

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

3.3 矩阵乘法的第三视角



本节你将掌握的核心技能：

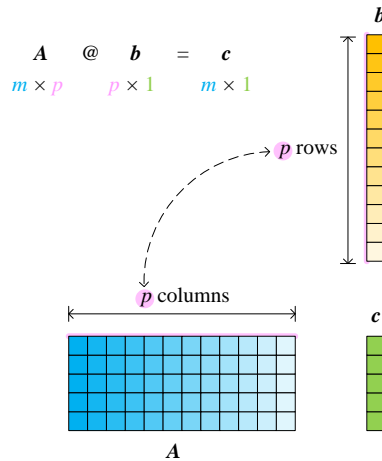
- ▶ 列向量线性组合视角：矩阵乘法 AB ，结果每一列列向量是左矩阵 A 列向量的加权和。
- ▶ Ab 就是 A 的列向量按 b 中系数加权组合成新列向量。
- ▶ 矩阵乘法 AB ， B 的每一列作为权重，作用于 A 的每一列，加权求和生成 C 的每一列。
- ▶ 矩阵乘法 AB ， B 有几列， AB 就有几列。

矩阵乘法的第三视角是一种将矩阵乘法视为列向量的线性组合的方式。这种视角从列向量的角度出发，揭示了矩阵乘法的另一种本质——对于矩阵乘法 $C = A @ B$ ， C 的每一列是矩阵 A 列向量线性组合的结果。

先看 $A @ b = c$

矩阵 A 的形状为 $m \times p$ ，列向量 b 的形状为 $p \times 1$ ；两者的乘积 $Ab = c$ 的形状为 $m \times 1$ 。

图 1 展示矩阵乘法 $Ab = c$ 的第一视角。

图 1. 矩阵乘法 $Ab = c$

列向量的线性组合

将 A 写成一组列向量

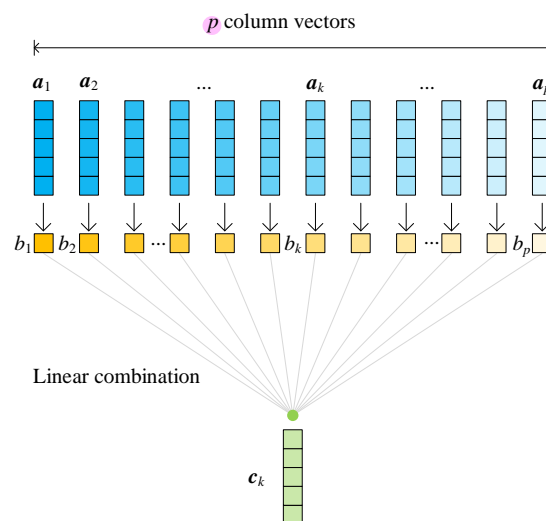
$$A_{m \times p} = [a_1 \ a_2 \ \cdots \ a_p] \quad (1)$$

其中，列向量 a_k ($k = 1, 2, \dots, p$) 的形状为 $m \times 1$ 。

矩阵乘法 $A @ b$ 可以写成

$$A_{m \times p} b_{p \times 1} = [a_1 \ a_2 \ \cdots \ a_p] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} = b_1 a_1 + b_2 a_2 + \cdots + b_p a_p = c \quad (2)$$

如图 2 所示， c 为矩阵 A 的 p 个列向量的线性组合 (加权和)。我们管它叫做矩阵乘法的第三视角，即列向量线性组合视角。

图 2. A 列向量的线性组合

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

再看 $A @ B = C$

同样，我们也可以用列向量线性组合视角来观察矩阵乘法 $C = A @ B$ 。

如图 3 所示，把矩阵 C 、 B 分别写成一组列向量，这样矩阵乘法 $C = A @ B$ 可以写成

$$C = [c_1 \ c_2 \ \cdots \ c_n] = A @ B = A @ [b_1 \ b_2 \ \cdots \ b_n] = [A @ b_1 \ A @ b_2 \ \cdots \ A @ b_n] \quad (3)$$

图 3 也告诉我们为什么矩阵 B 有 p 列，矩阵乘法 $A @ B$ 也有 p 列。

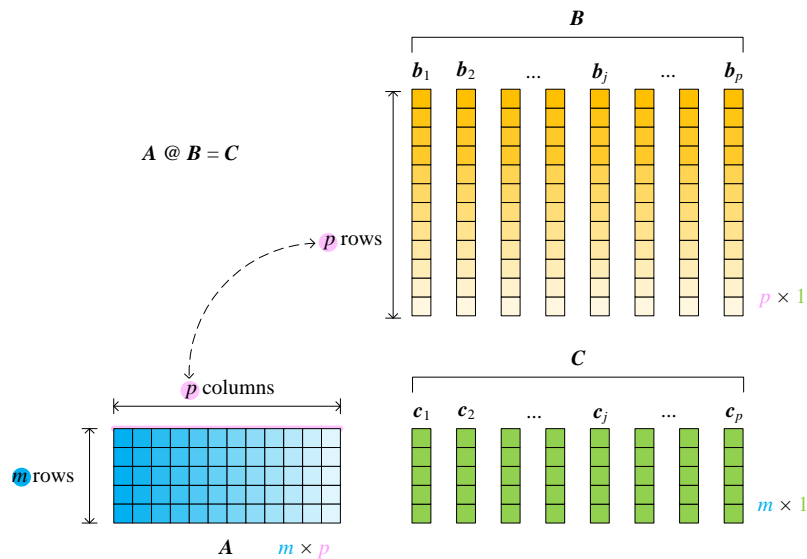


图 3. 把矩阵 C 、 B 分别写成一组列向量

如图 4 所示，矩阵 C 的任意一列可以通过下式计算得到

$$c_j = A @ b_j \quad (4)$$

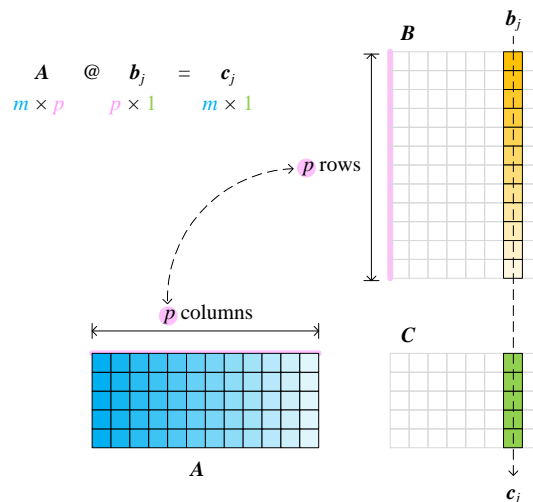
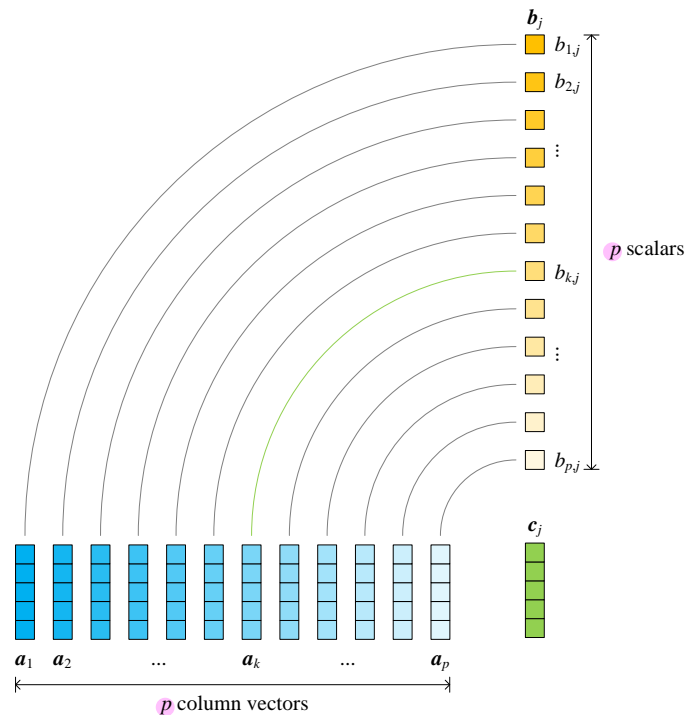


图 4. $c_j = Ab_j$ 对应的矩阵乘法

因此，如图 5 所示，矩阵 C 的第 j 列列向量 c_j 可以进一步写成矩阵 A 的列向量的线性组合，组合系数由矩阵 B 的第 j 列列向量 b_j 的每个分量提供，即

$$c_j = A @ b_j = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix} \begin{bmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{p,j} \end{bmatrix} = b_{1,j}a_1 + b_{2,j}a_2 + \cdots + b_{p,j}a_p \quad (5)$$

图 5. 矩阵 A 的列向量线性组合视角看 $c_j = Ab_j$

反过来看，分别计算了 Ab_1 、 Ab_2 、 Ab_3 等等之后，再把他们按顺序排列成一行，便得到矩阵 C 即

$$[A @ b_1 \quad A @ b_2 \quad \cdots \quad A @ b_n] = [c_1 \quad c_2 \quad \cdots \quad c_n] \quad (6)$$

第一个例子

先看本书前文讲过的矩阵乘法的例子，让我们先用符号展开 $A @ b_1$

$$c_1 = A @ b_1 = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} @ \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ b_{3,1} \end{bmatrix} = b_{1,1} \cdot a_1 + b_{2,1} \cdot a_2 + b_{3,1} \cdot a_3 \quad (7)$$

代入具体值，计算 c_1

$$\mathbf{c}_1 = \mathbf{A} @ \mathbf{b}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \times \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 2 \times \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 3 \times \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 1+4+9 \\ 4+10+18 \end{bmatrix} = \begin{bmatrix} 14 \\ 32 \end{bmatrix} \quad (8)$$

用符号展开 $\mathbf{A} @ \mathbf{b}_2$

$$\mathbf{c}_2 = \mathbf{A} @ \mathbf{b}_2 = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3] @ \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ b_{3,2} \end{bmatrix} = b_{1,2} \cdot \mathbf{a}_1 + b_{2,2} \cdot \mathbf{a}_2 + b_{3,2} \cdot \mathbf{a}_3 \quad (9)$$

代入具体值，计算 \mathbf{c}_2

$$\mathbf{c}_2 = \mathbf{A} @ \mathbf{b}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = 4 \times \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 5 \times \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 6 \times \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 4+10+18 \\ 16+25+36 \end{bmatrix} = \begin{bmatrix} 32 \\ 77 \end{bmatrix} \quad (10)$$

$$\begin{array}{l} b_{1,1} \times \begin{bmatrix} a_1 \\ 1 \\ 4 \end{bmatrix} + b_{2,1} \times \begin{bmatrix} a_2 \\ 2 \\ 5 \end{bmatrix} + b_{3,1} \times \begin{bmatrix} a_3 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} c_1 \\ 14 \\ 32 \end{bmatrix} \\ b_{1,2} \times \begin{bmatrix} a_1 \\ 1 \\ 4 \end{bmatrix} + b_{2,2} \times \begin{bmatrix} a_2 \\ 2 \\ 5 \end{bmatrix} + b_{3,2} \times \begin{bmatrix} a_3 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} c_2 \\ 32 \\ 77 \end{bmatrix} \end{array}$$

图 6. 通过线性组合计算 \mathbf{c}_1 、 \mathbf{c}_2

如图 7 所示，然后再把列向量 \mathbf{c}_1 、 \mathbf{c}_2 顺序摆放得到矩阵 \mathbf{C}

$$\mathbf{C} = \mathbf{A} @ \mathbf{B} = [\mathbf{c}_1 \quad \mathbf{c}_2] = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix} \quad (11)$$

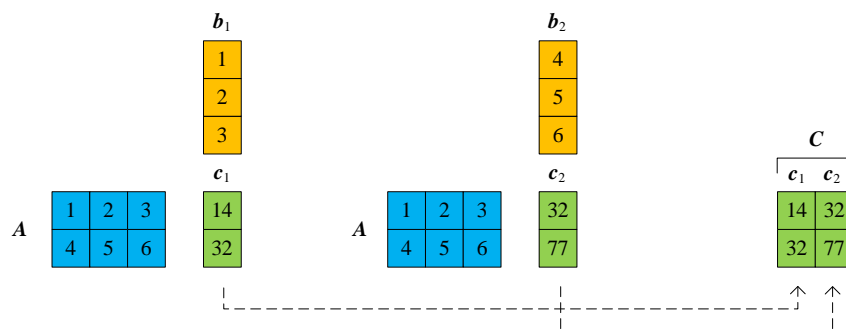


图 7. 把列向量 \mathbf{c}_1 、 \mathbf{c}_2 顺序摆放得到矩阵 \mathbf{C}

第二个例子

让我们用线性组合视角看第二个矩阵乘法的例子。代入具体值，计算 \mathbf{d}_1

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

$$\mathbf{d}_1 = \mathbf{B} @ \mathbf{a}_1 = 1 \times \mathbf{b}_1 + 4 \times \mathbf{b}_2 = 1 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 4 \times \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1+16 \\ 2+20 \\ 3+24 \end{bmatrix} = \begin{bmatrix} 17 \\ 22 \\ 27 \end{bmatrix} \quad (12)$$

计算 \mathbf{d}_2

$$\mathbf{d}_2 = \mathbf{B} @ \mathbf{a}_2 = 2 \times \mathbf{b}_1 + 5 \times \mathbf{b}_2 = 2 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 5 \times \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2+20 \\ 4+25 \\ 6+30 \end{bmatrix} = \begin{bmatrix} 22 \\ 29 \\ 36 \end{bmatrix} \quad (13)$$

计算 \mathbf{d}_3

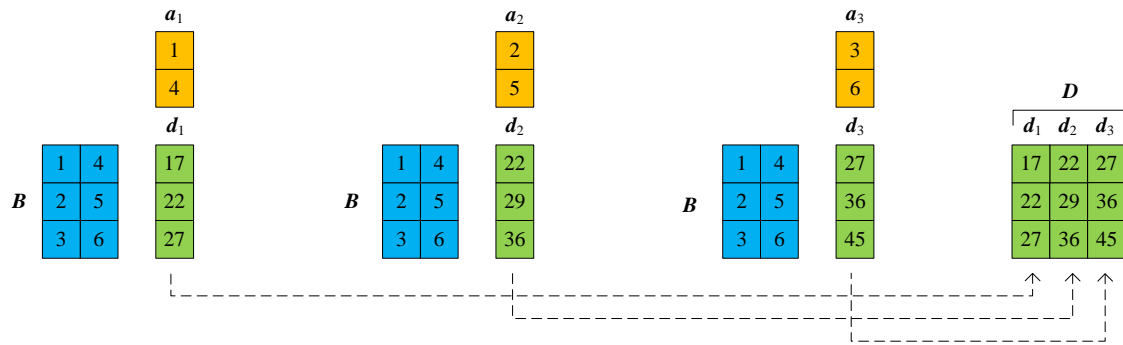
$$\mathbf{d}_3 = \mathbf{B} @ \mathbf{a}_3 = 3 \times \mathbf{b}_1 + 6 \times \mathbf{b}_2 = 3 \times \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 6 \times \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3+24 \\ 6+30 \\ 9+36 \end{bmatrix} = \begin{bmatrix} 27 \\ 36 \\ 45 \end{bmatrix} \quad (14)$$

$$\begin{array}{l} \begin{array}{c} a_{1,1} \\ 1 \end{array} \times \begin{array}{c} b_1 \\ 1 \\ 2 \\ 3 \end{array} + \begin{array}{c} a_{2,1} \\ 4 \end{array} \times \begin{array}{c} b_2 \\ 4 \\ 5 \\ 6 \end{array} = \begin{array}{c} d_1 \\ 17 \\ 22 \\ 27 \end{array} \\ \\ \begin{array}{c} a_{1,2} \\ 2 \end{array} \times \begin{array}{c} b_1 \\ 1 \\ 2 \\ 3 \end{array} + \begin{array}{c} a_{2,2} \\ 5 \end{array} \times \begin{array}{c} b_2 \\ 4 \\ 5 \\ 6 \end{array} = \begin{array}{c} d_2 \\ 22 \\ 29 \\ 36 \end{array} \\ \\ \begin{array}{c} a_{1,3} \\ 3 \end{array} \times \begin{array}{c} b_1 \\ 1 \\ 2 \\ 3 \end{array} + \begin{array}{c} a_{2,3} \\ 6 \end{array} \times \begin{array}{c} b_2 \\ 4 \\ 5 \\ 6 \end{array} = \begin{array}{c} d_3 \\ 27 \\ 36 \\ 45 \end{array} \end{array}$$

图 8. 通过线性组合计算 \mathbf{d}_1 、 \mathbf{d}_2 、 \mathbf{d}_3

如图 9 所示，把 \mathbf{d}_1 、 \mathbf{d}_2 、 \mathbf{d}_3 左右排列，便得到矩阵 \mathbf{D}

$$\mathbf{D} = \mathbf{B} @ \mathbf{A} = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \mathbf{d}_3] = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix} \quad (15)$$

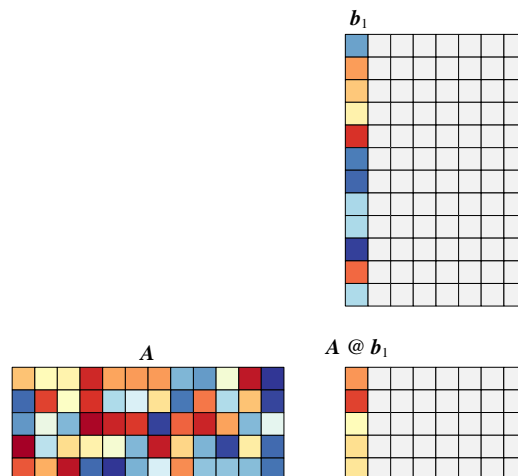
图 9. 把列向量 d_1 、 d_2 、 d_3 顺序摆放得到矩阵 D 

LA_03_03_01.ipynb 完成以上矩阵乘法第三视角计算，请大家自学。

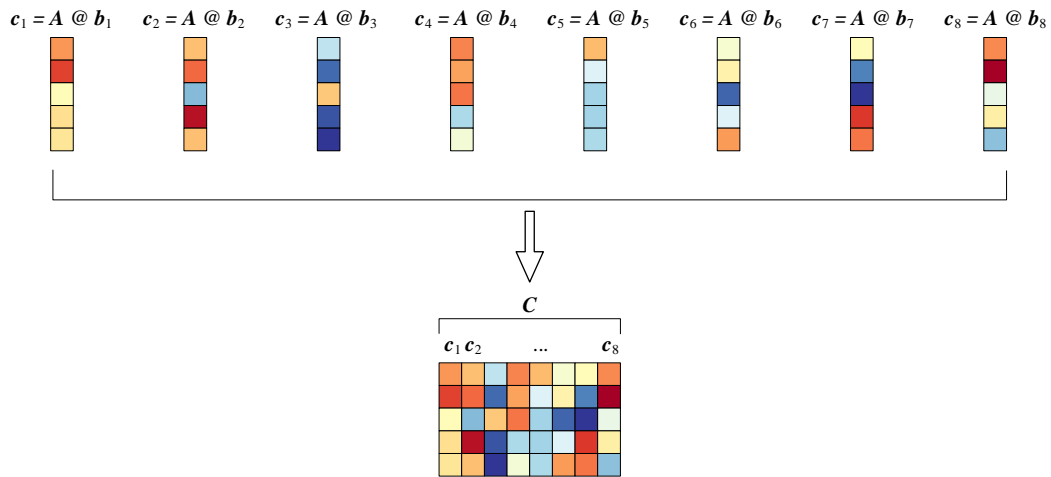
热图展示矩阵乘法第三视角

下面让我们用热图来可视化矩阵乘法第三视角。

图 10 用热图展示如何计算 $A @ b_1$ 。

图 10. 热图展示 $A @ b_1$

由于 B 有 8 列，类似 $A @ b_1$ 矩阵乘法一共有 8 个，结果如图 11 所示。它们的结果顺序排列便得到矩阵 $C = A @ B$ 。

图 11. 8 个 $A @ b_i$ 左右排列得到矩阵 $C = A @ B$ 

LA_03_03_02.ipynb 利用 `seaborn.heatmap()` 绘制上述热图可视化矩阵乘法第三视角，请大家自学。

自定义 Python 函数计算矩阵乘法：列向量线性组合视角

代码 1 通过自定义函数，利用列向量线性组合视角完成矩阵乘法。下面聊聊其中关键语句。

- a** 定义一个函数，叫 `col_LC`，意思是“列向量线性组合”。函数就是一段我们可以重复使用的代码。这个函数需要两个输入：一个矩阵 A 和一个系数列表 `coefficients`，最终会返回一个向量，表示用这些系数组合矩阵的每一列。
- b** 开始一个循环，从第 0 列一直到第 $p-1$ 列。`A[:, k]` 取出 A 的第 k 列；然后我们用对应的系数 `coefficients[k]` 乘这个列向量，再加到最终结果上。这个动作重复做多次，最后就得到了列向量线性组合的结果。
- c** 又定义了一个新函数，叫 `matrix_multiplication_col_LC`，它的作用是实现“矩阵乘法”，但方式比较特别：是通过“列向量组合”来实现的。输入是两个矩阵 A 和 B ，最后返回一个新矩阵 C ，表示 A 乘以 B 的结果。
- d** 开始一个循环，从第 0 列到第 $n-1$ 列，表示我们要一列一列地计算矩阵乘法结果 C 的列。`C[:, j] = col_LC(A, coeffs)` 调用前面写好的 `col_LC()` 函数，把 A 和当前这列系数传进去。函数会帮我们把 A 的每列乘上对应的系数加在一起，结果就是新矩阵 C 的第 j 列。

代码 1. 自定义 Python 函数计算矩阵乘法，列向量线性组合视角 | LA_03_03_03.ipynb

```

## 初始化
import numpy as np

## 定义列向量线性组合函数
def col_LC(A, coefficients):

    m, p = A.shape
    col_LC_result = np.zeros(m) # 初始化组合结果

    for k in range(p):
        col_LC_result += coefficients[k] * A[:, k]
        # A (左侧) 每列乘系数并相加
    return col_LC_result

## 定义矩阵乘法函数，列向量组合
def matrix_multiplication_col_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 j in range(n):
        coeffs = B[:, j]
        # 取出 B 的第 j 列，作为线性组合的系数
        C[:, j] = col_LC(A, coeffs)
        # 计算 A 的列的线性组合

    return C

## 矩阵乘法

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

## 矩阵乘法
matrix_multiplication_col_LC(A, B)
matrix_multiplication_col_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

$$\blacktriangleright \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\blacktriangleright \mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\blacktriangleright \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\blacktriangleright \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Q2. 用矩阵乘法第三视角展开如下成对矩阵乘法

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

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

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

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

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。

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Q5. 把以下三个矩阵乘法写成一个，并给出结果。

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} @ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Q6. 比较 Q4、Q5 的结果，解释两者结果的异同。

Q7. 请自学置换矩阵。

Q8. 一只鸡一个头、两只脚；一只兔一个头、四只脚。笼子里现有，5 只鸡、8 只兔，请利用矩阵乘法计算整个笼子里的头数、脚数。