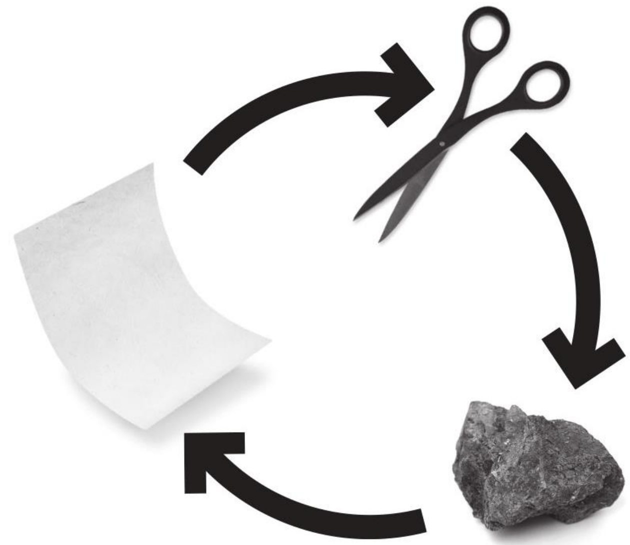
Learning Dynamics for Nash and Coarse Correlated Equilibria in Bimatrix Games

Ioannis Panageas (UC Irvine)

Playing Rock-Paper-Scissors

			
	0, 0	-1, 1	1,-1
	1,-1	0, 0	-1, 1
	-1, 1	1, -1	0, 0



Solution concepts

			
	0, 0	-1, 1	1,-1
	1,-1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Nash Equilibrium (NE)

- No incentive to unilaterally deviate
- Each agent throws her own coins (decentralized).

E.g., $(1/3, 1/3, 1/3)$

Solution concepts

			
	0, 0	-1, 1	1,-1
	1,-1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Nash Equilibrium (NE)

- No incentive to unilaterally deviate
- Each agent throws her own coins. (decentralized).

E.g., $(1/3, 1/3, 1/3)$.

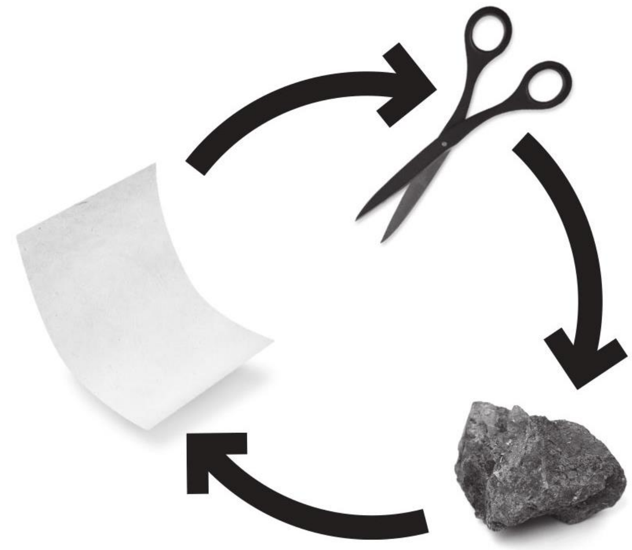
Coarse Correlated Equilibrium (CCE)

- No incentive to unilaterally deviate
- All agents use same coins. (centralized)

E.g., (R,P), (R,S), (P,R), (P,S), (S,R), (S,P) with probability $1/6$.

Playing Rock-Paper-Scissors

0, 0	-1, 1	1,-1
1,-1	0, 0	-1, 1
-1, 1	1, -1	0, 0



Let’s repeat the game multiple times

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities $x_t \in \Delta_n$.*
- *Gets expected utility $u_t(x_t)$.*

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities* $x_t \in \Delta_n$.
- *Gets expected utility* $u_t(x_t)$.

Players' choices induce a process (dynamics).
Behaviors include:

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities* $x_t \in \Delta_n$.
- *Gets expected utility* $u_t(x_t)$.

Players' choices induce a process (dynamics).
Behaviors include:

- Convergence of time average

([Robinson51], [Freund-Schapire99], … a lot of works)

$$\frac{\sum_{t=1}^{T} x_t}{T}$$ close to solution of interest.

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities* $x_t \in \Delta_n$.
- *Gets expected utility* $u_t(x_t)$.

Players' choices induce a process (dynamics).
Behaviors include:

- Best iterate or Last iterate convergence

([Daskalakis-Ilyas-Syrgkanis-Zeng17], [Daskalakis-Panageas19], [Mertikopoulos-Lecouat-Zenati-Foo-Chandrasekhar-Piliouras19], [Anagnostides-Panageas-Farina-Sandholm22], …)

$$\exists t^* \in [T] \text{ s.t } x_{t^*} \text{ or } x_T$$

close to solution of interest.

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities* $x_t \in \Delta_n$.
- *Gets expected utility* $u_t(x_t)$.

Players' choices induce a process (dynamics).

Behaviors include:

- Cycling or recurrent behavior

([Bailey-Piliouras18], [Mertikopoulos-Papadimitriou-Piliouras18], [Mai-Mihail-Panageas-Ratcliff-Vazirani-Yunker18], …)

- Even chaotic ([Palaiopanos-Panageas-Piliouras17], …)

From [PP16]

From [SFG18]

Players use learning dynamics

Full feedback. *At each time step t.*

- *Each player chooses probabilities* $x_t \in \Delta_n$.
- *Gets expected utility* $u_t(x_t)$.

Players' choices induce a process (dynamics).
Behaviors include:

Many applications in Game Theory, Optimization, Machine Learning (GANs), even in evolution.

Behaviors appear in two player zero-sum or identical payoff.

No-regret learning dynamics

- Aim at minimizing regret.

Regret.

$$\mathrm{Reg} := \max_{\mathbf{x}^* \in \Delta_n} \left\{ \sum_{t=1}^{T} \mathbf{u}_t(\mathbf{x}^*) \right\} - \sum_{t=1}^{T} \mathbf{u}_t(\mathbf{x}_t).$$

No-regret learning dynamics

- Aim at minimizing regret.

Regret.

$$\text{Reg} := \max_{\mathbf{x}^* \in \Delta_n} \left\{ \sum_{t=1}^{T} \mathbf{u}_t(\mathbf{x}^*) \right\} - \sum_{t=1}^{T} \mathbf{u}_t(\mathbf{x}_t).$$

No-regret learning algorithms (make it sublinear):

- Multiplicative Weights Update (MWU), Online Gradient Descent (GD)
- Follow-The-Regularized-Leader, Follow-The-Perturbed-Leader etc

2-player zero-sum games

Two-player zero-sum. *Player* $\boldsymbol{y}$ *gets payoff* $\boldsymbol{x}^\top A\boldsymbol{y}$ *and* $\boldsymbol{x}$ *gets* $-\boldsymbol{x}^\top A\boldsymbol{y}$. *A Nash equilibrium is a solution to*

$$\min_{\boldsymbol{x}\in\Delta_n}\max_{\boldsymbol{y}\in\Delta_m} \boldsymbol{x}^\top A\boldsymbol{y}.$$

Rock-Paper-Scissors.

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

NE is $(1/3, 1/3, 1/3)$ for both players.

2-player zero-sum games

Two-player zero-sum. *Player* $\boldsymbol{y}$ *gets payoff* $\boldsymbol{x}^\top A\boldsymbol{y}$ *and* $\boldsymbol{x}$ *gets* $-\boldsymbol{x}^\top A\boldsymbol{y}$. *A Nash equilibrium is a solution to*

$$\min_{\boldsymbol{x}\in\Delta_n}\max_{\boldsymbol{y}\in\Delta_m} \boldsymbol{x}^\top A\boldsymbol{y}.$$

Rock-Paper-Scissors.

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

NE is $(1/3, 1/3, 1/3)$ for both players.

Time average converges to a NE when players update according to

- MWUA, GDA
- More generally FTRL, FTPL ...

2-player zero-sum games

Two-player zero-sum. *Player* $\boldsymbol{y}$ *gets payoff* $\boldsymbol{x}^\top A\boldsymbol{y}$ *and* $\boldsymbol{x}$ *gets* $-\boldsymbol{x}^\top A\boldsymbol{y}$. *A Nash equilibrium is a solution to*

$$\min_{\boldsymbol{x}\in\Delta_n}\max_{\boldsymbol{y}\in\Delta_m} \boldsymbol{x}^\top A\boldsymbol{y}.$$

Rock-Paper-Scissors.

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

NE is $(1/3, 1/3, 1/3)$ for both players.

Time average converges to a NE when players update according to

- MWUA, GDA
- More generally FTRL, FTPL ...

Why?

2-player zero-sum games (cont.)

- No-regret learning algorithms converge to CCEs in general games.

2-player zero-sum games (cont.)

- No-regret learning algorithms converge to CCEs in general games.
- Phenomenon of equilibrium collapse → Marginals of CCEs are NE in 2-player zero-sum games.

			
	0, 0	-1, 1	1,-1
	1,-1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Recall CCE: (R,P), (R,S), (P,R), (P,S), (S,R), (S,P) with probability 1/6. Marginalizing yields NE.

2-player zero-sum games (cont.)

- No-regret learning algorithms converge to CCEs in general games.
- Phenomenon of equilibrium collapse → Marginals of CCEs are NE in 2-player zero-sum games.

Do we need to take averages though?

2-player zero-sum games (cont.)

- No-regret learning algorithms converge to CCEs in general games.
- Phenomenon of equilibrium collapse ⇒ Marginals of CCEs are NE in 2-player zero-sum games.

Do we need to take averages though?

For MWUA and more generally FTRL dynamics YES.
[Mertikopoulos-Papadimitriou-Piliouras18]

2-player zero-sum games (cont.)

- No-regret learning algorithms converge to CCEs in general games.
- Phenomenon of equilibrium collapse ➡ Marginals of CCEs are NE in 2-player zero-sum games.

Do we need to take averages though?

For MWUA and more generally FTRL dynamics YES.
[Mertikopoulos-Papadimitriou-Piliouras18]

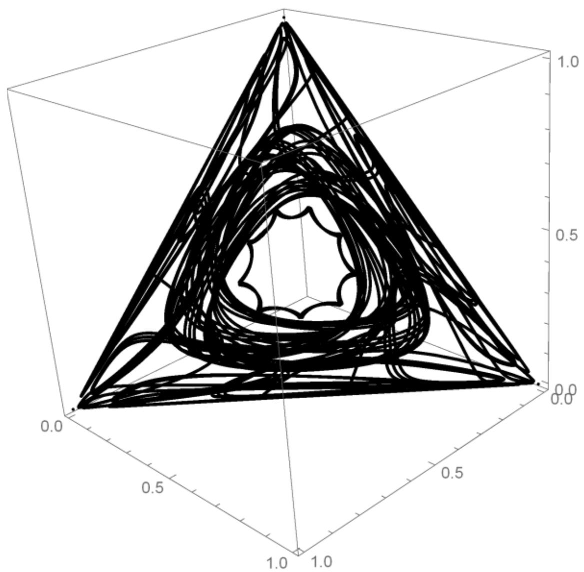
If A changes with time, cycles persist?

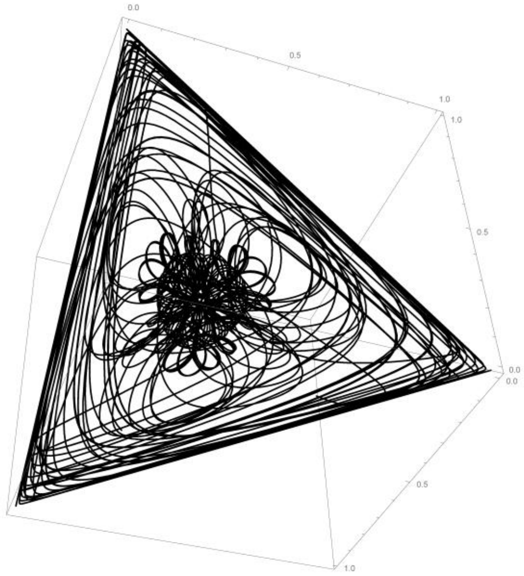
In [Mai-Mihail-Panageas-Ratcliff-Vazirani-Yunker18] we show recurrent behavior for a biological model in which the species update according to replicator dynamics.

2-player zero-sum games (cont.)

If A changes with time, cycles persist?

In [Mai-Mihail-Panageas-Ratcliff-Vazirani-Yunker18] we show recurrent behavior for a biological model in which the species update according to replicator dynamics.





Toy example

GDA for $f(x, y) = xy$

$$x_{t+1} = x_t - \eta_t y_t,$$
$$y_{t+1} = y_t + \eta_t x_t.$$

Can $(0, 0)$ be reached?

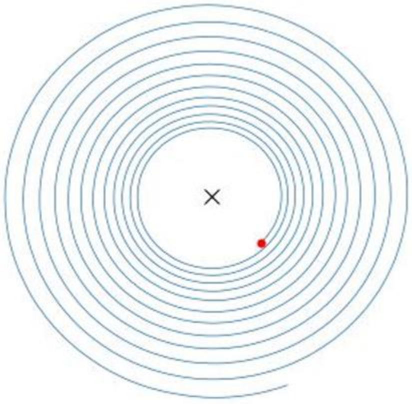
Toy example

GDA for $f(x, y) = xy$

$$x_{t+1} = x_t - \eta_t y_t,$$
$$y_{t+1} = y_t + \eta_t x_t.$$

Can $(0, 0)$ be reached?

No since $x_{t+1}^2 + y_{t+1}^2 = (1 + \eta_t^2) \cdot (x_t^2 + y_t^2)$, i.e., norm is expanding.



Optimism avoids cycles

GDA for $f(x,y) = xy$

$$x_{t+1} = x_t - \eta_t y_t,$$
$$y_{t+1} = y_t + \eta_t x_t.$$

Can $(0,0)$ be reached?

No since $x_{t+1}^2 + y_{t+1}^2 = (1+\eta_t^2)\cdot(x_t^2 + y_t^2)$, i.e., norm is expanding.

Can fix this behavior?

We can use "optimism" (negative momentum).

$$x_{t+1} = x_t - \eta \cdot \nabla_x f(x_t, y_t) + \boldsymbol{\eta/2 \cdot \nabla_x f(x_{t-1}, y_{t-1})}$$

$$y_{t+1} = y_t + \eta \cdot \nabla_y f(x_t, y_t) - \boldsymbol{\eta/2 \cdot \nabla_y f(x_{t-1}, y_{t-1})}$$

~~$x_{t+1} = x_t - \eta \cdot \nabla_x f(x_t, y_t)$~~

~~$y_{t+1} = y_t + \eta \cdot \nabla_y f(x_t, y_t)$~~

Optimism avoids cycles (cont.)

Introduced by Popov in the 80s.

$$x_{t+1} = x_t - \eta \cdot \nabla_x f(x_t, y_t) + \boldsymbol{\eta/2 \cdot \nabla_x f(x_{t-1}, y_{t-1})}$$

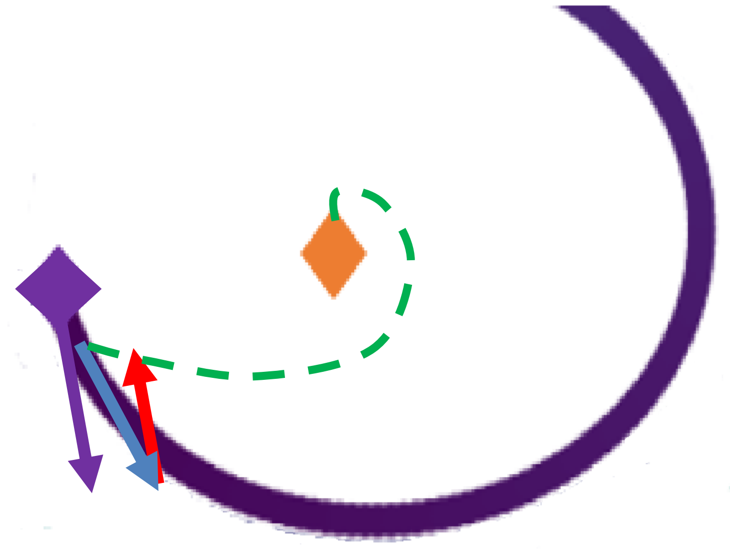
$$y_{t+1} = y_t + \eta \cdot \nabla_y f(x_t, y_t) - \boldsymbol{\eta/2 \cdot \nabla_y f(x_{t-1}, y_{t-1})}$$

Optimism avoids cycles (cont.)

Introduced by Popov in the 80s.

$$x_{t+1} = x_t - \eta \cdot \nabla_x f(x_t, y_t) + \boldsymbol{\eta/2 \cdot \nabla_x f(x_{t-1}, y_{t-1})}$$

$$y_{t+1} = y_t + \eta \cdot \nabla_y f(x_t, y_t) - \boldsymbol{\eta/2 \cdot \nabla_y f(x_{t-1}, y_{t-1})}$$

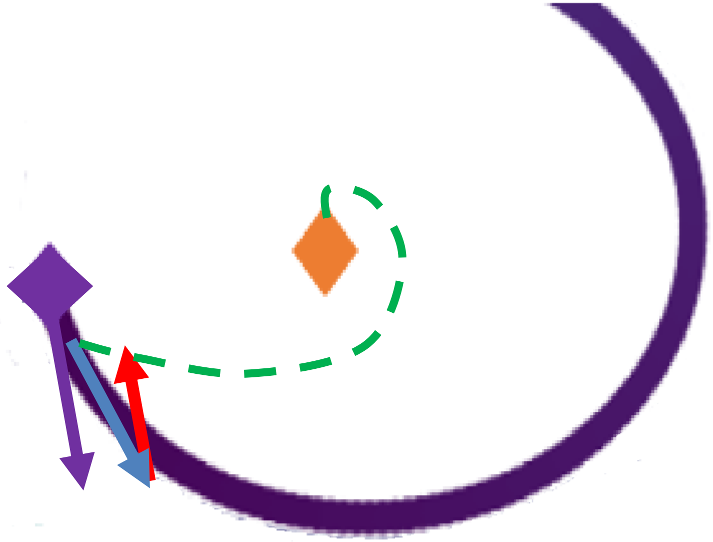


Optimism avoids cycles (cont.)

Introduced by Popov in the 80s.

$$x_{t+1} = x_t - \eta \cdot \nabla_x f(x_t, y_t) + \boldsymbol{\eta/2 \cdot \nabla_x f(x_{t-1}, y_{t-1})}$$

$$y_{t+1} = y_t + \eta \cdot \nabla_y f(x_t, y_t) - \boldsymbol{\eta/2 \cdot \nabla_y f(x_{t-1}, y_{t-1})}$$



[Daskalakis-Ilyas-Syrgkanis-Zeng17], [Daskalakis-Panageas19], [Mertikopoulos-Lecouat-Zenati-Foo-Chandrasekhar-Piliouras19], [Wei-Lee-Zhang-Luo21], [Golowich-Pattathil-Daskalakis21], [Anagnostides-Panageas-Farina-Sandholm22], [Cai-Oikonomou-Zheng22], [Diakonikolas-Daskalakis-Jordan22], and many more.

Use optimism for convergence to NE in general-sum games?

General Sum Bimatrix Games

Player $\mathbf{x}$ gets payoff $\mathbf{x}^\top A\mathbf{y}$ and $\mathbf{y}$ gets $\mathbf{x}^\top B\mathbf{y}$. A Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies the variational inequalities

$$\mathbf{x}^{*\top} A\mathbf{y}^* \geq \mathbf{x}^\top A\mathbf{y}^* \text{ for all } \mathbf{x} \in \Delta_n$$
$$\mathbf{x}^{*\top} B\mathbf{y}^* \geq \mathbf{x}^{*\top} B\mathbf{y} \text{ for all } \mathbf{y} \in \Delta_m$$

General Sum Bimatrix Games

Player $\mathbf{x}$ gets payoff $\mathbf{x}^\top A\mathbf{y}$ and $\mathbf{y}$ gets $\mathbf{x}^\top B\mathbf{y}$. A Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies the variational inequalities

$$\mathbf{x}^{*\top} A\mathbf{y}^* \geq \mathbf{x}^\top A\mathbf{y}^* \text{ for all } \mathbf{x} \in \Delta_n$$
$$\mathbf{x}^{*\top} B\mathbf{y}^* \geq \mathbf{x}^{*\top} B\mathbf{y} \text{ for all } \mathbf{y} \in \Delta_m$$

Equilibrium collapse does not hold! Avoid cycles?

General Sum Bimatrix Games

Player $\mathbf{x}$ gets payoff $\mathbf{x}^\top A\mathbf{y}$ and $\mathbf{y}$ gets $\mathbf{x}^\top B\mathbf{y}$. A Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies the variational inequalities

$$\mathbf{x}^{*\top} A\mathbf{y}^* \geq \mathbf{x}^\top A\mathbf{y}^* \text{ for all } \mathbf{x} \in \Delta_n$$
$$\mathbf{x}^{*\top} B\mathbf{y}^* \geq \mathbf{x}^{*\top} B\mathbf{y} \text{ for all } \mathbf{y} \in \Delta_m$$

Equilibrium collapse does not hold! Avoid cycles?

No, last/best-iterate is inherently tied with NE.

Let's add some structure then…

General Sum Bimatrix Games

Player $\mathbf{x}$ gets payoff $\mathbf{x}^\top A\mathbf{y}$ and $\mathbf{y}$ gets $\mathbf{x}^\top B\mathbf{y}$. A Nash equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies the variational inequalities

$$\mathbf{x}^{*\top} A\mathbf{y}^* \geq \mathbf{x}^\top A\mathbf{y}^* \text{ for all } \mathbf{x} \in \Delta_n$$
$$\mathbf{x}^{*\top} B\mathbf{y}^* \geq \mathbf{x}^{*\top} B\mathbf{y} \text{ for all } \mathbf{y} \in \Delta_m$$

Equilibrium collapse does not hold! Avoid cycles?

No, last/best-iterate is inherently tied with NE.

Let's add some structure then…

Rank-k. *The payoff matrices A, B are such that $\text{rank}(A + B) = k$.*

Learning in rank-k games

Rank-k. *The payoff matrices A, B are such that rank$(A+B) = k$.*

- Rank-0 are zero-sum games.
- Induces an hierarchy of games.
- Rank-1 games are in P [Adsul, Garg, Mehta, Sohoni, von Stengel18].
- Finding approximate NE for constant k is tractable [Kannan-Theobald05] but exact is PPAD-hard even for $k = 3$ [Mehta14] .

Learning in rank-k games

Rank-k. *The payoff matrices A, B are such that rank$(A + B) = k$.*

- Rank-0 are zero-sum games.
- Induces an hierarchy of games.
- Rank-1 games are in P [Adsul, Garg, Mehta, Sohoni, von Stengel18].
- Finding approximate NE for constant k is tractable [Kannan-Theobald05] but exact is PPAD-hard even for $k = 3$ [Mehta14] .

Theorem (Patris, Panageas 23). *Let (A, B) be a rank-1 game. There exists a modification of optimistic GDA such that when run for $O\left(\frac{1}{\epsilon^2}\log(\frac{1}{\epsilon}) \cdot (\log n + \log m)\right)$ iterations, it returns an ϵ-approximate Nash equilibrium of (A, B).*

Learning in rank-1 games

Definition (λ-parametrized zero-sum)**.** *Let* $A + B = ab^\top$. *Zero-sum game with payoff* $U_\lambda := A - \lambda \mathbf{1} b^\top$ *is called* *λ-parametrized*.

Learning in rank-1 games

Definition (λ-parametrized zero-sum). *Let* $A + B = ab^\top$. *Zero-sum game with payoff* $U_\lambda := A - \lambda \mathbf{1} b^\top$ *is called* *λ-parametrized*.

Lemma (λ-parametrized zero-sum). *Let* $A + B = ab^\top$. *If* $(\boldsymbol{x}, \boldsymbol{y})$ *is an* $O(\epsilon)$*-NE for the game* (A, B) *then there exists a* λ *such that*

- $(\boldsymbol{x}, \boldsymbol{y})$ *is a* $O(\epsilon)$*-NE of the* *λ-parametrized* *zero-sum game with payoff* U_λ.
- $\left|\boldsymbol{x}^\top a - \lambda\right|$ *is* $O(\epsilon)$.

Learning in rank-1 games

Lemma (λ-parametrized zero-sum). *Let $A + B = ab^\top$. If $(\boldsymbol{x}, \boldsymbol{y})$ is an $O(\epsilon)$-NE for the game (A, B) then there exists a λ such that*

- *$(\boldsymbol{x}, \boldsymbol{y})$ is a $O(\epsilon)$-NE of the λ-parametrized zero-sum game with payoff U_λ.*
- *$|\boldsymbol{x}^\top a - \lambda|$ is $O(\epsilon)$.*

Main idea: (similar to [Adsul, Garg, Mehta, Sohoni, von Stengel18])

Run optimistic GDA on the time-varying game $\mathbf{x}^\top \left(A - \lambda_t \mathbf{1} b^\top\right) \mathbf{y} + (\mathbf{x}^\top a - \lambda_t)^2$.

Learning in rank-1 games

Lemma (λ-parametrized zero-sum). *Let $A + B = ab^\top$. If $(\boldsymbol{x}, \boldsymbol{y})$ is an $O(\epsilon)$-NE for the game (A, B) then there exists a λ such that*

- $(\boldsymbol{x}, \boldsymbol{y})$ *is a $O(\epsilon)$-NE of the λ-parametrized zero-sum game with payoff U_λ.*
- $\left|\boldsymbol{x}^\top a - \lambda\right|$ *is $O(\epsilon)$.*

Main idea: (similar to [Adsul, Garg, Mehta, Sohoni, von Stengel18])

Run optimistic GDA on the time-varying game $\mathbf{x}^\top \left(A - \lambda_t \mathbf{1} b^\top\right) \mathbf{y} + (\mathbf{x}^\top a - \lambda_t)^2$.

Update λ_t so that it gets closer to $x_t^\top a$.

Learning in rank-1 games

Lemma (λ-parametrized zero-sum). *Let $A + B = ab^\top$. If $(\boldsymbol{x}, \boldsymbol{y})$ is an $O(\epsilon)$-NE for the game (A, B) then there exists a λ such that*

- *$(\boldsymbol{x}, \boldsymbol{y})$ is a $O(\epsilon)$-NE of the λ-parametrized zero-sum game with payoff U_λ.*
- *$|\boldsymbol{x}^\top a - \lambda|$ is $O(\epsilon)$.*

Main idea: (similar to [Adsul, Garg, Mehta, Sohoni, von Stengel18])

Run optimistic GDA on the time-varying game $\mathbf{x}^\top \left(A - \lambda_t \mathbf{1} b^\top\right) \mathbf{y} + (\mathbf{x}^\top a - \lambda_t)^2$.

Update λ_t so that it gets closer to $x_t^\top a$.

Learning in rank-1 games

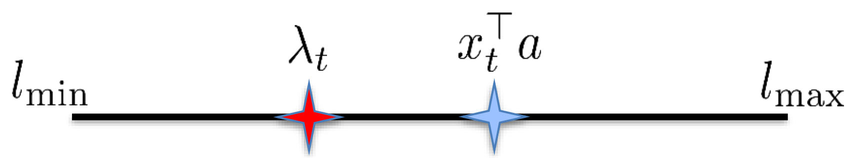
Lemma (λ-parametrized zero-sum)**.** *Let* $A + B = ab^\top$. *If* $(\boldsymbol{x}, \boldsymbol{y})$ *is an* $O(\epsilon)$*-NE for the game* (A, B) *then there exists a* λ *such that*

- $(\boldsymbol{x}, \boldsymbol{y})$ *is a* $O(\epsilon)$*-NE of the* λ*-parametrized zero-sum game with payoff* U_λ.
- $|\boldsymbol{x}^\top a - \lambda|$ *is* $O(\epsilon)$.

Main idea: (similar to [Adsul, Garg, Mehta, Sohoni, von Stengel18])

Run optimistic GDA on the time-varying game $\mathbf{x}^\top \left(A - \lambda_t \mathbf{1} b^\top\right) \mathbf{y} + (\mathbf{x}^\top a - \lambda_t)^2$.

Update λ_t so that it gets closer to $x_t^\top a$.



Learning in rank-1 games

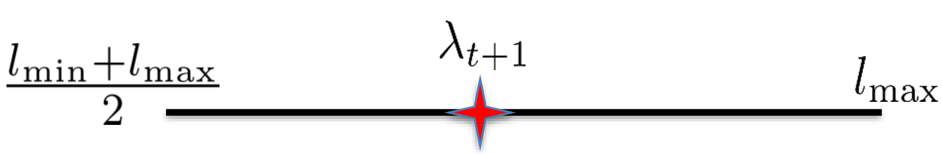
Lemma (λ-parametrized zero-sum). *Let $A + B = ab^\top$. If $(\boldsymbol{x}, \boldsymbol{y})$ is an $O(\epsilon)$-NE for the game (A, B) then there exists a λ such that*

- *$(\boldsymbol{x}, \boldsymbol{y})$ is a $O(\epsilon)$-NE of the λ-parametrized zero-sum game with payoff U_λ.*
- *$|\boldsymbol{x}^\top a - \lambda|$ is $O(\epsilon)$.*

Main idea: (similar to [Adsul, Garg, Mehta, Sohoni, von Stengel18])

Run optimistic GDA on the time-varying game $\mathbf{x}^\top \left(A - \lambda_t \mathbf{1} b^\top\right) \mathbf{y} + (\mathbf{x}^\top a - \lambda_t)^2$.

Update λ_t so that it gets closer to $x_t^\top a$.



Learning in constant rank games

Open Question. *Can we generalize for higher ranks? The main challenge is the update of λ. Maybe consider other parametrization?*

Learning in constant rank games

Open Question. *Can we generalize for higher ranks? The main challenge is the update of λ. Maybe consider other parametrization?*

Open Question. *Get learning algorithms for other classes of bimatrix games?*

Remark:

- We focus on rank games because the computation is tractable.
- Looking for computationally tractable settings.
- We have convergence for potential and strategically zero-sum games (Anagnostides-Panageas-Farina-Sandhold22).

General sum Bimatrix Games (cont.)

Is cycling that bad? Not necessarily…

General sum Bimatrix Games (cont.)

Is cycling that bad? Not necessarily…

Theorem (Anagnostides, Farina, Panageas, Sandholm 22). *Let (A, B) be a bimatrix game and suppose agents follow optimistic gradient (ascent) with stepsize $\eta \approx \epsilon^2$. Then after $poly(1/\epsilon)$ iterations:*

- *Either the dynamics reaches an ϵ-approximate NE*
- *Or the average distribution consists a $poly(\epsilon)$-strong CCE.*

General sum Bimatrix Games (cont.)

Is cycling that bad? Not necessarily…

Theorem (Anagnostides, Farina, Panageas, Sandholm 22). *Let* (A, B) *be a bimatrix game and suppose agents follow optimistic gradient (ascent) with stepsize* $\eta \approx \epsilon^2$. *Then after* $poly(1/\epsilon)$ *iterations:*

- *Either the dynamics reaches an* ϵ*-approximate NE*
- *Or the average distribution consists a* $poly(\epsilon)$*-strong CCE.*

The techniques for the above show an interesting interplay between regret and convergence.

General sum Bimatrix Games (cont.)

Is cycling that bad? Not necessarily…

Theorem (Anagnostides, Farina, Panageas, Sandholm 22). *Let (A, B) be a bimatrix game and suppose agents follow optimistic gradient (ascent) with stepsize $\eta \approx \epsilon^2$. Then after $poly(1/\epsilon)$ iterations:*

- *Either the dynamics reaches an ϵ-approximate NE*
- *Or the average distribution consists a $poly(\epsilon)$-strong CCE.*

The techniques for the above show an interesting interplay between regret and convergence.

If sum of regrets is non-negative ⇒ Optimistic Gradient exhibits best iterate
Sum of regrets is negative ⇒ strong-CCE.

Take away messages and future directions

Learning dynamics can have all kinds of behaviors.

- Cycling even for 2player zero-sum which is fixable
- However not much is known if one wants to go beyond zero-sum.
- Learning in rank-1 games.
- If one is happy with cycles, will get an exact CCE in constant steps.

Can we go beyond two players? E.g., team games.

Lower bounds on rates of convergence? For Fictitious Play we have for both zero-sum and potential games (Daskalakis, Pan14) and (Panageas, Patris, Skoulakis, Cevher23).

Take away messages and future directions

Learning dynamics can have all kinds of behaviors.

- Cycling even for 2player zero-sum which is fixable
- However not much is known if one wants to go beyond zero-sum.
- Learning in rank-1 games.
- If one is happy with cycles, will get an exact CCE in constant steps.

Can we go beyond two players? E.g., team game

Lower bounds on rates of convergence? For Fict
both zero-sum and potential games (Daskalakis,
(Panageas, Patris, Skoulakis, Cevher23).