Chapter 9: Two-player zero-sum games

A *game*, in the mathematical sense, consists of several players (at least 2), each of which can choose among a set of *strategies*. When all players pick a strategy, the game is played and results in *payoffs* being distributed to every player. By convention the larger the payoff, the better and the general goal for the players is to maximize their payoffs.

In this section we consider two-player zero sum games. Zero-sum games are the games closest to what we think as game and the term zero-sum refers that when the game is played, if one player receives a payoff (or a *gain*) of a, then the other player incurs a *loss* of a,

To describe a zero-sum game we need to specify all the strategies for the two players. We call the two players R (also stands for "row") and C (also stands for "column"). If player R use strategies i and palyer C use strategies j then we denote

$$a_{ij} = \text{ gain for player } R = \text{ loss for player } R$$

We see then that we can specify a game completely by using a $m \times n$ matrix A with entries a_{ij}. Here m is the number of rows of A, which are indexed by all the strategies at R disposal and n is the number of columns of A which are indexed by all the strategies at C disposal.

Matching pennies: This children is game is played as follows. The players R=Robert and C=Collin each hold a penny which they may show either as heads or tail. If both pennies coincide then Robert wins and takes Collin's penny while if they do not match Collin wins and takes Robert's penny. The payoffs are given by the matrix

$$A = \begin{array}{c} \\ H \\ T \end{array} \begin{array}{c} \begin{array}{cc} H & T \end{array} \\ \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array}$$

Playing head or tail is not a winning strategy and it should be intuitively clear that the best option is is to play head or tail at random.

Rock-Scissor-Paper: This other well-known children game has three strategies and we have

Scissors cuts Paper cover Rocks crushes Scissors

which means that Rock wins against scissor, scissor wins against paper, paper wins over rock, that is the strategies cyclically dominate each other.

We shall encounter other situations where this strategic structure occur (including in nature) but for now you may also think of

Man eats chicken eats worm eats .....

and the game matrix is

$$A = \begin{array}{c} \\ R \\ S \\ P \end{array} \begin{array}{c} \begin{array}{ccc} R & S & P \end{array} \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array}$$

Every child will tell you that the best way to play is to pick a random strategy.

Non-transitive dice: Non-transitive dice are dice with non-standard face which display a strategic structure similar to rock-scissor-paper. There are various versions of such dice. For example consider three dice whose faces are marked wiht the following numbers.

$$\text{RED} = 2, 2, 4, 4, 9, 9 \quad \text{BLUE} = 1, 1, 6, 6, 8, 8 \quad \text{GREEN} = 3, 3, 5, 5, 7, 7$$

The game consists of tow players picking each a dice and rolling them. The one woth the higher number wins. If the loser pays 1 to the winner we use as payoff the expected gain. By conidtioning on the outcome of red dice we find

$$P(\text{ RED wins against BLUE}) = \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times 1 = \frac{5}{9}$$

and similarly you may check that

$$P(\text{ BLUE wins against GREEN}) = P(\text{ GREEN wins against RED}) = \frac{5}{9}$$

and so the expected gains will be given by

$$A = \begin{array}{c} \\ R \\ B \\ G \end{array} \begin{array}{c} \begin{array}{ccc} R & B & G \end{array} \\ \begin{pmatrix} 0 & \frac{1}{9} & -\frac{1}{9} \\ -\frac{1}{9} & 0 & \frac{1}{9} \\ \frac{1}{9} & -\frac{1}{9} & 0 \end{pmatrix} \end{array}$$

which is the same as Rock-Scissor-Paper up to a factor 1/9.

The Penalty Kick Game: Let us consider penalty kick in soccer games. We can think of the penalty kicker having two options, either kick $RIGHT$ or $LEFT$, meaning he will aim for the right part or the left part of the net (we assume for simplicity that he never kicks the ball in the middle). In order to catch the ball the goalie must decide to jump either *RIGHT* or jump *LEFT* (by convention if the goalie jumps *RIGHT* it means he will jump to his left to meet the ball that the kicker kicked to his right...) It is reasonable to argue that the kicker and the goalie must decide their strategies essentially simultaneously and to think this as a game theory exercise.

In order to test game theory empirically , two economists decided to watch very many penalty kicks and record empirical payoffs for strategy choices. The found the following payoff matrix where the Kicker is the Row player and the goalie is the Column player.

$$A = \begin{array}{c} \\ L \\ R \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} .58 & .95 \\ .93 & .70 \end{pmatrix} \end{array}$$

with the meaning for example that if the goalie jumps left and the kicker kicks the goal was scored %58 percent of the time. Note that there is an asymmetry between right and left since the majority of people are right-handed (right-footed more precisely...).

We will study what game theory tells us about this game and compare it with reality in a little bit.

Safety strategies and the Von Neumann minimax theorem We will imagine that the players are very *conservative* and want to minimize the risks involved in playing these games. Consider the game payoff matrix

$$\begin{array}{c} \\ \alpha \\ \beta \end{array} \begin{array}{c} \begin{array}{cc} a & b \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \end{array}$$

Since one doesn't know what the other player will do let us argue as follows:

If R chooses α then his worst-case gain is 0 (if C plays b)
If R chooses β then his worst-case gain 0 (if C plays a)

so no matter that C does, R can guarantee himself a payoff of 0.

From the point of view of C (remembering that the payoff matrix entries are C's losses)

If C chooses a then his worst case loss is 1 (if R plays α)
If R chooses b then his worst case loss is 2 (if R plays β)

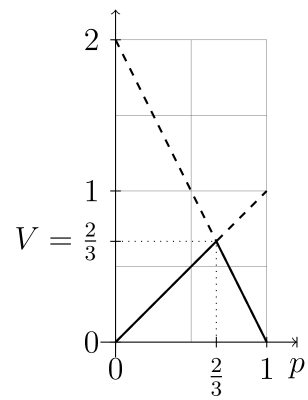
so C can guarantees himself a loss of no more than 1 by choosing strategy a.

We may also wonder what happens and if the players can do better if they *randomize strategies*. From the point of view of R he assumes that he chooses α with probability p and β with probability $1-p$. Then

$$\begin{aligned}&\text{If } C \text{ chooses } a \text{ then } R\text{'s expected gain is } p\times 1+(1-p)\times 0=p\\&\text{If } C \text{ chooses } b \text{ then } R\text{'s expected gain is } p\times 0+(1-p)\times 2=2(1-p)\end{aligned}$$

In that case

$$R\text{'s worst expected gain} = \min(p, 2(1-p))$$



The safety strategy for R is pick the q which maximizes the function $\min(p, 2(1-p))$ and it is easy to see that the maximum is attained exactly when

$$p = 2(1-p) \text{ or } p = \frac{2}{3}$$

in which case his expected gain will be equal to $\frac{2}{3}$. This means that R can guarantee himself a gain of $\frac{2}{3}$ irrespective of what C does.

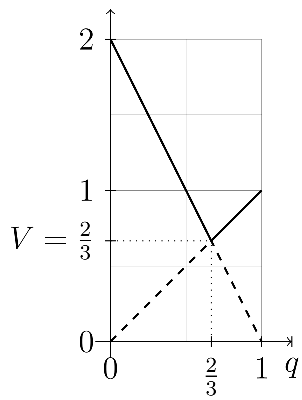
Putting oneself is C's shoes now, we assume that C picks straregy a with probability q and b with probability $1-q$ then

$$\begin{aligned}&\text{If } R \text{ chooses } \alpha \text{ then } C\text{'s expected loss is } q\times 1+(1-q)\times 0=q\\&\text{If } R \text{ chooses } \beta \text{ then } C\text{'s expected loss is } q\times 0+(1-q)\times 2=2(1-q)\end{aligned}$$

In that case

$$R\text{'s worst expected loss} = \max(q, 2(1-q))$$

and C wants to mimimizes his loss given by the function $\max(q, 2(1-q))$.



This is achieved if

$$q = 2(1-q) \text{ or } q = \frac{2}{3}$$

in which case C's expected loss is equal to $\frac{2}{3}$.

This is *remarkable*! By using safety strategies R managed to get a gain of at least $2/3$ while C manages to limits his loss ot $2/3$, exactly what R gains.

$$V = \frac{2}{3} \text{ is called the value of the game.}$$

Also one should realize that have reached some kind of *equilibrium* and that the value of the game is the best possible outcome. Indeed let us assume for example that C is not using his safety strategy of $q^* = 2/3$ but another $q \neq q^*$. Then the expected gain to play α is q and and the payoff to play β is $2(1-q)$. If $q > 2/3$ then $q > 2(1-q)$ and for R' it is better to play α rather than β and therefore R should always play α and use the strategy $p = 1$. In this case the gain for R is equal to $1 \times q + 0 \times 2(1-q) = q$. That is the gain for R and thus the loss for C is $q > 2/3$. Similarly if R knowns that C uses the strategy $q < 2/3$ then he should always play β and ensures himself a gain of $2(1-q) > 2/3$ which causes a loss of more than $2/3$ to C. Therefore C should never move away from q^*. And one can argue in the same way about R that he should not deviate from his safety strategy. Therefore we have learned

Safety strategies are optimal!

What we have demonstrated in this example is actually a general fact. Consider a two-player zero-sum game with Row player R with m strategies and and Column player

C with n strategies. The payoff matrix is

$$A = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ m \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & n \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \end{array} = \text{ Gain for } R = \text{ Loss for } R\,.$$

The set of *mixed or randomized strategies* for R and C are described by probability vectors which assign a certain probability to a certain strategy:

$$\begin{aligned} \Delta_R &= \left\{ p = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix}, p_i \geq 0\,, \sum_{i=1}^m p_i = 1 \right\} \\ \Delta_R &= \left\{ p = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}, q_i \geq 0\,, \sum_{j=1}^n q_j = 1 \right\} \end{aligned}$$

Then the column vector

$$Aq = \begin{pmatrix} (Aq)_1 \\ \vdots \\ (Aq)_m \end{pmatrix}$$

is the vector of payoff for R, i.e. $(Aq)_i$ is the gain for R to play i against the mixed strategy q. Similarly the row

$$p^T A = \big((p^T A)_1, \cdots, (p^T A)_n \big)$$

describes loss for C to play a strategy j against the mixed strategy p and

$$p^T A q$$

is the gain for R (Loss for C) to play strategy p against q.

Definition of safety stragies

1. A strategy p^* is called a *safety strategy* for player R if it maximizes the worst case gain, i.e. the function
$$f(p) = \min_q p^T A q$$
attains its maximum at $p = p^*$. The value
$$f(p^*) = \max_p \min_q p^T A q$$
is called the *safety value* for player R.

2. A strategy q^* is called a *safety strategy* for player C if it minimizes the worst case loss , i.e. the function
$$g(q) = \max_p p^T A q$$
attains its minimum at $q = q^*$. The value
$$g(q^*) = \min_q \max_q p^T A q$$
is called the *safety value* for player C.

The fundamental result of two players zero-sum game is

Von Neumann minimax theorem: *For any two-player zero-sum game with payoff matrix A we have*
$$\max_p \min_q p^T A q = V = \min_q \max_q p^T A q$$
and the number V is called the value of the game.

If the players R uses a safety strategy p^* then R will have an expected gain at most equal to V, no matter which strategy C plays. Moreover if C plays a strategy which is not a safety strategy q^* then R can obtain an expected gain strictly greater than V by using an appropriate strategy.

Conversely if C uses a safety strategy q^* then C will have an expected loss nor more than V, no matter which strategy R plays. Moreover if R plays a strategy which is not a safety strategy p^* then C can obtain an expected loss strictly less than V.

We turn next to several examples where we introduce several techniques to compute game solutions efficiently.

Example 1: The Penalty Kick Game: As discussed before the empirical game payoff matrix for the game is
$$A = \begin{array}{c} \\ L \\ R \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} .58 & .95 \\ .93 & .70 \end{pmatrix} \end{array}$$
If the Kicker use the strategy $(p,(1-p))^T$ then his expected gain will be
$$\begin{aligned} 0.58p + 0.93(1-p) &= 0.93 - 0.35p \quad \text{if Goalie chooses Left} \\ 0.95p + 0.70(1-p) &= 0.70 + 0.25p \quad \text{if Goalie chooses Right} \end{aligned}$$
So his worst case expected gain is $\min(0.93 - 0.35p, 0.70 + 0.25p)$ and the maximum over p is attained if
$$0.93 - 0.35p = 0.70 + 0.25p \Longleftrightarrow p^* = \frac{23}{60} = 0.38333...$$

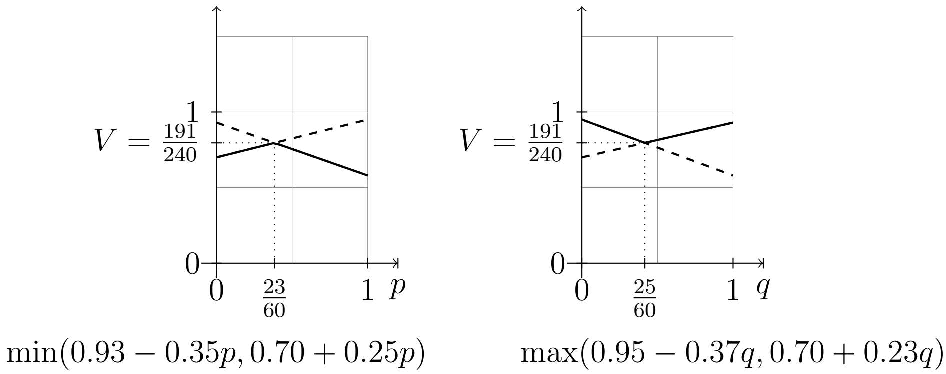


Figure 1: The Penalty-Kick safety strategies

On the other hand if Goalie chooses $(q,(1-q))^T$ then his expected loss is going to be

$$\begin{array}{ll} 0.58q + 0.95(1-q) = 0.95 - 0.37q & \text{if Kicker chooses Left} \\ 0.93q + 0.70(1-q) = 0.70 + 0.23q & \text{if Kicker chooses Right} \end{array}$$

Then Goalie worst case loss is $\max(0.95-0.37q, 0.70+0.23q)$ and the maximum is attained if

$$0.95 - 0.37q = 0.70 + 0.23q \Longleftrightarrow q^* = \frac{25}{60} = 0.416666...$$

The value of the game is the expected gain for *Kicker* and expected loss for *Goalie*

$$\begin{aligned} V &= 0.93 - 0.35p^* = 0.93 - 0.35\frac{23}{60} \\ &= .0.95 - 0.37q^* = 0.95 - 0.37\frac{25}{60} \\ &= \frac{191}{240} = 0.7958333 \end{aligned}$$

which means that about 80 percent of goals are scored.

It is *remarkable* that the empirical values obtained from data is $\widehat{p} = 0.40...$ and $\widehat{q} = 0.42..$ which is very close to the values from minimax. So

Soccer players do play game theory!

Example 2: Rock Paper Scissor It can become complicated to compute safety strategies if there are more than 2 strategies at play. To help we use the fiollowing

Equality of payoff Theorem: Suppose p^* and q^* are optimal strategies. Then we must have

$$\text{If } p_i^* > 0 \text{ and } p_j^* > 0 \text{ then } (Aq^*)_i = (Aq^*)_j\,.$$

and

$$\text{If } p_j^* > 0 \text{ and } p_j^* = 0 \text{ then } (Aq^*)_i \geq (Aq^*)_j \,.$$

proof:. To see why this is true we argue by contradiction. If strategies i and j are played with positive probability but $Aq^*(i) > Aq^*(j)$ then is more advantageous for R to play i rather than j and thus it cannot be an equlibrium! The second statement is argued in a similar way. If the strategy j is not played by R in an optimal strategy then its payoff cannot exceed the payoff played in optimal strategies, otherwise it would be better for R to play it, thus contradicting optimality. ∎

We apply this idea to the Rock Paper Scissor Talking the point of view of Row player we argue as follow: if Collin use the defensive strategy $q = (q_1, q_2, 1 - q_1 - q_2)^T$ then the expected gains for R are

$$Aq = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ 1 - q_1 - q_2 \end{pmatrix} = \begin{pmatrix} 2q_1 + q_2 - 1 \\ 1 - q_1 - 2q_2 \\ q_1 - q_2 \end{pmatrix}$$

Now if Row uses a defensive strategy $p = (p_1, p_2, p_3)$ which assigns positive probability to all three strategies then all payoffs must be equal. Then we must have

$$\begin{aligned} 2q_1 + q_2 - 1 = 1 - q_1 - 2q_2 \quad &\Longrightarrow \quad q_1 + q_2 = \frac{2}{3} \\ 2q_1 + q_2 - 1 = q_1 - q_2 \quad &\Longrightarrow \quad q_1 + 2q_2 = 1 \end{aligned}$$

from which we deduce that

$$q_1 = q_2 = q_3 = \frac{1}{3}$$

as expected. By symmetry between the plkayer we likewise find that $p_1 = p_2 = p_3 = \frac{1}{3}$ and the value of the game is then 0.

Example 3 The Plus One game: Eliminating dominated strategies Consider the following game where two players give a number between say 1 and 10 (or1 to 100...). If the numbers match there is a win; if the numbers differ by 1 the player with the higher number wins \$1 for the other player; if the number differ by 2 or more, then the player with the higher number pays \$2 to the other player. The game. matrix is given by

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{pmatrix} 0 & -1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 0 & -1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & 1 & 0 & -1 & 2 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & 1 & 0 & -1 & 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & 1 & 0 & -1 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & 1 & 0 & -1 & 2 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & 1 & 0 & -1 & 2 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 & 1 & 0 & -1 & 2 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & 0 & -1 \\ -2 & -2 & -2 & -2 & -2 & -2 & -2 & -2 & 1 & 0 \end{pmatrix} \end{array}$$

A priori this looks like a daunting task to compute an optimal strategy but we can use the concept of *domination*. For example from the point of view of R if we compare strategy 1 to any strategy 4 (or higher) we see that *no matter what* C plays the payoff for R is always better if he plays 1 rather than 4. (compare one by one the entries in the first and forth rows of A). But then R should never play strategies 4 and so we can ignore it for all practicakl purpose. In the same way R should never play strategy 4 or higher. Since the game is symmetric C should never play 4 or higher either and so we can reduce ourselves to a 3×3 game

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -1 \\ -2 & 1 & 0 \end{pmatrix} \end{array}$$

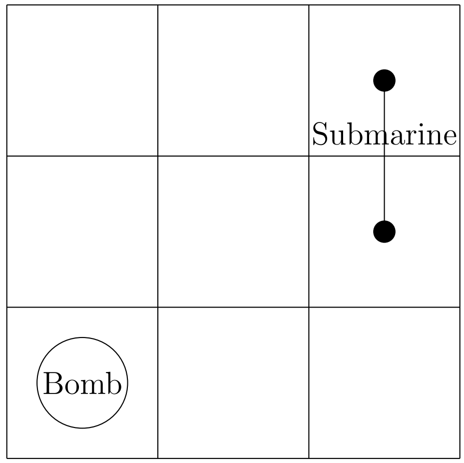
Using now the equality of payoff theorem we look first for optimal strategies p^* and q^* where all $p_i^* > 0$ and so all payoff must be equal and we must have

$$-q_2 + 2q_3 = q_1 - q_3 = -2q_1 + q_2$$

Using $q_3 = 1 - q_1 - q_2$ the first inequality gives $q_1 + q_2 = 3/4$ and so $q_3 = 1/4$. The second inequality gives then $3q_1 - q_2 = 1/4$ and this give the optimal

$$q_1 = \frac{1}{4}, \quad q_2 = \frac{1}{2}, \quad q_3 = \frac{1}{4}.$$

Example 4 Submarine game: Using symmetries Let us imagine the following game between a bomber and a submarine. The game is played on a three by three grid. The submarine is large and so occupies two adjacent square on the grid (see for example the picture below) and the bomber can hit only one square on the grid (the circle in the picture below). The bomber win 1 if it hits the submarine.



The bomber has 9 strategies (one for each square) while the submarine has 12 ways to position the submarine. In order to reduce the complexity of the game we can use some symmetries of the problem. For the comber there are only three sorts of square where to bomb, either in the *center*, in a *corner* or in the *midside*. For the submarine up to symmetries there are only two positions: either the the submarine intersects the c*enter* square (4 such position) or it intersects a *corner* (8 such position). So we write a reduced game matrix. whose entries are the probability for the bomber to hit the submarine. Say if the bombers hits a corner and the submarine intersects a corner the probability to be hit is 1/4, and so on.... We find the game matrix

$$\begin{array}{c} \\ \text{corner} \\ \text{midside} \\ \text{center} \end{array} \begin{array}{c} \begin{array}{cc} \text{center} & \text{corner} \end{array} \\ \begin{pmatrix} 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \\ 1 & 0 \end{pmatrix} \end{array}$$

from which we see that for the bomber the strategy *midside* dominates *corner* which we can then eliminate leading to

$$\begin{array}{c} \\ \text{midside} \\ \text{center} \end{array} \begin{array}{c} \begin{array}{cc} \text{center} & \text{corner} \end{array} \\ \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ 1 & 0 \end{pmatrix} \end{array}$$

Then if submarine knows that the bomber will not bomb *corner* it is better for submarine to go in a corner: *corner* dominate *center* so we have reduced the game to

$$\begin{array}{c} \\ \text{midside} \\ \text{center} \end{array} \begin{array}{c} \text{corner} \\ \begin{pmatrix} \frac{1}{4} \\ 0 \end{pmatrix} \end{array}$$

So the the opitimal strategies are *midside* for the bomber and *corner* for the submarine and thus

$$\text{The value of the game is } V = \frac{1}{4}\,.$$