IE 521: Convex Optimization **Spring 2017, UIUC**

Lecture 8: Convex Programming – February 15

Instructor: Niao He *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- Basics of Convex Programs
- Convex Theorem on Alternatives
- Lagrange Duality

References: Ben-Tal & Nemirovski, Chapter 3.1-3.3

8.1 Basics of Convex Programs

The standard form of an optimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, i = 1, ..., m \quad (P) \\ & h_j(x) = 0, j = 1, ..., k \end{aligned}$$

The optimal value of (P) is

$$p^* = \begin{cases} +\infty, & \text{if no feasible solution} \\ \inf_{x: g_i(x) \leq 0, \forall i, h_j(x) = 0\ \forall j} f(x), & \text{if exists feasible solution.} \end{cases}$$

- (P) is <u>infeasible</u>, if $p^* = +\infty$
- (P) is <u>unbounded below</u>, if $p^* = -\infty$
- (P) is <u>solvable</u>. if $\exists$ a feasible solution x^*, s.t. $p^* = f(x^*)$
- (P) is <u>unattainable</u>, if $|p^*| < \infty$ but $\nexists$ feasible x^*, s.t. $p^* = f(x^*)$. For example, $\min_{x \in (0,1)} e^x$

Given a solution x^*,

- x^* is a <u>global optimum</u> for (P) if x^* is feasible and $f(x^*) \leq f(x), \forall x$ feasible

- x^* is a <u>local optimum</u> for (P) if x^* is feasible and $\exists r > 0$, s.t. $f(x^*) \le f(x), \forall$ feasible $x \in B(x^*, r)$

Proposition 8.1 *For convex programs, a local optimum is a global optimum.*

Proof: Let C denote the feasible set, and x^* be local optimum. We want to show that $\forall x \in C,\ f(x^*) \le f(x)$. Let $z = \epsilon x^* + (1-\epsilon)x$. Then $z \in C \cap B(x^*, r)$ when ϵ is small enough. Hence,

$$f(x^*) \le f(z) \le \epsilon f(x^*) + (1-\epsilon) f(x)$$

Hence, $f(x^*) \le f(x), \forall x \in C$. ■

Question: How to verify whether a solution x^* is optimal?

In the linear program case, we have shown strong duality between (P) & (D)

$$(P) \quad \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \ge 0 \end{aligned} \qquad\qquad (D) \quad \begin{aligned} \max \quad & b^T y \\ \text{s.t.} \quad & A^T y \le c \end{aligned}$$

We know that

$$\begin{aligned} x^* \text{ is optimal} \Leftrightarrow & Ax^* = b, x \ge 0 \text{ (primal feasibility)} \\ & \exists y^*, A^T y^* \le c \text{ (dual feasibility)} \\ & c^T x^* = b^T y^* \text{ (zero-duality gap)} \end{aligned}$$

The strong duality is based on the Farkas' Lemma (or separation theorem). As an analogy, we can derive duality and optimality condition for general convex programs.

8.2 Convex Theorems on Alternatives

Consider the general form of convex program

$$\begin{aligned} \min_{x \in X} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \le 0, i = 1, ..., m \qquad (P) \end{aligned}$$

where X is a convex set, $f, g_1, \ldots, g_m$ are convex functions.

Theorem 8.2 *Assume $g_1, ..., g_m$ satisfy the <u>Slater condition:</u> $\exists \bar{x} \in X$, s.t. $g_i(\bar{x}) < 0$. Then exactly one of the following two systems must be empty.*

(I) $\{x \in X : f(x) < 0, g_i(x) \leq 0, i = 1, ..., m\}$

(II) $\{\lambda \in \mathbf{R}^m : \lambda \geq 0, \inf_{x \in X} \{f(x) + \sum_{i=1}^m \lambda_i g_i(x)\} \geq 0\}$

Proof:

1. We first show that system (II) feasible $\Rightarrow$ system (I) infeasible.

   Suppose (I) is also feasible, $\exists x_0$, s.t. $f(x_0) < 0, g_i(x_0) \leq 0$. Then

   $$\forall \lambda \geq 0, \inf_{x \in X} \left\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) \right\} < 0$$

   Contradiction!

2. We then show that system (I) infeasible $\Rightarrow$ system (II) feasible.

   Consider the following two sets:

   $$S = \left\{ u = (u_0, u_1, ..., u_m) \in \mathbf{R}^{m+1} : \exists x \in X, f(x) \leq u_0, g_1(x) \leq u_1, ..., g_m(x) \leq u_m \right\}$$
   $$T = \left\{ u = (u_0, u_1, ..., u_m) \in \mathbf{R}^{m+1} : u_0 < 0, u_1 \leq 0, ..., u_m \leq 0 \right\}$$

   Note that S is convex (why?) and nonempty, T is convex and nonempty and $S \cap T = \emptyset$. By separation theorem, $\exists a = (a_0, a_1, ..., a_m) \in \mathbf{R}^{m+1}$ and $a \neq 0$, s.t.

   $$\sup_{u \in T} a^T u \leq \inf_{u \in S} a^T u$$

   $$\text{i.e.} \quad \sup_{u_0 < 0, u_i \leq 0, \forall i = 1,..,m} \sum_{i=0}^m a_i u_i \leq \inf_{x, u_0, ..., u_m : u_0 \geq f(x), u_i \geq g_i(x), \forall i = 1,..,m} \sum_{i=0}^m a_i u_i$$

   Observe $a \geq 0$, hence:

   $$0 \leq \inf_x a_0 f(x) + a_1 g_1(x) + ... + a_m g_m(x)$$

   Note that $a_0 > 0$, otherwise: we have $(a_1, ..., a_m) \geq 0$ and $(a_1, \ldots, a_m) \neq 0$,

   $$\inf_x \{a_1 g_1(x) + ... + a_m g_m(x)\} \geq 0$$

   However, $\exists \bar{x}$, s.t. $g_i(x) < 0, \forall i = 1, ..., m$. This implies

   $$\inf_{x \in X} a_1 g_1(x) + ... + a_m g_m(x) < 0$$

   Contradiction!
   Hence, setting $\lambda_i = \frac{a_i}{a_0}, i = 1, ..., m$. we have

   $$\lambda \geq 0 \text{ and } \inf_x \left\{ f(x) + \sum \lambda_i g_i(x) \right\} \geq 0$$

   i.e. system (II) is feasible

■

Remark (relaxed Slater condition): The Slater condition can be relaxed to accommodate for linear equalities: $\exists x \in \text{rint}(X)$ s.t. $g_i(x) < 0$ for all $i = \{1, ..., m\}$ such that $g_i(x)$ is not affine.

Note that in the general case,

$$x^* \text{ is optimal to } (P) \Leftrightarrow \begin{cases} \{x \in X : f(x) \leq f(x^*), g_i(x) \leq 0, i = 1, ..., m\} \text{ is feasible} \\ \{x \in X : f(x) < f(x^*), g_i(x) \leq 0, i = 1, ..., m\} \textbf{ is infeasible} \end{cases}$$

$$\Leftrightarrow \begin{cases} \{x \in X : f(x) \leq f(x^*), g_i(x) \leq 0, i = 1, ..., m\} \text{ is feasible} \\ \{\lambda \in \mathbf{R}^m : \lambda \geq 0, \inf_{x \in X} \{f(x) + \sum \lambda_i g_i(x)\} \geq f(x^*)\} \textbf{is feasible} \end{cases}$$

Observe that the function $\inf_{x \in X} \{f(x) + \sum \lambda_i g_i(x)\} \leq f(x^*)$ for any $\lambda \geq 0$. Therefore, the optimality of x^* implies that there must exsit $\lambda^* \geq 0$ such that $\inf_{x \in X} \{f(x) + \sum \lambda_i^* g_i(x)\} = f(x^*)$.

8.3 Lagrange Duality

Definition 8.3 *The function $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ is called the Lagrange function. This induces two related functions:*

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

$$\bar{L}(x) = \sup_{\lambda \geq 0} L(x, \lambda)$$

The Lagrange dual *of the problem*

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ s.t. & g_i(x) \leq 0, i = 1, ..., m \\ & x \in X \end{array}$$

is defined as

$$(D) \qquad \begin{array}{ll} \max & \underline{L}(\lambda) \\ s.t. & \lambda \geq 0 \end{array}$$

Theorem 8.4 (Duality Theorem) *Denote Opt(P) and Opt(D) as the optimal values to (P) and (D), we have*

(a) **(Weak Duality)** $\forall \lambda \geq 0, \underline{L}(\lambda) \leq \text{Opt}(D)$. *Moreover, Opt$(D) \leq$ Opt(P).*

(b) **(Strong Duality)** *If (P) is convex and below bounded, and satisfies Slater condition, then (D) is solvable, and*

$$\text{Opt}(D) = \text{Opt}(P)$$

Proof:

(a) If (P) is infeasible, $\text{Opt}(P) = \infty$, $\text{Opt}(D) \leq \text{Opt}(P)$ always hold.
If (P) is feasible, let x_0 be feasible solution, i.e. $g_i(x_0) \leq 0, x_0 \in X$

$$\forall \lambda \geq 0, \underline{L}(\lambda) = \inf_{x \in X} \left\{ f(x) + \sum \lambda_i g_i(x) \right\} \leq f(x_0) + \lambda_i g_i(x_0) \leq f(x_0)$$

Hence,

$$\forall \lambda \geq 0, \underline{L}(\lambda) \leq \inf \{ f(x_0) : x_0 \in X, g_i(x_0) \leq 0, i = 1, ..., m \} = \text{Opt}(P)$$

Furthermore, $\text{Opt}(D) = \sup_{\lambda \geq 0} \underline{L}(\lambda) \leq \text{Opt}(P)$

(b) By optimality, we know that $\{x \in X : f(x) < \text{Opt}(P), g_i(x) \leq 0, i = 1, ..., m\}$ has no solution. By convex theorem on alternative and Slater condition, the system $\{\lambda \geq 0 : \ \underline{L}(\lambda) \geq \text{Opt}(P)\}$ has a solution. This implies that $\text{Opt}(D) = \sup_{\lambda \geq 0} \underline{L}(\lambda) \geq \text{Opt}(P)$. Combining with part (a), we have $\text{Opt}(P) = \text{Opt}(D)$.

■

Example: The Slater condition does not hold in the following example.

$$\begin{aligned} \min_{x \in X} \quad & e^{-x} \\ \text{s.t.} \quad & \frac{x^2}{y} \leq 0 \end{aligned}$$

where $X = \{(x, y) : y > 0\}$. Note that there exists no solution in $x \in X$ such that $x^2/y < 0$.