The background features a dark gray base with several wavy, organic shapes in shades of brown and gold. A prominent grid pattern is visible in the upper left corner, transitioning into the wavy lines. The overall aesthetic is modern and technical.

Machine Learning Specialization

The Cost Function + Gradient Descent

$f_w(x)$ = function of x , fixed w

$J(w)$ = function of w

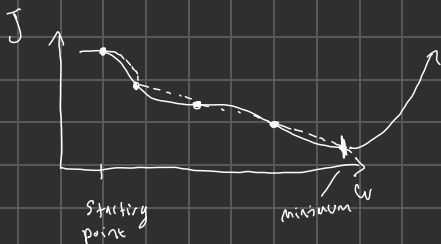
$$J(w) = \frac{1}{2M} \sum_{i=1}^M (f_w(x^{(i)}) - y^{(i)})^2$$

$$= \frac{1}{2M} \sum_{i=1}^M (w x^{(i)} - y^{(i)})^2 = \frac{1}{2M} (0^2 + 0^2 + 0^2)$$

min $J(w, b)$ Minimize the cost fn

min $J(w_1, w_2, \dots, w_n, b)$
 w_1, w_2, \dots, w_n, b

- Start with some w, b , then iteratively change values to minimize the cost fn



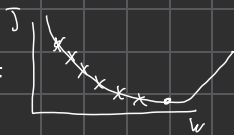
Main update formula: $w = w - \alpha \frac{d}{dw} J(w, b)$

$x, y \in \mathbb{R}$

learning rate α

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b)$$

α is small: then lots of small updates



fail to converge and even diverge

α is big, lots of big updates

