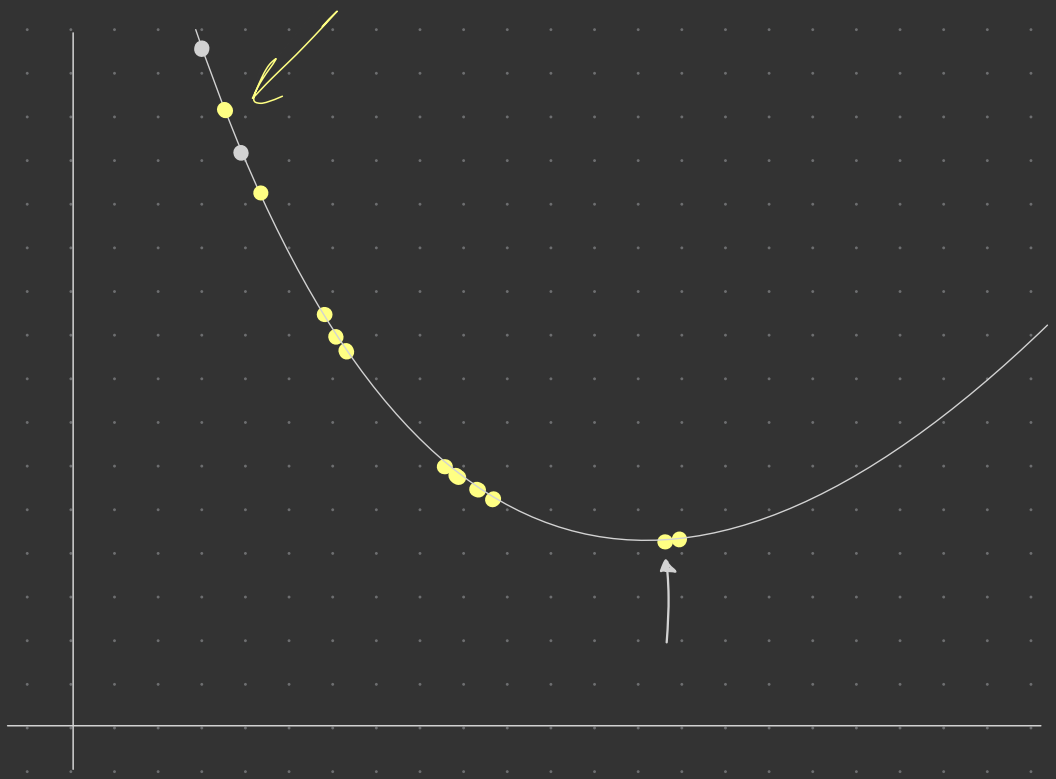
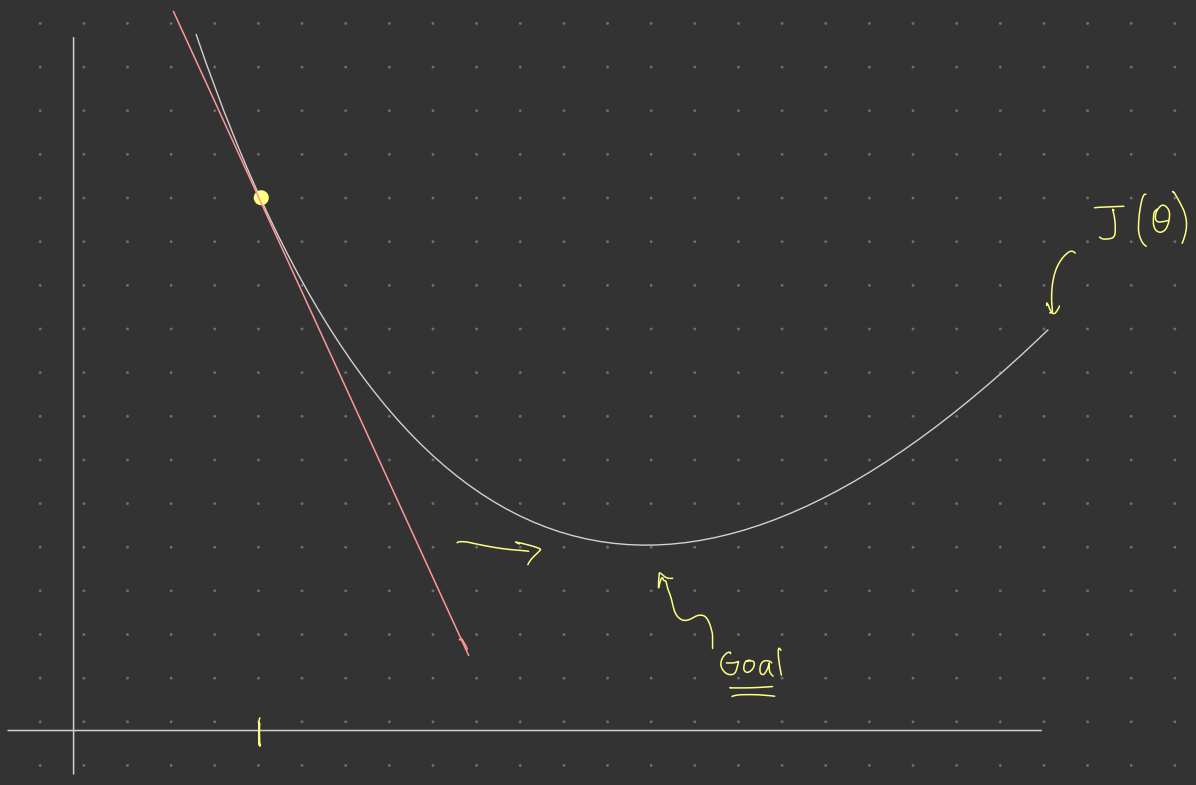
Optimization

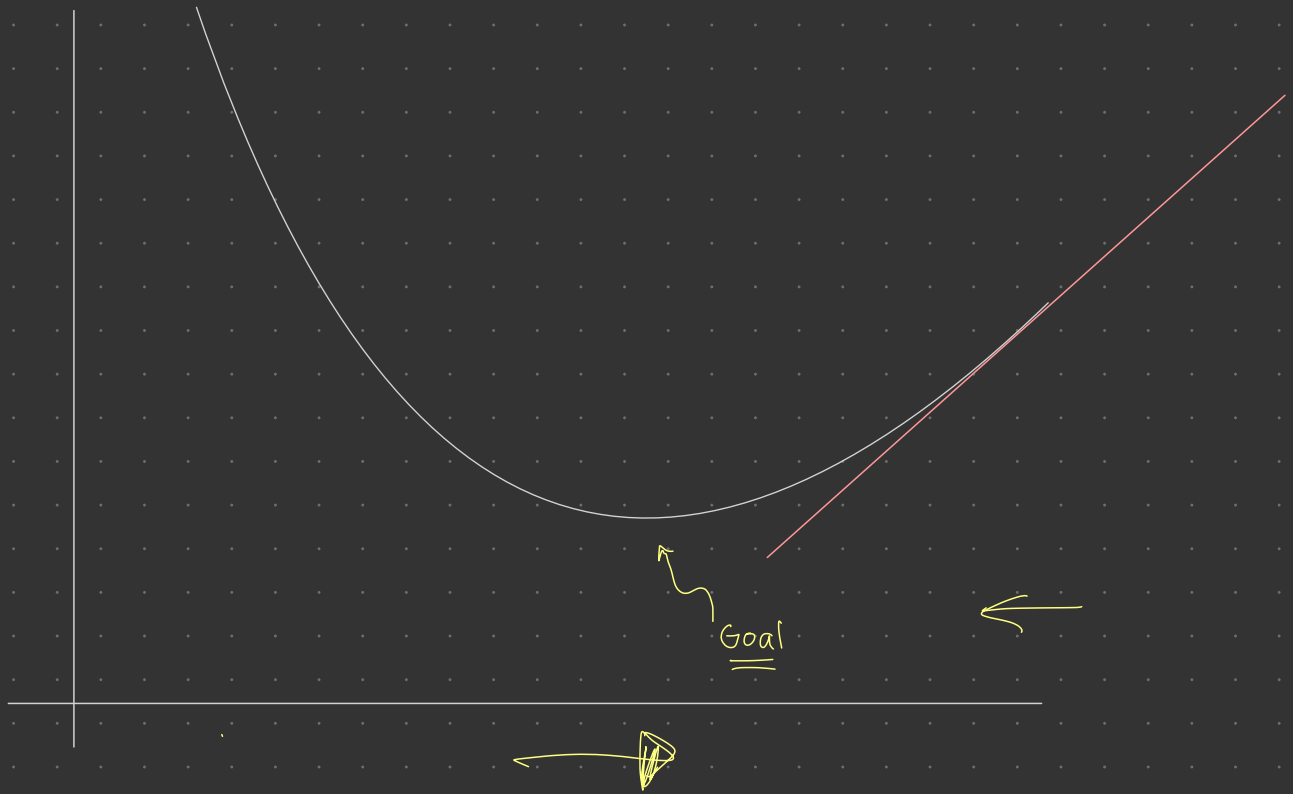




- Start with random point
- Compute derivative $\frac{d}{dx} J(\theta)$
- if derivative is negative, increase x

$$x := x - \left[\frac{d}{dx} J(x) \right]$$

why?



- if derivative is positive, decrease x

$$x := x - \left[\frac{d}{dx} J(\theta) \right]$$

why?

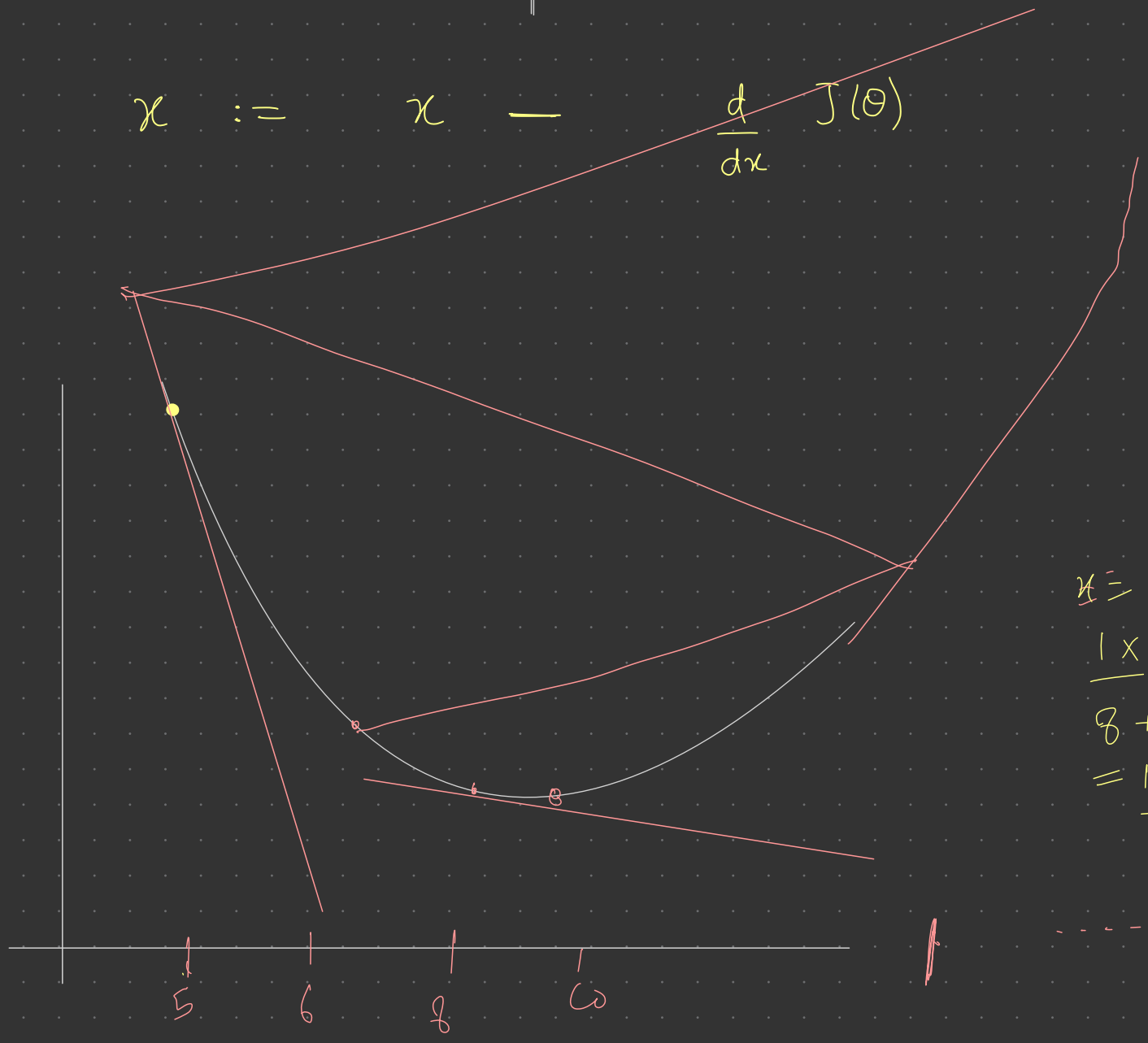
- if derivative is negative, increase x

$$x := x - \left[\frac{d}{dx} J(\theta)\right]$$

- if derivative is positive, decrease x

$$x := x - \left[\frac{d}{dx} J(\theta)\right]$$

$$x := x - \frac{d}{dx} J(\theta)$$



$x = 8$

1×2

$8 + 2$

$= 10$

$6 \leftarrow 5 + 1$

$$x := x - \underbrace{0.001}_{\alpha} \frac{d}{dx} J(\theta)$$

learning rate (hyperparam)