Functional Analysis

Ahmed Shaaban, ams2551

December 2020

Introduction

Hausdorfness, and compactness are some of the most useful properties a topological space can possess. Compactness in particular generalizes the notion of 'finiteness' of a space, and makes it easier to carry out certain constructions, and argue about the properties of the space. For example, the compactness of the unit sphere in a finite dimensional normed space is key in proving that all norms are equivalent.

However, not every topology gives us a compact space, and in fact, as we will prove, for an infinite dimensional vector spaces it is impossible to find a topology where the the space is Hausdorff, and locally compact.

Nevertheless, we shall construct a topology on certain infinite dimensional spaces, called the weak* topology where important subspaces like the unit ball are compact. This is the statement of the Banach-Alaoglu theorem.

Constructing the weak* topology will require us to consider the dual of a vector space which is the set of all continuous linear functionals on a space. However, before we work with a topology on the dual space we need to know that the dual space is non-trivial. To show this we will state, and prove the Hahn-Banach Theorem which, aside from being a fundamental theorem in functional analysis, will enable us to argue the existence of non-zero linear functional for any non-trivial topological vector space.

We will also construct another topology on a space called the weak topology, which will be useful in problems related to convexity. In particular, we will prove the Krein-Milman theorem, which in essence states that a convex set can be recovered using its extreme points. Weak topologies are useful in problems related to convexity because they enable us argue that a space is locally convex, subject to some minor conditions. The Krein-Milman theorem combined with the Banach-Alaolgu theorem are useful tools in proving the celebrated Stone-Weierstrass theorem.

1 Vector Spaces

We begin our exposition by introducing topological vector spaces. We will also prove some basic lemmas along the way.

Definition 1.1. A vector space X over a field $\mathbb{F}$ is a set consisting of elements called vectors that satisfy the following properties:

1. The elements of X are an abelian group under the addition operation $(+)$.
2. Each pair (α, x) $a \in \mathbb{F}, x \in X$, is associated with a vector αx such that:
   (a) $1x = x$.
   (b) $\alpha(\beta x) = (\alpha\beta)x$.
   (c) $\alpha(x_1 + x_2) = \alpha x_1 + \alpha x_2$, $\forall \alpha \in \mathbb{F}, x_1, x_2 \in X$.
   (d) $(\alpha + \beta)x = \alpha x + \beta x$.

Definition 1.2. A basis for a vector space is a collection of vectors $\{u_1, \cdots, u_n\}$ such that every x in X can be expressed uniquely as

$$x = \alpha_1 u_1 + \cdots + \alpha_n u_n,\ \alpha_i \in \mathbb{F}.$$

Definition 1.3. A topological vector space is a vector space endowed with a topology $\mathcal{T}$ such that :

1. The addition operation on vectors is continuous.
2. The operation of multiplying a vector by a scalar is continuous.
3. Every singleton is closed.

Lemma 1.1. The operations $T_a(x) = a + x$, and $M_\lambda(x) = \lambda x$, $\lambda \neq 0$ are homeomorphisms:

proof: It is clear that both T and M are bijective. We prove continuity.

Let $(+)$ be the addition operation. If U is open in X, then

$$(+)^{-1}(U) = \bigcup_{x \in x} (x, U - x)$$

is open because X is a topological vector space.

By restricting $+$ to (a, X), we see that $T_a = (+)_{|(a,X)} \circ i$, where $i(a) = (a, X)$. So T_a is the composition of continuous functions and therefore continuous.

We need to show that $T_a^{-1} = T_{-a}$ is also continuous, but this just follows by the same logic as above.

The continuity of M_λ is proven similarly.

This lemma shows us that we need to only consider neighbourhoods at the point 0, since any open set is homeomorphic to an open set containing 0 using translation.

Definition 1.4. A neighbourhood of a point x in a topological vector space is an open set containing x.

Definition 1.5. A local base $\mathcal{B}_x$ around a point x is a collection of neighbourhoods of x such that every neighbourhood of x contains an element of $\mathcal{B}_x$.

Definition 1.6. A map, f, from a vector spaces X to Y over the same field is said to be linear if:

$$f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2)$$

Definition 1.7. A linear functional Λ is a map from the vector space X to the field of scalars such that:

1. $\Lambda(x_1 + x_2) = x_1 + x_2$.
2. $\Lambda(cx) = c\Lambda(x)$.

2 Norms and Compactness

For the remainder of this paper every vector space is over the reals or complex numbers. We will still use the symbol $\mathbb{F}$ to represent the field of complex or real numbers.

Definition 2.1. A norm on a vector space is a map $N : X \to \mathbb{R}$ such that:

1. $N(x) = 0 \iff x = 0$.
2. $N(cx) = |c|N(x)$.
3. $N(x_1 + x_2) \leq N(x_1) + N(x_2)$.

Using the norm we can define a metric on a vector space given by $|x_1 - x_2| = N(x_1 - x_2)$.

For a finite dimensional vector space we can prove that the unit ball is compact by showing it is bounded, and closed. This, however, does not extend to an infinite dimensional vector space. In an infinite dimensional we are no longer able to achieve bounds on the size of the unit ball, and as we shall soon see the norm topology is too 'big'.

The compactness of the unit ball in finite dimensions is a useful result because it enables us to conclude that all norms on a finite dimensional vector space induce the same topology. With this fact in mind we prove the following lemma.

Lemma 2.1. Every finite dimensional subspace Y in a normed space is closed.

proof: Let $\{u_1, \cdots, u_n\}$ be the basis of Y.

First note that all norms in a finite dimensional vector space give the same topology, and recall that this equivalent to saying that for any two norms N_1, N_2 there exists C_1, C_2 such that

$$C_1 N_1(x) \leq N_2(x) \leq C_2 N_1(x), \forall x. \text{ (1)}$$

To use this fact consider a sequence of points $\{y_i\}_{i\in\mathbb{N}} \in Y$ that converge to a point p, and first note that if $y_i \to p$, then we have that $\|y_i\| \to \|p\|$.

Now let us consider L_1 norm which for a vector $y = \alpha_1 u_1 + \cdots + \alpha_n u_n$ is defined as:

$$\|y\|_1 = \sum_{j=1}^{n} |\alpha_j|.$$

Now using (1) we have that $\|y_i\|_1 \leq C_k \|y_i\|$ for some constant C_k, and since the sequence $\{y_i\}_{i\in\mathbb{N}}$ converges we have that $\|y_i\| \leq C$, and so $\|y_i\|_1 \leq C_i$ for some constant C_i.

Since the sequence $\{\|y_i\|\}_{i\in\mathbb{N}}$ is bounded we can apply the Bolzano-Weierstrass Theorem to select a convergent subsequence of $\{\|y_i\|_1\}_{i\in\mathbb{N}}$. Let this subsequence be $\{\|y_p\|_1\}_{p\in\mathbb{N}}$.
We may write the subsequence $\{\|y_p\|_1\}_{p\in\mathbb{N}}$ as:

$$\{\sum_{j=1}^{n} |\alpha_{jp}|\}_{p\in\mathbb{N}}.$$

Here is where we see why the choice of L_1 norm helps. Since the subsequence $\{\|y_p\|\}_{p\in\mathbb{N}}$ is convergent we have that the sequence $\{\alpha_{jp}\}_{p\in\mathbb{N}}$ converges to some constant which we will denote by α_j. This is true for all $j \in [n]$.

Finally note that our sequence, and subsequence converge to the same point, and so we have that the sequence $\{y_i\} = \{\sum_{j=1}^{n} \alpha_{ji} y_j\}$ converges to

$$\sum_{j=1}^{n} \alpha_j u_j \in Y.$$

So we have shown that every convergent sequence of points in Y converges to a point in Y. This implies that Y is closed. ■

Definition 2.2. (Sequential Compactness) A space X is said to be sequentially compact if every sequence contains a convergent subsequence.

See Munkres Theorem 28.2 for a proof that compactness and sequential compactness are equivalent notions on a metric space X.

Theorem 2.1. The unit ball in an infinite dimensional normed space is not compact.

proof: To prove that the unit ball is not compact, we will construct a sequence of unit vectors $\{z_i\}_{i\in\mathbb{N}}$ such that $\|z_i - z_j\| \geq \frac{1}{2}$. For such a sequence, it is clear that no subsequence in $\{z_i\}_{i\in\mathbb{N}}$ can converge in the Cauchy sense, and so no subsequence can converge in the normal sense, and so we have that the unit ball is not sequentially compact.

Now to construct this sequence we first prove the following lemma.

Lemma 2.2. If $Y \neq X$ is a closed subspace, then there exists $z \in X$ such that: $\forall \epsilon > 0$ $\|z\| = 1$, and $\|z - y\| > 1 - \epsilon$, $\forall y \in Y$.

proof: Choose a vector w not in Y. Define the smallest distance, d, from w to y by:

$$d = \inf\{\|w - y\| : y \in Y\}$$

We have that $d > 0$, or else w would be a limit point of Y, which contradicts the fact the Y is a closed set.

Let y_0 be a vector in Y such that $\|w - y_0\| = (1 + \delta)d$, and consider the vector $z = \frac{w - y_0}{\|w - y_0\|}$. We have that for any y in Y:

$$\|z - y\| = \left\| \frac{w - y_0}{\|w - y_0\|} - y \right\| = \frac{1}{\|w - y_0\|} \|w - y_0 - y(\|w - y_0\|)\|$$

$$= \frac{1}{\|w - y_0\|} \|w - y_1\|.$$

Where $y_1 \in Y$ (since Y is a vector space), but $\|w - y_1\| \geq d$, so $\|z - y\| \geq \frac{1}{1+\delta}$, and by taking δ of appropriate size we get the required claim.

We are now ready to construct our wanted sequence one vector at a time by induction.
Base case: Let $w \in X$, and take $z_1 = \frac{w}{\|w\|}$.

Inductive step: Let $z_1, \cdots, z_n$ be a collection of linearly independent vectors that satisfy $\|z_i - z_j\| \geq \frac{1}{2}$ for $i \neq j$. Now consider the set $Y_n = \text{span}\{z_1, \cdots, z_n\}$.

We have shown that every finite subspace in a normed vector space is closed, so Y_n is also closed. Now we can use Lemma 2.2 to select a unit vector $z_{n+1} \notin Y_n$ such that $\|z_n - z_i\| \geq \frac{1}{2}$. We have constructed the sequence that we needed so we see that the unit ball in a infinite dimensional vector space is not compact. ■

3 Local Compactness and Finite Dimensions

We will later construct other topologies where the unit ball is compact, but we now show that it is futile to try and look for a topology on an infinite dimensional space that makes the space locally compact, and Hausdorff.

Local compactness will also serve as a useful in distinguishing between finite, and infinite dimensional vector spaces.

Definition 3.1. A space is said to be locally compact if every point has a neighbourhood whose closure is compact.

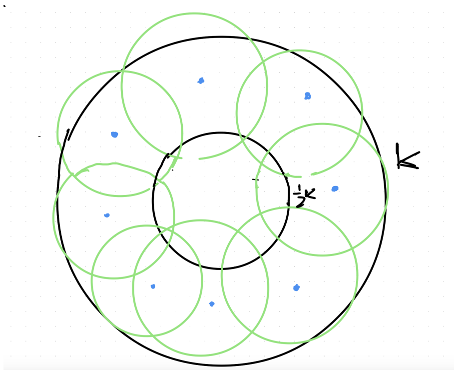
Definition 3.2. A subset A is said to be absorbing if $\forall x \in V$ there exists a scalar t which is a function of x such that $x \in tA$. Furthermore, A is called balanced if $cA \subset A$ for $|c| \leq 1$.

Lemma 3.1. Every neighbourhood of 0 is absorbing.

proof: Let V be a neighbourhood of 0, and consider $f^{-1}(V) = \bigcup A_i \times B_j$, where f represents scalar multiplication, A_i is open in $\mathbb{F}$, and B_j is open in X. For any vector $x \in X$, we have that $0x = 0$, so by continuity of scalar multiplication, there exists some open sets A_x, B_x containing $(0, x)$ such that $A_x \times B_x \in f^{-1}(V)$. So there is some scalar k such that $kx \in V \implies x \in \frac{1}{k}V$.

Theorem 3.1. Every locally compact Hausdorff vector space is finite dimensional.

proof: Let X be a locally compact Hausdorff space, since it is locally compact there exists a neighbourhood, K of 0 whose closure is compact. We have that $\frac{1}{2}K$ is also a neighbourhood of 0 (since scaling is continuous). Now we can cover K using the translates of $\frac{1}{2}K$. In the drawing below, we only need finitely many translates of $\frac{1}{2}K$, shown below in green, to cover K.



Since K is is contained in a compact set, we need only finitely many points $x_1, \cdots, x_n$ by which we need to translate $\frac{1}{2}K$ so that it covers K. So we have that:

$$K \subset \bigcup_{i=1}^{n} \frac{1}{2}K + x_i.$$

Now let $W = \text{span}\{x_1, \cdots, x_n\}$, we have that $K \subset \frac{1}{2}K + W$.
We can repeat these steps, but instead of scaling K by $\frac{1}{2}$, we scale K by 2^{-n}. So we get that

$$K \subset 2^{-n}K + W.$$

By Lemma 2.1, we know that every neighbourhood U of 0 is absorbing, so for sufficiently large n, we have that $2^{-n}K \subset U$. So $K \subset W + U$. This tells us that K is in the closure of W (Rudin's Functional analysis Theorem 1.10).

Recall Lemma 2.1, that says every finite dimensional vector space is closed, applying it here we get that $K \subset W$.
Finally, we have that K is an absorbing set, so for every vector x in X we have that $x \in CK$ for

sufficiently large C. But $K \subset W$, so $x \in CW = W$, where the final equality is because W is a subspace. ■

4 Dual of a Vector Space

Definition 4.1. The dual of a topological vector space X is the vector space X^*, which consists of the set of all **continuous** linear functionals on X.
The vector space operations on X^* are defined as follows:

1. Addition: $(\Lambda_1 + \Lambda_2)(x) = \Lambda_1(x) + \Lambda_2(x)$
2. Scalar Multiplication: $(\alpha f_1)(x) = \alpha f_1(x)$

We prove that $(\Lambda_1 + \Lambda_2)(x)$ is indeed a continuous linear functional:
proof:

1. Continuity: if Λ_1, Λ_2 are continuous linear functionals then map $F : X \to \mathbb{F} \times \mathbb{F}$ defined by $F(x) = (\Lambda_1(x), \Lambda_2(x))$ is continuous. We have that $(\Lambda_1 + \Lambda_2)(x) = S \circ F$, where $S : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ is the addition operation. So $\Lambda_1 + \Lambda_2$ is the composition of continuous functions and therefore continuous.
2. Addition: $(\Lambda_1 + \Lambda_2)(x_1 + x_2) = \Lambda_1(x_1 + x_2) + \Lambda_2(x_1 + x_2) = \Lambda_1(x_1) + \Lambda_2(x_1) + \Lambda_1(x_2) + \Lambda_2(x_2) = (\Lambda_1 + \Lambda_2)(x_1) + (\Lambda_1 + \Lambda_2)(x_2)$
3. Scalars: $(\Lambda_1 + \Lambda_2)(\alpha x) = \alpha(\Lambda_1(x) + \Lambda_2(x)) = \alpha((\Lambda_1 + \Lambda_2)(x))$

Proving that $\alpha\Lambda(x)$ is a linear functional is very similar to the above.

5 The Hahn-Banach Theorem

We now state and prove the Hahn-Banach Theorem, which is a fundamental result in functional analysis. We will then see how the Hahn-Banach will allow us to argue that the dual of a non-empty vector space non-trivial.

Definition 5.1. A map $p : X \to \mathbb{R}$ is positive homogeneous, and subadditive if:

1. $p(\alpha x) = \alpha p(x)\ \forall x \in X$, $\alpha > 0$. (positive homogeneous)
2. $p(x + y) \leq p(x) + p(y), \forall x, y \in X$. (subadditive)

We call a real valued function p a seminorm if instead of 1. we have that $p(\alpha x) \leq |\alpha|x, \forall \alpha \in \mathbb{F}$.

Theorem 5.1. The Hahn-Banach Theorem:
Let Y be a subspace of the vector space X, if $f : Y \to \mathbb{R}$ is a linear map, and if there exists a positive homogeneous, and subadditive map $p : X \to R$, such that $f(y) \leq p(y)\ \forall y \in Y$, then we have that there exists a linear extension $F : X \to \mathbb{R}$ such that:

1. F is linear
2. $F_{|Y}(x) = f(x)$

3. $F(x) \leq p(x)$, $\forall x$.

proof:

Part I: We begin the extension process of f as follows:
Assume Y is a proper subset of X, and pick an element $z \notin Y$. Define Y_z to be:

$$Y_z = \{y + \alpha z : y \in Y, \alpha \in \mathbb{R}\}.$$

We note that elements in this set are uniquely represented, because $y_1 + a_1 z = y_2 + \alpha_2 z \implies y_1 - y_2 = (\alpha_1 - \alpha_2)z$, but $y_1 - y_2 \in Y$, so $\alpha_1 - \alpha_2 = 0$, since $z \notin Y$.

If we wish to extend f to X, then a good start would be to extend f to Y_z so that $F(y + \alpha z) = F(y) + \alpha F(z) = f(y) + \alpha F(z)$.

Now note that if we wish the inequality $F(y+\alpha z) \leq p(y+\alpha z)$ to hold for $\alpha > 0$ then we have that:

$$f(y) + \alpha F(z) \leq p(y + \alpha z) \iff \frac{p(y + \alpha z)}{\alpha} - f\left(\frac{1}{\alpha} y\right) \geq F(z) \text{ ,} \forall \alpha > 0.$$

The above inequality is equivalent to $p(y + z) - f(y) \geq F(z)$ $\forall y$.
So:

$$F(z) \leq \inf\{p(y + z) - f(y)\} \iff f(y) + \alpha F(z) \leq p(y + \alpha z) \ \forall \alpha > 0.$$

Similarly for $\alpha < 0$ we have that:

$$F(z) \geq \sup\{f(y) - p(y - z)\} \iff \alpha F(z) \leq p(y + \alpha z) \ \forall \alpha < 0.$$

Lemma: $\sup\{f(y) - p(y - z)\} \leq \inf\{p(y + z) - f(y)\}$.
To prove this lemma it suffices to show that:

$$f(y_1) - p(y_1 - z) \leq p(y_2 + z) - f(y_2) \iff f(y_1) + f(y_2) \leq p(y_2 + z) + p(y_1 - z)$$

We have that $f(y_1) + f(y_2) = f(y_1 + y_2) \leq p(y_1 + y_2) = p((y_1 - z) + (y_2 + z))$ (assumption)
We also have that $p((y_1 - z) + (y_2 + z)) \leq p(y_1 - z) + p(y_2 + z)$ (subaddivity)
So the lemma is true.

We can choose $F(z)$ such that it lies between the supremum and infimum we defined.

Part II
Now that we have managed to extend f to the space Y_z while fulfilling the conditions of linearity and being bounded by p. We now extend f to the entire space X.

Define the following set Ω.

$$\Omega = \{(Z, F_Z)\}$$

Where $Y \subset Z$ is a subspace, and F_Z meets the following conditions:

(i) $F_Z : Z \to \mathbb{R}$ is a linear functional on Z.

(ii) $F_z(y) = f(y)$,$\forall y \in Y$.

(iii) $F_Z(z) \leq p(z)\ \forall z \in Z$.

Note that the extension to Y_z we constructed from **Part I** satisfied all of the above conditions.

We can impose a partial order the elements of Ω by $(Z_1, F_{Z_1}) \leq (Z_2, F_{Z_2})$, if $Z_1 \subset Z_2$, and F_{Z_2} is an extension of F_{Z_1}.

Now to apply Zorn's lemma we need to confirm that every totally ordered subcollection

$$\Omega' = \{(Z_I, F_{Z_I})\}$$

is bounded, but this just follows from considering the union of all $Z_I \in \Omega'$.
We can then apply to to Zorn's to find a maximal element.

Once we have found the maximal element (Z, F_Z) what remains to be shown is that $Z = X$, but this follows from the first part since if they are not equal we can apply the extension process of **Part I** to get a bigger pair $(Z', F_{Z'})$. This proves that $Z = X$, and that F_Z is the linear functional fulfilling the conditions of our hypothesis.$\blacksquare$

Corollary 5.1. If there exists a linear functional f defined on a subspace M, and a semi-norm on X such that:

$$|f(x)| \leq p(x)\ \forall x \in M.$$

Then we can extend f to the functional Λ on X, while ensuring that $|\Lambda(x)| \leq p(x)$.
This is Theorem 3.3. in Rudin's Functional Analysis.

Lemma 5.2. If X is non-trivial, then dual space X^* is non-trivial.

proof: Take $x \in X$, and define $f : \mathrm{span}(x) \to \mathbb{F}$ by $f(\alpha x) = \alpha f(x)$, and $f(x) = 1$.
It is clear that $|f(\alpha x)| \leq |\alpha|$, and it is clear that $p(\alpha x) = |\alpha|$ is a seminorm. So we can apply the previous theorem to show us that f can be extended to the non-zero function $\Lambda \in X^*$.

Λ is continuous because it is bounded by a seminorm.

6 Weak Topologies

Recall our result about the non-compactness of the unit ball in an infinite dimensional normed space. We were able to obtain this result, by showing that there are enough open sets in the norm topology such that it is possible to construct an open cover of the unit ball, without a finite cover. We thus stand a better chance of creating a topology where the unit ball is compact if there are 'fewer' open sets in our topology.
The next definition makes precise the notion of a 'bigger', and 'smaller' topology.

Definition 6.1. Let $\mathcal{T}_1$, and $\mathcal{T}_2$ be two topologies on a vector space V. If $\mathcal{T}_1 \subset \mathcal{T}_2$, then $\mathcal{T}_1$ is said to be weaker than $\mathcal{T}_2$, or alternatively $\mathcal{T}_2$ is said to be stronger than $\mathcal{T}_1$. It is clear that the discrete topology is the strongest topology on a vector space.

Although we wish to throw away open sets, we do so systematically, and without breaking the continuity of certain maps.

Consider a set of functions $\{f_i\}_{i\in I}$ from a vector space X, to the topological vector spaces $\{Y_i\}_{i\in I}$. We wish to endow X with a topology $\mathcal{T}$ such that each of the f_i maps is continuous. Clearly, the discrete topology is one trivial example that achieves this task, but now if we require that $\mathcal{T}$ be the weakest topology on X we get a much more interesting question.

To find the weakest topology $\mathcal{T}$, we note that $\mathcal{T}$ must contain elements of the form

$$\mathcal{S} = \{f_1^{-1}(U_1) \cap \cdots \cap f_1^{-1}(U_n) : f_i \in \mathcal{F}, U_i \text{ open in } Y_i\}.$$

We then see that the elements of S form a basis, and it is clear that the topology under this basis is the weakest topology such that every $f_i \in \mathcal{F}$ is continuous.
The weakest topology induced by a collection of maps $\mathcal{F}$ is usually called the $\mathcal{F}$ topology on X.

Lemma 6.1. If a collection, $\mathcal{F}$ of maps, from X into Hausdorff spaces separates points on X, then the $\mathcal{F}$ topology is Hausdorff.

proof: Let x_1, x_2 be two distinct points in X. Since our collection $\mathcal{F}$ is separating, we have that there exists a function $f_i : X \to Y_i$, where Y_i is Hausdorff such that $f_i(x_1) \neq f_i(x_2)$. Since Y_i is Hausdorff, there exists opens disjoint sets V_1, V_2 containing $f_i(x_1), f_i(x_2)$ respectively. It follows that $f_i^{-1}(V_1), f_i^{-1}(V_2)$ are disjoint open sets containing x_1, x_2 respectively. These two sets are elements of the $\mathcal{F}$ topology.

Since Hausdorffness is a desirable property we will direct most our attention to spaces where X^* separate points on X.

Definition 6.2. The weakest topology induced by X^* on X is called the weak topology on X.

An important observation to make is that every $x \in X$ induces a linear functional g_x on X^*, defined by:

$$g_x \Lambda = \Lambda(x).$$

To see that every g_x is linear note that:

$$g_x(\Lambda_1 + \Lambda_2) = \Lambda_1(x) + \Lambda_2(x) = g_x(\Lambda_1) + g_x(\Lambda_2).$$

Scalar multiplication is demonstrated similarly. So we see that X can also induce a topology on its dual, called the weak* star topology.

The subbasis that generates the weak* topology is of the form:

$$\{\Lambda : |\Lambda(x) - r| < \epsilon\}.$$

For any arbitrary choice of $x \in X$, $r \in \mathbb{C}$, and $\epsilon \in \mathbb{R}$.,

7 The Banach-Alaoglu Theorem

It is important to note that the product topology is the weakest topology on a product space such that the projection maps are continuous. This just follows from considering the subbasis that generates the product topology.

This idea is powerful because it allows us to use the tools we have for the product topology, and apply them to the weak topology.

Before we begin proving the Banach-Alagolu Theorem, we state without proof the Tychonoff theorem.

Theorem 7.1. (Tychonoff's Theorem) The product of compact spaces is compact.
The proof for the Tychonoff theorem can be found in section 37 of Munkres. Tychonoff's theorem will enable to us to create a compact set under the product topology, and from there will use the close link between product topologies, and weak topologies to prove the Alaoglu theorem.

Theorem 7.2. (Banach-Alaoglu)
If V is a neighbourhood of 0, then we have that the set:

$$K = \{\Lambda \in X^* : |\Lambda(x)| \leq 1,\ \forall x \in V\}$$

is compact in the weak* topology.

proof:
Recall that every neighbourhood V of 0 is absorbing. So for each $x \in X$ there exists a scalar $\gamma(x)$ such that $x \in \gamma(x)V$.

For each x define the set D_x by:

$$D_x = \{a : |a| \leq \gamma(x)\}$$

Now each D_x is bounded, and therefore compact (Heine-Borel) so by Tychonoff's theorem we have that the product

$$P = \prod_{x \in X} D_x$$

is compact.
Observe now that the elements of P can be thought of as functions, f, that satisfy $|f(x)| \leq \gamma(x)$. This is because every $p \in P$ can be viewed as a function f defined by $f(x) = p_x$

Using the above understanding of the elements of P we see that $K \subset X^* \cap P$.

Now note that K is a subset of two different topological spaces, X^*, and P. So as a subset we can endow K with two, potentially, different topologies, but we will now show that these two topologies actually coincide.

Claim: The topological spaces $K \cap X^*$ and $K \cap P$ coincide.

proof: Informally this claim makes sense, since the elements of $K \cap P$ are the functions f where $f(x) \leq 1$, and we know that the product topology P is the weakest topology where the projection maps are continuous, but by design we see that projection maps correspond nicely with the evaluation of a functional at a coordinate, and so it makes sense that these two topologies coincide.

To formally prove this we need to consider the local base around a functional $\Lambda_0 \in X^*$. This local base $\mathcal{B}_w$ consists of finite intersections of sets of the form

$$B_{x,\epsilon} = \{\Lambda : |\Lambda(x) - \Lambda_0(x)| < \epsilon\}$$

To see why this is a local base, first note that the set $\{\Lambda : |\Lambda(x) - \Lambda_0(x)| < \epsilon\}$; constitutes a neighbourhood of Λ_0.

Now if Λ_0 is contained in some neighbourhood of the form

$$U_1 = \{\Lambda : |\Lambda(x) - r| < \epsilon\},$$

then $|\Lambda_0(x) - r| = \alpha < \epsilon$. By the triangle inequality the set

$$U_2 = \{\Lambda : |\Lambda_0(x) - \Lambda(x)| < \epsilon - \alpha\}$$

is an open neighbourhood of Λ_0, which is also a subset of U_1. Taking finite intersections proves that indeed $\mathcal{B}_w$ is a local basis for Λ_0.

Similarly a local base at p_0 for the product topology P consists of finite intersection of sets of the form

$$C_{x,\epsilon} = \{p : |p_x - (p_0)_x| < \epsilon\}.$$

It is not hard to see that $B_{x,\epsilon} \cap K = C_{x,\epsilon} \cap K$.
We've show now that these two topologies on K coincide, what remains to be shown is that K is closed in P. Which would imply that K is compact, since the closed subset of a compact space is compact.

Claim: K is closed in P.
proof: Let f_0 be a limit point of K. We need to show two things:

1. f_0 is linear.
2. $|f_0(x)| \leq 1, \forall x \in V$

To show 1. select the points $x, y, \alpha x + \beta y \in X$, and consider the following neighbourhood of f_0:

$$U = \{f : |f(z) - f_0(z)| < \epsilon\},$$

where z is any of $x, y, \alpha x + \beta y$.

Since f_0 is a limit point of k we can select a linear functional $\Lambda \in K$ that intersects U.

Consider the function $(\Lambda - f_0)$ defined by $(\Lambda - f_0)(x) = \Lambda(x) - f_0(x)$.

It is easy to check, using the linearity of Λ, that

$$(\Lambda - f_0)(\alpha x + \beta y) - (\Lambda - f_0)(\alpha x) - (\Lambda - f_0)(\beta y) = f_0(\alpha x + \beta y) - f_0(\alpha x) - f_0(\beta y).$$

Since $\Lambda \in U$, it follows that $|f_0(\alpha x + \beta y) - f_0(\alpha x) - f_0(\beta y)| < \epsilon|\alpha + \beta + 1|$. We see that the RHS of the inequality is arbitrarily close to 0, and so must be 0. So f_0 is linear. This proves 1.

To show 2. Choose an $x \in V$, and consider the neighbourhood of $\{f : |f_0(x) - f(x)| < \epsilon\}$. Again, since f_0 is a limit point of k pick a functional $\Lambda \in K$. We have that $|f(x) - \Lambda(x)| < \epsilon$, but since $\Lambda \in K$, we have that $|\Lambda(x)| \leq 1$. This means that $f_0(x) \leq 1$, for an appropriate choice of ϵ. This proves 2.

Since P is compact, and K is closed we have that K is compact in P. This implies that K is compact in X^* by the first claim. ■

At a first glance, the Banach-Alaoglu theorem seems to contradict Theorem 3.1, but this is not the case because for an infinite dimensional vector space our set K is not a neighbourhood of 0.

8 Convexity

We now study weak topologies, and their relationship with convex sets, an important class of sets that often appear in problems related to optimization. We will see how weak topologies let us prove results about the extreme points of a convex set. But first we will need to state some definitions.

Definition 8.1. A subset A of a vector space X is said to be convex if:

$$tA + (1-t)A \subset A, \ \forall t \in [0,1].$$

Informally, the above definition just means that the set contains every line segment connecting two points.

Lemma 8.1. The intersection of convex sets is convex.
proof:
If x_1, x_2 are vectors in the intersection of a collection of convex sets $\{E_i\}_{i \in I}$ then we have that $tx_1 + (1-t)x_2 \in E_i, \forall i \in I$. So

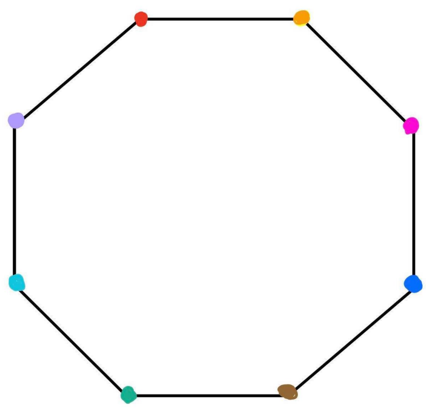
$$tx_1 + (1-t)x_2 \in \bigcap_{i \in I} E_i \implies \bigcap_i E_i \text{ is convex.}$$

Definition 8.2. The convex hull of a set E is the intersection of all convex sets that contain E. We will denote the convex hull of E by co(E).

Definition 8.3. Let K be a subset of X. A nonempty set $S \subset K$ is called an extreme subset of K if for $x, y \in X$, $t \in (0,1)$

$$(1-t)x + ty \in S \implies x \in S, y \in S.$$

An extreme point of is an extreme set of X with only one point. We will denote the set of all extreme points of a set K by $E(K)$.
In some sense we may view the extreme points of a subset K as the 'vertices' of the set K. The picture on the next page shows a convex shape with its extreme points in colour.



Definition 8.4. A space X is said to be locally convex if there exists a local base around 0 whose members are convex.

Theorem 8.1. If X is a locally convex topological vector space, and if A, B convex, disjoint, and compact subsets of X, then there exists a continuous linear functional Λ such that:

$$\sup_{x \in A} \mathrm{Re}(\Lambda(x)) < \inf_{x \in B} \mathrm{Re}(\Lambda(x))$$

This a corollary of the the Hahn-Banach separation theorem, we will omit the proof here. For a proof see reference number 7.

The usefulness of the weak topology on X, will be made evident once we prove that if X^* separates points on X then the weak topology on X makes X a locally convex space. This will enable us to apply the previous theorem, and reach our goal of proving the Krein-Milman theorem.

Lemma 8.2. If X^* separates points on X, then there exists a continous linear function $\Lambda \in X^*$ such that

$$\sup_{x \in A} \mathrm{Re}(\Lambda(x)) < \inf_{x \in B} \mathrm{Re}(\Lambda(x)).$$

proof:
Let X_w be X endowed with the weak topology. We first show that X_w is locally convex.

The local base around 0, for X_w consists of the elements which are the finite intersection of sets of the form:

$$\{x : |\Lambda(x)| < \epsilon\}.$$

To prove local convexity, pick an element $B \in \mathcal{B}$. If $x_1, x_2 \in B$, then we have that:

$$\Lambda(tx_1 + (1-t)x_2) = t\Lambda(x_1) + (1-t)\Lambda(x_2) < \epsilon.$$

This proves that $tx_1 + (1-t)x_2 \in B$.

Now if A, B are both compact in X they must also be compact in X_w, since X_w is weaker than X. We are then able to apply Theorem on X_w to conclude that there exists a linear functional Λ that satisfies $\sup_{x \in A}\mathrm{Re}(\Lambda(x)) < \inf_{x \in B}\mathrm{Re}(\Lambda(x))$. Since every continuous functional on X_w is also continuous on X, Λ satisfies the conditions of the hypothesis.

Theorem 8.3. Every compact extreme set of K contains an extreme point of K.

proof: Let $\mathcal{P}$ be the collection of all compact extreme sets of K.
Choose some element S in $\mathcal{P}$, and let $\mathcal{P}_S$ be set the set of all elements in $\mathcal{P}$ that are subsets of S.

Impose a partial order on the elements of $\mathcal{P}_S$ by inclusion, and note that by Hausdorff's Maximal Principle, there must exist a maximally ordered subcollection. We will denote this maximally ordered subcollection by Ω'. Now let M be the intersection of all elements in Ω'.

Note that M is a compact extreme set of K. It is compact because X is a Hausdorff space, and so every compact set is closed. Since M is the intersection of closed sets, it is closed in any of the elements of Ω' which implies that M is compact. It is easy to verify that M is an extreme set. Finally, note that M is not empty, this follows from a property of compact sets called the finite intersection property.

Lemma 8.3. Every $\Lambda \in X^*$ is constant on M.
proof: Consider the set

$$S_\Lambda = \{x \in S : \mathrm{Re}(\Lambda(x)) = \sup_{a \in S}\mathrm{Re}(\Lambda(a))\}$$

Let $\sup_{a \in S}\mathrm{Re}(\Lambda(a)) = \mu$.
S_Λ an extreme compact set of K. To see why take $x, y \in K$ such that

$$tx + (1-t)y \in S_\Lambda$$

We have $tx + (1-t)y \in S$, and since S is an extreme set of K, we have that $x \in S$, and $y \in S$. This implies that $\mathrm{Re}(\Lambda(x)), \mathrm{Re}(\Lambda(y)) \leq \mu$.

Finally, note that $\Lambda(tx + (1-t)y) = t\Lambda(x) + (1-t)\Lambda(y)$, so

$$\mathrm{Re}(\Lambda(tx + (1-t)y) \leq t\mu + (1-t)\mu \leq \mu.$$

So $\mathrm{Re}(\Lambda(x)), \mathrm{Re}(\Lambda(x)) = \mu \implies x, y \in S_\Lambda$.

Since S_Λ is an element of $\mathcal{P}$, and $S_\Lambda \subset S$, we have that $M_\Lambda \subset M$, so $\Lambda \in X^*$ is constant on M. Since X^* separates points it must be the case that M contains only one point, and is therefore an extreme point of K.

Since M depends on the choice of $S \in \mathcal{P}$, we conclude that every compact extreme set of K contains an extreme point of K.

Theorem 8.4. (Krein-Milman) Suppose X is a topological space on which X^* separates points. If K is a compact, convex nonempty set in X, then K is the closure of the convex hull of its extreme points. Or more succinctly $K = \overline{\text{co}(E(K))}$.

First note that K is convex and compact. Convexity, and compactness imply that K is a closed convex set, so it must contain $\overline{\text{co}(E(k))}$.

To show the reverse inclusion take a point $x_0 \in K$ that is not in $\overline{\text{co}(E(k))}$. Applying Lemma 8.2. we find a $\Lambda \in X^*$ such that $\text{Re}(\Lambda(x)) < \Lambda(x_0)$ for all $x \in \overline{\text{co}(E(k))}$. Consider the set

$$K_\Lambda = \{x \in K : \text{Re}(\Lambda(x)) = \text{sup}_{a \in K} \text{Re}(\Lambda(a))\}$$

which by Lemma 8.3. is a compact extreme set of K. Then by Theorem 8.2. it must contain an extreme point of K, but note that $K_\Lambda \cap \overline{\text{co}(E(k))} = \emptyset$. A contradiction. ■

9 References

These are the resources I used for every section of the paper.

1. Walter Rudin, Functional Analysis. Sections 1, 2, 3, 4, 6, 7, 8.

2. John B. Conway, Functional Analysis. Sections 7, 8.

3. IMPA Lectures on Functional Analysis. Sections 2, 5.

4. Terry Tao's blog on local compactness and vector spaces.

5. Keith Conrad's notes on Topological Vector Spaces. Section 1.

6. Brezis Haim, Functional Analysis, Sobolev Spaces and Partial Differential Equations. Section 7.

7. TUM Notes on Hahn-Banach Separation theorem: https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MGMQM2012/Hahn-Banach.pdf