

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

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

$\frac{d}{dx}$ 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 $\sigma \rightarrow 0$ 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 $\delta x \rightarrow 0$ 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.

$$\frac{d}{dx}g(y(x)) = \frac{d}{dx}h(x)$$

$$\frac{dg}{dy} \cdot \frac{dy}{dx} = \frac{dh}{dx}$$

$$\frac{dy}{dx} = \frac{dh}{dx} / \frac{dg}{dy}$$

Reciprocal Rule

From the geometric interpretation 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.

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

$$= \left(\lim_{\delta x \rightarrow 0} \frac{\delta y}{\delta x} \right)^{-1}$$

$$= \frac{1}{dy/dx}$$

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.

Graph Sketching

We want to study the derivative because we figure out the properties of a curve.

Definition

A **stationary point** is where the derivative is zero.

A stationary point can be a maximum, minimum, or an inflection point.

- If $y'' > 0$, then it is a local minimum.
- If $y'' < 0$, then it is a local maximum.
- If $y'' = 0$, then it can be a maximum or negative, or an inflection point.

Consider the Taylor series of the function, the only term that matters is the lowest power term.

- If the first nonzero $y^n(x)$ has n odd then the stationary point is a point of inflection.
- If the first nonzero $y^n(x)$ has n even and positive then the stationary point is a local minimum.
- If the first nonzero $y^n(x)$ has n even and negative then the stationary point is a local maximum.

Points of Inflection

Points of inflection does not have to be stationary points.

- The second derivative has meaning of curvature.
- If the curvature changes sign, then it has to be zero at that point.

When sketching a graph, we want to find

- The x and y intercept: what are the zeros of the function.
- If the function is symmetric.
- The asymptotic behaviour.
- Stationary points and curvature.

The Exponential Function

Definition

The **exponential function** is the only function that is equal to its derivative.

- e^x goes to infinity faster than any power of x .

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^m} = \infty$$

- Conversely, e^{-x} goes to 0 faster than any powers of m , regardless of m .

This is a case of **superpolynomial scaling**.

The Logarithm

Definition

The **logarithm function** is the inverse of the exponential function.

Because the exponential function grows faster than any polynomial, log grows slower than any polynomial.

$$\lim_{\{x \rightarrow \infty\}} \frac{\ln x}{x^m} = 0$$

Independent of m .

END Differentiation