Distributed Learning in Routing Games

Convergence, Estimation of Player Dynamics, and Control

Walid Krichene Alexandre Bayen

Electrical Engineering and Computer Sciences, UC Berkeley

IPAM Workshop on Decision Support for Traffic
November 18, 2015

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Communication networks
- Nash equilibrium quantifies efficiency of network in steady state.

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Communication networks
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Communication networks
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

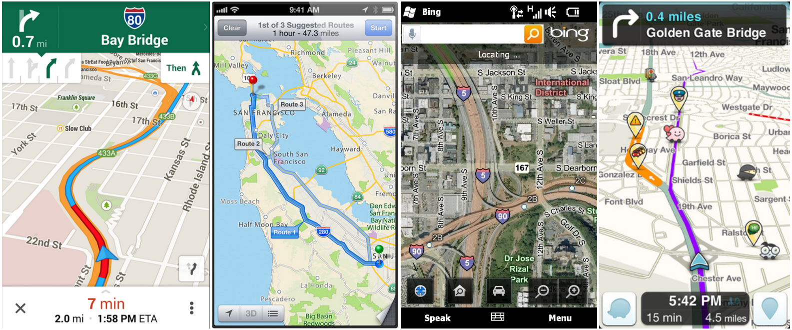
- A realistic model for decision dynamics is essential for prediction, optimal control.

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Communication networks
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

- A realistic model for decision dynamics is essential for prediction, optimal control.



Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.

Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.
- Consistent with the full information Nash equilibrium.

$$x^{(t)} \to \mathcal{X}^\star$$

Convergence rates?

Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.
- Consistent with the full information Nash equilibrium.

$$x^{(t)} \to \mathcal{X}^\star$$

  Convergence rates?
- Robust to stochastic perturbations.
  - Observation noise
  - (Bandit feedback)

Outline

Outline

1. Introduction
2. Convergence of agent dynamics
3. Routing Examples
4. Estimation and optimal control

Interaction of K decision makers

Decision maker k faces a sequential decision problem

At iteration t

(1) chooses probability distribution $x^{(t)}_{\mathcal{A}_k}$ over action set $\mathcal{A}_k$

(2) discovers a loss function $\ell^{(t)}_{\mathcal{A}_k} : \mathcal{A}_k \to [0, 1]$

(3) updates distribution

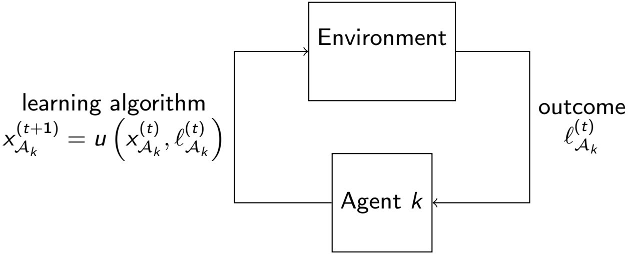


Figure: Sequential decision problem.

Interaction of K decision makers

Decision maker k faces a sequential decision problem
At iteration t

(1) chooses probability distribution $x^{(t)}_{\mathcal{A}_k}$ over action set $\mathcal{A}_k$
(2) discovers a loss function $\ell^{(t)}_{\mathcal{A}_k} : \mathcal{A}_k \to [0, 1]$
(3) updates distribution

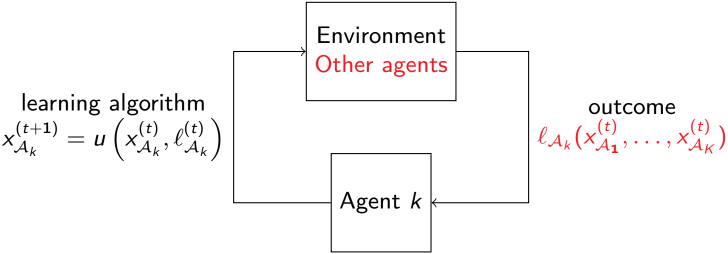


Figure: Sequential decision problem.

Loss of agent k affected by strategies of other agents.
Does not know this function, only observes its value.
Write $x^{(t)} = (x^{(t)}_{\mathcal{A}_1}, \dots, x^{(t)}_{\mathcal{A}_K})$.

Examples of decentralized decision makers

Routing game

- Player drives from source to destination node
- Chooses path from $\mathcal{A}_k$
- Mass of players on each edge determines cost on that edge.

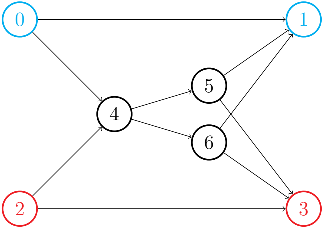


Figure: Routing game

Examples of decentralized decision makers

Routing game

- Player drives from source to destination node
- Chooses path from $\mathcal{A}_k$
- Mass of players on each edge determines cost on that edge.

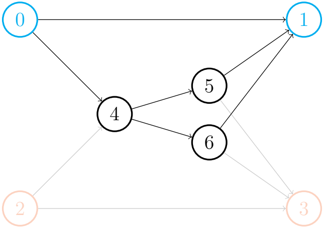


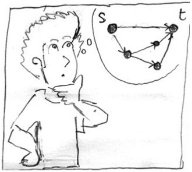
Figure: Routing game

Online learning model

Online Learning Model

```
for $t \in \mathbb{N}$ do
    Play $p \sim x^{(t)}_{\mathcal{A}_k}$
    Discover $\ell^{(t)}_{\mathcal{A}_k}$
    Update
        $x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$
end for
```

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$

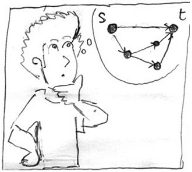


Online learning model

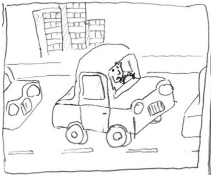
Online Learning Model

```
for $t \in \mathbb{N}$ do
    Play $p \sim x^{(t)}_{\mathcal{A}_k}$
    Discover $\ell^{(t)}_{\mathcal{A}_k}$
    Update
        $x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$
end for
```

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$



Sample $p \sim x^{(t)}_{\mathcal{A}_1}$

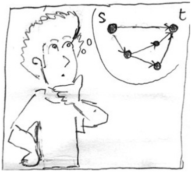


Online learning model

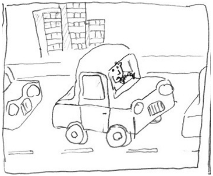
Online Learning Model

```
for $t \in \mathbb{N}$ do
    Play $p \sim x^{(t)}_{\mathcal{A}_k}$
    Discover $\ell^{(t)}_{\mathcal{A}_k}$
    Update
        $$x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$$
end for
```

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$



Sample $p \sim x^{(t)}_{\mathcal{A}_1}$

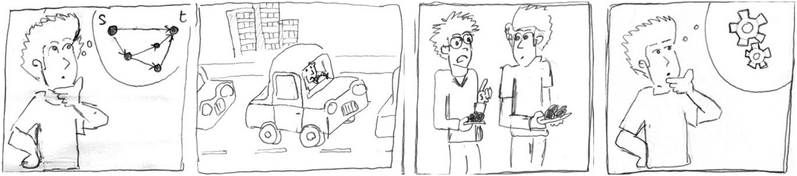


Discover $\ell^{(t)}_{\mathcal{A}_1}$

Online learning model

Online Learning Model

```
for $t \in \mathbb{N}$ do
    Play $p \sim x^{(t)}_{\mathcal{A}_k}$
    Discover $\ell^{(t)}_{\mathcal{A}_k}$
    Update
        $$x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$$
end for
```

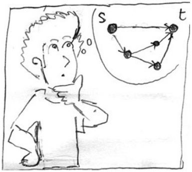


$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$ Sample $p \sim x^{(t)}_{\mathcal{A}_1}$ Discover $\ell^{(t)}_{\mathcal{A}_1}$ Update $x^{(t+1)}_{\mathcal{A}_1}$

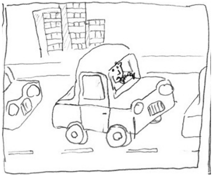
Online learning model

Online Learning Model

```
for $t \in \mathbb{N}$ do
    Play $p \sim x^{(t)}_{\mathcal{A}_k}$
    Discover $\ell^{(t)}_{\mathcal{A}_k}$
    Update
        $x^{(t+\mathbf{1})}_{\mathcal{A}_k} = u_k \left( x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k} \right)$
end for
```

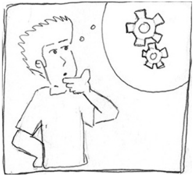


$x^{(t)}_{\mathcal{A}_\mathbf{1}} \in \Delta^{\mathcal{A}_\mathbf{1}}$



Sample $p \sim x^{(t)}_{\mathcal{A}_\mathbf{1}}$

Discover $\ell^{(t)}_{\mathcal{A}_\mathbf{1}}$



Update $x^{(t+\mathbf{1})}_{\mathcal{A}_\mathbf{1}}$

Main problem

Define class of dynamics $\mathcal{C}$ such that

$$u_k \in \mathcal{C} \ \forall k \Rightarrow x^{(t)} \rightarrow \mathcal{X}^\star$$

A brief review

Discrete time:

- Hannan consistency: [10]
- Hedge algorithm for two-player games: [9]
- Regret based algorithms: [11]
- Online learning in games: [7]
- Potential games: [19]

Continuous-time:

---

[10]James Hannan. Approximation to Bayes risk in repeated plays.
Contributions to the Theory of Games, 3:97–139, 1957

[9]Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights.
Games and Economic Behavior, 29(1):79–103, 1999

[11]Sergiu Hart and Andreu Mas-Colell. A general class of adaptive strategies.
Journal of Economic Theory, 98(1):26 – 54, 2001

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

[19]Jason R Marden, Gürdal Arslan, and Jeff S Shamma. Joint strategy fictitious play with inertia for potential games.
Automatic Control, IEEE Transactions on, 54(2):208–220, 2009

Outline

1. Introduction
2. Convergence of agent dynamics
3. Routing Examples
4. Estimation and optimal control

Nash equilibria, and the Rosenthal potential

Write
$x = (x_{\mathcal{A}_1}, \dots, x_{\mathcal{A}_K}) \in \Delta^{\mathcal{A}_1} \times \cdots \times \Delta^{\mathcal{A}_K}$
$\ell(x) = (\ell_{\mathcal{A}_1}(x), \dots, \ell_{\mathcal{A}_K}(x))$

Nash equilibrium

$x^\star$ is a Nash equilibrium if

$$\langle \ell(x^\star), x - x^\star \rangle \geq 0 \ \forall x \Leftrightarrow \forall k, \forall x_{\mathcal{A}_k}, \langle \ell_{\mathcal{A}_k}(x^\star), x_{\mathcal{A}_k} - x^\star_{\mathcal{A}_k} \rangle \geq 0$$

In words, for all k, paths in the support of $x^\star_{\mathcal{A}_k}$ have minimal loss.

Nash equilibria, and the Rosenthal potential

Write

$x = (x_{\mathcal{A}_1}, \dots, x_{\mathcal{A}_K}) \in \Delta^{\mathcal{A}_1} \times \dots \times \Delta^{\mathcal{A}_K}$

$\ell(x) = (\ell_{\mathcal{A}_1}(x), \dots, \ell_{\mathcal{A}_K}(x))$

Nash equilibrium

$x^\star$ is a Nash equilibrium if

$$\langle \ell(x^\star), x - x^\star \rangle \geq 0 \ \forall x \Leftrightarrow \forall k, \forall x_{\mathcal{A}_k}, \langle \ell_{\mathcal{A}_k}(x^\star), x_{\mathcal{A}_k} - x^\star_{\mathcal{A}_k} \rangle \geq 0$$

In words, for all k, paths in the support of $x^\star_{\mathcal{A}_k}$ have minimal loss.

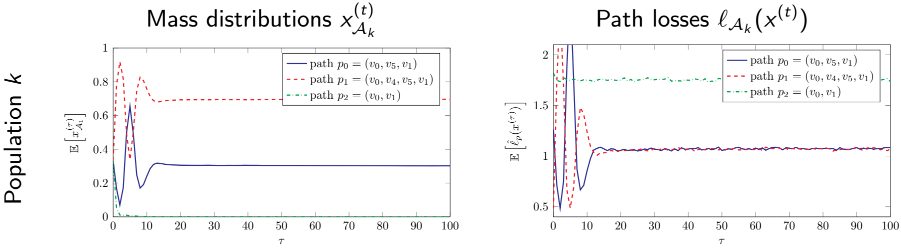


Figure: Population distributions and path losses

Nash equilibria, and the Rosenthal potential

Rosenthal potential

$\exists f$ convex such that

$$\nabla f(x) = \ell(x)$$

Then the set of Nash equilibria is

$$\mathcal{X}^\star = \operatorname*{arg\,min}_{x \in \Delta^{\mathcal{A}_1} \times \cdots \times \Delta^{\mathcal{A}_K}} f(x)$$

Nash equilibria, and the Rosenthal potential

Rosenthal potential

$\exists f$ convex such that

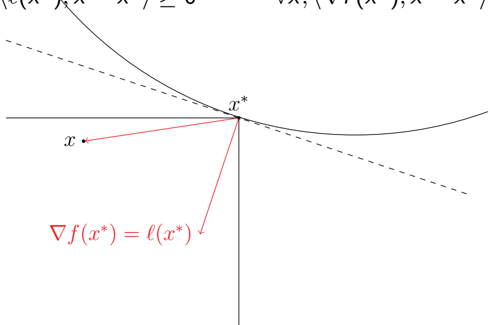
$$\nabla f(x) = \ell(x)$$

Then the set of Nash equilibria is

$$\mathcal{X}^\star = \operatorname*{arg\,min}_{x \in \Delta^{\mathcal{A}_1} \times \dots \times \Delta^{\mathcal{A}_K}} f(x)$$

Nash condition $\Leftrightarrow$ first order optimality

$\forall x, \langle \ell(x^\star), x - x^\star \rangle \geq 0$ $\qquad$ $\forall x, \langle \nabla f(x^\star), x - x^\star \rangle \geq 0$



Regret analysis

Technique 1: Regret analysis

Regret analysis

Cumulative regret

$$R_{\mathcal{A}_k}^{(t)} = \sup_{x_{\mathcal{A}_k} \in \Delta^{\mathcal{A}_k}} \sum_{\tau \le t} \left\langle x_{\mathcal{A}_k}^{(t)} - x_{\mathcal{A}_k}, \ell_{\mathcal{A}_k}(x^{(t)}) \right\rangle$$

“Online” optimality condition. Sublinear if $\limsup_t \frac{R_{\mathcal{A}_k}^{(t)}}{t} \le 0$.

Convergence of averages

$$\left[\forall k, R_{\mathcal{A}_k}^{(t)} \text{ is sublinear}\right] \Rightarrow \bar{x}^{(t)} \to \mathcal{X}^\star$$

$\bar{x}^{(t)} = \frac{1}{t} \sum_{\tau=1}^{t} x^{(\tau)}$. ▸ proof

Convergence of $\bar{x}^{(t)}$ Vs. convergence of $x^{(t)}$

Routing game example

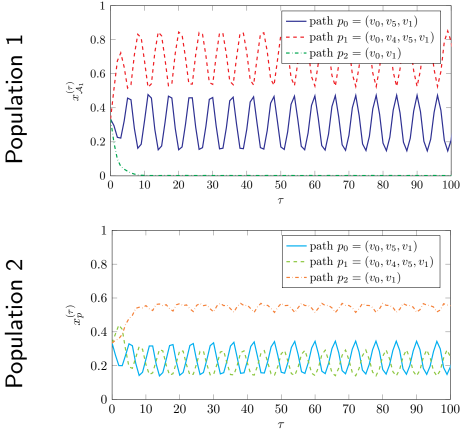


Figure: Population distributions

▸ More

Convergence of $\bar{x}^{(t)}$ Vs. convergence of $x^{(t)}$

Routing game example

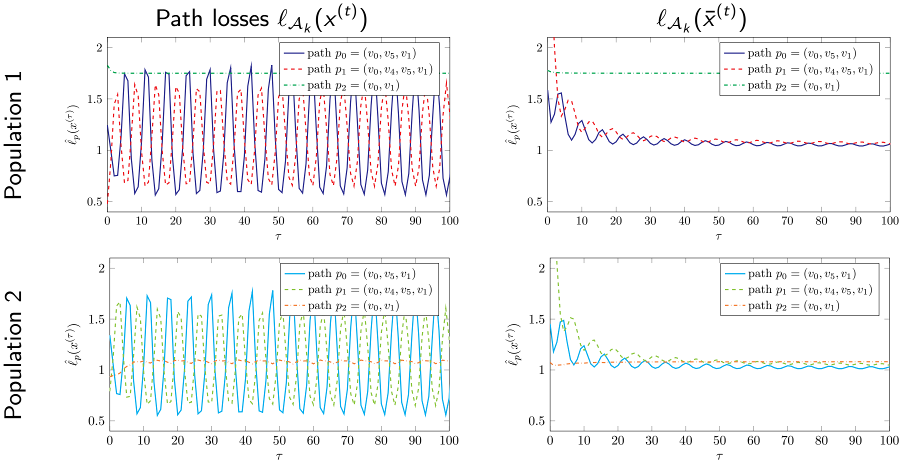


Figure: Path losses

▸ More

Stochastic approximation

Technique 2: Stochastic approximation

Stochastic approximation

Idea:

- View the learning dynamics as a discretization of an ODE.
- Study convergence of ODE.
- Relate convergence of discrete algorithm to convergence of ODE.

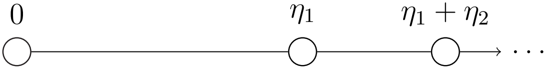


Figure: Underlying continuous time

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications.
Theory of Computing, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

Also known as

- Exponentially weighted average forecaster [7].

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications.
Theory of Computing, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

Also known as

- Exponentially weighted average forecaster [7].
- Multiplicative weights update [1].

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications.
Theory of Computing, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

Also known as

- Exponentially weighted average forecaster [7].
- Multiplicative weights update [1].
- Exponentiated gradient descent [12].

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.* Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, May 2003

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

Also known as

- Exponentially weighted average forecaster [7].
- Multiplicative weights update [1].
- Exponentiated gradient descent [12].
- Entropic descent [2].

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.* Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, May 2003

Example: the Hedge algorithm

Hedge algorithm

Update the distribution according to observed loss

$$x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t^k \ell_a^{(t)}}$$

Also known as

- Exponentially weighted average forecaster [7].
- Multiplicative weights update [1].
- Exponentiated gradient descent [12].
- Entropic descent [2].
- Log-linear learning [5], [18]

---

[7]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications.
Theory of Computing, 8(1):121–164, 2012

[12]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

The replicator ODE

In Hedge $x_p^{(t+1)} \propto x_p^{(t)} e^{-\eta_t^k \ell_p^{(t)}}$, take $\eta_t \to 0$.

Replicator equation [29]

$$\forall a \in \mathcal{A}_k, \frac{dx_a}{dt} = x_a \left(\langle \ell_{\mathcal{A}_k}(x), x_{\mathcal{A}_k} \rangle - \ell_a(x) \right) \tag{1}$$

---

[29] Jörgen W Weibull. *Evolutionary game theory.*
MIT press, 1997

[8] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing.
In *Algorithms–ESA 2004*, pages 323–334. Springer, 2004

The replicator ODE

In Hedge $x_p^{(t+1)} \propto x_p^{(t)} e^{-\eta_t^k \ell_p^{(t)}}$, take $\eta_t \to 0$.

Replicator equation [29]

$$\forall a \in \mathcal{A}_k, \frac{dx_a}{dt} = x_a \left(\langle \ell_{\mathcal{A}_k}(x), x_{\mathcal{A}_k} \rangle - \ell_a(x) \right) \quad (1)$$

Theorem: [8]

Every solution of the ODE (1) converges to the set of its stationary points.

---

[29] Jörgen W Weibull. *Evolutionary game theory.*
MIT press, 1997

[8] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing.
In *Algorithms–ESA 2004*, pages 323–334. Springer, 2004

AREP dynamics: Approximate REPlicator

Discretization of the continuous-time replicator dynamics

$$x_a^{(t+1)} - x_a^{(t)} = \eta_t x_a^{(t)} \left(\left\langle \ell_{\mathcal{A}_k}(x^{(t)}), x_{\mathcal{A}_k}^{(t)} \right\rangle - \ell_a(x^{(t)}) \right) + \eta_t U_a^{(t+1)}$$

---

[3] Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de probabilités XXXIII*, pages 1–68. Springer, 1999

AREP dynamics: Approximate REPlicator

Discretization of the continuous-time replicator dynamics

$$x_a^{(t+1)} - x_a^{(t)} = \eta_t x_a^{(t)} \left(\left\langle \ell_{\mathcal{A}_k}(x^{(t)}), x_{\mathcal{A}_k}^{(t)} \right\rangle - \ell_a(x^{(t)}) \right) + \eta_t U_a^{(t+1)}$$

- η_t discretization time steps.

---

[3] Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de probabilités XXXIII*, pages 1–68. Springer, 1999

AREP dynamics: Approximate REPlicator

Discretization of the continuous-time replicator dynamics

$$x_a^{(t+1)} - x_a^{(t)} = \eta_t x_a^{(t)} \left(\left\langle \ell_{\mathcal{A}_k}(x^{(t)}), x_{\mathcal{A}_k}^{(t)} \right\rangle - \ell_a(x^{(t)}) \right) + \eta_t U_a^{(t+1)}$$

- η_t discretization time steps.
- $(U^{(t)})_{t \geq 1}$ perturbations that satisfy for all $T > 0$, $\lim_{\tau_1 \to \infty} \max_{\tau_2 : \sum_{t=\tau_1}^{\tau_2} \eta_t < T} \left\| \sum_{t=\tau_1}^{\tau_2} \eta_t U^{(t+1)} \right\| = 0$

(a sufficient condition is that $\exists q \geq 2$: $\sup_\tau \mathbb{E} \| U^{(\tau)} \|^q < \infty$ and $\sum_\tau \eta_\tau^{1+\frac{q}{2}} < \infty$)

---

[3] Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de probabilités XXXIII*, pages 1–68. Springer, 1999

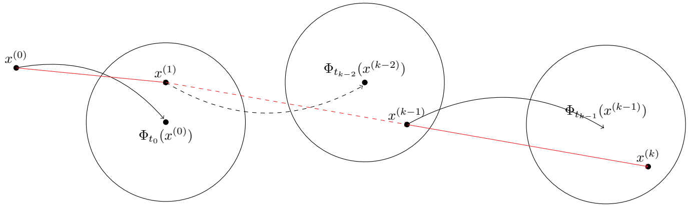
Convergence to Nash equilibria

Theorem [15]

Under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory.



- Use f as a Lyapunov function. ▸ proof details

---

[15] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. Learning nash equilibria in congestion games. *SIAM Journal on Control and Optimization (SICON), to appear*, 2014

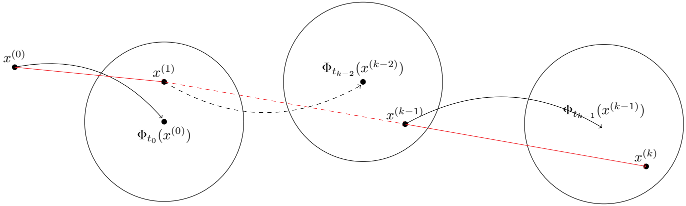
Convergence to Nash equilibria

Theorem [15]

Under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory.



- Use f as a Lyapunov function. proof details

However, No convergence rates.

---

[15] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. Learning nash equilibria in congestion games.
SIAM Journal on Control and Optimization (SICON), to appear, 2014

Stochastic convex optimization

Technique 3: (Stochastic) convex optimization

Stochastic convex optimization

Idea:

- View the learning dynamics as a distributed algorithm to minimize f.
- (More generally: distributed algorithm to find zero of a monotone operator).

Stochastic convex optimization

Idea:

- View the learning dynamics as a distributed algorithm to minimize f.
- (More generally: distributed algorithm to find zero of a monotone operator).
- Allows us to analyze convergence rates.

Stochastic convex optimization

Idea:

- View the learning dynamics as a distributed algorithm to minimize f.
- (More generally: distributed algorithm to find zero of a monotone operator).
- Allows us to analyze convergence rates.

Here:
Class of distributed optimization methods: stochastic mirror descent.

Stochastic Mirror Descent

$$\begin{array}{lll} \text{minimize} & f(x) & \text{convex function} \\ \text{subject to} & x \in \mathcal{X} \subset \mathbb{R}^d & \text{convex, compact set} \end{array}$$

---

[21]A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization.*
Wiley-Interscience series in discrete mathematics. Wiley, 1983

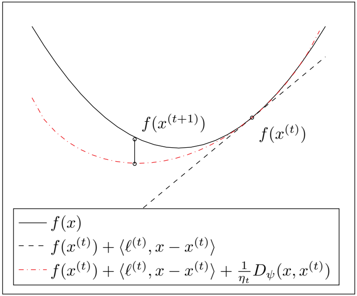
[20]A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming.
SIAM Journal on Optimization, 19(4):1574–1609, 2009

Stochastic Mirror Descent

$$
\begin{array}{lll}
\text{minimize} & f(x) & \text{convex function} \\
\text{subject to} & x \in \mathcal{X} \subset \mathbb{R}^d & \text{convex, compact set}
\end{array}
$$

Algorithm 2 MD Method with learning rates (η_t)

```
for $t \in \mathbb{N}$ do
    observe $\ell^{(t)} \in \partial f(x^{(t)})$
    $x^{(t+1)} = \arg\min_{x \in \mathcal{X}} \left\langle \ell^{(t)}, x \right\rangle + \frac{1}{\eta_t} D_\psi(x, x^{(t)})$
end for
```



- η_t: learning rate
- D_ψ: ▸ Bregman divergence

---

[21]A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization.*
Wiley-Interscience series in discrete mathematics. Wiley, 1983

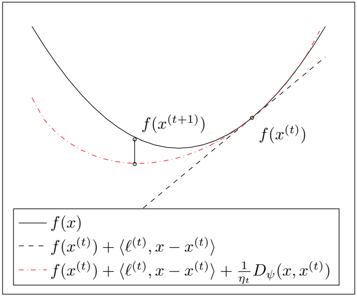
[20]A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming.
SIAM Journal on Optimization, 19(4):1574–1609, 2009

Stochastic Mirror Descent

$$\begin{array}{lll} \text{minimize} & f(x) & \text{convex function} \\ \text{subject to} & x \in \mathcal{X} \subset \mathbb{R}^d & \text{convex, compact set} \end{array}$$

Algorithm 2 MD Method with learning rates (η_t)

```
for $t \in \mathbb{N}$ do
    observe $\ell^{(t)}_{\mathcal{A}_k} \in \partial_{\mathcal{A}_k} f(x^{(t)})$
    $x^{(t+1)}_{\mathcal{A}_k} = \arg\min_{x \in \mathcal{X}_{\mathcal{A}_k}} \left\langle \ell^{(t)}_{\mathcal{A}_k}, x \right\rangle + \frac{1}{\eta^k_t} D_{\psi_k}(x, x^{(t)}_{\mathcal{A}_k})$
end for
```



- η_t: learning rate
- D_ψ: ▸ Bregman divergence

---

[21]A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization.*
Wiley-Interscience series in discrete mathematics. Wiley, 1983

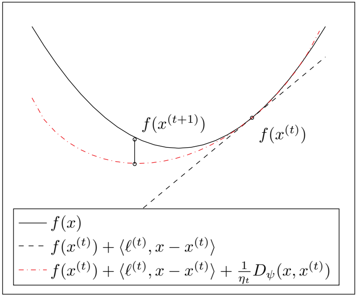
[20]A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming.
SIAM Journal on Optimization, 19(4):1574–1609, 2009

Stochastic Mirror Descent

$$\begin{array}{lll} \text{minimize} & f(x) & \text{convex function} \\ \text{subject to} & x \in \mathcal{X} \subset \mathbb{R}^d & \text{convex, compact set} \end{array}$$

Algorithm 2 SMD Method with learning rates (η_t)

```
for $t \in \mathbb{N}$ do
    observe $\hat{\ell}^{(t)}_{\mathcal{A}_k}$ with $\mathbb{E}\left[\hat{\ell}^{(t)}_{\mathcal{A}_k} | \mathcal{F}_{t-1}\right] \in \partial_{\mathcal{A}_k} f(x^{(t)})$
    $x^{(t+1)}_{\mathcal{A}_k} = \arg\min_{x \in \mathcal{X}_{\mathcal{A}_k}} \left\langle \hat{\ell}^{(t)}_{\mathcal{A}_k}, x \right\rangle + \frac{1}{\eta^k_t} D_{\psi_k}(x, x^{(t)}_{\mathcal{A}_k})$
end for
```



- η_t: learning rate
- D_ψ: ▸ Bregman divergence

---

[21]A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization.*
Wiley-Interscience series in discrete mathematics. Wiley, 1983

[20]A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming.
SIAM Journal on Optimization, 19(4):1574–1609, 2009

Convergence

$$d_\tau = D_\psi(\mathcal{X}^\star, x^{(\tau)}).$$

Main ingredient

$$\mathbb{E}\left[d_{\tau+1} | \mathcal{F}_{\tau-1}\right] \leq d_\tau - \eta_\tau (f(x^{(\tau)}) - f^\star) + \frac{\eta_\tau^2}{2\mu} \mathbb{E}\left[\|\hat{\ell}^{(\tau)}\|_*^2 | \mathcal{F}_{\tau-1}\right]$$

---

[22]H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications.
Optimizing Methods in Statistics, 1971

[6]Léon Bottou. Online algorithms and stochastic approximations.
1998

Convergence

$$d_\tau = D_\psi(\mathcal{X}^\star, x^{(\tau)}).$$

Main ingredient

$$\mathbb{E}\left[d_{\tau+1} | \mathcal{F}_{\tau-1}\right] \leq d_\tau - \eta_\tau \left(f(x^{(\tau)}) - f^\star\right) + \frac{\eta_\tau^2}{2\mu}\, \mathbb{E}\left[\|\hat{\ell}^{(\tau)}\|_*^2 | \mathcal{F}_{\tau-1}\right]$$

From here,

- Can show a.s. convergence $x^{(t)} \to \mathcal{X}^\star$ if $\sum \eta_t = \infty$ and $\sum \eta_t^2 < \infty$

---

[22]H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications.
Optimizing Methods in Statistics, 1971

[6]Léon Bottou. Online algorithms and stochastic approximations.
1998

Convergence

$$d_\tau = D_\psi(\mathcal{X}^\star, x^{(\tau)}).$$

Main ingredient

$$\mathbb{E}\left[d_{\tau+1} | \mathcal{F}_{\tau-1}\right] \le d_\tau - \eta_\tau (f(x^{(\tau)}) - f^\star) + \frac{\eta_\tau^2}{2\mu} \mathbb{E}\left[\|\hat{\ell}^{(\tau)}\|_*^2 | \mathcal{F}_{\tau-1}\right]$$

From here,

- Can show a.s. convergence $x^{(t)} \to \mathcal{X}^\star$ if $\sum \eta_t = \infty$ and $\sum \eta_t^2 < \infty$
  d_τ is an almost super martingale [22], [6]

---

[22]H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications.
Optimizing Methods in Statistics, 1971

[6]Léon Bottou. Online algorithms and stochastic approximations.
1998

Convergence

$$d_\tau = D_\psi(\mathcal{X}^\star, x^{(\tau)}).$$

Main ingredient

$$\mathbb{E}\left[d_{\tau+1} | \mathcal{F}_{\tau-1}\right] \leq d_\tau - \eta_\tau (f(x^{(\tau)}) - f^\star) + \frac{\eta_\tau^2}{2\mu} \mathbb{E}\left[\|\hat{\ell}^{(\tau)}\|_*^2 | \mathcal{F}_{\tau-1}\right]$$

From here,

- Can show a.s. convergence $x^{(t)} \to \mathcal{X}^\star$ if $\sum \eta_t = \infty$ and $\sum \eta_t^2 < \infty$ d_τ is an almost super martingale [22], [6]

  Deterministic version:
  If $d_{\tau+1} \leq d_\tau - a_\tau + b_\tau$, and $\sum b_\tau < \infty$, then (d_τ) converges.

---

[22]H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. *Optimizing Methods in Statistics*, 1971

[6]Léon Bottou. Online algorithms and stochastic approximations. 1998

Convergence

- To show convergence $\mathbb{E}\left[f(x^{(t)})\right] \to f^\star$, generalize the technique of Shamir et al. [25] (for SGD, $\alpha = \frac{1}{2}$).

Convergence of Distributed Stochastic Mirror Descent

For $\eta_t^k = \frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0, 1)$,

$$\mathbb{E}\left[f(x^{(t)})\right] - f^\star = \mathcal{O}\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$$

Non-smooth, non-strongly convex.

▸ More details

---

[25]Ohad Shamir and Tong Zhang. Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes.
In *ICML*, pages 71–79, 2013

[14] Syrine Krichene, Walid Krichene, Roy Dong, and Alexandre Bayen. Convergence of heterogeneous distributed learning in stochastic routing games.
In *53rd Allerton Conference on Communication, Control and Computing*, 2015

Summary

- Regret analysis: convergence of $\bar{x}^{(t)}$
- Stochastic approximation: almost sure convergence of $x^{(t)}$
- Stochastic convex optimization: almost sure convergence, $\mathbb{E}\left[f(x^{(t)})\right] \to f^\star$, $\mathbb{E}\left[D_\psi(x^\star, x^{(t)})\right] \to 0$, convergence rates.

Outline

1. Introduction
2. Convergence of agent dynamics
3. Routing Examples
4. Estimation and optimal control

Application to the routing game

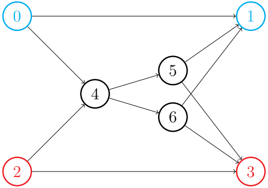


Figure: Example with strongly convex potential.

- Centered Gaussian noise on edges.
- Population 1: Hedge with $\eta_t^1 = t^{-1}$
- Population 2: Hedge with $\eta_t^2 = t^{-1}$

Routing game with strongly convex potential

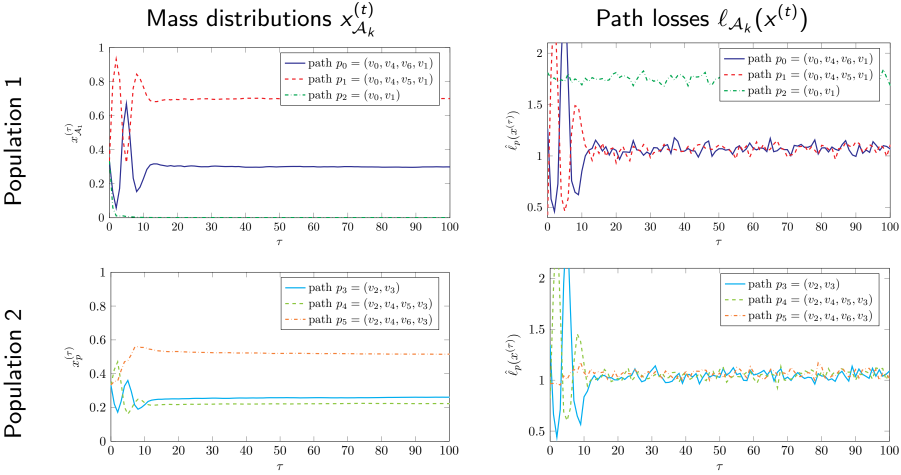


Figure: Population distributions and noisy path losses

Routing game with strongly convex potential

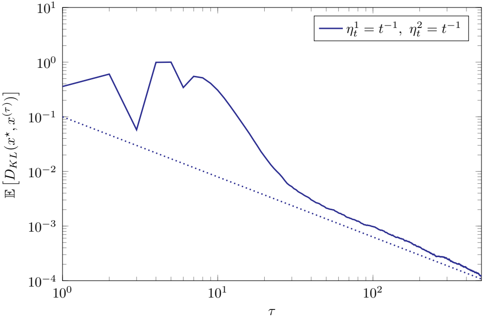


Figure: Distance to equilibrium.

For $\eta_t^k = \frac{\theta_k}{\ell_f t^{\alpha_k}},\ \alpha_k \in (0,1]$, $\mathbb{E}\left[D_\psi(x^\star, x^{(t)})\right] = O(\sum_k t^{-\alpha_k})$

Routing game with weakly convex potential

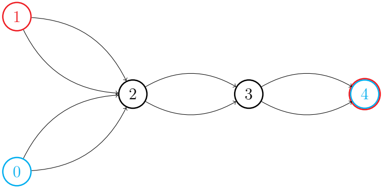


Figure: A weakly convex example.

Routing game with weakly convex potential

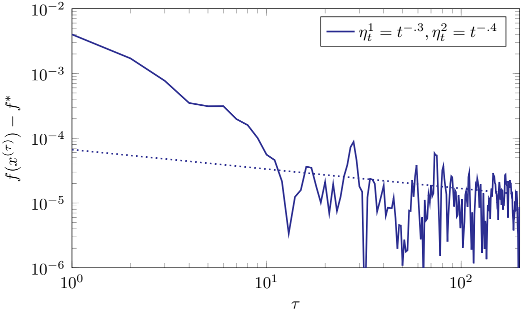


Figure: Potential values.

For $\frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0,1)$, $\mathbb{E}\left[f(x^{(t)})\right] - f^\star = O\left(\sum_k \frac{\log t}{t^{\mathbf{min}(\alpha_k, \mathbf{1}-\alpha_k)}}\right)$

Routing game with weakly convex potential

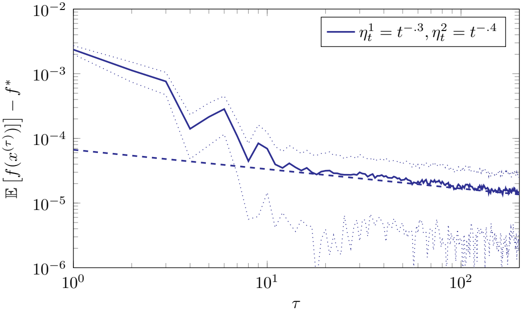


Figure: Potential values.

For $\frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0,1)$, $\mathbb{E}\left[f(x^{(t)})\right] - f^\star = O\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$

Routing game with weakly convex potential

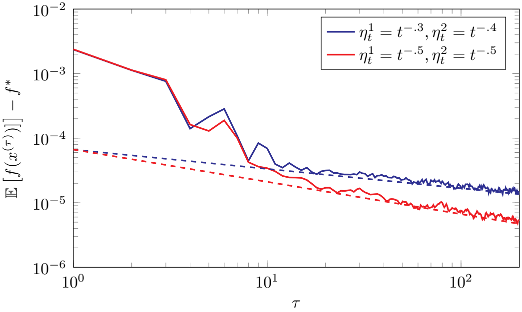


Figure: Potential values.

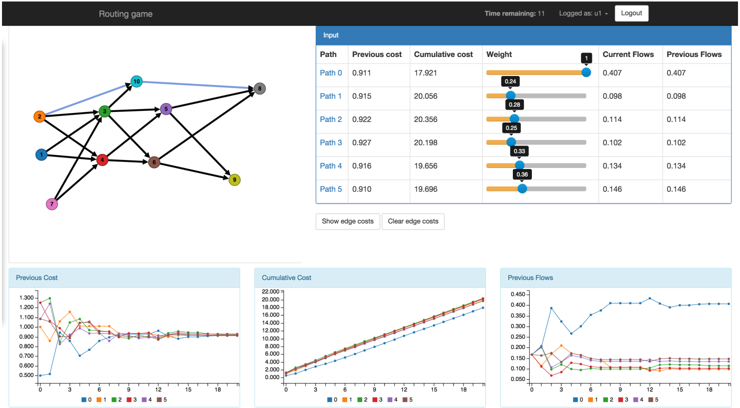
For $\frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0,1)$, $\mathbb{E}\left[f(x^{(t)})\right] - f^\star = O\left(\sum_k \frac{\log t}{t^{\mathbf{min}(\alpha_k,\mathbf{1}-\alpha_k)}}\right)$

Outline

1. Introduction
2. Convergence of agent dynamics
3. Routing Examples
4. Estimation and optimal control

A routing experiment

- Interface for the routing game.
- Used to collect sequence of decisions $\bar{x}^{(t)}$.



Routing game

Time remaining: 11 Logged as: u1

Input

Path	Previous cost	Cumulative cost	Weight	Current Flows	Previous Flows
Path 0	0.911	17.921	1	0.407	0.407
Path 1	0.915	20.056	0.24	0.098	0.098
Path 2	0.922	20.356	0.28	0.114	0.114
Path 3	0.927	20.198	0.25	0.102	0.102
Path 4	0.916	19.656	0.33	0.134	0.134
Path 5	0.910	19.696	0.36	0.146	0.146

Figure: Interface for the routing game experiment.

A routing experiment

- Interface for the routing game.
- Used to collect sequence of decisions $\bar{x}^{(t)}$.

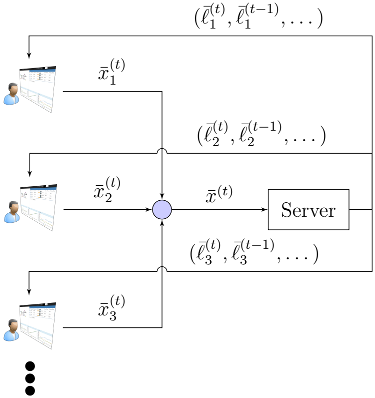


Figure: Interface for the routing game experiment.

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

---

[17]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

Mirror descent model

Estimate the learning rate in the mirror descent model

$$x^{(t+1)}(\eta) = \underset{x \in \Delta^{\mathcal{A}_k}}{\arg\min} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

---

[17]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

Mirror descent model

Estimate the learning rate in the mirror descent model

$$x^{(t+1)}(\eta) = \arg\min_{x \in \Delta^{\mathcal{A}_k}} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

Then $d(\eta) = D_{KL}(\bar{x}^{(t+1)}, x^{(t+1)}(\eta))$ is a convex function. Can minimize it to estimate $\eta_k^{(t)}$.

---

[17]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

Preliminary results

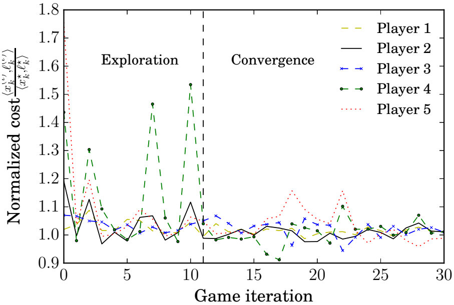


Figure: Costs of each player (normalized by the equilibrium cost)

Preliminary results

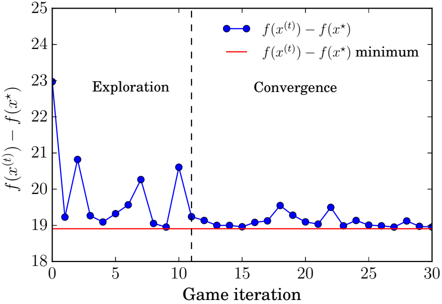


Figure: Potential function $f(x^{(t)}) - f^\star$.

Preliminary results

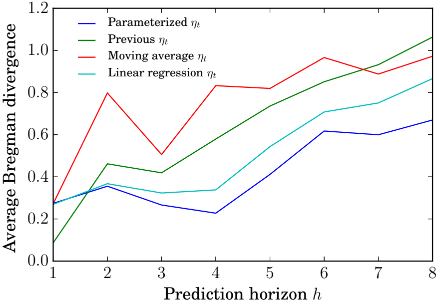


Figure: Average KL divergence between predicted distributions and actual distributions, as a function of the prediction horizon h.

Optimal routing with learning dynamics

Assumptions

- A central authority has control over a fraction of traffic: $u^{(t)} \in \alpha_1 \Delta^{\mathcal{A}_1} \times \cdots \times \alpha_K \Delta^{\mathcal{A}_K}$
- Remaining traffic follows learning dynamics: $x^{(t)} \in (1-\alpha_1)\Delta^{\mathcal{A}_1} \times \cdots \times (1-\alpha_K)\Delta^{\mathcal{A}_K}$

Optimal routing under selfish learning constraints

$$\begin{array}{ll} \text{minimize}_{u^{(1:T)}, x^{(1:T)}} & \sum_{t=1}^{T} J(x^{(t)}, u^{(t)}) \\ \text{subject to} & x^{(t+1)} = u(x^{(t)} + u^{(t)}, \ell(x^{(t)} + u^{(t)})) \end{array}$$

Solution methods

- Greedy method: Approximate the problem with a sequence of convex problems.

$$\text{minimize}_{u^{(t)}} J(u(x^{(t-1)}, u^{(t-1)}), u^{(t)})$$

Solution methods

- Greedy method: Approximate the problem with a sequence of convex problems.

$$\text{minimize}_{u^{(t)}} J(u(x^{(t-1)}, u^{(t-1)}), u^{(t)})$$

- Mirror descent with the adjoint method.

Adjoint method

$$\begin{aligned} \text{minimize}_{u} \quad & J(u, x) \\ \text{subject to} \quad & H(x, u) = 0 \end{aligned}$$

equivalent to

$$\text{minimize} \quad J(u, X(u))$$

Then perform mirror descent on this function of u.

A simple example

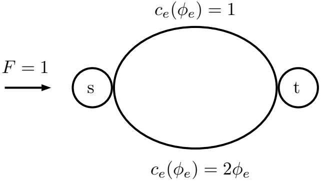


Figure: Simple Pigou network used for the numerical experiment.

- Social optimum: $(\frac{3}{4}, \frac{1}{4})$
- Nash equilibrium $(\frac{1}{2}, \frac{1}{2})$
- Control over $\alpha = \frac{1}{2}$ of traffic

A simple example

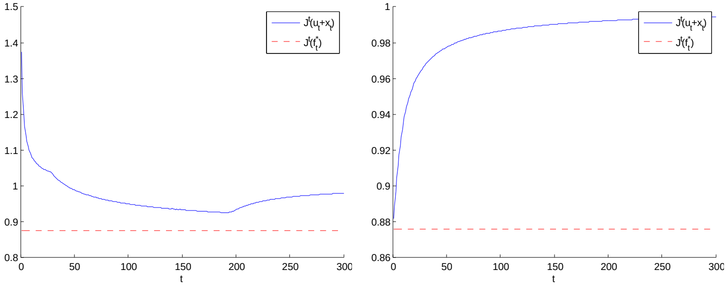


Figure: Social cost $J^{(t)}$ over time induced by adjoint solution (left) and the greedy solution (right). The dashed line shows the social optimal allocation.

A simple example

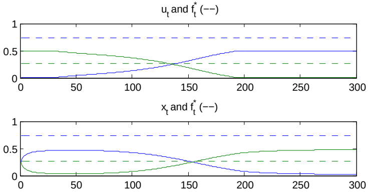


Figure: Adjoint controlled flows (top), selfish flows (bottom). The green lines correspond to the top path, and the blue lines to the bottom path. The dashed lines show the social optimal flows $x^{SO} = (\frac{3}{4}, \frac{1}{4})$.

Application to the L.A. highway network

- Simplified model of the L.A. highway network.
- Cost functions uses the B.P.R. function, calibrated using the work of [28].

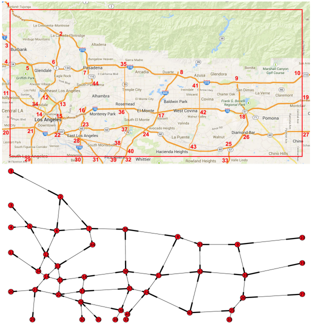


Figure: Los Angeles highway network.

---

[28]J. Thai, R. Hariss, and A. Bayen. A multi-convex approach to latency inference and control in traffic equilibria from sparse data.
In *American Control Conference (ACC), 2015*, pages 689–695, July 2015

Application to the L.A. highway network

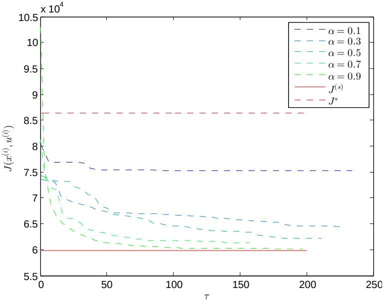
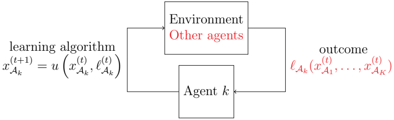


Figure: Average delay without control (dashed), with full control (solid), and different values of α.

---

[27]Milena Suarez, Walid Krichene, and Alexandre Bayen. Optimal routing under hedge response.
Transactions on Control of Networked Systems (TCNS), in preparation, 2015

Summary



learning algorithm

$x_{\mathcal{A}_k}^{(t+1)} = u\left(x_{\mathcal{A}_k}^{(t)}, \ell_{\mathcal{A}_k}^{(t)}\right)$

Environment

Other agents

Agent k

outcome

$\ell_{\mathcal{A}_k}(x_{\mathcal{A}_1}^{(t)}, \dots, x_{\mathcal{A}_K}^{(t)})$

Figure: Coupled sequential decision problems.

- Simple model for distributed learning.
- Techniques for design / analysis of learning dynamics.
- Estimation of learning dynamics.
- Optimal control.

Ongoing / future work

- Larger classes of games.
- Acceleration of learning dynamics (Nesterov's method).
    - In continuous time: accelerated replicator dynamics.
    - In discrete time: convergence in $\mathcal{O}(1/t^2)$ instead of $\mathcal{O}(1/t)$.
    - Heuristic to remove oscillations.
- Applications to load balancing.
- Other control schemes: tolling / incentivization.

Ongoing / future work

- Larger classes of games.
- Acceleration of learning dynamics (Nesterov's method).
  - In continuous time: accelerated replicator dynamics.
  - In discrete time: convergence in $\mathcal{O}(1/t^2)$ instead of $\mathcal{O}(1/t)$.
  - Heuristic to remove oscillations.
- Applications to load balancing.
- Other control schemes: tolling / incentivization.

Thank you!

`eecs.berkeley.edu/∼walid/`

References I

[1] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[2] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3): 167–175, May 2003.

[3] Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de probabilités XXXIII*, pages 1–68. Springer, 1999.

[4] Avrim Blum, Eyal Even-Dar, and Katrina Ligett. Routing without regret: on convergence to nash equilibria of regret-minimizing algorithms in routing games. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, PODC '06, pages 45–52, New York, NY, USA, 2006. ACM.

[5] Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387 – 424, 1993. ISSN 0899-8256.

[6] Léon Bottou. Online algorithms and stochastic approximations. 1998.

[7] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

References II

[8] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing. In *Algorithms–ESA 2004*, pages 323–334. Springer, 2004.

[9] Yoav Freund and Robert E Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1):79–103, 1999.

[10] James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957.

[11] Sergiu Hart and Andreu Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98(1):26 – 54, 2001.

[12] Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132 (1):1 – 63, 1997.

[13] Robert Kleinberg, Georgios Piliouras, and Eva Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 533–542. ACM, 2009.

References III

[14] Syrine Krichene, Walid Krichene, Roy Dong, and Alexandre Bayen. Convergence of heterogeneous distributed learning in stochastic routing games. In *53rd Allerton Conference on Communication, Control and Computing*, 2015.

[15] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. Learning nash equilibria in congestion games. *SIAM Journal on Control and Optimization (SICON), to appear*, 2014.

[16] Walid Krichene, Syrine Krichene, and Alexandre Bayen. Convergence of mirror descent dynamics in the routing game. In *European Control Conference (ECC)*, 2015.

[17] Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game. In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015.

[18] Jason R. Marden and Jeff S. Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75(2):788–808, 2012.

References IV

[19] Jason R Marden, Gürdal Arslan, and Jeff S Shamma. Joint strategy fictitious play with inertia for potential games. *Automatic Control, IEEE Transactions on*, 54(2):208–220, 2009.

[20] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.

[21] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience series in discrete mathematics. Wiley, 1983.

[22] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. *Optimizing Methods in Statistics*, 1971.

[23] William H Sandholm. Potential games with continuous player sets. *Journal of Economic Theory*, 97(1):81–108, 2001.

[24] William H. Sandholm. *Population games and evolutionary dynamics*. Economic learning and social evolution. Cambridge, Mass. MIT Press, 2010. ISBN 978-0-262-19587-4.

References V

[25] Ohad Shamir and Tong Zhang. Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes. In *ICML*, pages 71–79, 2013.

[26] Weijie Su, Stephen Boyd, and Emmanuel Candes. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights. In *NIPS*, 2014.

[27] Milena Suarez, Walid Krichene, and Alexandre Bayen. Optimal routing under hedge response. *Transactions on Control of Networked Systems (TCNS), in preparation*, 2015.

[28] J. Thai, R. Hariss, and A. Bayen. A multi-convex approach to latency inference and control in traffic equilibria from sparse data. In *American Control Conference (ACC), 2015*, pages 689–695, July 2015.

[29] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.

Continuous time model

▸ Back

Continuous-time learning model

$$\dot{x}_{\mathcal{A}_k}(t) = v_k\left(x^{(t)}_{\mathcal{A}_k}, \ell_{\mathcal{A}_k}(x^{(t)})\right)$$

- Evolution in populations: [24]
- Convergence in potential games under dynamics which satisfy a positive correlation condition [23]
- Replicator dynamics for the congestion game [8] and in evolutionary game theory [29]
- No-regret dynamics for two player games [11]

---

[24]William H. Sandholm. *Population games and evolutionary dynamics.* Economic learning and social evolution. Cambridge, Mass. MIT Press, 2010. ISBN 978-0-262-19587-4

[23] William H Sandholm. Potential games with continuous player sets. *Journal of Economic Theory*, 97(1):81–108, 2001

[29] Jörgen W Weibull. *Evolutionary game theory.* MIT press, 1997

[8] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing. In *Algorithms–ESA 2004*, pages 323–334. Springer, 2004

[11] Sergiu Hart and Andreu Mas-Colell. A general class of adaptive strategies. *Journal of Economic Theory*, 98(1):26 – 54, 2001

Oscillating example

▸ Back

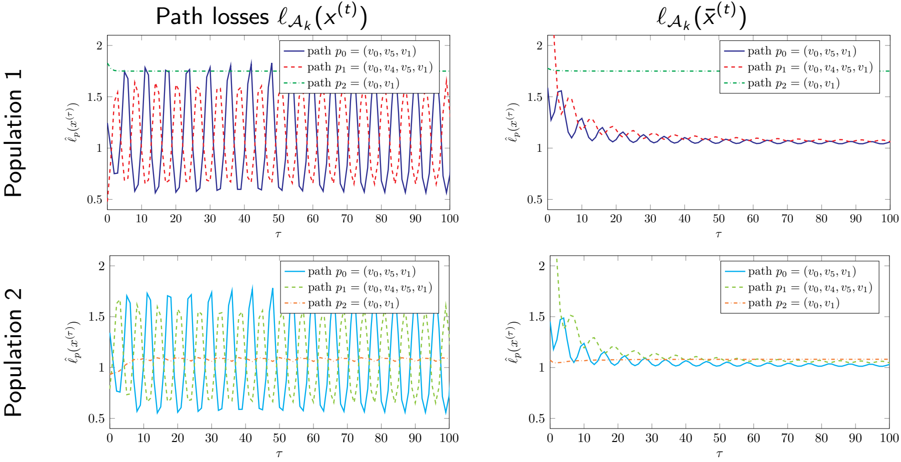


Figure: Path losses

Oscillating example

▸ Back

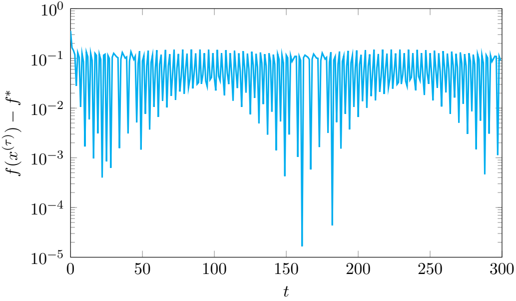


Figure: Potentials

Oscillating example

Back

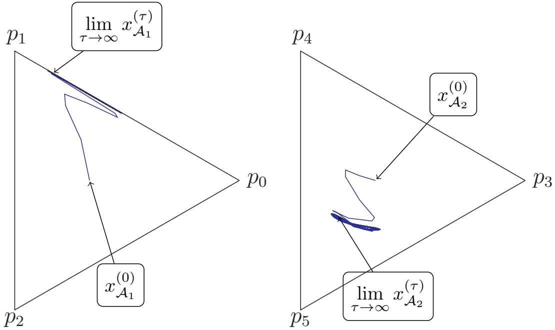


Figure: Trajectories in the simplex

Regret [10]

▸ Back Cumulative regret

$$R_{\mathcal{A}_k}^{(t)} = \sup_{x_{\mathcal{A}_k} \in \Delta^{\mathcal{A}_k}} \sum_{\tau \leq t} \left\langle x_{\mathcal{A}_k}^{(t)} - x_{\mathcal{A}_k}, \ell_{\mathcal{A}_k}(x^{(t)}) \right\rangle$$

Convergence of averages

$$\forall k, \ \limsup_t \frac{R_{\mathcal{A}_k}^{(t)}}{t} \leq 0 \Rightarrow \bar{x}^{(t)} = \frac{1}{t} \sum_{\tau \leq t} x^{(\tau)} \to \mathcal{X}^\star$$

By convexity of f,

$$f\left(\frac{1}{t}\sum_{\tau \leq t} x^{(\tau)}\right) - f(x) \leq \frac{1}{t}\sum_{\tau \leq t} f(x^{(\tau)}) - f(x)$$

$$\leq \frac{1}{t}\sum_{\tau \leq t} \left\langle \ell(x^{(t)}), x^{(t)} - x \right\rangle = \sum_{k=1}^{K} \frac{R_{\mathcal{A}_k}^{(t)}}{t}$$

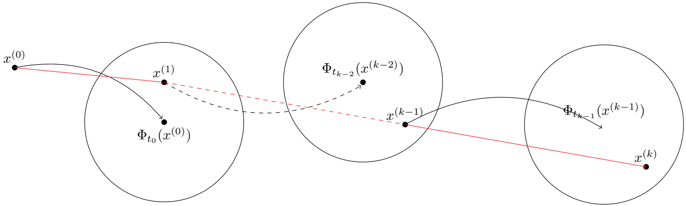
---

[10] James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957

AREP convergence proof

▸ Back

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory.



- The set of limit points of an APT is internally chain transitive ICT.
- If Γ is compact invariant, and has a Lyapunov function f with $\operatorname{int} f(\Gamma) = \emptyset$, then $\forall L$ ICT, Γ, and f is constant on L.
- In particular, f is constant on $L(x^{(t)})$, so $f(x^{(t)})$ converges.

Bregman Divergence

▸ Back

Bregman Divergence

Strongly convex function ψ

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y\rangle$$

---

[2] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

Bregman Divergence

▸ Back

Bregman Divergence

Strongly convex function ψ

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle$$

Example [2]: when $\mathcal{X} = \Delta^d$

- $\psi(x) = -H(x) = \sum_a x_a \ln x_a$
- $D_\psi(x, y) = D_{KL}(x, y) = \sum_a x_a \ln \frac{x_a}{y_a}$
- The MD update has closed form solution

$$x^{(t+1)} \propto x_a^{(t)} e^{-\eta_t g_a^{(t)}}$$

A.k.a. Hedge algorithm, exponential weights.

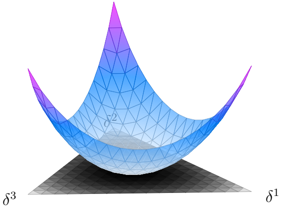


Figure: KL divergence

---

[2] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

A bounded entropic divergence

▸ Back

- $\mathcal{X} = \Delta$
- $D_{KL}(x, y) = \sum_{i=1}^{d} x_i \ln \frac{x_i}{y_i}$ is unbounded.

A bounded entropic divergence

Back

- $\mathcal{X} = \Delta$
- $D_{KL}(x,y) = \sum_{i=1}^{d} x_i \ln \frac{x_i}{y_i}$ is unbounded.
- Define $D_{KL}^{\epsilon}(x,y) = \sum_{i=1}^{d} (x_i + \epsilon) \ln \frac{x_i+\epsilon}{y_i+\epsilon}$

Proposition

- *D_{KL}^{ϵ} is $\frac{1}{1+d\epsilon}$-strongly convex w.r.t. $\|\cdot\|_1$*
- *D_{KL}^{ϵ} is bounded by $(1+d\epsilon)\ln\frac{1+\epsilon}{\epsilon}$.*

Convergence of DMD

▸ Back

Theorem: Convergence of DMD [16]

Suppose f has L Lipschitz gradient. Then under the MD class with $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$,

$$f(x^{(t)}) - f^\star = O\left(\frac{\sum_{\tau \le t} \eta_\tau}{t} + \frac{1}{\eta_t} + \frac{1}{t}\right)$$

$$\frac{1}{t}\sum_{\tau \le t} f(x^{(t)}) - f^\star \le \sum_k \frac{L_k^2}{2\ell_{\psi_k}} \sum_{\tau \le t} \eta_\tau^k + \frac{D_k}{\eta_t^k}$$

and

$$f(x^{(t)}) - f^\star \le \frac{1}{t}\sum_{\tau \le t} f(x^{(t)}) - f^\star + O\left(\frac{1}{t}\right)$$

Convergence in DSMD

▸ Back

Regret bound [14]

SMD method with (η_t). $\forall t_2 > t_1 \geq 0$ and $\mathcal{F}_{t_1}$-measurable x,

$$\sum_{\tau=t_1}^{t_2} \mathbb{E}\left[\left\langle g^{(\tau)}, x^{(\tau)} - x \right\rangle\right] \leq \frac{\mathbb{E}\left[D_\psi(x, x^{(t_1)})\right]}{\eta_{t_1}} + D\left(\frac{1}{\eta_{t_2}} - \frac{1}{\eta_{t_1}}\right) + \frac{G}{2\ell_\psi} \sum_{\tau=t_1}^{t_2} \eta_\tau$$

Strongly convex case:

$$\mathbb{E}[D_\psi(x^\star, x^{(t+1)})] \leq (1 - 2\ell_f \eta_t)\, \mathbb{E}[D_\psi(x^\star, x^{(t)})] + \frac{G}{2\ell_\psi}\eta_t^2$$

Convergence in DSMD

▸ Back Weakly convex case:

Theorem [14]

Distributed SMD such that $\eta_t^p = \frac{\theta_p}{t^{\alpha_p}}$ with $\alpha_p \in (0,1)$. Then

$$\mathbb{E}\left[f(x^{(t)})\right] - f(x^\star) \le \left(1 + \sum_{i=1}^{t} \frac{1}{i}\right) \sum_{k \in \mathcal{A}} \left(\frac{1}{t^{1-\alpha_k}} \frac{D}{\theta_k} + \frac{\theta_k G}{2\ell_\psi(1-\alpha_k)} \frac{1}{t^{\alpha_k}} \right)$$
$$= O\left(\frac{\log t}{t^{\min(\min_k \alpha_k, 1 - \max_k \alpha_k)}} \right)$$

Define $S_i = \frac{1}{i+1} \sum_{t-i}^{t} \mathbb{E}[f(x^{(\tau)})]$

Show $S_{i-1} \le S_i + \left(\frac{D}{\theta} \frac{1}{t^{\alpha - 1}} + \frac{\theta G}{2\ell_\psi(1-\alpha)} \frac{1}{t^\alpha} \right) \frac{1}{i}$

---

[14] Syrine Krichene, Walid Krichene, Roy Dong, and Alexandre Bayen. Convergence of heterogeneous distributed learning in stochastic routing games.
In *53rd Allerton Conference on Communication, Control and Computing*, 2015

Stochastic mirror descent in machine learning

▸ Back

Large scale learning:

$$\begin{aligned} \text{minimize}_x \quad & \sum_{i=1}^{N} f_i(x) \\ \text{subject to} \quad & x \in \mathcal{X} \end{aligned}$$

N very large. Gradient prohibitively expensive to compute exactly. Instead, compute

$$\hat{g}(x^{(t)}) = \sum_{i \in \mathcal{I}} \nabla f_i(x^{(t)})$$

with $\mathcal{I}$ random subset of $\{1, \dots, N\}$.