Multivariable Calculus

Avery Karlin

Fall 2015

Contents

1 Multivariable Calculus Basics **5**
1.1 Green's Theorem . . . 5
1.2 Generalized FTC . . . 5

2 Real Number Set **6**
2.1 Definitions of Number Structures . . . 6
2.1.1 $\mathbb{N}$ . . . 6
2.1.2 $\mathbb{Z}$ . . . 6
2.1.3 $\mathbb{Q}$ . . . 6
2.1.4 $\mathbb{R}$ . . . 6
2.2 Basic Structure of $\mathbb{R}^1$ . . . 7
2.3 Properties of $\mathbb{R}^n$ . . . 7
2.4 Basic Axioms for $\mathbb{R}^1$ . . . 7
2.5 The Seperation Axiom . . . 8
2.5.1 Proof of Irrationals . . . 8

3 Sequence Theorems **9**
3.1 The Least Upper Bound Theorem . . . 9
3.1.1 Proof . . . 9
3.2 Bounded Monotone Sequence Theorem . . . 9
3.2.1 Proof . . . 9
3.3 Archimedean Property . . . 10
3.3.1 Proof . . . 10
3.3.2 Consequences . . . 10
3.4 Sunrise Lemma . . . 10
3.4.1 Well-Ordering Property . . . 10
3.4.2 Proof . . . 10
3.5 Bolzano-Weierstrass Theorem . . . 11
3.5.1 Proof . . . 11

4 Extreme Value Theorem **11**
4.1 Proof . . . 11
4.1.1 $\exists M > 0$ such that $f(x) \leq< M \forall x \in [a, b]$ (f(x) is bounded above) . . . 11
4.1.2 The function reaches the supremum and infimum . . . 11

5 Higher Dimensional Mathematics **12**
5.1 Higher Dimensions . . . 12
5.1.1 Cartesian Product . . . 12
5.1.2 Shapes . . . 12
5.1.3 Boundary . . . 12
5.1.4 Interior . . . 12
5.1.5 Exterior . . . 12
5.1.6 Closure . . . 12
5.1.7 Ordered Pair . . . 13
5.1.8 Vector and Points . . . 13
5.1.9 Space and Lines . . . 13

5.2 Dot Product . . . 13
5.2.1 Definition and Perpendicularity . . . 13
5.2.2 Properties . . . 14
5.3 Functions in Higher Dimensions . . . 14
5.3.1 Domain, Range, and One-to-One . . . 14
5.3.2 Bolzano-Weirstrauss in Higher Dimensions . . . 14
5.3.3 Cauchy Sequence . . . 14
5.3.4 Convergence as Cauchy . . . 14
5.3.5 Cauchy's Convergence Theorem . . . 15
5.3.6 Cauchy Sequences in Higher Dimensions . . . 15
5.4 Metric Topology in $\mathbb{R}^{\times}$ . . . 16
5.4.1 Continuity . . . 16
5.4.2 Directional Limits . . . 16
5.4.3 Limits in Higher Dimensions . . . 16
5.5 Properties of Domain . . . 16
5.5.1 Bounded . . . 16
5.5.2 Closed . . . 17
5.5.3 Open Set Theorem . . . 17
5.6 Extreme-Value Theorem . . . 17

6 Distance Functions **18**
6.1 Euclidean Distance . . . 18
6.2 Geometric Distance . . . 19
6.3 Basic Distance Bounds Lemma . . . 19
6.4 Other Distance . . . 20
6.5 Distances Between Sets . . . 20

7 Inequalities **20**
7.1 Level of Operations . . . 20
7.2 AM-GM . . . 20
7.2.1 AM-GM Inequality . . . 21
7.2.2 Proof . . . 21
7.3 Young's Inequality . . . 22
7.3.1 Hölder Conjugate . . . 22
7.3.2 Young's Inequality . . . 22
7.3.3 Proof . . . 22
7.4 Hölder's Inequality . . . 23
7.4.1 p-norm . . . 23
7.4.2 Hölder's Inequality . . . 23
7.4.3 Proof . . . 23
7.5 Cauchy-Schwarz Inequality . . . 24
7.6 Triangle Inequality . . . 24
7.7 Reverse Triangle Inequality . . . 24
7.8 Minkowski's Inequality . . . 25
7.8.1 Rational Power Proof . . . 25
7.8.2 Proof . . . 25

8 Heine-Borel Theorem and Domain Properties **26**

8.1 Heine-Borel Theorem . . . 26
8.1.1 Proof . . . 26
8.2 Uniform Continuity . . . 26
8.3 Uniform Continuity Theorem . . . 26
8.4 Connected . . . 27
8.4.1 Definition . . . 27
8.4.2 Theorem . . . 27

9 Intermediate Value Theorem **27**
9.1 Theorem in $\mathbb{R}$ . . . 27
9.1.1 Proof . . . 28
9.2 Theorem in $\mathbb{R}^{\ltimes}$ . . . 28
9.2.1 Proof . . . 28

10 Vector Valued Functions **28**
10.1 Definition . . . 28
10.2 Continuity . . . 28
10.2.1 Component-wise Nature of Continuity . . . 29
10.2.2 Composition of Continuous Functions . . . 29
10.3 Compactness Theorem . . . 30
10.3.1 Theorem . . . 30
10.3.2 Proof . . . 30
10.4 Connectedness Theorem . . . 30
10.4.1 Theorem . . . 30
10.4.2 Proof . . . 30

11 Sequences of Functions **31**
11.1 Infinity Norm . . . 31
11.2 Pointwise Convergence: . . . 31
11.3 Uniform Convergence . . . 31
11.3.1 In $\mathbb{R}$ . . . 31
11.3.2 In $\mathbb{R}^{\ltimes}$ . . . 31
11.4 Uniform Convergence Theorem . . . 32
11.4.1 Theorem . . . 32
11.4.2 Proof . . . 32

12 Differentiation **32**
12.1 Differentiable in $\mathbb{R}$ . . . 32
12.2 Differentiable in $\mathbb{R}^{\ltimes}$ . . . 32
12.3 Gradient . . . 33
12.3.1 Gradient Representation . . . 33
12.3.2 Gradient Calculation . . . 33
12.4 Directional Derivative . . . 34
12.5 Vector Valued Directional Derivative . . . 35

Teacher: Stern

1 Multivariable Calculus Basics

Multivariable calculus is the study of the analogs, if any, of the fudemental theorem of calculus to higher dimensions, and any restrictions that may exist on higher dimensions

Let R be a simple closed curve, where simple means that each point is crossed by the curve at most once and closed means that the curve does not have a unqiue starting point and end point.

When integrating in 2 dimensions, we also need to pick an "interval." In this case, R, the bounded region, would be an interval (analogous to (a, b) in single variable calculus) and C would be the boundary (analogous to a, b).

If we were to integrate a function $f(x, y)$ over R, it is denoted by:

$$\iint_R f(x, y) dA_{xy}$$

The dA_{xy} is the area element, an infinitesimally small piece of area, analogous to dx, the length element in single variable calculus.

Note that in single variable calculus, there is an implied orientation, going left to right is the positive "direction," In multivariable calculus, it is accepted that the positive direction for the curve to go in is the counterclockwise direction, such that the opposite direction adds a negative.

1.1 Green's Theorem

This is one of the FTC's generalization to higher dimensions. The Green's Theorem works with functions that take in 2 variables.

Suppose there exists $f(x, y)$ and $g(x, y)$, and a region R bounded by a positively oriented, simple closed curve C. Then Green's theorem states that:

$$\iint_R (\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}) dA_{xy} = \int_C f dx + g dy = \int_C (f x'(t) + g y'(t)) dt$$

1.2 Generalized FTC

The goal of this course in multivariable calculus is to reach the following conclusion:

For some function ω evaluated over the region M, where ∂M is the boundry of M:

$$\int_M d\omega = \int_{\partial M} \omega$$

2 Real Number Set

2.1 Definitions of Number Structures

2.1.1 $\mathbb{N}$

We can define the natural number system by sets, like the following:

$$0 = \emptyset$$

And from there we introduce a succession operation:

$$n + 1 = n \cup \{n\}$$

So for example, $1 = \{0\} = \{\emptyset\}$, $1 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$, etc.

Set theory is the deepest concept in mathematics, such that it must be assumed as postulates.

2.1.2 $\mathbb{Z}$

Positive integers are defined as an ordered pair of natural numbers, for example, $2 = (2, 0)$, and $-2 = (0, 2)$

2.1.3 $\mathbb{Q}$

Rationals are defined as an infinite set of ordered pairs of integers. A rational number $q = \frac{n}{m}$, then $q = \{(n, m), (2n, 2m), (3n, 3m) \dots (-2n, -2m), (-3, -3m) \dots\}$

2.1.4 $\mathbb{R}$

There are several definitions of the real number system

- Each real number can be thouhgt of as an infinite sequence in the following format:

  $$(s, N, d_1, d_2, d_3 \dots)$$

  Where $s = \pm 1$, $N \in \mathbb{N}$, and $d_n \in 0, 1, 2, 3, \dots 8, 9$. It is also not the case that $d_n = d_{n+1} = \dots = 0$. If there is a terminal decimal, we express it as 9 repeated.

- Each real number forms a subdivision of $\mathbb{Q}$ into two disjoint sets that cover the entirety of $\mathbb{Q}$, one of which lies entirely to the left of the other.

2.2 Basic Structure of $\mathbb{R}^1$

$\mathbb{R}$ is an instance of many kinds of mathematical structures, such as:

- Field
  - A set closed under addition and multiplication (results are within set)
  - Obeys all associated laws of addition and multiplication
- Ordered Set
  - The set has an ordering that reflect the operations in the field (such that x, y ¿ 0, then x + y and xy ¿ 0)
  - This structure is what allows for comparisons, like the $<$ function
- Metric Space
  - There exists a standard distance operation between numbers
  - In $\mathbb{R}$, $dist(x, y) := |x - y|$ ($:=$ means "is defined as")
- Vector Space
  - Elements can be thought of as vectors from the origin
- Geometric Space
  - This structure means that you can measure angle in a meaningful way
  - In $\mathbb{R}^1$, the angle measure can be either 0 or π
  - In higher dimensions there are more angles possible

2.3 Properties of $\mathbb{R}^n$

In higher dimensions, several of the properties of $\mathbb{R}$ no longer hold. $\mathbb{R}^n, n > 1$ is not a field, and not an ordered set, but it is a vector, metric, and geometric space.

2.4 Basic Axioms for $\mathbb{R}^1$

$\mathbb{R}$ is a field under $+$ and $\cdot$ $(x, y \in \mathbb{R})$

1. Additive closure: $x + y \in \mathbb{R}$
2. Associative Property of Addition: $x + (y + z) = (x + y) + z$
3. Communicative Property of Addition: $x + y = y + x$
4. 0 is the identity element of addition: $x + 0 = x$
5. Every element has an additive inverse: $x + (-x) = 0$
6. Multiplicative closure: $xy \in \mathbb{R}$
7. Associative Property of Multiplication: $x(yz) = (xy)z$

8. Communicative Property of Multiplication: $xy = yx$
9. 1 is the identity element of multiplication: $x(1) = x$
10. Every element (except 0) has a multiplicative inverse: $x \cdot \frac{1}{x} = 1$
    - Theorem: $x(0) = 0$
    - Proof: $x(0) = x(0+0)$, then we apply the distributive law, and get $x(0) = x(0) + x(0)$. Now we add $-x(0)$ to both side, and we get: $0 = x(0)$
11. Distributive Law: $x(y+z) = xy + xz$

$\mathbb{R}$ is an ordered field and has a proper subset (aka not the entire set) $\mathbb{R}^+$ (the <u>positives</u>) such that:

1. $\mathbb{R}^+$ is closed under $+$ and $\cdot$
2. $1 \in \mathbb{R}^+, 0 \notin \mathbb{R}^+$
3. **Trichotomy Property**: for any $x \in \mathbb{R}$, x is either 0, $\in \mathbb{R}^+$ or $\notin \mathbb{R}^+$
4. Definition of $<$ and $>$:
    - $x < y$ means $y - x \in \mathbb{R}^+$
    - $x > y$ means $y < x$

2.5 The Seperation Axiom

The main difference between the $\mathbb{Q}$ and the $\mathbb{R}$ is the separation axiom which the rationals do not have. If $\mathcal{A} \subseteq \mathbb{R}$ and $\mathcal{B} \subseteq \mathbb{R}$, and:

1. $\mathcal{A} \cap \mathcal{B} = \emptyset$
2. $\mathcal{A} \neq \emptyset$, $\mathcal{B} \neq \emptyset$
3. $\mathcal{A} < \mathcal{B}$ "$\mathcal{A}$ is to the left of $\mathcal{B}$"
    - $\forall$ (for all) $a \in \mathcal{A}, \forall b \in \mathcal{B}, a < b$

2.5.1 Proof of Irrationals

$$\mathcal{A} = \mathbb{Q}^- \cup \{0\} \cup \{q \in \mathbb{Q}^+ | q^2 < 2\}$$
$$\mathcal{B} = \{q \in \mathbb{Q}^+ | q^2 \geq 2\}$$

Then, $\exists$ (there exists at least 1) $c \in \mathbb{R}$ such that $\mathcal{A} \leq c \leq \mathcal{B}$

We know that $\mathcal{A} \neq \emptyset$ and $\mathcal{B} \neq \emptyset$ because 0 is in $\mathcal{A}$ and 2 is in $\mathcal{B}$. We also know that $\mathcal{A} \cup \mathcal{B} = \mathbb{Q}$ and $\mathcal{A} \cap \mathcal{B} = \emptyset$. As well as the fact that $\mathcal{A} < \mathcal{B}$ (all elements of A are less than all elements of B.

Now, if we want to find the boundary element, q_0, which separates $\mathcal{A}$ and $\mathcal{B}$. We know that $\mathcal{A} \leq q_0 \leq \mathcal{B}$. So that must mean $q_0 = \sqrt{2}$. However, $\sqrt{2} \notin \mathbb{Q}$. Therefore, we know that the set of rational numbers do not follow the separation axiom.

3 Sequence Theorems

3.1 The Least Upper Bound Theorem

If $\mathcal{A} \subseteq \mathbb{R}$ is non-empty, and is <u>bounded above</u> (So $\exists b_1 \in \mathbb{R}$ such that $\mathcal{A} < b_1$), then $\mathcal{A}$ has a <u>least upper bound</u>, i.e. a number $b_0 \in \mathbb{R}$ such that $\mathcal{A} \leq b_0$ and for any b with $\mathcal{A} \leq b$, $b_0 \leq b$

$$\mathcal{A} \subseteq \mathbb{R}, \mathcal{A} \neq \emptyset, (\exists\ b_1 \in \mathbb{R} : \mathcal{A} \leq b_1) \rightarrow [\exists\ b_0 \in \mathbb{R} : \mathcal{A} < b_0, \forall b \in \mathbb{R}(\mathcal{A} \leq b \rightarrow b_0 \leq b)]$$

b_0 is known as the least possible upper bound, or the *supremum* of $\mathcal{A}$, we write $b_0 = \sup \mathcal{A}$.

Similarly, for any non-empty set $\mathcal{A}$ bounded below, it has a <u>greatest lower bound</u>, $\inf \mathcal{A}$, called the *infimum* of $\mathcal{A}$

3.1.1 Proof

Define $\mathcal{B}$ to be the set of all upper bounds of $\mathcal{A}$. Let $\mathcal{C} = \mathbb{R} \setminus \mathcal{B}$. Clearly $\mathcal{B}$ is nonempty; also $\mathcal{C}$ is non-empty because it contains $x_0 - 1$, where $x_0 \in \mathcal{A}$. By the way in which we defined $\mathcal{C}$, $\mathcal{B} \cap \mathcal{C} = \emptyset$. Pick any $c \in \mathcal{C}$ and $b \in \mathcal{B}$. By the definition of $\mathcal{C}$, $\exists\ x_1 \in \mathcal{A} : c < x_1$. But $x_1 \leq b$ by the definition of $\mathcal{B}$. Therefore $\mathcal{C} < \mathcal{B}$. By the separation postulate, $\exists\ b_0 \in \mathbb{R} : \mathcal{C} \leq b_0 \leq \mathcal{B}$. Note that $\mathcal{A} \setminus \{b_0\} \subseteq \mathcal{C}$. Thus, b_0 is an upper bound for $\mathcal{A}$. Morever, it is the least upper bound because $b_0 \leq \mathcal{B}$.

3.2 Bounded Monotone Sequence Theorem

For any sequence $\{a_n\}$

1. If $a_n \leq a_{n+1}$ for all $n \geq 1$, and $\exists\ b \in \mathbb{R}$ such that $a_n \leq b$ for all $n \geq 1$, then $\lim_{n\to\infty} a_n$ exists and is less than or equal to b
2. If $a_n \geq a_{n+1}$ for all $n \geq 1$, and $\exists\ b \in \mathbb{R}$ such that $a_n \geq b$ for all $n \geq 1$, then $\lim_{n\to\infty} a_n$ exists and is greater than or equal to b

3.2.1 Proof

We first convert the sequence $\{a_n\}$, which is bounded by b into the set $\mathcal{A} = \{a_n | n \geq 1\}$. We know that $\mathcal{A} \neq \emptyset$ because the sequence has some terms. We also know that $\mathcal{A}$ is bounded above by b; $\mathcal{A} < b$ By the Least Upper Bound Theorem, $\exists\ b_0 = sup\mathcal{A}$. We now show that $b_0 = \lim_{n\to\infty} a_n$. By the definition of limits, to say $b_0 = \lim_{n\to\infty} a_n$ means to say $\forall \epsilon > 0, \exists\ N > 0, \forall n \geq N$, $|a_n - b_0| < \epsilon$

If we look at the number $b_0 - \epsilon$, it is not an upper bound on $\mathcal{A}$ because b_0 is the least upper bound and $\epsilon > 0$. Therefore, $\exists\ a_N > b_0 - \epsilon$. Since $\{a_n\}$ is increasing, $\forall n > N$, $a_n > b_0 - \epsilon$. If we rearrange the terms, we get $b_0 - a_n < \epsilon$. Therefore, b_0 (which exists by the least upper bound theorem) is the limit of a_n as $n \to \infty$.

3.3 Archimedean Property

For any positive numbers x and y, it is possible to find some $n \in \mathbb{N}$ such that $nx > y$.

$$\forall x, y > 0, \exists\ n \in \mathbb{N} : nx > y$$

3.3.1 Proof

Assume that $\neg\exists\ n : nx > y$, this is logically equivalent to $\forall n : nx \leq y$. Let $\mathcal{C} = \{nx | n \in \mathbb{N}\}$. Then $\mathcal{C} \leq y$, let $c = \sup \mathcal{C}$. We claim that $\exists N : c - \frac{1}{2}x < Nx \leq c$. This is true because if such N does not exist, then $c - \frac{1}{2}x$ would be an upper bound, but c is the least upper bound, so such N must exist. Now we've established the existence of N, let us consider $(N+1)x$. $(N+1)x = Nx + x > (c - \frac{1}{2}x) + x = c + \frac{1}{2}x > c$. But $(N+1)x \in \mathcal{C}$, so it should be $< c$. We have a contradiction. This shows that the original assumption is false, so $\forall x, y > 0, \exists n \in \mathbb{N} : nx > y$

3.3.2 Consequences

This property can be used to show that $\lim_{n\to\infty} \frac{1}{n} = 0$. If we consider the definition of limits, the statement is equivalent to saying that $\forall \epsilon \in \mathbb{R} > 0, \exists N \in \mathbb{N}, \forall n > N, n \in \mathbb{R}, \frac{1}{n} < \epsilon$. If we rearrange the term, we get that we need to show $1 < \epsilon N$ for any ϵ. This is true because of the Archimedean Property. $\forall n > N$, since $\epsilon > 0$, $1 < \epsilon N < \epsilon n$. Therefore we know the limit is truely 0.

3.4 Sunrise Lemma

The Sunrise Lemma states for any sequence $(a_n)_n^\infty = 1$ in $\mathbb{R}$, $\exists$ monotone subsequence $(a_{n_k})_{k=1}^\infty$, where $(n_k)_{k=1}^\infty$ is a strictly increasing sequence in $\mathbb{N}$ and $n_k \geq k$ for all $k \in \mathbb{N}$

Vistas are points in a sequence, a_n, where $N \in \mathbb{N}$, such that $a_N > a_n$ for all $n > N$

This means that for any sequence, there exists a subset of points within, such that within that sequence, the sequence is monotone

3.4.1 Well-Ordering Property

For any set $A \subseteq \mathbb{N}, A \neq \emptyset, min(A)$ exists

3.4.2 Proof

<u>Case I:</u> The set V of vistas, is infinite, such that $n_1 = min(v)$ and $n_k = min(V \cap n_{k-1}^\infty$, where $k \geq 2$, then a_{n_k} is strictly decreasing
<u>Case II:</u>

$$n_1 = \begin{cases} 1 & if v = \emptyset \\ 1 + max(v) & if v \neq \emptyset \end{cases}$$

$n_k = choice\{n > n_{k-1} | a_n \geq a_{n_{k-1}}$, thus $n_k \neq \emptyset$ because V is finite, thus a_{n_k} is increasing

3.5 Bolzano-Weierstrass Theorem

Every bounded sequence in $\mathbb{R}$ has at least one convergent subsequence

3.5.1 Proof

Let $(a_n)_{n=1}^{\infty}$ be a bounded sequence. For any monotone sequence $(a_{n_k})_{k=1}^{\infty}$, then $(a_{n_k})_{k=1}^{\infty}$ is both bounded and monotone, so it converges.

4 Extreme Value Theorem

For some function $f : [a, b] \to \mathbb{R}$ (for the codomain, not the range) that is continuous ($f(x_0) = \lim_{x \to x_o} f(x)$ for any x $\in (a, b)$, $f(a) = \lim_{x \to a^+} f(x)$, $f(b) = \lim_{x \to b^-} f(x)$), then $\exists c, d \in [a, b]$ such that $f(c) \leq f(x) \leq f(d)$.

4.1 Proof

4.1.1 $\exists M > 0$ such that $f(x) \leq< M \forall x \in [a, b]$ (f(x) is bounded above)

Lets assume that f is not bounded above, such that for any $n \in \mathbb{N}, \exists x_n \in [a, b]$ such that $f(x_n) > n$. The sequence $(x_n)_{n=1}^{\infty}$ is bounded between a and b, such that it has a convergent subsequence $(x_{n_k})_{k=1}^{\infty}$ converging to some point t, when $\lim_{k \to \infty}(x_{n_k})$, by the below claim.

<u>Claim:</u> $t \in [a, b]$. Let $t > b$, then $\epsilon > 0$ such that $[a, b] \cap (t - \epsilon, t + \epsilon) = \emptyset$. But $\exists N$ such that $x_N \in (t - \epsilon, t + \epsilon)$ and $x_N \in [a, b]$, which is a contradiction.

By the previous claim, $lim_{k \to \infty} f(x_{n_k}) = f(t)$ by the assumed continuity of f. Thus, $\exists$ K such that for all $k \geq K, f(x_{n_k}) < f(t) + 1 \in \mathbb{R}$. On the other hand, $f(x_{n_k}) > n_k > f(t) + 1$ when k is sufficiently large. Thus, there is a contradiction, and f(x) is bounded from above. This can be reversed to show it is bounded from below, as well.

4.1.2 The function reaches the supremum and infimum

We now know $R := f([a, b]) = \{f(x) | x \in [a, b]\}$ is bounded, such that S := sup(R) and I := inf(R). By the definition of infimum and supremum, $\exists (y_n)_{n=1}^{\infty}$ such that $y_n \in R \forall n$, and $\lim_{n \to \infty} y_n = S$. Since $y_n \in R, \exists x_n [a, b]$ such that $f(x_n) = y_n$. Now $(x_n)_{n=1}^{\infty}$ is bounded between a and b so it has a convergent subsequence, $(x_{n_k})_{k=1}^{\infty}$, converging to $t \in [a, b]$. Also, by continuity of f, $\lim_{k \to \infty} f(x_{n_k}) = f(t)$. Thus, f(t) = S. This can be reversed to apply to the infimum.

5 Higher Dimensional Mathematics

5.1 Higher Dimensions

5.1.1 Cartesian Product

The Cartesian Plane represents the set $\mathbb{R}^2 := \{(x, y) + | + x, y \in (R)\}$. This is known as the **Cartesian Product** of $\mathbb{R}$ with itself. The Cartesian Product of two sets $\mathcal{S}$ and $\mathcal{T}$, $\mathcal{S} \times \mathcal{T} := \{(s, t) + | + s \in \mathcal{S}, t \in \mathcal{T}\}$.

5.1.2 Shapes

- <u>Open-ball</u>: $B_r(P) := \{X + | + dist(X, P) < r\}$
- <u>Closed-ball</u>: $\bar{B}_r := \{X + | + dist(X, P) \leq r\}$
- <u>Sphere</u>: $S_r := \{X + | + dist(X, P) = r\}$

5.1.3 Boundary

Given $D \subseteq \mathbb{R}^2, D \neq \emptyset$. We say $(a, b) \in \partial D$, i.e. (a, b) is on a <u>boundry point</u> of D if $\forall \epsilon > 0$, there are points $(x, y) \in D$ and $(u, v) \in D^c$ ($D^c := \mathbb{R}^2 \setminus D$) such that $dist(x, y+; +a, b) < \epsilon$ and $dist(u, v+; +a, b) < \epsilon$. Since the definition is symmetrical, $\partial D^c = \partial D$.

5.1.4 Interior

Let $D \subseteq \mathbb{R}^2$. We say (a, b) is an <u>interior point</u> of D if $\exists + r > 0 : B_r(a, b) \subseteq D$. The set of all interior points is called the <u>interior</u> of D and is written as int D.

5.1.5 Exterior

Let $D \subseteq \mathbb{R}^2$. We say (a, b) is an <u>exterior point</u> for D if it is an interior point of D^c. $\exists r > 0 : B_r(a, b) \subseteq D^c$. The set of all exterior points for D is the <u>exterior</u> of D, written as ext D.

<u>Thm</u>: For any $D \subseteq \mathbb{R}^2$, $\mathbb{R}^2 = \text{int } D \cup \partial D \cup \text{ext } D$. And $\text{int } D \cap \partial D = \emptyset$, $\text{int } D \cap \text{ext } D = \emptyset$, $\partial D \cap \text{ext } D = \emptyset$.

<u>Thm</u>: $\text{int } D = \text{ext } D^c$ and $\text{ext } D = \text{int } D^c$

<u>Thm</u>: $\text{ext } D \subseteq D^c$

5.1.6 Closure

The **closure** of $D \subseteq \mathbb{R}^2$:

$$\bar{D} := D \cup \partial D$$

<u>Thm</u>: $\partial \bar{D} = \partial D$, $\text{int } \bar{D} = \text{int } D$, and $\text{ext } \bar{D} = \text{ext } D$

Thm: int $D \subseteq D \subseteq \bar{D}$.

5.1.7 Ordered Pair

The ordered pair (a, b) can be thought of as a set, but a set is inheritly unordered. To express the order, we can do the following: $(a, b) = \{\{a\}, \{a, b\}\}$. Now we know that a is the first element because it appears in both subsets. We can then expand this into higher dimensions like the following: $(a, b, c) = ((a, b), c)$. Note that this means that $((a, b), c) \neq (a, (b, c))$. But this does not matter to us.

Fundamental Postulate of Ordered Pairs: $(a_1, a_2, a_3, \ldots, a_n) = (b_1, b_2, b_3, \ldots, b_n)$ if and only if $a_1 = b_1 \wedge a_2 = b_2 \wedge \cdots \wedge a_n = b_n$.

5.1.8 Vector and Points

Vectors are quantities of directionality and length, its location does not matter. Points are just positions in space. In higher dimensions with no ambiance space (flat space surrounding the surface, i.e. the shortest distance in the ambiant space is the straight line), we define a vector as all the lines with the same direction at a certain point.

However, the nice thing about $\mathbb{R}^d$ is that there is always ambiance space, so we will not make any notational distinction between a point and a vector.

The length of a vector in d space is defined as:

$$||\vec{a}|| := dist(\vec{0}, \vec{a}) = \sqrt{\sum_{i=1}^{d} a_i^2}$$

5.1.9 Space and Lines

$\mathbb{R} = \{(x_1, x_2, ..., x_d) | x_1, x_2, ...x_d \in \mathbb{R}\}$ A line, l, can be defined such that l $= \{(a_1 + tb_1, a_2 + tb_2, ..., a_d, tb_d) | t \in \mathbb{R}\}$

5.2 Dot Product

5.2.1 Definition and Perpendicularity

The dot product arises naturally through the idea of geometric distance, such that if $a \neq \emptyset, b \neq \emptyset$, then $a \perp bi$ iff $dist(a; b)^2 = dist(\emptyset, a)^2 + dist(\emptyset, b)^2$, where $a = (a_1, a_2), b = (b_1, b_2)$. Thus, by expanding out, $a \perp b$ iff $a_1b_1 + a_2b_2 = 0$ where $a \neq \emptyset, b \neq \emptyset$. In addition, orthoganal refers to both perpendicular vectors and where $a = \emptyset$ and/or $b = \emptyset$, so that no vector can be perpendicular to itself.

By extension, in $\mathbb{R}$, $a\dot{b} = a_1b_1 + a_2b_2 + a_3b_3 + ... + a_db_d$, such that it forms a scalar, rather than a vector.

5.2.2 Properties

The dot product is:

- Commutative
- Distributive over Vector Sums

5.3 Functions in Higher Dimensions

5.3.1 Domain, Range, and One-to-One

The domain is a subset of $\mathbb{R}^{\ltimes}$. Let the function of f(x, y) be an ordered pair within some curve, such that $(x, y) \in D$. Thus, the range of f, G $= x, y, f(x, y)|(x, y) \in D \subseteq R^{n+1}$. Functions are defined as one-to-one if for f(x, y), $(x, y), (u, v) \in D, f(x, y) = f(u, v)$ iff (if and only if) x = u $\wedge$ (and) y = v (such that for every z value, there is only one point that will create it.

5.3.2 Bolzano-Weirstrauss in Higher Dimensions

<u>Theorem:</u> A bounded sequence $\in \mathbb{R}$ has a convergent subsequence.

<u>Lemma:</u> If $(x_n)_{n=1}^{\infty}$ converges to $x \in \mathbb{R}$, then every subsequence $(x_{n_k})_{k=1}^{\infty}$ also converges to x, following from the definition of convergence and limits ($\forall \epsilon > 0, \exists N, \forall n \geq N : |x_n - x| < \epsilon$, thus $\exists K, \forall k \geq K : n_k \geq N$, since $n_k \to \infty$ as $k \to \infty$, such that $|x_{n_k} - x| < \epsilon$).

<u>Proof for $d = 2$:</u> Let $P_n = (x_n, y_n)$. Consider $(x_n)_{n=1}^{\infty} \in \mathbb{R}$. Since $\forall x_n, -M \leq x_n \leq M, (x_n)_{n=1}^{\infty}$ is bounded. For some $(x_{n_k})_{k=1}^{\infty}$, converging to some $x \in \mathbb{R}$. Consider $(y_{n_k})_{k=1}^{\infty}$ is bounded by the same rationale, thus $(y_{n_{k_j}})_{j=1}^{\infty}$ converges to some $y \in \mathbb{R}$. Since $(x_{n_{k_j}})_{j=1}^{\infty}$ is a subsequence of a converging sequence, it converges to the same value, x. Thus, $P_{n_{k_j}} = (x_{n_{k_j}}, y_{n_{k_j}}) \to (x, y) = P$.

5.3.3 Cauchy Sequence

A cauchy sequence in $\mathbb{R}$ is a sequence $(x_n)_{n=1}^{\infty}$ such that: $\forall \epsilon > 0, \exists N, \forall n, m \geq N : |x_n - x_m| < \epsilon$. This defines a sequence where as $n \to \infty$, the distance between values of points on the sequence decreases.

5.3.4 Convergence as Cauchy

Note that any convergent sequence is cauchy, because as terms get together to a limit, they also go very closely together. <u>Proof:</u> Let $(x_n)_{n=1}^{\infty}$ be convergent, with limit $x \in \mathbb{R}$. Then, by definition, $\forall \epsilon > 0, \exists N_\epsilon, \forall n \geq N_\epsilon : |x_n - x| < \epsilon$. Note that we can replace ϵ with $\frac{\epsilon}{2}$, all we have to change is the cutoff point from N_ϵ to $N_{\frac{\epsilon}{2}}$. Now if we take two subscripts $n, m \geq N_{\frac{\epsilon}{2}} \longrightarrow |x_n - x_m| = |(x_n - x) + (x - x_m)| \leq |x_n - x| + |x_m - x|$ because of the Triangle Inequality for Absolute Values. However, note that $|x_n - x| \leq \frac{\epsilon}{2}$ and $|x_m - x| \leq \frac{\epsilon}{2}$. Therefore, $|x_n - x_m| \leq |x_n - x| + |x_m - x| \leq \epsilon$.

5.3.5 Cauchy's Convergence Theorem

In $\mathbb{R}$, every cauchy sequence converges to a limit in $\mathbb{R}$.

<u>Lemma #1</u>: Every cauchy sequence is bounded.

Let us take $\epsilon = 1$, then the definition of "cauchiness" becomes:

$$\exists + N_1, \forall n, m \geq N_1 : |x_n - x_m| < 1$$

Let $M := \max\{|x_1|, |x_2|, \ldots, |x_{N_1-1}|, |x_{N_1}|+1\}$. We claim that $|x_n| \leq M$, for all $n \geq 1$. This is true because when $n \in \{1, 2, \ldots, N_1 - 1\}$, the statement is true by definition of M. When $n \geq N_1$, we know that $|x_n| \leq |x_{N_1}|+1 \leq M$ because we can let $m = N_1$, then by the definition of "cauchiness," we know that $|x_n - x_{N_1}| < 1$.

Now we see that M is a bound on the sequence for all $n \geq 1$. Therefore the sequence is bounded.

<u>Lemma #2</u>: If a subsequence of a cauchy sequence converges to $x \in \mathbb{R}$, the whole sequence must converge to x.

Say $(x_n)_{n=1}^{\infty}$ is cauchy, and $(x_{n_k})_{k=1}^{\infty}$ converges to x. For any arbitrary $\epsilon > 0$, we try to find N such that $\forall n \geq N : |x_n - x| < \epsilon$. If we prove the existence of N for all ϵ, we will have proven that the original sequence converges.

We know that $\forall \epsilon > 0, \exists K_\epsilon, \forall k \geq K_\epsilon : |x_{n_k} - x| < \epsilon$. We add and subtract x_{n_k} and group terms, and use the Triangle Inequality: $|x_n - x| = |(x_n - x_{n_k}) + (x_{n_k} - x)| \leq |x_n - x_{n_k}| + |x_{n_k} - x|$. Note that $|x_{n_k} - x| < \epsilon$ provided $k \geq K_\epsilon$ from the convergent subsequence condition. We also know that $|x_n - x_{n_k}| < \epsilon$ provided that $n, n_k \geq N_\epsilon$, which we call the "cauchy cutoff." This is true from the "cauchiness" condition.

We know that $k \to \infty$ implies $n_k \to \infty$. This means eventually $n_k > N_\epsilon$ provided that $k > L_{N_\epsilon}$. Now let $k = \max\{L_{N_\epsilon}, K_\epsilon\}$ and $n \geq N_\epsilon$, which implies $|x_n - x_{n_k}| < \epsilon$ and $|x_{n_k} - x| < \epsilon$. Now we know: $|x_n - x| \leq 2\epsilon$ provided $n \geq N_\epsilon$. Therefore the cauchy sequence converges to x.

With these two lemmas, the theorem becomes very easy to prove:

Because of Lemma #1 and the Bolzano-Weierstrass Theorem, we know that for all cauchy sequences, there is a bounded subsequence that converges to some value x. Then by Lemma #2, we know that the entire cauchy sequence converges to x as well, therefore the sequence converges.

5.3.6 Cauchy Sequences in Higher Dimensions

$(P_n)_{n=1}^{\infty}$ is cauchy if $\forall \epsilon > 0, \exists N_\epsilon, \forall n, m \geq N_\epsilon : dist(P_n, P_m) < \epsilon$. This is easy to prove due to the coordinate nature of $\mathbb{R}^d$.

5.4 Metric Topology in $\mathbb{R}^{\ltimes}$

5.4.1 Continuity

Let $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^2$, $D \neq \emptyset$. Let $(a, b) \in D$. We say that f is <u>continuous</u> at (a, b) if:

$$f(a, b) = \lim_{\substack{(x,y)\to(a,b) \\ (x,y)\in D}} f(x, y)$$

Or in other terms:

$$\forall \epsilon > 0, \exists + \delta > 0, \forall (x, y) \in D : dist(x, y+; +a, b) < \delta \to |f(x, y) - f(a, b)| < \epsilon$$

$f : D \to \mathbb{R}, D \subseteq \mathbb{R}^{\nvDash}, D \neq \emptyset$ is continuous if f is continuous at $(a, b) \forall (a, b) \in D$.

5.4.2 Directional Limits

Let $D \subseteq \mathbb{R} and a \in D \cup \partial D$ (the boundry, both already included in D, and not), $\lim_{x\to a^+} f(x) = L$ means $\forall \epsilon > 0, \exists \delta(\epsilon) > 0 : \forall x \in D \cap (a, \infty), |x - a| < \delta(\epsilon) \Rightarrow |f(x) - L| < \epsilon$. The limit only exists if both directions equal the same value.

5.4.3 Limits in Higher Dimensions

The same theory can be applied to higher dimensions, such that if two arbitrary approaches are not the same, it doesn't exist, but if several approaches yield the same result, the definition of a limit is used. Polar coordinate substitutions can be used to give format to directions of approach.

Due to difficulty defining approaching through lines, it is said that $(x, y) \to (a, b) iff dist(x, y : a, b) \to 0$.

Let $f : D \to \mathbb{R}, D \subseteq \mathbb{R}^{\nvDash}, D \neq \emptyset, and (a, b) \in D \cup \delta D$. Then, $L = \lim_{(x,y)\in D\to(a,b)} f(x) iff \forall \epsilon > 0, \exists \delta > 0, \forall (x, y) \in D : dist(x, y : a, b) < \delta \to |f(x, y) - L| < \epsilon$. As a corollary, when (a, b) is on the boundry, the approach can only be from the domain.

5.5 Properties of Domain

For the extreme value theorem to apply to a domain, the set must be compact, such that it must be bounded and closed over limits. On the $\mathbb{R}$ dimension, this applies to all closed intervals, as well as the empty set, though functions except the empty set cannot accept it as a domain.

5.5.1 Bounded

$If D \subseteq (R^2)$ is bounded if $\exists M > 0 : D \subseteq [-M, M]x[-M, M]$. Thus, a sequence is considered bounded if the set of all values within the sequence is bounded.

5.5.2 Closed

The term closed is used to apply to sets which are closed under limits. On $mathbbR$, if $x_n \in [a,b] for \forall n \in mathbbN, and x_n \to x \in \mathbb{R}, then x \in [a,b]$.

$D \subseteq \mathbb{R}^{\nvDash} is closed if for any points (x_n, y_n) \in D(for all n \in \mathbb{N}, if (x_n, y_n) \to (x,y) \in \mathbb{R}^{\nvDash}, then (x,y) \in D.(x_n, y_n) \to (a,b) as n \to \infty$ means $d_n = \sqrt[2]{(x_n - a)^2 + (y_n - b)^2} \to 0 as n \to \infty$.

This applies the definition of limits to sequences, such that $(x_n, y_n) \to (a,b) if dist(x_n, y_n : a, b) \to 0 as n \to \infty$.

Thus, $D \subseteq \mathbb{R}^{\nvDash}$ is closed if for any sequence $((x_n, y_n))_{n=1}^{\infty}$ in D that converges, the limit poiint (a, b) of the sequence also lies in D.

5.5.3 Open Set Theorem

$D \subseteq \mathbb{R}^{\nvDash}$ is open if $\forall (a,b) \in D, \exists r > 0$: the disk of radius r, $B_r(a,b) \subseteq D$, such that D = Interior of D

It follows that for any open set, the complement set within the space is a closed set.

<u>Proof</u>: Assume D is open, we prove D^c is closed. Choose any convergent sequence $((x_n, y_n))_{n=1}^{\infty}$, converging to (a,b), where $(x_n, y_n) \in D^c$ for all $n \geq 1$. We prove this by contradiction. Assume that $(a,b) \in D$. Since D is open, $\exists + r > 0 : B_r(a,b) \subseteq D$. $\exists + N, \forall n \geq \mathbb{N}, (x_n, y_n) \in B_r(a,b)$ since $(x_n, y_n) \to (a,b)$. But we assumed that $(x_n, y_n) \in D^c$, and $(x_n, y_n) \in D$. But $D \cap D^c = \emptyset$. Therefore D^c is closed under taking limits.

For the other direction, we can pick some point within D, then assume there is no ball, such that any ball contains some ball not in D, even as radius $\to 0$, creating a sequence of points converging on the point, P, a contradiction.

<u>Theorem</u>: An open-ball $B_r(\vec{p}) = \{\vec{x} \in \mathbb{R} + | + dist(\vec{x}, \vec{p}) < r\}$, $r > 0$, is an open set.

<u>Proof:</u> $\forall \vec{q} \in B_r(\vec{p})$, $B_\epsilon(\vec{q}) \subseteq B_r(\vec{p})$, $\epsilon > 0$.

$d = dist(\vec{p}, \vec{q}) < r$. THerefore, $r - d > 0$. Take $\epsilon = \frac{1}{2}(r - d) > 0$. Let $\vec{x} = B_\epsilon(\vec{q})$, show $\vec{x} \in B_r(\vec{p})$.

$dist(\vec{x}, \vec{q}) < \epsilon$

$dist(\vec{x}, \vec{p}) \leq dist(\vec{x}, \vec{q}) + dist(\vec{q}, \vec{p}) < \epsilon + d = d + \frac{1}{2}r - \frac{1}{2}d = \frac{1}{2}r + \frac{1}{2}d < \frac{1}{2} \cdot 2 \cdot r = r$

5.6 Extreme-Value Theorem

Let $f : D \to \mathbb{R}$ be continuous, where D $\subseteq \mathbb{R}$ is compact. Then $\exists P, Q \in D$, which do not need to be unique, such that $\forall X \in D : f(P) \leq f(x) \leq f(Q)$.

<u>Proof</u>: Assume that f is not bounded above, such that $\forall n \geq 1, f(P_n) > n$, where $P_n \in D$. For some subsequence $P_{n_k} \in D$ converging to P by the Balzano-Weirstrauss, by closure of D, $P \in D$. This is a contradiction since $f(P_{n_k}) \to \infty$ and $\to f(P)$, such that it must be bounded from above.

Thus, $\exists M = sup_{x \in D} f(x)$. We can find $P_n \in D$, with $f(P_n) \to M$. For some convergent subsequence $P_{n_k} \in D, P_{n_k} \to Q$.

6 Distance Functions

In axiomatic geometry, certain axioms including the definition of euclidean distance are taken as assumed. In actuality, standard distance functions must qualify under several non-geometric requirements, of which only the Euclidean distance qualifies.

Distance functions must be translation-iniant, or for any translation of two points, the distance must remain the same, such that $T_{h,k} : (x, y) \mapsto (mapsto)(x+h, y+k), thendist(x+h, y+k; x'+h, y'+k) = dist(x, y; x', y')$.

Thus, $dist(x, y; \tilde{(x)}, \tilde{(y)}) = f(|x-\tilde{(x)}|, |y-\tilde{(y)}|)$, where f the distance function defined on $[0, \infty)x[0, \infty)$. As a result, it must be symmetrical, such that $dist(x, y; \tilde{(x)}, \tilde{(y)}) = dist(\tilde{(x)}, \tilde{(y)}; x, y)$.

In addition, it must have basic reflection symmetry (isotropy), such that $dist(x, y; 0, 0) = dist(y, x; 0, 0)$. Thus, f(u, v) = f(v, u) for any $u \geq 0, v \geq 0$. It must also have the self-distance of (0, 0), such that dist(0,0; 0,0) = 0.

It must recreate the standard distance function on each axis, such that $dist(x, 0; \tilde{(x)}, 0) = |x - \tilde{(x)}|, dist(0, y; 0, \tilde{(y)}) = |y - \tilde{(y)}|. Therefore, f(u, 0) = u, f(0, v) = v \forall u \geq 0, v \geq 0$.

As a result, it must have asymptotic flatness, where if a line is drawn to (x, y), where y is fixed, such that $dist(0, 0; 0, y) = v_0, \overline{whiledist(0, 0; x, 0)} = u. Then, \lim_{u\to\infty} f(u, v_0)/u = 1$. This also applies in the opposite direction, where x is fixed.

It must be continuous in its variables, such that with a minute movement of a point, the distance changes minutely as well.

The set of isometries (distance preserving one-to-one functions) that fix the origin onto itself (f(0) = 0) is an infinite set.

Based on these requirements, an ansatz (educated guess, verified by later results) is made, such that $f(u, v) = F(G(u) + G(v)), where F : [0, \infty) \to \mathbb{R} and G : [0, \infty) \to \mathbb{R}$. The use of G(u) and G(v) is needed to assure symmetry. The use of addition is mandated by symmetry, using addition rather than another symmetrical operation simply due to ease of calculations.

Theorem: $\exists only one suitable pair F, G; G(x) = x^2, F(x) = \sqrt{x}, that fits all requirements. If G(x) = x^n, F(x) = \sqrt[n]{x}$, it would have all required properties except infinite set of isometries.

Property: Iff $dist(p; q) = 0$, then $p = q$, where p and q are asome vector $\in \mathbb{R}$

6.1 Euclidean Distance

The distance function in one space between two points a and b is simply $|a - b|$. However, we can also write it in the following way: $\sqrt{(a - b)^2}$

In $\mathbb{R}^2$, the distance function is:

$$dist(x, y + +; +a, b) := \sqrt{(x - a)^2 + (y - b)^2}$$

And in $\mathbb{R}^3$, the distance function is:

$$dist(x, y, z+; +a, b, c) := \sqrt{(x - a)^2 + (y - b)^2 + (c - z)^2}$$

The generalized form of Euclidean Distance in N space is:

$$dist(\vec{p}, \vec{q}) = \sqrt{\sum_{j=1}^{N} (p_j - q_j)^2}$$

This is known as the <u>Euclidean Distance</u>. We use this specific definition of distance because this is preserved under an infinite set of rigid or isometric motions, such as rotation, reflection, translation, etc.

6.2 Geometric Distance

Basic Transformations:

- $T_h : x \mapsto x + h$
- $R : x \mapsto -x$

Properties:

1. $dist(\vec{p}, \vec{q}) = dist(\vec{q}, \vec{p})$
2. $dist(\vec{p}, \vec{q}) \geq 0$
3. $dist(\vec{p}, \vec{q}) = 0 \leftrightarrow \vec{p} = \vec{q}$

6.3 Basic Distance Bounds Lemma

$\forall \vec{p}, \vec{q} \in \mathbb{R}^d$, and $\forall j \in \{1, 2, 3, \ldots, d\}$:

$$|p_j - q_j| \leq dist(\vec{p}, \vec{q}) \leq \sqrt{d} \max_{1 \leq k \leq d} |p_k - q_k|$$

Proof Note that $(p_j - q_j)^2 \leq \sum_{k=1}^{d} (p_k - q_k)^2$ is trivial, because you can only add positive nunder when you add squares. Now let's take the square root, and we get

$$\sqrt{(p_j - q_j)^2} = |p_j - q_j| \leq \sqrt{\sum_{k=1}^{d} (p_k - q_k)^2} = dist(\vec{p}, \vec{q})$$

To prove the other inequality, it is trivial as well. We can just factor out the length of the vector d and multiply that with the maximum value of the distance vector. Then we get:

$$dist(\vec{p}, \vec{q}) = \sqrt{\sum_{k=1}^{d} (p_k - q_k)^2} \leq \sqrt{d \max_{1 \leq k \leq d} (p_k - q_k)^2} = \sqrt{d} \max_{1 \leq k \leq d} |p_k - q_k|$$

Cor: Componentwise Nature of Convergence

Let $(\vec{p}_n)_{n=1}^{\infty}$ be a sequence in $\mathbb{R}^d$, and let $\vec{p} \in \mathbb{R}^d$. Then $\vec{p}_n \to \vec{p}$ if and only if $p_{n|j} \to p_j$ ($\vec{p} = (p_1, p_2, p_3, \ldots, p_d)$ and $\vec{p}_n = (p_{n|1}, p_{n|2}, \ldots, p_{n|d})$). Otherwise known as convergence of points can be reduced to conversion of dimensions.

This follows directly from the inequality, because if the total distance goes to 0, then $|p_j - q_j|$ goes to 0. Therefore if the points converge, the corresponding coordinates must converge.

To prove the converse, we prove using the other side of the distance bounds. If all d coordinates are going to 0, then if we take the maximum, that would be going to 0. (the maximum of a sequence is less than the sum of the sequence, but if every term of the sum is going to 0, then the sum is going to 0, then the maximum is going to 0). Therefore the distance must also be going to 0. Thus the two points converges.

6.4 Other Distance

Of course, there are other distance formulas, like the <u>Minkowski Distance</u>

$$((x-a)^p + (y-b)^p)^{\frac{1}{p}} + + + +(p > 1)$$

This is another distance formula, but under this, only reflection preserves distance.

6.5 Distances Between Sets

We define the distance between a point $\vec{p}$ and a set D as:

$$dist(\vec{p}; D) = \inf_{\vec{d} \in D} dist(\vec{d}; \vec{p})$$

We also define the distance between two sets D_1 and D_2 as:

$$dist(D_1; D_2) = \inf_{\substack{\vec{p} \in D_1 \\ \vec{q} \in D_2}} dist(\vec{p}; \vec{q})$$

7 Inequalities

7.1 Level of Operations

Powers/root $\to$ Multiplation/division $\to$ addition/subtraction $\to$ succession/pretrition

7.2 AM-GM

$$\mu = [x_1, x_2, \ldots, x_n] \text{ and } x_1, x_2, x_3, \ldots, x_n \geq 0$$

"Multiset" $\mu = \{(x, n), (y, m), \ldots\}$

We define the arithmetic mean of a multiset as:

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

And the geometric mean as:

$$G(\mu) = \sqrt[n]{x_1 x_2 x_3 \dots x_n}$$

7.2.1 AM-GM Inequality

$A(\mu) \geq G(\mu)$, with equality iff all elements of μ are the same.

7.2.2 Proof

This is done by mathematical induction. Base case is $n = 2$, then $\mu = [x, y]$. Then $A(\mu) = \frac{x+y}{2}, G(\mu) = \sqrt{xy}$

We know by the trivial inequality that $(\sqrt{x} + \sqrt{y})^2 \geq 0$, with equality case happening iff $x = y$. Then we get:

$$\begin{aligned} x - 2\sqrt{xy} + y &\geq 0 \\ \frac{x+y}{2} &\geq \sqrt{xy} \\ A(\mu) &\geq G(\mu) \end{aligned}$$

Now we induce on n, we seek to prove that case n implies case $2n$.

$\mu = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]$

Then we know that

$$\begin{aligned} A(\mu) &= \frac{A(\mu_x) + A(\mu_y)}{2} \\ &\geq \frac{G(\mu_x) + G(\mu_y)}{2} \\ &\geq \sqrt{G(\mu_x)G(\mu_y)} \\ &= G(\mu) \end{aligned}$$

Note that in all inequalities used, the equality case is always when all x_n and y_n are the same element, therefore the equality case holds in all cases where the length of the list is 2^n.

Now we prove that case n implies $n - 1$

$\mu = [x_1, x_2, x_3, \dots, x_{n-1}]$

Note that we can construct $\mu' = [x_1, x_2, x_3, \dots, x_{n-1}, A(\mu)]$

Note that $A(\mu') = A(\mu)$, and since the AM-GM inequality is true for μ' by the assumption, we know

$$\begin{aligned} A(\mu') = A(\mu) &\geq \sqrt[n]{x_1 x_2 x_3 \dots x_{n-1} A(\mu)} \\ &\geq \sqrt[n]{G(\mu)^{n-1} A(\mu)} \\ &\geq \sqrt[n]{G(\mu)^{n-1} G(\mu)} \\ &\geq G(\mu) \end{aligned}$$

7.3 Young's Inequality

7.3.1 Hölder Conjugate

q is said to be the Hölder Conjugate of p:

$$q := p^* := \frac{p}{p-1}$$

Note that $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$

7.3.2 Young's Inequality

$a, b \geq 0$; $p > 1$; $q = p^*$, then Young's Inequality states that:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

With equality case iff $a^p = b^q$

7.3.3 Proof

We first proof Young's Inequality assuming that $p, q \in \mathbb{Q}$. We can rewrite $p = \frac{n+m}{n}$ and $q = \frac{n+m}{m}$ for some $n, m \in \mathbb{N}$. Now Young's Inequality turns into:

$$ab \leq \frac{na^{\frac{n+m}{n}}}{n+m} + \frac{mb^{\frac{m+n}{m}}}{n+m}$$

If we let $x = a^{\frac{1}{n}}$ and $y = b^{\frac{1}{m}}$. Then the inequality turns into:

$$ab = x^n y^m =\leq \frac{nx^{n+m} + my^{n+m}}{n+m}$$

And that is true by weighted AM-GM, with equality iff $x = y$, which equals to $a^{\frac{1}{n}} = b^{\frac{1}{m}}$, which equals to

We can prove the inequality for irrational by taking limits, because $\forall n \geq n_0 : f(n) \leq g(n)$, and the limits of both $f(x)$ and $g(x)$ as $n \to \infty$ exists and are finite, then we know that $\lim_{n\to\infty} f(n) \leq \lim_{n\to\infty} g(n)$. When p and q are irrational, we construct $\{p_n\}$ and $\{q_n\}$ as two sequences of rationals that approaches p and q, the left hand side of Young's Inequality is unaffected by the limit, and by what we've just said about limits, we know that:

$$ab \leq \lim_{n\to\infty} \frac{a^{p_n}}{p_n} + \frac{b^{q_n}}{q_n} = \frac{a^p}{p} + \frac{b^q}{q}$$

7.4 Hölder's Inequality

7.4.1 p-norm

In $\mathbb{R}^2$: let $||(a,b)||_p = (|a^p| + |b^p|)^{1/p}$ for any $p \in \mathbb{R} > 1$, this is known as the p-norm of a vector. Note that $||(a,b)||_2 = ||(a,b)|| = \sqrt{a^2+b^2}$

7.4.2 Hölder's Inequality

$\forall (a,b),(c,d) \in \mathbb{R}^{\not=}, p > 1, q = p^* = \frac{p}{p-1}$:

$$0 \leq |ac| + |bd| \leq ||(a,b)||_p ||(c,d)||_q$$

Equality happens iff $(\frac{a}{s})^p = (\frac{c}{t})^q$ and $(\frac{b}{s})^p = (\frac{d}{t})^q$ where $s = ||(a,b)||_p$ and $t = ||(c,d)||_q$, such that $||(a,b)||_p ||(c,d)||_q = (|a|^p + |b|^p)^{1/p}(|c|^q + |d|^q)^{1/q}$

7.4.3 Proof

We know that the absolute value of the product is equal to the product of the absolute value. If we apply Young's Inequality to $|\frac{a}{s}||\frac{c}{t}|$ and $|\frac{b}{s}||\frac{d}{t}|$, we get:

$$|\frac{a}{s}||\frac{c}{t}| \leq \frac{|a|^p}{p|s|^p} + \frac{|c|^q}{q|t|^q}$$

$$|\frac{b}{s}||\frac{d}{t}| \leq \frac{|b|^p}{p|s|^p} + \frac{|d|^q}{q|t|^q}$$

Now we add:

$$\frac{1}{st}(|ac| + |bd|) \leq \frac{|a|^p + |b|^p}{p|s|^p} + \frac{|c|^q + |d|^q}{q|t|^q}$$

Note that $|a|^p + |b|^p = |s|^p$ and $|c|^q + |d|^q$, so everything cancels

$$\frac{1}{st}(|ac| + |bd|) \leq \frac{1}{p} + \frac{1}{q}$$

But we know that $q = p^*$, therefore, $\frac{1}{p} + \frac{1}{q} = 1$, and we get:

$$|ac| + |bd| \leq st = ||(a,b)||_p \cdot ||(c,d)||_q$$

The equality case occurs at basically the same way as Young's Inequality's equality case.

7.5 Cauchy-Schwarz Inequality

This is a special case of Hölder's Inequality, where $p = q = 2$. (This is very important, 2 is the *only* value that is its own conjugate, this is why Euclidean distance is so special)

If we plug in 2 for p and q and use the Triangle Inequality: $|ac+bd| \leq |ac|+|bd| \leq \sqrt{a^2+b^2}\sqrt{c^2+d^2}$, or $|v\dot{u}| \leq ||u||||v||$ with equality iff $ab \geq 0$.

Then, by Young's inequality, $|\frac{ac}{st}| = |\frac{a}{s}||\frac{c}{t}| \leq \frac{|a|^p}{p|s|^p} + \frac{|c|^q}{q|t^q}$, and the same is true for bd.

It follows that $\frac{1}{st}(|ac| + |bd|) \leq \frac{|a|^p+|b|^p}{p|s|^p} + \frac{|c|^q+|d|^q}{q|t|^q} = \frac{1}{p} + \frac{1}{q} = 1$, or $|ac| + |bd| \leq st$.

7.6 Triangle Inequality

$$dist(\vec{p}, \vec{q}) + dist(\vec{q}, \vec{r}) \geq dist(\vec{p}, \vec{r})$$

This can be generalized by mathematical induction to $dist(\vec{p_0}, \vec{q_n}) \leq \sum_{j=1}^{n} dist(\vec{p}_{j-1}, \vec{p}_j)$ (Otherwise known that the shortest distance between two points is the straight line, or the **Generalized Triangle Inequality** or the "Broken Line Inequality")

This can be thought of algebraically, such that $|a + b| \leq |a| + |b|$ with equality iff $ab \geq 0$

7.7 Reverse Triangle Inequality

From the Triangle Inequality we know:

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

Let $\vec{z} = \vec{x} + \vec{y}$, then we subtract, we get:

$$|| + \vec{z} + || \leq || + \vec{x} + || + || + \vec{z} - \vec{x} + ||$$

$$|| + \vec{z} + || - || + \vec{x} + || \leq || + \vec{z} - \vec{x} + ||$$

Because $\vec{x}$ and $\vec{z}$ are just variables, we can switch them, and we get:

$$|| + \vec{x} + || - || + \vec{z} + || \leq || + \vec{x} - \vec{z} + || = || + \vec{z} - \vec{x} + ||$$

Since the right hand side is greater than both of the above qualities, we can just say it's greater than the absolute value of the difference. Hence we get the Reverse Triangle Inequality:

$$|| + \vec{z} - \vec{x} + || \geq ||| + \vec{z} + || - || + \vec{x} + |||$$

7.8 Minkowski's Inequality

Mikowski's states that $||u+v||_p \leq ||u||_p + ||v||_p$, with equality if v = tu or u = tv for some $t \geq 0$.

This can be thought of as the triangle inequality for the p-norm, rather than the ordinary norm.

7.8.1 Rational Power Proof

The rational power of some nunder, m, exists if there is some sequence, qn, where $n \to \infty, qn \to$ the rational nunder, only true if for any sequence which does this, the limit is equal. This is proven by for any two sequences, qn and rn, $m^{qn}/m^{rn} = m^{qn-rn}$, such that as $n \to \infty$, it equals 1.

7.8.2 Proof

The calculation works in any dimension, for simplicity's sake, let's work in $\mathbb{R}^2$, let $\vec{u} = (a, b)$ and $\vec{v} = (c, d)$

$$||+\vec{u}+\vec{v}+||_p^p = |+a+c+|^p + |+b+d+|^p = |+a+c+||+a+c+|^{p-1} + |+b+d+||+b+d+|^{p-1}$$

Now we factor and use the Triangle Inequality for Absolute Value:

$$\leq (|a|+|c|)|+a+c+|^{p-1} + (|b|+|d|)|+b+d+|^{p-1}$$

Now we rearrange the terms:

$$= (|a||a+c|^{p-1} + |b||b+d|^{p-1}) + (|c||a+c|^{p-1} + |d||b+d|^{p-1})$$

Now we apply Hölder's Inequality, we get:

$$\leq (|a|^p + |b|^p)^{\frac{1}{p}}(|a+c|^{(p-1)q} + |b+d|^{(p-1)q})^{\frac{1}{q}} + (|c|^p + |d|^p)^{\frac{1}{p}}(\dots)$$

Note that $q = p^*$, therefore $(p-1)q = p$:

$$= (||+\vec{u}+||_p + ||+\vec{v}+||_p)||+\vec{u}+\vec{v}+||_p^{\frac{p}{q}}$$

Since $q = p^*$, we know that $\frac{p}{q} = p-1$, and if we bring the inequality to the original left hand side:

$$\leq (||+\vec{u}+||_p + ||+\vec{v}+||_p) + ||+\vec{u}+\vec{v}+||_p^{p-1}$$

Now we divide:

$$||+\vec{u}+\vec{v}+||_p \leq ||+\vec{u}+||_p + ||+\vec{v}+||_p$$

Now let's consider the equality cases. If one of the vectors is 0, then the inequality is trivially true.

If neither vectors are the 0 vector, we see the equality cases of all the inequalities used to prove Minkowski's. First we used the triangle inequality, which only has equality when $ac \geq 0$ and $bd \geq 0$. Next we applied Hölder's, which has equality case when both coordinates are proportional. Therefore, the two vectors must be positive multiples of one another.

8 Heine-Borel Theorem and Domain Properties

8.1 Heine-Borel Theorem

Let K be a compact set in $\mathbb{R}$, and let $U_\lambda | \lambda \in \Lambda$ be a family of open sets in $\mathbb{R}$, which covers K, such that $K \subseteq \cup_{x \in \Lambda} U_\lambda$. Then $\exists \lambda_1, \lambda_2, \lambda_3 ...$ such that$K \subseteq U_{\lambda_1} \cup U_{\lambda_2} \cup ... \cup U_{\lambda_n}$.

8.1.1 Proof

Since K is bounded, it can be fully contained within some rectangle (R), which can then be split into 4 congruent parts, $R_{i_1}(R_1, R_2, R_3, R_4)$. Suppose one quadrant cannot be covered by any finite collection of U_λ. By extension, if that quadrant is divided further, one of the subquadrants ($R_{i_1 i_2}$) cannot be covered. This can continue for countable infinity divisions (able to be counted with an infinite amount of integer). Since $R_{i_1 i_2 ... i_n} \neq \emptyset$, let P_n be any point in $K \cap R_{i_1 i_2 ... i_n}$, such that there is a bounded sequence of points in K. Thus it must have a convergent subsequence such that $P_{n_k} \to P, P \in K$. Thus, $P \in U_{\lambda^*}$ for some $\lambda^* \in \Lambda$. Since U_{λ^*} is open, $\exists r > 0,$ such that$B_r(P) \subseteq U_{\lambda^*}$. Diameter/diagonal length (diam) of $R_{i_1 i_2 ... i_n} = \frac{diam(R)}{2^n} < r$ as $n \to \infty$. Thus, $R_{i_1 i_2 ... i_n} \subseteq B_r(P)$ as $n_k \to \infty$. $dist(P, P_{n_k}) < \frac{r}{2}$ and $dist(P_{n_k}, x) \leq diam(R_{i_1 i_2 ... i_n})$ and by the triangle inequality, $diam(R_{i_1 i_2 ... i_n}) < \frac{r}{2}$. Thus, there is a contradiction, and it must be covered by a finite number of sets.

8.2 Uniform Continuity

Let $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^d$. We say f is uniformally continuous on D if:

$$\forall \epsilon > 0, \exists + \delta > 0 : \forall \vec{x}, \vec{y} \in D, || + \vec{x} - \vec{y} + || < \delta \to |f(\vec{x}) - f(\vec{y})| < \epsilon$$

The difference between this and regular continuity is that the δ in regular continuity is defined by both ϵ and the specific point we are considering. Uniform continuity, however, the value δ is independent to the point you chose within the domain and is just dependent on ϵ.

For example, consider $y = \tan x$ where $D = (-\frac{\pi}{2}, \frac{\pi}{2})$. the value required for δ for a fixed ϵ gets smaller and smaller as x approaches both endpoints. This function is continuous but not uniformally so. If it were uniformally continuous, that δ value would NOT change.

<u>Theorem</u>: If f is uniformally continuous on D, then f is continuous for every point in D.

8.3 Uniform Continuity Theorem

<u>Theorem:</u> If $f : K \to \mathbb{R}$, where $K \subseteq \mathbb{R}^d$ is compact, and f is continuous at each $\vec{x} \in K$. Then f is uniformally continuous on K.

<u>Proof:</u> Fix $\epsilon > 0$. For each $\vec{x} \in K$, let $u_{\vec{x}}$ be an open ball, centered at $\vec{x}$ such that for any $\vec{y} \in K \cap 2u_{\vec{x}}$ [$2u_{\vec{x}}$ is an open ball centered at $\vec{x}$ with twice the radius of $u_{\vec{x}}$], $|f(\vec{x}) - f(\vec{y})| < \frac{\epsilon}{2}$. Because the function is continuous at $\vec{x}$, there is a radius 2δ around $\vec{x}$ such that $|f(\vec{x}) - f(\vec{y})| < \frac{\epsilon}{2}$. Now we see that every $\vec{x} \in K$ is covered by at least one such open ball, namely $u_{\vec{x}}$. The collection $\{u_{\vec{x}} + | + \vec{x} \in K\}$ is an open covering of K. By Heine-Borel, we can select a finite set of points

$\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ such that K is covered by $u_{\vec{x}_1} \cup u_{\vec{x}_2} \cup \cdots \cup u_{\vec{x}_n}$. Take $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_n\} > 0$ where δ_j is the radius of $u_{\vec{x}_j}$ for $j = 1, 2, 3, \ldots, n$.

Let $\vec{p}, \vec{q} \in K$ such that $|| + \vec{p} - \vec{q} + || < \delta$. We need to prove that $|f(\vec{p}) - f(\vec{q})| < \epsilon$. $\vec{p} \in K \to \exists + j : \vec{p} \in u_{\vec{x}_j} \to || + \vec{p} - \vec{x}_j + || < \delta_j$. But we know that $|| + \vec{p} - \vec{q} + || < \delta_j$. Now by the Triangle Inequality, we get:

$$|| + \vec{q} - \vec{x}_j + || \leq || + \vec{q} - \vec{p} + || + || + \vec{p} - \vec{x}_j + || \leq 2\delta_{\vec{x}_j}$$

This means that $\vec{q} \in 2u_{\vec{x}_j}$. By the way we picked our δ, we know that $|f(\vec{q}) - f(\vec{x}_j)| < \frac{\epsilon}{2}$. Similarly, we also have $|f(\vec{p}) - f(\vec{x}_j)| < \frac{\epsilon}{2}$. Now if we apply the triangle inequality again, we get:

$$|f(\vec{p}) - f(\vec{q})| = |f(\vec{p}) - f(\vec{x}_j) + f(\vec{x}_j) - f(\vec{q})| \leq |f(\vec{p}) - f(\vec{x}_j)| + |f(\vec{q}) - f(\vec{x}_j)| < \epsilon$$

Therefore, f is uniformally continuous over K.

8.4 Connected

8.4.1 Definition

$D \subseteq \mathbb{R}^d$ is said to be connected if $\forall \vec{p}, \vec{q} \in D$, $\exists$ a continuous function $\vec{\gamma} : [a, b] \to \mathbb{R}^d$ that $\vec{\gamma}(a) = \vec{p}$, $\vec{\gamma}(b) = \vec{q}$, and $\forall t \in [a, b], \vec{\gamma}(t) \in D$

8.4.2 Theorem

Let $D \subseteq \mathbb{R}^d$ be connected, $D \neq \emptyset$. Suppose $D = A \cup B$, where A and B are open. Such that $A \cap B = \emptyset$. Then $A = \emptyset$ or $B = \emptyset$.

<u>Proof</u>:

Assume that D can be broken up into two open, non-empty sets A and B such that $D = A \cup B$. Because $A \neq \emptyset$, we can pick $\vec{p} \in A$ and similarly we can pick $\vec{q} \in B$. Clearly, $\vec{p}, \vec{q} \in D$. Since D is connected, $\exists$ continuous $\vec{\gamma} : [a, b] \to D$ such that $\vec{\gamma}(a) = \vec{p}$, $\vec{\gamma}(b) = \vec{q}$. Let $S = \{t \in [a, b] + | + \vec{\gamma}(t) \in A\}$. We know that $a \in S$, so $S \neq \emptyset$. Note that $S \leq b$, and since it has a bound, it has a supremum. Let $t_0 = \sup S$, note that $a \leq t_0 \leq b$. Since $\vec{\gamma}(t_0) \in D$, it is either in A or in B (since $A \cap B = \emptyset$). Suppose $\vec{\gamma}(t_0) \in A$. Since A is open, we can draw an open ball (B_ϵ) around $\vec{\gamma}(t_0)$. Now consider $\vec{\gamma}(t_0 + \delta)$ for some small positive δ. If δ is small enough, $\delta \in B_\epsilon \in A$. This is a contradiction, as then it means that $t_0 + \delta \in S$. This is a contradiction, as then $t_0 \neq \sup S$. A similar contradiction can be made for the case that $\vec{\gamma}(t_0) \in B$. Therefore the original assumption was false.

9 Intermediate Value Theorem

9.1 Theorem in $\mathbb{R}$

If f: [a, b] $\to \mathbb{R}$ is continuous, and $f(a) \neq f(b)$, $\forall y \in (f(a), f(b))$, then $\exists c \in (a, b) : f(c) = y$.

9.1.1 Proof

Let there exist y, without loss of generality, such that $f(a) < y < f(b)$. Take the set $S = \{x \in [a,b] | f(x) \leq y\}$. Since $a \in S$, $S \neq \emptyset$, we also know that $S \leq b$. Now take $c = \sup S$, $a \leq c \leq b$ ($c \in [a,b]$). There are three possible cases, $f(c)$ is either $> y, < y$, or $= y$.

Consider the case in which $f(c) < y$. If this is the case, then we can chose an arbitrarily small $\delta > 0$ such that $f(c + delta) < y$. However, then $c + \delta \in S$. But this causes a contradiction because $c = \sup S$ and there should not be any element of S that's larger than c. Therefore, this case is impossible.

Now consider the case in which $f(c) > y$. Take some arbitrarily small $u \in [0, \delta], \delta > 0$ such that $f(c - u) > y$. However, then $c - \delta$ is therefore an upper bound of the set S, we once again reach the same contradiction.

Therefore, since c exists, $f(c) = y$.

9.2 Theorem in $\mathbb{R}^{\ltimes}$

If $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}^d$ is connected and continuous. $\forall \vec{p}, \vec{q} \in D$ such that $f(\vec{p}) \neq f(\vec{q})$ and if $y \in \mathbb{R}$ is between $f(\vec{p})$ and $\vec{q}$, then $\exists \vec{r} \in D$ such that $f(\vec{r}) = y$.

9.2.1 Proof

Without loss of generality, let us assume $f(\vec{p}) < f(\vec{q})$.

Since D is connected, there is some path $\vec{\gamma}(t) = (x(t), y(t)) \in D$, $a \leq t \leq b$ such that $\vec{\gamma}(a) = \vec{p}$ and $\vec{\gamma}(b) = \vec{q}$ and is continuous over $[a, b]$. Now we construct $g(t) = f(\vec{r}(t)) = (f \cdot \vec{r})(t) \in mathbbR$. Because $g(t)$ is a composition of continuous functions, $g(t)$ is also continuous. $g(a) = f(\vec{\gamma}(a)) = f(\vec{p})$ and $g(b) = f(\vec{\gamma}(b)) = f(\vec{q})$. Now we see that $g(a) < y < g(b)$. Therefore, by the IVT for single-variable functions, $\exists + t_0$ such that $g(t_0) = y$. Now we plug t_0 into $\vec{\gamma}$, $\vec{r} = \vec{\gamma}(t_0) \in D$. Then $f(\vec{r}) = f(\vec{\gamma}(t_0)) = g(t_0) = y$.

10 Vector Valued Functions

10.1 Definition

$f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^d$ is known as <u>real-valued</u> functions.

$\vec{f} : D \to \mathbb{R}^e$, $D \subseteq \mathbb{R}^d$ is known as <u>vector-valued/point-valued</u> functions.

$\vec{f}(\vec{p}) = (f_1\vec{p}, f_2\vec{p}, \ldots, f_e(\vec{p}))$ where $f_1, f_2, \ldots, f_e$ are real-valued and are called the <u>component functions</u> of f.

10.2 Continuity

$f : \mathbb{R}^d \to \mathbb{R}^e$ is continuous at $\vec{a} = (a_1, a_2, a_3, \ldots, a_d) \in D$ iff $\forall \epsilon, \exists + \delta > 0 : \forall \vec{x} \in D$:

$$|| + \vec{x} - \vec{a} + ||_d < \delta \to || + f(\vec{x}) - f(\vec{a}) + ||_e < \epsilon$$

10.2.1 Component-wise Nature of Continuity

$f : D \to \mathbb{R}^e$ is continuous at a point $\vec{a} \in D$ iff $f_1, f_2, \ldots, f_e$ are all continuous at $\vec{a}$.

<u>Proof</u>:

First fix an ϵ, then by the basic distances bound lemma, we get a bunch of inequalities:

$$\begin{aligned}
|f_1(\vec{x}) - f_1(\vec{p})| &< \frac{\epsilon}{\sqrt{e}}, \forall \vec{x} \in D \cap B_{\delta_1}(\vec{p}) \\
|f_2(\vec{x}) - f_2(\vec{p})| &< \frac{\epsilon}{\sqrt{e}}, \forall \vec{x} \in D \cap B_{\delta_2}(\vec{p}) \\
&\ldots\ \ldots\ \ \ldots\ \ldots \\
|f_e(\vec{x}) - f_e(\vec{p})| &< \frac{\epsilon}{\sqrt{e}}, \forall \vec{x} \in D \cap B_{\delta_e}(\vec{p})
\end{aligned}$$

Then let $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_e\} > 0$, which must exist and satisfy all the distance inequalities, specifically $\max_{1 \le j \le e}(|f_j(\vec{x}) - f_j(\vec{p})|)$. Then again, by the basic distance bounds lemma, we know that $|| + \vec{f}(\vec{x}) - \vec{f}(\vec{p}) + ||_e \le \max_{1 \le j \le e}(|f_j(\vec{x}) - f_j(\vec{p})|)$. Therefore we get that for any fixed ϵ, we can find a δ such that $|| + \vec{x} - \vec{p} + ||_d < \delta \to || + f(\vec{x}) - f(\vec{p}) + ||_e < \epsilon$

10.2.2 Composition of Continuous Functions

If $\vec{f} : D \to E$, where $D \subseteq \mathbb{R}^d$, $E \subseteq \mathbb{R}^e$ and $\vec{g} : E \to \mathbb{R}^k$ are both continuous on their respective domains. Then $\vec{h} = \vec{g} \circ \vec{f}$ is continuous on D

<u>Proof</u>:

To prove this, fix $\epsilon > 0$. There's a $\eta > 0$ such that $\forall \vec{y} \in E \cap B_\eta(\vec{f}(\vec{p}))$, because we know that $\vec{g}$ is continuous at $\vec{f}(\vec{p})$, we get:

$$|| + \vec{g}(\vec{y}) - \vec{g}(\vec{f}(\vec{p})) + || < \epsilon$$

To guarantee that $\vec{y} = \vec{f}(\vec{x})$ lies within η units of $\vec{f}(\vec{p})$ i.e. $|| + \vec{f}(\vec{x}) - \vec{f}(\vec{p}) + || < \eta$, we can take $\vec{x} \in D \cap B_\delta(\vec{p})$ where $\delta > 0$ corresponding to η [using the continuity of $\vec{f}$ at $\vec{p} \in D$].

Now, as long as $\vec{x} \in D \cap B_\delta(\vec{p})$, we have $\vec{f}(\vec{x}) \in E \cap B_\eta(\vec{f}(\vec{p}))$. Thus, if we take $\vec{y} = \vec{f}(\vec{p})$, we get: $|| + \vec{g}(\vec{f}(\vec{x})) - \vec{g}(\vec{f}(\vec{p})) + || < \epsilon$, or $|| + \vec{h}(\vec{x}) - \vec{h}(\vec{p}) + || < \epsilon$. So $\vec{h}$ is continuous at $\vec{p}$.

10.3 Compactness Theorem

10.3.1 Theorem

Let $\vec{f} : D \to \mathbb{R}^e$ be a contnuous function, where $D \subseteq \mathbb{R}^d$ is compact. Then its *range* $\vec{f}(D) := \{f(\vec{p}) + | + \vec{p} \in D\}$ is also compact. In other words: compactness is preserved under continuous mappings. Note that this is the generalization of the Extreme Value Theorem.

10.3.2 Proof

To prove this, write $R := f(D)$. We need to show that R is closed and bounded in $\mathbb{R}^e$. Boundedness is easy. Since each component function f_j of f is real valued, by EVT each component function f_j has an absolute bound M_j, so that $| + f_j(\vec{p}) + | \leq M_j$ for all $\vec{p} \in D$. Take $M := \max\{M_1, M_2, \ldots, M_e\}$. Then for all $\vec{p} \in D$, we have

$$|| + \vec{f}(\vec{p}) + || = \sqrt{f_1(\vec{p})^2 + \cdots + f_e(\vec{p})^2} \leq \sqrt{e \times M^2} = M\sqrt{e}$$

This says that the range R lies within the closed ball of radius $M\sqrt{e}$ centered at $\vec{0}$ in $\mathbb{R}^e$. It therefore certainly lies within some closed cube centered at $\vec{0}$, and hence is bounded.

To prove closedness, let $(\vec{y}_n)_{n=1}^{\infty}$ be a convergent sequence in $\mathbb{R}^e$ with limit $\vec{y}$, such that $\vec{y}_n \in R$ for each $n \geq 1$. We need to prove that $\vec{y} \in R$. Since R is the range of f, we must have $\vec{y} = \vec{f}(\vec{x_n})$, where $\vec{x_n} \in D$. Because D is bounded, we can pick a convergent subsequence $\vec{x}_{n_k} \to \vec{x}$. But because D is closed, we know that $\vec{x} \in D$. Because $\vec{f}$ is continuous, we get that $\vec{y}_{n_k} = \vec{f}(\vec{x}_{n_k}) \to \vec{f}(\vec{x})$. However, we know that $\vec{y}_{n_k} \to \vec{y}$. But since each sequence converges to one point, we know that $\vec{y} = \vec{f}(\vec{x})$, where $\vec{x} \in D$. Therefore, $\vec{y} \in R$.

Therefore, R is closed and bounded.

10.4 Connectedness Theorem

10.4.1 Theorem

$\vec{f} : D \to \mathbb{R}^e$, $D \in \mathbb{R}^d$ is continuous on D. If D is connected, $E = \vec{f}(D)$ (the range of the domain), is also connected. Note that this is the generalization of the Intermediate Value Theorem.

10.4.2 Proof

$\forall \vec{u}, \vec{v} \in E$, we can find two points $\vec{p}, \vec{q}$ such that $\vec{f}(\vec{p}) = \vec{u}$ and $\vec{f}(\vec{q}) = \vec{v}$. Now because D is connected, $\exists + \vec{\gamma} : [a, b] \to \mathbb{R}^d$, $\vec{\gamma}([a, b]) \subseteq D$. Now we consider $\vec{\delta} = \vec{f}(\vec{\gamma}(t))$, $\vec{\delta} : [a, b] \to \mathbb{R}^e$. Since it's a composition of continuous functions, $\vec{\delta}$ is also continuous. And $\forall t \in [a, b]$, $\vec{\delta}(t) = \vec{f}(\vec{\gamma}(t)) \in E$. We also know that $\vec{\delta}(a) = \vec{f}(\vec{\gamma}(a)) = \vec{f}(\vec{p}) = \vec{u}$, and $\vec{\delta}(b) = \vec{f}(\vec{\gamma}(b)) = \vec{f}(\vec{q}) = \vec{v}$. Therefore, $\forall \vec{u}, \vec{v} \in E$, there is a continuous path that connects $\vec{u}$ to $\vec{v}$ and stays within E. Therefore, E is connected.

11 Sequences of Functions

11.1 Infinity Norm

For $f : [a, b] \to \mathbb{R}$ that is bounded, we say:

$$||f||_\infty := ||f||_D := \sup_{x \in [a,b]} |f(x)|$$

11.2 Pointwise Convergence:

<u>Definition:</u> For $f_n : [a, b] \to \mathbb{R}$ for $n \geq 1$, and assume that they are all bounded on $[a, b]$. ($\exists + M_n > 0 : |f_n(x)| \leq M_n$ for all $x \in [a, b]$), and $f : [a, b] \to \mathbb{R}$, f is bounded ($\exists + M > 0 : |f(x)| \leq M$ for all $x \in [a, b]$).

If $\forall x \in [a, b] : \lim_{n\to\infty} f_n(x) = f(x)$, then we say that $f_n \overset{p}{\to} f(x)$ (f_n converges "pointwise" to f).

<u>Theorem:</u> If $\vec{f_n} \overset{u(\text{uniform})}{\to} \vec{f} : \mathbb{R} \to \mathbb{R}$, then $\vec{f_n} \overset{p(\text{pointwise})}{\to} \vec{f}$ on D: $\vec{f_n}(\vec{x}) \to \vec{f}(\vec{x})$ for every point $\vec{x} \in D$ (for any converging sequences, any point will also converge similarly, such that uniform convergence implies pointwise convergence, though the converse is untrue)

<u>Proof:</u> $\vec{f_n} \overset{u}{\to} \vec{f}$ on D (uniform convergence): $||\vec{f_n} - \vec{f}||_D = sup_{\vec{x}\in D}||\vec{f_n}(\vec{x}) - \vec{f}(\vec{x})|| \to 0$ as $n \to \infty$. $Then 0 \leq ||\vec{f_n}(\vec{x}) - \vec{f}(\vec{x})|| \leq ||\vec{f_n} - \vec{f}|| \to 0$.

Then, by the squeeze theorem, for a uniform sequences, $f_n(x) \to f(x)$ on D, giving pointwise convergence.

11.3 Uniform Convergence

11.3.1 In $\mathbb{R}$

$f_n : [a, b] \to \mathbb{R}$ for $n \geq 1$, and assume that they are all bounded on $[a, b]$. ($\exists + M_n > 0 : |f_n(x)| \leq M_n$ for all $x \in [a, b]$).

$f : [a, b] \to \mathbb{R}$, f is bounded ($\exists + M > 0 : |f(x)| \leq M$ for all $x \in [a, b]$).

We then claim that $f_n \overset{u}{\to} f$ as $n \to \infty$ if $\forall \epsilon$, $\exists + N_\epsilon$ such that $\forall n \geq N_\epsilon : || + f_n - f + ||_\infty < \epsilon$. In other words, this forces that the greatest vertical difference between the two functions will be arbitrarily small after N_ϵ. This forces the two functions to be "close" as a whole.

11.3.2 In $\mathbb{R}^\ltimes$

Let $D \in \mathbb{R}^d$ be a non-empty set, and let $\vec{f} : D \to \mathbb{R}^e$ and $\vec{f_n} : D \to \mathbb{R}^e$ for $n \geq 1$. We say that $\vec{f_n} - \vec{f}$ on D if:

$$|| + f_n - f + ||_D \to 0 \text{ as } n \to \infty$$

11.4 Uniform Convergence Theorem

11.4.1 Theorem

If $f_n \xrightarrow{u} f$ on $[a, b]$, where each f_n is continuous on $[a, b]$. Then f is also continuous on $[a, b]$

11.4.2 Proof

Pick any $\vec{p} \in D$, and fix $\epsilon > 0$. We need to find δ such that $\forall \vec{x} \in D$ with $||\vec{x} - \vec{p}|| < \delta$, then $||\vec{f}(\vec{x}) - \vec{f}(\vec{p})|| < \epsilon$. We can apply the triangle inequality and we get:

$$||\vec{f}(\vec{p}) - \vec{f}(\vec{p})|| \leq ||\vec{f}(\vec{x})\vec{f_n}(\vec{x})|| + ||\vec{f_n}(\vec{x}) - \vec{f_n}(\vec{p})|| + ||\vec{f_n}(\vec{p}) - \vec{f}(\vec{p})||$$

For large enough n, we know that $||\vec{f_n} - \vec{f}||_D < \frac{\epsilon}{3}$ because $\vec{f_n} - \vec{f}$. Now we know that the first and third ter are bounded by $\frac{\epsilon}{3}$. The second term is bounded b $\frac{\epsilon}{3}$ because $\vec{f_n}$ is uniformally continuous.

Therefore, we know that $||\vec{f}(\vec{p}) - \vec{f}(\vec{p})|| \leq \epsilon$ for any δ we pick. $\therefore \vec{f}$ is continuous on D

12 Differentiation

12.1 Differentiable in $\mathbb{R}$

Let $f : [a, b] \to \mathbb{R}$, let $p \in (a, b) = [a, b]^o(int[a, b])$,, then f is <u>differentiable</u> at p if $\exists a \in \mathbb{R}$ such that $a = \lim h \to 0 \frac{f(p+h)-f(p)}{h}$, called the <u>derivative</u> at point p ($a = f'(p) = \frac{df}{dx}(p)$).

12.2 Differentiable in $\mathbb{R}^{\ltimes}$

This is generalized to high dimensions, such that for $f : D \subseteq \mathbb{R} \to \mathbb{R}$, and $\vec{p} \in D^o(\exists r > 0 : B_r(\vec{p} \subseteq D)$. Thus, it can be approached from any given direction, due to the ball existing in all directions.

$\vec{h} \to \vec{0}$ iff $h_d \to 0$, such that it can be substituted for the limit. Due to the lack of vector division though, the definition of the derivative has to be changed.

By the previous definition, $\lim h \to 0|\frac{f(p+h)-f(p)}{h} - a| = 0$, rewritten $\lim h \to 0\gamma(p, h) = 0$.

$$|h|\gamma(p, h) = |f(p + h) - f(p) - ah| < \epsilon|h| \text{ as } h \to 0(|h| < \delta).$$

$\exists a \in \mathbb{R}$, such that $\forall \epsilon > 0, \exists \delta > 0 : |h| < \delta$, then $|f(p + h) - f(p) - ah| < \epsilon|h|$, defining a as the derivative at p.

This is due to the idea that for some function, g(x), with linear approximation l(x), $\frac{|g(x)-l(x)|}{|x|} \to 0$ as $x \to 0$, such that the y distance decreases far faster than the x decay, called super-linear (suplinear) decay. Then, there must at most be 1 line that can superlinearly approximate g at x = 0. The superlinear decay curve is then the derivative function, where f(p + h) - f(p) = g(x) and ah = l(x).

Proof: Assume there are two function, l(x) = ax and m(x) = bx, then $\frac{|g(x)-ax|}{|x|} \to 0$ as $x \to 0$ and $\frac{|g(x)-bx|}{|x|} \to 0$ as $x \to 0$. Then $0 \leq |a-b|\frac{|ax-bx|}{|x|} = \frac{|(g(x)-bx)-(g(x)-ax|)}{|x|} \leq \frac{|g(x)-bx|}{|x|} + \frac{|g(x)-ax|}{|x|} \to 0$ as $x \to 0$, such that a = b. This proof can be done in each component for higher dimensions, using the dot product to remove all other components, by the properties of the dot product.

The definition can now be easily moved to higher dimensions, by superlinear decay, such that there is only one object in $\mathbb{R}^{\ltimes - \nVdash}$, a hyperplane, or the set of all vectors orthoganal to the non-zero normal vector, all anchored to a specific point, $\vec{p_0}$, such that superlinear decay takes place.

Definition: A hyperplane in $\mathbb{R}^{+\nVdash}$ is a set of the form, Q = $\{\vec{x} \in \mathbb{R}^{+\nVdash} | (\vec{x} - \vec{p_0} \cdot \vec{n} = 0)$, where $\vec{p_0}, \vec{n} \in \mathbb{R}^{+\nVdash}$ with $\vec{n} \neq \vec{0}$, where n is a normal of Q ($\vec{n} \perp Q$).

Thus to summerize differentiability for real valued functions, $\exists!$ (exists exactly one) or $\exists \vec{a} \in \mathbb{R}$ such that $\forall \epsilon > 0, \exists \delta > 0 : ||\vec{h}|| < \delta$, then $|f(\vec{p} + \vec{h}) - f(\vec{p}) - \vec{a}\vec{h}| < \epsilon ||\vec{h}||$. If the latter is true, f(x) is said to be differentiable at $\vec{p}$, since it is a unique value, denoted f'(p).

12.3 Gradient

This is then defined specifically such that $\vec{a} = (\vec{\nabla} f)(p)$, or the gradient of f at p.

Claim: The graph of $y = f(\vec{p}) + \vec{a} \cdot (\vec{x} - \vec{p})$ is a hyperplane in $\mathbb{R}^{+\nVdash}$, such that the gradient is a hyperplane.

Proof: Let $\vec{N}$ (the normal vector to the hyperplane$\in \mathbb{R}^{+\nVdash} = (-a_1, -a_2, ..., -a_d, 1), and \vec{P_0} \in \mathbb{R}^{+\nVdash} = (p_1, p_2, ..., p_d, f(\vec{p}))$and $\vec{X} \in \mathbb{R}^{+\nVdash} = (x_1, x_2, ..., x_d, y)$. Thus, $\vec{(X)} - \vec{P_0} = (x_1 - p_1, x_2 - p_2, ...x_d - p_d, y - f(\vec{p}))$, giving the equation of a hyperplane $(\vec{X} - \vec{P_0}) \cdot \vec{N} = 0$.

If $\vec{p} + \vec{h} = \vec{X}$, then $|f(\vec{X}) - [f(\vec{p}) + \vec{a} \cdot (\vec{X} - \vec{p})|| < \epsilon ||\vec{h}||$. This is equal to the difference between the point on the graph and the function approximation. This can also be seen to be easily equal to the equation of the hyperplane from above, such that the gradient is the hyperplane.

12.3.1 Gradient Representation

Since $\vec{N}$ is the normal vector to the plane at that point, equal to $(-\vec{\nabla} f(\vec{p}), 1)$, or the projection of the negation of the normal vector onto the domain, such that it is the vector of the direction and magnitude of the fastest increase of the function at $\vec{p}$.

12.3.2 Gradient Calculation

Take $\vec{h} = h\vec{e_j}$, where $h \to 0$ and the set of $\vec{e_j}$ is known as the *standard basis vectors* in $\mathbb{R}^d$:

$$\vec{e_j} = \begin{cases} \vec{e}_1 & = (1, 0, 0, \ldots, 0) \\ \vec{e}_2 & = (0, 1, 0, \ldots, 0) \\ \vdots & \vdots \quad \vdots \quad \vdots \quad \vdots \\ \vec{e}_d & = (0, 0, 0, \ldots, 1) \end{cases}$$

Note that because $||\vec{e}_j|| = 1$, $||\vec{h}|| = |h|||\vec{e}_j|| = |h|$. Now if we fix a $j \in \{1, 2, 3, \ldots, d\}$ and apply the definition of differentiability, the unique gradient ($\vec{a}$) must satisfy:

$$\forall \epsilon > 0, \exists \delta > 0 : \forall h \text{ with } |h| < \delta \implies |f(\vec{p} + h\vec{e}_j) - f(\vec{p}) - \vec{a} \cdot h\vec{e}_j| < \epsilon |h|$$

However, note that when we dot $\vec{a}$ with $\vec{e}_j$, the result is the jth component of $\vec{a}$, or a_j. Now if we divide through by $|h|$, we get:

$$\left| \frac{f(\vec{p} + h\vec{e}_j) - f(\vec{p})}{h} - a_j \right| < \epsilon$$

This says is that $\left| \frac{f(\vec{p}+h\vec{e}_j)-f(\vec{p})}{h} \right|$ approaches a_j indefinitely, therefore, we can rewrite the relationship as a limit statement:

$$\boxed{a_j = \lim_{h \to 0} \left| \frac{f(\vec{p} + h\vec{e}_j) - f(\vec{p})}{h} \right|}$$

We call this a_j as a **partial derivative** of $f(\vec{x})$ at j^{th} component, which can be written as $\partial_{x_j} f(\vec{p})$ or $\frac{\partial f}{\partial x_j}(\vec{p})$.

Note that:

$$\partial_{x_j} f(\vec{p}) = \frac{d}{dx_j} |_{x_j = p_j} f(p_1, p_2, \ldots, p_{j-1}, x_j, p_{j+1}, \ldots, p_d)$$

In other words, we can hold all other components of f constant and differentiate based on only one component, and plug in the value p_j after the differentiation. Now we know how to compute the gradient of f, it is simply the vector of all the partial derivatives:

$$\vec{\nabla} f = (\partial_{x_1} f(\vec{p}), \partial_{x_2} f(\vec{p}), \ldots, \partial_{x_d} f(\vec{p}))$$

12.4 Directional Derivative

Take $\vec{u} \in \mathbb{R}^d$, $\vec{u} \neq \vec{0}$ and take a point on the function f, $\vec{p}$, we define the <u>directional derivative</u> $\partial_{\vec{u}} f(\vec{p})$ as:

$$\begin{aligned} \partial_{\vec{u}} f(\vec{p}) &= \lim_{h \to 0} \frac{f(\vec{p} + h\vec{u}) - f(\vec{p})}{h} \\ &= \frac{d}{dt}|_{t=0} f(\vec{p} + t\vec{u}) \end{aligned}$$

This quality describes how fast the function f is changing at $\vec{p}$ in the direction of $\vec{u}$.

The latter definition gives the single variable function, such that g(t) $= \vec{p} + t\vec{u}$, where $\vec{u} \neq \vec{0}$, is called the uniform rectilinear motion curve, due to being a line with constant curve speed.

12.5 Vector Valued Directional Derivative

Given a vectored valued function $\vec{\gamma}(t)$:

$$\vec{\gamma}(t) = \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \\ \vdots \\ \gamma_d(t) \end{bmatrix}$$

We define the "speed" vector of $\vec{\gamma}$ as its derivative, which is defined as:

$$\frac{d\vec{\gamma}}{dt}(t) := \lim_{h \to 0} \frac{\vec{\gamma}(t+h) - \vec{\gamma}(t)}{h}$$

But because of the componentwise nature of limits, we can distribute the limit into each component of $\vec{\gamma}$, so we can rewrite the speed vector as:

$$\frac{d\vec{\gamma}}{dt}(t) = \begin{bmatrix} \gamma_1'(t) \\ \gamma_2'(t) \\ \vdots \\ \gamma_d'(t) \end{bmatrix}$$