Differentiation

First Principles

We are interested in the slope of the function at every point.

We want to create a linear function that is tangent to the point, if δx becomes very small, and the function is smooth, then the slope is

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{y(x + \delta x) - y(x)}{\delta x} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}$$

$\frac{d}{dx}$ is an operator, because it *maps* a function $y(x)$ into its derivative $y'(x)$.

Non-differentiable Functions

Functions don't have to be differentiable everywhere.

- $\frac{d|x|}{dx} = 1$ for $x > 0$ and -1 for $x < 0$.
- The derivative does not exist when $x = 0$.
  - If we take the derivative by first principle at $x = 0$, we will get the derivative is 1 at $x = 0$.
  - But that's because we are using the $+\delta x$ definition of the limit, if we do that by the $-\delta x$ definition which is equally valid, we will get -1.
  - The derivative requires both definitions to agree with each other.

The Heaviside Step Function

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases}$$

Consider a graph of a normal distribution so the area under the graph is 1.

- The Heaviside step function is the ingeral of the function from $-\infty$ to x when standard deviation $\sigma \to 0$ so the peak becomes infinitely tall.
- So we have $H'(x) = 0$ everywhere but

$$\int_{-\infty}^{\infty} H'(x) dx = 1$$

Definition

Differentiability: the derivative exists iff

- The function is continuous.
- f' exists and is finite and defined.

Implicit differentiation says if you can get the derivative of a function, you can also very easily take the derivative of the inverse function. We can see this is true by rotating the graph.

Differentiation Rules

We can do **first order approximation**

$$u(x + \delta x) \approx u(x) + \frac{du}{dx} \cdot \delta x = u + \delta u$$

As $\delta x \to 0$ the approximation becomes exact.

$$\frac{df}{du} = \frac{f(u+\delta u) - f(u)}{\delta u}$$
$$\frac{df}{dx} = \left(\frac{f(u+\delta u) - f(u)}{\delta u}\right)\left(\frac{u(x+\delta x) - u(x)}{\delta x}\right)$$

The rules we will use are

Chain rule	$\frac{d}{dx}(f(u(x))) = \frac{df}{du}\frac{du}{dx}$
Product rule	$\frac{d}{dx}(fg) = \frac{df}{dx}g + f\frac{dg}{dx}$
Quotient rule	$\frac{d}{dx}\frac{f}{g} = \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2}$

You are expected to know the proof to these rules.

When differentiating functions like $\frac{\sin x}{x}$, analysis won't be useful because the function is not defined at 0, but we can differentiate it using the quotient rule.

Implicit Differentiation

$$g(y) = h(x)$$

Where y is a function over x, using the chain rule.

$$\frac{d}{dx}g(y(x)) = \frac{d}{dx}h(x)$$
$$\frac{dg}{dy} \cdot \frac{dy}{dx} = \frac{dh}{dx}$$
$$\frac{dy}{dx} = \frac{dh}{dx} \Big/ \frac{dg}{dy}$$

Reciprocal Rule

From the geometric intepretation of the alternative, if we know the derivative of a function at a point, we automatically know the derivative of the inverse at that point.

$$\frac{dx}{dy} = \lim_{\delta x \to 0} \frac{\delta x}{\delta y} = \lim_{\delta y \to 0} \left(\frac{\delta y}{\delta x}\right)^{-1}$$
$$= \left(\lim_{\delta x \to 0} \frac{\delta y}{\delta x}\right)^{-1}$$
$$= \frac{1}{dy/dx}$$

Higher Derivatives

- The first derivative gives the slope of the graph.
- The second derivative gives the curvature of the graph.

Since the derivative is an operator, we can differentiate the function again and have as many derivative as we want, as long as it is differentiable.

Approximation

Higher derivatives is useful for approximating the formula.

With the first derivative

$$y(x + \delta x) \approx y(x) + \delta x \cdot \frac{dy}{dx}$$

This can be extended to a Taylor series.

Leibnitz's Formula

How does the product rule extend if we take multiple derivatives?

$$\begin{aligned} y &= fg \\ y' &= f'g + fg' \\ y'' &= f''g + 2f'g' + fg'' \\ y''' &= f'''g' + 3f''g' + 3f'g'' + fg''' \end{aligned}$$

This looks like **Pascal's triangle**, which is given by the binomial coefficient, so we are suggesting

$$\frac{d^n(fg)}{dx^n} = \sum_{m=0}^{n} \binom{n}{m} f^{n-m} g^m$$

Which can be proved by induction.
