MA3101

Analysis III

Autumn 2021

Satvik Saha
19MS154

*Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.*

Contents

1 Euclidean spaces

1.1 $\mathbb{R}^n$ as a vector space

We are familiar with the vector space $\mathbb{R}^n$, with the standard inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + \cdots + x_n y_n.$$

The standard norm is defined as

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle = \sum_{k=1}^{n} (x_i - y_i)^2.$$

Exercise 1.1. What are all possible inner products on $\mathbb{R}^n$?

Solution. Note that an inner product is a bilinear, symmetric map such that $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$. Thus, an product map on $\mathbb{R}^n$ is completely and uniquely determined by the values $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = a_{ij}$. Let A be the $n \times n$ matrix with entries a_{ij}. Note that A is a real symmetric matrix with positive entries. Now,

$$\langle \boldsymbol{x}, \boldsymbol{e}_j \rangle = x_1 a_{1j} + \cdots + x_n a_{nj} = \boldsymbol{x}^\top \boldsymbol{a}_j,$$

where $\boldsymbol{a}_j$ is the j^{th} column of A. Thus,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{a}_1 y_1 + \cdots + \boldsymbol{x}^\top \boldsymbol{a}_n y_n = \boldsymbol{x}^\top A \boldsymbol{y}.$$

Furthermore, any choice of real symmetric A with positive entries produces an inner product.

Theorem 1.1 (Cauchy-Schwarz)**.** *Given two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$, we have*

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \|\boldsymbol{v}\| \|\boldsymbol{w}\|.$$

Proof. This is trivial when $\boldsymbol{w} = \boldsymbol{0}$. When $\boldsymbol{w} \neq \boldsymbol{0}$, set $\lambda = \langle \boldsymbol{v}, \boldsymbol{w} \rangle / \|\boldsymbol{w}\|^2$. Thus,

$$0 \leq \|\boldsymbol{v} - \lambda \boldsymbol{w}\|^2 = \|\boldsymbol{v}\|^2 - 2\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \lambda^2 \|\boldsymbol{w}\|^2.$$

Simplifying,

$$0 \leq \|\boldsymbol{v}\|^2 - \frac{|\langle \boldsymbol{v}, \boldsymbol{w} \rangle|^2}{\|\boldsymbol{w}\|^2}.$$

This gives the desired result. Clearly, equality holds if and only if $\boldsymbol{v} = \lambda \boldsymbol{w}$. □

Theorem 1.2 (Triangle inequality)**.** *Given two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$, we have*

$$\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|.$$

Proof. Write

$$\|\boldsymbol{v} + \boldsymbol{w}\|^2 = \|\boldsymbol{v}\|^2 + 2\langle \boldsymbol{v}, \boldsymbol{w} \rangle + \|\boldsymbol{w}\|^2 \leq \|\boldsymbol{v}\|^2 + 2|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| + \|\boldsymbol{w}\|^2.$$

Applying Cauchy-Schwarz gives

$$\|\boldsymbol{v} + \boldsymbol{w}\|^2 \leq (\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2.$$

Equality holds if and only if $\boldsymbol{v} = \lambda \boldsymbol{w}$ for $\lambda \geq 0$. □

1.2 $\mathbb{R}^n$ as a metric space

Our previous observations allow us to define the standard metric on $\mathbb{R}^n$, seen as a point set.

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|.$$

Definition 1.1. For any $\delta > 0$, the set

$$B_\delta(\boldsymbol{x}) = \{\boldsymbol{y} \in \mathbb{R}^n : d(\boldsymbol{x}, \boldsymbol{y}) < \delta\}$$

is called the open ball centred at $\boldsymbol{x} \in \mathbb{R}^n$ with radius δ. This is also called the δ neighbourhood of $\boldsymbol{x}$.

Definition 1.2. A set U is open in $\mathbb{R}^n$ if for every $\boldsymbol{x} \in U$, there exists an open ball $B_\delta(\boldsymbol{x}) \subset U$.

Remark. Every open ball in $\mathbb{R}^n$ is open.

Remark. Both $\emptyset$ and $\mathbb{R}^n$ are open.

Definition 1.3. A set F is closed in $\mathbb{R}^n$ if its complement $\mathbb{R}^n \setminus F$ is open in $\mathbb{R}^n$.

Remark. Both $\emptyset$ and $\mathbb{R}^n$ are closed.

Remark. Finite sets in $\mathbb{R}^n$ are closed.

Theorem 1.3. *Unions and finite intersections of open sets are open.*

Corollary 1.3.1. *Intersections and finite unions of closed sets are closed.*

Definition 1.4. An interior point x of a set $S \subseteq \mathbb{R}^n$ is such that there is a neighbourhood of x contained within S.

Example. Every point in an open set is an interior point by definition. The interior of a set is the largest open set contained within it.

Definition 1.5. An exterior point x of a set $S \subseteq \mathbb{R}^n$ is an interior point of the complement $\mathbb{R}^n \setminus S$.

Definition 1.6. A boundary point of a set is neither an interior point, nor an exterior point.

Example. The boundary of the unit open ball $B_1(0) \subset \mathbb{R}^n$ is the sphere S^{n-1}.

Definition 1.7. A limit point x of a set $S \subseteq \mathbb{R}^n$ is such that every neighbourhood of x contains a point from S other than itself.

Definition 1.8. The closure of a set $S \subseteq \mathbb{R}^n$ is the union of S and its limit points.

Remark. The closure of a set is the smallest closed set containing it.

Lemma 1.4. *Every open set in $\mathbb{R}^n$ is a union of open balls.*

Proof. Let $U \subseteq \mathbb{R}^n$ be open. Thus, for every $\boldsymbol{x} \in \mathbb{R}^n$, we can choose $\delta_x > 0$ such that $B_{\delta_x}(\boldsymbol{x}) \subset U$. The union of all such open balls is precisely the set U. □

1.3 $\mathbb{R}^n$ as a topological space

Definition 1.9. A topology on a set X is a collection τ of subsets of X such that

1. $\emptyset \in \tau$
2. $X \in \tau$
3. Arbitrary union of sets from τ belong to τ.
4. Finite intersections of sets from τ belong to τ.

Sets from τ are called open sets.

Example. The Euclidean metric induces the standard topology on $\mathbb{R}^n$.

Example. The discrete topology on a set X is one where every singleton set is open. This is the topology induced by the discrete metric,

$$d_{\text{discrete}} \colon X \times X \to \mathbb{R}, \qquad (x, y) \mapsto \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } x \neq y. \end{cases}$$

Example. Let X be an infinite set. The collection of sets consisting of $\emptyset$ along with all sets A such that $X \setminus A$ is finite is a topology on X. This is called the Zariski topology.

Example. Consider the set of real numbers, and let τ be the collection $\emptyset$, $\mathbb{R}$, and all intervals $(-x, +x)$ for $x > 0$. This constitutes a topology on $\mathbb{R}$, very different from the usual one.

This topology cannot be induced by a metric; it is not metrizable.

Consider the constant sequence of zeros. In this topology $(\mathbb{R}, \tau)$, this sequence converges to *every* point in $\mathbb{R}$. Given any $\ell \in \mathbb{R}$, the open neighbourhoods of ℓ are precisely the sets $\mathbb{R}$ and the open intervals $(-x, +x)$ for $x > |\ell|$. The tail of the constant sequence of zeros is contained within every such neighbourhood of ℓ, hence $0 \to \ell$. Indeed, the element zero belongs to every open set apart from $\emptyset$ in this topology.

Definition 1.10. A topological space is called Hausdorff if for every distinct $x, y \in X$, there exist disjoint neighbourhoods of x and y.

Example. Every metric space is Hausdorff. Given distinct x, y in a metric space (X, d), set $\delta = d(x, y)/3$ and consider the open balls $B_\delta(x)$ and $B_\delta(y)$.

Lemma 1.5. *Every convergent sequence in a Hausdorff space has exactly one limit.*

Proof. Consider a sequence $\{x_n\}_{n\in\mathbb{N}}$, and suppose that it converges to distinct x_1 and x_2. Construct disjoint neighbourhoods U_1 and U_2 around x_1 and x_2. Now, convergence implies that both U_1 and U_2 contain the tail of $\{x_n\}$, which is impossible since they are disjoint and hence contain no elements in common. □

Definition 1.11. Given a topological space (X, τ) and a subset $Y \subseteq X$, the collection of sets $U \cap Y$ where $U \in \tau$ is a topology τ_Y on Y. We call this collection the subspace topology on Y, induced by the topology on X.

1.4 Compact sets in $\mathbb{R}^n$

Definition 1.12. A set $K \subset X$ in a topological space is compact if every open cover of K has a finite sub-cover. That is, for every collection if $\{U_\alpha\}_{\alpha\in A}$ of open sets such that K is contained in their union, there exists a finite sub-collection $U_{\alpha_1}, \dots, U_{\alpha_k}$ such that K is also contained in their union.

Example. All finite sets are compact.

Example. Given a convergent sequence of real numbers $x_n \to x$, the collection $\{x_n\}_{n\in\mathbb{N}} \cup \{x\}$ is compact.

Example. In $\mathbb{R}^n$, compact sets are precisely those sets which are closed and bounded. This is the Heine-Borel Theorem.

Theorem 1.6. *The closed intervals $[a, b] \subset \mathbb{R}$ are compact.*

Remark. This can be extended to show that any k-cell $[a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ is compact.

Proof. Let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of $[a, b]$, and suppose that $I_1 = [a, b]$ has no finite sub-cover. Then, at least one of the intervals $[a, (a + b)/2]$ and $[(a + b)/2, b]$ must not have a finite sub-cover; pick one and call it I_2. Similarly, one of the halves of I_2 must not have a finite sub-cover; call it I_3. In this process, we generate a sequence of closed intervals $I_1 \supset I_2 \supset \dots$, none of which have a finite sub-cover. The length of each interval is given by

$$|I_n| = 2^{-n+1}\|b-a\| \to 0.$$

Now, pick a sequence of points $\{x_n\}$ where each $x_n \in I_n$. Then, $\{x_n\}$ is a Cauchy sequence. To see this, given any $\epsilon > 0$, we can find sufficiently large n_0 such that $2^{-n_0+1}\|b-a\| < \epsilon$. Thus, $x_n \in I_n \subset I_{n_0}$ for all $n \geq n_0$, which means that for any $m, n \geq n_0$, we have $x_m, x_n \in I_{n_0}$ forcing[1]

$$\|x_m - x_n\| \leq |I_{n_0}| = 2^{-n_0+1}\|b-a\| < \epsilon.$$

From the completeness of $\mathbb{R}$, this sequence must converge in $\mathbb{R}$, specifically in $[a, b]$. Thus, $x_n \to x$ for some $x \in [a, b]$. It can also be seen that the limit $x \in I_n$ for all $n \in \mathbb{N}$; if not, say $x \notin I_{n_0}$, then $x \in [a, b] \setminus I_{n_0}$ which is open, hence there is an open interval such that $(x-\delta, x+\delta) \cap I_{n_0} = \emptyset$. However, I_{n_0} contains all $x_{n \geq n_0}$, thus this δ-neighbourhood of x would miss out a tail of $\{x_n\}$.

Now, pick the open set $U \in \{U_\alpha\}$ which covers the point x. Thus, $x \in U$ so U contains some non-empty open interval $(x-\delta, x+\delta)$ around x. Choose n_0 such that $2^{-n_0+1}\|b-a\| < \delta$; this immediately gives $I_{n_0} \subseteq (x-\delta, x+\delta) \subset U$. This contradicts that fact that I_{n_0} has no finite sub-cover from $\{U_\alpha\}$, completing the proof. □

Remark. The fact that Cauchy sequences in $\mathbb{R}^n$ converge isn't immediately obvious; it is a consequence of the completeness of $\mathbb{R}^n$. Start by noting that $\mathbb{R}$ has the Least Upper Bound property, from which the Monotone Convergence Theorem follows; every monotonic, bounded sequence of reals converges. It can also be shown that any sequence of reals with contain a monotone subsequence, from which it follows that every bounded sequence contains a convergent subsequence (Bolzano-Weierstrass). Finally, it can be shown that if a subsequence of a Cauchy sequence converges, then the entire sequence also converges to the same limit, giving us the desired result for $\mathbb{R}$. For sequence in $\mathbb{R}^n$, we may apply this coordinate-wise to obtain the result.

Lemma 1.7. *Compact sets in $\mathbb{R}^n$ are closed and bounded.*

Proof. Consider a compact set $K \subset \mathbb{R}^n$. Let $x \in \mathbb{R}^n \setminus K$, and let $y \in K$. Since $x \neq y$, we choose open balls U_y around y and V_y around x such that $U_y \cap V_y = \emptyset$. Repeating this for all $y \in K$, we generate an open cover $\{U_y\}$ of K consisting of open balls. The compactness of K guarantees that this has a finite sub-cover, i.e. there is a finite set Y such that the collection $\{U_y\}_{y \in Y}$ covers X. As a result, the finite intersection of all V_y for $y \in Y$ is contained within $\mathbb{R}^n \setminus K$. Thus, x is in the exterior of K. Since x was chosen arbitrarily from $\mathbb{R}^n \setminus K$, we see that K is closed.

Now, consider the open cover $\{B_1(x)\}_{x \in K}$, and extract a finite sub-cover of unit open balls. The distance between any two points in K is at most the maximum distance between the centres of any two balls in our sub-cover, plus two. □

Lemma 1.8. *The intersection of a closed set and a compact set is compact.*

[1] If $x_1, x_2 \in [a, b]$ with $x_1 < x_2$, note that $a \leq x_1 < x_2 \leq b$, so

$$|x_2 - x_1| = x_2 - x_1 \leq b - a.$$

Proof. Let $F \subseteq \mathbb{R}^n$ be closed and let $K \subseteq \mathbb{R}^n$ be compact. Suppose that the open cover $\{U_\alpha\}$ of $F \cap K$ has no finite sub-cover. Now the complement $U = F^c$ is open in $\mathbb{R}^n$, hence the collection $\{U_\alpha\} \cup \{U\}$ is an open cover of K, and hence must admit a finite sub-cover of K. In particular, this must be a finite sub-cover of $F \cap K$. However, we can remove the set U from this sub-cover since it shares no element with $F \cap K$; as a result, our sub-cover must be a finite sub-collection of sets U_α, contradicting our assumption. This shows that $F \cap K$ is compact. □

Lemma 1.9 (Finite intersection property)**.** *Let $\{K_\alpha\}$ be a collection of compact sets in $\mathbb{R}^n$ which have the property that any finite intersection of them is non-empty. Then,*

$$\bigcap_\alpha K_\alpha \neq \emptyset.$$

Proof. Suppose to the contrary that the intersection of all K_α is empty. Fix an index β, and note that no element of K_β lies in every K_α. Set $J_\alpha = K_\alpha^c$, whence the collection $\{J_\alpha : \alpha \neq \beta\}$ is an open cover of K_β. This must admit a finite sub-cover $\{J_{\alpha_1}, \dots, J_{\alpha_k}\}$ of K_β. Thus, we must have

$$K_\beta^c \cup J_{\alpha_1} \cup \cdots \cup J_{\alpha_k} = \mathbb{R}^n.$$

This immediately gives the contradiction

$$K_\beta \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_k} = \emptyset.$$

□

Theorem 1.10 (Heine-Borel)**.** *Compact sets in $\mathbb{R}^n$ are precisely those that are closed and bounded.*

Proof. Given a compact set in $\mathbb{R}^n$, we have already shown that it must be closed and bounded. Next, if $F \subset \mathbb{R}^n$ is closed and bounded, it can be enclosed within a k-cell which we know is compact. Thus, F is the intersection of the closed set F and the compact k-cell, proving that F must be compact. □

1.5 Continuous maps

Definition 1.13. A map $f: X \to Y$ is continuous if the pre-image of every open set from Y is open in X.

Lemma 1.11. *A map $f: X \to Y$ is continuous if the pre-image of every closed set from Y is closed in X.*

Theorem 1.12. *The projection maps $\pi_i: \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x} \mapsto x_i$ are continuous.*

Proof. Let $U \subseteq \mathbb{R}$ be open; we claim that $\pi_i^{-1}(U)$ is open. Pick $\boldsymbol{x} \in \pi_i^{-1}(U)$, and note that $\pi_i(\boldsymbol{x}) = x_i \in U$. Thus, there exists $\delta > 0$ such that $(x_i - \delta, x_i + \delta) \subset U$. Now examine $B_\delta(\boldsymbol{x})$; for any point $\boldsymbol{y}$ within this open ball, we have $d(\boldsymbol{x}, \boldsymbol{y}) < \delta$ hence

$$|x_i - y_i|^2 \leq \sum_{k=1}^n (x_k - y_k)^2 = d(\boldsymbol{x}, \boldsymbol{y})^2 < \delta^2.$$

In other words, $\pi_i(\boldsymbol{y}) = y_i \in (x_i - \delta, x_i + \delta)$, hence $\pi_i\, B_\delta(\boldsymbol{x}) \subseteq (x_i - \delta, x_i + \delta) \subset U$. Thus, given arbitrary $\boldsymbol{x} \in \pi_i^{-1}(U)$, we have found an open ball $B_\delta(\boldsymbol{x}) \subset \pi_i^{-1}(U)$. □

Lemma 1.13. *Finite sums, products, and compositions of continuous functions are continuous.*

Corollary 1.13.1. *A function $f\colon [a,b] \to \mathbb{R}^n$ is continuous if and only if the components, $\pi_i \circ f$, are continuous.*

Theorem 1.14. *All polynomial functions of the coordinates in $\mathbb{R}^n$ are continuous.*

Example. The unit sphere $S^{n-1} \subset \mathbb{R}^n$ is closed. It is by definition the pre-image of the singleton closed set $\{1\}$ under the continuous map

$$\boldsymbol{x} \mapsto x_1^2 + \cdots + x_n^2.$$

Theorem 1.15. *The continuous image of a compact set is compact.*

Proof. Let $f\colon X \to Y$ be continuous, where Y is the image of the compact set X, and let $\{U_\alpha\}$ be an open cover of Y. Then, the collection $\{f^{-1}(U_\alpha)\}$ is an open cover of X. Using the compactness of X, extract a finite sub-cover $f^{-1}(U_{\alpha_1}), \ldots, f^{-1}(U_{\alpha_k})$ of X. It follows that the collection $U_{\alpha_1}, \ldots, U_{\alpha_k}$ is a finite sub-cover of Y. □

1.6 Connectedness

Definition 1.14. Let X be a topological space. A separation of X is a pair U, V of non-empty disjoint open subsets such that $X = U \cup V$.

Definition 1.15. A connected topological space is one which cannot be separated.

Lemma 1.16. *A topological space X is connected if and only if the only sets which are both open and closed are $\emptyset$ and X.*

Example. The intervals $(a,b) \subset \mathbb{R}$ are connected. To see this, suppose that U, V is a separation of (a,b). Pick $x \in U$, $y \in V$, and without loss of generality let $x < y$. Define $S = [x,y] \cap U$, and set $c = \sup S$. It can be argued that $c \in (a,b)$, but $c \notin U$, $c \notin V$, using the properties of the supremum.

Theorem 1.17. *The continuous image of a connected set is connected.*

Proof. Let f be a continuous map on the connected set X, and let Y be the image of X. If U, V is a separation of Y, then it can be shown that $f^{-1}(U)$, $f^{-1}(V)$ constitutes a separation of X, which is a contradiction. $\square$

Definition 1.16. A path γ joining two points $x, y \in X$ is a continuous map $\gamma\colon [a,b] \to X$ such that $\gamma(a) = x$, $\gamma(b) = y$.

Definition 1.17. A set in X is path connected if given any two distinct points in X, there exists a path joining them.

Lemma 1.18. *Every path connected set is connected.*

Proof. Let X be path connected, and suppose that U, V is a separation of X. Then, pick $x \in U$, $y \in V$, and choose a path $\gamma\colon [0,1] \to X$ between x and y. The sets $f^{-1}(U)$ and $f^{-1}(V)$ separate the interval $[0,1]$, which is a contradiction. $\square$

Example. All connected sets are not path connected. Consider the topologist's sine curve,

$$\left\{\left(x, \sin\frac{1}{x}\right) : 0 < x \leq 1\right\} \cup \{(0,0)\}.$$

Definition 1.18. The ϵ neighbourhood of a set K in a metric space X is defined as

$$\bigcup_{a\in K} B_\epsilon(a) = \bigcup_{a\in K} \{x \in X : d(x,a) < \epsilon\}.$$

Exercise 1.2. Let $K \subseteq \mathbb{R}^n$ be compact, and define $f\colon \mathbb{R}^n \to \mathbb{R}$,

$$f(x) = \operatorname{dist}(x, K) = \inf_{a\in K} d(x,a).$$

Show that f is continuous on $\mathbb{R}^n$, and $f^{-1}(\{0\}) = K$.

Exercise 1.3. If $K \subseteq \mathbb{R}^n$ is compact and $K \cap L = \emptyset$, then

$$\operatorname{dist}(K, L) = \inf_{a\in K} \operatorname{dist}(a, L) > 0.$$

Exercise 1.4. If $K \subseteq \mathbb{R}^n$ is compact and U is an open set containing K, then there exists $\epsilon > 0$ such that U contains the ϵ neighbourhood of K.

Is the compactness of K necessary?

1.7 Differentiability

Definition 1.19. Let $f\colon (a,b) \to \mathbb{R}^n$, and let $f_i = \pi_i \circ f$ be its components. Then, f is differentiable at $t_0 \in (a,b)$ if the following limit exists.

$$f'(t_0) = \lim_{h\to 0} \frac{f(t_0+h) - f(t_0)}{h}.$$

Remark. The vector $f'(t_0)$ represents the tangent to the curve f at the point $f(t_0)$. The full tangent line is the parametric curve $f(t) + f'(t_0)(t - t_0)$.

Definition 1.20. Let $U \subseteq \mathbb{R}^n$ be open, and let $f\colon U \to \mathbb{R}^m$. Then, f is differentiable at $x \in U$ if there exists a linear transformation $\lambda\colon \mathbb{R}^n \to \mathbb{R}^m$ such that

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

The derivative of f at x is denoted by $\lambda = Df(x)$.

Remark. In a neighbourhood of x, we may approximate

$$f(x+h) \approx f(x) + Df(x)(h).$$

Remark. The statement that this quantity goes to zero means that each of the m components must also go to zero. For each of these limits, there are n axes along which we can let $h \to 0$. As a result, we obtain $m \times n$ limits, which allow us to identify the $m \times n$ components of the matrix representing the linear transformation λ (in the standard basis). These are the partial derivatives of f, and the matrix of λ is the Jacobian matrix of f evaluated at x.

Example. Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. By choosing $\lambda = T$, we see that T is differentiable everywhere, with $DT(x) = T$ for every choice of $x \in \mathbb{R}^n$. This is made obvious by the fact that the best linear approximation of a linear map at some point is the ma itself; indeed, the 'approximation' is exact.

Lemma 1.19. *If $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \mathbb{R}^n$, with derivative $Df(x)$, then*

1. *f is continuous at x.*
2. *The linear transformation $Df(x)$ is unique.*

Proof. We prove the second part. Suppose that λ, μ satisfy the requirements for $Df(x)$; it can be shown that $\lim_{h\to 0}(\lambda - \mu)h/\|h\| = 0$. Now, if $\lambda v \neq \mu v$ for some non-zero vector $v \in \mathbb{R}^n$, then

$$\lambda v - \mu v = \frac{\lambda(tv) - \mu(tv)}{\|tv\|} \cdot \|v\| \to 0,$$

a contradiction. □

Definition 1.21. Let $U \subseteq \mathbb{R}^n$ be open, and let $f\colon U \to \mathbb{R}$. The partial derivative of f with respect to the coordinate x_j at some $a \in$ is defined by the following limit, if it exists.

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h \to 0} \frac{f(a + h\boldsymbol{e}_j) - f(a)}{h}.$$

Lemma 1.20. *If $f\colon U \to \mathbb{R}$ is differentiable at a point $a \in \mathbb{R}^n$, then*

$$Df(a)(x_1, \dots, x_n) = x_1 \frac{\partial f}{\partial x_1}(a) + \cdots + x_n \frac{\partial f}{\partial x_n}(a).$$

Example. Consider

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto \begin{cases} xy/(x^2 + y^2), & \text{if } (x, y) \neq (0, 0), \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$$

Note that f is not differentiable at $(0, 0)$; it is not even continuous there. However, both partial derivatives of f exist at $(0, 0)$.

Definition 1.22. Let $C^1(\mathbb{R}^n)$ be the set of real-valued differentiable functions on $\mathbb{R}^n$. Fix a point $a \in \mathbb{R}^n$, then fix a tangent vector $v \in \mathbb{R}^n$. Then, the map

$$\nabla_v\colon C^1(\mathbb{R}^n) \to \mathbb{R}, \qquad f \mapsto Df(a)(v)$$

is a linear functional. The quantity $\nabla_v f$ is called the directional derivative of f in the direction v at the point a.

Remark. We can represent ∇_v as the map

$$\nabla_v(\cdot) = D(\cdot)(a)(v) = \sum_i v_i \frac{\partial}{\partial x_i}\Big|_a = v \cdot \nabla(\cdot).$$

Lemma 1.21. *The directional derivatives ∇_v form a vector space called the tangent space, attached to the point $a \in \mathbb{R}^n$. This can be identified with the vector space $\mathbb{R}^n$ by the natural map $\nabla_v \mapsto v$. The standard basis can be informally denoted by the vectors*

$$\nabla_{\boldsymbol{e}_1} \equiv \frac{\partial}{\partial x_1}, \dots, \nabla_{\boldsymbol{e}_n} \equiv \frac{\partial}{\partial x_n}.$$

1.7.1 Differentiation on manifolds

Definition 1.23. A homeomorphism is a continuous, bijective map whose inverse is also continuous.

Lemma 1.22. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be continuous. Denote the graph of f as*

$$\Gamma(f) = \{(x, f(x)) : x \in \mathbb{R}^n\}.$$

Then, $\Gamma(f)$ is a smooth manifold.

Proof. Consider the homeomorphism

$$\varphi\colon \Gamma(f) \to \mathbb{R}^n, \qquad (x, f(x)) \mapsto x.$$

This is clearly bijective, continuous (restriction of a projection map), with a continuous inverse (from the continuity of f). Call this homeomorphism φ a coordinate map on $\Gamma(f)$. $\square$

Definition 1.24. Let $f\colon M \to \mathbb{R}$ where M is a smooth manifold, with a coordinate map $\varphi\colon M \to \mathbb{R}^n$. We say that f is differentiable at a point $a \in M$ if $f \circ \varphi^{-1}\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\varphi(a)$.

Definition 1.25. Let $f\colon M \to \mathbb{R}$ where M is a smooth manifold, let $\varphi\colon M \to \mathbb{R}^n$ be a coordinate map, and let $a \in M$. Let $\gamma\colon \mathbb{R} \to M$ be a curve such that $\gamma(0) = a$, and further let γ be differentiable in the sense that $\varphi \circ \gamma\colon \mathbb{R} \to \mathbb{R}^n$ is differentiable. The directional derivative of f at a along γ is defined as

$$\frac{d}{dt} f(\gamma(t))\Big|_{t=0} = \lim_{h\to 0} \frac{f(\gamma(t+h)) - f(\gamma(t))}{h}\Big|_{t=0}.$$

Note that we are taking the derivative of $f \circ \gamma\colon \mathbb{R} \to \mathbb{R}$ in the conventional sense.

Lemma 1.23. *Let γ_1 and γ_2 be two curves in M such that $\gamma_1(0) = \gamma_2(0) = a$, and*

$$\frac{d}{dt}\varphi \circ \gamma_1(t)\Big|_{t=0} = \frac{d}{dt}\varphi \circ \gamma_2(t)\Big|_{t=0}.$$

In other words, γ_1 and γ_2 pass through the same point a at $t = 0$, and have the same velocities there. Then, the directional derivatives of f at a along γ_1 and γ_2 are the same.

Definition 1.26. Let M be a smooth manifold, and let $a \in M$. Consider the following equivalence relation on the set of all curves γ in M such that $\gamma(0) = a$.

$$\gamma_1 \sim \gamma_2 \iff \frac{d}{dt}\varphi \circ \gamma_1(t)\Big|_{t=0} = \frac{d}{dt}\varphi \circ \gamma_2(t)\Big|_{t=0}.$$

Each resultant equivalence class of curves is called a tangent vector at $a \in M$. Note that all these curves in a particular equivalence class pass through a with the same velocity vector.

The collection of all such tangent vectors, i.e. the space of all curves through a modulo the equivalence relation which identifies curves with the same velocity vector through a, is called the tangent space to M at a, denoted T_aM.

Remark. Each tangent vector $v \in T_aM$ acts on a differentiable function $f\colon M \to \mathbb{R}$ yielding a (well-defined) directional derivative at a.

$$v\colon C^1(M) \to \mathbb{R}, \qquad f \mapsto \frac{d}{dt} f(\gamma_v(t))\Big|_{t=0}.$$

Thus, the tangent space represents all the directions in which taking a derivative of f makes sense.

Remark. The tangent space T_aM is a vector space. Upon fixing f, the map $Df(a)\colon T_aM \to \mathbb{R}$, $v \mapsto vf(a)$ is a linear functional on the tangent space.

Remark. Given a tangent vector $v \in T_aM$, it can be identified with its corresponding velocity vector in $\mathbb{R}^n$. Thus, the tangent space T_aM can be identified with the geometric tangent plane drawn to the manifold M at the point a.