

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

Multiclass Classification

Review

Solving the Problem of Overfitting

Review



Decision Boundary

In order to get our discrete 0 or 1 classification, we can translate the output of the hypothesis function as follows:

$$\begin{aligned} h_{\theta}(x) &\geq 0.5 \rightarrow y = 1 \\ h_{\theta}(x) &< 0.5 \rightarrow y = 0 \end{aligned}$$

The way our logistic function g behaves is that when its input is greater than or equal to zero, its output is greater than or equal to 0.5:

$$\begin{aligned} g(z) &\geq 0.5 \\ \text{when } z &\geq 0 \end{aligned}$$

Remember.

$$\begin{aligned} z = 0, e^0 = 1 &\Rightarrow g(z) = 1/2 \\ z \rightarrow \infty, e^{-\infty} &\rightarrow 0 \Rightarrow g(z) = 1 \\ z \rightarrow -\infty, e^{\infty} &\rightarrow \infty \Rightarrow g(z) = 0 \end{aligned}$$

So if our input to g is $\theta^T X$, then that means:

$$\begin{aligned} h_{\theta}(x) = g(\theta^T x) &\geq 0.5 \\ \text{when } \theta^T x &\geq 0 \end{aligned}$$

From these statements we can now say:

$$\begin{aligned} \theta^T x &\geq 0 \Rightarrow y = 1 \\ \theta^T x &< 0 \Rightarrow y = 0 \end{aligned}$$

The **decision boundary** is the line that separates the area where $y = 0$ and where $y = 1$. It is created by our hypothesis function.

Example:

$$\begin{aligned} \theta &= \begin{bmatrix} 5 \\ -1 \\ 0 \end{bmatrix} \\ y &= 1 \text{ if } 5 + (-1)x_1 + 0x_2 \geq 0 \\ 5 - x_1 &\geq 0 \\ -x_1 &\geq -5 \\ x_1 &\leq 5 \end{aligned}$$

In this case, our decision boundary is a straight vertical line placed on the graph where $x_1 = 5$, and everything to the left of that denotes $y = 1$, while everything to the right denotes $y = 0$.

Again, the input to the sigmoid function $g(z)$ (e.g. $\theta^T X$) doesn't need to be linear, and could be a function that describes a circle (e.g. $z = \theta_0 + \theta_1 x_1^2 + \theta_2 x_2^2$) or any shape to fit our data.