Hahn-Banach Theorem

It's one of the most powerful tools to study dual space, with surprising connections with machine learning ...

Thm. (Hahn - Banach):
Let X be a vector space over $\mathbb{R}$, $p: X \to [0, +\infty)$ a quasi-norm:
1) $p(\lambda x) = \lambda p(x) \quad \forall x \in X, \ \forall \lambda \in \mathbb{R}, \lambda \geqslant 0$ (positive homogeneity)
2) $p(x+y) \leqslant p(x) + p(y) \quad \forall x, y \in X$

Let Y be a proper vector subspace of X, $T: Y \to \mathbb{R}$ linear s.t. $T(x) \leqslant p(x) \ \forall x \in Y$. Then there $\exists \tilde{T}: X \to \mathbb{R}$ linear extension of T s.t. $\tilde{T}(x) \leqslant p(x) \ \forall x \in X$

In particular, if $(X, \|\cdot\|)$ is a normed space and $T \in Y'$, then we can extend it to $\tilde{T}: X \to \mathbb{R}$ linear with $\|\tilde{T}\|_{X'} = \|T\|_{Y'}$
(this corollary follows by choosing $p(x) = \|T\|_{Y'} \cdot \|x\|$, so we have $\tilde{T}(x) \leqslant \|T\|_{Y'} \cdot \|x\|$)

Corollary:
Given $(X, \|\cdot\|)$ normed space, Y proper vector subspace of X, $T: Y \to \mathbb{R}$ linear and continuous. Then $\exists \tilde{T}: X \to \mathbb{R}$ linear extension of T s.t. $\|\tilde{T}\|_{X'} = \|T\|_{Y'}$

Corollary:
Given $(X, \|\cdot\|)$ a normed space, we have the following:
1) $\forall x_0 \in X \ \exists T \in X'$ s.t. $\|T\|_{X'} = 1$ and $T(x_0) = \|x_0\|$
2) $\forall x \in X, \ \|x\| = \max\{T(x) : T \in X', \|T\|_{X'} \leqslant 1\}$
3) $x, y \in X, \ x \neq y \Rightarrow \exists T \in X'$ s.t. $T(x) \neq T(y)$ (The Dual Space X' separates points of X)

Proof:
1) Let $Y = \mathbb{R}\{x_0\} = \{t x_0 : t \in \mathbb{R}\}$, define $T: Y \to \mathbb{R}$ s.t.
$T(t x_0) = t\|x_0\| \Rightarrow \|T\|_{X'} = 1 \Rightarrow$ extend by Hahn-Banach:
$$\exists \tilde{T} \in X' \text{ s.t. } \|\tilde{T}\|_{X'} = 1 \ \wedge \ \tilde{T}(x_0) = T(x_0) = \|x_0\|$$
2) $(\leqslant)$: $T(x) \leqslant \underbrace{\|T\|_{X'}}_{\leqslant 1} \cdot \|x\| \leqslant \|x\|$

$(\geqslant)$: thanks to 1.

$\square$

Proof (Hahn - Banach):

CLAIM:

Z proper vector subspace of X, $T: Z \to \mathbb{R}$ linear s.t. $T(x) \leq p(x)$ $\forall x \in Z$ $\Rightarrow$ $\exists \tilde{T}: \tilde{Z} \to \mathbb{R}$ with $\tilde{Z}$ strictly larger than Z s.t. $\tilde{T}(x) \leq p(x)$ $\forall x \in \tilde{Z}$, $\tilde{T}(x) = T(x)$ $\forall x \in Z$

PROOF OF THE CLAIM:

Let $x_0 \in X \setminus Z$ $\Rightarrow$ we extend $T: Z \to \mathbb{R}$ to $\tilde{Z} = Z \oplus \mathbb{R}\{x_0\}$ $= \{x + tx_0 : x \in Z, t \in \mathbb{R}\}$. Then:

$$\tilde{T}(x + tx_0) = \tilde{T}(x) + t\tilde{T}(x_0) = T(x) + t\underbrace{\tilde{T}(x_0)}_{= \alpha}$$

$\Rightarrow$ The condition we need to prove is:

$$(*) \quad T(x) + t\alpha \leq p(x + tx_0) \quad \forall x \in Z, \forall t \in \mathbb{R}$$

$t>0)$

$(*) \Leftrightarrow \cancel{t}\left(T\left(\frac{x}{t}\right) + \alpha\right) \leq \cancel{t}\, p\left(\frac{x}{t} + x_0\right) \Leftrightarrow T(x) + \alpha \leq p(x + x_0)$ $\forall x \in Z$

$\Leftrightarrow \alpha \leq T(x) - p(x + x_0)$

$t<0)$

$(*) \Leftrightarrow -\cancel{t}\left(T\left(-\frac{x}{t}\right) - \alpha\right) \leq -\cancel{t}\, p\left(-\frac{x}{t} - x_0\right) \Leftrightarrow T(y) - \alpha \leq p(y - x_0)$ $\forall y \in Z$

$\Leftrightarrow \alpha \geq T(y) - p(y - x_0)$

$\Rightarrow T(y) - p(y - x_0) \leq \alpha \leq T(x) - p(x + x_0) \quad \forall x, y \in Z$

We now show that $T(y) - p(y - x_0) \leq T(x) - p(x + x_0)$ $\forall x, y \in Z$:

$\Leftrightarrow T(x) + T(y) \leq p(x + tx_0) \Leftrightarrow T(x+y) \leq p(x+y) \leq p(x+x_0) + p(y-x_0)$

So the claim is proved. We now prove the main statement of the Hahn - Banach Thm. Let $F = \{S: Z \to \mathbb{R} : S$ linear, $Z \supset Y$, Z a vector subspace, S extends T and $S(x) \leq p(x)$ $\forall x \in Z\}$. We put a partial order relation on F:

$$(S_1 : Z_1 \to \mathbb{R}) \leq (S_2 : Z_1 \to \mathbb{R}) \Leftrightarrow S_2 \text{ extends } S_1$$

Thanks to the Hausdorff Maximal Principle $\exists\, \mathcal{G} \subset F$ s.t. $\mathcal{G}$ is a maximal totally ordered set:

$$\mathcal{G} = \{S_\alpha : Z_\alpha \to \mathbb{R}, \alpha \in I\}$$

$\Rightarrow \exists$ upper bound of $\mathcal{G}$: $S: Z \to \mathbb{R}$ with $Z = \bigcup_{\alpha \in I} Z_\alpha$ s.t.

$$S(x) := S_\alpha(x) \quad \forall x \in Z_\alpha$$

$\square$

Dual Space of the L^p space

We prove now that the L^p-norm $\|\cdot\|_{L^p(\mu)}$ is, indeed, a norm:

<u>Lemma</u> (Young's inequality):
Let $p \in (1,+\infty)$, q conjugate exponent of p (i.e. $\frac{1}{p}+\frac{1}{q}=1$ $\Leftrightarrow q=\frac{p}{p-1}$). Then $\forall a,b \in \mathbb{R}$, $a,b \geq 0$ we have $ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$

<u>Proof</u>:
$x \longmapsto \log x$ is concave
$\log(ab) = \log a + \log b = \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) \leq \log\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right)$
and we take the exp on both sides.
$\square$

<u>Lemma</u> (Hölder's inequality):
Let $p \in [1,+\infty]$, q conjugate exponent of p, μ outer measure of a set X, let $u, v : X \longrightarrow \overline{\mathbb{R}}$ μ-meas. Then we have:

$$\int_X |u(x)v(x)|\,d\mu(x) \leq \|u\|_{L^p(\mu)} \cdot \|v\|_{L^q(\mu)}$$

<u>Proof</u>:
Obvious if either one of the 2 norms are $+\infty$ and the other is positive, or when one of the norms is 0. Remember that $\|u\|_{L^p(\mu)} = \left(\int_X |u(x)|^p d\mu\right)^{\frac{1}{p}}$. Assume then WLOG $0 < \|u\|_{L^p(\mu)}, \|v\|_{L^p(\mu)} < +\infty$
$\Leftrightarrow \int_X \left|\frac{u}{\|u\|_{L^p}} \cdot \frac{v}{\|v\|_{L^q}}\right| d\mu \leq 1 \Rightarrow$ it's enough to prove the inequality when $\|u\|_{L^p} = \|v\|_{L^q} = 1$:

$$\int_X |u(x)|\cdot \underbrace{|v(x)|}_{\leq\, \frac{1}{p}|u|^p + \frac{1}{q}|v|^q}\,d\mu(x) \leq \int_X \frac{1}{p}|u|^p + \frac{1}{q}|v|^q\,d\mu(x) = \frac{1}{p}\cdot 1 + \frac{1}{q}\cdot 1 = 1.$$

Case $p=1$, $q=+\infty$ (the other case is symmetric):
$\Rightarrow |v(x)| \leq \|v\|_{L^\infty}$ μ-a.e. $x \in X$
$\Rightarrow \int_X |u(x)|\cdot \underbrace{|v(x)|}_{\leq\, \|v(x)\|_{L^\infty(\mu)}}\,d\mu(x) \leq \|v\|_{L^\infty}\int_X |u(x)|\,d\mu(x) = \|u\|_{L^1}\|v\|_{L^\infty}$
$\square$

<u>Lemma</u> (Minkowski's inequality):
$u, v : X \longrightarrow \overline{\mathbb{R}}$ μ-meas. $\Rightarrow \|u+v\|_{L^p(\mu)} \leq \|u\|_{L^p(\mu)} + \|v\|_{L^p(\mu)}$ $\forall p \in [1,+\infty]$

<u>Proof</u>:
Case $p=1, +\infty$ are trivial.

Case $p \in (1, +\infty)$:

$\Rightarrow$ Remark: $u, v \in L^p(\mu) \Rightarrow u+v \in L^p(\mu)$ (comes from the fact that $y \longmapsto |y|^p$ is convex, so $|\frac{u+v}{2}|^p \leq \frac{1}{2}(|u|^p + |v|^p)$)

We have:

$$\int_X |u(x) + v(x)|^p d\mu = \int_X |u(x)+v(x)|^{p-1} \cdot \underbrace{|u(x)+v(x)|}_{\leq\, |u(x)|+|v(x)|} d\mu$$

$$\leq \int_X |u(x)| \cdot |u(x)+v(x)|^{p-1} d\mu + \int_X |v(x)| \cdot |u(x)+v(x)|^{p-1} d\mu$$

$$\leq \|u\|_{L^p(\mu)} \left(\int_X |u(x)+v(x)|^p d\mu \right)^{\frac{p-1}{p}} + \|v\|_{L^p(\mu)} \left(\int_X |u(x)+v(x)|^p d\mu \right)^{\frac{p-1}{p}}$$

(Hölder's ineq.)

$$= \left(\|u\|_{L^p(\mu)} + \|v\|_{L^p(\mu)} \right) \left(\int_X |u(x)+v(x)|^p d\mu \right)^{\frac{p-1}{p}}$$

$\Rightarrow$ so we have $\left(\int_X |u(x)+v(x)|^p d\mu \right)^{\frac{1}{p}} \leq \|u\|_{L^p(\mu)} + \|v\|_{L^p(\mu)}$

$\square$

Thm. (Riesz – Fischer):

$L^p(\mu)$ is Banach. Moreover, if $\{u_n\}_{n \in \mathbb{N}} \subset L^p(\mu) \wedge u_n \xrightarrow{L^p} u$ then $\exists \{u_{n_k}\}_{k \in \mathbb{N}}$ subsequence s.t. $u_{n_k}(x) \xrightarrow[n \to +\infty]{} u(x)$ for μ-a.e. $x \in X$

(No need to extract a subsequence if $p = +\infty$).

Proof:

Case $p \in [1, +\infty)$:

Let $\{u_n\}_n$ be a Cauchy seq. in $L^p(\mu)$: $\forall \varepsilon > 0 \; \exists \nu \in \mathbb{N}$ s.t. $\forall n, m \geq \nu \quad \|u_n - u_m\|_{L^p} < \varepsilon$. We can find n_k increasing seq. of indexes s.t. $\|u_{n_{k+1}} - u_{n_k}\|_{L^p(\mu)} < \frac{1}{2^k}$. Consider the series

$$g(x) = \sum_{k=1}^{+\infty} |u_{n_{k+1}}(x) - u_{n_k}(x)| \in L^p(\mu), \quad g_K(x) = \sum_{k=1}^{K} |u_{n_{k+1}}(x) - u_{n_k}(x)|$$

$\Rightarrow \|g_K\|_{L^p(\mu)} < 1 \quad \forall K \in \mathbb{N} \Rightarrow \|g_K\|_{L^p(\mu)} \xrightarrow[K \to +\infty]{} \|g\|_{L^p(\mu)}$ (Beppo – Levi)

$\Rightarrow g(x) \in \mathbb{R}$ for μ-a.e. $x \in X \Rightarrow \sum_{k=1}^{+\infty} (u_{n_{k+1}}(x) - u_{n_k}(x))$ converges absolutely for a.e. $x \in X$ and it also is telescopic:

$$\sum_{k=1}^{K} (u_{n_{k+1}}(x) - u_{n_k}(x)) = u_{n_{K+1}}(x) - u_{n_1}(x) \longrightarrow u(x) - u_{n_1}(x)$$

$\Rightarrow u_{n_{K+1}}(x) \longrightarrow u(x)$ for μ-a.e. $x \in X$

$\Rightarrow |u_{n_{K+1}}(x)| \leq (g(x) + |u_{n_1}(x)|) \in L^p(\mu) \Rightarrow u_{n_k} \xrightarrow{L^p(\mu)} u$ (Dominated conv.)

$\Rightarrow$ the subsequence converges in $L^p(\mu) \Rightarrow$ the whole sequence converges in $L^p(\mu)$

$\Rightarrow \{u_n\}$ is Cauchy because it converges $\Rightarrow \exists\, u_{n_k}(x) \xrightarrow[k\to+\infty]{L^p(\mu)} v(x)$ μ-a.e.

$\Rightarrow u = v$ (uniqueness of the limit in L^p)

Case $p = +\infty$:

Let $\{u_n\}_{n\in\mathbb{N}} \subset L^\infty(\mu)$ be a Cauchy seq. $\Rightarrow \forall k \in \mathbb{N}$, $k \neq 0$ fixed, $\exists\, n_k \in \mathbb{N}$ s.t.:

$$\|u_n - u_m\|_{L^\infty(\mu)} \leq \frac{1}{k} \quad \forall n, m \geq n_k$$

(Def. of Cauchy seq. with $\varepsilon = \frac{1}{k}$). But:

$\forall n, m \geq n_k$ $|u_n(x) - u_m(x)| \leq \frac{1}{k}$ is true exept for a set of measure 0. Define:

$$A_k = \{x \in X : |u_n(x) - u_m(x)| > \frac{1}{k},\ n, m \geq n_k\} \Rightarrow \mu(A_k) = 0$$

$\Rightarrow A = \bigcup_{k=1}^{+\infty} A_k \Rightarrow \mu(A) = 0 \Rightarrow$ if $x \in X \setminus A$, then $\{u_n(x)\}_n$ is Cauchy in $\mathbb{R} \Rightarrow u_n(x) \xrightarrow[n]{} u(x)$ and $u_n \xrightarrow[\text{unif.}]{} u$ in $X\setminus A$.

$\square$