7 Dual Spaces and the Hahn-Banach Theorem

Let $X$ be a vector space over $\mathbb{R}$. A map $f : X \to \mathbb{R}$ is a linear functional if $\forall x, y \in X$, $\forall a \in \mathbb{R}$ there holds $f(x + y) = f(x) + f(y)$ and $f(ax) = af(x)$.

The linear functionals on a vector space may themselves be regarded as elements of a vector space. Given two linear functionals $f_1, f_2$ on $X$, their sum is a functional on $X$ given by $(f_1 + f_2)(x) = f_1(x) + f_2(x)$. Similarly, one can define $(\alpha f)(x) = \alpha f(x)$ and $(f_1f_2)(x) = f_1(x)f_2(x)$. The space of linear functionals defined in this way is called the algebraic dual of $X$.

Consider a normed space $(X, \| \cdot \|)$. Its dual space

$$X^* = \{ f : X \to \mathbb{R} : \text{linear, continuous functional} \}$$

The norm of a functional $f$ can be defined in the following equivalent ways as:

$$\|f\| = \inf \{ M : |f(x)| \leq M\|x\|, \forall x \in X \}$$
$$= \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}$$
$$= \sup_{\|x\| \leq 1} |f(x)|$$
$$= \sup_{\|x\| = 1} |f(x)| \tag{2}$$

The norm defined as above satisfies the requirement of a norm as: $\|f\| > 0$; $\|f\| = 0$ if and only if $f = 0$; $\|\alpha f\| = |\alpha|\|f\|$ and $\|f_1 + f_2\| = \sup_{x \neq 0} \frac{|f_1(x) + f_2(x)|}{\|x\|} \leq \sup_{x \neq 0} \frac{|f_1(x)| + |f_2(x)|}{\|x\|} \leq \|f_1\| + \|f_2\|$

**Lemma 7.1** Let $T : X \to Y$ be a linear functional where $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$ are normed space. Then, $T$ is continuous everywhere if and only if it is continuous at some point $x_0 \in X$.

**Proof:** First, assume $T$ is continuous at $x_0 \in X$. Let $\{x_n\}$ be a sequence in $X$ converging to $x \in X$. By linearity of $T$, we have for $\tilde{x}_n = x_n - x + x_0$,

$$T(x_n) = T(\tilde{x}_n + x - x_0) = T(\tilde{x}_n) + T(x) - T(x_0)$$

But since $\tilde{x}_n \to x_0$ and $T$ is continuous at $x_0$, we have $T(\tilde{x}_n) \to T(x_0)$. Hence, $T(x_n) \to T(x)$. The other direction in the proof is straightforward. ■

**Theorem 7.1** $X^*$ is a Banach space

**Proof:** Since we already established that $X^*$ is a normed linear space, it remains only to show that $X^*$ is complete.

Let $(f_n)_{n \geq 1}$ be a Cauchy sequence in $X^*$, i.e., $\forall \varepsilon > 0$, there exists $n_0$ s.t. $\|f_m - f_n\| \leq \varepsilon$, for all $m, n \geq n_0$. We then have $|f_n(x) - f_m(x)| \leq \varepsilon\|x\|, \forall x \in X$. Thus, $(f_n(x))_{n \geq 1}$ is a Cauchy sequence of real number and that for each $x$, there is a $f(x) = \lim f_n(x)$. Next, we need to show $f \in X^*$:

First, $f$ is linear since

$$f(\alpha x + \beta y) = \lim f_n(\alpha x + \beta y) = \lim (\alpha f_n(x) + \beta f_n(y)) = \alpha f(x) + \beta f(y) \tag{3}$$
Second, since \((f_n)\) is Cauchy, given \(\varepsilon > 0\), there is a \(n_0\) such that \(|f_n(x) - f_m(x)| \leq \varepsilon\|x\|\) for all \(m, n \geq n_0\) and all \(x\). But since \(f_n(x) \to f(x)\), we then have

\[
|f(x)| = |f(x) - f_n(x) + f_n(x)| \leq |f(x) - f_n(x)| + |f_n(x)| \\
= \lim_{m \to \infty} |f_m(x) - f_n(x)| + |f_n(x)| \\
< \varepsilon\|x\| + |f_n(x)|\|x\| \\
= (\varepsilon + \|f_n\|)\|x\|
\]

and \(f\) is a bounded linear functional. Also from \(|f(x) - f_n(x)| \leq \varepsilon\|x\|, \forall n > n_0\), it follows that \(\|f - f_n\| \leq \varepsilon\). Hence, \(f\) is in \(X^*\).

It will be convenient to use the following notation: \(x^*\) for a generic element of \(X^*\) and \(\langle x^*, x \rangle\) for \(x^*(x)\) (the action of \(x^*\) on a generic element \(x\) of \(X\)). Note that this is not an inner product, since \(x^*\) and \(x\) belong to different spaces (\(X^*\) and \(X\), respectively); however, the map \((x^*, x) \mapsto \langle x^*, x \rangle\) is evidently linear in both of its arguments; we will refer to it as a pairing between \(X\) and its dual \(X^*\).

**Example 7.1 (The Dual of Hilbert Space)** Hilbert space is self-dual: i.e., \(X = \mathcal{H}, X^* = \mathcal{H}\) (Riesz representation theorem)

**Example 7.2 (The Dual of \(L^p\))** For every \(p > 1\), define \(q = p/(p - 1)\); if \(p = 1\), take \(q = \infty\). We will show the dual space of \(L^p\) is \(L^q\) in the following theorem:

**Theorem 7.2** \((L^p)^* = L^q\), where \(1/p + 1/q = 1\); Moreover, \(f \in (L^p)^*\) iff there exists \(y = (\eta_1, \eta_2, \ldots)\) such that

\[
\|f\| = \|y\|_{L^q} = \left\{ \begin{array}{ll}
\left( \sum_{i=1}^{\infty} |\eta_i|^q \right)^{1/q} & \text{if } 1 < p < \infty \\
\max_i |\eta_i| & \text{if } p = 1
\end{array} \right.
\]

**Proof:**

Define the element \(e_i \in L^p, i = 1, 2, \ldots\) as the sequence that is identically 0 except for a 1 in the \(i\)-th component. For any \(x = (\xi_1, \xi_2, \ldots) \in L^p\), we can write \(x = \sum_{i}^{\infty} \xi_i e_i\).

First, consider the case where \(1 < p < \infty\). Let \(X = L^p\) and \(f \in X^*\). Define \(y = (\eta_1, \eta_2, \ldots)\), where \(\eta_i = f(e_i)\), we have \(f(x) = \sum_{i=1}^{\infty} \eta_i \xi_i\).

For each \(n = 1, 2, \ldots\), define the vector \(x_n = (\xi_{1n}, \xi_{2n}, \ldots)\) having components

\[
\xi_{in} = \left\{ \begin{array}{ll}
|\eta_i|^{q/p} \text{sgn } \eta_i & i \leq n \\
0 & i > n
\end{array} \right.
\]

Then

\[
\|x_n\|_{L^p} = \left( \sum_{i=1}^{n} |\eta_i|^q \right)^{1/p}
\]
and since $|f(x_n)| \leq \|f\| \|x_n\|$, it follows that:

$$f(x_n) = \sum_{i=1}^{n} |\eta_i|^{(q/p)+1} = \sum_{i=1}^{n} |\eta_i|^q$$

We then have $\sum_{i=1}^{n} |\eta_i|^q = |f(x_n)| \leq \|f\|\|x_n\|_p = \|f\|\left(\sum_{i=1}^{n} |\eta_i|^q\right)^{1/q}$. Hence, $\left(\sum_{i=1}^{n} |\eta_i|^q\right)^{1/q} \leq \|f\|$ for all $n$. Hence $y$ is an element of $l^q$ and $\|y\| \leq \|f\|$.

For the other direction, suppose $y = (\eta_1, \eta_2, \ldots)$ is an element of $l^q$. If $x = (\xi_1, \xi_2, \ldots) \in l^p$, then by Hölder’s inequality we have $|f(x)| \leq \sum_{i=1}^{\infty} |\xi_i|\eta_i \leq \|x\|_p\|y\|_q$, thus $f(x) = \sum_{i=1}^{\infty} \xi_i\eta_i$ is a bounded linear functional on $l^p$ and that $\|f\| \leq \|y\|_q$. Since $f(e_i) = \eta_i$, which implies $\|y\|_q \leq \|f\|$, we conclude that $\|f\| = \|y\|_q$.

Finally, for the case where $p = 1, q = \infty$, we can define $x_n$ by

$$\xi_i = \begin{cases} 0 & i \neq n \\ \text{sgn} \eta_i & i = n \end{cases}$$

Then $\|x_n\| \leq 1$ and

$$|\eta_n| = f(x_n) \leq \|f\| \|x_n\| \leq \|f\|$$

Thus, $\|y\|_{\infty} \leq \|f\|$. Conversely, if $y = (\eta_1, \eta_2, \ldots) \in l^\infty$, following the same line of reasoning as in the previous case, we have $\|y\|_{\infty} \leq \|f\|$ and thus $\|f\| = \|y\|_{\infty}$.

**Remark 7.1** Note that the dual of $(l^\infty)^*$ is not $l^1$. In fact, $(l^\infty)^*$ is much larger than $l^1$.

**Example 7.3** The dual of $L^p[0, 1]$ is $L^q$, where $1/p + 1/q = 1$ for $1 \leq p < \infty$.

### 7.1 Hahn-Banach Theorem

**Definition 7.1** A real valued function $p : X \rightarrow \mathbb{R}$ is a sublinear functional on $X$ if:

$$p(x + y) \leq p(x) + p(y), \quad x, y \in X$$

$$p(\alpha x) = \alpha p(x), \quad \forall \alpha \geq 0, x \in X$$

**Theorem 7.3 (Hahn-Banach Theorem, Extension Form)** Let $L$ be a subspace of a vector space $X$, and let $F : L \rightarrow \mathbb{R}$ be a linear functional such that $F(x) \leq p(x), \forall x \in L$. Then, there exists an extension $f$ of $F$ from $L$ to $X$ which is linear and $f(x) \leq p(x)$ on $X$.

**Proof:** (a special case)

Assume a nested sequence of subspaces $L_1 \subseteq L_2 \ldots$ s.t: $L_1 = L, L_{n+1} = L_n + \operatorname{span}\{v_n\}, v_n \notin L_n \equiv \{u + \lambda v_n : u \in L_n, \lambda \in \mathbb{R}\}$ and $X = \bigcup_{n=1}^{\infty} L_n$.

(By induction) $X = L + \operatorname{span}\{v\}, v \notin L, x = u + \lambda v, u \in L, \lambda \in \mathbb{R}$. Define $f(x) := F(u) + \lambda c$ where $c$ is a constant. Next, we show that $c$ can be chosen as follows:
For any two elements $u, u' \in L$, we have

$$F(u) + F(u') = F(u + u') = p(u + u') = p(u - v + v - u') \leq p(u - v) + p(u' + v)$$

$$\implies F(u) - p(u - v) \leq p(u' + v) - F(u'), \quad \forall u, u'$$

$$\implies \exists c : \sup_{u \in L} \{ F(u) - p(u - v) \} \leq c \leq \inf_{u \in L} \{ p(u' + v) - F(u') \}$$

To complete the proof, we must show that $f(u + \lambda v = F(u) + \lambda c \leq p(u + \lambda c), \forall u \in L, \lambda \in \mathbb{R}$. Since $\lambda \neq 0$, first assume $\lambda > 0$. Then we have

$$F(u) + \lambda c \leq F(u) + \lambda (p(u' + v) - F(u')) = F(u) + p(\lambda u' + v) - F(\lambda u'), \quad \forall u' \in L$$

Choosing $u' = u/\lambda$ gives:

$$F(u) + \lambda c \leq F(u) + p(u + \lambda v) - F(u) \equiv p(u + \lambda v)$$

Similarly, if $\lambda < 0$, we have $F(u) + \lambda c \leq p(u + \lambda v)$. By induction, the functional $F$ can be extended from $L_1$ to $L_1 + \text{span}(v_1)$ then to $\{ L_1 + \text{span}(v_1) \} + \text{span}(v_2)$ and so on.

Theorem 7.4 (Hahn–Banach Theorem for normed spaces) Let $L$ be a subspace of a separable normed space $(X, \| \cdot \|)$. Let $F : L \to \mathbb{R}$ be linear s.t. $|F(x)| \leq \alpha \|x\|, x \in L$ for some $\alpha \geq 0$. Then, there exists $f \in X^*$ s.t., $f = F$ on $L$ and $\|f\| \leq \alpha$.

Proof: Let $\{ y_1, y_2, \ldots \}$ be a basis of $X$. $X^o = \text{span}\{ y_1, y_2, \ldots \}$. Take $p(x) = \alpha \|x\|$ in the Hahn-Banach Extension Theorem, $F(x) \leq p(x)$ on $L$. Then we can extend $F$ to $f$ on $X^o$ and $|f(x)| \leq \alpha \|x\|, x \in X^o$. Since $X^o$ is dense in $X$, we can extend $F$ to all of $X$ by continuity.

Corollary 7.1 For any $x_0 \neq 0$ an element of $X$, there exists a $f \in X^*$ s.t. $f(x_0) = \|x_0\|$ and $\|f\| = 1$. Therefore, for any $x \in X$,

$$\|x\| = \sup\{ f(x) : f \in X^*, \|f\| = 1 \}.$$

Proof: We construct a subspace $L \subset X$ containing the element $x_0$ as $L = \text{span}\{ x_0 \}$. We can also define a linear functional $F \in L^*$ s.t. $F(\lambda x_0) = \lambda \|x_0\|$ and $|F(\lambda x_0)| = \| \lambda \| x_0 \| = \| \lambda x_0 \|$. Then the previous Theorem implies the existence of $f \in X^*$.

7.2 Hyperplanes, half-spaces, separation of sets

In this section, we will introduce hyperplane, half-spaces and a theorem about the separation of a point and subspace. We first start with some definitions.

Definition 7.2 A hyperplane $H$ in a Hilbert space $\mathcal{H}$ is defined as follows:

$$H := \{ x \in \mathcal{H} : \langle x, h \rangle = c \},$$

for some $h \in \mathcal{H}$ and $c \in \mathbb{R}$. Similarly, the hyperplane $H$ in a normed space $X$ is defined as follows:

$$H := \{ x \in X : \langle x^*, x \rangle = c \},$$

for a fixed $x^* \in X^*$ and $c \in \mathbb{R}$. 
If $c = 0$, then $H$ is a subspace of $\mathcal{H}$ or $X$ with codimension equals to 1.

**Definition 7.3** A **half-space** in $H \leq$ in a normed space $X$ is defined as follows:

\[ H \leq := \{ x \in X : \langle x^*, x \rangle \leq c \}, \]

for some $x^* \in X^*$ and $c \in \mathbb{R}$. Similarly, half-space $H >$ is defined as:

\[ H > := \{ x \in X : \langle x^*, x \rangle > c \}, \]

By definition, it is obvious that $X = H \leq \cup H >$ and $H \leq \cap H > = \emptyset$.

### 7.3 Separating Point from Subspace

Consider $A, B \subset X$, and $H \subset X$ is a hyperplane. We say $H$ strictly separates $A$ and $B$ if $A \subset H \leq$ and $B \subset H >$.

**Theorem 7.5** (Separating a point from a subspace) Let $L$ be a subspace of $(X, \| \cdot \|)$, and $x_0 \in X$ be given such that $\text{dist}(x_0, L) = \inf_{x \in L} \| x_0 - x \| > 0$. Then there exists $f^* \in X^*$ such that the following holds simultaneously:

1. $\| f^* \| = 1$,
2. $f^*(x_0) = \text{dist}(x_0, L) > 0$,
3. if $x \in L$, then $f^*(x) = 0$.

**Proof:** Consider the linear subspace $M = L + \text{span}\{x_0\} = \{ x + \lambda x_0, x \in L, \lambda \in \mathbb{R} \}$. Define a linear functional $F : M \to \mathbb{R}$:

\[ F(x + \lambda x_0) = \lambda \text{dist}(x, L). \]

Clearly, if $x \in L$, then $F(x) = 0$. $F(x_0) = 1$. Next we show that $\forall y \in M$, we have $|F(y)| \leq \|y\|$.

Indeed, we have, for $\lambda \neq 0$,

\[
F(x + \lambda x_0) \leq |\lambda| \text{dist}(x, L)
\begin{align*}
&\leq |\lambda| \| x_0 - x' \| \quad \text{(for some } x' \in L) \\
&= |\lambda| \| x_0 - (-\frac{x}{\lambda}) \| \quad (x' = -\frac{x}{\lambda} \in L) \\
&= \| \lambda x_0 + x \|.
\end{align*}
\]

Now, by the Hahn–Banach Theorem, we can extend $F$ to $f \in X^*$ such that $f = F$ on $M$ and $\| f \| \leq 1$. To show $\| f \| = 1$, noticing that $f(x_0) = \text{dist}(x_0, L) = \inf_{x \in L} \| x_0 - x \|$. For any $\epsilon > 0$, $\exists x \in L$ s.t. $\| x_0 - x \| < \text{dist}(x_0, L) + \epsilon$. Therefore, we have

\[
\frac{f(x_0 - x)}{\| x_0 - x \|} = \frac{\text{dist}(x_0, L)}{\| x_0 - x \|} > \frac{\text{dist}(x_0, L)}{\text{dist}(x_0, L) + \epsilon}.
\]

Since $\epsilon > 0$ is arbitrary, we can let $\epsilon \to 0$ to get $\| f \| \geq 1$. Hence $\| f \| = 1$. \(\blacksquare\)
7.4 Separating Point from Closed Convex Set

Consider a normed space \((X, \| \cdot \|)\), let \(M\) be a closed convex subset of \(X\) and a point \(x_0 \notin M\). Then we know that \(\text{dist}(x_0, M) > 0\). Now we want to construct \(f \in X^*\), such that \(f(x_0) > 1\) and \(f(x) \leq 1, \forall x \in M\).

If \(M\) is a unit ball, then we can directly apply the Hahn–Banach theorem to construct such \(f\). For general \(M\), we will need some preliminaries first. We assume that a subset \(K\) of \(X\) has the following properties:

1. Convex;
2. Non empty interior (\(\text{int}(K) \neq \emptyset\));
3. \(0 \in \text{int}(K)\).

**Definition 7.4** The Minkowski functional of the set \(K\) satisfying the above conditions is defined as:

\[
p_K : X \to [0, \infty], \quad p_K(x) = \inf\{r > 0, \frac{x}{r} \in K\}.
\]

Before we introduce the properties of Minkowski functional, let us first look at a motivating example. Consider a normed vector space \((X, \| \cdot \|)\), with the norm \(\| \cdot \|\) and let \(K\) be the unit ball in \(X\), i.e., \(K = \{x : \|x\| \leq 1\}\). Then for every \(x \in X\), \(p_K(x) = \|x\|\). Thus the Minkowski functional \(p_K\) is just the norm on \(X\).

**Proposition 7.1** The functional \(p_K\) has the following properties:

1. \(\exists c > 0 \text{ s.t. } 0 \leq p_K(x) \leq c\|x\|, \quad \forall x \in X\).
2. \(p_K(\alpha x) = \alpha p_K(x), \quad \forall x \in X, \alpha \geq 0\).
3. \(p_K(x + y) \leq p_K(x) + p_K(y), \quad \forall x, y \in X\).
4. \(p_K(x)\) is continuous.

**Proof:** We prove point by point as follows:

1. By definition, we have \(p_K(x) \geq 0 \text{ and } p_K(0) = 0\) holds. For the boundedness, since \(\text{int}(K) \neq \emptyset\), \(\exists r > 0 \text{ s.t. } \|x\| \leq r \text{ implies } x \in K\). For any \(x \in X\), consider the mapping from \(x\) to \(\bar{x}\):
   \[
x \to \bar{x} = \frac{rx}{\|x\|}, \text{ then } \|\bar{x}\| = r \to \bar{x} \in K, \text{ hence } p_K(x) \leq \frac{\|x\|}{r}.
\]
2. Fix \(x \in K\), clearly \(p_K(\alpha x) = \alpha p_K(x)\) holds for \(\alpha = 0\) as \(p_K(0) = 0\). For \(\alpha > 0\). We have
   \[
p_K(\alpha x) = \inf\{r > 0, \frac{\alpha x}{r} \in K\} = \alpha \inf\{r > 0, \frac{x}{r/\alpha} \in K\} = \alpha p_K(x).
\]

6
3. By definition, $\forall \epsilon > 0, \exists \alpha, \beta > 0$ s.t.

$$p_K(x) < \alpha < p_K(x) + \epsilon,$$
$$p_K(y) < \beta < p_K(y) + \epsilon.$$ 

Define $\gamma := \alpha + \beta$. Then we have $\frac{\alpha}{\gamma} + \frac{\beta}{\gamma} = 1$. In addition, $\frac{x+y}{\gamma} = \frac{\alpha}{\gamma} x + \frac{\beta}{\gamma} y \in K$ by convexity. Therefore $p_K(x+y) \leq \gamma < p_K(x) + \epsilon + p_K(y) + \epsilon = p_K(x) + p_K(y) + 2\epsilon$. Since $\epsilon > 0$ is arbitrary, $p_K(x+y) \leq p_K(x) + p_K(y)$.

4. $\forall x, y \in K$, we have $p_K(x) = p_K(x+y-y) \leq p_K(x-y) + p_K(y) \leq c\|x-y\| + p_K(y)$. Switching $x$ and $y$, we also have $p_K(y) \leq c\|x-y\| + p_K(x)$. Therefore, $|p_K(x) - p_K(y)| \leq c\|x-y\|$. We show that $p_K(x)$ is in fact Lipschitz continuous.

Now we are ready to state the following theorem:

**Theorem 7.6 (Separating a point from a closed convex set)** Let $M$ be a closed convex set in $(X, \| \cdot \|)$. Then $\forall x_0 \notin M, \exists x^*_0 \in X^*$ s.t.

1. $\langle x^*_0, x_0 \rangle > 1$.

2. $\langle x^*_0, x \rangle \leq 1$, for $x \in M$.

**Proof:** Assume $0 \in M$ without loss of generality. If not, we just need to shift everything such that $0 \in M$. Define $d := \text{dist}(x_0, M) > 0$, $M_r := \{ x \in X, \text{dist}(x, M) < r \}$, and $K := \text{closure}(M_{d/2})$. By such construction, we have $x_0 \notin K$, $M \subset K$, $K$ is convex, and the interior of $K$ is nonempty. Now let $p_K$ be the Minkowski functional of $K$, $L := \text{span}\{x_0\}$. Consider $F(\lambda x_0) = \lambda p_K(x_0)$, then

$$\lambda > 0 : F(\lambda x_0) = \lambda p_K(x_0) = p_K(\lambda x_0)$$
$$\lambda < 0 : F(\lambda x_0) \geq 0 \leq p_K(\lambda x_0)$$

Therefore, we have $F(y) \leq p_K(y)$ for $y \in L$. By applying the Hahn–Banach Theorem, we can extend $F$ to $f : X \to \mathbb{R}$. We have $|f(x)| \leq p_K(x) \leq c\|x\|$, $f$ is linear continuous. In addition, we have:

$$f(x_0) = p_K(x_0) > 1 \text{ (since } x_0 \notin K),$$
$$f(x) \leq 1 \text{ for } x \in K.$$