Analysis and Topology

Shing Tak Lam

May 11, 2022

Contents

1 Topological spaces

Definition 1.1 (Topological space)

Given a set X, a topology on X is a collection $\mathscr{T}$ of subsets of X such that

- $\varnothing, X \in \mathscr{T}$
- If $U_i \in \mathscr{T}$ for all $i \in I$, then $\bigcup_{i \in I} U_i \in \mathscr{T}$
- If $U_1, U_2 \in \mathscr{T}$, then $U_1 \cap U_2 \in \mathscr{T}$.

and a topological space is a pair $(X, \mathscr{T})$.

Definition 1.2 (Open and closed sets)

For $A \subseteq X$, A is open if $A \in \mathscr{T}$, and A is closed if $X \setminus A \in \mathscr{T}$

Proposition 1.3.

- $\varnothing, X$ are closed
- If U_i are closed for all $i \in I$, then so is $\bigcap i \in IU_i$
- If U_1, U_2 are closed, then so is $U_1 \cup U_2$

Remark 1.4. This gives us an equivalent way to defining a topology, by specifying the closed sets. This can be useful, eg. for the Zariski topology.

Definition 1.5 (Interior)
For $A \subseteq X$, define the interior of A to be

$$\operatorname{Int} A = \bigcup_{U \subseteq A, U \text{ open}} U$$

Definition 1.6 (Closure)
For $A \subseteq X$, define the closure of A to be

$$\overline{A} = \bigcap_{U \subseteq X, U \text{ closed}} U$$

Proposition 1.7. For all A, we have that $\operatorname{Int} A \subseteq A \subseteq \overline{A}$, with equality holding if and only if A open (closed resp.).

Proposition 1.8. For all $x \in X$, we have that $x \in \operatorname{Int} A$ if and only if there exists U open such that $x \in U \subseteq A$.

Proposition 1.9. For all $x \in X$, $x \in \overline{A}$ if and only if for all U open such that $x \in U$, $U \cap A \neq \varnothing$.

Proposition 1.10. For any $A \subseteq X$, $U \subseteq A$ open, $A \subseteq B$ closed, then

$$U \subseteq \operatorname{Int} A \subseteq A \subseteq \overline{A} \subseteq B$$

Definition 1.11 (Dense)
$A \subseteq X$ is dense in X if $\overline{A} = X$.

Definition 1.12 (Separable)
X is separable if it has a countable dense subset.

Definition 1.13 (Hausdorff)
A topological space X is Hausdorff if for all $x, y \in X$ distinct, we have open sets U, V such that $x \in U$, $y \in V$ and $U \cap V = \varnothing$.

Definition 1.14 (Convergence)
Let X be a topological space, $x \in X$. Then $x_n \to x$ if for all neighbourhoods U of x, there exists N such that for all $n \geq N$, $x_n \in U$.

Definition 1.15 (Continuity)
Let X, Y be topological spaces. $f : X \to Y$ is continuous if for all $U \subseteq Y$ open, $f^{-1}(U) \subseteq X$ is open.

Definition 1.16 (Open map)
Let $f : X \to Y$ is open if for all $U \subseteq X$ open, $f(U) \subseteq Y$ is open.

Definition 1.17 (Homeomorphism)
$f : X \to Y$ is a homeomorphism if f is a bijection, f and f^{-1} are both continuous.

1.1 Constructions

Subspaces

Definition 1.18 (Subspace topology)
Suppose X is a topological space, $Y \subseteq X$. Then

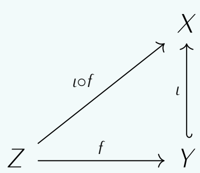
$$\{V \cap Y : V \text{ open in } X\}$$

defines a topology on Y, known as the subspace topology.

Proposition 1.19. The inclusion map $Y \hookrightarrow X$ is continuous.

Proof. Let $V \subseteq X$ be open, $\iota : Y \hookrightarrow X$ be the inclusion map. Then $\iota^{-1}V = V \cap Y$ is open by definition of the subspace topology. □

Proposition 1.20 (Universal property of subspaces). Suppose $Y \subseteq X$ and Z is any topological space, $f : Z \to Y$. Then



f is continuous if and only if $\iota \circ f$ is continuous.

Proof. Suppose f is continuous. Then $\iota \circ f$ is a composition of continuous functions, so continuous.

Now suppose $\iota \circ f$ is continuous. Let $U \subseteq Y$ be open. Then there exists $V \subseteq X$ open such that $U = V \cap Y$. Then

$$(\iota \circ f)^{-1}(V) = f^{-1}\left(\iota^{-1}(V)\right) = f^{-1}(U)$$

is open. So f is continuous. □

Product topology

Definition 1.21 (Basis)
For a set X, a basis $\mathcal{B}$ for a topology on X is any collection of subsets of X such that

- $\mathcal{B}$ covers X, that is, $\bigcup_{B \in \mathcal{B}} B = X$,
- For any $B_1, B_2 \in \mathcal{B}$, $x \in B_1 \cap B_2$, there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Definition 1.22 (Generated topology)
Given a basis $\mathcal{B}$, the collection

$$\left\{ \bigcup_{i \in I} B_i : B_i \in \mathcal{B} \right\}$$

forms a topology on X.

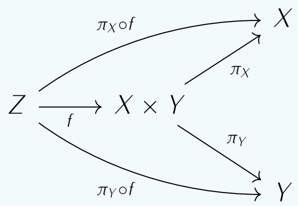
Definition 1.23 (Box topology)
Let X, Y be topological spaces. The topology generated by the basis

$$\{U \times V : U \text{ open in } X, V \text{ open in } Y\}$$

is known as the box topology (or the product topology for finite products) on $X \times Y$.

Proposition 1.24. The projection maps $X \times Y \to X$ and $X \times Y \to Y$ are continuous.

Proposition 1.25 (Universal property of product topology)**.** Suppose we have



Then f is continuous if and only if $\pi_X \circ f$ and $\pi_Y \circ f$ are continuous.

Proof. If f is continuous then $\pi_X \circ f$ and $\pi_Y \circ f$ are compositions of continuous functions. Now suppose $\pi_X \circ f$ and $\pi_Y \circ f$ are continuous.

Let $W \subseteq X \times Y$ be open. Without loss of generality, we may assume that $W = U \times V$ is a basis element. Then we have that

$$f^{-1}(U \times V) = f^{-1}\left((U \times Y) \cap (X \times V)\right) = f^{-1}\left(\pi_X(W)\right) \cap f^{-1}\left(\pi_Y(W)\right)$$

is open. So f is continuous. □

Proposition 1.26. $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ and $y_n \to y$.

Proof. Suppose $(x_n, y_n) \to (x, y)$. For any neighbourhood U of x, then $U \times Y$ is a neighbourhood of (x, y). So by convergence we have N such that for all $n \geq N$, $x_n \in U$ for all $n \geq N$.

Conversely, let W be a neighbourhood of (x, y). Without loss of generality, we may assume $W = U \times V$. Then for n sufficiently large, $(x_n, y_n) \in W$. □

Quotient topology

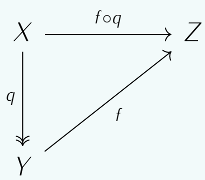
Definition 1.27 (Quotient topology)
Let X be a topological space, $q : X \twoheadrightarrow Y$ surjective. Then the collection

$$\left\{V : q^{-1}(V) \text{ open in } X\right\}$$

defines a topology on Y.

Proposition 1.28. The quotient map $q : X \twoheadrightarrow Y$ is continuous.

Proposition 1.29 (Universal property of quotient topology)**.** Suppose we have



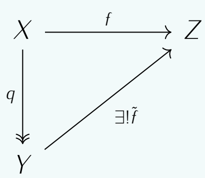
where Y has the quotient topology, then f is continuous if and only if $f \circ q$ is continuous.

Proof. Suppose $f \circ q$ continuous. Let $V \subseteq Z$ be open. Then

$$(f \circ q)^{-1}(V) = q^{-1}\left(f^{-1}(V)\right)$$

is open in X, which means that $f^{-1}(V)$ is open in Y, so f is continuous. □

Proposition 1.30. Suppose we have



Then if f is open, so is $\tilde{f}$.

Proposition 1.31. Let X be a topological space, R be an equivalence relation on X. Suppose X/R is Hausdorff. Then $R \subseteq X \times X$ is closed.

Proof. Let $W = X \times X \setminus R$. Then for any $(x, y) \in W$, $q(x) \neq q(y)$. Then we have neighbourhoods S and S of $q(x), q(y)$ respectively such that $S \cap T = \varnothing$. Now let $U = q^{-1}(S)$ and $V = q^{-1}(T)$. Then $(x, y) \in U \times V$, and $U \times V \cap R = \varnothing$. So W is open, and R is closed. □

Proposition 1.32. Let X be a topological space, and suppose R is an equivalence relation on X, $R \subseteq X \times X$ closed, and $q : X \to X/R$ open. Then X/R is Hausdorff.

Proof. Let z, w be distinct points in X/R. Choose $x, y \in X$ such that $q(x) = z$ and $q(y) = w$. Let $W = X \times X \setminus R$. Then $(x, y) \in W$, which is open as R is closed. So we have U, V such that $(x, y) \in U \times V \subseteq W$. Then $q(U)$ and $q(V)$ are disjoint neighbourhoods of z and w respectively. □

2 Connectedness

Definition 2.1 (Connected)

A topological space X is disconnected by A, B if $A, B \subseteq X$ open, A, B nonempty, $A \cap B = \varnothing$ and $X = A \cup B$.
X is connected if no such A, B exists.

Definition 2.2 (Interval)

$I \subseteq \mathbb{R}$ is an interval if for all $x, z \in I$, if $x < y < z$ then $y \in I$.

Theorem 2.3. For a topological space X, the following are equivalent.

- X is connected.
- For any continuous function $f : X \to \mathbb{R}$, $f(X)$ is an interval.
- For any continuous function $f : X \to \mathbb{Z}$, f is constant.

Proof. For (i) implies (ii), we prove the contrapositive. Suppose $f(X)$ is not connected. Then we have $x, z \in f(X)$, $x < y < z$, $y \notin f(X)$. Then $f^{-1}((-\infty, y))$ and $f^{-1}((y, \infty))$ disconnect X.

For (ii) implies (iii), then consider $X \to \mathbb{Z} \hookrightarrow \mathbb{R}$, the image is an interval which only contains integers, which must be a singleton.

For (iii) implies (i), suppose $X = A \cup B$ is not connected. Then $f = \mathbb{1}_A$ is continuous, but not constant. □

Proposition 2.4. X is connected if and only if the only clopen sets are X and $\varnothing$.

Proposition 2.5. $X \subseteq \mathbb{R}$ is connected if and only if it is an interval.

Proposition 2.6. Let $Y \subseteq X$, then Y is disconnected if and only if we have U, V open such that

- $U \cap Y, V \cap Y \neq \varnothing$,
- $U \cap V \cap Y = \varnothing$,
- $Y \subseteq U \cup V$.

Proposition 2.7. If $Y \subseteq X$ is connected, then $\overline{Y}$ is connected.

Proof. Suppose U, V disconnect $\overline{Y}$. If $U \cap Y$ and $V \cap Y$ are both nonempty, then U and V will disconnect Y. So without loss of generality, we may assume $V \cap Y = \varnothing$. Then $Y \subseteq X \setminus V$, which is closed, so $\overline{Y} \subseteq X \setminus V$, contradiction. □

Corollary 2.8. If $Y \subseteq Z \subseteq \overline{Y}$, then if Y is connected, Z is connected.

Theorem 2.9. If $f : X \to Y$ is continuous, X is connected, then $f(X)$ is connected.

Proof. If U, V disconnect $f(X)$, then $f^{-1}(U), f^{-1}(V)$ disconnect X. □

Corollary 2.10. If X is connected, then X/R is connected.

Lemma 2.11. Let $\mathcal{A}$ be a family of connected subsets of X. Suppose B is a connected subset of X such that for all $A \in \mathcal{A}$, $A \cap B \neq \varnothing$. Then

$$Y = \left(\bigcup_{A \in \mathcal{A}} A \right) \cup B$$

is connected.

Proof. Let $f : Y \to \mathbb{Z}$ be continuous. For all A, $f|_A$ is constant. Similarly, $f|_B$ is constant. But for all $A \in \mathcal{A}$, $A \cap B \neq \varnothing$, so f is contant. □

Theorem 2.12. If X, Y are connected, then $X \times Y$ is connected.

Proof. Fix $x_0 \in X$. Let $B = \{x_0\} \times Y$, and $\mathcal{A} = \{X \times \{y_0\} : y_0 \in Y\}$. Then apply above lemma. □

2.1 Path connectedness

Definition 2.13 (Path)
For $x, y \in X$, a path from x to y is a continuous function

$$\gamma : [0, 1] \to X$$

such that $\gamma(0) = x$ and $\gamma(1) = y$.

Definition 2.14 (Path connected)
A topological space X is path connected if for all $x, y \in X$, there exists a path from x to y.

Theorem 2.15. If X is path connected, then X is connected.

Proof. Suppose $X = U \cup V$ is disconnected. Let $x \in U$, $y \in V$, and $\gamma : x \to y$ path. Then $\gamma^{-1}(U)$ and $\gamma^{-1}(V)$ disconnect $[0, 1]$. □

2.2 Components

Connected components

Definition 2.16 (Connected components)
Let X be a topological space, define the equivalence relation $\sim$ on X where

$$x \sim y \iff \exists A \text{ connected, s.t. } x, y \in A$$

The elements of $X/\sim$ are known as connected components of X.

Proposition 2.17. The connected components are nonempty, maximal, connected, closed subsets of X.

Proof. Nonempty is clear as it is an element of a partition. Now suppose if C is a connected component, $C \subseteq A$ where A is connected. Then for any $x \in C$, $y \in A$, we have that $x \sim y$. So $A = C$ and C is maximal.

Fix $x \in C$. For any $y \in C$, we have A_y connected such that $x, y \in A_y$. Then $A = \bigcup_{y \in C} A_y$ is connected, by maximality, $A = C$, so C is connected.

Since C is connected, so is $\overline{C}$, so by maximality $C = \overline{C}$, so C is closed. □

Path components

Definition 2.18 (Path components)
Let X bs a topological space, define the equivalence relation $\sim$ on X by

$$x \sim y \iff \exists \text{ path } \gamma : x \to y$$

Theorem 2.19. Let $U \subseteq \mathbb{R}^n$ be open. Then U is connected if and only if U is path connected.

Proof. One implication is always true. Therefore, assume U connected. Without loss of generality, assume U nonempty. Fix $x_0 \in U$, and let $P = \{x \in U : \exists \gamma : x \to x_0\}$. We will show that P is clopen, and thus must be the whole set U.

Fix $x \in P$, then there exists $r > 0$ such that $D(x; r) \subseteq U$. Then $D(x; r)$ is path connected, so for any $y \in D(x; r)$, we have a path $x_0 \to x \to y$. Thus $D(x; r) \subseteq P$, and P is open.

Now fix $x \in U \setminus P$, for any $y \in D(x; r)$, if $y \in P$ then we have a path $x_0 \to y \to x$. So $D(x; r) \subseteq U \setminus P$, and P is closed. □

3 Compactness

Definition 3.1 (Compact space)
A topological space X is compact if every open cover has a finite subcover.

Lemma 3.2. Let X be compact, (U_i) be a sequence of open sets, $U_1 \subseteq U_2 \subseteq \cdots$, and $\bigcup_i U_i = X$. Then there exists N such that for all $n \geq N$, $U_n = U_N = X$.

Proof. The U_i form an open cover. Take a finite subcover. □

Lemma 3.3. Let X be compact, (T_i) be a sequence of closed sets, $T_1 \supseteq T_2 \supseteq \cdots$, all $T_i \neq \varnothing$. Then

$$\bigcap_{i=1}^{\infty} T_i \neq \varnothing$$

Proof. Suppose not. Let $U_i = X \setminus T_i$. Then the U_i form an ascending chain, which covers X. So by previous lemma, there exists N such that $U_N = X$. But this means $T_N = \varnothing$. Contradiction. ☐

Theorem 3.4. Suppose $f : X \to \mathbb{R}$ continuous, X compact, then f is bounded and attains its bounds.

Proof. For $n \in \mathbb{N}$, let $U_n = \{x \in X : |f(x)| < n\}$. By above lemma, U_n is eventually constant, so the function is bounded. Let $\alpha = \inf f(X)$, and suppose there is no $x \in X$ such that $f(x) = \alpha$. So for all $x \in X$, there exists n such that $f(x) > \alpha + 1/n$. Let $V_n = \{x \in X : f(x) > \alpha + 1/n\}$. Again this is an increasing sequence, so eventually constant. Thus we have some N such that for all $x \in X$, $f(x) > \alpha + 1/N$. But α is the infimum of the image. Contradiction. ☐

Lemma 3.5. Suppose $Y \subseteq X$, then Y is compact if and only if for every collection $\mathscr{U}$ of open sets of X such that $Y \subseteq \bigcup_{U \in \mathscr{U}} U$, we have a finite subcollection which covers Y.

Theorem 3.6. $[0, 1] \subseteq \mathbb{R}$ is compact.

Proof. For $A \subseteq [0, 1]$, we say that A is finitely covered by $\mathscr{U}$ if there is a finite subcollection of open sets in $\mathscr{U}$ which cover $\mathscr{U}$. Thus, we want to show that $[0, 1]$ is finitely covered.

Suppose not. Then at least one of $[0, 1/2]$ and $[1/2, 1]$ is not finitely covered. Say $[a_1, b_1]$. Let $c = \frac{1}{2}(a_1 + b_1)$. Then at least one of $[a_1, c]$ and $[c, b_1]$ will not be finitely covered. Repeating this, we get

$$[a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots$$

Such that $[a_n, b_n]$ are all not finitely covered, and $b_n - a_n = 2^{-n}$. Then $a_n \to x$ and $b_n \to x$ for some $x \in [0, 1]$. Choose $U \in \mathscr{U}$ such that $x \in U$. Since U is open, there must be $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq U$. Since $a_n, b_n \to x$, choose N such that $x - \varepsilon \leq a_N \leq x \leq b_N \leq x + \varepsilon$. Contradiction. ☐

Theorem 3.7. Let $Y \subseteq X$, suppose X compact, Y closed in X. Then Y is compact.

Proof. Let $\mathscr{U}$ be an open cover of Y by sets open in X. Then $\mathscr{U} \cup \{X \setminus Y\}$ is an open cover of X, so we have a finite subcover, which must cover Y. ☐

Theorem 3.8. Let $Y \subseteq X$, suppose X is Hausdorff, Y is compact, then Y is closed in X.

Proof. Fix $x \in X \setminus Y$. By the Hausdorff property, for all $y \in Y$, we have U_y, V_y neighbourhoods of x, y respectively such that $U_y \cap V_y \neq \varnothing$. The V_ys cover Y, so by compactness we have a finite subcover,

$$V_{y_1} \cup \cdots \cup V_{y_n} \supseteq Y$$

Thus, $U_{y_1} \cap \cdots \cap U_{y_n}$ is a neighbourhood of x, which is disjoint from Y. ☐

Theorem 3.9. Let $f : X \to Y$, X is compact, then $f(X)$ is compact.

Corollary 3.10. Any quotient of a compact space is compact.

Theorem 3.11 (Topological inverse function theorem)**.** Let $f : X \to Y$ be a continuous bijection, X is compact and Y Hausdorff. Then $f^{-1} : Y \to X$ is continuous. Equivalently, f is open, f is a homeomorphism.

Proof. Let $U \subseteq X$ be open. Then $K = X \setminus U$ is closed, so K is compact. So $f(K)$ is compact, and $f(K)$ is closed in Y. Thus $f(U) = Y \setminus f(K)$ is open in Y. □

Theorem 3.12 (Tychonoff)**.** Let X, Y be compact. Then $X \times Y$ is compact.

Proof. Let $\mathscr{U}$ be an open cover. Without loss of generality, we may assume that all elements of $\mathscr{U}$ are basis elements, ie. $U \times V$. Then for all $z \in X \times Y$, we have U_z, V_z such that $z \in U_z \times V_z \in \mathscr{U}$.

Fix $x \in X$. Then $\{x\} \times Y$ is compact, and we have an open cover by elements of $\mathscr{U}$, so we have a finite subcover, say

$$\left(U_{x,i} \times V_{x,i}\right)_{i=1}^{n_x}$$

Let $U_x = \bigcap_{i=1}^{n_x} U_{x_i}$. Then we have that

$$U_x \times Y \subseteq \bigcup_{i=1}^{n_x} \left(U_{x,i} \times V_{x,i}\right)$$

Now $(U_x)_{x \in X}$ is an open cover of X, so we have a finite subcover, say $U_{x_1}, \dots, U_{x_k}$. Then

$$X \times Y = \bigcup_{i=1}^{k} \bigcup_{j=1}^{n_{x_i}} (U_{x_i,j} \times V_{x_i,j})$$

is finitely covered. □

Theorem 3.13 (Heine-Borel)**.** $K \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proof. $\mathbb{R}^n$ is Hausdorff, so if K is compact it must be closed. Furthermore, $x \mapsto ||x||$ is continuous, so K is bounded.

Conversely, if K is bounded, then $K \subseteq [-M, M]^n$ for some $M > 0$. So it is a closed subspace of a compact space, hence compact. □

4 Metric spaces

Definition 4.1 (Metric space)

For a set X, a metric $d : X \times X \to \mathbb{R}$ is a function such that

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

A pair (X, d) where d is a metric on X is known as a metric space.

Proposition 4.2. A normed space $(X, \|\cdot\|)$ has a metric given by $d(x, y) = \|x - y\|$.

Proposition 4.3. Any metric space (X, d) is a topological space, with topology generated by the basis

$$\mathcal{B} = \{D(x; r) : x \in X, r > 0\}$$

Proposition 4.4. Any metric space is Hausdorff.

Definition 4.5 (Subspace)
For $Y \subseteq X$, $d|_{Y\times Y}$ defines a metric on Y.

Proposition 4.6. For any metric space (X, d), and $Y \subseteq X$, we have that

$$\begin{array}{ccc} (X, d) & \xrightarrow{\text{subspace metric}} & (Y, d|_{Y\times Y}) \\ \downarrow{\scriptstyle \text{metric topology}} & & \downarrow{\scriptstyle \text{metric topology}} \\ (X, \mathscr{T}_X) & \xrightarrow[\text{subspace topology}]{} & (Y, \mathscr{T}_Y) \end{array}$$

commutes.

Definition 4.7 (Product)
For $p \in [1, \infty]$,

$$d((x_1, y_1), (x_2, y_2)) = \left(d_X(x_1, x_2)^p + d_Y(y_1, y_2)^p\right)^{1/p}$$

defines a metric on $X \times Y$. We write the product as $X \oplus_p Y$.

Proposition 4.8. For a metric space (X, d), $x \in X$, we have that $x_n \to x$ if and only if

$$\forall \varepsilon > 0, \exists N \text{ s.t. } \forall n \geq N, d(x_n, x) < \varepsilon$$

Proposition 4.9. For metric spaces $(X, d_X), (Y, d_Y)$, $f : X \to Y$ is continuous if and only if

$$\forall x \in X, \forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. } d_X(x, y) < \delta \implies d_Y(x, y)$$

Definition 4.10 (Isometric)
$f : X \to Y$ is isometric if for all x, y, we have that

$$d(f(x), f(y)) = d(x, y)$$

Definition 4.11 (Lipschitz)
$f : X \to Y$ is Lipschitz with constant C if for all x, y, we have that

$$d(f(x), f(y)) \leq Cd(x, y)$$

Definition 4.12 (Uniformly continuous)
$f : X \to Y$ is uniformly continuous if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in M$,

$$d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon$$

Proposition 4.13. For fixed y, $x \mapsto (x, y)$ is an isometric map.

Proposition 4.14. Projection maps are 1-Lipschitz.

Proposition 4.15.

$$\text{Isometric} \implies \text{Lipschitz} \implies \text{Uniform continuous} \implies \text{Continuous}$$

Definition 4.16 (Equivalent metrics)
Two metrics d and d' are equivalent if they induce the same topology. Equivalently, if the maps

$$\mathrm{id} : (X, d) \to (X, d') \quad \text{and} \quad \mathrm{id} : (X, d') \to (X, d)$$

are continuous.

Definition 4.17 (Uniformly equivalent metrics)
Two metrics d and d' are uniformly equivalent if the maps

$$\mathrm{id} : (X, d) \to (X, d') \quad \text{and} \quad \mathrm{id} : (X, d') \to (X, d)$$

are uniformly continuous.

Definition 4.18 (Lipschitz equivalent metrics)
Two metrics d and d' are Lipschitz equivalent if the maps

$$\mathrm{id} : (X, d) \to (X, d') \quad \text{and} \quad \mathrm{id} : (X, d') \to (X, d)$$

are Lipschitz.

5 Uniform convergence

Definition 5.1 (Uniform metric)
For a set S, and a metric space (X, d), define a metric[a] d_∞ on the set of all functions $S \to \mathbb{R}$ by

$$d_\infty(f, g) = \sup_{x \in S} d(f(x), g(x))$$

[a]Strictly speaking this can also take value ∞

Definition 5.2 (Bounded functions)
Let S be a set, X be a metric space, then define

$$\ell_\infty(S, X) = \{f : S \to X : f \text{ bounded}\}$$

Definition 5.3 (Uniform convergence)
We say that a sequence f_n converges uniformly to f on S if $d_\infty(f_n, f) \to 0$. Equivalently,

$$\forall \varepsilon > 0, \exists N, \forall n \geq N, \forall x \in S, d(f_n(x), f(x)) < \varepsilon$$

Definition 5.4 (Pointwise convergence)
We say f_n converges to f pointwise on S if for all $x \in S$, $f_n(x) \to f(x)$.

Proposition 5.5. Uniform convergence implies pointwise convergence.

Proposition 5.6. Suppose S is a topological space. $f_n : S \to X$ for all n, and $f_n \to f$ uniformly on S. Suppose f_n is continuous. Then f is continuous.

Proof. Let $U \subseteq X$ be open, and let $t \in f^{-1}(U)$. Since $f(t) \in U$, there exists $\varepsilon > 0$ such that $D(f(t), 3\varepsilon) \subseteq U$. As $f_n \to f$ uniformly, there exists N such that for all $s \in S$, $d(f(x), f_N(x)) < \varepsilon$. Let $V = D(f_N(t), \varepsilon)$, and $f_N^{-1}(V)$ is open. Furthermore, $t \in f_N^{-1}(V)$, so suffices to show $f_N^{-1}(V) \subseteq f^{-1}(U)$.

Let $w \in f_N^{-1}(V)$. Then

$$d(f(w), f(t)) \leq d(f(w), f_N(w)) + d(f_N(w), f_N(t)) + d(f_N(t), f(t)) < 3\varepsilon$$

So $f(w) \in U$. □

Lemma 5.7. If $f_n \to f$ uniformly on S, and f_n is bounded for every n, then f is bounded.

Proof. Fix N such that for all $x \in S$, $d(f(x), f_N(x)) < 1$. Since f_N is bounded, there exists M, $z \in X$ such that $d(f(x), z) < M$ for all $x \in S$. Then for all $x \in S$, we have that

$$d(f(x), z) \leq d(f(x), f_N(x)) + d(f_N(x), z) < M + 1$$

□

Theorem 5.8. Suppose $f_n : [a, b] \to \mathbb{R}$ is integrable for every n, and $f_n \to f$ uniformly on $[a, b]$. Then f is integrable. Furthermore, we have that

$$\int_a^b f_n \to \int_a^b f$$

Proof. We have already shown that the uniform limit of bounded functions is bounded. Fix $\varepsilon > 0$, we have $N \in \mathbb{N}$ such that for all $x \in [a, b]$, $|f_N(x) - f(x)| < \varepsilon$. Since f_N is integrable, we have a dissection $\mathscr{D}$ such that

$$\mathcal{U}_{\mathscr{D}}(f) - \mathcal{L}_{\mathscr{D}}(f) < \varepsilon$$

For any $x, y \in [x_{k-1}, x_k]$ in $\mathscr{D}$, we have that

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < |f_N(x) - f_N(y)| + 2\varepsilon$$

Which means that

$$\sup_{x,y\in[x_{k-1},x_k]} |f(x) - f(y)| \leq \sup_{x,y\in[x_{k-1},x_k]} |f_N(x) - f_N(y)| + 2\varepsilon$$

So

$$\sum_{k=1}^{m}(x_k - x_{k-1}) \sup_{x,y\in[x_{k-1},x_k]} |f(x) - f(y)| \leq \mathcal{U}_{\mathscr{D}}(f) - \mathcal{L}_{\mathscr{D}}(f) + 2\varepsilon(b - 1) \leq (2(b - a) + 1)\varepsilon$$

Thus f is integrable. Furthermore, we have that

$$\left|\int_a^b f_n - \int_a^b f\right| \leq \int_a^b |f_n - f| \leq (b - a) \sup_{x\in[a,b]} |f_n(x) - f(x)| \to 0$$

□

Corollary 5.9. Let f_n be integrable for every n, and suppose $\sum f_n$ converges uniformly on $[a, b]$. Then $x \mapsto \sum f_n$ is integrable, with

$$\int_a^b \sum_n f_n(x)\mathrm{d}x = \sum_n \int_a^b f_n(x)\mathrm{d}x$$

Proof. For $x \in [a, b]$, $n \in N$, define $F_n(x) = \sum_{k=1}^{n} f_k(x)$, and $F(x) = \sum f_k(x)$. Then $F_n \to F$ uniformly on $[a, b]$. Each F_n is integrable, and result follows by above. □

Theorem 5.10. Let $f_n : [a, b] \to \mathbb{R}$ be C^1, and suppose

- $\sum f'_k$ converges uniformly on $[a, b]$
- There exists $c \in [a, b]$ such that $\sum f_k(c)$ converges.

Then $\sum f$ converges uniformly on $[a, b]$ to a continuously differentiable function f, and

$$f'(x) = \left(\sum_k f_k(x)\right)' = \sum_k f'_k(x)$$

Proof. Let $g(x) = \sum_k f'_k(x)$, and $\lambda = \sum f_k(c)$. Define

$$f(x) = \lambda + \int_c^x g(t)\mathrm{d}t$$

Since g is the uniform limit of continuous functions, it is continuous. Furthermore, by FTC, we have that $f' = g$ on $[a, b]$. So f' is continuous, with $f(c) = \lambda$.

Also by FTC, $f_k(x) = f_k(c) + \int_c^x f_k'(t)\mathrm{d}t$. Fix $\varepsilon > 0$. We have N such that

- $\left|\lambda - \sum_{k=1}^n f_k(c)\right| < \varepsilon$ for all $n \geq N$, and
- $\left|g(t) - \sum_{k=1}^n f_k'(t)\right| < \varepsilon$ for all $n \geq N$ and $t \in [a, b]$.

Then for $n \geq N$, $x \in [a, b]$, we have that

$$\begin{aligned}\left|f(x) - \sum_{k=1}^n f_k(x)\right| &= \left|\lambda + \int_c^x g(t)\mathrm{d}t - \sum_{k=1}^n \left(f_k(c) + \int_c^x f_k'(t)\mathrm{d}t\right)\right| \\ &\leq \left|\lambda - \sum_{k=1}^n f_k(c)\right| + \left|\int_x^c g(t) - \sum_{k=1}^n f_k'(t)\mathrm{d}t\right| \\ &\leq \varepsilon + |x - c|\varepsilon \\ &\leq (b - a + 1)\varepsilon\end{aligned}$$

□

Definition 5.11 (Uniformly Cauchy)
A sequence (f_n) of functions is uniformly Cauchy if

$$\forall \varepsilon > 0, \exists N, \forall n, m \geq N, \forall x \in S, d(f_n(x), f_m(x)) < \varepsilon$$

Theorem 5.12 (General principle of uniform convergence). Suppose X is a complete metric space, f_n is uniformly Cauchy on S. Then f_n converges uniformly on S.

Proof. Fix $x \in S$. Then $f_n(x)$ is a Cauchy sequence, so by completeness we have f such that $f_n(x) \to f(x)$. Then $f_n \to f$ pointwise.

Fix $\varepsilon > 0$, we have N such that for all $n, m \geq N$, $x \in S$, $d(f_n(x), f_m(x)) < \varepsilon$. Fix $x \in S$, $n \geq N$. Then for all $m \geq N$, $d(f_m(x), f_n(x)) < \varepsilon$, taking $m \to \infty$[1], we have that $d(f(x), f_n(x)) \leq \varepsilon$, so $f_n \to f$ uniformly. □

Theorem 5.13 (Weierstrass M-test). Let X be a complete normed space[a], $f_n : S \to X$. Assume for all $n \geq N$, there exists $M_n \geq 0$ such that $\|f_n(x)\| \leq M_n$ for all $x \in S$. Furthermore, assume $\sum M_n$ converges. Then $\sum f_n$ converges uniformly.

[a] a Banach space

Proof. Let $F_n(x) = \sum_{k=1}^n f_k(x)$. Fix $n > m$, then

$$\|F_n(x) - F_m(x)\| = \left\|\sum_{k=m+1}^n f_k(x)\right\| \leq \sum_{k=m+1}^n \|f_k(x)\| \leq \sum_{k=m+1}^n M_n$$

Thus, given $\varepsilon > 0$, we have N such that $\sum_{k=N+1}^\infty M_n < \varepsilon$, and for $m, n \geq N$, we have that $\|F_n(x) - F_m(x)\| < \varepsilon$. So F_n is uniformly Cauchy. □

Theorem 5.14. Suppose $\sum c_n z^n$ has radius of convergence R. Then for $0 \leq r < R$, the power series converges uniformly on $D(a; r)$.

[1] Strictly speaking we are assuming continuity of d.

Proof. Fix $w \in \mathbb{C}$ such that $r < |w| < R$. Since $\sum c_n w^n$ converges, $c_n w^n \to 0$, which means that $c_n w^n$ is bounded. Let $\rho = r/|w|$. Fix $z \in D(a; r)$ and $n \in \mathbb{N}$. Then

$$|c_n z^n| = |c_n w^n| \left| \frac{z}{w} \right|^n \leq M\rho^n$$

Since $\sum M\rho^n$ converges, by the M-test, $\sum c_n z^n$ converges uniformly on $D(a; r)$. □

Definition 5.15 (Local uniform convergence)
Let $U \subseteq \mathbb{C}$ be open, $f_n \to f$ locally uniformly on U if for all $w \in U$, there exists $D(w, \delta) \subseteq U$ such that $f_n \to f$ uniformly on $D(w, \delta)$.

Proposition 5.16. A power series converges locally uniformly within its radius of convergence.

6 Uniform continuity

Definition 6.1 (Uniform continuity)
Let X, Y be metric spaces. Then $f : X \to Y$ is uniformly continuous if

$$\forall \varepsilon > 0, \exists \delta > 0, d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon$$

Theorem 6.2. Let X be a compact metric space, Y be any metric space, $f : X \to Y$ continuous. Then f is uniformly continuous.

Proof. Given $\varepsilon > 0$, for all $x \in X$, by continuity we have δ_x such that $f(D(x, 2\delta_x)) \subseteq D(f(x), \varepsilon)$. Then $D(x, \delta_x)$ form an open cover of X, so we have a finite subcover $(D(x_i, \delta_{x_i}))_i$. Let $\delta = \min_i \delta_{x_i}$. Then for $x, y \in X$, $d(x, y) < \delta$, since we have a cover, we have x_i such that $d(x_i, x) < \delta_{x_i}$. Then

$$d(y, x_i) \leq d(y, x) + d(x, x_i) < \delta + \delta_{x_i} \leq 2\delta_{x_i}$$

Hence $f(x), f(y) \in D(f(x_i), \varepsilon)$, so $d(f(x), d(y)) < 2\varepsilon$. □

Corollary 6.3. A continuous function $f : [a, b] \to \mathbb{R}$ is integrable.

Proof. See IA Analysis. □

7 Completeness

Definition 7.1 (Cauchy sequence)
Let X be a metric space, a sequence x_n in X is Cauchy if

$$\forall \varepsilon > 0, \exists N, \forall n, m \geq N, d(x_n, x_m) < \varepsilon$$

Definition 7.2 (Complete metric space)

A metric space is complete if all Cauchy sequences converge in M.

Proposition 7.3. If X and Y are complete metric spaces, then so is $X \oplus_p Y$.

Proof. Let (x_n, y_n) be a Cauchy sequence in $X \oplus_p Y$. Since a sequence in the product converges if and only if each component converges, suffices to show that x_n is Cauchy. But this is immediate since

$$d(x_n, x_m) \leq d((x_n, y_n), (x_m, y_m))$$

□

Proposition 7.4. Let X be a metric space, $Y \subseteq X$. If Y is complete, then Y is closed in X.

Proposition 7.5. Let X be a complete metric space, $Y \subseteq X$ closed. Then Y is complete.

Theorem 7.6. Let S be any set, X be a complete metric space. Then $\ell_\infty(S, X)$ is complete with respect to the uniform metric.

Proof. Let f_n be a Cauchy sequence. Then it is uniformly Cauchy, so converges uniformly to say $f : S \to X$. The uniform limit of bounded functions is bounded, so $f \in \ell_\infty(S, X)$. □

Definition 7.7 (Continuous bounded functions)
Let X be a topological space, Y be a metric space. Then define

$$C_B(X, Y) = \{f : X \to Y : f \text{ continuous, bounded}\} \leq \ell_\infty(X, Y)$$

Theorem 7.8. Let X be a topological space, Y be a complete metric space. Then $C_B(X, Y)$ is complete.

Proof. Suffices to show it is closed in $\ell_\infty(X, Y)$. This follows as the uniform limit of continuous functions is continuous. □

8 Sequential compactness

Definition 8.1 (Sequential compactness)
A topological space X is sequentially compact if every sequence has a convergent subsequence.

Definition 8.2 (Net)
For a metric space X, a subset $F \subseteq X$ is an ε-net for X if

$$X = \bigcup_{y \in F} \overline{D}(y, \varepsilon)$$

Definition 8.3 (Totally bounded)

A metric space X is totally bounded if there exists a finite ε-net for all $\varepsilon > 0$.

Definition 8.4 (Diameter)

The diameter of $A \subseteq X$ is

$$\operatorname{diam}(A) = \sup_{x,y \in A} d(x, y) \in [0, \infty]$$

Lemma 8.5. Assume that X is totally bounded, $A \subseteq M$ nonempty closed, and $\varepsilon > 0$. Then there exists closed sets $B_1, \ldots, B_k$ such that $A = B_1 \cup \cdots \cup B_k$, and $\operatorname{diam}(B_i) \leq \varepsilon$ for all i.

Proof. Let F be a finite $\varepsilon/2$-net. Then $\left(A \cap \overline{D}(x, \varepsilon/2) : x \in F, A \cap \overline{D}(x, \varepsilon/2) \neq \varnothing\right)$ works. ☐

Theorem 8.6. For a metric space X, the following are equivalent.

(i) X is compact.

(ii) X is sequentially compact.

(iii) X is complete and totally bounded.

Proof. (i) $\implies$ (ii). First suppose X is compact, and let (x_n) be a sequence in X. Let $T_n = \{x_k : k > n\}$, then the limit of any subsequence must be in $T = \bigcap_n \overline{T_n}$, which is nonempty by compactness. Let $x \in T$. Then $x \in \overline{T_1}$ means that we have $k_1 > 1$ such that $x_{k_1} \in T_1 \cap D(x; 1)$, and $x \in \overline{T_{k_1}}$ means we have $k_2 > k_1$ such that $x_{k_2} \in T_{k_1} \cap D(x; 1/2)$, and so on. This gives us $k_1 < k_2 < \ldots$ such that $d(x, x_{k_n}) < 1/n$, so $x_{k_n} \to x$.

(ii) $\implies$ (iii). Now suppose X is sequentially compact. Let (x_n) be a Cauchy sequence in X, then it has a convergent subsequence, so converges. So X is complete.

Now suppose if X is not totally bounded. Let ε be such that there is no finite ε-net. Fix $x_1 \in X$. Then we have x_2 such that $x_2 \notin \overline{D}(x_1, \varepsilon)$. More generally, we always have

$$x_{n+1} \notin \bigcup_{i=1}^{n} \overline{D}(x_i, \varepsilon)$$

Then this sequence has no Cauchy subsequence, and so cannot have a convergent subsequence.

(iii) $\implies$ (i). Let $\mathscr{U}$ be an open cover for X, and suppose $\mathscr{U}$ does not finitely cover X. Let $A_0 = X$. Then by lemma, we have closed sets $B_1, \ldots, B_k$ such that $A_0 = B_1 \cup \cdots \cup B_k$ and $\operatorname{diam}(B_i) < 1$. These can't all be finitely covered, without loss of generality assume B_k cannot be finitely covered. Then set $A_1 = B_k$. Inductively, we have $A_1 \supseteq A_2 \supseteq \ldots$ such that each A_i is closed, $\operatorname{diam}(A_n) < 2^{-n}$, and A_i is not finitely covered by $\mathscr{U}$.

For each n, choose $x_n \in A_n$. Then this is a Cauchy sequence, so by completeness converges to say x. Now choose $U \in \mathscr{U}$ such that $x \in U$. Since U is open, we must have some n such that $D(x; 1/n) \subseteq U$. But then $\mathscr{U}$ finitely covers (say) A_n. Contradiction. ☐

9 Contraction mapping

Definition 9.1 (Contraction map)

Let X, Y be metric spaces. Then $f : X \to Y$ is a contraction map if f is λ-Lipschitz for some $0 \leq \lambda < 1$.

Theorem 9.2 (Contraction mapping theorem, Banach's fixed point theorem)**.** Let X be a nonempty complete metric space, $f : M \to M$ is a contraction map. Then f has a unique fixed point.

Proof. Suppose f is λ-Lipschitz, for some $0 \leq \lambda < 1$. Fix $x_0 \in X$, and define $x_n = f^n(x_0)$. Inductively, we have that

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \lambda d(x_{n-1}, x_n) \leq \cdots \leq \lambda^n d(x_0, x_1)$$

and for $m \geq n$, we have that

$$d(x_m, x_n) \leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}) \leq \sum_{k=n}^{m-1} \lambda^n d(x_0, x_1) \leq \frac{\lambda^n}{1-\Lambda} d(x_0, x_1) \to 0 \quad \text{as} \quad n \to \infty$$

So (x_n) is Cauchy, and by completeness, $x_n \to z$ for some $z \in X$. By continuity, $f(x_n) \to f(z)$, but $f(x_n) = x_{n+1} \to z$, so by uniqueness of limits, we must have that $f(z) = z$.

For uniqueness, suppose z and w are fixed points. Then

$$d(z, w) = d(f(z), f(w)) \leq \lambda d(z, w)$$

So we must have $d(z, w) = 0$, i.e. $z = w$. □

Lemma 9.3. For $f : \mathbb{R} \to \mathbb{R}^n$ integrable (say componentwise), we have that

$$\left\| \int_a^b f(t) \mathrm{d}t \right\| \leq \int_a^b \left\| f(t) \right\| \mathrm{d}t \leq (b-a) \sup_{t \in [c,d]} \left\| f(t) \right\|$$

Theorem 9.4 (Picard-Lindelöf)**.** Suppose we have $y_0 \in \mathbb{R}^n$ and $R > 0$, and

$$\phi : [a, b] \times \overline{D}(y_0; R) \to \mathbb{R}^n$$

where $\phi(t, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is k-Lipschitz for all t. Then there exists $\varepsilon > 0$ such that for all $t_0 \in [a, b]$, the initial value problem

$$f'(t) = \phi(t, f(t)) \quad \text{with initial value} \quad f(t_0) = y_0$$

has a unique solution on $[c, d] = [t_0 - \varepsilon, t_0 + \varepsilon] \cap [a, b]$.

Proof. ϕ is a continuous function from a compact set, so it is bounded. Let $C = \sup_{t \in [a,b], x \in \overline{D}(y_0;R)} \left\| \phi(t, x) \right\|$. Let $\varepsilon = \min\left(\frac{R}{c}, \frac{\delta}{k}\right)$ for any $\delta \in (0, 1)$. Fix $t_0 \in [a, b]$, and let $[c, d] = [t_0 - \varepsilon, t_0 + \varepsilon] \cap [a, b]$.

Now let $X = C([c, d], \overline{D}(y_0; R))$. This is complete as $\overline{D}(y_0; R)$ is a compact metric space. Define $T : X \to X$ by

$$Tg(t) = y_0 + \int_{t_0}^{t} \phi(s, g(s)) \mathrm{d}s$$

We first need to check that T is well defined. For any $t \in [c, d]$, we have that

$$\left\| Tg(t) - y_0 \right\| = \left\| \int_{t_0}^{t} \phi(s, g(s)) \mathrm{d}x \right\| \leq |t - t_0| \sup_{s \in [t_0, t]} \left\| \phi(s, g(s)) \right\| \leq \varepsilon C \leq R$$

So $Tg \in M$. Now we will show that T is a contraction. Let $g, h \in M$. For $t \in [c, d]$, we have that

$$\begin{aligned}\|Tg(t) - Th(t)\| &\le \left\|\int_{t_0}^{t} \phi(s, g(s)) - \phi(s, h(s))ds\right\| \\ &\le |t - t_0| \sup_{s\in[t_0,t]} \|\phi(s, g(s)) - \phi(s, h(s))\| \\ &\le \varepsilon k\|g - h\|_\infty \\ &\le \delta\|g - h\|_\infty\end{aligned}$$

So $\|Tg - Th\|_\infty \le \delta\|g - h\|_\infty$. Thus, by the contraction mapping theorem, T has a unique fixed point, say f. Then

$$f(t) = y_0 + \int_{t_0}^{t} phi(s, f(s))ds$$

and the fundamental theorem of calculus shows that this satisfies the differential equation. □

10 Differentiation

Definition 10.1 (Matrix norm)
For $T \in L(\mathbb{R}^m, \mathbb{R}^n)$, we can identify T with a matrix (T_{ij}) with respect to the standard bases. Then we define the (elementwise) norm as

$$\|T\| = \left(\sum_{j=1}^{m}\sum_{i=1}^{n} T_{ij}^2\right)^{1/2} = \left(\sum_{j=1}^{m} \|Te_i\|^2\right)^{1/2}$$

Lemma 10.2. For $T \in L(\mathbb{R}^m, \mathbb{R}^n)$, $x \in \mathbb{R}^m$, we have that

$$\|Tx\| \le \|T\| ||x||$$

Proof.

$$\begin{aligned}\|Tx\| &= \left\|\sum_{i=1}^{m} x_i Te_i\right\| \\ &\le \sum_{i=1}^{m} |x_i|\|Te_i\| \quad \text{by Triangle-Ineq.} \\ &\le \left(\sum_{i=1}^{m} x_i^2\right)^{1/2} \left(\sum_{i=1}^{m} \|Te_i\|^2\right)^{1/2} \quad \text{by Cauchy-Schwarz} \\ &= \|T\| ||x||\end{aligned}$$

□

Corollary 10.3. Linear maps are Lipschitz, hence (uniform) continuous.

Lemma 10.4. For $S \in L(\mathbb{R}^n, \mathbb{R}^p)$, $T \in L(\mathbb{R}^m, \mathbb{R}^n)$, we have that

$$\|ST\| \le \|S\|\|T\|$$

Proof.

$$\begin{aligned}\|ST\| &= \left(\sum_{i=1}^{m} \|STe_i\|^2\right)^{1/2} \\ &\le \left(\sum_{i=1}^{m} \|S\|^2 \|Te_i\|^2\right)^{1/2} \\ &= \|S\|\|T\|\end{aligned}$$

□

Definition 10.5 (Differentiable)
Let $f : \mathbb{R}^m \to \mathbb{R}^n$, $a \in \mathbb{R}^m$. We say f is differentiable at a, with derivative $Df(a) = f'(a) = T \in L(\mathbb{R}^m, \mathbb{R}^n)$ if

$$\frac{f(a+h) - f(a) - T(h)}{\|h\|} \to 0 \quad \text{as} \quad h \to 0$$

Proposition 10.6. f is differentiable if and only if there exists $T \in L(\mathbb{R}^m, \mathbb{R}^n)$, $\varepsilon : \mathbb{R}^m \to \mathbb{R}^n$ continuous, $\varepsilon(0) = 0$ such that

$$f(a+h) = f(a) + T(h) + \|h\|\varepsilon(h)$$

Proposition 10.7. The derivative, if it exists, is unique.

Proof. Suppose we have S and T both satisfying the equation in the definition of a derivative. Fix $x \in \mathbb{R}^m$ nonzero. Then for all $k \in \mathbb{R}$, we have that $x/k \to 0$, so

$$\frac{S(x) - T(x)}{||x||} = \frac{S(x/k) - T(x/k)}{\|x/k\|} \to 0$$

□

Proposition 10.8. Suppose $f \in L(\mathbb{R}^m, \mathbb{R}^n)$. Then f is differentiable, with $Df = f$.

Proposition 10.9. Suppose $f \in \text{Bil}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p)$. Then f is differentiable, with

$$Df(a,b)(h,k) = f(a,k) + f(h,b)$$

Definition 10.10 (Differentiable in a set)
Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^n$, $a \in U$. Then f is differentiable at a, with derivative $Df = f'(a) = T \in L(\mathbb{R}^m, \mathbb{R}^n)$ if

$$\frac{f(a+h) - f(a) - T(h)}{\|h\|} \to 0 \quad \text{as} \quad h \to 0$$

whenever $a + h \in U$.

Remark 10.11. For $m = 1$, we note that $L(\mathbb{R}, \mathbb{R}^n) = (\mathbb{R}^n)^* \cong \mathbb{R}^n$, so in this case we can write $Df(a)(h) = v \cdot h$ for some $v \in \mathbb{R}^n$.

Proposition 10.12. Let $f : U \to R^n$ be differentiable at a. Then f is continuous at a.

Proof. For $x \in U$, we have that

$$f(x) = f(a) + f'(a)(x - a) + \|x - a\|\varepsilon(x - a)$$

which is sum/product/composition of continuous function, hence continuous (at a). ☐

Proposition 10.13 (Chain rule)**.** Let $f : U \to \mathbb{R}^n$, $g : V \to \mathbb{R}^p$, $f(V) \subseteq V$, and $a \in U$. Suppose that f is differentiable at a, and g is differentiable at $b = f(a)$. Then $g \circ f$ is differentiable at a, with

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

Proof. Let $S = Df(a)$, $T = Dg(f(a))$. Then we have that

$$f(a + h) = f(a) + S(h) + \|h\|\varepsilon(h) \quad \text{and} \quad g(b + k) = g(b) + T(k) + \|k\|\zeta(k)$$

So we get that

$$\begin{aligned}
(g \circ f)(a + h) &= g(f(a + h)) \\
&= g(f(a) + S(h) + \|h\|\varepsilon(h)) \\
&= g(f(a)) + T(S(h) + \|h\|\varepsilon(h)) + \|S(h) + \|h\|\varepsilon(h)\|\zeta(S(h) + \|h\|\varepsilon(h)) \\
&= g(f(a)) + (T \circ S)(h) + \underbrace{\|h\|T(\varepsilon(h)) + \|k(h)\|\zeta(k(h))}_{=\eta(h)}
\end{aligned}$$

Suffices to show that $\eta(h)/\|h\| \to 0$ as $h \to 0$. Since T is continuous, $\|h\|T(\varepsilon(h))/\|h\| = T(\varepsilon(h)) \to 0$ as $h \to 0$.

$$\frac{\|k\|}{\|h\|} \le \frac{\|S(h)\| + \|h\|\|\varepsilon(h)\|}{\|h\|} \le \frac{\|S\|\|h\| + \|h\|\|\varepsilon(h)\|}{\|h\|} = \|S\| + \|\varepsilon(h)\|$$

is bounded as $h \to 0$, $k(h) = S(h) + \|h\|\varepsilon(h) \to 0$ as $h \to 0$, so we have that $\zeta(k(h)) \to 0$ as $h \to 0$, which means that $\eta(h)/\|h\| \to 0$ as $h \to 0$. ☐

Proposition 10.14 (Components of derivatives)**.** Let $f : U \to \mathbb{R}^n$, $a \in U$. Let $f_j = \pi_j \circ f$ be the j-th component of f. Then f is differentiable at a if and only if each f_j is differentiable at a, with

$$Df(a)(h) = \sum_{j=1}^{n} Df_j(a)(h)e_j$$

Equivalently,

$$\pi_j(Df(a)) = D(\pi_j \circ f)(a)$$

Proof. First suppose f is differentiable. π_j is linear, so differentiable. The chain rule gives the required result.

Conversely, suppose each f_j is differentiable, with

$$f_j(a + h) = f_j(a) + Df_j(a)(h) + \|h\|\varepsilon_j(h)$$

Then

$$f(a+h) = f(a) + \sum_{j=1}^{n} Df_j(a)(h)e_j + \|h\| \sum_{j=1}^{n} \varepsilon_j(h)e_j$$

and we have that $\sum_{j=1}^{n} \varepsilon_j(h)e_j \to 0$ as $h \to 0$. □

Proposition 10.15 (Linearity of derivative)**.** For $\lambda, \mu \in \mathbb{R}$, $f, g : U \to \mathbb{R}^n$ differentiable at $a \in U$, we have that

$$D(\lambda f + \mu g)(a) = \lambda Df(a) + \mu Dg(a)$$

Proposition 10.16 (Product rule)**.** Let $f : U \to \mathbb{R}^n$ and $\phi : U \to \mathbb{R}$ be differentiable at a. Then

$$D(\phi f)(a)(h) = \phi(a)[Df(a)(h)] + [D\phi(a)(h)]f(a)$$

Proof. Define $F : U \to \mathbb{R} \times \mathbb{R}^n$ by $F(x) = (\phi(x), f(x))$ and $G : \mathbb{R} \times \mathbb{R}^n$ by $G(\lambda, v) = \lambda v$. Then $\phi f = G \circ F$. $DF = (D\phi, Df)$ by projections, and $DG(\lambda, v)(\mu, w) = \lambda w + \mu v$ by bilinearity. The result follows by chain rule. □

Definition 10.17 (Directional derivative)

Let $U \subseteq \mathbb{R}^m$ be open, $f : U \to \mathbb{R}^n$, $a \in U$. Fix $u \in \mathbb{R}^m \setminus \{0\}$. If

$$\lim_{t \to 0} \frac{f(a+tu) - f(a)}{t}$$

exists, we call this the directional derivative of f at a, and write $D_u f(a)$ for the limit.

Definition 10.18 (Partial derivative)

We write $D_i = D_{e_i}$ for the i-th partial derivative of a function.

Proposition 10.19. If f is differentiable at a, then for all u, $D_u f(a)$ exists, with

$$D_u f(a) = Df(a)(u)$$

Proof. Suppose $f(a+h) = f(a) + Df(a)(h) + \|h\|\varepsilon(h)$. Then

$$\frac{f(a+tu) - f(a)}{t} = Df(a)(u) + \frac{|t|}{t}\|u\|\varepsilon tu \to Df(a)(u)$$

So $D_u f(a) = Df(a)(u)$. □

Corollary 10.20.

$$Df(a)(h) = \sum_{i=1}^{m} h_i D_i f(a)$$

Definition 10.21 (Jacobian)
The Jacobian matrix of a function f is $Jf(a) = [Df(a)]$, i.e. Df with respect to standard bases.

Proposition 10.22. The i-th column of $Jf(a)$ is $D_i f(a)$, and

$$(Jf(a))_{ij} = \frac{\partial f_j}{\partial x_i}$$

Theorem 10.23. Suppose $D_i f(x)$ exists for $x \in V$, where $V \subseteq U$ is an open neighbourhood of a. Moreover, $x \mapsto D_i f(x)$ is continuous at a. Then f is differentiable at a.

Proof. Considering components of f, without loss of generality $f : \mathbb{R}^m \to \mathbb{R}$. We will prove this by induction on m. The case $m = 1$ is trivial. For $h \in \mathbb{R}^m$, define

$$h^{(1)} = \sum_{i=1}^{m-1} h_i e_i \quad \text{and} \quad h^{(2)} = h_m e_m$$

Let

$$\psi(h) = f(a+h) - f(a) + \sum_{i=1}^{m} h_i D_i f(a)$$

We wish to show that $\psi(h)$ is $o(h)$. We have that

$$\psi(h) = \underbrace{f(a+h) - f(a+h^{(1)}) - h_m D_m f(a)}_{(i)} + \underbrace{f(a+h^{(1)}) - f(a) - \sum_{i=1}^{m-1} h_i D_i f(a)}_{(ii)}$$

By the inductive hypothesis, (ii) is $o(\|h^{(1)}\|)$ so $o(\|h\|)$. Now let

$$\phi(t) = f(a + h^{(1)} + th^{(2)})$$

ϕ is differentiable, with $\phi'(t) = h_m D_m f(a + h^{(1)} + h^{(2)})$. Then by the mean value theorem, we have $t \in (0,1)$ such that $\phi(1) - \phi(0) = \phi'(t)$. Which means that

$$(i) = \phi(1) - \phi(0) - h_m D_m f(a) = h_m(D_m f(a + h^{(1)} + th^{(2)}) - D_m f(a))$$

As $h \to 0$, the part in the brackets tends to zero by continuity. So (i) is $o(|h_m|)$, and hence $o(\|h\|)$. □

Theorem 10.24 (Mean value inequality)**.** Suppose $f : U \to \mathbb{R}^n$ differentiable, and M is such that $\|f'(z)\| \leq M$ for all $z \in [a,b] \subseteq U$. then

$$\|f(b) - f(a)\| \leq M\|b - a\|$$

Proof. Let $u = b - a$, $v = f(b) - f(a)$. Without loss of generality, $u \neq 0$. Then define $\gamma(t) = a + tu$. $f \circ \gamma$ is differentiable, with

$$D(f \circ \gamma)(t) = Df(\gamma(t))(D\gamma(t)) = Df(a + tu)(u)$$

Furthermore, let $\phi(t) = \langle f(\gamma(t)), v \rangle$. Then $\|f(b) - f(a)\|^2 = \phi(1) - \phi(0)$, and $\phi'(t) = \langle Df(a + tu)(u), v \rangle$. Then by the mean value theorem, we have $\theta \in (0,1)$ such that $\phi(1) - \phi(0) = \phi'(\theta)$. So

$$
\begin{aligned}
\left\|f(b)-f(a)\right\|^2 &= \phi(1)-\phi(0) \\
&= \phi'(\theta) \\
&= \langle Df(a+\theta u)(u), v\rangle \\
&\le \left\|Df(a+\theta u)(u)\right\| \|v\| \\
&\le \left\|Df(a+\theta u)\right\| \|u\| \|v\| \\
&\le M\left\|b-a\right\|\left\|f(b)-f(a)\right\|
\end{aligned}
$$

□

Corollary 10.25. Let U be open, $f : U \to \mathbb{R}^n$ differentiable, $Df = 0$ on U. Then U is locally constant.

Proof. For any $x \in U$, we have a neighbourhood $D(x; r) \subseteq U$. Then for any $y \in D(x; r)$, the segment $[x, y]$ is contained in U. Applying the mean value inequality we get the required result. □

Corollary 10.26. Suppose U is open and connected, $f : U \to \mathbb{R}^n$ differentiable, $Df = 0$ on U. Then U is constant.

Proof. A locally constant function on a connected space is constant. □

Proposition 10.27. Let $f : V \to W$ be a bijection, f differentiable at a and f^{-1} differentiable at $f(a)$, where $V \subseteq \mathbb{R}^m$ and $W \subseteq \mathbb{R}^n$. Then $m = n$.

Proof. Let $S = Df(a)$, and $T = D(f^{-1})(f(a))$. Then by the chain rule, $TS = \mathrm{id}_{\mathbb{R}^m}$ and $ST = \mathrm{id}_{\mathbb{R}^n}$. So $m = \mathrm{tr}(TS) = \mathrm{tr}(ST) = n$. □

Theorem 10.28 (Inverse function theorem)**.** Let $U \subseteq \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^n$ is C^1, $a \in U$, $f'(a)$ is invertible. Then there exists open neighbourhoods V, W of $a, f(a)$ respectively, such that $f|_V : V \to W$ is a bijection, with inverse $g : W \to V$ that is C^1.

Furthermore, $Dg(y) = (Df(g(y)))^{-1}$.

Proof. Let $T = Df(a)$, $h(x) = T^{-1}(f(a+x) - f(a))$. By the chain rule, h is differentiable, with $h'(x) = T^{-1} \circ Df(a+x)$, which is a composition of continuous functions, so continuous. Furthermore, we have that $h(0) = 0$, and $Dh(0) = \mathrm{id}$. Since $f(x) = T(h(x-a)) + f(a)$, suffices to prove the result for h. So without loss of generality, we may assume $a = f(a) = 0$, and $Df(0) = \mathrm{id}$.

Since Df is continuous, we have $r > 0$ such that $\overline{D}(0; r) \subseteq U$ and for all $x \in \overline{D}(0; r)$, $\left\|Df(x) - \mathrm{id}\right\| \le \frac{1}{2}$. For $x \in \overline{D}(0; r)$, let $p(x) = f(x) - x$. Then $Dp(x) = Df(x) - \mathrm{id}$. So for all $x \in \overline{D}(0; r)$, $\left\|Dp(x)\right\| \le \frac{1}{2}$. So by mean value inequality, $\left\|Dp(x) - Dp(y)\right\| \le \frac{1}{2}\left\|x - y\right\|$. As a result,

$$
\left\|f(x)-f(y)\right\| = \left\|p(x)-p(y)+x-y\right\| \ge \left\|x-y\right\| - \left\|p(x)-p(y)\right\| \ge \frac{1}{2}\left\|x-y\right\|
$$

Let $W = D(0; \frac{r}{2})$ and fix $w \in W$. Let $q(x) = w - f(x) + x = w - p(x)$. Since $p(0) = f(0) = 0$, we have that

$$
\left\|q(x)\right\| \le \|w\| + \left\|p(x)\right\| \le \|w\| + \frac{1}{2}\|x\| < 2 \cdot \frac{r}{2} = r
$$

So $q(\overline{D}(0; r)) \subseteq \overline{D}(0; r)$. Furthermore, we have that

$$
\left\|q(x)-q(y)\right\| = \left\|p(x)-p(y)\right\| \le \frac{1}{2}\left\|x-y\right\|
$$

So q is a contraction mapping on the complete metric space $\overline{D}(0;r)$, and has a unique fixed point. That is, a unique x such that $f(x) = w$. Since w is arbitrary, we have that $D(0; r/2) \subseteq f(D(0; r))$.

Let $V = f^{-1}(W) \cap D(0; r)$. Then V and W satisfy the requirements of the theorem, as V, W open, $f|_V : V \to W$ is a bijection. Let g be the inverse, we will show that g is continuous.

$$\|g(u) - g(v)\| \leq 2\|f(g(u)) - f(g(v))\| = 2\|u - v\|$$

So g is Lipschitz, hence continuous. The proof that g is C^1 is non-examinable. □

10.1 Second derivative

Definition 10.29 (Second derivative)
Let $f : U \to \mathbb{R}^n$ be differentiable on V open, where $a \in V \subseteq U$. Then we say that f is twice differentiable at a if

$$Df : V \to L(\mathbb{R}^m, \mathbb{R}^n)$$

is differentiable at a. We write the result as $f''(a) = D^2 f(a)$, where

$$D^2 f : V \to L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n))$$

Remark 10.30.

$$L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n)) \cong \mathrm{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$$

via $T(h)(k) \leftrightarrow T(h, k)$, so we do not make a distinction between the two.

Proposition 10.31. Let $f : U \to \mathbb{R}^n$ be differentiable on V open, where $a \in V \subseteq U$. Then f is twice differentiable at a if and only if there exists $T \in \mathrm{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ such that

$$Df(a + h)(k) = Df(a)(k) + T(h, k) + o(\|h\|)$$

for each fixed k.

Proof. Assume f is twice differentiable at a. Then

$$Df(a + h) = Df(a) + D^2 f(a)(h) + \|h\|\varepsilon(h)$$

Fix $k \in \mathbb{R}^m$, and evaluating the above at k, we find that

$$Df(a + h)(k) = Df(a)(k) + D^2 f(a)(h, k) + \|h\|\varepsilon(h)(k)$$

Letting $T = D^2 f(a)$, $\|\varepsilon(h)(k)\| \leq \|\varepsilon(h)\|\|k\| \to 0$ as $h \to 0$, so the error is $o(\|h\|)$.
Now suppose T exists. Let

$$\varepsilon(h) = \frac{Df(a + h) - Df(a) - T(h)}{\|h\|}$$

Suffices to show that $\varepsilon(h) \to 0$ as $h \to 0$. For each fixed $k \in \mathbb{R}^m$, $\varepsilon(h)(k) \to 0$ as $h \to 0$. So we have that

$$\|\varepsilon(h)\| = \left(\sum_{i=1}^{m} \|\varepsilon(h)(e_i)\|^2 \right)^{1/2} \to 0 \quad \text{as} \quad h \to 0$$

□

Proposition 10.32. If $f : \mathbb{R}^m \to \mathbb{R}^n$ linear, then $D^2 f \equiv 0$.

Proposition 10.33. If $f \in \text{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$, then $D^2 f = Df$ (up to identifying linear/bilinear forms).

Proof.

$$Df(a, b) = ((h, k) \mapsto f(a, k) + f(h, b))$$

is linear in (a, b). □

10.2 Partial derivatives

Proposition 10.34. Suppose $f : U \to \mathbb{R}^n$ is twice differentiable at a. Then

$$D_u D_v f(a) = D^2 f(a)(u, v)$$

Proof. For each fixed $k \in \mathbb{R}^m$, we have

$$Df(a + h)(k) = Df(a)(k) + D^2 f(a)(h, k) + o(\|h\|)$$

Putting $k = v$, we get that

$$D_v f(a + h) = D_v f(a) + D^2 f(a)(h, v) + o(\|h\|)$$

Which then gives us that $D_v f : V \to \mathbb{R}^n$ is differentiable at a, with

$$D(D_v f)(a)(h) = D^2(a)(h, v)$$

Seeting $h = u$ gives the required result. □

Theorem 10.35 (Symmetry of mixed partial derivatives). Suppose $f : U \to \mathbb{R}^n$ is twice differentiable, with $D^2 f : V \to \text{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ continuous at $a \in V \subseteq U$. Then

$$D_u D_v f(a) = D_v D_u f(a)$$

Equivalently, $D^2 f(a)$ is a symmetric bilinear map.

Proof. Since

$$(D_u f)_j(x) = (D_u f(x))_j = (Df(x)(u))_j = Df_j(x)(u) = D_u f_j(x)$$

without loss of generality $n = 1$. Define

$$\phi(s, t) = f(a + su + tv) - f(a + tv) - f(a + su) + f(a)$$

Fix s, t. Define $\psi(y) = f(a + yu + tv) - f(a + yu)$. Then $\phi(s, t) = \psi(s) - \psi(0)$. By the mean value theorem, we have $\alpha \in (0, 1)$ such that

$$\phi(s, t) = \psi(s) - \psi(0) = s\psi'(\alpha s) = s\,(D_u f(a + \alpha su + tv) - D_u f(a + \alpha su))$$

Apply the mean value theorem to $z \mapsto D_u f(a + \alpha su + zv)$, we have some $\beta \in (0, 1)$ such that

$$\phi(s, t) = st D_v D_u f(a + \alpha su + \beta tv) = st D^2 f(a + \alpha su + \beta tv)(v, u)$$

By continuity, we have that

$$\frac{\phi(s, t)}{st} = D^2 f(a + \alpha su + \beta tv)(v, u) \to D^2 f(a)(v, u) \quad \text{as} \quad s, t \to 0$$

If instead we used $\tilde{\psi}(y) = f(a + su + yv) - f(a + yv)$, then we would get

$$\frac{\phi(s, t)}{st} \to D^2 f(u, v) \quad \text{as} \quad s, t \to 0$$

Uniqueness of limits gives the required result. □