Mathematics Bootcamp Lecture Notes

Department of Statistical Sciences, University of Toronto

Emma Kroell

Last updated: July 9, 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 taught 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. *A Taste of Topology* by Volker Runde
3. *Linear Algebra Done Right* by Sheldon Axler
4. *Linear Algebra Done Wrong* by Sergei Treil
5. *Introduction to Real Analysis* by William F. Trench
6. *Real Mathematical Analysis* by Charles C. Pugh
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) is used as reference 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. 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.

Contents

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.

ex. "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.

Examples:

- 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.

1.1.1 Truth values

Example 1.1 *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 statement is true or not, we use a truth table

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

1.1.2 Logical equivalence

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

What is $\neg(P \implies Q)$?

1.1.3 Quantifiers

There are two important logical operators that we have not yet discussed. They are 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 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 set, 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 as $(\exists x \in \mathbb{R})$ $(x^2 = 4)$.

How do we prove a there exists statement? We need to find an element of the set 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,

$$\begin{gathered}
\neg \forall x P(x) = \exists \neg P(X) \\
\neg \exists x P(x) = \forall \neg P(X)
\end{gathered}$$

The negations of the statements above are (note that we use De Morgan's laws, as well as the negation of an if, then statement).

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 $\lvert f(x) - f(x_0)\rvert \geq \epsilon$

What do these mean in English?

1.2 Types of proof

1.2.1 Direct Proof

Approach: Use the definition and known results.

Example 1.2 *The product of an even number with another integer is even.*

Approach: use the definition of even.

Definition 1.3 *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 $n = 2j + 1$.*

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. □

Definition 1.4 *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 1.5 *Let $a, b, c \in \mathbb{Z}$ with $a \neq 0$. Prove that if $a|b$ and $b|c$, then $a|c$.*

Proof. 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

Example 1.6 *If an integer squared is even, then the integer is itself even.*

How would you approach this proof?
$P \implies Q$ $\qquad\qquad$ $\neg P \implies \neg Q$

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

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

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

Example 1.7 *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. $\square$

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

Theorem 1.8 (Well-ordering principle for $\mathbb{N}$) *Every nonempty set of natural numbers has a least element.*

Theorem 1.9 (Principle of mathematical induction) *Let n_0 be a non-negative integer. Suppose P is a property 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 $n \geq n_0$

Note: Principle of strong mathematical induction: For every integer $k \geq n_0$, if $P(n)$ is true for every $n = n_0, \ldots, k$, then $P(k+1)$ is true.

Example 1.10 $n! > 2^n$ *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

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

$\square$

Example 1.11 *Every integer $n \geq 2$ can be written as the product of primes.*

Proof. We prove this by 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.
First, if $k + 1$ is prime then we are done.

Otherwise, if $k + 1$ is not prime, by definition it can be written as a product of some integers a, b such that $1 < a, b < k + 1$. By the induction hypothesis, a and b can both be written as products of primes, so we are done. $\square$

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 6 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 $a, n \in \mathbb{Z}_{>0}$ (positive integers), then $a \leq b$.
   (ii) If $a|b$ and $a|c$, then $a|(xb + yc)$, where $x, y \in \mathbb{Z}$.
   (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 Gerstein [1], though many of the examples are my own. Lakins [2] is also a great resource, but sadly it 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 and 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 $a < b \cup \{-\infty, \infty\}$.*
Open interval: $(a, b) := \{x \in \mathbb{R} : a < x < b\}$
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*

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

$$\bigcap_{\alpha \in A} S_\alpha := \{x : x \in S_\alpha \forall \alpha \in A\}$$

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 $x \in S_\alpha^c \ \forall \alpha \in A$. 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 is 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.

2.2 Ordered sets

Definition 2.12 *A* relation *R on a set X is a subset of $X \times X$. 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* $C([0, 1], \mathbb{R}) := f : [0, 1] \to \mathbb{R} : f$ *is continuous.*
For two function $f, g \in C([0, 1], \mathbb{R})$*, 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. $\mathbb{R}$ does not have this property.

2.3 Functions

One way to define a function is as follows [3, Definition 1.1.14]:

Definition 2.18 *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?

Example 2.19 *For a set* X*, the identity function is:*

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

Definition 2.20 (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* $f^{-1}(B) := \{x : f(x) \in B\}$

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 2.21 (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 2.22 *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 2.23 *Let* $f : X \to Y$ *and* $A \subseteq X$*. Prove that* $A \subseteq f^{-1}(f(A))$*, with equality iff* 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

Definition 2.24 *The* cardinality *of a set A, denoted $|A|$, is the number of elements in the set.*

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

Proposition 2.25 *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. □

Definition 2.26 *Two sets A and B have same cardinality, $|A| = |B|$, if there exists bijection $f : A \to B$.*

Example 2.27 *Which is bigger, $\mathbb{N}$ or $\mathbb{N}_0$?*
Intuitively (at least to me), 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 2.28 *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 2.29 (Cantor-Schröder-Bernstein) *Let A, B, be sets. If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.*

Proof is omitted. See [3, Theorem 1.2.7]

Example 2.30 $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$

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 a 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. □

Definition 2.31 *Let A be a set.*

1. *A is* finite *if there exists an $n \in \mathbb{N}$ and a bijection $f : \{1, \ldots, n\} \to A$*

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 2.32 *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:*

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

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* $\mathbb{N} \times \mathbb{N}$ *(exercise), 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 2.33 *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 $r = 0.\sigma_1\sigma_2\sigma_3\dots$ 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 $f(n) = 0.\sigma_1(n)\sigma_2(n)\sigma_3(n)\dots$. This means we can list out the binary expansions, for example like

$$\begin{aligned}
f(1) =& 0.00000000\dots \\
f(2) =& 0.11111111111\dots \\
f(3) =& 0.0101010101\dots \\
f(4) =& 0.1010101010\dots
\end{aligned}$$

We will construct a number $\tilde{r} \in (0, 1)$ that is not in the image of f. Define $\tilde{r} = 0.\tilde{\sigma}_1\tilde{\sigma}_2\dots$, 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. $\square$

Proposition 2.34 *(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. $\square$

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 $\mathfrak{c}$. There is even a relationship between them:

Proposition 2.35 $\mathfrak{c} = 2^{\aleph_0}$*, i.e. the cardinality of* $\mathbb{R}$ *is the same as the cardinality of* $\mathcal{P}(\mathbb{N})$*.*

This proof is omitted; see [3, Proposition 1.2.9].

2.5 Exercises

1. Let $A = \{x \in \mathbb{R} : x < 100\}$, $B = \{x \in \mathbb{Z} : |x| \geq 20\}$, and $C = \{y \in \mathbb{N} : y \text{ is prime}\}$. 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 $x_1 \leq y_1$ a partially ordered set?
3. [3, 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 $(x, y) \in \mathbb{R}$. Show that two equivalence classes are either disjoint or identical.
4. Let $f : X \to Y$ be defined by the map $x \mapsto \sin(x)$. For what choices of X and Y is f injective, surjective, bijective, or neither?
5. Show that for sets $A, B \subseteq X$ and $f : X \to Y$, $f(A \cap B) \subseteq f(A) \cap f(B)$.
6. Let $f : X \to Y$ and $B \subseteq Y$. Prove that $f(f^{-1}(B)) \subseteq B$, with equality iff f is surjective.
7. Prove that $f(\cup_{i \in I} A_i) = \cup_{i \in I} f(A_i)$.
8. Show that $\mathbb{N}$ and $\mathbb{Z}$ have the same cardinality.
9. Show that $|(0,1)| = |(1, \infty)|$.

2.6 References

The content in this section comes following texts:
A Taste of Topology [3]
The first chapter in Laurent Marcoux's Real Analysis notes (University of Waterloo) [4]
The first chapter of Piotr Zwiernik's *Lecture notes in Mathematics for Economics and Statistics* [5]

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 $x, y, z \in X, d(x, z) \leq (d, y) + d(y, z)$

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

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

$$d(x, y) = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}$$

is a metric space

Definition 3.3 *A* norm *on a linear 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 linear space with a norm is called a normed space. A normed space can be turned into 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 $x = (x_1, \ldots, x_n)$ 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 is

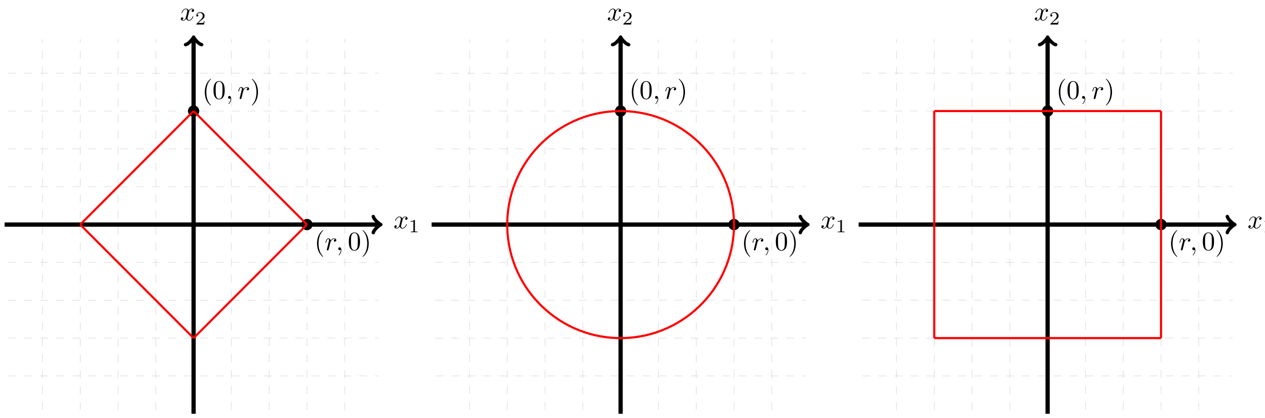
$$||f||_p = \max_{x \in [0,1]} |f(x)|.$$

Definition 3.5 *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 3.6 *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.*

Definition 3.7 (Open and closed sets) *Let (X, d) be a metric space.*



(a) 1-norm (taxicab metric) (b) 2-norm (Euclidean metric) (c) ∞-norm

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

- *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 $K \subseteq X$ is* closed *if $K^c := X \setminus K$ is open.*

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

Proposition 3.8 *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.

(i) 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.

□

We immediately have the following corollary:

Corollary 3.9 *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.

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

Definition 3.10 (Interior and closure) *Let $A \subseteq X$ where (X, d) is a metric space.*

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

Proposition 3.11 *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 A and also by definition, $x \notin \partial A$.

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 $B_\epsilon(x) \cap A \neq \emptyset$ or $B_\epsilon(x) \cap A^c \neq \emptyset$. Since $x \in B_\epsilon(x) \cap A$ for any ϵ, the former cannot be true for any ϵ. Therefore $\exists \epsilon > 0$ such that $B_\epsilon(x) \cap A^c \neq \emptyset$, i.e. $B_\epsilon(x) \subseteq A$. Thus $x \in \mathring{A}$. □

Definition 3.12 *A subset A of a metric space (X, d) is* bounded *if there exists $M > 0$ such that $d(x, y) < M$ for all $x, y \in A$.*

3.2 Sequences

Definition 3.13 *Let (X, d) be a metric space. A* sequence *is a 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 3.14 *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 there $x_n \in B_\epsilon(x)$, and since $x_n \in A$, $B_\epsilon(x) \cap A \neq \emptyset$. Thus $x \in \overline{A}$. □

Corollary 3.15 *A set $K \subseteq X$, where (X, d) is a metric space, is closed if and only if every sequence in K which converges in X converges to a point in K.*

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

Definition 3.16 *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 uncountably many points in A.*

The following result also follows from Proposition 3.14.

Corollary 3.17 *$x \in X$ 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 $x_n \to x$.*

Proposition 3.18 *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 3.19 (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 3.20 *Let (X, d) be a metric space, and let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence in X. The $(x_n)_{n\in\mathbb{N}}$ 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. □

Definition 3.21 *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 3.22 ([3, 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 3.14, 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 3.20, $(y_n)_{n\in\mathbb{N}}$ 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 (Section 3.6, exercise 6), $y = y'$. Thus Y is closed by Corollary 3.15. □

3.2.2 Subsequences

Definition 3.23 *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 3.24 *The sequence $\left((-1)^n\right)_{n\in\mathbb{N}}$ 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 limit points 1 and -1, respectively.*

Proposition 3.25 *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 3.26 *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 3.27 ([3, 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 have 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 a $N \in \mathbb{N}$ 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 3.28 *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. *$f^{-1}(A) \subseteq f^{-1}(B)$*
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 3.27(iii), there exists a δ_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 3.27(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".

□

Definition 3.29 *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 $d_Y(f(x_1), f(x_2))) < K d_X(x_1, x_2)$*

Proposition 3.30 *Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$.*

$$f \text{ is Lipschitz continuous } \Rightarrow f \text{ is uniformly continuous } \Rightarrow f \text{ is continuous}$$

The proof is left as an exercise.

Definition 3.31 *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 3.32 *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 iff it is Lipschitz continuous with constant $K < 1$.

Theorem 3.33 *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 [6, p.240] for the proof as well as more details on how to find the fixed point.

Example 3.34 *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 3.22) 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 Equivalence of metrics

Definition 3.35 (Equivalent metrics) *Two metrics d_1 and d_2 on a set X are* equivalent *if the identity maps from (X, d_1) to (X, d_2) and from (X, d_2) to (X, d_1) are continuous.*

Proposition 3.36 *Two metrics d_1, d_2 on a set X are equivalent iff they have the same open sets or the same closed sets.*

Definition 3.37 *Two metrics d_1 and d_2 on a set X are* strongly equivalent *if for every $x, y \in X$, there exists constants $\alpha > 0$ and $\beta > 0$ such*

$$\alpha d_1(x,y) \leq d_2(x,y) \leq \beta d_1(x,y).$$

If two metrics are strongly equivalent then they are equivalent. The proof of this is one of the exercises.

Example 3.38 *We show that the Euclidean distance (induced by 2-norm) and the metric induced by the ∞-norm are equivalent on $\mathbb{R}^n$.*

Let $||x-y||_2 = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}$ be the Euclidean metric and $||x-y||_\infty = \max_{j=1,\dots,n} |x_j - y_j|$ be the ∞-norm metric. We have

$$||x-y||_2 = \sqrt{\sum_{j=1}^n (x_j - y_j)^2} \leq \sqrt{n \max_{j=1,\dots,n} (x_j - y_j)^2} = \sqrt{n} \max_{j=1,\dots,n} |x_j - y_j| = \sqrt{n}\, ||x-y||_\infty$$

and

$$||x-y||_\infty = \max_{j=1,\dots,n} |x_j - y_j| = \sqrt{\max_{j=1,\dots,n} (x_j - y_j)^2} \leq \sqrt{\sum_{j=1}^n (x_j - y_j)^2} = ||x-y||_2\,.$$

Thus the two metrics are strongly equivalent.

3.5 Extra properties of $\mathbb{R}^n$

[**EK:** *left- and right- continuous on $\mathbb{R}$,*
series, Conditional and absolute convergence, Convergence tests
sup, inf lim sup, lim inf, sup norm?]

3.6 Exercises

1. Show that the infinity norm $||x||_\infty$, $x \in \mathbb{R}^n$, defined in Example 3.4 is a norm.

2. Let (X, d) be any metric space, and define $\tilde{d} : X \times X \to \mathbb{R}$ by

$$\tilde{d}(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \quad x, y \in X.$$

   Show that $\tilde{d}$ is a metric on X.

3. Let X be a set and define $d\colon X \times X \to \mathbb{R}$ by $d(x, x) = 0$ and $d(x, y) = 1$ for $x \neq y \in X$. Prove that d is a metric on X. What do open balls look like for different radii $r > 0$? What does an arbitrary open set look like?

4. Following up on Proposition 3.8 and Corollary 3.9: Show that the infinite intersection of open sets may not be open and that the infinite union of closed sets may not be closed.

5. Find the closure, interior, and boundary of the following sets using Euclidean distance:

   (i) $\{(x, y) \in \mathbb{R}^2 : y < x^2\} \subseteq \mathbb{R}^2$

   (ii) $[0, 1) \times [0, 1) \subseteq \mathbb{R}^2$

   (iii) $\{0\} \cup \{1/n\colon n \in \mathbb{N}\} \subseteq \mathbb{R}$

6. Prove the following: Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space (X, d) that converges to a point $x \in X$. Then x is unique.

7. Prove the following: Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}$ such that $x_n \to x$ and and $y_n \to y$, with $\alpha, x, y, \in \mathbb{R}$.

   (i) Show that $\alpha\, x_n \to \alpha\, x$.

   (i) Show that $x_n + y_n \to x + y$.

8. Show that discrete metric spaces (i.e. those with the metric from exercise 3) are complete.

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

$$f \text{ is Lipschitz continuous } \Rightarrow f \text{ is uniformly continuous } \Rightarrow f \text{ is continuous.}$$

   Provide examples to show that the other directions does not hold.

10. Show that the function $f(x) = \frac{1}{2}\left(x + \frac{5}{x}\right)$ has a unique fixed point on $(0, \infty)$. What is it? (Hint: you will have to restrict the interval.)

11. Prove the following: If two metrics are strongly equivalent then they are equivalent.

3.7 References

The content in this section comes following texts:
A Taste of Topology [3]
Real Mathematical Analysis [6]
The first chapter of Piotr Zwiernik's *Lecture notes in Mathematics for Economics and Statistics* [5]

4 Topology

4.1 Basic definitions

Let X be a set. If X is not a metric space, can we still have open and closed sets? This motivates the concept of a topology. One can think of a topology on X as a specification what subsets of X are open. From the previous section (metric spaces), we already saw some properties of open and closed sets. This motivates the following definition.

Definition 4.1 *Let $\mathcal{T} \subseteq \mathcal{P}(X)$. We call $\mathcal{T}$ a* topology *on X if the following holds:*

(i) $\emptyset, X \in \mathcal{T}$

(ii) If $\{U_\alpha\}_{\alpha \in A} \subseteq \mathcal{T}$, then $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{T}$ ($\mathcal{T}$ is closed under arbitrary unions)

(iii) Let $n \in \mathbb{N}$. If $U_1, \dots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^n U_i \in \mathcal{T}$ ($\mathcal{T}$ is closed under finite intersections)

If $U \in \mathcal{T}$, we call U open. *We call $U \subseteq X$* closed*, if $U^c \in \mathcal{T}$. We call $(X, \mathcal{T})$ a* topological space.

Alternatively we could have specified closed sets, and obtained similar axioms using De Morgan's rules.

Example 4.2 *For a set X, the following $\mathcal{T} \subseteq \mathcal{P}(X)$ are examples of topologies on X.*

- *Trivial topology: $\mathcal{T} = \{\emptyset, X\}$,*
- *Discrete topology: $\mathcal{T} = \mathcal{P}(X)$,*
- *Topology induced by metric: i.e. if d is a metric on X we can define*

  $$\mathcal{T}_d = \{U \subseteq X \mid \forall x \in U \, \exists \epsilon > 0 \text{ such that } B_\epsilon(x) \subseteq U\}.$$

  The discrete topology is also induced by a metric, can you guess which one?
- $\mathcal{T} = \{U \subseteq X \,:\, U^c \text{ is finite}\} \cup \emptyset$

Given a topological space $(X, \mathcal{T})$ and a subset $Y \subseteq X$ we can restrict the topology on X to Y which leads to the next definition.

Definition 4.3 (Relative topology) *Given a topological space $(X, \mathcal{T})$ and an arbitrary non-empty subset $Y \subseteq X$, we define the relative topology on Y as follows*

$$\mathcal{T}|_Y = \{U \cap Y \,:\, U \in \mathcal{T}\}.$$

Recall that in the metric space setting, we had set theoretic descriptions of closures and interiors of sets. We will generalize this in the next definition.

Definition 4.4 *Let $(X, \mathcal{T})$ be a topological space and let $A \subseteq X$ be any subset.*

- *The* interior *of A is $\mathring{A} := \{a \in A \colon \exists U \in \mathcal{T} \text{ s.t. } U \subseteq A \text{ and } a \in U\}$.*
- *The* closure *of A is $\overline{A} := \{x \in X \colon \forall U \in \mathcal{T} \text{ with } x \in U, U \cap A \neq \emptyset\}$.*
- *The* boundary *of A is $\partial A := \{x \in X \colon \forall U \in \mathcal{T} \text{ with } x \in U, U \cap A \neq \emptyset \text{ and } U \cap A^c \neq \emptyset\}$.*

One can see that the definitions are taken fairly verbatim from the metric space setting, except that we are now looking at arbitrary open sets given by the topology instead of balls of the form $B_\epsilon(x)$. Similar to the metric space setting, the interior is the largest open set contained in A, whereas the the closure is the smallest closed set that contains A.

Example 4.5 *Let $X = \{a, b, c\}$ and $\mathcal{T} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, X\}$. Then the following holds*

- $\mathring{\{a\}} = \{a\}$,

- $\mathring{\{c\}} = \emptyset$,
- $\overline{\{a\}} = \{a, c\}$,
- $\overline{\{c\}} = \{c\}$.

Note that even though we do not necessarily have a characterization of closures in terms of limits of sequences as in metric spaces for arbitrary topological spaces, there exists an alternative characterization that still holds in this general setting (and thus in particular also for metric spaces).

Proposition 4.6 (Proposition 3.1.18 in [3]) *Let $(X, \mathcal{T})$ be a topological space and $A \subseteq X$. Then,*

$$\overline{A} = \bigcap \{F \colon F \text{ is closed and } A \subseteq F\}.$$

Proof. For convenience define $A' = \bigcap\{F \colon F \text{ is closed and } A \subseteq F\}$. We will show $\overline{A} \subseteq A'$ by showing that $(A')^c \subseteq \overline{A}^c$ (contrapositive). Suppose $x \notin A'$. Then, since the closure is closed, $(A')^c$ is an open set containing x. But since $A \subseteq A'$ we have $A \cap (A')^c = \emptyset$, showing that $x \notin \overline{A}$.
Conversely, assume $X \notin \overline{A}$. Then there exists an open set U with $x \in U$ such that $U \cap A = \emptyset$. Thus, $A \subseteq U^c$. Since U^c is closed, we have by the definition of closure $A' \subseteq U^c$ and since $x \notin U^c$, we have $x \notin A'$. Thus, $A' \subseteq \overline{A}$. □

Similarly, one can show $\mathring{A} = \bigcup\{U : U \text{ is open and } U \subseteq A\}$. Hence, we see that the interior of A is the largest open set contained in A and the closure is the smallest closed set that contains A.

Another important concept in topology (and thus also in metric spaces) is density.

Definition 4.7 *Let $(X, \mathcal{T})$ be a topological space. A subset $A \subseteq X$ is called* dense, *if $\overline{A} = X$*

Using the definition of closure, we see that $A \subseteq X$ is dense if and only if for all $U \in \mathcal{T}$, $U \cap A \neq \emptyset$.

Example 4.8

- *The rationals $\mathbb{Q}$ are dense in the reals $\mathbb{R}$.*
- *The only dense subset in $(X, \mathcal{P}(X))$ is X itself.*
- *Any non-empty subset is dense in $(X, \{\emptyset, X\})$.*

The concept of a dense subset allows us to look at it instead of the whole space. In the metric space setting, this means that elements in X can be approximated arbitrarilyw well with elements from the dense subset.

Definition 4.9 *A topological space $(X, \mathcal{T})$ is* separable *if there exists a countable dense subset.*

As stated in the previous example the rationals are dense in $\mathbb{R}$, and since $\mathbb{Q}$ is countable, $\mathbb{R}$ is separable. We could extend this example to $\mathbb{R}^n$. However, if we look at all bounded real-valued sequences with the metric induced by the supremums norm, this space fails to be countable.

Example 4.10 *Define $\ell_\infty = \{(x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{R},\ \sup_{n\in\mathbb{N}} |x_n| < \infty\}$, the space of bounded real valued sequences. We can endow ℓ_∞ with a metric induced by the supremum norm, namely $d((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \sup_{n\in\mathbb{N}} |x_n - y_n|$. Then ℓ_∞ is not separable with respect to the topology induced by this metric. To see this, for each $M \subseteq \mathbb{N}$ define*

$$e_n^M = \begin{cases} 1 & \text{if } n \in A, \\ 0 & \text{otherwise}, \end{cases}$$

for $n \in \mathbb{N}$. Then if $M_1, M_2 \subseteq \mathbb{N}$ with $M_1 \neq M_2$, $d((e_n^{M_1})_{n\in\mathbb{N}}, (e_n^{M_2})_{n\in\mathbb{N}}) = 1$. Thus the open balls $B_{1/3}((e_n^{M_1})_{n\in\mathbb{N}})$, $B_{1/3}((e_n^{M_2})_{n\in\mathbb{N}})$ are disjoint for all $M_1, M_2 \subseteq \mathbb{N}$ with $M_1 \neq M_2$ (check using contradiction and triangle inequality).
Now suppose towards contradiction $A \subseteq \ell^\infty$ is dense and countable. Then by density, for all open sets U, $U \cap A \neq \emptyset$. In particular for all $M \subseteq \mathbb{N}$, $B_{1/3}((e_n^M)_{n\in\mathbb{N}}) \cap A \neq \emptyset$. However, there are uncountably many such M (see the cardinality of $\mathcal{P}(\mathbb{N})$ in the set theory section), but only countably many elements in A. Since the balls are disjoint, this is a contradiction.

4.2 Compactness

We will start off by giving the definition of an important separation axiom.

Definition 4.11 *A topological space $(X, \mathcal{T})$ is called Hausdorff if for all $x \neq y \in X$ there exist open sets U_x, U_y with $x \in U_x$ and $y \in U_y$ such that $U_x \cap U_y = \emptyset$.*

Example 4.12

- *Let (X, d) be a metric space. Then $(X, \mathcal{T}_d)$ is Hausdorff, where $\mathcal{T}_d$ is the topology induced by the metric d. Why? If $x \neq y \in X$, then choose $\epsilon := d(x, y) > 0$ and thus $U_x = B_{\epsilon/2}(x)$ and $U_y = B_{\epsilon/2}(y)$ are disjoint.*

- *Let X be an infinite set and $\mathcal{T} = \{U \subseteq X : U^c \text{ is finite }\}$. Then $(X, \mathcal{T})$ is not Hausdorff. Why? Suppose in order to derive a contradiction that it is Hausdorff and take $x \neq y \in X$. Then there exist open sets U_x, U_y with $x \in U_x$ and $y \in U_y$ such that $U_x \cap U_y = \emptyset$. However since $U_x \cap U_y$ is open, its complement X is finite; a contradiction.*

Definition 4.13 *Let $(X, \mathcal{T})$ be a topological space and $K \subseteq X$. An collection $\{U_i\}_{i \in I}$ of open sets is called* open cover *of K if $K \subseteq \cup_{i \in I} U_i$. The set K is called* compact *if for all open covers $\{U_i\}_{i \in I}$ there exists a finite subcover, meaning there exits an $n \in \mathbb{N}$ and $\{U_1, \ldots, U_n\} \subseteq \{U_i\}_{i \in I}$ such that $K \subseteq \cup_{i=1}^n U_i$.*

Proposition 4.14 ([3, Proposition 3.3.6]) *Let $(X, \mathcal{T})$ be a topological space and take a non-empty subset $K \subseteq X$. The following holds:*

1. *If X is compact and K is closed then K is compact (i.e. closed subsets of compact sets are compact).*
2. *If $(X, \mathcal{T})$ is Hausdorff, then K being compact implies that K is closed.*

Proof. 1. We need to show that any open cover of K has a finite subcover. Let $\{U_i\}_{i \in I}$ be an open cover of K. Then, since K^c is open, $\{U_i\}_{i \in I} \cup K^c$ is an open cover of X. Since X is compact there exists a finite subcover. There are two possibilities, the finite subcover is either of the form $\{U_1, \ldots, U_n, K^c\}$ or $\{U_1, \ldots, U_n\}$. In either case, $\{U_1, \ldots, U_n\}$ is a finite subcover for K. Hence, K is compact.
2. We will show that K^c is open. Let $x \in K^c$. Then, since X is Hausdorff, for all $y \in K$ there exist disjoint open sets $U_{x,y}$ and U_y with $x \in U_{x,y}, y \in U_y$. We would like to take the intersection over the $U_{x,y}$ to obtain an open set containing x and contained in K^c (note that the union would not work since we want to guarantee that the resulting set is still contained in K^c). However, since K is compact and $\{U_y\}_{y \in K}$ is an open cover of K there exist $y_1, \ldots, y_n$ such that $K \subseteq \cup_{i=1}^n U_{y_i}$. Thus $\tilde{U} = \cap_{i=1}^n U_{x,y_i}$ is open and $x \in \tilde{U}$ with $\tilde{U} \subseteq K^c$. Thus, K^c is open. ☐

In undergraduate math classes you may have seen an equivalent definition for compactness on $\mathbb{R}^n$. This is a nice feature of Euclidean space.

Theorem 4.15 (Heine-Borel Theorem) *Let $K \subseteq \mathbb{R}^n$. Then K is compact with respect to the topolgy induced by the Euclidean distance if and only if it is closed and bounded.*

The proof is omitted. See [3, Corollary 2.5.12].

Just as we had a sequential characterization of the closure of a set in metric spaces, we similarly have a sequential characterization of compactness.

Theorem 4.16 *Let (X, d) be a metric space. Then $K \subset X$ is compact with respect to the metric induced by d if and only if every sequence in K admits a subsequence converging to some point in K.*

Again the proof is omitted. See [3, Theorem 2.5.10]. A corollary of this statement together with Heine-Borel is the Bolzano-Weierstrass theorem.

Corollary 4.17 (Bolzano-Weierstrass) *Any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

4.3 Continuity

Lastly, we will discuss continuity in this general setting.

Definition 4.18 *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A map $f: X \to Y$ is called* continuous *if for all $U \in \mathcal{T}_Y$, $f^{-1}(U) \in \mathcal{T}_X$, i.e. the preimage of open sets is open.*

We can also specify continuity at a point $x_0 \in X$.

Definition 4.19 *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. A map $f: X \to Y$ is called* continuous at $x_0 \in X$ *if for all $U \in \mathcal{T}_Y$ with $f(x_0) \in U$, $f^{-1}(U) \in \mathcal{T}_X$, i.e. the preimage of open sets containing $f(x_0)$ is open (and contains x_0).*

The next proposition is, in a certain sense, a generalization of the extreme value theorem to topological spaces.

Proposition 4.20 *Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces. Suppose $K \subset X$ is compact and let $f: K \to Y$ be continuous. Then $f(K)$ is compact.*

Proof. Let $\{U_i\}_{i \in I}$ be an open cover of $f(K)$, i.e. $f(K) \subseteq \bigcup_{i\in I} U_i$. Then $f^{-1}(f(K)) \subseteq f^{-1}(\bigcup_{i \in I} U_i)$ (check!). By Exercise 6 in **??**, we have $f^{-1}(\bigcup_{i\in I} U_i) = \bigcup_{i \in I} f^{-1}(U_i)$ and by Proposition 2.23, we have $K \subseteq f^{-1}(f(K))$. Hence, we obtain $K \subseteq \bigcup_{i\in I} f^{-1}(U_i)$. Since f is continuous, each $f^{-1}(U_i)$ is open and thus $\{f^{-1}(U_i)\}_{i \in I}$ is an open cover of K. Since K is compact, there exist $f^{-1}(U_1), \dots, f^{-1}(U_n)$ such that $K \subseteq \bigcup_{i=1}^n f^{-1}(U_i)$. Then $f(K) \subseteq f(\bigcup_{i=1}^n f^{-1}(U_i)) = \bigcup_{i=1}^n f(f^{-1}(U_i)) \subseteq \bigcup_{i=1}^n U_i$, where we use that images preserve set inclusions (check!), and Exercises 5 and 6 in **??**. Thus, $\{U_1, \dots, U_n\}$ is a finite subcover for $f(K)$ and $f(K)$ is compact. □

As you can see, a lot of results from introductory real analysis or calculus have extensions to a more general topological setting. However, topology is a large field with many powerful tools that we do not have time to cover. The final result in this section is an important result in topology.

Definition 4.21 *A topological space $(X, \mathcal{T})$ is* normal *if the following hold:*

(i) For all $x \in X$, $\{x\}$ is closed.

(ii) For all disjoint closed sets $F_1, F_2 \subseteq X$ there exist disjoint open sets U_1, U_2 such that $F_1 \subseteq U_1$ and $F_2 \subseteq U_2$.

Example 4.22

- *Any metric space (X, d) is normal.*
- *Let $(X, \mathcal{T})$ be Hausdorff and compact. Then X is normal.*
- *If $(X, \mathcal{T})$ is normal, then it is Hausdorff.*

Theorem 4.23 (Urysohn's Lemma) *Let $(X, \mathcal{T})$ be normal and $F_1, F_2 \subseteq X$ be closed with $F_1 \cap F_2 = \emptyset$. Then there exists a continuous function $f: X \to [0, 1]$ such that $f(F_1) = \{0\}$ and $f(F_2) = \{1\}$. (Here the topology on $[0, 1]$ is the relative topology inherited from the usual metric topology on $\mathbb{R}$.)*

Proof omitted. See [3, Theorem 4.1.2].

4.4 Exercises

1. Let $(X, \mathcal{T})$ be a topological space. Prove that $A \subseteq X$ is closed if and only if $\overline{A} = A$.
2. Let $(X, \mathcal{T})$ be a topological space and $\{A_i\}_{i\in I}$ be a collection of subsets of X. Show that

$$\bigcup_{i \in I} \overline{A_i} \subseteq \overline{\bigcup_{i\in I} A_i}.$$

   Show that if the collection is finite, the two sets are equal.

3. Let $(X, \mathcal{T})$ be a topological space and $\{A_i\}_{i \in I}$ be a collection of subsets of X. Prove that

$$\overline{\bigcap_{i \in I} A_i} \subseteq \bigcap_{i \in I} \overline{A_i}.$$

   Find a counterexample that shows that equality is not necessarily the case.

4. Let $(X, \mathcal{T})$ be a topological space and $A \subseteq X$ be dense. Show that if $A \subseteq B \subseteq X$, then B is dense as well.

5. Let $(X, \mathcal{T})$ be a Hausdorff topological space. Show that the singleton $\{x\}$ is closed for all $x \in X$. Hint: Show that the complement is open.

6. Let $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ and $(Z, \mathcal{T}_Z)$ be topological spaces and let $f \colon X \to Y$, $g \colon Y \to Z$ be continuous. Show that $g \circ f \colon X \to Z$ is continuous as well.

7. Let (X, d) be a metric space and $K \subset X$ compact. Show that for all $\epsilon > 0$ there exists $\{x_1, x_2, \ldots, x_n\} \subseteq K$ such that for all $y \in K$ we have $d(y, x_i) < \epsilon$ for some $i = 1, \ldots, n$.

4.5 References

The content in this section comes following texts:
A Taste of Topology [3]
The first chapter in Laurent Marcoux's Real Analysis notes (University of Waterloo) [4]

5 Linear Algebra

5.1 Vector spaces

5.1.1 Axioms of a vector space

Let V be a set and let $\mathbb{F}$ be a field.

Definition 5.1 *We call V a **vector space** if the following hold:*
Addition:

- *(A) Commutativity in addition: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in V$*
- *(B) Associativity in addition: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$*
- *(C) Existence of a neutral element, addition: There exists a vector $\mathbf{0}$ such that for any $\mathbf{v} \in V$, $\mathbf{0} + \mathbf{v} = \mathbf{v}$*
- *(D) Additive inverse: For every $\mathbf{v} \in V$, there exists another vector, which we denote $-\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.*

Multiplication by a scalar:

- *(E) Existence of a neutral element, multiplication: For any $\mathbf{v} \in V$, $1 \times \mathbf{v} = \mathbf{v}$*
- *(F) Associativity in multiplication: Let $\alpha, \beta \in \mathbb{F}$. For any $\mathbf{v} \in V$, $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$*

Associativity:

- *(G) Let $\alpha \in \mathbb{F}, \mathbf{u}, \mathbf{v} \in V$. $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \beta\mathbf{v}$.*
- *(H) Let $\alpha, \beta \in \mathbb{F}, \mathbf{v} \in V$. $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$.*

Elements of the vector space are called vectors.
Most often we will assume $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$.
Examples of vector spaces: $\mathbb{R}^n$. $\mathbb{C}^n$, $M_{m \times n}$ (matrices of size $m \times n$), $\mathbb{P}_n$ (polynomials of degree n, $p(x) = a_0 + a_1 x + \ldots + a_n x^n$).

Lemma 5.2 *For every* $\mathbf{v} \in V$*, we have* $-\mathbf{v} = (-1) \times \mathbf{v}$*.*

Proof. Our goal is to show that $(-1) \times \mathbf{v}$ is the additive inverse of $\mathbf{v}$. We show this as follows:

$$\mathbf{v} + (-1) \times \mathbf{v} = \mathbf{v} \times (1 + (-1)) = \mathbf{v} \times 0 = 0$$

The last step uses Exercise 5.8. [**EK:** *Do by hand in class*] ☐

5.1.2 Subspaces

Definition 5.3 *A subset* U *of* V *is called a* ***subspace*** *of of* V *if* U *is also a vector space (using the same addition and scalar multiplication as on* V*).*

Proposition 5.4 *A subset* U *of* V *is a subspace of* V *if and only if* U *satisfies the following three conditions:*

1. $\mathbf{0} \in U$
2. *Closed under addition:* $u, w \in U$ *implies* $\mathbf{u} + \mathbf{v} \in U$
3. *Closed under scalar multiplication:* $\alpha \in \mathbb{F}$ *and* $u \in U$ *implies* $\alpha \mathbf{u} \in U$

Proof. $\Rightarrow$ If U is a subspace of V, then U satisfies these 3 properties by Definition 5.1.
$\Leftarrow$ Suppose U satisfies the given 3 conditions. Then for any $\mathbf{v} \in U$, there must exist $-\mathbf{v} \in U$ by property 3, since $-\mathbf{v} = (-1) \times \mathbf{v}$ by Lemma 5.2 (property D). Property 1 assures property C. Properties 2 and 3, and the fact that $U \subset V$, assure the remaining properties hold. ☐

This characterisation allows us to easily show that the intersection of subspaces is again a subspace.

Proposition 5.5 *Let* V *be a vector space and let* $U_1, U_2 \subseteq V$ *be subspaces. Then* $U_1 \cap U_2$ *is also a subspace of* V*.*

Proof. We use the characterization in Proposition 5.4. First, since $\mathbf{0} \in U_1$ and $\mathbf{0} \in U_2$, we have $\mathbf{0} \in U_1 \cap U_2$. Second, for $\mathbf{u}, \mathbf{v} \in U_1 \cap U_2$, since in particular $\mathbf{u}, \mathbf{v} \in U_1$ and $\mathbf{u}, \mathbf{v} \in U_2$ and U_1, U_2 are subspaces, $\mathbf{u} + \mathbf{v} \in U_1$ and $\mathbf{u} + \mathbf{v} \in U_2$. Thus, $\mathbf{u} + \mathbf{v} \in U_1 \cap U_2$. Similarly, one shows $\alpha \mathbf{u} \in U_1 \cap U_2$ for $\alpha \in \mathbb{F}$. ☐

On the contrary the union of two subspaces is not a subspace in general (see Exercise 5.12). However, the next definition introduces the smallest subspace containing the union.

Definition 5.6 *Suppose* $U_1, ..., U_m$ *are subsets of* V*. The sum of* $U_1, ..., U_m$*, denoted* $U_1 + ... + U_m$*, is the set of all possible sums of elements of* $U_1, ..., U_m$*. More precisely,*

$$U_1 + ... + U_m = \{\mathbf{u}_1 + ... + \mathbf{u}_m : \mathbf{u}_1 \in U_1, ..., \mathbf{u}_m \in U_m\}$$

Proposition 5.7 *Suppose* $U_1, ..., U_m$ *are subspaces of* V*. Then* $U_1 + ... + U_m$ *is the smallest subspace of* V *containing* $U_1, ..., U_m$*.*

5.1.3 Exercises

Exercise 5.8 (1.1.7 in [7]) Show that $0\mathbf{v} = \mathbf{0}$ for $\mathbf{v} \in V$.

Exercise 5.9 (1.B.1 in [8]) Show that $-(-v) = v$ for $\mathbf{v} \in V$.

Exercise 5.10 (1.B.2 in [8]) Suppose that $\alpha \in \mathbb{F}, \mathbf{v} \in V$, and $\alpha \mathbf{v} = 0$. Prove that $a = 0$ or $v = 0$.

Exercise 5.11 (1.B.4 in [8]) Why is the empty space not a vector space?

Exercise 5.12 (7.4.1 in [7]) Let U_1 and U_2 be subspaces of a vector space V. Prove that $U_1 \cup U_2$ is a subspace of V if and only if $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$.

5.2 Linear (in)dependence and bases

Definition 5.13 *A linear combination of vectors* $\mathbf{v}_1, ..., \mathbf{v}_n$ *in* V *is a vector of the form*

$$\alpha_1\mathbf{v}_1 + ... + \alpha_n\mathbf{v}_n = \sum_{k=1}^{n} \alpha_k\mathbf{v}_k$$

where $\alpha_1, ..., \alpha_m \in \mathbb{F}$.

Definition 5.14 *The set of all linear combinations of a list of vectors* $v_1, ..., v_m$ *in* V *is called the* ***span*** *of* $v_1, ..., v_m$, *denoted span*$\{\mathbf{v}_1, ..., \mathbf{v}_n\}$. *In other words,*

$$span\{\mathbf{v}_1, ..., \mathbf{v}_n\} = \{\alpha_1\mathbf{v}_1 + ... + \alpha_m\mathbf{v}_n : \alpha_1, ..., \alpha_n \in \mathbb{F}\}$$

The span of the empty list is defined to be $\{\mathbf{0}\}$.

Definition 5.15 *A system of vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *is called a basis (for the vector space* V *) if any vector* $\mathbf{v} \in V$ *admits a unique representation as a linear combination*

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \sum_{k=1}^{n} \alpha_k\mathbf{v}_k$$

In undergrad, you likely thought about this as: the equation $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$, where the x_i are unknown, has a unique solution.
Example of bases:
For $\mathbb{R}^n$: $e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ e_n = (0, \ldots, 0, 1)$
For $\mathbb{P}^n$: $1, x, x^2, \ldots, x^n$

Definition 5.16 *The linear combination* $\alpha_1\mathbf{v}_1 + ... + \alpha_n\mathbf{v}_n$ *is called trivial if* $\alpha_k = 0$ *for every* k.

Proposition 5.17 *A system of vectors* $\mathbf{v}_1, \ldots \mathbf{v}_n \in V$ *is a basis if and only if it is linearly independent and complete (generating).*

[**EK:** *Proof done by hand*]

5.2.1 Exercises

From Harvard: Exercise: Suppose v_1, v_2, v_3, v_4 (a) spans V and (b) is linearly independent. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also (a) spans V and (b) is linearly independent.

Exercise: Suppose $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, ..., v_m + w$ is linearly dependent, then $w \in \text{span}(v_1, ..., v_m)$.
Exercise: Suppose that $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Show that $v_1, ..., v_m, w$ is linearly independent if and only if

$$w \notin \text{span}(v_1, ..., v_m)$$

onExercises Exercise: Suppose v_1, v_2, v_3, v_4 (a) spans V and (b) is linearly independent. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also (a) spans V and (b) is linearly independent.
Exercise: Suppose $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, ..., v_m + w$ is linearly dependent, then $w \in \text{span}(v_1, ..., v_m)$.
Exercise: Suppose that $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Show that $v_1, ..., v_m, w$ is linearly independent if and only if

$$w \notin \text{span}(v_1, ..., v_m)$$

[**EK:** *Add a few from books*]

5.3 Linear transformations

Definition 5.18 *A **map** T from domain X to codomain Y is a rule that assigns an output $y = T(x) \in Y$ to each input $x \in X$*

Definition 5.19 *A map from a vector space U to a vector space V is **linear** if*

$$T(\alpha \mathbf{u} + \beta \mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v}) \quad \textit{for any } \mathbf{u}, \mathbf{v} \in V,\ \alpha, \beta \in \mathbb{F}$$

Let's denote the set of all linear maps from vector space U to vector space V by $\mathcal{L}(U, V)$.

Example 5.20 (Differentiation is a linear map) *Let $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$, (i.e. D is a linear map from the polynomials on $\mathbb{R}$ to the polynomials on $\mathbb{R}$), defined as $Dp = p'$. The fact that such a map is linear follows from basic facts about derivatives, i.e. $\frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha f'(x) + \beta g'(x)$.*

Other examples: integration, rotation of vectors, reflection of vectors

Lemma 5.21 *Let $T \in \mathcal{L}(U, V)$. Then $T(0) = 0$.*

Proof. By linearity, $T(0) = T(0 + 0) = T(0) + T(0)$. Add $-T(0)$ to both sides to obtain the result. □

Theorem 5.22 *Let $S, T \in \mathcal{L}(U, V)$ and $\alpha \in \mathbb{F}$. $\mathcal{L}(U, V)$ is a vector space with addition defined as the sum $S + T$ and multiplication as the product αT.*

Definition 5.23 (Product of linear maps) *Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. We define the product $ST \in \mathcal{L}(U, W)$ for $\mathbf{u} \in U$ as $ST(\mathbf{u}) = S(T(\mathbf{u}))$.*

Definition 5.24 *Let $T : U \to V$ be a linear transformation. We define the following important subspaces:*

- Kernel or null space*:* $\ker T = \{\mathbf{u} \in U : T\mathbf{u} = 0\}$
- Range $\text{range } T = \{\mathbf{v} \in V : \exists \mathbf{u} \in U \textit{ such that } \mathbf{v} = T\mathbf{u}\}$

The dimensions of these spaces are often called the following:

- Nullity $\text{nullity}(T) = \dim(\ker(T))$
- Rank $\text{rank}(T) = \dim(\text{range}(T))$

Example 5.25 *The null space of the differentiation map (see Example 5.20) is the set of constant functions.*

Definition 5.26 (Injective and surjective) *Let $T : U \to V$. T is* injective *if $T\mathbf{u} = T\mathbf{v}$ implies $\mathbf{u} = \mathbf{v}$ and T is* surjective *if $\forall \mathbf{u} \in U$, $\exists \mathbf{v} \in V$ such that $\mathbf{v} = T\mathbf{u}$, i.e. if $\text{range}\, T = V$.*

Theorem 5.27 *$T \in \mathcal{L}(U, v)$ is injective $\iff \ker T = 0$.*

Proof. $\Rightarrow$ Suppose T is injective. By Lemma 5.21, we know that 0 is in the null space of T, i.e. $T(0) = 0$. Suppose $\exists \mathbf{v} \in \ker T$. Then $T(\mathbf{v}) = 0 = T(0)$, and by injectivity, $\mathbf{v} = 0$.
$\Leftarrow$ Suppose $\ker T = 0$. Let $T\mathbf{u} = T\mathbf{v}$; we want to show $\mathbf{v} = \mathbf{u}$.
$T\mathbf{u} = T\mathbf{v} \implies T(\mathbf{u} - \mathbf{v}) = 0$, which implies $\mathbf{u} - \mathbf{v} \in \ker T$. But $\ker T = 0$, so then $\mathbf{u} - \mathbf{v} = 0 \implies \mathbf{u} = \mathbf{v}$. □

Theorem 5.28 (Rank Theorem) *For a matrix A or equivalently a linear transformation $A : \mathbb{F}^n \to \mathbb{F}^m$:*

$$\text{rank}\, A = \text{rank}\, A^T$$

Theorem 5.29 *Rank Nullity Theorem Let $T : U \to V$ be a linear transformation, where U and V are finite-dimensional vector spaces. Then*

$$\text{rank}\, T + \text{nullity}\, T = \dim U.$$

5.3.1 Exercises

Exercise 5.30 Let $T \in \mathcal{L}(\mathbb{P}(\mathbb{R}), \mathbb{P}(\mathbb{R}))$ be the map $T(p(x)) = x^2 p(x)$ (multiplication by x^2).

(i) Show that T is linear.

(ii) Find $\ker T$.

5.4 Linear maps and matrices

We can use matrices to represent linear maps.

Definition 5.31 *Let $T \in \mathcal{L}(U, V)$ where U and V are vector spaces. Let $u_1, \ldots, u_n$ and $v_1, \ldots, v_m$ be bases for U and V respectively. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ with entries $A_{i,j}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ defined by*

$$Tu_k = A_{1,k}v_1 + \cdots + A_{m,k}v_m$$

i.e. the kth column of A is the scalars needed to write Tu_k as a linear combination of the basis of V:

$$Tu_k = \sum_{i=1}^{m} A_{i,k}v_i$$

Example 5.32 *Let $D \in \mathcal{L}(\mathcal{P}_5(\mathbb{R}), \mathcal{P}_4(\mathbb{R}))$ be the diiferentiation map, $Dp = p'$. Find the matrix of D with respect to the standard bases of $\mathcal{P}_4(\mathbb{R})$ and $\mathcal{P}_5(\mathbb{R})$.*
Standard basis: $1, x, x^2, x^3, x^4, (x^5)$
$T(u_1) = (1)' = 0$
$T(u_2) = (x)' = 1$
$T(u_3) = (x^2)' = 2x$
$T(u_4) = (x^3)' = 3x^2$
$T(u_5) = (x^4)' = 4x^3$

$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

This way of looking at matrices gives us an intuitive explanation for why we do matrix multiplication the way we do! Let $T : U \to V$ and $S : V \to W$, where T, S are linear maps and U, V, W are vector spaces with bases $u_1, \ldots, u_n$, $v_1, \ldots, v_m$, and $w_1, \ldots, w_p$. If we want to have

$$\mathcal{M}(ST) := \mathcal{M}(S)\mathcal{M}(T),$$

how would we need to define matrix multiplication?
Let $A = \mathcal{M}(S)$ and $B = \mathcal{M}(T)$. Then

$$(ST)u_k = S(T(u_k)) = S(Bu_k) = S(b_k) = Ab_k,$$

where b_k is the kth column of B.
We also have $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$ when $S, T \in \mathcal{L}(U, V)$.

5.5 Determinants

5.6 Inner product spaces

[EK:

- *transpose, adjoint*

- *inner products and norms*
- *Orthogonal matrices and unitary matrices*
- *Orthonormalization and Gram-Schmidt*
- *Orthogonal projections and decompositions*

]

5.7 Spectral theory

Note: here we will assume $\mathbb{F} = \mathbb{C}$, so that we are working on an algebraically closed field.
Let $T\colon V \to V$ be a linear map, where V is a vector space. We would like to describe the action of this linear map in a particularly "nice" way. For example, if there exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V such that $T\mathbf{v}_i = \alpha_i \mathbf{v}_i$ where $\alpha_i \in \mathbb{F}$ for $i = 1, \ldots, n$, then T acts on this basis merely by scaling the basis vectors. If we look at the matrix of T with respect to this basis, T is a diagonal matrix with α_i in the diagonal.

Definition 5.33 *Let V be a vector space. Given a linear map $T\colon V \to V$ and $\alpha \in \mathbb{F}$, α is called an **eigenvalue** of T if there exists a non-zero vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ such that $Tv = \alpha v$. We call such $\mathbf{v}$ an **eigenvector** of T with eigenvalue α. We call the set of all eigenvalues of T **spectrum** of T and denote it by $\sigma(T)$.*

[**EK:** *Define just for matrices?*]
Note that $Tv = \alpha v$ can be rewritten as $(T - \alpha I)v = 0$. Thus, if α is an eigenvalue, the map $T - \alpha I$ is not invertible, since it must have non-trivial kernel. Using the known characterizations of invertability, this gives the following characterization for eigenvalues.

Theorem 5.34 *Let V be a vector space and $T\colon V \to V$ be a linear map and let A_T be a matrix representation of T. The following are equivalent*

1. *$\alpha \in \mathbb{F}$ is an eigenvalue of T,*
2. *$(A_T - \alpha I)\mathbf{x} = 0$ has a non-trivial solution,*
3. *$\det(A_T - \alpha I) = 0$.*

Theorem 5.35 *Suppose A is a square matrix with distinct eigenvalues $\alpha_1, \ldots, \alpha_k$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors corresponding to these eigenvalues. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent.*

Proof. Induction on k. □

Hence, if all the eigenvalues are distinct, there exists a basis of eigenvectors. This gives the next result.

Corollary 5.36 *If a $A \in M_n(\mathbb{C})$ has n distinct eigenvalues, then A is diagonalizable. That is there exists an invertible matrix $U \in M_n(\mathbb{C})$ such that $A = UDU^{-1}$, where D is a diagonal matrix with the eigenvalues of A in the diagonal.*

[**EK:** *mention problem of eigenspaces not having high enough dimension when eigenvalues are repeated, possibly mention geometric and algebraic multiplicity*]

Theorem 5.37 *Let $A \in M_n(\mathbb{C})$ be a Hermitian matrix. Then, there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that $A = UDU^*$, where D is a diagonal matrix with the eigenvalues of A in the diagonal. Furthermore, all eigenvalues of A are real.*

Proof. It suffices to show that there exists an orthogonal basis of eigenvectors and that the eigenvalues are real. We will prove the former by induction. □

Note that in the previous theorem, the orthogonality of the eigenvectors is special. In general, even if a matrix is diagonalizable, there might not exists a orthogonal eigenbasis. The next theorem states a characterization of matrices that exhibit an orthogonal eigenbasis.

Theorem 5.38 *A matrix A is diagonalizable by a unitary matrix if and only if $AA^* = A^*A$. We call such a matrix **normal**.*

Proof omitted.
[**EK:**

- *Characteristic and minimal polynomials*
- *Cayley-Hamilton theorem*
- *Jordan Canonical Form (and block matrices)*
- *Matrix norms?*

]

5.7.1 Exercises

Exercise 5.39 Let $A, U \in M_n(\mathbb{F})$ be matrices, where U is invertible. Show that $\sigma(A) = \sigma(UAU^{-1})$.

Exercise 5.40 Let $A \in M_n(\mathbb{C})$ be an invertible matrix with $\sigma(A) = \{\alpha_1, \ldots, \alpha_n\}$ counted with multiplicities. Determine $\sigma(A^{-1})$, $\sigma(A^T)$ and $\sigma(A^*)$.

5.8 Matrix decomposition

[**EK:**

- *singular value decomposition*
- *LU & QR decompositions*
- *PCA*

]

5.9 References

The following texts:
Linear Algebra Done Right [8]
Linear Algebra Done Wrong [7]

6 Calculus

6.1 Differentiation

[**EK:** *Topics:*

- *k-th order differentiability and smoothness*
- *Local extrema and derivative tests*
- *L'Hopital's rule*
- *Mean value theorem*
- *Ordinary differential equations*

]

6.2 Integration

6.3 Exercises

1. Consider the space of continuous functions on the unit interval, $C([0,1])$. Prove that there exists a unique $f \in C([0,1])$ such that for all $x \in [0,1]$

$$f(x) = x + \int_0^x s f(s) \mathrm{d}s.$$

Hint: You can use that $C([0,1])$ is a complete metric space with respect to the supremums metric $d_\infty(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|$ for $f, g \in C([0,1])$.

[**EK:** *Topics:*

- *The Riemann integral*
- *Construction via Riemann sums*
- *Fundamental theorem of Calculus*
- *Riemann-Stieltjes integrals*

]

6.4 Exercises

6.5 References

7 Multivariable calculus

7.1 Differentiation

[**EK:** *Topics:*

- *Gradients, directional derivatives, Lagrange multipliers*
- *Partial derivatives, Hessians, Jacobians*
- *Taylor's theorem*
- *Inverse function theorem*
- *Multivariate chain rule*

]

7.2 Implicit Function Theorem

7.3 Integration

[**EK:** *Topics:*

- *Change-of-variables*
- *Exchange of integrals*

]

7.4 Exercises

7.5 References

References

1. Gerstein LJ. Introduction to Mathematical Structures and Proofs. Undergraduate Texts in Mathematics. 2012. Available from: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/978-1-4614-4265-3l
2. Lakins TJ. The Tools of Mathematical Reasoning. Pure and Applied Undergraduate Texts. 2016
3. Runde V. A Taste of Topology. Universitext. 2005. Available from: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/0-387-28387-0
4. Marcoux LW. PMATH 351 Notes. 2019. Available from: https://www.math.uwaterloo.ca/~lwmarcou/notes/pmath351.pdf
5. Zwiernik P. Lecture notes in Mathematics for Economics and Statistics. 2022. Available from: http://84.89.132.1/~piotr/docs/RealAnalysisNotes.pdf
6. Pugh CC. Real Mathematical Analysis. Undergraduate Texts in Mathematics. 2015. Available from: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/978-3-319-17771-7
7. Treil S. Linear Algebra Done Wrong. 2017. Available from: https://www.math.brown.edu/streil/papers/LADW/LADW.html
8. Axler S. Linear Algebra Done Right. 3rd ed. Undergraduate Texts in Mathematics. Springer, 2015. Available from: https://link-springer-com.myaccess.library.utoronto.ca/book/10.1007/978-3-319-11080-6