Lecture 7

- Covering numbers and total boundedness

- Topological Notions in Metric Spaces: sequential convergence

We have "plausibly" strengthened the def. of boundedness to total boundedness:

"$\forall \varepsilon > 0$, $\exists$ a finite cover of open balls of radius ε"

which is negated by:

"$\exists \varepsilon > 0$, $\forall$(finite collection of open balls of radius ε is not a cover)"

- Negation is "easier" to hold → easier to construct counter examples

- We will try to construct examples using (without much rigour) the concept of covering numbers – Metric Entropy.

Covering Numbers and Metric Entropy

(Our approach: not very rigorous)

We will use without loss of generality covers by closed balls.

Framework: X, d, $A \subseteq X$ ($\neq \emptyset$), $\varepsilon > 0$

Definition: The covering number of A w.r.t. ε and d, $N(\varepsilon, d, A) :=$ min # of closed balls of radius ε that cover A.

Remark: it is possible to prove that $N(\varepsilon, d, A)$ is unique $\forall \varepsilon, d, A$.
$\hookrightarrow$ well defined as a function of ε, d, A

- $\ln N(\varepsilon, d, A)$ is the metric entropy number w.r.t. ε and d.

- It is easy to show that A is t.b. w.r.t. d iff $N(\varepsilon, d, A) \in \mathbb{N}$ (i.e. it is $< +\infty$) $\forall \varepsilon > 0$.

  (dually A is not t.b. w.r.t. d iff $\exists \varepsilon > 0$: $N(\varepsilon, d, A) = +\infty$)

- It is easy to see that $N(\varepsilon, d, A)$ is decreasing w.r.t. ε for fixed d, A.

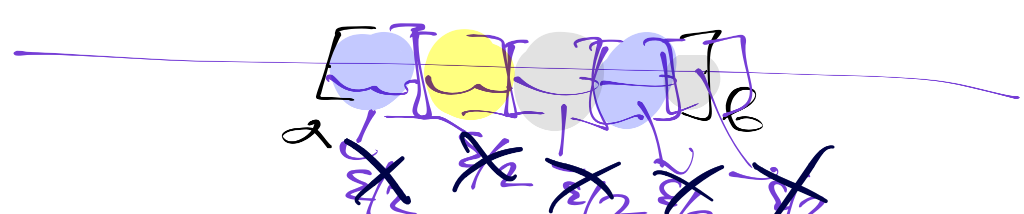
- Generally, as $\varepsilon \downarrow 0$ $\left(\begin{smallmatrix}\varepsilon \to 0 \\ \varepsilon > 0\end{smallmatrix}\right)$, $N(\varepsilon, d, A) \to +\infty$ for fixed d, A

  even when A is t.b.

- What is very informative on the complexity of A as a metric (sub-)space is the rate at which $N(\varepsilon, d, A)$ (or equivalently $\ln N(\varepsilon, d, A)$) diverges as $\varepsilon \downarrow 0$.

E.g. $X = \mathbb{R}$, $d = d_u$ $\quad A = [\alpha, \beta] = O_{d_u}\left[\frac{\alpha+\beta}{2}, \frac{\beta-\alpha}{2}\right]$

closed ball



$$N(\varepsilon, d_u, [\alpha,\beta]) = \begin{cases} 1, & \varepsilon \geq \frac{\beta-\alpha}{2} \\ \text{ο μικρότερος φυσικός} \geq \frac{\beta-\alpha}{2\varepsilon}, & \varepsilon < \frac{\beta-\alpha}{2} \end{cases} < +\infty$$

the formula is valid also in the first case

$\forall \varepsilon > 0$. Hence $[\alpha, \beta]$ is t.b. w.r.t. d_u

Hence we have shown that $\forall x \in \mathbb{R}, \varepsilon > 0$, $O_{d_u}[x, \varepsilon]$ is totally bounded w.r.t. d_u. Hence if $A \subseteq \mathbb{R}$ is bounded w.r.t. d_u then it is also t.b. w.r.t. d_u.

Hence Boundness $\Longleftrightarrow$ Total Boundness inside $\mathbb{R}, d_u$

[redacted]

- If (X,d) a metric space, $A \subseteq X$, $\varepsilon > 0$

$N(\varepsilon, d, A) :=$ the "smallest" number of closed balls of ε radius that covers A. [Covering number of A corresponding to ε].

- Variations of the above are definable if we use open balls and/or if we restrict the centers to lie in A. Those variations will not necessarily equal each other but it is possible to prove that they have equivalent asymptotic behavior as $\varepsilon \downarrow 0$, hence they convey the same information on the complexity of A.
- A is t.b. iff $N(\varepsilon, d, A) \in \mathbb{R}$ $\forall \varepsilon > 0$.
- $\ln N(\varepsilon, d, A)$ is termed as metric entropy number of A.

E.g. $X = \mathbb{R}$, $d = d_u$, $A = [\alpha, \beta] = O_{d_u}[\frac{\alpha+\beta}{2}, \frac{\beta-\alpha}{2}]$
$\alpha, \beta \in \mathbb{R}$

we have easily shown $N(\varepsilon, d_u, [\alpha,\beta]) \sim \frac{c}{\varepsilon}$ depends on α, β

Since $\frac{c}{\varepsilon} \in \mathbb{R}$ $\forall \varepsilon > 0$ this essentially showed us that boundness is equivalent to total boundness in $\mathbb{R}, d_u$

- Metric Entropy: $\ln N(\varepsilon, d_u, [\alpha,\beta]) \sim \ln c - \ln \varepsilon$ dominant term as $\varepsilon \downarrow 0$
$\sim -\ln \varepsilon$

E.g. $X = \mathbb{R}^d$ $d = d_I$ it is possible to prove that:
$\forall x \in \mathbb{R}^d, \delta > 0$, $A = O_{d_I}[x, \delta]$, $N(\varepsilon, d_I, O_{d_I}[x,\delta]) \sim \frac{c}{\varepsilon^d}$ C is independent of ε
for some suitable $c > 0$. Since $\frac{c}{\varepsilon^d} \in \mathbb{R}$ $\forall \varepsilon > 0$ every closed ball in this metric space is totally bounded.

Hence boundness is equivalent to total boundness in $\mathbb{R}^d, d_I$.

– Metric Entropy: $\ln N(\varepsilon, d_I, O_{d_I}[x,\delta]) \sim -d \ln \varepsilon$

E.g. $X = B([0,1], \mathbb{R})$, $d = d_{sup}$, $A = \{f : [0,1] \to \mathbb{R}, f(0) = 0,$
$L > 0$ $\quad |f(x) - f(y)| \leq L|x-y|\}$

We know that A is bounded (uniform boundness)

but is it totally bounded?

$\rightarrow$ using piecewise linear approximations

We can show that $N(\varepsilon, d_{sup}, A) \sim \exp(c/\varepsilon)$ and

c is a positive constant independent of ε that depends on L.

Since $\exp(c/\varepsilon)$ is finite $\forall \varepsilon > 0$ we conclude that A is totally bounded w.r.t. d_{sup}. By comparing the asymptotic behavior of $e^{c/\varepsilon}$ with c/ε^d we can conclude that the particular A is "more complex" than any Euclidean ball

Metric Comparison and total boundness

X, d_1, d_2 well defined metrics, for some $c > 0$, $\underline{d_1 \leq c d_2}$.

$\Leftrightarrow \forall x \in X, \varepsilon > 0 \quad O_{d_2}(x, \varepsilon) \subseteq O_{d_1}(x, c\varepsilon) \Leftrightarrow \forall x \in X, \forall \delta > 0$

$\updownarrow$ $\quad \delta = c\varepsilon$ $\rightarrow$ bijection since $c > 0$.

$(*)$ $\quad O_{d_2}(x, \delta/c) \subseteq O_{d_1}(x, \delta)$ ✓

Suppose that A is t.b. w.r.t. d_2. In order to see whether A is t.b. w.r.t. d_1, let $\delta > 0$, since A is d_2 t.b. we have that there exists a finite cover of open balls or radius δ/c of A w.r.t. d_2, hence due to $(*)$ there exists a finite cover of open balls of radius δ of A w.r.t. d_1 (Just use the centers of the previous cover). Hence since δ is arbitrary A is t.b. w.r.t. d_1. Hence we have proven:

Lemma. Total boundness w.r.t. $d_2 \Rightarrow$ Total boundness w.r.t. d_1.

Furthermore due to $(*)$ we have that

$$N(\delta/c, d_2, A) \geq N(\delta, d_1, A).$$ ← Exercise!

↓ Corollary of the previous

Lemma. Suppose that for $c_1, c_2 > 0$: $c_1 d_2 \leq d_1 \leq c_2 d_2$. $(**)$

Then Total boundness w.r.t. $d_1 \Leftrightarrow$ Total boundness w.r.t. d_2.

For example in $\mathbb{R}^d$ boundness is equivalent to total boundness w.r.t. any of the metrics d_I, d_{max}, d_{II}. $(*)$

(Exercise: provide the details)

Exercise: what is the relation between the respective covering numbers when $(**)$ holds?

– Total boundness can be very useful in applications related with the issue of convergence in function spaces.

[redacted]

it seems that t.b. "matters" in function spaces

$(*)$ Implies that in $(\mathbb{R}^d, d_*)$ tot. boundedness $\Leftrightarrow$ boundedness $\forall * = I, II, max$

Topological Notions in Metric Spaces

- A topology on a set X is a "specification" of which subsets of X are considered open and dually closed.
- Topologies facilitate the examination of notions of convergence (limits) and continuity.
- In metric spaces balls "produce" topologies.

We will try to largely avoid the constructions implied above and examine convergence and continuity by solely restricting ourselves to the use of open (closed) balls.
↦ you will take a glimpse of the general approach of the tutorials

Convergence. (of sequences in metric spaces)

Preparation: Let X be a non-empty set. A sequence of elements of X (X-valued sequence) is a function $\mathbb{N} \to X$ or equivalently a vector of elements of X that has an initial element, does not have a final element and has as many elements as $\mathbb{N}$.
generalization of real sequences

E.g. $X = \mathbb{R}$, $f : \mathbb{N} \to \mathbb{R}$, $f(n) = \frac{1}{n+1}$, or equivalently it is the vector $(1, 1/2, 1/3, \dots, 1/n+1, \dots)$
this is an example of a real sequence

E.g. $X = \mathcal{B}([0,1], \mathbb{R})$, and consider the sequence

$(1, x, x^2, \dots, x^n, \dots)$, $x \in [0,1]$

- A property P is said to hold almost everywhere for a given sequence iff every element of the sequence obeys P except for a finite number of elements.

E.g. A X-valued sequence is termed almost everywhere (or eventually) constant iff for some $c \in X$ we have that $x_n = c$ $\forall n \in \mathbb{N}$ except for a finite number of n's

is the $n+1^{th}$ element of the sequence $(x_0, x_1, x_2, \dots, x_n, \dots)$

This is equivalent to that $\exists n^* \in \mathbb{N}$:

$$(x_0, x_1, x_2, \dots, x_{n^*-1}, c, c, \dots, c, \dots) \checkmark$$

Consider now a metric space X, d and $(x_0, x_1, \dots, x_n, \dots)$ is a X-valued sequence.

Definition. The sequence above is convergent (w.r.t. d) iff $\exists x \in X$ w.r.t. which we have that:
$\forall \varepsilon > 0$ the sequence lies almost everywhere in $O_d(x, \varepsilon)$ (i.e. it is allowed that a finite part of the sequence can lie outside $O_d(x,\varepsilon)$)
x is termed limit of the sequence (w.r.t. d) and it is denoted with $\lim x_n$ or $d\text{-}\lim x_n$ etc.
(the convergence is also usually denoted with $x_n \to x$)

Remark: the definition above allows that the elements and their number of the sequence that lie outside $O_d(x,\varepsilon)$ may depend on ε.

Remark: Suppose that $x_n \to x$ (w.r.t. d), i.e. due to the definition above $\forall \varepsilon>0$, almost every element of the sequence lies in $O_d(x,\varepsilon) \Leftrightarrow d(x_n,x)<\varepsilon$ for almost every n (i.e. this holds $\forall n\in\mathbb{N}$ except for a finite number of n's). Now consider the vector

$$(d(x_0,x), d(x_1,x), d(x_2,x), \dots, d(x_n,x), \dots) \checkmark$$

this is a $\mathbb{R}$-valued sequence, and notice that

$$d(x_n,x)<\varepsilon \Leftrightarrow |d(x_n,x)|<\varepsilon \Leftrightarrow |d(x_n,x)-0|<\varepsilon$$

$\Leftrightarrow d(x_n,x) \in O_{d_u}(0,\varepsilon)$, and since for any $\varepsilon>0$ $\to d_u(d(x_n,x),0)<\varepsilon$

this holds $\forall n\in\mathbb{N}$ except for a finite number of n's

the previous is equivalent to that $\lim\limits_{n\to+\infty} d(x_n,x)=0$.

Hence $x_n \to x \Leftrightarrow d(x_n,x)\to 0$. $\square$

Remark: In the definition above we can equivalently replace open balls with the respective closed balls.

This is due to the following: (def. based on open balls $\Rightarrow$ def. based on closed balls):

If $x_n \to x$ via the original definition then $\forall \varepsilon>0$ almost every element of the sequence lies $O_d(x,\varepsilon)$ but since $O_d(x,\varepsilon) \subseteq O_d[x,\varepsilon]$ then almost every element of the sequence lies in $O_d[x,\varepsilon]$.

(def. based on closed balls $\Rightarrow$ def. based on open balls)

If $x_n \to x$ via the "closed-balls" based definition. Let $\varepsilon>0$. Due to the closed balls definition almost every element of the sequence lies $O_d[x,\varepsilon/2]$. But we have that $O_d[x,\varepsilon/2] \subseteq O_d(x,\varepsilon)$ and thereby almost every element of the sequence lie in $O_d(x,\varepsilon)$. The result follows since ε is arbitrary.

Lemma. If the sequence is convergent the its limit is unique.

Proof.

$\to$ Next lecture

[see next page: the result heavily depends on the separation by open (closed) balls property of metric spaces.

Definition Denial

* $x_n \to x$ iff $\forall \varepsilon>0$, $x_n \in O_d(x,\varepsilon)$ for almost every n.
* $x_n \not\to x$ iff $\exists \varepsilon>0$, $x_n \notin O_d(x,\varepsilon)$ for infinite n.

End of lecture 7

[redacted] [redacted]

Reminder: Sequential convergence in Metric spaces

- (X,d) a M.S., (x_n) a sequence of elements of X, x is a limit of (x_n) (w.r.t. d) iff $\forall \varepsilon > 0$, $x_n \in O_d(x,\varepsilon)$ $\forall n \in \mathbb{N}$ except for a finite number of n (that may depend on ε).
- Equivalent if closed balls are used.

$(x_0, x_1, x_2, \dots, x_n, \dots)$

Uniqueness: If (x_n) has a limit (w.r.t. d) then it is unique.

Proof. Suppose that (x_n) is convergent. Suppose that its limit is not unique. This means that it has at least two limits, say $x,y \in X$ $(x \neq y)$. Since $x \neq y \Leftrightarrow d(x,y) > 0$, hence due to the separation with open balls property in metric spaces, $\exists\, \varepsilon, \delta > 0$: $O_d(x,\varepsilon) \cap O_d(y,\delta) = \emptyset$ $\left(\varepsilon = \delta = \frac{d(x,y)}{2}\right)$. Since $x_n \to x \Rightarrow x_n \in O_d(x,\varepsilon)$ $\forall n \in \mathbb{N}$ except for a finite number of n. Hence only a finite number of x_n's lie in $O_d^c(x,\varepsilon)$. But $O_d(y,\delta) \subseteq O_d^c(x,\varepsilon)$ hence only a finite number of x_n's can lie inside $O_d(y,\delta)$. This is impossible because $x_n \to y$ by assumption. □

Remark: In pseudo-metric spaces uniqueness may fail.

↦ Could you construct an example?

Lemma (Boundness). If (x_n) is convergent (w.r.t. d) then (x_n) is bounded (w.r.t. d).
↳ as a subset of X

Reminder: (x_n) is considered bounded w.r.t. d iff $\exists y \in X$, $\delta > 0$: $x_n \in O_d(y, \delta)$ (equiv. $x_n \in O_d[y,\delta]$), $\forall n \in \mathbb{N}$.

Proof. We have that $x_n \to x$ for some $x \in X$ (we use closed balls without loss of generality). We know that $x_n \in O_d[x,1]$ $\forall n \in \mathbb{N}$ except for a finite number of n. Suppose that $x_{n_1}, x_{n_2}, \dots, x_{n_k} \notin O_d[x,1]$ for some $k > 0$. Since k is finite $0 < \max(\underbrace{d(x,x_{n_1}), d(x,x_{n_2}), \dots, d(x,x_{n_k}), 1}_{k+1 \text{ reals}}) < +\infty$

Hence we can choose $\delta := \max(d(x,x_{n_1}), d(x,x_{n_2}), \dots, d(x,x_{n_k}), 1)$ as a new radius and consider $O_d[x,\delta] \supseteq O_d[x,1]$ hence every element of the sequence that belong in $O_d[x,1]$ also belong in $O_d[x,\delta]$, but also for $i = 1,\dots,k$ we have that $d(x, x_{n_i}) \leq \max(d(x,x_{n_1}), \dots, d(x,x_{n_k}), 1) = \delta$ hence $x_{n_i} \in O_d[x,\delta]$ $\forall i = 1,\dots,k$. Hence $x_n \in O_d[x,\delta]$ $\forall n \in \mathbb{N}$.

Hence (x_n) is bounded. ▫

Remark. The converse does not generally hold.

Does convergence imply total boundedness?

Lemma. Every (x_n) that is eventually constant (say at $x \in X$) is convergent (to c) w.r.t. every d. (universal property)

Proof. Since (x_n) is eventually constant (at $c \in X$)

(x_n) will have the form $(x_0, x_1, \dots, x_k, c, c, \dots, c, \dots)$

where $x_i \neq c$ possibly for some $i = 0, 1, \dots, k$, for some k.

Consider $O_d(c, \varepsilon)$ for any d, $\varepsilon > 0$. Then $x_n \in O_d(c, \varepsilon)$ $\forall n \in \mathbb{N}$ except perhaps for some $n = 0$, and/or 1, ..., and/or k.

Hence $x_n \to c$ $\forall d$. $\square$

Examples:

1. $X = \mathbb{R}$, $d = d_u$, $x_n = \frac{1}{n+1}$, $n \in \mathbb{N}$, $(1, 1/2, 1/3, \dots, 1/{n+1}, \dots)$.

We have that $\frac{1}{n+1} \to 0$ w.r.t. d_u since for $\varepsilon > 0$ we have

$O_{d_u}(0, \varepsilon) = (-\varepsilon, \varepsilon) \ni \frac{1}{n+1}$ $\forall n >$ smallest natural greater than or equal $1/\varepsilon - 1$.

2. X arbitrary, $d = d_d$ (hence we are examining the issue of sequential convergence in discrete spaces). Remember that

$\forall x \in X$, $\varepsilon > 0$, $O_{d_d}(x, \varepsilon) = \begin{cases} X, & \varepsilon > 1 \\ \{x\}, & \varepsilon \leq 1 \end{cases}$. Hence in order

for (x_n) (with $x_n \in X$ $\forall n \in \mathbb{N}$) to converge to x, x_n must equal x $\forall n \in \mathbb{N}$ except for a finite number of n. Hence in order for (x_n) to be convergent w.r.t. d_d it has to be eventually constant. Hence combining this with the previous Universal Property we have proven that:

In a discrete space a sequence is convergent iff it is eventually constant.

This showed us:

- The issue of convergence or divergence of a sequence depends among others on d (e.g. $(1/n+1)$ in $\mathbb{R}$ converges to 0 w.r.t. d_u, yet diverges a non-eventually constant of).
- Convergence in discrete spaces seems trivial!

3. $X = B(Y, \mathbb{R})$, $d = d_{sup}$, if (f_n) is a sequence in X (i.e. $f_n \in B(Y, \mathbb{R})$ $\forall n \in \mathbb{N}$), and it is convergent w.r.t. d_{sup} to some limit $f \in B(Y, \mathbb{R})$ - we say that f_n is uniformly convergent to f.

Compare the above with: if $x \in Y$, then $(f_n(x)) := (f_0(x), f_1(x), \dots, f_n(x), \dots)$ is a sequence of real numbers. ← real sequence We can ask for which $x \in Y$ the real sequence $(f_n(x))$ converges w.r.t. d_u.

We can thus define another concept of convergence for (f_n).

We say that $f: Y^* \to \mathbb{R}$ (Y^* is a suitable subset of Y - see below) is the pointwise limit of (f_n) iff $\forall x \in Y^*$, the real sequence $(f_n(x))$ is convergent w.r.t. d_u, and $f(x) := \lim f_n(x)$ (w.r.t. d).

[the pointwise limit exists iff $Y^* \neq \emptyset$ i.e. iff there exists at least a $x \in Y$ for which the real sequence $(f_n(x))$ is convergent w.r.t. d_u].

What is the relation between pointwise and uniform convergence:

Pointwise convergence of (f_n) to f is defined by the convergence of $|f_n(x) - f(x)| \to 0 \ \forall x \in Y^*$

Uniform convergence of (f_n) to f is defined by the convergence of $d_{sup}(f_n, f) = \sup_{x \in Y} |f_n(x) - f(x)| \to 0$

We have that $\forall x \in X \quad |f_n(x) - f(x)| \leq \underbrace{\sup_{x \in Y} |f_n(x) - f(x)|}_{\to d_{sup}}$

hence if $\sup_{x \in Y} |f_n(x) - f(x)| \xrightarrow{d_u} 0 \ \forall x \Rightarrow |f_n(x) - f(x)| \to 0 \ \forall x \in Y$

hence if (f_n) converges uniformly to f the necessarily it converges to the same limit pointwisely.

Does the converse hold? The following counter example shows that uniform convergence is genuinely stronger than pointwise convergence:

$Y=[0,1]$, $X=B([0,1],\mathbb{R})$, and $f_n(x):=x^n$, $x\in[0,1]$

(hence we are considering the following sequence in

$$(\underset{\substack{\| \\ \text{constant} \\ \text{at } 1}}{x^0}, x, x^2, \dots, x^n, \dots))$$

Is (f_n) pointwise convergent?

$$f_n(x)=x^n \xrightarrow[n\to+\infty]{\text{Pointwisely}} f(x):=\begin{cases}0, & x\in[0,1)\\ 1, & x=1\end{cases} \in B([0,1],\mathbb{R})$$

Does (f_n) converge uniformly to f? We have that

$$d_{sup}(f_n,f)=\sup_{x\in[0,1]}|f_n(x)-f(x)|=$$

$$=\sup_{x\in[0,1]}|x^n-f(x)|=1\underset{n\to+\infty}{\not\to}0$$

$$\sup_{x\in[0,1)}|x^n|$$

$$\boxed{\begin{aligned}&\sup_{x\in[0,1]}|x^n-f(x)|\\ &=\sup_{x\in[0,1)}|x^n-f(x)|=\\ &=\sup_{x\in[0,1)}|x|^n=\\ &=\left(\sup_{x\in[0,1)}|x|\right)^n=1^n=1\\ &\qquad\forall n\in\mathbb{N}\end{aligned}}$$

hence f_n does not converge to f uniformly (this directly informs us also that (f_n) does not have a

Uniform limit) – hence this also constitutes an example of a divergent sequence. $\square$

Sufficient Conditions for Uniform conv.?
Pointwise conv. + ?

Sequential Convergence and Metrics Comparison

$X \neq \emptyset$, d_1, d_2 are metrics definable on X,

$\exists c > 0: \quad d_1 \leq c\, d_2.$

Let (x_n), $(x_n \in X\ n \in \mathbb{N})$ and $x_n \to x$ w.r.t. d_2.

Hence $d_2(x_n, x) \to 0$ as $n \to +\infty$. We know that

for all $n \in \mathbb{N}$ $\quad d_1(x_n, x) \leq c\, d_2(x_n, x)$ $\to 0$

$\to 0$ as $n \to +\infty$ since $d_1 \geq 0$.

Hence if $x_n \to x$ w.r.t. d_2 then $x_n \to x$ w.r.t. d_1 $\square$