Convex Analysis

We'll assume throughout, without always saying so, that we're in the finite-dimensional Euclidean vector space $\mathbb{R}^n$, although sometimes, for statements that hold in any vector space, we'll say explicitly that we're in a vector space V.

Definition: A set S in a vector space V is **convex** if for any two points x and y in S, and any λ in the unit interval $[0, 1]$, the point $(1 - \lambda)x + \lambda y$ is in S.

Theorem: The intersection of any collection of convex sets is convex — *i.e.*, if for each α in some set A the set S_α is convex, then the set $\bigcap_{\alpha \in A} S_\alpha$ is convex.

Theorem: The closure and the interior of a convex set in $\mathbb{R}^n$ are both convex.

Theorem: If $X_1, X_2, \ldots, X_m$ are convex sets, then $\sum_1^m X_i$ is convex.

Theorem: For any sets $X_1, X_2, \ldots, X_m$ in $\mathbb{R}^n$, $\sum_{i=1}^m cl\, X_i \subseteq cl \sum_{i=1}^m X_m$ — *i.e.*, the sum of the sets' closures is a subset of the closure of their sum.

Exercise: Provide proofs of the above theorems and a counterexample to show that the sum of sets' closures need not be *equal to* the closure of their sum.

Definition: Let $\mathbf{p} \neq \mathbf{0} \in \mathbb{R}^n$ and let b be a real number. The set of solutions of the linear equation $p_1x_1 + \cdots + p_nx_n = b$ is a **hyperplane** in $\mathbb{R}^n$, and is denoted $H(\mathbf{p}, b)$ — *i.e.*, $H(\mathbf{p}, b) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} = b\}$.

Remark: For any non-zero real number λ, we have $H(\lambda\mathbf{p}, \lambda b) = H(\mathbf{p}, b)$. Therefore for any norm $\|\cdot\|$ on $\mathbb{R}^n$ any hyperplane can be represented by a $\mathbf{p}$ for which $\|\mathbf{p}\| = 1$: given a hyperplane $H(\mathbf{p}, b)$ with $\|\mathbf{p}\| \neq 1$, let $\lambda = 1/\|\mathbf{p}\|$ and let $\mathbf{p}' = \lambda\mathbf{p}$ and $b' = \lambda b$; then $H(\mathbf{p}', b') = H(\mathbf{p}, b)$ and $\|\mathbf{p}'\| = 1$.

Definition: A **closed half-space** is a set of the form $\{x \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} \leqq b\}$ for some $\mathbf{p} \neq \mathbf{0} \in R^n$ and $b \in \mathbb{R}$. An **open half-space** is a set of the form $\{x \in \mathbb{R}^n \mid \mathbf{p} \cdot \mathbf{x} < b\}$ for some $\mathbf{p} \neq \mathbf{0} \in R^n$ and $b \in \mathbb{R}$.

Remark: A closed (resp. open) half-space is the set of all points on one side of a hyperplane, including (resp. not including) the hyperplane itself.

If $\mathbf{p} = (p_1, \ldots, p_n)$ is a list of prices of some items and $\mathbf{x} = (x_1, \ldots, x_n)$ is a "bundle" of the items, then $\mathbf{p} \cdot \mathbf{x}$ is the value of the bundle $\mathbf{x}$. The hyperplane $H(\mathbf{p}, b)$ is the set of all bundles that have value b, and the half-spaces on each side of the hyperplane $H(\mathbf{p}, b)$ are the bundles whose value is greater than b and the bundles whose value is less than b.

Separating and Supporting Hyperplanes

Definition: The hyperplane $H(\mathbf{p}, b)$ **separates** sets X and Y in $\mathbb{R}^n$ if for all $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, we have $\mathbf{p} \cdot \mathbf{x} \leqq b \leqq \mathbf{p} \cdot \mathbf{y}$. We also say that a hyperplane separates a set X and a point $\mathbf{y}$ if it separates the sets X and $\{\mathbf{y}\}$. We say that $H(\mathbf{p}, b)$ **strictly separates** X and Y, or strictly separates X and y, if the inequalities are both strict. See Figures 1 and 2.

Definition: A hyperplane $H(\mathbf{p}, b)$ is **bounding** for a set S, or **bounds** S, if S lies entirely on one side of H — *i.e.*, if either $\forall \mathbf{x} \in S : \mathbf{p} \cdot \mathbf{x} \geqq b$ or $\forall \mathbf{x} \in S : \mathbf{p} \cdot \mathbf{x} \leqq b$.

Definition: A hyperplane $H(\mathbf{p}, b)$ is **supporting** for a set S, or **supports** S, if it is bounding for S and also contains a boundary point of S. See Figures 3 and 4.

Remark: If $H(\mathbf{p}, b)$ is a supporting hyperplane for a set S, then either $b = \sup\{\mathbf{p} \cdot \mathbf{x} \,|\, \mathbf{x} \in S\}$ (if $\mathbf{p} \cdot \mathbf{x} \leqq b$ for all $\mathbf{x} \in S$) or $b = \inf\{\mathbf{p} \cdot \mathbf{x} \,|\, \mathbf{x} \in S\}$ (if $\mathbf{p} \cdot \mathbf{x} \geqq b$ for all $\mathbf{x} \in S$). If S is closed, then $b = \max\{\mathbf{p} \cdot \mathbf{x} \,|\, \mathbf{x} \in S\}$ or $b = \min\{\mathbf{p} \cdot \mathbf{x} \,|\, \mathbf{x} \in S\}$.

The classical theorems of convex analysis are existence theorems, theorems that guarantee the existence of a supporting hyperplane or a separating hyperplane $H(\mathbf{p}, b)$ — *i.e.*, they're theorems that guarantee the existence of a price-list $\mathbf{p}$ for which all bundles in some convex set (for example, a feasible set, or an upper-contour set) cost more or cost less than some specified amount.

The following theorem is essential for establishing supporting and separating hyperplane theorems. Let's write $d(\mathbf{x}, \mathbf{y})$ for the Euclidean distance $\|\mathbf{x} - \mathbf{y}\|$ in $\mathbb{R}^n$.

Theorem: Let S be a nonempty closed set in $\mathbb{R}^n$, and let $\overline{\mathbf{x}}$ be a point which is not in S. Then there is a point $\widehat{\mathbf{x}} \in S$ that is closest to $\overline{\mathbf{x}}$ — *i.e.*, such that $d(\widehat{\mathbf{x}}, \overline{\mathbf{x}}) \leqq d(\mathbf{x}, \overline{\mathbf{x}})$ for all $\mathbf{x} \in S$.

Proof: Let $\mathbf{y} \in S$ and let $r = d(\mathbf{y}, \overline{\mathbf{x}})$. Note that $\mathbf{y} \neq \overline{\mathbf{x}}$, and therefore $r > 0$. Define $\overline{B} = clB(\overline{\mathbf{x}}, r)$, the closed ball of radius r about $\overline{\mathbf{x}}$. Note that $\mathbf{y} \in \overline{B}$ and $\overline{B}$ is compact. Therefore $\overline{B} \cap S$ is nonempty ($\mathbf{y} \in \overline{B} \cap S$) and compact (because S is closed). The function $d(\cdot, \overline{\mathbf{x}})$ is continuous, therefore it attains a minimum on the compact set $\overline{B} \cap S$, say at a (not necessarily unique) point $\widehat{\mathbf{x}} \in \overline{B} \cap S$. Now suppose there is a point $\widetilde{\mathbf{x}} \in S$ that is closer to $\overline{\mathbf{x}}$ — *i.e.*, $d(\widetilde{\mathbf{x}}, \overline{\mathbf{x}}) < d(\widehat{\mathbf{x}}, \overline{\mathbf{x}}) \leqq r$, so we have $\widetilde{\mathbf{x}} \in \overline{B}$ as well, and therefore $\widetilde{\mathbf{x}} \in \overline{B} \cap S$ and $d(\widetilde{\mathbf{x}}, \overline{\mathbf{x}}) < d(\widehat{\mathbf{x}}, \overline{\mathbf{x}})$, contradicting the fact that $\widehat{\mathbf{x}}$ minimizes $d(\mathbf{x}, \overline{\mathbf{x}})$ on $\overline{B} \cap S$. ■

Now we can prove two theorems guaranteeing that a convex set S can be separated by a hyperplane from any point that's either not in S or is in the boundary of S. Then we'll use the theorems to establish the classical Minkowski Theorem, which guarantees that two disjoint convex sets in $\mathbb{R}^n$ can be separated by a hyperplane.

Closed-set Supporting Hyperplane Theorem: Let S be a nonempty, closed, convex set. For any point $\overline{\mathbf{x}}$ which is not in S, there is a hyperplane $H(\mathbf{p}, b)$ that supports S and separates $\overline{\mathbf{x}}$ and S.

Proof: Let $\overline{\mathbf{x}} \notin S$ and let $\widehat{\mathbf{x}}$ be a point in S that is closest to $\overline{\mathbf{x}}$, the existence of which is guaranteed by the preceding proposition. Let $\mathbf{p} = \widehat{\mathbf{x}} - \overline{\mathbf{x}}$ and let $b = \mathbf{p} \cdot \widehat{\mathbf{x}}$. We will show that $H(\mathbf{p}, b)$ supports S and separates $\overline{\mathbf{x}}$ and S.

We first show that $\mathbf{p} \cdot \overline{\mathbf{x}} < b$. We have $\mathbf{p} \cdot \widehat{\mathbf{x}} - \mathbf{p} \cdot \overline{\mathbf{x}} = \mathbf{p} \cdot (\widehat{\mathbf{x}} - \overline{\mathbf{x}}) = \mathbf{p} \cdot \mathbf{p}$. Since $\widehat{\mathbf{x}} \neq \overline{\mathbf{x}}$, we have $\mathbf{p} \neq \mathbf{0}$, and therefore $\mathbf{p} \cdot \mathbf{p} > 0$, so we have $\mathbf{p} \cdot \widehat{\mathbf{x}} - \mathbf{p} \cdot \overline{\mathbf{x}} > 0$ — *i.e.*, $\mathbf{p} \cdot \overline{\mathbf{x}} < \mathbf{p} \cdot \widehat{\mathbf{x}} = b$.

In order to show that $\mathbf{x} \in S \Rightarrow \mathbf{p} \cdot \mathbf{x} \geqq b$, let $\mathbf{x} \in S$ and assume, by way of contradiction, that $\mathbf{p} \cdot \mathbf{x} < b$. For each $\lambda \in (0, 1)$ define $\mathbf{x}^\lambda = (1 - \lambda)\widehat{\mathbf{x}} + \lambda\mathbf{x}$. (See Figure 9.) Since S is convex and $\widehat{\mathbf{x}}, \mathbf{x} \in S$, we have $\mathbf{x}^\lambda \in S$ for all $\lambda \in (0, 1)$. We will show that for small values of λ we have $d(\overline{\mathbf{x}}, \mathbf{x}^\lambda) < d(\overline{\mathbf{x}}, \widehat{\mathbf{x}})$, which will provide our contradiction, since $\widehat{\mathbf{x}}$ minimizes $d(\overline{\mathbf{x}}, \mathbf{x})$ on S.

First note that we have

$$\begin{aligned} \mathbf{x}^\lambda - \overline{\mathbf{x}} &= (1 - \lambda)\widehat{\mathbf{x}} + \lambda\mathbf{x} - \overline{\mathbf{x}} \\ &= (\widehat{\mathbf{x}} - \overline{\mathbf{x}}) + \lambda(\mathbf{x} - \widehat{\mathbf{x}}) \\ &= \mathbf{p} + \lambda(\mathbf{x} - \widehat{\mathbf{x}}). \end{aligned}$$

We show that $d(\overline{\mathbf{x}}, \mathbf{x}^\lambda) < d(\overline{\mathbf{x}}, \widehat{\mathbf{x}})$, or equivalently, $\|\overline{\mathbf{x}} - \mathbf{x}^\lambda\|^2 < \|\overline{\mathbf{x}} - \widehat{\mathbf{x}}\|^2$:

$$\begin{aligned} \|\overline{\mathbf{x}} - \mathbf{x}^\lambda\|^2 &= (\mathbf{x}^\lambda - \overline{\mathbf{x}}) \cdot (\mathbf{x}^\lambda - \overline{\mathbf{x}}) \\ &= \big[\, \mathbf{p} + \lambda(\mathbf{x} - \widehat{\mathbf{x}}) \,\big] \cdot \big[\, \mathbf{p} + \lambda(\mathbf{x} - \widehat{\mathbf{x}}) \,\big] \\ &= \mathbf{p} \cdot \mathbf{p} + 2\lambda\mathbf{p} \cdot (\mathbf{x} - \widehat{\mathbf{x}}) + \lambda^2\|\mathbf{x} - \widehat{\mathbf{x}}\|^2 \\ &= \|\widehat{\mathbf{x}} - \overline{\mathbf{x}}\|^2 + \lambda g(\lambda), \end{aligned}$$

where $g(\lambda) := 2(\mathbf{p} \cdot \mathbf{x} - \mathbf{p} \cdot \widehat{\mathbf{x}}) + \lambda\|\mathbf{x} - \widehat{\mathbf{x}}\|^2$. Because, by assumption, $\mathbf{p} \cdot \mathbf{x} < \mathbf{p} \cdot \widehat{\mathbf{x}}$, we have $\lim_{\lambda \to 0} g(\lambda) < 0$, and therefore for small values of λ we have $\|\overline{\mathbf{x}} - \mathbf{x}^\lambda\|^2 < \|\overline{\mathbf{x}} - \widehat{\mathbf{x}}\|^2$ — *i.e.*, $d(\overline{\mathbf{x}}, \mathbf{x}^\lambda) < d(\overline{\mathbf{x}}, \widehat{\mathbf{x}})$, the contradiction we wished to show, thereby establishing that $\mathbf{p} \cdot \mathbf{x} \geqq b$ for all $\mathbf{x} \in S$. ■

Corollary: Let S be a nonempty, closed, convex set. For any point $\overline{\mathbf{x}}$ which is not in S, there is a hyperplane $H(\mathbf{p}, c)$ containing $\overline{\mathbf{x}}$ which separates $\overline{\mathbf{x}}$ and S.

Proof: Let $H(\mathbf{p}, b)$ be a hyperplane for which $\mathbf{p} \cdot \overline{\mathbf{x}} \leqq b$ and $\mathbf{x} \in S \Rightarrow \mathbf{p} \cdot \mathbf{x} \geqq b$, the existence of which is guaranteed by the theorem, and let $c = \mathbf{p} \cdot \overline{\mathbf{x}}$. (See Figure 9.) ■

Corollary: A closed convex set S is the intersection of the closed half-spaces that contain S.

Boundary-point Supporting Hyperplane Theorem: If S is a nonempty convex set and $\overline{\mathbf{x}}$ is in the boundary of S, then there is a hyperplane that supports S and contains $\overline{\mathbf{x}}$.

Proof: Let $\overline{S}$ denote the closure of S; $\overline{S}$ is a nonempty closed convex set. Because $\overline{\mathbf{x}}$ is a boundary point of S, for every $n \in \mathbb{N}$ the open ball $B(\overline{\mathbf{x}}, \frac{1}{n})$ contains a point $\mathbf{x}_n \notin \overline{S}$. Note that $\lim \mathbf{x}_n = \overline{\mathbf{x}}$. The previous theorem ensures that for each of these points $\mathbf{x}_n$ there is a $\mathbf{p}_n \neq \mathbf{0} \in \mathbb{R}^n$ such that

$$\forall \mathbf{x} \in S: \ \mathbf{p}_n \cdot \mathbf{x} \geqq \mathbf{p}_n \cdot \mathbf{x}_n; \quad i.e., \ \forall \mathbf{x} \in S: \ \mathbf{p}_n \cdot (\mathbf{x} - \mathbf{x}_n) \geqq 0,$$

and we may assume without loss of generality that $\|\mathbf{p}_n\| = 1$. The set of all $\mathbf{p} \in \mathbb{R}^n$ such that $\|\mathbf{p}\| = 1$ is the unit sphere in $\mathbb{R}^n$, a compact set, so the sequence $\{\mathbf{p}_n\}$ has a convergent subsequence. We restrict attention to this subsequence, which we also denote by $\{\mathbf{p}_n\}$, and we denote its limit by $\overline{\mathbf{p}}$. Note that $\|\overline{\mathbf{p}}\| = 1$; in particular, $\overline{\mathbf{p}} \neq \mathbf{0}$. We also denote the corresponding subsequence of $\{\mathbf{x}_n\}$ by $\{\mathbf{x}_n\}$, and for each $n \in \mathbb{N}$ we now have $\{\mathbf{x}_n\} \to \overline{\mathbf{x}}$, $\{\mathbf{p}_n\} \to \overline{\mathbf{p}}$, and

$$\forall \mathbf{x} \in S: \ \mathbf{p}_n \cdot (\mathbf{x} - \mathbf{x}_n) \geqq 0.$$

Therefore

$$\forall \mathbf{x} \in S: \ \overline{\mathbf{p}} \cdot (\mathbf{x} - \overline{\mathbf{x}}) \geqq 0; \quad i.e., \ \forall \mathbf{x} \in S: \ \overline{\mathbf{p}} \cdot \mathbf{x} \geqq \overline{\mathbf{p}} \cdot \overline{\mathbf{x}}) \geqq 0. \ \blacksquare$$

Combining the two previous theorems gives us a theorem about any convex set and any point not in the set.

Theorem: Let S be a nonempty convex set. For any point $\overline{\mathbf{x}}$ which is not in S, there is a hyperplane that supports S and separates S and $\overline{\mathbf{x}}$.

Proof: If $\overline{\mathbf{x}}$ is a boundary point of S then the Boundary-point Supporting Hyperplane Theorem yields the desired result. If $\overline{\mathbf{x}}$ is not a boundary point of S, then $\overline{\mathbf{x}} \notin \overline{S}$, a nonempty, closed, convex set, and the Closed-set Supporting Hyperplane Theorem yields the desired result. $\blacksquare$

Note that a separating or supporting hyperplane might be unique, as in Figures 2 and 3, but these theorems don't guarantee uniqueness, as shown in Figures 1 and 4. The theorems require that the set(s) are convex and (in some cases) closed, but these are not necessary conditions for the existence of a separating or supporting hyperplane: in Figure 1 the set X is not convex and either set might not be closed, but there is still a separating hyperplane. On the other hand, we can't dispense with the condition that the set(s) are convex; otherwise, as in Figures 5 and 6, there might not exist a separating or supporting hyperplane. Figure 1 also shows that sets need not be bounded.

Minkowski's Theorem: Let S_1 and S_2 be nonempty disjoint convex sets. Then there exist $\mathbf{p} \neq \mathbf{0} \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that $\forall \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2 : \mathbf{p} \cdot \mathbf{x}_2 \leqq b \leqq \mathbf{p} \cdot \mathbf{x}_1$ — *i.e.*, there is a hyperplane that separates the sets.

Proof: We first show that since S_1 and S_2 are disjoint, $\mathbf{0} \notin S_1 - S_2 = S_1 + (-S_2)$. Suppose instead that $\mathbf{0} \in S_1 - S_2$. Then there exist $\mathbf{x}_1 \in S_1$ and $\mathbf{x}_2 \in S_2$ such that $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$. But then $\mathbf{x}_2 = \mathbf{x}_1$, and we therefore have $\mathbf{x}_1 \in S_1$ and $\mathbf{x}_1 \in S_2$, *i.e.*, S_1 and S_2 are not disjoint, a contradiction. Therefore $\mathbf{0} \notin S_1 - S_2$. Since $S_1 - S_2$ is nonempty and convex (as the sum of convex sets), there is a hyperplane that separates the set $S_1 - S_2$ and the point $\mathbf{0}$, so we have

$$\forall \mathbf{z} \in S_1 - S_2 : \ \mathbf{p} \cdot \mathbf{z} \geqq \mathbf{p} \cdot \mathbf{0} = 0; \ \textit{i.e.},$$

$$\forall \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2 : \ \mathbf{p} \cdot (\mathbf{x}_1 - \mathbf{x}_2) \geqq 0; \ \textit{i.e.},$$

$$\forall \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2 : \ \mathbf{p} \cdot \mathbf{x}_1 \geqq \mathbf{p} \cdot \mathbf{x}_2.$$

Clearly $\inf_{\mathbf{x}_1 \in S_1} \mathbf{p} \cdot \mathbf{x}_1$ and $\sup_{\mathbf{x}_2 \in S_2} \mathbf{p} \cdot \mathbf{x}_2$ both exist and satisfy $\sup_{\mathbf{x}_2 \in S_2} \mathbf{p} \cdot \mathbf{x}_2 \leqq \inf_{\mathbf{x}_1 \in S_1} \mathbf{p} \cdot \mathbf{x}_1$. Let b be any real number that satisfies

$$\sup_{\mathbf{x}_2 \in S_2} \mathbf{p} \cdot \mathbf{x}_2 \leqq b \leqq \inf_{\mathbf{x}_1 \in S_1} \mathbf{p} \cdot \mathbf{x}_1,$$

and then we have

$$\forall \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2 : \ \mathbf{p} \cdot \mathbf{x}_2 \leqq \sup_{\mathbf{x} \in S_2} \mathbf{p} \cdot \mathbf{x} \leqq b \leqq \inf_{\mathbf{x} \in S_1} \mathbf{p} \cdot \mathbf{x} \leqq \mathbf{p} \cdot \mathbf{x}_1. \ \blacksquare$$

Convex Optimization

Here's an example of the kind of theorem we can often obtain by using convex analysis in place of differentiability.

Theorem: Let $f : X \to \mathbb{R}$ be a continuous quasiconcave function on a convex domain $X \subseteq \mathbb{R}^n$; and let S be a convex subset of X. Let $\overline{\mathbf{x}}$ be an element of S at which f does not attain a local maximum on X. Then $\overline{\mathbf{x}}$ maximizes f on S if and only if there is a $\mathbf{p} \neq \mathbf{0} \in \mathbb{R}^n$ such that

(a) $\overline{\mathbf{x}}$ maximizes $f(\mathbf{x})$ s.t. $\mathbf{p} \cdot \mathbf{x} \leqq \mathbf{p} \cdot \overline{\mathbf{x}}$ and (b) $\overline{\mathbf{x}}$ maximizes $\mathbf{p} \cdot \mathbf{x}$ s.t. $\mathbf{x} \in S$.

Proof: It's easy to see that (a) and (b) together imply that $\overline{\mathbf{x}}$ maximizes f on S: if $\mathbf{x} \in S$, then $\mathbf{p} \cdot \mathbf{x} \leqq \mathbf{p} \cdot \overline{\mathbf{x}}$ according to (b); and therefore $f(\mathbf{x}) \leqq f(\overline{\mathbf{x}})$ according to (a).

To prove the converse, we assume that $\overline{\mathbf{x}}$ maximizes f on S, and we will show that there is a $\mathbf{p} \neq \mathbf{0} \in \mathbb{R}^n$ that satisfies (a) and (b). Let $U = \{\mathbf{x} \in X \mid f(\mathbf{x}) > f(\overline{\mathbf{x}})\}$, the strict f-upper-contour set of $\overline{\mathbf{x}}$. U is nonempty and convex, and is disjoint from S; and since S is also nonempty and convex, the Minkowski Theorem guarantees the existence of a $\mathbf{p} \neq \mathbf{0} \in \mathbb{R}^n$ and a real number $b > 0$ such that the hyperplane $H(\mathbf{p}, b)$ separates the two sets — *i.e.*,

$$(*) \quad \forall \mathbf{x} \in S : \mathbf{p} \cdot \mathbf{x} \leqq b \qquad \text{and} \qquad (**) \quad \forall \mathbf{x} \in U : b \leqq \mathbf{p} \cdot \mathbf{x}.$$

We first show that $b = \mathbf{p} \cdot \overline{\mathbf{x}}$. Because $\overline{\mathbf{x}} \in S$, we have $\mathbf{p} \cdot \overline{\mathbf{x}} \leqq b$, according to $(*)$. We show that $\mathbf{p} \cdot \overline{\mathbf{x}} \geqq b$ as follows: for each $n \in \mathbb{N}$ let $\mathbf{x}_n$ be such that $\mathbf{x}_n \in B(\overline{\mathbf{x}}, \frac{1}{n})$ and $f(\mathbf{x}_n) > f(\overline{\mathbf{x}})$ — *i.e.*, $\mathbf{x}_n \in U$ for each n (these points $\mathbf{x}_n$ exist because $\overline{\mathbf{x}}$ is not a local maximum of f). Therefore $(**)$ implies that $\mathbf{p} \cdot \mathbf{x}_n \geqq b$ for each n. Since $\{\mathbf{x}_n\}$ converges to $\overline{\mathbf{x}}$, we have $\mathbf{p} \cdot \overline{\mathbf{x}} \geqq b$. Therefore we have established that $\mathbf{p} \cdot \overline{\mathbf{x}} = b$.

Conclusion (b) of the theorem now follows immediately: we have $\overline{\mathbf{x}} \in S$, and according to $(*)$ we have $\mathbf{p} \cdot \mathbf{x} \leqq b = \mathbf{p} \cdot \overline{\mathbf{x}}$ for all $\mathbf{x} \in S$.

In order to establish (a), suppose that (a) fails to hold — *i.e.*, there is a point $\widehat{\mathbf{x}}$ that satisfies both $\mathbf{p} \cdot \widehat{\mathbf{x}} \leqq \mathbf{p} \cdot \overline{\mathbf{x}} = b$ and $f(\widehat{\mathbf{x}}) > f(\overline{\mathbf{x}})$ — *i.e.*, $\widehat{\mathbf{x}} \in U$. Since U is open (because f is continuous) and nonempty, there is an open ball $B(\widehat{\mathbf{x}}, \epsilon) \subseteq U$; and since $\mathbf{p} \cdot \widehat{\mathbf{x}} \leqq \mathbf{p} \cdot \overline{\mathbf{x}} = b$, the open ball $B(\widehat{\mathbf{x}}, \epsilon)$ contains points $\mathbf{x}$ that satisfy $\mathbf{p} \cdot \mathbf{x} < \mathbf{p} \cdot \overline{\mathbf{x}} \leqq b$, a violation of $(**)$. ■

Here are several observations about the convex optimization theorem we've just proved:

(1) The theorem is an existence theorem: its conclusion (in one direction) says that *there exists* a vector $\mathbf{p}$ that satisfies (a) and (b). And the $\mathbf{p}$ that exists is often interpretable as a list (a vector) of prices, which should be clear in both (a) and (b) — which would then say (a) that $\overline{\mathbf{x}}$ maximizes f among all the alternatives $\mathbf{x}$ whose value (at prices $\mathbf{p}$) does not exceed the value of $\overline{\mathbf{x}}$; and (b) that $\overline{\mathbf{x}}$ maximizes the value of $\mathbf{x}$ (for example, the profit it yields) among all the "feasible" alternatives $\mathbf{x} \in S$. Note that the theorem fits exactly our Robinson Crusoe example, ensuring the existence of "decentralizing" or efficiency prices.

(2) No mention is made of differentiability of f or of differentiability of any functions that might define the set S (see (4) below). So the theorem ensures the existence of a separating $\mathbf{p}$ in a broad range of situations where differentiability is not present.

(3) The function f could be replaced with a quasiconcave, locally nonsatiated, continuous preference relation $\succsim$. In that case (a) would say that $\mathbf{x}$ is maximal in the set $\{\mathbf{x} \,|\, \mathbf{p}\cdot\mathbf{x} \leqq \mathbf{p}\cdot\overline{\mathbf{x}}\}$.

(4) The constraint set S, or feasible set, is often defined by a set of inequality constraints, such as

$$g_i(\mathbf{x}) \leqq c_i \quad (i = 1, \ldots, m)$$

where each function $g_i(\cdot)$ is quasiconvex, as in Figure 7. In particular, the constraints could be linear, as in Figure 8.

Our next theorem is the Second Welfare Theorem (more precisely, the Second Welfare Theorem is the corollary of the next theorem).

Theorem: For each $i \in N = \{1, \ldots, n\}$, let $u^i : X \to \mathbb{R}$ be a continuous quasiconcave function on a convex set $X \subseteq \mathbb{R}^\ell$ that is unbounded above. Assume that for at least one $i \in N$, u^i is strictly increasing (wlog, let this be u^1). Let $(\widehat{\mathbf{x}}^i)_{i \in N}$ be a Pareto allocation for the allocation problem $\big((u^i)_{i \in N}, \mathring{\mathbf{x}}\big)$, where $\mathring{\mathbf{x}} \in \mathbb{R}^\ell_+$. Then there is a price-list $\mathbf{p} \in R^\ell_+$ such that $\forall i \in N : \widehat{\mathbf{x}}^i$ minimizes $\mathbf{p} \cdot \mathbf{x}$ on the upper-contour set $U_i = \{\mathbf{x} \in X \,|\, u^i(\mathbf{x}) \geqq u^i(\widehat{\mathbf{x}}^i)\}$.

Corollary: If $\mathring{\mathbf{x}}^i = \widehat{\mathbf{x}}^i$ for each $i \in N$; if the economy $E = \big(u^i, \mathring{\mathbf{x}}^i\big)_1^n$ satisfies the assumptions of the theorem; and if, for each $i \in N$ there is a bundle $\mathbf{x}^i \in X$ that satisfies $\mathbf{p} \cdot \mathbf{x}^i < \mathbf{p} \cdot \mathring{\mathbf{x}}^i$, then $\big(\mathbf{p}, (\widehat{\mathbf{x}}^i)_N\big)$ is a Walrasian equilbrium of the economy E.

Proof of the Theorem: Because $(\widehat{\mathbf{x}}^i)_N$ is Pareto and u^1 is strictly increasing, we have $\sum_1^n \widehat{\mathbf{x}}^i = \mathring{\mathbf{x}}$. Let U_1° denote the strict upper-contour set $\{\mathbf{x} \in X \,|\, u^1(\mathbf{x}) > u^1(\widehat{\mathbf{x}}^1)\}$, and let U and U° denote the sets

$$U = \sum_{i=1}^n U_i \qquad \text{and} \qquad U^\circ = U_1^\circ + \sum_{i=2}^n U_i.$$

Clearly $\mathring{\mathbf{x}} \in U$, and because $(\widehat{\mathbf{x}}^i)_{i \in N}$ is Pareto, we also have $\mathring{\mathbf{x}} \notin U^\circ$.

We first show that $\mathring{\mathbf{x}} \in bdy\, U^\circ$. Each u^i is continuous, therefore each U_i is closed, and we therefore have $cl\, U_i = U_i$. It's easy to show that $cl\, U_1^\circ = U_1$. Since the sum of the closures of sets is always a subset of the closure of the sum of the sets (by a theorem above), we have

$$U = \sum_{i=1}^n U_i = cl\, U_1^\circ + \sum_{i=2}^n cl\, U_i \subseteq cl\, \big(U_1^\circ + \sum_{i=2}^n U_i\big) = cl\, U^\circ.$$

Since $\mathring{\mathbf{x}} \in U \subseteq cl\, U^\circ$ — *i.e.*, $\mathring{\mathbf{x}} \in cl\, U^\circ$ — and $\mathring{\mathbf{x}} \notin U^\circ$, we have $\mathring{\mathbf{x}} \in bdy\, U^\circ$.

We next show that U° is nonempty and convex. Each u^i is quasiconcave, therefore each set U_i is convex (and is obviously nonempty); U_1° is also convex, and is nonempty because u^1 is strictly increasing and X has no upper bound. Therefore U° is nonempty and convex, as the sum of nonempty convex sets.

Since U° is nonempty and convex and $\mathring{\mathbf{x}} \in bdy\, U^\circ$, the Supporting Hyperplane Theorem ensures that there is a hyperplane that supports U° and contains $\mathring{\mathbf{x}}$ — *i.e.*, there exists a $\mathbf{p} \neq \mathbf{0} \in R^\ell$ such that $\mathbf{p} \cdot \mathbf{x} \geqq \mathbf{p} \cdot \mathring{\mathbf{x}}$ for all $\mathbf{x} \in U^\circ$. Since U is unbounded above, $\mathbf{p} \in R^\ell_+$.

Now let $\mathbf{x} \in U$; we've just shown that $U \subseteq cl\, U^\circ$, so we have $\mathbf{x} \in cl\, U^\circ$, and therefore there is a sequence $\{\mathbf{x}(k)\}$ in U° that converges to $\mathbf{x}$. Since each term $\mathbf{x}(k)$ is in U°, it satisfies $\mathbf{p} \cdot \mathbf{x}(k) \geqq \mathbf{p} \cdot \mathring{\mathbf{x}}$, and therefore the sequence's limit, $\mathbf{x}$, satisfies $\mathbf{p} \cdot \mathbf{x} \geqq \mathbf{p} \cdot \mathring{\mathbf{x}}$. Thus, every $\mathbf{x} \in U$ satisfies $\mathbf{p} \cdot \mathbf{x} \geqq \mathbf{p} \cdot \mathring{\mathbf{x}}$ — *i.e.*, $\mathring{\mathbf{x}}$ minimizes $\mathbf{p} \cdot \mathbf{x}$ on U. The Sum-of-Sets Maximization Theorem therefore guarantees that for every $i \in N$, $\widehat{\mathbf{x}}^i$ minimizes $\mathbf{p} \cdot \mathbf{x}$ on the set U_i. ■

Some Additional Theorems

Here are some additional definitions and theorems that are important and useful.

Definition: The **convex hull** of a set S, denoted conv$S$, is the intersection of all convex sets that contain S.

Remark: For any set S, conv$S$ is convex (as the intersection of a collection of convex sets) and is therefore the smallest convex set containing S.

Definition: A **convex combination** of a finite set $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of points is a linear combination of the points, with the restriction that the coefficients are non-negative and sum to 1 — *i.e.*, a point $\mathbf{x}$ of the form $\mathbf{x} = \lambda_1\mathbf{x}_1 + \cdots + \lambda_m\mathbf{x}_m$ for some scalars $\lambda_1, \ldots, \lambda_m$ that satisfy $\lambda_i \geqq 0$ for all i and $\sum \lambda_i = 1$. (Important note: As with linear combinations, convex combinations are only defined for finite sets of vectors.)

Remark: The convex hull of a set S is the set of all convex combinations of points in S.

A point $\mathbf{x} \in \text{conv}S$ is therefore a convex combination of a finite number of points in S; but the required number of points for a particular $\mathbf{x}$ might be very large. The following theorem ensures that any point in the convex hull of S can actually be expressed as a convex combination of a small number of points in S.

Caratheodory's Theorem: Any point in the convex hull of a set $S \subseteqq \mathbb{R}^n$ is a convex combination of at most $n+1$ points in S.

This result is easy to see for any finite set S in $\mathbb{R}^2$: if S has m elements, the convex hull of S is a polygon with no more then m sides, and it's easy to see that any point in the polygon can be expressed as a convex combination of no more than three of the vertices of the polygon — *i.e.*, no more than three of the members of S.

Definition: An **extreme point** of a convex set S is a point in S that is not a convex combination of any other points in S.

Examples: In $\mathbb{R}^2$, if S is a polygon, then the extreme points of S are the vertices of S. In $\mathbb{R}^n$, if S is a closed ball (using the Euclidean norm!), then every point in S is an extreme point; and if S is an open convex set, then S has no extreme points.

Krein-Milman Theorem: A compact convex set is the convex hull of its extreme points.

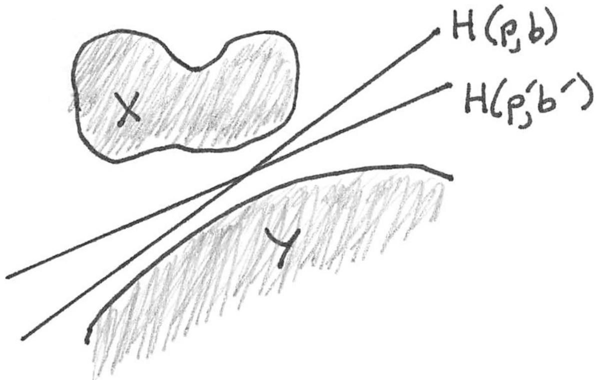

Figure 1

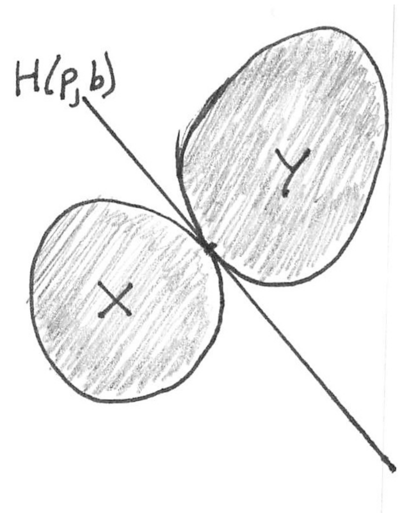

Figure 2

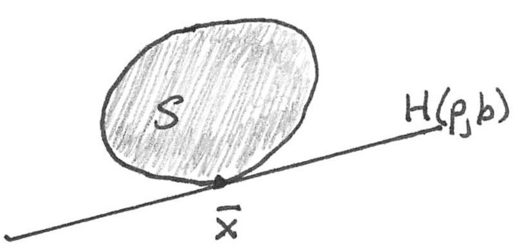

Figure 3

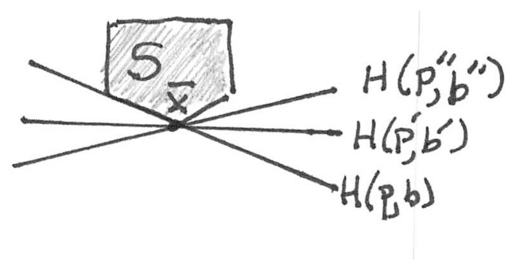

Figure 4

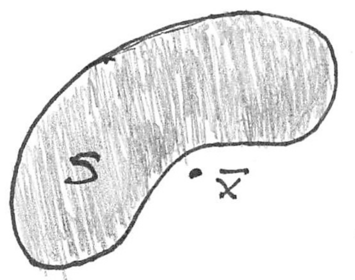

Figure 5

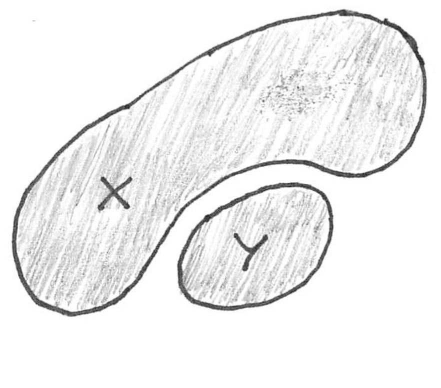

Figure 6

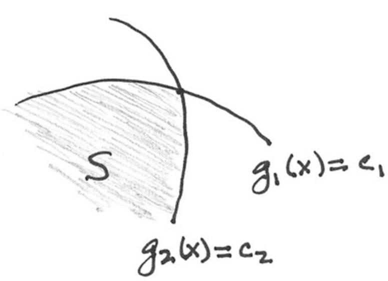

Figure 7

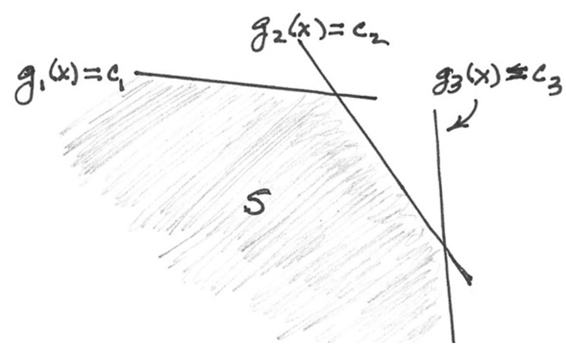

Figure 8

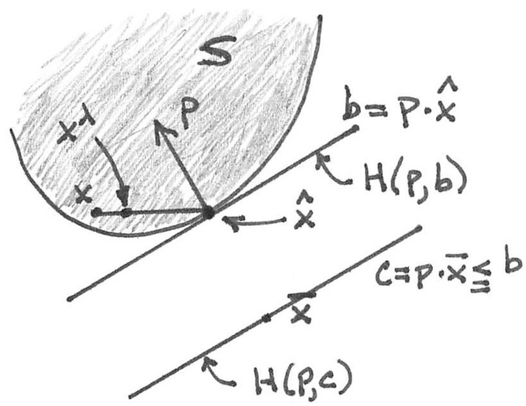

Figure 9