Algorithmic game theory

Martin Balko

7th lecture

November 18th 2025

Applications of regret minimization

Our notation

Our notation

- We have an agent A in an adversary environment.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X,

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$,

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \dots, N\}$.

- At each step $t = 1, \dots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \dots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \dots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^{N} p_i^t \ell_i^t$.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. This is the expected loss of A in step t.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. This is the expected loss of A in step t.

- After T steps, the cumulative loss of action i is $L_i^T = \sum_{t=1}^T \ell_i^t$.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. This is the expected loss of A in step t.

- After T steps, the cumulative loss of action i is $L_i^T = \sum_{t=1}^T \ell_i^t$.
- The cumulative loss of A is $L_A^T = \sum_{t=1}^T \ell_A^t$.

Our notation

- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. This is the expected loss of A in step t.

- After T steps, the cumulative loss of action i is $L_i^T = \sum_{t=1}^T \ell_i^t$.
- The cumulative loss of A is $L_A^T = \sum_{t=1}^T \ell_A^t$.
- Given a comparison class $\mathcal{A}_X$ of agents A_i that select a single action i in all steps, we let $L_{min}^T = \min_{i \in X}\{L_{A_i}^T\}$ be the minimum cumulative loss of an agent from $\mathcal{A}_X$.

Our notation

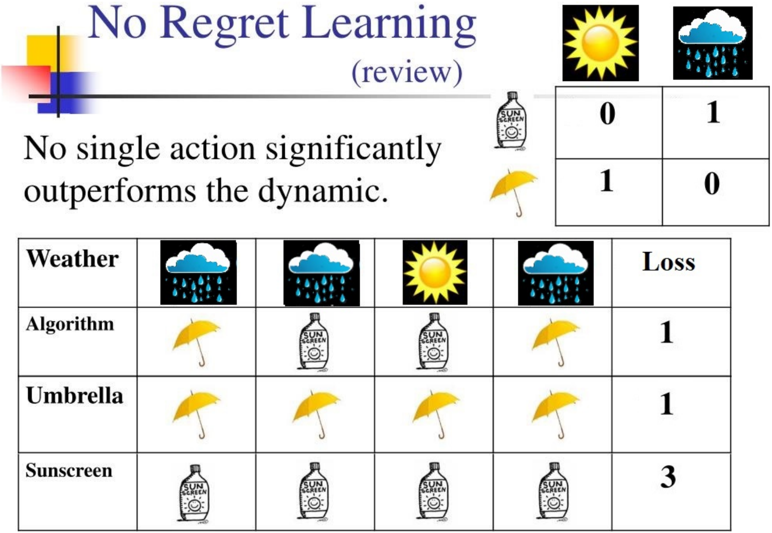
- We have an agent A in an adversary environment.
- There are N available actions for A in the set $X = \{1, \ldots, N\}$.

- At each step $t = 1, \ldots, T$:
  - Our agent A selects a probability distribution $p^t = (p_1^t, \ldots, p_N^t)$ over X, where p_i^t is the probability that A selects i in step t.
  - Then, the adversary chooses a loss vector $\ell^t = (\ell_1^t, \ldots, \ell_N^t)$, where $\ell_i^t \in [-1, 1]$ is the loss of action i in step t.
  - The agent A then experiences loss $\ell_A^t = \sum_{i=1}^N p_i^t \ell_i^t$. This is the expected loss of A in step t.

- After T steps, the cumulative loss of action i is $L_i^T = \sum_{t=1}^T \ell_i^t$.
- The cumulative loss of A is $L_A^T = \sum_{t=1}^T \ell_A^t$.
- Given a comparison class $\mathcal{A}_X$ of agents A_i that select a single action i in all steps, we let $L_{min}^T = \min_{i \in X}\{L_{A_i}^T\}$ be the minimum cumulative loss of an agent from $\mathcal{A}_X$.
- Our goal is to minimize the external regret $R_A^T = L_A^T - L_{min}^T$.

Example

Example


No Regret Learning
(review)
No single action significantly outperforms the dynamic.
0
1
1
0
Weather
Loss
Algorithm
1
Umbrella
1
Sunscreen
3

The Polynomial weights algorithm (PW algorithm)

The Polynomial weights algorithm (PW algorithm)

Algorithm 0.2: POLYNOMIAL WEIGHTS ALGORITHM(X, T, η)

Input : A set of actions $X = \{1, \ldots, N\}$, $T \in \mathbb{N}$, and $\eta \in (0, 1/2]$.
Output : A probability distribution p^t for every $t \in \{1, \ldots, T\}$.
$w_i^1 \leftarrow 1$ for every $i \in X$,
$p^1 \leftarrow (1/N, \ldots, 1/N)$,
for $t = 2, \ldots, T$

$$\textbf{do} \begin{cases} w_i^t \leftarrow w_i^{t-1}(1 - \eta \ell_i^{t-1}), \\ W^t \leftarrow \sum_{i \in X} w_i^t, \\ p_i^t \leftarrow w_i^t / W^t \text{ for every } i \in X. \end{cases}$$

Output $\{p^t : t \in \{1, \ldots, T\}\}$.

The Polynomial weights algorithm (PW algorithm)

Algorithm 0.3: POLYNOMIAL WEIGHTS ALGORITHM(X, T, η)

Input : A set of actions $X = \{1, \ldots, N\}$, $T \in \mathbb{N}$, and $\eta \in (0, 1/2]$.
Output : A probability distribution p^t for every $t \in \{1, \ldots, T\}$.
$w_i^1 \leftarrow 1$ for every $i \in X$,
$p^1 \leftarrow (1/N, \ldots, 1/N)$,
for $t = 2, \ldots, T$
do $\begin{cases} w_i^t \leftarrow w_i^{t-1}(1 - \eta \ell_i^{t-1}), \\ W^t \leftarrow \sum_{i \in X} w_i^t, \\ p_i^t \leftarrow w_i^t / W^t \text{ for every } i \in X. \end{cases}$
Output $\{p^t \colon t \in \{1, \ldots, T\}\}$.

- For any sequence of loss vectors, we have $R_{\mathrm{PW}}^T \leq 2\sqrt{T \ln N}$.

The Polynomial weights algorithm (PW algorithm)

Algorithm 0.4: POLYNOMIAL WEIGHTS ALGORITHM(X, T, η)

Input : A set of actions $X = \{1, \dots, N\}$, $T \in \mathbb{N}$, and $\eta \in (0, 1/2]$.
Output : A probability distribution p^t for every $t \in \{1, \dots, T\}$.
$w_i^1 \leftarrow 1$ for every $i \in X$,
$p^1 \leftarrow (1/N, \dots, 1/N)$,
for $t = 2, \dots, T$

do $\begin{cases} w_i^t \leftarrow w_i^{t-1}(1 - \eta \ell_i^{t-1}), \\ W^t \leftarrow \sum_{i \in X} w_i^t, \\ p_i^t \leftarrow w_i^t / W^t \text{ for every } i \in X. \end{cases}$

Output $\{p^t : t \in \{1, \dots, T\}\}$.

- For any sequence of loss vectors, we have $R^T_{\mathrm{PW}} \leq 2\sqrt{T \ln N}$.
- So the average regret $\frac{1}{T} \cdot R^T_{\mathrm{PW}}$ goes to 0 with $T \to \infty$.

Applications of regret minimization

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \ldots, C_n)$, where $C_i \colon A \to [-1, 1]$.

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \ldots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $= -$utility.

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \dots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $= -$utility.
- This will be done via the so-called No-regret dynamics:

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \ldots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $=$ $-$utility.
- This will be done via the so-called No-regret dynamics:
  - "Players play against each other by selecting actions according to an algorithm with small external regret."

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \ldots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $= -$utility.
- This will be done via the so-called No-regret dynamics:
  - "Players play against each other by selecting actions according to an algorithm with small external regret."
  - Each player $i \in P$ chooses a mixed strategy $p_i^t = (p_i^t(a_i))_{a_i \in A_i}$ using some algorithm with small external regret such that actions correspond to pure strategies.

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \ldots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $= -$utility.
- This will be done via the so-called No-regret dynamics:
  - "Players play against each other by selecting actions according to an algorithm with small external regret."
  - Each player $i \in P$ chooses a mixed strategy $p_i^t = (p_i^t(a_i))_{a_i \in A_i}$ using some algorithm with small external regret such that actions correspond to pure strategies.
  - Then, i receives a loss vector $\ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}$, where

$$\ell_i^t(a_i) = \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)]$$

for the product distribution $p_{-i}^t = \prod_{j \neq i} p_j^t$.

Applications of regret minimization

- Today, we will see how to apply regret minimization in the theory of normal-form games.
- Let $G = (P, A, C)$ be a normal-form game of n players with a cost function $C = (C_1, \dots, C_n)$, where $C_i \colon A \to [-1, 1]$. Cost $=$ $-$utility.
- This will be done via the so-called No-regret dynamics:
  - "Players play against each other by selecting actions according to an algorithm with small external regret."
  - Each player $i \in P$ chooses a mixed strategy $p_i^t = (p_i^t(a_i))_{a_i \in A_i}$ using some algorithm with small external regret such that actions correspond to pure strategies.
  - Then, i receives a loss vector $\ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}$, where

$$\ell_i^t(a_i) = \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)]$$

    for the product distribution $p_{-i}^t = \prod_{j \neq i} p_j^t$.
  - That is, $\ell_i^t(a_i)$ is the expected cost of the pure strategy a_i given the mixed strategies chosen by the other players.

The No-regret dynamics

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

Algorithm 0.7: NO-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

Algorithm 0.8: NO-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \dots, T\}$.
for every step $t = 1, \dots, T$

do $\Bigg\{$

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

Algorithm 0.9: NO-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.
for every step $t = 1, \ldots, T$
do { Each player $i \in P$ independently chooses a mixed strategy p_i^t using an algorithm with average regret at most ε, with actions corresponding to pure strategies.

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

Algorithm 0.10: NO-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.
for every step $t = 1, \ldots, T$
do
- Each player $i \in P$ independently chooses a mixed strategy p_i^t using an algorithm with average regret at most ε, with actions corresponding to pure strategies.
- Each player $i \in P$ receives a loss vector $\ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}$, where $\ell_i^t(a_i) \leftarrow \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)]$ for the product distribution $p_{-i}^t = \prod_{j \neq i} p_j^t$.

The No-regret dynamics

- "Players play against each other by selecting actions according to an algorithm with small external regret."

Algorithm 0.11: NO-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.
for every step $t = 1, \ldots, T$

do $\begin{cases} \text{Each player } i \in P \text{ independently chooses a mixed strategy } p_i^t \\ \text{using an algorithm with average regret at most } \varepsilon \text{, with actions} \\ \text{corresponding to pure strategies.} \\ \text{Each player } i \in P \text{ receives a loss vector } \ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i} \text{, where} \\ \ell_i^t(a_i) \leftarrow \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)] \text{ for the product distribution} \\ p_{-i}^t = \prod_{j \neq i} p_j^t. \end{cases}$

Output $\{p^t : t \in \{1, \ldots, T\}\}$.

Application: Modern proof of the Minimax Theorem

Application: Modern proof of the Minimax Theorem

- A new proof of the Minimax theorem.

Application: Modern proof of the Minimax Theorem

- A new proof of the Minimax theorem.
- A zero-sum game $G = (\{1, 2\}, A, C)$ with $A_1 = \{a_1, \dots, a_m\}$, $A_2 = \{b_1, \dots, b_n\}$ is represented with an $m \times n$ matrix M where $M_{i,j} = -C_1(a_i, b_j) = C_2(a_i, b_j) \in [-1, 1]$.

Application: Modern proof of the Minimax Theorem

- A new proof of the Minimax theorem.
- A zero-sum game $G = (\{1,2\}, A, C)$ with $A_1 = \{a_1, \ldots, a_m\}$, $A_2 = \{b_1, \ldots, b_n\}$ is represented with an $m \times n$ matrix M where $M_{i,j} = -C_1(a_i, b_j) = C_2(a_i, b_j) \in [-1, 1]$.
- The expected cost $C_2(s)$ for player 2 equals $x^\top My$, where x and y are the mixed strategy vectors.

Application: Modern proof of the Minimax Theorem

- A new proof of the Minimax theorem.
- A zero-sum game $G = (\{1, 2\}, A, C)$ with $A_1 = \{a_1, \ldots, a_m\}$, $A_2 = \{b_1, \ldots, b_n\}$ is represented with an $m \times n$ matrix M where $M_{i,j} = -C_1(a_i, b_j) = C_2(a_i, b_j) \in [-1, 1]$.
- The expected cost $C_2(s)$ for player 2 equals $x^\top My$, where x and y are the mixed strategy vectors.
- The Minimax theorem then states

$$\max_{x \in S_1} \min_{y \in S_2} x^\top My = \min_{y \in S_2} \max_{x \in S_1} x^\top My.$$

Source: https://www.privatdozent.co/

Application: Modern proof of the Minimax Theorem

- A new proof of the Minimax theorem.
- A zero-sum game $G = (\{1,2\}, A, C)$ with $A_1 = \{a_1, \ldots, a_m\}$, $A_2 = \{b_1, \ldots, b_n\}$ is represented with an $m \times n$ matrix M where $M_{i,j} = -C_1(a_i, b_j) = C_2(a_i, b_j) \in [-1, 1]$.
- The expected cost $C_2(s)$ for player 2 equals $x^\top My$, where x and y are the mixed strategy vectors.
- The Minimax theorem then states

$$\max_{x \in S_1} \min_{y \in S_2} x^\top My = \min_{y \in S_2} \max_{x \in S_1} x^\top My.$$

Source: https://www.privatdozent.co/

- We can prove it without LP!

Modern proof of the Minimax Theorem I

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily,

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0, 1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0, 1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4 \ln(\max\{m, n\})/\varepsilon^2$.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \le \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \ge \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0, 1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4 \ln(\max\{m, n\})/\varepsilon^2$.
- Let $p^1, \ldots, p^T$ and $q^1, \ldots, q^T$ be strategies played by players 1 and 2.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0, 1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4 \ln(\max\{m, n\})/\varepsilon^2$.
- Let $p^1, \dots, p^T$ and $q^1, \dots, q^T$ be strategies played by players 1 and 2.
- We let $\overline{x} = \frac{1}{T}\sum_{t=1}^{T} p^t$ and $\overline{y} = \frac{1}{T}\sum_{t=1}^{T} q^t$ be the time-averaged strategies of players 1 and 2.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0,1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4\ln(\max\{m,n\})/\varepsilon^2$.
- Let $p^1, \ldots, p^T$ and $q^1, \ldots, q^T$ be strategies played by players 1 and 2.
- We let $\overline{x} = \frac{1}{T}\sum_{t=1}^{T} p^t$ and $\overline{y} = \frac{1}{T}\sum_{t=1}^{T} q^t$ be the time-averaged strategies of players 1 and 2.
- The payoff vector revealed to each no-regret algorithm after step t is the expected payoff of each strategy, given the mixed strategy played by the other player.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0, 1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4\ln(\max\{m, n\})/\varepsilon^2$.
- Let $p^1, \dots, p^T$ and $q^1, \dots, q^T$ be strategies played by players 1 and 2.
- We let $\overline{x} = \frac{1}{T}\sum_{t=1}^T p^t$ and $\overline{y} = \frac{1}{T}\sum_{t=1}^T q^t$ be the time-averaged strategies of players 1 and 2.
- The payoff vector revealed to each no-regret algorithm after step t is the expected payoff of each strategy, given the mixed strategy played by the other player.
- Thus, players 1 and 2 get the payoff vectors Mq^t and $-(p^t)^\top M$.

Modern proof of the Minimax Theorem I

- First, the inequality $\max_x \min_y x^\top My \leq \min_y \max_x x^\top My$ follows easily, since it is only worse to go first.
- Second, we prove the inequality $\max_x \min_y x^\top My \geq \min_y \max_x x^\top My$.
- We choose a parameter $\varepsilon \in (0,1]$ and run the No-regret dynamics for a sufficient number T of steps so that both players have average expected external regret at most ε.
- With the PW algorithm, we can set $T = 4\ln(\max\{m,n\})/\varepsilon^2$.
- Let $p^1,\dots,p^T$ and $q^1,\dots,q^T$ be strategies played by players 1 and 2.
- We let $\overline{x} = \frac{1}{T}\sum_{t=1}^{T} p^t$ and $\overline{y} = \frac{1}{T}\sum_{t=1}^{T} q^t$ be the time-averaged strategies of players 1 and 2.
- The payoff vector revealed to each no-regret algorithm after step t is the expected payoff of each strategy, given the mixed strategy played by the other player.
- Thus, players 1 and 2 get the payoff vectors Mq^t and $-(p^t)^\top M$.
- The time-averaged expected payoff of 1 is then $v = \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top Mq^t$.

Modern proof of the Minimax Theorem II

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i.

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M \overline{y}$$

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top M q^t$$

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top M q^t \leq \frac{1}{T}\sum_{t=1}^{T} (p^t)^\top M q^t + \varepsilon$$

Modern proof of the Minimax Theorem II

- For $i = 1, \dots, m$, let $e_i = (0, \dots, 0, 1, 0 \dots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top Mq^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top Mq^t + \varepsilon = v + \varepsilon.$$

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top Mq^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top Mq^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$.

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top Mq^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top Mq^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$. Analogously, $(\overline{x})^\top My \geq v - \varepsilon$ for every $y \in S_2$.

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top M q^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top M q^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$. Analogously, $(\overline{x})^\top My \geq v - \varepsilon$ for every $y \in S_2$.
- Putting everything together, we get

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top M q^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top M q^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$. Analogously, $(\overline{x})^\top My \geq v - \varepsilon$ for every $y \in S_2$.
- Putting everything together, we get

$$\begin{aligned}\max_{x\in S_1}\min_{y\in S_2} x^\top My &\geq \min_{y\in S_2}(\overline{x})^\top My \geq v - \varepsilon \\ &\geq \max_{x\in S_1} x^\top M\overline{y} - 2\varepsilon \geq \min_{y\in S_2}\max_{x\in S_1} x^\top My - 2\varepsilon.\end{aligned}$$

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top M q^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top M q^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$. Analogously, $(\overline{x})^\top My \geq v - \varepsilon$ for every $y \in S_2$.
- Putting everything together, we get

$$\begin{aligned}\max_{x\in S_1}\min_{y\in S_2} x^\top My &\geq \min_{y\in S_2}(\overline{x})^\top My \geq v - \varepsilon \\ &\geq \max_{x\in S_1} x^\top M\overline{y} - 2\varepsilon \geq \min_{y\in S_2}\max_{x\in S_1} x^\top My - 2\varepsilon.\end{aligned}$$

- For $T \to \infty$, we get $\varepsilon \to 0$ and we obtain the desired inequality.

Modern proof of the Minimax Theorem II

- For $i = 1, \ldots, m$, let $e_i = (0, \ldots, 0, 1, 0 \ldots, 0)$ be the mixed strategy vector for the pure strategy a_i. Since the external regret of player 1 is at most ε, we have

$$e_i^\top M\overline{y} = \frac{1}{T}\sum_{t=1}^{T} e_i^\top Mq^t \leq \frac{1}{T}\sum_{t=1}^{T}(p^t)^\top Mq^t + \varepsilon = v + \varepsilon.$$

- Since every strategy $x \in S_1$ is a convex combination of the vectors e_i, the linearity of expectation gives $x^\top M\overline{y} \leq v + \varepsilon$. Analogously, $(\overline{x})^\top My \geq v - \varepsilon$ for every $y \in S_2$.
- Putting everything together, we get

$$\begin{aligned}\max_{x\in S_1}\min_{y\in S_2} x^\top My &\geq \min_{y\in S_2}(\overline{x})^\top My \geq v - \varepsilon \\ &\geq \max_{x\in S_1} x^\top M\overline{y} - 2\varepsilon \geq \min_{y\in S_2}\max_{x\in S_1} x^\top My - 2\varepsilon.\end{aligned}$$

- For $T \to \infty$, we get $\varepsilon \to 0$ and we obtain the desired inequality. $\square$

Application: Coarse correlated equilibria

Application: Coarse correlated equilibria

- Recall: a prob. distribution p on A is a correlated equilibrium (CE) if

$$\sum_{a_{-i} \in A_{-i}} C_i(a_i; a_{-i}) p(a_i; a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a_i'; a_{-i}) p(a_i; a_{-i})$$

  for every player $i \in P$ and all pure strategies $a_i, a_i' \in A_i$.

Application: Coarse correlated equilibria

- Recall: a prob. distribution p on A is a correlated equilibrium (CE) if

$$\sum_{a_{-i} \in A_{-i}} C_i(a_i; a_{-i}) p(a_i; a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a_i'; a_{-i}) p(a_i; a_{-i})$$

for every player $i \in P$ and all pure strategies $a_i, a_i' \in A_i$.

  - In other words,

$$\mathbb{E}_{a \sim p}[C_i(a) \mid a_i] \leq \mathbb{E}_{a \sim p}[C_i(a_i'; a_{-i}) \mid a_i].$$

Application: Coarse correlated equilibria

- Recall: a prob. distribution p on A is a correlated equilibrium (CE) if

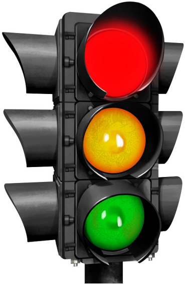
$$\sum_{a_{-i} \in A_{-i}} C_i(a_i; a_{-i}) p(a_i; a_{-i}) \leq \sum_{a_{-i} \in A_{-i}} C_i(a_i'; a_{-i}) p(a_i; a_{-i})$$

for every player $i \in P$ and all pure strategies $a_i, a_i' \in A_i$.

- In other words,

$$\mathbb{E}_{a \sim p}[C_i(a) \mid a_i] \leq \mathbb{E}_{a \sim p}[C_i(a_i'; a_{-i}) \mid a_i].$$

- We define an even more tractable concept and use no-regret dynamics to converge to it.



Coarse correlated equilibrium

Coarse correlated equilibrium

- For a normal-form game $G = (P, A, C)$ of n players, a probability distribution p on A is a coarse correlated equilibrium (CCE) in G if

Coarse correlated equilibrium

- For a normal-form game $G = (P, A, C)$ of n players, a probability distribution p on A is a coarse correlated equilibrium (CCE) in G if

$$\sum_{a \in A} C_i(a)p(a) \leq \sum_{a \in A} C_i(a'_i; a_{-i})p(a)$$

for every player $i \in P$ and every $a'_i \in A_i$.

Coarse correlated equilibrium

- For a normal-form game $G = (P, A, C)$ of n players, a probability distribution p on A is a coarse correlated equilibrium (CCE) in G if

$$\sum_{a \in A} C_i(a)p(a) \leq \sum_{a \in A} C_i(a'_i; a_{-i})p(a)$$

for every player $i \in P$ and every $a'_i \in A_i$.

- CCE can be expressed as

$$\mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(a'_i; a_{-i})]$$

for every $i \in P$ and each $a'_i \in A_i$.

Coarse correlated equilibrium

- For a normal-form game $G = (P, A, C)$ of n players, a probability distribution p on A is a coarse correlated equilibrium (CCE) in G if

$$\sum_{a \in A} C_i(a)p(a) \leq \sum_{a \in A} C_i(a'_i; a_{-i})p(a)$$

  for every player $i \in P$ and every $a'_i \in A_i$.
- CCE can be expressed as

$$\mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(a'_i; a_{-i})]$$

  for every $i \in P$ and each $a'_i \in A_i$.
- The difference between CCE and CE is that CCE only requires that following your suggested action a_i when a is drawn from p is only a best response in expectation before you see a_i.

Coarse correlated equilibrium

- For a normal-form game $G = (P, A, C)$ of n players, a probability distribution p on A is a coarse correlated equilibrium (CCE) in G if

$$\sum_{a \in A} C_i(a)p(a) \leq \sum_{a \in A} C_i(a'_i; a_{-i})p(a)$$

  for every player $i \in P$ and every $a'_i \in A_i$.
- CCE can be expressed as

$$\mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(a'_i; a_{-i})]$$

  for every $i \in P$ and each $a'_i \in A_i$.
- The difference between CCE and CE is that CCE only requires that following your suggested action a_i when a is drawn from p is only a best response in expectation before you see a_i. This makes sense if you have to commit to following your suggested action or not upfront, and do not have the opportunity to deviate after seeing it.

Example: Coarse correlated equilibrium

Example: Coarse correlated equilibrium

- Giving probability $1/6$ to each red outcome gives coarse correlated equilibrium in the Rock-Paper-Scissors game.

	Rock	Paper	Scissors
Rock	(0,0)	(-1,1)	(1,-1)
Paper	(1,-1)	(0,0)	(-1,1)
Scissors	(-1,1)	(1,-1)	(0,0)

Example: Coarse correlated equilibrium

- Giving probability $1/6$ to each red outcome gives coarse correlated equilibrium in the Rock-Paper-Scissors game.

	Rock	Paper	Scissors
Rock	(0,0)	(-1,1)	(1,-1)
Paper	(1,-1)	(0,0)	(-1,1)
Scissors	(-1,1)	(1,-1)	(0,0)

- Then, the expected payoff of each player is 0 and deviating to any pure strategy gives the expected payoff 0.

Example: Coarse correlated equilibrium

- Giving probability $1/6$ to each red outcome gives coarse correlated equilibrium in the Rock-Paper-Scissors game.

	Rock	Paper	Scissors
Rock	(0,0)	(-1,1)	(1,-1)
Paper	(1,-1)	(0,0)	(-1,1)
Scissors	(-1,1)	(1,-1)	(0,0)

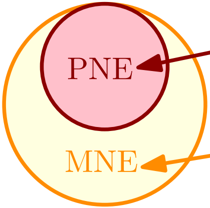
- Then, the expected payoff of each player is 0 and deviating to any pure strategy gives the expected payoff 0.
- It is not a correlated equilibrium though.

Hierarchy of Nash equilibria

Hierarchy of Nash equilibria

Pure Nash equilibria,
not always exist

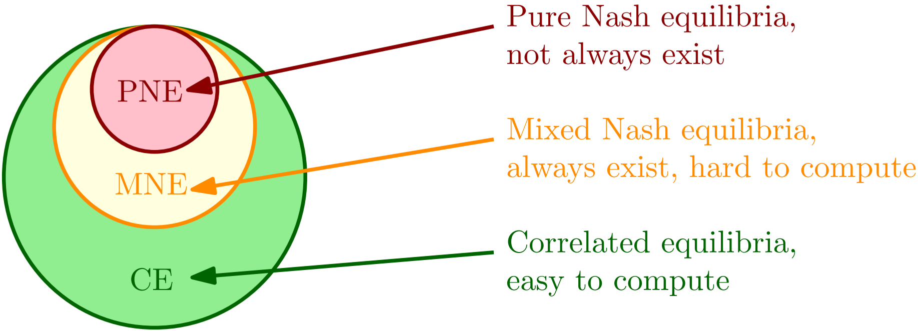
Hierarchy of Nash equilibria



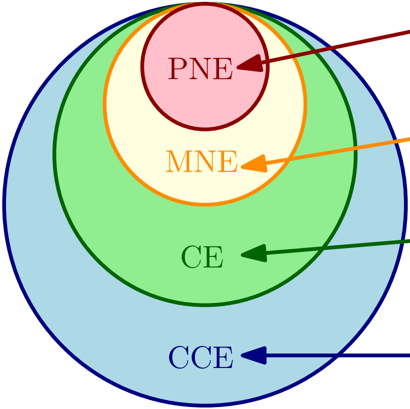
Pure Nash equilibria,
not always exist

Mixed Nash equilibria,
always exist, hard to compute

Hierarchy of Nash equilibria



Hierarchy of Nash equilibria



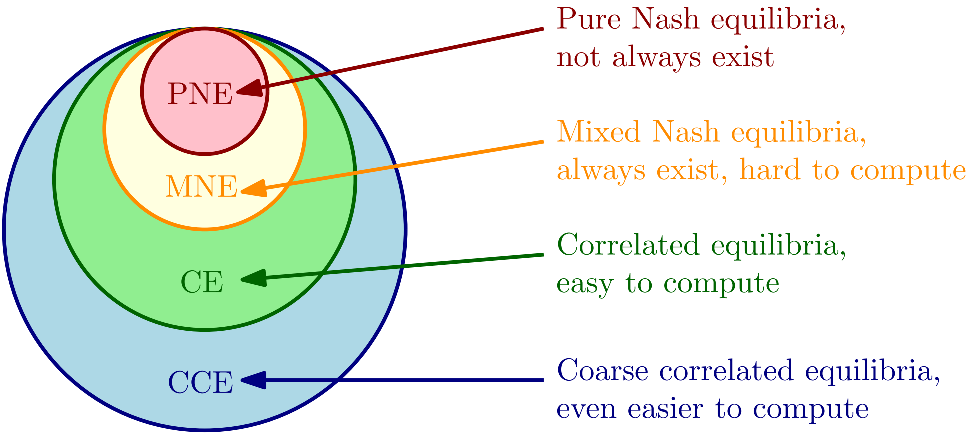
Pure Nash equilibria,
not always exist

Mixed Nash equilibria,
always exist, hard to compute

Correlated equilibria,
easy to compute

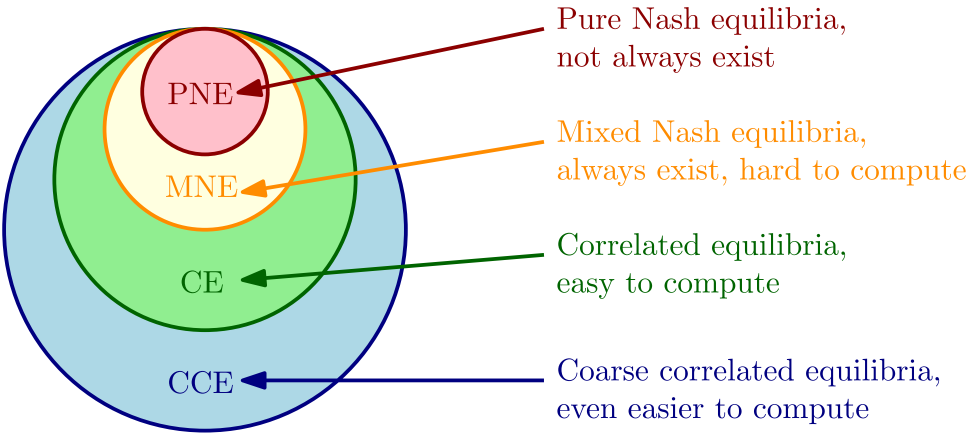
Coarse correlated equilibria,
even easier to compute

Hierarchy of Nash equilibria



- In general normal-form game, no-regret dynamics converges to a coarse correlated equilibrium.

Hierarchy of Nash equilibria



- In general normal-form game, no-regret dynamics converges to a coarse correlated equilibrium.
- For $\varepsilon > 0$, a probability distribution p on A is an ε-coarse correlated equilibrium (ε-CCE) if $\mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.

Converging to CCE

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

- Proof:

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T} \sum_{t=1}^{T} p^t$.

- Proof: We want to prove $\mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)]$$

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^n p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^T p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

- The right-hand sides are time-averaged expected costs of i when playing according to the algorithm with small external regret and when playing a_i' every iteration.

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

- The right-hand sides are time-averaged expected costs of i when playing according to the algorithm with small external regret and when playing a_i' every iteration. Since every player has regret at most ε, we obtain

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^n p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^T p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

- The right-hand sides are time-averaged expected costs of i when playing according to the algorithm with small external regret and when playing a_i' every iteration. Since every player has regret at most ε, we obtain

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{a\sim p^t}[C_i(a)] \leq \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{a\sim p^t}[C_i(a_i'; a_{-i})] + \varepsilon.$$

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^n p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^T p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \le \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

- The right-hand sides are time-averaged expected costs of i when playing according to the algorithm with small external regret and when playing a_i' every iteration. Since every player has regret at most ε, we obtain

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a\sim p^t}[C_i(a)] \le \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a\sim p^t}[C_i(a_i'; a_{-i})] + \varepsilon.$$

- This verifies the ε-CCE condition for $p = \frac{1}{T}\sum_{t=1}^T p_t$.

Converging to CCE

Theorem 2.54

For every $G = (P, A, C)$, $\varepsilon > 0$, and $T = T(\varepsilon) \in \mathbb{N}$, if after T steps of the No-regret dynamics, each player $i \in P$ has time-averaged expected regret at most ε, then p is ε-CCE where $p^t = \prod_{i=1}^{n} p_i^t$ and $p = \frac{1}{T}\sum_{t=1}^{T} p^t$.

- Proof: We want to prove $\mathbb{E}_{a\sim p}[C_i(a)] \leq \mathbb{E}_{a\sim p}[C_i(a_i'; a_{-i})] + \varepsilon$.
- By the definition of p, we have, for every player $i \in P$ and $a_i' \in A_i$,

$$\mathop{\mathbb{E}}_{a\sim p}[C_i(a)] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a)] \text{ and } \mathop{\mathbb{E}}_{a\sim p}[C_i(a_i'; a_{-i})] = \frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{a\sim p^t}[C_i(a_i'; a_{-i})].$$

- The right-hand sides are time-averaged expected costs of i when playing according to the algorithm with small external regret and when playing a_i' every iteration. Since every player has regret at most ε, we obtain

$$\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a\sim p^t}[C_i(a)] \leq \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{a\sim p^t}[C_i(a_i'; a_{-i})] + \varepsilon.$$

- This verifies the ε-CCE condition for $p = \frac{1}{T}\sum_{t=1}^{T} p_t$. $\square$

Other notions of regret

Other notions of regret

- Converging to CCE is nice, but how about converging to CE?

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an “internal setting” when we compare our agent to its modifications.

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an "internal setting" when we compare our agent to its modifications.

- A modification rule is a function $F : X \to X$.

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an "internal setting" when we compare our agent to its modifications.

- A modification rule is a function $F \colon X \to X$.
- We modify a sequence $(p^t)_{t=1}^T$ with F by replacing it with a sequence $(f^t)_{t=1}^T$, where $f^t = (f_1^t, \ldots, f_N^t)$ and $f_i^t = \sum_{j\colon F(j)=i} p_j^t$.

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an "internal setting" when we compare our agent to its modifications.

- A modification rule is a function $F\colon X \to X$.
- We modify a sequence $(p^t)_{t=1}^T$ with F by replacing it with a sequence $(f^t)_{t=1}^T$, where $f^t = (f_1^t, \ldots, f_N^t)$ and $f_i^t = \sum_{j\colon F(j)=i} p_j^t$.
  - "The modified agent plays $F(i)$ whenever A plays i."

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an "internal setting" when we compare our agent to its modifications.

- A modification rule is a function $F\colon X \to X$.
- We modify a sequence $(p^t)_{t=1}^T$ with F by replacing it with a sequence $(f^t)_{t=1}^T$, where $f^t = (f_1^t, \ldots, f_N^t)$ and $f_i^t = \sum_{j\colon F(j)=i} p_j^t$.
  - "The modified agent plays $F(i)$ whenever A plays i."
- The cumulative loss of A modified by F is $L_{A,F}^T = \sum_{t=1}^T \sum_{i=1}^N f_i^t \ell_i^t$.

Other notions of regret

- Converging to CCE is nice, but how about converging to CE? We can do that with a different notion of regret!
- We consider an "internal setting" when we compare our agent to its modifications.

- A modification rule is a function $F\colon X \to X$.
- We modify a sequence $(p^t)_{t=1}^T$ with F by replacing it with a sequence $(f^t)_{t=1}^T$, where $f^t = (f_1^t, \ldots, f_N^t)$ and $f_i^t = \sum_{j\colon F(j)=i} p_j^t$.
  - "The modified agent plays $F(i)$ whenever A plays i."
- The cumulative loss of A modified by F is $L_{A,F}^T = \sum_{t=1}^T \sum_{i=1}^N f_i^t \ell_i^t$.

- Given a set of modification rules $\mathcal{F}$, we can compare our agent to his modifications by rules from $\mathcal{F}$, obtaining different notions of regret.

Internal and swap regret

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\}$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p^t_i \ell^t_i \right) - \ell^t_j \right) \right\}.$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p^t_i \ell^t_i \right) - \ell^t_j \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p^t_i \ell^t_i \right) - \ell^t_j \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \left\{ L^T_A - L^T_{A,F} \right\}$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p_i^t \ell_i^t \right) - \ell_j^t \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{i,j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{L^T_A - L^T_{A,F}\right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p_i^t \ell_i^t \right) - \ell_j^t \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \left\{L^T_A - L^T_{A,F}\right\} = \max_{i,j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

- For the set $\mathcal{F}^{sw}$ of all modification rules, we get the swap regret:

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p^t_i \ell^t_i \right) - \ell^t_j \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \left\{ L^T_A - L^T_{A,F} \right\} = \max_{i,j \in X} \left\{ \sum_{t=1}^{T} p^t_i (\ell^t_i - \ell^t_j) \right\}.$$

- For the set $\mathcal{F}^{sw}$ of all modification rules, we get the swap regret:

$$R^T_{A,\mathcal{F}^{sw}} = \max_{F \in \mathcal{F}^{sw}} \left\{ L^T_A - L^T_{A,F} \right\}$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \left\{L^T_A - L^T_{A,F}\right\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p_i^t \ell_i^t \right) - \ell_j^t \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \left\{L^T_A - L^T_{A,F}\right\} = \max_{i,j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

- For the set $\mathcal{F}^{sw}$ of all modification rules, we get the swap regret:

$$R^T_{A,\mathcal{F}^{sw}} = \max_{F \in \mathcal{F}^{sw}} \left\{L^T_A - L^T_{A,F}\right\} = \sum_{i=1}^{N} \max_{j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

Internal and swap regret

- For a set $\mathcal{F}^{ex} = \{F_i \colon i \in X\}$ of rules where F_i always outputs action i, we obtain exactly the external regret:

$$R^T_{A,\mathcal{F}^{ex}} = \max_{F \in \mathcal{F}^{ex}} \{L^T_A - L^T_{A,F}\} = \max_{j \in X} \left\{ \sum_{t=1}^{T} \left(\left(\sum_{i \in X} p_i^t \ell_i^t \right) - \ell_j^t \right) \right\}.$$

- For $\mathcal{F}^{in} = \{F_{i,j} \colon (i,j) \in X \times X, i \neq j\}$ where $F_{i,j}$ is defined by $F_{i,j}(i) = j$ and $F_{i,j}(i') = i'$ for each $i' \neq i$, we get the internal regret:

$$R^T_{A,\mathcal{F}^{in}} = \max_{F \in \mathcal{F}^{in}} \{L^T_A - L^T_{A,F}\} = \max_{i,j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

- For the set $\mathcal{F}^{sw}$ of all modification rules, we get the swap regret:

$$R^T_{A,\mathcal{F}^{sw}} = \max_{F \in \mathcal{F}^{sw}} \{L^T_A - L^T_{A,F}\} = \sum_{i=1}^{N} \max_{j \in X} \left\{ \sum_{t=1}^{T} p_i^t (\ell_i^t - \ell_j^t) \right\}.$$

- Since $\mathcal{F}^{ex}, \mathcal{F}^{in} \subseteq \mathcal{F}^{sw}$, we immediately have $R^T_{A,\mathcal{F}^{ex}}, R^T_{A,\mathcal{F}^{in}} \leq R^T_{A,\mathcal{F}^{sw}}$.

The No-swap-regret dynamics

The No-swap-regret dynamics

- Using swap regret instead of external regret, we will get:

The No-swap-regret dynamics

- Using swap regret instead of external regret, we will get:

Algorithm 0.14: NO-SWAP-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.
for every step $t = 1, \ldots, T$
do { Each player $i \in P$ independently chooses a mixed strategy p_i^t using an algorithm with average swap regret at most ε, with actions corresponding to pure strategies.
Each player $i \in P$ receives a loss vector $\ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}$, where $\ell_i^t(a_i) \leftarrow \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)]$ for the product distribution $p_{-i}^t = \prod_{j \neq i} p_j^t$.
Output $\{p^t : t \in \{1, \ldots, T\}\}$.

The No-swap-regret dynamics

- Using swap regret instead of external regret, we will get:

Algorithm 0.15: NO-SWAP-REGRET DYNAMICS(G, T, ε)

Input : A normal-form game $G = (P, A, C)$ of n players, $T \in \mathbb{N}$, and $\varepsilon > 0$.
Output : A prob. distribution p_i^t on A_i for each $i \in P$ and $t \in \{1, \ldots, T\}$.
for every step $t = 1, \ldots, T$
do Each player $i \in P$ independently chooses a mixed strategy p_i^t using an algorithm with average swap regret at most ε, with actions corresponding to pure strategies.
Each player $i \in P$ receives a loss vector $\ell_i^t = (\ell_i^t(a_i))_{a_i \in A_i}$, where $\ell_i^t(a_i) \leftarrow \mathbb{E}_{a_{-i}^t \sim p_{-i}^t}[C_i(a_i; a_{-i}^t)]$ for the product distribution $p_{-i}^t = \prod_{j \neq i} p_j^t$.
Output $\{p^t : t \in \{1, \ldots, T\}\}$.

- No-swap-regret dynamics then converges to a correlated equilibrium.


WHEN YOU FIND THE NASH EQUILIBRIUM
NO RAGRETS
memegenerator.net

Thank you for your attention.