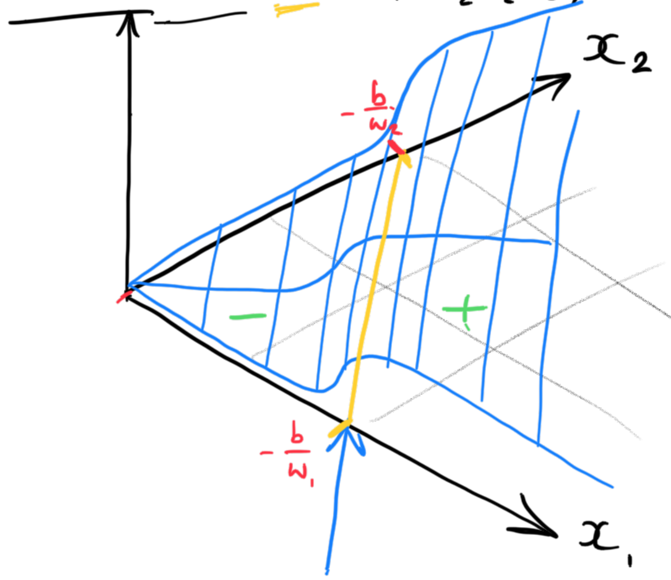
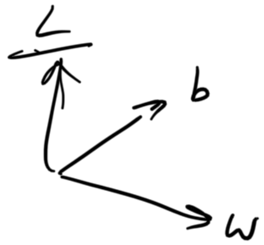
$$h(x) = \sigma(z) = \sigma(\overbrace{w_1x_1 + w_2x_2 + b}^{\text{logit}})$$

$$w_1\cancel{x_1} + w_2x_2 + b = 0$$

if $x_1 = 0$ $\quad w_2x_2^{*} + b = 0$

$$x_2^{*} = -\frac{b}{w_2}$$





Linear regression

$$\hat{y} = XW + b$$

$$L_{MSE} = \mathbb{E}(\hat{y} - y)^2$$

$$\frac{\partial L}{\partial W} = 2\mathbb{E}(XW + b)X$$

$$\frac{\partial L}{\partial b} = 2\mathbb{E}(XW + b)$$

$$W \leftarrow W - \alpha\frac{\partial L}{\partial W}$$

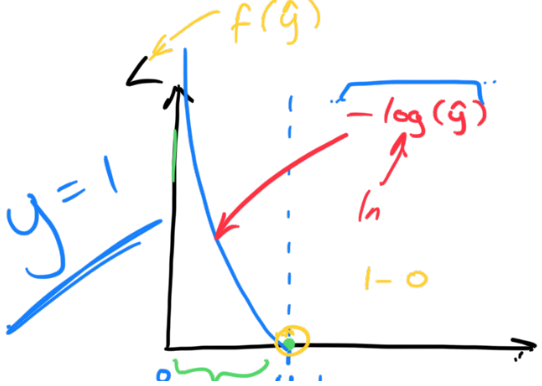
Logistic regression

$$\hat{y} = \sigma(XW + b)$$

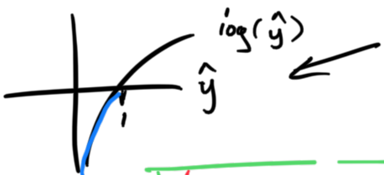
$$L_{BCE} = -[y\log(\hat{y}) + (1-y)\log(1-$$

$$L(\hat{y} @ y=1)$$

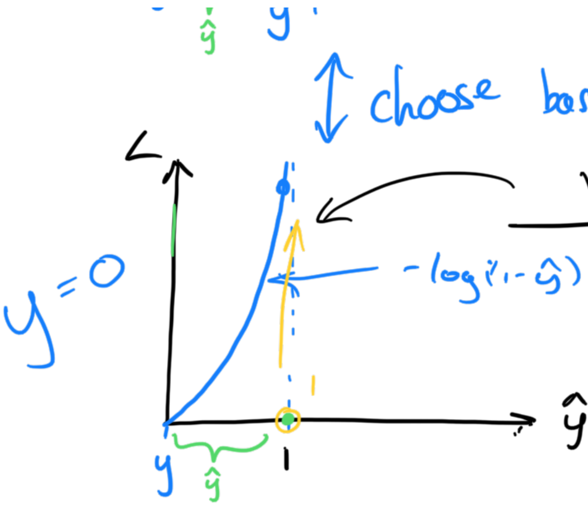
$$\neq 1 - L$$

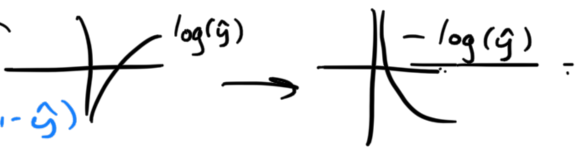


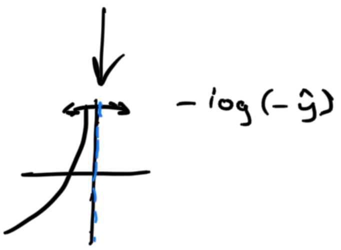
$\hat{y} = \sigma(z) \in (1, 0)$



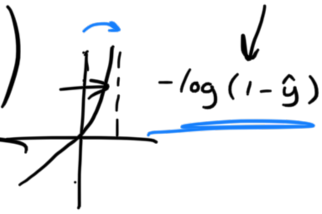
$\log_n(x)$ = the power o[f] n which gives x







$$L_{BCE} = -\underbrace{\Big(\overbrace{y\log(\hat y)}^{\text{if } y=1} + \overbrace{(1-y)\log(1-\hat y)}^{\text{if } y=0}\Big)}$$



$$\underbrace{x \xrightarrow{xw+b} z} \xrightarrow{\sigma} \hat y \xrightarrow{BCE} L$$

$$L = -(y\log\hat y + (1-y)\log(1-\hat y)) \leftarrow \qquad \frac{\partial L}{\partial w}$$

$$\hat y = \sigma(z)$$

$$z = xw + b$$

$$\boxed{\frac{dy}{dx} = \frac{dy}{dz} \times \frac{dz}{dx}}$$

$y = f($

$z = f($

chain rule

$$\frac{\partial L}{\partial w} = \underbrace{\frac{\partial L}{\partial \hat y}}_{-\left(y \times \frac{1}{\hat y} - (1-y)\frac{1}{1-\hat y}\right)} \underbrace{\frac{\partial \hat y}{\partial z}}_{\sigma(z)(1-\sigma(z))} \underbrace{\frac{\partial z}{\partial w}}_{x} \in \mathbb{R}^n$$

$$\frac{d}{dz}\log(z) = \frac{1}{z}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial \hat y}\frac{\partial \hat y}{\partial z}\frac{\partial z}{\partial b} \in \mathbb{R}$$