Making Optimistic Gradient Adaptive and Robust to Noise

Yu-Guan Hsieh

Université Grenoble Alpes, France

April 8th, 2023

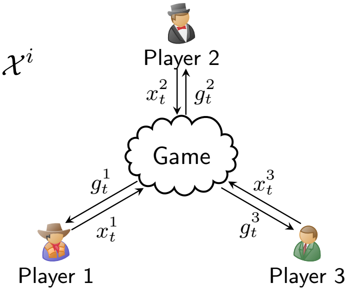
Outline

Learning in Continuous Games with Gradient Feedback

At each round $t = 1, 2, \ldots$, each player $i \in \{1, \ldots, N\}$

- Plays an action $x_t^i \in \mathcal{X}^i$
- Suffers loss $\ell^i(\mathbf{x}_t)$ and receives first order feedback $g_t^i \approx \nabla_i \ell^i(\mathbf{x}_t)$

- Each player i is associated with a convex closed action set $\mathcal{X}^i$ and a loss function $\ell^i : \mathcal{X}^1 \times \ldots \times \mathcal{X}^N \to \mathbb{R}$
- Joint action of all players $\mathbf{x} = (x^i)_{i \in \mathcal{N}} = (x^i, \mathbf{x}^{-i})$
- $\ell^i(\cdot, \mathbf{x}^{-i})$ is convex and $\nabla_i \ell^i(\mathbf{x}_t)$ is Lipschitz continuous

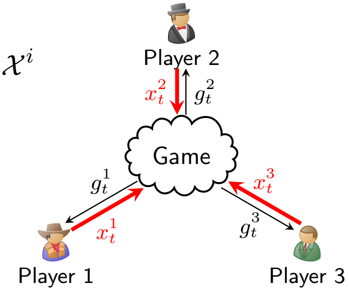


Learning in Continuous Games with Gradient Feedback

At each round $t = 1, 2, \ldots$, each player $i \in \{1, \ldots, N\}$

- Plays an action $x_t^i \in \mathcal{X}^i$
- Suffers loss $\ell^i(\mathbf{x}_t)$ and receives first order feedback $g_t^i \approx \nabla_i \ell^i(\mathbf{x}_t)$

- Each player i is associated with a convex closed action set $\mathcal{X}^i$ and a loss function $\ell^i : \mathcal{X}^1 \times \ldots \times \mathcal{X}^N \to \mathbb{R}$
- Joint action of all players $\mathbf{x} = (x^i)_{i \in \mathcal{N}} = (x^i, \mathbf{x}^{-i})$
- $\ell^i(\cdot, \mathbf{x}^{-i})$ is convex and $\nabla_i \ell^i(\mathbf{x}_t)$ is Lipschitz continuous

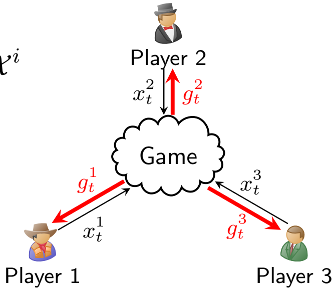


Learning in Continuous Games with Gradient Feedback

At each round $t = 1, 2, \ldots$, each player $i \in \{1, ..., N\}$

- Plays an action $x_t^i \in \mathcal{X}^i$
- Suffers loss $\ell^i(\mathbf{x}_t)$ and receives first order feedback $g_t^i \approx \nabla_i \ell^i(\mathbf{x}_t)$

- Each player i is associated with a convex closed action set $\mathcal{X}^i$ and a loss function $\ell^i : \mathcal{X}^1 \times \ldots \times \mathcal{X}^N \to \mathbb{R}$
- Joint action of all players $\mathbf{x} = (x^i)_{i \in \mathcal{N}} = (x^i, \mathbf{x}^{-i})$
- $\ell^i(\cdot, \mathbf{x}^{-i})$ is convex and $\nabla_i \ell^i(\mathbf{x}_t)$ is Lipschitz continuous



Evaluating Learning in Games Algorithms

Two interaction scenarios

- Self-play: all the players use the same algorithm
- Adversarial: the actions of the other players are arbitrary and even adversarial

Two evaluation criteria

- Regret of player i with respect to comparator set $\mathcal{P}^i$

$$\mathrm{Reg}_T^i(\mathcal{P}^i) = \max_{p^i \in \mathcal{P}^i} \sum_{t=1}^{T} \Big(\underbrace{\ell^i(x_t^i, \mathbf{x}_t^{-i}) - \ell^i(p^i, \mathbf{x}_t^{-i})}_{\text{cost of not playing } p^i \text{ in round } t} \Big).$$

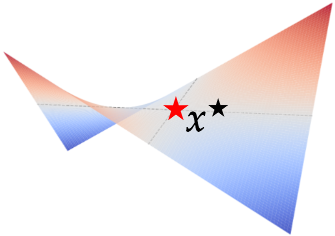
- Whether the sequence of play $\mathbf{x}_t$ converges to a Nash equilibrium $\mathbf{x}_\star$

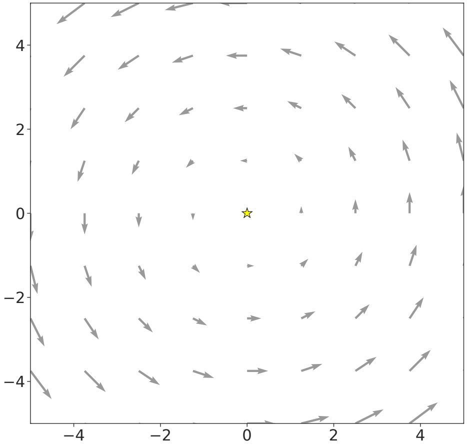
The Failure of Vanilla Gradient Method in Bilinear Games

- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

Unique Nash equilibrium: $(0, 0)$





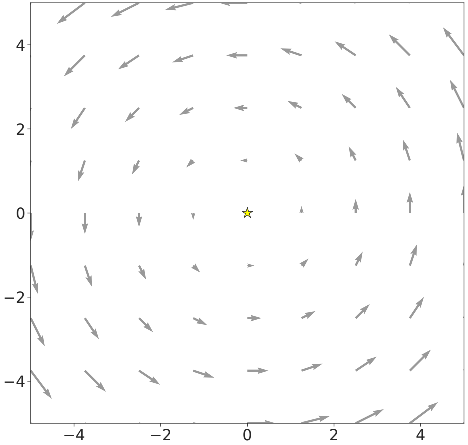
The Failure of Vanilla Gradient Method in Bilinear Games

- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \ell^1(\mathbf{x}), \nabla_\phi \ell^2(\mathbf{x})) = (\phi, -\theta)$$



The Failure of Vanilla Gradient Method in Bilinear Games

- Two-player planar bilinear zero-sum game

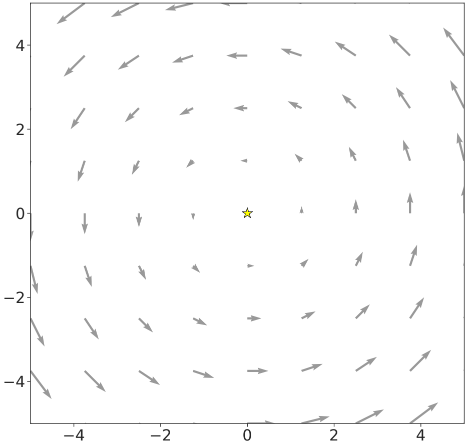
$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \ell^1(\mathbf{x}), \nabla_\phi \ell^2(\mathbf{x})) = (\phi, -\theta)$$

- Gradient descent

$$\mathbf{X}_{t+1} = \mathbf{X}_t - \eta_t \mathbf{V}(\mathbf{X}_t)$$



The Failure of Vanilla Gradient Method in Bilinear Games

- Two-player planar bilinear zero-sum game

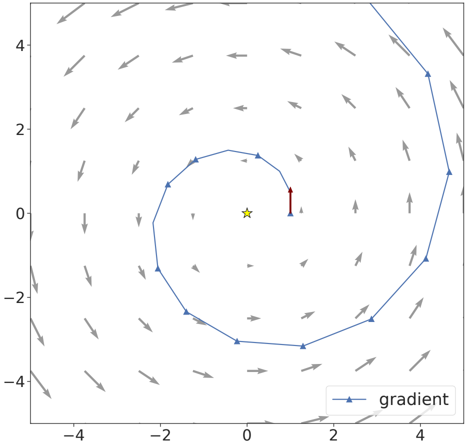
$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \ell^1(\mathbf{x}), \nabla_\phi \ell^2(\mathbf{x})) = (\phi, -\theta)$$

- Gradient descent

$$\mathbf{X}_{t+1} = \mathbf{X}_t - \eta_t \mathbf{V}(\mathbf{X}_t)$$



Optimistic Gradient to the Rescue

- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

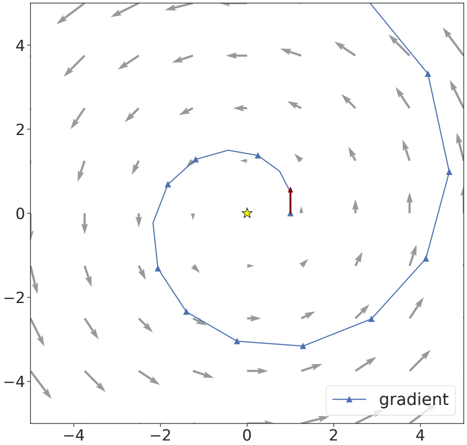
- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \, \ell^1(\mathbf{x}), \nabla_\phi \, \ell^2(\mathbf{x})) = (\phi, -\theta)$$

- Optimistic gradient [Popov 1980]

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \, \mathbf{V}(\mathbf{X}_{t-\frac{1}{2}})$$

$$\mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$



Optimistic Gradient to the Rescue

- Two-player planar bilinear zero-sum game

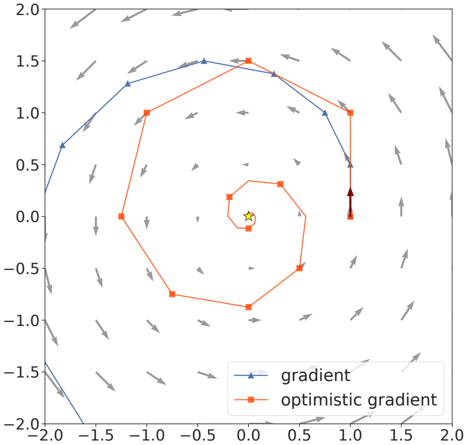
$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \ell^1(\mathbf{x}), \nabla_\phi \ell^2(\mathbf{x})) = (\phi, -\theta)$$

- Optimistic gradient [Popov 1980]

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}(\mathbf{X}_{t-\frac{1}{2}})$$
$$\mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$



Optimistic Gradient to the Rescue

- Two-player planar bilinear zero-sum game

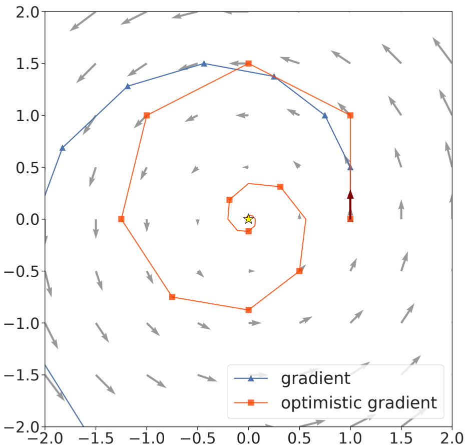
$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad [\mathbf{x} = (\theta, \phi) \in \mathbb{R}^2]$$

- Joint vector field

$$\mathbf{V}(\mathbf{x}) \coloneqq (\nabla_\theta \ell^1(\mathbf{x}), \nabla_\phi \ell^2(\mathbf{x})) = (\phi, -\theta)$$

- Optimistic gradient [Popov 1980]

$$X^i_{t+\frac{1}{2}} = X^i_{t-\frac{1}{2}} - 2\eta^i_t g^i_t + \eta^i_{t-1} g^i_{t-1} \quad (x^i_t = X^i_{t+\frac{1}{2}})$$



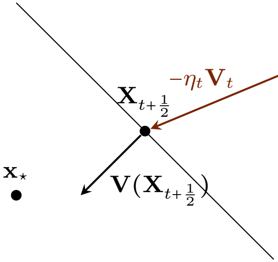
Outline

Intuitions Behind Optimistic Gradient

- Better discretization of the continuous flow [Lu 21]
- Look into the future, anticipating the landscape
  - Learning with recency bias [Rakhlin and Sridharan 13, Syrgkanis et al. 15]
    Compare with methods that 'knows the future'
  - Approximation of proximal point (PP) methods [Mokhtari et al. 20]
- Relaxed projection onto a separating hyperplane [Tseng 00, Facchinei and Pang 03]

Optimistic Gradient as Relaxed projection onto Separating Hyperplane

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

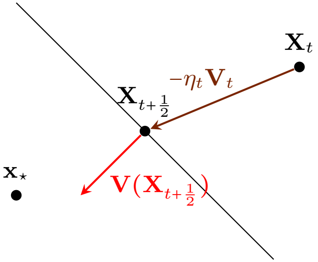


Optimistic Gradient as Relaxed projection onto Separating Hyperplane

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$



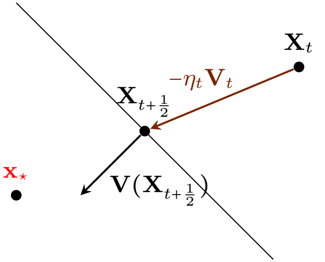
Optimistic Gradient as Relaxed projection onto Separating Hyperplane

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$

- Assumption: $\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x}_\star \rangle \geq 0$



Optimistic Gradient as Relaxed projection onto Separating Hyperplane

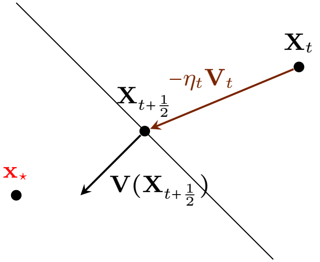
$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$

- Assumption: $\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x}_\star \rangle \geq 0$

  Monotone: $\langle \mathbf{V}(\mathbf{x}') - \mathbf{V}(\mathbf{x}), \mathbf{x}' - \mathbf{x} \rangle \geq 0$



Optimistic Gradient as Relaxed projection onto Separating Hyperplane

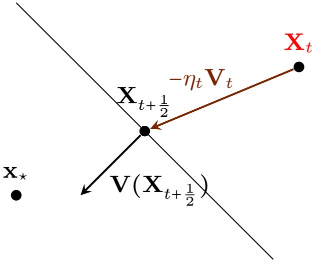
$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$

- Assumption: $\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x}_\star \rangle \geq 0$
- If $\mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}) \approx \mathbf{V}_t$ then

$$\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{X}_t \rangle = -\eta_{t+1} \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{V}_t \rangle \leq 0$$



Optimistic Gradient as Relaxed projection onto Separating Hyperplane

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

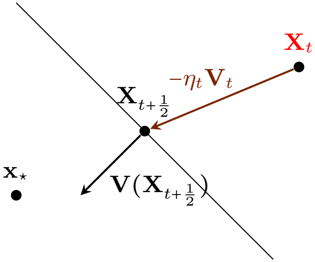
- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$

- Assumption: $\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x}_\star \rangle \geq 0$
- If $\mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}) \approx \mathbf{V}_t$ then

$$\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{X}_t \rangle = -\eta_{t+1} \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{V}_t \rangle \leq 0$$

This is why we require Lipschitz continuity



Optimistic Gradient as Relaxed projection onto Separating Hyperplane

$$\mathbf{X}_{t+\frac{1}{2}} = \mathbf{X}_t - \eta_t \mathbf{V}_t, \quad \mathbf{X}_{t+1} = \mathbf{X}_t - \eta_{t+1} \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$$

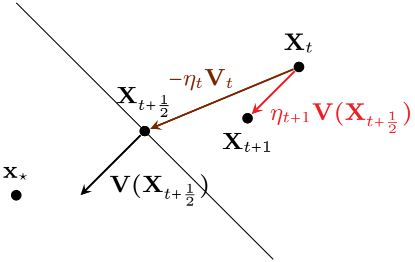
- Consider the hyperplan

$$\mathcal{H} \coloneqq \{\mathbf{x} : \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x} \rangle = 0\}$$

- Assumption: $\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{x}_\star \rangle \geq 0$
- If $\mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}) \approx \mathbf{V}_t$ then

$$\langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{X}_{t+\frac{1}{2}} - \mathbf{X}_t \rangle = -\eta_{t+1} \langle \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}}), \mathbf{V}_t \rangle \leq 0$$

- The update step moves the iterate closer to the solutions



Variational Stability

Definition [Variationally stable games]

Let $\mathbf{V} = (\nabla_1 \ell^1, \ldots, \nabla_M \ell^M)$. A continuous convex game is variationally stable (VS) if the set $\mathcal{X}_\star$ of Nash equilibria of the game is nonempty and

$$\langle \mathbf{V}(\mathbf{x}), \mathbf{x} - \mathbf{x}_\star \rangle = \sum_{i=1}^{N} \langle \nabla_i \ell^i(\mathbf{x}), x^i - x^i_\star \rangle \geq 0 \quad \text{for all } \mathbf{x} \in \mathcal{X},\ \mathbf{x}_\star \in \mathcal{X}_\star \tag{1}$$

The game is strictly variationally stable if (1) holds as a strict inequality whenever $\mathbf{x} \notin \mathcal{X}_\star$.

Especially, a game is variationally stable if $\mathbf{V}$ is monotone

Examples:
- Convex-concave zero-sum games
- Zero-sum polymatrix games
- Cournot oligopolies
- Kelly auctions

Theoretical Guarantees of Optimistic Gradients

- Bounded gradient feedback: $\exists\, G > 0,\ \forall\, t,\ g_t^i \leq G$
- Learning rate $\eta_t = \Theta(1/\sqrt{t})$
- $\mathcal{O}\left(\sqrt{t}\right)$ minimax-optimal regret in the adversarial regime

	Adversarial	Same algorithm + Variational Stability			
	Bounded feedback	-	-	Strongly M	Error bound
	$\mathrm{Reg}_t\,/t$	$\mathrm{Reg}_t\,/t$	$\|\mathbf{V}(\mathbf{x}_t)\|$	$\mathrm{dist}(\mathbf{x}_t, \mathcal{X}_\star)$	$\mathrm{dist}(\mathbf{x}_t, \mathcal{X}_\star)$
GD	$1/\sqrt{t}$	✗	✗	$e^{-\rho_1 t}$	✗
OG	$1/\sqrt{t}$ Chiang et al. 12	$1/t$ H. et al. 19	$1/\sqrt{t}$ Cai et al. 22	$e^{-\rho_2 t}$ $(\rho_2 \geq \rho_1)$ Mokhtari et al. 20	$e^{-\rho t}$ Wei et al. 21

Theoretical Guarantees of Optimistic Gradients

- Joint vector field: $\mathbf{V}(\mathbf{X}) = (\nabla_i \ell^i(\mathbf{X}))_{i\in\mathcal{N}}$; Nash equilibria: $\mathcal{X}_\star$
- Strongly monotone: $\exists\, \alpha > 0$, $\langle \mathbf{V}(\mathbf{x}') - \mathbf{V}(\mathbf{x}), \mathbf{x}' - \mathbf{x}\rangle \geq \alpha \|\mathbf{x}' - \mathbf{x}\|^2$
- Error bound / Metric (sub-)regularity: $\exists\, \tau > 0$, $\|\mathbf{V}(\mathbf{x})\| \geq \tau\, \mathrm{dist}(\mathbf{x}, \mathcal{X}_\star)$

	Adversarial	Same algorithm + Variational Stability			
	Bounded feedback	-	-	Strongly M	Error bound
	Reg_t/t	Reg_t/t	$\|\mathbf{V}(\mathbf{x}_t)\|$	$\mathrm{dist}(\mathbf{x}_t, \mathcal{X}_\star)$	$\mathrm{dist}(\mathbf{x}_t, \mathcal{X}_\star)$
GD	$1/\sqrt{t}$	✗	✗	$e^{-\rho_1 t}$	✗
OG	$1/\sqrt{t}$ Chiang et al. 12	$1/t$ H. et al. 19	$1/\sqrt{t}$ Cai et al. 22	$e^{-\rho_2 t}$ $(\rho_2 \geq \rho_1)$ Mokhtari et al. 20	$e^{-\rho t}$ Wei et al. 21

Limitations of Vanilla Optimistic Methods I: Learning Rate Turning

Fast convergence of sequence of play is mostly proved for **suitably tuned** learning rates

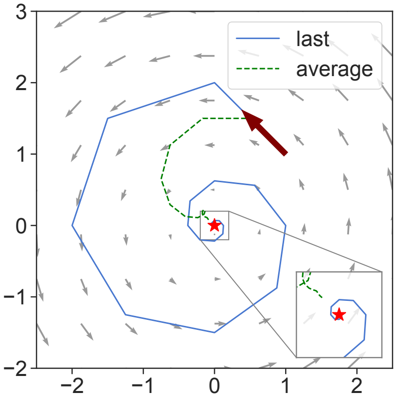
- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad \text{where} \quad \mathcal{X}^1 = \mathcal{X}^2 = [-4, 8]$$

- The two players play optimistic gradient with constant stepsize $\eta = 0.5$ and $T = 100$

Property

OG converges in bilinear zero-sum games



Limitations of Vanilla Optimistic Methods I: Learning Rate Turning

Fast convergence of sequence of play is mostly proved for **suitably tuned** learning rates

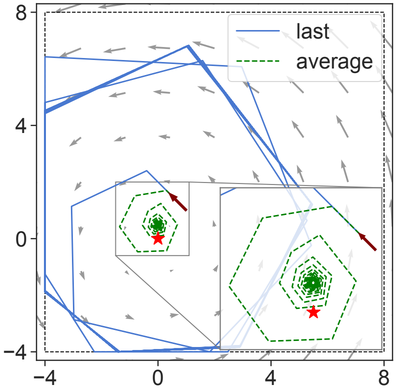
- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad \text{where} \quad \mathcal{X}^1 = \mathcal{X}^2 = [-4, 8]$$

- The two players play optimistic gradient with constant stepsize $\eta = 0.7$ and $T = 100$

Problem

This only holds when η is small enough



Limitations of Vanilla Optimistic Methods I: Learning Rate Turning

Fast convergence of sequence of play is mostly proved for **suitably tuned** learning rates

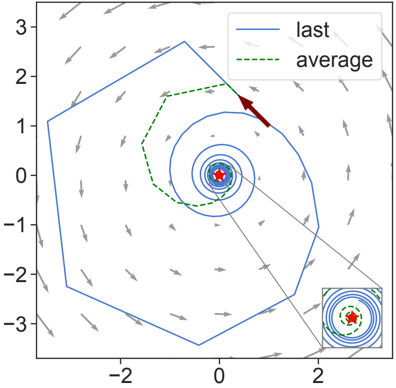
- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad \text{where} \quad \mathcal{X}^1 = \mathcal{X}^2 = [-4, 8]$$

- The two players play optimistic gradient with decreasing stepsize $\eta_t = 1/\sqrt{t}$ and $T = 100$

Solution?

$\eta_t \propto 1/\sqrt{t} \rightarrow$ slow convergence



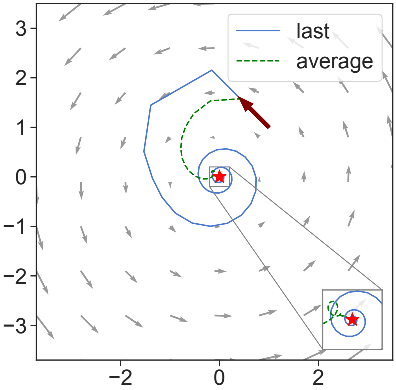
Limitations of Vanilla Optimistic Methods I: Learning Rate Turning

Fast convergence of sequence of play is mostly proved for **suitably tuned** learning rates

- Two-player planar bilinear zero-sum game

$$\ell^1(\mathbf{x}) = -\ell^2(\mathbf{x}) = \theta\phi \quad \text{where} \quad \mathcal{X}^1 = \mathcal{X}^2 = [-4, 8]$$

- The two players play optimistic gradient with adaptive stepsize and $T = 100$



Solution

Adaptive learning ← focus of part I

Limitations of Vanilla Optimistic Methods II: Noisy Feedback

All the favorable guarantees break if feedback is **noisy**

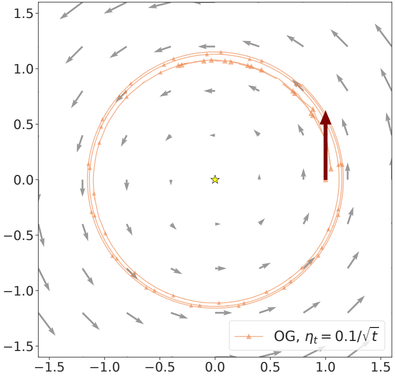
- Stochastic estimate $\mathbb{E}[\hat{\mathbf{V}}_{t+\frac{1}{2}}] = \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$

$$\hat{\mathbf{V}}_{t+\frac{1}{2}} = \begin{cases} (3\phi_{t+\frac{1}{2}}, -3\theta_{t+\frac{1}{2}}) & \text{with prob. } 1/2 \\ (-\phi_{t+\frac{1}{2}}, \theta_{t+\frac{1}{2}}) & \text{with prob. } 1/2 \end{cases}$$

- The two players play optimistic gradient with decreasing stepsize $\eta_t = 0.1/\sqrt{t}$

Problem

We observe non-convergence and linear regret



Limitations of Vanilla Optimistic Methods II: Noisy Feedback

All the favorable guarantees break if feedback is **noisy**

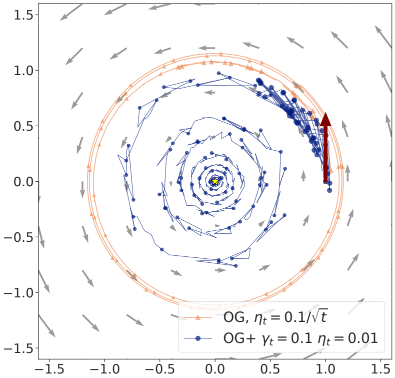
- Stochastic estimate $\mathbb{E}[\hat{\mathbf{V}}_{t+\frac{1}{2}}] = \mathbf{V}(\mathbf{X}_{t+\frac{1}{2}})$

$$\hat{\mathbf{V}}_{t+\frac{1}{2}} = \begin{cases} (3\phi_{t+\frac{1}{2}}, -3\theta_{t+\frac{1}{2}}) & \text{with prob. } 1/2 \\ (-\phi_{t+\frac{1}{2}}, \theta_{t+\frac{1}{2}}) & \text{with prob. } 1/2 \end{cases}$$

- The two players play optimistic gradient with decreasing stepsize $\eta_t = 0.1/\sqrt{t}$

Solution

Scale separation → focus of part II



Outline

Mirror Descent versus Dual Averaging

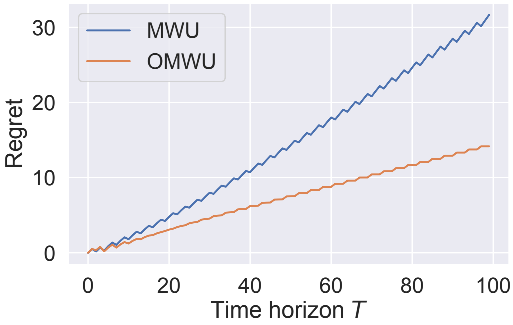
Mirror descent type methods with dynamic learning rates **may incur regret**

Assume that player 1 has a linear loss and simplex-constrained action set.

- $\mathcal{X}^1 = \Delta^1 = \{(w_1, w_2) \in \mathbb{R}^2_+, w_1 + w_2 = 1\}$
- Feedback sequence:

$$[\underbrace{-e_1, \ldots, -e_1}_{\lceil T/3 \rceil}, \underbrace{-e_2, \ldots, -e_2}_{\lfloor 2T/3 \rfloor}]$$

- Adaptive (Optimistic) Multiplicative Weight Update



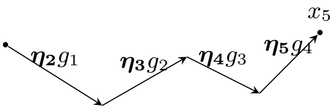
(Example from [Orabona and Pal 16])

Mirror Descent versus Dual Averaging

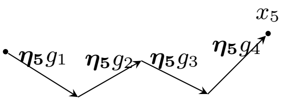
Mirror descent type methods with dynamic learning rates **may incur regret**

- Cause: new information enters MD with a decreasing weight
- Solution: enter each feedback with equal weight

  E.g. Dual averaging or stabilization technique



MD-type update



DA-type update

Mirror Descent versus Dual Averaging

Mirror descent type methods with dynamic learning rates **may incur regret**

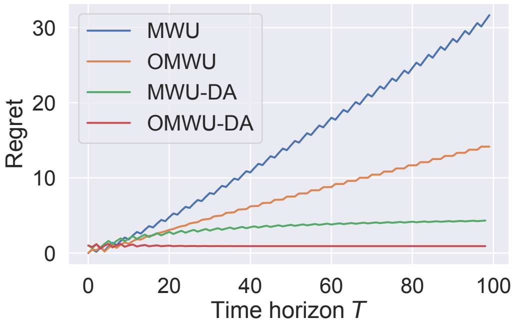
Assume that player 1 has a linear loss and simplex-constrained action set.

- $\mathcal{X}^1 = \Delta^1 = \{(w_1, w_2) \in \mathbb{R}^2_+, w_1 + w_2 = 1\}$
- Feedback sequence:

$$[\underbrace{-e_1, \ldots, -e_1}_{\lceil T/3 \rceil}, \underbrace{-e_2, \ldots, -e_2}_{\lfloor 2T/3 \rfloor}]$$

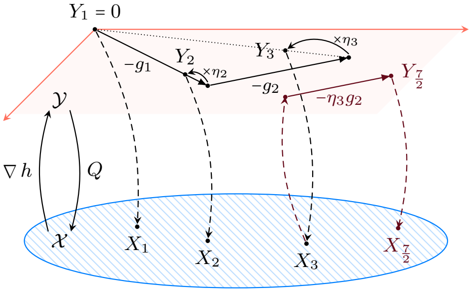
- Adaptive (Optimistic) Multiplicative Weight Update with Dual Averaging

(Example from [Orabona and Pal 16])



Optimistic Dual Averaging (OptDA)

$$X_t^i = \arg\min_{x\in\mathcal{X}^i} \sum_{s=1}^{t-1}\langle g_s^i, x\rangle + \frac{h^i(x)}{\eta_t^i} = Q\Big(\underbrace{-\eta_t^i \sum_{s=1}^{t-1} g_t^s}_{Y_t^i}\Big), \qquad X_{t+\frac{1}{2}}^i = \arg\min_{x\in\mathcal{X}^i} \langle g_{t-1}^i, x\rangle + \frac{D^i(x, X_t^i)}{\eta_t^i}$$



Regularizer h^i: 1-strongly convex and C^1

Mirror map: $Q^i(y) = \arg\max_{x\in\mathcal{X}^i} \langle y, x\rangle - h^i(x)$

Bregman divergence:

$D^i(p, x) = h^i(p) - h^i(x) - \langle \nabla h^i(x), p - x\rangle$

Optimistic Dual Averaging: Examples

- OG-OptDA ▸ $\mathcal{X}^i$ convex closed ▸ $h^i(x) = \frac{\|x\|_2^2}{2}$ ▸ Q: Euclidean projection $\Pi_{\mathcal{X}}$

$$X_t^i = \Pi_{\mathcal{X}}(-\eta_t^i \sum_{s=1}^{t-1} g_t^s), \qquad X_{t+\frac{1}{2}}^i = \Pi_{\mathcal{X}}(X_t^i - \eta_t^i g_{t-1}^i)$$

- Stabilized OMWU ▸ $\mathcal{X}^i = \Delta^{d^i-1}$ ▸ $h^i(x) = \sum_{k=1}^{d_i} x_{[k]} \log x_{[k]}$ ▸ Q: Softmax

$$X_{t+\frac{1}{2},[k]}^{(i)} = \frac{\exp(-\eta_t^i(\sum_{s=1}^{t-1} g_{s,[k]} + g_{t-1,[k]}))}{\sum_{l=1}^{d_i} \exp(-\eta_t^i(\sum_{s=1}^{t-1} g_{s,[l]} + g_{t-1,[l]}))}$$

Energy Inequality

$$\frac{F^i(p^i, Y^i_{t+1})}{\eta^i_{t+1}} \leq \frac{F^i(p^i, Y^i_t)}{\eta^i_t} - \langle g^i_t, X^i_{t+\frac{1}{2}} - p^i \rangle + \left(\frac{1}{\eta^i_{t+1}} - \frac{1}{\eta^i_t} \right) (h^i(p^i) - \min h^i)$$

$$+ \langle g^i_t - g^i_{t-1}, X^i_{t+\frac{1}{2}} - X^i_{t+1} \rangle - \frac{D^i(X^i_{t+1}, X^i_{t+\frac{1}{2}})}{\eta^i_t} - \frac{D^i(X^i_{t+\frac{1}{2}}, X^i_t)}{\eta^i_t}$$

$F^i(p, y) = h^i(p) + (h^i)^*(y) - \langle y, p \rangle$ is Fenchel coupling

1. $F^i(p, y) \geq \frac{1}{2} \|Q^i(y) - p\|^2$
2. Reciprocity condition: if $X^i_t \to p^i$ then $F^i(p^i, Y^i_t) \to 0$

Energy Inequality

$$\frac{F^i(p^i, Y^i_{t+1})}{\eta^i_{t+1}} \le \frac{F^i(p^i, Y^i_t)}{\eta^i_t} - \langle g^i_t, X^i_{t+\frac{1}{2}} - p^i \rangle + \left(\frac{1}{\eta^i_{t+1}} - \frac{1}{\eta^i_t} \right) (h^i(p^i) - \min h^i)$$
$$+ \langle g^i_t - g^i_{t-1}, X^i_{t+\frac{1}{2}} - X^i_{t+1} \rangle - \frac{D^i(X^i_{t+1}, X^i_{t+\frac{1}{2}})}{\eta^i_t} - \frac{D^i(X^i_{t+\frac{1}{2}}, X^i_t)}{\eta^i_t}$$

Sum the energy inequality from $t = 1$ to T gives

$$\sum_{t=1}^{T} \langle g^i_t, X^i_{t+\frac{1}{2}} - p^i \rangle \le \frac{h^i(p^i) - \min h^i}{\eta^i_{T+1}} + \sum_{t=1}^{T} \eta^i_t \| g^i_t - g^i_{t-1} \|^2_{(i),*} - \sum_{t=2}^{T} \frac{1}{8\eta^i_{t-1}} \| X^i_{t+\frac{1}{2}} - X^i_{t-\frac{1}{2}} \|^2_{(i)}$$

Adaptive Learning Rate

$$\sum_{t=1}^{T}\langle g_t^i, X_{t+\frac{1}{2}}^i - p^i\rangle \leq \frac{h^i(p^i) - \min h^i}{\eta_{T+1}^i} + \sum_{t=1}^{T} \eta_t^i \|g_t^i - g_{t-1}^i\|_{(i),*}^2 - \sum_{t=2}^{T} \frac{1}{8\eta_{t-1}^i}\|X_{t+\frac{1}{2}}^i - X_{t-\frac{1}{2}}^i\|_{(i)}^2$$

Take the adaptive learning rate

$$\eta_t^i = \frac{1}{\sqrt{\tau^i + \sum_{s=1}^{t-1}\|g_t^i - g_{t-1}^i\|_{(i),*}^2}}$$

- $\tau^i > 0$ can be chosen freely by the player
- η_t^i is thus computed solely based on local information available to each player

Theoretical Guarantees for General Convex Games

Let player i plays OptDA or DS-OptMD with (Adapt):

- No-regret: If $\mathcal{P}^i \subseteq \mathcal{X}^i$ is bounded and $G = \sup_t \|g_t^i\|$, the regret incurred by the player is bounded as $\mathrm{Reg}_T^i(\mathcal{P}^i) = \mathcal{O}(G\sqrt{T} + G^2)$.
- Consistent: If $\mathcal{X}^i$ is compact and the action profile $\mathbf{x}_t^{-i}$ of all other players converges to some limit profile $\mathbf{x}_\infty^{-i}$, the trajectory of chosen actions of player i converges to the best response set $\arg\min_{x^i \in \mathcal{X}^i} \ell^i(x^i, \mathbf{x}_\infty^{-i})$.

Theoretical Guarantees for Variationally Stable Games

If all players use AdaOptDA in a variationally stable game:

- Constant regret For all $i \in \mathcal{N}$ and every bounded comparator set $\mathcal{P}^i \subseteq \mathcal{X}^i$, the individual regret of player i is bounded as $\mathrm{Reg}_T^i(\mathcal{P}^i) = \mathcal{O}(1)$.
- Convergence to Nash equilibrium The induced trajectory of play converges to a Nash equilibrium provided that either of the following is satisfied:
  - a The game is strictly variationally stable
  - b The game is variationally stable and h^i is (sub)differentiable on all $\mathcal{X}^i$
  - c The players of a two-player finite zero-sum game follow stabilized OMWU

Outline

Stochastic Oracle

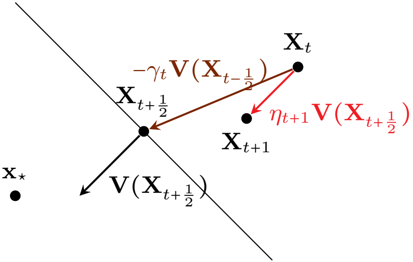
We focus on the unconstrained setup $\mathcal{X}^i = \mathbb{R}^{d^i}$ in this part

- Stochastic feedback $g_t^i = V^i(\mathbf{x}_t) + \xi_t^i$ with noise satisfying
  1. *Zero-mean:* For all $i \in \mathcal{N}$ and $t \in \mathbb{N}$, $\mathbb{E}_t[\xi_t^i] = 0$.
  2. *Variance control:* For all $i \in \mathcal{N}$ and $t \in \mathbb{N}$, $\mathbb{E}_t[\|\xi_t^i\|^2] \leq \sigma_A^2 + \sigma_M^2 \|V^i(\mathbf{x}_t)\|^2$.
- We say that the noise is multiplicative if $\sigma_A^2 = 0$:
  randomized coordinate descent, physical measurement, finite sum of operators whose solution sets intersect

Scale Separation of Learning Rates

$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_t - \eta^i_{t+1} \, g^i_t \tag{OG+}$$

$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_1 - \eta^i_{t+1} \sum_{s=1}^{t} g^i_s \tag{OptDA+}$$



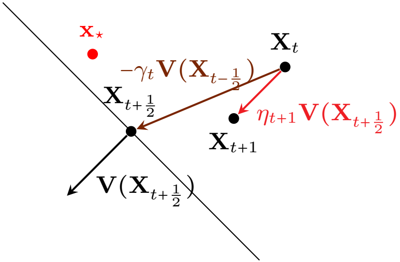
Scale Separation of Learning Rates

$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_t - \eta^i_{t+1} \, g^i_t \qquad \text{(OG+)}$$

$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_1 - \eta^i_{t+1} \sum_{s=1}^{t} g^i_s \qquad \text{(OptDA+)}$$

- Weak minty solution, for some $\rho \in (-1/2\beta, 0)$

$$\langle \mathbf{V}(\mathbf{x}), \mathbf{x} - \mathbf{x}_\star \rangle \geq \rho \| \mathbf{V}(\mathbf{x}) \|^2$$



Scale Separation of Learning Rates

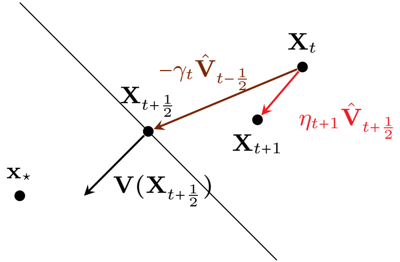
$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_t - \eta^i_{t+1} \, g^i_t \tag{OG+}$$

$$X^i_{t+\frac{1}{2}} = X^i_t - \gamma^i_t \, g^i_{t-1}, \qquad X^i_{t+1} = X^i_1 - \eta^i_{t+1} \sum_{s=1}^{t} g^i_s \tag{OptDA+}$$

- Weak minty solution, for some $\rho \in (-1/2\beta, 0)$

$$\langle \mathbf{V}(\mathbf{x}), \mathbf{x} - \mathbf{x}_\star \rangle \geq \rho \|\mathbf{V}(\mathbf{x})\|^2$$

- Stochastic update: relaxation of an approximate projection step with relaxation factor of the order of $\eta_{t+1}/\gamma_t \to$ the ratio η_{t+1}/γ_t should go to 0



Energy Inequality

$$\mathbb{E}_{t-1}\left[\frac{\|X^i_{t+1}-p^i\|^2}{\eta^i_{t+1}}\right] \le \mathbb{E}_{t-1}\left[\frac{\|X^i_t-p^i\|^2}{\eta^i_t} + \left(\frac{1}{\eta^i_{t+1}}-\frac{1}{\eta^i_t}\right)\|u^i_t-p^i\|^2\right.$$

(linearized regret) $\quad -2\langle V^i(\mathbf{X}_{t+\frac{1}{2}}), X^i_{t+\frac{1}{2}}-p^i\rangle$

(negative drift) $\quad -\gamma^i_t(\|V^i(\mathbf{X}_{t+\frac{1}{2}})\|^2+\|V^i(\mathbf{X}_{t-\frac{1}{2}})\|^2)$

(use smoothness) $\quad -\frac{\|X^i_t-X^i_{t+1}\|^2}{2\eta^i_t}+\gamma^i_t\|V^i(\mathbf{X}_{t+\frac{1}{2}})-V^i(\mathbf{X}_{t-\frac{1}{2}})\|^2$

(noise) $\quad \left.+(\gamma^i_t)^2\beta\|\xi^i_{t-\frac{1}{2}}\|^2+\beta\|\boldsymbol{\xi}_{t-\frac{1}{2}}\|^2_{(\boldsymbol{\eta}_t+\boldsymbol{\gamma}_t)^2}+2\eta^i_t\|g^i_t\|^2\right]$

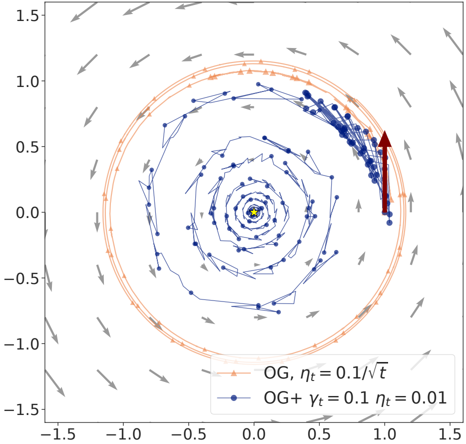
Last-iterate convergence

- OG+ is guaranteed to converge to Nash equilibrium under VS if

$$\sum_{t=1}^{+\infty} \gamma_t \eta_{t+1} = +\infty,$$

$$\sum_{t=1}^{+\infty} \gamma_t^2 \eta_{t+1} < +\infty, \quad \sum_{t=1}^{+\infty} \eta_t^2 < +\infty$$

- We can take constant learning rates if the noise is multiplicative

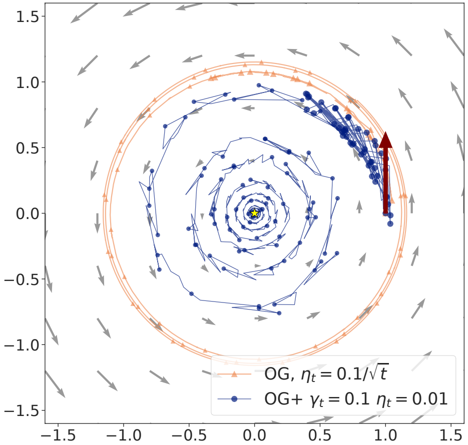


Adaptive Optimistic Dual Averaging with Scale Separation

- AdaOptDA+ uses learning rate

$$\gamma_t^i = \frac{1}{\left(1 + \sum_{s=1}^{t-2} \|g_s^i\|^2\right)^{\frac{1}{4}}}$$

$$\eta_t^i = \frac{1}{\sqrt{1 + \sum_{s=1}^{t-2} \left(\|g_s^i\|^2 + \|X_s^i - X_{s+1}^i\|^2\right)}}$$

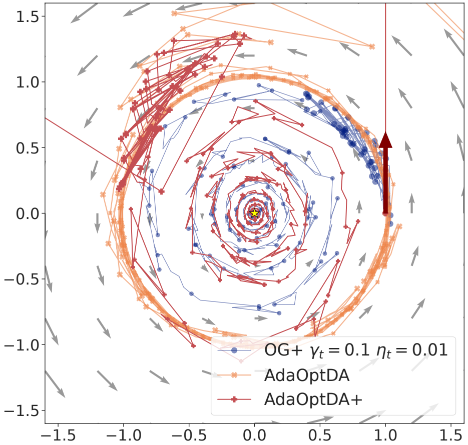


Adaptive Optimistic Dual Averaging with Scale Separation

- AdaOptDA+ uses learning rate

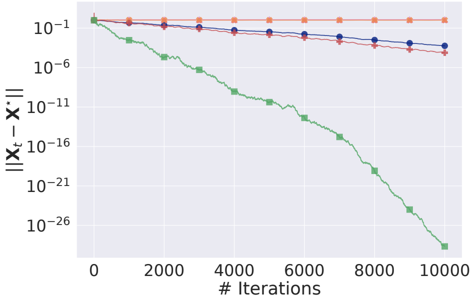
$$\gamma_t^i = \frac{1}{\left(1 + \sum_{s=1}^{t-2} \|g_s^i\|^2\right)^{\frac{1}{4}}}$$

$$\eta_t^i = \frac{1}{\sqrt{1 + \sum_{s=1}^{t-2} \left(\|g_s^i\|^2 + \|X_s^i - X_{s+1}^i\|^2\right)}}$$

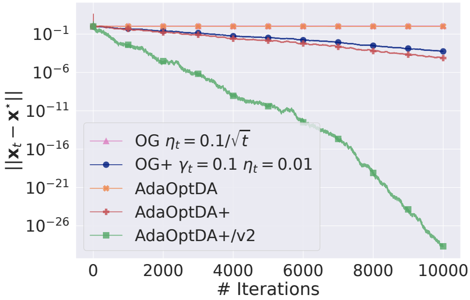


Convergence to Solution Under Multiplicative Noise

- $\hat{\mathbf{V}}_{t+\frac{1}{2}}$ is $(3\phi_{t+\frac{1}{2}}, -3\theta_{t+\frac{1}{2}})$ or $(-\phi_{t+\frac{1}{2}}, \theta_{t+\frac{1}{2}})$ with probability one half for each



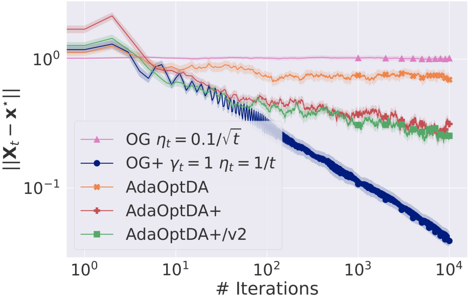
Base state $\mathbf{X}_t$



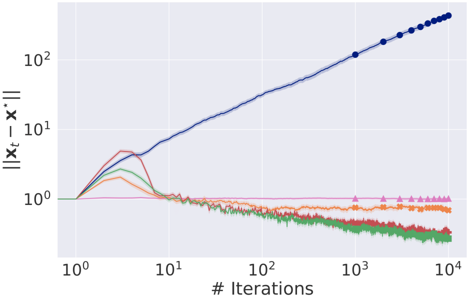
Played action $\mathbf{x}_t = \mathbf{X}_{t+\frac{1}{2}}$

Convergence to Solution Under Additive Noise

- $\hat{\mathbf{V}}_{t+\frac{1}{2}} = (\phi_{t+\frac{1}{2}} + \xi_t^1, -\theta_{t+\frac{1}{2}} + \xi_t^2)$ where $\xi_t^1, \xi_t^2 \sim \mathcal{N}(0,1)$



Base state $\mathbf{X}_t$



Played action $\mathbf{x}_t = \mathbf{X}_{t+\frac{1}{2}}$

Theoretical Guarantees Under Uncertainty (in Expectation)

		Adversarial	Same algorithm + Variational Stability			
		Bounded feedback	-	-	Strongly M	Error bound
		Reg_t / t	Reg_t / t	Cvg?	$\text{dist}(\mathbf{X}_t, \mathcal{X}_\star)$	$\text{dist}(\mathbf{X}_t, \mathcal{X}_\star)$
OG+	Mul.	✗	$1/t$	✓	$e^{-\rho t}$	$e^{-\rho t}$
	Add.		$1/\sqrt{t}$	✓	$1/\sqrt{t}$	$1/t^{1/6}$
OptDA+	Add.	$1/\sqrt{t}$	$1/\sqrt{t}$	-	-	-
AdaOptDA+	Mul.	$1/t^{1/4}$	$1/t$	✓	-	-
	Add.		$1/\sqrt{t}$	-	-	-

Outline

Summary

- We can make optimistic gradient adaptive with player-wise Adagrad type learning rate
- We can make optimistic gradient effective under noise feedback with scale separation of the optimistic and the update steps
- We can put the previous two points together

Perspectives

- Change feedback type: Bandit feedback
- Change interaction scenario: Partial adherence to the algorithm
- Change evaluation criterion: Policy regret
- Dealing with constraints under stochastic feedback
- Trajectory convergence for dual averaging under additive noise

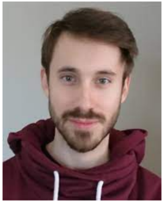
Perspectives

- Change feedback type: Bandit feedback
- Change interaction scenario: Partial adherence to the algorithm
- Change evaluation criterion: Policy regret
- Dealing with constraints under stochastic feedback
- Trajectory convergence for dual averaging under additive noise

Thank you for your attention

References

Jérôme Malick



Franck Iutzeler

Panayotis Mertikopoulos

Kimon Antonakopoulos

Volkan Cevher

1. Y-G H., F. Iutzeler, J. Malick, and P. Mertikopoulos. *Explore Aggressively, Update Conservatively: Stochastic Extragradient Methods with Variable Stepsize Scaling.* In **NeurIPS**, 2020.
2. Y-G H., K. Antonakopoulos., and P. Mertikopoulos. *Adaptive Learning in Continuous Games: Optimal Regret Bounds and Convergence to Nash Equilibrium.* In **COLT**, 2021.
3. Y-G H., K. Antonakopoulos, V. Cevher, and P. Mertikopoulos. *No-Regret Learning in Games with Noisy Feedback: Faster Rates and Adaptivity via Learning Rate Separation.* In **NeurIPS**, 2022.