

The Knowledge

CALLUM CASSIDY-NOLAN*

November 5, 2021

Dedicated to the proofs that were left as exercises to the readers

Contents

1	Analysis	5
1.1	Limits	6
1.2	Differentiation	8
1.3	Sequences	11
1.4	Multi Variable	11
2	Linear Algebra	12
2.1	Vectors	12
2.2	Matrices	13
3	First Order Logic	15
3.1	Languages	15
3.2	Deductions	17
3.3	Completeness	18
3.4	Incompleteness	20
4	Topology	21
4.1	Topological Spaces and Continuous Functions	21
4.1.1	Basis for a Topology	21
4.1.2	The Subspace Topology	21
4.1.3	The Product Topology	21
4.1.4	The Metric Topology	22
4.2	Connectedness and Compactness	25
4.2.1	Compact Spaces	26
5	Computer Science	27
6	Probability	28
6.1	Bayesian Interference	28
7	Programming	29

List of Theorems

1.0.1 Triangle Inequality	5
1.2.1 Chain Rule	8
3.3.1 Completeness Theorem	19
4.1.1 Basis for the Box Topology	22
4.1.2 Basis for the Product Topology	22

List of Propositions

1.1.1 Limit of Constant	6
1.1.2 Constant in Limit	6
1.1.3 Sum of Limits	7
1.2.1 Little O and Differentiability Equivalence	9
2.1.1 Algebraic Def Implies Geometric Def of Dot Product	12
3.3.1 Contradiction has no Model	18
3.3.2 Implying a Contradiction Means no Model	18
3.3.3 Logical Contradiction	18
4.1.1 Epsilon Balls Form A Basis	23
4.2.1 Connected Implies Closure Connected	26

List of Lemmas

1.0.1 Absolute Value is Equal to max	5
3.2.1 Universal connection to Variable Assignment Function	17
3.3.1 Constant Extension still Consistent	19
4.1.1 Epsilon Ball Contains Another	23
4.2.1 Covering Yields Finite Covering if and only if Compact	26

List of Definitions

1.1.1 Real Valued Limit	6
1.3.1 Bounded Sequence	11
1.4.1 Directional Derivative	11
2.1.1 Algebraic Dot Product	12
2.1.2 Geometric Dot Product	12
2.2.1 Matrix Multiplication	13
3.1.1 First-Order Language	15
3.1.2 A Term's Variable Replaced by a Term	15
3.1.3 A Term is Substitutable for a Variable	16
3.1.4 Free Variable in a Formula	16
3.2.1 Deduction of a Formula	17
3.2.2 Logical Axioms	17
3.3.1 Consistent Set of L Formulas	18
3.4.1 Representable Set	20
3.4.2 Weakly Representable Set	20
3.4.3 Total Function	20
3.4.4 Partial Function	20
3.4.5 Representable Function	20
3.4.6 Definable Set	20
4.1.1 Basis	21
4.1.2 Subspace Topology	21
4.1.3 Product Topology	21
4.1.4 Box Topology	22
4.1.5 R Omega	22
4.1.6 A metric	22
4.1.7 Epsilon Ball	23
4.1.8 Metric Topology	23
4.1.9 Bounded Subset of a Metric Space	24
4.2.1 Connected Space	25
4.2.2 Totally Disconnected	26
4.2.3 Covering	26
4.2.4 Compact Space	26
5.0.1 Big-O	27
5.0.2 Little-o	27
6.0.1 Conditional Independence	28

List of Examples

4.1.1 Discrete Metric	23
4.1.2 Function on \mathbb{R}^ω	24
4.2.1 Closed and Bounded, not Compact	25
4.2.2 \mathbb{R}^ω Connected?	25

Preface

This book contains knowledge that that me or my peers have obtained, the purpose is to explain things fundamentally and in full detail so that someone who has never touched the subject may be able to understand it. It will focus on conveying the ideas that are involved in synthesizing the new knowledge with less of a focus on the results themselves.

It differs from a normal textbook in that it is open source and will fall under more continuous development rather than having editions that periodically come out. It also differs in the sense that it welcomes other users to improve the book.

Here is a link to the book: <https://github.com/cuppajoeman/knowledge-book/blob/main/build/book.pdf>

Structure of book

The book is partitioned into different sections based on the domain it is involved with. There may be shared definitions and theorems throughout the chapters, but in general it will start more elementary and get more advanced.

Knowledge

In this book you will find many results, will will characterize them as being one of the following

- Theorems – Results that are of importance and who's proof is not easily found (maybe using a novel idea)
- Propositions – Results of less importance who's proof could be constructed without a novel idea
- Lemmas – Results that are technical intermediate steps which has no standing as an independent result on first observation *
- Corollaries – Results which follow readily from an existing result of greater importance

Recommendations

By now you might know that in order to actually get better at mathematics you have to engage with it. This book may be used as a reference at times, but I highly recommend trying to re-prove statements or coming up with your own ideas before instantly looking at the solutions.

About the companion website

The website[†] for this file contains:

- A link to (freely downloadable) latest version of this document.
- Link to download \LaTeX source for this document.

*But sometimes they escape, as their usage becomes more than just an intermediate step, as Zorn's or Fatou's Lemmas did

[†]<https://github.com/cuppajoeman/knowledge-book>

Acknowledgements

- A special word of thanks to professors who wanted to make sure I understood and learned as much as possible Alfonso Gracia-Saz[‡], Jean-Baptiste Campesato[§], Valentine Chiche-Lapierre and Gal Gross[¶]
- Thanks to Z-Module, riv, PlanckWalk, franciman, qergle from #math on <https://libera.chat/>.

[‡]<https://www.math.toronto.edu/cms/alfonso-memorial/>

[§]<https://math.univ-angers.fr/~campesato/>

[¶]<https://www.galgr.com/>

Contributing

Contributions to the project are very welcome, let's delve into how to get started with this.

If you want to contribute to the project it's most likely that a contribution will fall into one of the following categories

- Content Based
 - Adding Definitions, Theorems, ...
 - Finishing TODO's
 - Formatting of the book
- Structural Layout of Project
 - Organization
 - Simplifying the existing structure of the directories
 - Making scripts which set up new structures
- External
 - Adding explanatory content to help onboard new users
 - Getting others involved
 - Creating infrastructure to support users (Github discussions)

Communication

All communication will occur through github discussions. You can access it here:

Content Based

If you want to add a new top level structure, the best thing to do is to verify with other members of the project if it warrants it's own top level structure, otherwise it can be added as a substructure of an existing one.

Supposing that you are on linux, the easiest way to make a new structure and start working on it:

```
user@machine:~# cp -r structure new_structure
user@machine:~# cd new_structure
user@machine:~# mv content.tex new_structure.tex
user@machine:~# nvim new_structure.tex
```

Otherwise if you're adding a new theorem, it could be:

```
user@machine:~# cd existing_structure/theorems
user@machine:~# cp theorem.tex my_new_theorem.tex
user@machine:~# nvim my_new_theorem.tex
user@machine:~# ...
user@machine:~# git add -A && git commit -m "add my new theorem" && git push
```

Creating Files

For example if want to create a new definition for topology we would go into the definition folder for topology and create a new file using pothole case and don't include any extraneous words, for example it is better not to append "the" to the front of your file names when not specifically required.

Analysis

Theorem 1.0.1: Triangle Inequality

Let $a, b \in \mathbb{R}$ then

$$|a + b| \leq |a| + |b|$$

Proof

- $a + b \leq |a| + b \leq |a| + |b| \Leftrightarrow a + b \leq |a| + |b|$
- $-(a + b) = -a - b \leq |a| - b \leq |a| + |b| \Leftrightarrow -(a + b) \leq |a| + |b|$
 - Which are both derived by the fact that $|x| = \max(x, -x) \geq x, -x$
- Finally we can see that $|a + b| = \max(a + b, -(a + b))$, so no matter which one of $\{a + b, -(a + b)\}$ $|a + b|$ is equal to we have that

$$|a + b| \leq |a| + |b|$$

■

Note that the triangle equality is intuitively telling us that the shortest distance between two points is a straight line

Lemma 1.0.1: Absolute Value is Equal to max

$\forall x \in \mathbb{R}$

$$|x| = \max(x, -x)$$

Proof

- If $x \geq 0$ then $|x| = x$ and $\max(x, -x) = x$
- If $x < 0$ then $|x| = -x$ and $\max(x, -x) = -x$ since $x < 0 \Leftrightarrow -x > 0$ and therefore $-x > x$

■

1.1 Limits

Proposition 1.1.1: Limit of Constant

Let $f(x) = \alpha \in \mathbb{R}$ be a constant function, then

$$\lim_{x \rightarrow a} f(x) = \alpha$$

Proof

- Let $\varepsilon \in \mathbb{R}^+$, let δ be fixed as any real number, then assume that $\forall x \in \text{dom}(f)$ that $0 < |x - a| < \delta$, we will show that $|f(x) - \alpha| < \varepsilon$.
- Recall that $f(x) = \alpha$ for all $x \in \text{dom}(f)$, therefore $|f(x) - \alpha| = |\alpha - \alpha| = 0 < \varepsilon$ as needed. ■

Notice that in the above proof that we didn't use our hypothesis in the proof of the consequent, which makes sense because no matter which interval you look in the function is constant there, thus it doesn't depend on the hypothesis at all.

Definition 1.1.1: Real Valued Limit

Suppose f is a real valued function, then we say that the limit of f at a is L and write $\lim_{x \rightarrow a} f(x) = L$ when the following holds:

$$\forall \varepsilon \in \mathbb{R}^+, \exists \delta \in \mathbb{R}^+ \text{ such that } \forall x \in \text{dom}(f), 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon$$

Proposition 1.1.2: Constant in Limit

Assume that the following limit exists $\lim_{x \rightarrow a} f(x)$ and define L to be it's value, then for any $c \in \mathbb{R}$

$$\lim_{x \rightarrow a} [cf(x)] = c \lim_{x \rightarrow a} f(x)$$

Proof

- If $c = 0$ then using the fact that the limit of a constant is that constant itself, we have:

$$\lim_{x \rightarrow a} [0f(x)] = \lim_{x \rightarrow a} 0 = 0 = 0 \lim_{x \rightarrow a} f(x)$$

- If $c \neq 0$, we must prove that for any $\varepsilon_c \in \mathbb{R}^{>0}$ there exists δ_c such that for all $x_c \in \text{dom}(f)$

$$|x_c - a| \leq \delta_c \Rightarrow |cf(x_c) - cL| \leq \varepsilon_c$$

- Notice that if we were to let ε in the original definition be equal to $\frac{\varepsilon_c}{|c|}$ then we could multiply the equation after the implication on both sides by $|c|$ (so that we can absorb it into the absolute value).
- Let $\varepsilon_c \in \mathbb{R}^{>0}$ since the original limit holds for any epsilon, bind ε to $\frac{\varepsilon_c}{|c|}$ and we get δ such that for all $x \in \text{dom}(f)$, the following holds:

$$|x - a| \leq \delta \Rightarrow |f(x) - L| \leq \frac{\varepsilon_c}{|c|} \quad (\alpha)$$

- Take $\delta_c = \delta$, let $x_c \in \text{dom}(f)$ and bind x in the original definition to x_c , and assume that $|x_c - a| \leq \delta_c$, because of our choice for δ_c we satisfy α 's hypothesis with x replaced by x_c and we get

$$|f(x_c) - L| \leq \frac{\varepsilon_c}{|c|} \Leftrightarrow |c| |f(x_c) - L| \leq \varepsilon_c$$

- Since for any $a, b \in \mathbb{R}$ we have $|ab| = |a| |b|$ we can conclude with distributivity in \mathbb{R} that

$$|cf(x_c) - cL| \leq \varepsilon_c$$

As required. ■

Proposition 1.1.3: Sum of Limits

If $\lim_{x \rightarrow a} f(x) = L$ and $\lim_{x \rightarrow a} g(x) = M$ both exist then

$$\lim_{x \rightarrow a} [f(x) + g(x)] = L + M$$

Proof

- Suppose antecedent holds, so we have existence of the limits $\lim_{x \rightarrow a} f(x) = L$ and $\lim_{x \rightarrow a} g(x) = M$, now we'd like to show that the limit $\lim_{x \rightarrow a} [f(x) + g(x)]$ equals $L + M$
- Let $\varepsilon \in \mathbb{R}^+$, now from the fact that the two prior limits exist we can substitute in $\frac{\varepsilon}{2}$ into them, to get δ_L and δ_M respectively, such that

$$\forall x \in \text{dom}(f), 0 < |x - a| < \delta_L \Rightarrow |f(x) - L| < \frac{\varepsilon}{2}$$

and

$$\forall x \in \text{dom}(g), 0 < |x - a| < \delta_M \Rightarrow |g(x) - M| < \frac{\varepsilon}{2}$$

- Take $\delta = \min(\delta_L, \delta_M)$, so that we may utilize the above inequalities as we assume that $\forall x \in \text{dom}(f + g)$ that $0 < |x - a| < \delta$, and proceed to show that $|(f(x) + g(x)) - (L + M)| < \varepsilon$, since our x is bounded by the min of the two deltas from the other limits, we utilize the inequalities in tandem with the triangle equality to obtain:

$$|(f(x) + g(x)) - (L + M)| = |(f(x) - L) + (g(x) - M)| \leq |f(x) - L| + |g(x) - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$
■

1.2 Differentiation

Theorem 1.2.1: Chain Rule

Given two functions f and g where g is differentiable at the point \bar{x} and f differentiable at the point $\bar{y} = g(\bar{x})$ then

$$(f \circ g)'(\bar{x}) = f'(g(\bar{x})) \cdot g'(\bar{x})$$

Proof

- We define the following two new functions

$$v(h) = \frac{g(\bar{x} + h) - g(\bar{x})}{h} - g'(\bar{x}) \text{ and } w(k) = \frac{f(\bar{y} + k) - f(\bar{y})}{k} - f'(\bar{y})$$

α : Note that $\lim_{h \rightarrow 0} v(h) = 0$ and $\lim_{k \rightarrow 0} w(k) = 0$, and that we can re-arrange the above to:

$$(v(h) + g'(\bar{x})) \cdot h + g(\bar{x}) = g(\bar{x} + h) \text{ and } (w(k) + f'(\bar{y})) \cdot k + f(\bar{y}) = f(\bar{y} + k)$$

- Now it follows that

$$\begin{aligned} f(g(\bar{x} + h)) &= f\left((v(h) + g'(\bar{x})) \cdot h + g(\bar{x})\right) \\ &= f\left(g(\bar{x}) + (v(h) + g'(\bar{x})) \cdot h\right) \\ &= f\left(\bar{y} + (v(h) + g'(\bar{x})) \cdot h\right) \\ &= \left(w\left((v(h) + g'(\bar{x})) \cdot h\right) + f'(\bar{y})\right) \cdot (v(h) + g'(\bar{x})) \cdot h + f(\bar{y}) \end{aligned}$$

- Now recall that we are interested in $(f \circ g)'(\bar{x})$, so we're going to have to look at:

$$\lim_{h \rightarrow 0} \frac{f(g(\bar{x} + h)) - f(g(\bar{x}))}{h}$$

- Notice that the term on the left of the sum in the numerator is something we already know, and therefore it becomes:

$$\lim_{h \rightarrow 0} \frac{\left(w\left((v(h) + g'(\bar{x})) \cdot h\right) + f'(\bar{y})\right) \cdot (v(h) + g'(\bar{x})) \cdot h + f(\bar{y}) - f(g(\bar{x}))}{h}$$

- By noting that $f(\bar{y}) = f(g(\bar{x}))$, then cancelling out the h we can see that the above simplifies to

$$\lim_{h \rightarrow 0} \left(w\left((v(h) + g'(\bar{x})) \cdot h\right) + f'(\bar{y})\right) \cdot (v(h) + g'(\bar{x}))$$

- By α on v and w with $k = g'(\bar{x}) \cdot h$, we have that the above limit is:

$$\lim_{h \rightarrow 0} \left(w\left(g'(\bar{x}) \cdot h\right) + f'(\bar{y})\right) \cdot g'(\bar{x}) = \lim_{h \rightarrow 0} f'(\bar{y}) \cdot g'(\bar{x}) = \boxed{f'(g(\bar{x})) \cdot g'(\bar{x})}$$

■

The chain rule is sometimes written as $\frac{df}{dx}(x) = \frac{df}{dg}(x) \frac{dg}{dx}(x)$, but notice that this extends liebniz notation as we only know that $\frac{df}{dx} = f'(x)$, and it is not defined for when we take the derivative with respect to another function. This motivates the definition

$$\frac{df}{dg}(x) = \frac{df}{dx}(g(x))$$

Proposition 1.2.1: Little O and Differentiability Equivalence

A function f is differentiable at the point \bar{x} if and only if there is a number $\alpha \in \mathbb{R}$ for any $h \in \mathbb{R}$ we have

$$f(\bar{x} + h) = f(\bar{x}) + h\alpha + E(h)$$

Where $E \in o(h)$

Proof

• \Rightarrow

– Assume that f is differentiable, therefore we have

$$f'(\bar{x}) = \lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x})}{h}$$

– Equivalently:

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x})}{h} - f'(\bar{x}) = 0$$

– Since $\lim_{h \rightarrow 0} f'(\bar{x}) = f'(\bar{x})$ as it's a constant, we may use the limit law to deduce:

$$\lim_{h \rightarrow 0} \left(\frac{f(\bar{x} + h) - f(\bar{x})}{h} - f'(\bar{x}) \right) = 0 \Leftrightarrow \lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x}) - hf'(\bar{x})}{h} = 0$$

– That is to say that $f(\bar{x} + h) - f(\bar{x}) - hf'(\bar{x}) \in o(h)$, set $E(h) = f(\bar{x} + h) - f(\bar{x}) - hf'(\bar{x})$ then we have that

$$f(\bar{x} + h) = f(\bar{x}) + hf'(\bar{x}) + E(h)$$

where we've taken $\alpha = f'(\bar{x})$

• \Leftarrow

– Assume that there is a number $\alpha \in \mathbb{R}$ and $E \in o(h)$, so that for any $h \in \mathbb{R}$ we have:

$$f(\bar{x} + h) = f(\bar{x}) + h\alpha + E(h) \Leftrightarrow l(h) = f(\bar{x} + h) - f(\bar{x}) - h\alpha$$

thus $f(\bar{x} + h) - f(\bar{x}) - h\alpha \in o(h)$.

– So then by definition of little-o, we get:

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x}) - h\alpha}{h} = 0$$

– Since $\lim_{h \rightarrow 0} \alpha = \alpha$ we may add it to both sides using the limit law on the left to obtain:

$$\lim_{h \rightarrow 0} \left(\frac{f(\bar{x} + h) - f(\bar{x}) - h\alpha}{h} + \alpha \right) = \alpha \Leftrightarrow \lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x})}{h} = \alpha$$

Meaning that $\alpha = f'(\bar{x})$ and that f is differentiable



Notice that the number α is equal to a limit and that limits are unique, therefore the solution $\alpha = f'(\bar{x})$ is the unique solution to the above proposition. Also since this is the unique solution, we know that there is no other real number besides $f'(\bar{x})$ where this holds, therefore $f'(\bar{x})$ is the best linear approximation to f at the point \bar{x} .

Finally notice that if you solve for $E(h)$ and then recall that $E(h) \in o(h)$ we have:

$$\lim_{h \rightarrow 0} \frac{\overbrace{f(\bar{x} + h) - (f(\bar{x}) + hf'(\bar{x}))}^{\text{error}}}{h} = 0$$

the error is going to zero at a rate which is faster than linear, as if the error were decreasing at a linear rate the limit would evaluate to some non-zero constant.

1.3 Sequences

Definition 1.3.1: Bounded Sequence

A sequence $\{x_n\}$ is said to be bounded if $\exists M \in \mathbb{R}$ such that $|x_n| \leq M$ for all $n \in \mathbb{N}$

1.4 Multi Variable

Definition 1.4.1: Directional Derivative

Let $U \subseteq \mathbb{R}^n$ be open, $f : U \rightarrow \mathbb{R}$ and $\vec{a} \in U$ and $\vec{v} \in \mathbb{R}^n$

$$\partial_{\vec{v}} f(\vec{a}) \stackrel{\text{D}}{=} \lim_{t \rightarrow 0 \in \mathbb{R}} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t}$$

2

Linear Algebra

2.1 Vectors

Definition 2.1.1: Algebraic Dot Product

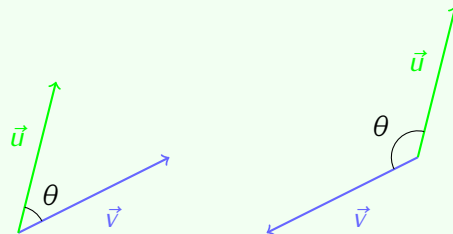
Let $u_1, u_2, \dots, u_{n-1}, u_n$ and $v_1, v_2, \dots, v_{n-1}, v_n$ denote the components of \vec{u} and \vec{v} respectively, then we have:

$$\vec{v} \cdot \vec{u} \stackrel{D}{=} \sum_{i=1}^n v_i u_i$$

Definition 2.1.2: Geometric Dot Product

Let $\vec{u}, \vec{v} \in \mathbb{R}^n$ and let θ be the angle between the two (the one in the range $[0, \pi]$), then we have the dot product:

$$\vec{v} \cdot \vec{u} \stackrel{D}{=} \|\vec{v}\| \|\vec{u}\| \cos(\theta)$$



Proposition 2.1.1: Algebraic Def Implies Geometric Def of Dot Product

$$\sum_{i=1}^n v_i u_i = \|\vec{v}\| \|\vec{u}\| \cos(\theta)$$

Proof

- Consider two vectors $\vec{x}, \vec{y} \in \mathbb{R}^2$, we know that from the polar coordinate system we can represent them as

$$\vec{x} = (\|\vec{x}\| \cos(\theta_x), \|\vec{x}\| \sin(\theta_x)) \quad \text{and} \quad \vec{y} = (\|\vec{y}\| \cos(\theta_y), \|\vec{y}\| \sin(\theta_y))$$

- The algebraic definition implies the following

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \|\vec{y}\| \cos(\theta_x) \cos(\theta_y) + \|\vec{x}\| \|\vec{y}\| \sin(\theta_x) \sin(\theta_y) = \|\vec{x}\| \|\vec{y}\| (\cos(\theta_x) \cos(\theta_y) + \sin(\theta_x) \sin(\theta_y))$$

- Then we can recall the sin angle difference formula to obtain:

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \|\vec{y}\| (\cos(\theta_x - \theta_y))$$

- And just noting that $\theta_x - \theta_y$ is the angle between the vectors \vec{x} and \vec{y} as needed.



2.2 Matrices

Definition 2.2.1: Matrix Multiplication

If \mathbf{A} is an $m \times n$ matrix and \mathbf{B} is an $n \times p$ matrix,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

the matrix product $\mathbf{C} = \mathbf{AB}$ (denoted without multiplication signs or dots) is defined to be the $m \times p$ matrix

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix}$$

such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}$$

for $i = 1, \dots, m$ and $j = 1, \dots, p$.

```

1 def multiply_matrices(matrix_A, matrix_B):
2     """
3     Given two matrices, we compute the matrix multiplication of the two of them
4
5     Supposing that matrix_A is MxN and that matrix_B is NxK
6     then the resulting matrix is M x K (1)
7     """
8
9     transpose_matrix_B = transpose_matrix(matrix_B)
10
11     num_rows_of_matrix_A = len(matrix_A) # This is M
12     num_columns_of_matrix_A = len(matrix_A[0]) #This is N
13
14     num_rows_of_matrix_B = len(matrix_B) #this is N
15     num_columns_of_matrix_B = len(matrix_B[0])# This is K
16
17     resulting_matrix = []
18
19     #initialize the matrix to be all zeros
20     # then we will fill in the entries with the correct values
21     for i in range(num_rows_of_matrix_A):
22         blank_row = []
23         for j in range(num_columns_of_matrix_B):
24             blank_row.append(0)
25         resulting_matrix.append(blank_row)

```

```
26 # Notice that the matrix now has dimenisions 'num_rowsof_matrix_A' x 'num_columns_of_matrix_B'
27 # Or equivalently M x K , as we noted by (1)
28
29 for row_index in range(num_rows_of_matrix_A):
30     for column_index in range(num_rows_of_matrix_B):
31         row_from_matrix_A = matrix_A[row_index]
32         transposed_column_from_matrix_B = transpose_matrix_B[column_index]
33         resulting_matrix[row_index][column_index] = dot_product(row_from_matrix_A,
34                             transposed_column_from_matrix_B)
35
36
37
38 return resulting_matrix
39
40
41
42
43
44 def transpose_matrix(matrix):
45     """Given a matrix turn it into a new matrix where it's rows are equal to it's columns"""
46     rows = len(matrix)
47     columns = len(matrix[0]) # assuming matrix is non-empty
48
49     transpose_matrix = []
50     for i in range(columns):
51         temp_new_row = []
52         for j in range(rows):
53             temp_new_row.append(matrix[j][i])
54         transpose_matrix.append(temp_new_row)
55
56 return transpose_matrix
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
```

3

First Order Logic

3.1 Languages

Definition 3.1.1: First-Order Language

A first order language \mathcal{L} is an infinite collection of distinct symbols, no one of which is properly contained in another, separated into the following categories:

1. Parentheses: $(,)$.
2. Connectives: \vee, \neg .
3. Quantifier: \forall .
4. Variables, one for each positive integer n : $v_1, v_2, \dots, v_n, \dots$. The set of variable symbols will be denoted Vars.
5. Equality symbol: $=$.
6. Constant symbols: Some set of zero or more symbols.
7. Function symbols: For each positive integer n , some set of zero or more n -ary function symbols.
8. Relation symbols: For each positive integer n , some set of zero or more n -ary relation symbols.

Definition 3.1.2: A Term's Variable Replaced by a Term

Suppose that u is a term, x is a variable, and t is a term. We define the term u_t^x (read " u with x replaced by t ") as follows:

1. If u is a variable not equal to x , then u_t^x is u .
2. If u is x , then u_t^x is t .
3. If u is a constant symbol, then u_t^x is u .
4. If $u \equiv fu_1u_2\dots u_n$, where f is an n -ary function symbol and the u_i are terms, then u_t^x is $f(u_1)_t^x(u_2)_t^x\dots(u_n)_t^x$

- In the fourth clause of the definition above and in the first two clauses of the next definition, the parentheses are not really there. Because $u_1)_t^x$ is hard to read so the parentheses have been added in the interest of readability.
- If we let t be $g(c)$ and we let u be $f(x, y) + h(z, x, g(x))$, then u_t^x is

$$f(g(c), y) + h(z, g(c), g(g(c)))$$

Definition 3.1.3: A Term is Substitutable for a Variable

Suppose that ϕ is an \mathcal{L} -formula, t is a term, and x is a variable. We say that t is substitutable for x in ϕ if

1. ϕ is atomic, or
2. $\phi \equiv \neg(\alpha)$ and t is substitutable for x in α , or
3. $\phi \equiv (\alpha \vee \beta)$ and t is substitutable for x in both α and β , or
4. $\phi \equiv (\forall y)(\alpha)$ and either
 - x is not free in ϕ , or
 - y does not occur in t and t is substitutable for x in α .

To understand the motivation behind the fourth clause, consider the formula:

$$\phi \equiv (\forall x) (\forall y) x = y$$

Then one might want to say that $\left((\forall y) x = y \right)_t^x$ where t is any term

Definition 3.1.4: Free Variable in a Formula

Suppose that v is a variable and ϕ is a formula. We will say that v is free in ϕ if

1. ϕ is atomic and v occurs in (is a symbol in) ϕ , or
2. $\phi \equiv (\neg\alpha)$ and v is free in α , or
3. $\phi \equiv (\alpha \vee \beta)$ and v is free in at least one of α or β , or
4. $\phi \equiv (\forall u)(\alpha)$ and v is not u and v is free in α .

Examples

- Thus, if we look at the formula

$$\forall v_2 \neg (\forall v_3) (v_1 = S(v_2) \vee v_3 = v_2)$$

the variable v_1 is free whereas the variables v_2 and v_3 are not free.

- A slightly more complicated example is

$$\left(\forall v_1 \forall v_2 (v_1 + v_2 = 0) \right) \vee v_1 = S(0)$$

- In this formula, v_1 is free whereas v_2 is not free. Especially when a formula is presented informally, you must be careful about the scope of the quantifiers and the placement of parentheses.

Notes

- We will have occasion to use the informal notation $\forall x \phi(x)$. This will mean that ϕ is a formula and x is among the free variables of ϕ .
- If we then write $\phi(t)$, where t is an \mathcal{L} -term, that will denote the formula obtained by taking ϕ and replacing each occurrence of the variable x with the term t .

3.2 Deductions

Definition 3.2.1: Deduction of a Formula

Suppose that Σ is a collection of \mathcal{L} -formulas and D is a finite sequence $(\phi_1, \phi_2, \dots, \phi_n)$ of \mathcal{L} -formulas. We will say that D is a deduction from Σ if for each $i, 1 \leq i \leq n$, either

1. $\phi_i \in \Lambda$ (ϕ_i is a logical axiom), or
2. $\phi_i \in \Sigma$ (ϕ_i is a nonlogical axiom), or
3. There is a rule of inference (Γ, ϕ_i) such that $\Gamma \subseteq \{\phi_1, \phi_2, \dots, \phi_{i-1}\}$.

If there is a deduction from Σ , the last line of which is the formula ϕ , we will call this a deduction from Σ of ϕ , and write $\Sigma \vdash \phi$.

Definition 3.2.2: Logical Axioms

Given a first order language \mathcal{L} the set Λ of logical axioms is the collection of all \mathcal{L} -formulas that fall into one of the following categories:

$$\begin{aligned}
 & x = x \text{ for each variable } x \\
 & \left[(x_1 = y_1) \wedge (x_2 = y_2) \wedge \dots \wedge (x_n = y_n) \right] \rightarrow \left(f(x_1, x_2, \dots, x_n) = f(y_1, y_2, \dots, y_n) \right) \\
 & \left[(x_1 = y_1) \wedge (x_2 = y_2) \wedge \dots \wedge (x_n = y_n) \right] \rightarrow \left(R(x_1, x_2, \dots, x_n) \rightarrow R(y_1, y_2, \dots, y_n) \right) \\
 & (\forall x \phi) \rightarrow \phi_t^x \text{ if } t \text{ is substitutable for } x \text{ in } \phi \\
 & \phi_t^x \rightarrow (\exists x \phi) \text{ if } t \text{ is substitutable for } x \text{ in } \phi
 \end{aligned}$$

To refer to them easily we label them by moving down the above list E1, E2, E3, Q1, Q2

Lemma 3.2.1: Universal connection to Variable Assignment Function

$$\Sigma \vdash \theta \text{ if and only if } \Sigma \vdash \forall x \theta$$

Proof

• \Rightarrow

– Suppose that $\Sigma \vdash \theta$, therefore we have a deduction \mathcal{D} of θ , then the proof

$$\begin{aligned}
 & \mathcal{D} \\
 & \left[\left(\forall y (y = y) \right) \wedge \neg \left(\forall y (y = y) \right) \right] \rightarrow \theta \quad \text{(taut. PC)} \\
 & \left[\left(\forall y (y = y) \right) \wedge \neg \left(\forall y (y = y) \right) \right] \rightarrow (\forall x) \theta \quad \text{(QR)} \\
 & (\forall x) \theta \quad \text{(PC)}
 \end{aligned}$$

• \Leftarrow

– Suppose that $\Sigma \vdash \forall x \theta$, so we have a deduction of it, call it \mathcal{D} , then the following deduction suffices

$$\begin{aligned}
 & \mathcal{D} \\
 & \forall x \theta \\
 & \forall x \theta \rightarrow \theta_x^x \\
 & \theta_x^x
 \end{aligned}$$



3.3 Completeness

Definition 3.3.1: Consistent Set of L Formulas

Let Σ be a set of \mathcal{L} -formulas. We will say that Σ is inconsistent if there is a deduction from Σ of $[(\forall x)x = x] \wedge \neg[(\forall x)x = x]$. We say that Σ is consistent if it is not inconsistent.

Proposition 3.3.1: Contradiction has no Model

There is no \mathcal{L} -Structure \mathfrak{A} such that $\mathfrak{A} \models \perp$

Proof

Suppose there was a model of \perp , that would mean that we have an \mathcal{L} -Structure \mathfrak{A} such that $\mathfrak{A} \models \perp$. Recall that $\perp := [(\forall x)x = x] \wedge \neg[(\forall x)x = x]$, so then we would have to have $\mathfrak{A} \models (\forall x)x = x$ and also not have $\mathfrak{A} \models (\forall x)x = x$ which is a contradiction. ■

Proposition 3.3.2: Implying a Contradiction Means no Model

Let Σ be a set of sentences, then if

$$\Sigma \models \perp$$

then Σ has no model

Proof

- Recall that $\Sigma \models \perp$ means that for any \mathcal{L} -Structure \mathfrak{A} we have that

$$\mathfrak{A} \models \Sigma \text{ implies } \mathfrak{A} \models \perp$$

but recall that $\mathfrak{A} \models \perp$ is always false, therefore for the implication to hold we require that $\mathfrak{A} \models \Sigma$ to be false, which means Σ has no model as \mathfrak{A} was arbitrary. ■

Proposition 3.3.3: Logical Contradiction

Let Σ be a set of sentences, and let ϕ be a sentence, then

$$\Sigma \models \phi \text{ implies } \Sigma \cup \{\neg\phi\} \models \perp$$

Proof

- Suppose that $\Sigma \models \phi$ that means for any \mathcal{L} -Structure \mathfrak{A} if $\mathfrak{A} \models \Sigma$ then $\mathfrak{A} \models \phi$ we want to show that $\Sigma \cup \{\neg\phi\} \models \perp$, recall that by the previous theorem that means that $\Sigma \cup \{\neg\phi\}$ has no model.
- Suppose for the sake of contradiction that $\Sigma \cup \{\neg\phi\}$ has a model \mathfrak{B} then certainly $\mathfrak{B} \models \Sigma$ and by assumption we have that $\mathfrak{B} \models \phi$ but also we know that $\mathfrak{B} \models \neg\phi$ which is a contradiction, so therefore $\Sigma \cup \{\neg\phi\}$ has no model and therefore

$$\Sigma \cup \{\neg\phi\} \models \perp$$

as needed.



Lemma 3.3.1: Constant Extension still Consistent

If Σ is a consistent set of \mathcal{L} -sentences and \mathcal{L}' is an extension by constants of \mathcal{L} , then Σ is consistent when viewed as a set of \mathcal{L}' -sentences.

Proof

- Suppose for the sake of contradiction that Σ is not consistent when viewed as a set of \mathcal{L}' -sentences, so $\Sigma \vdash \perp$, considering the set of all deductions of \perp from Σ we may find a deduction \mathcal{D} which uses the least number of the newly added constants by the well ordering principle, let this number be $n \in \mathbb{N}$ and that $n > 0$ or else we would have a deduction from \mathcal{L} of \perp which would be a contradiction.
- Let v be a variable that isn't used in \mathcal{D} and let c be one of the newly added constants which is used in \mathcal{D} and let \mathcal{D}_v be created where for each line $\phi \in \mathcal{D}$ we create ϕ_v by replacing all occurrences of c in ϕ with v , and note that the last line of \mathcal{D}_v is still \perp , at this point we don't know if \mathcal{D}_v is a deduction, so we have to check that it is.
- If ϕ is an equality axiom or an element of Σ then $\phi_v \equiv \phi$ because equality axioms only contain variables and not constants, also Σ is a set of \mathcal{L} sentences and so it can't contain any of the new constants, and so ϕ_v is a valid step in the deduction since it is still an equality axiom or an element of Σ
- If ϕ is $(\forall x) \theta \rightarrow \theta_t^x$ then ϕ_v is $(\forall x) \theta_v \rightarrow (\theta_v)_{t_v}^x$, to see why we use t_v as well as θ_v try $\theta \equiv c = x$ and $t \equiv c + 3$



Notice that if we write $\Sigma \models \perp$ it means that for any \mathcal{L} -Structure \mathfrak{A} if $\mathfrak{A} \models \Sigma$ then $\mathfrak{A} \models \perp$ by the above discussion that forces $\mathfrak{A} \models \Sigma$ to be false, and therefore Σ has no model.

Theorem 3.3.1: Completeness Theorem

Suppose that Σ is a set of \mathcal{L} -formulas, where \mathcal{L} is a countable language and ϕ is an \mathcal{L} -formula. If $\Sigma \models \phi$, then $\Sigma \vdash \phi$.

Setup

- We start by assuming that $\Sigma \models \phi$, we must show that $\Sigma \vdash \phi$.
- If ϕ is not a sentence then we can always prove ϕ' which is the same as ϕ with all of its variables bound
 - We can do that by appending $(\forall x_i)$ where each x_i is a free variable of ϕ to the front of ϕ
- Therefore we will prove it for all sentences ϕ

3.4 Incompleteness

Definition 3.4.1: Representable Set

A set $A \subseteq \mathbb{N}^k$ is said to be representable (in N) if there is an \mathcal{L}_{NT} -formula $\phi(\underline{x})$ such that

$$\begin{aligned} \forall a \in A \quad N \vdash \phi(\bar{a}) \\ \forall \underline{b} \notin A \quad N \vdash \neg \phi(\bar{b}) \end{aligned}$$

In this case we will say that the formula ϕ represents the set A .

Definition 3.4.2: Weakly Representable Set

A set $A \subseteq \mathbb{N}^k$ is said to be weakly representable (in N) if there is an \mathcal{L}_{NT} -formula $\phi(\underline{x})$ such that

$$\begin{aligned} \forall a \in A \quad N \vdash \phi(\bar{a}) \\ \forall \underline{b} \notin A \quad N \not\vdash \phi(\bar{b}) \end{aligned}$$

In this case we will say that the formula ϕ weakly represents the set A .

Definition 3.4.3: Total Function

Suppose that $A \subseteq \mathbb{N}^k$ and suppose that $f : A \rightarrow \mathbb{N}$. If $A = \mathbb{N}^k$ we will say that f is a total function

Definition 3.4.4: Partial Function

Suppose that $A \subsetneq \mathbb{N}^k$ and suppose that $f : A \rightarrow \mathbb{N}$, then we will call f a partial function.

Definition 3.4.5: Representable Function

Suppose that $f : \mathbb{N}^k \rightarrow \mathbb{N}$ is a total function. We will say that f is a representable function (in N) if there is an \mathcal{L}_{NT} formula $\phi(x_1, \dots, x_{k+1})$ such that, for all $a_1, a_2, \dots, a_{k+1} \in \mathbb{N}$

- If $f(a_1, \dots, a_k) = a_{k+1}$, then $N \vdash \phi(\bar{a}_1, \dots, \bar{a}_{k+1})$
- If $f(a_1, \dots, a_k) \neq a_{k+1}$, then $N \vdash \neg \phi(\bar{a}_1, \dots, \bar{a}_{k+1})$

Definition 3.4.6: Definable Set

We will say that a set $A \subseteq \mathbb{N}^k$ is definable if there is a formula $\phi(\underline{x})$ such that

$$\begin{aligned} \forall a \in A \quad \mathfrak{N} \models \phi(\bar{a}) \\ \forall \underline{b} \notin A \quad \mathfrak{N} \models \neg \phi(\bar{b}) \end{aligned}$$

In this case, we will say that ϕ defines the set A .

Corollary 3.4.1: I

$A \subseteq \mathbb{N}^k$ is definable by a Δ formula, then it is representable

4

Topology

4.1 Topological Spaces and Continuous Functions

4.1.1 Basis for a Topology

Definition 4.1.1: Basis

If X is a set, a basis for a topology on X is a collection \mathcal{B} of subsets of X (called basis elements) such that

1. For each $x \in X$, there is at least one basis element B containing x .
2. If x belongs to the intersection of two basis elements B_1 and B_2 , then there is a basis element B_3 containing x such that $B_3 \subset B_1 \cap B_2$.

If \mathcal{B} satisfies these two conditions, then we define the topology \mathcal{T} generated by \mathcal{B} as follows: A subset U of X is said to be open in X (that is, to be an element of \mathcal{T}) if for, each $x \in U$, there is a basis element $B \in \mathcal{B}$ such that $x \in B$ and $B \subset U$. Note that each basis element is itself an element of \mathcal{T} .

4.1.2 The Subspace Topology

Definition 4.1.2: Subspace Topology

Let X be a topological space with topology \mathcal{T} . If Y is a subset of X , the collection

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\}$$

is a topology on Y , called the subspace topology. With this topology, Y is called a subspace of X ; its open sets consist of all intersections of open sets of X with Y .

4.1.3 The Product Topology

Definition 4.1.3: Product Topology

Let \mathcal{S}_β denote the collection

$$\mathcal{S}_\beta = \left\{ \pi_\beta^{-1}(U_\beta) \mid U_\beta \text{ open in } X_\beta \right\}$$

and let \mathcal{S} denote the union of these collections,

$$\mathcal{S} = \bigcup_{\beta \in I} \mathcal{S}_\beta$$

The topology generated by the subbasis \mathcal{S} is called the product topology. In this topology $\prod_{\alpha \in I} X_\alpha$ is called a product space.

Definition 4.1.4: Box Topology

Let $\{X_\alpha\}_{\alpha \in J}$ be an indexed family of topological spaces. Let us take as a basis for a topology on the product space

$$\prod_{\alpha \in J} X_\alpha$$

the collection of all sets of the form

$$\prod_{\alpha \in J} U_\alpha$$

where U_α is open in X_α , for each $\alpha \in J$. The topology generated by this basis is called the box topology.

Theorem 4.1.1: Basis for the Box Topology

Suppose the topology on each space X_α is given by a basis \mathcal{B}_α . The collection of all sets of the form

$$\prod_{\alpha \in J} B_\alpha$$

where $B_\alpha \in \mathcal{B}_\alpha$ for each α , will serve as a basis for the box topology on $\prod_{\alpha \in J} X_\alpha$.

Theorem 4.1.2: Basis for the Product Topology

Suppose the topology on each space X_α is given by a basis \mathcal{B}_α . The collection of all sets of the form

$$\prod_{\alpha \in J} B_\alpha$$

where $B_\alpha \in \mathcal{B}_\alpha$ for finitely many indices α and $B_\alpha = X_\alpha$ for all the remaining indices, will serve as a basis for the product topology $\prod_{\alpha \in J} X_\alpha$.

Definition 4.1.5: \mathbb{R}^ω

\mathbb{R}^ω , the countably infinite product of \mathbb{R} with itself. Recall that

$$\mathbb{R}^\omega = \prod_{n \in \mathbb{N}} X_n$$

with $X_n = \mathbb{R}$ for each n

4.1.4 The Metric Topology**Definition 4.1.6: A metric**

A metric on a set X is a function

$$d : X \times X \rightarrow \mathbb{R}$$

having the following properties:

1. $d(x, y) \geq 0$ for all $x, y \in X$; equality holds if and only if $x = y$.
2. $d(x, y) = d(y, x)$ for all $x, y \in X$.
3. Triangle Inequality: $d(x, y) + d(y, z) \geq d(x, z)$, for all $x, y, z \in X$.

Example 4.1.1: Discrete Metric

$d : X \times X \rightarrow \mathbb{R}$ given by

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

Definition 4.1.7: Epsilon Ball

Given $\epsilon > 0$, consider the set

$$B_d(x, \epsilon) = \{y \mid d(x, y) < \epsilon\}$$

of all points y whose distance from x is less than ϵ . It is called the ϵ -ball centered at x . Sometimes we omit the metric d from the notation and write this ball simply as $B(x, \epsilon)$, when no confusion will arise.

Lemma 4.1.1: Epsilon Ball Contains Another

Let $x \in X$ then for every $B(x, \epsilon)$ there is $y \in B(x, \epsilon)$ and $\delta \in \mathbb{R}^+$ such that

$$B(y, \delta) \subseteq B(x, \epsilon)$$

Proof

- Let $y \in B(x, \epsilon)$ we claim $B(y, \delta)$ where $\delta = \epsilon - d(x, y)$ works.
- Note $B(y, \delta) \subseteq B(x, \epsilon)$, since given any $z \in B(y, \delta)$ we have that $d(y, z) < \epsilon - d(x, y)$, rearranging gives us $d(x, y) + d(y, z) < \epsilon$ and by the triangle inequality we get $d(x, z) < \epsilon$ therefore $z \in B(x, \epsilon)$

■

Proposition 4.1.1: Epsilon Balls Form A Basis

The collection \mathcal{B} of all ϵ -balls $B_d(x, \epsilon)$, for $x \in X$ and $\epsilon > 0$, is a basis for a topology on X

Proof

- Let $x \in X$ then clearly $x \in B_d(x, \epsilon)$ works for any $\epsilon \in \mathbb{R}^+$
- Let $B_1, B_2 \in \mathcal{B}$ and let $y \in B_1 \cap B_2$, for each of B_1 and B_2 we have δ_1 and δ_2 respectively with $B(y, \delta_1) \subseteq B_1$ and $B(y, \delta_2) \subseteq B_2$, by taking $\delta = \min(\delta_1, \delta_2)$ then we have $B(y, \delta) \subseteq B_1 \cap B_2$ as needed.

■

Definition 4.1.8: Metric Topology

If d is a metric on the set X , then the collection of all ϵ -balls $B_d(x, \epsilon)$, for $x \in X$ and $\epsilon > 0$, is a basis for a topology on X , called the metric topology induced by d .

Definition 4.1.9: Bounded Subset of a Metric Space

Let X be a metric space with metric d . A subset A of X is said to be bounded if there is some number $M \in \mathbb{R}$ such that

$$d(a_1, a_2) \leq M$$

for every pair of points $a_1, a_2 \in A$

Example 4.1.2: Function on \mathbb{R}^ω

Consider a function $h : \mathbb{R}^\omega \rightarrow \mathbb{R}^\omega$ defined by

$$h(x_1, x_2, \dots) = (\alpha_1 x_1, \alpha_2 x_2, \dots)$$

with $\alpha_1, \alpha_2, \dots \in \mathbb{R} \setminus \{0\}$

1. Is h continuous, when \mathbb{R}^ω is given the product topology?
2. Is h continuous, when \mathbb{R}^ω is given the box topology?
3. Is h continuous, when \mathbb{R}^ω is given the uniform topology?

Proof

1. • Take an arbitrary basis element from the product topology, that is:

$$\prod_{\alpha \in \mathcal{J}} B_\alpha$$

where $B_\alpha \in \mathcal{B}_\mathbb{R}$ for finitely many indices α and $B_\alpha = \mathbb{R}$ for all the remaining indices.

α : Now note the following

- The inverse image of each of these intervals is a scaled version of itself therefore it is still an interval of \mathbb{R} , and it is therefore still open with respect to \mathbb{R}
- The inverse image of \mathbb{R} under any scaling is still \mathbb{R}
- Therefore $h^{-1}\left(\prod_{\alpha \in \mathcal{J}} B_\alpha\right)$ is a product of finitely many intervals, and infinitely many \mathbb{R} 's and so it is open with respect to \mathbb{R}^ω as needed.

2. • Given an arbitrary basis element of the box topology we have

$$\prod_{\alpha \in \mathcal{J}} B_\alpha$$

where $B_\alpha \in \mathcal{B}_\mathbb{R}$ for each α

- Due to α we know that each of the intervals become new scaled intervals and so $h^{-1}\left(\prod_{\alpha \in \mathcal{J}} B_\alpha\right)$ is a product of open intervals and is therefore open in with respect \mathbb{R}^ω equipped with the box topology

3. Recall that a basis for the uniform topology are epsilon balls with radius less than 1 if that's the case, then so long as every α_i is greater than 1 then the inverse image scales them to be even smaller, resulting in an open set, otherwise the function h wouldn't be continuous, since all the α_i 's are greater than one, it is continuous.

■

4.2 Connectedness and Compactness

Definition 4.2.1: Connected Space

Let X be a topological space. A separation of X is a pair U, V of disjoint nonempty open subsets of X whose union is X . The space X is said to be connected if there does not exist a separation of X .

Notice that U, V are actually clopen, as $X \setminus U = V$ and $X \setminus V = U$ stating that V and U are closed as well.

Example 4.2.1: Closed and Bounded, not Compact

A metric space X and a closed and bounded subspace Y of X that is not compact.

- Consider the set $X = \left\{ \frac{1}{n} : n \in \mathbb{N}^+ \right\}$, with the discrete metric, it is bounded because for any two points $a, b \in X$, $d(a, b) \leq 1$
- Let X be an infinite set and let consider the discrete metric on that set, the metric topology which it induces (call it \mathcal{T}) is the discrete topology of X . Therefore if we consider any subset Y of X it is closed, as $X \setminus Y \in \mathcal{T}$ (remember it's the discrete topology). But the open covering $\{\{x\} : x \in X\}$ has no finite subcollection which also covers X .

Example 4.2.2: \mathbb{R}^ω Connected?

Consider the product, uniform, and box topologies on \mathbb{R}^ω . In which topologies is \mathbb{R}^ω connected?

Proof

- Consider the product topology, we will show that \mathbb{R}^ω is not connected. Consider the set A of bounded sequences in \mathbb{R}^ω , and the complement of A , namely the unbounded sequence of \mathbb{R}^ω let's label this set as $B = \mathbb{R}^\omega \setminus A$.
 - Note that A is open in the box topology as if we fix any $\varepsilon > 0$ we may define for any $\vec{x} \in \mathbb{R}^\omega$

$$U_{\vec{x}}^D = (x_1 - \varepsilon, x_1 + \varepsilon) \times (x_2 - \varepsilon, x_2 + \varepsilon) \times \dots$$
 - If \vec{a} is a bounded sequence then U is a set of bounded sequences, as it only ever adds constants to the terms of \vec{a} , therefore $\vec{a} \in U_{\vec{a}}$ and $U_{\vec{a}} \subseteq A$, therefore A is open.
 - If \vec{b} is an unbounded sequence then U is a set of unbounded sequences, as adding a constant to every element to an unbounded sequence yields an unbounded sequence. So for every $\vec{b} \in B$ there we have $\vec{b} \in U_{\vec{b}}$ and $U_{\vec{b}} \subseteq B$ so B is open.
 - Therefore A non-trivial set which is both open and closed in \mathbb{R}^ω and so it is disconnected.
- Note that \mathbb{R}^ω is closed with the uniform topology by following the same proof with $\varepsilon = 1$
- Now if we consider \mathbb{R}^ω with the product topology we will see that \mathbb{R}^ω is connected.
 - Let $\mathbb{R}_0^n = \left\{ (x_1, x_2, \dots) : \text{where for } i > n \text{ we have that } x_i = 0 \right\}$
 - \mathbb{R}_0^n is homeomorphic to \mathbb{R}^n and since \mathbb{R} is connected, and the finite cartesian product of connected spaces is connected, we get that \mathbb{R}_0^n is connected.
 - We now will show that the closure of \mathbb{R}^∞ is equal to \mathbb{R}^ω and then by the fact that the closure of a connected set is closed we obtain that \mathbb{R}^ω is connected.
 - To show that \mathbb{R}^ω is equal to the closure of \mathbb{R}^∞ we proceed by using the fact that a point is part of the closure of a set if and only if every basis element intersects it:
 - Let $\vec{a} \in \mathbb{R}^\omega$ and let $B = \prod_{i \in \mathbb{N}} B_i$ be some basis element for the product topology with $\vec{a} \in U$, now we have to show this set intersects \mathbb{R}^∞ . Since only finitely many of the B_i are equal to basis elements and the rest are equal to \mathbb{R} that means there is some $N \in \mathbb{N}$ such that $\forall j \in \mathbb{N}^{\geq N}, B_j = \mathbb{R}$, therefore the point $\vec{b} = (a_1, a_2, \dots, a_{N-1}, a_N, 0, 0, 0, \dots) \in B \cap \mathbb{R}^\infty$



Proposition 4.2.1: Connected Implies Closure Connected

Let $A \subseteq X$ be a connected subspace of X then \overline{A} is also connected.

Proof

- Suppose for the sake of contradiction that there is a separation B, C of \overline{A} , then $B \cup C = \overline{A}$ and note that that means that $A \subseteq B \cup C$ since $A \subseteq \overline{A}$ so $(B \cup C) \cap A = \overline{A} \cap A = A$.
- Therefore $(B \cap A) \cup (C \cap A) = A$ is a separation of A noting that $B \cap A$ and $C \cap A$ are non-empty because ...



Definition 4.2.2: Totally Disconnected

A topological space is totally disconnected if it's only connected subspaces are one-point sets.

Consider \mathbb{R}_ℓ if we have $\{a\}$ and $\{b\}$ then the open sets $(-\infty, b)$ and $[b, \infty)$ is a separation of \mathbb{R}_ℓ , therefore it's disconnected. Similarly for any two points a, b in \mathbb{Q} we have some irrational number r between the two, and thus $(-\infty, r)_\mathbb{Q}, (r, \infty)_\mathbb{Q}$ is a separation thus \mathbb{Q} is totally disconnected under this topology.

Let's look at \mathbb{R} with the finite complement topology. Right off the bat, we note that if a set is finite in \mathbb{R} it's complement must be infinite therefore if \mathbb{R} was completely disconnected it would mean for any singleton sets, we have a separation U, V , but that means that U and V must be infinite, but we then get a contradiction as $\mathbb{R} = U \cup V$ so $\mathbb{R} \setminus U = V$, now since U was open this implies that V is finite, which is a contradiction. This idea may be extended to \mathbb{R}^2 .

4.2.1 Compact Spaces

Definition 4.2.3: Covering

A collection A of subsets of a space X is said to cover X , or to be a covering of X , if the union of the elements of A is equal to X . It is called an open covering of X if its elements are open subsets of X .

Definition 4.2.4: Compact Space

A space X is said to be compact if every open covering A of X contains a finite subcollection that also covers X .

Lemma 4.2.1: Covering Yields Finite Covering if and only if Compact

Let Y be a subspace of X . Then Y is compact if and only if every covering of Y by sets open in X contains a finite subcollection covering Y .

Computer Science

Definition 5.0.1: Big-O

Let $f, g : \mathbb{N} \rightarrow \mathbb{R}^{\geq 0}$, then we define the set

$$\mathcal{O}(g) \stackrel{\text{D}}{=} \left\{ j : \exists c, B \in \mathbb{R}^+, \forall n \in \mathbb{N}, n \geq B \rightarrow j(n) \leq cg(n) \right\}$$

And we say that f is in the big-O of g when $f \in \mathcal{O}(g)$

Definition 5.0.2: Little-o

We define the following set:

$$o(g) = \left\{ j : \lim_{x \rightarrow 0} \frac{j(x)}{g(x)} = 0 \right\}$$

We say that f is in the little-o of g when $f \in o(g)$

6

Probability

Definition 6.0.1: Conditional Independence

Suppose A and B are independent events, then we would want $P(A|C) = P(A|B \cap C)$:

$$P(A|C) = \frac{P(A \cap C)}{P(C)} \text{ and } P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)}$$

So then we want:

$$\begin{aligned} P(A|C) &= P(A|B \cap C) \\ \Updownarrow \\ \frac{P(A \cap C)}{P(C)} &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ \Updownarrow \\ \frac{P(A \cap B \cap C)}{P(C)} &= \frac{P(A \cap C) P(B \cap C)}{P(C) P(C)} \\ \Updownarrow \\ P(A \cap B|C) &= P(A|C) P(B|C) \end{aligned}$$

6.1 Bayesian Inference

7

Programming

Working in the Shell

Convert all files names to lowercase

```
user@machine:~# for f in `find`; do mv -v "$f" "$(echo $f | tr 'A-Z' 'a-z')"; done
```