

On the Convergence of No-regret Learning in Selfish Routing

ICML 2014 - Beijing

Walid Krichene ¹ Benjamin Drighès ² Alexandre Bayen ³

UC Berkeley

Ecole Polytechnique

June 23, 2014

¹walid@cs.berkeley.edu

²benjamin.drighes@polytechnique.edu

³bayen@berkeley.edu

Introduction

Routing game: players choose routes.

Population distributions: $\mu^{(t)} \in \Delta^{\mathcal{P}_1} \times \dots \times \Delta^{\mathcal{P}_k}$

Nash equilibria: \mathcal{N}

Classes of dynamics such that $\mu^{(t)} \rightarrow \mathcal{N}$

Introduction

Routing game: players choose routes.

Population distributions: $\mu^{(t)} \in \Delta^{\mathcal{P}_1} \times \dots \times \Delta^{\mathcal{P}_K}$

Nash equilibria: \mathcal{N}

Classes of dynamics such that $\mu^{(t)} \rightarrow \mathcal{N}$

- Under no-regret dynamics, $\bar{\mu}^{(t)} = \frac{1}{t} \sum_{\tau \leq t} \mu^{(\tau)} \rightarrow \mathcal{N}$.
- New classes with $\mu^{(t)} \rightarrow \mathcal{N}$?

Outline

- 1 Online learning in the routing game
- 2 Convergence of average strategies $\bar{\mu}^{(t)}$
- 3 Convergence of strategies $\mu^{(t)}$

Routing game

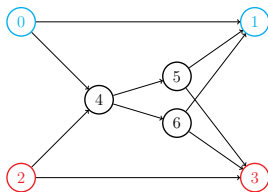


Figure: Example network

- Directed graph (V, E)
- Population \mathcal{X}_k : paths \mathcal{P}_k

Routing game

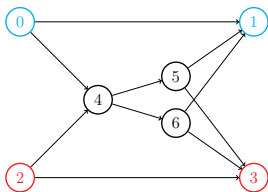


Figure: Example network

- Directed graph (V, E)
- Population \mathcal{X}_k : paths \mathcal{P}_k
- Player $x \in \mathcal{X}_k$: distribution over paths $\pi(x) \in \Delta^{\mathcal{P}_k}$,
- Population distribution over paths $\mu^k \in \Delta^{\mathcal{P}_k}$, $\mu^k = \int_{\mathcal{X}_k} \pi(x) dm(x)$
- Loss on path p : $\ell_p^k(\mu)$

Routing game

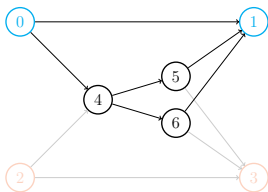
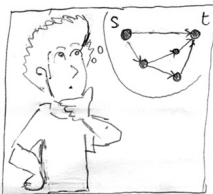


Figure: Example network

- Directed graph (V, E)
- Population \mathcal{X}_k : paths \mathcal{P}_k
- Player $x \in \mathcal{X}_k$: distribution over paths $\pi(x) \in \Delta^{\mathcal{P}_k}$,
- Population distribution over paths $\mu^k \in \Delta^{\mathcal{P}_k}$, $\mu^k = \int_{\mathcal{X}_k} \pi(x) dm(x)$
- Loss on path p : $\ell_p^k(\mu)$

Online learning model

$\pi^{(t)} \in \Delta^{\mathcal{P}_1}$



Sample $p \sim \pi^{(t)}$



Discover $\ell^{(t)} \in [0, 1]^{\mathcal{P}_1}$



Update $\pi^{(t+1)}$



The Hedge algorithm

Hedge algorithm

- Update the distribution according to observed loss

$$\pi_p^{(t+1)} \propto \pi_p^{(t)} e^{-\eta_t \ell_p^{k(t)}}$$

Nash equilibria

Nash equilibrium

$\mu \in \mathcal{N}$ if $\forall k, \forall p \in \mathcal{P}_k$ with positive mass,

$$\ell_p^k(\mu) \leq \ell_{p'}^k(\mu) \quad \forall p' \in \mathcal{P}_k$$

- How to compute Nash equilibria?

Nash equilibria

Nash equilibrium

$\mu \in \mathcal{N}$ if $\forall k, \forall p \in \mathcal{P}_k$ with positive mass,

$$\ell_p^k(\mu) \leq \ell_{p'}^k(\mu) \quad \forall p' \in \mathcal{P}_k$$

- How to compute Nash equilibria? **Convex formulation**

Nash equilibria

Convex potential function

$$V(\mu) = \sum_e \int_0^{(M\mu)_e} c_e(u) du$$

- V is convex.
 - $\nabla_{\mu^k} V(\mu) = \ell^k(\mu)$.
 - Minimizer not unique.
- How do players find a Nash equilibrium?
 - Iterative play.

Nash equilibria

Convex potential function

$$V(\mu) = \sum_e \int_0^{(M\mu)_e} c_e(u) du$$

- V is convex.
 - $\nabla_{\mu^k} V(\mu) = \ell^k(\mu)$.
 - Minimizer not unique.
- How do players find a Nash equilibrium?
 - Iterative play.
 - Ideally: **distributed**, and has reasonable information requirements.

Assume sublinear regret dynamics

- Losses are in $[0, 1]$.
- Expected loss is $\langle \pi^{(t)}(x), \ell^k(\mu^{(t)}) \rangle$
- Discounted regret

$$\bar{r}^{(T)}(x) = \frac{\sum_{t \leq T} \gamma_t \langle \pi^{(t)}(x), \ell^k(\mu^{(t)}) \rangle - \min_p \sum_{t \leq T} \gamma_t \ell_p^{k(t)}}{\sum_{t \leq T} \gamma_t}$$

Assume sublinear regret dynamics

- Losses are in $[0, 1]$.
- Expected loss is $\langle \pi^{(t)}(x), \ell^k(\mu^{(t)}) \rangle$
- Discounted regret

$$\bar{r}^{(T)}(x) = \frac{\sum_{t \leq T} \gamma_t \langle \pi^{(t)}(x), \ell^k(\mu^{(t)}) \rangle - \min_p \sum_{t \leq T} \gamma_t \ell_p^{k(t)}}{\sum_{t \leq T} \gamma_t}$$

Assumptions

- $\gamma(t) > 0$
- $\gamma(t) \downarrow 0$
- $\sum_t \gamma(t) = \infty$

Convergence to Nash equilibria

Population regret

$$\bar{r}^{k(T)} = \frac{1}{m(\mathcal{X}_k)} \int_{\mathcal{X}_k} \bar{r}^{(T)}(x) dm(x)$$

Convergence of averages to Nash equilibria

If an update has **sublinear population regret**, then

$\bar{\mu}^{(T)} = \sum_{t \leq T} \gamma_t \mu^{(t)} / \sum_{t \leq T} \gamma_t$ converges

$$\lim_{T \rightarrow \infty} d(\bar{\mu}^{(T)}, \mathcal{N}) = 0$$

Convergence to Nash equilibria

Population regret

$$\bar{r}^{k(T)} = \frac{1}{m(\mathcal{X}_k)} \int_{\mathcal{X}_k} \bar{r}^{(T)}(x) dm(x)$$

Convergence of averages to Nash equilibria

If an update has **sublinear population regret**, then

$\bar{\mu}^{(T)} = \sum_{t \leq T} \gamma_t \mu^{(t)} / \sum_{t \leq T} \gamma_t$ converges

$$\lim_{T \rightarrow \infty} d(\bar{\mu}^{(T)}, \mathcal{N}) = 0$$

Proof: show

$$V(\bar{\mu}^{(T)}) - V(\mu^*) \leq \sum_k \bar{r}^{k(T)}$$

Convergence of a dense subsequence

Proposition

Under any algorithm with sublinear discounted regret, a dense subsequence of $(\mu^{(t)})_t$ converges to \mathcal{N}

- Subsequence $(\mu^{(t)})_{t \in \mathcal{T}}$ converges
- $\lim_{T \rightarrow \infty} \frac{\sum_{t \in \mathcal{T}: t \leq T} \gamma_t}{\sum_{t \leq T} \gamma_t} = 1$

Convergence of a dense subsequence

Proposition

Under any algorithm with sublinear discounted regret, a dense subsequence of $(\mu^{(t)})_t$ converges to \mathcal{N}

- Subsequence $(\mu^{(t)})_{t \in \mathcal{T}}$ converges
- $\lim_{T \rightarrow \infty} \frac{\sum_{t \in \mathcal{T}: t \leq T} \gamma_t}{\sum_{t \leq T} \gamma_t} = 1$

Proof.

- Absolute Cesàro convergence implies convergence of a dense subsequence.

Example: Hedge with learning rates γ_T

$$\pi_p^{(t+1)} \propto \pi_p^{(t)} e^{-\eta_t \ell_p^{k(t)}}$$

Regret bound

Under Hedge with $\eta_t = \gamma_t$,

$$\bar{r}^{(T)}(x) \leq \rho \frac{\ln \pi_{\min}^{(0)}(x) + c \sum_{t \leq T} \gamma_t^2}{\sum_{t \leq T} \gamma_t}$$

Simulations

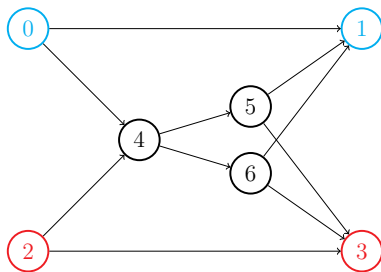


Figure: Example network

Simulations

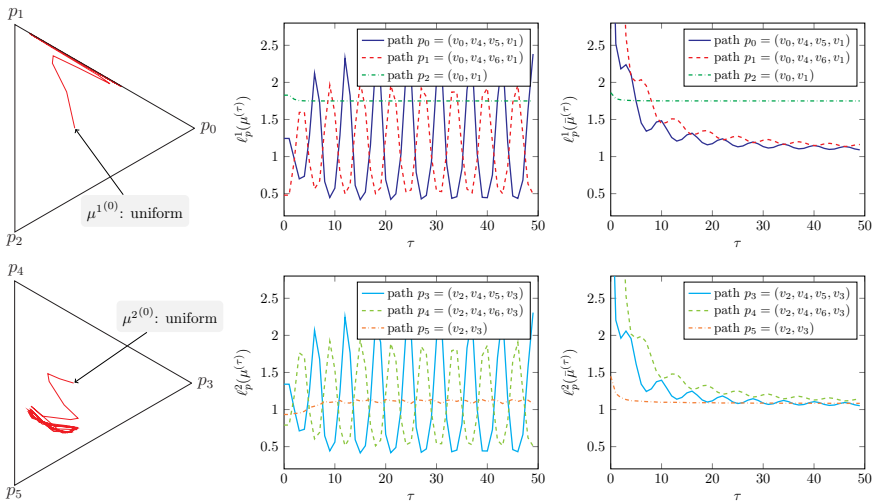


Figure: Path losses and strategies for the Hedge algorithm with constant γ_τ

Simulations

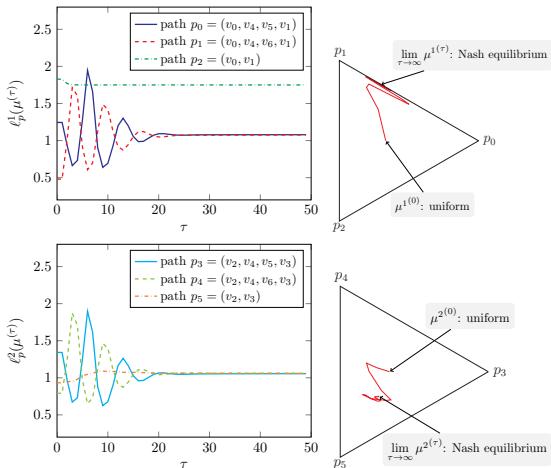


Figure: Path losses and strategies for the Hedge algorithm with $\gamma_\tau = 1/(10 + \tau)$

Sufficient conditions for convergence of $(\mu^{(t)})_t$

- Have $\bar{\mu}^{(t)} \rightarrow \mathcal{N}$.

Sufficient conditions for convergence of $(\mu^{(t)})_t$

- Have $\bar{\mu}^{(t)} \rightarrow \mathcal{N}$.

Sufficient condition

If $V(\mu^{(t)})$ converges ($\mu^{(t)}$ need not converge), then

- $V(\mu^{(t)}) \rightarrow V_*$
- $\mu^{(t)} \rightarrow \mathcal{N}$ (V is continuous, $\mu^{(t)} \in \Delta$ compact)

Replicator dynamics

Imagine an underlying continuous time. Updates happen at $\gamma_1, \gamma_1 + \gamma_2, \dots$

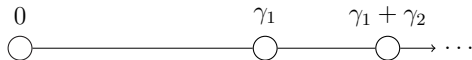


Figure: Underlying continuous time

Replicator dynamics

Imagine an underlying continuous time. Updates happen at $\gamma_1, \gamma_1 + \gamma_2, \dots$

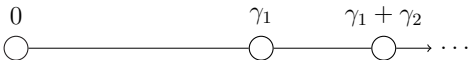


Figure: Underlying continuous time

In the update equation $\mu_p^{(t+1)} \propto \mu_p^{(t)} e^{-\gamma_t \ell_p(t)}$, take $\gamma_t \rightarrow 0$
We obtain the autonomous ODE:

Replicator equation

$$\forall p \in \mathcal{P}_k, \frac{d\mu_p^k}{dt} = \mu_p^k (\langle \ell^k(\mu), \mu^k \rangle - \ell_p^k(\mu)) \quad (1)$$

Also in evolutionary game theory.

Replicator dynamics

Replicator equation

$$\forall p \in \mathcal{P}_k, \frac{d\mu_p^k}{dt} = \mu_p^k (\langle \ell^k(\mu), \mu^k \rangle - \ell_p^k(\mu))$$

Replicator dynamics

Replicator equation

$$\forall p \in \mathcal{P}_k, \frac{d\mu_p^k}{dt} = \mu_p^k (\langle \ell^k(\mu), \mu^k \rangle - \ell_p^k(\mu))$$

Theorem (Fischer and Vöcking (2004))

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

Replicator dynamics

Replicator equation

$$\forall p \in \mathcal{P}_k, \frac{d\mu_p^k}{dt} = \mu_p^k (\langle \ell^k(\mu), \mu^k \rangle - \ell_p^k(\mu))$$

Theorem (Fischer and Vöcking (2004))

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

Proof: V is a Lyapunov function.

AREP update

Discretization of the continuous-time replicator dynamics

$$\pi_p^{(t+1)} - \pi_p^{(t)} = \eta_t \pi_p^{(t)} \left(\langle \ell^k(\mu^{(t)}), \pi^{(t)} \rangle - \ell_p^k(\mu^{(t)}) \right) + \eta_t U_p^{k(t+1)}$$

$(U^{(t)})_{t \geq 1}$ perturbations that satisfy for all $T > 0$,

$$\lim_{\tau_1 \rightarrow \infty} \max_{\tau_2: \sum_{t=\tau_1}^{\tau_2} \eta_t < T} \left\| \sum_{t=\tau_1}^{\tau_2} \eta_t U^{(t+1)} \right\| = 0$$

Benaïm (1999)

Convergence to Nash equilibria

Theorem

Under any no-regret algorithm which is approximate REP, $\mu^{(t)} \rightarrow \mathcal{N}$.

Convergence to Nash equilibria

Theorem

Under any no-regret algorithm which is approximate REP, $\mu^{(t)} \rightarrow \mathcal{N}$.

Proof uses two facts

- Affine interpolation of $\mu^{(t)}$ is an asymptotic pseudo trajectory for the ODE.
- V is a Lyapunov function for Nash equilibria.

REP update

In particular

- REP update: take $U = 0$

$$\pi_p^{(t+1)} - \pi_p^{(t)} = \eta_t \pi_p^{(t)} \left(\left\langle \ell^k(\mu^{(t)}), \pi^{(t)} \right\rangle - \ell_p^k(\mu^{(t)}) \right)$$

REP update

In particular

- REP update: take $U = 0$

$$\pi_p^{(t+1)} - \pi_p^{(t)} = \eta_t \pi_p^{(t)} \left(\left\langle \ell^k(\mu^{(t)}), \pi^{(t)} \right\rangle - \ell_p^k(\mu^{(t)}) \right)$$

- Hedge

$$\pi_p^{(t+1)} - \pi_p^{(t)} = \eta_t \pi_p^{(t)} \frac{e^{-\eta_t \ell_p^k(\mu^{(t)})} - 1}{\eta_t \sum_{p'} e^{-\eta_t \ell_{p'}^k(\mu^{(t)})}}$$

Mirror Descent

Consider the convex problem
 minimize $_{\mu \in \Delta} V(\mu)$

Algorithm 1 Mirror Descent Method

- 1: **for** $t \in \mathbb{N}$ **do**
 - 2: $\mu^{(t+1)} = \arg \min_{\mu \in \Delta} \langle \nabla V(\mu^{(t)}), \mu \rangle + \frac{1}{\eta_t} D_\psi(\mu, \mu^{(t)})$
 - 3: **end for**
-

where D_ψ is a Bregman divergence

$$D_\psi(\mu, \nu) = \psi(\mu) - \psi(\nu) - \langle \nabla \psi(\nu), \mu - \nu \rangle$$

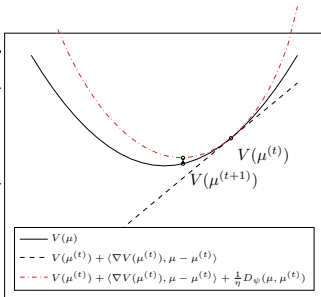


Figure: Mirror Descent iteration.

Mirror Descent

Hedge = Mirror descent on V

- Take $D_\psi(\mu, \nu) = \sum_k D_{KL}(\mu^k, \nu^k)$
- Update:

$$\mu^{(t+1)} = \arg \min_{\mu \in \Delta^1 \times \dots \times \Delta^K} \sum_k \left(\langle \ell^k(\mu^{(t)}), \mu^k \rangle + \frac{1}{\eta_t} D_{KL}(\mu^k, \mu^{k(t)}) \right)$$

Mirror Descent

Hedge = Mirror descent on V

- Take $D_\psi(\mu, \nu) = \sum_k D_{KL}(\mu^k, \nu^k)$
- Update:

$$\mu^{(t+1)} = \arg \min_{\mu \in \Delta^1 \times \dots \times \Delta^K} \sum_k \left(\langle \ell^k(\mu^{(t)}), \mu^k \rangle + \frac{1}{\eta_t} D_{KL}(\mu^k, \mu^{k(t)}) \right)$$

- Solution: Hedge algorithm with learning rate η

$$\mu_p^{k(t+1)} \propto \mu_p^{k(t)} e^{-\eta \ell_p^{k(t)}}$$

Mirror Descent

Hedge = Mirror descent on V

- Take $D_\psi(\mu, \nu) = \sum_k D_{KL}(\mu^k, \nu^k)$
- Update:

$$\mu^{(t+1)} = \arg \min_{\mu \in \Delta^1 \times \dots \times \Delta^K} \sum_k \left(\langle \ell^k(\mu^{(t)}), \mu^k \rangle + \frac{1}{\eta_t} D_{KL}(\mu^k, \mu^{k(t)}) \right)$$

- Solution: Hedge algorithm with learning rate η

$$\mu_p^{k(t+1)} \propto \mu_p^{k(t)} e^{-\eta \ell_p^{k(t)}}$$

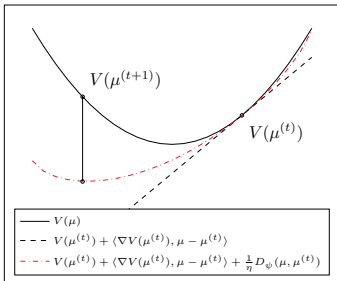
General result

Convergence of $\bar{\mu}^{(T)} = \frac{\sum_{t \leq T} \eta_t \mu^{(t)}}{\sum_{t \leq T} \eta_t}$ to \mathcal{N} for **Any Mirror Descent method**

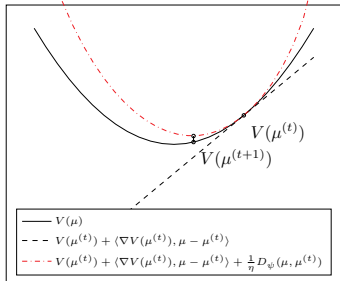
Strong convergence of Mirror Descent

Convex V with L -Lipschitz gradient

If η_t small enough, MD update guarantees $V(\mu^{(t+1)}) \leq V(\mu^{(k)})$.



(a) Large η



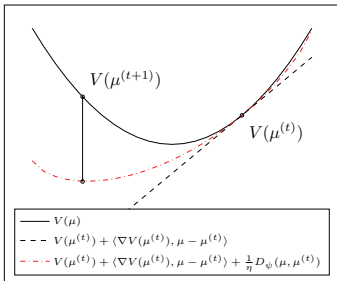
(b) Small η

Figure: Mirror Descent iteration for a function with L -Lipschitz gradient.

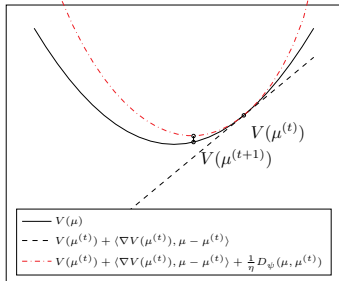
Strong convergence of Mirror Descent

Convex V with L -Lipschitz gradient

If η_t small enough, MD update guarantees $V(\mu^{(t+1)}) \leq V(\mu^{(k)})$.



(a) Large η



(b) Small η

Figure: Mirror Descent iteration for a function with L -Lipschitz gradient.

$V(\mu^{(t)})$ is monotone, converges, so $\mu^{(t)} \rightarrow \mathcal{N}$.

Summary

Summary

- Convergence of $\bar{\mu}^{(t)}$ under **no-regret** updates.
- Convergence of a dense subsequence $(\mu^{(t)})_{t \in \mathcal{T}}$.
- Convergence of $\mu^{(t)}$ for **no-regret AREP** updates.
 - Hedge, REP
- Convergence of $\mu^{(t)}$ for **MD** updates (+ convergence rate)
 - Hedge

Summary

Summary

- Convergence of $\bar{\mu}^{(t)}$ under **no-regret** updates.
- Convergence of a dense subsequence $(\mu^{(t)})_{t \in \mathcal{T}}$.
- Convergence of $\mu^{(t)}$ for **no-regret AREP** updates.
 - Hedge, REP
- Convergence of $\mu^{(t)}$ for **MD** updates (+ convergence rate)
 - Hedge

Future work

- Bandit setting.
- Stochastic perturbations on the losses.

Thank you.

walid@cs.berkeley.edu

benjamin.drighes@polytechnique.edu

bayen@berkeley.edu