Cost function with Regularizarion

small numbers $w_1, w_2, \cdots, w_n, b$

Simpler model
less likely to overfit

size x_1	bedrooms x_2	floors x_3	age x_4	Avg income x_5	...	distance to coffee shop x_{100}	price y	
						$n = 100$		

n features

$$w_1, w_1, w_2, \cdots, w_{100}, b$$

Regularization term

can include or exclude b

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

$\lambda \to$ lambda
regulization parameter $\lambda > 0$

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

Regulized linear regression

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

repeat {
$j = 1 \cdots n$

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} J(\vec{w}, b) \qquad \frac{\partial}{\partial w_j} J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^{m} (f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}) x_j^{(i)} + \frac{\lambda}{m} w_j$$

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

} simultaneous updates

Don't have to regulaiize b

Regulized logistic regression

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

$$\min_{\vec{w},b} J(\vec{w}, b) = \quad \to w_j$$

repeat {
$j = 1 \cdots n$

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} J(\vec{w}, b) \qquad \frac{\partial}{\partial w_j} J(\vec{w}, b) = \frac{1}{m} \sum_{i=1}^{m} (f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}) x_j^{(i)} + \frac{\lambda}{m} w_j$$

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

} simultaneous updates