Zorn's Lemma

Partial Order

A partial order $(\mathcal{S}, \leq)$ is a nonempty set together with a relation $\leq$ on $\mathcal{S} \times \mathcal{S}$ which is:

reflexive: $A \leq A \quad \forall A \in \mathcal{S}$

antisymmetric: $A \leq B \;\&\; B \leq A \Rightarrow A = B$

Transitive: $A \leq B \;\&\; B \leq C \Rightarrow A \leq C$.

A partial ordering s.t. every pair $A, B \in \mathcal{S}$ is comparable is a linear or total ordering.

A nonempty subset of $\mathcal{S}$ that is linearly ordered by $\leq$ is a chain in $\mathcal{S}$.

An element $A \in \mathcal{S}$ is maximal in $\mathcal{S}$ if

$$B \in \mathcal{S} \;\&\; B \text{ comparable to } A \Longrightarrow B \leq A.$$

A maximal element need not be comparable to all elements of $\mathcal{S}$, & need not be unique.

An upper bound for a collection $\mathcal{U} \subset \mathcal{S}$ is an element $A \in \mathcal{S}$ s.t. $B \leq A \;\forall B \in \mathcal{U}$.

Zorn's Lemma: If $(\mathcal{S}, \leq)$ is partially ordered & every chain in $\mathcal{S}$ has an upper bound in $\mathcal{S}$, then $\mathcal{S}$ contains a maximal element.

Proof of the Hahn-Banach Theorems

Hahn-Banach for $\mathbb{F} = \mathbb{R}$

Let X be a real vector space
$q: X \to \mathbb{R}$ sublinear
M a subspace of X
$\lambda: M \to \mathbb{R}$ linear functional with $\langle x, \lambda \rangle \leq q(x)$ $\forall x \in M$.

Then $\exists$ linear functional $\Lambda: X \to \mathbb{R}$ s.t.

$$\Lambda|_M = \lambda \quad \& \quad \langle x, \Lambda \rangle \leq q(x) \ \forall x \in X.$$

Proof

Case 1: Suppose M is a hyperplane, i.e., $\dim(X/M) = 1$.

Since $M \subsetneq X$, we can choose $x_0 \in X/M$. Define

$$M_1 = \text{span}(M, x_0).$$

Exercise: Show that $M_1 = X$, and that every $x \in X$ can be written

$$x = m + c x_0$$

for a unique choice of $m \in M$, $c \in \mathbb{R}$.

Now, given $x = m + cx_0 \in X$, we know that $\langle m, \lambda \rangle$ exists. In order to define Λ, we will set $\langle m, \Lambda \rangle = \langle m, \lambda \rangle$. Therefore, the issue is how to define $\langle x_0, \Lambda \rangle = \alpha_0 \in \mathbb{R}$ so that the required properties of Λ will hold.

<u>Motivation</u>

Suppose that we have found Λ. Then $\forall m \in M$,

$$\begin{aligned} \langle x_0, \Lambda \rangle &= \langle x_0 + m, \Lambda \rangle - \langle m, \Lambda \rangle && \Lambda \text{ linear} \\ &\leq q(x_0 + m) - \langle m, \lambda \rangle && \Lambda \leq q \ \& \ \Lambda|_M = \lambda \end{aligned}$$

and also

$$\begin{aligned} \langle x_0, \Lambda \rangle &= \langle m, \Lambda \rangle - \langle m - x_0, \Lambda \rangle && \Lambda \text{ linear} \\ &\geq \langle m, \lambda \rangle - q(m - x_0) && \Lambda \leq q \ \& \ \Lambda|_M = \lambda. \end{aligned}$$

Therefore,

(*) $$\sup_{m\in M}\left[\langle m,\lambda\rangle - q(m-x_0)\right] \le \langle x_0,\Lambda\rangle \le \inf_{m\in M}\left[q(x_0+m)-\langle m,\lambda\rangle\right].$$

This is what our scalar $\alpha_0=\langle x_0,\Lambda\rangle$ must necessarily satisfy, if we are to be able to construct Λ.

Is there any such scalar? The answer is yes <u>if</u>:

(**) $$\forall m_1, m_2 \in M,\quad \langle m_1,\lambda\rangle - q(m_1-x_0) \le q(x_0+m_2)-\langle m_2,\lambda\rangle.$$

Let us verify that (**) is true:

$$\begin{aligned}
\langle m_1,\lambda\rangle + \langle m_2,\lambda\rangle &= \langle m_1+m_2,\lambda\rangle \\
&\le q(m_1+m_2) && \lambda\le q \\
&= q(m_1-x_0+x_0-m_2) \\
&\le q(m_1-x_0)+q(x_0-m_2) && q \text{ sublinear.}
\end{aligned}$$

Thus (**) is true, and therefore we can choose a scalar $\langle x_0,\Lambda\rangle$ so that (*) holds — just let $\langle x_0,\Lambda\rangle$ be <u>any</u> scalar s.t. (*) holds. This then defines Λ for all $x \in X$: write $x = m + c x_0$ uniquely and set $\langle x, \Lambda \rangle = \langle m, \lambda \rangle + c \langle x_0, \Lambda \rangle$.

This gives a well-defined definition of Λ, & we have $\Lambda|_M = \lambda$.

Exercise: Show Λ is linear.

It therefore only remains to show that $\Lambda \le q$.

Suppose $x = m + c x_0$ with $c > 0$. Then

$$\langle x, \Lambda \rangle = \langle m, \lambda \rangle + c \langle x_0, \Lambda \rangle$$

$$\le \langle m, \lambda \rangle + c \left[q\left(x_0 + \tfrac{m}{c}\right) - \left\langle \tfrac{m}{c}, \lambda \right\rangle \right] \qquad \text{since } \tfrac{m}{c} \in M$$

$$= \langle m, \lambda \rangle + q(c x_0 + m) - \langle m, \lambda \rangle \qquad q \text{ sublinear}$$

$$= q(x).$$

A similar argument applies if $c < 0$. If $c = 0$ then $x \in M$, and we already know $\lambda \le q$ on M.

<u>Case 2:</u> M arbitrary.

We will apply Zorn's Lemma. Define

$$\mathcal{S} = \left\{ (M_1, \lambda_1) : \begin{array}{l} M \subseteq M_1,\ M_1 \text{ subspace of } X, \\ \lambda_1 : M_1 \to \mathbb{R} \text{ linear},\ \lambda_1|_M = \lambda,\ \lambda_1 \le q \text{ on } M_1 \end{array} \right\}$$

Order by

$$(M_1, f_1) \le (M_2, f_2) \iff M_1 \subseteq M_2 \ \& \ \lambda_2|_{M_2} = \lambda_1$$

Exercise: Show that $\le$ is a partial order on $\mathcal{S}$.

Suppose now that $\mathcal{C} = \{(M_i, f_i) : i \in I\}$ is a chain in $\mathcal{S}$, i.e., $\forall i, j \in I$ either $(M_i, f_i) \le (M_j, f_j)$ or $(M_j, f_j) \le (M_i, f_i)$. We must show that this chain has an upper bound.

Define

$$N = \bigcup_{i \in I} M_i.$$

Exercise: Show N is a subspace of X, & $M \subseteq N$.

Define
$$\Lambda : N \longrightarrow \mathbb{R}$$
by defining $\langle x, \Lambda \rangle = \langle x, \lambda_i \rangle$ if $x \in M_i$.

Exercise: Show that the fact that $\mathcal{C}$ is a chain implies that Λ is well-defined, linear, & satisfies $\Lambda \leq q$ on N.

Therefore $(N, \Lambda) \in \mathcal{S}$. Also $(N, \Lambda) \geq (M_i, \lambda_i)$ for all $i \in I$, so (N, Λ) is an upper bound for $\mathcal{C}$.

Thus, we've shown that every chain in $\mathcal{S}$ has an upper bound. Zorn's Lemma therefore implies that $\mathcal{S}$ has a maximal element, say (Y, Λ).

Suppose that $Y \neq X$. In this case, we can choose $x_0 \in X \setminus Y$. But then, if we define $Y_1 = \operatorname{span}(Y, x_0)$, then $\dim(Y_1 / Y) = 1$, so Case 1 implies that $\exists \Lambda_1 : Y_1 \rightarrow \mathbb{R}$ s.t. $\Lambda_1|_Y = \Lambda$.

But then $(Y, \Lambda) \lneq (Y_1, \Lambda_1)$, contradicting the fact (Y, Λ) is maximal.

Hence we must have $Y = X$, and so Λ is our desired functional. $\blacksquare$

Hahn Banach for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$

Let X be a vector space over $\mathbb{F}$

$p : X \longrightarrow [0, \infty)$ a seminorm

M a subspace of X

$\lambda : M \longrightarrow \mathbb{F}$ linear with $|\langle x, \lambda \rangle| \leq p(x) \ \forall x \in M$.

Then $\exists$ linear functional $\Lambda : X \to \mathbb{F}$ s.t.

$$\Lambda|_M = \lambda \quad \& \quad |\langle x, \Lambda \rangle| \leq p(x) \quad \forall x \in X.$$

Proof:

Case $\mathbb{F} = \mathbb{R}$

We have $\lambda \leq p$ & p sublinear, so by Hahn-Banach for $\mathbb{R}$, $\exists$ linear $\Lambda : X \longrightarrow \mathbb{R}$ s.t.

$\Lambda|_M = \lambda$ & $\Lambda \leq p$. However, since p is a seminorm we have

$$-\langle x, \Lambda \rangle = \langle -x, \Lambda \rangle \leq p(-x) = p(x),$$

so $|\langle x, \Lambda \rangle| \leq p(x) \quad \forall x \in X$.

Case $\mathbb{F} = \mathbb{C}$

Note that since X is a vector space over $\mathbb{C}$, it is also a vector space over $\mathbb{R}$. Define $g = \operatorname{Re} \lambda$, i.e.,

$$\langle x, g \rangle = \operatorname{Re} \langle x, \lambda \rangle, \quad x \in X.$$

Then $g : X \to \mathbb{R}$, and g is $\mathbb{R}$-linear, i.e.,

$$\langle \alpha x + \beta y, g \rangle = \alpha \langle x, g \rangle + \beta \langle y, g \rangle \quad \forall x, y \in X, \ \forall \alpha, \beta \in \mathbb{R}.$$

Further,

$$|\langle x, g \rangle| \leq |\langle x, \lambda \rangle| \leq p(x) \quad \forall x \in M.$$

Hence by Hahn-Banach for $\mathbb{R}$, $\exists$ $\mathbb{R}$-linear functional $G : X \to \mathbb{R}$ s.t.

$$G|_M = g \quad \& \quad |\langle x, G \rangle| \leq p(x) \ \forall x \in X.$$

<u>Claim:</u> $\operatorname{Im} \langle x, \lambda \rangle = -\langle ix, g \rangle \quad \forall x \in M.$

<u>Note:</u> Since we only know that g is $\underline{\mathbb{R}}$-linear, we cannot say that $\langle ix, g \rangle = i \langle x, g \rangle$. We know that

maps all of M to $\mathbb{R}$, so if $x \in M$ then

$\langle x, g \rangle$ & $\langle ix, g \rangle$ are each real scalars, but we don't know know their relationship.

To prove the claim, let $h = \operatorname{Im} \lambda$, so

$\langle x, \lambda \rangle = \langle x, g \rangle + i \langle x, h \rangle$ with $\langle x, g \rangle \in \mathbb{R}$, $\langle x, h \rangle \in \mathbb{R}$.

Then since λ is $\mathbb{C}$-linear,

$$i \langle x, \lambda \rangle = \langle ix, \lambda \rangle = \langle ix, g \rangle + i \langle ix, h \rangle.$$

Therefore

$$\langle x, \lambda \rangle = \langle ix, h \rangle - i \langle ix, g \rangle.$$

But $\langle ix, h \rangle$ & $\langle ix, g \rangle$ are real, so this implies that

$$\operatorname{Re} \langle x, \lambda \rangle = \langle x, g \rangle = \langle ix, h \rangle = \operatorname{Im} \langle ix, \lambda \rangle$$

$$\operatorname{Im} \langle x, \lambda \rangle = \langle x, h \rangle = - \langle ix, g \rangle = - \operatorname{Re} \langle ix, \lambda \rangle.$$

This is valid for $x \in M$.

We have already extended the real part of λ to X.

The extension of the imaginary part is therefore already determined, i.e., we must define $\Lambda : X \to \mathbb{C}$ by

$$\langle x, \Lambda \rangle = \langle x, G \rangle - i \langle ix, G \rangle.$$

Exercise: Show Λ defined in this way is $\mathbb{C}$-linear.

Remember that we only know that G is $\mathbb{R}$-linear.

Claim: $\Lambda|_M = \lambda$

If $x \in M$, then

$$\begin{aligned} \langle x, \Lambda \rangle &= \langle x, G \rangle - i \langle ix, G \rangle \\ &= \langle x, g \rangle - i \langle ix, g \rangle \quad \text{since } G|_M = g \\ &= \langle x, \lambda \rangle. \end{aligned}$$

Claim: $|\langle x, \Lambda \rangle| \leq p(x)$.

If $x \in X$, then

$$
\begin{aligned}
|\langle x, \Lambda\rangle| &= c\,\langle x, \Lambda\rangle \qquad \text{where } |c| = 1 \\
&= \langle cx, \Lambda\rangle \\
&= \operatorname{Re}\langle cx, \Lambda\rangle \\
&= \langle cx, G\rangle \\
&\leq p(cx) \\
&= |c|\,p(x) \qquad \text{since } p \text{ is a seminorm} \\
&= p(x).
\end{aligned}
$$

Real!

Thus Λ has all the required properties. ∎