Logistic Regression: Theory & Visualizations

1. Logistic Regression Formula

Logistic regression predicts the probability of a binary outcome using the sigmoid function:

P(y = 1 | X) = 1 / (1 + e^-(b0 + b1*x1 + ... + bn*xn))

The output is a probability between 0 and 1.

2. Sigmoid Function

The sigmoid function maps any real-valued number to (0, 1):

sigmoid(z) = 1 / (1 + e^(-z))

This makes it suitable for modeling probabilities.

3. Gradient Descent in Logistic Regression

To train logistic regression, we use gradient descent to minimize the binary cross-entropy loss:

Loss = -1/m * sum [y * log(y_hat) + (1 - y) * log(1 - y_hat)]

Update rule (each epoch):

theta := theta - alpha * gradient

Where:

- alpha is the learning rate
- gradient = derivative of the loss with respect to parameters

4. Decision Boundary

Logistic Regression: Theory & Visualizations

Logistic regression draws a linear decision boundary between classes in feature space.

This boundary corresponds to the threshold where $P(y=1) = 0.5$.