Mathematics Bootcamp Lecture Notes

Department of Statistical Sciences, University of Toronto

Emma Kroell

Last updated: July 15, 2022

Preface

These lecture notes were prepared for the Mathematics course at the inaugural Department of Statistical Sciences Graduate Student Bootcamp at the University of Toronto. The course teaches an overview of necessary mathematics prerequisites to incoming statistics graduate students, with an emphasis on proofs.

These lectures are based on the following books or lecture notes:

1. *An Introduction to Mathematical Structures and Proofs* by Larry J. Gerstein
2. *The Tools of Mathematical Reasoning* by Tamara J. Lakins
3. *A Taste of Topology* by Volker Runde
4. *Real Mathematical Analysis* by Charles C. Pugh
5. *Linear Algebra Done Right* by Sheldon Axler
6. *Linear Algebra Done Wrong* by Sergei Treil
7. *Lecture notes in Mathematics for Economics and Statistics* by Piotr Zwiernik
8. *Real Analysis Lecture Notes* by Laurent Marcoux

Chapter 1 of Gerstein (2012) and the first three chapters of Lakins (2016) are used as references for the proof technique section. Runde (2005) is the main text for the sections on set theory, metric spaces, and topology, which follow chapters 1, 2, and 3 of his book, respectively. Some additional topics come from Pugh (2015). The linear algebra content comes mostly from Axler (2015), with Treil (2017) used in some sections for an alternate perspective.

Most of the material in these notes belongs to these texts. All of these texts are available online to University of Toronto users (some to everyone).

I would like to acknowledge the assistance of Jesse Gronsbell, Stanislav Volgushev, Piotr Zwiernik, and Robert Zimmerman in developing the list of topics for the course.

Please notify me of any typos or corrections at emma.kroell@mail.utoronto.ca .

Contents

A short note on notation:

$\mathbb{N}$ denotes the whole numbers, i.e. $\mathbb{N} = \{1, 2, \ldots\}$

$\mathbb{N}_0$ denotes the non-negative integers, i.e. $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$

$\mathbb{Z}$ denotes the integers, i.e. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$

$\mathbb{Q}$ denotes the rational numbers, i.e. $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z} \text{ and } q \neq 0\}$

$\mathbb{R}$ denotes the real numbers

$\mathbb{C}$ denotes the complex numbers

1 Review of proof techniques

1.1 Propositional logic

Propositions are statements that could be true or false. They have a corresponding *truth value*. We will use capital letters to denote propositions. For example, "n is odd" and "n is divisible by 2" are propositions. Let's call them P and Q. Whether they are true or not (i.e. their truth value) depends on what n is.

We can negate statements: $\neg P$ is the statement "n is not odd".

We can combine statements:

- $P \wedge Q$ is the statement "n is odd and n is divisible by 2".
- $P \vee Q$ is the statement "n is odd or n is divisible by 2".

We always assume the inclusive or unless specifically stated otherwise.

Example 1.1 *Here are some statements, which we want to write in propositional logic.*

- *If it's not raining, I won't bring my umbrella.*
- *I'm a banana or Toronto is in Canada.*
- *If I pass this exam, I'll be both happy and surprised.*

For the first one, let A be the statement "it's raining" and B be the statement "I will bring my umbrella". In logic, the statement is $\neg A \implies \neg B$.

For the second, let C be the statement "I'm a banana" and let D be the statement "Toronto is in Canada". We write this as $C \vee D$.

For the third, let P be the statement "I pass this exam", let Q be the statement "I am happy", and let R be the statement "I am surprised". This one is written $P \implies (Q \wedge R)$.

1.1.1 Truth values

Example 1.2 *Write the following using propositional logic:*
If it is snowing, then it is cold out.
It is snowing.
Therefore, it is cold out.

Solution.
$P \implies Q$
P
Conclusion: Q

To examine if a statement is true or not, we use a truth table, where we write out all the possibilities.

Example 1.3 *The truth table for $P \implies Q$ where P, Q are propositions is:*

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

1.1.2 Logical equivalence

We say that two statements are *logically equivalent* if they have the same truth tables.

Example 1.4 *Let P, Q be propositions. $P \implies Q$ is logically equivalent to $\neg P \vee Q$.*

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

P	Q	$\neg P$	$\neg P \vee Q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

Theorem 1.5 (De Morgan's Laws) *Let P, Q be propositions.*

(i) $\neg(P \wedge Q)$ is logically equivalent to $\neg P \vee \neg Q$.

(ii) $\neg(P \vee Q)$ is logically equivalent to $\neg P \wedge \neg Q$.

Proving this is your first exercise.

The following fact is often useful.

Example 1.6 *$\neg(P \implies Q)$ is logically equivalent to $P \wedge \neg Q$. This follows from Example 1.4 and Theorem 1.5.*

1.1.3 Quantifiers

There are two important logical operators that we have not yet discussed. They are denoted using the following symbols: $\forall$, read as "for all" or "for each", and $\exists$, read as "there exists". We will explore their meanings, how they can help us simplify statements we need to prove, and how we prove such statements.

For all

"for all", $\forall$, is also called the universal quantifier. If $P(x)$ is some property that applies to x from some domain, then $\forall x P(x)$ means that the property P holds for every x in the domain. An example is the statement "Every real number has a non-negative square." We write this as $\forall x \in \mathbb{R}$, $x^2 \geq 0$. In logic, people often use brackets to separate parts of the logical expression, ex. $(\forall x \in \mathbb{R})(x^2 \geq 0)$.

How do we prove a for all statement? We need to take an arbitrary element of the domain, and show the property holds for that element.

There exists

"there exists", $\exists$, is also called the existential quantifier. If $P(x)$ is some property that applies to x from some domain, then $\exists x P(x)$ means that the property P holds for some x in the domain. An example is the statement that 4 has a square root in the reals. We write this as $\exists x \in \mathbb{R}$ such that $x^2 = 4$ or in proper logic notation, $(\exists x \in \mathbb{R})$ $(x^2 = 4)$.

How do we prove a there exists statement? We need to find an element in the domain for which the property holds (find an example).

There is also a special way of writing when there exists a unique element. We use $\exists!$ for this case. For example, the statement "there exists a unique positive integer such that the integer squared is 64" is written $\exists! z \in \mathbb{N}$ such that $z^2 = 64$.

Combining quantifiers

Often we will need to prove statements where we combine quantifiers. Here are some examples:

Statement	Logical expression
Every non-zero rational number has a multiplicative inverse	$\forall q \in \mathbb{Q} \setminus \{0\}$, $\exists s \in \mathbb{Q}$ such that $qs = 1$
Each integer has a unique additive inverse	$\forall x \in \mathbb{Z}\,, \exists! y \in \mathbb{Z}$ such that $x + y = 0$
$f : \mathbb{R} \to \mathbb{R}$ is continuous at $x_0 \in \mathbb{R}$	$\forall \epsilon > 0\ \exists \delta > 0$ such that whenever $\lvert x - x_0 \rvert < \delta$, $\lvert f(x) - f(x_0) \rvert < \epsilon$

The order of quantifiers is important! Changing the order changes the meaning. Consider the following example. Which are true? Which are false?

$$\begin{gathered}\forall x \in \mathbb{R}\, \forall y \in \mathbb{R}\ x + y = 2 \\ \forall x \in \mathbb{R}\, \exists y \in \mathbb{R}\ x + y = 2 \\ \exists x \in \mathbb{R}\, \forall y \in \mathbb{R}\ x + y = 2 \\ \exists x \in \mathbb{R}\, \exists y \in \mathbb{R}\ x + y = 2\end{gathered}$$

It's also important to know how to negate logical statements that include quantifiers, as it will often help us prove or disprove the statements. The results are intuitive, but things can get complicated when we have more complex statements. The negation of a statement being true for all x is that is isn't true for at least one x. The negation of a statement being true for at least one x is that is isn't true for any x.

In summary,

$$\begin{aligned}\neg \forall x P(x) &= \exists x (\neg P(x)) \\ \neg \exists x P(x) &= \forall x (\neg P(x))\end{aligned}$$

The negations of the statements above are:

Logical expression	Negation
$\forall q \in \mathbb{Q} \setminus \{0\}$, $\exists s \in \mathbb{Q}$ such that $qs = 1$	$\exists q \in \mathbb{Q} \setminus \{0\}$ such that $\forall s \in \mathbb{Q}$, $qs \neq 1$
$\forall x \in \mathbb{Z}\,, \exists! y \in \mathbb{Z}$ such that $x + y = 0$	$\exists x \in \mathbb{Z}$ such that $(\forall y \in \mathbb{Z}, x + y \neq 0) \vee (\exists y_1, y_2 \in \mathbb{Z}$ such that $y_1 \neq y_2 \wedge x + y_1 = 0 \wedge x + y_2 = 0\ )$
$\forall \epsilon > 0\ \exists \delta > 0$ such that whenever $\lvert x - x_0 \rvert < \delta$, $\lvert f(x) - f(x_0) \rvert < \epsilon$	$\exists \epsilon > 0$ such that $\forall \delta > 0$, $\lvert x - x_0 \rvert < \delta$ and $\lvert f(x) - f(x_0) \rvert \geq \epsilon$

Note that we use De Morgan's laws (Theorem 1.5), as well as the negation of an implication (Example 1.6). What do these negations mean in English?

1.2 Types of proof

1.2.1 Direct proof

In a direct proof, our approach is to use the definition and known results.

Example 1.7 *The product of an even number with another integer is even.*

To prove this statement, we will use the definition of even. First we state that definition.

Definition 1.8 *We say that an integer n is* even *if there exists another integer j such that $n = 2j$. We say that an integer n is* odd *if there exists another integer j such that $n = 2j + 1$.*

Now we prove the example directly.

Proof. Let $n, m \in \mathbb{Z}$, with n even. By definition, there $\exists\, j \in \mathbb{Z}$ such that $n = 2j$. Then

$$nm = (2j)m = 2(jm)$$

Therefore nm is even by definition. □

Here is another example, which uses the concept of divisibility.

Definition 1.9 *Let $a, b \in \mathbb{Z}$. We say that "a divides b", written $a|b$, if the remainder is zero when b is divided by a, i.e. $\exists j \in \mathbb{Z}$ such that $b = aj$.*

Example 1.10 *Let $a, b, c \in \mathbb{Z}$ with $a \neq 0$. Prove that if $a|b$ and $b|c$, then $a|c$.*

Proof. Let $a, b, c \in \mathbb{Z}$. Suppose $a|b$ and $b|c$. Then by definition, there exists $j, k \in \mathbb{Z}$ such that $b = aj$ and $c = kb$. Combining these two equations gives $c = k(aj) = a(kj)$. Thus $a|c$ by definition. □

1.2.2 Proof by contrapositive

Sometimes instead of proving an implication $P \implies Q$ directly, it is easier to prove $\neg Q \implies \neg P$. This is called the contrapositive. First, we show that these two statements are logically equivalent using truth tables.

$P \implies Q$

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

$\neg Q \implies \neg P$

P	Q	$\neg P$	$\neg Q$	$\neg Q \implies \neg P$
T	T	F	F	T
T	F	F	T	T
F	T	T	F	F
F	F	T	T	T

Note that $\neg P \implies \neg Q$ is *not* logically equivalent to $P \implies Q$ (can you think of an example?). This is a common mistake.

Here is an example of a statement that is easier to prove using the contrapositive as opposed to directly.

Example 1.11 *If an integer squared is even, then the integer is itself even.*

Proof. We prove the contrapositive. Let n be odd. Then there exists $k \in \mathbb{Z}$ such that $n = 2k + 1$. We compute

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

Thus n^2 is odd. □

1.2.3 Proof by contradiction

Another proof technique is to assume something we know to be (or think to be) false, and then try to derive a contradiction. A contradiction is something that is impossible, like 0=1 or showing that a number is both odd and even.

In sum, to prove that a statement P is true by contradiction, we assume $\neg P$ is true, derive a contradiction, and conclude that P is true. Here is an example.

Example 1.12 *The sum of a rational number and an irrational number is irrational.*

Proof. Let $q \in \mathbb{Q}$ and $r \in \mathbb{R} \setminus \mathbb{Q}$. Suppose in order to derive a contradiction that their sum is rational, i.e. $r + q = s$ where $s \in \mathbb{Q}$. But then $r = s - q \in \mathbb{Q}$. Contradiction. Therefore the sum of a rational number and an irrational number is irrational. □

1.2.4 Summary

In sum, to prove $P \implies Q$:

Direct proof:	assume P, prove Q
Proof by contrapositive:	assume $\neg Q$, prove $\neg P$
Proof by contradiction:	assume $P \wedge \neg Q$ and derive something that is impossible

1.2.5 Induction

Finally, we consider a special proof technique for proving statements about the natural numbers (or subsets of them of certain forms). It is based on the following theorem, which we state without proof.

Theorem 1.13 (Well-ordering principle for $\mathbb{N}$) *Every nonempty set of natural numbers has a least element.*

Because of the well-ordering principle, we can prove something holds for the natural numbers by proving it holds for the smallest one, and then creating a logical ladder linking them together as follows:

Theorem 1.14 (Principle of mathematical induction) *Let n_0 be a non-negative integer. Suppose P is a statement about positive integers n such that*

1. *(base case) $P(n_0)$ is true*
2. *(induction step) For every integer $k \geq n_0$, if $P(k)$ is true, then $P(k+1)$ is true.*

Then $P(n)$ is true for every integer $n \geq n_0$

Here is an example of a proof by induction.

Example 1.15 $n! > 2^n$ *if* $n \geq 4$.

Proof. We prove this by induction on n.
Base case: Let $n = 4$. Then $n! = 4! = 24 > 16 = 2^4$.
Inductive hypothesis: Suppose for some $k \geq 4$, $k! > 2^k$.
Then

$$(k+1)! = (k+1)k! > (k+1)2^k > 2(2^k) = 2^{k+1}.$$

Thus the statement holds by induction on n. □

Sometimes we use a different version of induction, called strong induction.

Theorem 1.16 (Principle of strong mathematical induction) *Let n_0 be a non-negative integer. Suppose P is a statement about positive integers n such that*

1. *(base case) $P(n_0)$ is true*
2. *(induction step) For every integer $k \geq n_0$, if $P(m)$ is true for every integer m with $n_0 \leq m \leq k$, then $P(m+1)$ is true.*

Then $P(n)$ is true for every integer $n \geq n_0$.

Next, we will consider an example where it is much simpler to use the strong version of induction than the regular one. First, we recall the definition of a prime number.

Definition 1.17 *A positive integer p is prime if p has exactly two positive integer factors: 1 and p. Note that 1 is not prime. We can write this as*

$$p > 1 \text{ is prime if } \forall\, a, b \in \mathbb{N},\, p = ab \implies (a = 1 \text{ or } b = 1).$$

We want to prove the existence part of the Fundamental Theorem of Arithmetic, that every integer greater than or equal to 2 has a prime factorization. The fact that such a factorization is unique is left as an exercise.

Example 1.18 *Every integer $n \geq 2$ can be written as the product of primes.*

Proof. We prove this using the Principle of Strong Mathematical Induction on n.

Base case: $n = 2$ is prime.

Inductive hypothesis: Suppose for some $k \geq 2$ that one can write every integer n such that $2 \leq n \leq k$ as a product of primes.

We must show that we can write $k + 1$ as a product of primes.

Case 1: if $k + 1$ is prime, then we are done.

Case 2: if $k + 1$ is not prime, then by Definition 1.17, there exists $a, b \in \mathbb{N}$ such that $k + 1 = ab$ where $a, b \neq 1$. Then it must also be the case that $a, b \leq k$.

By the inductive hypothesis, we can write a and b as products of primes, i.e. $\exists p_1, \ldots p_\ell, q_1, \ldots q_m$, all prime, such that

$$a = p_1 \cdots p_\ell, \qquad b = q_1 \cdots q_m.$$

Then

$$k + 1 = ab = p_1 \cdots p_\ell \, q_1 \cdots q_m,$$

therefore we can write $k + 1$ as a product of primes.

Thus the claim holds by strong induction. □

The Principle of Strong Mathematical Induction and the Principle of Mathematical Induction are logically equivalent, but sometimes it is easier to use one or the other, as we saw.

1.3 Exercises

1. Prove De Morgan's Laws for propositions: $\neg(P \wedge Q) = \neg P \vee \neg Q$ and $\neg(P \vee Q) = \neg P \wedge \neg Q$ (Hint: use truth tables).
2. Write the following statements and their negations using logical quantifier notation and then prove or disprove them:
   (i) Every odd integer is divisible by three.
   (ii) For any real number, twice its square plus twice itself plus six is greater than or equal to five. *(You may assume knowledge of calculus.)*
   (iii) Every integer can be written as a unique difference of two natural numbers.
3. Prove the following statements:
   (i) If $a|b$ and $a, b \in \mathbb{N}$ (positive integers), then $a \leq b$.
   (ii) If $a|b$ and $a|c$, then $a|(xb + yc)$, where $a, b, c, x, y \in \mathbb{Z}$.
   (iii) Let $a, b, n \in \mathbb{Z}$. If n does not divide the product ab, then n does not divide a and n does not divide b.
4. Prove that for all integers $n \geq 1$, $3|(2^{2n} - 1)$.
5. Prove the Fundamental Theorem of Arithmetic, that every integer $n \geq 2$ has a unique prime factorization (i.e. prove that the prime factorization from the last proof is unique).

1.4 References

Most of this content may be found in Chapter 1 of (Ger12), though many of the examples are my own. (Lak16) is also a great resource. In particular, the content on induction comes from there. Unfortunately the latter text is not freely available online or at U of T.

2 Set theory

2.1 Basics

For our purposes, we define a *set* to be a collection of mathematical objects. If S is a set and x is one of the objects in the set, we say x is an element of S and denote it by $x \in S$. The set of no elements is called empty set and is denoted by $\emptyset$.

Definition 2.1 (Subsets, Union, Intersection) *Let S, T be sets.*

- *We say that S is a* subset *of T, denoted $S \subseteq T$, if $s \in S$ implies $s \in T$.*
- *We say that $S = T$ if $S \subseteq T$ and $T \subseteq S$.*
- *We define the* union *of S and T, denoted $S \cup T$, as all the elements that are in* either *S or T.*
- *We define the* intersection *of S and T, denoted $S \cap T$, as all the elements that are in* both *S and T.*
- *We say that S and T are* disjoint *if $S \cap T = \emptyset$.*

Example 2.2 $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

Example 2.3 *Let $a, b \in \mathbb{R}$ such that $a < b$.*
Open interval: $(a, b) := \{x \in \mathbb{R} : a < x < b\}$ (a, b may be $-\infty$ or $+\infty$)
Closed interval: $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$
We can also define half-open intervals.

Example 2.4 *Let $A = \{x \in \mathbb{N} : 3|x\}$ and $B = \{x \in \mathbb{N} : 6|x\}$ Show that $B \subseteq A$.*

Proof. Let $x \in B$. Then $6|x$, i.e. $\exists j \in \mathbb{Z}$ such that $x = 6j$. Therefore $x = 3(2j)$, so $3|x$. Thus $x \in A$. $\square$

Definition 2.5 *Let $A, B \subseteq X$. We define the* set-theoretic difference *of A and B, denoted $A \setminus B$ (sometimes $A - B$) as the elements of X that are in A but* not *in B.*

The complement of a set $A \subseteq X$ is the set $A^c := X \setminus A$.

We extend the definition of union and intersection to an arbitrary family of sets as follows:

Definition 2.6 *Let S_α, $\alpha \in A$, be a family of sets. A is called the* index set. *We define*

$$\bigcup_{\alpha \in A} S_\alpha := \{x : \exists \alpha \text{ such that } x \in S_\alpha\},$$

$$\bigcap_{\alpha \in A} S_\alpha := \{x : x \in S_\alpha \text{ for all } \alpha \in A\}.$$

Example 2.7

$$\bigcup_{n=1}^{\infty} [-n, n] = \mathbb{R}$$

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$

Theorem 2.8 (De Morgan's Laws) *Let $\{S_\alpha\}_{\alpha \in A}$ be an arbitrary collection of sets. Then*

$$\left(\bigcup_{\alpha \in A} S_\alpha\right)^c = \bigcap_{\alpha \in A} S_\alpha^c \quad \text{and} \quad \left(\bigcap_{\alpha \in A} S_\alpha\right)^c = \bigcup_{\alpha \in A} S_\alpha^c$$

Proof. For the first part: Let $x \in \left(\bigcup_{\alpha \in A} S_\alpha\right)^c$. This is true if and only if $x \notin \left(\bigcup_{\alpha \in A} S_\alpha\right)$, or in other words $x \in S_\alpha^c$ for all $\alpha \in A$. This is true if and only if $x \in \bigcap_{\alpha \in A} S_\alpha^c$, which gives the result.
The second part is similar and is left as an exercise. $\square$

Since a set is itself a mathematical object, a set can itself contain sets.

Definition 2.9 *The power set $\mathcal{P}(S)$ of a set S is the set of all subsets of S.*

Example 2.10 *Let $S = \{a, b, c\}$. Then $\mathcal{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, S\}$.*

Another way of building a new set from two old ones is the Cartesian product of two sets.

Definition 2.11 *Let S, T be sets. The* Cartesian product *$S \times T$ is defined as the set of tuples with elements from S, T, i.e*

$$S \times T = \{(s, t) : \ s \in S \ \text{ and } \ t \in T\}.$$

This can also be extended inductively to a finite family of sets.

2.2 Ordered sets

Definition 2.12 *A* relation *R on a set X is a subset of $X \times X$. We say that $x \leq y$ if $(x, y) \in \mathbb{R}$. A relation $\leq$ is called a* partial order *on X if it satisfies*

1. *Reflexivity: $x \leq x$ for all $x \in X$*
2. *Transitivity: for $x, y, z \in X$, $x \leq y$ and $y \leq z$ implies $x \leq z$*
3. *Anti-symmetry: for $x, y \in X$, $x \leq y$ and $y \leq x$ implies $x = y$*

The pair $(X, \leq)$ is called a partially ordered set*.*

A chain *or* totally ordered set *$C \subseteq X$ is a subset with the property $x \leq y$ or $y \leq x$ for any $x, y \in C$.*

Example 2.13 *The real numbers with the usual ordering, $(\mathbb{R}, \leq)$, are totally ordered.*

Example 2.14 *The power set of a set X with the ordering given by subsets, $(\mathcal{P}(X), \subseteq)$ is partially ordered set.*

Example 2.15 *Let $X = \{a, b, c, d\}$. What is $\mathcal{P}(X)$? Find a chain in $\mathcal{P}(X)$.*

$\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{b, c\}, \{c, d\}, \{b, d\}, \{a, c\}, \{a, d\}, \{a, b, c\}, \{b, c, d\}, \{a, b, d\}, \{a, c, d\}, X\}$

An example of a chain $C \subseteq \mathcal{P}(X)$ is $C = \{\emptyset, \{b\}, \{b, c\}, \{a, b, c\}, X\}$

Example 2.16 *Consider the set $C([0, 1], \mathbb{R}) := \{f : [0, 1] \to \mathbb{R} : f \text{ is continuous}\}$.*

For two functions $f, g \in C([0, 1], \mathbb{R})$, we define the ordering as $f \leq g$ if $f(x) \leq g(x)$ for $x \in [0, 1]$. Then $(C([0, 1], \mathbb{R}), \leq)$ is a partially ordered set. Can you think of a chain that is a subset of $(C([0, 1], \mathbb{R})$?

Definition 2.17 *A non-empty partially ordered set $(X, \leq)$ is* well-ordered *if every non-empty subset $A \subseteq X$ has a mimimum element.*

Recall that we already saw that $\mathbb{N}$ is well-ordered, as we used it to prove the principle of mathematical induction. $\mathbb{R}$ with the usual order does not have this property.

Having a partially ordered set allows us to talk about upper and lower bounds.

Definition 2.18 *Let $(X, \leq)$ be a partially ordered set and $S \subseteq X$. Then $x \in X$ is an* upper bound *for S if for all $s \in S$ we have $x \leq s$. Similarly $y \in X$ is a* lower bound *for S if for all $s \in S$, $y \leq s$. If there exists an upper bound for S, we call S* bounded above *and if there exists a lower bound for S, we call S* bounded below*. If S is bounded above and bounded below, we say S is* bounded*.*

We can also ask if there exists a least upper bound or a greatest lower bound.

Definition 2.19 *Let $(X, \leq)$ be a partially ordered set and $S \subseteq X$. We call $x \in X$* least upper bound *or* supremum*, denoted $x = \sup S$, if x is an upper bound and for any other upper bound $y \in X$ of S we have $x \leq y$. Likewise $x \in X$ is the* greatest lower bound *or* infimum *for S, denoted $x = \inf S$, if it is a lower bound and for any other lower bound $y \in X$, $y \leq x$.*

Note that the supremum and infimum of a bounded set do not necessarily need to exist. However, if they do exists they are unique, which justifies the article *the* (see Exercise 4). Nevertheless, the reals have a remarkable property, which we will take as an axiom.

Axiom 2.20 *[Completeness Axiom] Let $S \subseteq \mathbb{R}$ be bounded above. Then there exists $r \in \mathbb{R}$ such that $r = \sup S$, i.e. S has a least upper bound.*

By setting $S' = -S := \{-s : s \in S\}$ and noting $\inf S = -\sup S'$, we obtain a similar statement for infima if S is bounded below. As mentioned above, this property is fairly special, for example it fails for the rationals.

Example 2.21 *Let $S = \{q \in \mathbb{Q} : x^2 < 7\}$. Then S is bounded above in $\mathbb{Q}$, but there exists no least upper bound in $\mathbb{Q}$.*

There is a nice alternative characterization for suprema in the real numbers.

Proposition 2.22 *Let $S \subseteq \mathbb{R}$ be bounded above. Then $r = \sup S$ if and only if r is an upper bound and for all $\epsilon > 0$ there exists an $s \in S$ such that $r - \epsilon < s$.*

Proof. ($\Rightarrow$) We will prove the forward direction ($\implies$) by contrapositive. Suppose r is either not an upper bound or there exists an $\epsilon > 0$ such that for all $s \in S$, $r - \epsilon \geq s$. In the first case, r is not the supremum by definition. In the second case, $r - \varepsilon$ is an upper bound which is smaller than r. Thus $r \neq \sup S$.

($\Leftarrow$) For the backward direction we will proceed by contradiction. Suppose r is an upper bound and for all $\epsilon > 0$ there exists an $s \in S$ such that $r - \epsilon < s$, but $r \neq \sup S$. Then $\sup S < r$ or equivalently $r - \sup S > 0$. Then by assumption there exists an $s \in S$ such that $\sup S = r - (r - \sup S) < s$, which contradicts the definition of supremum. ☐

Using the same trick, we may obtain a similar result for infima.

Proposition 2.23 *Let $S \subseteq \mathbb{R}$ be bounded below. Then $r = \inf S$ if and only if r is a lower bound and for all $\epsilon > 0$ there exists an $s \in S$ such that $r + \epsilon > s$.*

Example 2.24 *Consider $S = \{1/n : n \in \mathbb{N}\}$. Then $\sup S = 1$ and $\inf S = 0$.*

2.3 Functions

One way to define a function is as follows :

Definition 2.25 ((Run05, Definition 1.1.14)) *A function f from a set X to a set Y is a subset of $X \times Y$ with the properties:*

1. *For every $x \in X$, there exists a $y \in Y$ such that $(x, y) \in f$*
2. *If $(x, y) \in f$ and $(x, z) \in f$, then $y = z$.*

X is called the domain *of f.*

How does this connect to other descriptions of functions you may have seen? Instead of writing $f \subseteq X \times Y$, we often write $f : X \to Y,\ x \mapsto y$, where $(x, y) \in f$.

Example 2.26 *For a set X, the identity function is:*

$$1_X : X \to X, \quad x \mapsto x$$

Definition 2.27 (Image and pre-image) *Let $f : X \to Y$ and $A \subseteq X$ and $B \subseteq Y$. The image of f is the set $f(A) := \{f(x) : x \in A\}$ and the pre-image of f is the set $f^{-1}(B) := \{x : f(x) \in B\}$*

The following re-statements of the above may be helpful way to think about it for proofs:
If $y \in f(A)$, then $y \in Y$, and there exists an $x \in A$ such that $y = f(x)$.
If $x \in f^{-1}(B)$, then $x \in X$ and $f(x) \in B$.

Definition 2.28 (Surjective, injective and bijective) *Let $f : X \to Y$, where X and Y are sets. Then*

- f *is* injective *if $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$*
- f *is* surjective *if for every $y \in Y$, there exists an $x \in X$ such that $y = f(x)$*
- f *is* bijective *if it is both injective and bijective*

Example 2.29 *Let $f : X \to Y$, $x \mapsto x^2$.*
If $X = \mathbb{R}$ and $Y = [0, \infty)$: f is surjective.
If $X = [0, \infty)$ and $Y = \mathbb{R}$: f is injective.
If $X = Y = [0, \infty)$: f is bijective.
If $X = Y = \mathbb{R}$, then f is neither surjective nor injective.

Proposition 2.30 *Let $f : X \to Y$ and $A \subseteq X$. Prove that $A \subseteq f^{-1}(f(A))$, with equality if f is injective.*

Proof. First we show $A \subseteq f^{-1}(f(A))$. Let $x \in A$. Let $B = f(A)$, $B \subseteq Y$. By definition, $f(x) \in B$. So then again by definition, $x \in f^{-1}(B)$. Thus $x \in f^{-1}(f(A))$.

Next, suppose f is injective. We have already shown that $A \subseteq f^{-1}(f(A))$, so it remains to show that $f^{-1}(f(A)) \subseteq A$. Let $x \in f^{-1}(f(A))$. Then $f(x) \in f(A)$ by the definition of the pre-image. This means that there exists a $\tilde{x} \in A$ such that $f(x) = f(\tilde{x})$. Since f is injective, we have $x = \tilde{x}$, and hence $x \in A$. □

2.4 Cardinality

Intuitively, the *cardinality* of a set A, denoted $|A|$, is the number of elements in the set. For sets with only a finite number of elements, this intuition is correct. We call a set with finitely many elements finite.

We say that the empty set has cardinality 0 and is finite.

Proposition 2.31 *If X is finite set of cardinality n, then the cardinality of $\mathcal{P}(X)$ is 2^n.*

Proof. We proceed by induction. First, suppose $n = 0$. Then $X = \emptyset$, and $\mathcal{P}(X) = \{\emptyset\}$ which has cardinality $1 = 2^0$.

Next, suppose that the claim holds for some $n \in \mathbb{N}_0$. Let X have $n + 1$ elements. Let's call them $\{x_1, \ldots, x_n, x_{n+1}\}$. Then we can split X up into subsets $A = \{x_1, \ldots, x_n\}$ and $B = \{x_{n_1}\}$. By the inductive hypothesis, $\mathcal{P}(A)$ has cardinality 2^n. Any subset of X must either be a subset of A or contain x_{n+1}. How many subsets are there for the latter form? Let's count them out. Each subset will be formed by taking elements from A and combining them with x_{n+1}. We start with no elements from A and count up to all of them:

$$
\begin{aligned}
&1 + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n-1} + \binom{n}{n} \\
=&\sum_{k=0}^{n} \binom{n}{k} \\
=&2^n
\end{aligned}
$$

Therefore the total number of elements in $\mathcal{P}(X)$ is the number of subsets of A (2^n) plus the number of mixed subsets (2^n), i.e. the cardinality of $\mathcal{P}(X)$ is $2^n + 2^n = 2^{n+1}$.

Thus the claim holds by induction. □

Note: you do not need to prove this by induction. There are other ways to do it. You can try to prove it without using induction as an exercise.

Definition 2.32 *Two sets A and B have same cardinality, $|A| = |B|$, if there exists bijection $f : A \to B$.*

Example 2.33 *Which is bigger, $\mathbb{N}$ or $\mathbb{N}_0$?*
Intuitively, it seems that $\mathbb{N}_0$ should be bigger, since it includes exactly one more element than $\mathbb{N}$, namely 0. However, clearly the function $f : \mathbb{N}_0 \to \mathbb{N}$ defined by $n \mapsto n+1$ is a bijection. Therefore $\mathbb{N}_0$ and $\mathbb{N}$ have the same cardinality! One way to think about this is that $\mathbb{N}_0$ and $\mathbb{N}$ are the "same size" of infinity.

It may sometimes be difficult to find such a bijection. However you can also use the following definition and theorem to instead show that two sets have the same cardinality by finding two injective functions between them.

Definition 2.34 *We say that the cardinality of a set A is less than the cardinality of a set B, denoted $|A| \leq |B|$ if there exists an injection $f : A \to B$.*

Theorem 2.35 (Cantor-Schröder-Bernstein) *Let A, B, be sets. If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.*

Proof is omitted. See (Run05, Theorem 1.2.7)

Example 2.36 $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$

Proof. First, we show $|\mathbb{N}| \leq |\mathbb{N} \times \mathbb{N}|$. The function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ defined by $n \mapsto (n, 1)$ is an injection, thus $|\mathbb{N}| \leq |\mathbb{N} \times \mathbb{N}|$.

Next, we show $|\mathbb{N} \times \mathbb{N}| \leq |\mathbb{N}|$. We define the function $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $(n, m) \mapsto 2^n 3^m$. Why is this an injection? Assume we have n_1, n_2, m_1, m_2 such that $2^{n_1}3^{m_1} = 2^{n_2}3^{m_2}$. We need to show $n_1 = n_2$ and $m_1 = m_2$. By the Fundamental Theorem of Arithmetic, every natural number greater than 1 has a unique prime factorization, so therefore the result must hold. $\square$

Definition 2.37 *Let A be a set.*

1. *A is* finite *if there exists an $n \in \mathbb{N}$ and a bijection $f : \{1, \dots, n\} \to A$*
2. *A is* countably infinite *if there exists a bijection $f : \mathbb{N} \to A$*
3. *A is* countable *if it is finite or countably infinite*
4. *A is* uncountable *otherwise*

Example 2.38 *The rational numbers are countable, and in fact $|\mathbb{Q}| = |\mathbb{N}|$.*
Let's look at $\mathbb{Q}^+ := \{x \in \mathbb{Q} : x > 0\}$. The fact that the rationals are countable relies on this famous way of listing the rational numbers:

$$\begin{array}{cccccc}
1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots \\
2 & \color{red}{\frac{2}{2}} & \frac{2}{3} & \color{red}{\frac{2}{4}} & \frac{2}{5} & \cdots \\
3 & \frac{3}{2} & \color{red}{\frac{3}{3}} & \frac{3}{4} & \frac{3}{5} & \cdots \\
4 & \color{red}{\frac{4}{2}} & \frac{4}{3} & \color{red}{\frac{4}{4}} & \frac{4}{5} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

This is a map from $\mathbb{N}$ to $\mathbb{Q}^+$. As long as we skip any fraction that is already in our list as we go along, it is injective. Since we can find an injection from $\mathbb{Q}^+$ to $\mathbb{N} \times \mathbb{N}$ (take $q/p \mapsto (q, p)$), and $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$ by Example 2.36, we have that $|\mathbb{Q}^+| = |\mathbb{N}|$.

We can extend this to $\mathbb{Q}$. To do so, let $f \colon \mathbb{N} \to \mathbb{Q}^+$ be a bijection (which exists by the previous part). Then we can define another bijection $g \colon \mathbb{N} \to \mathbb{Q}$ by setting $g(1) = 0$ and

$$g(n) = \begin{cases} f(n) & \text{if } n \text{ is even,} \\ -f(n) & \text{if } n \text{ is odd,} \end{cases}$$

for $n > 1$.

Next we show that $\mathbb{N}$ is "smaller" than $(0,1)$.

Theorem 2.39 *The cardinality of* $\mathbb{N}$ *is smaller than that of* $(0,1)$.

Proof. First, we show that there is an injective map from $\mathbb{N}$ to $(0,1)$. The map $n \to \frac{1}{n}$ fulfils this.

Next, we show that there is no surjective map from $\mathbb{N}$ to (0, 1). We use the fact that every number $r \in (0,1)$ has a binary expansion of the form $r = 0.\sigma_1\sigma_2\sigma_3\ldots$ where $\sigma_i \in \{0,1\}$, $i \in \mathbb{N}$.

Now we suppose in order to derive a contradiction that there does exist a surjective map f from $\mathbb{N}$ to (0, 1)., i.e. for $n \in \mathbb{N}$ we have $f(n) = 0.\sigma_1(n)\sigma_2(n)\sigma_3(n)\ldots$. This means we can list out the binary expansions, for example like

$$\begin{aligned} f(1) =& 0.\textcolor{red}{0}00000000\ldots \\ f(2) =& 0.1\textcolor{red}{1}111111111\ldots \\ f(3) =& 0.01\textcolor{red}{0}1010101\ldots \\ f(4) =& 0.101\textcolor{red}{0}101010\ldots \end{aligned}$$

We will construct a number $\tilde{r} \in (0,1)$ that is not in the image of f. Define $\tilde{r} = 0.\tilde{\sigma}_1\tilde{\sigma}_2\ldots$, where we define the nth entry of $\tilde{r}$ to be the the opposite of the nth entry of the nth item in our list:

$$\tilde{\sigma}_n = \begin{cases} 1 & \text{if } \sigma_n(n) = 0, \\ 0 & \text{if } \sigma_n(n) = 1. \end{cases}$$

Then $\tilde{r}$ differs from $f(n)$ at least in the nth digit of its binary expansion for all $n \in \mathbb{N}$. Hence, $\tilde{r} \notin f(\mathbb{N})$, which is a contradiction to f being surjective. This technique is often referred to as Cantor's diagonal argument. ☐

Proposition 2.40 *(0,1) and* $\mathbb{R}$ *have the same cardinality.*

Proof. The map $f : \mathbb{R} \to (0,1)$ defined by $x \mapsto \frac{1}{\pi}\left(\arctan(x) + \frac{\pi}{2}\right)$ is a bijection. ☐

We have shown that there are different sizes of infinity, as the cardinality of $\mathbb{N}$ is infinite but still smaller than that of $\mathbb{R}$ or $(0,1)$. In fact, we have

$$|\mathbb{N}| = |\mathbb{N}_0| = |\mathbb{Z}| = |\mathbb{Q}| < |\mathbb{R}|.$$

Because of this, there are special symbols for these two cardinalities: The cardinality of $\mathbb{N}$ is denoted $\aleph_0$, while the cardinality of $\mathbb{R}$ is denoted $\mathfrak{c}$.

2.5 Exercises

1. Let $A = \{x \in \mathbb{R} : x < 100\}$, $B = \{x \in \mathbb{Z} : |x| \geq 20\}$, and $C = \{y \in \mathbb{N} : y \text{ is prime}\}$ $(A, B, C \subseteq \mathbb{R})$. Find $A \cap B$, $B^c \cap C$, $B \cup C$, and $(A \cup B)^c$.
2. Is $\mathbb{R} \times \mathbb{R}$ with the ordering $(x_1, y_1) \preceq (x_2, y_2)$ if $x_1 \leq x_2$ a partially ordered set?
3. (Run05, Exercise 1.3.1) Let S be a non-empty set. A relation R on S is called an equivalence relation if it is
   (i) Reflexive: $(x,x) \in R$ for all $x \in S$
   (ii) Symmetric: if $(x,y) \in R$ then $(y,x) \in R$ for all $x, y \in S$
   (iii) Transitive: if $(x,y), (y,z) \in R$ then $(x,z) \in R$ for all $x, y, z \in S$

Given $x \in S$, the equivalence class of x (with respect to a given equivalence relation R) is defined to consist of those $y \in S$ for which $(x, y) \in R$. Show that two equivalence classes are either disjoint or identical.

4. Let $(X, \leq)$ be a partially ordered set and $S \subseteq X$ be bounded. Show that the infimum and supremum of S are unique (if they exist).
5. Let $S, T \subseteq \mathbb{R}$ and suppose both are bounded above. Define $S + T = \{s + t : s \in S, t \in T\}$. Show that $S + T$ is bounded above and $\sup(S + T) = \sup S + \sup T$.
6. Let $f : X \to Y$, $X, y \subseteq \mathbb{R}$ be defined by the map $x \mapsto \sin(x)$. For what choices of X and Y is f injective, surjective, bijective, or neither?
7. Show that for sets $A, B \subseteq X$ and $f : X \to Y$, $f(A \cap B) \subseteq f(A) \cap f(B)$.
8. Let $f : X \to Y$ and $B \subseteq Y$. Prove that $f(f^{-1}(B)) \subseteq B$, with equality iff f is surjective.
9. Prove that $f(\cup_{i \in I} A_i) = \cup_{i \in I} f(A_i)$, where $f : X \to Y$, $A_i \subseteq Y \; \forall i \in I$.
10. Show that $\mathbb{N}$ and $\mathbb{Z}$ have the same cardinality.
11. Show that $|(0, 1)| = |(1, \infty)|$.

2.6 References

The content in this section mostly follows (Run05), but is supplemented by (Mar19) (ordered sets) and (Zwi22) (introductory set theory).

3 Metric spaces and sequences

3.1 Metric spaces

Definition 3.1 *A* metric *on a set X is a function $d : X \times X \to \mathbb{R}$ that satisfies:*

(a) Positive definiteness: $d(x, y) \geq 0$ for $x, y \in X$ and $d(x, y) = 0 \Leftrightarrow x = y$

(b) Symmetry: for $x, y \in X$, $d(x, y) = d(y, x)$

(c) Triangle inequality: for $x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$

A set together with a metric is called a metric space.

Example 3.2 $\mathbb{R}^n$ *with the Euclidean distance*

$$d(x, y) = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2} \quad \text{for } x, y \in \mathbb{R}^n$$

is a metric space.

Many metric spaces we know are in fact normed spaces, which have more structure than metric spaces. We will briefly discuss normed spaces. This assumes some knowledge of vector spaces, which we will discuss in a further section. In particular, we denote a field by $\mathbb{F}$. For now, we can think of this as $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.3 *A* norm *on an $\mathbb{F}$-vector space E is a function $\| \cdot \| : E \to \mathbb{R}$ that satisfies:*

(a) Positive definiteness: $\|x\| \geq 0$ for $x \in E$ and $||x|| = 0 \Leftrightarrow x = 0$

(b) Homogeneity: for $x \in E$ and $\alpha \in \mathbb{F}$, $||\alpha x|| = |\alpha| ||x||$

(c) Triangle inequality: for $x, y \in E, ||x + y|| \leq ||x|| + ||y||$

A vector space with a norm is called a normed space. A normed space is a metric space using the metric $d(x, y) = \|x - y\|$.

Example 3.4 *The p-norm is defined for $p \geq 1$ for a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ as*

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

The infinity norm is the limit of the p-norm as $p \to \infty$, defined as

$$||x||_\infty = \max_{i=1,\ldots,n} |x_i|.$$

If we look at the space of continuous functions $C([0,1];\mathbb{R})$, the p-norm is

$$||f||_p = \left(\int_0^1 |f(x)|^p dx \right)^{1/p}$$

and the $\infty-$norm (or sup norm) is

$$||f||_\infty = \max_{x \in [0,1]} |f(x)|.$$

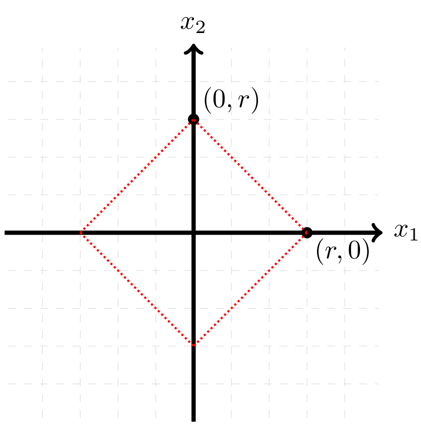
Definition 3.5 *A subset A of a metric space (X, d) is* bounded *if there exists $M > 0$ such that $d(x, y) < M$ for all $x, y \in A$.*

Definition 3.6 *Let (X, d) be a metric space. We define the* open ball *centred at a point $x_0 \in X$ of radius $r > 0$ as*

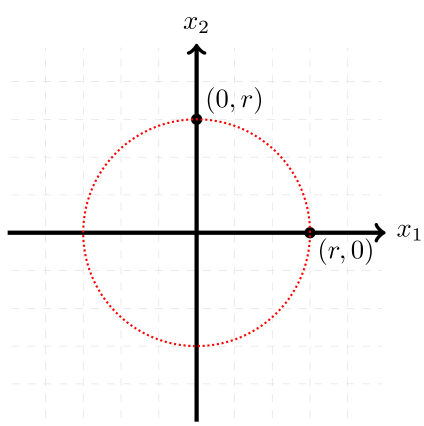
$$B_r(x_0) := \{x \in X : d(x, x_0) < r_0\}.$$

Example 3.7 *In $\mathbb{R}$ with the usual norm (absolute value), open balls are symmetric open intervals, i.e. $B_r(x_0) = (x_0 - r, x_0 + r)$.*

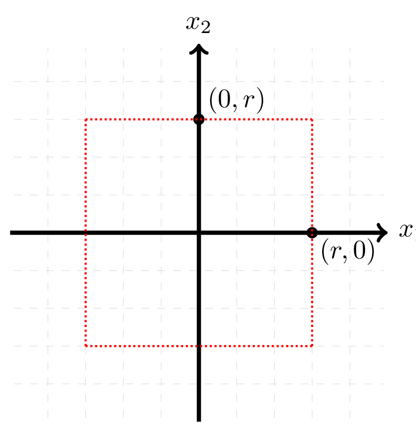
Example 3.8 *Consider $\mathbb{R}^2$ with the taxicab or Manhattan metric (1-norm) $d(x, y) = \sum_{i=1}^{2} |x_i - y_i|$, the usual Euclidean distance (2-norm) $d(x, y) = \sqrt{\sum_{j=1}^{2} (x_j - y_j)^2}$, and the ∞-norm $d(x, y) = \max_{j=1,2} |x_j - y_j|$. The open ball $B_r(0)$ in these three metric spaces is shown in Fig. 1.*



(a) 1-norm (taxicab metric)



(b) 2-norm (Euclidean metric)



(c) ∞-norm

Figure 1: $B_r(0)$ for different metrics

Definition 3.9 (Open and closed sets) *Let (X, d) be a metric space.*

- *A set $U \subseteq X$ is* open *if for every $x \in U$ there exists $\epsilon > 0$ such that $B_\epsilon(x) \subseteq U$.*
- *A set $F \subseteq X$ is* closed *if $F^c := X \setminus F$ is open.*

We note that $\emptyset$ and X are both open and closed!

Proposition 3.10 *Let (X, d) be a metric space.*

(i) Let $A_1, A_2 \subseteq X$. If A_1 and A_2 are open, then $A_1 \cap A_2$ is open.

(ii) If $A_i \subseteq X$, $i \in I$ are open, then $\cup_{i\in I} A_i$ is open.

Proof. (i) Since A_1 is open, for each $x \in A_1$, there exists an $\epsilon_1 > 0$ such that $B_{\epsilon_1}(x) \subseteq A_1$. Since A_2 is open, for each $x \in A_2$, there exists an $\epsilon_2 > 0$ such that $B_{\epsilon_2}(x) \subseteq A_2$. Let $x \in A_1 \cap A_2$. Choose $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. Then $B_\epsilon(x) \subseteq A_1 \cap A_2$ as required.

(ii) Let $x \in \cup_{i\in I} A_i$. Then there exists $i \in I$ such that $x \in A_i$, and since A_i is open there exists $\epsilon > 0$ such that $B_\epsilon(x) \subseteq A_i$. Since $A_i \subseteq \cup_{i\in I} A_i$, we are done. □

Using DeMorgan, we immediately have the following corollary:

Corollary 3.11 *Let (X, d) be a metric space.*

(i) Let $A_1, A_2 \subseteq X$. If A_1 and A_2 are closed, then $A_1 \cup A_2$ is closed.

(ii) If $A_i \subseteq X$, $i \in I$ are closed, then $\cap_{i\in I} A_i$ is closed.

Definition 3.12 (Interior and closure) *Let $A \subseteq X$ where (X, d) is a metric space.*

- *The* closure *of A is $\overline{A} := \{x \in X : \forall\epsilon > 0 \ B_\epsilon(x) \cap A \neq \emptyset\}$*
- *The* interior *of A is $\mathring{A} := \{x \in X : \exists\epsilon > 0 \text{ s.t. } B_\epsilon(x) \subseteq A\}$*
- *The* boundary *of A is $\partial A := \{x \in X : \forall\epsilon > 0, \ B_\epsilon(x) \cap A \neq \emptyset \text{ and } B_\epsilon(x) \cap A^c \neq \emptyset\}$*

The closure of a set is the smallest closed set that contains it while the interior of a set is the largest open set contained by it.

Example 3.13 *Let $X = (a, b] \subseteq \mathbb{R}$ with the ordinary (Euclidean) metric. Then $\overline{X} = [a, b]$, $\mathring{X} = (a, b)$ and $\partial X = \{a, b\}$.*

Proposition 3.14 *Let $A \subseteq X$ where (X, d) is a metric space. Then $\mathring{A} = A \setminus \partial A$.*

Proof. First, we show $\mathring{A} \subseteq A \setminus \partial A$. Let $x \in \mathring{A}$. Then by definition $\exists\epsilon > 0$ s.t. $B_\epsilon(x) \subseteq A$. Clearly $x \in A$ and also $\exists\epsilon > 0$ such that $B_\epsilon(x) \cap A^c = \emptyset$. Thus by definition, $x \notin \partial A$. Thus $x \in A \setminus \partial A$.

Next, we show $A \setminus \partial A \subseteq \mathring{A}$. Let $x \in A \setminus \partial A$. Then $x \in A$ and $x \notin \partial A$. The latter means that $\exists\epsilon > 0$ such that $B_\epsilon(x) \cap A = \emptyset$ or $B_\epsilon(x) \cap A^c = \emptyset$. Since $x \in A$, $x \in B_\epsilon(x) \cap A$ for any $\epsilon > 0$, so the former cannot be true. Therefore $\exists\epsilon > 0$ such that $B_\epsilon(x) \cap A^c = \emptyset$, i.e. $B_\epsilon(x) \subseteq A$. Thus $x \in \mathring{A}$. □

3.2 Sequences

Definition 3.15 *Let (X, d) be a metric space. A* sequence *is an ordered list of points x_n, $n \in \mathbb{N}$, in X, denoted $(x_n)_{n\in\mathbb{N}}$. We say that a sequence $(x_n)_{n\in\mathbb{N}}$* converges *to a point $x \in X$ if*

$$\forall\epsilon > 0 \, \exists \, n_\epsilon \in \mathbb{N} \text{ s.t. } d(x_n, x) < \epsilon \text{ for all } n \geq n_\epsilon.$$

Proposition 3.16 *Let (X, d) be a metric space, and let $A \subseteq X$. Then $\overline{A}$ is equal to the set of points in X which are limits of a sequence in A.*

Proof. Let $x \in \overline{A}$. Then by definition, for every $\epsilon > 0$, $B_\epsilon(x) \cap A \neq \emptyset$. In particular this is true for $\epsilon = 1/n$. Thus, for any $n \in N$, we can choose an $x_n \in A$ such that $x_n \in B_{1/n}(x)$, which means $d(x, x_n) < \frac{1}{n}$ by the definition of an open ball. Since $1/n$ decreases monotonically to zero, we must have $x_n \to x$.

Let $x \in X$ be the limit of a sequence $(x_n)_{n\in\mathbb{N}} \in A$. Then for $\epsilon > 0$, $\exists\, n_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq n_\epsilon$. This means $x_n \in B_\epsilon(x)$, and since $x_n \in A$, $B_\epsilon(x) \cap A \neq \emptyset$. Thus $x \in \overline{A}$. $\square$

Combining this result with the fact that $A \subseteq X$ is closed if and only if $A = \overline{A}$ (exercise), gives the following useful way to characterize closed sets.

Corollary 3.17 *A set $F \subseteq X$, where (X, d) is a metric space, is closed if and only if every sequence in F which converges in X converges to a point in F.*

We also define a concept related to the closure of a set: a cluster or accumulation point.

Definition 3.18 *Let (X, d) be a metric space and $A \subseteq X$. A point $x \in X$ is a* cluster point *of A (also called accumulation point) if for every $\epsilon > 0$, $B_\epsilon(x)$ contains infinitely many points in A.*

Proposition 3.19 *$x \in X$ is a cluster point of $A \subseteq X$ where (X, d) is a metric space if and only if there exists a sequence of points $x_n \in A$, $n \in \mathbb{N}$, such that $x_n \to x$.*

Proof. ($\Leftarrow$) Suppose there exists a sequence $(x_n)_{n\in N}$ in A such that $x_n \to x$. Then for every $\epsilon > 0$, by the definition of a convergent sequence, $B_\epsilon(x)$ contains infinitely many elements of the sequence x_n (in particular, $\exists n_0 \in \mathbb{N}$ such that $x_n \in B_\epsilon(x)$ for all $n \geq n_0$). Since each $x_n \in A$, x is a cluster point of A.

($\Rightarrow$) Suppose x is a cluster point of A. Then for any $\epsilon > 0$, $\exists x_\epsilon \in A$ such that $x_\epsilon \in B_\epsilon(x)$. In particular, take $\epsilon = 1/n$. Then $\exists x_n \in A$ such that $x_n \in B_{1/n}(x)$. By construction, such x_n form a sequence in A that converges to x.

$\square$

Combining Proposition 3.16 and Proposition 3.19 gives the following:

Corollary 3.20 *For $A \subseteq X$, (X, d) a metric space, we have $\overline{A} = A \cup \{x \in X : x \text{ is a cluster point of } A\}$.*

3.2.1 Cauchy sequences

Definition 3.21 (Cauchy sequence) *Let (X, d) be a metric space. A sequence denoted $(x_n)_{n\in\mathbb{N}} \in X$ is called a* Cauchy sequence *if*

$$\forall \epsilon\, \exists\, n_\epsilon \in \mathbb{N} \text{ s.t. } d(x_n, x_m) < \epsilon \text{ for all } n, m \geq n_\epsilon.$$

Proposition 3.22 *Let (X, d) be a metric space, and let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence in X. Then $(x_n)_{n\in\mathbb{N}}$ is Cauchy.*

Proof. Let $\epsilon > 0$ be arbitrary. Let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence in a metric space (X, d). Then there exists $n_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \frac{\epsilon}{2}$ for all $n \geq n_\epsilon$. Then for $n, m \geq n_\epsilon$, using the triangle inequality we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus $(x_n)_{n\in\mathbb{N}}$ is Cauchy. $\square$

Definition 3.23 *A metric space where every Cauchy sequence converges (to a point in the space) is called* complete.

In addition, a normed space that is complete with respect to the metric induced by the norm is called a *Banach space*. $\mathbb{R}^n$ with the Euclidean distance is complete (and is, in fact, a Banach space).

Proposition 3.24 ((Run05, Proposition 2.4.5)) *Let (X, d) be a metric space, and let $Y \subseteq X$.*

(i) If X is complete and if Y is closed in X, then Y is complete.

(ii) If Y is complete, then it is closed in X.

Proof. (i) Let X be a complete metric space and Y be a closed subset of X. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in Y. Since $Y \subseteq X$, $(x_n)_{n\in\mathbb{N}}$ is also a Cauchy sequence in X. Therefore $(x_n)_{n\in\mathbb{N}}$ converges to an $x \in X$ since X is complete. But since Y is closed, by Proposition 3.16, we must have $x \in Y$. Therefore Y is complete.

(ii) Let (X, d) be a metric space and let $Y \subseteq X$ be complete. Let $(y_n)_{n\in\mathbb{N}}$ be a sequence in Y that converges to some point $y \in X$. By Proposition 3.22, $(y_n)_{n\in\mathbb{N}}$ is Cauchy in X and therefore also in Y. Since Y is complete, $(y_n)_{n\in\mathbb{N}}$ converges to a point $y' \in Y$. Since sequences in metric spaces converge to unique points (see exercises), $y = y'$. Thus Y is closed by Corollary 3.17. □

3.2.2 Subsequences

Definition 3.25 *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space (X, d). Let $(n_k)_{k\in\mathbb{N}}$ be a sequence of natural numbers with $n_1 < n_2 < \cdots$. The sequence $(x_{n_k})_{k\in\mathbb{N}}$ is called a* subsequence *of $(x_n)_{n\in\mathbb{N}}$. If $(x_{n_k})_{k\in\mathbb{N}}$ converges to $x \in X$, we call x a* subsequential limit*.*

Example 3.26 *The sequence $((-1)^n)_{n\in\mathbb{N}}$ diverges but the subsequences $\left((-1)^{2n}\right)_{n\in\mathbb{N}}$ and $\left((-1)^{2n-1}\right)_{n\in\mathbb{N}}$ converge to subsequential limits 1 and -1, respectively.*

Proposition 3.27 *A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) converges to $x \in X$ if and only if every subsequence of $(x_n)_{n\in\mathbb{N}}$ also converges to x.*

Proof. ($\Leftarrow$) If every subsequence of $(x_n)_{n\in\mathbb{N}}$ converges to $x \in X$, then $(x_n)_{n\in\mathbb{N}}$ must converge to it as well, since a sequence is a subsequence of itself.

($\Rightarrow$) Suppose $(x_n)_{n\in\mathbb{N}}$ converges to $x \in X$ and let $(x_{n_k})_{k\in\mathbb{N}}$ be an arbitrary subsequence of $(x_n)_{n\in\mathbb{N}}$. Let $\epsilon > 0$ be arbitrary. There exists $n_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq n_\epsilon$. Choose k_ϵ such that $n_{k_\epsilon} \geq n_\epsilon$, which must exist since $(n_k)_{k\in\mathbb{N}}$ is strictly increasing. Then for all $k \geq k_\epsilon$, $d(x_{n_k}, x) < \epsilon$. Thus $(x_{n_k})_{k\in\mathbb{N}}$ converges to x. □

3.3 Continuity

Definition 3.28 *Let (X, d_X) and (Y, d_Y) be metric spaces, let $x_0 \in X$, and let $f : X \to Y$. f is* continuous *at x_0 if for every sequence $(x_n)_{n\in\mathbb{N}}$ in X that converges to x_0, we have $\lim_{n\to\infty} f(x_n) = f(x_0)$. We say that f is* continuous *if it is continuous at every point in X.*

Theorem 3.29 ((Run05, Theorem 2.3.7.)) *Let (X, d_X) and (Y, d_Y) be metric spaces, let $x_0 \in X$, and let $f : X \to Y$. The following are equivalent:*

(i) f is continuous at x_0

(ii) for all $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(x_0))) < \epsilon$ for all $x \in X$ with $d_X(x, x_0) < \delta$

(iii) for each $\epsilon > 0$, there is $\delta > 0$ such that $B_\delta(x_0) \subseteq f^{-1}(B_\epsilon(f(x_0)))$

Proof. (i) $\Rightarrow$ (ii) We prove the contrapositive. Assume

$$\exists \epsilon_0 \text{ such that } \forall \delta > 0 \text{ there exists an } x_\delta \in X \text{ with } d_X(x_\delta, x_0) < \delta \text{ and } d_Y(f(x_\delta), f(x_0))) \geq \epsilon_0 \qquad (\star)$$

We need to find a sequence in X that converges to x_0 but the sequence of images does not converge. Let's construct such a sequence.

Let $\delta = \frac{1}{n}$ in $(\star)$ for $n \in \mathbb{N}$. Then we can pick a sequence $x_n := x_{1/n}$ given by $(\star)$ which converges to x_0. However, for each $n \in \mathbb{N}$, we have $d_Y(f(x_n), f(x_0))) \geq \epsilon_0$, so we cannot have $\lim_{n\to\infty} f(x_n) = f(x_0)$.

(ii) $\Rightarrow$ (iii) Follows from the definitions of the pre-image and open balls.

(iii) $\Rightarrow$ (i) Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that converges to x_0. Let $\epsilon > 0$. Then by (iii), there exists $\delta > 0$ such that $B_\delta(x_0) \subseteq f^{-1}(B_\epsilon(f(x_0)))$, i.e. if x is such that $d_X(x, x_0) < \delta$, then x is such that $d_Y(f(x), f(x_0)) < \epsilon$. By the definition of convergence, there exists an $N \in \mathbb{N}$ such that $d(x_n, x_0) < \delta$ for all $n \geq N$. Then by (iii), $d(f(x_n), f(x_0)) < \epsilon$ for all $n \geq N$. Thus $\lim_{n\to\infty} f(x_n) = f(x_0)$. □

Corollary 3.30 *Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$. The following are equivalent:*

(i) f is continuous

(ii) if $U \subseteq Y$ is open, then $f^{-1}(U)$ is open

(iii) if $F \subseteq Y$ is closed, then $f^{-1}(F)$ is closed

Note: the following proof uses the following results, which you may wish to prove as an exercise using techniques from the set theory section if they are not clear to you: Let X and Y be sets and $f : X \to Y$. Let $A, B \subseteq Y$. Then

1. $A \subseteq B \implies f^{-1}(A) \subseteq f^{-1}(B)$
2. $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$

Proof. Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$.

(i) $\Rightarrow$ (ii): Suppose f is continuous (on every point in X) and let $U \subseteq Y$ be open. Let $x \in f^{-1}(U)$, then $f(x) \in U$, and since U is open, there exists $\epsilon_0 > 0$ such that $B_{\epsilon_0}(f(x)) \subseteq U$. By Theorem 3.29(iii), there exists a $\delta_0 > 0$ such that $B_{\delta_0}(x) \subseteq f^{-1}(B_{\epsilon_0}(f(x)))$. Since $B_{\epsilon_0}(f(x)) \subseteq U$, $f^{-1}(B_{\epsilon_0}(f(x))) \subseteq f^{-1}(U)$. Thus for each $x \in f^{-1}(U)$, there exists δ_0 such that $B_{\delta_0}(x) \subseteq f^{-1}(B_{\epsilon_0}(f(x))) \subseteq f^{-1}(U)$, so $f^{-1}(U)$ is open.

(ii) $\Rightarrow$ (i): We want to prove that f is continuous at every $x \in X$ using the definition from Theorem 3.29(iii), i.e. we must show that for $x \in X$, for each $\epsilon > 0$, there is $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(B_\epsilon(f(x)))$.

Let $x \in X$ and let $\epsilon > 0$ be arbitrary. Since $B_\epsilon(f(x))$ is an open set, by (ii), $f^{-1}(B_\epsilon(f(x)))$ is also open. Since $x \in f^{-1}(B_\epsilon(f(x)))$, there exists a $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(B_\epsilon(f(x)))$ by the definition of a set being open, so we are done.

(ii) $\Rightarrow$ (iii): Let $F \subseteq Y$ be closed. Then $Y \setminus F$ is open, so by (ii), $f^{-1}(Y \setminus F)$ is open as well. Since $f^{-1}(Y \setminus F) = X \setminus f^{-1}(F)$, $f^{-1}(F)$ is closed.

(iii) $\Rightarrow$ (ii) follows from the above, exchanging "open" and "closed".

$\square$

Definition 3.31 *Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$.*

- *f is* uniformly continuous *if for all $\epsilon > 0$, there exists $\delta > 0$ such that for every $x_1, x_2 \in X$ with $d_X(x_1, x_2) < \delta$, we have $d_Y(f(x_1), f(x_2))) < \epsilon$*
- *f is* Lipschitz continuous *if there exists a $K > 0$ such that for every $x_1, x_2 \in X$ we have $d_Y(f(x_1), f(x_2))) \leq K d_X(x_1, x_2)$*

Proposition 3.32 *Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$.*

f is Lipschitz continuous $\Rightarrow$ f is uniformly continuous $\Rightarrow$ f is continuous

The proof is left as an exercise.

Definition 3.33 *Let (X, d) be a metric space and let $f : X \to X$. We say that $x^* \in X$ is a* fixed point *of f if $f(x^*) = x^*$.*

Definition 3.34 *Let (X, d) be a metric space and let $f : X \to X$. f is a* contraction *if there exists a constant $k \in [0, 1)$ such that for all $x, y \in X$, $d(f(x), f(y))) \leq kd(x, y)$.*

Observe that a function is a contraction if and only if it is Lipschitz continuous with constant $K < 1$.

Theorem 3.35 *Suppose that $f : X \to X$ is a contraction and the metric space X is complete. Then f has a unique fixed point x^*.*

We omit the proof here; see (Pug15, p.240) for the proof as well as more details on how to find the fixed point.

Example 3.36 *Let* $f : \left[-\frac{1}{3}, \frac{1}{3}\right] \to \left[-\frac{1}{3}, \frac{1}{3}\right]$ *be defined by the mapping* $x \mapsto x^2$. *Assume we use the standard Euclidean metric,* $d(x, y) = |x - y|$. f *has a unique fixed point because* $\left[-\frac{1}{3}, \frac{1}{3}\right]$ *is a complete metric space (see Proposition 3.24) and* f *is a contraction with Lipschitz constant 2/3.*

To see that it is a contraction, let $x, y \in \left[-\frac{1}{3}, \frac{1}{3}\right]$. *Then*

$$|x^2 - y^2| = |x + y||x - y| \leq \frac{2}{3}|x - y|.$$

3.4 References

The content in this section comes mostly from (Run05). (Pug15) is used to supplement, and in particular the content on accumulation points and contractions comes from there. (Zwi22) is also used for some additional topics.

References

[Ger12] Larry J. Gerstein. *Introduction to Mathematical Structures and Proofs*. Undergraduate Texts in Mathematics. 2012. URL: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/978-1-4614-4265-3l.

[Lak16] Tamara J. Lakins. *The Tools of Mathematical Reasoning*. Pure and Applied Undergraduate Texts. 2016.

[Mar19] Laurent W. Marcoux. Pmath 351 notes, 2019. URL: https://www.math.uwaterloo.ca/~lwmarcou/notes/pmath351.pdf.

[Pug15] Charles C. Pugh. *Real Mathematical Analysis*. Undergraduate Texts in Mathematics. 2015. URL: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/978-3-319-17771-7.

[Run05] Volker Runde. *A Taste of Topology*. Universitext. 2005. URL: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/0-387-28387-0.

[Zwi22] Piotr Zwiernik. Lecture notes in mathematics for economics and statistics, 2022. URL: http://84.89.132.1/~piotr/docs/RealAnalysisNotes.pdf.