DLI Accelerated Data Science Teaching Kit

Lecture 14.3 - Linear Model

Linear Supervised Model

Linear Model

$$Y = w_0 + w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + \cdots + w_n \times x_n$$

- **Linear Classification**
- **Linear Regression**

Linear Classification

Logistic regression is a simple linear classification algorithm that predicts the probability of a binary response belonging to one class or the other.

- If the estimated probability is greater than 50%, then the model predicts that the instance belongs to that class (called the positive class, labeled “1”)
- Or else it predicts that it does not (i.e., it belongs to the negative class, labeled “0”).

Because of its simplicity, logistic regression is commonly used as a starting point for binary classification problems.

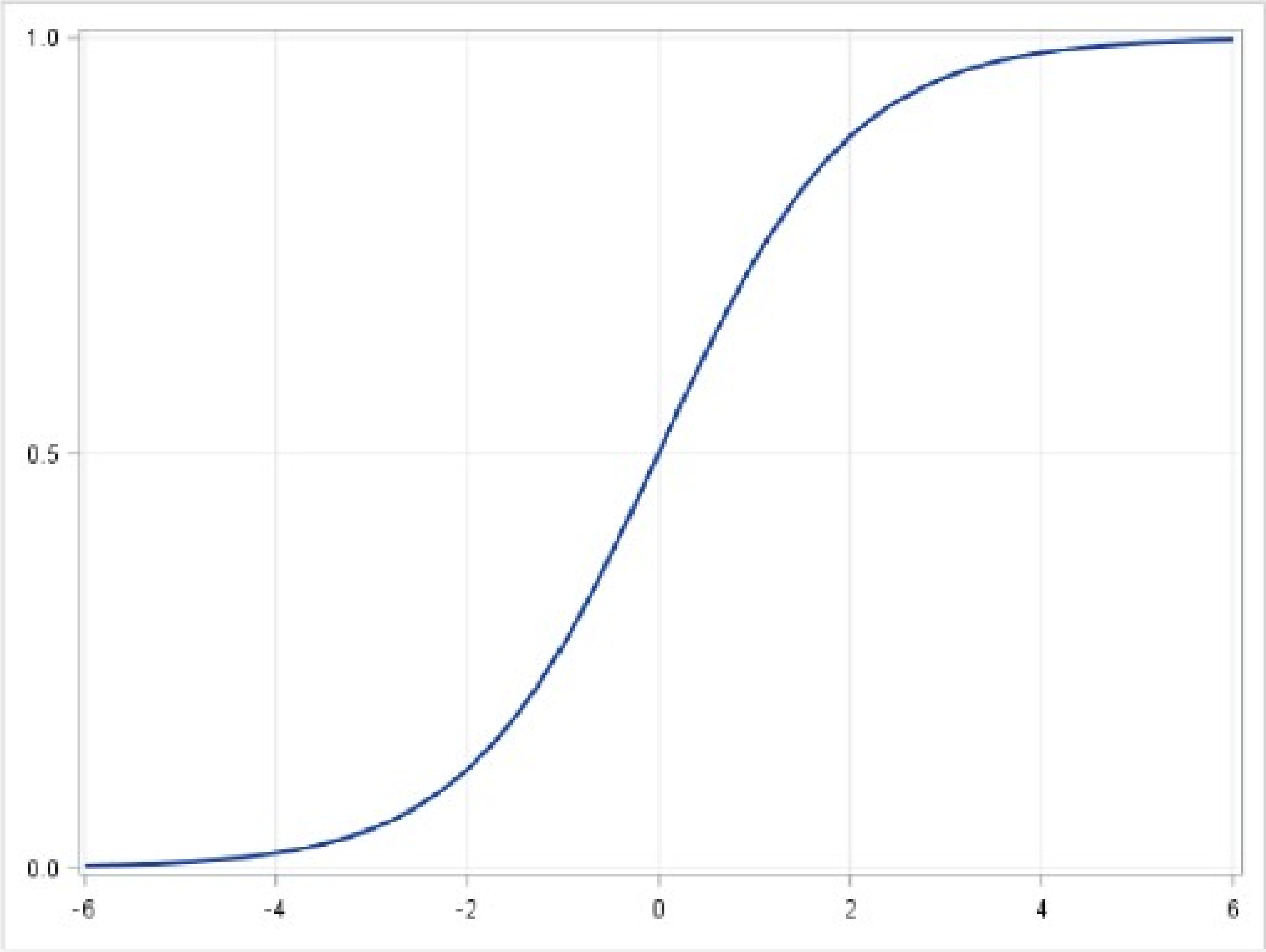
Logistic regression can be used as a baseline for evaluating more complex classification method.

Logistic Regression

It’s a classification rather than regression algorithm.

It uses an 'S'-shaped curve instead of a straight line makes it a natural fit for dividing data into groups.

In this model, the probabilities describing the possible outcomes of a single trial are modeled using a logistic function.



Logistic regression

Logistic Regression Hypothesis

$$h_\theta(x) = \theta^T x$$

When our hypothesis $h_\theta(x)$ outputs some number, we can treat that number as estimated probability that $y = 1$ on input x.

In a binary classification we use a different hypothesis

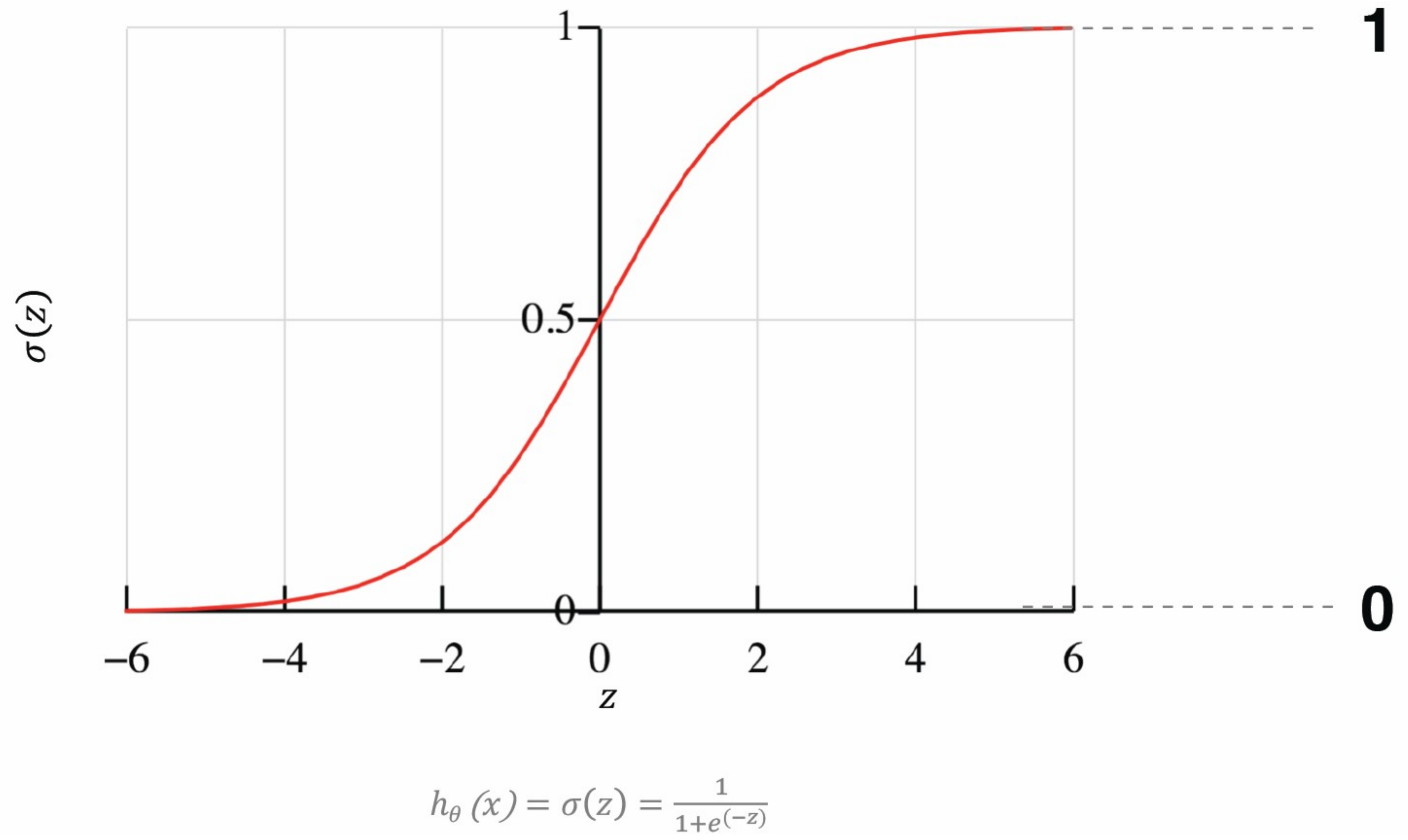
- Predict the probability that a given example belongs to the “1” class versus the probability that it belongs to the “0” class
  - 1 (positive) : e.g. malignant tumor
  - 0 (negative) : e.g. non-malignant tumor

Sigmod Function

Use **Sigmod function** to transform the linear calculation to probability.

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

$$z = h_\theta(x) = \theta^T x$$



Squash z into (0,1)

Parameter Learning

Search for a value of θ so that model estimates high probabilities for positive instances (y =1) and low probabilities for negative instances (y = 0)

Logistic regression model uses a different approach with logistic— Maximum Likelihood Estimation.

Loss Function: cross entropy loss

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y^i log\sigma\left(x^i\right) + \left(1 - y^i\right)\log(1 - \sigma(x^i))$$

Linear Regression

A supervised machine learning algorithm

Linear relationship between the samples and labels

A linear relationship between a dependent variable and independent variable(s).

The value of the dependent variable of a linear regression model is a continuous value i.e. real numbers.

Linear Regression

Simple Linear Regression

$$Y = w_0 + w_1 \times x_1$$

Multiple Linear Regression

$$Y = w_0 + w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 + \cdots + w_n \times x_n$$

Loss Function: mean square error

$$J(w) = -\frac{1}{m}\sum_{i=1}^{m}(y^i - \hat{y}^i)^2$$

DLI Accelerated Data Science Teaching Kit

Thank You