Lecture 7: The Hahn-Banach Theorem.

Hyperplanes in Linear Spaces.

Recall from linear algebra the definition of a plane in $\mathbf{R}^3$. Such a plane was determined by a point in the plane, (x_0, y_0, z_0) and a vector normal to the plane $\vec{\eta}$. Then an arbitrary point (x, y, z) in the plane must satisfy

$$\langle \vec{p} - \vec{p_0}, \vec{\eta} \rangle = 0$$

or equivalently

$$\langle \vec{p}, \vec{\eta} \rangle = \langle \vec{p_0}, \vec{\eta} \rangle$$

where $\vec{p} = \langle x, y, z \rangle$ and $\vec{p_0} = \langle x_0, y_0, z_0 \rangle$. Of course, this definition can be extended to $\mathbf{R}^n$ where the same formula is used but all the vectors are n-vectors.

The question is: How do we extend the notion of a hyperplane to infinite dimensional linear spaces where we have no notion of "orthogonality" or of "inner product?" We make the following observations about the $\mathbf{R}^n$ case.

1. $f_{\vec{\eta}}(\vec{x}) = \langle \vec{x}, \vec{\eta} \rangle$ defines a linear functional on $\mathbf{R}^n$.

2. The set

$$null(f) = \{\vec{x} \colon \langle \vec{x}, \vec{\eta} \rangle = 0\} = \{\vec{x} \colon f(\vec{x}) = 0\}$$

is a linear subspace of $\mathbf{R}^n$.

3. The hyperplane P through $\vec{p_0}$ orthogonal to $\vec{\eta}$ can be written

$$P = \vec{p_0} + null(f) = \{\vec{x} \colon \vec{x} = \vec{p_0} + \vec{y},\ \vec{y} \in null(f)\}.$$

4. Since *all* linear functionals on $\mathbf{R}^n$ can be written as $f_{\vec{\eta}}(\vec{x}) = \langle \vec{x}, \vec{\eta} \rangle$ for some $\vec{\eta} \in \mathbf{R}^n$, all hyperplanes in $\mathbf{R}^n$ can be written as $P = \vec{p_0} + null(f)$ for some linear functional f and some fixed vector $\vec{p_0}$.

Definition 59. *A subspace L' of a real linear space L has* codimension 1 *if there exists a vector $x_0 \in L$, $x_0 \neq 0$, such that every $x \in L$ can be written uniquely as $x = \alpha x_0 + y$ for some $\alpha \in \mathbf{R}$ and $y \in L'$.*

Remark 60. Let f be a linear functional on a real linear space L. Then by a previous theorem, $null(f)$ has codimension 1.

Definition 61. *Let L' be a codimension 1 subspace of the linear space L. A set $M' \subseteq L$ is a* hyperplane parallel to L' *if there exists an element $x_0 \in L$ such that*

$$M' = L' + x_0 = \{x \colon x_0 + y,\ y \in L'\}.$$

(note the typo on p. 127 of Kolmogorov).

Theorem 62.

(a) *Let f be a nonzero linear functional on a linear space L. Then the set*

$$M_f = \{x\colon f(x) = 1\}$$

is a hyperplane parallel to $L' = null(f)$.

(b) *Let $M' = L' + x_0$ be any hyperplane parallel to L' not containing* 0 (*that is, such that $x_0 \notin L'$*)*. Then there exists a unique nonzero linear functional f on L such that*

$$M' = \{x\colon f(x) = 1\}.$$

Proof: (a) We must show that $M_f = L' + x_0$ for some $x_0 \in L$. Let x_0 be any vector such that $f(x_0) = 1$ (such an x_0 always exists since f is not the identically zero functional). We will show that this x_0 is the one we want. Let $x \in M_f$. Any vector of the form $x = x_0 + y$ where $y \in L'$ is in M_f since $f(x) = f(x_0 + y) = f(x_0) + f(y) = f(x_0) = 1$. Therefore, $L' + x_0 \subseteq M_f$. Let $x \in M_f \subseteq L$. Since L' is the nullspace of a linear functional, x can be written $x = \alpha x_0 + y$ for some $\alpha \in \mathbf{R}$ and $y \in L'$. But since $x \in M_f$,

$$1 = f(x) = f(\alpha x_0 + y) = \alpha f(x_0) + f(y) = \alpha$$

so that $x \in L' + x_0$ as required.

(b) We would like to show first that every $x \in L$ can be written uniquely as $x = \alpha x_0 + y$ for some $\alpha \in \mathbf{R}$ and $y \in L'$. In order to show this, note first that since L' has codimension 1, there exists $y_0 \in L$ such that every $x \in L$ can be written uniquely as $x = \beta y_0 + y$ where $\beta \in \mathbf{R}$ and $y \in L'$. Since $x_0 \in L$ we have that for some $\beta' \in \mathbf{R}$ and $y' \in L'$, $x_0 = \beta' y_0 + y'$ and $\beta' \neq 0$ since $x_0 \notin L'$. Therefore, $y_0 = \frac{1}{\beta'} x_0 - \frac{1}{\beta'} y'$, and any $x \in L$ can be written as

$$x = \beta y_0 + y = \frac{\beta}{\beta'} x_0 + \left(y - \frac{\beta}{\beta'} y' \right) = \alpha x_0 + y''$$

where $\alpha \in \mathbf{R}$ and $y'' \in L'$. It is easy to show that this representation is unique.

Define the linear functional f on L as follows. Given $x \in L$ write $x = \alpha x_0 + y$ with $y \in L'$ and let $f(x) = \alpha$. It is easy to verify that f is linear. Also, $x \in M'$ if and only if $x = x_0 + y$ for some $y \in L'$ if and only if $f(x) = 1$.

The functional f is unique for suppose that g were another such functional. Then since $x_0 \in M'$ and g is identically 1 on M', $g(x_0) = 1$. If $y \in L'$ then y has the unique representation as $y = \alpha x_0 + y'$ where $\alpha = 0$. Thus, $g(y) = 0$. Thus, for any $x \in L$, $g(x) = g(\alpha x_0 + y) = \alpha = f(x)$. ☐

Convex Functionals and The Hahn-Banach Theorem.

Definition 63. *Let L be a real vector space. The functional p defined on L is a* convex functional *if p satisfies*

$$p(\alpha x + \beta y) \leq \alpha p(x) + \beta p(y),$$

for all x, $y \in L$, and all real numbers $0 \leq \alpha$, $\beta \leq 1$ with $\alpha + \beta = 1$. Clearly any linear functional is also a convex functional. (Compare this definition with Kolmogorov, p. 130, Definition 3.)

Note that if $L = \mathbf{R}$ then a convex functional on L is just a function that is concave upward at every $x \in \mathbf{R}$.

Theorem 64. (Hahn-Banach Theorem.) *Suppose that f is a linear functional defined on a subspace L' of L which satisfies*

$$f(x) \leq p(x), \quad \forall x \in L'.$$

Then there is a linear functional F defined on L such that

1. *$F(x) = f(x)$ for all $x \in L'$ (that is, F is an* extension *of f), and*
2. *$F(x) \leq p(x)$ for all $x \in L$.*

Proof:
Claim 1. Given $z \in L \setminus L'$, we can find an extension $\widetilde{f}$ of f from L' to the subspace $\{\alpha z + y \colon y \in L'\}$.
Proof of Claim 1. Let $z \in L \setminus L'$. The extension $\widetilde{f}$ will be completely determined by the value of $\widetilde{f}(z)$ because,

$$\widetilde{f}(\alpha z + y) = \alpha \widetilde{f}(z) + \widetilde{f}(y) = \alpha \widetilde{f}(z) + f(y).$$

We must only ensure that $\widetilde{f}(x) \leq p(x)$ for all $x \in \{\alpha z + y \colon y \in L'\}$. That is, we must have for all $\alpha > 0$ and $y \in L'$,

$$\begin{aligned} \widetilde{f}(\alpha z + y) &= \alpha \widetilde{f}(z) + f(y) \leq p(y + \alpha z) \\ \widetilde{f}(-\alpha z + y) &= -\alpha \widetilde{f}(z) + f(y) \leq p(y - \alpha z). \end{aligned}$$

Rearranging, this becomes

$$\frac{1}{\alpha}(f(y) - p(y - \alpha z)) \leq \widetilde{f}(z) \leq \frac{1}{\alpha}(p(y + \alpha z) - f(y)),$$

for all $\alpha > 0$ and $y \in L'$.

Now, for any α, $\beta > 0$, and y_1, $y_2 \in L'$,

$$\begin{aligned} \beta f(y_1) + \alpha f(y_2) &= f(\beta y_1 + \alpha y_2) \\ &= (\alpha + \beta) f \left(\frac{\beta}{\alpha + \beta} y_1 + \frac{\alpha}{\alpha + \beta} y_2 \right) \end{aligned}$$

$$\begin{aligned}
&\leq (\alpha+\beta)p\left(\frac{\beta}{\alpha+\beta}y_1+\frac{\alpha}{\alpha+\beta}y_2\right)\\
&= (\alpha+\beta)p\left(\frac{\beta}{\alpha+\beta}(y_1-\alpha z)+\frac{\alpha}{\alpha+\beta}(y_2+\beta z)\right)\\
&\leq (\alpha+\beta)\left(\frac{\beta}{\alpha+\beta}p(y_1-\alpha z)+\frac{\alpha}{\alpha+\beta}p(y_2+\beta z)\right)\\
&= \beta p(y_1-\alpha z)+\alpha p(y_2+\beta z).
\end{aligned}$$

Therefore, for all α, $\beta > 0$ and y_1, $y_2 \in L'$,

$$\beta(f(y_1)-p(y_1-\alpha z)) \leq \alpha(p(y_2+\beta z)-f(y_2)),$$

or

$$\frac{1}{\alpha}(f(y_1)-p(y_1-\alpha z)) \leq \frac{1}{\beta}(p(y_2+\beta z)-f(y_2)).$$

Therefore,

$$\sup_{\alpha>0, y\in L'}\frac{1}{\alpha}(f(y_1)-p(y_1-\alpha z)) \leq \inf_{\alpha>0, y\in L'}\frac{1}{\alpha}(p(y_2+\alpha z)-f(y_2)).$$

Now, choose $\widetilde{f}(z)$ to be between these two numbers. Thus, Claim 1 is proved.

The proof of the existence of F required a Zorn's Lemma argument which we will not reproduce here. See Kolmogorov, p. 134 for this argument. ☐