John Nachbar
Washington University in St. Louis
November 4, 2017

$\mathbb{R}^\omega$ Completeness and Compactness.[1]

1 Metrics on $\mathbb{R}^\omega$.

In contrast to $\mathbb{R}^N$, where the default metric is based on the Euclidean norm, there is no default norm and no default metric for $\mathbb{R}^\omega$. Instead, I consider two standard metrics, each of which can be useful, depending on the application.

The theme of these notes is: your intuition about compactness in $\mathbb{R}^\omega$ is probably wrong.

2 The space $(\ell^\infty, d_{\text{sup}})$.

Recall that ℓ^∞ is the subset of $\mathbb{R}^\omega$ such that $x \in \ell^\infty$ iff there is an M (which can depend on x) such that $|x_n| < M$ for all n. By LUB, this implies that, for any $x \in \ell_\infty$,

$$\|x\|_{\text{sup}} = \sup_n |x_n|$$

is well defined. $\|\cdot\|_{\text{sup}}$ is called the *sup norm.* I have already shown (in the notes on Vector Spaces and Norms) that $\|\cdot\|_{\text{sup}}$ satisfies the norm properties and hence the associated metric d_{sup} is indeed a metric.

2.1 Pointwise convergence in $(\ell^\infty, d_{\text{sup}})$.

An important fact about the sup metric, formalized in the next theorem, is that convergence under d_{sup} is more demanding than pointwise convergence in the sense that every sequence that converges under d_{sup} converges pointwise but not necessarily conversely. This is in contrast to the situation with $d_{\max}$ in $\mathbb{R}^N$, where convergence was equivalent to pointwise convergence.

Theorem 1. *Consider any sequence (x_t) in $(\ell^\infty, d_{\text{sup}})$.*

1. *If $x_t \to x^*$ then $x_{tn} \to x^*_n$ for all n.*
2. *If (x_t) is Cauchy then, for all n, (x_{tn}) is Cauchy.*

The converses of these statements are not true.

Proof. The proofs that convergence under d_{sup} implies pointwise convergence and that if (x_t) is Cauchy then the coordinate sequences are Cauchy are both essentially the same as for $d_{\max}$ in $\mathbb{R}^{\mathbb{N}}$, so I omit them.

To show that the converse statements are not true, it suffices to give an example for each. Consider the sequence (x_t) where

$$x_t = (0, \ldots, 0, 1, 0, \ldots),$$

with the 1 appearing in coordinate $n = t$. Then (x_t) converges pointwise to the origin; indeed, for every coordinate n, that coordinate is 0 for every $t \geq n + 1$. But this sequence does not converge under the sup metric: for every t, $d_{\text{sup}}(x_t, 0) = 1$.

Likewise, using the same example, this sequence is Cauchy in each coordinate but not Cauchy under d_{sup}. ■

2.2 $(\ell^\infty, d_{\text{sup}})$ is complete.

Theorem 2. *$(\ell^\infty, d_{\text{sup}})$ is complete.*

Proof. $(\ell^\infty, d_{\text{sup}})$. Fix any Cauchy sequence (x_t) in ℓ^∞. If (x_t) is Cauchy under d_{sup} then for each n each coordinate sequence (x_{tn}) is Cauchy. Since $\mathbb{R}$ is complete, for each n there is an x_n^* such that $x_{tn} \to x_n^*$. That is, (x_t) converges pointwise to $x^* \in \mathbb{R}^\omega$. I claim (1) $x_t \to x^*$ and (2) $x^* \in \ell^\infty$.

1. Because of Theorem 1, I cannot simply claim, as was the case in $\mathbb{R}^N$, that if x_t converges to x^* pointwise then it converges in the max/sup metric.

   Consider any $\varepsilon > 0$ and choose any $\hat{\varepsilon} \in (0, \varepsilon)$. Since (x_t) is Cauchy, there is a T such that for all $t, s > T$, $d_{\text{sup}}(x_t, x_s) < \hat{\varepsilon}/2$. I claim that for any $t > T$, $d_{\text{sup}}(x_t, x^*) < \varepsilon$. Therefore, fix any $t > T$.

   Consider any n. Since $x_{sn} \to x_n^*$, there is an $s > T$ (with possibly different s for different n) such that $|x_{sn} - x_n^*| < \hat{\varepsilon}/2$. Then by the triangle inequality,

   $$|x_{tn} - x_n^*| \leq |x_{tn} - x_{sn}| + |x_{sn} - x_n^*| < \hat{\varepsilon}.$$

   Since this holds for all n, $d_{\text{sup}}(x_t, x^*) \leq \hat{\varepsilon} < \varepsilon$, as was to be shown. Hence $x_t \to x^*$.

2. Consider any n. By the triangle inequality, for any t, $|x_n^*| \leq |x_n^* - x_{tn}| + |x_{tn}|$. Since $x_t \to_{\text{sup}} x^*$, there is a T such that for all $t > T$, $\sup_n |x_n^* - x_{tn}| < 1$. Also, for any t, $x_t \in \ell^\infty$, hence $\|x_t\|_{\text{sup}}$ is finite. Therefore, taking any $t > T$, $|x_n^*| < 1 + \|x_t\|_{\text{sup}}$, which implies $x^* \in \ell^\infty$.

■

2.3 Compactness in $(\ell^\infty, d_{\text{sup}})$.

Since $(\ell^\infty, d_{\text{sup}})$ is complete, a set in $(\ell^\infty, d_{\text{sup}})$ is compact iff it is closed and totally bounded. One would like to replace totally bounded with bounded, as we were able to do in $\mathbb{R}^N$, but this is not possible: Heine-Borel fails in $(\ell^\infty, d_{\text{sup}})$. In particular, in $\mathbb{R}^N$ the canonical example of a compact set is a closed ball. In $(\ell^\infty, d_{\text{sup}})$, closed balls are not compact.

Theorem 3. *In $(\ell^\infty, d_{\text{sup}})$, for any $x \in \ell^\infty$ and any $\varepsilon > 0$, $\overline{N_\varepsilon(x)}$ is not totally bounded and hence not compact.*

Proof. In the notes on Metric Spaces, I noted that the set $A \subseteq \overline{N_1(0)}$ with elements $(1, 0, 0, \dots)$, $(0, 1, 0, 0, \dots)$ is not totally bounded, which implies that $\overline{N_1(0)}$ is not totally bounded, hence not compact. The argument generalizes easily to any closed ball. ■

Remark 1. A slightly different way to see that $\overline{N_1(0)}$ is not compact is the following. Form the sequence (x_t) in $\overline{N_1(0)}$, with $x_t = (0, \dots, 0, 1, 0, \dots)$, where the 1 appears in coordinate $n = t$. Any pair of terms is distance 1 apart in the sup metric, hence the sequence has no Cauchy subsequence, hence no convergent subsequence. Therefore $\overline{N_1(0)}$ is not sequentially compact, hence not compact. □

So what *is* compact in $(\ell^\infty, d_{\text{sup}})$? The short answer is: not much. The easiest non-trivial example is the *Hilbert cube*: $A = [0, 1] \times [0, \frac{1}{2}] \times [0, \frac{1}{3}] \times \dots$. In $\mathbb{R}^2$, $[0, 1] \times [0, \frac{1}{2}]$ a rectangle with one side of length 1 and the other of length 1/2. In $\mathbb{R}^3$, $[0, 1] \times [0, \frac{1}{2}] \times [0, \frac{1}{3}]$ is a rectangular solid with one side of length 1, the second side of length 1/2, and the third side of length 1/3. And so on. The Hilbert cube is the infinite-dimensional version of this.

More generally, a rectangular solid in $(\ell^\infty, d_{\text{sup}})$ is compact iff for every $\varepsilon > 0$, there are only finitely many sides of length greater than ε. Any closed set that is contained in such a rectangular solid is likewise compact.

2.4 A Generalization of Theorem 3.

Theorem 3 has an extreme generalization: there is *no* infinite-dimensional normed vector space in which closed balls are compact (in the metric generated by the norm). Put differently, the Heine-Borel theorem fails catastrophically in every infinite-dimensional normed vector space. Because this runs so sharply counter to finite-dimensional intuition, I give the argument in detail here.

Theorem 4. *Let X be a normed vector space. If X is infinite dimensional then for any $x \in X$ and any $\varepsilon > 0$, $\overline{N_\varepsilon(x)}$ is not totally bounded and hence not compact.*

Proof. Let the norm be f and the associated metric d. For the moment, focus on the unit ball centered at the origin.

$$B = \overline{N_1(0)} = \{x \in X : f(x) \leq 1\}.$$

I claim that there is a countably infinite set $A = \{a_1, a_2, \dots\} \subseteq B$ such that, for any $a, \hat{a} \in A$, $d(a, \hat{a}) > 1/2$. Assuming that the claim is true, the proof follows, since A is not totally bounded (it cannot be covered by any finite set of open 1/2 balls), hence X is not totally bounded. It remains to show the existence of the set A.

I define A recursively as follows. Let a_1 be any element of B, $a_1 \neq 0$. Let M_1 be the vector subspace spanned by a_1. M_1 is closed since, as discussed in the notes on Completeness in $\mathbb{R}^N$, any finite-dimensional normed vector space is complete, and hence closed as a vector subspace of a normed vector space.

Suppose that we are at stage T in this construction, with $\{a_1, \dots, a_T\} \subseteq B$ and with M_T the finite-dimensional vector subspace spanned by $\{a_1, \dots, a_T\}$. Again, M_T is closed. I claim that there is an a_{T+1} such that $f(a_{T+1}) = 1$ (hence $a_{T+1} \in B$) and $d(a_{T+1}, x) > 1/2$ for every $x \in M_T$ (hence, in particular, $d(a_{T+1}, a_t) > 1/2$ for every $t \leq T$). I am done if I can show that this a_{T+1} exists.

The existence of a_{T+1} follows from a result called the Riesz Lemma. Here is a proof for this special case.

Since M_T is finite-dimensional and X is infinite dimensional, $X \setminus M_T \neq \emptyset$. Take any $y \in X \setminus M_T$ and consider the problem $\min_{x \in M_T} d(y, x)$. In a general infinite dimensional setting, this minimization problem may not have a solution. However, we can finesse this issue as follows.

The set $\{d(y, x)\}_{x \in M_T}$ is bounded below (by 0) and hence has an infimum. Let $\alpha = \inf\{d(y, x)\}_{x \in M_T}$ and note that $\alpha > 0$.[2] There is an $x^* \in M_T$ such that $d(y, x^*) < 2\alpha$ (since otherwise, 2α would be a lower bound of $\{d(y, x\}_{x \in M_T}$ and hence α wouldn't be the greatest lower bound). Let

$$a_{T+1} = (y - x^*) \frac{1}{d(y, x^*)}.$$

This is well defined since $d(y, x^*) > 0$ (since $d(y, x^*) \geq \alpha > 0$). Then $f(a_{T+1}) = 1$

[2]By contraposition. Suppose that for any $t \in \{1, 2, \dots\}$ there is an $x_t \in M_T$ such that $d(y, x_t) < 1/t$. Then $x_t \to y$, which implies, since M_T is closed, that $y \in M_T$.

and for any $x \in M_T$,

$$\begin{aligned}
d(a_{T+1}, x) &= f(a_{T+1} - x) \\
&= f\left((y - x^*)\frac{1}{d(y, x^*)} - x\right) \\
&= \frac{1}{d(y, x^*)} f(y - x^* - d(y, x^*)x) \\
&= \frac{1}{d(y, x^*)} d(y, x^* + d(y, x^*)x) \\
&\geq \frac{\alpha}{d(y, x^*)} \\
&> \frac{\alpha}{2\alpha} \\
&= \frac{1}{2},
\end{aligned}$$

where the first inequality comes from the fact that $x^* + d(y, x^*)x \in M_T$, since $x, x^* \in M_T$ and M_T is a vector space. This completes the proof for the case $x = 0$, $\varepsilon = 1$.

For a general radius $\varepsilon > 0$, take

$$a_{T+1} = (y - x^*)\frac{\varepsilon}{d(y, x^*)}.$$

Then $f(a_{T+1}) = \varepsilon$ and $d(a_{T+1}, x) > \varepsilon/2$ for any $x \in M_T$. Then once again A, and hence $N_\varepsilon(0)$, is not totally bounded because A cannot be covered by any finite set of open $\varepsilon/2$ balls.

Finally, the proof of the general case, for any center x, follows by the preceding arguments, since, in a normed vector space, $\overline{N_\varepsilon(x)}$ is totally bounded iff $\overline{N_\varepsilon(0)}$ is. In particular, in a normed vector space, a set of open balls, call it $\mathcal{O}$, covers $\overline{N_\varepsilon(0)}$ iff the open balls $O + \{x\}$, $O \in \mathcal{O}$, cover $\overline{N_\varepsilon(x)}$. ■

3 The space $(\mathbb{R}^\omega, d_{\mathbf{pw}})$.

Since the problems with compactness in $(\ell^\infty, d_{\text{sup}})$ seem to have something to do with the failure of equivalence between d_{sup} convergence and pointwise convergence, it is natural to look at metrics for which convergence *is* equivalent to pointwise convergence. It is indeed the case that compactness is less of a rarity for such metrics, but there are still issues.

3.1 The d_{pw} metric.

For any $a, b \in \mathbb{R}^\omega$, let

$$d_{\text{pw}}(a, b) = \sup_n \frac{\min\{1, |a_n - b_n|\}}{n}.$$

I first introduced this metric in the Metric Space notes. Under d_{pw}, convergence is pointwise and the Cauchy property holds iff it holds in each coordinate.

Theorem 5. *Consider any sequence (x_t) in $(\mathbb{R}^\omega, d_{\text{pw}})$.*

1. *$x_t \to x^*$ iff $x_{tn} \to x^*_n$ for all n.*
2. *(x_t) is Cauchy iff for all n, (x_{tn}) is Cauchy.*

Proof.

1. $\Rightarrow$ Consider any sequence (x_t) and suppose that $x_t \to x$ under d_{pw}. Fix any $\varepsilon \in (0, 1)$ and any coordinate n. Then there is a T such that for all $t > T$, $d_{\text{pw}}(x_t, x) < \varepsilon/n$ which implies that, in particular,

   $$\frac{\min\{1, |x_{tn} - x_n|\}}{n} < \frac{\varepsilon}{n},$$

   hence

   $$\min\{1, |x_{tn} - x_n|\} < \varepsilon.$$

   Since $\varepsilon < 1$, this implies $|x_{tn} - x_n| < \varepsilon$, as was to be shown.

   $\Leftarrow$. Suppose that $x_t \to x$ pointwise. Fix $\varepsilon > 0$. Choose N sufficiently large that $1/N < \varepsilon$. For each $n \leq N$, choose T_n such that for all $t > T_n$, $|x_{tn} - x_n| < \varepsilon$. Let $T = \max\{T_1, \ldots, T_N\}$; the max operation is well defined since the set is finite. Then for all $t > T$ and all $n \leq N$,

   $$\frac{\min\{1, |x_{tn} - x_n|\}}{n} < \varepsilon.$$

   The construction of N then implies that $d_{\text{pw}}(x_t, x) < \varepsilon$, as was to be shown.

2. The proof is very similar to the one for convergence, so I omit it.

■

3.2 $(\mathbb{R}^\omega, d_{\text{pw}})$ is complete.

Theorem 6. $(\mathbb{R}^\omega, d_{\text{pw}})$ *is complete.*

Proof. $(\mathbb{R}^\omega, d_{\text{pw}})$. The proof is essentially the same as the proof for $\mathbb{R}^N$. By Theorem 5, if (x_t) is Cauchy under d_{pw} then for each n, (x_{tn}) is Cauchy. This implies that (x_t) converges to x^*_t pointwise. By Theorem 5, this implies $x_t \to x^*$. ■

3.3 Compactness in $(\mathbb{R}^\omega, d_{\mathbf{pw}})$, Part A.

Since $(\mathbb{R}^\omega, d_{\text{pw}})$ is complete, a set in $(\mathbb{R}^\omega, d_{\text{pw}})$ is compact iff it is closed and totally bounded. Once again one would like to replace totally bounded with bounded but once again this is not possible. In particular, in $(\mathbb{R}^\infty, d_{\text{pw}})$, closed balls are not compact.

Theorem 7. *In $(\mathbb{R}^\omega, d_{\text{pw}})$, for any $x \in \mathbb{R}^\omega$ and any $\varepsilon > 0$, $\overline{N_\varepsilon(x)}$ is not compact.*

Proof. Consider first the case $x = 0$. Choose any coordinate n such that $1/n \leq \varepsilon$. Consider the sequence (x_t) defined by $x_t = (0, \dots, 0, t, 0, \dots)$, with the value t appearing in coordinate n. Thus $x_1 = (0, \dots, 0, 1, 0, \dots)$, $x_2 = (0, \dots, 0, 2, 0, \dots)$, and so on. For any such x_t, $d_{\text{pw}}(x_t, 0) = 1/n \leq \varepsilon$, hence $x_t \in \overline{N_\varepsilon(0)}$. But no subsequence of (x_t) converges (since no subsequence converges in coordinate n), hence $\overline{N_\varepsilon(0)}$ is not sequentially compact, hence $\overline{N_\varepsilon(0)}$ is not compact.

The argument for a general center x is almost identical, since $x_t + x \in \overline{N_\varepsilon(x)}$ iff $x_t \in \overline{N_\varepsilon(0)}$. ■

3.4 Pointwise convergence on $\mathbb{R}^\omega$ is weird.

By a *pointwise convergence metric* on $\mathbb{R}^\omega$, I mean any metric for which convergence is equivalent to pointwise convergence. Theorem 7 looks like an artifact of the particular pointwise convergence metric that we used. Is there another pointwise convergence metric that doesn't have the defects of d_{pw}? The answer is, not really. The problems are with pointwise convergence rather than with d_{pw}.

Let Z^n be the set of all points of the form $z^n = (0, \dots, 0, z^n_n, 0, \dots)$. Thus $z^1 = (z^1_1, 0, 0, \dots)$, $z^2 = (0, z^2_2, 0, 0, \dots)$, and so on. For any pointwise convergent metric and any set $\{z^1, z^2, \dots\}$, the associated sequence $(z^1, z^2, \dots)$ converges to the origin, *regardless of the values of the z^n_n.* This observation has a number of consequences.

1. *For any pointwise convergence metric, any x, and any $\varepsilon > 0$, $N_\varepsilon(x)$ contains $x + z^n$ for every $z^n \in Z^n$, for all but at most a finite set of n.*

   For example, for d_{pw}, $x = 0$ and $\varepsilon = 1/3$,

   $$N_{1/3}(0) = (-1, 1) \times (-2, 2) \times (-3, 3) \times \mathbb{R} \times \mathbb{R} \times \dots$$

   In particular, every $z^n \in N_{1/3}(0)$ for every $n \geq 4$.

   The proof of the general case is by contraposition. Suppose that for a metric d there is some $\varepsilon > 0$ and infinitely many n such that, for each such n, there is a z^n such that $x + z^n \notin N_\varepsilon(x)$. Then the infinite sequence formed by taking one such $x + z^n$ for each such n does not converge to x under the d metric. But this sequence does converge to x pointwise, and hence d is not a pointwise convergence metric.

2. *For any pointwise convergence metric, any x, and any $\varepsilon > 0$, $\overline{N_\varepsilon(x)}$ is not compact.*

   The argument is similar to that for Theorem 7. By point (1) above, given $\varepsilon > 0$, there is a coordinate n such that every point of the form $x + z^n$ is in $\overline{N_\varepsilon(x)}$. Form the sequence (x_t) with $x_t = x + z_t^n$ and $z_t^n = (0, \ldots, 0, t, 0, \ldots)$, with the t in coordinate n (i.e., $z_{tn}^n = t$). No subsequence of (x_t) converges since no subsequence converges in coordinate n, hence $\overline{N_\varepsilon(x)}$ is not sequentially compact, hence not compact.

3. *A pointwise convergence metric cannot be generated by a norm.*

   Let d be any pointwise convergence metric and consider any norm f. By point (1) above, for any $\varepsilon > 0$ there is an n such that $d(z^n, 0) < \varepsilon$ for any $z^n \in Z^n$. Take any $z^n \neq 0$. By the first property of a norm, $f(z^n) > 0$. By the second property of a norm, for any $\gamma > 0$, $f(\gamma z^n) = \gamma f(z^n)$. On the other hand, for any γ, $\gamma z^n = (0, \ldots, 0, \gamma z^n, 0, \ldots)$, hence $\gamma z^n \in Z^n$, hence $d(\gamma z^n, 0) < \varepsilon$. Take $\gamma > 0$ such that $\gamma f(z^n) > \varepsilon$. Therefore,

   $$d(\gamma z^n, 0) < \varepsilon < \gamma f(z^n) = f(\gamma z^n).$$

   Since $d(\gamma z^n, 0) \neq f(\gamma z^n)$, d was not generated by f. Since f was arbitrary, d cannot be generated by any norm.

   A related point is that if I restrict attention to ℓ^∞, then I can define the metric d_{pw^*} by,

   $$d_{\mathrm{pw}^*}(a, b) = \sup_n \frac{|a_n - b_n|}{n}.$$

   Although similar in appearance to the pw metric, the pw* metric is not a pointwise convergence metric. One way to see this is to note that the pw* metric is generated by the pw* norm,

   $$\|x\|_{\mathrm{pw}^*} = \sup_n \frac{|x_n|}{n}.$$

   I mentioned this norm briefly in the Normed Vector Space notes, in the section on equivalent norms. A pointwise convergence metric cannot be generated by a norm, whether the pw* norm or any other norm.

4. *For any pointwise convergence metric, $\mathbb{R}^\omega_{++} = \{x \in \mathbb{R}^\omega : x_n > 0 \ \forall n\}$ is not open.*

   Consider any point $x \in \mathbb{R}^\omega_{++}$ and any $\varepsilon > 0$. By property (1) above, there is an n such that $x + z^n \in N_\varepsilon(x)$ for every $z^n \in Z^n$. Then, in particular, $x + z^n \in N_\varepsilon(x)$ for z^n such that $x_n + z_n^n < 0$. Hence $N_\varepsilon(x) \not\subseteq \mathbb{R}^\omega_{++}$.

Finally, the next theorem records that under any pointwise convergence metric, ℓ^∞ is not closed as a subset of $\mathbb{R}^\omega$, hence not complete. In contrast, recall that $(\ell^\infty, d_{\sup})$ is complete (Theorem 2).

Theorem 8. *ℓ^∞ is not complete under any pointwise convergence metric.*

Proof. Consider the sequence (x_t) in ℓ^∞ where $x_t = (1, 2, \dots, t-1, t, 0, 0 \dots)$. Thus $x_1 = (1, 0, 0, \dots)$, $x_2 = (1, 2, 0, 0, \dots)$, and so on. Then x_t converges pointwise to $x^* = (1, 2, 3, \dots) \in \mathbb{R}^\omega$, but x^* is not an element of ℓ^∞. ■

3.5 Compactness in $(\mathbb{R}^\omega, d_{\mathbf{pw}})$, Part B.

There is, however, some good news about compactness in $(\mathbb{R}^\omega, d_{\text{pw}})$. Even though closed balls in $(\mathbb{R}^\omega, d_{\text{pw}})$ are not compact (Theorem 7), many other sets are.

Theorem 9 (Tychonoff's Theorem for $(\mathbb{R}^\omega, d_{\text{pw}})$)**.** *Consider $A \subseteq \mathbb{R}^\omega$, with $A = A_1 \times A_2 \times \dots$, each $A_n \subseteq \mathbb{R}$. If A_n is compact for all n then A is compact.*

Proof. Let (x_t) be a sequence in A. I will show that (x_t) has a convergent subsequence and hence is sequentially compact.

Since A_1 is compact, there is a subsequence (x_{tk}) and an $x_1^* \in A_1$ such that $x_{t_k 1} \to x_1^*$. Let y_1 be the any term in this subsequence.

Since A_2 is compact, there is subsequence of (x_{tk}), a sub-subsequence of the original sequence, and an $x_2^* \in A_2$ such that the second coordinate of this sub-subsequence converges to x_2^*. (I'm trying to avoid introducing too much notation here; the idea, I hope, is clear.) Let y_2 be any term in this sub-subsequence that comes after y_1 in the original sequence.

Continuing in this way, I construct a sequence $(y_k) = (y_1, y_2, \dots)$, which is a subsequence of the original sequence (x_t), and a point $x^* = (x_1^*, x_2^*, \dots) \in A$ such that (y_k) converges pointwise to x^*. ■

Example 1. The "anti"-Hilbert cube

$$[0, 1] \times [0, 2] \times [0, 3] \times \dots$$

is compact in $(\mathbb{R}^\omega, d_{\text{pw}})$. □

We also have the following.

Theorem 10. *Let $C \subseteq \ell^\infty$ be a closed sup-ball: $C = \{x \in \ell^\infty : \|x - a\|_{\text{sup}} \leq \varepsilon\}$ for some $\varepsilon > 0$ and some $a \in \ell^\infty$. Then C is compact as a subset of $(\mathbb{R}^\omega, d_{\text{pw}})$.*

Proof. For ease of exposition, consider a closed ε sup-ball centered on the origin. Then

$$C = \{x \in \ell^\infty : |x_n| \leq \varepsilon \; \forall \, n\} = [-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon] \times \dots,$$

which is a countable product of compact sets. By Tychonoff, C is compact. ■

Note the "mix-and-match" flavor of Theorem 10: we use *different* metrics to define the ball and to define compactness. By Theorem 4, if we define both the ball and compactness in terms of a single norm-based metric, then there is no infinite dimensional space in which the ball is compact.

A related point is that Theorem 3 and Theorem 10 together show that there are sets in ℓ^∞ that are compact under d_{pw} but not under d_{sup}. In contrast, any set that is compact under d_{sup} is compact under d_{pw}, since any sequence that converges under d_{sup} converges under d_{pw} (by Theorem 1 and Theorem 5).

Theorem 10 is a special case of an important result called the Banach-Alaoglu theorem. Stating Banach-Alaoglu carefully would take me afield, but the underlying intuition is the same as in Theorem 10.