GAME THEORY

László Gulyás
Associate Professor, ELTE, AI Department
✉ lgulyas@inf.elte.hu

Tamás Takács
PhD student, ELTE, AI Department
✉ tamastheactual@inf.elte.hu
tamastheactual.github.io

Game Theory

Lecture 5

Joint Policies, Expected Return, and Minimax

1 Joint Policies

2 Expected Return

3 Minimax Algorithm

Complete Formal Framework

For imperfect information, an extensive form game is:

$$\Gamma = \langle N, H, P, \{I_i\}_{i \in N}, A, \rho, \{u_i\}_{i \in N} \rangle$$

Where:

- $N = \{1, \ldots, n\}$: finite set of players
- H: set of all histories (sequences of actions)
- $Z \subseteq H$: terminal histories (leaves of the game tree, where no further action is taken)
- $P: H \to N \cup \{c\}$: player function (c = chance)
- I_i: collection of information sets for player i
- $A(h)$ or $A(I)$: actions available at history / information set
- ρ: probability distribution for chance moves
- $u_i: Z \to \mathbb{R}$: payoff function for player i ($\mathbb{R}$ = real numbers)

Histories in Detail

A **history** $h = (a_1, \dots, a_k)$ where:

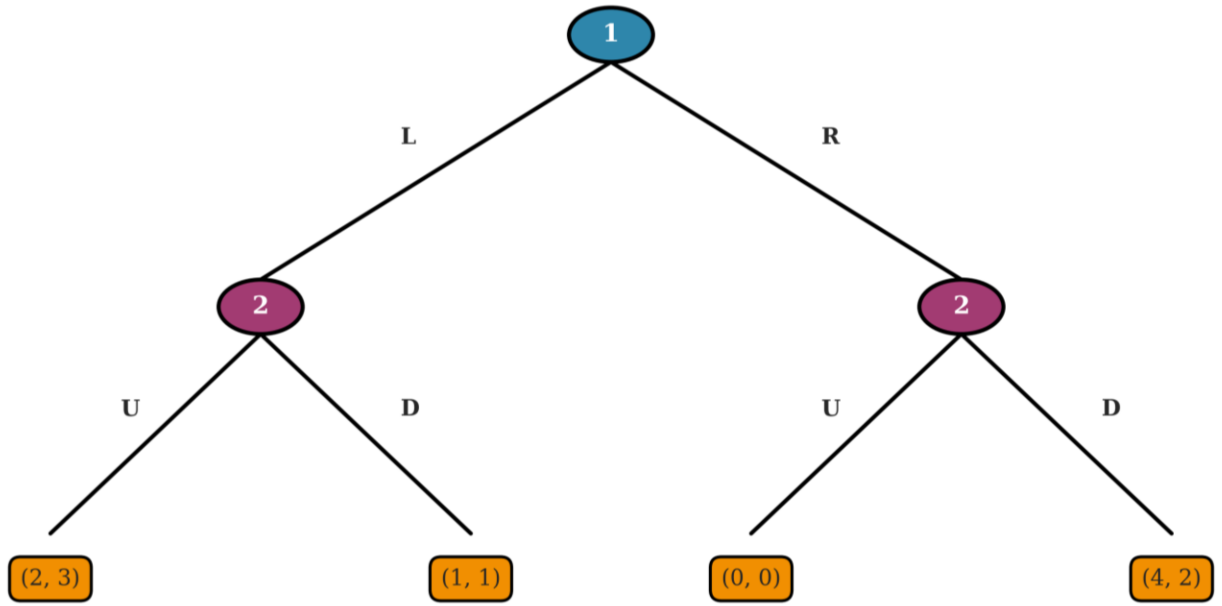
- Empty history $\emptyset$ is the root
- Each $a_i \in A(h_{i-1})$ where $h_{i-1} = (a_1, \dots, a_{i-1})$
- $Z \subseteq H$ is the set of **terminal histories**

Precedence: $h' \sqsubseteq h$ if h' is a prefix of h

Immediate successors: (h, a) denotes history obtained by adding action a to h

Tree property: For any two histories h, h', either $h \sqsubseteq h'$, $h' \sqsubseteq h$, or they are incomparable

Example: Sequential Game Tree



Backward induction solution:

- At left node, Player 2 chooses U (payoff 3 vs 1)
- At right node, Player 2 chooses D (payoff 2 vs 0)
- At root, Player 1 anticipates and chooses R (payoff 4 vs 2)
- **Outcome**: (R, D) with payoffs (4, 2)

Subgames: Rigorous Definition

A **subgame** of Γ starting at $h \in H$ consists of:

- All histories $h' \sqsupseteq h$
- For each information set I:
- If $I \cap \{h' : h' \sqsupseteq h\} \neq \emptyset$
- Then $I \subseteq \{h' : h' \sqsupseteq h\}$

Intuition: Information sets cannot "straddle" the boundary

Consequence: Under perfect information, every h starts a subgame

Restriction: Strategies restricted to subgame form valid strategies for subgame

Stackelberg: Full Analysis

Stackelberg duopoly: two-stage quantity competition. Firm 1 (leader) commits to a quantity first; Firm 2 (follower) observes $q1$ and responds. Solved by backward induction (SPE)

Model:

- Leader chooses $q1 \geq 0$
- Follower observes $q1$, chooses $q2 \geq 0$
- Price: $P(Q) = a - b(q1 + q2)$
- Cost: c per unit

Follower's problem: $\max_{q2} [(a - b(q1+q2))q2 - c \cdot q2]$

FOC (first-order condition, set $\partial\pi/\partial q2 = 0$): $a - bq1 - 2bq2 - c = 0$

Best response: $q2^*(q1) = (a - c - bq1) / 2b$

Pareto Efficiency: Formal Definition

Outcome z is **Pareto dominated** by z' if:

$$u_i(z') \geq u_i(z) \quad \forall i \in N$$

with strict inequality for at least one i

Outcome z is **Pareto efficient** (or Pareto optimal) if no z' Pareto dominates it

Set of Pareto efficient outcomes: Pareto frontier

Computing Pareto Frontier

For finite games:

1. List all possible outcomes $z_1, \dots, z_m$
2. For each outcome z_k:
   - Check if $\exists z_j$ such that $u_i(z_j) \geq u_i(z_k)$ for all i with strict inequality for some i
   - If yes, z_k is dominated; if no, z_k is Pareto efficient
3. Pareto frontier $=$ all non-dominated outcomes

Computational complexity: $O(m^2 \cdot n)$ for m outcomes, n players

Price of Anarchy (PoA)

Definition: Ratio of welfare at worst Nash equilibrium to optimal welfare

For maximization problems:

$$\mathrm{PoA} = \frac{\max_{z \in Z} W(z)}{\min_{z \in \mathrm{NE}} W(z)}$$

For minimization problems (e.g., congestion):

$$\mathrm{PoA} = \frac{\max_{z \in \mathrm{NE}} C(z)}{\min_{z \in Z} C(z)}$$

where W is welfare and C is cost

Game Theory

Lecture 5

Joint Policies, Expected Return, and Minimax

1 Joint Policies

2 Expected Return

3 Minimax Algorithm

Why Sequential Games?

Stage game: Row payoff matrix $A \in R^{m\times n}$, Column payoff $-A$.
Round $\boldsymbol{t = 0, 1, 2, \ldots}$: players choose pure actions $i_t \in [m], j_t \in [n]$.
Instant payoff: $r_t = A_{i_t j_t}$ to Row (and $-r_t$ to Column).

A **joint policy** $\Pi = (\pi_1, \pi_2)$ specifies, for each round t, a (possibly history-dependent) mixed action for each player:

$$\pi_1(\cdot|\, h_t) \in \Delta^m, \quad \pi_2(\cdot|\, h_t) \in \Delta^n, \qquad h_t = (i_0, j_0, \ldots, i_{t-1}, j_{t-1}).$$

- **Stationary (memoryless) policy:** $\pi_1(\cdot|\, h_t) \equiv p \in \Delta^m, \pi_2(\cdot|\, h_t) \equiv q \in \Delta^n$ for all t.
- Unless stated otherwise, draws are **independent across players and across time** under a stationary policy.

Notation: e_i denotes the i-th standard basis vector.
$\mathbf{1}$ denotes an all-ones vector.
$\Delta^k = \{x \in R^k : x \geq 0,\ \mathbf{1}^\top \mathbf{x} = 1\}$ is the probability simplex.

Discounted return under stationary independent mixing

Fix $\gamma \in (0,1)$. Under **stationary independent mixing** (p, q) each round,

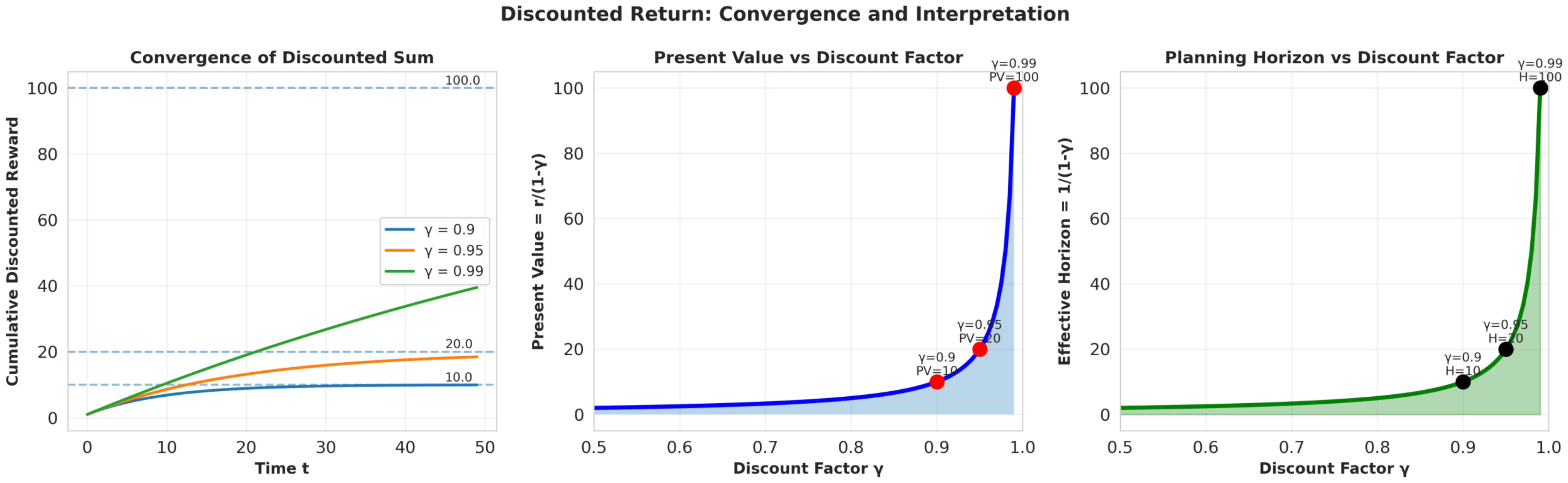
$$E[r_t] = \sum_{i,j} p_i, A_{ij}, q_j = p^\top A q \quad \text{for all } t,$$

and, by linearity of expectation with i.i.d. draws,

$$J_\gamma(p,q) := E\left[\sum_{t=0}^{\infty} \gamma^t r_t\right] = \sum_{t=0}^{\infty} \gamma^t E[r_t] = \frac{p^\top A q}{1-\gamma}$$

- The independence across time is sufficient; no ergodic or martingale machinery is required here.
- If you add a constant c to all entries of A, then J_γ shifts by $c/(1-\gamma)$; optimal mixes are unchanged.

Discounted return under stationary independent mixing (Cont.)



Game Theory

Lecture 5

Joint Policies, Expected Return, and Minimax

1 Joint Policies

2 Expected Return

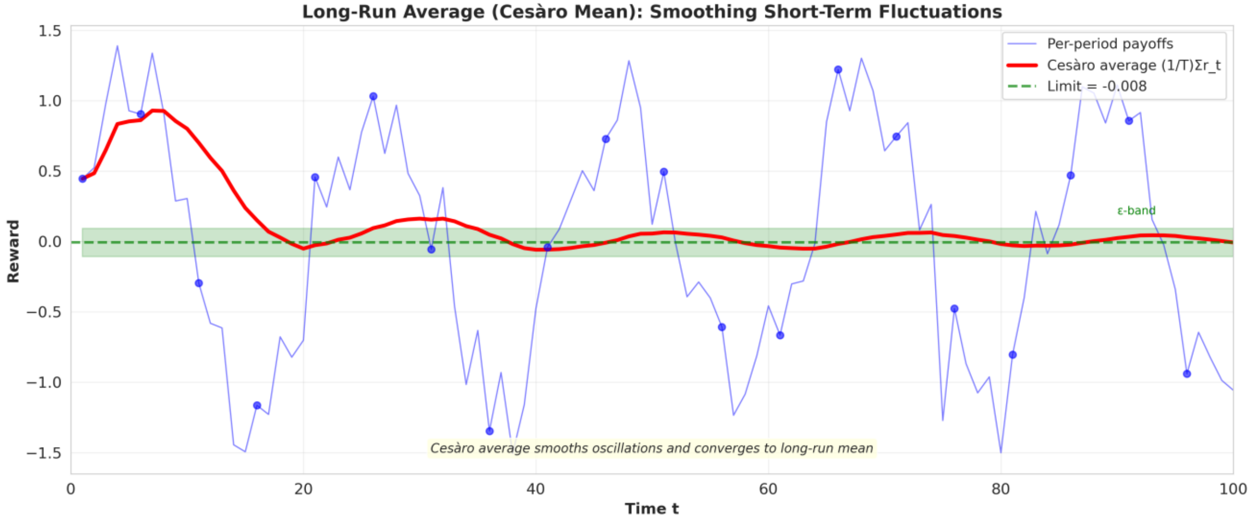
3 Minimax Algorithm

Long-run average reward (Cesàro average)

Under stationary independent mixing (p, q),

$$\bar{J}(p,q) := \lim_{T\to\infty} \frac{1}{T} E\left[\sum_{t=0}^{T-1} r_t\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} E[r_t] = p^\top A q.$$

Interpretation: The one-shot value $p^\top A q$ is the **per-period** expected return of the stationary joint policy; discounting just rescales it by $(1-\gamma)^{-1}$.



Long-run average reward (Cesàro average) (Cont.)

*If players **correlate across time** (e.g., contingent punishments), per-period **expectation** under stationary* (p, q) *is still* $p^\top Aq$*, but non-stationary history-dependent strategies can implement different **paths**. In **matrix zero-sum games**, however, these paths cannot raise the secure average payoff above the minimax value (see Minimax section).*

Security levels: maximin vs minimax

Row's **maximin** (security guarantee):

$$v^- := \max_{p\in\Delta^m} \min_{q\in\Delta^n} p^\top A q$$

Column's **minimax** (Row's worst-case given Column's choice):

$$v^+ := \min_{q\in\Delta^n} \max_{p\in\Delta^m} p^\top A q$$

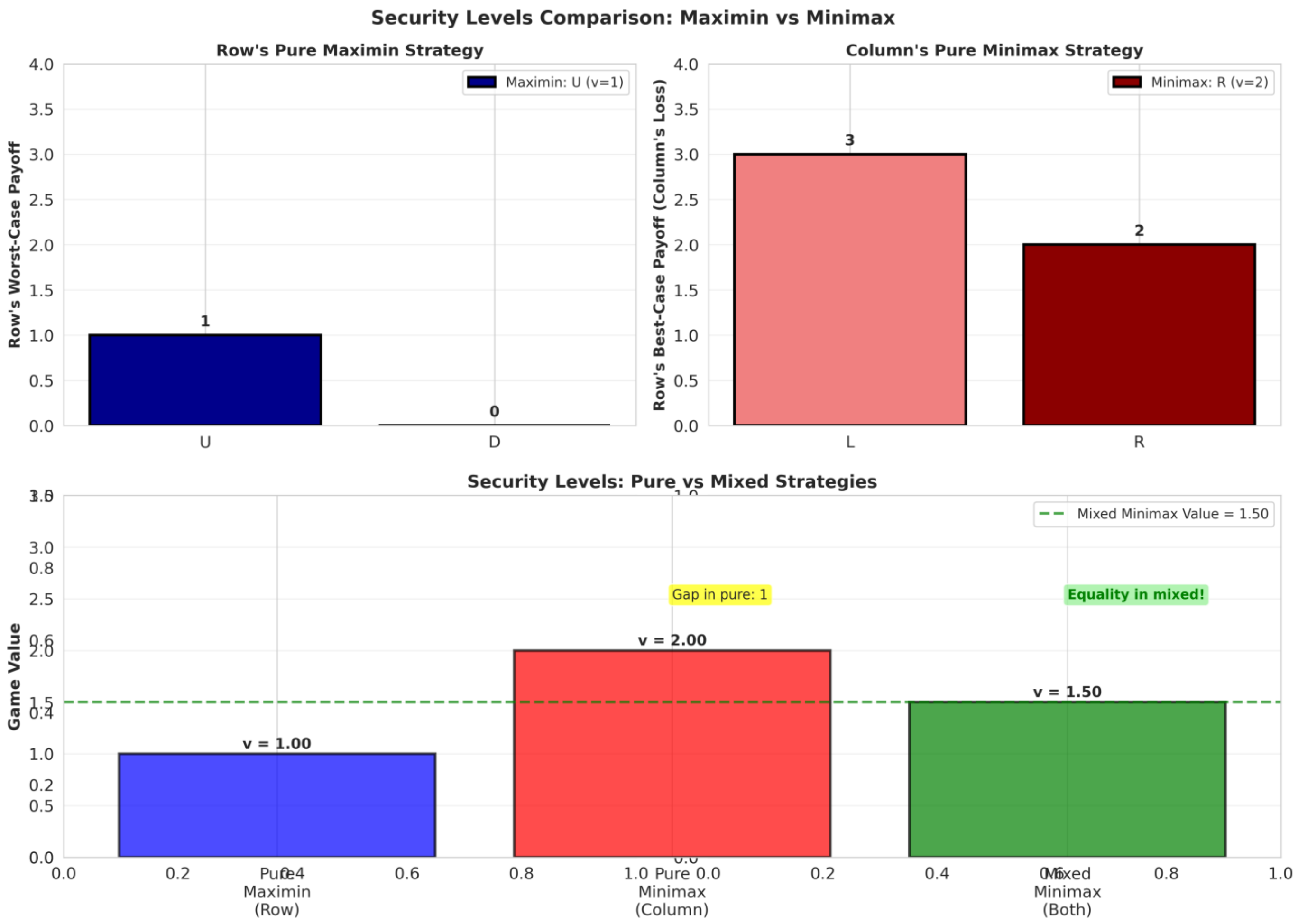
Weak minimax inequality
For any bilinear form over compact convex sets

$$v^- \le v^+$$

- v^-: Row can guarantee at least v^- regardless of Column ($\Rightarrow$ **security level)**
- v^+: Column can hold Row down to at most v^+ regardless of Row

In finite zero-sum matrix games, the **Minimax Theorem** (next section) states $v^- = v^+ =: v$, the **value** of the game.

Security levels: maximin vs minimax (Cont.)



Side notes & edge cases

- Finite horizon T: Under stationary independent (p, q),

$$E\left[\sum_{t=0}^{T-1} r_t\right] = Tp^\top Aq.$$

- **Time-varying stationary but independent** (piecewise constant p_t, q_t): the per-period mean is $\frac{1}{T}\sum_{t=0}^{T-1} p_t^\top A\, q_t$.
- **History-dependent strategies (threats/punishments):** In general-sum, these matter (Folk theorems). In **zero-sum matrix games**, they cannot beat the minimax value v in expected average payoff.
- **Nonzero-sum:** Expected returns are bilinear in (p, q) per player; many results above carry but **security equality** (minimax) does not.

Micro-check (2×2 stationary)

Let

$$A = \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix}, \quad p = (p, 1-p),\ q = (q, 1-q)$$

Then

$$p^{\top} A q = p[2q - 1(1-q)] + (1-p)[-3q + 4(1-q)] = (10p - 7)q + (-5p + 4)$$

- For fixed q, Row's best response is $\arg m\, ax_p$ of a linear function in p.
- For fixed p, Column's best response is $\arg m\, in_q$ of the same bilinear expression.

Indifference equalization yields $p^* = 0.7,\ q^* = 0.5$ (derived later), hence $\bar{J} = p^{*\top} A q^* = 0.5$ and $J_\gamma = 0.5/(1-\gamma)$.

Questions

1. Show $J_\gamma(p,q) = \frac{p^\top A q}{1-\gamma}$ using only independence and linearity of expectation.
2. Prove $v^- \le v^+$ for any compact convex P, Q and continuous bilinear payoff.
3. Give a general-sum 2×2 where non-stationary correlation across time changes the *distribution of outcomes* relative to stationary play, even though the stage expectation formula holds under stationarity.

From General Repeated Games to Strictly Competitive Games

So far in this lecture:

- Joint policies and expected returns (general framework)
- Discounted and average returns apply to **any** repeated game
- Security levels exist but may not equate (maximin $\neq$ minimax in general-sum)

Now: Focus on **zero-sum matrix games** where $u_1 + u_2 = 0$ (strict competition)

Key questions for zero-sum games:

1. What is the **value** of the game?
2. How do we compute **optimal mixed strategies**?
3. What role does **LP duality** play?
4. Can players guarantee the value against any opponent strategy?

Why zero-sum is special: Minimax Theorem guarantees security equality $v^- = v^+$, enabling LP formulation.

Lecture 5

Joint Policies, Expected Return, and Minimax

1 Joint Policies

2 Expected Return

3 Minimax Algorithm

Minimax Theorem

Theorem (von Neumann). For any finite two-player zero-sum matrix game with Row payoff $A \in R^{m \times n}$,

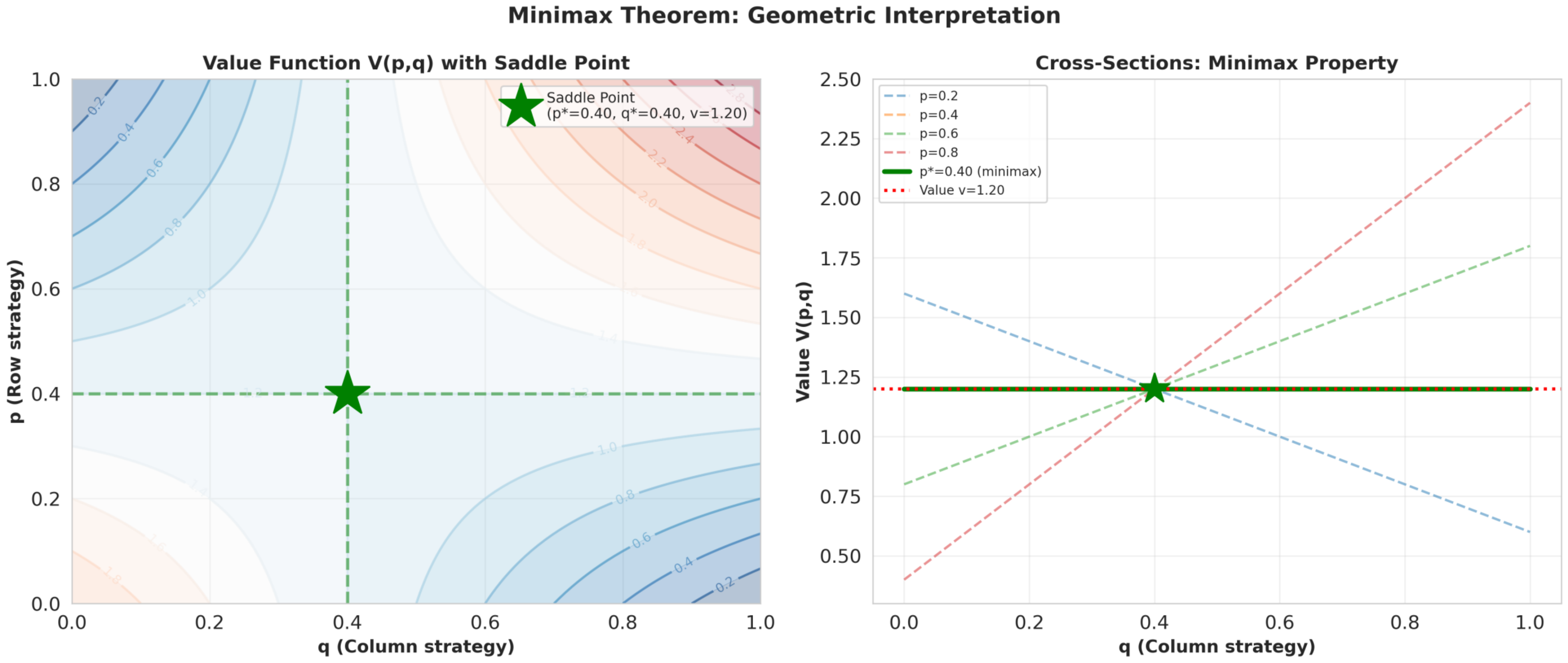
$$\max_{p \in \Delta^m} \min_{q \in \Delta^n} p^\top A\, q = \min_{q \in \Delta^n} \max_{p \in \Delta^m} p^\top A\, q = v$$

There exist optimal mixes $p^* \in \Delta^m,\ q^* \in \Delta^n$ such that

$$p^{*\top} A q \geq v\ \ \forall q, \qquad p^\top A q^* \leq v\ \ \forall p$$

Row can **guarantee** at least v (security), Column can **hold** Row down to at most v; equality pins down the **value** and optimal mixed strategies.

Intuition (finite matrix games)



Saddle-Point View (equivalent characterization)

$$(p^*, q^*) \text{ is minimax } \Leftrightarrow p^{*\top} A q \geq p^{*\top} A q^* \geq p^\top A q^* \quad \forall p, q$$

At a saddle point, neither player can profitably deviate; the common value is $v = p^{*\top} A q^*$.

Weak Minimax Inequality

For any bilinear payoff over compact convex sets,

$$\max_{p} \min_{q} p^\top A\, q \;\leq\; \min_{q} \max_{p} p^\top A\, q$$

Proof sketch: For all $p, q, \min_{q'} p^\top A\, q' \leq p^\top A q \leq \max_{p'} p'^\top A\, q$. Take max_p on the left, min_q on the right.

LP Formulation (Row / "primal")

$$max_{p,v}\, v$$
$$s.t. A^{\top}p \ \geq \ v1,$$
$$1^{\top}p \ = \ 1, \quad p \geq 0.$$

Meaning: Choose p so that every column yields **at least** v.

LP Formulation (Row / “primal”)

$$max_{\mathrm{q},v}\, v$$
$$s.t.\, \mathrm{Aq} \;\geq\; v\mathbf{1},$$
$$\mathbf{1}^T\mathrm{q} \;=\; 1, \quad \mathrm{q} \geq 0.$$

Meaning: Choose q so that every **row** yields **at most** v.

Consequence: LP strong duality ⇒ optimal values match ⇒ minimax equality.

Complementary Slackness (support equalization)

At an optimal pair (p^*, q^*, v):

If $p_i^* > 0$, then the i-th row payoff equals the value: $(Aq^*)_i = v$
If $q_j^* > 0$, then the j-th column payoff equals the value: $(A^\top p^*)_j = v$

Takeaway: Supported pure actions are **equalized at value v**; excluded actions satisfy the corresponding inequality strictly

Computing by Indifference (support method)

Given supports $I \subseteq [m], J \subseteq [n]$:

Row equalization on $\boldsymbol{I}$: $(Aq)_i = v$ for all $i \in I$
Column equalization on $\boldsymbol{J}$: $(A^\top p)_j = v$ for all $j \in J$
Normalization: $\sum_{i \in I} p_i = 1,\ p_i \geq 0;\ \ \sum_{j \in J} q_j = 1,\ q_j \geq 0$
Verify inequalities: $(Aq)_i \leq f\ or\ i \notin I; (A^\top p)_j \geq v$ for $j \notin J$

If feasible, (p, q, v) is a solution. Otherwise try different supports

Example 1 (2×2 by indifference)

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

Let $p = \Pr[U]$, $q = \Pr[L]$.

Row indiff. (equalize Column's payoffs):

$$2q + (-3)(1-q) = -1 \cdot q + 4(1-q) \;\Rightarrow\; 3q - 1 = -7q + 4 \;\Rightarrow\; q^* = 0.5$$

Column indiff. (equalize Row's payoffs):

$$2p + (-1)(1-p) = -3p + 4(1-p) \;\Rightarrow\; 3p - 1 = -7p + 4 \;\Rightarrow\; p^* = 0.7$$

Value:

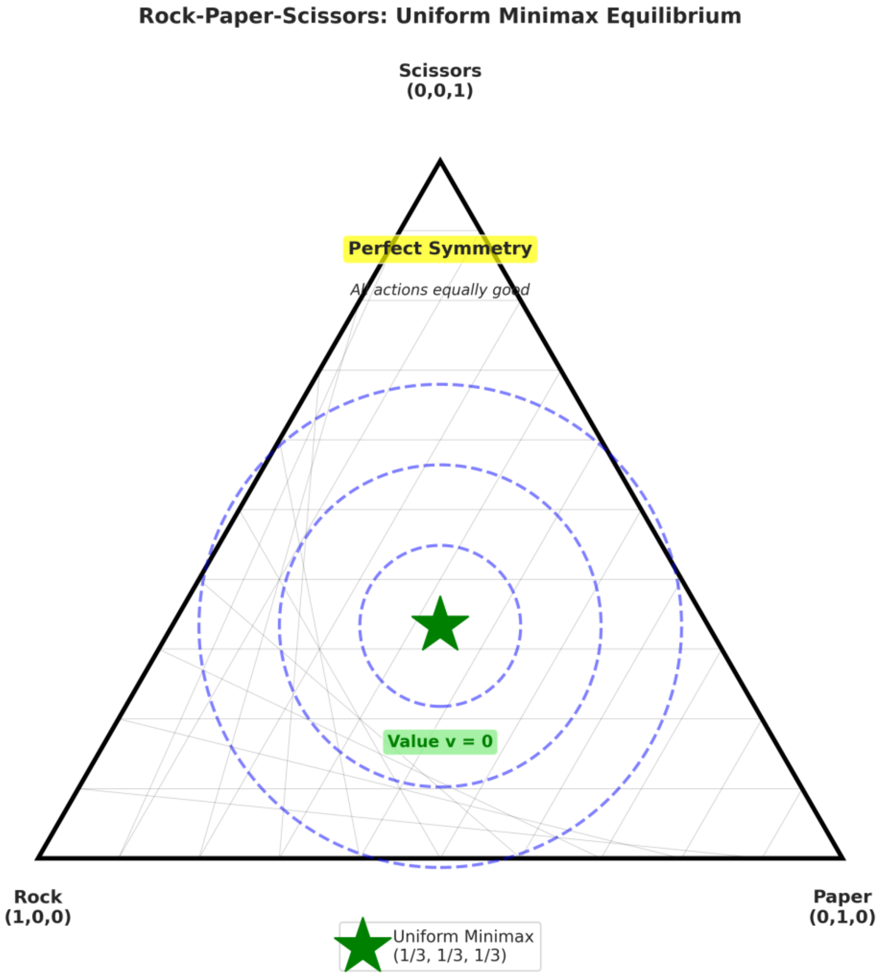
$$v = p^{*\top} A q^* = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 0.5 & 0.5 \end{pmatrix} = 0.05.$$

Example 2 (3×3 RPS)

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

By symmetry, $p^* = q^* = (1/3,1/3,1/3), v = 0$. **Check:** $Aq^* = 0 \cdot \mathbf{1}, A^\top p^* = 0 \cdot \mathbf{1}$.

Example 2 (3×3 RPS) (Cont.)



Example 3 (weighted RPS, full support system)

$$A = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Solve

$$Aq = v,\mathbf{1}, \quad A^\top p = v,\mathbf{1}, \quad \mathbf{1}^\top \mathbf{p} = \mathbf{1}^\top \mathbf{q} = 1$$

Check $p, q \geq 0$; the solution yields full-support mixes and v.

Example 4 (dominance pruning)

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Row 3 is dominated by a mixture of Rows $1–2$. Remove it, solve 2×2 by indifference; verify Row 3 remains unprofitable at the solution.

Geometry & Invariances

- **Hyperplanes**: Equal-payoff sets $(Aq)_i = (Aq)_{i'}$ and $(A^\top p)_j = (A^\top p)_{j'}$ are linear (hyperplanes in the simplexes)
- **Polytopes**: Best-response regions are intersections of half-spaces ⇒ polytopes; equilibria are at polytope intersections
- **Affine transforms**: $A \mapsto \alpha A + c\mathbf{1}\mathbf{1}^\top$: mixes unchanged; value scales by α and shifts by c

Computational Tools

```python
import numpy as np, nashpy as nash

A = np.array([[2,-1],[-3,4]])
G = nash.Game(A)  # zero-sum shorthand
list(G.vertex_enumeration())  # returns (p*, q*)
```

LP template for Row

Variables: probabilities on rows $p \in R^m$ and value $v \in R$
Constraints: $A^\top p \geq v\mathbf{1}, \mathbf{1}^\top \mathbf{p} = 1, p \geq 0$
Objective: maximize v

The **dual** is Column's problem automatically (minimize v with $Aq \leq v\mathbf{1}, \mathbf{1}^\top \mathbf{q} = 1, q \geq 0$)

Use any LP solver that handles linear inequalities (cvxopt, PuLP, SciPy 'linprog', CVXPy, …)

Complementary Slackness

At $(p^*, q^*, v) = (0.7, 0.5, 0.5)$ for

$$A = \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix},$$

- **Row** supported actions U, D both achieve value exactly 0.5 against q^*: $(Aq^*)_U = (Aq^*)_D = v$
- **Column** supported actions L, R yield value exactly 0.5 against p^*: $(A^\top p^*)_L = (A^\top p^*)_R = v$
- In a 2×2 there are no excluded pure actions; feasibility is immediate

Supported actions are equalized at v; any excluded actions (in larger games) must satisfy strict inequality

Invariances: value shifts & scalings

Shift: $A \mapsto A + c\,\mathbf{1}\mathbf{1}^\top$ mixes p^*, q^* unchanged; value shifts by c

Scale: $A \mapsto \alpha A$ with $\alpha > 0$ mixes unchanged; value scales by α

Use these to simplify arithmetic (e.g., make entries nonnegative for LP stability)

ε –security

For any $\varepsilon > 0$, there exists p_ε s.t.

$$\min_q p_\varepsilon^\top A\, q \;\geq\; v - \varepsilon,$$

and q_ε s.t.

$$\max_p p^\top A\, q_\varepsilon \;\leq\; v + \varepsilon$$

Practice: When you compute $(\hat{p}, \hat{q})$ numerically, report the **deviation incentives**

$$\varepsilon_{\text{row}} = \max_i (A\hat{q})_i - \widehat{p^\top} A\hat{q}, \qquad \varepsilon_{\text{col}} = \widehat{p^\top} A\hat{q} - \min_j \left(\widehat{p^\top} A\right)_j$$

and use $\max(\varepsilon_{\text{row}}, \varepsilon_{\text{col}})$ as a conservative error bound.

Relation to Nash Equilibrium (zero-sum)

In two-player zero-sum games, **minimax strategies** (p^*, q^*) are **exactly** Nash equilibria, and the **equilibrium payoff** equals the **value** v. Conversely, any NE mixed profile is minimax.

Sion’s Minimax Theorem

Let $X \subset R^m, Y \subset R^n$ be nonempty compact convex sets

If $f: X \times Y \to R$ is **quasi-convex and lower semicontinuous** in x for each y, and **quasi-concave and upper semicontinuous** in y for each x, then

$$\min_{y \in Y} \max_{x \in X} f(x, y) = \max_{x \in X} \min_{y \in Y} f(x, y)$$

This generalizes the matrix-game minimax equality

Proof Sketch A: LP strong duality

1. Write Row's problem as an LP; Column's is the **dual**
2. Feasibility & boundedness ⇒ **strong duality**: optimal values match
3. Optimal primal/dual solutions yield (p^*, q^*, v)
4. **Complementary slackness** explains equalization of supported actions

Proof Sketch B: Fixed-point route

1. Mixed strategy spaces are simplexes (compact, convex).
2. Best-response correspondences are nonempty, convex-valued, upper-hemicontinuous (Berge).
3. Existence of NE (Kakutani) in zero-sum ⇒ value-attaining equilibrium; Row's secured payoff equals Column's held-down payoff ⇒ minimax equality.

JP and Minimax Repeated Play

- In **static** zero-sum matrix games, you computed $(p^*, q^*, v)$by **minimax**
- In **repeated** play (finite/infinite horizon), if each round is the *same* stage game and players **mix independently** each round, then:
  - Per-period payoff is $p^\top Aq$
  - Discounted return $J_\gamma(p, q) = \frac{p^\top Aq}{1-\gamma}$
  - Stationary minimax p^*, q^* **secure** value v each round → the repeated game's value is v (per period)

Joint policies in repeated matrix games

- Stage game payoffs: Row $A \in R^{m \times n}$, Column $-A$ (zero-sum)
- **Joint policy** (possibly history-dependent): mapping from histories $\mathcal{H}_t$ to mixed actions
- **Stationary independent** policy: fixed $p \in \Delta^m, q \in \Delta^n$ each round

Discounted return (stationary, independent)

$$J_\gamma(p, q) = E\left[\sum_{t=0}^{\infty} \gamma^t a_{i_t j_t}\right] = \frac{p^\top A q}{1 - \gamma}$$

Average reward (Cesàro)

$$\lim_{T \to \infty} \frac{1}{T} E\left[\sum_{t=0}^{T-1} a_{i_t j_t}\right] = p^\top A q$$

Security levels in repeated play

Let

$$v^- = \max_p \min_q p^\top A\, q, \quad v^+ = \min_q \max_p p^\top A\, q, \quad (v^- \leq v^+)$$

By the **minimax theorem** $v^- = v^+ = v$

- Against any opponent policy (even history-dependent), Row can play p^* i.i.d. each round and **guarantee** at least v per period
- Symmetrically Column can hold Row to **at most** v
- Thus, the **repeated zero-sum game** (with the same stage game) has **per-period value** v; using history cannot beat v in expectation

Example 5: uniform security in a centered symmetric game

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \frac{2}{3}\mathbf{1}\mathbf{1}^{\top}$$

- Each row/column sum is 0 $\rightarrow$ **uniform** $p = q = (1/3,1/3,1/3)$
- Value $v = 0$. Verify $Aq = \mathbf{0}$ and $A^{\top}p = \mathbf{0}$

Example 6: dominance pruning first

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

Row 3 is dominated by a mixture of Rows $1\text{–}2$. Remove, solve the 2×2, then **reinsert** Row 3 to confirm it remains unprofitable at $(p^\wedge *, q^\wedge *, v)$

Geometry refresher

Equal-payoff hyperplanes: $e_i^\top Aq = e_{i'}^\top Aq$ are linear constraints in q

Best response regions: intersections of halfspaces (polyhedral)

Equilibria: at intersections where both players are indifferent on **their supported actions** and inequalities hold for excluded ones

Algorithms in practice

- 2×2 / **some** 3×3: support guessing + indifference + inequality checks
- **Small/medium**: vertex enumeration of BR polytopes (e.g., NashPy for zero-sum)
- **Larger**: LP (sparse) or first-order primal-dual methods

ε-security & Solver tolerance

Given numerical $(\hat{p}, \hat{q})$, define

$$\varepsilon_{\text{row}} = \max_i (A\hat{q})_i - \widehat{p^\top A\hat{q}}, \quad \varepsilon_{\text{col}} = \widehat{p^\top A\hat{q}} - \min_j \left(\widehat{p^\top A}\right)_j$$

Report $\max\left(\varepsilon_{\text{row}}, \varepsilon_{\text{col}}\right)$ as a conservative suboptimality bound

Exercise 1

For

$$A = \begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix},$$

1. Compute (p^*, q^*, v) by support equalization
2. Remove any dominated actions if found and recompute
3. Verify equalization/inequalities

Exercise 2

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

1. Enumerate size-2 supports, keep feasible ones. If none feasible, try full support and solve the linear system $Aq = v\mathbf{1}$, $A^\top p = v\mathbf{1}$, $\mathbf{1}^\top \mathbf{p} = \mathbf{1}^\top \mathbf{q} = 1$

Exercise 3

1. Add i.i.d. noise $\xi_{ij} \sim \text{Unif}[-0.1,0.1]$ to a 3×3 with known (p^*, q^*, v)
2. Recompute $(\tilde{p}, \tilde{q}, \tilde{v})$
3. Summarize how supports and v change

Exercise 4

Let

$$A(\theta) = \begin{pmatrix} 2 & -1 \\ -3 & 1+\theta \end{pmatrix}, \quad \theta \in [-1,2]$$

Derive $\left(p^*(\theta), q^*(\theta), v(\theta)\right)$ **piecewise** in θ; identify breakpoints where supports change

Exercise 5

1. Prove rigorously that stationary (p^*, q^*) yields discounted return $v/(1-\gamma)$.

2. Argue any history-dependent deviation cannot improve the **per-period** value above v.

Summary

- In zero-sum matrix games, **minimax** $=$ **Nash** and yields value v.
- Repeated play with stationary independent minimax achieves **per-period** v; history dependence cannot beat it in expectation.
- **Computation**: supports + indifference for small games; LP/vertex enumeration otherwise.
- **Diagnostics**: complementary slackness and ε-security quantify solution quality.

Course Textbooks

- Bonanno, G. (2024). ***Game Theory (3rd ed.)***. University of California, Davis. Received from: GT Book
- Axelrod, R. (1984). ***The Evolution of Cooperation***. Basic Books. Received from: Axelrod Article
- Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. V. (2007). ***Algorithmic Game Theory***. Cambridge University Press. Received from: AGT Book
- Myerson, R. B. (1991). ***Game Theory: Analysis of Conflict.*** Harvard University Press. Received from: GT Book 2F.
- Christianos et al., ***Multi-Agent Reinforcement Learning: Foundations and Modern Approaches***, 2023. Received from: MARL Book.pdf
- Shoham, Y., & Leyton-Brown, K. (2008). ***Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations***. Cambridge University Press. Received from: MARL Book.pdf
- `***nashpy***` documentation (readthedocs). Link: NashPy Docs

That’s All for Today!