Sequencial characterization of Closeness.

Lemma. $A \subseteq X$ is d-closed iff for any $(x_n)_{n\in\mathbb{N}}$, with $x_n \in A$ $\forall n\in\mathbb{N}$, with $x = d\text{-}\lim x_n$ then $x \in A$.

Proof. $(\Rightarrow)$ Suppose that A is not d-closed, i.e. A' is not d-open. Then $\exists\, x \in A'$ such that $O_d(x,\varepsilon)\cap A \neq \emptyset$ $\forall \varepsilon>0$. For any $n\in\mathbb{N}$, let $x_n \in O_d(x, 1/_{n+1})\cap A \Rightarrow x_n \in A$ $\forall n\in\mathbb{N}$. Then if $\varepsilon>0$ $x_n \in O_d(x,\varepsilon)$ $\forall n \geq \frac{1}{\varepsilon}-1$ (why?)*. Hence $x = d\text{-}\lim x_n$. But due to the assumption $x\in A$. Contradiction since also $x \in A'$.
$(\Leftarrow)$. Suppose that A is d-closed, and let $(x_n)_{n\in\mathbb{N}}$ be a d-convergent sequence. If $x\in A'$, and since A' is d-open, $\exists\, \varepsilon>0$: $O_d(x,\varepsilon)\cap A = \emptyset$. Since $x_n \to x$, almost every member of $(x_n)_{n\in\mathbb{N}}$ must lie in $O_d(x,\varepsilon)$. This is a contradiction, since $x_n \in A$ $\forall n\in\mathbb{N}$. □

Remark. 1. The previous lemma gives a characterization of closeness (and thereby of openess) using the behavior of convergent sequences. It will be useful in what follows.

2. The $(\Rightarrow)$ part crucially depends on first countability (why?). In more general topologies the limiting behavior of more complex notions (topological nets) is needed.

3. The $(\Leftarrow)$ part would be true in any topological space. □

Sequential Convergence in finite Product Spaces

Prove the following: For I a finite index set,

* such n exist, since $n^*(\varepsilon)$ = the smallest $n \geq \frac{1}{\varepsilon}-1$ is well defined. Hence the property holds for almost every $n\in\mathbb{N}$.

(X_i, d_i) metric spaces $\forall i \in I$, $X := \prod_{i \in I} X_i$, $(x_n)_{n \in \mathbb{N}}$, with $x_n \in X$ $\forall n \in \mathbb{N}$, $x = d_\rho\text{-}\lim x_n$ iff $x_i = d_i \lim x_{i,n}$ $\forall i \in I$, $\rho = \prod_{\max}, \prod_I, \prod_{1,1}$. $\square$

Continuity of Functions Between Metric Spaces

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

Definition [Continuity at a point] f is d_Y/d_X-continuous at $x \in X$, iff $\forall (x_n)_{n \in \mathbb{N}}$, $x_n \in X$ $\forall n \in \mathbb{N}$, with $x = d_X\text{-}\lim x_n$ then $d_Y\text{-}\lim (f(x_n)) = f(x)$. $\square$

Hence continuity at a point is essentially invariance w.r.t. convergence to this point. It obviously is a topological notion (why?) and it also characterized by the behavior of the preimage set function f^{-1} on d_Y-open sets.

Lemma [Open Set Characterization] f is d_Y/d_X-continuous at x iff either one of the following equivalent conditions hold:

1. $\forall \delta > 0, \exists \varepsilon(\delta) : f(O_{d_X}(x, \varepsilon)) \subseteq O_{d_Y}(f(x), \delta)$.
2. If $A \in \tau_{d_Y}(f(x)) \Rightarrow \exists B \in \tau_{d_X}(x) : B \subseteq f^{-1}(A)$.

Proof. • 1. $(\Rightarrow)$ For $x_n \to x$ consider $(f(x_n))_{n \in \mathbb{N}}$. Then if $\delta > 0$, we have that $x_n \in O_{d_X}(x, \varepsilon(\delta))$ $\forall n \geq n^*(\varepsilon)$. But then $f(x_n) \in f(O_{d_X}(x, \varepsilon(\delta)))$ $\forall n \geq n^*(\varepsilon) \Rightarrow f(x_n) \in O_{d_Y}(f(x), \delta)$ $\forall n \geq n^*(\varepsilon)$ and thereby since δ is arbitrary, then $f(x_n) \to f(x)$.

• $(\Leftarrow)$ Suppose that f is d_Y/d_X-continuous at x. Suppose that $\exists \delta > 0 : \forall \varepsilon > 0$, $f(O_{d_X}(x, \varepsilon)) \cap O'_{d_Y}(f(x), \delta) \neq \emptyset$.

This implies that $\forall n \in \mathbb{N}$, $f(O_{d_X}(x, \frac{1}{n+1})) \cap O'_{d_Y}(f(x), \delta) \neq \emptyset$.

For each $n \in \mathbb{N}$ choose $x_n \in O_{d_X}(x, \frac{1}{n+1}) \cap f^{-1}(O'_{d_Y}(f(x), \delta))$.

$x_n \to x$ (why?), yet $f(x_n) \in O'_{d_Y}(f(x), \delta)$ $\forall n \in \mathbb{N}$ which implies that $f(x_n) \not\to f(x)$ which is a contradiction.

- $2 \Rightarrow 1$. For $\delta > 0$, choose $A = O_{d_Y}(f(x), \delta)$. Then $\exists B \in \tau_{d_X}(x) : B \subseteq f^{-1}(A)$. Since $B \in \tau_{d_X}(x)$ $\exists \varepsilon > 0$:

$$O_{d_X}(x, \varepsilon) \subseteq B \subseteq f^{-1}(A) = f^{-1}(O_{d_Y}(f(x), \delta)) \Rightarrow$$
$$f(O_{d_X}(x, \varepsilon)) \subseteq f[f^{-1}(O_{d_Y}(f(x), \delta))] \subseteq O_{d_Y}(f(x), \delta).$$

- $1 \Rightarrow 2$. Suppose that $\exists A \in \tau_{d_Y}(f(x)) : \forall B \in \tau_{d_X}(x)$, $B \cap (f^{-1}(A))' \neq \emptyset$. Hence $\forall \varepsilon > 0$ $O_{d_X}(x, \varepsilon) \cap (f^{-1}(A')) \neq \emptyset \Rightarrow f[O_{d_X}(x, \varepsilon) \cap f^{-1}(A')] \neq \emptyset$ $\Rightarrow f(O_{d_X}(x, \varepsilon)) \cap f[f^{-1}(A')] \neq \emptyset \Rightarrow f(O_{d_X}(x, \varepsilon)) \cap A' \neq \emptyset$. Since $A \in \tau_{d_Y}(f(x))$, $\exists \delta > 0 : O_{d_Y}(f(x), \delta) \subseteq A$. Hence $\forall \varepsilon > 0$

$f(O_{d_X}(x, \varepsilon)) \cap O'_{d_Y}(f(x), \delta) \neq \emptyset$ contradicting 1. $\square$

Remarks. 1. The condition 2 involves only open neighborhoods, hence defines continuity at a point for general topological spaces. 2. The previous have analogous representations w.r.t. closed balls and closed neighborhoods (Derive them!)

The global extension of the previous notion concerns the following definition.

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

The following lemma again characterizes continuity for general topological spaces.

Lemma. f is d_Y/d_X-continuous iff $\forall A\in\tau_{d_Y}$, $f^{-1}(A)\in\tau_{d_X}$.
Proof. [Prove it]

Example. Suppose that $d_X=d_\delta$. Then any $f:X\to Y$ is d_Y-d_δ continuous, since $f^{-1}(A)\subseteq X$, $\forall A\in\tau_{d_Y}$, $\tau_{d_X}=2^X$ and due to the previous lemma.

Continuity and Comparison of Metrics

Lemma. Suppose that f is d_{2Y}/d_X continuous at x, and $\exists c>0: d_{1Y}\leq c\, d_{2Y}$ (as functions). Then f is also d_{1Y}/d_X-continuous.
Proof. We have already proven that in this framework if $y=d_{2Y}$-$\lim(y_n) \Rightarrow y=d_{1Y}$-$\lim(y_n)$. Combine this with the sequential definition of continuity. $\square$

Corollary. If for $c_1,c_2>0$ $c_1d_2\leq d_1\leq c_2d_2$ (as functions) then f is d_{1Y}/d_X-continuous iff it is d_{2Y}/d_X-continuous.

Application: Approximation of Optimization Problems [Corrected and updated 04/2017]

Consider $(B(Y,\mathbb{R}),d_{sup})$ for $Y\neq\emptyset$. Notice that $\sup_{x\in Y} f(x)\in\mathbb{R}$ $\forall f\in B(Y,\mathbb{R})$. For $p\geq 0$ set

$$K^p:=\left\{f\in B(Y,\mathbb{R}) : \forall\, 0<u\leq p\ \exists x_{u,f} : f(x_{u,f})\geq \sup_{x\in Y} f(x)-u\right\}.$$

Example. If (Y,d_Y) is a totally bounded and complete metric space the it is possible to prove that $K^0\neq\emptyset$ (Consider this as a pending example also depending on notions that are to be examined, e.g. completeness).

Lemma. If for $p\geq 0$, $K^p\neq\emptyset$, then $\sup: K^p\to\mathbb{R}$ is d/d_{sup}-continuous, where d is the usual metric on $\mathbb{R}$, and d_{sup} is the uniform metric restricted to $K^p\times K^p$. $\square$

Proof. Suppose that $f_n, f \in K^p$, $\forall n \in \mathbb{N}$ and that f is the uniform limit of $(f_n)_{n\in\mathbb{N}}$, i.e. $d_{sup}(f_n, f) \to 0$. Towards a contradiction suppose that $\sup_{x\in Y} f_n(x) \not\to \sup_{x\in Y} f(x)$ [in what follows abbreviate the expression

[1] for an infinite subset of $\mathbb{N}$ with [finn]].

The previous hypothesis is equivalent to the existence of $\delta > 0$: $\left|\sup_{x\in Y} f_n(x) - \sup_{x\in Y} f(x)\right| > \delta$ [finn].

Correction: For $p_n \begin{cases} = 0 & \text{if } p = 0 \\ \in (0, p] & \text{if } p > 0 \end{cases}$ and such that $p_n \to 0$ as $n \to \infty$ we have that

$$\left|\sup_{x\in Y} f_n(x) - \sup_{x\in Y} f(x)\right| = \left|\left(\sup_{x\in Y} f_n(x) - p_n\right) - \left(\sup_{x\in Y} f(x) - p_n\right)\right| = |A_n - B_n| = (*)$$

For x_n: $f_n(x_n) \geq \sup_{x\in Y} f_n(x) - p_n$, and x_n^*: $f(x_n^*) \geq \sup_{x\in Y} f(x) - p_n$, which exist due to the definition of K^p and p_n, notice that

a. if $A_n - B_n \geq 0$, then $0 \leq A_n - B_n \leq f_n(x_n) - \left(\sup_{x\in Y} f(x) - p_n\right) = [f_n(x_n) - f(x_n)]$
$+ \underbrace{\left(f(x_n) - \sup_{x\in Y} f(x)\right)}_{\leq 0} + p_n \leq f_n(x_n) - f(x_n) + p_n = |f_n(x_n) - f(x_n) + p_n| \leq |f_n(x_n) - f(x_n)| + p_n$
$\leq \sup_{x\in Y} |f_n(x) - f(x)| + p_n$

b. if $B_n - A_n \geq 0$, then analogously (show it!) $0 \leq B_n - A_n \leq \ldots \leq \sup_{x\in Y} |f_n(x) - f(x)| + p_n$.

Hence $|A_n - B_n| \leq d_{sup}(f_n, f) + p_n$ and thereby we obtain that under the hypothesis that $\sup f_n \not\to \sup f$ that for some $\delta > 0$

$$d_{sup}(f_n, f) + p_n > \delta \quad \text{[finn]}$$

which establishes the required contradiction since we know that by assumption

$$\text{(why?)} \quad \left.\begin{cases} \exists n^* \in \mathbb{N} : & d_{sup}(f_n, f) < \delta/2 \quad \forall n \geq n^* \\ \exists n^{**} \in \mathbb{N} : & p_n < \delta/2 \quad \forall n \geq n^{**} \end{cases}\right\} \Rightarrow$$

$$d_{sup}(f_n, f) + p_n < \delta \quad \forall n \geq \max(n^*, n^{**}). \quad \square$$

Exercise. State and prove, by defining an analogous appropriate subset of $B(Y,\mathbb{R})$ the dual result about the relevant continuity of the inf functional.

The following result also enables convergence of sequences of approximate maximizers when the limit function has sufficient properties and the optimization error vanishes asymptotically.

Lemma. Suppose that for some $p \geq 0$, $K^p \neq \emptyset$. For $(f_n)_{n\in\mathbb{N}}$, $f_n, f \in K^p$ $\forall n \in \mathbb{N}$, $d_{\sup}(f_n, f) \to 0$. Furthermore, $\exists y \in Y$: $\{y\} = \underset{x\in Y}{\arg\max} f(x)$, and for a metric d_Y with which Y is endowed y is d_Y-distinguishable, i.e. $\forall \varepsilon > 0$ $\sup_{x \in O'_{d_Y}(y,\varepsilon)} f(x) < f(y)$. For $0 < p_n \leq p$, $\forall n \in \mathbb{N}$ let $y_n \in Y$: $f(y_n) = \sup_{x\in Y} f(x) - p_n$.

Then if $p_n \to 0$, $d_Y(y_n, y) \to 0$. □

Remark. Due to uniqueness and d_Y-distinguishability of y, and if $\bar{\mathbb{N}}$ is an is an infinite subset of $\mathbb{N}$, $\forall \varepsilon > 0$, $\exists \delta > 0$: if $d_Y(x_n, y) > \varepsilon$ $\forall n \in \bar{\mathbb{N}} \Leftrightarrow x_n \in O'_{d_Y}(y,\varepsilon)$ $\forall n \in \bar{\mathbb{N}}$ $\Rightarrow f(y) - f(x_n) \geq f(y) - \sup_{x \in O'_{d_Y}(y,\varepsilon)} f(x) > \delta > 0$, $\forall n \in \bar{\mathbb{N}}$. □

Proof. Suppose that $\exists \varepsilon > 0$: $d_Y(y, y_n) > \varepsilon$ for an infinite number of elements of $(y_n)_{n\in\mathbb{N}}$. Due to the previous **Remark** there exists

$\delta > 0$: $|f(y_n) - f(y)| > \delta$ [fim]

This implies that $|f_n(y_n) - f(y_n)| + |f_n(y_n) - f(y)| > \delta$ [fim] (why?)

$\Rightarrow \sup_{x\in Y} |f_n(x) - f(x)| + |\sup_{x\in Y} f_n(x) - p_n - \sup_{x\in Y} f(x)| > \delta$ [fim]

$\Rightarrow d_{sup}(f_n,f) + |\sup_{x\in Y} f_n(x) - \sup_{x\in Y} f(x)| + p_n > \delta$ [f.inn]

$\Rightarrow \exists\, \varepsilon_1,\varepsilon_2,\varepsilon_3>0$: $d_{sup}(f_n,f)>\varepsilon_1$ [f.inn], and/or
$|\sup_{x\in Y} f_n(x) - \sup_{x\in Y} f(x)|>\varepsilon_2$ [f.inn],
and/or
$p_n>\varepsilon_3$ [f.inn].

Either one is impossible since by assumptions $d_{sup}(f_n,f)\to 0$, $p_n\to 0$, and due to the previous lemma

$\sup_{x\in Y} f_n(x) \to \sup_{x\in Y} f(x)$. Contradiction. $\square$

Hence p_n-argmax f_n approximates argmax f, and the former optimization problem might be easier to solve. Furthermore a probabilistic "extension" of the previous result, in many cases provides consistency for M-estimators.

Notice also that the previous result encompasses the case where $p = p_n = 0\ \forall n\in\mathbb{N}$ if $K^{o}\neq\emptyset$.

Remark. In the above framework we saw that uniform convergence implies convergence of the relevant optimizers. It is possible to prove that analogous results would hold under a weaker topology called the topology of hypoconvergence (epiconvergence for minimizers).

[The notes are in a state of perpetual correction. They do not substitute the lectures. Please report any typos to stelios@aueb.gr or the course's e-class.]

Updates and further remarks:

1. In the previous we work under the convention that the system $0 < z \leq p$ means that $z = p$ when $p = 0$.

2. The definition of $x_{u,f}$: $f(x_{u,f}) \geq \sup_{x \in Y} f(x) - u$ implies that

$x_{u,f}$ is an u-approximate maximizer of f, i.e. $u\text{-}\arg\max_{x \in Y} f(x) \neq \emptyset$

where $u\text{-}\arg\max_{x \in Y} f(x) = \{ y \in Y : f(y) \geq \sup_{x \in Y} f(x) - u \}$. Approximate maximizers

may exist for $u > 0$, even if $\arg\max_{x \in Y} f(x) \left(:= 0\text{-}\arg\max_{x \in Y} f(x) \right) = \emptyset$. Notice that

if $u\text{-}\arg\max_{x \in Y} f(x) \neq \emptyset$ for some $u \geq 0$ then $u^*\text{-}\arg\max_{x \in Y} f(x) \neq \emptyset$ $\forall u^* \geq u$ since

$u\text{-}\arg\max_{x \in Y} f(x) \subseteq u^*\text{-}\arg\max_{x \in Y} f(x)$, $\forall u^* \geq u$. Hence a sufficient condition for $K^p \neq \emptyset$

is that $K^0 \neq \emptyset$.

3. The previous are easily extended to the dual concept of approximate minimizers (exercise!). Notice that the concept of approximate optimizers are particularly relevant to numerical optimization.

4. The final lemma above is more general to the one proven in the classroom since it is obtained for $p > 0$ as long as f has a maximizer that satisfies the conditions above. Notice that it enables the approximation of the latter by <u>selections</u> from $p_n\text{-}\arg\max_{x \in Y} f_n(x)$ (even in cases where $\arg\max_{x \in Y} f_n(x) = \emptyset$) as long as the "optimization error" p_n becomes asymptotically negligible.