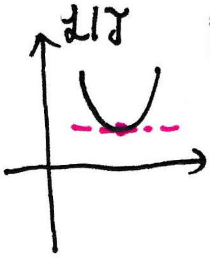
09/11/2024

Pg 13. $\frac{\partial J}{\partial w_j} = \frac{\partial}{\partial w_j} \frac{1}{N} \sum_i^N \mathcal{L}_i = \frac{1}{N} \cdot \sum_i^N (y^i - t^i) x_j^i = 0$ ← at critical point

$\frac{\partial J}{\partial b} = \frac{\partial}{\partial b} \frac{1}{N} \sum_i^N \mathcal{L}_i = \frac{1}{N} \cdot \sum_i^N (y^i - t^i) = 0$

Basically, we want to reach a point such that, you have a w_j and a b (the parameters) and the partial derivatives / slopes wrt w_j & b are 0.



partial derivative wrt a parameter is 0 ⇒ you have reached a critical point } ⇒ you have achieved a parameter that minimizes the loss ⇒ bingo!

if $\frac{\partial J}{\partial w_j} \neq 0$ ⇒ Not at critical pt ⇒ can always improve by changing w_j

(Similarly, $\frac{\partial J}{\partial b} \neq 0$ ⇒ could improve by changing b)

Pg 14. $\nabla f(w)$ = gradient

$= \frac{\partial}{\partial w_j} f(w)$. When $f = J$, your gradient is $\frac{\partial J}{\partial w_j}$

* we also went over several topics covered in the previous Notes. Pls see the previous Notes for reference.