Logistic Regression

David Robinson

Logistic regression is a supervised learning algorithm that predicts the probability of a binary outcome based on input features.

Sigmoid Function

The **sigmoid function** transforms input values into output values that lie between 0 and 1.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Cost Function

In linear regression, mean squared error is used but will give results with local minima, which will not help with non-linear tasks, like logistic regression. Logistic regression uses log loss, or **cross-entropy loss**, which is derived from the maxmimum likelihood estimation method.

$$\text{Cost}(h_\theta(x), y) = -\log(1 - y + (2y - 1)h_\theta(x))$$

Loss over a batch size of m,

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \text{Cost}(h_\theta(x), y)$$

At each epoch,

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

Linear Regression	**Logistic Regression**
Predicts a **continuous** dependent variable using independent variables.	Predicts a **categorical** (binary) dependent variable using independent variables.
Solves **regression** problems.	Solves **classification** problems.
Uses the **Least Squares** estimation method for accuracy.	Uses the **Maximum Likelihood** estimation method for accuracy.
Fits a **straight line** (best-fit line) to predict the output.	Fits an **S-curve** (logistic function) to classify the samples.
Requires a **linear relationship** between the dependent and independent variables.	Does **not** require a linear relationship between the dependent and independent variables.

Key Points

1. The decision boundary is linear if the model includes only linear terms.
2. Logistic regression handles multi-class classification by using multiple one-vs-all (OvR) classifiers.