CS497 Planning and Decision Making

WEEK 7 NOTES : MARCH 4,6 2003

Game Theory

Scribe: Shamsi Tamara Iqbal

Topics Covered

- Overview
- Zero Sum Games
  - Saddle Points
  - Mixed Strategies
  - Computing Equilibria
  - Sequential Games
- Non Zero Sum Games
  - Definitions
  - Nash Equilibria

1 Overview

In a game, several agents strive to maximize their (expected) utility by choosing particular courses of action, and each agent's final utility payoffs depend on the courses of action chosen by all agents. The interactive situation specified by the set of participants, the possible courses of action of each agent, and the set of all possible utility payoffs, is called a **game**; the agents 'playing' a game are called the **players**.

Game theory is a set of analytical tools designed to help us understand the phenomena that we observe when decision-makers interact. The basic assumptions that underlie the theory are that decision-makers pursue well-defined exogenous objectives (they are rational) and take into account their knowledge or expectations of other decision-makers' behavior (they reason strategically).

Some of the areas of game theory that we are going to look into are:

- **Multiple Decision Makers:** There will be two or more decision makers, trying to make decisions at the same time.
- **Single stage v Multiple stage**
- **Zero sum v Non zero sum games:** Zero-sum games are games where the amount of "winnable goods" (or resources) is fixed. Whatever is gained by one agent, is therefore lost by the other agent: the sum of gained (positive) and lost (negative) is zero.
  In Non-zero-sum games there is no universally accepted solution. That is, there is no single

optimal strategy that is preferable to all others, nor is there a predictable outcome. Non-zero-sum games are also non-strictly competitive, as opposed to the completely competitive zero-sum games, because such games generally have both competitive and cooperative elements. Players engaged in a non-zero sum conflict have some complementary interests and some interests that are completely opposed.

- **Different Information States for each player:** Each player has an information set corresponding to the decision nodes, which are used to represent situations where the player may not have complete knowledge about everything that happens in a game. Information sets are unique for each player.

- **Deterministic v Randomized Strategies:** When the player uses a deterministic or pure strategy, the player specifies a choice from his information set. When a player uses a mixed strategy, he plays unpredictably in order to keep the opponent guessing.

- **Cooperative v Noncooperative:** A player may be interpreted as an individual or as a group of individuals making a decision. Once we define the set of players, we may distinguish between two types of models: those in which the sets of possible actions of individual players are primitives and those in which the sets of possible joint actions of groups of players are primitives. Models of the first type can be referred to as "noncooperative", while those of the second type can be referred to as "cooperative".

The following table summarizes some of the above mentioned features. We make the following two assumptions:

- Players know each other's loss functionals.
- Players are rational decision makers.

# of players	# of steps	Nature ?	Loss Functionals	Example
1	1	N	1	Classical Optimization
1	1	Y	1	Decision Theory
> 1	1	N	> 1	Matrix Games
> 1	1	Y	> 1	Markov Games (probabilistic)
1	>1	N	1	Optimal Control Theory
1	>1	Y	1	Stochastic Control
>1	>1	N/Y	> 1	Dynamic Game Theory
1	1	N	> 1	Multi-objective Optimality
>1	>1	N/Y	1	Team Theory

2 Single Stage two Player Zero Sum Games

The most elementary type of two-player zero sum games are *matrix games*. The main features of such games are:

- There are two players P_1 and P_2 and an $(\mathrm{m} \times \mathrm{n})$ dimensional loss matrix $\mathrm{A} = \{a_{ij}\}$.
- Each entry of the matrix is an outcome of the game, corresponding to a particular pair of decisions made by the players.

- For P_1, the alternatives are the m rows of the matrix and for P_2, the alternatives are the n columns of the matrix. These are also known as the strategies of the players and can be expressed in the following way:

$$\begin{aligned} U^1 &= u_1^1, u_2^1, \ldots, u_m^1 \\ U^2 &= u_1^2, u_2^2, \ldots, u_n^2 \end{aligned}$$

- Both players play simultaneously.
- If P_1 chooses the ith row and P_2 chooses the jth column, then a_{ij} is the outcome of the game and P_1 pays this amount to P_2. In case a_{ij} is negative, this should be interpreted as P_2 paying P_1 the positive amount corresponding to this entry.

  More formally, for each pair $< U_i^1, U_j^2 >$,
  P_1 has loss $L_1(U_i^1, U_j^2)$ and
  P_2 has loss $L_2(U_i^1, U_j^2) = -L_1(U_i^1, U_j^2)$
  We can write the loss functional as simply L, where P_1 tries to minimize L and P_2 tries to maximize L.

Example:

Suppose the loss matrix for players P_1 and P_2 is as below:

# of players	# of steps	Nature ?	Loss Functionals	Example
1	1	N	1	Classical Optimization
1	1	Y	1	Decision Theory
> 1	1	N	> 1	Matrix Games
> 1	1	Y	> 1	Markov Games (probabilistic)
1	>1	N	1	Optimal Control Theory
1	>1	Y	1	Stochastic Control
>1	>1	N/Y	> 1	Dynamic Game Theory
1	1	N	> 1	Multi-objective Optimality
>1	>1	N/Y	1	Team Theory

In order to illustrate the above mentioned features of matrix games, let us consider the following (3×4) matrix.

$$\begin{array}{cc} & P_2 \\ P_1 & \begin{array}{|c|c|c|c|} \hline 1 & 3 & 3 & 2 \\ \hline 0 & -1 & 2 & 1 \\ \hline -2 & 2 & 0 & 1 \\ \hline \end{array} \end{array}$$

In this case, P_2, who is the maximizer, has a unique *security strategy*, "column 3" ($j^* = 3$), securing him a gain-floor $\underline{V} = 0$. P_1, who is the minimizer, has two security strategies, "row 2" and "row 3" ($i_1^* = 2, i_2^* = 3$) yielding him a loss ceiling of $\overline{V} = \max_j a_{2j} = \max_j a_{3j} = 2$ which is above the security level of P_2.
We can express this more formally in the following notation:
Security strategy for $P_1 = argmin_i \max_j L(U_i^1, U_j^2)$

Therefore, loss-ceiling or upper value $\overline{V} = \min_i \max_j L(U_i^1, U_j^2)$

Security strategy for $P_2 = argmax_i \min_j L(U_i^1, U_j^2)$
Therefore, gain-floor or lower value $\underline{V} = \max_i \min_j L(U_i^1, U_j^2)$

2.1 Regret

If P_2 plays first, then he chooses column 3 as his security strategy and P_1's unique response would be row 3, yielding an outcome of $\underline{V} = 0$. However, if P_1 plays first, then he is indifferent between his two security strategies. In case he chooses row 2, P_2 will respond with choosing column 2 and if P_1 chooses row 3, then P_2 chooses column 2, both strategies resulting in in outcome of $\overline{V} = 2$. This means that when there is a definite order of play, security strategies of the player who acts first make complete sense and they can be considered to be in equilibrium with the corresponding response strategies of the other player. By the two strategies being in equilibrium, it is meant that after the game is over and its outcome is observed, the players should have no ground to regret their past actions. Therefore, in a matrix game with a fixed order of play, for example, there is no justifiable reason for a player who acts first to regret his security strategy.

In matrix games where players *arrive at their decisions independently* the security strategies cannot possibly possess any sort of equilibrium. To illustrate this, we look at the following matrix:

	P_2		
	4	0	-1
P_1	0	-1	3
	1	2	1

We assume that the players act independently and the game is to be played only once. Both players have unique security strategies, "row 3" for P_1 and "column 1" for P_2, with the upper and lower values of the game being $\overline{V} = 2$ and $\underline{V} = 0$ respectively. If both players play according to their security strategies, then the outcome of the game is 1, which is midway between the security strategies of the players. But after the game is over, both P_1 and P_2 might have regrets. This indicates that in this matrix game, the security strategies of the players cannot possibly possess any equilibrium property.

2.2 Saddle Points

For a class of matrix games with equal upper and lower values, a dilemma regarding regret does not arise. If there exists a matrix game where $\overline{V} = \underline{V} = V$ then we say that the strategies are in equilibrium, since each one is optimal against the other.The strategy pair (row x, col y), possessing all the favorable features is clearly the only candidate that can be considered as the equilibrium of the matrix game.

Such equilibrium strategies are known as *saddle point strategies* and the matrix game in question is said to have a *saddle point* in *pure strategies*.

There can also be multiple saddle points as shown in the following figure:

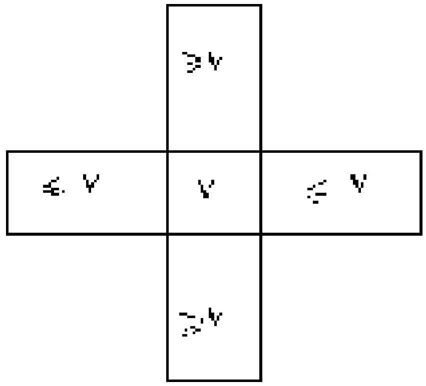

Figure 1: Saddle point

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

2.3 Mixed Strategies

Another approach to obtain equilibrium in a matrix game that does not possess a saddle point and in which players act independently is to enlarge the strategy spaces so that the players can base their decisions on the outcome of the random events - this strategy is called *mixed strategy* or *randomized strategy*. Unlike the pure strategy case, here the same game is allowed to be played over and over again, and the final outcome, sought to be minimized by P_1 or maximized by P_2 is determined by averaging the outcomes of the individual outcomes.

A strategy of a player can be represented by probability vectors. Suppose the strategy for P_1 is represented by

$y = [y_1, y_2, \ldots, y_n]^T$ where $y_i \geq 0$ and $\sum y_i = 1$

and the strategy for P_2 is represented by

$z = [z_1, z_2, \ldots, z_n]^T$ where $z_i \geq 0$ and $\sum z_i = 1$

Let A be the loss matrix. Therefore,
Expected loss for $\mathbf{P_1}$ is,

$$
\begin{aligned}
E[L_1] &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} y_i z_j \\
&= y^T A z
\end{aligned}
$$

Note: Az is the expected losses over nature's choices, given P_1's actions. Az makes P_2 look like nature(probabilistic) to P_1.

Expected loss for $\mathbf{P}_2$ is,

$$
E[L_2] = -E[L_1]
$$

It turns out that we can always find a saddle point in the space of mixed strategies.

2.3.1 Mixed Security Strategy

A vector $y^* \epsilon Y$ is called a *mixed security strategy* for P_1 in the matrix game A, if the following inequality holds $\forall_Y$:

$$
\overline{V}_m(A) = \max_{z \epsilon Z} y^{*\prime} A z \leq \max_{z \epsilon Z} y' A z \qquad y \epsilon Y \tag{1}
$$

Here, the quantity $\overline{V}_m$ is known as the average security level of P_1.

Analogously, a vector $z^* \epsilon Z$ is called a *mixed security strategy* for P_2 in the matrix game A, if the following inequality holds $\forall_Z$:

$$
\underline{V}_m(A) = \min_{y \epsilon Y} y^{*\prime} A z \leq \min_{y \epsilon Y} y' A z \qquad z \epsilon Z \tag{2}
$$

Here, the quantity $\underline{V}_m$ is known as the average security level of P_2.

From eq. (1), we have,

$$
\overline{V}_m \leq \overline{V} \tag{3}
$$

Similarly, from eq. (2):

$$
\underline{V} \leq \underline{V}_m \tag{4}
$$

Therefore, combining eq. (3) and eq. (4), we have:

$$
\underline{V} \leq \underline{V}_m \leq \overline{V}_m \leq \overline{V} \tag{5}
$$

According to Von Neumann, $\underline{V}_m$ and $\overline{V}_m$ always equal. So eq. (5) can be written as :

$$
\underline{V} \leq \underline{V}_m = \overline{V}_m \leq \overline{V} \tag{6}
$$

which essentially means that there always exists a saddle point for mixed strategies.

2.4 Computation of Equilibria

It has been shown that a two person zero-sum matrix game always admits a saddle point equilibrium in mixed strategies. An important property of mixed saddle point strategies is that, for each player there is a mixed security strategy and for each mixed security strategy there is a corresponding mixed saddle point strategy. Using this property, there is a possible way of obtaining the saddle point solution of a matrix game, which can be used to determine the mixed security strategies for each player.

Let us consider the following (2×2) matrix game:

$$\begin{array}{cc} & \mathrm{P}_2 \\ \mathrm{P}_1 & \begin{array}{|c|c|} \hline 3 & 0 \\ \hline -1 & 1 \\ \hline \end{array} \end{array}$$

Let the mixed strategies of y and z be defined as follows:

$$\begin{aligned} y &= [y_1, y_2]^T \\ z &= [z_1, z_2]^T \end{aligned}$$

For P_1, our goal is to find the y^* that optimizes $y^T Az$ while P_2 is trying to do his best, i.e. P_2 uses only pure strategies. Therefore, P_2 can be expected to play either $(z_1 = 1, z_2 = 0)$ or $(z_1 = 0, z_2 = 1)$ and under different choices of mixed strategies for P_1, we can determine the average outcome of the game as shown in Fig 3 by the bold line, which forms the upper envelope to the straight lines drawn. Now, if the mixed strategy $(y_1^* = \frac{2}{5}, y_2^* = \frac{3}{5})$ corresponds to the lowest point of that envelope adopted by P_1, then the average outcome will be no greater than $\frac{3}{5}$. This implies that the strategy $(y_1^* = \frac{2}{5}, y_2^* = \frac{3}{5})$ is a mixed security strategy for P_1 (and his only one), and thereby, it is his mixed saddle point strategy. From the figure, we can see that the mixed saddle point value is $\frac{3}{5}$.

In order to find z^*, we assume the P_1 adopts pure strategies. Therefore for different mixed strategies of P_2, the average outcome of the game can be determined to be the bold line, shown in Fig. 4, which forms the lower envelope to the straight lines drawn. Since P_2 is the maximizer, the highest point on this envelope is his average security level. This he can guarantee by playing the mixed strategy which is also his saddle point strategy.

2.4.1 Solving matrix games with larger dimensions

One alternative to the graphical solution described above when the dimensions are large (i.e. $n \times m$ games) is to convert the original matrix game into a linear programming model and make use of the powerful algorithms for linear programming in order to obtain the saddle point solutions.

This equivalency of games and LP may be surprising, since a LP problem involves just one decision-maker, but it should be noted that with each LP problem there is an associated problem called the dual LP. The optimal values of the objective functions of the two LPs are equal, corresponding to the value of the game. When solving LP by simplex-type methods, the optimal solution of the dual problem also appears as part of the final tableau.

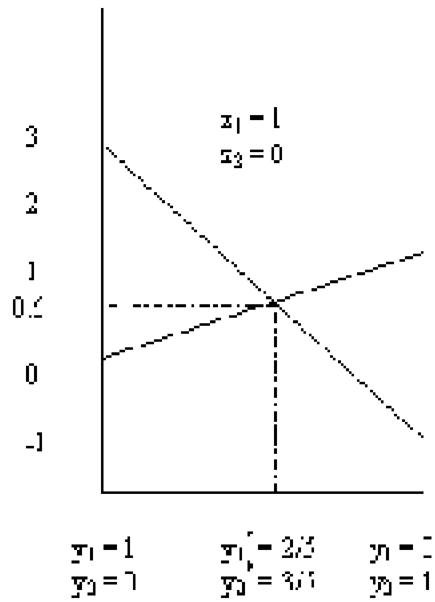

Figure 2: Mixed Security strategy for P_1 for the matrix game

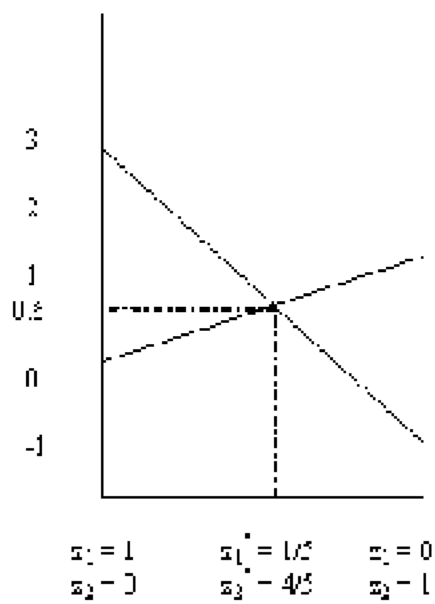

Figure 3: Mixed Security strategy for P_2 for the matrix game

3 Non Zero Sum Games

The branch of Game Theory that better represents the dynamics of the world we live in is called the theory of non-zero-sum games. Non-zero-sum games differ from zero-sum games in that there is no universally accepted solution. That is, there is no single optimal strategy that is preferable to all others, nor is there a predictable outcome. Non-zero-sum games are also non-strictly competitive, as opposed to the completely competitive zero-sum games, because such games generally have both competitive and cooperative elements. Players engaged in a non-zero sum conflict have some complementary interests and some interests that are completely opposed.

3.1 Nash Equilibria

A bi-matrix game is comprised of two $(m \times n)$ dimensional matrices A $= \{a_{ij}\}$ and B $= \{b_{ij}\}$ where each pair of entries $\{a_{ij}\, b_{ij}\}$ denote the outcome of the game corresponding to a particular pair of decisions made by the players. Being a rational decision maker each player will strive for an outcome which provides him with the lowest possible loss.

Assuming that there are no cooperations between the players and the players make their decisions independently, we now try to find out a noncooperative equilibrium solution. The notion of saddle points in zero sum games is also relevant in non zero sum games, where the equilibrium solution is expected to exist if there is no incentive for any unilateral deviation for the players. Therefore, we have the following definition:

Definition 3.1 *A pair of strategies {row i^*, column j^* } is said to constitute a* ***Nash Equilibrium*** *if the following pair of inequalities is satisfied for all i = 1,...,m and all j = 1,...,n:*

$$\begin{aligned} a_{i^*j^*} &\leq a_{ij^*} \\ b_{i^*j^*} &\leq b_{ij^*} \end{aligned}$$

We use a 2 player, single stage game to illustrate the features of a non zero sum game. A and B are the two players, each of them have individual loss functions P_1 and P_2 respectively. The loss functions are represented by the following two matrices:
For A:

$$\begin{array}{cc} & P_2 \\ P_1 & \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 2 & -1 \\ \hline \end{array} \end{array}$$

and for B:

$$\begin{array}{cc} & P_2 \\ P_1 & \begin{array}{|c|c|} \hline 2 & 3 \\ \hline 1 & 0 \\ \hline \end{array} \end{array}$$

It admits two Nash equilibria, {row 1, col 1} and {row 2, col 2}. The corresponding Nash equilibria is (1,2) and (-1,0).

3.2 Betterness and Admissibility

The previous example shows that a bi-matrix game can admit more than one Nash equilibrium solution, with the equilibrium outcomes being different in each case. This raises the question whether there is a way of choosing one equilibrium over the other. We introduce the concepts of betterness and admissibility as follows:

Betterness

A pair of strategies {row i_1, column j_1}is said to be better than another pair of strategies {row i_2, column j_2} if $a_{i_1 j_1} \leq a_{i_2 j_2}$ and $b_{i_1 j_1} \leq b_{i_2 j_2}$ and if at least one of these inequalities is strict.

Admissibility

A Nash equilibrium strategy pair is said to be admissible if there exists no better Nash equilibrium strategy pair.

In the given example, {row 2 ,column 2} is the one that is admissible out of the two Nash equilibrium solutions, since it provides lower costs for both players. This pair of strategies can be described as the most reasonable noncooperative equilibrium solution of the bi-matrix game. In the case when a bimatrix game admits more than one admissible Nash equilibrium the choice becomes more difficult. If the two matrices are as follows:
For A:

$$\begin{array}{cc} & \mathrm{P}_2 \\ \mathrm{P}_1 & \begin{array}{|c|c|}\hline -2 & 1 \\ \hline -1 & -1 \\ \hline \end{array} \end{array}$$

and for B:

$$\begin{array}{cc} & \mathrm{P}_2 \\ \mathrm{P}_1 & \begin{array}{|c|c|}\hline -1 & 1 \\ \hline 2 & -2 \\ \hline \end{array} \end{array}$$

there are two admissible Nash equilibrium solutions{ row 1, column 1}, {row 2, column 2} with the equilibrium outcomes being (-2,-1) and (-1,-2)respectively. This game can lead to regret unless some communication and negotiation is possible.

However if the equilibrium strategies are interchangeable then the ill-defined equilibrium solution accruing from the existence of multiple admissible Nash equilibrium solution can be resolved. This necessarily requires the corresponding outcomes to be the same. Since zero sum matrix games are special types of bi-matrix games (in which case the equilibrium solutions are known to be interchangeable), it follows that there exists some non empty class of bi-matrix games whose equilibrium solutions possess such a property. More precisely :

Multiple Nash equilibria of a bimatrix game (A,B) are interchangeable if (A,B) is strategically equivalent to (A,-A).

3.3 The Prisoner's Dilemma

The following example shows how by using Nash's equilibrium, the prisoners can achieve results that yield no regrets, but how by cooperating, they could have done much better. We show the cost of cooperation and denial of wrong doing in form of the following two matrices:
For A:

$$\begin{array}{cc} & \mathrm{P}_2 \\ \mathrm{P}_1 & \begin{array}{|c|c|}\hline 8 & 0 \\ \hline 30 & 2 \\ \hline \end{array} \end{array}$$

and for B:

		P_2	
P_1		8	30
		0	2

Using Nash equilibrium, the choice is (8,8) which yields no regret for either A or B. However, if the prisoners had cooperated then they would have ended up with (2,2) which is much better for both of them.