Game Theory

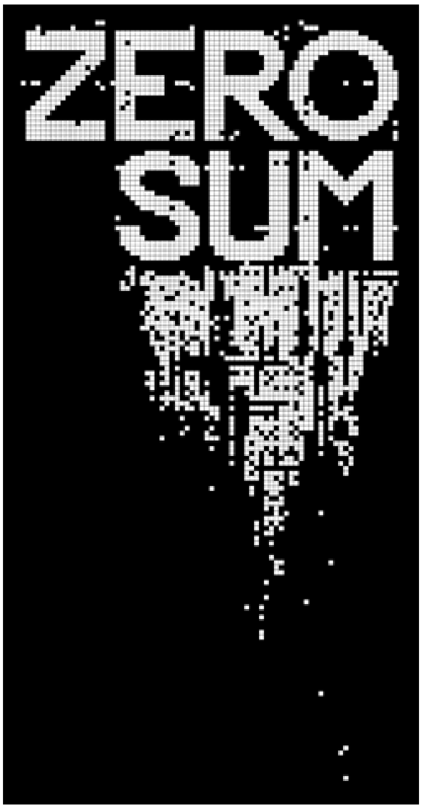
Lecture notes for MATH11090 & MATH09002

Peter Richtárik

University of Edinburgh

November 9, 2010

2-Person Zero-Sum Game Theory: Summary



Slides 13–28 of Lecture 2 cover the **basic theory of zero-sum games**

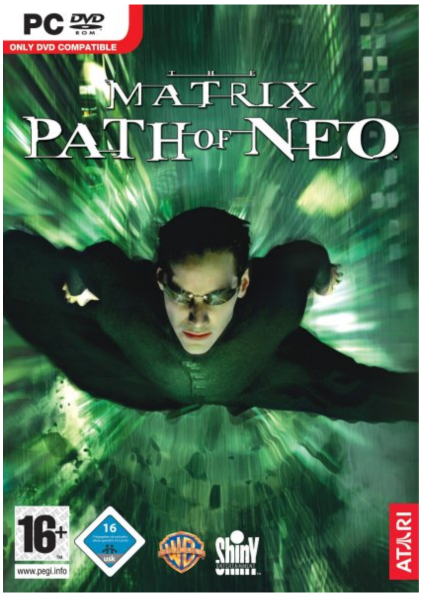
- **No examples were given** to introduce the concepts yet!
- **Little explanation was given** into what the various results and assumptions mean

Let us remedy that now!

We will first illustrate the essential concepts on two games

- "Farmer vs Nature" (Slides 5–15): a **finite** 2-person zero-sum game (i.e., a **matrix game**)
- "Quadratic vs Linear" (Slides 16–22): an **infinite** 2-person zero-sum game (turns out to be a **convex-concave game**)

Matrix Games: 3 Notations



Recall: **Matrix game** = finite 2-person zero-sum game

We now have 3 different notations:

	2-Person Games		
	general	**finite**	**zero-sum**
strategy of P_1	$s_1 \in \Sigma_1$	$p \in R^m$	$s_1 \in \Sigma_1$
strategy of P_2	$s_2 \in \Sigma_2$	$q \in R^n$	$s_2 \in \Sigma_2$
payoff of P_1	$\pi_1(s_1, s_2)$	$p^T Aq$	$-f(s_1, s_2)$
payoff of P_2	$\pi_2(s_1, s_2)$	$p^T Bq$	$f(s_1, s_2)$

The Basic Theorem about Matrix Games

Theorem (Matrix Games)

- *Each matrix game has a value.*
- *Nash equilibrium strategies in a matrix game are the conservative strategies of the players.*

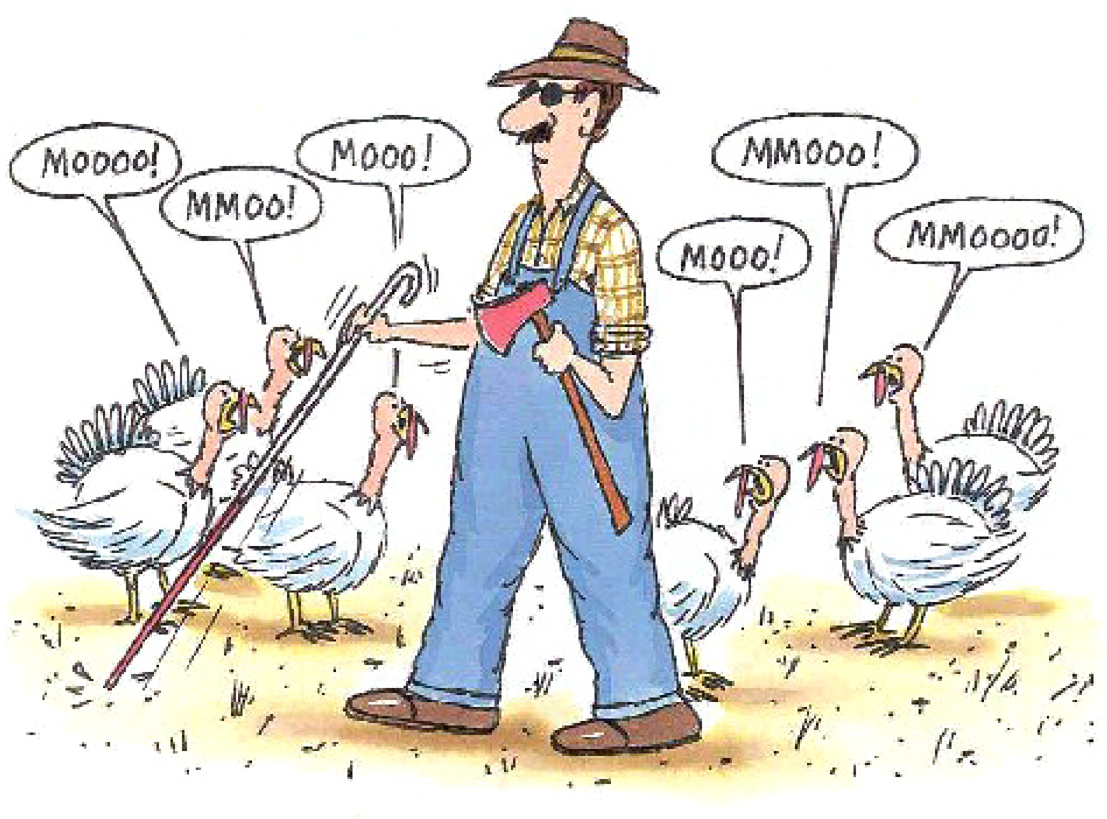
Proof.

Matrix games are convex-concave games if we replace the pure strategy sets S_1, S_2 in the definition of a convex-concave game with the mixed strategy sets Σ_1, Σ_2. This is because

- the sets Σ_1, Σ_2 are convex and
- $f(p, q) = p^T Bq$ is convex in p and concave in q (in fact, it is linear in both).

Matrix games also satisfy the 3 additional conditions of the "Existence of Saddle Points" theorem: Σ_i are bounded and closed (compact), f is continuous and well-defined on $\Sigma_1 \times \Sigma_2$. The result now follows from the "Minimax Equality" theorem, since Saddle Points are Nash Equilibria! $\square$

Farmer vs Nature



Game: Farmer vs Nature

Farmer wants to decide whether to plant **cactuses** or **rice.** His profit per hectare depends on whether the year is **dry** or **wet,** as in the table below:

	Cactuses	Rice
Dry weather	5	1
Wet weather	2	8

Assume that any given year will be dry with probability α and that the farmer decides to use $0 \leq \beta \leq 1$ portion of his land for cactuses and the rest for rice.

Note:

- Farmer is the column player, "nature" is the row player
- Farmer wants to maximize his payoff/profit
- Nature's payoff is not defined, but
  - we can view nature as an adversary who "wants" to minimize farmer's profit
  - nature's payoff = farmer's loss
- This is a **finite** 2-person zero-sum game (i.e. a **matrix game**)

Farmer vs Nature: Questions

Questions:

(a) Find farmer's profit, as a function of β, in a dry year and in a wet year.

(b) Find farmer's expected profit as a function of α and β.

(c) Find the farmer's maximum achievable profit as a function of the weather (α).

(d) What weather will make the farmer's best-response profit as small as possible?

(e) Find the farmer's worst-case profit as a function of his decision (β).

(f) What should the farmer do to guarantee as large a profit as possible?

The Answers

"I don't have any answers. I'm a non-prophet."

Farmer vs Nature: Question (a)

Question: Find farmer's profit, as a function of β, in a dry year and in a wet year.

Solution: We let $q = \binom{\beta}{1-\beta}$ and

$$B = \begin{pmatrix} 5 & 1 \\ 2 & 8 \end{pmatrix}.$$

Then the profit function for each weather is

$$\pi_2(\text{dry}, q) = e_1^T Bq = (5,1)\binom{\beta}{1-\beta} = 5\beta + 1(1-\beta) = 4\beta + 1 \qquad (1)$$

$$\pi_2(\text{wet}, q) = e_2^T Bq = (2,8)\binom{\beta}{1-\beta} = 2\beta + 8(1-\beta) = 8 - 6\beta \qquad (2)$$

Farmer vs Nature: Question (b)

Question: Find the expected profit as a function of α and β.

Solution: Let $p = \binom{\alpha}{1-\alpha}$. Then the expected profit is given by

$$\begin{aligned} f(p,q) = \pi_2(p,q) = p^T Bq &= (\alpha e_1 + (1-\alpha)e_2)^T Bq \\ &= \alpha e_1^T Bq + (1-\alpha) e_2^T Bq \\ &\overset{(1)+(2)}{=} \alpha(4\beta+1) + (1-\alpha)(8-6\beta). \end{aligned}$$

Clearly, this function is **linear** in β for fixed α and **linear** in α for fixed β.

Remarks: In a finite two-person game, the expected payoff is always a bilinear function of p and q. See Slide 18 of Lecture 1:

$$p \mapsto p^T Aq, \quad p \mapsto p^T Bq \qquad \text{are linear for fixed} \quad q$$

$$q \mapsto p^T Aq, \quad q \mapsto p^T Bq \qquad \text{are linear for fixed} \quad p$$

Farmer vs Nature: Question (c)

Question: Find the farmer's maximum achievable profit as a function of the weather (α).

Solution: Look at Slide 15 from Lecture 2, where the worst-case loss of the row player (=best-response profit of the column player) was defined:

$$
\begin{aligned}
u_1(p) &\stackrel{\text{def}}{=} \max_{q \in \Sigma_2} f(p,q) \\
&= \max_{0 \le \beta \le 1} (\alpha, 1-\alpha) B \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} \\
&= \max_{0 \le \beta \le 1} (10\alpha - 6)\beta + (8 - 7\alpha) \\
&= \begin{cases} 3\alpha + 2 & \text{if } \alpha > \frac{3}{5} \quad (\text{since then } \beta^* = 1) \\ \frac{19}{5} & \text{if } \alpha = \frac{3}{5} \quad (\text{since then } \beta^* \in [0,1]) \\ 8 - 7\alpha & \text{if } \alpha < \frac{3}{5} \quad (\text{since then } \beta^* = 0) \end{cases}
\end{aligned}
$$

Farmer vs Nature: Question (d)

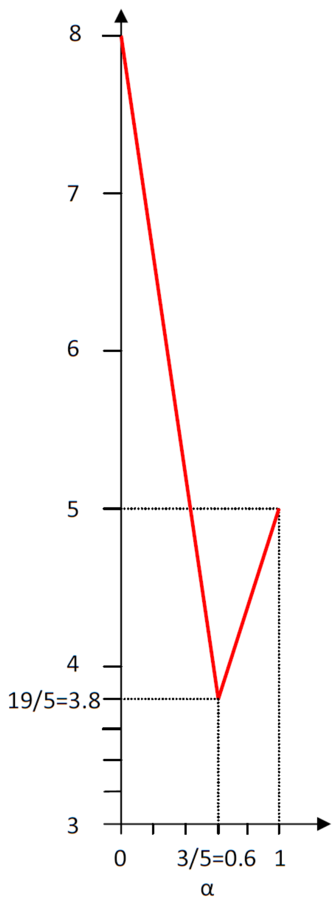
Question: What weather will make the farmer's best-response profit as small as possible?

Solution:

- We need to minimize the function from the previous slide (depicted in the pic).
- The minimum is attained at $\alpha = \frac{3}{5} = 0.6$ **(minimax strategy)**, and is equal to $\frac{19}{5} = 3.8$.

Remarks:

- By choosing weather $\alpha = 0.6$, nature is minimizing the best-response profit of the farmer (i.e., maximizing her worst-case payoff).
- This is the largest profit the nature can guarantee to get.
- $\hat{u}_1 = 3.8$ is the **conservative value** of the row player (nature), defined on Slide 15 of Lecture 2.



Farmer vs Nature: Question (e)

Question: Find the farmer's worst-case profit as a function of his decision (β).

Solution: Look at slide 16 from Lecture 2, where the worst-case profit function was defined:

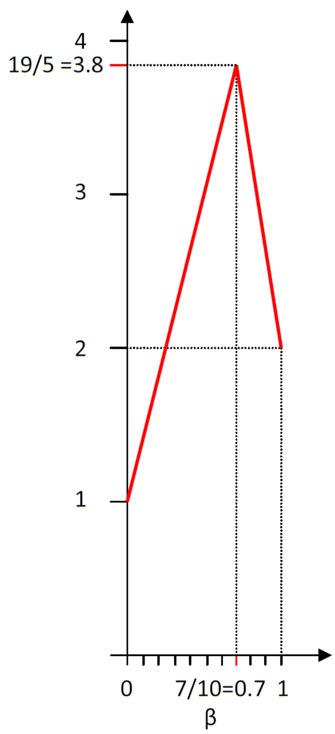
$$
\begin{aligned}
u_2(q) &\stackrel{\text{def}}{=} \min_{p \in \Sigma_1} f(p,q) \\
&= \min_{0 \le \alpha \le 1} (\alpha, 1-\alpha) B \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} \\
&= \min_{0 \le \alpha \le 1} (10\alpha - 6)\beta + (8 - 7\alpha) \\
&= \min_{0 \le \alpha \le 1} (10\beta - 7)\alpha + (8 - 6\beta) \\
&= \begin{cases} 8 - 6\beta & \text{if } \beta > \frac{7}{10} \quad (\text{since then } \alpha^* = 0) \\ \frac{19}{5} & \text{if } \beta = \frac{7}{10} \quad (\text{since then } \alpha^* \in [0,1]) \\ 4\beta + 1 & \text{if } \beta < \frac{7}{10} \quad (\text{since then } \alpha^* = 1) \end{cases}
\end{aligned}
$$

Farmer vs Nature: Question (f)

Question: What should the farmer do to guarantee as large a profit as possible?

Solution: In other words, we need to find β which maximizes the farmer's worst-case profit.

- We need to maximize the function from the previous slide (depicted in the picture).
- It is clear that the maximum is attained at $\beta = 0.7$ **(maximin strategy)**, and is equal to 3.8.



Remarks:

- The farmer can **guarantee** to get a profit of 3.8 by making the planting decision $\beta = 0.7$.
- This is the largest profit he can guarantee to get.
- $\hat{u}_2 = 3.8$ is the **conservative value** of the column player (farmer), defined on Slide 16 of Lecture 2.

Farmer vs Nature: Summary

Comparing the results of (d) and (f):
The conservative value of nature **is equal** to the conservative value of the farmer.

$$\hat{u}_1 = 3.8 = \hat{u}_2$$

- This means that the **Farmer vs Nature game has a value.**
- This was expected as follows from the "Matrix Games" theorem.

An Infinite Convex-Concave Game

Game: Quadratic vs Linear

Consider a 2-person zero-sum game where the strategy set

- of player P_1 is $P = [0, 1]$,
- of player P_2 is $Q = [0, 1]$,

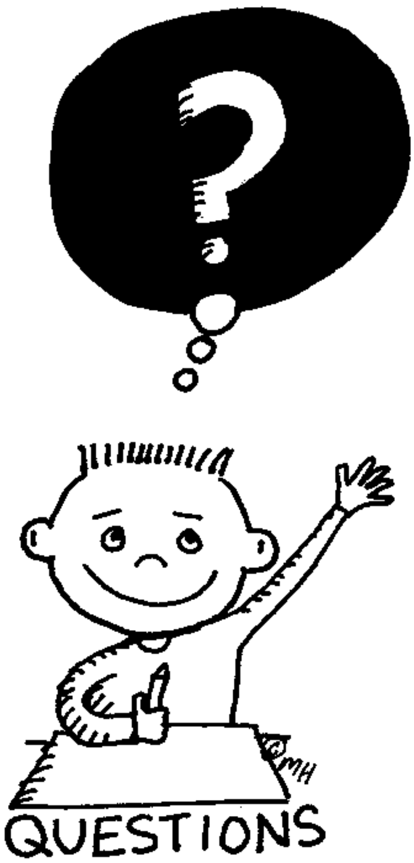
and the payoff of P_2 (i.e., loss of P_1) is given by

$$f(p, q) = (p + \tfrac{1}{2})^2(q + \tfrac{1}{2}), \qquad p \in P, q \in Q.$$

Note:

1. P_1 has control over $(p + \frac{1}{2})^2$ **(quadratic)**
2. P_2 has control over $q + \frac{1}{2}$ **(linear)**
3. This is an **infinite** 2-person zero-sum game since the strategy sets of the players are infinite (intervals)

Quadratic vs Linear: Questions



Questions:

(a) Show that this game must have value.

(b) Compute the minimax strategy of P_1 and his conservative value. What payoff can P_1 guarantee to obtain?

(c) Compute the maximin strategy of P_2 and his conservative value. What payoff can P_2 guarantee to obtain?

(d) Find a saddle point of this game.

Quadratic vs Linear: Question (a)

Question: Show that this game must have value.

Step 1. This is a two-person zero-sum game,

- the strategy sets P, Q are convex (intervals are convex),
- $p \mapsto f(p,q)$ is convex for all q,
- $q \mapsto f(p,q)$ is concave for all p.

Therefore, this is a **convex-concave game** (Lec2, Slide 26).

Step 2. The assumptions of the "Existence of Saddle Points" theorem (Lec2, Slide 27) hold:

(i) P, Q are closed and bounded sets,
(ii) f is defined on $P \times Q$,
(iii) f is continuous.

Therefore, a **saddle point exists!**

Step 3. The "Minimax Equality" theorem (Lec2, Slide 25) now implies that **the game has a value,** i.e.,

$$\hat{u}_1 \stackrel{\text{def}}{=} \min_{p \in P} \max_{q \in Q} f(p,q) = \max_{q \in Q} \min_{p \in P} f(p,q) \stackrel{\text{def}}{=} \hat{u}_2.$$

Quadratic vs Linear: Question (b)

Question: Compute the minimax strategy of P_1 and his conservative value. What payoff can P_1 guarantee to obtain?

Solution:

$$u_1(p) \stackrel{\text{def}}{=} \max_{0 \leq q \leq 1} f(p,q) = \max_{0 \leq q \leq 1} (p + \tfrac{1}{2})^2 (q + \tfrac{1}{2}) = \tfrac{3}{2}(p + \tfrac{1}{2})^2,$$

- Clearly, $u_1(p)$ is minimized at $p^* = 0$: this is the **minimax strategy.**
- Therefore, the **conservative value** of P_1 is

$$\hat{u}_1 \stackrel{\text{def}}{=} \min_{0 \leq p \leq 1} u_1(p) = u_1(p^*) = \tfrac{3}{8}.$$

P_1 can **guarantee** to obtain a payoff of $-\frac{3}{8}$.

Quadratic vs Linear: Question (c)

Question: Compute the maximin strategy of P_2 and his conservative value. What payoff can P_2 guarantee to obtain?

Solution:

$$u_2(q) \stackrel{\text{def}}{=} \min_{0 \le p \le 1} f(p,q) = \min_{0 \le p \le 1} (p + \tfrac{1}{2})^2 (q + \tfrac{1}{2}) = \tfrac{1}{4}(q + \tfrac{1}{2})$$

- Clearly, $u_2(q)$ is maximized at $q^* = 1$: this is the **maximin strategy.**
- Therefore, the **conservative value** of P_2 is

$$\hat{u}_2 \stackrel{\text{def}}{=} \max_{0 \le q \le 1} u_2(q) = u_2(q^*) = \tfrac{3}{8}.$$

P_2 can **guarantee** to obtain a payoff of $\frac{3}{8}$.

Quadratic vs Linear: Question (d)

Question: Find a saddle point of this game.

Solution: The pair of conservative strategies

- $p^* = 0$ (minimax strategy), and
- $q^* = 1$ (maximin strategy),

calculated in problems (b) and (c), respectively, forms a saddle point. To show this, wee need to check whether

$$\min_{0 \le p \le 1} f(p, q^*) = f(p^*, q^*) = \max_{0 \le q \le 1} f(p^*, q).$$

Indeed,

$$\min_{0 \le p \le 1} f(p, 1) = \min_{0 \le p \le 1} (p + \tfrac{1}{2})^2 (1 + \tfrac{1}{2}) = \tfrac{3}{8}$$

$$f(0, 1) = (0 + \tfrac{1}{2})^2 (1 + \tfrac{1}{2}) = \tfrac{3}{8}$$

$$\max_{0 \le q \le 1} f(0, q) = \max_{0 \le q \le 1} (0 + \tfrac{1}{2})^2 (q + \tfrac{1}{2}) = \tfrac{3}{8}$$

6 Insights into 2-Person Zero-Sum Game Theory

On Slides 24–29 we will give

- **6 insights**
- into the theory of **two-person zero-sum games**

- as covered in Lecture 2

Pure vs Mixed Strategy Sets

Insight 1: The theory of two-person zero-sum games developed in Lecture 2 on Slides 13–28 holds for **arbitrary sets** S_1 and S_2 and **arbitrary function** f (go back and see!)

Some of the results will not be very useful when S_1 and S_2 are the sets of pure strategies. For example,

- Matrix games would not belong into the convex-concave category (they do!) because the sets S_i, since finite, cannot be convex!
- Matrix games would not necessarily have a saddle point.

An interesting option is to use

- the **mixed strategy** sets Σ_1, Σ_2 instead of S_1, S_2, and
- the **expected payoff function** of P_2:

$$f(s_1, s_2) = E\pi_2(s_1, s_2), \quad s_1 \in \Sigma_1, s_2 \in \Sigma_2$$

instead of the payoff function $f = \pi_2 : S_1 \times S_2 \to R$

Saddle Points are Nash Equilibria

Insight 2: If we replace S_i by Σ_i and f by the expected payoff in the definition of saddle points on Slide 24 of Lect2, then

Saddle points = Nash equilibria!

That is, the notion of a saddle point coincides with the notion of a NE.

- Check this as an exercise!
- The "Minimax Equality" theorem then says: If (s_1^*, s_2^*) is a NE, then s_1^* (resp. s_2^*) is the conservative strategy of P_1 (resp. of P_2) and $\hat{u}_1 = \hat{u}_2$ (the game has a value)
  - Check this as an exercise.
- In fact, the converse is also true: If s_1^* (resp. s_2^*) is the conservative strategy of P_1 (resp. of P_2) and $\hat{u}_1 = \hat{u}_2$ (the game has a value), then (s_1^*, s_2^*) is a NE
  - Can you prove this?

Conservative Strategies May Not Exist

Insight 3: In the general case, **conservative strategies may not exist!**

Consider a game given by $f(s_1, s_2) = (s_1)^2 s_2, \;\; S_1 = S_2 = (0, 1)$.

- $u_1(s_1) = \sup_{s_2 \in S_2} f(s_1, s_2) = s_1^2 \quad \Rightarrow \quad \hat{u}_1 = \inf_{s_1 \in S_1} u_1(s_1) = 0$
  The **infimum is not attained,** i.e., there is no $s_1 \in S_1$ such that $u_1(s_1) = \hat{u}_1$. That is, P_1 **does not have a conservative strategy.**
- $u_2(s_2) = \inf_{s_1 \in S_1} f(s_1, s_2) = 0 \quad \Rightarrow \quad \hat{u}_2 = \sup_{s_2 \in S_2} u_2(s_2) = 0$
  The **supremum is not attained,** i.e., there is no $s_2 \in S_2$ such that $u_2(s_2) = \hat{u}_2$. That is, P_2 **does not have a conservative strategy.**
- In this example the game has a value though as $\hat{u}_1 = \hat{u}_2 = 0$.
- Nonexistence of conservative strategies can be viewed as a **pathological situation** which should be avoided by proper assumptions about the game (the sets S_1, S_2 and the payoff f).
  - In this game, a simple replacement of the open intervals $(0, 1)$ by the closed intervals $[0, 1]$ will solve the problem.
  - The problem does not occur in convex-concave games satisfying the 3 regularity assumptions in the "Existence of Saddle Points" theorem.

The Trouble with Games Without a Value

Insight 4: Even if conservative strategies exist, the **game might not have a value:** it may be that $\hat{u}_1 > \hat{u}_2$.

What does this mean?

- Recall that by playing their conservative strategies, P_1 can guarantee to get a payoff at least $-\hat{u}_1$ and player P_2 at least $\hat{u}_2$
- Therefore, we can expect that the real payoff of P_1 will be $-\hat{u}_1 + \epsilon_1$ and of P_2 will be $\hat{u}_2 + \epsilon_2$, where $\epsilon_i \geq 0$
- Since the sum of payoffs must be 0, it must be that $-\hat{u}_1 + \epsilon_1 + \hat{u}_2 + \epsilon_2 = 0$, i.e.,

$$\epsilon_1 + \epsilon_2 = \hat{u}_1 - \hat{u}_2 > 0$$

- Therefore, if a game does not have a value, there are infinitely many ways in which the sum of the epsilons can be equal to the difference between the conservative values. This means that:
  - The game is **not solved** by merely computing the conservative values.
  - The players will need to somehow **divide the remaining non-guaranteed** payoff of $\hat{u}_1 - \hat{u}_2$.

Information Advantage in Games Without a Value

Insight 5: Consider a zero-sum game where conservative strategies exist but which does not have a value ($\hat{u}_1 > \hat{u}_2$).

If P_2 **knew that P_1 would use her conservative strategy** (call it s_1^*), then P_2 **could increase his guaranteed payoff** from $\hat{u}_2$ (his conservative strategy guarantees this to him) by

$$\delta \stackrel{\text{def}}{=} \hat{u}_1 - \hat{u}_2 > 0 \qquad \text{to the value of} \qquad \hat{u}_2 + \delta = \hat{u}_1$$

by using his **best response strategy** to s_1^* (instead of using his conservative strategy).

Proof.

Let s_2' be the best response of P_2 to s_1^*. Then by definition

$$\hat{u}_1 = \min_{s_1 \in \Sigma_1} u_1(s_1) = u_1(s_1^*) = \max_{s_2 \in \Sigma_2} f(s_1^*, s_2) = f(s_1^*, s_2').$$

Therefore, the payoff of P_2 when P_1 plays her conservative strategy s_1^* and P_2 plays his best response s_2' to it is equal to $\hat{u}_1$. □

By symmetry the theorem holds when we swap the players!

When Saddle Points Exist, Life is Beautiful

Insight 6: The theory of 2-person zero-sum games is **especially insightful** in the case when a **saddle point exists.**

- Saddle points do exist in **convex-concave games** satisfying the 3 **regularity assumptions** in the “Existence of Saddle Points” theorem (L2, Slide 27)
- The main insight comes from the “Minimax Equality” theorem which says: if a saddle point (s_1^*, s_2^*) exists, then
  - s_1^* is the conservative (minimax) strategy of P_1, guaranteeing him a payoff of $-\hat{u}_1$
  - s_2^* is the conservative (maximin) strategy of P_2, guaranteeing her a payoff of $\hat{u}_2$
  - the **game has a value:** $\hat{u}_1 = \hat{u}_2$
- A saddle point is **the only reasonable solution** of a game in which it exists since
  - P_2 cannot hope to get a payoff x larger than $\hat{u}_2$, since then P_1 would then get $-x < -\hat{u}_2 = \hat{u}_1$, i.e., less than what he can guarantee!
  - P_1 knows that P_2 knows this...
  - The same is true for P_1 by symmetry

Game Zoo