

# Derivative of the Sigmoid function - a worked example

Aug. 10, 2017, 10:43 p.m.

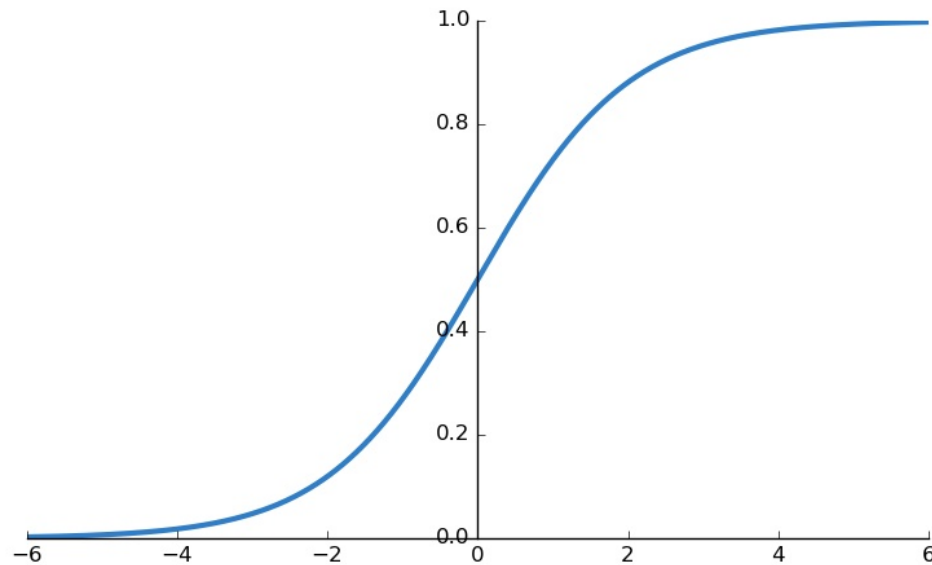
## Summary

Although I have practical experience with machine learning, I do have gaps in my knowledge regarding the mathematics behind deep learning. So today I worked on calculating the derivative of a sigmoid function, which is something that had perplexed me previously.

This post is mostly for my own reference so I can find my notes on this issue easily. But hopefully it will be of value to other people too.

## Sigmoid Function

The sigmoid function is one of many possible functions that are used as a nonlinear activation function between layers of a neural network. One of the properties that makes it appealing is that for any input value, it maps it to an output value between 0 and 1, which makes it useful if you want to perform binary classification.



The formula that specifies this curve is:

$$a = \frac{1}{1 + e^{-z}}$$

## Derivative of Sigmoid Function

The reason why calculating the derivative of this function is important, is because the learning process for neural networks involves making small changes to parameters, proportional to the partial derivatives of those parameter values, and relative to the loss function. This requires stepping backwards through the layers, applying the chain rule along each step. So,

one of those components that require calculation, is the derivative of the activation function.

One of the interesting properties of the Sigmoid function is that the derivative can be expressed in terms of the function itself. The reason this is appealing is because it can save computation time, as we can just reuse the same calculated values for the forward pass, and for the gradient.

The derivative is commonly expressed as follows (where  $a$  represents the original sigmoid function):

$$\frac{da}{dz} = a(1-a)$$

## Calculating the Derivative

Below, I will go step by step showing how the derivative is calculated.

Firstly, it is useful if we represent the fraction as a power.

$$a = (1 + e^{-z})^{-1}$$

This representation is really just a personal preference, since it means I only ever have to remember the chain rule in calculus, and not worry about remembering the quotient rule. We can therefore think of the formula as being composed of as a function of a function, as follows:

$$a = \underbrace{\left( \underbrace{1 + e^{-z}}_{g(z)} \right)^{-1}}_{a(g)}$$

We can now continue on with the chain rule to calculate the rest as follows:

$$\begin{aligned}
\frac{dg}{dz} &= \frac{da(g)}{g} \cdot \frac{dg(z)}{dz} \\
&= \frac{d}{dg} g^{-1} \cdot \left( \frac{d}{dz} 1 + e^{-z} \right) \\
&= -1 g^{-2} \cdot \left( \frac{d}{dz} 1 + \frac{d}{dz} e^{-z} \right) \quad \text{h} = -z \\
&= \boxed{\frac{-1}{g^2}} \cdot \left( \boxed{0} + \frac{d}{dh} e^h \cdot \frac{d}{dz} -z \right) \\
&\quad \quad \quad e^h \cdot -1 \\
&\quad \quad \quad \boxed{-e^h} \\
&\quad \quad \quad \boxed{-e^{-z}} \\
&= \frac{-1}{g^2} \cdot (0 + -e^{-z}) \\
&= \frac{-1}{g^2} \cdot -e^{-z} \\
&= \frac{e^{-z}}{g^2}
\end{aligned}$$

When we substitute  $g$  back in we get:

$$\frac{da}{dz} = \frac{e^{-z}}{(1+e^{-z})^2}$$

This doesn't look anything like the derivative that was shown previously. You can rest assured that it is actually representing the same thing. And technically it is the correct answer. But, we would like to make use of the property of specifying the derivative in terms of the function itself, and we want to know how to get to there. So let's continue to do some algebra.

We can do an algebraic operation that seems pointless at first, but which will help us later. We can add a  $+1-1$  to the numerator. This does not affect the value, since the two values cancel themselves out.

$$\frac{da}{dz} = \frac{e^{-z} + 1 - 1}{(1+e^{-z})^2}$$

We can now perform a little bit more algebra as follows:

$$\begin{aligned}
 \frac{da}{dz} &= \frac{1 + e^{-z} - 1}{(1 + e^{-z})^2} \\
 &= \frac{\cancel{1 + e^{-z}}}{(1 + e^{-z})^2} - \frac{1}{(1 + e^{-z})^2} \\
 &= \frac{1}{1 + e^{-z}} - \frac{1}{(1 + e^{-z})^2} \\
 &= a - a^2
 \end{aligned}$$

Which can easily be represented as:

$$\frac{da}{dz} = a(1 - a)$$

And.. voila! We are done!

## What the derivative looks like

We can create a plot that shows the relationship between the sigmoid function and its derivative as follows:

```
import matplotlib.pyplot as plt
import numpy as np

def sigmoid(x):
```

```

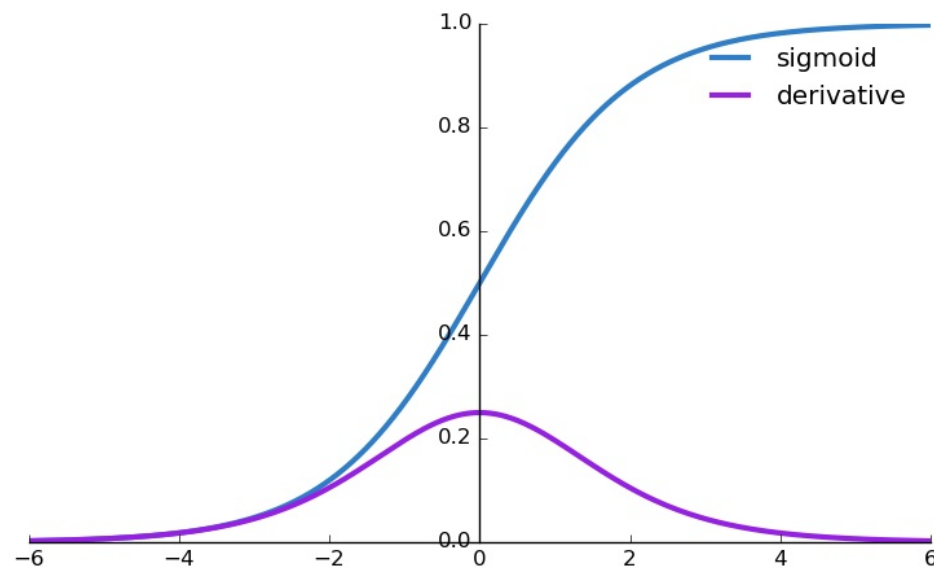
    return 1./(1+np.e**(-x))

# Calculate plot points
x = np.arange(-6., 6., 0.01)
y = sigmoid(x)
dx = y*(1-y)

# Setup centered axes
fig, ax = plt.subplots(figsize=(9, 5))
ax.spines['left'].set_position('center')
ax.spines['right'].set_color('none')
ax.spines['top'].set_color('none')
ax.xaxis.set_ticks_position('bottom')
ax.yaxis.set_ticks_position('left')

# Create and show plot
ax.plot(x,y, color="#307EC7", linewidth=3, label="sigmoid")
ax.plot(x,dx, color="#9621E2", linewidth=3, label="derivative")
ax.legend(loc="upper right", frameon=False)
fig.show()

```





## Comments

Note you can comment without any login by:

1. Typing your comment
2. Selecting "sign up with Disqus"
3. Then checking "I'd rather post as a guest"

3 Comments

ronny.rest

1 Login ▾

♥ Recommend 12

🔗 Share

Sort by Best ▾



Join the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS ?



Name



**Harri Väyrynen** • 5 months ago

Thank you for this page. I tried to understand myself (with help of google) how  $a(1-a)$  is got. Now you give us nice example. Sharing is caring!

1 ^ | v • Reply • Share ›



**Nika** • 7 months ago

Thank you very much! I finally understood it!

1 ^ | v • Reply • Share ›



**Ronny Restrepo** Mod → **Nika** • 7 months ago

Very happy to hear it was helpful for you :)

1 ^ | v • Reply • Share ›

✉ Subscribe

D Add Disqus to your site

🔒 Privacy

DISQUS