→ gradient descent로 연산 (weight update)

bias

↳ z

w_{11}

w_{31} w_{21} w_{12}

w_{13}

w_{14}

w_{11}
w_{21} . . . bias
w_{31}
w_{41}

태그 단위에 관련지 보기

버선 활용할 수

H

→ feature

x_1 x_2 x_3 x_4

→ 데이터 1개

데이터 개수

→ dimension

e → activate function 적용 후
다음 전향하기까지 결정.

dimension

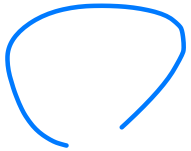
e → number of
Hidden layer

linear → non linear

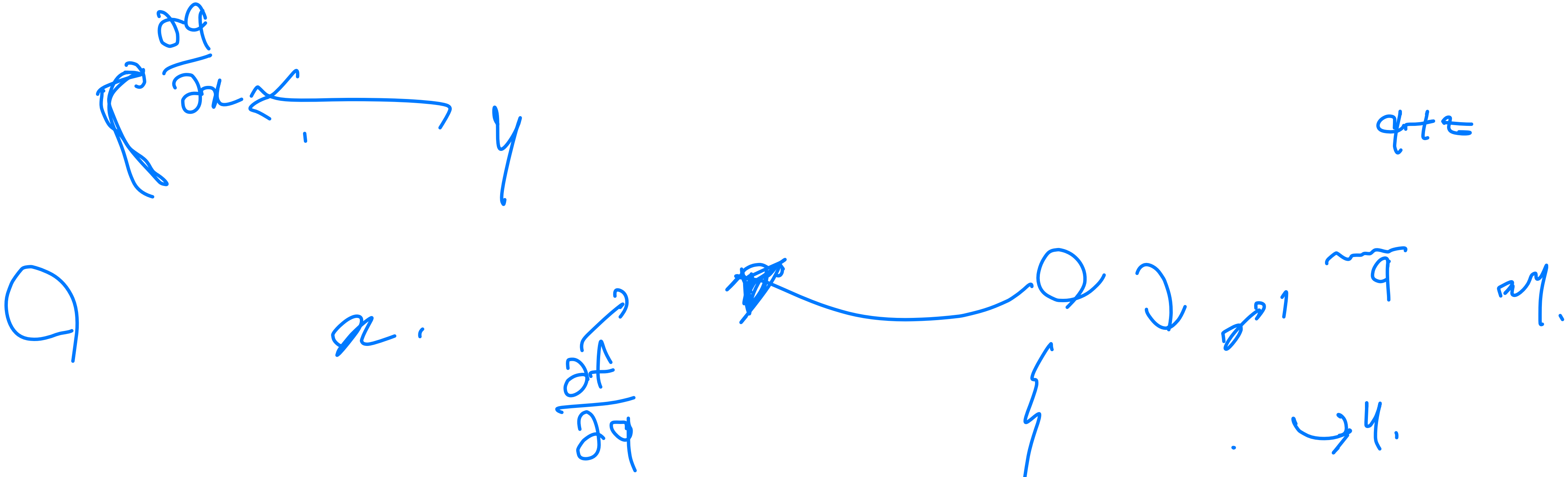
$z=0$ or ∞

→ 양 끝 값에서 gradient가 거의
없기 때문에

맨 끝에 달려서 binary classification 문제에서만 사용.



↳ Lemma. cost가 굉장히 높아진다.


$\frac{\partial q}{\partial x}$
y
q+z
q
x
$\frac{\partial f}{\partial q}$
q
xy
y

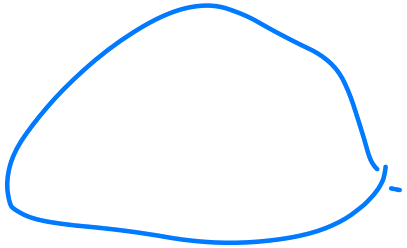
$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial q}\frac{\partial q}{\partial z} = 1 \times 1$$

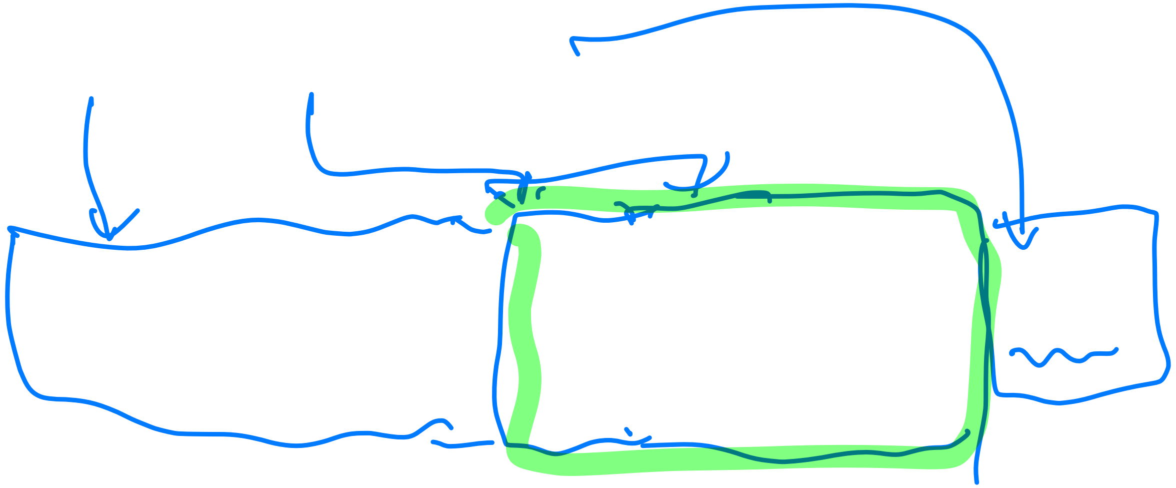
bias

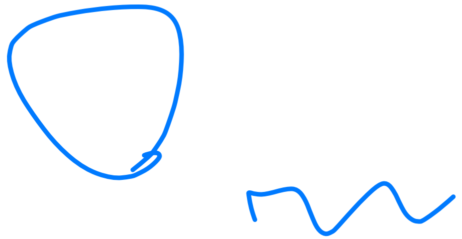
θ_2 $z^{(3)}$

y

activate func



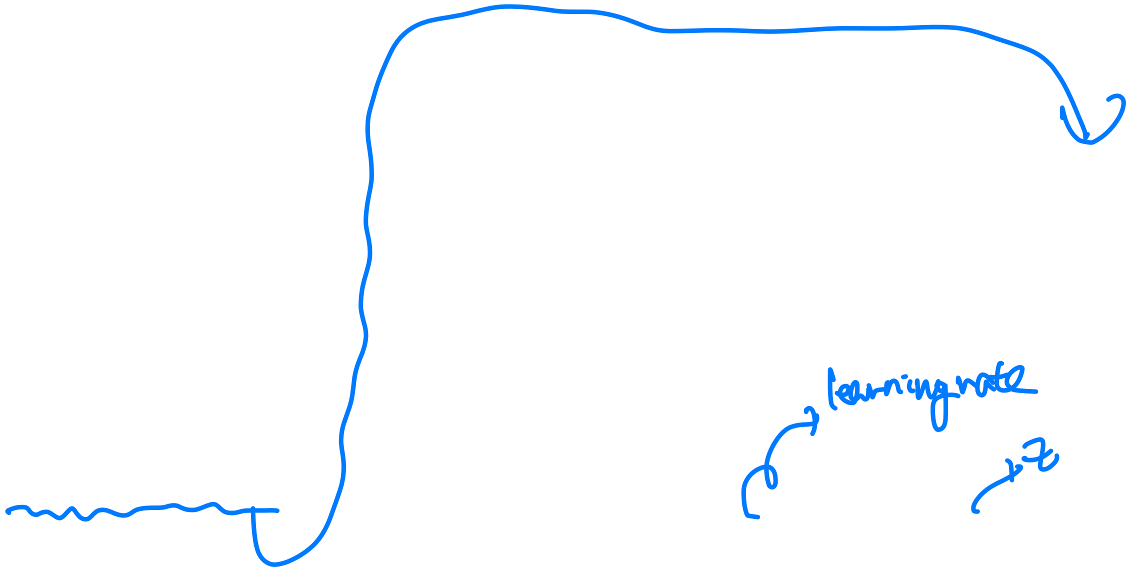




$z(4-z)(1-z)$

0.332

0.525



$$w_{ij} := w - \alpha \frac{\partial E}{w_{jk}}$$

learning rate

z