Mathematical Methods and Tools

Fynn Krebser–f.krebser@olympiad.ch

IAAO Training 2026

1 Analysis

1.1 Linearization and the Derivative

The simplest functions we know how to work with are LINEAR FUNCTIONS of the form

$$f(x) = ax + b \quad \text{with } a, b \in \mathbb{R}.$$

Unfortunately, many functions describing the real world are not linear at all. They can be quadratic, exponential or periodic. Our goal in this first part will be to learn how to simplify these functions so we can work with them as if they were linear.

Let us start with a simple example. Consider the function $f(x) = x^2$ depicted in Figure 1. This function is not linear. Our goal is now to find a linear function that approximates f around a point x_0. This is called LINEARIZATION.

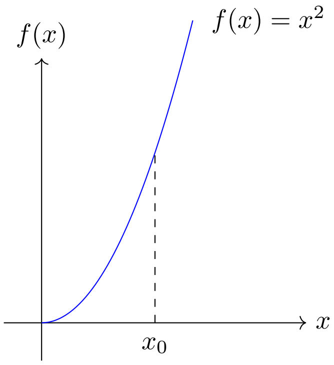


Figure 1: The function $f(x) = x^2$.

One way to achieve this is to find the tangent line to the function at the point x_0. For this we can approximate the slope of the function at x_0 by calculating the slope of the secant line through the points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$ for a small h. The slope of this secant line is given by

$$m = \frac{f(x_0 + h) - f(x_0)}{h}.$$

To get the exact slope of the tangent line, we take the limit as h approaches zero. This is the definition of the DERIVATIVE of f at x_0.

Definition 1.1: Derivative

Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathbb{R}$. The derivative of f at x_0 is defined as

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

provided this limit exists. If the limit does not exist, we say that f is not differentiable at x_0.

With this result, we can now find the equation of the tangent line to f at x_0.

Proposition 1.2: Tangent Line

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function at x_0. The equation of the tangent line to the graph of f at the point $(x_0, f(x_0))$ is given by

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Proof. We use the Ansatz $y = mx + p$ for the equation of the line. Since the slope is $f'(x_0)$, we have $m = f'(x_0)$. The line passes through the point $(x_0, f(x_0))$, so we can plug in these values to find p:

$$f(x_0) = f'(x_0)x_0 + p \Rightarrow p = f(x_0) - f'(x_0)x_0.$$

Hence, the equation of the tangent line is

$$y = f'(x_0)x + f(x_0) - f'(x_0)x_0 = f'(x_0)(x - x_0) + f(x_0).$$

□

We now want to apply this to our function $f(x) = x^2$. We have

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2x\cancel{h} + \cancel{h}^{h^2} - x^2}{\cancel{h}}$$
$$= \lim_{h \to 0} \frac{2xh + h^2}{h} = \lim_{h \to 0} (2x + h) = 2x.$$

The general result for polynomials is the following:

Proposition 1.3: Derivative of a Polynomial

Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial of degree n. Then the derivative of f is given by

$$f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1.$$

Exercise 1.4:

Compute the derivative of the following functions:

1. $f(x) = 3x^3 - 5x^2 + 2x - 7$
2. $g(x) = x^4 + 4x^3 - 2x + 1$
3. $h(x) = 5x^5 - 3x^4 + x^2 - 6$

Solution.

1. $f'(x) = 9x^2 - 10x + 2$
2. $g'(x) = 4x^3 + 12x^2 - 2$
3. $h'(x) = 25x^4 - 12x^3 + 2x$

Some other important rules for derivatives are the following:

Lemma 1.5: Derivative Rules

Let f and g be differentiable functions. Then the following rules hold:

1. **Sum Rule:**
$$(f+g)'(x) = f'(x) + g'(x)$$

2. **Product Rule:**
$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

3. **Quotient Rule:**
$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$
, provided $g(x) \neq 0$

4. **Chain Rule:** If $h(x) = f(g(x))$, then
$$h'(x) = f'(g(x))g'(x)$$

Other derivatives, often seen involve

Lemma 1.6: Common Derivatives

It holds that

1. $\frac{d}{dx}e^x = e^x$
2. $\frac{d}{dx}\ln(x) = \frac{1}{x}$, for $x > 0$
3. $\frac{d}{dx}\sin(x) = \cos(x)$
4. $\frac{d}{dx}\cos(x) = -\sin(x)$
5. $\frac{d}{dx}\tan(x) = \sec^2(x)$, for $x \neq \frac{\pi}{2} + k\pi, k \in \mathbb{Z}$

1.2 Numerical Methods

Let's see an application of this. In particular, we want to find zeroes of a function f. For example the function $f(x) = \cos(x) - x$ has a zero somewhere between $x = 0$ and $x = 1$. We want to find this zero more precisely using numerical methods.

Proposition 1.7: Bisection Method

Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a)f(b) < 0$. Then there exists a zero $c \in (a, b)$ of f. The bisection method is an iterative method to approximate this zero. The algorithm is as follows:

1. Compute the midpoint $m = \frac{a+b}{2}$.
2. If $f(m) = 0$, then m is the zero and we are done.
3. If $f(a)f(m) < 0$, then set $b = m$; otherwise, set $a = m$.
4. Repeat steps 1-3 until the desired accuracy is achieved.

Theorem 1.8: Newton's Method

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and let x_0 be an initial guess for a zero of f. Then the sequence defined by
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
converges to a zero of f, provided that $f'(x_n) \neq 0$ for all n.

This method is an application of the tangent line, we discussed before. We use the fact that it is much simpler to find the zero of a linear function than a non-linear one. The idea is to start with an initial guess x_0 and then use the tangent line at that point to find a better approximation x_1. We repeat this process until we reach a satisfactory approximation of the zero.

Example 1.9:

We want to find a zero of the function $f(x) = \cos(x) - x$ using Newton's method. We start with an initial guess $x_0 = 3$. The derivative of f is $f'(x) = -\sin(x) - 1$. We compute the next approximations:

$$\begin{aligned}
x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} = 3 - \frac{\cos(3) - 3}{-\sin(3) - 1} \approx -0.4966 \\
x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \approx 2.131 \\
x_3 &= x_2 - \frac{f(x_2)}{f'(x_2)} \approx 0.6897 \\
x_4 &= x_3 - \frac{f(x_3)}{f'(x_3)} \approx 0.7397 \\
x_5 &= x_4 - \frac{f(x_4)}{f'(x_4)} \approx 0.7391 \\
x_6 &= x_5 - \frac{f(x_5)}{f'(x_5)} \approx 0.7391.
\end{aligned}$$

It is crucial to note, that the convergence of Newton's method is strongly dependent on the choice of the initial guess x_0. A poor choice can lead to divergence or convergence to a different zero than intended.

Exercise 1.10:

Find all zeros of the function $f(x) = x^3 - x - 1$.

Solution. The only real solution is $x_0 \approx 1.3247$.

1.3 Area under the curve and the Riemann Integral

One method to compute the area under a curve is to plot the curve, and count squares. This is a very crude method, but it can give us a rough estimate of the area. Our goal here is to develop a more precise method to compute the area under a curve. We will do this by approximating the area under the curve by a sum of rectangles. This is called a RIEMANN SUM.

Definition 1.11: Riemann Sum

Let $f : [a,b] \to \mathbb{R}$ be a function and let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of the interval $[a,b]$, where $a = x_0 < x_1 < \cdots < x_n = b$. For each $i = 1, \ldots, n$, let $\xi_i \in [x_{i-1}, x_i]$ be a sample point. The Riemann sum of f with respect to the partition P and the sample points ξ_i is defined as

$$S(f, P, \xi) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

Notice, that for n large, this becomes a very good approximation of the area under the curve. In fact, we can make this approximation arbitrarily good by taking the limit as the number of rectangles goes to infinity. This leads us to the definition of the RIEMANN INTEGRAL.

Definition 1.12: Riemann Integral

Let $f : [a,b] \to \mathbb{R}$ be a function. We say that f is Riemann integrable on $[a,b]$ if the limit

$$\int_a^b f(x)\,dx = \lim_{||P|| \to 0} S(f, P, \xi)$$

exists, where $||P||$ is the norm of the partition P. The value of this limit is called the Riemann integral of f from a to b.

In general, using Riemann sums to compute the integral is not very practical. However, the following theorem simplifies the computation of integrals significantly.

Theorem 1.13: Fundamental Theorem of Calculus

Let $f : [a,b] \to \mathbb{R}$ be continuous. Then, $\forall C \in \mathbb{R}$, the function

$$F(x) = \int_a^x f(t)\,dt + C$$

is a primitive of f, i.e. $F'(x) = f(x)$. Also, all primitives of f are of this form.

Using this theorem, we can compute the integral of a function by finding its primitive. This is often much easier than computing the limit of Riemann sums.

Exercise 1.14:

Compute the following integrals:

1. $\int_0^1 x^2\,dx$
2. $\int_0^\pi \sin(x)\,dx$
3. $\int_1^2 e^x\,dx$

Solution.

1. $\frac{1}{3}$
2. 2
3. $e^2 - e$

1.4 Numeric Integration

In many cases, it is not possible to find a primitive of a function, or the function is only known through data points. In these cases, we can use numerical methods to approximate the integral. The most common method is the TRAPEZOIDAL RULE.

Proposition 1.15: Trapezoidal Rule

Let $f : [a,b] \to \mathbb{R}$ be a continuous function and let n be a positive integer. We divide the interval $[a,b]$ into n subintervals of equal length $\Delta x = \frac{b-a}{n}$. The trapezoidal rule states that

$$\int_a^b f(x)\,dx \approx \frac{\Delta x}{2}\left(f(a) + 2\sum_{i=1}^{n-1} f(a + i\Delta x) + f(b)\right).$$

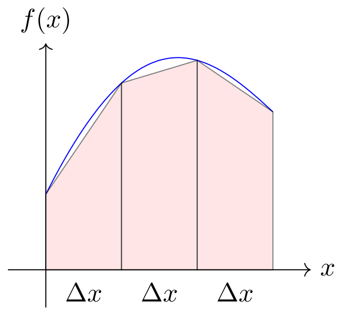


Figure 2: Visualization of the trapezoidal rule using 3 subintervals.

Exercise 1.16:

Approximate the integrals from exercise 1.14 using the trapezoidal rule.

1.5 Taylor Series

We started this chapter by discussing linearization. We can extend this idea to approximate a function by a polynomial of higher degree. This leads us to the concept of Taylor series.

Definition 1.17: Taylor Series

Let $f : \mathbb{R} \to \mathbb{R}$ be a function that is infinitely differentiable at a point $a \in \mathbb{R}$. The Taylor series of f at a is given by

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Exercise 1.18:

Find the Taylor series up to degree 5 of the following functions at $a = 0$:

1. $f(x) = e^x$
2. $g(x) = \sin(x)$
3. $h(x) = \cos(x)$

Solution.

1. $T_f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
2. $T_g(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$
3. $T_h(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$

Oftentimes, it is sufficient to compute the first few terms of the Taylor series to get a good approximation of the function. The more terms we include, the better the approximation will (in general) be. For example, the Taylor series of $\log(1+x)$ at $a = 0$ is given by

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots .$$

However, this series only converges for $|x| < 1$. For $|x| \geq 1$, the series diverges. This is an important point to keep in mind when working with Taylor series.

> **Remark 1.19:**
>
> The Taylor series computed in the exercise above, all have infinite radius of convergence. This means that the Taylor series converges to the function for all $x \in \mathbb{R}$.

2 Statistics

2.1 Important Quantities

Before doing statistics, we first have to gather data. This can be done by conducting experiments, surveys or by collecting data from existing sources. Once we have the data, we can compute some important quantities that will help us to understand the data better.

> **Definition 2.1: Data**
>
> A DATA SET is a collection of data points $\mathcal{X}$. Each data point can be a number, a string or any other type of data.

In general, before doing data analysis, we will tabularize the data for example as follows: Especially for time dependent data, it is often useful to plot the data points in a graph. This can help us to visualize the data and to see patterns that might not be obvious from the table alone. Some important quantities that we can compute from the data are the MEAN, the MEDIAN and the MODE.

Time t (s)	**Position x (m)**
0	0
1.0	1.1
1.4	1.5
2.0	3.8
3.1	9.2

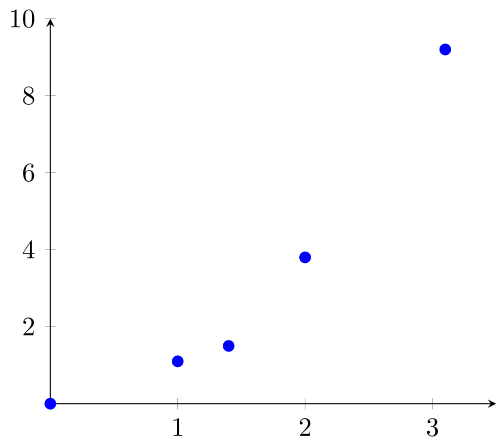


Figure 3: Plot of the data points.

> **Definition 2.2: Mean**
>
> The MEAN of a data set $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ is defined as
> $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

> **Definition 2.3: Median**
>
> The MEDIAN of a data set $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ is the middle value when the data points are sorted in ascending order. If n is odd, the median is the middle value. If n is even, the median is the average of the two middle values.

Definition 2.4: Mode

The MODE of a data set $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ is the value that occurs most frequently in the data set. A data set can have no mode, one mode or multiple modes.

These quantities all describe some sort of average of the data set. The mean is the most commonly used measure of central tendency, but it can be affected by outliers. The median is more robust to outliers, while the mode can be useful for categorical data.

To describe the spread of the data we typically use the STANDARD DEVIATION

Definition 2.5: Standard Deviation

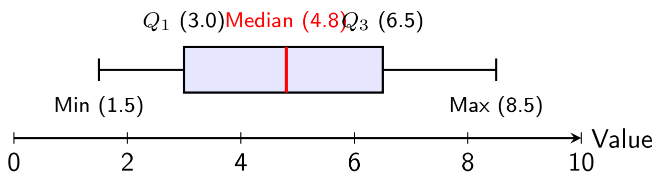
The STANDARD DEVIATION of a data set $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ is defined as

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Where $\bar{x}$ is the mean of the data set.

Another common way to describe the spread of the data is using PERCENTILES. The p-th percentile is the value below which $p\%$ of the data points fall. For example, the 25th percentile is the value below which 25% of the data points fall. The 50th percentile is the median, and the 75th percentile is the value below which 75% of the data points fall.

Using percentiles also allows us to create BOX PLOTS, which are a graphical representation of the data set. A box plot shows the minimum, maximum, median, 25th percentile and 75th percentile of the data set. The box represents the interquartile range (IQR), which is the range between the 25th and 75th percentiles. The whiskers represent the minimum and maximum values.



2.2 Error Calculations

If one calculates a quantity Q from measured quantities $x, y, z, \dots$ then estimating the error of Q is not trivial. To do this, we can use the Gaussian error propagation. For this we first determine our error in $x, y, z, \dots$ by any method we like. Then, we can calculate the error of Q by the following formula

Theorem 2.6: Gaussian Error Propagation

Given a quantity Q which is a function of measured quantities $x, y, z, \dots$. Then, the error of Q is given by

$$\Delta Q = \sqrt{\left(\frac{\partial Q}{\partial x}\Delta x\right)^2 + \left(\frac{\partial Q}{\partial y}\Delta y\right)^2 + \dots}$$

The notation $\frac{\partial Q}{\partial x}$ is the partial derivative of Q with respect to x. For our purposes, it is just a derivative of Q treating everything other than x as a constant. Δx is the error of x.

Exercise 2.7:

A mass m is attached to a spring with spring constant k. The period of the oscillation is given by

$$T = 2\pi\sqrt{\frac{m}{k}}.$$

If we measure T and m with errors ΔT and Δm, how can we determine the error of k?