On Minimax Theorems

Jyunyi Liao

In this article, we study the minimax theorems, which provide conditions ensuring that the max-min inequality is also an equality. For any function $f : X \times Y \to \mathbb{R}$, the *max-min inequality* asserts

$$\sup_{x\in X} \inf_{y\in Y} f(x,y) \leq \inf_{y\in Y} \sup_{x\in X} f(x,y)$$

This is also called *weak duality* in optimization. We wonder if the equality holds under certain conditions.

1 Quasi-convex Functions

Definition 1 (Quasi-convexity)**.** *Let V be a vector space. A function $\varphi : V \to \mathbb{R}$ is said to be **quasi-convex**, if for each $x, y \in V$ and each $0 \leq \lambda \leq 1$,*

$$\varphi(\lambda x + (1-\lambda)y) \leq \max\{\varphi(x), \varphi(y)\}.$$

*A function $\varphi : V \to \mathbb{R}$ is said to be **quasi-concave** if its negative is quasi-convex. In other words, φ is quasi-concave if for each $x, y \in V$ and each $0 \leq \lambda \leq 1$,*

$$\varphi(\lambda x + (1-\lambda)y) \geq \min\{\varphi(x), \varphi(y)\}.$$

We have the following level-set characterization of quasi-convex functions.

Theorem 2. *Let V be a vector space. A function $\varphi : V \to \mathbb{R}$ is quasi-convex if and only if each sublevel set*

$$\varphi^{-1}((-\infty, \alpha]) = \{x \in V : \varphi(x) \leq \alpha\}, \quad \alpha \in \mathbb{R}$$

is a convex set in V.

Proof. Assume $\varphi : V \to \mathbb{R}$ is quasi-convex, and fix $\alpha \in \mathbb{R}$. For any pair $x, y \in \varphi^{-1}((-\infty, \alpha])$,

$$\varphi(\lambda x + (1-\lambda)y) \leq \max\{\varphi(x), \varphi(y)\} \leq \alpha, \quad \forall 0 \leq \lambda \leq 1.$$

Hence $\varphi^{-1}((-\infty, \alpha])$ is convex. Conversely, assume each sublevel sets of φ is convex. Fix $x, y \in V$, and take $\alpha = \max\{\varphi(x), \varphi(y)\}$. Then $x, y \in \varphi^{-1}((-\infty, \alpha])$, and by convexity

$$\lambda x + (1-\lambda)y \in \varphi^{-1}((-\infty, \alpha]) \quad \Rightarrow \quad \varphi(\lambda x + (1-\lambda)y) \leq \alpha = \max\{\varphi(x), \varphi(y)\}, \quad \forall 0 \leq \lambda \leq 1.$$

Hence φ is quasi-convex. □

Remark. Similarly, a function $\varphi : V \to \mathbb{R}$ is quasi-concave if and only if the each superlevel set

$$\varphi^{-1}([\alpha, \infty)) = \{x \in V : \varphi(x) \geq \alpha\}, \quad \alpha \in \mathbb{R}$$

is a convex set in V.

2 Von-Neumann Minimax Theorem

We first introduce an intersection property of convex sets.

Lemma 3. *Let $C_1, \cdots, C_n$ and C be compact convex sets in a Euclidean space such that*

(i) $C \cap \bigcap_{i=1, i\neq j}^{n} C_i \neq \emptyset$ *for each* $j = 1, \cdots, n$*; and*

(ii) $C \cap \bigcap_{i=1}^{n} C_i = \emptyset$.

Then C is not contained in $\bigcup_{i=1}^{n} C_i$.

Proof. We proceed by induction on n. If $n = 1$, then (i) implies that C is nonempty, and (ii) implies that C and C_1 are disjoint. The result $C \not\subset C_1$ is then clear.

Suppose our assertion holds for $n-1$, and consider the case n. According to (ii), $C \cap C_n$ and $\bigcap_{i=1}^{n-1} C_i$ are disjoint compact convex sets, which can be strictly separated by a hyperplane H. Let $D = C \cap H$, and $D_i = C_i \cap H$ for $i = 1, \cdots, n$. We verify that the sets $D_1, \cdots, D_{n-1}$ and D satisfies conditions (i) and (ii).

(i) Given any $j \in \{1, \cdots, n-1\}$, we can take $x_j \in C \cap \bigcap_{i=1, i\neq j}^{n} C_i$, and take $x_k \in C \cap \bigcap_{i=1}^{n-1} C_i$ by (i). Since $x_j \in C \cap C_n$ and $x_k \in \bigcap_{i=1}^{n-1} C_i$, they are separated by hyperplane H. We take y_j to be the intersection of the line segment $[x_j, x_k]$ and the hyperplane H. Clearly, $[x_j, x_k]$ lies in the convex set $C \cap \bigcap_{i=1, i\neq j}^{n-1} C_i$. Hence $y_j \in D \cap \bigcap_{i=1, i\neq j}^{n-1} D_i \neq \emptyset$.

(ii) Since $\bigcap_{i=1}^{n-1} C_i$ and $C \cap C_n$ are strictly separated by H, we have $D \cap \bigcap_{i=1}^{n-1} D_i = \emptyset$.

By the induction hypothesis, there exists $x_0 \in D$ such that $x_0 \notin \bigcup_{i=1}^{n-1} D_i$. Since $D \cap C_n = (C \cap H) \cap C_n = \emptyset$, it follows $x_0 \notin C_n$. Hence $C \ni x_0 \notin \bigcup_{i=1}^{n} C_k$, which completes the proof. □

Remark. This proof is due to [1]. In fact, one can assume that $C_1, \cdots, C_n$ and C to be closed convex sets.

Theorem 4 (Von Neumann Minimax theorem)**.** *Let U and V be topological vector spaces, and let X and Y be compact convex subsets of U and V, respectively. Let $f : X \times Y \to \mathbb{R}$ be a function such that*

(i) f is continuous on $X \times Y$;

(ii) $f(x, y)$ is quasi-concave in variable x; and

(iii) $f(x, y)$ is quasi-convex in variable y.

Then there exists a saddle point $(x_0, y_0) \in X \times Y$ such that

$$f(x, y_0) \leq f(x_0, y_0) \leq f(x_0, y), \quad \forall x \in X, y \in Y.$$

Moreover, the max-min inequality for f is also an equality:

$$f(x_0, y_0) = \max_{x\in X} \min_{y\in Y} f(x, y) = \min_{y\in Y} \max_{x\in X} f(x, y).$$

The proof of this result needs some technical lemmata.

Lemma 5. *Define the sets*

$$E_\lambda = \{x \in X : f(x, y) \geq \lambda \textit{ for all } y \in Y\},$$
$$F_\mu = \{x \in X : f(x, y) \leq \mu \textit{ for all } x \in X\},$$

where λ and μ are arbitrary real numbers. Define

$$\lambda_0 = \sup\{\lambda : E_\lambda \neq \emptyset\}, \quad \textit{and} \quad \mu_0 = \inf\{\mu : F_\mu \neq \emptyset\}.$$

Then

$$\lambda < \infty, \quad \mu > -\infty, \quad \textit{and} \quad E_{\lambda_0} \neq \emptyset, \quad F_{\mu_0} \neq \emptyset.$$

Proof. We define $g(x) = \inf_{y \in Y} f(x,y)$. Then $g(x) \geq \lambda$ if and only if $f(x,y) \geq \lambda$ for each $y \in Y$. Hence

$$E_\lambda = \{x \in X : g(x) \geq \lambda\} = \bigcap_{y \in Y} \{x \in X : f(x,y) \geq \lambda\}$$

is a superlevel set of function g, and it is convex. Furthermore,

$$\inf_{y \in Y} [f(x,y) - f(x',y)] \leq g(x) - g(x') \leq \sup_{y \in Y} [f(x,y) - f(x',y)] .$$

By uniform continuity of f on the compact set $X \times Y$, the function g is also (uniformly) continuous on X. Then g is bounded, and $\lambda < \infty$. Also, the level set $E_\lambda = g^{-1}([\lambda, \infty))$ is closed in X. Furthermore,

$$E_{\lambda_0} = g^{-1}([\lambda_0, \infty)) = g^{-1}\left(\bigcap_{\lambda < \lambda_0} [\lambda, \infty)\right) = \bigcap_{\lambda < \lambda_0} g^{-1}([\lambda, \infty)) = \bigcap_{\lambda < \lambda_0} E_\lambda.$$

Since the closed sets E_λ are nonempty for all $\lambda < \lambda_0$, by compactness of X, their intersection E_{λ_0} is also nonempty. A similar assertion also holds for F_{μ_0}. $\square$

Lemma 6. $\lambda_0 \geq \mu_0$.

Proof. Fix arbitrarily $\epsilon > 0$. By definition of λ_0, we have $E_{\lambda_0+\epsilon} = \emptyset$. Hence for every $x \in X$, there exists $y \in Y$ such that $f(x,y) < \lambda_0 + \epsilon$. We define

$$U_y^\epsilon = \{x \in X : f(x,y) < \lambda_0 + \epsilon\} .$$

Since f is continuous, U_y^ϵ is an open set. Furthermore, $\bigcup_{y \in Y} U_y^\epsilon$ is an open cover of X. By compactness of X, there exists finitely many $y_1, \cdots, y_n \in Y$ such that

$$X \subset \bigcup_{i=1}^n U_{y_i}^\epsilon .$$

Similarly, we can find finitely many $x_1, \cdots, x_m \in X$ that form an open cover $\bigcup_{i=1}^n V_{x_i}^\epsilon$ of Y, where V_x^ϵ is

$$V_x^\epsilon = \{x \in X : f(x,y) > \mu_0 - \epsilon\} .$$

Let $C = \mathrm{conv}(x_1, \cdots, x_m)$, and let $L = x_1 + \mathrm{span}\{x_2 - x_1, \cdots, x_m - x_1\}$. Clearly, C is compact, and L is the minimal affine subspace containing C, which is homeomorphic to an Euclidean space. Then, $X \cap L$ is covered by $\{U_i = U_{y_i}^\epsilon \cap L, i = 1, \cdots, n\}$. By dropping possibly redundant elements, one may assume the cover $\{U_i, i = 1, \cdots, n\}$ is minimal, in the sense that $C \subset \cap_{i=1}^n U_i$ and $C \not\subset \cap_{i=1, i \neq j}^n U_i$ for each $j = 1, \cdots, n$. Define

$$C_i = L \backslash U_i = \{x \in X \cap L : f(x, y_i) \geq \lambda_0 + \epsilon\} , \quad i = 1, \cdots, n.$$

By continuity and quasi-concavity of $f(\cdot, y_i)$, this is a compact convex set. Since the cover $\{U_i, i = 1, \cdots, n\}$ is minimal, the two conditions in Lemma 3 are satisfied. Hence there exists $x_0 \in C$ such that

$$f(x_0, y_i) < \lambda_0 + \epsilon, \quad \forall i = 1, \cdots, n.$$

By quasi-convexity of $f(x_0, \cdot)$, we have

$$f(x_0, y) < \lambda_0 + \epsilon, \quad y \in D := \mathrm{conv}(y_1, \cdots, y_n).$$

Similarly, there exists $y_0 \in D$ such that

$$f(x, y_0) > \mu_0 - \epsilon, \quad x \in C := \mathrm{conv}(x_1, \cdots, x_m).$$

Therefore,

$$\mu_0 - \epsilon < f(x_0, y_0) < \lambda_0 + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have $\mu_0 \leq \lambda_0$. □

Now we are prepared to prove the von Neumann Minimax theorem.

Proof of Theorem 4. By Lemma 5, we take $x_0 \in E_{\lambda_0}$ and $y_0 \in F_{\mu_0}$. Then

$$\lambda_0 \leq f(x_0, y_0) \leq \mu_0.$$

By Lemma 6, $\lambda_0 = \mu_0$. Then for all $x \in X$ and $y \in Y$,

$$f(x, y_0) \leq \mu_0 = f(x_0, y_0) = \lambda_0 \leq f(x_0, y)$$

Hence (x_0, y_0) is a saddle point. Furthermore,

$$\lambda_0 = \sup\left\{\lambda : \exists x \in X \textit{ such that } \inf_{y\in Y} f(x, y) \geq \lambda\right\} = \sup_{x\in X} \inf_{y\in Y} f(x, y),$$
$$\mu_0 = \inf\left\{\lambda : \exists y \in Y \textit{ such that } \sup_{x\in X} f(x, y) \geq \mu\right\} = \inf_{y\in Y} \sup_{x\in X} f(x, y).$$

Since both X and Y are compact, and both the mappings $x \mapsto \inf_{y\in Y} f(x, y)$ and $y \mapsto \sup_{x\in X} f(x, y)$ are continuous, we have

$$\max_{x\in X} \min_{y\in Y} f(x, y) = \lambda_0 = f(x_0, y_0) = \mu_0 = \min_{y\in Y} \max_{x\in X} f(x, y).$$

Thus we complete the proof. □

Application: Matrix Game. Consider a two-player zero-sum matrix game, which is defined by a triplet $(\mathcal{A}, \mathcal{B}, F)$. where $\mathcal{A} = \{1, 2, \cdots, m\}$ is a finite set of actions that the max player can take, $\mathcal{B} = \{1, 2, \cdots, n\}$ is the set of actions that the max player can take, and $F : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ is utility function. The zero-sum game can be formulated as the following max-min problem

$$\max_{\xi\in\Delta(\mathcal{A})} \min_{\eta\in\Delta(\mathcal{B})} \xi^\top F\eta,$$

where $\xi \in \Delta(\mathcal{A})$ and $\eta \in \Delta(\mathcal{B})$ are strategies for each player:

$$\Delta(\mathcal{A}) = \{\xi = (\xi_1, \cdots, \xi_m) : \xi_1, \cdots, \xi_m \geq 0,\ \xi_1 + \cdots + \xi_m = 1\},$$
$$\Delta(\mathcal{B}) = \{\eta = (\eta_1, \cdots, \eta_n) : \eta_1, \cdots, \eta_n \geq 0,\ \eta_1 + \cdots + \eta_n = 1\}$$

and $F = (F(a, b))_{a\in\mathcal{A}, b\in\mathcal{B}} \in \mathcal{R}^{m\times n}$ is the utility matrix. Clearly, the simplexes $\Delta(\mathcal{A})$ and $\Delta(\mathcal{B})$ are compact convex subsets of Euclidean spaces. By von Neumann minimax theorem, there exists strategies $\xi^0 \in \Delta(\mathcal{A})$ and $\eta^0 \in \Delta(\mathcal{B})$ such that

$$\xi^{0\top} F\eta^0 = \max_{\xi\in\Delta(\mathcal{A})} \xi^\top F\eta^0 = \min_{\eta\in\Delta(\mathcal{B})} \xi^{0\top} F\eta.$$

In fact, the last display implies

$$\xi^{0\top} F \eta^0 = \max_{i \in \{1, \cdots, m\}} \sum_{j=1}^{n} \eta_j^0 F(i,j) = \min_{j \in \{1, \cdots, n\}} \sum_{i=1}^{m} \xi_i^0 F(i,j).$$

References

[1] Klee, V.L. (1951) ‘On Certain Intersection Properties of Convex Sets’, *Canadian Journal of Mathematics*, 3, pp. 272–275.

[2] Ben-El-Mechaiekh, H. and Dimand, R. W. (2011) ‘A Simpler Proof of the Von Neumann Minimax Theorem’, *The American Mathematical Monthly*, 118(7), pp. 636–641.

[3] Fan, Ky. (1953) ‘Minimax Theorems’, *Proceedings of the National Academy of Sciences of the United States of America*, 39(1), pp. 42–47.