THE MINIMAX THEOREM

FEDERICO ECHENIQUE
ECON 201B – UC BERKELEY

Let $G = (\{1, 2\}, (S_1, S_2), (u_1, u_2))$ be a finite two-player normal-form game. We say that G is a ***zero-sum game*** if $u_1 + u_2 = 0$.

We may describe the game using the payoff function $u = u_1$, which we interpret as a payment from player 2 to player 1.

Minimax Theorem. *There is a strategy profile* (σ_1^*, σ_2^*) *such that*

$$\max_{\sigma_1 \in \Delta(S_1)} \min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1, \sigma_2) = u(\sigma_1^*, \sigma_2^*) = \min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1 \in \Delta(S_1)} u(\sigma_1, \sigma_2).$$

The number $v = u(\sigma_1^*, \sigma_2^*)$ is the ***value*** of the game.

Observe that, if the profile (σ_1^*, σ_2^*) is as in the theorem, then

$$u(\sigma_1, \sigma_2^*) \leq u(\sigma_1^*, \sigma_2^*) \leq u(\sigma_1^*, \sigma_2)$$

for all $\sigma_1 \in \Delta(S_1)$ and $\sigma_2 \in \Delta(S_2)$. This means that (σ_1^*, σ_2^*) is a ***saddle point*** of u. In particular, (σ_1^*, σ_2^*) is a Nash equilibrium of G.

Proof of the theorem.

Observe first that

$$(1) \qquad \max_{\sigma_1 \in \Delta(S_1)} \min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1, \sigma_2) \leq \min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1 \in \Delta(S_1)} u(\sigma_1, \sigma_2)$$

which we may describe as "moving second confers an advantage." To show inequality (1), note that for any $\sigma_1 \in \Delta(S_1)$ and $\sigma_2 \in \Delta(S_2)$ we have

$$u(\sigma_1, \sigma_2) \leq \max_{\sigma_1' \in \Delta(S_1)} u(\sigma_1', \sigma_2).$$

Since this inequality holds for all σ_2, we obtain that

$$\min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1, \sigma_2) \leq \min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1' \in \Delta(S_1)} u(\sigma_1', \sigma_2).$$

And since this inequality holds for all σ_1, we obtain inequality (1).

I'm grateful to SangMok Lee for comments on this note.

The main point of the Minimax Theorem is that inequality (1) is actually an equality — which we now show by establishing the reverse inequality.

Let $n_i = |S_i|$ and write, for each fixed $\sigma_2 \in \Delta(S_2)$, the function $s_1 \mapsto u(s_1, \sigma_2)$ as a vector $\vec{u}(\sigma_2) \in \mathbf{R}^{n_1}$.[1]

Let

$$\mathcal{C} = \{\vec{u}(\sigma_2) : \sigma_2 \in \Delta(S_2)\} \subseteq \mathbf{R}^{n_1},$$

and observe that $\mathcal{C}$ is a compact and convex set.

Define the function $m : \mathbf{R}^{n_1} \to \mathbf{R}$ by $m(x) = \max\{x_i : 1 \leq i \leq n_1\}$ and define

$$v = \inf\{m(x) : x \in \mathcal{C}\} = \inf\{\max\{u(s_1, \sigma_2) : s_1 \in S_1\} : \sigma_2 \in \Delta(S_2)\}.$$

Because $\mathcal{C}$ is compact, there exists $\sigma_2^* \in \Delta(S_2)$ for which

$$v = m(\vec{u}(\sigma_2^*)) = \max\{u(s_1, \sigma_2^*) : s_1 \in S_1\}$$

Note that $u(s_1, \sigma_2^*) \leq v$ for all $s_1 \in S_1$. Hence, for any $\sigma_1 \in \Delta(S_1)$,

$$u(\sigma_1, \sigma_2^*) = \sigma_1 \cdot \vec{u}(\sigma_2^*) \leq v. \tag{2}$$

Consider the set $\mathcal{A} = \{x \in \mathbf{R}^{n_1} : x \ll (v, \ldots, v)\}$. Then by definition of v, $\mathcal{A} \cap \mathcal{C} = \emptyset$. The set A is convex so there exists, by the separating hyperplane theorem, a vector $p \in \mathbf{R}^{n_1}$, $p \neq 0$, so that

$$p \cdot x \leq p \cdot \vec{u}(\sigma_2)$$

for all $x \in \mathcal{A}$ and $\sigma_2 \in \Delta(S_2)$.

Now, the set $\mathcal{A}$ contains vectors with arbitrarily small entries in any dimension. So we must in fact have $p \geq 0$. Since $p \neq 0$ we conclude that $p > 0$. Thus

$$\sigma_1^* = \frac{1}{\sum_{i=1}^{n_1} p_i} p \in \Delta(S_1)$$

is well-defined because $\sum_{i=1}^{n_1} p_i > 0$.

Given that σ_1^* is a positive scalar multiple of p, we have that

$$\sigma_1^* \cdot x \leq \sigma_1^* \cdot \vec{u}(\sigma_2) = u(\sigma_1^*, \sigma_2)$$

for all $x \in \mathcal{A}$ and $\sigma_2 \in \Delta(S_2)$.

[1] The notation here should remind you of our proof of the domination theorem. Other aspects of the proof will also remind you of the proof of the domination theorem.

Note that $(v, \ldots, v) - \varepsilon(1, \ldots, 1) \in \mathcal{A}$ for any $\varepsilon > 0$. Hence, for any $\sigma_2 \in \Delta(S_2)$,

$$\sigma_1^* \cdot [(v, \ldots, v) - \varepsilon(1, \ldots, 1)] \leq \sigma_1^* \cdot \vec{u}(\sigma_2).$$

Since this inequality holds for arbitrarily small ε, we conclude that

$$v = \sigma_1^* \cdot (v, \ldots, v) \leq \sigma_1^* \cdot \vec{u}(\sigma_2). \tag{3}$$

If we use $\sigma_1 = \sigma_1^*$ in Equation (2) we see that $u(\sigma_1^*, \sigma_2^*) \leq v$, while $\sigma_2 = \sigma_2^*$ in Equation (3) yields $v \leq u(\sigma_1^*, \sigma_2^*)$.

Hence, $v = u(\sigma_1^*, \sigma_2^*)$, and

$$u(\sigma_1, \sigma_2^*) \leq v = u(\sigma_1^*, \sigma_2^*) \leq u(\sigma_1^*, \sigma_2)$$

for any $\sigma_1 \in \Delta(S_1)$ and $\sigma_2 \in \Delta(S_2)$. The first inequality is due to Equation (2) and the second to Equation (3).

Finally, we claim that

$$\max_{\sigma_1 \in \Delta(S_1)} \min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1, \sigma_2) \geq \min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1 \in \Delta(S_1)} u(\sigma_1, \sigma_2).$$

Indeed,

$$\max_{\sigma_1 \in \Delta(S_1)} \min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1, \sigma_2) \geq \min_{\sigma_2 \in \Delta(S_2)} u(\sigma_1^*, \sigma_2) = u(\sigma_1^*, \sigma_2^*).$$

And

$$\min_{\sigma_2 \in \Delta(S_2)} \max_{\sigma_1 \in \Delta(S_1)} u(\sigma_1, \sigma_2) \leq \max_{\sigma_1 \in \Delta(S_1)} u(\sigma_1, \sigma_2^*) = u(\sigma_1^*, \sigma_2^*).$$

□

Wait, what happened?

The key idea in this proof is the work done by σ_1^*. We started from looking at the function m, which provides a worst-case scenario for player 2 for each σ_2. You could think of m by imagining that player 1 moves second, after a choice of σ_2 by player 2 that fixes a vector $\vec{u}(\sigma_2)$. Then σ_2^* is optimal for 2 when they imagine that 1 moves after them.

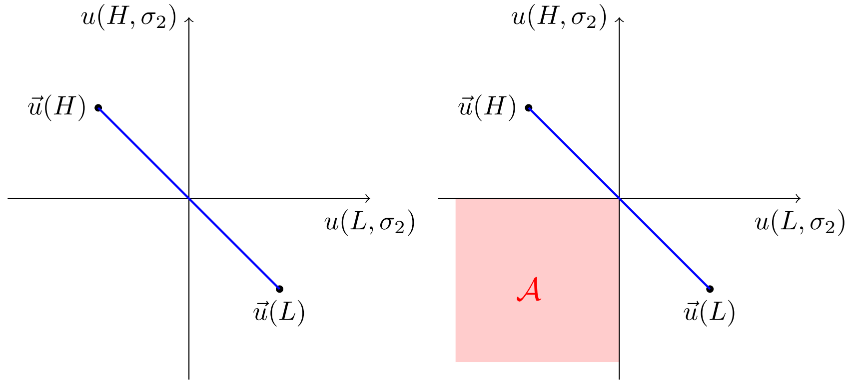
When we get the σ_1^* vector from the SHT, we can substitute the "player 1 moves second" idea implicit in function m with the expected payoff $\sigma_2 \mapsto \sigma_1^* \cdot \vec{u}(\sigma_2)$. The level curves of this expected payoff function are the parallel lines to the hyperplane that separate $\mathcal{A}$ and $\mathcal{C}$. And the function is minimized by choosing $\sigma_2 = \sigma_2^*$, even when 1's strategy is fixed at σ_1^* **before** 2's choice of strategy. In other words, if "player 1 moves first" by setting $\sigma_1 = \sigma_1^*$ then it is **still optimal** for player 2 to choose $\sigma_2 = \sigma_2^*$.

Examples

Example: Matching Pennies

	H	T
H	1	-1
T	-1	1

The next figure represents $\mathcal{C}$ as the blue set: a blue line segment in this case, obtained as the convex combinations of $(1,-1)$ and $(-1,1)$. We see that $v = 0$ and the separating hyperplane will be parallel to $\mathcal{C}$ (actually it will contain $\mathcal{C}$).



Example: Rock-Paper-Scissors

	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

I'm not going to try to draw this in $\mathbf{R}^3$! So I'll instead use the following game in which P1 only has two strategies. Note that P1 is at a disadvantage here because she cannot play Scissors.

	R	P	S
R	0	-1	1
P	1	0	-1

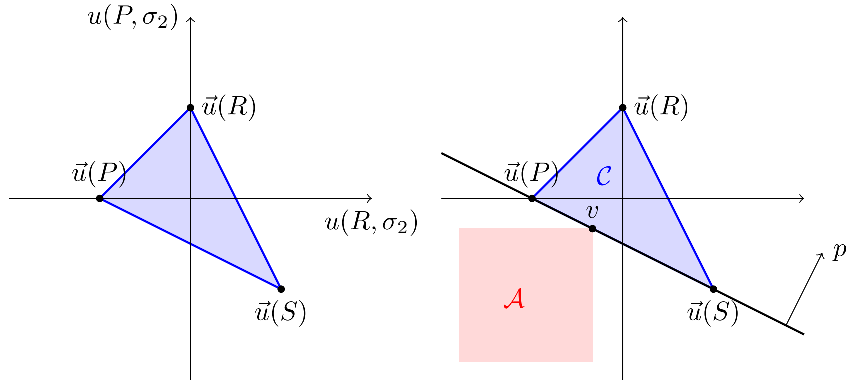
Here the blue set is $\mathcal{C}$ and the pink set is $\mathcal{A}$. It's easy to see that $v = -1/3$ because if we take a convex combination of the payoffs from when P2 plays P and S (clearly P2 is not going to play R, which can only win against the strategy for P1 that we eliminated) we get

$$a(1,-1) + (1-a)(-1,0) = (a-(1-a), -a) = (2a-1, -a)$$

So if we set $2a - 1 = a$ then we get $a = 1/3$ and therefore

$$(1/3)(1,-1) + 2/3(-1,0) = (1/3 - 2/3, -1/3) = -(1/3, 1/3) = (v, v).$$

Player 1's disadvantage is reflected in the game's value being negative.



Remarks

I wrote this note while teaching graduate students at Berkeley and Caltech. The Minimax theorem seems magical. Most popular books on game theory for economists don't seem to include a proof of the Minimax Theorem based on the separating hyperplane theorem. Instead, they present it as a consequence of the existence of Nash equilibrium. This hides, in my view, the main ideas behind its magic. It's also nice for students to see yet another argument using the separation theorem, which is so useful, in so many different contexts, in economics.

Finally, I should mention that the Minimax Theorem is due to John Von Neumann: see [1] for an interesting discussion of the history behind the theorem.

1. Application: Expert Testing

Let Ω be a finite set of states of the world. For example,

$$\Omega = \{\text{rain}, \text{shine}\}^n$$

could be a record of whether it rains or not in the next n days. Today we are uncertain about the weather in the next n days, and we may be interested in forecasting the weather.

Let $\Delta(\Omega)$ be the set of all possible probabilistic ***theories*** about the state of the world. In the weather example, a theory $\mu \in \Delta(\Omega)$ could describe the probabilistic law that governs whether it rains or not in the next n days.

Nature chooses a "true" $\mu \in \Delta(\Omega)$ from which a state of the world will be drawn. Think of μ as the data generating process (dgp). Then theories are simply different possible dgps.

A putatitve expert claims to know the true data generating process. The expert may actually know the correct μ chosen by Nature, or they may be completely ignorant. A charlatan.

An agent, or the public, proposes a test to determine whether the expert knows what they claim. The test consists of a function $T : \Delta(\Omega) \to 2^{\Omega}$. If the expert claims that the true dgp is ν, and the realized state of the world $\omega \in \Omega$ is in $T(\nu)$, then the expert passes the test. Otherwise the expert fails.

Let $\varepsilon > 0$. A test is ***ε-accurate*** if, when the expert knows the true dgp, they are guaranteed to pass the test with probability at least $1-\varepsilon$. Formally, this happens when $\nu(T(\nu)) \geq 1 - \varepsilon$ for all $\nu \in \Delta(\Omega)$.

Theorem 1. *Any ε-accurate test can be passed with probability at least $1 - \varepsilon$ by an expert who does not know the dgp at all, and chooses randomly which theory to report. In symbols: For any ε-accurate test T, there exists $\xi \in \Delta(\Delta(\Omega))$ so that for every true dgp μ,*

$$\xi(\{\nu \in \Delta(\Omega) : \mu\{\omega \in \Omega : \omega \in (T(\nu)\}\}) \geq 1 - \varepsilon.$$

Proof. Fix an ε-accurate test T. Since 2^{Ω} is a finite set, there is a finite set $V \subseteq \Delta(\Omega)$ with $T(V) = T(\Delta(\Omega))$, meaning that for any $\nu \in \Delta(\Omega)$ there is $\nu' \in V$ with $T(\nu') = T(\nu)$.

Consider then a zero-sum game between Player 1, the expert, choosing $\nu \in V$ while Player 2, "Nature," chooses $\omega \in \Omega$. Player 2 pays Player 1 one dollar (or one "util") if ω lies in $T(\nu)$. The payment from 2 to 1 is then $\mathbf{1}_{T(\nu)}(\omega)$. If 1 chooses a mixed strategy $\xi \in \Delta(Y)$ and 2 $\mu \in \Delta(\Omega)$ then the expected payment is

$$\int_Y \int_\Omega \mathbf{1}_{T(\nu)}(\omega) \, \mathrm{d}\mu \, \mathrm{d}\xi = \int_Y \mu(T(\nu)) \, \mathrm{d}\xi.$$

By the minimax theorem, the exists $\xi^* \in \Delta(Y)$ (and hence, $\xi^* \in \Delta(\Delta(\Omega))$) and so that

$$\min_{\mu\in\Delta(\Omega)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi^* = \max_{\xi\in\Delta(Y)} \min_{\mu\in\Delta(\Omega)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi = \min_{\mu\in\Delta(\Omega)} \max_{\xi\in\Delta(Y)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi,$$

But for any μ, Player 2 can choose $\nu \in V$ with $T(\nu) = T(\mu)$, and hence, by the assumption of the test being ε-accurate, $\mu(T(\nu)) = \mu(T(\nu)) \geq 1-\varepsilon$. This implies that

$$\min_{\mu\in\Delta(\Omega)} \max_{\xi\in\Delta(Y)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi \geq 1-\varepsilon.$$

Thus,

$$\min_{\mu\in\Delta(\Omega)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi^* = \min_{\mu\in\Delta(\Omega)} \max_{\xi\in\Delta(Y)} \int_Y \mu(T(\nu))\,\mathrm{d}\xi \geq 1-\varepsilon.$$

Meaning that, when Player 1 (the expert) chooses mixed strategy ξ^*, then for any dgp $\mu \in \Delta(\Omega)$ that Player 2 (nature) may choose, the probability of 1 passing the test is at least $1-\varepsilon$. □

Note: This proof is taken from [2]. I learned about it from Eran Shmaya.

References

[1] Tinne Hoff Kjeldsen. John von neumann's conception of the minimax theorem: A journey through different mathematical contexts. *Archive for history of exact sciences*, 56(1):39–68, 2001.

[2] Alvaro Sandroni. The reproducible properties of correct forecasts. *International Journal of Game Theory*, 32(1):151–159, 2003.