richardwu.ca

PMATH 351 Course Notes

Real Analysis

Brian Forrest • Fall 2018 • University of Waterloo

Last Revision: September 28, 2018

Table of Contents

Abstract

These notes are intended as a resource for myself; past, present, or future students of this course, and anyone interested in the material. The goal is to provide an end-to-end resource that covers all material discussed in the course displayed in an organized manner. These notes are my interpretation and transcription of the content covered in lectures. The instructor has not verified or confirmed the accuracy of these notes, and any discrepancies, misunderstandings, typos, etc. as these notes relate to course's content is not the responsibility of the instructor. If you spot any errors or would like to contribute, please contact me directly.

1 September 10, 2018

1.1 Basic notation

We denote

$$\begin{aligned}
\mathbb{N} &= \{1, 2, 3, \ldots\} \\
\mathbb{Z} &= \{\ldots, -2, -1, 0, 1, 2, \ldots\} \\
\mathbb{Q} &= \{\frac{n}{m} \mid n \in \mathbb{Z}, m \in \mathbb{N}\} \\
\mathbb{R} &= \text{real numbers}
\end{aligned}$$

We use $\subset$ and $\subseteq$ interchangeably, and use $\subsetneq$ for strict subsets. $\subset$ or $\subseteq$ is called "inclusion", and $\supset$ or $\supseteq$ is called "containment".

1.2 Basic set theory

We denote X as our universal set. If $\{A_\alpha\}_{alpha \in I}$ is such that $A_\alpha \subset X$ for all $\alpha \in I$ (index set), then

$$\bigcup_{\alpha \in I} A_\alpha = \{x \in X \mid x \in A_\alpha \text{ for some } \alpha \in I\} \qquad \text{(union)}$$

$$\bigcap_{\alpha \in I} A_\alpha = \{x \in X \mid x \in A_\alpha \text{ for all } \alpha \in I\} \qquad \text{(intersection)}$$

Define for $A, B \subseteq X$

$$\begin{aligned}
A \setminus B &= \{x \in X \mid x \in A, x \notin B\} && \text{(set difference)} \\
A \Delta B &= \{x \in X \mid x \in A \text{ and } x \notin B\} \text{ OR } x \in B \text{ and } x \notin A\} && \text{(semantic difference)} \\
A^c &= X \setminus A = \{x \in X \mid x \notin A\} && \text{(complement)} \\
&\varnothing && \text{(empty set)} \\
P(X) &= \{A \mid A \subset X\} \quad \varnothing \in P(X), X \in P(X) && \text{(power set)}
\end{aligned}$$

1.3 De Morgan's laws

De Morgan's laws states that given $\{A_\alpha\}_{\alpha\in I} \subset P(X)$

$$\left(\bigcup_{\alpha\in I} A_\alpha\right)^c = \bigcap_{\alpha\in I} A_\alpha^c$$
$$\left(\bigcap_{\alpha\in I} A_\alpha\right)^c = \bigcup_{\alpha\in I} A_\alpha^c$$

Question: what if $I = \varnothing$, what is $\bigcup_{\alpha\in\varnothing} A_\alpha$? It is in fact $\bigcup_{\alpha\in\varnothing} A_\alpha = \varnothing$.
Note that $\bigcap_{\alpha\in\varnothing} A_\alpha = X$ (from De Morgan's Law, and also $A_\alpha = A_\alpha^c$).

1.4 Products of sets, relations, and functions

Given X, Y define the product

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

If $X = \{x_1, \dots, x_n\}$, $Y = \{y_1, \dots, y_m\}$ then $X \times Y = \{(x_i, y_j) \mid i = 1, \dots, n \quad j = 1, \dots, m\}$ containing nm elements.

Definition 1.1 (Relation)**.** A **relation** on X, Y is a subset R of the product $X \times Y$.
We write xRy if $(x, y) \in R$. The **domain** of R is

$$\{x \in X \mid \exists y \in Y \text{ with } (x, y) \in R\}$$

which need not cover our universal set.
The **range** of R is

$$\{y \in Y \mid \exists x \in X \text{ with } (x, y) \in R\}$$

Definition 1.2 (Function (as a relation))**.** A **function** from X into Y is a relation R such that for every $x \in X$, there exists exactly one $y \in Y$ with $(x, y) \in R$.

Suppose that we have $X_1, X_2, \dots, X_n$ non-empty sets. Define

$$X_1 \times X_2 \times \ldots \times X_n = \prod_{i=1}^{n} X_i = \{(x_1, x_2, \dots, x_n) \mid x_i \in X_i\}$$

or a set of *n-tuples*.
If $X_i = X_j = X$ for all $i, j = 1, \dots, n$, then

$$\prod_{i=1}^{n} X_i = \prod_{i=1}^{n} X = X^n$$

Problem 1.1. Given a collection $\{X_\alpha\}_{\alpha\in I}$ of non-empty sets, what do we mean by $\prod_{\alpha\in I} X_\alpha$?
Motivation: consider $X_1 \times \ldots \times X_n = \{(x_1, \dots, x_n) \mid x_i \in X_i\}$. We choose some $(x_1, \dots, x_n) \in \prod_{i\in\{1,\dots,n\}} = I$.
This point induces *a function*

$$f_{(x_1,\dots,x_n)} : \{1, \dots, n\} \to \bigcup_{i=1}^{n} X_i$$

with $f(1) = x_1 \in X_1$, $f(i) = x_i \in X_i$, $f(n) = x_n \in X_n$, etc. Assume we have have $f : \{1, \ldots, n\} \to \bigcup_{i=1}^{n} X_i$ such that $f(i) \in X_i$. Then

$$(f(1), f(2), \ldots, f(n)) = \prod_{i=\{1,\ldots,n\}} X_i$$

Definition 1.3 (Product of sets). Given a collection $\{X_\alpha\}_{\alpha\in I}$ of non-empty sets we let

$$\prod_{\alpha\in I} X_\alpha = \{f : I \to \bigcup_{\alpha\in I} X_\alpha\}$$

such that $f(\alpha) \in X_\alpha$ (i.e $\prod_{\alpha\in I} X_\alpha$ is a "set of functions"). f is called a **choice function**.
Question: If $X_\alpha \neq \varnothing$, is $\prod_{\alpha\in I} X_\alpha \neq \varnothing$?

2 September 12, 2018

2.1 Zermelo's Axiom of Choice

Question: If $\{X_\alpha\}_{\alpha\in I}$ is a non-empty collection of non-empty sets is

$$\prod_{\alpha\in I} X_\alpha \neq \varnothing$$

This is analogous to saying: given a collection of non-empty sets in $\mathbb{R}$, how would you choose an element from each subset of $\mathbb{R}$? This is easy if they were subsets of $\mathbb{N}$ (take the least element which exists by the *well-ordering principle*) but much more difficult in $\mathbb{R}$.

Axiom 2.1 (Zermelo's Axiom of Choice). If $\{X_\alpha\}_{\alpha\in I}$ is a non-empty collection of non-empty sets, then $\prod_{\alpha\in I} \neq \varnothing$.

Equivalently we have an analogous version:

Axiom 2.2 (Axiom of Choice V2). If $X \neq \varnothing$, then there exists a function

$$f : P(X) \setminus \{\varnothing\} \to X$$

such that $f(A) \in A$ for all $A \in P(X) \setminus \{\varnothing\}$ (we can always pick out a subset ($e \in P(X)$) from a non-empty set A).

2.2 Properties of relations

Definition 2.1 (Relation properties). A relation R on X (i.e. $R \subseteq X \times X$) is

1. **reflexive** if $x\ R\ x$ for all $x \in X$
2. **symmetric** if $x\ R\ y \Rightarrow y\ R\ x$
3. **anti-symmetric** if $x\ R\ y$ and $y\ R\ x$, then $x = y$
4. **transitive** if $x\ R\ y$ and $y\ R\ z$ implies $x\mathbb{R}z$

2.3 Partially and totally ordered sets

Example 2.1. Let $X = \mathbb{R}$. We have $x\ R\ y$ iff $x \leq y$.
Note that $\leq$ is reflexive, anti-symmetric, and transitive.

Example 2.2. Let $Y \neq \varnothing$ and $X = P(Y)$.
We write $A\ R\ B$ iff $A \subseteq B$.
Note that $\subseteq$ is reflexive, anti-symmetric, and transitive.

Example 2.3. Let $Y \neq \varnothing$ and $X = P(Y)$.
We write $A\ R\ B$ iff $B \subseteq A$.
Note that $\subseteq$ is reflexive, anti-symmetric, and transitive.

Definition 2.2 (Partially ordered sets)**.** A set X with a relation R on X is called a **partially ordered set** if R is

1. reflexive
2. anti-symmetric
3. transitive

(R is a **partial order** on X if it satisfies these three conditions).
We write (X, R) and call this a **poset**.

Definition 2.3 (Totally ordered sets)**.** If (X, R) is a poset, then if $A \subseteq X$ and $R_1 = R_{|A \times A}$ then (A, R_1) is a poset. We say (A, R_1) is **totally ordered** if for each $x, y \in A$ *either* $x\ R\ y$ or $y\ R\ x$. We also call totally ordered sets **chains**.

How many partial orderings can we have for a given set X (i.e. the number of ways to define partial order relations)?

Example 2.4. Let $X = \{x\}$. We have one relation $R = \{(x, x)\}$ (from $X \times X$) and thus 1 partial ordering.

Example 2.5. Let $X = \{x, y\}$. We know posets $(X, \preceq)$ must be reflexive, thus we have one relation where $x \preceq x$ and $y \preceq y$.
We can also have a poset with the reflexive relations above as well as $x \preceq y$. Similarly we can have a poset with $y \preceq x$.

Example 2.6. Let $X = \{x, y, z\}$.

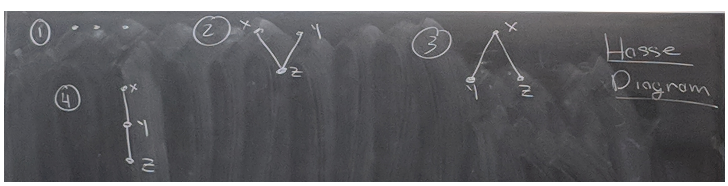

Figure 2.1: Hasse diagrams for the possible $(X, \preceq)$ posets (an edge downwards from a to b denotes $a \preceq b$; note reflexive $a \preceq a$ is assumed automatically).

We have the poset with just the reflexive relations $e \preceq e$ for $e \in X$.
We have the poset with the reflexive relations and $x \preceq z$ and $y \preceq z$ (3 posets with permutations).
We have the poset with the reflexive relations and $x \preceq y$ and $x \preceq z$ (3 posets with permutations).
We have the poset with the reflexive relations and $x \preceq y$ and $y \preceq z$ (6 posets with permutations).
We have the poset with the reflexive relations and $y \preceq z$ (6 posets with permutations, not shown in diagram above).

Note that when identifying these posets isomorphisms, we should not draw lines between two elements $a \leq b$ if the transitive property already implies that. For example if we had the chain $a \leq b \leq c$, the diagram with a line from a to c would be redundant (thus we will end up double counting).

2.4 Bounds on posets

Definition 2.4 (Upper and lower bounds)**.** Let $(X, \preceq)$ be a partially ordered set.
Let $A \subset X$. We say that $x_0 \in X$ is an **upper bound** for A if $x \preceq x_0$ for all $x \in A$.
If A has an upper bound, we say it is **bounded above**.
If A is bounded above then x_0 is the **least upper bound** if

1. x_0 is an upper bound of A
2. If y is an upper bound of A then $x_0 \preceq y$.

We write $x_0 = \text{lub}(A)$ or $x_0 \sup(A)$ (supremum).
If $x_0 = \text{lub}(A) \in A$, then x_0 is the *maximum* in A.
Similarly we define the same for lower bounds (infimum).

Example 2.7. Let $X = \mathbb{R}$ and $\preceq$ the usual ordering.

Fact 2.1. Every non-empty subset that is bounded above has a least upper bound (LUBP (lub property) for $\mathbb{R}$).

Example 2.8. Let $Y \neq \varnothing$, $X = P(Y)$, and $\preceq$ be $\subseteq$ (ordering by inclusion).
Y is the maximum element of $(X, \subseteq)$.
If $\{A_\alpha\}_{\alpha \in I} \subset P(X)$ is bounded above by Y, but note that

$$\text{lub}(\{A_\alpha\}_{\alpha \in I}) = \bigcup_{\alpha \in I} A_\alpha$$
$$\text{glb}(\{A_\alpha\}_{\alpha \in I}) = \bigcap_{\alpha \in I} A_\alpha$$

Recall that if $I = \varnothing$, then the glb is all of $\mathbb{R}$: this is in fact correct (it's the greatest set that is a lower bound for relation $\subseteq$).

3 September 14, 2018

3.1 Maximal

Definition 3.1. Let $(X, \preceq)$ be a partially ordered set. An element $x \in X$ is **maximal** if whenver $y \in X$ such that $x \preceq y$, we must have $x = y$.

Example 3.1. Suppose we have $x \preceq x$, $y \preceq y$, and $z \preceq z$. Then all of x, y, z are maximal.
Suppose we have $x \preceq z$ and $y \preceq z$ (as well as the reflexive relations). Then only z is maximal.
Suppose we have $x \preceq y$ and $x \preceq z$ (as well as the reflexive relations). Then y and z are maximal.
Suppose $x \preceq y \preceq z$ (and transitives). Only z is maximal.
Suppose $x \preceq y$ (and transitives). Then both y and z are maximal.
For $X \neq \varnothing$ and $(P(X), \subseteq)$, X is maximal.
For $X \neq \varnothing$ and $(P(X), \supseteq)$, $\varnothing$ is maximal.
For $(\mathbb{R}, \leq)$ has no maximal element.

3.2 Zorn's Lemma

Axiom 3.1 (Zorn's Lemma)**.** If $(X, \preceq)$ is a non-empty partially ordered set such that every chain $S \subset X$ has an upper bound. Then $(X, \preceq)$ has a maximal element.

We can apply Zorn's Lemma to prove a fundamental linear algebra theorem:

Theorem 3.1. Every non-zero vector space V has a basis.

Proof. Let $= \{A \subset X \mid A \text{ is linear indep.}\}$. Note $\neq \varnothing$ because $V \neq \{0\}$.
Order with $\subseteq$.
A basis is a maximal element in $(, \subseteq)$ (if we add vector to this basis, it would be a linear combination of the basis vectors by definition of a basis).
Let $S = \{A_\alpha\}_{\alpha \in I}$ be a chain in . Let $A_0 = \bigcup_{\alpha \in I} A_\alpha$.
Choose $x_1, \ldots, x_n \in A_0$ distinct elements. Assume that $\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$. But $x_i \in A_{\alpha_i}$ and we can assume that

$$A_{\alpha_1} \subseteq A_{\alpha_2} \subseteq \ldots \subseteq A_{\alpha_n} \Rightarrow \{x_1, \ldots, x_n\} \subset A_{\alpha_n}$$

So $\alpha_i = 0$ for all $i = 1, \ldots, n$, thus A_0 is an upper bound of S. By Zorn's Lemma we have a basis. □

3.3 Well-ordered

Definition 3.2 (Well-ordered)**.** We say that a partially ordered set $(X, \preceq)$ is **well-ordered** if every non-empty subset A of X has a least element in A.
For example, $(\mathbb{N}, \preceq)$ is well-ordered.
Note that if a set is well-ordered it must also be totally ordered (how would you compare some arbitrary element to the least element if the set was not well-ordered?)

Axiom 3.2 (Well-Ordering Principle)**.** Every non-empty set of $\mathbb{Z}^+$ can be well-ordered.

Theorem 3.2. The following are equivalent:

1. Axiom of Choice
2. Zorn's Lemma
3. Well-Ordering Principle

Example 3.2. Let $X = \mathbb{Q}$. Define the function ϕ

$$\phi(\frac{m}{n}) = \begin{cases} 2^m 5^n & \text{if } m > 0 \\ 1 & \text{if } m = 0 \\ 3^{-m} 7^n & \text{if } m < 0 \end{cases}$$

Note that $\phi : \mathbb{Q} \to \mathbb{N}$ is 1-1. (we could have used any combination of unique primes, as long as we ensure there is a 1-1 mapping).
Note that we can map the rationals to a subset of $\mathbb{N}$, thus the rationals are well-ordered by the Well-Ordering Principle.
Note that we also have $r \leq s \iff \phi(r) \leq \phi(s)$ (ϕ is an order isomorphism).

3.4 Equivalence relations and partitions

Definition 3.3 (Equivalence relation)**.** Let X be non-empty. A relation $\sim$ on X is an **equivalence relation** if the relation is

1. reflexive
2. symmetric
3. transitive

Observation 3.1. Let $[x] = \{y \in X \mid x \sim y\}$ or the **equivalence class** of X.
Then

1. Either $[x] = [y]$ or $[x] \cap [y] = \varnothing$
2. $X = \bigcup_{x \in X} [x]$

Definition 3.4. Let $X \neq \varnothing$. A **partition** of X is a collection $\{A_\alpha\}_{\alpha \in I} \subset P(X)$ such that

1. $A_\alpha \neq \varnothing$
2. $A_\alpha \cap A_\beta = \varnothing$ if $\alpha \neq \beta$
3. $X = \bigcup_{\alpha \in I} A_\alpha$

Observation 3.2. If $\{A_\alpha\}_{\alpha \in I}$ is a partition of X and $x \sim y$ *iff* $x, y \in A_\alpha$, then $\sim$ is an equivalence relation (i.e. if we start with a partition based on some relation $\sim$, we can show $\sim$ is an equivalence relation).

Example 3.3. How many equivalence relations are there on $X = \{1, 2, 3\}$? We can count the number of partitions:

1. $\{\{1\}, \{2\}, \{3\}\}$
2. $\{\{1, 2, 3\}\}$
3. $\{\{1, 2\}\{3\}\}$ (3 permutations since $\binom{3}{2}$)

Example 3.4. Let X be any set (empty or non-empty). Define $\sim$ on $P(X)$ by $A \sim B$ *iff* there exists $f : A \to B$ that is $1-1$ and onto.
$\sim$ has properties:

reflexive Take $\mathrm{id} : A \to A$ where $\mathrm{id}(x) = x$

symmetric If we have $f : A \to B$ then we have $f^{-1} : B \to A$ since f is bijective.

transitive If we have $f : A \to B$ and $g : B \to C$, then we have $g \circ f : A \to B$

thus $\sim$ is an equivalence relation.
For $X = \{1, 2, 3\}$, we have four equivalence classes on $P(X)$: one for every possible subset size $(0, \ldots, 3)$.

4 September 17, 2018

4.1 Cardinality

Definition 4.1 (Equivalence of sets)**.** We say that two sets X and Y are **equivalent** if there exists a *1-1 and onto* function $f : X \to Y$. We write $X \sim Y$.

Definition 4.2 (Cardinality)**.** If $X \sim Y$, we say that the two sets have the same **cardinality** and write $|X| = |Y|$.

Definition 4.3 (Finite sets)**.** X is **finite** if $X = \varnothing$ or if $X \sim \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. If $X \sim \{1, \ldots, n\}$ we say X has cardinality n and write $|X| = n$. We let $|\varnothing| = 0$.

Definition 4.4 (Infinite sets)**.** X is **infinite** if it is not finite.

Example 4.1. We know $\mathbb{N}$ is infinite. We claim $\{2, 3, \ldots\}$ is also infinite.
Note that $f : \mathbb{N} \to \{2, 3, \ldots\}$ where $f(n) = f(n+1)$ is a 1-1 and onto map, thus $\mathbb{N} \sim \{2, 3, \ldots\}$ so $\{2, 3, \ldots\}$ is infinite as well.

4.2 Pigeonhole Principle

Question 4.1. If $n \neq m$, can $\{1, \ldots, n\} \sim \{1, \ldots, m\}$?

Theorem 4.1 (Pigeonhole Principle)**.** The set $\{1, \ldots, n\}$ is **not** equivalent to any proper subset.

Proof. We prove this by induction on n.

Base case Note that $\{1\} \not\sim \varnothing$.

Inductive step Assume the statement holds for $\{1, \ldots, k\}$ for some k.

Suppose that we had a 1-1 function $f : \{1, 2, \ldots, k, k+1\} \to \{1, 2, \ldots, k, k+1\} \backslash \{m\}$ for some $m \in \{1, \ldots, k+1\}$. We have one of two possibilities:

$m = k + 1$ Then

$$f_{|\{1,\ldots,k\}} : \{1, \ldots, k\} \overset{1-1}{\to} \{1, \ldots, k\} \setminus \{f(k+1)\}$$

where $f_{|A}$ is restrict of f to A.

Thus $f_{|\{1,\ldots,k\}}$ is a 1-1 onto function to a proper subset of $\{1, \ldots, k\}$ (since $f(k+1)$ must map to one of $\{1, 2, \ldots, k, k+1\} \setminus \{m\} = \{1, \ldots, k\}$), which is a contradiction of inductive hypothesis.

$m \neq k + 1$ Assume that $f(j_0) = k + 1$ and also $m \in \{1, \ldots, k\}$.

Note if $j_0 = k+1$, then $f_{|\{1,\ldots,k\}} : \{1, \ldots, k\} \to \{1, \ldots, k\} \setminus \{m\}$, which is a contradiction of the inductive hypothesis. Thus $j_0 \neq k + 1$ so $f(k + 1) \neq k + 1$.

Let $g : \{1, \ldots, k + 1\} \to \{1, \ldots, k + 1\} \setminus \{m\}$ where

$$g(i) = \begin{cases} k+1 & \text{if } i = k+1 \\ f(k+1) & \text{if } i = j_0 \\ f(i) & \text{if } i \neq k+1, j_0 \end{cases}$$

so g is a 1-1 function where $g(k + 1) = k + 1$, but we already know that such a function cannot exist thus this is impossible.

□

Corollary 4.1. If X is finite, then X is not equivalent to any proper subset.

Proof. If we assume there is a 1-1 and onto $g : X \to A \subsetneq X$, then for some $m \neq n$ we could apply $f(\{1,\dots,m\}) = X$ and $f^{-1}(A) = \{1,\dots,n\}$, thus

$$\{1,\dots,m\} \xrightarrow{f} X \to g \to A \xrightarrow{f^{-1}} \{1,\dots,n\}$$

which would contradict the Pigeonhole principle since $n < m$. □

4.3 Countable

Definition 4.5 (Countable)**.** We say that X is **countable** if either X is finite or if $X \sim \mathbb{N}$.
If $X \sim \mathbb{N}$ we can say that X is **countably infinite** and we write $|X| = |N| = \aleph_0$ or **aleph naught**.

4.4 Infinite sets has countably infinite subset

Proposition 4.1 (Infinite set has countably infinite subset)**.** Every infinite set contains a subset $A \sim \mathbb{N}$.

Proof. Assume X is infinite. Let $f : P(X) \setminus \{\varnothing\} \to X$ where for every $A \subset X$ the Axiom of Choice permits $f(A) \in A$.
Let $x_1 = f(X)$. We define recursively

$$x_{n+1} = f(X \setminus \{x_1,\dots,x_n\})$$

This gives us a sequence $\{x_n\} = \{x_1, x_2, \dots, x_n, \dots\} = A \sim \mathbb{N}$. □

Corollary 4.2. Every infinite set X is equivalent to a proper subset.

Proof. Given X construct $\{x_n\}$ as above. Define $f : X \to X \setminus \{x_1\}$ by

$$f(x) = \begin{cases} x_{n+1} & \text{if } x = x_n \\ x & \text{if } x \notin \{x_n\} \end{cases}$$

thus we have a 1-1 and onto function to a proper subset of X. □

4.5 1-1 and onto duality

Proposition 4.2. The follow are equivalent (TFAE):

1. There exists $f : X \to Y$ that is 1-1
2. There exists $g : Y \to X$ that is onto

Proof. $1 \to 2$ Assume $f : X \to Y$ is 1-1. Define $g : Y \to X$ by

$$g(y) = \begin{cases} x & \text{if } y = f(x) \text{ for some } x \\ x_0 & \text{for some arbitrary } x_0 \in X \end{cases}$$

$2 \to 1$ Let $g : Y \to X$ be onto and let $h : P(Y) \setminus \{\varnothing\}$ be a choice function.

For each $x \in X$ define

$$f(x) = h(g^{-1}(\{x\}))$$

where $g^{-1}(\{x\}) = \{y \in Y_{|g}\}$.

□

4.6 Partial order on cardinalities

Definition 4.6 ($\leq$ relation on cardinalities)**.** Given X, Y we write $|X| \leq |Y|$ if there exists a **1-1 function** $f : X \to Y$ ($f(X) \sim X$).

Observation 4.1. Note that $|\mathbb{N}| \leq |\mathbb{Q}|$ since $f(n) = \frac{n}{1}$ is a 1-1 function $f : \mathbb{N} \to \mathbb{Q}$.
Also $|\mathbb{Q}| \to |\mathbb{N}|$ since

$$g\left(\frac{m}{n}\right) = \begin{cases} 2^m 3^n & \text{if } m > 0 \\ 1 & \text{if } m = 0 \\ 5^{-m} 7^n & \text{if } m < 0 \end{cases}$$

where g is still a function by unique prime factorization of the integers.
Does this imply $|\mathbb{N}| = |\mathbb{Q}|$, that is does $|X| \leq |Y|$ and $|Y| \leq |X|$ imply $|X| = |Y|$?

5 September 19, 2018

Note: theorems marked with (*****) are important and one should be familiar with the proof.

5.1 Cantor-Schröder-Bernstein theorem (*****)

Theorem 5.1 (Cantor-Schröder-Bernstein theorem)**.** Let $A_2 \subset A_1 \subset A_0 = A$. Assume that $A_2 \sim A_0$. Then $A_0 \sim A_1$.
(aside: support $f : X \to Y$ is 1-1 and onto. Let $A \subset B$, then $f(B \setminus A) = f(B) \setminus f(A)$).

Proof. Let $\phi : A_0 \to A_2$ be 1-1 and onto. Let $A_3 = \phi(A_1)$ and $A_4 = \phi(A_2)$.
In fact, we let $A_{n+2} = \phi(A_n)$.
Notice that $A_{n+1} \subset A_n$ for all $n \in Z^+$.
Key observation:

$$A_0 = (A_0 \setminus A_1) \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \ldots \cup \bigcap_{n=0}^{\infty} A_n$$

Similarly, we have

$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \ldots \cup \bigcap_{n=1}^{\infty} A_n$$

We want to show there is a 1-1 and onto mapping between the two expressions for A_0 and A_1.
Notice that the two $\bigcap A_n$ are equivalent since $A_0 \cap \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} A_n$.
We can map $(A_1 \setminus A_2)$ in A_0 to $(A_1 \setminus A_2)$ in A_1, $(A_3 \setminus A_4)$ in both, etc. Note ϕ maps $A_0 \setminus A_1$ to $\phi(A_0) \setminus \phi(A_1) = A_2 \setminus A_3$ (from aside before).
More formally, we define $f : A_0 \to A_1$ by

$$f(x) = \begin{cases} x & \text{if } x \in \bigcap_{n=0}^{\infty} A_n = \bigcap_{n=1}^{\infty} A_n \\ x & \text{if } x \in A_{2k+1} \setminus A_{2k+2} \text{ for } k = 0, 1, \ldots \\ \phi(x) & \text{if } x \in A_{2k} \setminus A_{2k+1} \text{ for } k = 0, 1, \ldots \end{cases}$$

Clearly f is 1-1 and onto, thus $A_0 \sim A_1$. □

Corollary 5.1. If $A_1 \subset A$, $B_1 \subset B$ and $A \sim B_1$ (i.e. $|A| \leq |B|$) $B \sim A_1$ (i.e. $|B| \leq |A|$), then $A \sim B$.

Proof. Let $f : A \to B_1$ and $g : B \to A_1$ be 1-1 and onto functions.
Let $A_2 = g(B_1)$, then $A_2 \subseteq A_1 \subseteq A$. Then $g \circ f : A \to A_2$ is 1-1 and onto so $A \sim A_2$, thus by CSB we have $A \sim A_1 \sim B$. ☐

Example 5.1. Back to the example where we have $|\mathbb{Q}| \leq |\mathbb{N}|$ and $|\mathbb{N}| \leq |\mathbb{Q}|$, by CSB we have $|\mathbb{Q}| = |\mathbb{N}|$.

5.2 Proving countability

Proposition 5.1. If X is infinite then $|X| = |\mathbb{N}| = \aleph_0$ **if and only if** there is a 1-1 function $f : X \to \mathbb{N}$.

Proof. If $|X| = |N|$, then there is a 1-1 and onto function from $f : X \to \mathbb{N}$ by definition.
Assume there exists a 1-1 $f : X \to \mathbb{N}$. Then $|X| \leq |\mathbb{N}|$.
Since X is infinite, there exists a countably infinite subset of cardinality $|\mathbb{N}|$, thus $|\mathbb{N}| \leq |X|$. By CSB we have $|X| = |\mathbb{N}|$. ☐

Example 5.2. Show that $\mathbb{N} \times \mathbb{N}$ is countable.
Let $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be defined as $f(n, m) = 2^n 3^m$. Thus we have a 1-1 function from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$, thus by the previous proposition $\mathbb{N} \times \mathbb{N}$ is countable.

5.3 Uncountability and Cantor's diagonal proof

Definition 5.1 (Uncountable sets)**.** A set X is **uncountable** if X is not countable.

Theorem 5.2 (Cantor)**.** $(0, 1)$ is uncountable.

Proof. Assume that $(0, 1)$ is countable.
We can write

$$\begin{aligned}
a_1 &= 0.a_{11}a_{12}a_{13}\dots \\
a_2 &= 0.a_{21}a_{22}a_{23}\dots \\
&\vdots \\
a_n &= 0.a_{n1}a_{n2}a_{n3}\dots
\end{aligned}$$

and these representations are unique if we do not allow the representations to end in a string of 9's.
We want to construct some number $b \in (0, 1)$ that is not within our countable set.
Let $b = 0.b_1b_2\dots$ where

$$b_n = \begin{cases} 7 & \text{if } a_{nn} \neq 7 \\ 3 & \text{if } a_{nn} = 7 \end{cases}$$

Thus $b \notin$ our set. ☐

Corollary 5.2. $\mathbb{R}$ is uncountable.
Note that $(0, 1) \sim \mathbb{R}$ since we have $f : (0, 1) \to \mathbb{R}$ where

$$f(x) = \tan(\pi x - \frac{\pi}{2})$$

which is a 1-1 and onto function.
We denote $|\mathbb{R}|$ by c.

Question 5.1. Given X, Y: is it always true that either

1. $|X| = |Y|$
2. $|X| < |Y|$
3. $|Y| < |X|$

If we accept AC, the answer is yes.
If we do not accept AC, the answer could be no.

6 September 21, 2018

6.1 Comparibility of cardinals

Theorem 6.1 (Comparibility of cardinals)**.** If X, Y are non-empty then either $|X| \preceq |Y|$ or $|Y| \preceq |X|$.

Proof. Let $S = \{(A, B, f) \mid A \subseteq X, B \subseteq Y, f : A \to B \text{ is 1-1 and onto}\}$ (note $S \neq \varnothing$; take singletons from each X,Y with trivial f).
Order S as follows: $(A_1, B_1, f_1) \preceq (A_2, B_2, f_2)$ if $A_1 \subseteq A_2$, $B_1 \subseteq B_2$ and $f_1 = f_{2|A_1}$ (this is possible since $A_1 \subseteq A_2$ so restriction exists) (if any of the three conditions fail, we cannot order the two triples: this is fine since we are only looking for a partial order).
Let $C = \{(A_\alpha, B_\alpha, f_\alpha)\}_{\alpha \in I}$ be a chain in $(S, \preceq)$.
Let $A_0 = \bigcup_{\alpha \in I} A_\alpha$, $B_0 = \bigcup_{\alpha \in I} B_\alpha$, and $f_0 : A_0 \to B_0$ by $f_0(x) = f_{\alpha_0}(x)$ if $x \in A_{\alpha_0}$ (we find the subset A_{α_0} which the point x we pick out from A_0 is found in: then we take the corresponding function f_{α_0} as our function for that point).
Note: if $x \in A_{\alpha_1}$ and $x \in A_{\alpha_2}$ we can assume taht $(A_{\alpha_1}, B_{\alpha_1}, f_{\alpha_1}) \preceq (A_{\alpha_2}, B_{\alpha_2}, f_{\alpha_2})$ then

$$f_{\alpha_1}(x) = f_{\alpha_2|A_{\alpha_1}}(x) = f_{\alpha_2}(x)$$

thus f is well-defined (it doesn't really matter which f_{α_0} we choose since they're all the same for a given point x).
We need to show

$f_0 : A_0 \to B_0$ **is 1-1** Let $x_1, x_2 \in A_0$, $x_1 \neq x_2$. We may assume that $x_1 \in A_{\alpha_1}$, $x_2 \in A_{\alpha_2}$ with $A_{\alpha_1} \subseteq A_{\alpha_2}$ thus $x_1, x_2 \in A_{\alpha_2}$. Since f_{α_2} is 1-1 ($f_{\alpha_2}(x_1) \neq f_{\alpha_2}(x_2)$) then $f_0(x_1) \neq f_0(x_2)$.

f_0 **is onto** Let $y_0 \in B_0 \Rightarrow y_0 \in B_{\alpha_0}$ for some α_0.

Then there exists $x_0 \in A_{\alpha_0}$ with $f_{\alpha_0}(x_0) = y_0$ (since f_{α_0} is onto), thus $f_0(x_0) = y_0$.

thus (A_0, B_0, f_0) belongs to our set S (since f_0 is 1-1 and onto) and it is an **upper bound** for C (A_0, B_0 are the unions so they're upper bounds for all A_α, B_α, and f restricted to any subset is equivalent to the function on that subset).
Let (A, B, f) be maximal in S by Zorn's Lemma: we have three cases

1. if $A = X$, then $|X| \leq |Y|$ (since we have a 1-1, onto function from X onto $B \subseteq Y$).
2. if $B = Y$, then $|Y| = |A||X|$
3. Suppose $X \setminus A \neq \varnothing$ and $Y \setminus B \neq \varnothing$. Let $x_0 \in X \setminus A$, $y_0 \in Y \setminus B$.
   Let $A^* = A \cup \{x_0\}$, $B^* = B \cup \{y_0\}$, and $f^* : A^* \to B^*$ by
   $$f^*(x) = \begin{cases} f(x) & \text{if } x \in A \\ y_0 & \text{if } x = x_0 \end{cases}$$
   Thus $(A, B, f) \npreceq (A^*, B^*, f^*)$ which is impossible (i.e. this case is impossible).

□

6.2 Cardinal arithmetic

Sum If $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_m\}$ and $X \cap Y = \varnothing$, then $|X| = n$, $|Y| = m$, and $|X \cup Y| = n + m$ obviously.

Definition 6.1. Assume that X and Y are such that $X \cap Y \neq \varnothing$, we define

$$|X| + |Y| = |X \cup Y|$$

(note if we had $X_1 \sim X_2$, $Y_1 \sim Y_2$ and $X_1 \cap Y_1 = \varnothing$ and $X_2 \cap Y_2 = \varnothing$. We have $X_1 \cup Y_1 \sim X_2 \cup Y_2$ since we always have a 1-1 and onto mapping: simply partition points in $X_1 \cup Y_1$ into $x_1 \in X_1$ and $x_1 \in Y_1$: we have 1-1 mappings to each of X_2 and Y_2, respectively).

Question 6.1. What is $\aleph_0 + \aleph_0$?

Let $X = \{2, 4, 6, \ldots\}$ and $Y = \{1, 3, 5, \ldots\}$ then $X \cup Y = \{1, 2, 3, \ldots\}$ thus $\aleph_0 + \aleph_0 = \aleph_0$ by definition.

Question 6.2. What is $c + c$ ($|\mathbb{R}| = c$)?

Let $X = (0, 1) \Rightarrow |X| = c$ and $Y = (1, 2) \Rightarrow |Y| = c$, then

$$c \leq |X| \leq |X| + |Y| \leq |\mathbb{R}| = c$$

thus by CSB we have $c + c = c$.

Theorem 6.2. Given X, Y if X is infinite, then

1. $|X| + |X| = |X|$
2. $|X| + |Y| = \max\{|X|, |Y|\}$

Proof. 1. Exercise. (Hint: for countably infinite, we can create two countably infinite sets indexed by even and odd numbers. For infinite sets, we simply take out a countably infinite set (by theorem) A_1. If $X \setminus A_1$ is finite, then we are done. Otherwise we keep taking out countably infinite sets to form a collection of disjoint countably infinite sets).

□

Multiplication Let $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_m\}$, and $X \times Y = \{(x_i, y_j) \mid i = 1, \ldots, n, j = 1, \ldots, m\}$. Then $|X \times Y| = n \times n$.

Definition 6.2. Given X, Y define

$$|X| \cdot |Y| = |X \times Y|$$

Example 6.1. $|\mathbb{N} \times \mathbb{N}| = \aleph_0$, where define $f(n, m) = 2^n 3^m$ and $g(n) = (n, n)$ (1-1 and onto functions).

Question 6.3. What is $c \cdot c$?

Theorem 6.3. If X is infinite and $Y \neq 0$, then

1. $|X \times X| = |X| \Rightarrow |X||X| = |X|$
2. $|X \times Y| = \max(|X|, |Y|)$

Exponentiation Recall if $\{Y_x\}_{x\in X}$ is a collection of non-empty sets, then

$$\Pi_{x\in X}Y_x = \{f : X \to \bigcup_{x\in X} Y_x \mid f(x) \in Y_x\}$$

If $Y = Y_x$ for all $x \in X$ we have

$$Y^x = \Pi_{x\in X}Y = \{f : X \to Y\}$$

Example 6.2. Let $Y = \{1, \ldots, m\}$, $X = \{1, \ldots, n\}$, then

$$Y^X = \{f : \{1, \ldots, n\} \to \{1, \ldots, m\}\}$$

What is $|Y^X|$? It is m^n (for each $1, \ldots, m$, you have n choices thus we have $m \cdot \ldots \cdot m$ or m^n).

Definition 6.3. We define

$$|Y|^|X| = |Y^X|$$

Theorem 6.4. If X, Y are non-empty then

1. $|Y|^{|X|} \cdot |Y|^|Z| = |Y|^{|X|+|Z|}$
2. $(|Y|^{|X|})^{|Z|} = |Y|^{(|X|\cdot|Z|)}$

7 September 24, 2018

7.1 $2^{\aleph_0} = c$

Theorem 7.1. $2^{\aleph_0} = c$.

Proof. Observation:

$$2^{\aleph_0} = |\{0, 1\}^{\mathbb{N}}| = |\{f : \mathbb{N} \to \{0, 1\}\}| = |\{\{a_n\}_{n=1}^{\infty} \mid a_n = 0, 1\}|$$

Given a sequence $\{a_n\} \in \{0, 1\}^{\mathbb{N}}$, define

$$\phi(\{a_n\}) = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$$

where $\phi : \{0, 1\}^{\mathbb{N}} \to (0, 1)$ is 1-1 (note that we cannot map to two of the same real numbers in base 3 **unless** we had trailing 2s: but in this case we can't have 2s).
So $2^{\aleph_0} \leq c$.
Given $\alpha \in (0, 1)$ let

$$\alpha = \sum_{n=1}^{\infty} \frac{b_n}{2^n} \qquad b_n = 0, 1$$

i.e. the binary representation of our α (there may be *multiple binary representations*, but we could just pick one).
Let $\psi : (0, 1) \to \{0, 1\}^{\mathbb{N}}$, where

$$\psi(\alpha) = \psi(\sum_{n=1}^{\infty} \frac{b_n}{2^n}) = \{b_n\}$$

so ψ is 1-1 which means $c \leq 2^{\aleph_0}$. □

7.2 Countable union of countable sets

Observation 7.1. Suppose that $\{X_\alpha\}_{\alpha\in I}$ is a countable collection of countable sets.

Claim. $\bigcup_{\alpha\in I} X_i$ is countable.

Note: we can assume that $X_i \cap X_j = \varnothing$ if $i \neq j$. Why? Otherwise we can let

$$\begin{aligned} &E_1 = X_1 \\ &E_2 = X_2 \setminus E_1 \\ &E_3 = X_3 \setminus (E_1 \cup E_2) \\ &\vdots \\ &E_{n+1} = X_{n+1} \setminus (\bigcup_{i=1}^{n} E_i) \end{aligned}$$

Note $\bigcup_{i=1}^{\infty} X_i = \bigcup_{i=1}^{\infty} E_i$.
Let $E_n = \{x_{n,1}, x_{n,2}, \ldots\}$ (we need to use the Axiom of Choice here to pick out an enumeration of our set). Define $f : \bigcup_{i=1}^{\infty} E_i \to \mathbb{N}$

$$f(x_{n,j}) = 2^n 3^j$$

7.3 Cardinality of power sets

Question 7.1. Show that $|P(X)| = 2^{|X|} = |2^X|$.

Solution. Given $f : X \to \{0, 1\}$ let $A = \{x \in X \mid f(x) = 1\} \subset X$.
Define $\Gamma : 2^X \to P(X)$ by $\Gamma(f) = f^{-1}(\{1\})$. Γ is 1-1 (since two functions f differ only if they map something differently, one to 0 and one to 1, thus they will map to different sets in $P(X)$).
Convsersely, given $A \subset X$ define

$$X_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

$X_A \in 2^X$ (X_A is the characteristic function of A). We define $\Phi(A) = X_A$, thus $\Phi : P(X) \to 2^X$ is 1-1.
By CSB we have $|P(X)| = |2^X|$.

7.4 Russell's Paradox

Theorem 7.2 (Russsell's Paradox)**.** For any X, $|X| < 2^{|X|}$.

Proof. Let $f : X \to P(X)$ defined by $f(x) = \{x\}$ (1-1) thus $|X| \leq |P(X)|$.

Claim. There is no onto function $g : X \to P(X)$. Suppose we had such a g. Let $A = \{x \in X \mid x \neq g(x)\}$.
Since $A \subseteq X$ and f is onto, there exists some $x_0 \in X$ with $g(x_0) = A$.
If $x_0 \in A$, then $x_0 \in g(x_0) \Rightarrow x_0 \notin A$ by definition of A. If $x_0 \notin A$, then $x_0 \notin g(x_0)$ so $x_0 \in A$ by definition of A.

$\square$

7.5 Continuum Hypothesis

Axiom 7.1 (Continuum Hypothesis)**.** If $\aleph_0 \leq \gamma \leq c = 2^{\aleph_0}$, then either $\gamma = \aleph_0$ or $\gamma = c = 2^{\aleph_0}$.

Axiom 7.2 (Generalized Contnuum Hypothesis)**.** If $|X| \leq \gamma \leq 2^{|X|}$ then either $\gamma = |X|$ or $\gamma = 2^{|X|}$.

8 September 25, 2018

8.1 Metric spaces

Definition 8.1 (Metric and metric space)**.** Given X: a **metric** on X is a function $d : X \times X \to \mathbb{R}$ such that

1. $d(x, y) \geq 0$ and $d(x, y) = 0$ **iff** $x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

The pair (X, d) is a called a **metric space**.

Example 8.1. If $X = \mathbb{R}$, let $d(x, y) = |x - y|$ (standard metric on $\mathbb{R}$).

Question 8.1. Can we define a metric on any X?
Yes: we have the **discrete metric** where

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

We can verify that all conditions for a metric is satisfied by d.

Example 8.2. Let $X = \mathbb{R}^n$ and $d_2(\vec{x}, \vec{y}) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$. This is the **Euclidean** or **2-metric** on $\mathbb{R}^n$.

8.2 Norms

Definition 8.2 (Norm)**.** Given a vector space V (over $\mathbb{R}$), a norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ such that

1. $\|v\| \geq 0$ and $\|v\| = 0$ iff $v = 0$
2. $\|\alpha \cdot v\| = |\alpha|\|v\|$
3. $\|v + w\| \leq \|v\| + \|w\|$

Example 8.3. Define $\|\cdot\|_2$ on $\mathbb{R}^2$ by

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

then $\|\cdot\|_2$ is a norm. Note if $n = 1$ we have $\|x\| = |x|$ or the absolute value.
Furthermore note that $d_2(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|_2$.

Definition 8.3 (Normed linear/vector space)**.** A pair $(V, \|\cdot\|)$ is called a **normed linear space** (nls) or normed vector space.

Remark 8.1. Given a nls $(V, \|\cdot\|)$ we can define a *metric* $d_{\|\cdot\|}$ on V by $d_{\|\cdot\|}(x, y) = \|x - y\|$.
Note that this is a well-defined metric. Positive definiteness and symmetry properties are obvious. Let $x, y, z \in V$

$$\begin{aligned} d_{\|\cdot\|}(x, y) = \|x - y\| &= \|(x - z) + (z - y)\| \\ &\leq \|x - z\| + \|z - y\| \\ &= \|x - z\| + \|y - z\| \\ &= d_{\|\cdot\|}(x, z) + d_{\|\cdot\|}(y, z) \end{aligned}$$

Other norms on $\mathbb{R}^n$:

1. $\|\cdot\|_1$ on $\mathbb{R}^n$ where $\|(x_1, \ldots, x_n)\|_1 = \sum_{i=1}^n |x_i|$. Note that

$$\begin{aligned}\|\vec{x} + \vec{y}\|_1 &= \sum_{i=1}^n |x_i + y_i| \\ &\leq \sum_{i=1}^n |x_i| + |y_i| \\ &= \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| \\ &= \|\vec{x}\|_1 + \|\vec{y}\|_1\end{aligned}$$

   So we define $d_1(\vec{x}, \vec{y}) = \sum_{i=1}^n |x_i - y_i|$.

2. Let $\|\cdot\|_\infty$ (infinity norm) by $\|\vec{x}\|_\infty = \max\{|x_i|\}$.

   Positive definiteness is straightforward. Scalar multiple is obvious too.

   Note for any i, $|x_i + y_i| \leq |x_i| + |y_i|$, thus $\max\{|x_i + y_i|\} \leq \max\{|x_i|\} + \max\{|y_i|\}$.

   We can thus define the metric $d_\infty(\vec{x}, \vec{y}) = \|_\|\infty(\vec{x} - \vec{y}) = \max\{|x_i - y_i|\}$.

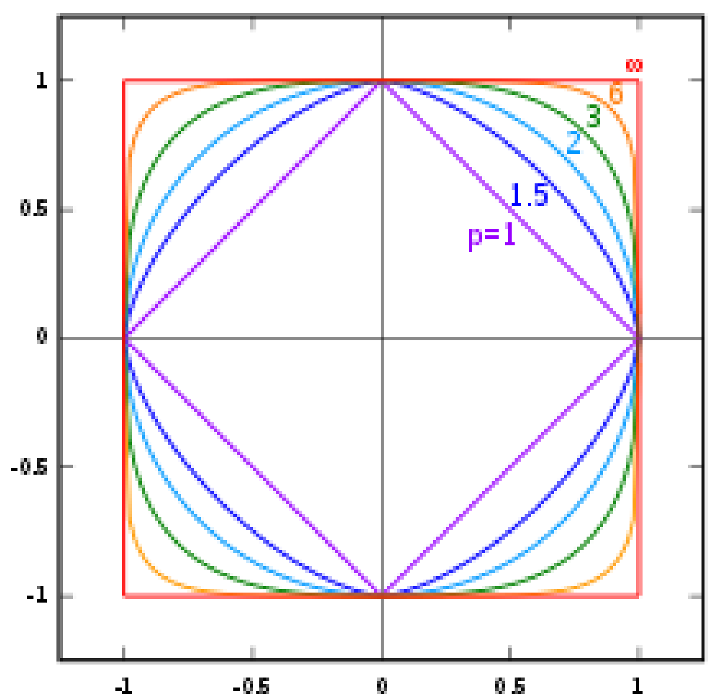

Figure 8.1: Diagrams for the l_p norms "balls" where $S_p = \{\vec{x} \in \mathbb{R}^2 \mid \|\vec{x}\|_p = 1\}$. In the diagram we have $p = 1, 1.5, 3, 6, \infty$.

We observe that $d_\infty \leq d_2 \leq d_1$: the number of points with distance ≤ 1 (inside their respective S_p balls) is the smallest for d_1, thus distances are "larger" for points in $\mathbb{R}^2$.

9 September 28, 2018

9.1 l_p norm

Definition 9.1 (l_p norm)**.** For $1 < p < \infty$, define on $\mathbb{R}^n$

$$\|\vec{x}\|_p = \big(\sum_{i=1}^{n} |x_i|^p\big)^{\frac{1}{p}}$$

we can also define the metric

$$d_p(\vec{x}, \vec{y}) = \big(\sum_{i=1}^{n} |x_i - y_i|^p\big)^{\frac{1}{p}}$$

Note that l_p for $0 < p < 1$ results in a non-convex ball: this means any convex combination of two points may result in a point outside the ball. This implies that the triangle inequality does not hold.

We can show that $\|\cdot\|_p$ is a norm.

9.2 Young's Inequality

Lemma 9.1 (Young's Inequality)**.** If $1 < p < \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and if $\alpha, \beta > 0$ then $\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$.

Proof. Let us draw $u = t^{p-1}$ where u is the y-axis and t is the y-axis.
We bound the area with $t = \alpha$ and $u = \beta$. Note that the inverse becomes $t = u^{\frac{1}{p-1}} = u^{q-1}$ (where $\frac{1}{p-1} = q - 1$ after a bit of algebraic manipulation).
We clearly see that the area above and below the curve is greater than the box, thus

$$\begin{aligned}
\alpha\beta &\leq \int_0^\alpha t^{p-1}\,\mathrm{d}t + \int_0^\beta u^{q-1}\,\mathrm{d}u \\
&= \frac{t^p}{p} \big|_0^\alpha + \frac{u^q}{q} \big|_0^\beta \\
&= \frac{\alpha^p}{p} + \frac{\beta^q}{q}
\end{aligned}$$

□

9.3 Holder's Inequality (*****)

Theorem 9.1 (Holder's Inequality)**.** Let $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$. Let $\vec{x} = (x_1, \ldots, x_n)$ $\vec{y} = (y_1, \ldots, y_n)$. Then

$$\sum_{i=1}^{n} |x_i y_i| \leq \big(\sum_{i=1}^{n} |x_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n} |y_i|^q\big)^{\frac{1}{q}}$$

Note $p = 2$ is the Cauchy-Schwarz Inequality (i.e. Holder's is a generalization of Cauchy Schwarz).

Proof. WLOG we may assume that $\vec{x}, \vec{y} \neq \vec{0}$.
Note if $\alpha, \beta \neq 0$ then

$$\sum_{i=1}^{n} |x_i y_i| \leq \big(\sum_{i=1}^{n} |x_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n} |y_i|^q\big)^{\frac{1}{q}}$$

holds **if and only if**

$$\sum_{i=1}^{n}|(\alpha x_i)\beta y_i| \leq \big(\sum_{i=1}^{n}|\alpha x_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n}|\beta y_i|^q\big)^{\frac{1}{q}}$$

(we can arbitrarily scale our vectors $\vec{x}, \vec{y}$). Hence we can assume that

$$\big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} = 1 = \big(\sum_{i=1}^{n}|y_i|^q\big)^{\frac{1}{q}}$$

(that is we scale our vectors so that the above equality holds). By Jensen's inequality we have

$$|x_i y_i| \leq \frac{|x_i|^p}{p} + \frac{|y_i|^q}{q}$$

Thus the sume over all $i = 1, \ldots, n$ is

$$\begin{aligned}
\sum_{i=1}^{n}|x_i y_i| &\leq \frac{\sum_{i=1}^{n}|x_i|^p}{p} + \frac{\sum_{i=1}^{n}|y_i|^q}{q} \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1 \\
&= \big(\sum_{i=1}^{n}|\alpha x_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n}|\beta y_i|^q\big)^{\frac{1}{q}}
\end{aligned}$$

□

9.4 Minkowski's Inequality

Theorem 9.2 (Minkowski's Inequality)**.** Let $1 < p < \infty$. If $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n)$ then

$$\big(\sum_{i=1}^{n}|x_i + y_i|^p\big)^{\frac{1}{p}} \leq \big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} + \big(\sum_{i=1}^{n}|y_i|^p\big)^{\frac{1}{p}}$$

i.e. the *triangle inequality* for l_p norm holds

$$\|_{\|}\vec{x} + \vec{y}_p \leq \|\vec{x}\|_p + \|\vec{y}\|_p$$

Proof. We can assume that $\vec{x} + \vec{y} \neq 0$. We have

$$\begin{aligned}
\sum_{i=1}^{n}|x_i + y_i|^p &= \sum_{i=1}^{n}|x_i + y_i||x_i + y_i|^{p-1} \\
&\overset{\triangle}{\leq} \sum_{i=1}^{n}|x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{n}|y_i||x_i + y_i|^{p-1} \\
&\leq \big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n}|x_i + y_i|^{(p-1)q}\big)^{\frac{1}{q}} + \big(\sum_{i=1}^{n}|y_i|^p\big)^{\frac{1}{p}} \cdot \big(\sum_{i=1}^{n}|x_i + y_i|^{(p-1)q}\big)^{\frac{1}{q}}
\end{aligned}$$

where the last line follows from Holder's inequality. Thus we have

$$\begin{aligned}
&\sum_{i=1}^{n}|x_i+y_i|^p \leq \Big(\big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} + \big(\sum_{i=1}^{n}|y_i|^p\big)^{\frac{1}{p}}\Big) \cdot \big(\sum_{i=1}^{n}|x_i+y_i|^p\big)^{\frac{1}{q}} \\
&\Rightarrow \sum_{i=1}^{n}|x_i+y_i|^{1-\frac{1}{q}} \leq \big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} + \big(\sum_{i=1}^{n}|y_i|^p\big)^{\frac{1}{p}} \\
&\Rightarrow \sum_{i=1}^{n}|x_i+y_i|^{\frac{1}{p}} \leq \big(\sum_{i=1}^{n}|x_i|^p\big)^{\frac{1}{p}} + \big(\sum_{i=1}^{n}|y_i|^p\big)^{\frac{1}{p}}
\end{aligned}$$

as desired. □

Remark 9.1. This shows that $\|\cdot\|_p$ is a norm on $\mathbb{R}^n$.

Observation 9.1. Given $1 \leq p \leq q \leq \infty$ we have $\|\cdot\|_\infty \leq \|\cdot\|_q \leq \|\cdot\|_p \leq \|\cdot\|_1$.

9.5 Sequence spaces

Definition 9.2 (Sequence space)**.**

1. Let the l_1 space be defined as

   $$l_1(\mathbb{N}) = l_1 = \{\{x_n\} \mid x_n = \mathbb{R}, \sum_{n=1}^{\infty}|x_n| < \infty\}$$

   (i.e. sequences that converge).

   We define a norm on l_1

   $$\|\{x_n\}\|_1 = \sum_{n=1}^{\infty}|x_n|$$

   Let $\{x_i\}, \{y_i\} \in l_1$. For all $n \in \mathbb{N}$

   $$\begin{aligned}
   \sum_{i=1}^{n}|x_i+y_i| &\overset{\triangle}{\leq} \sum_{i=1}^{n}|x_i| + \sum_{i=1}^{n}|y_i| \\
   &= \|\{x_i\}\|_1\|\{y_i\}\|_1
   \end{aligned}$$

   hence $\sum_{i=1}^{\infty}|x_i+y_i| \leq \|\{x_i\}\|_1 + \|\{y_i\}\|_1$, thus $\{x_i+y_i\} \in l_1$ (finite sum) and the triangle inequality holds i.e. $\|\{x_i+y_i\}\|_1 \leq \|\{x_i\}\|_1 + \|\{y_i\}\|_1$.

   Let $\{x_i\} \in l_1$, $\alpha \in \mathbb{R}$. We know for a convergent sequence

   $$\sum_{i=1}^{\infty}|\alpha x_i| = |\alpha|\sum_{i=1}^{\infty}|x_i|$$

   thus $\{\alpha x_i\} \in l_1$ and $\|\{\alpha x_i\}\|_1 = |\alpha|\|\{x_i\}\|_l$.

   Positive definiteness is trivial, thus l_1 is a vector space and $(l_1, \|\cdot\|_1)$ is a normed linear space.

2. Let

   $$l_\infty(\mathbb{N}) = l_\infty = \{\{x_i\} \mid \{x_i\} \text{ is bounded}\}$$

   Define the norm on l_∞

   $$\|\{x_i\}\|_\infty = \text{lub}\{|x_i|\} \qquad i \in \mathbb{N}$$

If $\{x_i\}, \{y_i\} \in l_\infty$ then

$$|x_i + y_i| \leq |x_i| + |y_i| \leq \|\{x_i\}\|_\infty + \|\{y_i\}\|_\infty$$

for all $i \in \mathbb{N}$. So $\{x_i + y_i\} \in l_\infty$ and $\|\{x_i + y_i\}\|_\infty \leq \|\{x_i\}\|_\infty + \|\{y_i\}\|_\infty$.

Similarly $\{\alpha x_i\} \in l_\infty$ and $\|\{\alpha x_i\}\|_\infty = |\alpha| \|\{x_i\}\|_\infty$. Therefore l_∞ is a vector space and $(l_\infty, \|\cdot\|_\infty)$ is a normed linear space.