CS 497: Planning & Decision Making **Spring 2003**

Lecture 13 — March 4

Lecturer: Steve LaValle *Scribe: Peng Cheng*

13.1 Overview of game theory

In a simple word, game theory is about how multiplayers make decisions without the intervention of nature, as compared with what we learned in the past lectures about how the single player makes decisions with or without nature. A game consists of a set of players $\mathcal{P} = \{P_i\}, i \geq 2$, a set of actions, $\mathcal{A} = \{\mathcal{A}_i\}, i \geq 1$, for each player, and a set of loss function $\mathcal{L} = \{L_i\}, i \geq 1$.

Research on game theory could be roughly divided into the following categories:

1. Single stage v.s. multiple stage. Single stage game means that players in the game only need to make one decision; In a multiple stage game, a sequence of decisions are necessary.

2. Zero sum v.s. non-zero sum. A zero sum game is a game in which one player's loss is the other player's gain, and the sum of loss of all players remains constant. The non-zero sum game is the opposite of the zero sum game.

3. Different information states for each player. This kind of game means that players might have different observation for the same state and/or each player might have his own set of actions such that the information space for each player is different.

4. Deterministic (pure) v.s. randomized (mixed) strategies. This topic is about what kind of strategy the player might take. With the deterministic strategy, the decision the player makes at one state is to choose an action with probability 1; while using

Research field	Optimality	Number of players	Number of loss functions	Number of stages	Nature involved?
Classical optimization	Optimal	1	1	1	No
Decision theory	Expected Optimal	1	1	1	Yes
Multiobjective optimization	Pareto optimal	1	>1	1	Yes/No
Classical(matrix) game theory	Saddle Nash Equil.	>1	>1	1	No
Markov game theory	N/A	>1	>1	1	Yes
Optimal control	Optimal	1	1	>1	No
Stochastic control	Expected optimal	1	1	>1	Yes
Dynamic game theory	Optimal	>1	>1	>1	No
Differential game theory (Team theory)	Optimal	>1	1	>1	Yes/No

Figure 13.1. Relation between various research fields.

randomized strategy, the decision of the player is a probability distribution among available actions.

5. Cooperative v.s. non-cooperative A cooperative game means that there is a common objective for all the players in the game. All the players cooperate with each other to either maximize or minimize the objective; a non-cooperative game means that each player makes decisions to either maximize or minimize his own objective.

Game theory has extensive relation with other research fields, such as classical optimization, decision theory, optimal control. Table 13.1 will be helpful to understand the relation between various related research fields.

		P_2			
		a	b	c	d
	1	1	3	3	-2
P_1	2	0	-1	2	1
	3	-2	2	0	1

Figure 13.2. Loss matrix for a zero-sum game.

13.2 Zero-sum games

An example of zero-sum game is as follows:

The game consists of two players P_1 and P_2. Action set of player 1 P_1 is $\mathcal{A}_1 = \{1, 2, 3\}$. Action set of player 2 P_2 is $\mathcal{A}_2 = \{a, b, c, d\}$. Because the sum of the loss is zero, we could represent loss functions for both players with a loss matrix A shown in Fig. 13.2. The result of using a single matrix A to represent the loss function is that P_1 becomes a minimizer who wants to minimize his loss, and P_2 becomes a maximizer who wants to maximize his gain. However, the sum of their loss and gain is always zero.

For the zero-sum game, we define *upper value*, denoted as $\overline{V}$, and *lower value*, denoted as $\underline{V}$ as follows:

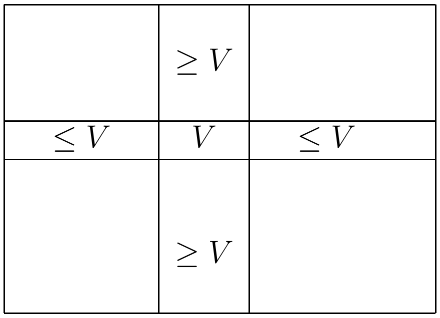
$$\overline{V} = \min_i \max_j L(x_i, y_j), x_i \in \mathcal{A}_1, y_j \in \mathcal{A}_2,$$

$$\underline{V} = \max_i \min_j L(x_i, y_j), x_i \in \mathcal{A}_1, y_j \in \mathcal{A}_2,$$

in which $L(x_i, y_j) = A(i, j)$ as defined in Fig. 13.2.

It is easy to see that $\underline{V} \leq \overline{V}$. For some zero-sum games, it is possible that $\underline{V} = \overline{V} = V$. In this case, we say that the game has a *saddle point.* The structure of the loss matrix with a saddle point is shown in Fig. 13.3, in which V is the smallest number in its row and the largest number in its column. Note, it is possible that some zero-sum games might have several saddle points, but the value of these saddle points is all the same.

When both players use the deterministic strategies and choose actions at the saddle point, they will not regret their decisions, i.e., even after the game, they obtain the best result with respect to their choices.



	$\geq V$	
$\leq V$	V	$\leq V$
	$\geq V$	

Figure 13.3. The structure of the loss matrix with a saddle point.

If randomized strategies are used by both players in the zero-sum game, given $y = [y_1, y_2, \ldots, y_n]^T$ and $z = [z_1, z_2, \ldots, z_m]^T$ as randomized strategies of P_1 and P_2, respectively. The expected loss is

$$E[L] = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} y_i z_j = y^T A z,$$

in which a_{ij} is the element of loss matrix A at i-th row and j-th column. Similarly, we define the upper and lower value for the randomized strategy as

$$\overline{V}_m = \min_y \max_z y^T A z.,$$

$$\underline{V}_m = \max_y \min_z y^T A z.$$

It is easy to verify that $\overline{V}_m \leq \overline{V}$ and $\underline{V}_m \geq \underline{V}$ by considering the deterministic strategy as a special case of the randomized strategy. It was proved by von Newman that $\overline{V}_m = \underline{V}_m$, which means that there always exists a saddle point when randomized strategies are used.

To calculate $\overline{V}_m$ and $\underline{V}_m$, we could use linear programming techniques, which will be shown in the following example.

In the zero-sum game, both players have two elements in their action sets. The loss matrix is shown in Fig. 13.4.

Let y and z are randomized strategies for P_1 and P_2 ,respectively, we could draw two lines in a 2D plane shown in Fig. 13.5, which correspond to line $E[L] = y^T A[0,1]^T$ and $E[L] = y^T A[1,0]^T$ with z fixed at $[0,1]$ and $[1,0]$, respectively. From the picture, we could see that for any given $y = [1 - y_2, y_2]$, $\max_z y^T A z$ is on the upper envelop of the arrangement of two

	P_2	a	b
P_1	1	3	0
	2	-1	1

Figure 13.4. The loss matrix for a zero-sum game.

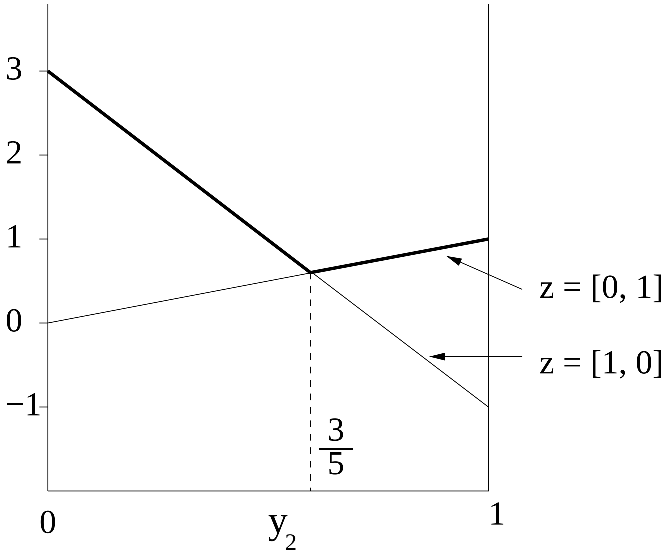


Figure 13.5. Use linear programming to find $\overline{V}_m$.

lines, shown as thick lines in Fig. 13.5. To find y^* such that $\max_z {y^*}^T Az \leq \max_z y^T Az, \forall y$, it is equivalently to find the smallest point on the upper envelop. For this example, the smallest point is at $y_2 = \frac{3}{5}$, and the randomized strategy for P_1 is $[\frac{2}{5}, \frac{3}{5}]$.

13.3 Nonzero sum games

Because the sum of the loss and gain is not zero anymore, two matrices are needed to represent loss functions for two players, and two players will make decisions to minimize their own loss.

An example of loss matrices is shown in Fig. 13.6, in which player P_1 has action set $\{1, 2\}$, and player P_2 has action set $\{a, b\}$.

Similarly to the zero sum games, people are interested in actions for which both players have no regret. For zero sum games, this kind of points are called saddle points. In nonzero

$$\begin{array}{cc|c|c} & & \multicolumn{2}{c}{P_2} \\ & & \text{a} & \text{b} \\ \hline P_1 & 1 & 1 & 0 \\ \hline & 2 & 2 & -1 \\ \hline \end{array} \qquad \begin{array}{cc|c|c} & & \multicolumn{2}{c}{P_2} \\ & & \text{a} & \text{b} \\ \hline P_1 & 1 & 2 & 3 \\ \hline & 2 & 1 & 0 \\ \hline \end{array}$$

A B

Figure 13.6. Loss matrices for a nonzero sum game.

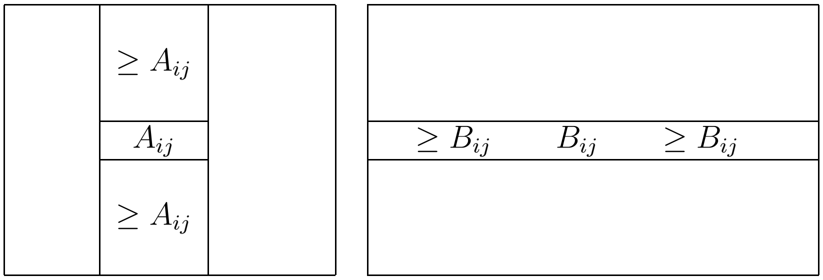


Figure 13.7. The structure of loss matrices with a Nash equilibrium.

sum games, they are called as *Nash equilibria.* The structure of matrices with Nash equilibrium is shown in Fig. 13.7, in which (i, j) is the Nash equilibrium if A_{ij} and B_{ij} are the largest number in j-th column of A and i-th row of B, respectively. When both players make their decisions according to Nash equilibria, they will not feel regret after the game. Note, a nonzero sum game might have either zero or more than one Nash equilibrium. If there are several Nash equilibria, a dominant one might be chosen.

In Fig. 13.6, both $(1, 2)$ and $(-1, 0)$ are Nash equilibria, and $(-1, 0)$ is the dominant equilibrium. An example of the nonzero sum game with no dominant equilibrium is shown in Fig. 13.8, in which both $(-2, -1)$ and $(-1, -2)$ are Nash equilibria, but no one is dominant.

One thing needs to mention is that Nash equilibria will lead to solution with no regret for both players in a non-cooperative game, but the solution does not provide the minimal loss. If cooperative is allowed, there might be better solutions than those corresponding to Nash equilibria. The following Prisoner's Dilemma provides such an example.

One version [1] of the Prisoner's Dilemma is given as follows: "Tanya and Cinque have been arrested for robbing the Hibernia Savings Bank and placed in separate isolation cells.

$$\begin{array}{cc|c|c} & & \multicolumn{2}{c}{P_2} \\ & & \text{a} & \text{b} \\ \hline P_1 & 1 & -2 & 1 \\ \cline{2-4} & 2 & -1 & -1 \end{array} \qquad \begin{array}{cc|c|c} & & \multicolumn{2}{c}{P_2} \\ & & \text{a} & \text{b} \\ \hline P_1 & 1 & -1 & 1 \\ \cline{2-4} & 2 & 2 & -2 \end{array}$$

$$A \qquad\qquad\qquad B$$

Figure 13.8. A nonzero sum game with no dominant equilibria.

$$\begin{array}{cc|c|c} & & \multicolumn{2}{c}{\text{Cinque}} \\ & & \text{C} & \text{S} \\ \hline \text{Tanya} & \text{C} & 8 & 0 \\ \cline{2-4} & \text{S} & 30 & 2 \end{array} \qquad \begin{array}{cc|c|c} & & \multicolumn{2}{c}{\text{Cinque}} \\ & & \text{C} & \text{S} \\ \hline \text{Tanya} & \text{C} & 8 & 30 \\ \cline{2-4} & \text{S} & 0 & 2 \end{array}$$

$$\text{Loss of Tanya} \qquad\qquad \text{Loss of Cinque}$$

Figure 13.9. Loss matrices for the game of Prisoner's Dilemma.

Both care much more about their personal freedom than about the welfare of their accomplice. A clever prosecutor makes the following offer to each. ”You may choose to confess, denoted as C, or remain silent, denoted as S. If you confess and your accomplice remains silent I will drop all charges against you and use your testimony to ensure that your accomplice does serious time. Likewise, if your accomplice confesses while you remain silent, they will go free while you do the time. If you both confess I get two convictions, but I'll see to it that you both get early parole. If you both remain silent, I'll have to settle for token sentences on firearms possession charges. If you wish to confess, you must leave a note with the jailer before my return tomorrow morning.”” Quantizing the loss for both players by number of years in the jail, we obtain the loss matrices in Fig. 13.9.

It is easy to verify that when no cooperation is allowed, both Tanya and Cinque will choose confess and be sentenced to the jail for 8 years, which is not an optimal result but will not cause regret for both players. However, when cooperation is allowed, Tanya and Cinque will both remain silence and only stay in the jail for 2 years.

Bibliography

[1] Prisoner's dilemma. http://plato.stanford.edu/entries/prisoner-dilemma/.