Mathematical Programming Glossary Supplement: Response Space†

Harvey J. Greenberg
University of Colorado Denver
hjgreenberg@gmail.com

June 8, 2014

Several structural questions have been successfully studied in modern times by examining the range of the given $(m+1)$-vector function (f, g) on S. One of these questions deals with the response of marginal changes in certain parameters, or more directly with the functions, themselves. Translated into mathematics this question becomes one of continuity, and several results in topology have been used by Berge (1959), Dantzig, Folkman and Shapiro (1967), Rockafellar (1966), Evans and Gould (1969), and by Greenberg and Pierskalla (1970,-72).

Another important question deals with convexity structure and providing a geometrical foundation for Lagrangian theory (c.f., Chapters 3 and 4). Rockafellar's book, *Convex Analysis* (1970), which is based largely upon his own research, describes a particularly rich theory for convex programs; Falk (1967), Greenberg (1968–73), and others have extended those results to study more general nonlinear programs.

Our intention in this chapter is to examine major elements of convexity, continuity and differentiability. To begin, define the *set of feasible right-hand-sides* as follows:

$$B = \{b \in \mathbb{R}^m \mid g(x)\,\rho\, b \text{ for some } x \in S\}.$$

Equivalently, B = Range of $g - K$, where K is the convex cone given by

$$K = \{y \in \mathbb{R}^m \mid y\,\rho\, 0\}.$$

Next define the *feasibility map*, $\varphi : B \mapsto 2^S$, where

$$\varphi(b) = \{x \in S \mid g(x)\,\rho\, b\},$$

and the *value function*, $f^* : B \mapsto \mathbb{R}_\infty$, where

$$f^*(b) = \mathrm{Sup}\{f(x) \mid x \in \varphi(b)\}.$$

The third fundamental response map is the *optimality map*, $\varphi^* : B \mapsto 2^S$, where

$$\varphi^*(b) = \{x^* \in \varphi(b) \mid f(x^*) = f^*(b)\}.$$

†This is a transcription of Chapter 5 of my 1969 NLP notes (with some references updated in 1974) for the mathematical program, $P: \ \mathrm{Max}\,\{f(x) \mid g(x)\,\rho\, 0,\ x \in S\}$, where $S \subseteq \mathbb{R}^n$, $f : S \mapsto \mathbb{R}$, $g : S \mapsto \mathbb{R}^m$, and ρ is an m-tuple of relations, $\{\leq, \geq, =\}$.

CONVEXITY THEOREM: If P is a convex program, then f^* is concave on B.

Proof: Let b^1, b^2 be arbitrary members of B, and let $b = \alpha b^1 + (1-\alpha)b^2$ for some arbitrary $\alpha \in (0,1)$. There are two members of S, say x^1 and x^2, such that $g(x^i)\,\rho\, b^i$ for $i = 1, 2$. Define $x = \alpha x^1 + (1-\alpha)x^2$. Since P is a convex program, $g(x)\,\rho\,\alpha g(x^1) + (1-\alpha)g(x^2)$. Hence, $g(x)\,\rho\, b$, which proves that $b \in B$, so B is convex. Further, concavity of f yields $f^*(b) \geq f(x) \geq \alpha f(x^1) + (l-\alpha)f(x^2)$. Let $\{u^k\} \subseteq \varphi(b^1)$ and $\{v^k\} \subseteq \varphi(b^2)$, with $f(u^k) \to f^*(b^1)$ and $f(v^k) \to f^*(b^2)$. Then, define $x^k = \alpha u^k + (1-\alpha)v^k$. As above, it follows that $\{x^k\} \subseteq \varphi(b)$ and $f(x^k) \geq \alpha f(u^k) + (l-\alpha)f(v^k)$ for all k; thus, $f^*(b) \geq \liminf_{k\to\infty} f(x^k) \geq \alpha f^*(b^1) + (1-\alpha)f^*(b^2)$, so f^* is concave on B. $\square$

Exercise. Suppose that g satisfies the convexity conditions of a convex program, but f is only quasi-concave on S. Prove that f^* is quasi-concave on B.

It is not necessary that P be a convex program in order that f^* be concave on B. This is evidenced by the following

Example: $S = \{x \in \mathbb{R}^2_+ \mid x_1 x_2 \neq 0\}$, $f(x) = -x_1/x_2$, ρ is $\leq$, and $g(x) = x_1 x_2$. Then, $B = (0, \infty)$ and $f^*(b) = 0$.

Observe that if we transform the variables (one-to-one) as $x_j = \exp(z_j)$, then S becomes all of $\mathbb{R}^2$ and we have a convex program in the new variables.[1] Thus, the Convexity Theorem can be extended to assume that P is equivalent to a convex program in the sense that a transformation of variables yields a convex program while the response functions remain the same.[2]

Recall from Chapter 3 that Slater's interiority condition assumes there exists x in R^n such that $g(x) < 0$, where g is assumed to be convex. Translating this into a condition in response space, Slater's condition becomes the assumption that b is in the interior of B: i.e., $\text{int}\,(B) = \{b \mid g(x) < b \text{ for some } x \in S\}$ when ρ_i is $\leq$ for each i. An important property of convex or concave functions is that they are continuous at interior points of their domain. Therefore, when P is a convex program satisfying Slater's interiority condition, then f^* is continuous at the specified resource limit (which we have conventionally taken to be zero for a fixed program, but here we allow variation).

To study continuity properties further we must define certain topology concepts associated with point-to-set maps (e.g, feasibility map). Let S be any subset of $\mathbb{R}^n$; we denote its ε-neighborhood by $\eta_\varepsilon(S)$, which is the union of the point neighborhoods as

$$\eta_\varepsilon(S) = \bigcup_{x\in S} N_\varepsilon(x).$$

Let $X, Y \subseteq \mathbb{R}^n$ and let $M : X \mapsto 2^Y$ be a point-to-set map from X into the set of subsets of Y. We say that M is upper (lower) semi-continuous, which we abbreviate usc (lsc), at $x \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $z \in N_\delta(x)$ implies $M(z) \subseteq \eta_\varepsilon(M(x))$ $\big(M(x) \subseteq \eta_\varepsilon(M(z))\big)$.

[1] Take log, so P is equivalent to: $\text{Max}\{-\exp^{z_1 - z_2} \mid z_1 + z_2 \leq \beta \stackrel{\text{def}}{=} \log b\}$.

[2] The point is that we may have favorable properties in response space even if those properties do not hold in decision space.

VALUE CONTINUITY THEOREM: Suppose φ is usc (lsc) at b in B and f is upper (lower) semi-continuous throughout S. Then, f^* is upper (lower) semi-continuous at b.

Proof: Suppose $b^k \to b$, where $\{b^k\} \subseteq B$. Given $\varepsilon > 0$, choose $\delta > 0$ (which depends on ε) such that $z \in N_\varepsilon(x)$ implies $f(x) \leq f(z) + \varepsilon$ (using the assumed upper semi-continuity of f). Then choose k^* (which depends on δ) such that $\varphi(b^k) \subseteq \eta_\varepsilon(\varphi(b))$ for $k \geq k^*$ (using the upper semi-continuity of φ at b). Then, for all $k \geq k^*$ there exists $z^k \in \varphi(b)$ such that $x^k \in \varphi(b^k)$ implies $\left|\left|x^k - z^k\right|\right| < \delta$ and $f(x^k) \leq f(z^k) + \varepsilon$ for all $k \geq k^*$. Since $f(z^k) \leq f^*(b) + \varepsilon$, we then have $f(x^k) \leq f^*(b) + \varepsilon$ for all $k \geq k^*$. Since ε was arbitrary, and since the above is true for all x^k in $\varphi(b^k)$, it follows that for any $\varepsilon > 0$ there exists k^* such that $k \geq k^*$ implies $f^*(b^k) \leq f^*(b) + \varepsilon$, so f^* is upper semi-continuous at b. The lower semi-continuity follows mutatis mutandis. □

We now wish to consider the continuity properties of the feasibility map since they are needed in the Value Continuity Theorem. The following two pathologies illustrate the sort of behavior that could lead to discontinuities.

Example A[3]: Let $f^*(b) = \text{Sup}\{x \mid \text{Min}\{-1 + 2|x+1|,\ e^{-x}\} \leq b\}$, so $B = [-1, \infty)$ and the response maps are as follows:

$$\varphi(b) = \begin{cases} [-\frac{b+3}{2}, \frac{b-1}{2}] & \text{if } b \in [-1, 0] \\ [-\frac{b+3}{2}, \frac{b-1}{2}] \cup [-\ln b, \infty) & \text{if } b > 0 \end{cases}$$

$$f^*(b) = \begin{cases} \frac{b-1}{2} & \text{if } b \in [-1, 0] \\ \infty & \text{if } b > 0 \end{cases}$$

$$\varphi^*(b) = \begin{cases} \{\frac{b-1}{2}\} & \text{if } b \in [-1, 0] \\ \emptyset & \text{if } b > 0 \end{cases}$$

Note that φ is lsc but not usc at 0, and f^* is lower but not upper semi-continuous at 0.

Example B: Let $f^*(b) = \text{Sup}\{x \mid \text{Min}\{x, 0\} \leq b \text{ and } x \in S\}$, where $S = [-10, 10]$. Then, $B = [-10, \infty)$ and the response maps are as follows:

$$\varphi(b) = \begin{cases} [-10,\ b] & \text{if } b < 0 \\ [-10,\ 10] & \text{if } b \geq 0 \end{cases}$$

$$f^*(b) = \begin{cases} b & \text{if } b < 0 \\ 10 & \text{if } b \geq 0 \end{cases}$$

$$\varphi^*(b) = \begin{cases} \{b\} & \text{if } b < 0 \\ \{10\} & \text{if } b \geq 0 \end{cases}$$

Note that φ is usc but not lsc at $b = 0$, and f^* is upper but not lower semi-continuous at $b = 0$.

[3]Edited

Before proceeding with continuity properties of the feasibility map let us point out that the continuity of f^* implies a more general continuity property. Suppose we let $\{g^k\}$ be a sequence of functions on S converging uniformly to g. Then, for any $\delta > 0$ there exists k^*, such that $k \geq k^*$ implies $F(g^k) \subseteq \varphi(b)$ for some $b \in \mathbb{N}_\delta(0)$. If φ is usc at 0 then $\varphi(b) \subseteq \eta_\varepsilon \varphi(0) = \eta_\varepsilon(F(g))$. We conclude that $F(g^k) \subseteq \eta_\varepsilon(F(g))$ for $k \geq k^*$, where k^* depends upon ε. Thus, F is usc at g if φ is usc at 0. A similar result is true for lower semi-continuity except an interiority condition is needed. We shall not detail that result here.

UPPER SEMI-CONTINUITY THEOREM FOR FEASIBILITY MAP:[4] Assume $\varphi(b)$ is bounded and g is ρ-continuous. Then, φ is usc at b if, and only if, there exists $\varepsilon > 0$ and a bounded set D such that $\varphi(\beta) \subseteq D$ for all $\beta \in N_\varepsilon(b)$.

Proof: Let $b^k \to b$, where $\{b^k\} \subseteq B$. Suppose the bounded set D exists to satisfy $\varphi(\beta) \subseteq D$ for all β in $N_\varepsilon(b)$. Then, assume to the contrary that there exists $\{x^k\}$ such that $x^k \in \varphi(b^k)$ and $\left|\left|x^k - z\right|\right| \geq \delta$ for all $z \in \varphi(b)$. There is a limit point x, and since g is ρ-continuous, x is in $\varphi(b)$. This yields the contradiction that $\left|\left|x^k - x\right|\right| \geq \delta$. Conversely, if φ is usc at b, then simply let $D = \eta_\delta(\varphi(b))$. □

Corollary: If P is a closed program and S is compact, then f^* is upper semi-continuous throughout B.

Exercise: Prove that B is a closed set if g is ρ-continuous and S is compact.

We now consider lower semi-continuity for the restricted case where ρ_i is $\leq$ for all $i = 1, \ldots, m$. Let $\varphi^0(b) = \{x \mid g(x) < b\}$ denote the *strict interior* of $\varphi(b)$. Recall that $\operatorname{int}(B) = \{b \mid \varphi^0(b) \neq \emptyset\}$. An extension to allow linear equations is stated as an exercise later.

LOWER SEMI-CONTINUITY OF FEASIBILITY MAP: Suppose ρ is $\leq$, b is in $\operatorname{int}(B)$ and $\varphi(b)$ is bounded. Then, φ is lsc at b if, and only if, $\operatorname{cl}(\varphi(b)) = \operatorname{cl}\left(\varphi^0(b)\right)$.

Proof: Let $b^k \to b$, where $\{b^k\} \subseteq B$. Suppose φ is not lsc at b, so there exists $\{x^k\} \subseteq \varphi(b)$ such that for some $\delta > 0$, we have $\left|\left|x^k - z\right|\right| \geq \delta$ for all z in $\varphi(b^k)$ for k sufficiently large. Since $\varphi(b)$ is bounded, there is a limit point x in $\operatorname{cl}(\varphi(b))$. Then, the closure condition in the hypothesis implies there exists $\{y^k\} \subseteq \varphi^0(b)$ such that $y^k \to x$. For k sufficiently large $y^k \in \varphi(b^k)$, so we obtain the contradiction that $\left|\left|x^k - y^k\right|\right| \geq \delta$. Conversely, if φ is lsc at b and x is in $\operatorname{cl}(\varphi(b))$, then merely choose $b^k = b - 1/k$ and note that $\varphi(b) \subseteq \eta_\delta(\varphi(b^k)) \subseteq \eta_\delta(\varphi^0(b))$ for $k \geq k^*$. This shows there exists $\{y^k\} \subseteq \varphi^0(b)$ such that $y^k \to x$, so $\operatorname{cl}\left(\varphi^0(b)\right) \supseteq \operatorname{cl}(\varphi(b))$. The reverse inclusion is always true, so $\operatorname{cl}\left(\varphi^0(b)\right) = \operatorname{cl}(\varphi(b))$. □

We shall now identify some classes of programs that satisfy the closure property in the Lower Semi-continuity Theorem for the Feasibility Map (viz., $\operatorname{cl}\left(\varphi^0(b)\right) = \operatorname{cl}(\varphi(b))$). Let us suppose g is explicitly quasi-convex on S; i.e., S is a convex set; x, y in S and w in $[x, y]$ imply $g_i(w) \leq \operatorname{Max}\{g_i(x), g_i(y)\}$ $(g_i(x) \neq g_i(y))$ for each $i = 1, \ldots, m$. Suppose x is in $\operatorname{cl}(\varphi(b))$ and y is in $\varphi^0(b)$. Choose $x^k = x + (y - x)/k$. Clearly, $x^k \to x$ and $g(x^k) \leq (<) \operatorname{Max}\{g(x), g(y)\} < (\leq) b$.

[4] g is ρ-continuous if g_i is upper semi-continuous for ρ_i equal to $\leq$, lower semi-continuous for ρ_i equal to $\geq$, and continuous for ρ_i equal to $=$.

Thus, $\{x^k\} \subseteq \varphi^0(b)$. This implies $x \in \mathrm{cl}\left(\varphi^0(b)\right)$, so the closure property holds. Note that this includes convex programs satisfying Slater's interiority condition.

Exercise: Let g be explicitly quasi-convex on S, but suppose Slater's condition fails (i.e., $\varphi^0(b) = \emptyset$), However, suppose $\{g_i \mid g_i(x) = b_i \text{ for all } x \in \varphi(b)\}$ contains only affine functions, so that $\varphi(b) = \{x \in X \mid g^1(x) \leq b^1,\ Ax + c = b^2\}$ and there exists $x^0 \in S$ to satisfy $g(x^1) < b^1$ and $Ax + c = b^2$. Prove that φ is lsc at $b = (b^1, b^2)$ relative to any feasible sequence (i.e., $\{b^k\} \subseteq B$).

Exercise: Suppose $S = \mathbb{R}^n$, $m = 1$ and g is strictly concave on $\mathbb{R}^n$ with ρ equal to $\leq$. Prove that the closure condition holds for any b in $\mathrm{int}\,(B)$.

Another class of nonlinear programs is where g is *strictly isotonic*; i.e., $x \leq y$ and $x \neq y$ imply $g(x) < g(y)$. If S is a convex, meet semi-lattice (i.e., x, y in S implies $[x, y] \subseteq S$ and $x \wedge y \in S$, where $x \wedge y$ is the *meet* of x and y, which is the n-vector whose i^{th} coordinate is $\mathrm{Min}\,\{x_i, y_i\}$). Then we shall show the closure condition holds.

Suppose $x \in \mathrm{cl}\,(\varphi(b))$ and $y \in \varphi^0(b)$. Define $w = x \wedge y$. Note that $w \leq y$ so $g(w) \leq g(y) < b$; hence, $w \in \varphi^0(b)$. Let $x^k = x + (w - x)/k$, so $x^k \to x$ and $\{x^k\} \subseteq S$. Further, $x^k \leq x$ for all k. If $x \in \varphi^0(b)$, then we are finished; otherwise $x^k \neq x$ for all k, so $g(x^k) < g(x) \leq b$; thus, x is in $\mathrm{cl}\left(\varphi^0(b)\right)$, so the closure condition is satisfied.

Exercise: Let S be a cone and let $g_i(tx) = t^{p_i} g_i(x)$ for all $t \geq 0$. Suppose $p_i > 0$ for all i or $p_i < 0$ for all i. Prove that the closure condition holds for any $b \in \mathrm{int}\,(B)$.

Exercise: Suppose g is quasi-convex and the closure condition holds for every $b \in \mathrm{int}\,(B)$. Prove that g is explicitly quasi-convex.

Exercise: Suppose g is isotonic (i.e., $x \leq y$ implies $g(x) \leq g(y)$) and the closure condition holds for every b in $\mathrm{int}\,(B)$. Prove that g is strictly isotonic.

The class of programs where P is restricted to be $\leq$ has special properties besides those described in our study of lower semi-continuity. Notice that φ is monotonic in that $b \geq a$ implies $\varphi(b) \supseteq \varphi(a)$; further, a in B and $b \geq a$ imply that b is in B. Monotonic maps are easier to study because of certain equivalent forms semi-continuity assumes. We can restrict our attention to monotone sequences converging to b because φ is usc (lsc) at b if, and only if, it is right (left) continuous; that is, φ is usc (lsc) iff $\bigcup_{k=1}^{\infty} \varphi(b^k) = \varphi(b)$ for all $\{b^k\} \downarrow b$ $\left(\bigcup_{k=1}^{\infty} \varphi(b^k) = \varphi(b)\right.$ for all $\left.\{b^k\} \uparrow b\right)$. In addition, f^* is isotonic on B, so f^* is upper (lower) semi-continuous at b iff f^* is continuous from the right (1eft); i.e., $f^*(b^k) \to f^*(b)$ for $b^k \downarrow b$ ($b^k \uparrow b$). The proofs of these assertions are left as exercises.

Now let us consider the optimality map. We do not have general conditions under which $\varphi*$ is lsc at b, and pathologies can occur even for convex programs. Of course, when $\varphi^*(\beta)$ is a singleton for each β in $N_\varepsilon(b)$, then $\varphi*$ is lsc at b iff φ^* is usc at b.

The usc of φ^* is related to convergence of certain algorithms. In particular, if $\{x^k\}$ and $\{b^k\}$ are sequences such that $x^k \in \varphi^*(b^k)$ and $b^k \to b$, then to guarantee a limit point of $\{x^k\}$ is in $\mathrm{cl}\,(\varphi^*(b))$ we need φ^* usc at b.

To prove that φ^* is usc under conditions stated in the theorem below we require the following

Lemma: Let X and Y be closed subsets of $\mathbb{R}^n$ such that X or Y is bounded and $X \cap Y \neq \emptyset$. Then, for all $\delta > 0$ there exists $\varepsilon > 0$ such that $\eta_\varepsilon(X) \cap \eta_\varepsilon(Y) \subset \eta_\delta(X \cap Y)$.

Exercise: Prove the above lemma.

USC OF OPTIMALITY MAP: Suppose f^* is continuous at b in B, $\varphi(b)$ is bounded and φ is usc at b. Then, φ^* is usc at b.

Proof: Let $b^k \to b$, where $\{b^k\} \subseteq B$. Let $\delta > 0$ be specified, where we must demonstrate there exists k^* (depending on δ) such that $\varphi^*(b^k) \subseteq \eta_\delta(\varphi^*(b))$ for $k \geq k^*$. Define $F^*(z) = \{x \in D \mid f(x) \geq z\}$, where D is a bounded set containing $\varphi(\beta)$ for all β in $N_\varepsilon(b)$. Clearly, F^* is usc at $f^*(b)$, so from our lemma above we have $\varphi^*(b^k) = \varphi(b^k) \cap F^*(f^*(b^k)) \subseteq \eta_\varepsilon(F^*(b)) \subseteq \eta_\delta(\varphi^*(b))$. (Note that the continuity of f^* is used in deducing that $f^*(b^k) \to f^*(b)$.) Thus, φ^* is usc at b. $\square$

Comment: To see that the complete continuity of f^* is needed consult Examp1e B.

We now consider forms of differentiability of f^* and its relation to Lagrangian theory (c.f., chapters 3 and 4). Our analysis begins with elements of Conjugate Function Theory and Subdifferentials. After presenting some basic theorems we revisit Lagrangian saddle point equivalence to gain a useful perspective on geometric foundations in response space.

Let $h : S \to \mathbb{R}_\infty$. Recall that the epigraph and hypograph of h are given respectively by[5]

$$\begin{aligned} \operatorname{epi}(h) &= \{(z,x) \mid x \in S \text{ and } z \geq h(x)\} \\ \operatorname{hypo}(h) &= \{(z,x) \mid x \in S \text{ and } z \leq h(x)\}. \end{aligned}$$

Let $y \in R^n$, and define the associated family of nonvertical hyperplanes $H(y;k) = \{(z,x) \in \mathbb{R}^{1+n} \mid (y,x) - z = k\}$. What is the value of k that causes the hyperplane to intersect the closure of the graph of h so as to support epi$\,(h)$? It is the supremum of $(y,x) - h(x)$ subject to x in S. In case there is no support with slope y, this supremum will be $+\infty$. This leads us to define the *convex conjugate*, where

$$h^\vee(y) = \operatorname{Sup}\{(y,x) - h(x) \mid x \in S\}.$$

By analogy, the *concave conjugate* of h is given by

$$h^\wedge(y) = \operatorname{Inf}\{(y,x) - h(x) \mid x \in S\}.$$

An important fact that motivated Fenchel's (1949) pioneering analysis is that we have the inequality $h(x) + h^\vee(y) \geq (x,y)$ for all x in S and y in $\mathbb{R}^n$. Different choices of h provide interesting special cases. For example, let $h(x) = \frac{1}{2}\,||x||^2$; then $h^\vee(y) = \frac{1}{2}\,||y||^2$, and we obtain $||x||^2 + ||y||^2 \geq 2(x,y)$ for all x, y in $\mathbb{R}^n$.

[5]The graph of h is their intersection: $\{(z,x) \mid x \in S \text{ and } z = h(x)\}$.

Further, we define the *effective domains* of $h^\vee$ and $h^\wedge$ respectively as

$$S^\vee \;=\; \{y \in \mathbb{R}^n \mid h^\vee(y) < +\infty\}$$

and

$$S^\wedge \;=\; \{y \in \mathbb{R}^n \mid h^\wedge(y) > -\infty\}.$$

Of particular importance to our analysis are the second conjugates. By definition we have

$$\begin{aligned} h^{\vee\vee}(x) &= \operatorname{Sup}\{(y,x) - h^\vee(y) \mid y \in \mathbb{R}^n\} \\ &= \operatorname{Sup}\{(y,x) - h^\vee(y) \mid y \in S^\vee\}. \end{aligned}$$

Replacing $h^\vee(y)$ by its definition we obtain

$$h^{\vee\vee}(x) \;=\; \operatorname*{Sup}_{y \in S^\vee} \operatorname*{Inf}_{w \,\in\, S} \{h(w) - (y, w-x) \mid y \in \mathbb{R}^n\}.$$

Exercise: Prove $h^{\wedge\wedge}(x) = \operatorname*{Inf}_{y \,\in\, S^\vee} \operatorname*{Sup}_{w \in S} \{h(w) - (y, w-x)\}$.

Exercise: Find first and second conjugates (both convex and concave) of the following:

1. $S = (-10, 10]$ and $h(x) = x^t$, where $t > 0$,
2. $S = \{0, 1, 2\}$ and $h(x) = x - 1$,
3. $S = \mathbb{R}$ and $h(x) = x^2$,
4. $S = \mathbb{R}$ and $h(x) = \frac{1}{p}x^p$, where $p > 1$,
5. $S = \mathbb{R}^n$ and $h(x) = (xQ, x)$, where Q is positive definite,
6. $S = \mathbb{R}^n$ and $h(x) = (xQ, x)$.

Here are some important properties of conjugates:

Property C1: $h^\vee$ is convex on $S^\vee$ and $h^\wedge$ is concave on $S^\wedge$.

Property C2: $h^{\vee\vee}(x) \le h(x) \le h^{\wedge\wedge}(x)$ for all x in S.

Property C3: $h^\vee$ is lower semi-continuous, and $h^\wedge$ is upper semi-continuous.

Property C4: If $S^\vee \neq \emptyset$, then $S^{\vee\vee} = \operatorname{cl}(\operatorname{convh}(S))$, and if $S^\wedge \neq \emptyset$, then $S^{\wedge\wedge} = \operatorname{cl}(\operatorname{convh}(S))$.

Property C5: $\operatorname{epi}(h^\vee)$ and $\operatorname{hypo}(h^\wedge)$ are closed sets.

Property C6: $\operatorname{epi}(h^{\vee\vee}) = \operatorname{cl}(\operatorname{convh}(\operatorname{epi}(h)))$, and $\operatorname{hypo}(h^{\wedge\wedge}) = \operatorname{cl}(\operatorname{convh}(\operatorname{hypo}(h)))$.

Property C7: $h(x) + h^\vee(y) \ge (x, y) \ge h(x) + h^\wedge(y)$.

Property C8: $h^{\vee\vee\vee} = h^\vee$ and $h^{\wedge\wedge\wedge} = h^\wedge$.

Property C9: $h^{\vee\vee}(x) = \operatorname*{Sup}_{y \in S^\vee} \operatorname*{Inf}_{w \,\in\, S} \{h^{\vee\vee}(w) - (y, w-x)\}$ and

$$h^{\wedge\wedge}(x) = \operatorname*{Inf}_{y \,\in\, S^{\wedge}} \ \operatorname*{Sup}_{w \,\in\, S} \{h^{\wedge\wedge}(w) - (y, w - x)\}.$$

We shall prove the first four properties for the convex conjugation; the remaining proofs are left as an exercise.

To begin we use the triangle inequality for the 'sup' operator, where we observe

$$\begin{aligned} h^{\vee}(\alpha u + (l-\alpha)v) &= \operatorname{Sup}\{\alpha((x,u) - h(x)) + (1-\alpha)((x,v) - h(x))\} \\ &\le \alpha \operatorname{Sup}\{(x,u) - h(x)\} + (1-\alpha)\operatorname{Sup}\{(x,v) - h(x)\} \\ &= \alpha h^{\vee}(u) + (1-\alpha)h^{\vee}(v). \end{aligned}$$

Thus, $h^{\vee}$ is convex on $S^{\vee}$.

Property C2 follows from the minimax inequality (c.f., chapter 4), where we note

$$\begin{aligned} h^{\vee\vee}(x) &= \operatorname*{Sup}_{y \in S^{\vee}} \ \operatorname*{Inf}_{w \,\in\, S} \{h(w) - (y, w - x)\} \\ &\le \operatorname*{Inf}_{w \,\in\, S} \ \operatorname*{Sup}_{y \in S^{\vee}} \{h(w) - (y, w - x)\} \\ &\le h(x) \ (= \text{if } S^{\vee} \ne \emptyset). \end{aligned}$$

Third, suppose $y^k \to y$. We shall prove $\liminf_{k \to \infty} h^{\vee}(y^k) \ge h^{\vee}(y)$. By definition $h^{\vee}(y^k) = \operatorname{Sup}\{(x, y^k) - h(x) \mid x \in S\}$, so $h^{\vee}(y^k) \ge (x, y^k) - h(x)$ for all x in S. Thus, $\liminf_{k \to \infty} h^{\vee}(y^k) \ge (x, y^k) - h(x)$ for all x in S, so the result follows.

Property C4 follows from Property C6, but we shall provide a separate proof. Notice that we need $S^{\vee} \ne \emptyset$; for example, if $S = [0, \infty)$ and $h(x) = -x^2$, then $h^{\vee}(y) = +\infty$ for all y in $\mathbb{R}$, so $S^{\vee} = \emptyset$. In that case $h^{\vee\vee}(x) = \operatorname{Sup}\{(x,y) - h^{\vee}(y) \mid y \in \mathbb{R}\} = -\infty$ for all x. Hence, $S^{\vee\vee} = \mathbb{R} \ne \operatorname{cl}(\operatorname{convh}(S))$. If we bound S to be the interval $[0, b)$, then $h^{\vee}(y) = \operatorname{Max}\{0,\ by - b^2\}$ for all y in $\mathbb{R} = S^{\vee}$. Thus, $h^{\vee\vee}(x) = \operatorname{Sup}\{(x,y) - h^{\vee}(y) \mid y \in \mathbb{R}\} = +\infty$ if, and only if, x is not in $[0, b]$, and $h^{\vee\vee}(x) = -bx$ if x is in $[0, b] = \operatorname{cl}(\operatorname{convh}(S)) = \operatorname{cl}(S)$. Further, if $S = (-1, 0) \cup (1, b)$, then $h^{\vee}(y) = \operatorname{Max}\{1 + |y|,\ by + b^2\}$ for y in $\mathbb{R} = S^{\vee}$, and $h^{\vee\vee}(x) = +\infty$ if, and only if, x is not in $[-1, b] = S^{\vee\vee} = \operatorname{cl}(\operatorname{convh}(S))$. Now let us prove the result.

Property C1 implies $S^{\vee\vee}$ is a convex set, and clearly $S \subseteq S^{\vee\vee}$. Further, Property C3 implies $S^{\vee\vee}$ is closed; thus, $\operatorname{cl}(\operatorname{convh}(S)) \subseteq S^{\vee\vee}$. Now suppose $x \notin \operatorname{cl}(\operatorname{convh}(S))$. Then, there exists $u \in \mathbb{R}^n$ and $\varepsilon > 0$ such that $(u, w - x) \le -\varepsilon$ for all w in S. Thus, we can choose any y in $S^{\vee}$ and deduce

$$\begin{aligned} h^{\vee\vee}(x) &\ge \operatorname*{Sup}_{t} \ \operatorname*{Inf}_{w \,\in\, S} \{h(w) - (y, w - x) - t(u, w - x)\} \\ &\ge (y, x) - h^{\vee}(y) + \operatorname*{Sup}_{t}\{t\varepsilon\}. \end{aligned}$$

Since y is in $S^{\vee}$, it follows that $h^{\vee}(y) < +\infty$, so $h^{\vee\vee}(x) = +\infty$. Therefore, $x \notin S^{\vee\vee}$. We have thus shown that $\operatorname{cl}(\operatorname{convh}(S)) \supseteq S^{\vee}$, which then completes our proof.

Notice that $h^{\vee\vee} = h$ if, and only if, h is convex on S and $\operatorname{epi}(h)$ is a closed set.

Suppose we have x, y such that Property C7 holds with equality; i.e., $x \in \text{argmax}\{(w, y) - h(w) \mid w \in S\}$. Then, y is called a *convex subgradient* of h at x. The set of convex subgradients is called the *subdifferential* of h at x and is denoted $\partial^\vee h(x)$. The subdifferential contains n-vectors that define slopes of non-vertical hyperplanes supporting epi (h); when h is convex, then the properties stated below imply that the support occurs at $(h(x), x)$ (if at all).

Similarly, the *concave subdifferential* is given by

$$\partial^\wedge h(x) = \{y \mid h(x) + h^\wedge(y) = (x, y)\}.$$

Example: Let $f(x) = |x|$ and $S = [-1, \infty)$. Then,

$$\partial^\vee f(x) = \begin{cases} [-1, \infty) & \text{if } x = -1 \\ \{-1\} & \text{if } -1 < x < 0 \\ [-1, 1] & \text{if } x = 0 \\ \{1\} & \text{if } x > 0. \end{cases}$$

Here are some important properties of subdifferentials:

Property D1: $\partial^\vee h(x)$ and $\partial^\wedge h(x)$ are closed, convex sets.

Property D2: $h \in \mathbb{C}^1(\mathbb{R}^n)$ implies $\partial^\vee h(x) \cup \partial^\wedge h(x) \subseteq \{\nabla h(x)\}$.[6]

Property D3: $y \in \partial^\vee h^{\vee\vee}(x) \leftrightarrow\ x \in \partial^\vee h^\vee(y)$, and $y \in \partial^\wedge h^{\wedge\wedge}(x) \leftrightarrow\ x \in \partial^\wedge h^\wedge(y)$.

Property D4: $x \in \text{int}\,(S^{\vee\vee})$ implies $\partial^\vee h^{\vee\vee}(x)$ is nonempty and bounded; $x \in \text{int}\,(S^{\wedge\wedge})$ implies $\partial^\wedge h^{\wedge\wedge}(x)$ is nonempty and bounded.

Property D5: If h is isotonic on S, then $\partial^\vee h^{\vee\vee}(x) \cup \partial^\wedge h^{\wedge\wedge}(x) \subseteq \mathbb{R}^n_+$.

Property D6: $\partial^\vee h^\vee(y) = \text{cl}\,(\text{convh}\,(\text{argmin}\{h(w) - (y, w) \mid w \in S\}))$.

The first five properties are straightforward, and we leave their proofs as an exercise. Let us consider property D6. For notational convenience let $X = \text{cl}\,(\text{convh}\,(\text{argmin}\{h(w) - (y, w) \mid w \in S\}))$. Clearly, $X \subseteq \partial^\vee h^\vee(y)$. Suppose $x \notin X$, so there exists $u \in \mathbb{R}^n$ and $\varepsilon > 0$ such that $(u, w - x) \leq -\varepsilon$ for all w in X. Thus, $h^{\vee\vee}(x) + h^\vee(y) \geq \operatorname*{Sup}_t \operatorname*{Inf}_{w \in S} \{h(w) - t(u, w - x) - (y, w - x)\} + h^\vee(y) = +\infty$. This implies $x \notin \partial^\vee h^\vee(y)$, and our proof is complete.

Exercise: Find $\partial^\vee h(x)$ and $\partial^\wedge h(x)$ at each x in S for the following:

1. $S = (-1, 1)$ and $h(x) = x^3$,
2. $S = [-1, 1]$ and $h(x) = -(1 - x^2)^{\frac{1}{2}}$,
3. $S = \mathbb{R}$ and $h(x) = e^x$,
4. $S = \mathbb{R}^2$ and $h(x) = x_1 + 2x_2$,

[6]Note: for $h = x^3$, $h^\vee(y) = \infty$ and $h^\wedge(y) = -\infty$ for all y, so $\partial^\vee h(x) = \partial^\wedge h(x) = \emptyset$. Thus, h can be differentiable but not subdifferentiable. If we restrict $S = R_+$, x^3 is convex and $\partial^\vee h(x) = 3x^2$ for $x > 0$.

5. $S = \mathbb{R}_+$ and $h(x) = \text{Max}\,\{x, 2x-1, 3x-3\}$.

Now let us return to nonlinear programming analysis. When f^* is concave on B, then $f^*(b) = f^{*\wedge\wedge}(b)$ at $b \in \text{int}\,(B)$; moreover, $\partial^\wedge f^*(b)$ is nonempty and compact. We shall now describe some of the relations these facts have to Lagrangian theory. In the process of so doing we shall obtain a more general theorem for Lagrangian saddle point equivalence and provide geometric insights into Lagrange multipliers.

Lemma l. For each λ in Λ the Lagrangian maximum equals the negative of the first concave conjugate of f^*; i.e.,

$$\sup_{x\in X}\{f(x) - (\lambda, g(x)\} = \sup_{b\in B}\{f^*(b) - (\lambda, b)\}.$$

Proof: We have denoted the left side of the above equation by $L^*(\lambda)$, and the right side is clearly $-f^{*\wedge}(\lambda)$. We see that $-f^{*\wedge}(\lambda) \geq \text{Sup}\{f^*(g(x)) - (\lambda, g(x)) \mid x \in S\}$ since λ is in Λ. Further, $f^*(g(x)) \geq f(x)$, so $-f^{*\wedge}(\lambda) \geq L^*(\lambda)$. On the other hand, we proved (c.f., chapter 4) that $f^*(b) \leq L^*(\lambda) + (\lambda, b)$ for all $b \in B$, so $-f^{*\wedge}(\lambda) \leq L^*(\lambda)$, which then implies $L^*(\lambda) = -f^{*\wedge}(\lambda)$. $\square$

We say *weak saddle point equivalence* holds at b (for the Lagrangian) if $f^*(b) = \inf_{\lambda\in\Lambda} \sup_{x\in S}\{L(x,\lambda) + (\lambda, b)\}$; if in addition there exists $\lambda^* \in \Lambda$ such that $f^*(b) = L^*(\lambda^*) + (\lambda^*, b)$, then we say *strong saddle point equivalence* holds at b. In chapter 4 we noted that if P is a closed, convex program satisfying the interiority condition, then strong saddle point equivalence holds at b. This means that $x^* \in \Omega(f, \varphi(b))$ if, and only if, there exists λ^* in Λ such that (x^*, λ^*) is a saddle point of the Lagrangian function, $L(x, \lambda) + (\lambda, b)$.

The first corollary to the following theorem generalizes that result because we only need f^* concave on B, and we are dealing directly with response space for our structural analysis.

LAGRANGIAN-RESPONSE THEOREM:

> Weak saddle point equivalence holds at b in B if, and only if, the optimal response function equals its second concave conjugate. Strong saddle point equivalence holds at b in B if, and only if, the optimal response function equals its second concave conjugate and is subdifferentiable there. In that case each Lagrange multiplier is a subgradient of f^* at b.

Corollary 1: If f^* is concave on B and $b \in \text{int}\,(B)$, then strong saddle point equivalence holds at b.

Coro11ary 2: If $\partial^\wedge f^*(b) = \{\lambda\}$, then $f^*(b) = f^{*\wedge\wedge}(b)$ and f^* is differentiable at b with $\nabla f^*(b) = \lambda$.

Corollary 3: If P is a convex program and the Lagrange multiplier is unique for $b \in \text{int}\,(B)$, (e.g., LCQ[7] holds and $\Omega(f, \varphi(b)) = \{x^*\}$), then f^* is differentiable at b.

[7]Lagrange Constraint Qualification

Corollary 4: If $x^* \in \underset{x \in S}{\operatorname{argmax}}\{L(x, y)\}$, then $x^* \in \Omega(f, \varphi(b))$ for any b such that $g(x^*)\, \rho\, b$ and $(\lambda, g(x^*) - b) = 0$.

Let us now show that Corollary 1 implies the Lagrangian Saddle Point Equivalence Theorem for Convex Programs (c.f., Chapter 4). We have already shown that f^* is concave on B when P is a convex program. Let us now show that the interiority condition implies $0 \in \operatorname{int}(B)$. It suffices to place 0 in a rectangle: i.e., that $R[a,b] \subseteq B$, where $a < 0 < b$. Partition the constraint function labels as

$$I^-\{i \mid \rho_i \text{ is } \le\},\ I^0\{i \mid \rho_i \text{ is } =\},\ \text{and } I^+\{i \mid \rho_i \text{ is } \ge\}.$$

Recall that the interiority condition ensures the existence of

1. $\{u^i\}_{i \in I^-} \subseteq S$ such that $g(u^i)\,\rho\,0$ and $g_i(u^i) < 0$;
2. $\{v^i\}_{i \in I^+} \subseteq S$ such that $g(v^i)\,\rho\,0$ and $g_i(v^i) > 0$;
3. $\{(w^i, x^i)\}_{i \in I^0} \subseteq S^2$ such that $g_j(w^i)\ \rho_j\ 0$ and $g_j(x^i)\ \rho_j\ 0$ for $j \neq i$, further that $g_i(w^i) < 0 < g_i(x^i)$.

Exercise: Prove that the interiority condition implies that the linear equations (from I^0) are linearly independent.

Define $M = \left|I^0\right|$ (i.e., the number of equality constraints) and

$$a_i = \begin{cases} \frac{1}{m+M}\, g_i(u^i) & \text{if } i \in I^- \\ \frac{1}{m+M}\, g_i(w^i) & \text{if } i \in I^0 \\ -1 & \text{if } i \in I^+ \end{cases}$$

$$b_i = \begin{cases} 1 & \text{if } i \in I^- \\ \frac{1}{m+M}\, g_i(v^i) & \text{if } i \in I^0 \\ \frac{1}{m+M}\, g_i(x^i) & \text{if } i \in I^+. \end{cases}$$

Note that $a < 0 < b$. The rectangle $R(a,b)$ is contained in B when g is ρ-convex[8], as we shall now show. Suppose $\beta \in R(a,b)$, and define

$$y = \sum_{i \in I^-} \alpha_i u^i + \sum_{i \in I^+} \alpha_i v^i + \sum_{i \in I^0} (\theta_i w^i + \delta_i x^i) + \gamma x,$$

where $x \in F(g)$, $(\alpha, \theta, \delta, \gamma) \ge 0$, and

$$\gamma = 1 - \sum_{i \in I^- \cup I^+} \alpha_i - \sum_{i \in I^0} (\theta_i + \delta_i).$$

We shall construct values of $(\alpha, \theta, \delta, \gamma)$ such that $g(y)\,\rho\,\beta$.

---

[8] g_i is convex for ρ_i equal to $\le$, concave for ρ_i equal to $\ge$, and affine for ρ_i equal to $=$.

The ρ-convexity of g implies

$$g_k(y)\,\rho_k \sum_{i\in I^-} \alpha_i g_k(u^i) + \sum_{i\in I^+} \alpha_i g_k(v^i) + \sum_{i\in I^0} \theta_i g_k(w^i) + \sum_{i\in I^0} \delta_i g_k(x^i) + \gamma g_k(x).$$

For $k \in I^-$ we have $g_k(y) \le \alpha_k g_k(u^k)$. Choose

$$\alpha_k = \frac{\beta_k^-}{g_k(u^k)}$$

and note $g_k(y) \le \beta_k$, $0 \le \alpha_k \le \frac{1}{m+M}$.

For $k \in I^+$ we have $g_k(y) \ge \alpha_k g_k(u^k)$. Choose

$$\alpha_k = \frac{\beta_k^+}{g_k(v^k)}$$

and note $g_k(y) \ge \beta_k$, $0 \le \alpha_k \le \frac{1}{m+M}$.

Notation: $z^+ = \text{Max}\,\{0, z\}$ and $z^- = \text{Min}\,\{0, z\}$.

For $k \in I^0$ we have $g_k(y) = \theta_k g_k(wk) + \delta_k g_k(x^k)$. Choose

$$\theta_k = \frac{\beta_k^-}{g_k(w^k)}, \quad \delta_k = \frac{\theta_k^+}{g_k(x^k)}$$

and note that $g_k(y) = \beta_k$, $\theta_k\delta_k = 0$, $0 \le \theta_k, \delta_k \le \frac{1}{m+M}$.

We have already defined γ for the feasible point, x. Thus, we have constructed $y \in S$ such that $g(y)\,\rho\,\beta$, where $\beta \in R(a,b)$, $a < 0 < b$. This proves $0 \in \text{int}\,(B)$.

The conditions of Corollary 1 are met and f^* is subdifferentiable at 0. This means there exists $\lambda^* \in \Lambda$ such that $f^*(0) = L^*(\lambda^*)$. If $g(x^*)\,\rho\,0$ and $f(x^*) = f^*(0)$, then $x^* \in \text{argmax}\{L^*(x,\lambda^*) \mid x \in S\}$ and $(\lambda^*, g(x^*)) = 0$. Thus, (x^*, λ^*) is indeed a saddle point of the Lagrangian.

When the convexity assumptions are dropped, then Lagrangian saddle point equivalence need not hold. (Keep in mind that it is the concavity of f^* on B that is pertinent, and there are important classes of problems for which this is true even though the program is not convex, e.g., geometric programs.) However, as we indicated earlier, if x^* maximizes the Lagrangian, $L(x,\lambda)$ on S, then x^* is optimal for any program with the same functions, but with g replaced by $g - b$ for b such that $g(x^*)\,\rho\,b$ and $(\lambda, g(x^*) - b) = 0$.[9] We thus define the *solution set* for a multiplier $\lambda \in \Lambda$:

$$\Psi(\lambda) = \left\{ b \in B \mid \exists x^* \in \underset{x\in S}{\text{argmax}}\{L(x,\lambda)\} \cap \varphi(b) \text{ such that } (\lambda, g(x^*) - b) = 0 \right\}.$$

[9] This is complementary slackness.

Note that $g(x^*) \in \Psi(\lambda)$ for any $x^* \in \operatorname{argmax} L(x,\lambda)$. The existence of $\lambda^* \in \Lambda$ such that $0 \in \Psi(\lambda^*)$ is precisely the question of strong saddle point equivalence; it is ensured for programs such that f^* is concave on B and $0 \in \operatorname{int}(B)$.

Exercise: Prove that $\operatorname{cl}(\operatorname{convh}(\Psi(\lambda))) = \partial^\vee L^{*\vee}(\lambda)$.

Exercise: Suppose $\lambda \in \Lambda$. Prove that $b^* \in \operatorname{argmax}\{f^*(b) - \lambda b \mid b \in B\}$ if, and only if, there exists $x^* \in \underset{x \in S}{\operatorname{argmax}}\{L(x,\lambda)\} \cap \varphi(b^*)$ such that $f(x^*) = f^*(b^*)$.

Exercise: Use the Lagrangian approach alluded to above to solve Max $(c,x) : ||x||^2 \le b$.

Let us conclude with a comparative analysis of two conjugate approximates that emphasize the response space structure underlying Lagrangian saddle point theory.

We have seen that the Lagrangian maximum provides a bound on $f^*(b)$ by the conjugate inequality

$$f^*(b) \le f^{*\vee\vee}(b) = \operatorname*{Inf}_{\lambda \in \Lambda}\{L^*(\lambda) + (\lambda, b)\},$$

where recall $L^*(\lambda) = -f^{*\wedge}(\lambda)$. In this sense the Lagrangian approach rests upon a conjugate approximation of our response function, f^* on B.

Suppose we approximate P by taking second conjugates of f and g. Specifically, let ρ_i be $\le$ for each $i = 1, \ldots, m$ and define the *conjugate approximate* of P as the convex program represented by

$$P^\vee : \text{ Max } f^{\wedge\wedge}(x) : g^{\vee\vee}(x) \le b, \ x \in \operatorname{cl}(\operatorname{convh}(S)).$$

Let

$$h(b) \equiv \operatorname{Sup}\{f^{\wedge\wedge}(x) \mid g^{\vee\vee}(x) \le b, \ x \in \operatorname{cl}(\operatorname{convh}(S))$$

and

$$B^* \equiv \operatorname{Range}\left(g^{\vee\vee}\right) + \mathbb{R}^m_+ = \{b \mid g^{\vee\vee}(x) \le b \text{ for some } x \in \operatorname{cl}(\operatorname{convh}(S))\}.$$

Then, $B \subseteq B^*$ and $f^*(b) \le h(b)$ for all $b \in B$.

Let us compare the bound $h(b)$ on $f^*(b)$ with our Lagrangian bound. Suppose the interiority condition holds for b, so there exists $\{u^i\}_1^m \subseteq F$ such that $g_i(u^i) < 0$ for $i = 1, \ldots, m$. Since $g^{\vee\vee}(u^i) \le g(u^i)$, it follows that the Slater interiority condition holds, so $b \in \operatorname{int}(B^*)$. This implies $h(b) = h^{\wedge\wedge}(b)$ (and in fact h is subdifferentiable at b). Then, we have

$$\begin{aligned} h(b) = h^{\wedge\wedge}(b) &= \operatorname*{Inf}_{\lambda \ge 0} \operatorname*{Sup}_{\beta \in B^*} \{h(\beta) - (\lambda, \beta - b)\} \\ &\ge \operatorname*{Sup}_{\beta \in B} \operatorname*{Inf}_{\lambda \ge 0} \{f^(\beta - (\lambda, \beta - b)\} = f^{*\wedge\wedge}(b). \end{aligned}$$

We have now proven the following result:[10]

[10]See Greenberg (1973) for the additional argument that it never requires more computation to solve the Lagrangian dual than to solve $P^\vee$ — i.e., response space conjugation provides a stronger (or equal) bound with less (or equal) computation.

CONJUGATE COMPARISON THEOREM:

> Suppose the interiority condition holds, so $b \in \operatorname{int}(B^*)$. Then, the Lagrangian bound is never weaker than the bound obtained from the conjugate approximation in decision space; i.e.,
>
> $$\begin{aligned} f^*(b) &\leq \operatorname{Inf}\{L^*(\lambda) + (\lambda, b)\} \\ &\leq \operatorname{Sup}\{f^{\wedge\wedge}(x) \mid x \in \operatorname{cl}(\operatorname{convh}(S)),\ g^{\vee\vee}(x) \leq b\}. \end{aligned}$$

References

C. Berge. *Topological Spaces*. MacMillan Company, New York, NY, 1959.

G. B. Dantzig, J. Folkman, and N. Shaprio. On the continuity of the minimum set of a continuous function. *Journal of Mathematical Analysis and Applications*, 17(3):519–548, 1967.

J. P. Evans and F. J. Gould. Stability in nonlinear programming. *Operations Research*, 18(1):107–118, 1970.

J. E. Falk. Lagrange multipliers and nonlinear programming. *Journal of Mathematical Analysis and Applications*, 19(1):141–159, 1967.

W. Fenchel. On conjugate convex functions. *Canadadian Journal of Mathematics*, 1:73–77, 1949.

H. J. Greenberg. Generalized Lagrange multiplier theory. Lecture note, Southern Methodist University, Dallas, TX, 1968.

H. J. Greenberg. Lagrangian duality gaps: Their source and resolution. Technical Report CP-69005, Southern Methodist University, Dallas, TX, 1969.

H. J. Greenberg. Bounding nonconvex programs by conjugates. *Operations Research*, 21(1):346–348, 1973.

H. J. Greenberg. The generalized penalty-function/surrogate model. *Operations Research*, 21(1):162–178, 1973.

H. J. Greenberg and W. P. Pierskalla. Surrogate mathematical programming. *Operations Research*, 18(5):924–939, 1970.

H. J. Greenberg and W. P. Pierskalla. Extensions of the Evans-Gould stability theorems for mathematical programs. *Operations Research*, 20(1):143–153, 1972.

H. J. Greenberg and W. P. Pierskalla. A review of quasi-convex functions. *Operations Research*, 19(1):1553–1570, 1972.

H. J. Greenberg and T. Robbins. Finding Everett's Lagrange multipliers by generalized linear programming. Technical Report CP-70008, Southern Methodist University, Dallas, TX, 1970.

R. T. Rockafellar. Level sets and continuity of conjugate convex functions. *Transactions of the American Mathematical Society*, 123(1):46–63, 1966.

R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

Acknowledgement and Commentary

I thank the referee who noted several ways to improve the presentation, including its organization. I would go even further with adding clarifications, but I want to maintain this as a *transcription* of 1969 notes with only a few edits, which I footnoted, to correct something. Moreover, there are many additional references I would add, though most of the classical results, notably from conjugate duality, assume convex programs. My starting point was with Everett's paper, which is not referenced by even recent texts when applying Lagrangian duality to non-convex programs.