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

Lecture 10: Minimax Problems – February 22

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

- Saddle Point and Minimax Problems
- Minimax Theorem (Sion-Kakutani Theorem)

References: Bental & Nemirovski Chapter 3.4

10.1 Saddle Ppint and Minimax Problem

Let $X \subseteq \mathbf{R}^n$ and $Y \subseteq \mathbf{R}^m$ be two non-empty sets. Let $L(x, y) : X \times Y \to \mathbf{R}$ be a real-valued function. We say (x^*, y^*) is a saddle point of $L(x, y)$ if

$$L(x, y^*) \geq L(x^*, y^*) \geq L(x^*, y), \forall x \in X, y \in Y$$

Game Theory Interpretation: This can be viewed as a two player zero-sum game.

- Player 1 chooses $x \in X$
- Player 2 chooses $y \in Y$
- After both players reveal their decisions, player 1 pays to player 2 the amount $L(x, y)$.
- Player 1 is interested in minimizing his loss, while player 2 is interested in maximizing his gain
- The strategy (x^*, y^*) is called a (Nash) equilibrium if no player has an interactive to change his chosen strategy after considering the opponent's action.

Question: What should players do to optimize their profit?

Considering the following two situations

- *Situation I:* Player 1 choose x first and player 2 choose y knowing the choice of player 2

$$\min_{x \in X} \bar{L}(x) := \max_{y \in Y} L(x, y)$$

- *Situation II:* Player 2 choose y first and player 1 choose x knowing the choice of player 1

$$\max_{y\in Y} \underline{L}(y) := \min_{x\in X} L(x,y)$$

Two Induced Problems:

$$\begin{aligned} (P): &\quad \min_{x\in X}\max_{y\in Y} L(x,y) = \min_{x\in X} \bar{L}(y) \\ (D): &\quad \max_{y\in Y}\min_{x\in X} L(x,y) = \max_{y\in Y} \underline{L}(x) \end{aligned}$$

Intuitively, it is better for players to play second. Indeed, there is an inequality:

Proposition 10.1

$$\sup_{y\in Y}\inf_{x\in X} L(x,y) \le \inf_{x\in X}\sup_{y\in Y} L(x,y).$$

Proof: This is because

$$\begin{aligned} &\forall x\in X: \inf_{x\in X} L(x,y) \le L(x,y), \forall y \\ \Rightarrow &\forall x \in X: \sup_{y\in Y}\inf_{x\in X} L(x,y) \le \sup_{y\in Y} L(x,y) \\ \Rightarrow &\sup_{y\in Y}\inf_{x\in X} L(x,y) \le \inf_{x\in X}\sup_{y\in Y} L(x,y) \end{aligned}$$

■

On the other hand, if (x^*,y^*) is a saddle point.

$$\begin{aligned} &L(x,y^*) \ge L(x^*,y^*) \ge L(x^*,y), \quad \forall x\in X, y\in Y \\ \Rightarrow &\inf_{x\in X} L(x,y^*) \ge L(x^*,y^*) \ge \sup_{y\in Y} L(x^*,y) \\ \Rightarrow &\sup_{y\in Y}\inf_{x\in X} L(x,y) \ge \inf_{x\in X}\sup_{y\in Y} L(x,y) \end{aligned}$$

Hence, if (x^*,λ^*) is a saddle point, then

$$\sup_{y\in Y}\inf_{x\in X} L(x,y) = \inf_{x\in X}\sup_{y\in Y} L(x,y)$$

Remark:

- If there exists a saddle point, there is no advantage to the players of knowing the opponent's choice. Moreover, the equilibrium, minimax and maximin all give the same solution.
- Saddle points do not always exist. For example. $L(x,y)=(x-y)^2, X=[0,1], Y=[0,1]$

$$\begin{aligned} &\bar{L}(x) = \sup_{y\in[0,1]} L(x,y) = \max\left\{x^2,(x-1)^2\right\}, &&\inf_{x\in X}\sup_{y\in Y} L(x,y) = \frac{1}{4} \\ &\underline{L}(y) = \inf_{x\in[0,1]} L(x,y) = 0, &&\sup_{y\in Y}\inf_{x\in X} L(x,y) = 0 \end{aligned}$$

- As already shown in last lecture: Saddle point exists if and only the induced problems (P) and (D) are both solvable and the optimal values equal to each other. Moreover, the saddle points are pairs (x^*,λ^*), where x^* is optimal to (P) and λ^* is optimal to (D).

10.2 Existence of Saddle Points

Lemma 10.2 *(Minimax Lemma) Let $f_i(x), i = 1, ..., m$ be convex and continuous on a convex compact set X. Then*

$$\min_{x\in X}\max_{1\le i\le m} f_i(x) = \min_{x\in X}\sum_{i=1}^{m}\lambda_i^* f_i(x)$$

for some $\lambda^* \in \mathbf{R}^m$ such that, $\lambda_i^* \ge 0, i = 1, ..., m$ and $\sum_{i=1}^m \lambda_i^* = 1$

Remark: Let $L(x,\lambda) = \sum_{i=1}^m \lambda_i f_i(x), \triangle = \{\lambda : \lambda \ge 0, \sum_{i=1}^m \lambda_i = 1\}$ be a standard simplex. The above lemma implies: $L(x,\lambda)$ has a saddle point on $X \times \triangle$. Since

$$\max_{\lambda\in\triangle}\min_{x\in X}\sum\lambda_i f_i(x) \ge \min_{x\in X}\sum\lambda_i^* f_i(x) = \min_{x\in X}\max_{\lambda\in\triangle}\sum\lambda_i f_i(x)$$

Proof: Consider the epigraph form of the problem $\min_{x\in X}\max_{1\le i\le m} f_i(x)$:

$$\begin{aligned}\min_{x,t}\quad & t\\ \text{s.t.}\quad & f_i(x) - t \le 0, i = 1, ..., m\\ & (x,t) \in X_t\end{aligned}$$

where $X_t = \{(x,t) : x \in X, t \in \mathbf{R}\}$. The optimal value $t^* = \min_{x\in X}\max_{1\le i\le m} f_i(x)$.
Note that the above problem satisfies the slater condition and is solvable (why ?)
Hence, there exists $(x^*, t^*) \in X_t$ and $\lambda^* \ge 0$, such that $(x^*, t^*; \lambda^*)$ is a saddle point of the Lagrange function:

$$L(x,t;\lambda) = t + \sum_{i=1}^m \lambda_i(f_i(x) - t) = (1 - \sum_{i=1}^m \lambda_i)t + \sum_{i=1}^m \lambda_i f_i(x)$$

Therefore:

$$\begin{cases}\frac{\partial L}{\partial t}L(x^*,t^*;\lambda^*) = 1 - \sum_{i=1}^m \lambda_i^* = 0\\ \sum_{i=1}^m \lambda_i^*(f_i(x^*) - t^*) = 0\end{cases} \Rightarrow \begin{cases}\sum_{i=1}^m \lambda_i^* = 1\\ \sum_{i=1}^m \lambda_i^* f_i(x^*) = t^*\end{cases}$$

This implies that $\exists\lambda_i^* \ge 0, \sum_{i=1}^m \lambda_i^* = 1$, such that

$$\min_{x\in X}\max_{1\le i\le m} f_i(x) = t^* = \min_{(x,t)\in X_t} L(x,t;,\lambda^*) = \min_{x\in X}\sum_{i=1}^m \lambda_i^* f_i(x)$$

■

Theorem 10.3 *(Minima Theorem, Sion-Kakutani)*
Let X and Y be two convex compact sets. Let $L(x,y) : X \times Y \to \mathbf{R}$ be a continuous function that is convex in $x \in X$ for every fixed $y \in Y$ and concave in $y \in Y$ for every fixed $x \in X$. Then $L(x,y)$ has a saddle point on $X \times Y$, and

$$\min_{x\in X}\max_{y\in Y} L(x,y) = \max_{y\in Y}\min_{x\in X} L(x,y)$$

Proof: We should prove that

$$\begin{aligned}(P): &\quad \min_{x\in X} \bar{L}(x) := \max_{y\in Y} L(x,y)\\ (D): &\quad \max_{y\in Y} \underline{L}(y) := \min_{x\in X} L(x,y)\end{aligned}$$

are both solvable with equal optimal values.

1. X and Y are compact, $L(x,y)$ is continuous, both $\bar{L}(x)$ and $\underline{L}(y)$ are continuous and attain their optimum on compact set. It is sufficient to show $\text{Opt}(D) \geq \text{Opt}(P)$,i.e.

$$\max_{y\in Y}\min_{x\in X} L(x,y) \geq \min_{x\in X}\max_{y\in Y} L(x,y)$$

   Consider the sets $X(y) = \{x \in X : L(x,y) \leq \text{Opt}(D)\}$

2. Note that $X(y)$ is <u>nonempty</u>, <u>compact</u> and <u>convex</u> for any $y \in Y$. We show that every collection of these sets has a point in common.
   Suppose $\exists y_1, ..., y_m$. s.t. $X(y_1) \cap ... \cap X(y_m) = \emptyset$. This implies:

$$\min_{x\in X}\max_{i=1,\ldots,m} L(x,y_i) > Opt(D)$$

   By Minimax Lemma, $\exists \lambda_i^* \geq 0$ and $\sum_{i=1}^m \lambda_i^* = 1$, such that

$$\begin{aligned}\min_{x\in X}\max_{i=1,\ldots,m} L(x,y_i) &= \min_{x\in X}\sum_{i=1}^{m}\lambda_i^* L(x,y_i)\\ &\leq \min_{x\in X} L(x, \sum_{i=1}^{m}\lambda_i^* y_i)\\ &= \underline{L}(\bar{y}) \leq \text{Opt}(D)\end{aligned}$$

   where $\bar{y} = \sum_{i=1}^m \lambda_i^* y_i$ and the first inequality is by concavity of $L(x,y)$. The result here leads to a contradiction! Hence, every finite collection of $X(y)$ has a common point.

3. By Helley's theorem, all of these sets $\{X(y) : y \in Y\}$ has a common point. Therefore, $\exists x^* \in X : x^* \in X(y), \forall y \in Y$, which means $\exists x^* \in X, L(x^*,y) \leq \text{Opt}(D), \forall y \in Y$. Hence $\text{Opt}(P) \leq \text{Opt}(D)$

∎