

# INTRODUCTION TO CALCULUS

MATH 1A

## Unit 27: Sigmoid function

**27.1.** What is  $\frac{d}{dx} \int_0^x f(t) dt$ ? If  $F$  is an anti-derivative, this is  $\frac{d}{dx}[F(x) - F(0)] = f(x)$ .

<sup>1</sup> There are two aspects of the **fundamental theorem**:

$$\int_0^x f'(t) dt = f(x) - f(0), \quad \frac{d}{dx} \int_0^x f(t) dt = f(x).$$

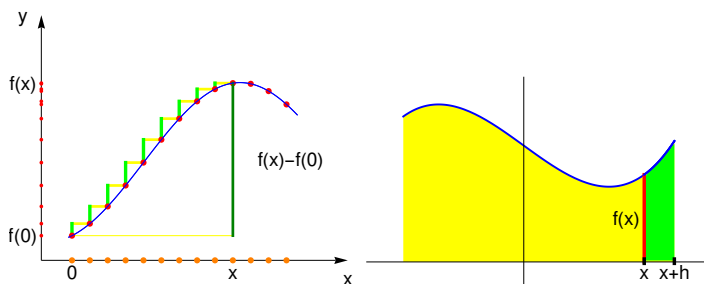


FIGURE 1. Integrate the derivative or differentiate the integral.

**27.2.** The **activation function** for **neural networks** is given by a differentiable function like  $\sigma(x) = (\tanh(x/2) + 1)/2 = e^x/(1 + e^x)$  rather than a **step function**  $(\text{sign}(x) + 1)/2$ . The first one is the **sigmoid function**. You work on this a bit in this homework.

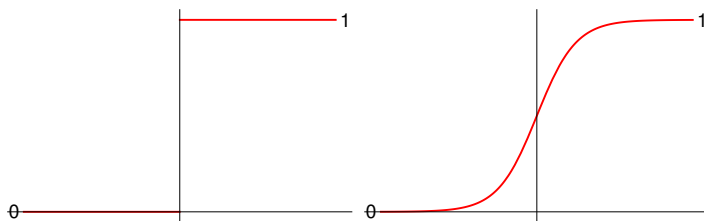


FIGURE 2. The **step function**  $(\text{sign}(x) + 1)/2$  is non-differentiable, the **sigmoid function**  $(\tanh(x/2) + 1)/2 = e^x/(1 + e^x)$  is differentiable.

The reason is that differentiability allows to use gradient descent minimum algorithms (GDM) similarly as the Newton method we have seen to find maxima or minima. Sometimes one sees  $\sigma(x) = \frac{1}{1+e^{-x}}$ . Why is this the same?

<sup>1</sup>All except one student got this wrong in the exam.

# Homework: due April 8, 2024

**Problem 27.1:** Verify  $\sigma(x) = \frac{\tanh(x/2)+1}{2}$ , where  $\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$  and  $\sinh(x) = \frac{e^x - e^{-x}}{2}$  and  $\cosh(x) = \frac{e^x + e^{-x}}{2}$ .

**Problem 27.2:** In this problem we work on the **logistic distribution** in statistics.  
a) Check that  $F(x) = (\tanh(\frac{x}{2}) + 1)/2$  (which by 27.1) is  $\sigma(x)$ ) has the derivative

$$f(x) = \frac{1}{4 \cosh^2(\frac{x}{2})}.$$

It is called the **logistic distribution**.

b) Why is  $\int_{-\infty}^{\infty} f(x) dx = 1$ ? Hint.  $\int_{-a}^a f(x) dx = F(a) - F(-a) = 2 \tanh(x/2)$ .

**Problem 27.3:** The function  $G(x) = \frac{\arctan(x)}{\pi} + \frac{1}{2}$  resembles  $F(x) = \sigma(x)$ . Plot both  $f = F'$  and  $g = G'$  (2 points) then complete the following table (2 points each):

$\frac{d}{dx} \int_0^x \frac{1}{4 \cosh^2(\frac{t}{2})} dt =$		$\int_0^x \frac{1}{4 \cosh^2(\frac{t}{2})} dt =$	
$\frac{d}{dx} \int_0^x \frac{1}{\pi(1+t^2)} dt =$		$\int_0^x \frac{1}{\pi(1+t^2)} dt =$	

**Problem 27.4:** The sigmoid function  $F(x)$  is also called the **standard logistic function** because it satisfies the **logistic equation**  $F'(x) = F(x)(1 - F(x))$ . Verify this. (Compute both  $F'(x)$  and  $F(x)(1 - F(x))$  and compare).

2

**Problem 27.5:** The function  $N(x) = (1 + \operatorname{erf}(x))/2$  looks similar to the sigmoid function. The error function  $\operatorname{erf}$  satisfies  $\operatorname{erf}'(x) = 2e^{-x^2}/\sqrt{\pi}$ .

a) Plot  $n(x) = \operatorname{erf}'(x)$ .

b) Which of the  $f(x), g(x), n(x)$  is the tallest at  $x = 0$ ?

c) Which of the  $f(x), g(x), n(x)$  is the tallest at  $x = 10$ ?

OLIVER KNILL, KNILL@MATH.HARVARD.EDU, MATH 1A, SPRING, 2024

---

<sup>2</sup>This is extremely important in machine learning as the derivative is given in terms of the same function. One uses this in backpropagation.