Exercise 3.1 Derivatives

$f : \mathbb{R}^N \to \mathbb{R}^M$
$g : \mathbb{R}^{M \times M} \to \mathbb{R}$
$h : \mathbb{R}^N \to \mathbb{R}, \qquad h(x) := g(f(x), f(x))$

using the information given on slide DL3.15/69:

$$f : \mathbb{R}^N \to \mathbb{R}^M, \quad g : \mathbb{R}^M \to \mathbb{R}^K, \quad h = g \circ f, \qquad \text{then}$$

$$Dh(x) = Dg(f(x))Df(x)$$

for $K = 1$:

$$\nabla h(x)^T = \nabla g(f(x))^T Df(x)$$

using the information given on slide DL3.22/69:

"the gradients being propagated back can just be added"

$$\Rightarrow \nabla h(x)^T = \nabla g(f_1(x))^T Df_1(x) + \nabla g(f_2(x))^T Df_2(x)$$

since $f_1 = f_2$

$$\nabla h(x)^T = 2\nabla g(f(x))^T Df(x)$$

Exercise 3.2 Backpropagation