Game Theory
Lecture #9 – Zero-Sum Games

Focus of Lecture:

- Zero-Sum Games

- Value

- Minimax and Maximin

1   Introduction

The last lecture introduced the idea of strategic decision-making in uncertain environments. Here, we focused on scenarios where an individual was tasked with making decisions to maximize her utility (such as whether or not to bring an umbrella). A significant complication that arose in this setting was that the individual was not able to fully assess the ramification of a decision, as her utility was dependent on both her choice and an environmental factor that was out of her control. How should an individual assess the quality of a given decision for such situations?

A central component of the decision-maker's analysis is a predictive model for the environment's behavior. For example, one could have a probabilistic forecast for the environment, e.g., it is going to rain with 50% probability, and evaluate the quality of a decision through an appropriately defined average or expectation. This may be a reasonable path forward in scenarios where an individual has a reliable and comprehensive model of the environment. For situations where such a model is not readily available, an alternative modeling choice is that of a worst-case model. Here, an individual models the environment from a worst-case perspective where she identifies the worst possible environmental choice for each of her possible decisions. Our last lecture demonstrated that worst-case models could be effective for strategic decision-making provided that the individual was able to randomize her choices.

In this lecture we are going to start the transition from systems with a single strategic decision-maker to systems with multiple strategic decision-makers. This lecture will focus on a special class of such systems, termed zero-sum games, that involve two decision-makers with opposed objectives. Here, we will continue to focus on worst-case models and ask the following question: what happens if both decision-makers in a zero-sum game employ worst-case models *against each other?*

2   Zero-Sum Games

In this section we introduce the formal model for zero-sum games. These strategic environments involve two decision-makers with diametrically opposed objectives. The specifics of the model are as follows:

- **Decision-makers:** There are two decision-makers, i.e., $N = \{1, 2\}$. We will use the terms decision-makers, players, and agents interchangeably throughout the text.

- **Choice Sets:** Each decision-maker $i \in N$ is associated with a given choice set $\mathcal{A}_i$. The set of joint choices is defined by $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$. We will use the terms choices and actions interchangeably throughout the text.

- **Utility Function:** Each decision-maker $i \in N$ is associated with a given utility function $U_i : \mathcal{A} \to \mathbb{R}$ that defines her preference over the joint actions $\mathcal{A}$. We will use the terms utilities, payoffs, rewards, and objectives interchangeably throughout the text.

- **Zero-Sum Property:** A zero sum game imposes the constraint that for any choice profile $a = (a_1, a_2) \in \mathcal{A}$ we have that $U_1(a) + U_2(a) = 0$. That is, the payoffs of the decision-makers always sums to 0 for any joint choice $a \in \mathcal{A}$.

In the following we look at two classic examples of zero-sum games and how we represent the components highlighted above.

Example 2.1 (Matching Pennies) *The game matching pennies involves two players each with a choice set $\mathcal{A}_1 = \mathcal{A}_2 = \{H, T\}$, where we refer to H as heads and T as tails. The goal of player 1 is to match the choice of player 2, while the goal of player 2 is to not match the choice of player 1. The set of joint choices is given by $\mathcal{A} = \{(H, H), (H, T), (T, H), (T, T)\}$, where the first entry in each tuple corresponds to the choice of player 1 and the second entry in each tuple corresponds to the choice of player 2. The following utility functions encode the players' preferences defined above where we associate a payoff of 1 if the player's goal is satisfied and -1 otherwise, i.e.,*

$$U_1(a_1 = H, a_2 = H) = 1 \qquad U_2(a_1 = H, a_2 = H) = -1$$
$$U_1(a_1 = H, a_2 = T) = -1 \qquad U_2(a_1 = H, a_2 = T) = 1$$
$$U_1(a_1 = T, a_2 = H) = -1 \qquad U_2(a_1 = T, a_2 = H) = 1$$
$$U_1(a_1 = T, a_2 = T) = 1 \qquad U_2(a_1 = T, a_2 = T) = -1$$

Rather than express each utility function exhaustively as above, here we can use a mild variation of the payoff matrix presented in the previous lecture given by

	H	T
H	$1, -1$	$-1, 1$
T	$-1, 1$	$1, -1$

The row indicates the decision of player 1, the column indicates the decision of player 2, and the cell corresponding to a particular choice for each player, i.e., the intersection of the chosen Row and Column, contains the payoffs to both player 1 (first number) and 2 (second number) for that joint choice. In a zero-sum game, note that the terms in any cell add to 0; hence, we will often only present the payoffs to player 1, which we will refer to as the ROW player, i.e.,

	H	T
H	1,	-1
T	-1	1

Given this minimal representation, we will ofter refer to the ROW *player as the maximizing player, i.e., goal is to maximize the cell number, and the* COL *player as the minimizing player, i.e., goal is to minimize the cell number. Note that minimizing the cell number is consistent with maximizing the negative of the cell number, which is the utility of* COL.

Example 2.2 (Rock - Paper - Scissors) *Consider the classic game Rock - Paper - Scissors which involves two players $N = \{1, 2\}$ and each players is associated with a choice set $\mathcal{A}_1 = \mathcal{A}_2 = \{R, P, S\}$, i.e., Rock, Paper, or Scissors. The set of joint choices is given by $\mathcal{A} = \{(R, R), (R, P), (R, S), \ldots, (S, R), (S, P), (S, S)\}$. If we associate a payoff of 1 to Win, 0 to Tie, and -1 to Loss, we have a zero-sum game with the payoff matrix*

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

As before, the ROW *player is the maximizing player whose goal is to maximize the cell number and the* COL *player is the minimizing player whose goal is to minimize the cell number.*

3 Strategic Behavior in Zero-Sum Games

What choice should each player make in a zero-sum game? Answering this question requires developing a reasonable model of the opposing player's behavior which is a difficult task. One model which seems reasonable in zero-sum games is a worst-case model as discussed extensively in the last lecture. Analyzing the performance of this model requires analyzing all the "what if" scenarios for each player, i.e., what would the opponent do if I play a_i? The performance guarantees attainable through such worst-case models are referred to as security levels, which we formally define below:

Definition 3.1 (Security Levels and Security Strategies) *Consider a two-player zero-sum game with action sets $\mathcal{A}_1$ and $\mathcal{A}_2$ and a utility function for* ROW *given by $U : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$. Note that the utility function for player 2 is given by $-U$. The security level of player 1, i.e., the highest achievable payoff when using a worst-case model, is given by*

$$\underline{v} = \max_{a_1 \in \mathcal{A}_1} \min_{a_2 \in \mathcal{A}_2} U(a_1, a_2). \tag{1}$$

We will refer to a security strategy of player 1 as any action $a_1^{ss} \in \mathcal{A}_1$ that guarantees a payoff of at least $\underline{v}$, i.e., $U_i(a_1^{ss}, a_2) \geq \underline{v}$ for all $a_2 \in \mathcal{A}_2$. The security level of player 2, i.e., the lowest achievable value of U when using a worst-case model, is given by

$$\overline{v} = \min_{a_2 \in \mathcal{A}_2} \max_{a_1 \in \mathcal{A}_1} U(a_1, a_2). \tag{2}$$

We will refer to a security strategy of player 2 as any action $a_2^{ss} \in \mathcal{A}_1$ that guarantees a penalty of at most $\bar{v}$, i.e., $U(a_1, a_2^{ss}) \leq \bar{v}$ for all $a_1 \in \mathcal{A}_1$.

Before proceeding, we will spend just a little time analyzing the $\max \min$ and $\min \max$ operators given above. First, these operators have an order commitment that reads from left to right. For example, in (1) player 1 first commits to a choice $a_1 \in \mathcal{A}_1$ and then agent 2, observing this choice a_1, selects an action a_2 with minimum utility value. Accordingly, we can rewrite (1) as

$$\underline{v} = \max_{a_1 \in \mathcal{A}_1} U(a_1, WC_2(a_1)).$$

where

$$WC_2(a_1) \in \arg\min_{a_2 \in \mathcal{A}_2} U(a_1, a_2)$$

is a worst-case choice of a_2 given knowledge of a_1. Similarly, one can rewrite (2) as

$$\bar{v} = \min_{a_2 \in \mathcal{A}_2} U(WC_1(a_2), a_2).$$

where

$$WC_1(a_2) \in \arg\max_{a_1 \in \mathcal{A}_1} U(a_1, a_2)$$

The following example will shed some light on the computation of security strategies and security levels.

Example 3.1 *Consider a two-player game with payoff matrix given by*

	L	R
T	3	0
B	1	2

Focusing on ROW, *if* ROW *plays T, then the worst-case outcome is 0. Alternatively, if* ROW *plays B, then the worst-case outcome is 1. Accordingly,* ROW*'s security strategy is B and the security level is $\underline{v} = 1$. Focusing on* COL, *if* COL *plays L, then the worst-case outcome (i.e., maximum penalty) is 3. Alternatively, if* COL *plays R, then the worst-case outcome is 2. Accordingly,* COL*'s security strategy is R and the security level is $\bar{v} = 2$.*

Example 3.2 *Consider an alternative two-player game with payoff matrix given by*

	L	R
T	3	0
B	2	1

Focusing on ROW, *if* ROW *plays T, then the worst-case outcome is 0. Alternatively, if* ROW *plays B, then the worst-case outcome is 1. Accordingly,* ROW*'s security strategy is B and the security level is $\underline{v} = 1$. Focusing on* COL, *if* COL *plays L, then the worst-case outcome (i.e., maximum penalty) is 2. Alternatively, if* COL *plays R, then the worst-case outcome is 1. Accordingly,* COL*'s security strategy is R and the security level is $\bar{v} = 1$.*

The above examples appear to demonstrate that $\underline{v} \le \overline{v}$; however, whether or not this is true in general will be addressed in the following section. In Example 3.2 in particular we had $\underline{v} = \overline{v}$; this turns out to be an important concept in zero-sum games. Whenever it is the case that a game has $\underline{v} = \overline{v}$, we will say that such a game has a *value:*

Definition 3.2 (Game Value) *Consider a two-player zero-sum game with action sets $\mathcal{A}_1$ and $\mathcal{A}_2$ with a payoff matrix given by $U : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$. If $\underline{v} = \overline{v}$, we say that the game has a value, and that value is defined as $v := \underline{v} = \overline{v}$.*

4   Comparing Security Levels

In this section we start to explore the relationship between $\underline{v}$ and $\overline{v}$. We begin by providing a lemma pertaining to the min max and max min operators defined above. Informally, this lemma demonstrates that the largest minimum (i.e., max min) is always smaller than the smallest maximum (i.e., min max).

Lemma 4.1 *Let $F : X \times Y \to \mathbf{R}$ be any function where X and Y are discrete sets. Then*

$$\max_{x \in X} \min_{y \in Y} \ F(x, y) \le \min_{y \in Y} \max_{x \in X} \ F(x, y).$$

Proof 4.1 *Suppose $(\tilde{x}, \tilde{y})$ and $(\tilde{x}', \tilde{y}')$ are the solutions associated with the max min and min max problems defined above, i.e.,*

$$F(\tilde{x}, \tilde{y}) \ = \ \max_{x \in X} \min_{y \in Y} \ F(x, y), \tag{3}$$

$$F(\tilde{x}', \tilde{y}') \ = \ \min_{y \in Y} \max_{x \in X} \ F(x, y). \tag{4}$$

Focusing on the pair $(\tilde{x}', \tilde{y})$, we have that

$$F(\tilde{x}', \tilde{y}) \ge \min_{x \in X} F(x, \tilde{y}) = F(\tilde{x}, \tilde{y}) = \max_{x \in X} \min_{y \in Y} \ F(x, y),$$

where the first equality comes from the definition of max min. Furthermore, we also have that

$$F(\tilde{x}', \tilde{y}) \le \max_{y \in Y} F(\tilde{x}', y) = F(\tilde{x}', \tilde{y}') = \min_{y \in Y} \max_{x \in X} \ F(x, y),$$

where the first equality comes from the definition of min max. Combining these two sets of inequalities gives us

$$\max_{x \in X} \min_{y \in Y} \ F(x, y) \le F(\tilde{x}', \tilde{y}) \le \min_{y \in Y} \max_{x \in X} \ F(x, y),$$

which completes the proof.

Given this lemma, we can now state the following proposition which shows how security levels compare in zero-sum games.

Proposition 4.1 *Consider a two-player zero-sum game with action sets $\mathcal{A}_1$ and $\mathcal{A}_2$ and a utility function for player 1 (ROW) given by $U : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$. Then the security levels of the two players satisfy $\underline{v} \leq \overline{v}$.*

This proposition confirms our intuition of the previous section.

5   What about mixed strategies?

In this section we want to explore how the use of mixed strategies impact the resulting security levels and security strategies of the players. To that end, it is convenient to express the payoff matrix as merely a matrix M with rows $\mathcal{I}$, columns $\mathcal{J}$, and game matrix elements m_{ij} which corresponds to the entry $M(i,j)$. Given this notation, the computation of the security levels of the two players is now of the form

$$\underline{v} = \max_{i \in \mathcal{I}} \min_{j \in \mathcal{J}} m_{ij}, \tag{5}$$

$$\overline{v} = \min_{j \in \mathcal{J}} \max_{i \in \mathcal{I}} m_{ij}, \tag{6}$$

and we know from Proposition 4.1 that $\underline{v} \leq \overline{v}$. We will refer to i^* as a maximizing security strategy (or maximinimizer) if

$$\underline{v} \leq m_{i^*j} \quad \text{for all } j$$

i.e., i^* assures a payoff of at least $\underline{v}$. Similarly, we will refer to j^* as a minimizing security strategy (or minimaximizer) if

$$m_{ij^*} \leq \overline{v} \quad \text{for all } i$$

i.e., j^* assures a penalty of at most $\overline{v}$.

The matrix notation above allows us to easily assess the impact of probabilistic strategies on the resulting security levels. Given a row strategy $p \in \Delta(\mathcal{I})$ and col strategy $q \in \Delta(\mathcal{J})$, the expected utility to the row player is defined as

$$\mathcal{U}(p,q) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} p_i \times q_j \times m_{ij} = p^{\mathrm{T}} M q,$$

which is the expected utility to row given that each player is choosing independently accordingly to their respective mixed strategies. While the computation of the expected utilities looks complicated, it can be computed using matrix multiplication as shown above. Accordingly, one can re-define the security levels of the two player when allowing mixed strategies,

$$\underline{v} = \max_{p \in \Delta(\mathcal{I})} \min_{q \in \Delta(\mathcal{J})} p^{\mathrm{T}} M q$$

$$\overline{v} = \min_{q \in \Delta(\mathcal{J})} \max_{p \in \Delta(\mathcal{I})} p^{\mathrm{T}} M q.$$

As before, we could easily show that $\underline{v} \leq \overline{v}$ and that the game has a value of v^* if $\underline{v} = \overline{v} = v^*$. Lastly, we will refer to p^* as a maximizing security strategy (or maximinimizer) if

$$\underline{v} \leq p^{*\mathrm{T}} M q \quad \text{for all } q,$$

and q^* as a minimizing security strategy (or minimaximizer) if

$$p^{\mathrm{T}} M q^* \leq \bar{v} \quad \text{for all } p.$$

Example 5.1 *Recall the previous two-player zero-sum game with payoff matrix given by*

	L	R
T	3	0
B	1	2

When restricting attention to pure strategies, i.e., non-probabilistic strategies, we found the security levels were $\underline{v} = 1$ and $\bar{v} = 2$. Suppose ROW *playing a strategy $(p, 1-p)$, i.e., play T with probability p and B with probability $(1-p)$. Given this strategy, the expected utility of* ROW *satisfies*

$$
\begin{aligned}
\mathcal{U}(p, L) &= 3p + 1(1-p) = 2p + 1 \\
\mathcal{U}(p, R) &= 2(1-p) = 2 - 2p
\end{aligned}
$$

Plotting these expected utilities as in
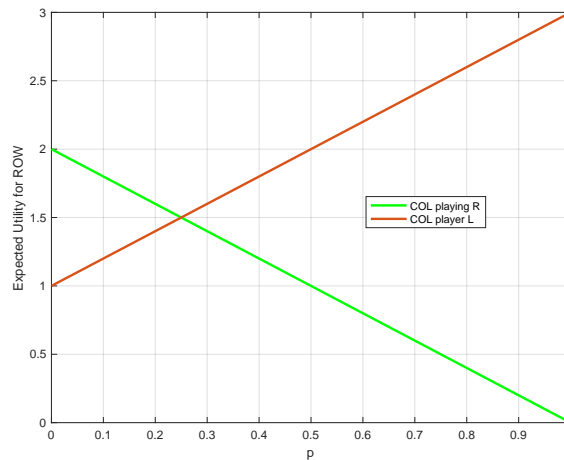
reveals a new security level of $\underline{v} = 1.5$ with a security strategy $p = 1/4$. Repeating this analysis for COL, *if* COL *is playing a strategy $(q, 1-q)$, i.e., play L with probability q and R with probability $(1-q)$, then the expected penalty of* COL *satisfies*

$$
\begin{aligned}
\mathcal{U}(T, q) &= 3q + 0(1-q) = 3q \\
\mathcal{U}(B, q) &= q + 2(1-q) = 2 - q,
\end{aligned}
$$

Plotting these expected penalities as in
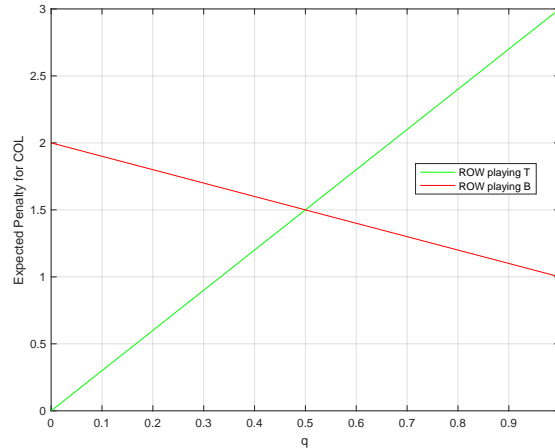
reveals a new security level of $\overline{v} = 1.5$ with a security strategy $q = 1/2$. Note that now over mixed strategies we have $\overline{v} = \underline{v}$.

Observe that $\overline{v} = \underline{v}$ in the example above, i.e., the game has a value. This is in contrast to our previously studied examples where $\overline{v} > \underline{v}$; however, the computation in those example involved pure strategies, i.e., non-probabilistic strategies, as opposed to mixed strategies. The following famous theorem demonstrates that a game will always have a value when considering mixed strategies.

Theorem 5.1 (Minimax Theorem) *Consider any two-player zero-sum game with player set $N = \{1, 2\}$, action sets $\mathcal{A}_1$ and $\mathcal{A}_2$, and payoff matrix $U : \mathcal{A} \to \mathbb{R}$. The game always has a value when considering mixed strategies, i.e.,*

$$\underline{v} = \max_{p \in \Delta(\mathcal{A}_1)} \min_{q \in \Delta(\mathcal{A}_2)} \mathcal{U}(p, q) = \min_{q \in \Delta(\mathcal{A}_2)} \max_{p \in \Delta(\mathcal{A}_1)} \mathcal{U}(p, q) = \overline{v}. \tag{7}$$

The proof of this theorem involves judicious use of the "separating hyperplane" theorem, which is beyond the scope of this class. Nonetheless, this remarkable theorem provides several noteworthy results. First, it establishes that every zero-sum matrix game has a value over mixed strategies. Second, in zero-sum games mixed strategies do often constitute a reasonable prediction of behavior. Lastly, computing the players' security strategies in zero-sum games is relatively straightforward and can be done in a tractable fashion.

6 Conclusion

This lecture established our first significant result pertaining to strategic decision-making in competitive environments. In particular, we demonstrated that any zero-sum game possesses a value, i.e., the security levels of the two players are the same. This also implies a stability condition regarding these security strategies. That is, while a security strategy optimizes worst-case guarantees, it is also a best response to the opponent's security strategy. The following lectures will focus on exploring this stability condition beyond zero-sum games.

7   Exercises

1. A zero-sum game has a payoff (for the row player) given by

$$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$$

   (a) Compute the security strategies for both players using *pure* actions, and conclude that the game does not have a value in this case.

   (b) Repeat using *mixed* strategies, and compute the value of the game.

2. A zero-sum game has a payoff (for the row player) given by

$$\begin{pmatrix} 0 & 3 & 1 \\ 4 & 1 & 2 \end{pmatrix}$$

   (a) Compute the security strategies for both players using *pure* actions, and conclude that the game does not have a value in this case.

   (b) Repeat using *mixed* strategies for the row player.

   (c) Compute the value of the game.

   (d) What is the column player's security level? Identify a security strategy for the column player.

3. Consider the following resource allocation problem known as a Colonel Blotto game

Colonel 1

v^1 v^2
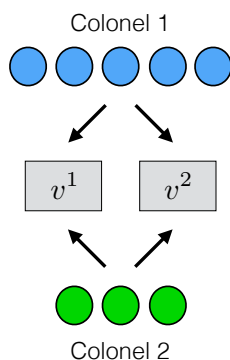
Colonel 2

   There are two Colonels, denoted by $N = \{1, 2\}$, and the number of assets (or soldiers) available to each Colonel is $B_1 = 5$ and $B_2 = 3$. In this problem, each Colonel must simultaneously decide how to allocate its assets over two distinct battlefields that have valuations $v^1 \geq 0$ and $v^2 \geq 0$. A Colonel wins a battlefield when it allocates strictly more assets to that battlefield than its opponent, and consequently the other Colonel loses. The payoff to a Colonel is the total value of the battlefields it won minus the total value of the battlefields it lost (ties give zero payoff to either side). For example, if Colonel 1 allocates 4 assets to battlefield 1 and 1 asset to battlefield 2, which we denote by $x_1 = (4, 1)$, and Colonel 2's allocation is $x_2 = (1, 2)$, the valuation to Colonel 1 is $v^1 - v^2$ and the valuation to Colonel 2 is $v^2 - v^1$.

(a) What are the set of actions available to each Colonel?

(b) What is the payoff matrix of the Colonel Blotto game?

(c) Suppose $v^1 = v^2 = 1$. What are the security levels of each Colonel under pure strategies?

(d) EXTRA CREDIT: What are the security levels of each Colonel under mixed strategies?