M.S. Math Bootcamp: Calculus

Noah Kochanski

Department of Statistics
University of Michigan

08/18/2024

Acknowledgements

This material has been adapted from

- Spivak's *Calculus*, 4th ed.
- Stewart's *Calculus*, 7th ed.
- Thomas's *Calculus*, 10th ed.
- Rudin's Principles of Mathematical Analysis
- Previous bootcamps offered by other PhD students.

Bootcamp Structure

Broadly, we'll split the topics into:

1. Day 1: Calculus
2. Day 2: Linear Algebra
3. Day 3: Statistics

Although the first two days will be a bit more "math-based", we'll try to stick to concepts/examples that you'll find especially useful in your studies in statistics.

We'll use about half the time for lecture, and half the time for problem-solving.

Outline

1. Functions
2. Continuity & Differentiability
3. Integration
4. Multivariable Calculus

Functions

Sets

- A *set* is an unordered collection of objects. The members of a set are called *elements*.

Sets

- A *set* is an unordered collection of objects. The members of a set are called *elements*.
- We often denote sets with curly braces, e.g.

$$A = \{x, y, z\} \tag{1}$$

Sets

- A *set* is an unordered collection of objects. The members of a set are called *elements*.
- We often denote sets with curly braces, e.g.

$$A = \{x, y, z\} \tag{1}$$

- We say that a set B is a *subset* of A if every element in B is also in A. We write this as $B \subseteq A$.

Sets

- A *set* is an unordered collection of objects. The members of a set are called *elements*.
- We often denote sets with curly braces, e.g.

$$A = \{x, y, z\} \tag{1}$$

- We say that a set B is a *subset* of A if every element in B is also in A. We write this as $B \subseteq A$.
- We use the "$\in$" symbol to denote set membership, e.g. $x \in A$ but $w \notin A$.

Set Operations

- The *intersection* of sets A, B is the set of elements x such that $x \in A$ and $x \in B$
  - Denoted $A \cap B$.

Set Operations

- The *intersection* of sets A, B is the set of elements x such that $x \in A$ and $x \in B$
  - Denoted $A \cap B$.
- The *union* of sets A, B is the set of elements x such that $x \in A$ or $x \in B$.
  - Denoted $A \cup B$.

Set Operations

- The *intersection* of sets A, B is the set of elements x such that $x \in A$ and $x \in B$
  - Denoted $A \cap B$.
- The *union* of sets A, B is the set of elements x such that $x \in A$ or $x \in B$.
  - Denoted $A \cup B$.
- The *difference* of sets A, B is the set of elements x such that $x \in A$ but $x \notin B$.
  - Denoted $A \setminus B$.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.
The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.
- The set of real-valued n-dimensional vectors, $\mathbb{R}^n$.

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.

The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.
- The set of real-valued n-dimensional vectors, $\mathbb{R}^n$.
- The set of natural numbers, $\mathbb{N}$.
    - E.g., 1,2,3,4,...

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.
The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.
- The set of real-valued n-dimensional vectors, $\mathbb{R}^n$.
- The set of natural numbers, $\mathbb{N}$.
    - E.g., $1,2,3,4,\ldots$
- The set of integers, $\mathbb{Z}$.

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.
The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.
- The set of real-valued *n*-dimensional vectors, $\mathbb{R}^n$.
- The set of natural numbers, $\mathbb{N}$.
    - E.g., 1,2,3,4,...
- The set of integers, $\mathbb{Z}$.
- The set of complex numbers $\mathbb{C}$.

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.
The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Important Sets To Know

- The set of real numbers, $\mathbb{R}$.
- The set of real-valued n-dimensional vectors, $\mathbb{R}^n$.
- The set of natural numbers, $\mathbb{N}$.
  - E.g., $1, 2, 3, 4, \ldots$
- The set of integers, $\mathbb{Z}$.
- The set of complex numbers $\mathbb{C}$.
- The set of rational numbers, $\mathbb{Q}$.

Often one of these sets will be the *universal set* for a problem of interest, i.e. the set that contains all elements under study.

The *complement* of a set A is the set of elements x in the universal set, but $x \notin A$.

- Denoted A^C.

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b, \text{ including } a \text{ and } b\}$

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b\text{, including } a \text{ and } b\}$
- $(\mathrm{a,b}) = \{\text{all real numbers between } a \text{ and } b\text{, excluding } a \text{ and } b\}$

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b\text{, including } a \text{ and } b\}$
- $(a, b) = \{\text{all real numbers between } a \text{ and } b\text{, excluding } a \text{ and } b\}$
- $[a, b) =$
  $\{\text{all real numbers between } a \text{ and } b\text{, including } a\text{, excluding } b\}$

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b\text{, including } a \text{ and } b\}$
- $(a, b) = \{\text{all real numbers between } a \text{ and } b\text{, excluding } a \text{ and } b\}$
- $[a, b) =$ $\{\text{all real numbers between } a \text{ and } b\text{, including } a\text{, excluding } b\}$
- $(a, b] =$ $\{\text{all real numbers between } a \text{ and } b\text{, excluding } a\text{, including } b\}$

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b\text{, including } a \text{ and } b\}$
- $(a, b) = \{\text{all real numbers between } a \text{ and } b\text{, excluding } a \text{ and } b\}$
- $[a, b) =$ $\{\text{all real numbers between } a \text{ and } b\text{, including } a\text{, excluding } b\}$
- $(a, b] =$ $\{\text{all real numbers between } a \text{ and } b\text{, excluding } a\text{, including } b\}$
- $(a, \infty) = \{\text{all real numbers between } > a\}$

Interval Notation

Some of the most common sets to come across are continuous regions of $\mathbb{R}$. We have special notation for these sets.

- $[a, b] = \{\text{all real numbers between } a \text{ and } b, \text{ including } a \text{ and } b\}$
- $(a, b) = \{\text{all real numbers between } a \text{ and } b, \text{ excluding } a \text{ and } b\}$
- $[a, b) =$
  $\{\text{all real numbers between } a \text{ and } b, \text{ including } a, \text{ excluding } b\}$
- $(a, b] =$
  $\{\text{all real numbers between } a \text{ and } b, \text{ excluding } a, \text{ including } b\}$
- $(a, \infty) = \{\text{all real numbers} > a\}$
- $(-\infty, b) = \{\text{all real numbers} < b\}$

Try It

Let $A = [1, 3]$, $B = [2, 4]$, $C = [3, 5]$, $D = [4, 6]$ (all of which are subsets of the universal set $\mathbb{R}$).

1. What is $A \setminus B$?
2. What is $A \cup B$? $A \cap B$?
3. What is $A \cap C$?
4. What is $A \cap D$?
5. What is $(A \cup D)^C$?

Try It

Let $A = [1, 3]$, $B = [2, 4]$, $C = [3, 5]$, $D = [4, 6]$ (all of which are subsets of the universal set $\mathbb{R}$).

1. What is $A \setminus B$? Answer: $[1, 2)$
2. What is $A \cup B$? $A \cap B$? Answer: $[1, 4]$, $[2, 3]$
3. What is $A \cap C$?
4. What is $A \cap D$?
5. What is $(A \cup D)^C$?

Try It

Let $A = [1,3]$, $B = [2,4]$, $C = [3,5]$, $D = [4,6]$ (all of which are subsets of the universal set $\mathbb{R}$).

1. What is $A \setminus B$? Answer: $[1,2)$
2. What is $A \cup B$? $A \cap B$? Answer: $[1,4]$, $[2,3]$
3. What is $A \cap C$?Answer: $\{3\}$
4. What is $A \cap D$?
5. What is $(A \cup D)^C$?

Try It

Let $A = [1, 3]$, $B = [2, 4]$, $C = [3, 5]$, $D = [4, 6]$ (all of which are subsets of the universal set $\mathbb{R}$).

1. What is $A \setminus B$? Answer: $[1, 2)$
2. What is $A \cup B$? $A \cap B$? Answer: $[1, 4]$, $[2, 3]$
3. What is $A \cap C$?Answer: $\{3\}$
4. What is $A \cap D$? Answer: $\emptyset$
5. What is $(A \cup D)^C$?

Try It

Let $A = [1,3]$, $B = [2,4]$, $C = [3,5]$, $D = [4,6]$ (all of which are subsets of the universal set $\mathbb{R}$).

1. What is $A \setminus B$? Answer: $[1,2)$
2. What is $A \cup B$? $A \cap B$? Answer: $[1,4]$, $[2,3]$
3. What is $A \cap C$?Answer: $\{3\}$
4. What is $A \cap D$? Answer: $\emptyset$
5. What is $(A \cup D)^C$? Answer: $(-\infty, 1) \cup (3,4) \cup (6, \infty)$

Functions

- A *function* is a correspondence between two sets A, B such that for every $x \in A$, $f(x) \in B$. We write $f : A \to B$.

Functions

- A *function* is a correspondence between two sets A, B such that for every $x \in A$, $f(x) \in B$. We write $f : A \to B$.
- We say A is the *domain* of f, and B is the *codomain* of f.

Functions

- A *function* is a correspondence between two sets A, B such that for every $x \in A$, $f(x) \in B$. We write $f : A \to B$.
- We say A is the *domain* of f, and B is the *codomain* of f.
- The *range* of f is the set

$$f(A) := \{f(x) | x \in A\}$$

  - Clearly, we have $f(A) \subseteq B$.

Vertical Line Rule

- A function $f : A \to B$ must assign a *single* element $f(x) \in B$ for each $x \in A$.
- A correspondence which assigns two or more possible values to $x \in A$ is not considered a function.

Vertical Line Rule

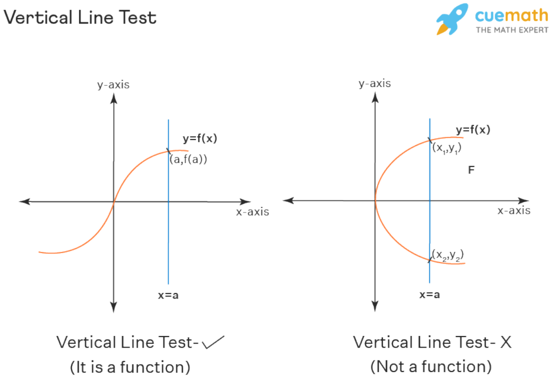


Figure: Vertical Line Test: if any vertical line touches the graph in two or more places, the graph is not a function.

Image Source: https://www.cuemath.com/algebra/vertical-line-test/

Functions You Know

- $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.

Functions You Know

- $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.
- $p : \mathbb{R} \to \mathbb{R}$ given by $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Functions You Know

- $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.
- $p : \mathbb{R} \to \mathbb{R}$ given by $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
- $g_i : \mathbb{R}^n \to \mathbb{R}$ given by

$$g_i(x_1, \ldots, x_n) = x_i$$

.

Injective, Surjective, Bijective

- A function $f : A \to B$ is *injective* (one-to-one) if for all $x_1, x_2 \in A$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.
    - You can check if a function is injective using the *horizontal line test*. Do you see how?

Injective, Surjective, Bijective

- A function $f : A \to B$ is *injective* (one-to-one) if for all $x_1, x_2 \in A$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.
  - You can check if a function is injective using the *horizontal line test*. Do you see how?
- A function $f : A \to B$ is *surjective* (onto) if $f(A) = B$.

Injective, Surjective, Bijective

- A function $f : A \to B$ is *injective* (one-to-one) if for all $x_1, x_2 \in A$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.
  - You can check if a function is injective using the *horizontal line test*. Do you see how?
- A function $f : A \to B$ is *surjective* (onto) if $f(A) = B$.
- A function $f : A \to B$ is *bijective* if f is both injective and surjective.

Try It

- Consider a function $f : \mathbb{Q} \to \mathbb{Z}$ which is given by the mapping

$$f\left(\frac{p}{q}\right) = (p - q)^2$$

- Is f injective? Surjective? Bijective?

Function Composition

If $f : A \to B$ and $g : B \to C$, then the function *composition* $g \circ f$ is defined as

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

Function Composition: Example

If $f : A \to B$ and $g : B \to C$, then the function *composition* $g \circ f$ is defined as
$$(g \circ f)(x) = g(f(x)).$$

For example, let $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$, $g(x) = \sin(x)$.

Then $(g \circ f)(x) = \sin(x^2)$.

Function Composition: Example

If $f : A \to B$ and $g : B \to C$, then the function *composition* $g \circ f$ is defined as

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

For example, let $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$, $g(x) = \sin(x)$.

In this case, we can also compute $(f \circ g)(x) = \left[\sin(x)\right]^2$.

Practice: Functions

Continuity & Differentiability

Continuity

- Intuitive definition: a continuous function is one for which you can draw the graph without lifting up your pencil.
- Math-ier definition: a function $f : A \to B$ is continuous if $x_1, x_2 \in A$ being close to each other implies $f(x_1), f(x_2)$ are also close to each other.

Continuity (or lack thereof)

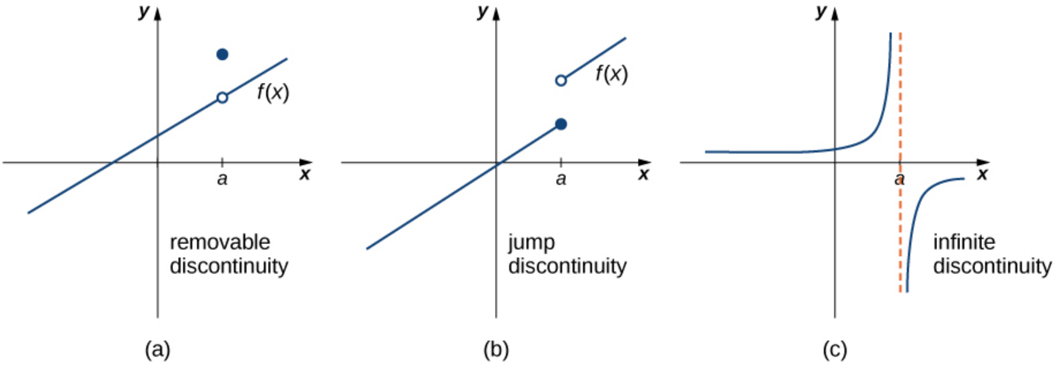


Figure: Different types of discontinuities in functions $f : \mathbb{R} \to \mathbb{R}$

Continuity

Formal Definition:

A function $f : A \to B$ is continuous at a point $a \in A$ if

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

If the function f is continuous at all points $a \in A$, then f as a whole is said to be a continuous function.

Continuity: Example

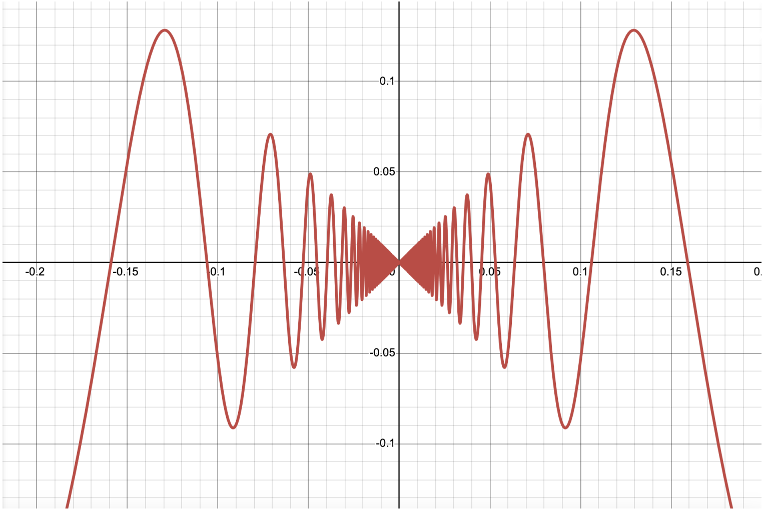
Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x\sin\left(\frac{1}{x}\right), x \neq 0 \\ 0, x = 0 \end{cases}$$

Is f continuous at $x = 0$?

This example adapted from Spivak's *Calculus*, 4th edition.

Continuity: Example



Discussion

Consider the function $f : \mathbb{R} \to \mathbb{Z}$ given by

$$f(x) = \begin{cases} 0, \text{if } x \text{ is irrational} \\ 1, \text{if } x \text{ is rational} \end{cases}$$

At which points x is f continuous?

This example adapted from Spivak's Calculus, 4th edition.

Derivative At A Point

A function $f : A \to B$ is differentiable at a point $a \in A$ if the limit

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

exists. If the limit above exists, we denote the limit by $f'(a)$ and call this the derivative of f at a.

Example

Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. What is the derivative of f at 2?

Let's write out the limit. We want

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

$$\lim_{x \to 2} \frac{x^2 - 4}{x - 2}$$

Example

Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. What is the derivative of f at 2?

Let's write out the limit. We want

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

$$\lim_{x \to 2} \frac{x^2 - 4}{x - 2}$$

$$\lim_{x \to 2} \frac{\cancel{(x-2)}(x+2)}{\cancel{x-2}}$$

$$\lim_{x \to 2} x + 2 = 4$$

Alternate Definition

A function $f : A \to B$ is differentiable at a point $a \in A$ if the limit

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

exists. If the limit exists, we denote it $f'(a)$ and call this the derivative of f at a.

Intuition

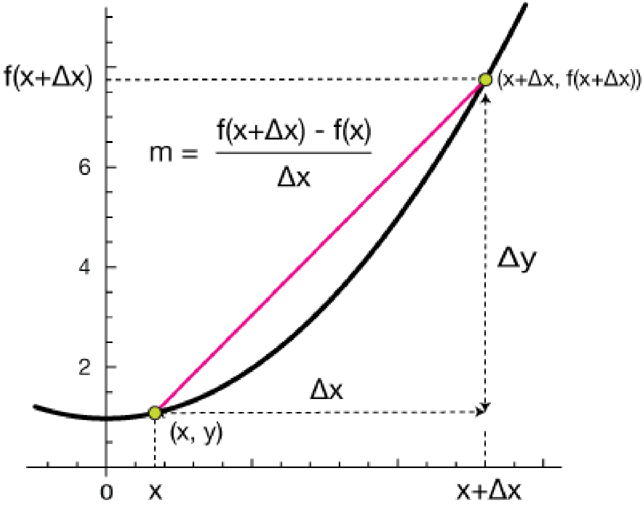


Image Source: `https://xaktly.com/Images/Mathematics/TheDerivative/DerivativeDefinition2.png`

Try It

Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Use the alternative definition to find the derivative of f at arbitrary point a.

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

Try It

Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Use the alternative definition to find the derivative of f at arbitrary point a.

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

$$\lim_{h \to 0} \frac{(a+h)^2 - a^2}{h}$$

$$\lim_{h \to 0} \frac{\cancel{a^2} + 2ah + h^2 - \cancel{a^2}}{h}$$

$$\lim_{h \to 0} \frac{2ah + h^2}{h}$$

$$\lim_{h \to 0} 2a = 2a$$

Derivative As Tangent Line

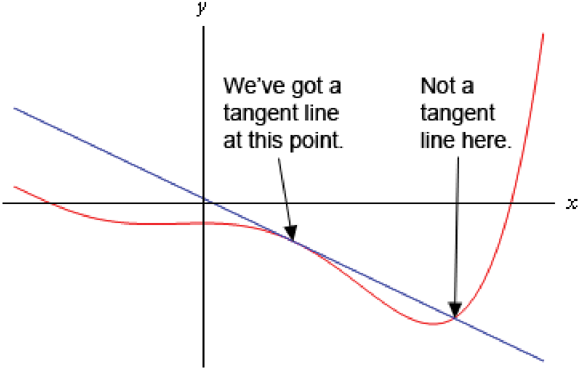


Figure: Tangent Line Visualization

Image Source: https://tutorial.math.lamar.edu/classes/calci/tangents$_r$*ates.aspx*

Maxima & Minima

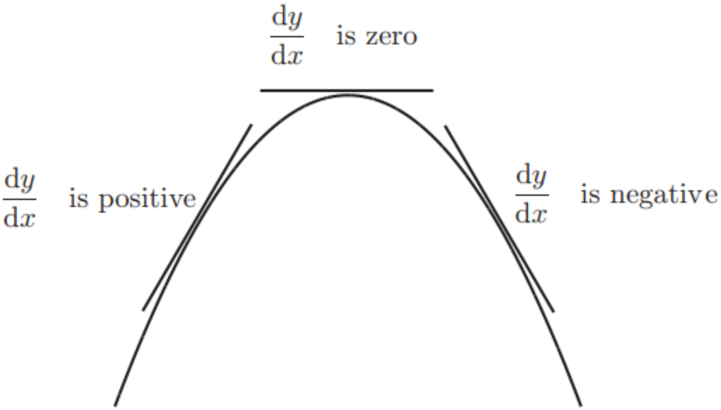


Figure: The derivative is zero at the local maximium pictured above.

Image Source: https://www.toppr.com/guides/maths/application-of-derivatives/maxima-and-minima/

Differentiation Rules

$$\frac{d}{dx} c = 0$$ Constant Rule

$$\frac{d}{dx} x^n = nx^{n-1}$$ Power Rule

$$\frac{d}{dx} \sin(x) = \cos(x)$$

$$\frac{d}{dx} \cos(x) = -\sin(x)$$

Trigonometric Rules

$$\frac{d}{dx} b^x = b^x \ln(b)$$ Exponential Rule

$$\frac{d}{dx} \ln(x) = \frac{1}{x}$$ Logarithmic Rule

Figure: Common Derivative Rules

Image Source: https://study.com/learn/lesson/function-differentiation-using-chain-rule-formulas.html

Differentiation Rules

Product Rule

$$\frac{d}{dx}\left[f(x)g(x)\right] = f(x)g'(x) + g(x)f'(x)$$

Quotient Rule

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{\left[g(x)\right]^2}$$

Image Source: https://iteducationlearning.com/quotient-and-product-rule-formula/

Differentiation Rules

Famously, we have

$$\frac{d}{dx}e^x = e^x$$

Chain Rule

The chain rule is the most important derivative rule and comes up frequently in most problems of interest. The chain rule is as follows:

$$\frac{d}{dx} f(g(x)) = f'(g(x))g'(x)$$

Chain Rule Example

Suppose that $f(x) = x^2$ and $g(x) = x^3$. We want to take the derivative of the function $f(g(x))$. Individually, we know $f'(x) = 2x, g'(x) = 3x^2$.

Chain Rule Example

Suppose that $f(x) = x^2$ and $g(x) = x^3$. We want to take the derivative of the function $f(g(x))$. Individually, we know $f'(x) = 2x, g'(x) = 3x^2$. Let's apply the chain rule.

$$\begin{aligned}\frac{d}{dx}f(g(x)) &= f'(g(x))g'(x) \\ &= 2g(x) \cdot 3x^2\end{aligned}$$

Chain Rule Example

Suppose that $f(x) = x^2$ and $g(x) = x^3$. We want to take the derivative of the function $f(g(x))$. Individually, we know $f'(x) = 2x, g'(x) = 3x^2$. Let's apply the chain rule.

$$\begin{aligned} \frac{d}{dx} f(g(x)) &= f'(g(x))g'(x) \\ &= 2g(x) \cdot 3x^2 \\ &= 2x^3 \cdot 3x^2 \\ &= 6x^5 \end{aligned}$$

Is this correct? How do you know?

Try It

Let $g(x) = x^2 + 2$ and $f(x) = \sin(x)$.
Find the derivative of $f(g(x) = \sin(x^2 + 2)$ using the chain rule.

Try It

Let $g(x) = x^2 + 2$ and $f(x) = \sin(x)$.
Find the derivative of $f(g(x) = \sin(x^2 + 2)$ using the chain rule.
Solution: We compute $g'(x) = 2x$ and $f'(x) = \cos(x)$.
Using the chain rule, we get

$$
\begin{aligned}
\frac{d}{dx} f(g(x)) &= f'(g(x))g'(x) \\
&= \cos(x^2 + 2) \cdot 2x
\end{aligned}
$$

Practice: Differentiation

Integration

Antiderivatives (Indefinite Integrals)

The antiderivative of a function f is another function F such that

$$F'(x) = f(x).$$

We write

$$\int f(x)dx$$

to denote the antiderivative of f.

Antiderivative Example

Clearly, we have

$$\int 2xdx = x^2 + C$$

The "$+C$" term denotes an arbitrary constant that may be added to the result, because the constant disappears in the derivative step.

In other words, $\frac{d}{dx}x^2 + C = 2x$ for all constants C.

Antiderivative Rules

$$\int k\,dx = k\,x + c$$

$$\int x^n\,dx = \tfrac{1}{n+1}x^{n+1} + c, n \neq -1$$

$$\int x^{-1}\,dx = \int \tfrac{1}{x}\,dx = \ln|x| + c$$

$$\int \frac{1}{ax+b}\,dx = \tfrac{1}{a}\ln|ax+b| + c$$

$$\int \ln u\,du = u\ln(u) - u + c$$

$$\int \mathbf{e}^u\,du = \mathbf{e}^u + c$$

Common Integrals

$$\int \cos u\,du = \sin u + c$$

$$\int \sin u\,du = -\cos u + c$$

$$\int \sec^2 u\,du = \tan u + c$$

$$\int \sec u \tan u\,du = \sec u + c$$

$$\int \csc u \cot u\,du = -\csc u + c$$

$$\int \csc^2 u\,du = -\cot u + c$$

Integration Techniques

- Integration is significantly more difficult than differentiation.
- Most derivative problems can be split up into smaller problems via the chain rule, but this can be much more difficult for antiderivatives.
- For example, how would you go about integrating

$$\int x\cos(x^2)dx?$$

U-Substituion (Informal Treatment)

- Many integral problems (at least, the ones given in courses) often can be rewritten in terms of another quantity u via the change-of-variable formula.
- How do you pick the u?
- You'll know you have the correct u **if you see both u and $\frac{du}{dx}$ inside your integral.**

U-Substitution (Example)

Compute

$$\int 2xe^{x^2}\,dx$$

U-Substitution (Example)

Compute
$$\int 2xe^{x^2}\,dx$$
We see both x^2 and it's derivative $2x$. This suggests we ought to take $u = x^2$.

U-Substitution (Example)

Compute

$$\int 2xe^{x^2} dx$$

We see both x^2 and it's derivative $2x$. This suggests we ought to take $u = x^2$.

Note that $\frac{du}{dx} = 2x \implies du = 2xdx$.

U-Substitution (Example)

Compute

$$\int 2xe^{x^2}dx$$

We see both x^2 and it's derivative $2x$. This suggests we ought to take $u = x^2$.

Note that $\frac{du}{dx} = 2x \implies du = 2xdx$.

Now, make the change-of-variable to u.

$$\int e^u du$$

U-Substitution (Example)

A good U-substitution should turn a tough problem into a more recognizable one. We know

$$\int e^u du = e^u + C$$

so now we just need to plug in for u. The final answer is

$$e^{x^2} + C$$

Correction Factor

Often, the u and $\frac{du}{dx}$ terms will be present in the integral, but with the wrong constant. This isn't a big deal, though, as we can always correct these. For example, suppose my problem was

$$\int 7xe^{x^2}\,dx$$

Now, instead of seeing $\frac{du}{dx} = 2x$ as I'd like to see, I only see $7x$.

Multiply By One

We can multiply by a clever form of 1. For example,

$$\frac{2}{7}\frac{7}{2}\int 7xe^{x^2}\,dx$$

is the same problem as before.
I can choose to pull in either constant term, such as the first one.

$$\begin{aligned} &= \frac{7}{2}\int \frac{2}{7}7xe^{x^2}\,dx \\ &= \frac{7}{2}\int 2xe^{x^2}\,dx \end{aligned}$$

Multiply By One

$$
\begin{aligned}
&= \frac{7}{2} \int \frac{2}{7} 7xe^{x^2}\, dx \\
&= \frac{7}{2} \int 2xe^{x^2}\, dx
\end{aligned}
$$

From here, we can proceed as we did on the previous slide. The answer will be $\frac{7}{2}e^{x^2} + C$

Try It

Compute

$$\int 3x\sin(x^2).$$

Compute

$$\int \cos(x)\sin^3(x)dx.$$

We Want The Area: Definite Integrals

We'll motivate definite integration initially as an area problem. How can we find the area of regions defined by functions?

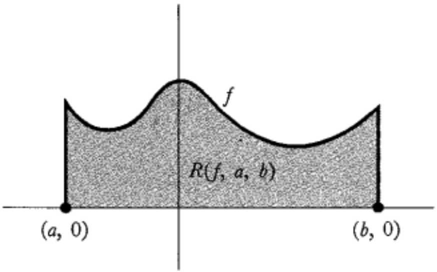


Figure: The area under the curve defined by $f(x)$ between $x = a$ and $x = b$.

Image from Spivak's Calculus, 4th ed.

Riemann Integrals

- In your undergraduate or high-school calculus course, you likely learned initially how to compute the area as the limit of a sum of the areas of small rectangles.

Riemann Integrals

- In your undergraduate or high-school calculus course, you likely learned initially how to compute the area as the limit of a sum of the areas of small rectangles.
- As the width of the rectangles becomes smaller and smaller, the approximation to the true area under the curve becomes better and better. Integrals computed in this way are called *Riemann integrals*.

Notation

We write the (signed) area under the curve of $f(x)$ between points $x = a$ and $x = b$ as

$$\int_a^b f(x)dx$$

Riemann Integrals

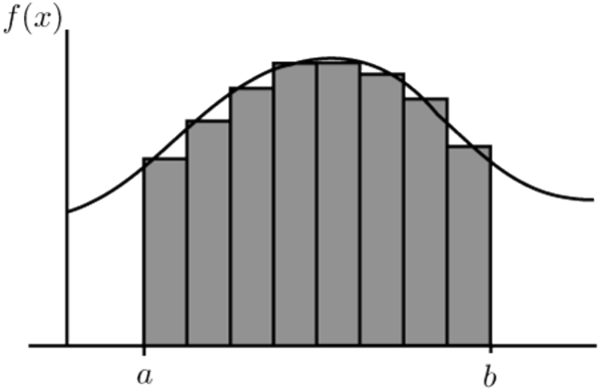
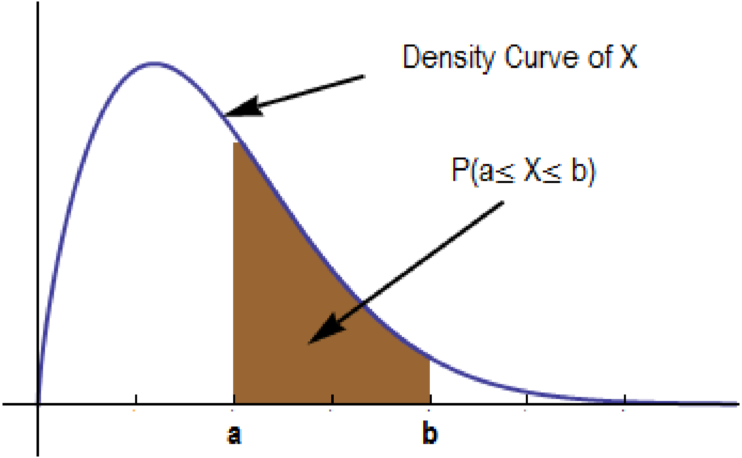


Image Source: https://commons.wikimedia.org/wiki/File:Riemann$_I$*ntegration*$_4$*.png*

Why Area?



Fundamental Theorem of Calculus

Given a function $f : \mathbb{R} \to \mathbb{R}$ continuous on $[a, b]$, suppose that $F(x)$ is an antiderivative of f. Then

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

Amazing!

Example

Find the area of the shaded region.

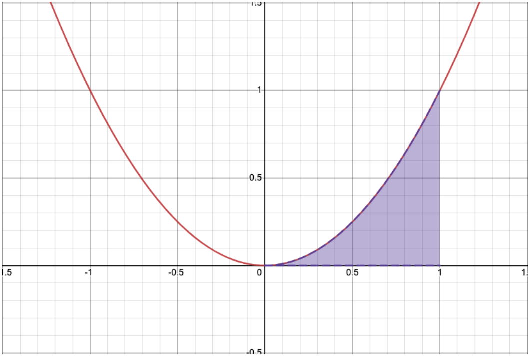


Figure: Find the area under the curve $f(x) = x^2$ between $x = 0$ and $x = 1$.

Example

We know an antiderivative of $f(x) = x^2$ is $F(x) = \frac{x^3}{3}$. Using the fundamental theorem of calculus, we have

$$\mathrm{Area} = F(1) - F(0) = \frac{1}{3} - 0 = \frac{1}{3}$$

Try It

Find the shaded area.

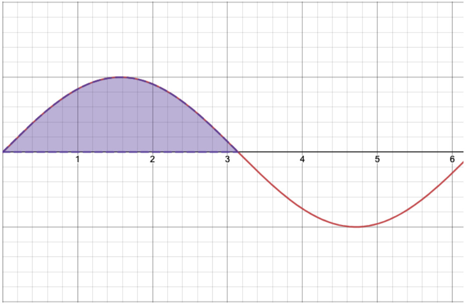


Figure: Find the area under $f(x) = sin(x)$ between $x = 0$ and $x = \pi$.

Try It

An antiderivative of $f(x) = sin(x)$ is $F(x) = -cos(x)$. Applying the FTC, we have

$$\text{Area} = -cos(\pi) - (-\cos(0)) = -(-1) - (-1) = 2$$

Practice: Integration

Multivariable Calculus

Partial Derivatives

- In the real-world, we often have functions of many variables, e.g.

$$f(x, y) = x^2 + y^2$$

Partial Derivatives

- In the real-world, we often have functions of many variables, e.g.

$$f(x, y) = x^2 + y^2$$
$$f(w, x, y, z) = \sqrt{2xy} + \sin(wz^y)$$

Partial Derivatives

- In the real-world, we often have functions of many variables, e.g.

$$f(x, y) = x^2 + y^2$$
$$f(w, x, y, z) = \sqrt{2xy} + \sin(wz^y)$$

- The *partial derivatives* of a multivariable function are one extension of the univariate derivative.

Partial Derivatives

- In the real-world, we often have functions of many variables, e.g.

$$f(x, y) = x^2 + y^2$$
$$f(w, x, y, z) = \sqrt{2xy} + \sin(wz^y)$$

- The *partial derivatives* of a multivariable function are one extension of the univariate derivative.
- These are often denoted $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$, etc.

Partial Derivatives

- In the real-world, we often have functions of many variables, e.g.

$$f(x, y) = x^2 + y^2$$
$$f(w, x, y, z) = \sqrt{2xy} + \sin(wz^y)$$

- The *partial derivatives* of a multivariable function are one extension of the univariate derivative.
- These are often denoted $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$, etc.
- Computing partial derivatives is surprisingly easy; just take the derivative as usual with respect to the variable of interest, treating all other variables as constants.

Partial Derivatives: Example

What is

$$\frac{\partial}{\partial x}(x+y)^2?$$

Partial Derivatives: Example

What is

$$\frac{\partial}{\partial x}(x+y)^2?$$

We treat y as constant and use the chain rule to get

$$2(x+y).$$

Partial Derivatives: Example

What is

$$\frac{\partial}{\partial z}\sin(xyz) + z^{x+y}?$$

Partial Derivatives: Example

What is

$$\frac{\partial}{\partial z}\sin(xyz) + z^{x+y}?$$

We treat x, y as constants and use the chain rule to get

$$\cos(xyz)\frac{\partial}{\partial z}xyz + (x+y)z^{x+y-1}$$
$$xy\cos(xyz) + (x+y)z^{x+y-1}$$

Gradient

The *gradient* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is a function $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1}(\vec{x}) \\ \frac{\partial}{\partial x_2}(\vec{x}) \\ \vdots \\ \frac{\partial}{\partial x_n}(\vec{x}) \end{bmatrix}$$

Gradient

The *gradient* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is a function $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1}(\vec{x}) \\ \frac{\partial}{\partial x_2}(\vec{x}) \\ \vdots \\ \frac{\partial}{\partial x_n}(\vec{x}) \end{bmatrix}$$

At a given point $\vec{x}$, the gradient contains information about how the function f is changing along each direction.

Gradient

- In applied problems, given a function $f : \mathbb{R}^n \to \mathbb{R}$, one will often be interested in finding vectors $\vec{x}_0$ such that

$$\nabla f(\vec{x}) \mid_{\vec{x}_0} = \vec{0}.$$

Gradient

- In applied problems, given a function $f : \mathbb{R}^n \to \mathbb{R}$, one will often be interested in finding vectors $\vec{x}_0$ such that

$$\nabla f(\vec{x}) \mid_{\vec{x}_0} = \vec{0}.$$

- These "points" may be maximizers (or minimizers) of the multivariate function f.

Gradient

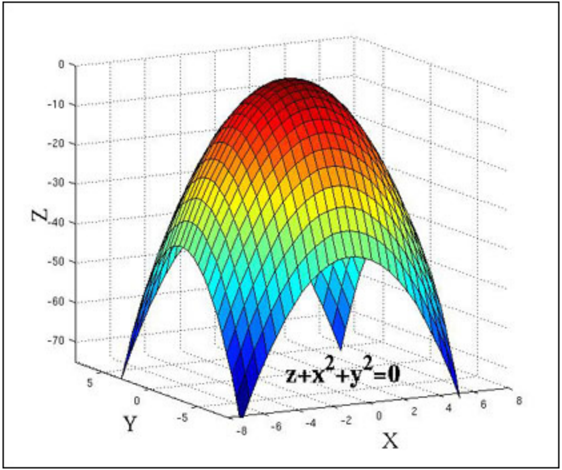


Image Source: `https://i.stack.imgur.com/hqYS5.png`

Gradient

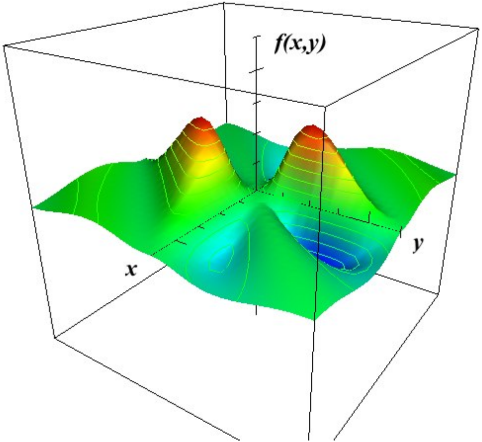


Image Source: `https://umdberg.pbworks.com/f/1322949891/GradientShape.jpg`

Gradient

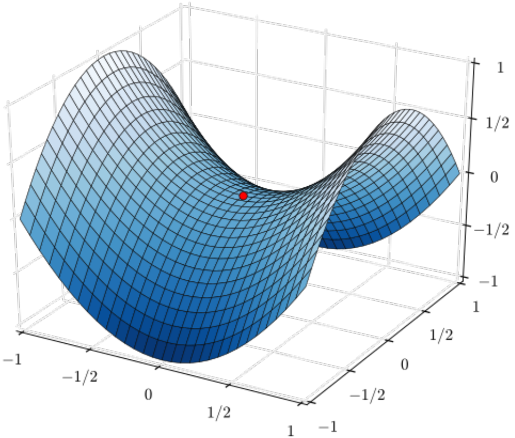


Image Source: `https://upload.wikimedia.org/wikipedia/commons/1/1e/Saddle_point.svg`

Gradient As A Direction

- Suppose we want to maximize a function $f : \mathbb{R}^n \to \mathbb{R}$, and we start with an initial guess $\vec{x}_0$.

Gradient As A Direction

- Suppose we want to maximize a function $f : \mathbb{R}^n \to \mathbb{R}$, and we start with an initial guess $\vec{x}_0$.
- The gradient *evaluated at any* point can be thought of as the "direction of steepest ascent" from that point.

Gradient As A Direction

- Suppose we want to maximize a function $f : \mathbb{R}^n \to \mathbb{R}$, and we start with an initial guess $\vec{x}_0$.
- The gradient *evaluated at any* point can be thought of as the "direction of steepest ascent" from that point.
- A common scheme in numerical optimization is to evaluate

$$\nabla f(\vec{x}) \mid_{\vec{x}_0}$$

and "move" your guess x_0 in this direction.

Gradient Ascent (GD)

$$\vec{x}_{t+1} \leftarrow \vec{x}_t + \eta \nabla f(\vec{x}_t)$$

Gradient Ascent (GD)

$$\underbrace{\vec{x}_{t+1}}_{\text{next iterate}} \leftarrow \underbrace{\vec{x}_t}_{\text{current iterate}} + \eta \nabla f(\vec{x}_t)$$

Gradient Ascent (GD)

$$\underbrace{\vec{x}_{t+1}}_{\text{next iterate}} \leftarrow \underbrace{\vec{x}_t}_{\text{current iterate}} + \underbrace{\eta}_{\text{step size}} \nabla f(\vec{x}_t)$$

Gradient Ascent (GD)

$$\underbrace{\vec{x}_{t+1}}_{\text{next iterate}} \leftarrow \underbrace{\vec{x}_t}_{\text{current iterate}} + \underbrace{\eta}_{\text{step size}} \underbrace{\nabla f(\vec{x}_t)}_{\text{gradient}}$$

Gradient Descent (GD)

$$\vec{x}_{t+1} \leftarrow \vec{x}_t - \eta \nabla f(\vec{x}_t)$$

Higher-Dimensional Functions

- Thus far, we've only considered multivariable functions of the form

$$f : \mathbb{R}^n \to \mathbb{R}.$$

Higher-Dimensional Functions

- Thus far, we've only considered multivariable functions of the form

$$f : \mathbb{R}^n \to \mathbb{R}.$$

- The natural next extension is functions

$$g : \mathbb{R}^n \to \mathbb{R}^m.$$

Higher-Dimensional Functions

- Thus far, we've only considered multivariable functions of the form

$$f : \mathbb{R}^n \to \mathbb{R}.$$

- The natural next extension is functions

$$g : \mathbb{R}^n \to \mathbb{R}^m.$$

- How can we differentiate such functions?

Examples

- Bijections in $\mathbb{R}^n$ are invertible functions $f : \mathbb{R}^n \to \mathbb{R}^n$, e.g.

$$f(x) = Ax$$

for some invertible matrix $A \in \mathbb{R}^{n\times n}$.

Examples

- Bijections in $\mathbb{R}^n$ are invertible functions $f : \mathbb{R}^n \to \mathbb{R}^n$, e.g.

$$f(x) = Ax$$

  for some invertible matrix $A \in \mathbb{R}^{n \times n}$.
- A dimensionality reduction procedure may apply a function $g : \mathbb{R}^n \to \mathbb{R}^q$ for $q << n$.

Jacobian

- For a function $f : \mathbb{R}^n \to \mathbb{R}^m$, consider the m real-valued functions $f_1, \ldots, f_m$ defined by

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix} \in \mathbb{R}^m$$

  In other words, each $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$.

Jacobian

- For a function $f : \mathbb{R}^n \to \mathbb{R}^m$, consider the m real-valued functions $f_1, \ldots, f_m$ defined by

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix} \in \mathbb{R}^m$$

  In other words, each $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$.
  These are sometimes called *coordinate functions*.

Jacobian

With this definition, the *Jacobian* is the matrix

$$Jf(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Jacobian

With this definition, the *Jacobian* is the matrix

$$
Jf(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} --- & \nabla f_1(\vec{x})^\top & --- \\ --- & \nabla f_2(\vec{x})^\top & --- \\ \vdots & \vdots & \vdots \\ --- & \nabla f_m(\vec{x})^\top & --- \end{bmatrix}
$$

Jacobian

With this definition, the *Jacobian* is the matrix

$$Jf(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} --- & \nabla f_1(\vec{x})^\top & --- \\ --- & \nabla f_2(\vec{x})^\top & --- \\ \vdots & \vdots & \vdots \\ --- & \nabla f_m(\vec{x})^\top & --- \end{bmatrix}$$

The Jacobian can be understood as the *derivative* of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ in the sense of the more general "Fréchet derivative" (you should read the Wikipedia for this derivative).

Application: Change-of-Variable

- Suppose X is a random vector with n-dimensional multivariate Gaussian distribution
$$X \sim \mathcal{N}(\mu, \Sigma).$$
What is the distribution (or density) of $Y = T(X)$ for bijective and differentiable function T?

Application: Change-of-Variable

- Suppose X is a random vector with n-dimensional multivariate Gaussian distribution
$$X \sim \mathcal{N}(\mu, \Sigma).$$
What is the distribution (or density) of $Y = T(X)$ for bijective and differentiable function T?
- It turns out that Y has density

$$p_Y(y) = p_X(T^{-1}(y)) \left| \det J T^{-1}(y) \right|$$

Multiple Integrals

- Just as one-dimensional integrals compute an area, multiple integrals compute a volume (or higher dimensional quantity).

Multiple Integrals

- Just as one-dimensional integrals compute an area, multiple integrals compute a volume (or higher dimensional quantity).

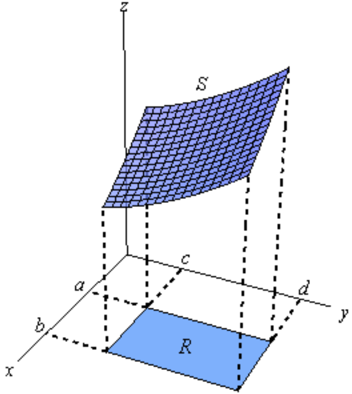


Image Source: https://tutorial.math.lamar.edu/classes/calciii/DoubleIntegrals.aspx

Multiple Integrals: Example

Evaluate

$$\int_0^2 \int_0^1 xy^2 dydx$$

Multiple Integrals: Example

Evaluate

$$\int_0^2 \int_0^1 xy^2 dydx$$

The inner integral

$$\int_0^1 xy^2 dy$$

can be evaluated by treating x as a constant. We get

$$x\frac{y^3}{3}\Big|_0^1 = \frac{x}{3} - 0 = \frac{x}{3}$$

Multiple Integrals: Example

Evaluate

$$\int_0^2 \int_0^1 xy^2 dydx$$

Plugging in yields

$$\int_0^2 \int_0^1 xy^2 dydx = \int_0^2 \left[\int_0^1 xy^2 dy \right] dx = \int_0^2 \frac{x}{3} dx$$

Multiple Integrals: Example

Evaluate

$$\int_0^2 \int_0^1 xy^2 dydx$$

Plugging in yields

$$\int_0^2 \int_0^1 xy^2 dydx = \int_0^2 \left[\int_0^1 xy^2 dy \right] dx = \int_0^2 \frac{x}{3} dx$$

and solving,

$$\frac{x^2}{6} \Big|_0^2 = \frac{4}{6} = \frac{2}{3}$$

The End