Logistic Regression

Some regression algorithms can be used for classification as well and vice versa. Logistic Regression is commonly used to **estimate the probability that an instance belongs to a particular class** (e.g. what is the probability that this email is spam?).

If the estimated probability is greater than 50%, then the model predicts that the instance belongs to the positive class, labeled ‘1’, otherwise, belongs to the negative class, labeled ‘0’. This makes it a binary classifier.

The logistic - noted $\sigma(\cdot)$ - is a sigmoid function (i.e., S-shaped) that outputs a number between 0 and 1.

Training and Cost Function

The objective of the training is to set the parameter vector θ so that the model estimates high probabilities for positive instances ($y = 1$) and low probabilities for negative instances ($y = 0$).