\2 & 1\end{bmatrix}$
- $\boldsymbol{A}=\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}1 & 3\\0 & 1\end{bmatrix}$
- $\boldsymbol{A}=\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix}$, $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{T}}$

Q2. 给定如下成对列向量，请大家把矩阵乘法 $\boldsymbol{a} @ \boldsymbol{b}^{\mathrm{T}}$ 写成行、列倍数关系。

► $\boldsymbol{a} = [1, 2]^{\mathrm{T}}, \boldsymbol{b} = [1, 2]^{\mathrm{T}}$

► $\boldsymbol{a} = [1, 2]^{\mathrm{T}}, \boldsymbol{b} = [1, 2, 3]^{\mathrm{T}}$

► $\boldsymbol{a} = [1, 2, 3]^{\mathrm{T}}, \boldsymbol{b} = [1, 2]^{\mathrm{T}}$

Q3. 什么是矩阵的秩，和矩阵的大小有关系吗？

Q4. 什么是线性相关？如何从几何角度理解线性相关？

Q5. 什么是线性无关？如何从几何角度理解线性无关？