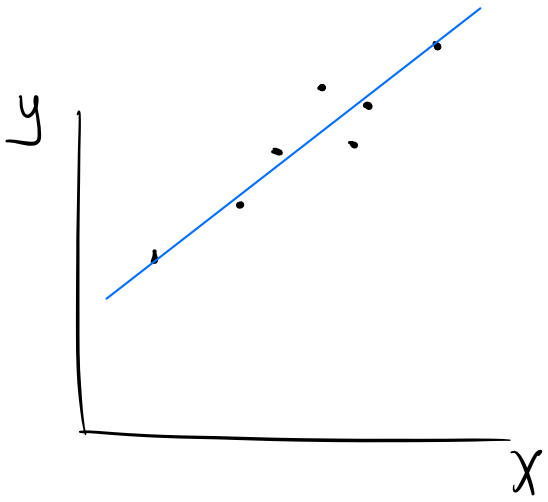
Model: $\hat{y} = w x_i + b$



Loss: $\frac{1}{2n} \sum (y_i - \hat{y}_i)$

More y's?

$y^1 = w_1 x_i + b_1$

$y^2 = w_2 x_i + b_2$

$\vec{y} = \vec{w} \; x + \vec{b}$

$[\;] = [\;] x + [\;]$

More x's?

$y = \vec{w}^T \vec{x} + b$

$\square = \square\!\!\square + \square$

More everything?

$\vec{y} = W^T \vec{x} + \vec{b}$

$10 \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 64 \\ 10 \; W^T \end{bmatrix}_{64} + \begin{bmatrix} \vec{b} \\ 10 \end{bmatrix}$

$$\hat{y} = \begin{bmatrix} W^T \end{bmatrix} \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix} \qquad \vec{\hat{y}} = W^T \vec{x}$$

Softmax $\sigma(\vec{z})_i = \dfrac{e^{z_i}}{\sum_j e^{z_j}}$

multiclass classification

$\rightarrow \begin{bmatrix} 5 \\ 2 \\ -8 \end{bmatrix} \rightarrow \begin{bmatrix} .6 \\ .3 \\ .1 \end{bmatrix}$

$c \times h \quad h \times d$

$$\hat{y} = W_2 \sigma (W_1 \vec{x})$$

$$W_2 W_1 = W \quad (c \times d)$$

activation function