3 Metric Spaces

3.1 Motivation

Consider a map $f : R \to R$. Then, we all know that

Definition 1 *f is said to be continuous at x if for every $\varepsilon > 0$ there exists some δ such that*

$$|f(x) - f(y)| < \varepsilon \text{ for every } y \text{ such that } |x - y| < \delta$$

Thinking about the proofs in R we note that:

1. Most any proof of a non trivial result used the fact that if $y \neq x$, then there is some $\varepsilon > 0$ such that $|x - y| > \varepsilon$. Hence, we suspect that it is crucial that the distance is zero if and only if $x = y$.
2. It seems sensible to require that distance is the same going from x to y as going back, which obviously is a property of $|\cdot|$ as well as $\sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$. It is not as obvious how crucial this property is.
3. Furthermore, both the Heine Borel result and the equivalence of Cauchy and standard convergence used the fact that the shortest path between two points is the direct path (as opposed to going from x to y via z being shorter). This is referred to as the triangle inequality.

In essence, these properties are the defining properties of a metric space.

Definition 2 *A metric space is a pair $M = \{X, \rho\}$, where X is a non empty set and $\rho : X \times X \to R$ (the metric) is a function satisfying:*

1. $\rho(x, y) \geq 0$ *for every* $(x, y) \in X \times X$
2. $\rho(x, y) = \rho(y, x)$ *for every* $(x, y) \in X \times X$
3. $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ *for every* $(x, y, z) \in X^3$

Definition 3 *Given metric spaces $M_1 = \{X_1, \rho_1\}$ and $M_2 = \{X_2, \rho_2\}$, a map $f : X_1 \to X_2$ and a point $x \in X_1$ we say that f is continuous at x (relative the metrics ρ_1 and ρ_2) if for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\rho_2 \left(f(x), f(y) \right) < \varepsilon \textit{ for every } y \textit{ such that } \rho_1 (x, y) < \delta.$$

Example 1 *Let X be arbitrary and let*

$$\rho(x, y) = \begin{cases} 0 & \textit{if } x = y \\ 1 & \textit{if } x \neq y \end{cases}$$

Obviously $\rho(x, y) \geq 0$ for every $(x, y) \in X \times X$ and $\rho(x, y) = \rho(y, x)$ for every $(x, y) \in X \times X$.

Finally, note that $\rho(x, z) + \rho(z, y) = \rho(x, x) + \rho(x, y) = \rho(x, y)$ if $x = z$ and (same argument) $\rho(x, z) + \rho(z, y) = \rho(x, y)$ if $y = z$, whereas

$$\rho(x, z) + \rho(z, y) = 2 > 1 \geq \rho(x, z)$$

if z is distinct from x and y. Hence, $(X, \rho\ )$ is a metric space.

Example 2 *Let $X = R^n$ and*

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Clearly, non-negativity and symmetry holds. We thus need only to establish that

$$\sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} + \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2}$$

or

$$\begin{aligned} \sum_{i=1}^{n} (x_i - y_i)^2 &\leq \left(\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} + \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2} \right)^2 \\ &= \sum_{i=1}^{n} (x_i - z_i)^2 + 2\sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2} + \sum_{i=1}^{n} (z_i - y_i)^2 \end{aligned}$$

But

$$\begin{aligned}\sum_{i=1}^{n}(x_i-y_i)^2 &= \sum_{i=1}^{n}(x_i-z_i+(z_i-y_i))^2\\ &= \sum_{i=1}^{n}\left[(x_i-z_i)^2+2(x_i-z_i)(z_i-y_i)+(z_i-y_i)^2\right]\\ &= \sum_{i=1}^{n}(x_i-z_i)^2+2\sum_{i=1}^{n}(x_i-z_i)(z_i-y_i)+\sum_{i=1}^{n}(z_i-y_i)^2\end{aligned}$$

Hence, we must show that

$$\sum_{i=1}^{n}(x_i-z_i)(z_i-y_i)\leq\sqrt{\sum_{i=1}^{n}(x_i-z_i)^2}\sqrt{\sum_{i=1}^{n}(z_i-y_i)^2}$$

Let $r_i=(x_i-z_i)$ *and* $s_i=(z_i-y_i)$. *Then, we want to establish that*

$$\left(\sum_{i=1}^{n}r_is_i\right)\leq\sqrt{\left(\sum_{i=1}^{n}r_i^2\right)\left(\sum_{i=1}^{n}s_i^2\right)}$$

But, for any $\lambda\in R$

$$\sum_{i=1}^{n}(r_i-\lambda s_i)^2\geq 0$$

Hence

$$f(\lambda)\equiv\left(\sum_{i=1}^{n}r_i^2\right)-2\lambda\sum_{i=1}^{n}r_is_i+\lambda^2\left(\sum_{i=1}^{n}s_i^2\right)\geq 0$$

for any λ *and*

$$2\sum_{i=1}^{n}r_is_i\leq\frac{1}{\lambda}\left(\sum_{i=1}^{n}r_i^2\right)+\lambda\left(\sum_{i=1}^{n}s_i^2\right)$$

for every $\lambda>0$.*Evaluate at*

$$\lambda=\frac{\sqrt{\left(\sum_{i=1}^{n}r_i^2\right)}}{\sqrt{\sum_{i=1}^{n}s_i^2}}$$

which gives

$$\begin{aligned}2\sum_{i=1}^{n}r_is_i &\leq \frac{\sqrt{\left(\sum_{i=1}^{n}s_i^2\right)}}{\sqrt{\sum_{i=1}^{n}r_i^2}}\left(\sum_{i=1}^{n}r_i^2\right)+\frac{\sqrt{\left(\sum_{i=1}^{n}r_i^2\right)}}{\sqrt{\sum_{i=1}^{n}s_i^2}}\left(\sum_{i=1}^{n}s_i^2\right)\\ &= 2\sqrt{\left(\sum_{i=1}^{n}r_i^2\right)\left(\sum_{i=1}^{n}s_i^2\right)},\end{aligned}$$

proving the triangle inequality. Hence, R^n *equipped with a metric given by Euclidean distance is a metric space.*

Example 3 *Let* $\{x_1, ..., x_K\} \subset R^2$ *be a set of towns on a map and let* $\rho(x_i, x_j)$ *denote the (shortest) millage between any two towns. This is different metric than the Euclidean...*

Example 4 $X = \{f : [a,b] \to R | f$ *is continuous*$\}$ *and*

$$\rho(x,y) = \sup_{a \le t \le b} |x(t) - y(t)|$$

which is an example of a functional space.

By the supremum property of R*, the* $\rho(x,y)$ *is well defined for every* $x, y \in X$*. Clearly, if* $x(t) = y(t)$, *then* $\rho(x,y) = 0$.*Also, if* $x \neq y$ *there exists* t^* *such that* $x(t^*) \neq y(t^*)$, *and* $\sup_{a \le t \le b} |x(t) - y(t)| \geq |x(t^*) - y(t^*)| > 0$, *so* $\rho(x,y) = 0$ *if and only if* $x = y$ *and* $\rho(x,y) \geq 0$ *always. Next*

$$\rho(x,y) = \sup_{a \le t \le b} |x(t) - y(t)| = \sup_{a \le t \le b} |y(t) - x(t)| = \rho(y,x),$$

so the distance measure is symmetric. Finally, we seek to demonstrate that

$$\sup_{a \le t \le b} |x(t) - y(t)| \leq \sup_{a \le t \le b} |x(t) - z(t)| + \sup_{a \le t \le b} |z(t) - y(t)|.$$

Suppose not, so that

$$\sup_{a \le t \le b} |x(t) - y(t)| > \sup_{a \le t \le b} |x(t) - z(t)| + \sup_{a \le t \le b} |z(t) - y(t)|$$

and let

$$\varepsilon = \sup_{a \le t \le b} |x(t) - y(t)| - \left[\sup_{a \le t \le b} |x(t) - z(t)| + \sup_{a \le t \le b} |z(t) - y(t)|\right] > 0$$

We have that there exists t_1

$$|x(t_1) - y(t_1)| > \sup_{a \le t \le b} |x(t) - y(t)| - \frac{\varepsilon}{3}$$

and

$$\begin{aligned} \sup_{a \le t \le b} |x(t) - z(t)| &\geq |x(t_1) - z(t_1)| \\ \sup_{a \le t \le b} |z(t) - y(t)| &\geq |z(t_1) - y(t_1)| \end{aligned}$$

Then

$$\begin{aligned}
\sup_{a\le t\le b}|x(t)-y(t)| &= \sup_{a\le t\le b}|x(t)-z(t)|+\sup_{a\le t\le b}|z(t)-y(t)|+\varepsilon \\
&\geq |x(t_1)-z(t_1)|+|z(t_1)-y(t_1)|+\varepsilon \\
&\geq |z(t_1)-y(t_1)| > \sup_{a\le t\le b}|x(t)-y(t)|-\frac{\varepsilon}{3}+\varepsilon \\
&\Rightarrow \\
0 &> \frac{2}{3}\varepsilon
\end{aligned}$$

a contradiction.

Example 5 $X=\{f:[a,b]\to R|f$ *is continuous*$\}$ *and*

$$\rho(x,y)=\sqrt{\int_a^b (x(t)-y(t))^2\,dt}$$

Positivity and symmetry obvious. For the triangle inequality note that

$$\begin{aligned}
\rho(x,y) &\leq \rho(x,z)+\rho(z,y)\Leftrightarrow[\rho(x,y)]^2\le[\rho(x,z)+\rho(z,y)]^2 \\
&\Leftrightarrow \\
&\quad \int_a^b (x(t)-y(t))^2\,dt \\
&= \int_a^b (x(t)-z(t))^2\,dt+2\int_a^b (x(t)-z(t))(z(t)-y(t))\,dt+\int_a^b (z(t)-y(t))^2\,dt \\
&\leq \int_a^b (x(t)-z(t))^2\,dt+2\sqrt{\int_a^b (x(t)-z(t))^2\,dt\int_a^b (z(t)-y(t))^2\,dt}+\int_a^b (z(t)-y \\
&\Leftrightarrow \\
\left[\int_a^b s(t)\,r(t)\,dt\right]^2 &\leq \int_a^b (s(t))^2\,dt\int_a^b (r(t))^2\,dt
\end{aligned}$$

This is the Schwarz inequality, which may be shown just like in the finite dimensional case.

That is

$$\begin{aligned}
\int_a^b (r(t) - \lambda s(t))^2 \, dt &\geq 0 \text{ for all } \lambda \Leftrightarrow \\
\int_a^b [r(t)]^2 \, dt - 2\lambda \int_a^b r(t) s(t) \, dt + \lambda^2 \int_a^b [s(t)]^2 \, dt &\geq 0 \text{ for all } \lambda \Rightarrow \\
\frac{1}{\lambda} \int_a^b [r(t)]^2 \, dt - 2 \int_a^b r(t) s(t) \, dt + \lambda \int_a^b [s(t)]^2 \, dt &\geq 0 \text{ for all } \lambda > 0 \Rightarrow \\
\frac{\sqrt{\int_a^b [s(t)]^2 \, dt}}{\sqrt{\int_a^b [r(t)]^2 \, dt}} \int_a^b [r(t)]^2 \, dt - 2 \int_a^b r(t) s(t) \, dt + \frac{\sqrt{\int_a^b [r(t)]^2 \, dt}}{\sqrt{\int_a^b [s(t)]^2 \, dt}} \int_a^b [s(t)]^2 \, dt &\geq 0 \\
&\Leftrightarrow \\
\sqrt{\int_a^b [r(t)]^2 \, dt} \sqrt{\int_a^b [s(t)]^2 \, dt} &\leq \int_a^b r(t) s(t) \, dt
\end{aligned}$$

Definition 4 *Given metric space* $M = \{X, \rho\}$ *and* $A \subset X$ *we call* $\{A, \widehat{\rho}\}$ *a metric subspace of* M *given that*

$$\rho(x, y) = \widehat{\rho}(x, y) \text{ for every } (x, y) \text{ in } A.$$

Example 6 *Consider* $f : R \to R$ *given by*

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is not rational} \end{cases}$$

Notice that $\{Q, |\cdot|\}$ *is a metric subspace of* $\{R, |\cdot|\}$ *and that* f *is continuous in* $\{Q, |\cdot|\}$ *but not in* $\{R, |\cdot|\}$

Example 7 *Let* $l^2 = \{X, \rho\}$ *where* X *is the set of all infinite sequences such that*

$$\sum_{i=1}^{\infty} x_i^2 < \infty$$

and

$$\rho(x, y) = \sqrt{\sum_i (x_i - y_i)^2}.$$

Since:

1. *$\sum_{i=1}^{n} x_i^2 \to \sum_{i=1}^{\infty} x_i^2$ and $\sum_{i=1}^{n} y_i^2 \to \sum_{i=1}^{\infty} y_i^2$ for any $x, y \in X$ it follows that $\sum_i (x_i - y_i)^2$ converges. This is because*

$$\begin{aligned} (x_i - y_i)^2 &= x_i^2 - 2x_iy_i + y_i^2 \leq x_i^2 + y_i^2 \Rightarrow \\ \sum_{i=1}^{\infty} (x_i - y_i)^2 &\leq \sum_{i=1}^{\infty} x_i^2 + \sum_{i=1}^{\infty} y_i^2 \end{aligned}$$

2. *Triangle inequality holds as*

$$\sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} + \sqrt{\sum_{i=1}^{n} (z_i - y_i)^2}$$

is true for any n, implying that

$$\sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^{\infty} (x_i - z_i)^2} + \sqrt{\sum_{i=1}^{\infty} (z_i - y_i)^2}$$

3.2 Open Sets in a Metric Space

Definition 5 *An open ball (or ε neighborhood), $B(x, \varepsilon)$, in a metric space $\{X, \rho\}$ is given by*

$$B(x, \varepsilon) = \{y \in X | \rho(x, y) < \varepsilon\}.$$

Example 8 *In the discrete metric where*

$$\rho(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

we have that

$$B(x, \varepsilon) = \begin{matrix} \{x\} & \text{if } \varepsilon < 1 \\ X & \text{if } \varepsilon \geq 1 \end{matrix}.$$

We can then rephrase the definition of continuity as:

Definition 6 *Given metric spaces $M_1 = \{X_1, \rho_1\}$ and $M_2 = \{X_2, \rho_2\}$, a map $f : X_1 \to X_2$ and a point $x \in X_1$ we say that f is continuous at x (relative the metrics ρ_1 and ρ_2) if given any open ball $B(f(x), \varepsilon)$ there is an open ball $B(x, \delta)$ such that*

$$f(B(x, \delta)) \subset B(f(x), \varepsilon)$$

Definition 7 *Let $A \subseteq X$ where $M = \{X, \rho\}$ is a metric space. We say that:*

1. *A is open if for every $y \in A$ there exists ε such that $B(x, \varepsilon) \subset A$.*
2. *A is closed if $X \backslash A$ is open.*

Example 9 *Let $X = \{x_1, x_2 ... x_K\}$ and*

$$\rho(x_i, x_j) = |i - j|$$

Now, for $\varepsilon < 1$ we have that

$$B(x_i, \varepsilon) = \{x_i\} \subset A$$

Hence any $A \subseteq X$ is open. Moreover, pick any $x_i \in X \backslash A$, then if $\varepsilon < 1$

$$B(x_i, \varepsilon) = \{x_i\} \subset X \backslash A.$$

So A is also closed. We conclude that any set in this metric space is both open and closed.

Definition 8 *Let $A \subseteq X$ where $M = \{X, \rho\}$ is a metric space. The closure of A, denoted $\overline{A}$, consists of every point $x \in X$ such that every open ball $B(x, \varepsilon)$ contains at least one point $y \in A$.*

Remark 1 *Unlike the definition of cluster point we don't require that $y \in A \backslash \{x\}$.*

Proposition 1 *The closure operator satisfies:*

1. $A \subset C \Rightarrow \overline{A} \subset \overline{C}$
2. $\overline{\overline{A}} = \overline{A}$
3. $\overline{A \cup C} = \overline{A} \cup \overline{C}$
4. $\overline{\varnothing} = \varnothing$

Proof. Part 1 is Obvious.

Part 2: If $x \in \overline{\overline{A}}$ it means that $x \in B(x_1, \varepsilon)$ for some $x_1 \in \overline{A}$ and every $\varepsilon > 0$. Consider $B(x_1, \varepsilon_1) \subset B(x_1, \varepsilon)$ with

$$\varepsilon_1 = \varepsilon - \rho(x, x_1).$$

As $x_1 \in \overline{A}$ there exists $x_2 \in A$ such that $x_2 \in B(x_1, \varepsilon_1)$. Moreover, the triangle inequality implies that

$$\begin{aligned} \rho(x, x_2) &\leq \rho(x, x_1) + \rho(x_1, x_2) < \rho(x, x_1) + \varepsilon - \rho(x, x_1) \\ &= \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary we conclude that $x \in \overline{A}$.

Part 3: By property 1

$$\begin{aligned} \overline{A} &\subset \overline{A \cup C} \\ \overline{C} &\subset \overline{A \cup C} \end{aligned}$$

Hence

$$\overline{A} \cup \overline{C} \subset \overline{A \cup C}$$

Also, if $x \in \overline{A \cup C}$ then, for every $\varepsilon > 0$, there exists $y \in A \cup C$ such that $y \in B(x, \varepsilon)$ meaning that for every $\varepsilon > 0$ there exists $y \in A$ such that $y \in B(x, \varepsilon)$ (in which case $y \in \overline{A}$) or there exists $y \in C$ such that $y \in B(x, \varepsilon)$ (in which case $y \in \overline{C}$). We conclude that every $x \in \overline{A \cup C}$ is also in $\overline{A} \cup \overline{C}$, so

$$\overline{A} \cup \overline{C} \supset \overline{A \cup C}$$

Hence

$$\overline{A} \cup \overline{C} = \overline{A \cup C}$$

Part 4: For any A

$$\overline{A} = \overline{A \cup \varnothing} = \overline{A} \cup \overline{\varnothing},$$

implying that $\overline{\varnothing} \subset \overline{A}$, which is possible for arbitrary A only if $\overline{\varnothing} = \varnothing$.

■

Proposition 2 *A is closed in $M = \{X, \rho\}$ if and only if $A = \overline{A}$.*

Proof. A is closed$\Rightarrow$ $X \backslash A$ is open, so for every $x \in X \backslash A$ there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \cap A = \varnothing$. Hence, $x \notin \overline{A}$. Which implies that

$$\overline{A} \subset A \subset \overline{A},$$

so $A = \overline{A}$. ■

Hence,

Corollary 1 *$\overline{A}$ is the smallest closed set containing A.*

Theorem 1 *In any metric space:*

1. *the union of any collection of open sets is open.*
2. *the intersection of a finite collection of open sets is open.*
3. *the intersection of an arbitrary collection of closed sets is closed.*
4. *the intersection of an arbitrary collection of closed sets is closed.*

Proof. Homework (just like in R). ■

3.3 Complete Metric Spaces

Definition 9 *A sequence $\langle x_n \rangle_{n=1}^{\infty}$ in metric space $M = \{X, \rho\}$is said to be a Cauchy sequence if given any $\varepsilon > 0$ there exists K such that*

$$\rho(x_n, x_m) < \varepsilon$$

for every $m, n \geq K$.

Definition 10 *A sequence $\langle x_n \rangle_{n=1}^{\infty}$ in metric space $M = \{X, \rho\}$ converges to $x \in X$ if given any $\varepsilon > 0$ there exists K such that*

$$\rho(x_n, x) < \varepsilon$$

for every $n \geq K$.

Proposition 3 *In a metric space M, every convergent sequence of real numbers $\langle x_n \rangle_{n=1}^{\infty}$ is a Cauchy sequence.*

Proof. Fix $\varepsilon > 0$ and let $\langle x_n \rangle_{n=1}^{\infty}$ have limit x^*. As $\langle x_n \rangle_{n=1}^{\infty}$ converges to x^* there exists $K\left(\frac{\varepsilon}{2}\right) \in N$ such that

$$\begin{aligned} \rho(x_n, x^*) &< \frac{\varepsilon}{2} \\ \rho(x^*, x_m) &< \frac{\varepsilon}{2} \end{aligned}$$

Since M is a metric space the triangle inequality holds, so

$$\rho(x_n, x_m) < \rho(x_n, x^*) + \rho(x^*, x_m) < \varepsilon,$$

which proves the claim as $\varepsilon > o$ was sarbitrary.

■

Definition 11 *A metric space $M = \{X, \rho\}$ is said to be complete if every Cauchy sequence converges (to a point in X).*

Example 10 *The real numbers (equipped with Euclidean metric) is a complete metric space $M = \{R, |\cdot|\}$. To see this, pick some $\varepsilon > 0$ and let K be such that $|x_n - x_m| < \varepsilon$ for every $n, m \geq K$. Then*

$$|x_n - x_K| < \varepsilon$$

for every n, implying that

$$\begin{aligned} x_n - x_K &< \varepsilon \text{ and } x_K - x_n < \varepsilon \\ x_K - \varepsilon &< x_n < x_K + \varepsilon \end{aligned}$$

for every $n \geq K$. The set

$$\{x_1, x_2, ..., x_K - \varepsilon, x_K + \varepsilon\}$$

is obviously bounded above and below and any bounds a, b to this set also bounds $x_n - \varepsilon$ and $x_n + \varepsilon$.

Because $\{x_n\}$ *is bounded we know from the Bolzano-Weierstrass Theorem that there exists a convergent subsequence* $\langle x_{k_n} \rangle_{k_n}$ *. Let the limit of the convergent subsequence be denoted by* x^**. Fix* $\varepsilon > 0$*. Since* $\langle x_{k_n} \rangle$ *converges to* x^* *there exists* $K_1 \in \{k_i\}_{i=1}^{\infty}$ *such that*

$$|x_n - x^*| < \frac{\varepsilon}{2}$$

for every $n \geq K_1$ *such that there exists an element in* $\{k_i\}_{i=1}^{\infty}$ *with* $k_i = n$*. This implies that*

$$|x_{K_1} - x^*| < \frac{\varepsilon}{2}.$$

But, since $\langle x_n \rangle_{n=1}^{\infty}$ *is Cauchy there exists* K_2 *such that*

$$|x_n - x_{K_1}| < \frac{\varepsilon}{2}$$

holds for every $n \geq K_2$*. Therefore*

$$\begin{aligned} |x_n - x^*| &= |(x_n - x_{K_1}) + (x_{K_1} - x^*)| \\ &\leq |x_n - x_{K_1}| + |x_{K_1} - x^*| < \varepsilon \end{aligned}$$

holds for every $n \geq K_2$*. We conclude that* x^* *is the limit of* $\langle x_n \rangle_{n=1}^{\infty}$*.* ■

Example 11 *Consider* $M = (R^n, \|\cdot\|)$ *for*

$$\|x - y\| \equiv \sqrt{\sum_i (x_i - y_i)^2}$$

Suppose that $\{x_k\}$ *is a Cauchy sequence. Then, for each* $\varepsilon > 0$ *there exists* K *such that*

$$\sum_i (x_{ik} - y_{ik})^2 < \varepsilon^2$$

for every $k \geq K$*. Hence*

$$(x_{ik} - y_{ik})^2 < \varepsilon^2 \Leftrightarrow |x_{ik} - y_{ik}| < \varepsilon$$

for every $k \geq K$ *and* $i \in \{1, .., n\}$*. By completeness of the reals there is some* x_i^* *and* K_i *such that*

$$|x_{ik} - x_i^*| < \frac{\varepsilon}{\sqrt{n}}$$

for every $k \geq K_i$*. Hence*

$$\sqrt{\sum_i (x_{ik} - x_i^*)^2} < \sqrt{n \left(\frac{\varepsilon}{\sqrt{n}}\right)^2} = \varepsilon.$$

for $k \geq \max\{K_1, .., K_n\}$*. Since* $\varepsilon > 0$ *was arbitrary we have proved completeness of* R^n*.*

3.4 Definition of Topological Spaces

We will not do much with general topological spaces for some time, but we need the terminology in order to characterize compactness for metric spaces.

Definition 12 *A topological space is a pair* $T = \{X, \mathcal{F}\}$, *where* X *is a non empty set and* $\mathcal{F}$ *is collection of subsets of* X *satisfying:*

1. $X, \varnothing \in \mathcal{F}$
2. *If* $A, B \in \mathcal{F}$, *then* $A \cap B \in \mathcal{F}$
3. *If* $\widehat{\mathcal{F}} = \left\{ \widehat{F_i} \right\}$ *is an arbitrary collection of sets from* $\mathcal{F}$, *then* $\cup_i \widehat{F_i} \in \mathcal{F}$

- $\mathcal{F}$ is called the open sets of T
- Note that open sets are defined in accordance with the most crucial properties PROVED for metric spaces.
- Pick any metric space and let $\mathcal{F}$ be the open sets of the metric space. This defines a topological space.
- Examples that are not metric spaces will have to wait.

recall the definition of compactness for R. Indeed, this is the general definition of compactness for any topological space (which includes metric spaces)

Definition 13 *An open cover of a set* A *in topological space* T *is a collection* $\mathcal{C} = \{C_i\}$ *of open sets such that* $A \subset \cup_{i \in I} C_i$

Definition 14 $\mathcal{C}'$ *is a subcover of* $\mathcal{C}$ *if each set* C_j *in* $\mathcal{C}'$ *also is in* $\mathcal{C}$ *and* $\mathcal{C}'$ *is a cover of* A.

Definition 15 *A in T is said to be compact if* ***every*** *open cover* $\mathcal{C}$ *has a finite open subcovering* $\mathcal{C}'$.

Note that we no longer have any notion of distance to refer to when thinking about continuity. Instead we define continuity directly in terms of open sets:

Definition 16 (Variant 1) *Consider* $f : X_1 \to X_2$, *where* $T_1 = \{X_1, \mathcal{F}_1\}$ *and* $T_2 = \{X_2, \mathcal{F}_2\}$ *are topological spaces. We say that*

1. f *is continuous at* $x_1 \in X_1$ *if for every* $B_2 \in \mathcal{F}_2$ *such that* $f(x_1) \in B_2$ *there exists some* $B_1 \in \mathcal{F}_1$ *such that* $x_1 \in B_1$ *and* $f(B_1) \subset B_2$.
2. f *is continuous if it it is continuous at every* $x_1 \in X_1$

An equivalent definition is:

Definition 17 (Variant 2) *Consider* $f : X_1 \to X_2$, *where* $T_1 = \{X_1, \mathcal{F}_1\}$ *and* $T_2 = \{X_2, \mathcal{F}_2\}$ *are topological spaces. We say that* f *is continuous if* $f^{-1}(B_2) \in \mathcal{F}_1$ *for every* $B_2 \in \mathcal{F}_2$ *.*

Proposition 4 *The two notions of continuity are equivalent*

Proof. [continuity 1⇒continuity 2] Pick any $B_2 \in \mathcal{F}_2$. Without loss, assume that $f^{-1}(B_2)$ is non-empty [empty set is open] and pick some $x_1 \in f^{-1}(B_2)$ .By **continuity 1** there exists $B_1(x_1) \in \mathcal{F}_1$ such that $f(B_1(x_1)) \subset B_2$. But, this is true for **any** $x_1 \in f^{-1}(B_2)$. So

$$f^{-1}(B_2) \subset \cup_{x_1 \in f^{-1}(B_2)} B_1(x_1)$$

since every $x_1 \in f^{-1}(B_2)$ is in at least one of the open sets in $\{B_1(x_1)\}_{x_1 \in f^{-1}(B_2)}$. Moreover, $\{B_1(x_1)\}_{x_1 \in f^{-1}(B_2)}$ is a collection of sets such that $f(B_1(x_1)) \subset B_2$ for every $x_1 \in f^{-1}(B_2)$, so

$$f^{-1}(B_2) \supset \cup_{x_a \in f^{-1}(B_2)} B_1(x_\alpha)$$

In conjunction we have that there exists a collection of sets $\{B_1(x_1)\}_{x_1 \in f^{-1}(B_2)}$ where each $B_1(x_1) \in \mathcal{F}_1$ such that

$$f^{-1}(B_2) = \cup_{x_a \in f^{-1}(B_2)} B_1(x_\alpha).$$

Hence,$f^{-1}(B_2)$ is open as a consequence of an arbitrary union of open sets being open.

[continuity 2⇒continuity 1] Pick an arbitrary $x_1 \in X_1$ and any $B_2 \in \mathcal{F}_2$ such that $f(x_1) \in B_2$. Pick $B_1 = f^{-1}(B_2)$ which is open by **continuity 2** and contains x_1 by construction. ■

Proposition 5 *If $f : X_1 \to X_2$ and X_1 is compact, then $f(X_1)$ is compact.*

Proof. Let $\mathcal{C}_2 = \{C_2^i\}_{i\in I}$ be an open cover of $f(X_1)$. Let $\mathcal{C}_1 = \{C_1^i\}_{i\in I}$ be given by

$$C_1^i = f^{-1}\left(C_2^i\right)$$

for every $i \in I$. We have that $\mathcal{C}_1$ is an open cover since:

1. Each C_1^i is open by continuity of f
2. For every $x_1 \in X_1$ there exists $i \in I$ such that $f(x_1) \in C_2^i$ and

$$x_1 = f^{-1}(f(x_1)) \subset f^{-1}\left(C_2^i\right) = C_1^i.$$

Since X_1 is compact there exists finite subcover

$$\left\{C_1^k\right\}_{k=1}^K = \left\{f^{-1}\left(C_2^k\right)\right\}_{k=1}^K$$

implying that

$$\left\{C_2^k\right\}_{k=1}^K$$

is a finite subcover of $\mathcal{C}_2 = \{C_2^i\}_{i\in I}$ [since $x_1 \in f^{-1}(C_2^i) \Leftrightarrow f(x_1) \in C_2^i$]

■

3.5 Compactness in Metric Spaces

In R we proved that that A is compact⇔ A is closed and bounded. For metric spaces the following two propositions show that closedness and boundedness are still properties that are very much related to compactness.

Proposition 6 *Let M be a compact metric space and suppose that $A \subset M$ is closed. Then, A is compact.*

Proof. Pick any open cover $\mathcal{C} = \{C_i\}$ of A. Since A is closed $M \backslash A$ is open. So

$$\widehat{\mathcal{C}} = \{M \backslash A, C_i\}$$

is a cover of M. By compactness of M there is a finite subcover $\left\{\widehat{C}_1, ..., \widehat{C}_K\right\}$ of

$$M = A \cup M \backslash A,$$

so $\left\{\widehat{C}_1, ..., \widehat{C}_K\right\}$ covers A. If $\widehat{C}_i = M \backslash A$ we can take it out from the list, implying that we are left with a finite subcovering of $\mathcal{C} = \{C_i\}$. ■

Definition 18 *A subset A of a metric space $M = \{X, \rho\}$ is bounded if there is some $x \in X$ and K in R such that $\rho(x, y) \leq K$ for every $y \in A$*

Proposition 7 *Let $M = \{X, \rho\}$ be a metric space and suppose that $A \subset M$ is compact. Then, A is bounded.*

Proof. Pick any $x \in A$ and consider the open cover $\{B(x, n)\}_{n=1}^{\infty}$ where $B(x, n)$ denotes the open ball around x with radius n. If A is compact there exists a finite subcover, which since

$$\cup_{n=1}^{K} B(x, n) = B(x, K)$$

implies that there exists some K such that $A \subset B(x, K)$ ■

It follows:

Corollary 2 *A continuous function from a compact space to a metric space is bounded.*

Proof. Continuity preserves compactness for all spaces and boundedness is necessary for compactness in a metric space ■

Corollary 3 (Weierstrass Maximum Theorem) *Let $T = \{X, \mathcal{F}\}$ be a topological space and let $A \subset X$ be compact and $f : X \rightarrow R$ be continuous. Then, there exists $x^* \in A$ such that $f(x^*) \geq f(x)$ for every $x \in A$.*

Remark 2 *Another way of saying this is that there exists V such that*

$$V = \max_{x \in A} f(x).$$

We often write

$$x^* \in \arg\max_{x \in A} f(x),$$

where we note that while the maximum is obviously uniquely determined, we need to be careful to observe that there may be multiple maximizers.

Proof. $f(A) \subset R$ is compact. By the Heine Borel theorem, $f(A)$ is bounded, so by the completeness of the real number system a supremum to $f(A)$ exists. Call the supremum $s \in R$. Again by the Heine Borel theorem, $f(A)$ is closed, so $R \backslash f(A)$ is open. Hence, if $s \notin f(A)$ there exists some $\varepsilon > 0$ such that $B(s, \varepsilon) \subset R \backslash f(A)$ implying that there is ε such that

$$s - \varepsilon > \sup f(A)$$

which contradicts that s is the least upper bound of $f(A)$. Thus $s \in f(A)$, which by definition of the direct image means that there exists $x \in A$ such that $f(x) = s$. ■

Theorem 2 $X_1 \times X_2$ *is compact if and only if* X_1 *and* X_2 *are compact.*

I will cheat in one step below to avoid discussing basis for topological spaces. For metric spaces this is much easier as we will show below that compactness has a nice characterization.

Proof. ($\Rightarrow$) Follows from the fact that continuity preserves compactness since $f : X_1 \times X_2 \rightarrow X_1$ given by

$$f(x_1, x_2) = x_1$$

for every (x_1, x_2) is continuous.

($\Leftarrow$) Let $\mathcal{C} = \{C_i\}$ be an open cover of $X_1 \times X_2$. Fix $x_1 \in X_1$ and observe that $\mathcal{C} = \{C_i\}$ covers $\{x_1\} \times X_2$.

We will base the proof on the intuitive claim below. To prove this requires some extra machinery that I want to avoid.

Claim *For every open set C_i such that $(x_1, x_2) \in C_i$ there exists $U(x_1)$ open in T_1 and $V(x_2)$ open in T_2 such that*

$$(x_1, x_2) \in U(x_2) \times V(x_2) \subset C_i$$

Since this is true for all x_2 we have a collection of sets $\{V(x_2)\}_{x_2 \in X_2}$ that is an open cover for T_2. Since T_2 is compact there exists a finite collection

$$\left\{V\left(x_2^i\right)\right\}_{i=1}^{K_2}$$

Let

$$U(x_1) = \cap_{i=1}^{K_2} U\left(x_2^i\right),$$

an open set which covers $\{x_1\} \times X_2$. Repeat process for every $x_1 \in X_1$ to get open cover of X_1 given by$\{U(x_1)\}_{x_1 \in X_1}$. As X_1 is compact there is a finite subcover

$$\left\{U\left(x_1^i\right)\right\}_{i=1}^{K_1}$$

Every point in $X_1 \times X_2$ is in some $U\left(x_1^i\right) \times V\left(x_2^j\right) \subset C_i$ so we can cover $X_1 \times X_2$ by a finite subcover ($K_1 \times K_2$ sets). ■

Theorem 3 *A set $A \subset R^n$ is compact if and only if it is closed and bounded.*

Proof. Necessity is immediate from the previous result as $R^n = R \times R^{n-1}$.

Pick any $A \subset R^n$ that is closed and bounded. Then, there exists $[a, b]$ such that

$$A \subset [a, b]^n.$$

$[a, b]^n$ is compact as every $[a, b]$ is compact by the Heine Borel theorem for R. Hence, A is a closed set in a compact metric space and therefore compact. ■

3.6 Sequential Compactness

Definition 19 *Let $A \subset X$ where $M = \{X, \rho\}$ is a metric space. We say that A is sequentially compact if every sequence $\langle x_n \rangle_{n=1}^{\infty}$ with $x_n \in A$ for every n has a convergent subsequence.*

Proposition 8 *Suppose that* $A \subset X$ *where* $M = \{X, \rho\}$ *is a metric space is compact. Then,* A *is sequentially compact.*

Proof. Suppose that A is compact and let $\langle x_n \rangle_{n=1}^{\infty}$ be a sequence with $x_n \in A$ for every n. We want to show that a convergent subsequence exists.

CASE 1: Suppose $V = \{x \in X | x = x_n \text{ for some } n\}$ is finite. Then there must be some constant subsequence $x_{n_k} = x \in B$ for every k in $\{n_k\}_{k=1}^{\infty}$, implying that $\{x_{n_k}\}$ is convergent

CASE 2: Suppose $V = \{x \in X | x = x_n \text{ for some } n\}$ is infinite. If no convergent subsequence exists it means that for every $y \in A$ there exists $\varepsilon > 0$ such that the set $V \cap B(y, \varepsilon)$ is finite [if infinite for every $\varepsilon > 0$ we can construct subsequence converging to y....you may want to think a bit about this].

$$\{B(y, \varepsilon)\}_{y \in A}$$

is an open cover of A and by compactness there exists finite subcover

$$\{B(y_1, \varepsilon), ..., B(y_K, \varepsilon)\}.$$

But each $B(y_i, \varepsilon)$ contains a finite set of points from V implying that

$$\cup_{i=1}^{K} B(y_i, \varepsilon)$$

contains a finite set of points in V which contradicts that $\cup_{i=1}^{K} B(y_i, \varepsilon)$ covers $A \supset V$. ■

Lemma 1 *Suppose that* A *is sequentially compact. Then, given any* $\varepsilon > 0$ *there exists a finite set* $\{a_1, ..., a_K\} \subset A$ *such that*

$$A \subset \cup_{i=1}^{K} B(a_i, \varepsilon)$$

Proof. For contradiction, pick a_1 arbitrarily and suppose that $\{a_1, ..., a_i\}$ have been picked so that

$$\rho(a_i, a_j) \geq \varepsilon$$

for all $i, j \leq k$ with $i \neq j$. By hypothesis we can then find $\varepsilon > 0$ so that $\rho(a_{i+1}, a_j) \geq \varepsilon$ for all $i, j \leq k$ with $i \neq j$. This defines a sequence so that

$$\rho(a_k, a_j) \geq \varepsilon$$

for every $k > j$, so $\{a_k\}_{k=1}^{\infty}$ is not Cauchy. Indeed, the inequality above implies that no subsequence is Cauchy as we will necessary have $\rho(a_{n_k}, a_j) \geq \varepsilon$ for every subsequence $\{n_k\}$ too, implying that no convergent subsequence can exist. ■

This is central enough for us to invent some terminology.

Definition 20 *In a metric space, A is totally bounded if given any $\varepsilon > 0$ there exists a finite set $\{a_1, ..., a_K\} \subset A$ such that*

$$A \subset \cup_{i=1}^{K} B(a_i, \varepsilon)$$

Definition 21 *Given open cover $\mathcal{C} = \{C_i\}_{i \in I}$ we say that $\varepsilon > 0$ is a Lebesgue number for $\mathcal{C}$ if given any $x \in A$ there exists $C(x)$ in $\mathcal{C}$ such that*

$$B(x, \varepsilon) \subset C(x).$$

Example 12 *Let $A = (0, 1)$ and take $\mathcal{C} = \left\{\left(\frac{1}{i}, 1\right)\right\}_{i=2}^{\infty}$. Fix $\varepsilon > 0$ and let $x = \varepsilon$. Then $B(\varepsilon, \varepsilon) = (0, 2\varepsilon)$. Hence there exists no i such that*

$$B(\varepsilon, \varepsilon) = (0, 2\varepsilon) \subset \left(\frac{1}{i}, 1\right),$$

which means that the open cover has no Lebesgue number.

Lemma 2 *Suppose that A is sequentially compact. Then any open cover has a Lebesgue number.*

Proof. Suppose not. Let $\mathcal{C} = \{C_i\}_{i \in I}$ be an open cover for which no Lebesgue number exists. Construct a sequence so that for every n the element x_n is such that $B\left(x_n, \frac{1}{n}\right)$ is not contained in any C_i. If such a point doesn't exist $\frac{1}{n}$ is a Lebesgue number for the cover.

A is sequentially compact, so there is a convergent subsequence $\langle x_{n_k} \rangle$ converging to some $x^* \in A$. $\mathcal{C}$ covers A so there is some C^* in the open cover such that $x^* \in C^*$. But then there is some integer m such that $B\left(x^*, \frac{2}{m}\right) \subset C^*$ and since $\langle x_{n_k} \rangle$ converges to x^* we have that $x_{n_k} \in B\left(x^*, \frac{1}{m}\right)$ for every $n_k \geq R$ for some $R \geq m$ (for indices that correspond to the subsequence). Let $y \in B\left(x_{n_k}, \frac{1}{n_k}\right)$. Then

$$\rho(x_{n_k}, y) \leq \frac{1}{m} \Rightarrow \rho(x^*, y) \leq \rho(x^*, x_{n_k}) + \rho(x_{n_k}, y) < \frac{1}{n_k} + \frac{1}{m} \leq \frac{2}{m},$$

implying that

$$B\left(x_{n_k}, \frac{1}{n_k}\right) \subset B\left(x^*, \frac{2}{m}\right) \subset C^*,$$

contradicting the hypothesis that $B\left(x_n, \frac{1}{n}\right)$ is not contained in any C_i. ■

Theorem 4 *Consider a metric space $M = \{X, \rho\}$. Then, $A \subset X$ is compact if and only if it is sequentially compact.*

Proof. We have shown that compactness⇒sequential compactness. For the other direction, let A be sequentially compact and let $\mathcal{C}$ be an open cover. Let $\varepsilon > 0$ be a Lebesgue number for the cover. Since A is totally bounded there is a finite set of points $\{a_1, ..., a_K\} \subset A$ such that $\{B(a_i, \varepsilon)\}_{i=1}^{K}$ cover A. By virtue of ε being a Lebesgue number there exists $C(a_i) \in \mathcal{C}$ such that $B(a_i, \delta) \subset C(a_i)$ for every $i \in \{1, ..., K\}$ in $\mathcal{C}$, implying that $\{C(a_i)\}_{i=1}^{K}$ is a finite subcover for A. ■

Theorem 5 *Consider a metric space $M = \{X, \rho\}$. Then, X is compact if and only if M is complete and X is totally bounded.*

Proof. (⇒)We've already established that compactness⇒sequential compactnesss⇒ X is totally bounded. Pick any Cauchy sequence $\langle x_n \rangle$ in X. Since compactnesssss⇒boundedness we know that there is a convergent subsequence (with limit in X). It is easy to establish that this implies that $\langle x_n \rangle$ must converge to the limit of the subsequence.

(⇐) Let $\langle x_n \rangle$ be an arbitrary sequence. Given each m let $\left\{x_1^m, ..., x_{K(m)}^m\right\}$ be a finite set of points such that $\left\{B\left(x_k^m, \frac{1}{m}\right)\right\}_{k=1}^{K(m)}$ covers X. Such a finite set exists for every m by the assumption that X is totally bounded. We note:

1. One ball in $\{B(x_k^1, 1)\}_{k=1}^{K(1)}$ must contain an infinite set of points from $\langle x_n \rangle$. Let B^1 denote such a ball and let $X^1 = \{x \in B^1 | x = x_n \text{ for some } n\}$

2. One ball in $\left\{B\left(x_k^2, \frac{1}{2}\right)\right\}_{k=1}^{K(2)}$ must contain an infinite set of points from X^1. Let B^2 denote such a ball and let

$$\begin{aligned} X^2 &= \left\{x \in B^2 | x \in X^1\right\} \\ &= \left\{x \in B^1 \cap B^2 | x = x_n \text{ for some } n\right\} \end{aligned}$$

3. One ball in $\left\{B\left(x_k^m, \frac{1}{m}\right)\right\}_{k=1}^{K(m)}$ must contain an infinite sets of points from X^{m-1}, where

$$X^{m-1} = \left\{x \in \cap_{j=1}^{m-1} B^j | x = x_n \text{ for some } n\right\}$$

so let B^m denote the ball and let

$$\begin{aligned} X^m &= \left\{x \in B^m | x \in X^{m-1}\right\} \\ &= \left\{x \in \cap_{j=1}^{m} B^j | x = x_n \text{ for some } n\right\} \end{aligned}$$

Now, Construct a subsequence $\langle x_{n_k} \rangle$ as follows

1. pick $x_{n_1} \in X^1$ arbitrarily
2. pick $x_{n_2} \in X^2$ in a way so that $n_2 > n_1$. Since n_1 is finite and X^2 is an infinite set this is possible.
3. By induction, n_{k-1} is finite for every k and X^k is infinite so we can always let $n_k > n_{k-1}$.

Clearly $\langle x_{n_k} \rangle$ is a subsequence and if $k, l \geq 2m$ then

$$\rho\left(x_{n_k}, x_{n_l}\right) \leq \frac{1}{m},$$

so $\langle x_{n_k} \rangle$ is Cauchy. Since X is complete the subsequence converges, implying that X is sequentially compact. ■

4 Contraction Mappings

Definition 22 *Let $M = \{X, \rho\}$ be a metric space and let $F : X \to X$. We say that F is a contraction if there exists $\beta \in (0, 1)$ such that*

$$\rho\left(F\left(x\right), F(y)\right) \leq \beta\rho\left(x, y\right)$$

for every $(x, y) \in X$.

Definition 23 *$x \in X$ is called a fixed point of F if $F\left(x\right) = x$.*

Remark 3 *Any contraction map is (uniformly) continuous.*

Proof. Pick $\varepsilon > 0$. Let $\delta = \frac{\varepsilon}{\beta}$. Then if $\rho(x, y) \leq \delta$ we have that

$$\rho(F(x), F(y)) \leq \beta\rho(x, y) \leq \beta\delta = \beta\frac{\varepsilon}{\beta} = \varepsilon.$$

■

Example 13 *Let* $F : [0, 1] \rightarrow [0, 1]$ *and* $\rho(x, y) = |x - y|$. *Then a contraction is a function* F *such that*

$$\frac{|F(x) - F(y)|}{|x - y|} \leq \beta$$

Absolute value of slope less than unity. Hence the function

$$H(x) = F(x) - x$$

is such that x *is a fixed point if and only if* $H(x) = 0$. *Also,* H *satisfies:*

1. $H[0] = F(0) - 0 \geq 0$
2. $H(1) = F(1) - 1 \leq 0$
3. H *is continuous. Hence, an existence of a fixed point is immediate from intermediate value theorem*
4. *Let* $x > y$. *Then* $(x - y) = |x - y| \geq \frac{|F(x) - F(y)|}{\beta} \Rightarrow$

   $$-(x - y) \leq -\frac{|F(x) - F(y)|}{\beta}$$

   so,

   $$\begin{aligned} H(x) - H(y) &= (F(x) - F(y)) - (x - y) \\ &\leq F(x) - F(y) - \frac{|F(x) - F(y)|}{\beta} \\ &\leq |F(x) - F(y)|\left[\frac{\beta - 1}{\beta}\right] < 0 \end{aligned}$$

Hence, if x is a fixed point and $x > y$. Then (by calculation above)

$$H(y) = H(y) - H(x) \geq |F(x) - F(y)| \left[\frac{1-\beta}{\beta}\right] > 0$$

and if $x < y$ (reversing roles)

$$H(y) = H(y) - H(x) \leq |F(y) - F(x)| \left[\frac{\beta-1}{\beta}\right] < 0$$

showing that the fixed point is unique.

General case almost identical. Recursively, define $F^n : X \to X$ as

$$F^n(x) = F\left(F^{n-1}(x)\right)$$

for every x.

Theorem 6 *If $M = \{X, \rho\}$ is a complete metric space and $F : X \to X$ is a contraction mapping with modulus β. Then,*

1. *F has exactly one fixed point $x \in X$.*
2. *Given any $x_0 \in X$ we have that $\rho(F^n(x_0), x) \leq \beta^n \rho(x_0, x)$ (eg we have an algoritm to compute a fixed point)*

Proof. Let x_0 be arbitrary and let

$$\begin{aligned}
x_1 &= F(x_0) \\
x_2 &= F(x_1) = F(F(x_0)) = F^2(x_0) \\
&.... \\
x_n &= F(x_{n-1}) = F\left(F^{n-1}(x_0)\right) = F^n(x_0) \\
&....
\end{aligned}$$

Note that

$$\begin{aligned}
\rho\left(x_n, x_{n-1}\right) &= \rho\left(F^{n-1}\left(x_1\right), F^{n-1}\left(x_0\right)\right) \\
&= \rho\left(F\left(F^{n-2}\left(x_1\right)\right), F\left(F^{n-2}\left(x_0\right)\right)\right) \leq \beta\rho\left(F^{n-2}\left(x_1\right), F^{n-2}\left(x_0\right)\right) \\
&= \rho\left(F\left(F^{n-3}\left(x_1\right)\right), F\left(F^{n-3}\left(x_0\right)\right)\right) \leq \beta^2\rho\left(F^{n-3}\left(x_1\right), F^{n-3}\left(x_0\right)\right) \\
& \quad .... \\
&\leq \beta^{n-1}\rho\left(x_1, x_0\right)
\end{aligned}$$

Hence, if $m > n$ we have that

$$\begin{aligned}
\rho\left(x_m, x_n\right) &\leq \rho\left(x_m, x_{m-1}\right) + \rho\left(x_{m-1}, x_{m-2}\right) + ... + \rho\left(x_{n+1}, x_n\right) \\
&\leq \beta^{m-1}\rho\left(x_1, x_0\right) + \beta^{m-2}\rho\left(x_1, x_0\right) + ... + \beta^n\rho\left(x_1, x_0\right) \\
&= \beta^n\rho\left(x_1, x_0\right)\left[\sum_{j=0}^{m-1-n}\beta^j\right] = \beta^n\rho\left(x_1, x_0\right)\frac{1-\beta^{m-1-n}}{1-\beta} \leq \frac{\beta^n\rho\left(x_1, x_0\right)}{1-\beta}
\end{aligned}$$

we conclude that $\{x_n\}$ is Cauchy. ($\frac{\beta^n\rho(x_1,x_0)}{1-\beta} \longrightarrow 0$ so given $\varepsilon > 0$ there exists K so that $\frac{\beta^n\rho(x_1,x_0)}{1-\beta} < \varepsilon$). Since the metric space is complete there exists a limit point x such that $x_n \longrightarrow x$.

Next, observe that for all n

$$\begin{aligned}
\rho\left(F\left(x\right), x\right) &\leq \rho\left(F\left(x\right), F^n\left(x_0\right)\right) + \rho\left(F^n\left(x_0\right), x\right) \\
&\leq \beta\rho\left(x, F^{n-1}\left(x_0\right)\right) + \rho\left(F^n\left(x_0\right), x\right)
\end{aligned}$$

If $F\left(x\right) \neq x$ it follows that

$$\beta\rho\left(x, F^{n-1}\left(x_0\right)\right) + \rho\left(F^n\left(x_0\right), x\right) \geq \rho\left(F\left(x\right), x\right) = \varepsilon > 0,$$

which contradicts the fact that $x_{n-1} = F^{n-1}\left(x_0\right) \longrightarrow x$ and $x_n = F^n\left(x_0\right) \longrightarrow x$.

For uniqueness, assume that there exists two fixed-points $\left(x, y\right)$. Then

$$d = \rho\left(x, y\right) = \rho\left(F\left(x\right), F\left(y\right)\right) \leq \beta\rho\left(x, y\right) = \beta d,$$

contradicting $\beta < 1$.

The formula follows by induction (see above).

■

Proposition 9 *Let* $A \subset R^n$ *and* $X = \{f : A \to R|\ f$ *is bounded and continuous*$\}$ *and let*

$$\rho(f, g) = \sup_{a \in A} |f(a) - g(a)|.$$

Then $\{X, \rho\}$ *is a complete metric space.*

Proof. Let $\{f_n\}$ be Cauchy. For every $a \in A$ the sequence of reals $\{f_n(a)\}$ is Cauchy because given $\varepsilon > 0$ we have that there exists K such that $\rho(f_n, f_m) = \sup_{a \in A} |f_n(a) - f_m(a)| < \varepsilon$ for $n, m \geq K$ so

$$|f_n(a) - f_m(a)| \leq \sup_{a \in A} |f_n(a) - f_m(a)| < \varepsilon$$

for $m, m \geq K$. By completeness of the reals there exists $f^*(a)$ such that $f_n(a) \to f^*(a)$. Let $f^* : A \to R$ given by $f^*(a)$ for every a be our candidate limit. Fix ε and let $K(\varepsilon)$ be given so that $\rho(f_n, f_m) < \frac{\varepsilon}{2}$ for every $m, n \geq K(\varepsilon)$ and note that **for any** $a \in A$

$$\begin{aligned} |f_n(a) - f(a)| &\leq |f_n(a) - f_m(a)| + |f_m(a) - f(a)| \\ &\leq \underbrace{\sup_{a \in A} |f_n(a) - f_m(a)|}_{\rho(f_n, f_m) < \frac{\varepsilon}{2}} + |f_m(a) - f(a)| \\ &< \frac{\varepsilon}{2} + |f_m(a) - f(a)| \end{aligned}$$

As $f_m(a) \to f(a)$ there is $M(\varepsilon)$ such that $f_m(a) - f(a) < \frac{\varepsilon}{2}$ if $m \geq M(\varepsilon)$. Hence

$$|f_n(a) - f(a)| < \varepsilon$$

for every $n \geq K(\varepsilon)$ because the triangle inequality applies for all $m, n \geq K(\varepsilon)$. Finally, the argument applied for any $a \in A$, so

$$\rho(f_n, f) = \sup_{a \in A} |f_n(a) - f(a)| < \varepsilon$$

for every $n \geq K(\varepsilon)$.

Next, to show completeness we must show that if each f_n is bounded and continuous, then it follows that f is bounded and continuous. Boundedness is immediate (if we can't bound the limit we can't bound the functions approaching the limit). For continuity, fix $\varepsilon > 0$ and pick K so that

$$\rho(f_n, f) = \sup_{a \in A} |f_n(a) - f(a)| < \frac{\varepsilon}{3}$$

if $n \geq K$, which is possible since $f_n \to f$ in the supnorm. Moreover, pick δ so that

$$\left|f_n(a) - f_n(b)\right| < \frac{\varepsilon}{3} \text{ for all } (a,b) \text{ such that } \|a-b\| < \delta,$$

which is possible since f_n is continuous. But then, for a, b such that for $\|a-b\| < \delta$

$$\begin{aligned}
\left|f(a) - f(b)\right| &\leq \left|f(a) - f_n(a)\right| + \left|f_n(a) - f(b)\right| \\
&\leq \left|f(a) - f_n(a)\right| + \left|f_n(a) - f_n(b)\right| + \left|f_n(b) - f(b)\right| \\
&\leq 2\sup_{a \in A}\left|f_n(a) - f(a)\right| + \left|f_n(a) - f_n(b)\right| < \varepsilon.
\end{aligned}$$

■

Theorem 7 (Blackwell's Sufficient Conditions) *Let $X \subset R^n$ and $B = \{f : X \to R|\ f$ is bounded and*

and let

$$\rho(f,g) = \sup_{a \in A}\left|f(a) - g(a)\right|.$$

Suppose that $F : B \to B$ satisfies:

1. *[monotonicity] $f(x) \leq g(x)$ for every $x \in X \Rightarrow F(f(x)) \leq F(g(x))$ for every $x \in X$.*
2. *[discounting] there exists $\beta \in (0,1)$ such that*

   $$F(f(x) + a) \leq F(f(x)) + \beta a$$

   for all $f \in B, a \geq 0$ and $x \in X$.

Then F is a contraction mapping.

Proof. For any $f, g \in B$ and $x \in X$ we have that

$$\begin{aligned}
f(x) - g(x) &\leq \sup_{y \in X}\left|f(x) - g(x)\right| \Rightarrow \\
f(x) &\leq g(x) + \sup_{y \in X}\left|f(x) - g(x)\right|
\end{aligned}$$

Hence,

$$F(f(x)) \underset{\text{mono}}{\leq} F\left(g(x) + \sup_{y \in X}\left|f(y) - g(y)\right|\right) \underset{\text{disc}}{\leq} F(g(x)) + \beta\sup_{y \in X}\left|f(y) - g(y)\right|$$

Symmetrically

$$F\left(g\left(x\right)\right)\underset{\text{mono}}{\leq}F\left(f\left(x\right)+\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right|\right)\underset{\text{disc}}{\leq}F\left(f\left(x\right)\right)+\beta\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right|$$

Hence, for all $x\in X$

$$\begin{aligned}-\beta\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right| &\leq F\left(f\left(x\right)\right)-F\left(g\left(x\right)\right)\leq\beta\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right|\\ \left|F\left(f\left(x\right)\right)-F\left(g\left(x\right)\right)\right| &\leq \beta\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right|\\ \sup_{x\in X}\left|F\left(f\left(x\right)\right)-F\left(g\left(x\right)\right)\right| &\leq \beta\sup_{y\in X}\left|f\left(y\right)-g\left(y\right)\right|,\end{aligned}$$

implying that F is a contraction.

■

Example 14 *Suppose that $u\left(c\right)$ and $f\left(k\right)$ are bounded continuous functions. The most standard optimal growth problem is then*

$$\begin{aligned}&\max_{\{k_{t+1}\}_{t=0}^{\infty}}\sum_{t=0}^{\infty}\beta^{t}u\left(f\left(k_{t}\right)-k_{t}\right)\\ s.t.\ 0 &\leq k_{t+1}\leq f\left(k_{t}\right)\end{aligned}$$

The associated functional equation is

$$\left(FV\right)\left(k\right)=\max_{0\leq y\leq f(k)}u\left(f\left(k\right)-y\right)+\beta V\left(y\right)$$

Suppose that $V\left(y\right)\leq W\left(y\right)$ for all y. Let

$$y^{*}\left(k\right)\in\arg\max_{0\leq y\leq f(k)}u\left(f\left(k\right)-y\right)+\beta V\left(y\right).$$

Then

$$\begin{aligned}\left(FW\right)\left(k\right) &= \max_{0\leq y\leq f(k)}u\left(f\left(k\right)-y\right)+\beta W\left(y\right)\\ &\geq u\left(f\left(k\right)-y^{*}\left(k\right)\right)+\beta W\left(y^{*}\left(k\right)\right)\\ &= u\left(f\left(k\right)-y^{*}\left(k\right)\right)+\beta V\left(y^{*}\left(k\right)\right)+\beta\left[W\left(y^{*}\left(k\right)\right)-V\left(y^{*}\left(k\right)\right)\right]\\ &\geq u\left(f\left(k\right)-y^{*}\left(k\right)\right)+\beta V\left(y^{*}\left(k\right)\right)=\max_{0\leq y\leq f(k)}u\left(f\left(k\right)-y\right)+\beta V\left(y\right),\end{aligned}$$

proving monotonicity. Furthermore

$$\begin{aligned} F\left(V\left(k\right)+a\right) &= \max_{0\leq y\leq f(k)} u\left(f\left(k\right)-y\right)+\beta\left[V\left(y\right)+a\right] \\ &= \max_{0\leq y\leq f(k)} u\left(f\left(k\right)-y\right)+\beta V\left(y\right)+\beta a \\ &= F\left(V\left(k\right)\right)+\beta a \end{aligned}$$

showing that the discounting hypothesis is satisfied. Hence, Blackwells conditions hold, implying that F is a contraction. By the contraction mapping result it follows that there exists a function $V\left(\cdot\right)$ such that

$$V\left(k\right)=\max_{0\leq y\leq f(k)} u\left(f\left(k\right)-y\right)+\beta V\left(y\right).$$

One then needs to show that the dynamic program and the sequential formulation are equivalent, which involves some work, but is highly intuitive (see chapter 4 in Stokey Lucas).

5 Theorem of the Maximum

Definition 24 *Given sets X and Y we say that G is a correspondence between X and Y (denoted $G : X :\rightrightarrows Y$) if $G\left(x\right)\subset Y$ for every $x\in X$.*

Example 15 *Let $g : X\rightarrow R$ be a function and let $G : X :\rightrightarrows R$ be defined as*

$$G\left(x\right)=\left\{x\in R|g\left(x\right)\geq 0\right\}.$$

then G is a correspondence.

Definition 25 (uhc 1) *$G : X :\rightrightarrows Y$ is upper hemi continuous at $x\in X$ if given any open set $B\subset Y$ such that $G\left(X\right)\subset B$ there exists an open set A with $x\in A$ such that $G\left(A\right)\subset B$ (or, equivalently, $G\left(x'\right)\subset B$ for every $x'\in A$.*

Definition 26 (lhc 1) *$G : X :\rightrightarrows Y$ is lower hemi continuous at $x\in X$ if given any $y\in G\left(X\right)$ and open set $B\subset Y$ such that $y\in B$ there exists an open set A with $x\in A$ such that $G\left(x'\right)\cap B\neq\varnothing$ (or, equivalently, so that for every $y\in G\left(X\right)$ and open B containing y*

there is an open A containing x such that there exists some $y' \in B$ with $y' \in G(x')$ for every $x' \in A$.

Definition 27 *$G : X :\rightrightarrows Y$ is continuous at x if it is upper and lower hemi continuous*

Lemma 3 *Suppose that $X \subset R^n$ and $Y \subset R^l$ and that $G : X :\rightrightarrows Y$ is upper hemi continuous at $x \in X$. Furthermore, suppose that $G(x)$ is compact and non-empty for every $x \in X$. Then, given any sequence $\langle x_n \rangle$ in X converging to x and every sequence $\langle y_n \rangle$ with $y_n \in G(x_n)$ there exists a converging subsequence $\langle y_{n_k} \rangle$ with limit $y \in G(x)$.*

Proof. Since $G(x_n)$ is compact for every n the sequence $\langle y_n \rangle$ must be bounded (otherwise there exists x_n and $y_n \in G(x_n)$ such that $\|y_n - y\| \geq M$ for every $y \in Y$ and M, implying that we can create a sequence $\langle y_n \rangle$ from $\langle G(x_n) \rangle$ with $y_n \to \infty$ or $y_n \to -\infty$, which would imply that $G(x)$ is not bounded, violating the assumption that $G(x)$ is compact). Hence, there is a convergent subsequence with limit y^*. If $y^* \notin G(x)$ we have that $y^* \in R^l \backslash G(x)$ which is open by the Heine Borel theorem. That implies that there exists $\varepsilon > 0$ such that

$$B(y, 2\varepsilon) \subset R^l \backslash G(x)$$

and, as $\langle y_{n_k} \rangle \to y^*$ there exists K_1 such that

$$\|y_{n_k} - y\| \geq \varepsilon$$

for every $n_k \geq K_1$ and $y \in G(x)$, which we see because

$$\begin{array}{rcl} \|y - y^*\| & \leq & \|y - y_{n_k}\| + \|y_{n_k} - y^*\| \\ & \Leftrightarrow & \\ \|y - y_{n_k}\| & \geq & \|y - y^*\| - \|y_{n_k} - y^*\| \\ & \geq & 2\varepsilon - \varepsilon \end{array}$$

if n_k chosen so that $\|y_{n_k} - y^*\| \leq \varepsilon$. But, let

$$B = \left\{ y \in Y | \text{there exists } y' \in G(x) \text{ such that } y \in B\left(y', \frac{\varepsilon}{2}\right) \right\}.$$

as the mapping is upper hemi-continous there exists open set A containing x s.t.

$$G(A) \subset B,$$

and since $x_n \to x$ there is some K_2 such that $x_n \in A$ for every $n \geq K_2$. In other words

$$\|y_n - y\| \leq \frac{\varepsilon}{2}$$

for every $n \geq K_2$. This is a contradiction since

$$\varepsilon \leq \|y_{n_k} - y\| \leq \frac{\varepsilon}{2}$$

for every n_k occurring in the subsequence with $n_k \geq \max\{K_1, K_2\}$.

■

Lemma 4 *Suppose that $X \subset R^n$ and $Y \subset R^l$ and that $G : X :\rightrightarrows Y$ is lower hemi continuous at $x \in X$. Furthermore, suppose that $G(x)$ is compact and non-empty for every $x \in X$. Then, given any sequence $\langle x_n \rangle$ in X converging to x and every $y \in G(x)$ there exists a sequence $\langle y_n \rangle$ in Y converging to y and some finite N such that $y_n \in G(x_n)$ for all $n \geq N$.*

Proof. Similar. ■

Now, consider

$$\begin{aligned} V(x) &= \max_{y \in G(x)} f(x,y) \\ P(x) &= \arg\max_{y e G(x)} f(x,y) \end{aligned}$$

In economic applications, we often have problems on form

$$\begin{aligned} &\max_{y \in R^n} f(x,y) \\ g_1(x) &\geq 0 \\ &... \\ g_l(x) &\geq 0, \end{aligned}$$

which clearly is covered by this as we can define

$$G(x) = \{x \in X | g_j(x) \geq 0 \text{ for all } j\}.$$

Theorem 8 *Suppose that* $X \subset R^n$ *and* $Y \subset R^l$ *and that* $G : X :\rightrightarrows Y$ *is a compact non-empty valued continuous correspondence. Let* $f : X \times Y \to R$ *be continuous. Then, there exists a uniquely defined function* $V : X \to R$ *and a uniquely defined correspondence* $P : X :\rightrightarrows Y$ *such that:*

1. V *is continuous.*
2. P *is upper hemi continuous and compact valued.*

Proof. $G(x)$ is compact for every x and $f(x, y)$ is continuous and therefore continuous in x. Hence, $P(x) \subset G(x)$ is non empty and $V(x)$ is well defined.

COMPACTNESS OF $G(x)$: Moreover

$$P(x) \subset G(x) \subset R^l$$

and $G(x)$ is compact$\Leftrightarrow$ $G(x)$ is closed and bounded$\Rightarrow$ $P(x)$ is bounded.

We want to show that $P(x)$ is closed. Suppose not, then there exists some $y \in \overline{P(x)}$ (the closure) which is not an element of $P(x)$. That implies that:

- there exists $y \notin P(x)$ such that for every $\varepsilon > 0$ there exists $y_n \in B\left(y, \frac{1}{n}\right)$ such that $y_n \in P(x)$
- but, $y \notin P(x)$ means that

$$f(y, x) < V(x) = f(y_n, x),$$

  which, as $y_n \to y$ contradicts continuity of $f(., x)$. Hence, $P(x)$ is closed and bounded$\Leftrightarrow$compact.

P **UHC:** Pick a sequence $\langle x_n \rangle$ converging to x and let $y_n \in P(x_n)$ for every n. Want to show that there exists a subsequence $\langle y_{n_k} \rangle$ of $\langle y_n \rangle$ such that $y_{n_k} \to y \in P(x)$. Now,

- G is upper hemi continuous. hence, there exists subsequence $\langle y_{n_k} \rangle$ of $\langle y_n \rangle$ such that $y_{n_k} \to y \in G(x)$.

- G is lower hemi continuous, which means that given any sequence $\langle x_n \rangle$ in X converging to x and every $y \in G(x)$ there exists a sequence $\langle y_n \rangle$ in Y converging to y and some finite N such that $y_n \in G(x_n)$ for all $n \geq N$.

- Applying this to the sequence $\langle x_{n_k} \rangle$ with every n_k corresponding to the subsequence $\langle y_{n_k} \rangle$ we have that $\langle x_{n_k} \rangle$ converges to x (every subsequence of a convergent sequence must converge to the limit of the sequence) so (LHC) for every $z \in G(x)$ there exists sequence $\langle z_{n_k} \rangle$ such that $z_{n_k} \in G(x_{n_k})$ for n_k large enough and

$$z_{n_k} \to z \in G(x)$$

- Moreover

$$f(y_{n_k}, x_{n_k}) \geq f(z_{n_k}, x_{n_k})$$

for every n_k and f is continuous. Hence

$$f(y, x) \geq f(z, x).$$

True for every $z \in G(x)$. Hence, $y \in P(x)$, implying that P is upper hemi continuous.

V **CONTINUOUS:** For contradiction, suppose that there exits x and sequence $\langle x_n \rangle$ with $x_n \to x$ such that

$$V(x_n) \nrightarrow V(x)$$

Let $\langle y_n \rangle$ be a sequence with $y_n \in P(x_n)$ for every n. Upper hemi continuity of P means that there exists a converging subsequence $\langle y_{n_k} \rangle$ with limit $y \in P(x)$. Hence, $\langle x_{n_k}, y_{n_k} \rangle \to (x, y)$ where

$$\begin{aligned} V(x, y) &= f(x, y) \\ V(x_{n_k}, y_{n_k}) &= f(x_{n_k}, y_{n_k}) \text{ for every } n_k \end{aligned}$$

and since f is continuous it follows that

$$V(x_{n_k}) = f(x_{n_k}, y_{n_k}) \to f(x, y) = V(x)$$

But

$$\begin{aligned} V(x_{n_k}) &\rightarrow V(x) \\ f(x_{n_k}, y) &\rightarrow f(x,y) = V\left(x\right) \end{aligned}$$

contradicts the discontinuity as $x_n \rightarrow x \Leftrightarrow x_{n_k} \rightarrow x$ for every subsequence

■