bowl

well-Conditioned

boat

Ill Conditioned.

$$\eta_t = G$$

$$\eta_t = \frac{c}{t} \qquad \frac{c}{\sqrt{t}}$$

Direction 1 : Safe to move (grad is small)

Direction 2 : Risky to move (grad is large)

Neural Network

$$(\vec{x} \cdot \vec{w} + b) \xrightarrow{v_1}$$
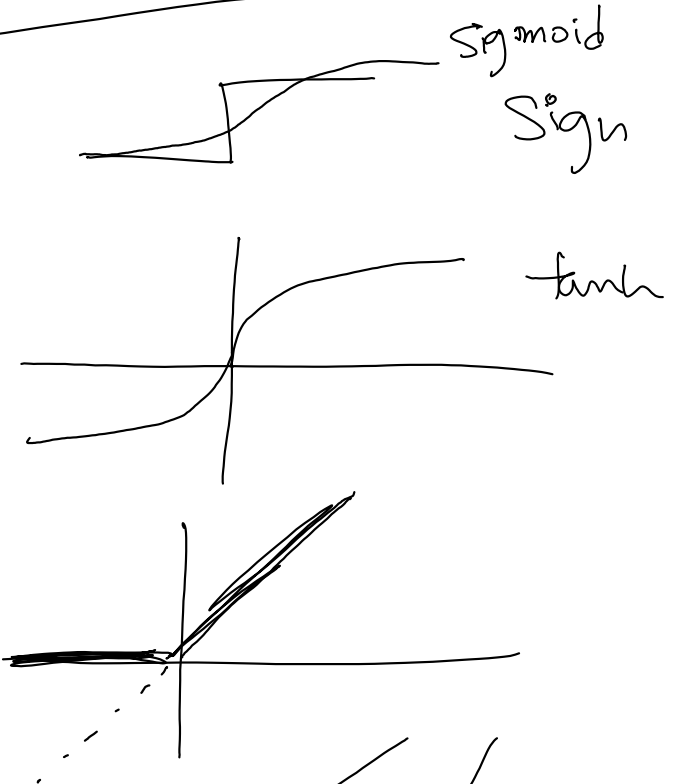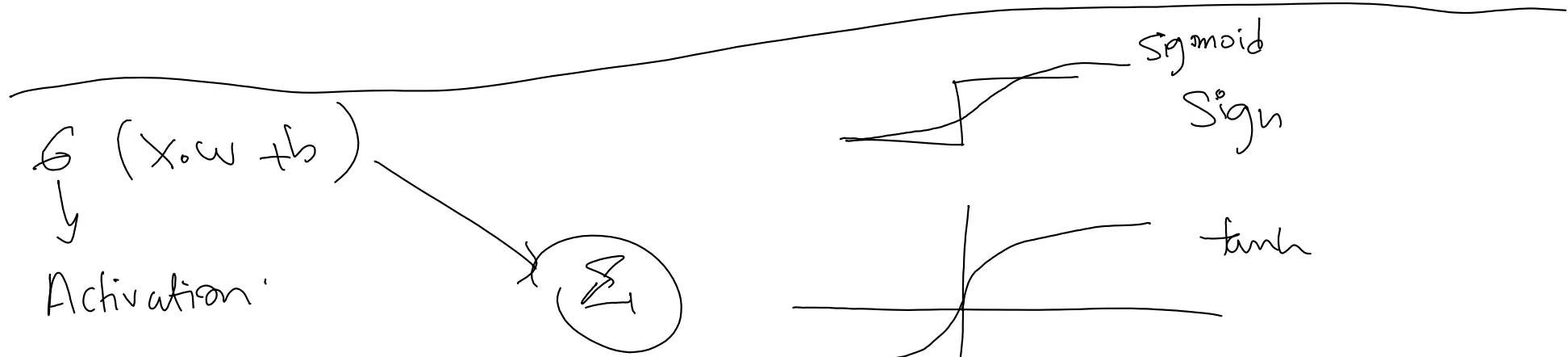
$$(\vec{x} \cdot \vec{w}' + b') \xrightarrow{v_2}$$

$$v_1 (\vec{x}_o \cdot \vec{w} + b) + v_2 (\vec{x}_o \cdot \vec{w}' + b')$$

$$= \vec{x}_o \cdot (v_1 \vec{w} + v_2 \vec{w}') + (v_1 b + v_2 b')$$
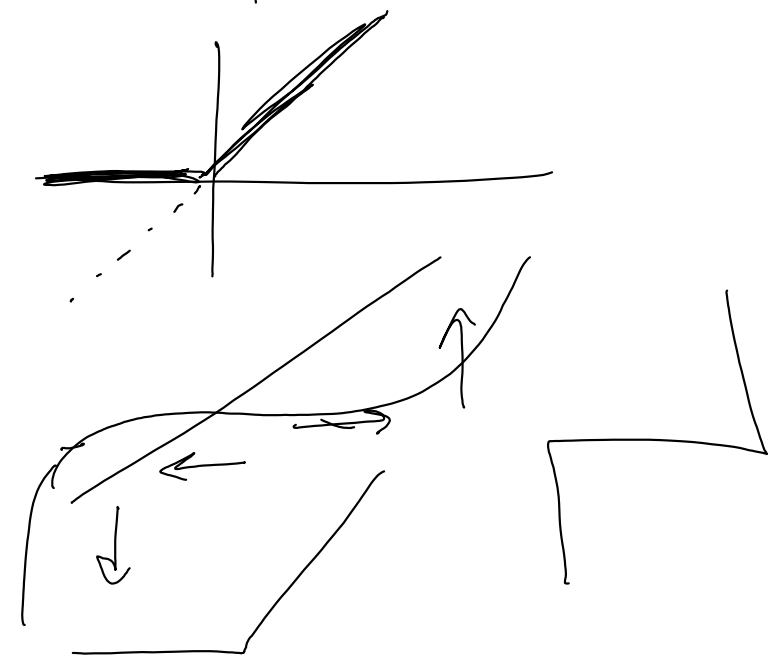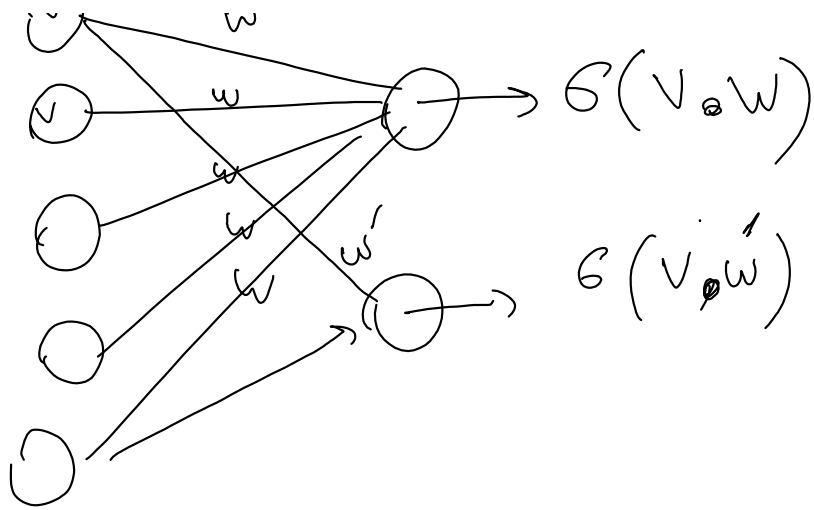
Need for non-linearity

$\vec{w}$" b" linear

$\sigma(X \cdot w + b)$
↓
Activation

ε₁

Sigmoid
Sign

tanh

$gd = 0$

Sign

∇ → allows to optimize.

$\sigma(V_\bullet W)$

$\sigma(V_\bullet \acute{W})$

XOR example

$\omega = (1,1)$

(0.1)

$b' = (0,1)$

$\text{Sign}(-x_1 - x_2 + 0.1) * 1$

$(-1,-1)$

$\text{Sign}(x_1 + x_2 + 0.1) \quad * 1$