Getting started: Review of Calculus

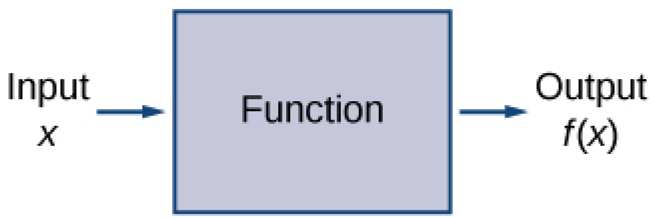
Calculus is a big branch of mathematics. If you are taking this course, then you have already taken Calculus I and II.

The following is a brief review of concepts needed in the context of Data Science and Machine Learning. For a complete review, please follow the sources provided at the end of this document.

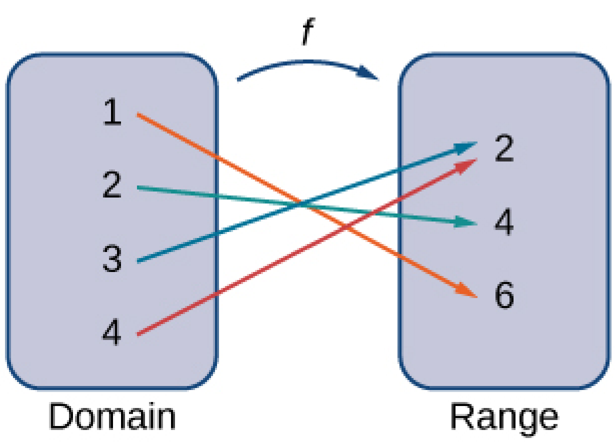
If there are any concepts in this document that you are unfamiliar or *rusty* about, please review them as soon as possible.

Functions and Graphs

A **function** consists of a set of inputs, a set of outputs, and a rule for assigning each input to exactly one output.



- The set of inputs is called the **domain** of the function. The set of outputs is called the **range** of the function.



A general function f with domain D, we often use x to denote the input and y to denote the output associated with x.

- When doing so, we refer to x as the **independent variable** and y as the **dependent variable**, because it depends on x. Using function notation, we write

  $$y = f(x),$$

  and we read this equation as y *equals* f *of* x.

Typically, a function is represented using one or more of the following tools:

- A table
- A graph
- An algebraic formula

Combining Functions

We can create a new function by composing two functions. Consider the functions $f(x) = 3x + 1$ and $g(x) = x^2$:

- The **sum** function: $(f + g)(x) = f(x) + g(x) = x^2 + 3x + 1$

- The **difference** function: $(f - g)(x) = f(x) - g(x) = -x^2 + 3x + 1$

- The **product** function: $(f \cdot g)(x) = f(x)g(x) = x^2(3x + 1)$

- The **quotient** function: $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} \text{ for } g(x) \neq 0 = \frac{3x+1}{x^2}$

- The **composite** function: $(f \circ g)(x) = f(g(x)) = f(x^2) = 3x^2 + 1$

Linear Function

Linear functions have the form:

$$y = mx + b,$$

where m and b are constants.

A linear function is defined by its slope, m and y-intercept, b.

For example, consider $y = 3x + 1$:

- The function y describes a **line** with **slope** $m = 3$.
  - The slope measures both the steepness and the direction of a line.

- The function's y-**intercept** is equal to $b = 1$.

The **standard form of a line** is given by the equation:

$$ax + by = c,$$

where a and b are both not zero. This form is more general because it allows for a vertical line, $x = k$.

Polynomials

A linear function is a special type of a more general class of functions: polynomials. A polynomial function is any function that can be written in the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

for some integer $n \geq 0$ and constants $a_n, a_{n-1}, \ldots, a_0$, where $a_n \neq 0$.

- The value n is called the **degree** (or **order**).

- A polynomial function of degree 2 is called a **quadratic function**: $f(x) = ax^2 + bx + c$, where $a \neq 0$.

- A polynomial function of degree 3 is called a **cubic function**.

Finding Zeros

The solutions x' that solve the equation $f(x') = 0$ are called **zeros** because they intersect the x-axis.

The zeros of the quadratic equation $f(x) = ax^2 + bx + c$ where $a \neq 0$ are given by the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Mathematical Models

In engineering, we often describe a large variety of real-world situations using **mathematical models**.

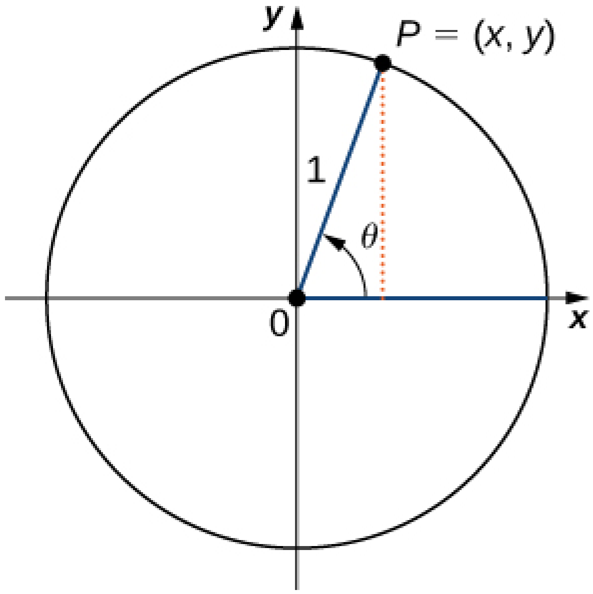
- A mathematical model is a method of simulating real-life situations with mathematical equations.

- Models are useful because they help predict future outcomes. Examples of mathematical models include the study of population dynamics, investigations of weather patterns, and predictions of product sales.

Trigonometric functions

Trigonometric functions are used to model many phenomena, including sound waves, vibrations of strings, alternating electrical current, and the motion of pendulums.

We first consider the *unit circle* centered at the *origin* and a point $P = (x, y)$ on the unit circle. Let θ be an angle between the positive x-axis and with the line segment OP.



- $\sin\theta = y$
- $\cos\theta = x$
- $\tan\theta = \frac{y}{x}$
- $\csc\theta = \frac{1}{y}$
- $\sec\theta = \frac{1}{x}$
- $\cot\theta = \frac{x}{y}$

On a circle of radius r with a corresponding angle θ, the coordinates x and y satisfy:

$$\cos\theta = \frac{x}{r}$$

$$x = r\cos\theta$$

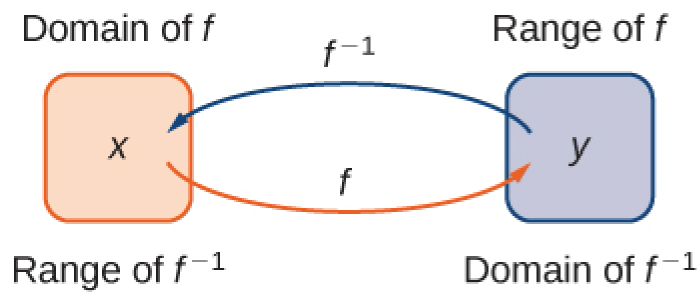
and

$$\sin\theta = \frac{y}{r}$$

$$y = r\sin\theta$$

Inverse functions

An inverse function reverses the operation done by a particular function.



Given a function f and an output $y = f(x)$, we are often interested in finding what value or values x were mapped to y by f. We denote the **inverse function** as f^{-1}.

- We say that the f is a **one-to-one function** (or **injective function**) if $f(x_1) \neq f(x_2)$ where $x_1 \neq x_2$.

Exponential and Logarithmic Functions

Any function of the form

$$f(x) = b^x,$$

where $b > 0, b \neq 1$, is an **exponential function** with **base** b and **exponent** x.

- The exponential $f(x) = b^x$ is one-to-one, with domain $(-\infty, \infty)$ and range $(0, \infty)$.

- We call the function

  $$f(x) = e^x$$

  the **natural exponential function**, where the **number** $e \approx 2.718282$.

- The **logarithmic function** with base b is the inverse function of the exponential function

- For any $b > 0, b \neq 1$, the logarithmic function with base b is defined as

$$y = \log_b(x)$$

with domain $(0, \infty)$ and range $(-\infty, \infty)$.

- The **natural logarithmic function** with base e is the inverse function of the natural exponential function

$$y = \log_e(x) = \ln(x)$$

Limits

Let $f(x)$ be a function defined at all values in an open interval containing a,, with the possible exception of a itself, and let L be a real number.

If all values of the function $f(x)$ approach the real number L as the values of x ($\neq a$) approach the number a, then we say that the limit of $f(x)$ as x approaches a is L.

Symbolically, we express this idea as

$$\lim_{x \to a} f(x) = L$$

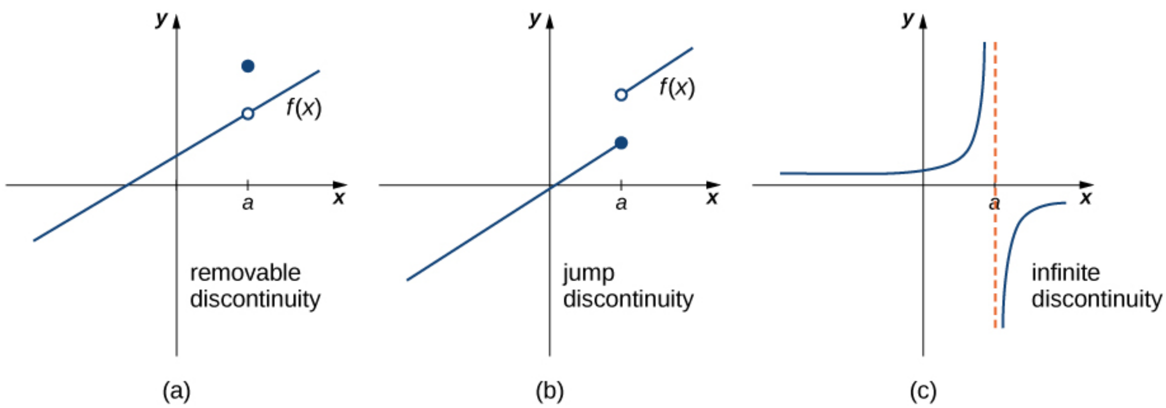
Continuity

A function $f(x)$ is **continuous at a point** a if and only if the following three conditions are satisfied:

1. $f(a)$ is defined
2. $\lim_{x \to a} f(x)$ exists
3. $\lim_{x \to a} f(x) = f(a)$

A function is **discontinuous at a point** a if it fails to be continuous at a.

Types of Discontinuities



Tangent Line

Let $f(x)$ be a function defined in an open interval containing a. The **tangent line** to $f(x)$ at a is the line passing through the point $(a, f(a))$ having slope

$$m_{\tan} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

Equivalently, we may define the tangent line to $f(x)$ at a to be the line passing through the point $(a, f(a))$ having slope

$$m_{\tan} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

provided this limit exists.

Derivative

Let $f(x)$ be a function defined in an open interval containing a. The **derivative of the function $f(x)$ at** a, denoted by $f'(a)$, is defined by

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

Alternatively, we may also define the **derivative of** $f(x)$ **at** a as

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

provided this limit exists.

For $y = f(x)$, each of the following notations represents the derivative of $f(x)$:

$$f'(x), \frac{dy}{dx}, y', \frac{d}{dx}(f(x))$$

- The **instantaneous rate of change** (or instantaneous velocity) of a function $f(x)$ at a value a is its derivative:

$$v_{\text{inst}} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

- The **average rate of change** (average velocity) over an interval $[a, x]$ if $x > a$ or $[x, a]$ if $x < a$ is given by the different quotient

$$v_{\text{ave}} = \frac{f(x) - f(a)}{x - a}$$

Differentiation Rules

There are a few differentiation rules, including:

- The Constant Rule
- The Power Rule
- The Sum, Difference, and Constant Multiple Rules
- The Product Rule
- The Quotient Rule

For definitions, please visit this section.

The Chain Rule

Let f and g be functions. For all x in the domain of g for which g is differentiable at x and f is differentiable at $g(x)$, the derivative of the composite function

$$h(x) = (f \circ g)(x) = f(g(x))$$

is given by

$$h'(x) = f'(g(x))g'(x).$$

Alternatively, if y is a function of u, and u is a function of x, then

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

Some applications of Derivatives

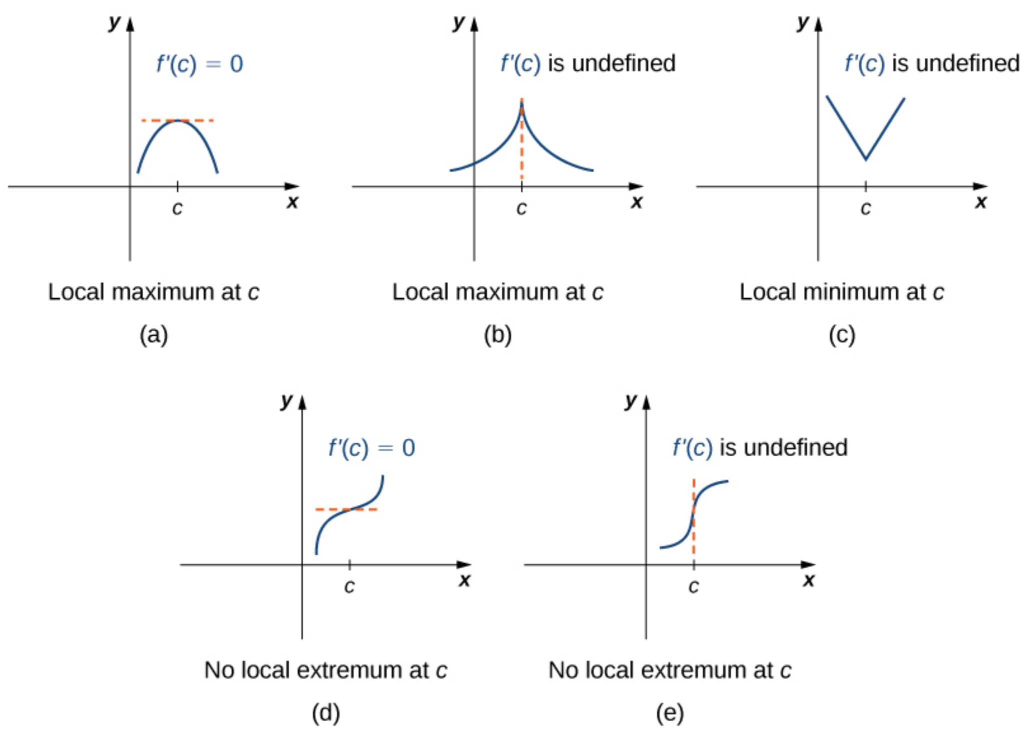
1. **Linear Approximations of a Function at a Point**

In general, for a differentiable function f, the equation of the tangent line to f at $x = a$ ($y = f(a) + f'(a)(x - a)$) can be used to approximate $f(x)$ for $xnear$a\$:

$$f(x) \approx f(a) + f'(a)(x - a)$$

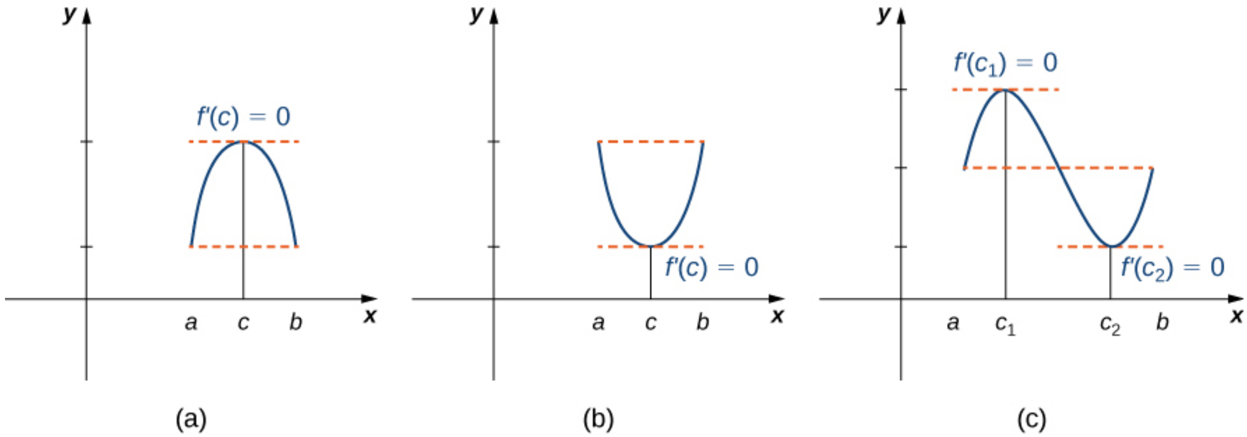
1. **Critical points**

Let c be an interior point in the domain of f. We say that c is a **critical point** of f if $f'(c) = 0$ or $f'(c)$ is undefined.



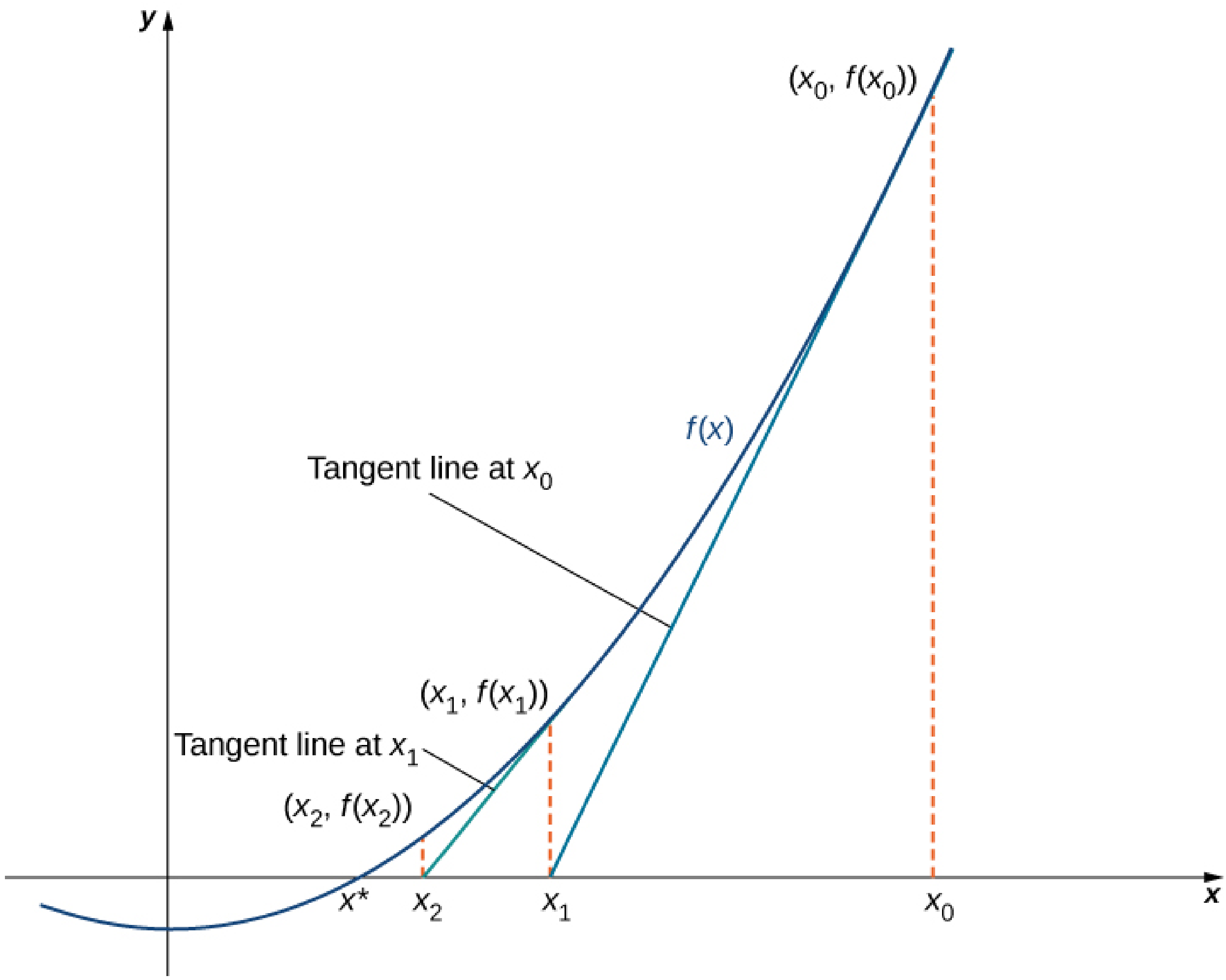
1. **Rolle's Theorem**

Let f be a continuous function over the closed interval $[a, b]$ and differentiable over the open interval (a, b) such that $f(a) = f(b)$. There then exists at least one $c \in (a, b)$ such that $f'(c) = 0$.



1. **Newton's Method**

It's an algorithm to approximate solutions of $f(x) = 0$.



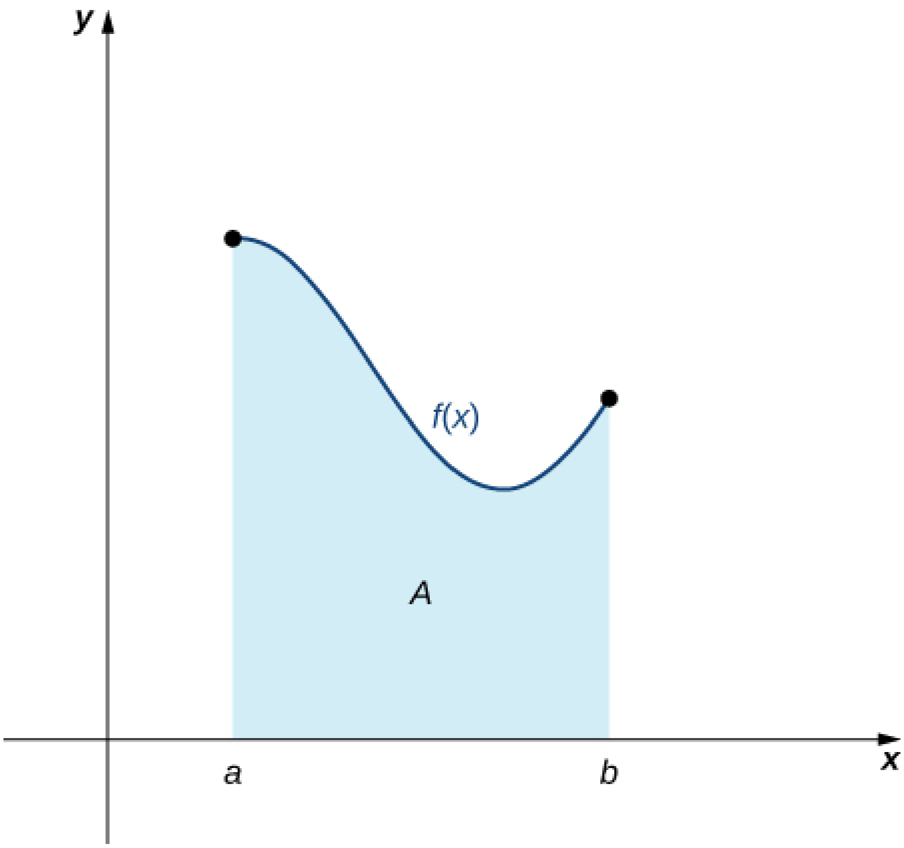
The equation of this tangent line is $y = f(x_0) + f'(x_0)(x - x_0)$

Therefore, x_1 must satisfy: $f(x_0) + f'(x_0)(x_1 - x_0) = 0$, then $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

In general, for $n > 0$, x_n satisfies

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

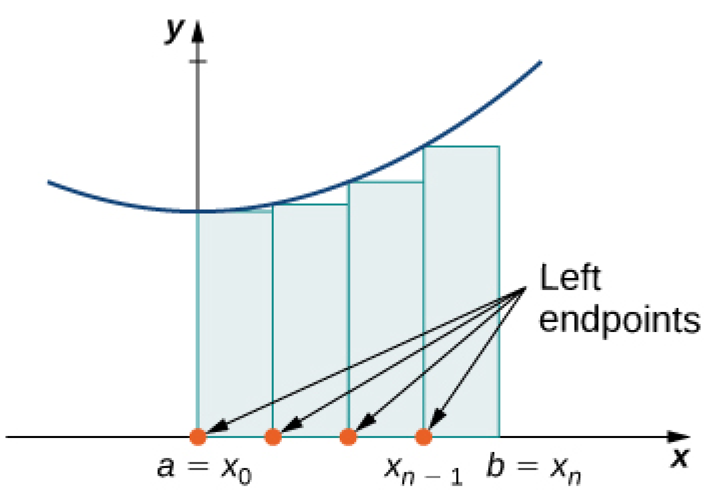
Approximating Areas



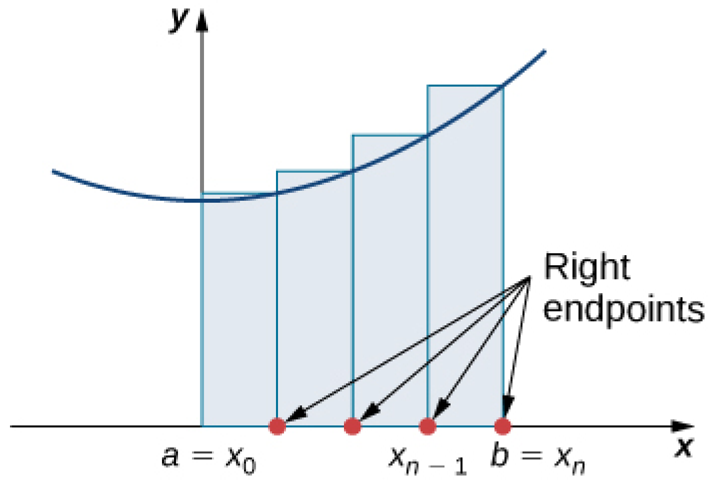
How do we approximate the area under this curve? The approach is a geometric one. By dividing a region into many small shapes that have known area formulas, we can sum these areas and obtain a reasonable estimate of the true area.

Denote the width of each subinterval as $\Delta x = \frac{b-a}{n}$.

1. **Left-endpoint approximation**



1. **Right-endpoint approximation**



Left-endpoint approximation:
$A \approx f(x_0)\Delta x + f(x_1)\Delta x + \cdots + f(x_{n-1})\Delta x = \sum_{i=1}^{n} f(x_{i-1})\Delta x$

Right-endpoint approximation: $A \approx f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x = \sum_{i=1}^{n} f(x_i)\Delta x$

In both cases, these are called **Riemann sums**.

Area under the Curve (AUC)

Let $f(x)$ be a continuous, nonnegative function on an interval $[a, b]$, and let $\sum_{i=1}^{n} f(x_i^*)\Delta x$ be a Riemmann sum for $f(x)$. The, the **area under th curve** $y = f(x)$ on $[a, b]$ is given by

$$A = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i^*)\Delta x$$

Definite Integral

If $f(x)$ is a function defined on an interval $[a, b]$, the **definite integral** of f from a to b is given by

$$\int_a^b f(x)dx = \lim_{b \to \infty} \sum_{i=1}^{n} f(x_i^*)\Delta x$$

provided the limit exists. If this limit exists, the function $f(x)$ is said to be **integrable** on $[a, b]$, or is an integrable function.

- Continuous functions are integrable.

Fundamental Theorem of Calculus

(It establishes the relationship between differentiation and integration.)

If $f(x)$ is continuous over an interval $[a, b]$, and the function $F(x)$ is defined by

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

then $F'(x) = f(x)$ over $[a, b]$.

The Evaluation Theorem

If f is continuous over the interval $[a, b]$ and $F(x)$ is any antiderivative of $f(x)$, then

$$\int_a^b f(x)dx = F(b) - F(a)$$

We often see the notation $F(x)|_a^b$ to denote $F(b) - F(a)$.

Note: The indefinite integral without bounds represents an antiderivative.

The Net Change Theorem

The new value of a changing quantity equals the initial value plus the integral of the **rate of change**:

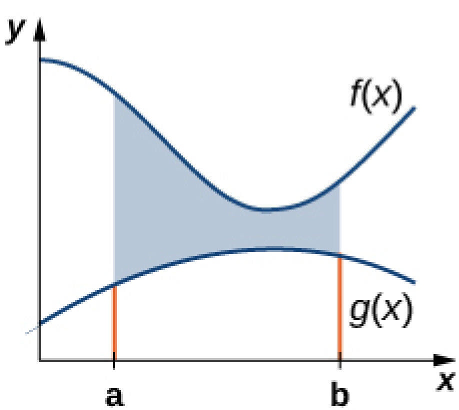
$$F(b) = F(a) + \int_a^b F'(x)dx$$

$$\text{or}$$

$$\int_a^b F'(x)dx = F(b) - F(a)$$

Applications of Integration

1. **Area between two curves**



Let $f(x)$ and $g(x)$ be continuous functions such that $f(x) \geq g(x)$ over an interval $[a, b]$. Let R denote the region bounded above by the graph of $f(x)$, below by the graph of $g(x)$, and on the left and right by the lines $x = a$ and $x = b$, respectively. Then, the area of R is given by

$$A = \int_a^b (f(x) - g(x))\, dx$$

1. **Arc Length of a Curve** $y = f(x)$

Let $f(x)$ be a smooth function over the interval $[a, b]$. Then the arc length of the portion of the graph of $f(x)$ from the point $(a, f(a))$ to the point $(b, f(b))$ is given by

$$\text{Arc Length} = \int_a^b \sqrt{1 + \left(f'(x)\right)^2} dx$$

1. and many more.

Parametric Equations

If x and y are continuous functions of t on an interval I, then the equations

$$x = x(t) \text{ and } y = y(t)$$

are called **parametric equations** and t is called a **parameter**. The set of points (x, y) obtained as t varies over the interval I is called the graph of the parametric equations. The graph of parametric equations is called a **parametric curve** or plane curve, and is denoted by C.

- In many Data Science and Machine Learning applications, we seek to find such parametrization.

Derivative of Parametric Equations

Consider the plane curve defined by the parametric equations $x = x(t)$ and $y = y(t)$. Suppose that $x'(t)$ and $y'(t)$ exist, and assume that $x'(t) \neq 0$. Then the derivative $\frac{dy}{dx}$ is given by

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{y'(t)}{x'(t)}$$

Second-order derivatives of Parametric Equations

The second derivative of a function $y = f(x)$ is defined to be the derivative of the first derivative; that is,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left[\frac{dy}{dx}\right].$$

Area under Parametric Curve

Consider the non-self-intersecting plane curve defined by the parametric equations

$$x = x(t), y = y(t), a \leq t \leq b$$

and assume that $x(t)$ is differentiable. The area under this curve is given by

$$A = \int_a^b y(t)x'(t)dt$$

Further review

If you are interested in reviewing these topics further, I recommend:

- 3Blue1Brown, "Essence of Calculus" YouTube series, link
- "Calculus Volume 1" book, available online
- "Calculus Volume 2" book, available online
- Gilbert Strang, 18.005 MITOpenCourseWare "Highlights of Calculus", link

(The images in this notebook were obtained from "Calculus Volume 1" and "Calculus Volume 2" books.)