Stats Paper on Sigmoid Function

Giovanni Lunetta

University of Connecticut

February 1, 2023

Abstract

A sigmoid function is a mathematical function having that looks like an "S"-shaped curve. This function will be explored and its relationship to machine learning will be stated.

1 Introduction

A very common example of the sigmoid function is logistic function. This function can be defined by the formula:

$$S(x) = \frac{1}{1 + e^{-x}}. \tag{1}$$

2 Definition

According to Wikipedia (wik, 2022), "A sigmoid function is a bounded, differentiable, real function that is defined for all real input values and has a non-negative derivative at each point and exactly one inflection point. A sigmoid 'function' and a sigmoid 'curve' refer to the same object.

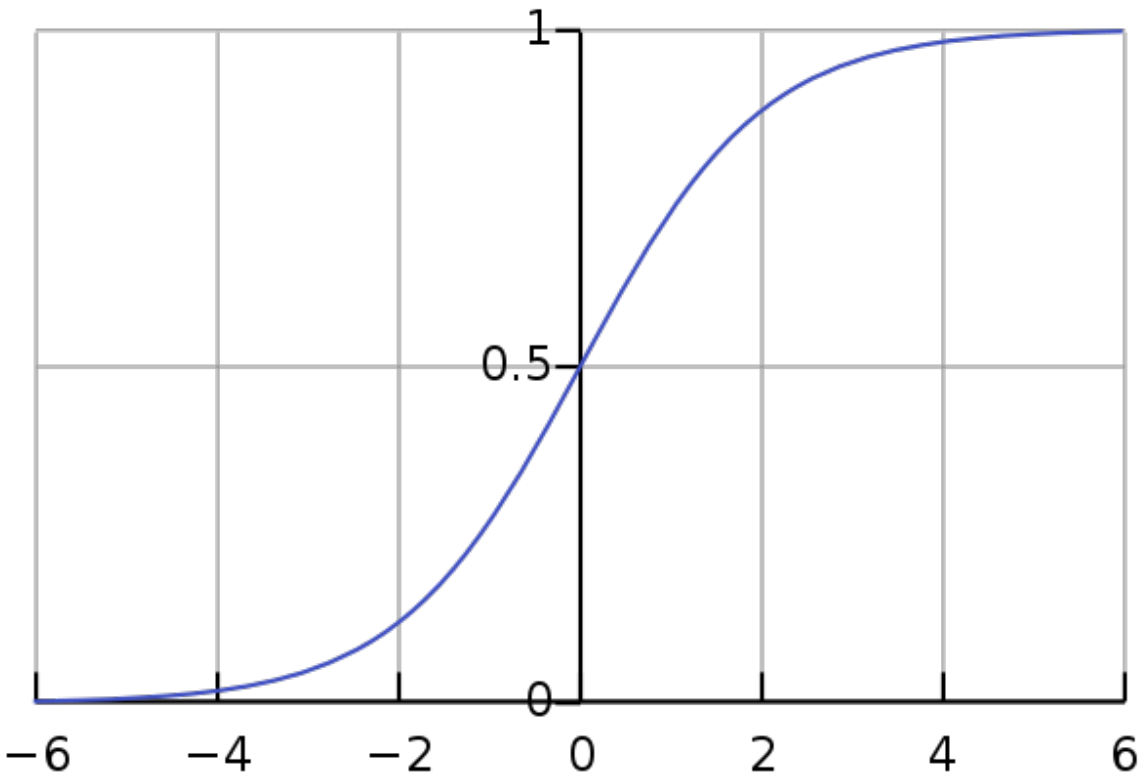

Figure 1: Here is a graph of the function shown in Equation (1).

3 Function

The equation in Section 1 can also be written as $1 - S(-x)$. As shown in the figure above, a sigmoid function is a smooth, upward sloping curve that has a bell shape and evens off at high and low values. Its integral is related to common probability distributions like the normal and Cauchy distributions. Sigmoid functions are either concave or convex around a certain point, usually 0.

4 Uses

The Sigmoid Function gets most of its use in the data-science world, particularly in classification machine learning problems. A good real world example of this functions use case would be attempted to create a machine learning model that predicts if an email is spam

or not. Here is a small picture of what a dataset might look like where a "yes" is coded as "1" and a "no" is coded as a "0." If we were to graph these points, we would see that the

Table 1: Small Spam Dataset

	Attachments	Emoji	Spam
User 1	1	1	1
User 2	1	0	0
User 3	0	0	1
User 4	1	0	1
User 5	0	1	1
User 6	0	0	0
User 7	0	1	0

data points would have a y value of either 0 or 1. This makes Figure 1 a perfect use-case for predicting the proper outcome.

5 Further Ideas

Data such as the data seen in Table 1 under Section 4 is not hard to come by. There are many websites that can provide you with data. Such as Kaggle (kag, 2023) and NYC Open Data City of New York (2023).

References

(2022), "Sigmoid function," .

(2023), "Find open datasets and Machine Learning Projects," .

City of New York, N. O. D. (2023), "NYC open data," .