Data-Driven Problem Solving in Mechanical Engineering

Logistic Regression

Masoud Masoumi

ME 364 - Spring 2022

Department of Mechanical Engineering
The Cooper Union for the Advancement of Science and Art

April, 2022

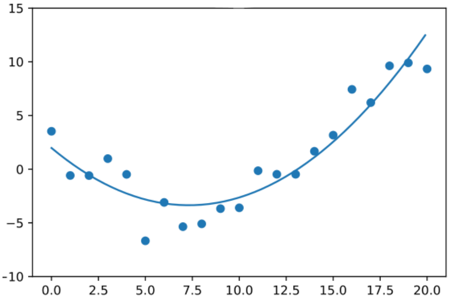
Regression

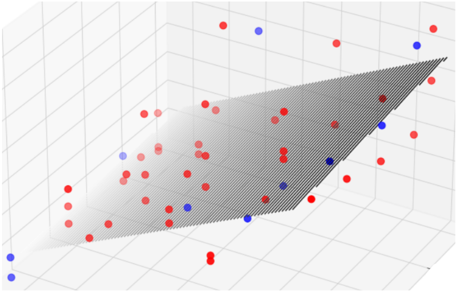
We can apply linear regression to classification problems by converting the class names of training examples to numbers, i.e. probabilities.

Recall Multiple **Linear Regression**

$f(\mathrm{x}) = \omega_o + \omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + ... \rightarrow f(\mathrm{x}) = \mathrm{w}^T \mathrm{x}$

with $\mathrm{w}^T = [\omega_o, \omega_1, \omega_2, \omega_3, ...]$ and $\mathrm{x} = [1, x_1, x_2, x_3, ...]^T$





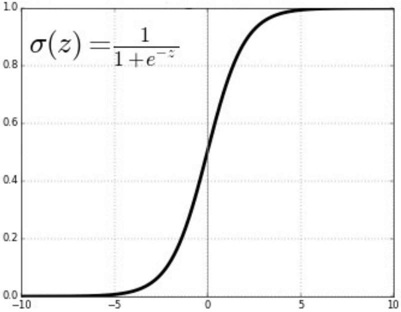
Regression to Logistic Regression

- In order to estimate the class of a data point, we need some sort of guidance on what would be the *most probable class* for that data point.

- Logistic regression produces a formula that predicts the probability of the class label as a function of independent variables.

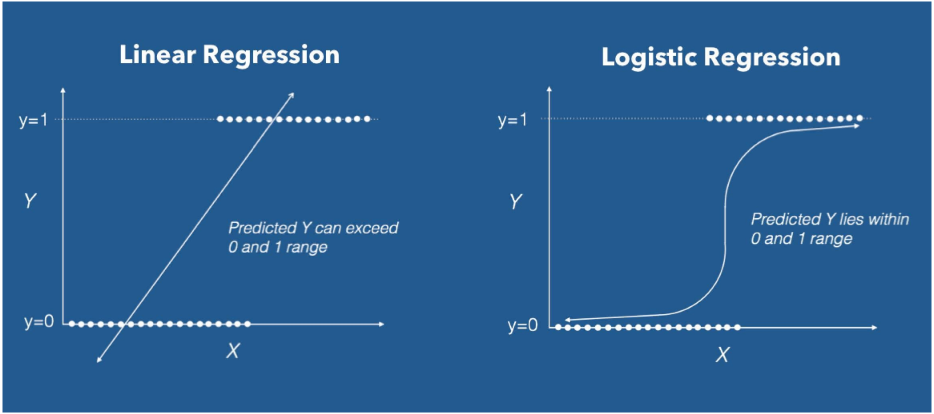
- In logistic regression, we need to output a probability, a value between 0 and 1. $P(y = 1|x)$ (Probability of y being class 1 given features x)

- One way to achieve this goal is to use logistic function or Sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$, whose output is between 0 and 1.



Logistic Regression

Logistic regression fits Sigmoid function by taking the linear regression and transforming the numeric estimate into a probability $\sigma(\mathrm{w}^T x) = \frac{1}{1+e^{-\mathrm{w}^T x}}$



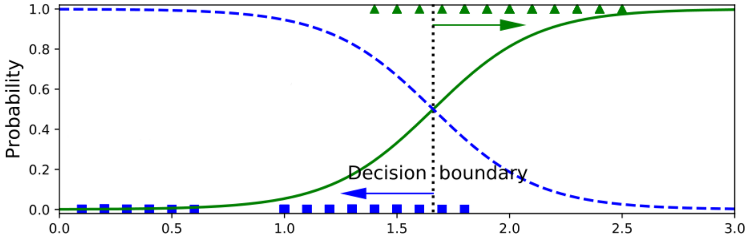
(ref: https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148)

Logistic Regression

In other words and in statistical language, we are looking for the probability of dependent variable being a specific class, given the selected features.

Probability of y being class 1 given features x: $P(y=1|x) = \sigma(\mathrm{w}^T x) = \frac{1}{1+e^{-\mathrm{w}^T x}}$.

The prediction function returns a probability score between 0 and 1. In order to map this to a discrete class (true/false, cat/dog, faulty/not faulty, 1/0), we select a *threshold value* above which we will classify values into class 1 and below which we classify values into class 0.



(*modified from* Murphy, Kevin P. Probabilistic machine learning: an introduction. MIT press, 2022)

Model Evaluation

To evaluate our model logistic regression model, we introduce a new metric for classifiers where the predicted output is a probability value between 0 and 1.

$$-\begin{cases} \log a & y = 1, \\ \log(1-a) & y = 0. \end{cases}$$

For one prediction, we can define **Log Loss** as

$$-y \times log[\sigma(\mathrm{w}^T x)] - (1-y) \times log[1 - \sigma(\mathrm{w}^T x)]$$

For all m samples

Log Loss= $\frac{1}{m}\sum_{i=1}^{m} -y^{(i)} \times log[\sigma(\mathrm{w}^T x^{(i)})] - (1-y^{(i)}) \times log[1 - \sigma(\mathrm{w}^T x^{(i)})]$