The Hahn-Banach Separation Theorem and Applications

Dylan Adams, mentored by André Martins Rodrigues

Department of Mathematics, University of California at Santa Barbara

Introduction

The Hahn-Banach Separation Theorem is a key theoretical tool used in mathematics and its applications to *separate* distinct sets, either of abstract vectors or concrete quantities such as asset prices in financial markets.
The geometric picture is clearest in $\mathbb{R}^2$, where we may have two disjoint polygons (which we view here as convex, compact, disjoint sets) which we can draw a line between, each polygon falling entirely on one side. A set C is **convex** if for all $x, y \in C$ and $t \in (0,1)$ we have $tx + (1-t)y \in C$. The geometric picture is thus: the line between any two points in C lies entirely in C. Thus "nice" shapes like a circle or regular polygon are convex, while a donut is not.

The Extension Theorem

We begin by giving the proof of the **Hahn-Banach extension theorem**.
Let V be a vector space over $\mathbb{R}$, and $\psi : V \to \mathbb{R}$ be positively homogenous and subadditive. Formally, for all $x, y \in V$ and $t \geq 0$:

$$\psi(tx) = t\psi(x), \text{ and } \psi(x+y) \leq \psi(x) + \psi(y).$$

Suppose W is a subspace of V and that $\phi : W \to \mathbb{R}$ is a linear functional dominated by ψ on W (i.e. $\phi(x) \leq \psi(x)$ for $x \in W$). Then there exists a linear functional Φ defined on all of V extending ϕ.

Proof. We prepare an appeal to Zorn's lemma by considering the collection of pairs $(U(h), h)$ where $W \subseteq U(h) \subseteq V$ is a subspace of V containing W, and $h : U(h) \to \mathbb{R}$ a linear functional extending ϕ to $U(h)$ dominated by ψ. The collection admits a natural partial order defined by

$$(U(h), h) \leq (U(h'), h') \iff U(h) \subseteq U(h') \text{ and } h' \text{ extends } h.$$

Then for any chain $(U(h_\alpha), h_\alpha)$ contained in the collection, the pair (U, h) with U the union over all subspaces inside the chain, and the functional h defined by $h(x) = h_\alpha(x)$ if $x \in U_\alpha$ is an upper bound, and satisfies the conditions on the collection. By Zorn's lemma, the collection hence admits a maximal element $(U(\Phi), \Phi)$.
Suppose by contradiction that $U(\Phi) \neq V$, so that there exists $y \in V$ with $y \notin U(\Phi)$. Then let $U(g)$ be the subspace generated by $U(\Phi) \cup \{y\}$, so that a vector in $U(g)$ takes form $x + cy$ with $x \in U(\Phi)$ and $c \in \mathbb{R}$. Define $g : U(g) \to \mathbb{R}$ by

$$g(x + cy) = \Phi(x) + \delta c$$

where δ is a yet-to-be determined scalar. Then this mapping extends Φ to $U(g)$, and we will choose δ so that it is dominated by ψ. Using the linearity of Φ and positive homogeneity of ψ, this can be shown equivalent to

$$\sup_{z \in D(\Phi)} \Phi(z) - \psi(z - y) \leq \delta \leq \inf_{k \in D(\Phi)} \psi(k + y) - \Phi(k),$$

which is itself equivalent to

$$\Phi(z + z') = \Phi(z) + \Phi(z') \leq \psi(z + y) + \psi(z' - y) \ \ \forall z, z' \in D(\Phi).$$

But the domination of Φ by ψ and the subadditivity of ψ imply that

$$\Phi(z + z') \leq \psi(z + z') = \psi(z + y + z' - y) \leq \psi(z + y) + \psi(z' - y).$$

Since the conditions hold, we see that a choice of δ is possible such that ψ dominates h. For such choice we see $(U(g), g)$ is in our collection, contradicting the maximality of $(U(\Phi), \Phi)$. We conclude that $U(\Phi) = V$ and that Φ is the desired extension of ϕ. □

The Separation Theorem

As a corollary, we derive the celebrated **Hahn-Banach separation theorem**: Suppose that A and B are nonempty, disjoint, and convex subsets of a normed space V. Then we may separate in either of two cases:

(i) If A is open, then there exist $\zeta \in V^*, \gamma \in \mathbb{R}$ satisfying

$$\langle \zeta, a \rangle < \gamma \leq \langle \zeta, b \rangle, \ \ \forall a \in A, \ b \in B.$$

(ii) If A is compact, B closed, then there exist $\zeta \in V^*$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ satisfying

$$\langle \zeta, a \rangle < \lambda_1 < \lambda_2 < \langle \zeta, b \rangle, \ \ \forall a \in A, \ b \in B.$$

Proof. We sketch the idea of part (a). For $a \in A$ and $b \in B$, we set $c = b - a$ and $C = A - B + c = \{x - y + c \ : \ x \in A, \ y \in B\}$. It follows that C is an open convex neighborhood of zero. We define the *gauge function* $\psi : V \to \mathbb{R}$ of C by $\psi(v) = \inf\{\delta > 0 \ : \ v \in \delta C\}$, where $\delta C = \{\delta z \ : \ z \in C\}$. Letting $W = \mathbb{R}c$ and defining $\phi(tz) = t$ on W, we deduce immediately that ϕ is dominated by ψ on W, and by the previous theorem there hence exists a functional ζ on V extending ϕ and dominated by ψ. Since $\zeta \leq 1$ on C, a neighborhood of zero, it follows by linearity that ζ is continuous, and if $x \in A, y \in B$, then $x - y + c \in C$ so that

$$\langle \zeta, x \rangle - \langle \zeta, y \rangle + 1 = \langle \zeta, x - y + z \rangle \leq \psi(x - y + z) < 1,$$

so that $\langle \zeta, x \rangle < \langle \zeta, y \rangle$. By linearity and this inequality, we see that $\zeta(A)$ and $\zeta(B)$ are disjoint intervals in $\mathbb{R}$, and by the general result that nonzero linear functionals on normed spaced are open mappings, we have $\zeta(A)$ an open interval. The result follows for $\gamma := \sup \zeta(A)$. □

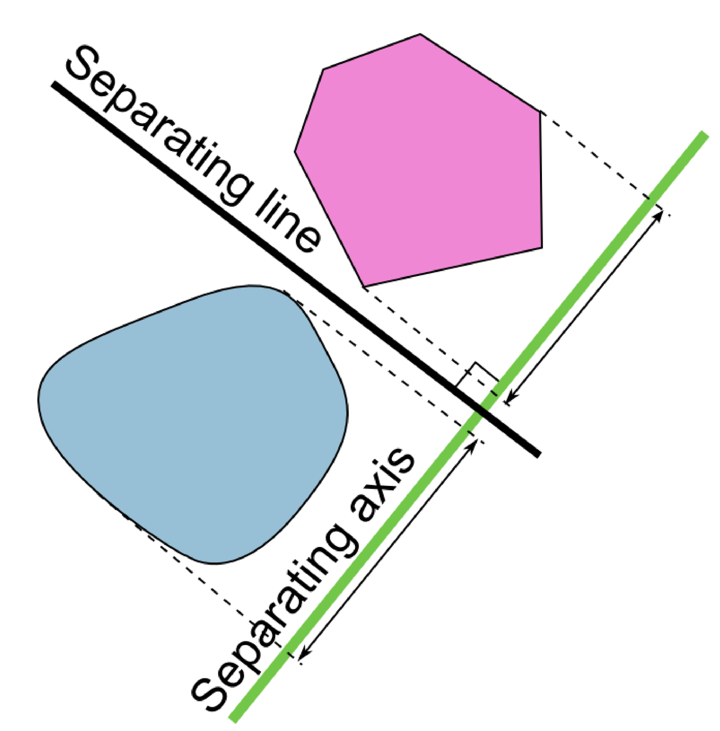

Fig. 1: Created by Oleg Alexandrov, Inkscape.

Theoretical Applications

Perhaps the most crucial application of the separation theorem is showing that the dual space V^* separates points in V – that is, for all $v, w \in V$ with $v \neq w$, there exists $\zeta \in V^*$ such that $\langle \zeta, x \rangle \neq \langle \zeta, y \rangle$. This simple observation allows one to define the weak topology on V as a special case of a *induced topology*. If $\mathcal{F}$ is a family of linear functionals defined on a vector space V which separates points in V as above, we define the induced topology to have a sub-base of open sets given by sets of the form

$$V(x, \phi, r) = \{y \in V \ : \ |\phi(x - y)| < r\}$$

for $x \in V, \phi \in \mathcal{F}$, and $r > 0$. This means every open set is the union of finite intersections of the above basic sets. The induced topology is the initial topology with respect to the family $\mathcal{F}$; i.e., it is the coarsest topology such that each functional $\phi \in \mathcal{F}$ is continuous in the topology. Intuitively, one desires a topology with fewer open sets so that more sets become compact. In addition, new sequences may converge; for example the sequence $\{\sqrt{2/\pi} \sin kx\}$ forms an orthonormal basis of $L^2(-\pi/2, \pi/2)$, and the strong limit as $k \to \infty$ does not exist while the weak limit as $k \to \infty$ is zero by the Riemann-Lebesque lemma (see Stein-Shakarchi, "Fourier Analysis").

Practical Applications

In the case of a finite dimensional vector space, say $\mathbb{R}^n$ (with no loss of generality since all finite dimensional normed spaces have an equivalent norm which is isometric to $(\mathbb{R}^n, |\cdot|)$), the general separation result implies any disjoint convex subsets can be separated by a hyperplane. In $\mathbb{R}^2$, this simply means that convex shapes such as those illustrated can be separated by a line. This constitutes a theoretical basis for the popular "Support Vector Machine" algorithm in machine learning, wherein one classifies data in high-dimensional feature space by separating hyperplanes.
In mathematical economics, **Fundamental Theorem of Asset Pricing** states that a market is arbitrage-free (meaning there are no risk-free ways to profit from buying and immediately selling goods) if and only if there is a risk neutral probability measure equivalent to the original. This essentially means that in the absense of arbitrage, one does not need to consider individual investor's discounted valuations of assets based on identified risk, but can once and for all price assets via a probability measure based on the probabilities of future outcomes.
To situate this in our theorem, consider the set of wealths attained via trading in a market starting from no wealth, and the wealth gained via arbitrage opportunities. Under certain conditions these sets satisfy the hypotheses of the Hahn-Banach separation theorem, and this observation leads to the fundamental theorem.

References

- Clarke, Francis. "Functional Analysis, Calculus of Variations, and Optimal Control." Springer. 2013.
- Stein, Elias; Shakarchi, Rami. "Fourier Analysis." Princeton University Press. 2003.
- 2020 Mathematics Directed Reading Program. University of California - Santa Barbara.
- Lewis, Keith. A Simple Proof of the Fundamental Theorem of Asset Pricing. arXiv. 2 Dec 2019. arXiv:1912.01091