Ariel Zetlin-Jones
September 20, 2007
Notes on Badly Behaved Concave Programming Problems

I present below a lemma which you may find useful in solving "badly" behaved programming problems. Following the lemma, I give an example. I encourage you to think about how you might apply the lemma to Problem 2 from Problem Set 1.

Lemma 1.1. *Let* $D \subset \mathbb{R}^n$ *and* $f : D \to \mathbb{R}$ *a function. Let*

$$\begin{aligned} f^* &\equiv \max\{f(x)|x \in D\} \\ D^* &\equiv \arg\max\{f(x)|x \in D\} \end{aligned}$$

Suppose $\exists\ D_1$ *and* D_2 *such that* $D = D_1 \cup D_2$*, and for* $i = 1, 2$*, let*

$$\begin{aligned} f_i^* &\equiv \max\{f(x)|x \in D_i\} \\ D_i^* &\equiv \arg\max\{f(x)|x \in D_i\}. \end{aligned}$$

Then,

$$f^* \equiv \max\{f_1^*, f_2^*\}$$

$$D^* \equiv \begin{cases} D_1^* & if \quad f_1^* > f_2^* \\ D_1^* \cup D_2^* & if \quad f_1^* = f_2^* \\ D_2^* & if \quad f_2^* > f_1^* \end{cases}$$

if f_i^*, D_i^* *exist for all* i*.*

Note that you can extend this to more than the union of two sets.

Proof. We break the proof down into 3 cases and then prove the relevant set containment (e.g. $D^* \subset D_1^*$ and vice versa:

Case 1 **($f_1^* > f_2^*$).** Let $x^* \in D^*$. Then $f(x^*) = \max\{f(x)|x \in D\}$. Let $x \in D_1$. Then trivially $x \in D$ and we have $f(x^*) \geq f(x)$. Notice that this implies $f_1^* \leq f(x^*)$. Now, suppose that $x^* \notin D_1$. Then $x^* \in D_2$. Thus, $f(x^*) \leq f_2^* < f_1^* \leq f(x^*)$, which is a contradiction. Thus, $x^* \in D_1$, and it follows that $x^* \in D_1^*$.

Now, let $x^* \in D_1^*$. Then, for all $x \in D_1, f(x^*) \geq f(x)$, and for all $x \in D_2$, we have that $f(x) \leq f_2^* < f_1^* = f(x^*)$. Hence, for all $x \in D$, we have $f(x) \leq f(x^*)$, which implies that $x^* \in D^*$. Thus, $D_1^* = D^*$.

Finally, let $\tilde{x} \in D^* = D_1^*$. Then

$$f^* = max\{f(x)|x \in D\} = f(\tilde{x}) = max\{f(x)|x \in D_1\} = f_1^* = \max\{f_1^*, f_2^*\}.$$

Case 2 **($f_2^* > f_1^*$).** This case is analgous to Case 1.

Case 3 ($f_1^* = f_2^*$). Let $x^* \in D^*$. Either $x^* \in D_1$ or $x^* \in D_2$. Without loss of generality, suppose $x^* \in D_1$. Then, for all $x \in D_1, x \in D$, and thus $f(x^*) = \max\{f(x)|x \in D\} \geq f(x)$. Hence, $x^* \in D_1^*$, and thus, we have that $x^* \in D_1^* \cup D_2^*$.

Let $x^* \in D_1^* \cup D_2^*$. Without loss of generality, suppose $x^* \in D_1^*$. Then $f(x^*) = \max\{f(x)|x \in D_1\} = f_1^* = f_2^* = \max\{f(x)|x \in D_2\} \geq f(x)$ for all $x \in D_1, x \in D_2$. Hence, for all $x \in D_1 \cup D_2 = D$, we have that $f(x^*) \geq f(x)$. Thus, $x^* \in D^*$. Hence, $D^* = D_1^* \cup D_2^*$. Thus,

$$f^* = \max\{f(x)|x \in D\} = \max\{f(x)|x \in D_1 \cup D_2\} = \max\{f(x)|x \in D_1\} = f_1^* = f_2^*.$$

□

Example 1.2. Find π^* and s^* given the following production technology:

$$Y = \left\{(y_1, y_2)|y_1 \leq 0, y_2 \leq \max(\sqrt{-y_1} - 2, 0)\right\}.$$

I will restrict my attention to positive prices (specifically, I assume that at least one of p_1 and p_2 is nonzero, else the problem is boring), and it is immediate that at an optimum, $y_2 = \max(\sqrt{-y_1} - 2, 0)$ (since otherwise we would be able to increase y_2 and strictly increase profits. Thus, we may consider the following problem:

$$\begin{aligned} &\max_{y_1} \quad p_1 y_1 + p_2 \max(\sqrt{-y_1} - 2, 0) \\ &\text{s.t.} \\ &y_1 \leq 0 \end{aligned}$$

Hence, re-interpreting this problem, let $x = y_1$, $f(x) = p_1 x + p_2 \max(\sqrt{-x} - 2, 0)$, and $D = \mathbb{R}_-$. The natural choice of D_1 and D_2 are

$$D_1 = [-4, 0], D_2 = (-\infty, -4].$$

Obviously, $D = D_1 \cup D_2$. Then, we can write the following:

$$\begin{aligned} f_1^* &= \max\{p_1 y_1 | y_1 \in D_1\} \\ D_1^* &= \arg\max\{p_1 y_1 | y_1 \in D_1\} \end{aligned}$$

$$\begin{aligned} f_2^* &= \max\{p_1 y_1 + p_2(\sqrt{-y_1} - 2) | y_1 \in D_2\} \\ D_2^* &= \arg\max\{p_1 y_1 + p_2(\sqrt{-y_1} - 2) | y_1 \in D_2\} \end{aligned}$$

Then, it is obvious that $f_1^* = 0$ for any positive price, and

$$D_1^* = \begin{cases} 0 & \text{if } p_1 > 0 \\ D_1 & \text{if } p_1 = 0 \end{cases}$$

The second set is a bit more complicated. First, it is obvious that no solution exists if $p_1 = 0$ and $p_2 > 0$. If $p_2 = 0$ and $p_1 > 0$, then it is clear that $D_2^* = -4$. Now, for strictly positive prices, I first drop the constraint that $y_1 \in D_2$. Using first order conditions, you'll find that

$$y_1^* = \frac{-p_2^2}{4p_1^2}.$$

Notice, this tells us that the constraint on y_1 is non-binding (that is, $y_1 < -4$) only if $p_2 > 4p_1$. If this is not the case, then the constraint binds and it must be that $y_1^* = -4$. Notice, then, that when the constraint on y_1 is non-binding, we have the following formula for profits:

$$f_2^* = p_1 \frac{-p_2^2}{4p_1^2} + p_2(\frac{p_2}{2p_1} - 2) = \frac{p_2^2 - 8p_1p_2}{4p_1}$$

when $p_1 > 4p_2$. Eventually, though, we'll want to compare profits under D_1 to profits under D_2, so let's point out that in this region, $f_2^* \geq 0$ only if $p_2 > 8p_1$. Notice, in this region of prices, we know that the constraint $y_1 < -4$ is non-binding. Now, we can write f_2^* as

$$f_2^* = \begin{cases} \frac{p_2^2 - 8p_1p_2}{4p_1} & \text{if } p_2 > 4p_1 \\ -4p_1 & \text{if } 4p_1 \geq p_2 \end{cases}$$

and

$$D_2^* = \begin{cases} \frac{-p_2^2}{4p_1^2} & \text{if } p_2 > 4p_1 \\ -4 & \text{if } 4p_1 \geq p_2 \end{cases} .$$

Now, combining the results (for strictly positive prices), we see that

$$\pi^*(p) = \begin{cases} \frac{p_2^2 - 8p_1p_2}{4p_1} & \text{if } p_2 > 8p_1 \\ 0 & \text{o/w} \end{cases}$$

and

$$s^*(p) = \begin{cases} \left(\frac{-p_2^2}{4p_1^2}, \frac{p_2}{2p_1} - 2\right) & \text{if } p_2 > 8p_1 \\ \left\{\left(\frac{-p_2^2}{4p_1^2}, \frac{p_2}{2p_1} - 2\right), (0,0)\right\} & \text{if } p_2 = 8p_1 \\ (0,0) & \text{o/w} \end{cases} .$$

Ariel Zetlin-Jones
September 13, 2007
Notes on Separating Hyperplanes

1 Preliminaries

1.1 Notation

Let X be an arbitrary subset of $\mathbb{R}^n$. I will denote the closure of X in $\mathbb{R}^n$ by $\overline{X}$ and the boundary of X by ∂X. Obviously, since we are in $\mathbb{R}^n$, multiplication represents scalar dot products (I have omitted the dots everywhere).

Fix $x, x' \in \mathbb{R}^n$. Then $x \geq x'$ if and only if $x_i \geq x'_i$ for all $i = 1, \ldots, n$, and $x >> x'$ if and only if $x_i > x'_i$ for all $i = 1, \ldots, n$.

When I define objects, I usually use $\equiv$ instead of $=$.

1.2 Definitions

Definition 1.1. *Hyperplane.* Let $p \in \mathbb{R}^n, p \neq 0, \|p\| < \infty$ and $\alpha \in \mathbb{R}$. The set $H = \{x \in \mathbb{R} | px = \alpha\}$ is called a *hyperplane* in $\mathbb{R}^n$ with normal p.

In $\mathbb{R}^2$, any line defines a hyperplane. Moreover, if x^* and y^* are elements of hyperplane H, then for some $\alpha \in \mathbb{R}$, $px^* = py^* = \alpha$. Thus, we have $p(x^* - y^*) = 0$, which implies that p is orthogonal to $(x^* - y^*)$.

Definition 1.2. *X, Y Separated by H.* Let X, Y be two nonempty subsets of $\mathbb{R}^n$. We say X and Y are *separated by the hyperplane* $H = \{x \in \mathbb{R} | px = \alpha\}$ if and only if

$$px \leq \alpha \qquad \forall x \in X$$
$$\text{and}$$
$$py \geq \alpha \qquad \forall y \in Y.$$

Note that the hyperplane H divides $\mathbb{R}^n$ into two closed half-spaces:

$$\{x \in \mathbb{R} | px \geq \alpha\} \text{ and } \{x \in \mathbb{R} | px \leq \alpha\}.$$

Definition 1.3. *Bounding Hyperplane.* Let X be a nonempty subset of $\mathbb{R}^n$. The hyperplane H is said to be *bounding for* X if and only if

$$X \subseteq \{x \in \mathbb{R} | px \geq \alpha\} \text{ or } X \subseteq \{x \in \mathbb{R} | px \leq \alpha\}.$$

Definition 1.4. *Supporting Hyperplane.* If H is bounding for X and $\exists z \in H$ such that $z \in \partial X$, then H is said to be a *Supporting Hyperplane* to X.

2 Separating Hyperplane Theorems

2.1 X nonempty, closed, convex, $x_0 \notin X$

Since we will consider X is closed, we have $X = \overline{X}$. Then, we have the following theorem.

Theorem 2.1. *Let $\overline{X}$ be a nonempty, closed, convex subset of $\mathbb{R}^n$ and let $x_0 \notin \overline{X}$. Then*

1. *$\exists a \in \overline{X}$ such that $d(x_0, a) \leq d(x_0, x) \forall x \in \overline{X}$ and $d(x_0, a) > 0$.*
2. *$\exists \alpha \in \mathbb{R}$ and $p \in \mathbb{R}^n$ with $p \neq 0, \|p\| < \infty$ such that $px \geq \alpha \ \forall x \in \overline{X}$ and $px_0 < \alpha$.*

Proof. We prove this in parts:

1. Let $\overline{B}(x_0)$ be a closed ball with center x_0. That is, $\overline{B}(x_0) = \{x \in \mathbb{R}^n | d(x_0, x) \leq \epsilon\}$ for some $\epsilon > 0$. Choose ϵ large enough such that $\overline{B}(x_0) \cap \overline{X} \neq \emptyset$. Define $A \equiv \overline{B}(x_0) \cap \overline{X}$. Then A is nonempty, closed (it is the intersection of closed sets), and bounded ($A \subseteq \overline{B}(x_0)$ which is trivially bounded). Hence, A is compact. As $d(x_0, x)$ is a continuous function of x, there exists $a \in A$ such that
$$d(x_0, a) = \min_{a' \in A} d(x_o, a')$$
by the Weierstrass theorem. Hence, for all $x \in A, d(x_0, a) \leq d(x_0, x)$. In fact, then, for all $x \in X, d(x_0, a) \leq d(x_0, x)$ since for any $x \in X$ with $x \notin A$, we must have $d(x_0, x) > \epsilon \geq d(x_0, a)$. Moreover, since $a \in \overline{X}$, $x_0 \notin \overline{X}$, it must be that $d(x_0, a) > 0$.

2. In what follows, a is precisely the the element of $\overline{X}$ we found in part 1 which had minimum distance from x_0. Now, define $p \equiv a - x_0$ and $\alpha \equiv pa$. Clearly, since $a \neq x_0, p \neq 0$ and $\|p\|$ is finite.

   Comment 2.2. Note that I could just as easily define $\tilde{p} \equiv \frac{a - x_0}{\|a - x_0\|}$ and $\tilde{\alpha} = \tilde{p}a$, and the same proof would work. Thus, we may, without loss of generality, restrict p such that $\|p\| = 1$. Formally, $\tilde{p} = \frac{p}{\|p\|}$. Then,
$$\|\tilde{p}\| = \sqrt{\sum_{i=1}^{n} \left(\frac{p_i}{\|p\|} \right)^2} = \sqrt{\frac{1}{\|p\|^2} \sum_{i=1}^{n} p_i^2} = \frac{1}{\|p\|} \sqrt{\sum_{i=1}^{n} p_i^2} = \frac{\|p\|}{\|p\|} = 1.$$

   Returning to the proof, we need to show $px \geq \alpha \ \forall x \in \overline{X}$ and $px_0 < \alpha$. First, note that
$$\begin{aligned} px_0 &= (a - x_0)x_0 \\ &= (a - x_0)x_0 - a(a - x_0) + a(a - x_0) \\ &= -(a - x_0)(a - x_0) + (a - x_0)a \\ &= -\|p\|^2 + \alpha \\ &< \alpha \text{ since } \|p\| \geq 0. \end{aligned}$$

Second, fix $x \in \overline{X}$. Since $a \in \overline{X}$ and $\overline{X}$ is convex, for all $t \in (0, 1]$

$$x(t) \equiv (1-t)a + tx \in \overline{X}.$$

By part 1, we immediately have that $d(x_0, a) \leq d(x_0, x(t))$. Thus,

$$\begin{aligned}
\|a - x_0\|^2 &\leq \|x(t) - x_0\|^2 \\
&= \|(1-t)(a - x_0) + t(x - x_0)\}^2 \\
&= (1-t)^2\|a - x_0\|^2 + 2t(1-t)(a - x_0)(x - x_0) + t^2\|x - x_0\|^2. \\
& \text{And, so} \\
0 &\leq (t^2 - 2t)\|a - x_0\|^2 + 2t(1-t)(a - x_0)(x - x_0) + t^2\|x - x_0\|^2 \\
&= (t-2)\|a - x_0\|^2 + 2(1-t)(a - x_0)(x - x_0) + t\|x - x_0\|^2.
\end{aligned}$$

Taking limits as $t \to 0$ and then manipulating the inequalities, yields

$$\begin{aligned}
& 0 \leq -2\|a - x_0\|^2 + 2(a - x_0)(x - x_0) \\
\Rightarrow\ & 0 \geq (a - x_0)(a - x_0) - (a - x_0)(x - x_0) \\
\Rightarrow\ & 0 \geq (a - x_0)a - (a - x_0)x \\
\Rightarrow\ & (a - x_0)x = px \geq \alpha = pa = (a - x_0)a
\end{aligned}$$

This completes the proof. □

Note that in our construction of a separating hyperplane, we in fact used $\alpha = pa$. Hence, $a \in \overline{X}$ and $a \in H$. I'm pretty sure that a must be on the boundary of $\overline{X}$ (that is, $a \in \partial\overline{X}$) since otherwise you should be able to get closer to x_0, but I am not proving that rigorously here (you should prove it!). Assuming that's correct, though, we have that H is in fact a supporting hyperplane to $\overline{X}$.

2.2 X nonempty, convex, $x_0 \notin X$

Theorem 2.3. *Let $X \subseteq \mathbb{R}^n$ be nonempty and convex (not necessarily closed). Fix $x_0 \in \mathbb{R}^n$ with $x_0 \notin X$. Then $\exists p \in \mathbb{R}^n$ with $p \neq 0, \|p\| < \infty$ such that $px \geq px_0$ $\forall x \in X$.*

Proof. Suppose $x_0 \notin \overline{X}$. Then, by the previous theorem, as $\overline{X}$ is closed, there exists H that separates $\overline{X}$ and x_0. Thus, there exists p satisfying the above properties such that for all $x \in \overline{X}$, $px > px_0$. In particular, as $X \subseteq \overline{X}$, for all $x \in X$, $px > px_0$. Hence, H also separates X and x_0.

Now, suppose $x_0 \in \overline{X}$. Then it must be that $x_0 \in \partial\overline{X}$ (since $x_0 \notin X$). Since x_0 is a boundary point of $\overline{X}$, any neighborhood of x_0, contains a point that is not in $\overline{X}$. So, we may define a sequence $\{\epsilon_n\}$ such that $\epsilon_n \to 0$ as $n \to \infty$. Then, define $B_{\epsilon_n}(x_0) \equiv \{x \in \mathbb{R}^n | d(x_0, x) < \epsilon_n\}$. For each ϵ_n, $\exists x_{\epsilon_n} \in B_{\epsilon_n}(x_0)$ such that $x_{\epsilon_n} \notin \overline{X}$. Notice, as $\epsilon_n \to 0$, we have $x_{\epsilon_n} \to x_0$. Moreover, since $x_{\epsilon_n} \notin \overline{X}$, we may again apply the previous theorem. Thus, for each ϵ_n, $\exists p_{\epsilon_n}$ with appropriate properties such that for all $x \in \overline{X}, p_{\epsilon_n}x > p_{\epsilon_n}x_{\epsilon_n}$. As stated in the previous proof, we may assume without loss of generality that $\|p_{\epsilon_n}\| = 1$ which ensures that each p_{ϵ_n} lies in the unit sphere in $\mathbb{R}^n$. So the sequence $\{p_{\epsilon_n}\}$ lies in a compact set, and, therefore, it has a convergent subsequence, $\{p_{\epsilon_{n_k}}\}$ such that $\{p_{\epsilon_{n_k}}\} \to p$ with $\|p\| = 1$. Now, everywhere along the subsequence $\{\epsilon_{n_k}\}$, we have

$$\forall x \in \overline{X}, p_{\epsilon_{n_k}} x > p_{\epsilon_{n_k}} x_{\epsilon_{n_k}}.$$

Taking limits makes the strict inequality weak and we have that for all $x \in \overline{X}, px \geq px_0$. Again, since we have succeeded in separating $\overline{X}$ from x_0, we have in fact separated X from x_0. □

2.3 X, Y nonempty, convex, empty intersection

Theorem 2.4. *(Minkowsky) Let X and Y be two nonempty, convex subsets of $\mathbb{R}^n$ with $X \cap Y = \emptyset$. Then $\exists \alpha \in \mathbb{R}$ and $p \in \mathbb{R}^n$ with $p \neq 0, \|p\| < \infty$ such that $px \geq \alpha \geq py$ for all $x \in X$ and $y \in Y$.*

Proof. Define the set $S \equiv X + (-Y)$ by

$$S \equiv \{z \in \mathbb{R}^n | z = x - y \text{ for some } x \in X, y \in Y\}.$$

Now, notice that $0 \notin S$. To see this, suppose it is. Then $0 = x - y$ for some $x \in X, y \in Y$. But then $x = y$ which implies that $x \in X \cap Y$, a contradiction. You can show that S is nonempty, and convex, and thus by the previous theorem, $\exists p \in \mathbb{R}^n$ with $p \neq 0, \|p\| < \infty$ such that $pz \geq p \cdot 0 = 0$ for all $z \in S$. Since $z = x - y$, we have $px \geq py$ for all $x \in X$, $y \in Y$. In particular, it must be that

$$\inf_{x \in X} px \geq \sup_{y \in Y} py.$$

Thus, we can choose an $\alpha \in \mathbb{R}$ such that $px \geq \alpha \geq py$ for all $x \in X, y \in Y$. □

3 Application to Efficient Production Plans

Recall the following definition of an efficient production plan:

Definition 3.1. *Efficient Production Plan.* Let Y be a production set. Then $y \in Y$ is said to be *efficient* if and only if $\nexists y' \in Y$ such that $y' \geq y$ and $y' \neq y$.

Then, we have the following proposition:

Proposition 3.2. *Suppose the production set Y is a nonempty and convex subset of $\mathbb{R}^L$. If $y \in Y$ is efficient, then $\exists p \in \mathbb{R}^L, p \geq 0$ such that for all $y' \in Y, py \geq py'$.*

Proof. Suppose $y \in Y$ is efficient. Then, consider the set

$$P_y \equiv \{y' \in \mathbb{R}^L | y' >> y\}.$$

Clearly, P_y is nonempty and convex. You can check for yourself that since y is efficient, $Y \cap P_y = \emptyset$. Thus, we may apply Minkowsky's Separating Hyperplane theorem. That is, $\exists p \in \mathbb{R}^L$ with $p \neq 0, \|p\| < \infty$ such that $py' \geq py''$ for all $y' \in P_y$ and $y'' \in Y$.

In particular, since $y \in Y$, we have that for all $y' \in P_y$, $py' \geq py$. Now, rewrite this inequality as

$$p_j(y'_j - y_j) + \sum_{i \neq j} p_i(y'_i - y_i) \geq 0.$$

Now, suppose that $p_j < 0$. Then, it is clear that, keeping y'_i fixed for $i \neq j$, we can increase y'_j arbitrarily until we violate the above constraint. So it must be the case that for each i, $p_i \geq 0$. Thus, $p \geq 0$.

Moreover, if we let $\epsilon^L = (\epsilon, \ldots, \epsilon)$ (that is, if ϵ^L denotes unit vector in $\mathbb{R}^L$ scaled by $\epsilon > 0$, then it is clear that the production plan $y + \epsilon^L \in P_y$. Thus, for all $y'' \in Y$, we have $p(y + \epsilon) \geq py''$. Now, taking limits as $\epsilon \to 0$, we find that $py \geq py''$ for all $y'' \in Y$ as desired. □

4 References

1. Mas-Collel

2. Takayama, Akira, *Mathematical Economics,* 1985, Ch. 0.B.