Further Analysis

Kexing Ying

May 15, 2020

Contents

1 Euclidean Spaces

For $n \geq 1$, the n-dimensional *Euclidean space* denoted by $\mathbb{R}^n$, is the set of ordered n-tuples $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ for $x_i \in \mathbb{R}$. Recall that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$, we can use the usual vector space operations, i.e. vector addition and scalar multiplication. Furthermore, $\mathbb{R}^n$ forms a inner product space with the operation,

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (\mathbf{x}, \mathbf{y}) \mapsto \sum_{i=1}^n x_i y_i.$$

Thus, as a inner product space induces a normed vector space, we find a natural norm defined for $\mathbb{R}^n$ by,

$$\| \cdot \| : \mathbb{R}^n \to \mathbb{R} : \mathbf{x} \mapsto \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^n x_i^2}.$$

By manually checking, we find that this norm satisfy the norm axioms, i.e. it satisfy the *triangle inequality*, is *absolutely scalable*, and *positive definite* (In fact, we do not need the norm to be non-negative as it can deduced from the other axioms).

1.1 Preliminary Concepts in $\mathbb{R}^n$

Sequences in $\mathbb{R}^n$ can be defined similarly to that of $\mathbb{R}$, and we carry over all notations in all suitable places.

Definition 1.1 (Convergence in $\mathbb{R}^n$)**.** We say a sequence $(\mathbf{x}_i)_{i=1}^{\infty} \subseteq \mathbb{R}^n$ converges to $\mathbf{x} \in \mathbb{R}^n$ if and only if for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $i \geq N$, $\|\mathbf{x}_i - \mathbf{x}\| < \epsilon$.

Proposition 1. *A sequence $(\mathbf{x}_i)_{i=1}^{\infty} \in \mathbb{R}^n$ converges to $\mathbf{x} \in \mathbb{R}^n$ if and only if each component of $\mathbf{x}_i$ converges to the corresponding component of $\mathbf{x}$.*

In the first dimension, we've considered the topology of $\mathbb{R}$ including the examination of open and closed sets. We extend this idea for higher dimensions. The most basic examples we have of an open set (or closed set for that matter) in $\mathbb{R}$ are the open and closed intervals respectively. This is extended in $\mathbb{R}^n$ to be sets of the form

$$\prod_{i=1}^{n}(a_i, b_i) := \{\mathbf{x} \mid a_i < \mathbf{x}_i < b_i, \forall 1 \leq i \leq n\},$$

and similarly for closed intervals. However, while this is nice to look at, it is not very useful. So for this reason, we again will extend the notion of open and closed sets for $\mathbb{R}^n$.

Definition 1.2 (Open Ball)**.** Let $\mathbf{x} \in \mathbb{R}^n$ and $r \in \mathbb{R}^+$, we define the open ball of radius r about $\mathbf{x}$ as the set

$$B_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\| < r\}$$

.

Definition 1.3 (Open)**.** A set $U \subseteq \mathbb{R}^n$ is open in $\mathbb{R}^n$ if and only if for all $\mathbf{x} \in U$, there is some $r \in \mathbb{R}^+$ such that $B_r(\mathbf{x}) \subseteq U$.

Definition 1.4 (Closed)**.** A set $U \subseteq \mathbb{R}^n$ is closed if and only if its complement is open.

Straight away from the definition, we can see that every open ball is open (see here). Furthermore, we find the union and intersection of two open sets is open. In fact, the union and any collection of open sets is also open, however, this is not necessarily true for closed sets.

Definition 1.5 (Continuity at a Point)**.** Let $A \subseteq \mathbb{R}^n$ be an open set, and let $f : A \to \mathbb{R}^m$. We say f is continuous at $p \in A$ if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in A \cap B_\delta(p)$, $\|f(x) - f(p)\| < \epsilon$.

If the function f is continuous at every point of A, then we say f is continuous on A.

Definition 1.6. Let $A \in \mathbb{R}^n$ be open in $\mathbb{R}^n$ and let $f : A \to \mathbb{R}^m$. For $p \in A$, we say that the limit of f as $\mathbf{x}$ tends to $\mathbf{p}$ in A is equal to $\mathbf{q} \in \mathbb{R}^m$ if and only if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A \cap B_\delta(p), x \neq p$, $\|f(x) - \mathbf{q}\| < \epsilon$.

This is the same notion we used for continuity in the first dimension to say that f is continuous at p if and only if $\lim_{x\to p} f(x) = q$.

Proposition 2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then f is continuous if and only if for all open subsets U of $\mathbb{R}^m$, $f^{-1}(U)$ is open in $\mathbb{R}^n$.*

Proof. See here for the proof. ☐

1.2 Derivative of Maps in Euclidean Spaces

Let Ω be a open in $\mathbb{R}^n$, and $f : \Omega \to \mathbb{R}^m$ be a "nice behaving map". We poses the question on how we should define the notion of derivatives for this mapping at some point $p \in \Omega$. We recall that the derivative at a point p in the first dimension is defined to be

$$f'(p) = \lim_{x \to p} \frac{f(x) - f(p)}{x - p}.$$

While we see that this equation makes no sense if we simply generalise this equation to higher dimensions, we see the following result.

Lemma 1.1. Let $f : S \to \mathbb{R}$ where $S \subseteq \mathbb{R}$, then f is differentiable at some $p \in S$ if and only if there exists some $\lambda \in \mathbb{R}$ such that

$$\lim_{x \to p} \left| \frac{f(x) - A_\lambda(x)}{x - p} \right| = 0,$$

where $A_\lambda(x) = \lambda(x - p) + f(p)$.

Proof. Follows from algebraic manipulation. $\square$

With this, we can conclude that $f(x) - A_\lambda(x)$ tends to zero faster than $x - p$. We will generalise this result to higher dimensions.

We may rewrite $A_\lambda(x) = \lambda(x - p) + f(p) = \lambda x + (f(p) + \lambda p)$, so, we see that λ is the translation of a linear map λx, i.e. $A_\lambda = (x \mapsto x + (f(p) + \lambda p)) \circ (x \mapsto \lambda x)$. We can easily generalise such maps to higher dimensions and we call such maps *affine maps*.

Definition 1.7 (Differentiable Functions in $\mathbb{R}^n$)**.** Recall the definition of linear maps for general vector spaces which we will use in the context of Euclidean spaces. Let $L(\mathbb{R}^n; \mathbb{R}^m)$ denote the set of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, $\Omega \subseteq \mathbb{R}^n$ be open, and $f : \Omega \to \mathbb{R}^m$ be a function. Then we say f is differentiable at some point $p \in \Omega$ if and only if there exists some $\Lambda \in L(\mathbb{R}^n; \mathbb{R}^m)$, such that

$$\lim_{x \to p} \frac{\|f(x) - (\Lambda(x - p) + f(p))\|}{\|x - p\|} = 0.$$

If this is true, we write $Df(p) = \Lambda$ and call Λ the derivative of f at p.

Remark. *Some book refers to $Df(p)$ as the total derivative of f at p.*

It is often useful to express the derivative criterion as

$$\lim_{h \to 0} \frac{\|f(p + h) - f(p) - \Lambda(h)\|}{\|h\|} = 0.$$

Proposition 3. *Let $f_i : (a, b) \to \mathbb{R}$ be differentiable for all i, then the function, $f : (a, b) \to \mathbb{R}^m : x \mapsto (f_i(x))_{i=1}^m$ is differentiable for all $p \in (a, b)$.*

Proof. Let the Jacobian $\Lambda = \text{diag}(\lambda_i)$ where λ_i is the derivative of f_i at p. Then I claim, $Df(p) = \Lambda p$.

Consider

$$\lim_{h \to 0} \frac{\|f(p+h) - f(p) - \Lambda h\|}{\|h\|} = \lim_{h \to 0} \frac{\sqrt{\sum_{i=0}^m |f_i(p+h) - f_i(p) - \lambda_i h|}}{\|h\|}.$$

However, for all i, $\|h\| \geq |h_i|$, so

$$\lim_{h \to 0} \frac{\sqrt{\sum_{i=0}^m |f_i(p+h) - f_i(p) - \lambda_i h|}}{\|h\|} \leq \lim_{h \to 0} \sum_{i=0}^m \sqrt{\frac{|f_i(p+h) - f_i(p) - \lambda_i h|}{|h_i|}} = 0$$

□

A lot of results from the first dimension generalises easily to higher dimensions. Similar to that of the first dimension, the chain rule in general Euclidean spaces states,

Theorem 1. *Let $\Omega \subseteq \mathbb{R}^n$ and $\Omega' \subseteq \mathbb{R}^m$ be open sets with $g : \Omega \to \Omega'$ be differentiable at $p \in \Omega$ and $f : \Omega' \to \mathbb{R}^l$ be derivatives at $g(p)$. Then $h = f \circ g$ is differentiable at p with derivative*

$$Dh(p) = Df(g(p)) \circ Dg(p).$$

Proof. Similar to the proof of the Chain rule in the first dimension using algebraic manipulation. □

Omitted many examples here, check official lecture notes for these examples.

1.3 Directional Derivatives

Although the definition of derivative in dimension one and higher is similar, it is different in that we verify whether a linear map is the total derivative at a point instead of computing the limit of some equation. This is difficult as often times, it is not easy to guess what the derivative of a function is. Thus, it is useful to somehow identify candidate linear maps for the derivative.

Assume $\Omega \subseteq \mathbb{R}^n$ is open and $f : \Omega \to \mathbb{R}^m$ is a differentiable function at some $p \in \Omega$. Let $v \in \mathbb{R}^n$ be a unit vector. We would like to identify $Df(p)[v] \in \mathbb{R}^m$.

By the definition of derivatives in higher dimensions,

$$\lim_{h \to 0} \frac{\|f(p+h) - f(p) - \Lambda h\|}{\|h\|} = 0.$$

So, by letting $t \in \mathbb{R}$, we have

$$\begin{aligned} 0 &= \lim_{t \to 0} \frac{\|f(p+tv) - f(p) - \Lambda(tv)\|}{\|tv\|} \\ &= \lim_{t \to 0} \frac{\|f(p+tv) - f(p) - t\Lambda v\|}{|t|} \\ &= \lim_{t \to 0} \frac{\|f(p+tv) - f(p)\|}{|t|} - \Lambda v, \end{aligned}$$

So, $\lim_{t\to 0} \|f(p+tv) - f(p)\| / |t| = \Lambda v$. Thus, by finding the limits of the above equation for each basis vector $v \in B$, we find the Jacobian $[\Lambda]_B$.

Remark. *For notation, we denote the limit as* $\lim_{t\to 0} \|f(p+tv) - f(p)\| / |t|$ *as* $\partial f / \partial v \mid_p$ *and we call it the directional derivative of* f *in the direction of* v *at* p*. We will normally consider the directional derivatives in the direction of the standard basis and we call them the partial derivatives of* f *at* p*.*

Theorem 2. *Let* $\Omega \subseteq \mathbb{R}^n$ *be open, and* $f : \Omega \to \mathbb{R}^m : x \mapsto [f_i(x)]$ *for* $i \in \{1, \cdots m\}$ *is differentiable at some* $p \in \Omega$*. Then the Jacobian of* f *at* p *is* $[\partial f_i / \partial e_j \mid_p]_{i,j}$*.*

Proof. We recall that the Jacobian is simply the matrix form of the linear map that is the derivative. So, for all $v \in \mathbb{R}^n$, we have $Jv = Df(p)(v)$. As, $v \in \mathbb{R}^n$, it can be represented as a sum of the standard basis, that is there exists $v_i \in \mathbb{R}$, $v = \sum_{i=1}^n v_i e_i$, so, $Df(p)(v) = Df(p)(\sum_{i=1}^n v_i e_i) = \sum_{i=1}^n v_i Df(p)(e_i) = \sum_{i=1}^n v_i \partial f / \partial e_i \mid_p = [\sum_{i=1}^n v_i \partial f_j / \partial e_i \mid_p]_j = [\partial f_i / \partial e_j \mid_p]_{i,j} v$. (We used the fact that $[\partial f / \partial e_i]_j = \partial f_j / \partial e_i$.) ☐

Remark. *We note that the reverse is not true, that is the existence of partial derivatives does not imply differentiability. A counter example of this is* $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto \textit{if } x = y = 0, \textit{ then } 0, \textit{ else}, \frac{xy}{\sqrt{x^2+y^2}}$.

Theorem 3. *Let* $\Omega \subseteq \mathbb{R}^n$ *is open, and* $f : \Omega \to \mathbb{R}$ *be a function. Suppose that the partial derivatives of* f*,* $D_i f(x)$ *exists for all* $i = 1 \cdots n$ *exists at all* $x \in \Omega$*. Furthermore, if the map* $x \mapsto D_i f(x)$ *is continuous for all* i *at some point* p*. Then* f *is differentiable at* p*.*

1.4 Higher Derivatives

Similar to the first dimension, we would like to think about how to differentiate more than once.

Let $\Omega \subseteq \mathbb{R}^n$ is open, and $f : \Omega \to \mathbb{R}^m$ be differentiable everywhere on Ω. Then we may consider the map

$$Df : \Omega \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) : p \mapsto Df(p).$$

As there is a trivial isomorphism between $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ and the matrices of dimension $m \times n$, we can represent every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ as an element of $\mathbb{R}^{m\times n}$. Thus, Df can be represented as a map from Ω to $\mathbb{R}^{m\times n}$. So, we may ask if Df is continuous at some p and furthermore, we can ask if Df is differentiable at some $p \in \Omega$. If this is the case, we have the second derivative

$$DDf : \Omega \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{m\times n}).$$

Definition 1.8 (Second Derivative)**.** Let $\Omega \subseteq \mathbb{R}^n$ be open and $f : \Omega \to \mathbb{R}^m$ be differentiable everywhere on Ω with derivative $Df : \Omega \to \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. Then the second derivative of f at some $p \in \Omega$ is the linear map $\Lambda \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{m\times n})$ such that

$$\lim_{x\to p} \frac{\|Df(x) - Df(p) - \Lambda(x - p)\|}{\|x - p\|} = 0.$$

Thus, with this definition, we can easily extend the notion of derivatives any number of times to get the k-th derivative of a function. However, this is formally difficult and requires the notion of multilinear maps. Luckily, instead of working with this difficult definition whenever we would like to work with higher derivatives, we can instead look at whether the k-th derivative exists and whether or not it is continuous by theorem 3.

Now that we have established the notion of higher derivatives we would like to ask how higher partial derivatives interacts. That is, when does $D_i D_j f(p) = D_j D_i f(p)$?

Theorem 4 (Schwartz' Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open and $f : \Omega \to \mathbb{R}$ be differentiable at every $p \in \Omega$. Suppose that for some $i, j \in \{1, \cdots, n\}$ the second partial derivatives $D_i D_j f$ and $D_j D_i f$ exists and is continuous for all $p \in \Omega$, then*

$$D_i D_j f(p) = D_j D_i f(p)$$

for all $p \in \Omega$.

If $f : \Omega \to \mathbb{R}$, we call the matrix of second partial derivatives of f at some point p the **Hessian** of f at p and we write $\operatorname{Hess} f(p) = [D_i D_j f(p)]_{i,j=1,\cdots,n}$. Given that the hypothesis of Schwartz's theorem holds, we find that $[\operatorname{Hess} f(p)]_{i,j} = [\operatorname{Hess} f(p)]_{j,i}$ so the Hessian is symmetric.

1.5 Taylor's Theorem

We recall that the first derivative of a map $f : \Omega \to \mathbb{R}^m$ allows us to find an affine map at some point $p \in \Omega$ such that the error decreases faster than that of $\|x - p\|$. The existence of higher derivatives allows us to obtain better estimates with error decreasing even faster.

Let us first introduce some notations. We define a multi-index α an element of $\mathbb{N}^n$ and we write $|\alpha| = \sum \alpha_i$. Furthermore, given some function $f : \Omega \to \mathbb{R}^n$, we write

$$D^\alpha f := (D_1)^{\alpha_1} \cdots (D_n)^{\alpha_n} f,$$

and given $h \in \mathbb{R}^n$, we write

$$h^\alpha := h_1^{\alpha_1} \cdots h_n^{\alpha_n}.$$

Lastly, we write $\alpha! := \alpha_1! \cdots \alpha_n!$.

With that, we can state the Taylor's theorem.

Theorem 5 (Taylor's Theorem)**.** *Suppose $p \in \mathbb{R}^n$ and $f : B_r(p) \to \mathbb{R}$ is k-times differentiable on $B_r(p)$ for some $r > 0$ and $k \geq 1$. Then for any $h \in \mathbb{R}^n$ with $\|h\| < r$, we have,*

$$f(p+h) = \sum_{\substack{\alpha \in \mathbb{N}^n, \\ |\alpha| \leq k-1}} D^\alpha f(p) \frac{h^\alpha}{\alpha!} + R_k(p,h),$$

where there exists some $x \in \mathbb{R}^n$ with $0 < \|x - p\| < \|h\|$ such that

$$R_k(p,h) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k}} D^\alpha f(x) \frac{h^\alpha}{\alpha!}.$$

Proof. See one dimensional version from year one. (Essentially boils down to finding the Taylor expansion of the restriction of f on $\{p + th \mid t \in \mathbb{R}\} \cap B_r(p)$ which is isomorphic to a open set in the real line.) $\square$

1.6 Inverse and Implicit Function Theorem

The inverse and implicit function theorem are two important theorems and we shall look at them in this section.

From last year, we looked at the inverse function theorem in the first dimension. Let $f : (a, b) \to \mathbb{R}$ be continuously differentiable, and suppose there exists $p \in (a, b)$ such that $f'(p) \neq 0$. Then there exists a neighbourhood I around p such that $f\mid_I : I \to f(I)$ is bijective and thus has a inverse on I, f^{-1} that is differentiable and

$$(f^{-1})'(y)) = \frac{1}{f'(f^{-1}(y))}.$$

This theorem can be generalised into higher dimensions.

Theorem 6 (C^1 Inverse Function Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open and $f : \Omega \to \mathbb{R}^m$ be continuously differentiable on Ω, and there exists some $q \in \Omega$ such that $Df(q) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ is invertible. Then, there exist open neighbourhoods around q and $f(q)$, namely U, V respectively, such that,*

- *$f : U \to V$ is a bijection;*
- *$f^{-1} : V \to U$ is continuously differentiable;*
- *for all $v \in V$, $Df^{-1}(v) = (Df(f^{-1}(v)))^{-1}$.*

Lemma 1.2 (Contraction Mapping Theorem)**.** Let X be a complete metric space and let $\phi : X \to X$ be a contraction of X. Then there exists an unique x such that $\phi(x) = x$.

Remark. *We shall examine exactly what this theorem states in the later sections on metric spaces and topology.*

Proof. (Part 1 of the C^1 Inverse Function Theorem). We denote $\|A\| = \sup_{\|x\| \le 1} \|Ax\|$ within this proof. Suppose $q \in \Omega$ and $Df(q) = A$ is invertible, then let us define $\epsilon := 1/\|Df(q)\|$. As f is continuously differentiable, there exists some open neighbourhood of $q - U$ such that for all $x \in U$

$$\|Df(x) - A\| < \epsilon,$$

(simply choose U to have diameter smaller than 1). Now, for all $y \in \mathbb{R}^m$, we define

$$\phi_y(x) = x + A^{-1}(y - f(x)).$$

It is easy to see that ϕ_y is differentiable with derivative $D\phi_y(x) = I - A^{-1}Df(x) = A^{-1}(A - Df(x))$, so $\|D\phi_y(x)\| = \|A^{-1}\|\|(A - Df(x))\| < \|A^{-1}\|\epsilon = 1$. By the mean value theorem, $\|\phi_y(x_1) - \phi_y(x_2)\| \le \|x_1 - x_2\|$, that is phi_y is a contraction on B, and hence has a unique fixed point. Now as $f(x) = y$ if and only if x is a fixed point of ϕ_y, we are done. $\square$

It is in general not easy to find the inverse of a function in the higher dimensions, so the inverse function theorem can help us obtain some properties about the inverse that is otherwise difficult or unobtainable.

The inverse function theorem can be used to show existence and uniqueness of solutions of non-linear system of equations. Given $f_i(x_1, \cdots, x_n) = y_i$ for $i = 1, \cdots, n$, we can define $F(\mathbf{x}) = (f_i(\mathbf{x}))^T$. Then, by looking at some open neighbourhood containing $\mathbf{y}$, it might be possible to determine $F^{-1}(\mathbf{y})$.

Let $\Omega, \Omega' \subseteq \mathbb{R}^n$ be open. Then, we say $f : \Omega \to \Omega'$ is a C^1-diffeomorphism is

- $f : \Omega \to \Omega'$ is a bijection;
- $f : \Omega \to \Omega'$ is continuously differentiable;
- for all $x \in \Omega$, $Df(x)$ is invertible.

Remark. *In fact, the set of all C^1-diffeomorphisms from some open $\Omega \subseteq \mathbb{R}^n$ to itself forms a group under composition.*

Theorem 7 (Implicit Function Theorem – Simple ver.)**.** *Let $\Omega \subseteq \mathbb{R}^2$ be open and $f : \Omega \to \mathbb{R}$ is continuously differentiable, moreover, suppose there exists $q = (a, b) \in \Omega$ such that $f(a, b) = 0$ and $D_2 f(a, b) \neq 0$. Then, there exists open $A, B \subseteq \mathbb{R}$ and $g : A \to B$ such that $a \in A$, $b \in B$ and $(x, y) \in A \times B$ satisfies $f(x, y) = 0$ if and only if $y = g(x)$. Furthermore, g is continuously differentiable.*

Proof. Wlog. We assume $D_2 f(p) > 0$, then as f is continuously differentiable, $D_2 f(p)$ is continuous and thus, there exists some open neighbourhoods around a and b – $A, B = (a - \delta_a, a + \delta_a), (b - \delta_b, b + \delta_b)$ respectively, such that for all $u \in A \times B$, $D_2 f(u) > 0$ (this can be obtained by drawing the a square insider the open ball). Now, suppose we define

$$h : B \to \mathbb{R} : y \mapsto f(a, y).$$

As $h'(y) = D_2 f(a, y) > 0$, we see that h is strictly increasing. Furthermore, as $h(b) = f(a, b) = 0$, we have $h(b - \delta_b/2) < h(b) = 0$, and similarly, $h(b + \delta_b/2) > 0$. Thus, there exists some $\delta_a > \delta' > 0$ such that $f(x, b - \delta_b/2) < 0$ and $f(x, b + \delta_b/2) > 0$ for all $x \in (a - \delta', a + \delta')$. Now, by the intermediate value theorem, for all $x \in (a - \delta', a + \delta')$ there exists (uniquely) some $y_x \in (b - \delta_b/2, b + \delta_b/2)$ such that $f(x, y_x) = 0$ (unique as $D_2 f(x, y) > 0$). Thus, we can define

$$g : (a - \delta', a + \delta') \to (b - \delta_b/2, b + \delta_b/2) : x \mapsto y_x.$$

We see straight away that g is continuously differentiable as f is. So we are done. □

There is a more general version of this theorem applying to arbitrary dimensions.

Theorem 8 (Implicit Function Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^n, \Omega' \subseteq \mathbb{R}^m$ be open, and $f : \Omega \times \Omega' \to \mathbb{R}^m$ be continuously differentiable on $\Omega \times \Omega'$. Suppose there exists $p = (a, b)\Omega \times \Omega'$ such that $f(p) = 0$ and $D_{n+j} f^i(p)$ is invertible for all $1 \leq i, j \leq m$. Then there are $A \subseteq \Omega, B \subseteq \Omega'$ with $a \in A, b \in B$ such that there exists a map $g : A \to B$ in which, $f(x, y) = 0$ if and only if $y = g(x)$ for some $x \in A$. Furthermore, g is continuously differentiable.*

2 Metric and Topological Spaces

2.1 Metric Spaces

Definition 2.1 (Metric)**.** Let X be some set. Then, a metric d on X is a function from X^2 to $\mathbb{R}$ such that,

- for all $x, y \in X$, $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$;
- for all $x, y \in X$, $d(x, y) = d(y, x)$;
- for all $x, y, z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$.

We call the ordered pair (X, d) where d is a metric on X, a metric space. In general, for short hand, we simply refer to the metric space (X, d) as X.

Definition 2.2 (Subspace)**.** Let (X, d) be a metric space. Then, for all $Y \subseteq X$, $(Y, d \mid_Y)$ forms a metric space where

$$d \mid_y : Y \times Y \to \mathbb{R} : (x, y) \mapsto d(x, y).$$

We call this metric space a metric subspace of (X, d).

Up to now, we have seen examples of metric spaces in terms of normed vector spaces. It is not difficult to see that all normed vector spaces induces a metric space (where $d(u, v) := \|u - v\|$) while the reverse is in general not true. We shall formally define the notion of norm here.

Definition 2.3 (Norm)**.** Let V be some vector space over a field $\mathbb{R}$. Then a function $\| \cdot \| : V \to \mathbb{R}$ is a norm on V if

- for all $v \in V$, $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$;
- for all $v \in V$, and $\lambda \in \mathbb{R}$, $\|\lambda v\| = |\lambda| \, \|v\|$;
- for all $u, v \in V$, $\|u + v\| \leq \|u\| + \|v\|$.

We see that this definition can be extending to vector spaces over arbitrary fields if and only if there is a notion of absolute value on that field.

Similar to metric spaces, we call the ordered pair $(V, \| \cdot \|)$ where $\| \cdot \|$ is a norm on V a normed vector space.

Definition 2.4 (n-Norm)**.** Given $\mathbb{R}^n$, the n-norm $\| \cdot \|_n$ on $\mathbb{R}^n$ is the norm such that for all $v \in \mathbb{R}^n$, $\|v\|_n = (\sum v_i^n)^{\frac{1}{n}}$. We also define the ∞-norm as $\|v\|_\infty = \max\{|v_i|\}$.

As one might expect, d_n is the metric induced by the n-norm and similarly, d_∞ is the metric induced by the ∞-norm.

Definition 2.5 (Open Ball)**.** Let (X, d) be a metric space and suppose $x \in X$. Then a ball of radius $\epsilon > 0$ at x is the set $B_\epsilon(x) = \{y \in X \mid d(x, y) < \epsilon\}$.

This is also called the ϵ-ball about x or the ϵ-neighbourhood of x.

Definition 2.6 (Open)**.** Let (X, d) be a metric space and let $U \subseteq X$. We say that U is open on (X, d) if and only if for all $u \in U$, there exists some $\delta > 0$ such that $B_\delta(u) \subseteq U$.

Lemma 2.1. For all metric spaces (X, d), $\varnothing$ and X are open.

Proof. Follows from definition. □

Lemma 2.2. Let (X, d) be a metric space and let $\mathcal{C}$ be a collection of open sets, then $\bigcup \mathcal{C}$ is open. On the other hand, if $\mathcal{C}$ is finite, then $\bigcap \mathcal{C}$ is also open.

Proof. Easy. □

Definition 2.7 (Topologically Equivalent)**.** Let d_1, d_2 be two metrics on X. We call d_1 and d_2 topologically equivalent if and only if for all $U \subset X$, U is open with respect to d_1 if and only if is it open with respect to d_2.

With this definition, we see that the family of $\{d_n \mid n \in \bar{\mathbb{N}}\}$ is all topologically equivalent to one another.

There is another notion of equivalence for metrics called the Lipschitz equivalence.

Definition 2.8 (Lipschitz Equivalence)**.** Let d_1, d_2 be two metrics on X. We call d_1 and d_2 Lipschitz equivalent if and only if there exists $M_0, M_1 \in \mathbb{R}^+$ such that for all $x, y \in X$,

$$M_0 d_1(x, y) \leq d_2(x, y) \leq M_1 d_1(x, y).$$

It is not difficult to see that Lipschitz equivalence is a stronger notion of equivalence since if two metrics are Lipschitz equivalent then they are topologically equivalent. This can be proved by choosing $\delta' = M_0 \delta$.

Sometimes, it is also useful to induce a new metric using an existing one by pushing it over some function. It turns out, this is possible over continuous and injective functions.

Theorem 9. *Let (X, d_X) be a metric space and let $f : Y \to X$ be continuous and injective. Then*

$$d_Y : Y \times Y \to \mathbb{R} : (y_1, y_2) \mapsto d_X(f(y_1), f(y_2))$$

is a metric on Y.

Proof. d_Y is trivially non-negative and symmetric so let us consider whenever $d_Y = 0$. Let $y_1, y_2 \in Y$ such that $d_Y(y_1, y_2) = 0$, then by definition, $f(y_1) = f(y_2)$, and so, as f is injective, $y_1 = y_2$.

Thus, it suffices to prove the triangle inequality for d_Y. let $y_1, y_2, y_3 \in Y$, then

$$\begin{aligned} d_Y(y_1, y_2) &= d_X(f(y_1), f(y_2)) \\ &\leq d_X(f(y_1), f(y_3)) + d_X(f(y_3), f(y_2)) \\ &= d_Y(y_1, y_3) + d_Y(y_3, y_2). \end{aligned}$$

□

2.2 Basic Notions in Metric Spaces

In this section, we will establish some basic notions regarding metric spaces.

Definition 2.9 (Convergence)**.** Let (X, d) be a metric space and suppose $(x_n)_{n=1}^{\infty}$ is a sequence in X. Then, we say x_n converges to some $x \in X$ is and only if for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $d(x, x_n) < \epsilon$.

As this definition of convergence is essentially the same as the one we looked at last year, it follows some similar properties.

Lemma 2.3. Let (X, d) be a metric space and suppose $(x_n)_{n=1}^{\infty}$ is a sequence in X. Then, if x_n converges in X, the its limit is unique.

Proof. Suppose $x_n \to a$ and $x_n \to b$ for some $a, b \in X$. Then for all ϵ, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $d(x_n, a), d(x_n, b) < \epsilon/2$ and thus, by triangle inequality, $d(a, b) < \epsilon/2 + \epsilon/2 = \epsilon$. As ϵ is an arbitrary positive number, $d(a, b) = 0$ and so $a = b$. $\square$

Definition 2.10 (Closed Set)**.** Let (X, d) be a metric space and $U \subseteq X$. Then we say U is closed in X is and only if for any convergent sequence $(x_n)_{n \geq 1}$ in U converges in U.

We see that this is the same definition we had for closed sets in $\mathbb{R}$ and similarly to what we had found last year, a set is closed if and only if its complement is open.

Proposition 4. *Let (X, d) be a metric space and let $U \subseteq X$. Then U is closed in X if and only if U^c is open.*

Proof. ($\implies$) We prove the contrapositive. Suppose U^c is not open, then there exists $x \in U^c$ such that for all $\epsilon > 0$, $B_\epsilon(x) \not\subseteq U^c$. So, for all $n \in \mathbb{N}$, $B_{1/n}(x) \cap U \neq \varnothing$. So, by defining the sequence $(x_n)_{n \geq 1}$ such that $x_n \in B_{1/n}(x) \cap U$ we have a sequence in U. Now, we see that $x_n \to x$ as for all ϵ, we can choose $N \geq 1/\epsilon$ so, for all $n \geq N$, $x_n \in B_{1/n}(x) \subseteq B_{1/N}(x) \subseteq B_\epsilon(x)$. But $x \notin U$ and thus, U is not closed.

($\impliedby$) The reverse is by similar argument. Suppose there exists a sequence $(x_n)_{n \geq 1}$ in U such that x_n converges to $x \in U^c$. Then, if U^c is open, there exists some $\epsilon > 0$ such that $B_\epsilon(x) \subseteq U^c$. But now as $x_n \to x$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \in B_\epsilon(x)$ implying $B_\epsilon(x) \not\subseteq U^c$ since $x_n \notin U^c$ # $\square$

We might sometimes find the above condition as the definition of closed sets instead, that is, a set is closed if and only if its complement is open. However, this does not matter by the above proposition. The question of which definition to use is rather pedagogical since the definition we presented resembles more of that we had defined for Euclidean spaces while the latter definition resembles the topological definition which we shall examine later.

Definition 2.11 (Interior, Isolated, Limit and Boundary Point)**.** Let (X, d) be a metric space and $V \subseteq X$. Then

- a point $x \in V$ is an interior point (or inner point) of V if and only if there exists some $\delta > 0$ such that $B_\delta \subseteq V$;

- $x \in V$ is an isolated point of V is and only if there exists there exists some $\delta > 0$ such that $B_\delta(x) \cap V = \{x\}$;
- $x \in X$ is a limit point (or accumulation point) of V if and only if for all $\epsilon > 0$, $(B_\epsilon \cap V) \setminus \{x\} \neq \varnothing$;
- $x \in X$ is a boundary point of V if and only if for all $\epsilon > 0$, $B_\epsilon(x) \cap V \neq \varnothing$ and $B_\epsilon(x) \cap V^c \neq \varnothing$.

Straight away, we see that every element of an open set is an interior point while having no boundary points.

Definition 2.12. Let (X, d) be a metric space and $V \subseteq X$. Then,

- the interior of V is the set of all interior points of V;
- the closure of V, $\bar{V}$ is the union of V and all limit points of V;
- the boundary of V is the set of boundary points of V.

Again, the notion of the closure might be different in other literatures. We might see the closure of V as the intersection of all closed sets that are greater or equal to V, that is the smallest closed set containing V. We once again find these two definitions to be equivalent.

Proposition 5. *Let (X, d) be a metric space and $V \subseteq X$. Then,*

$$\bar{V} = \bigcap_{\substack{V \subseteq U \subseteq X \\ U \text{ closed}}} U$$

Proof. It is obvious that the intersection of collections of closed sets is closed, so to prove this proposition, it suffices to show that $\bar{V}$ is closed in X and for all closed $V \subseteq U \subseteq X$, U contains the limit points of V.

Let $x \in V$ be a limit point of V, then, for all $\epsilon > 0$, $B_\epsilon(x) \cap V \neq \varnothing$. Thus, we can construct a sequence in V converging to x by the same method as proposition 4. Now, as $V \subseteq U$, this is also a sequence in U, to its limit x, is also in U.

Now, we show $\bar{V}$ is closed. Suppose there exists some sequence $(x_n)_{n \geq 1}$, $x_n \to x$, we need to show $x \in \bar{V}$. But, we see straight away x is a limit point of V since for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \in B_\epsilon(x)$, and as $x_n \in V$, $x_n \in B_\epsilon(x) \cap V$, so the intersection is not empty and we are done! $\square$

2.3 Maps Between Metric Spaces

Just as how we can consider properties about functions of real numbers, we can define equivalent properties for function between metric spaces.

Definition 2.13 (Continuity at a Point)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. Then, we say that f is continuous at some $x \in X$ if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $f(B_\delta(x)) \subseteq B_\epsilon(f(x))$.

Definition 2.14 (Uniformly Continuous)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. Then, we say that f is uniformly continuous if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x \in X$, $f(B_\delta(x)) \subseteq B_\epsilon(f(x))$.

This is essentially the same idea as the definition of continuity in Euclidean spaces. While, it makes sense to define it this way, we shall encounter a different notion of continuity later in this course – continuity on topological spaces. While we shall not talk about it too much here, it turns out that these notions of continuity is consistent on metric spaces, that is, a function is continuous on some metric spaces if and only if it is continuous on their induced topologies.

Proposition 6. *Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. Then f is continuous on X if and only if $f^{-1}(U)$ is open in X for all open $U \subset Y$*

Proof. ($\implies$) Suppose f is continuous, then, for all $v \in f^{-1}(U)$, where U is open in Y, we have $f(v) \in U$. So, as U is open, there exists some $\epsilon > 0$ such that $B_\epsilon(f(v)) \subseteq U$. By continuity of f, there exists some $\delta > 0$ such that $f(B_\delta(v)) \subseteq B_\epsilon(f(v)) \subset U$ so, $B_\delta(v) \subset f^{-1}(U)$, and thus, $f^{-1}(U)$ is open.

($\impliedby$) Suppose now that $f^{-1}(U)$ is open in X for all open $U \subset Y$, then for all $x \in X$, $\epsilon > 0$, as $B_\epsilon(f(x))$ is open, $f^{-1}(B_\epsilon(f(x)))$ is also open. Now, as $x \in f^{-1}(B_\epsilon(f(x)))$, there exists some $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(B_\epsilon(f(x)))$ and so, $f(B_\delta(x)) \subseteq B_\epsilon(f(x))$. □

Proposition 7. *Let $(X, d_X), (Y, d_Y), (Z, d_Z)$ be metric spaces and let f and $g : Y \to Z$ be maps, then, if f, g are both continuous at some $x \in X$, then $g \circ f$ is continuous at x.*

Proof. Trivial. □

Definition 2.15 (Homeomorphism)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. Then, we call f a homeomorphism between X and Y if and only if f is a bijection and both f and f^{-1} are continuous. If such a homeomorphism exists between two metric spaces, then we call them homeomorphic.

A classic example of homeomorphisms between metric spaces is the homeomorphism between $(-1, 1)$ and $\mathbb{R}$ by scaling tan (of course this is not the only homeomorphism between the two, we see that $x \mapsto \log(1 + x) - \log(1 - x)$ is also a homeomorphism). With this, we can conclude that $\mathbb{R}$ is homeomorphic to any open intervals. Furthermore, we find that $(a_1, b_1]$ is homeomorphic to $[a_2, b_2)$ by considering that $(-1, 1] \cong [-1, 1)$ by $x \mapsto -x$.

Definition 2.16 (Lipschitz Continuous)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. We call f Lipschitz continuous if and only if there exists some $M > 0$ such that for all $x_1, x_2 \in X$,

$$d_Y(f(x_1), f(x_2)) \leq M d_X(x_1, x_2).$$

Definition 2.17 (bi-Lipschitz)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. We call f bi-Lipschitz if and only if there exists $M_0, M_1 > 0$ such that for all $x_1, x_2 \in X$, we have

$$M_2 d_X(x_1, x_2) \leq d_Y(f(x_1), f(x_2)) \leq M_1 d_X(x_1, x_2).$$

We see that bi-Lipschitz is stronger than Lipschitz continuity which is in turn a stronger condition than uniform continuity by simply choosing $\delta = \epsilon/2M$.

Definition 2.18 (Isometry)**.** Let $(X, d_X), (Y, d_Y)$ be two metric spaces and let $f : X \to Y$ be a map. We call f an isometry if and only if for all $x_1, x_2 \in X$,

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2).$$

3 Extras

This is a temporary section and all contents within this section will be moved to their appropriate sections once the notes is complete.

3.0.1 Completeness is not a Topological Property

Completeness is not a topological property. Heuristically this makes sense as in general, it does not make sense to call a metric space complete as completeness is a property of metric spaces and not all topologies are metric spaces.

Theorem 10. *Completeness is not a Topological Property.*

Proof. To show this, we need find two metrics on some set that induces the same topology but one is complete while the other is not. We see that $(0, 1)$ is not complete with respect to the Euclidean metric but $\mathbb{R}$ is with $\mathbb{R}$ being homeomorphic to $(0, 1)$. Thus, the main idea is to pull back the metric on $\mathbb{R}$ on to $(0, 1)$ over the homeomorphism.

By theorem 9, as $\tan : (0, 1) \to \mathbb{R}$ is injective we have a induced metric on $(0, 1)$, $d_{\tan}$. I claim the topologies induced by $(X, d_{\|\cdot\|})$ and $(X, d_{\tan})$ are the same, that is for all $U \subseteq (0, 1)$, U is open in $d_{\|\cdot\|}$ if and only if it is open in $d_{\tan}$. But this is true as tan and arctan are both continuous, so we are done! $\square$

3.0.2 Properties of the Quotient Map

Theorem 11. *Let $(X, \mathcal{T})$ be a topological space and let $\sim$ be an equivalence relation on X. Let us denote the quotient map from X to $X/\sim$ as q where $q(x) = [x] = \{x' \in X \mid x \sim x'\}$. Then, q is continuous and surjective.*

Proof. Surjectivity is trivial while q being continuous is true by definition. $\square$