Cost Function

Logistic Regression Model

Logistic Regression

Training set: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})\}$

m examples $\quad x \in \begin{bmatrix} x_0 \\ x_1 \\ \dots \\ x_n \end{bmatrix} \quad x_0 = 1, y \in \{0, 1\}$

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}}$$

How to choose parameters θ ?

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

Linear regression: $J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

- Linear regression: $J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}$

$\mathrm{cost}(h_\theta(x^{(i)}), y)$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

- Linear regression: $J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}$

$$\underline{\mathrm{cost}(h_\theta(x^{(i)}), y)}$$

$$\mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)}) = \frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

- Linear regression: $J(\theta) = \boxed{\frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}}$

  $\mathrm{cost}(h_\theta(x^{(i)}), y)$

$$\mathrm{Cost}(h_\theta(x), y) = \frac{1}{2}\left(h_\theta(x) - y\right)^2$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

→ ~~Linear~~ regression: $J(\theta) = \boxed{\frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2} \quad \mathrm{cost}(h_\theta(x^{(i)}), y)}$

logistic

$$\rightarrow \mathrm{Cost}(\underbrace{h_\theta(x^{\cancel{(i)}})}, \underbrace{y^{\cancel{(i)}}}) = \frac{1}{2}\underbrace{\left(h_\theta(x^{\cancel{(i)}}) - y^{\cancel{(i)}}\right)^2}$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

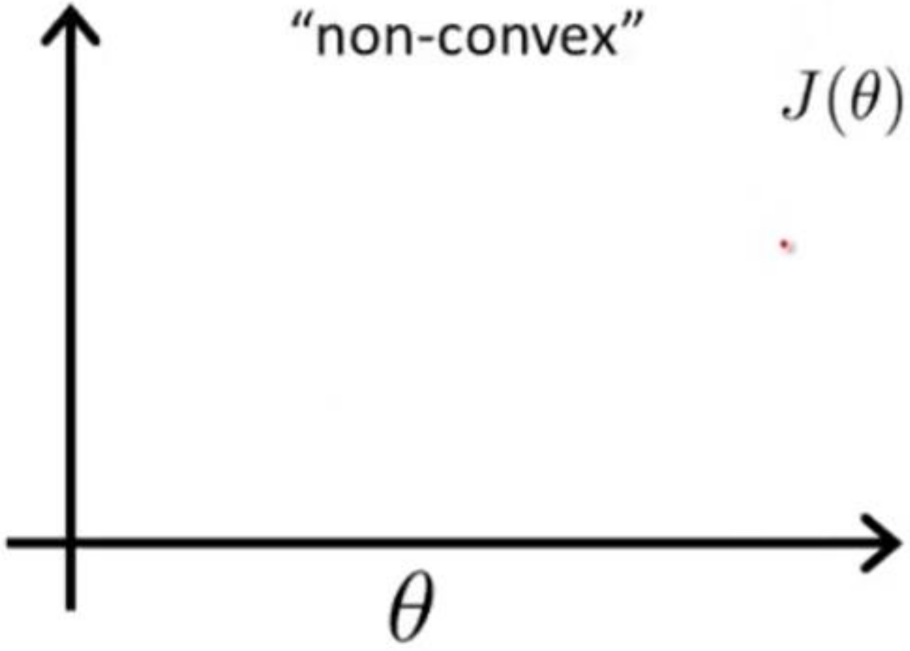
→ ~~Linear~~ regression: $J(\theta) = \boxed{\frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}}$

logistic

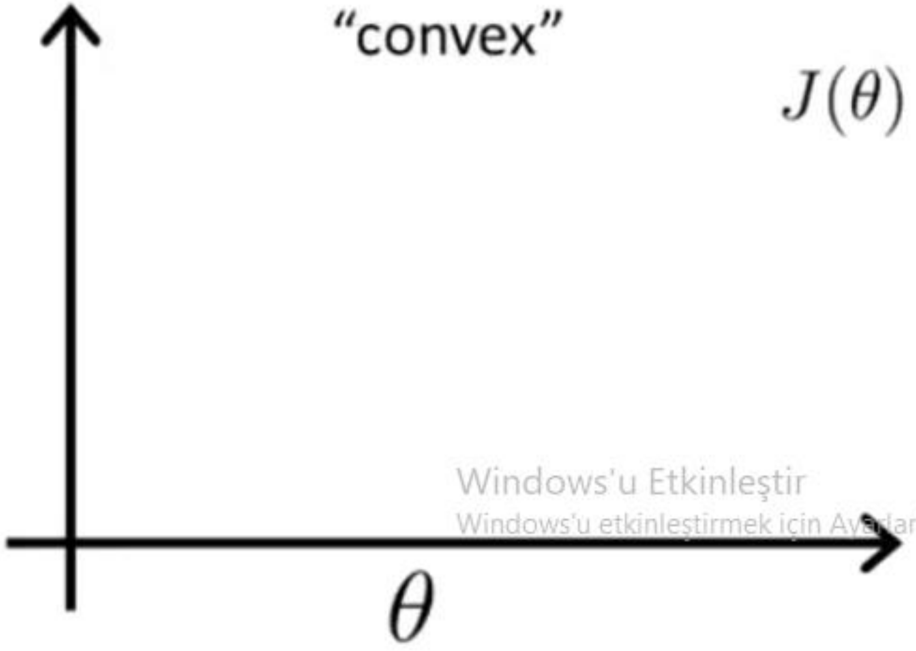
$\mathrm{cost}(h_\theta(x^{(i)}), y)$

→ $\mathrm{Cost}(\underbrace{h_\theta(x)}, \underbrace{y}) = \frac{1}{2}\underline{\left(h_\theta(x) - y\right)^2}$ ←

“non-convex”



“convex”



Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

→ ~~Linear~~ regression: $J(\theta) = \boxed{\frac{1}{m} \sum_{i=1}^{m} \cancel{\frac{1}{2} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}}$

logistic

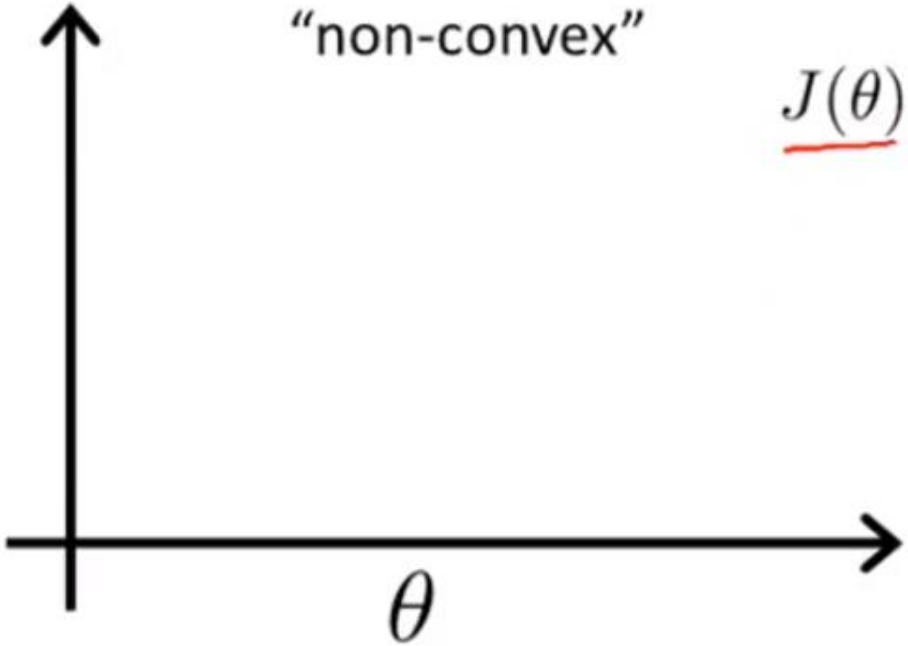
$\mathrm{cost}(h_\theta(x^{(i)}), y)$

$\frac{1}{1+e^{-\theta^T x}}$

→ $\mathrm{Cost}(h_\theta(x^{\cancel{(i)}}), y^{\cancel{(i)}}) = \frac{1}{2} \left(h_\theta(x^{\cancel{(i)}}) - y^{\cancel{(i)}}\right)^2$ ←

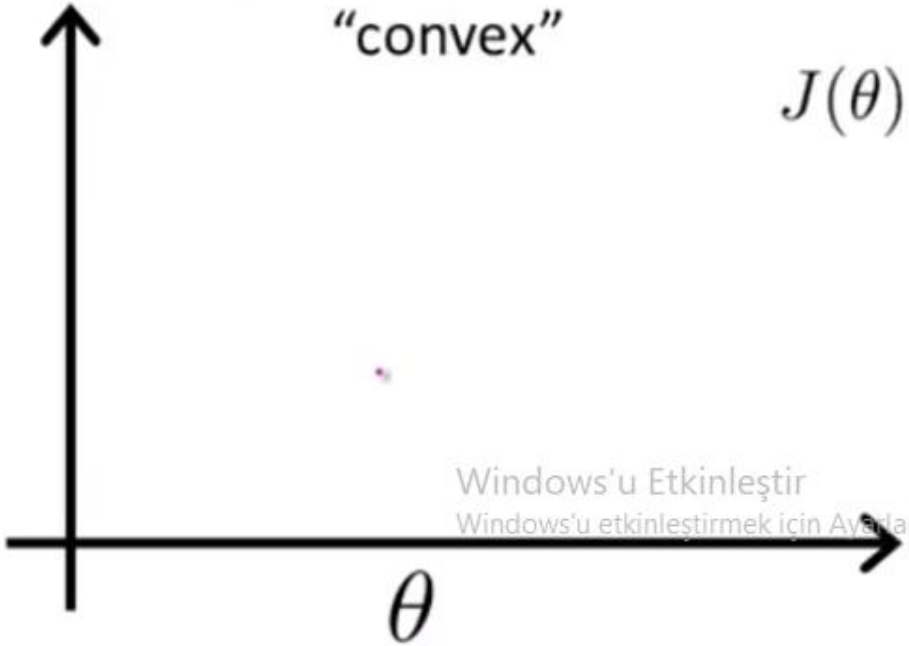
"non-convex"

$J(\theta)$



"convex"

$J(\theta)$



Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Cost function

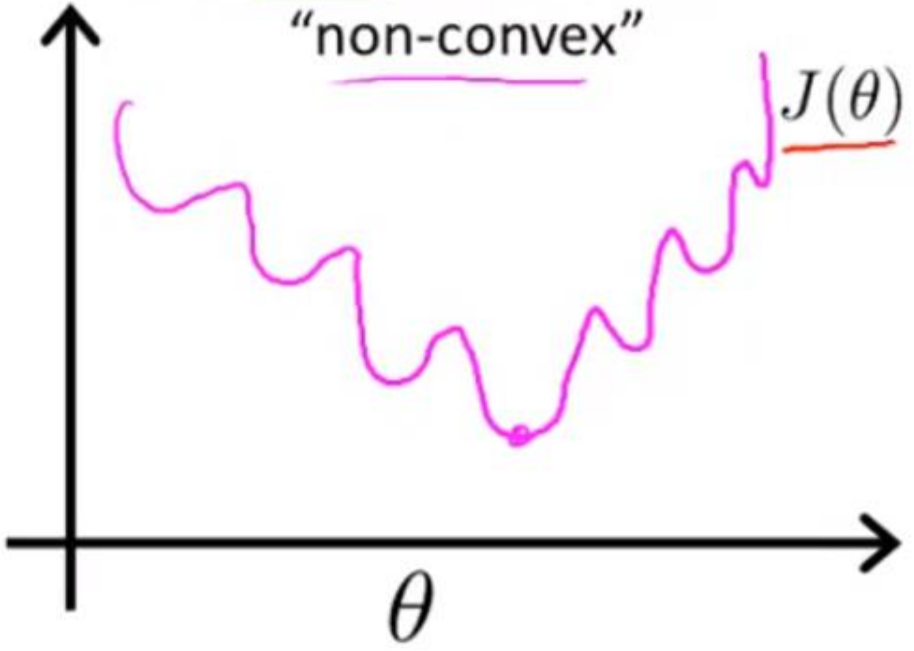
→ ~~Linear~~ regression: $J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2}$

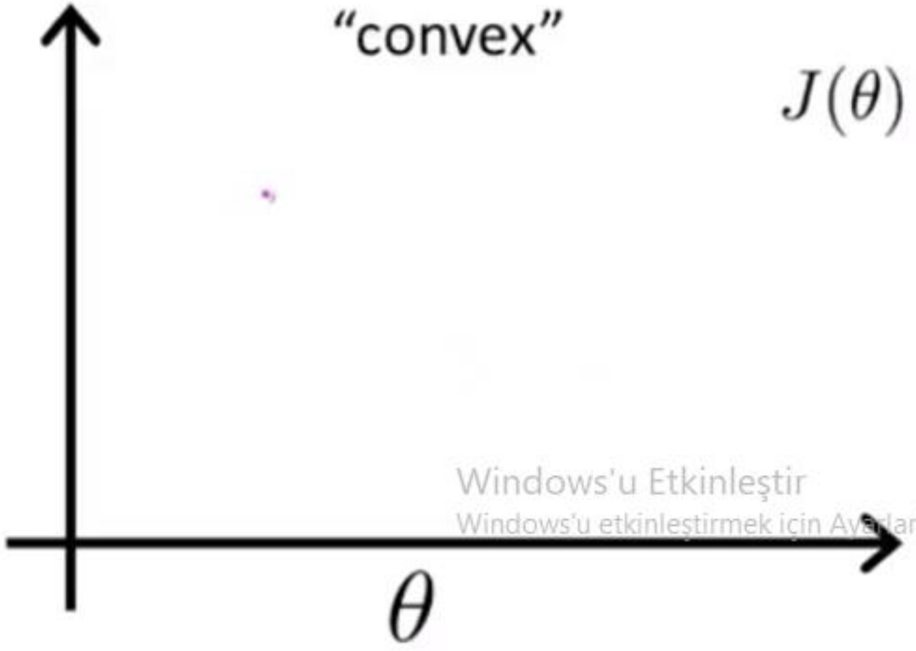
logistic

$\rightarrow \mathrm{cost}(h_\theta(x^{(i)}), y)$

$\rightarrow \mathrm{Cost}(h_\theta(x), y) = \frac{1}{2}\left(h_\theta(x) - y\right)^2 \leftarrow$

$\frac{1}{1+e^{-\theta^T x}}$





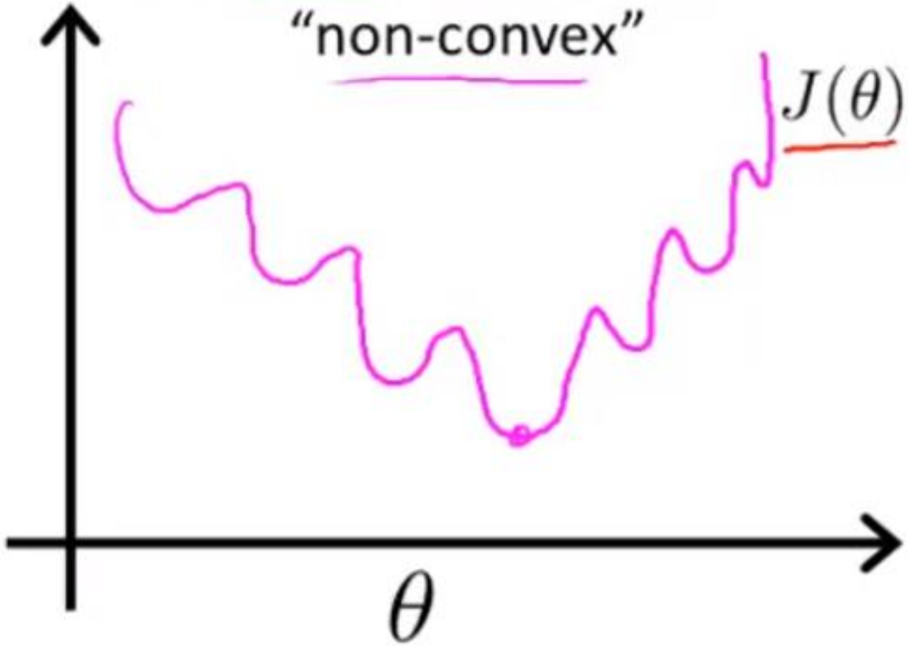
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

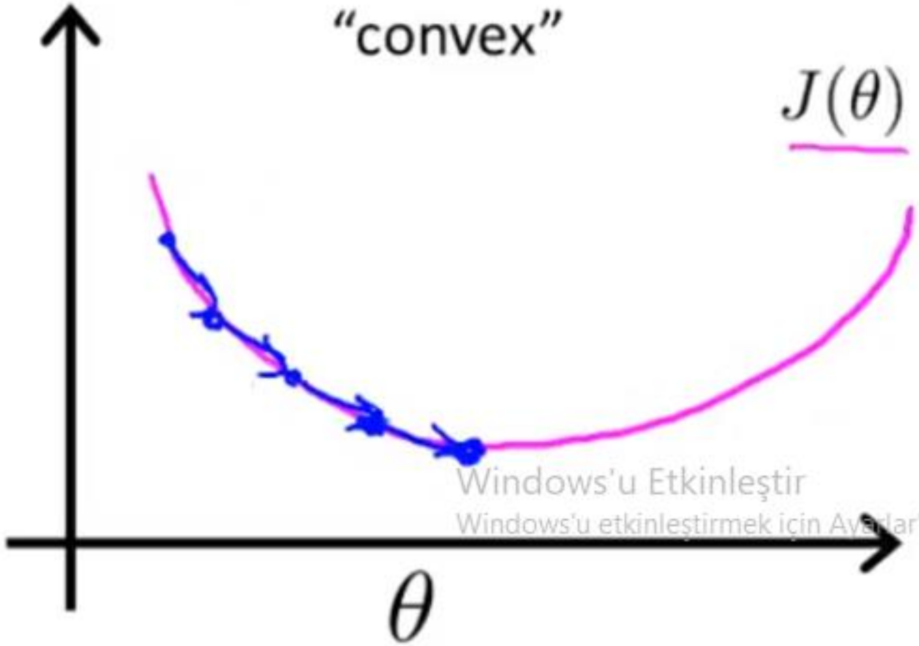
Cost function

→ ~~Linear~~ regression: logistic

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} \cancel{\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2} \; \mathrm{cost}(h_\theta(x^{(i)}), y)$$

$$\to \mathrm{Cost}(h_\theta(x), y) = \frac{1}{2}\left(h_\theta(x) - y\right)^2 \leftarrow \qquad \frac{1}{1+e^{-\theta^T x}}$$





Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Andrew Ng

Non-convexity???

- So what I would like to do is come up with a convex cost function so that the GD works well.
  - By working well we mean it finds the global minimum

Logistic regression cost function

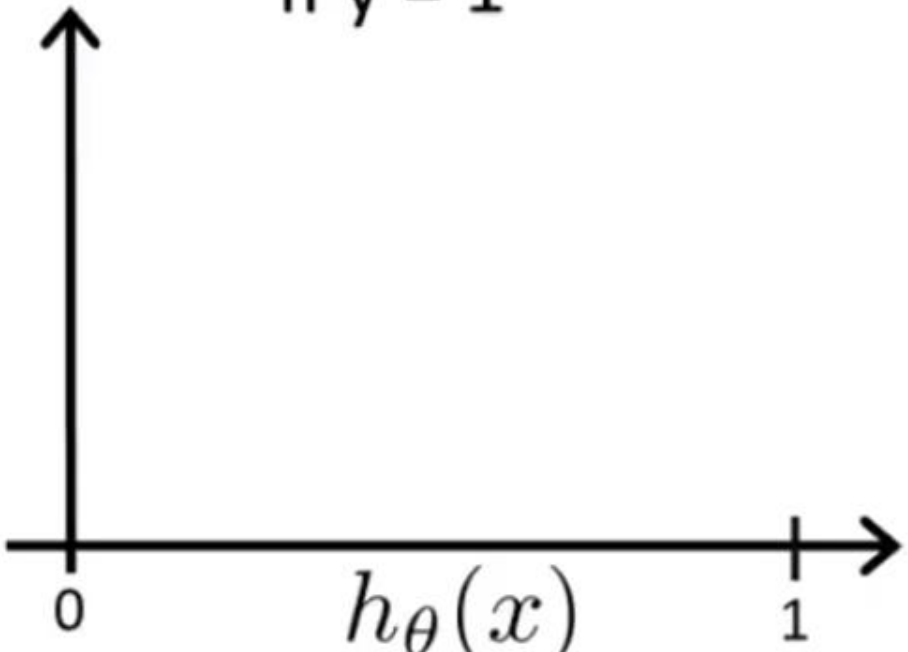
$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 1

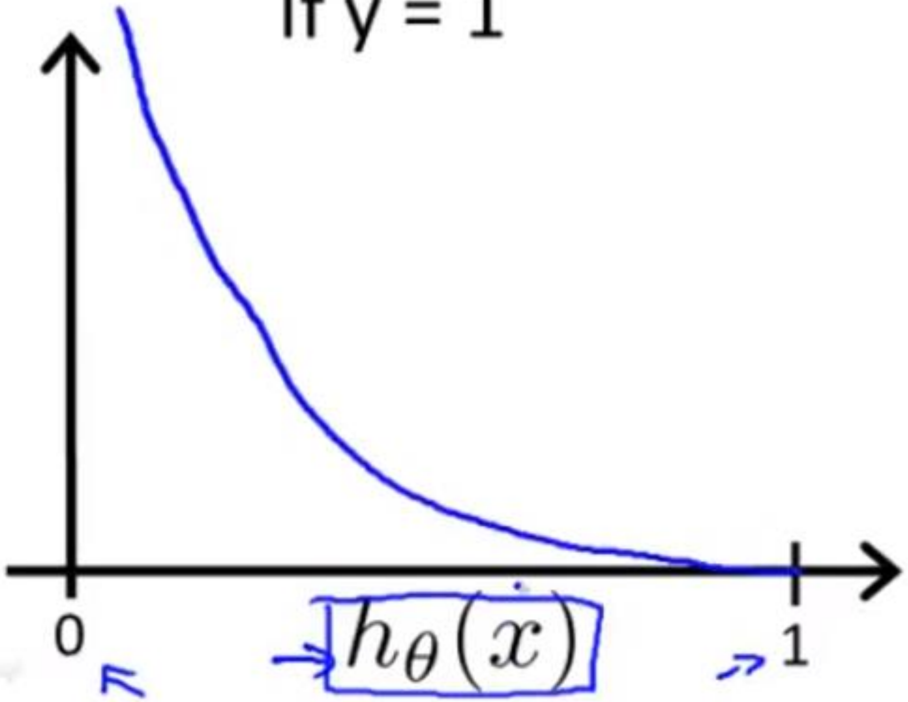


Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} \boxed{-\log(h_\theta(x))} & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 1

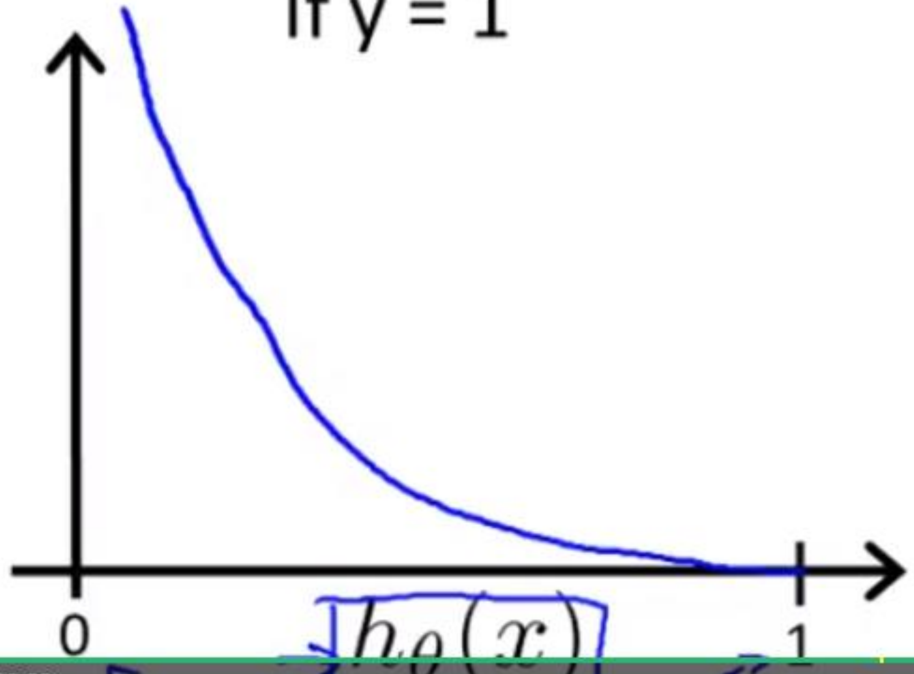


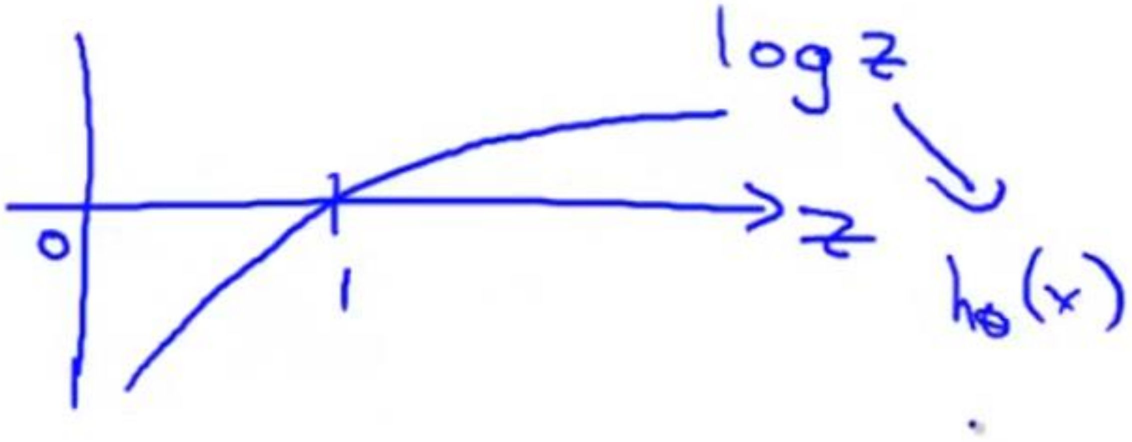
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} \boxed{-\log(h_\theta(x))} & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 1



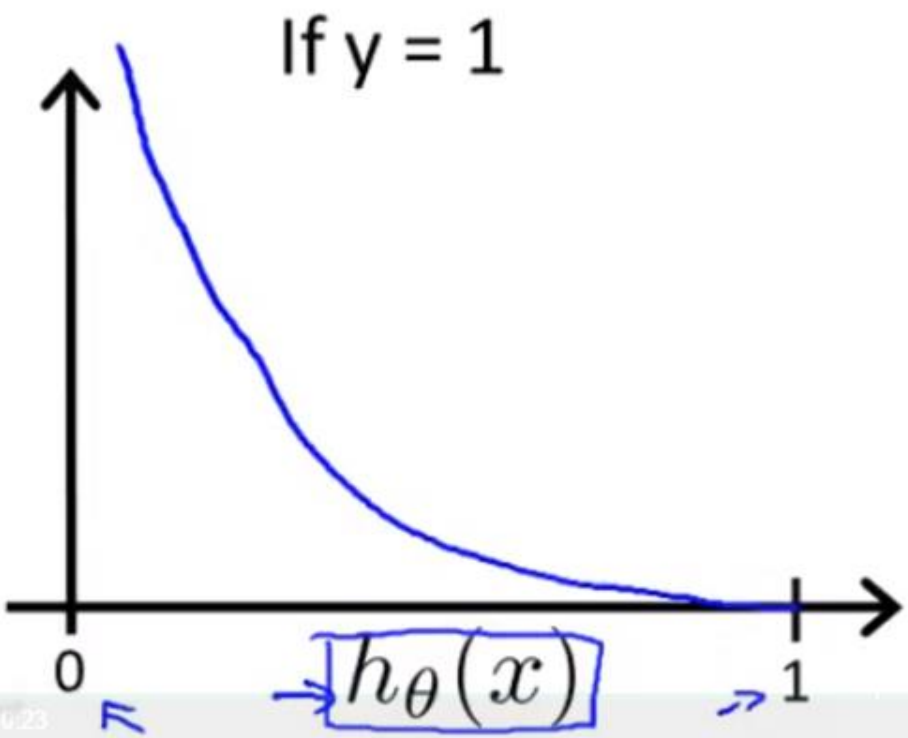


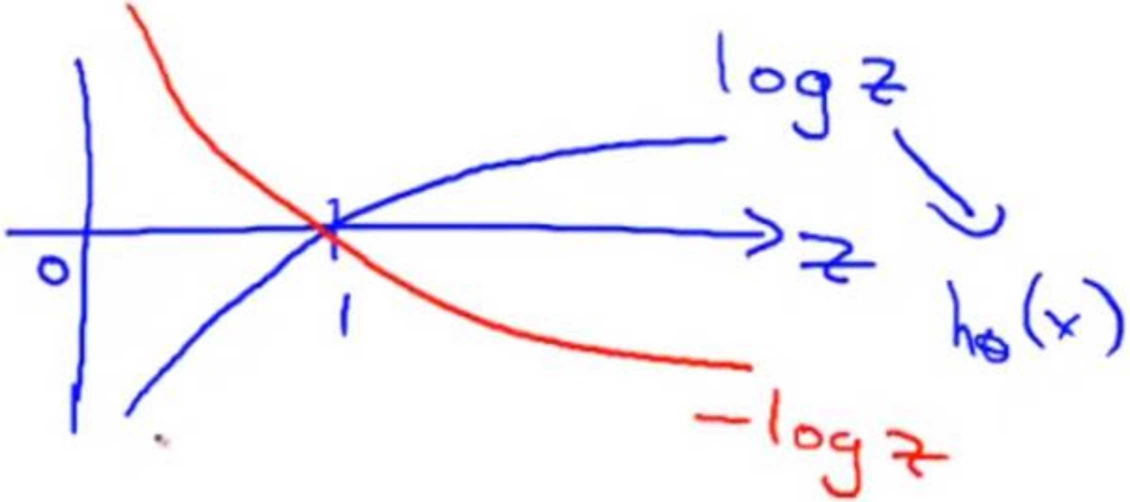
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

5:59 / 10:23
Andrew Ng

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} \boxed{-\log(h_\theta(x))} & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$



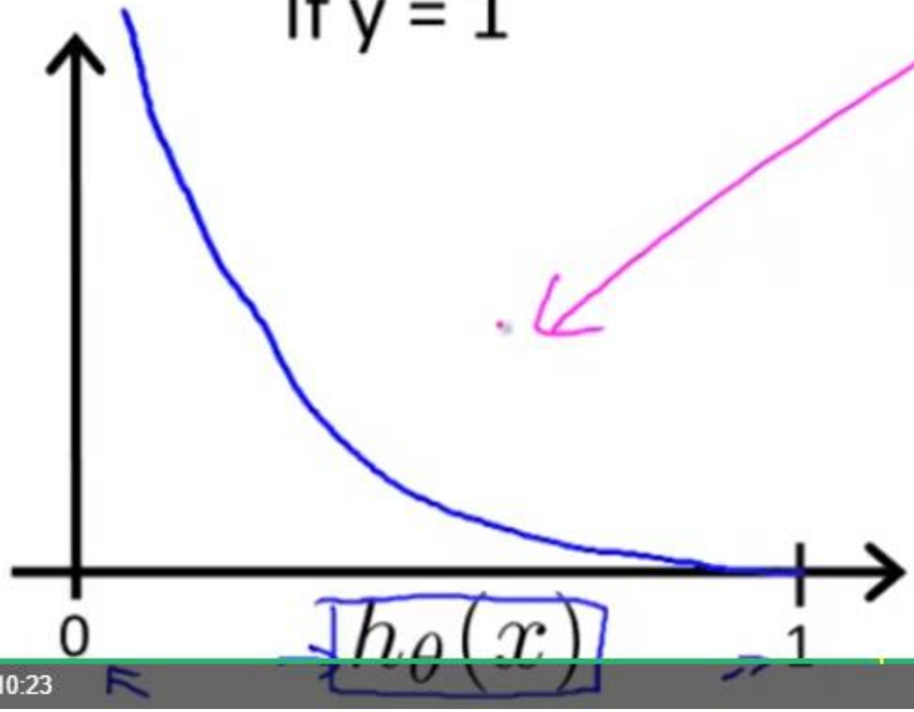


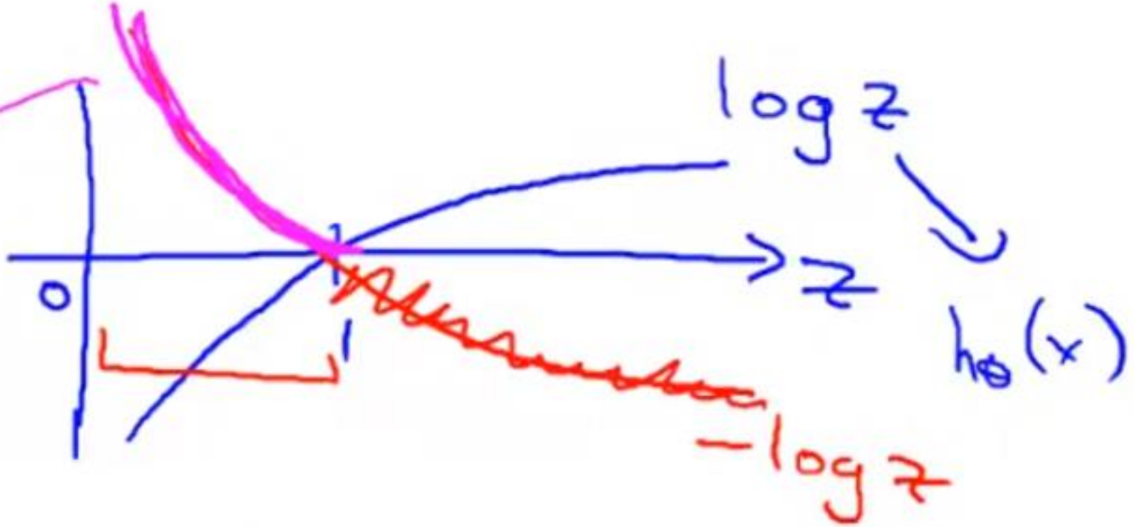
Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} \boxed{-\log(h_\theta(x))} & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 1



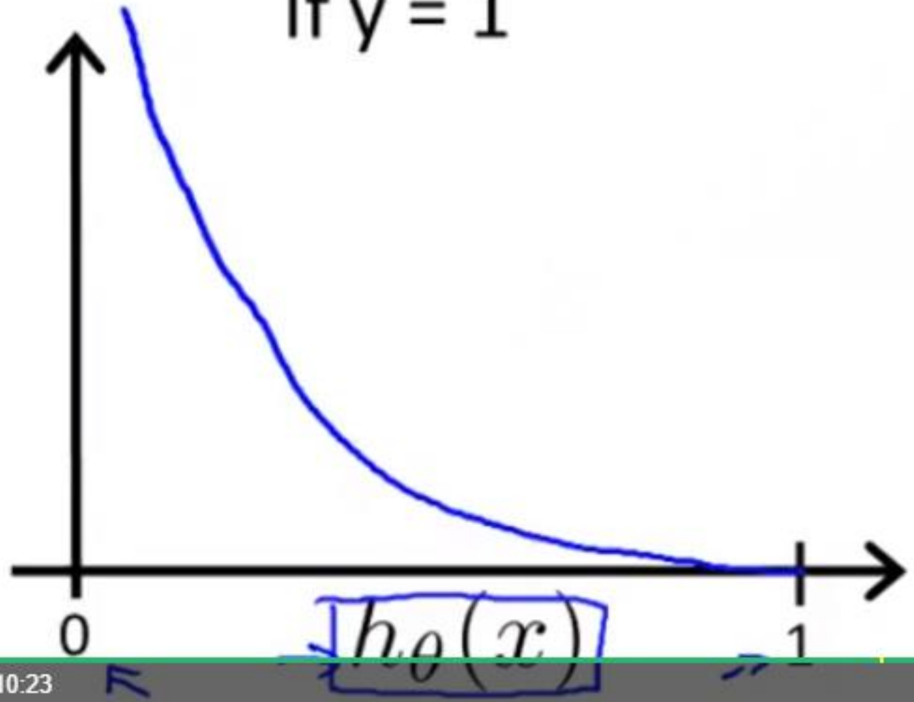


Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.
6:21 / 10:23
Andrew Ng

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} \boxed{-\log(h_\theta(x))} & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 1



$\mathrm{Cost} = 0$ if $y = 1, h_\theta(x) = 1$
But as $h_\theta(x) \rightarrow 0$
$Cost \rightarrow \infty$

Note that we are at y=1

Captures intuition that if $h_\theta(x) = 0$, (predict $P(y = 1|x;\theta) = 0$), but $y = 1$, we'll penalize learning algorithm by a very large cost.

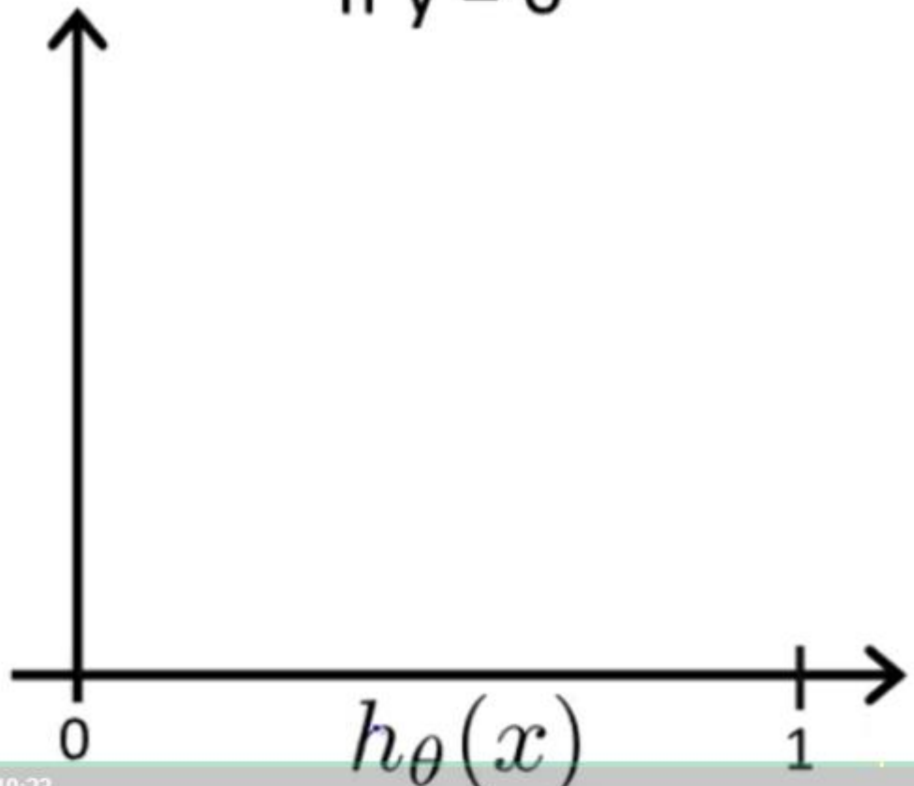
Hence we have to arrange theta's so that at this point h(x) !=0

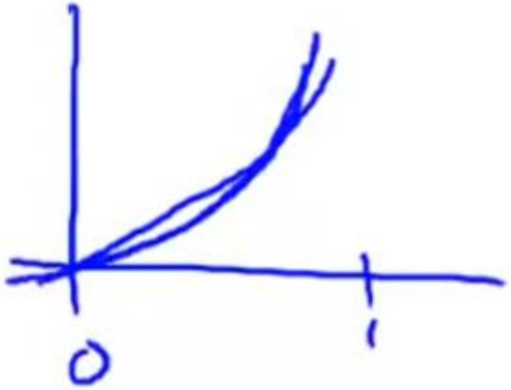
6:31 / 10:23

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x^{(i)}, y^{(i)}) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 0





$-\log(1-z)$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

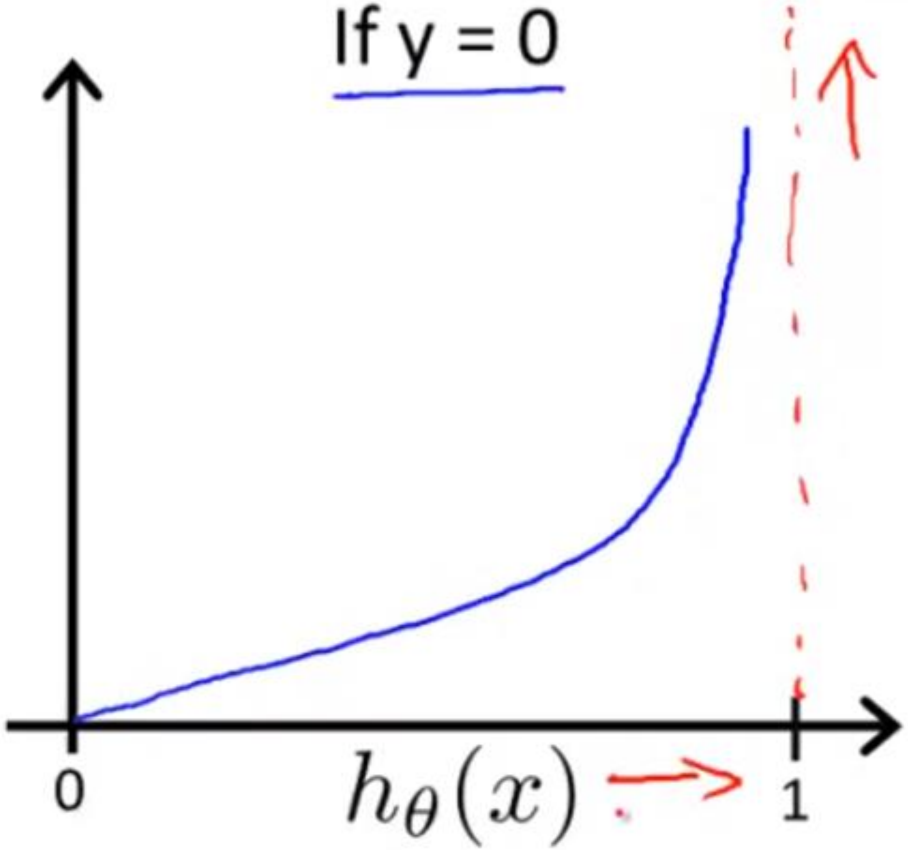
8:53 / 10:23

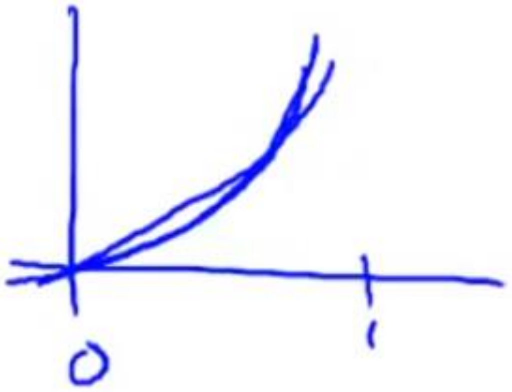
Andrew Ng

Logistic regression cost function

$$\mathrm{Cost}(h_\theta(x^{(i)}, y^{(i)}) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

If y = 0





$-\log(1-z)$

Windows'u Etkinleştir
Windows'u etkinleştirmek için Ayarlar'a gidin.

Andrew Ng

Exercise

In logistic regression, the cost function for our hypothesis outputting (predicting) $h_\theta(x)$ on a training example that has label $y \in \{0,1\}$ is:

$$\text{Cost}(h_\theta(x), y) = \begin{cases} -\log h_\theta(x) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

Which of the following are true? Check all that apply.

- ○ If $h_\theta(x) = y$, then cost$(h_\theta(x), y) = 0$ (for $y = 0$ and $y = 1$).
- ○ If $y=0$ then cost$(h_\theta(x), y) \rightarrow \infty$ as $h_\theta(x) \rightarrow 1$
- ○ If y=0 then cost$(h_\theta(x), y) \rightarrow \infty$ as $h_\theta(x) \rightarrow \infty$
- ○ Regardless of whether y=0 or y=1, if $h_\theta(x) = 0.5$, then cost$(h_\theta(x), y) > 0$