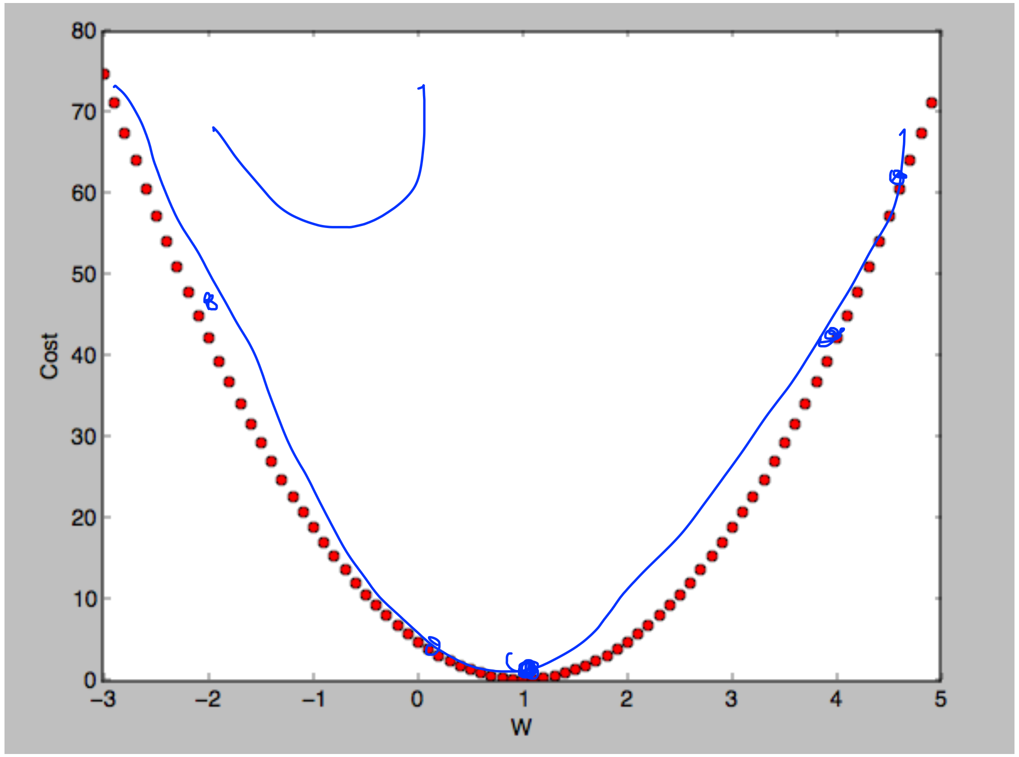
Lecture 5-2

Logistic (regression) classification: cost function & gradient decent

Sung Kim <hunkim+mr@gmail.com>

Cost

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2 \quad \text{when} \quad \underline{H(x) = Wx + b}$$



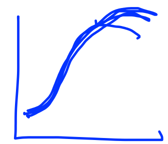
Cost function

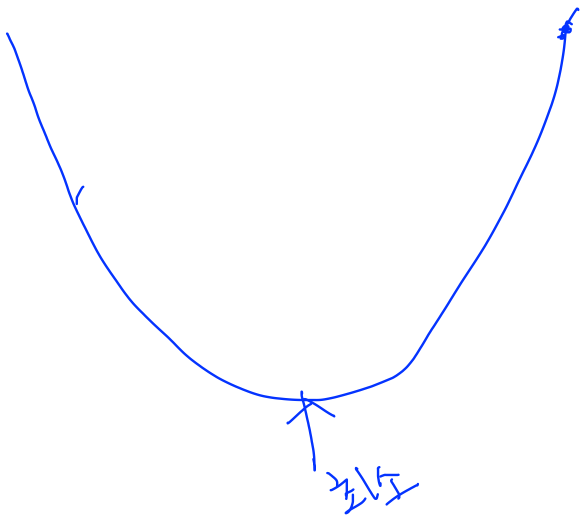
$$cost(W, b) = \frac{1}{m} \sum_{i=1}^{m} (H(x^{(i)}) - y^{(i)})^2$$

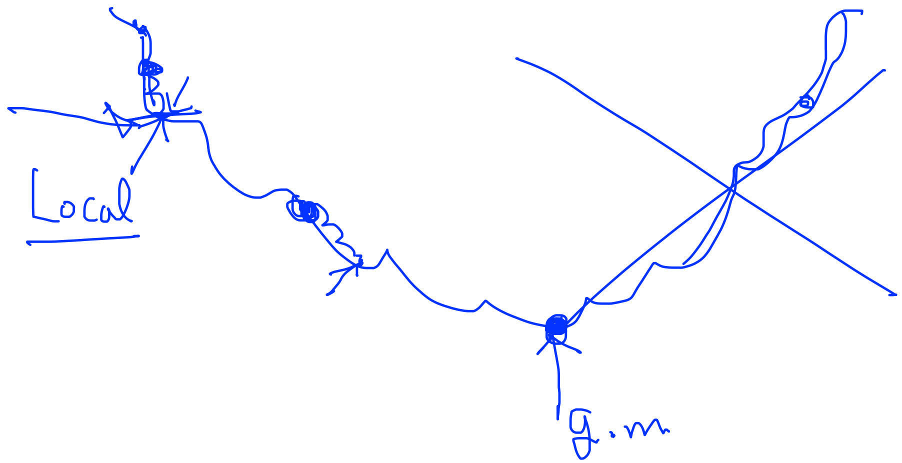
$0 < \sim < 1$

$$H(x) = Wx + b$$

$$H(X) = \frac{1}{1 + e^{-W^T X}}$$







New cost function for logistic

$$cost(W) = \frac{1}{m} \sum c(H(x), y)$$

$$c(H(x), y) = \begin{cases} -log(H(x)) & : y = 1 \\ -log(1 - H(x)) & : y = 0 \end{cases}$$

understanding cost function

$$C(H(x), y) = \begin{cases} -log(H(x)) & : y = 1 \\ -log(1 - H(x)) & : y = 0 \end{cases}$$

$\frac{1}{1+e^{-z}}$

log

~~Cost~~ $\boxed{y=1}$

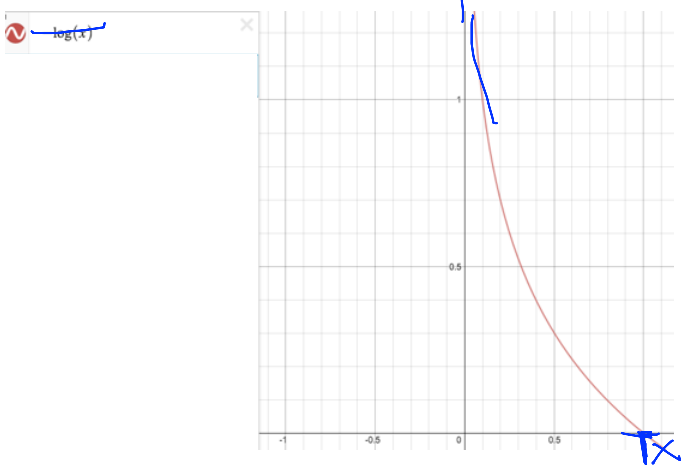
$H(x) = 1 \rightarrow cost(1) = 0$

$H(x) = 0 \rightarrow cost = \infty \uparrow$

cost

=

$g(z) = -\log(z)$

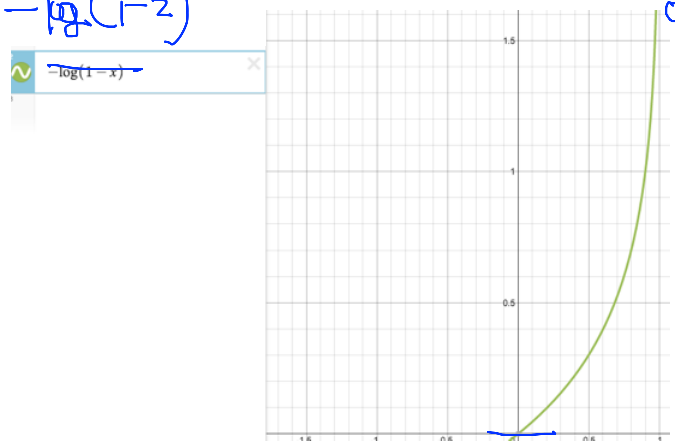


$y = 0$

$H(x) = 0, \; cost = 0$

$H(x) = 1, \; cost = \infty \uparrow$

$-\log(1-z)$



Cost function

$$cost(W) = \frac{1}{m} \sum c(H(x), y)$$

$$C(H(x), y) = \begin{cases} -log(H(x)) & : y = 1 \\ -log(1 - H(x)) & : y = 0 \end{cases}$$

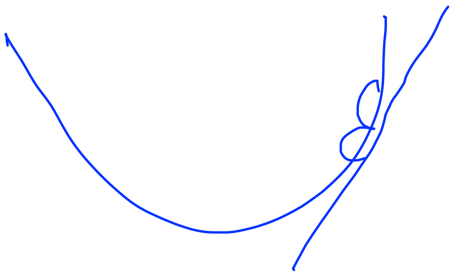
$$\boxed{C(H(x), y) = -ylog(H(x)) - (1 - y)log(1 - H(x))}$$

$y=1$, $c = -log(H(x))$

$y=0$, $c = -1 * log(1-H(x))$

Minimize cost - Gradient decent algorithm

$$cost(W) = -\frac{1}{m}\sum y\log(H(x)) + (1-y)\log(1-H(x))$$



$$W := W - \alpha\frac{\partial}{\partial W}cost(W)$$

Gradient decent algorithm

$$cost(W) = -\frac{1}{m}\sum y log(H(x)) + (1-y) log(1-H(x))$$

$$W := W - \alpha \frac{\partial}{\partial W} cost(W)$$

```
# cost function
cost = tf.reduce_mean(-tf.reduce_sum(Y*tf.log(hypothesis) + (1-Y)*tf.log(1-hypothesis)))

# Minimize
a = tf.Variable(0.1) # Learning rate, alpha
optimizer = tf.train.GradientDescentOptimizer(a)
train = optimizer.minimize(cost)
```

Next

Multinomial classification (Softmax)