Fun Summary

November 2, 2023

Contents

1 metric spaces

1.1 metric spaces

Definition 1.1.1. A metric space is a non-empty set X together with a map

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

$$(x, y) \mapsto d(x, y)$$

such that

1. $d(x, y) = 0$ iff $x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)\ \forall x, y, z \in X$

Remark 1.1.2. (d admits only positive values)

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y)$$

Example 1.1.3. *1.* $d_2(x, y) = ||x - y||_2$

2. $d(x, y) = \begin{cases} 0 & \textit{if } x = y \\ 1 & \textit{else} \end{cases}$

Definition 1.1.4. (convergence)
A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) is said to be convergent to $x \in X$ if

$$x_n \to x \text{ in } (X, d)$$

or

$$\lim_{n\to\infty} x_n = x \text{ in } (x, d)$$

1.2 Topology in metric spaces

Let (X, d) be a metric space.

Definition 1.2.1. 1. an open ball is defined by

$$B_r(x) = \{y \in X : \ d(x, y) < r\}$$

2. $O \subset X$ is called open if $\forall y \in O$ there is $r > 0$ such that $B_r(y) \subset O$

3. $A \subset X$ is closed if $X \setminus A$ is open.

Theorem 1.2.2. *(metric spaces are topological spaces)*
Let $\mathcal{T}$ be the set of open subsets of X. Then

1. $\varnothing, X \in \mathcal{T}$

2. *if* $U, V \in \mathcal{T}$, *then* $U \cup V \in \mathcal{T}$

3. *if* $\{U_i\}_{i\in I} \subset \mathcal{T}$, *then* $\bigcup_{i\in I} \in \mathcal{T}$

Remark 1.2.3. 1. $\varnothing, X$ are closed

2. finite union of closed sets is closed

3. arbitrary intersections of closed sets is closed

Lemma 1.2.4. *$A \subset X$ is closed iff $\forall$ convergent sequences $(x_n)_{n\in\mathbb{N}} \subset A$ the limit point is in A.*

Definition 1.2.5. For $M \subset X$ we define

$$\overline{M} = \bigcap_{A \supset M,\ A \text{ closed}}$$

as the closure of M and

$$M = \bigcup_{O \subset M,\ O \text{ open}}$$

as the interior of M.
$\partial M = \overline{M} \setminus M$ is the boundary of M

<u>Attention:</u>
Define the closed ball as $\overline{B}_r(a) = \{y \in X :\ d(y,a) \leq r\}$. Then in general $\overline{B_r(a)} \neq \overline{B}_r(a)$.
Example: Take $X \neq \varnothing$ and the trivial metric d. Then

$$B_1(a) = \{a\} = \overline{B_1(a)}$$

but $\overline{B}_1(a) = X$.

1.3 separability and completion

Let (X, d) be a metric space.

Definition 1.3.1. 1. $M \subset X$ is called dense in X if $\overline{M} = X$.

2. X is called separable if X has a countable dense subset.

Remark 1.3.2. M is dens in X iff

$$\forall x \in X \ \forall \varepsilon > 0 \ \exists y \in M \text{ s.t. } d(x, y) < \varepsilon$$

Definition 1.3.3. 1. $(x_n)_{n\in\mathbb{N}} \subset X$ is called a Cauchy sequence if

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ s.t. } m, n > N \text{ implies } d(x_n, x_m) < \varepsilon$$

2. A metric space in which all Cauchy sequences converge is called complete.

Example 1.3.4. *1.* $(C^0([a, b], \mathbb{R}), d_\infty)$ *with* $d_\infty(f, g) = \max_{x\in[a,b]} |f(x) - g(x)|$ *is complete.*

2. $(\mathbb{R}^n, d_2)$ *with* $d_2(x, y) = ||x - y||_2$ *is complete.*

Lemma 1.3.5. *Let* (X, d) *be a complete metric space and* $\varnothing \neq A \subset X$*. Then* (A, d) *is complete iff* A *is closed.*

Definition 1.3.6. $A \subset X$ is called bounded if its diameter

$$diam(A) = \sup\{d(x, y) : \ x, y \in A\}$$

is finite.

Theorem 1.3.7. (X, d) *is complete iff* $\forall (F_n)_{n\in\mathbb{N}}$ *sequences of closed subsets such that* $F_{n+1} \subset F_n$ *and* $diam(F_n) \to 0$ *then*

$$\exists! x_0 \in X \ s.t. \ \bigcap_{n\in\mathbb{N} F_n = \{x_0\}}$$

1.4 Continuity

Definition 1.4.1. Let $(X, d_x), (Y, d_y)$ be metric spaces and $f : X \to Y$. f is continuous in x_0 if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \text{ s.t. } \forall x \in X \ d_x(x, x_0) < \delta \text{ implies } d_y(f(x), f(x_0)) < \varepsilon$$

With sequences:

$$\forall (x_n)_{n\in\mathbb{N}} \subset X \; x_n \to x_0 \text{ in } (X, d_x) \text{ if it holds } (f(x_n))_{n\in\mathbb{N}} \subset Y,\; f(x_n) \to f(x_0) \text{ in } (Y, d_y)$$

f is continuous if f is continuous in x_0 for all $x_0 \in X$.
In other words f is continuous if for all $O \subset Y$ open (closed) $f^{-1}(O)$ is open (closed) in X.
Special case: f is Lipschitz continuous if $\exists L > 0$ s.t.

$$d_y(f(x), f(y)) \leq L d_x(x, y) \;\forall x, y \in X$$

f is an isometric if $\forall x, y \in X$ it holds that $d_Y(f(y), f(x)) = d_x(x, y)$.

1.5 Compact sets

Definition 1.5.1. Let (X, d) be a metric space and $A \subset X$.

1. an open cover of A is a collection $\{U_i\}_{i\in I}$ where $I \neq \varnothing$ is an arbitrary index set of open subsets of X s.t. $A \subset \bigcup\limits_{i\in I} U_i$.
2. A is compact if every open cover of A contains a finite subcover i.e. there is $N \in \mathbb{N}$ and indices $i_1, ..., i_N$ such that

$$A \subset U_1 \cup ... \cup U_N$$

3. A is sequentially compact if every sequence in A has a convergence subsequence in A.

4. A is called precompact or totally bounded if $\forall \varepsilon > 0 \; \exists N \in \mathbb{N}$ and $\exists x_1, ..., x_N \in X$ such that $A \subset \bigcup_{i=1}^{N} B_\varepsilon(x_i)$.

Theorem 1.5.2. *Let (X, d) be a metric scape and $A \subset X$. The following are equivalent:*

1. *A is compact*
2. *A is sequentially compact*
3. *(A, d) is complete and A is precompact.*

Remark 1.5.3. If A is precompact, then $\overline{A}$ is precompact. Further, if (X, d) is complete and $A \subset X$ then A is precompact $\Leftrightarrow$ $\overline{A}$ is compact.

Recall: A compact $\Rightarrow$ bounded and closed and $f : X \to Y$ continuous with $A \subset X$ compact, then $f(A)$ is compact as well. Further, if $f : A \to \mathbb{R}$ is continuous and A is compact, then

$$\exists x_1, x_2 \in A \ \text{ s.t. } f(x_1) \leq f(x) \leq f(x_2) \ \forall x \in A$$

Theorem of Heine-Borel: $A \subset \mathbb{R}^n$ is compact iff A is closed and bounded.

1.6 Theorem of Baire

Theorem 1.6.1. *Let (X, d) be a complete metric space and $\forall n \in \mathbb{N}$ consider $U_n \subset X$ open and dense. Then*

$$\bigcap_{n \in \mathbb{N}} U_n$$

is dense in X.

Remark 1.6.2. 1. Completeness is in general necessary. Consider $(\mathbb{Q}, d)$ and $d(x, y) = |x - y|$. Define a sequence x_n such that $\mathbb{Q} = \{x_n \ n \in \mathbb{N}\}$. Take $U_n = \mathbb{Q} \setminus \{x_n\}$ which is open and dense. Then

$$\bigcap_{n \in \mathbb{N}} U_n = \varnothing$$

Corollary 1.6.3. *Let (X, d) be a complete metric space. Let $\forall n \in \mathbb{N}$, $A_n \subset X$ be closed and*

$$X = \bigcup_{n \in \mathbb{N}} A_n$$

Then $\exists N \in \mathbb{N}$ s.t. A_N has an interior point.

Remark 1.6.4. Theorem 1.6.1 is also called Baire category theory.

- In a metric space (X, d) $A \subset X$ is called nowhere dense if $\overline{A}$ has no interior points.
- A is called of first category if $\exists (M_n)_{n \in \mathbb{N}}$ where $M_n \subset A$ nowhere dense s.t. $A = \bigcup_{n \in \mathbb{N}} M_n$
- A is called of second category if it is not of first category

Hence the theorem of Baire implies that every complete metric space is of second category.

2 Normal spaces and Banach spaces

Let X be a $\mathbb{K}$-vector space where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

2.1 definitions

Definition 2.1.1. A map $||\cdot|| : X \to \mathbb{R}$ is called a norm on X if

1. $\forall x \in X$, $||x|| \geq 0$ and $||x|| = 0$ iff $x = 0$
2. $\forall \lambda \in \mathbb{K}$ and $\forall x \in X$ it holds that $||\lambda x|| = |\lambda| \cdot ||x||$
3. $\forall x, y \in X$ it holds $||x + y|| \leq ||x|| + ||y||$

The pair $(X, ||\cdot||)$ is called an normed space.

$p : X \to \mathbb{R}$ is called a seminorm if $p(x) \geq 0$ $\forall x \in X$ and 2. and 3. are also satisfied.

Example 2.1.2. *1. $C^0([0,1];\mathbb{R})$ with $||f||_\infty = \max_{x\in[0,1]} |f(x)|$*

2. more general for a compact metric space K: $C^0(K,\mathbb{R})$ with $||f||_\infty = \max_{x\in K} |f(x)|$

3. $C^1([0,1];\mathbb{R})$ with $p(f) = \max_{x\in[0,1]} |f'(x)|$

4. $\Omega \subset \mathbb{R}^n$ measurable. $L^1(\Omega) = \{f : \Omega \to \mathbb{R} :\ f$ integrable $\}$ with

$$p : L^(\Omega) \to \mathbb{R} :\ p(f) = \int_\Omega |f(x)|\, dx$$

then p is a seminorm.

Remark 2.1.3. Any normed space is a metric space via

$$d(x, y) = ||x - y||$$

All concepts from chapter 1 apply.

Lemma 2.1.4. *Let $(X, ||\cdot||)$ be a normed space. Then X is called separable iff $\exists A \subset X$ countable such that s.t. $\overline{span\{A\}} = X$ where $span\{A\} = \{\sum_{i=1}^n \lambda_i x_i\}$ with $n \in \mathbb{N}$, $\lambda_i \in K$ and $x_i \in A$. Here the colusre is defined w.r.t the norm.*

Definition 2.1.5. A complete normed space is called a Banach space.

2.2 Example: l^p-spaces

We consider the vector space $\mathbb{K}^\mathbb{N}$ of sequences in in $\mathbb{K}$. Let $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$. Define $x + y = (x_n + y_n)_{n\in\mathbb{N}}$ and $\lambda x = (\lambda x_n)_{n\in\mathbb{N}}$.
For $x \in \mathbb{K}^\mathbb{N}$ define

$$||x||_{l^p} = \left(\sum_{n=1}^{\infty} |x|^p\right)^{1/p}$$

for $1 \leq p < \infty$ and

$$||x||_{l^\infty} = \sup_{n \in \mathbb{N}} |x_n|$$

else.

Define $l^p = \{x = (x_n)_{n\in\mathbb{N}} : ||x||_{l^p} < \infty\}$ for $1 \leq p \leq \infty$. We find that l^p is a subspace of $\mathbb{K}^{\mathbb{N}}$ and l^p is a normed space (for the triangle inequality use the Hölder inequality).

Theorem 2.2.1. *For $1 \leq p \leq \infty$ l^p is a Banach space.*

Lemma 2.2.2. *For finite p, l^p is separable while l^∞ is not.*

2.3 Finite dimensional normed spaces

Let X be a vector space over $\mathbb{K}$. $\exists e_1, ..., e_n \in X$ s.t.

$$\forall x \in X;\ \exists \lambda_1, ..., \lambda_n \in \mathbb{K} :\ x = \sum_{i=1}^{n} \lambda_i x_i$$

For $p \in [1, \infty)$ we define

$$||x||_p = \left(\sum_{i=1}^{n} |\lambda_i|^p \right)^{1/p}$$

and for $p = \infty$

$$||x||_\infty = \max_{1 \leq i \leq n} |\lambda_i|$$

Definition 2.3.1. Two norms are equivalent in that

$$\alpha || \cdot ||_1 \leq || \cdot ||_2 \leq \beta || \cdot ||_1$$

Theorem 2.3.2. *In a finite dimensional space, all norms are equivalent.*

Theorem 2.3.3. *Finite dimensional normed spaces are Banach spaces.*

2.4 On the closure of $\overline{B_1(0)}$

Lemma 2.4.1 (Lemma of Riesz, Lemma of the almost orthogonal element)**.** *Let X be a normed space. $U \subset X$ a closed subspace of X s.t. $U \neq X$. Then $\forall \lambda \in (0,1) \exists x_\lambda \in X$ s.t. $||x_\lambda|| = 1$ and $dist(x_\lambda, U) \geq \lambda$.*

Theorem 2.4.2. *In a normed space X, $\overline{B_1(0)}$ is compact iff X is finite dimensional.*

3 A question from approximation theory

3.1 Theorem of Stone-Weierstrass

Let X be a compact metric space. Then $(C^0(X), \mathbb{K}), ||\cdot||_\infty$, where $||f||_\infty = \max\limits_{x\in X} |f(x)|$ is a Banach space.
Which property of $A \subset C^0(X, \mathbb{K})$ ensures that A is dense.

Definition 3.1.1. $A \subset C^0(X, \mathbb{K})$ is called subalgebra, if $\forall f, g, \in A$

1. $\lambda f + \mu g \in A$ (subspace)

2. $f \cdot g \in A$

Example 3.1.2.

- $\{p : [0, 1] \to \mathbb{R}\}$ *is a subalgebra of* $C^0([0, 1]; \mathbb{R})$.

- $\{f : [-1, 1] \to \mathbb{R}; f$ *continuous and even*$\}$ *is a subalgebra.*

Remark 3.1.3. If A is a subalgebra, then $\overline{A}$ is also a subalgebra.

Definition 3.1.4. Let $A \subset C^0(X)$ be a subalgebra.

1. A is called unital if $1 \in A$

2. A separates point if $x, y \in X$, $x \neq y$, $\exists f \in A$ s.t. $f(x) \neq f(y)$.

3. (if $\mathbb{K} = \mathbb{C}$) A is stable under conjuguation if from $f \in A$ we conclude that also $\overline{f} \in A$.

Remark 3.1.5. If A is unital then all constant functions are in A.

Lemma 3.1.6. *Consider* $f : [-1, 1] \to \mathbb{R}$ *where* $f(x) = |x|$. *Then* $\exists$ *sequence of polynomials* $(p_n)_{n\in\mathbb{N}}$ *s.t.*

$$p_n \to f$$

uniformly in $[-1, 1]$.

Lemma 3.1.7. *Let* $A \subset C^0(X, \mathbb{R})$ *be a unital subalgebra. Then*

1. *if* $f \in A$ *then* $|f| \in \overline{A}$.

2. *if* $f, g \in A$ *then* $\max\{f, g\} \in \overline{A}$ *and* $\min\{f, g\} \in \overline{A}$

Theorem 3.1.8 (Stone-Weierstrass)**.** *Let A be a compact metric space.* $A \subset C^0(X, \mathbb{K})$ *is a unital subalgebra that separates points and if* $\mathbb{K} = \mathbb{C}$ *is stable under conjugation, then* A *is dense in* $C^0(X, \mathbb{K})$ *w.r.t* $||\cdot||_\infty$.

Theorem 3.2.1 (Theorem of Weierstraß). *Let $[a, b]$ be a compact interval in $\mathbb{R}$, $f : [a, b] \to \mathbb{R}$ be a continuous function and $\varepsilon > 0$. Then $\exists p : [a, b] \to \mathbb{R}$ a polynomial s.t.*

$$||p - f||_\infty = \sup_{x \in [a,b]} |p(x) - f(x)| < \varepsilon$$

Definition 3.2.2. A function $f : \mathbb{R} \to \mathbb{C}$ is periodic if

$$f(x + t) = f(x)$$

for a $t \in \mathbb{R}$ and all $x \in \mathbb{R}$.

Remark 3.2.3. If f is periodic with period t then $\tilde{f} : \mathbb{R} \to \mathbb{C}$ where $\tilde{f}(x) = f(t\frac{x}{2\pi})$ is periodic of period 2π.
Consider $C^0_{2\pi}(\mathbb{R}, \mathbb{C})$ the space of continuous 2π-periodic functions. We consider the span of $\{e^{ikx} = \cos(kx) + i\sin(kx), k \in \mathbb{Z}\}$.

Definition 3.2.4. A trigonometric polynomial is a function $f : \mathbb{R} \to \mathbb{C}$

$$f(x) = \sum_{k=-N}^{N} c_k \cdot e^{ikx}$$

with $c_k \in \mathbb{C}$

Theorem 3.2.5 (Approximation of periodic functions). *Trigonometric polynomials are dense in $(C^0_{2\pi}(\mathbb{R}, \mathbb{C}), || \cdot ||_\infty)$*

Application to neural networks

The simplest case of a neural network has d inputs $x_1, ..., x_d$ and one output Z called a *feed forward* network. Each input influences the output and x_i might have a weight α_i associated to it. The output is a function in $x = (x_1, ..., x_d)$ and the weights $\alpha = (\alpha_1, ..., \alpha_d)$. For instance, the output is often of the form

$$Z = \sum_{i=1}^{d} \alpha_i x_i + b$$

where b is the bias of the network. To make the network slightly stronger, we add a intermediate layer $y = (y_1, ..., y_r)$ where each x_i is connected to each y_j with the associated weight $\gamma_{i,j}$. The y layer (often called activation) is the connected to the output Z as above with weights α_j. We introduce the realtion

$$y_j = \Phi(\sum_{i=1}^{d} \gamma_{j,i} x_i + b)$$

for a measurable function Φ. Lastly, the output is then given by

$$Z = \sum_{j=1}^{r} \alpha_j y_j$$

Definition 3.2.6. 1. $A^d = \{a : \mathbb{R}^d \to \mathbb{R} : \ a(x9 = w^T x + b)\}$ where $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$.

2. given $\Phi : \mathbb{R} \to \mathbb{R}$ measurable $d \in \mathbb{N}$ define $\Sigma^d(\Phi) = \{f : \mathbb{R}^d \to \mathbb{R} : \ f(x) = \sum_{j=1}^{N} \alpha_j \Phi(a_j(x))$ wtih $N \in \mathbb{N}, \alpha_j \in \mathbb{R}, a_j \in A^d\}$ as the set of single hidden layer feed forward networks.

3. A squashing function is a measurable non-decreasing function $\Phi : \mathbb{R} \to \mathbb{R}$ s.t. $\lim_{x \to -\infty} \Phi(x) = 0$ and $\lim_{x \to \infty} \Phi(x) = 1$.