# CONVEX OPTIMIZATION THEORY

# ATHENA SCIENTIFIC, BELMONT, MASS

# A SUMMARY

**by**

**Dimitri P. Bertsekas**

**SPRING 2010**

We provide a summary of theoretical concepts and results relating to convex analysis, convex optimization and duality theory. In particular, we list definitions and propositions (without proofs) of the author's book "Convex Optimization Theory," Athena Scientific, 2009 (abbreviated COT). The proposition numbers have been preserved from this book.

## CHAPTER 1: Basic Concepts of Convex Analysis

### Section 1.1. Convex Sets and Functions

**Definition 1.1.1:** A subset $C$ of $\Re^n$ is called *convex* if

$$\alpha x + (1-\alpha)y \in C, \qquad \forall\ x, y \in C,\ \forall\ \alpha \in [0,1].$$

**Proposition 1.1.1:**

(a) The intersection $\cap_{i\in I} C_i$ of any collection $\{C_i \mid i \in I\}$ of convex sets is convex.

(b) The vector sum $C_1 + C_2$ of two convex sets $C_1$ and $C_2$ is convex.

(c) The set $\lambda C$ is convex for any convex set $C$ and scalar $\lambda$. Furthermore, if $C$ is a convex set and $\lambda_1$, $\lambda_2$ are positive scalars,

$$(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C.$$

(d) The closure and the interior of a convex set are convex.

(e) The image and the inverse image of a convex set under an affine function are convex.

A *hyperplane* is a set of the form $\{x \mid a'x = b\}$, where $a$ is a nonzero vector and $b$ is a scalar. A *halfspace* is a set specified by a single linear inequality, i.e., a set of the form $\{x \mid a'x \le b\}$, where $a$ is a nonzero vector and $b$ is a scalar. A set is said to be *polyhedral* if it is nonempty and it has the form $\{x \mid a_j'x \le b_j,\ j = 1, \ldots, r\}$, where $a_1, \ldots, a_r$ and $b_1, \ldots, b_r$ are some vectors in $\Re^n$ and scalars, respectively. A set $C$ is said to be a *cone* if for all $x \in C$ and $\lambda > 0$, we have $\lambda x \in C$.

**Definition 1.1.2:** Let $C$ be a convex subset of $\Re^n$. We say that a function $f : C \mapsto \Re$ is *convex* if

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha) f(y), \qquad \forall\ x, y \in C,\ \forall\ \alpha \in [0,1].$$

A convex function $f : C \mapsto \Re$ is called *strictly convex* if $f\big(\alpha x + (1 - \alpha)y\big) < \alpha f(x) + (1-\alpha) f(y)$ for all $x, y \in C$ with $x \ne y$, and all $\alpha \in (0,1)$. A function $f : C \mapsto \Re$, where $C$ is a convex set, is called *concave* if the function $(-f)$ is convex.

The *epigraph* of a function $f : X \mapsto [-\infty, \infty]$, where $X \subset \Re^n$, is defined to be the subset of $\Re^{n+1}$ given by

$$\mathrm{epi}(f) = \big\{(x, w) \mid x \in X,\ w \in \Re,\ f(x) \le w\big\}.$$

The *effective domain* of $f$ is defined to be the set

$$\mathrm{dom}(f) = \big\{x \in X \mid f(x) < \infty\big\}.$$

We say that $f$ is *proper* if $f(x) < \infty$ for at least one $x \in X$ and $f(x) > -\infty$ for all $x \in X$, and we say that $f$ *improper* if it is not proper.

**Definition 1.1.3:** Let $C$ be a convex subset of $\Re^n$. We say that an extended real-valued function $f : C \mapsto [-\infty, \infty]$ is *convex* if $\mathrm{epi}(f)$ is a convex subset of $\Re^{n+1}$.

**Definition 1.1.4:** Let $C$ and $X$ be subsets of $\Re^n$ such that $C$ is nonempty and convex, and $C \subset X$. We say that an extended real-valued function $f : X \mapsto [-\infty, \infty]$ is *convex over* $C$ if $f$ becomes convex when the domain of $f$ is restricted to $C$, i.e., if the function $\tilde{f} : C \mapsto [-\infty, \infty]$, defined by $\tilde{f}(x) = f(x)$ for all $x \in C$, is convex.

We say that a function $f : X \mapsto [-\infty, \infty]$ is *closed* if $\mathrm{epi}(f)$ is a closed set. 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 $x_k \to x$. We say that $f$ *is lower semicontinuous* if it is lower semicontinuous at each point $x$ in its domain $X$. We say that $f$ is *upper semicontinuous* if $-f$ is lower semicontinuous.

**Proposition 1.1.2:** For a function $f : \Re^n \mapsto [-\infty, \infty]$, the following are equivalent:

(i) The level set $V_\gamma = \big\{x \mid f(x) \leq \gamma\big\}$ is closed for every scalar $\gamma$.

(ii) $f$ is lower semicontinuous.

(iii) $\mathrm{epi}(f)$ is closed.

**Proposition 1.1.3:** Let $f : X \mapsto [-\infty, \infty]$ be a function. If $\mathrm{dom}(f)$ is closed and $f$ is lower semicontinuous at each $x \in \mathrm{dom}(f)$, then $f$ is closed.

**Proposition 1.1.4:** Let $f : \Re^m \mapsto (-\infty, \infty]$ be a given function, let $A$ be an $m \times n$ matrix, and let $F : \Re^n \mapsto (-\infty, \infty]$ be the function

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

If $f$ is convex, then $F$ is also convex, while if $f$ is closed, then $F$ is also closed.

**Proposition 1.1.5:** Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be given functions, let $\gamma_1, \ldots, \gamma_m$ be positive scalars, and let $F : \Re^n \mapsto (-\infty, \infty]$ be the function

$$F(x) = \gamma_1 f_1(x) + \cdots + \gamma_m f_m(x), \qquad x \in \Re^n.$$

If $f_1, \ldots, f_m$ are convex, then $F$ is also convex, while if $f_1, \ldots, f_m$ are closed, then $F$ is also closed.

**Proposition 1.1.6:** Let $f_i : \Re^n \mapsto (-\infty, \infty]$ be given functions for $i \in I$, where $I$ is an arbitrary index set, and let $f : \Re^n \mapsto (-\infty, \infty]$ be the function given by

$$f(x) = \sup_{i \in I} f_i(x).$$

If $f_i$, $i \in I$, are convex, then $f$ is also convex, while if $f_i$, $i \in I$, are closed, then $f$ is also closed.

**Proposition 1.1.7:** Let $C$ be a nonempty convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be differentiable over an open set that contains $C$.

(a) $f$ is convex over $C$ if and only if

$$f(z) \geq f(x) + \nabla f(x)'(z - x), \qquad \forall\ x, z \in C.$$

(b) $f$ is strictly convex over $C$ if and only if the above inequality is strict whenever $x \neq z$.

**Proposition 1.1.8:** Let $C$ be a nonempty convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be convex and differentiable over an open set that contains $C$. Then a vector $x^* \in C$ minimizes $f$ over $C$ if and only if

$$\nabla f(x^*)'(z - x^*) \geq 0, \qquad \forall\ z \in C.$$

**Proposition 1.1.9: (Projection Theorem)** Let $C$ be a nonempty closed convex subset of $\Re^n$, and let $z$ be a vector in $\Re^n$. There exists a unique vector that minimizes $\|z-x\|$ over $x \in C$, called the projection of $z$ on $C$. Furthermore, a vector $x^*$ is the projection of $z$ on $C$ if and only if

$$(z - x^*)'(x - x^*) \leq 0, \qquad \forall\, x \in C.$$

**Proposition 1.1.10:** Let $C$ be a nonempty convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be twice continuously differentiable over an open set that contains $C$.

(a) If $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$, then $f$ is convex over $C$.

(b) If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then $f$ is strictly convex over $C$.

(c) If $C$ is open and $f$ is convex over $C$, then $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$.

## Section 1.2. Convex and Affine Hulls

The *convex hull* of a set $X$, denoted $\text{conv}(X)$, is the intersection of all convex sets containing $X$. A *convex combination* of elements of $X$ is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where $m$ is a positive integer, $x_1, \ldots, x_m$ belong to $X$, and $\alpha_1, \ldots, \alpha_m$ are scalars such that

$$\alpha_i \geq 0, \quad i = 1, \ldots, m, \qquad \sum_{i=1}^m \alpha_i = 1.$$

The convex hull $\text{conv}(X)$ is equal to the set of all convex combinations of elements of $X$. Also, for any set $S$ and linear transformation $A$, we have $\text{conv}(AS) = A\,\text{conv}(S)$. From this it follows that for any sets $S_1, \ldots, S_m$, we have $\text{conv}(S_1 + \cdots + S_m) = \text{conv}(S_1) + \cdots + \text{conv}(S_m)$.

If $X$ is a subset of $\Re^n$, the *affine hull* of $X$, denoted $\text{aff}(X)$, is the intersection of all affine sets containing $X$. Note that $\text{aff}(X)$ is itself an affine set and that it contains $\text{conv}(X)$. The dimension of $\text{aff}(X)$ is defined to be the dimension of the subspace parallel to $\text{aff}(X)$. It can be shown that $\text{aff}(X) = \text{aff}\big(\text{conv}(X)\big) = \text{aff}\big(\text{cl}(X)\big)$. For a convex set $C$, the *dimension* of $C$ is defined to be the dimension of $\text{aff}(C)$.

Given a nonempty subset $X$ of $\Re^n$, a *nonnegative combination* of elements of $X$ is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where $m$ is a positive integer, $x_1, \ldots, x_m$ belong to $X$, and $\alpha_1, \ldots, \alpha_m$ are nonnegative scalars. If the scalars $\alpha_i$ are all positive, $\sum_{i=1}^m \alpha_i x_i$ is said to be a *positive combination*. The *cone generated by* $X$, denoted $\mathrm{cone}(X)$, is the set of all nonnegative combinations of elements of $X$.

**Proposition 1.2.1: (Caratheodory's Theorem)** Let $X$ be a nonempty subset of $\Re^n$.

(a) Every nonzero vector from $\mathrm{cone}(X)$ can be represented as a positive combination of linearly independent vectors from $X$.

(b) Every vector from $\mathrm{conv}(X)$ can be represented as a convex combination of no more than $n+1$ vectors from $X$.

**Proposition 1.2.2:** The convex hull of a compact set is compact.

**Section 1.3. 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 the affine hull of $C$. 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 $\mathrm{cl}(C)$ that are not relative interior points are said to be *relative boundary points* of $C$, and their collection is called the *relative boundary* of $C$.

**Proposition 1.3.1: (Line Segment Principle)** Let $C$ be a nonempty convex set. 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 $\mathrm{ri}(C)$.

**Proposition 1.3.2: (Nonemptiness of Relative Interior)** Let $C$ be a nonempty convex set. Then:

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

(b) If $m$ is the dimension of $\text{aff}(C)$ and $m > 0$, there exist vectors $x_0, x_1, \ldots, x_m \in \text{ri}(C)$ such that $x_1 - x_0, \ldots, x_m - x_0$ span the subspace parallel to $\text{aff}(C)$.

**Proposition 1.3.3: (Prolongation Lemma)** Let $C$ be a nonempty convex set. 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 $x + \gamma(x - \overline{x}) \in C$].

**Proposition 1.3.4:** 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., $X^* = X$.

**Proposition 1.3.5:** Let $C$ be a nonempty convex set.

(a) $\text{cl}(C) = \text{cl}\big(\text{ri}(C)\big)$.

(b) $\text{ri}(C) = \text{ri}\big(\text{cl}(C)\big)$.

(c) Let $\overline{C}$ be another nonempty convex set. Then the following three conditions are equivalent:

(i) $C$ and $\overline{C}$ have the same relative interior.

(ii) $C$ and $\overline{C}$ have the same closure.

(iii) $\text{ri}(C) \subset \overline{C} \subset \text{cl}(C)$.

**Proposition 1.3.6:** Let $C$ be a nonempty convex subset of $\Re^n$ and let $A$ be an $m \times n$ matrix.

(a) We have $A \cdot \mathrm{ri}(C) = \mathrm{ri}(A \cdot C)$.

(b) We have $A \cdot \mathrm{cl}(C) \subset \mathrm{cl}(A \cdot C)$. Furthermore, if $C$ is bounded, then $A \cdot \mathrm{cl}(C) = \mathrm{cl}(A \cdot C)$.

**Proposition 1.3.7:** Let $C_1$ and $C_2$ be nonempty convex sets. We have

$$\mathrm{ri}(C_1 + C_2) = \mathrm{ri}(C_1) + \mathrm{ri}(C_2), \qquad \mathrm{cl}(C_1) + \mathrm{cl}(C_2) \subset \mathrm{cl}(C_1 + C_2).$$

Furthermore, if at least one of the sets $C_1$ and $C_2$ is bounded, then

$$\mathrm{cl}(C_1) + \mathrm{cl}(C_2) = \mathrm{cl}(C_1 + C_2).$$

**Proposition 1.3.8:** Let $C_1$ and $C_2$ be nonempty convex sets. We have

$$\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \subset \mathrm{ri}(C_1 \cap C_2), \qquad \mathrm{cl}(C_1 \cap C_2) \subset \mathrm{cl}(C_1) \cap \mathrm{cl}(C_2).$$

Furthermore, if the sets $\mathrm{ri}(C_1)$ and $\mathrm{ri}(C_2)$ have a nonempty intersection, then

$$\mathrm{ri}(C_1 \cap C_2) = \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2), \qquad \mathrm{cl}(C_1 \cap C_2) = \mathrm{cl}(C_1) \cap \mathrm{cl}(C_2).$$

**Proposition 1.3.9:** Let $C$ be a nonempty convex subset of $\Re^m$, and let $A$ be an $m \times n$ matrix. If $A^{-1} \cdot \mathrm{ri}(C)$ is nonempty, then

$$\mathrm{ri}(A^{-1} \cdot C) = A^{-1} \cdot \mathrm{ri}(C), \qquad \mathrm{cl}(A^{-1} \cdot C) = A^{-1} \cdot \mathrm{cl}(C),$$

where $A^{-1}$ denotes inverse image of the corresponding set under $A$.

**Proposition 1.3.10:** Let $C$ be a convex subset of $\Re^{n+m}$. For $x \in \Re^n$, denote

$$C_x = \{y \mid (x, y) \in C\},$$

and let

$$D = \{x \mid C_x \neq \emptyset\}.$$

Then

$$\mathrm{ri}(C) = \big\{(x, y) \mid x \in \mathrm{ri}(D),\, y \in \mathrm{ri}(C_x)\big\}.$$

#### Continuity of Convex Functions

**Proposition 1.3.11:** If $f : \Re^n \mapsto \Re$ is convex, then it is continuous. More generally, if $f : \Re^n \mapsto (-\infty, \infty]$ is a proper convex function, then $f$, restricted to $\mathrm{dom}(f)$, is continuous over the relative interior of $\mathrm{dom}(f)$.

**Proposition 1.3.12:** If $C$ is a closed interval of the real line, and $f : C \mapsto \Re$ is closed and convex, then $f$ is continuous over $C$.

#### Closures of Functions

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* $f$ and denoted $\mathrm{cl}\, f : \Re^n \mapsto [-\infty, \infty]$, is given by

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

The closure of the convex hull of the epigraph of $f$ is the epigraph of some function, denoted $\check{\mathrm{cl}}\, f$ 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

$$F(x) = \inf\big\{w \mid (x, w) \in \mathrm{conv}\big(\mathrm{epi}(f)\big)\big\}, \qquad x \in \Re^n. \tag{B.0}$$

It is easily shown that $F$ is 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).

**Proposition 1.3.13:** Let $f : X \mapsto [-\infty, \infty]$ be a function. Then

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

where $F$ is given by Eq. (B.0). Furthermore, any vector that attains the infimum of $f$ over $X$ also attains the infimum of $\mathrm{cl}\, f$, $F$, and $\check{\mathrm{cl}}\, f$.

**Proposition 1.3.14:** Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function.

(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 $g(x) \le f(x)$ for all $x \in \Re^n$, then $g(x) \le (\mathrm{cl}\, f)(x)$ 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 $g(x) \le f(x)$ for all $x \in \Re^n$, then $g(x) \le (\check{\mathrm{cl}}\, f)(x)$ for all $x \in \Re^n$.

**Proposition 1.3.15:** Let $f : \Re^n \mapsto [-\infty, \infty]$ be a convex function. Then:

(a) We have

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

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

Furthermore, $\mathrm{cl}\, f$ is proper if and only if $f$ is proper.

(b) If $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$, we have

$$(\mathrm{cl}\, f)(y) = \lim_{\alpha\downarrow 0} f\big(y + \alpha(x-y)\big), \qquad \forall\, y \in \Re^n.$$

**Proposition 1.3.16:** Let $f : \Re^m \mapsto [-\infty, \infty]$ be a convex function and $A$ be an $m \times n$ matrix such that the range of $A$ contains a point in $\mathrm{ri}\big(\mathrm{dom}(f)\big)$. The function $F$ defined by

$$F(x) = f(Ax),$$

is convex and

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

**Proposition 1.3.17:** Let $f_i : \Re^n \mapsto [-\infty, \infty]$, $i = 1, \ldots, m$, be convex functions such that

$$\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i)\big) \neq \emptyset.$$

The function $F$ defined by

$$F(x) = f_1(x) + \cdots + f_m(x),$$

is convex and

$$(\mathrm{cl}\, F)(x) = (\mathrm{cl}\, f_1)(x) + \cdots + (\mathrm{cl}\, f_m)(x), \qquad \forall\, x \in \Re^n.$$

## Section 1.4. 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 a cone containing the origin, called the *recession cone* of $C$, and denoted by $R_C$.

**Proposition 1.4.1: (Recession Cone Theorem)** Let $C$ be a nonempty closed convex set.

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

(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 $\alpha \geq 0$.

**Proposition 1.4.2: (Properties of Recession Cones)** Let $C$ be a nonempty closed convex set.

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

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

(c) 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

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

(d) Let $W$ be a compact and convex subset of $\Re^m$, and let $A$ be an $m \times n$ matrix. The recession cone of the set

$$V = \{x \in C \mid Ax \in W\}$$

(assuming this set is nonempty) is $R_C \cap N(A)$, where $N(A)$ is the nullspace of $A$.

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

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

**Proposition 1.4.3: (Properties of Lineality Space)** Let $C$ be a nonempty closed convex subset of $\Re^n$.

(a) $L_C$ is a subspace of $\Re^n$.

(b) $L_C = L_{\mathrm{ri}(C)}$.

(c) 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

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

(d) Let $W$ be a compact and convex subset of $\Re^m$, and let $A$ be an $m \times n$ matrix. The lineality space of the set

$$V = \{x \in C \mid Ax \in W\}$$

(assuming it is nonempty) is $L_C \cap N(A)$, where $N(A)$ is the nullspace of $A$.

**Proposition 1.4.4: (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 notion of direction of recession of a convex function $f$ can be described in terms of its epigraph via the following proposition.

**Proposition 1.4.5:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function and consider the level sets

$$V_\gamma = \big\{x \mid f(x) \leq \gamma\big\}, \qquad \gamma \in \Re.$$

Then:

(a) All the nonempty level sets $V_\gamma$ have the same recession cone, denoted $R_f$, and given by

$$R_f = \big\{d \mid (d, 0) \in R_{\text{epi}(f)}\big\},$$

where $R_{\text{epi}(f)}$ is the recession cone of the epigraph of $f$.

(b) If one nonempty level set $V_\gamma$ is compact, then all of these level sets are compact.

For a closed proper convex function $f : \Re^n \mapsto (-\infty, \infty]$, the (common) recession cone $R_f$ of the nonempty level sets is called the *recession cone of* $f$. A vector $d \in R_f$ is called a *direction of recession of* $f$.

The lineality space of the recession cone $R_f$ of a closed proper convex function $f$ is denoted by $L_f$, and is the subspace of all $d \in \Re^n$ such that both $d$ and $-d$ are directions of recession of $f$, i.e.,

$$L_f = R_f \cap (-R_f).$$

We have that $d \in L_f$ if and only if

$$f(x + \alpha d) = f(x), \qquad \forall\, x \in \text{dom}(f),\ \forall\, \alpha \in \Re.$$

Consequently, any $d \in L_f$ is called a *direction in which $f$ is constant*, and $L_f$ is called the *constancy space of* $f$.

**Proposition 1.4.6:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function. Then the recession cone and constancy space of $f$ are given in terms of its recession function by

$$R_f = \big\{d \mid r_f(d) \leq 0\big\}, \qquad L_f = \big\{d \mid r_f(d) = r_f(-d) = 0\big\}.$$

**Proposition 1.4.7:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function. Then, for all $x \in \mathrm{dom}(f)$ and $d \in \Re^n$,

$$r_f(d) = \sup_{\alpha > 0} \frac{f(x + \alpha d) - f(x)}{\alpha} = \lim_{\alpha \to \infty} \frac{f(x + \alpha d) - f(x)}{\alpha}.$$

**Proposition 1.4.8: (Recession Function of a Sum)** Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be closed proper convex functions such that the function $f = f_1 + \cdots + f_m$ is proper. Then

$$r_f(d) = r_{f_1}(d) + \cdots + r_{f_m}(d), \qquad \forall\ d \in \Re^n.$$

**Nonemptiness of Set Intersections**

A sequence of nonempty closed sets $\{C_k\}$ in $\Re^n$ with $C_{k+1} \subset C_k$ for all $k$ (such a sequence is said to be *nested*. We are concerned with the question whether $\cap_{k=0}^{\infty} C_k$ is nonempty.

**Definition 1.4.1:** Let $\{C_k\}$ be a nested sequence of nonempty closed convex sets. We say that $\{x_k\}$ is an *asymptotic sequence of* $\{C_k\}$ if $x_k \neq 0$, $x_k \in C_k$ for all $k$, and

$$\|x_k\| \to \infty, \qquad \frac{x_k}{\|x_k\|} \to \frac{d}{\|d\|},$$

where $d$ is some nonzero common direction of recession of the sets $C_k$,

$$d \neq 0, \qquad d \in \cap_{k=0}^{\infty} R_{C_k}.$$

A special case is when all the sets $C_k$ are equal. In particular, for a nonempty closed convex $C$, we say that $\{x_k\} \subset C$ is an asymptotic sequence of $C$ if $\{x_k\}$ is asymptotic for the sequence $\{C_k\}$, where $C_k \equiv C$.

Given any unbounded sequence $\{x_k\}$ such that $x_k \in C_k$ for each $k$, there exists a subsequence $\{x_k\}_{k\in\mathcal{K}}$ that is asymptotic for the corresponding subsequence $\{C_k\}_{k\in\mathcal{K}}$. In fact, any limit point of $\{x_k/\|x_k\|\}$ is a common direction of recession of the sets $C_k$.

**Definition 1.4.2:** Let $\{C_k\}$ be a nested sequence of nonempty closed convex sets. We say that *an asymptotic sequence* $\{x_k\}$ *is retractive* if for the direction $d$ corresponding to $\{x_k\}$ as per Definition 1.4.1, there exists an index $\overline{k}$ such that

$$x_k - d \in C_k, \qquad \forall\ k \geq \overline{k}.$$

We say that *the sequence* $\{C_k\}$ *is retractive* if all its asymptotic sequences are retractive. In the special case $C_k \equiv C$, we say that *the set* $C$ *is retractive* if all its asymptotic sequences are retractive.

A closed halfspace is retractive. Intersections and Cartesian products (involving a finite number of sets) preserve retractiveness. In particular, if $\{C_k^1\},\ldots,\{C_k^r\}$ are retractive nested sequences of nonempty closed convex sets, the sequences $\{N_k\}$ and $\{T_k\}$ are retractive, where

$$N_k = C_k^1 \cap C_k^2 \cap \cdots \cap C_k^r, \qquad T_k = C_k^1 \times C_k^2 \times \cdots \times C_k^r, \qquad \forall\ k,$$

and we assume that all the sets $N_k$ are nonempty. A simple consequence is that a polyhedral set is retractive, since it is the nonempty intersection of a finite number of closed halfspaces.

**Proposition 1.4.9:** A polyhedral set is retractive.

The importance of retractive sequences is motivated by the following proposition.

**Proposition 1.4.10:** A retractive nested sequence of nonempty closed convex sets has nonempty intersection.

**Proposition 1.4.11:** Let $\{C_k\}$ be a nested sequence of nonempty closed convex sets. Denote

$$R = \cap_{k=0}^{\infty} R_{C_k}, \qquad L = \cap_{k=0}^{\infty} L_{C_k}.$$

(a) If $R = L$, then $\{C_k\}$ is retractive, and $\cap_{k=0}^{\infty} C_k$ is nonempty. Furthermore,

$$\cap_{k=0}^{\infty} C_k = L + \tilde{C},$$

where $\tilde{C}$ is some nonempty and compact set.

(b) Let $X$ be a retractive closed convex set. Assume that all the sets $\overline{C}_k = X \cap C_k$ are nonempty, and that

$$R_X \cap R \subset L.$$

Then, $\{\overline{C}_k\}$ is retractive, and $\cap_{k=0}^{\infty} \overline{C}_k$ is nonempty.

**Proposition 1.4.12: (Existence of Solutions of Convex Quadratic Programs)** Let $Q$ be a symmetric positive semidefinite $n \times n$ matrix, let $c$ and $a_1, \ldots, a_r$ be vectors in $\Re^n$, and let $b_1, \ldots, b_r$ be scalars. Assume that the optimal value of the problem

$$\begin{aligned} &\text{minimize} \quad x'Qx + c'x \\ &\text{subject to} \quad a_j'x \le b_j, \quad j = 1, \ldots, r, \end{aligned}$$

is finite. Then the problem has at least one optimal solution.

### Closedness under Linear Transformation and Vector Sum

The conditions of Prop. 1.4.11 can be translated to conditions guaranteeing the closedness of the image, $AC$, of a closed convex set $C$ under a linear transformation $A$.

**Proposition 1.4.13:** Let $X$ and $C$ be nonempty closed convex sets in $\Re^n$, and let $A$ be an $m \times n$ matrix with nullspace denoted by $N(A)$. If $X$ is a retractive closed convex set and

$$R_X \cap R_C \cap N(A) \subset L_C,$$

then $A(X \cap C)$ is a closed set.

A special case is the following.

**Proposition 1.4.14:** Let $C_1, \ldots, C_m$ be nonempty closed convex subsets of $\Re^n$ such that the equality $d_1 + \cdots + d_m = 0$ for some vectors $d_i \in R_{C_i}$ implies that $d_i \in L_{C_i}$ for all $i = 1, \ldots, m$. Then $C_1 + \cdots + C_m$ is a closed set.

When specialized to just two sets, the above proposition implies that if $C_1$ and $-C_2$ are closed convex sets, then $C_1 - C_2$ is closed if there is no common nonzero direction of recession of $C_1$ and $C_2$, i.e.

$$R_{C_1} \cap R_{C_2} = \{0\}.$$

This is true in particular if either $C_1$ or $C_2$ is bounded, in which case either $R_{C_1} = \{0\}$ or $R_{C_2} = \{0\}$, respectively.

Some other conditions asserting the closedness of vector sums can be derived from Prop. 1.4.13. For example, we can show that the vector sum of a finite number of polyhedral sets is closed, since it can be viewed as the image of their Cartesian product (clearly a polyhedral set) under a linear transformation. Another useful result is that if $X$ is a polyhedral set, and $C$ is a closed convex set, then $X + C$ is closed if every direction of recession of $X$ whose opposite is a direction of recession of $C$ lies also in the lineality space of $C$. In particular, $X + C$ is closed if $X$ is polyhedral, and $C$ is closed.

## Section 1.5. Hyperplanes

A *hyperplane* in $\Re^n$ is a set of the form $\{x \mid a'x = b\}$, where $a$ is nonzero vector in $\Re^n$ and $b$ is a scalar. The sets

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

are called the *closed halfspaces* associated with the hyperplane (also referred to as the *positive and negative halfspaces*, respectively). The sets

$$\{x \mid a'x > b\}, \qquad \{x \mid a'x < b\},$$

are called the *open halfspaces* associated with the hyperplane.

**Proposition 1.5.1: (Supporting Hyperplane Theorem)** Let $C$ be a nonempty convex subset of $\Re^n$ and let $\overline{x}$ be a vector in $\Re^n$. If $\overline{x}$ is not an interior point of $C$, there exists a hyperplane that passes through $\overline{x}$ and contains $C$ in one of its closed halfspaces, i.e., there exists a vector $a \neq 0$ such that

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

**Proposition 1.5.2: (Separating Hyperplane Theorem)** Let $C_1$ and $C_2$ be two nonempty convex subsets of $\Re^n$. If $C_1$ and $C_2$ are disjoint, there exists a hyperplane that separates them, i.e., there exists a vector $a \neq 0$ such that

$$a'x_1 \leq a'x_2, \qquad \forall\ x_1 \in C_1,\ \forall\ x_2 \in C_2.$$

**Proposition 1.5.3: (Strict Separation Theorem)** Let $C_1$ and $C_2$ be two disjoint nonempty convex sets. There exists a hyperplane that strictly separates $C_1$ and $C_2$ under any one of the following five conditions:

(1) $C_2 - C_1$ is closed.

(2) $C_1$ is closed and $C_2$ is compact.

(3) $C_1$ and $C_2$ are polyhedral.

(4) $C_1$ and $C_2$ are closed, and

$$R_{C_1} \cap R_{C_2} = L_{C_1} \cap L_{C_2},$$

where $R_{C_i}$ and $L_{C_i}$ denote the recession cone and the lineality space of $C_i$, $i = 1, 2$.

(5) $C_1$ is closed, $C_2$ is polyhedral, and $R_{C_1} \cap R_{C_2} \subset L_{C_1}$.

The notion of direction of recession of a convex function $f$ can be described in terms of its epigraph via the following proposition.

**Proposition 1.5.4:** The closure of the convex hull of a set $C$ is the intersection of the closed halfspaces that contain $C$. In particular, a closed convex set is the intersection of the closed halfspaces that contain it.

Let $C_1$ and $C_2$ be two subsets of $\Re^n$. We say that a hyperplane *properly separates $C_1$ and $C_2$* if it separates $C_1$ and $C_2$, and does not fully contain both $C_1$ and $C_2$. If $C$ is a subset of $\Re^n$ and $\overline{x}$ is a vector in $\Re^n$, we say that a hyperplane *properly separates $C$ and $\overline{x}$* if it properly separates $C$ and the singleton set $\{\overline{x}\}$.

**Proposition 1.5.5: (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 1.5.6: (Proper Separation of Two Convex Sets)** Let $C_1$ and $C_2$ be two nonempty convex subsets of $\Re^n$. There exists a hyperplane that properly separates $C_1$ and $C_2$ if and only if

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

**Proposition 1.5.7: (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

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

Consider hyperplanes in $\Re^{n+1}$ and associate them with nonzero vectors of the form $(\mu, \beta)$, where $\mu \in \Re^n$ and $\beta \in \Re$. We say that such a hyperplane is *vertical* if $\beta = 0$.

**Proposition 1.5.8: (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 $\gamma$ 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 $C$.

**Section 1.6. Conjugate Functions**

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

$$f^\star(y) = \sup_{x \in \Re^n} \big\{ x'y - f(x) \big\}, \qquad y \in \Re^n. \tag{B.1}$$

**Proposition 1.6.1: (Conjugacy Theorem)** Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function, let $f^\star$ be its conjugate, and consider the double conjugate $f^{\star\star}$. Then:

(a) We have

$$f(x) \geq f^{\star\star}(x), \qquad \forall\ x \in \Re^n.$$

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

(c) If $f$ is closed proper convex, then

$$f(x) = f^{\star\star}(x), \qquad \forall\ x \in \Re^n.$$

(d) The conjugates of $f$ and its convex closure $\check{\mathrm{cl}}\, f$ are equal. Furthermore, if $\check{\mathrm{cl}}\, f$ is proper, then

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

### Positively Homogeneous Functions and Support Functions

Given a nonempty set $X$, consider the *indicator function* of $X$, defined by

$$\delta_X(x) = \begin{cases} 0 & \text{if } x \in X, \\ \infty & \text{if } x \notin X. \end{cases}$$

The conjugate of $\delta_X$ is given by

$$\sigma_X(y) = \sup_{x \in X} y'x$$

and is called the *support function of* $X$.

Let $C$ be a convex cone. The conjugate of its indicator function $\delta_C$ is its support function,

$$\sigma_C(y) = \sup_{x \in C} y'x.$$

The support/conjugate function $\sigma_C$ is the indicator function $\delta_{C^*}$ of the cone

$$C^* = \{y \mid y'x \le 0,\ \forall\, x \in C\},$$

called the *polar cone of* $C$. By the Conjugacy Theorem [Prop. 1.6.1(c)] $C^*$ is equal to $\mathrm{cl}\,\delta_C$. Thus the polar cone of $C^*$ is $\mathrm{cl}(C)$. In particular, if $C$ is closed, the polar of its polar is equal to the original. This is a special case of the *Polar Cone Theorem*, given in Section 2.2.

A function $f : \Re^n \mapsto [-\infty, \infty]$ is called *positively homogeneous* if its epigraph is a cone in $\Re^{n+1}$. Clearly, the support function $\sigma_X$ of a nonempty set $X$ is closed proper convex and positively homogeneous. The reverse is also true, namely that the closure of any proper convex positively homogeneous function is the support function of some set. In particular, we claim that if $\sigma : \Re^n \mapsto (-\infty, \infty]$ is a proper convex positively homogeneous function, then the conjugate of $\sigma$ is the indicator function of the closed convex set

$$X = \big\{x \mid y'x \le \sigma(y),\ \forall\, y \in \Re^n\big\},$$

and $\mathrm{cl}\,\sigma$ is the support function of $X$. For a proof, let $\delta$ be the conjugate of $\sigma$:

$$\delta(x) = \sup_{y \in \Re^n} \big\{y'x - \sigma(y)\big\}.$$

Since $\sigma$ is positively homogeneous, we have for any $\gamma > 0$,

$$\gamma\, \delta(x) = \sup_{y \in \Re^n} \big\{\gamma y'x - \gamma\, \sigma(y)\big\} = \sup_{y \in \Re^n} \big\{(\gamma y)'x - \sigma(\gamma y)\big\}.$$

The right-hand sides of the preceding two relations are equal, so we obtain

$$\delta(x) = \gamma\, \delta(x), \qquad \forall\ \gamma > 0,$$

which implies that $\delta$ takes only the values 0 and $\infty$ (since $\sigma$ and hence also its conjugate $\delta$ is proper). Thus, $\delta$ is the indicator function of a set, call it $X$, and we have

$$\begin{aligned} X &= \big\{x \mid \delta(x) \le 0\big\} \\ &= \left\{x \;\Big|\; \sup_{y\in\Re^n} \big\{y'x - \sigma(y)\big\} \le 0\right\} \\ &= \big\{x \mid y'x \le \sigma(y),\ \forall\ y \in \Re^n\big\}. \end{aligned}$$

Finally, since $\delta$ is the conjugate of $\sigma$, we see that $\mathrm{cl}\,\sigma$ is the conjugate of $\delta$; cf. the Conjugacy Theorem [Prop. 1.6.1(c)]. Since $\delta$ is the indicator function of $X$, it follows that $\mathrm{cl}\,\sigma$ is the support function of $X$.

We now discuss a characterization of the support function of the 0-level set of a closed proper convex function $f : \Re^n \mapsto (-\infty, \infty]$. The closure of the cone generated by $\mathrm{epi}(f)$, is the epigraph of a closed convex positively homogeneous function, called the *closed function generated by* $f$, and denoted by $\mathrm{gen}\, f$. The epigraph of $\mathrm{gen}\, f$ is the intersection of all the halfspaces that contain $\mathrm{epi}(f)$ and contain 0 in their boundary. Alternatively, the epigraph of $\mathrm{gen}\, f$ is the intersection of all the closed cones that contain $\mathrm{epi}(f)$.

We first note that if the level set $\big\{y \mid h(y) \le 0\big\}$ is nonempty, the generated function $\mathrm{gen}\, f$ is proper. Indeed if $(\mathrm{cl}\, f)(0) > 0$, by the Nonvertical Hyperplane Theorem (Prop. 1.5.8), there exists a nonvertical hyperplane passing through the origin and containing $\mathrm{epi}(f)$ in one of its closed halfspaces, implying that the epigraph of $\mathrm{gen}\, f$ does not contain a line, so $\mathrm{gen}\, f$ is proper. Any $y$ such that $f^\star(y) \le 0$, or equivalently $y'x \le f(x)$ for all $x$, defines a nonvertical hyperplane that separates the origin from $\mathrm{epi}(f)$.

Now let $f^\star$ be the conjugate of $f$. We will show that if the level set $\big\{y \mid f^\star(y) \le 0\big\}$ [respectively $\big\{x \mid f(x) \le 0\big\}$] is nonempty, its support function is the closed function generated by $f$ (respectively $f^\star$). Indeed, let $\sigma$ be the closed function generated by $f$. Then, since $\big\{y \mid f^\star(y) \le 0\big\} \ne \emptyset$, $\sigma$ is proper, and by our earlier analysis, $\sigma$ is the support function of the set

$$Y = \big\{y \mid y'x \le \sigma(x),\ \forall\ x \in \Re^n\big\}.$$

Since $\mathrm{epi}(\sigma)$ is the intersection of all the halfspaces that contain $\mathrm{epi}(f)$ and contain 0 in their boundary, the set $Y$ can be written as

$$Y = \big\{y \mid y'x \le f(x),\ \forall\ x \in \Re^n\big\},$$

from which we obtain

$$Y = \left\{y \;\Big|\; \sup_{x\in\Re^n} \big\{y'x - f(x)\big\} \le 0\right\} = \big\{y \mid f^\star(y) \le 0\big\}.$$

Note that the method used to characterize the 0-level set of $f$ can be applied to any level set. In particular, a nonempty level set $L_\gamma = \big\{x \mid f(x) \leq \gamma\big\}$ is the 0-level set of the function $f_\gamma$ defined by $f_\gamma(x) = f(x) - \gamma$, and its support function is the closed function generated by $f_\gamma^\star$, the conjugate of $f_\gamma$, which is given by $f_\gamma^\star(y) = f^\star(y) + \gamma$.

## CHAPTER 2: Basic Concepts of Polyhedral Convexity

### Section 2.1. Extreme Points

In this chapter, we discuss polyhedral sets, i.e., nonempty sets specified by systems of a finite number of affine inequalities

$$a_j'x \leq b_j, \qquad j = 1, \ldots, r,$$

where $a_1, \ldots, a_r$ are vectors in $\Re^n$, and $b_1, \ldots, b_r$ are scalars.

Given a nonempty convex set $C$, a vector $x \in C$ is said to be an *extreme point* of $C$ if it does not lie strictly between the endpoints of any line segment contained in the set, i.e., if there do not exist vectors $y \in C$ and $z \in C$, with $y \neq x$ and $z \neq x$, and a scalar $\alpha \in (0,1)$ such that $x = \alpha y + (1-\alpha)z$.

**Proposition 2.1.1:** Let $C$ be a convex subset of $\Re^n$, and let $H$ be a hyperplane that contains $C$ in one of its closed halfspaces. Then the extreme points of $C \cap H$ are precisely the extreme points of $C$ that belong to $H$.

**Proposition 2.1.2:** A nonempty closed convex subset of $\Re^n$ has at least one extreme point if and only if it does not contain a line, i.e., a set of the form $\{x + \alpha d \mid \alpha \in \Re\}$, where $x$ and $d$ are vectors in $\Re^n$ with $d \neq 0$.

**Proposition 2.1.3:** Let $C$ be a nonempty closed convex subset of $\Re^n$. Assume that for some $m \times n$ matrix $A$ of rank $n$ and some $b \in \Re^m$, we have

$$Ax \geq b, \qquad \forall\ x \in C.$$

Then $C$ has at least one extreme point.

**Proposition 2.1.4:** Let $P$ be a polyhedral subset of $\Re^n$.

(a) If $P$ has the form

$$P = \{x \mid a_j'x \le b_j,\ j = 1, \ldots, r\},$$

where $a_j \in \Re^n$, $b_j \in \Re$, $j = 1, \ldots, r$, then a vector $v \in P$ is an extreme point of $P$ if and only if the set

$$A_v = \{a_j \mid a_j'v = b_j,\ j \in \{1, \ldots, r\}\}$$

contains $n$ linearly independent vectors.

(b) If $P$ has the form

$$P = \{x \mid Ax = b,\ x \ge 0\},$$

where $A$ is an $m \times n$ matrix and $b$ is a vector in $\Re^m$, then a vector $v \in P$ is an extreme point of $P$ if and only if the columns of $A$ corresponding to the nonzero coordinates of $v$ are linearly independent.

(c) If $P$ has the form

$$P = \{x \mid Ax = b,\ c \le x \le d\},$$

where $A$ is an $m \times n$ matrix, $b$ is a vector in $\Re^m$, and $c$, $d$ are vectors in $\Re^n$, then a vector $v \in P$ is an extreme point of $P$ if and only if the columns of $A$ corresponding to the coordinates of $v$ that lie strictly between the corresponding coordinates of $c$ and $d$ are linearly independent.

**Proposition 2.1.5:** A polyhedral set in $\Re^n$ of the form

$$\{x \mid a_j'x \le b_j,\ j = 1, \ldots, r\}$$

has an extreme point if and only if the set $\{a_j \mid j = 1, \ldots, r\}$ contains $n$ linearly independent vectors.

## Section 2.2. Polar Cones

We return to the notion of *polar cone* of nonempty set $C$, denoted by $C^*$, and given by

$$C^* = \{y \mid y'x \leq 0,\ \forall\, x \in C\}.$$

**Proposition 2.2.1:**

(a) For any nonempty set $C$, we have

$$C^* = \big(\mathrm{cl}(C)\big)^* = \big(\mathrm{conv}(C)\big)^* = \big(\mathrm{cone}(C)\big)^*.$$

(b) (*Polar Cone Theorem*) For any nonempty cone $C$, we have

$$(C^*)^* = \mathrm{cl}\big(\mathrm{conv}(C)\big).$$

In particular, if $C$ is closed and convex, we have $(C^*)^* = C$.

**Section 2.3. Polyhedral Sets and Functions**

We recall that a polyhedral cone $C \subset \Re^n$ has the form

$$C = \{x \mid a_j'x \leq 0,\ j = 1, \ldots, r\},$$

where $a_1, \ldots, a_r$ are some vectors in $\Re^n$, and $r$ is a positive integer. We say that a cone $C \subset \Re^n$ is *finitely generated*, if it is generated by a finite set of vectors, i.e., if it has the form

$$C = \mathrm{cone}\big(\{a_1, \ldots, a_r\}\big) = \left\{ x \;\middle|\; x = \sum_{j=1}^{r} \mu_j a_j,\ \mu_j \geq 0,\ j = 1, \ldots, r \right\},$$

where $a_1, \ldots, a_r$ are some vectors in $\Re^n$, and $r$ is a positive integer.

**Proposition 2.3.1: (Farkas' Lemma)** Let $a_1, \ldots, a_r$ be vectors in $\Re^n$. Then, $\{x \mid a_j'x \leq 0,\ j = 1, \ldots, r\}$ and $\mathrm{cone}\big(\{a_1, \ldots, a_r\}\big)$ are closed cones that are polar to each other.

**Proposition 2.3.2: (Minkowski-Weyl Theorem)** A cone is polyhedral if and only if it is finitely generated.

**Proposition 2.3.3: (Minkowski-Weyl Representation)** A set $P$ is polyhedral if and only if there is a nonempty finite set $\{v_1, \ldots, v_m\}$ and a finitely generated cone $C$ such that $P = \text{conv}\big(\{v_1, \ldots, v_m\}\big) + C$, i.e.,

$$P = \left\{ x \;\middle|\; x = \sum_{j=1}^{m} \mu_j v_j + y,\ \sum_{j=1}^{m} \mu_j = 1,\ \mu_j \geq 0,\ j = 1, \ldots, m,\ y \in C \right\}.$$

**Proposition 2.3.4: (Algebraic Operations on Polyhedral Sets)**

(a) The intersection of polyhedral sets is polyhedral, if it is nonempty.

(b) The Cartesian product of polyhedral sets is polyhedral.

(c) The image of a polyhedral set under a linear transformation is a polyhedral set.

(d) The vector sum of two polyhedral sets is polyhedral.

(e) The inverse image of a polyhedral set under a linear transformation is polyhedral.

We say that a function $f : \Re^n \mapsto (-\infty, \infty]$ is *polyhedral* if its epigraph is a polyhedral set in $\Re^{n+1}$. Note that a polyhedral function $f$ is, by definition, closed, convex, and also proper [since $f$ cannot take the value $-\infty$, and $\text{epi}(f)$ is closed, convex, and nonempty (based on our convention that only nonempty sets can be polyhedral)].

**Proposition 2.3.5:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a convex function. Then $f$ is polyhedral if and only if $\text{dom}(f)$ is a polyhedral set and

$$f(x) = \max_{j=1,\ldots,m} \{a_j' x + b_j\}, \qquad \forall\ x \in \text{dom}(f),$$

where $a_j$ are vectors in $\Re^n$, $b_j$ are scalars, and $m$ is a positive integer.

**Proposition 2.3.6:** The sum of two polyhedral functions $f_1$ and $f_2$, such that $\text{dom}(f_1) \cap \text{dom}(f_2) \neq \varnothing$, is a polyhedral function.

**Proposition 2.3.7:** If $A$ is a matrix and $g$ is a polyhedral function such that $\mathrm{dom}(g)$ contains a point in the range of $A$, the function $f$ given by $f(x) = g(Ax)$ is polyhedral.

### Section 2.4. Polyhedral Aspects of Optimization

Polyhedral convexity plays a very important role in optimization. The following are two basic results related to linear programming, the minimization of a linear function over a polyhedral set.

**Proposition 2.4.1:** Let $C$ be a closed convex subset of $\Re^n$ that has at least one extreme point. A concave function $f : C \mapsto \Re$ that attains a minimum over $C$ attains the minimum at some extreme point of $C$.

**Proposition 2.4.2: (Fundamental Theorem of Linear Programming)** Let $P$ be a polyhedral set that has at least one extreme point. A linear function that is bounded below over $P$ attains a minimum at some extreme point of $P$.

## CHAPTER 3: Basic Concepts of Convex Optimization

### Section 3.1. Constrained Optimization

Let us consider the problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X, \end{array}$$

where $f : \Re^n \mapsto (-\infty, \infty]$ is a function and $X$ is a nonempty subset of $\Re^n$. Any vector $x \in X \cap \mathrm{dom}(f)$ is said to be a *feasible solution* of the problem (we also use the terms *feasible vector* or *feasible point*). If there is at least one feasible solution, i.e., $X \cap \mathrm{dom}(f) \neq \emptyset$, we say that the problem is *feasible*; otherwise we say that the problem is *infeasible*. Thus, when $f$ is extended real-valued, we view only the points in $X \cap \mathrm{dom}(f)$ as candidates for optimality, and we view $\mathrm{dom}(f)$ as an implicit constraint set. Furthermore, feasibility of the problem is equivalent to $\inf_{x \in X} f(x) < \infty$.

We say that a vector $x^*$ is a *minimum of $f$ over $X$* if

$$x^* \in X \cap \text{dom}(f), \qquad \text{and} \qquad f(x^*) = \inf_{x \in X} f(x).$$

We also call $x^*$ a *minimizing point* or *minimizer* or *global minimum of $f$ over $X$*. Alternatively, we say that $f$ *attains a minimum over $X$ at $x^*$*, and we indicate this by writing

$$x^* \in \arg\min_{x \in X} f(x).$$

If $x^*$ is known to be the unique minimizer of $f$ over $X$, with slight abuse of notation, we also write

$$x^* = \arg\min_{x \in X} f(x).$$

We use similar terminology for maxima.

Given a subset $X$ of $\Re^n$ and a function $f : \Re^n \mapsto (-\infty, \infty]$, we say that a vector $x^*$ is a *local minimum of $f$ over $X$* if $x^* \in X \cap \text{dom}(f)$ and there exists some $\epsilon > 0$ such that

$$f(x^*) \leq f(x), \qquad \forall\, x \in X \text{ with } \|x - x^*\| < \epsilon.$$

A local minimum $x^*$ is said to be *strict* if there is no other local minimum within some open sphere centered at $x^*$. Local maxima are defined similarly.

**Proposition 3.1.1:** If $X$ is a convex subset of $\Re^n$ and $f : \Re^n \mapsto (-\infty, \infty]$ is a convex function, then a local minimum of $f$ over $X$ is also a global minimum. If in addition $f$ is strictly convex, then there exists at most one global minimum of $f$ over $X$.

## 3.2. Existence of Optimal Solutions

**Proposition 3.2.1: (Weierstrass' Theorem)** Consider a closed proper function $f : \Re^n \mapsto (-\infty, \infty]$, and assume that any one of the following three conditions holds:

(1) $\text{dom}(f)$ is bounded.

(2) There exists a scalar $\overline{\gamma}$ such that the level set

$$\{x \mid f(x) \leq \overline{\gamma}\}$$

is nonempty and bounded.

(3) $f$ is coercive, i.e., if for every sequence $\{x_k\}$ such that $\|x_k\| \to \infty$, we have $\lim_{k\to\infty} f(x_k) = \infty$..

Then the set of minima of $f$ over $\Re^n$ is nonempty and compact.

**Proposition 3.2.2:** Let $X$ be a closed convex subset of $\Re^n$, and let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed convex function with $X \cap \mathrm{dom}(f) \neq \emptyset$. The set of minima of $f$ over $X$ is nonempty and compact if and only if $X$ and $f$ have no common nonzero direction of recession.

**Proposition 3.2.3: (Existence of Solution, Sum of Functions)** Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be closed proper convex functions such that the function $f = f_1 + \cdots + f_m$ is proper. Assume that the recession function of a single function $f_i$ satisfies $r_{f_i}(d) = \infty$ for all $d \neq 0$. Then the set of minima of $f$ is nonempty and compact.

### Section 3.3. Partial Minimization of Convex Functions

Functions obtained by minimizing other functions partially, i.e., with respect to some of their variables, are especially useful in the treatment of duality and minimax theory. It is then useful to be able to deduce properties of the function obtained, such as convexity and closedness, from corresponding properties of the original.

**Proposition 3.3.1:** 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

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

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

**Proposition 3.3.2:** 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 closed proper convex. Furthermore, for each $x \in \mathrm{dom}(f)$, the set of minima in the definition of $f(x)$ is nonempty and compact.

**Proposition 3.3.3:** 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 real-valued convex function over $X$.

**Proposition 3.3.5:** 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 closed proper convex. Furthermore, for each $x \in \text{dom}(f)$, the set of minima in the definition of $f(x)$ is nonempty.

**Section 3.4. Saddle Point and Minimax Theory**

Let us consider a function $\phi : X \times Z \mapsto \Re$, where $X$ and $Z$ are nonempty subsets of $\Re^n$ and $\Re^m$, respectively. Of great interest is to derive conditions guaranteeing that

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z), \tag{B.2}$$

and that the infima and the suprema above are attained.

**Definition 3.4.1:** A pair of vectors $x^* \in X$ and $z^* \in Z$ is called a *saddle point* of $\phi$ if

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

**Proposition 3.4.1:** A pair $(x^*, z^*)$ is a saddle point of $\phi$ if and only if the minimax equality (B.2), and $x^*$ is an optimal solution of the problem

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

while $z^*$ is an optimal solution of the problem

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

## CHAPTER 4: Geometric Duality Framework

### Section 4.1. Min Common/Max Crossing Duality

We introduce a geometric framework for duality analysis, which aims to capture the most essential characteristics of duality 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 $\big\{(0,w) \mid w \geq 0\big\}$. We want to find the maximum crossing point of the $(n+1)$st axis with such a hyperplane.

We refer to the two problems as the *min common/max crossing (MC/MC) framework*, and we will show that it can be used to develop much of the core theory of convex 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.,

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

The max crossing problem is to maximize over all $\mu \in \Re^n$ the maximum crossing level corresponding to $\mu$, i.e.,

$$\begin{aligned} &\text{maximize} \quad \inf_{(u,w)\in M} \{w + \mu' u\} \\ &\text{subject to} \quad \mu \in \Re^n. \end{aligned} \tag{B.3}$$

We also refer to this as the *dual problem*, we denote by $q^*$ its optimal value,

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

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

**Proposition 4.1.1:** The dual function $q$ is concave and upper semicontinuous.

The following proposition states that we always have $q^* \leq w^*$; we refer to this as *weak duality*. When $q^* = w^*$, we say that *strong duality holds* or that *there is no duality gap*.

**Proposition 4.1.2: (Weak Duality Theorem)** We have $q^* \leq w^*$.

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 4.1.3:** 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 $\overline{M}$.

**Section 4.2. Some Special Cases**

There are several interesting special cases where the set $M$ is the epigraph of some function. For example, 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

$$f(x) = F(x, 0), \qquad \forall\ x \in \Re^n, \tag{B.4}$$

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

$$p(u) = \inf_{x \in \Re^n} F(x, u). \tag{B.5}$$

and consider the MC/MC framework with

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

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

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

The max crossing problem (B.3) can be written as

$$\begin{array}{ll} \text{maximize} & q(\mu) \\ \text{subject to} & \mu \in \Re^r, \end{array}$$

where the dual function is

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} = \inf_{u\in\Re^r} \big\{p(u) + \mu' u\big\} = \inf_{(x,u)\in\Re^{n+r}} \big\{F(x,u) + \mu' u\big\}. \tag{B.6}$$

Note that from Eq. (B.6), an alternative expression for $q$ is

$$q(\mu) = -\sup_{(x,u)\in\Re^{n+r}} \big\{-\mu' u - F(x,u)\big\} = -F^\star(0,-\mu),$$

where $F^\star$ is the conjugate of $F$, viewed as a function of $(x,u)$. Since

$$q^* = \sup_{\mu\in\Re^r} q(\mu) = -\inf_{\mu\in\Re^r} F^\star(0,-\mu) = -\inf_{\mu\in\Re^r} F^\star(0,\mu),$$

the strong duality relation $w^* = q^*$ can be written as

$$\inf_{x\in\Re^n} F(x,0) = -\inf_{\mu\in\Re^r} F^\star(0,\mu).$$

Different choices of function $F$, as in Eqs. (B.4) and (B.5), yield corresponding MC/MC frameworks and dual problems. An example of this type is minimization with inequality constraints:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X, \quad g(x) \le 0, \end{array} \tag{B.7}$$

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

$$C_u = \big\{x \in X \mid g(x) \le u\big\}, \qquad u \in \Re^r, \tag{B.8}$$

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 $x \in C_0$ [cf. Eq. (B.4)].

The function $p$ of Eq. (B.5) is given by

$$p(u) = \inf_{x\in\Re^n} F(x,u) = \inf_{x\in X,\, g(x)\le u} f(x), \tag{B.9}$$

and is known as the *primal function* or *perturbation function*. It captures the essential structure of the constrained minimization problem, relating to duality and other properties, such as sensitivity. Consider now the MC/MC framework corresponding to $M = \text{epi}(p)$. From Eq. (B.6), we obtain with some calculation

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

The following proposition derives the primal and dual functions in the minimax framework.

**Proposition 4.2.1:** 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. Assume that $(-\hat{\text{cl}}\, \phi)(x, \cdot)$ is proper for all $x \in X$, and consider the MC/MC framework corresponding to $M = \text{epi}(p)$, where $p$ is given by

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

Then the dual function is given by

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

## 4.3. Strong Duality Theorem

The following propositions give general results for strong duality.

**Proposition 4.3.1: (MC/MC Strong Duality)** 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 $w^* \leq \liminf_{k \to \infty} w_k$.

**Proposition 4.3.2:** Consider the MC/MC framework, assuming that $w^* < \infty$.

(a) Let $M$ be closed and convex. Then $q^* = w^*$. Furthermore, the function

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

is convex and its epigraph is the set

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

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

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

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 $\mathrm{conv}(M)$.

**Section 4.4. Existence of Dual Optimal Solutions**

The following propositions give general results for strong duality, as well existence of dual optimal solutions.

**Proposition 4.4.1: (MC/MC Existence of Max Crossing Solutions)** Consider the MC/MC framework and assume the following:

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

(2) The set

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

is convex.

(3) The origin is a relative interior point of the set

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

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

**Proposition 4.4.2:** Let the assumptions of Prop. 4.4.1 hold. Then $Q^*$, the set of optimal solutions 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 an interior point of $D$.

## Section 4.5. Duality and Polyhedral Convexity

The following propositions address special cases where the set $M$ has partially polyhedral structure.

**Proposition 4.5.1:** Consider the MC/MC framework, 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 $\tilde{M}$ and $P$ are convex sets.

(3) Either $\mathrm{ri}(\tilde{D}) \cap \mathrm{ri}(P) \neq \emptyset$, or $P$ is polyhedral and $\mathrm{ri}(\tilde{D}) \cap P \neq \emptyset$, where $\tilde{D}$ is the set given by

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

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

**Proposition 4.5.2:** Consider the MC/MC framework, and assume that:

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

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

# CHAPTER 5: Duality and Optimization

## Section 5.1. Nonlinear Farkas' Lemma

A nonlinear version of Farkas' Lemma 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 use the following assumption.

**Assumption 5.1:** 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 5.1.1: (Nonlinear Farkas' Lemma)** Let Assumption 5.1 hold and let $Q^*$ be the subset of $\Re^r$ given by

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

Assume that one of the following two conditions holds:

(1) There exists $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all $j = 1, \ldots, r$.

(2) The functions $g_j$, $j = 1, \ldots, r$, are affine, and there exists $\overline{x} \in \text{ri}(X)$ such that $g(\overline{x}) \leq 0$.

Then $Q^*$ is nonempty, and under condition (1) it is also compact.

By selecting $f$ and $g_j$ to be linear, and $X$ to be the entire space in the Nonlinear Farkas' Lemma, we obtain a version of Farkas' Lemma (cf. Section 2.3) as a special case.

**Proposition 5.1.2: (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

$$A'x \leq 0 \qquad \Rightarrow \qquad c'x \leq 0.$$

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

$$A'x = 0,\ x \geq 0 \qquad \Rightarrow \qquad c'x \leq 0.$$

**Section 5.2. Linear Programming Duality**

One of the most important results in optimization is the linear programming duality theorem. Consider the problem

$$\begin{aligned} &\text{minimize} \ \ c'x \\ &\text{subject to} \ \ a_j'x \geq b_j, \quad j = 1, \ldots, r, \end{aligned}$$

where $c \in \Re^n$, $a_j \in \Re^n$, and $b_j \in \Re$, $j = 1, \ldots, r$. We refer to this as the *primal problem*. We consider the *dual problem*

$$\begin{aligned} &\text{maximize} \ \ b'\mu \\ &\text{subject to} \ \ \sum_{j=1}^{r} a_j \mu_j = c, \quad \mu \geq 0, \end{aligned}$$

which can be derived from the MC/MC duality framework in Section 4.2. We denote the primal and dual optimal values by $f^*$ and $q^*$, respectively.

**Proposition 5.2.1: (Linear Programming Duality Theorem)**

(a) If either $f^*$ or $q^*$ is finite, then $f^* = q^*$ and both the primal and the dual problem have optimal solutions.

(b) If $f^* = -\infty$, then $q^* = -\infty$.

(c) If $q^* = \infty$, then $f^* = \infty$.

Another related result is the following necessary and sufficient condition for primal and dual optimality.

**Proposition 5.2.2: (Linear Programming Optimality Conditions)** A pair of vectors $(x^*, \mu^*)$ form a primal and dual optimal solution pair if and only if $x^*$ is primal-feasible, $\mu^*$ is dual-feasible, and

$$\mu_j^*(b_j - a_j' x^*) = 0, \qquad \forall\ j = 1, \ldots, r.$$

**Section 5.3. Convex Programming Duality**

We first focus on the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \quad g(x) \le 0, \end{aligned} \tag{B.10}$$

Consider the Lagrangian function

$$L(x, \mu) = f(x) + \mu' g(x), \qquad x \in X,\ \mu \in \Re^r,$$

and the dual problem

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

For this and other similar problems, we denote the primal and dual optimal values by $f^*$ and $q^*$, respectively.

**Proposition 5.3.1: (Convex Programming Duality - Existence of Dual Optimal Solutions)** Consider problem (B.10). Assume that $f^*$ is finite, and that one of the following two conditions holds:

(1) There exists $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all $j = 1, \ldots, r$.

(2) The functions $g_j$, $j = 1, \ldots, r$, are affine, and there exists $\overline{x} \in \mathrm{ri}(X)$ such that $g(\overline{x}) \leq 0$.

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

**Proposition 5.3.2: (Optimality Conditions)** Consider problem (B.10). There holds $q^* = f^*$, and $(x^*, \mu^*)$ are a primal and dual optimal solution pair if and only if $x^*$ is feasible, $\mu^* \geq 0$, and

$$x^* \in \arg\min_{x \in X} L(x, \mu^*), \qquad \mu_j^* g_j(x^*) = 0, \quad j = 1, \ldots, r.$$

The analysis for problem (B.10) can be refined by making more specific assumptions regarding available polyhedral structure in the constraint functions and the abstract constraint set $X$. Here is an extension of problem (B.10) where there are additional linear equality constraints:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \ \ x \in X, \quad g(x) \leq 0, \quad Ax = b, \end{aligned} \tag{B.11}$$

where $X$ is a convex set, $g(x) = \big(g_1(x), \ldots, g_r(x)\big)'$, $f : X \mapsto \Re$ and $g_j : X \mapsto \Re$, $j = 1, \ldots, r$, are convex functions, $A$ is an $m \times n$ matrix, and $b \in \Re^m$. The corresponding Lagrangian function is

$$L(x, \mu, \lambda) = f(x) + \mu' g(x) + \lambda'(Ax - b),$$

and the dual problem is

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

In the special case of a problem with just linear equality constraints:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \ \ x \in X, \quad Ax = b, \end{aligned} \tag{B.12}$$

the Lagrangian function is

$$L(x, \lambda) = f(x) + \lambda'(Ax - b),$$

and the dual problem is

$$\begin{array}{ll} \text{maximize} & \inf_{x \in X} L(x, \lambda) \\ \text{subject to} & \lambda \in \Re^m. \end{array}$$

**Proposition 5.3.3: (Convex Programming - Linear Equality Constraints)** Consider problem (B.12).

(a) Assume that $f^*$ is finite and that there exists $\overline{x} \in \text{ri}(X)$ such that $A\overline{x} = b$. Then $f^* = q^*$ and there exists at least one dual optimal solution.

(b) There holds $f^* = q^*$, and $(x^*, \lambda^*)$ are a primal and dual optimal solution pair if and only if $x^*$ is feasible and

$$x^* \in \arg\min_{x \in X} L(x, \lambda^*).$$

**Proposition 5.3.4: (Convex Programming - Linear Equality and Inequality Constraints)** Consider problem (B.11).

(a) Assume that $f^*$ is finite, that the functions $g_j$ are linear, and that there exists $\overline{x} \in \text{ri}(X)$ such that $A\overline{x} = b$ and $g(\overline{x}) \leq 0$. Then $q^* = f^*$ and there exists at least one dual optimal solution.

(b) There holds $f^* = q^*$, and $(x^*, \mu^*, \lambda^*)$ are a primal and dual optimal solution pair if and only if $x^*$ is feasible, $\mu^* \geq 0$, and

$$x^* \in \arg\min_{x \in X} L(x, \mu^*, \lambda^*), \qquad \mu_j^* g_j(x^*) = 0, \quad j = 1, \ldots, r.$$

**Proposition 5.3.5: (Convex Programming - Linear Equality and Nonlinear Inequality Constraints)** Consider problem (B.11). Assume that $f^*$ is finite, that there exists $\overline{x} \in X$ such that $A\overline{x} = b$ and $g(\overline{x}) < 0$, and that there exists $\tilde{x} \in \text{ri}(X)$ such that $A\tilde{x} = b$. Then $q^* = f^*$ and there exists at least one dual optimal solution.

**Proposition 5.3.6: (Convex Programming - Mixed Polyhedral and Nonpolyhedral Constraints)** Consider problem (B.11), where $X$ is the intersection of a polyhedral set $P$ and a convex set $C$,

$$X = P \cap C,$$

$g(x) = \big(g_1(x), \ldots, g_r(x)\big)'$, the functions $f : \Re^n \mapsto \Re$ and $g_j : \Re^n \mapsto \Re$, $j = 1, \ldots, r$, are defined over $\Re^n$, $A$ is an $m \times n$ matrix, and $b \in \Re^m$. Assume that $f^*$ is finite and that for some $\overline{r}$ with $1 \leq \overline{r} \leq r$, the functions $g_j$, $j = 1, \ldots, \overline{r}$, are polyhedral, and the functions $f$ and $g_j$, $j = \overline{r} + 1, \ldots, r$, are convex over $C$. Assume further that:

(1) There exists a vector $\tilde{x} \in \mathrm{ri}(C)$ in the set

$$\tilde{P} = P \cap \big\{x \mid Ax = b,\, g_j(x) \leq 0,\, j = 1, \ldots, \overline{r}\big\}.$$

(2) There exists $\overline{x} \in \tilde{P} \cap C$ such that $g_j(\overline{x}) < 0$ for all $j = \overline{r}+1, \ldots, r$.

Then $q^* = f^*$ and there exists at least one dual optimal solution.

We will now give a different type of result, which under some compactness assumptions, guarantees strong duality and that there exists an optimal primal solution (even if there may be no dual optimal solution).

**Proposition 5.3.7: (Convex Programming Duality - Existence of Primal Optimal Solutions)** Assume that problem (B.10) is feasible, that the convex functions $f$ and $g_j$ are closed, and that the function

$$F(x, 0) = \begin{cases} f(x) & \text{if } g(x) \leq 0,\ x \in X, \\ \infty & \text{otherwise,} \end{cases}$$

has compact level sets. Then $f^* = q^*$ and the set of optimal solutions of the primal problem is nonempty and compact.

We now consider another important optimization framework, the problem

$$\begin{aligned} &\text{minimize} \quad f_1(x) + f_2(Ax) \\ &\text{subject to} \ \ x \in \Re^n, \end{aligned} \tag{B.13}$$

where $A$ is an $m \times n$ matrix, $f_1 : \Re^n \mapsto (-\infty, \infty]$ and $f_2 : \Re^m \mapsto (-\infty, \infty]$ are closed convex functions. We assume that there exists a feasible solution.

**Proposition 5.3.8: (Fenchel Duality)**

(a) If $f^*$ is finite and $\big(A \cdot \mathrm{ri}\big(\mathrm{dom}(f_1)\big)\big) \cap \mathrm{ri}\big(\mathrm{dom}(f_2)\big) \neq \emptyset$, then $f^* = q^*$ and there exists at least one dual optimal solution.

(b) There holds $f^* = q^*$, and $(x^*, \lambda^*)$ is a primal and dual optimal solution pair if and only if

$$x^* \in \arg\min_{x \in \Re^n} \big\{ f_1(x) - x' A' \lambda^* \big\} \quad \text{and} \quad A x^* \in \arg\min_{z \in \Re^n} \big\{ f_2(z) + z' \lambda^* \big\}. \tag{B.14}$$

An important special case of Fenchel duality involves the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in C, \end{aligned} \tag{B.15}$$

where $f : \Re^n \mapsto (-\infty, \infty]$ is a closed proper convex function and $C$ is a closed convex cone in $\Re^n$. This is known as a *conic program*, and some of its special cases (semidefinite programming, second order cone programming) have many practical applications.

**Proposition 5.3.9: (Conic Duality Theorem)** Assume that the optimal value of the primal conic problem (B.15) is finite, and that $\mathrm{ri}\big(\mathrm{dom}(f)\big) \cap \mathrm{ri}(C) \neq \emptyset$. Then, there is no duality gap and that the dual problem

$$\begin{aligned} &\text{minimize} \quad f^\star(\lambda) \\ &\text{subject to} \quad \lambda \in \hat{C}, \end{aligned}$$

where $f^\star$ is the conjugate of $f$ and

$$\hat{C} = -C^* = \{\lambda \mid \lambda' x \geq 0,\ \forall\ x \in C\}.$$

has an optimal solution.

## Section 5.4. Subgradients and Optimality Conditions

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 $x \in \mathrm{dom}(f)$ if

$$f(z) \geq f(x) + g'(z - x), \qquad \forall\ z \in \Re^n. \tag{B.16}$$

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 $x \notin \text{dom}(f)$. Generally, $\partial f(x)$ is closed and convex, since based on the subgradient inequality (B.16), it is the intersection of a collection of closed halfspaces. Note that we restrict attention to proper functions (subgradients are not useful and make no sense for improper functions).

**Proposition 5.4.1:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function. For every $x \in \text{ri}\big(\text{dom}(f)\big)$,

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

where $S$ is the subspace that is parallel to the affine hull of $\text{dom}(f)$, and $G$ is a nonempty convex and compact set. In particular, if $x \in \text{int}\big(\text{dom}(f)\big)$, then $\partial f(x)$ is nonempty and compact.

It follows from the preceding proposition that *if $f$ is real-valued, then $\partial f(x)$ is nonempty and compact for all $x \in \Re^n$*. An important property is that *if $f$ is differentiable at some $x \in \text{int}\big(\text{dom}(f)\big)$, its gradient $\nabla f(x)$ is the unique subgradient at $x$.*

**Proposition 5.4.2: (Subdifferential Boundedness and Lipschitz Continuity)** Let $f : \Re^n \mapsto \Re$ be a real-valued convex function, and let $X$ be a nonempty compact subset of $\Re^n$.

(a) The set $\cup_{x \in X} \partial f(x)$ is nonempty and bounded.

(b) The function $f$ is Lipschitz continuous over $X$, i.e., there exists a scalar $L$ such that

$$\big|f(x) - f(z)\big| \leq L\, \|x - z\|, \qquad \forall\ x, z \in X.$$

### 5.4.1 Subgradients of Conjugate Functions

We will now derive an important relation between the subdifferentials of a proper convex function $f : \Re^n \mapsto (-\infty, \infty]$ and its conjugate $f^\star$. Using the definition of conjugacy, we have

$$x'y \leq f(x) + f^\star(y), \qquad \forall\ x \in \Re^n,\ y \in \Re^n.$$

This is known as the *Fenchel inequality*. A pair $(x, y)$ satisfies this inequality as an equation if and only if $x$ attains the supremum in the definition

$$f^\star(y) = \sup_{z \in \Re^n} \big\{y'z - f(z)\big\}.$$

Pairs of this type are connected with the subdifferentials of $f$ and $f^\star$, as shown in the following proposition.

**Proposition 5.4.3: (Conjugate Subgradient Theorem)** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function and let $f^\star$ be its conjugate. The following two relations are equivalent for a pair of vectors $(x, y)$:

(i) $x'y = f(x) + f^\star(y)$.

(ii) $y \in \partial f(x)$.

If in addition $f$ is closed, the relations (i) and (ii) are equivalent to

(iii) $x \in \partial f^\star(y)$.

For an application of the Conjugate Subgradient Theorem, note that the necessary and sufficient optimality condition (B.14) in the Fenchel Duality Theorem can be equivalently written as

$$A'\lambda^* \in \partial f_1(x^*), \qquad \lambda^* \in -\partial f_2(Ax^*).$$

The following proposition gives some useful corollaries of the Conjugate Subgradient Theorem:

**Proposition 5.4.4:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function and let $f^\star$ be its conjugate.

(a) $f^\star$ is differentiable at a vector $y \in \operatorname{int}\big(\operatorname{dom}(f^\star)\big)$ if and only if the supremum of $x'y - f(x)$ over $x \in \Re^n$ is uniquely attained.

(b) The set of minima of $f$ is given by

$$\arg\min_{x \in \Re^n} f(x) = \partial f^\star(0),$$

### 5.4.2 Subdifferential Calculus

We now give generalizations of some of the basic theorems of ordinary differentiation.

**Proposition 5.4.5: (Chain Rule)** Let $f : \Re^m \mapsto (-\infty, \infty]$ be a convex function, let $A$ be an $m \times n$ matrix, and assume that the function $F$ given by

$$F(x) = f(Ax)$$

is proper. Then

$$\partial F(x) \supset A' \partial f(Ax), \qquad \forall\ x \in \Re^n.$$

Furthermore, if either $f$ is polyhedral or else the range of $A$ contains a point in the relative interior of $\mathrm{dom}(f)$, we have

$$\partial F(x) = A' \partial f(Ax), \qquad \forall\ x \in \Re^n.$$

As a special case of Prop. 5.4.5, we obtain the following.

**Proposition 5.4.6: (Subdifferential of Sum of Functions)** Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be convex functions, and assume that the function $F = f_1 + \cdots + f_m$ is proper. Then

$$\partial F(x) \supset \partial f_1(x) + \cdots + \partial f_m(x), \qquad \forall\ x \in \Re^n.$$

Furthermore, if $\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i)\big) \neq \emptyset$, we have

$$\partial F(x) = \partial f_1(x) + \cdots + \partial f_m(x), \qquad \forall\ x \in \Re^n.$$

More generally, the same is true if for some $\overline{m}$ with $1 \leq \overline{m} \leq m$, the functions $f_i$, $i = 1, \ldots, \overline{m}$, are polyhedral and

$$\Big( \cap_{i=1}^{\overline{m}} \mathrm{dom}(f_i) \Big) \cap \Big( \cap_{i=\overline{m}+1}^{m} \mathrm{ri}\big(\mathrm{dom}(f_i)\big) \Big) \neq \emptyset.$$

### 5.4.3 Optimality Conditions

It can be seen from the definition of subgradient that a vector $x^*$ minimizes $f$ over $\Re^n$ if and only if $0 \in \partial f(x^*)$. We give the following generalization of this condition to constrained problems.

**Proposition 5.4.7:** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function, let $X$ be a nonempty convex subset of $\Re^n$, and assume that one of the following four conditions holds:

(1) $\text{ri}\big(\text{dom}(f)\big) \cap \text{ri}(X) \neq \emptyset$.

(2) $f$ is polyhedral and $\text{dom}(f) \cap \text{ri}(X) \neq \emptyset$.

(3) $X$ is polyhedral and $\text{ri}\big(\text{dom}(f)\big) \cap X \neq \emptyset$.

(4) $f$ and $X$ are polyhedral, and $\text{dom}(f) \cap X \neq \emptyset$.

Then, a vector $x^*$ minimizes $f$ over $X$ if and only if there exists $g \in \partial f(x^*)$ such that $-g$ belongs to the normal cone $N_X(x^*)$, or equivalently,

$$g'(x - x^*) \geq 0, \qquad \forall\, x \in X. \tag{B.17}$$

When $f$ is real-valued, the relative interior condition (1) of the preceding proposition is automatically satisfied [we have $\text{dom}(f) = \Re^n$]. If in addition, $f$ is differentiable, the optimality condition (B.17) reduces to the one of Prop. 1.1.8:

$$\nabla f(x^*)'(x - x^*) \geq 0, \qquad \forall\, x \in X.$$

### 5.4.4 Directional Derivatives

For a proper convex function $f : \Re^n \mapsto (-\infty, \infty]$, the directional derivative at any $x \in \text{dom}(f)$ in a direction $d \in \Re^n$, is defined by

$$f'(x; d) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha d) - f(x)}{\alpha}. \tag{B.18}$$

An important fact here is that the ratio in Eq. (B.18) is monotonically nonincreasing as $\alpha \downarrow 0$, so that the limit above is well-defined. To verify this, note that for any $\overline{\alpha} > 0$, the convexity of $f$ implies that for all $\alpha \in (0, \overline{\alpha})$,

$$f(x + \alpha d) \leq \frac{\alpha}{\overline{\alpha}} f(x + \overline{\alpha} d) + \left(1 - \frac{\alpha}{\overline{\alpha}}\right) f(x) = f(x) + \frac{\alpha}{\overline{\alpha}}\big(f(x + \overline{\alpha} d) - f(x)\big),$$

so that

$$\frac{f(x + \alpha d) - f(x)}{\alpha} \leq \frac{f(x + \overline{\alpha} d) - f(x)}{\overline{\alpha}}, \qquad \forall\, \alpha \in (0, \overline{\alpha}). \tag{B.19}$$

Thus the limit in Eq. (B.18) is well-defined (as a real number, or $\infty$, or $-\infty$) and an alternative definition of $f'(x; d)$ is

$$f'(x; d) = \inf_{\alpha > 0} \frac{f(x + \alpha d) - f(x)}{\alpha}, \qquad d \in \Re^n. \tag{B.20}$$

The directional derivative is related to the support function of the subdifferential $\partial f(x)$, as indicated in the following proposition.

**Proposition 5.4.8: (Support Function of the Subdifferential)** Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function, and let $(\mathrm{cl}\, f')(x; \cdot)$ be the closure of the directional derivative $f'(x; \cdot)$.

(a) For all $x \in \mathrm{dom}(f)$ such that $\partial f(x)$ is nonempty, $(\mathrm{cl}\, f')(x; \cdot)$ is the support function of $\partial f(x)$.

(b) For all $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$, $f'(x; \cdot)$ is closed and it is the support function of $\partial f(x)$.

**Section 5.5. Minimax Theory**

We will now provide theorems regarding the validity of the minimax equality and the existence of saddle points. These theorems are obtained by specializing the MC/MC theorems of Chapter 4. We will assume throughout this section the following.

**Assumption 5.5.0: (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 $x \in X$.

**Proposition 5.5.1:** Assume that the function $p$ given by

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

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 $u = 0$.

**Proposition 5.5.2:** Assume that $0 \in \mathrm{ri}\big(\mathrm{dom}(p)\big)$ and $p(0) > -\infty$. 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 $\mathrm{dom}(p)$.

**Proposition 5.5.3: (Classical Saddle Point Theorem)** Let the sets $X$ and $Z$ be compact. Then the set of saddle points of $\phi$ is nonempty and compact.

To formulate more general saddle point theorems, 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

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

The next two propositions provide 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 5.5.4:** Assume that $t$ is proper and that the level sets $\big\{x \mid t(x) \leq \gamma\big\}$, $\gamma \in \Re$, 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.

**Proposition 5.5.5:** Assume 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.

**Proposition 5.5.6:** Assume that either $t$ is proper or $r$ is proper.

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

**Proposition 5.5.7: (Saddle Point Theorem)** The set of saddle points of $\phi$ 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

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

is nonempty and compact.

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

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

is nonempty and compact.

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

$$\{x\in X \mid \phi(x,\overline{z})\le\overline{\gamma}\},\qquad \{z\in Z \mid \phi(\overline{x},z)\ge\overline{\gamma}\},$$

are nonempty and compact.

## Section 5.6. Theorems of the Alternative

Theorems of the alternative are important tools in optimization, which address the feasibility (possibly strict) of affine inequalities. These theorems can be viewed as special cases of MC/MC duality, as discussed in [Ber09].

**Proposition 5.6.1: (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 $\mu \in \Re^m$,

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

(iii) Any polyhedral set of the form

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

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

**Proposition 5.6.2: (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:

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

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

**Proposition 5.6.3: (Stiemke's Transposition Theorem)** Let $A$ be an $m \times n$ matrix, and let $c$ be a vector in $\Re^m$. The system

$$Ax = c, \qquad x > 0$$

has a solution if and only if

$$A'\mu \geq 0 \text{ and } c'\mu \leq 0 \qquad \Rightarrow \qquad A'\mu = 0 \text{ and } c'\mu = 0.$$

The theorems of Gordan and Stiemke can be used to provide necessary and sufficient conditions for the compactness of the primal and the dual optimal solution sets of linear programs. We say that the primal linear program

$$\begin{array}{ll} \text{minimize} & c'x \\ \text{subject to} & a_j'x \geq b_j, \quad j = 1, \ldots, r, \end{array} \tag{B.21}$$

*is strictly feasible* if there exists a primal-feasible vector $x \in \Re^n$ with $a_j'x > b_j$ for all $j = 1, \ldots, r$. Similarly, we say that the dual linear program

$$\begin{array}{ll} \text{maximize} & b'\mu \\ \text{subject to} & \displaystyle\sum_{j=1}^{r} a_j \mu_j = c, \quad \mu \geq 0, \end{array} \tag{B.22}$$

*is strictly feasible* if there exists a dual-feasible vector $\mu$ with $\mu > 0$. We have the following proposition.

**Proposition 5.6.4:** Consider the primal and dual linear programs (B.21) and (B.22), and assume that their common optimal value is finite. Then:

(a) The dual optimal solution set is compact if and only if the primal problem is strictly feasible.

(b) Assuming that the set $\{a_1, \ldots, a_r\}$ contains $n$ linearly independent vectors, the primal optimal solution set is compact if and only if the dual problem is strictly feasible.