August 2008 **LIDS-P-2796**

Revised Jan. 2009

Min Common/Max Crossing Duality: A Geometric View of Conjugacy in Convex Optimization[1]

Dimitri P. Bertsekas[2]

Abstract

We provide a unifying framework for the visualization and analysis of duality, and other issues in convex optimization. It is based on two simple optimization problems that are dual to each other: the *min common point* problem and the *max crossing point* problem. Within the insightful geometry of these problems, several of the core issues in convex analysis become apparent and can be analyzed in a unified way. These issues are related to conditions for strong duality in constrained optimization and zero sum games, existence of dual optimal solutions and saddle points, existence of subgradients, and theorems of the alternative.

The generality and power of our framework is due to its close connection with the Legendre/Fenchel conjugacy framework. However, the two frameworks offer complementary starting points for analysis and provide alternative views of the geometric foundation of duality: conjugacy emphasizes functional/algebraic descriptions, while min common/max crossing emphasizes set/epigraph descriptions. The min common/max crossing framework is simpler, and seems better suited for visualizing and investigating questions of strong duality and existence of dual optimal solutions. The conjugacy framework, with its emphasis on functional descriptions, is more suitable when mathematical operations on convex functions are involved, and the calculus of conjugate functions can be brought to bear for analysis or computation.

[1] Supported by NSF Grant ECCS-0801549. The MC/MC framework was initially developed in joint research with A. Nedic, and A. Ozdaglar. This research is described in the book by Bertsekas, Nedic, and Ozdaglar [BNO03]. The present account is an improved and more comprehensive development. In particular, it contains some more streamlined proofs and some new results, particularly in connection with minimax problems and separable problems. While this paper is self-contained, a fuller presentation may be found in the author's convex optimization theory book [Ber09].

[2] Dimitri Bertsekas is with the Dept. of Electr. Engineering and Comp. Science, M.I.T., Cambridge, Mass., 02139.

Contents

1. INTRODUCTION

Duality in optimization is often viewed as a manifestation of a fundamental dual/conjugate description of a closed convex set as the intersection of all closed halfspaces containing the set. When specialized to the epigraph of a function $f : \Re^n \mapsto [-\infty, \infty]$, this description leads to the formalism of the conjugate convex function of f, which is defined by

$$h(y) = \sup_{x \in \Re^n} \big\{ x'y - f(x) \big\}, \tag{1.1}$$

and permeates much of convex optimization theory.

In this paper, we focus on a framework referred to as *min common/max crossing* (MC/MC for short). It is related to the conjugacy framework, but does not involve an algebraic definition such as Eq. (1.1), and it is structured to emphasize optimization duality aspects. For this reason it is simpler, and seems better suited for geometric visualization and analysis in many important convex optimization contexts. Our framework aims to capture the most essential optimization-related features of the preceding conjugate description of closed convex sets in two simple geometrical problems, defined by a nonempty subset M of $\Re^{n+1}$.

(a) *Min Common Point Problem*: Consider all vectors that are common to 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, i.e., the closed halfspace whose recession cone contains the vertical halfline $\big\{(0, w) \mid w \geq 0\big\}$ (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, and that under favorable circumstances the two optimal values are equal. Our purpose in this paper is to formalize the analysis of the two problems, to provide conditions that guarantee equality of their optimal values and the existence of their optimal solutions, and to show that they can be used to develop much of the core theory of convex analysis and optimization in a unified way.

Mathematically, the min common problem is

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

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

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

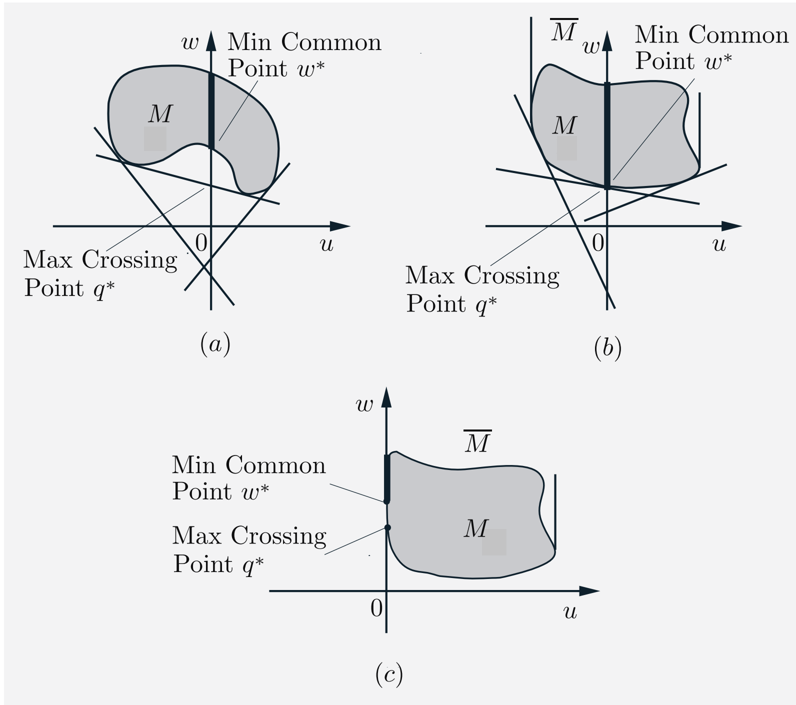

Figure 1.1. Illustration of the optimal values of the min common and max crossing problems. In (a), the two optimal values are not equal. In (b), when M is "extended upwards" along the $(n+1)$st axis, it yields the set

$$\overline{M} = M + \big\{(0, w) \mid w \geq 0\big\} = \big\{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in M\big\},$$

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 $q^* < w^*$.

To describe mathematically the max crossing problem, we recall that a nonvertical hyperplane in $\Re^{n+1}$ is specified by its normal vector $(\mu, 1) \in \Re^{n+1}$, and a scalar ξ as

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

Such a hyperplane crosses the $(n+1)$st axis at $(0, \xi)$. For M to be contained in the closed halfspace that corresponds to $H_{\mu,\xi}$ and contains the vertical halfline $\big\{(0, w) \mid w \geq 0\big\}$ in its recession cone, we must have

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

or equivalently

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

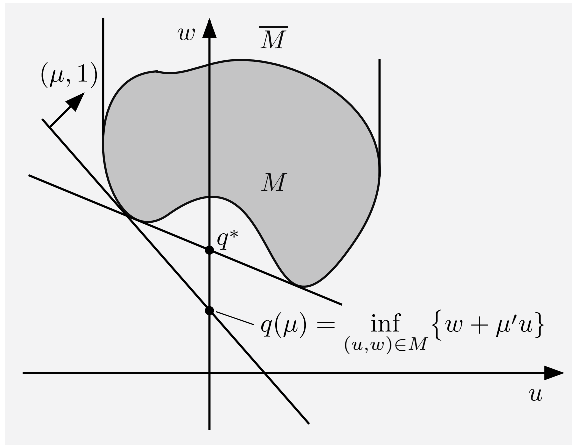

Figure 1.2. Mathematical specification of the max crossing problem. When considering crossing points by nonvertical hyperplanes, it is sufficient to restrict attention to hyperplanes with normals of the form $(\mu, 1)$, $\mu \in \Re^n$. For each $\mu \in \Re^n$, we consider $q(\mu)$, the highest crossing 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 $q(\mu)$.

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

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\}; \tag{1.2}$$

(see Fig. 1.2). Thus the max crossing problem is to maximize over all $\mu \in \Re^n$ the maximum crossing level corresponding to μ, i.e.,

$$\begin{aligned} &\text{maximize} \quad q(\mu) \\ &\text{subject to} \quad \mu \in \Re^n. \end{aligned} \tag{1.3}$$

We denote by q^* the corresponding optimal value,

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

and we refer to $q(\mu)$ as the *crossing* or *dual* function.

Note that both w^* and q^* remain unaffected if M is replaced by its "upwards extension"

$$\overline{M} = M + \big\{(0, w) \mid w \geq 0\big\} = \big\{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in M\big\} \tag{1.4}$$

(cf. Fig. 1.1). It is often more convenient to work with $\overline{M}$ because in many cases of interest $\overline{M}$ is convex while M is not. However, on occasion M has some interesting properties (such as compactness) that are masked when passing to $\overline{M}$, in which case it may be preferable to work with M. We will show in the next section that generically we have $q^* \leq w^*$; this is referred to as *weak duality*. When $q^* = w^*$, we say that *strong duality holds* or that *there is no duality gap*.

Note also that we do not exclude the possibility that either w^* or q^* (or both) are infinite. In particular, we have $w^* = \infty$ if the min common problem has no feasible solution $[M \cap \{(0,w) \mid w \in \Re\} = \emptyset]$. Similarly, we have $q^* = -\infty$ if the max crossing problem has no feasible solution, which occurs in particular if $\overline{M}$ contains a vertical line, i.e., a set of the form $\{(x,w) \mid w \in \Re\}$ for some $x \in \Re^n$.

The paper is organized as follows. In Section 2, we develop some general results, and we discuss the connection with the conjugacy framework. We also consider some special cases, including constrained optimization and minimax, and derive the corresponding forms of the crossing function. In Section 3, we establish conditions for strong duality, and also conditions under which the max crossing problem has a nonempty and/or compact optimal solution set. In Section 4, we apply the results of Section 3, and we develop as special cases the core theory of minimax theory and constrained optimization duality. We also establish connections with subdifferential theory, and with theorems of the alternative, such as the Gordan and Motzkin theorems. Finally, in Section 5 we focus on nonconvex cases where there is a nonzero duality gap $w^* - q^*$ and discuss techniques that may be used to estimate its size, with a focus on separable-type problems. A considerable part of the analysis is new and has not been presented in [BNO03] or elsewhere. This includes most of Section 2 (including Prop. 2.4), parts of Section 3 and 4 (including Props. 3.4 and 3.5, and the connections to the Gordan and Motzkin theorems), and the duality gap material of Section 5.

Generally, we assume that the reader is familiar with the standard notions of convex analysis, as established in the book by Rockafellar [Roc70], and also described in other books such as Auslender and Teboulle [AuT03], Barvinok [Bar02], Bertsekas, Nedic, and Ozdaglar [BNO03], Ben-Tal and Nemirovski [BeN01], Bonnans and Shapiro [BoS00], Borwein and Lewis [BoL00], Ekeland and Temam [EkT76], Hiriart-Urrutu and Lemarechal [HiL93], Stoer and Witzgall [StW70], Webster [Web94]. For easy reference, we collect in an appendix some of the convex analysis results that we will be using.

Regarding notation, all 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 $\mathrm{cl}(C)$, $\mathrm{int}(C)$, and $\mathrm{conv}(C)$ the closure, the interior, and the convex hull of a set C, respectively. We denote by $\mathrm{aff}(C)$ the affine hull of a set 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 $\mathrm{aff}(C)$. We denote by $\mathrm{epi}(f)$ and $\mathrm{dom}(f)$ the epigraph and the effective domain, respectively, of an extended real-valued function $f : X \mapsto [-\infty, \infty]$:

$$\mathrm{epi}(f) = \{(x,w) \mid f(x) \leq w,\, x \in X,\, w \in \Re\}, \qquad \mathrm{dom}(f) = \{x \mid f(x) < \infty,\, x \in X\}.$$

We say that f is convex if $\mathrm{epi}(f)$ is convex. We say that f is 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 $x \in X$. We say that f is *lower semicontinuous* at a vector $x \in X$ if $f(x) \leq \liminf_{k\to\infty} f(x_k)$ for every sequence $\{x_k\} \subset X$ with

$x_k \to x$. If f is lower semicontinuous at every x in a subset Y of X, we say that f is *lower semicontinuous over* Y. We say that f *is lower semicontinuous* if it is lower semicontinuous over its entire domain X. We use a similar terminology for upper semicontinuous functions. We say that f is closed if $\mathrm{epi}(f)$ is closed.† We remind the reader that if $\mathrm{dom}(f)$ is closed and f is lower semicontinuous over $\mathrm{dom}(f)$, then f is closed (the converse need not be true). Furthermore, a function $f : X \mapsto [-\infty, \infty]$ is lower semicontinuous if and only if it is closed. The function whose epigraph is $\mathrm{cl}\big(\mathrm{epi}(f)\big)$ is called the closure of f and is denoted by $\mathrm{cl}\, f$. These definitions are elaborated on in the appendix at the end of the paper, together with additional convex analysis definitions and background, relating to recession cones, hyperplanes, polyhedral convexity, conjugate functions, and Fenchel duality.

2. GENERAL RESULTS AND SOME SPECIAL CASES

In this section we establish some generic properties of the MC/MC framework, and its connection with conjugacy. We also derive the form of the dual function on some important special cases. The following proposition gives a basic semicontinuity property of the crossing function q.

Proposition 2.1: The crossing function q is concave and upper semicontinuous over $\Re^n$.

Proof: By definition [cf. Eq. (1.2)], q is the infimum of a collection of affine functions, from which the result follows. **Q.E.D.**

We next establish the weak duality property, which is intuitively apparent from Fig. 1.1.

Proposition 2.2: (Weak Duality Theorem) We have $q^* \le w^*$.

Proof: For every $(u, w) \in M$ and $\mu \in \Re^n$, we have

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

† In Rockafellar [Roc70], the definition of a closed function is somewhat different. For proper and convex functions, our definition of a closed function and the one of [Roc70] (p. 52) coincide.

so by taking the supremum of the left-hand side over $\mu \in \Re^n$, we obtain $q^* \leq w^*$. **Q.E.D.**

As Fig. 1.2 indicates, the feasible solutions of the max crossing problem are restricted by the horizontal directions of recession of $\overline{M}$. This is the essence of the following proposition.

Proposition 2.3: Assume that the set

$$\overline{M} = M + \big\{(0, w) \mid w \geq 0\big\}$$

is convex. Then the set of feasible solutions of the max crossing problem, $\big\{\mu \mid q(\mu) > -\infty\big\}$, is contained in the cone

$$\big\{\mu \mid \mu' d \geq 0 \text{ for all } d \text{ with } (d, 0) \in R_{\overline{M}}\big\},$$

where $R_{\overline{M}}$ is the recession cone of $\overline{M}$.

Proof: Let $(\overline{u}, \overline{w}) \in \overline{M}$. If $(d, 0) \in R_{\overline{M}}$, then $(\overline{u} + \alpha d, \overline{w}) \in \overline{M}$ for all $\alpha \geq 0$, so that for all $\mu \in \Re^n$,

$$q(\mu) = \inf_{(u,w) \in M} \{w + \mu' u\} = \inf_{(u,w) \in \overline{M}} \{w + \mu' u\} \leq \overline{w} + \mu' \overline{u} + \alpha \mu' d, \qquad \forall\ \alpha \geq 0.$$

Thus, if $\mu' d < 0$, we must have $q(\mu) = -\infty$, implying that $\mu' d \geq 0$ for all μ with $q(\mu) > -\infty$. **Q.E.D.**

As an example, consider the case where $\overline{M}$ is the vector sum of a convex set and the nonnegative orthant of $\Re^{n+1}$. Then it can be seen that the set $\big\{d \mid (d, 0) \in R_{\overline{M}}\big\}$ contains the nonnegative orthant of $\Re^n$, so the preceding proposition implies that $\mu \geq 0$ for all μ such that $q(\mu) > -\infty$. This case arises in optimization problems with inequality constraints (see Sections 3 and 4).

Connection to Conjugate Convex Functions

Consider the case where the set M is the epigraph of a function $p : \Re^n \mapsto [-\infty, \infty]$. Then M coincides with the set $\overline{M}$ of Eq. (1.4) (cf. Fig. 1.1), and the optimal min common value is

$$w^* = p(0).$$

The crossing function q of Eq. (1.2) is given by

$$q(\mu) = \inf_{(u,w) \in \text{epi}(p)} \{w + \mu' u\} = \inf_{\{(u,w) \mid p(u) \leq w\}} \{w + \mu' u\},$$

and finally

$$q(\mu) = \inf_{u \in \Re^m} \big\{ p(u) + \mu' u \big\}. \tag{2.1}$$

Thus, we have $q(\mu) = -h(-\mu)$, where

$$h(\mu) = \sup_{u \in \Re^n} \big\{ \mu' u - p(u) \big\}$$

is the conjugate convex function of p [cf. Eq. (1.1)]. Since the conjugate of the conjugate of a closed proper convex function p is p itself (cf. Prop. 7.15), this means that if we consider the min common problem defined by the epigraph of $-q$, the corresponding max crossing function is $-p$, provided p is closed proper and convex.

General Optimization Duality

Consider the problem of minimizing a function $f : \Re^n \mapsto [-\infty, \infty]$. We introduce a function $F : \Re^{n+r} \mapsto [-\infty, \infty]$ of the pair (x, u), which satisfies

$$f(x) = F(x, 0), \qquad \forall\ x \in \Re^n.$$

Let the function $p : \Re^n \mapsto [-\infty, \infty]$ be defined by

$$p(u) = \inf_{x \in \Re^n} F(x, u). \tag{2.2}$$

If we view $u \in \Re^r$ as a perturbation, then $p(u)$ has the classical interpretation of a perturbation function. It represents the optimal value of an optimization problem whose cost function is perturbed by u. The perturbed problem coincides with the original problem of minimizing f when $u = 0$.

Consider the MC/MC framework with

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

The min common value w^* is the minimal value of f, since

$$w^* = p(0) = \inf_{x \in \Re^n} F(x, 0) = \inf_{x \in \Re^n} f(x).$$

By Eq. (2.1), the crossing function is

$$q(\mu) = \inf_{u \in \Re^r} \big\{ p(u) + \mu' u \big\} = \inf_{(x,u) \in \Re^{n+r}} \big\{ F(x, u) + \mu' u \big\}, \tag{2.3}$$

and the max crossing problem is

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

Optimization with Inequality Constraints

Different choices of perturbation structure and function F yield corresponding MC/MC frameworks, and corresponding dual problems. An example of this type that we will also consider later is minimization with inequality constraints. Here the problem is

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in C \end{aligned}$$

where $f : \Re^n \mapsto \Re$ is a given function, and C is a constraint set of the form

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

where X is a nonempty subset of $\Re^n$, and $g(x) = \big(g_1(x), \ldots, g_r(x)\big)$ with $g_j : \Re^n \mapsto \Re$ being given functions. We introduce a "perturbed constraint set" of the form

$$C_u = \big\{x \in X \mid g(x) \leq u\big\}, \qquad u \in \Re^r,$$

and the function

$$F(x,u) = \begin{cases} f(x) & \text{if } x \in C_u, \\ \infty & \text{otherwise,} \end{cases}$$

which satisfies the condition $F(x,0) = f(x)$ for all $x \in C$.

Note that the function p of Eq. (2.2) is given by

$$p(u) = \inf_{x \in X,\, g(x) \leq u} f(x),$$

so $p(u)$ is the well-known *primal function* or *perturbation function*, which captures the essential structure of the constrained minimization problem, relating to duality and other properties, such as sensitivity. Furthermore, from Eq. (2.3),

$$q(\mu) = \inf_{u \in \Re^r} \big\{p(u) + \mu' u\big\} = \begin{cases} \inf_{x \in X} \big\{f(x) + \mu' g(x)\big\}, & \text{if } \mu \geq 0, \\ -\infty, & \text{otherwise,} \end{cases}$$

so q is the standard dual function, obtained by minimizing over $x \in X$ the Lagrangian function $f(x) + \mu' g(x)$.

Fenchel Duality Framework

Let us also consider the Fenchel duality framework, involving the problem

$$\begin{aligned} &\text{minimize} \quad f_1(x) - f_2(Qx) \\ &\text{subject to} \quad x \in \Re^n, \end{aligned} \tag{2.4}$$

where Q is an $m \times n$ matrix, and $f_1 : \Re^n \mapsto (-\infty, \infty]$ and $f_2 : \Re^m \mapsto [-\infty, \infty)$ are extended real-valued functions. The dual problem within this framework is

$$\begin{aligned} &\text{maximize} \quad h_2(\mu) - h_1(Q'\mu) \\ &\text{subject to} \quad \mu \in \Re^m, \end{aligned} \tag{2.5}$$

where

$$h_1(Q'\mu) = \sup_{y \in \Re^n} \big\{ y'Q'\mu - f_1(y) \big\}, \qquad h_2(\mu) = \inf_{z \in \Re^m} \big\{ z'\mu - f_2(z) \big\},$$

(see e.g., [Roc70], [BNO03]). Note that h_1 is the conjugate convex function of f_1, and h_2 is the conjugate concave function of f_2, i.e., h_2 is given by $h_2(\mu) = -f_2^*(-\mu)$, where f_2^* is the conjugate function of $-f_2$.

Consider the function

$$F(x, u) = f_1(x) - f_2(Qx + u),$$

where $u \in \Re^m$ represents the perturbation, the corresponding function p of Eq. (2.2), and the MC/MC framework with $M = \text{epi}(p)$. Then, by using Eqs. (2.2) and (2.3), it can be verified that the crossing function q is

$$q(\mu) = h_2(\mu) - h_1(Q'\mu),$$

so the corresponding max crossing problem coincides with the Fenchel dual problem.

Minimax Problems

Consider a function $\phi : X \times Z \mapsto \Re$, where X and Z are nonempty subsets of $\Re^n$ and $\Re^m$, respectively. We wish to either

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

or

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

An important question is whether the minimax equality

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z), \tag{2.6}$$

holds, and whether the infimum and the supremum above are attained. This is significant in a zero sum game context, as well as in optimization duality theory (we will return to this issue later).

We introduce the function $p : \Re^m \mapsto [-\infty, \infty]$ given by

$$p(u) = \inf_{x \in X} \sup_{z \in Z} \big\{ \phi(x, z) - u'z \big\}, \qquad u \in \Re^m, \tag{2.7}$$

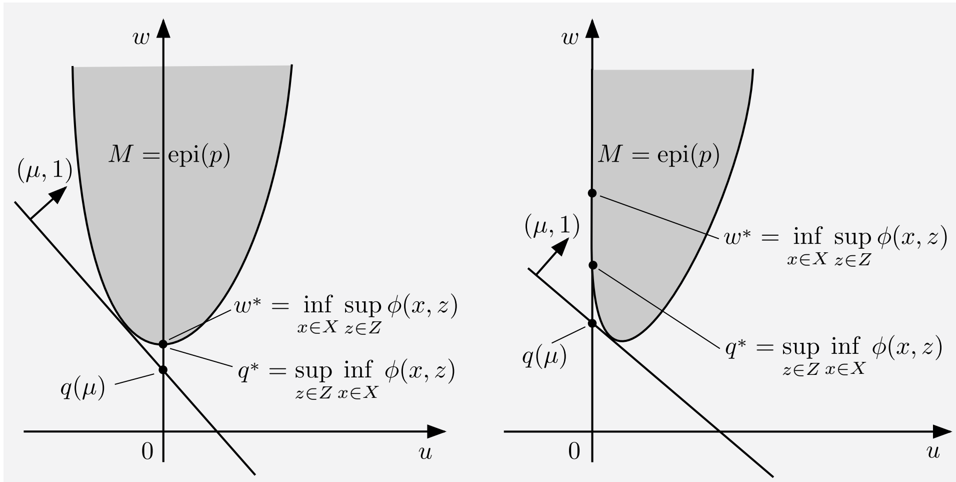

Figure 2.1. MC/MC framework for minimax theory. The set M is taken to be the epigraph of the function

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

The "infsup" value of ϕ is equal to the min common value w^*. Under suitable assumptions, the "supinf" values of ϕ will turn out to be equal to the max crossing value q^*. The figures illustrate cases where p is convex. On the left, the minimax equality holds, while on the right it does not because p is not lower semicontinuous at 0.

which can be viewed as a perturbation function. It characterizes how the "infsup" of the function ϕ changes when the linear perturbation term $u'z$ is subtracted from ϕ. We consider the MC/MC framework with

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

so that the min common value is

$$w^* = p(0) = \inf_{x\in X} \sup_{z\in Z} \phi(x,z). \tag{2.8}$$

We will show that the max crossing value q^* is equal to the "supinf" value of a function derived from ϕ via a convexification operation. Given a set $X \subset \Re^n$ and a function $f : X \mapsto [-\infty, \infty]$, the *convex closure of* f, denoted by $\check{\mathrm{cl}}\, f$, is the function whose epigraph is the closure of the epigraph of f (see the Appendix for analysis and properties of $\check{\mathrm{cl}}\, f$). The *concave closure of* f, denoted by $\hat{\mathrm{cl}}\, f$, is the opposite of the convex closure of $-f$, i.e.,

$$\hat{\mathrm{cl}}\, f = -\check{\mathrm{cl}}\,(-f).$$

It is the smallest concave and upper semicontinuous function that majorizes f, i.e., $\hat{\mathrm{cl}}\, f \le g$ for any $g : X \mapsto [-\infty, \infty]$ that is concave and upper semicontinuous with $g \ge f$. Note that we have

$$\sup_{x\in X} f(x) = \sup_{x\in\Re^n} (\hat{\mathrm{cl}}\, f)(x) \tag{2.9}$$

(cf. Prop. 7.10).

Proposition 2.4: Let X and Z be nonempty subsets of $\Re^n$ and $\Re^m$, respectively, and let $\phi : X \times Z \mapsto \Re$ be a function. For each $x \in X$, let $(\hat{\text{cl}}\,\phi)(x,\cdot)$ be the concave closure of the function $\phi(x,\cdot)$, and assume that $(\hat{\text{cl}}\,\phi)(x,\mu) < \infty$ for all $x \in X$ and $\mu \in Z$. Then, we have

$$q(\mu) = \inf_{x\in X} (\hat{\text{cl}}\,\phi)(x,\mu), \qquad \forall\ \mu \in \Re^m. \tag{2.10}$$

Proof: Let us write

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

where

$$p_x(u) = \sup_{z\in Z}\big\{\phi(x,z) - u'z\big\}, \qquad x \in X,$$

and note that

$$\inf_{u\in\Re^m} \big\{p_x(u) + u'\mu\big\} = -\sup_{u\in\Re^m} \big\{u'(-\mu) - p_x(u)\big\} = -h_x(-\mu), \tag{2.11}$$

where h_x is the conjugate of p_x. Since $p_x(u) = \zeta_x(-u)$, where ζ_x is the conjugate of $(-\phi)(x,\cdot)$, from the Conjugacy Theorem [Prop. 7.15(d)], using also the assumption $(\hat{\text{cl}}\,\phi)(x,\mu) < \infty$ for all $x \in X$ and $\mu \in Z$, we obtain

$$h_x(-\mu) = -(\hat{\text{cl}}\,\phi)(x,\mu). \tag{2.12}$$

Now for every $\mu \in \Re^m$, using Eq. (2.1) for q, and Eqs. (2.11), (2.12), we obtain

$$\begin{aligned} q(\mu) &= \inf_{u\in\Re^m} \big\{p(u) + u'\mu\big\} \\ &= \inf_{u\in\Re^m}\inf_{x\in X} \big\{p_x(u) + u'\mu\big\} \\ &= \inf_{x\in X}\inf_{u\in\Re^m} \big\{p_x(u) + u'\mu\big\} \\ &= \inf_{x\in X} \big\{-h_x(-\mu)\big\} \\ &= \inf_{x\in X} (\hat{\text{cl}}\,\phi)(x,\mu)\big\}. \end{aligned} \tag{2.13}$$

Q.E.D.

Proposition 2.4 leads to several conclusions regarding the "infsup" and "supinf" values of $\phi(x,z)$, and the associated MC/MC values w^* and q^*.

(a) In general, we have

$$\sup_{z\in Z}\inf_{x\in X} \phi(x,z) \le q^* \le w^* = \inf_{x\in X}\sup_{z\in Z}\phi(x,z), \tag{2.14}$$

where the first inequality follows by writing for all $\mu \in Z$,

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

and taking supremum over μ, the second inequality is the weak duality relation, and the last equality follows by the definition of w^*. Therefore *the minimax equality (2.6) always implies the strong duality relation* $q^* = w^*$.

(b) If

$$\phi(x,z) = (\hat{\mathrm{cl}}\,\phi)(x,z), \qquad \forall\, x \in X,\, z \in Z,$$

as in the case where $-\phi(x,\cdot)$ is closed and convex for all $x \in X$, then using Prop. 2.4,

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

where the second equality follows from the fact $(\hat{\mathrm{cl}}\,\phi)(x,z) = -\infty$ for $x \in X$ and $z \notin Z$. In view of Eq. (2.14), it follows that the strong duality relation $q^* = w^*$ is equivalent to the minimax equality (2.6).

(c) From Eq. (2.9), we have $\sup_{z\in Z} \phi(x,z) = \sup_{z\in\Re^m} (\hat{\mathrm{cl}}\,\phi)(x,z)$, so the min common value w^* is equal to $\inf_{x\in X} \sup_{z\in\Re^m} (\hat{\mathrm{cl}}\,\phi)(x,z)$. The max crossing value q^* is equal to $\sup_{z\in\Re^m} \inf_{x\in X} (\hat{\mathrm{cl}}\,\phi)(x,z)$, assuming that $(\hat{\mathrm{cl}}\,\phi)(x,\mu) < \infty$ for all $x \in X$ and $\mu \in Z$, so that Prop. 2.4 applies. Thus, w^* and q^* are the "infsup" and "supinf" values of $\hat{\mathrm{cl}}\,\phi$, respectively. If $(\hat{\mathrm{cl}}\,\phi)(\cdot,z)$ is closed and convex for each z, then the "infsup" and "supinf" values of the convex/concave function $\hat{\mathrm{cl}}\,\phi$ will ordinarily be expected to be equal (see Section 4). In this case we will have strong duality ($q^* = w^*$), but this will not necessarily imply that the minimax equality (2.6) holds, because strict inequality may hold in the first relation of Eq. (2.14). If $(\hat{\mathrm{cl}}\,\phi)(\cdot,z)$ is not closed and convex for some z, then the strong duality relation $q^* = w^*$ should not ordinarily be expected. The size of the duality gap $w^* - q^*$ will depend on the difference between $(\hat{\mathrm{cl}}\,\phi)(\cdot,z)$ and its convex closure, and may be investigated in some cases by using methods to be presented in Section 5.

Example 2.1: (Finite Set Z)

Let

$$\phi(x,z) = z'f(x),$$

where $f : X \mapsto \Re^m$, X is a subset of $\Re^n$, and $f(x)$ is viewed as a column vector whose components are functions $f_j : X \mapsto \Re$, $j = 1,\ldots,m$. Suppose that Z is the finite set $Z = \{e_1,\ldots,e_m\}$, where e_j is the jth column of the $m \times m$ identity matrix. Then, we have

$$\sup_{z\in Z} \inf_{x\in X} \phi(x,z) = \max\left\{\inf_{x\in X} f_1(x),\ldots,\inf_{x\in X} f_m(x)\right\},$$

and

$$w^* = \inf_{x \in X} \sup_{z \in Z} \phi(x, z) = \inf_{x \in X} \max\{f_1(x), \ldots, f_m(x)\}. \tag{2.15}$$

Let $\overline{Z}$ denote the unit simplex in $\Re^m$ (the convex hull of Z). We have

$$(\hat{\mathrm{cl}}\, \phi)(x, z) = \begin{cases} z' f(x), & \text{if } z \in \overline{Z}, \\ -\infty, & \text{if } z \notin \overline{Z}, \end{cases}$$

and by Prop. 2.4,

$$q^* = \sup_{z \in Z} \inf_{x \in X} (\hat{\mathrm{cl}}\, \phi)(x, z) = \sup_{z \in \overline{Z}} \inf_{x \in X} z' f(x).$$

If $f_1, \ldots, f_m$ are convex functions, by considering the convex program

$$\begin{aligned} &\text{minimize} \quad \xi \\ &\text{subject to} \quad x \in X, \qquad f_j(x) \le \xi, \quad j = 1, \ldots, m, \end{aligned}$$

associated with the optimization in Eq. (2.15), we may show that $q^* = w^*$. On the other hand, if $f_1, \ldots, f_m$ are not convex, we may have $q^* < w^*$.

Under certain conditions, q^* can be associated with a special minimax problem derived from the original by introducing mixed (randomized) strategies. This is so in the following classical game context.

Example 2.2: (Finite Zero Sum Games)

Consider a minimax problem where the sets X and Z are finite:

$$X = \{d_1 \ldots, d_n\}, \qquad Z = \{e_1, \ldots, e_m\},$$

where d_i is the ith column of the $n \times n$ identity matrix, and e_j is the jth column of the $m \times m$ identity matrix. Let

$$\phi(x, z) = x' A z,$$

where A is an $n \times m$ matrix. This corresponds to the classical game context, where upon upon selection of $x = d_i$ and $z = e_j$, the payoff is the ijth component of A. Let $\overline{X}$ and $\overline{Z}$ be the unit simplexes in $\Re^n$ and $\Re^m$, respectively. Then it can be seen, similar to the preceding example that

$$q^* = \max_{z \in \overline{Z}} \min_{x \in \overline{X}} x' A z = \min_{x \in \overline{X}} \max_{z \in \overline{Z}} x' A z.$$

Note that $\overline{X}$ and $\overline{Z}$ can be interpreted as sets of mixed strategies, so q^* is the value of the corresponding mixed strategy game.

3. STRONG DUALITY THEOREMS

We will now establish conditions implying that strong duality holds in the MC/MC framework, i.e., $q^* = w^*$, and then obtain conditions that guarantee existence of max crossing hyperplanes. To avoid degenerate cases, we will often exclude the case $w^* = \infty$, which corresponds to an infeasible min common problem.

Conditions for Strong Duality

An important point, around which much of our analysis revolves, is that when w^* is finite, the vector $(0, w^*)$ is a closure point of the set $\overline{M}$ of Eq. (1.4), so *if we assume that $\overline{M}$ is convex and admits a nonvertical supporting hyperplane at $(0, w^*)$, then we have $q^* = w^*$ while 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: (MC/MC Theorem I) Consider the min common and max crossing problems, and assume the following:

(1) Either $w^* < \infty$, or else $w^* = \infty$ and M contains no vertical lines.

(2) The set

$$\overline{M} = M + \big\{(0, w) \mid w \geq 0\big\}$$

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 $w^* \leq \liminf_{k\to\infty} w_k$.

Proof: Consider first the case $w^* = -\infty$. Then, by the Weak Duality Theorem (Prop. 2.2), we also have $q^* = -\infty$, so the conclusion trivially follows. Consider next the case where $w^* = \infty$ and M contains no vertical lines. For every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, we have

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} \leq w_k + \mu' u_k, \qquad \forall\, k, \quad \forall\, \mu \in \Re^n. \tag{3.1}$$

This implies that $q(\mu) \leq \liminf_{k\to\infty} w_k$ and

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

Hence if $q^* = w^* = \infty$, it follows that $\liminf_{k\to\infty} w_k = w^*$. Conversely, if we have $\liminf_{k\to\infty} w_k = w^* = \infty$ for all sequences $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, it follows that the vertical axis of $\Re^{n+1}$ contains no closure points of M. Hence by the Nonvertical Hyperplane Theorem (Prop. 7.12), for any vector $(0, w) \in \Re^{n+1}$, there exists a nonvertical hyperplane strictly separating $(0, w)$ and M. The crossing point of this hyperplane with the vertical axis lies between w and q^*, so $w < q^*$ for all $w \in \Re$, which implies that $q^* = w^* = \infty$.

Consider finally 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 $\mathrm{ri}(\overline{M})$. Because $(0, w^*)$ is a closure point of $\overline{M}$, it is also a closure point of $\mathrm{ri}(\overline{M})$, 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 the Nonvertical Separating Hyperplane Theorem (Prop. 7.12), 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 ($q^* \le w^*$), implies that $w^* - \epsilon \le q^* \le w^*$. Since ϵ can be arbitrarily small, it follows that $q^* = w^*$.

Conversely, assume that $q^* = w^*$. Consider any sequence $\big\{(u_k, w_k)\big\} \subset M$ such that $u_k \to 0$. Then, by taking the limit as $k \to \infty$ in Eq. (3.1), we obtain $q(\mu) \le \liminf_{k\to\infty} w_k$, and

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

Q.E.D.

For an example where assumption (1) of the preceding proposition is violated, let M consist of a vertical line that does not pass through the origin. Then we have $w^* = \infty$, $q^* = -\infty$, and yet the condition $w^* \le \liminf_{k\to\infty} w_k$, for every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, trivially holds.

An important corollary of Prop. 3.1 is that if $M = \mathrm{epi}(p)$ where $p : \Re^n \mapsto [-\infty, \infty]$ is a closed convex function with $p(0) = w^* < \infty$, then we have $q^* = w^*$ if and only if p is lower semicontinuous at 0.

Existence of Dual Optimal Solutions

We now discuss the nonemptiness and the structure of the optimal solution set of the max crossing problem. The following proposition, in addition to the equality $q^* = w^*$, guarantees the attainment of the maximum crossing point by a nonvertical hyperplane under an additional assumption.

Proposition 3.2: (MC/MC Theorem II) Consider the min common and max crossing problems, and assume the following:

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

(2) The set

$$\overline{M} = M + \big\{(0, w) \mid w \ge 0\big\}$$

is convex.

(3) The origin is in the relative interior of the set

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

Then $q^* = w^*$, and there exists at least one optimal solution of the max crossing problem.

Proof: We first note that condition (3) implies that the vertical axis contains a point of M, so that $w^* < \infty$. Hence in view of condition (1), w^* is finite.

We note that $(0, w^*)$ is not a relative interior point of $\overline{M}$, since the line $\big\{(0, w) \mid w \in \Re\big\}$ is contained in the affine hull of $\overline{M}$, and w^* is the minimum value corresponding to vectors in the intersection of this line and $\overline{M}$. Therefore, by the Proper Separation Theorem (Prop. 7.13), 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 (μ, β) such that

$$\beta w^* \le \mu' u + \beta w, \qquad \forall\ (u, w) \in \overline{M}, \tag{3.2}$$

$$\beta w^* < \sup_{(u,w)\in \overline{M}} \{\mu' u + \beta w\}. \tag{3.3}$$

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. (3.2) that $\beta \geq 0$. If $\beta = 0$, then from Eq. (3.2), 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 condition (2), it follows (see Prop. 7.2) that $\mu' u$ is constant over D, i.e.,

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

This, however, contradicts Eq. (3.3). Therefore, we must have $\beta > 0$, and by appropriate normalization if necessary, we can assume that $\beta = 1$. From Eq. (3.2), 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 we also have $q^* \leq w^*$ by the Weak Duality Theorem (Prop. 2.2), equality holds throughout in the above relation, and we must have $q(\mu) = q^* = w^*$. Thus μ is an optimal solution of the max crossing problem. **Q.E.D.**

Note that if $w^* = -\infty$, by weak duality, we have $q^* \leq w^*$, so that $q^* = w^* = -\infty$. This means that $q(\mu) = -\infty$ for all $\mu \in \Re^n$, and that the dual problem is infeasible. The following proposition supplements the preceding one, and characterizes the optimal solution set of the max crossing problem.

Proposition 3.3: Let the assumptions of Prop. 3.2 hold. Then the set of optimal solutions Q^* of the max crossing problem has the form

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

where $\tilde{Q}$ is a nonempty, convex, and compact set. In particular, Q^* is compact if and only if the origin is in the interior of D.

Proof: By Prop. 3.2, Q^* is nonempty. Since $Q^* = \big\{\mu \mid q(\mu) \geq q^*\big\}$ and q is concave and upper semicontinuous (cf. Prop. 2.1), it follows that Q^* is also convex and closed. We now prove that the recession cone R_{Q^*}

and the lineality space L_{Q^*} of Q^* are both equal to $\big(\text{aff}(D)\big)^\perp$ [note here that $\text{aff}(D)$ is a subspace since it contains the origin]. The proof of this is based on the generic relation $L_{Q^*} \subset R_{Q^*}$ and the following two relations

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

which we show next.

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

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

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

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

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

Since $0 \in \text{ri}(D)$, for any $u \in \text{aff}(D)$, 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

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

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

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

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

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

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

Therefore

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

Based on the decomposition result of Prop. 7.4, we have

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

Since $L_{Q^*} = \big(\text{aff}(D)\big)^\perp$, we obtain

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

where $\tilde{Q} = Q^* \cap \mathrm{aff}(D)$. Furthermore, by Prop. 7.3, 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.

From the formula $Q^* = \big(\mathrm{aff}(D)\big)^\perp + \tilde{Q}$, it follows that Q^* is compact if and only if $\big(\mathrm{aff}(D)\big)^\perp = \{0\}$, or equivalently $\mathrm{aff}(D) = \Re^n$. Since 0 is a relative interior point of D by assumption, this is equivalent to 0 being an interior point of D. **Q.E.D.**

Special Cases Involving Convexity and/or Compactness

We now consider the cases where M is closed and convex, and also the case where M has some compactness structure but is not necessarily convex.

Proposition 3.4: Consider the MC/MC framework, assuming that $w^* < \infty$.

(a) Let M be closed and convex. Then $q^* = w^*$. Furthermore, $\overline{M}$ is the epigraph of the convex function

$$p(u) = \inf\big\{w \mid (u, w) \in M\big\}, \qquad u \in \Re^n.$$

If in addition $-\infty < w^*$, then p is closed and proper.

(b) q^* is equal to the optimal value of the min common problem corresponding to $\mathrm{cl}\big(\mathrm{conv}(M)\big)$.

(c) If M is of the form

$$M = \tilde{M} + \big\{(u, 0) \mid u \in C\big\},$$

where $\tilde{M}$ is a compact set and C is a closed convex set, then q^* is equal to the optimal value of the min common problem corresponding to $\mathrm{conv}(M)$.

Proof: (a) Since M is closed, for each $u \in \mathrm{dom}(p)$, the infimum defining $p(u)$ is either $-\infty$ or else it is attained. In view of the definition of $\overline{M}$, this implies that $\overline{M}$ is the epigraph of p. Furthermore, $\overline{M}$ is convex, being the vector sum of two convex sets, so p is convex.

If $w^* = -\infty$, then $q^* = w^*$ by weak duality. Thus for the remainder of the proof, we may assume that $-\infty < w^*$. This implies that $(0, -1)$ is not a direction of recession of M. On the other hand, the only nonzero direction of recession of the halfline $\{(0,\alpha) \mid \alpha \geq 0\}$ is $(0,1)$. Since $\overline{M}$ is the vector sum of M and $\{(0,\alpha) \mid \alpha \geq 0\}$, it follows from Prop. 7.5 that $\overline{M}$ is closed as well as convex. This in turn implies that p is closed and convex, and by Prop. 3.1, we have $q^* = w^*$. Since w^* is finite, it follows that p is proper (an improper closed and convex function cannot take finite values, see [BNO03], p. 29).

(b) The max crossing value q^* is the same for M and $\text{cl}\big(\text{conv}(M)\big)$, since the closed halfspaces containing M are the ones that contain $\text{cl}\big(\text{conv}(M)\big)$. Since $\text{cl}\big(\text{conv}(M)\big)$ is closed and convex, the result follows from part (a).

(c) The max crossing value is the same for M and $\text{conv}(M)$, since the closed halfspaces containing M are the ones that contain $\text{conv}(M)$. It can be seen that

$$\text{conv}(M) = \text{conv}(\tilde{M}) + \big\{(u,0) \mid u \in C\big\},$$

from which it follows that the upwards extension of $\text{conv}(M)$ is given by

$$\overline{\text{conv}(M)} = \text{conv}(\tilde{M}) + \big\{(u,w) \mid u \in C,\, w \geq 0\big\}.$$

Since $\tilde{M}$ is compact, $\text{conv}(\tilde{M})$ is also compact, so $\overline{\text{conv}(M)}$ is the vector sum of a compact set and a closed convex set. Hence, by Prop. 7.5, $\overline{\text{conv}(M)}$ is closed, and the result follows from part (a). **Q.E.D.**

Note that if M is convex and closed, and $w^* = -\infty$, then p is convex, but need not be closed [cf. Prop. 3.4(a)]. For an example in $\Re^2$, consider the closed and convex set

$$M = \big\{(u,w)\} \mid w \leq -1/(1-|u|),\, |x| < 1\big\}.$$

Then, $\overline{M} = \big\{(u,w) \mid |u| < 1\big\}$, so $\overline{M}$ is convex but not closed, implying that p (which is $-\infty$ for $|u| < 1$ and ∞ otherwise) is not closed. Note also if M is closed but is neither convex nor compact (e.g., when M is the epigraph of some nonconvex function) the property of part (c) above may not hold. For an example in $\Re^2$, consider the set

$$M = \big\{(0,0)\big\} \cup \big\{(u,w) \mid u > 0,\, w \leq -1/u\big\}.$$

Then

$$\text{conv}(M) = \big\{(0,0)\big\} \cup \big\{(u,w) \mid u > 0,\, w < 0\big\}.$$

We have $q^* = -\infty$ but the min common value corresponding to $\text{conv}(M)$ is $w^* = 0$ [the min common value corresponding to $\text{cl}\big(\text{conv}(M)\big)$ is equal to $-\infty$, consistent with part (b)].

MC/MC and Polyhedral Convexity

We now provide an extension of MC/MC Theorem II (Prop. 3.2) for the case where the "upwards extension"

$$\overline{M} = M + \big\{(0, w) \mid w \geq 0\big\}$$

of the set M has a partially polyhedral structure. In particular, we will consider the special case where $\overline{M}$ is the vector difference of two sets of the form

$$\overline{M} = \tilde{M} - \big\{(u, 0) \mid u \in P\big\}, \tag{3.4}$$

where $\tilde{M}$ is convex and P is polyhedral. Then the corresponding set

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

can be written as

$$D = \tilde{D} - P,$$

where

$$\tilde{D} = \big\{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in \tilde{M}\big\}. \tag{3.5}$$

To understand the nature of the following proposition, we note that from Props. 3.1-3.3, assuming that $-\infty < w^*$, we have:

(a) $q^* = w^*$, and Q^* is nonempty, if $0 \in \mathrm{ri}(D)$, which is equivalent to

$$\mathrm{ri}(\tilde{D}) \cap \mathrm{ri}(P) \neq \emptyset, \tag{3.6}$$

since $\mathrm{ri}(D) = \mathrm{ri}(\tilde{D}) - \mathrm{ri}(P)$.

(b) $q^* = w^*$, and Q^* is nonempty and compact, if $0 \in \mathrm{int}(D)$, which is true in particular if either

$$\mathrm{int}(\tilde{D}) \cap P \neq \emptyset,$$

or

$$\tilde{D} \cap \mathrm{int}(P) \neq \emptyset.$$

(c) Every $\mu \in Q^*$ satisfies

$$\mu' y \geq 0, \qquad \forall\, y \text{ such that } (y, 0) \in R_{\overline{M}},$$

where $R_{\overline{M}}$ is the recession cone of $\overline{M}$, from which it follows that

$$Q^* \subset R_P^*,$$

where R_P^* is the polar of the recession cone of P.

The following proposition shows that when P is polyhedral, these results can be strengthened, and in particular, the condition $\mathrm{ri}(\tilde{D}) \cap \mathrm{ri}(P) \neq \emptyset$ [Eq. (3.6)] can be replaced by the condition

$$\mathrm{ri}(\tilde{D}) \cap P \neq \emptyset.$$

The proof is very similar to the proofs of Props. 3.2 and 3.3, with essentially the only difference being the use of the Polyhedral Proper Separation Theorem (Prop. 7.14) in place of the Proper Separation Theorem (Prop. 7.13).

Proposition 3.5: (MC/MC Theorem III) Consider the min common and max crossing problems, and assume the following:

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

(2) The set $\overline{M}$ has the form

$$\overline{M} = \tilde{M} - \big\{(u, 0) \mid u \in P\big\},$$

where P is a polyhedral set and $\tilde{M}$ is a convex set.

(3) We have

$$\mathrm{ri}(\tilde{D}) \cap P \neq \emptyset,$$

where $\tilde{D}$ is the set given by Eq. (3.5).

Then $q^* = w^*$, and Q^*, the set of optimal solutions of the max crossing problem, is a nonempty subset of R_P^*, the polar cone of the recession cone of P. Furthermore, Q^* is compact if $\mathrm{int}(\tilde{D}) \cap P \neq \emptyset$.

Proof: We consider the sets

$$C_1 = \big\{(u, v) \mid v > w \text{ for some } (u, w) \in \tilde{M}\big\},$$

$$C_2 = \big\{(u, w^*) \mid u \in P\big\},$$

(cf. Fig. 3.1). It can be seen that C_1 and C_2 are nonempty and convex, and C_2 is polyhedral. Furthermore, C_1 and C_2 are disjoint. To see this, note that if $\overline{u}$ is such that $\overline{u} \in P$ and there exists $(\overline{u}, w)$ with $w^* > w$, then $(0, w) \in \overline{M}$, which is impossible since w^* is the min common value. Therefore, by Polyhedral Proper Separation Theorem (Prop. 7.14), there exists a hyperplane that separates C_1 and C_2, and does not contain C_1, i.e., a vector $(\overline{\mu}, \beta)$ such that

$$\beta w^* + \overline{\mu}' z \leq \beta v + \overline{\mu}' u, \qquad \forall\ (u, v) \in C_1,\ \forall\ z \in P, \tag{3.7}$$

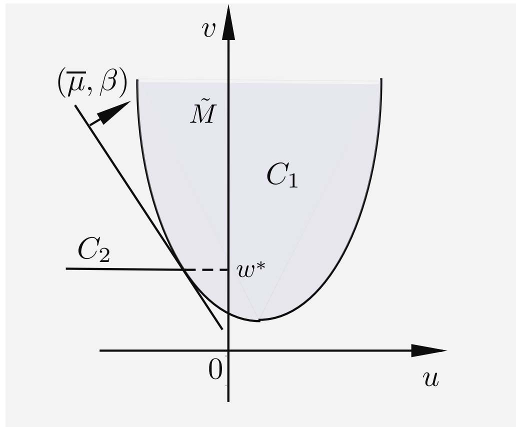

Figure 3.1. Illustration of the sets

$$C_1 = \Big\{(u, v) \mid v > w \text{ for some } (u, w) \in \tilde{M}\Big\}, \qquad C_2 = \Big\{(u, w^*) \mid u \in P\Big\},$$

and the hyperplane separating them in the proof of Prop. 3.5.

$$\inf_{(x,v)\in C_1} \big\{\beta v + \overline{\mu}' x\big\} < \sup_{(x,v)\in C_1} \big\{\beta v + \overline{\mu}' x\big\}. \tag{3.8}$$

From Eq. (3.7), since $(0, 1)$ is a direction of recession of C_1, we see that $\beta \geq 0$. If $\beta = 0$, then for a vector $\overline{u} \in \mathrm{ri}(\tilde{D}) \cap P$, we have $\overline{\mu}'\overline{u} \leq \inf_{u\in\tilde{D}} \overline{\mu}' u$, so $\overline{u}$ attains the minimum of the linear function $\overline{\mu}' u$ over $\tilde{D}$, and from Prop. 7.2, it follows that $\overline{\mu}' u$ is constant over $\tilde{D}$. On the other hand, by Eq. (3.8) we have $\inf_{u\in\tilde{D}} \overline{\mu}' u < \sup_{u\in\tilde{D}} \overline{\mu}' u$, a contradiction. Hence $\beta > 0$, and by normalizing $(\overline{\mu}, \beta)$ if necessary, we may assume that $\beta = 1$.

Thus, from Eq. (3.7), we have

$$w^* + \overline{\mu}' z \leq \inf_{(u,v)\in C_1} \{v + \overline{\mu}' u\}, \qquad \forall\ z \in P, \tag{3.9}$$

which in particular implies that $\overline{\mu}' d \leq 0$ for all $d \in R_P$. Hence $\overline{\mu} \in R_P^*$. From Eq. (3.9), we also obtain

$$\begin{aligned} w^* &\leq \inf_{(u,v)\in C_1,\, z\in P} \big\{v + \overline{\mu}'(u - z)\big\} \\ &= \inf_{(u,v)\in \tilde{M}-P} \{v + \overline{\mu}' u\} \\ &= \inf_{(u,v)\in \overline{M}} \{v + \overline{\mu}' u\} \\ &= q(\overline{\mu}). \end{aligned}$$

Using the weak duality relation $q^* \leq w^*$, we have $q(\overline{\mu}) = q^* = w^*$. The proofs that $Q^* \subset R_P^*$ and that Q^* is compact if $\mathrm{int}(\tilde{D}) \cap P \neq \emptyset$ are similar to the ones of Props. 3.2 and 3.3 (see the discussion preceding the proposition). **Q.E.D.**

We now discuss an interesting special case of the preceding theorem. It corresponds to $\tilde{M}$ being a linearly transformed epigraph of a convex function f, and P being a polyhedral cone, such as the nonpositive orthant. This special case applies to constrained optimization duality, and will be used in the proof of the subsequent Nonlinear Farkas' Lemma.

Proposition 3.6: Consider the min common and max crossing problems, and assume that:

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

(2) The set $\overline{M}$ is defined in terms of a convex function $f : \Re^m \mapsto (-\infty, \infty]$, an $r \times m$ matrix A, a vector $b \in \Re^r$, and a polyhedral cone P as follows:

$$\overline{M} = \big\{(u, w) \mid Ax - b - u \in P \text{ for some } (x, w) \in \text{epi}(f)\big\}.$$

(3) There is a vector $\overline{x} \in \text{ri}\big(\text{dom}(f)\big)$ such that $A\overline{x} - b \in P$.

Then $q^* = w^*$ and Q^*, the set of optimal solutions of the max crossing problem, is a nonempty subset of

$$R_P^* = \{\mu \mid \mu' y \leq 0, \forall\, y \in R_P\},$$

the polar cone of the recession cone of P. Furthermore, Q^* is compact if the matrix A has rank r and there is a vector $\overline{x} \in \text{int}\big(\text{dom}(f)\big)$ such that $A\overline{x} - b \in P$.

Proof: Let

$$\tilde{M} = \big\{(Ax - b, w) \mid (x, w) \in \text{epi}(f)\big\}.$$

The following calculation relates $\tilde{M}$ and $\overline{M}$:

$$\begin{aligned}
&\tilde{M} - \big\{(z, 0) \mid \ z \in P\big\} \\
&\qquad = \big\{(u, w) \mid u = Ax - b - z \text{ for some } (x, w) \in \text{epi}(f) \text{ and } z \in P\big\} \\
&\qquad = \big\{(u, w) \mid Ax - b - u \in P \text{ for some } (x, w) \in \text{epi}(f)\big\} \\
&\qquad = \overline{M}.
\end{aligned}$$

Thus the framework of MC/MC Theorem III (Prop. 3.5) applies. Furthermore, the set $\tilde{D}$ of Eq. (3.5) is given by

$$\tilde{D} = \big\{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in \tilde{M}\big\} = \big\{Ax - b \mid x \in \text{dom}(f)\big\}.$$

Hence, the relative interior assumption (3) implies that the corresponding relative interior assumption of Prop. 3.5 is satisfied. Furthermore, if A has rank r, we have $A\overline{x} - b \in \text{int}(\tilde{D})$ for all $\overline{x} \in \text{int}\big(\text{dom}(f)\big)$. The result follows from the conclusion of Prop. 3.5. **Q.E.D.**

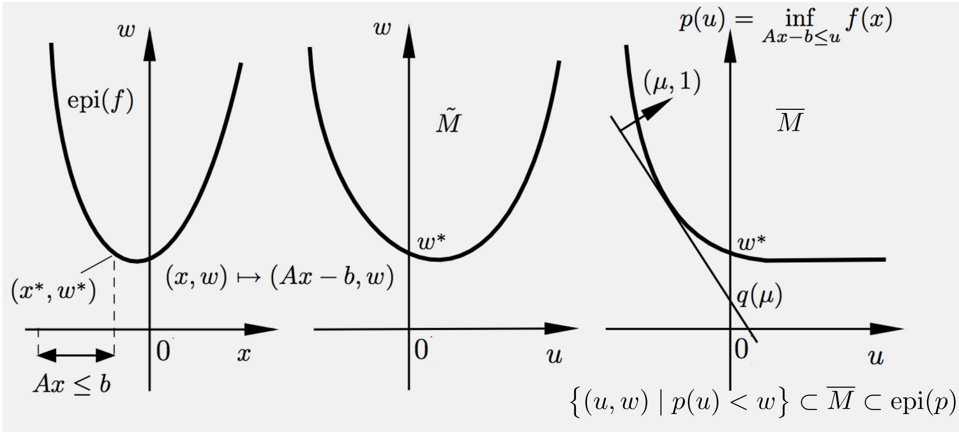

Figure 3.2. The MC/MC framework for minimizing $f(x)$ subject to $Ax \le b$. The set $\overline{M}$ is

$$\overline{M} = \big\{(u,w) \mid Ax - b \le u \text{ for some } (x,w) \in \text{epi}(f)\big\}$$

(cf. Prop. 3.6) and is the vector sum

$$\overline{M} = \tilde{M} + \big\{(u,0) \mid u \ge 0\big\},$$

where $\tilde{M}$ is obtained by transformation of epi(f),

$$\tilde{M} = \big\{(Ax - b, w) \mid (x,w) \in \text{epi}(f)\big\},$$

and P is the negative orthant. Also $\overline{M}$ relates to the perturbation function $p(u) = \inf_{Ax-b\le u} f(x)$ as follows:

$$\big\{(u,w) \mid p(u) < w\big\} \subset \overline{M} \subset \text{epi}(p).$$

In particular, we have $w^* = p(0) = \inf_{Ax\le b} f(x)$.

The special case where P is the negative orthant

$$P = \{u \mid u \le 0\},$$

is noteworthy as it corresponds to a major convex optimization model. In this case, the set $\overline{M}$ in Prop. 3.6 is equal to M and relates to the epigraph of the function

$$p(u) = \inf_{Ax-b\le u} f(x),$$

as shown in Fig. 3.2. The min common value is equal to $p(0)$, the optimal value of the problem of minimizing $f(x)$ subject to $Ax \le b$,

$$w^* = p(0) = \inf_{Ax\le b} f(x).$$

The max crossing problem is to maximize over $\mu \in \Re^r$ the function q given by

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} = \inf_{u\in\Re^r} \big\{p(u) + \mu' u\big\} = \inf_{u\in\Re^r} \inf_{Ax-b\le u} \big\{f(x) + \mu' u\big\},$$

and finally

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

Thus, the max crossing problem is to maximize q over $\mu \geq 0$. As shown by Prop. 3.6, assuming that there exists $\overline{x} \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$ such that $A\overline{x} \leq b$, strong duality holds and there exists a maximizing μ.

4. APPLICATIONS

In this section we show how the MC/MC duality results of the preceding section apply to a broad set of contexts, including minimax theory and optimization duality. However, because of its geometric structure, the MC/MC framework also applies to some settings where conjugacy ordinarily does not play a role, such as for example theorems of the alternative and subdifferential theory.

Minimax Theorems

We will prove two theorems regarding the validity of the minimax equality (2.6), which correspond to the MC/MC theorems of Section 3. We will assume the following.

Assumption 4.1: (Convexity/Concavity and Closedness) X and Z are nonempty convex subsets of $\Re^n$ and $\Re^m$, respectively, and $\phi : X \times Z \mapsto \Re$ is a function such that $\phi(\cdot, z) : X \mapsto \Re$ is convex and closed for each $z \in Z$, and $-\phi(x, \cdot) : Z \mapsto \Re$ is convex and closed for each $x \in X$.

We will use the MC/MC Theorem I (Prop. 3.1) to prove the following proposition.

Proposition 4.1: (Minimax Theorem I) Let Assumption 4.1 hold. Assume further that the function p of Eq. (2.7) satisfies either $p(0) < \infty$, or else $p(0) = \infty$ and $p(u) > -\infty$ for all $u \in \Re^m$. Then

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

if and only if p is lower semicontinuous at $u = 0$, i.e., for all sequences $\{u_k\}$ with $u_k \to 0$, we have $p(0) \leq \liminf_{k\to\infty} p(u_k)$.

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 MC/MC Theorem I.

We choose the set M in the MC/MC 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 Prop. 7.6. Thus, condition (2) of the MC/MC 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 $p(0) < \infty$, or else $p(0) = \infty$ and $p(u) > -\infty$ for all u, is equivalent to condition (1) of the MC/MC Theorem I.

Finally, we have

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

for all $\{u_k\}$ with $u_k \to 0$ if and only if for every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, there holds $w^* \le \liminf_{k \to \infty} w_k$. Thus, by the conclusion of the MC/MC Theorem I, the condition $p(0) \le \liminf_{k \to \infty} p(u_k)$ holds if and only if $q^* = w^*$, which in turn holds if and only if the minimax equality holds [cf. Prop. 2.4, which applies because of the assumed closedness and convexity of $-\phi(x, \cdot)$]. **Q.E.D.**

As an example illustrating the need for the finiteness assumptions on p in the preceding proposition, consider the case where x and z are scalars and

$$\Phi(x, z) = x + z, \qquad X = \{x \mid x \le 0\}, \qquad Z = \{z \mid z \ge 0\}.$$

We have

$$p(u) = \inf_{x \le 0} \sup_{z \ge 0} \{x + z - uz\} = \begin{cases} \infty & \text{if } u < 1, \\ -\infty & \text{if } u \ge 1, \end{cases}$$

so the assumptions of the proposition are violated, and we also have

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

The proof of the preceding proposition can be easily modified to use the second MC/MC Theorem of the preceding section [cf. Props. 3.2 and 3.3, and Eq. (2.10)]. What is needed is an assumption that 0 lies in the relative interior or the interior of the effective domain of p and that $p(0) > -\infty$. We then obtain the following result, which also asserts that the supremum in the minimax equality is attained [this follows from the corresponding attainment assertion of Prop. 3.2 and Eq. (2.10)].

Proposition 4.2: (Minimax Theorem II) Let Assumption 4.1 hold, and assume that 0 lies in the relative interior of the effective domain of the function p of Eq. (2.7) and that $p(0) > -\infty$. Then

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

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.

The preceding two minimax theorems indicate that, aside from convexity and semicontinuity assumptions, the properties of the function p around $u = 0$ are critical to guarantee the minimax equality. Here is an illustrative example:

Example 4.1:

Let

$$X = \left\{x \in \Re^2 \mid x \geq 0\right\}, \qquad Z = \{z \in \Re \mid z \geq 0\},$$

and let

$$\phi(x,z) = e^{-\sqrt{x_1 x_2}} + z x_1,$$

which can be shown to satisfy the convexity/concavity and closedness assumptions of Prop. 4.1. [For every $z \in Z$, the Hessian matrix of the function $\phi(\cdot,z)$ is positive definite within $\mathrm{int}(X)$, so it can be seen that $\phi(\cdot,z)$ is convex over $\mathrm{int}(X)$. Since $\phi(\cdot,z)$ is continuous, it is also convex over X.] For all $z \geq 0$, we have

$$\inf_{x\geq 0}\phi(x,z) = \inf_{x\geq 0}\left\{e^{-\sqrt{x_1 x_2}} + z x_1\right\} = 0,$$

since the expression in braces is nonnegative for $x \geq 0$ and can approach zero by taking $x_1 \to 0$ and $x_1 x_2 \to \infty$. Hence,

$$\sup_{z\geq 0}\inf_{x\geq 0}\phi(x,z) = 0.$$

We also have for all $x \geq 0$,

$$\sup_{z\geq 0}\phi(x,z) = \sup_{z\geq 0}\left\{e^{-\sqrt{x_1 x_2}} + z x_1\right\} = \begin{cases} 1 & \text{if } x_1 = 0, \\ \infty & \text{if } x_1 > 0. \end{cases}$$

Hence,

$$\inf_{x\geq 0}\sup_{z\geq 0}\phi(x,z) = 1,$$

so

$$\inf_{x\geq 0}\sup_{z\geq 0}\phi(x,z) > \sup_{z\geq 0}\inf_{x\geq 0}\phi(x,z).$$

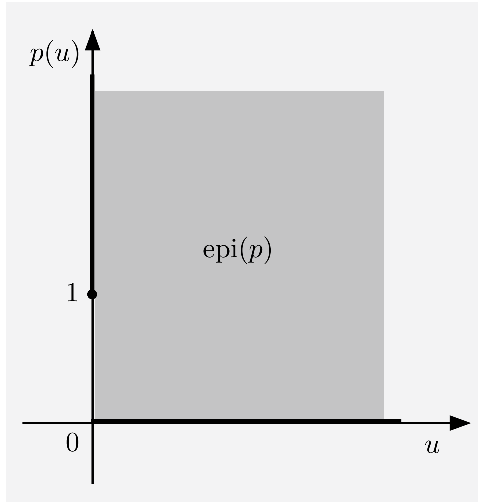

Figure 4.1. The function p for Example 4.1:

$$p(u) = \begin{cases} \infty & \text{if } u < 0, \\ 1 & \text{if } u = 0, \\ 0 & \text{if } u > 0. \end{cases}$$

Here p is not lower semicontinuous at 0, and the minimax equality does not hold.

Here, the function p is given by

$$p(u) = \inf_{x \geq 0} \sup_{z \geq 0} \left\{ e^{-\sqrt{x_1 x_2}} + z(x_1 - u) \right\} = \begin{cases} \infty & \text{if } u < 0, \\ 1 & \text{if } u = 0, \\ 0 & \text{if } u > 0, \end{cases}$$

(cf. Fig. 4.1). Thus, it can be seen that even though $p(0)$ is finite, p is not lower semicontinuous at 0. As a result the assumptions of Props. 4.1 are violated, and the minimax equality does not hold.

The following example illustrates how the minimax equality may hold, while the supremum over $z \in Z$ is not attained because the relative interior assumption of Prop. 4.2 is not satisfied.

Example 4.2:

Let

$$X = \Re, \qquad Z = \{z \in \Re \mid z \geq 0\},$$

and let

$$\phi(x, z) = x + zx^2,$$

which satisfy the convexity/concavity and closedness assumptions of Prop. 4.2. For all $z \geq 0$, we have

$$\inf_{x \in \Re} \phi(x, z) = \inf_{x \in \Re} \left\{ x + zx^2 \right\} = \begin{cases} -1/(4z) & \text{if } z > 0, \\ -\infty & \text{if } z = 0. \end{cases}$$

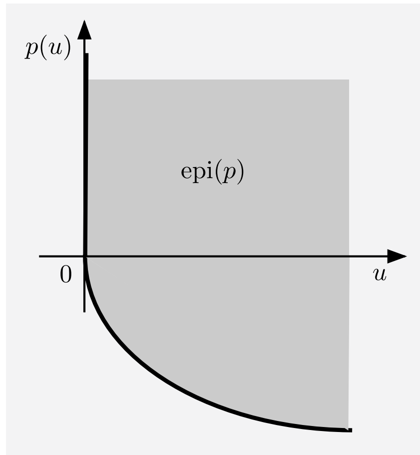

Figure 4.2. The function p for Example 4.2:

$$p(u) = \begin{cases} -\sqrt{u} & \text{if } u \geq 0, \\ \infty & \text{if } u < 0. \end{cases}$$

Here p is lower semicontinuous at 0 and the minimax equality holds. However, 0 is not a relative interior point of dom(p) and the supremum of $\inf_{x \in \Re} \phi(x, z)$, which is

$$\inf_{x \in \Re} \phi(x, z) = \begin{cases} -1/(4z) & \text{if } z > 0, \\ -\infty & \text{if } z = 0, \end{cases}$$

is not attained.

Hence,

$$\sup_{z \geq 0} \inf_{x \in \Re} \phi(x, z) = 0.$$

We also have for all $x \in \Re$,

$$\sup_{z \geq 0} \phi(x, z) = \sup_{z \geq 0} \left\{x + zx^2\right\} = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0. \end{cases}$$

Hence,

$$\inf_{x \in \Re} \sup_{z \geq 0} \phi(x, z) = 0,$$

and the minimax equality holds. However, the problem

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

does not have an optimal solution. Here we have

$$F(x, u) = \sup_{z \geq 0} \{x + zx^2 - uz\} = \begin{cases} x & \text{if } x^2 \leq u, \\ \infty & \text{if } x^2 > u, \end{cases}$$

and

$$p(u) = \inf_{x \in \Re} \sup_{z \geq 0} F(x, u) = \begin{cases} -\sqrt{u} & \text{if } u \geq 0, \\ \infty & \text{if } u < 0, \end{cases}$$

(cf. Fig. 4.2). It can be seen that 0 is not a relative interior point of the effective domain of p, thus violating the relative interior assumption of Prop. 4.2.

Saddle Point Theorems

We will now use the two minimax theorems (Props. 4.1 and 4.2) to obtain more specific conditions for the validity of the minimax equality and the existence of saddle points. The preceding analysis has underscored the importance of the function

$$p(u) = \inf_{x \in \Re^n} F(x, u), \tag{4.1}$$

where

$$F(x, u) = \begin{cases} \sup_{z \in Z} \{\phi(x, z) - u'z\} & \text{if } x \in X, \\ \infty & \text{if } x \notin X. \end{cases} \tag{4.2}$$

It suggests a two-step process to ascertain the validity of the minimax equality and the existence of a saddle point:

(1) Show that p is closed and convex, thereby showing that the minimax equality holds by using the first minimax theorem (Prop. 4.1).

(2) Verify that the infimum of $\sup_{z \in Z} \phi(x, z)$ over $x \in X$, and the supremum of $\inf_{x \in X} \phi(x, z)$ over $z \in Z$ are attained, thereby showing that the set of saddle points is nonempty (cf. Prop. 7.16).

Step (1) requires two types of assumptions:

(a) The convexity/concavity and closedness Assumption 4.1. Then, the MC/MC framework applies (cf. Prop. 3.1). In addition, F is convex and closed (being the pointwise supremum over $z \in Z$ of closed convex functions), which is important in ascertaining that p is closed.

(b) Conditions that guarantee that the partial minimization in the definition of p preserves closedness.

Step (2) requires that either Weierstrass' Theorem can be applied, or else that some other suitable condition for existence of an optimal solution is satisfied. Fortunately, conditions that guarantee that the partial minimization in the definition of F preserves closedness as in (b), also guarantee the existence of corresponding of optimal solutions.

As an example of this line of analysis, we obtain the following classical result.

Proposition 4.3: (Classical Saddle Point Theorem) Let Assumption 4.1 hold, and assume further that X and Z and compact. Then the set of saddle points of ϕ is nonempty and compact.

Proof: From Prop. 7.6, we see that the MC/MC framework applies and the function p of Eq. (4.1) is convex, and the function F of Eq. (4.2) is closed by Prop. 7.8. Using the compactness of X and Z, it follows that F is real-valued over $X \times \Re^m$, and that p is also real-valued and hence continuous. Hence, the minimax equality holds by Prop. 4.1.

Finally, the function $\sup_{z\in Z} \phi(x,z)$ is equal to $F(x,0)$, so it is closed, and the set of its minima over $x \in X$ is nonempty and compact by Weierstrass' Theorem. Similarly the set of maxima of the function $\inf_{x\in X} \phi(x,z)$ over $z \in Z$ is nonempty and compact. It follows that the set of saddle points is nonempty and compact. **Q.E.D.**

We will now derive alternative and more general saddle point theorems, using a similar line of analysis. To formulate the corresponding results, we consider the functions $t : \Re^n \mapsto (-\infty,\infty]$ and $r : \Re^m \mapsto (-\infty,\infty]$ given by

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

and

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

Note that by Assumption 4.1, t is closed and convex, being the supremum of closed and convex functions. Furthermore, since $t(x) > -\infty$ for all x, we have

$$t \text{ is proper if and only if } \inf_{x\in X} \sup_{z\in Z} \phi(x,z) < \infty.$$

Similarly, the function r is closed and convex, and

$$r \text{ is proper if and only if } -\infty < \sup_{z\in Z} \inf_{x\in X} \phi(x,z).$$

The next two propositions provide alternative conditions for the minimax equality to hold. These propositions are subsequently used to prove results about nonemptiness and compactness of the set of saddle points.

Proposition 4.4: Let Assumption 4.1 hold, and assume further that t is proper and that the level sets $\{x \mid t(x) \le \gamma\}$, $\gamma \in \Re$, are compact. Then

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

and the infimum over X in the right-hand side above is attained at a set of points that is nonempty and compact.

Proof: The function p is defined by partial minimization of the function F of Eq. (4.2), i.e.,

$$p(u) = \inf_{x \in \Re^n} F(x,u)$$

[cf. Eq. (4.1)]. We note that

$$t(x) = F(x,0),$$

so F is proper, since t is proper and F is closed [by the Recession Cone Theorem (Prop. 7.3), $\mathrm{epi}(F)$ contains a vertical line if and only if $\mathrm{epi}(t)$ contains a vertical line]. Furthermore, the compactness assumption on the level sets of t can be translated to the compactness assumption of Prop. 7.7 (with 0 playing the role of the vector $\overline{x}$). It follows from the result of that proposition that p is closed and proper, and that $p(0)$ is finite. By the Minimax Theorem I (Prop. 4.1), it follows that the minimax equality holds. Finally, the infimum over X in the right-hand side of the minimax equality is attained at the set of minima of t, which is nonempty and compact since t is proper and has compact level sets. **Q.E.D.**

Example 4.3:

Let us show that

$$\min_{\|x\| \le 1} \max_{z \in S+C} x'z = \max_{z \in S+C} \min_{\|x\| \le 1} x'z,$$

where S is a subspace, and C is a nonempty, convex, and compact subset of $\Re^n$. By defining

$$X = \big\{x \mid \|x\| \le 1\big\}, \qquad Z = S + C,$$

$$\phi(x,z) = x'z, \qquad \forall\ (x,z) \in X \times Z,$$

we see that Assumption 4.1 is satisfied, so we can apply Prop. 4.4. We have

$$\begin{aligned} t(x) &= \begin{cases} \sup_{z \in S+C} x'z & \text{if } \|x\| \le 1, \\ \infty & \text{otherwise,} \end{cases} \\ &= \begin{cases} \sup_{z \in C} x'z & \text{if } x \in S^\perp,\ \|x\| \le 1, \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Since $\sup_{z \in C} x'z$, viewed as a function of x over $\Re^n$, is continuous, the level sets of t are compact. It follows from Prop. 4.4 that the minimax equality holds. It also turns out that a saddle point exists. We will show this shortly, after we develop some additional machinery.

Proposition 4.5: Let Assumption 4.1 hold, and assume further that t is proper, and that the recession cone and the constancy space of t are equal. Then

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

and the infimum over X in the right-hand side above is attained.

Proof: The proof is similar to the one of Prop. 4.4. We use Prop. 7.9 in place of Prop. 7.7. **Q.E.D.**

By combining the preceding two propositions, we obtain conditions for existence of a saddle point.

Proposition 4.6: Let Assumption 4.1 hold, and assume that either t is proper or r is proper.

(a) If the level sets $\{x \mid t(x) \leq \gamma\}$ and $\{z \mid r(z) \leq \gamma\}$, $\gamma \in \Re$, of t and r are compact, the set of saddle points of ϕ is nonempty and compact.

(b) If the recession cones of t and r are equal to the constancy spaces of t and r, respectively, the set of saddle points of ϕ is nonempty.

Proof: We assume that t is proper. If instead r is proper, we reverse the roles of x and z.

(a) From Prop. 4.4, it follows that the minimax equality holds, and that the infimum over X of $\sup_{z \in Z} \phi(x, z)$ is finite and is attained at a nonempty and compact set. Therefore,

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

and we can reverse the roles of x and z, and apply Prop. 4.4 again to show that the supremum over Z of $\inf_{x \in X} \phi(x, z)$ is attained at a nonempty and compact set.

(b) The proof is similar to the one of part (a), except that we use Prop. 4.5 instead of Prop. 4.4. **Q.E.D.**

The following example illustrates Props. 4.4 and 4.6(a).

Example 4.4: (continued)

Consider the case where

$$\phi(x, z) = \tfrac{1}{2} x' Q x + x' z - \tfrac{1}{2} z' R z, \qquad X = Z = \Re^n,$$

and Q and R are symmetric positive semidefinite matrices. Then it can be seen that

$$\sup_{z \in \Re^n} \phi(x, z) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0, \end{cases} \qquad \inf_{x \in \Re^n} \phi(x, z) = \begin{cases} 0 & \text{if } z = 0, \\ -\infty & \text{if } z \neq 0, \end{cases}$$

and

$$X^* = Z^* = \{0\}.$$

Since $\inf_x \sup_z \phi(x, z) = \sup_z \inf_x \phi(x, z) = 0$, it follows that $(0, 0)$ is the unique saddle point.

Assume now that Q is not positive semidefinite, but R is. Then it can be seen that

$$\sup_{z\in\Re^n} \phi(x,z) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0, \end{cases} \qquad \inf_{x\in\Re^n} \phi(x,z) = -\infty, \quad \forall\ z \in \Re^n.$$

Hence $0 = \inf_x \sup_z \phi(x,z) > \sup_z \inf_x \phi(x,z) = -\infty$, so there are no saddle points, even though the sets X^* and Z^* are nonempty ($X^* = \{0\}$ and $Z^* = \Re$). Then,

$$t(x) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0, \end{cases}$$

so t has compact level sets. Also $r = t$, so r has compact level sets, and Prop. 4.6(a) applies. Here $(0,0)$ is the unique saddle point, consistently with the conclusions of the proposition.

Example 4.2: (continued)

To illustrate the difference between Props. 4.4 and 4.6(a), let

$$X = \Re, \qquad Z = \{z \in \Re \mid z \geq 0\}, \qquad \phi(x,z) = x + zx^2.$$

A straightforward calculation yields

$$t(x) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0, \end{cases} \qquad r(z) = \begin{cases} \frac{1}{4z} & \text{if } z > 0, \\ \infty & \text{if } z \leq 0. \end{cases}$$

Here the assumptions of Prop. 4.4 are satisfied, and the minimax equality holds, but the assumptions of Prop. 4.6(a) are violated, and there is no saddle point.

Example 4.3: (continued)

Let

$$X = \big\{x \mid \|x\| \leq 1\big\}, \qquad Z = S + C, \qquad \phi(x,z) = x'z,$$

where S is a subspace, and C is a nonempty, convex, and compact subset of $\Re^n$. We have

$$\begin{aligned} t(x) &= \begin{cases} \sup_{z\in S+C} x'z & \text{if } \|x\| \leq 1, \\ \infty & \text{otherwise}, \end{cases} \\ &= \begin{cases} \sup_{z\in C} x'z & \text{if } x \in S^\perp,\ \|x\| \leq 1, \\ \infty & \text{otherwise}, \end{cases} \end{aligned}$$

$$\begin{aligned} r(z) &= \begin{cases} \sup_{\|x\|\leq 1} -x'z & \text{if } z \in S + C, \\ \infty & \text{otherwise}, \end{cases} \\ &= \begin{cases} \|z\| & \text{if } z \in S + C, \\ \infty & \text{otherwise}. \end{cases} \end{aligned}$$

Since the level sets of both t and r are compact, it follows from Prop. 4.6(a) that the minimax equality holds, and that the set of saddle points is nonempty and compact.

The compactness of the level sets of t can be guaranteed by simpler sufficient conditions. In particular, under Assumption 4.1, the level sets $\{x \mid t(x) \leq \gamma\}$ are compact if any one of the following two conditions holds:

(1) The set X is compact [since $\{x \mid t(x) \leq \gamma\}$ is closed, by the closedness of t, and is contained in X].

(2) For some $\overline{z} \in Z$, $\overline{\gamma} \in \Re$, the set

$$\{x \in X \mid \phi(x, \overline{z}) \leq \overline{\gamma}\}$$

is nonempty and compact [since then all sets $\{x \in X \mid \phi(x, \overline{z}) \leq \gamma\}$ are compact, and a nonempty level set $\{x \mid t(x) \leq \gamma\}$ is contained in $\{x \in X \mid \phi(x, \overline{z}) \leq \gamma\}$].

Furthermore, any one of above two conditions also guarantees that r is proper; for example under condition (2), the infimum over $x \in X$ in the relation

$$r(\overline{z}) = -\inf_{x \in X} \phi(x, \overline{z})$$

is attained by Weierstrass' Theorem, so that $r(\overline{z}) < \infty$.

By a symmetric argument, we also see that, under Assumption 4.1, the level sets of r are compact under any one of the following two conditions:

(1) The set Z is compact.

(2) For some $\overline{x} \in X$, $\overline{\gamma} \in \Re$, the set

$$\{z \in Z \mid \phi(\overline{x}, z) \geq \overline{\gamma}\}$$

is nonempty and compact.

Furthermore, any one of above two conditions also guarantees that t is proper.

Thus, by combining the preceding discussion and Prop. 4.6(a), we obtain the following result, which generalizes the classical saddle point theorem (Prop. 4.3), and provides sufficient conditions for the existence of a saddle point.

Proposition 4.7: (Saddle Point Theorem) Let Assumption 4.1 hold. The set of saddle points of ϕ is nonempty and compact under any one of the following conditions:

(1) X and Z are compact.

(2) Z is compact, and for some $\overline{z} \in Z$, $\overline{\gamma} \in \Re$, the level set

$$\big\{x \in X \mid \phi(x, \overline{z}) \leq \overline{\gamma}\big\}$$

is nonempty and compact.

(3) X is compact, and for some $\overline{x} \in X$, $\overline{\gamma} \in \Re$, the level set

$$\big\{z \in Z \mid \phi(\overline{x}, z) \geq \overline{\gamma}\big\}$$

is nonempty and compact.

(4) For some $\overline{x} \in X$, $\overline{z} \in Z$, $\overline{\gamma} \in \Re$, the level sets

$$\big\{x \in X \mid \phi(x, \overline{z}) \leq \overline{\gamma}\big\}, \qquad \big\{z \in Z \mid \phi(\overline{x}, z) \geq \overline{\gamma}\big\},$$

are nonempty and compact.

Proof: From the discussion preceding the proposition, it is seen that, under Assumption 4.1, t and r are proper, and the level sets of t and r are compact. The result follows from Prop. 4.6(a). **Q.E.D.**

A Nonlinear Version of Farkas' Lemma

We will now use the MC/MC duality results of the preceding subsection to prove a nonlinear version of Farkas' Lemma, which among others, captures the essence of convex programming duality. The lemma involves a nonempty convex set $X \subset \Re^n$, and functions $f : X \mapsto \Re$ and $g_j : X \mapsto \Re$, $j = 1, \ldots, r$. We denote $g(x) = \big(g_1(x), \ldots, g_r(x)\big)'$, and assume the following.

Assumption 4.2: The functions f and g_j, $j = 1, \ldots, r$, are convex, and

$$f(x) \geq 0, \qquad \forall\ x \in X \text{ with } g(x) \leq 0.$$

Proposition 4.8: (Nonlinear Farkas' Lemma) Let Assumption 4.2 hold and let Q^* be the subset of $\Re^r$ given by

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

Then:

(a) Q^* is nonempty and compact if and only if there exists a vector $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all $j = 1, \ldots, r$.

(b) Q^* is nonempty if the functions g_j, $j = 1, \ldots, r$, are affine and there exists a vector $\overline{x} \in \mathrm{ri}(X)$ such that $g(\overline{x}) \leq 0$.

Proof: (a) Assume that there exists a vector $\overline{x} \in X$ such that $g(\overline{x}) < 0$. We will apply the MC/MC Theorem II (Props. 3.2 and 3.3) to the subset of $\Re^{r+1}$ given by

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

(cf. Fig. 4.3). 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,

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

satisfies $-\infty < w^*$.

(ii) The set

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

is convex. (Note here that $\overline{M} = M + \{(0, w) \mid w \geq 0\} = M$.)

(iii) The condition that there exists a vector $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all j is equivalent to the set

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

containing the origin in its interior.

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

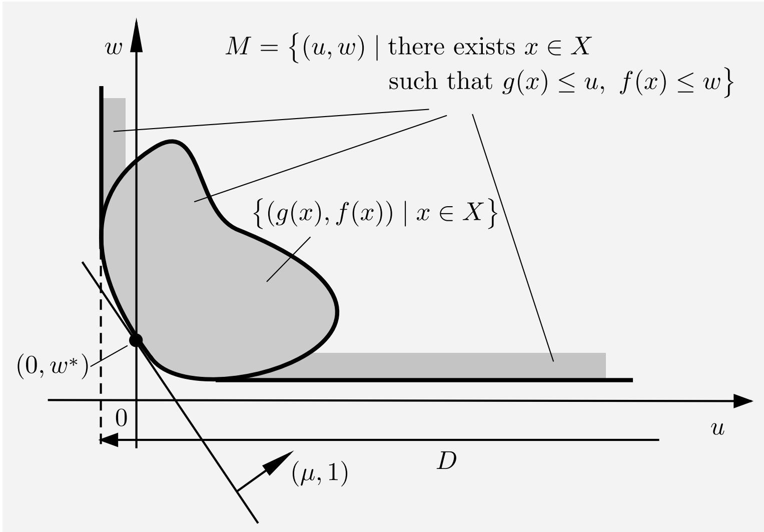

Figure 4.3. Illustration of the sets used in the MC/MC framework of the proof of Prop. 4.8(a). We have

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

and

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

The existence of $\overline{x} \in X$ such that $g(\overline{x}) < 0$ is equivalent to $0 \in \text{int}(D)$.

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

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

If $g(\overline{x}) < 0$ for some $\overline{x} \in X$, then since D contains the set $g(\overline{x}) + \{u \mid u \ge 0\}$, we have $0 \in \text{int}(D)$. Conversely, if $0 \in \text{int}(D)$, there exists $\epsilon > 0$ such that the vector $(-\epsilon, \ldots, -\epsilon)$ belongs to D, so that $g_j(\overline{x}) \le -\epsilon$ for some $\overline{x} \in X$ and all j.

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 $x \in X$ and $\tilde{x} \in X$, 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

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

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

By convexity of X, we have $\alpha x + (1-\alpha)\tilde{x} \in X$ 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 MC/MC 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

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

From the definition of Q^*, we have

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

so Q^* and $\tilde{Q}$ are level sets of the proper convex function $-q$, which is closed by Prop. 2.1. Therefore, if one of them is compact, so is the other. Since $Q^* \supset \tilde{Q}$, and $\tilde{Q}$ is nonempty and compact, so is Q^*.

(b) Let the constraints $g_j(x) \leq 0$ be written as

$$Ax - b \leq 0,$$

where A is an $r \times n$ matrix and b is a vector in $\Re^n$. We apply Prop. 3.6, with P being the nonpositive orthant and the set $\overline{M}$ defined by

$$\overline{M} = \big\{(u,w) \mid \text{for some } (x,w) \in \text{epi}(\tilde{f}),\ Ax - b - u \leq 0\big\},$$

where

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in X, \\ \infty & \text{if } x \notin X. \end{cases}$$

Since by assumption, $f(x) \geq 0$ for all $x \in X$ with $Ax - b \leq 0$, the optimal min common value satisfies

$$w^* = \inf_{Ax - b \leq 0} \tilde{f}(x) \geq 0.$$

By Prop. 3.6 and the discussion following its proof, there exists $\mu \geq 0$ such that

$$q^* = q(\mu) = \inf_{x\in\Re^n} \big\{\tilde{f}(x) + \mu'(Ax - b)\big\}.$$

Since $q^* = w^* \geq 0$, it follows that $\tilde{f}(x) + \mu'(Ax - b) \geq 0$ for all $x \in \Re^n$, or $f(x) + \mu'(Ax - b) \geq 0$ for all $x \in X$. **Q.E.D.**

By selecting f and g_j to be linear, and X to be the entire space in the Nonlinear Farkas' Lemma, we obtain the following classical result.

Proposition 4.9: (Linear Farkas' Lemma) Let A be an $m \times n$ matrix and c be a vector in $\Re^m$.

(a) The system $Ay = c$, $y \geq 0$ has a solution if and only if

$$A'x \leq 0 \qquad \Rightarrow \qquad c'x \leq 0. \tag{4.3}$$

(b) The system $Ay \geq c$ has a solution if and only if

$$A'x = 0,\ x \geq 0 \qquad \Rightarrow \qquad c'x \leq 0. \tag{4.4}$$

Proof: (a) Using part (b) of the Nonlinear Farkas' Lemma with $f(x) = -c'x$, $g(x) = A'x$, and $X = \Re^m$, we see that the relation (4.3) implies the existence of $\mu \geq 0$ such that

$$-c'x + \mu' A'x \geq 0, \qquad \forall\ x \in \Re^m,$$

or equivalently $(A\mu - c)'x \geq 0$ for all $x \in \Re^m$, or $A\mu = c$.

(b) This part follows by writing the system $Ay \geq c$ in the equivalent form

$$Ay^+ - Ay^- - z = c, \qquad y^+ \geq 0,\ y^- \geq 0,\ z \geq 0,$$

and by applying part (a). **Q.E.D.**

Convex Programming Duality

Consider the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \ \ x \in X, \quad g(x) \leq 0, \end{aligned} \tag{4.5}$$

where X is a convex set, $g(x) = \big(g_1(x), \ldots, g_r(x)\big)'$, and $f : X \mapsto \Re$ and $g_j : X \mapsto \Re$, $j = 1, \ldots, r$, are convex functions. We refer to this as the *primal problem*, and we will use the minimax theory of Section 3.3, and the Nonlinear Farkas' Lemma to derive the connection of this problem with an associated dual problem. We assume that the problem is feasible, i.e., that the optimal value

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

satisfies $f^* < \infty$.

Let us consider the Lagrangian function defined by

$$L(x, \mu) = f(x) + \mu' g(x), \qquad x \in X, \ \mu \geq 0,$$

and the minimax problem involving $L(x, \mu)$, and minimization over $x \in X$ and maximization over $\mu \geq 0$. Consider also the function $t : X \mapsto (-\infty, \infty]$ given by

$$t(x) = \sup_{\mu \geq 0} L(x, \mu) = \begin{cases} f(x) & \text{if } g(x) \leq 0, \ x \in X, \\ \infty & \text{otherwise,} \end{cases} \tag{4.6}$$

and the function q given by

$$q(\mu) = \inf_{x \in X} L(x, \mu), \qquad \mu \geq 0. \tag{4.7}$$

We refer to q as the *dual function* and to the problem

$$\begin{aligned} &\text{maximize} \quad q(\mu) \\ &\text{subject to} \quad \mu \in \Re^r, \ \mu \geq 0, \end{aligned} \tag{4.8}$$

as the *dual problem*. The dual optimal value is

$$q^* = \sup_{\mu \geq 0} q(\mu),$$

and the generic relation

$$\sup_{\mu \geq 0} \inf_{x \in X} L(x, \mu) \leq \inf_{x \in X} \sup_{\mu \geq 0} L(x, \mu)$$

is equivalent to the weak duality relation $q^* \leq f^*$ [cf. Eqs. (4.6), (4.7)].

Proposition 4.10: (Convex Programming Duality I) Assume that the convex functions f and g_j are closed, and the function t of Eq. (4.6) has compact level sets. Then $f^* = q^*$ and the set of optimal solutions of the primal problem is nonempty and compact.

Proof: We have

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

where inf and sup can be interchanged since Prop. 4.4 applies because f and g_j are closed, and t is proper (since the problem is feasible) and has compact level sets. Also, the set of optimal solutions of the primal problem is the set of minima of t, and is nonempty and compact since t has compact level sets. **Q.E.D.**

Note that the compactness assumption of Prop. 4.10 is satisfied if either X is compact or if X is closed and f has compact level sets. We now turn to the question whether the dual problem has an optimal solution. Assume that the optimal value f^* is finite. Then, we have

$$0 \le f(x) - f^*, \qquad \forall\, x \in X \text{ with } g(x) \le 0,$$

so by replacing $f(x)$ by $f(x) - f^*$ and by applying the Nonlinear Farkas' Lemma [assuming that the assumption of part (a) or part (b) of the lemma holds], we see that there exists a vector $\mu^* \ge 0$ such that

$$f^* \le f(x) + \mu^{*\prime} g(x), \qquad \forall\, x \in X.$$

It follows that

$$f^* \le \inf_{x \in X} \bigl\{ f(x) + \mu^{*\prime} g(x) \bigr\} \le \inf_{x \in X,\, g(x) \le 0} f(x) = f^*.$$

Thus equality holds throughout in the above relation, and from the definition (4.7) of the dual function, we have

$$f^* = \inf_{x \in X} \bigl\{ f(x) + \mu^{*\prime} g(x) \bigr\} = q(\mu^*),$$

or using also the weak duality relation $q^* \le f^*$,

$$f^* = q(\mu^*) = q^*. \tag{4.9}$$

Thus μ^* is an optimal solution of the dual problem, and its existence is guaranteed if one of the conditions of the Nonlinear Farkas' Lemma is satisfied. We state this conclusion as a proposition:

Proposition 4.11: (Convex Programming Duality II) Assume that f^* is finite, and that one of the following two conditions holds:

(1) There exists $\overline{x} \in X$ such that $g(\overline{x}) < 0$.

(2) The functions g_j are affine, and there exists $\overline{x} \in \mathrm{ri}(X)$ such that $g(\overline{x}) \le 0$.

Then $q^* = f^*$ and the set of optimal solutions of the dual problem is nonempty. Under condition (1) this set is also compact.

Note that we have used two distinct approaches for establishing $q^* = f^*$. The first is based on minimax theory, and also guarantees that there exists an optimal primal solution (even if there may be no dual optimal solution; cf. Prop. 4.10). The second is based on the Nonlinear Farkas' Lemma, and guarantees that the dual problem has an optimal solution (even if there may be no primal optimal solution; cf. Prop. 4.11).

Theorems of the Alternative

Theorems of the alternative are important tools in optimization, which address the feasibility (possibly strict) of affine inequalities. Interestingly, theorems of the alternative can be viewed as special cases of the MC/MC Theorem II, and the supplementary proposition that asserts the equivalence of compactness of the set of max crossing solutions with the origin being an interior point of the "domain" of the set M in the MC/MC framework (cf. Props. 3.2, 3.3). This connection is demonstrated in the proofs of the following two classical theorems.

Proposition 4.12: (Gordan's Theorem) Let A be an $m \times n$ matrix and b be a vector in $\Re^m$. The following are equivalent:

(i) There exists a vector $x \in \Re^n$ such that

$$Ax < b.$$

(ii) For every vector $\mu \in \Re^m$,

$$\mu \geq 0, \quad A'\mu = 0, \quad \mu'b \leq 0 \qquad \Rightarrow \qquad \mu = 0.$$

(iii) Any polyhedral set of the form

$$\{\mu \mid A'\mu = c,\, \mu'b \leq d,\, \mu \geq 0\}, \tag{4.10}$$

where $c \in \Re^n$ and $d \in \Re$, is compact.

Proof: We will show that (i) and (ii) are equivalent, and then that (ii) and (iii) are equivalent. The equivalence of (i) and (ii) is geometrically evident, once the proper MC/MC framework is considered (see Fig. 4.4). For completeness we provide the details. Consider the set

$$M = \big\{(u, w) \mid w \geq 0,\, Ax - b \leq u \text{ for some } x \in \Re^n\big\},$$

its projection on the x axis

$$D = \{u \mid Ax - b \leq u \text{ for some } x \in \Re^n\},$$

and the corresponding MC/MC framework. Let w^* and q^* be the min common and max crossing values, respectively. Clearly, the system $Ax \leq b$ has a solution if and only if $w^* = 0$. Also, if (ii) holds, we have

$$A'\mu = 0, \quad \mu \geq 0 \qquad \Rightarrow \qquad b'\mu \geq 0,$$

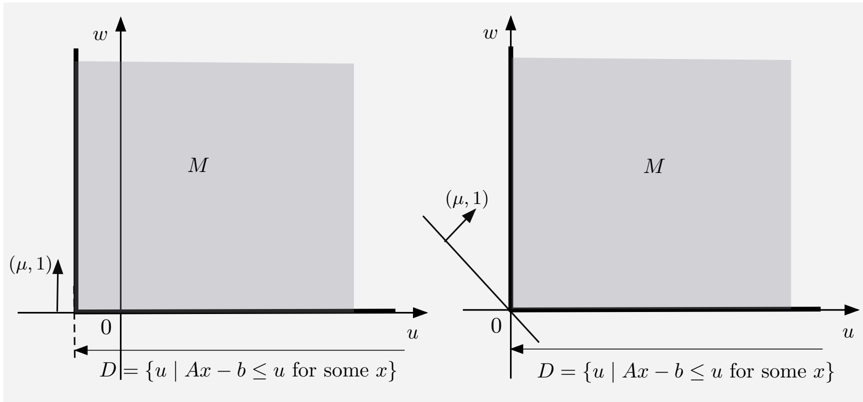

Figure 4.4. MC/MC framework for showing the equivalence of conditions (i) and (ii) of Gordan's Theorem. We consider the set

$$M = \Big\{(u, w) \mid w \ge 0,\ Ax - b \le u \text{ for some } x \in \Re^n\Big\},$$

its projection on the x axis

$$D = \{u \mid Ax - b \le u \text{ for some } x \in \Re^n\},$$

and the corresponding MC/MC framework. Condition (i) of Gordan's Theorem is equivalent to $0 \in \text{int}(D)$, and is also equivalent to 0 being the unique solution of the max crossing problem (see the figure on the left). It can be seen that the latter condition can be written as

$$\mu \ge 0, \quad 0 \le \mu'(Ax - b) + w, \quad \forall\, x \in \Re^n,\ w \ge 0 \qquad \Rightarrow \qquad \mu = 0,$$

which is equivalent to condition (ii). In the figure on the right both conditions (i) and (ii) are violated.

which by the linear version of Farkas' Lemma [Prop. 4.9(a)], implies that the system $Ax \le b$ has a solution. In conclusion, both (i) and (ii) imply that the system $Ax \le b$ has a solution, which is in turn equivalent to $w^* = q^* = 0$. Thus in proving the equivalence of (i) and (ii), we may assume that the system $Ax \le b$ has a solution and $w^* = q^* = 0$.

The cost function of the max crossing problem is

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

Since M contains vectors (u, w) with arbitrarily large components [for each $(u, w) \in M$, we have $(\overline{u}, \overline{w}) \in M$ for all $\overline{u} \ge u$ and $\overline{w} \ge w$], it follows that $q(\mu) = -\infty$ for all μ that are not in the nonnegative orthant, and we have

$$q(\mu) = \begin{cases} \inf_{Ax-b\le u} \mu' u & \text{if } \mu \ge 0, \\ -\infty & \text{otherwise,} \end{cases}$$

or equivalently,

$$q(\mu) = \begin{cases} \inf_{x\in\Re^n} \mu'(Ax - b) & \text{if } \mu \ge 0, \\ -\infty & \text{otherwise.} \end{cases} \tag{4.11}$$

We now note that condition (i) is equivalent to 0 being an interior point of D (to see this note that if $\overline{x}$ satisfies $A\overline{x} - b < 0$, then the set $\{u \mid A\overline{x} - b \leq u\}$ contains 0 in its interior and is contained in the set D). By Prop. 3.3, it follows that (i) is equivalent to $w^* = q^* = 0$ and the set of optimal solutions of the max crossing problem being nonempty and compact. Thus, using the form (4.11) of q, we see that (i) is equivalent to $\mu = 0$ being the only $\mu \geq 0$ satisfying $q(\mu) \geq 0$ or equivalently, satisfying $A'\mu = 0$ and $\mu' b \leq 0$. It follows that (i) is equivalent to (ii).

To show that (ii) is equivalent to (iii), we note that the recession cone of the set (4.10) is

$$\{\mu \geq 0 \mid A'\mu = 0,\ \mu' b \leq 0\}.$$

Thus (ii) states that the recession cone of the set (4.10) consists of just the origin, which is equivalent to (iii). **Q.E.D.**

There are several theorems of the alternative involving strict inequalities. Among the ones involving linear constraints, the following is the most general.

Proposition 4.13: (Motzkin's Transposition Theorem) Let A and B be $p \times n$ and $q \times n$ matrices, and let $b \in \Re^p$ and $c \in \Re^q$ be vectors. The system

$$Ax < b, \qquad Bx \leq c$$

has a solution if and only if for all $\mu \in \Re^p$ and $\nu \in \Re^q$, with $\mu \geq 0$, $\nu \geq 0$, the following two conditions hold:

$$A'\mu + B'\nu = 0 \qquad \Rightarrow \qquad b'\mu + c'\nu \geq 0, \tag{4.12}$$

$$A'\mu + B'\nu = 0,\ \mu \neq 0 \qquad \Rightarrow \qquad b'\mu + c'\nu > 0. \tag{4.13}$$

Proof: Consider the set

$$M = \big\{(u, w) \mid w \geq 0,\ Ax - b \leq u \text{ for some } x \text{ with } Bx \leq c\big\},$$

its projection on the x axis

$$D = \{u \mid Ax - b \leq u \text{ for some } x \text{ with } Bx \leq c\},$$

and the corresponding MC/MC framework. The cost function of the max crossing problem is

$$q(\mu) = \begin{cases} \inf_{(u,w) \in M}\{w + \mu' u\} & \text{if } \mu \geq 0, \\ -\infty & \text{otherwise,} \end{cases}$$

or

$$q(\mu)=\begin{cases}\inf_{Ax-b\le u,\,Bx\le c}\mu'u & \text{if } \mu\ge 0,\\ -\infty & \text{otherwise.}\end{cases} \tag{4.14}$$

Similar to the proof of Gordan's Theorem (Prop. 4.12), and using the linear version of Farkas' Lemma [Prop. 4.9(b)], we may assume that the system $Ax\le b$, $Bx\le c$ has a solution.

Now the system $Ax<b$, $Bx\le c$ has a solution if and only if 0 is an interior point of D, so by Prop. 3.3, it follows that the system $Ax<b$, $Bx\le c$ has a solution if and only if the set of optimal solutions of the max crossing problem is nonempty and compact. Since $q(0)=0$ and $\sup_{\mu\in\Re^m}q(\mu)=0$, using also Eq. (4.14), we see that the set of dual optimal solutions is nonempty and compact if and only if $q(\mu)<0$ for all $\mu\ge 0$ with $\mu\ne 0$ [if $q(\mu)=0$ for some nonzero $\mu\ge 0$, it can be seen from Eq. (4.14) that we must have $q(\gamma\mu)=0$ for all $\gamma>0$]. We will complete the proof by showing that these conditions hold if and only if conditions (4.12) and (4.13) hold for all $\mu\ge 0$ and $\nu\ge 0$.

We calculate the infimum of the linear program in (x,u) in the right-hand side of Eq. (4.14) by computing the optimal dual value using linear programming duality. In particular, we have for all $\mu\ge 0$, after a straightforward calculation,

$$q(\mu)=\inf_{Ax-b\le u,\,Bx\le c}\mu'u=\sup_{A'\mu+B'\nu=0,\,\nu\ge 0}(-b'\mu-c'\nu).$$

Using this equation, we see that $q(0)=0$ and $q(\mu)<0$ for all $\mu\ge 0$ with $\mu\ne 0$ if and only if conditions (4.12) and (4.13) hold for all $\mu\ge 0$ and $\nu\ge 0$. **Q.E.D.**

Subdifferential Theory

We finally provide a connection between dual optimal solutions in the MC/MC framework and subgradients. In particular, we will show that the subdifferential of a convex function at a point can be identified with the set of max crossing hyperplanes in a suitable MC/MC framework. Through this connection, the basic result on nonemptiness and compactness of the subdifferential will follow as a special case of the MC/MC theory developed in Section 3 (Props. 3.2 and 3.3).

Let $f:\Re^n\mapsto(-\infty,\infty]$ be a proper convex function. We say that a vector $g\in\Re^n$ is a *subgradient* of f at a point $x\in\text{dom}(f)$ if

$$f(z)\ge f(x)+g'(z-x),\qquad \forall\ z\in\Re^n. \tag{4.15}$$

The set of all subgradients of f at x is called the *subdifferential of f at x*, and is denoted by $\partial f(x)$. By convention, $\partial f(x)$ is considered empty for all $x\notin\text{dom}(f)$.

Note that g is a subgradient of f at x if and only if the hyperplane in $\Re^{n+1}$ that has normal $(-g,1)$ and passes through $\big(x,f(x)\big)$ supports the epigraph of f, as shown in the left-hand side of Fig. 4.5. From this

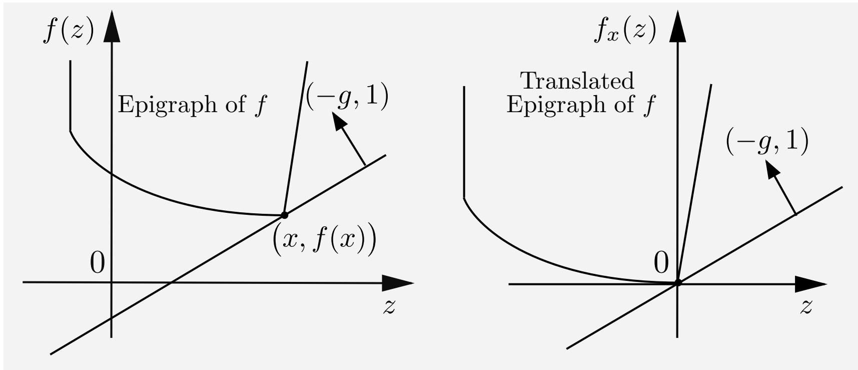

Figure 4.5. Illustration of the subdifferential $\partial f(x)$ of a convex function f (left figure) and its connection with the MC/MC framework. The defining relation (4.15) can be written as

$$f(z) - z'g \geq f(x) - x'g, \qquad \forall\, z \in \Re^n.$$

Thus, g is a subgradient of f at x if and only if the hyperplane in $\Re^{n+1}$ that has normal $(-g, 1)$ and passes through $\big(x, f(x)\big)$ supports the epigraph of f, as shown in the left-hand side figure.

The right-hand side figure shows that $\partial f(x)$ is the set of max crossing solutions in the min MC/MC framework where M is the epigraph of f_x, the x-translation of f.

geometric view, it is evident that there is a strong connection with the MC/MC framework. In particular, for any $x \in \mathrm{dom}(f)$, consider the *x-translation of f*, which is the function f_x, whose epigraph is the epigraph of f translated so that $(x, f(x))$ is moved to the origin of $\Re^{n+1}$:

$$f_x(z) = f(x + z) - f(x), \qquad z \in \Re^n.$$

Then the subdifferential $\partial f(x)$ is the set of all max crossing solutions for the MC/MC framework corresponding to the set

$$M = \mathrm{epi}(f_x) = \mathrm{epi}(f) - \big\{(x, f(x)\big\}, \tag{4.16}$$

as illustrated in Fig. 4.5. Based on this fact, we can use the max common/min crossing theory to obtain results regarding the existence of subgradients, as in the following proposition.

Proposition 4.14: Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function. For every $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$, the subdifferential $\partial f(x)$ has the form

$$\partial f(x) = S^\perp + G,$$

where S is the subspace that is parallel to the affine hull of $\mathrm{dom}(f)$, and G is a nonempty and compact set. Furthermore, $\partial f(x)$ is nonempty and compact if and only if x is in the interior of $\mathrm{dom}(f)$.

Proof: The result follows by applying MC/MC Theorem II (Props. 3.2 and 3.3) to the set M given by Eq. (4.16). **Q.E.D.**

5. NONCONVEX PROBLEMS AND ESTIMATES OF THE DUALITY GAP

In this section we focus on the MC/MC framework in the absence of the convex structure that guarantees strong duality, and we aim to estimate the duality gap. Let us consider the case where M has the form

$$M = \tilde{M} + \big\{(u, 0) \mid u \in C\big\},$$

where $\tilde{M}$ is a compact set and C is a closed convex set, or more generally, where $\mathrm{conv}(M)$ is closed and $q^* > -\infty$. Then, by Prop. 3.4, the duality gap $w^* - q^*$ is equal to $w^* - \overline{w}^*$, where w^* and $\overline{w}^*$ are the min common values corresponding to the sets M and $\mathrm{conv}(M)$. Thus estimates of the duality gap should depend on how much M differs from its convex hull, at least along the vertical axis. A special case where interesting estimates of this type can be obtained arises in separable optimization problems, as we now discuss.

Separable Problems and their Geometry

Suppose that x has m components $x_1, \ldots, x_m$ of dimensions $n_1, \ldots, n_m$, respectively, and the problem has the form

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^m f_i(x_i) \\ &\text{subject to} \quad \sum_{i=1}^m g_i(x_i) \le 0, \quad x_i \in X_i, \quad i = 1, \ldots, m, \end{aligned} \tag{5.1}$$

where $f_i : \Re^{n_i} \mapsto \Re$ and $g_i : \Re^{n_i} \mapsto \Re^r$ are given functions, and X_i are given subsets of $\Re^{n_i}$. Besides linear programs, an important special case is when the functions f_i and g_i are linear, and the sets X_i are finite, in which case we obtain a discrete/integer programming problem.

We call a problem of the form (5.1) *separable*. Its salient feature is that the minimization involved in the calculation of the dual function

$$q(\mu) = \inf_{\substack{x_i \in X_i \\ i=1,\ldots,m}} \left\{ \sum_{i=1}^m \big(f_i(x_i) + \mu' g_i(x_i)\big) \right\} = \sum_{i=1}^m q_i(\mu),$$

is decomposed into the m simpler minimizations

$$q_i(\mu) = \inf_{x_i \in X_i} \big\{ f_i(x_i) + \mu' g_i(x_i) \big\}, \qquad i = 1, \ldots, m.$$

These minimizations are often conveniently done either analytically or computationally, in which case the dual function can be easily evaluated.

When the cost and/or the constraints are not convex, the separable structure is helpful in another, somewhat unexpected way. In particular, in this case *the duality gap turns out to be relatively small and can often be shown to diminish to zero relative to the optimal primal value as the number m of separable terms increases*. As a result, one can often obtain a near-optimal primal solution, starting from a dual-optimal solution. In integer programming problems, this may obviate the need for costly branch-and-bound procedures.

The small duality gap size is a consequence of the structure of the set of achievable constraint-cost pairs

$$S = \big\{ \big(g(x), f(x)\big) \mid x \in X \big\},$$

which plays a central role in the corresponding application of the Nonlinear Farkas Lemma (see Fig. 4.3). In the case of a separable problem, this set can be written as a vector sum of m sets, one for each separable term, i.e.,

$$S = S_1 + \cdots + S_m,$$

where

$$S_i = \big\{ \big(g_i(x_i), f_i(x_i)\big) \mid x_i \in X_i \big\}.$$

Generally, a set that is the vector sum of a large number of possibly nonconvex but roughly similar sets "tends to be convex" in the sense that any vector in its convex hull can be closely approximated by a vector in the set. As a result, the duality gap tends to be relatively small. The analytical substantiation is based on the following theorem, which roughly states that at most $r+1$ out of the m convex sets "contribute to the nonconvexity" of their vector sum, regardless of the value of m. A proof of the Shapley-Folkman Theorem is given is several sources, including [Zho93] and [Ber99].

Proposition 5.1: (Shapley-Folkman Theorem) Let S_i, $i = 1, \ldots, m$, be nonempty subsets of $\Re^{r+1}$, with $m > r+1$, and let $S = S_1 + \cdots + S_m$. Then every vector $s \in \mathrm{conv}(S)$ can be represented as $s = s_1 + \cdots + s_m$, where $s_i \in \mathrm{conv}(S_i)$ for all $i = 1, \ldots, m$, and $s_i \notin S_i$ for at most $r+1$ indices i.

As an illustration, let us use the Shapley-Folkman Theorem to estimate the duality gap in the case of a linear program with integer 0-1 constraints:

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{m} c_i x_i \\ &\text{subject to} \quad \sum_{i=1}^{m} a_{ji} x_i \le b_j, \quad j = 1, \ldots, r, \\ &\qquad\qquad\quad x_i = 0 \text{ or } 1, \quad i = 1, \ldots, m. \end{aligned} \tag{5.2}$$

Let f^* and q^* denote the optimal primal and dual values, respectively. Note that the "relaxed" linear program, where the integer constraints are replaced by $x_i \in [0,1]$, $i = 1, \ldots, m$, has the same dual, so its optimal value is q^*. The set S can be written as

$$S = S_1 + \cdots + S_m - (b, 0),$$

where $b = (b_1, \ldots, b_r)$ and each S_i consists of just two elements corresponding to $x_i = 0, 1$, i.e.,

$$S_i = \big\{(0, \ldots, 0, 0), (a_{1i}, \ldots, a_{ri}, c_i)\big\}.$$

Thus, S consists of a total of 2^m points. A natural idea is to solve the "relaxed" program, and then try to suitably "round" the fractional (i.e., noninteger) components of the relaxed optimal solution to an integer, thereby obtaining a suboptimal solution of the original integer program (5.2).

Let us denote

$$\gamma = \max_{i=1,\ldots,m} |c_i|, \qquad \delta = \max_{i=1,\ldots,m} \delta_i,$$

where

$$\delta_i = \begin{cases} 0 & \text{if } a_{1i}, \ldots, a_{ri} \ge 0, \\ 0 & \text{if } a_{1i}, \ldots, a_{ri} \le 0, \\ \max_{j=1\ldots,r} |a_{ji}| & \text{otherwise.} \end{cases}$$

Note that δ_i is an upper bound to the maximum amount of constraint violation that can result when a fractional variable $x_i \in (0,1)$ is rounded suitably (up or down). The following proposition shows what can be achieved with such a rounding procedure.†

Proposition 5.2: Assume that the relaxed version of the linear/integer problem (5.2) is feasible. Then there exists $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_m)$, with $\overline{x}_i \in \{0,1\}$ for all i, which violates the inequality constraints of problem (5.2) by at most $(r+1)\delta$, and has cost that is at most $q^* + (r+1)\gamma$.

† A slightly stronger bound [$r\gamma$ and $r\delta$ in place of $(r+1)\gamma$ and $(r+1)\delta$, respectively] can be shown with an alternative proof that uses the theory of the simplex method. However, the proof given here generalizes to the case where f_i and g_i are nonlinear (see the subsequent Prop. 5.4).

Proof: We note that the relaxed problem, being a feasible linear program with compact constraint set, has an optimal solution with optimal value q^*. Since

$$\mathrm{conv}(S) = \mathrm{conv}(S_1) + \cdots + \mathrm{conv}(S_m) - (b, 0),$$

we see that $\mathrm{conv}(S)$ is the set of constraint-cost pairs of the relaxed problem, so it contains a vector of the form (u^*, q^*) with $u^* \leq 0$. By the Shapley-Folkman Theorem, there is an index subset I with at most $r+1$ elements such that (u^*, q^*) can be written as

$$u^* = \sum_{i \in I} \overline{u}_i + \sum_{i \notin I} u_i, \qquad q^* = \sum_{i \in I} \overline{w}_i + \sum_{i \notin I} w_i,$$

where $(\overline{u}_i, \overline{w}_i) \in \mathrm{conv}(S_i)$ for $i \in I$ and $(u_i, w_i) \in S_i$ for $i \notin I$. Each pair (u_i, w_i), $i \notin I$ corresponds to an integer component $\overline{x}_i \in \{0, 1\}$. Each pair $(\overline{u}_i, \overline{w}_i)$, $i \in I$ may be replaced/rounded by one of the two elements (u_i, w_i) of S_i, yielding again an integer component $\overline{x}_i \in \{0, 1\}$, with an increase in cost of at most γ and an increase of the level of each inequality constraint of at most δ_i. We thus obtain an integer vector $\overline{x}$ that violates the inequality constraints by at most $(r+1)\delta$, and has cost that is at most $q^* + (r+1)\gamma$. **Q.E.D.**

The proof of the proposition also indicates the rounding mechanism for obtaining the vector $\overline{x}$ of the above proposition. In practice, the simplex method provides a relaxed problem solution with no more than r noninteger components, which are then rounded as indicated in the proof to obtain $\overline{x}$. Note that $\overline{x}$ may not be feasible, and indeed it is possible that the relaxed problem is feasible, while the original integer problem is not (consider for example the constraints $x_1 - x_2 \leq -1/2$ and $x_1 + x_2 \leq 1/2$). On the other hand if for each j, all the constraint coefficients $a_{j1}, \ldots, a_{jm}$ are either 0 or have the same sign, we have $\delta = 0$, and a feasible solution that is $(r+1)\gamma$-optimal can be found. Assuming this condition, let us consider now a sequence of similar problems where the number of inequality constraints r is kept constant, but the dimension m grows to infinity. Assuming that $|f^*| = O(n)$, we see that the rounding error $(r+1)\delta$ (which bounds the duality gap $f^* - q^*$) diminishes in relation to f^*, i.e., its ratio to f^* tends to 0 as $n \to \infty$. In particular, we have

$$\lim_{n \to \infty} \frac{f^* - q^*}{f^*} \to 0.$$

The line of analysis of the preceding proposition can be generalized to the nonlinear problem (5.1), and similar results can be obtained. In particular, it can be shown under assumptions that generalize the condition $\delta = 0$ discussed earlier, that the duality gap satisfies

$$f^* - q^* \leq (r+1) \max_{i=1,\ldots,m} \gamma_i,$$

where for each i, γ_i is a nonnegative scalar that depends on the structure of the functions f_i, g_i, and the set X_i (see the following proposition). This estimate suggests that as $m \to \infty$, the duality gap diminishes relative to f^* as $m \to \infty$.

Estimates of Duality Gap in Separable Problems

We will now generalize the duality gap estimate of Prop. 5.2. We first derive a general duality gap estimate for the case where the set M in the MC/MC framework is a vector sum, and then we specialize to the case of a separable problem.

Proposition 5.3: Consider the MC/MC framework corresponding to a set $M \subset \Re^{n+1}$ given by

$$M = M_1 + \cdots + M_m,$$

where the sets M_i have the form

$$M_i = \tilde{M}_i + \big\{(u,0) \mid u \in C_i\big\}, \qquad i = 1,\ldots,m,$$

and for each i, $\tilde{M}_i$ is a compact set and C_i is convex, and such that $C_1 + \cdots + C_m$ is closed. Let w^* and q^* be the min common and max crossing values, respectively, and assume that they are finite. For $i = 1,\ldots,m$, assume that for each vector $(u_i, w_i) \in \mathrm{conv}(M_i)$, there exists a vector of the form $(\overline{u}_i, \overline{w}_i) \in M_i$, with $u_i - \overline{u}_i \in C_i$, and let

$$\gamma_i = \sup_{(u_i,w_i)\in \mathrm{conv}(M_i)} \ \inf_{\substack{(\overline{u}_i,\overline{w}_i)\in M_i \\ u_i-\overline{u}_i\in C_i}} (\overline{w}_i - w_i).$$

Then

$$w^* - q^* \le (n+1) \max_{i=1,\ldots,m} \gamma_i.$$

Proof: By the result of Prop. 3.4(c), q^* is equal to the optimal value of the MC/MC problem corresponding to $\mathrm{conv}(M)$. Since the sets M_i are assumed compact and $C_1 + \cdots + C_m$ is closed, $\mathrm{conv}(M)$ is closed, and it follows that the vector $(0, q^*)$ belongs to $\mathrm{conv}(M)$. By the Shapley-Folkman Theorem (Prop. 5.1), there is an index subset I with at most $n+1$ elements such that $(0, q^*)$ can be expressed as

$$0 = \sum_{i\in I} u_i + \sum_{i\notin I} \overline{u}_i, \qquad q^* = \sum_{i\in I} w_i + \sum_{i\notin I} \overline{w}_i,$$

where $(u_i, w_i) \in \mathrm{conv}(M_i)$ for $i \in I$ and $(\overline{u}_i, \overline{w}_i) \in M_i$ for $i \notin I$. By the hypothesis, for any $\epsilon > 0$ and $i \in I$, there also exist vectors of the form $(\overline{u}_i, \overline{w}_i) \in M_i$ with $\overline{w}_i \le \gamma_i + \epsilon$ and $u_i - \overline{u}_i \in C_i$. Hence, the vector $(\overline{u}, \overline{w})$, where

$$\overline{u} = \sum_{i\in I} \overline{u}_i + \sum_{i\notin I} \overline{u}_i, \qquad \overline{w} = \sum_{i\in I} \overline{w}_i + \sum_{i\notin I} \overline{w}_i,$$

belongs to M, so that $w^* \leq \overline{w}$. Thus, we have

$$w^* \leq \overline{w} = \sum_{i \in I} \overline{w}_i + \sum_{i \notin I} \overline{w}_i \leq \sum_{i \in I} (w_i + \gamma_i + \epsilon) + \sum_{i \notin I} \overline{w}_i \leq q^* + (n+1) \max_{i=1,\ldots,m} (\gamma_i + \epsilon).$$

Since ϵ can be taken arbitrarily small, the result follows. **Q.E.D.**

Proposition 5.4: Consider the separable problem of Eq. (5.1):

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^m f_i(x_i) \\ &\text{subject to} \quad \sum_{i=1}^m g_i(x_i) \leq 0, \quad x_i \in X_i, \quad i = 1, \ldots, m, \end{aligned}$$

where the sets

$$\big\{\big(g_i(x_i), f_i(x_i)\big) \mid x_i \in X_i\big\}, \qquad i = 1, \ldots, m,$$

are compact. Assume that the primal and dual optimal values q^* and f^* are finite, and that for each i, given any $\tilde{x}_i \in \text{conv}(X_i)$, there exists $x_i \in X_i$ such that

$$g_i(x_i) \leq (\check{\text{cl}}\, g_i)(\tilde{x}_i),$$

where $\check{\text{cl}}\, g_i$ is the function whose components are the convex closures of the corresponding components of g_i. Then

$$f^* - q^* \leq (n+1) \max_{i=1,\ldots,m} \gamma_i,$$

where

$$\gamma_i = \sup\big\{\hat{f}_i(\tilde{x}_i) - (\check{\text{cl}}\, f_i)(\tilde{x}_i) \mid \tilde{x}_i \in \text{conv}(X_i)\big\},$$

$\check{\text{cl}}\, f_i$ is the convex closure of f_i, and $\hat{f}_i$ is given by

$$\hat{f}_i(\tilde{x}_i) = \inf\big\{f_i(x_i) \mid g_i(x_i) \leq (\check{\text{cl}}\, g_i)(\tilde{x}_i),\, x_i \in X_i\big\}, \qquad \forall\, \tilde{x}_i \in \text{conv}(X_i).$$

Proof: Let

$$\tilde{M}_i = \big\{\big(g_i(x_i), f_i(x_i)\big) \mid x_i \in X_i\big\}, \qquad i = 1, ..., m,$$

and

$$M_i = \tilde{M}_i + \big\{(u, 0) \mid u \geq 0\big\}.$$

Each M_i is the vector sum of a compact set and a closed convex set. Their vector sum becomes

$$M = M_1 + M_2 + \cdots + M_m.$$

Using Prop. 5.3 and a standard duality argument, we have

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

To complete the proof, we now prove the following lemma:

Lemma 5.1: Let $f : \Re^n \mapsto \Re$ and $g : \Re^n \mapsto \Re^r$ be functions, and let X be a set. Denote

$$\tilde{M} = \big\{\big(g(x), f(x)\big) \mid x \in X\big\},$$

and

$$M = \tilde{M} + \big\{(u, 0) \mid u \geq 0\big\}.$$

We have

$$\mathrm{conv}(M) \subset \big\{(\check{\mathrm{cl}}\, g(\tilde{x}), \check{f}(\tilde{x})) \mid \tilde{x} \in \mathrm{conv}(X)\big\} + \big\{(u, 0) \mid u \geq 0\big\},$$

where $\check{f}$ is the convex closure of f and $\check{g}$ is the function whose components are the convex closures of the corresponding components of g.

Proof: By Caratheodory's Theorem, any $(u, w) \in \mathrm{conv}(M)$ is the convex combination of at most $r + 2$ elements of M,

$$(u, w) = \sum_{j=1}^{r+2} \alpha_j (u_j, w_j), \qquad (u_j, w_j) \in M,\ j = 1, ..., r + 2,$$

where $\sum_{j=1}^{r+2} \alpha_j = 1$ and $\alpha_j \geq 0$ for all j. Since $(u_j, w_j) \in M$, there exists $x_j \in M$ such that $g(x_j) \leq u_j, f(x_j) = w_j$. Let $\tilde{x} = \sum_{j=1}^{r+2} \alpha_j x_j$, so that

$$\tilde{x} \in \mathrm{conv}(X), \qquad (\check{\mathrm{cl}}\, g)(\tilde{x}) \leq \sum_{j=1}^{r+2} \alpha_j g(x_j) \leq u, \qquad (\check{\mathrm{cl}}\, f)(\tilde{x}) \leq \sum_{j=1}^{r+2} \alpha_j f(x_j) = w.$$

It follows that $(u, w) \in \big\{((\check{\mathrm{cl}}\, g)(\tilde{x}), \check{f}(\tilde{x})) \mid \tilde{x} \in \mathrm{conv}(X)\big\} + \big\{(u, 0) \mid u \geq 0\big\}$. **Q.E.D.**

For $i = 1, ..., m$, for each vector $(u_i, w_i) \in \mathrm{conv}(M_i)$, using the preceding lemma, we have that there exists $\tilde{x}_i \in \mathrm{conv}(X_i)$ satisfying

$$(\check{\mathrm{cl}}\, g_i)(\tilde{x}_i) \leq u_i, \qquad \check{f}_i(\tilde{x}_i) = w_i.$$

It follows that there exists $x_i \in X_i$ such that

$$g_i(x_i) \leq (\check{\mathrm{cl}}\, g_i)(\tilde{x}_i) \leq u_i.$$

Letting $\bar{u}_i = g_i(x_i)$ and $\bar{w}_i = f_i(x_i)$ we have that for each vector $(u_i, w_i) \in \mathrm{conv}(M_i)$, there exists a vector of the form $(\bar{u}_i, \bar{w}_i) \in M_i$ with $u_i - \bar{u}_i \in C_i$, where $C_i = \{u \mid u \geq 0\}$. We now apply the result of Prop. 5.3 and obtain

$$w^* - q^* \leq (n+1) \max_{i=1,\ldots,m} \gamma_i.$$

Q.E.D.

Estimates of Duality Gap in Minimax

Let us now consider duality gap issues in the context of minimax problems.† We saw in Sections 2 and 4 that minimax theory can be developed within the context of the MC/MC framework. In particular, consider a function $\phi : X \times Z \mapsto \Re$ defined over nonempty sets X and Z, and the MC/MC framework involving the set

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

where $p : \Re^m \mapsto [-\infty, \infty]$ is the function given by

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

Let us assume that $(\hat{\mathrm{cl}}\, \phi)(x, \mu) < \infty$ for all $x \in X$ and $\mu \in Z$, so that the corresponding crossing function is (cf. Prop. 2.4)

$$q(\mu) = \inf_{x \in X} (\hat{\mathrm{cl}}\, \phi)(x, \mu). \tag{5.3}$$

Then we have

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) \leq \sup_{z \in Z} \inf_{x \in X} (\hat{\mathrm{cl}}\, \phi)(x, z) = \sup_{\mu \in \Re^m} q(\mu) = q^*. \tag{5.4}$$

Let us also assume that

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

so that $w^* < \infty$ in the corresponding MC/MC framework. Then, using also the discussion in Sections 2-4 [cf. Prop. 3.4(b)], we see that

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) \leq q^* = (\check{\mathrm{cl}}\, p)(0) \leq p(0) = w^* = \inf_{x \in X} \sup_{z \in Z} \phi(x, z), \tag{5.5}$$

where $\check{\mathrm{cl}}\, p$ is the convex closure of p.

† The ideas of this subsection on minimax problems are due to Mengdi Wang.

Thus the gap between "infsup" and "supinf" can be decomposed into the sum of two terms: $w^* - q^* = g_1 + g_2$, where

$$g_1 = \inf_{x \in X} \sup_{z \in Z} \phi(x,z) - (\check{\mathrm{cl}}\, p)(0) = p(0) - (\check{\mathrm{cl}}\, p)(0),$$

$$g_2 = (\check{\mathrm{cl}}\, p)(0) - \sup_{z \in Z} \inf_{x \in X} \phi(x,z) = \sup_{z \in \Re^m} \inf_{x \in X} (\hat{\mathrm{cl}}\, \phi)(x,z) - \sup_{z \in Z} \inf_{x \in X} \phi(x,z).$$

The term g_1 can be attributed to the lack of convexity and/or closure of ϕ with respect to x; it would be 0 if ϕ were closed and convex with respect to x for each z. Similarly, the term g_2 can be attributed to the lack of concavity and/or upper semicontinuity of ϕ with respect to z.

In cases where ϕ has a separable form in x and ϕ is concave and upper semicontinuous with respect to z, we have $g_2 = 0$, while the Shapley-Folkman Theorem can be used to estimate g_1, similar to Prop. 5.4. Similarly, we can estimate the duality gap if ϕ has a separable form in z and ϕ is convex and lower semicontinuous with respect to x, or if ϕ separable in both x and z.

6. REFERENCES

[AuT03] Auslender, A., and Teboulle, M., 2003. Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer, N. Y.

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

[Bar02] Barvinok, A., 2002. A Course in Convexity, American Math. Society, Providence, RI.

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

[Ber09] Bertsekas, D. P., 2009. Convex Optimization Theory, Athena Scientific, Belmont, MA.

[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.

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

[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.

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

[Web94] Webster, R., 1994. Convexity, Oxford Univ. Press, Oxford, UK.

[Zho93] Zhou, L., 1993. "A Simple Proof of the Shapley-Folkman Theorem," Economic Theory, Vol. 3, pp. 371-372.

7. APPENDIX: CONVEX ANALYSIS BACKGROUND

We summarize here some convex analysis backgound, with few proofs or detailed explanations. All the results for which we give no proofs can be found in Rockafellar [Roc70], as well as other convex analysis sources. We give page references to our text [BNO03]. The reader may wish to skip this section and return to it later as needed.

Relative Interior and Closure

Let C be a nonempty convex set. We say that x is a *relative interior point* of C if $x \in C$ and there exists an open sphere S centered at x such that $S \cap \mathrm{aff}(C) \subset C$, i.e., x is an interior point of C relative to $\mathrm{aff}(C)$. The set of all relative interior points of C is called the *relative interior of* C, and is denoted by $\mathrm{ri}(C)$. The set C is said to be *relatively open* if $\mathrm{ri}(C) = C$. The vectors in the closure of C that are not relative interior points are said to be *relative boundary points* of C, and their collection is called the *relative boundary* of C. A central fact about relative interiors is given in the following proposition ([BNO03], p. 40).

Proposition 7.1: Let C be a nonempty convex set.

(a) (Line Segment Principle) If $x \in \mathrm{ri}(C)$ and $\overline{x} \in \mathrm{cl}(C)$, then all points on the line segment connecting x and $\overline{x}$, except possibly $\overline{x}$, belong to $\mathrm{ri}(C)$.

(b) $\mathrm{ri}(C)$ is a nonempty convex set, and has the same affine hull as C.

(c) (Prolongation Lemma) A vector x is a relative interior point of C if and only if every line segment in C having x as one endpoint can be prolonged beyond x without leaving C [i.e., for every $\overline{x} \in C$, there exists a $\gamma > 0$ such that $x + \gamma(x - \overline{x}) \in C$].

A relative interior condition is central in the following useful characterization of the set of optimal solutions in the case where the cost function is linear or, more generally, concave. For a proof, see [BNO03], p. 43.

Proposition 7.2: Let X be a nonempty convex subset of $\Re^n$, let $f : X \mapsto \Re$ be a concave function, and let X^* be the set of vectors where f attains a minimum over X, i.e.,

$$X^* = \left\{ x \in X \;\middle|\; f(x^*) = \inf_{x \in X} f(x) \right\}.$$

If X^* contains a relative interior point of X, then f must be constant over X, i.e., $X^* = X$.

Recession Cones

Given a nonempty convex set C, we say that a vector d is a *direction of recession* of C if $x + \alpha d \in C$ for all $x \in C$ and $\alpha \geq 0$. The set of all directions of recession is the *recession cone* of C and it is denoted by R_C. The following is the basic result relating to recession cones (see [BNO03], p. 50).

Proposition 7.3: (Recession Cone Theorem) Let C be a nonempty closed convex set.

(a) The recession cone R_C is a closed convex cone.

(b) A vector d belongs to R_C if and only if there exists a vector $x \in C$ such that $x + \alpha d \in C$ for all $\alpha \geq 0$.

(c) R_C contains a nonzero direction if and only if C is unbounded.

(d) $R_C = R_{\mathrm{ri}(C)}$.

(e) For any collection of closed convex sets C_i, $i \in I$, where I is an arbitrary index set and $\cap_{i \in I} C_i \neq \emptyset$, we have

$$R_{\cap_{i \in I} C_i} = \cap_{i \in I} R_{C_i}.$$

Given a closed proper convex function $f : \Re^n \mapsto (-\infty, \infty]$, the level sets of f are the sets $\{x \mid f(x) \leq \gamma\}$, $\gamma \in \Re$. It can be shown that all the nonempty level sets of f have the same recession cone. In particular, if one of them is compact, all of them are compact (see [BNO03], p. 93).

The *lineality space* of a convex set C, denoted by L_C. It is defined as the set of directions of recession d whose opposite, $-d$, are also directions of recession:

$$L_C = R_C \cap (-R_C).$$

The following is a useful decomposition result (see [BNO03], p. 56).

Proposition 7.4: (Decomposition of a Convex Set) Let C be a nonempty convex subset of $\Re^n$. Then, for every subspace S that is contained in the lineality space L_C, we have

$$C = S + (C \cap S^\perp).$$

The following is a useful closedness criterion for a vector sum (see [BNO03], p. 67). As a special case, it yields that the vector sum of two sets is closed if both sets are closed and convex, and one of the two is bounded, so that its recession cone is just the origin.

Proposition 7.5: Let $C_1, \ldots, C_m$ be nonempty closed convex subsets of $\Re^n$ such that the equality $y_1 + \cdots + y_m = 0$ with $y_i \in R_{C_i}$ implies that each y_i belongs to the lineality space of C_i. Then, the vector sum $C_1 + \cdots + C_m$ is a closed set and

$$R_{C_1+\cdots+C_m} = R_{C_1} + \cdots + R_{C_m}.$$

Partial Minimization of Convex Functions

There are some useful conditions guaranteeing that closedness is preserved under partial minimization, while simultaneously the partial minimum is attained. Starting with a function $F : \Re^{n+m} \mapsto [-\infty, \infty]$ of two vectors $x \in \Re^n$ and $z \in \Re^m$, we consider the function

$$f(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n.$$

Convexity of F implies convexity of f, as stated in the following proposition (see [BNO03], Section 2.3.3).

Proposition 7.6: Consider a function $F : \Re^{n+m} \mapsto (-\infty, \infty]$ and the function $f : \Re^n \mapsto [-\infty, \infty]$ defined by

$$f(x) = \inf_{z \in \Re^m} F(x, z).$$

Then:

(a) If F is convex, then f is also convex.

(b) We have

$$P\big(\mathrm{epi}(F)\big) \subset \mathrm{epi}(f) \subset \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big), \tag{7.1}$$

where $P(\cdot)$ denotes projection on the space of (x, w), i.e., for any subset S of $\Re^{n+m+1}$, $P(S) = \big\{(x, w) \mid (x, z, w) \in S\big\}$.

It is not necessarily true that closedness of F implies closedness of f, as we have discussed in Section 1.3.3. Indeed, as Prop. 7.6(b) suggests, if the projection operation does not preserve closedness of epi(F), then epi(f) may not be closed. The following propositions provide conditions guaranteeing closedness (see [BNO03], Section 2.3.3).

Proposition 7.7: Let $F : \Re^{n+m} \mapsto (-\infty, \infty]$ be a closed proper convex function, and consider the function f given by

$$f(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n.$$

Assume that for some $\overline{x} \in \Re^n$ and $\overline{\gamma} \in \Re$ the set

$$\big\{z \mid F(\overline{x}, z) \leq \overline{\gamma}\big\}$$

is nonempty and compact. Then f is convex, closed, and proper. Furthermore, for each $x \in \mathrm{dom}(f)$, the set of minima in the definition of $f(x)$ is nonempty and compact.

The assumption that there exists a vector $\overline{x} \in \Re^n$ and a scalar $\overline{\gamma}$ such that the level set

$$\big\{z \mid F(\overline{x}, z) \leq \overline{\gamma}\big\}$$

is nonempty and compact is equivalent to assuming that all the nonempty level sets of the form $\big\{z \mid F(x, z) \leq \gamma\big\}$ are compact. This is so because all these sets have the same recession cone, namely $\big\{d_z \mid (0, d_z) \in R_F\big\}$.

A simple but useful special case of the preceding proposition is the following:

Proposition 7.8: Let X and Z be nonempty convex sets of $\Re^n$ and $\Re^m$, respectively, let $F : X \times Z \mapsto \Re$ be a closed convex function, and assume that Z is compact. Then the function f given by

$$f(x) = \inf_{z \in Z} F(x, z), \qquad x \in X,$$

is a closed real-valued convex function over X.

The following is a slight generalization of Prop. 7.7.

Proposition 7.9: Let $F : \Re^{n+m} \mapsto (-\infty, \infty]$ be a closed proper convex function, and consider the function f given by

$$f(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n.$$

Assume that for some $\overline{x} \in \Re^n$ and $\overline{\gamma} \in \Re$ the set

$$\big\{z \mid F(\overline{x}, z) \leq \overline{\gamma}\big\}$$

is nonempty and its recession cone is equal to its lineality space. Then f is convex, closed, and proper. Furthermore, for each $x \in \mathrm{dom}(f)$, the set of minima in the definition of $f(x)$ is nonempty.

Closures of Convex Functions

A nonempty subset E of $\Re^{n+1}$ is the epigraph of some function if for every $(\overline{x}, \overline{w}) \in E$, the set $\big\{w \mid (\overline{x}, w) \in E\big\}$ is either the real line or else it is a halfline that is bounded below and contains its (lower) endpoint. Then E is the epigraph of any function $f : X \mapsto [-\infty, \infty]$ such that

$$X \supset \big\{x \mid \text{there exists } (x, w) \in E\big\},$$

and

$$f(x) = \inf\big\{w \mid (x, w) \in E\big\}, \qquad \forall\, x \in X.$$

If E is the empty set, it is the epigraph of the function that is identically equal to ∞.

The closure of the epigraph of a function $f : X \mapsto [-\infty, \infty]$ can be seen to be a legitimate epigraph of another function. This function, called the *closure of* f, is denoted by $\mathrm{cl}\, f$, and is given by†

$$(\mathrm{cl}\, f)(x) = \inf\big\{w \mid (x, w) \in \mathrm{cl}\big(\mathrm{epi}(f)\big)\big\}, \qquad x \in \Re^n.$$

Similarly, the closure of the convex hull of the epigraph of f is the epigraph of some function, denoted $\check{\mathrm{cl}}\, f$, necessarily closed and convex, which is called the *convex closure of* f. It can be seen that $\check{\mathrm{cl}}\, f$ is the closure of the function $F : \Re^n \mapsto [-\infty, \infty]$ given by

$$F(x) = \inf\big\{w \mid (x, w) \in \mathrm{conv}\big(\mathrm{epi}(f)\big)\big\}. \tag{7.2}$$

It is easily seen that F convex, but it need not be closed and its domain may be strictly contained in $\mathrm{dom}(\check{\mathrm{cl}}\, f)$ (it can be seen though that the closures of the domains of F and $\check{\mathrm{cl}}\, f$ coincide).

From the point of view of optimization, an important property is that the minimal values of f, $\mathrm{cl}\, f$, F, and $\check{\mathrm{cl}}\, f$ coincide, as stated in the following proposition:

Proposition 7.10: Let $f : X \mapsto [-\infty, \infty]$ be a function. Then

$$\inf_{x \in X} f(x) = \inf_{x \in \Re^n} (\mathrm{cl}\, f)(x) = \inf_{x \in \Re^n} F(x) = \inf_{x \in \Re^n} (\check{\mathrm{cl}}\, f)(x),$$

wheer F is given by Eq. (7.2). Furthermore, any vector that attains the infimum of f also attains the infimum of $\mathrm{cl}\, f$, F, and $\check{\mathrm{cl}}\, f$.

Proof: If $\mathrm{epi}(f)$ is empty, i.e., $f(x) = \infty$ for all x, the results trivially hold. We thus may assume that $\mathrm{epi}(f)$ is nonempty. Let $f^* = \inf_{x \in \Re^n} (\mathrm{cl}\, f)(x)$. For any sequence $\big\{(x_k, w_k)\big\} \subset \mathrm{cl}\big(\mathrm{epi}(f)\big)$ with $w_k \to f^*$, we can construct a sequence $\big\{(\overline{x}_k, \overline{w}_k)\big\} \subset \mathrm{epi}(f)$ such that $\overline{x}_k \in X$, and $|w_k - \overline{w}_k| \to 0$, so that $\overline{w}_k \to f^*$. This shows that

$$f^* = \inf_{x \in X} f(x) = \inf_{x \in \Re^n} (\mathrm{cl}\, f)(x).$$

To show that $\inf_{x \in \Re^n} F(x) = f^*$, we note that clearly $F(x) \le f(x)$ for all $x \in \Re^n$. If $f^* = -\infty$, then $\inf_{x \in \Re^n} F(x) = f^*$. Assume to arrive at a contradiction that $-\infty < f^*$ and $\inf_{x \in \Re^n} F(x) < f^*$. Then there exists $(\overline{x}, \overline{w}) \in \mathrm{conv}\big(\mathrm{epi}(f)\big)$ such that $F(\overline{x}) \le \overline{w} < f^*$, and $\overline{x} = \sum_{i=1}^m \alpha_i x_i$, $\overline{w} = \sum_{i=1}^m \alpha_i w_i$ for some $(x_i, w_i) \in \mathrm{epi}(f)$ and $\alpha_i \ge 0$ with $\sum_{i=1}^m \alpha_i = 1$. It follows that

$$\sum_{i=1}^m \alpha_i f_i(x_i) \le \sum_{i=1}^m \alpha_i w_i < f^*,$$

† In Rockafellar [Roc70], $\mathrm{cl}\, f$ is called the "lower semicontinuous hull of f" while in Rockafellar and Wets [RoW98], $\mathrm{cl}\, f$ is called the "lower closure of f."

while $f(x_i) \geq f^*$ for all i – a contradiction. Hence $\inf_{x\in\Re^n} F(x) = f^*$. Since $\check{\mathrm{cl}}\, f$ is the closure of F, it follows also that $\inf_{x\in\Re^n} (\check{\mathrm{cl}}\, f)(x) = f^*$.

We have $f(x) \geq (\mathrm{cl}\, f)(x)$ for all x, so if x^* attains the infimum of f,

$$\inf_{x\in\Re^n} (\mathrm{cl}\, f)(x) = \inf_{x\in X} f(x) = f(x^*) \geq (\mathrm{cl}\, f)(x^*),$$

showing that x^* attains the infimum of $(\mathrm{cl}\, f)$. Similarly, x^* attains the infimum of F and $\check{\mathrm{cl}}\, f$. **Q.E.D.**

Working with the closure of a convex function is often useful because in some sense the closure "differs minimally" from the original. In particular, a proper convex function coincides with its closure on the relative interior of its domain. This and other properties of closures are derived in the following proposition.

Proposition 7.11: Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function. Then:

(a) $\mathrm{cl}\, f$ is the greatest closed function majorized by f, i.e., if $g : \Re^n \mapsto [-\infty, \infty]$ is closed and satisfies $g(x) \leq f(x)$ for all $x \in \Re^n$, then $g(x) \leq (\mathrm{cl}\, f)(x)$ for all $x \in \Re^n$.

(b) $\check{\mathrm{cl}}\, f$ is the greatest closed and convex function majorized by f, i.e., if $g : \Re^n \mapsto [-\infty, \infty]$ is closed and convex, and satisfies $g(x) \leq f(x)$ for all $x \in \Re^n$, then $g(x) \leq (\mathrm{cl}\, f)(x)$ for all $x \in \Re^n$.

(c) Let f be convex. Then $\mathrm{cl}\, f$ is convex, and it is proper if and only if f is proper. Furthermore,

$$(\mathrm{cl}\, f)(x) = f(x), \qquad \forall\, x \in \mathrm{ri}\big(\mathrm{dom}(f)\big),$$

and if $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$ and $y \in \mathrm{dom}(\mathrm{cl}\, f)$, then

$$(\mathrm{cl}\, f)(y) = \lim_{\alpha \downarrow 0} f\big(y + \alpha(x - y)\big).$$

Proof: (a) Let $g : \Re^n \mapsto [-\infty, \infty]$ be closed and such that $g(x) \leq f(x)$ for all $x \in \Re^n$. Choose any $x \in \mathrm{dom}(\mathrm{cl}\, f)$. Since $\mathrm{epi}(\mathrm{cl}\, f) = \mathrm{cl}\big(\mathrm{epi}(f)\big)$, there is a sequence $\big\{(x_k, w_k)\big\} \in \mathrm{epi}(f)$ such that $x_k \to x$, $w_k \to (\mathrm{cl}\, f)(x)$. Using the lower semicontinuity of g, we have

$$g(x) \leq \liminf_{k\to\infty} g(x_k) \leq \liminf_{k\to\infty} f(x_k) \leq \liminf_{k\to\infty} w_k = (\mathrm{cl}\, f)(x).$$

Since $\mathrm{epi}(f) \subset \mathrm{epi}(\mathrm{cl}\, f)$, we have $(\mathrm{cl}\, f)(x) \leq f(x)$ for all $x \in \Re^n$, so $\mathrm{cl}\, f$ is the greatest closed function majorized by f.

(b) Similar to the proof of part (a).

(c) If $\mathrm{epi}(f)$ is convex, then $\mathrm{cl}\big(\mathrm{epi}(f)\big)$ is convex, implying that $\mathrm{cl}\, f$ is convex. Furthermore, $\mathrm{epi}(f)$ contains a vertical line if and only if $\mathrm{cl}\big(\mathrm{epi}(f)\big)$ contains a vertical line (since these two sets share the same relative interior). Thus $\mathrm{cl}\, f$ is proper if and only if f is. For a proof of the remaining statement, see [Roc70], Ch. 7. **Q.E.D.**

Hyperplane Theorems

An important tool for our analysis is the following proposition (see [BNO03], p. 118).

Proposition 7.12: (Nonvertical Hyperplane Theorem) 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 $\beta \neq 0$, 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 $(\overline{u}, \overline{w})$ and C.

The following two hyperplane separation theorems will be important tools in the analysis (see [BNO03], pp. 115 and 193).

Proposition 7.13: (Proper Separation Theorem) Let C be a nonempty convex subset of $\Re^n$ and let $\overline{x}$ be a vector in $\Re^n$. There exists a hyperplane that properly separates C and $\overline{x}$ if and only if

$$\overline{x} \notin \mathrm{ri}(C).$$

Proposition 7.14: (Polyhedral Proper Separation Theorem) Let C and P be two nonempty convex subsets of $\Re^n$ such that P is polyhedral. There exists a hyperplane that separates C and P, and does not contain C if and only if

$$\mathrm{ri}(C) \cap P = \emptyset.$$

Conjugate Functions

Consider an extended real-valued function $f : \Re^n \mapsto [-\infty, \infty]$. The *conjugate function* of f is the function $h : \Re^n \mapsto [-\infty, \infty]$ defined by

$$h(y) = \sup_{x \in \Re^n} \big\{x'y - f(x)\big\}, \qquad y \in \Re^n. \tag{7.3}$$

We have the following basic result, which involves the conjugate of the conjugate function (or *double conjugate*, see [BNO03], p. 426).

Proposition 7.15: (Conjugacy Theorem) Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function, let h be its conjugate, and consider the conjugate of h, denoted by $f^\ddagger$:

$$f^\ddagger(x) = \sup_{y \in \Re^n} \big\{y'x - h(y)\big\}, \qquad x \in \Re^n.$$

(a) We have

$$f(x) \geq f^\ddagger(x), \qquad \forall\, x \in \Re^n.$$

(b) If f is convex, then properness of any one of the functions f, h, and $f^\ddagger$ implies properness of the other two.

(c) If f is closed proper convex, then

$$f(x) = f^\ddagger(x), \qquad \forall\, x \in \Re^n.$$

(d) Let $\check{\mathrm{cl}}\, f$ be the convex closure of f. If $(\check{\mathrm{cl}}\,)f(x) > -\infty$ for all $x \in \Re^n$, then

$$(\check{\mathrm{cl}}\,)f(x) = f^\ddagger(x), \qquad \forall\, x \in \Re^n.$$

Minimax Theory

Consider a function $\phi : X \times Z \mapsto \Re$, where X and Z are nonempty subsets of $\Re^n$ and $\Re^m$, respectively. We will focus on deriving conditions guaranteeing that

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z), \tag{7.4}$$

and that the infimum and the supremum above are attained. We note the definition and a basic property of saddle points (see [BNO03], Section 2.6).

Definition 7.1: A pair of vectors $x^* \in X$ and $z^* \in Z$ is called a *saddle point* of ϕ if

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

Proposition 7.16: A pair (x^*, z^*) is a saddle point of ϕ if and only if the minimax equality (7.4) holds, and x^* is an optimal solution of the problem

$$\begin{aligned} &\text{minimize} && \sup_{z \in Z} \phi(x, z) \\ &\text{subject to} && x \in X, \end{aligned} \tag{7.5}$$

while z^* is an optimal solution of the problem

$$\begin{aligned} &\text{maximize} && \inf_{x \in X} \phi(x, z) \\ &\text{subject to} && z \in Z. \end{aligned} \tag{7.6}$$