

Classification and Representation

- ✓

Video: Classification

8 min
- ✓

Reading: Classification

2 min
- ✓

Video: Hypothesis Representation

7 min
- ✓

Reading: Hypothesis Representation

3 min
- ✓

Video: Decision Boundary

14 min
- ✓

Reading: Decision Boundary

3 min

Logistic Regression Model

- ✓

Video: Cost Function

10 min
- ✓

Reading: Cost Function

3 min
- ✓

Video: Simplified Cost Function and Gradient Descent

10 min
- ✓

Reading: Simplified Cost Function and Gradient Descent

3 min
- ▶

Video: Advanced Optimization

14 min
- 📖

Reading: Advanced Optimization

3 min

Multiclass Classification

- ▶

Video: Multiclass Classification: One-vs-all

6 min
- 📖

Reading: Multiclass Classification: One-vs-all

3 min

Review

- ✓

Reading: Lecture Slides

10 min
- 📋

Quiz: Logistic Regression

5 questions

Solving the Problem of Overfitting

Review

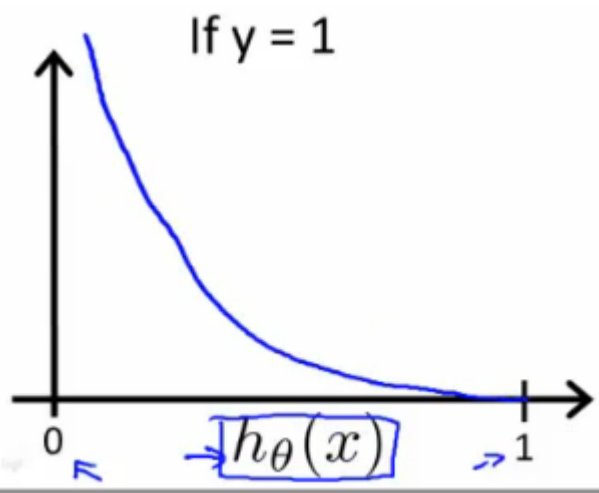
Cost Function

We cannot use the same cost function that we use for linear regression because the Logistic Function will cause the output to be wavy, causing many local optima. In other words, it will not be a convex function.

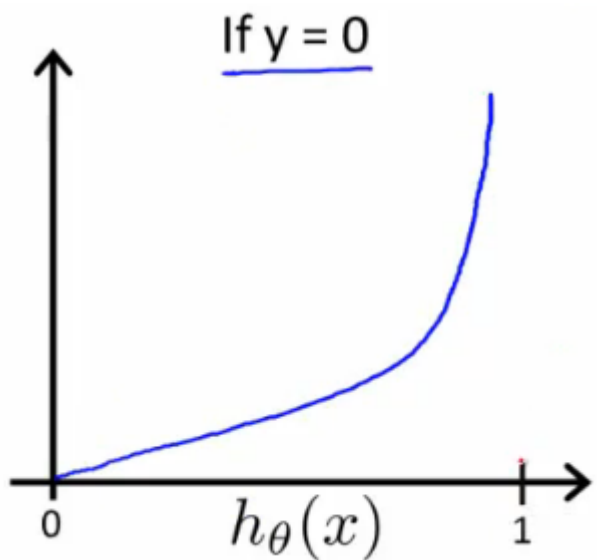
Instead, our cost function for logistic regression looks like:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m \text{Cost}(h_{\theta}(x^{(i)}), y^{(i)})$$
$$\begin{aligned} \text{Cost}(h_{\theta}(x), y) &= -\log(h_{\theta}(x)) && \text{if } y = 1 \\ \text{Cost}(h_{\theta}(x), y) &= -\log(1 - h_{\theta}(x)) && \text{if } y = 0 \end{aligned}$$

When $y = 1$, we get the following plot for $J(\theta)$ vs $h_{\theta}(x)$:



Similarly, when $y = 0$, we get the following plot for $J(\theta)$ vs $h_{\theta}(x)$:



$$\begin{aligned} \text{Cost}(h_{\theta}(x), y) &= 0 \text{ if } h_{\theta}(x) = y \\ \text{Cost}(h_{\theta}(x), y) &\rightarrow \infty \text{ if } y = 0 \text{ and } h_{\theta}(x) \rightarrow 1 \\ \text{Cost}(h_{\theta}(x), y) &\rightarrow \infty \text{ if } y = 1 \text{ and } h_{\theta}(x) \rightarrow 0 \end{aligned}$$

If our correct answer 'y' is 0, then the cost function will be 0 if our hypothesis function also outputs 0. If our hypothesis approaches 1, then the cost function will approach infinity.

If our correct answer 'y' is 1, then the cost function will be 0 if our hypothesis function outputs 1. If our hypothesis approaches 0, then the cost function will approach infinity.

Note that writing the cost function in this way guarantees that $J(\theta)$ is convex for logistic regression.