Metric Spaces

A. Banerji

January 4, 2017

1 Introduction

Definition 1 *A metric space is a set S and a metric $\rho : S \times S \to \Re$ satisfying*

1. $\rho(x, y) = 0 \; iff \; x = y$
2. $\rho(x, y) = \rho(y, x)$
3. *For all* $x, y, z \in S$, $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$

Note. $\rho(x, y) \geq 0$. Indeed, consider the 3 points x, x, y. Applying the triangle inequality,
$0 = \rho(x, x) \leq \rho(x, y) + \rho(y, x) = 2\rho(x, y)$ ■
Example: d_2 in $\Re^k$, defined by
$d_2(x, y) = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$

Definition 2 *A mapping $\| \; \| : \Re^k \to \Re$ is called a norm on the vector space $\Re^k$ if*

1. $\|x\| = 0 \; iff \; x = 0$
2. *For all* $\gamma \in \Re$, $\|\gamma x\| = |\gamma| \|x\|$
3. *For all* $x, y \in \Re^k$, $\|x + y\| \leq \|x\| + \|y\|$

Note. A norm induces the metric $\|x - y\|$. Verify that this satisfies properties 1 and 2 of a metric. For the triangle inequality, note that
$\|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\|$
(by Property 3 of the norm)

Note. $\|z\| \geq 0$ for every z. Indeed, 3 implies that for $z, -z \in \Re^k$, $0 = \|z + (-z)\| \leq \|z\| + \|(-1)z\|$. By 2, this equals $2\|z\|$.

Examples. $\|x\|_p = (\sum_{i=1}^{k} |x_i|^p)^{1/p}$ where $p \in [1, \infty)$. $p = 2$ is the Euclidean norm which induces the Euclidean metric d_2. Minkowski for triangle inequality. $p \to \infty \Rightarrow$ Max norm.

Functions and Sup Norm

Consider the vector space bU of all bounded functions from a set U to $\Re$ (over the field of real numbers). For any function f, define its norm $\|f\| = sup\{|f(x)| : x \in U\}$

- Claim:

$$d_\infty(f,g) = \sup_{x\in U} |f(x) - g(x)|$$

  is an induced metric. (bU, d_∞) is a metric space.

- For the triangle inequality, let $f, g, h \in bU$
- Take any $x \in U$. $|f(x) - g(x)| =$

  $|f(x) - h(x) + h(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)|$

  $\leq d_\infty(f,h) + d_\infty(h,g)$

- Since x is arbitrary, $d_\infty(f,h) + d_\infty(h,g)$ is an upper bound of the set $\{|f(x) - g(x)| : x \in U\}$. So its supremum

  $d_\infty(f,g) \leq d_\infty(f,h) + d_\infty(h,g)$

2 Sequences

Definition 3 *A sequence $(x_n) \subset S$ converges to $x \in S$ if for all $\epsilon > 0$, there exists a positive integer N such that $n \geq N$ implies $\rho(x_n, x) < \epsilon$.*

Note:So the sequence $(\rho(x_n, x)) \subset \Re$ converges to 0, in the metric space $(\Re, |\quad|)$.

Theorem 1 *A sequence in (S, ρ) can have at most one limit.*

Proof.

Suppose $x, x^{'}$ are limits of (x_n). By the triangle inequality, we have for all n, $0 \leq \rho(x, x^{'}) \leq \rho(x, x_n) + \rho(x_n, x^{'})$

Since the RHS tends to zero, $\rho(x, x^{'}) < \epsilon$ for every positive ϵ. So $\rho(x, x^{'}) = 0$, so $x = x^{'}$. ∎

Open ball $B(x, \epsilon) = \{y \in S | \rho(x, y) < \epsilon\}$

What does $B(f, \epsilon)$ look like if $f \in (b[a,b], d_\infty)$?

Definition 4 *A subset $E \subset S$ is bounded if $E \subset B(x, n)$ for some $x \in S$ and $n \in \mathbb{N}$. A sequence (x_n) is bounded if its range $\{y | y = x_n \ for \ some \ n \in \mathbb{N}\}$ is a bounded set.*

Example 1 *JS Ex3.1.3. Every convergent sequence in S is bounded.*

Proof 1 *Let $x_n \to x$. Choose $\epsilon > 0$. Let N be such that $n \geq N$ implies $\rho(x, x_n) < \epsilon$. Let $r = max\{\rho(x, x_1), \ldots, \rho(x, x_{N-1}), \epsilon\}$. Then $(x_n) \subset B(x, k)$, for $k > r$.* ■

Subsequences.

Definition 5 *If (x_n) is a sequence and $n_1 < n_2 < \ldots < n_k < \ldots$ is an increasing sequence of positive integers, then (x_{n_k}) is called a subsequence of (x_n).*

Definition 6 *x is a limit point of the sequence (x_n) if it is the limit of some subsequence of (x_n).*

Note: A limit point x has infinitely many members of (x_n) ϵ-close to it (for every $\epsilon > 0$).

Example 2 *Non-convergent sequences may nevertheless have limit points.*

$(x_n) = (-1)^n$ is a sequence with limit points $1, -1$ but no limit.

$$(x_n) = \begin{cases} 2 + (-1/n)^n & n = 1, 2, 5, 6, ... \\ -2 + (-1/n)^n & n = 3, 4, 7, 8, \end{cases}$$

$2, -2$ are the limit points of this sequence, which again has no limit.

JS Ex3.1.4. (x_n) converges to x if and only if every subsequence of (x_n) does so.

Indeed, if every subsequence converges to x, so does (x_n) since it is itself a subsequence. Conversely, suppose $x_n \to x$ but that some subsequence (x_{n_k}) does not. So there exists $\epsilon > 0$ such that for every $N \in \mathbb{N}$, there exists $k \geq N$ such that $\rho(x, x_{n_k}) \geq \epsilon$. Thus for $n = n_k$, $n \geq N$ and $\rho(x, x_n) \geq \epsilon$. So x_n does not converge to x. Contradiction.

Recall the **Completeness Property** of the set of Real numbers: Every set of reals bounded above has a supremum, every set of reals bounded below has an infimum.

This follows from the way that the real numbers are constructed from the rationals: e.g., by means of Dedekind cuts from the set of rational numbers, or equivalence classes of Cauchy sequences of rational numbers. Note that the completeness property does not hold for, say, the rational numbers. For instance, the successive decimal expansions approximating $\sqrt{2}$, $\{1.4, 1.41, 1.414, \ldots\}$ is bounded above (say by the rational number 1.42), but it does not have a rational number supremum. Hence the rationals are *completed* by extending them to the larger set of real numbers.

Recall the property of the supremum of a set A: for every $\epsilon > 0$, there's an element $a \in A$ s.t. $a \in (\sup A - \epsilon, \sup A]$. If this were not true for some ϵ, then $\sup A - \epsilon$ would be an upper bound of A, contradicting that $\sup A$ is the least upper bound of A. We then have

Theorem 2 *Every bounded and increasing sequence of real numbers converges (to its supremum).*

Proof 2 *Let $a \equiv \sup(x_n)$. Take any $\epsilon > 0$. By the above discussion, there exists some $x_N \in (a - \epsilon, a]$. And since (x_n) is an increasing sequence, we have that for all $k \geq N$, $x_k \in (a - \epsilon, a]$. So $(x_n) \to a$.* ■

A similar conclusion holds for decreasing bounded sequences. And:

Theorem 3 *Every sequence of real numbers has a monotone subsequence.*

Proof 3 *For the bounded sequence* (x_k), *let* $A_n = \{x_k | k \geq n\}, n = 1, 2, \ldots.$ *If any one of these sets* A_n *does not have a maximum, we can pull out an increasing sequence. For instance, suppose* A_1 *does not have a max. Then let* $x_{k_1} = x_1$. *Let* x_{k_2} *be the first member of the sequence* (x_k) *that is greater than* x_1, *and so on.*

On the other hand, if all A_n *have maxes, then we can pull out a decreasing subsequence. Let* $x_{k_1} = \max\{A_1\}$, $x_{k_2} = \max\{A_{k_1} + 1\}, x_{k_3} = \max\{A_{k_2} + 1\}$ *and so on.* ■

It follows from the above two theorems, that we have

Theorem 4 Bolzano-Weierstrass Theorem.

Every bounded sequence of real numbers has a convergent subsequence.

Finally, as an application to the ideas of monotone sequences, we have

Theorem 5 Cantor's Nested Intervals theorem.

If $[a_1, b_1] \supseteq [a_2, b_2] \supseteq \ldots$ *is a nested sequence of closed intervals, then* $\cap_{m=1}^{\infty}[a_m, b_m]$ *is nonempty. Moreover, if* $b_m - a_m \to 0$, *then this intersection is a single point.*

Proof 4 *Because of the nesting,* $a_1 \leq a_2 \leq \ldots \leq b_2 \leq b_1$. *So,* (a_k) *is bounded and increasing and so has a supremum, say* a; (b_k) *is bounded and decreasing and has an infimum, say* b; *and* $a \leq b$. *So,* $[a, b] \subseteq [a_m, b_m], m = 1, 2, \ldots,$ *and therefore lies in the intersection* $\cap_{m=1}^{\infty}[a_m, b_m]$; *which is therefore nonempty. Moreover, if* $b_m - a_m \to 0$, *then by sandwiching,* $a = b$ *and the intersection is a single point.* ■

Note. We've seen examples of non-convergent sequences of *real numbers* with multiple limit points. If there is a largest and smallest limit point, these are called the *limit superior* and *limit inferior* of (x_n). Let's discuss these a little, starting with the alternative definition given in Stachurski, Appendix A.

Definition 7 *Let (x_n) be a bounded sequence of real numbers.*

$\limsup x_n = \lim_{n\to\infty}(\sup_{k\geq n} x_k)$

$\liminf x_n = \lim_{n\to\infty}(inf_{k\geq n} x_k)$

The set $A_n = \{x_k | k \geq n\}$ is bounded and so has a supremum and an infimum. Call these y_n and z_n respectively.

Notice that for every n,

$A_{n+1} = \{x_k | k \geq n+1\} \subset \{x_k | k \geq n\} = A_n$

(as we are removing an element from the first set), so $y_{n+1} \leq y_n$. So (y_n) is a decreasing bounded sequence and therefore converges to its infimum.

Thus

$\limsup x_n = \inf(y_n) = \inf_n(\sup_{k\geq n} x_k)$

Similarly,

$\liminf x_n = \sup(z_n) = \sup_n(\inf_{k\geq n} x_k)$

Note. For each n, y_n and z_n are the sup and inf of the same set. So $y_n \geq z_n$. Taking inf on the y_ns and sup on the z_ns preserves this inequality, so

$\limsup x_n \geq \liminf x_n$.

Theorem 6 $\limsup x_n$ *is the largest limit point of* (x_n).

Proof 5 *Let $y = \limsup x_n$. Since $y_n \downarrow y$, $y_n \geq y$, for all n. Now, $y_n = \sup\{x_k | k \geq n\}$, so for each n, pick an element (calling it w_n) in $\{x_k | k \geq n\}$ 'very close to' y_n. Note $y_n \geq w_n \geq y$, $\forall n$. So, since $y_n \to y$, we have $w_n \to y$; also all w_n are members of the original sequence x_n. So y is indeed a limit point of (x_n). y is also the largest limit point. For suppose $y' > y$ is another limit point. Let $\epsilon = (y' - y)/2$. So there exists N such that $n \geq N$ implies $y_n < y' - \epsilon$. Since $y_n = \sup\{x_k | k \geq n\}$, at most a finite number of* *points (those from $(x_1, ..., x_{N-1})$) can be ϵ-close to y'. Thus y' is not a limit point of (x_n).* ■

Similarly, $\liminf x_n$ is the smallest limit point of (x_n).

Sequences and Function Continuity

Theorem 7 *A sequence (x_n) in $(\Re^k, d_p)$ converges to a point $x = (x^1, ..., x^k) \in \Re^k$ if and only if each coordinate sequence (x_n^j) converges to x^j.*

Remark: Thus if (x_n) converges for *some* metric d_p, $p \in [1, \infty)$, all coordinate sequences converge in $\Re$, which therefore implies that (x_n) converges for *every* metric d_p.

Definition 8 *Let S, Y be two metric spaces, and $A \subset S$. A function $f : A \to Y$ is continuous at $a \in A$ if for every sequence (x_n) converging to a, $(f(x_n))$ converges to $f(a)$. f is continuous on A if this is true for all $a \in A$. (Diagram)*

Lemma 1 *Let $\bar{x} \in S$. Define $f : S \to \Re$ by $f(x) = \rho(x, \bar{x})$, for all $x \in S$. Then f is continuous on S.*

Proof 6 *$\rho(x_n, \bar{x}) \leq \rho(x_n, x) + \rho(x, \bar{x})$ by the triangle inequality (draw a diagram also). So, $\rho(x_n, \bar{x}) - \rho(x, \bar{x}) \leq \rho(x_n, x)$. Similarly, $\rho(x, \bar{x}) - \rho(x_n, \bar{x}) \leq \rho(x_n, x)$. And so, $|\rho(x_n, \bar{x}) - \rho(x, \bar{x})| \leq \rho(x_n, x)$. Since the RHS converges to 0, the result follows.* ■

Second Proof. ($\epsilon - \delta$ continuity defn). Fix any $\epsilon > 0$. We want to find $\delta > 0$ s.t. whenever $\rho(x, y) < \delta$, $|\rho(x, \bar{x}) - \rho(y, \bar{x})| < \epsilon$. Set $\delta = \epsilon$ and notice that

$\rho(x, \bar{x}) - \rho(y, \bar{x}) \leq \rho(x, y) < \epsilon$

The first inequality is the triangle inequality and the 2nd is by choice of y. Similarly, we can show that

$\rho(y, \bar{x}) - \rho(x, \bar{x}) \leq \rho(y, x) < \epsilon$
so the result follows.

Exercises

JS Exercises 3.1.5-6 can be done easily assuming we know corresponding results for limits of sequences. For example, if $a_n \to a$ and $b_n \to b$ then $a_n b_n \to ab$. Using this we can immediately show that if f and g are continuous at x, then so is fg. Because by continuity, $x_n \to x$ implies $f(x_n) \to f(x)$ and $g(x_n) \to g(x)$, so by the product of limits rule, $f(x_n)g(x_n) \to f(x)g(x)$.

Exercise 3.1.7. $f : S \to \Re$ is upper-semicontinuous (usc) at x if for every $x_n \to x$, $\limsup f(x_n) \leq f(x)$; and lsc if $\liminf f(x_n) \geq f(x)$. Suppose f is usc at x. Thinking of largest and smallest limit points, note that $\liminf -f(x_n) = -\limsup f(x_n) \geq -f(x)$, so $-f$ is lsc at x. Suppose f is both usc and lsc at x: so $\liminf f(x_n) \geq f(x) \geq \limsup f(x_n)$. The reverse inequality always holds for $\limsup$ and $\liminf$. So $\liminf f(x_n) = \limsup f(x_n) = \lim f(x_n)$ and this equals $f(x)$. The converse follows because if a sequence $(f(x_n))$ is convergent, then so are all its subsequences.

3 Open, Closed, Compact Sets

Open and Closed Sets

Definition 9 *x is an adherent (or contact or closure) point of $E \subset S$ if for every $\epsilon > 0$, $B(x, \epsilon)$ contains at least one point of E.*

JS Ex3.1.8. x adheres to E iff it is the limit of a sequence (x_n) contained in E.

Proof 7 *If $(x_n) \subset E$ converges to x, then for every $B(x, \epsilon)$, there is some $x_n \in E$ that belongs to $B(x, \epsilon)$; so x adheres to E. Conversely, if x adheres to E, then every $B(x, 1/n)$, $n = 1, 2, ...$, contains a point, say x_n, that belongs to E. By construction, the sequence (x_n) converges to x.* ■

Definition 10 *x is an interior point of E if $B(x,\epsilon) \subset E$, for some $\epsilon > 0$.*

x is an accumulation (or limit or cluster) point of E if for every $\epsilon > 0$, $B(x,\epsilon)$ contains a point of E distinct from x.

Every open ball around a limit point contains an infinity of points of E.

Definition 11 *A set $E \subset S$ is open if all its points are interior points. A set E is closed if it contains all its adherent points.*

Theorem 8 *A set $F \subset S$ is closed iff for every sequence $(x_n) \in F$ that converges in S, $lim_{n\to\infty} x_n \in F$.*

Proof 8 *Suppose F is closed and $(x_n) \subset F$ is a sequence converging to x. So by Ex.3.1.8, x adheres to F. Since F is closed, $x \in F$. Conversely, suppose all convergent $(x_n) \subset F$ have limits in F. To show F is closed, take an arbitrary adherent point of F, say x. By Ex. 3.1.8, there is a sequence $(x_n) \subset F$ that converges to x. So by assumption, $x \in F$, and hence F is closed by definition.* ■

Theorem 9 *E is open iff E^c is closed.*

Proof 9 *Suppose E is open, x adheres to E^c and $x \notin E^c$. So $x \in E$, and so $x \in int(E)$. So for some $\epsilon > 0$, $B(x,\epsilon) \subset E$. So x does not adhere to E^c. Contradiction. So $x \in E^c$ and E^c is closed. Conversely, suppose E^c is closed and $x \in E$. So x does not adhere to E^c; $B(x,\epsilon)$ does not intersect with E^c, for some $\epsilon > 0$. So $B(x,\epsilon) \subset E$, so $x \in int(E)$.* ■

Theorem 10 *Arbitrary unions and finite intersections of open sets are open. Arbitrary intersections and finite unions of closed sets are closed.*

The 2nd statement follows from the first using the preceding result and De-Morgan's laws. Let $G_n = (-1/n, 1/n), n = 1, 2, ...$ be a collection of open sets. $\cap_n G_n = \{0\}$, which is closed.

Definition 12 *The closure of E, $cl(E)$, is the set of all points that adhere to E.*

JS Ex 3.1.18. $cl(E)$ is a closed set. Also, if $E \subset F$, and F is closed, then $cl(E) \subset F$.

Suppose x adheres to $cl(E)$. Want to show that then $x \in cl(E)$, i.e. that x adheres to E. Since x adheres to $cl(E)$, do this: take $\epsilon > 0$, and have a $y \in cl(E)$ s.t. $y \in B(x, \epsilon)$. Since $y \in cl(E)$, it adheres to E, so take $\delta = \epsilon - \rho(x, y)$, and a $z \in E$ s.t. $z \in B(y, \delta) \subseteq B(x, \epsilon)$. So, x adheres to E.

Now suppose $E \subset F$ and F is closed. Suppose x adheres to E. So some $(x_n) \subset E \subset F$ converges to x. So x adheres to F, and since F is closed, $x \in F$.

Similarly, $int(E)$ is the largest open set contained in E.

Open Sets and Continuity

Theorem 11 *A function $f : S \to Y$ is continuous iff the preimage $f^{-1}(G)$ of every open set $G \subset Y$ is open in S.*

Example of discontinuous function violating this definition.

Proof 10 *Suppose f is continuous, $G \subset Y$ is open, and $x \in f^{-1}(G)$ (so $f(x) \in G$). Then $x \in int(f^{-1}(G))$. For suppose not. Then we can construct a sequence $(x_n) \notin f^{-1}(G)$ converging to x. So, the sequence $(f(x_n)) \notin G$, and $f(x_n) \to f(x)$. So $f(x) \notin G$. Contradiction.*

Conversely, suppose that preimages of all open sets of Y are open. Suppose that in S, $(x_n) \to x$. To show $f(x_n) \to f(x)$, let B be any ϵ-open ball around $f(x)$. Since $f^{-1}(B)$ is open and contains x, for large enough N we have $n \geq N$ implies $x_n \in f^{-1}(B)$. So $f(x_n) \in B$. ■

Theorem 12 *Let $f : \Re^k \to \Re^m$. f is d_p-continuous iff it is d_q-continuous, for $p, q \in [1, \infty)$.*

Proof 11 *(Sketch). d_p-closed and d_q-closed sets are identical, and therefore so are d_p and d_q-open sets. Indeed, suppose A is d_p-closed. Suppose $(x_n) \subset A$ d_q-converges to x. Then (x_n^j) converges to x^j, for all coordinate sequences $j = 1, ..., k$. So (x_n) d_p-converges to x. Since A is d_p-closed, $x \in A$. So A is d_q-closed. And so on.* ■

Note. JS Exercise 3.1.24 shows that for metrics other than d_p-metrics, however, the open and closed sets, and therefore the sets of continuous functions and convergent sequences, can be quite different.

Compactness

An open cover of A is a collection $\{G_\alpha\}$ of open sets whose union contains A. A subcover is a subset of sets from this family whose union still contains A. We say a subcover is finite if it has a finite number of sets.

Definition 13 *A set $A \subset S$ is compact if every open cover of A has a finite subcover.*

Examples:

(1). $(0, 1)$ is not compact in the metric space $(\Re, |\ |)$. For consider the family of open sets $\mathcal{O} = \{(\frac{1}{n}, 1) | n = 1, 2, \ldots\}$. Evidently, every $x \in (0, 1)$ belongs to one of these sets, so $\mathcal{O}$ is an open cover of $(0, 1)$. However, there is no finite subcover, because such a subcover would have a largest interval $(\frac{1}{k}, 1)$, and we can find $0 < x < \frac{1}{k}$ for some x.

(2). Any finite subset $S = \{a_1, \ldots, a_K\}$ of a metric space is compact. Suppose $\mathcal{O}$ is an open cover of S. So for each $a_i \in S$, there is a set A_i belonging the family $\mathcal{O}$ that contains it. So, (up to non-uniqueness), $S \subseteq \cup_{i=1}^{K} A_i$; we have a finite subcover.

(3). $[0, 1]$ is compact in $\Re$.

For suppose not. Then there exists an open cover $\mathcal{O}$ such that either $[0, \frac{1}{2}]$ or $[\frac{1}{2}, 1]$ has no finite subcover. Call the relevant interval $[a_1, b_1]$. Then either $[a_1, a_1 + \frac{b_1 - a_1}{2}]$ or $[a_1 + \frac{b_1 - a_1}{2}, b_1]$ does not have a finite subcover. Call the relevant set $[a_2, b_2]$ and note that it's nested within $[a_1, b_1]$. Proceeding recursively, we get a nested sequence of closed intervals $[a_m, b_m]$, none of which has a finite subcover, and with $b_m - a_m \to 0$. By Cantor's nested intervals theorem, it has a single point intersection at some y. Since $y \in [0, 1]$, there is some $A \in \mathcal{O}$ that contains it. Since A is open, there exists $(y-\epsilon, y+\epsilon) \subseteq A$; and for m large enough, $[a_m, b_m] \subseteq (y - \epsilon, y + \epsilon)$. But then, such a $[a_m, b_m]$ is covered by A, and hence has a finite subcover. Contradiction.

Compactness has some of the structure of finite sets (Efe Ok); for example, a finite set $E = \{x_1, \ldots, x_k\}$is bounded. Let $n > \max\{\rho(x_1, x_i)|i = 2, \ldots, k\}$. Then $E \subseteq B(x_1, n)$. For an infinite set A, we can't necessarily find a maximum distance as in a finite set. however, if A is compact, we can cover the entire set with a *finite* number of open balls centered at some $x \in A$; as shown below.

Theorem 13 *Every compact subset of a metric space is closed and bounded.*

Proof 12 *Boundedness: Let A be compact. Choose some $x \in A$ and consider the following open cover of A: $\{B(x, n)|n \in \mathbb{N}\}$. If $\{B(x, r_1), ..., B(x, r_k)\}$ is a finite subcover, then $A \subset B(x, r)$, where $r = max\{r_1, ..., r_k\}$.*

Closedness: We show $(S-A)$ is open. Let $x \in (S-A)$. So for every $z \in A$, $B(x, r_z)$ and $B(z, r_z)$ are disjoint, where $r_z = \frac{1}{2}\rho(x, z)$. Since A is compact, the open cover $\{B(z, r_z)|z \in A\}$ has a finite subcover $\{B(z_1, r_{z_1}), ..., B(z_k, r_{z_k})\}$. Let $r = \min\{r_{z_1}, ..., r_{z_k}\}$. Since $B(x, r)$ does not intersect with any of the sets in the finite subcover, it does not intersect with A. Thus $B(x, r) \subset (S - A)$.

■

Note in the proof of closedness that we could take the minimum of the radii because there were only a finite number of open balls due to compactness.

The converse is not true in general for metric spaces, and we'll give an example or two later.

Proposition 1 *A closed subset A of a compact metric space S is compact.*

Proof 13 *Let $\mathcal{O}$ be an open cover of A. So $\mathcal{O} \cup (S - A)$ is an open cover of S; since A is closed. So S has a finite subcover $\mathcal{B}$, which is also a finite subcover of $A \subset S$.* ∎

While the above was the Heine-Borel way to compactness, for metric spaces, there is a very useful alternative definition of compactness, as seen in the following theorem. This alternative notion is called *sequential compactness.*

Theorem 14 *A subset E of a metric space S is compact iff it is sequentially compact, i.e. every sequence in E has a subsequence that converges to a point in E.*

Proof 14 *Suppose E is compact, and suppose $(x_m) \subset E$ is a sequence that has no convergent subsequence with limit in E. Then, (x_m) must have infinitely many distinct points (else, some point will repeat an infinite number of times, forming a convergent subsequence).*

Around each distinct point x_i, there exists an open ball $B(x_i, \epsilon_i)$ that does not contain any other point x_j; (for, if every open ball around x_i had another point x_j from (x_m), then it would be a limit point of (x_m)).

So, $\{B(x_i, \epsilon_i), i = 1, 2, \ldots\}$ is an open cover of (x_m).

Now, since (x_m) does not have a limit point, it's a closed set; and since it's a subset of the compact set E, it's compact itself. So it has a finite subcover from $\{B(x_i, \epsilon_i), i = 1, 2, \ldots\}$; but then, (x_m) must have only finitely many distinct points. Contradiction.

Conversely, suppose E is sequentially compact, and suppose $\{U_\alpha | \alpha \in A\}$ is an open cover of E.

(i). Note first that E is totally bounded, i.e., for every $\epsilon > 0$, E can be covered with finitely many open ϵ-balls. For if this were not so, we could start with

some $x_1 \in E$, pick $x_2 \in E - B(x_1, \epsilon)$, pick $x_3 \in E - (B(x_1, \epsilon) \cup B(x_2, \epsilon))$, and so on. The resulting sequence $(x_m) \subseteq E$ would have every pair of points at a distance of at least ϵ, and so not have any convergent subsequence.
(ii). Now, there exists $r > 0$ such that for every point $x \in E$, $B(x, r) \subseteq U_\alpha$, for some $\alpha \in A$. For if this were not the case, then for every $1/m$, we would have a $x_m \in E$ s.t. $B(x_m, 1/m)$ is not a subset of any U_α. Since $(x_m) \subseteq E$, it has a convergent subsequence, call it (z_k), converging to $z \in E$. Since $z \in U_\alpha$, for some α, and this is open, $B(z, \epsilon) \subseteq U_\alpha$ for some $\epsilon > 0$. For k sufficiently large, z_k is very close to z, and we can get $B(z_k, \frac{\epsilon}{2}) \subseteq B(z, \epsilon) \subseteq U_\alpha$ (why?). Now $z_k = x_m$ for some m, and for k large enough, $1/m < \epsilon/2$. But then $(B(x_m, 1/m) \subseteq U_\alpha$. Contradiction.
(iii). For the r in Step (ii), E is covered by a finite number of r-balls. So there exist $x_1, \ldots, x_K$ in E s.t. $E \subseteq \cup_{i=1}^K B(x_i, r)$. And since for each i, $B(x_i, r) \subseteq U_i$ for some $i \in A$; so $\{U_i | i = 1, \ldots, K\}$ is a finite subcover of E.
■

From the Bolzano-Weierstrass theorem, it follows that $[a, b]$ is compact. Because, every sequence in $[a, b]$ is a bounded sequence, hence has a limit point in $[a, b]$. By the same token, the rectangle $[a, b]^k$ is compact in $\Re^k$. Indeed, take any sequence $(x_m) \subset [a, b]^k$. Then the coordinate sequence (x_m^1) is bounded in $[a, b]$, hence the coordinate sequence from some subsequence $(x_{m_1}^1)$ converges to some x^1 in $[a, b]$. The second coordinate sequence $(x_{m_1}^2)$ from this sequence is bounded in $[a, b]$, hence has a subsequence converging to some x^2. And so forth; so some subsequence converges to $x \in [a, b]^k$. We can now prove:

Theorem 15 *Suppose $A \subset \Re^k$ is closed and bounded. Then A is compact.*

Proof 15 *For some $x \in A$ and some $\epsilon > 0$ we have $A \subset B(x, \epsilon)$, since A is bounded. Choose some $a < \min\{x^j - \epsilon | j = 1, \ldots, k\}$ and some $b >$*

$\max\{x^j + \epsilon | j = 1, \ldots, k\}$. *Then* $A \subset B(x, \epsilon) \subset [a, b]^k$. *Since* A *is a closed subset of the compact set* $[a, b]^k$, *it is compact.* ■

For an arbitrary metric space, a closed and bounded subset may not be compact. As an example, consider the space $\mathcal{C}[0, 1]$ of all real-valued continuous functions defined on $[0, 1]$. Consider the closed unit ball $\bar{B}(\mathbf{0}, 1)$ or $\bar{B}$ for short, in this space. This is the subset of all functions f with $||f|| \leq 1$. Where $||f|| = \sup\{|f(x)| : x \in [0, 1]\}$.

In this space, $\bar{B}$ is closed, bounded, but not compact. Clearly, $\bar{B}(\mathbf{0}, 1) \subseteq B(\mathbf{0}, 2)$, so it's bounded. To see closedness, suppose the function f is an adherent point of $\bar{B}$. So, for an open ball $B(f, \epsilon)$, there is a function $g \in \bar{B}$ that lies in $B(f, \epsilon)$. So, $||f - g|| < \epsilon$. So,

$$||f|| = ||f - g + g|| \leq ||f - g|| + ||g|| < \epsilon + 1$$

(where the first inequality is the triangle inequality; also, $||g|| \leq 1$ as $g \in \bar{B}$). Since this is true for every $\epsilon > 0$, we have $||f|| \leq 1$, so $f \in \bar{B}$. So, $\bar{B}$ is closed.

To show $\bar{B}$ is not compact, consider the sequence $(f_n) \subseteq \bar{B}$ defined by $f_n(x) = x^n$, $n = 1, 2, \ldots$. No function $f \in \mathcal{C}[0, 1]$ can be the limit of any subsequence of (f_n). Suppose f is such a limit, and let $\epsilon = 1/4$, and let (f_{n_k}) be a subsequence. So, for each $k = 1, 2, ...$, $f_{n_k}(x) = x^{n(k)}$, where $n()$ is an increasing function. So, for $0 \leq x < 1$, $f_{n_k}(x) \to 0$. And for $x = 1$, $f_{n_k}(x) = 1$. Suppose $f(1) > 1/2$. By continuity, for high enough k, $||f_{n_k} - f|| \geq 1/2$. If $f(1) \leq 1/2$, the same conclusion holds. (Picture). So f is not a limit.

In fact, for $\mathcal{C}[a, b]$, the problem is that it's not *equicontinuous*. The Arzela-Ascoli theorem says that a subset E of $\mathcal{C}[a, b]$ is compact if and only if it is closed, bounded and equicontinuous. (See for example, Efe Ok).

3.1 Weierstrass' Theorem

Compact Sets and Optimization

You've seen the theorem below proved the Bolzano-Weierstrass way, so see the Heine-Borel argument below.

Theorem 16 *(Weierstrass) Let $f : S \to Y$, where S and Y are metric spaces and f is continuous. If $K \subset S$ is compact, then so is $f(K)$.*

Proof 16 *Suppose $\{G_\alpha\}$ is an open cover of $f(K)$. So $\{f^{-1}(G_\alpha)\}$ is an open cover of K. (Because $f^{-1}(\cup G_\alpha) = \cup f^{-1}(G_\alpha)$; 'open' follows from continuity). Since K is compact, there is a finite subcover $\{f^{-1}(G_1), ..., f^{-1}(G_k)\}$. So $\{G_1, ..., G_k\}$ is a finite subcover of $f(K)$.* ■

Since we've done the corollary which we use all the time, that is, if Y above is the set $\Re$ of real numbers, then f has a maximum and a minimum on K, I will forego that discussion here.

4 Complete Metric Spaces

Definition 14 *(x_n) is a Cauchy sequence in S if for every $\epsilon > 0$, there exists N such that $n, m \geq N$ implies $\rho(x_n, x_m) < \epsilon$*

Every convergent sequence is obviously Cauchy (since $\rho(x_n, x_m) \leq \rho(x_n, x) + \rho(x_m, x)$, where $x = \lim x_n$). But Cauchy sequences may not converge in a metric space. For instance, in the metric space $(Q, |\ \ |)$, Cauchy sequences of rationals need not converge. (Consider for instance the Cauchy sequence $(1.4, 1.41, 1.414, ...)$ does not converge in Q). Hence the need to "complete" the space to include the irrationals, which are all limits of rational Cauchy sequences.

Definition 15 *A metric space is complete if every Cauchy sequence in it converges to a point in the metric space.*

So $(\Re, |\ \ |)$ is a complete metric space.

To labor the point, note that a Cauchy sequence (x_m) is bounded. (Let $\epsilon > 0$, then there is an N s.t. $(x_m) \subset B(x_N, r)$, where

$r = \max\{\epsilon, \rho(x_1, x_N), \ldots, \rho(x_{N-1}, x_N)\}$).

So every Cauchy sequence (x_m) in $\Re$, being bounded, has a convergent subsequence (x_{m_k}) converging to a point $x \in \Re$ (by the B-W theorem). But then, $(x_m) \to x$. Indeed

$$\rho(x_m, x) \leq \rho(x_m, x_{m_k}) + \rho(x_{m_k}, x)$$

and the RHS becomes as small as we want; the first term due to Cauchy-ness.

Theorem 17 *(x_n) is Cauchy in $(\Re^k, d_p)$ iff all its coordinate sequences are Cauchy in $(\Re, |\ \ |)$.*

Proof 17 *(Sketch). Suppose all coordinate sequences $(x_n^j)$are Cauchy, and let $\epsilon > 0$. Let N be such that for all the coordinate sequences, $n, m \geq N$ implies $|x_n^j - x_m^j| < (\epsilon/k^{1/p})$. So $(\sum_{j=1}^k |x_n^j - x_m^j|^p)^{1/p} < \epsilon$. And so forth.* ■

Since $\Re$ is complete, it follows from above that so is $\Re^k$.

Convergence of Functions, Continuity.

Definition 16 *A sequence of real-valued functions (f_n) defined on U converges pointwise to a function f if for all $x \in U$, $|f_n(x) - f(x)| \to 0$.*

Continuity is not necessarily preserved by pointwise convergence; e.g., (f_n) defined on the domain $[0, 1]$ by $f_n(x) = x^n$, for all x and n, converges to the f defined by $f(x) = 0$ for $x \in [0, 1)$ and $f(1) = 1$, which is discontinuous at 1. We need uniform (d_∞ or sup norm) convergence for this.

Theorem 18 *Let (f_n) be real-valued continuous functions on U, and $d_\infty(f_n, f) \to 0$. Then f is continuous on U.*

Proof 18 *Let $\epsilon > 0$, and let $(x_k) \to \bar{x}$ in U. So for some N, $n \geq N$ implies $|f_n(x) - f(x)| < (1/3)\epsilon$, for all $x \in U$, by sup norm convergence of (f_n) to f. And for some K, $k \geq K$ implies $|f_n(x_k) - f_n(\bar{x})| < (1/3)\epsilon$, by continuity of f_n. So for $k \geq \max\{N, K\}$, by the triangle inequality,*

$|f(x_k) - f(\bar{x})| \leq$

$|f(x_k) - f_n(x_k)| + |f_n(x_k) - f_n(\bar{x})| + |f_n(\bar{x}) - f(\bar{x})|$

$< \epsilon$

∎

Theorem 19 *The metric space (bU, d_∞) is complete, for any domain U of functions.*

Proof. Let (f_n) be a Cauchy sequence in bU. So for each $x \in U$, $(f_n(x))$ is a Cauchy sequence of real numbers, and so converges to some real number: call this $f(x)$, so we have defined a function with pointwise convergence. Then $f \in bU$. Indeed, take any $\epsilon > 0$, so for $n \geq N$, some N, $|f(x) - f_n(x)| < \epsilon$. Take any such n. Since $f_n \in bU$, $|f_n(x)| < K$ for some $K > 0$.

So, $|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| < \epsilon + K$.

Second, we show (f_n) d_∞-converges to f. Take any $\epsilon > 0$. By Cauchyness, there exists N s.t. $n \geq N$ implies $d_\infty(f_n, f_N) < \epsilon/2$.

Therefore also for every $x \in U$, $|f_n(x) - f_N(x)| \leq d_\infty(f_n, f_N) < \epsilon/2$. Since $|f_n(x) - f_N(x)| \to |f(x) - f_N(x)|$ as $n \to \infty$, we have $|f(x) - f_N(x)| \leq \epsilon/2$. This is true for all x, so $d_\infty(f, f_N) \leq \epsilon/2$.

So we have that for $n \geq N$,

$$d_\infty(f, f_n) \leq d_\infty(f, f_N) + d_\infty(f_n, f_N) < \epsilon$$

∎

Note. (i) The continuous functions in (bU, d_∞) are a closed subset and therefore form a complete metric space (bcU, d_∞).

(ii). The sup norm plays a key role in the space being complete here. By way of a counterexample, the space $\mathcal{C}[a,b]$ with the L_2 norm, $||g|| = \left(\int_a^b |g(x)|^2 dx\right)^{1/2}$ is not complete. As an aside which you can ignore, note that the L_2 norm still has the virtue that it is induced by an inner product, namely $\langle f, g\rangle = \left(\int_a^b |f(x)g(x)|dx\right)^{1/2}$; whereas, the sup norm is not induced by any inner product.

4.1 Contraction Mapping Theorem

Definition 17 *Let (S, ρ) be a metric space and $T : S \to S$. T is nonexpansive if $\rho(Tx, Ty) \leq \rho(x, y)$ for all $x, y \in S$. T is a strict contraction if $\rho(Tx, Ty) < \rho(x, y)$ whenever $x \neq y$. T is a (uniformly strict) contraction with modulus λ if $\lambda < 1$ and for all $x, y \in S$, $\rho(Tx, Ty) \leq \lambda\rho(x, y)$.*

Note: 1. Nonexpansiveness itself is sufficient for continuity. If $(x_n) \to x$, $\rho(x_n, x) \to 0$. Since $\rho(Tx_n, Tx) \leq \rho(x_n, x)$, it tends to 0 as well, so $Tx_n \to Tx$.

2. If T is a strict contraction, it has at most 1 fixed point. Suppose x and y are fixed points. So $\rho(Tx, x) = \rho(Ty, y) = 0$. By the triangle inequality,

$\rho(x, y) \leq \rho(x, Tx) + \rho(Tx, Ty) + \rho(Ty, y) = \rho(Tx, Ty)$.

On the other hand, contraction implies $\rho(Tx, Ty) < \rho(x, y)$ if $x \neq y$. So $x = y$.

Examples. Talk about successive approximations.

Example 1. $f : [0, 1] \to [0, 1]$ defined by $f(x) = a + bx$, where $0 \leq a \leq 1$ and $|b| \leq 1$. f is nonexpansive; and a uniformly strict contraction if $|b| < 1$.

Example 2. $f : [1, \infty) \to [1, \infty)$ defined by $f(x) = x + \frac{1}{x}$, $\forall x \in [1, \infty)$. So,

$$f(x) - f(y) = x - y + \frac{1}{x} - \frac{1}{y} = (x - y)(1 - \frac{1}{xy})$$

.

So for all $x \neq y$, $|f(x) - f(y)| < |x - y|$, but f is not uniformly contracting, and in this example has no fixed point.

Theorem 20 *(Banach) If (S, ρ) is a complete metric space and T a contraction with modulus λ, T has exactly one fixed point.*

Proof 19 *Let $x_0 \in S$. Define (x_n) by $x_n = T^n x_0$, where T^n operates with T n successive times. (x_n) is Cauchy. Indeed, let $m \geq n$.*

$\rho(x_n, x_m) \leq \lambda \rho(x_{n-1}, x_{m-1}) \leq \ldots \leq$

$\lambda^n \rho(x_0, x_{m-n}) = \lambda^n \rho(x_0, T^{m-n} x_0)$

By the triangle inequality,

$\rho(x_0, x_{m-n}) \leq \rho(x_0, x_1) + \ldots + \rho(x_{m-n-1}, x_{m-n}) \leq$

$\rho(x_0, x_1)[1 + \lambda + \ldots + \lambda^{m-n-1}] \leq \frac{1}{1-\lambda} \rho(x_0, x_1)$

So $\rho(x_n, x_m) \leq \frac{\lambda^n}{1-\lambda} \rho(x_0, x_1)$, which can be made arbitrarily close to 0. Since (x_n) is Cauchy and S is complete, $x_n \to x \in S$.

x is a fixed point of T. Indeed, by continuity of T, $x_n \to x$ implies $T x_n \to T x$. On the other hand, $T x_n = T(T^n x_0) = x_{n+1}$, which clearly tends to x. Since $(T x_n)$ has a single limit, $Tx = x$. Uniqueness follows since T is a strict contraction. ■

4.2 Blackwell's Sufficient Conditions

Theorem 21 *(Blackwell). Let $M \subset bU$ s.t. $u \in M$ and $a \in [0, \infty) \Rightarrow u + a \in M$ (where $u + a$ is the function defined by $u(x) + a, \forall x \in U$). Let $T : M \to M$ be an operator satisfying (i) Monotonicity: $u \leq v \Rightarrow Tu \leq Tv$, and (ii) Discounting: There exists $\lambda \in [0, 1)$ s.t. $\forall u \in M$ and $a \in [0, \infty)$, $T(u + a) \leq Tu + \lambda a$. Then T is a uniformly strict contraction with modulus λ on the metric space (M, d_∞).*

Proof of Blackwell's theorem. Interpret the inequalities as holding pointwise for all $x \in U$.

$u(x) = v(x) + u(x) - v(x) \leq v(x) + |u(x) - v(x)| \leq v(x) + ||u - v||_\infty$.

Since T is monotone, applying it to both sides we have

$Tu \leq T(v + ||u - v||_\infty) \leq Tv + \lambda ||u - v||_\infty$.

The last inequality is due to the discounting property of T.

Thus $Tu - Tv \leq \lambda ||u - v||_\infty$.

Reversing the roles of u and v, we get

$Tv - Tu \leq \lambda ||u - v||_\infty$.

Combining the two, $|Tu - Tv| \leq \lambda ||u - v||_\infty$.

Since this inequality holds when the left hand is evaluated at arbitrary $x \in U$, it holds for the sup. Thus

$||Tu - Tv||_\infty \leq \lambda ||u - v||_\infty$. That is, T is a uniformly strict contraction.

■

4.3 Correspondences and the Maximum Theorem

At present, notes on this topic are not written. I will follow Section 3.3 of Stokey and Lucas for this part.