Banach-Alaoglu theorems

László Erdős

Jan 23, 2007

1 Compactness revisited

In a topological space a fundamental property is the **compactness (Kompaktheit)**. We recall the definition:

Definition 1.1 *A subset $K \subset X$ of a topological space is* **compact***, if from any open covering one can always select a finite covering, i.e. whenever*

$$K \subset \bigcup_{\alpha \in A} O_\alpha$$

(the index set A is of arbitrary cardinality), then there exist finitely many $\alpha_1, \ldots \alpha_m$ such that

$$K \subset \bigcup_{j=1}^{m} O_{\alpha_j}$$

Lemma 1.2 *The continuous image of a compact set is compact*

We remark, that the compactness is an internal property, while closedness is not. Being internal means the following. Suppose we have a topological space Y and let $X \subset Y$. Then X inherits a natural topology from Y (called **subspace** topology (Unterraumtopologie) – but there is no vectorspace here), namely one can simply define

$$\mathcal{O}_X := \{O \cap X \ : \ O \in \mathcal{O}_Y\}$$

i.e. the open sets in X are declared to be the traces (intersections with X) of the open sets in Y. One can easily check that this collection satisfies (i)–(iii) above. Now if $K \subset X$, then K is also a subset of Y, so K lies in two different topological spaces. When one asks a question whether K is closed, open or compact, in principle one has to specify in which space K lies. A property of K is called internal, if the answer does not depend on whether we consider K as a subset of Y or X. Check that compactness is such a property, but closedness and openness are not, i.e. these two concepts depend on the ambient space. Also think over whether the concept of convergent sequences and continuous functions are internal or not.

Compactness in many (most) cases are used in a somewhat different form: instead of selecting an open cover, one would like to select a convergent subsequence from any sequence.

Definition 1.3 *A subset $K \subset X$ of a topological space X is called* **sequentially compact (Folgenkompakt)** *if for any sequence $\{x_n\} \subset K$ there exists a convergent subsequence with limit in K, i.e. there is a sequence n_k, $k = 1, 2, \dots$ and $x \in K$ such that $x_{n_k} \to x$.*

Compactness and sequential compactness are not the same in general topological spaces, but in real applications they usually coincide.

Theorem 1.4 *In every metric space compactness and sequential compactness are the same.*

The easy direction (compactness implies sequential compactness) goes exactly as in the real case (Bolzano-Weierstrass). The other direction takes a bit more effort, see Werner Satz B.1.7.

In real applications one would like to check compactness. This is hard, since the definition itself is almost uncheckable directly. In $\mathbf{R}^n$ we know that compact is the same as closed and bounded, and we know that this does not hold in infinite dimensional normed spaces. The hope is that one can still save this characterization, but maybe in a different topology. This will be the main message of the various versions of the Banach-Alaoglu theorem.

2 Tychonov theorem

2.1 Direct product

Let X_α, $\alpha \in A$ be a family of sets, parametrized by a set A which can have arbitrary cardinality.

Definition 2.1 *The direct product of X_α nonempty sets is given by*

$$\times_{\alpha \in A} X_\alpha = \{f : A \to \bigcup_{\alpha \in A} X_\alpha,\ f(\alpha) \in X_\alpha\}$$

i.e. it is the set of maps with domain A such that for each $\alpha \in A$ the map selects an element of X_α.

It is a nontrivial claim that $\times_{\alpha\in A} X_\alpha$ is not empty (if X_α is not empty for all α). It is so nontrivial, that (surprisingly) one cannot even prove it from the standard axioms of set theory. It is an additional axiom:

Axiom of choice (Auswahlaxiom): If X_α is not empty for all $\alpha \in A$, then $\times_{\alpha\in A} X_\alpha$ is not empty.

It is fairly easy to show the axiom of choice is equivalent to the Zorn lemma.

Now suppose that all X_α are topological spaces. We would like to define a "natural" topology on the direct product. Our requirement is that all projection maps

$$\Pi_\alpha : \times_{\alpha\in A} X_\alpha \to X_\alpha, \qquad \Pi_\alpha(f) := f(\alpha)$$

be continuous, but we want to do it as "economically" as possible. This construction works in general; given a set and a family of maps from the set to topological spaces, there is a natural way to define an "economical" topology on the set so that all functions be continuous.

2.2 Topology generated by maps

So far we have seen that if somebody gives us a topology (collection of open sets) we can decide whether a function is continuous. We can reverse this argument. Suppose we have a collection $\mathcal{F}$ of functions defined on Y with range in some topological space:

$$\mathcal{F} \subset \{f : Y \to X\}$$

or even it can happen that the target space for each f is different

$$\mathcal{F} \subset \{f : Y \to X_f\}$$

where X_f is a topological space.

We would like to define a topology on Y such that all $f \in \mathcal{F}$ be continuous. We thus need to specify the collection of open sets, $\mathcal{O}_Y$. This is not hard, if we choose $\mathcal{O}_Y = P(Y)$ (power set), then we are done, but this is a useless definition. In this topology nothing will converge (apart from the constant sequence), it is too strong.

So we aim at the weakest (=least number of open sets) topology so that all $f \in \mathcal{F}$ are still continuous. To construct such a topology is not straightforward. Of course every preimage $f^{-1}(O)$ of every open set $O \in X_f$ under any $f \in \mathcal{F}$ must be included in $\mathcal{O}_Y$, but we need many more sets to satisfy (i)–(iii), so we can keep on adding finite intersections and arbitrary unions of elements of $\mathcal{O}_Y$. But it is very easy to see that such topology exists. Simply take all families $\mathcal{O}$ that satisfy (i)–(iii) and that contain all sets of the form $f^{-1}(O)$, $O \in X_f$, open, $f \in \mathcal{F}$. There is at least such family, namely $\mathcal{O} = P(Y)$. Now take the intersection of all these families. This is a family, that still satisfies (i)–(iii) (think over) and still contains all $f^{-1}(O)$, and it is the smallest for these properties.

The upshot is that for any given collection of functions, there is the weakest topology so that all these functions are continuous. Of course it does not mean that only these functions are continous, clearly the sums, products etc. of such functions are still continuous.

Finally a definition that will be used in the general version of the Banach-Alaoglu theorem:

Definition 2.2 *X is a Banach space, then the weak* topology on its dual, X^* is the weakest topology that makes all functionals of the form*

$$x : \ell \mapsto \ell(x)$$

from X^* to the scalars, continuous.

2.3 Direct product of compact sets

The best example to see the concept of the topology of the direct product is the space $L^\infty(X)$, which is the set of everywhere defined bounded functions. Note that there is no measure on X, there is no "almost everywhere". With the usual supremum norm it is a normed space, hence metric, in particular continuity can be checked via sequences.

We can define a topology different from the norm topology on this set $L^\infty(X)$ by requiring that all functions $x : f \mapsto f(x)$ be continuous (and consider the weakest such topology). Don't get misled by the letters, here x denotes the function (on the space of functions f) and f is its variable. For example the following sets are open

$$S_{x,c,\varepsilon} := \{f \in L^\infty(X) \ : \ |f(x) - c| < \varepsilon\}$$

and by (i)–(iii) all finite intersections and arbitrary unions of such sets are also open.

When restricted to the unit ball, $[-1, 1]^X$, you can see that this is weakest topology on the direct product that makes all projections continuous, because the map

$$x : f \in [-1, 1]^X \mapsto f(x)$$

is exactly the projection map onto the x-component of the direct product.

This new topology is not the norm topology. This is not easy to see directly. One argument is that in the norm topology the unit ball is not compact, while in the new topology the unit ball is compact. To see this, we need Tychonoff theorem:

Theorem [Tychonoff] Let K_α, $\alpha \in A$ be a family of compact topological spaces. Then their direct product, $\times_{\alpha\in A} K_\alpha$ is compact in the product topology, i.e. in the weakest topology that makes all projections continuous.

An equivalent version is:

Theorem [Tychonoff], second version Let X_α, $\alpha \in A$ be topological spaces and $K_\alpha \subset X_\alpha$ be compact. Then $\times_{\alpha\in A} K_\alpha$ is a compact subset of $\times_{\alpha\in A} X_\alpha$ in the product topology, i.e. in the weakest topology that makes all projections continuous.

This theorem is proved in Reed-Simon (Thm IV.5 and at the end: Supplement to IV.3) or in Werner Thm. B. 2.10, but it is not short. For countable direct product a Cantor diagonalization trick works (EXERCISE).

3 Banach-Alaoglu Theorem

The goal is to find a checkable characterization of compact sets. Actually, for applications, we will rather need the sequential compactness, because eventually we want to select convergent subsequence.

We will mention three versions of Banach-Alaoglu Theorem of different generality (and difficulty). We will prove only the simplest version (Version II), which is actually the most often used. We start presenting the version on the "middle" level, which does not require the concept of w^*-convergence.

3.1 Banach-Alaoglu theorem for reflexive spaces

We first need a natural definition, it is self-explanatory:

Definition 3.1 *A subset K of a Banach space is* **weakly sequentially compact** *(in short, w.s.c.) if for any sequence $\{x_n\} \subset K$ there is a weakly convergent subsequence with limit in K, i.e. there is n_k numerical sequence and $x \in K$ such that $x_{n_k} \rightharpoonup x$.*

Theorem 3.2 (Version I.) *Let X be a reflexive Banach space (i.e. $X = X^{**}$). Let $B = \{x \in X \,:\, \|x\| \leq 1\}$ be its closed unit ball. Then B is w.s.c.*

The proof is in Werner III.3.7. and it is based on the Cantor diagonalization trick. A similar, but somewhat simpler proof will be presented that will prove Version II (Helly's theorem).

The most important application is when $X = L^p(M, \mu)$ for any measure space M, if $1 < p < \infty$.

It sounds that the unit ball plays a special role, but it is not really the case. First, by scaling it is clear that the same holds for any closed ball of radius R, i.e. for $B_R = \{x \in X : \|x\| \leq R\}$. More importantly, the ball itself plays no role, but its convexity does. The following corollary is really a trivial generalization of the above theorem:

Corollary 3.3 (to Version I.) *Let X be a reflexive Banach space and let $K \subset X$ be closed, bounded and convex. Then K is w.s.c.*

This Corollary trivially follows from Version I. plus Mazur's theorem (recall: closed and convex sets are weakly closed). Simply take a sufficiently big R such that $K \subset B_R$ (by boundedness). If $(x_n) \subset K$, then $(x_n) \subset B_R$, so there is a weakly convergent subsequence. But the limit stays in K since K is weakly closed by Mazur.

The reflexivity requirement is not really the optimal one. The "right" concept is to introduce the weak* convergence.

3.2 Weak* convergence

We already defined the weak* topology on the dual of a Banach space (Definition 2.2). Here we explicitly show what the convergence in the weak* topology means:

Definition 3.4 *Suppose that the normed space where we want to work, U, itself is the dual of some Banach space, i.e. $U = X^*$ (recall that not everything is a dual space). A sequence $(u_n) \subset U$* **converges in weak*-sense or in the weak* topology** *to $u \in U$, if $u_n(x) \to u(x)$ for any $x \in X$.*

Viewing x as an element $\hat{x} \in X^{**}$ by the canonical embedding $\hat{x}(u) = u(x)$, we see that weak and weak* convergences are equivalent for reflexive spaces, but in general this is not the case.

An example, when they are different, is an approximate Dirac delta function viewed as a measure. More precisely, let $K = [-1, 1]$ and consider the set of continuous functions, $C(K)$. We need the following theorem (without proof, see Reed-Simon IV.14 and the supplement of RS for a special case).

Theorem 3.5 (Riesz-Markov) *If K is a compact (Hausdorff) topological space, then the dual of $C(K)$ is isomorphic to the space $\mathcal{M}(K)$ of (complex) measures on K, equipped with the total variation norm.*

(Note: A topological space is Hausdorff if any two different points can be separated, i.e. for any $x \neq y$ there are disjoint open neighborhoods U_x and U_y of x and y, respectively.)

Armed with this theorem, we consider the measure μ_n on $[-1, 1]$ with density function $f_n(x) := \frac{n}{2}\mathbf{1}(|x| \leq 1/n)$. The action of the measure on elements of $C(K)$ is simply integration:

$$\mu_n(f) := \int f \mathrm{d}\mu_n$$

It is clear that μ_n converges in *weak*$*$-sense to the Dirac-delta measure at the origin, since for any continuous function

$$\mu_n(f) = \int f \mathrm{d}\mu_n \to f(0) = \delta(f)$$

However, μ_n does not converge to δ weakly. To see this, check that $L^\infty(K) \subset \mathcal{M}^*(K)$ (again, L^∞ is understood as functions defined at all points not just for almost all points, since there is no canonical measure on K anyway) and test the presumed weak convergence $\mu_n \rightharpoonup \delta$ on the characteristic function of the point zero (as element of $L^\infty \subset \mathcal{M}^*$).

Since $X \subset X^{**}$, it is clear that weak* convergence (on $U = X^*$) weaker than weak convergence, it requires checking less limits. The standard boundedness theorems are nevertheless true:

Theorem 3.6 *(i) Every weak* convergent sequence* $(u_n) \subset U = X^*$ *is bounded.*

(ii) If u_n *converges to* u *in* $w*$*-sense, then* $\|u\| \leq \liminf \|u_n\|$.

Proof (i) follows directly from Banach-Steinhaus. For part (ii), for any $\delta > 0$ take $x \in X$, $\|x\| = 1$ with $|u(x)| \geq \|u\| - \delta$ (by the def. of the norm). Since $u_n(x) \to u(x)$, we have

$$\|u\| - \delta \leq |u(x)| \leq \liminf |u_n(x)| \leq \liminf \|u_n\|$$

and this is true for any $\delta > 0$. $\square$

Definition 3.7 *Thet set* $K \subset U = X^*$ *is weak* sequentially compact (w*sc) if from any sequence* $\{u_n\} \subset K$ *one can select a weak*-convergent subsequence with limit in* K.

3.3 Banach-Alaoglu in duals of separable spaces

We have the following, easiest version of Banach-Alaoglu.

Theorem 3.8 (Helly's theorem: Version II of Banach-Alaoglu) . *Let* X *be a separable Banach space, let* $U = X^*$. *Then the closed unit ball* B *of* U *is w*sc.*

(Note the difference with version I: here we do not assume reflexivity but assume separability. In particular this version works for $U = L^\infty(M,\mu) = (L^1(M,\mu))^*$ as well, while the previous one did not.)

Proof. Let $\{u_n\} \subset U$, $\|u_n\| \le 1$. Let x_k be a dense subset of X. By the Cantor procedure, there is a subsequence, u_{n_k} such that $u_{n_m}(x_k)$ converges for every fixed k as $m \to \infty$. Call

$$v(x_k) := \lim_{m\to\infty} u_{n_m}(x_k)$$

note that

$$|v(x_k)| = \lim_m |u_{n_m}(x_k)| \le \liminf_m \|u_{n_m}\|\|x_k\| \le \|x_k\|$$

i.e. v is a bounded (linear) map on a dense subset of X (namely on the linear span of x_k's). By the bounded extension principle, v extends to all X and

$$\lim_m u_{n_m}(x) = v(x)$$

holds for all $x \in X$. This is the weak* limit map. $\square$

Theorem 3.9 (Banach-Alaoglu theorem: the original version) *Let X be a Banach space. Then the unit ball B of X^* is compact in the $w*$ topology.*

Note that the statement is compactness and not w*sc. If X is separarable, then it is known that B with the w* topology is metrizable (interestingly enough, not the whole space X^*), hence compactness implies sequential compactness (Theorem 1.4). But this we already knew in Version II.

Sketch of the proof. Note that

$$B \subset Y := \times_{x\in X}\{\lambda \in \mathbf{C} \;:\; |\lambda| \le \|x\|\}$$

because every element of B is an assignment b of a number $b(x) \in \mathbf{C}$ to every $x \in X$ satisfying $|b(x)| \le \|x\|$. So B canonically sits in the direct product (indexed by $x \in X$) of closed disks with radii $\|x\|$. Of course B is much smaller than the direct product Y, since elements of B must also be linear maps.

What is the topology induced on B by the topology of Y? This is exactly the weakest topology on B that makes all maps $b \to b(x)$ (for every $x \in X$) continuous, i.e. the w* topology. (Here one needs to think it over that in the subset B there is no even weaker topology to do the job than the one inherited from Y).

From Tychonov we know that Y is compact. To conclude that B is compact, we need to show that B is closed (in the product topology), because the closed subset of a compact set is compact. I.e. we need to show that the linearity (the property specifying B within Y) survives the limit. If we knew that it is sufficient to check closedness by sequences, then we would only need to argue that if $\ell_n \in B$ is a sequence converging to $\ell \in Y$, then $\ell \in B$, i.e. if all ℓ_n are linear than so is the limit. But this is obvious, for any $x, y \in X$ and scalars $\alpha, \beta \in \mathbf{C}$

$$\ell(\alpha x + \beta y) = \lim_n \ell_n(\alpha x + \beta y) = \lim_n \alpha \ell_n(x) + \beta \ell_n(y) = \alpha \ell(x) + \beta \ell(y)$$

The only problem is that if the space too big, then closedness cannot be checked by showing that limits of sequences lie in the set. To overcome this problem, one introduces the concept of nets, i.e. generalized sequences, where the label set is not the natural numbers, but an arbitrary ordered set. We skipped to introduce them, so we leave to proof as it is, if you are interested, you can learn nets from Reed-Simon IV.2.