

$$\hat{y} = y(\omega) = \sum_{i=1}^{m} b_i \sigma(\alpha; x)$$

$$L(w) = \frac{1}{2} ||y - y(w)||^2$$
 $O(\frac{1}{m})$

$$\frac{\partial \hat{y}}{\partial a_i} = b_i \sigma'(a_i x) \cdot x$$

$$\frac{\partial^2 \hat{y}}{\partial a_i \partial a_j} = \begin{cases} b_i x^2 \sigma''(a_i x) & i=j \\ 0 & 0.w. \end{cases}$$

$$\frac{\partial \dot{y}}{\partial b_i} = \sigma(o_i \times) \qquad \frac{\partial^2 \dot{y}}{\partial b_i \partial b_j} = 0$$

$$\frac{\partial^2 \hat{y}}{\partial \alpha_i \partial b_j} = \begin{cases} \sigma'(\alpha_i x) \cdot x & i=j \\ 0 & o.w. \end{cases}$$

$$\frac{\partial^2 \hat{y}}{\partial \alpha_i \partial b_j} = \begin{cases} \sigma'(\alpha_i x) \cdot \alpha_i & i=j \\ 0 & o.w. \end{cases}$$

Hessian =

.

spectral norm O(1)

E 11 / w y (w) 11 = m

 $\mathbb{E} \| \nabla_{w}^{2} y(w) \| = O(1)$

