6 Separation Theorems

6.1 Linear (Vector) Spaces

Definition 1 *A Set L is called a linear (vector) space if:*

1. *For any pair $(x, y) \in L$ the sum of x and y denoted $(x + y) \in L$ is uniquely defined and satisfies:*

   (a) $x + y = y + x$ *(communativity)*

   (b) $(x + y) + z = x + (y + z)$ *(associativity)*

   (c) *there exists an element $0 \in L$ such that $x + 0 = x$ for every $x \in L$*

   (d) *for every $x \in L$ there exists a negative of x denoted $-x \in L$ with the property that* $x + (-x) = 0$

2. *Any $x \in L$ and number uniquely defines an element $\alpha x \in L$ called the product of α and x, such that:*

   (a) $\alpha (\beta x) = (\alpha\beta) x$

   (b) $1x = x$

3. *Addition and multiplication obeys distributive laws*

   (a) $(\alpha + \beta) x = \alpha x + \beta x$

   (b) $\alpha (x + y) = \alpha x + \alpha y$

Remark 1 *α is usually called a scalar. If $\alpha \in R$ we say that L is a real linear space. If α is complex we say that L is a complex linear space.*

Example 1 *R (with the arithmetic operations of addition and multiplication)*

Example 2 R^n *with*

$$\begin{aligned}(x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n) &= (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \\ \alpha (x_1, x_2, ..., x_n) &= (\alpha x_1, \alpha x_2, ..., \alpha x_n)\end{aligned}$$

Example 3 $\Phi = \{f : [0,1] \to R\}$ *together with operations*

$$\begin{aligned}[f+g](x) &= f(x) + g(x) \\ [\lambda f](x) &= \lambda f(x) \\ 0(x) &= 0 \\ [-f](x) &= -f(x)\end{aligned}$$

Example 4 *Hilbert Space,* l^2*, the set of sequences* $\langle x_k \rangle_1^\infty$ *such that*

$$\sum_1^\infty x_k^2 < \infty$$

equipped with operations

$$\begin{aligned}(x_1, x_2, ..., x_k, ...) + (y_1, y_2, ..., y_k, ...) &= (x_1 + y_1, x_2 + y_2, ..., x_k + y_k......) \\ \alpha (x_1, x_2, ..., x_k, ...) &= (\alpha x_1, \alpha x_2, ..., \alpha x_k, ...) .\end{aligned}$$

Here we note that

$$\begin{aligned}& 2\left(x_k^2 + y_k^2\right) - (x_k + y_k)^2 \\ =\ & 2\left(x_k^2 + y_k^2\right) - \left(x_k^2 + 2x_k y_k + y_k^2\right) \\ =\ & \left(x_k^2 + y_k^2\right) - 2x_k y_k = (x_k - y_k)^2 \geq 0,\end{aligned}$$

so

$$\sum_1^\infty (x_k + y_k)^2 \leq \sum_1^\infty 2\left(x_k^2 + y_k^2\right) < \infty.$$

6.2 Linear Functions, Functionals and Hyperplanes

Definition 2 *Given vector spaces X and Y we say that $f : X \to Y$ is linear if for all $x, x' \in X$*

$$\begin{aligned} f(x) + f(x') &= f(x + x') \\ \lambda f(x) &= f(\lambda x) \end{aligned}$$

Definition 3 *If X is a vector space and $f : X \to R$ is a linear function we say that f is a linear functional.*

Example 5 *$I : X \to R$ where for every $x \in X = \{g : [a, b] \to R\}$ the value is defined by the integral $I(x) = \int_a^b x(t) f(t) \, dt$ where the continouus function f is held fixed. Clearly, I is a linear functional.*

Example 6 *$a : R^n \to R$ defined by usual scalar multiplication,*

$$ax = \sum_{i=1}^{n} a_i x_i$$

is obviously also a linear functional.

Definition 4 *Let L be a linear space and $f : L \to R$ is a linear functional. Then, the set $N \subset L$ defined as*

$$N = \{x \in L | f(x) = 0\}$$

is called the nullspace of the functional f.

Remark 2 *N is a subspace of L since if $(x, y) \in N$ then*

$$f(x + y) = f(x) + f(y) = 0 + 0 = 0.$$

Definition 5 (Sometimes taken as a result) *Let $f : L \to R$ be a linear functional satisfying $f(x) \neq 0$ for some ($\Rightarrow$infinitely many) $x \in L$ and let k be a real number. Then, the set*

$$H = \{x \in L | f(x) = k\}$$

is called a hyperplane in L with normal f.

6.3 Convex Sets

Definition 6 *A set $S \subset L$ is said to be convex if $\lambda x + (1-\lambda) y \in S$ for every $x, y \in S$ and $\lambda \in [0,1]$.*

Example 7 *A hyperplane $H = \{x \in L | f(x) = k\}$ is convex since given any $x, y \in H$ we have that using the two defining properties of linearity*

$$\begin{aligned} f(\lambda x + (1-\lambda) y) &= f(\lambda x) + f((1-\lambda) y) \\ &= \lambda f(x) + (1-\lambda) f(y) \\ &= \lambda k + (1-\lambda) k = k \end{aligned}$$

Example 8 *The closed upper halfspace $S_+ = \{x \in L | f(x) \geq k\}$ is convex since given any $x, y \in S_+$ we have that using the two defining properties of linearity*

$$\begin{aligned} f(\lambda x + (1-\lambda) y) &= f(\lambda x) + f((1-\lambda) y) \\ &= \lambda f(x) + (1-\lambda) f(y) \\ &\geq \lambda k + (1-\lambda) k \geq k \end{aligned}$$

Obviously, an identical argument works for the lower halfspace as well.

Proposition 1 *The interior of a convex set is convex.*

Proof. Suppose that S is convex. If the interior is empty, convexity is trivial. If the interior is non-empty, take any x, y in the interior and let $\lambda \in [0,1]$ be arbitrary. Let $a \in L$ be arbitrary and note that, since x is in the interior of S there exists $\varepsilon_1 > 0$ such that

$$x + t_1 a \in S$$

if $|t_1| < \varepsilon_1$ and $\varepsilon_2 > 0$ such that

$$y + t_2 a \in S$$

if $|t_2| < \varepsilon_2$. Then, by convexity of S

$$\lambda (x + ta) + (1-\lambda)(y + ta) = \lambda x + (1-\lambda) y + ta \in S$$

for every $t < \min\{\varepsilon_1, \varepsilon_2\}$, implying that $\lambda x + (1-\lambda) y$ is in the interior of S. ■

Proposition 2 *Let $\{S_\alpha\}_{\alpha \in A}$ be a collection of convex sets. Then $\cap_{\alpha \in A} S_\alpha$ is convex.*

Proof. Pick $\lambda \in [0,1]$ arbitrarily and suppose that $x, y \in \cap_{\alpha \in A} S_\alpha$. Then $x, y \in S_\alpha$ for every $\alpha \in A$. Since each S_α is convex it follows that

$$\lambda x + (1-\lambda) y \in S_\alpha$$

for every $\alpha \in A$. Hence, $\lambda x + (1-\lambda) y \in \cap_{\alpha \in A} S_\alpha$. ■

Definition 7 *Let S and T be non empty sets in a linear space L and let $f : L \to R$ be a linear functional. The hyperplane $H = \{x \in L | f(x) = k\}$ is said to separate S and T is:*

1. *$f(x) \leq k$ for every $x \in S$*

2. *$f(x) \geq k$ for every $x \in T$.*

If both inequalities are strict we say that H strictly separates the sets.

Remark 3 *If H separates S and T we have that*

$$f(x) \leq k \leq f(y)$$

for every $(x, y) \in S \times T$. Consider the equivalent hyperplane

$$H' = \{x \in L | -f(x) = -k\}$$

then

$$-f(x) \geq k \geq -f(y)$$

for every $(x, y) \in S \times T$. Hence, the direction of the inequalities are irrelevant.

Example 9 *In consumer theory, the budget set is usually given by*

$$B = \left\{x \in R^n_+ | px \leq m\right\}$$

Let S_- be the closed halfspace below the hyperplane defined by normal p and constant m and observe

$$B = S_- \cap R^n_+.$$

It is obvious that R_+^n is convex (convex combination of two positive numbers is positive). We know that S_- is convex. Hence, as the intersection of convex sets is convex, the budget set is convex.

6.4 Separation Results in Euclidean Space

Proposition 3 (Simplest Separation Theorem) *Let $S \subset R^n$ be a closed convex set and suppose that $y \notin S$. Then there exists a hyperplane $H = \{x \in R^n | ax = k\}$ that separates S and $\{y\}$ strictly, eg $ax < k < ay$ for every $x \in S$.*

Proof. Let

$$z \in \arg\min_{x \in S} \|x - y\|.$$

We note that if $\widehat{x} \in S$ is picked arbitrarily, then it has to be that

$$\begin{aligned} & \min_{x \in S} \|x - y\| \\ = & \min_{x} \|x - y\| \\ \text{s.t. } x \in & S \\ x \in & \overline{B(y, \|\widehat{x} - y\|)}. \end{aligned}$$

That is, we may "compactify" the feasible set. Hence, since $z \in S \cap \overline{B(y, \|\widehat{x} - y\|)}$ where S and $B(y, \|\widehat{x} - y\|)$ are closed (⇒intersection closed) and $B(y, \|\widehat{x} - y\|)$ is bounded and since the Euclidean norm is continuous it follows that z is well defined by Weierstrass maximum theorem. Let

$$\begin{aligned} a &= y - z \\ k &= \frac{1}{2}(az + ay) = \frac{1}{2}((y - z) z + (y - z) y) \\ &= \frac{1}{2}(y \cdot y - z \cdot z) \end{aligned}$$

and note that

$$\begin{aligned} ay &= (y-z)\,y = \frac{1}{2}\left((y-z)\,y + (y-z)\,y\right) \\ &= \frac{1}{2}\left((y-z)\,(y+z-z) + (y-z)\,y\right) \\ &= \frac{1}{2}\left((y-z)\,z + (y-z)\,y\right) + \frac{1}{2}(y-z)\,(y-z) \\ &= k + \frac{1}{2}(y-z)\,(y-z) > k \end{aligned}$$

Also

$$\begin{aligned} az &= (y-z)\,z = \frac{1}{2}\left((y-z)\,z + (y-z)\,z\right) \\ &= \frac{1}{2}\left((y-z)\,(z+y-y) + (y-z)\,z\right) \\ &= \frac{1}{2}\left((y-z)\,z + (y-z)\,z\right) + \frac{1}{2}\left((y-z)\,(z-y)\right) \\ &= k - \frac{1}{2}(y-z)\,(y-z) < k \end{aligned}$$

Now, pick any $x \in S$. By convexity of S it follows that

$$(1-\lambda)\,z + \lambda x \in S$$

for any $\lambda \in [0,1]$. Given any λ we have that (because z minimizes the distance to y)

$$\begin{aligned} \|z-y\|^2 &\leq \|(1-\lambda)\,z + \lambda x - y\|^2 = \|(1-\lambda)\,(z-y) + \lambda\,(x-y)\|^2 \\ &= \sum_{i=1}^{n}\left((1-\lambda)\,(z_i-y_i) + \lambda\,(x_i-y_i)\right)^2 \\ &= \sum_{i=1}^{n}\left[(1-\lambda)\,(z_i-y_i)\right]^2 + 2\,(1-\lambda)\,(z_i-y_i)\,\lambda\,(x_i-y_i) + \left[\lambda\,(x_i-y_i)\right]^2 \\ &= (1-\lambda)^2\sum_{i=1}^{n}(z_i-y_i)^2 + 2\,(1-\lambda)\,\lambda\sum_{i=1}^{n}(z_i-y_i)\,\lambda\,(x_i-y_i) + \lambda^2\sum_{i=1}^{n}(x_i-y_i)^2 \\ &= (1-\lambda)^2\,\|z-y\|^2 + 2\,(1-\lambda)\,\lambda\,(z-y)\cdot(x-y) + \lambda^2\,\|x-y\|^2 \end{aligned}$$

Hence

$$\begin{aligned} 0 &\leq \lambda\,(\lambda-2)\,\|z-y\|^2 + 2\,(1-\lambda)\,\lambda\,(z-y)\cdot(x-y) + \lambda^2\,\|x-y\|^2 \\ &\Rightarrow \text{(if } \lambda > 0\text{)} \\ 0 &\leq (\lambda-2)\,\|z-y\|^2 + 2\,(1-\lambda)\,(z-y)\cdot(x-y) + \lambda\,\|x-y\|^2 \end{aligned}$$

which is true for every $\lambda \in (0,1)$. Taking the limit as $\lambda \to 0$

$$\begin{aligned}
0 &\leq -2\|z-y\|^2 + 2(z-y)\cdot(x-y) \\
&\Leftrightarrow \\
0 &\leq (z-y)(x-y) - (z-y)(z-y) \\
&= (z-y)x - (z-y)z \\
&= (y-z)z - (y-z)x = az - ax \\
&\Leftrightarrow \\
ax &\leq az < k < ay,
\end{aligned}$$

which completes the proof. ■

In the first part of the result above we made active use of the assumption that S was closed. To deal with any convex sets we define a boundary point:

Definition 8 *We say that x is a boundary point of S if every open ball $B(x,\varepsilon)$ contains at least one point $y \in S$ and one point $z \notin S$.*

We note that the boundary points of a set is a subset of the closure. We can now show:

Proposition 4 (Simple Supporting Hyperplane Theorem) *Let $S \subset R^n$ be convex and let y be a boundary point of S. Then, there exists a supporting hyperplane $H = \{x|ax = k\}$, meaning that $ax \leq ay$ for every $x \in S$.*

Proof. Let $\overline{S}$ be the closure of S. We'll take as granted that $\overline{S}$ is convex (homework). Moreover, y is a boundary point of $\overline{S}$. We can therefore construct a sequence $\langle y_n \rangle_1^\infty$ such that $y_n \notin \overline{S}$ for every k and $y_n \to y$. By application of the simple separation theorem there exists a hyperplane

$$H_n = \{x|a_n x = k_n\}$$

such that

$$a_n x < k_n < a_n y_n$$

for every n. Without loss of generality, let $\|a_n\| \leq 1$ for every n. Since $\{a| \|a\| \leq 1\}$ is compact we know that there exists a convergent subsequence $\langle a_{n_k} \rangle$ of $\langle a_n \rangle$. Let a be the limit of the subsequence and note that $\langle y_{n_k} \rangle$ converges to y. Hence, for every $x \in \overline{S}$

$$\begin{aligned} a_{n_k} x &< a_{n_k} y_{n_k} \\ &\Rightarrow \\ ax &= \lim_{n_k \to \infty} a_{n_k} x \leq \lim_{n_k \to \infty} a_{n_k} y_{n_k} = ay. \end{aligned}$$

The argument is completed by picking $k = ay$ and noting that $S \subset \overline{S}$. ■

Definition 9 *Let S, T be convex sets in a linear space L, then a linear combination of S and T is given by*

$$aS + bT = \{z = L | \textit{such that there exists } x \in S, y \in T \textit{ with } z = \alpha x + by\}$$

Lemma 1 *If S, T are convex, then $aS + bT$ is convex.*

Proof. Consider any pair $x', x'' \in X = aS + bT$. Since $x' \in X$ there exists $s' \in S$ and $t' \in T$ such that

$$x' = as' + bt'$$

and, by the same argument, there is $s'' \in S$ and $t'' \in T$ such that

$$x'' = as'' + bt''.$$

Take any λ and note that

$$\begin{aligned} (1-\lambda) x' + \lambda x'' &= a [(1-\lambda) s' + \lambda s''] \\ &\quad + b [(1-\lambda) t' + \lambda t''] \end{aligned}$$

Since S is convex $(1-\lambda) s' + \lambda s'' \in S$ and since T is convex $(1-\lambda) t' + \lambda t'' \in T$. Hence, $(1-\lambda) x' + \lambda x'' \in aS + bT$. ■

Theorem 1 (Minkowski) *Let S and T be convex sets in R^n with $S \cap T = \varnothing$. Then there exists a hyperplane H that separates S and T.*

Proof. Consider the set $X = S - T$. Suppose $x, y \in X$, which is convex by the argument above. Moreover, the closure $\overline{X}$ is convex and

$$0 \notin X = S - T,$$

as the two sets are disjoint.

CASE 1: $0 \notin \overline{X}$. The simple separation theorem can then be used to conclude that there exists $H = \{x|ax = k\}$ such that

$$ax < k < a0 = 0.$$

for every $x \in X$.

CASE 2: : $0 \in \overline{X}$. The simple supporting hyperplane can then be used to conclude that there exists $H = \{x|ax = k\}$ such that

$$ax \leq a0 = 0.$$

Pick $s \in S$ and $t \in T$ and note that (combining the two cases) that $(s - t) \in X$, so

$$a(s - t) \leq 0 \Leftrightarrow as \leq at.$$

Since this is true for all $s \in S$ and $t \in T$ we can pick k so that

$$as \leq k \leq at.$$

■

Example 10 *Consider a pure exchange economy* $\left\{I, \{u_i\}_{i\in I}, \{\omega\}_{i\in I}\right\}$ *where* $u_i : X \to R$ *and* $X \subset R^n_+$ *(consumption space) and* $\omega \in R^n_{++}$ *is the aggregate endowment.*

Definition 10 *Consider a map* $f : X \to R$ *where* X *is a convex set. Then, we say that* f *is quasi-concave if*

$$f(\lambda x' + (1 - \lambda) x'') \geq \min\{f(x'), f(x'')\}$$

for every $(x', x'') \in X$ *and* $\lambda \in [0, 1]$.

We note that q-concavity implies that the upper countor set

$$\{x \in X | f(x) \geq f(x^*)\}$$

is a convex set, as otherwise there would exist some x', x'' and λ such that;

1. $x' \in \{x \in X | f(x) \geq f(x^*)\} \Rightarrow f(x') \geq f(x^*)$
2. $x'' \in \{x \in X | f(x) \geq f(x^*)\} \Rightarrow f(x'') \geq f(x^*)$
3. $x^\lambda = \lambda x' + (1-\lambda) x'' \notin \{x \in X | f(x) \geq f(x^*)\} \Rightarrow f\left(x^\lambda\right) < f(x^*) \Rightarrow$

$$f\left(x^\lambda\right) < f(x^*) \leq \min\{f(x'), f(x'')\},$$

violating q-concavity.

Definition 11 *A competitive equilibrium in the pure exchange economy is a price vector $p \in R^n_+$ and an allocation x^* such that:*

1. $x_i^* \in \arg\max_{x \in X} u_i(x)$ *subject to* $x \in B(p, w_i)$
2. $\sum x_i^* = \sum_i w_i$

Proposition 5 *Suppose that u_i is quasi-concave for every $i \in I$ and that for every $x \in X$ and $\varepsilon > 0$ there exists some $y \in B(x, \varepsilon)$ such that $u_i(y) > u_i(x)$. Then, for every Pareto optimal allocation x^O there exists a distribution of property rights $(\omega_1, ..., \omega_n)$ and a price vector $p \neq 0$ such that $\left(p, x^O\right)$ is a competitive equilibrium in economy $\left\{I, \{u_i\}_{i \in I}, \{\omega\}_{i \in I}\right\}$.*

Proof. *Let $\left\{x^O\right\}$ be a Pareto optimal allocation. Define*

$$U_i\left(x_i^O\right) = \left\{x \in R^n_+ | u_i(x) \geq u_i\left(x_i^0\right)\right\},$$

which is convex by the assumption of quasi-concave utility functions. Also, let

$$U\left(x^O\right) = \sum_{i=1}^{I} U_i\left(x_i^O\right) = \left\{x \in R^n_+ | \text{exists } (x_1, ..., x_n) \in R^{n \times I}_+ | u_i(x_i) \geq u_i\left(x_i^0\right)\right\}.$$

Since $U\left(x^O\right)$ *is a linear combination of convex sets it is convex. Moreover, by local nonsatiation we have that the point* $\sum_i x_i^O$ *is a boundary point of* $U\left(x^O\right)$ *(if interior we can make everyone strictly better off which violates Pareto optimality). Hence, we may use the supporting hyperplane result to conclude that there exists* $p \in R^n, m \in R$ *such that* $H = \{x|px = m\}$ *supports* x^O*, meaning that*

$$px \geq r = px^0 \text{ for every } x \in U(x^0).$$

Now, consider some x_i *such that* $u_i(x_i) > u_i\left(x_i^O\right)$. *Then,* $px_i \geq px_i^O$. *This is so b/c if* $px_i < px_i^O$ *we have that*

$$p\left(x_i + \sum_{j \neq i} x_j^O\right) < p\left(x_i^O + \sum_{j \neq i} x_j^O\right),$$

which contradicts the separation result above since $x_i + \sum_{j \neq i} x_j^O \in U\left(x^O\right)$. *Hence, we may simply let*

$$\omega_i = px_i^O,$$

which completes the result since then any x_i *that is preferred to* x_i^O *is unaffordble.*

■

6.5 Separation Theorems and Linear Programming

Definition 12 *A convex cone is a nonempty set* $K \subset R^n$ *satisfying:*

1. $x, y \in K \Rightarrow x + y \in K$
2. $x \in K$ *and* $\alpha \geq 0s \Rightarrow \alpha x \in K$

One way to generalize this type of an object is by a set of "spanning vectors".

Definition 13 *Let* $A = \{a_1, ..., a_k\}$ *with* $a_i \in R^n$ *for each* $i \in \{1, ..., k\}$. *Then, the set*

$$K(A) = \left\{x \in R^n | \exists \lambda \in R_+^k \text{ such that } x = \sum_{i=1}^k \lambda_i a_i\right\}$$

is called the convex polyhedral cone spanned by A.

Theorem 2 (Farkas Lemma) *Let $\{a_1, ..., a_k\}$ and $b \neq 0$ be points in R^n. Suppose that $bx \geq 0$ for every x such that $a_i x \geq 0$ for $i = 1, .., m$. Then, there exists $\lambda \in R^k_+ \setminus \{0\}$ such that $b = \sum_{i=1}^{k} \lambda_i a_i$ (eg b is in the convex cone spanned by $\{a_1, .., a_k\}$)*

Lemma 2 *Let K be a convex cone. Suppose that there exists M such that $px \geq M$ for every $x \in K$. Then, $px \geq 0$ for every $x \in K$.*

Proof. Suppose that $px \geq M$ for every and there exists $x' \in K$ such that $px' < 0$. Consider

$$\alpha = 2\frac{M}{px'} > 0$$

and note that

$$p\left(2\frac{M}{px'}x'\right) = 2M < M.$$

■

Proof. Let $K(A)$ be the convex cone generated by $\{a_1, ..., a_k\}$. For contradiction, suppose that $b \notin K(A)$. Since $K(A)$ is convex we can apply the simplest separation theorem to conclude that there is hyperplane $H = \{x | px = k\}$ such that

$$px > k > pb$$

for every $x \in K(A)$. Hence, px is bounded below by k so we can apply the lemma above to conclude that

$$px \geq 0$$

for every $x \in K(A)$ implying that:

1. $pb < 0$

2. $pa_i \geq 0$ since every $a_i \in K(A)$ and $px \geq 0$ for every $x \in K(A)$ by assumption. But this means that $pb = bp < 0$ for the vector p which is such that $pa_i = a_i p \geq 0$ for every i, which contradicts the hypothesis of the Theorem.

■

Theorem 3 (Theorem of the alternatives) *Let* $\{a_1, ..., a_k\}$ *and* $b \neq 0$ *be points in* R^n. *Then, exactly one of the two following alternatives is true:*

1. **ALTERNATIVE 1:** There exists $\lambda \in R^k_+$ such that $b = \sum_{i=1}^k \lambda_i a_i$

2. **ALTERNATIVE 2:** There exists $x \in R^n$ such that $bx < 0$ and $\alpha_i x \geq 0$ for all i.

Proof. We've proved that NOT ALTERNATIVE 2⇒ALTERNATIVE 1. Need to show that ALTERNATIVE 1⇒NOT ALTERNATIVE 2. To see this, suppose that there exists $\lambda \in R^k_+$ such that $b = \sum_{i=1}^k \lambda_i a_i$ and x is such that $\alpha_i x \geq 0$ for all i. Then

$$bx = \left(\sum_{i=1}^k \lambda_i a_i\right) x = \left(\sum_{i=1}^k \lambda_i a_i x\right) \geq 0$$

■

Consider a linear programming problem on the form

$$\begin{aligned} &\min_x bx \\ \text{s.t. } a_i x &\geq 0 \text{ for } i = 1, ..., k \end{aligned}$$

where $p, a_i \in R^n$, or

$$\begin{aligned} &\min_x bx \\ \text{s.t. } Ax &\geq 0 \ (k \text{ inequalities}) \end{aligned} \tag{1}$$

Also, consider the program

$$\begin{aligned} &\max 0 \cdot \lambda \\ A'\lambda &= b \ (n \text{ equalities}) \\ \lambda &\geq 0. \end{aligned} \tag{2}$$

Then, we have:

Theorem 4 (a duality result) *The primal has a solution if and only if the dual has a solution and:*

1. $bx^* = 0 = 0\lambda^*$ *where* x^* *solves the primal and* λ^* *solves the dual.*

2. $b = A'\lambda^* = \sum_{i=1}^k a_i \lambda_i^*$

Proof. [SOLUTIONS EXIST] Suppose that x^* solves the primal. Then, there exists no x such that

$$\begin{aligned} bx &< bx^* \leq 0 \\ Ax &\geq 0 \end{aligned}$$

Where $bx^* \leq 0$ follows as 0 is a feasible solution to the program. If $bx^* < 0$ then bx is not bounded below, so no solution exists to the program. Hence, we can use the theorem of the alternatives to conclude that there exists λ^* such that

$$\lambda^* \in R_+^k \text{ such that } b = \sum_{i=1}^k \lambda_i^* a_i \Leftrightarrow p = A'\lambda,$$

which shows that the there exist a feasible, and therefore optimal, solution to the dual.

[SOLUTIONS DON'T EXIST] Suppose no solution exists to primal. Then, there exists $x \in K(A)$ such that $bx < 0$. By the theorem of alternativives it follows that there is no $\lambda \in R_+^k$ such that $b = \sum_{i=1}^k \lambda_i a_i$. Hence, there is no feasible solution to the dual. ■

Consider the program

$$\begin{aligned} &\max_x f(x) \\ \text{s.t. } g_i(x) &\geq 0 \text{ for } i = 1, .., k \end{aligned}$$

where $f, g_1, .., g_k : R^n \to R$ are continuously differentiable functions. An obvious idea with respect to solving the non-linear program is to replace it with

$$\begin{aligned} &\max_x \nabla f(x^*)(x - x^*) \\ \text{s.t. } \nabla g_i(x^*)(x - x^*) &\geq 0 \text{ for } i = 1, .., k \end{aligned}$$

Usually, this works (but there are pathological cases)

Definition 14 *We say that the non-linear program satisfies constraint qualification if the solutions to the non-linear program solves the linear program.*

There are a bunch of sufficient conditions. I will use one of them below.

Theorem 5 (version of Kuhn-Tucker) *Suppose that x^* is a local maximum and suppose that constraints $i = 1, ..., m$ are binding ($g_j(x^*) > 0$ for $j = m+1, .., k$). Moreover, suppose that $f, g_1, ..., g_k$ are continuously differentiable and that*

$$\left(\nabla g_1(x^*), ..., \nabla g_m(x^*)\right)$$

are linearly independent. Then there exists $\lambda \in R_+^k$ such that

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^{K} \lambda_i \nabla g_i(x^*) &= 0 \\ \lambda_i g_i(x^*) &= 0 \text{ for } i = 1, ..k \end{aligned}$$

.

Proof. Suppose that x^* is a local maximum in the constrained optimization problem. Then, there exists $\varepsilon > 0$ such that

$$f(x^*) \geq f(x)$$

for all $x \in B(x^*, \varepsilon)$ such that

$$g_j(x) \geq 0.$$

By contiinuity we may pick ε so that

$$g_j(x) > 0$$

for every $x \in B(x^*, \varepsilon)$ and $j = m+1, ..., k$. We assign $\lambda_i = 0$ to these constraints and note that we can rephrase the assumption that x^* is a local max as there being no $x \in B(x^*, \varepsilon)$ such that

$$\begin{aligned} f(x) &> f(x^*) \\ g_j(x) &\geq g_j(x^*) = 0 \text{ for } j = 1, ..., m \end{aligned}$$

This implies that there is no perturabation so that

$$\begin{aligned} f(x^*) + \nabla f(x^*)(x - x^*) &> f(x^*) \text{ and} \\ g_j(x^*) + \nabla f(x^*)(x - x^*) &\geq g_j(x^*) = 0 \end{aligned}$$

orThe full rank condition assures that there is a direction $(x - x^*)$ such that $\nabla g_j(x^*)(x - x^*) \geq 0$ (otherwise constraint qualification fails and Kuhn-Tucker conditions are no longer necessary for an optimum). By the theorem of the alternatives it follows that there is $\lambda \in R_+^m$ such that

$$-\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x^*) = 0,$$

where the second inquality follows from the fact that $\lambda_i = 0$ for $i > m$. Moreover as $g_i(x^*) = 0$ for $i \leq m$ it follows that

$$\lambda_i g_i(x^*) = 0$$

for every $i \in I$. ■

6.6 The Hahn Banach Theorem

Definition 15 *A (finite) functional $f : L \to R$ where L is a linear space is said to be sublinear (aka convex functional) if:*

1. *$f(x) \geq 0$ for every $x \in L$*
2. *$f(\alpha x) = \alpha f(x)$ for every $x \in L, \alpha \geq 0$*
3. *$f(x + y) \leq f(x) + f(y)$ for every $x, y \in L$.*

Proposition 6 *If $f : L \to R$ is sublinear, then $\{x \in L | f(x) \leq k\}$ is convex for every k.*

Proof. $f(\lambda x + (1 - \lambda) y) \leq f(\lambda x) + f((1 - \lambda) y) = \lambda f(x) + (1 - \lambda) f(y) \leq k$ if $f(x) \leq k$ and $f(y) \leq k$ ■

Proposition 7 *If $f : L \to R$ is (finite and) sublinear, and if there exists x such that $f(x) > 0$ then $\{x \in L | f(x) \leq k\}$ has a nonempty interior for every $k > 0$ (most importantly, 0 is in the interior.*

Proof. Let x be such that $0 < f(x) < k$. Pick $y \in L$ ($-y \in L$ since L is a linear space). Suppose that $f(y) = f(-y) = 0$. Then,

$$f(x + ty) \leq f(x) + tf(y) = f(x) < k$$

for every t. Without loss assume that $f(y) \geq f(-y)$. Then, if $f(y) > 0$ we have that for $t < \frac{k-f(x)}{f(y)}$

$$\begin{aligned} f(x+ty) &\leq f(x) + tf(y) < f(x) + \left(\frac{k-f(x)}{f(y)}\right) f(y) = k \\ f(x+ty) &\leq f(x) + tf(-y) < f(x) + \left(\frac{k-f(x)}{f(y)}\right) f(-y) \\ &\leq f(x) + k - f(x) = k \end{aligned}$$

Hence, the interior of $\{x \in L | f(x) \leq k\}$ s nonempty. ■

6.6.1 Bases in Abstract Spaces

In Euclidean spaces we define a basis as follows. First we define a set of vectors to be linearly dependent if one of them can be written as a linear combination of the others. This can be done also if we have a countably infinite set of vectors:

Definition 16 *Let $\{x_i\}_{i \in I}$ be a set of points with $x_i \in L$, where L is a linear space. Then,*

1. *$\{x_i\}_{i \in I}$ are said to be linearly dependent if there exists coefficients $\{a_i\}_{i \in I}$ with $a_i \neq 0$ for some $i \in I$ such that*
$$\sum_{i \in I} a_i x_i = 0.$$
2. *$\{x_i\}_{i \in I}$ are said to be linearly independent if not linearly dependent.*

Definition 17 *Let $\{x_i\}_{i \in I}$ be a set of points with $x_i \in L$, where L is a linear space and let $L(\{x_i\}_{i \in I})$ be the smallest subspace that contains $\{x_i\}_{i \in I}$. Then, $L(\{x_i\}_{i \in I})$ is called the linear hull of $\{x_i\}_{i \in I}$ ($L(\{x_i\}_{i \in I}) \subset L$, and $L(\{x_i\}_{i \in I})$ can be thought of as the intersection of all subspaces containing $\{x_i\}_{i \in I}$, so existence is no problem)*

Definition 18 *If the smallest subspace of L containing $\{x_i\}_{i \in I}$ is L we say that $\{x_i\}_{i \in I}$ is a Hamel basis of L.*

Example 11 *In l^2 we have that*

$$\begin{aligned}&(1,0,0,....)\\&(0,1,0,....)\\&(0,0,1,0,...)\\&....\\&(...0,1,0,...)\end{aligned}$$

is a Hamel basis.

The problem with the concept of a Hamel basis is that it requires infinite summations when the vector space is infinite dimensional. For general linear spaces L such infinite sums are ill-defined. Hence, we use the following notion

Definition 19 *$B = \{x_i\}_{i \in I}$ is called a basis for a real vector space L if;*

1. *If every finite subset $B_0 \subset B$ consists of linearly independent vectors.*
2. *For every $x \in L$ there exists finite set $\{x_1, ..., x_n\}$ and real coefficients $\{a_1, ..., a_n\}$such that $x = \sum a_i x_i$.*

Theorem 6 (Hahn Banach) *Let $f : L \to R$ be a sublinear functional on a real linear space L. Let S be a suspace of L and suppose that that $g_s : S \to R$ is a linear functional such that $g_s(x) \leq f(x)$ for every $x \in S$. Then there exists an extension of g_s, a linear functional $g : L \to R$ such that*

$$\begin{aligned}g(x) &\leq f(x) \text{ for every } x \in L\\ g(x) &= g_s(x) \text{ for every } x \in S\end{aligned}$$

Proof. If$L = S$ the result is trivial. If not, let B be a basis for L and pick $z \in B \backslash S$, which must be non empty if $L \neq S$. Let $x', x'' \in S$ be arbitrary and note that

$$\begin{aligned}&g_s(x'') - g_s(x') \underset{\text{linear}}{=} g_s(x'') + g_s(-x') \underset{\text{linear}}{=} g_s(x'' - x') \underset{\text{assumption}}{\leq} f(x'' - x')\\ = \ &f(x'' + z - (x' + z)) \underset{\text{sublinear}}{\leq} f(x'' + z) + f(-(x' + z))\end{aligned}$$

Hence,

$$g_s(x'') - f(x'' + z) \leq g_s(x') + f(-(x' + z))$$

for any $(x', x'') \Rightarrow$

$$c'' = \sup_{x'' \in S} [g_s(x'') - f(x'' + z)] \leq \inf_{x' \in S} [g_s(x') + f(-(x' + z))] = c'$$

Now, let

$$\widetilde{g}_s(z) = \frac{c'' + c'}{2}.$$

Then, consider any point in the expanded space, that is a point on form $x + tz$ where $x \in S$ and note that if $t > 0$ then

$$\begin{aligned}
\widetilde{g}_s(x + tz) &= g_s(x) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x + tz) - f(x + tz) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x + tz) - tf\left(\frac{x}{t} + z\right) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x + tz) + t\left[\widetilde{g}_s(z) - f\left(\frac{x}{t} + z\right)\right] \\
&= g_s(x) + f(x + tz) + t\left[\widetilde{g}_s(z) - g_s\left(\frac{x}{t}\right) + \underbrace{g_s\left(\frac{x}{t}\right) - f\left(\frac{x}{t} + z\right)}_{\leq \sup_{x \in S}[g_s(x) - f(x+z)] = c' \leq \widetilde{g}_s(z)}\right] \\
&\leq g_s(x) + f(x + tz) + t\left[-g_s\left(\frac{x}{t}\right)\right] = g_s(x) + f(x + tz) - g_s(x) = f(x + tz)
\end{aligned}$$

Moreover, if $t < 0$, then

$$
\begin{aligned}
\widetilde{g}_s(x+tz) &= g_s(x) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x+tz) - f(x+tz) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x+tz) - f\left(-\frac{|t|}{t}x - |t|\,z\right) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x+tz) - |t|\,f\left(-\left(\frac{x}{t}+z\right)\right) + t\widetilde{g}_s(z) \\
&= g_s(x) + f(x+tz) - |t|\left\{f\left(-\left(\frac{x}{t}+z\right)\right) - \widetilde{g}_s(z)\right\} \\
&= g_s(x) + f(x+tz) - |t|\left\{\underbrace{g_s\left(\frac{x}{t}\right) + f\left(-\left(\frac{x}{t}+z\right)\right)}_{\geq \inf_{x\in S}[g_s(x)+f(-(x+z))]\geq \widetilde{g}_s(z)} - g_s\left(\frac{x}{t}\right) - \widetilde{g}_s(z)\right\} \\
&\leq g_s(x) + f(x+tz) - |t|\left\{-g_s\left(\frac{x}{t}\right)\right\} = g_s(x) + f(x+tz) - tg_s\left(\frac{x}{t}\right) \\
&= f(x+tz)
\end{aligned}
$$

Hence, we've shown that

$$\widetilde{g}_s(x) \leq f(x)$$

for every point in the larger space which consists of the original subspace and all points which are linear combinations of points from the orignal space and point z.

- Now, suppose that L is spanned by a countable set $\{x_k\} = X$ and let

  $$\{z_k\} = \{x_k \in X | x_k \notin S\}$$

  Argument by induction.

- If no countable set of elements span L, then consider the set of extensions satisfying

  $$\widetilde{g}_s(x+tz) = f(x) + t\widetilde{g}_s(z)$$

  I'll skip details. The result follows from the axiom of choice/zorns lemma. ■

Definition 20 *Let E be a convex set such that 0 is in the interior of E, then $f_M : L \to R$ given by*

$$f_M(x) = \inf\left\{r \in R_+ | \frac{x}{r} \in E\right\}$$

is called the Minkowski functional of E.

Note that $x \in E$ and $0 \in E \Rightarrow$

$$\frac{x}{r} = \frac{1}{r}x + \frac{r-1}{r}0 \in E$$

if $r > 1$. Hence, $f_M(x) \leq 1$ for points in the set that defines the Minkowski functional.

Proposition 8 *The Minkowski functional is sublinear*

Proof. Note first that $\frac{x}{r} \to 0$, and there exists $\varepsilon > 0$ such that $B(0,\varepsilon) \subset E$, and hence $\frac{x}{r} \in B(0,\varepsilon) \subset E$ for r large enough, so $f_M(x) \geq 0$ [property 1] is well-defined for every $x \in L$. Let $\alpha > 0$. Then

$$f_M(\alpha x) = \inf\left\{r \in R_+ | \frac{\alpha x}{r} \in E\right\}$$

Let $s = \frac{s}{\alpha}$. Then, $\alpha s = r$, so

$$\begin{aligned} f_M(\alpha x) &= \inf\left\{r \in R_+ | \frac{\alpha x}{r} \in E\right\} \\ &= \inf\left\{\alpha s \in R_+ | \frac{\alpha x}{\alpha s} = \frac{x}{s} \in E\right\} \\ &= \alpha \inf\left\{s \in R_+ | \frac{x}{s} \in E\right\} = \alpha f(x) \end{aligned}$$

[property 2] Now, pick any $x, y \in L$ and let $\varepsilon > 0$ be arbitrary and r_x, r_y be numbers such that

$$\begin{aligned} f_M(x) &< r_x < f_M(x) + \frac{\varepsilon}{2} \\ f_M(y) &< r_x < f_M(y) + \frac{\varepsilon}{2} \end{aligned}$$

Then given any $\varepsilon > 0$ we have that

$$\begin{aligned} \frac{x}{r_x} &\in E \text{ and } \frac{y}{r_y} \in E \Rightarrow \text{(convexity)} \\ \frac{x+y}{r_x+r_y} &= \frac{r_x}{r_x+r_y}\left(\frac{x}{r_x}\right) + \frac{r_y}{r_x+r_y}\left(\frac{y}{r_y}\right) \in E \end{aligned}$$

for each ε. Hence,

$$\begin{aligned} f_M(x+y) &= \inf\left\{r \in R_+ | \frac{x+y}{r} \in E\right\} \leq r_x + r_y \\ &< f_M(x) + f_M(y) + \varepsilon \Rightarrow \\ f_M(x+y) &\leq f_M(x) + f_M(y) \end{aligned}$$

since $\varepsilon > 0$ was arbitrary. ■

Proposition 9 *Let A and B be disjoint convex sets in a real linear space L and assume that A has a nonempty interior. Then there exists a non-trivial linear functional f that separates A and B.*

PROOF: Without loss, assume that 0 is in the interior of A. Let $b \in B$, which implies that

$$\begin{aligned} -b &\in A - B \\ &= \{x \in L | \text{there is } y \in A \text{ and } z \in B \text{ such that } x = y - z\}. \end{aligned}$$

Indeed, $-b$ is in the interior of the set as there exists $\varepsilon > 0$ such that the neighborhood $N(0, \varepsilon) \subset A$, so for every $x \in N(-b, \varepsilon)$ pick

$$\begin{aligned} y &= x - b \in N(0, \varepsilon) \\ z &= -b. \end{aligned}$$

Also $0 \notin A - B$, implying that

$$b \notin A - B - \{b\}$$

Consider the Minkowski functional for $A - B - \{b\}$. Then

$$f_M(b) = \inf\left\{r \in R_+ | \frac{b}{r} \in A - B - \{b\}\right\} \geq 1$$

since

$$\begin{aligned} f_M(b) &= \inf\left\{r \in R_+ | \frac{b}{r} \in A - B - \{b\}\right\} < 1 \Rightarrow \\ \frac{b}{1} &= b \in A - B - \{b\}, \end{aligned}$$

which is a contradiction. Consider the one-dimensional subspace

$$S = \{x \in L | \text{there exists } \alpha \in R \text{ with } x = \alpha b\}$$

and let $g_s : S \to R$ be

$$g_s(\alpha b) = \alpha f_M(b)$$

Hence

$$\begin{aligned} g_s(\alpha b) &= \alpha f_M(b) = f_M(\alpha b) \text{ for } \alpha \geq 0 \\ g_s(\alpha b) &= \alpha g_s(b) < 0 \leq f_M(\alpha b) \text{ for } \alpha < 0 \\ &\Rightarrow \\ g_s(x) &\leq f_M(x) \end{aligned}$$

for all $x = \alpha b \in S$. By the Hahn-Banach theorem $g_s : S \to R$ can be extended to a linear functional $g : L \to R$ such that

$$g(x) \leq f_M(x) \text{ for all } x \in L$$

But $y \in A - B - \{b\} \Rightarrow f_M(y) \leq 1$ since the Minkowski functional always takes on values in $[0, 1]$ at points in the convex set that defines it. Hence,

$$\begin{aligned} g(x) &\leq f_M(x) \text{ for all } x \in A - B - \{b\} \\ f_M(b) &= g(b) \geq 1 \end{aligned}$$

so:

- g separates $A - B - \{b\}$ and $\{b\} \Leftrightarrow$
- g separates $A - B$ and $\{0\} \Leftrightarrow$
- g separates A and B. ■