Algorithmic game theory

Laurea Magistrale in Computer Science

2024/25

Lecture 9

Learning in finite games

Can equilibria be dynamically learnt?

Learning in finite games

Can equilibria be dynamically learnt?

General framework

- mixed strategies
- successive repetitions of the (same) game
- players have complete knowledge of past actions

Learning in finite games

Can equilibria be dynamically learnt?

General framework

- mixed strategies
- successive repetitions of the (same) game
- players have complete knowledge of past actions

Fictitious play (Brown 1949, 1951)

Learning in finite games

Can equilibria be dynamically learnt?

General framework

- mixed strategies
- successive repetitions of the (same) game
- players have complete knowledge of past actions

Fictitious play (Brown 1949, 1951)

players:

- keep track of all the previous rounds
- compute the average behaviour of the other players
- best respond to the average behaviour

average behaviour = average of the (mixed) strategies players chose in all rounds

Learning through averaging

$$S_i = \{1, \ldots, m_i\} \quad (i = 1, \ldots, n)$$

Ficititious play process

1. $\sigma^1 = (\sigma_1^1, \ldots, \sigma_n^1) \in \Delta_{m_1} \times \cdots \Delta_{m_n},\ k = 1$
2. *compute* $\hat{\sigma}_{-i}^k = \left(\sum_{\ell=1}^{k} \sigma_{-i}^{\ell}\right)/k$
3. σ_i^{k+1} *is a best response to* $\hat{\sigma}_{-i}^k$ $(\sigma_i^{k+1} \in R_i(\hat{\sigma}_{-i}^k))$
4. $k = k + 1$ *and go back to 2*

(Steps 2–3) $\Longrightarrow \sigma^{k+1} \in R(\hat{\sigma}^k)$

Knowledge of other players' utility functions is not required

Lack of a stopping criterion (other than *"σ^k is a Nash equilibrium"*)

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)		

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)
3	(0,0,1)	(1,0,0)		

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)
3	(0,0,1)	(1,0,0)	(1/3,0,2/3)	(1/3,2/3,0)

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)
3	(0,0,1)	(1,0,0)	(1/3,0,2/3)	(1/3,2/3,0)
4	(0,1,0)	(1,0,0)		

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)
3	(0,0,1)	(1,0,0)	(1/3,0,2/3)	(1/3,2/3,0)
4	(0,1,0)	(1,0,0)	(1/4,1/4,1/2)	(1/2,1/2,0)
5	....		....	....

Fictitious play in rock-paper-scissors

I/II	paper	scissors	rock
paper	0	-1	1
scissors	1	0	-1
rock	-1	1	0

$$h(p,q) = p^T A q = \begin{cases} p_1(q_3 - q_2) + p_2(q_1 - q_3) + p_3(q_2 - q_1) \leftarrow \max_p \\ q_1(p_2 - p_3) + q_2(p_3 - p_1) + q_3(p_1 - p_2) \leftarrow \min_q \end{cases}$$

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(0,0,1)	(0,1,0)	(1/2,0,1/2)	(0,1,0)
3	(0,0,1)	(1,0,0)	(1/3,0,2/3)	(1/3,2/3,0)
4	(0,1,0)	(1,0,0)	(1/4,1/4,1/2)	(1/2,1/2,0)
5	....		....	....

↓

convergence to the unique Nash equilibrium ((1/3,1/3,1/3),(1/3,1/3,1/3))?

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(1,0,0)	(0,0,1)		

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/2,1/2)

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/2,1/2)
3	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/3,2/3)

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/2,1/2)
3	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/3,2/3)
4	(0,1,0)	(0,0,1)	(3/4,1/4,0)	(0,1/4,3/4)
5	....		....	....

↓

convergence to the unique Nash equilibrium ((1/3,1/3,1/3),(1/3,1/3,1/3))?

Fictitious play in a gentle rock-paper-scissors (Shapley 1964)

I/II	Left	Middle	Right
Top	(0,0)	(1,0)	(0,1)
Middle	(0,1)	(0,0)	(1,0)
Down	(1,0)	(0,1)	(0,0)

no actual loss → not a zero-sum game

it. k	chosen strategies σ^k		average behaviours $\hat{\sigma}^k$	
1	(1,0,0)	(0,1,0)	(1,0,0)	(0,1,0)
2	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/2,1/2)
3	(1,0,0)	(0,0,1)	(1,0,0)	(0,1/3,2/3)
4	(0,1,0)	(0,0,1)	(3/4,1/4,0)	(0,1/4,3/4)
5	....		....	....

↓

convergence to the unique Nash equilibrium ((1/3,1/3,1/3),(1/3,1/3,1/3))?

not really! ⟶ asympotically stable limit cycling

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* *(Robinson 1951)*

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* *(Robinson 1951)*
- *two player 2×2 games satisfying the diagonal property* *(Miyasawa 1961)*

diagonal property:

$a_{11} + a_{22} \neq a_{12} + a_{21}$ and $b_{11} + b_{22} \neq b_{12} + b_{21}$ where

I/II	1	2
1	(a_{11}, b_{11})	(a_{12}, b_{12})
2	(a_{21}, b_{21})	(a_{22}, b_{22})

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* *(Robinson 1951)*
- *two player 2×2 games satisfying the diagonal property* *(Miyasawa 1961)*
- *two player 2×n nondegenerate games* *(Berger 2005)*

diagonal property:

$a_{11} + a_{22} \neq a_{12} + a_{21}$ and $b_{11} + b_{22} \neq b_{12} + b_{21}$ where

I/II	1	2
1	(a_{11}, b_{11})	(a_{12}, b_{12})
2	(a_{21}, b_{21})	(a_{22}, b_{22})

nondegenerate = unique best response to pure strategies

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* *(Robinson 1951)*
- *two player 2×2 games satisfying the diagonal property* *(Miyasawa 1961)*
- *two player 2×n nondegenerate games* *(Berger 2005)*
- *games that are solvable by the IESDS algorithm* *(Milgrom-Roberts 1991)*

diagonal property:

$a_{11} + a_{22} \neq a_{12} + a_{21}$ and $b_{11} + b_{22} \neq b_{12} + b_{21}$ where

I/II	1	2
1	(a_{11}, b_{11})	(a_{12}, b_{12})
2	(a_{21}, b_{21})	(a_{22}, b_{22})

nondegenerate = unique best response to pure strategies

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* *(Robinson 1951)*
- *two player 2×2 games satisfying the diagonal property* *(Miyasawa 1961)*
- *two player 2×n nondegenerate games* *(Berger 2005)*
- *games that are solvable by the IESDS algorithm* *(Milgrom-Roberts 1991)*
- *ordinal potential games* *(Monderer-Shapley 1996)*

diagonal property:

$a_{11} + a_{22} \neq a_{12} + a_{21}$ and $b_{11} + b_{22} \neq b_{12} + b_{21}$ where

I/II	1	2
1	(a_{11}, b_{11})	(a_{12}, b_{12})
2	(a_{21}, b_{21})	(a_{22}, b_{22})

nondegenerate = unique best response to pure strategies

Convergence of the fictitious play process

A bunch of games

The fictitious play process converges to a Nash equilibrium for

- *two player zero-sum games* (Robinson 1951)
- *two player 2×2 games satisfying the diagonal property* (Miyasawa 1961)
- *two player 2×n nondegenerate games* (Berger 2005)
- *games that are solvable by the IESDS algorithm* (Milgrom-Roberts 1991)
- *ordinal potential games* (Monderer-Shapley 1996)

diagonal property:

$a_{11} + a_{22} \neq a_{12} + a_{21}$ and $b_{11} + b_{22} \neq b_{12} + b_{21}$ where

I/II	1	2
1	(a_{11}, b_{11})	(a_{12}, b_{12})
2	(a_{21}, b_{21})	(a_{22}, b_{22})

nondegenerate = unique best response to pure strategies

rate of convergence in two player zero-sum games (Robinson 1951)

$$0 \leq w_2(\sigma_2^k) - w_1(\sigma_1^k) = O(1/\sqrt[m]{k}) \quad \textit{with} \quad m = m_1 + m_2 - 2$$

$$\big(w_1(\sigma_1^k) \leq \text{value of the game} \leq w_2(\sigma_2^k)\big)$$

Another learning approach

General framework

- successive repetitions of the (same) game
- pure strategies are played (no fictitious environment)
- players have complete knowledge of past actions

Another learning approach: avoid regretting

General framework

- successive repetitions of the (same) game
- pure strategies are played (no fictitious environment)
- players have complete knowledge of past actions

Regret matching (Hart-Mas Colell 2000)

Another learning approach: avoid regretting

General framework

- successive repetitions of the (same) game
- pure strategies are played (no fictitious environment)
- players have complete knowledge of past actions

Regret matching (Hart-Mas Colell 2000)

players:

- keep track of all the previous rounds
- measure the regret of not having played other strategies
- choose a pure strategy in a probabilistic fashion according to regrets

higher regret calls for lower probability

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

	past rounds	
it. k	strategies	utility
1	③ ③	5
2	② ①	5
3	③ ①	3
4	③ ②	4

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

	past rounds	
it. k	strategies	utility
1	③ ③	5
2	② ①	5
3	③ ①	3
4	③ ②	4

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

past rounds			what if		regret wrt	
it. k	strategies	utility	①	②	①	②
1	③ ③	5				
2	② ①	5				
3	③ ①	3				
4	③ ②	4				

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

	past rounds			what if		regret wrt	
it. k	strategies		utility	①	②	①	②
1	③	③	5	5	1	0	-4
2	②	①	5				
3	③	①	3				
4	③	②	4				

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

	past rounds		what if		regret wrt	
it. k	strategies	utility	①	②	①	②
1	③ ③	5	5	1	0	-4
2	② ①	5	5	0	0	0
3	③ ①	3				
4	③ ②	4				

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

	past rounds			what if		regret wrt	
it. k	strategies		utility	①	②	①	②
1	③	③	5	5	1	0	-4
2	②	①	5	5	0	0	0
3	③	①	3	5	5	2	2
4	③	②	4				

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

past rounds			what if		regret wrt	
it. k	strategies	utility	①	②	①	②
1	③ ③	5	5	1	0	-4
2	② ①	5	5	0	0	0
3	③ ①	3	5	5	2	2
4	③ ②	4	7	5	3	1

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

past rounds			what if		regret wrt	
it. k	strategies	utility	①	②	①	②
1	③ ③	5	5	1	0	-4
2	② ①	5	5	0	0	0
3	③ ①	3	5	5	2	2
4	③ ②	4	7	5	3	1
					5	-~~1~~

up to now what overall regret of choosing ③ for player I?

Measuring regrets

I/II	①	②	③	④
①	(5,3)	(7,4)	(5,2)	(3,4)
②	(5,5)	(5,7)	(1,1)	(2,5)
③	(3,4)	(4,2)	(5,5)	(6,3)

past rounds it. k	strategies		utility	what if ①	②	regret wrt ①	②
1	③	③	5	5	1	0	-4
2	②	①	5	5	0	0	0
3	③	①	3	5	5	2	2
4	③	②	4	7	5	3	1
						5	-~~1~~

up to now what overall regret of choosing ③ for player I?

draw a pure strategy for round $k = 5$ from the probability distribution

p(①)=$5/\mu_k$, p(②)=0, p(③)=$1\text{-}5/\mu_k$

(for some suitable $\mu_k > 0$)

Learning through no regrets

Regret matching process

1. $x^1 = (x_1^1, \ldots, x_n^1) \in S_1 \times \cdots S_m$, $k = 1$ $\quad (S_i = \{1, \ldots, m_i\}\ i = 1, \ldots, n)$
2. *compute regrets* $R_i^k(s, x_i^k)$ *for all* $s \in S_i$, $s \neq x_i^k$
3. *draw* x_i^{k+1} *from* $p_i^k \in \Delta_{m_i}$

   (steps 2–3) $\Longrightarrow x^{k+1}$
4. $k = k + 1$ *and go back to 2*

Learning through no regrets

Regret matching process

1. $x^1 = (x_1^1, \ldots, x_n^1) \in S_1 \times \cdots S_m,\ k = 1$ $\quad (S_i = \{1, \ldots, m_i\}\ i = 1, \ldots, n)$
2. *compute regrets* $R_i^k(s, x_i^k)$ *for all* $s \in S_i$, $s \neq x_i^k$
3. *draw* x_i^{k+1} *from* $p_i^k \in \Delta_{m_i}$

$\Big\} \implies x^{k+1}$

4. $k = k + 1$ *and go back to 2*

previous rounds where player i chose strategy $x_i^k \in S_i$: $T_i^k(x_i^k) = \{t \leq k \ : \ x_i^t = x_i^k\}$

regret of choosing x_i^k over any other $s \in S_i$ in the previous rounds:

$$R_i^k(s, x_i^k) = \Big[\sum_{t \in T_i^k(x_i^k)} \big(u_i(s, x_{-i}^t) - u_i(x_i^k, x_{-i}^t)\big) \Big]^+ \qquad ([a]^+ = \max\{a, 0\})$$

probability distribution:

$$p_i^k(s) = R_i^k(s, x_i^k)/\mu_i^k \quad s \neq x_i^k, \quad p_i^k(x_i^k) = 1 - \sum_{x_i^k \neq s \in S_i} p_i^k(s)$$

$$\mu_i^k \geq k(m_i - 1) \max\{\, |u_i(x_i, x_{-i}) - u_i(x_i', x_{-i})| \ : \ x_i, x_i' \in S_i,\ x_{-i} \in S_{-i}\}$$

Approaching a larger class of equilibria

Empirical distribution $z_k : S \to \mathbb{R}$ provided by the process after k rounds:

$$z_k(x) = |\{t \leq k \ : \ x^t = x\}|/k$$

($z_k(x)$ = frequency of the strategy profile x in the first k rounds)

Approaching a larger class of equilibria

Empirical distribution $z_k : S \to \mathbb{R}$ provided by the process after k rounds:

$$z_k(x) = |\{t \leq k \ : \ x^t = x\}|/k$$

($z_k(x)$ = frequency of the strategy profile x in the first k rounds)

Convergence of the regret matching process

The sequence of empirical distributions z_k converges almost surely as $k \to +\infty$ to the set of correlated equilibria

Approaching a larger class of equilibria

Empirical distribution $z_k : S \to \mathbb{R}$ provided by the process after k rounds:

$$z_k(x) = |\{t \leq k \ : \ x^t = x\}|/k$$

($z_k(x)$ = frequency of the strategy profile x in the first k rounds)

Convergence of the regret matching process

The sequence of empirical distributions z_k converges almost surely as $k \to +\infty$ to the set of correlated equilibria

Correlated equilibria are probability distribution over S providing some conditions of equilibrium for strategies not necessarily "independent" of each other (Aumann 1974)

		q_1	q_2
	I/II	L	R
p_1	T	$p_1 q_1$	$p_1 q_2$
p_2	D	$p_2 q_1$	$p_2 q_2$

mixed strategies $\subsetneq$ correlation device

I/II	L	R
T	z_{11}	z_{12}
D	z_{21}	z_{22}

Approaching a larger class of equilibria

Empirical distribution $z_k : S \to \mathbb{R}$ provided by the process after k rounds:

$$z_k(x) = |\{t \leq k \ : \ x^t = x\}|/k$$

($z_k(x)$ = frequency of the strategy profile x in the first k rounds)

Convergence of the regret matching process

The sequence of empirical distributions z_k converges almost surely as $k \to +\infty$ to the set of correlated equilibria

Correlated equilibria are probability distribution over S providing some conditions of equilibrium for strategies not necessarily "independent" of each other (Aumann 1974)

$$\{\text{Nash equilibria in mixed strategies}\} \subseteq \{\text{correlated equilibria}\}$$

		q_1	q_2
	I/II	L	R
p_1	T	$p_1 q_1$	$p_1 q_2$
p_2	D	$p_2 q_1$	$p_2 q_2$

mixed strategies $\subsetneq$ correlation device

I/II	L	R
T	z_{11}	z_{12}
D	z_{21}	z_{22}

Correlation devices and correlated equilibria

Let $(N, (S_i)_{i \in N}, (u_i)_{i \in N})$ be a finite strategic game.

Definition

A correlation device is a probability measure/distribution over S, namely any $z : S \to \mathbb{R}_+$ such that
$$\sum_{x \in S} z(x) = 1.$$

Definition

A correlated equilibrium is a correlation device z such that any players $i \in N$ satisfies the incentive constraints
$$\sum_{x_{-i} \in S_{-i}} z(x_i, x_{-i}) \big(u_i(x_i, x_{-i}) - u_i(x'_i, x_{-i}) \big) \geq 0$$
for any $x_i, x'_i \in S_i$.

x_i is a best response to the mixed strategies of the other players induced by the correlation device provided that the pure strategy x_i is played (induced mixed strategy=conditional probability)

Mixed strategies and the geometry of correlated equilibria

Mixed strategies as correlated equilibria

Let $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ be a mixed strategy profile. Then, σ^* is a Nash equilibrium in mixed strategies if and only if the correlation device z given by $z(x) = \sigma^*(x_1) \ldots \sigma^*(x_n)$ is a correlated equilibrium.

(the incentive constraints collpase to the definition of Nash equilibrium)

Mixed strategies and the geometry of correlated equilibria

Mixed strategies as correlated equilibria

Let $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ be a mixed strategy profile. Then, σ^* is a Nash equilibrium in mixed strategies if and only if the correlation device z given by $z(x) = \sigma^*(x_1) \ldots \sigma^*(x_n)$ is a correlated equilibrium.

(the incentive constraints collpase to the definition of Nash equilibrium)

- The set *CE* of correlated equilibria is always a polytope $\rightarrow$ linear programming
- The convex hull of the set of Nash equilibria can be a proper subset of *CE*

Mixed strategies and the geometry of correlated equilibria

Mixed strategies as correlated equilibria

Let $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ be a mixed strategy profile. Then, σ^* is a Nash equilibrium in mixed strategies if and only if the correlation device z given by $z(x) = \sigma^*(x_1) \ldots \sigma^*(x_n)$ is a correlated equilibrium.

(the incentive constraints collpase to the definition of Nash equilibrium)

- The set *CE* of correlated equilibria is always a polytope $\rightarrow$ linear programming
- The convex hull of the set of Nash equilibria can be a proper subset of *CE*

he/she	football	dancing
football	(2,1)	(0,0)
dancing	(0,0)	(1,2)

Mixed strategies and the geometry of correlated equilibria

Mixed strategies as correlated equilibria

Let $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ be a mixed strategy profile. Then, σ^* is a Nash equilibrium in mixed strategies if and only if the correlation device z given by $z(x) = \sigma^*(x_1) \ldots \sigma^*(x_n)$ is a correlated equilibrium.

(the incentive constraints collpase to the definition of Nash equilibrium)

- The set *CE* of correlated equilibria is always a polytope → linear programming
- The convex hull of the set of Nash equilibria can be a proper subset of *CE*

he/she	football	dancing
football	(2,1)	(0,0)
dancing	(0,0)	(1,2)

h/s	F	D
F	1	0
D	0	0

h/s	F	D
F	0	0
D	0	1

		1/3	2/3
	h/s	F	D
2/3	F	2/9	4/9
1/3	D	1/9	2/9

← Nash equilibria →

h/s	F	D
F	1/4	0
D	1/4	1/2

correlated equilibrium