January 2002 (revised April 2003) **LIDS-P-2536**

Min Common/Max Crossing Duality: A Simple Geometric Framework for Convex Optimization and Minimax Theory[1]

by

Dimitri P. Bertsekas, Angelia Nedić, and Asuman E. Ozdaglar[2]

Abstract

We provide a simple unifying framework for the visualization and analysis of convex programming duality and minimax (saddle point) theory. In particular, we introduce two geometrical problems that are dual to each other: the *min common point* problem and the *max crossing point* problem. Within the simple geometry of these problems, the fundamental constraint qualifications needed for strong duality are quite apparent, and admit straightforward proofs. We develop the relevant theorems, and we then obtain as special cases the major results of Lagrangian duality theory for constrained optimization and of convex/concave minimax theory.

[1] Research supported by NSF under Grant ACI-9873339.

[2] Dept. of Electrical Engineering and Computer Science, Lab. for Information and Decision Systems, M.I.T., Cambridge, Mass., 02139.

1. INTRODUCTION

Duality in optimization is often considered to be a manifestation of a fundamental dual description of a closed convex set:

(a) As the closure of the union of all line segments connecting the points of the set.

(b) As the intersection of all closed halfspaces containing the set.

This is largely true but it is also somewhat misleading, because the strongest duality theorems in optimization require assumptions such as the Slater condition and other constraint qualifications, whose connection to the dual description of closed convex sets given above is not readily apparent. As a result, one often observes a dichotomy in various developments of optimization duality theory found in the literature: some suggestive geometrical insight may be given, but the main proof lines are not clearly connected with the fundamental underlying geometry (except perhaps in the eyes of a skilled mathematician). For example, the proof of the main duality theorem of linear programming is often developed based on Farkas' lemma whose relation with the preceding dual closed convex set description is not readily apparent, and in other cases it is developed based on the termination properties of the simplex method, with hardly any geometrical insight resulting.

In this paper, we aim to capture the most essential optimization-related aspect of the preceding dual characterization of closed convex sets in two easily visualized geometrical optimization problems that are defined in terms of a nonempty subset M of $\Re^{n+1}$.

(a) *Min Common Point Problem*: Among all points that are common to both M and the $(n+1)$st axis, we want to find one whose $(n+1)$st component is minimum.

(b) *Max Crossing Point Problem*: Consider nonvertical hyperplanes that contain M in their corresponding "upper" closed halfspace [the halfspace that contains the vertical halfline $\big\{(0,w) \mid w \geq 0\big\}$ in its recession cone, see Fig. 1.1]. We want to find the maximum crossing point of the $(n+1)$st axis with such a hyperplane.

Figure 1.1 suggests that the optimal value of the max crossing problem is no larger than the optimal value of the min common problem; we refer to this relation as *weak duality*. Furthermore, under favorable circumstances the two optimal values are equal; we refer to this relation as *strong duality*. Our objective is to establish conditions for strong duality, and to characterize circumstances under which the max crossing problem has an optimal solution.

There are many works that treat duality, including the textbooks by Rockafellar [Roc70],

Figure 1.1. Illustration of the optimal values w^* and q^* of the min common and max crossing problems. In (a), the two optimal values are not equal. In (b), when M is "extended upwards" along the $(n+1)$st axis it yields the set

$$\overline{M} = \Big\{(u,w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \le w \text{ and } (u,\overline{w}) \in M\Big\},$$

which is convex and admits a nonvertical supporting hyperplane passing through $(0,w^*)$. As a result, the two optimal values are equal. In (c), the set $\overline{M}$ is convex but not closed, and there are points $(0,\overline{w})$ on the vertical axis with $\overline{w} < w^*$ that lie in the closure of $\overline{M}$. Here, q^* is equal to the minimum such value of $\overline{w}$, and we have $q^* < w^*$.

Stoer and Witzgall [StW70], Ekeland and Temam [EkT76], Rockafellar [Roc84], Hiriart-Urruty and Lemarechal [HiL93], Rockafellar and Wets [RoW98], Bertsekas [Ber99], Bonnans and Shapiro [BoS00], and Borwein and Lewis [BoL00]. The constructions involved in the min common and max crossing problems are implicit in these duality analyses, and in fact have been used earlier for visualization purposes (see Bertsekas [Ber99], Ch. 5). However, the two problems have never been explicitly analyzed, to our knowledge, nor have they been used as a unifying theoretical framework for analysis of important special cases arising in constrained optimization duality, saddle-point theory, or other contexts. There is an important benefit from the analysis of these problems: within their simple geometry, the fundamental constraint qualifications needed for strong duality are quite apparent, and admit straightforward proofs. This allows us to develop essentially all of duality theory within the simple min common/max crossing framework, and then to apply it to optimization (Lagrangian) duality and obtain (as special cases) all of the major strong duality theorems. Note that while we focus on the Lagrangian duality framework (see Section 4), our analysis also applies to the Fenchel duality framework, which is equivalent to the Lagrangian framework (see e.g., Bertsekas [Ber99], Section 5.4). Thus, the standard Fenchel duality results can also be recovered as special cases of the min common/max crossing results.

There is also another important unification benefit from our min common/max crossing framework. The duality theorems that we prove within this framework can be used not only to develop optimization duality, but also to develop saddle point and minimax theory (under convexity/concavity assumptions), including the fundamental von Neuman Theorem of zero sum game theory [Neu28]. Our min common/max crossing line of development is related to the approach of Rockafellar [Roc70] (Section 33), which is based on convex bifunctions and conjugate saddle functions. Pedagogically, however, it appears desirable to develop minimax theory without resort to this complicated machinery, as we have done.

It is well known that saddle point theory and optimization duality are strongly connected; for example, they have both been treated within the unifying framework of convex bifunctions ([Roc70], Section 33). In principle, optimization duality can be viewed as a special case of saddle point/minimax theory. On the other hand, it is not always convenient or possible to use minimax results to prove optimization duality results. In our approach, rather than trying to build a closer connection between duality and saddle point theory, we show, with quite elementary proofs, how they both stem from a common geometrical root: the min common/max crossing duality, which is in turn transparently connected to the dual characterization of a closed convex set as the closure of the union of line segments and as an intersection of closed halfspaces.

The paper is organized as follows. In Section 2, we provide some terminology and background on supporting and separating hyperplane theory, and we develop some preliminary results on the existence of nonvertical supporting hyperplanes of convex sets. In Section 3, we analyze the min common and max crossing problems, and we develop several conditions that guarantee the equality of their optimal values and/or the existence of their optimal solutions. In Section 4, we apply the results of Section 3 to minimax and saddle point theory, we recover a number of known results, including a version of the classical von Neuman Theorem. The proofs of several of our results, however, are considerably simpler than the ones found in the literature, and also admit insightful visualization. In Section 5, we apply the results of Sections 3 and 4 to the Lagrangian duality framework for constrained optimization. We recover, as special cases of the min common/max crossing results, the major strong duality theorems.

Regarding notation, all of the vectors are column vectors and a prime denotes transposition. We write $x \geq 0$ or $x > 0$ when a vector x has nonnegative or positive components, respectively. Similarly, we write $x \leq 0$ or $x < 0$ when a vector x has nonpositive or negative components, respectively. We use throughout the paper the standard Euclidean norm in $\Re^n$, $\|x\| = (x'x)^{1/2}$, where $x'y$ denotes the inner product of any $x, y \in \Re^n$. We denote by $\mathrm{cl}(C)$ and $\mathrm{int}(C)$ the closure and the interior of a set C, respectively. We also use some of the standard notions of convex analysis. In particular, for a convex set C, we denote by $\mathrm{aff}(C)$ the affine hull of C, i.e., the smallest affine set containing C, and by $\mathrm{ri}(C)$ the relative interior of C, i.e., its interior relative to $\mathrm{aff}(C)$. The epigraph $\big\{(x, w) \mid f(x) \leq w,\, x \in X,\, w \in \Re\big\}$ of an extended real-valued function $f : X \mapsto [-\infty, \infty]$ is denoted by $\mathrm{epi}(f)$. Following Rockafellar [Roc70], the function f is said to be convex or closed if $\mathrm{epi}(f)$ is convex or closed, respectively. It is said to be proper, if its epigraph is nonempty and does not contain a vertical line, i.e., if $f(x) > -\infty$ for all $x \in X$ and

$f(x) < \infty$ for al least one $x \in X$.

2. NONVERTICAL HYPERPLANES

We recall that a *hyperplane* in $\Re^n$ is a set of the form $\{x \mid a'x = b\}$, where $a \in \Re^n$, $a \neq 0$, and $b \in \Re$. The sets

$$\{x \mid a'x \geq b\}, \qquad \{x \mid a'x \leq b\},$$

are called the *closed halfspaces* associated with the hyperplane (also referred to as the *positive and negative closed halfspaces*, respectively). For the purpose of easy reference, we list some of the separating and supporting hyperplane results that we will use in our analysis. The first four propositions can be found in many textbooks, including Rockafellar [Roc70], Section 11, and our recent book [BNO03].

Proposition 2.1: (Supporting Hyperplane Theorem) Let C be a nonempty convex subset of $\Re^n$ and let $\overline{x}$ be a vector in $\Re^n$. If either C has empty interior or, more generally, if $\overline{x}$ is not an interior point of C, there exists a hyperplane that passes through $\overline{x}$ and contains C in one of its closed halfspaces, i.e., there exists a vector $a \neq 0$ such that

$$a'\overline{x} \leq a'x, \qquad \forall\ x \in C. \tag{2.1}$$

If the vector $\overline{x}$ in the above proposition belongs to the closure of the set C, the corresponding hyperplane is said to be *supporting* C at $\overline{x}$.

Proposition 2.2: (Strict Separation Theorem) Let C_1 and C_2 be nonempty disjoint convex subsets of $\Re^n$ such that C_1 is closed and C_2 is compact. Then, there exists a hyperplane that strictly separates them, i.e., a vector $a \in \Re^n$ and a scalar b such that

$$a'x_1 < b < a'x_2, \qquad \forall\ x_1 \in C_1, \quad \forall\ x_2 \in C_2. \tag{2.2}$$

Proposition 2.3: (Proper Separation Theorem) Let C_1 and C_2 be nonempty convex subsets of $\Re^n$ such that

$$\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \emptyset.$$

Then there exists a hyperplane that properly separates C_1 from C_2, i.e., a vector a such that

$$\sup_{x \in C_2} a'x \le \inf_{x \in C_1} a'x, \qquad \inf_{x \in C_2} a'x < \sup_{x \in C_1} a'x.$$

Proposition 2.4: A closed and convex set is the intersection of the closed halfspaces that contain it.

The next proposition is important in the context of duality when polyhedral sets are involved. It is given as Theorem 20.2 of Rockafellar [Roc70].

Proposition 2.5: (Polyhedral Proper Separation Theorem) Let C_1 and C_2 be nonempty convex subsets of $\Re^n$ such that C_2 is polyhedral and

$$\mathrm{ri}(C_1) \cap C_2 = \emptyset.$$

Then there exists a hyperplane that properly separates them and does not contain C_1, i.e., a vector a such that

$$\sup_{x \in C_2} a'x \le \inf_{x \in C_1} a'x, \qquad \inf_{x \in C_1} a'x < \sup_{x \in C_1} a'x.$$

We now discuss hyperplanes in a special context that involves among others, the epigraph $\mathrm{epi}(f)$ of a convex function $f : \Re^n \mapsto [-\infty, \infty]$, i.e., the subset of $\Re^{n+1}$ defined by

$$\mathrm{epi}(f) = \big\{(u, w) \mid f(u) \le w\big\}.$$

We will represent the normal vector of a hyperplane in $\Re^{n+1}$ as a nonzero vector of the form (μ, β), where $\mu \in \Re^n$ and $\beta \in \Re$. We say that the hyperplane is *horizontal* if $\mu = 0$ and we say that it is *vertical* if $\beta = 0$.

Note that if a hyperplane with normal (μ, β) is nonvertical (i.e., $\beta \neq 0$), then it crosses the $(n+1)$st axis (the axis associated with w) at a unique point. If $(\overline{u}, \overline{w})$ is any vector on the hyperplane, the crossing point has the form $(0, \xi)$, where

$$\xi = \frac{\mu'}{\beta}\overline{u} + \overline{w}, \tag{2.3}$$

since from the hyperplane equation, we have $(0, \xi)'(\mu, \beta) = (\overline{u}, \overline{w})'(\mu, \beta)$. On the other hand, it can be seen that if the hyperplane is vertical, it either contains the entire $(n+1)$st axis, or else it does not cross it at all; see Fig. 2.1.

Figure 2.1. Illustration of vertical and nonvertical hyperplanes in $\Re^{n+1}$. A hyperplane with normal (μ, β) is nonvertical if $\beta \neq 0$, or equivalently, if it intersects the $(n+1)$st axis at the unique point $\xi = (\mu/\beta)'\overline{u} + \overline{w}$, where $(\overline{u}, \overline{w})$ is any vector on the hyperplane.

Vertical lines in $\Re^{n+1}$ are sets of the form $\{(\overline{u}, w) \mid w \in \Re\}$, where $\overline{u}$ is a fixed vector in $\Re^n$. It can be seen that vertical hyperplanes, as well as their corresponding closed halfspaces, consist of the union of the vertical lines that pass through their points. If $f(u) > -\infty$ for all u, then $\mathrm{epi}(f)$ cannot contain a vertical line, and it appears plausible that $\mathrm{epi}(f)$ is contained in some closed halfspace corresponding to a nonvertical hyperplane. We prove this fact in greater generality in the following proposition, which will also be useful as a first step in the subsequent development.

Proposition 2.6: Let C be a nonempty convex subset of $\Re^{n+1}$ that contains no vertical lines. Let the vectors in $\Re^{n+1}$ be denoted by (u, w), where $u \in \Re^n$ and $w \in \Re$. Then:

(a) C is contained in a closed halfspace corresponding to a nonvertical hyperplane, i.e., there exist a vector $\mu \in \Re^n$, a scalar β with $\beta \neq 0$, and a scalar γ such that

$$\mu' u + \beta w \geq \gamma, \qquad \forall\ (u, w) \in C.$$

(b) If $(\overline{u}, \overline{w})$ does not belong to $\mathrm{cl}(C)$, there exists a nonvertical hyperplane strictly separating $(\overline{u}, \overline{w})$ from C.

Proof: We first note that if C contains no vertical lines, then $\mathrm{ri}(C)$ contains no vertical lines, which implies that $\mathrm{cl}(C)$ contains no vertical lines, since the recession cones of $\mathrm{cl}(C)$ and $\mathrm{ri}(C)$

coincide (cf. Rockafellar [Roc70], Cor. 8.3.1). Thus, if we prove the result assuming that C is closed, the proof for the case where C is not closed will readily follow by replacing C with $\mathrm{cl}(C)$. Hence, we may assume without loss of generality that C is closed.

(a) By Prop. 2.4, C is the intersection of all closed halfspaces that contain it. If every hyperplane containing C in one of its closed halfspaces is vertical, we must have

$$C = \cap_{i\in I}\big\{(u,w) \mid \mu_i' u \geq \gamma_i\big\}$$

for a collection of nonzero vectors μ_i, $i \in I$, and scalars γ_i, $i \in I$. Then, for every $(\overline{u},\overline{w}) \in C$, the vertical line $\big\{(\overline{u},w) \mid w \in \Re\big\}$ also belongs to C, a contradiction. It follows that if no vertical line belongs to C, there exists a nonvertical hyperplane containing C.

(b) If $(\overline{u},\overline{w}) \notin C$, then since C is assumed to be closed, there exists a hyperplane strictly separating $(\overline{u},\overline{w})$ from C (cf. Prop. 2.2). If this hyperplane is nonvertical, we are done, so assume otherwise. Then, we have a nonzero vector $\overline{\mu}$ and a scalar $\overline{\gamma}$ such that

$$\overline{\mu}' u > \overline{\gamma} > \overline{\mu}'\,\overline{u}, \qquad \forall\ (u,w) \in C.$$

Consider a nonvertical hyperplane containing C in one of its subspaces [which exists by part (a)], so that for some (μ,β) and γ, with $\beta \neq 0$, we have

$$\mu' u + \beta w \geq \gamma, \qquad \forall\ (u,w) \in C.$$

By multiplying this relation with any $\epsilon > 0$ and adding it to the preceding relation, we obtain

$$(\overline{\mu} + \epsilon\mu)' u + \epsilon\beta w > \overline{\gamma} + \epsilon\gamma, \qquad \forall\ (u,w) \in C.$$

Since $\overline{\gamma} > \overline{\mu}'\,\overline{u}$, there is a small enough ϵ such that

$$\overline{\gamma} + \epsilon\gamma > (\overline{\mu} + \epsilon\mu)'\overline{u} + \epsilon\beta\overline{w}.$$

From the above two relations, we obtain

$$(\overline{\mu} + \epsilon\mu)' u + \epsilon\beta w > (\overline{\mu} + \epsilon\mu)'\overline{u} + \epsilon\beta\overline{w}, \qquad \forall\ (u,w) \in C,$$

implying that there is a nonvertical hyperplane with normal $(\overline{\mu} + \epsilon\mu, \epsilon\beta)$ that strictly separates $(\overline{u},\overline{w})$ from C. **Q.E.D.**

3. MIN COMMON/MAX CROSSING DUALITY

We now consider the min common and max crossing problems introduced in Section 1. Let M be a nonempty subset of $\Re^{n+1}$. The min common problem is

$$\begin{aligned} &\text{minimize} \quad w \\ &\text{subject to} \quad (0,w) \in M, \end{aligned} \tag{3.1}$$

and its optimal value is denoted by w^*, i.e.,

$$w^* = \inf_{(0,w)\in M} w.$$

Given a nonvertical hyperplane in $\Re^{n+1}$, multiplication of its normal vector (μ,β) by a nonzero scalar produces a vector that is also normal to the same hyperplane. Hence, the set of nonvertical hyperplanes, where $\beta \neq 0$, can be equivalently described as the set of all hyperplanes with normals of the form $(\mu,1)$. A hyperplane of this type crosses the $(n+1)$st axis at some vector $(0,\xi)$ and is of the form

$$H_{\mu,\xi} = \big\{(u,w) \mid w + \mu' u = \xi\big\}.$$

In order for M to be contained in the closed halfspace that corresponds to the hyperplane $H_{\mu,\xi}$ and contains the vertical halfline $\big\{(0,w) \mid w \geq 0\big\}$ in its recession cone, we must have

$$\xi \leq w + \mu' u, \qquad \forall\ (u,w) \in M.$$

The maximum crossing level ξ over all hyperplanes $H_{\mu,\xi}$ with the same normal $(\mu,1)$ is given by

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\}; \tag{3.2}$$

(see Fig. 3.1). The problem of maximizing the crossing level over all nonvertical hyperplanes is to maximize over all $\mu \in \Re^n$ the maximum crossing level corresponding to μ, i.e.,

$$\begin{aligned} &\text{maximize} \quad q(\mu) \\ &\text{subject to} \quad \mu \in \Re^n. \end{aligned} \tag{3.3}$$

Note that q is concave and upper semicontinuous over $\Re^n$, since it is defined as the infimum of a collection of affine functions. We denote by q^* the corresponding optimal value,

$$q^* = \sup_{\mu\in\Re^n} q(\mu).$$

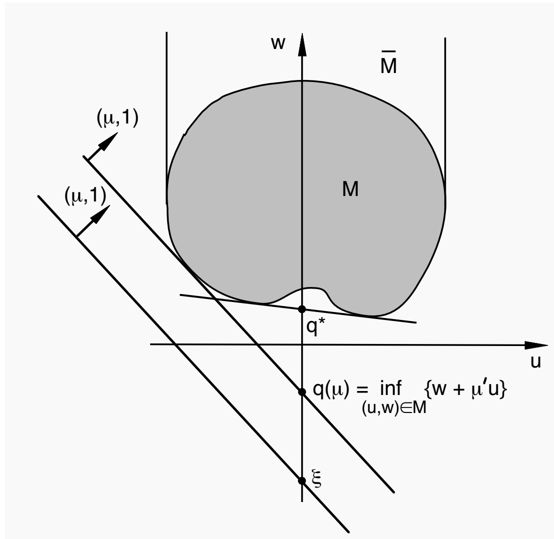

Figure 3.1. Mathematical specification of the max crossing problem. For each $\mu \in \Re^n$, we consider $q(\mu)$, the highest crossing level over hyperplanes, which have normal $(\mu, 1)$ and are such that M is contained in their positive halfspace [the one that contains the vertical halfline $\big\{(0, w) \mid w \geq 0\big\}$ in its recession cone]. The max crossing point q^* is the supremum over $\mu \in \Re^n$ of the crossing levels $q(\mu)$.

Note that for every $(u, w) \in M$ and every $\mu \in \Re^n$, we have

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu'u\} \leq \inf_{(0,w)\in M} w = w^*,$$

so by taking the supremum of the left-hand side over $\mu \in \Re^n$, we obtain

$$q^* \leq w^*, \tag{3.4}$$

i.e., the max crossing point is no higher than the min common point, as suggested by Fig. 1.1. We will refer to relation (3.4) as *weak duality.*

We now turn to establishing conditions under which we have $q^* = w^*$, in which case we say that *strong duality holds* or that *there is no duality gap*. To avoid degenerate cases, we will generally exclude the case $w^* = \infty$, when the min common problem is infeasible, i.e., $\big\{w \mid (0, w) \in M\big\} = \emptyset$.

An important point, around which much of our analysis revolves, is that when w^* is a scalar, the vector $(0, w^*)$ is a closure point of the set M, so *if we assume that M is convex and closed, and admits a nonvertical supporting hyperplane at* $(0, w^*)$*, then we have $q^* = w^*$ and the optimal values q^* and w^* are attained*. Between the "unfavorable" case where $q^* < w^*$, and the "most favorable" case where $q^* = w^*$ while the optimal values q^* and w^* are attained, there are several intermediate cases. The following proposition provides a necessary and sufficient condition for $q^* = w^*$, but does not address the attainment of the optimal values.

Proposition 3.1 (Min Common/Max Crossing Theorem I): Consider the min common and max crossing problems, and assume the following:

(1) $w^* < \infty$.

(2) The set

$$\overline{M} = \big\{(u,w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \le w \text{ and } (u,\overline{w}) \in M\big\}$$

is convex.

Then, we have $q^* = w^*$ if and only if for every sequence $\big\{(u_k,w_k)\big\} \subset M$ with $u_k \to 0$, there holds $w^* \le \liminf_{k\to\infty} w_k$.

Proof: If $w^* = -\infty$, by weak duality, we also have $q^* = -\infty$ and $q(\mu) = -\infty$ for all $\mu \in \Re^n$, so the conclusion trivially follows. We thus focus on the case where w^* is a real number. Assume that for every sequence $\big\{(u_k,w_k)\big\} \subset M$ with $u_k \to 0$, there holds $w^* \le \liminf_{k\to\infty} w_k$. We first note that $(0,w^*)$ is a closure point of $\overline{M}$, since by the definition of w^*, there exists a sequence $\big\{(0,w_k)\big\}$ that belongs to M, and hence also to $\overline{M}$, and is such that $w_k \to w^*$.

We next show by contradiction that $\overline{M}$ does not contain any vertical lines. If this were not so, by convexity of $\overline{M}$, the direction $(0,-1)$ would be a direction of recession of $\mathrm{cl}(\overline{M})$ (although not necessarily a direction of recession of $\overline{M}$), and hence also a direction of recession of $\mathrm{ri}(\overline{M})$ [cf. [BNO03], Prop. 1.5.1(d)]. Because $(0,w^*)$ is a closure point of $\overline{M}$, it is also a closure point of $\mathrm{ri}(\overline{M})$ [cf. [BNO03], Prop. 1.4.3(a)], and therefore, there exists a sequence $\big\{(u_k,w_k)\big\} \subset \mathrm{ri}(\overline{M})$ converging to $(0,w^*)$. Since $(0,-1)$ is a direction of recession of $\mathrm{ri}(\overline{M})$, the sequence $\big\{(u_k,w_k-1)\big\}$ belongs to $\mathrm{ri}(\overline{M})$ and consequently, $\big\{(u_k,w_k-1)\big\} \subset \overline{M}$. In view of the definition of $\overline{M}$, there is a sequence $\big\{(u_k,\overline{w}_k)\big\} \subset M$ with $\overline{w}_k \le w_k - 1$ for all k, so that $\liminf_{k\to\infty} \overline{w}_k \le w^* - 1$. This contradicts the assumption $w^* \le \liminf_{k\to\infty} w_k$, since $u_k \to 0$.

We now prove that the vector $(0,w^*-\epsilon)$ does not belong to $\mathrm{cl}(\overline{M})$ for any $\epsilon > 0$. To arrive at a contradiction, suppose that $(0,w^*-\epsilon)$ is a closure point of $\overline{M}$ for some $\epsilon > 0$, so that there exists a sequence $\big\{(u_k,w_k)\big\} \subset \overline{M}$ converging to $(0,w^*-\epsilon)$. In view of the definition of $\overline{M}$, this implies the existence of another sequence $\big\{(u_k,\overline{w}_k)\big\} \subset M$ with $u_k \to 0$ and $\overline{w}_k \le w_k$ for all k, and we have that $\liminf_{k\to\infty} \overline{w}_k \le w^*-\epsilon$, which contradicts the assumption $w^* \le \liminf_{k\to\infty} w_k$.

Since, as shown above, $\overline{M}$ does not contain any vertical lines and the vector $(0,w^*-\epsilon)$ does not belong to $\mathrm{cl}(\overline{M})$ for any $\epsilon > 0$, by Prop. 2.6(b), it follows that there exists a nonvertical hyperplane strictly separating $(0,w^*-\epsilon)$ and $\overline{M}$. This hyperplane crosses the $(n+1)$st axis at a unique vector $(0,\xi)$, which must lie between $(0,w^*-\epsilon)$ and $(0,w^*)$, i.e., $w^*-\epsilon \le \xi \le w^*$.

Furthermore, ξ cannot exceed the optimal value q^* of the max crossing problem, which, together with weak duality ($q^* \leq w^*$), implies that $w^* - \epsilon \leq q^* \leq w^*$. Since ϵ can be arbitrarily small, it follows that $q^* = w^*$.

Conversely, assume that $q^* = w^*$. Let $\big\{(u_k, w_k)\big\}$ be any sequence in M, which is such that $u_k \to 0$. Then,

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} \leq w_k + \mu' u_k, \qquad \forall\ k, \quad \forall\ \mu \in \Re^n.$$

Taking the limit as $k \to \infty$, we obtain

$$q(\mu) \leq \liminf_{k\to\infty} w_k, \qquad \forall\ \mu \in \Re^n,$$

implying that

$$w^* = q^* = \sup_{\mu\in\Re^n} q(\mu) \leq \liminf_{k\to\infty} w_k.$$

Q.E.D.

We now provide another version of the Min Common/Max Crossing Theorem, which in addition to the equality $q^* = w^*$, guarantees the attainment of the maximum crossing point by a nonvertical hyperplane under an additional relative interior assumption [see condition (3) of the proposition].

Proposition 3.2 (Min Common/Max Crossing Theorem II): Consider the min common and max crossing problems, and assume the following:

(1) $-\infty < w^*$.

(2) The set

$$\overline{M} = \big\{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in M\big\}$$

is convex.

(3) The set

$$D = \big\{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in \overline{M}\big\}$$

contains the origin in its relative interior.

Then $q^* = w^*$ and the optimal solution set of the max crossing problem, $Q^* = \big\{\mu \mid q(\mu) = q^*\big\}$, has the form

$$Q^* = \big(\mathrm{aff}(D)\big)^\perp + \tilde{Q},$$

where $\tilde{Q}$ is a nonempty, convex, and compact set, and $\big(\mathrm{aff}(D)\big)^\perp$ is the orthogonal complement of $\mathrm{aff}(D)$ [which is a subspace by assumption (3)]. Furthermore, Q^* is nonempty and compact if and only if D contains the origin in its interior.

Proof: We first show that $q^* = w^*$ and that Q^* is nonempty. We note that condition (3) implies that $w^* < \infty$, so in view of condition (1), w^* is a real number. Since w^* is the optimal min common value and the line $\big\{(0, w) \mid w \in \Re\big\}$ is contained in the affine hull of $\overline{M}$, it follows that $(0, w^*)$ is not a relative interior point of $\overline{M}$. Therefore, by the Proper Separation Theorem (cf. Prop. 2.3), there exists a hyperplane that passes through $(0, w^*)$, contains $\overline{M}$ in one of its closed halfspaces, but does not fully contain $\overline{M}$, i.e., there exists a vector (μ, β) such that

$$\beta w^* \leq \mu' u + \beta w, \qquad \forall\ (u, w) \in \overline{M}, \tag{3.5}$$

$$\beta w^* < \sup_{(u,w)\in\overline{M}} \{\mu' u + \beta w\}. \tag{3.6}$$

Since for any $(\overline{u}, \overline{w}) \in M$, the set $\overline{M}$ contains the halfline $\big\{(\overline{u}, w) \mid \overline{w} \leq w\big\}$, it follows from Eq. (3.5) that $\beta \geq 0$. If $\beta = 0$, then from Eq. (3.5), we have

$$0 \leq \mu' u, \qquad \forall\ u \in D.$$

Thus, the linear function $\mu' u$ attains its minimum over the set D at 0, which is a relative interior point of D by condition (3). Since D is convex, being the projection on the space of u of the set

$\overline{M}$, which is convex by assumption (2), it follows that $\mu' u$ is constant over D, i.e.,

$$\mu' u = 0, \qquad \forall\, u \in D,$$

(see [BNO03], Prop. 1.4.2). This, however, contradicts Eq. (3.6). Therefore, we must have $\beta > 0$, and by appropriate normalization if necessary, we can assume that $\beta = 1$. From Eq. (3.5), we then obtain

$$w^* \leq \inf_{(u,w)\in \overline{M}} \{\mu' u + w\} \leq \inf_{(u,w)\in M} \{\mu' u + w\} = q(\mu) \leq q^*.$$

Since the inequality $q^* \leq w^*$ holds always [cf. Eq. (3.4)], equality holds throughout in the above relation, and we must have $q(\mu) = q^* = w^*$. Thus Q^* is nonempty, and since $Q^* = \big\{\mu \mid q(\mu) \geq q^*\big\}$ and q is concave and upper semicontinuous, it follows that Q^* is also convex and closed.

We next show that $Q^* = \big(\mathrm{aff}(D)\big)^\perp + \tilde{Q}$. We first prove that the recession cone R_{Q^*} and the lineality space L_{Q^*} of Q^* are both equal to $\big(\mathrm{aff}(D)\big)^\perp$. The proof of this is based on the generic relation $L_{Q^*} \subset R_{Q^*}$ and the following two relations

$$\big(\mathrm{aff}(D)\big)^\perp \subset L_{Q^*}, \qquad R_{Q^*} \subset \big(\mathrm{aff}(D)\big)^\perp,$$

which we show next.

Let y be a vector in $\big(\mathrm{aff}(D)\big)^\perp$, so that $y'u = 0$ for all $u \in D$. For any vector $\mu \in Q^*$ and any scalar α, we then have

$$q(\mu + \alpha y) = \inf_{(u,w)\in \overline{M}} \big\{(\mu + \alpha y)' u + w\big\} = \inf_{(u,w)\in \overline{M}} \{\mu' u + w\} = q(\mu),$$

implying that $\mu + \alpha y$ is in Q^*. Hence $y \in L_{Q^*}$, and it follows that $\big(\mathrm{aff}(D)\big)^\perp \subset L_{Q^*}$.

Let y be a vector in R_{Q^*}, so that for any $\mu \in Q^*$ and $\alpha \geq 0$,

$$q(\mu + \alpha y) = \inf_{(u,w)\in \overline{M}} \big\{(\mu + \alpha y)' u + w\big\} = q^*.$$

Since $0 \in \mathrm{ri}(D)$, for any $u \in \mathrm{aff}(D)$, there exists a positive scalar γ such that the vectors γu and $-\gamma u$ are in D. By the definition of D, there exist scalars w^+ and w^- such that the pairs $(\gamma u, w^+)$ and $(-\gamma u, w^-)$ are in $\overline{M}$. Using the preceding equation, it follows that for any $\mu \in Q^*$, we have

$$(\mu + \alpha y)'(\gamma u) + w^+ \geq q^*, \qquad \forall\, \alpha \geq 0,$$

$$(\mu + \alpha y)'(-\gamma u) + w^- \geq q^*, \qquad \forall\, \alpha \geq 0.$$

If $y'u \neq 0$, then for sufficiently large $\alpha \geq 0$, one of the preceding two relations will be violated. Thus we must have $y'u = 0$, showing that $y \in \big(\mathrm{aff}(D)\big)^\perp$ and implying that

$$R_{Q^*} \subset \big(\mathrm{aff}(D)\big)^\perp.$$

This relation, together with the generic relation $L_{Q^*} \subset R_{Q^*}$ and the relation $\big(\mathrm{aff}(D)\big)^\perp \subset L_{Q^*}$ shown earlier, shows that

$$\big(\mathrm{aff}(D)\big)^\perp \subset L_{Q^*} \subset R_{Q^*} \subset \big(\mathrm{aff}(D)\big)^\perp.$$

Therefore

$$L_{Q^*} = R_{Q^*} = \big(\mathrm{aff}(D)\big)^\perp.$$

We decompose the convex set Q^* along its lineality space and its orthogonal complement as

$$Q^* = L_{Q^*} + (Q^* \cap L_{Q^*}^\perp),$$

(see [BNO03], Prop. 1.5.4). Since $L_{Q^*} = \big(\mathrm{aff}(D)\big)^\perp$, we obtain

$$Q^* = \big(\mathrm{aff}(D)\big)^\perp + \tilde{Q},$$

where $\tilde{Q} = Q^* \cap \mathrm{aff}(D)$. Furthermore, we have

$$R_{\tilde{Q}} = R_{Q^*} \cap R_{\mathrm{aff}(D)}.$$

Since $R_{Q^*} = \big(\mathrm{aff}(D)\big)^\perp$, as shown earlier, and $R_{\mathrm{aff}(D)} = \mathrm{aff}(D)$, the recession cone $R_{\tilde{Q}}$ consists of the zero vector only, implying that the set $\tilde{Q}$ is compact.

Finally, to show the last statement in the proposition, we note that 0 is an interior point of D if and only if $\mathrm{aff}(D) = \Re^n$, which in turn is equivalent to Q^* being equal to the compact set $\tilde{Q}$. **Q.E.D.**

We next provide a min common/max crossing duality theorem, which involves polyhedral convexity assumptions. In particular, the definition of the set M will involve a linear mapping and a polyhedral cone of $\Re^r$, i.e., a set of the form

$$P = \{y \mid Ey \le 0\},$$

where E is a matrix with r columns. Note that the polar cone of P, denoted P^*, is given by

$$P^* = \{z \mid z'y \le 0, \ \ \forall\, y \in P\} = \{z \mid z = E'\zeta, \ \zeta \ge 0\}.$$

The theorem is new in the form given here, although its assumptions are related to constructions that are implicit in classical duality analyses under polyhedral convexity assumptions.

Proposition 3.3 (Min Common/Max Crossing Theorem III): Consider the min common and max crossing problems, and assume the following:

(1) The set M is defined in terms of a convex set $V \subset \Re^{m+1}$, an $r \times m$ matrix A, an $r \times n$ matrix B, a vector b in $\Re^r$, and a polyhedral cone $P \subset \Re^r$ as follows:

$$M = \big\{(u, w) \mid u \in \Re^n, \text{ and there is a vector } (x, w) \in V \\ \text{such that } Ax - b - Bu \in P\big\}.$$

(2) There exists a vector $(\overline{x}, \overline{w})$ in the relative interior of V such that $A\overline{x} - b \in P$.

Then $q^* = w^*$ and there exists a vector μ in the polar cone P^* such that $q(B'\mu) = q^*$.

Proof: If $w^* = -\infty$, then the conclusion holds since, by weak duality, we have $q^* \leq w^*$, so that $q^* = w^* = -\infty$, and $q(B'\mu) = q^*$ for all μ, including the vector $\mu = 0$, which belongs to P^*. We may thus assume that $-\infty < w^*$, which also implies w^* is finite, since the min common problem has a feasible solution in view of the assumptions (1) and (2). Consider the convex subsets of $\Re^{m+1}$ defined by

$$C_1 = \big\{(x, v) \mid \text{there is a vector } (x, w) \in V \text{ such that } v > w\big\},$$

$$C_2 = \big\{(x, w^*) \mid Ax - b \in P\big\}$$

(cf. Fig. 3.2). The set C_1 is nonempty since $(\overline{x}, v) \in C_1$ for all $v > \overline{w}$, while the set C_2 is nonempty since $(\overline{x}, w^*) \in C_2$. Finally, C_1 and C_2 are disjoint. To see this, note that

$$w^* = \inf_{(x,w) \in V,\, Ax - b \in P} w, \tag{3.7}$$

and if $(x, w^*) \in C_1 \cap C_2$, by the definition of C_2, we must have $Ax - b \in P$, while by the definition of C_1, we must have $w^* > w$ for some w with $(x, w) \in V$, contradicting Eq. (3.7).

Figure 3.2. Illustration of the sets C_1 and C_2, and the hyperplane separating them in the proof of Prop. 3.3. Note that since $(\overline{x}, \overline{w})$ is a relative interior point of V, all the vectors $(\overline{x}, v)$ with $v > \overline{w}$ belong to the relative interior of C_1.

Since $C_1 \cap C_2 = \emptyset$ and C_2 is polyhedral, by Prop. 2.5, there exists a hyperplane that separates C_1 and C_2, and does not contain C_1, i.e., a vector (ξ, β) such that

$$\beta w^* + \xi' z \leq \beta v + \xi' x, \qquad \forall\ (x, v) \in C_1,\ \forall\ z \text{ such that } Az - b \in P, \tag{3.8}$$

$$\inf_{(x,v)\in C_1} \big\{\beta v + \xi' x\big\} < \sup_{(x,v)\in C_1} \big\{\beta v + \xi' x\big\}. \tag{3.9}$$

If $\beta = 0$, then from Eq. (3.8), we have

$$\xi' \overline{x} \leq \sup_{Az-b\in P} \xi' z \leq \inf_{(x,v)\in C_1} \xi' x \leq \xi' \overline{x}.$$

Thus, equality holds throughout in the preceding relation, implying that all the vectors $(\overline{x}, v)$ with $v > \overline{w}$ minimize the linear function $(\xi, 0)'(x, v)$ over the set C_1. Since $(\overline{x}, \overline{w})$ is a relative interior point of V, all these vectors are relative interior points of C_1 (cf. Fig. 3.2). It follows that the linear function $(\xi, 0)'(x, v)$ is constant over C_1 (see [BNO03], Prop. 1.4.2). This, however, contradicts Eq. (3.9). Therefore, we must have $\beta \neq 0$.

By using Eq. (3.8) with $z = \overline{x}$ and $v > \overline{w}$, we obtain $\beta w^* + \xi' \overline{x} \leq \beta v + \xi' \overline{x}$, or $\beta w^* \leq \beta v$. Since $w^* \leq \overline{w} < v$ and $\beta \neq 0$, it follows that $\beta > 0$, and by normalizing (ξ, β) if necessary, we may assume that $\beta = 1$. Thus, from Eq. (3.8) and the definition of C_1, we have

$$\sup_{Az-b\in P} \big\{w^* + \xi' z\big\} \leq \inf_{(x,w)\in V} \big\{w + \xi' x\big\}. \tag{3.10}$$

Let $\{y \mid Dy \leq 0\}$ be a representation of the polyhedral cone P in terms of a matrix D. Then the maximization problem on the left-hand side of Eq. (3.10) involves the linear program

$$\begin{aligned} &\text{maximize} \quad \xi' z \\ &\text{subject to} \quad DAz - Db \leq 0. \end{aligned} \tag{3.11}$$

Since the minimization problem in the right-hand side of Eq. (3.10) is feasible by assumption, the linear program (3.11) is bounded (as well as feasible) and therefore has an optimal solution, which is denoted by z^* (see [BNO03], Prop. 2.3.4). Let c_j' be the rows of DA, let $(Db)_j$ denote the corresponding components of Db, and let

$$J = \big\{j \mid c_j' z^* = (Db)_j\big\}.$$

If $J = \emptyset$, then z^* lies in the interior of the constraint set of problem (3.11), so we must have $\xi = 0$. If $J \neq \emptyset$ and y is such that $c_j' y \leq 0$ for all $j \in J$, then there is a small enough $\epsilon > 0$ such that $A(z^* + \epsilon y) - b \in P$, and the optimality of z^* implies that $\xi'(z^* + \epsilon y) \leq \xi' z^*$ or $\xi' y \leq 0$. Hence, by Farkas' Lemma, there exist scalars $\zeta_j \geq 0$, $j \in J$, such that

$$\xi = \sum_{j\in J} \zeta_j c_j.$$

Thus, by defining $\zeta_j = 0$ for $j \notin J$, we see that for the vector ζ, we have

$$\xi = A'D'\zeta, \qquad \zeta'(DAz^* - Db) = 0.$$

Let $\mu = D'\zeta$, and note that since $\zeta \geq 0$, by Farkas' Lemma, we have $\mu \in P^*$. Furthermore, the preceding relations can be written as

$$\xi = A'\mu, \qquad \mu'(Az^* - b) = 0,$$

from which we obtain

$$\xi' z^* = \mu' A z^* = \mu' b.$$

Thus, from Eq. (3.10) and the equalities $\xi = A'\mu$ and $\xi'z^* = \mu'b$, we have $w^* + \mu'b \leq \inf_{(x,w)\in V}\big\{w + \mu'Ax\big\}$ or equivalently,

$$w^* \leq \inf_{(x,w)\in V}\big\{w + \mu'(Ax - b)\big\}.$$

Since $\mu \in P^*$, we have $\mu'(Ax - b) \leq \mu'Bu$ for all (x, u) such that $u \in \Re^n$ and $Ax - b - Bu \in P$, so that

$$\begin{aligned}
\inf_{(x,w)\in V}\big\{w + \mu'(Ax - b)\big\} &\leq \inf_{\substack{(x,w)\in V,\, u\in\Re^n \\ Ax-b-Bu\in P}} \big\{w + \mu'(Ax - b)\big\} \\
&\leq \inf_{\substack{(x,w)\in V,\, u\in\Re^n \\ Ax-b-Bu\in P}} \{w + \mu'Bu\} \\
&= \inf_{(u,w)\in M} \{w + \mu'Bu\} \\
&= q(B'\mu) \\
&\leq q^*.
\end{aligned}$$

By combining the preceding relations, we obtain $w^* \leq q(B'\mu) \leq q^*$. On the other hand, by the weak duality relation, we have $q^* \leq w^*$, so that $q(B'\mu) = q^* = w^*$. **Q.E.D.**

4. MINIMAX AND SADDLE POINT THEOREMS

Suppose that we are given a function $\phi : X \times Z \mapsto \Re$, where $X \subset \Re^n$, $Z \subset \Re^m$, and we want to either

$$\begin{aligned}
&\text{minimize} \quad \sup_{z\in Z} \phi(x, z) \\
&\text{subject to} \quad x \in X
\end{aligned} \tag{4.1}$$

or

$$\begin{aligned}
&\text{maximize} \quad \inf_{x\in X} \phi(x, z) \\
&\text{subject to} \quad z \in Z.
\end{aligned} \tag{4.2}$$

We want to derive conditions guaranteeing that the optimal values of these two problems are equal, i.e.,

$$\sup_{z\in Z}\inf_{x\in X}\phi(x,z)=\inf_{x\in X}\sup_{z\in Z}\phi(x,z), \tag{4.3}$$

and that a *saddle point* of ϕ exists, i.e., a pair of vectors (x^*,z^*) exists such that $x^*\in X$ and $z^*\in Z$, and

$$\phi(x^*,z)\le\phi(x^*,z^*)\le\phi(x,z^*),\qquad\forall\ x\in X,\quad\forall\ z\in Z.$$

It is well known that saddle points are related to optimal solutions of problems (4.1) and (4.2) as in the following proposition.

Proposition 4.1: A pair (x^*,z^*) is a saddle point of ϕ if and only if x^* and z^* are optimal solutions of problems (4.1) and (4.2), respectively, and the minimax equality (4.3) holds.

In the following analysis, a critical role is played by the min common/max crossing framework of Section 3 and by the function $p:\Re^m\mapsto[-\infty,\infty]$ given by

$$p(u)=\inf_{x\in X}\sup_{z\in Z}\big\{\phi(x,z)-u'z\big\},\qquad\forall\ u\in\Re^m. \tag{4.4}$$

This function, whose significance is well understood in both minimax theory and Lagrangian duality (see e.g., Rockafellar [Roc70], Borwein and Lewis [BoL00]), defines how the "infsup" of the function ϕ changes when the linear perturbation term $u'z$ is subtracted from ϕ. It turns out that if p changes in a "regular" manner to be specified shortly, the minimax equality (4.3) is guaranteed.

In the subsequent applications of the min common/max crossing framework, the set M will be taken to be the epigraph of p,

$$M=\mathrm{epi}(p),$$

so that the min common value w^* will be equal to $p(0)$, which by the definition of p, is also equal to the "infsup" value

$$w^*=p(0)=\inf_{x\in X}\sup_{z\in Z}\phi(x,z). \tag{4.5}$$

Under some convexity assumptions with respect to x (see the subsequent Lemma 4.1), we will show that p is convex, so that M is convex, which satisfies a major assumption for the application of the min common/max crossing theorems of the preceding section (with M equal to an epigraph of a function, the sets M and $\overline{M}$ appearing in the min common/max crossing theorems coincide).

The corresponding max crossing problem is [cf. Eqs. (3.3) and (3.2)]

$$\begin{aligned}&\text{maximize}\quad q(\mu)\\&\text{subject to}\quad \mu\in\Re^n,\end{aligned}$$

where

$$q(\mu) = \inf_{(u,w)\in \mathrm{epi}(p)} \{w + \mu' u\} = \inf_{\{(u,w)\mid p(u)\le w\}} \{w + \mu' u\} = \inf_{u\in\Re^m} \big\{p(u) + \mu' u\big\}.$$

By using this relation and the definition of p, we obtain

$$q(\mu) = \inf_{u\in\Re^m} \inf_{x\in X} \sup_{z\in Z} \big\{\phi(x,z) + u'(\mu - z)\big\}.$$

For every $\mu \in Z$, by setting $z = \mu$ in the right-hand side above, we obtain

$$\inf_{x\in X} \phi(x,\mu) \le q(\mu), \qquad \forall\ \mu \in Z.$$

Thus, using also Eq. (4.5) and the weak duality relation $q^* \le w^*$, we have

$$\sup_{z\in Z} \inf_{x\in X} \phi(x,z) \le \sup_{\mu\in\Re^m} q(\mu) = q^* \le w^* = p(0) = \inf_{x\in X} \sup_{z\in Z} \phi(x,z). \tag{4.6}$$

This inequality indicates a generic connection of the minimax and the min common/max crossing frameworks. In particular, if the minimax equality

$$\sup_{z\in Z} \inf_{x\in X} \phi(x,z) = \inf_{x\in X} \sup_{z\in Z} \phi(x,z)$$

holds, then $q^* = w^*$, i.e., that the optimal values of the min common and max crossing problems are equal.

Figure 4.1. Min common/max crossing framework for minimax theory. The set M will be taken to be the epigraph of the function

$$p(u) = \inf_{x\in X} \sup_{z\in Z} \{\phi(x,z) - u'z\}.$$

Under suitable assumptions, the "infsup" and "supinf" values of ϕ will turn out to be equal to the min common value w^* and the max crossing value q^*, respectively. Figures (a) and (b) illustrate the cases where the minimax equality (4.3) holds and does not hold, respectively.

An even stronger connection between the minimax and the min common/max crossing frameworks holds under some convexity and semicontinuity assumptions, as shown in the following two lemmas. Loosely phrased, these lemmas assert that:

(a) Convexity with respect to x [convexity of X and $\phi(\cdot, z)$ for all $z \in Z$] guarantees that epi(p) is a convex set, thereby allowing the use of the two Min Common/Max Crossing Theorems of the preceding section (Props. 3.1 and 3.2).

(b) Concavity and semicontinuity with respect to z [convexity of Z, and concavity and upper semicontinuity of $\phi(x, \cdot)$ for all $x \in X$] guarantee that

$$q(\mu) = \inf_{x \in X} \phi(x, \mu), \qquad q^* = \sup_{z \in Z} \inf_{x \in X} \phi(x, z),$$

as indicated in Fig. 4.1. Thus, under these conditions, the minimax equality (4.3) is equivalent to the equality $q^* = w^*$ in the corresponding min common/max crossing framework.

Thus, if ϕ is convex with respect to x, and concave and upper semicontinuous with respect to z, as specified in the following two lemmas, the min common/max crossing framework applies in its most powerful form and provides the answers to the most critical questions within the minimax framework.

Lemma 4.1: Let X be a nonempty convex subset of $\Re^n$, let Z be a nonempty subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $\phi(\cdot, z) : X \mapsto \Re$ is convex. Then the function p of Eq. (4.4) is convex.

Proof: Let u and v be such that $p(u) < \infty$ and $p(v) < \infty$. Rewriting $p(u)$ as

$$p(u) = \inf_{x \in X} l(x, u),$$

where $l(x, u) = \sup_{z \in Z} \big\{\phi(x, z) - u'z\big\}$, we have that there exist sequences $\{x_k\}$ and $\{y_k\}$ in X such that

$$l(x_k, u) \to p(u), \qquad l(y_k, v) \to p(v).$$

By convexity of X, we have $\alpha x_k + (1-\alpha) y_k \in X$ for all $\alpha \in [0, 1]$ and all k. Using the convexity of $\phi(\cdot, z)$ for each $z \in Z$, we obtain

$$\begin{aligned}
p\big(\alpha u + (1-\alpha) v\big) &\le l\big(\alpha x_k + (1-\alpha) y_k, \alpha u + (1-\alpha) v\big) \\
&= \sup_{z \in Z} \Big\{\phi\big(\alpha x_k + (1-\alpha) y_k, z\big) - \big(\alpha u + (1-\alpha) v\big)' z\Big\} \\
&\le \sup_{z \in Z} \Big\{\alpha \phi(x_k, z) + (1-\alpha) \phi(y_k, z) - \big(\alpha u + (1-\alpha) v\big)' z\Big\} \\
&\le \alpha \sup_{z \in Z} \big\{\phi(x_k, z) - u'z\big\} + (1-\alpha) \sup_{z \in Z} \big\{\phi(y_k, z) - v'z\big\} \\
&= \alpha l(x_k, u) + (1-\alpha) l(y_k, v).
\end{aligned}$$

Since

$$\alpha l(x_k, u) + (1-\alpha) l(y_k, v) \to \alpha p(u) + (1-\alpha) p(v),$$

it follows that

$$p\big(\alpha u + (1-\alpha)v\big) \le \alpha p(u) + (1-\alpha)p(v),$$

implying that p is convex. **Q.E.D.**

The following lemma shows that under some convexity and semicontinuity assumptions, p defines not only the "infsup" of the function ϕ [cf. Eq. (4.5)], but through its epigraph, it also defines the "supinf" (cf. Fig. 4.1).

Lemma 4.2: Let X be a nonempty subset of $\Re^n$, let Z be a nonempty convex subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $x \in X$, the function $-\phi(x,\cdot) : Z \mapsto \Re$ is closed and convex. Then the function $q : \Re^m \mapsto [-\infty, \infty]$ given by

$$q(\mu) = \inf_{(u,w)\in \mathrm{epi}(p)} \big\{w + u'\mu\big\}, \qquad \mu \in \Re^m,$$

where p is given by Eq. (4.4), satisfies

$$q(\mu) = \begin{cases} \inf_{x\in X} \phi(x,\mu) & \text{if } \mu \in Z, \\ -\infty & \text{if } \mu \notin Z. \end{cases} \tag{4.7}$$

Furthermore, we have $q^* = w^*$ if and only if the minimax equality (4.3) holds.

Proof: For every $\mu \in \Re^m$, we have

$$q(\mu) = \inf_{(u,w)\in \mathrm{epi}(p)} \{w + \mu'u\} = \inf_{\{(u,w)\mid p(u)\le w\}} \{w + \mu'u\} = \inf_{u\in\Re^m} \big\{p(u) + \mu'u\big\}.$$

By using this relation and the definition of p, we obtain for every $\mu \in \Re^m$,

$$\begin{aligned} q(\mu) &= \inf_{u\in\Re^m} \big\{p(u) + u'\mu\big\} \\ &= \inf_{u\in\Re^m} \inf_{x\in X} \sup_{z\in Z} \big\{\phi(x,z) + u'(\mu - z)\big\} \\ &= \inf_{x\in X} \inf_{u\in\Re^m} \sup_{z\in Z} \big\{\phi(x,z) + u'(\mu - z)\big\}. \end{aligned} \tag{4.8}$$

For $\mu \in Z$, we have

$$\sup_{z\in Z} \big\{\phi(x,z) + u'(\mu - z)\big\} \ge \phi(x,\mu), \qquad \forall\ x \in X, \quad \forall\ u \in \Re^m,$$

implying that

$$q(\mu) \ge \inf_{x\in X} \phi(x,\mu), \qquad \forall\ \mu \in Z.$$

Thus, to prove Eq. (4.7), we must show that

$$q(\mu) \le \inf_{x\in X} \phi(x,\mu), \qquad \forall\ \mu \in Z, \tag{4.9}$$

and

$$q(\mu) = -\infty, \qquad \forall\ \mu \notin Z.$$

For all $x \in X$ and $z \in Z$, denote

$$r_x(z) = -\phi(x, z),$$

so that the function $r_x : Z \mapsto \Re$ is closed and convex by assumption. We will consider separately the two cases where $\mu \in Z$ and $\mu \notin Z$. We first assume that $\mu \in Z$. We fix an arbitrary $x \in X$, and we note that by assumption, $\mathrm{epi}\big(r_x\big)$ is a closed convex set. Since $\mu \in Z$, the point $\big(\mu, r_x(\mu)\big)$ belongs to $\mathrm{epi}(r_x)$. For some $\epsilon > 0$, we consider the point $\big(\mu, r_x(\mu) - \epsilon\big)$, which does not belong to $\mathrm{epi}(r_x)$. By the definition of r_x, $r_x(z)$ is finite for all $z \in Z$, Z is nonempty, and $\mathrm{epi}(r_x)$ is closed, so that $\mathrm{epi}(r_x)$ does not contain any vertical lines. Therefore, by Prop. 2.6(b), there exists a nonvertical hyperplane that strictly separates the point $\big(\mu, r_x(\mu) - \epsilon\big)$ from $\mathrm{epi}(r_x)$, i.e., a vector $(\overline{u}, \zeta)$ with $\zeta \neq 0$, and a scalar c such that

$$\overline{u}'\mu + \zeta\big(r_x(\mu) - \epsilon\big) < c < \overline{u}'z + \zeta w, \qquad \forall\ (z, w) \in \mathrm{epi}(r_x).$$

Since w can be made arbitrarily large, we have $\zeta > 0$, and without loss of generality, we can take $\zeta = 1$. In particular for $w = r_x(z)$, with $z \in Z$, we have

$$\overline{u}'\mu + \big(r_x(\mu) - \epsilon\big) < \overline{u}'z + r_x(z), \qquad \forall\ z \in Z,$$

or equivalently,

$$\phi(x, z) + \overline{u}'(\mu - z) < \phi(x, \mu) + \epsilon, \qquad \forall\ z \in Z.$$

Letting $\epsilon \downarrow 0$, we obtain for all $x \in X$

$$\inf_{u \in \Re^m} \sup_{z \in Z} \big\{\phi(x, z) + u'(\mu - z)\big\} \leq \sup_{z \in Z} \big\{\phi(x, z) + \overline{u}'(\mu - z)\big\} \leq \phi(x, \mu).$$

By taking the infimum over $x \in X$ in the above relation, and by using Eq. (4.8), we see that Eq. (4.9) holds.

We now assume that $\mu \notin Z$. We consider a sequence $\{w_k\}$ with $w_k \to \infty$ and we fix an arbitrary $x \in X$. Since $\mu \notin Z$, the sequence $\big\{(\mu, w_k)\big\}$ does not belong to $\mathrm{epi}(r_x)$. Therefore, similar to the argument above, there exists a sequence of nonvertical hyperplanes with normals $(u_k, 1)$ such that

$$w_k + u_k'\mu < -\phi(x, z) + u_k'z, \qquad \forall\ z \in Z, \quad \forall\ k,$$

implying that

$$\phi(x, z) + u_k'(\mu - z) < -w_k, \qquad \forall\ z \in Z, \quad \forall\ k.$$

Thus, we have

$$\inf_{u\in\Re^m}\sup_{z\in Z}\big\{\phi(x,z)+u'(\mu-z)\big\}\le\sup_{z\in Z}\big\{\phi(x,z)+u_k'(\mu-z)\big\}\le -w_k,\qquad \forall\ k,$$

and by taking the limit in the preceding inequality as $k\to\infty$, we obtain

$$\inf_{u\in\Re^m}\sup_{z\in Z}\big\{\phi(x,z)+u'(\mu-z)\big\}=-\infty,\qquad \forall\ x\in X.$$

Using Eq. (4.8), we see that $q(\mu)=-\infty$. Thus, $q(\mu)$ has the form given in Eq. (4.7). The equality $q^*=w^*$ and the minimax equality are equivalent in view of the discussion following Eq. (4.6). **Q.E.D.**

The assumption that $-\phi(x,\cdot)$ is closed and convex in Lemma 4.2 is essential for Eq. (4.7) to hold. This can be seen by considering the special case where ϕ is independent of x, and by noting that q is concave and upper semicontinuous.

We now use the preceding two lemmas and the Min Common/Max Crossing Theorem I (cf. Prop. 3.1) to prove the following proposition.

Proposition 4.2: (Minimax Theorem I) Let X and Z be nonempty convex subsets of $\Re^n$ and $\Re^m$, respectively, and let $\phi : X\times Z\mapsto\Re$ be a function. Assume that for each $z\in Z$, the function $\phi(\cdot,z):X\mapsto\Re$ is convex, and for each $x\in X$, the function $-\phi(x,\cdot):Z\mapsto\Re$ is closed and convex. Assume further that

$$\inf_{x\in X}\sup_{z\in Z}\phi(x,z)<\infty.$$

Then, the minimax equality holds, i.e.,

$$\sup_{z\in Z}\inf_{x\in X}\ \phi(x,z)=\inf_{x\in X}\sup_{z\in Z}\phi(x,z),$$

if and only if the function p of Eq. (4.4) is lower semicontinuous at $u=0$, i.e., $p(0)\le\liminf_{k\to\infty}p(u_k)$ for all sequences $\{u_k\}$ with $u_k\to 0$.

Proof: The proof consists of showing that with an appropriate selection of the set M, the assumptions of the proposition are essentially equivalent to the corresponding assumptions of the Min Common/Max Crossing Theorem I.

We choose the set M (as well as the set $\overline{M}$) in the Min Common/Max Crossing Theorem I to be the epigraph of p,

$$M=\overline{M}=\big\{(u,w)\mid u\in\Re^m,\ p(u)\le w\big\},$$

which is convex in view of the assumed convexity of $\phi(\cdot, z)$ and Lemma 4.1. Thus, assumption (2) of the Min Common/Max Crossing Theorem I is satisfied.

From the definition of p, we have

$$w^* = p(0) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z).$$

It follows that the assumption

$$\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty$$

is equivalent to the assumption $w^* < \infty$ of the Min Common/Max Crossing Theorem I.

Finally, the condition

$$p(0) \leq \liminf_{k \to \infty} p(u_k)$$

for all sequences $\{u_k\}$ with $u_k \to 0$ is equivalent to the condition of the Min Common/Max Crossing Theorem I that for every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, there holds $w^* \leq \liminf_{k \to \infty} w_k$. Thus, by the conclusion of that theorem, the condition $p(0) \leq \liminf_{k \to \infty} p(u_k)$ holds if and only if $q^* = w^*$, which in turn holds if and only if the minimax equality holds [cf. Lemma 4.2, which applies because of the assumed closedness and convexity of $-\phi(x, \cdot)$]. **Q.E.D.**

The assumptions of Prop. 4.2 are satisfied under some easily verified conditions, as shown in the following proposition, which is related to results due to Rockafellar that were developed using considerably more complicated mathematical machinery (see [Roc70], Theorems 37.3 and 37.6). Our line of analysis is related to the one of Borwein and Lewis [BoL00], p. 96.

Proposition 4.3: Let X be a nonempty convex subset of $\Re^n$, let Z be a nonempty convex subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $t_z : \Re^n \mapsto (-\infty, \infty]$ defined by

$$t_z(x) = \begin{cases} \phi(x,z), & \text{if } x \in X, \\ \infty, & \text{if } x \notin X, \end{cases}$$

is closed and convex, and that for each $x \in X$, the function $r_x : \Re^m \mapsto (-\infty, \infty]$ defined by

$$r_x(z) = \begin{cases} -\phi(x,z) & \text{if } z \in Z, \\ \infty & \text{otherwise,} \end{cases}$$

is closed and convex. Assume further that

$$\inf_{x \in X} \sup_{z \in Z} \phi(x,z) < \infty$$

and that the set of common directions of recession of all the functions t_z, $z \in Z$, consists of the zero vector only. Then, the minimax equality holds, i.e.,

$$\sup_{z \in Z} \inf_{x \in X} \phi(x,z) = \inf_{x \in X} \sup_{z \in Z} \phi(x,z),$$

and the infimum over X in the right-hand side above is finite and is attained at a set of points that is nonempty and compact. Furthermore, the function p of Eq. (4.4) is closed, proper, and convex.

Proof: Consider the function $\bar{t} : \Re^n \mapsto (-\infty, \infty]$ defined by

$$\bar{t}(x) = \sup_{z \in Z} t_z(x),$$

which is closed and convex since all the functions t_z, $z \in Z$, are closed and convex, and has some nonempty level sets because of the assumption $\inf_{x \in X} \sup_{z \in Z} \phi(x,z) < \infty$, which can be written as $\inf_{x \in X} \bar{t}(x) < \infty$. A nonempty level set $\{x \mid \bar{t}(x) \leq \gamma\}$, where $\gamma \in \Re$, is equal to the intersection of the level sets $\{x \mid t_z(x) \leq \gamma\}$, $z \in Z$. Since the latter level sets are closed (in view of the closedness of t_z), the recession cone of their intersection, is equal to the intersection of their recession cones ([Roc70], Cor. 8.3.2), which consists of just the zero vector by assumption. Hence the recession cone of a nonempty level set of $\bar{t}$ consists of the zero vector only. It follows that the level sets of $\bar{t}$ are compact and by Weierstrass' Theorem, its set of minimizing points is nonempty and compact. In particular, $p(0)$ is finite.

We will prove that the epigraph of p,

$$\text{epi}(p) = \big\{(u, w) \mid p(u) \le w\big\},$$

is a closed set in $\Re^{m+1}$. This will imply that the assumptions of Prop. 4.2 are satisfied, thereby showing the desired assertions, except for the properness of p (the convexity of p follows from Lemma 4.1).

Let $\big\{(u_k, w_k)\big\}$ be a sequence in epi(p) that converges to some $(u, w) \in \Re^{m+1}$. Since $\big\{(u_k, w_k)\big\}$ is in epi(p), we have

$$p(u_k) = \inf_{x\in X} \sup_{z\in Z} \big\{\phi(x, z) - u_k' z\big\} \le w_k, \qquad \forall\ k.$$

Let ϵ be any positive scalar. Then, from the preceding relation we see that for each k, there exists a vector $x_k \in X$ such that

$$\sup_{z\in Z} \big\{\phi(x_k, z) - u_k' z\big\} \le w_k + \epsilon,$$

implying that

$$\phi(x_k, z) \le u_k' z + w_k + \epsilon, \qquad \forall\ k, \quad \forall\ z \in Z. \tag{4.10}$$

Therefore,

$$\phi(x_k, z) \le \overline{\gamma}_z, \qquad \forall\ k, \quad \forall\ z \in Z, \tag{4.11}$$

where the scalar $\overline{\gamma}_z$ defined by

$$\overline{\gamma}_z = \sup_k \{u_k' z + w_k + \epsilon\}$$

is finite, because $(u_k, w_k) \to (u, w)$. In view of Eq. (4.11), it follows that the sequence $\{x_k\}$ is contained in

$$\cap_{z\in Z} \big\{x \mid t_z(x) \le \overline{\gamma}_z\big\}. \tag{4.12}$$

By repeating the argument given in the beginning of the proof, we see that this set is compact. Hence the sequence $\{x_k\}$ belongs to the compact set (4.12), and all its limit points belong to this set and therefore also to X.

Let $\overline{x} \in X$ be a limit point of $\{x_k\}$, and without loss of generality assume that $x_k \to \overline{x}$. Taking the limit as $k \to \infty$ in Eq. (4.10), and using the fact $(u_k, w_k) \to (u, w)$ and the closedness of t_z for each $z \in Z$ [which implies lower semicontinuity of $\phi(\cdot, z)$], we obtain

$$\phi(\overline{x}, z) \le \liminf_{k\to\infty} \phi(x_k, z) \le u' z + w + \epsilon, \qquad \forall\ z \in Z.$$

This relation implies that

$$p(u) = \inf_{x\in X} \sup_{z\in Z} \big\{\phi(x, z) - u' z\big\} \le \sup_{z\in Z} \big\{\phi(\overline{x}, z) - u' z\big\} \le w + \epsilon.$$

Letting $\epsilon \downarrow 0$, we have $p(u) \leq w$, showing that $(u, w) \in \mathrm{epi}(p)$ and that $\mathrm{epi}(p)$ is closed.

Finally, to show that p is proper, i.e., $-\infty < p(u)$ for all $u \in \Re^n$, we argue by contradiction. If $p(\overline{u}) = -\infty$ for some $\overline{u} \in \Re^m$, then by the finiteness of $p(0)$ and the convexity of p (cf. Lemma 4.1), we must have $p(\alpha\overline{u}) \leq \alpha p(\overline{u}) + (1 - \alpha)p(0) = -\infty$ for all $\alpha \in (0, 1]$. This contradicts the condition $p(0) \leq \liminf_{k \to \infty} p(u_k)$ for all sequences $\{u_k\}$ with $u_k \to 0$, which was implied by the fact shown earlier that p is closed. **Q.E.D.**

The assumption of Prop. 4.3 that there is no nonzero common direction of recession of all the functions t_z, $z \in Z$, is satisfied under any one of the following conditions:

(1) The set X is compact.

(2) There exists a vector $\overline{z} \in Z$ such that all the sets $\{x \in X \mid \phi(x, \overline{z}) \leq \gamma\}$, $\gamma \in \Re$, are compact.

The conclusions of Prop. 4.3 using condition (2) in place of our directions of recession assumption have been given by Borwein and Lewis [BoL00], p. 96.

As a corollary of the preceding proposition, we have the following result, which contains as special cases the classical saddle point theorems of von Neumann [Neu28] and Kakutani [Kak41].

Proposition 4.4: (Saddle Point Theorem) Let X be a nonempty convex subset of $\Re^n$, let Z be a nonempty convex subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function such that either

$$-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z),$$

or

$$\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty.$$

Assume that for each $z \in Z$, the function $t_z : \Re^n \mapsto (-\infty, \infty]$ defined by

$$t_z(x) = \begin{cases} \phi(x, z), & \text{if } x \in X, \\ \infty, & \text{if } x \notin X, \end{cases}$$

is closed and convex, and that for each $x \in X$, the function $r_x : \Re^m \mapsto (-\infty, \infty]$ defined by

$$r_x(z) = \begin{cases} -\phi(x, z) & \text{if } z \in Z, \\ \infty & \text{otherwise}, \end{cases}$$

is closed and convex. Assume further that the set of common directions of recession of all the functions t_z, $z \in Z$, consists of the zero vector only, and that the set of common directions of recession of all the functions r_x, $x \in X$, consists of the zero vector only. Then, the set of saddle points of ϕ is nonempty and compact.

Proof: If $\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty$, we apply Prop. 4.3 to show that the minimax equality holds and that the infimum over X is attained at a nonempty and compact set. We then reverse the roles of x and z and the sign of ϕ, and apply Prop. 4.3 again to show that the supremum over Z is attained.

If $\inf_{x \in X} \sup_{z \in Z} \phi(x, z) = \infty$, we have $-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z)$ by condition (1). We then reverse the roles of x and z and the sign of ϕ, and apply the preceding argument in conjunction with Prop. 4.3. **Q.E.D.**

Note that the assumptions of Prop. 4.4, which relate to directions of recession of t_z and r_x, as well as the condition that either $-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z)$, or $\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty$, are satisfied under any one of the following four conditions:

(1) X and Z are compact.

(2) Z is compact and there exists a vector $\overline{z} \in Z$ such that all the sets $\{x \in X \mid \phi(x, \overline{z}) \le \gamma\}$,

$\gamma \in \Re$, are compact.

(3) X is compact and there exists a vector $\overline{x} \in X$ such that all the sets $\{z \in Z \mid \phi(\overline{x}, z) \geq \gamma\}$, $\gamma \in \Re$, are compact.

(4) There exist vectors $\overline{x} \in X$ and $\overline{z} \in Z$ such that all the level sets $\{x \in X \mid \phi(x, \overline{z}) \leq \gamma\}$, $\gamma \in \Re$, and all the level sets $\{z \in Z \mid \phi(\overline{x}, z) \geq \gamma\}$, $\gamma \in \Re$, are compact.

A proof of the saddle-point theorem under each of the above four conditions (and somewhat stronger assumptions on ϕ) is given by Hiriart-Urruty and Lemarechal [HiL93]. Their line of proof involves a complex argument, which is fundamentally different than ours.

Propositions 4.3 and 4.4 constitute variations of the corresponding minimax theorems of Rockafellar [Roc70] (Theorems 37.3 and 37.6), where instead of assuming that there is no nonzero common direction of recession of all the functions t_z, $z \in Z$ (and/or r_x, $x \in X$), a stronger condition is assumed, namely that there is no nonzero common direction of recession of all the functions t_z, $z \in \mathrm{ri}(Z)$ [and/or r_x, $x \in \mathrm{ri}(X)$, respectively], but some of the other assumptions of Props. 4.3 and 4.4 are not made.

The proof of the Minimax Theorem I (Prop. 4.2) can be easily modified to use the Min Common/Max Crossing Theorem II [cf. Prop. 3.2 and Eq. (4.7)]. What is needed is an assumption that $p(0)$ is finite and that 0 lies in the relative interior of the effective domain of p. We then obtain the following result, which asserts that the supremum in the minimax equality is attained [this follows from the corresponding attainment assertion of the Min Common/Max Crossing Theorem II and Eq. (4.7)].

Proposition 4.5: (Minimax Theorem II) Let X and Z be nonempty convex subsets of $\Re^n$ and $\Re^m$, respectively, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $\phi(\cdot, z) : X \mapsto \Re$ is convex, and for each $x \in X$, the function $-\phi(x, \cdot) : Z \mapsto \Re$ is closed and convex. Assume further that

$$-\infty < \inf_{x \in X} \sup_{z \in Z} \phi(x, z),$$

and that 0 lies in the relative interior of the effective domain of the function p of Eq. (4.4). Then, the minimax equality holds, i.e.,

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z),$$

and the supremum over Z in the left-hand side is finite and is attained. Furthermore, the set of $z \in Z$ attaining this supremum is compact if and only if 0 lies in the interior of the effective domain of p.

By using instead the Min Common/Max Crossing Theorem III (cf. Prop. 3.3), we obtain the following minimax theorem, which relates to problems where x and z are linearly coupled.

Proposition 4.6: (Minimax Theorem III) Let $\phi : X \times Z \mapsto \Re$ be a function of the form

$$\phi(x, z) = f(x) + z'Qx - h(z),$$

where X and Z are convex subsets of $\Re^n$ and $\Re^m$, respectively, Q is an $m \times n$ matrix, $f : X \mapsto \Re$ is a convex function, and $h : Z \mapsto \Re$ is a closed convex function. Consider the function

$$h^*(\zeta) = \sup_{z \in Z} \big\{ z'\zeta - h(z) \big\}, \qquad \zeta \in \Re^m$$

and assume the following:

(1) X is the intersection of a polyhedron P_1 and a convex set C_1, and f can be extended to a real-valued convex function over C_1 [i.e., there exists a convex function $\overline{f} : C_1 \mapsto \Re$ such that $f(x) = \overline{f}(x)$ for all $x \in X$].

(2) $\mathrm{dom}(h^*)$ is the intersection of a polyhedron P_2 and a convex set C_2, and h^* can be extended to a real-valued convex function over C_2 [i.e., there exists a convex function $\overline{h}^* : C_2 \mapsto \Re$ such that $h^*(\zeta) = \overline{h}^*(\zeta)$ for all $\zeta \in \mathrm{dom}(h^*)$].

(3) The sets $Q \cdot \big(X \cap \mathrm{ri}(C_1) \big)$ and $\mathrm{dom}(h^*) \cap \mathrm{ri}(C_2)$ have nonempty intersection.

Then the minimax equality holds, i.e.,

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z),$$

and the supremum over Z in the left-hand side above is attained by some vector $z \in Z$.

Proof: The function $p(u) = \inf_{x \in X} \sup_{z \in Z} \big\{ \phi(x, z) - u'z \big\}$ is given by

$$\begin{aligned} p(u) &= \inf_{x \in X} \sup_{z \in Z} \big\{ f(x) + z'Qx - h(z) - u'z \big\} \\ &= \inf_{x \in X} \left\{ f(x) + \sup_{z \in Z} \big\{ z'(Qx - u) - h(z) \big\} \right\} \\ &= \inf_{x \in X,\ (Qx - u) \in \mathrm{dom}(h^*)} \big\{ f(x) + h^*(Qx - u) \big\}. \end{aligned} \tag{4.13}$$

Because f is convex, and h is convex and closed, Lemmas 4.1 and 4.2, which connect the min common/max crossing framework with minimax problems apply. Thus, if we can use the Min Common/Max Crossing Theorem III in conjunction with the set

$$M = \mathrm{epi}(p),$$

the minimax equality will be proved, and the supremum over $z \in Z$ will be attained. We will thus show that the assumptions of the Min Common/Max Crossing Theorem III are satisfied under the assumptions of the present proposition.

Let P_1 and P_2 be represented in terms of linear inequalities as

$$P_1 = \{x \mid a_j'x - b_j \le 0,\ j = 1, \ldots, r_1\},$$

and

$$P_2 = \{\zeta \mid c_k'\zeta - d_k \le 0,\ k = 1, \ldots, r_2\}.$$

We have from the preceding expressions

$$p(u) = \inf_{\substack{x \in C_1,\ \zeta \in C_2 \\ A(x,\zeta) - b - Bu \le 0}} \big\{f(x) + h^*(\zeta)\big\},$$

where $A(x,\zeta) - b - Bu \le 0$ is a matrix representation of the inequalities

$$a_j'x - b_j \le 0,\ j = 1, \ldots, r_1, \qquad c_k'\zeta - d_k \le 0,\ k = 1, \ldots, r_2,$$

$$\zeta - Qx + u = 0,$$

for suitable matrices A and B, and vector b. Consider the convex set

$$V = \big\{(x,\zeta,w) \mid x \in C_1,\ \zeta \in C_2,\ w \in \Re,\ \overline{f}(x) + \overline{h}^*(\zeta) \le w\big\},$$

where $\overline{f}$ and $\overline{h}^*$ are real-valued convex functions, which are the extended versions of f and h^* over C_1 and C_2, respectively [cf. conditions (1) and (2)]. Thus, the epigraph of p has the form

$$\begin{aligned} \text{epi}(p) = \big\{(u,w) \mid\ & \text{there is a vector } (x,\zeta,w) \text{ such that} \\ & x \in C_1,\ \zeta \in C_2,\ A(x,\zeta) - b - Bu \le 0,\ \overline{f}(x) + \overline{h}^*(\zeta) \le w\big\}, \end{aligned}$$

so it can be written as

$$\begin{aligned} \text{epi}(p) = \big\{(u,w) \mid & \text{ there is a vector } (x,\zeta,w) \in V \\ & \text{such that } A(x,\zeta) - b - Bu \le 0\big\}. \end{aligned}$$

Therefore, epi(p) has the appropriate form for the application of the Min Common/Max Crossing Theorem III with the polyhedral cone P being the nonpositive orthant.

Finally, consider condition (3), which in conjunction with conditions (1) and (2), says that there exists a vector $\overline{x} \in P_1 \cap \text{ri}(C_1)$ such that $Q\overline{x} \in P_2 \cap \text{ri}(C_2)$. By letting $\overline{\zeta} = Q\overline{x}$, this holds if and only if there exist $\overline{x} \in \text{ri}(C_1)$ and $\overline{\zeta} \in \text{ri}(C_2)$ such that

$$\overline{x} \in P_1, \qquad \overline{\zeta} \in P_2, \qquad \overline{\zeta} = Q\overline{x},$$

or, in view of the definition of A and b,

$$A(\overline{x}, \overline{\zeta}) - b \le 0.$$

Thus, given conditions (1) and (2), condition (3) is equivalent to the existence of $\overline{x} \in \mathrm{ri}(C_1)$ and $\overline{\zeta} \in \mathrm{ri}(C_2)$ such that $A(\overline{x}, \overline{\zeta}) - b \le 0$.

On the other hand, since V is the epigraph of the function $\overline{f}(x) + \overline{h}^*(\zeta)$, whose domain is $C_1 \times C_2$, we have

$$\mathrm{ri}(V) = \big\{(x, \zeta, w) \mid x \in \mathrm{ri}(C_1),\, \zeta \in \mathrm{ri}(C_2),\, w \in \Re,\, \overline{f}(x) + \overline{h}^*(\zeta) < w\big\}.$$

Therefore, condition (3) is equivalent to the second assumption of the Min Common/Max Crossing Theorem III, i.e., that there exists a vector of the form $(\overline{x}, \overline{\zeta}, \overline{w})$ in $\mathrm{ri}(V)$ such that $A(\overline{x}, \overline{\zeta}) - b \le 0$. Thus, all the assumptions needed for application of the Min Common/Max Crossing Theorem III are satisfied, and the proof is complete. **Q.E.D.**

We note some special cases of the above proposition. Its assumptions are satisfied if one of the following two conditions holds:

(a) There exists a vector $\overline{x} \in \mathrm{ri}\big(X\big)$ such that $Q\overline{x} \in \mathrm{dom}(h^*)$, and the function h^* is polyhedral. [Take $C_1 = X$, $C_2 = \Re^m$, and $P_1 = \Re^n$, $P_2 = \mathrm{dom}(h^*)$ in Prop. 4.6.]

(b) There exists a vector $\overline{x} \in \mathrm{ri}\big(X\big)$ such that $Q\overline{x} \in \mathrm{ri}\big(\mathrm{dom}(h^*)\big)$. [Take $C_1 = X$, $C_2 = \mathrm{dom}(h^*)$, and $P_1 = \Re^n$, $P_2 = \Re^m$ in Prop. 4.6.]

Note that a minimax problem involving a function ϕ of the form

$$\phi(x, z) = f(x) + z'Qx - h(z)$$

is closely related to Fenchel duality (see e.g., Rockafellar [Roc70], Bertsekas [Ber99]). Indeed, from Eq. (4.13), we have

$$p(0) = \inf_{x \in X,\, Qx \in \mathrm{dom}(h^*)} \big\{f(x) + h^*(Qx)\big\},$$

which is the standard problem arising within the Fenchel framework. In fact the conditions (1) and (2) of Prop. 4.6 guaranteeing that the minimax equality holds are typical of the conditions used to guarantee that there is no duality gap in the Fenchel duality context.

5. LAGRANGIAN DUALITY

We now consider the constrained minimization problem

$$\begin{aligned}&\text{minimize} \quad f(x)\\&\text{subject to} \quad x \in X, \quad h(x) = 0, \quad g(x) \leq 0,\end{aligned} \tag{P}$$

where X is a subset of $\Re^n$, $f : \Re^n \mapsto \Re$, $h_i : \Re^n \mapsto \Re$, $i = 1, \ldots, m$, and $g_j : \Re^n \mapsto \Re$, $j = 1, \ldots, r$, are functions, and we use the notation

$$h(x) = \big(h_1(x), \ldots, h_m(x)\big), \qquad g(x) = \big(g_1(x), \ldots, g_r(x)\big).$$

We refer to this as the *primal problem* and we denote by f^* its optimal value, i.e.,

$$f^* = \inf_{\substack{x \in X \\ h(x)=0,\ g(x) \leq 0}} f(x).$$

Let $L : \Re^{n+m+r} \mapsto \Re$ be the Lagrangian function

$$L(x, \lambda, \mu) = f(x) + \lambda' h(x) + \mu' g(x).$$

Following Rockafellar [Roc70], we use the following definition of a multiplier vector.

Definition 5.1: A vector $(\lambda^*, \mu^*) \in \Re^{m+r}$ is said to be a *geometric multiplier* for the primal problem (P) if $\mu^* \geq 0$ and

$$f^* = \inf_{x \in X} L(x, \lambda^*, \mu^*).$$

We consider the *dual function* q defined for $(\lambda, \mu) \in \Re^{m+r}$ by

$$q(\lambda, \mu) = \inf_{x \in X} L(x, \lambda, \mu).$$

The *dual problem* is

$$\begin{aligned}&\text{maximize} \quad q(\lambda, \mu)\\&\text{subject to} \quad \lambda \in \Re^m, \quad \mu \geq 0.\end{aligned} \tag{D}$$

It is well-known that the effective domain of q, i.e., the set

$$\big\{(\lambda, \mu) \mid q(\lambda, \mu) > -\infty\big\},$$

is convex and that q is concave over its effective domain. It is also well known that $q^* \leq f^*$. This is the *Weak Duality Theorem*. If $q^* = f^*$ we say that *there is no duality gap* and if $q^* < f^*$ we say that *there is a duality gap*. Existence of a geometric multiplier guarantees that there is no duality gap, as asserted in the following proposition (see e.g., Bertsekas [Ber99], Prop. 5.1.4).

Proposition 5.1:

(a) If there is no duality gap, the set of geometric multipliers is equal to the set of optimal solutions of the dual problem.

(b) If there is a duality gap, the set of geometric multipliers is empty.

Duality theory is intimately connected with the minimax theory of the preceding section. In particular, suppose we identify z with the multiplier vector (λ, μ) and we choose ϕ to be the Lagrangian function

$$L(x, \lambda, \mu) = f(x) + \lambda' h(x) + \mu' g(x).$$

Then, we have

$$\sup_{\lambda \in \Re^m,\ \mu \geq 0} L(x, \lambda, \mu) = \begin{cases} f(x) & \text{if } h(x) = 0 \text{ and } g(x) \leq 0, \\ \infty & \text{otherwise,} \end{cases}$$

so the primal problem (P) is equivalent to the problem

$$\begin{aligned} &\text{minimize} \quad \sup_{\lambda \in \Re^m,\ \mu \geq 0} L(x, \lambda, \mu) \\ &\text{subject to} \ \ x \in X. \end{aligned}$$

Furthermore, by the definition of the dual function, we have

$$q(\lambda, \mu) = \inf_{x \in X} L(x, \lambda, \mu),$$

so the dual problem (D) is equivalent to the problem

$$\begin{aligned} &\text{maximize} \quad \inf_{x \in X} L(x, \lambda, \mu) \\ &\text{subject to} \ \ \lambda \in \Re^m,\ \mu \geq 0. \end{aligned}$$

Thus, with the preceding identifications, the presence of no duality gap between the primal and dual problems is equivalent to the minimax equality when ϕ is the Lagrangian function as per the preceding identifications. Moreover, we have the following well-known proposition, which shows that a saddle point of the Lagrangian function is an optimal solution-geometric multiplier pair of problem (P) (see e.g., Bertsekas [Ber99], Ch. 5).

Proposition 5.2: (Lagrangian Saddle Point Theorem) The vector (x^*, λ^*, μ^*) is an optimal solution-geometric multiplier pair of problem (P) if and only if (x^*, λ^*, μ^*) is a saddle point of the Lagrangian in the sense that $x^* \in X$, $\mu^* \geq 0$, and

$$L(x^*, \lambda, \mu) \leq L(x^*, \lambda^*, \mu^*) \leq L(x, \lambda^*, \mu^*), \qquad \forall\ x \in X,\ \lambda \in \Re^m,\ \mu \geq 0.$$

The classical von Neumann result of Prop. 4.4, which guarantees the minimax equality (4.3), as well as the attainment of the inf and sup, assuming convexity/concavity assumptions on ϕ, and compactness assumptions on X and Z, is not adequate for the development of duality theory. The reason is that compactness of Z and, to some extent, compactness of X are restrictive assumptions [in particular, Z corresponds to the constraint set $\big\{(\lambda,\mu) \mid \lambda \in \Re^m,\, \mu \geq 0,\, q(\lambda,\mu) > -\infty\big\}$ of the dual problem, which is not necessarily compact]. However, the other propositions of Section 4 are more relevant to duality theory and give conditions guaranteeing that there is no duality gap, although they need not always guarantee the attainment of the infimum and the supremum. Similar duality and multiplier existence results have been known for a long time, but our objective here is to show how they can be proved as special cases of our min common/max crossing framework.

Conditions for no Duality Gap

We first derive conditions under which there is no duality gap. As mentioned earlier, we can analyze this question by applying the minimax theorems of Section 4, with the function ϕ equal to the Lagrangian function. For simplicity, we consider problem (P) in the absence of the equality constraint $h(x) = 0$. We introduce the family of problems

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \quad g_j(x) \leq u_j, \quad j = 1,\ldots,r, \end{aligned} \tag{5.1}$$

parameterized by $u = (u_1,\ldots,u_r)$, and we let $p(u)$ denote the corresponding optimal value, i.e.,

$$p(u) = \inf_{\substack{x\in X \\ g_j(x)\leq u_j,\, j=1,\ldots,r}} f(x). \tag{5.2}$$

The function p is known as the *primal function* of the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \quad g_j(x) \leq 0, \quad j = 1,\ldots,r. \end{aligned} \tag{5.3}$$

We now note that

$$\sup_{\mu\geq 0}\big\{L(x,\mu) - u'\mu\big\} = \begin{cases} f(x) & \text{if } g_j(x) \leq u_j,\ j = 1,\ldots,r, \\ \infty & \text{otherwise,} \end{cases}$$

so that the primal function of Eq. (5.2) coincides with the function p of Eq. (4.4),

$$p(u) = \inf_{x\in X}\sup_{\mu\geq 0}\big\{L(x,\mu) - u'\mu\big\}, \qquad u \in \Re^m,$$

where Z is the nonnegative orthant, i.e., $Z = \{\mu \mid \mu \geq 0\}$, and $\phi(x, \mu)$ is the Lagrangian function $L(x, \mu)$. Thus, the Minimax Theorem I of Section 4 applies, and we obtain the following well-known proposition.

Proposition 5.3: Let X be convex, let f and the g_j be convex over X, and assume that $-\infty < p(0) < \infty$. Then there is no duality gap if and only if p is lower semicontinuous at 0.

To guarantee the lower semicontinuity of p at 0, we may use Prop. 4.3. Thus, if we assume that the problem is feasible, that X is convex and compact, and that the functions f and g_j are convex over $\Re^n$ (and are therefore continuous), then using Props. 4.2 and 4.3, we have that there is no duality gap and the primal problem (P) has at least one optimal solution.

Conditions for Existence of a Geometric Multiplier

It is possible to analyze the question of existence of a geometric multiplier in the context of the minimax theory of the preceding section. In particular, a geometric multiplier exists if and only if there is no duality gap and the supremum of $\inf_{x \in X} L(x, \mu)$ [which is $q(\mu)$] over $\mu \geq 0$ is attained. Thus, the Minimax Theorems II and III of the preceding section can be applied.

We will follow an alternative and more direct approach. In particular, we will use the min common/max crossing duality framework to prove the following nonlinear version of Farkas' Lemma (if $C = \Re^n$ and the functions f and g_j are linear in the statement below, we obtain the classical Farkas' Lemma). Versions of this lemma are available in the literature, dating to Fan, Glicksberg, and Hoffman [FGH57], and including Berge and Ghouila-Houri [BeG62], and Rockafellar [Roc70]. From this lemma, the existence of a geometric multiplier vector under various constraint qualifications follows easily.

Note that the conditions (a) of the following lemma correspond to the relative interior conditions of the Min Common/Max Crossing Theorem II and the Minimax Theorem II. The conditions (b) of the lemma correspond to the alternative polyhedral/relative interior conditions of the Min Common/Max Crossing Theorem III and the Minimax Theorem III.

Proposition 5.4: (Nonlinear Farkas' Lemma) Let C be a nonempty convex subset of $\Re^n$, and let $f : C \mapsto \Re$ and $g_j : C \mapsto \Re$, $j = 1, \ldots, r$, be convex functions. Consider the set F given by

$$F = \big\{x \in C \mid g(x) \leq 0\big\},$$

where $g(x) = \big(g_1(x), \ldots, g_r(x)\big)$, and assume that

$$f(x) \geq 0, \qquad \forall\, x \in F. \tag{5.4}$$

Consider the subset Q^* of $\Re^r$ given by

$$Q^* = \big\{\mu \mid \mu \geq 0,\ f(x) + \mu' g(x) \geq 0,\ \forall\, x \in C\big\}.$$

Then:

(a) Q^* is nonempty and compact if and only if there exists a vector $\overline{x} \in C$ such that

$$g_j(\overline{x}) < 0, \qquad \forall\, j = 1, \ldots, r.$$

(b) Q^* is nonempty if the functions g_j, $j = 1, \ldots, r$, are affine, and F contains a relative interior point of C.

Proof: (a) Assume that there exists a vector $\overline{x} \in C$ such that $g(\overline{x}) < 0$. We will apply the Min Common/Max Crossing Theorem II (Prop. 3.2) to the subset of $\Re^{r+1}$ given by

$$M = \big\{(u, w) \mid \text{there exists } x \in C \text{ such that } g(x) \leq u,\ f(x) \leq w\big\}$$

(cf. Fig. 5.1). To this end, we verify that the assumptions of the theorem are satisfied for the above choice of M.

In particular, we will show that:

(i) The optimal value w^* of the corresponding min common problem,

$$w^* = \inf\big\{w \mid (0, w) \in M\big\},$$

is such that $-\infty < w^*$.

(ii) The set

$$\overline{M} = \big\{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in M\big\},$$

is convex. (Note here that $\overline{M} = M$.)

Figure 5.1. Illustration of the sets

$$M = \overline{M} = \Big\{(u,w) \mid \text{there exists } x \in C \text{ such that } g(x) \le u,\ f(x) \le w\Big\}$$

and

$$D = \Big\{u \mid g(x) \le u \text{ for some } x \in C\Big\} = \Big\{u \mid (u,w) \in \overline{M} \text{ for some } w \in \Re\Big\}$$

that are used in the proof of Prop. 5.4. Note that if there exists $\overline{x} \in C$ such that $g_j(\overline{x}) < 0$ for all $j = 1,\ldots,r$, then 0 is an interior point of D.

(iii) The set

$$D = \Big\{u \mid \text{there exists } w \in \Re \text{ such that } (u,w) \in \overline{M}\Big\}$$

contains the origin in its interior.

To show (i), note that since $f(x) \ge 0$ for all $x \in F$, we have $w \ge 0$ for all $(0,w) \in M$, so that $w^* \ge 0$.

To show (iii), note that the set D can also be written as

$$D = \Big\{u \mid \ \text{there exists } x \in C \text{ such that } g(x) \le u\Big\}.$$

If $g(\overline{x}) < 0$ for some $\overline{x} \in C$, then since D contains the set $g(\overline{x}) + \{u \mid u \ge 0\}$, we have $0 \in \text{int}(D)$.

There remains to show (ii), i.e., that the set $\overline{M}$ is convex. Since $\overline{M} = M$, we will prove that M is convex. To this end, we consider vectors $(u,w) \in M$ and $(\tilde{u},\tilde{w}) \in M$, and we show that their convex combinations lie in M. By the definition of M, for some $x \in C$ and $\tilde{x} \in C$, we have

$$f(x) \le w, \qquad g_j(x) \le u_j, \quad \forall\, j = 1,\ldots,r,$$

$$f(\tilde{x}) \le \tilde{w}, \qquad g_j(\tilde{x}) \le \tilde{u}_j, \quad \forall\, j = 1,\ldots,r.$$

For any $\alpha \in [0,1]$, we multiply these relations with α and $1-\alpha$, respectively, and add them. By using the convexity of f and g_j for all j, we obtain

$$f\big(\alpha x + (1-\alpha)\tilde{x}\big) \le \alpha f(x) + (1-\alpha) f(\tilde{x}) \le \alpha w + (1-\alpha)\tilde{w},$$

$$g_j\big(\alpha x + (1-\alpha)\tilde{x}\big) \le \alpha g_j(x) + (1-\alpha) g_j(\tilde{x}) \le \alpha u_j + (1-\alpha)\tilde{u}_j, \quad \forall\, j = 1,\ldots,r.$$

By convexity of C, we have $\alpha x+(1-\alpha)\tilde{x} \in C$ for all $\alpha \in [0,1]$, so the preceding inequalities imply that the convex combination $\alpha(u,w)+(1-\alpha)(\tilde{u},\tilde{w})$ belongs to M, showing that M is convex.

Thus all the assumptions of the Min Common/Max Crossing Theorem II hold, and by the conclusions of the theorem, we have $w^* = \sup_\mu q(\mu)$, where

$$q(\mu) = \inf_{(u,w)\in M}\{w+\mu' u\}.$$

Furthermore, the optimal solution set $\tilde{Q} = \big\{\mu \mid q(\mu) \geq w^*\big\}$ is nonempty and compact. Using the definition of M, it can be seen that

$$q(\mu) = \begin{cases} \inf_{x\in C}\big\{f(x)+\mu' g(x)\big\} & \text{if } \mu \geq 0, \\ -\infty & \text{otherwise.} \end{cases}$$

From the definition of Q^*, we have

$$Q^* = \big\{\mu \mid \mu \geq 0,\ f(x)+\mu' g(x) \geq 0,\ \forall\ x \in C\big\} = \big\{\mu \mid q(\mu) \geq 0\big\},$$

so Q^* and $\tilde{Q}$ are level sets of the proper convex function $-q$, which is closed. Therefore, since $\tilde{Q}$ is nonempty and compact, Q^* is compact. Furthermore, Q^* is nonempty since $Q^* \supset \tilde{Q}$.

Conversely, assuming that Q^* is nonempty and compact, we will show that there exists a vector $\overline{x} \in C$ such that $g(\overline{x}) < 0$. Indeed, if this were not so, then 0 would not be an interior point of the set D. Since D is convex, there exists a hyperplane that passes through 0 and contains D in its positive halfspace, i.e., there is a nonzero vector $\nu \in \Re^r$ such that $\nu' u \geq 0$ for all $u \in D$. From the definition of D, it follows that $\nu \geq 0$. Since $g(x) \in D$ for all $x \in C$, we obtain

$$\nu' g(x) \geq 0, \qquad \forall\ x \in C.$$

Thus, for any $\mu \in Q^*$, we have

$$f(x) + (\mu+\gamma\nu)' g(x) \geq 0, \qquad \forall\ x \in C,\ \forall\ \gamma \geq 0.$$

Since we also have $\nu \geq 0$, it follows that $(\mu+\gamma\nu) \in Q^*$ for all $\gamma \geq 0$, which contradicts the boundedness of Q^*.

(b) We apply the Min Common/Max Crossing Theorem III (cf. Prop. 3.3), with the polyhedral cone P in that theorem being equal to the nonpositive orthant, i.e., $P = \{u \mid u \leq 0\}$, and with the matrix B being the identity matrix. The assumptions of the theorem for this choice of P and B are:

(i) The subset M of $\Re^{r+1}$ is defined in terms of a convex set V of $\Re^{n+1}$, an $r \times n$ matrix A, and a vector b in $\Re^r$ as follows:

$$M = \big\{(u,w) \mid \text{there is a vector } (x,w) \in V \text{ such that } Ax - b \leq u\big\}.$$

(ii) There exists a vector $(\overline{x}, \overline{w})$ in the relative interior of V such that $A\overline{x} - b \leq 0$.

Let the affine functions g_j have the form

$$g_j(x) = a_j' x - b_j,$$

where a_j are some vectors in $\Re^n$ and b_j are some scalars. We choose the matrix A to have as rows the vectors a_j', and the vector b to be equal to $(b_1, \ldots, b_r)'$. We also choose the convex set V to be

$$V = \big\{(x, w) \mid x \in C,\ f(x) \leq w\big\}.$$

To prove that (ii) holds, note that by our assumptions, there exists a vector $\overline{x}$ in $F \cap \mathrm{ri}(C)$, i.e., $\overline{x} \in \mathrm{ri}(C)$ and $A\overline{x} - b \leq 0$. Then, the vector $(\overline{x}, \overline{w})$ with $\overline{w} > f(\overline{x})$ belongs to $\mathrm{ri}(V)$. Hence, all the assumptions of the Min Common/Max Crossing Theorem III are satisfied, and by using this theorem, similar to the proof of part (a), we have

$$\inf_{x \in C} \big\{ f(x) + \mu^{*\prime} g(x) \big\} = q(\mu^*) = q^* = w^*$$

for some $\mu^* \geq 0$. Since $w^* \geq 0$, it follows that $f(x) + \mu^{*\prime} g(x) \geq 0$ for all $x \in C$. **Q.E.D.**

The existence of a vector $\overline{x} \in C$ such that $g(\overline{x}) < 0$ [part (a) of the lemma] is known as the *Slater condition*, and will be reencountered in the next section. We now use the preceding Nonlinear Farkas' Lemma to assert the existence of geometric multipliers under some specific assumptions.

Convex Constraints

We now consider the nonlinearly constrained problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \ \ x \in X, \quad g_j(x) \leq 0,\ j = 1, \ldots, r, \end{aligned} \tag{5.5}$$

under the following assumption.

> **Assumption 5.1: (Slater Condition)** The optimal value f^* of problem (5.5) is finite, X is convex, and the functions $f : \Re^n \mapsto \Re$ and $g_j : \Re^n \mapsto \Re$ are convex over X. Furthermore, there exists a vector $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all $j = 1, \ldots, r$.

We have the following proposition.

Proposition 5.5: (Strong Duality Theorem - Nonlinear Constraints) Let Assumption 5.1 hold for problem (5.5). Then, there is no duality gap, and the set of geometric multipliers is nonempty and compact.

Proof: The result follows by applying the Nonlinear Farkas' Lemma [Prop. 5.4, condition (a)] with $C = X$ and assuming that $f^* = 0$ [otherwise, we replace $f(x)$ by $f(x) - f^*$]. In particular, the set $\{u \mid g(\overline{x}) \leq u\}$ is a subset of the set

$$D = \{u \mid \text{ there exists } x \in X \text{ with } g(x) \leq u\},$$

and contains 0 in its interior. Hence, D also contains 0 in its interior, and condition (a) of the Nonlinear Farkas' Lemma is satisfied. By weak duality, it can be seen that the set of nonempty and compact vectors μ^*, whose existence is asserted by the lemma, is the set of geometric multipliers. **Q.E.D.**

Linear Constraints

We first consider the linearly constrained problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \;\; x \in X, \;\; e_i'x - d_i = 0, \; i = 1, \dots, m, \;\; a_j'x - b_j \leq 0, \; j = 1, \dots, r, \end{aligned} \tag{5.6}$$

where $f : \Re^n \mapsto \Re$ is a convex function and X is the intersection of a polyhedral set with some other convex set. An important special case is when X itself is a polyhedral set.

Assumption 5.2: (Linear Constraints) The optimal value f^* of problem (5.6) is finite, and the following hold:

(1) The set X is the intersection of a polyhedral set P and a convex set C.

(2) The cost function $f : \Re^n \mapsto \Re$ is convex over C.

(3) There exists a feasible solution of the problem that lies in the relative interior of C.

Proposition 5.6: (Strong Duality Theorem - Linear Constraints) Let Assumption 5.2 hold for problem (5.6). Then, there is no duality gap and there exists at least one geometric multiplier.

Proof: The proof is based on the Nonlinear Farkas' Lemma [Prop. 5.4, condition (b)]. Without loss of generality, we can assume that there are no equality constraints, so we are dealing with the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \;\; x \in X, \quad a_j' x - b_j \leq 0, \; j = 1, \ldots, r, \end{aligned}$$

(each equality constraint can be converted into two inequality constraints). Also without loss of generality, we can assume that $f^* = 0$ [otherwise, we replace $f(x)$ by $f(x) - f^*$].

We have $X = P \cap C$, where P is a polyhedral set that can be expressed in terms of linear inequalities as

$$P = \{x \mid a_j' x - b_j \leq 0, \; j = r+1, \ldots, p\},$$

where p is an integer with $p > r$. By applying the Nonlinear Farkas' Lemma [Prop. 5.4, condition (b)] with F being the set

$$\{x \in C \mid a_j' x - b_j \leq 0, \; j = 1, \ldots, p\},$$

we see that there exist nonnegative $\mu_1^*, \ldots, \mu_p^*$ such that

$$f(x) + \sum_{j=1}^{p} \mu_j^* (a_j' x - b_j) \geq 0, \qquad \forall \; x \in C.$$

Since for $x \in P$, we have $\mu_j^*(a_j' x - b_j) \leq 0$ for all $j = r+1, \ldots, p$, the above equation yields

$$f(x) + \sum_{j=1}^{r} \mu_j^* (a_j' x - b_j) \geq 0, \qquad \forall \; x \in P \cap C = X,$$

implying that

$$\inf_{x \in X} L(x, \mu^*) = q(\mu^*) \geq 0 = f^*.$$

By weak duality, we have $q(\mu^*) \leq q^* \leq f^*$, so it follows that $q^* = f^*$ and that μ^* is a geometric multiplier. **Q.E.D.**

Convex and Linear Constraints

We finally consider a generalization of problems (5.6) and (5.5), where there are linear equality and inequality constraints, as well as convex inequality constraints:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \;\; x \in X, \quad g_j(x) \leq 0, \;\; j = 1, \ldots, \overline{r}, \\ &\qquad\qquad e_i' x - d_i = 0, \; i = 1, \ldots, m, \;\; a_j' x - b_j \leq 0, \; j = \overline{r}+1, \ldots, r. \end{aligned} \tag{5.7}$$

To cover this case, we suitably modify Assumptions 5.2 and 5.1.

Assumption 5.3: (For Linear and Nonlinear Constraints) The optimal value f^* of problem (5.7) is finite, and the following hold:

(1) X is the intersection of a polyhedral set P and a convex set C.

(2) The functions $f : \Re^n \mapsto \Re$ and $g_j : \Re^n \mapsto \Re$ are convex over C.

(3) There exists a feasible vector $\bar{x}$ such that $g_j(\bar{x}) < 0$ for all $j = 1, \ldots, \overline{r}$.

(4) There exists a vector that satisfies the linear constraints [but not necessarily the constraints $g_j(x) \leq 0$, $j = 1, \ldots, \overline{r}$] and belongs to X and to the relative interior of C.

Note that part (4) of the preceding assumption is slightly weaker than the corresponding assumption of Th. 28.2 of Rockafellar [Roc70], which requires that the relative interior point of C must satisfy the nonlinear as well as the linear constraints.

Proposition 5.7: (Strong Duality Theorem - Linear and Nonlinear Constraints) Let Assumption 5.3 hold for problem (5.7). Then, there is no duality gap and there exists at least one geometric multiplier.

Proof: Using Prop. 5.5, we argue that there exist $\mu_j^* \geq 0$, $j = 1, \ldots, \overline{r}$, such that

$$f^* = \inf_{\substack{x \in X,\, a_j' x - b_j \leq 0,\, j = \overline{r}+1, \ldots, r \\ e_i' x - d_i = 0,\, i = 1, \ldots, m}} \left\{ f(x) + \sum_{j=1}^{\overline{r}} \mu_j^* g_j(x) \right\}.$$

Then, we apply Prop. 5.6 to the minimization problem in the right-hand side of the above equation to show that there exist λ_i^*, $i = 1, \ldots, m$, and $\mu_j^* \geq 0$, $j = \overline{r}+1, \ldots, r$, such that

$$f^* = \inf_{x \in X} \left\{ f(x) + \sum_{i=1}^{m} \lambda_i^* (e_i' x - d_i) + \sum_{j=\overline{r}+1}^{r} \mu_j^* (a_j' x - b_j) + \sum_{j=1}^{\overline{r}} \mu_j^* g_j(x) \right\}.$$

Q.E.D.

Let us finally note that one can extend some of the analysis of this section to the more general constrained optimization problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \quad h(x) = 0, \quad g(x) \in C, \end{aligned}$$

where X is a subset of $\Re^n$, $f : \Re^n \mapsto \Re$, $h_i : \Re^n \mapsto \Re$, $i = 1, \ldots, m$, and $g_j : \Re^n \mapsto \Re$, $j = 1, \ldots, r$, are functions, and C is a polyhedral cone in $\Re^r$. When C is the nonpositive orthant, this problem reduces to problem (P). We can modify the definitions of geometric multiplier and the dual problem, and adapt the analysis of this section to cover the general optimization problem above. In particular, we can use the Min Common/Max Crossing Theorem III to analyze the case where the functions h and g are linear.

6. CONCLUSIONS

The simple duality framework of this paper provides an intuitive intermediate step between the fundamental dual characterization of closed convex sets, and the constraint qualifications and related assumptions needed to prove minimax and strong duality theorems. We have shown that the major minimax and Lagrangian duality results (under convexity assumptions) are special cases of three theorems related to our min common and max crossing problems. Because of its geometric and fundamental character, our framework may prove useful in contexts beyond the ones discussed in this paper.

7. REFERENCES

[BNO03] Bertsekas, D. P., with Nedic, A., and Ozdaglar, A. E., 2003. Convex Analysis and Optimization, Athena Scientific, Belmont, MA.

[Ber99] Bertsekas, D. P., 1999. Nonlinear Programming, 2nd Edition, Athena Scientific, Belmont, MA.

[BeG62] Berge, C., and Ghouila-Houri, A., 1962. Programmes, Jeux et Reseau de Transport, Dunod, Paris.

[BoL00] Borwein, J. M., and Lewis, A. S., 2000. Convex Analysis and Nonlinear Optimization, Springer-Verlag, N. Y.

[BoS00] Bonnans, J. F., and Shapiro, A., 2000. Perturbation Analysis of Optimization Problems, Springer-Verlag, N. Y.

[EkT76] Ekeland, I., and Temam, R., 1976. Convex Analysis and Variational Problems, North-Holland Publ., Amsterdam.

[FGH57] Fan, K., Glicksberg, I., and Hoffman, A. J., 1957. "Systems of Inequalities Involving Convex Functions," Proc. Amer. Math. Soc., Vol. 8, pp. 617–622.

[HiL93] Hiriart-Urruty, J.-B., and Lemarechal, C., 1993. Convex Analysis and Minimization Algorithms, Vols. I and II, Springer-Verlag, Berlin and N. Y.

[Kak41] Kakutani, S., 1941. "A Generalization of Brouwer's Fixed Point Theorem," Duke Mathematical Journal, Vol. 8, pp. 457-459.

[Neu28] Neumann, J. von, 1928. "Zur Theorie der Gesellschaftspiele," Math. Ann., Vol. 100, pp. 295–320.

[RoW98] Rockafellar, R. T., and Wets, R. J.-B., 1998. Variational Analysis, Springer-Verlag, Berlin.

[Roc70] Rockafellar, R. T., 1970. Convex Analysis, Princeton Univ. Press, Princeton, N. J.

[Roc84] Rockafellar, R. T., 1984. Network Flows and Monotropic Optimization, Wiley, N. Y.; republished by Athena Scientific, Belmont, MA, 1998.

[StW70] Stoer, J., and Witzgall, C., 1970. Convexity and Optimization in Finite Dimensions, Springer-Verlag, Berlin.