Lecture 8

- Sequential Convergence
- Continuity
- Application: Approximation of optimization problems

[redacted] [redacted]

Definition denial:
$\exists \varepsilon>0 : \forall n^* \in \mathbb{N}\ \exists n \geq n^* : x_n \notin O_d(x,\varepsilon)$

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).

$\forall \varepsilon>0, \exists n^*(\varepsilon) \in \mathbb{N} : x_n \in O_d(x,\varepsilon)\ \forall n \geq n^*(\varepsilon)$

- 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. $\square$

Remark: In pseudo-metric spaces uniqueness may fail.

$\mapsto$ Could you construct an example?

Lemma (Boundedness). 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 $x_{n_1}, x_{n_2}, \ldots, x_{n_k} \notin O_d[x, 1]$ for some $k > 0$.

Since k is finite $0 < \max\big(\underbrace{d(x, x_{n_1}), d(x, x_{n_2}), \ldots, d(x, x_{n_k}), 1}_{k+1 \text{ reals}}\big) < +\infty$

Hence we can choose $\delta := \max(d(x, x_{n_1}), d(x, x_{n_2}), \ldots, d(x, x_{n_k}), 1)$ 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 $i = 1, \ldots, k$ we have that $d(x, x_{n_i}) \leq \max(d(x, x_{n_1}), \ldots, d(x, x_{n_k}), 1) = \delta$ hence $x_{n_i} \in O_d[x, \delta]$ $\forall i = 1, \ldots, k$. Hence $x_n \in O_d[x, \delta]$ $\forall n \in \mathbb{N}$.

Hence (x_n) is bounded. $\square$

Remark. The converse does not generally hold.

Does convergence imply total boundedness?

$\{x_0, x_1, x_2, \dots, x_n, \dots\} \subseteq X$

— Of course it is! (x_n) and $x_n \to x \in X$ — it is totally bounded (w.r.t. d)

Let $\varepsilon > 0$, consider $O_d(x, \varepsilon)$. $\exists\, n^*(\varepsilon)$:

$\{x_{n^*(\varepsilon)}, x_{n^*(\varepsilon)+1}, \dots\} \subseteq O_d(x, \varepsilon)$.

Also,

$$x_0 \in O_d(x_0, \varepsilon)$$

$$x_1 \in O_d(x_1, \varepsilon)$$

$$\vdots$$

$$x_{n^*(\varepsilon)-1} \in O_d(x_{n^*(\varepsilon)-1}, \varepsilon)$$

$n^*(\varepsilon)+1$ members

hence $\{O_d(x, \varepsilon), O_d(x_i, \varepsilon), i = 0, \dots, n^*(\varepsilon)-1\}$

is finite and covers $(x_0, x_1, \dots, x_n, \dots)$

Hence (x_n) is totally bounded since ε is arbitrary.

Convergence $\Rightarrow$ Total boundedness $\Rightarrow$ 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 $1, \dots,$ 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 [redacted] $1/\varepsilon - 1$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overset{\shortparallel}{n^*(\varepsilon)}$

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

$\forall x \in X$, $\varepsilon > 0$, $\quad O_{d_d}(x, \varepsilon) = \begin{cases} X, & \varepsilon > 1 \\ \{x\}, & \varepsilon \leq 1 \end{cases}$. 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. $(\frac{1}{n+1})$ 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 $(f_n) = (f_0, f_1, \dots, f_n, \dots)$ 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 $|f_n(x) - f(x)| \to 0 \quad \forall x \in Y^*$

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

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

Does the converse hold? The following counterexample 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

$$(\underbrace{x^0}_{\substack{\parallel \\ \text{Constant} \\ \text{at } 1}}, x, x^2, \dots, x^n, \dots))$$

$|f_n(x) - f(x)|\big|_{x=1}$

$= |1^n - 1|$

$= 0$

$Y^* = Y = [0,1]$

Is (f_n) pointwise convergent?

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

Pointwisely

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

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

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

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

$\sup\limits_{x\in[0,1]} |x^n - f(x)|$
$= \sup\limits_{x\in[0,1)} |x^n - f(x)| =$
$= \sup\limits_{x\in[0,1)} |x|^n =$
$= (\sup\limits_{x\in[0,1)} |x|)^n = 1^n = 1$
$\forall n \in \mathbb{N}$

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 $n \in \mathbb{N}$ $\quad \underbrace{d_1(x_n, x)}_{\to 0 \text{ as } n \to +\infty \text{ since } d_1 \geq 0.} \leq c \, \overbrace{d_2(x_n, x)}^{\to 0}$

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

$\longleftrightarrow$

$\varepsilon > 0$, $\quad O_{d_1}(x, \varepsilon) \supseteq O_{d_2}(x, \varepsilon/c) \xRightarrow{x_n \to x \; (\text{-}d_2)} x_n$

for almost every n

[redacted]

- We have shown that if $\exists c>0: d_1 \leq c d_2$, and $x_n \to x$ w.r.t. d_2 $\Rightarrow$ $x_n \to x$ w.r.t. d_1

(hence the set of convergent in X sequences w.r.t. d_2 $\subseteq$ to // w.r.t. d_1)

- Thereby, (why), if $\exists\, c_1, c_2 > 0$: $c_1 d_2 \leq d_1 \leq c_2 d_2$ then

$x_n \to x$ w.r.t. d_1 $\Leftrightarrow$ $x_n \to x$ w.r.t. d_2

(hence the aforementiones sets coincide)

E.g. $X = \mathbb{R}^p$, $d = d_I, d_{max}, d_{II}$ are not "distinguishable" w.r.t. sequential convergence.

Closeness [Sequential characterization]

Definition: $A \subseteq X$ is closed (w.r.t. d) iff for any sequence of elements of A ((x_n), $x_n \in A$ $\forall n \in \mathbb{N}$) that has a d-limit, its limit lies in A (i.e. every convergent sequence of elements of A converges inside A).

Dually: $A \subseteq X$ is open (w.r.t. d) iff A^c is closed.

- It is easy to show that: i. $X, \emptyset$ are simultaneously open and closed.

ii. If τ_c denotes the collection of the closed subsets of X then $X, \emptyset \in \tau_c$, arbitrary intersections of elements of τ_c lie in τ_c (i.e. arbitrary intersections of closed subsets of X are closed.

ii. (dual) If τ_o denotes the collection of the open subsets of X, then $X, \phi \in \tau_o$, arbitrary unions of elements of τ_o lie in τ_o (i.e. arbitrary intersections of open subsets of X are open).

τ_o is term topology on X generated by d.

E.g. It is possible to show that for any $x \in X$ and $\varepsilon > 0$ $O_d(x, \varepsilon)$ is open (i.e. $\in \tau_o$), and $\{x\}$, $O_d[x, \varepsilon]$ are closed (i.e. $\in \tau_c$).

Continuity of functions between Metric Spaces.

Suppose $f: X \to Y$, (X, d_X), (Y, d_Y) are Metric spaces.

Pair of Metric Spaces

- Suppose that (x_n) is a X-sequence (i.e. $x_n \in X, \forall n \in \mathbb{N}$). By evaluation via f we obtain the Y-sequence $(f(x_n))$. Continuity of f at a point involves "the comparison of" such sequences:

$(f(x_0) \overset{\in Y}{}, f(x_1) \overset{\in Y}{}, \ldots f(x_n) \overset{\in Y}{}, \ldots)$

Definition. Let $x \in X$, f is (d_Y/d_X) continuous at x iff $\forall (x_n)$, with $x_n \in X$ $\forall n \in \mathbb{N}$, that converges (w.r.t. d_X) to x, $(f(x_n))$ converges (w.r.t. d_Y) to $f(x)$. (f must not destroy the convergence of sequences in X that converge to x).

f is (d_Y/d_X-) continuous iff it is (d_Y/d_X-) continuous at x, $\forall x \in X$.

E.g. X, $d_X = d_d$ (discrete metric). We know that $x_n \to x$ w.r.t. d_d iff (x_n) is eventually constant at x (i.e. it is of the form $(x_0, x_1, \dots, x_k, x, x, x, \dots, x, \dots)$ for some k).

We then have that $(f(x_n))$ is eventually constant at $f(x)$ (since it must of the form $(f(x_0), f(x_1), \dots, f(x_k), f(x), \dots, f(x), \dots, f(x), \dots)$)

Thereby $f(x_n) \to f(x)$ w.r.t. d_Y (for any d_Y)

Hence if (X, d_X) is discrete EVERY function $X \to Y$ (with whatever d_Y) is continuous

Remark. We can characterize continuity at x using appropriately open (or closed) balls, or more generally using open (or closed) sets (see tutorial). For example it is possible to show that f is continuous at x iff

$$\forall \delta > 0, \exists \varepsilon > 0: \underbrace{f^{-1}(O_{d_Y}(f(x), \delta))}_{= \{x \in X : f(x) \in O_{d_Y}(f(x), \delta)\}} \subseteq O_{d_X}(x, \varepsilon).$$

ε-δ definition [further more primitive characterizations via open/closed sets]

Application: Approximation of Optimization Problems

(or some suitable subset)

- $X = B(Z, \mathbb{R})$, $d_X = d_{sup}$

$\neq \emptyset$

- $Y = \mathbb{R}$, $d_Y = d_u$

End of Lecture 20

- $\sup : B(Z,\mathbb{R}) \longrightarrow \mathbb{R}$ (for any $g \in B(Z,\mathbb{R})$

$\sup(g)$ exists as a real number hence sup is well defined as a function from $B(Z,\mathbb{R}) \to \mathbb{R}$). The evaluation of sup at $g \in B(Z,\mathbb{R})$ is equivalent to the Maximization of g.

- Is sup (possibly restricted to some subset of $B(Z,\mathbb{R})$ d_u/d_{sup}-continuous)? (Using the definition above this is equivalent to asking whether $\forall f_n, f \in B(Z,\mathbb{R})$ such that $d_{sup}(f_n, f) \to 0$, $\sup f_n \to \sup f$ w.r.t. d_u, i.e. $|\underline{\sup f_n} - \underline{\sup f}| \to 0$.)

If this is true then we can "approximate" the issue of Maximization of f by the Maximization of f_n.

- Dually we can work with inf (Exercise!!!)

- Let $g \in B(Z, \mathbb{R})$. It is possible that even though $\sup(g)$ is well-defined, that $\max g$ (i.e. $\max_{x \in Z} g(x)$)

$\sup(g) = \sup_{x \in Z}(g(x))$

does not exist due to that $\arg\max_{x \in Z} g(x)$ is empty.

However it is possible to prove that: if $\rho > 0$, $\exists\, x_{g,\rho} \in Z$:

$$\underline{g(x_{g,\rho})} \geq \sup_{x \in Z} g(x) - \rho = \sup_{x \in Z}(g(x) - \rho)$$

ρ can be perceived as an optimization error.

$x_{g,\rho}$ is an approximate maximizer

We denote the set of ρ-approximate maximizers of g with $\rho\text{-}\arg\max_{x \in Z} g(x)$.

$\boxed{\text{if } g \in B(Z,\mathbb{R})\ \rho\text{-}\arg\max(g) \neq \emptyset\ \forall \rho > 0. \text{ Maybe empty for } \rho = 0.}$

* Obviously if we let $\rho = 0$ then $\rho\text{-}\arg\max_{x \in Z} g(x) = \arg\max_{x \in Z} g(x)$ when it exists. Moreover if we let Z be a metric space that is also totally bounded and complete then it is possible to prove that $\exists\, g \in B(Z, \mathbb{R})$ such that $\rho\text{-}\arg\max g$ exist even for $\rho = 0$. (e.g. if g continuous)

next lecture

- We keep the fact that when $\rho > 0$, $\rho\text{-}\arg\max_{x \in Z} g(x) \neq \emptyset$ $\forall\, g \in B(Z, \mathbb{R})$.

Theorem. Given the above the $\sup : B(Z,\mathbb{R}) \to \mathbb{R}$ function is $\left(d_u / d_{\sup} - \right)$ continuous.

Proof. Suppose that continuity is not the case. This is equivalent to the existence of $f \in B(Z,\mathbb{R})$ and $f_n \in B(Z,\mathbb{R})$ such that $d_{\sup}(f_n, f) = \sup_{x\in Z} |f_n(x) - f(x)| \to 0$ (as $n \to +\infty$) yet $d_u(\sup f_n, \sup f) = |\sup_{x\in Z} f_n(x) - \sup_{x\in Z} f(x)| \not\to 0$.

The latter means that $\exists\ \delta > 0$:

$$\left|\sup_{x\in Z} f_n(x) - \sup_{x\in Z} f(x)\right| > \delta \quad \text{for an infinite number of } n \ (*)$$

[f.inf.n]
↳ abbreviation

Let $p_n > 0$ such that $p_n \to 0$ as $n \to +\infty$ (e.g. $p_n = \frac{1}{n+1}$) ✓

Since $p_n > 0\ \forall n$, we have that p_n-argmax f_n and p_n-argmax f are not empty. $x_{f_n, p_n} \in p_n\text{-argmax}(f_n)$, $x_{f, p_n} \in p_n\text{-argmax}(f)$.

Consider $|\sup f_n(x) - \sup f(x)| = |\underbrace{(\sup f_n(x) - p_n)}_{A_n} - \underbrace{(\sup f - p_n)}_{B_n}|$

We examine $|A_n - B_n|$:

α. $A_n \geq B_n$ $\quad |A_n - B_n| = A_n - B_n =$

$$= (\sup f_n(x) - p_n) - \underbrace{(\sup f - p_n)}_{\geq 0} \leq$$

$$f_n(x_{f_n,p_n}) - (\sup f - p_n) =$$

$$f_n(x_{f_n,p_n}) \pm f(x_{f_n,p_n}) - (\sup f - p_n) =$$

$$= \left(f_n(x_{f_n,p_n}) - f(x_{f_n,p_n})\right) + \cancel{\left(f(x_{f_n,p_n}) - \sup f\right)} + p_n$$

$$\leq \left(f_n(x_{f_n,p_n}) - f(x_{f_n,p_n})\right) + p_n$$

$$= \left| f_n(x_{f_n,p_n}) - f(x_{f_n,p_n}) + p_n \right|$$

tr. ineq.

$$\leq \left| f_n(x_{f_n,p_n}) - f(x_{f_n,p_n}) \right| + p_n$$

$$\leq \sup_{x \in Z} \left| f_n(x) - f(x) \right| + p_n = d_{\sup}(f_n, f) + p_n$$

b. $B_n > A_n \Rightarrow |A_n - B_n| = B_n - A_n$

then using a similar reasoning (exercise – swap in the previous f for f_n, f_n for f, and x_{f,p_n} for x_{f_n,p_n})

we have that $B_n - A_n \leq d_{\sup}(f_n, f) + p_n$

Hence by a, b we have that:

$$|\sup f_n - \sup f| \leq d_{\sup}(f_n, f) + p_n \quad \forall n \in \mathbb{N}$$

hence (*) $\Rightarrow$ $\cancel{d_{\sup}(f_n, f)}^{\to 0} + \cancel{p_n}^{\to 0} > \delta$ for all n if this is true this implies that $d_{\sup}(f_n, f) + p_n \not\to 0$

which is impossible since $d_{sup}(f_n, f) \to 0$ and $p_n \to 0$. $\square$

Remark: A dual result holds for the inf function (Exe!!!)

Remark. The previous imply that uniform convergence implies the approximability of optimization problems.
(it thus pays to study uniform convergence!!!)
↳ there exist dominated metrics to d_{sup} for which sup is continuous: e.g. hypo-convergence.

What about optimizers?

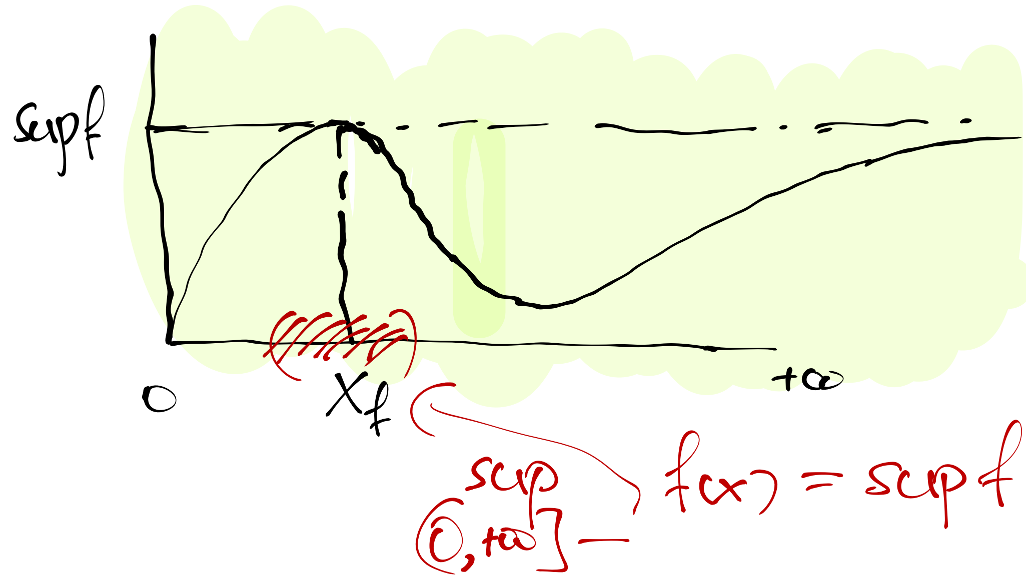
In order to study the issue of approximation of the (approximate) optimizers (those belong to Z) we need to be able to examine convergence in Z. Hence we now assume that Z, d_Z is a metric space. We assume the following:

i. $f_n, f \in B(Z, \mathbb{R})$, $\forall n \in \mathbb{N}$ such that $d_{sup}(f_n, f) \to 0$.

ii. There exists a unique $x_f \in Z$ such that $\{x_f\} = \underset{x \in Z}{\arg\max}\, f(x)$ and furthermore, x_f is "distinguishable" which means that, $\forall \varepsilon > 0$

$$\sup_{x \in O^c_{d_Z}(x_f, \varepsilon)} f(x) < f(x_f) = \sup_{x \in Z} f(x)$$

the distinguishability assumption precludes behaviours like the following:



[redacted] . . .

i.e. distinguishability precludes
approximability of sup f asymptotically

We will show that under the above, if $x_n \in \varepsilon_n$-argmax f_n and $\varepsilon_n \to 0$ then $d_2(x_n, x_f) \to 0$. $\square$