

activations

$x$

$$\frac{\partial L}{\partial x}$$

$$= \frac{\partial L}{\partial z} \frac{\partial z}{\partial x}$$

“local gradient”

$$\frac{\partial z}{\partial x}$$

$f$

$$\frac{\partial z}{\partial y}$$

$y$

$$\frac{\partial L}{\partial y}$$

$$= \frac{\partial L}{\partial z} \frac{\partial z}{\partial y}$$

$z$

$$\frac{\partial L}{\partial z}$$

gradients

