Mathematica Compendium

Miguel Antonio Méndez Hernández

October 31, 2025

Contents

1 Propositional Logic

Propositional logic is also called Boolean logic as it works on 0 and 1. In propositional logic, we use symbolic variables to represent the logic, and we can use any symbol for a representing a proposition, such A, B, C, P, Q, R, etc. Propositions can be either *true false*, but it cannot be both.

1.1 Logic Operators and Truth Tables

NOT ($\neg, \sim$)

A	$\neg A$
0	1
1	0

AND ($\wedge$)

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

OR ($\vee$)

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

IMPLIES ($\implies$)

A	B	$A \implies B$
0	0	1
0	1	1
1	0	0
1	1	1

IFF ($\iff$)

A	B	$A \iff B$
0	0	1
0	1	0
1	0	0
1	1	1

XOR ($\oplus$)

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

NOR ($\downarrow$)

A	B	$A \downarrow B$
0	0	1
0	1	0
1	0	0
1	1	0

NAND ($\uparrow$)

A	B	$A \uparrow B$
0	0	1
0	1	1
1	0	1
1	1	0

As a side not for the implication we can think about it as *if, then*. For the cases false, true; and false false; we call them *vacously* true. From a false premise can come everything out.

1.2 Tautology and Contradiction

Tautology: A logical formula that is always true.

Contradiction: A formula that is always false.

1.3 Logical Equivalences

Commutative Laws

$$p \wedge q \iff q \wedge p \qquad p \vee q \iff q \vee p$$

Associative Laws

$$(p \wedge q) \wedge r \iff p \wedge (q \wedge r) \qquad (p \vee q) \vee r \iff p \vee (q \vee r)$$

Distributive Laws

$$p \wedge (q \vee r) \iff (p \wedge q) \vee (p \wedge r) \qquad p \vee (q \wedge r) \iff (p \vee q) \wedge (p \vee r)$$

Identity Laws

$$p \wedge \mathrm{T} \iff p \qquad p \vee \mathrm{F} \iff p$$

Negation Laws

$$p \vee \sim p \iff \mathrm{T} \qquad p \wedge \sim p \iff \mathrm{F}$$

Double Negation Law

$$\sim (\sim p) \iff p$$

Idempotent Laws

$$p \wedge p \iff p \qquad p \vee p \iff p$$

Universal Bound Laws

$$p \vee \mathrm{T} \iff \mathrm{T} \qquad p \wedge \mathrm{F} \iff \mathrm{F}$$

De Morgan's Laws

$$\sim (p \wedge q) \iff (\sim p) \vee (\sim q) \qquad \sim (p \vee q) \iff (\sim p) \wedge (\sim q)$$

Absorption Laws

$$p \vee (p \wedge q) \iff p \qquad p \wedge (p \vee q) \iff p$$

Conditional Laws

$$(p \implies q) \iff (\sim p \vee q) \qquad \sim (p \implies q) \iff (p \wedge \sim q)$$

Complement Law

$$p \vee \neg p \iff \mathrm{T} \qquad p \wedge \neg p \iff \mathrm{F}$$

Biconditional

$$p \iff q \iff (p \implies q) \wedge (q \implies p)$$

Transitivity

$$(p \implies q) \land (q \implies r) \implies (p \Rightarrow r)$$

Indirect Proof (Contrapositive)

$$(p \implies q) \iff (\neg q \implies \neg p)$$

Disjunctive Syllogism (Disjunctive Exclusion)

$$p \nabla q \equiv (p \lor q) \land \neg p \implies q$$

$$p \nabla q \equiv (p \land q) \lor \neg(p \land q)$$

1.4 Truth Tables

Truth tables are a fundamental tool in logic that systematically show the truth value (true or false) of a compound statement for every possible combination of the truth values of its individual component statements.

Essentially, they lay out all the scenarios and the resulting truth of the overall logical expression. This helps determine if an argument is valid, if statements are logically equivalent, or the circumstances under which a complex statement is true or false.

Example:

p	q	$\neg p$	$\neg q$	$\neg p \implies q$	$(\neg p \implies q) \land \neg p$	$[(\neg p \implies q) \land \neg p] \Rightarrow q$
T	T	F	F	T	F	T
T	F	F	T	T	F	T
F	T	T	F	T	T	T
F	F	T	T	F	F	T

1.4.1 Filling a truth table

To fill a truth table for a logical expression with truth values (True or False), you follow a specific order for the input variables. This order ensures that all possible combinations of truth values for the variables are covered.

1.4.2 General Procedure:

1. **List all possible combinations of truth values for the input variables**: If you have n variables, the number of rows in the truth table will be 2^n. Each variable can be either True (T) or False (F).

2. **Order of the input variables**:

   - Start by filling in the truth values for the first variable. It alternates between True and False every 2^{n-1} rows.

   - Then for the second variable, it alternates every 2^{n-2} rows, and so on.

   - In short: the first variable alternates every other row, the second variable every two rows, the third every four rows, and so on.

Example:

For 3 variables, there are $2^3 = 8$ possible combinations of truth values. The truth values are filled in the following order:

A	B	C	Expression Result
T	T	T	–
T	T	F	–
T	F	T	–
T	F	F	–
F	T	T	–
F	T	F	–
F	F	T	–
F	F	F	–

The pattern for filling the input truth values

- The first column (A) alternates every 4 rows: 'T, T, F, F, T, T, F, F'.

- The second column (B) alternates every 2 rows: 'T, T, F, F, T, T, F, F'.

- The third column (C) alternates every row: 'T, F, T, F, T, F, T, F'.

This ensures that all combinations of A, B, and C are covered, and you can then evaluate the logical expression for each combination.

1.4.3 Truth Table for the Expression $(A \land B) \lor C$

A	B	C	$(A \land B) \lor C$
T	T	T	T
T	T	F	T
T	F	T	T
T	F	F	F
F	T	T	T
F	T	F	F
F	F	T	T
F	F	F	F

1.5 Disjunctive Normal Form (DNF)

Disjunctive Normal Form (DNF) is a standard way of writing a logical expression as a *disjunction(OR)* of *conjunctions (ANDs)*. A DNF expression consists of a series of conjunctions of literals, where each conjunction is connected by disjunctions.

To find the DNF in a truth table take the rows of the final result where there are true statements and bind the propositions that generated it with an AND inside parenthesis. Repeat it with each of true rows and connect all parenthesis with OR's

Example of DNF:

Consider the logical expression:

$$(A \land B) \lor (\neg A \land C) \lor (B \land \neg C)$$

This is in DNF because it is a disjunction (OR) of conjunctions (ANDs) of literals.

1.6 Conjunctive Normal Form (CNF)

Conjunctive Normal Form (CNF) is a standard way of writing a logical expression as a *conjunction (AND)* of *disjunctions (ORs)*. A CNF expression consists of a series of disjunctions of literals, where each disjunction is connected by conjunctions.

To find the CNF proceed just as the DNF but with the false rows and instead of ANDs inside the parenthesis use OR and connect the terms with OR. Also add a negation before each parenthesis.

Example of CNF:

Consider the logical expression:

$$\neg(A \vee B) \wedge \neg(\neg A \vee C) \wedge \neg(B \vee \neg C)$$

This is in CNF because it is a conjunction (AND) of disjunctions (ORs) of literals.

1.7   Basis

A *Basis* in the field of logic is set of operations with which you can represent all possible logic operations.

Example: The set $\{\neg, \wedge\}$

1.8   Karnaugh Maps

Karnaugh Maps (K-Maps) are a graphical method used to simplify Boolean expressions. The main goal of a K-map is to group adjacent cells that contain 1's in order to simplify the expression. A K-map helps identify common terms, allowing the Boolean expression to be reduced to its simplest form.

1.8.1   Karnaugh Map for Two Variables

Consider the Boolean expression $(A \vee (B \wedge \neg A \wedge \neg B))$.

We first construct a K-map for two variables, A and B. The truth table for this expression gives the following values:

A	B	$(A \vee (B \wedge \neg A \wedge \neg B))$
0	0	0
0	1	1
1	0	1
1	1	1

The corresponding K-map is:

AB	00	01	11	10
Value	0	1	1	1

Here, we group the ones together to simplify the Boolean expression. The simplified expression is:

$$A \vee B$$

1.8.2   Karnaugh Map for Three Variables

Now, let's consider the expression $\neg C$. This expression only depends on one variable, but for illustration, we will use a 3-variable K-map with variables A, B, and C.

The truth table for $\neg C$ is as follows:

A	B	C	$\neg C$
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

The corresponding K-map for three variables A, B, and C is:

$AB\backslash C$	0	1
00	1	0
01	1	0
11	1	0
10	1	0

We see that the ones are grouped in a column, leading to the simplified Boolean expression:

$$\neg C$$

1.8.3 Solving a Karnaugh Map

To solve a Karnaugh Map and simplify a Boolean expression, follow these steps:

1. **Determine the Number of Variables:** Decide how many variables are in the Boolean function. This determines the size of the K-Map:

   - 2 variables: 2×2

   - 3 variables: 2×4

   - 4 variables: 4×4

   - etc.

2. **Fill in the K-Map:** Place 1's in the cells that correspond to the min-terms (where the function outputs 1). You may also include do not care conditions (usually denoted as X).

3. **Group the 1's:** Form groups (called *implicants*) of 1's. The groups must follow these rules:

   - Each group must contain $1, 2, 4, 8, \ldots$ (powers of 2) 1's.

   - Groups must be rectangular (e.g., 1×2, 2×2).

   - Groups can wrap around the edges of the K-Map.

   - Try to form the largest groups possible to simplify the expression.

   - Each 1 should be included in at least one group.

4. **Write the Simplified Expression:**
   For each group:

   - Identify the variables that are constant (either always 0 or always 1) across the group.

   - Write a product term (AND) using only the constant variables.

   - Combine all product terms with OR operations to get the final simplified SOP (Sum of Products) expression.

1.9   Mathematical Quantifiers

1.9.1   Universal Quantifier $\forall$

Means **"for all"** or **"for every"**.

Negation: $\neg(\forall x)\,P(x) \equiv (\exists x)\,\neg P(x)$ "Not all" is the same as "There exists one that does not".

Example:

$\forall x \in \mathbb{R},\ x^2 \geq 0$ For all real numbers, the square is greater than or equal to zero.

Negated Example:

$\exists x \in \mathbb{R},\ x^2 < 0$ There exists a real number whose square is less than zero — this is false.

1.9.2   Existential Quantifier ∃

Means **"there exists at least one"**.

Negation: $\neg(\exists x)\,P(x) \equiv (\forall x)\,\neg P(x)$ "There does not exist" is the same as "For all, not".

Example:

$\exists x \in \mathbb{N},\ x > 10$ There exists a natural number greater than 10.

Negated Example:

$\forall x \in \mathbb{N},\ x \leq 10$ All natural numbers are less than or equal to 10 — this is false.

1.9.3   Unique Existential Quantifier ∃!

Means **"there exists exactly one"**.

Negation: "Not exactly one" means: $\neg(\exists! x)\,P(x) \equiv (\forall x)\,\neg P(x)\ \lor\ (\exists x_1 \neq x_2)\,P(x_1) \land P(x_2)$ Either no such x exists, or more than one does.

Example:

$\exists! x \in \mathbb{R},\ x + 5 = 0$ There exists exactly one real number such that $x + 5 = 0$.

Negated Example:

$\exists x_1 \neq x_2 \in \mathbb{R},\ x_1^2 = 4 \land x_2^2 = 4$ There are multiple solutions to $x^2 = 4$.

1.10   Common Symbols Used in Mathematical Expressions

- $>$ (greater than)
- $<$ (less than)
- $\geq$ (greater than or equal to)
- $\leq$ (less than or equal to)
- $=$ (equals)
- $\neq$ (not equal)
- $\in$ (element of a set)
- $\notin$ (not an element of)
- $\subset$ (proper subset)
- $\subseteq$ (subset)
- $\supset$ (proper superset)
- $\supseteq$ (superset)
- $\land$ (logical AND)
- $\lor$ (logical OR)
- $\implies$ (implies)
- $\iff$ (if and only if)

2   Set Theory

Set theory is a foundational branch of mathematics that studies *sets*, which are simply collections of objects. At its core, it deals with the fundamental concepts of membership (whether an object belongs to a set), equality (when two sets are the same), and relationships between sets (like subsets, intersections, and unions).

It might seem simple, but set theory provides the basic language and tools to define and reason about almost all mathematical objects, from numbers and functions to more complex structures. It helps us understand the concept of infinity, organize mathematical ideas logically, and resolve paradoxes that arise from dealing with collections.

In essence, set theory provides the building blocks upon which much of modern mathematics is constructed.

2.1   Basics

To be concise, a set is a collection of mathematical objects that can also be other sets. Here is a list of common symbols for set theory.

- *Empty set ($\emptyset$)*: The set that contains no elements. It is the unique set with zero elements. Every statement with $\forall x \in \emptyset$ is true, because there are no counterexamples.

- $A \subseteq B$: Set A is a subset of B. This means every element of A is also in B.

- $A \subseteq B$ or $A = B$: This notation already includes the possibility that A equals B since a set is always a subset of itself.

- $A \cup B$: The union of sets A and B. It includes all elements that are in A, in B, or in both.

- $A \cap B$: The intersection of sets A and B. It includes only the elements that are in both sets.

- $A \setminus B$: The difference of sets. Elements in A that are not in B.

- A^c or $\overline{A}$: The complement of set A. All elements not in A, relative to a universal set.

- $|A|$: The cardinality of set A, which is the number of elements in the set.

2.1.1   Visuals

1. $\emptyset$

1. Empty Set

2. $\{1, 2\}$
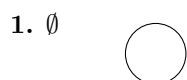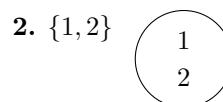
2. Set with Two Elements

3. $A \subseteq B$

3. A is a subset of B

5. $A \cup B$
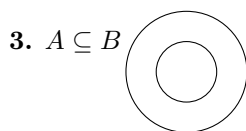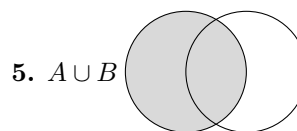
5. Union of A and B

6. $A \cap B$
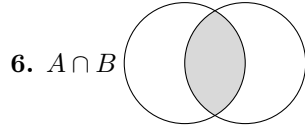
6. Intersection of A and B

7. $A \setminus B$

7. A minus B

8. A^c
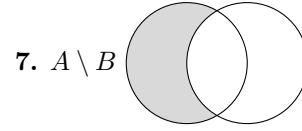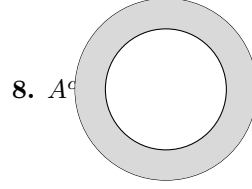
8. Complement of A

2.2 Set Complement Proof

Given $A, B \subseteq M$ then $A \subseteq B \implies B^c \subseteq A^c$

For some $x \in B^c$ we know that $x \notin B$. If $x \in A$ then $x \in B$ which is a contradiction. Hence $x \notin A$ therefore, for every $x \in B^c$ we conclude $B^c \subseteq A^c$.

$\mathfrak{QED}$

2.3 Equality of Sets

To prove that two sets are equal the following property is very useful.

$$A = B \iff A \subseteq B \wedge B \subseteq A$$

Proof:

For the case $A \subseteq B$

$$x \in A \implies x \in B$$

Also for the case $B \subseteq A$

$$x \in B \implies x \in A$$

If both statements are true on the same time then every element in A is also in B and vice versa. The other way of the equivalence is trivial.

$\mathfrak{QED}$

2.4 Axioms of Set Theory (Zermelo Fraenkel)

These are Zermelo Fraenkel axioms of set theory including *The Axiom of Choice*.

   I. *Axiom of Extensionality:* $\quad \forall A, B : A = B \implies (\forall C : C \in A \iff C \in B)$

  II. *Empty-Set Axiom:* $\quad \exists \emptyset : \forall X : X \notin \emptyset$

 III. *Axiom of Pairing:* $\quad \forall A, B : \exists C : \forall D : D \in C \iff (D = A \vee D = B)$

 IV. *Axiom of Union:* $\quad \forall A : \exists B : \forall C : C \in B \iff (\exists D : C \in D \wedge D \in A)$

  V. *Axiom of Infinity:* $\quad \exists N : \emptyset \in N \wedge (\forall x : x \in N \implies x \cup \{x\} \in N)$

 VI. *Axiom Schema of Specification:* $\quad \forall A : \exists B : \forall C : C \in B \iff (C \in A \wedge P(C))$

VII. *Axiom Schema of Replacement:* $\quad \forall A : \exists B : \forall y : y \in B \implies \exists x \in A : y = F(x)$

VIII. *Power Set Axiom:* $\quad \forall A : \exists B : \forall C : C \subseteq A \implies C \in B$

IX. *Foundation Axiom:* $\quad \forall A \neq \emptyset : \exists B \in A : A \cap B = \emptyset$

X. *Axiom of Choice:* $\quad \forall X : ([\forall A \in X : A \neq \emptyset] \wedge [\forall B, C \in X : B \neq C \implies B \cap C = \emptyset])$
$\implies \exists Y : \forall I \in X : \exists! J \in Y : J \in I$

2.5 The Cartesian Product

The Cartesian product of two sets A and B, written $A \times B$, is the set of all ordered pairs in which the first element belongs to A and the second belongs to B:

$$A \times B = \{(a, b) : a \in A, \wedge\ b \in B\}$$

Example:

Table 1: Cartesian Product of $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$

	$b \in B$		
$A \times B$	4	5	6
$a \in A$			
1	$(1, 4)$	$(1, 5)$	$(1, 6)$
2	$(2, 4)$	$(2, 5)$	$(2, 6)$
3	$(3, 4)$	$(3, 5)$	$(3, 6)$

The general Cartesian product of n sets can be written as:

$$X_{i=1}^{n+1} A_i = (X_{i=1}^{n} A_i) \times A_{n+1} \quad \text{with} \quad X_{i=1}^{1} A_i = A_1$$
$$\text{When } A_i = M \text{ for all } i :$$
$$M^n := M \times M \times \cdots \times M = X_{i=1}^{n} M \quad \text{with} \quad M^1 = M$$

2.5.1 Ordered Pair

Let A and B be two non-empty sets and $m \in A, n \in B$. We define

$$(m, n) := \{\{m, n\}, m\}$$

and

$$(n, m) := \{\{m, n\}, n\}$$

as ordered pairs which are elements of $A \times B$ and $B \times A$ respectively.

2.5.2 Distributive Law for the Cartesian Product

Let $A, A' \subseteq X$ and $B, B' \subseteq Y$. Then

$$(A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B').$$

We prove the equality by showing both inclusions.

1. Direction: $(A \times B) \cap (A' \times B') \subseteq (A \cap A') \times (B \cap B')$.

Let $(x, y) \in (A \times B) \cap (A' \times B')$.

$\Rightarrow (x, y) \in A \times B$ and $(x, y) \in A' \times B'$

$\Rightarrow x \in A, y \in B$ and $x \in A', y \in B'$

$\Rightarrow x \in A \cap A'$ and $y \in B \cap B'$

$\Rightarrow (x, y) \in (A \cap A') \times (B \cap B')$.

2. Direction: $(A \cap A') \times (B \cap B') \subseteq (A \times B) \cap (A' \times B')$.

Let $(x, y) \in (A \cap A') \times (B \cap B')$.

$\Rightarrow x \in A \cap A'$ and $y \in B \cap B'$

$\Rightarrow x \in A, x \in A'$ and $y \in B, y \in B'$

$\Rightarrow (x, y) \in A \times B$ and $(x, y) \in A' \times B'$

$\Rightarrow (x, y) \in (A \times B) \cap (A' \times B')$.

Thus, both inclusions hold and the equality follows.

$\mathfrak{QED}$

2.6 Laws of Set Algebra

Let X be the universal set and $A, B, C \subseteq X$.

- $\emptyset \subseteq A$

- $A \subseteq B \iff A \cap B = A \iff A \cup B = B \iff X \setminus B \subseteq X \setminus A \iff B \subseteq A$

- $A \cup B = B \cup A$

- $A \cap B = B \cap A$

- $(A \cup B) \cup C = A \cup (B \cup C)$

- $(A \cap B) \cap C = A \cap (B \cap C)$

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- $A \cup A = A$ and $A \cap A = A$

- $A \setminus B = A \cap (X \setminus B) = A \cap \overline{B}$

- $B = \overline{A} \iff (A \cup B = X \wedge A \cap B = \emptyset)$

- $\overline{A} \cap \overline{B} = \overline{A \cup B}$

- $\overline{A} \cup \overline{B} = \overline{A \cap B}$

- $\overline{\overline{A}} = A$

- $(A \setminus B) \cup (B \setminus) A = A \Delta B$

- $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$

- $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

2.6.1 Proof of the Distributive Law

$$\text{Let } x \in A \cap (B \cup C).$$
$$\Rightarrow x \in A \text{ and } x \in B \cup C$$
$$\Rightarrow x \in A \text{ and } (x \in B \text{ or } x \in C)$$
$$\Rightarrow (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C)$$
$$\Rightarrow x \in A \cap B \text{ or } x \in A \cap C$$
$$\Rightarrow x \in (A \cap B) \cup (A \cap C).$$

Thus,

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C).$$

$$\text{Conversely, let } x \in (A \cap B) \cup (A \cap C).$$
$$\Rightarrow x \in A \cap B \text{ or } x \in A \cap C$$
$$\Rightarrow (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C)$$
$$\Rightarrow x \in A \text{ and } (x \in B \text{ or } x \in C)$$
$$\Rightarrow x \in A \text{ and } x \in B \cup C$$
$$\Rightarrow x \in A \cap (B \cup C).$$

Thus,

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C).$$

$$\mathfrak{QED}$$

2.6.2 Proof of De Morgans's Law for sets and logic

The complement of $A \cup B$ is $\overline{(A \cup B)}$, and the law on disjoint decomposition states:

$$B = \overline{A} \iff (A \cup B = X) \wedge (A \cap B = \emptyset)$$

So define $\overline{C} := A \cup B$ and $D := \overline{A} \cap \overline{B}$, and use Law (11) to show the disjoint decomposition:

$$D = C \iff A \cap B = A \cup B$$

To show:

$$D \cup C = X \iff (\overline{A} \cap \overline{B}) \cup (A \cup B) = X$$

$$(\overline{A} \cap \overline{B}) \cup (A \cup B) = (\overline{A} \cup A \cup B) \cap (\overline{B} \cup A \cup B) \quad \text{(Law (8))}$$
$$= (X \cup B) \cap (X \cup A)$$
$$= X \cap X$$
$$= X$$

To show:

$$D \cap C = \emptyset \iff (\overline{A} \cap \overline{B}) \cap (A \cup B) = \emptyset$$

$$(\overline{A} \cap \overline{B}) \cap (A \cup B) = (A \cap B \cap A) \cup (A \cap B \cap B) \quad \text{(Law (7))}$$
$$= (\overline{A} \cap A \cap \overline{B}) \cup (\overline{A} \cap \overline{B} \cap B)$$
$$= (\emptyset \cap \overline{B}) \cup (\overline{A} \cap \emptyset)$$
$$= \emptyset \cup \emptyset$$
$$= \emptyset$$

$$\mathfrak{QED}$$

2.7 Indexed Sets

Let X be a set, and $A_i \subseteq X$ for all $i \in J$, where J is the *index set*.

If $J = \{1, 2, \ldots, n\}$:

$$\bigcup_{i=1}^{n} A_i := A_1 \cup A_2 \cup \cdots \cup A_n = \{x \mid \exists i \in J \ (x \in A_i)\}$$

$$\bigcap_{i=1}^{n} A_i := A_1 \cap A_2 \cap \cdots \cap A_n = \{x \mid \forall i \in J \ (x \in A_i)\}$$

$$X_{i=1}^{n} A_i = \{(a_1, \ldots, a_n) \mid a_i \in A_i\}$$

If J is any set:

$$\bigcup_{i \in J} A_i := \{x \mid \exists i \in J \ (x \in A_i)\}$$

$$\bigcap_{i \in J} A_i := \{x \mid \forall i \in J \ (x \in A_i)\}$$

If J is any set, then $(A_i)_{i \in J}$ are pairwise disjoint if and only if:

$$\forall i_1, i_2 \in J, \ i_1 \neq i_2 \implies A_{i_1} \cap A_{i_2} = \emptyset$$

If J is any set, then $(A_i)_{i \in J}$ forms a (disjoint) decomposition of X if and only if:

$$(A_i)_{i \in J} \text{ are pairwise disjoint and } \bigcup_{i \in J} A_i = X$$

2.7.1 Partitions Laws

Let $A_i, B_j \subseteq X$ for $i \in I$ and $j \in J$. Then the following holds:

- *De Morgan's Laws:*

$$\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \overline{A_i} \quad \text{and} \quad \overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \overline{A_i}$$

- $\bigcap_{i \in I} A_i \cup \bigcap_{j \in J} B_j = \bigcap_{i,j} (A_i \cup B_j) \quad \text{with} \quad \bigcap_{i,j} = \bigcap_{(i,j) \in I \times J}$
- $\bigcup_{i \in I} A_i \cap \bigcup_{j \in J} B_j = \bigcup_{i,j} (A_i \cap B_j) \quad \text{with} \quad \bigcup_{i,j} = \bigcup_{(i,j) \in I \times J}$

Here, $I = \{1, 2, 3, \ldots, n\}$ and $J = \{1, 2, 3, \ldots, m\}$. Then:

$$
\begin{aligned}
I \times J &= \{(i,j) \mid 1 \leq i \leq n, 1 \leq j \leq m, i, j \in \mathbb{N}\} \\
&= \{(1,1), (1,2), \ldots, (1,m), (2,1), (2,2), \ldots, (2,m), \ldots, (n,1), (n,2), \ldots, (n,m)\}
\end{aligned}
$$

2.8 Cardinality

The cardinality of a set is the number of elements in that set.

Let A and B be finite sets with $|A| = n$, $|B| = m$, and let X be the finite universal set. Then the following holds:

2.8.1 Cardinality of a Set

$$
A = (A \cap B) \cup (A \setminus B), \quad |A| = |A \cap B| + |A \setminus B|
$$

2.8.2 Cardinality of the complements

$$
|A| = |X \setminus A| = |X| - |A|
$$
$$
A \setminus B = A \cap (X \setminus B), \quad |A \setminus B| = |A| - |A \cap B|
$$

2.8.3 Cardinality of the Cartesian Product

$$
|A \times B| = |A| \cdot |B| = n \cdot m
$$

2.8.4 Inclusion-Exclusion Formula for Two Disjoint Sets

$$
|A \cup B| = |A| + |B| = n + m
$$

2.8.5 Inclusion-Exclusion Formula for Two Non-Disjoint Sets

Let $|A \cap B| = k$, then:

$$
|A \cup B| = |A| + |B| - |A \cap B| = n + m - k \quad \text{(since we do not count the intersection twice)}
$$

2.8.6 Inclusion-Exclusion Formula for Three Non-Disjoint Sets

$$
\begin{aligned}
|A \cup B \cup C| &= |(A \cup B) \cup C| \\
&= |A \cup B| + |C| - |(A \cup B) \cap C| \\
&= |A| + |B| - |A \cap B| + |C| - |(A \cap C) \cup (B \cap C)| \\
&= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
\end{aligned}
$$

2.8.7 General Formula for the cardinality of the union of sets

$$
\left| \bigcup_{i=1}^{n} M_i \right| = \sum_{I \subseteq \{1, \ldots, n\}, I \neq \emptyset} (-1)^{|I|-1} \left| \bigcap_{i \in I} M_i \right|
$$

2.9 Definition of Countable and Non-countable

A set A is *infinite countable* or just *countable* if there exist a bijection from $\mathbb{N}$ to A. If A has infinite elements and there is no such bijection, we call it *non-countable*

2.10   Countable Union

Given $A_n, n \in \mathbb{N}$ countable sets, then

$$M := \bigcup_{n \in \mathbb{N}} A_n$$

is also countable.

2.11   Countability of Subsets

Given $A \subseteq B$, where B is countable. Then A is also countable. The other case is that if A is non-countable then B must be non-countable as well.

2.12   The Power Set

The Power Set of a set is the set of all subsets of a given set

$$\mathcal{P}(x) := \{M : M \subseteq X\}$$

Its cardinality is 2^n with n being the number of elements in the original set X.

2.13   Set System

Let X be a non-empty set. A subset $\mathscr{F}$ of the power set of X is called a *set system* of X.

We also define

$$\bigcap_{i \in I} T_i := \{x \in X \mid \forall i \in I : x \in T_i\}$$

and

$$\bigcup_{i \in I} T_i := \{x \in X \mid \exists i \in I : x \in T_i\}$$

where $T_i \in \mathcal{P}(X)$ as the *intersection* and *union* of the set system.

Example:

Let $M \neq M$, if $\mathcal{U} := \emptyset \subseteq \mathcal{P}(M)$ then

$$\bigcap_{T \in \mathcal{U}} T = M \text{ and } \bigcup_{T \in \mathcal{U}} T = \emptyset$$

Proof:

The set intersection of the empty-set gives us the set of all elements that are in every set of the empty-set. We have defined our definition

$$\bigcap_{i \in I} T_i := \{x \in X \mid \forall i \in I : x \in T_i\}$$

for our case

$$
\begin{aligned}
\bigcap_{T \in \mathcal{U}} T &= \{m \in M \mid \forall T \in \mathcal{U} : m \in T\} \\
&= \{m \in M \mid T \in \mathcal{U} \implies m \in T\} \\
&= M
\end{aligned}
$$

Because the statement $T \in \mathcal{U}$ in the implication $T \in \mathcal{U} \implies m \in T$ is always true. In other words, the condition says, for all elements of an empty-set, $m \in T$ holds true, but there are no elements in the empty-set therefore there is no counter example. Ergo the condition is true.

For the union

$$\bigcup_{T \in \mathcal{U}} T = \{m \in M \mid \exists T \in \mathcal{U} : m \in T\}$$

The condition is that there exists an element in T such that, that element is also in every other set. Therefore, the only element which satisfies such a requirement is the empty-set.

$$\mathfrak{QED}$$

2.14 Partition

Let X be a non-empty set. A subset $\mathscr{F}$ of the power set of X is called a partition if:

1. $M \neq \emptyset \; \forall M \in \mathscr{F}$
2. $\bigcap \mathscr{F} = X$
3. $M_1 \bigcap M_2 \neq \emptyset \implies M_1 = M_2 \; \forall \; M_1, M_2 \in \mathscr{F}$

2.14.1 Properties of Partitions

- Every equivalence relation corresponds to a partition:

$$J : \{R : R \text{ is an equivalence relation on } X\} \to \{F : F \text{ is a partition of } X\}$$

where $J(R) := X/R$ is a bijection.

- If F is a partition of X, then we can define an equivalence relation R_F by:

$$R_F := \{(x, y) \in X \times X : \exists M \in F \text{ such that } x, y \in M\}$$

Then R_F is an equivalence relation on X.

2.15 Family of Subsets Operations

Let $\mathscr{F}$ be a system of sets (a family of subsets) on the set X. We define:

$$\bigcup_{M \in F} M := \bigcup F := \{x \in X : \text{there exists } M \in F \text{ such that } x \in M\},$$

$$\bigcap_{M \in F} M := \bigcap F := \{x \in X : x \in M \text{ for all } M \in F\}.$$

2.16 Pigeon Principle

For $g : A \to B$ and $|A| = n + 1$, $|B| = n$ then there exists $b \in B$ such that $|g^{-1}(\{b\})| \geq 2$.

In other words if you have n boxes and $n + 1$ objects, then at least one of those boxes will have more than 2 elements.

3 Relations, Maps and Functions

3.1 Map Definition

A *map* for two sets A and B is defined as a subset of $A \times B$. And it is denoted by

$$f : A \to B \subseteq \{(a, b) \mid (a, b) \in A \times B\}$$

3.2 The Set of All Maps

For the sets M and N, we define the sets off all maps $f : M \to N$ as

$$N^M := \{f \mid f : M \to N\}$$

3.3 The Characteristic Map

Given $T \subseteq M$ we define the *characteristic map* as

$$\mathcal{X}_T : M \to \{0, 1\} : m \to \begin{cases} 1, & m \in T \\ 0, & m \notin T \end{cases}$$

This map tell us if an element is in the subset. Also each map $\alpha : M \to \{0, 1\}$ defines a unique subset T of M, $T := M \mid \alpha(m) = 1$.

3.4 Relations

A *relation* in mathematics is a connection or relationship between elements of two sets. It's formally defined as a subset of the Cartesian product of the sets.

For example, if we have sets A and B, a relation R from A to B consists of ordered pairs (a, b) where $a \in A$ and $b \in B$, such that a is related to b according to some rule or property.

Common types of relations include:

- Functions (special relations where each input has exactly one output)

- Equivalence relations (reflexive, symmetric, and transitive)

- Partial orders (reflexive, antisymmetric, and transitive)

Relations can be represented using diagrams, matrices, or sets of ordered pairs, and they're fundamental to many areas of mathematics including algebra, calculus, and discrete mathematics.

3.5 Types of relations

Let A be a set and X be a relation on A.

- *Reflexive:* $\forall a \in A : (a, a) \in X$ (or written as $a \sim a$)

- *Irreflexive:* $\forall a \in A : (a, a) \notin X$

- *Symmetric:* $\forall a, b \in A : (a, b) \in X \to (b, a) \in X$ (or written as $(a \sim b) \Rightarrow (b \sim a)$)

- *Antisymmetric:* $\forall a, b \in A : (a, b) \in X$ and $(b, a) \in X \to a = b$

- *Transitive:* $\forall a, b, c \in A : (a, b) \in X$ and $(b, c) \in X \to (a, c) \in X$ (or written as $(a \sim b)$ and $(b \sim c) \Rightarrow (a \sim c)$)

- *Total:* $\forall a, b \in A : a \neq b \to (a, b) \in X$ or $(b, a) \in X$

3.6 Equivalence relation

An *equivalence relation* is a relation that is symmetric, transitive and reflexive.

Example:

$$R := \{(a, b) \in \mathbb{N} \times \mathbb{N} : a = b\}$$

3.7 Functions

A subset $f \subseteq A \times B$ is called a **function** if:

$$(x, y) \in f \text{ and } (x, y') \in f \implies y = y'.$$

That is, for each $x \in A$, there exists at most one $y \in B$ such that $(x, y) \in f$.

3.8 The Graph

$$\text{graph}(f) := \{(x, f(x)) : x \in X\}$$

3.9 The identity

$$\text{id}(f) := idx := \{(x, x) : x \in X\}$$

3.10 Image

The *image (or range)* of a relation R from set A to set B is the set of all elements in B that are related to at least one element in A. Formally, if $R \subseteq A \times B$ is a relation, then the image of R is defined as:

$$\text{img}(R) = \{b \in B \mid \exists a \in A \text{ such that } (a, b) \in R\}$$

In other words, the image consists of all the output values that appear in the ordered pairs of the relation. For example, if $R = \{(1, 4), (2, 5), (3, 4), (2, 6)\}$, then $\text{img}(R) = \{4, 5, 6\}$.

3.11 Domain

The *domain* of a relation R from set A to set B is the set of all elements in A that are related to at least one element in B. Formally, if $R \subseteq A \times B$ is a relation, then the domain of R is defined as:

$$\text{Dom}(R) = \{a \in A \mid \exists b \in B \text{ such that } (a, b) \in R\}$$

In other words, the domain consists of all the input values that appear in the ordered pairs of the relation. For example, if $R = \{(1, 4), (2, 5), (3, 4), (2, 6)\}$, then $\text{Dom}(R) = \{1, 2, 3\}$.

3.12 Equivalence Class

Let $\sim$ be an equivalence relation on a set A. For an element $x \in A$, the *equivalence class* of x, denoted by $[x]$, is the set of all elements in A that are equivalent to x. Formally, it is defined as:

$$[x] := \{y \in A \mid x \sim y\} \subseteq A$$

In other words, the equivalence class of x contains all elements y in A such that x is related to y under the equivalence relation $\sim$.

3.13   Quotient Space

Let $\sim$ be an equivalence relation on a set A. The *quotient space* of A by $\sim$, denoted by $A/\sim$ (or sometimes A/R), is the set of all distinct equivalence classes of elements in A. Formally, it is defined as:

$$A/\sim := \{[x] \mid x \in A\}$$

The quotient space $A/\sim$ is a partition of the original set A into disjoint equivalence classes. Each element of the quotient space is an equivalence class $[x]$, which itself is a subset of A.

3.14   Composition of Maps

The composition of two functions is the operation of applying one function to the result of another function. If we have functions $f : A \to B$ and $g : B \to C$, then the composition of g and f, denoted as $g \circ f$ (read as g composed with f), is a function from A to C defined by:

$$(g \circ f)(a) = g(f(a)) \quad \text{for all } a \in A$$

The composition applies f first, then applies g to the result. Note that the co-domain of f must match the domain of g for the composition to be defined.

The formal definition is

$$g \circ f := \{(s, u) \in S \times U \mid \exists t \in T \text{ such that } (s, t) \in f \land (t, u) \in g\}$$

Example:

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$ and $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = x + 3$. Then:

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 3$$

$$(f \circ g)(x) = f(g(x)) = f(x + 3) = (x + 3)^2 = x^2 + 6x + 9$$

Note that $g \circ f \neq f \circ g$ in general, which shows that function composition is not commutative.

3.15   Types of Functions

3.15.1   Injective Functions

An injective function (also called a one-to-one function) is a function that maps distinct elements from the domain to distinct elements in the co-domain.

Formally, a function $f : A \to B$ is injective if for all $a_1, a_2 \in A$:

$$a_1 \neq a_2 \to f(a_1) \neq f(a_2)$$

Equivalently, using the contrapositive:

$$f(a_1) = f(a_2) \to a_1 = a_2$$

3.15.2   Surjective Functions

A surjective function (also called an onto function) is a function whose image equals its co-domain, meaning that every element in the co-domain has at least one pre-image in the domain.

Formally, a function $f : A \to B$ is surjective if:

$$\forall b \in B, \exists a \in A \text{ such that } f(a) = b$$

3.15.3   Bijection Functions

A bijective function (also called a one-to-one correspondence) is a function that is both injective and surjective. In other words, every element in the co-domain is mapped to by exactly one element in the domain.

Formally, a function $f : A \to B$ is bijective if it is both:

- Injective: $\forall a_1, a_2 \in A, a_1 \neq a_2 \to f(a_1) \neq f(a_2)$

- Surjective: $\forall b \in B, \exists a \in A$ such that $f(a) = b$

Bijective functions establish a perfect pairing between elements of the domain and co-domain, where each element in the domain corresponds to exactly one element in the co-domain, and vice versa. A bijection allows us to define an inverse function $f^{-1} : B \to A$.

3.16   Propositions on Images and Pre-images under Set Operations

Let $f : X \to Y$ be a function.

3.16.1   Union and Cut Sets

- For subsets $A_1, A_2 \subseteq X$, we have:

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2) \quad \text{and} \quad f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$$

- For subsets $B_1, B_2 \subseteq Y$, we have:

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2) \quad \text{and} \quad f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$$

Proof:

We prove the second part of 1 and the first part of 2.

Let $y \in f(A_1 \cap A_2)$. Then there exists $x \in A_1 \cap A_2$ such that $f(x) = y$. Since $x \in A_1$ and $x \in A_2$, it follows that $y \in f(A_1)$ and $y \in f(A_2)$, hence $y \in f(A_1) \cap f(A_2)$. Therefore, every element of $f(A_1 \cap A_2)$ is also an element of $f(A_1) \cap f(A_2)$, so:

$$f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$$

Let $x \in f^{-1}(B_1 \cup B_2)$. Then $f(x) \in B_1 \cup B_2$, which means $f(x) \in B_1$ or $f(x) \in B_2$. Thus, $x \in f^{-1}(B_1)$ or $x \in f^{-1}(B_2)$, which implies:

$$x \in f^{-1}(B_1) \cup f^{-1}(B_2)$$

Hence, both sets contain the same elements and are therefore, equal:

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$$

$$\mathfrak{QED}$$

3.16.2   Union and Cut of the whole Domain and Range

Let $f : X \to Y$ be a function.

- Let $\mathcal{F}$ be a collection of subsets of X. Then:

$$f\left(\bigcup_{A \in \mathcal{F}} A\right) = \bigcup_{A \in \mathcal{F}} f(A) \quad \text{and} \quad f\left(\bigcap_{A \in \mathcal{F}} A\right) \subseteq \bigcap_{A \in \mathcal{F}} f(A)$$

- Let $\mathcal{G}$ be a collection of subsets of Y. Then:

$$f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right) = \bigcup_{B\in\mathcal{G}} f^{-1}(B) \quad \text{and} \quad f^{-1}\left(\bigcap_{B\in\mathcal{G}} B\right) = \bigcap_{B\in\mathcal{G}} f^{-1}(B)$$

Proof (partial):

We show the first statement of part (ii); the rest follows analogously.

Let $x \in f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right)$. Then:

$$f(x) \in \bigcup_{B\in\mathcal{G}} B \quad \Leftrightarrow \quad \exists B \in \mathcal{G} \text{ such that } f(x) \in B \quad \Leftrightarrow \quad \exists B \in \mathcal{G} \text{ such that } x \in f^{-1}(B)$$

Hence:

$$x \in \bigcup_{B\in\mathcal{G}} f^{-1}(B)$$

It follows that:

$$f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right) = \bigcup_{B\in\mathcal{G}} f^{-1}(B)$$

$$\mathfrak{QED}$$

Example:

We define $A = \{1, 0\}$ and $B = \{2\}$ and $f(1) = 3$, $f(2) = 3$ $f(0) = 1$. Then

$$f(A \cap B) = f(\emptyset) = \emptyset$$

And

$$f(A) \cap f(B) = 3$$

3.17 Even Function

A real function f is *even* if

- $\forall x \in \mathrm{Dom}(f)\exists - x \in \mathrm{Dom}(f)$
- $\forall x \in \mathrm{Dom}(f)$ the following is true $f(x) = f(-x)$

3.18 Odd Function

A real function f is *odd* if

- $\forall x \in \mathrm{Dom}(f)\exists - x \in \mathrm{Dom}(f)$
- $\forall x \in \mathrm{Dom}(f)$ the following is true $f(x) = -f(x)$

3.19 Inverse of a Function

The *inverse* of a function $f : A \to B$ is a function $f^{-1} : B \to A$ that reverses the operation of f. That is, if f maps an element $a \in A$ to an element $b \in B$, then the inverse function f^{-1} maps b back to a.

Formally, a function $f : A \to B$ has an inverse $f^{-1} : B \to A$ if and only if f is bijective (both injective and surjective). The inverse function satisfies the following properties:

$$f^{-1}(f(a)) = a \quad \text{for all } a \in A$$

$$f(f^{-1}(b)) = b \quad \text{for all } b \in B$$

In other words, composing a function with its inverse yields the identity function. That is:

$$f^{-1} \circ f = id_A \quad \text{and} \quad f \circ f^{-1} = id_B$$

Where id_A and id_B are the identity functions on sets A and B, respectively.

3.19.1 Steps to Find the Inverse of a Function

To find the inverse of a function $f(x)$, follow these steps:

1. Replace $f(x)$ with y: $y = f(x)$

2. Interchange the variables x and y: $x = f(y)$

3. Solve for y in terms of x: $y = f^{-1}(x)$

4. Verify that the resulting function is indeed the inverse by checking that $f^{-1}(f(x)) = x$ and $f(f^{-1}(x)) = x$

Example:

Let's apply the steps above to find the inverse of $f(x) = 2x + 3$.

Step 1: Replace $f(x)$ with y

$$y = 2x + 3$$

Step 2: Interchange the variables x and y

$$x = 2y + 3$$

Step 3: Solve for y in terms of x

$$x = 2y + 3$$
$$x - 3 = 2y$$
$$\frac{x - 3}{2} = y$$

So, the inverse function is:

$$f^{-1}(x) = \frac{x - 3}{2}$$

Step 4: Verify that $f^{-1}(f(x)) = x$ and $f(f^{-1}(x)) = x$

Let's verify $f^{-1}(f(x)) = x$:

$$f^{-1}(f(x)) = f^{-1}(2x + 3)$$
$$= \frac{(2x + 3) - 3}{2}$$
$$= \frac{2x}{2}$$
$$= x$$

And let's verify $f(f^{-1}(x)) = x$:

$$f(f^{-1}(x)) = f\left(\frac{x - 3}{2}\right)$$
$$= 2\left(\frac{x - 3}{2}\right) + 3$$
$$= (x - 3) + 3$$
$$= x$$

Since both compositions yield the identity function, $f^{-1}(x) = \frac{x-3}{2}$ is indeed the inverse of $f(x) = 2x + 3$.

3.19.2 Properties of the Inverse Function

Let $f : X \to Y$ be a function.

- Assume $x \sim y$ if $f(x) = f(y)$ so is $\sim$ an equivalence relation.

- Consider the quotient set $X_f := X/\sim$, where $\sim$ is the equivalence relation defined by $x \sim x' \iff f(x) = f(x')$. Let $q_f : X \to X_f$ be the canonical projection defined by $q_f(x) = [x]$, and let $\iota_f : f(X) \to Y$ be the inclusion map, $y \mapsto y$. Then the function

$$\hat{f} : X_f \to f(X), \quad [x] \mapsto f(x)$$

  is a bijection, and the original map f can be written as the composition

$$f = \iota_f \circ \hat{f} \circ q_f.$$

  Here, q_f is surjective, $\hat{f}$ is bijective, and ι_f is injective. This yields the following commutative diagram:

$$
\begin{array}{ccc}
X & \xrightarrow{\;f\;} & Y \\
\downarrow{\scriptstyle q_f} & & \uparrow{\scriptstyle \iota_f} \\
X_f & \xrightarrow{\;\hat{f}\;} & f(X)
\end{array}
$$

3.20 Transformations of a Function

Transformations modify the appearance of a function's graph without altering its basic shape. Here, we examine how different algebraic changes to a function $f(x)$ affect its graph:

Vertical Translation: $f(x) + a$

- Shifts the graph *upward* if $a > 0$, and *downward* if $a < 0$.

- Each point on the graph moves vertically by a units.

Horizontal Translation: $f(x + a)$

- Shifts the graph *left* if $a > 0$, and *right* if $a < 0$.

- This is opposite of what might be expected: adding to x shifts the graph in the negative direction.

Vertical Scaling (Stretch/Compression): $af(x)$

- If $|a| > 1$: the graph is *stretched* vertically (taller and narrower).
- If $0 < |a| < 1$: the graph is *compressed* vertically (shorter and wider).
- If $a < 0$: includes a reflection across the *x-axis*.

Horizontal Scaling (Stretch/Compression): $f(ax)$

- If $|a| > 1$: the graph is *compressed* horizontally (narrower).
- If $0 < |a| < 1$: the graph is *stretched* horizontally (wider).
- If $a < 0$: includes a reflection across the *y-axis*.

Reflection across the x-axis: $-f(x)$

- Flips the graph upside-down over the x-axis.
- Each point (x, y) becomes $(x, -y)$.

Reflection across the y-axis: $f(-x)$

- Flips the graph left-to-right over the y-axis.
- Each point (x, y) becomes $(-x, y)$.

4 Mathematical Proofs

In this section I will provide with some examples of different types of proofs.

4.1 Proof by Direct Argument

$$A \implies F$$

For any integer n, if n is even, then n^2 is even. We will prove this theorem by direct argument.

Assume n is an even integer. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even, as it can be expressed as $2m$ where $m = 2k^2$ is an integer. Therefore, we conclude that if n is even, then n^2 is even.

<div align="right">𝖰𝖤𝖣</div>

4.2 Proof by Contradiction

$$A \iff \neg A \implies F$$

If n is an integer such that n^2 is even, then n is even.

We will prove this theorem by contradiction. Assume that n is an integer such that n^2 is even, but n is odd. Then we can write $n = 2k + 1$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

This shows that n^2 is odd, which contradicts our assumption that n^2 is even. Therefore, our assumption that n is odd must be false, and thus, n must be even.

<div align="right">𝖰𝖤𝖣</div>

Another example is the proof that there are infinite prime numbers.

Let us define the prime numbers $p_1, \ldots, p_n$ where we assume p_n is the biggest prime number. If we construct a number

$$q = 1 + p_1 \cdot \ldots \cdot p_n$$

This number is bigger than p_n but it is not divisible by any of the other prime numbers will give us always a residual of 1. Therefore the assumption that there is a biggest prime number is false.

<div align="right">𝖰𝖤𝖣</div>

4.3 Proof by Induction

For all $n \in \mathbb{N}$, the sum of the first n positive integers is given by:

$$S(n) = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

We will prove this theorem by induction on n.

Base Case: For $n = 1$:

$$S(1) = 1 = \frac{1(1+1)}{2}$$

The base case holds.

Inductive Step: Assume that the statement holds for some $n = k$, i.e., assume that:

$$S(k) = 1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}$$

We need to show that the statement holds for $n = k + 1$:

$$S(k+1) = S(k) + (k+1)$$

By the inductive hypothesis, we have:

$$\begin{aligned}
S(k+1) &= \frac{k(k+1)}{2} + (k+1) \\
&= \frac{k(k+1)}{2} + \frac{2(k+1)}{2} \\
&= \frac{k(k+1) + 2(k+1)}{2} \\
&= \frac{(k+1)(k+2)}{2}
\end{aligned}$$

Thus, the statement holds for $n = k + 1$. By the principle of mathematical induction, the statement holds for all $n \in \mathbb{N}$.

<div align="right">𝔔𝔈𝔇</div>

4.4   Proof by Exhaustion

The only integer solutions to the equation $x^2 + y^2 = 1$ are $(0, 1), (1, 0), (0, -1), (-1, 0)$.

We will prove this theorem by exhaustion. We will check all possible integer values of x and y such that $x^2 + y^2 = 1$.

The possible integer values for x and y are $-1, 0, 1$. We will check each case:

If $x = 0$:

Then $y^2 = 1$ gives $y = 1$ or $y = -1$.

Solutions: $(0, 1), (0, -1)$.

If $x = 1$:

Then $y^2 = 0$ gives $y = 0$.

Solution: $(1, 0)$.

If $x = -1$:

Then $y^2 = 0$ gives $y = 0$.

Solution: $(-1, 0)$.

Thus, the only integer solutions to the equation are:

$$(0, 1), (1, 0), (0, -1), (-1, 0)$$

<div align="right">𝔔𝔈𝔇</div>

4.5   Proof by Cases

For any integer n, n^2 is even if and only if n is even.

We will prove this theorem by cases.

Case 1: Assume n is even. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even.

Case 2: Assume n is odd. Then we can write $n = 2k + 1$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

This shows that n^2 is odd.

Since both cases have been considered, we conclude that n^2 is even if and only if n is even.

$\mathfrak{QED}$

4.6   Proof by Construction

There exists an irrational number x such that x^2 is rational.

We will construct an irrational number x such that x^2 is rational.

Let $x = \sqrt{2}$. We know that $\sqrt{2}$ is irrational. Now, we compute x^2:

$$x^2 = (\sqrt{2})^2 = 2$$

Since 2 is a rational number, we have constructed an irrational number $x = \sqrt{2}$ such that $x^2 = 2$ is rational. Therefore, the theorem is proved.

$\mathfrak{QED}$

4.7   Proof by Counterexample

The statement all prime numbers are odd is false.

To prove this theorem, we will provide a counterexample.

The number 2 is a prime number, as its only divisors are 1 and 2. However, 2 is even, which contradicts the statement that all prime numbers are odd.

Therefore, the statement All prime numbers are odd is false.

$\mathfrak{QED}$

4.8   Proof by Contrapositive

$$A \implies B \iff \neg B \implies A$$

If n is an integer such that n^2 is odd, then n is odd.

We will prove this theorem by contrapositive. The contrapositive of the statement is: If n is an integer such that n is even, then n^2 is even.

Assume n is even. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even.

Since the contrapositive statement is true, the original statement If n^2 is odd, then n is odd is also true.

<div align="right">𝔔𝔈𝔇</div>

4.9   Proof by Reduction to Absurdity

The square root of 2 is irrational.

We will prove this theorem by reduction to absurdity. Assume that $\sqrt{2}$ is rational. Then we can write:

$$\sqrt{2} = \frac{p}{q}$$

Where p and q are integers with no common factors (i.e., the fraction is in the simplest form).

Squaring both sides gives:

$$2 = \frac{p^2}{q^2}$$

Rearranging gives:

$$p^2 = 2q^2$$

This implies that p^2 is even, and therefore, p must be even (since the square of an odd number is odd). Let $p = 2k$ for some integer k. Substituting this back into the equation gives:

$$(2k)^2 = 2q^2$$
$$4k^2 = 2q^2$$
$$2k^2 = q^2$$

This implies that q^2 is even, and therefore, q must also be even. Since both p and q are even, they have a common factor of 2, which contradicts our assumption that p and q have no common factors. Therefore, our assumption that $\sqrt{2}$ is rational must be false, and thus, $\sqrt{2}$ is irrational.

<div align="right">𝔔𝔈𝔇</div>

4.10   Proof by Analogy

The set of rational numbers is dense in the set of real numbers.

We will prove this theorem by analogy.

Consider the set of rational numbers $\mathbb{Q}$ and the set of real numbers $\mathbb{R}$. The density of $\mathbb{Q}$ in $\mathbb{R}$ means that between any two real numbers, there exists a rational number. For example, between the real numbers 1 and 2, we can find the rational number $\frac{3}{2} = 1.5$. Similarly, between any two real numbers a and b (where $a < b$), we can find a rational number $r = \frac{a+b}{2}$.

This shows that the set of rational numbers is dense in the set of real numbers. Therefore, the theorem is proved by analogy.

<div align="right">𝔔𝔈𝔇</div>

5   The Natural Numbers

In this section we will take a look at the natural numbers, which are the numbers we use for counting. The natural numbers are defined as follows:

We will now define the set of natural numbers, $\mathbb{N}$, via the following 9 axioms. These axioms are known as the *Peano Axioms*. The first 4 axioms define equality on the set $\mathbb{N}$.

   I. For every $x \in \mathbb{N}$, we have $x = x$.   (Reflexivity)

  II. For every $x, y \in \mathbb{N}$, if $x = y$ then $y = x$.   (Symmetry)

 III. For every $x, y, z \in \mathbb{N}$, if $x = y$ and $y = z$ then $x = z$.   (Transitivity)

 IV. For all x, y, if $x \in \mathbb{N}$ and $x = y$, then $y \in \mathbb{N}$.   (Closure of Equality)

The remaining 5 axioms define the structure of $\mathbb{N}$:

  V. $1 \in \mathbb{N}$

 VI. If $x \in \mathbb{N}$, then the successor $S(x) \in \mathbb{N}$.

 VII. There is no $x \in \mathbb{N}$ such that $S(x) = 1$.

VIII. For all $x, y \in \mathbb{N}$, if $S(x) = S(y)$, then $x = y$.

 IX. Let $P(x)$ be a statement about the natural number x. If:

- $P(1)$ is true, and
- for all $n \in \mathbb{N}$, if $P(n)$ is true, then $P(S(n))$ is also true,

    then $P(x)$ is true for all $x \in \mathbb{N}$.   (Mathematical Induction)

As shorthand, we denote:

$$S(1) = 2, \quad S(S(1)) = 3, \quad S(S(S(1))) = 4, \quad \text{and so on.}$$

5.1   Principle of the Ordering

Every non-empty set of natural numbers has a minimum.

5.2   Order in Fields

An *ordered field* is a field F equipped with a total order $<$ that is compatible with the field operations $+$ and $\cdot$. That is, the order satisfies both algebraic and ordering properties.

5.2.1   Order Axioms for Fields

A field F is called an *ordered field* if it satisfies the following properties for all $a, b, c \in F$:

  I. *Trichotomy:* Exactly one of the following holds:

$$a < b, \quad a = b, \quad a > b$$

  II. *Transitivity:* If $a < b$ and $b < c$, then $a < c$.

 III. *Additive Compatibility:* If $a < b$, then $a + c < b + c$.

 IV. *Multiplicative Compatibility:* If $0 < a$ and $0 < b$, then $0 < ab$.

These properties ensure that arithmetic operations respect the order structure.

5.2.2 Consequences of the Order Axioms

- $a < b \Rightarrow -b < -a$

- $a < b \wedge c < 0 \Rightarrow ac > bc$

- Squares are always non-negative: $a^2 \geq 0$

- The order is total: any two elements are comparable

5.2.3 Examples of Ordered Fields

- $\mathbb{Q}$: Rational numbers with the usual order

- $\mathbb{R}$: Real numbers with the usual order

- $\mathbb{C}$: Complex numbers are **not** an ordered field, since $i^2 = -1 < 0$ would violate positivity of squares

5.3 Propositions and Proofs

5.3.1 Proposition 1: $n \neq m \implies S(n) \neq S(m)$

Proof:

We will prove this by contradiction. Assume $n \neq m$ and $S(n) = S(m)$. By Axiom 8, we have $n = m$, which is a contradiction. Therefore, $S(n) \neq S(m)$.

5.3.2 Proposition 2: For any $n \in \mathbb{N}$, $n \neq S(n)$

Proof:

$M = \{n \in \mathbb{N} \| n \neq S(n)\}$

By $1 \neq S(n)$ for any $n \in \mathbb{N}$, this implies that 1 is part of the set M. This implies that $S(n) \neq S(S(n)) \implies S(n) \in M$ By Axiom 9 $M = \mathbb{N}$

5.3.3 Proposition 3: $n \neq 1 \ \exists m \in \mathbb{N} \mid n = S(m)$

Proof:

$$M = \{1\} \cup \{n \in \mathbb{N} | \text{Proposition 3 is true}\}$$

We know that 1 ins in the set M. And by proposition 1 we know that

$$S(n) = S(S(m)) \implies S(n) \in M$$

And by Axiom 9 we know that $M = \mathbb{N}$

5.4 Definition of Addition in $\mathbb{N}$

For any pair n, m $\in \mathbb{N}$ there is a unique way to define

$$\text{Add}(n, m) = n + m$$

1. **Base Case:** $n + 1 = S(n)$

2. **Inductive Step:** $n + S(m) = S(n + m) \iff S(n + m)$

Uniqueness:

Suppose: A & B satisfy our conditions. Fix n and then let $M = \{m \in \mathbb{N} | A(n, m) = B(n, m)\}$. Then

$$A(n, 1) = S(n) = B(n, 1) \implies 1 \in M$$

$$m \in M \implies A(n,m) = B(n,m) \implies A(n,S(m)) = S(A(n,m)) = S(B(n,m)) = B(n,S(m))$$

$$\implies A(n,S(m)) = B(n,S(m))$$

by Axiom 9 we know that $M = \mathbb{N}$ and A = B

Construction: For $n = 1$ Define $A(n,m) = S(m)$

1. $A(n,1) = S(n) = S(1)$

2. $A(n,S(m)) = S(A(n,m)) = S(S(m))$

Define: $A(S(n),S(m)) = S(A(n,m))$

1. $A(S(n),1) = S(A(n,1)) = S(S(n))$

2. $A(S(n),S(m)) = S(A(n,m)) = S(S(m))$

Commutativity of Addition:

The proposition says $n + m = m + n$ for any $n, m \in \mathbb{N}$. We will prove this by induction on m. Fix n and consider $M = \{n \in \mathbb{N} | A(n,m) = A(n,m)\}$

Now recall that $A(n,1) = S(n)$

For $n = 1$:

$$A(n,1) = S(1) = A(1,n) \implies 1 \in \mathbb{N}$$

also $A(n,k) = 1 + k \implies 1 + m = S(m) \implies 1 + m = m + 1 \implies 1 \in \mathbb{N}$

Suppose: $n \in \mathbb{N} \implies n + m = m + n$ or $A(n,m) = A(m,n)$

By construction $A(S(n),m) = S(A(n,m))$ and by definition $A(S(n),m) = S(A(n,m)) = A(m,S(n)) = S(n) + m = m + S(n) \implies S(n) \in M \implies$ by induction $M = \mathbb{N}$.

6 The Archimedean Principle

For any real number $x \in \mathbb{R}$, there exists a natural number $n \in \mathbb{N}$ such that $n > x$.

In other words, no matter how large a real number you choose, there is always a natural number that is larger. Similarly, for any positive real number $\epsilon > 0$, there exists a natural number $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$.

Proof:

We prove the Archimedean Principle by contradiction.

Assume that there exists some real number $x \in \mathbb{R}$ such that $n \leq x$ for all $n \in \mathbb{N}$. That is, x is an upper bound for the set $\mathbb{N} \subset \mathbb{R}$.

Let $S = \sup(\mathbb{N})$, the least upper bound of $\mathbb{N}$. Then $S - 1 < \sup(\mathbb{N})$, so $S - 1$ is not an upper bound of $\mathbb{N}$. Hence, there exists $n_0 \in \mathbb{N}$ such that:

$$n_0 > S - 1 \Rightarrow n_0 + 1 > S$$

But $n_0 + 1 \in \mathbb{N}$, which contradicts the assumption that S is an upper bound of $\mathbb{N}$. Therefore, our assumption must be false, and the theorem is proven.

$$\mathfrak{QED}$$

6.1 Equivalent Formulations

The Archimedean Principle is often stated in different but equivalent ways:

- For any $\epsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$.
- For any $a, b \in \mathbb{R}$ with $a > 0$, there exists $n \in \mathbb{N}$ such that $na > b$.

6.2 Applications

1. **Density of Rational Numbers:** The Archimedean Principle helps in proving that between any two real numbers, there exists a rational number.

2. **Limits and Infinitesimals:** It ensures that sequences like $\left\{ \frac{1}{n} \right\}$ converge to 0, foundational in real analysis and calculus.

3. **Bounding Functions:** It is used in analysis to show that functions do not grow faster than natural numbers in certain contexts.

4. **Non-Existence of Infinitely Small Numbers:** The principle implies that real numbers do not contain infinitesimals (nonzero numbers smaller than all $\frac{1}{n}$), distinguishing $\mathbb{R}$ from non-standard number systems.

7 Fundamental Theorem of Arithmetic

- The Fundamental Theorem of Arithmetic says that every integer greater than 1 can be factored uniquely into a product of primes.

- Euclid's lemma says that if a prime divides a product of two numbers, it must divide at least one of the numbers.

- The least common multiple $[a, b]$ of nonzero integers a and b is the smallest positive integer divisible by both a and b.

7.1 Fundamental Theorem of Arithmetic

Every integer greater than 1 can be written in the form

$$p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

Where $n_i \geq 0$ and the p_i are distinct primes. The factorization is unique, except possibly for the order of the factors.

Example.

$$4312 = 2 \cdot 2156 = 2 \cdot 2 \cdot 1078 = 2 \cdot 2 \cdot 2 \cdot 539 = 2 \cdot 2 \cdot 2 \cdot 7 \cdot 77 = 2 \cdot 2 \cdot 2 \cdot 7 \cdot 7 \cdot 11$$

That is,

$$4312 = 2^3 \cdot 7^2 \cdot 11$$

7.2 Lemmas

Lemma

If $m \mid pq$ and $\gcd(m, p) = 1$, then $m \mid q$.

Proof:

Write $1 = \gcd(m, p) = am + bp$ for some $a, b \in \mathbb{Z}$. Then

$$q = amq + bpq$$

Since $m \mid amq$ and $m \mid bpq$ (because $m \mid pq$), we conclude $m \mid q$.

$$\mathfrak{QED}$$

Lemma

If p is prime and $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_i$ for some i.

Proof (Case $n = 2$):

Suppose $p \mid a_1 a_2$, and $p \nmid a_1$. Then $\gcd(p, a_1) = 1$, and by the previous lemma, $p \mid a_2$.

For general $n > 2$: Assume the result is true for $n-1$. Suppose $p \mid a_1 a_2 \cdots a_n$. Group as $(a_1 a_2 \cdots a_{n-1}) a_n$.

By the $n = 2$ case, either $p \mid a_n$ or $p \mid a_1 a_2 \cdots a_{n-1}$, and by induction, $p \mid a_i$ for some i.

$$\mathfrak{QED}$$

7.3 Proof of the Fundamental Theorem

Existence:

Use induction on $n > 1$. Base case: $n = 2$ is prime.

Inductive step: If n is prime, done. Otherwise, $n = ab$, with $1 < a, b < n$. By induction, both a and b factor into primes, so n does too.

Uniqueness:

Suppose:

$$p_1^{m_1} \cdots p_j^{m_j} = q_1^{n_1} \cdots q_k^{n_k}$$

With all p_i and q_i distinct primes.

Since p_1 divides the LHS, it divides the RHS. So $p_1 \mid q_i^{n_i}$ for some i, hence $p_1 = q_i$. Re-order so $p_1 = q_1$. Then:

If $m_1 > n_1$, divide both sides by $q_1^{n_1}$:

$$p_1^{m_1 - n_1} \cdots p_j^{m_j} = q_2^{n_2} \cdots q_k^{n_k}$$

But then p_1 divides LHS but not RHS, contradiction. So $m_1 = n_1$. Cancel and repeat.

Eventually, all p_i match with some q_i, and the exponents are equal. So the factorizations are the same up to order.

$$\mathfrak{QED}$$

7.4 Least Common Multiple

The least common multiple of a and b, denoted $[a, b]$, is the smallest positive integer divisible by both.

Example:

$$[6, 4] = 12, \quad [33, 15] = 165$$

Fact:

$$[a, b] \cdot \gcd(a, b) = ab$$

Let:

$$a = p_1 \cdots p_l q_1 \cdots q_m, \quad b = q_1 \cdots q_m r_1 \cdots r_n$$

Then:

$$
\begin{aligned}
\gcd(a, b) &= q_1 \cdots q_m \\
[a, b] &= p_1 \cdots p_l q_1 \cdots q_m r_1 \cdots r_n \\
ab &= p_1 \cdots p_l q_1^2 \cdots q_m^2 r_1 \cdots r_n
\end{aligned}
$$

So:

$$[a, b] \cdot \gcd(a, b) = ab$$

Example:

$$\gcd(36, 90) = 18, \quad [36, 90] = 180, \quad 36 \cdot 90 = 32400 = 18 \cdot 180$$

8 Real Numbers

Let K be an ordered field. On the set

$$\mathrm{ch}(K) := \{x : \mathbb{N} \to K \mid x \text{ is a Cauchy sequence}\}$$

and on the set

$$c(K) := \{x : \mathbb{N} \to K \mid x \text{ is a convergent sequence}\}$$

we can define an addition and a multiplication using 1.2.55 and 1.2.57 as follows:

If $x = (x_n)_{n \in \mathbb{N}}$ and $y = (y_n)_{n \in \mathbb{N}}$ are Cauchy sequences (respectively, convergent sequences), then their sum is defined as

$$x + y := (x_n)_{n \in \mathbb{N}} + (y_n)_{n \in \mathbb{N}} := (x_n + y_n)_{n \in \mathbb{N}}$$

and their product is defined as

$$x \cdot y := (x_n)_{n \in \mathbb{N}} \cdot (y_n)_{n \in \mathbb{N}} := (x_n \cdot y_n)_{n \in \mathbb{N}}$$

The sum and product satisfy all field axioms except for the existence of the multiplicative inverse. The zero element is $0_\mathbb{N} = (0, 0, \ldots)$, the unit element is $1_\mathbb{N} = (1, 1, \ldots)$, and the additive inverse of $x = (x_n)_{n \in \mathbb{N}}$ is $-x = (-x_n)_{n \in \mathbb{N}}$. We demonstrate the distributive law as an example:

Let $x = (x_n)_{n \in \mathbb{N}}, y = (y_n)_{n \in \mathbb{N}}, z = (z_n)_{n \in \mathbb{N}}$ be Cauchy sequences (convergent sequences). Then we have:

$$
\begin{aligned}
x(y + z) &= (x_n)_{n \in \mathbb{N}} \cdot \big((y_n)_{n \in \mathbb{N}} + (z_n)_{n \in \mathbb{N}}\big) \\
&= (x_n)_{n \in \mathbb{N}} \cdot (y_n + z_n)_{n \in \mathbb{N}} \\
&= (x_n(y_n + z_n))_{n \in \mathbb{N}} \\
&= (x_n y_n + x_n z_n)_{n \in \mathbb{N}} \\
&= (x_n y_n)_{n \in \mathbb{N}} + (x_n z_n)_{n \in \mathbb{N}} \\
&= xy + xz.
\end{aligned}
$$

We now aim to construct the ordered field $\mathbb{R}$ of the real numbers; it will have the following properties:

- There exists an injective mapping $j : \mathbb{Q} \to \mathbb{R}$ which respects addition, multiplication, and order, such that the following holds:

  For all $z, w \in \mathbb{R}$ with $z < w$, there exists an $x \in \mathbb{Q}$ such that

$$z < j(x) < w$$

- Every Cauchy sequence in $\mathbb{R}$ converges.

Via j, we identify $\mathbb{Q}$ with $j(\mathbb{Q})$ and consider $\mathbb{Q}$ as a subset of $\mathbb{R}$. In $\mathbb{R}$, the following will additionally hold:

- For all $y > 0$ and $n \in \mathbb{N}$, the equation $x^n = y$ has a solution.

- Every bounded above subset of $\mathbb{R}$ has a supremum.

We define the following relation on the set $\mathrm{ch}(\mathbb{Q})$ of all Cauchy sequences in $\mathbb{Q}$:

$$x \sim y \quad \text{if and only if} \quad x - y \text{ is a null sequence}$$

That is, $(x_n)_{n \in \mathbb{N}} \sim (y_n)_{n \in \mathbb{N}}$ if and only if

$$x_n - y_n \to 0 \quad (n \to \infty)$$

8.1 Definition

The set

$$\mathbb{R} := \{[x]_\sim : x \in \mathrm{ch}(\mathbb{Q})\}$$

is called the set of real numbers.

Analogous to the construction of the rational numbers, the real numbers consist of equivalence classes. Roughly speaking, an equivalence class consists of those Cauchy sequences in $\mathbb{Q}$ that exhibit the same limit behavior.

Equipped with the addition

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad [x], [y] \mapsto [x] + [y] := [x + y]$$

and the multiplication

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad [x], [y] \mapsto [x] \cdot [y] := [xy]$$

$\mathbb{R}$ is a field. The zero element is $[0_\mathbb{N}]$, and the unit element is $[1_\mathbb{N}]$.

9 Complex Numbers

A complex number is a number of the form:

$$z = a + bi,$$

where $a, b \in \mathbb{R}$, and i is the imaginary unit defined by $i^2 = -1$. The set of all complex numbers is denoted by $\mathbb{C}$.

9.1 The Complex Plane

Complex numbers can be represented graphically in the *complex plane*, where the horizontal axis represents the real part and the vertical axis the imaginary part.
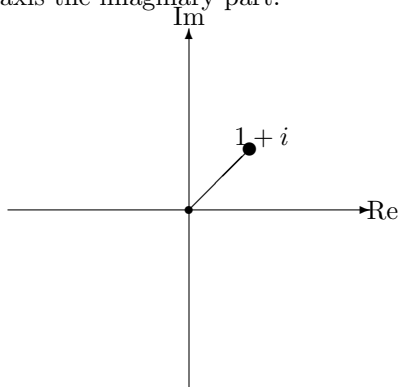
The point $1 + i$ is located at (1,1), showing 1 unit on the real axis and 1 unit on the imaginary axis.

9.2 Conjugate of a Complex Number

The *conjugate* of a complex number $z = a + bi$ is denoted $\overline{z}$ and is defined as:

$$\overline{z} = a - bi$$

Geometrically, it reflects the point z across the real axis in the complex plane. Conjugates are useful in division and in finding the modulus, since:

$$z \cdot \overline{z} = a^2 + b^2 = |z|^2$$

9.3 Operations in Cartesian Coordinates

Let $z_1 = a + bi$ and $z_2 = c + di$ be two complex numbers.

Addition:

$$z_1 + z_2 = (a + c) + (b + d)i$$

Multiplication:

$$z_1 \cdot z_2 = (ac - bd) + (ad + bc)i$$

Quotient:

$$\frac{z_1}{z_2} = \frac{(a + bi)}{(c + di)} \frac{(c - di)}{(c - di)} = \frac{(a + bi)(c - di)}{c^2 + d^2} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2}$$

9.4   Polar Coordinates

A complex number can also be expressed in polar form as:

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta},$$

where:

$$r = |z| = \sqrt{a^2 + b^2} \quad \text{(modulus)}$$
$$\theta = \arg(z) = \tan^{-1}\left(\frac{b}{a}\right) \quad \text{(argument)}$$

The value of θ depends on the quadrant where the complex number lies:

- Quadrant I: $a > 0, b > 0$ - use $\tan^{-1}(b/a)$
- Quadrant II: $a < 0, b > 0$ — add π to $\tan^{-1}(b/a)$
- Quadrant III: $a < 0, b < 0$ — add π to $\tan^{-1}(b/a)$
- Quadrant IV: $a > 0, b < 0$ — use $\tan^{-1}(b/a)$

9.5   Multiplication and Division in Polar Coordinates

Given:

$$z_1 = r_1 e^{i\theta_1}, \quad z_2 = r_2 e^{i\theta_2},$$

Multiplication:

$$z_1 \cdot z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$

Division:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}$$

This polar form is especially useful in simplifying powers and roots of complex numbers using De Moivre's Theorem.

9.6   Exponentiation and Roots (De Moivre´s Theorem)

Let $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$ be a complex number in polar form.

9.6.1   Exponentiation

To raise z to the power $n \in \mathbb{N}$, we use De Moivre's Theorem:

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta)) = r^n e^{in\theta}$$

9.6.2   Roots of Complex Numbers

To find the nth roots of a complex number $z = re^{i\theta}$, we use the formula:

$$z^{1/n} = r^{1/n}\left(\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right)\right), \quad k = 0, 1, \ldots, n-1$$

This yields n distinct roots, each separated by an angle of $\frac{2\pi}{n}$ in the complex plane.

9.7 Example: Solve $z^4 = 1 + \sqrt{3}i$

Step 1: Convert RHS to polar form.

Let $w = 1 + \sqrt{3}i$. Real part: $a = 1$, Imaginary part: $b = \sqrt{3}$

$$r = |w| = \sqrt{1^2 + \left(\sqrt{3}\right)^2} = \sqrt{1 + 3} = 2$$
$$\theta = \arg(w) = \tan^{-1}\left(\frac{\sqrt{3}}{1}\right) = \frac{\pi}{3}$$

So,

$$w = 2\left(\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)\right)$$

Step 2: Solve $z^4 = w \to z = w^{1/4}$

Using the root formula:

$$z_k = 2^{1/4}\left(\cos\left(\frac{\pi + 2k\pi}{12}\right) + i\sin\left(\frac{\pi + 2k\pi}{12}\right)\right), \quad k = 0, 1, 2, 3$$

So the four roots are:

$$z_0 = 2^{1/4}\left(\cos\left(\frac{\pi}{12}\right) + i\sin\left(\frac{\pi}{12}\right)\right)$$
$$z_1 = 2^{1/4}\left(\cos\left(\frac{5\pi}{12}\right) + i\sin\left(\frac{5\pi}{12}\right)\right)$$
$$z_2 = 2^{1/4}\left(\cos\left(\frac{9\pi}{12}\right) + i\sin\left(\frac{9\pi}{12}\right)\right) = 2^{1/4}\left(\cos\left(\frac{3\pi}{4}\right) + i\sin\left(\frac{3\pi}{4}\right)\right)$$
$$z_3 = 2^{1/4}\left(\cos\left(\frac{13\pi}{12}\right) + i\sin\left(\frac{13\pi}{12}\right)\right)$$

These represent the four complex 4th roots of $1 + \sqrt{3}i$, equally spaced around the circle of radius $2^{1/4}$ in the complex plane.

9.8 Solving Equations with Complex Numbers

Solving equations in $\mathbb{C}$ can involve various forms. Here are the most common cases:

9.8.1 Linear Equations:

Solve for z in $az + b = 0$, where $a, b \in \mathbb{C}$, $a \neq 0$:

$$z = -\frac{b}{a}$$

9.8.2 Equations Involving the Conjugate:

Solve for z in equations like $z + \bar{z} = 4$. Let $z = x + iy$, then $\bar{z} = x - iy$. So:

$$z + \bar{z} = 2x \quad \to \quad x = 2 \quad \Rightarrow \quad z = 2 + iy$$

The imaginary part remains free unless further constraints are given.

9.8.3   Modulus Equations:

Solve $|z| = r$. Let $z = x + iy$, then:

$$\sqrt{x^2 + y^2} = r \to x^2 + y^2 = r^2$$

This is a circle of radius r centered at the origin in the complex plane.

9.8.4   Equations Involving $z \cdot \overline{z}$

Recall $z \cdot \overline{z} = |z|^2$. For example, solve:

$$z \cdot \overline{z} = 9 \to |z| = 3$$

Again, a circle in the complex plane of radius 3.

9.8.5   Quadratic Equations:

Complex roots occur naturally. For example:

$$z^2 + 1 = 0 \to z^2 = -1 \Rightarrow z = \pm i$$

9.8.6   General Polynomial Equations:

Use De Moivre's Theorem or polar form. Example:

$$z^n = w \to z_k = \sqrt[n]{|w|} \cdot e^{i\left(\frac{\arg(w) + 2k\pi}{n}\right)}, \quad k = 0, 1, \ldots, n-1$$

Example 1:

Solve:

$$\left(\frac{2 + 3i}{1 + i} + \frac{4 + 5i}{2 - 2i}\right)\hat{z} = \frac{i + 2}{i}$$
$$\left(\frac{-3 - i}{4}\right)\hat{z} = \frac{i + 2}{i}$$
$$\hat{z} = \frac{i + 2}{i} : \frac{-3 - i}{4}$$

Example 2:

Solve:

$$z - 3i + (2 - i)\hat{z} + 2 = 0$$

In this case we let $z = x + iy$ and $\hat{z} = x - iy$.

$$z = 3i - (2 - i)\hat{z} - 2$$
$$x + iy = 3i - (2 - i)(x - iy) - 2$$
$$x + iy = 3i - [2x - 2yi - xi + yi^2] - 2$$
$$x + yi = 3i - 2 + 2x + 2yi + xi + y$$
$$x + yi = (y - 2 - 2x) + i(3 + 2y + x)$$
$$x = y - 2 - 2x \quad y = 3 + 2y + x$$
$$x = \frac{y - 2}{3} \quad y = 3 + 2y + \frac{y - 2}{3} = -7$$
$$x = \frac{-7 - 2}{3} = -3$$
$$z = -3 - 7i \quad \hat{z} = -3 + 7i$$

9.9 The Complex Logarithm

The logarithm of a complex number is multivalued due to the periodic nature of the complex exponential.

Let $z = re^{i\theta}$ with $r > 0$, $\theta \in \mathbb{R}$. Then:

$$\log z = \ln r + i(\theta + 2\pi k), \quad k \in \mathbb{Z}$$

Here:

- $\ln r$ is the natural (real) logarithm of the modulus.
- θ is the principal argument $\arg(z) \in (-\pi, \pi]$.
- The term $2\pi k$ accounts for the infinitely many branches of the logarithm in $\mathbb{C}$.

Principal Value:

The principal value of the complex logarithm is often written:

$$\text{Log } z = \ln |z| + i \,\text{Arg}(z), \quad \text{where Arg}(z) \in (-\pi, \pi]$$

Example:

Let $z = -1$. Then:

$$|z| = 1, \quad \arg(z) = \pi, \quad \rightarrow \log(-1) = i(\pi + 2\pi k), \quad k \in \mathbb{Z}$$

$$\rightarrow \text{Log}(-1) = i\pi$$

The multivalued nature of $\log z$ is crucial in advanced complex analysis, especially in defining analytic continuations and branch cuts.

9.10 Complex Exponents

$a^x \approx 1\left(a + \alpha \frac{x}{N}\right)^N \rightarrow e^z := \lim_{N \to \inf} \left(1 + \frac{z}{N}\right)^N$ The process above is called linearization of the exponential function by zooming $\alpha \approx \frac{dy}{d}$

For $e^{ci} = \cos \theta + i \sin \theta$ every exponentiation of a complex number is a rotation in the complex plane.

$$e^{ic} = \lim_{N \to \inf} \left(1 + \frac{ic}{N}\right)^N$$

Now imagine that in a sector of a circumference you put triangles one above the other with base of length one and a height of $\frac{C}{N}$ and an angle of δ

$$\tan\delta \approx \delta \text{ for } \delta \ll 1$$

$$1 + \frac{ci}{N} = 1\angle\frac{c}{N}, N \gg 1$$

$$e^{ic} = \lim_{N\to\inf}\left(1 + \frac{c}{N}\right)^N \to e^{ci} = 1\angle c = \cos c + i\sin c$$

9.11 Euler's Formula Proof

We know that

$$e^{i\pi} = -1 \text{ and } e^{i\theta} = |r|(\cos\theta + i\sin\theta)$$

$$e^z = \lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n \implies e^{i\pi} = \lim_{n\to\infty}\left(1 + \frac{i\pi}{n}\right)^n = -1$$

$$\implies \lim_{n\to\infty}|r_n| = 1 \quad \lim_{n\to\infty}\theta = 0+$$

Now we can demonstrate the formula.

$$|r_n| = \left(1 + \left|\frac{z}{n}\right|^n\right) \implies \left(\sqrt{1 + \frac{\pi^2}{n^2}}\right)$$

$$\theta = \sum_{k=1}^{n} n\arctan\frac{\pi}{2} = n\arctan\frac{\pi}{n}$$

$$\lim_{n\to\infty}\left(\sqrt{1 + \frac{\pi^2}{n^2}}\right)^n = \lim_{n\to\infty}\left(1 + \frac{\pi^2}{2n}\right)^{\frac{n}{2}} = \lim_{n\to\infty}e^{\ln\left(1 + \frac{\pi^2}{2}\right)\frac{n}{2}} = e^0 = 1$$

$$\lim_{n\to\infty}\theta = \lim_{n\to\infty}n\arctan\frac{\pi}{n} = \lim_{n\to\infty}n^{-1}\arctan\frac{\pi}{n} = 0$$

Thus, for $e^{i\pi} = 1$ for $x = \pi \forall x \in \lim r_n(x) = 1$ and $\lim\theta(x) = x$

$$\mathfrak{QED}$$

10 Topology

In this section, we introduce essential vocabulary used in topology. Each term is accompanied by a brief explanation and its formal mathematical definition.

10.1 Introduction to topological nomenclature

10.1.1 Open Set

A subset $U \subseteq X$ of a topological space is called *open* if for every point $x \in U$, there exists an $\varepsilon > 0$ such that the open ball $B_\varepsilon(x) \subseteq U$. Intuitively, an open set contains none of its boundary points and every point has some wiggle room around it.

10.1.2 Closed Set

A subset $A \subseteq X$ is called *closed* if its complement $X \setminus A$ is open. Equivalently, A contains all its limit points. That is, A is closed if it includes its boundary.

10.1.3 Interior Point

A point $x \in A$ is an *interior point* of $A \subseteq X$ if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq A$. The set of all interior points of A is called the *interior* of A, denoted $\text{int}(A)$.

10.1.4 Boundary Point

A point $x \in X$ is a *boundary point* of a set $A \subseteq X$ if every open ball around x contains both points in A and in $X \setminus A$. The set of all boundary points is called the *boundary* of A, denoted ∂A.

10.1.5 Accumulation Point / Limit Point

A point $x \in X$ is an *accumulation point* of a set $A \subseteq X$ if every open ball $B_\varepsilon(x)$ contains a point of $A \setminus \{x\}$. In other words, points of A cluster arbitrarily close to x, even if $x \notin A$.

10.1.6 Isolated Point

A point $x \in A$ is an *isolated point* if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \cap A = \{x\}$. That is, x stands alone in A without other points of A nearby.

10.1.7 Compact Set

A set $K \subseteq X$ is *compact* if every open cover of K has a finite sub-cover. In $\mathbb{R}^n$, this is equivalent to K being closed and bounded (by the Heine–Borel theorem).

10.1.8 Dense Set

A subset $D \subseteq X$ is *dense* in X if every point $x \in X$ is either in D or is a limit point of D. Equivalently, the closure of D is X, i.e., $\overline{D} = X$.

10.1.9 Open Ball

$(B_\varepsilon(x))$

For a metric space (X, d), the *open ball* centered at $x \in X$ with radius $\varepsilon > 0$ is defined as:

$$B_\varepsilon(x) := \{y \in X \mid d(x, y) < \varepsilon\}$$

It represents the set of all points within distance ε from x, excluding the boundary.

11 Inequalities with and without Absolute Value

Inequalities are like normal *equalities* but with ranges of solutions instead of one particular solution. The inequalities without absolute value are straightforward to solve, just like equations with the particularity of when dividing both sides by a negative number the relation $<, >\geq, \leq$ has to be converted to its opposite.

The with absolute value is similar but here we have to account for different cases depending on the expression inside the $|\circ|$.

Example:

Solve $\|x + 1\| \leq 10 + \|2x - 6\|$

Main cases:

$$F_1 : x + 1 \geq 0 \implies x \geq 1$$

$$F_2 : x + 1 < 0 \implies x < -1$$

Subcases:

$$F_{1,1} : 2x - 6 \geq 0 \implies x \geq 3$$

$$x + 1 \geq 10 + 2x - 6$$
$$x \geq -3$$
$$\mathcal{L} = [3; \infty)$$

$$F_{1,2} : 2x - 6 < 0 \implies x < 3$$

$$x + 1 \geq 10 - (2x - 6)$$
$$x \geq 5$$
$$\mathcal{L} = (-\infty; 3]$$

$$F_{2,1} : 2x - 6 \geq 0 \implies x \geq 3$$

$$-(x + 1) \geq 10 + 2x - 6$$
$$x \geq \frac{-5}{3}$$
$$\mathcal{L} = \emptyset$$

$$F_{2,2} : 2x - 6 < 0 \implies x < 3$$

$$-(x + 1) \geq 10 - (2x - 6)$$
$$x \leq 17$$
$$\mathcal{L} = (-\infty, 3)$$

Union of the solutions:

$$\mathcal{L} = (-\infty; 3] \cup [3; \infty) \cup \emptyset \cup (-\infty, 3) = (-\infty, \infty)$$

11.1 Reciprocal Inequality

$$a < b < c \implies \frac{1}{a} > \frac{1}{b} > \frac{1}{c}$$

11.2 Bernoulli Inequality

Given $x \in \mathbb{R}, x > -1, x \neq 0$. Then for $n \in \mathbb{N}, n \geq 2$

$$(1+x)^n > 1 + nx$$

Proof:

For $n = 2$

$$(1+x)^2 > 1 + 2x$$

Then for $n + 1$ with $n > 2$

$$(1+x)^{n+1} > 1 + (n+1)x$$
$$(1+x)^n > 1 + (n+1)x$$

$\mathfrak{QED}$

12   Means and Proofs

In math there are a lot of means. In this section I will show some of them with the corresponding proofs.

Arithmetic Mean: The arithmetic mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$A = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Geometric Mean: The geometric mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

Harmonic Mean: The harmonic mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

Quadratic Mean: The quadratic mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$Q = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$$

12.1   Proof of the Arithmetic Mean-Geometric Mean Inequality

Let $a_1, a_2, \ldots, a_n > 0$. We will prove by induction that:

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}$$

With equality if and only if $a_1 = a_2 = \cdots = a_n$.

Base Case: $n = 2$

We want to prove:

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2}$$

Let $a_1, a_2 > 0$. Then by the identity

$$\left(\frac{a_1 - a_2}{2} \right)^2 \geq 0,$$

we get

$$\frac{a_1^2 - 2a_1 a_2 + a_2^2}{4} \geq 0 \Rightarrow a_1^2 + a_2^2 \geq 2a_1 a_2.$$

So,

$$(a_1 + a_2)^2 \geq 4a_1 a_2 \Rightarrow \left(\frac{a_1 + a_2}{2} \right)^2 \geq a_1 a_2,$$

and taking square roots gives the desired result:

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2}.$$

For $n \geq 2$

$$A_{n+1} := (\sum_{i=1}^{n+1} a_i)/(n+1) = \frac{a_1 + a_2 + \cdots + a_n + a_{n+1}}{n+1}$$

$$G_{n+1} := \sqrt[n+1]{a_1 a_2 \cdots a_n a_{n+1}} = \sqrt[n+1]{(a_1 a_2 \cdots a_n) a_{n+1}}$$

$$A_{n-1}^{n+1} = (\frac{a_1 + a_2 + \cdots + a_{n+1}}{n+1})^{n-1} = A_{n+1}^{n-1}$$

$$G_{n+1}^{n+1} := (\sqrt[n+1]{a_1 a_2 \cdots a_n a_{n+1}})^{n+1} = (a_1 a_2 \cdots a_n) a_{n+1} = \sqrt[n]{(a_1 a_2 \cdots a_n)^n} a_{n+1}^{n+1} = G_n^n a_{n+1}^{n+1}$$

Then

$$G_{n+1}^{n+1} A_{n+1}^{n-1} = G_n^n a_{n+1}^{n+1} A_{n+1}^{n-1} \leq A_n^n A_{n+1}^{n-1}$$

This comes from

$$G_n \leq A_n$$
$$G_n^n \leq A_n^n$$
$$G_n^n a_{n+1} \leq A_n^n a_{n+1}$$
$$G_n^n a_{n+1}^{n+1} A_{n+1}^{n-1} \leq A_n^n A_{n+1}^{n-1} = \left(A_n^n (a_{n+1} A_{n+1})^{n-1} \right)^{\frac{n}{n}}$$
$$\leq A_n^n \left(\frac{a_{n+1} + A_{n+1} + \cdots A_{n+1}}{n} \right)^n$$
$$A_n^n \left(\frac{a_{n+1} + (n-1) A_{n+1}}{n} \right)^n$$
$$\left(A_n \frac{a_{n+1} + (n-1) A_{n+1}}{n} \right)^n$$

Note that

$$\left(A_n \frac{a_{n+1} + (n-1) A_{n+1}}{n} \right) \rightarrow \left(\sqrt{A_n \frac{a_{n+1} + (n-1) A_{n+1}}{n}} \right)^2$$
$$\leq \left(\frac{A_n + \frac{a_{n+1} + (n-1) A_{n+1}}{n}}{2} \right)^{2n}$$

Now with power of n we have

$$\left(A_n + \frac{a_{n+1} + (n-1) A_{n+1}}{n} \right)^n \leq \left(\frac{A_n + \frac{a_{n+1} + (n-1) A_{n+1}}{n}}{2} \right)^{2n}$$
$$= \left(\frac{A_n}{2} + \frac{a_{n+1} + (n-1) A_{n+1}}{2n} \right)^{2n}$$
$$= \left(\frac{A_n n}{2n} + \frac{a_{n+1} + (n-1) A_{n+1}}{2n} \right)^{2n}$$
$$= \left(\frac{A_n n + a_{n+1} + (n-1) A_{n+1}}{2n} \right)^{2n}$$
$$= \left(\frac{A_{n+1}(n+1) + (n-1) A_{n+1}}{2n} \right)^{2n}$$
$$= \left(\frac{2n A_{n+1}}{2n} \right)^{2n} = (A_{n+1})^{2n}$$

65

Now we have proven that $G_{n+1}^{n+1} A_{n+1}^{n-1} \leq (A_{n+1})^{2n}$ by dividing both sides by A_{n+1}^{n-1} we get

$$G_{n+1}^{n+1} \leq A_{n+1}^{n+1}$$

$\mathfrak{QED}$

12.2   Proof of the Harmonic Mean Geometric Mean Inequality

Let $a_1, a_2, \ldots, a_n > 0$. We will prove the inequality.

We know that $G_n \leq A_n$

$$G_n \leq A_n$$
$$\sqrt[n]{x_1 \ldots x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$
$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} \leq \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

This concludes the proof of the harmonic mean-geometric mean inequality.

$\mathfrak{QED}$

13 Solving Polynomial Equations

In this section, we discuss the solution formulas for polynomial equations of degrees 2 and 3: the PQ formula, the ABC formula, and the Cubic formula. We also derive each of them step by step.

13.1 The PQ Formula

The PQ formula solves quadratic equations of the form:

$$x^2 + px + q = 0$$

13.1.1 Derivation

To derive the PQ formula, we complete the square:

$$
\begin{aligned}
x^2 + px + q &= 0 \\
x^2 + px &= -q \\
x^2 + px + \left(\frac{p}{2}\right)^2 &= -q + \left(\frac{p}{2}\right)^2 \\
\left(x + \frac{p}{2}\right)^2 &= \left(\frac{p}{2}\right)^2 - q \\
x + \frac{p}{2} &= \pm\sqrt{\left(\frac{p}{2}\right)^2 - q} \\
x &= -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}
\end{aligned}
$$

13.1.2 PQ formula:

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

13.2 The Quadratic Formula

The general quadratic equation is:

$$ax^2 + bx + c = 0 \quad \text{with } a \neq 0$$

13.2.1 Derivation of the PQ-Formula

We normalize the equation by dividing through by a and complete the square:

$$
\begin{aligned}
ax^2 + bx + c &= 0 \\
x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\
x^2 + \frac{b}{a}x &= -\frac{c}{a} \\
x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 &= -\frac{c}{a} + \left(\frac{b}{2a}\right)^2 \\
\left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \\
x + \frac{b}{2a} &= \pm\frac{\sqrt{b^2 - 4ac}}{2a} \\
x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\end{aligned}
$$

13.2.2 ABC formula:

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

13.3 The Cubic Formula

To solve a general cubic equation:

$$ax^3 + bx^2 + cx + d = 0,$$

we first reduce it to a depressed cubic using a substitution.

Step 1: Depress the cubic

Let $x = t - \frac{b}{3a}$, then the equation becomes:

$$t^3 + pt + q = 0$$

with:

$$p = \frac{3ac - b^2}{3a^2}, \quad q = \frac{2b^3 - 9abc + 27a^2d}{27a^3}$$

Step 2: Solve the depressed cubic using Cardano's method

Assume a solution of the form:

$$t = u + v$$

Then substitute and simplify:

$$(u + v)^3 + p(u + v) + q = 0$$

Expanding and setting:

$$u^3 + v^3 + (3uv + p)(u + v) + q = 0$$

To eliminate the $(u + v)$ term, set:

$$3uv + p = 0 \quad \Rightarrow \quad uv = -\frac{p}{3}$$

Now:

$$u^3 + v^3 = -q$$

Let:

$$u^3 = A, \quad v^3 = B \quad \Rightarrow \quad A + B = -q, \quad AB = -\frac{p^3}{27}$$

These are the roots of the quadratic:

$$z^2 + qz - \frac{p^3}{27} = 0$$

Solve for A and B, then take cube roots to get u and v. The final solution is:

$$x = u + v$$

13.3.1 Cardano's Formula (for depressed cubic)

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

This formula gives one real root. The other roots (if real) can be found using trigonometric or complex methods depending on the discriminant.

14 The Binomial Coefficient

In this short section we will cover the definition and Properties of the binomial coefficient without going to deep into the combinatorics meaning or Pascal's Triangle.

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

or as Product

$$\prod_{j=1}^{n} \frac{n-k+j}{j},$$

with n and k as natural numbers (for the moment).

14.1 Properties

- $\binom{n}{k} = \binom{n}{n-k}$
- $\binom{n}{k-1} = \binom{n}{k} \frac{n-k}{k+1}$
- $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
- $\sum_{k=0}^{n} \binom{m}{k} = \binom{n+1}{m+1}$
- $\binom{n}{0} = 1 = \binom{n}{n}$
- $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$
- $\binom{n}{k} = (-1)^k \binom{k-n-1}{k}$

14.2 The binomial Theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Proof:

We will prove that the inequality holds for any n.

For n we have $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$

For $n+1$ we get

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k$$

Now let us manipulate the expression to demonstrate our theorem. Assume that the statement is true for some integer $n \geq 0$, i.e.

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Now for some greater n

$$(a+b)^{n+1} = (a+b)(a+b)^n.$$

Using the induction hypothesis,

$$(a+b)^{n+1} = (a+b) \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k}.$$

Distribute:

$$(a+b)^{n+1} = \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k} b^{k} + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1}.$$

In the second sum, change the index of summation by setting $j = k+1$, then rename j back to k:

$$\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} = \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n+1-k} b^{k}.$$

Now combine both sums:

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \left(\binom{n}{k} + \binom{n}{k-1} \right) a^{n+1-k} b^{k},$$

where we use the convention $\binom{n}{k} = 0$ for $k < 0$ or $k > n$.

By *Pascal's Identity*,

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

Therefore,

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^{k}.$$

<div align="right">𝔔𝔈𝔇</div>

15 Proportionality and the Rule of Three

15.1 Proportionality

Two quantities are said to be *proportional* if their ratio remains constant.

15.1.1 Direct Proportionality

Two quantities a and b are in *direct proportion* if:

$$\frac{a}{b} = k \quad \Rightarrow \quad a = k \cdot b$$

Where k is the constant of proportionality.

Example:

If 2 pencils cost 1, then 4 pencils cost 2. The ratio is constant: $\frac{2}{1} = \frac{4}{2}$.

15.1.2 Inverse Proportionality

Two quantities a and b are in *inverse proportion* if their product is constant:

$$a \cdot b = k$$

Example:

If 4 workers finish a job in 6 hours, then 2 workers would need 12 hours:

$$4 \cdot 6 = 2 \cdot 12 = 24$$

15.2 The Rule of Three (Simple)

The *Rule of Three* is a method to find a fourth value when three values are known and a proportional relationship is assumed.

15.2.1 Direct Rule of Three (Simple)

Given: $a : b = c : x$, solve for x:

$$x = \frac{b \cdot c}{a}$$

Example:

If 3 apples cost 6, how much do 5 apples cost?

$$x = \frac{6 \cdot 5}{3} = 10$$

15.2.2 Inverse Rule of Three (Simple)

If the relationship is inverse:

$$a : b = x : c \quad \Rightarrow \quad x = \frac{a \cdot c}{b}$$

Example:

If 5 people finish a task in 8 hours, how long will 10 people need?

$$x = \frac{5 \cdot 8}{10} = 4$$

15.3   The Rule of Three (Compound)

The *Compound Rule of Three* (or *composed rule of three*) involves more than two variables.

Example:

If 4 machines produce 120 items in 5 hours, how many items will 6 machines produce in 8 hours?

Step 1: Set up proportionally:

$$\text{Items} \propto \text{Machines} \quad (\text{direct}) \qquad \text{Items} \propto \text{Time} \quad (\text{direct})$$

Step 2: Adjust the quantity:

$$\text{Initial: 4 machines, 5 hrs } \rightarrow 120 \text{ items}$$
$$\text{New: 6 machines, 8 hrs } \rightarrow x \text{ items}$$

Step 3: Use proportionality:

$$x = 120 \cdot \frac{6}{4} \cdot \frac{8}{5} = 120 \cdot 1.5 \cdot 1.6 = 288$$

Answer: 288 items.

Summary Table

Type	Formula
Direct Proportion	$x = \frac{b \cdot c}{a}$
Inverse Proportion	$x = \frac{a \cdot c}{b}$
Compound Rule of Three	Multiply by all direct and divide by inverse ratios

16 Factorization Techniques

Factorization is the process of writing a mathematical expression as a product of its factors. This is a fundamental technique in algebra used to simplify expressions, solve equations, and analyze functions.

16.1 Common Factor

Factor out the greatest common divisor (GCD) of all terms.

Example:

$$6x^2 + 9x = 3x(2x + 3)$$

16.2 Difference of Squares

A difference of squares follows the identity:

$$a^2 - b^2 = (a - b)(a + b)$$

Example:

$$x^2 - 16 = (x - 4)(x + 4)$$

16.3 Sum of Cubes:

$$a^3 + b^3 = (a + b)(a^2 - ab + b^2)$$

16.4 Difference of Cubes:

$$a^3 - b^3 = (a - b)(a^2 + ab + b^2)$$

Example:

$$x^3 + 8 = (x + 2)(x^2 - 2x + 4)$$

16.5 Trinomial: Special Case $x^2 + bx + c$

This is the case where $a = 1$. Find two numbers whose product is c and sum is b.

Example:

$$x^2 + 5x + 6 = (x + 2)(x + 3)$$

16.6 Trinomial: General Form $ax^2 + bx + c$

Like in the previous case we are going to use a similar procedure, but we start by multiplying and diving by $\frac{a}{a}$ and then proceed to do the rest. Look at the example

$$
\begin{aligned}
3x^2 - 5x - 2 &= \frac{3(3x^2 - 5x - 2)}{3} \\
&= \frac{(3x)^2 - 5(3x) - 6}{3} \\
&= \frac{(3x - 6)(3x + 1)}{3} \\
&= (x - 2)(3x + 1)
\end{aligned}
$$

16.7 Perfect Square Trinomial

These follow the identities:

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$(a - b)^2 = a^2 - 2ab + b^2$$

Example:

$$x^2 + 6x + 9 = (x + 3)^2$$

16.8 Substitution

Substitute a more complex expression with a single variable, factor, then back-substitute.

Example:

$$x^4 + 2x^2 + 1 \Rightarrow \text{Let } y = x^2 \Rightarrow y^2 + 2y + 1 = (y + 1)^2 \Rightarrow (x^2 + 1)^2$$

16.9 Rationalization of Radicals

Rationalizing removes radicals from the denominator.

Example (Single Radical):

$$\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

Example (Binomial):

$$\frac{1}{\sqrt{3} + 1} = \frac{\sqrt{3} - 1}{(\sqrt{3} + 1)(\sqrt{3} - 1)} = \frac{\sqrt{3} - 1}{2}$$

16.10 Horner's Method

Used to divide a polynomial by a binomial of the form $(x - r)$.

Steps:

1. Write coefficients of the polynomial $P_1(x)$.

2. Find the factors of the term with x^0.

3. Bring down the first coefficient.

4. Multiply it by r, add to next coefficient. Repeat until you get to the term with x^0.

5. If the final addition gives zero then you have found a zero of the polynomial which Now can be written in the following form $P_2(x)(x - r)$. With $P(x)$ being the new polynomial with the coefficients which were computed during the process to get the zero.

Example:

Given is $P_1(x) = x^3 - 6x^2 + 11x - 6$.

The factors of -6 are $\{\pm 1, \pm 2, \pm 3, \pm 6\}$. We use $r = 1$:

$$
\begin{array}{c|cccc}
1 & 1 & -6 & 11 & -6 \\
  &   & 1 & -5 & 6 \\
\hline
  & 1 & -5 & 6 & 0
\end{array}
\Rightarrow P_2(x) = (x^2 - 5x + 6)(x - 1)
$$

Now by repeating the process with $r = 2$

$$\begin{array}{c|rrr} 2 & 1 & -5 & 6 \\ & & 2 & -6 \\ \hline & 1 & -3 & 0 \end{array} \Rightarrow P_3(x) = (x-3)(x-2)(x-1)$$

The result final result is $(x-3)(x-2)(x-1)$

16.11   Long Division of Polynomials

Use the same algorithm as numerical long division.

Example:

$$\begin{array}{r} x^2 + 2x + 2 \\ x - 1 \overline{)\ x^3\ + x^2\qquad -1} \\ \underline{-\ x^3\ + x^2} \\ 2x^2 \\ \underline{-\ 2x^2 + 2x} \\ 2x - 1 \\ \underline{-\ 2x + 2} \\ 1 \end{array}$$

with the result: $x^2 + 2x + 2 + \frac{1}{x-1}$

17   Partial Fractions

17.1   The Simplest Case

In the most common partial fraction decomposition, we split:

$$\frac{N(x)}{(x - a_1)(x - a_2) \cdots (x - a_d)}$$

into a sum of the form:

$$\frac{A_1}{x - a_1} + \cdots + \frac{A_d}{x - a_d}$$

We now show that this decomposition can always be achieved, under the assumption that the a_i are all different and $N(x)$ is a polynomial of degree at most $d - 1$.

17.1.1   Lemma 1

Let $N(x)$ and $D(x)$ be polynomials of degree n and d, respectively, with $n \leq d$. Suppose that a is not a root of $D(x)$. Then there exists a polynomial $P(x)$ of degree $< d$ and a number A such that:

$$\frac{N(x)}{D(x)(x - a)} = \frac{P(x)}{D(x)} + \frac{A}{x - a}$$

Proof:

Let $z = x - a$. Define:

$$\tilde{N}(z) = N(z + a), \quad \tilde{D}(z) = D(z + a)$$

Then:

$$\frac{\tilde{N}(z)}{\tilde{D}(z)z} = \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A}{z} \Rightarrow \frac{\tilde{P}(z)z + A\tilde{D}(z)}{\tilde{D}(z)z}$$

We equate:

$$\tilde{P}(z)z + A\tilde{D}(z) = \tilde{N}(z)$$

Choosing $A = \frac{\tilde{N}(0)}{\tilde{D}(0)}$, the constant terms match. The remainder has no constant term and is divisible by z, so:

$$\tilde{P}(z)z = \tilde{N}(z) - A\tilde{D}(z)$$

Thus, $\tilde{P}(z)$ is a polynomial of degree $< d$.

17.1.2   Recursive Decomposition

Now, consider:

$$\frac{N(x)}{(x - a_1)(x - a_2) \cdots (x - a_d)}$$

Apply Lemma 1 recursively:

$$\frac{N(x)}{(x-a_1)\cdots(x-a_d)} = \frac{A_1}{x-a_1} + \frac{P(x)}{(x-a_2)\cdots(x-a_d)}$$

Then:

$$\frac{P(x)}{(x-a_2)\cdots(x-a_d)} = \frac{A_2}{x-a_2} + \frac{Q(x)}{(x-a_3)\cdots(x-a_d)}$$

Continue until:

$$\frac{N(x)}{(x-a_1)\cdots(x-a_d)} = \frac{A_1}{x-a_1} + \cdots + \frac{A_d}{x-a_d}$$

17.2 Lemma 2

Let $N(x)$ and $D(x)$ be polynomials of degree n and d respectively, with $n < d + m$. Suppose that a is NOT a zero of $D(x)$. Then there is a polynomial $P(x)$ of degree $p < d$ and numbers $A_1, \ldots, A_m$ such that

$$\frac{N(x)}{D(x)(x-a)^m} = \frac{P(x)}{D(x)} + \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_m}{(x-a)^m}$$

Proof:

To save writing, let $z = x - a$. Then $\tilde{N}(z) = N(z+a)$ and $\tilde{D}(z) = D(z+a)$ are polynomials of degree n and d respectively, $\tilde{D}(0) = D(a) \neq 0$ and we have to find a polynomial $\tilde{P}(z)$ of degree $p < d$ and numbers $A_1, \ldots, A_m$ such that

$$\frac{\tilde{N}(z)}{\tilde{D}(z)z^m} = \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A_1}{z} + \frac{A_2}{z^2} + \cdots + \frac{A_m}{z^m}$$

$$= \frac{\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_m\tilde{D}(z)}{\tilde{D}(z)z^m}$$

or equivalently, such that

$$\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_{m-1}z\tilde{D}(z) + A_m\tilde{D}(z) = \tilde{N}(z)$$

Now look at the polynomial on the left-hand side. Every single term on the left-hand side, except for the very last one, $A_m\tilde{D}(z)$, has at least one power of z. So the constant term on the left-hand side is exactly the constant term in $A_m\tilde{D}(z)$, which is $A_m\tilde{D}(0)$. The constant term on the right-hand side is $\tilde{N}(0)$. So the constant terms on the left and right-hand sides are the same if we choose $A_m = \frac{\tilde{N}(0)}{\tilde{D}(0)}$. Recall that $\tilde{D}(0) \neq 0$. Now move $A_m\tilde{D}(z)$ to the right-hand side.

$$\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_{m-1}z\tilde{D}(z) = \tilde{N}(z) - A_m\tilde{D}(z)$$

The constant terms in $\tilde{N}(z)$ and $A_m\tilde{D}(z)$ are the same, so the right-hand side contains no constant term and the right-hand side is of the form $\tilde{N}_1(z)z$ with $\tilde{N}_1$ a polynomial of degree at most $d + m - 2$. (Recall that $\tilde{N}$ is of degree at most $d + m - 1$ and $\tilde{D}$ is of degree at most d.) Divide the whole equation by z.

$$\tilde{P}(z)z^{m-1} + A_1 z^{m-2}\tilde{D}(z) + A_2 z^{m-3}\tilde{D}(z) + \cdots + A_{m-1}\tilde{D}(z) = \tilde{N}_1(z)$$

Now, we can repeat the previous argument. The constant term on the left-hand side, which is exactly $A_{m-1}\tilde{D}(0)$ matches the constant term on the right-hand side, which is $\tilde{N}_1(0)$ if we choose $A_{m-1} = \frac{\tilde{N}_1(0)}{\tilde{D}(0)}$. With this choice of A_{m-1}

$$\tilde{P}(z)z^{m-1} + A_1 z^{m-2}\tilde{D}(z) + A_2 z^{m-3}\tilde{D}(z) + \cdots + A_{m-2}z\tilde{D}(z) = \tilde{N}_1(z) - A_{m-1}\tilde{D}(z) = \tilde{N}_2(z)z$$

With $\tilde{N}_2$ a polynomial of degree at most $d + m - 3$. Divide by z and continue. After m steps like this, we end up with

$$\tilde{P}(z)z = \tilde{N}_{m-1}(z) - A_1\tilde{D}(z)$$

After having chosen $A_1 = \frac{\tilde{N}_{m-1}(0)}{\tilde{D}(0)}$. There is no constant term on the right side so that $\tilde{N}_{m-1}(z) - A_1\tilde{D}(z)$ is of the form $\tilde{N}_m(z)z$ with $\tilde{N}_m$ a polynomial of degree $d - 1$. Choosing $\tilde{P}(z) = \tilde{N}_m(z)$ completes the proof.

Now back to

$$\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}}$$

Apply Lemma 2, with $D(x) = (x - a_2)^{n_2} \times \cdots \times (x - a_d)^{n_d}$, $m = n_1$ and $a = a_1$. It says

$$\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}} = \frac{P(x)}{(x - a_2)^{n_2} \times \cdots \times (x - a_d)^{n_d}} + \frac{A_{1,1}}{x - a_1} + \frac{A_{1,2}}{(x - a_1)^2} + \cdots + \frac{A_{1,n_1}}{(x - a_1)^{n_1}}$$

Apply Lemma 2 a second time, with $D(x) = (x - a_3)^{n_3} \times \cdots \times (x - a_d)^{n_d}$, $N(x) = P(x)$, $m = n_2$ and $a = a_2$. And so on. Eventually, we end up with

$$\frac{N(x)}{(x - a_1)^{n_1} \times \cdots \times (x - a_d)^{n_d}} = \left[\frac{A_{1,1}}{x - a_1} + \cdots + \frac{A_{1,n_1}}{(x - a_1)^{n_1}} \right] + \cdots + \left[\frac{A_{d,1}}{x - a_d} + \cdots + \frac{A_{d,n_d}}{(x - a_d)^{n_d}} \right]$$

18 Recursion

Recursion is term used to refer to a thing define in terms of itself. An example the factorial $n! = (n-1)!$ with $0! = 1$ being the base case where the recursion stops.

18.1 From Recursive to Closed Form

A sequence defined recursively can often be rewritten in an explicit or closed form.

Recursive form defines a_n based on previous terms (e.g., $a_n = a_{n-1} + 2$).

Closed form expresses a_n directly in terms of n, with no reference to previous terms.

18.1.1 General Method (when linear):

1. Calculate some terms and watch the difference or quotient between them maybe the difference is clear, and you can build a geometric or arithmetic sequence directly.

2. Identify the type (e.g., linear, homogeneous, constant coefficients).

3. Solve the characteristic equation (if linear).

4. Use initial conditions to determine constants.

18.2 Linear Recursion

A linear recurrence has the form:

$$a_n = \lambda a_{n-1} + c$$
$$a_1 = \lambda a_0 + c \implies a_2 = \lambda(\lambda a_0 + c) + c$$
$$\implies a_n = \lambda^n a_0 + (1 + \lambda + \lambda^2 + \cdots + \lambda^{n-1})c$$

Therefore,

$$a_n = \lambda^n a_0 + \frac{\lambda^n - 1}{\lambda - 1}c$$

18.3 Recursive Inequalities

A *recursive inequality* bounds a sequence rather than defines it exactly.

Example:

Suppose

$$a_{n+1} \leq \frac{1}{2}a_n + 3$$

To analyze:

- Guess a bound (e.g., show $a_n \leq M$ for all n)
- Use induction or iteration to justify convergence or boundedness
- Compare with a simpler sequence (e.g., geometric series)

These techniques are common in analysis, approximation, and numerical algorithms.

18.4 Fibonacci Numbers and Closed Form

The Fibonacci sequence is defined recursively:

$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2}$$

The characteristic equation is:

$$x^2 - x - 1 = 0 \Rightarrow x = \varphi = \frac{1 + \sqrt{5}}{2}, \quad \psi = \frac{1 - \sqrt{5}}{2}$$

The closed form (Binet's Formula) is:

$$F_n = \frac{1}{\sqrt{5}} \left(\varphi^n - \psi^n \right)$$

- $\varphi \approx 1.618$ is the golden ratio
- This formula demonstrates exponential growth of F_n

19 Sum and Product Notation

In mathematics, the sum and product notations are compact ways to represent repeated addition and multiplication, respectively. These notations are essential for working with sequences, series, and algebraic expressions.

19.1 Sum Notation $\sum$

The summation symbol $\sum$ represents the addition of a sequence of terms:

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + \cdots + a_n$$

Where i is the index of summation, m is the lower bound, and n is the upper bound.

Linearity:

$$\sum_{i=m}^{n} (a_i + b_i) = \sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i$$

$$\sum_{i=m}^{n} c \cdot a_i = c \cdot \sum_{i=m}^{n} a_i$$

Splitting:

$$\sum_{i=m}^{n} a_i = \sum_{i=m}^{k} a_i + \sum_{i=k+1}^{n} a_i \quad (m \le k < n)$$

19.1.1 Change of Index

Let $j = i + k$, then:

$$\sum_{i=m}^{n} a_i = \sum_{j=m+k}^{n+k} a_{j-k}$$

Example:

$$\sum_{i=1}^{4} a_i = \sum_{j=2}^{5} a_{j-1}$$

19.1.2 Power Sums and Their Formulas

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$
$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$
$$\sum_{i=1}^{n} i^3 = \left[\frac{n(n+1)}{2} \right]^2$$
$$\sum_{i=1}^{n} i^k = \text{(Higher-order polynomial in } n)$$

19.1.3 Derivation of $\sum_{i=1}^{n} i^2$

We use the method of finite differences or induction. Assume a quadratic form:

$$\sum_{i=1}^{n} i^2 = An^3 + Bn^2 + Cn$$

Plug in small values of n (e.g., 1, 2, 3), solve the system of equations to find:

$$A = \frac{1}{3}, \quad B = \frac{1}{2}, \quad C = \frac{1}{6} \Rightarrow \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

19.1.4 Telescoping Sum

A telescoping sum is a sum where intermediate terms cancel out, leaving only the first and last terms.

$$\sum_{k=m}^{n} (a_k - a_{k+1}) = a_m - a_{n+1}$$

Example:

$$\sum_{i=1}^{n} \left(\frac{1}{i} - \frac{1}{i+1} \right) = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n} - \frac{1}{n+1} = 1 - \frac{1}{n+1}$$

19.2 Geometric Series

19.2.1 Finite Geometric Series:

For a geometric sequence $a, ar, ar^2, \ldots, ar^{n-1}$:

$$\sum_{i=0}^{n-1} ar^i = a \cdot \frac{1 - r^n}{1 - r}, \quad r \neq 1$$

19.2.2 Infinite Geometric Series:

If $|r| < 1$, then:

$$\sum_{i=0}^{\infty} ar^i = \frac{a}{1 - r}$$

19.3 Product Notation $\prod$

The product notation $\prod$ represents repeated multiplication:

$$\prod_{i=m}^{n} a_i = a_m \cdot a_{m+1} \cdot \cdots \cdot a_n$$

Multiplication:

$$\prod_{i=m}^{n} (a_i \cdot b_i) = \left(\prod_{i=m}^{n} a_i \right) \cdot \left(\prod_{i=m}^{n} b_i \right)$$

Power Rule:

$$\prod_{i=m}^{n} a^k = a^{k(n-m+1)}$$

19.3.1   Change of Index

Let $j = i + k$, then:

$$\prod_{i=m}^{n} a_i = \prod_{j=m+k}^{n+k} a_{j-k}$$

Example:

$$\prod_{i=1}^{3} a_i = \prod_{j=2}^{4} a_{j-1}$$

19.3.2   Telescoping Product

A telescoping product occurs when consecutive terms simplify or cancel.

$$\prod_{k=m}^{n} \frac{a_k}{a_{k+1}} = \frac{a_m}{a_{n+1}}$$

Example:

$$\prod_{i=1}^{n} \frac{i}{i+1} = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{n}{n+1} = \frac{1}{n+1}$$

19.4   Proof of the geometric Series

Let $S_n = \sum_{k=0}^{n} aq^k$ with $|q| < 1$.

Then $S_n q = a(\sum_{k=0}^{n+1})q^k = a\left(\sum_{k=0}^{n} q^k + q^{n+1}\right)$

$$a + S_n q = \sum_{k=0}^{n} aq^k + aq^{n+1} = S_n + aq^{n+1}$$

Finally, we subtract S_n and a and after that divide by $q - 1$

$$S_n = \frac{a - aq^{n+1}}{q - 1}$$

$$\mathfrak{QED}$$

20   Logarithms

The *Logarithm* of a number x with base b is the exponent to which the base must be raised to produce that number. It is denoted as:

$$\log_b(x) = y \iff b^y = x$$

Where $b > 0$, $b \neq 1$, and $x > 0$.

$$\log_2(8) = 3 \quad \text{because} \quad 2^3 = 8$$

We can also approximate a logarithm by dividing our target x by the basis b repeatedly until we get 1 or less.

20.1   Properties of Logarithms

- $\log_b(xy) = \log_b(x) + \log_b(y)$
- $\log_b\left(\frac{x}{y}\right) = \log_b(x) - \log_b(y)$
- $\log_b(x^k) = k \cdot \log_b(x)$
- $\log_b(b) = 1$
- $\log_b(1) = 0$
- $\log_{b^k}(x^w) = \frac{1}{k} \cdot \log_b(x)$
- $\log_b\left(\frac{1}{x}\right) = -\log_b(x)$
- $\log_b(b^x) = x$
- $\log_b(x) = \frac{\log_k(x)}{\log_k(b)}$ for any positive $k \neq 1$
- $e^{\ln(x)} = x$
- If $0 < a < 1$ then $\ln(a)$ is a negative number.

20.2   Fundamental Identity of Logarithms

$$a^{\log_a(x)} = x$$

20.3   Change of Basis Formula

$$\log_b(x) = \frac{\log_k(x)}{\log_k(b)}$$

Where k is any positive number different from 1.

20.4   The Chain Rule

$$\log_y(a) \log_a(b) = \log_y(b)$$

20.5   The derivative of the Natural Logarithm

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

Let $f(x) = \ln(x)$, then

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

$$= f'(x) = \lim_{h \to 0} \frac{\ln\left(\frac{x+h}{x}\right)}{h}$$

$$= f'(x) = \lim_{h \to 0} \frac{\ln\left(1 + \frac{h}{x}\right)}{h}$$

$$= f'(x) = \lim_{h \to 0} \ln\left(1 + \frac{h}{x}\right)^{\frac{1}{h}}$$

Now let $n = \frac{h}{x}$

$$f'(x) = \lim_{n \to 0} \ln(1 + n)^{\frac{x}{h}\frac{1}{x}}$$

$$f'(x) = \lim_{n \to 0} \frac{1}{x} \ln(1 + n)^{\frac{1}{n}}$$

$$f'(x) = \frac{1}{x} \ln\left(\lim_{n \to 0}(1 + n)^{\frac{1}{n}}\right)$$

$$f'(x) = \frac{1}{x}$$

Therefore, the derivative of the natural logarithm is:

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

$$\mathfrak{QED}$$

20.6 Demonstration of the Properties

20.6.1 Product

Let $a^n a^m = a^{n+m}$ and $a^n = x$ and $a^m = y$ then

$$\log_a x = n \text{ and } \log_a y = m \to \log_a(xy) = n + m$$

$$\log_a xy = \log_a x + \log_a y$$

$$\mathfrak{QED}$$

20.6.2 Quotient

Let $\frac{a^n}{a^m} = a^{n-m}$ and $a^n = x$ and $a^m = y$ then like in the previous proof

$$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y$$

$$\mathfrak{QED}$$

20.6.3 Power

Let $(a^n)^m = a^{nm}$ and $a^n = x$ and $a^{nm} = y$

$$\log_x y = m \text{ and } \log_a x = n \text{ thus, } \log_a y = (\log_a x)m$$

$$\mathfrak{QED}$$

20.6.4 Power Rule General Case:

Let $(a^n)^m = b^x$ then $\log_{a^n} b^x = m$ now let us manipulate the original expression

$$\log_a (a^n)^m = \log_a b^x$$
$$nm \log_a a = x \log_a b$$
$$m = \frac{x}{n} \log_a b = \log_{a^n} b^x$$

$$\mathfrak{QED}$$

21 Fractions, Roots, and Exponents

This is a small chapter to remember the properties of fractions, roots, and exponents. It is not a complex chapter, but it is useful to have it in the compendium.

21.1 Fractions

- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$

- $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$

- $\frac{a}{b} \div \frac{c}{d} = \frac{ad}{bc}$

- $\frac{a}{b} = \frac{c}{d} \iff ad = bc$

- $\frac{ac}{bc} = \frac{a}{b}$

21.2 Roots and Exponents

- $\sqrt{a} \cdot \sqrt{b} = \sqrt{ab}$

- $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$

- $\sqrt[nm]{a} = \sqrt[n]{\sqrt[m]{a}}$

- $a \sqrt[n]{b} = \sqrt[n]{a^n b}$

- $a^m \cdot a^n = a^{m+n}$

- $a^m \div a^n = a^{m-n}$

- $(a^m)^n = a^{mn}$

- $a^{-n} = \frac{1}{a^n}$

- $a^{\frac{m}{n}} = \sqrt[n]{a^m}$

- $a^{\frac{1}{2}} = \sqrt{a}$

- $a^{\frac{1}{3}} = \sqrt[3]{a}$

- $a^{\frac{1}{n}} = \sqrt[n]{a}$

- $a^0 = 1$

- $a^1 = a$

22 The Fundamental Theorem of Algebra

Given any positive integer $n \geq 1$ and any choice of complex numbers $a_0, a_1, \ldots, a_n$, such that $a_n \neq 0$, the polynomial equation

$$a_n z^n + \cdots + a_1 z + a_0 = 0$$

has at least one solution $z \in \mathbb{C}$.

22.1 Gist of the Proof

For readers familiar with Newton's method for solving equations, one starts with a reasonably close approximation to a root, then adjusts the approximation by moving closer in an appropriate direction. We will employ the same strategy here, showing that if one assumes that the argument where the polynomial function achieves its minimum absolute value is not a root, then there is a nearby argument where the polynomial function has an even smaller absolute value, contradicting the assumption that the argument of the minimum absolute value is not a root.

22.2 Definitions and Axioms

In the following, $p(z)$ will denote the n-th degree polynomial

$$p(z) = p_0 + p_1 z + p_2 z^2 + \cdots + p_n z^n,$$

where the coefficients p_i are any complex numbers, with neither p_0 nor p_n equal to zero (otherwise the polynomial is equivalent to one of lesser degree).

We will utilize a fundamental completeness property of real and complex numbers, namely that a continuous function on a closed set achieves its minimum at some point in the domain. This can be taken as an axiom, or can be easily proved by applying other well-known completeness axioms, such as the Cauchy sequence axiom or the nested interval axiom.

22.3 Theorem I

Every polynomial with real or complex coefficients has at least one complex root.

Proof:

Suppose that $p(z)$ has no roots in the complex plane. First note that for large z, say $|z| > 2 \max_i |p_i/p_n|$, the z^n term of $p(z)$ is greater in absolute value than the sum of all the other terms. Thus, given some $B > 0$, for any sufficiently large s, we have $|p(z)| > B$ for all z with $|z| \geq s$. We will take $B = 2|p(0)| = 2|p_0|$.

Since $|p(z)|$ is continuous on the interior and boundary of the circle with radius s, it follows by the completeness axiom that $|p(z)|$ achieves its minimum value at some point t in this circle. But since $|p(0)| < \frac{1}{2}|p(z)|$ for all z on the circumference of the circle, it follows that $|p(z)|$ achieves its minimum at some point t in the interior.

Now rewrite the polynomial $p(z)$ by translating the argument z by t, thus, producing a new polynomial

$$q(z) = p(z + t) = q_0 + q_1 z + q_2 z^2 + \cdots + q_n z^n,$$

and similarly translate the circle. Presumably the polynomial $q(z)$, defined on some circle centered at the origin, has a minimum absolute value $M > 0$ at $z = 0$. Note that $M = |q(0)| = |q_0|$.

Our proof strategy is to construct some point x, close to the origin, such that $|q(x)| < |q(0)|$, thus, contradicting the assumption that $|q(z)|$ has a minimum nonzero value at $z = 0$.

Construction of x such that $|q(x)| < |q(0)|$

Let the first nonzero coefficient of $q(z)$ following q_0 be q_m, so that

$$q(z) = q_0 + q_m z^m + q_{m+1} z^{m+1} + \cdots + q_n z^n.$$

We choose

$$x = r\left(-\frac{q_0}{q_m}\right)^{1/m},$$

where r is a small positive real value, and $\left(-\frac{q_0}{q_m}\right)^{1/m}$ denotes any m-th root of $\left(-\frac{q_0}{q_m}\right)$.

Unlike the real numbers, in the complex number system the m-th roots of any complex number are guaranteed to exist. If $z = z_1 + iz_2$, then the m-th roots of z are given by

$$\left\{ R^{1/m} \cos\left(\frac{\theta + 2k\pi}{m}\right) + iR^{1/m} \sin\left(\frac{\theta + 2k\pi}{m}\right) \, \middle| \, k = 0, 1, \ldots, m-1 \right\},$$

where $R = \sqrt{z_1^2 + z_2^2}$ and $\theta = \arctan(z_2/z_1)$.

Proof:

$$|q(x)| < |q(0)|$$

With the definition of x, we can write

$$q(x) = q_0 - q_0 r^m + q_{m+1} r^{m+1}\left(-\frac{q_0}{q_m}\right)^{(m+1)/m} + \cdots + q_n r^n \left(-\frac{q_0}{q_m}\right)^{n/m} = q_0 - q_0 r^m + E$$

Where the extra terms E can be bounded as follows. Assume $q_0 \le q_m$, and define $s = r\left|\frac{q_0}{q_m}\right|^{1/m}$. Then

$$|E| \le r^{m+1} \max_i |q_i| \left|\frac{q_0}{q_m}\right|^{(m+1)/m} (1 + s + s^2 + \cdots + s^{n-m-1}) \le \frac{r^{m+1} \max_i |q_i|}{1-s} \left|\frac{q_0}{q_m}\right|^{(m+1)/m}.$$

Thus, $|E|$ can be made arbitrarily small compared to $|q_0 r^m| = |q_0| r^m$ by choosing r small enough. For example, select r so that $|E| < \frac{|q_0| r^m}{2}$. Then:

$$|q(x)| = |q_0 - q_0 r^m + E| < |q_0 - \frac{q_0 r^m}{2}| = |q_0|\left(1 - \frac{r^m}{2}\right) < |q_0| = |q(0)|,$$

which contradicts the assumption that $|q(z)|$ has a minimum nonzero value at $z = 0$.

<div align="right">Q𝔈𝔇</div>

22.4 Theorem II

Every polynomial of degree n with real or complex coefficients has exactly n complex roots, when counting multiplicities.

Proof:

If α is a root of the polynomial $p(z)$ of degree n, then by dividing $p(z)$ by $(z - \alpha)$, we get:

$$p(z) = (z - \alpha)q(z) + r,$$

where $q(z)$ has degree $n - 1$ and r is a constant. But since $p(\alpha) = r = 0$, we conclude:

$$p(z) = (z - \alpha)q(z).$$

Continuing by induction, we conclude that the original polynomial $p(z)$ has exactly n complex roots, counted with multiplicities.

$\mathfrak{QED}$

23 Generating Functions

Generating functions are a central tool in combinatorics and discrete mathematics. At their core, a generating function is a formal power series in one or more variables, typically used to encode sequences of numbers.

23.1 Definition

Given a sequence $\{a_n\}_{n \geq 0}$, its (ordinary) generating function is defined as:

$$G(x) = \sum_{n=0}^{\infty} a_n x^n$$

Here, x is an indeterminate, and the series is usually treated formally, meaning we do not worry about convergence unless the context requires it.

23.2 Applications

Generating functions are used to:

- Encode and manipulate sequences algebraically.
- Solve recurrence relations.
- Find closed-form expressions for sequences.
- Prove combinatorial identities.
- Model and solve counting problems in partitions, compositions, and permutations.

For example, if a_n counts the number of ways to form a sum n using coins of certain denominations, the generating function helps derive a formula for a_n or analyze its properties.

23.3 Connection to Subsets

There is a deep similarity between generating functions and the process of forming subsets. Consider the generating function for the set $\{1, x\}$:

$$(1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$

This expansion encodes all subsets of an n-element set, where each term x^k corresponds to selecting a subset of size k. The coefficient $\binom{n}{k}$ gives the number of such subsets.

In this way, generating functions can be viewed as a powerful algebraic analogue of subset selection and counting, with the exponents representing sizes or weights, and the coefficients representing counts.

Generating functions serve as a bridge between algebra and combinatorics, allowing abstract operations to yield concrete combinatorial results.

23.4 Closed form of the Fibonacci Sequence

Let us use generating function to find a closed form for the Fibonacci number.

We know that $F_0 = 1$ and $F_1 = 1$ and for $n \geq 2$ $F_n = F_{n-1} + F_{n-2}$

Now let us use the following function

$$F(x) = \sum_{n=0}^{\infty} F_n x^n$$

Which is going to have the Fibonacci numbers as the coefficients of the power series.

$$1 + x + 2x^2 + 3x^3 + 5x^4 + \ldots$$

But before that we have to look at when our recursive formula is defined. In this case $n = 2$. Thus, we have to write our expression in the following way.

$$F(x) = 1 + x + \sum_{n=2}^{\infty} F_n x^n$$

$$= 1 + x + \sum_{n=2}^{\infty} F_{n-1} x^n + \sum_{n=2}^{\infty} F_{n-2} x^n$$

Here we are going to adjust our expression by making our n be equal.

$$F(x) = 1 + x + \sum_{n=2}^{\infty} F_{n-1} x^n + \sum_{n=2}^{\infty} F_{n-2} x^n$$

$$F(x) = 1 + x + x \left(\sum_{n=2}^{\infty} F_{n-1} x^{n-1} \right) + x^2 \left(\sum_{n=2}^{\infty} F_{n-2} x^{n-2} \right)$$

And now we are going to reset our sum to start from zero by adding a smart zero to our term with $n-1$. This allows us to shift the sum to start from 0 because when $n = 1$ for $F_{n-1} x^{n-1}$ turns into $F_0 x^0$.

$$F(x) = 1 + x + x \left(-F_0 x^0 + F_0 x^0 + \sum_{n=2}^{\infty} F_{n-1} x^{n-1} \right) + x^2 \left(\sum_{n=2}^{\infty} F_n x^n \right)$$

$$F(x) = 1 + x + x \left(-F_0 x^0 + \sum_{n=1}^{\infty} F_{n-1} x^{n-1} \right) + x^2 \left(\sum_{n=2}^{\infty} F_n x^n \right)$$

$$F(x) = 1 + x + x \left(\sum_{n=0}^{\infty} F_n x^n \right) + x^2 \left(\sum_{n=2}^{\infty} F_n x^n \right)$$

$$F(x) = 1 + x + x F(x) - x + x^2 F(x)$$

Which at end simplifies to:

$$F(x) = \frac{1}{1 - x - x^2}$$

If you start with $F_0 = 0$ then the result is

$$F(x) = \frac{x}{1 - x - x^2}$$

Now we can use partial fraction decomposition to make this expression even more explicit by using the following expression where α and β are zeroes of the polynomial in the denominator.

$$\frac{1}{(1 - \alpha x)(1 - \beta x)}$$

The zeros of the polynomial are:

$$x_1 = \frac{1 + \sqrt{5}}{2} \text{ and } x_2 = \frac{1 - \sqrt{5}}{2}$$

Now we can write our expression like:

$$F(x) = \frac{1}{1 - x - x^2} = \frac{1}{\left(1 - x \cdot \frac{1+\sqrt{5}}{2}\right)\left(1 - x \cdot \frac{1-\sqrt{5}}{2}\right)}$$

Which after performing the partial fraction decomposition left us with:

$$F(x) = \left(\left(\frac{1}{\sqrt{5}}(1 + \sqrt{5})\right)^n - \left(\frac{1}{\sqrt{5}}(1 - \sqrt{5})\right)^n\right)$$

$\mathfrak{QED}$

24 Modular Arithmetic

We define an equivalence relation

$$a \cong b \mod n \iff n|(a - b) \wedge a, b \in \mathbb{Z}$$

Means that both have the same residual when divided by n.

24.1 Residual-class and Residual-class Rings

Because we have an equivalence relation we can classify the partitions of our set $\mathbb{Z}$ into classes which we will call *Residual-classes*.

Example:

$$[13]_7 + \{m \in \mathbb{Z} \mid m \cong 13 \mod 7\}$$

The set of all Residual-classes can be written as $\mathbb{Z}/n\mathbb{Z}$ where n is the modulo.

24.2 Properties

Addition

$$(a + b) \mod n = (a \mod n + b \mod n) \mod n$$

Multiplication

$$(a * b) \mod n = (a \mod n * b \mod n) \mod n$$

Exponentiation

$$(a^m) \mod n = ((a \mod n)^n) \mod n$$

Division

$$\left(\frac{a}{b}\right) \mod n \neq \frac{a \mod n}{b \mod n}$$

24.3 General Division Rules

24.3.1 Sum Rules

If the number a can be expressed as a sum and b is a common divisor of all the terms, then b is also a divisor of a.

Example:

$742 = 700 + 42$. Then 7, as a common divisor of 700 and 42, is a divisor of 742.

24.3.2 Product Rule

If the number a can be written as a product of two factors and b is a divisor of one of these factors, then b is also a divisor of a.

Example:

$1500 = 15 \cdot 100$. Then 3 and 5 are divisors of 15, and 2, 4, 5, 10, 25, and 50 are divisors of 100—thus, all are divisors of 1500.

24.3.3 Divisor Pair Rules

If the number a has b as a divisor, then $\frac{a}{b}$ is also a divisor of a.

Example:

3 is a divisor of 315. Then $105 = \frac{315}{3}$ is also a divisor of 315.

24.3.4 Divisor Rule

If the number a has b as a divisor, then every divisor of b is also a divisor of a.

Example:

45 is a divisor of 450. Then 3, 5, 9, and 15 (divisors of 45) are also divisors of 450.

24.3.5 Difference Rule

If the number a can be expressed as a difference and b is a common divisor of the minuend and subtrahend, then b is also a divisor of a.

Example:

$686 = 700 - 14$. Then 7, as a common divisor of 700 and 14, is a divisor of 686.

24.3.6 Relative Prime Rule

If the number a has two relatively prime numbers b and c as divisors, then their product $b \cdot c$ is also a divisor of a.

Example:

1540 has the relatively prime numbers 7 and 11 as divisors. Then $77 = 7 \cdot 11$ is also a divisor of 1540.

24.4 Using Modular Arithmetic to find the n digit of a big number

What is the last digit of the number 13^{2023}

Note that $13 \mod 10 = 3$ and $13^2 \mod 10 = ((13 \mod 10)(13 \mod 10)) \mod 10 = 9$

We can continue this process, and we will notice a pattern. After some iteration we will now that

$$(13^{2020}13) \mod 10 = 7$$

24.5 Gaussian Brackets

The Gaussian brackets refer to two closely related functions used to round real numbers to integers in specific ways:

Floor Function

The floor function, denoted by $\lfloor x \rfloor$, gives the greatest integer less than or equal to a real number x.

Example:

$\lfloor 3.7 \rfloor = 3$, $\lfloor -1.2 \rfloor = -2$

Ceiling Function

The ceiling function, denoted by $\lceil x \rceil$, gives the smallest integer greater than or equal to a real number x.

Example:

$\lceil 3.7 \rceil = 4$, $\lceil -1.2 \rceil = -1$

These functions are particularly useful: in number theory, computer science and rounding operations, where precise integer bounds of real numbers are required.

25 Geometry

25.1 Formulas for Area and Perimeter of Geometric Figures

- **Square**
  - Area: $A = a^2$
  - Perimeter: $P = 4a$
- **Rectangle**
  - Area: $A = l \cdot w$
  - Perimeter: $P = 2(l + w)$
- **Circle**
  - Area: $A = \pi r^2$
  - Circumference: $C = 2\pi r$
- **Triangle (General)**
  - Area: $A = \frac{1}{2} b \cdot h$
  - Perimeter: $P = a + b + c$
- **Equilateral Triangle**
  - Area: $A = \frac{\sqrt{3}}{4} a^2$
  - Perimeter: $P = 3a$
- **Isosceles Triangle**
  - Area: $A = \frac{b}{4} \sqrt{4a^2 - b^2}$
- **Trapezoid**
  - Area: $A = \frac{1}{2}(a + b)h$
  - Perimeter: $P = a + b + c + d$
- **Parallelogram**
  - Area: $A = b \cdot h$
  - Perimeter: $P = 2(a + b)$

25.2 Formulas for Area and Volume of 3D Geometric Figures

- **Cube**
  - Surface Area: $A = 6a^2$
  - Volume: $V = a^3$
- **Cylinder**
  - Surface Area: $A = 2\pi r(h + r)$
  - Volume: $V = \pi r^2 h$
- **Cone**
  - Surface Area: $A = \pi r(r + l)$ with $l =$ slant height
  - Volume: $V = \frac{1}{3} \pi r^2 h$
- **Sphere**
  - Surface Area: $A = 4\pi r^2$
  - Volume: $V = \frac{4}{3} \pi r^3$

- **Square Pyramid**
  - Surface Area: $A = a^2 + 2a \cdot l$
  - Volume: $V = \frac{1}{3} a^2 h$
- **Triangular Pyramid (Tetrahedron)**
  - Volume: $V = \frac{1}{3} A_b \cdot h$
- **Prism (General)**
  - Surface Area: $A = 2A_b + P_b \cdot h$
  - Volume: $V = A_b \cdot h$

25.3 Thales' Theorem

Thales' Theorem states: If A, B, and c are points on a circle where the line AB is the diameter, then the angle $\angle \alpha + \beta$ is a right angle.

$$\angle \alpha + \beta = 90° \quad \text{if } AB \text{ is a diameter of the circle}$$
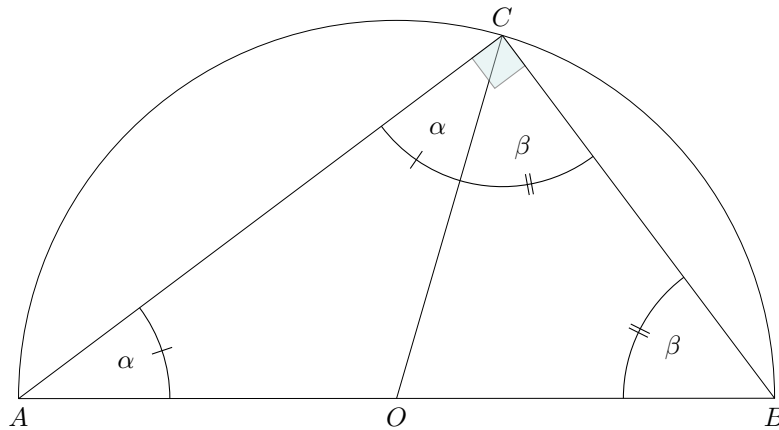
Proof:

Note that we have two Isosceles triangles where the sides AO, AC and OC have all the same length, being this the radius of the circle.

Because they are Isosceles then two of their angles have to be equal and the sum of the angles of this two triangle has to add up to the be $180°$. Therefore,

$$\alpha + \alpha + \beta + \beta = 180$$
$$2\alpha + 2\beta = 180$$
$$\alpha + \beta = 90$$

$$\mathfrak{QED}$$

25.4 Conversion Between Radians and Degrees

To convert between radians and degrees, use the following formulas:

$$\text{Degrees} = \text{Radians} \times \left(\frac{180°}{\pi} \right)$$
$$\text{Radians} = \text{Degrees} \times \left(\frac{\pi}{180°} \right)$$

25.5 Geometric Figures with Formulas (2D)

Square and Rectangle

$A = a^2 \quad P = 4a$

$A = l \cdot w \quad P = 2(l + w)$

Square

a
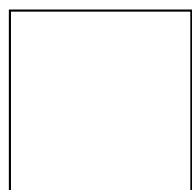
w
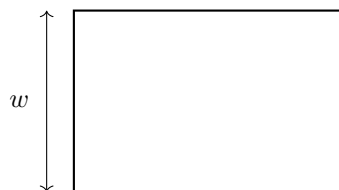
Rectangle

l

Circle and Triangle Types

r
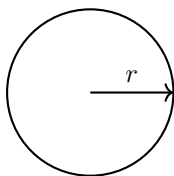
Circle

$A = \pi r^2 \quad C = 2\pi r$
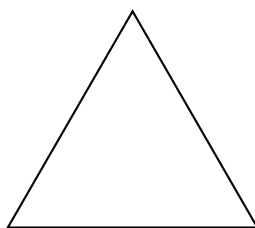
Equilateral

$A = \frac{\sqrt{3}}{4}a^2$
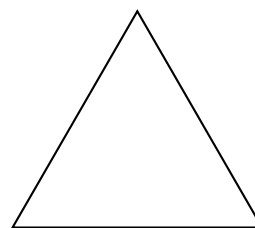
Isosceles

$A = \frac{b}{4}\sqrt{4a^2 - b^2}$

Scalene Triangle, Trapezoid, Parallelogram
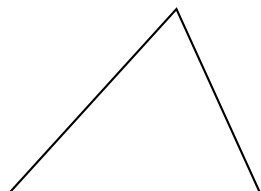
Scalene Triangle

$A = \frac{1}{2}b \cdot h$

Trapezoid

$A = \frac{1}{2}(a + b)h$
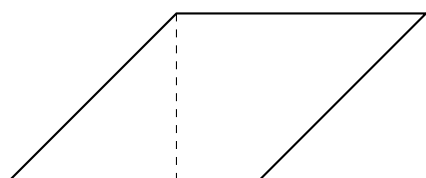
Parallelogram

$A = b \cdot h \quad P = 2(a + b)$

25.6 3D Geometric Figures with Formulas

25.6.1 Cube and Cylinder
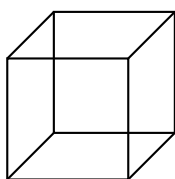
Cube
$$A = 6a^2 \quad V = a^3$$
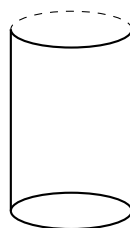
Cylinder
$$A = 2\pi r(h + r)$$
$$V = \pi r^2 h$$

25.6.2 Cone and Sphere
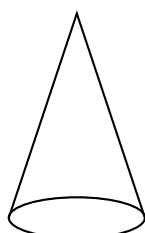
Cone
$$A = \pi r(r + l)$$
$$V = \frac{1}{3}\pi r^2 h$$
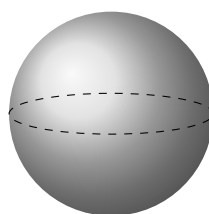
Sphere
$$A = 4\pi r^2 \quad V = \frac{4}{3}\pi r^3$$

25.6.3 Pyramid and Prism
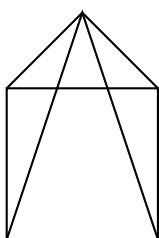
Square Pyramid
$$A = a^2 + 2al$$
$$V = \frac{1}{3}a^2 h$$
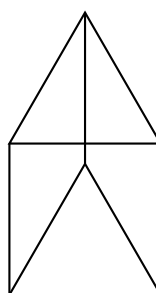
Triangular Prism
$$A = 2A_b + P_b h$$
$$V = A_b \cdot h$$

25.7 Axioms of Euclidean Geometry

Euclidean geometry is founded on a set of basic assumptions known as axioms or postulates, originally formulated by the Greek mathematician Euclid in his work *Elements*. These axioms form the logical basis for plane geometry and include the following:

I. *Line Postulate:* A straight line segment can be drawn joining any two points.

II. *Extension Postulate:* Any straight line segment can be extended indefinitely in a straight line.

III. *Circle Postulate:* Given any straight line segment, a circle can be drawn having the segment as radius and one endpoint as center.

IV. *Right Angle Postulate:* All right angles are congruent.

V. *Parallel Postulate:* If a line segment intersects two straight lines and makes the interior angles on the same side less than two right angles, then the two lines, if extended indefinitely, meet on that side.

These five postulates define the framework of classical geometry in a flat (Euclidean) space. Notably, the fifth postulate—the parallel postulate—has been the subject of much scrutiny and led to the development of non-Euclidean geometries when it was altered or replaced.

25.8   Types of Angles and Angle Relationships

An angle is the distance between two rays or intercepting lines.

- If angle is less than 90° it is an *acute* angle

- If it is exactly 90° it is a *right* angle

- It is more than 90° but less than 180° it is called an *obtuse* angle

Now let us look at other types of angles

25.8.1   Vertical Angles

When two lines intersect, they form two pairs of opposite (or vertical) angles. Vertical angles are always congruent. That is, if two angles are vertical, their measures are equal:

$$\angle A = \angle B$$

For example, if two lines intersect and form angles of 60° and 120°, the angles opposite each other are both 60° and both 120°.

25.8.2   Adjacent Angles

Adjacent angles are two angles that share a common side and a common vertex, and do not overlap. When two lines intersect, each pair of adjacent angles forms a straight angle (or a linear pair), summing to 180°:

$$\angle 1 + \angle 2 = 180°$$

This is known as the Linear Pair Postulate.

25.8.3   Complementary Angles

Complementary angles are two angles whose measures add up to 90°:

$$\angle X + \angle Y = 90°$$

Complementary angles do not have to be adjacent, but if two intersecting lines form a right angle, the two acute angles adjacent to the right angle may be complementary.

25.8.4   Angles Generated by Traversals

When a transversal crosses two or more lines, it forms several pairs of angles with specific relationships. Some key types include:

- *Corresponding Angles:* Angles in the same relative position at each intersection. If the lines are parallel, corresponding angles are congruent:

$$\angle 1 \cong \angle 2$$

- *Alternate Interior Angles:* Angles that lie between the two lines but on opposite sides of the transversal. They are congruent if the lines are parallel:

$$\angle 3 \cong \angle 4$$

- *Alternate Exterior Angles:* Angles outside the two lines and on opposite sides of the transversal. Also, congruent when the lines are parallel:

$$\angle 5 \cong \angle 6$$

- *Consecutive (Same-Side) Interior Angles:* Angles on the same side of the transversal and inside the parallel lines. These are supplementary:

$$\angle 7 + \angle 8 = 180°$$

25.9 Triangle Similarity and Congruence

Triangles can be compared using criteria of similarity or congruence, depending on whether their shapes are the same or both shape and size are identical.

25.9.1 Triangle Congruence

Two triangles are *congruent* if all corresponding sides and angles are equal. This means the triangles are identical in shape and size. Common criteria for triangle congruence include:

- *SSS (Side-Side-Side):* All three pairs of corresponding sides are equal.
- *SAS (Side-Angle-Side):* Two pairs of sides and the angle between them are equal.
- *ASA (Angle-Side-Angle):* Two pairs of angles and the included side are equal.
- *AAS (Angle-Angle-Side):* Two angles and a non-included side are equal.
- *HL (Hypotenuse-Leg):* For right triangles, if the hypotenuse and one leg are equal.

25.9.2 Triangle Similarity

Two triangles are *similar* if their corresponding angles are equal and their corresponding sides are in proportion. This means they have the same shape but not necessarily the same size. Criteria for similarity include:

- *AA (Angle-Angle):* Two pairs of corresponding angles are equal.
- *SAS (Side-Angle-Side):* One pair of corresponding angles is equal, and the sides including those angles are in proportion.
- *SSS (Side-Side-Side):* All three pairs of corresponding sides are in proportion.

25.10 Visualizations of Triangle Similarity and Congruence

25.10.1 SSS (Side-Side-Side)

A triangle is uniquely determined if all three sides are known.
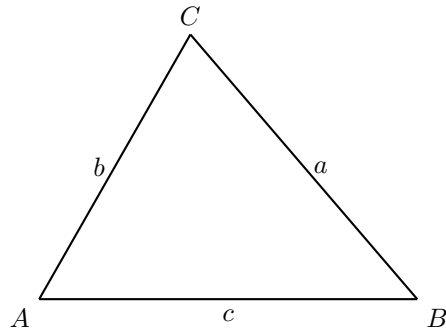
25.10.2 SAS (Side-Angle-Side)

A triangle is uniquely determined if two sides and the included angle are known.
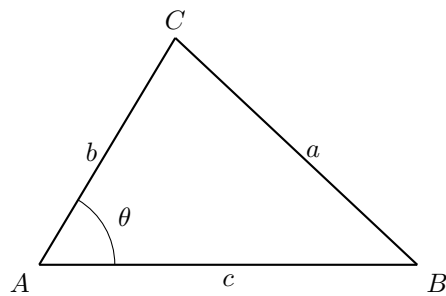
25.10.3 ASA (Angle-Side-Angle)

A triangle is uniquely determined if two angles and the included side are known.
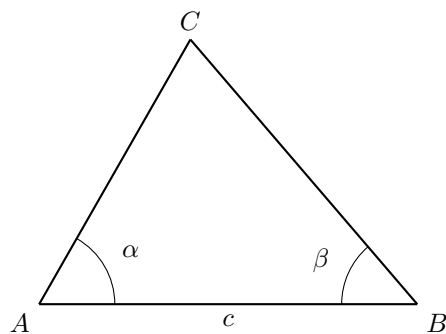
25.10.4 AAS (Angle-Angle-Side)

A triangle is uniquely determined if two angles and a non-included side are known.
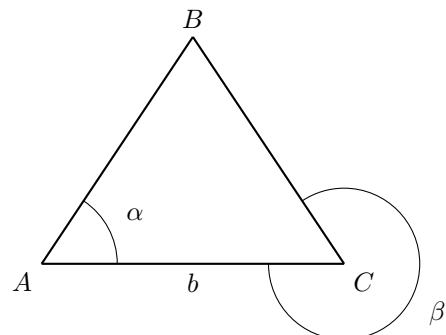
25.10.5 SSA (Side-Side-Angle)

Two sides and a non-included angle do *not always* determine a unique triangle. This case is ambiguous and can result in:

- One triangle (when the opposite side is long enough),
- Two triangles (ambiguous case),
- No triangle (when the side is too short).
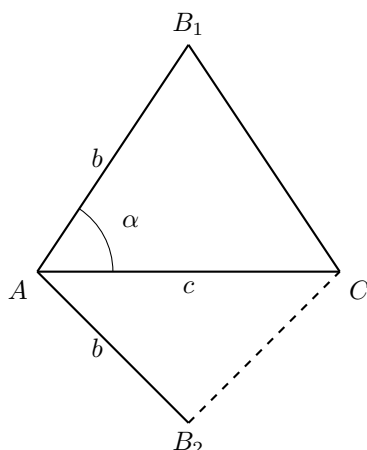
25.11 Triangle Medians, Angle Bisectors, and Altitudes

In any triangle, there are special line segments that connect vertices to specific points on the opposite sides. These segments play key roles in triangle geometry and have unique properties.

Median

A *median* of a triangle is a line segment that connects a vertex to the midpoint of the opposite side. Every triangle has three medians, and they intersect at a single point called the *centroid*, which is the triangle's center of mass. The centroid divides each median in a ratio of 2 : 1, with the longer part closer to the vertex.

Angle Bisector

An *angle bisector* is a line segment that divides an angle of the triangle into two equal parts and extends to the opposite side. All three angle bisectors intersect at a point called the *incenter*, which is the center of the triangle's incircle (the largest circle that fits inside the triangle and touches all three sides).

Altitude

An *altitude* of a triangle is a perpendicular segment from a vertex to the line containing the opposite side (not necessarily within the triangle). Altitudes intersect at a point called the *orthocenter*. Unlike medians and bisectors, altitudes can fall outside the triangle in obtuse triangles.

$$\text{Centroid} \rightarrow \text{Intersection of medians}$$
$$\text{Incenter} \rightarrow \text{Intersection of angle bisectors}$$
$$\text{Orthocenter} \rightarrow \text{Intersection of altitudes}$$

25.12 The Midsegment Theorem

The *Midsegment Theorem* states that the segment connecting the midpoints of two sides of a triangle is:

- *Parallel* to the third side of the triangle
- *Half the length* of the third side

Let $\triangle ABC$ be a triangle, and let D and E be the midpoints of sides AB and AC, respectively. Then the segment $\overline{DE}$ is the *midsegment*, and we have:

$$\overline{DE} \parallel \overline{BC} \quad \text{and} \quad DE = \frac{1}{2}BC$$

This theorem is useful in coordinate geometry, triangle proofs, and for simplifying complex geometric relationships within triangles.

25.13 Intercept Theorems

The intercept theorems describe relationships between segment lengths when two rays from a point intersect two parallel lines. They are based on similar triangles and allow us to calculate unknown lengths using proportions.

25.13.1 First Intercept Theorem

If two rays start from a common point and are intersected by two parallel lines, then:

$$\frac{a}{a'} = \frac{b}{b'},$$

where a and b are segments on one ray, and a', b' are the corresponding segments on the other ray.
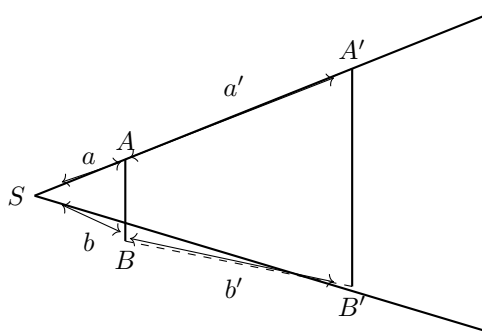
25.13.2 Second Intercept Theorem (General Form)

If a ray intersects two parallel lines, the segments from the origin point to the lines are in the same ratio as the segments along the parallels:

$$\frac{SA}{SA'} = \frac{SB}{SB'} \quad \text{and} \quad \frac{AB}{A'B'} = \frac{SA}{SA'}$$
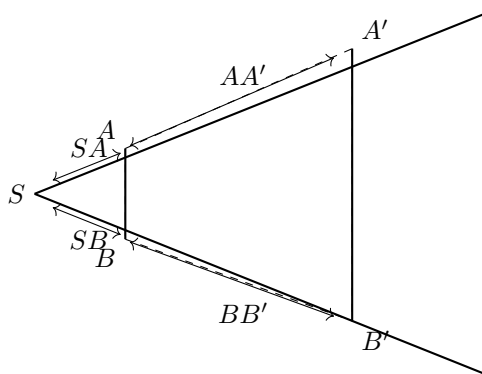
25.14   Proof of the Pythagorean Theorem

We want to prove that $c^2 = a^2 + b^2$ is a true statement.

Let us a build a Square with area c^2 and take our original triangle and put it on the sides on the square.

This will create a square of sides $(a+b)^2$ which is bigger than the inner square c^2.

Now let us find the area c^2 using what we know about this shape.

$$(a+b)^2 = 4(\frac{ab}{2}) + c^2$$
$$(a+b)^2 = 2ab + c^2$$
$$a^2 + 2ab + b^2 = 2ab + c^2$$
$$a^2 + b^2 = c^2$$

$$\mathfrak{QED}$$

26 Trigonometry

In this section, we will cover a variety of trigonometric functions and their properties.

26.1 Trigonometric Functions

- $\sin(x) = \frac{1}{\csc(x)}$

- $\cos(x) = \frac{1}{\sec(x)}$

- $\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{1}{\cot(x)}$

- $\csc(x) = \frac{1}{\sin(x)}$

- $\sec(x) = \frac{1}{\cos(x)}$

- $\cot(x) = \frac{\cos(x)}{\sin(x)} = \frac{1}{\tan(x)}$

26.2 SOH-CAH-TOA

- $\sin(x) = \frac{\text{opposite}}{\text{hypotenuse}}$

- $\cos(x) = \frac{\text{adjacent}}{\text{hypotenuse}}$

- $\tan(x) = \frac{\text{opposite}}{\text{adjacent}}$

26.3 Pythagorean Identities

- $\sin^2(x) + \cos^2(x) = 1$

- $1 + \tan^2(x) = \sec^2(x)$

- $1 + \cot^2(x) = \csc^2(x)$

26.4 Sum and Difference Formulas

- $\sin(a + b) = \sin(a)\cos(b) + \cos(a)\sin(b)$

- $\sin(a - b) = \sin(a)\cos(b) - \cos(a)\sin(b)$

- $\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b)$

- $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$

- $\tan(a + b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)}$

- $\tan(a - b) = \frac{\tan(a) - \tan(b)}{1 + \tan(a)\tan(b)}$

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$
$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$
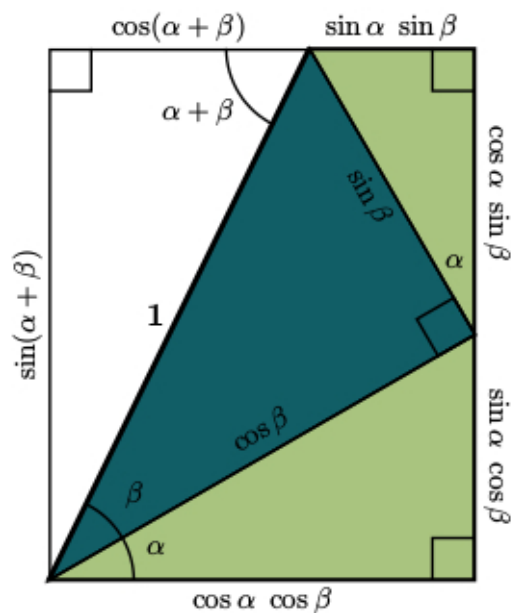
26.5 Double Angle Formulas

- $\sin(2x) = 2\sin(x)\cos(x)$

- $\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$

- $\tan(2x) = \frac{2\tan(x)}{1-\tan^2(x)}$

- $\csc(2x) = \frac{2\csc(x)\sec(x)}{1-\tan^2(x)}$

- $\sec(2x) = \frac{1+\tan^2(x)}{2\tan(x)}$

- $\cot(2x) = \frac{1-\tan^2(x)}{2\tan(x)}$

26.6 Half Angle Formulas

- $\sin\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1-\cos(x)}{2}}$

- $\cos\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1+\cos(x)}{2}}$

- $\tan\left(\frac{x}{2}\right) = \frac{\sin(x)}{1+\cos(x)} = \frac{1-\cos(x)}{\sin(x)} = \frac{\tan(x)}{1+\tan^2(x)}$

- $\csc\left(\frac{x}{2}\right) = \frac{1}{\sin\left(\frac{x}{2}\right)} = \pm\sqrt{\frac{2}{1-\cos(x)}}$

- $\sec\left(\frac{x}{2}\right) = \frac{1}{\cos\left(\frac{x}{2}\right)} = \pm\sqrt{\frac{2}{1+\cos(x)}}$

- $\cot\left(\frac{x}{2}\right) = \frac{1}{\tan\left(\frac{x}{2}\right)} = \frac{1+\cos(x)}{\sin(x)} = \frac{1-\tan^2(x)}{2\tan(x)}$

26.7 Product to Sum Formulas

- $\sin(a)\sin(b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)]$

- $\cos(a)\cos(b) = \frac{1}{2}[\cos(a-b) + \cos(a+b)]$

- $\sin(a)\cos(b) = \frac{1}{2}[\sin(a+b) + \sin(a-b)]$

- $\cos(a)\sin(b) = \frac{1}{2}[\sin(a+b) - \sin(a-b)]$

26.8 Power Reducing Formulas

- $\sin^2(x) = \frac{1-\cos(2x)}{2}$

- $\cos^2(x) = \frac{1+\cos(2x)}{2}$

- $\tan^2(x) = \frac{1-\cos(2x)}{1+\cos(2x)}$

- $\csc^2(x) = \frac{1}{\sin^2(x)} = \frac{2}{1-\cos(2x)}$

- $\sec^2(x) = \frac{1}{\cos^2(x)} = \frac{2}{1+\cos(2x)}$

- $\cot^2(x) = \frac{1}{\tan^2(x)} = \frac{1+\cos(2x)}{1-\cos(2x)}$

26.9 Even and Odd Functions

- $\sin(-x) = -\sin(x)$

- $\cos(-x) = \cos(x)$

- $\tan(-x) = -\tan(x)$

- $\csc(-x) = -\csc(x)$

- $\sec(-x) = \sec(x)$

- $\cot(-x) = -\cot(x)$

26.10 Graph Identity

$\cos(x) = \sin(x + \frac{\pi}{2})$

26.11 Trigonometric Values

θ (radians)	$\sin(\theta)$	$\cos(\theta)$	$\tan(\theta)$
0	0	1	0
$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
$\frac{\pi}{2}$	1	0	undefined
$\frac{2\pi}{3}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$-\sqrt{3}$
$\frac{3\pi}{4}$	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	-1
$\frac{5\pi}{6}$	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{3}}$
π	0	-1	0
$\frac{7\pi}{6}$	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$\frac{5\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	1
$\frac{4\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$\sqrt{3}$
$\frac{3\pi}{2}$	-1	0	undefined
$\frac{5\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$-\sqrt{3}$
$\frac{7\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	-1
$\frac{11\pi}{6}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{3}}$
2π	0	1	0

26.12 Unit Circle and Trigonometric Functions

The unit circle is a circle with a radius of 1 centered at the origin. Trigonometric functions like $\sin(\theta)$, $\cos(\theta)$, and $\tan(\theta)$ can be defined based on the coordinates of a point on the circle corresponding to angle θ.
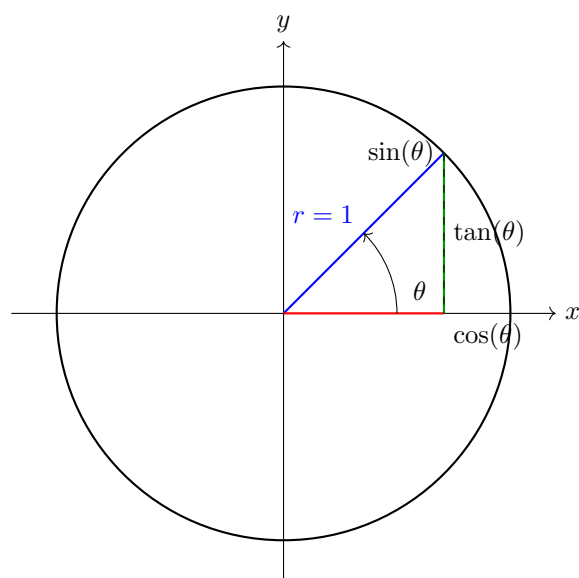
26.13 Definitions on the Unit Circle:

- $\cos(\theta)$ is the *x-coordinate* of the point on the circle.
- $\sin(\theta)$ is the *y-coordinate*.

- $\tan(\theta) = \dfrac{\sin(\theta)}{\cos(\theta)}$ is the ratio of the opposite side to the adjacent side in the triangle.

26.14 Trigonometric Rules

- $\frac{\sin \alpha}{a} = \frac{\sin \beta}{b} = \frac{\sin \gamma}{c}$
- $c^2 = a^2 + b^2 - 2ab \cos \gamma$
- $\frac{a-b}{a+b} = \frac{\tan \frac{1}{2}(\alpha - \beta)}{\tan \frac{1}{2}(\alpha + \beta)}$

26.15 Proof of the Cosine Rule

The *Cosine Rule* relates the lengths of the sides of a triangle to the cosine of one of its angles. For any triangle ABC with sides $a = BC$, $b = AC$, and $c = AB$, and angle $\gamma = \angle ACB$, the cosine rule states:

$$c^2 = a^2 + b^2 - 2ab \cos(\gamma)$$

Proof:

Consider triangle ABC, and place it on the coordinate plane such that point c is at the origin $(0,0)$, point B is at $(a, 0)$, and point A lies somewhere in the plane. Let angle $\gamma = \angle ACB$.

Let the coordinates of point A be $(b \cos \gamma, b \sin \gamma)$, since $b = AC$ and we're using polar coordinates to define its location relative to point c.

Using the distance formula to find side $c = AB$, we get:

$$
\begin{aligned}
c^2 &= (a - b \cos \gamma)^2 + (0 - b \sin \gamma)^2 \\
&= a^2 - 2ab \cos \gamma + b^2 \cos^2 \gamma + b^2 \sin^2 \gamma \\
&= a^2 + b^2 (\cos^2 \gamma + \sin^2 \gamma) - 2ab \cos \gamma \\
&= a^2 + b^2 - 2ab \cos \gamma
\end{aligned}
$$

Since $\cos^2 \gamma + \sin^2 \gamma = 1$.

Thus, we have proven that:

$$c^2 = a^2 + b^2 - 2ab \cos \gamma$$

$$\mathfrak{QED}$$

27 Linear Systems of Equations

In this section, we will discuss the solution of linear systems of equations. A linear system of equations is a set of equations that can be expressed in the form:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m,$$

where a_{ij} are the coefficients of the variables x_j, and b_i are the constants on the right-hand side of the equations. The goal is to find the values of $x_1, x_2, \ldots, x_n$ that satisfy all equations simultaneously.

27.1 Matrix Representation

A linear system can be represented in matrix form as:

$$Ax = b$$

Where A is the coefficient matrix, x is the vector of variables, and b is the vector of constants. The coefficient matrix A is an $m \times n$ matrix, where m is the number of equations and n is the number of variables. The vector x is an $n \times 1$ column vector, and the vector b is an $m \times 1$ column vector.

27.2 Solution Set of a Linear System

In the context of Definition 6.1, the solution set $L(A, b)$ of the system of equations associated with (A, b) is defined as

$$L(A, b) := \{x \in \mathbb{F}^n \mid Ax = b\}.$$

27.3 Equality of the Rank $\mathrm{rg}(A) = \mathrm{rg}_S(A)$

We have $\mathrm{rg}(A) = \mathrm{rg}_S(A)$.

Proof:

For $x = (x_1, \ldots, x_n)^T \in \mathbb{F}^n$, it holds that

$$Ax = \sum_{i=1}^{n} x_i a_i.$$

It follows that

$$\mathrm{rg}(A) = \mathrm{rg}(L_A) = \dim(\mathrm{img}(L_A)) = \dim\left(\{L_A(x) \mid x \in \mathbb{F}^n\}\right) = \dim\left(\{Ax \mid x \in \mathbb{F}^n\}\right)$$

$$= \dim\left\{\sum_{i=1}^{n} x_i a_i \,\middle|\, x_i \in K\right\} = \dim(L(a_1, \ldots, a_n)) = \mathrm{rg}_S(A).$$

27.4 Solvability of the Linear System $Ax = b$

The linear system $Ax = b$ is solvable if and only if

$$\text{rg}(a_1, \ldots, a_n) = \text{rg}(a_1, \ldots, a_n, b).$$

More concisely, we write

$$\text{rg}(A) = \text{rg}(A, b).$$

Proof:

We have

$$Ax = (a_1, \ldots, a_n)x = x_1 a_1 + \cdots + x_n a_n = b.$$

Therefore, a solution vector x exists if and only if

$$b \in L(a_1, \ldots, a_n) \quad \Leftrightarrow \quad \text{rg}(A) = \text{rg}(A, b),$$

where $L(a_1, \ldots, a_n)$ denotes the linear span of $a_1, \ldots, a_n$.

27.5 Solution Set of the Linear System $Ax = b$

Let $x_s \in \mathbb{F}^n$ be a solution of $Ax = b$. Then the solution set is given by

$$L(A, b) = x_s + \ker(A) = \{x_s + x \mid x \in \ker(A)\}.$$

Proof:

$$x \in \ker(A) \iff Ax = 0 \iff A(x_s + x) = Ax_s + Ax = Ax_s = b.$$

27.6 Free Variables

In a system of linear equations, a *free variable* is a variable whose column in the coefficient matrix does not correspond to a pivot column after row reduction. Free variables are not directly solved by any equation in the system; instead, they can take any real value, and the values of the leading (or dependent) variables are expressed in terms of them.

Free variables arise naturally in systems with infinitely many solutions and play a key role in describing the solution set. Given a system of with p equations and q unknowns then if $q > p$ then there are $q - p$ free variables.

Example:

Consider the system

$$\begin{cases} x_1 + 2x_2 - x_3 = 4 \\ 3x_1 + 6x_2 - 3x_3 = 12 \end{cases}$$

which row reduces to

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here, the pivot is in the first column corresponding to x_1, while x_2 and x_3 are free variables. Assigning parameters to x_2 and x_3, the general solution can be expressed in terms of these free variables.

27.7 Gaussian Elimination

Gaussian elimination is a method for solving linear systems by transforming the system into an upper triangular form. The steps involved in Gaussian elimination are:

1. Forward elimination: Transform the system into an upper triangular form by eliminating the variables from the equations.

2. Back substitution: Solve for the variables starting from the last equation and substituting back into the previous equations.

The forward elimination process involves performing row operations on the augmented matrix $[A|\mathbf{b}]$ to create zeros below the diagonal. The row operations include:

- Swapping two rows

- Multiplying a row by a non-zero scalar

- Adding or subtracting a multiple of one row from another row

Once the matrix is in upper triangular form, back substitution is used to find the values of the variables. The last equation gives the value of the last variable, which can then be substituted into the previous equations to find the other variables.

27.8 Gauss-Jordan Elimination

Gauss-Jordan elimination is an extension of Gaussian elimination that transforms the matrix into reduced row echelon form (RREF). In RREF, each leading entry in a row is 1, and all entries above and below the leading entry are zeros. The steps involved in Gauss-Jordan elimination are:

1. Forward elimination: Transform the system into an upper triangular form.

2. Back substitution: Transform the upper triangular matrix into RREF by eliminating the entries above the leading 1s.

3. Solve for the variables directly from the RREF matrix.

The Gauss-Jordan elimination method is particularly useful for finding the inverse of a matrix, as it 7 can be applied to the augmented matrix $[A|I]$, where I is the identity matrix. If the left side of the augmented matrix becomes I, then the right side will be the inverse of A.

Example:

Solve the following system of equations using Gaussian elimination:

$$2x + 3y + z = 1$$
$$4x + y - z = 2$$
$$-2x + y + 3z = 3$$

Solution: The augmented matrix for the system is:

$$\begin{bmatrix} 2 & 3 & 1 & | & 1 \\ 4 & 1 & -1 & | & 2 \\ -2 & 1 & 3 & | & 3 \end{bmatrix}$$

Performing row operations to eliminate the variables, we can transform the matrix into upper triangular form:

$$\begin{bmatrix} 1 & \frac{3}{2} & \frac{1}{2} & | & \frac{1}{2} \\ 0 & -5 & -3 & | & 0 \\ 0 & 0 & 1 & | & 1 \end{bmatrix}$$

Now, we can perform back substitution to find the values of x, y, and z:

$$z = 1$$

$$-5y - 3z = 0 \implies y = -\frac{3}{5}$$

$$2x + 3y + z = 1 \implies x = \frac{1}{5}$$

Thus, the solution to the system is:

$$x = \frac{1}{5}$$
$$y = -\frac{3}{5}$$
$$z = 1$$

27.9 Homogeneous Linear Equations

A linear equation is said to be *homogeneous* if its constant term is zero. That is, it can be written in the form:

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

Such equations always have at least the trivial solution $x_1 = x_2 = \cdots = x_n = 0$.

27.10 Particular Solution

A particular solution to a linear system of equations is a specific solution that satisfies the system. It can be found using various methods, including substitution, elimination, or matrix methods. A particular solution is not unique; there may be multiple particular solutions depending on the system. A particular solution can be found by substituting specific values for the variables and solving for the remaining variables. For example, in the system

$$2x + 3y = 5$$
$$4x - y = 1$$

we can substitute $x = 1$ into the first equation to find y:

$$2(1) + 3y = 5$$
$$3y = 3$$
$$y = 1$$

Thus, $(x, y) = (1, 1)$ is a particular solution to the system. However, this is not the only solution; other values of x may yield different values of y.

27.11 General = Particular + Homogeneous

The general solution of a linear system of equations is the complete set of solutions that satisfy the system. It can be expressed as the sum of a particular solution and the general solution of the associated homogeneous system.

The general solution can be written as:

$$x = x_p + x_h$$

Where x_p is a particular solution to the non-homogeneous system, and x_h is the general solution to the homogeneous system. The homogeneous system is obtained by setting the right-hand side of the equations to zero:

$$Ax = 0$$

The theorem says:

Any linear system's solution set has the form:

$$\left\{ \vec{p} + c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \;\middle|\; c_1, \ldots, c_k \in \mathbb{R} \right\},$$

where $\vec{p}$ is a particular solution to the system, and the vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ form a basis of the solution space to the corresponding homogeneous system. The number k equals the number of **free variables** the system has after applying Gaussian elimination.

27.12   Linear Combination Lemma

Any linear combination of linear combinations is a linear combination.

27.13   Example: Gaussian Elimination with 3 Equations and 4 Unknowns

Consider the following system of linear equations:

$$x_1 + 2x_2 + x_3 + x_4 = 4$$
$$2x_1 + 5x_2 + x_3 + 3x_4 = 10$$
$$x_1 + 3x_2 + 2x_3 + 2x_4 = 7$$

Step 1: Augmented Matrix

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 2 & 5 & 1 & 3 & 10 \\ 1 & 3 & 2 & 2 & 7 \end{bmatrix}$$

Step 2: Eliminate below pivot in column 1

- Row 2 = Row 2 $- 2 \times$ Row 1
- Row 3 = Row 3 $-$ Row 1

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 1 & 1 & 1 & 3 \end{bmatrix}$$

Step 3: Eliminate below pivot in column 2

- Row 3 = Row 3 $-$ Row 2

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 0 & 2 & 0 & 1 \end{bmatrix}$$

Step 4: Back Substitution

From Row 3:

$$2x_3 = 1 \Rightarrow x_3 = \frac{1}{2}$$

From Row 2:
$$x_2 - x_3 + x_4 = 2 \Rightarrow x_2 = 2 + x_3 - x_4 = 2 + \frac{1}{2} - x_4 = \frac{5}{2} - x_4$$

From Row 1:
$$x_1 + 2x_2 + x_3 + x_4 = 4 \Rightarrow x_1 = 4 - 2x_2 - x_3 - x_4$$

Substitute:
$$x_1 = 4 - 2\left(\frac{5}{2} - x_4\right) - \frac{1}{2} - x_4 = 4 - 5 + 2x_4 - \frac{1}{2} - x_4 = -1 - \frac{1}{2} + x_4 = -\frac{3}{2} + x_4$$

General Solution

Let $x_4 = t$ (free variable), then:

$$
\begin{aligned}
x_1 &= -\frac{3}{2} + t \\
x_2 &= \frac{5}{2} - t \\
x_3 &= \frac{1}{2} \\
x_4 &= t
\end{aligned}
\quad \text{with } t \in \mathbb{R}
$$

Solution Set:

$$\left\{ \begin{bmatrix} -\frac{3}{2} \\ \frac{5}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}$$

27.14 The Determinant, the cross product and the solutions of linear Systems of Equations

A linear system of three equation has the following properties:

- There is a unique solution if the determinant of the coefficient matrix is non-zero.

$$\langle (a \times b), c \rangle = \det(a, b, c) \neq 0$$

- There are infinitely many solutions if the determinant of the coefficient matrix is zero.

$$\langle (a \times b), c \rangle = \det(a, b, c) = 0$$

- There is no solution if the determinant of the coefficient matrix is zero and the system is inconsistent.

$$\langle (a \times b), c \rangle = 0$$

27.15 Chart for the number of solutions

```
                              ┌─────────────────────┐
                              │ Linear n × n System │
                              └─────────────────────┘
                                        │
                                        ▼
┌──────────────────────┐    No      ╱─────────────╲
│  Exactly one solution │ ◀─────────  det(A) = 0?  
└──────────────────────┘            ╲─────────────╱
                                        │ Yes
                                        ▼
                                  ╱───────────────────────╲
                                   rank(A) = rank(A, b)?
                                  ╲───────────────────────╱
                           Yes ╱                    │ No
                              ╱                      ▼
┌──────────────────────────────────────┐    ┌─────────────┐
│ Inf. sol. with (n − rank(A)) parameters │   │ No solution │
└──────────────────────────────────────┘    └─────────────┘


                              ┌─────────────────────┐
                              │ Linear m × n System │
                              └─────────────────────┘
                                        │
                                        ▼
┌──────────────┐    No      ╱───────────────────────╲
│ No solution  │ ◀─────────  rank(A) = rank(A, b)?
└──────────────┘            ╲───────────────────────╱
                                        │ Yes
                                        ▼
                                  ╱─────────────╲
                                   rank(A) = n?
                                  ╲─────────────╱
                           No ╱                 ╲ Yes
                             ╱                   ╲
┌──────────────────────────────────────┐   ┌──────────────────────┐
│ Inf. sol. with (n − rank(A)) parameters │  │  Exactly one solution │
└──────────────────────────────────────┘   └──────────────────────┘
```
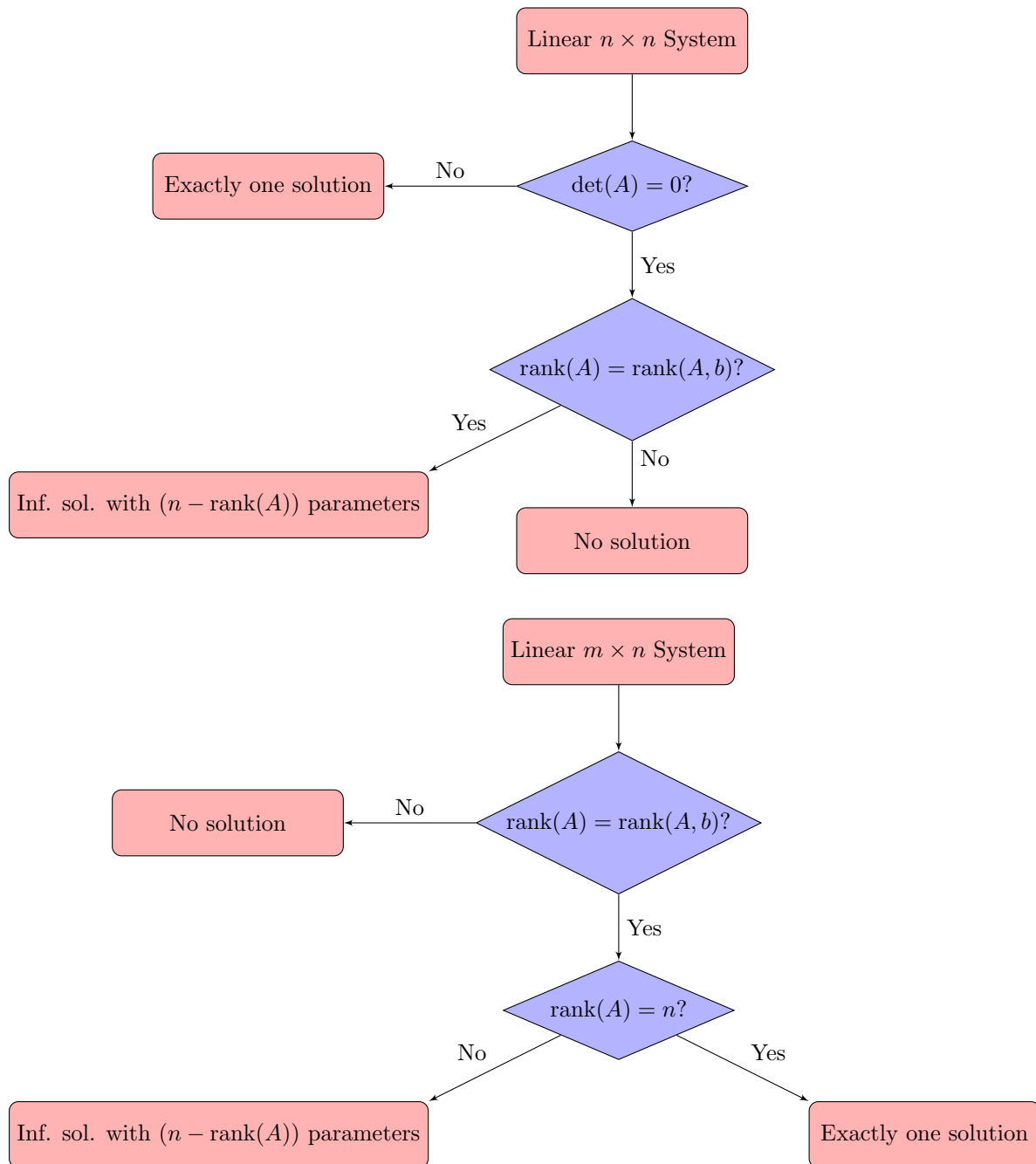
27.16 Proof of the Gaussian Elimination

The row operations do not change the solution set.

Proof:

A linear system of equations of the form $Ax = b$, $A \in \mathbb{F}^{m \times n}$ with some solution x can be multiplied by some elementary matrix C_1, C_2, C_3 which correspond to a row operation $CAx = Cb$.

The Equality is kept and therefore, $L(A, b) \subset L(CA, Cb)$. If $x \in L(CA, Cb)$ then

$$CAx = Cb \implies C^{-1}CAx = C^{-1}Cb \implies Ax = b$$

Because of the invertibility of elementary matrices we can be sure that $L(CA, Cb) \subset L(A, b)$.

Now for the case that $L(A, b) = \emptyset$. If we assume that $L(CA, CB) \neq \emptyset$, then there would be some $x \in L(CA, CB)$ with $CAx = Cb$. By multiplying both side by C^{-1} we get $Ax = b$ and therefore, $L(A, b) \neq \emptyset$, which is a contradiction to our original assumption $L(A, b) = \emptyset$.

<div align="right">𝔔𝔈𝔇</div>

Now we only have to prove that each regular matrix can be transformed in the reduced echelon form or even row reduced echelon form. This is the equivalent of saying that the row operations do not change the rank of the matrix.

Proof:

If C is an elementary matrix and $A \in \mathbb{F}^{m \times n}$ some matrix. The statement $rg(CA) = rg(A)$ is true because of the invertibility of the elementary matrices and the isomorphism in vector spaces.

Given $A = (a_1, a_2, \ldots, a_n)$. The exchange of two columns preserves $L := L(a_1, \ldots, a_n)$, and also the $rg_s(A)$. Concerning the scaling of a column by some λ $L := L(a_1, \ldots, \lambda a_k, \ldots, a_n)$ also preserves $dim(L) = rg_s(A)$. And finally, the addition $L = L(a_1, \ldots, a_i + \lambda a_k, \ldots, a_n)$ also preserves $rg_s(A) = rg(A)$. Thus, $rg(CA) = rg(A)$.

<div align="right">𝔔𝔈𝔇</div>

The only part left is to prove that if A has an inverse, then with use of the row operations we can transform it in to a matrix with an upper triangular with also no zero element on the main diagonal.

Proof:

First we eliminate all the elements under the main diagonal for all columns until the column $k-1$. With this, we have created a new matrix $A^{(\ell)} = C_\ell \cdots \dot{C}_1 A$ which looks like:

$$A^{(\ell)} = \begin{bmatrix} a_{11}^{(\ell)} & \cdots & a_{1,k-1}^{(\ell)} & a_{1k}^{(\ell)} & a_{1,k+1}^{(\ell)} & \cdots & a_{1n}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{k-1,k-1}^{(\ell)} & a_{k-1,k}^{(\ell)} & a_{k-1,k+1}^{(\ell)} & \cdots & a_{k-1,n}^{(\ell)} \\ 0 & \cdots & 0 & a_{kk}^{(\ell)} & a_{k,k+1}^{(\ell)} & \cdots & a_{kn}^{(\ell)} \\ 0 & \cdots & 0 & \vdots & a_{k+1,k+1}^{(\ell)} & \cdots & a_{k+1,n}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nk}^{(\ell)} & \cdots & \cdots & a_{nn}^{(\ell)} \end{bmatrix}$$

Because of $rg(A) = rg(A^{(\ell)})$, $a_{ii}^{(\ell)} \neq 0$ for $1 \leq i < k$. All of these elements will not change for the rest of the Gaussian Elimination. Now we want to build another matrix with row operations $A^{(\mu)}$ where $a_{kk}^{(\mu)} \neq 0$ and $a_{ik}^{(\mu)} = 0$ for $i > k$. If $a_{kk}^{(\ell)} \neq 0$ then (i, k)-th element will become 0 by subtracting $\frac{a_{ik}^{(\ell)}}{a_{kk}^{(\ell)}}$ time the k-th row of the i-th row. So can the k-th column be transformed with a series of row operations to the desired form. But if $a_{kk}^{(\ell)} = 0$, then you have to swap it with another row under it. This is the critical point of the algorithm: What if there is no row, also $a_{ik}^{(\ell)} = 0$ for $i \geq k$? Then $A^{(\ell)}$ would be:

$$A^{(\ell)} = \begin{bmatrix} a_{11}^{(\ell)} & \cdots & a_{1,k-1}^{(\ell)} & a_{1k}^{(\ell)} & a_{1,k+1}^{(\ell)} & \cdots & a_{1n}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_{k-1,k-1}^{(\ell)} & a_{k-1,k}^{(\ell)} & a_{k-1,k+1}^{(\ell)} & \cdots & a_{k-1,n}^{(\ell)} \\ 0 & \cdots & 0 & 0 & a_{k,k+1}^{(\ell)} & \cdots & a_{kn}^{(\ell)} \\ 0 & \cdots & 0 & \vdots & a_{k+1,k+1}^{(\ell)} & \cdots & a_{k+1,n}^{(\ell)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & a_{nn}^{(\ell)} \end{bmatrix}$$

But, that would imply that the k-th column is a linear combination of the other $k-1$ columns, and then $rg(A^{(\ell)}) \leq n - 1$ which is a contradiction to the assumption $rg(A^{(\ell)}) = n$

<div align="right">𝔔𝔈𝔇</div>

27.17 Alternative Method for finding the Rank of a matrix

For a matrix $A \in \mathbb{F}^{(m \times n)}$. The biggest submatrix of the size $r \times r$ with a non-zero determinant gives us the *rank* of the matrix. $rg(A) = r$.

27.18 Cramer's Rule

For a square matrix $A = (a_1, \ldots, a_n)$ and $x, b \in \mathbb{F}^n$ with $Ax = b$ with $\det A \neq 0$ then

$$A_i = (a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n), i = 1, \ldots, n.$$

The solution i is then

$$x_i = \frac{\det(A_i)}{\det(A)}$$

Proof:

Given a matrix $X_i \in \mathbb{F}^{n \times n}$ by

$$\begin{bmatrix} 1 & & \cdots & x_1 & \cdots & \\ \vdots & & \ddots & \vdots & & 0 \\ & & & x_{i-1} & & \\ & 0 & & x_i & & \\ & & & x_{i+1} & & \\ & & & \vdots & \ddots & \\ & & & x_n & & 1 \end{bmatrix}$$

$X_i = (e_1, \ldots, e_{i-1}, e_i, e_{i+1}, \ldots, e_n)$. By using an n-times the Laplace formula in the first row shows $|X_i| = x_i$.

$$AX_i = A(e_1, \ldots, e_{i-1}, e_i, e_{i+1}, \ldots, e_n)$$
$$AX_i = A(a_1, \ldots, a_{i-1}, Ax, a_{i+1}, \ldots, a_n)$$
$$AX_i = A(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) = A_i$$

This implies $|A_i| = |AX_i| = |A||X_i|$ therefore, $x_i = \frac{|A_i|}{A}$.

$$\mathfrak{QED}$$

Another way of thinking about this in a more geometric way is to first think at the determinant as the area/volume/etc. that is spanned by the basis vector and the stretching factor of the transformation $\det(A)$. Now visualize your matrix A as a linear transformation with a non-zero determinant and think about $\det(A_i)$ as the area/volume/etc. which is spanned by the basis vector but with one of the vector substituted by b.

Now in two dimensions, think of the area spanned by the basis vectors $A = xy$. The signed area of the original parallelogram gets stretched by $\det(A)$ this gives us $Area = \det(A)y \Rightarrow y = \frac{Area}{\det(A)}$. This also works for x.

28 Over/Underdetermined Systems of Equations

Let us start with a little motivation for this chapter. Imagine that you have collected some data points, and you want to relate them via a linear function $Ax = (\alpha, \beta, \dots)^T$. This would probably not work because it will be impossible to connect all the points, but you could make an approximation which is the line $\alpha x_1 + \beta x_2 + \dots$ that is the closest to all the points in your data set. Thus, we will find an approximation b that is close to the image of A ($Im(A)$).

For this purpose will use a theorem that says that the best approximation is the orthogonal projection of $p_A(b)$. This projection can be found if there exist an orthonormal basis of the subspace we want to project our vector onto.

28.1 Normal Equations

Given a projection $p_A(b)$ onto the subspace U generated by $A = (a_1, \dots, a_n)$ then there exists an $x \in \mathbb{R}^n$ such that

$$p_A(b) = \sum_{k=1}^{n} x_k a_k = Ax$$

Also, because of the orthogonal property then

$$
\begin{aligned}
b - p_A(b) \perp U &\implies b - Ax \in U^{\perp} \\
&\implies b - Ax \perp a_k \quad \forall k \\
&\implies \langle a_k, b - Ax \rangle = 0 \quad \forall k \\
&\iff A^T(b - Ax) = \vec{0} \\
&\implies A^T b = Ax
\end{aligned}
$$

This equation $A^T b = Ax$ is called the *normal equation.*

A *normal equation* of a matrix $A \in \mathbb{R}^{(m \times n)}$ has a solution. If $rg(A) = n$ then

$$x = (A^T A)^{-1} A^T b$$

is a unique solution to the system. The term $(A^T A)^{-1} A^T$ is called the *Generalized Inverse* of A.

Proof:

The statement is true because of the existence of the orthogonal projection which give us a concrete basis. Also, the fact that $rg(A) = n$ also implies that. And also because of the invertibility of $(A^T A)^{-1}$.

$$\mathfrak{QED}$$

28.2 Method of the Least Squares

For a linear system of equations $Ax = b$ with $A \in \mathbb{R}^{(m \times n)}$, $b \in \mathbb{B}^m$ and $m \geq n$. For the case $rg(A) = n$ we have

$$x_s = (A^T A)^{-1} A^T b$$

and

$$\|b - Ax_s\| = min_{z \in \mathbb{R}^n} \|b - Az\|$$

Here the vector z is called *approximation solution via the method of the least squares.*

Proof:

We know that because of $rg(A) = n$ our solution is unique. And that $p_A(b) = Ax$ is our orthogonal projection. We can use Pythagoras theorem to

$$\|b - Az\|^2 = \|(b - p_A(b)) + (p + p_A(b))\|^2$$
$$= \|(b - p_A(b))\|^2 + \|(p + p_A(b))\|^2$$
$$\geq \|(b - p_A(b))\|^2$$
$$= \|b - Ax\|^2$$

Thus, $p_A(b)$ gives us the smallest distance to the $Img(A)$

$$\mathcal{QED}$$

Example I:

$$x = 1, x = -1$$

$$\|Ax - b\|^2 = (x + 1)^2 + (x - 1)^2 = 2x^2 + 2$$

Then $\|Ax - b\|^2$ gives us the smallest distance for $x = 0$.

Example II:

We can also take the last example and solve it with an alternative method.

$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, x = (x), b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Let us use the normal equations

$$A^T Ax = A^T b$$

$$(1, 1) \begin{bmatrix} 1 \\ 1 \end{bmatrix} x = (1, 1) \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\implies 2x = 0$$

Example 3:

Given are

$$x_1 = 1, x_2 = 2, x_1 + x_2 = -1$$

with the corresponding matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

and solution vector

$$b = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Now let us build the normal equation

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Now with Gaussian elimination we get $x = (-\frac{1}{3}, \frac{2}{3})^T$.

Finally, we can use the $\vec{x}$ to solve $Ax - b$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \frac{4}{3} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

And therefore, $\|Ax - b\|^2 = \frac{48}{9}$.

28.3   Underdetermined Systems of Equations

For a linear system $Ax = b$ such that $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $m \leq n$. For $rg(A) = m$ we get the solution vector

$$x_s = A^T (AA^T)^{-1} b$$

Also, x_s is the solution with the smallest length.

$$\|x_s\| = min_{Ax=b} \|x\|,$$

and x_s is orthogonal to the homogenous system $Az = 0$ which means

$$x_s \perp ker(A)$$

Finally, we call for $rg(A) = m$, $A^T(AA^T)^{-1}$ the *Generalized Inverse* of A.

Proof:

First, we show the invertibility of AA^T because this implies that x_s is defined.

Given $x \in kern(AA^T)$. This implies that

$$AA^T x = 0$$

which implies

$$x^T AA^T x = 0$$

because of $x^T A = A^T x$ and this gives us

$$\|A^T x\| = 0$$

therefore, $x \in kern(A^T)$.

Now because of $rg(A^T) = rg(A) = m$ we get $dim(kern(A)^T) = 0$, thus $x = \vec{0}$. Because of that $kern(AA^T) = 0$ and $rg(AA^T) = m$ thus, the invertibility is proven.

x_s is a solution of $Ax = b$ because of

$$Ax_s = A(A^T(AA^T)^{-1}b) = AA^T(AA^T)^{-1}b = b$$

$x_s \perp kern(A)$ because for $z \in kern)(A)$

$$\langle z, x_s \rangle = z^T A^T (AA^T)^{-1} b = \langle Az, (AA^T)^{-1} b \rangle = 0$$

doe to $Az = 0$.

Finally, thanks to the general solution theorem we know that $x = x_s + z$. Using Pythagoras and by taking into consideration $x_s \perp kern(A)$

$$\|x\|^2 = \|x\|^2 + \|z\|^2 \geq \|x_s\|^2$$

Example:

We have the equation $2x_1 + 2x_2 - x_3 = 6$.

In this case the solution build a plane in $\mathbb{R}^3$, thus the solution is the normal vector that cuts the plane. We build the hessian normal form.

$$\frac{1}{3}(2x_1 + 2x_2 - x_3) = 2$$

whose solution is $x = \frac{2}{3}(2, 2, -1)^T$.

Now with our method: $A = (2, 2, -1)$, $x = (x_1, x_2, x_3)$, $b = 6$

$$x = A^T(AA^T)^{-1}b = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \left((2, 2, -1) \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \right)^{-1} 6$$

$$x = \frac{2}{3}(2, 2, -1)^T$$

29 Analytical Geometry

In this section, we will cover the topics for the geometry of $\mathbb{R}^2$ and $\mathbb{R}^3$. And maybe also in higher dimensions.

29.1 Vectors and Points

In analytical geometry, points and vectors are the basic elements.

A point in $\mathbb{R}^3$ is represented as $\vec{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Or as $(x_1, x_2, \ldots, x_n)$.

A vector is an object with direction and magnitude (in this case), also represented as $\vec{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$.

Or as $(x_1, x_2, \ldots, x_n)^T$.

29.2 Vector Addition and Scalar Multiplication

Given two vectors $\vec{a}$ and $\vec{b}$:

$$\vec{a} \pm \vec{b} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \pm \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \begin{bmatrix} a_x \pm b_x \\ a_y \pm b_y \end{bmatrix}$$

For a scalar λ and vector $\vec{a}$:

$$\lambda\vec{a} = \lambda \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} \lambda a_x \\ \lambda a_y \end{bmatrix}$$
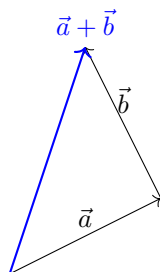
Figure 1: Vector addition: $\vec{a} + \vec{b}$

29.3 Equation of a Line

A line is defined by a point $\vec{P}$ and a direction vector $\vec{v}$:

$$\vec{r}(t) = \vec{P} + t\vec{v}, \quad t \in \mathbb{R}$$

29.4 Equation of a Plane

A plane is defined by a point $\vec{P}$ and a normal vector $\vec{n}$:

$$P : \vec{x} = \vec{s} + \lambda_1 \vec{v_1} + \lambda_2 \vec{v_2}$$

29.5   Dot Product

The inner product of vectors $\vec{a}$ and $\vec{b}$ is a function $\langle x, y \rangle := V \times V \to \mathbb{K}$ with the following properties:

$\forall \vec{a}, \vec{b} \in V$:

- $\langle \vec{a}, \vec{b} \rangle = \langle \vec{b}, \vec{a} \rangle$

- $\langle \vec{a}, \vec{b} + \vec{c} \rangle = \langle \vec{a}, \vec{b} \rangle + \langle \vec{a}, \vec{c} \rangle$

- $\langle \vec{a}, \lambda \vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle$

- $\langle \vec{a}, \vec{b} \rangle = 0 \Leftrightarrow \vec{a} = \vec{0} \vee \vec{b} = \vec{0}$

Formula:

$$\langle \vec{a}, \vec{b} \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Proof:

Let us look at the following Triangle:
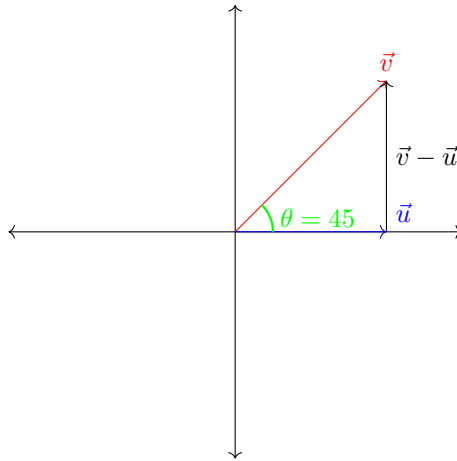
We can use the Cosine law: $\|v - u\|^2 = \|v\|^2 + \|u\|^2 - 2\|u\|\|v\| \cos \theta$.

Now we expand:

$$\|u\|^2 = u_1^2 + u_2^2 \ \text{ and } \ \|v\|^2 = v_1^2 + v_2^2$$

Then: $\|v - u\|^2 = (v_1 - u_1)^2 + (v_2 - u_2)^2$ and by re-grouping after expansion we get

$$\langle v, u \rangle = v_1 u_1 + v_2 u_2 = \|v\|\|u\| \cos \theta$$

$$\mathfrak{QED}$$

And if we manipulates this expression we can also prove the angle formulas. And with that information we can prove that

$$\cos \frac{\pi}{2} = \frac{\langle v, u \rangle}{\|v\|\|u\|} = 0$$

Only if the Dot Product is equal to zero so, like that we have also proved that only orthogonal vector have a dot product of zero.

29.5.1 Dot Product as Projection

The dot product of two vectors can be interpreted as projecting one onto another and then multiplying the length of the projecting to the magnitude of the other. For real vectors.

This comes from the fact that taking a dot product of real vectors is the same as multiplying one of them with the linear transformation formed by the transpose version of the other vector.

The transformation will take a vector from our space and take it to the number line and it is also linear. In our case,

$$\text{Proj} = \begin{bmatrix} u_{x_1} & u_{x_2} & \cdots & u_{x_n} \end{bmatrix}$$

is defined this way because when we project our basis vector $\vec{b}_i$ onto the line our vector $\vec{u}$ because of the symmetry we will get the same value as if we were projecting $\vec{u}$ onto the line $\vec{b}_i$ lives in.

29.6 Norm of a Vector

The norm of a vector $\vec{a}$ is described by different norms:

It also has the following properties:

- $\|\vec{a}\| \geq 0$ and $\|\vec{a}\| = 0 \Leftrightarrow \vec{a} = \vec{0}$

- $\|\lambda \vec{a}\| = |\lambda| \|\vec{a}\|$

- $\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|$ (Triangle inequality)

- $\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\langle \vec{a}, \vec{b} \rangle$ (Pythagorean theorem)

These are the most common norms, although we will be using primarily the euclidean norm:

- *Euclidean norm*:
$$\|\vec{a}\| = \sqrt{\langle \vec{a}, \vec{a} \rangle} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

- *Manhattan norm*:
$$\|\vec{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n|$$

- *Maximum norm*:
$$\|\vec{a}\|_\infty = \max(|a_1|, |a_2|, \ldots, |a_n|)$$

29.7 Angle Relations

The angle θ between two vectors is:

$$\cos\theta = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|\|\vec{b}\|}, \quad \theta = \arccos\left(\frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|\|\vec{b}\|}\right)$$

29.7.1 Line-Line Angle

Use direction vectors $\vec{v}_1$ and $\vec{v}_2$:

$$\theta = \arccos\left(\frac{\langle \vec{v}_1, \vec{v}_2 \rangle}{\|\vec{v}_1\|\|\vec{v}_2\|}\right)$$

29.7.2 Line-Plane Angle

Let $\vec{v}$ be the line's direction and $\vec{n}$ the plane's normal:

$$\theta = \arcsin\left(\frac{\langle \vec{v}, \vec{n} \rangle}{\|\vec{v}\|\|\vec{n}\|}\right)$$

29.7.3 Plane-Plane Angle

Angle between planes is angle between their normals:

$$\theta = \arccos\left(\frac{\langle \vec{n}_1, \vec{n}_2 \rangle}{\|\vec{n}_1\|\|\vec{n}_2\|}\right)$$

29.8 Line Relations

Two lines can be:

- *Identical*: same direction vector and point
- *Parallel*: direction vectors are proportional
- *Intersecting*: one solution for t_1, t_2 such that $\vec{r}_1(t_1) = \vec{r}_2(t_2)$
- *Skew*: not parallel, do not intersect

To find the relation:

1. Check if direction vectors are scalar multiples $\Rightarrow$ parallel
2. Solve $\vec{P}_1 + t\vec{v}_1 = \vec{P}_2 + s\vec{v}_2$ for t and s:
   - Solution exists $\Rightarrow$ intersect
   - No solution $\Rightarrow$ skew
3. If same point and direction vector $\Rightarrow$ identical

29.9 Normalization of a vector

To normalize a vector $\vec{a}$, we divide it by its length:

$$\hat{\vec{a}} = \frac{\vec{a}}{\|\vec{a}\|} = \frac{\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}}{\sqrt{a_x^2 + a_y^2 + a_z^2}} = \begin{bmatrix} \frac{a_x}{\|\vec{a}\|} \\ \frac{a_y}{\|\vec{a}\|} \\ \frac{a_z}{\|\vec{a}\|} \end{bmatrix}$$

29.10 Orthogonal Vectors and the Orthogonal Projection

Two vectors $\vec{a}$ and $\vec{b}$ are orthogonal if:

$$\langle \vec{a}, \vec{b} \rangle = 0$$

The orthogonal projection of vector $\vec{a}$ onto vector $\vec{b}$ is given by:

$$\mathrm{p}_{\vec{b}}(\vec{a}) = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{b}\|^2} \cdot \vec{b}$$

Proof of the projection formula

$$p\|q \implies \alpha q = p \;\forall \alpha \in \mathbb{K} \text{ and } \forall\, p, q \in V$$
$$\langle a - p, q \rangle = 0 \implies \langle a - \alpha q, q \rangle = 0 \implies \langle a, q \rangle - \alpha\langle q, q \rangle = 0$$
$$\alpha = \frac{\langle a, q \rangle}{\langle q, q \rangle}$$

Therefore, the orthogonal projection of vector $\vec{a}$ on $\vec{b}$ is given by multiplying $\vec{b}$ by the scalar α.
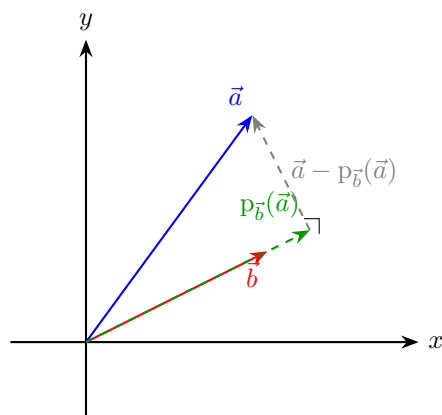
Figure 2: Orthogonal projection of $\vec{a}$ onto $\vec{b}$

29.11 The Cross Product

The cross product of two vectors $\vec{a}$ and $\vec{b}$ in $\mathbb{R}^3$ is defined as:

$$\vec{a} \times \vec{b} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

or

$$\det \left(\begin{bmatrix} \imath & v_1 & w_1 \\ \jmath & v_2 & w_2 \\ \hat{k} & v_3 & w_3 \end{bmatrix} \right) = \imath(v_2 w_3 - v_3 w_2) - \jmath(v_1 w_3 - v_3 w_1) + \hat{k}(v_1 w_2 - v_2 w_1)$$

The cross product has the following properties:

- $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$
- $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$
- $(\lambda_1 \vec{a}) \times (\lambda_2 \vec{b}) = \lambda_1 \lambda_2 (\vec{a} \times \vec{b})$
- $\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \|\vec{b}\| \sin(\theta)$
- $\langle \vec{a}, (\vec{b} \times \vec{c}) \rangle = 0$ (scalar triple product)
- $\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a} = \lambda \vec{b}$ for some $\lambda \in \mathbb{R}$ (parallel vectors)
- $\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a} = \vec{0} \vee \vec{b} = \vec{0}$ (zero vector)

The cross product is not defined in $\mathbb{R}^2$. The cross product is not commutative, but it is associative:

$$(\vec{a} \times \vec{b}) \times \vec{c} = \vec{a} \times (\vec{b} \times \vec{c})$$

The cross product is distributive over vector addition:

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$$

The cross product is anti-commutative:

$$\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$$

The cross product is not associative:

$$(\vec{a} \times \vec{b}) \times \vec{c} \neq \vec{a} \times (\vec{b} \times \vec{c})$$

The cross product is not distributive over scalar multiplication:

$$\lambda(\vec{a} \times \vec{b}) \neq (\lambda\vec{a}) \times \vec{b}$$

The length of the cross product in $\mathbb{R}^3$ is the area of the parallelogram spanned by the two vectors.

29.11.1 Determinant Trick

For two vectors $\vec{a}\ \vec{b}$ in $\mathbb{C}^3$ we can compute the cross product by:

$$\vec{n} = \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

29.11.2 Connection to the Volume of the Parallelepiped

We want to define a linear transformation f which relates the volume of the parallelepiped formed by three vectors, the determinant; and the dot product of a input vector of the transformation with some vector $\vec{p}$.

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \det\left(\begin{bmatrix} x & v_1 & w_1 \\ y & v_2 & w_2 \\ z & v_3 & w_3 \end{bmatrix}\right) = \left\langle \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}, \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\rangle$$

This gives us another version of the determinant trick view previously and we know that this gives us a vector perpendicular to $\vec{v}$ and $\vec{w}$. The motivation behind this weird formula is that if you compute both sides of the equation you will get the coordinates $\vec{p}$ from the determinant trick.

This whole equality can be related to the geometric meaning of the cross product. Remember that the dot product of two vectors can be interpreted as projecting one onto another and then multiplying the length of the projecting to the magnitude of the other. This is exactly the same as projecting our input vector to a line which is perpendicular to the parallelogram and then multiplying it by the area which will gives us the volume.

29.11.3 Derivation of the formula

We define $\vec{n}$ as a vector perpendicular to $\vec{v}$ and $\vec{w}$. Therefore,

$$\vec{v} \cdot \vec{n} = 0 \quad \text{and} \quad \vec{w} \cdot \vec{n} = 0$$

Writing this out in components, let $\vec{v} = (v_1, v_2, v_3)$, $\vec{w} = (w_1, w_2, w_3)$, and $\vec{n} = (n_1, n_2, n_3)$. Then:

$$v_1 n_1 + v_2 n_2 + v_3 n_3 = 0$$
$$w_1 n_1 + w_2 n_2 + w_3 n_3 = 0$$

Now we multiply the first equation by $-w_1$, the second by v_1 and add them

$$n_2(w_2 v_1 - v_2 w_1) + n_3(w_3 v_1 - v_3 w_1) = 0$$

We know that the equation $pc_1 + qc_2 = 0$ has the results $c_1 = -q$ and $c_2 = p$. Thus,

$$n_2 = v_3 w_1 + w_3 v_1$$

$$n_3 = v_1 w_2 - w_1 v_2$$

Finally, by substituting into the original equation and solving we get

$$n_1 = v_2 w_3 - v_3 w_2$$

Thus, we have gotten the formula for the cross product.

$$\mathfrak{QED}$$

29.12   Orthogonal vectors in $\mathbb{R}^2$ $\mathbb{R}^3$

29.12.1   Orthogonal vectors in $\mathbb{R}^2$

1. Interchange the components

2. Change the sign of one component

29.12.2   Orthogonal vectors in $\mathbb{R}^3$

1. Interchange two components

2. the one that was not changed, set to zero

3. Change the sign of first component

Example:

$$\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

29.13   Hessian Normal Form

The Hessian normal form is a way of expressing the equation of a plane in three-dimensional space using a normalized normal vector. It is particularly useful in computational geometry and physics, where signed distances from points to planes are important.

29.13.1   Geometric Interpretation

The Hessian normal form represents a plane by specifying:

- A unit normal vector $\vec{n} = (a, b, c)$ to the plane,

- and the shortest distance d from the origin to the plane.

This form is derived by normalizing the general plane equation. A plane in 3D can be written as:

$$ax + by + cz + d = 0,$$

where (a, b, c) is a normal vector to the plane and d is the dot product of the normal vector with a point p. If we divide all terms by $\sqrt{a^2 + b^2 + c^2}$, we normalize the normal vector:

$$\frac{a}{\sqrt{a^2 + b^2 + c^2}} x + \frac{b}{\sqrt{a^2 + b^2 + c^2}} y + \frac{c}{\sqrt{a^2 + b^2 + c^2}} z + \frac{d}{\sqrt{a^2 + b^2 + c^2}} = 0.$$

Let:

$$\vec{n} = \left(\frac{a}{\sqrt{a^2 + b^2 + c^2}}, \frac{b}{\sqrt{a^2 + b^2 + c^2}}, \frac{c}{\sqrt{a^2 + b^2 + c^2}} \right), \quad d' = \frac{d}{\sqrt{a^2 + b^2 + c^2}},$$

then the equation becomes:

$$\vec{n} \cdot \vec{r} + d' = 0,$$

which is the Hessian normal form. Here, $\vec{r} = (x, y, z)$ is any point on the plane, and $\vec{n}$ is the unit normal.

29.13.2 Signed Distance to the Plane

This form allows easy calculation of the signed distance from any point $\vec{p}$ to the plane:

$$\text{distance} = \vec{n} \cdot \vec{p} + d',$$

which is positive if $\vec{p}$ lies on the same side of the plane as the normal vector.

29.13.3 Derivation Illustration

To visualize the derivation, imagine a plane with unit normal vector $\vec{n}$, and a point P in space. The shortest distance from P to the plane is the projection of the vector $\vec{p} - \vec{q}$ onto $\vec{n}$, where $\vec{q}$ is any point on the plane. This leads to:

$$\text{distance} = (\vec{p} - \vec{q}) \cdot \vec{n}.$$

This gives the signed distance formula and thus, motivates the Hessian form.

29.14 Converting from the parametric form to the Hessian normal form

Steps:

1. Find the normal vector $\vec{n}$ of the plane
2. Normalize the normal vector
3. Find the distance d from the origin to the plane
4. Write the Hessian normal form

29.15 Converting from the Hessian normal form to the parametric form

Steps:

1. Find a point on the plane
2. Find two direction vectors in the plane
3. Write the parametric form

29.16 Properties of lines and planes

- Two planes are parallel if their normal vectors are scalar multiples of each other.
- A line and a plane are parallel if the direction vector of the line is orthogonal to the normal vector of the plane.
- A line intersects a plane if there exists a point on the line that satisfies the equation of the plane.
- Two planes intersect in a line if their normal vectors are not parallel.
- Three planes can intersect in a point, a line, or not at all.
- If we have a line G and a point on the line, for every vector $\vec{n}$ that is orthogonal to the direction vector of the line: $x \in G \iff \langle x, n \rangle$
- If p and q are two points in the line G with a normal vector then $\langle p, n \rangle = \langle q, n \rangle$
- Let E be a plane with the origin p and the direction vector $\vec{v}$ and $\vec{w}$, then there exist a normal vector and $x \in E \iff \langle x, n \rangle = \langle p, n \rangle$

29.17 Convert Normal Vector in Two Direction Vectors

Steps:

1. Given the normal vector $\vec{n} = (a, b, c)$, interchange a and b and multiply b by -1

2. Set the other component to 0. This gives you the first direction vector $\vec{v} = (-b, a, 0)$

3. Take the original normal vector $\vec{n}$ and interchange a and c and multiply c by -1

4. Set the other component to 0. This gives you the second direction vector $\vec{w} = (-c, 0, a)$

29.18 Intersection between Line and Plane

To find the intersection between a line and a plane, we can use the following steps:

1. Write the parametric form of the line: $\vec{r}(t) = \vec{P} + t\vec{v}$, where $\vec{P}$ is a point on the line and $\vec{v}$ is the direction vector.

2. Write the equation of the plane in Hessian normal form: $\langle \vec{n}, \vec{x} - \vec{P} \rangle = 0$, where $\vec{n}$ is the normal vector and $\vec{P}$ is a point on the plane.

3. Substitute the parametric form of the line into the equation of the plane.

4. Solve for t to find the intersection point.

5. Substitute t back into the parametric form of the line to find the intersection point.

Example:

Given the line:

$$g(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

and the plane:

$$E(u, m) = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + u \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + m \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

we want to find the intersection point.

Step 1: Determine the normal vector of the plane using the cross product of the two direction vectors:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{n} = \vec{v}_1 \times \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Step 2: Use the normal vector and a point on the plane to write the plane equation:

$$\langle \vec{n}, \vec{x} - \vec{Q} \rangle = 0, \quad \vec{Q} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Step 3: Plug the line into the plane equation:

$$\vec{x}(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \vec{x}(t) - \vec{Q} = \begin{bmatrix} t \\ t-1 \\ t-1 \end{bmatrix}$$

Now compute the dot product:

$$\langle \vec{n}, \vec{x}(t) - \vec{Q} \rangle = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} t \\ t-1 \\ t-1 \end{bmatrix} = -t + (t-1) + (t-1) = t - 2$$

Step 4: Solve for t:

$$t - 2 = 0 \Rightarrow t = 2$$

Step 5: Substitute $t = 2$ into the line:

$$\vec{x}(2) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

Result: The line intersects the plane at the point

$$\begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

Proof:

$$E : \langle x, n \rangle = \langle p, n \rangle$$
$$G : x = p + t \cdot v$$
$$\langle x, n \rangle = c$$
$$\langle p + t \cdot v, n \rangle = c$$
$$\langle p, n \rangle + t \cdot \langle v, n \rangle = c$$
$$t = \frac{c - \langle p, n \rangle}{\langle v, n \rangle}$$

$$\mathfrak{QED}$$

29.19   Distances between points, lines and planes

29.19.1   Distance between two points

The distance between two points $\vec{P_1}$ and $\vec{P_2}$ in $\mathbb{R}^n$ is given by:

$$d(\vec{P_1}, \vec{P_2}) = \|\vec{P_1} - \vec{P_2}\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + \cdots + (z_1 - z_2)^2}$$

where $\vec{P_1} = (x_1, y_1, \ldots, z_1)$ and $\vec{P_2} = (x_2, y_2, \ldots, z_2)$.

29.19.2   Distance between a point and a hyperplane

The distance between a point $\vec{P}$ and a hyperplane defined by the equation $\langle \vec{n}, \vec{x} - \vec{P_0} \rangle = 0$ is given by:

$$d(\vec{P}, \text{hyperplane}) = \frac{|\langle \vec{n}, \vec{P} - \vec{P_0} \rangle|}{\|\vec{n}\|}$$

where $\vec{P_0}$ is a point on the hyperplane and $\vec{n}$ is the normal vector of the hyperplane.

29.19.3 Distance between two lines

The distance between two lines in $\mathbb{R}^3$ can be calculated using the formula:

$$d = \frac{|\langle \vec{v_1} \times \vec{v_2}, \vec{P_2} - \vec{P_1}\rangle|}{\|\vec{v_1} \times \vec{v_2}\|}$$

where $\vec{P_1}$ and $\vec{P_2}$ are points on the two lines, and $\vec{v_1}$ and $\vec{v_2}$ are the direction vectors of the lines.

29.19.4 Distance between a point and a line

The distance between a point $\vec{P}$ and a line defined by the parametric equation $\vec{r}(t) = \vec{P_0} + t\vec{v}$ is given by:

$$d(\vec{P}, \text{line}) = \frac{\|\vec{v} \times (\vec{P} - \vec{P_0})\|}{\|\vec{v}\|}$$

where $\vec{P_0}$ is a point on the line and $\vec{v}$ is the direction vector of the line.

29.19.5 Distance between two planes

The distance between two parallel planes defined by the equations $\langle \vec{n}, \vec{x} - \vec{P_1}\rangle = 0$ and $\langle \vec{n}, \vec{x} - \vec{P_2}\rangle = 0$ is given by:

$$d = \frac{|\langle \vec{n}, \vec{P_2} - \vec{P_1}\rangle|}{\|\vec{n}\|}$$

where $\vec{P_1}$ and $\vec{P_2}$ are points on the two planes, and $\vec{n}$ is the normal vector of the planes.

29.19.6 Distance between a point and a plane

The distance between a point $\vec{P}$ and a plane defined by the equation $\langle \vec{n}, \vec{x} - \vec{P_0}\rangle = 0$ is given by:

$$d(\vec{P}, \text{plane}) = \frac{|\langle \vec{n}, \vec{P} - \vec{P_0}\rangle|}{\|\vec{n}\|}$$

where $\vec{P_0}$ is a point on the plane and $\vec{n}$ is the normal vector of the plane.

Example:

Distance Between Two Skew Lines

To find the shortest distance between two skew lines, we use the formula:

$$\text{distance} = \frac{|\langle (\vec{P_2} - \vec{P_1}), (\vec{v_1} \times \vec{v_2})\rangle|}{\|\vec{v_1} \times \vec{v_2}\|}$$

Where:

- $\vec{P_1}$ and $\vec{P_2}$ are points on each line,
- $\vec{v_1}$ and $\vec{v_2}$ are the direction vectors,
- $\vec{v_1} \times \vec{v_2}$ is the cross product of the direction vectors.

Given:

$$g_1 : \vec{r_1}(a) = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + a \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad g_2 : \vec{r_2}(b) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Step 1: Set

$$\vec{P}_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\vec{P}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Step 2: Compute the vector between base points:

$$\vec{P}_2 - \vec{P}_1 = \begin{bmatrix} 1 - 2 \\ 2 - 2 \\ 3 - 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Step 3: Compute the cross product:

$$\vec{v}_1 \times \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} (1)(1) - (1)(2) \\ (1)(3) - (0)(1) \\ (0)(2) - (1)(3) \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -3 \end{bmatrix}$$

Step 4: Compute scalar triple product:

$$\langle (\vec{P}_2 - \vec{P}_1), (\vec{v}_1 \times \vec{v}_2) \rangle = \langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -3 \end{bmatrix} \rangle = (-1)(-1) + (0)(3) + (1)(-3) = 1 + 0 - 3 = -2 \Rightarrow |\ldots| = 2$$

Step 5: Magnitude of the cross product:

$$\|\vec{v}_1\vec{v}_2\| = \sqrt{(-1)^2 + 3^2 + (-3)^2} = \sqrt{1 + 9 + 9} = \sqrt{19}$$

Final Answer:

$$\text{distance} = \frac{2}{\sqrt{19}} \approx 0.458$$

$$textShortestdistancebetweenthelinesis\frac{2}{\sqrt{19}}$$

29.20 Foot of the Perpendicular and Mirror Point

29.20.1 Foot of the Perpendicular

The *foot of the perpendicular* from a point $\vec{P}$ to a line (or plane) is the point on the line (or plane) where the perpendicular from $\vec{P}$ meets it.

Line case:

Given a line in parametric form:

$$g : \vec{r}(t) = \vec{A} + t\vec{v}$$

and a point $\vec{P}$ not on the line, the foot of the perpendicular $\vec{F}$ satisfies:

$$(\vec{P} - \vec{F}) \perp \vec{v} \quad \Rightarrow \quad (\vec{P} - (\vec{A} + t\vec{v})) \cdot \vec{v} = 0$$

Solve this inner product for t, then compute:

$$\vec{F} = \vec{A} + t\vec{v}$$

Plane case:

Given a plane in normal form:

$$\langle \vec{n}, \vec{x} - \vec{Q} \rangle = 0$$

then the foot of the perpendicular from point $\vec{P}$ to the plane is:

$$\vec{F} = \vec{P} - ((\vec{P} - \vec{Q}) \cdot \vec{n}) \cdot \vec{n}$$

29.20.2   Mirror Point

The *mirror point* (or reflected point) of $\vec{P}$ across a line or plane is the point $\vec{P}'$ such that the midpoint between $\vec{P}$ and $\vec{P}'$ is the foot of the perpendicular.

Formula:

$$\vec{P}' = 2\vec{F} - \vec{P}$$

Where $\vec{F}$ is the foot of the perpendicular from $\vec{P}$ to the line or plane.

30 Algebraic Structures

Algebraic structures are mathematical systems consisting of a set equipped with one or more operations that satisfy certain axioms. They provide a unified language to study various objects in mathematics, from numbers and matrices to functions and vector spaces. Understanding these structures is fundamental in abstract algebra and has applications in computer science, cryptography, coding theory, and physics.

30.1 Operations: Internal and External

An *internal composition law* is a binary operation that takes two elements from a set and returns another element in the same set. Formally, for a set S and operation $\circ$, we have:

$$\circ : S \times S \to S$$

An *external composition law* involves a second set acting on the structure, such as scalar multiplication in vector spaces:

$$\cdot : \mathbb{F} \times V \to V$$

where $\mathbb{F}$ is a field and V is a vector space.

30.2 Properties of Operations

Let $*$ be a binary operation on a set S. The most important properties include:

- *Associativity:* $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$
- *Commutativity:* $a * b = b * a$ for all $a, b \in S$
- *Identity Element:* There exists $e \in S$ such that $a * e = e * a = a$ for all $a \in S$
- *Inverse Element:* For every $a \in S$, there exists $a^{-1} \in S$ such that $a * a^{-1} = a^{-1} * a = e$
- *Distributivity:* $a \circ (b \bullet c) = (a \circ b) \bullet (a \circ c)$ and/or $(b \bullet c) \circ a = (b \circ a) \bullet (c \circ a)$

30.3 Homomorphisms and Isomorphisms

Let $(G, \oplus)$ and $(H, \oplus')$ be two algebraic structures.

A *homomorphism* is a function $\varphi : G \to H$ such that:

$$\varphi(a \oplus b) = \varphi(a) \oplus' \varphi(b), \quad \forall a, b \in G$$

An *isomorphism* is a bijective homomorphism. If such a map exists, we say the structures are *isomorphic*, written as $G \cong H$.

30.4 Common Algebraic Structures

30.4.1 Semigroup $(S, \oplus)$

A *semigroup* is a set S equipped with a binary operation $\oplus$ that is *associative*. This means:

$$(a \oplus b) \oplus c = a \oplus (b \oplus c) \quad \text{for all } a, b, c \in S.$$

There is no requirement for an identity element or inverses. Semigroups capture the essence of combining elements consistently (e.g., string concatenation).

30.4.2 Monoid $(M, \oplus)$

A *monoid* builds on a semigroup by adding an *identity element e* such that:

$$a \oplus e = e \oplus a = a \quad \text{for all } a \in M.$$

This structure is useful when an operation must have a "do nothing" element, like 0 for addition or 1 for multiplication.

30.4.3 Group $(G, \oplus)$

A *group* is a monoid where every element has an *inverse*:

$$\text{For each } a \in G, \text{ there exists } a^{-1} \in G \text{ such that } a \oplus a^{-1} = a^{-1} \oplus a = e.$$

Groups model reversible processes and are foundational in symmetry and abstract algebra.

30.4.4 Abelian Group $(A, \oplus)$

An *Abelian group* (or commutative group) is a group where the operation is *commutative*:

$$a \oplus b = b \oplus a \quad \text{for all } a, b \in A.$$

This property makes Abelian groups especially important theory and linear algebra.

30.4.5 Ring $(R, \oplus, \odot)$

A *ring* is a set with two operations:

- $(R, \oplus)$ is an Abelian group.
- $(R, \odot)$ is a semigroup (associative multiplication).
- Multiplication distributes over addition:

$$a \odot (b \oplus c) = a \odot b \oplus a \odot c.$$

Rings generalize arithmetic of integers, where addition and multiplication interact in a structured way.

30.4.6 Commutative Ring $(R, \oplus, \odot)$

A *commutative ring* is a ring where multiplication is also *commutative*:

$$a \odot b = b \odot a \quad \text{for all } a, b \in R.$$

This is the kind of ring most often encountered in basic algebra, such as the integers $\mathbb{Z}$.

30.4.7 Field $(\mathbb{F}, \oplus, \odot)$

A *field* is a commutative ring where every nonzero element has a *multiplicative inverse*:

$$(\mathbb{F} \setminus \{0\}, \odot) \text{ is an Abelian group.}$$

Fields like $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ allow division (except by zero), enabling the full range of arithmetic operations.

30.4.8   Vector Space $(V, \oplus, \cdot)$

A *vector space* over a field $\mathbb{F}$ consists of:

- An Abelian group $(V, \oplus)$ for vector addition.
- A scalar multiplication $\cdot : \mathbb{F} \times V \to V$ satisfying:
    - Distributivity over vector addition: $a \cdot (v_1 + v_2) = a \cdot v_1 + a \cdot v_2$
    - Distributivity over field addition: $(a + b) \cdot v = a \cdot v + b \cdot v$
    - Associativity: $a \cdot (b \cdot v) = (ab) \cdot v$
    - Identity: $1 \cdot v = v$

Vector spaces provide the foundation for linear algebra, where scalars from a field combine with vectors from a set.

31 Vector Spaces

A *vector space* is a set V with two operations, vector addition and scalar multiplication, such that:

I. The set V is closed under vector addition.

II. The set V is closed under scalar multiplication.

III. Vector addition is commutative.

IV. Vector addition is associative.

V. There exists a zero vector $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$.

VI. For every vector $\vec{v} \in V$, there exists a vector $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = \vec{0}$.

VII. Scalar multiplication is distributive with respect to vector addition: $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ for all $a \in F$ and $\vec{u}, \vec{v} \in V$.

VIII. Scalar multiplication is distributive with respect to field addition: $(a+b)\vec{v} = a\vec{v} + b\vec{v}$ for all $a, b \in F$ and $\vec{v} \in V$.

IX. Scalar multiplication is associative: $a(b\vec{v}) = (ab)\vec{v}$ for all $a, b \in F$ and $\vec{v} \in V$.

X. The multiplicative identity acts as a scalar: $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.

The set F is a field, and the elements of V are called *vectors*.

Examples:

1. The set of all n-tuples of real numbers $\mathbb{R}^n$ is a vector space over the field of real numbers $\mathbb{R}$.

2. The set of all polynomials of degree less than or equal to n is a vector space over the field of real numbers $\mathbb{R}$.

3. The set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space 7 over the field of real numbers $\mathbb{R}$.

4. The set of all $m \times n$ matrices with real entries is a vector space over the field of real numbers $\mathbb{R}$.

31.1 Subspaces

A subset W of a vector space V is a *subspace* of V if:

I. The zero vector $\vec{0} \in W$.

II. For all $\vec{u}, \vec{v} \in W$, $\vec{u} + \vec{v} \in W$.

III. For all $a \in F$ and $\vec{v} \in W$, $a\vec{v} \in W$.

If W is a subspace of V, we write $W \subseteq V$.

The intersection of two subspaces is also a subspace.

31.2 Linear Combinations

A *linear combination* of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in a vector space V is an expression of the form:

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$$

Where $a_1, a_2, \ldots, a_n$ are scalars from the field F. The set of all linear combinations of a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is called the **span** of those vectors, denoted by $\text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$. The span of a set of vectors is a subspace of the vector space V.

The span of a set of vectors is the smallest subspace containing those vectors.

31.3   Properties of the subspaces

- The intersection of two subspaces is a subspace.

- The union of two subspaces is not necessarily a subspace.

- The sum of two subspaces U and W is defined as:

$$U + W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

  The sum of two subspaces is a subspace.

The sum of two subspaces is the smallest subspace containing both subspaces.

- The direct sum of two subspaces U and W is defined as:

$$U \oplus W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

  The direct sum of two subspaces is a subspace.

- The direct sum of two subspaces is the smallest subspace containing both subspaces, such that $U \cap W = \{\vec{0}\}$.

- The direct sum of two subspaces is denoted by $U \oplus W$.

31.4   Linear Independence

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ in a vector space V is said to be *linearly independent* if the only solution to the equation:

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_n \vec{v}_n = 0$$

or

$$\sum_{i=1}^{n} \lambda_i \vec{v}_i = 0$$

is $a_1 = a_2 = \cdots = a_n = 0$. If there exists a non-trivial solution to this equation, then the set of vectors is said to be linearly dependent. A set of vectors is linearly independent if and only if the only linear combination of those vectors that equals the zero vector is the trivial combination where all coefficients are zero.

31.4.1   Properties of the linear independence

- A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent if and only if the only linear combination of those vectors that equals the zero vector is the trivial combination where all coefficients are zero.

- If a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent, then any subset of that set is also linearly independent.

- If a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly dependent, then at least one vector in that set can be expressed as a linear combination of the others.

31.5   Base

- A *base* of a vector space V is a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ that is linearly independent and spans the vector space V.

- The number of vectors in a base of a vector space is called the *dimension* of the vector space.

- The dimension of a vector space V is denoted by $\dim(V)$.

- If V has a finite base, then it is said to be *finite-dimensional*.

- If V does not have a finite base, then it is said to be *infinite-dimensional*.

31.6 Dimension

The *dimension* of a vector space V is the number of vectors in a base of V.

- The dimension of a vector space is denoted by $\dim(V)$.
- The dimension of a vector space can be finite or infinite.
- If the dimension of a vector space is finite, then it is said to be *finite-dimensional*.
- If the dimension of a vector space is infinite, then it is said to be *infinite-dimensional*.

31.6.1 How to find the base of a set vector

To find the base of a set of vectors, we can use the following steps:

1. Write the vectors as columns of a matrix.
2. Row-reduce the matrix to echelon form.
3. The non-zero rows of the echelon form matrix correspond to the base of the vector space spanned by the original set of vectors.

The number of non-zero rows in the echelon form matrix is equal to the dimension of the vector space spanned by the original set of vectors.

The base of a vector space is not unique. Different bases can span the same vector space.

31.7 Basis Extension Theorem

Let V be a vector space over a field K, and let

$$v_1, \ldots, v_r, \quad w_1, \ldots, w_s \in V.$$

Suppose that $(v_1, \ldots, v_r)$ is a linearly independent tuple and that

$$\text{span}(v_1, \ldots, v_r, w_1, \ldots, w_s) = V.$$

Then it is possible to extend $(v_1, \ldots, v_r)$ to a basis of V by possibly adding suitable vectors from the set $\{w_1, \ldots, w_s\}$.

Proof:

If $\text{span}(v_1, \ldots, v_r) = V$, the statement is obvious. So assume

$$\text{span}(v_1, \ldots, v_r) \neq V.$$

Then there exists at least one w_i such that $w_i \notin \text{span}(v_1, \ldots, v_r)$; otherwise, if all $w_i \in \text{span}(v_1, \ldots, v_r)$, then

$$\text{span}(v_1, \ldots, v_r, w_1, \ldots, w_s) = \text{span}(v_1, \ldots, v_r) = V,$$

which contradicts our assumption that $\text{span}(v_1, \ldots, v_r) \neq V$.

The tuple $(w_i, v_1, \ldots, v_r)$ is linearly independent, because of

$$\sum_{j=1}^{r} \lambda_j v_j + \lambda w_i = 0$$

It follows that $\lambda = 0$ (since $w_i \notin \text{span}(v_1, \ldots, v_r)$), and then also $\lambda_j = 0$ for all j because the v_j are linearly independent.

Possibly, $(w_i, v_1, \ldots, v_r)$ is still not a basis of V. Then we repeat the previous step and keep adding further w_i until the tuple extends $(v_1, \ldots, v_r)$ to a basis of V. This process terminates after finitely many steps, since

$$\text{span}(v_1, \ldots, v_r, w_1, \ldots, w_s) = V.$$

<div align="right">𝔔𝔈𝔇</div>

Every finitely generated vector space V has a basis.

31.8 Exchange Lemma

Let $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ be bases of a vector space V. Then, for every v_i, there exists a w_j such that if we replace v_i by w_j in the tuple $(v_1, \ldots, v_n)$, it still forms a basis of V.

Proof:

Let $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ be two bases of V. Suppose we remove v_i from the first basis. The truncated tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ satisfies

$$\text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \neq V,$$

because if $\text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) = V$, then v_i would lie in the span of the remaining vectors and could be written as a linear combination of them. This would contradict the assumption that $(v_1, \ldots, v_n)$ is linearly independent and a basis of V.

By the Basis Extension Theorem, we can extend the truncated tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ to a basis of V by adding vectors from $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_1, \ldots, w_m)$. Therefore, by the Basis Extension Theorem, there exists a w_j such that

$$w_j \notin \text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n),$$

and the tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_j)$ is linearly independent.

If this tuple does not form a basis, we can again apply the Basis Extension Theorem and add one of the vectors $v_1, \ldots, v_n$ to complete the basis. Clearly, the only possibility is to add v_i, but this would imply that the tuple $(v_1, \ldots, v_n, w_j)$ is not a basis, as w_j would then be linearly dependent on the other vectors. Therefore, $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_j)$ must form a basis of V.

<div align="right">𝔔𝔈𝔇</div>

31.9 Dimension of a sum of subspaces

Let U and W be two subspaces of a vector space V. Then the dimension of the sum of the two subspaces is given by:

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

31.10 Linear Independence of polynomials

Let P_n be the vector space of polynomials of degree at most n. The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is a basis for P_n. The dimension of P_n is $n + 1$.

- The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.
- The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ spans the vector space P_n.
- The dimension of P_n is $n + 1$.

To prove that the set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent, we can use the following steps:

1. Assume that there exists a linear combination of the polynomials that equals zero:

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0,$$

where $a_0, a_1, \ldots, a_n$ are scalars.

2. Since the left-hand side is a polynomial of degree at most n, it can only be equal to zero if all coefficients are zero.

3. Therefore, we have $a_0 = a_1 = \cdots = a_n = 0$, which proves that the set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.

So you only have to prove that the set of coefficients vector is linearly independent.

31.11 Interpolation Polynomial

Given the $n + 1$ points (x_k, y_k), with $0 \leq k \leq n$ and all x_k distinct, there exists exactly one polynomial $p_n \in P_n$ such that $y_k = p_n(x_k)$ for all $0 \leq k \leq n$. This polynomial is called the interpolation polynomial.

Proof:

The uniqueness follows immediately. We prove the existence by induction on n. For $n = 0$, choose $p_0(x) = y_0$.

Now assume the statement is true for $n-1$. Let the polynomial p_{n-1} interpolate the points $(x_0, y_0), \ldots, (x_{n-1}, y_{n-1})$.

Define

$$p_n(x) = p_{n-1}(x) + q(x),$$

where

$$q(x) = \frac{(x - x_0)(x - x_1) \cdots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \cdots (x_n - x_{n-1})}(y_n - p_{n-1}(x_n)).$$

We have $q \in P_n$, and it follows that $p_n \in P_n$. Furthermore, $q(x_k) = 0$ for $k \leq n - 1$ because a linear factor in the numerator always vanishes at x_k. Therefore, $p_n(x_k) = y_k$ for $k \leq n - 1$. Additionally, we have

$$q(x_n) = y_n - p_{n-1}(x_n),$$

so that $p_n(x_n) = y_n$.

$$\mathfrak{QED}$$

Example:

Consider the three points $(-2, 1)$, $(-1, -1)$, and $(1, 1)$. These points uniquely define an interpolating parabola p_2. This parabola can be determined using the definition of p_n. For hand calculations and a few points to interpolate, the following approach is also useful. The general form of the polynomial is

$$p_2(x) = ax^2 + bx + c.$$

Substituting the three points into this form gives the system of equations:

$$1 = a + b + c \quad \text{(from the point (1, 1))}$$
$$-1 = a - b + c \quad \text{(from the point (-1, -1))}$$
$$1 = 4a - 2b + c \quad \text{(from the point (-2, 1))}$$

This leads to the system of equations:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 4 & -2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Solving this system gives $a = 1$, $b = 1$, and $c = -1$, so the interpolation polynomial is

$$p_2(x) = x^2 + x - 1.$$

32 Dot Product, Euclidean and Unitary Space

32.1 Inner Product

Let V be a vector space over a field K. A mapping $\langle \cdot, \cdot \rangle : V \times V \to K$ is called an inner product (or inner product) if the following conditions are satisfied:

Symmetry

For all $a, b \in V$:

$$\langle a, b \rangle = \begin{cases} \langle b, a \rangle & \text{if } K = \mathbb{R}, \\ \overline{\langle b, a \rangle} & \text{if } K = \mathbb{C}. \end{cases}$$

Linearity in the First Argument

For all $a, b, c \in V$:

$$\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$$

and

$$\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle.$$

Homogeneity in the First Argument

For all $\alpha \in K$, we have:

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle = \begin{cases} \langle a, \alpha b \rangle & \text{if } K = \mathbb{R}, \\ \langle a, \overline{\alpha} b \rangle & \text{if } K = \mathbb{C}. \end{cases}$$

Positive Definiteness

For all $a \in V \setminus \{0\}$:

$$\langle a, a \rangle > 0,$$

and

$$\langle 0, 0 \rangle = 0.$$

It also can be interpreted as a linear transformation that takes a vector and maps it to a real number via matrix vector multiplication. Also, a more practical way of thinking about the *dot product* is the question: How much are two vectors pointing in the same direction?

32.2 Trigonometric Definition

Think about the orthogonal projection as a triangle. Now, remember that to find the adjacent side x you need to take $x = \cos(\theta)h$, which is the value α in our projection formula $p_a(b) = \frac{\langle a, b \rangle}{\langle a, a \rangle} \alpha a$. Here we add $\|a\|\|b\|$ to make this function linear.

$$\langle a, b \rangle = \|a\|\|b\| \cos(\theta)$$

32.3   Standard Inner Product for Complex Number

Let $a = (a_i)_{i=1}^n$ and $b = (b_i)_{i=1}^n$ be vectors in $\mathbb{C}^n$. The standard inner product is defined by

$$\langle a, b \rangle := \sum_{i=1}^n a_i \overline{b_i} = a^T \overline{b}$$

and by the *hermitian inner product*

$$\langle a, b \rangle := \overline{a^T} b$$

Here the *symmetry* is given by $\langle a, b \rangle = \overline{\langle b, a \rangle}$.

32.4   Inner Product on $C[a, b]$

Let $f, g \in C[a, b]$. The inner product on $C[a, b]$ is defined by

$$\langle f, g \rangle := \int_a^b f(x) \cdot g(x) \, dx.$$

32.5   Euclidean and Unitary Vector Spaces

A real vector space equipped with an inner product is called a *Euclidean vector space*, while a complex vector space with an inner product is called a *unitary vector space*.

32.6   Norms in Vector Spaces

Let V be a K-vector space and $a, b \in V$. A function $\| \cdot \| : V \to \mathbb{R}$ is called a norm if and only if the following conditions hold:

- $\|a\| \in \mathbb{R}$,
- $\|a\| \geq 0$,
- $\|a\| = 0 \iff a = 0$,
- $\forall \lambda \in K, \ \|\lambda a\| = |\lambda| \|a\|$,
- $\|a + b\| \leq \|a\| + \|b\|$.

32.6.1   Induced Norm by a Inner Product

As in the special case $V = \mathbb{R}^n$, an inner product induces a norm. In a unitary (or Euclidean) space, the inner product induces a (standard) norm defined by

$$\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}.$$

32.7   Cauchy-Schwarz Inequality in Unitary Vector Spaces

In all unitary vector spaces V, the Cauchy-Schwarz inequality holds:

$$|\langle a, b \rangle| \leq \|a\| \|b\| \quad \forall a, b \in V.$$

32.7.1 Proof of the Triangle Inequality

Both sides of the triangle inequality are real and, in particular, non-negative. Therefore, it is sufficient to prove that the squares of both sides satisfy the desired inequality, i.e., we need to show:

$$\langle a+b, a+b \rangle \leq (\|a\| + \|b\|)^2.$$

First, we expand the left-hand side:

$$\langle a+b, a+b \rangle = \langle a,a \rangle + \langle a,b \rangle + \langle b,a \rangle + \langle b,b \rangle.$$

Since $\langle b,a \rangle = \langle a,b \rangle$, we have:

$$\langle a,b \rangle + \langle b,a \rangle = 2\operatorname{Re}\langle a,b \rangle.$$

Now, we know that the absolute value of a complex number is always greater than or equal to its real part, so:

$$2\operatorname{Re}\langle a,b \rangle \leq 2|\langle a,b \rangle|.$$

Using the Cauchy-Schwarz inequality, we can further bound this by:

$$2\operatorname{Re}\langle a,b \rangle \leq 2\|a\|\|b\|.$$

Thus, we have:

$$\langle a+b, a+b \rangle \leq \langle a,a \rangle + 2\|a\|\|b\| + \langle b,b \rangle.$$

Using the definition of the norm, $\|a\|^2 = \langle a,a \rangle$ and $\|b\|^2 = \langle b,b \rangle$, we obtain:

$$\langle a+b, a+b \rangle \leq \|a\|^2 + 2\|a\|\|b\| + \|b\|^2.$$

This is exactly the expansion of $(\|a\| + \|b\|)^2$, which completes the proof.

$$\mathfrak{QED}$$

32.8 Definiteness of Inner Products

The concept of the *inner product* plays a major rule in linear algebra, and there are various manifestations of this element. We are going to see how inner products are constructed based on the properties we are interested in. Those being linearity for the columns, symmetry and positive definiteness.

Let us start with a matrix $A \in \mathbb{R}^{n \times n}$ and a map $(\cdot,\cdot)_A \mathbb{R}^n \times \mathbb{R}^n :\to \mathbb{R}$ more specific $(x,y)_A := \langle x, Ay \rangle$ using the standard Euclidean inner product.

For our first property, the linearity; it is granted for every matrix $A = (a_1, a_2, \ldots, a_n)$ because of the rules of matrix-vector multiplication. For example for the basis vectors e_i, e_j we get

$$(e_i, e_j)_A = \langle e_i, Ae_j \rangle = e_i^T A e_j = e_i^T a_j = a_{ij}$$

Therefore, $(e_i, e_j)_A = (e_j, e_i)_A \iff a_{ij} = a_{ji}$. This means that A has to be symmetric. As a side note, the choice for the Euclidean inner product for the structure of our map is optional. Any kind of inner product would also do the job. This implies that each inner product can be described via a matrix.

32.9 Definiteness and Quadratic Forms

Given a symmetric $A \in \mathbb{R}^{n \times n}$ then:

1. The map $x \to \langle x, Ax \rangle$ is called the *Quadratic Form*

2. A is *positive definite* if $\langle x, Ax \rangle > 0 \forall x \in \mathbb{R}^n \setminus \{0\}$

3. A is *negative definite* if $\langle x, Ax \rangle < 0 \forall x \in \mathbb{R}^n \setminus \{0\}$

4. A is *positive semidefinite* if $\langle x, Ax \rangle \geq 0 \forall x \in \mathbb{R}^n \setminus \{0\}$

5. A is *negative semidefinite* if $\langle x, Ax \rangle \leq 0 \forall x \in \mathbb{R}^n \setminus \{0\}$

6. A is *indefinite* if $\exists x, y \in \mathbb{R}^n \setminus \{0\} : \langle x, Ax \rangle < 0 \wedge \langle y, Ay \rangle > 0$

$(\cdot, \cdot)_A$ is only an inner product if A is symmetric and positive definite.

Example:

Given is the term $x_1^2 - x_3^2 + x_1 x_4$, show that it is a quadratic form.

First write the formula for the quadratic form $\langle x, Ax \rangle$ where A is real symmetric matrix

$$
\left\langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & q \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \right\rangle
$$

$$
\left\langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \begin{bmatrix} ax_1 + bx_2 + cx_3 + dx_4 \\ ex_1 + fx_2 + gx_3 + hx_4 \\ ix_1 + jx_2 + kx_3 + lx_4 \\ mx_1 + nx_2 + ox_3 + px_4 \end{bmatrix} \right\rangle
$$

$$x_1(ax_1+bx_2+cx_3+dx_4)+x_2(ex_1+fx_2+gx_3+hx_4)+x_3(ix_1+jx_2+kx_3+lx_4)+x_4(mx_1+nx_2+ox_3+px_4)$$

Now we have to compare the coefficients with the original expression $x_1^2 + (0)x_2^2 - x_3^2 + x_1 x_4$.

$$
\begin{aligned}
& ax_1^2 + bx_1 x_2 + cx_1 x_3 + dx_4 x_4 + \\
& ex_2 x_1 + fx_2^2 + gx_2 x_3 + hx_2 x_4 + \\
& ix_3 x_1 + jx_3 x_2 + kx_3^2 + lx_3 x_4) + \\
& mx_4 x_1 + nx_4 x_2 + ox_4 x_3 + px_4^2
\end{aligned}
$$

Note that the only quadratic terms present in the expression are x_1^2 and $-x_3^2$ while the others are zero. This means that they have to disappear in the main diagonal. Also, the only term that x_1 is also multiplying is x_4 therefore all the others have to be zero like all other non-quadratic terms which are not present. Thus

$$
\begin{bmatrix}
1 & 0 & 0 & d \\
0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 \\
m & 0 & 0 & 0
\end{bmatrix}
$$

Now the question is what are d and m because A has to be symmetric, but they can not be 1 because then we would have $2x_1 x_4$. Because of that, we will take the only available option $\frac{1}{2}$.

$$
\begin{bmatrix}
1 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 \\
\frac{1}{2} & 0 & 0 & 0
\end{bmatrix}
$$

And now we are done.

32.10   Minors

For a matrix $A \in \mathbb{R}^{n \times n}$ the starting from the left $k \times k$-submatrices $A_k = (a_{ij})_{i,j=1}^{k}$ of A. The $D_k = \det(A_k)$ is called a *main minor* or *minor* of A.

32.10.1   Minors and Definiteness

Let A be real, square symmetric matrix, then

- A is positive definite if all main minors are positive.

- A is negative definite if all main minors are negative.

- A is positive semidefinite if all main minors are greater or equal to zero but not all zero.

- A is negative semidefinite if all main minors are less or equal to zero but not all zero.

- A is indefinite if all main minors are equal to zero.

32.11   Connection to Symmetric Matrices

For a real symmetric matrix A the following statements are equivalent.

- A is positive definite.

- All eigenvalues of A are positive.

- All main minors of A are positive.

Continuing, A is positive semidefinite, if all eigenvalues of A are not negative and the main minors are all not positive.

Example: Symmetry and Positive Definiteness of BAB^T and $B^T AB$

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix and $B \in \mathbb{R}^{m \times n}$ be a real matrix. We study the symmetry and positive definiteness of the matrices $C = BAB^T \in \mathbb{R}^{m \times m}$ and $D = B^T AB \in \mathbb{R}^{n \times n}$.

(i) Symmetry of BAB^T:

$$(BAB^T)^T = (B^T)^T A^T B^T = BAB^T,$$

since $(B^T)^T = B$ and $A^T = A$. Thus, BAB^T is symmetric.

(ii) Symmetry of $B^T AB$:

$$(B^T AB)^T = B^T A^T B = B^T AB,$$

since A is symmetric. Hence, $B^T AB$ is symmetric as well.

To determine when BAB^T and $B^T AB$ are positive definite, we use the quadratic form criterion.

(i) Positive definiteness of $C = BAB^T$:

Let $x \in \mathbb{R}^m$, $x \neq 0$. Consider:

$$x^T C x = x^T BAB^T x.$$

Let $y = B^T x \in \mathbb{R}^n$. Then:

$$x^T BAB^T x = y^T Ay.$$

Since A is positive definite, $y^T Ay > 0$ for all $y \neq 0$. Therefore, $x^T C x > 0$ if and only if $B^T x \neq 0$ for all $x \neq 0$, which holds if and only if $\ker(B^T) = \{0\}$, or equivalently, B has full column rank.

(ii) Positive definiteness of $D = B^T AB$:

Let $z \in \mathbb{R}^n$, $z \neq 0$. Then:

$$z^T D z = z^T B^T A B z = (Bz)^T A (Bz).$$

Let $w = Bz$. Since A is positive definite, $w^T A w > 0$ if $w \neq 0$. Thus, D is positive definite if and only if $Bz \neq 0$ for all $z \neq 0$, i.e., if B has full rank (i.e., $\ker(B) = \{0\}$).

In summary:

- BAB^T is positive definite if B has full column rank.
- $B^T AB$ is positive definite if B has full rank.

$$\mathfrak{QED}$$

32.12   Sign of the D_k

For the main minors D_k of a negative definite matrix it applies that

$$(-1)^k D_k = (-1)^k \det(A_k) = \det(-A_k) > 0,$$

because $-A$ is positive definite.

32.13   Negative Quadratic Form

If A is negative definite then

$$-\langle x, Ax \rangle = \langle x - Ax \rangle > 0$$

32.14   Theorems for Non-squares matrices

32.14.1   Theorem I

Given $m \geq n$ and $A \in \mathbb{R}^{m \times n}$ with $rg(A) = n$. Then $A^T A$ is positive definite. For $rg(A) < n$ is $A^T A$ positive semidefinite.

Proof:

Because of $(A^T A)^T = A^T A$ is $A^T A$ symmetric. Therefore, it applies that

$$\langle x, A^T Ax \rangle = x^T A^T Ax = (Ax)^T Ax = \langle Ax, Ax \rangle = \|Ax\|^2 \geq 0$$

Also, $A^T A$ is positive semidefinite, and it is also invertible for $rg(A) = n$. If there were an $x \neq 0$ such that $\|Ax\| = 0$ this would imply that $Ax = 0$, also $A^T(Ax) = 0$ therefore $x \in ker(A^T A)$. And $A^T A$ would not be invertible. Finally, $\langle x, A^T Ax \rangle > 0 \forall x \neq 0$, and $A^T A$ is symmetric and positive definite.

32.15   Theorem II

Given $m \leq n$ and $A \in \mathbb{R}^{m \times n}$. Then $A^T A$ is positive semidefinite and even positive definite if $rg(A) = m$.

33 Orthogonality

33.1 Orthogonality and Projection

Let $a, b \in V$. The vectors a and b are orthogonal to each other if

$$\langle a, b \rangle = 0.$$

This is written as $a \perp b$.

With the same proof as in Theorem 2.25, the Pythagorean Theorem holds

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2,$$

for $a \perp b$ in all unitary vector spaces.

For the orthogonal projection $p_b(a)$ of a vector a onto b, with $b \neq 0$, the formula in every unitary vector space is:

$$p_b(a) = \frac{\langle a, b \rangle}{\langle b, b \rangle} b.$$

33.2 Orthogonal Projection and Orthogonal Complement

Let U be a finitely generated subspace of V and $a \in V$. A vector $p_U(a) \in U$ is called the orthogonal projection of a onto U if

$$a - p_U(a) \perp u \quad \forall u \in U$$

The question arises about well-definedness, i.e., whether such a vector $p_U(a)$ always exists and if it is unique. The following concept helps in the discussion of uniqueness.

For $M \subseteq V$, the *orthogonal complement* of M is defined as

$$M^{\perp} = \{ v \in V \mid v \perp u \, \forall u \in M \}.$$

- $M^{\perp}$ is a subspace of V.
- Let U be a subspace of V. Then, we have $U \cap U^{\perp} = \{0\}$.

Proof:

We need to check the closure property. For $u \in M$, $x, y \in M^{\perp}$, and $\lambda \in \mathbb{R}$, we have:

$$\langle x + y, u \rangle = \langle x, u \rangle + \langle y, u \rangle = 0$$
$$\langle \lambda x, u \rangle = \lambda \langle x, u \rangle = 0.$$

Let $a \in U \cap U^{\perp}$. Then, we have $\langle a, u \rangle = 0$ for all $u \in U$, because $a \in U^{\perp}$, and in particular, $\langle a, a \rangle = 0$ since $a \in U$, which implies $a = 0$.

33.3 Orthogonality of the basis to the Complement

Let U be as before, and let $(u_1, \ldots, u_m)$ be a basis of U. For $v \in V$, we have:

$$v \in U^{\perp} \quad \text{if and only if} \quad \langle v, u_i \rangle = 0 \quad \forall 1 \leq i \leq m.$$

33.4 How to find the orthogonal complement

Let U be a finitely generated subspace of V with basis $(u_1, \ldots, u_m)$. To find the orthogonal complement $U^\perp$, we can solve the system of equations:

$$\langle v, u_i \rangle = 0 \quad \forall 1 \le i \le m.$$

This system can be expressed in matrix form as $A \cdot v = 0$, where A is the matrix whose rows are the vectors u_i and v is the vector we want to find in $U^\perp$.

33.5 Orthogonal and Orthonormal Systems

Let $B = (v_1, \ldots, v_m)$ be an m-tuple of vectors in $V \setminus \{0\}$.

- B is called an orthogonal system in V if all the vectors v_i are pairwise orthogonal.

- An orthogonal system is called an orthonormal system if, in addition, $\|v_i\| = 1$ for all $i = 1, \ldots, m$.

- An orthogonal system that forms a basis of V is called an orthogonal basis of V.

- An orthonormal system that forms a basis of V is called an orthonormal basis of V.

Notation:

$$\text{OG-System} \quad \text{OG-Basis} \quad \text{ON-System} \quad \text{ON-Basis}$$

Using the Kronecker delta symbol $\delta_{i,j}$:

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \ne j, \end{cases}$$

we have $\langle v_i, v_j \rangle = \delta_{i,j}$ for any orthonormal system.

33.6 Writing a vector in terms of the Orthogonal Basis

Example:

The vectors

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

are orthogonal and normalized, but do not form a basis. Thus, they constitute an *orthonormal system.*

Example:

The vectors

$$a_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis, are orthogonal and normalized. Therefore, they form an *orthonormal basis.*

33.7 Linear Independence of Orthogonal Systems

An orthogonal system $(v_1, \ldots, v_n)$ is linearly independent.

Proof:

Let $\sum_{i=1}^{n} \lambda_i v_i = 0$. Then, for any v_j, it follows that

$$\left\langle \sum_{i=1}^{n} \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^{n} \lambda_i \langle v_i, v_j \rangle = 0.$$

Due to the orthogonality of the system, all terms vanish except $\lambda_j \langle v_j, v_j \rangle$, so

$$\lambda_j \langle v_j, v_j \rangle = 0.$$

Since $v_j \neq 0$ in any orthogonal system, we have $\langle v_j, v_j \rangle = \|v_j\|^2 > 0$, implying $\lambda_j = 0$ for all $1 \leq j \leq n$.

$$\mathfrak{QED}$$

33.8 Representation with Respect to an Orthogonal Basis

Let $B = (v_1, \ldots, v_n)$ be an orthogonal basis of a vector space V. Then for every $v \in V$, we have:

$$v = \sum_{k=1}^{n} \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k,$$

that is, the coordinates of v with respect to the basis B are given by

$$\left(\frac{\langle v, v_k \rangle}{\|v_k\|^2} \right)^T_{1 \leq k \leq n}.$$

33.9 Orthogonal Projection and Direct Sum Decomposition

Let $B = (v_1, \ldots, v_m)$ be an orthogonal system in V, and let $U = L(B)$ be the subspace of V spanned by B.

- For every $v \in V$, the orthogonal projection of v onto U is given by:

$$p_U(v) = \sum_{i=1}^{m} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

- Every vector $v \in V$ can be uniquely written as the sum $v = p_U(v) + w$, with $w \in U^\perp$. In this case, we have:

$$w = v - p_U(v).$$

- $V = U \oplus U^\perp$, which means that every vector in V can be uniquely decomposed into a sum of a vector from U and a vector from $U^\perp$.

- If $\dim(V) = n$, then:

$$\dim(U) + \dim(U^\perp) = n \quad \text{for every subspace } U.$$

33.10 The Gram-Schmidt Process

The Gram-Schmidt process is a method for orthonormalizing a set of vectors in an inner product space. Given a finite set of linearly independent vectors $(v_1, \ldots, v_n)$, the process generates an orthonormal basis $(u_1, \ldots, u_n)$ as follows:

1. Set $u_1 = \frac{v_1}{\|v_1\|}$.

2. For $k = 2, \ldots, n$:

   (a) Set $w_k = v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j$.

   (b) Set $u_k = \frac{w_k}{\|w_k\|}$.

3. The resulting set $(u_1, \ldots, u_n)$ is an orthonormal basis of the subspace spanned by $(v_1, \ldots, v_n)$.

The Gram-Schmidt process can be applied to any finite set of linearly independent vectors in an inner product space, and it is particularly useful for constructing orthonormal bases in Euclidean spaces.

Proof:

Let W be a non-zero finite-dimensional subspace of an inner product space and let $(u_1, \ldots, u_n)$ be a basis for W. We will show that an *orthonormal basis* exists.

Let $v_1 = u_1$, then compute the vector u_2 orthogonal to the space W_1 spanned by v_1

$$v_2 = u_2 - proj_{w_1} u_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|} u_2$$

Then normalize the vector.

We can continue this approach to of projecting the vector of u_i orthogonal to the space W_i spanned by all basis previously computed v_i basis vectors and then normalize it until we get our orthonormal basis. Note, that because our previous vectors formed a basis we can be sure that during this process no zero vector will be generated because then we will have some u_i that is a linear combination of the other vectors. Thus, this would be a contradiction.

$$\mathfrak{QED}$$

Example:

Let $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$, and $v_3 = (1, 1, 1)$. The Gram-Schmidt process yields:

$$u_1 = \frac{v_1}{\|v_1\|} = (1, 0, 0), \quad u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\|v_2 - \langle v_2, u_1 \rangle u_1\|} = (0, 1, 0),$$

$$u_3 = \frac{v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2}{\|v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2\|} = (0, 0, 1).$$

Thus, the orthonormal basis is $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.

33.11 Existence of Orthonormal Bases and Orthogonal Projections

- Every finitely generated unitary vector space has an orthonormal basis.

- This is in general false for vector spaces that are not finitely generated.

- We now provide the existence proof of the orthogonal projection.

- Let V be a unitary vector space and U a finitely generated subspace. Then for every $v \in V$, the orthogonal projection $p_U(v)$ of v onto U exists.

- Let V be a finitely generated unitary vector space and U any subspace. Then we have $V = U \oplus U^{\perp}$, and

$$\dim(V) = \dim(U) + \dim(U^{\perp}).$$

- Every hyperplane in $\mathbb{R}^n$ admits a normal form; the normal vector is unique up to scalar multiplication.

- Let V be as above and let $v_1, \ldots, v_m \in V$. If it is possible to construct orthonormal vectors $w_1, \ldots, w_m$ from them using the Gram-Schmidt process, then $(v_1, \ldots, v_m)$ are linearly independent.

33.11.1 Visualization of the Gram-Schmidt process in 2 dimension
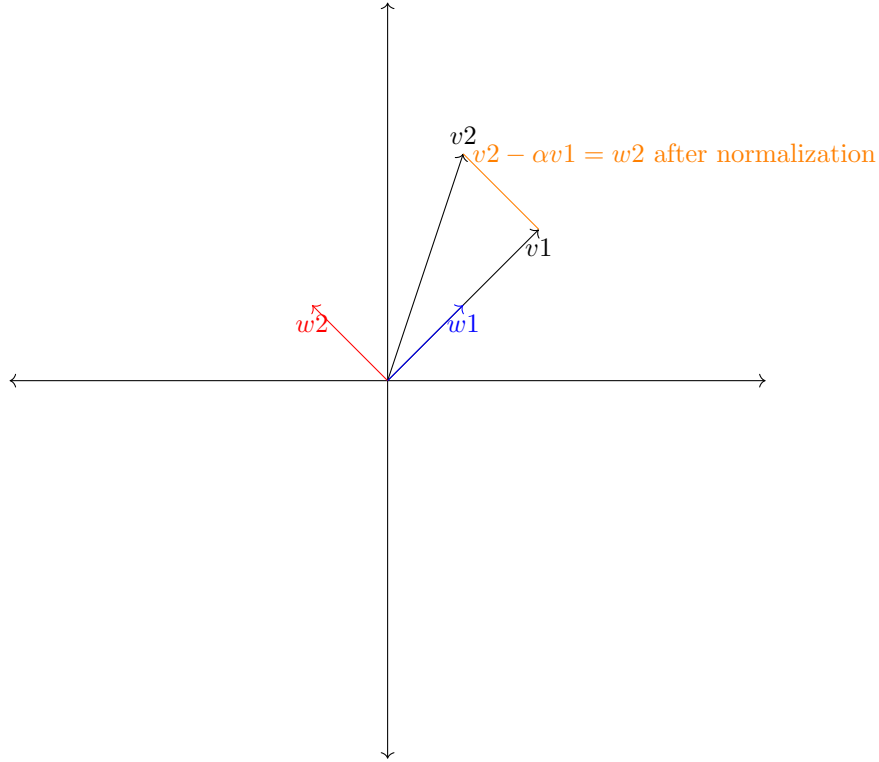
33.12 Best Approximation

Let V be a unitary vector space and U a finitely generated subspace of V. A vector $v* \in U$ is called Best Approximation in U on v, if:

$$\|v*-v\| = \inf_{x \in U} \|x - v\|$$

also

$$\|v* - P_u(v)\| = \inf_{u \in U} \|u - v\|$$

Proof:

Let $x \in V$, v be some point and $v* = proj_V v$. Then

$$v - x = v - v* + v* - x$$

where $v - v* \in V^{\perp}$ and $v* - x \in V$. By the Pythagorean Theorem

$$\|x - v\|^2 = \|v* - v\|^2 + \|v* - x\|^2$$

Therefore

$$\|v* - v\|^2 \leq \|v* - x\|^2$$

$\mathfrak{QED}$

34 Matrices

In this section, we explore the fundamental concepts of matrices, their operations, and important properties that form the foundation of linear algebra.

34.1 Definition of a Matrix

A matrix is a rectangular array of numbers, symbols, or expressions arranged in rows and columns. Formally, an $m \times n$ matrix A consists of mn elements a_{ij} where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m \times n}$. Special cases include:

- Square matrix: matrix with the same number of rows and columns ($m = n$)

- Column vector: $m \times 1$ matrix

- Row vector: $1 \times n$ matrix

- Identity matrix I_n: An $n \times n$ matrix with ones on the main diagonal and zeros elsewhere

- Zero matrix: A matrix where all entries are zero

34.2 Matrix Addition and Subtraction

Matrix addition and subtraction are defined for matrices of the same dimensions.

$$c_{ij} = a_{ij} \pm b_{ij} \quad \text{for all } i = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, n$$

Matrix addition/subtraction satisfies the following properties:

$$A \pm B = B \pm A \quad \text{(Commutativity)}$$
$$(A \pm B) + C = A \pm (B \pm C) \quad \text{(Associativity)}$$
$$A \pm O = A \quad \text{(Identity element)}$$
$$A + (-A) = O \quad \text{(Inverse element)}$$

Where O is the zero matrix.

34.3 Matrix Multiplication

Matrix multiplication is defined between matrices where the number of columns in the first matrix equals the number of rows in the second matrix.

For $A \in \mathbb{C}^{m \times p}$ and $B \in \mathbb{C}^{p \times n}$, their product $C = AB \in \mathbb{C}^{m \times n}$ is defined as:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ip} b_{pj}$$

Matrix multiplication is defined this way to think as each operation as dot product of two vectors.

An example of this is thinking about the dot product as the total price of a purchase in store. The total price is the sum of the amounts which are given by the number of items per price. Now to extend this idea for matrices we think about the number of items as a vector of size i and the prices of each item as another vector p_1 with dot product $\langle i, p_1 \rangle$. If we want to by other items with different prices, but in the

same amounts given by the vector i we get a matrix $A = (p_1, \ldots, p_n)$. Now it makes sense that, if we take three different arrays of products we get a vector of three components which represent each of the total amounts to pay.

Thus, each operation in matrix multiplication is taking the dot product of the *transpose* of the current row with the column vector of the other matrix.

$$A \circ \vec{v} = \begin{bmatrix} \langle a_1^T, v \rangle \\ \langle a_2^T, v \rangle \\ \vdots \\ \langle a_n^T, v \rangle \end{bmatrix}$$

As for the case with multiple columns

$$A \circ B = \begin{bmatrix} \langle a_1^T, b_j \rangle & \ldots & \langle a_1^T, b_n \rangle \\ \vdots & \ldots & \vdots \\ \langle a_n^T, b_j \rangle & \ldots & \langle a_n^T, b_n \rangle \end{bmatrix}$$

Matrix multiplication satisfies the following properties:

$$\begin{aligned} A(BC) &= (AB)C \quad \text{(Associativity)} \\ A(B+C) &= AB + AC \quad \text{(Left distributivity)} \\ (A+B)C &= AC + BC \quad \text{(Right distributivity)} \\ AI_n &= A \quad \text{and} \quad I_m A = A \quad \text{(Identity)} \end{aligned}$$

Note that matrix multiplication is generally not commutative, i.e., $AB \neq BA$ in most cases.

Example:

Consider the matrices A and B given by:

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 0 & -2 \end{bmatrix} \in \mathbb{R}^{2 \times 3} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 4 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$$

To compute the product $C = AB \in \mathbb{R}^{2 \times 2}$, we calculate each entry c_{ij} using the formula:

$$c_{ij} = \sum_{k=1}^{3} a_{ik} b_{kj}$$

Let's calculate each entry of C:

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ &= 2 \cdot 1 + 3 \cdot (-1) + 1 \cdot 4 \\ &= 2 - 3 + 4 = 3 \end{aligned}$$

$$\begin{aligned} c_{12} &= a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ &= 2 \cdot 2 + 3 \cdot 3 + 1 \cdot 0 \\ &= 4 + 9 + 0 = 13 \end{aligned}$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}$$
$$= 1 \cdot 1 + 0 \cdot (-1) + (-2) \cdot 4$$
$$= 1 + 0 - 8 = -7$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}$$
$$= 1 \cdot 2 + 0 \cdot 3 + (-2) \cdot 0$$
$$= 2 + 0 + 0 = 2$$

Therefore, the product $C = AB$ is:

$$C = AB = \begin{bmatrix} 3 & 13 \\ -7 & 2 \end{bmatrix}$$

Let's verify that matrix multiplication is not generally commutative by attempting to compute BA:

Since B is a 3×2 matrix and A is a 2×3 matrix, the product BA would be a 3×3 matrix. However, this calculation cannot be performed since the number of columns in B (which is 2) does not equal the number of rows in A (which is 2). Thus, BA is undefined, demonstrating that matrix multiplication is not always commutative.

34.4 The Transpose of a Matrix

The transpose of a matrix $A \in \mathbb{R}^{m \times n}$, denoted $A^T \in \mathbb{R}^{n \times m}$, is obtained by interchanging rows and columns:

$$(A^T)_{ij} = a_{ji} \quad \text{for all } i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, m$$

34.4.1 Properties of the Transpose

$$\left(A^T\right)^T = A$$
$$(A + B)^T = A^T + B^T$$
$$(AB)^T = B^T A^T$$
$$(\alpha A)^T = \alpha A^T \quad \text{for any scalar } \alpha$$
$$(A)^{-1^T} = (A^T)^{-1}$$
$$\text{rank}(A) = \text{rank}(A^T)$$
$$\det(A) = \det(A^T)$$
$$eig(A) = eig(A^T) \text{ where } eig() \text{ are the eigenvalues}$$

The transpose can also help us interpret the dot product of two vectors being 0 as a matrix multiplication. Here are given $A = (a_1, \ldots, a_n)$ with Ax as some vector and b some other vector and a_k is some row of A.

$$\langle a_k, b - Ax \rangle = 0 \quad \forall k \iff A^T(b - Ax) = \vec{0}$$

34.5   The Equivalence of Matrices

Two matrices A and B are said to be equivalent if one can be transformed into the other through a finite sequence of elementary row operations. We write $A \sim B$ to denote this equivalence.

The elementary row operations are:

- Interchanging two rows: $R_i \leftrightarrow R_j$

- Multiplying a row by a non-zero scalar: $R_i \mapsto \alpha R_i$ where $\alpha \neq 0$

- Adding a multiple of one row to another: $R_i \mapsto R_i + \alpha R_j$ where $i \neq j$

Matrix equivalence is an equivalence relation, satisfying reflexivity, symmetry, and transitivity. Equivalent matrices represent the same linear system in different bases.

34.5.1   Row Echelon Form (REF)

A matrix is in row echelon form if:

- All rows consisting entirely of zeros are at the bottom of the matrix.

- The leading entry (first non-zero element) of each non-zero row is to the right of the leading entry of the row above it.

- All entries in a column below a leading entry are zeros.

34.5.2   Reduced Row Echelon Form (RREF)

A matrix is in reduced row echelon form if:

- It is in row echelon form.

- Each leading entry is 1.

- Each leading entry is the only non-zero entry in its column.

The Gauss-Jordan elimination algorithm proceeds as follows:

- Start with the leftmost non-zero column.

- Find the pivot (non-zero element) in this column. If necessary, swap rows to move a non-zero element to the pivot position.

- Divide the pivot row by the pivot value to make the pivot equal to 1.

- Eliminate all other entries in the pivot column by subtracting appropriate multiples of the pivot row.

- Cover the pivot row and column, and repeat steps 1-4 on the submatrix until all rows are processed.

- For RREF, eliminate all entries above each pivot as well.

Example:

We'll use Gauss-Jordan elimination to solve the linear system:

$$2x + y - z = 8$$
$$-3x - y + 2z = -11$$
$$x + y + z = 3$$

First, we set up the augmented matrix:

$$\begin{bmatrix} 2 & 1 & -1 & | & 8 \\ -3 & -1 & 2 & | & -11 \\ 1 & 1 & 1 & | & 3 \end{bmatrix}$$

Now we apply Gauss-Jordan elimination to transform this into reduced row echelon form:

Step 1: We'll choose the first element in the first row as our pivot. Let's first swap row 1 and row 3 to get a simpler pivot:

$$\begin{bmatrix} 1 & 1 & 1 & | & 3 \\ -3 & -1 & 2 & | & -11 \\ 2 & 1 & -1 & | & 8 \end{bmatrix}$$

Step 2: Eliminate the first elements in rows 2 and 3:

Row 2 + 3 × Row 1:

$$\begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 2 & 5 & | & -2 \\ 2 & 1 & -1 & | & 8 \end{bmatrix}$$

Row 3 - 2 × Row 1:

$$\begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 2 & 5 & | & -2 \\ 0 & -1 & -3 & | & 2 \end{bmatrix}$$

Step 3: Make the pivot in row 2 equal to 1 by dividing the entire row by 2:

$$\begin{bmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & -1 & -3 & | & 2 \end{bmatrix}$$

Step 4: Eliminate the second element in rows 1 and 3:

Row 1 - Row 2:

$$\begin{bmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & -1 & -3 & | & 2 \end{bmatrix}$$

Row 3 + Row 2:

$$\begin{bmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & -\frac{1}{2} & | & 1 \end{bmatrix}$$

Step 5: Make the pivot in row 3 equal to 1 by multiplying the entire row by -2:

$$\begin{bmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

Step 6: Eliminate the third element in rows 1 and 2:

Row 1 + $\frac{3}{2}$× Row 3:

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

Row 2 - $\frac{5}{2}$× Row 3:

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & -1+5=4 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

The matrix is now in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & 4 \\ 0 & 0 & 1 & | & -2 \end{bmatrix}$$

This corresponds to the system:

$$x = 1$$
$$y = 4$$
$$z = -2$$

Therefore, the solution to the original system is $x = 1$, $y = 4$, and $z = -2$.

34.6 The Inverse of a Matrix

For a square matrix $A \in \mathbb{R}^{n \times n}$, the inverse matrix A^{-1} (if it exists) satisfies:

$$AA^{-1} = A^{-1}A = I_n$$

34.6.1 Properties of the Inverse

$$(A^{-1})^{-1} = A$$
$$(AB)^{-1} = B^{-1}A^{-1}$$
$$(A^T)^{-1} = (A^{-1})^T$$
$$\det(A^{-1}) = \frac{1}{\det(A)}$$
$$(cA)^{-1} = \frac{1}{c}A^{-1}$$
$$\det(A^{-1}) = \det(A)^{-1}$$
$$\text{rank}(A^{-1}) = \text{rank}(A)$$
$$eig(A^{-1}) = eig(A)^{-1}$$

34.6.2 Finding the Inverse

There are several methods to find the inverse of a matrix:

Gauss-Jordan Method

Form the augmented matrix $[A|I_n]$ and apply Gauss-Jordan elimination to transform it into $[I_n|A^{-1}]$:

1. Create the augmented matrix $[A|I_n]$
2. Apply row operations to transform the left side into I_n
3. The right side will be A^{-1}

Adjoint Method

For an $n \times n$ matrix A:

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

Where adj(A) is the adjoint (or adjugate) of A, defined as the transpose of the co-factor matrix.

A matrix is invertible if and only if its determinant is non-zero. Such matrices are called non-singular or regular matrices.

Example:

Let's find the inverse of the matrix:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$$

We form the augmented matrix $[A|I_3]$:

$$\begin{bmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 3 & 2 & 1 & | & 0 & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{bmatrix}$$

Now we apply Gauss-Jordan elimination:

Step 1: Make the first pivot equal to 1 by dividing the first row by 2:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 3 & 2 & 1 & | & 0 & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{bmatrix}$$

Step 2: Eliminate the first element in rows 2 and 3:

Row 2 - 3 × Row 1:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & | & -\frac{3}{2} & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{bmatrix}$$

Row 3 - 2 × Row 1:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & | & -\frac{3}{2} & 1 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

Step 3: Make the second pivot equal to 1 by multiplying the second row by 2:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & | & -3 & 2 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

Step 4: Eliminate the second element in row 1 and the third element in row 2:

Row 1 - $\frac{1}{2}$ × Row 2:

$$\begin{bmatrix} 1 & 0 & 1 & | & 2 & -1 & 0 \\ 0 & 1 & -1 & | & -3 & 2 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

Row 2 + Row 3:

$$\begin{bmatrix} 1 & 0 & 1 & | & 2 & -1 & 0 \\ 0 & 1 & 0 & | & -4 & 2 & 1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

Step 5: Eliminate the third element in row 1:

Row 1 - Row 3:

$$\begin{bmatrix} 1 & 0 & 0 & | & 3 & -1 & -1 \\ 0 & 1 & 0 & | & -4 & 2 & 1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

The right side of the augmented matrix now gives us A^{-1}:

$$A^{-1} = \begin{bmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

To verify, we can check that $AA^{-1} = I_3$:

$$\begin{aligned} AA^{-1} &= \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot 3 + 1 \cdot (-4) + 1 \cdot (-1) & 2 \cdot (-1) + 1 \cdot 2 + 1 \cdot 0 & 2 \cdot (-1) + 1 \cdot 1 + 1 \cdot 1 \\ 3 \cdot 3 + 2 \cdot (-4) + 1 \cdot (-1) & 3 \cdot (-1) + 2 \cdot 2 + 1 \cdot 0 & 3 \cdot (-1) + 2 \cdot 1 + 1 \cdot 1 \\ 2 \cdot 3 + 1 \cdot (-4) + 2 \cdot (-1) & 2 \cdot (-1) + 1 \cdot 2 + 2 \cdot 0 & 2 \cdot (-1) + 1 \cdot 1 + 2 \cdot 1 \end{bmatrix} \\ &= \begin{bmatrix} 6 - 4 - 1 & -2 + 2 + 0 & -2 + 1 + 1 \\ 9 - 8 - 1 & -3 + 4 + 0 & -3 + 2 + 1 \\ 6 - 4 - 2 & -2 + 2 + 0 & -2 + 1 + 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3 \end{aligned}$$

This confirms that we have correctly found the inverse of matrix A.

34.7   The Rank of a Matrix

The rank of a matrix A, denoted rank(A) or rg(A), is the dimension of the column space (or equivalently, the row space) of A.

Equivalent definitions of rank include:

- The maximum number of linearly independent columns of A
- The maximum number of linearly independent rows of A
- The order of the largest non-zero minor of A
- The number of non-zero rows in any row echelon form of A

34.8   How to isolate Matrices

Like in other types of equations we can isolate a matrix by doing the proper operations.

- $A \pm B = C \implies A = C \mp B$
- $AB = C \implies ABB^{-1} = CB^{-1} \implies A = CB^{-1}$
- $AB = C \implies A^{-1}AB = A^{-1}C \implies B = A^{-1}C$

34.8.1 Finding the Rank

To find the rank of a matrix:

1. Transform the matrix into row echelon form using Gauss-Jordan elimination

2. Count the number of non-zero rows in the resulting matrix

34.8.2 Properties of Rank

$$\text{rank}(A) \leq \min(m, n) \text{ for } A \in \mathbb{R}^{m \times n}$$
$$\text{rank}(A^T) = \text{rank}(A)$$
$$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$$
$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$$

For a square matrix $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

- A is invertible

- $\text{rank}(A) = n$

- $\det(A) \neq 0$

- The columns of A are linearly independent

- The rows of A are linearly independent

- $Ax = 0$ has only the trivial solution $x = 0$

34.9 Column Space

The column space of A, denoted $\text{Col}(A)$, is the span of the columns of A:

$$\text{Col}(A) = \{\vec{y} \in \mathbb{R}^m : \vec{y} = A\vec{x} \text{ for some } \vec{x} \in \mathbb{R}^n\}$$

This is also called the range or image of the linear transformation represented by A. The dimension of the column space equals the rank of A.

34.10 Row Space

The *row space* of A, denoted $\text{Row}(A)$, is the span of the rows of A: It is also perpendicular to the *Null space*.

$$\text{Row}(A) = \text{Col}(A^T)$$

The dimension of the *row space* also equals the rank of A.

34.11 Null Space

The null space (or kernel) of A, denoted $\text{Null}(A)$ or $\text{Ker}(A)$, is the set of all vectors that A maps to zero:

$$\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$$

The dimension of the null space is called the nullity of A, denoted $\text{nullity}(A)$.

34.11.1 Left Null Space

The left null space of A is the null space of A^T:

$$\text{Null}(A^T) = \{\vec{y} \in \mathbb{R}^m : A^T \vec{y} = \vec{0}\} = \{\vec{y} \in \mathbb{R}^m : \vec{y}^T A = \vec{0}^T\}$$

The Rank-Nullity Theorem connects these spaces:

$$\text{rank}(A) + \text{nullity}(A) = n$$

To find a basis for these spaces:

- *Column space*: Take the linearly independent columns of A.
- *Row space*: Take the non-zero rows from any row echelon form of A.
- *Null space*: Solve the homogeneous system $A\vec{x} = \vec{0}$ and express the general solution in terms of free variables.

34.12 Linear Maps as matrices

Linear maps between vector spaces V and can be represented by matrices. In the following way

$$A\vec{x} = \sum_{i=1}^{n} x_i \vec{a}_i$$

where $\vec{a}_i$ are images of the basis vector of V as columns of A and $\vec{x} = (x_1, \ldots, x_n)^T$.

34.13 Matrix Representation of Linear Maps on Basis vectors

Let V and W be two $\mathbb{F}$-vector spaces, $\beta = \{v_1, \ldots, v_n\}$ a basis of V and $\gamma = \{w_1, \ldots, w_m\}$ a basis of W, and let $f : V \to W$ be linear. Then there exists a unique matrix $A = (a_{ij}) \in \mathbb{F}^{m \times n}$ with

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \forall j = 1, \ldots, n$$

Example:

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, and define a linear map $f : V \to W$ by:

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + 2y \\ 3x + y \\ 4y \end{bmatrix}$$

We aim to find the matrix representation $[M]_\beta^\gamma(f) \in \mathbb{R}^{3 \times 2}$ of f, with respect to the standard bases:

$$\beta = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \subseteq V$$

$$\gamma = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \subseteq W$$

Step 1: Apply f to the basis vectors of V

$$f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1 + 2 \cdot 0\\3 \cdot 1 + 0\\4 \cdot 0\end{bmatrix} = \begin{bmatrix}1\\3\\0\end{bmatrix}$$

$$f\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0 + 2\\0 + 1\\4\end{bmatrix} = \begin{bmatrix}2\\1\\4\end{bmatrix}$$

Step 2: Form the matrix $[M]_\beta^\gamma(f)$

The columns of the matrix are the coordinate vectors of $f(v_1)$ and $f(v_2)$ in basis γ:

$$[M]_\beta^\gamma(f) = \begin{bmatrix}1 & 2\\3 & 1\\0 & 4\end{bmatrix}$$

Step 3: Verify the matrix representation

We check that:

$$f(v_j) = \sum_{i=1}^{3} a_{ij} w_i \quad \text{for } j = 1, 2$$

$$f(v_1) = 1w_1 + 3w_2 + 0w_3$$
$$f(v_2) = 2w_1 + 1w_2 + 4w_3$$

This confirms the matrix representation is correct.

34.14 Matrices with respect to a basis

Expressing a linear map in terms of a different basis means that given two vector-spaces V and W with the basis $\beta = \{v_1, \ldots, v_n\}$ and $\gamma = \{w_1, \ldots, w_n\}$ respectively. Then the map $f : V \to W$ with a matrix representation A whose columns are vectors in the basis β can be expressed in terms of the basis γ using the following steps:

1. Determine the images of the basis vectors of V under f which will give us the matrix A_β. If not already given.

2. For every basis vector w_i of W apply $A_\beta w_i$. This is going to give you a vector k_i

3. Express each k_i as a linear combination of the basis vectors of W. Which means finding coefficients λ_k such that $\sum_{k=i}^{n} \lambda_k w_k = k_j$. This can be found using Gaussian elimination.

4. Form the matrix A_γ using each collection of coefficients λ_k as columns, where each column corresponds to the image of a basis vector v_j of V.

Another much faster, method to express a linear map in terms of another basis is to use the change of basis formula. If φ_γ is the change of basis matrix from γ to the canonical basis formed by the basis vector of W, and φ_γ^{-1} is the change of basis matrix from the canonical basis to γ, then:

$$A_\gamma = \varphi_\gamma^{-1} A_\beta \varphi_\gamma$$

The intuition behind is that in the previous method we are solving the system of equations of $A_\gamma \vec{x}_i = A_\beta w_i$ for each basis vector w_i in γ. Which is equivalent of $\vec{x}_i = \varphi_\gamma^{-1} A_\beta w_i$ when using Gaussian elimination. If you then put all $\vec{x}_i$ in order in a matrix then you get A_γ.

34.15 Example of a complete exercise

1. Determination of the kernel

2. Determination of the dimension of the kernel

3. Determination of the rank (dimension formula)

4. Determination of the image

Example:

Given

$$f \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - x_2 + x_3 \\ 4x_1 - x_2 + 2x_3 \end{bmatrix}.$$

It should be shown that f is linear, and $\ker(f)$, $\mathrm{img}(f)$ and their dimensions should be determined.

A direct proof of linearity is easily possible. Instead, we give the transformation matrix A. The images of the (canonical) basis vectors are

$$f \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad f \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \quad f \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

We obtain

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{bmatrix}.$$

But now it must be shown that indeed $f(x) = Ax \quad \forall x \in \mathbb{R}^3$ holds, by, for example, calculating both Ax and $f(x)$ for a general x and showing equality: Here, with $x = (x_1, x_2, x_3)^T$

$$A \cdot x = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - x_2 + x_3 \\ 4x_1 - x_2 + 2x_3 \end{bmatrix}.$$

This obviously corresponds to $f(x)$, so that by Theorem 4.36, the map f is linear. To determine the kernel, one has to solve the system of linear equations

$$\begin{array}{cccc} 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 4 & -1 & 2 & 0 \end{array}$$

Which corresponds to the equation $Ax = f(x) = 0$. Gaussian elimination yields $x_3 = \lambda'$; $x_2 = \frac{2}{3} \cdot \lambda'$; $x_1 = -\frac{1}{3} \cdot \lambda'$, thus, with $\lambda = \frac{1}{3}\lambda'$:

$$\ker(f) = \left\{ x = \lambda \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \,\middle|\, \lambda \in \mathbb{R} \right\}$$

It follows that $\dim(\ker(f)) = 1$ and because of $\dim(V) = 3$ from the dimension formula $\dim(\mathrm{img}(f)) = 2$.

$\mathrm{img}(f)$ corresponds to the linear span of the columns of the matrix. One chooses consequently $\dim(\mathrm{img}(f))$ column vectors, e.g., the first ones, and tests if they are linearly independent. In the concrete case, this is obvious, because the second column is not a multiple of the first. It follows therefore,

$$\text{img}(f) = \left\{ x \, \middle| \, x = \lambda \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \quad \lambda, \mu \in \mathbb{R} \right\}$$

34.16 Diagonal and Triangular Matrices

A matrix is called *Diagonal* if the only non-zero elements are the elements on the main diagonal. A matrix is *Over Triangular* if all elements under or above the main diagonal are all zeroes.

34.16.1 Properties of Diagonal and Triangular Matrices

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

$$eig(A) = \{a_{1,1}, a_{2,2}, \ldots, a_{n,n}\}$$

34.17 Complex Matrices

This kind of matrices can have complex numbers in some or all of its entries. Something special about them is they can be written as the sum of a real matrix $A \in \mathbb{R}^{n \times m}$ and a matrix $B \in \mathbb{C}^{n \times m}$ which is made of the imaginary part of the entries of the original matrix.

Complex matrices also have complex conjugates.

34.17.1 The Hermitian Transpose

The *Hermitian Transpose* A^H is the transpose of the conjugate matrix of a matrix A.

34.17.2 Hermitian Matrices

Are matrices with the following property:

$$M^H = M$$

34.17.3 Unitary Matrices

Unitary Matrices are matrices whose columns form orthonormal vectors and

$$U^H = U^{-1}$$

35 Linear Maps

A linear map $f : V \to W$ is a function that satisfies the following properties:

- $f(v_1 + v_2) = f(v_1) + f(v_2)$ for all $v_1, v_2 \in V$.
- $f(\lambda v) = \lambda f(v)$ for all $v \in V$ and $\lambda \in \mathbb{R}$.
- $f(0) = 0$.

This kind of function is also called *Homomorphism*. If $V = W$ is it called *Endomorphism*.

35.1 Types of Linear Maps

- *Injective or Monomorphism* (One-to-One): A linear map $f : V \to W$ is injective if $f(v_1) = f(v_2)$ implies $v_1 = v_2$.
- *Surjective or Epimorphism* (Onto): A linear map $f : V \to W$ is surjective if for every $w \in W$, there exists a $v \in V$ such that $f(v) = w$.
- *Bijective or Isomorphism*: A linear map $f : V \to W$ is bijective if it is both injective and surjective.

Isomorphic if there exists a bijective linear map between them. In this case, we can say that the two vector spaces are *isomorphic*, and we write $V \cong W$.

35.2 Properties of Linear Maps

- The composition of two linear maps is a linear map.
- The inverse of a bijective linear map is also a linear map.
- The zero map $f : V \to W$ defined by $f(v) = 0$ for all $v \in V$ is a linear map.
- The identity map $\text{id}_V : V \to V$ defined by $\text{id}_V(v) = v$ for all $v \in V$ is a linear map.
- The sum of two linear maps $f : V \to W$ and $g : V \to W$ is a linear map defined by $(f + g)(v) = f(v) + g(v)$.
- The scalar multiplication of a linear map $f : V \to W$ by a scalar c is a linear map defined by $(cf)(v) = c(f(v))$.
- The composition of linear maps is associative, i.e., $(f \circ g) \circ h = f \circ (g \circ h)$.
- The composition of linear maps is distributive over addition, i.e., $f \circ (g + h) = f \circ g + f \circ h$.
- The composition of linear maps is compatible with scalar multiplication, i.e., $(cf) \circ g = c(f \circ g)$.
- Linear Maps compose a vector space over the field of scalars.

$$(Hom(V, W), +, \cdot)$$

35.3 Linear Extension Theorem

Let $\beta = \{v_1, \ldots, v_n\}$ be a basis of the finite dimensional vector space V and $w_1, \ldots, w_n$ be arbitrary vector in W. Then $!\exists L$ with $L : V \to W$ such that $T(v_i) = w_i$.

Proof:

Define T as follows: Whenever $a = \lambda_1 v_1 + \cdots + \lambda_n v_n$ then $T(a) = \lambda_1 w_1 + \cdots + \lambda_n w_n$.

We know it is unique because for $\lambda_1, \ldots, \lambda_n \neq \gamma_1, \ldots, \gamma_n$ if we assume $a - b = \vec{0}$ then:

$$a - b = (\lambda_1 - \gamma_1)v_1 + \cdots + (\lambda_n - \gamma_n)v_n = \vec{0}$$

If and only if $\lambda_1, \ldots, \lambda_n = \gamma_1, \ldots, \gamma_n$. Thus, it works uniquely for each vector.

For the linearity, using the definition of linearity and T this part becomes trivial.

Finally, for the uniqueness of T itself, let us assume there is a map U which $U(v_i) = T(v_i)$. Because we can write a vector a linear combination we will arrive at the fact that

$$U(x) = \lambda_1 U(v_1) + \cdots + \lambda_n U(v_n) = \lambda_1 T v_1 + \cdots + \lambda_n T(v_n)$$

Therefore, $U = T$.

<div align="right">𝔔𝔈𝔇</div>

35.4 The Kernel of a Linear Map

The kernel of a linear map $f : V \to W$ is the set of all vectors in V that are mapped to the zero vector in W:

$$\ker(f) = \{v \in V \mid f(v) = 0\}$$

To compute the *kernel* of the linear map, we just solve the homogenous system of equations $Ax = \vec{0}$ where A is the matrix representation of f. The number of free variables is the $\dim(\ker(f))$.

35.5 The Image of a Linear Map

The image of a linear map $f : V \to W$ is the set of all vectors in W that can be expressed as $f(v)$ for some $v \in V$:

$$\text{img}(f) = \{w \in W \mid w = f(v) \text{ for some } v \in V\}$$

To find $\dim(\text{im}(f))$, compute the rank of the matrix representing f. This is the number of pivot columns in the row-reduced form of A.

35.6 The Rank of a Linear Map

The rank of a linear map $f : V \to W$ is the dimension of its image:

$$\text{rank}(f) = \text{img}(f)$$

35.7 The Nullity of a Linear Map

The nullity of a linear map $f : V \to W$ is the dimension of its kernel:

$$\text{nullity}(f) = \ker(f)$$

35.8 Proof of the injectivity of the Kernel

Suppose we have a linear map f and two vectors v_1 and v_2 which are not equal. We are going to assume that they both map to the $\vec{0}$ therefore:

$$f(v_1) = f(v_2) = \vec{0}$$

Because of the linearity of the map we can write:

$$f(v_1 - v_2) = f(v_1) - f(v_2) = \vec{0} - \vec{0} = \vec{0}$$

And that contradicts the assumption that v_1 and v_2 are not equal. Therefore, the kernel of a linear map is injective.

35.9 Dimension Formula for Linear Mappings

Let $f : V \to W$ be linear and $\dim(V) = n$. Then

$$\dim(\ker(f)) + \mathrm{rank}(f) = n.$$

Remember that $\ker(f)$ forms a subspace of V and therefore, $\dim(\ker(f)) := r \leq n$. We extend an arbitrary basis $(v_1, \ldots, v_r)$ of $\ker(f)$ to a basis $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ of V. Setting $w_{r+i} = f(v_{r+i})$ for $i = 1, \ldots, n - r$, we have $\forall v \in V$:

$$\begin{aligned}
f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n) \\
&= \lambda_1 \underbrace{f(v_1)}_{=0} + \cdots + \lambda_r \underbrace{f(v_r)}_{=0} + \lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n) \\
&= \lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n) \\
&= \lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n
\end{aligned}$$

Thus, $\mathrm{img}(f) = \mathrm{span}(w_{r+1}, \ldots, w_n)$. We now show that $w_{r+1}, \ldots, w_n$ are linearly independent. Let

$$\lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n = 0$$

From

$$0 = \lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n = f(\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n)$$

it follows that

$$\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n \in \ker(f)$$

Therefore,

$$\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n = \lambda_1 v_1 + \cdots + \lambda_r v_r$$

For some $\lambda_1, \ldots, \lambda_r$. Since $v_1, \ldots, v_n$ are linearly independent, we have $\lambda_1 = \cdots = \lambda_n = 0$. Thus, the vectors $w_{r+1}, \ldots, w_n$ are linearly independent.

It follows that $\dim(\mathrm{img}(f)) = \mathrm{rank}(f) = n - r$ and therefore,

$$\dim(\ker(f)) + \dim(\mathrm{img}(f)) = \dim(\ker(f)) + \mathrm{rank}(f) = r + n - r = n$$

$$\mathfrak{QED}$$

35.10 Linear Maps and Matrices

Let V and W be finite-dimensional vector spaces over the same field K. If $\dim V = n$ and $\dim W = m$, then a linear map $f : V \to W$ can be represented by an $m \times n$ matrix. The action of the linear map on a vector can be expressed as:

$$f(x) = A \cdot x,$$

where A is the matrix representation of f and v is the vector represented in a column format. The columns of the matrix A are the images of the basis vectors of V under the linear map f.

35.11 How to identify the type of Linear Map Summary

We want to summarize how to identify the type of linear map given $f : V \to W$,

- *Injective*: A linear map f is injective if and only if

$$\dim(\ker(f)) = 0.$$

- *Surjective*: The map f is surjective if every element of W is the image of some element in V. This occurs when

$$\dim(\operatorname{img}(f)) = \dim(W).$$

  That is, the image of f are the linear independent vector of f.

- *Bijective*: A linear map is bijective if it is both injective and surjective. Therefore, it must satisfy both conditions:

$$\dim(\ker(f)) = 0 \quad \text{and} \quad \dim(\operatorname{im}(f)) = \dim(W).$$

- *Neither*: If neither of the above conditions is satisfied, the map is neither injective nor surjective.

Example:

Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \end{bmatrix}.$$

To identify the type of this linear map:

- Compute $\dim(\ker(f))$: By reducing the matrix A to row echelon form, we find that $\dim(\ker(f)) = 1$. This means f is not injective.

- Compute $\dim(\operatorname{im}(f))$: The image corresponds to the column space of A. Since there are 2 linearly independent rows in the row echelon form, we have $\dim(\operatorname{im}(f)) = 2$, which equals $\dim(W) = 2$. Therefore, f is surjective.

Thus, f is surjective but not injective.

36   Changing between Basis

The change of basis is a very important concept in linear algebra which allows us to change the way we describe vectors and linear transformations. This is particularly useful when we want to express vectors in a different coordinate systems, or we want to apply a linear transformation, in an easy-to-work basis. A good metaphor for this is to think of it as translating between different languages, where each basis is a different language. While keeping in mind that from the perspective of a user of a different basis, the basis vectors are not different from the canonical basis.

36.1   Understanding the change of basis

Given a vector $\vec{x} = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n$ with $\vec{v}_1, \ldots, \vec{v}_n$ being the basis vectors, we can write $\vec{x}$ in terms of its components $(x_1, \ldots, x_n)^T$.

Now imagine having another basis $\vec{b}_1, \ldots, \vec{b}_n$ and the same vector $\vec{x}$ with all of its components. The key here is that we want to know how our original vector is described in terms of other basis. More precisely what are its coordinates

So our vector $\vec{x}$ is described:

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n = \vec{x} = \mu_1 \vec{b}_1 + \cdots + \mu_n \vec{b}_n,$$

in the corresponding bases. Note, that the vectors $\vec{v}_i$ are written in the standard basis. And same goes for the vectors $\vec{b}_i$.

Now let us write then as matrix vector multiplication

$$(\vec{v}_1, \ldots, \vec{v}_n) \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = (\vec{b}_1, \ldots, \vec{b}_n) \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}$$

Now let us use the following notation for the basis and the coefficient vectors.

$$P_\beta (\vec{x})_\beta = \vec{x} = P_\gamma (\vec{x})_\gamma$$

Here P_β is the matrix of the basis V vectors and $(\vec{x})_\beta$ the coefficients of the vector described by that basis. The same goes for P_γ and $(\vec{x})_\gamma$.

Now that conversion of basis is just a matter of solving for our desired vector by multiplying by the inverse of the corresponding basis matrix. We can also ignore the vector x in the middle and just use the equality where we can solve for the vector in the desired basis.

$$P_\beta (\vec{x})_\beta = P_\gamma (\vec{x})_\gamma$$

Now let us understand what really is happening in the step where we multiply by the inverse matrix.

$$\begin{aligned} P_\gamma^{-1} P_\beta &= P_\gamma^{-1} (\vec{b}_1, \ldots, \vec{b}_n) \\ &= (P_\gamma^{-1} \vec{b}_1, \ldots, P_\gamma^{-1} \vec{b}_n) \\ &= ((\vec{b}_1)_\gamma, \ldots, (\vec{b}_n)_\gamma) \end{aligned}$$

When we start from some basis, our perspective is that our basis look like the standard basis. Given a matrix P_β which consist of our basis vectors from the perspective of the 'real' standard basis, when we multiply a vector of coefficients in our basis with P_β we are going from our basis to the standard basis. If we then apply P_β^{-1} this takes us back to our basis.

Knowing that, then if instead of P_β^{-1} we apply the inverse matrix of another basis P_γ^{-1} then it will translate our vector from the canonical basis to the basis γ. Which at the end is simply getting the coordinate of the basis vectors of β in the basis γ.

36.2 Coordinate mapping

Let V be a $\mathbb{F}$-vector space with a basis $\beta = \{v_1, \ldots, v_n\}$. Then there exists exactly one isomorphism

$$\varphi_\beta : \mathbb{F}^n \to V$$

such that

$$\varphi_\beta(e_i) = v_i, \quad \text{for } 1 \leq i \leq n.$$

which is called the *coordinate mapping* of V with respect to the basis β.

36.3 Coordinates with respect to a basis

The isomorphism φ_β from the previous theorem is called the *coordinate mapping*, and for $v \in V$, we define

$$(\vec{v})_\beta := \varphi_\beta^{-1}(v) \in \mathbb{F}^n$$

as the *coordinates of v with respect to β*.

36.4 Translation of Coordinates

We define the *Translation* of a vector to another basis as:

$$T_\beta^\gamma = \varphi_\beta^{-1} \circ \varphi_\gamma$$

where φ_β and φ_γ are the matrices made of the basis vector of the bases β and A with respect to the standard basis, respectively.

36.5 Change of Basis Theorem

Let $v \in V$ be arbitrary, with $(\vec{v})_\gamma = (x_1, \ldots, x_n)^T$ and $(\vec{v})_\beta = (y_1, \ldots, y_n)^T$. Then the following holds:

$$(\vec{v})_\beta = \varphi_\beta^{-1} \circ \varphi_\gamma(\vec{v})_\gamma$$

In vector notation with the translation:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = T_\beta^A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

36.6 Change of Basis in Linear Maps

Let V and W be finitely generated $\mathbb{F}$-vector spaces with bases γ and β, respectively, and let $f \in \text{Hom}(V, W)$. Then we have

$$M_\beta^\gamma(f) = \varphi_\beta^{-1} \circ f \circ \varphi_\gamma$$

Thus, $M^{\gamma}_{\beta}(f)$ tells you to take vector in the basis γ, translate it to the canonical basis, apply the linear map f and then convert the result to the basis β.

36.7 Transformations involving multiple bases

Let V and W be finitely generated vector spaces with bases γ and γ' for V, and β and β' for W, respectively. Furthermore, let $f : V \to W$ be a linear map. Then the following holds:

$$M^{\gamma'}_{\beta'}(f) = T^{\beta}_{\beta'} \cdot M^{\gamma}_{\beta}(f) \cdot \left(T^{\gamma}_{\gamma'}\right)^{-1}$$

More explicitly, we can write this as:

$$\begin{aligned}
M^{\gamma'}_{\beta'}(f) &= T^{\beta}_{\beta'} \cdot M^{\gamma}_{\beta}(f) \cdot \left(T^{\gamma}_{\gamma'}\right)^{-1} \\
&= (\varphi^{-1}_{\beta'} \circ \varphi_{\beta}) \circ (\varphi^{-1}_{\beta} \circ f \circ \varphi_{\gamma}) \circ (\varphi^{-1}_{\gamma'} \circ \varphi_{\gamma})^{-1}. \\
&= (\varphi^{-1}_{\beta'} \circ \varphi_{\beta}) \circ (\varphi^{-1}_{\beta} \circ f \circ \varphi_{\gamma}) \circ (\varphi^{-1}_{\gamma} \circ \varphi_{\gamma'}).
\end{aligned}$$

We can visualize this as a commutative diagram where when take an inverse we reverse the direction of the arrows:

37 The Determinant of a Matrix

The determinant of a square matrix $A \in \mathbb{R}^{n \times n}$, denoted $\det(A)$ or $|A|$, is a scalar value that provides important information about the matrix, including whether it is invertible and the volume scaling factor of the linear transformation represented by A. The determinant can be computed using various methods, including the Laplace expansion, row reduction, or the Leibniz formula.

The determinant of a 2×2 matrix is given by:

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

For a 3×3 matrix, the determinant can be computed using the rule of Sarrus or the co-factor expansion:

$$\det(A) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh$$

The determinant of larger matrices can be computed using co-factor expansion along any row or column:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}),$$

where A_{ij} is the $(n-1) \times (n-1)$ sub-matrix obtained by deleting the i-th row and j-th column of A.

37.1 Properties of the determinant

- $\det(A) = 0$ if and only if A is singular (not invertible).
- $\det(AB) = \det(A) \cdot \det(B)$ for any square matrices A and B of the same size.
- $\det(A^T) = \det(A)$.
- If a row (or column) of A is multiplied by a scalar α, then $\det(A)$ is multiplied by α.
- If two rows (or columns) of A are swapped, then $\det(A)$ changes sign.

$$\det(a, b, c) = -\det(b, a, c)$$

- If a row (or column) of A is added to another row (or column), then $\det(A)$ remains unchanged.
- If one of the columns is a linear combination of the others, then $\det(A) = 0$.
- The determinant of the identity matrix I_n is 1.
- The determinant of can split into the sum of more determinants:

$$\det(a, b, c + d) = \det(a, b, c) + \det(a, b, d)$$

- The determinant of a diagonal matrix is the product of its diagonal entries.
- For elementary matrices $\det C1 = 1$, $\det C2 = -1$, $\det C3 = \lambda$.
- For $A \in \mathbb{R}^{n \times n}$ $\det(\lambda A) = \lambda^n \det(A)$

37.2 Proof of the multiplication of determinants

For $A, B \in \mathbb{F}^{n \times n}$, it holds that

$$\det(AB) = \det(A)\det(B).$$

Proof:

If $\det(A) = 0$ or $\det(B) = 0$, then $\operatorname{rank}(L_A) < n$ or $\operatorname{rank}(L_B) < n$.

Thus, $\operatorname{rank}(L_{AB}) < n$, and the statement is trivial.

So let us assume that A and B are invertible, i.e., $\operatorname{rank}(A) = \operatorname{rank}(B) = n$.

We now assume that in such a case, a matrix can always be transformed into reduced row echelon form using the row operations from Gaussian elimination.

Then, $A = C_k \cdots\cdots C_1$ for certain elementary matrices C_i, and therefore:

$$
\begin{aligned}
\det(AB) &= \det(C_k \cdots\cdots C_1 B) \\
&= \det(C_k)\det(C_{k-1} \cdots\cdots C_1 B) \\
&\quad\vdots \\
&= \det(C_k)\cdots\det(C_1)\det(B) \\
&= \det(C_k \cdots\cdots C_1)\det(B) \\
&= \det(A)\det(B).
\end{aligned}
$$

Therefore, also if A is invertible, then

$$\det(A^{-1}) = (\det(A))^{-1}.$$

Proof:

It holds that

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(E) = 1.$$

37.3 The Leibniz Formula

Let $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ be a square matrix over a field K. The determinant of A is defined via the **Leibniz formula** as:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

where:

- S_n is the set of all permutations of $\{1, 2, \ldots, n\}$,

- $\operatorname{sgn}(\sigma)$ is the sign of the permutation σ, equal to $+1$ for even permutations and -1 for odd ones.

37.3.1 Derivation

Step 1: Expansion using multilinearity

Each column vector a_j of the matrix A can be written as a linear combination of the standard basis vectors $e_1, \ldots, e_n$:

$$a_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Using the multilinearity of the determinant function (i.e., linearity in each column), we expand:

$$\det(a_1, a_2, \ldots, a_n) = \det\left(\sum_{i_1} a_{i_1 1} e_{i_1}, \sum_{i_2} a_{i_2 2} e_{i_2}, \ldots, \sum_{i_n} a_{i_n n} e_{i_n}\right).$$

By multilinearity, this becomes a sum over all n-tuples $(i_1, i_2, \ldots, i_n)$:

$$= \sum_{i_1, i_2, \ldots, i_n = 1}^{n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} \cdot \det(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

Step 2: Eliminate repeated indices

If two indices $i_j = i_k$ for $j \neq k$, then the corresponding determinant is zero (because the determinant is alternating and thus, zero for repeated columns). Therefore, only terms with pairwise distinct indices remain. These correspond to permutations of $\{1, 2, \ldots, n\}$.

Let $\sigma \in S_n$ be such a permutation. Then we can write:

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \cdot \det(e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}).$$

Step 3: Sign of the permutation

Since the determinant of $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ is $\mathrm{sgn}(\sigma)$, we obtain:

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Equivalently, by re-indexing:

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i\sigma(i)}.$$

This formula expresses the determinant as a sum over all permutations of the indices $\{1, \ldots, n\}$, where each term is a product of matrix entries taken from different rows and columns, weighted by the sign of the corresponding permutation.

37.4 Laplace's Method (Co-factor Expansion)

The determinant of an $n \times n$ matrix $A = [a_{ij}]$ can be computed using the *Laplace expansion*, also known as co-factor expansion. This method expands the determinant along a specific row or column.

Expansion Along Row i:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

Expansion Along Column j:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

Where:

- a_{ij} is the element in the i-th row and j-th column of matrix A.

- M_{ij} is the *minor* of a_{ij}, i.e., the determinant of the submatrix formed by removing the i-th row and j-th column from A.

- $(-1)^{i+j}$ is the *sign factor*, determining the sign of each co-factor term.

Steps:

1. **Choose a Row or Column:** Select any row or column of the matrix. It's often easiest to choose one with many zeros.

2. **For Each Element:** For each element, a_{ij}, in the chosen row or column:

Find the Minor, M_{ij}: The minor M_{ij} is the determinant of the sub-matrix formed by deleting the i-th row and the j-th column of the original matrix.

Find the Co-factor, C_{ij}: The co-factor C_{ij} is the minor multiplied by a sign factor:

$$C_{ij} = (-1)^{i+j} M_{ij}$$

The term $(-1)^{i+j}$ gives a checkerboard pattern of signs:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

3. **Calculate the Determinant:** The determinant of the matrix, A, is the sum of the products of the elements in the chosen row or column and their corresponding co-factors.

Expansion along the i-th row:

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} = a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in}$$

Expansion along the j-th column:

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} = a_{1j} C_{1j} + a_{2j} C_{2j} + \cdots + a_{nj} C_{nj}$$

Both expansions give the same result.

Example (3×3 Matrix):

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Expanding along the first row:

1. a_{11}: $M_{11} = \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$, $C_{11} = +M_{11}$

2. a_{12}: $M_{12} = \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$, $C_{12} = -M_{12}$

3. a_{13}: $M_{13} = \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$, $C_{13} = +M_{13}$

Therefore,

$$\det(A) = a_{11} C_{11} + a_{12} C_{12} + a_{13} C_{13}$$

Example of Determinant Calculation

Let's calculate the determinant of the matrix:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

Using the rule of Sarrus for 3×3 matrices:

$$\det(A) = 2 \cdot 0 \cdot 1 + 1 \cdot 2 \cdot 3 + 3 \cdot 1 \cdot 1 - (3 \cdot 0 \cdot 0 + 1 \cdot 2 \cdot 2 + 2 \cdot 1 \cdot 1)$$
$$= 0 + 6 + 3 - (0 + 4 + 2)$$
$$= 9 - 6 = 3$$

Thus, the determinant of matrix A is $\det(A) = 3$.

Now consider the following 4×4 matrix:

$$A = \begin{bmatrix} 2 & 1 & 3 & 2 \\ 4 & 0 & -1 & 3 \\ -2 & 3 & 1 & 5 \\ 1 & -1 & 0 & 2 \end{bmatrix}$$

For a 4×4 matrix, we can use Laplace Method along the first row:

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + a_{14}C_{14}$$
$$= 2 \cdot \det \begin{bmatrix} 0 & -1 & 3 \\ 3 & 1 & 5 \\ -1 & 0 & 2 \end{bmatrix} - 1 \cdot \det \begin{bmatrix} 4 & -1 & 3 \\ -2 & 1 & 5 \\ 1 & 0 & 2 \end{bmatrix}$$
$$+ 3 \cdot \det \begin{bmatrix} 4 & 0 & 3 \\ -2 & 3 & 5 \\ 1 & -1 & 2 \end{bmatrix} - 2 \cdot \det \begin{bmatrix} 4 & 0 & -1 \\ -2 & 3 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\det(A) = 145$$

37.5 Determinant of the reduced echelon form

The determinant has another trait that states that if your matrix A is REF then its determinant is given by the product of the main diagonal.

37.6 Gaussian Method for finding determinants

Another method for finding determinants is using the row operations we already knew plus also column operations but with the catch that for the column case we have to take into consideration how operating with columns affect the determinant, in contrast to the row operations, except for scaling a row by some λ in this case we would need to undo the change by multiplying the determinant by $\frac{1}{\lambda}$.

Example:

We are given the matrix:

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

We apply Gaussian elimination:

Swap $R_2 \leftrightarrow R_3$:

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Now the matrix is upper triangular. The determinant is the product of diagonal elements times -1 for one row swap:

$$\det(A) = (-1)^1 \cdot (1) \cdot (-1) \cdot (-1) \cdot (3) = -3$$

$$\det(A) = -3$$

38 Elementary Matrix

We are going to define 3 matrices that can be obtained from the Elementary Row Operations on the *Identity Matrix* and call them *Elementary Matrices.*

C1: Row Addition

$$C_1 := \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \lambda & 1 & \\ & & & & \ddots \end{bmatrix} \in \mathbb{F}^{n \times n} \quad \text{or} \quad C_1 := \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \lambda & \\ & & & 1 & \\ & & & & \ddots \end{bmatrix} \in \mathbb{F}^{n \times n}$$

C2: Row swap

$$C_2 := \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & 1 & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \in \mathbb{F}^{n \times n}$$

C3: Row Scaling

$$C_3 := \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \in \mathbb{F}^{n \times n}$$

Each of them also encodes the row Operations from where they came from. So now Gaussian Elimination can be encoded as matrix multiplication.

38.1 Invertibility

Because every row operation is invertible now we can also claim that Elementary Matrices are invertible and thy also represented that inverted row operation.

38.2 Determinants

The determinants corresponding to each matrix are:

$$\det C1 = 1$$
$$\det C2 = -1$$
$$\det C3 = \lambda$$

39 Geometry of Linear Maps

In this section we are going to take a look at some special type of linear maps that are called *orthogonal* and at the term *isometry*. It will be useful to remember some concepts about the dot product such as $\langle x, x \rangle = x^T \cdot x = \|x\|^2$.

39.1 Isometry

Given two vector spaces $(V, \| \cdot \|_V)$ and $(W, \| \cdot \|_W)$. Then a map $f : V \to W$ is called *Isometric* if and only if $\|x - y\|_V = \|f(x) - f(y)\|_W$.

It also has the following traits:

1. *Length-Preserving*: $\|x\|_V = \|y\|_W \quad \forall x \in V$.

2. *Angle-Preserving*: $\angle(x, y) = \angle(f(x), f(y)) \quad \forall x, y \in V$

3. *Linear*: If $f(x + \lambda y) = f(x) + \lambda f(y)$

39.2 The Rotation Matrix

This matrix is a common example of an isometric linear map, because of the way it transforms the space.

For the $2D$ we have

$$\begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$$

The derivation comes from the addition theorems for $\sin(\phi)$ and $\cos(\phi)$.

Given two vectors $\vec{v}$ and $\vec{g}$ whose angles are α and $\theta = \alpha + \beta$. We can rotate the vector $\vec{v}$ by some β to transform into $\vec{g}$.

$$x = r \cos(\alpha)$$
$$y = r \sin(\alpha)$$

Then

$$x' = r' \cos(\alpha + \beta)$$
$$y' = r' \sin(\alpha + \beta)$$

And therefore

$$x' = r' \cos(\alpha) \cos(\beta) - r' \sin(\alpha) \sin(\beta)$$
$$y' = r' \cos(\alpha) \sin(\beta) + r' \cos(\beta) \sin(\alpha)$$

Because, a rotation does not change the magnitude we can say that $r = r'$. Thus

$$x' = r \cos(\alpha) \cos(\beta) - r \sin(\alpha) \sin(\beta)$$
$$y' = r \cos(\alpha) \sin(\beta) + r \cos(\beta) \sin(\alpha)$$

Now notice that we have our x and y inside the expression

$$x' = x\cos(\beta) - y\sin(\beta)$$
$$y' = x\sin(\beta) + y\cos(\beta)$$

which gives us

$$\begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$\mathfrak{QED}$

39.2.1 Other rotation matrices

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix}$$

$$R_y(\phi) = \begin{bmatrix} 1 & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix}$$

$$R_z(\phi) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

39.3 Orthogonal Matrices

A matrix $A \in \mathbb{R}^{n \times n}$ is *orthogonal* if its columns vectors form an orthonormal basis.

The set of all *orthogonal* matrices is $O(n)$.

39.3.1 Properties of orthogonal matrices

1. $A \in O(n)$
2. If A has an inverse then $A^{-1} = A^T$
3. $A^T \in O(n)$

Proof 1-2:

$$\begin{aligned} A^T = A^{-1} &\iff A^T A = E \\ &\iff (A^T A)_{ij} = \langle a_i, a_j \rangle = \delta_{ij} 1 \leq i, j \leq n \\ &\iff a_i \perp a_j, i \neq j, \text{ and } \|a_i\| = 1, 1 \leq i \leq n \\ &\iff A \in O(n) \end{aligned}$$

Proof 1-3:

$$\begin{aligned} A \in O(n) &\iff \\ &\iff A^{-1} = A^T \\ &\iff (A^T)^{-1} = (A^T)^T \\ &\iff (A^1)^T = A \\ &\iff A^T \in O(n) \end{aligned}$$

Example of a proof

Prove that H_n is *orthogonal* for all $\vec{u} \in \mathbb{R}^n \backslash \{0\}$.

$$H_n = I_n - 2\frac{uu^T}{u^T u}$$

I_n is the identity matrix.

Proof:

$$
\begin{aligned}
(H_n)^2 &= I \\
&= \left(I_n - 2\frac{uu^T}{u^T u}\right)^2 \\
&= I_n^2 - 4\frac{uu^T}{u^T u} + 4\left(\frac{uu^T}{u^T u}\right)^2 \\
&= I_n - 4\frac{uu^T}{u^T u} + 4\frac{uu^T uu^T}{(u^T u)^2} \\
&= I_n - 4\frac{uu^T}{u^T u} + 4\frac{u(u^T u)u^T}{(u^T u)^2} \\
&= I_n - 4\frac{uu^T}{u^T u} + 4\frac{uu^T}{u^T u} \\
&= I_n
\end{aligned}
$$

$\mathfrak{QED}$

39.4   The Determinants of Orthogonal Matrices

The *determinant* of this kind of matrices can only be either positioning 1 or negative 1.

Proof::

We know that $\det(I) = I$, therefore

$$
\begin{aligned}
\det(I) &= \det(AA^{-1}) \\
&= \det(AA^T) \\
&= \det(A)\det(A^T) \\
&= (\det(A))^2 \\
&= \|\det(A)\| = 1
\end{aligned}
$$

$\mathfrak{QED}$

39.5   The Orthogonal Group

$O(n)$ with the operation matrix multiplication forms the *orthogonal group*.

39.6   Dot Product

Given $f : \mathbb{R}^n \to \mathbb{R}^n$ then this two statements are equivalent.

$$\langle f(x), f(y)\rangle = \langle x, y\rangle \quad \forall x, y \in \mathbb{R}^n$$

and

f preserves the angles and length.

f is also called *orthogonal*.

Proof ⇒:

$$\|x\|^2 = \langle x, x \rangle = \langle f(x), f(x) \rangle = \|f(x)\|^2$$

This implies the length preservation. The isometry gives us:

$$\begin{aligned}
\|f(x) - f(y)\|^2 &= \langle f(x) - f(y), f(x) - f(y) \rangle \\
&= \langle f(x), f(x) \rangle - 2\langle f(x), f(y) \rangle + \langle f(y), f(y) \rangle \\
&= \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle \\
&= \langle x - y, x - y \rangle = \|x - y\|^2
\end{aligned}$$

Then

$$\begin{aligned}
\angle(f(x), f(y)) &= \arccos\left(\frac{\langle f(x), f(y) \rangle}{\|f(x)\| \cdot \|f(y)\|} \right) \\
&= \arccos\left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right) \\
&= \angle(x, y)
\end{aligned}$$

$$\mathfrak{QED}$$

Proof ⇐:

With $\alpha = \angle(x, y) = \angle(f(x), f(y))$ gilt:

$$\langle f(x), f(y) \rangle = \cos(\alpha) \cdot \|f(x)\| \cdot \|f(y)\| = \cos(\alpha) \cdot \|x\| \cdot \|y\| = \langle x, y \rangle$$

$$\mathfrak{QED}$$

39.7   Theorem of the Matrix Representation

The mapping $f : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if the associated matrix representation Q (with respect to the standard basis) is an orthogonal matrix, i.e., $Q \in O(n)$.

Proof ⇒:

f is linear, and thus has a matrix representation Q. It follows that

$$\langle Qx, Qy \rangle = \langle Q^\top Qx, y \rangle = \langle x, y \rangle \quad \forall x, y \in \mathbb{R}^n,$$

thus,

$$\langle Q^\top Qx - x, y \rangle = 0 \quad \forall x, y.$$

Choosing $y = Q^\top Qx - x$, we get

$$\langle Q^\top Qx - x, Q^\top Qx - x \rangle = 0 \quad \forall x,$$

and hence $(Q^\top Q - I)x = 0 \quad \forall x$. It follows that $Q^\top Q = I$, and by Theorem 7.3, $Q \in O(n)$.

$$\mathfrak{QED}$$

Proof ⇐:

Follows by the same calculation as above.

$$\mathfrak{QED}$$

39.8 Change of Basis

Let f be orthogonal. Then the matrix representation of f with respect to any orthonormal basis is orthogonal.

Proof:

Let E be the standard basis and B an orthonormal basis. Let $Q = M_E^E(f)$ and $S = T_E^B$ be the change-of-basis matrix from B to E. Since the columns of S are the basis vectors of B, we have $S \in O(n)$.

$$M_B^B(f) = S^{-1}QS \in O(n),$$

due to the group property of $O(n)$.

$$\mathfrak{QED}$$

39.9 QR-Decomposition

A use case of orthogonal matrices is the *QR-Decomposition*. The goal is that given a system $Ax = b$ with a matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$ and $rg(A) = n$ we can find two matrices Q and R such that $A = QR$ with $R \in \mathbb{R}^{n \times n}$ and $Q \in \mathbb{R}^{m \times n}$. R is matrix with only an upper triangle and Q is orthogonal for the columns.

We know that the best approximation by the method of the least squares is given by

$$x_s = (A^T A)^{-1} A^T b$$

Because of the orthogonality of the columns $rg(Q) = rg(A)$ therefore it can be inverted. Therefore, we get:

$$\begin{aligned}
x_s &= (A^T A)^{-1} A^T b \\
&= ((QR)^T (QR))^{-1}(QR)^t b \\
&= (R^T Q^T QR)^{-1}(QR)^t b \\
&= (R^T R)^{-1} R^T Q^T b \\
&= R^{-1}(R^T)^{-1} R^T Q^T b \\
&= R^{-1} Q^T b \\
Rx_s &= Q^T b
\end{aligned}$$

Thus, we can solve the linear system by finding Q which is matrix with an orthonormal basis of the vector space which can be found via the Gram-Schmidt algorithm. The matrix R consists of the matrix of the inner products our A matrix and Q

$$R = \begin{bmatrix} \langle q_1, a_1 \rangle & \langle q_1, a_2 \rangle & \cdots & \langle q_1, a_n \rangle \\ 0 & \langle q_2, a_2 \rangle & \cdots & \langle q_2, a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle q_n, a_n \rangle \end{bmatrix}$$

Or even simpler $R = Q^T A$.

Example:

Find the *QR-Decomposition* of the matrix A given by:

$$\begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$$

First, we find the orthonormal basis of the columns of our matrix A

$$q_1 = \frac{a_1}{\|a_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$r_2 = a_2 - \langle a_2, q_1 \rangle q_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \langle \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rangle \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$q_2 = \frac{1}{\sqrt{8}} \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{8}} \\ \frac{1}{\sqrt{2}} & -\frac{2}{\sqrt{8}} \end{bmatrix}$$

$$R = \begin{bmatrix} \sqrt{2} & sqrt2 \\ 0 & \sqrt{8} \end{bmatrix}$$

40 Eigenvectors and Eigenvalues

Eigenvector:

An eigenvector of a square matrix A is a non-zero vector $\vec{v}$ that, when multiplied by A, results in a vector that is a scalar multiple of itself. In other words, the direction of the vector **v** remains unchanged (up to scaling) when the linear transformation represented by A is applied to it.

Eigenvalue:

The scalar multiple, denoted by λ, is called the eigenvalue associated with the eigenvector **v**. It represents the factor by which the eigenvector is scaled when transformed by the matrix A.

Mathematically, the relationship between a square matrix A, an eigenvector $\vec{v}$, and its corresponding eigenvalue λ is expressed by the following equation:

$$A\vec{v} = \lambda\vec{v}$$

Where $A = f : V \to V$ is an *endomorphism*, $\vec{v} \neq \vec{0}$ and $\lambda \in \mathbb{C}$.

The way we find a method to find them is by manipulating the equation above

$$A\vec{v} = \lambda\vec{v}$$
$$A\vec{v} - \lambda\vec{v} = \vec{0}$$
$$A\vec{v} - \lambda I\vec{v} = \vec{0}$$
$$(A\vec{v} - \lambda I)\vec{v} = \vec{0}$$

Here note that if $(A\vec{v} - \lambda I)$ were invertible then we could multiply both sides by its inverse and then get $\vec{v} = 0$. But we do not want that therefore, $(A\vec{v} - \lambda I)$ must be linear transformation whose determinant is 0 which means that it can not ve invertible.

$$(A\vec{v} - \lambda I)\vec{v} = \vec{0}$$
$$\det(A\vec{v} - \lambda I) = 0$$

40.1 Characteristic Polynomial

The polynomial which we get from the determinant $\det(A\vec{v} - \lambda I) = 0$ is called the characteristic polynomial of A.

$$\mathcal{X}_A = \det(A\vec{v} - \lambda I) = 0$$

40.1.1 Eigenspace

The *Eigenspace* of an eigenvalue are the solution of the homogeneous system $A - \lambda I$.

$$Eig(f; \lambda) = ker(A - \lambda I)$$

40.2 Theorems of Eigenvalues and Eigenvectors

40.2.1 Theorem I

Given an eigenvalue λ of f and $v_1, \ldots, v_n$ being the eigenvectors of f with λ then $v \in L(v_1, \ldots, v_k)\backslash\{0\}$ is an eigenvector of f with λ.

40.2.2 Theorem II

For $\lambda \in \mathbb{C}$ is $Eig(f; \lambda) := \{c \in V | f(v) = \lambda v\}$ the *Eigenspace* of f with λ and, it is a subspace of V.

40.2.3   Theorem III

For $\lambda \neq \gamma \quad Eig(f; \lambda) \cap Eig(f; \gamma) = \{0\}$.

40.2.4   Theorem IV

Eigenvectors with different eigenvalues are linearly independent.

40.2.5   Theorem V

Given the following eigenvalues $\lambda_1, \ldots, \lambda_n$, then the determinant of the matrix A is given by

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

Proof:

Given the eigenvalues $\lambda_1, \ldots, \lambda_n$ are the roots of the characteristic polynomial

$$\begin{aligned}
\det(A - \lambda I) = p(\lambda) &= (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots \\
&= (-1)(\lambda - \lambda_1)(-1)(\lambda - \lambda_2) \cdots \\
&= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots
\end{aligned}$$

Now, by setting $\lambda = 0$ we get

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

Or by using the definition of the diagonal matrix:

$$\begin{aligned}
A &= XDX^{-1} \\
\det(A) &= \det(XDX^{-1}) \\
\det(A) &= \det(X)\det(D)\det(X^{-1}) \\
\det(A) &= \det(D) = \prod_{i=1}^{n} \lambda_i
\end{aligned}$$

$$\mathfrak{QED}$$

40.2.6   Theorem VI

For all linear isometric map f with some eigenvalue λ it applies that

$$\|x\| = \|f(x)\| = \|\lambda x\| = \|\lambda\|\|x\|$$

Also, $\lambda = 1$.

An example for this are rotation matrices which map to the same space and preserve the length and whose eigenvalue are complex with a modulus of length of one.

40.2.7   Theorem VII

For some eigenvector x with an eigenvalue λ of A. It applies that

$$A^k x = \lambda^k x$$

40.3 How to find the Eigenvectors and Eigenvalues

To find the eigenvalues and eigenvectors of a square matrix A, we solve the eigenvalue equation:

1. Form the characteristic equation:

Rewrite the equation $A\vec{v} = \lambda\vec{v}$ as $(A - \lambda I)\vec{v} = \vec{0}$, where λI is the identity matrix times λ because this matrix encodes the multiplication by some scalar.

To have a non-trivial solution 0 for $\vec{v}$, the matrix $(A - \lambda I)$ must be singular like said before, which means its determinant must be zero because otherwise $\vec{v}$ would be 0. Thus, we have the characteristic equation:

$$\det(A - \lambda I) = 0$$

2. Solve for the eigenvalues:

Solve the characteristic equation for λ. The solutions $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of the matrix A.

3. Find the eigenvectors:

For each eigenvalue λ_i, substitute it back into the equation $(A - \lambda_i I)\vec{v} = \vec{0}$ and solve for the vector $\vec{v}$. The non-zero solutions for $\vec{v}$ are the eigenvectors corresponding to the eigenvalue λ_i. Note that we are going to get infinite solutions thus, we have to write our vector with dependence on the free parameters.

Example:

Find the eigenvalues and vectors of

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$$

Let us build the characteristic equation and solve it

$$\det \begin{bmatrix} 1 - \lambda & 1 \\ 4 & 1 - \lambda \end{bmatrix} = 0$$

$$(1 - \lambda)^2 - 4 = 0$$
$$1 - 2\lambda + \lambda^2 - 4 = 0$$
$$\lambda^2 - 2\lambda - 3 = 0$$
$$(\lambda - 3)(\lambda + 1) = 0$$

We have got the eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -1$. Now let us continue by substituting our eigenvalues in the original equation.

$$\begin{bmatrix} 1 - \lambda_2 & 1 \\ 4 & 1 - \lambda - 2 \end{bmatrix} = \vec{0}$$
$$\begin{bmatrix} 1 - 2 & 1 \\ 4 & 1 - 2 \end{bmatrix} = \vec{0}$$
$$\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \vec{0}$$

Using Gaussian elimination we get:

$$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

Thus, let $v_2 = -2$ and $v_1 = \frac{2}{2} = 1$. This give us the first eigenvector and all of it scalar versions.

$$e_1 = \left\{ c \begin{bmatrix} 1 \\ -2 \end{bmatrix} \,|\, c \in \mathbb{R} \right\}$$

The process for the other eigenvalue is exactly the same.

40.3.1 Complex Eigenvalues

This is a special case of *eigenvalues* which are not as easy to visualize.

Example:

Find the eigenvalues and vectors of

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\det \left(\begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} \right) = 0$$

$$\lambda^2 + 1 = 0$$

$$\lambda = \pm i$$

Now the corresponding eigenvectors.

For $\lambda = i$

$$\begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \rightarrow \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

Therefore, our first eigenvector is

$$c \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

For $\lambda = -i$

$$\begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$$

Therefore, the second eigenvector is

$$c \begin{bmatrix} i \\ 1 \end{bmatrix}$$

40.4 Similar Matrices

$A \sim B$ if there is some X such that

$$A = XDX^{-1} \iff X^{-1}AX = D$$

If our matrices are really similar then:

$$\det(A - \lambda I) = \det(XDX^{-1} - \lambda I)$$
$$\det(A - \lambda I) = \det(XDX^{-1} - \lambda XIX^{-1})$$
$$\det(A - \lambda I) = \det(X[D - \lambda I]X^{-1})$$
$$\det(A - \lambda I) = \det(X)\det(D - \lambda X)\det(X^{-1})$$
$$\det(A - \lambda I) = \det(D - \lambda X)$$

Thus, similar matrices have the same eigenvalues with multiplicity.

40.5 Diagonal Matrices

A *diagonal matrix* is matrix of dimensions $n \times n$ where only the entries of the main diagonal are not zero. This kind of matrices make a lot of the computations in linear algebra trivial. Therefore, if our matrix is not diagonal we are interested in finding a way of transforming the original matrix into a diagonal matrix.

A matrix is diagonalizable if and only if the eigenvectors of it are linear independent.

40.5.1 Complex Diagonalization

Given a matrix $A \in \mathbb{C}^{n \times n}$ with n distinct eigenvalues then is A diagonalizable. But this does not guaranty the diagonalizibility for $\mathbb{R}$.

40.5.2 Geometric and Algebraic Multiplicity

Given a matrix $A \in \mathbb{C}^{n \times n}$ with λ as eigenvalue. The multiple occurrences of the zero λ of $\mathcal{X}_A$ is called *algebraic multiplicity* $\text{alg}(\lambda)$. Continuing, $\text{geo}(\lambda) := dim(Eig(A; \lambda))$ is the *geometrical multiplicity* of λ.

Also, A is diagonalizable if and only if the *geometric multiplicity* of every eigenvalue is equal to its *algebraic multiplicity*.

Example:

This example shows a matrix which is not diagonalizable:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The corresponding multiplicities are:

$$\mathcal{X} = \det(A - \lambda I) = 0$$
$$(1 - \lambda)^2 = 0$$

$\text{alg}(1) = 2$ and the $\text{geo}(1) = 1$ because of

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

More precisely $\dim(Eig(A; 1)) = 1$.

Example of Complex Eigenvalues:

Given is the rotation matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

With the characteristic polynomial

$$\mathcal{X} = (\cos(\theta) - \lambda)^2 + \sin^2(\theta) = 0$$
$$= \cos^2(\theta) - 2\cos(\theta)\lambda + \lambda^2 + \sin^2(\theta) = 0$$
$$= \lambda^2 - 2\cos(\theta)\lambda + 1 = 0$$

Using the quadratic formula gives us

$$\lambda_{1,2} = \cos(\theta) \pm \sqrt{\cos^2(\theta) - 1}$$
$$= \cos(\theta) \pm \sqrt{-\sin^2(\theta)}$$
$$= \cos(\theta) \pm i\sin(\theta) = e^{\pm i\theta}$$

Therefore, our eigenvalues are complex. Now, is the matrix diagonalizable? The answer is yes. Because of $(\cos(\theta) - \lambda)^2 \, \text{alg}(\lambda) = 2$

$$\text{alg}(e^{i\theta}) = 2, \, \text{alg}(e^{-i\theta}) = 2$$

And

$$\text{geo}(e^{i\theta}) = 2, \, \text{geo}(e^{-i\theta}) = 2$$

For $\lambda = e^{i\theta}$ we get two linearly independent vectors

$$\begin{bmatrix} \cos(\theta) - \cos(\theta) - i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - \cos(\theta) - i\sin(\theta) \end{bmatrix}$$

$$\begin{bmatrix} -i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & -i\sin(\theta) \end{bmatrix}$$

For $\lambda = e^{-i\theta}$ we also get two linearly independent vectors

$$\begin{bmatrix} \cos(\theta) - \cos(\theta) + i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - \cos(\theta) + i\sin(\theta) \end{bmatrix}$$

$$\begin{bmatrix} i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & i\sin(\theta) \end{bmatrix}$$

40.5.3 Eigenvalues of Real Symmetric Matrices

Eigenvalues of real symmetric matrices are real and also they respective eigenvectors.

Proof:

By definition, we have that $Ax = \lambda x$. The let us multiply both sides by $\bar{x}^T$. Note that because x is a real vector then the conjugate does not change anything.

$$Ax = \lambda x$$
$$\bar{x}^T A x = \bar{x}^T \lambda x$$
$$\bar{x}^T A x = \lambda \bar{x}^T x$$
$$\bar{x}^T A x = \lambda \langle \bar{x} x \rangle$$
$$\bar{x}^T A x = \lambda \|x\|^2$$
$$\frac{\bar{x}^T A x}{\|x\|^2} = \lambda$$

With this we know that the denominator is a real number thus, if we prove that the numerator is also real we are done. For this we will use the properties of the conjugate.

$$\begin{aligned}
\bar{x}^T A x &= \overline{\bar{x}^T A x} \\
&= \overline{\bar{x}^T}\, \overline{Ax} \\
&= x^T\, \overline{Ax} \\
&= (\overline{Ax})^T x \\
&= (\bar{A}\,\bar{x})^T x \\
&= \bar{x}^T \bar{A}^T x
\end{aligned}$$

Notice that we ended just were we started. Therefore, the conjugate had no effect on whole term, thus it is real-valued.

$$\mathfrak{QED}$$

40.5.4 Orthogonality of Eigenvectors

Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with $\lambda \neq \mu$ as two eigenvalues of A with eigenvectors v and w. Then v and w are *orthogonal* to each other.

Proof:

$$\langle w, Av \rangle = w^T A v = (A^T w)^T v = (Aw)^T v = \langle Aw, v \rangle$$

Therefore, $\lambda \langle v, w \rangle = \langle Av, w \rangle = \langle v, Aw \rangle = \mu \langle v, w \rangle$. Because $\mu \neq \lambda$ and $\lambda - \mu \langle v, w \rangle = 0$ the vectors must be orthogonal.

$$\mathfrak{QED}$$

40.5.5 Eigenvalue Decomposition of Symmetric Matrices

Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then because of the orthogonality of its eigenvectors and the fact that they can be normalized write the matrix Q as an orthogonal matrix, and therefore

$$A = QDQ^T$$

40.5.6 Matrix Diagonalization

Diagonalizing a matrix involves finding a diagonal matrix that is similar to the given matrix. A square matrix A is diagonalizable if there exists an invertible matrix X made of eigenvectors of A such that $X^{-1}AX = D$, where D is a diagonal matrix made of eigenvalues.

The order of the eigenvalue and eigenvectors matters if you choose $\lambda_1, \lambda_2, \ldots, \lambda_n$ then the order of the eigenvectors $v_1, v_2, \ldots, v_n$ must also correspond. This process only works if A has unique eigenvalues or if they are duplicates they have to be linearly independent.

The process of diagonalization is as follows:

1. Find the eigenvalues and eigenvectors of A.

2. Form the matrix X:

Create a matrix X whose columns are the linearly independent eigenvectors of A.

3. Form the diagonal matrix D:

Create a diagonal matrix D whose diagonal entries are the eigenvalues of A, corresponding to the order of the eigenvectors in P. That is, if the i-th column of P is the eigenvector corresponding to the eigenvalue λ_i, then the i-th diagonal entry of D is λ_i.

4. Verify the diagonalization:

Check that $X^{-1}AX = D$.

A matrix A is diagonalizable if and only if it has n linearly independent eigenvectors, where n is the size of the matrix.

Example:

Consider the matrix

$$A = \begin{bmatrix} -3 & -4 \\ 5 & 6 \end{bmatrix}$$

The characteristic equation is

$$\det(A - \lambda I) = \det \begin{bmatrix} -3-\lambda & -4 \\ 5 & 6-\lambda \end{bmatrix} = (-3-\lambda)(6-\lambda) + 20 = \lambda^2 - 3\lambda + 2 = 0$$

$$(\lambda - 1)(\lambda - 2)$$

Solving for λ, we get $\lambda_1 = 1$ and $\lambda_2 = 2$.

For $\lambda_1 = 1$:

$$(A - I)\vec{v} = \begin{bmatrix} -4 & -4 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -4 & -4 & | & 0 \\ 5 & 5 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & | & 0 \\ 5 & 5 & | & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$$

Our free variable is y therefore

$$\vec{v} = \left\{ y \begin{bmatrix} -1 \\ 1 \end{bmatrix} \mid y \in \mathbb{R} \right\}$$

Now choose $y = 1$ which gives us $x = -1$. Now we get the vector

$$\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 2$ the process is similar, so we are going to skip it. We get the vector

$$\vec{y} = \begin{bmatrix} -\frac{4}{5} \\ 1 \end{bmatrix}$$

Now we will complete the process by writing $D, X,$ and X^{-1}.

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} X = \begin{bmatrix} -1 & -\frac{4}{5} \\ 1 & 1 \end{bmatrix} X^{1} = \begin{bmatrix} -1 & \frac{4}{5} \\ -1 & -1 \end{bmatrix}$$

40.5.7   Normal Matrices

A matrix $A \in \mathbb{C}^{n \times n}$ is considered *normal* if and only if

$$A\overline{A}^T = \overline{A}^T A$$

Also, then there exist with the euclidean inner product in $\mathbb{C}^n$ an orthonormal basis of eigenvector and therefore, A is diagonalizable.

This implies that

1. All real symmetric matrices are diagonalizable, and their eigenvalues are real.

2. All real antisymmetric matrices $A^T = -A$ are diagonalizable, and their eigenvalues are purely imaginary or 0.

3. All orthogonal matrices are diagonalizable.

40.6   Singular Value Decomposition (SVD)

The *Singular Value Decomposition (SVD)* is a fundamental matrix factorization technique in linear algebra. It expresses any real matrix $A \in \mathbb{R}^{n \times n}$ as a product of three specific matrices:

$$A = USV^T$$

where:

- $U \in \mathbb{R}^{n \times n}$ is an orthogonal matrix whose columns are called the *left singular vectors* of A,

- $V \in \mathbb{R}^{n \times n}$ is an orthogonal matrix whose columns are the *right singular vectors* of A,

- $S \in \mathbb{R}^{n \times n}$ is a diagonal matrix with non-negative entries $\sigma_i \geq 0$ on the diagonal, called the *singular values* of A.

The SVD has numerous applications in mathematics, science, and engineering, including:

- Dimensionality reduction (e.g., Principal Component Analysis),

- Solving ill-posed or overdetermined systems of linear equations,

- Low-rank matrix approximations,

- Signal processing and image compression,

- Latent Semantic Analysis (LSA) in natural language processing.

Example:

Let $A \in \mathbb{R}^{n \times n}$ be a real square matrix with SVD given by:

$$A = USV^T$$

We show that:

1. The columns of U are eigenvectors of AA^T,

2. The columns of V are eigenvectors of $A^T A$,

3. The eigenvalues of both AA^T and $A^T A$ are the squares of the singular values σ_i^2.

Proof:

For AA^T, we compute:

$$AA^T = (USV^T)(VS^TU^T) = USV^TVS^TU^T = USS^TU^T = US^2U^T$$

Since SS^T is diagonal with entries σ_i^2, this is an eigenvalue decomposition of AA^T. Thus, the columns of U are eigenvectors of AA^T, and the corresponding eigenvalues are σ_i^2.

For A^TA, we have:

$$A^TA = (VS^TU^T)(USV^T) = VS^TU^TUSV^T = VS^TSV^T = VS^2V^T$$

Again, S^TS is diagonal with entries σ_i^2, so the matrix A^TA is diagonalized by V, and the columns of V are eigenvectors of A^TA, with eigenvalues σ_i^2.

<div align="right">𝔔𝔈𝔇</div>

41 Dual Spaces

Given a $\mathbb{F}$-vector space V, we define V^* *Dual Space* of V, as the set of linear transformations from $f : V \to K$. f is called the *Linear Functional*.

Example:

$V = \mathbb{R}^3$, with $f : V \to \mathbb{R}$ and $f(x, y, z) = x + y + z$

41.1 The Dual Basis

Given a finite dimensional $\mathbb{F}$-vector space V. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis of V let $f \in V*$. The set of functionals $\beta^* = \{f_1, \ldots, f_n\}$ of V^*. Defined by:

$$f_i(v_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases} = \delta_{ij}$$

Is called a *Dual Basis*. The definition above is saying that only one of the basis vector of V is going to be associated with a function of the basis β.

Proof:

To prove a set of vectors is basis we have to prove that they are linearly independent and that they span the entire vector space.

Linear Independence:

We want that

$$\sum_{i=1}^{n} \lambda_i f_i(v) = 0 \text{ only if all } \lambda_i = 0 \text{ for all } v \in V$$

Now just choose n-times $v = v_j, 1 \leq j \leq n$ and you are going to realize that because $f_i(v_j) = 1$ only if $j = i$ you are going to get $\lambda_j f_j(v_j) + 0 + \cdots + 0 = 0$. Therefore, by repeating this process you will arrive at the conclusion that all lambdas have to zero. Thus, the linear Independence is shown.

Span:

Suppose $f \in V^*$ is arbitrary and $f = \sum_{i=1}^{n} \lambda_i f_i(v)$ for all $v \in V$. Using the previously collected knowledge, which tells us that for $v = v_j$ the linear combination of $\sum_{i=1}^{n} \lambda_i f_i(v_j)$ is

$$\sum_{i=1}^{n} \lambda_i f_i(v_j) = \lambda_j f_j(v_j) = \lambda_j$$

Now let us guess $f = f(v_1)f_1 + \cdots f(v_n)f_n = g$. If $f(v_i) = g(v_i)$ for all basis vector of V.

Note for

$$g(v_i) = f(v_1)f_1(v_i) + \cdots + f(v_i)f_i(v_i) + \cdots f(v_i)f_n(v_i)$$

All terms are zero except for $f(v_i)f_i(v_i)$ which equals $f(v_i) = g(v_i)$.

$$\mathfrak{QED}$$

41.2 How to find a Dual Basis

We know that to find the coefficients of a linear combination $f = a_1 f_1 + a_2 f_2 + \cdots + a_n f_n$ in the space $V*$ with the basis $\beta = \{f_1, \ldots, f_n\}$ we just need to evaluate the function f at each of the basis vectors of the space V which are $\{v_1, \ldots, v_n\}$.

The process for solving this kind of problems can be summarized into the following steps using multiple systems of equations.

1. For each f_i repeat the following steps except for the last one.

2. Use the linearity property of vector to write non-canonical basis vectors as a linear combination $f_i(\vec{k}) = \sum \lambda_k f_i(\vec{v}_k)$, do this for every vector in the basis of V with the current f_i. This will set up the left-hand side of the system of equations.

3. Use the definition of the dual basis for setting up the right-hand side of linear system

4. Solve the system of equations

5. If a f was given then use the found basis to write $f = \sum_{i=1}^{n} f(v_i) f_i$

Example:

Given the space $V = \mathbb{R}^n$ with the basis $\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$ find $\beta^* = \{f_1, f_2\}$.

Because of linearity

$$f_1(2,1) = 2f_1(\vec{e}_1) + 1f_1(\vec{e}_2) = 1$$
$$f_1(3,1) = 3f_1(\vec{e}_1) + 1f_1(\vec{e}_2) = 0$$

We get a linear system of equations. By solving the system we get $f_1(1,0) = -1$ and $f_2(0,1) = 3$. Thus, $f_1(x,y) = x(-1) + y(3)$.

And for f_2 it is the same process

$$f_2(2,1) = 2f_1(\vec{e}_1) + 1f_1(\vec{e}_2) = 0$$
$$f_2(3,1) = 3f_1(\vec{e}_1) + 1f_1(\vec{e}_2) = 1$$

This time we get $f_2 = 2x - 5y$. Thus, our basis is $\beta^* = \{-x + 3y, 2x - 5y\}$, and if we let f be $f = 2x - 5$ we get that it can be decomposed into:

$$f = f(2,1)f_1 + f(3,1)f_2$$
$$= (-1)f_1 + 1f_2$$

41.3   Change of Basis in the Dual Space

Given some finite $\mathbb{F}$-vector space V with basis β and β'. Then matrix for the change of basis $\beta^* \to \beta'^*$ is given by

$$(T^{-1})^T = (T^T)^{-1}$$

Where $T = \beta^{-1}\beta'$.

Example:

Given are $\beta = \{(1,1); (-1,1)\}$ and $\beta' = \{(2,3), (1,1)\}$.

$$T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}$$

Then

$$(T^T)^{-1} = \begin{pmatrix} 0 & 1 \\ 2 & -5 \end{pmatrix}$$

41.4   Dimension Theorem

If the dimension of V is finite then V and V^* are *isomorphic*.

41.5   The Transpose of a Linear Transformation

Let V and W be finite dimensional vector spaces with the field K. Let $T : V \to W$ be a linear transformation. Now define

$$T^T : W^* \to V^*$$

with an input $g \in W^*$ where $g : W \to K$; and an output $T^T(g) \in V^*$ with $T^T(g) : V \to K$. Therefore, $T^T(g)(\vec{x}) = g(T(\vec{x}))$ for all $\vec{x} \in V$.

41.6   Connection to the Transpose Matrix

Let $T : V \to W$ be a linear transformation, $\beta = \{v_1, \ldots, v_n\}$ be a basis of V, $\gamma = \{w_1, \ldots, w_n\}$ be a basis of W. The matrix of the change of the transformation is given by $A = [T]_\beta^\gamma$ which takes vector from β to γ.

Then let $\beta^* = \{f_1, \ldots, f_n\}$ be the dual basis of V^* and $\gamma^* = \{g_1, \ldots, g_n\}$ the dual basis of W^*. Then

$$[T^T]_{\gamma^*}^{\beta^*} = A^T$$

Which takes functions from γ^* to β^*.

Proof:

Because $T(v_i) \in W$ we can write

$$T(v_i) = sum_{k=1}^n A_{ki} w_k$$

as a vector in terms of the basis γ.

For $T^T : W^* \to V^*$ we can write $T^T(g)$ as a linear combination of the basis vectors β^*

$$T^T(g_j) = \sum_{i=1}^n T^T(g_j)(v_i) f_i = \sum_{i=1}^n g(T(v_i)) f_i = \sum_{i=1}^n A_{ij} f_i$$

where $A = (T^T(g_1), \ldots, T^T(g_n)))$.

We have assumed that $T^T(g)(\vec{x}) = g(T(\vec{x}))$. Now let us apply this to any $v_j \in V$ and $g_i \in W^*$

$$T^T(g_j)(v_i) = g_j(T(v_i))$$

$$= g_j\left(\sum_{k=1}^n A_{ki} w_k\right) = \sum_{k=1}^n A_{ki} g_j(w_k)$$

$$= A_{ji}$$

This sum is equals to A_{ji} for when $g_j(w_j)$. Therefore, we have flipped the columns and rows. Thus, we have shown that

$$[T^T]_{\gamma^*}^{\beta^*} = A^T$$

Because we have shown it for the basis vectors it gets extended to all of it linear combinations.

$$\mathfrak{QED}$$

Example:

Let $P_1 \mathbb{R}$ be the space of polynomials of degree at most 1 and $W = \mathbb{R}^2$. Also, let $T : V \to W$ be defined by $T(p) = \begin{bmatrix} p(0) - 2p(1) \\ p(0) + p'(0) \end{bmatrix}$

For example $T(1 + 2x) = \begin{bmatrix} 1 - 2(1+2) \\ 1+2 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

If $f(a, b) = a - 2b \in W^*$ find $T^T(f)$.

The output of $T^T(f)$ is in V^* and because we are using the transpose we can write it as $T^T(f)(p) = f(T(p))$ and

$$
\begin{aligned}
f(a, b) &= a - 2b \text{ then} \\
f(T(p)) &= f((p(0) - 2p(1), p(0) + p'(0))) \\
&= (p(0) - 2p(1)) - 2(p(0) + p'(0)) \\
&= p(0) - 2p(1) - 2p(0) - 2p'(0) \\
T^T(f)(p) &= -p(0) - 2p(1) - 2p'(0)
\end{aligned}
$$

Now calculate the matrix of T and T^T. The basis of V is $\{1, x\}$ and the basis of W is $\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \}$.

First we know that $\beta^* = \{f_1, f_2\}$ and $\gamma^* = \{g_1, g_2\}$, and we remember that

$$
f_i(v_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases} = \delta_{ij}
$$

To find the matrix of T^T which is a 2×2

$$
T^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}
$$

We need to evaluate $T^T(g_1)$ and $T^T(g_2)$ and then write them in terms of β^*.

$$
T^T(g_1) = af_1 + cf_2
$$

Now we evaluate at the basis vectors of V

$$
T^T(g_1)(1) = af_1(1) + cf_2(1) = a
$$

Now we use the definition of T

$$
\begin{aligned}
g_1(T(1)) = g_1((\begin{bmatrix} 1 - 2 \\ 1 + 0 \end{bmatrix}) = g_1(\begin{bmatrix} -1 \\ 1 \end{bmatrix}) \\
= g_1(-1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}) = -1
\end{aligned}
$$

Thus, we have $a = -1$.

Now lets repeat this for x

$$
T^T(g_1)(x) = af_1(1) + cf_2(x) = c
$$

$$g_1(T(x)) = g_1((\begin{bmatrix} 0-2 \\ 0+1 \end{bmatrix}) = g_1(\begin{bmatrix} -2 \\ 1 \end{bmatrix})$$

$$= g_1(-2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = -2$$

Thus, we have $c = -2$.

Now if you repeat this for g_2 you will get $b = 1$ and $d = 0$.

Thus, the matrix of T^T is

$$T^T = \begin{bmatrix} -1 & 1 \\ -2 & 0 \end{bmatrix}$$

And the matrix of T is calculated by applying $T(1)$ and $T(x)$ and then writing them in terms of the basis γ.

This gives us

$$T(1) = \begin{bmatrix} 1-2 \\ 1+0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Which is $-1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

And

$$T(x) = \begin{bmatrix} 0-2 \\ 0+0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Which is $-2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Thus, the matrix of T is

$$T = \begin{bmatrix} -1 & -2 \\ 1 & 0 \end{bmatrix}$$

41.7   The Dual of the Dual Space

Given a dual space $V^* = L(V, \mathbb{F})$ then

$$(V^*)^* = L(V^*, \mathbb{F})$$

We also define the linear functional $\varphi : V \to V^{**}$ where $\varphi(x) = \hat{x}$ and $\hat{x}(f) = f(x)$ for all $f \in V$.

41.7.1   Isomorphism of the Dual Space

If V is finite dimensional then V^{**} and V are isomorphic.

41.7.2   Isomorphism from V to V^{**}

If V is finite dimensional then the map $\phi : V \to V^{**}$ is isomorphic.

41.8   Basis Theorem for the Dual Space

Every basis β^* of V^* is the dual basis of some basis β of V.

Proof:

Let $\beta^* = \{f_1, \ldots, f_n\}$ be a basis of V^*. We can define a basis β of V with help of the following:

Let us find a basis $\beta^{**} = \{\mathcal{F}_1, \ldots, \mathcal{F}_n\}$ of V^{**}. The functional $\varphi : V \to V^{**}$ is an isomorphism and $\mathcal{F}_i = \hat{v}_i$ for some $v_i \in V$.

Let us choose a basis $\beta = \{v_1, \ldots, v_n\}$ of V and use β^*. By definition

$$f_i(v_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases} = \delta_{ij}$$

Using the isomorphism φ we get

$$f_j(v_i) = \hat{v}_i(f_j) = \mathcal{F}_i(f_j) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases} = \delta_{ij}$$

<div align="right">𝔔𝔈𝔇</div>

41.9   Transpose of the Transpose is T

We define $T^{TT} V^{**} \to W^{**}$ by

$$T^{TT}(\hat{x}) = \hat{x}(T^T)$$

and claim $T^{TT} = T$.

Proof:

In this proof we know that $T^T : W^* \to V^*$. Now let us declare our basis $\beta = \{v_1, \ldots, v_n\}$ for V, $\gamma = \{w_1, \ldots, w_n\}$ for W, $\beta^* = \{f_1, \ldots, f_n\}$ for V^*, $\gamma^* = \{g_1, \ldots, g_n\}$ for W^*, $\beta^{**} = \{\mathcal{F}_1, \ldots, \mathcal{F}_n\}$ for V^{**} and finally $\gamma^{**} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ for W^**.

When we evaluate at $g : W \to \mathbb{F}$ we get

$$\begin{aligned} T^{TT}(\hat{x})(g) &= \hat{x}(T^T)(g) \\ &= \hat{x}(T^T(g)) \text{ by associativity} \\ &= T^T(g)(x) \text{ using the definition of } \hat{x} \\ &= g(T)(x) \text{ by associativity} \\ &= g(T(x)) \\ &= \hat{T(x)}(g) \end{aligned}$$

Here $\hat{T(x)} \in W^{**}$. Finally

$$\begin{aligned} T^{TT}(\hat{x})(g) &= \hat{T(x)}(g) \\ T^{TT}(\hat{x}) &= \hat{T(x)} \\ T^{TT} &= \hat{T} \\ T^{TT} &= T \text{ by ignoring the hats} \end{aligned}$$

They are not strictly the same, but because of the isomorphism we can set them equal.

<div align="right">𝔔𝔈𝔇</div>

41.10   The Annihilator

We define the *Annihilator* of a subspace of V as

$$W^\circ = \{f \in V^* \mid f(w) = 0, \forall w \in W\}$$

We also get the following formula

$$\dim(W^\circ) + \dim(W) = \dim(V)$$

whose proof is similar to the one of the rank nullity theorem.

Proof:

Let $\{w_1, \ldots, w_{k-1}, w_k, w_{k+1}, \ldots, w_n\}$ be a basis of W with dimension k. Let us extend the basis to a basis of V by $\beta = \{w_1, \ldots, w_{k-1}, w_k, w_{k+1}, \ldots, w_n\}$ with dimension n

Now let us claim that the vector $f_{k+1}, \ldots, f_n$ is a basis of W°. This implies that for every linear combination of vector from v_1 to v_k which equals w.

$$f_{k+1}(w) = \sum_{i=1}^{k} f_{k+1}(\lambda_i v_i) = 0$$

Therefore, $f_{k+1}, \ldots, f_n$ are in W°. Now we have to show that they are linearly independent.

For $f \in W^\circ$

$$f(x) = \sum_{i=0}^{n} f(v_i)f_i(x) = f(v_{k+1})f_{k+1}(x) + \ldots f(v_n)f_n(x)$$

Therefore we get that for some f in the annihilator we get a linear combination of the our chosen basis.

$$\mathfrak{QED}$$

41.11   The Duality Argument

$$T \text{ is surjective} \iff T^T \text{ is injective}$$

Proof:

Let us assume the opposite

$$\neg T \text{ is surjective} \iff \neg T^T \text{ is injective}$$

If T is not surjective we can assume that the image of it is a strict subset of W which we will call W'. We can also define a function $g \in W^*$ which is not 0 everywhere except for the vector in W'.

Then if we use the definition of T^T which tells us that when evaluating at g which is not the zero functional gives us

$$T^T(g)(x) = g(T(x)) = 0 \text{ beacause } T(x) \in W'$$

This concludes the proof in one way because we have found some some linear combination for the zero vector different from the trivial solution.

For the other way of the if and only if, we know that T^T is not injective. This means that multiple vector get transformed into the zero vector under T^T. Therefore, if suppose T is surjective and

$$T^T(g)(x) = 0 \text{ for some } g \neq 0$$
$$g(T(x)) = 0, \quad \forall x \in V$$

Then for $y = T(x)$ is always 0 but because we have assume that $g \neq 0$ we have a contradiction. T must not surjective and by getting rid of the negation on both sides we have proven the statement.

<div align="right">𝔔𝔈𝔇</div>

41.12 The Rank A transpose is equal to the rank of A

For this proof we will claim:

$$(T^T) = (\text{img}(T))^\circ$$

The right part tell us that for some $g \in (T^T)$, $T^T(g) = f_0$. But because we can write

$$T^T(g)(x) = g(T(x)) = 0 \text{ where } T(x) \in \text{img}(T)$$

We know that g is in the annihilator of the image of T.

With this knowledge we can know proof $\text{rank}(A) = \text{rank}(A^T)$ using the rank nullity theorem

$$\dim((T^T)) + \dim(\text{img}(T^T)) = \dim(W^*) = \dim(W)$$
$$\dim((\text{img}(T^\circ)) + \dim(\text{img}(T^T)) = \dim(W)$$

For $\dim(W)$ we get

$$\dim((\text{img}(T)^\circ)) + \dim(\text{img}(T)) = \dim(W)$$
$$\dim((\text{img}(T)^\circ)) = \dim(W) - \dim(\text{img}(T))$$

By substituting back into the original equation

$$\dim((\text{img}(T)^\circ)) + \dim(\text{img}(T^T)) = \dim(W)$$
$$\dim(W) - \dim(\text{img}(T)) + \dim(\text{img}(T^T)) = \dim(W)$$
$$\dim(\text{img}(T^T)) = \dim(\text{img}(T))$$

<div align="right">𝔔𝔈𝔇</div>

42 The Exponential Function and Relatives

The exponential function and its related functions such as sine, cosine, and the hyperbolic functions are fundamental in both pure and applied mathematics. This section explores these functions through their series expansions and interrelationships.

42.1 Exponential Function as a Series

The exponential function e^x is defined as an infinite series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

This series converges for all real and complex values of x.

42.2 Sine and Cosine Definitions with e

Using Euler's identities and Taylor series, sine and cosine can be defined in terms of exponential functions:

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}$$

These definitions are consistent with their Taylor expansions:

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

42.3 Euler's Formula

One of the most elegant formulas in mathematics is Euler's formula:

$$e^{ix} = \cos x + i \sin x$$

This identity provides a deep connection between exponential and trigonometric functions and is valid for all real (and complex) x.

Special Case:

$$e^{i\pi} + 1 = 0$$

This is known as Euler's Identity and is often celebrated as one of the most beautiful equations in mathematics.

42.4 Reason for the Infinite Series of e (Differentiation)

The exponential series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

has the remarkable property that its derivative is the same function:

$$\frac{d}{dx} e^x = \sum_{n=1}^{\infty} \frac{n x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!}$$

Now replace the index $m = n - 1$:

$$= \sum_{m=0}^{\infty} \frac{x^m}{m!} = e^x$$

This shows that e^x is the unique function (up to constant multiples) that is equal to its own derivative.

42.5 Hyperbolic Functions with e

The hyperbolic functions are analogs of the trigonometric functions but defined using the exponential function:

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

$$\cosh x = \frac{e^x + e^{-x}}{2}$$

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

These functions satisfy identities similar to those of trigonometric functions, for example:

$$\cosh^2 x - \sinh^2 x = 1$$

They naturally arise in the study of differential equations, special relativity, and the geometry of hyperbolas.

42.6 Inverse of the Hyperbolic Trigonometric Functions

- $\sinh^{-1}(x) = \ln(x + \sqrt{x^2 + 1})$
- $\cosh^{-1}(x) = \ln(x + \sqrt{x^2 - 1})$
- $\tanh^{-1}(x) = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)$
- $\coth^{-1}(x) = \frac{1}{2 \ln\left(\frac{x+1}{x-1}\right)}$
- $\text{sech}^{-1} = \ln\left(\frac{1+\sqrt{1-x^2}}{x}\right)$
- $\text{csch}^{-1} = \ln\left(\frac{1}{x} + \frac{\sqrt{1+x^2}}{|x|}\right)$

43 Growth

Growth models are fundamental in understanding population dynamics, resource consumption, and many natural and artificial processes. Two primary models of growth are the **exponential** and the **logistic** growth models. Each of these has discrete and continuous forms.

43.1 Exponential Growth

Exponential growth describes a process that increases at a rate proportional to its current value.

43.1.1 Discrete Exponential Growth

$$P_n = P_0 \cdot (1 + r)^n$$

Where:

- P_n is the population at time step n
- P_0 is the initial population
- r is the growth rate per time step
- n is the number of time steps

Example:

If $P_0 = 100$ and $r = 0.1$, after 5 steps:

$$P_5 = 100 \cdot (1 + 0.1)^5 = 100 \cdot 1.61051 \approx 161.05$$

43.1.2 Continuous Exponential Growth

$$P(t) = P_0 \cdot e^{rt}$$

Where:

- $P(t)$ is the population at time t
- e is Euler's number (≈ 2.718)
- r is the continuous growth rate

Example:

If $P_0 = 100$, $r = 0.1$, and $t = 5$:

$$P(5) = 100 \cdot e^{0.5} \approx 100 \cdot 1.64872 \approx 164.87$$

Continuous Exponential Growth
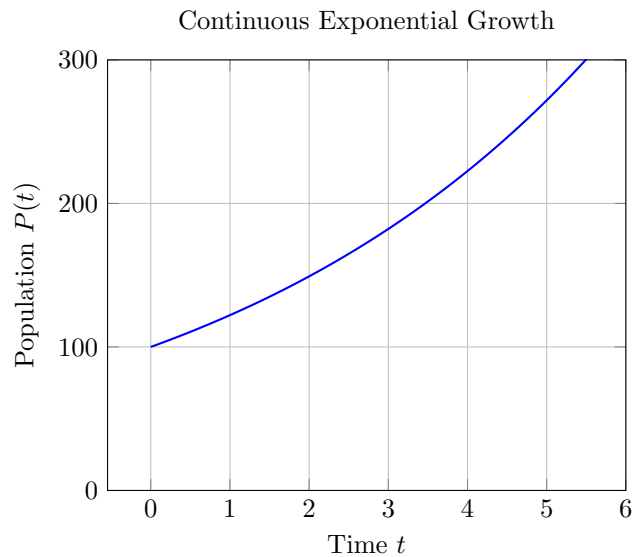
43.2 Logistic Growth

Logistic growth occurs when a population grows rapidly at first and then slows as it approaches a maximum sustainable size (carrying capacity).

43.2.1 Discrete Logistic Growth

$$P_{n+1} = P_n + r \cdot P_n \cdot \left(1 - \frac{P_n}{K}\right)$$

Where:

- K is the carrying capacity
- r is the intrinsic growth rate

Example:

With $P_0 = 10$, $r = 0.5$, $K = 100$:

$$P_1 = 10 + 0.5 \cdot 10 \cdot \left(1 - \frac{10}{100}\right) = 10 + 0.5 \cdot 10 \cdot 0.9 = 10 + 4.5 = 14.5$$

Discrete Logistic Growth
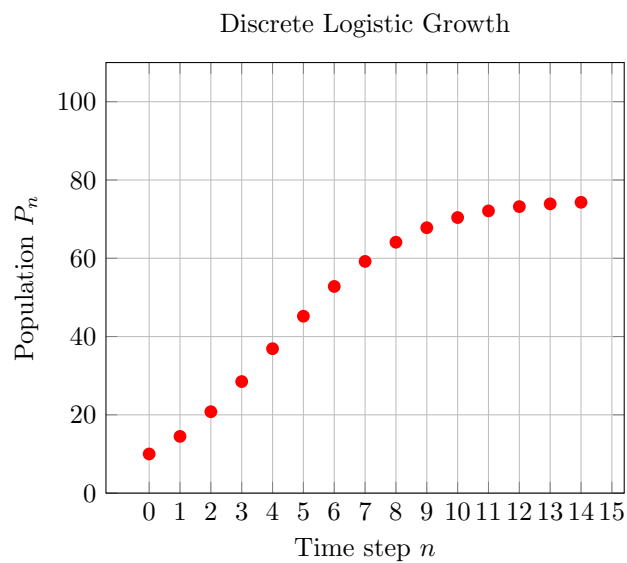
43.2.2 Continuous Logistic Growth

$$P(t) = \frac{K}{1 + \left(\frac{K - P_0}{P_0}\right) e^{-rt}}$$

Example:

With $P_0 = 10$, $K = 100$, and $r = 0.5$, at $t = 5$:

$$P(5) = \frac{100}{1 + \left(\frac{90}{10}\right) e^{-0.5 \cdot 5}} = \frac{100}{1 + 9e^{-2.5}} \approx \frac{100}{1.7387} \approx 57.52$$

Continuous Logistic Growth
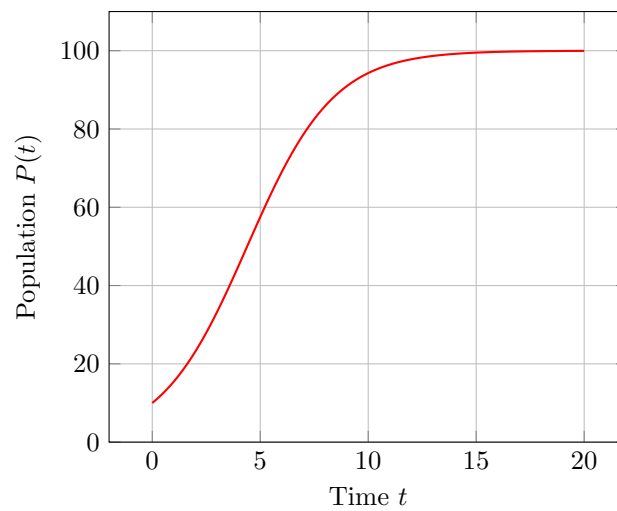
44   Asymptotes

An *asymptote* of a curve is a line that the curve approaches but never touches as the input or output grows infinitely large in magnitude. There are three main types of asymptotes: vertical, horizontal, and oblique (slant). Each one describes a different long-term behavior of a function.

44.1   Vertical Asymptotes

A *vertical asymptote* is a vertical line $x = a$ where the function grows without bound as x approaches a from the left or right.

$$\lim_{x \to a^-} f(x) = \pm\infty \quad \text{or} \quad \lim_{x \to a^+} f(x) = \pm\infty$$

44.1.1   How to Find Vertical Asymptotes

1. Find values of x that make the denominator of a rational function zero.

2. Eliminate removable discontinuities (common factors).

3. For each remaining zero in the denominator, check if the limit tends to infinity.

Example:

$$f(x) = \frac{1}{x - 2} \quad \Rightarrow \quad x = 2 \text{ is a vertical asymptote}$$

44.2   Horizontal Asymptotes

A *horizontal asymptote* is a horizontal line $y = L$ that the function approaches as x tends to infinity or negative infinity:

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

44.2.1   How to Find Horizontal Asymptotes (Rational Functions)

1. Compare the degrees of the numerator (n) and denominator (m):

   - If $n < m$: Horizontal asymptote at $y = 0$
   - If $n = m$: Horizontal asymptote at $y = \frac{\text{leading coefficient of numerator}}{\text{leading coefficient of denominator}}$
   - If $n > m$: No horizontal asymptote (check for slant)

Example:

$$f(x) = \frac{2x^2 + 3}{x^2 - 1} \quad \Rightarrow \quad y = \frac{2}{1} = 2 \text{ is the horizontal asymptote}$$

44.3   Oblique (Slant) Asymptotes

An *oblique or slant asymptote* occurs when the degree of the numerator is exactly one more than the degree of the denominator.

$$\lim_{x \to \infty} [f(x) - (mx + b)] = 0$$

44.3.1 How to Find Oblique Asymptotes

1. Perform polynomial long division (or synthetic division) on the rational function.

2. The quotient (ignoring the remainder) is the equation of the slant asymptote.

Example:

$$f(x) = \frac{x^2 + 1}{x - 1}$$

Long division gives:

$$\frac{x^2 + 1}{x - 1} = x + 1 + \frac{2}{x - 1} \Rightarrow \text{Slant asymptote: } y = x + 1$$

44.4 Summary Table

Type	Equation	Occurs When
Vertical	$x = a$	Denominator $\to 0$ (non-removable)
Horizontal	$y = L$	$\lim_{x \to \pm\infty} f(x) = L$
Oblique	$y = mx + b$	Degree numerator = degree denominator + 1

45 Sequences and Series

A sequence is an ordered list of numbers, usually defined by a formula or a recurrence relation. It is one of the fundamental objects of study in mathematical analysis.

45.1 Arithmetic Sequence

An *arithmetic sequence* is a sequence where the difference between consecutive terms is constant.

$$a_n = a_1 + (n-1)d$$

- a_1 is the first term
- d is the common difference $d = a_{n+1} - a_n$

Example:

$$a_n = 3 + (n-1) \cdot 2 = 2n + 1 \quad \text{(Odd numbers)}$$

45.2 Geometric Sequence

A *geometric sequence* is a sequence where each term is obtained by multiplying the previous term by a fixed non-zero constant.

$$a_n = a_1 \cdot r^{n-1}$$

- a_1 is the first term
- r is the common ratio $r = \frac{a_{n+1}}{a_n}$

Example:

$$a_n = 2 \cdot 3^{n-1} = 2, 6, 18, 54, \ldots$$

45.3 Function Sequence

A *function sequence* is sequence composed of functions and its noted like:

$$(f)_n \text{with } x^t := \{f : D \to R | \ f \text{ is a function}\}$$

45.4 Convergence

A sequence (a_n) *converges* to a limit L if for every $\varepsilon > 0$, there exists an integer n_0 such that for all $n \geq n_0$:

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} : \forall i \geq n \ |a_i - L| < \varepsilon$$

In this case, we write:

$$\lim_{n \to \infty} a_n = L$$

45.5   Cauchy Sequence

A sequence (a_n) is a *Cauchy sequence* if for every $\varepsilon > 0$, there exists an integer n_0 such that for all $n, m \geq n_0$:

$$\forall \varepsilon > 0 \ \exists n, m \in \mathbb{N} : |a_n - a_m| < \varepsilon \text{ with } n < m$$

Every convergent sequence is a Cauchy sequence. In complete metric spaces (like $\mathbb{R}$), the converse also holds.

Example: Convergence Exercise

Let $a_i = \frac{1}{i}$ and $\varepsilon = 0.01$. We want to find i_0 such that:

$$\left| \frac{1}{n} - 0 \right| < \varepsilon \implies \frac{1}{i} \leq \frac{1}{n} < \varepsilon \to \frac{1}{\varepsilon} < n$$

$$\Rightarrow \frac{1}{i} < 0.01 \Rightarrow i > 100 \Rightarrow i_0 = 101$$

45.6   Definition of ε-Neighborhood

The *ε-neighborhood* of a point $a \in \mathbb{R}$ is the set:

$$B_\varepsilon(a) = \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

This is the open interval $(a - \varepsilon, a + \varepsilon)$.

45.7   Supremum, Infimum, Maximum and Minimum

The *supremum* of a set $A \subseteq \mathbb{R}$ is the least upper bound: the smallest real number s such that $a \leq s$ for all $a \in A$.

The *infimum* of a set A is the greatest lower bound: the largest number t such that $a \geq t$ for all $a \in A$. The *maximum* is the largest element in the set, if it exists.

The *minimum* is the smallest element in the set, if it exists.

Example:

For $A = (0, 1)$:

- $\sup A = 1$ (not in A)
- $\inf A = 0$ (not in A)
- $\max A$ and $\min A$ do not exist

45.8   Limit and Accumulation Point

- A *limit* of a sequence (a_n) is the value the terms get arbitrarily close to as $n \to \infty$.
- An *accumulation point* (or limit point) of a set A is a point x such that every neighborhood of x contains infinitely many points of A.

Example:

The sequence $a_n = (-1)^n + \frac{1}{n}$ has two accumulation points: 1 and -1.

45.9 Monotonicity by Difference and Quotient

Monotonicity by Difference:

- If $a_{n+1} - a_n \geq 0$ for all n, then the sequence is non-decreasing.
- If $a_{n+1} - a_n < 0$, it is non-increasing.

Monotonicity by Quotient: (Typically for positive sequences)

- If $\frac{a_{n+1}}{a_n} \geq 1$, then the sequence is non-decreasing.
- If $\frac{a_{n+1}}{a_n} < 1$, it is non-increasing.

45.10 Divergence (Definition)

A sequence *diverges* if it does not converge. That is, there is no real number L such that:

$$\lim_{n \to \infty} a_n = L$$

- A sequence can diverge to ∞ or $-\infty$
- A sequence can also oscillate without approaching any limit

Example:

The sequence $a_n = (-1)^n$ diverges since it oscillates between 1 and -1.

45.11 Subsequences and Their Properties

A *subsequence* of a sequence (a_n) is a sequence (a_{n_k}) where (n_k) is a strictly increasing sequence of natural numbers.

Properties:

- Every subsequence of a convergent sequence converges to the same limit.
- A sequence converges if and only if all of its subsequences converge to the same limit.
- A bounded sequence always has at least one convergent subsequence (Bolzano-Weierstraß theorem).

45.12 Series

A *series* is the sum of the terms of a sequence:

$$\sum_{n=1}^{\infty} a_n$$

We define the partial sums $S_n = \sum_{k=1}^{n} a_k$. If the sequence (S_n) converges to a limit S, then the series is said to converge:

$$\sum_{n=1}^{\infty} a_n = S$$

45.13 Power Series

A *power series* is a series of the form:

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

Where x_0 is the center of the expansion. Power series converge within a radius R:

$$R = \frac{1}{\displaystyle\limsup_{n \to \infty} \sqrt[n]{|a_n|}}$$

45.14 Geometric Series

There two types of *geometric series.*

1. Finite geometric series:

$$\sum_{n=0}^{r-1} aq^n = a \cdot \frac{1 - q^r}{1 - q} \quad \text{for } q \neq 1$$

2. Infinite geometric series:

$$\sum_{n=0}^{\infty} aq^n = \frac{a}{1 - q} \quad \text{for } |q| < 1$$

45.15 Point-wise Convergence

A sequence of functions (f_n) converges *point-wise* to a function f on a set A if:

$$\forall x \in A, \quad \lim_{n \to \infty} f_n(x) = f(x)$$

Point-wise convergence does not preserve properties like continuity or differentiability.

45.16 Uniform Convergence

A sequence of functions (f_n) converges *uniformly* to f on A if:

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} \text{ such that } \forall n \geq n_0, \forall x \in A, \quad |f_n(x) - f(x)| < \varepsilon$$

Key Property: Uniform convergence preserves continuity, integration, and differentiation under certain conditions.

45.17 Cauchy Criterion for Series

A series $\sum a_n$ converges if and only if for every $\varepsilon > 0$ there exists n_0 such that:

$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon \quad \forall n > m \geq n_0$$

45.18 Interval Nesting and the Interval Nesting Theorem

Let (I_n) be a sequence of closed intervals:

$$I_n = [a_n, b_n] \quad \text{with } I_{n+1} \subseteq I_n, \text{ and } \lim_{n \to \infty} (b_n - a_n) = 0$$

Then, the intersection contains exactly one point:

$$\bigcap_{n=1}^{\infty} I_n = \{x\}$$

This is useful in proving the existence of limits, roots, and fixed points.

45.19 Weierstraß Approximation Theorem

The *Weierstraß Approximation Theorem* states:

Every continuous function $f : [a, b] \to \mathbb{R}$ can be uniformly approximated by a polynomial $P(x)$, i.e., for every $\varepsilon > 0$ there exists a polynomial $P(x)$ such that:

$$|f(x) - P(x)| < \varepsilon \quad \forall x \in [a, b]$$

This theorem is foundational in numerical analysis and approximation theory.

45.20 Zero Sequence

A sequence (a_n) is called a *null sequence* or *zero sequence* if:

$$\lim_{n \to \infty} a_n = 0$$

Every zero sequence is convergent (to 0), and plays an essential role in convergence proofs and epsilon-delta arguments.

45.21 The Monotony Principle

Every bounded monotonic sequence converges.

- If (a_n) is non-decreasing and bounded above, then $\lim a_n$ exists and equals $\sup\{a_n\}$.
- If (a_n) is non-increasing and bounded below, then $\lim a_n = \inf\{a_n\}$.

To prove the convergence of a sequence via this principle we follow the next steps

1. Prove the Monotonicity by induction
2. Calculate the limit
3. Prove the limitedness via induction
4. Write the conclusion

Example:

$a_{n+1} = \sqrt{2a_n - 3} + 1$

Let us assume that the series is decreasing or $a_{n+1} < a_n$

$$a_{n+2} < a_{n+1}$$
$$\sqrt{2a_{n+1} - 3} + 1 < \sqrt{2a_n - 3} + 1$$
$$\sqrt{2a_{n+1} - 3} < \sqrt{2a_n - 3}$$
$$2a_{n+1} - 3 < 2a_n - 3$$
$$a_{n+1} < a_n$$

$$\mathfrak{QED}$$

Now let us calculated the limit

$$\lim_{n \to \infty} a_{n+1} = \lim_{n \to \infty} \sqrt{2a_n - 3} + 1$$

Because of the recursion we know that both a_{n+1} and a_n have the same value, and we call it a.

$$a = \sqrt{2a - 3} + 1$$
$$a = 2$$

We only have to prove the limitedness now:

For $n = 0$ we have $a_1 = 4 > 2$ then let us assume that the a is never going bellow 2.

$$a_{n+1} = \sqrt{2a_n - 3} + 1 \geq \sqrt{2(2) - 3} + 1 = 2$$

By setting $a_n = 2$ we get the limit again

$$\mathfrak{QED}$$

45.22 Operations on Sequences

45.22.1 Sum of Sequences

The sum of two sequences (a_n) and (b_n) is a new sequence (c_n) where each term c_n is the sum of the corresponding terms of (a_n) and (b_n):

$$c_n = a_n + b_n, \quad \forall n \in \mathbb{N}.$$

If $\lim_{n \to \infty} a_n = A$ and $\lim_{n \to \infty} b_n = B$, then the limit of the sum sequence is:

$$\lim_{n \to \infty} (a_n + b_n) = A + B.$$

45.22.2 Multiplication of Sequences

The product of two sequences (a_n) and (b_n) is a new sequence (d_n) where each term d_n is the product of the corresponding terms of (a_n) and (b_n):

$$d_n = a_n \cdot b_n, \quad \forall n \in \mathbb{N}.$$

If $\lim_{n \to \infty} a_n = A$ and $\lim_{n \to \infty} b_n = B$, then the limit of the product sequence is:

$$\lim_{n \to \infty} (a_n \cdot b_n) = A \cdot B.$$

45.22.3 Absolute Value of a Sequence

The absolute value of a sequence (a_n) is a new sequence (e_n) where each term e_n is the absolute value of the corresponding term of (a_n):

$$e_n = |a_n|, \quad \forall n \in \mathbb{N}.$$

If $\lim_{n \to \infty} a_n = A$, then the limit of the absolute value sequence is:

$$\lim_{n \to \infty} |a_n| = |A|.$$

45.22.4 Conjugate of a Complex Sequence

If (a_n) is a sequence of complex numbers, where $a_n = x_n + iy_n$ with $x_n, y_n \in \mathbb{R}$, then the conjugate of (a_n) is a new sequence $(\overline{a_n})$ where each term $\overline{a_n}$ is the complex conjugate of a_n:

$$\overline{a_n} = x_n - iy_n, \quad \forall n \in \mathbb{N}.$$

If $\lim_{n \to \infty} a_n = A = X + iY$, then the limit of the conjugate sequence is:

$$\lim_{n \to \infty} \overline{a_n} = \overline{A} = X - iY.$$

45.22.5 Complex Limit (Real and Imaginary Parts)

For a sequence of complex numbers (a_n) where $a_n = x_n + iy_n$, the limit of the sequence $\lim_{n\to\infty} a_n = A = X + iY$ exists if and only if the limits of the real part sequence (x_n) and the imaginary part sequence (y_n) exist individually:

$$\lim_{n\to\infty} a_n = A \iff \left(\lim_{n\to\infty} \operatorname{Re}(a_n) = \operatorname{Re}(A) = X \quad \text{and} \quad \lim_{n\to\infty} \operatorname{img}(a_n) = \operatorname{img}(A) = Y \right).$$

This means we can analyze the convergence of a complex sequence by examining the convergence of its real and imaginary parts separately.

45.22.6 Asymptotic Equality

Two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are said to be asymptotically equal, denoted by $a_n \sim b_n$, if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1,$$

provided that $b_n \neq 0$ for all sufficiently large n. Asymptotic equality implies that the sequences behave similarly as n approaches infinity.

46 Limits

Limits are a foundational concept in calculus and analysis. They describe the behavior of functions and sequences as the input approaches a particular value or infinity.

46.1 Definition of the Limit

Let f be a function defined near a point a. We say:

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

if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that:

$$0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon$$

This is the precise ε-δ *definition* of a limit.

46.2 Limit Calculation Rules

Sum Rule: $\lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$

Difference Rule: $\lim_{x \to a} (f(x) - g(x)) = \lim_{x \to a} f(x) - \lim_{x \to a} g(x)$

Product Rule: $\lim_{x \to a} (f(x)g(x)) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x)$

Quotient Rule: $\lim_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim f(x)}{\lim g(x)}$ if $\lim g(x) \neq 0$

Power Rule: $\lim_{x \to a} f(x)^n = \left(\lim_{x \to a} f(x) \right)^n$

Root Rule: $\lim_{x \to a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x \to a} f(x)}$ (when defined)

46.3 Limits at Infinity

$$\lim_{x \to \infty} f(x) = L \quad \text{means that } f(x) \text{ approaches } L \text{ as } x \to \infty$$

$$\lim_{x \to \infty} \frac{1}{x} = 0, \quad \lim_{x \to \infty} e^x = \infty, \quad \lim_{x \to \infty} \frac{1}{x^2} = 0$$

46.4 Indeterminate Forms

Indeterminate forms arise in limits when the expression does not directly imply a limit value:

$$\frac{0}{0}, \quad \frac{\infty}{\infty}, \quad 0 \cdot \infty, \quad \infty - \infty, \quad 1^\infty, \quad \infty^0, \quad 0^0$$

46.5 Limit of a Composite Function

If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to L} g(x) = g(L)$ (i.e., g is continuous at L), then:

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

46.6 Right and Left Limits

$$\lim_{x \to a^-} f(x) = L_-, \quad \lim_{x \to a^+} f(x) = L_+$$

The two-sided limit $\lim_{x \to a} f(x)$ exists if and only if $L_- = L_+$.

46.7 L'Hôpital's Rule

If $\lim\limits_{x \to a} f(x) = \lim\limits_{x \to a} g(x) = 0$ or $\pm\infty$, and $\lim\limits_{x \to a} \dfrac{f'(x)}{g'(x)}$ exists, then:

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

This rule is used to resolve indeterminate forms $\frac{0}{0}$ and $\frac{\infty}{\infty}$.

46.8 Limit of a fraction

For expressions like:

$$\lim_{n \to \infty} \left(\frac{a}{b}\right)^n$$

- If $\frac{a}{b} < 1$: limit is 0
- If $\frac{a}{b} > 1$: limit diverges to ∞

46.9 Limit of a Recursive Sequence

Let (a_n) be defined recursively:

$$a_{n+1} = f(a_n)$$

If (a_n) converges to L and f is continuous, then:

$$\lim_{n \to \infty} a_n = L \Rightarrow L = f(L)$$

This is used to find fixed points of recursive definitions.

46.10 Important Limits

- $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1$

- $\lim\limits_{x \to 0} \dfrac{1 - \cos x}{x^2} = \dfrac{1}{2}$

- $\lim\limits_{x \to 0} \dfrac{\ln(1 + x)}{x} = 1$

- $\lim\limits_{x \to 0} \dfrac{e^x - 1}{x} = 1$

- $\lim\limits_{n \to \infty} \dfrac{n!}{n^n} = 0$

- $\lim\limits_{n \to \infty} \sqrt[n]{a} = 1$ (for $a > 0$)

- $\lim\limits_{x \to 0} \dfrac{1}{\cos x} = 1$

46.11 Squeeze Theorem (Sandwich Theorem)

If $g(x) \le f(x) \le h(x)$ near a, and:

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

Proof:

We know that $g(x) \leq f(x) \leq h(x)$ therefore, for $g(x)$ and $h(x)$ we can define the following inequalities.

$$|g(x)_{N_1} - L| < \varepsilon$$
$$|h(x)_{N_2} - L| < \varepsilon$$

Then let us define $n = \max(N_1, N_2)$. And we can also write

$$-\varepsilon < g(x)_n - L < \varepsilon$$
$$-\varepsilon < h(x)_n - L < \varepsilon$$

Now, we write the following because we are assuming they all have the same limit

$$-\varepsilon < g(x)_n - L \leq f(x)_n - L \leq h(x)_n - L < \varepsilon$$
$$-\varepsilon <\leq f(x)_n - L < \varepsilon$$

from our original inequality. Thus, there has been no contradiction.

$$\mathfrak{QED}$$

Example:

$$\lim_{x \to 0} x^2 \cos\left(\frac{1}{x}\right) = 0$$

46.12   The Number e as a Limit

The number e is defined as:

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

More generally:

$$\lim_{n \to \infty} \left(1 + \frac{a}{bn^d}\right)^{n^d} = e^{\frac{a}{b}}$$

Example I:

$$\lim_{x \to \infty} \left(\frac{x^2 + 1}{x^2 + 2}\right)^{x^2}$$

$$\lim_{x \to \infty} \left(\left(\frac{x^2 + 1}{x^2 + 2}\right)^x\right)^x$$

$$\lim_{x \to \infty} \left(\left(\frac{1 + \frac{1}{x^2}}{1 + \frac{2}{x^2}}\right)^x\right)^x$$

$$\lim_{x \to \infty} \left(\frac{e}{e^2}\right)^x = \left(\frac{1}{e}\right)^x = 0$$

Example II:

$$\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^{x^2}$$

$$\lim_{x \to \infty} \left(\left(1 + \frac{1}{x}\right)^{x^2}\right)^{\frac{x}{x}}$$

$$\lim_{x \to \infty} \left(\left(1 + \frac{1}{x}\right)^{x}\right)^{x}$$

$$\lim_{x \to \infty} \left(\lim_{n \to \infty} \left(1 + \frac{1}{x}\right)^{x}\right)^{x}$$

$$\lim_{x \to \infty} (e)^x = \infty$$

46.13 Limits of arctan

- $\lim_{x \to \infty} \arctan x = \frac{\pi}{2}$
- $\lim_{x \to -\infty} \arctan x = -\frac{\pi}{2}$
- $\arctan x$ is continuous and differentiable everywhere, with a horizontal asymptote at $y = \pm\frac{\pi}{2}$

46.14 Bolzano-Weierstraß Theorem

Every bounded, convergent sequence in $\mathbb{C}$ has a convergent subsequence. More precisely the sequence a has biggest and smallest accumulation point h^* and h_*

$$\forall \varepsilon > 0 : h_* - \varepsilon < a_n < h^* + \varepsilon \text{ for all-most all } b \in \mathbb{N}$$

$$h^* := \limsup_{n \to \infty} a_n$$

$$h_* := \liminf_{n \to \infty} a_n$$

46.15 Uniqueness of the Limit

Let $f : D \subset \mathbb{R} \to \mathbb{R}$, and suppose a is a limit point of D. Then: If $\lim_{x \to a} f(x) = L_1$ and $\lim_{x \to a} f(x) = L_2$, then $L_1 = L_2$.

In other words, a function can have at most one limit at a given point. This is the *uniqueness of limits*.

Proof (by contradiction)

Assume $L_1 \neq L_2$, and let $\varepsilon = \frac{|L_1 - L_2|}{3} > 0$.

- Since $\lim_{x \to a} f(x) = L_1$, there exists $\delta_1 > 0$ such that for all $x \in D$, $0 < |x - a| < \delta_1$ implies:

$$|f(x) - L_1| < \varepsilon$$

- Since $\lim_{x \to a} f(x) = L_2$, there exists $\delta_2 > 0$ such that for all $x \in D$, $0 < |x - a| < \delta_2$ implies:

$$|f(x) - L_2| < \varepsilon$$

Let $\delta = \min(\delta_1, \delta_2)$, and choose $x \in D$ such that $0 < |x - a| < \delta$.

Then:

$$|f(x) - L_1| < \varepsilon, \quad |f(x) - L_2| < \varepsilon$$

Using the triangle inequality:

$$|L_1 - L_2| = |L_1 - f(x) + f(x) - L_2| \leq |f(x) - L_1| + |f(x) - L_2| < \varepsilon + \varepsilon = 2\varepsilon$$

But we chose $\varepsilon = \frac{|L_1 - L_2|}{3}$, so:

$$|L_1 - L_2| < \frac{2}{3}|L_1 - L_2| \Rightarrow \frac{1}{3}|L_1 - L_2| < 0$$

Which is a contradiction since absolute values are always non-negative.

Conclusion: Our assumption was false. Hence, $L_1 = L_2$. The limit is unique.

$$\mathfrak{QED}$$

46.16　Every Convergent Sequence is a Cauchy Sequence

Let (a_n) be a sequence in $\mathbb{R}$. We say:

If (a_n) converges, then it is a Cauchy sequence.

A sequence (a_n) is a *Cauchy sequence* if:

$$\forall \varepsilon > 0, \ \exists N \in \mathbb{N} \text{ such that } \forall n, m \geq N, \quad |a_n - a_m| < \varepsilon$$

Proof:

Assume $\lim_{n \to \infty} a_n = L$. Then:

- For every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that:

$$\forall n \geq N, \quad |a_n - L| < \frac{\varepsilon}{2}$$

- Now, for any $n, m \geq N$, use the triangle inequality:

$$|a_n - a_m| = |a_n - L + L - a_m| \leq |a_n - L| + |a_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Conclusion:

$|a_n - a_m| < \varepsilon$ for all $n, m \geq N \Rightarrow (a_n)$ is Cauchy

This completes the proof that every convergent sequence is Cauchy.

$$\mathfrak{QED}$$

47 Convergence Criteria

In this section, we present various tools used to determine whether an infinite series or sequence converges. These criteria are essential for analyzing the behavior of sequences and series in analysis.

47.1 Quotient (Ratio) Test

Let $\sum a_n$ be a series with positive terms, and suppose the limit

$$L = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

exists.

- If $L < 1$, the series converges absolutely.
- If $L > 1$ or $L = \infty$, the series diverges.
- If $L = 1$, the test is inconclusive.

Proof of the Ratio Test (for $L < 1$)

Let $r < 1$ such that $L < r < 1$. Then for sufficiently large n, we have:

$$\left| \frac{a_{n+1}}{a_n} \right| \leq r \Rightarrow a_{n+1} \leq r a_n$$

By induction:

$$a_{n+k} \leq a_n r^k$$

So,

$$\sum_{n=N}^{\infty} a_n \leq a_N \sum_{k=0}^{\infty} r^k = \frac{a_N}{1 - r}$$

Hence, the series converges by comparison with a geometric series.

47.2 Zero Sequence Criterion

If a series $\sum a_n$ converges, then:

$$\lim_{n \to \infty} a_n = 0$$

The converse is not true — the harmonic series is a classic counterexample.

47.3 Majorant (Comparison Test)

If $0 \leq a_n \leq b_n$ for all n and $\sum b_n$ converges, then $\sum a_n$ converges.

47.4 Minorant

If $0 \leq b_n \leq a_n$ and $\sum b_n$ diverges, then $\sum a_n$ diverges.

47.5 Root Test

Let $\sum a_n$ be a series with non-negative terms, and suppose:

$$L = \limsup_{n \to \infty} \sqrt[n]{|a_n|}$$

- If $L < 1$, the series converges absolutely.
- If $L > 1$, the series diverges.
- If $L = 1$, the test is inconclusive.

Proof of the Root Test (for $L < 1$)

Let r be such that $L < r < 1$. Then for large n:

$$\sqrt[n]{|a_n|} \leq r \Rightarrow |a_n| \leq r^n$$

Since $\sum r^n$ converges, by comparison $\sum |a_n|$ converges absolutely.

47.6 Leibniz Criterion (Alternating Series Test)

Let $\sum (-1)^n a_n$ be a series with:
- a_n decreasing: $a_{n+1} \leq a_n$
- $\lim a_n = 0$

Then the series converges.

Example:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

47.7 Compactness Criterion

A sequence (x_n) in $\mathbb{R}^n$ has a convergent subsequence if and only if it is bounded. This is a consequence of the **Bolzano-Weierstraß theorem**, which is a form of sequential compactness in $\mathbb{R}^n$.

47.8 Cauchy Criterion I

A sequence (a_n) converges if and only if it is a *Cauchy sequence*:

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n, m \geq n_0 \Rightarrow |a_n - a_m| < \varepsilon$$

This is an internal (non-limit based) criterion for convergence, valid in complete metric spaces.

47.9 Cauchy Criterion II

A series $a_n = \sum_{i=0}^{k} x_i$ converges $\iff \forall d \in \mathbb{R} \, d > 0 \exists N \in \mathbb{N} \forall n, m > N : |\sum_{i=0}^{n} x_i + \sum_{i=0}^{m}|$

$$= \left| \sum_{i=0}^{n} x_i \right| < d$$

47.10 The Cauchy Product

Let $\sum a_n$, $\sum b_n$ be convergent series. The *Cauchy product* is:

$$\sum_{n=0}^{\infty} c_n, \quad \text{where } c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

If $\sum a_n$ and $\sum b_n$ converge absolutely, then their Cauchy product converges absolutely and:

$$\sum c_n = \left(\sum a_n \right) \left(\sum b_n \right)$$

47.10.1 Cauchy Product Example

Prove that $e^x e^y = e^{x+y}$
Note that $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ and $e^y = \sum_{n=0}^{\infty} \frac{y^n}{n!}$

Therefore,

$$e^x e^y = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

And because of the absolute convergence

$$e^x e^y = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k}{n!} \frac{y^{n-k}}{(n-k)!}$$

Finally,

$$e^x e^y = \sum_{n=0}^{\infty} \sum_{k=0}^{n} x^k y^{n-k} \binom{n}{k} \frac{1}{n!} = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}$$

$$\mathfrak{QED}$$

47.11 Geometric Series

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \quad \text{for } |r| < 1$$

If $|r| \geq 1$, the series diverges.

47.12 Harmonic Series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

This series diverges, even though its terms approach zero.

Proof:

Group terms:

$$\left(\frac{1}{2} \right) + \left(\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \cdots + \frac{1}{8} \right) + \cdots > \sum_{k=1}^{\infty} \frac{1}{2}$$

The lower bound is a divergent series, hence the harmonic series diverges.

48 Convergence Radius

Consider a power series of the form

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n,$$

where c_n are the coefficients, x is the variable, and a is the center of the series. The radius of convergence, denoted by R, is a non-negative real number (or ∞) such that the series converges if $|x-a| < R$ and diverges if $|x-a| > R$. In other words, the interval of convergence is $(a-R, a+R)$, possibly including one or both endpoints.

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}$$

or

$$R = \limsup_{n\to\infty} \frac{a_n}{a_{n+1}}$$

48.1 Test of the Borders

When the absolute value of the difference between x and a is equal to the radius of convergence (i.e., $|x-a| = R$), the convergence of the power series must be checked separately. This involves substituting the values $x = a \pm R$ into the series and determining whether the resulting series converges or diverges using other convergence tests (e.g., the comparison test, ratio test, alternating series test).

48.2 Approximating the Error

When a power series is used to approximate a function, it is often necessary to estimate the error in the approximation. If the power series converges, the error can be made arbitrarily small by including a sufficient number of terms. For an alternating series that satisfies the conditions of the Alternating Series Test, the error in approximating the sum by the first n terms is bounded by the absolute value of the $(n+1)$-th term:

$$\left| \sum_{k=0}^{\infty} c_k(x-a)^k - \sum_{k=0}^{n} c_k(x-a)^k \right| \le |c_{n+1}(x-a)^{n+1}|.$$

For other power series, techniques such as Taylor's Theorem with remainder can be used to estimate the error.

48.3 Complete Example

Find the convergence radius of

$$\sum_{n=1}^{\infty} \frac{x^n}{n2^n}$$

at $x = 0$.

Step 1: Extract the sequence a_n

The general form is

$$\sum_{0}^{\infty} a_n(x-x_0)^n$$

233

in our case

$$= \sum_{n=1}^{\infty} x^n \frac{1}{n2^n}$$

with $\frac{1}{n2^n}$ as a_n.

Step 2: Compute the radius

$$R = \frac{1}{\limsup_{n \to \infty} \sqrt[n]{\frac{1}{n2^n}}}$$

$$= \frac{1}{\limsup_{n \to \infty} \frac{\sqrt[n]{1}}{\sqrt[n]{n}\sqrt[n]{2^n}}}$$

$$= \frac{1}{\frac{1}{2}} = 2$$

Step 3: Check Borders

The borders to check are $0 - 2$ to $0 + 2$.

For -2:

$$\sum_{n=1}^{\infty} \frac{(-2)^n}{n2^n} = \sum_{n=1}^{\infty} \frac{(-1)^n (2)^n}{n2^n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

by Leibniz, it converges.

For 2:

$$\sum_{n=1}^{\infty} \frac{(2)^n}{n2^n} = \sum_{n=1}^{\infty} \frac{1}{n}$$

This is the harmonic series therefore it diverges.

Step 4: Write the interval

The final interval of convergence is $[-2, 2)$.

49 Calculating Series Values

49.1 Telescoping Sums and Partial Fraction Decomposition

Telescoping sums are series where most of the terms cancel out, leaving only a few terms. Partial fraction decomposition is a technique used to rewrite rational functions as a sum of simpler fractions, which can often lead to telescoping sums.

Example: $\sum_{k=1}^{\infty} \frac{1}{4k^2-1}$

Consider the infinite series

$$S = \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1}.$$

We can decompose the fraction using partial fraction decomposition:

$$\frac{1}{4k^2 - 1} = \frac{1}{(2k-1)(2k+1)} = \frac{A}{2k-1} + \frac{B}{2k+1}.$$

Multiplying both sides by $(2k - 1)(2k + 1)$ gives

$$1 = A(2k + 1) + B(2k - 1).$$

To solve for A and B, we can use the following values of k:

- For $k = \frac{1}{2}$: $1 = A(1 + 1) + B(0) \implies A = \frac{1}{2}$.
- For $k = -\frac{1}{2}$: $1 = A(0) + B(-1 - 1) \implies B = -\frac{1}{2}$.

Thus, we have

$$\frac{1}{4k^2 - 1} = \frac{1}{2} \left(\frac{1}{2k-1} - \frac{1}{2k+1} \right).$$

Now we can write the series as a telescoping sum:

$$S = \frac{1}{2} \sum_{k=1}^{\infty} \left(\frac{1}{2k-1} - \frac{1}{2k+1} \right).$$

Let S_n be the n-th partial sum:

$$S_n = \frac{1}{2} \left[\left(\frac{1}{1} - \frac{1}{3} \right) + \left(\frac{1}{3} - \frac{1}{5} \right) + \left(\frac{1}{5} - \frac{1}{7} \right) + \cdots + \left(\frac{1}{2n-1} - \frac{1}{2n+1} \right) \right]$$
$$= \frac{1}{2} \left[1 - \frac{1}{2n+1} \right].$$

Taking the limit as n approaches infinity, we get

$$S = \lim_{n \to \infty} S_n = \lim_{n \to \infty} \frac{1}{2} \left(1 - \frac{1}{2n+1} \right) = \frac{1}{2}(1 - 0) = \frac{1}{2}.$$

Therefore, the value of the series is $\frac{1}{2}$.

Example II:

Now an example using the geometric series

$$\sum_{n=0}^{\infty} \frac{4}{3^n} = \sum_{n=0}^{\infty} 4 \left(\frac{1}{3} \right)^n$$
$$= \frac{4}{1 - \frac{1}{3}}$$

50 Transforming Functions into Power Series

Many functions can be represented as power series, which is useful for various applications in calculus and analysis. This section discusses some common techniques for transforming functions into power series.

50.1 Common Cases

Here are some common cases where functions can be easily transformed into power series:

$\frac{1}{1-x}$

The function $\frac{1}{1-x}$ has the well-known geometric series representation:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \quad |x| < 1.$$

$\frac{1}{1-(x-x_0)^n}$

A slight generalization is:

$$\frac{1}{1-(x-x_0)^n} = \sum_{k=0}^{\infty} (x-x_0)^{nk}, \quad |x-x_0| < 1.$$

$\frac{1}{1-d(x-x_0)^n}$

For a constant d:

$$\frac{1}{1-d(x-x_0)^n} = \sum_{k=0}^{\infty} d^k (x-x_0)^{nk}, \quad |d(x-x_0)^n| < 1.$$

Example: $\frac{x+1}{3-x}$

Let's find the power series representation of the function $f(x) = \frac{x+1}{3-x}$.

50.1.1 Partial Fractions

First, we perform partial fraction decomposition. We can rewrite the function as:

$$\frac{x+1}{3-x} = \frac{-(3-x)+4}{3-x} = -1 + \frac{4}{3-x}.$$

50.1.2 Manipulating the Expression

Now, we manipulate the fraction to fit the form $\frac{1}{1-u}$:

$$\frac{4}{3-x} = \frac{4}{3(1-\frac{x}{3})} = \frac{4}{3} \cdot \frac{1}{1-\frac{x}{3}}.$$

50.1.3 Power Series Representation

Using the geometric series formula, we have

$$\frac{1}{1-\frac{x}{3}} = \sum_{n=0}^{\infty} \left(\frac{x}{3}\right)^n, \quad \left|\frac{x}{3}\right| < 1.$$

Thus,

$$\frac{x+1}{3-x} = -1 + \frac{4}{3} \sum_{n=0}^{\infty} \left(\frac{x}{3}\right)^n$$

$$= -1 + \frac{4}{3} \sum_{n=0}^{\infty} \frac{x^n}{3^n}$$

$$= -1 + \sum_{n=0}^{\infty} \frac{4}{3^{n+1}} x^n, \quad |x| < 3.$$

Therefore, the power series representation of the function is

$$\frac{x+1}{3-x} = -1 + \sum_{n=0}^{\infty} \frac{4}{3^{n+1}} x^n, \quad |x| < 3.$$

51 Continuity

Continuity describes functions whose values change smoothly, without abrupt jumps or breaks. This concept is fundamental in calculus, analysis, and fixed point theory.

51.1 Definition of Continuity at a Point

A function f is *continuous* at a point x_0 if:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$$

Equivalently,

$$\lim_{x \to x_0} f(x) = f(x_0)$$

To solve continuity problems the general approach is to:

1. Write down the definition.

2. Substitute f for the concrete function into the definition.

3. Modify the expression with the concrete f $|f(x) - f(y)|$ in way that you get $|x - y|$ inside the expression.

4. Choose a δ if not given. On the one side it is often $\delta = |x - y|$ and then substitute into the expression. On the other side if δ is given make an approximation which is greater than the current expression and then simplify. In both cases after simplifying you should get a ε is the function is continuous.

Example:

$\varepsilon - \delta$ Exercise for $f(x) = \frac{1}{x}$ at $x_0 > 0$

We want to find δ such that:

$$|x - x_0| < \delta \Rightarrow \left| \frac{1}{x} - \frac{1}{x_0} \right| < \varepsilon$$

Start by rewriting:

$$\left| \frac{1}{x} - \frac{1}{x_0} \right| = \left| \frac{x_0 - x}{x x_0} \right| = \frac{|x - x_0|}{|x||x_0|}$$

Assume:

$$x \in \left[\frac{x_0}{2}, \frac{3x_0}{2} \right] \Rightarrow |x| \geq \frac{x_0}{2}$$

Then:

$$\left| \frac{1}{x} - \frac{1}{x_0} \right| \leq \frac{|x - x_0|}{\frac{x_0}{2} \cdot x_0} = \frac{2}{x_0^2} |x - x_0|$$

Choose:

$$\delta = \min \left\{ \frac{x_0}{2}, \frac{x_0^2 \varepsilon}{2} \right\}$$

51.2   Rules of Continuity

Let f and g be continuous at x_0:

- $f + g$, $f - g$, $f \cdot g$ are continuous at x_0
- $\frac{f}{g}$ is continuous at x_0 if $g(x_0) \neq 0$
- Compositions: if g is continuous at x_0, and f is continuous at $g(x_0)$, then $f \circ g$ is continuous at x_0

51.3   Lipschitz Continuity

A function $f : D \subset \mathbb{R} \to \mathbb{R}$ is *Lipschitz continuous* if:

$$\exists L > 0 \text{ such that } |f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in D$$

- Every Lipschitz continuous function is uniformly continuous.
- The smallest such L is called the *Lipschitz constant.*

51.4   Banach Fixed Point Theorem

Let (X, d) be a complete metric space, and $f : X \to X$ a *contraction mapping*, i.e.,

$$\exists L < 1 \text{ such that } d(f(x), f(y)) \leq L \cdot d(x, y)$$

Then:

- f has a unique fixed point $x^* \in X$, such that $f(x^*) = x^*$
- Iterating $x_{n+1} = f(x_n)$ converges to x^*

Interpretation and Meaning

The Banach Fixed Point Theorem ensures:

- Existence and uniqueness of solutions (fixed points)
- Convergence of approximation by iteration
- Powerful tool in numerical methods and differential equations

51.4.1   A Priori and A Posteriori Approximations

A priori estimate:

$$|x_n - x^*| \leq \frac{L^n}{1 - L}|x_1 - x_0| < \epsilon$$

A posteriori estimate:

$$|x_n - x^*| \leq \frac{L}{1 - L}|x_n - x_{n-1}|\epsilon$$

51.4.2   Steps for a Fixed Point Exercise

Given $f : [a, b] \to [a, b]$, to prove existence and convergence:

1. Show monotonicity: f is increasing or decreasing.
2. Check interval preservation: $f([a, b]) \subseteq [a, b]$
3. Check Lipschitz continuity: Find $L < 1$
4. Fixed point iteration: Choose x_0, compute $x_{n+1} = f(x_n)$
5. Apply a priori or a posteriori bound

Example:

$f(x) = \frac{1}{x} + 3$ on $[2, 5]$

Step 1: Monotonicity

In this case we will assume that it is monotonous.

Step 2: Domain check

$x \in [2, 5] \Rightarrow \frac{1}{x} \in [0.2, 0.5] \Rightarrow f(x) \in [3.2, 3.5] \subset [2, 5]$

Step 3: Lipschitz constant

$$f'(x) = -\frac{1}{x^2} \Rightarrow |f'(x)| \leq \frac{1}{2^2} = \frac{1}{4} < 1 \Rightarrow \text{Lipschitz with } L = \frac{1}{4} < 1$$

Step 4: Fixed point iteration

$$x_0 = 3, \quad x_1 = f(3) = \frac{1}{3} + 3 = \frac{10}{3}, \quad x_2 = f(x_1), \ldots$$

Step 5: Approximations

A priori

$$|x_n - x^*| \leq \frac{L^n}{1 - L} |x_1 - x_0|$$

$$|x_n - x^*| \leq \frac{\frac{1}{4}^n}{1 - \frac{1}{4}} |\frac{10}{3} - 3|$$

$$|x_n - x^*| \leq \frac{\left(\frac{1}{4}\right)^n}{1 - \frac{1}{4}} \cdot \left|\frac{10}{3} - 3\right|$$

$$= \frac{\left(\frac{1}{4}\right)^n}{\frac{3}{4}} \cdot \frac{1}{3}$$

$$= \frac{4}{3} \cdot \left(\frac{1}{4}\right)^n \cdot \frac{1}{3}$$

$$= \frac{4}{9} \cdot \left(\frac{1}{4}\right)^n$$

For $\varepsilon = 0,1$

$$\frac{4}{9} \cdot \left(\frac{1}{4}\right)^n < \varepsilon$$

$$\left(\frac{1}{4}\right)^n < \frac{9}{4}\varepsilon$$

$$\ln\left(\left(\frac{1}{4}\right)^n\right) < \ln\left(\frac{9}{4}\varepsilon\right)$$

$$n \cdot \ln\left(\frac{1}{4}\right) < \ln\left(\frac{9}{4}\varepsilon\right)$$

$$n > \frac{\ln\left(\frac{9}{4}\varepsilon\right)}{\ln\left(\frac{1}{4}\right)}$$

$$n > \frac{\ln\left(\frac{9}{4}\varepsilon\right)}{-\ln 4}$$

A posteriori

$$|x_n - x^*| \leq \frac{L}{1-L}|x_n - x_{n-1}| = \frac{\frac{1}{4}}{1 - \frac{1}{4}}|x_n - x_{n-1}| = \frac{1}{3}|x_n - x_{n-1}|$$

52 Differentiation

Differentiation is a core concept in calculus. It describes the rate at which a quantity changes and provides tools for analyzing and modeling change in real-world phenomena.

52.1 Definition of the Derivative

The derivative of a function f at a point x is defined as the limit:

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

or

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

provided this limit exists.

Geometrically, $f'(a)$ is the slope of the tangent line to the graph of f at $x = a$.

52.2 Derivative Rules

Power Rule:

$$\frac{d}{dx}x^n = nx^{n-1}, \quad \text{for } n \in \mathbb{R}$$

Constant Rule:

$$\frac{d}{dx}c = 0, \quad \text{for constant } c$$

Constant Multiple Rule:

$$\frac{d}{dx}[c \cdot f(x)] = c \cdot f'(x)$$

Sum and Difference Rule:

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

Product Rule:

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

Quotient Rule:

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

Chain Rule:

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

52.3 Extrema and Critical Points Test

A function f has a critical point at $x = c$ if:

$$f'(c) = 0 \quad \text{or} \quad f'(c) \text{ does not exist}$$

To determine the nature of the critical point:

- *First Derivative Test:* Check sign changes of f' around c
- *Second Derivative Test:*

$$f''(c) > 0 \Rightarrow \text{local minimum} \quad f''(c) < 0 \Rightarrow \text{local maximum}$$

This works, because the second derivative is telling us about the concavity of the function. If it is positive then we will have function which is concave up a minimum; and if negative concave down a maximum.

52.4 Trigonometric Derivatives

$$\frac{d}{dx}\sin x = \cos x$$
$$\frac{d}{dx}\cos x = -\sin x$$
$$\frac{d}{dx}\tan x = \sec^2 x$$
$$\frac{d}{dx}\cot x = -\csc^2 x$$
$$\frac{d}{dx}\sec x = \sec x \tan x$$
$$\frac{d}{dx}\csc x = -\csc x \cot x$$

52.5 Derivatives of Inverse Trigonometric Functions

$$\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1 - x^2}}$$
$$\frac{d}{dx}\arccos x = -\frac{1}{\sqrt{1 - x^2}}$$
$$\frac{d}{dx}\arctan x = \frac{1}{1 + x^2}$$
$$\frac{d}{dx}arcotx = -\frac{1}{1 + x^2}$$
$$\frac{d}{dx}arcsecx = \frac{1}{|x|\sqrt{x^2 - 1}}$$
$$\frac{d}{dx}arccscx = -\frac{1}{|x|\sqrt{x^2 - 1}}$$

52.6 The Differential

The differential dy of a function $y = f(x)$ is defined as:

$$dy = f'(x)\, dx$$

This linear approximation estimates the change in y for a small change in x.

52.7   The Secant Equation and Graphical Meaning

The *secant line* through points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$ has slope:

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

As $h \to 0$, the secant line approaches the tangent line. Graphically, this means:

$$\lim_{h \to 0} \text{slope of secant} = \text{slope of tangent} = f'(x) = \Delta_h$$

The secant line equation is given by:

$$L(x) = f(x) + \Delta_h f(x_0)(x - x_0)$$

52.8   Optimization via the Derivative

To find local or global extrema:

1. Compute $f'(x)$
2. Solve $f'(x) = 0$ to find critical points
3. Use the first or second derivative test to classify
4. Evaluate endpoints if optimizing over a closed interval

Example I:

Maximize area $A = x(10 - 2x)$

$$A = 10x - 2x^2, \quad A' = 10 - 4x, \quad A' = 0 \Rightarrow x = 2.5$$

Example II:

Given a number $c > 0$, we want to factor it into two positive numbers x and y such that their product is c, and their sum $S = x + y$ is minimal.

Since $x \cdot y = c$, we can express one variable in terms of the other, e.g.,

$$y = \frac{c}{x}$$

Substituting into the sum:

$$S(x) = x + \frac{c}{x}$$

To minimize $S(x)$, we take the derivative:

$$S'(x) = 1 - \frac{c}{x^2}$$

Setting the derivative to zero gives:

$$1 - \frac{c}{x^2} = 0 \quad \Rightarrow \quad x^2 = c \quad \Rightarrow \quad x = \sqrt{c}$$

Then:

$$y = \frac{c}{x} = \sqrt{c}$$

So the minimal sum occurs when $x = y = \sqrt{c}$, and the minimum value of the sum is:

$$x + y = 2\sqrt{c}$$

Let $c = 25$. Then the two positive factors that minimize the sum are:

$$x = y = \sqrt{25} = 5$$

and the minimal sum is:

$$x + y = 5 + 5 = 10$$

52.9   Power Rule Derivation

We derive the power rule:

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

for $n \in \mathbb{N}$ using the limit definition:

$$\frac{d}{dx} x^n = \lim_{h \to 0} \frac{(x+h)^n - x^n}{h}$$

Use the Binomial Theorem:

$$(x+h)^n = x^n + n x^{n-1} h + \cdots + h^n \Rightarrow \text{Only the linear term survives after dividing by } h$$

52.10   Implicit Differentiation

Given:

$$9x^2 + 4y^2 = 25$$

Differentiate both sides implicitly:

$$18x + 8y \frac{dy}{dx} = 0 \Rightarrow \frac{dy}{dx} = -\frac{18x}{8y} = -\frac{9x}{4y}$$

This method is used when y is not isolated and is defined implicitly.

53 Integration

Integration is the reverse process of differentiation and is used to calculate areas, volumes, accumulated change, and more.

53.1 Definition (Riemann Integral)

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We define the *definite integral* as:

$$\int_a^b f(x)\, dx = \lim_{\|P\| \to 0} \sum_{i=1}^n f(\xi_i) \Delta x_i$$

Where $P = \{x_0, \ldots, x_n\}$ is a partition, $\xi_i \in [x_{i-1}, x_i]$, and $\Delta x_i = x_i - x_{i-1}$.

53.2 Upper and Lower Sums

Let f be bounded on $[a, b]$ and P a partition:

$$\underline{S}(f, P) = \sum_{i=1}^n m_i \Delta x_i, \quad \overline{S}(f, P) = \sum_{i=1}^n M_i \Delta x_i$$

with:

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x), \quad M_i = \sup_{x \in [x_{i-1}, x_i]} f(x)$$

If $\sup \underline{S} = \inf \overline{S}$, then f is Riemann integrable.

53.3 Concrete and Non-Concrete Integrals

- *Concrete:* Can be explicitly evaluated with anti-derivatives.
- *Non-concrete:* Cannot be integrated in elementary terms (e.g., $\int e^{-x^2} dx$).

53.4 Rules of Integration

- $\int 0\, dx = C$
- $\int c\, dx = cx$
- $\int x^n\, dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1)$
- $\int \frac{1}{x}\, dx = \ln |x| + C$
- $\int e^x\, dx = e^x + C$
- $\int \sin x\, dx = -\cos x + C$
- $\int \cos x\, dx = \sin x + C$
- Linearity: $\int (af + bg)\, dx = a \int f\, dx + b \int g\, dx$

53.5 Substitution Rule (U-Substitution)

If $u = g(x)$ and $f = F'$, then:

$$\int f(g(x)) g'(x)\, dx = \int f(u)\, du$$

Example:

$$\int \frac{1}{(x-1)^4}\,dx, \quad u = x - 1 \Rightarrow \int \frac{1}{u^4}du = -\frac{1}{3u^3} + C = -\frac{1}{3(x-1)^3} + C$$

53.6 Integration by Parts

$$\int u\,dv = uv - \int v\,du$$

ILATE: Use this order to choose u:

1. Inverse Trig
2. Logarithmic
3. Algebraic
4. Trigonometric
5. Exponential

Cyclic Integrals: Some integrals cycle back to the original (e.g., $e^x \sin(x)$).

Example:

$\int e^x \sin(x)dx$

Let:

$$u = \sin(x),\; dv = e^x \Rightarrow du = \cos(x)dx,\; v = e^x$$

$$\int e^x \sin(x)dx = e^x \sin(x) - \int e^x \cos(x)\,dx$$

$$u = \cos(x),\; dv = e^x \Rightarrow du = -\sin(x),\; v = e^x$$

$$\int e^x \sin(x)dx = e^x \sin(x) - \left[\cos(x)e^x - \int \sin(x)e^x dx\right]$$
$$\int e^x \sin(x)dx = e^x \sin(x) - \cos(x)e^x - \int \sin(x)e^x dx$$
$$2\int e^x \sin(x)dx = e^x \sin(x) - \cos(x)e^x$$
$$\int e^x \sin(x)dx = \frac{e^x \sin(x) + \cos(x)e^x + C}{2}$$

53.7 Integration by Partial Fractions

Applies to rational functions. Four cases:

- *Linear factors*
  Denominator: $(x - a)(x - b)$
  Decomposition:
  $$\frac{P(x)}{(x-a)(x-b)} = \frac{A}{x-a} + \frac{B}{x-b}$$

- *Repeated linear factors*
  Denominator: $(x - a)^n$
  Decomposition:
  $$\frac{P(x)}{(x-a)^n} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_n}{(x-a)^n}$$

- *Irreducible quadratic factors*
  Denominator: $(x^2 + px + q)$, where $x^2 + px + q$ cannot be factored over the reals
  Decomposition:
$$\frac{P(x)}{x^2 + px + q} = \frac{Ax + B}{x^2 + px + q}$$

- *Repeated irreducible quadratic factors*
  Denominator: $(x^2 + px + q)^n$
  Decomposition:

$$\frac{P(x)}{(x^2 + px + q)^n} = \frac{A_1 x + B_1}{x^2 + px + q} + \frac{A_2 x + B_2}{(x^2 + px + q)^2} + \cdots + \frac{A_n x + B_n}{(x^2 + px + q)^n}$$

Example:

$$\int \frac{1}{x^2 - 1}\,dx = \int \left(\frac{1}{2(x-1)} - \frac{1}{2(x+1)} \right) dx = \frac{1}{2}\ln\left| \frac{x-1}{x+1} \right| + C$$

53.8   Trigonometric Substitution

Trigonometric substitution is useful for evaluating integrals involving square roots of quadratic expressions.

- $\sqrt{a^2 - x^2}$: use $x = a\sin\theta$
- $\sqrt{a^2 + x^2}$: use $x = a\tan\theta$
- $\sqrt{x^2 - a^2}$: use $x = a\sec\theta$

Example I:

$\int \frac{1}{\sqrt{4-x^2}}\,dx$

We use: $x = 2\sin\theta \Rightarrow dx = 2\cos\theta\,d\theta$

$$\sqrt{4 - x^2} = \sqrt{4 - 4\sin^2\theta} = \sqrt{4\cos^2\theta} = 2\cos\theta$$

Substitute:

$$\int \frac{1}{\sqrt{4-x^2}}\,dx = \int \frac{1}{2\cos\theta} \cdot 2\cos\theta\,d\theta = \int d\theta = \theta + C$$

Return to x using:

$$x = 2\sin\theta \Rightarrow \theta = \arcsin\left(\frac{x}{2}\right) \Rightarrow \int \frac{1}{\sqrt{4-x^2}}\,dx = \arcsin\left(\frac{x}{2}\right) + C$$

Example II:

$\int \frac{x^3}{\sqrt{x^2+4}}\,dx$

Use: $x = 2\tan\theta \Rightarrow dx = 2\sec^2\theta\,d\theta$

$$\sqrt{x^2 + 4} = \sqrt{4\tan^2\theta + 4} = \sqrt{4\sec^2\theta} = 2\sec\theta$$

Substitute:
$$\int \frac{x^3}{\sqrt{x^2+4}}\,dx = \int \frac{(2\tan\theta)^3}{2\sec\theta} \cdot 2\sec^2\theta\,d\theta = \int \frac{8\tan^3\theta}{\sec\theta} \cdot 2\sec^2\theta\,d\theta$$

Simplify:

$$8 \int \tan^2 \theta \sec \theta \tan \theta \, dx = 8 \int (\sec^2 \theta - 1) \sec \theta \tan \theta \, dx$$

$$8 \int (\sec^2 \theta \sec \theta - \sec \theta) \tan \theta \, dx$$

$$8 \left(\int \sec^2 \theta \sec \theta \tan \theta \, dx - \int \sec \theta \tan \theta \, dx \right)$$

$$8 \left(\int \sec^2 \theta \sec \theta \, dx - \int \sec \theta \tan \theta \, dx \right)$$

For the first one let $u = \sec^3 \theta$ and $du = 3 \sec^2 \theta \, d\theta$

$$\int u \, du = \frac{\sec^3 \theta}{3} + c$$

For the second one

$$\int \sec \theta \tan \theta = \sec \theta + c$$

Together

$$8 \left(\frac{\sec^3 \theta}{3} - \sec \theta + c \right)$$

Then solve the integral using trigonometric identities and reduction formulas. Return to x using:

$$\tan \theta = \frac{x}{2} \quad \sec \theta = \frac{\sqrt{x^2 + 4}}{2}$$

This leaves us with

$$\frac{8}{3} \left(\frac{\sqrt{x^2 + 4}}{2} \right)^3 - 8 \left(\frac{\sqrt{x^2 + 4}}{2} \right) + c$$

Example III:

$\int \frac{x^3}{\sqrt{x^4 - 4}} \, dx$

Use: $x = 2 \sec \theta$, $dx = 2 \sec \theta \tan \theta \, d\theta$

Then

$$\int \frac{(2 \sec \theta)^2 2 \sec \theta \tan \theta \, d\theta}{2 \tan \theta} = \int 8 \sec^4 \theta \, d\theta$$

$$8 \int (\tan^2 \theta + 1) \sec^2 \theta = 8 \left(\int \tan^2 \theta \sec^2 \theta + \int \sec^2 \theta \right)$$

Now let $u = \tan \theta$ and $du = \sec^2 \theta \, d\theta$

$$8 \left(\int u \, du + \int \sec^2 \theta \, d\theta \right) = \frac{8}{3} \tan^3 \theta + 8 \tan \theta + c$$

Now we return to x via our triangle:

$$\tan \theta = \frac{\sqrt{x^2 + 4}}{2}$$

$$\frac{8}{3}\left(\frac{\sqrt{x^2 + 4}}{2}\right)^3 + 8\frac{\sqrt{x^2 + 4}}{2} + c$$

53.9   Average Value of a Function

$$\bar{f} = \frac{1}{b - a} \int_a^b f(x)\,dx$$

53.10   Mean Value Theorem for Integrals

If f is continuous on $[a, b]$, then:

$$\exists c \in [a, b] \text{ such that } \int_a^b f(x)\,dx = f(c)(b - a)$$

53.11   Length, Area, Volume Formulas

Arc length:

$$L = \int_a^b \sqrt{1 + f'(x)^2}\,dx$$

Area between curves:

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

Volume (disk):

$$V = \pi \int_a^b f(x)^2\,dx$$

Shell:

$$V = 2\pi \int_a^b x f(x)\sqrt{1 + f'(x)^2}\,dx$$

53.12   Differentiation of Integrals with Variable Bounds

If:

$$F(x) = \int_{a(x)}^{b(x)} f(t)\,dt$$

Then:

$$F'(x) = f(b(x)) \cdot b'(x) - f(a(x)) \cdot a'(x)$$

53.13   Parameter-Dependent Integrals

Let $F(p) = \int_a^b f(x, p)\, dx$. If f is continuous and differentiable in p, then:

$$\frac{d}{dp} \int_a^b f(x, p)\, dx = \int_a^b \frac{\partial f}{\partial p}(x, p)\, dx$$

53.14   Leibniz Rule for Differentiating Under the Integral Sign

$$\frac{\partial F}{\partial x} \int_{a(x)}^{b(x)} f(x, t)\, dx = f(x, b(x)) \cdot b'(x) - f(x, a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f}{\partial t}(x, t)\, dt$$

53.15   Improper Integrals with infinite Bounds

$$\int_a^\infty f(x)\, dx = \lim_{R \to \infty} \int_a^R f(x)\, dx$$

The value when the bounds look like $\int_{-\infty}^\infty f(x)\, dx$ is called the *Cauchy-Main-Value*.

$$\int_{-\infty}^\infty f(x)\, dx = \lim_{R \to \infty} \int_{-R}^c f(x)\, dx + \int_c^R f(x)\, dx$$

53.16   Improper Integrals with Discontinuity

These integrals have some kind of asymptote between the bounds we want to evaluate. More concrete:

$$\int_a^b f(x)\, dx \text{ improper if } f \text{ has an infinite discontinuity on } [a, b]$$

To tackle this type of problems we just like before use a limit to approach the discontinuity without reaching it.

$$\int_a^b dx = \lim_{\varepsilon \to 0} \int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx$$

Here c is the value of the discontinuity.

Example:

$$\int_0^1 \frac{1}{2\sqrt{x}}$$

Discontinuity at $c = 0$

$$
\begin{aligned}
\int_0^1 \frac{1}{2\sqrt{x}} &= \lim_{\varepsilon \to 0} \int_\varepsilon^1 \frac{1}{2\sqrt{x}} dx \\
&= \lim_{\varepsilon \to 0} \sqrt{x}\big|_\varepsilon^1 \\
&= 1 - \lim_{\varepsilon \to 0} \sqrt{\varepsilon} = 1
\end{aligned}
$$

53.17   Convergence Criteria for Improper Integrals

- *Majorant:* If $0 \leq f(x) \leq g(x)$ and $\int g(x)\, dx$ converges, then so does $\int f(x)\, dx$.
- *Minorant:* If $f(x) \geq g(x) \geq 0$ and $\int g(x)\, dx$ diverges, then $\int f(x)\, dx$ diverges.

53.18 Integral Test for Series

Let $f(n) = a_n$ where f is continuous, decreasing, and positive for $n \geq N$:

$$\sum_{n=N}^{\infty} a_n \text{ converges} \iff \int_{N}^{\infty} f(x)\, dx \text{ converges}$$

Example:

$\sum \frac{1}{n^p}$ converges $\iff p > 1$

54 Important Theorems of Single Variable Calculus

54.1 Intermediate Value Theorem

Let f be a continuous function on the closed interval $[a, b]$. If $f(a) \neq f(b)$ and N is any number between $f(a)$ and $f(b)$, then there exists $c \in (a, b)$ such that $f(c) = N$.

Proof:

Without loss of generality, assume $f(a) < N < f(b)$. Define:

$$S = \{x \in [a, b] \mid f(x) \leq N\}$$

Since $a \in S$, S is non-empty and bounded above by b. Let $c = \sup S$.

By the continuity of f, we have:

$$\lim_{x \to c^-} f(x) = f(c) = \lim_{x \to c^+} f(x)$$

Since $f(x) \leq N$ for all $x < c$, and for any $\varepsilon > 0$, there exists $x > c$ with $f(x) > N$, then by the definition of supremum and continuity:

$$f(c) = N$$

$$\mathfrak{QED}$$
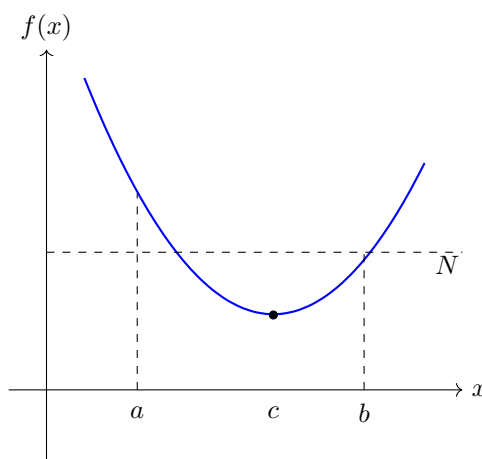
54.2 Extreme Value Theorem

If f is continuous on the closed interval $[a, b]$, then f attains a maximum and a minimum value on $[a, b]$. That is, there exist $c, d \in [a, b]$ such that:

$$f(c) \leq f(x) \leq f(d) \quad \text{for all } x \in [a, b]$$

Proof:

Since f is continuous on the compact interval $[a, b]$, the image $f([a, b]) \subset \mathbb{R}$ is also compact. A compact subset of $\mathbb{R}$ is closed and bounded, and hence attains its supremum and infimum.

Therefore, there exist $c, d \in [a, b]$ such that:

$$f(c) = \min\{f(x) : x \in [a, b]\}, \quad f(d) = \max\{f(x) : x \in [a, b]\}$$

$$\mathfrak{QED}$$

54.3   Rolle's Theorem

Let f be continuous on $[a, b]$, differentiable on (a, b), and suppose $f(a) = f(b)$. Then, there exists $c \in (a, b)$ such that $f'(c) = 0$.

Proof:

By the Extreme Value Theorem, f attains a maximum or minimum at some point $c \in [a, b]$.

If $c \in (a, b)$, then by Fermat's theorem, $f'(c) = 0$. If the maximum or minimum occurs at a or b, since $f(a) = f(b)$, f must be constant, so $f'(x) = 0$ everywhere on (a, b), and in particular for some $c \in (a, b)$.

$$\mathfrak{QED}$$
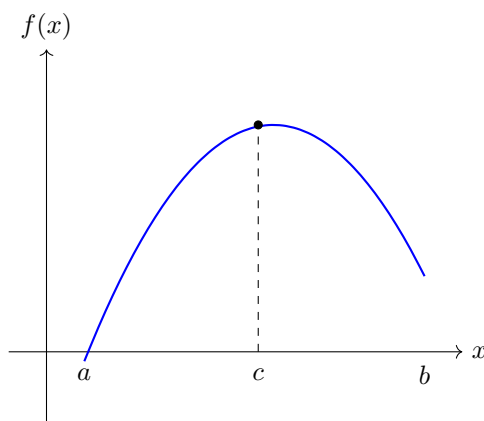
54.4   Mean Value Theorem

Let f be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof:

Define the auxiliary function:

$$g(x) = f(x) - \left(\frac{f(b) - f(a)}{b - a} \right)(x - a)$$

Then $g(a) = f(a)$, $g(b) = f(b) - m(b - a) = f(a) \Rightarrow g(a) = g(b)$

Apply Rolle's theorem to $g(x)$: Since $g(a) = g(b)$, and g is continuous and differentiable, there exists $c \in (a, b)$ such that $g'(c) = 0$

Then:

$$g'(x) = f'(x) - \left(\frac{f(b) - f(a)}{b - a} \right) \Rightarrow f'(c) = \frac{f(b) - f(a)}{b - a}$$

$$\mathfrak{QED}$$
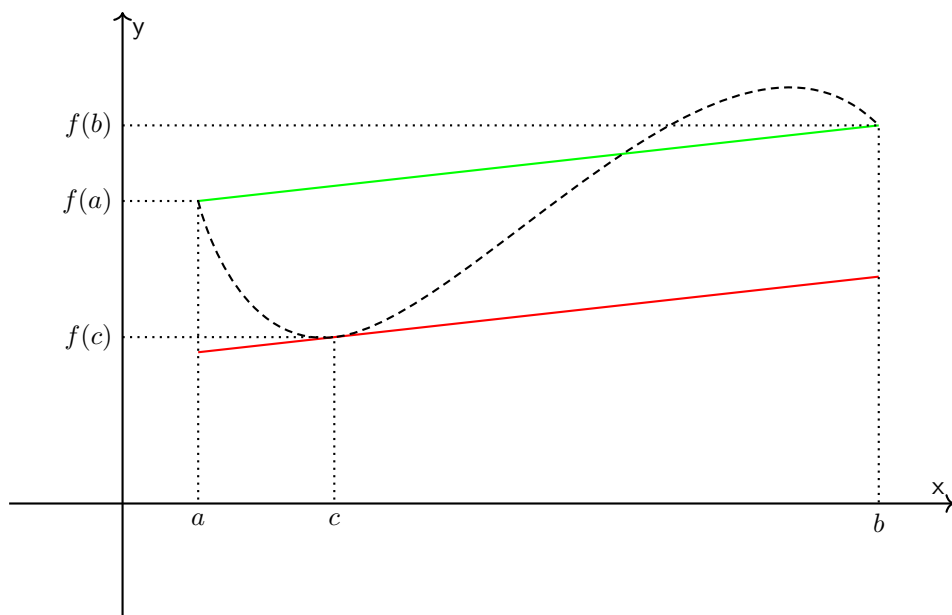
55 Coordinate Systems

Coordinate systems provide different ways of describing points in space, adapted to the geometry of the problem. Below are the most common systems used in multi-variable calculus.

55.1 Cartesian Coordinates

The *Cartesian coordinate system* uses mutually orthogonal axes. A point in $\mathbb{R}^2$ is given by:

$$(x, y) \quad \text{and in } \mathbb{R}^3: \quad (x, y, z)$$

This is the most natural system for rectangular domains and is the basis for most definitions in vector calculus. The coordinate axes are orthogonal and equidistant.

55.2 Polar Coordinates

In polar coordinates, a point $(x, y) \in \mathbb{R}^2$ is described by:

$$r \geq 0 \text{ (radius from origin)}, \quad \varphi \in [0, 2\pi) \text{ (angle from x-axis)}$$

Transformations:

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

This comes from our right triangle with $r = 1$ where

$$\cos\varphi = \frac{x}{r} \quad \sin\varphi = \frac{y}{r}$$

Inverse:

$$r = \sqrt{x^2 + y^2}, \quad \varphi = \arctan\left(\frac{y}{x}\right)$$

This system is useful for circular or radial symmetry. The Jacobian determinant is:

$$|J| = r\, dr\, d\theta$$

55.3 Cylindrical Coordinates

Cylindrical coordinates extend polar coordinates by adding a height z component:

$$(r, \varphi, z) \quad \text{where } r \geq 0, \ \varphi \in [0, 2\pi), \ z \in \mathbb{R}$$

Transformations:

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z$$

Inverse:

$$r = \sqrt{x^2 + y^2}, \quad \varphi = \arctan\left(\frac{y}{x}\right), \quad z = z$$

This coordinate system is ideal for objects with rotational symmetry about the z-axis. The Jacobian determinant is:

$$|J| = r\, dr\, d\theta\, dz$$

55.4   Spherical Coordinates

In spherical coordinates, a point $(x, y, z) \in \mathbb{R}^3$ is represented using:

$$\rho \geq 0 \text{ (radial distance)}, \quad \theta \in [0, \pi] \text{ (inclination)}, \quad \varphi \in [0, 2\pi) \text{ (azimuth)}$$

Transformations:

$$x = \rho \sin \theta \cos \varphi, \quad y = \rho \sin \theta \sin \varphi, \quad z = \rho \cos \theta$$

Inverse:

$$\rho = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \arccos\left(\frac{z}{\rho}\right), \quad \varphi = \arctan\left(\frac{y}{x}\right)$$

Spherical coordinates are useful for spherical symmetry (e.g., in gravitational or electric fields). The Jacobian determinant is:

$$|J| = \rho^2 \sin \theta \, d\rho \, d\phi \, d\theta$$

Each coordinate system is tailored to specific geometries and greatly simplifies integrals when used appropriately. Coordinate transformations require careful application of the Jacobian determinant to account for stretching or distortion of space.

56 Parameterized Curves

56.1 Definition

A *parameterized curve* is a mathematical object where each point on the curve is described as a function of one or more parameters, typically denoted by t. Formally, a parameterized curve in $\mathbb{R}^n$ is a continuous function:

$$\mathbf{r}(t) = (x(t), y(t), z(t), \ldots)$$

Where t ranges over some interval $I \subseteq \mathbb{R}$, and each component function (e.g., $x(t)$, $y(t)$) is real-valued.

56.2 Purpose

The purpose of parameterizing a curve is to provide a flexible and detailed way of describing geometric objects, particularly when they cannot be easily expressed as a simple function $y = f(x)$. Parameterizations allow us to:

- Describe curves that loop, cross themselves, or have vertical segments.
- Compute quantities like arc length, tangent vectors, and curvature.
- Easily model motion along a curve, where t can represent time.

56.3 From parametric to explicit form

This process is rather simple because in the case where it is possible to change from one form tho the other we only need to write the parameter in terms of one of the variables and then substitute.

56.4 How to Parameterize a Function

To parameterize a curve or a function:

1. Introduce a new parameter t, often representing time or distance along the curve.
2. Express the coordinates (x, y) (or (x, y, z)) as functions of t.
3. Ensure that as t varies over a chosen interval, the set of points $(x(t), y(t))$ traces the desired curve.

Sometimes, parameterizations are natural (e.g., circles using trigonometric functions), while in other cases, one must carefully design a suitable parameterization based on the curve's properties.

Example:

Parameterizing a Circle:

Consider the unit circle defined by the equation:

$$x^2 + y^2 = 1$$

A natural parameterization uses the trigonometric functions sine and cosine:

$$x(t) = \cos(t), \quad y(t) = \sin(t)$$

For $t \in [0, 2\pi]$.

As t varies from 0 to 2π, the point $(x(t), y(t))$ traces out the entire circle exactly once.

56.5   Arclength and t Parameter

Recall that the arclength of a curve in 3 dimensions is

$$\int_a^b \sqrt{f'(t9)^2 + g'(t)^2 + h'(t)^2} dt,$$

which is norm of our tangent vector.

$$\int_a^b \sqrt{|\vec{v}(t)|} dt$$

56.5.1   Arclength Parameter

This is given by

$$s(t) = \int_a^t \sqrt{|\vec{\tau}(t)|} dt,$$

where τ is the length parameter.

This parameter is intrinsic to the curve, and represents one choice because there only one valid interpretation of the length. The downside is that this is hard to compute.

56.5.2   The t Parameter

For the t parameter, we can say that this is intrinsic to the object along the curve, it can be interpreted in multiple ways therefore, many choices and is easy to compute.

57 Multi-variable Functions

57.1 Open Sets

Let $\| \cdot \|$ be a norm on $\mathbb{R}^n$. The ε-neighborhood of a point $\vec{x}_0 \in \mathbb{R}^n$ is defined as:

$$U_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{x}_0\| < \varepsilon\}$$

A point $\vec{x}_0 \in D \subseteq \mathbb{R}^n$ is called an *interior point* of D if there exists $\varepsilon > 0$ such that $U_\varepsilon(\vec{x}_0) \subseteq D$. A set D is called *open* if all its points are interior points.

Although ε-neighborhoods depend on the chosen norm, due to the equivalence of norms in $\mathbb{R}^n$, openness is norm-independent.

57.2 Closed Sets

Let $D \subseteq \mathbb{R}^n$ and $\| \cdot \|$ a norm. A point $\vec{x}_0$ is called an *accumulation point* (or limit point) of D if for every $\varepsilon > 0$, the neighborhood $U_\varepsilon(\vec{x}_0)$ contains a point $\vec{x} \neq \vec{x}_0$ such that $\vec{x} \in D$.

A set $D \subseteq \mathbb{R}^n$ is called *closed* if it contains all its accumulation points.

57.3 Boundedness and Order

A set $D \subseteq \mathbb{R}^n$ is called *bounded* if there exists a real number $M > 0$ such that:

$$\|\vec{x}\| < M \quad \text{for all } \vec{x} \in D$$

If no such bound M exists, the set is called *unbounded*.

57.4 Sequences

Let $(\vec{x}_n) \subseteq \mathbb{R}^n$ be a sequence of vectors. We say that $\vec{x}_n \to \vec{x} \in \mathbb{R}^n$ as $n \to \infty$ if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that:

$$\|\vec{x}_n - \vec{x}\| < \varepsilon \quad \text{for all } n > n_0$$

The sequence $(\vec{x}_n)$ is called *convergent* if such $\vec{x}$ exists.

The sequence $(\vec{x}_n)$ is called a *Cauchy sequence* if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that:

$$\|\vec{x}_n - \vec{x}_m\| < \varepsilon \quad \text{for all } n, m > n_0$$

Example I:

$$\lim_{n \to \infty} \vec{X}_n = \lim_{n \to \infty} \begin{bmatrix} \frac{1}{n} \\ 1 + \frac{1}{n^2} \\ \frac{\sin n}{n} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

57.5 Different Norms

From linear algebra it is known what a *Norm* is and its properties:

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A function $\| \cdot \| : V \to \mathbb{R}$ is called a norm if for all $x, y \in V$ and all scalars α, the following properties hold:

1. *Non-negativity:* $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$.

2. *Absolute homogeneity:* $\|\alpha x\| = |\alpha| \cdot \|x\|$.

3. *Triangle inequality:* $\|x + y\| \leq \|x\| + \|y\|$.

Commonly used norms in $\mathbb{R}^n$ (or $\mathbb{C}^n$) include:

- *Sum Norm* (also known as the ℓ_1 norm):

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

- *Max Norm* (also known as the ℓ_∞ norm):

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$

- *Euclidean Norm* (also known as the ℓ_2 norm):

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$$

Example II:

Given is

$$\vec{X}_n = \begin{bmatrix} 5 - \frac{10}{n} \\ \frac{1}{n} \end{bmatrix}$$

we want to find n_0 for $\varepsilon = 0,1$

$$\lim_{n \to \infty} \vec{X}_n = \begin{bmatrix} 5 \\ 0 \end{bmatrix} = \vec{X}_0$$

Now use the definition for convergence and *Euclidean Norm*

$$\|\vec{X}_n - \vec{X}_0\| < \varepsilon$$

$$\left\| \begin{bmatrix} -\frac{10}{n} \\ \frac{1}{n} \end{bmatrix} \right\|$$

$$\frac{\sqrt{101}}{n} < 0,1$$
$$n > 0,1\sqrt{101} = 100,5$$

Therefore, $n_0 = 101$.

57.5.1 Equivalence of Norms

In finite-dimensional vector spaces, all norms are equivalent. That is, if $\| \cdot \|_a$ and $\| \cdot \|_b$ are two norms on a finite-dimensional vector space V, then there exist constants $c, C > 0$ such that for all $x \in V$,

$$c\|x\|_a \le \|x\|_b \le C\|x\|_a$$

This implies that while different norms may measure length differently, they all induce the same topology and convergence behavior in finite-dimensional spaces.

57.5.2 Equivalence in the Reals

In the $\mathbb{R}^n$ all norms are equivalent and

$$\|\vec{x}\|_\infty \le \|\vec{x}\|_2 \le \sqrt{n}\|\vec{x}\|_\infty$$
$$\|\vec{x}\|_\infty \le \|\vec{x}\|_1 \le n\|\vec{x}\|_\infty$$
$$\|\vec{x}\|_2 \le \|\vec{x}\|_1 \le n\|\vec{x}\|_2$$

57.6 Accumulation Point

A vector $\vec{x}_0 \in \mathbb{R}^n$ is called an *accumulation point* of a set $D \subseteq \mathbb{R}^n$ if for every $\varepsilon > 0$, the neighborhood $U_\varepsilon(\vec{x}_0)$ contains a point $\vec{x} \in D \setminus \{\vec{x}_0\}$.

In the context of sequences, an accumulation point of a sequence $(\vec{x}_n)$ is a point $\vec{x} \in \mathbb{R}^n$ for which there exists a subsequence $(\vec{x}_{n_k})$ such that $\vec{x}_{n_k} \to \vec{x}$.

57.7 Bolzano–Weierstraß Theorem

Every infinite, bounded sequence in $\mathbb{R}^n$ has at least one accumulation point.

- Every bounded infinite sequence has at least one convergent subsequence.

- A bounded sequence is *convergent* if and only if it has exactly one accumulation point.

- Every Cauchy sequence in $\mathbb{R}^n$ is convergent (since $\mathbb{R}^n$ is complete).

57.8 Limits

A function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ has the limit $g \in \mathbb{R}$ at a point $\vec{X}_0 \in U$ if:

$$\lim_{\vec{X} \to \vec{X}_0} f(\vec{X}) = g$$

means that for every sequence $(\vec{X}_n) \subset U$ converging to $\vec{X}_0$, we have:

$$\lim_{n \to \infty} f(\vec{X}_n) = g$$

The convergence must hold along all possible paths to the point $\vec{X}_0$, making the multi-variable limit path-independent and unique (if it exists).

Example:

$$\vec{X}_n = \begin{bmatrix} \frac{1}{n} \\ 1 + \frac{1}{n^2} \\ \frac{\sin(n)}{n} \end{bmatrix}$$

The limit towards infinity is then

$$\lim_{n \to \infty} \begin{bmatrix} \frac{1}{n} \\ 1 + \frac{1}{n^2} \\ \frac{\sin(n)}{n} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

57.9 Partial Functions

For a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, a *partial function* is a restriction of f to a variable by fixing all others.

Examples in $\mathbb{R}^3$:

$$f(x_1, x_2, x_3) = 5x_1 + x_2 x_3$$

Fixing values:

$$x_2 = 3, \quad x_3 = 5 \Rightarrow f_1(x_1) = 5x_1 + 15$$

$$x_1 = 2, \quad x_3 = 5 \Rightarrow f_2(x_2) = 10 + 5x_2 \quad x_1 = 2, \quad x_2 = 3 \Rightarrow f_3(x_3) = 10 + 3x_3$$

Each f_i reduces the multi-variable function to a single-variable slice (cross-section), allowing analysis along coordinate axes.

57.10   Continuity

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, and $\vec{X}_0 \in U$. The function f is said to be *continuous at* $\vec{X}_0$ if:

$$\lim_{\vec{X} \to \vec{X}_0} f(\vec{X}) = f(\vec{X}_0)$$

This means that for any sequence $(\vec{X}_n) \to \vec{X}_0$, we have:

$$\lim_{n \to \infty} f(\vec{X}_n) = f(\vec{X}_0)$$

To check if a multi-variable function is continuous it is enough to prove that it is continuous towards the origin

$$\lim_{\vec{X} \to \vec{X}_0} f(\vec{X}) = f(\vec{X}_0)$$

$$\iff \lim_{\vec{X} \to \vec{X}_0} \left(f(\vec{X}) - f(\vec{X}_0) \right) = 0$$

$$\iff \lim_{\vec{h} \to 0} \left(f(\vec{X}_0 + \vec{h}) - f(\vec{X}_0) \right) = 0$$

$$\iff \lim_{\vec{h} \to 0} g(\vec{h}) = 0$$

If we are checking if a function is continuous and get an *indeterminate form* there are three main approaches.

1. *Factorizing*: Either normal factorization, adding a clever zero or multiplying by a clever one.

2. *Change of Coordinates*: Sometimes changing the coordinate system allows us to simplify the expression.

3. *Different Paths*: Trying different paths can be pretty straightforward. If wet two different values for the same point across different paths then the function is not continuous.

Example:

Given is $\frac{x^3 + 2yx^2 + y^2x + 2y^3}{x + 2y}$ for $(x, y) \neq (0, 0)$ and $0 for (0, 0)$.

$$\frac{x^3 + 2yx^2 + y^2x + 2y^3}{x + 2y} = \left(\frac{x^2(x + 2y) + y(x + 2y)}{x + 2y} \right.$$

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

$$= x^2 + y^2$$

Now by taking the limit we see that

$$\lim_{(x,y) \to (0,0)} x^2 + y^2 = 0$$

Thus, the function is continuous.

Example:

Check for continuity towards the origin of $f(x, y) = \frac{xy^2}{x^2+y^4}$.

For the path $x = y$ and $y = 0$

$$f(x, 0) = \frac{x0}{x^2 + 0} = 0$$

For the path $x = y^2$

$$f(y^2, y) = \frac{y^4}{2y^4} = \frac{1}{2}$$

Because we got two different values the function is not continuous.

57.10.1 Continuity in Polar Coordinates for 2D

Let $\vec{X} = (x, y)$ be written in polar coordinates:

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

Then:

$$\vec{X}_n \to (0, 0) \iff r_n \to 0$$

This alternative approach helps evaluate limits and continuity at the origin using radial convergence.

Example:

Given is

$$f(xy) = \begin{cases} \frac{xy}{x^2+y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{else} \end{cases}$$

Let $x = r \cos(\theta)$ and $y = r \sin(\theta)$

$$\lim_{r \to 0} \frac{r \cos(\theta) r \sin(\theta)}{r^2 \cos^2(\theta) + r^2 \sin^2(\theta)}$$
$$= \lim_{r \to 0} \frac{r^2 \cos(\theta) \sin(\theta)}{r^2(\cos^2(\theta) + \sin^2(\theta))}$$
$$= \lim_{r \to 0} \frac{\cos(\theta) \sin(\theta)}{1}$$

Because of the dependence on θ which can give us for $\theta = \frac{\pi}{4}$ $\frac{1}{2}$ this functions is not continuous.

57.10.2 Substitution

When dealing with multi-variable functions we can sometimes use *substitution* for converting our function into a single variable function to use *L'Hospital's Rule*.

Example:

Given is $\frac{e^{x^2+y^2}-1}{x^2+y^2}$.

See, that we have repeated terms we can substitute. $p^2 = x^2 + y^2$.

$$\frac{e^{p^2} - 1}{p^2}$$

Now differentiate

$$\frac{2pe^{p^2}}{2p} = e^{p^2}$$

Now take the limit

$$\lim_{p \to 0} e^{p^2} = 1$$

57.11 Uniform and Lipschitz Continuity

57.11.1 Uniform Continuity:

A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is called *uniformly continuous* on D if:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that } \|\vec{X} - \vec{X}_0\| < \delta \Rightarrow |f(\vec{X}) - f(\vec{X}_0)| < \varepsilon$$

The key feature is that δ depends only on ε, not on the points $\vec{X}, \vec{X}_0$ themselves. If f is continuous on a compact set (closed and bounded), then f is uniformly continuous.

57.11.2 Lipschitz Continuity:

A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is *Lipschitz continuous* if there exists a constant $L > 0$ such that:

$$|f(\vec{X}) - f(\vec{Y})| \leq L\|\vec{X} - \vec{Y}\| \quad \text{for all } \vec{X}, \vec{Y} \in D$$

If $L < 1$, then f is called a *contraction*.

57.12 Epsilon–Delta Criterion

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, and $\vec{X}_0 \in D$. The function f is continuous at $\vec{X}_0$ if:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that } \|\vec{X} - \vec{X}_0\| < \delta \Rightarrow |f(\vec{X}) - f(\vec{X}_0)| < \varepsilon$$

This generalizes the familiar epsilon-delta criterion from single-variable calculus to multi-variable settings using norms.

57.13 Fixpoints

A point $\vec{X}_0 \in D \subseteq \mathbb{R}^n$ is called a *fix-point* of a function $\vec{\varphi} : D \to \mathbb{R}^n$ if:

$$\vec{\varphi}(\vec{X}_0) = \vec{X}_0$$

That is, applying the function does not change the point — it maps to itself.

57.14 Banach Fixed Point Theorem

Let $\vec{\varphi} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a *contraction mapping*, i.e., there exists a constant $L < 1$ such that:

$$\|\vec{\varphi}(\vec{X}) - \vec{\varphi}(\vec{Y})\| \leq L \|\vec{X} - \vec{Y}\| \quad \text{for all } \vec{X}, \vec{Y} \in D$$

Then:

- There exists a unique fix-point $\vec{X}^* \in D$ such that $\vec{\varphi}(\vec{X}^*) = \vec{X}^*$
- Iteratively defined sequences $\vec{X}_{n+1} = \vec{\varphi}(\vec{X}_n)$ converge to the fix-point

This result is fundamental in nonlinear analysis and iterative numerical methods.

57.15 Vector fields

A *vector field* is a function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^n$. Also, just a function with vectors as inputs and outputs.

$$\vec{f}(x_1, x_2, \ldots, x_n) = A_1(x_1, \ldots, x_n)\vec{e}_1 + \cdots + A_n(x_1, \ldots, x_n)\vec{e}_n$$

We call a vector field *continuous/differentiable* if all functions

$$A_1, A_2, \ldots, A_n$$

are continuous/differentiable.

As a side node, the *gradient* of a multi-variable function is a vector field.

58 Multi-variable Differentiation

58.1 Partial Derivatives

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ and $\vec{X}_0 = (x_1^{(0)}, \ldots, x_n^{(0)}) \in U$. Then f is *partially differentiable with respect to* x_i at $\vec{X}_0$ if the limit:

$$\frac{\partial f}{\partial x_i}(\vec{X}_0) := \lim_{h \to 0} \frac{f(x_1^{(0)}, \ldots, x_i^{(0)} + h, \ldots, x_n^{(0)}) - f(\vec{X}_0)}{h}$$

exists. This derivative measures the rate of change of f in the direction of the x_i-axis while keeping all other variables fixed.

A function is *(partially) differentiable* at $\vec{X}_0$ if all partial derivatives $\frac{\partial f}{\partial x_i}(\vec{X}_0)$ exist.

58.2 Differentiability

We call a function f differentiable at $(x_1, x_2, \cdots, x_n)$ if and only if:

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots, x_n + \Delta x_n) - f(x_1, x_2, \ldots, x_n) =$$

$$\frac{\partial f(x_1, y_2, \ldots, x_n)}{\partial x_1} \Delta x_1 + \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_2} \Delta x_2 + \cdots + E_1(\Delta x_1) + E_2(\Delta x_n) + \cdots + E_n(\Delta x_n)$$

where $\lim_{\Delta x_1 \to 0} \frac{E_1(\Delta x_1)}{\Delta x_1} = \lim_{\Delta x_2 \to 0} \frac{E_2(\Delta x_2)}{\Delta x_2} = \cdots = 0$

58.3 The Gradient

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\vec{X}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$. The *gradient* of f at $\vec{X}_0$ is:

$$\nabla f(\vec{X}_0) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\vec{X}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{X}_0) \end{bmatrix}$$

The angle of the maximal slope is $\phi = \frac{\pi}{2}$. This comes from the fact that the directional derivative which is defined as $\langle \nabla f, \vec{v} \rangle = \|\nabla f\| \|\vec{v}\| \cos(\phi)$ is orthogonal to the gradient. And for smallest slope $\phi = \frac{3\pi}{4}$. Also, the *gradient* forms a vector field.

Formally, we have a vector d^* which points in the direction of the max slope defined by

$$d^* = \arg_{\|d\|=1} \max D_{\vec{v}} f(\vec{x})$$

where $\vec{v}$ is a unit vector.

The directional derivative written as the inner product wit the gradient gives us

$$d^* = \arg_{\|d\|=1} \max \langle \nabla f, \vec{v} \rangle$$

Then using the definition of the dot product

$$d^* = \arg_{\|d\|=1} \max \|\nabla f\| \|\vec{v}\| \cos(\phi)$$

Finally, because the magnitude of $\vec{v}$ is one and to maximize cosine we set $\phi = \frac{\pi}{2}$

$$d^* = \|\nabla f\|$$

Therefore, the magnitude of the gradient is the maximal slope.

58.3.1 Origin of the formula

Suppose along the curve $r(t) = x_1(t)\hat{v}_1 + x_2(t)\hat{v}_2 + \cdots + x_n(t)\hat{v}_n$ that $f(x_1(t), \ldots, x_n(t)) = C$

$$\frac{\partial}{\partial t} f(x_1(t), \ldots, x_n(t)) = \frac{\partial}{\partial t} C$$

$$\frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt} = 0$$

$$\left\langle \left(\frac{\partial f}{\partial x_1}\hat{v}_1 + \cdots + \frac{\partial f}{\partial x_n}\hat{v}_n \right), \left(\frac{dx_1}{dt}\hat{v}_1 + \cdots + \frac{dx_n}{dt}\hat{v}_n \right) \right\rangle = 0$$

From the formula we see that: ∇f is the normal and $\frac{d\vec{r}}{dt}$ the tangent vector of the curve

$$\mathfrak{QED}$$

58.3.2 Gradient Operations:

- $\nabla(f + g) = \nabla f + \nabla g$

- $\nabla(\lambda f) = \lambda \nabla f$ for $\lambda \in \mathbb{R}$

- $\nabla(fg) = f\nabla g + g\nabla f$

The $\frac{\operatorname{grad} f}{\|\operatorname{grad} f\|}$ points in the direction of the steepest increase of f and $-\frac{\operatorname{grad} f}{\|\operatorname{grad} f\|}$ in the lowest increase.

58.4 The Tangent Plane

The tangent plane to a differentiable surface $z = f(x, y)$ at the point (x_0, y_0) is the plane that best approximates the surface near that point. It is given by the linearization of f:

$$T(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

This plane touches the graph of the function at a single point
and shares its slope in both x- and y-directions.

It can also be written in the parameterized form:

$$T(x, y) = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \\ f_x(x_0, y_0) \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 1 \\ f_y(x_0, y_0) \end{bmatrix}$$

58.4.1 Generalization

$$T(\vec{X}) = f(\vec{x_0}) + \sum_{i=1}^{n} f_{x_i}(\vec{x_0})(x_i - x_i^0)$$

$$= f(\vec{x_0}) + \langle \nabla f, (\vec{x} - \vec{x_0}) \rangle$$

Example: $f(x, y) = 4x^2 y + xy^2 + 1$

Compute partial derivatives:

$$f_x(x, y) = 8xy + y^2, \quad f_y(x, y) = 4x^2 + 2xy$$

At the point $(1, 1)$:

$$f(1, 1) = 4(1)^2(1) + 1(1)^2 + 1 = 6 f_x(1, 1) = 8(1)(1) + 1 = 9, \quad f_y(1, 1) = 4(1)^2 + 2(1)(1) = 6$$

Thus, the tangent plane is:

$$T(x, y) = 6 + 9(x - 1) + 6(y - 1) \Rightarrow T(x, y) = 9x + 6y - 9$$

58.5 The Directional Derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\vec{x}_0$, and let $\vec{v} \in \mathbb{R}^n$ be a direction vector with $\|\vec{v}\| = 1$. The *directional derivative* of f at $\vec{x}_0$ in the direction $\vec{v}$ is defined as:

$$D_{\vec{v}} f(\vec{x}_0) := \lim_{h \to 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h}$$

Formula:

$$D_{\vec{v}} f(\vec{x}_0) = \langle \nabla f(\vec{x}_0), \vec{v} \rangle$$

Schwarz Inequality (Cauchy–Schwarz):

$$|\langle \vec{a}, \vec{b} \rangle| \leq \|\vec{a}\| \cdot \|\vec{b}\|$$

This implies:

$$|D_{\vec{v}} f(\vec{x}_0)| \leq \|\nabla f(\vec{x}_0)\|$$

58.5.1 Derivation of the Directional Derivative Formula

Let $f(x, y)$ be differentiable, and let $\vec{v} = (e_1, e_2)$ be a direction vector. We aim to derive the formula for the directional derivative $D_{\vec{v}} f(x_0, y_0)$.

We begin by considering the tangent plane to the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$. This plane is spanned by the vectors:

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ f_x(x_0, y_0) \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ f_y(x_0, y_0) \end{bmatrix}$$

These correspond to directional derivatives along the x- and y-axes, respectively.

Now consider the direction vector $\vec{v} = (e_1, e_2)$. Lift it into 3D space (onto the tangent plane) as:

$$\vec{u}_3 = \begin{bmatrix} e_1 \\ e_2 \\ D_{\vec{v}} f \end{bmatrix}$$

Since all three vectors lie in the same plane, the determinant of the matrix formed by these vectors as columns must vanish:

$$\det(\vec{u}_1, \vec{u}_2, \vec{u}_3) = 0$$

Explicitly:

$$\det \begin{bmatrix} 1 & 0 & e_1 \\ 0 & 1 & e_2 \\ f_x(x_0, y_0) & f_y(x_0, y_0) & D_{\vec{v}} f \end{bmatrix} = 0$$

Expanding the determinant gives:

$$f_x(x_0, y_0) \cdot e_2 - f_y(x_0, y_0) \cdot e_1 + D_{\vec{v}}f = 0$$

Solving for $D_{\vec{v}}f$, we obtain:

$$D_{\vec{v}}f = f_x(x_0, y_0) \cdot e_1 + f_y(x_0, y_0) \cdot e_2$$

Vector notation:

$$D_{\vec{v}}f = \langle \nabla f(x_0, y_0), \vec{v} \rangle$$

This is the desired result: the directional derivative equals the dot product of the gradient of f and the direction vector $\vec{v}$.

58.6 The Total Differential

If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{x}_0$, then the *total differential* of f at $\vec{x}_0$ is the linear approximation:

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}_0) \, dx_i = \langle \nabla f(\vec{x}_0), d\vec{x} \rangle$$

This expresses how small changes in the input variables propagate into changes in the function values.

58.7 The Chain Rule

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, and suppose $\vec{x} = \vec{x}(t) \in \mathbb{R}^n$ is a differentiable path. Then:

$$\frac{d}{dt}f(\vec{x}(t)) = \left\langle \nabla f(\vec{x}(t)), \frac{d\vec{x}(t)}{dt} \right\rangle$$

This is the chain rule in vector form.

Or to be more explicit, the small change in t makes a change in $x_1, x_2, \cdots$ and the addition of all this changes sum up to the total change in the original function.

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots + \frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial t}$$

58.8 Absolute and Relative Error

Given an approximate value $\tilde{x}$ for the exact value x:

- *Absolute error:* $\Delta z_{\max} \leq |f_{x_1}||\Delta x_1| + \cdots + |f_{x_n}||\Delta x_n|$
- *Relative error:* $\varepsilon = \frac{\Delta z_{\max}}{z}$ with dependence on the error of the inputs $\frac{\Delta x}{x}$ etc.

In multi-variable contexts, similar formulas apply using vector norms.

Absolute Error Example

We consider the function:

$$z = \sqrt{x^2 + y^2}$$

with the measured values:

$$x = 4\,\text{cm}, \quad y = 3\,\text{cm}$$

Both x and y are measured with a precision of $\Delta x = \Delta y = 0{,}1\,\text{cm}$. The side length calculated is $z = 5\,\text{cm}$. We now determine the maximum possible absolute error in z.

Step 1: Partial derivatives of z:

$$\frac{\partial z}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$

Step 2: Evaluate at the given point:

$$\frac{\partial z}{\partial x}(4,3) = \frac{4}{5}, \quad \frac{\partial z}{\partial y}(4,3) = \frac{3}{5}$$

Step 3: Use the absolute error formula:

$$\Delta z \leq \left|\frac{\partial z}{\partial x}\right| \cdot \Delta x + \left|\frac{\partial z}{\partial y}\right| \cdot \Delta y$$

$$\Delta z \leq \frac{4}{5} \cdot 0{,}1 + \frac{3}{5} \cdot 0{,}1 = \frac{7}{5} \cdot 0{,}1 = 0{,}14\,\text{cm}$$

Conclusion:

The result for z is accurate to within $0,14$cm.

Relative Error Example

The relative error of a function $z = f(x_1, \ldots, x_n)$ is defined by:

$$\varepsilon = \frac{\Delta z_{\max}}{z}$$

Assuming z depends on variables with known relative measurement errors, and the function is differentiable, we estimate:

$$\varepsilon = \frac{\left|\frac{\partial f}{\partial x_1}\right| \cdot \Delta x_1 + \cdots + \left|\frac{\partial f}{\partial x_n}\right| \cdot \Delta x_n}{f(x_1, \ldots, x_n)}$$

or more directly using relative errors:

$$\varepsilon \approx \left|\frac{\Delta x_1}{x_1}\right| + \cdots + \left|\frac{\Delta x_n}{x_n}\right|$$

Example: Volume of a Cuboid

Let:

$$z = f(a, b, c) = a \cdot b \cdot c$$

Assume the side lengths $a = b = c$ are measured with a uniform relative error $\varepsilon_{\text{in}} = \frac{\Delta a}{a} = \frac{\Delta b}{b} = \frac{\Delta c}{c} = \varepsilon$

Then the partial derivatives are:

$$\frac{\partial z}{\partial a} = bc, \quad \frac{\partial z}{\partial b} = ac, \quad \frac{\partial z}{\partial c} = ab$$

Applying the relative error formula:

$$\frac{\Delta z_{\max}}{z} \leq \frac{bc \cdot \Delta a + ac \cdot \Delta b + ab \cdot \Delta c}{abc} = \frac{\Delta a}{a} + \frac{\Delta b}{b} + \frac{\Delta c}{c} = 3\varepsilon$$

Target accuracy:

If the final result z must be accurate within $\varepsilon_z = 0{,}01$ (i.e., 1), then we must satisfy:

$$3\varepsilon \leq 0{,}01 \Rightarrow \varepsilon \leq \frac{0{,}01}{3} = 0{,}0033 \quad (\text{or } 0{,}33\%)$$

Conclusion:

Each input measurement must be made with a maximum relative error of $0{,}33\%$ to ensure the output $z = abc$ is accurate to within 1 percent.

58.9   Implicit Differentiation

If a function $F(x, y) = 0$ defines y implicitly as a function of x, and F is differentiable, then:

$$\frac{dz}{dx} = -\frac{F_x}{F_z}$$

This generalizes to higher dimensions using the total differential and the implicit function theorem.

Example:

Given is $x^2 + y^2 + z^2 = 1$ differentiate with respect to x. See z or the target function as function in terms of the other variables. In this case $z(x, y)$.

$$\frac{d}{dx} x^2 + y^2 + z^2 - 1 = 0$$

$$2x + 2z\frac{dz}{dx} = 0$$

$$\frac{dz}{dx} = -\frac{x}{z}$$

Example:

Given the implicitly defined function:

$$yx^2 + \frac{x}{y} + x + y - 4 = 0$$

We are asked to compute the derivative $y'(1)$, that is, the derivative of y with respect to y at the point $(1, 1)$.

Step 1: Differentiate implicitly

Differentiate both sides of the equation with respect to x, treating $y = y(x)$ as a function of x:

$$\frac{d}{dx}(yx^2) + \frac{d}{dx}\left(\frac{x}{y}\right) + \frac{d}{dx}(x) + \frac{d}{dx}(y) - \frac{d}{dx}(4) = 0$$

$$y'x^2 + y \cdot 2x + \frac{y - xy'}{y^2} + 1 + y' = 0$$

Collect terms involving y'. Group all y' terms on one side:

$$y'x^2 - \frac{xy'}{y^2} + y' + 2xy + \frac{1}{y} + 1 = 0$$

Factor y':

$$y' \left(x^2 - \frac{x}{y^2} + 1 \right) = - \left(2xy + \frac{1}{y} + 1 \right)$$

Evaluate at $x = 1, y = 1$

Substitute $x = 1$, $y = 1$:

$$y'(1 - 1 + 1) = -(2 + 1 + 1)$$
$$y'(1) = -4$$
$$y' = -4$$

Final Answer

$$y'(1) = -4$$

58.10   Schwarz's Theorem (Clairaut's Theorem)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable. Then for any $i \neq j$, the mixed partial derivatives satisfy:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

That is, the order of partial differentiation does not matter if all second derivatives are continuous.

58.11   The Jacobian

Let $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$, where

$$\vec{f}(\vec{x}) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix}$$

The *Jacobian matrix* of $\vec{f}$ is the $m \times n$ matrix:

$$J_{\vec{f}}(\vec{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

In the case $n = m$, the *Jacobian determinant* $\det J_{\vec{f}}(\vec{x})$ describes local invertibility and orientation.

58.12   Taylor's Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be k-times continuously differentiable at $\vec{x}_0$. Then the Taylor polynomial of degree 2 around $\vec{x}_0$ is:

$$f(\vec{x}) \approx f(\vec{x}_0) + \langle \nabla f(\vec{x}_0), \vec{x} - \vec{x}_0 \rangle + \frac{1}{2}(\vec{x} - \vec{x}_0)^T H_f(\vec{x}_0)(\vec{x} - \vec{x}_0) + \cdots,$$

where $H_f(\vec{x}_0)$ is the *Hessian matrix* of second partial derivatives.

Example:

Let $f(x, y) = e^y \sin(x + 2y)$. We expand around the point $(a, b) = (0, 0)$.

First, compute the necessary derivatives at $(0,0)$:

$$f(0,0) = \sin(0) = 0$$
$$f_x = e^y \cos(x+2y) \quad \Rightarrow \quad f_x(0,0) = 1$$
$$f_y = 2ye^y \cos(x+2y) + e^y \sin(x+2y) \quad \Rightarrow \quad f_y(0,0) = 0 + 2 = 2$$
$$f_{xx} = -e^y \sin(x+2y) \quad \Rightarrow \quad f_{xx}(0,0) = 0$$
$$f_{xy} = f_{yx} = 2e^y \sin(x+2y) + e^y \cos(x+2y) \quad \Rightarrow \quad f_{xy}(0,0) = 3$$
$$f_{yy} = -3e^y \sin(x+2y) + 4e^y \cos(x+2y) \quad \Rightarrow \quad f_{yy}(0,0) = 0 + 4 = 4$$

Now, the second order Taylor polynomial at $(0,0)$ is:

$$f(x,y) \approx 0 + 1 \cdot x + 2 \cdot y$$
$$+ \frac{1}{2} \cdot 0 \cdot x^2 + 3 \cdot xy + \frac{1}{2} \cdot 4 \cdot y^2$$
$$= x + 2y + 3xy + 2y^2$$

58.13 Relative Extrema

To determine local extrema for $f : \mathbb{R}^n \to \mathbb{R}$, follow this procedure:

Step 1: Find critical points by solving:

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

Step 2: Compute the Hessian matrix:

$$H_f(\vec{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

Step 3: Analyze the Hessian at critical points:

- If H is positive definite $f_{xx}(x_0, y_0) > 0, \det(H) > 0 \implies$ local minimum
- If H is negative definite $f_{xx}(x_0, y_0) < 0, \det(H) > 0 \implies$ local maximum
- If H has mixed signs (indefinite) $\det(H) < 0 \implies$ saddle point
- If $\det(H) = 0 \implies$ Next derivative decides

The determinant is needed as part of the test because, while the second derive tells us about the concavity, the determinant help us determine if all the concavities given by the axis second derivatives with respect to $x_1, x_2, \ldots, x_n$ and the mix partials are up, down or neither. If the signs are mixed then it is a saddle.

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

$$\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \Rightarrow \text{Critical point at } (0,0)$$

$$H_f = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow \text{positive definite} \Rightarrow (0,0) \text{ is a local minimum}$$

58.14 Finding Candidate Points for Zeros While Seeking Relative Extrema

To find local minima or maxima of a function $f : \mathbb{R}^n \to \mathbb{R}$, we first identify *critical points*, which are the candidates for relative extrema. These are the points where the gradient of f vanishes or does not exist.

Step 1: Compute the Gradient

The gradient of f, denoted ∇f, collects all the first partial derivatives:

$$\nabla f(x_1, \ldots, x_n) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Step 2: Solve the System $\nabla f = 0$

Find all points $\vec{x}_0 \in \mathbb{R}^n$ such that:

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

This is a nonlinear system of n equations in n variables. The solutions $\vec{x}_0$ are the *candidate points for extrema*, and also referred to as the **stationary points** or *critical points* of the function.

You use the following techniques to find the points:

- Factorization
- Addition of rows
- Substitution

There are more but which one to use depends on the kind problem.

Step 3: Analyze the Nature of the Critical Points

Once all critical points $\vec{x}_0$ satisfying $\nabla f(\vec{x}_0) = 0$ are found, analyze each using the *second derivative test* involving the Hessian matrix H_f:

$$H_f(\vec{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

Evaluate H_f at each critical point and use the eigenvalues or definiteness to classify:

- If $f_{xx}(x_0, y_0) > 0$ and $D > 0$, then f has a *local minimum* at (x_0, y_0).
- If $f_{xx}(x_0, y_0) < 0$ and $D > 0$, then f has a *local maximum* at (x_0, y_0).
- If $D < 0$, then (x_0, y_0) is a *saddle point*.
- If $D = 0$, the *second derivative test is inconclusive*; higher-order derivatives must be considered.

58.14.1 Note: Non-differentiable Points

If the function f is not differentiable at some point $\vec{x}_0$, but defined there, this point may also be a candidate for extremum and should be examined separately using limit-based analysis or directional behavior.

58.15 Lagrange Multipliers

Let $f : \mathbb{R}^n \to \mathbb{R}$ be the function to optimize, subject to the constraint $g(\vec{x}) = 0$, where $g : \mathbb{R}^n \to \mathbb{R}$.

If we look at the constraint intersection with the original function we are going to notice that along the intersection there are maxima and minima with their respective tangent lines and gradients that are normal to them. We also are going to see that both the gradient of f and the gradient of g are just scalar version of each other so:

This leads to the *Lagrange system*:

$$\begin{cases} \nabla f(\vec{x}) = \lambda \nabla g(\vec{x}) \\ g(\vec{x}) = 0 \end{cases}$$

For more conditions we can use $g_1, \ldots, g_n$ and $\lambda_1, \ldots, \lambda_n$

$$\begin{cases} \nabla f(\vec{x}) = \lambda_1 \nabla g_1(\vec{x}) + \cdots + \lambda_n \nabla g_n(\vec{c}) \\ g_1(\vec{x}) = 0 \\ \vdots \\ g_n(\vec{x}) = 0 \end{cases}$$

We can also compact these equations in the *Lagrangian*

$$\mathcal{L}(x, y, \ldots, \lambda_1, \ldots, \lambda_n) = f(x, y, \ldots) + \lambda_1 g_1(x, y, \ldots) + \cdots + \lambda_n g_n(x, y, \ldots)$$

58.15.1 Determinant method

Sometimes it is a good way to solve a problem of this topic is to:

1. Build the *Lagrangian*

2. $\nabla L = 0$ and maybe compute the determinant to generate another equation to solve

3. Get rid of λ via factorization, substitution, addition of rows, etc.

4. Solve the system without λ

58.15.2 The Bordered Hessian

The Bordered Hessian matrix is:

$$H_B = \begin{bmatrix} 0 & (\nabla g(x_1, \ldots, x_n))^T \\ \nabla g(x_1, \ldots, x_n) & \nabla^2 \mathcal{L}(x, \ldots, x_n, \lambda_1, \ldots, \lambda_n) \end{bmatrix}$$

Where:

- $\nabla g(x_1, \ldots, x_n)$ is the $m \times n$ Jacobian matrix of the constraints.
- $\nabla^2 \mathcal{L}(x, \ldots, x_n, \lambda_1, \ldots, \lambda_n)$ is the $n \times n$ Hessian of the Lagrangian with respect to x.

Example I:

Find the point on the circle $x^2 + y^2 = 4$ that is the closest to the point $(3, 4)$.

The function $f(x, y)$ is this case the euclidean norm:

$$f(x, y) = \sqrt{(x - 3)^2 + (y - 4)^2}$$

subject to the constraint:

$$g(x, y) = x^2 + y^2 - 4 = 0$$

The gradient of the constraint is:

$$\nabla g = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

Since $\nabla g \neq 0$, the rank condition is fulfilled, and no further critical points need to be analyzed.

Step 1: Build the Lagrangian

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

In this case we can use a property of the root the function to our advantage. Notice that the minimum of the root function is the minimum of the term under root therefore, we can simplify some work by writing $f(x, y)$ without the root.

$$\mathcal{L}(x, y, \lambda) = (x - 3)^2 + (y - 4)^2 + \lambda(x^2 + y^2 - 4)$$

Step 2: Compute the necessary conditions

$$\mathbf{I} \quad \frac{\partial \mathcal{L}}{\partial x} = 2(x - 3) + 2\lambda x = 0$$

$$\mathbf{II} \quad \frac{\partial \mathcal{L}}{\partial y} = 2(y - 4) + 2\lambda y = 0$$

$$\mathbf{III} \quad \frac{\partial \mathcal{L}}{\partial \lambda} = x^2 + y^2 - 4 = 0$$

Now, this is the difficult part. The best approach in this case is to use substitution to with a top-bottom approach:

$$x = \frac{3}{1 + \lambda}$$

$$y = \frac{4}{1 + \lambda}$$

Solving for λ is also an option but in these kinds of systems of equations where both the partial with respect to x and y are very similar it is a good idea to solve for them because they tend to have a similar form when solving for them.

$$\left(\frac{3}{1 + \lambda}\right)^2 + \left(\frac{4}{1 + \lambda}\right)^2 - 4 = 0$$

$$\frac{9}{(1 + \lambda)^2} + \frac{16}{(1 + \lambda)^2} - 4 = 0$$

$$\frac{25}{(1 + \lambda)^2} = 4$$

$$25 = 4(1 + \lambda)^2$$

$$\sqrt{\frac{25}{4}} = 1 + \lambda$$

$$\lambda = -1 \pm \frac{5}{2}$$

$$\lambda_1 = \frac{-7}{2} \quad \lambda_2 = \frac{3}{2}$$

Now, let us substitute backwards for λ_1:

$$x = \frac{3}{1 - \frac{7}{2}} = \frac{6}{-5}$$

278

$$y = \frac{4}{1 - \frac{7}{2}} = \frac{8}{-5}$$

Now, for λ_2:

$$x = \frac{3}{1 + \frac{3}{2}} = \frac{6}{5}$$

$$y = \frac{4}{1 + \frac{3}{2}} = \frac{8}{5}$$

Step 3: Hessian Test with Constraint

We construct the Hessian:

$$H = \begin{bmatrix} 2 + 2\lambda & 0 & 2x \\ 0 & 2 + 2\lambda & 2 \\ 2x & 2 & 0 \end{bmatrix}$$

For λ_1 we get $\det(H) = \frac{244}{5} > 0$, a minimum and for λ_2 $\det(H) = \frac{-244}{5}$ a maximum.

Example II:

Optimize the distance to the origin $f(x, y, z)$ with respect to the constraints $g_1(x, y, z) = z^2 - x^2 - y^2 = 0$ and $g_2(x, y, z) = x - 2z - 3 = 0$

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$$
$$g_1 = 0$$
$$g_2 = 0$$

We know that $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. Like with other minimization problems we can just find the minimum of the part under root thus, $f(x, y, z) = x^2 + y^2 + z^2$. Now we can put the build our system of equations

$$\textbf{I } 2x = \lambda(-2x) + \mu$$
$$\textbf{II } 2y = \lambda(-2y)$$
$$\textbf{III } 2z = \lambda(2z) + \mu(-2)$$
$$\textbf{IV } z^2 = x^2 + y^2$$
$$\textbf{V } x = 2z + 3$$

For the equation **II** we can already see that the equation is equal to zero only if $\lambda = -1$ or $y = 0$. Thus, we have got our first numbers to work with.

Case $\lambda = -1$:

$$\textbf{I } 2x = -1(-2x) + \mu \implies \mu = 0$$
$$\textbf{III } 2z = -1(2z) + 0(-2) \implies z = 0$$
$$\textbf{IV } 0^2 = x^2 + y^2 \implies x = y = 0$$
$$\textbf{V } 0 = 0 + 3 \implies 0 = 3$$

With the last contradiction we know that the whole case $\lambda = -1$ is not valid.

Case $y = 0$:

$$\textbf{IV } z^2 = x^2 + y^2 \implies z^2 = x^2 \implies z = \pm x$$
$$\textbf{V } x = 2x + 3 \implies x = -3(z = x)$$
$$\textbf{V } x = -2x + 3 \implies x = 1(z = -x)$$

Now with this new information we can get two different points $P_1(1, 0, -1)$ and $P_2(-3, 0, -3)$. We do not need to write the Hessian in this case because with geometric intuition the problem can be solved.

$$f(P_1) = 1^2 + 0^2 + (-1)^2 = 2$$

So the distance is $\sqrt{2}$.

$$f(P_2) = -3^2 + 0^2 - 3^2 = 18$$

Thus, the distance is $3\sqrt{2}$.

Our final Answer is that P_1 is the closest point to the intersection of the two planes with the function f.

58.16 The Tangent Vector

Let $\vec{X}(t) \in \mathbb{R}^n$ be a differentiable, parameterized curve. The *tangent vector* at the point $\vec{X}(t_0)$ is given by:

$$\vec{X}'(t_0) = \begin{bmatrix} x_1'(t_0) \\ x_2'(t_0) \\ \vdots \\ x_n'(t_0) \end{bmatrix}$$

This vector points in the direction in which the curve is moving at t_0, and its magnitude corresponds to the instantaneous speed. The unit tangent vector is:

$$\vec{X}(t) = \frac{\vec{X}'(t)}{\|\vec{X}'(t)\|}$$

The speed of the vector in $\mathbb{R}^2$ is given by:

$$\vec{X}(t) = \begin{bmatrix} t \\ f(t) \end{bmatrix}$$

and the tangent is:

$$T(t) = \begin{bmatrix} t \\ f(t) \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ f'(t) \end{bmatrix}$$

58.16.1 Derivation of the Tangent Vector Formula

Given a position vector function $\vec{r}(t)$, the tangent vector to the curve described by $\vec{r}(t)$ is obtained by taking the derivative of $\vec{r}(t)$ with respect to time. This derivative is defined as the following limit:

$$\frac{d\vec{r}}{dt} = \lim_{\Delta t \to 0} \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t}$$

Assuming that $\vec{r}(t)$ is expressed in terms of the canonical basis vectors $\hat{\imath}, \hat{\jmath}, \hat{k}$, we can write:

$$\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$$

Then, the increment becomes:

$$\vec{r}(t + \Delta t) = x(t + \Delta t)\hat{\imath} + y(t + \Delta t)\hat{\jmath} + z(t + \Delta t)\hat{k}$$

So the difference in the numerator of the derivative is:

$$
\begin{aligned}
\vec{r}(t + \Delta t) - \vec{r}(t) = {}& [x(t + \Delta t) - x(t)]\,\hat{\imath} \\
& + [y(t + \Delta t) - y(t)]\,\hat{\jmath} \\
& + [z(t + \Delta t) - z(t)]\,\hat{k}
\end{aligned}
$$

Dividing by Δt and taking the limit:

$$
\begin{aligned}
\frac{d\vec{r}}{dt} &= \lim_{\Delta t \to 0} \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t} \\
&= \left(\lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \right) \hat{\imath} \\
&\quad + \left(\lim_{\Delta t \to 0} \frac{y(t + \Delta t) - y(t)}{\Delta t} \right) \hat{\jmath} \\
&\quad + \left(\lim_{\Delta t \to 0} \frac{z(t + \Delta t) - z(t)}{\Delta t} \right) \hat{k} \\
&= \frac{dx}{dt}\hat{\imath} + \frac{dy}{dt}\hat{\jmath} + \frac{dz}{dt}\hat{k}
\end{aligned}
$$

This vector $\frac{d\vec{r}}{dt}$ points in the direction of motion and is tangent to the curve at each point t.

59 Multidimensional Integrals

59.1 Arclength of a Curve

We will use the famous technique of approximating the length/area/volume of something by dividing it in to equally smaller section that are straight lines in this case. In other cases they will be prisms, squares, etc. Now if we look at the difference between two points in our 3D space, for this Example we notice that there are three differences $\Delta i \; \Delta j \; \Delta k$.j Now let us take the length of the vector $\sqrt{(\Delta i)^2 + (\Delta j)^2 + (\Delta k)^2}$

The total Arclength is going to be the sum of all of these sectors

$$L = \sum_{i=1}^{n} \sqrt{(\Delta i)^2 + (\Delta j)^2 + (\Delta k)^2}$$

if we use a clever one $\left(\frac{\Delta t}{\Delta t}\right)^2$ we get

$$L = \sum_{i=1}^{n} \sqrt{\left(\frac{\Delta i}{\Delta t}\right)^2 + \left(\frac{\Delta j}{\Delta t}\right)^2 + \left(\frac{\Delta k}{\Delta t}\right)^2} \, \Delta t$$

Now if we take the limit we get

$$L = \lim_{n \to \infty} \sum_{i=1}^{n} \sqrt{\left(\frac{\Delta i}{\Delta t}\right)^2 + \left(\frac{\Delta j}{\Delta t}\right)^2 + \left(\frac{\Delta k}{\Delta t}\right)^2} \, \Delta t$$

which is

$$L = \int_{a}^{b} \sqrt{\left(\frac{\Delta i}{\Delta t}\right)^2 + \left(\frac{\Delta j}{\Delta t}\right)^2 + \left(\frac{\Delta k}{\Delta t}\right)^2} \, dt$$

$$L = \int_{a}^{b} \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2} \, dt$$

For differentiable function on the interval $t \in [a; b]$

59.2 Line Integrals

59.2.1 Over Vector Field

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a *vector field*, and let γ be a smooth curve parametrized by $\vec{r}(t)$, $t \in [a, b]$. The *line integral* of F along γ is:

$$\int_{\gamma} \left\langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle dt = \int_{a}^{b} \left\langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}}{dt} \right\rangle dt = \vec{F}(\vec{r}(b)) - \vec{F}(\vec{r}(a))$$

In the context of physics this integral measures the work done by the field $\vec{F}$ along the path γ, such as force over distance.

In compact form

$$\int_{\gamma} \langle \vec{F}, d\vec{r} \rangle$$

Example: Line Integral over a Vector Field

Let $f(x, y) = 2x + 2y$ be our vector field.

Let the curve c be the quarter-circle of radius 1 from $(1,0)$ to $(0,1)$, parametrized by:

$$\vec{r}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \quad t \in \left[0, \frac{\pi}{2}\right]$$

We compute the line integral of $\vec{F}$ along c:

$$\int_C \langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \rangle dt = \int_0^{\frac{\pi}{2}} \langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \rangle dt$$

First, compute $\vec{F}(\vec{r}(t))$:

$$\vec{F}(\vec{r}(t)) = \begin{bmatrix} 2\cos(t) \\ 2\sin(t) \end{bmatrix}$$

Then the tangent vector of $\vec{r}(t)$:

$$\frac{d\vec{r}(t)}{dt} = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}$$

Now compute the inner product:

$$\langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \rangle = -2\cos(t)\sin(t) + \cos(t)2\sin(t) = 0$$

Therefore, the line integral is:

$$\int_C \langle \vec{F}, \frac{d\vec{r}}{dt} \rangle dt$$
$$= \int_0^{\frac{\pi}{2}} 0 dt$$
$$= 0$$

The line integral of the gradient vector field $f(x, y) = 2x + 2y$ over this curve is zero.

59.2.2 Over Scalar field

Let $f(x_1, x_2, \ldots, x_n)$ be a scalar field defined on a smooth curve c, which is parametrized by a vector function.

$$\vec{r}(t) = (x_1(t), x_2(t), \ldots, x_n(t)), \quad t \in [a, b]$$

Then, the line integral of f over c is given by:

$$\int_C f(x_1, x_2, \ldots, x_n) \, ds = \int_a^b f(\vec{r}(t)) \left\| \frac{\vec{r}(t)}{dt} \right\| dt$$
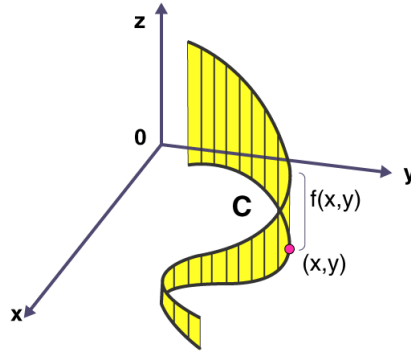
This formula comes from approximating the area of the squares under the curve with height $f(x_k, y_k)$ and the arclength Δs_k, we will simplify to three dimensions for the sake of simplicity, but like seen before this works for n dimensions. $A_k = f(x_k, y_k)\Delta s_k$, where $\Delta s_k = \sqrt{(\Delta s_k)^2 + (\Delta y_k)^2}$. Now by using or formula for the arclength we already know, we get $\|\frac{\vec{r}(t)}{dt}\|dt = ds$.

Example: Scalar Field

Compute the line integral of $f(x, y) = x^2 + y^2$ over the straight line segment going from $(1, 1)$ to $(3, 3)$

This time or curve is $y = x$ to the parametrized curve is $\begin{bmatrix} x(t) = t \\ y(t) = t \end{bmatrix}$ this left us with:

$$\int_1^3 2t^2 \sqrt{2}dt = \frac{52\sqrt{2}}{3}$$

59.3   Line Integrals with respect to a specific variable

We are going to take another look at line integrals but from the perspective of just one variable and how this affects the formulas we know. For the next explanation we will use only functions of two variables to keep the examples simple.

59.3.1   Over Scalar Field with respect to x or y

Imagining how this works in not specially difficult. Imagine taking a line integral where we do not move along the y-axis and only care about the plane xz. Here our line integral only changes in the fact that in the norm of our curve the component that was responsible for y will be set 0 and the ds changes to the variable we are integrating with respect to. In this case dx.

$$\int_c f(x, y)dx = \int_c f(g(t), h(t))\sqrt{g'(t)^2}dt$$

59.3.2   Over Vector Field with respect to x or y

In the case of a field the story is different in the fact that now our output vectors change depending on the variable we are choosing to left out.

Now, given a field $\vec{F}(x, y) = M(x, y)\imath + N(x, y)\jmath$ and the curve $\vec{r}(t) = g(t)\imath + h(t)\jmath$. The line integral of $\vec{F}$ along C is

$$\int_\gamma \left\langle \nabla \vec{F}, \frac{\vec{r(t)}}{dt} \right\rangle dt$$

$$\int_a^b M(x, y)g'(t)dt + \int_a^b N(x, y)h'(t)dt$$

Here $g'(t)dt = dx$ and $h'(t)dt = dy$

$$\int_a^b M dx + \int_a^b N dy$$

These two integrals give us both the integral with respect to x and y.

Example:

Curve C is a portion of the parabola $y = x^2$ that goes through the points (1,1) to (2,4)

Compute: $\int_a^b \frac{x}{y} dy$

Step 1: Parameterize C

The trick for this cases where the curve is the graph of a function is to set $x = t$ and them solve for y in terms of x.

In this case we get

$$\vec{r}(t) = t\imath + t^2\jmath \text{ with } 1 \le t \le 2$$

Step 2: Substitute into the integral

$$h'(t)dt = 2t dt$$

$$\int_a^b \frac{t}{t^2} 2t dt = \int_1^2 \frac{1}{t} 2t dt = 2$$

59.4   Flow and Circulation

For velocity $\vec{F} = M(x,y)\hat{\imath} + N(x,y)\hat{\jmath}$ and a smooth curve C parameterized by $\vec{r}(t)$

$$Flow = \int_C \langle \vec{F}, \vec{T} \rangle ds$$

If the curve C is closed, the it is called *Circulation*.

$$Circulation = \oint_C \langle \vec{F}, \vec{T} \rangle ds$$

Flow is a measurement of the degree of which a curve goes along the vector field.

59.4.1   Circulation Density

The *circulation density* of a vector field $\vec{F}(x,y) = M(x,y)\hat{\imath} + N(x,y)\hat{\jmath}$ is given by

$$CD = \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right)$$

It derives from imagining having a square as our curve where from our starting point (x,y) we move $(x+\Delta x, y), (x+\Delta x, y+\Delta y), (x, y+\Delta y)$ generating four different curves. Therefore our total circulation is given by

$$\oint_C \langle \vec{F}, d\vec{r} \rangle = \sum_{i=1}^{4} \int_{C_i} \langle \vec{F}, \vec{T} \rangle d\vec{r}$$

The dot products are given by

- Bottom: $\langle \vec{F}(x,y)\Delta x \hat{\imath} \rangle = M(x,y)\Delta x$

- Right: $\langle \vec{F}(x+\Delta x, y)\Delta y\hat{j}\rangle = N(x+\Delta x, y)\Delta y$

- Top: $\langle \vec{F}(x, y+\Delta y) - \Delta x\hat{i}\rangle = -M(x, y+\Delta y)\Delta x$

- Left: $\langle \vec{F}(x, y) - \Delta y\hat{j}\rangle = N(x, y)\Delta y$

Then the Top minus Bottom is $-(M(x, y+\Delta y) - M(x, y)\Delta x)$ and notice that inside the parenthesis we have something similar to the partial

$$\frac{\partial M}{\partial y} = \lim_{\Delta y \to 0} \frac{M(x, y+\Delta y) - M(x, y)}{\Delta y}$$

Thus, by multiplying a clever one $\frac{\Delta y}{\Delta y}$ we get $-\frac{\partial M}{\partial y}\Delta y\Delta x$. Similarly, we get for the Right minus left case $\frac{\partial N}{\partial x}\Delta y\Delta x$

Finally by putting all together

$$\oint_C \langle \vec{F}, d\vec{r}\rangle = -\frac{\partial M}{\partial y}\Delta y\Delta x + \frac{\partial N}{\partial x}\Delta y\Delta x = \left(-\frac{\partial M}{\partial y} + \frac{\partial N}{\partial x}\right)\Delta y\Delta x$$

$\Delta y\Delta x$ is the area of the rectangle.

59.5 Green's Theorem Circulation-Curl Form

For C, a smooth, simple, closed curve enclosing R, $\vec{F} = M\hat{i} + N\hat{j}$ having M, N continuous first partials in an open region containing R

$$\oint_C \langle \vec{F}, d\vec{r}\rangle = \int\int_R \left(-\frac{\partial M}{\partial y} + \frac{\partial N}{\partial x}\right) dxdy$$

59.6 Flux

For a continuous field $\vec{F} = M(x,y)\hat{i} + N(x,y)\hat{j}$, a smooth curve with no intersections C and $\vec{n}$ the outward normal

$$Flux = \oint_C \langle \vec{F}, \vec{n}\rangle ds$$

The derivation of the formula for actually computing the *flux* goes as follows: First, we find ourselves is 2 dimensions and we want to find a normal vector to the curve. We can pretend that we are in 3 dimensions and define a vector

$$\vec{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

This gives us $\vec{n} = \vec{T} \times \vec{k}$. Now either by using the determinant trick or formally computing the cross product we get

$$\vec{T} \times \vec{k} = \vec{n} = \begin{bmatrix} \frac{dy}{ds} \\ \frac{dx}{ds} \\ 0 \end{bmatrix}$$

The final formula is then

$$Flux = \oint_C \langle \vec{F}, \vec{n}\rangle ds = \oint_C \left\langle \vec{F}, \begin{bmatrix} \frac{dy}{ds} \\ \frac{dx}{ds} \\ 0 \end{bmatrix} \right\rangle ds = \int_a^b M dy - \int_a^b N dx$$

One extra note about the derivation is that if the curve goes counter clockwise we use $\vec{n} = \vec{T} \times \vec{k}$ but if it goes clockwise $\vec{n} = \vec{k} \times \vec{T}$ because of the right-hand-rule.

59.7 Flux Density

Just like for the *Circulation Density* we will use the small square to examine the divergence in a small region. Following a similar process with the formula for flux we can get a similar result.

$$FD = \left(-\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} \right)$$

59.8 Green's Theorem Divergence-Flux Form

For C, a smooth, simple, closed curve enclosing R, $\vec{F} = M\hat{i} + N\hat{j}$ having M, N continuous first partials in an open region containing R

$$\oint_C \langle \vec{F}, \vec{n} \rangle ds = \int\int_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} \right) dxdy$$

59.9 Conservative Vectors Fields

A field $\vec{F}$ is *conservative* on an open domain if

$$\int_C \langle \vec{F}, d\vec{r} \rangle$$

is the same for all paths C between points A and B in the domain. In other words it only depends on the endpoints.

59.10 The Potential Function

A vector field $\vec{F} : \mathbb{R}^n \to \mathbb{R}^n$ is called *conservative* if there exists a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ such that:

$$\vec{F} = \nabla f$$

The function f is then called a *potential function* of $\vec{F}$ and also

$$\int_C \langle \vec{F}, d\vec{r} \rangle = f(B) - f(A)$$

Proof:

$$\int_C \langle \vec{F}, d\vec{r} \rangle = \int_a^b \left\langle \vec{F}(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle dt$$

By assumption

$$= \int_a^b \left\langle \nabla f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle dt$$

This this like the chain rule. Therefore,

$$\left\langle \nabla f(\vec{r}(t)), \frac{d\vec{r}(t)}{dt} \right\rangle = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t} = \frac{d}{dt} f(\vec{r}(t))$$

$$= \int_a^b \frac{d}{dt} \phi(\vec{r}(t)) dt$$

By the fundamental theorem of calculus

$$= f(B) - f(A)$$

<div align="right">𝔔𝔈𝔇</div>

59.10.1   Check for conservative vector fields

Using the definition $\vec{F} = \nabla f$ and by writing

$$\vec{F} = f_1(x_1, x_2, \dots)\hat{i}_1 + f_2(x_1, x_2, \dots)\hat{i}_2 + \cdots + f_n(x_1, x_2, \dots)\hat{i}_n$$
$$= \frac{\partial f}{\partial x_1}\hat{i}_1 + \frac{\partial f}{\partial x_2}\hat{i}_2 + \cdots \frac{\partial f}{\partial x_n}\hat{i}_n$$

We get the following and reducing to 3 variables and dimensions for the sake of simplicity

$$\frac{\partial f_1}{\partial y} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial f_2}{\partial x}$$

We can also keep making pairs like this, but checking one is enough.

59.10.2   Finding the Potential Function of a Vector Field

To find the potential function $f(x, y)$ of a vector field $\vec{F}(x, y) = (P(x, y), Q(x, y))$, we must check if the field is conservative and then integrate accordingly. A vector field is conservative if there exists a scalar potential function f such that:

$$\vec{F} = \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Given:

$$\vec{F}(x, y) = \begin{bmatrix} 2xy + e^x \\ x^2 + \frac{1}{2\sqrt{y}} \end{bmatrix}$$

Let $P(x, y) = 2xy + e^x$, and $Q(x, y) = x^2 + \frac{1}{2\sqrt{y}}$.

First, check if the field is conservative by verifying:

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y}(2xy + e^x) = 2x$$

$$\frac{\partial Q}{\partial x} = \frac{\partial}{\partial x}\left(x^2 + \frac{1}{2\sqrt{y}}\right) = 2x$$

Since $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, the field is conservative.

Step 1: Integrate $P(x, y)$ with respect to x:

$$f(x, y) = \int (2xy + e^x)\, dx = x^2 y + e^x + h(y)$$

Here, $h(y)$ is a function of y only.

Step 2: Differentiate $f(x, y)$ with respect to y:

$$\frac{\partial f}{\partial y} = x^2 + h'(y)$$

Set this equal to $Q(x, y) = x^2 + \frac{1}{2\sqrt{y}}$, so:

$$x^2 + h'(y) = x^2 + \frac{1}{2\sqrt{y}} \Rightarrow h'(y) = \frac{1}{2\sqrt{y}}$$

Step 3: Integrate to find $h(y)$:

$$h(y) = \int \frac{1}{2\sqrt{y}}\, dy = \sqrt{y} + C$$

Therefore, the potential function is:

$$f(x, y) = x^2 y + e^x + \sqrt{y} + C$$

59.11 Double Integral over a Rectangular Region

Let $A = [x_0, x_1] \times [y_0, y_1]$ be a rectangular region in $\mathbb{R}^2$. The double integral of a function $f(x, y)$ over A is:

$$\iint_A f(x, y)\, dA = \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y)\, dy\, dx$$

For rectangular domains, the order of integration does not affect the result:

$$\int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y)\, dy\, dx = \int_{y_0}^{y_1} \int_{x_0}^{x_1} f(x, y)\, dx\, dy$$

59.12 Steps for finding the Region of Integration

When Integrals are not defined over regions defined from constants then finding the bounds of integration can become a tricky task depending on the region we want to integrate.

The general process for setting up iterated integrals over a non-rectangular region is as follows:

1. **Sketch the region if possible:** Begin by sketching the region of integration. A visual representation helps you understand the boundaries, the shape of the region, and the relationships between variables.

2. **Choose the order of integration:** Decide the order in which you'll integrate the variables. Remember that integration proceeds from the inside out, so the inner integral corresponds to the variable you integrate first.

3. **Determine the bounds for each variable:** For the chosen order of integration, identify the limits of integration for each variable. The bounds for the inner variable are often functions of the outer variable. For example, if the region is described by the inequality $x + y < 1$ and you choose to integrate with respect to y first, you must express y's bounds in terms of x. This means solving the inequality for y:

$$y < 1 - x \quad \Rightarrow \quad y \in [c(x), 1 - x]$$

where $c(x)$ is the lower bound for y, possibly another function of x. Then the full double integral would look like:

$$\int_a^b \int_{c(x)}^{1-x} f(x, y)\, dy\, dx$$

Repeat this process for all variables involved. If you've sketched the region, you can often find these bounds by "slicing" the region—holding one variable constant and observing how the others vary within the region.

59.13 Integration over Curvilinear Domains

If the region $A \subset \mathbb{R}^2$ is not rectangular, it is divided into subregions ΔA_k, and the integral is defined as the limit:

$$\iint_A f(x,y)\, dA = \lim_{n \to \infty} \sum_{k=1}^n f(x_k, y_k)\, \Delta A_k$$

In practice, for regions bounded by curves, we integrate in the form:

$$\int_a^b \int_{u(x)}^{o(x)} f(x,y)\, dy\, dx$$

Here, $y \in [u(x), o(x)]$ describes the vertical bounds as functions of x.

Example:

Compute the volume under $f(x,y) = 1 + x + y^2$ bounded by x-axis, $x = 1$ and y-axis $y = \sqrt{x}$
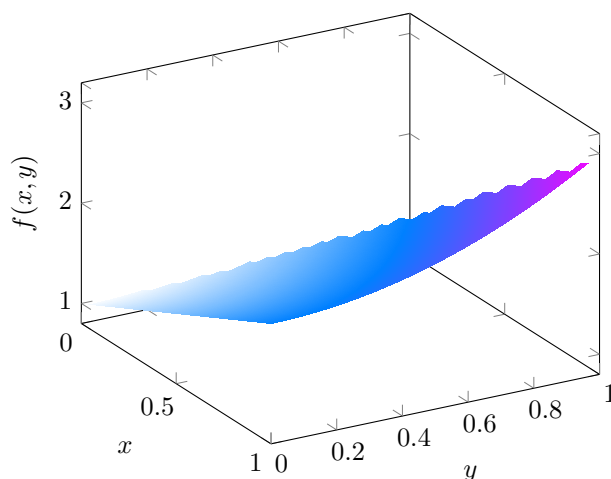
First we are going to imagine slicing the area under this sector vertically along the x-axis from 0 up to 1 which is the volume of an Area under a curve that is represented by going along the y-axis.

$$\int_0^1 \int_0^{\sqrt{x}} 1 + x + y^2\, dy\, dx$$

$$\int_0^1 \left. y + xy + \frac{y^3}{3} \right|_0^{\sqrt{x}} dx$$

$$\int_0^1 \left. \sqrt{x} + x\sqrt{x} + \frac{\sqrt{x}^3}{3} \right|_0^{\sqrt{x}} dx$$

$$= \left. \frac{2}{3} x^{\frac{3}{2}} + \frac{4}{3} \frac{2}{5} x^{\frac{5}{2}} \right|_0^1 = \frac{6}{5}$$

If we choose the other way then we have an area with a fix y in the inside that goes from y^2 to 1 with respect to x.

$$\int_0^1 \int_{y^2}^1 1 + x + y^2\, dx\, dy$$

290

59.14   Changing the order of integration to solve tricky integrals

Given $\int_0^8 \int_{\sqrt[3]{x}}^2 \frac{1}{y^4+1} dy dx$ which is very hard to integrate, let us change the order to make it easier.

$$\int_0^2 \int_0^{y^3} \frac{1}{y^4+1} dx dy$$

The steps are the same as in the previous Example.

1. Identify the area you want to integrate, what are the bounds respective to x and y

2. Write the same expression without bounds and change the order of the differentials

3. Imagine slicing the area underneath vertically for x and horizontally for y with their respective bounds corresponding to the perspective.

4. Write the new integral with the opposite of order of integration

Visualization of the area underneath the surface
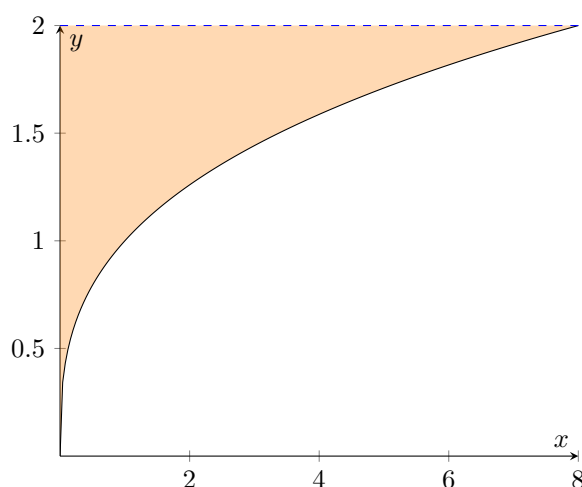
Here the original bounds are: x goes from 0 to 8 and y goes from $\sqrt[3]{x}$ to 2. As you can see in the diagram by slicing vertically the value of x is fixed, and we move towards the upper-bound of y which is 2. Therefore, the smallest x can get is 0 and the biggest is 2. While for y by slicing vertically we get that x goes from 0 to $\sqrt[3]{x}$, but we do not want to leave x, so we solve for y and get y^3. Now we can put all of this together and integrate.

In general changing the order of integration is only about using the logic to set the bounds of integration, but starting from a different perspective.

59.15   The Jacobian

The Jacobian matrix is a fundamental concept in multi-variable calculus. The Jacobian is the best approximation of a linear transformation for a point.

For the intuition of why, we use partial derivatives, just think about how we use them as this tiny changes in the component of original basis vectors.

Given a vector-valued function

$$f : \mathbb{R}^n \to \mathbb{R}^m, \quad f(x_1, \ldots, x_n) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ f_2(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix},$$

the *Jacobian matrix* of f is the $m \times n$ matrix of all first-order partial derivatives:

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

59.15.1 Local linearity

While performing non-linear transformations on a space. We can see that in a tiny neighborhood around a point, the transformation appears to be linear. And the matrix that encodes that transformation is the *Jacobian*.

For intuition, just imagine that in the tiny neighborhood the transformation is composed of small steps in the $x_1, \ldots, x_n$ directions which give of the partial derivatives that represent this small steps.

59.15.2 Determinant

When $m = n$, i.e., the function maps from $\mathbb{R}^n$ to $\mathbb{R}^n$, the Jacobian matrix is square. In this case, the *Jacobian determinant* is defined as:

$$\det(J_f(x)),$$

and provides important information about the local behavior of the function. For instance:

- If $\det(J_f(x)) \neq 0$, the function is locally invertible at x by the inverse function theorem.
- The sign and magnitude of the determinant gives us information about orientation and volume scaling of the transformation.

When we change our systems of coordinates it is important that our differential make sense in the original system we were. Because of that, we take the determinant with respect to the new coordinates to tell us the scaling factor that we have to use for our result to make sense.

Use Cases:

The Jacobian matrix and its determinant appear in many areas, including:

- **Change of variables in integrals:** The absolute value of the Jacobian determinant is used to adjust the measure when changing coordinates in multiple integrals.

- **Nonlinear optimization:** The Jacobian is crucial for computing gradients and performing techniques like Newton's method in multiple dimensions.

- **Differential equations:** Stability analysis of dynamical systems often involves evaluating the Jacobian at equilibrium points.

59.15.3 List of Common Jacobians

- *Cartesian Coordinates*:
  - Definition: The position of a point is given by (x, y, z) in a 3D space along orthogonal axes.
  - Jacobian:

$$J_{\text{Cartesian}} = 1dx$$

- *Polar Coordinates (2D)*:
  - Definition: The position of a point in a plane is given by (r, θ), where r is the radial distance and θ is the angle from the positive x-axis.
  - Jacobian:

$$J_{\text{Polar}} = rdrd\theta$$

- *Cylindrical Coordinates*:
  - Definition: The position of a point in 3D space is given by (r, θ, z), where (r, θ) are the polar coordinates in the xy-plane and z is the height along the z-axis.
  - Jacobian:

$$J_{\text{Cylindrical}} = r \; dr d\theta dz$$

- *Spherical Coordinates*:
  - Definition: The position of a point in 3D space is given by (ρ, θ, ϕ), where:
    * ρ is the distance from the origin,
    * θ is the azimuthal angle (in the xy-plane from the positive x-axis),
    * ϕ is the polar angle (from the positive z-axis).
  - Jacobian:

$$J_{\text{Spherical}} = \rho^2 \sin(\phi) d\rho d\theta d\phi$$

59.16   Integration in Polar Coordinates

For radially symmetric regions or functions, switching to polar coordinates simplifies the computation.

Transformation:

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

The Jacobian of the transformation gives the area element:

$$dA = r \, dr \, d\varphi$$

Integral:

$$\iint_D f(x, y) \, dA = \int_{\varphi_0}^{\varphi_1} \int_0^{R(\varphi)} f(r \cos \varphi, r \sin \varphi) \, r \, dr \, d\varphi$$

59.16.1   Origin of the Area

When we divide our circle in the sections we get both a difference in then angle and the radius which generate another area. The difference $\Delta A = $ Area of big wedge $-$ Area of small wedge

Radius: $r_k + \frac{\Delta r}{2}$ for the big wedge and $r_k - \frac{\Delta r}{2}$ for the small wedge

Angle: $\frac{\Delta \theta}{2\pi}$ which is the angle of the fraction of the circle because the formula for the area is $r^2 \pi$ and the whole circle would be 2π

Final Formula for the difference in the Area

$$\Delta A = \frac{\Delta \theta}{2\pi} \left(\left[r_k + \frac{\Delta r}{2} \right]^2 - \left[r_k - \frac{\Delta r}{2} \right]^2 \right)$$

$$= r_k \Delta r \Delta \theta$$

Which multiplied by the function value and by taking the limit of these sections we get that the volume:

$$V = \int_{\theta_1}^{\theta_2} \int_{r_1(\theta)}^{r_2(\theta)} f(r,\theta) r \Delta r \Delta \theta$$

Example:

$f(r,\theta) = r$ above cardioid $r = 1 - \sin\theta$

Then

$$V = \int_0^{2\pi} \int_0^{1-\sin\theta} r r \, dr \, d\theta$$

$$= \int_0^{2\pi} \frac{r^3}{3} \Big|_0^{1-\sin\theta} d\theta = \int_0^{2\pi} \frac{(1-\sin\theta)^3}{3} d\theta = \frac{5\pi}{3}$$

59.17 Improper Integrals over Unbounded Regions

If the domain of integration is unbounded (e.g., the entire plane $\mathbb{R}^2$), the double integral is defined via a limit process:

$$\iint_{\mathbb{R}^2} f(x,y)\, dx\, dy := \lim_{M_1, M_2 \to \infty} \lim_{N_1, N_2 \to \infty} \int_{M_1}^{M_2} \int_{N_1}^{N_2} f(x,y)\, dy\, dx$$

Often, using polar coordinates simplifies such cases, reducing multiple limits to one:

$$\iint_{\mathbb{R}^2} f(x,y)\, dx\, dy = \int_0^{2\pi} \int_0^\infty f(r\cos\varphi, r\sin\varphi)\, r\, dr\, d\varphi$$

59.17.1 Gaussian Integral

We are going to find the integral of the *unbounded integral* e^{-x^2} using a combination of techniques.

$$S = \int_{-\infty}^\infty e^{-x^2} dx$$

$$S^2 = \left(\int_{-\infty}^\infty e^{-x^2} dx \right) \left(\int_{-\infty}^\infty e^{-x^2} dx \right)$$

$$S^2 = \left(\int_{-\infty}^\infty e^{-x^2} dx \right) \left(\int_{-\infty}^\infty e^{-y^2} dy \right)$$

Now let us translate to polar coordinates

$$S^2 = \left(\int_0^{2\pi} \int_0^\infty r e^{-(r\cos\theta)^2} r e^{-(r\sin\theta)^2} dr d\theta \right)$$

$$S^2 = \left(\int_0^{2\pi} \int_0^\infty r e^{-(r\cos\theta)^2 + (r\sin\theta)^2} dr d\theta \right)$$

$$S^2 = \left(\int_0^{2\pi} \int_0^\infty r e^{-(r^2)} dr d\theta \right)$$

$$S^2 = \left(\int_0^{2\pi} \int_0^\infty \frac{1}{2} e^{-(s)} ds d\theta \right)$$

$$S^2 = \left(\int_0^{2\pi} d\theta \right)$$

$$S^2 = \pi$$

$$S = \sqrt{\pi}$$

59.18 Triple Integrals

The triple integral allows the computation of volume or mass in three-dimensional space.

Let $V \subset \mathbb{R}^3$ be a bounded region, then:

$$\iiint_V f(x, y, z)\, dV = \lim_{n \to \infty} \sum_{k=1}^{n} f(x_k, y_k, z_k)\, \Delta V_k,$$

where ΔV_k are small sub-volumes approximating V.

In practice, evaluate:

$$\iiint_V f(x, y, z)\, dz\, dy\, dx,$$

with limits determined by the geometry of the volume. The order of integration may be rearranged to simplify computation.

$$\int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\,dy\,dx$$

59.18.1 Triple Integrals in Cartesian Coordinates

In this short section we, will look at an example of how to calculate the volume enclose in between two functions using triple integrals.

Given are the functions $f_1(x, y) = x^2 + y^2$ and $f_2(x, y) = 3 - x^2 - y^2$
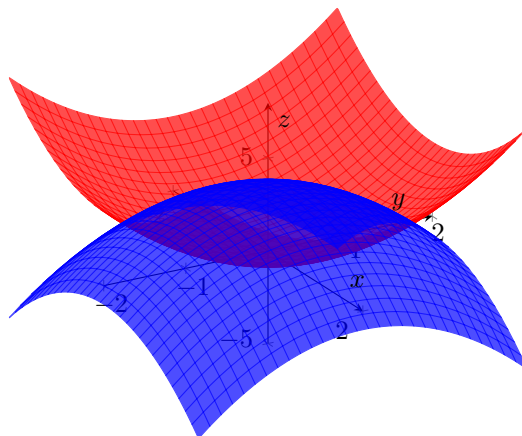
Let us find the intersection

$$x^2 + y^2 = 3 - x^2 - y^2$$

$$x^2 + y^2 = \frac{3}{2}$$

Now we have to find the limits of integration for

$$\int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} \int_{f_1(x,y)}^{f_2(x,y)} dz\,dy\,dx,$$

using our constraint and our original functions.

$$\int_{x=a}^{x=b} \int_{y=-\sqrt{\frac{3}{2}-x^2}}^{\sqrt{\frac{3}{2}-x^2}} \int_{x^2+y^3}^{3-x^2-y^2} dz\,dy\,dx$$

Here z means the height which is just the output of the original functions.

The bounds of y just by looking at out constraint and solving for it. And x only depends on the concrete limit of integration.

59.18.2 Coordinate Transformation and the Jacobian Determinant

Let a coordinate transformation be defined by:

$$x = x(u, v, w), \quad y = y(u, v, w), \quad z = z(u, v, w)$$

Then the triple integral transforms as:

$$\iiint_{(x,y,z)} f(x, y, z)\,dx\,dy\,dz = \iiint_{(u,v,w)} f(x(u, v, w), y(u, v, w), z(u, v, w)) \cdot |\det J|\,du\,dv\,dw$$

where J is the *Jacobian matrix* of the transformation:

$$J = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{bmatrix}$$

Example: Transformation to Cylindrical Coordinates

The standard cylindrical coordinate transformation is:

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z$$

The Jacobian determinant is:

$$|\det J| = \begin{vmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

Thus, the triple integral in cylindrical coordinates becomes:

$$\iiint f(x, y, z)\,dx\,dy\,dz = \iiint f(r \cos \varphi, r \sin \varphi, z)\,r\,dr\,d\varphi\,dz$$

59.18.3 Calculating the center of mass with triple integrals

We can make use of triple integration to calculate the center of mass of a body with respect to x, y or z with the following formulas:

$$x_s = \iiint_V xp(x, z, y)dz\ dy\ dx$$

$$y_s = \iiint_V yp(x, z, y)dz\ dy\ dx$$

$$z_s = \iiint_V zp(x, z, y)dz\ dy\ dx$$

Where $p(x, y, z)$ is our mass function.

59.19 Surfaces

In vector calculus sometimes, we want to examine the surface generated by a function. An approach to analyze this surfaces is via parameterization, and another is using explicit and implicit forms.

59.19.1 Explicit and Implicit Forms

For a 3d surfaces the *Implicit form* is given by

$$F(x, y, z) = c$$

and the *Explicit* by

$$z = f(x, y)$$

59.19.2 Parametric Form

It is given by

$$\vec{r}(u, v) = f(u, v)\hat{i} + g(u, v)\hat{j} + h(u, v)\hat{k}$$

where $a \le u \le b$ and $c \le v \le d$.

Example:

For the example we take the cone formed by the function $z = \sqrt{x^2 + y^2}$ which is in the explicit form.

The implicit form is then $0 = z - \sqrt{x^2 + y^2}$ and the parametric form can be described very naturally by

$$\vec{r}(r, \theta) = r\cos(\theta)\hat{i} + r\sin(\theta)\hat{j} + r\hat{k}$$

using something similar to cylindrical coordinates with $r = \sqrt{x^2 + y^2}$.

59.19.3 Surface Area Formula

Continuing with the surfaces in 3d we can compute the area using the following formula

$$A = \iint_D \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| du\, dv.$$

Notice that we are using the fact that the length of a cross product is the area of the parallelogram formed by vectors it is orthogonal to. Using the this on a small scale with the partial of u and v and then adding up results in the formula.

Example I: Area of the sphere

We want to compute the surface area of a sphere of radius a and show that it equals $4\pi a^2$. Let us proceed step by step.

The equation of a sphere of radius a centered at the origin is given implicitly by

$$x^2 + y^2 + z^2 = a^2.$$

To parametrize the sphere, one possible way is to solve for z:

$$z = \pm\sqrt{a^2 - x^2 - y^2}, \quad x^2 + y^2 \le a^2.$$

This representation works, but it splits the sphere into two surfaces (upper and lower hemispheres). This makes the parametrization less convenient for computing surface area.

A more natural parametrization comes from spherical coordinates. Any point on the sphere can be written as

$$\begin{cases} x = a \sin \theta \cos \varphi, \\ y = a \sin \theta \sin \varphi, \\ z = a \cos \theta, \end{cases}$$

where the parameters range over

$$0 \leq \theta \leq \pi, \quad 0 \leq \varphi < 2\pi.$$

Here, θ is the polar angle (measured from the z-axis) and φ is the azimuthal angle in the xy-plane.

The surface area of a parametrized surface $\vec{r}(u, v)$ is given by

$$A = \iint_D \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| \, du \, dv.$$

For the parametrization $\vec{r}(\theta, \varphi) = (a \sin \theta \cos \varphi, \, a \sin \theta \sin \varphi, \, a \cos \theta)$, we compute:

$$\frac{\partial \vec{r}}{\partial \theta} = (a \cos \theta \cos \varphi, \, a \cos \theta \sin \varphi, \, -a \sin \theta),$$

$$\frac{\partial \vec{r}}{\partial \varphi} = (-a \sin \theta \sin \varphi, \, a \sin \theta \cos \varphi, \, 0).$$

Their cross product is

$$\frac{\partial \vec{r}}{\partial \theta} \times \frac{\partial \vec{r}}{\partial \varphi} = (a^2 \sin^2 \theta \cos \varphi, \, a^2 \sin^2 \theta \sin \varphi, \, a^2 \sin \theta \cos \theta).$$

The magnitude is

$$\left\| \frac{\partial \vec{r}}{\partial \theta} \times \frac{\partial \vec{r}}{\partial \varphi} \right\| = a^2 \sin \theta.$$

Therefore, the surface area is

$$A = \int_0^{2\pi} \int_0^\pi a^2 \sin \theta \, d\theta \, d\varphi.$$

Compute the integrals:

$$A = a^2 \left(\int_0^{2\pi} d\varphi \right) \left(\int_0^\pi \sin \theta \, d\theta \right).$$

$$A = a^2 (2\pi)(2) = 4\pi a^2.$$

Thus, the surface area of the sphere is

$$4\pi a^2$$

$$\mathfrak{QED}$$

Example II:

Compute the surface area of the intersection of the plane $z = -x$ and the cylinder $x^2 + y^2 = 4$.

We will use polar coordinates because the intersection is circular.

$$\vec{r}(r, \theta) = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \\ -r\cos(\theta) \end{bmatrix}$$

Take the partials

$$\vec{r}_r = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \\ -\cos(\theta) \end{bmatrix}$$

$$\vec{r}_\theta = \begin{bmatrix} -r\sin(\theta) \\ r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}$$

Compute the cross product:

$$\vec{r}_r \times \vec{r}_\theta = \begin{bmatrix} r \\ 0 \\ r \end{bmatrix}, \qquad \|\vec{r}_r \times \vec{r}_\theta\| = \sqrt{r^2 + r^2} = \sqrt{2}\, r$$

Now integrate over the disk $0 \le r \le 2, \ 0 \le \theta \le 2\pi$:

$$\text{Area} = \int_0^{2\pi} \int_0^2 \sqrt{2}\, r \, dr \, d\theta = \int_0^{2\pi} \left[\tfrac{\sqrt{2}}{2} r^2 \right]_0^2 d\theta = \int_0^{2\pi} 2\sqrt{2} \, d\theta = 4\pi\sqrt{2}.$$

The final result is

$$4\pi\sqrt{2}$$

59.20   Implicit Surface Area

Given a smooth, continuous surface $F(x, y, z) = c$ where $\langle \nabla F, \vec{p} \rangle \neq 0$ the surface area is given by

$$SA = \iint_R \frac{\|\nabla F\|}{|\langle \nabla F, \vec{p} \rangle|} dA$$

$\vec{p}$ can be $\hat{i}, \hat{j}$ or $\hat{k}$.

59.21   Explicit Surface Area

Given a smooth, continuous surface $F(x, y, z) = f(x, y) - z = 0$ where $\nabla F = f_x \hat{i} + f_y \hat{j} - \hat{k}$, $|\langle \nabla F, \hat{k} \rangle| = |-1| = 1$ and $\|\nabla F\| = \sqrt{f_x^2 + f_y^2 + 1}$

$$SA = \iint_R \sqrt{f_x^2 + f_y^2 + 1} \, dA$$

59.22 Surface Integrals

We define a function $G(x, y, z)$ define on a smooth surface S with an area $d\sigma$ whose integral is

$$SI = \iint_R G(x, y, z) d\sigma$$

The function G associates each point on the surface to a value in a higher dimension.

Surface integral in different forms:

Parametric

$$SI = \iint_R G(f(u, v), g(u, v), h(u, v)) \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| du\, dv.$$

Explicit

$$SI = \iint_R G(x, y, z) \sqrt{f_x^2 + f_y^2 + 1}\, dA$$

Implicit

$$SI = \iint_R G(x, y, f(x, y)) \frac{\|\nabla F\|}{|\langle \nabla F, \vec{p} \rangle|} dA$$

59.23 The Orientation of a Surface

A smooth surface is *orientable* if there exists a continuous field of unit normal vectors to the surface.

59.24 The Flux of a Surface

For a continuous field $\vec{F}$ on a smooth surface S oriented with normal unit vectors $\vec{n}$,

$$Flux = \iint_S \langle \vec{F}, \vec{n} \rangle d\sigma$$

where σ is the surface area in the corresponding form. And the normal unit $\vec{n}$ is for parametric forms

$$\vec{n} = \frac{\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}}{\left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\|}$$

and for the implicit surfaces

$$\vec{n} = \frac{\nabla g}{\|\nabla g\|}$$

59.25 Curl

Let $\vec{F} = (F_1, F_2, F_3) : \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field.

$$\operatorname{curl} \vec{F} = \nabla \times \vec{F} = \begin{bmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{bmatrix}$$

These describe how the field spreads or rotates around a point. As a side not, the curl of the gradient of f is 0 if f has continuous mixed second derivatives.

There exists a connection between the curl and conservative vector fields. Notice that the components of the curl vector are the conditions similar to the conditions for a vector field to be conservative. Therefore, a conservative vector field has no curl because if the conditions are true then the curl is the zero-vector.

A point where curl $\vec{f} \neq 0$ is called a whirlpool.

59.26 Stokes's Theorem

For S a piecewise smooth oriented surface with piecewise smooth boundary C and $\vec{F}$ a field wit continuous first partials for each component on an open region containing S. The circulation around the curve is equal to

$$C = \oint_C \langle \vec{F}, d\vec{r} \rangle = \iint_S \langle (\nabla \times F), \vec{n} \rangle d\sigma$$

59.27 Divergence

We define *divergence* as

$$\operatorname{div} \vec{F} = \langle \nabla, \vec{F} \rangle = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}$$

A point where div $\vec{f} > 0$ is a called a source.

59.28 The Divergence Theorem

For S a piecewise smooth oriented, closed surface with piecewise smooth boundary C and $\vec{F}$ a field wit continuous first partials for each component on an open region containing S. The flux around the curve is equal to

$$\iint_S \langle \vec{F}, \vec{n} \rangle ds = \iiint_D \langle \nabla, \vec{F} \rangle dV$$

59.29 Unified View of Vector Calculus

Integrating a differential operator acting on a field over a domain is the same as adding the field components along the boundary.

Formally:

The integral of a differential form over a boundary $\partial \Omega$ of some orientable mainfold Ω is equal to the integral of its exterior derivative over the whole of Ω.

$$\int_{\partial \Omega} w = \int_\Omega dw$$

60 Differential Equations

A Differential Equation is an equation that relates a function (the dependent variable) with a variable or multiple variables (the independent variables) and its derivative.

They can be categorized in *Ordinary* and *Partial* Differential Equations.

60.1 Ordinary Differential Equations

An equation of the form

$$y^n = f(x, y, x', y', \ldots, y^{n-1})$$

is called *Ordinary Explicit Differential Equations of order n*

If $y^n = 0$ then it is *implicit*

If the exponent of the dependent variable is 1, and it is build like a linear equation then it is categorized as *Linear*

60.2 Initial Value Problems and Solutions

The specification of an explicit differential equation and the values

$$x_0, \quad y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \ldots, \quad y^{(n-1)}(x_0) = y_{n-1}$$

is called an *initial value problem (IVP)*.

An n-times differentiable function $y(x)$ that satisfies the explicit differential equation is called the *general solution* of the differential equation. If it also satisfies the conditions of the initial value problem from the previous definition, it is called the *particular solution* of the IVP.

60.3 Volterra Integral

The function $y(x)$ is a solution of the IVP $y' = f(x, y), \quad y(x_0) = y_0$ if and only if

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt$$

Proof:

$$\int_{x_0}^{x} f(t, y(t))dt = \int_{x_0}^{x} y'$$
$$y(x) = y(x_0)$$

The other way would be

$$\left(\int_{x_0}^{x} f(t, y(t))dt + y_0 \right)' = f(x, y(x)) = y'$$

$$\mathfrak{QED}$$

60.4 The Picard-Lindelöf Iteration Method

We can approximate the desired solution using the Volterra integral equation as follows. Start with an arbitrary initial function, for example a constant — preferably the initial value:

$$g_0(x) = y_0$$

Then form the next approximation to the desired function $y(x)$ as:

$$g_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, g_n(t)) \, dt$$

60.5 Integration case

To solve a ODE of the form $y' = f(x)$ it is as easy as integrating $f(x)$.

60.6 Separation of Variables

An equation $F(x, y, y', \dots) = 0$ is a called an equation of *separable variables* if it can be written in one of the following forms:

- $\frac{dy}{dx} = f(x)g(y)$

- $\frac{dy}{dx} = \frac{f(x)}{g(y)}$

- $\frac{dy}{dx} = \frac{g(y)}{f(x)}$

60.6.1 Steps to solve a ODE of separable variables

1. Solve for the derivative

2. Check that it is indeed an ODE of separable variables

3. Separate the variables

4. Integrate with respect to the independent variable

Example:

Given is $(y - 2)y' = 3x - 5$

It is an ODE of separable variable

$$y' = \frac{3x - 5}{y - 2}$$

Now we separate the variables and Integrate

$$\int (y - 2)y' dx = \int 3x - 5 dx$$

Note that $y' = \frac{dy}{dx}$ and $\frac{dy}{dx} dx = dy$ (only as a convenient fiction in reality we are changing the variables) thus,

$$\int (y - 2) dy = \int 3x - 5 dx$$

$$\frac{y^2}{2} - 2y + c = \frac{3x^2}{2} - 5x + c$$

The solution we get is an implicit solution. To make it look nice solve for y if you can.

60.6.2   Understanding the fiction

The previous fiction we used can explain in the following manner.

Given $f(y)\frac{dy}{dx}dx = \int f(y)dy$. We want to prove that this equality holds via the *chain rule* $\frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx}$.

$$\frac{d}{dx}\left(\int f(y)dy\right) = \frac{d}{dx}\left(\int f(y)dy\right)\frac{dy}{dx}$$

$$= f(y)dy\frac{dy}{dx}$$

$$\implies \int f(y)\frac{dy}{dx} = \int f(y)dy$$

$$\mathfrak{QED}$$

60.7   Geometry of ODE

You can imagine the plot of an ODE as a *Slope field* where the slope (derivative) of a function obeys a certain condition, this being the differential equation $\frac{dy}{dx} = expr$.

A point on this field is called an *Initial Condition* and the curve that goes through that point in the slope field is called an *Integral curve* which a solution to the differential equation with respect to that initial condition.

60.8   Existence and Uniqueness

Let $f(x, y)$ be continuous on a rectangle $[a, b] \times \mathbb{R}$. Then the differential equation

$$y' = f(x, y)$$

has at least one solution. A unique solution is only guaranteed under additional conditions.

A function $f : \mathbb{R}^2 \to \mathbb{R}$ is called *Lipschitz continuous with respect to y* if there exists a constant $L > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2| \quad \text{for all } (x, y_1), (x, y_2) \in D.$$

Let $x_0 \in [a, b]$, and let $f : [a, b] \times \mathbb{R} \to \mathbb{R}$ be continuous and bounded, and Lipschitz continuous in y. Then the initial value problem

$$y' = f(x, y), \quad y(x_0) = y_0$$

has a unique solution on $[a, b]$.

Or in easier words: If f and $\frac{\partial f}{\partial y}$ are continuous near (x_0, y_0) then there is a unique solution on an interval $\alpha < x_0 < \beta$ to the IVP

$$y' = f(x, y), \quad y(x_0) = y_0$$

60.9   How to deal with initial conditions

After solving a differential equation we may have the problem that certain initial conditions were set at the start, thus, making our general solution not specific.

Solving this problem is pretty easy, because we just have to solve for the constant terms in our general solution with respect to our initial conditions. Like for example $y(0) = 1$ we would plug 0 for all x's in our general solution the whole equation will be set equal to 1.

60.10 Solving DE with Substitution

In some cases a complex function can be simplified via a substitution.

60.10.1 Case I

$$y' = f(ax + by(x) + c)$$

With a substitution like $z(x) = ax + by(x) + c$ the right-hand side will become $f(z)$ and for the left-hand side we get $y(x) = \frac{z(x) - ax - c}{b}$.

Now we can substitute in the original equation, and we get

$$y'(x) = \frac{1}{b}(z'(x) - a) = f(z)$$

$$z'(x) = a + bf(z)$$

Now we can integrate and find the solution and then return to our original variables.

Example:

$$y' = (x + y(x))^2$$

Here $a = 1$, $b = 1$, $c = 0$ and $f(z) = z^2$

1. Substitute

$$z(x) = x + y(x)$$
$$y(x) = z(x) - x$$
$$y'(x) = z'(x) - 1$$

2. Plug in the DE

$$y' = (x + y(x))^2$$
$$z'(x) - 1 = z(x)^2$$

3. Integrate

$$z'(x) = 1 + z(x)^2$$
$$\frac{z'(x)}{1 + z(x)^2}$$
$$\int \frac{1}{1 + z^2} dz = \int 1 dx$$
$$\arctan(z) = x + c$$
$$z = \tan(x + c)$$

4. Return to the original variables

$$x + y(x) = \tan(x + c)$$
$$y(x) = \tan(x + c) - x$$

60.10.2 Case II

$$y' = f(\frac{y}{x})$$

For this to be valid $z(x) = \frac{y}{x}$ our $y(x)$ needs to look like $y(x) = z(x)x$ and so $y' = z(x) + z'(x)x$.

Now we substitute in the differential equation using $f(z) = z(x) + z'(x)x$

$$z'(x) = \frac{f(z) - z(x)}{x}$$

Now we can integrate and then return to our original variables. $y(x) = z(x)x$, where $z(x)$ is our result of the integration.

Example:

$$y' = \frac{y(x)}{x} + 1$$

This in converted to $y' = f(z) = z + 1$

1. Substitute

$$z(x) = \frac{y(x)}{x}$$
$$y(x) = z(x)x$$
$$y'(x) = z(x) + z'(x)x$$

2. Plug in the DE

$$z(x) + z'(x)x = z(x) + 1$$

3. Integrate

$$z'(x)x = 1$$
$$\int z'\,dz = \int \frac{1}{x}dx$$
$$z = (\ln(x) + c)x$$

60.11 Linear Differential Equations

An ODE is called linear if all occurrences of the dependent variable and its derivatives have an exponent of at most 1.

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

If such an Equation has also the property $b(x) = 0$ it is called *Homogeneous*

60.11.1 How to solve a Linear Differential Equation

In this case we will focus on first Order ODE.

Write it in the standard form $y' + p(x)y = f(x)$

Now use *Integrating Factor Method*

60.12   Integrating Factor Method

We are going to multiply our expression by function $r(x)$

$$r(x)y' + r(x)p(x)y = r(x)f(x)$$

Now we would like to write the left side as $\frac{d}{dx}yr(x)$ because now if we were to integrate it would take us to the expression in the right.

When we evaluate this new expression we get that

$$\frac{d}{dx}yr(x) = y'r(x) + r'(x)y$$

which is not quite what we actually have, but it is close. Now note that we can say

$$r'(x) = r(x)p(x)$$

This is a differential equation of separable variables thus, we can

$$\frac{r'(x)}{r(x)} = p(x)$$

$$\int \frac{r'(x)}{r(x)}dx = \int p(x)dx$$

$$\ln(r(x)) = \int p(x)dx$$

after using e

$$r(x) = e^{\int p(x)dx}$$

Now we have found an $r(x)$ and now that we know that such a function exists we can return to our initial problem and say

$$\frac{d}{dx}r(x)y = r(x)f(x)$$

$$y = \frac{1}{r(x)}\int r(x)f(x)dx$$

and

$$r(x) = e^{\int p(x)dx}$$

To summarize, given $y' + f(x)y = g(x)$ the steps are:

1. Find $c(x) = e^{\int f(x)dx}$.

2. Multiply the original expression with $c(x)$.

$$y'c(x) + f(x)c(x)y = g(x)c(x)$$

3. Use the power rule of differentiation to simplify.

$$y'c(x) + f(x)c(x)y = \frac{d}{dx}(c(x), y)$$

4. Integrate.

$$\int \frac{d}{dx}\left(c(x), y\right) dx = \int g(x)c(x)dx$$

$$c(x)y = \int g(x)c(x)dx$$

$$y = \frac{\int g(x)c(x)dx}{c(x)}$$

Example:

Given are $y' + 4y = e^{-x}$ and $y(0) = \frac{4}{3}$.

Here $4 = p(x)$ and $e^{-x} = f(x)$.

Now let us use the formula for $r(x) = e^{\int p(x)}$. This gives us $r(x) = e^{\int 4dx} = e^{4x}$.

Now recall what we saw earlier

$$e^{4x}y' + e^{4x}4y = e^{4x}e^{-x} = e^{3x}$$

Our left side is just $\frac{d}{dx}e^{4x}y$

Now let us complete the exercise with our last formula by integrating both sides

$$\frac{d}{dx}e^{4x}y = e^{3x}$$

$$y = \frac{1}{e^4x}\int e^{3x}$$

$$y = \frac{1}{e^4x}\frac{1}{3}e^{3x} + c$$

Which is our general solution

To find our specific solution we just plug 0 in to our general solution that gives us $\frac{1}{3} + c = \frac{4}{3}$ thus, $c = 1$

60.13 Homogeneous Differential Equations

A differential equation

$$M(x, y)dx + N(x, y)dy = 0$$

is considered *Homogeneous* if and only if $M(x)$ and $N(x)$ are *homogeneous*, and they have the same degree.

The equation $y' + f(x)y = 0$ is called *homogeneous linear differential equation of first order*.

60.13.1 Homogeneous Equation

A function $f(x, y)$ is called *homogeneous* if and only if it can be written

$$f(xt, yt) = t^n f(x, y)$$

Example:

$$f(x, y) = x^3 + y^3$$
$$f(tx, ty) = (tx)^3 + (ty)^3$$
$$t^3(x^3 + y^3) = t^3 f(x, y)$$

It is important to note that the key is that each term have a total degree of n. For example: $x^2y + x3y^2$

60.13.2 Homogeneous Functions Theorem

If $f(x,y)$ is homogeneous of degree 0 in x and y then f is a function of $\frac{y}{x}$

Or $f(x,y,y')$ is homogeneous if it can be written as $y' = f(\frac{x}{y})$ or $y' = f(\frac{y}{x})$

60.14 Solving Homogeneous DE I

1. Write in the form $M(x,y)dx + N(x,y)dy = 0$

2. Check if it is homogeneous

3. Change the variables to $y = ux$ or $x = uy$ depending on the situation

4. Solve using the method of separable variables method

5. Rewrite the answer in terms of the original variables again

Example:

Give is the function $(x - y)dx = -xdy$

Step 1. We write it in the desired form $(x - y)dx + xdy = 0$

Step 2. We notice that both are of degree 1. This time we skip the rigorous check.

Step 3. We are going to choose the simpler function to do the change of variable depending on the *differential!*

The simpler function is xdy thus, we are to change y

$$y = ux$$
$$dy = ux' + xu'$$

and

$$u = \frac{y}{x}$$

Now substitute

$$(x - ux)dx + x(udx + xdu) = 0$$
$$xdx - uxdx + xudx + x^2du = 0$$
$$(x - ux + xu)dx + x^2du = 0$$
$$xdx + x^2du = 0$$
$$xdx = -x^2du$$
$$dx = -xdu$$

Step 4. Now we can integrate both sides

$$\int \frac{1}{x}dx = \int du$$
$$\ln x + c = -u + c$$

Step 5. No we use $u = \frac{y}{x}$ to return to our original variables

$$\ln x + c = -\frac{y}{x} + c$$
$$-x\ln x - xc = y$$

60.15   Solving Homogeneous DE I

Another method for solving this kind of equations $y' + f(x)y = 0$ is to follow the next steps.

Step 1: Separate the variables

$$\frac{dy}{dx} = -f(x)y$$

Step 2: Use the separation of variables

$$\int \frac{\frac{dy}{dx}}{y} = \int -f(x)dx$$

$$\int \frac{dy}{y} = \int -f(x)dx$$

$$\ln|y| = \int -f(x)dx$$

$$y = Ce^{\int -f(x)dx}$$

This method works more directly.

60.16   Inhomogeneous Equation

The equation $y' + f(x)y = g(x)$ is called *linear inhomogeneous differential equation of first order* $g(x)$ is called the *distortion function*. The correspondent IVP is called *linear* IVP.

60.17   Solutions of LDE

Every linear IVP with continuous functions $f(x)$ and $g(x)$ plus a bounded $f(x)$ has exactly one solution.

Proof:

$$y' = -f(x)y + g(x) \text{ therefore, } |-f(x)y_1 + g(x) + -f(x)y_2 + g(x)| = |f(x)||y_1 - y_2|$$
$$\leq M|y_1 - y_2|$$

Therefore, the function is Lipschitz continuous, thus, a solution is possible.

$$\mathfrak{QED}$$

60.17.1   Variation of the Constants

Recall that for the homogeneous case $y' + f(x)y = 0$ case we use $y = ce^{\int -f(x)dx}$, but now we have $y' + f(x)y = g(x)$, thus, let us use the following approach

$$y = c(x)e^{\int -f(x)dx}$$

Now use the product rule to derivate

$$y' = c'(x)e^{\int -f(x)dx} + c(x)e^{\int -f(x)dx}(-f(x))$$
$$= c'(x)e^{\int -f(x)dx} - f(x)y$$

thus,

$$y' + f(x)y = c'(x)e^{\int -f(x)dx}$$

therefore, $g(x) = c'(x)e^{\int -f(x)dx}$

Then we can get a solution for $c(x)$ from

$$g(x) = c'(x)e^{\int -f(x)dx}$$
$$g(x) = c'(x)e^{-\int f(x)dx}$$
$$c'(x) = \frac{g(x)}{e^{-\int f(x)dx}}$$
$$c'(x) = g(x)e^{\int f(x)dx}$$

Example:

Given is $y' - \cos(x)y = xe^{\sin(x)}$ with $y(0) = 3$.

1. Solve as if it were homogeneous

$$y' - \cos(x)y = 0$$
$$y' = \cos(x)y$$
$$\frac{y'}{y} = \cos(x)$$
$$\int \frac{1}{y}dy = \int \cos(x)dx$$
$$\ln|y| + c = \sin(x) + c$$
$$y = e^{\sin(x)}c$$

2. Use $y = c(x)e^{\sin(x)dx}$

$$y' = c(x)\cos(x)e^{\sin(x)} + c'(x)e^{\sin(x)}$$
$$y' = \cos(x)y + c'(x)e^{\sin(x)}$$
$$y' - \cos(x)y = c'(x)e^{\sin(x)}$$

3. Compare with the original equation

$$y' - \cos(x)y = xe^{\sin(x)}$$
$$y' - \cos(x)y = c'(x)e^{\sin(x)}$$

4. Solve for $c'(x)$

$$xe^{\sin(x)} = c'(x)e^{\sin(x)}$$
$$x = c'(x)$$

5. Integrate

$$\int x \, dx = \int c'(x) \, dx$$

$$\frac{x^2}{2} + c = c(x)$$

6. Substitute in our approach $y = c(x)e^{\sin(x)}$ from earlier

$$y = \left(\frac{x^2}{2} + c \right) e^{\sin(x)}$$

7. Solve for c using our initial condition if asked

$$y(0) = 3$$

$$3 = \left(\frac{0^2}{2} + c \right) e^{\sin(0)}$$

$$3 = c$$

60.18   Existence and Uniqueness Differential Equations

For the IVP

$$y^n + p_{n-1}y^{n-1} + \cdots + p_0(x)y = f(x)$$

$$y(a) = b_0, y'(a) = b_1, \ldots, y^{n-1}(a) = b_{n-1}$$

If all p_i and f are continuous on the interval I about a, then there exists a unique solution on I.

60.19   Theorem of Superposition and the General Solution

Suppose $y_1, \ldots, y_n$ solve

$$y^n + p_{n-1}y^{n-1} + \cdots + p_0(x)y = 0$$

Then $y = c_1y_1 + \cdots + c_ny_n$ also solves the equation, and it is the *General Solution* if and only if the *Wronskian* $W(t) \neq 0$ for some t_0.

60.20   Undetermined Coefficients

Given an equation $y' + ay = g(x)$ notice that a does not depend on x. The homogeneous equation to this case will be $y' + ay = 0$ with a solution y_h and some partial solution y_p. Now remember from *Linear Algebra* the *general = particular + homogeneous* more precisely $y = y_p + y_h$ therefore, adding a particular solution to a homogeneous does not change the general solution of our DE.

Now let us substitute in our differential equation.

$$(y_h + y_p)' + a(y_h + y_p) = y_h' + y_p' + ay_h + ay_p$$
$$= (y_h' + ay_h) + (y_p' + ay_p) = 0 + g(x) = g(x)$$

60.20.1 Steps for solving Constant Coefficients problems

1. Find the *homogeneous* solution for $y' + ay = 0$. y_h is going to have a free parameter.

2. Find a partial solution for $y' + ay = g(x)$

3. Use $y = y_h + y_p$

4. If you have an IVP solve for the free parameter in y_h

Now we have to find a way to solve for y_p and y_h

60.20.2 Find the partial solution

We will use the method called based on guessing the type of the function on the right side.

Example:

Given is the following IVP: $y' + 2y = 2x + 13$ with $y(0) = 8$

By looking carefully we see that on the right-hand side we just have a normal polynomial, thus, we can substitute $y = bx + c$ and derivate $y' = b$. In the case for when we have higher order derivative we would differentiate for each of them.

Now substitute

$$b + 2(bx + c) = 2x + 13$$

$$2bx + b + 2c = 2x + 13$$

Now by comparing the Coefficients with the other side

$$2b = 2$$
$$b + 2c = 13$$

We get: $b = 1$, $c = 6$. And we have found a *particular* solution for our DE $y_p = x + 6$

Now we have to find a *homogeneous* solution with the formula we already know $y_h = ce^{\int f(x)dx}$ in our case $y_h = ce^{\int 2dx} = ce^{-2x}$

$$y = y_h + y_p = ce^{-2x} + x + 6$$

Now considering the condition $y(0) = 8$, this gives:

$$y(0) = c + 6 = 8 \Rightarrow c = 2$$

and we obtain the particular solution:

$$y = 2e^{-2x} + x + 6$$

60.20.3 Table of reference

Type of Forcing Function	Disturbance Function $g(x)$	Approach for y_p
Constant	k_0	c_0
Linear	$k_0 + k_1 x$	$c_0 + c_1 x$
Polynomial	$\sum_{i=0}^{n} k_i x^i$	$\sum_{i=0}^{n} c_i x^i$
Exponential	$ke^{bx},\ b \neq a$	$c_0 e^{bx}$
	ke^{ax}	$c_0 x e^{ax}$
Trigonometric	$k\sin(bx) + l\cos(bx)$	$c_0 \sin(bx) + c_1 \cos(bx)$

It is important to point out that sometimes it is necessary to combine different types of functions to get the solution.

Example:

$$y'' - 2x'3y = t^2 + 3e^{-t}\cos(4t)$$

Here the approach would be

$$y_p = (A + Bt + Ct^2) + D(e^{-t}\cos(4t)) + E(e^{-t}\sin(4t))$$

The best way is to tackle down each of the types of functions separately.

An alternative to way of using this method is to

1. Find the homogeneous solution.

2. Multiply both sides of the equation by it.

3. Integrate

This will give you the exact same result and is faster.

60.21   Bernoulli Equation

This is used for equations of the form

$$y' + P(x)y = Q(x)y^n$$

We can not use the methods we already know for this equation because it is linear but with a clever substitution we can turn it into a linear differential equation.

Let us use a substitution

$$u = y^{1-n}$$

then

$$u' = (1 - n)y^{-n}y'$$

Now this looks kinda similar to the original equation multiply by y^{-n} that looks like

$$y^{-n}y' + P(x)y^{-n}y = Q(x)y^{-n}yn$$

$$y^{-n}y' + P(x)y^{1-n} = Q(x)$$

Notice that $y^{-n} = \frac{u'}{(1-n)y'}$, thus, after substitution

$$\frac{1}{1 - n}u' + P(x)u = Q(x)$$

Now our equation is *linear*, and we can use the *integrating factor* method.

Example:

Given is $y' - 5y = \frac{-5}{2}xy^3$.

$$u = y^{-2} \text{ and } u' = -2y^{-3}y'$$

$$y^{-3} = \frac{u'}{-2y'}$$

And our equation after multiplying by y^{-3}

$$y^{-3}y' - 5yy^{-3} = -\frac{5}{2}x$$

$$y^{-3}y' - 5y^{-2} = -\frac{5}{2}x$$

$$\frac{u'}{-2y'}y' - 5u = -\frac{5}{2}x$$

Now we can use the *Integrating Factor Method*. First bring to the standard form. $y' + p(x)y = f(x)$

$$u' + 10u = 5x$$

Remember that $r(x) = e^{\int p(x)dx} = e^{\int 10dx} = e^{10x}$ therefore, we have

$$e^{10x}u' + 10ue^{10x} = e^{10x}5x$$

Now let us integrate

$$\frac{d}{dx}\left(e^{10x}u\right) = e^{10x}5x$$

$$Ce^{10x}u = \int e^{10x}5xdx = 5x\frac{e^{10x}}{10} - \int e^{10x}5dx$$

$$Ce^{10x}u = \int e^{10x}5xdx = 5x\frac{e^{10x}}{10} - \frac{5}{10}\int e^{10x}dx$$

$$Ce^{10x}u = \frac{xe^{10x}}{2} - \frac{5}{100}e^{10x} + c$$

$$Ce^{10x} = \frac{x}{2} - \frac{1}{20} + \frac{c}{e^{10x}} = y^{-2}$$

60.22   Autonomous Equations

An *Autonomous Differential Equation* only depends on the dependent variable y. **Example:** $\frac{dy}{dt} = (1+y)(1-y)$.

60.22.1   Equilibrium Solutions

Values where $f(y) = 0$ are *Equilibrium Solutions*.

Asymptotically stable

An equilibrium solution $y = a$ is *asymptotically stable* if solutions that start near a tend towards a as $t \to \infty$.

Asymptotically unstable

An equilibrium solution $y = a$ is *asymptotically unstable* if solutions that start near a leave as $t \to \infty$

60.23 Linear Inhomogeneous DEs with Non-Constant Coefficients

The method of *variation of constants* is often used alongside the *superposition principle.* In this method, the integration constant in the function $c(x)$ is omitted as it is part of the particular solution.

The superposition principle also holds for linear inhomogeneous differential equations with non-constant coefficients:

$$y'_p + f(x)y_p = g(x) \quad \text{(particular solution)}$$
$$y'_h + f(x)y_h = 0 \quad \text{(homogeneous solution)}$$

For the total solution $y = y_h + y_p$, we get:

$$\begin{aligned} y' + f(x)y &= (y_h + y_p)' + f(x)(y_h + y_p) \\ &= y'_h + y'_p + f(x)y_h + f(x)y_p \\ &= g(x) \end{aligned}$$

Thus, the solution can be decomposed into the general solution of the homogeneous DE and a particular solution based on the form of the right-hand side.

Example:

$$y' = -\frac{y}{x} + 1$$

The homogeneous DE is $y' = -\frac{y}{x}$ which can be solved by separating the variables. $y_h = \frac{c}{x}$

Now let use $y_p = \frac{c(x)}{x}$.

$$y'_p = \frac{c'(x)x - c(x)}{x^2}$$

Let us break down the fraction to find where to put our homogeneous solution we have already found

$$\frac{c'(x)}{x} - \frac{c(x)}{x^2} = \frac{c'(x)}{x} + \frac{1}{x}\frac{c}{x} = \frac{c'(x)}{x} + \frac{1}{x}y$$

and we get

$$y'_p = -\frac{y}{x} + \frac{c'(x)}{x}$$

Here we compare with the inhomogeneous part of the original equation and get

$$\frac{c'(x)}{x} = 1 \implies c'(x) = x$$

then $c(x)$ is equal to $\frac{x^2}{2}$ and therefore, $y_p = \frac{x^2}{x}\frac{1}{x} = \frac{x}{2}$

As our final result we get $y = \frac{c}{x} + \frac{x}{2}$

60.24   Power Series Approach

Another method to solve a certain type of differential equation is to find the function given an initial conditions. This can be better understood with an example.

$$y' = y$$
$$y(0) = 1$$

We say $y = \sum_{n=0}^{\infty} a_n x^n$ and $y(0) = a_0 = 1$.

Now let us derivate you y

$$y = \sum_{n=0}^{\infty} a_n x^n = a_0 x^0 + a_1 x^1 + \cdots = 1 + a_1 x^1 + a_2 x^2 + \cdots$$

$$y' = 0 + a_1 + 2a_2 x + 3a_3 x^2 + \cdots$$

Thus, we can rewrite the sum as

$$y' = \sum_{n=1}^{\infty} a_n n x^{n-1}$$

Starting from 1 because 0 does not contribute to the sum. After shifting the index we get

$$y' = \sum_{n=0}^{\infty} a_{n+1}(n+1)x^n$$

Now we build our original DE with our new definitions for y' and y

$$\sum_{n=0}^{\infty} a_{n+1}(n+1)x^n = \sum_{n=0}^{\infty} a_n x^n$$

If we compare the coefficients we get

$$a_n = a_{n+1}(n+1)$$

$$\frac{a_n}{(n+1)} = a_{n+1}$$

after iterating backwards we get

$$a_{n+1} = \frac{a_n}{n+1} = \frac{\frac{a_{n-1}}{(n-1+1)}}{n+1} = \frac{\frac{\frac{a_{n-2}}{(n-1)}}{(n-1+1)}}{n+1}$$

$$a_n = \frac{a_0}{(n+1)!} = \frac{1}{(n+1)!} = 1$$

therefore, $a_n = \frac{1}{n!}$ which gives us the solution

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n = e^x$$

60.24.1 Theorem for the Power Series of DE

For a differential equation of the form:

$$A(x)y^n + \cdots + B(x)y' + C(x)y = 0$$

more specific after dividing by $A(X)$

$$y^n + \cdots + P(x)y' + Q(x)y = 0,$$

with $x = a$ as an *Ordinary Point* if $\ldots, P(x)$ and $Q(x)$ are *Analytic* at $x = a$. Otherwise, a *Singular Point*.

Then this equation has n linearly independent solutions of the form

$$y(x) = \sum_{n=0}^{\infty} c_n(x - a)^n$$

The radius of convergence is at least as large as distance to the nearest *Singular Point*.

60.25 Exact Differential Equations

We will take a short look at functions with two inputs via implicit differentiation $F(x, y(x)) = 0$. With $\frac{\partial F}{\partial x} = P$ and $\frac{\partial F}{\partial y} = Q$.

$$P(x, y)dx + Q(x, y)dy = 0$$

after dividing by dx

$$P(x, y) + Q(x, y)y' = 0$$

We define this kind of differential equations as *exact*. And the $P_y = Q_x$ as *Integrability-Condition*.

By finding the *potential* we can solve this kind of differential equations.

Example:

Step 1: Solve for y'

$$(12xy + 3)dx + 6x^2dy = 0$$
$$(12xy + 3) + 6x^2y' = 0$$
$$y' = -\frac{12xy + 3}{6x^2}$$

Step 2: Test the Integrability-Condition

$$P(x, y) = 12xy + 3$$
$$Q(x, y) = 6x^2$$
$$P_y = 12x$$
$$Q_x = 12x$$

Step 3: Integrate P with respect to x and add $h(y)$

$$F(x, y) = \int P(x, y)dx = \int (12xy + 3)dx$$
$$= 6x^2 y + 3x + h(y)$$

Step 4: Differentiate F with respect to y and set equal to Q

$$\frac{\partial}{\partial y} \left(6x^2 y + 3x + h(y) \right) = 6x^2 + c'(y)$$

Step 5: Integrate with respect to y

$$6x^2 + c'(y) = 6x^2$$
$$c'(y) = 0$$
$$\int c'(y) = \int 0 = C$$

But $C = 0$ because, C in Q is zero. Therefore, $F(x, y) = 6x^2 y + 3x$.

60.25.1 Euler Multipliers

Sometimes a seemingly *exact* differential equations is not possible to solve do the partial derivatives of $P(x, y)$ and $Q(x, y)$ not being equal. An approach to this problem are the so-called *Euler Multipliers*, which are some factor $\mu(x)$ or $\mu(y)$ that when found. Can help solve the differential equation.

$$P(x, y)\mu(x) + Q(x, y)y'\mu(x) = 0$$

or

$$P(x, y)\mu(y) + Q(x, y)y'\mu(y) = 0$$

The steps for solving this kind of problems are

1. Choose either $\mu(x)$ or $\mu(y)$ and multiply the whole equation by it.

2. Find the new partial derivatives of P and Q.

3. Set both partial equals and solve for μ in the newly created differential equation using the necessary method. Also make sure that the solution has no zeroes. This is necessary because otherwise, we would have multiplied our original equation times zero.

4. Substitute μ with the solution. And if now both partials are equal then you can now solve the exact differential equation. If not try with another μ using the other variable. And if this also does not work then the differential equation can not be solved. More specific, if $\mu(x)$ depends on y and $\mu(y)$ on x then there is no way to make the differential equation exact.

Example:

$$3y^2 + 2xyy' = 0, x > 0$$

Here we have $P(x, y) = 3y^2$ and $Q(x, y) = 2xy$, whose partial derivatives are different

$$\frac{\partial P}{\partial y} = 6y \neq 2y = \frac{\partial Q}{\partial x}$$

Now we try to use a multiplier $\mu(x)$.

$$P_2(x, y) = 3y^2\mu(x) \text{ and } Q_2(x, y) = 2xy\mu(x)$$

Now by taking the partials again

$$\frac{\partial P_2}{\partial y} = 6y\mu(x)$$
$$\frac{\partial Q_2}{\partial x} = 2y(\mu(x) + x\mu'(x))$$

Now by setting them equal $6y\mu(x) = 2y(\mu(x) + x\mu'(x))$ we can solve a new differential equation to find $\mu(x)$ and $\mu'(x)$.

$$6y\mu(x) = 2y(\mu(x) + x\mu'(x))$$
$$3\mu(x) = \mu(x) + x\mu'(x)$$
$$2\mu(x) = x\mu'(x)$$
$$2\mu = x\frac{d\mu}{dx}$$
$$\int \frac{2}{x}dx = \int \frac{d\mu}{\mu}$$
$$2\ln(x) = \ln(\mu)$$
$$\mu = x^2$$

Now finally we check if now the partials are equal, which in this case is true.

$$\frac{\partial P_2}{\partial y} = 6yx^2$$
$$\frac{\partial Q_2}{\partial x} = 6yx^2$$

Now the equation can be solved like normal.

60.26 Constant Coefficients ODE

Consider the second-order linear differential equation with constant coefficients:

$$ay'' + by' + cy = 0,$$

where $a, b, c \in \mathbb{R}$, and $a \neq 0$.

We look for solutions of the form $y = e^{rt}$. Substituting into the equation:

$$y = e^{rt},$$
$$y' = re^{rt}$$
$$y'' = r^2 e^{rt}$$

Substituting into the original equation:

$$ar^2 e^{rt} + bre^{rt} + ce^{rt} = 0$$

Factoring out e^{rt} (which is never zero):

$$e^{rt}(ar^2 + br + c) = 0 \quad \Rightarrow \quad ar^2 + br + c = 0.$$

This is a quadratic equation in r. Solving using the quadratic formula:

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

Suppose there are two linearly independent solution y_1 and y_2 to

$$y'' + p(x)y' + q(x)y = 0,$$

then the general solution is $y = c_1 y_1 + c_2 y_2$.

We analyze the nature of the roots r based on the discriminant $D = b^2 - 4ac$. Also, for dealing with initial conditions of the form $y^n = a, \dots$ we only have to plug the values into our solutions $y_h, y'_h, \dots$ to get the constants.

60.26.1 Case 1: Two Distinct Real Roots ($D > 0$)

There not much not to just plug the solutions of the root.

$$y = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

60.26.2 Case 2: Repeated Real Root ($D = 0$)

For this case we need to use a trick.

$$y = c_1 e^{rt} + c_2 t e^{rt}$$

Note that this could be contra intuitive but te^{rt} is a solution and is linearly independent. For more repeated roots you can use higher powers of t $t, t^2, t^3, \dots$.

60.26.3 Case 3: Complex Conjugate Roots ($D < 0$)

When the discriminant $D = b^2 - 4ac < 0$, the roots of the characteristic equation are complex:

$$r = \alpha \pm i\beta,$$

where $\alpha = -\frac{b}{2a}$ and $\beta = \frac{\sqrt{4ac - b^2}}{2a}$.

The general solution to the differential equation is:

$$y(t) = c_1 e^{(\alpha + i\beta)t} + c_2 e^{(\alpha - i\beta)t}.$$

To express this in real form, we use Euler's formula:

$$e^{i\beta t} = \cos(\beta t) + i\sin(\beta t).$$

Now rewrite each exponential term:

$$e^{(\alpha + i\beta)t} = e^{\alpha t} \cdot e^{i\beta t} = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t)),$$
$$e^{(\alpha - i\beta)t} = e^{\alpha t} \cdot e^{-i\beta t} = e^{\alpha t}(\cos(\beta t) - i\sin(\beta t)).$$

Now to get rid of the imaginary part we can do the following trick

$$\frac{y_1 + y_2}{2} = e^{\alpha t} \cos \beta t$$

$$\frac{y_1 - y_2}{2i} = e^{\alpha t} \sin \beta t$$

This might feel wrong, but these are in fact linear combinations of our original solution, and they are also linearly independent, because of that we can just use them for our general solution.

$$y = c_1 e^{\alpha t} \cos \beta t + c_2 e^{\alpha t} \sin \beta t$$

Example:

Given is $y''' + y' = 0$.

$$y = e^{rt}$$

$$r^3 + r = 0$$
$$r(r^2 + 1) = 0$$
$$r(r + i)(r - i) = 0$$

Now $r_1 = 0 \quad r_2 = -i \quad r_3 = i$

$$y = c_1 e^{0t} + c_2 e^{-it} + c_3 e^{it}$$

$$y = c_1 + c_2 \cos t + c_3 \sin t$$

60.26.4   Inhomogeneous Case

For the inhomogeneous case we would apply the method of the *Undetermined Coefficients*. But differentiating two times or more depending on the order of the differential equation and add the homogeneous solution to the solution we get from that.

Example:

Solve the differential equation:

$$y' + \frac{1}{x}y = \sin(x)$$

The integrating factor is:

$$\mu(x) = e^{\int \frac{1}{x} dx} = e^{\ln |x|} = |x|$$

Assuming $x > 0$, we take $\mu(x) = x$. Multiply through by the integrating factor:

$$xy' + y = x \sin(x)$$

The left-hand side is the derivative of xy:

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

$$\int \frac{d}{dx}(xy) \, dx = \int x \sin(x) \, dx$$

Use integration by parts on the right:

$$\int x \sin(x) dx = -x \cos(x) + \int \cos(x)\, dx = -x \cos(x) + \sin(x)$$

So:

$$xy = -x \cos(x) + \sin(x) + C$$

$$y = -\cos(x) + \frac{\sin(x)}{x} + \frac{C}{x}$$

60.27 The Wronskian

Let $y_1, y_2, \ldots, y_n$ be n functions that are at least $(n-1)$-times differentiable on some interval. The *Wronskian* of these functions is defined as:

$$W(y_1, y_2, \ldots, y_n)(x) = \begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{vmatrix}$$

If $W(A) \neq 0$ then they are linearly independent.

60.28 Linearity Property

Let $y_p^{(1)}$ be a particular solution of

$$y'' + ay' + by = g_1(x)$$

and $y_p^{(2)}$ a particular solution of

$$y'' + ay' + by = g_2(x),$$

then

$$y_p = y_p^{(1)} + y_p^{(2)}$$

is a particular solution of

$$y'' + ay' + by = g_1(x) + g_2(x).$$

60.29 Variation of Parameters

For a linear differential equation of the form

$$y'' + p(x)y' + q(x)y = f(x)$$

The homogeneous equation has a solution of the linearly independent y_1 and y_2 of the form $y = u_1 y_1 + u_2 y_2$. We can guess the functions y_1 and y_2.

So, for finding a solution of these types of equations we use the following formulas.

$$u_1 = -\int \frac{y_2 g}{y_1 y_2' - y_2 y_1'}\, dx$$

$$u_2 = -\int \frac{y_1 g}{y_1 y_2' - y_2 y_1'} dx$$

These formulas come from:

Differentiating two times:

$$y = u_1 y_1 + u_2 y_2$$

$$y' = u_1 y_1' + y_1 u_1' + u_2 y_2' + y2 u_2'$$

Here we can add another constraint that $y_1 u_1' + y2 u_2' = 0$. This allows us to simplify the expression to:

$$y' = u_1 y_1' + u_2 y_2'$$

Now we can derivate one more time:

$$y'' = u_1 y_1'' + y_1' u_1' + u_2 y_2'' + y_2' u_2'$$

Substitute in the original equation

$$y'' + p(x)y' + q(x)y = f(x)$$

$$u_1 y_1'' + y_1' u_1' + u_2 y_2'' + y_2' u_2' + p(x)(u_1 y_1' + u_2 y_2') + q(x)(u_1 y_1 + u_2 y_2) = g(x)$$

$$u_1 y_1'' + y_1' u_1' + u_2 y_2'' + y_2' u_2' + p(x)(u_1 y_1') + p(x)(u_2 y_2') + q(x)(u_1 y_1) + q(x)(u_2 y_2) = g(x)$$

$$u_1 y_1'' + y_1' u_1' + u_2 y_2'' + y_2' u_2' + p(x)(u_1 y_1') + p(x)(u_2 y_2') + q(x)(u_1 y_1) + q(x)(u_2 y_2) = g(x)$$

With non-generic function a lot of the terms would cancel out. After that we add multiply by y_1 one of the rows and y_2 so that we can cancel them in a row operation. Which will left us with only two integrals to solve.

60.30   Which method to use depending on the type of differential equations

60.30.1   First Order

- Separable $\Rightarrow$ Separation of Variables.
- Form $y' = f(ax + by(x) + c) \Rightarrow$ Substitution $z = ax + by(x) + c$ and $f(z)$.
- Form $y' = \frac{y}{x} \Rightarrow$ Substitution $z = \frac{y}{x}$
- Linear $\Rightarrow$ Integrating Factor or Variation of Constants.
- Bernoulli (Non-linear) $\Rightarrow$ Bernoulli Substitution.
- Autonomous $\Rightarrow$ Equilibrium Analysis.
- Non-homogeneous $\Rightarrow$ Undetermined Coefficients or integrating factor.

60.30.2   Second Order and higher

- Constant Coefficients $\Rightarrow$ Guess e^{rx}.
- Non-homogeneous $\Rightarrow$ Undetermined Coefficients.
- Linear $\Rightarrow$ Series Solutions.

60.30.3 Exact

Using the method for exact differential equations.

60.31 Linear Systems of Differential Equations

Let us start by looking at a system of first order homogeneous differential equations

$$y_1' = ay_1 + by_2$$
$$y_2' = cy_1 + dy_2$$

This can be written as

$$\frac{d}{dt} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$Ax = \vec{y}'$$

Then if we use the guess $y(t) = \vec{v}e^{\lambda t}$ and differentiate with respect to t we get

$$\lambda \vec{v} e^{\lambda t} = A\vec{v} e^{\lambda t}$$

$$\lambda \vec{v} = A\vec{v},$$

which is our *Eigenvalue* problem. This can be solved via

$$\det(A - \lambda I) = 0$$

$$\lambda^2 - (a + c)\lambda + ad - bc = 0$$

And this solution can lead to the three cases we already know: distinct real roots, complex Conjugate roots and repeated roots.

We can also use substitution to solve for one of the y_i, which are functions of x like

$$y_1' = ay_1 + by_2$$
$$y_2' = cy_1 + dy_2$$
$$by_2 = y_1' - ay_1$$
$$y_2 \frac{y_1' - ay_1}{b}$$

Then we can differentiate our first equation with respect to x

$$by_2 = y_1' - ay_1$$
$$by_2' = y_1'' - ay_1'$$
$$= b(cy_2 + dy_1)$$
$$= bcy_2 + bdy_1$$
$$= c(y_1' - ay_1) + bdy_1$$
$$y_1'' - (a + c)y_1' + (ac - bd)y_1 + bdy_1 = 0$$

Example:

We are given the system of differential equations:

$$\begin{cases} y' = y + z \\ z' = -2y + 3z \end{cases}$$

We write this system in matrix form:

$$u' = Au, \quad \text{where } u = \begin{bmatrix} y \\ z \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix}$$

To solve, we find the eigenvalues of A by solving the characteristic equation:

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 1 \\ -2 & 3 - \lambda \end{vmatrix} = (1 - \lambda)(3 - \lambda) + 2 = \lambda^2 - 4\lambda + 5 = 0$$

Solving the quadratic equation:

$$\lambda = \frac{4 \pm \sqrt{(-4)^2 - 4(1)(5)}}{2} = \frac{4 \pm \sqrt{-4}}{2} = \frac{4 \pm 2i}{2} = 2 \pm i$$

Now we find an eigenvector corresponding to $\lambda = 2 + i$ by solving:

$$(A - (2 + i)I)v = 0$$

$$A - (2 + i)I = \begin{bmatrix} 1 - (2 + i) & 1 \\ -2 & 3 - (2 + i) \end{bmatrix} = \begin{bmatrix} -1 - i & 1 \\ -2 & 1 - i \end{bmatrix}$$

From the first row of the system:

$$(-1 - i)v_1 + v_2 = 0 \Rightarrow v_2 = (1 + i)v_1$$

So, an eigenvector corresponding to $\lambda = 2 + i$ is:

$$v = \begin{bmatrix} 1 \\ 1 + i \end{bmatrix}$$

After repeating the previous computation for $\lambda = 2 - i$ we get the complex solution to the system:

$$u(t) = c_1 e^{(2+i)t} \begin{bmatrix} 1 \\ 1 + i \end{bmatrix} + c_2 e^{(2-i)t} \begin{bmatrix} 1 \\ 1 - i \end{bmatrix}$$

61 Probability

61.1 Basics

A probability experiment is considered ideal if and only the following properties:

- We know the possible outcomes.
- It can be repeated multiple times under the same conditions.
- It will be executed under strict conditions.

61.2 The Birthday Paradox

This is a small question that says: What is the probability of at least two people having their birthday on the same day in a group of x persons.

$$\mathbb{P}(x) = 1 - \prod_{n=0}^{x} \frac{365 - n}{365}$$

This formula gives us the probability of total minus all persons having it on different days. Which what we are looking for.

61.3 The Result Set

The set of all possible results of an experiment is denoted as Ω and its cardinality is the number of possible outcomes.

For experiments with Ω as the result set of one repetition then we define

$$\Omega = \Omega \times \cdots \times \Omega$$

As the *result set* of the experiment.

Finally, each subset of Ω is called an *outcome*, each subset with only one element in it is called an *elemental outcome* The set of all possible outcomes is $\mathcal{P}(\Omega)$

61.4 Laplace Distribution and Experiment

A *Laplace distribution* is a probability experiment with finite outcomes where each elemental outcome has the same probability of occurring.

$$\mathbb{P}(\omega) = \frac{\omega}{|\Omega|}$$

Then an experiment with this property is considered a *Laplace experiment*.

$$\mathbb{P}(E) = \frac{E}{|\Omega|}$$

61.5 Sigma Algebra

Given a non-empty set Ω and $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ a set of subsets of *Omega*. $\mathcal{A}$ is a σ-Algebra if and only if

  I. $\Omega \in \mathcal{A}$

  II. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$

  III. $A_1, A_2, A_3, \cdots \in \mathcal{A} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

61.5.1 The Power-set is a Sigma Algebra

This proof is pretty straight forward.

For the first axiom, we know that $\Omega \in \mathcal{P}$.

For the second axiom, if A is a subset of the power-set then $\Omega \setminus A$ is also int the power-set.

And lastly, for a sequence of sets in the power-set their, union is also in the power-set, because the power-set contains all subsets of Ω.

<div align="right">QED</div>

61.5.2 Properties of Sigma Algebras

If $\mathcal{A}$ is a sigma algebra then

- $\emptyset \in \mathcal{A}$
- $A_1, A_2, \cdots \in \mathcal{A} \implies \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$
- $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$

Proof:

For the first, because $\mathcal{A}$ is a subset of the power-set of Ω and the empty set is always a subset of a collection, this is trivial.

Given a number of sets in $\mathcal{A}$, using the laws of De Morgan we get

$$\bigcap_{n \in \mathbb{N}} A_n = \bigcup A_n^c$$

and we know that the complements are part of $mathcal A$ by definition.

For the last part, $A \setminus B = A \cap B^c$, because $B^c \in \mathcal{A}$ we can apply De Morgan $(A^c \cup (B^c)^c)^c = (A^c \cup B)^c$ Now because this is the union of two sets in the sigma algebra, and we know by definition that their union if also in the set.

<div align="right">QED</div>

61.6 Probability Measurement

Given a non-empty set Ω and $\mathcal{A} \subseteq \mathcal{P}$ a sigma algebra. We define the function $\mathbb{P} : \mathcal{A} \to \mathbb{R}$ as a *probability measurement* if

I. $\forall A \in \mathcal{A} : \mathbb{P}(A) \geq 0$

II. $\mathbb{P}(\Omega) = 1$

III. $A_n \in \mathcal{A}, n \in \mathbb{N}$ disjunctive in pairs $\implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$

Also, the triplet $\Omega, \mathcal{A}, \mathbb{P}$ is called a *probability space*.

61.6.1 Probability space of the Power-set

We want to prove that the probability space with the power-set of omega and $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ is in fact a probability space

- $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ is always greater or equal to zero.
- $\mathbb{P}(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$
- If all of our subsets in the power-set are pairwise disjunctive this mean that their intersection is the empty-set. In other wise there therefore, their independent of each other which means that their probabilities are also independent.

<div align="center">328</div>

61.6.2 Discrete Probability Spaces

A probability space $\Omega, \mathcal{P}(\Omega), \mathbb{P}$ is considered *discrete* if Ω is finite or infinite countable. In that case $\Omega = \{\omega_i : i \in I\}$ for $I = \{1, \ldots, n\}$

For $A \in \mathcal{A}$

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i \in I : \omega_i \in A} \{\omega_i\}\right) = \sum_{i \in I : \omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{i \in I : \omega_i \in A} p(\omega_i)$$

The function $\mathbb{P}(\{\omega_i\}) = p(\omega_i)$ is called a *probability function*

To test if specific function or distribution is actually a probability one we have to check.

- $\forall p(\omega) \geq 0 \mid \omega \in \Omega$
- $\sum_{\omega \in \Omega} p(\omega) = 1$

61.6.3 Properties of Probability-Measurements

Given $(\Omega, \mathcal{A}, \mathbb{P})$ and $A, B \in \mathcal{A}$

- $\mathbb{P}(\emptyset) = 0$
- $A_1, \ldots, A_n \in \mathcal{A}$ pairwise disjoint: $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$
- $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$
- $0 \leq \mathbb{P}(A) \leq 1$
- $A_1, \ldots, A_n \in \mathcal{A}$: $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$
- $A_1, A_2, \ldots \in \mathcal{A}$: $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

Proof:

- Define $A_n = \emptyset$ and by the properties of the probability spaces we know that $A_n \in \mathcal{A}, n \in \mathbb{N}$ disjunctive in pairs $\implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$

  Therefore, $\mathbb{P}(\emptyset) = \sum_{i=1}^{n} \mathbb{P}(A_i)$ and the only real number with the property $x + x = x$ is 0.

- By definition

- $B = (A \cap B) \cup B \setminus A$ and by the definition of the addition of probabilities

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) - \mathbb{P}(A \cap B)$$

- Using the addition property and by writing the union as $(A \setminus B) \cup (B \setminus A) \cup (A \cap B)$ we get that

$$\mathbb{P}(A \cup B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cup B)$$

  which results in $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

- $\Omega = A^c \cup A$ and $1 = \mathbb{P}(\Omega)$, thus by the addition $P(A) + P(A^c) = 1$

- Given that if $A = B$ then the equality holds the case where A is a subset of B comes from the fact that it is a strict subset of the other than its cardinality is smaller.

- By the definition that $\mathbb{P}(\Omega) = 1$ and the property $\mathbb{P}(\emptyset) = 0$ we are given an upper and lower bound for our probability distribution. Therefore, every subset of Ω will have a probability less or equal to it. For the lower bound the logic is that it has no other subset with fewer elements than it, but it is part of sets with more elements.

- Let us define $B_n = A_n \setminus \bigcap i = 1_{n-1} A_i$ which are pairwise disjoint.

$$\mathbb{P}\left(\bigcup A_n\right) = \mathbb{P}\left(\bigcup B_n\right) = \sum \mathbb{P}(B_i) \leq \sum \mathbb{P}(A_i)$$

Due to the fact that the B_i are pairwise disjoint.

- This is practically the same but for an infinite case.

$$\mathfrak{QED}$$

61.7 Inclusion-Exclusion Formula for Two Non-Disjoint Sets

Let $|A \cap B| = k$, then:

$$|A \cup B| = |A| + |B| - |A \cap B| = n + m - k \quad \text{(since we do not count the intersection twice)}$$

61.8 Inclusion-Exclusion Formula for Three Non-Disjoint Sets

$$\begin{aligned}
|A \cup B \cup C| &= |(A \cup B) \cup C| \\
&= |A \cup B| + |C| - |(A \cup B) \cap C| \\
&= |A| + |B| - |A \cap B| + |C| - |(A \cap C) \cup (B \cap C)| \\
&= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
\end{aligned}$$

61.9 General Formula for the Cardinality of the union of sets

$$\left|\bigcup_{i=1}^{n} M_i\right| = \sum_{I \subseteq \{1,\ldots,n\}, I \neq \emptyset} (-1)^{|I|-1} \left|\bigcap_{i \in I} M_i\right|$$

61.10 Stochastic Independence

Two results $A, B \in \mathcal{A}$ are considered *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Example:

When throwing two times a dice is the probability of throwing during the first throw an even number and when throwing the second dice the sum gives us an even number.

$$\Omega = \{1, \ldots, 6\}^2 \quad A = \{2, 4, 6\} \times \{1, \ldots, 6\} \quad B = \{2, 4, 6\} \cup \{1, 3, 5\}^2$$

Therefore,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{18}{36} = \frac{1}{2}$$

$$\mathbb{P}(B) = \frac{|B|}{|\Omega|} = \frac{18}{36} = \frac{1}{2}$$

And

$$\mathbb{P}(A \cap B) = \frac{|\{2, 4, 6\}^2|}{|\Omega|} = \frac{1}{4}$$

This is just $\mathbb{P}(A) \cdot \mathbb{P}(B)$. Therefore, they are independent.

$$\mathfrak{QED}$$

61.10.1 Independence for multiple results

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdots \mathbb{P}(A_n)$$

For all pairs A_i, A_j where $i \leq j$

61.10.2 Independence in Laplace Distribution

Give a Laplace distribution $\mathbb{P} = \frac{|A|}{|\Omega|}$ with two independent results A and B.

$$\frac{|A \cap B|}{|A|} = \frac{B}{|\Omega|}$$

This means that B is a probable in A as in the greater picture.

61.10.3 Complement Property

Given the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $A, B \in \mathcal{A}$ as two independent results. Then A^c and B are also independent.

Proof:

The sets $A \cap B$ and $A^c \cap B$ are disjoint and their union is

$$B = (A \cap B) \cup (A^c \cap B)$$

If we calculate the probability

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)$$

This gives us

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

Using the Independence of A and B

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B)(1 - \mathbb{P}(A))$$

$$\mathbb{P}(B) = \mathbb{P}(B)(1 - \mathbb{P}(A)) = \mathbb{P}(B) \cdot \mathbb{P}(A^c)$$

$$\mathfrak{QED}$$

61.10.4 Independence of Blocks

Given $A_1, \ldots, A_n$ independent events, divided into two blocks $A_1, \ldots, A_k$ and $A_{k+1}, \ldots, A_n$. Let B and C be events resulted of the union, intersection or complement of the first two blocks. Then this results are independent

61.11 The Product Space

Given two or more probability spaces with their respective tuples of two independent experiments. We define their combined probability space as the *product space* where

$$\Omega := \Omega_1 \times \cdots \times \Omega_n$$

and

$$p(\omega_1, \ldots, \omega_n) = p(\omega_1) \cdot \ldots \cdot p(\omega_n)$$

An example is throwing two dices at the same time.

61.12 The Product Space of Discrete Probability Spaces

Given the product space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ of the spaces $(\Omega_i, \mathcal{P}(\Omega_i), \mathbb{P}_i)$ and $A_i \in \mathcal{P}(\Omega_i)$. We define

$$
\begin{aligned}
A_i^{'} &:= \{\omega \in \Omega : \omega_i \in A_i\} \\
&= \{(w_1, \ldots, w_n) : \omega_j \in \Omega_j, \text{ for } j \neq i \text{ and } \omega_i \in A_i\} \\
&= \Omega_1 \times \cdots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \cdots \times \Omega_n
\end{aligned}
$$

For $i = 1, \ldots, n$. Then it holds that $\mathbb{P}(A_i^{'}) = \mathbb{P}_i(A_i)$ and the events $A_i^{'}$ are independent.

61.13 Bernoulli-Experiments

An experiment where $\Omega = \{0, 1\}, \mathbb{P}(\{1\}) = p$ and $\mathbb{P}(\{0\}) = 1 - p$ for $p \in (0, 1)$ is called a *Bernoulli-Experiment*.

61.13.1 Bernoulli Schema

If we perform a Bernoulli-Experiment n-times with a success probability of p. Where $\mathbb{P}(A_i) = p$. We define

$$C_k = \bigcup_{I \in \{1, \ldots, n\} \wedge |I| = k} \left(\bigcap_{i \in I} A_i \cap \bigcap_{i \in I} A_i^c \right)$$

For the probability we get

$$\mathbb{P}(C_k) = \sum_{I \in \{1, \ldots, n\} \wedge |I| = k} \mathbb{P}\left(\bigcap_{i \in I} A_i \cap \bigcap_{i \notin I} A_i^c \right) = \sum_{I \in \{1, \ldots, n\} \wedge |I| = k} \prod_{i \in I} \mathbb{P}(A_i) = \prod_{i \notin I} = \binom{n}{k} p^k (1-p)^{n-k}$$

61.14 Conditional Probability

Given the events A and B where $\mathbb{P}(A) > 0$. The *conditional probability* of B under the condition A is defined as

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Example:

In words, what is the probability of B if we limit ourself only to the case where A happens. Also what is the ratio between $A \cap B$ and A. From this definition we can also derive these two statements which are also equivalent.

$$\mathbb{P}(B \mid A) = \mathbb{P}(B) \iff A \text{ and } B \text{ are independent}$$

Proof:

$$
\begin{aligned}
\mathbb{P}(B \cap A) = \mathbb{P}(A)\mathbb{P}(B) &= \mathbb{P}(A \cap B)\mathbb{P}(A) \\
&= \mathbb{P}(B)
\end{aligned}
$$

$$\mathfrak{QED}$$

61.14.1 Conditional Probability of a Subset

Given the events A and B where $\mathbb{P}(A) > 0$. If $A \subset B$ then $\mathbb{P}(B \mid A) = 1$

Proof:

$A \subset B \implies A \cap B = A$ therefore

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$$

<div align="right">𝔔𝔈𝔇</div>

61.14.2 Probability Measurement of Conditional Probability

Given the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $A \in \mathcal{A}$ where $\mathbb{P}(A) > 0$. The map $B \to \mathbb{P}(B \mid A)$ is a probability measurement with the following properties.

- $\forall \in \mathcal{A} : \mathbb{P}(B \mid A) \geq 0$
- $\mathbb{P}(\Omega \mid A) = 1$
- $B_n \in A, n \in \mathbb{N}$ pairwise disjoint $\implies \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} B_n \mid A\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(B_n \mid A)$

We also get

$$\mathbb{P}(\emptyset \mid A) = 0$$

$$\mathbb{P}(B^c \mid A) = 1 - \mathbb{P}(B \mid A)$$

Proof:

- Because our sets do not have negative cardinalities and the lowest we can reach is the probability of the empty-set we can not get below 0.
- Similar to another proof we get $\frac{\mathbb{P}(A)}{\mathbb{P}(A)}$
- By the definition of probability space.

<div align="right">𝔔𝔈𝔇</div>

61.14.3 Multiplication Theorem

Given $A_1, dots, A_n$ events with $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) \neq 0$. Then

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \cdots \cdot (A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

61.14.4 Theorem of the Total Probability

Given our standard probability space, the disjoint partitions of Ω, $B_1, \ldots, B_n$ where $B_1 \cup \cdots \cup B_n = \Omega$, $B_i \cap B_j = \emptyset$ for $i \neq j$ and $\mathbb{P}(B_i) > 0$ then

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i) \forall A \in \mathcal{A}$$

Example:

You are given three bags

- Bag A has 2 red balls and 3 green balls
- Bag B has 3 red balls and 2 green balls
- Bag C has 1 red ball and 4 green balls

What is the probability of a ball being red?

The answer is given by

$$\mathbb{P}(R) = \mathbb{P}(R \cap A) + \mathbb{P}(R \cap B) + \mathbb{P}(R \cap C)$$

Which gives us

$$\mathbb{P}(R) = \mathbb{P}(R \mid A)\mathbb{P}(A) + \mathbb{P}(R \mid B)\mathbb{P}(B) + \mathbb{P}(R \mid C)\mathbb{P}(C)$$

Finally we plug in the individual numbers

$$\mathbb{P}(R) = \frac{1}{3}\left(\frac{2}{5} + \frac{3}{2} + \frac{1}{4}\right) = \frac{6}{15}$$

61.15   Multistage Random Experiments

Let $\Omega_1, \ldots, \Omega_n$ be discrete sample spaces of n sub-experiments, and let $\Omega = \Omega_1 \times \cdots \times \Omega_n$. Furthermore, let an initial distribution be given by the probability function p_1 with

$$\sum_{\omega_1 \in \Omega_1} p_1(\omega_1) = 1$$

and transition probabilities p_j for $j = 2, \ldots, n$ satisfying

$$\sum_{\omega_j \in \Omega_j} p_j(\omega_j \mid \omega_1, \ldots, \omega_{j-1}) = 1,$$

for $(\omega_1, \ldots, \omega_{j-1}) \in \Omega_1 \times \cdots \times \Omega_{j-1}$.

This is used in the case when we model our experiment as a tree of probabilities.

For the multistage random experiment, the following two *path rules* apply:

1. Let $(\omega_1, \ldots, \omega_n) \in \Omega$. Then we have:

$$p(\omega_1, \ldots, \omega_n) = p_1(\omega_1) \cdot p_2(\omega_2 \mid \omega_1) \cdot \ldots \cdot p_n(\omega_n \mid \omega_1, \ldots, \omega_{n-1})$$

2. For $A \subset \Omega$, it holds that
$$\mathbb{P}(A) = \sum_{(\omega_1, \ldots, \omega_n) \in A} p(\omega_1, \ldots, \omega_n)$$

Note that for independent sub-experiments, we obtain the product space with

$$p(\omega_1, \ldots, \omega_n) = p_1(\omega_1) \cdot p_2(\omega_2) \cdot \ldots \cdot p_n(\omega_n).$$

61.16   Bayes Theorem

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Furthermore, let $B_1, \ldots, B_n$ be a disjoint partition of the sample space Ω with $\mathbb{P}(B_i) > 0$ for all $i = 1, \ldots, n$. Then, for every event $A \in \mathcal{A}$, the following holds:

$$\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i)}{\sum_{k=1}^{n} \mathbb{P}(A \mid B_k)\mathbb{P}(B_k)}.$$

In words and by changing A with E of evidence and B with H of hypothesis, we want to know how much the actual evidence changes our initial belief. The process can be described as: you have a hypothesis, you have observed some evidence and finally, you want to know the probability that your hypothesis is true given the observed evidence.

$$\mathbb{P}(H \mid E) = \frac{\mathbb{P}(E \mid H)\mathbb{P}(H)}{\mathbb{P}(H)\mathbb{P}(E \mid H) + \mathbb{P}(\neg H)\mathbb{P}(E \mid \neg H)}$$

- $\mathbb{P}(H)$ is called the *prior*. Our initial hypothesis, without seeing the evidence.

- $\mathbb{P}(E \mid H)$ is called the *likelihood*. This is the probability that we will see our hypothesis reflected on the evidence given that our hypothesis is true.

- $\mathbb{P}(E \mid \neg H)$ represents the probability that our hypothesis will be reflected on the sample given our hypothesis is false.

With this knowledge, breaking down the formula is very intuitive

- $\mathbb{P}(H \mid E)$ is the *posterior*. It describes the following question: How often is H true among the cases where H is true.probability

- $\mathbb{P}(H)\mathbb{P}(E \mid H)$ is the proportion of our hypothesis in the sample being true times the actual proportion of the fitting samples inside our hypothesis.

- $\mathbb{P}(H)\mathbb{P}(E \mid H) + \mathbb{P}(\neg H)\mathbb{P}(E \mid \neg H)$ consist of our expression above plus the proportion of the counter hypothesis times the proportion of the representatives of our hypothesis which will show up in the evidence in the counter group.

Example:

A disease affects 1 percent of the population and you were tested positive with a test which has a 99.9 percent accuracy. What is the probability for you truly being infected?

$$\mathbb{P}(I \mid +) = \frac{\mathbb{P}(I)\mathbb{P}(+ \mid I)}{\mathbb{P}(I)\mathbb{P}(+ \mid I) + \mathbb{P}(\neg I)\mathbb{P}(+ \mid \neg I)}$$

Our numbers are:

- $\mathbb{P}(I) = 0.001$

- $\mathbb{P}(+ \mid I) = 0.99$

- $\mathbb{P}(\neg I) = 0.999$

- $\mathbb{P}(+ \mid \neg I) = 0.1$

Thus, our formula gives us a 9 percent chance of actually being infected.

61.17   Random Variables

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we call a function $f : \Omega \to \mathbb{R}$ *measurable* if for all $z \in \mathbb{R}$ the following applies

$$f^{-1}((-\infty, z]) = \{\omega \in \Omega : f(\omega) \le z\} \in \mathcal{A}$$

This measurable function $X : \Omega \to \mathbb{R}$ is a *random variable*.

Note that for discrete probability spaces $\{\omega : X(\omega) \le z\} \subseteq \Omega \in \mathcal{P}$

Also, it considered discrete if it has only a finite or an infinite countable number of values.

Finally, we define the probability for random variables as

$$\mathbb{P}(X = k) := \mathbb{P}(\{X = k\})$$

61.17.1 Operations with Random Variables

Given two R.V. X,Y and some number $\alpha \in \mathbb{R}$, then

- $X + Y$
- $X - Y$
- $X \cdot Y$
- $\max(X, Y)$
- $\min(X, Y)$
- αX or αY

are also R.V.

61.18 Indicator Function

Given $A \subseteq \Omega$, we define the *indicator function* as

$$\mathrm{ind}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

We also the get the following rules:

- $\mathrm{ind}_\emptyset = 0$
- $\mathrm{ind}_A^2 = \mathrm{ind}_A$
- $\mathrm{ind}_{A^c} = 1 - \mathrm{ind}_A$
- $\mathrm{ind}_{A \cap B} = \mathrm{ind}_A \mathrm{ind}_B$
- $\mathrm{ind}_{A \cup B} = \mathrm{ind}_A + \mathrm{ind}_B - \mathrm{ind}_{A \cap B}$
- $\mathrm{ind}_{A+B} = \mathrm{ind}_A + \mathrm{ind}_B$

61.18.1 Indicator Sum

Given the events $A_1, \ldots A_n \in \Omega$. We define the random variable

$$X := \sum_{j=1}^{n} \mathrm{ind} - A_j$$

as the *indicator sum* or *counting variable*.

The realization of X gives us how much A_j have occurred.

$$\{X = k\} = \sum_{T \subseteq \{1,\ldots,n\}:|T|=k} \left(\bigcap_{j \in T} A_j \cap \bigcap_{j \notin T} A_j^c \right), k \in \{0, 1, \ldots, n\}$$

61.19 Distributions

Given our probability space and a R.V. X. The probability measurement $\mathbb{P}_X$ to $\mathbb{R}$ defined by

$$\mathbb{P}_X(A) := \mathbb{P}(\{\omega : X(w) \in A\}), \text{ for } A \subseteq \mathbb{R}$$

is called as the *distribution* of X. Also we write $\mathbb{P}(X \in A)$ as a abbreviation for $\mathbb{P}(\{\omega : X(w) \in A\})$.

61.20 Distribution Function

Given a probability space, a probability variable and a distribution $\mathbb{P}_X$ we define the function

$$\mathbb{F}_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(\{\omega : X(w) \le x\})$$

as the *distribution function*.

61.20.1 Properties of Distribution Functions

- F_X is monotonic growing
- It is right-sided continuous
- $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X = 1$

61.21 Formulas for Distributions

- $\mathbb{P}(X = a) = \mathbb{P}(X = a)$
- $\mathbb{P}(X \le a) = \mathbb{P}(X \le a)$
- $\mathbb{P}(X < a) = \mathbb{P}(X \le a - 1)$
- $\mathbb{P}(X > a) = 1 - \mathbb{P}(X \le a)$
- $\mathbb{P}(X \ge a) = 1 - \mathbb{P}(X \le -1)$
- $\mathbb{P}(a \le X \le b) = \mathbb{P}(X \le b) - \mathbb{P}(X \le a)$

61.22 Uniform Distribution

A random variable $X \in \{1, 2, \ldots, n\}$ with $\mathbb{P}(X = i) = \frac{1}{n}$ for $i = 1, \ldots, n$ is uniformly distributed. The distribution $\mathbb{P}_X$ is called the (discrete) *uniform distribution*.

$$\frac{1}{n}$$

The expected value is given by

$$\mathbb{E}(X) = \frac{n + 1}{2}$$

and the standard deviation is given by

$$\sigma = \frac{n^2 - 1}{12}$$

where a and b are the boundaries of our distribution.

61.23 Bernoulli Distribution

Distribution for binary events with only one trial.

$$\mathbb{P}(X = k) = p^k (1 - p)^{1-k}, \quad k \in \{0, 1\}$$

61.24 Binomial Distribution

Distribution for discrete binary experiments. Either an event occurs or it does not.

Formulas:

- *Probability:* $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

- *Expected Value:* $\mathbb{E}(X) = np$

- *Variance:* $\text{var}(X) = \mathbb{E}(x)(1-p)$

- *Standard Deviation:* $\sigma = \sqrt{\mathbb{E}(X)(1-p)}$

We also get the standard deviation and our distribution function

$$\sigma = \sqrt{\frac{(\bar{x} - x_1)^2 + \cdots + (\bar{x} - x_n)^2}{n}}$$

$$\mathbb{F}_X = \sum_{i=P(X=k)}^{P(X=n)(P(X=i))}$$

61.24.1 Sigma Rules

- $P(\mu - \sigma \leq x\mu + \sigma) \approx 68,3\%$

- $P(\mu - 2\sigma \leq x\mu + 2\sigma) \approx 95,4\%$

- $P(\mu - 3\sigma \leq x\mu + 3\sigma) \approx 99,7\%$

61.25 Geometric Distribution

A random variable $X \in \mathbb{N}_0$ with $\mathbb{P}(X = n) = p(1-p)^n$ for some $p \in (0,1)$ and $n \in \mathbb{N}_0$ is geometrically distributed. The distribution $\mathbb{P}_X$ is called the *geometric distribution*.

Similar to the binomial distribution, the *geometric* distribution is used for an experiment with two outcomes and we we want to know the probability of $k-1$ non-successes until a success happens

$$\mathbb{P}(X = k) = (1-p)^{k-1} p$$

We also get

$$\mathbb{E}(X) = \mu = \frac{1-p}{p}$$

$$\text{Var} = \frac{1-p}{p^2}$$

and

$$\sigma = \sqrt{\text{Var}(X)}$$

61.26 Poisson Distribution

A random variable $X \in \mathbb{N}_0$ with

$$\mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad \text{for some } \lambda > 0,$$

We also define the distribution function

$$\mathbb{F}_X(n) = \sum_{i=0}^{n} \frac{\lambda^i}{i!} i^{-\lambda}$$

is Poisson distributed. The distribution $\mathbb{P}_X$ is called the *Poisson distribution*. It simple words, it is the number of events x, in an specific interval.

We also get our expected result and variance

$$\mathrm{E}(X) = \lambda$$

and

$$\mathrm{Var}(X) = \lambda$$

61.27   Exponential Distribution

We define the *exponential distribution* as a real defined random variable, used mostly for the duration of time intervals.

$$\mathbb{P}(X = n) = \begin{cases} \lambda e^{-\lambda n} & , n \geq 0 \\ 0 & , \text{ else} \end{cases}$$

We also define the distribution function

$$\mathbb{F}_X(X = n) = \int_{-\infty}^{n} \mathbb{P}(X = n) = \begin{cases} 1 - e^{-\lambda n} & , n \geq 0 \\ 0 & , \text{ else} \end{cases}$$

We also get our expected result and variance

$$\mathrm{E}(X) = \frac{1}{\lambda}$$

and

$$\mathrm{Var}(X) = \frac{1}{\lambda^2}$$

61.28   Normal Distribution

- *Probability:* $\mathbb{P}(X = x) \dfrac{1}{\sqrt{2\pi\sigma^2} e^{\frac{1}{2}\left(\frac{x-u}{\sigma}\right)^2}}$
- *Expected Value:* $\mathrm{E}(X) = np = \mu$
- *Variance:* $\mathrm{Var}(X) = \mathrm{E}(X)(1 - p)$
- *Standard Deviation:* $\sqrt{\mathrm{Var}(X)}$

61.29   Hyper-geometric Distribution

Similar to the binomial distribution, but we keep track of the changes after each choice. In other words, binomial distribution without repetition.

$$\mathbb{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

Nomenclature:

- N: Total number of elements

- M: Elements with the trait A

- $N - M$: Elements without the trait A

- k: Number to elements to take

- n: Number of iterations

Each part of the formula has a different meaning. $\binom{N}{n}$ represents the total number of ways to choose the elements from the total. $\binom{M}{k}$ gives us the ways of choose k elements with the trait from its total M and finally, $\binom{N-M}{n-k}$ represents the remaining ways of choosing elements we are not interested in.

Formulas:

- *Expected Values:* $\mathrm{E}(x) = n\frac{M}{N}$

- *Variance:* $\mathrm{Var}(X) = \mathrm{E}(x)\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$

61.30   The Expected Value

Given $\Omega, \mathcal{A}, \mathbb{P}$, a probability space and X a discrete random variable. We define the *expected result* as

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x)$$

which exists only if

$$\sum_{x \in X(\Omega)} |x|\mathbb{P}(X = x) < \infty$$

Where x is an event $\mathbb{P}(X = x)$ its respective probability. It represents the mean of the random variable.

61.30.1   Transformation of the Expected Value

Given $u : X(\Omega) \to \mathbb{R}$ where $\sum_{x \in X(\Omega)} |u(x)|p(x) < \infty$ then

$$\mathbb{E}[u(X)] = \sum_{x \in X(\Omega)} u(x)\mathbb{P}(X = x)$$

This represents the expectation of the function.

61.31   Triangle Inequality for Expected Values

$$|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$$

Proof:

$$|\mathbb{E}[X]| = \left|\sum_{x \in X(\Omega)} x\mathbb{P}(X = x)\right| \leq \sum_{x \in X(\Omega)} |x\mathbb{P}(X = x)|$$
$$= \sum_{x \in X(\Omega)} |x|\mathbb{P}(X = x) = \mathbb{E}[|X|]$$

$$\mathfrak{QED}$$

61.31.1 Properties of Expected Values

Given two random variables X and Y, if both have an expected results then

- $\mathbb{E}[\alpha X] = \alpha \mathbb{E}[X]$
- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[b] = b$
- $X \leq Y = \mathbb{E}[X] \leq \mathbb{E}[Y]$

61.32 Variance

Given a random variable X for which $\mathbb{E}[(X - \mathbb{E})^2]$ exists. The we define the *variance* of X as

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E})^2] = \sum_{x \in X}(x - \mathbb{E}^2)\mathbb{P}(X = x)$$

And its root is the *standard deviation.*

61.32.1 Properties of the Variance

Let X be a (discrete) random variable for which the variance exists. Then the following hold:

- $\text{Var}(aX + b) = a^2 \text{var}(X)$, for $a, b \in \mathbb{R}$
- $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- For $a \in \mathbb{R}$, we have $\mathbb{E}[(X - a)^2] = \text{var}(X) + (\mathbb{E}[X] - a)^2$, and in particular, $\mathbb{E}[(X - a)^2] \geq \text{var}(X)$. Equality holds if and only if $a = \mathbb{E}[X]$.

Example:

Given $\mathbb{E}(X) = 1.2$ and $\mathbb{P}(X = 0) = 0.16, \quad \mathbb{P}(X = 1) = 0.48, \mathbb{P}(X = 2) = 0.36$.

$$\mathbb{E}[X^2] = 0^2(0.16) + 1^2(0.48) + 2^2(0.36) = 1.92$$

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E})^2] = (0 - 1.2)^2(0.16) + (1 - 1.2)^2(0.48) + (2 - 1.2)^2(0.36)$$

61.33 Moments

For a random variable X, we define $k.$ *moment* as $\mathbb{E}[X^k]$

61.33.1 Markov Inequality

Given a random variable X, $a \in \mathbb{R}$ and $h : D \to [0, \infty)$ a monotonic growing function with the properties $h(a) > 0$ and $X(\Omega) \subseteq D$ then

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[h(X)]}{h(a)}$$

There is a special case for $Y = |X|, a > 0$ and $h(x) = x$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[h(X)]}{a}$$

61.33.2 Chebychev Inequality

Given a random variable X with a variance and $a > 0$ then

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\text{var}(X)}{a^2}$$

62 Combinatorics

62.1 Three Principles of Combinatorics

These principles are *Addition*, *Multiplication* and *Inclusion-Exclusion*.

62.2 Permutation

Number of ways of ordering n distinct elements. In other words the number of tuples of n distinct elements.

$$n!$$

62.3 Permutation with repetition

Number of ways of ordering n non-distinct elements. Here $m_1, \ldots, m_n$ is the number a specific item is repeated in the original set. By dividing by the respective amount of the repeated non-distinct element, we eliminate the permutations, where the identical element swap their place.

$$\frac{n!}{m_1! \cdots m_n!}$$

62.4 Variation

Ways of put n objects in k slots with repetition.

$$n^k$$

62.5 Variation without repetition

Ways of put n objects in k slots without repetition. Number of tuples with k elements which can be selected from n elements, without repetition.

$$\frac{n!}{(n-k)!}$$

62.6 Combination without repetition

Ways of choose k objects of n elements.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Here $n!$ is the number of permutations of the original set, the tuples. $(n-k)!$ has the function of eliminating the permutations with more elements we are not interested in, and $k!$ has the function to eliminate the duplicates because for a combination the order does not matter, contrary to the permutations. It is chosen that way because for each choice of k objects there are $k!$ possible orderings, and we only want one.

62.7 Combination with repetition

This focuses not on the slots but in the separations between the slots. More specific the number of ways to distribute k identical objects into n identical boxes.

$$\binom{k+n-1}{n-1} = \binom{k+n-1}{k} = \frac{(k+n-1)!}{k!(n-1)!}$$

This comes from modeling our possible combinations. If we have a tuple $(a_1, a_2, \ldots, a_n)$ where a_i gives us the number of occurrences of the object i, then we can imagine using $|$ a number of times for the occurrences of each object and $\cdot$ as the separator between each set of $|$. For each possible combination we have k lines, $k + n - 1$ positions, and $n - 1$ points. Therefore, the question becomes: How many ways are there of putting k lines into $k + n - 1$ positions.

62.8   Multinomial Coefficients

When using the binomial coefficient repeatedly one after the other we can abbreviate this consecutive uses as

$$\binom{n}{k_1, \ldots, k_n}$$

This is like from n choose k_1 then from $n - k_1$ choose k_2 and so on.

62.8.1   Multinomial Theorem

Given $x_1, \ldots, x_p \in \mathbb{R}$ and $\mathbb{N}$ then

$$(x_1 + \cdots, x_n)^n = \sum_{k_1 + \cdots + k_{p=n}} \binom{n}{k_1, \ldots, k_p} x_1^{k_1} \cdots x_p^{k_p}$$

62.9   Disarray

Number of permutation in which no object ends in the same initial spot. Also, know as the *sub-factorial*

$$!n = n! \sum_{k=0}^{n} \left(\frac{(-1)^k}{k!} \right)$$

62.10   Bell Numbers

Number of ways of grouping n objects in an arbitrary number of slots.

$$B(N) = \sum_{k=0}^{N-1} \binom{N}{K} B(K)$$

with $B(1) = 1$ and $B(2) = 2$

62.11   Ramanujan Numbers

The same as the Bell Numbers but, all objects are equal. As an example the number of ways to decompose a number.

$$R(N) \approx \frac{1}{4N\sqrt{3}} e^{2\pi \sqrt{\frac{N}{6}}}$$

62.12   Stirling Numbers I

Ways of permuting N items with K exchanges.

$$\begin{bmatrix} N+1 \\ K \end{bmatrix} = N \begin{bmatrix} N \\ K \end{bmatrix} + \begin{bmatrix} N \\ K-1 \end{bmatrix}$$

62.13   Stirling Numbers II

Ways of grouping n items in k slots.

$$\left\{{N \atop K}\right\} = \frac{1}{K!}\sum_{j=0}^{K}(-1)^{K-j}\binom{K}{j}j^N$$

62.14   Lah Numbers

Ways of building K lists with a set N elements.

$$\left\lfloor{N \atop K}\right\rfloor = \binom{N-1}{K-1}\frac{N!}{K!} = \sum_{j=0}^{K}\left[{N \atop K}\right]\left\{{N \atop K}\right\}$$

62.15   Euler's Numbers

Number of permutations of a set of N elements of different size in which K elements are bigger than their previous element.

$$\left\langle{N \atop K}\right\rangle = \sum_{j=0}^{K}(-1)^j\binom{N+1}{j}(K+1-j)^N$$

62.16   Catalan Numbers

These have different applications like the number of ways of the triangulations of a polygon or the number of trees with N leafs.

$$\mathfrak{C}(N) = \binom{2N}{N} - \binom{2N}{N+1}$$

or

$$\mathfrak{C}(N) = \sum_{k=1}^{n}(k-1)\mathfrak{C}(N-k)$$

62.17   Pascals Triangle

This is a triangle build by the formula $\binom{n+1}{k+1} = \binom{n}{k}\binom{n}{k+1}$

$n=0$							1						
$n=1$						1		1					
$n=2$					1		2		1				
$n=3$				1		3		3		1			
$n=4$			1		4		6		4		1		
$n=5$		1		5		10		10		5		1	
$n=6$	1		6		15		20		15		6		1

Here is the triangle but with the Binomial Coefficients

$n=0$				$\binom{0}{0}$		
$n=1$			$\binom{1}{0}$		$\binom{1}{1}$	
$n=2$		$\binom{2}{0}$		$\binom{2}{1}$		$\binom{2}{2}$
$n=3$	$\binom{3}{0}$		$\binom{3}{1}$		$\binom{3}{2}$		$\binom{3}{3}$

It can be used with the Binomial Theorem to get the Coefficients of a polynomial of degree n.

Example:

$$(a + b)^2 = 1a^2b^0 + 2a^1b^1 + 1a^0b^2$$

or

$$(a - b)^2 = 1a^2b^0 - 2a^1b^1 + 1a^0b^2$$