MAT4010: Functional Analysis

Tutorial 4: The Hahn-Banach Theorem*

Mou, Minghao

The Chinese University of Hong Kong, Shenzhen

February 22, 2023

1 Exercises

Problem 1.1 (Banach Limits)**.** *Let ℓ^∞ be the Banach space of bounded sequences of real numbers with the supremum norm and define the shift operator $T : \ell^\infty \to \ell^\infty$ by*

$$Tx := (x_{n+1})_{n\in\mathbb{N}}; \quad \forall x = (x_n)_{n\in\mathbb{N}} \in \ell^\infty.$$

Consider the subspace

$$Y := im(id - T) = \{x - Tx | x \in \ell^\infty\}.$$

Prove the following.
(a). *The subspace $c_0 \subset \ell^\infty$ is contained in the closure of Y.*
(b). *Let $\mathbf{1} = (1, 1, 1, ...) \in \ell^\infty$ be the constant sequence with entries 1. Prove that $\sup_{n\in\mathbb{N}} |1 + x_{n+1} - x_n| \geq 1$ for all $x \in \ell^\infty$ and deduce that*

$$d(\mathbf{1}, Y) = \inf_{y\in Y} \|\mathbf{1} - y\|_\infty = 1.$$

(c). *By the Hahn-Banach Theorem, there exists a bounded linear functional $\Lambda : \ell^\infty \to \mathbb{R}$ such that*

$$\Lambda(\mathbf{1}) = 1, \quad \|\Lambda\| = 1, \quad \Lambda(x - Tx) = 0, \forall x \in \ell^\infty.$$

Prove that any such functional has the following properties
(i). $\Lambda(Tx) = \Lambda(x), \forall x \in \ell^\infty$.
(ii). $\liminf_{n\to\infty} x_n \leq \Lambda(x) \leq \limsup_{n\to\infty} x_n$ *for all $x \in \ell^\infty$.*
(iii). *If $x \in \ell^\infty$ satisfies $x_n \geq 0$ for all $n \in \mathbb{N}$, then $\Lambda(x) \geq 0$.*
(iv). *If $x \in \ell^\infty$ converges, then $\Lambda(x) = \lim_{n\to\infty} x_n$.*
(d). *Let Λ be as in (c). Find $x, y \in \ell^\infty$ such that $\Lambda(xy) \neq \Lambda(x)\Lambda(y)$.*
(e). *Let Λ be as in (c). Prove that there does not exist a sequence $y \in \ell^1$ such that $\Lambda(x) = \sum_{n=1}^\infty x_n y_n$ for all $x \in \ell^\infty$.*

*The presentation of this note is based on [BS18] and [Kre91].

Problem 1.2 (Schatten's Projective Tensor Product)**.** *Let X and Y be real normed vector spaces.*
(a). *For every normed vector space Z, the space $\mathcal{B}(X,Y;Z)$ of bounded bilinear maps $B : X \times Y \to Z$ is a normed vector space with the norm*

$$\|B\| := \sup_{x \in X \setminus \{0\}, Y \setminus \{0\}} \frac{\|B(x,y)\|_Z}{\|x\|_X \|y\|_Y}, \forall B \in \mathcal{B}(X,Y;Z).$$

(b). *The map*

$$\mathcal{B}(X,Y;Z) \to \mathcal{L}(X, \mathcal{L}(Y,Z)) : B \mapsto (x \mapsto B(x,\cdot))$$

is an isometric isomorphism.
(c). *Associated to each pair $(x,y) \in X \times Y$ is a linear functional*

$$x \otimes y \in \mathcal{B}(X,Y,\mathbb{R})^*$$

defined by $\langle x \otimes y, B \rangle := B(x,y)$ for all $B \in \mathcal{B}(X,Y;\mathbb{R}$. It satisfies

$$\|x \otimes y\| = \|x\|_X \|y\|_Y.$$

(d). *Let $X \otimes Y \subset \mathcal{B}(X,Y;\mathbb{R})^*$ be the smallest closed subspace containing the image of the bilinear map $X \times Y \to \mathcal{B}(X,Y;\mathbb{R})^* : (x,y) \mapsto x \otimes y$ in (c). Then, for every normed vector space Z, the map*

$$\mathcal{L}(X \otimes Y, Z) \to \mathcal{B}(X,Y;Z) : A \mapsto B_A$$

defined by $B_A(x,y) := A(x \otimes y)$ for all $x, y \in X$ and $A \in \mathcal{L}(X \otimes Y, Z)$ is an isometric isomorphism.

Problem 1.3. *This exercise shows that the hypothesis that one of the convex sets has nonempty interior cannot be removed in theorem 3.3. Consider the Hilbert space $H = \ell^2$ and define*

$$A := \left\{ x \in \ell^2 \middle| \exists n \in \mathbb{N}, \forall i \in \mathbb{N}, i < n \Rightarrow x_i > 0; i \geq n \Rightarrow x_i = 0 \right\}$$
$$B := \left\{ x \in \ell^2 \middle| \exists n \in \mathbb{N}, \forall i \in \mathbb{N}, i < n \Rightarrow x_i = 0; i \geq n \Rightarrow x_i > 0 \right\}.$$

Show that A and B are nonempty disjoint convex subsets of ℓ^2 with empty interior whose closures agree. If $\Lambda : \ell^2 \to \mathbb{R}$ is a bounded linear functional and c is a real number such that $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) \leq c$ for all $x \in B$, show that $\Lambda = 0$ and $c = 0$.

Problem 1.4. *Show that, for a real separable normed vector space X, we can prove the Hahn-Banach Theorem (cf. theorem 2.1) without using Zorn's lemma.*

Problem 1.5. *Show that $H \subset X$ is an affine hyperplane if and only if there exist a nonzero bounded linear functional $\Lambda : X \to \mathbb{R}$ and a real number $c \in \mathbb{R}$ such that $H = \Lambda^{-1}(c)$.*

Problem 1.6. *Use corollary 2.3 to show that: Let X be a real normed vector space and let $x_0 \in X$ be a nonzero vector. Then there exists a bounded linear functional $x^* \in X^*$ such that*

$$\|x^*\| = 1, \quad \langle x^*, x_0 \rangle = \|x_0\|.$$

Problem 1.7. *If x_0 in a normed space X is such that $|\langle x^*, x_0 \rangle| \leq c$ for all $x^* \in X^*$ of norm 1, show that $\|x_0\| \leq c$.*

Problem 1.8. *Prove 3.1 for topological vector spaces.*

2 The Hahn-Banach Theorem

The Hahn-Banach Theorem deals with bounded linear functionals on a subspace of a Banach space X and asserts that every such functional extends to a bounded linear functional on all of X. This theorem continues to hold in the more general setting where X is any real vector space and boundedness is replaced by a bound relative to a given quasi-seminorm on X.

Definition 2.1 (Quasi-Seminorm)**.** *Let X be a real vector space. A function $p : X \to \mathbb{R}$ is called a **quasi-seminorm** if it satisfies*

$$p(x+y) \leq p(x) + p(y), \quad p(\lambda x) = \lambda p(x)$$

*for all $x, y \in X$ and all $\lambda \geq 0$. It is called a **seminorm** if it is a quasi-seminorm and $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$ and all $\lambda \in \mathbb{R}$. A seminorm has nonnegative values, because $2p(x) = p(x) + p(-x) \geq p(0) = 0$ for all $x \in X$. Thus a seminorm satisfies all the axioms of a norm except the nondegeneracy.*

Theorem 2.1 (Hahn-Banach)**.** *Let X be a normed vector space and let $p : X \to \mathbb{R}$ be a quasi-seminorm. Let $Y \subset X$ be a linear subspace and let $\phi : Y \to \mathbb{R}$ be a linear functional such that $\phi(x) \leq p(x)$ for all $y \in Y$. Then there exists a linear functional $\Phi : X \to \mathbb{R}$ such that*

$$\Phi|_Y = \phi, \quad \Phi(x) \leq p(x), \quad \forall x \in X.$$

Proof. Define the set

$$\mathscr{P} := \left\{ (Z, \psi) \middle| Z \leq X; \psi : Z \to \mathbb{R} \text{ is such that } Y \subset Z, \psi|_Y = \phi, \psi(x) \leq p(x), \forall x \in Z \right\}.$$

This set is partially ordered by the relation

$$(Z, \psi) \leq (Z', \psi') \quad \Leftrightarrow \quad Z \subset Z' \ \& \ \psi'|_Z = \psi.$$

for all $(Z, \psi), (Z', \psi') \in \mathscr{P}$. A chain in $\mathscr{P}$ is a totally ordered subset $\mathscr{C} \subset \mathscr{P}$. Every nonempty chain $\mathscr{C} \subset \mathscr{P}$ has a supremum (Z_0, ψ_0) given by

$$Z_0 := \bigcup_{(Z,\psi) \in \mathscr{C}} Z, \quad \psi_0(x) := \psi(x), \forall (Z, \psi) \in \mathscr{C}, \forall x \in Z.$$

Hence it follows from Zorn's lemma that $\mathscr{P}$ has a maximal element (Z, ψ). By lemma 2.2 every such maximal element satisfies $Z = X$ (**Why?**) and the proof is complete. $\square$

Lemma 2.2. *Let X, p, y and ϕ be as in theorem 2.1. Let $x_0 \in X \backslash Y$ and define $Y' := Y \oplus \mathbb{R} x_0$. Then there exists a linear functional $\phi' : Y' \to \mathbb{R}$ such that $\phi'|_Y = \phi$ and $\phi'(x) \leq p(x)$ for all $x \in Y'$.*

Proof. An extension $\phi' : Y' \to \mathbb{R}$ of the linear functional $\phi : Y \to \mathbb{R}$ is uniquely determined by its value $a := \phi'(x_0) \in \mathbb{R}$ on x_0. This extension satisfies the required condition $\phi'(x) \leq p(x)$ for all $x \in Y'$ if and only if

$$\phi(y) + \lambda a \leq p(y + \lambda x_0), \forall y \in Y, \forall \lambda \in \mathbb{R}. \tag{1}$$

If this holds, then

$$\phi(y) \pm a \leq p(y \pm x_0), \forall y \in Y. \tag{2}$$

Conversely, if (2) holds for $\lambda > 0$, then

$$\begin{aligned}\phi(y) + \lambda a &= \lambda(\phi(\lambda^{-1}y) + a) \leq \lambda p(\lambda^{-1}y + x_0) = p(y + \lambda x_0),\\ \phi(y) - \lambda a &= \lambda(\phi(\lambda^{-1}y) - a) \leq \lambda p(\lambda^{-1}y - x_0) = p(y - \lambda x_0).\end{aligned}$$

This shows that (1) is equivalent to (2). Thus it remains to find a real number $a \in \mathbb{R}$ that satisfies (2). Equivalently, a must satisfy

$$\phi(y) - p(y - x_0) \leq a \leq p(y + x_0) - \phi(y), \forall y \in Y. \tag{3}$$

To see such a number exists, fix two vectors $y, y' \in Y$. Then

$$\begin{aligned}\phi(y) + \phi'(y) = \phi(y + y') &\leq p(y + y')\\ &= p(y + x_0 + y' - x_0) \leq p(y + x_0) + p(y' - x_0).\end{aligned}$$

Thus

$$\phi(y') - p(y' - x_0) \leq p(y + x_0) - \phi(y), \forall y, y' \in Y$$

□

and this implies

$$\sup_{y' \in Y} \left(\phi(y') - p(y' - x_0)\right) \leq \inf_{y \in Y} \left(p(y + x_0) - \phi(y)\right).$$

Hence there exists a real number $c \in \mathbb{R}$ that satisfies (3) and this proves the lemma.

Corollary 2.3. *Let X be a real normed vector space and $Y \subset X$ be a linear subspace, and let $x_0 \in X \backslash \overline{Y}$. Then*

$$\delta := d(x_0, Y) := \inf_{y \in Y} \|x_0 - y\| > 0$$

and there exists a bounded linear functional $x^* \in Y^\perp$ such that

$$\|x^*\| = 1, \quad \langle x^*, x_0 \rangle = \delta.$$

Proof. We first prove that the number δ is positive. Suppose by contradiction that $\delta = 0$. Then, by the axiom of countable choice, there exists a sequence $(y_n)_{n \in \mathbb{N}} \subset Y$ such that $\|x_0 - y_n\| < 1/n$ for all $n \in \mathbb{N}$. This implies that y_n converges to x_0 and hence $x_0 \in \overline{Y}$, in contradiction to our assumption. This shows that $\delta > 0$ as claimed.

Now define the subspace $Z \subset X$ by

$$Z := Y \oplus \mathbb{R}x_0 := \{y + tx_0 | y \in Y, t \in \mathbb{R}\}$$

and define the linear functional $\psi : Z \to \mathbb{R}$ by

$$\psi(y + tx_0) := \delta t, \forall y \in Y, \forall t \in \mathbb{R}.$$

This functional is well defined because $x_0 \notin Y$. It satisfies $\psi(y) = 0$ for all $y \in Y$ and $\psi(x_0) = \delta$. Moreover, if $y \in Y$ and $t \in \mathbb{R}\backslash\{0\}$, then

$$\frac{|\psi(y + tx_0)|}{\|y + tx_0\|} = \frac{|t|\delta}{\|y + tx_0\|} = \frac{\delta}{\|t^{-1}y + x_0\|} \leq 1.$$

With this understood, it follows from Hahn-Banach Theorem (cf. 2.1) that there is a bounded linear functional $x^* \in X^*$ such that

$$\|x^*\| \leq 1$$

and

$$\langle x^*, x\rangle = \psi(x), \forall x \in Z.$$

The norm of x^* is actually equal to one because

$$\|x^*\| \geq \sup_{y\in Y} \frac{|\psi(x_0 + y)|}{\|x_0 + y\|} = \sup_{y\in Y} \frac{|\delta|}{\|x_0 + y\|} = 1,$$

by definition of δ. Moreover,

$$\langle x^*, x_0\rangle = \psi(x_0) = \delta$$

and

$$\langle x^*, y\rangle = \psi(y) = 0, \forall y \in Y.$$

□

An immediate corollary is

Corollary 2.4. *Let X be a real normed vector space and let $x_0 \in X$ be a nonzero vector. Then there exists a bounded linear functional $x^* \in X^*$ such that*

$$\|x^*\| = 1, \quad \langle x^*, x_0\rangle = \|x_0\|.$$

The next corollary characterizes the closure of a linear subspace and gives rise to a criterion for a linear subspace to be dense.

Corollary 2.5 (Closure of a Subspace)**.** *Let X be a real normed vector space, let $Y \subset X$ be a linear subspace, and let $x \in X$. Then*

$$x \in \overline{Y} \quad \Leftrightarrow \quad \langle x^*, x\rangle = 0, \forall x^* \in Y^\perp.$$

Proof. Please refer to the last exercise of Tutorial 3. □

Corollary 2.6 (Dense Subspaces)**.** *Let X be a real normed vector space and let $Y \subset X$ be a linear subspace. Then Y is dense in X if and only if $Y^\perp = \{0\}$.*

Proof. Use corollary 2.5. Refer to the last exercise of Tutorial 3. □

3 Separation of Convex Sets

One application of the Hahn-Banach Theorem concerns a pair of disjoint convex sets in a normed vector space. They can be separated by a hyperplane whenever one of them has nonempty interior. The result and its proof carry over to any general topological vector spaces.

Lemma 3.1. *Let X be a normed vector space and let $A \subset X$ be a convex set. Then $int(A)$ and $\overline{A}$ are convex sets. Moreover, if $int(A) \neq \emptyset$, then $A \subset \overline{int(A)}$.*

Proof. A general version of this lemma is proved under the framework of topological vector space, please refer to problem 1.8. We only prove the **moreover** part here. Let $x_0 \in int(A)$ and choose $\delta > 0$ such that $B_\delta(x_0) \subset A$. If $x \in A$, then the set $U_x := \{tx + (1-t)y | y \in B_\delta(x_0), 0 < t < 1\} \subset A$ is open and hence $x \in \overline{U_x} \subset \overline{int(A)}$. □

Lemma 3.2. *Let X be a normed vector sapce, let $A \subset X$ be a convex set with nonempty interior, let $\Lambda : X \to \mathbb{R}$ be a nonzero bounded linear functional, and let $c \in \mathbb{R}$ such that $\Lambda(x) \geq c$ for all $x \in int(A)$. Then $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) > c$ for all $x \in int(A)$.*

Proof. Since A is convex and has nonempty interior, we have $A \subset \overline{int(A)}$ by lemma 3.1, and so $\Lambda(x) \geq c$ for all $x \in A$ by continuity. Now let $x \in int(A)$, choose $x_0 \in X$ such that $\Lambda(x_0) = 1$, and choose $t > 0$ such that $x - tx_0 \in A$. Then $\Lambda(x) = t + \Lambda(x - tx_0) \geq t + c > c$. □

Theorem 3.3 (Separation of Convex Sets)**.** *Let X be a real normed vector space and let $A, B \subset X$ be nonempty disjoint convex sets such that $int(A) \neq \emptyset$. Then there exist a nonzero bounded linear functional $\Lambda : X \to \mathbb{R}$ and a constant $c \in \mathbb{R}$ such that $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) \leq c$ for all $x \in B$. Moreover, every such bounded linear functional satisfies $\Lambda(x) > c$ for all $x \in int(A)$.*

Proof. The proof has three steps.
Step 1. Let X be a real normed vector space, let $U \subset X$ be a nonempty open convex set such that $0 \notin U$, and define $P := \{tx | x \in U, t \in \mathbb{R}, t \geq 0\}$. Then P is a convex subset of X and satisfies the following.
(P1). If $x \in P$ and $\lambda \geq 0$, then $\lambda x \in P$.
(P2). If $x, y \in P$, then $x + y \in P$.
(P3). If $x \in P$ and $-x \in P$, then $x = 0$.
If $x, y \in P \backslash \{0\}$, choose $x_0, x_1 \in U$ and $t_0, t_1 > 0$ such that $x = t_0 x_0$ and $y = t_1 x_1$; then $z = \frac{t_0}{t_0+t_1} x_0 + \frac{t_1}{t_0+t_1} x_1 \in U$ and hence $x + y = (t_0 + t_1) z \in P$. This proves (P2). That P satisfies (P1) is obvious and that it satisfies (P3) follows from the fact that $0 \notin U$. By (P1) and (P2) the set P is convex.

Step 2. Let X and U be as in Step 1. Then there exists a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x) > 0$ for all $x \in U$.
Let P be as in Step 1. Then it follows from (P1,2,3) that the relation

$$x \leq y \quad \Leftrightarrow \quad y - x \in P$$

defines a partial order $\leq$ on X that satisfies
(O1). If $0 \leq x$ and $0 \leq \lambda$, then $0 \leq \lambda x$.

(O2). If $x \leq y$, then $x + z \leq y + z$.
Let $x_0 \in U$. Then the linear subspace $Y := \mathbb{R}x_0$ satisfies
(O3). If $x \in P$ and $-x \in P$, then $x = 0$.
Moreover, the linear functional $Y \to \mathbb{R} : tx_0 \mapsto t$ is positive. Hence by a theorem about the extension of positive linear functional[1], there is a linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(tx_0) = t$ for all $t \in \mathbb{R}$ and $\Lambda(x) \geq 0$ for all $x \in P$. We prove that this functional is bounded. Choose $\delta > 0$ such that $\overline{B}_\delta(x_0) \subset P$, and let $x \in X$ with $\|x\| \leq 1$. Then $x_0 - \delta x \in P$, hence $\Lambda(x_0 - \delta x) \geq 0$, and so $\Lambda(x) \leq \delta^{-1}\Lambda(x_0) = \delta^{-1}$. Thus $|\Lambda(x)| \leq \delta^{-1}\|x\|$ for all $x \in X$. Since $U \subset P$, we have $\Lambda(x) \geq 0$ for all $x \in U$, and so $\Lambda(x) > 0$ for all $x \in U$ by lemma 3.2.

Step 3. We prove the theorem.
Let X, A, B be as in theorem 3.3. Then $U := int(A) - B$ is a nonempty open convex set and $0 \notin U$. Hence by step 2 there is a bounded linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x) > 0$ for all $x \in U$. Thus $\Lambda(x) > \Lambda(y)$ for all $x \in int(A)$ and for all $y \in B$. This implies $\Lambda(x) \geq c := \sup_{y \in B} \Lambda(y)$ for all $x \in int(A)$. Hence $\Lambda(x) \geq c$ for all $x \in A$ and $\Lambda(x) > c$ for all $x \in int(A)$ by lemma 3.2. □

Definition 3.1 (Hyperplane)**.** *Let X be a real normed vector space. A* ***hyperplane*** *in X is a closed linear subspace of codimension one. A* ***affine hyperplane*** *is a translate of a hyperplane. An* ***open half-space*** *is a set of the form $\{x \in X | \Lambda(x) > c\}$ where $\Lambda : X \to \mathbb{R}$ is a nonzero bounded linear functional and $c \in \mathbb{R}$.*

Let X, A, B, Λ, c be as in theorem 3.3. Then $H = \Lambda^{-1}(c)$ is an affine hyperplane that separates the convex sets A and B. It divides X into two connected components such that the interior of A is contained in one of them and B is contained in the closure of the other.

References

[BS18] Theo Bühler and Dietmar A Salamon. *Functional analysis*, volume 191. American Mathematical Soc., 2018.

[Kre91] Erwin Kreyszig. *Introductory functional analysis with applications*, volume 17. John Wiley & Sons, 1991.

[1]This is also an corollary of Hahn-Banach Theorem, for more details, please refer to theorem 2.3.7 in [BS18].