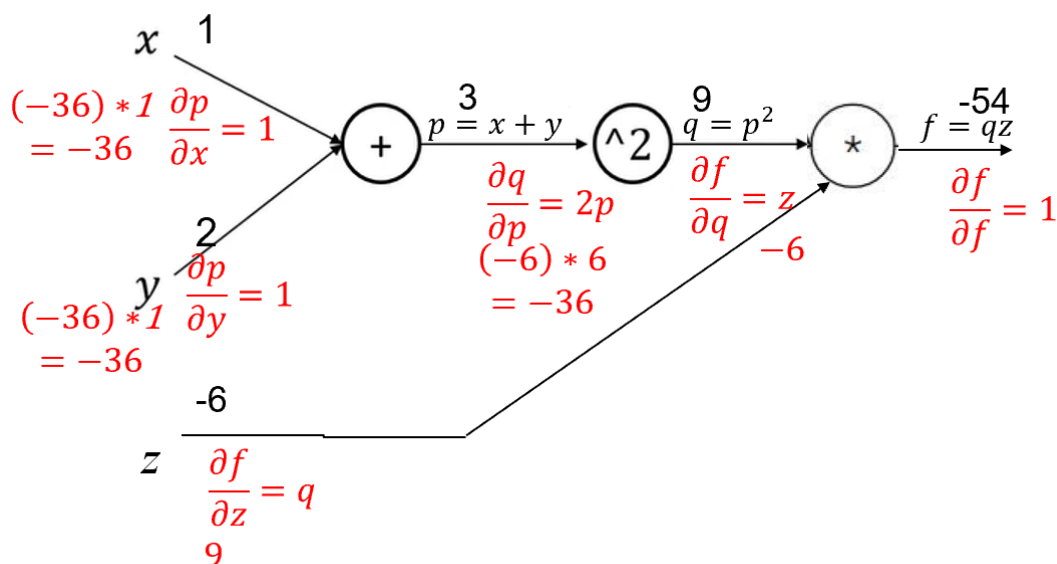


3, 画出  $(x+y)^2 z$  的计算图, 当  $x=1, y=2, z=-6$  时, 写出前向传播的数值和反向传播的梯度值。

解:



4, 计算 Sigmoid 函数、双曲正切函数和 ReLU 函数的导数函数, 分析这三个函数作为激活函数时的优缺点。

解: (1) 对于 Sigmoid 函数,  $\theta(x) = \frac{1}{1+e^{-x}}$

$$\frac{\partial \theta(x)}{\partial x} = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1+e^{-x}-1}{1+e^{-x}} \frac{1}{1+e^{-x}} = (1-\theta(x))\theta(x)$$

Sigmoid 函数作为激活函数的优点是连续、单调、可导且具有非线性特点。但其缺点是非中心对称, 输出不是 0 均值的, 同时它的导数函数曲线见右图,  $x$  变化很小的范围内, 导数才有值且不大, 在神经

