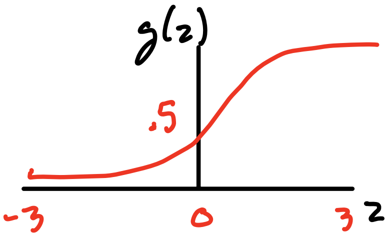
Logistic Regression

Sigmoid function: $g(z) = \frac{1}{1+e^{-z}}$



$$z = \vec{w} \cdot \vec{x} + b$$

$$f_{w,b}(\vec{x}) = g(\vec{w} \cdot \vec{x} + b) = \frac{1}{1+e^{-(\vec{w} \cdot \vec{x} + b)}}$$

Decision boundary:

ex) Is $f_{w,b}(\vec{x}) \geq 0.5$?

Yes: $\hat{y} = 1$ No: $\hat{y} = 0$

when is $f_{w,b}(\vec{x}) \geq 0.5$?

$$g(z) \geq 0.5$$

$$z \geq 0$$

$\vec{w} \cdot \vec{x} + b \geq 0$ $\vec{w} \cdot \vec{x} + b < 0$

Cost Function: Define log loss function $L(f_{w,b}(x^{(i)}), y^{(i)})$

$$L(f_{w,b}(x^{(i)}), y^{(i)}) = \begin{cases} -\log(f_{w,b}(x^{(i)})) & \text{if } y^{(i)} = 1 \\ -\log(1 - f_{w,b}(\vec{x}^{(i)})) & \text{if } y^{(i)} = 0 \end{cases}$$

$$J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^{m} L(f_{w,b}(x^{(i)}), y^{(i)})$$

Simplified Loss Function:

$$L(f_{w,b}(x^{(i)}), y^{(i)}) = -y^{(i)} \log(f_{w,b}(x^{(i)})) - (1-y^{(i)}) \log(1 - f_{w,b}(x^{(i)}))$$

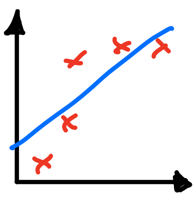
Simplified Cost Function:

$$J(w,b) = -\frac{1}{m} \sum_{i=1}^{m} [y^{(i)} \log(f_{w,b}(x^{(i)})) + (1-y^{(i)}) \log(f_{w,b}(x^{(i)}))$$

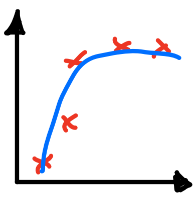
Gradient Descent for Logistic Regression

repeat
$$\begin{cases} w_j = w_j - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right) x_j^{(i)} \right] \\ b = b - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right) \right] \end{cases}$$

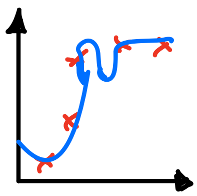
Overfitting



underfit
high bias



well fit
generalization



overfit
high variance

Addressing Overfitting:

- collect more training examples
- feature selection
- regularization (reduce size of params w_j)

Cost Function w/ Regularization

$$J(\vec{w}, b) = \frac{1}{2m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right)^2 + \frac{\lambda}{2m} \sum_{j=1}^{n} w_j^2$$

$$J(\vec{w}, b) = -\frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log\left(f_{w,b}(x^{(i)})\right) + (1 - y^{(i)}) \log\left(1 - f_{w,b}(x^{(i)})\right) + \frac{\lambda}{2m} \sum_{j=1}^{n} w_j^2 \right.$$

Regularized Linear Regression

Gradient Descent

simul update

$$\begin{cases} w_j = w_j - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)} \right) x_j^{(i)} + \frac{\lambda}{m} w_j \right] \\ b = b - \alpha \left[\frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right) \right] \end{cases}$$

Regularized Logistic Regression

Gradient Descent

simul update

$$\begin{cases} w_j = w_j - \alpha [\frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}(\vec{x}^{(i)}) - y^{(i)} \right) x_j^{(i)} + \frac{\lambda}{m} w_j \\ b = b - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(f_{w,b}(x^{(i)}) - y^{(i)} \right) \end{cases}$$