Lecture 9

Application: Approximability of optimization problems

Further Properties of Metric Spaces: Cauchy Sequences and Completeness

Application: Approximation of Optimization Problems

Structure Reminder

- $X = B(Z,\mathbb{R})$ (or some suitable subset) $\neq \emptyset$, $d_X = d_{sup}$
- $Y = \mathbb{R}$, $d_Y = d_u$
- sup: $B(Z,\mathbb{R}) \to \mathbb{R}$ (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. $|\sup f_n - \sup f| \to 0$.)

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!!!)

Auxiliary notion: Approximate Maximizers (relevant to numerical optimization)

- Let $g \in B(Z, \mathbb{R})$. It is possible that even though $\sup(g)$ is well-defined, that $\max g$ (i.e. $\max_{x \in Z} g(x)$) does not exist due to that $\arg\max_{x \in Z} g(x)$ is empty.

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

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

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

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 (next lecture) 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)

- 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 $\exists\ \delta > 0$:

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

[f.i.n.]

↳ abbreviation

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

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

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

We examine $|A_n - B_n|$:

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

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

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

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

$$= \left(f_n(x_{f_n,p_n}) - f(x_{f_n,p_n})\right) + \cancel{\underbrace{\left(f(x_{f_n,p_n}) - \sup f\right)}_{\leq 0}} + p_n$$

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

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

$$\overset{\text{tr. ineq.}}{\leq} \left| f_n(x_{f_n,p_n}) - f(x_{f_n,p_n}) \right| + p_n$$

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

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 $(*) \Rightarrow \cancel{d_{sup}(f_n, f)}^{\to 0} + \cancel{p_n}^{\to 0} > \delta$ finn. 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!!!)

$\hookrightarrow$ there exist dominated metrics to d_{sup} for which sup is continuous: e.g. hypo-convergence.

What about optimizers?

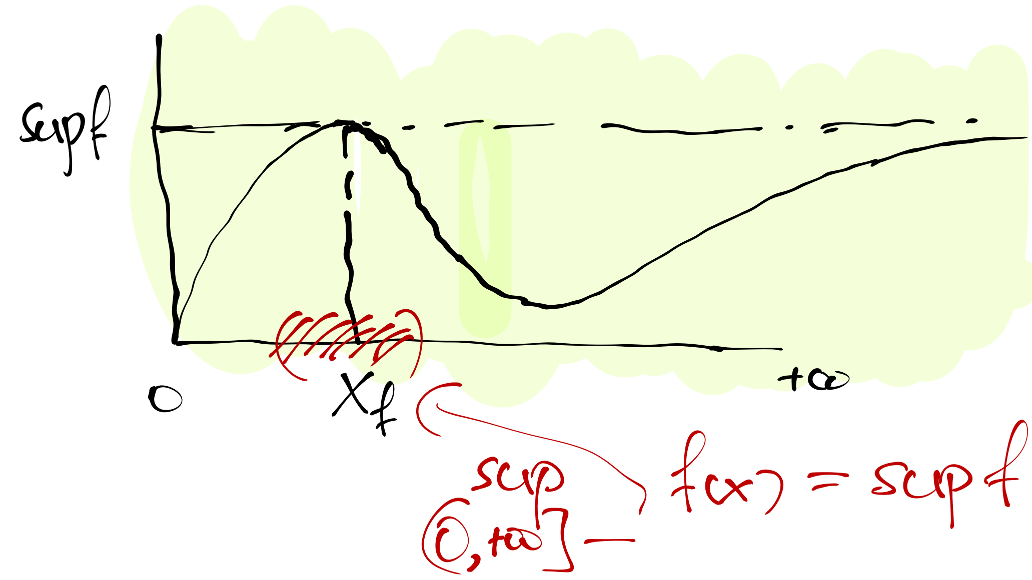
In order to study the issue of approximation of the (approximate) optimizers (those belong to Z) we need to be able to examine convergence in Z. Hence we now assume that Z, d_Z is a metric space. We assume the following:

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 $d_2(x_n, x_f) \to 0$. ∎

[redacted]

(Z,d_Z) a m.s.

Theorem (Convergence of (approximate) Maximizers): Suppose that

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

ii. $Z \ni x_f$ is the unique maximizer of f and $\forall \varepsilon>0$ $\sup\limits_{x\in O^c_{d_Z}(x_f,\varepsilon)} f(x) < f(x_f)$ (Uniqueness and distinguishability)

iii. $\rho_n>0$ with $\rho_n\to 0$.

Then if $x_{f_n,\rho_n} \in \rho_n\text{-}\arg\max\limits_{x\in Z} f_n(x)$, $d_Z(x_{f_n,\rho_n},x_f)\to 0$.

Remark: Uniqueness and distinguishability $\Rightarrow$ $\forall \varepsilon>0$ and if $d_Z(x_n,x_f)>\varepsilon$ ($\Leftrightarrow x_n\in O^c_{d_Z}(x_f,\varepsilon)$) finn $\Rightarrow$

$\exists\, \delta>0$: $f(x_f)-f(x_n)>\delta$ finn.

$\hookrightarrow$ i.e. it is independent of the n's for which $d_Z(x_n,x_f)>\varepsilon$.

the existence of δ is ensured by uniqueness. Its independence of the aforementioned n's is ensured by distinguishability.

Proof. Suppose that the above is not true, i.e. $\exists\, x_{f_n,\rho_n}\in$ $\rho_n\text{-}\arg\max f_n$ such that $d_Z(x_{f_n,\rho_n},x_f)\not\to 0$, i.e. $\exists\,\varepsilon>0$

$d_Z(X_{f_n,p_n}, X_f) > \varepsilon$ finn. $\overset{(ii)}{\Longrightarrow}$ $\exists\, \delta > 0$:

see Remark

$(*)$ $|f(x_{f_n,p_n}) - f(x_f)| > \delta$ finn. We have that:

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

$$= |(f(x_{f_n,p_n}) - f_n(x_{f_n,p_n})) + (f_n(x_{f_n,p_n}) - f(x_f))|$$

$$\overset{\text{tr.in}}{\leq} |f(x_{f_n,p_n}) - f_n(x_{f_n,p_n})| + |f_n(x_{f_n,p_n}) - f(x_f)|$$

$$\leq \sup_{x\in Z} |f(x) - f_n(x)| + \underbrace{\sup_{x\in Z} |f_n(x) - f(x_f)| + p_n}_{\text{from the proof of the previous theorem}}$$

$$= 2\, d_{sup}(f_n, f) + p_n. \qquad \forall n \in \mathbb{N} \quad (**)$$

Hence by $(*)$, $(**)$ we have that

$$2 d_{sup}(f_n, f) + p_n > \delta \text{ finn.}$$

Since $d_{sup}(f_n, f) \to 0$, and $p_n \to 0 \Rightarrow 2d_{sup}(f_n, f) + p_n \to 0$.

Hence it is impossible for $2 d_{sup}(f_n, f) + p_n > \delta$ finn to hold. $\square$

Remark: The above is quite general. It allows for the existence of optimization errors (φ_n) due to computational procedures of optimization, among others.

Exercise: State and prove the dual form of the above concerning minimization.

End of lecture 9

Further Non topological notions in Metric Spaces

1. Completeness

A. Cauchy Sequences (Βασικές Ακολουθίες)

Definition. (X,d) a m.s. and (x_n) is such that $x_n \in X$ $\forall n \in \mathbb{N}$. Then (x_n) is Cauchy iff $\forall \varepsilon>0$, $\exists n^*(\varepsilon)$: $\forall n,m \geq n^*(\varepsilon)$ $d(x_n,x_m)<\varepsilon$. [Definition Denial: $\exists \varepsilon>0$: $\forall n^* \in \mathbb{N}$, $\exists n,m \geq n^*$: $d(x_n,x_m)>\varepsilon$]

Remark: The above describes a property of "asymptotic concentration" for (x_n). $\forall \varepsilon>0$ $\exists n^*(\varepsilon)$: distance $<\varepsilon$

$$\left(x_0, x_1, x_2, \ldots, x_{n^*(\varepsilon)}, x_{n^*(\varepsilon)+1}, \ldots, x_n, x_m, \ldots\right)$$

distance $<\varepsilon$ distance $<\varepsilon$

etc.

Example. $X=\mathbb{R}$, $d=d_u$ $x_n=\frac{1}{n+1}$. Is (x_n) Cauchy?

We examine $(*)$ $|x_n-x_m|<\varepsilon \Leftrightarrow \left|\frac{1}{n+1}-\frac{1}{m+1}\right|<\varepsilon$. Let $m^*=\max(n,m)$

$n^*=\min(n,m) \Rightarrow m^*-n^*=k\in\mathbb{N} \Leftrightarrow m^*=n^*+k$.

$$\left|\frac{1}{n+1}-\frac{1}{m+1}\right| = \frac{1}{n^*+1}-\frac{1}{m^*+1} = \frac{m^*+1-(n^*+1)}{(n^*+1)(m^*+1)} = \frac{m^*-n^*}{(n^*+1)(n^*+1+k)} =$$

$$= \frac{k}{(n^*+1)^2+k(n^*+1)}.$$

$$(*) \Leftrightarrow \frac{k}{(n^*+1)^2+k(n^*+1)}<\varepsilon \Leftrightarrow \frac{k}{\varepsilon}<(n^*+1)^2+k(n^*+1) \Leftrightarrow$$

$$k(n^*+1)+(n^*+1)^2>\frac{k}{\varepsilon} \quad (**)$$ (when $k=0$ $n=m$ hence $(*)$ is in any case valid)

Hence suppose $k>0$, hence $(**) \Leftrightarrow n^*+1+\frac{(n^*+1)^2}{k}>\frac{1}{\varepsilon}$

$\Leftrightarrow n^* \cancel{+\frac{(n^*+1)^2}{k}} > \frac{1}{\varepsilon}-1$. $(***)$. The $(***)$ is valid whenever $n^*>\frac{1}{\varepsilon}-1$ and the latter holds

$\forall n^*\in\mathbb{N}$: $n^*\geq$ smallest natural greatest than or equal to $\frac{1}{\varepsilon}-1$

Hence if we define $n^*(\varepsilon):=$ smallest natural $\geq\frac{1}{\varepsilon}-1$ we have that $\forall n,m\geq n^*(\varepsilon)$ $|x_n-x_m|<\varepsilon$. Since ε is arbitrary this implies that $\left(\frac{1}{n+1}\right)$ is Cauchy. $\square$

Counter-Example. $X=\mathbb{R}$, $d=d_u$, $x_n=n$. Is (n) Cauchy?

Suppose that it is. Let $\varepsilon=1/3$. Then $\exists n^*(1/3)$: $\forall n,m \geq n^*(1/3)$

$|x_n - x_m| < 1/3$. Set $n := n^*(1/3)$, $m := n^*(1/3)+1$. So we must

have that $|x_n - x_m| = |n^*(1/3) - (n^*(1/3)+1)| = 1$ and

$1 < 1/3$, impossible. Hence (n) is not Cauchy. $\square$

Lemma. If (x_n) converges then it is Cauchy. (w.r.t. d)

Proof. Let $\varepsilon>0$. Let x be the limit of (x_n). $\exists\, n(\varepsilon/2)$:

$\forall n \geq n(\varepsilon/2)$ $d(x,x_n) < \varepsilon/2$. $\forall n,m \geq n(\varepsilon/2)$, $d(x_n,x_m) \leq \underbrace{d(x,x_n)}_{<\varepsilon/2} + \underbrace{d(x,x_m)}_{<\varepsilon/2}$

$< \varepsilon/2 + \varepsilon/2 = \varepsilon$. So $\forall n,m \geq n(\varepsilon/2)$, $d(x_n,x_m) < \varepsilon$. Set $n^*(\varepsilon) := n(\varepsilon/2)$.

The result follows since ε is arbitrary. $\square$

Remark. We have shown that the set of convergent sequences $\subseteq$ the set of Cauchy sequences in (X,d).

The reverse implication does not hold as the following example shows:

Example. $X=(0,1]$, $d=d_u$ restricted to $(0,1]$. $x_n = \frac{1}{n+1} \in (0,1]$

then. ((x_n) also belongs to the particular METRIC SUBSPACE of $(\mathbb{R}, d_u)$).

It is easy to see that $(\frac{1}{n+1})$ is Cauchy as a sequence inside $(0,1]$ — use the exact same arguments with the first example.

But $x_n \to 0 \notin X$, hence (x_n) cannot be considered as convergent in the particular X. Hence it is possible that a metric space (X,d) "does not contain the limits" of some of its Cauchy sequences. Hence it is possible that the set of convergent sequences $\subsetneq$ the set of Cauchy sequences in (X,d).

B. Complete Metric Spaces

Definition. (X,d) is complete iff every Cauchy sequence in X (w.r.t. d) is also convergent in X (w.r.t. d)

[i.e. iff the set of convergent sequences = the set of Cauchy sequences in (X,d)].

Example. It can be shown that $(\mathbb{R}, d_u)$ is complete.

Counter Example. $(\underbrace{[0,1], d_u}_{\text{restriction}})$ is not complete as the previous example shows. [hence restriction may destroy completeness]

Example. X is general $d = d_d$. (x_n) is Cauchy (w.r.t. d_d) iff it is eventually constant. (Suppose that it is eventually constant i.e. it is of the form $(x_0, x_1, \ldots, x_k, c, c, c, \ldots, c, \ldots)$ for some $k \in \mathbb{N}$ and $c \in X$. $\forall \varepsilon > 0$ $d(x_n, x_m) = 0 < \varepsilon$ $\forall n, m \geq k+1$, hence it is Cauchy) (Suppose that (x_n) is (d_d)-Cauchy then for $\varepsilon = 1/2$, $\exists n^*(1/2): \forall n,m \geq n^*(1/2), d_d(x_n, x_m) < 1/2$

$\Leftrightarrow x_n = x_m \ \forall n,m \geq n^*(1/2)$ hence (x_n) is eventually constant). We have already proven that (x_n) is $(d_d\text{-})$ convergent iff it is eventually constant. Hence

Set of Cauchy Sequences = Set of eventually constant sequences

= Set of convergent sequences in (X, d_d).

Hence every discrete space is complete!

(Remember Cauchyness and Completeness also depend on the metric!)

When is the completeness property inherited by metric subspaces?

Let (X,d) is a complete metric space. $A \subseteq X$, we consider the metric subspace (A, d). The subspace is complete
$A \neq \emptyset$
restriction

iff A is a closed subset of X.

Proof. (if (A,d) is complete then A is closed) Let (x_n) be such that $x_n \in A \ \forall n \in \mathbb{N}$, and $x_n \to x \in X$. Since (x_n) is convergent in X it is Cauchy in X. Hence it is Cauchy in A. Since (A,d) is complete (x_n) converges in A. Hence by the uniqueness of limits $x \in A$. Since (x_n) is arbitrary A is closed.

(if A is closed then (A,d) is complete). Suppose that (x_n) is Cauchy in A. Hence (x_n) is Cauchy in X. Since (X,d) is complete then (x_n) converges in X. Since A is closed (x_n) converges in A. Since (x_n) is arbitrary every Cauchy sequence in A converges in A. Hence (A,d) is complete. $\square$

[completeness inheritance]

Example. Let $X = B(Z,Y)$, (Y,d_Y), $d^Y_{sup}(f,g) = \sup_{x\in Z} d_Y(f(x),g(x))$. It is possible to show that if Y, d_Y is complete then $(B(Z,Y), d^Y_{sup})$ is complete. We know that $(\mathbb{R}, d_u)$ is complete then by the previous shows that $(B(Z,\mathbb{R}), d_{sup})$ is complete. (see tutorial for the proof).

C. Cauchyness – Completeness and Metrics Comparison

X general, d_1, d_2 metrics and $\exists c>0 : d_1 \leq c d_2$.

- If (x_n) is Cauchy w.r.t. d_2 then it is Cauchy w.r.t. d_1 (Exercise)
- Hence set of Cauchy sequences in $(X,d_2) \subseteq$ Set of Cauchy sequences in (X,d_1) $\square$