向量偏导意义：

假设 W 是一个 $n\times1$ 向量：$W=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$，一个以 W 为自变量函数为 $L(W)$（注意：W 是向量，$L(W)$ 是一个数）。

定义：

$$\frac{\partial L}{\partial W}=\begin{bmatrix} \frac{\partial L}{\partial w_1} \\ \frac{\partial L}{\partial w_2} \\ \vdots \\ \frac{\partial L}{\partial w_n} \end{bmatrix}$$

，（注意：$\frac{\partial L}{\partial W}$ 是一个向量）

有了这个定义，我们可以求偏导了。

① 证明：$\frac{\partial\left[\frac{1}{2}\|W\|^2\right]}{\partial W}=W$

证明：假设 $W=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$，则有：

$$\frac{1}{2}\|W\|^2=\frac{1}{2}\left(w_1^2+w_2^2\cdots+w_n^2\right)$$

根据定义：

$$\frac{\partial\left(\frac{1}{2}\|W\|^2\right)}{\partial W}=\begin{bmatrix} \frac{\partial\left(\frac{1}{2}\|W\|^2\right)}{\partial w_1} \\ \frac{\partial\left(\frac{1}{2}\|W\|^2\right)}{\partial w_2} \\ \vdots \\ \frac{\partial\left(\frac{1}{2}\|W\|^2\right)}{\partial w_n} \end{bmatrix}=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}=W$$

其它情况自己试着推导吧。