Learning Nash Equilibria

Dr. Jun Wang
Computer Science, UCL

Content

- Lecture 1: Multiagent AI and basic game theory
- Lecture 2: Potential games, and extensive form and repeated games
- Lecture 3: Solving ("Learning") Nash Equilibria
- Lecture 4: Auction theory and mechanism design
- Lecture 5: Learning and deep neural networks
- Lecture 6: Single-agent Learning (1)
- Lecture 7: Single-agent Learning (2)
- Lecture 8: Multi-agent Learning (1)
- Lecture 9: Multi-agent Learning (2)
- Lecture 10: Multi-agent Learning (3)

Subject to change

Exercise: The Stackelberg model of Duopoly

- The Stackelberg model of Duopoly (1934):
  - one player, called the dominant player or leader, moves first and the outcome of that player's choice is made known to the other player (follower) before the other player's choice is made
  - e.g., General Motors, at times big enough in U.S. history to play such a dominant role in the automobile industry

Exercise: The Stackelberg model of Duopoly

- Firm 1 chooses an amount to produce, q_1, at a cost *c* per unit.
- This amount is then told to Firm 2 which then chooses an amount q_2 to produce also at a cost of *c* per unit
- Then the price *P* per unit is determined by

$$P(Q) = \begin{cases} a - Q & \text{if } 0 \leq Q \leq a \\ 0 & \text{if } Q > a \end{cases} = (a - Q)^+$$

where $Q = q_1 + q_2$ and *a* is a constant

- the players receive the payoff

$$u_1(q_1, q_2) = q_1 P(q_1 + q_2) - cq_1 = q_1(a - q_1 - q_2)^+ - cq_1$$
$$u_2(q_1, q_2) = q_2 P(q_1 + q_2) - cq_2 = q_2(a - q_1 - q_2)^+ - cq_2$$

where cost per unit $c<a$.

Answer: The Stackelberg model of Duopoly

- Observations:
  - Firm 1's pure strategy space is X = [0, ∞).
  - Firm 2's pure strategy space, Y , is now a set of functions mapping q_1 into q_2
- Can be solved by backward induction:
  - Since Firm 2 moves last, we first find the optimal q_2 as a function of q_1.
  - That is, we maximise Firm 2's payoff with respect to q_2 (conditioned on q_1)

$$\frac{\partial}{\partial q_2} u_2(q_1, q_2) = a - q_1 - 2q_2 - c = 0$$

  - This gives us Firm 2's strategy as

$$q_2(q_1) = (a - q_1 - c)/2.$$

- Since Firm 1 now knows that Firm 2 will choose this best response, Firm 1 now wishes to choose q_1 to maximize their payoff:

$$\begin{aligned} u_1(q_1, q_2(q_1)) &= q_1(a - q_1 - (a - q_1 - c)/2) - cq_1 \\ &= -\frac{1}{2}q_1^2 + \frac{a-c}{2}q_1. \end{aligned}$$

- which gives $q_1 = q_1^* = (a-c)/2.$ and then $q_2^* = q_2(q_1^*) = (a-c)/4.$

Content

- Our discussion about Game Theory so far largely ignored issues of computation
  - How hard is it to compute the Nash equilibria of a game?
- The answer depends on the class of games being considered.
  - Compute the Nash equilibria of simple games (two-player two-action game)
  - Linear programing solution for simple two-player, zero-sum normal-form games
  - The Lemke-Howson algorithm for general-sum normal-form games

Nash's Theorem

Theorem (Nash, 1951): Every finite game (finite number of players, finite number of pure strategies) has at least one mixed-strategy Nash equilibrium.

Nash, John (1951) "Non-Cooperative Games" *The Annals of Mathematics* 54(2):286-295.

(John Nash did not call them "Nash equilibria", that name came later.)

He shared the 1994 Nobel Memorial Prize in Economic Sciences with game theorists Reinhard Selten and John Harsanyi for his work on Nash equilibria.

He suffered from schizophrenia in the 1950s and 1960s, as depicted in the 1998 film, "A Beautiful Mind". He nevertheless recovered enough to return to academia and continue his research.

Zero-Sum Games

Minimax Theorem (John von Neumann, 1928): For every two-person, zero-sum game with finitely many pure strategies, there exists a mixed strategy for each player and a value V such that:

- Given player 2's strategy, the best possible payoff for player 1 is V
- Given player 1's strategy, the best possible payoff for player 2 is –V.

The existence of strategies part is a special case of Nash's theorem, and a precursor to it.

This basically says that player 1 can guarantee himself a payoff of at least V, and player 2 can guarantee himself a payoff of at least –V. If both players play optimally, that's exactly what they will get.

It's called "minimax" because the players get this value by pursuing a strategy that tries to minimize the maximum payoff of the other player. We'll come back to this.

Definition: The value V is called the ***value (reward or payoff)*** of the game.

Eg: The value of Rock-paper-scissors is 0; the best that P1 can hope to achieve, assuming P2 plays optimally (1/3 probability of each action), is a payoff of 0.

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2	
		col 1	col 2
Player 1	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

This game has no pure-strategy Nash equilibria.

By Nash's theorem, it must have a mixed-strategy Nash equilibrium.

How can we find it?

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
Player 1	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Let’s start by making some definitions.

Let x_1 be the probability that player 1 (row player) plays action row 1, in the Nash equilibrium. So with probability $x_2=1-x_1$, player 1 will play action row 2

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
Player 1	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Let's start by making some definitions.

Likewise, let y_1 be the probability that player 2 (col player) plays action col 1, in the Nash equilibrium. So with probability $y_2 = 1 - y_1$, player 2 will play action col 2

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Next, let's write down what we know about the outcomes, in terms of x_1 and y_1.

In equilibrium, player 1's expected payoff is:

$x_1 y_1(-2) +$
$x_1 (1- y_1)(+3) +$
$(1- x_1)y_1 (+3) +$
$(1-x_1)(1-y_1)(-4)$

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Next, let's write down what we know about the outcomes, in terms of x_1 and y_1.

In equilibrium, player 2's expected payoff is:

$x_1 \; y_1 (+2) +$

$x_1 (1- y_1)(-3) +$

$(1- x_1) y_1 (-3) +$

$(1- x_1)(1- y_1)(+4)$

Computing Nash Equilibria: 2-person, Zero-Sum Games

$-2y_1 + 3y_2$

		Player 2	
		col 1	col 2
Player 1	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

$+3y_1 - 4y_2$

A 2-person, zero-sum game

Observation:

If player 2 selects y_1 so that player 1 gets a higher utility by playing row 1 instead of row 2, then player 1 will *always* select row 1.

But that can't be an equilibrium!

(Why not?)

Computing Nash Equilibria: 2-person, Zero-Sum Games

$+2x_1 - 3x_2$

$-3x_1 + 4x_2$

Player 2

Player 1	col 1	col 2
row 1	-2, +2	+3, -3
row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Observation:

Likewise, if player 1 selects x_1 so that player 2 gets a higher utility by playing col 2 instead of col 1, then player 2 will *always* select col 2

But that can't be an equilibrium, either!

Computing Nash Equilibria: 2-person, Zero-Sum Games

$+2x_1 -3\ x_2$

$-3x_1 +4\ x_2$

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
Player 1	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Observation:

So, the only possible equilibrium has player 1 selecting x_1 so that player 2's payoff for selecting col 1 ***equals*** player 2's payoff for selecting col 2.

Computing Nash Equilibria: 2-person, Zero-Sum Games

$+2x_1 -3\, x_2$

$-3x_1 +4\, x_2$

Player 2

Player 1	col 1	col 2
row 1	-2, +2	+3, -3
row 2	+3, -3	-4, +4

A 2-person, zero-sum game

In algebra:

Player 2's payoff when player 1 plays row 1 with probability x_1, and player 2 always plays col 1:

$$x_1(+2) + (1- x_1)(-3)$$

Player 2's payoff when player 1 plays row 1 with probability x_1, and player 2 always plays col 2:

$$x_1(-3) + (1- x_1)(+4)$$

Computing Nash Equilibria: 2-person, Zero-Sum Games

$+2x_1 - 3x_2$ $\quad$ $-3x_1 + 4x_2$

		Player 2	
		col 1	col 2
Player 1	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

Our observation says these should be equal:

$x_1 (+2) + (1 - x_1)(-3)$

$= x_1 (-3) + (1 - x_1)(+4)$

$\Rightarrow$

$+2x_1 - 3 + 3x_1 =$

$-3x_1 + 4 - 4x_1$

$\Rightarrow 7 = 12x_1$

$\Rightarrow x_1 = 7/12$

Computing Nash Equilibria: 2-person, Zero-Sum Games

$-2y_1 + 3\,y_2$

Player 1 \ Player 2	col 1	col 2
row 1	-2, +2	+3, -3
row 2	+3, -3	-4, +4

$+3y_1 - 4\;y_2$

A 2-person, zero-sum game

We could have done this for either player; here it is from player 1's perspective (its value):

$y_1(-2) + (1 - y_1)(+3)$
$= y_1(+3) + (1 - y_1)(-4)$

=>

$-2\,y_1 + 3 - 3\,y_1 =$
$+3\,y_1 - 4 + 4\,y_1$

=>

$12\,y_1 = 7$

=>

$y_1 = 7/12$

Computing Nash Equilibria: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
Player 1	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

So now we know a mixed-strategy Nash equilibrium:

$P_{player1}$ (row 1) = 7/12
$P_{player1}$(row 2) = 5/12

$P_{player2}$(col 1) = 7/12
$P_{player2}$(col 2) = 5/12

Quiz: 2-person, Zero-Sum Games

		Player 2 col 1	Player 2 col 2
Player 1	row 1	-2, +2	+3, -3
Player 1	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

What is the *value* of this game for player 1?

(Remember, the *value* of the game is the expected payoff for the player in equilibrium.)

Likewise, what is the *value* of the game for player 2?

Answer: 2-person, Zero-Sum Games

Player 2

Player 1	col 1	col 2
row 1	-2, +2	+3, -3
row 2	+3, -3	-4, +4

A 2-person, zero-sum game

You can get the *value* for Even three ways:

Recall: In equilibrium, Player 2's expected payoff is:

$x_1 \; y_1 (+2) +$
$x_1 (1- y_1)(-3) +$
$(1- x_1) \; y_1 (-3) +$
$(1- x_1)(1- y_1)(+4)$

or, $x_1 (+2) + (1- x_1)(-3)$
or, $x_1 (-3) + (1- x_1)(+4)$

These all equal: -1/12

Answer: 2-person, Zero-Sum Games

<u>Player 2</u>

<u>Player 1</u>		col 1	col 2
	row 1	-2, +2	+3, -3
	row 2	+3, -3	-4, +4

A 2-person, zero-sum game

You can get the *value* for player 1 the same three ways, or you can just say that this is a zero-sum game, so the value for player 1 must be opposite the value for player 2:

+1/12

In other words, it's better to be the player 1 than the player 2, since player 1 will win, on average.

Quiz: Computing an Equilibrium for Zero-Sum Games

		Player 1: col 1	Player 1: col 2
Player 2	row 1	+5, -5	+2, -2
Player 2	row 2	+3, -3	+6, -6

In equilibrium,

1. What is the probability that P_1 plays row 1?
2. What is the probability that P_2 plays col 1?
3. What is the value of the game for P1?

Answer: Computing an Equilibrium for Zero-Sum Games

		Player 1: X	Player 1: Y
Player 2	X	+5, -5	+2, -2
Player 2	Y	+3, -3	+6, -6

In equilibrium,

1. What is the probability that P_1 plays X? **2/3**
2. What is the probability that P_2 plays X? **0.5**
3. What is the value of the game for P_1? **4**

2-person zero-sum games with more actions

- When there are more actions available than 2 per person, the simple algorithm will no longer work

- However, it is still possible to compute Nash equilibria for zero-sum games in polynomial time using a technique called Linear Programming

Minimax Game example

- Two candidates for presidency have two possible strategies on which to focus their campaigns.
- Depending on their decisions they may win the favour of a number of voters. The table below describes their gains in millions of voters

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Row player's strategy $<x_1, x_2>$, where x_1 and x_2 are the probabilities of taking row 1 and row 2, respectively
- Column player's strategy $<y_1, y_2>$, where y_1 and y_2 are the probabilities of taking col 1 and col 2, respectively.

On the basis of Lecture 2 - Zero-Sum Games; Minimax Theorem via Linear Programming
https://stellar.mit.edu/S/course/6/sp17/6.853/courseMaterial/topics/topic3/lectureNotes/lec2/lec2.pdf

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Suppose now that the row player announces a strategy of ⟨1/2 , 1/2 ⟩ in advance.
- This is *presumably* a “suboptimal” way to play the game, since it exposes the row player’s strategy to the column player who can now react optimally
- Given the row player’s announced strategy, the column player will choose the strategy that maximizes his payoff (value) given the row player’s declared strategy;
- in this case, what the column player will choose?

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Suppose now that the row player announces a strategy of $\langle 1/2 , 1/2 \rangle$ in advance.
- This is *presumably* a "suboptimal" way to play the game, since it exposes the row player's strategy to the column player who can now react optimally
- Given the row player's announced strategy, the column player will choose the strategy that maximizes his payoff (value) given the row player's declared strategy;
- in this case, what the column player will choose?
  - he will choose "Tax-cuts"

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- More generally, if the row player announces strategy $\langle x_1, x_2 \rangle$, then the column player has the following expected payoffs:

  $V(\text{“Morality”}) = -3x_1 + 2x_2,$

  $V(\text{“Tax-cuts”}) = x_1 - x_2$

- Thus the column player’s optimal reaction to the row player’s announced strategy is $\max(-3x_1 + 2x_2, x_1 - x_2)$
- Since this is a zero-sum game, the resulting row player’s payoff is $-\max(-3x_1 + 2x_2, x_1 - x_2) \equiv \min(3x_1 - 2x_2, -x_1 + x_2)$.

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- So if the row player is forced to announce his strategy in advance he will choose $(x_1, x_2) = \arg max_{(x_1,x_2)}$ min ($3x_1$ - $2x_2$, -x_1 + x_2)
- In other words, the row player's optimal announced strategy (x_1,x_2) can be computed via the following linear program

$$\max z$$

$$\begin{array}{rrcl} \text{s.t.} & 3x_1 - 2x_2 & \geq & z \\ & -x_1 + x_2 & \geq & z \\ & x_1 + x_2 & = & 1 \\ & x_1, x_2 & \geq & 0. \end{array}$$

The optimal *z* will be equal to the row player's payoff after the column player's optimal response to (x_1, x_2). Solving this LP yields

x_1 = 3/7,x_2 = 4/7 and z = 1/7.

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Conversely, if the column player were forced to announce his strategy in advance, he would aim to solve the following LP.

$$
\begin{aligned}
&\max w \\
\text{s.t.} \quad & -3y_1 + y_2 \geq w \\
& 2y_1 - y_2 \geq w \\
& y_1 + y_2 = 1 \\
& y_1, y_2 \geq 0,
\end{aligned}
$$

- which yields the following solution:

$y_1 = 2/7, y_2 = 5/7$ and $w = -1/7$

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Here is a remarkable observation. If the row player plays **x** = <3/7,4/7> and the column player plays **y**=<2/7,5/7> then
  - The payoff of the row player is 1/7 and the payoff of the column payer is −1/7; in particular, the sum of their payoffs is zero;
  - Given strategy **x** = ⟨ 3/7 , 4/7 ⟩ for the row player, the column player cannot obtain a payoff greater -1/7 since the row player gets a payoff of at least 1/7 irrespective of the column player’s strategy (by the definition of the first LP) and the game is zero-sum;
  - Similarly, given **y** = ⟨ 2/7 , 5/7 ⟩ for the column player, the row player cannot obtain a payoff greater than 1/7 since the column player gets a payoff of at least −1/7 irrespective of the row player’s strategy.

Minimax Game example

	Morality	Tax-cuts
Economy	3, -3	-1, 1
Society	-2, 2	1, -1

- Hence, the pair of strategies **x** = $\langle 3/7, 4/7 \rangle$ and **y** = $\langle 2/7, 5/7 \rangle$ is a Nash equilibrium!
- This was somewhat unexpected since the two LPs were formulated independently by the two players without any care/guessing about what their opponent might do.
  - In particular, it seemed a priori suboptimal for a player to announce her strategy and let the other player respond
  - Moreover, there was no reason to expect that the suboptimal strategies for the two players are in fact best responses to each other.
  - As we show next, this coincidence happens for a deep reason, it is true in all two-player zero-sum games, and is a ramification of the strong LP duality.

Duality in Linear Programming

- One of the most important discoveries in the early development of linear programming was the concept of duality
- Every linear programming problem is associated with another linear programming problem called the **dual**
- The relationships between the dual problem and the original problem (called the **primal**) prove to be extremely useful in a variety of ways

Primal and Dual Problems in Linear Programming

Primal Problem	**Dual Problem**
Max $Z = \sum_{j=1}^{n} c_j x_j,$	Min $W = \sum_{i=1}^{m} b_i y_i,$
s.t. $\sum_{j=1}^{n} a_{ij} x_j \le b_i,$	s.t. $\sum_{i=1}^{m} a_{ij} y_i \ge c_j,$
for $i = 1, 2, \ldots, m.$	for $j = 1, 2, \ldots, n.$
$x_j \ge 0,$ for $j = 1, 2, \ldots, n.$	$y_i \ge 0,$ for $i = 1, 2, \ldots, m.$

The dual problem uses exactly the same parameters as the primal problem, but in different location.

In matrix notation

Primal Problem	Dual Problem
Maximize $Z = cx,$	**Minimize** $W = yb,$
subject to	**subject to**
$Ax \leq b$	$yA \geq c$
$x \geq 0.$	$y \geq 0.$

where c and $y = [y_1, y_2, \ldots, y_m]$ are row vectors but b and x are column vectors.

Example

Primal Problem in Algebraic Form

Max $Z = 3x_1 + 5x_2,$

s.t.

$$x_1 \le 4$$
$$2x_2 \le 12$$
$$3x_1 + 2x_2 \le 18$$
$$x_1 \ge 0, x_2 \ge 0$$

Dual Problem in Algebraic Form

Min $W = 4y_1 + 12y_2 + 18y_3,$

s.t.

$$y_1 + 3y_3 \ge 3$$
$$2y_2 + 2y_3 \ge 5$$
$$y_1 \ge 0, y_2 \ge 0, y_3 \ge 0$$

Primal Problem in Matrix Form

Max

s.t.

$$Z = [3,5]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \leq \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Dual Problem in Matrix Form

Min

s.t.

$$W = [y_1, y_2, y_3]\begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix}$$

$$[y_1, y_2, y_3]\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 2 \end{bmatrix} \geq [3,5]$$

$$[y_1, y_2, y_3] \geq [0,0,0].$$

Primal-dual table for linear programming

			Primal Problem					
			Coefficient of:				Right Side	
			x_1	x_2	$\cdots$	x_n		
Dual Problem	Coefficient of:	y_1	a_{11}	a_{12}	$\cdots$	a_{1n}	$\leq b_1$	Coefficients for Objective Function (Minimize)
		y_2	a_{21}	a_{22}	$\cdots$	a_{2n}	$\leq b_2$	
		$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\vdots$	
		y_m	a_{m1}	a_{m2}	$\cdots$	a_{mn}	$\leq b_m$	
	Right Side		VI	VI	$\cdots$	VI		
			c_1	c_2	$\cdots$	c_n		
			Coefficients for Objective Function (Maximize)					

Definitions

- The feasible solutions for a dual problem are those that satisfy the condition of optimality for its primal problem.
- A maximum value of Z in a primal problem equals the minimum value of W in the dual problem.

Week Duality and Strong Duality Theorems of LP

The following relation is always maintained

$$\mathbf{yAx} \leq \mathbf{yb} \quad (\text{from Primal: } \mathbf{Ax} \leq \mathbf{b}) \qquad (1)$$

$$\mathbf{yAx} \geq \mathbf{cx} \quad (\text{from Dual : } \mathbf{yA} \geq \mathbf{c}) \qquad (2)$$

From (1) and (2), we have (Weak Duality)

$$\mathbf{cx} \leq \mathbf{yAx} \leq \mathbf{yb} \qquad (3)$$

At optimality

$$\mathbf{cx^* = y^*Ax^* = y^*b} \qquad (4)$$

*** denotes optimal value**

is always maintained (Strong Duality).

Complementary Slackness Conditions

- derived from LP Strong Duality Theorem (4)

$$(c - y^*A)\, x^* = 0 \quad (5)$$

$$y^*(b - Ax^*) = 0 \quad (6)$$

$$x_j^* > 0 \longrightarrow y^*a_j = c_j, \quad y^*a_j > c_j \longrightarrow x_j^* = 0$$

$$y_i^* > 0 \longrightarrow a_i x^* = b_i, \quad a_i x^* < b_i \longrightarrow y_i^* = 0$$

The dual of a dual

- Any pair of primal and dual problems can be converted to each other.
- The dual of a dual problem always is the primal problem.

Back to general two-player zero-sum games

- Recall that a two-player game is zero-sum iff the payoff matrices (R, C) satisfy R + C = 0.
- The LP for the row player if he is forced to announce his strategy in advance.

$$
\begin{aligned}
& \max \quad z \\
\text{s.t.} \quad & \boldsymbol{x}^T R \geq z \mathbf{1}^T \\
& \boldsymbol{x}^T \mathbf{1} = 1 \\
& \forall i, \; x_i \geq 0.
\end{aligned}
\qquad \text{(LP1)}
$$

The maximum of *z* is actually the same as:

$$\max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^T R \boldsymbol{y}.$$

The Dual of LP1

Consider its dual as

$$
\begin{aligned}
& \min \;\; z' \\
\text{s.t.} \quad & -\boldsymbol{y}^T R^T + z'\mathbf{1}^T \geq \mathbf{0} \qquad\qquad LP(2) \\
& \boldsymbol{y}^T\mathbf{1} = 1 \\
& \forall j,\;\; y_j \geq 0.
\end{aligned}
$$

- Let us define z'' = −z' and use the fact that C = −R to change LP(2) into LP(3) below, whose value has flipped sign.

$$
\begin{aligned}
& \max \;\; z'' \\
\text{s.t.} \quad & C\boldsymbol{y} \geq z''\mathbf{1} \qquad\qquad LP(3) \\
& \boldsymbol{y}^T\mathbf{1} = 1 \\
& \forall j \;\; y_j \geq 0.
\end{aligned}
$$

The maximum of z'' is equal to

$$\max_{\boldsymbol{y}} \min_{\boldsymbol{x}} \boldsymbol{x}^T C \boldsymbol{y} = -\min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^T R \boldsymbol{y}.$$

Here is the interesting fact: LP(3) is nothing but the LP that the column player would solve if he were forced to announce her strategy in advance!

Optimal Value

- Since LP (2) is the dual of LP (1), LP strong duality theorem implies the following:
  - If ($\mathbf{x}$, z) is optimal for LP (1) and ($\mathbf{y}$,z′) is optimal for LP(2), then z = z′.
  - On the other hand, given our construction of LP(3) it follows that, if ($\mathbf{y}$,z′) is optimal for LP(2), then ($\mathbf{y}$,−z′) is optimal for LP(3), and vice versa.
  - Thus, if ($\mathbf{x}$, z) is optimal for LP (1) and ($\mathbf{y}$, z′′) is optimal for LP (3), then **z = −z′′.**

Solution of LP is Nash equilibrium

- **Theorem 1.** If ($\mathbf{x}$,z) is optimal for LP(1), and ($\mathbf{y}$,z'') is optimal for LP(3), then ($\mathbf{x}$,$\mathbf{y}$) is a Nash equilibrium of (R,C). Moreover, the payoffs of the row/column player in this Nash equilibrium are z and z'' = −z respectively.

Goal: from strong duality of LP, we have z = −z'' (optimal values) -> (x, y) is NE

Solution of LP is Nash equilibrium

- **Theorem 1.** If (**x**,z) is optimal for LP(1), and (**y**,z'') is optimal for LP(3), then (**x,y**) is a Nash equilibrium of (R,C). Moreover, the payoffs of the row/column player in this Nash equilibrium are z and z'' = −z respectively.

Proof: Since (**x**, z) is feasible for LP (1), and (**y**, z'') is feasible for LP (3):

$$\boldsymbol{x}^T R \geq z\mathbf{1}^T \Rightarrow \boldsymbol{x}^T R\boldsymbol{y} \geq z \tag{1}$$

$$C\boldsymbol{y} \geq z''\mathbf{1} \Rightarrow \boldsymbol{x'}^T C\boldsymbol{y} \geq z'', \forall \boldsymbol{x'}$$

$$\Rightarrow \boldsymbol{x'}^T R\boldsymbol{y} \leq -z'', \forall \boldsymbol{x'}. \tag{2}$$

- From eq (2), if the column player is playing **y**, the row player's payoff is at most −z'' ≡ z
  - (since z'' = −z from strong LP duality). $x'^T R y \leq z$ (the max)
- Eq (1) implies that if the row player plays **x** against **y**, his payoff will be at least z; in fact, given (2) it is exactly z. $x^T R y \geq z$, but $x^T R y \leq z \Rightarrow x^T R y = z$
- Thus, **x** is a best response for the row player against strategy y of the column player, giving payoff of z to the row player $x^T R y = z \Rightarrow x \in BR(y)$
- Similarly, **y** is a best response for the column player against strategy **x** of the row player, giving payoff z'' to the column player. $y \in BR(x)$
- Hence, (x, y) is a Nash equilibrium in which the players' payoffs are z and z'' = −z respectively.

Can every Nash Equilibrium of the two-player zero sum game can be found by LP?

- **Theorem 2.** If (**x**,**y**) is a Nash equilibrium of (R,C), then ($\mathbf{x}, \mathbf{x}^T R\mathbf{y}$) is an optimal solution of LP(1), and ($\mathbf{y}, -\mathbf{x}^T C\mathbf{y}$) is an optimal solution of LP (2).

Proof: Let $z = \mathbf{x}^T R\mathbf{y} = -\mathbf{x}^T C\mathbf{y}$. Since (**x, y**) is a Nash equilibrium, we have

$$
\begin{aligned}
& \boldsymbol{x}^T C \boldsymbol{y} && \geq \boldsymbol{x}^T C \boldsymbol{e_j} \quad \forall j \\
\Rightarrow\quad & -\boldsymbol{x}^T R \boldsymbol{y} && \geq -\boldsymbol{x}^T R \boldsymbol{e_j} \quad \forall j \\
\Rightarrow\quad & \boldsymbol{x}^T R && \geq z\mathbf{1}.
\end{aligned}
$$

- So (**x**,z) is a feasible solution for LP(1);
  - similarly we can argue that (**y**,z) is a feasible solution for LP (2).
- Since LP (2) is the dual of LP (1), by weak duality, we know that the value achieved by any feasible solution of LP(2) upper bounds the value achieved by any feasible solution of LP(1)
- But, these values are equal here. So, (**x**,z) and (**y**,z) must be optimal solutions for LP(1) and LP(2) respectively.

LP for Zero Sum Games

- Assume we have a 2 × 2 zero-sum matrix game given as

$$R_1 = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}, \; R_2 = -R_1$$

where R_1 is player 1's reward matrix and R_2 is player 2's reward matrix. We denote p_j (j = 1, 2) as the probability distribution over player 1's jth action and q_j as the probability distribution over player 2's jth action.

- The LP for player 1 is

Find (p_1, p_2) to maximize V_1

subject to

$$r_{11}p_1 + r_{21}p_2 \geq V_1$$
$$r_{12}p_1 + r_{22}p_2 \geq V_1$$
$$p_1 + p_2 = 1$$
$$p_j \geq 0, \quad j = 1, 2$$

The LP for player 2 is

Find (q_1, q_2) to maximize V_2

subject to

$$-r_{11}q_1 - r_{12}q_2 \geq V_2$$
$$-r_{21}q_1 - r_{22}q_2 \geq V_2$$
$$q_1 + q_2 = 1$$
$$q_j \geq 0, \quad j = 1, 2$$

LP is typically solved by **simplex method**

LP for Matching Pennies Games

- Assume we have a 2 × 2 zero-sum matrix game given as

$$R_1 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

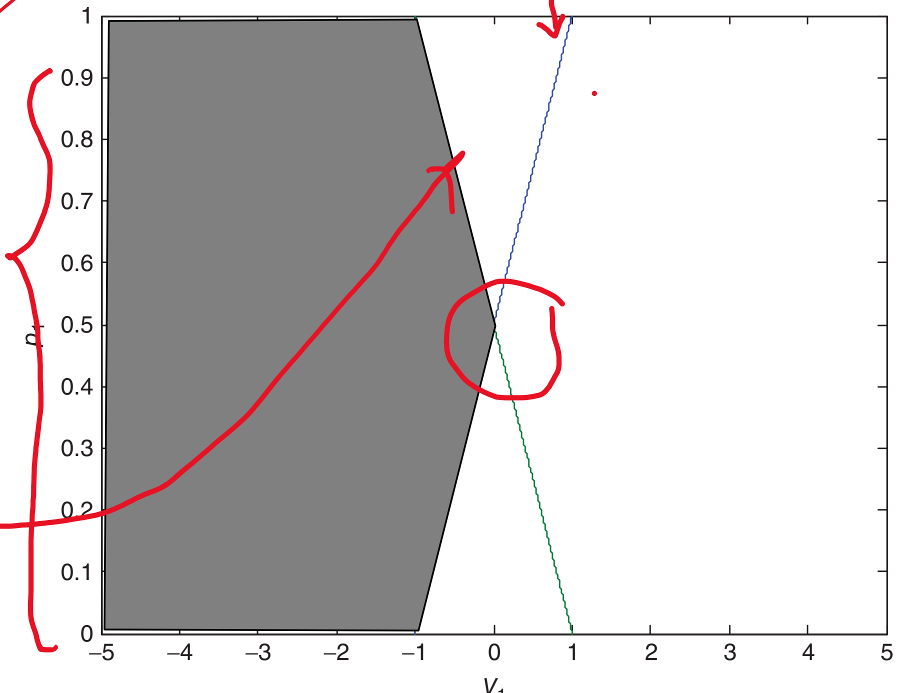
- Since p_2=1-p_1 ,the LP for player 1 is

Player 1: find p_1 to maximize V_1

subject to

$$2p_1 - 1 \geq V_1$$
$$-2p_1 + 1 \geq V_1$$
$$0 \leq p_1 \leq 1$$

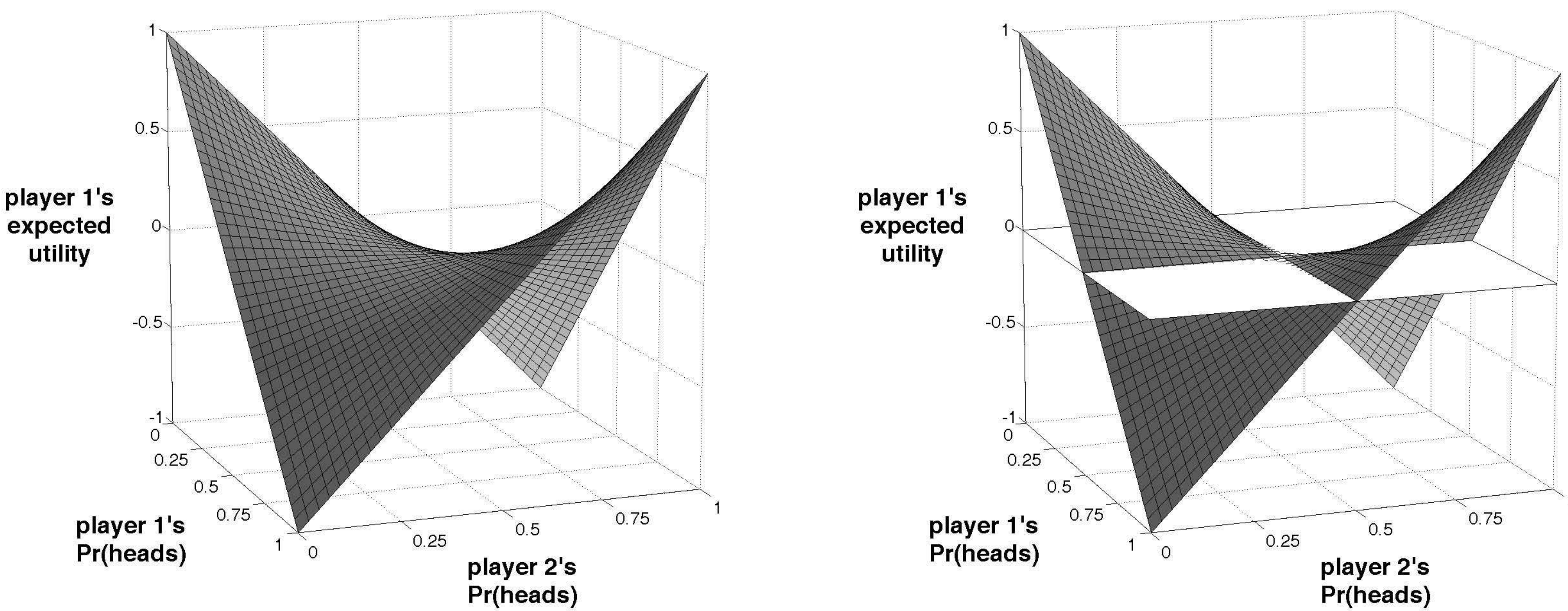
Similarly, we can use the simplex method to find the Nash equilibrium strategy for player 2



The plot of p_1 over V_1 where the grey area satisfies the constraints. Solution (p_1 = 0.5&V_1=0)

The saddle point in Matching Pennies

- At a saddle point, any deviation of the agent lowers his utility and increases the utility of the other agent



with and without a plane at utility z = 0.

LP for Zero Sum Games (in scalar form)

- Consider a two-player, zero-sum game

$$G = (\{1, 2\}, A_1 \times A_2, (u_1, u_2))$$

- Let U_i^* be the expected utility for player i in equilibrium (the value of the game);
- s_i^j is the prob. of player i choosing action j
- If the game is zero-sum, then $U_1^* = -U_2^*$

(Primal) The LP for player 1 is

$$\begin{aligned} \text{maximize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{j \in A_1} u_1(a_1^j, a_2^k) \cdot s_1^j \geq U_1^* \quad \forall k \in A_2 \\ & \sum_{j \in A_1} s_1^j = 1 \\ & s_1^j \geq 0 \qquad \forall j \in A_1 \end{aligned}$$

(Dual)The LP for player 2 is

$$\begin{aligned} \text{minimize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k \leq U_1^* \quad \forall j \in A_1 \\ & \sum_{k \in A_2} s_2^k = 1 \\ & s_2^k \geq 0 \qquad \forall k \in A_2 \end{aligned}$$

LP is typically solved by **simplex method**

LP for Zero Sum Games (in scalar form)

Note 1: all that the utility terms $u_1(\cdot)$ are constants in the linear program, while the mixed strategy terms s_2^k and U_1^* are variables

- s_i^j is the prob. of player i choosing action j
- If the game is zero-sum, then $U_1^* = -U_2^*$

(Primal) The LP for player 1 is

$$\begin{aligned} \text{maximize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{j \in A_1} u_1(a_1^j, a_2^k) \cdot s_1^j \geq U_1^* \quad \forall k \in A_2 \\ & \sum_{j \in A_1} s_1^j = 1 \\ & s_1^j \geq 0 \qquad \forall j \in A_1 \end{aligned}$$

(Dual)The LP for player 2 is

$$\begin{aligned} \text{minimize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k \leq U_1^* \quad \forall j \in A_1 \\ & \sum_{k \in A_2} s_2^k = 1 \\ & s_2^k \geq 0 \qquad \forall k \in A_2 \end{aligned}$$

LP is typically solved by **simplex method**

LP for Zero Sum Games (in scalar form)

Note 2: The first constraint states that for every pure strategy j of player 1, his expected utility for playing any action $j \in A_1$ given player 2's mixed strategy s_2 is at most U_1^*. optimise worse case.

Those pure strategies for which the expected utility is exactly U_1^* will be in player 1's **best response set**, while those pure strategies leading to lower expected utility will not.

$$
\begin{aligned}
\text{maximize} \quad & U_1^* \\
\text{subject to} \quad & \sum_{j \in A_1} u_1(a_1^j, a_2^k) \cdot s_1^j \geq U_1^* \quad \forall k \in A_2 \\
& \sum_{j \in A_1} s_1^j = 1 \\
& s_1^j \geq 0 \quad \forall j \in A_1
\end{aligned}
$$

$$
\begin{aligned}
\text{minimize} \quad & U_1^* \\
\text{subject to} \quad & \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k \leq U_1^* \quad \forall j \in A_1 \\
& \sum_{k \in A_2} s_2^k = 1 \\
& s_2^k \geq 0 \quad \forall k \in A_2
\end{aligned}
$$

LP is typically solved by simplex method

LP for Zero Sum Games (in scalar form)

Note 3: the LP will choose player 2's mixed strategy in order to minimize U_1^* subject to the constraint. Thus, the minimization and the first constraint state that player 2 plays the mixed strategy that minimizes the utility player 1 can gain by playing his best response.

Minimax U_1^* to obtain player 2's optimal strategy

- s_i^j is the prob. of player i choosing action j
- If the game is zero-sum, then $U_1^* = -U_2^*$

(Primal) The LP for player 1 is

$$\begin{aligned} &\text{maximize} \quad U_1^* \\ &\text{subject to} \quad \sum_{j \in A_1} u_1(a_1^j, a_2^k) \cdot s_1^j \geq U_1^* \quad \forall k \in A_2 \\ &\qquad \sum_{j \in A_1} s_1^j = 1 \\ &\qquad s_1^j \geq 0 \quad \forall j \in A_1 \end{aligned}$$

(Dual)The LP for player 2 is

$$\begin{aligned} &\text{minimize} \quad U_1^* \\ &\text{subject to} \quad \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k \leq U_1^* \quad \forall j \in A_1 \\ &\qquad \sum_{k \in A_2} s_2^k = 1 \\ &\qquad s_2^k \geq 0 \quad \forall k \in A_2 \end{aligned}$$

LP is typically solved by simplex method

LP for Zero Sum Games (in scalar form)

Note 4: This LP reverses the roles of player 1 and player 2 in the constraints; the objective is to *maximize* U_1^*, as player 1 wants to maximize his own payoffs.

Maximin U_1^* to obtain player 1's optimal strategy

(the value of the game);

- s_i^j is the prob. of player i choosing action j
- If the game is zero-sum, then $U_1^* = -U_2^*$

(Primal) The LP for player 1 is

$$\begin{aligned} \text{maximize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{j \in A_1} u_1(a_1^j, a_2^k) \cdot s_1^j \geq U_1^* \quad \forall k \in A_2 \\ & \sum_{j \in A_1} s_1^j = 1 \\ & s_1^j \geq 0 \qquad\qquad \forall j \in A_1 \end{aligned}$$

(Dual)The LP for player 2 is

$$\begin{aligned} \text{minimize} \quad & U_1^* \\ \text{subject to} \quad & \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k \leq U_1^* \quad \forall j \in A_1 \\ & \sum_{k \in A_2} s_2^k = 1 \\ & s_2^k \geq 0 \qquad\qquad \forall k \in A_2 \end{aligned}$$

LP is typically solved by **simplex method**

LP with *slack variables*

- We give a formulation equivalent to our LP (player 2) by introducing *slack variables* r_1^j for every $j \in A_1$ and
- then replacing the inequality constraints with equality constraints

The LP for player 2 is

$$\begin{cases} r_1^j = 0, \text{ while } s_1^j > 0 \rightarrow j \in A_1 \text{ BR set} \\ r_1^j > 0 \text{ while } s_1^j = 0 \quad j \notin \end{cases}$$

$$\begin{array}{lll}
\text{minimize} & U_1^* & \\
\text{subject to} & \sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* & \forall j \in A_1 \\
& \sum_{k \in A_2} s_2^k = 1 & \\
& s_2^k \geq 0 & \forall k \in A_2 \\
& r_1^j \geq 0 & \forall j \in A_1
\end{array}$$

It is equivalent as each slack variable must be positive.

Exercise : Odd or Even

- Players I and II simultaneously call out one of the numbers one or two.
  - Player I wins if the sum of the numbers is odd.
  - Player II wins if the sum of the numbers is even.
  - The amount paid to the winner by the loser is always the sum of the numbers in dollars.
  - Action for Player I: X = {1, 2}, Action for Player II: Y = {1, 2}, and Payoff for Player I is

$$
\begin{array}{ccc}
 & & \text{II (even)} \quad y \\
 & & \begin{array}{cc} 1 & 2 \end{array} \\
\text{I (odd)} & x \quad \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} -2 & +3 \\ +3 & -4 \end{pmatrix}
\end{array}
$$

$A(x, y) = \text{I's winnings} = \text{II's losses.}$

Game Theory, Second Edition, 2014. Thomas S. Ferguson

Computing Nash equilibria of two-player, general-sum games

- Unfortunately, finding a Nash equilibrium of a two-player, general- sum game cannot be formulated as a linear program
  - The two players' interests are no longer *completely* opposed
  - However, we can still state our problem as some optimization problem

Computing Nash equilibria of two-player, general-sum games

- We first consider an inner or totally mixed Nash equilibrium (x^*,y^*), i.e., $x_i^* > 0$ for all i and $y_j^* > 0$ for all j (all pure strategies are used with positive probability).
- Let a_i denote the rows of payoff matrix A and b_j denote the columns of payoff matrix B.
- Using the fact that all pure strategies in the support of a Nash equilibrium strategy yields the same payoff, which is also greater than or equal to the payoffs for strategies outside the support), we have
  - $a_1 y^* = a_i y^*$, $i = 2,\ldots,n$,
  - $(x^*)^T b_1 = (x^*)^T b_j$, $j = 2,\ldots,m$.

Assume every pure strategy is played with positive prob.

- The preceding is a system of linear equations which can be solved efficiently.

Computing Nash equilibria of two-player, general-sum games

- However, the assumption that every strategy is played with positive probability is restrictive. Most games do not have totally mixed Nash equilibria; for them,
- We compute all the Nash equilibria of a finite two-player game:
  - A mixed strategy profile $(x^*,y^*) \in X \times Y$ is a Nash equilibrium with support $\underline{S}_1 \subset S_1$ and $\underline{S}_2 \subset S_2$ if and only if

$$u = a_i\, y* \,,\ \forall\, i \in \underline{S}_1 \,,\ u \geq a_i\, y* \,,\ \forall\, i \notin \underline{S}_1 \,,$$
$$v = (x\, *)^T\, b_j \,,\ \forall\, j \in \underline{S}_2 \,,\ v \geq (x\, *)^T b_j \,,\ \forall\, j \notin \underline{S}_2$$
$$x^*_i = 0,\ \forall i \notin \underline{S}_1,\ y_j = 0,\ \forall\, j \notin \underline{S}_2.$$

(u,v) are the values in NE.

- To find the right supports for the above procedure to work, we need to search over all possible supports. Since there are 2^{n+m} different supports, this procedure leads to an exponential complexity in the number of pure strategies of the players.
- Remark: computational complexity of computing Nash equilibrium for finite games lies in finding the right support sets.

Optimization Formulation

- A general method for the solution of a bimatrix game is to transform it into a nonlinear (in fact, a bilinear) programming problem
- A mixed strategy profile (x∗,y∗) is a mixed Nash equilibrium of the bimatrix game (A,B) if and only if there exists a pair (p∗,q∗) such that (x∗,y∗,p∗,q∗) is a solution to the following bilinear programming problem:

$$
\begin{array}{ll}
\textit{maximize} & \{x^T Ay + x^T By - p - q\} \\
\textit{subject to} & Ay \le p\mathbf{1}_n, \qquad B^T x \le q\mathbf{1}_m, \\
& \sum_i x_i = 1, \qquad \sum_j y_j = 1, \\
& x \ge 0, \ \ y \ge 0,
\end{array}
$$

where 1_n (1_m) denotes the n (m)-dimensional vector with all components equal to 1.

Computing Nash equilibria of two-player, general-sum games

- Alternatively, finding a Nash equilibrium of a two-player, general-sum game can be formulated as a *linear complementarity problem* (LCP)
- Our LCP will have no objective function at all
  - It is a constraint satisfaction problem, or a *feasibility program*, rather than an optimization problem.
- Also, we can no longer determine one player's equilibrium strategy by only considering the other player's payoff;
  - instead, need to discuss both players explicitly

Linear Complementarity Problem

- The LCP for computing the Nash equilibrium of a general-sum two-player game follows:

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

$$s_1^j \geq 0, \quad s_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \geq 0, \quad r_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \cdot s_1^j = 0, \quad r_2^k \cdot s_2^k = 0 \qquad \forall j \in A_1, \forall k \in A_2$$

Linear Complementarity Problem

Note 1: This formulation bears a strong resemblance to the LP formulation with slack variables, but **without objective function**

of a general-sum two-player game follows.

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

$$s_1^j \geq 0, \quad s_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \geq 0, \quad r_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \cdot s_1^j = 0, \quad r_2^k \cdot s_2^k = 0 \qquad \forall j \in A_1, \forall k \in A_2$$

Linear Complementarity Problem

- The LCP for computing the Nash equilibrium of a general-sum two-player game follows.

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

Note 2: The first constraint is the same as that in our LP formulation with slack variables; however, here we also include the second constraint which constrains player 1's actions in the same way.

Linear Complementarity Problem

- The LCP for computing the Nash equilibrium

Note 3: also give the standard constraints that probabilities sum to one, that probabilities are nonnegative, and that slack variables are non- negative, but now state these constraints for both players rather than only for player 1.

$j \in A_1$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

$$s_1^j \geq 0, \quad s_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \geq 0, \quad r_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \cdot s_1^j = 0, \quad r_2^k \cdot s_2^k = 0 \qquad \forall j \in A_1, \forall k \in A_2$$

Linear Complementarity Problem

- The LCP for computing the Nash equilibrium of a general-sum two-player game follows.

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum s^j = 1 \qquad \sum s^k = 1$$

Problem: with the constraints we have described so far, U_1^* and U_2^* would be allowed to take unboundedly large values, because all of these constraints remain satisfied when both U_i^* and r_i^j are increased by the same constant, for any given i and j.

Solution: We solve this problem by adding the nonlinear constraint, called the ***complementarity condition***. The addition of this constraint means that we no longer have a linear program; instead, we have a ***linear complementarity problem***.

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

$$s_1^j \geq 0, \quad s_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \geq 0, \quad r_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \cdot s_1^j = 0, \quad r_2^k \cdot s_2^k = 0 \qquad \forall j \in A_1, \forall k \in A_2$$

Solution: This constraint requires that

- whenever an action is played by a given player with positive probability (i.e., whenever an action is in the support of a given player's mixed strategy) then the corresponding slack variable must be zero.
- Under this requirement, each slack variable can be viewed as the player's incentive to deviate from the corresponding action.
- Thus, the complementarity condition captures the fact that, in equilibrium, all strategies that are played with positive probability must yield the same expected payoff, while all strategies that lead to lower expected payoffs are not played.
- Taking all of our constraints together, each player plays a best response to the other player's mixed strategy: the definition of a Nash equilibrium.

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

$$\sum_{j \in A_1} s_1^j = 1, \quad \sum_{k \in A_2} s_2^k = 1$$

$$s_1^j \geq 0, \quad s_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \geq 0, \quad r_2^k \geq 0 \qquad \forall j \in A_1, \forall k \in A_2$$

$$r_1^j \cdot s_1^j = 0, \quad r_2^k \cdot s_2^k = 0 \qquad \forall j \in A_1, \forall k \in A_2$$

Recall: Complementary Slackness Conditions (in standard LP matrix form)

- $(cx^* = y^*Ax^* = y^*b)$ LP Strong Duality Theorem gives

$$(c - y^*A)\, x^* = 0 \quad (5)$$

$$y^*(b - Ax^*) = 0 \quad (6)$$

$$x_j^* > 0 \longrightarrow y^*a_j = c_j, \quad y^*a_j > c_j \longrightarrow x_j^* = 0$$

$$y_i^* > 0 \longrightarrow a_ix^* = b_i, \quad a_i x^* < b_i \longrightarrow y_i^* = 0$$

This is consistent with the theory from LP!

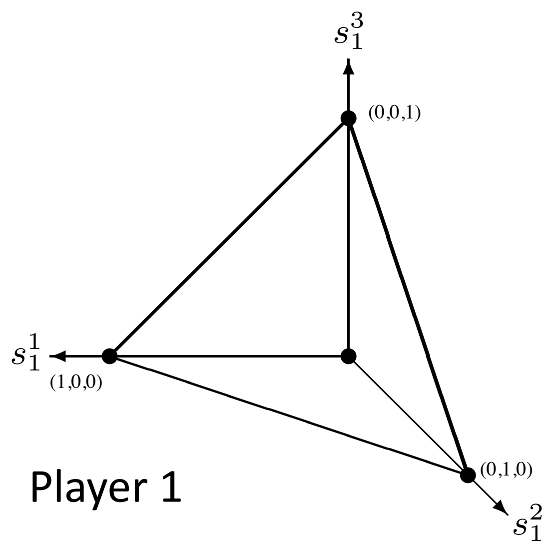
The Lemke–Howson algorithm

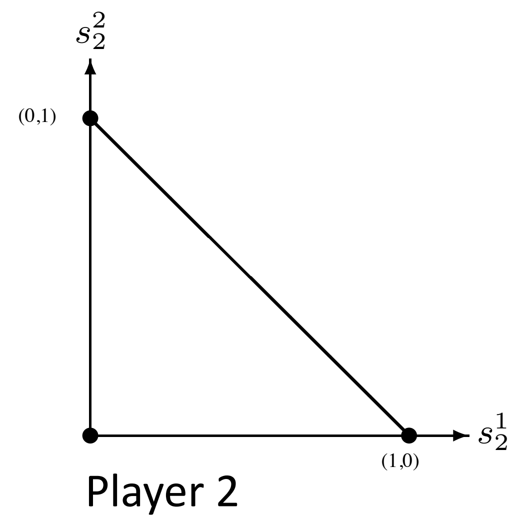
- The best-known algorithm to solve this LCP formulation is

  the *Lemke–Howson algorithm*

- We will explain it through a graphical exposition

Player 1 \ Player 2		
	$0, 1$	$6, 0$
	$2, 0$	$5, 2$
	$3, 4$	$3, 3$





Each axis corresponds to one of the corresponding player's pure strategies and the region spanned by these axes represents all the mixed strategies

Labeling on the strategies

- Every possible mixed strategy s_i is given a set of labels
  - $L(s^j_i) \subseteq A_1 \cup A_2$ drawn from the set of available actions for both players
- A given player as i and the other player as $-i$
- Mixed strategy s_i for player i is labeled as follows:
  - with each of player i's actions a^j_i that is *not* in the support of s_i(with zero prob. of choosing it) and
  - with each of player $-i$'s actions a^j_{-i} that *is* a best response by player $-i$ to s_i.

Labeling on the strategies

- This labeling is useful because a pair of strategies (s_1 , s_2) is a Nash equilibrium if and only if it is completely labeled
  - i.e., L(s1) ∪ L(s2) = A1 ∪ A2
- For a pair to be completely labeled, each action a^j_i for agent i must
  - either played by player i with zero probability, or
  - a^j_{-i}:be a best response by opponent player -i to the mixed strategy of himself (player i)

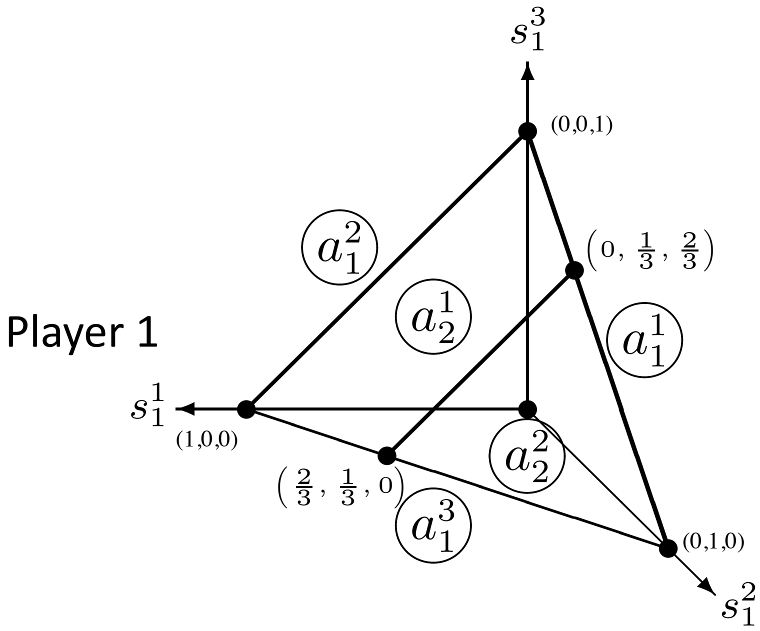
Labeling on the strategies

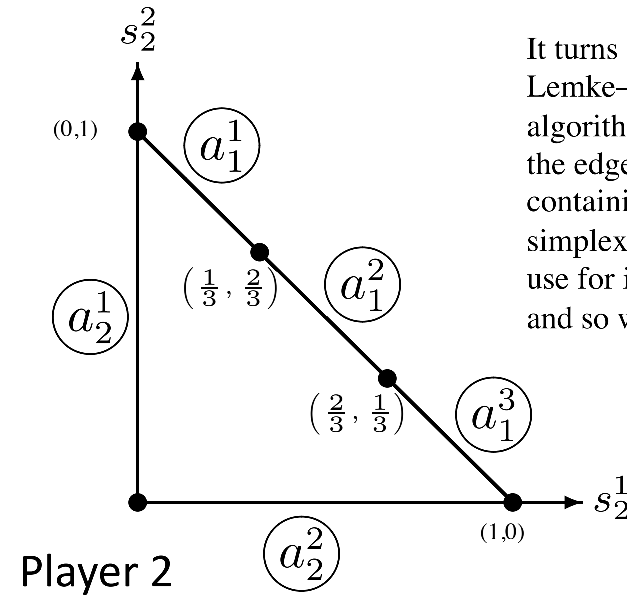
- This is a restatement of the complementarity condition given in constraint in our LCP
  - because the slack variable r_i^j is zero exactly when its corresponding action a_i^j is a best response to the mixed strategy s_{-i}.

$$\sum_{k \in A_2} u_1(a_1^j, a_2^k) \cdot s_2^k + r_1^j = U_1^* \qquad \forall j \in A_1$$

$$\sum_{j \in A_1} u_2(a_1^j, a_2^k) \cdot s_1^j + r_2^k = U_2^* \qquad \forall k \in A_2$$

Labeled strategy space

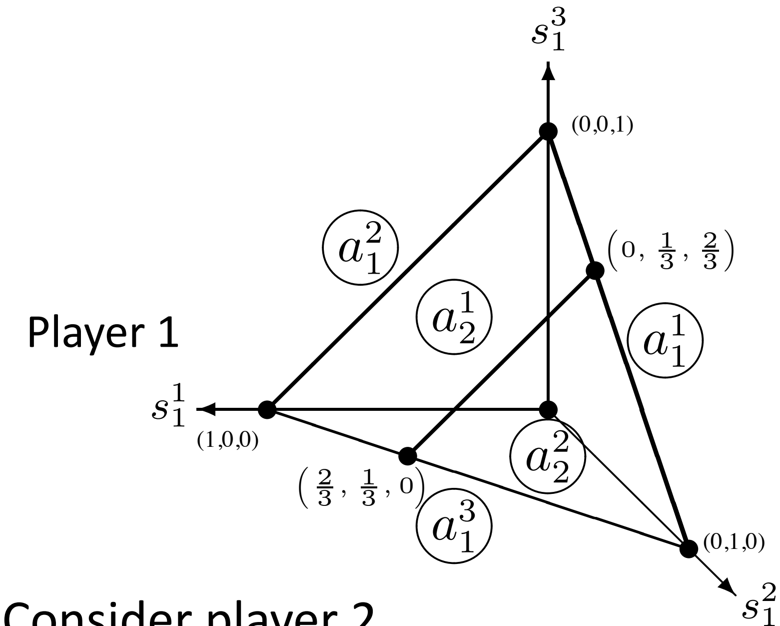


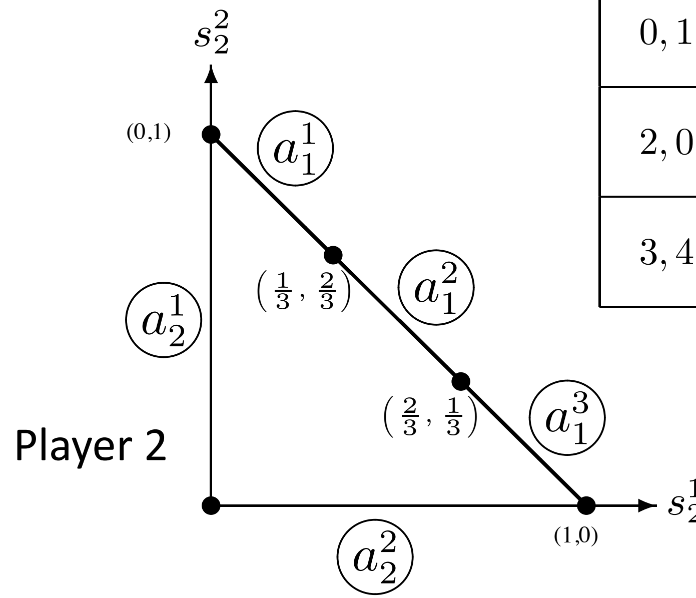


It turns out that the Lemke–Howson algorithm traverses only the edges of the polygon containing the simplexes and has no use for interior points, and so we ignore them.

- It is convenient to add one *fictitious point* in the strategy space of each agent, the origin; that is, (0,0,0) for player 1 and (0,0) for player 2.
  - player 2's strategy space is a triangle with vertices (0, 0), (1, 0), and (0, 1), while player 1's strategy space is a pyramid with vertices (0, 0, 0), (1, 0, 0), (0, 1, 0), and (0, 0, 1)

Labeled strategy space

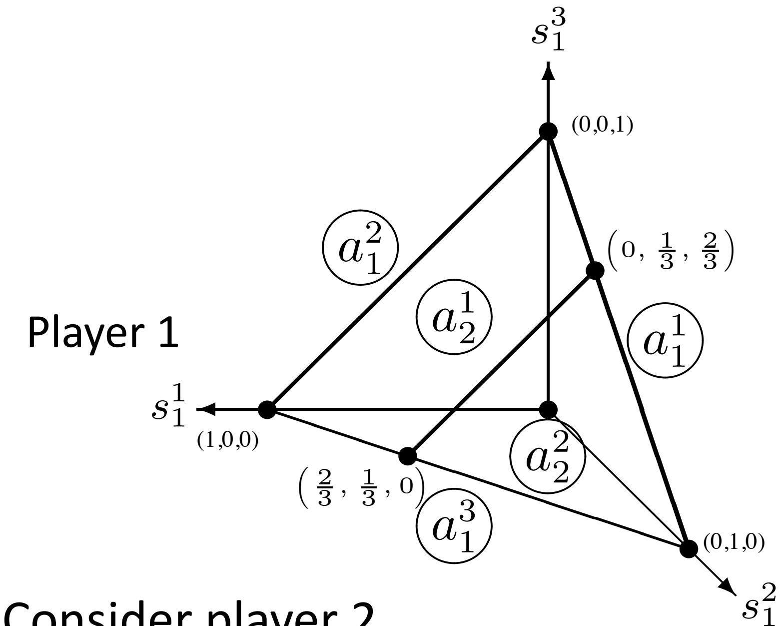


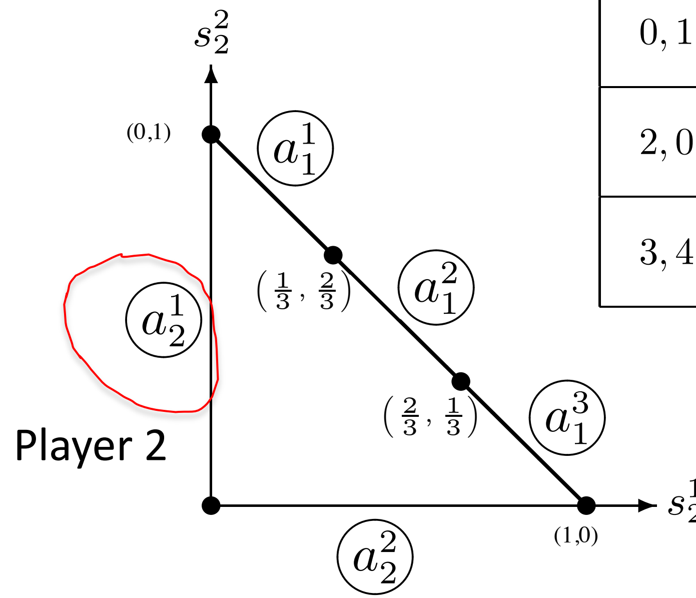


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

- Consider player 2
  - The line from (0,0) to (0,1) is labeled with $a^1{}_2$,
    - As none of them assign any probability to playing action $a^1{}_2$
  - Similarly, the line from (0, 0) to (1, 0) is labeled with $a^2{}_2$
  - Action $a_1{}^1$ is a best response by player 1 to any of the mixed strategies represented by the line from (0, 1) to (1/3 , 2/3)
  - the point (1/3 , 2/3) is labeled by both $a_1{}^1$ and $a_1{}^2$, because both of these actions are best responses by player 1 to the mixed strategy (1/3 , 2/3) by player 2

Labeled strategy space



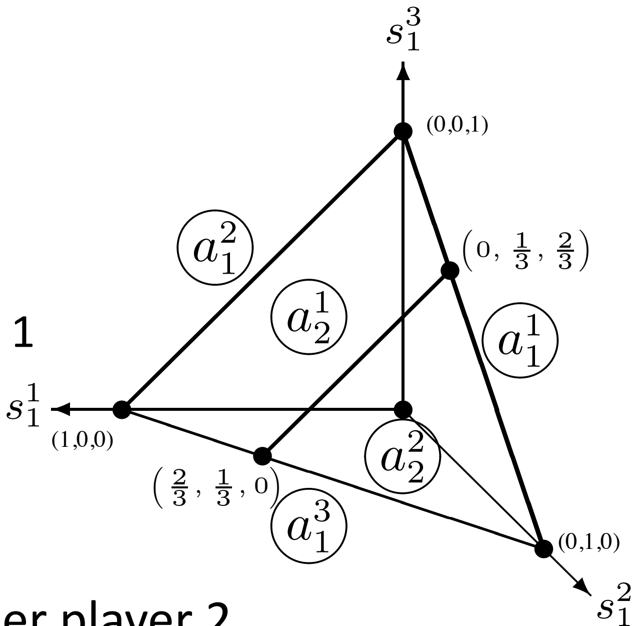


0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

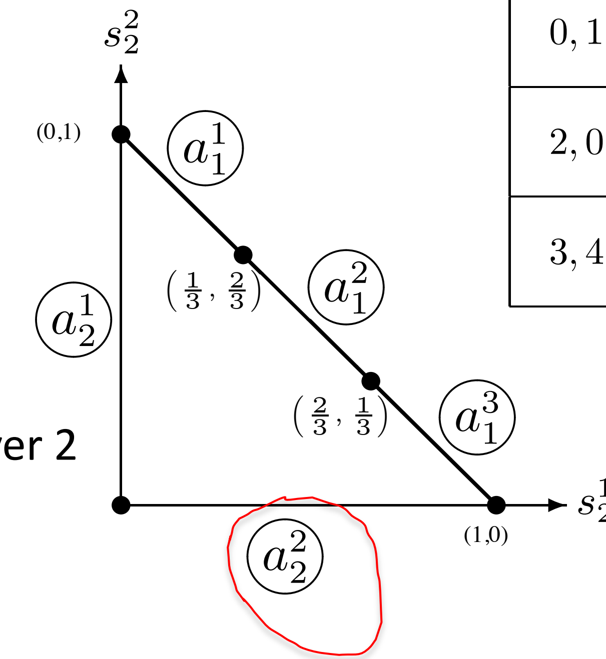
- Consider player 2
  - The line from (0,0) to (0,1) is labeled with $a^1{}_2$,
    - As none of them assign any probability to playing action $a^1{}_2$
  - Similarly, the line from (0, 0) to (1, 0) is labeled with $a^2{}_2$
  - Action a_1^1 is a best response by player 1 to any of the mixed strategies represented by the line from (0, 1) to (1/3 , 2/3)
  - the point (1/3 , 2/3) is labeled by both a_1^1 and a_1^2, because both of these actions are best responses by player 1 to the mixed strategy (1/3 , 2/3) by player 2

Labeled strategy space

Player 1



Player 2

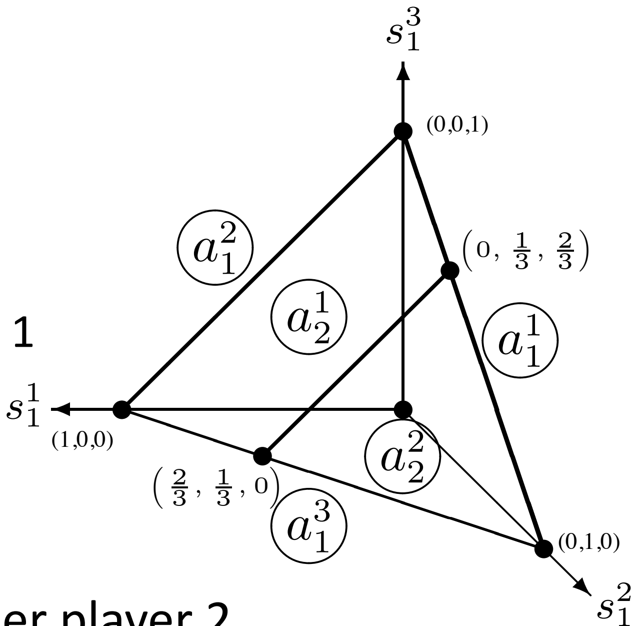


0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

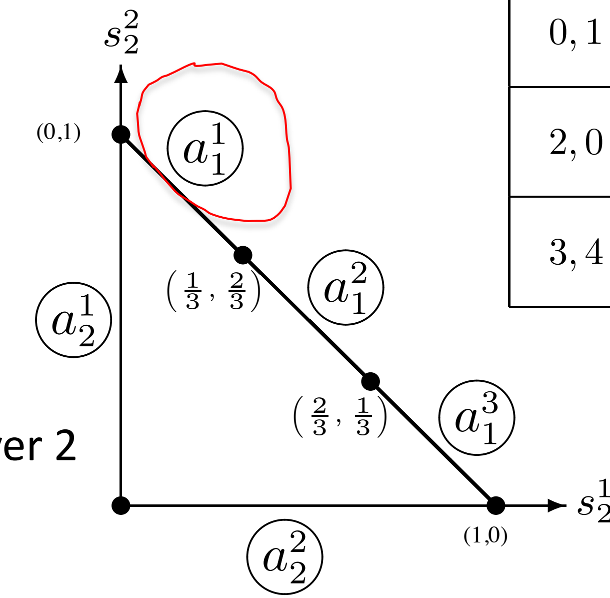
- Consider player 2
  - The line from (0,0) to (0,1) is labeled with $a^1{}_2$,
    - As none of them assign any probability to playing action $a^1{}_2$
  - Similarly, the line from (0, 0) to (1, 0) is labeled with $a^2{}_2$
  - Action $a_1{}^1$ is a best response by player 1 to any of the mixed strategies represented by the line from (0, 1) to (1/3 , 2/3)
  - the point (1/3 , 2/3) is labeled by both $a_1{}^1$ and $a_1{}^2$, because both of these actions are best responses by player 1 to the mixed strategy (1/3 , 2/3) by player 2

Labeled strategy space

Player 1



Player 2

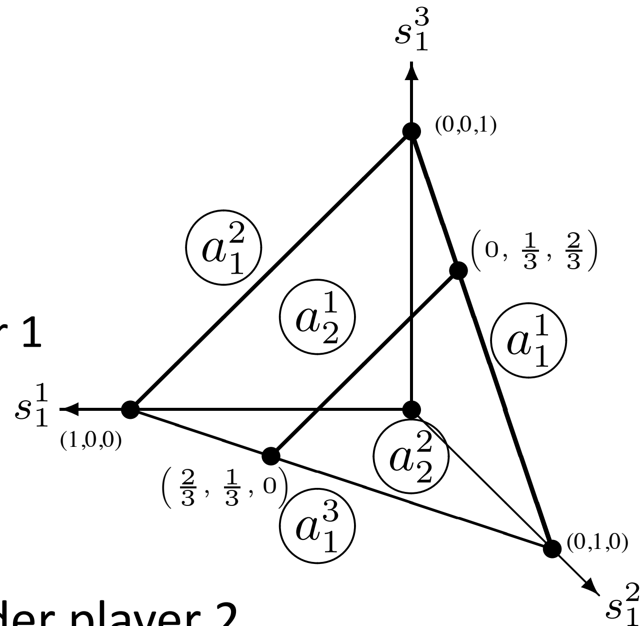


0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

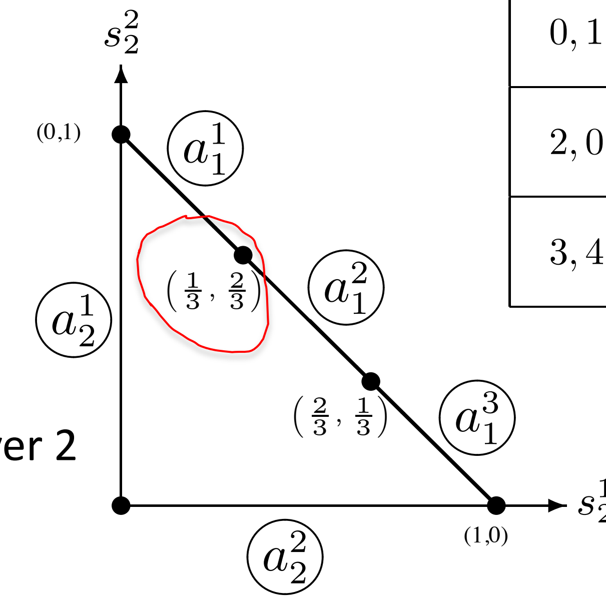
- Consider player 2
  - The line from (0,0) to (0,1) is labeled with $a^1{}_2$,
    - As none of them assign any probability to playing action $a^1{}_2$
  - Similarly, the line from (0, 0) to (1, 0) is labeled with $a^2{}_2$
  - Action $a_1{}^1$ is a best response by player 1 to any of the mixed strategies represented by the line from (0, 1) to (1/3 , 2/3)
  - the point (1/3 , 2/3) is labeled by both $a_1{}^1$ and $a_1{}^2$, because both of these actions are best responses by player 1 to the mixed strategy (1/3 , 2/3) by player 2

Labeled strategy space

Player 1



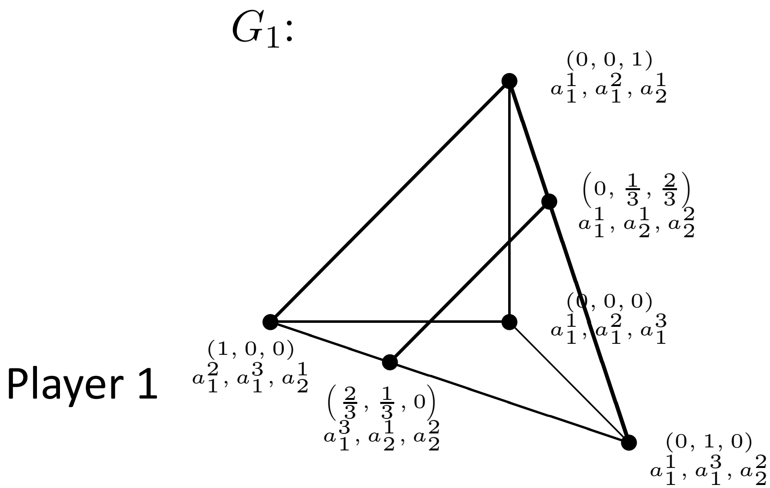
Player 2

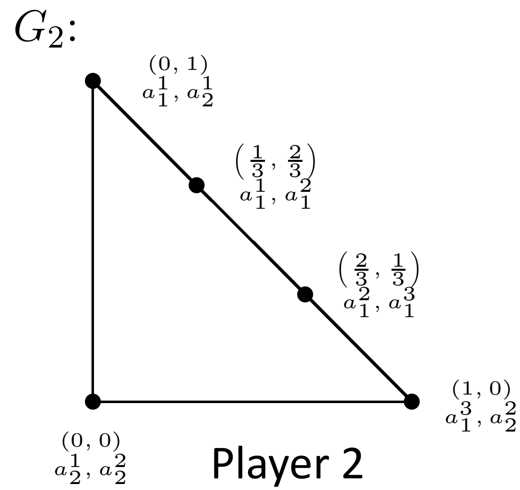


0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

- Consider player 2
  - The line from (0,0) to (0,1) is labeled with $a^1{}_2$,
    - As none of them assign any probability to playing action $a^1{}_2$
  - Similarly, the line from (0, 0) to (1, 0) is labeled with $a^2{}_2$
  - Action $a_1{}^1$ is a best response by player 1 to any of the mixed strategies represented by the line from (0, 1) to (1/3 , 2/3)
  - the point (1/3 , 2/3) is labeled by both $a_1{}^1$ and $a_1{}^2$, because both of these actions are best responses by player 1 to the mixed strategy (1/3 , 2/3) by player 2

The Lemke–Howson algorithm



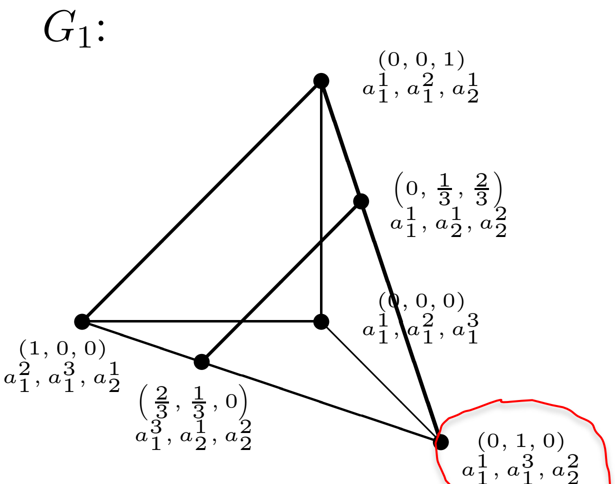


0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

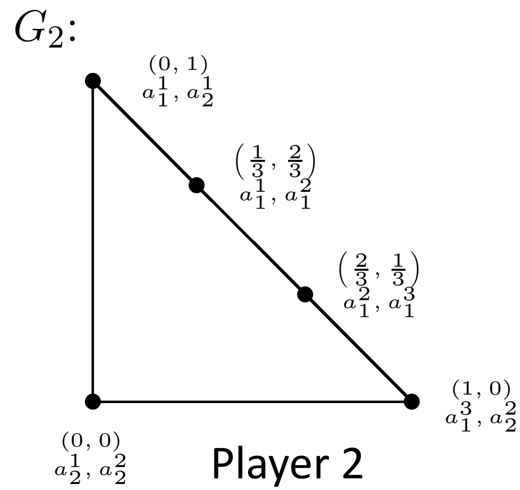
Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- Define G1 and G2 to be graphs, for players 1 and 2 respectively.
- The nodes in the graph are fully labeled points in the labeled space, that is, triply labeled points in G1 and doubly labeled points in G2.
- An edge exists between pairs of points that differ in exactly one label.
- The Lemke–Howson algorithm is to search these pairs of labeled spaces for a completely labeled pair (Nash equilibrium)
- In our example, the three Nash equilibria of the game:
  [(0, 0, 1), (1, 0)], [(0, 1/3, 2/3) , (2/3, 1/30)] , and [(2/3, 1/3,0) , (1/3, 2/3)]

The Lemke–Howson algorithm



Player 1

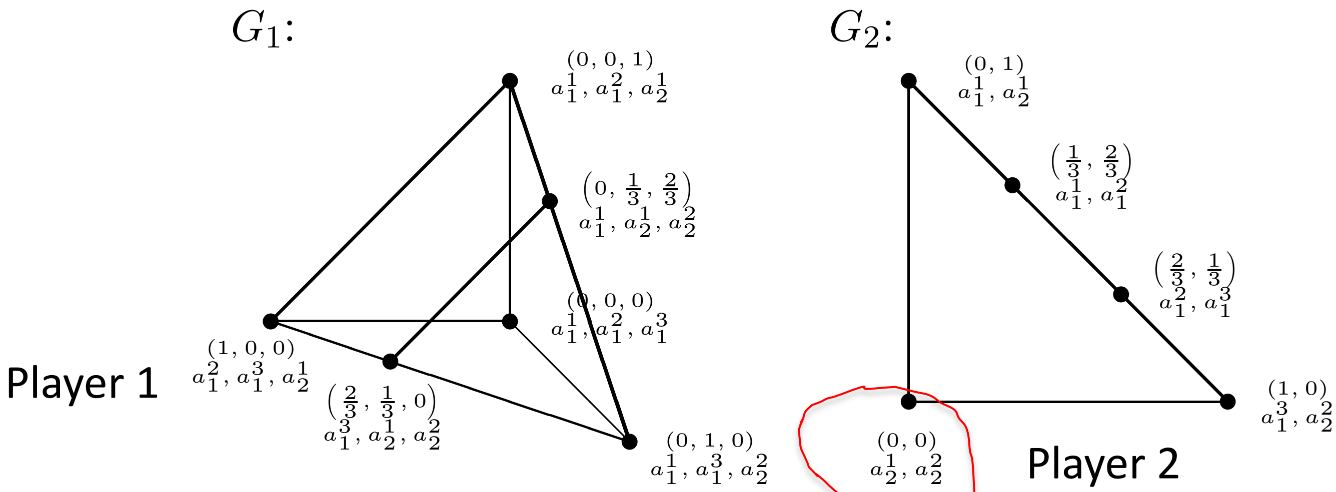


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- Define G1 and G2 to be graphs, for players 1 and 2 respectively.
- The nodes in the graph are fully labeled points in the labeled space, that is, triply labeled points in G1 and doubly labeled points in G2.
- An edge exists between pairs of points that differ in exactly one label.
- The Lemke–Howson algorithm is to search these pairs of labeled spaces for a completely labeled pair (Nash equilibrium)
- In our example, the three Nash equilibria of the game:
  [(0, 0, 1), (1, 0)], [(0, 1/3, 2/3) , (2/3, 1/30)] , and [(2/3, 1/3,0) , (1/3, 2/3)]

The Lemke–Howson algorithm

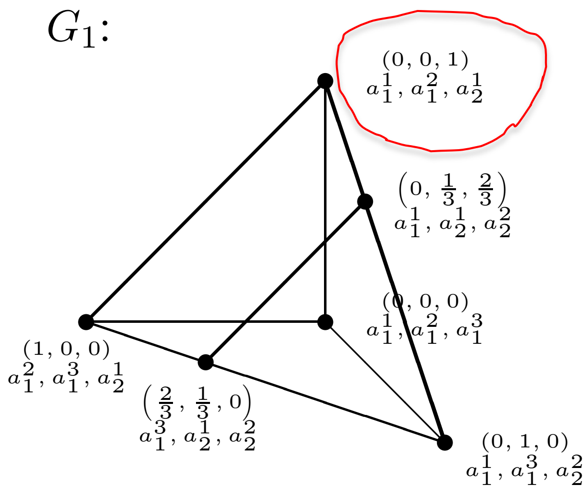


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

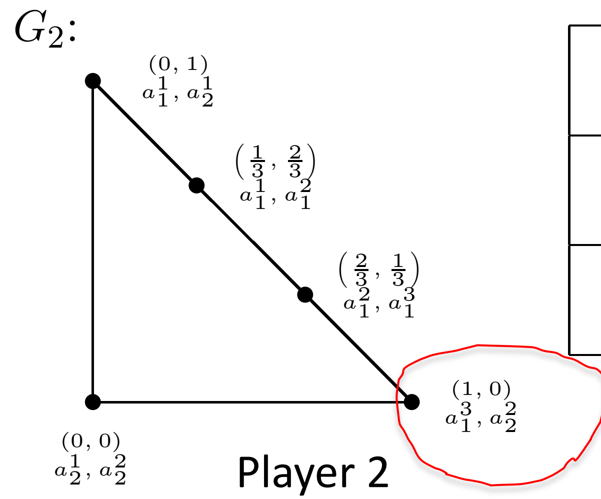
Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- Define G1 and G2 to be graphs, for players 1 and 2 respectively.
- The nodes in the graph are fully labeled points in the labeled space, that is, triply labeled points in G1 and doubly labeled points in G2.
- An edge exists between pairs of points that differ in exactly one label.
- The Lemke–Howson algorithm is to search these pairs of labeled spaces for a completely labeled pair (Nash equilibrium)
- In our example, the three Nash equilibria of the game:
  [(0, 0, 1), (1, 0)], [(0, 1/3, 2/3) , (2/3, 1/30)] , and [(2/3, 1/3,0) , (1/3, 2/3)]

The Lemke–Howson algorithm



Player 1

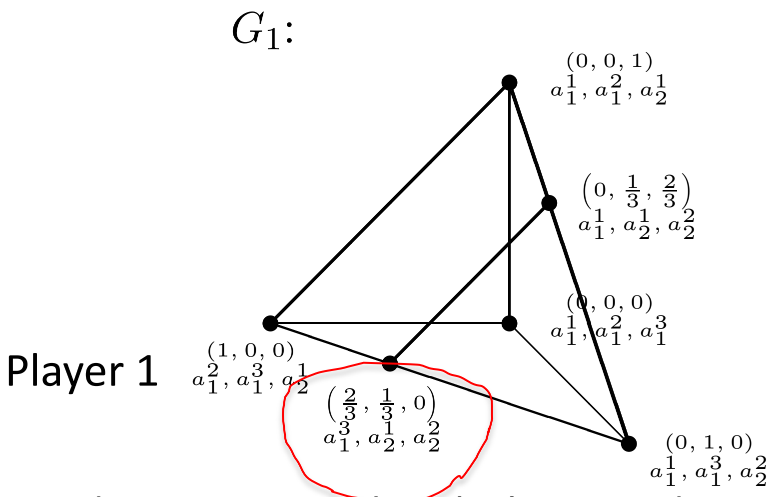


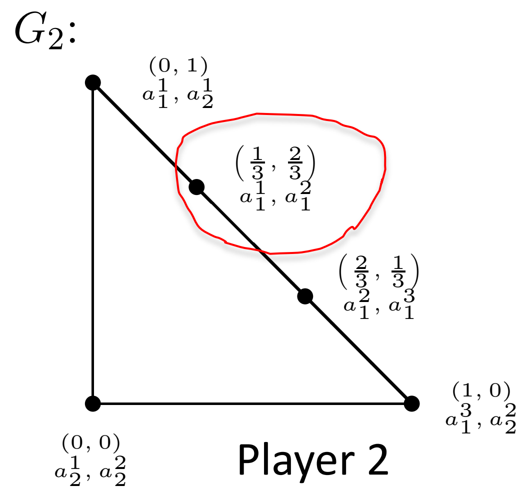
$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- Define G1 and G2 to be graphs, for players 1 and 2 respectively.
- The nodes in the graph are fully labeled points in the labeled space, that is, triply labeled points in G1 and doubly labeled points in G2.
- An edge exists between pairs of points that differ in exactly one label.
- The Lemke–Howson algorithm is to search these pairs of labeled spaces for a completely labeled pair (Nash equilibrium)
- In our example, the three Nash equilibria of the game:
  [(0, 0, 1), (1, 0)], [(0, 1/3, 2/3) , (2/3, 1/30)] , and [(2/3, 1/3,0) , (1/3, 2/3)]

The Lemke–Howson algorithm



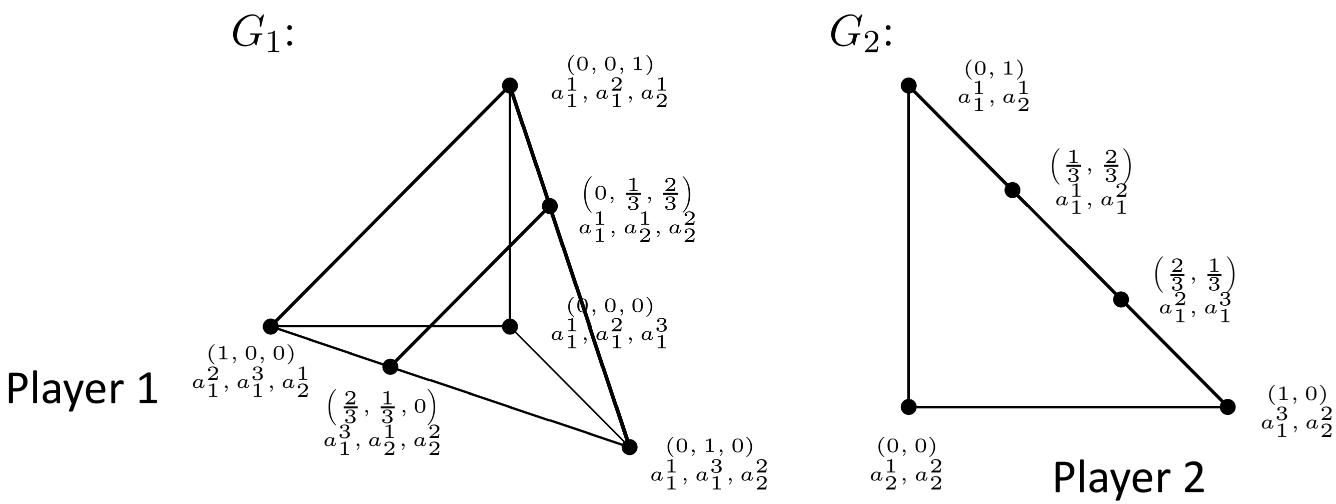


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- Define G1 and G2 to be graphs, for players 1 and 2 respectively.
- The nodes in the graph are fully labeled points in the labeled space, that is, triply labeled points in G1 and doubly labeled points in G2.
- An edge exists between pairs of points that differ in exactly one label.
- The Lemke–Howson algorithm is to search these pairs of labeled spaces for a completely labeled pair (Nash equilibrium)
- In our example, the three Nash equilibria of the game:
  [(0, 0, 1), (1, 0)], [(0, 1/3, 2/3) , (2/3, 1/30)] , and [(2/3, 1/3,0) , (1/3, 2/3)]

The Lemke–Howson algorithm

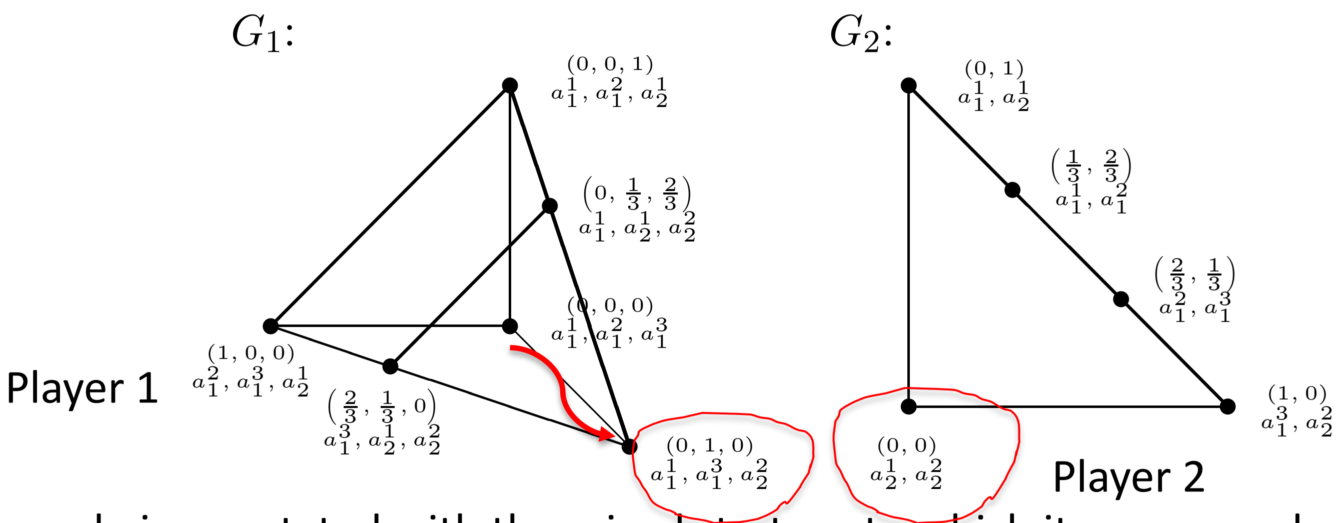


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- The Lemke–Howson algorithm finds an equilibrium by following a path through pairs $(s_1, s_2) \in G_1 \times G_2$ in the cross product of the two graphs.
- Alternating between the two graphs, each iteration changes one of the two points to a new point that is connected by an edge to the original point.
  - Starting from (**0**, **0**), which is completely labeled, the algorithm picks one of the two graphs and moves from **0** in that graph to some adjacent node x.
  - The node x, together with the **0** from the other graph, form an almost completely labeled pair, in that between them they miss exactly one label.
  - The algorithm then moves from the remaining **0** to a neighboring node that picks up that missing label, but in the process loses a different label.
  - The process thus proceeds, alternating between the two graphs, until an equilibrium (i.e., a totally labeled pair) is reached.

The Lemke–Howson algorithm

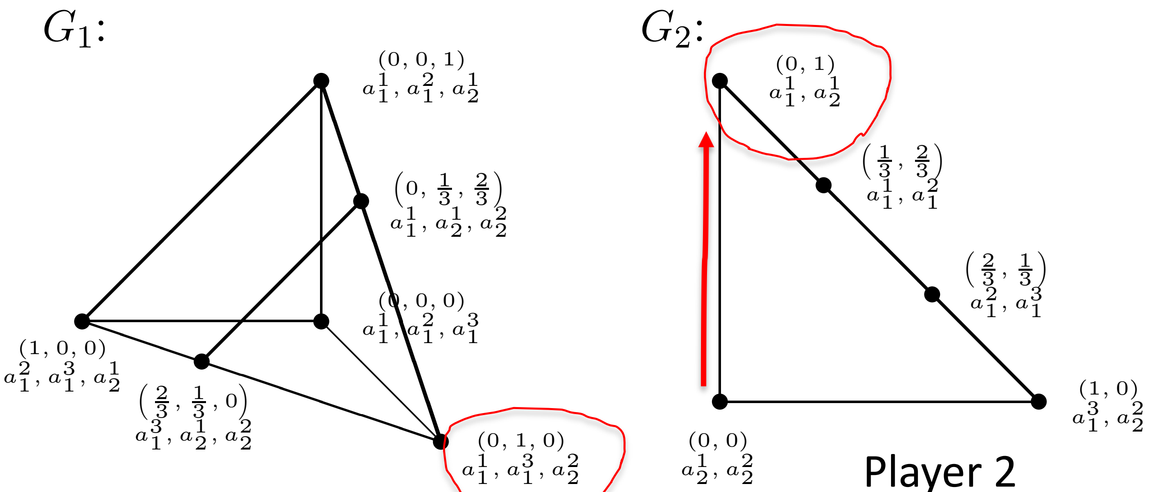


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- In our running example, a possible execution of the algorithm starts at (**0**, **0**) and
  - (1) then changes s_1 to (0, 1, 0). Now, our pair (0, 1, 0), (0, 0) is $a^2{}_1$-almost completely labeled, and the duplicate label is $a_2{}^2$.

The Lemke–Howson algorithm

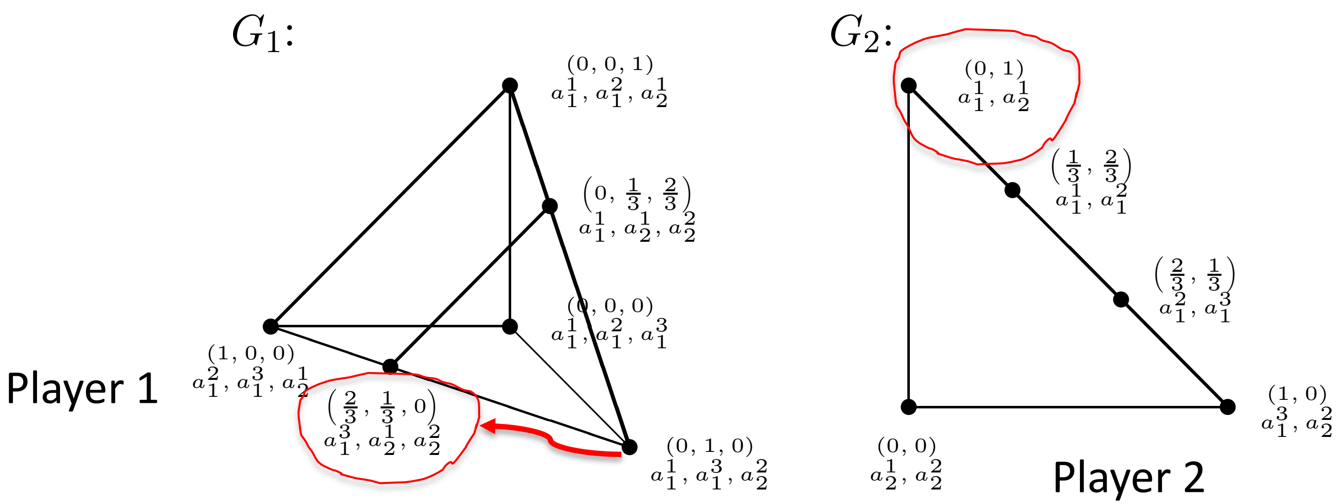


$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- In our running example, a possible execution of the algorithm starts at (**0**, **0**) and
  - (2) For its next step in G2 moves to (0, 1) because the other possible choice (1, 0) has label a^2_2

The Lemke–Howson algorithm



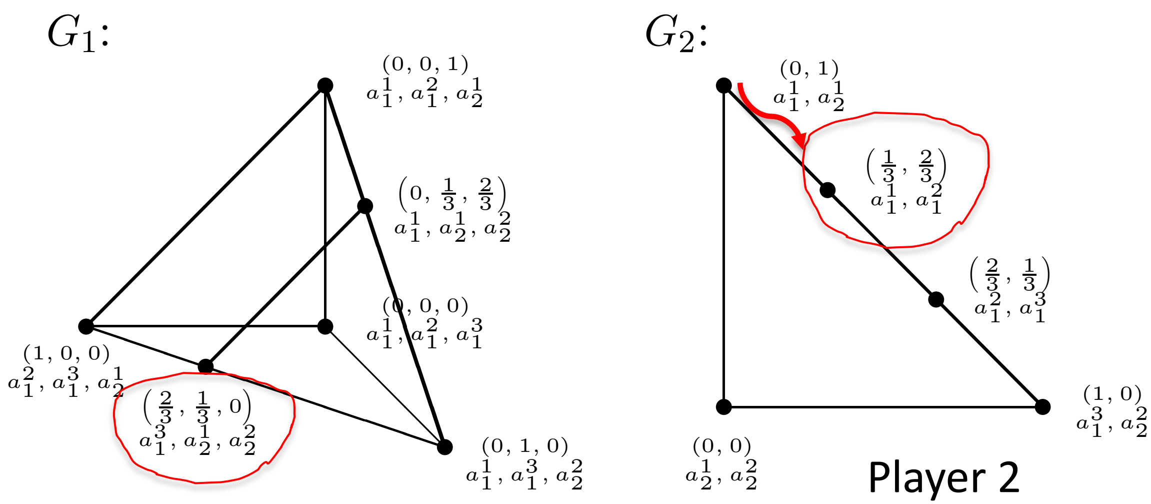
0, 1	6, 0
2, 0	5, 2
3, 4	3, 3

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- In our running example, a possible execution of the algorithm starts at (**0**, **0**) and
  - (3) Returning to G1 for the next iteration, we move to (2/3 , 1/3 , 0) because it is the point adjacent to (0, 1, 0) that does not have the duplicate label a_1^1

The Lemke–Howson algorithm

Player 1



$0, 1$	$6, 0$
$2, 0$	$5, 2$
$3, 4$	$3, 3$

Each node is annotated with the mixed strategy to which it corresponds as well as the actions with which it is labeled.

- In our running example, a possible execution of the algorithm starts at (**0**, **0**) and
  - (4) the final step is to change s_2 to (1/3,2/3) in order to move away from the label a^1_2
- Execution trace :((0, 0, 0), (0, 0)) → ((0,1,0),(0,0)) → ((0,1,0),(0,1)) → ((2/3, 1/3,0),(0,1)) → ((2/3, 1/3,0),(1/3, 2/3)).

Conclusions

- We have learned
  - A simple solution for finding Nash Equilibrium
  - Linear programming solutions for Finding Nash Equilibrium in zero-sum games
  - The Lemke–Howson algorithm for finding sample Nash Equilibrium in general-sum games
- The Minimax theorem and the Strong Duality theorem of LP are equivalent.

References

The slides are based on the following materials：

- Shoham, Yoav, and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.
- Schwartz, Howard M. *Multi-agent machine learning: A reinforcement approach*. John Wiley & Sons, 2014.
- Lecture 2 - Zero-Sum Games; Minimax Theorem via Linear Programming https://stellar.mit.edu/S/course/6/sp17/6.853/courseMaterial/topics/topic3/lectureNotes/lec2/lec2.pdf
- Von Stengel, Bernhard. "Computing equilibria for two-person games." *Handbook of game theory with economic applications*3 (2002): 1723-1759.
- Game Theory and Algorithms Lecture 6: The Lemke-Howson Algorithm http://ints.io/daveagp/gta/lecture6.pdf