

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

3.4 矩阵乘法的第四视角



本节你将掌握的核心技能：

- ▶ 行向量线性组合视角：矩阵乘法 AB ，每一行行向量结果向量是右侧矩阵 B 行向量的加权和。
- ▶ 矩阵乘法 AB ， A 的每一行作为权重，作用于 B 的每一行，加权求和生成 C 的每一行。
- ▶ 矩阵乘法的行向量视角与列向量视角通过转置相互转换。
- ▶ 矩阵乘法 AB ， A 有几行， AB 就有几行。

矩阵乘法的第四视角将矩阵乘法视为行向量的线性组合。这种视角从行向量的角度出发，揭示了矩阵乘法的另一种观察视角——对于矩阵乘法 $C = A @ B$ ， C 的每一行是矩阵 B 行向量线性组合的结果。

矩阵 A 的每一行元素作为系数，对矩阵 B 的行向量进行加权求和，生成结果矩阵 C 的对应行。

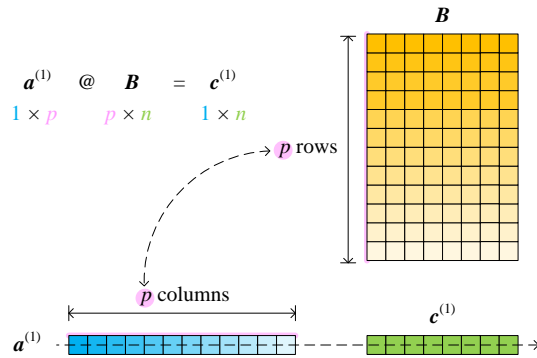
矩阵 C 的第一行

大家是否好奇，既然我们可以把矩阵 C 的每一列看成是矩阵 A 列向量的线性组合，是否也可以把矩阵 C 看成矩阵 B 的某种线性组合？

答案是肯定的！

以矩阵 C 的第一行为例，如图 1 所示， $c^{(1)}$ 对应的矩阵乘法为

$$c^{(1)} = a^{(1)} @ B \quad (1)$$

图 1. $c^{(1)} = a^{(1)}B$ 对应的矩阵乘法

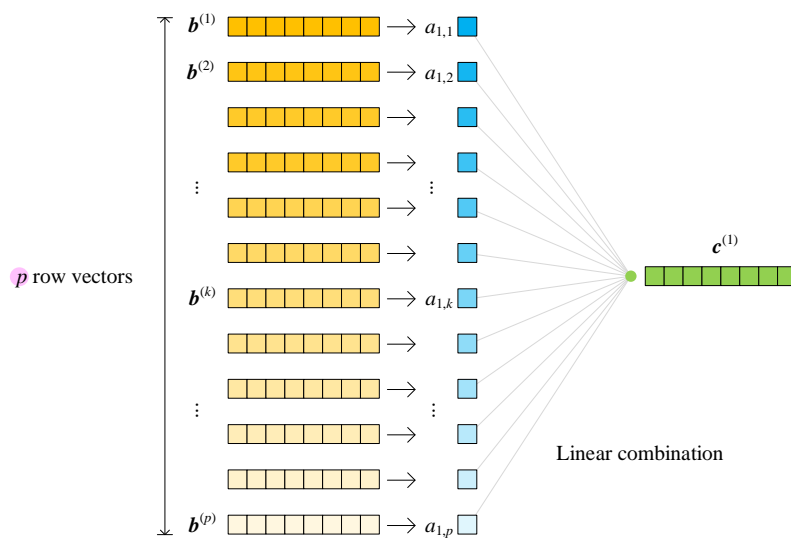
把矩阵 B 写成一组行向量

$$B = \begin{bmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(D)} \end{bmatrix} \quad (2)$$

如图 2 所示, $c^{(1)}$ 为矩阵 B 的 p 个行向量的线性组合 (加权和)

$$c^{(1)} = a^{(1)} @ B = \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,D} \end{bmatrix}}_{a^{(1)}} \underbrace{\begin{bmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(D)} \end{bmatrix}}_B = a_{1,1}b^{(1)} + a_{1,2}b^{(2)} + \dots + a_{1,D}b^{(D)} \quad (3)$$

我们管它叫做矩阵乘法的第四视角, 即行向量线性组合视角。

图 2. B 行向量的线性组合

转置

再换个视角，把图 1 矩阵乘法 $\mathbf{c}^{(1)} = \mathbf{a}^{(1)}\mathbf{B}$ 转置，我们得到图 3，对应

$$\mathbf{c}^{(1)\mathrm{T}} = \mathbf{B}^{\mathrm{T}} @ \mathbf{a}^{(1)\mathrm{T}} \quad (4)$$

这样，我们又回到了列向量的线性组合。下图也从一个侧面再次印证了矩阵乘法 $\mathbf{AB} = \mathbf{C}$ 的转置表示为 $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}$ 。

⚠ $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ ，大家千万别忘了调转顺序。

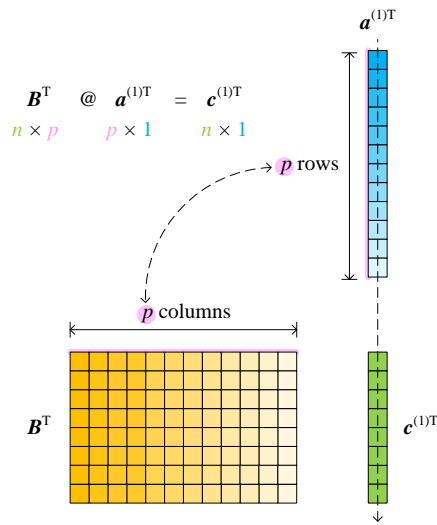


图 3. 矩阵乘法 $\mathbf{a}^{(1)}\mathbf{B} = \mathbf{c}^{(1)}$ 转置

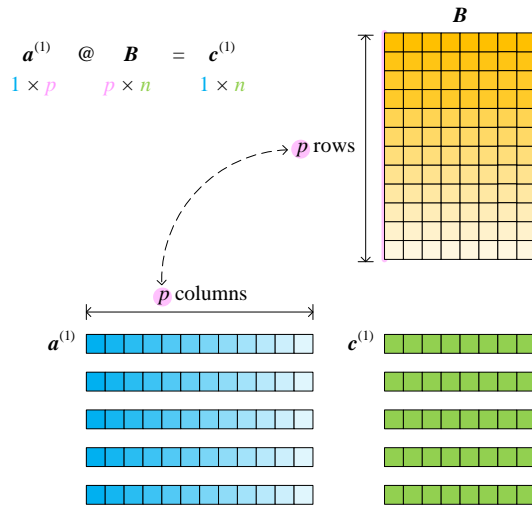
再看 $\mathbf{A} @ \mathbf{B} = \mathbf{C}$

同样，我们也可以出行向量线性组合视角来观察矩阵乘法 $\mathbf{C} = \mathbf{A} @ \mathbf{B}$ 。

如图 4 所示，将矩阵 \mathbf{A} 拆成一组行向量，就把矩阵乘法 $\mathbf{AB} = \mathbf{C}$ 拆成了 m 个行向量线性组合！

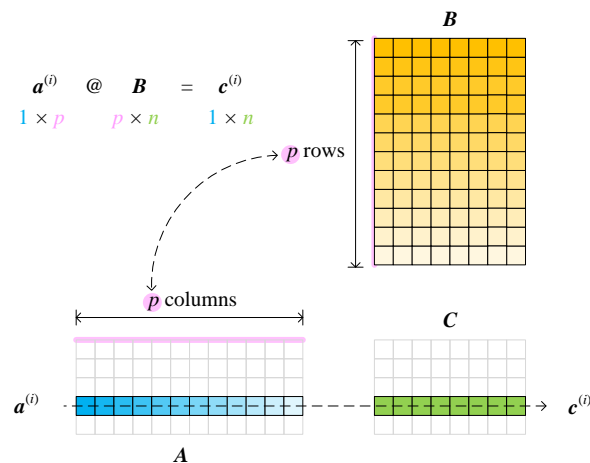
$$\mathbf{C} = \begin{bmatrix} \mathbf{c}^{(1)} \\ \mathbf{c}^{(2)} \\ \vdots \\ \mathbf{c}^{(m)} \end{bmatrix} = \begin{bmatrix} \mathbf{a}^{(1)} \\ \mathbf{a}^{(2)} \\ \vdots \\ \mathbf{a}^{(m)} \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{a}^{(1)}\mathbf{B} \\ \mathbf{a}^{(2)}\mathbf{B} \\ \vdots \\ \mathbf{a}^{(m)}\mathbf{B} \end{bmatrix} \quad (5)$$

图 4 告诉我们矩阵 \mathbf{A} 有 m 行，矩阵乘法 $\mathbf{A} @ \mathbf{B}$ 的结果也有 m 行。

图 4. 把矩阵 C 、 A 分别写成一组列向量

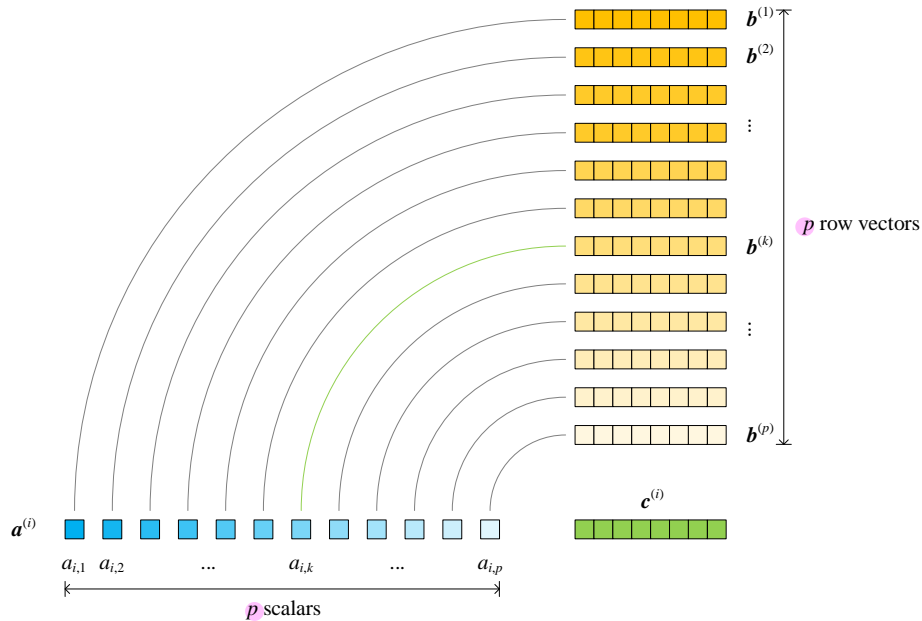
如图 5 所示，矩阵 C 的任意一行可以通过下式计算得到

$$c^{(i)} = a^{(i)} @ B \quad (6)$$

图 5. $c^{(i)} = a^{(i)}B$ 对应的矩阵乘法

因此，如图 6 所示，矩阵 C 的第 i 行行向量 $c^{(i)}$ 可以进一步写成矩阵 B 的行向量的线性组合，组合系数由矩阵 A 的第 i 行行向量 $a^{(i)}$ 的每个分量提供，即

$$c^{(i)} = a^{(i)} @ B = \underbrace{\begin{bmatrix} a_{i,1} & a_{i,2} & \cdots & a_{i,D} \end{bmatrix}}_{a^{(i)}} \begin{bmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(D)} \end{bmatrix}_B = a_{i,1}b^{(1)} + a_{i,2}b^{(2)} + \cdots + a_{i,D}b^{(D)} \quad (7)$$

图 6. 矩阵 B 的行向量线性组合视角看 $c^{(i)} = a^{(i)}B$

反过来看，分别计算了 $a^{(1)}B$ 、 $a^{(2)}B$ 、 $a^{(3)}B$ 等等之后，再把他们按顺序排列成一列，便得到矩阵 C ，即

$$\begin{bmatrix} a^{(1)}B \\ a^{(2)}B \\ \vdots \\ a^{(m)}B \end{bmatrix} = \begin{bmatrix} c^{(1)} \\ c^{(2)} \\ \vdots \\ c^{(m)} \end{bmatrix} \quad (8)$$

第一个例子

先看本书前文讲过的矩阵乘法的例子，利用行向量线性组合计算 $c^{(1)}$

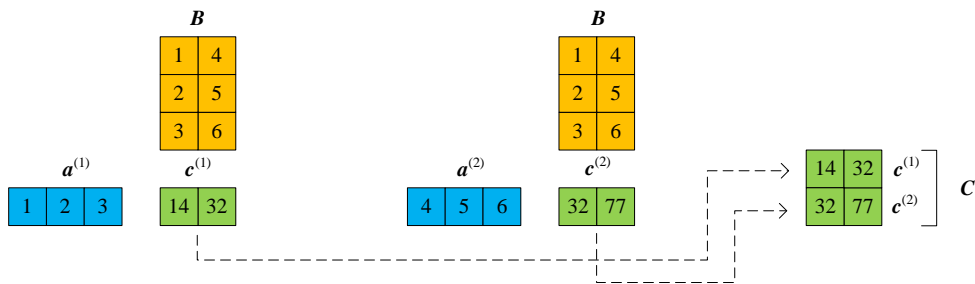
$$c^{(1)} = a^{(1)} @ B = [1 \ 2 \ 3] @ \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = 1 \times [1 \ 4] + 2 \times [2 \ 5] + 3 \times [3 \ 6] = [14 \ 32] \quad (9)$$

先计算 $c^{(2)}$

$$c^{(2)} = a^{(2)} @ B = [4 \ 5 \ 6] @ \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = 4 \times [1 \ 4] + 5 \times [2 \ 5] + 6 \times [3 \ 6] = [32 \ 77] \quad (10)$$

如图 7 所示，再把行向量 $c^{(1)}$ 、 $c^{(2)}$ 顺序摆放得到矩阵 C

$$C = A @ B = \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} = \begin{bmatrix} [14 \ 32] \\ [32 \ 77] \end{bmatrix} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix} \quad (11)$$

图 7. 把行向量 $c^{(1)}$ 、 $c^{(2)}$ 顺序摆放得到矩阵 C

第二个例子

用行向量线性组合计算 $d^{(1)}$

$$d^{(1)} = b^{(1)} @ A = [1 \ 4] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = 1 \times [1 \ 2 \ 3] + 4 \times [4 \ 5 \ 6] = [17 \ 22 \ 27] \quad (12)$$

先计算 $d^{(2)}$

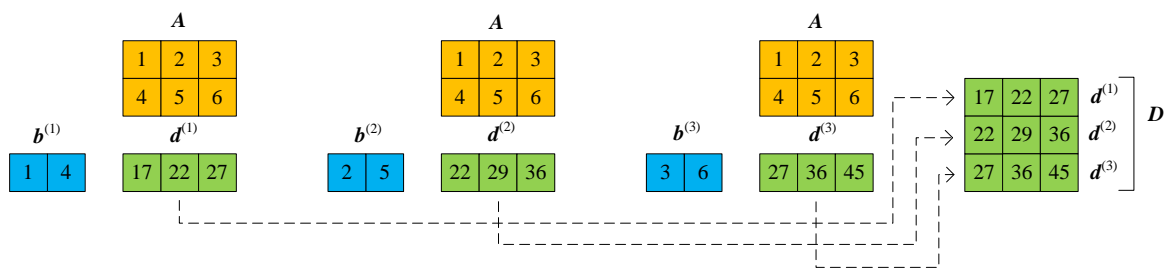
$$d^{(2)} = b^{(2)} @ A = [2 \ 5] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = 2 \times [1 \ 2 \ 3] + 5 \times [4 \ 5 \ 6] = [22 \ 29 \ 36] \quad (13)$$

最后计算 $d^{(3)}$

$$d^{(3)} = b^{(3)} @ A = [3 \ 6] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = 3 \times [1 \ 2 \ 3] + 6 \times [4 \ 5 \ 6] = [27 \ 36 \ 45] \quad (14)$$

如图 8 所示，把行向量 $d^{(1)}$ 、 $d^{(2)}$ 、 $d^{(3)}$ 顺序摆放得到矩阵 D

$$D = B @ A = \begin{bmatrix} d^{(1)} \\ d^{(2)} \\ d^{(3)} \end{bmatrix} = \begin{bmatrix} [17 \ 22 \ 27] \\ [22 \ 29 \ 36] \\ [27 \ 36 \ 45] \end{bmatrix} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix} \quad (15)$$

图 8. 把行向量 $d^{(1)}$ 、 $d^{(2)}$ 、 $d^{(3)}$ 顺序摆放得到矩阵 D 

LA_03_04_01.ipynb 完成以上矩阵乘法第三视角计算，请大家自学。

热图展示矩阵乘法第四视角

下面让我们用热图来可视化矩阵乘法第四视角。

图 9 用热图展示如何计算 $a^{(1)} @ B$ 。

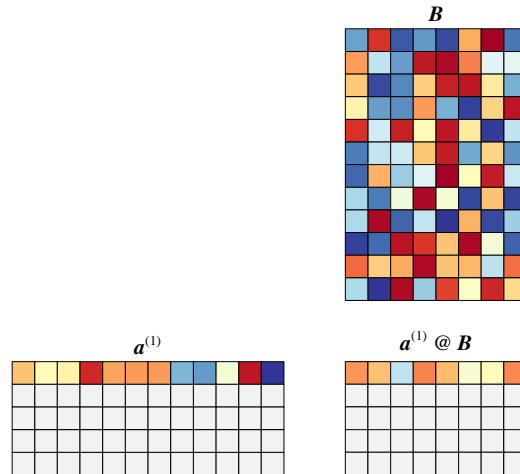


图 9. 热图展示 $a^{(1)} @ B$

由于 A 有 5 行，类似 $a^{(1)} @ B$ 矩阵乘法一共有 5 个，结果如图 10 所示。它们的结果上下顺序排列便得到矩阵 $C = A @ B$ 。

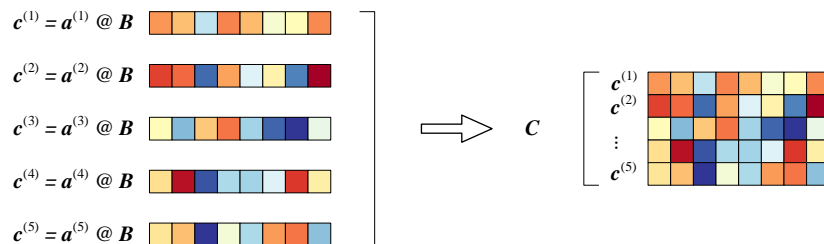



图 10. 5 个 $a^{(i)} @ B$ 上下排列得到矩阵 $C = A @ B$



LA_03_04_02.ipynb 利用 `seaborn.heatmap()` 绘制上述热图可视化矩阵乘法第四视角，请大家自学。

代码 1 通过自定义函数，利用行向量线性组合视角完成矩阵乘法。代码和上一节“列向量线性组合视角”类似，请大家对照分析这段代码。

代码 1. 自定义 Python 函数计算矩阵乘法，行向量线性组合视角 |  LA_03_04_03.ipynb

```
## 初始化
import numpy as np

## 定义行向量线性组合函数
def row_LC(B, coefficients):
    p, n = B.shape
    row_LC_result = np.zeros(n)
    # 初始化行向量组合结果

    for k in range(p):
        row_LC_result += coefficients[k] * B[k, :]
        # 每行乘系数并相加
    return row_LC_result

## 定义矩阵乘法函数（行线性组合法）

def matrix_multiplication_row_LC(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 i in range(m):
        coeffs = A[i, :]
        # 取出 A 的第 i 列，作为线性组合的系数
        C[i, :] = row_LC(B, coeffs)
        # 计算 B 行向量的线性组合

    return C

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

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




请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 用列向量线性组合计算如下成对矩阵乘法

► $\mathbf{a}^{(1)} = [1 \ 2], \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

► $\mathbf{a}^{(1)} = [1 \ 2 \ 3], \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

► $\mathbf{a}^{(1)} = [1 \ 2], \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

► $\mathbf{a}^{(1)} = [1 \ 2 \ 3], \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

Q2. 用矩阵乘法第四视角展开如下成对矩阵乘法

► $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

► $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

► $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \mathbf{B} = \mathbf{A}^T$

► $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{B} = \mathbf{A}^T$

Q3. 把以下两个矩阵乘法写成一个。

$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} @ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} @ \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Q4. 把以下三个矩阵乘法写成一个，并给出结果。

$[1 \ 0 \ 0] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, [0 \ 1 \ 0] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, [0 \ 0 \ 1] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

Q5. 把以下三个矩阵乘法写成一个，并给出结果。

$[0 \ 0 \ 1] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, [1 \ 0 \ 0] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, [0 \ 1 \ 0] @ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

Q6. 比较 Q4、Q5 的结果，解释两者结果的异同。