Alternation makes the adversary weaker

Stratis Skoulakis

joint work with

V. Cevher, EPFL

A. Cutkosky, BU

G. Piliouras, DeepMind

A. Kavis, UT Austin

L. Viano, EPFL

Normal Form Games

Rock-Paper-Scissors

Payoﬀ matrix of Alice

Alice/Bob	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Payoﬀ matrix of Bob

Alice/Bob	Rock	Paper	Scissors
Rock	0	1	-1
Paper	-1	0	1
Scissors	1	-1	0

Normal Form Games

Rock-Paper-Scissors

Alice plays Rock

Alice / Bob	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Alice gets -1

Bob plays Scissors

Alice / Bob	Rock	Paper	Scissors
Rock	0	1	-1
Paper	-1	0	1
Scissors	1	-1	0

Bob gets 1

Normal Form Games

Normal-Form Games

Matrices A and B

Payoff matrix A of Alice

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

Payoff matrix B of Bob

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Normal Form Games

Normal-Form Games

Matrices A and B

Alice plays action i

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{1j}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{nj}	$\cdots$	A_{nm}

Alice gets A_{ij}

Bob plays action j

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{1j}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{nj}	$\cdots$	B_{nm}

Bob gets B_{ij}

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \quad \Delta_n$

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \quad \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \quad \Delta_n$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

Alice's expected cost

$x \quad Ay$

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \quad \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

Bob's expected cost

$x \quad By$

Normal Form Games

Normal-Form Games - Mixed Strategies

Matrices A and B

Alice plays a prob. distr.

$x = (x_1, \ldots, x_n) \quad \Delta_n$

A/B	1	$\cdots$	j	$\cdots$	m
1	A_{11}	$\cdots$	A_{12}	$\cdots$	A_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	A_{i1}	$\cdots$	A_{ij}	$\cdots$	A_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	A_{n1}	$\cdots$	A_{n2}	$\cdots$	A_{nm}

$$x^\top A y$$

Bob plays a prob. distr.

$y = (y_1, \ldots, y_m) \quad \Delta_m$

A/B	1	$\cdots$	j	$\cdots$	m
1	B_{11}	$\cdots$	B_{12}	$\cdots$	B_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
i	B_{i1}	$\cdots$	B_{ij}	$\cdots$	B_{im}
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
n	B_{n1}	$\cdots$	B_{n2}	$\cdots$	B_{nm}

$$x^\top B y$$

Normal-Form Games over Time

- Alice and Bob play the normal-form game (A, B) over T rounds.
  - *Simultaneous play*: Agents simultaneously select their strategies at each round (very well-studied).
  - *Alternating play*: Agents alternatingly update their strategies (this work).

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\dashrightarrow$ Alice gets $x_1 A y_1$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\dashrightarrow$ Alice gets $x_1 A y_1$
- Alice selects x_2 / Bob selects y_2

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\longrightarrow$ Alice gets $x_1 A y_1$
- Alice selects x_2 / Bob selects y_2 $\longrightarrow$ Alice gets $x_2 A y_2$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play $(A; B)$ for T rounds.

- Alice selects x_1 / Bob selects y_1 ! Alice gets $x_1^{>} A y_1$
- Alice selects x_2 / Bob selects y_2 ! Alice gets $x_2^{>} A y_2$
- Alice selects x_3 / Bob selects y_3

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Alice selects x_3 / Bob selects y_3 $\to$ Alice gets $x_3^\top A y_3$

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$

Alice selects x_3 / Bob selects y_3 $\rightarrow$ Alice gets $x_3^\top A y_3$

How shoud Alice select her actions over time?

Simultaneous Play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects y_3 $\rightarrow$ Alice gets $x_3^\top A y_3$

How shoud Alice select her actions over time? No Regret algorithms

- Regret Matching [Blackwell '65]
- Hedge [Freund et al. '97]
- Online Gradient Descent [Zinkevic '03]
- Follow the Regularized Leader [Abernethy et al. '10]
- ....

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play $(A; B)$ for T rounds.

- Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Hedge Algorithm [Freund and Schapire '97] Godel Prize '03

$$x_{t+1}(i) = \frac{x_t(i)\, e^{[Ay_t]_i}}{\sum_{j=1}^{n} x_t(j)\, e^{[Ay_t]_j}}$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play $(A; B)$ for T rounds.

Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$

Hedge Algorithm [Freund and Schapire '97] Godel Prize '03

$$x_{t+1}(i) = \frac{x_t(i) \cdot e^{-\eta [A y_t]_i}}{\sum_{j=1}^{n} x_t(j) \cdot e^{-\eta [A y_t]_j}}$$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $R(T)$

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \leq O\left(\sqrt{T}\right)$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$
- Alice selects x_3 / Bob selects y_3 $\rightarrow$ Alice gets $x_3^\top A y_3$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies,

$$\underbrace{\frac{1}{T}\sum_{t=1}^{T} x_t^\top A y_t}_{\text{time-average cost}} \le \underbrace{\frac{1}{T}\min_{x \in \Delta_n}\sum_{t=1}^{T} x^\top A y_t}_{\text{best fixed action}} + O\left(\frac{\sqrt{T}}{T}\right) \rightarrow 0$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $R(T)$

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \approx O\left(\sqrt{T}\right)$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \dots, y_T$, the regret of Alice $R(T)$

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \leq O\left(\sqrt{T}\right)$$

Can Alice do better?

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$

Theorem (Freund and Schapire JCSS '97)

No matter Bob's strategies $y_1, \ldots, y_T$, the regret of Alice $R(T)$

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \leq \tilde{O}\left(\sqrt{T}\right)$$

Can Alice do better? No!

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$

Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Theorem (folkore)

Bob can always select $y_1, \ldots, y_T$, the regret of Alice

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \geq \Omega(\sqrt{T})$$

$O(\sqrt{T})$ regret is the best Alice can get in simultaneous play!

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\to$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\to$ Alice gets $x_2^\top A y_2$

Theorem (folkore)

Bob can always select $y_1, \ldots, y_T$, the regret of Alice

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \geq \Omega(\sqrt{T})$$

Simultaneous play

Simultaneous Play

Bob challenges Alice to play (A, B) for T rounds.

- Alice selects x_1 / Bob selects y_1 $\rightarrow$ Alice gets $x_1^\top A y_1$
- Alice selects x_2 / Bob selects y_2 $\rightarrow$ Alice gets $x_2^\top A y_2$

Theorem (folkore)

Bob can always select $y_1, \ldots, y_T$, the regret of Alice

$$R(T) := \underbrace{\sum_{t=1}^{T} x_t^\top A y_t}_{\text{cost of Alice}} - \underbrace{\min_{i \in [n]} \sum_{t=1}^{T} [A y_t]_i}_{\text{cost of best action}} \geq \Omega(\sqrt{T})$$

What if Alice and Bob play in alternating turns?

Solving Heads'up Poker

Online Learning + Simultaneous Play
Polaris [Bowling et al. AAMAS '09] Decent Performance

Solving Heads'up Poker

Online Learning + Simultaneous Play

Polaris [Bowling et al. AAMAS '09]! Decent Performance

Online Learning + Alternating Play

Cepheus [Oskari et al. IJCAI '15], Libratus [Brown et al. IJCAI '17]! Beat Human Experts!

Solving Heads'up Poker

Online Learning + Simultaneous Play

Polaris [Bowling et al. AAMAS '09]! Decent Performance

Online Learning + Alternating Play

Cepheus [Oskari et al. IJCAI '15], Libratus [Brown et al. IJCAI '17]! Beat Human Experts!

Regret Guarantees ! Faster Training!

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns.

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns.

Bob selects y_0

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 ! Alice gets $x_1^\top A y_0$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 $\to$ Alice gets $x_1^\top A y_0$
- Bob selects y_2

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 $\rightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects y_2 $\rightarrow$ Alice gets $x_1^\top A y_2$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 $\rightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects y_2 $\rightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects x_3

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A;B)$ for T rounds in alternating turns .

- Bob selects y_0
- Alice selects x_1 ! Alice gets $x_1^{>}Ay_0$
- Bob selects y_2 ! Alice gets $x_1^{>}Ay_2$
- Alice selects x_3 ! Alice gets $x_3^{>}Ay_2$

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns .

- Bob selects y_0
- Alice selects x_1 ! Alice gets $x_1^{>} A y_0$
- Bob selects y_2 ! Alice gets $x_1^{>} A y_2$
- Alice selects x_3 ! Alice gets $x_3^{>} A y_2$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 → Alice gets $x_1^\top A y_0$
- Bob selects y_2 → Alice gets $x_1^\top A y_2$
- Alice selects x_3 → Alice gets $x_3^\top A y_2$

Can Alice have regret $R(T)$ better than $O(\sqrt{T})$?

$$R(T) := \underbrace{\sum_{k=0}^{T/2} x_{2k+1}^\top A(y_{2k} + y_{2k+2})}_{\text{Alice's cost}} - \underbrace{\min_{x \in \Delta_n} \sum_{k=0}^{T/2} x^\top A(y_{2k} + y_{2k+2})}_{\text{best fixed action}}$$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in alternating turns.

- Bob selects y_0
- Alice selects $x_1 \to$ Alice gets $x_1^\top A y_0$
- Bob selects $y_2 \to$ Alice gets $x_1^\top A y_2$
- Alice selects $x_3 \to$ Alice gets $x_3^\top A y_2$

Theorem ([Cevher, Cutksovsky*, Piliouras, Skoulakis*, Viano NeurIPS '23 spotlight])

In alternating play, Alice can always guarantee

- $\tilde{O}(T^{1/3})$ regret for general games.
- $O(\log T)$ for $n = 2$ actions (different algorithm).

Alternating Turns

Previous Results

- If both Alice and Bob use Gradient Descent in unconstrained zero-sum games $\rightarrow O(1)$ regret [Bailey et al. COLT 2020]
- If both Alice and Bob use Hedge in zero-sum games $\rightarrow O(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Alternating Turns

Previous Results

- If both Alice and Bob use Gradient Descent in unconstrained zero-sum games $\rightarrow O(1)$ regret [Bailey et al. COLT 2020]
- If both Alice and Bob use Hedge in zero-sum games $\rightarrow O(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Previous results make assumptions on Bob's behavior.

Alternating Turns

Previous Results

- If both Alice and Bob use Gradient Descent in unconstrained zero-sum games $\to O(1)$ regret [Bailey et al. COLT 2020]
- If both Alice and Bob use Hedge in zero-sum games $\to O(T^{1/3})$ regret [Wibisono et al. NeurIPS 2022]

Previous results make assumptions on Bob's behavior.

Theorem ([Cevher, Cutksovsky[?], Piliouras, Skoulakis[?], Viano NeurIPS '23 spotlight])

In alternating play , Alice can always guarantee

- $\tilde{O}(T^{1/3})$ regret for general games.
- $O(\log T)$ for $n = 2$ actions (different algorithm).

no matter Bob's behavior!

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns**.

- Bob selects y_0
- Alice selects x_1 $\rightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects y_2 $\rightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects x_3 $\rightarrow$ Alice gets $x_3^\top A y_2$

	Simultaneous Play	Alternating Play
General Games	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(T^{1/3})$
2 actions for Alice	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(\log T)$

Alternating Turns

Alternating Play

Bob challenges Alice to play $(A; B)$ for T rounds in alternating turns.

- Bob selects y_0
- Alice selects x_1 $\rightarrow$ Alice gets $x_1^\top A y_0$
- Bob selects y_2 $\rightarrow$ Alice gets $x_1^\top A y_2$
- Alice selects x_3 $\rightarrow$ Alice gets $x_3^\top A y_2$

Our algorithm (general simplex)

$$x_{2k+1} \in \operatorname{argmin}_{x \in \Delta_n} \left[\underbrace{3\, x^\top A y_{2k}}_{\text{exploits } y_{2k}} + \underbrace{2 \sum_{k'=0}^{k-1} x^\top A y_{2k'}}_{\text{reinforces good past actions}} - \underbrace{\sum_{i=1}^{n} \log x_i}_{\text{prevents overfitting}} \right]$$

Alternating Turns

Alternating Play

Bob challenges Alice to play (A, B) for T rounds in **alternating turns.**

- Bob selects y_0
- Alice selects x_1 $\longrightarrow$ Alice gets $x_1 A y_0$
- Bob selects y_2 $\longrightarrow$ Alice gets $x_1 A y_2$
- Alice selects x_3 $\longrightarrow$ Alice gets $x_3 A y_2$
- . . .

Our algorithm (general simplex)

$$x_{2k+1} \leftarrow \operatorname{argmin}_{x \in \Delta_n} \left\{ \underbrace{3\, x^\top A y_{2k}}_{\text{exploits } y_{2k}} + \underbrace{2 \sum_{k'=0}^{k-1} x^\top A y_{2k'}}_{\text{reinforces good past actions}} - \underbrace{y \cdot \sum_{i=1}^{n} \log x_i}_{\text{prevents overfitting}} \right\}$$

Setting $y = O(T^{1/3})$ $\rightarrow$ $O(T^{1/3})$ alternating regret

2 action simplex

- If Alice admits 2 actions $\rightarrow O(\log T)$ in alternating play!

2 action simplex

- If Alice admits 2 actions $\rightarrow O(\log T)$ in alternating play!
- At round $t = 2k + 1$

2 action simplex

- If Alice admits 2 actions → $O(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response → $w_k = \arg\min_{x \in \Delta_n} \langle x, Ay_{2k} \rangle$

2 action simplex

- If Alice admits 2 actions → $O(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response → $w_k = \arg\min_{x \in \Delta_n} \left(x^\top A y_{2k} \right)$
  - Follow the Regularized Leader →

$$z_k = \arg\min_{x \in \Delta_n} \left(\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2 \right)$$

2 action simplex

- If Alice admits 2 actions → $O(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response → $w_k = \arg\min_{x \in \Delta_n} \left(x^\top A y_{2k} \right)$
  - Follow the Regularized Leader →

$$z_k = \arg\min_{x \in \Delta_n} \left(\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2 \right)$$

  - If FTRL (z_k) admits $O(\sqrt{T})$ regret → Greedy BR (w_k) admits $O(1)$ regret

2 action simplex

- If Alice admits 2 actions → $O(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response → $w_k = \arg\min_{x \in \Delta_n} \left(x^\top A y_{2k} \right)$
  - Follow the Regularized Leader →

$$z_k = \arg\min_{x \in \Delta_n} \left(\sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2 \right)$$

  - If FTRL (z_k) admits $O(\sqrt{T})$ regret → Greedy BR (w_k) admits $O(1)$ regret
  - If Greedy BR (w_k) admits $O(\sqrt{T})$ regret → FTRL (z_k) admits $O(1)$ regret

2 action simplex

- If Alice admits 2 actions → $O(\log T)$ in alternating play!
- At round $t = 2k + 1$
  - Greedy Best Response → $w_k = \arg\min_{x \in \Delta_n} \left\{ x^\top A y_{2k} \right\}$
  - Follow the Regularized Leader →
    $$z_k = \arg\min_{x \in \Delta_n} \left\{ \sum_{k=1}^{t/2} x^\top A y_{2k} + \gamma ||x||^2 \right\}$$
  - If FTRL (z_k) admits $O(\sqrt{T})$ regret → Greedy BR (w_k) admits $O(1)$ regret
  - If Greedy BR (w_k) admits $O(\sqrt{T})$ regret → FTRL (z_k) admits $O(1)$ regret
- Alice plays x_{2k+1} in the convex hull of z_k and w_k.

Take-Away Message

Algorithms beyond $O(\sqrt{T})$ regret lower bounds of simultaneous play.

	Simultaneous Play	Alternating Play
General Games	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(T^{1/3})$
2 actions for Alice	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(\log T)$

Take-Away Message

Algorithms beyond $O(\sqrt{T})$ regret lower bounds of simultaneous play.

	Simultaneous Play	Alternating Play
General Games	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(T^{1/3})$
2 actions for Alice	$\tilde{O}(\sqrt{T})$ tight	$\tilde{O}(\log T)$

Thank you!!

References