

矩阵 $A \in \mathbb{R}^{N \times N}$ ，那么 $d(Ax)dx$ 是什么

回答：

此题涉及到对向量求导。

逐项展开分析为：向量求导与标量求导法则不同的是，向量的求导还要注意结果的排法问题。注意排法是因为当一个目标函数是多个成分相加时，如果不注意排法可能导致有些结果是行，有些是列，无法继续进行运算。

所以向量求导的基本推导准则是（以 $f(x)$ 对向量 $x = (x_1, \dots, x_n)^T$ 求导为例）：

1. 先对 $f(x)$ 求向量分量的导数，也就是标量 x_i 的导数，该导数可能是标量，也可能是向量；
2. 如果第一步求导结果为标量，直接将结果排列成 x 的形状（本例为列）；如果第一步求出的结果是向量，在按照 x 排列结果时：

如果求导结果与 x 形状相同，则将结果取转置排成矩阵；

如果求导结果与 x 形状不同，则直接将结果排成矩阵。

向量 x 是列向量， $A = [a_1, a_2, \dots, a_n]$ 是矩阵， $A^T = [a'_1, \dots, a'_n]$ 是 A 的转置

$$\text{求分量导数 } \frac{\partial(Ax)}{\partial x_k} = \frac{\partial \sum_i a_i x_i}{\partial x_k} = a_k$$

排列：求导结果与 x 形状相同，将 a_k 取转置按列排好，即 A 的转置：
$$\frac{\partial Ax}{\partial x} = A^T$$