On the Convergence of No-regret Learning in Selfish Routing

Walid Krichene Benjamin Drighès Alexandre Bayen

Electrical Engineering and Computer Sciences, UC Berkeley

Applied Mathematics, Ecole Polytechnique, Palaiseau, France

walid@eecs.berkeley.edu, benjamin.drighès@polytechnique.edu, bayen@berkeley.edu

Routing Game

- Graph (V, E)
- source-sink pairs, (s_k, t_k): paths $\mathcal{P}_k$
- Congestion on edge e: $c_e : \phi_e \mapsto c_e(\phi_e)$, increasing
- Player $x \in \mathcal{X}_k$: distribution $\pi(x)$.
- Population $\mathcal{X}_k$: distribution $\mu^k = \frac{1}{m(\mathcal{X}_k)} \int_{\mathcal{X}_k} \pi(x) dm_k(x)$
- Loss on path p: $\ell_p(\mu) = \sum_{e \in p} c_e(\phi_e)$ where $\phi = M\mu$

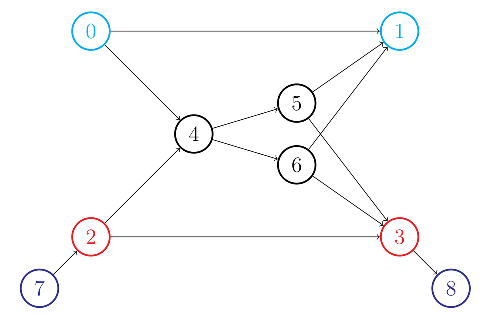


Figure : Example network

Nash equilibrium

- $\mu \in \mathcal{N}$ if for all k, for all $p \in \mathcal{P}_k$ with positive mass, $\ell^k_p(\mu)$ is minimal on $\mathcal{P}_k$
- How to compute Nash equilibria? Solution to a convex optimization problem

$$\text{minimize}_{\mu \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_K}} V(\mu) = \sum_e \int_0^{(M\mu)_e} c_e(u) du$$

- How do players find Nash equilibria? Online learning.

Online Learning

- Every player $x \in \mathcal{X}_k$ maintains distribution $\pi^{(t)}(x) \in \Delta^{\mathcal{P}_k}$ over paths.
- Example: Hedge algorithm: update the distribution according to observed loss

$$\pi_p^{k(t+1)} \propto \pi_p^{k(t)} \exp\left(-\eta_t \ell^k_p(\pi^{(t)})\right)$$

- Discounted Regret

$$R_p^{(T)}(x) = \sum_{t=1}^{T} \eta_t \left\langle \pi(x), \ell^k(\mu^{(t)}) \right\rangle - \sum_{t=1}^{T} \eta_t \ell^k_p(\mu^{(t)})$$

- Regret bound

$$\frac{\max_p R_p^{(T)}(x)}{\sum_{t=1}^{T} \eta_t} \le \frac{-\rho \ln \pi^{(0)}_{\min}(x)}{\sum_{t \le T} \eta_t} + \rho \frac{\sum_{t \le T} \eta_t^2}{\sum_{t \le T} \eta_t}$$

- Potential bound

$$V\left(\frac{\sum_{t \le T} \eta_t \mu^{(t)}}{\sum_{t \le T} \eta_t}\right) - V_* \le \sum_k \frac{\max_p R_p^{k(T)}}{\sum_{t \le T} \eta_t}$$

Average distributions

Convergence of average distributions

If discounted regret is sublinear, $\bar{\mu}^{(t)} = \frac{\sum_{\tau \le t} \eta_\tau \mu^{(\tau)}}{\sum_{\tau \le t} \eta_\tau}$ converges to $\mathcal{N}$.

Convergence of a dense subsequence

Convergence of a dense subsequence

For a potential game with convex potential V, under a MD algorithm with appropriate η_t, a dense subsequence of $(\mu^{(t)})_t$ converges to $\mathcal{N}$.

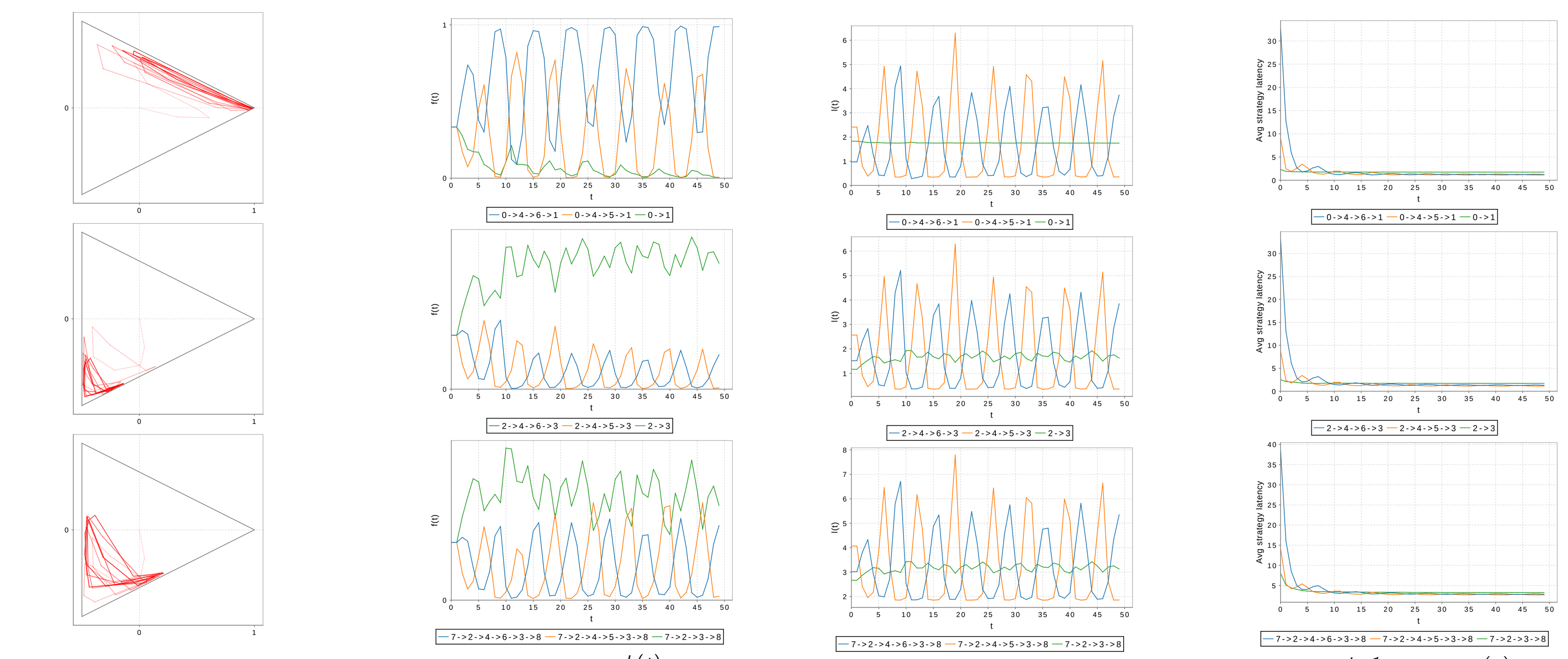
- subsequence $(\mu^{(t)})_{t \in \mathcal{T}}$ converges, and $\lim_{T \to \infty} \frac{\sum_{t \in \mathcal{T}: t \le T} \eta_t}{\sum_{t \le T} \eta_t} = 1$

Does $(\mu^{(t)})_t$ converge?

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

- If $V(\mu^{(t)})$ converges ($\mu^{(t)}$ need not converge), then
  - $V(\mu^{(t)}) \to V_*$
  - $\mu^{(t)} \to \mathcal{N}$ (V is continuous, $\mu \in \Delta$ compact)

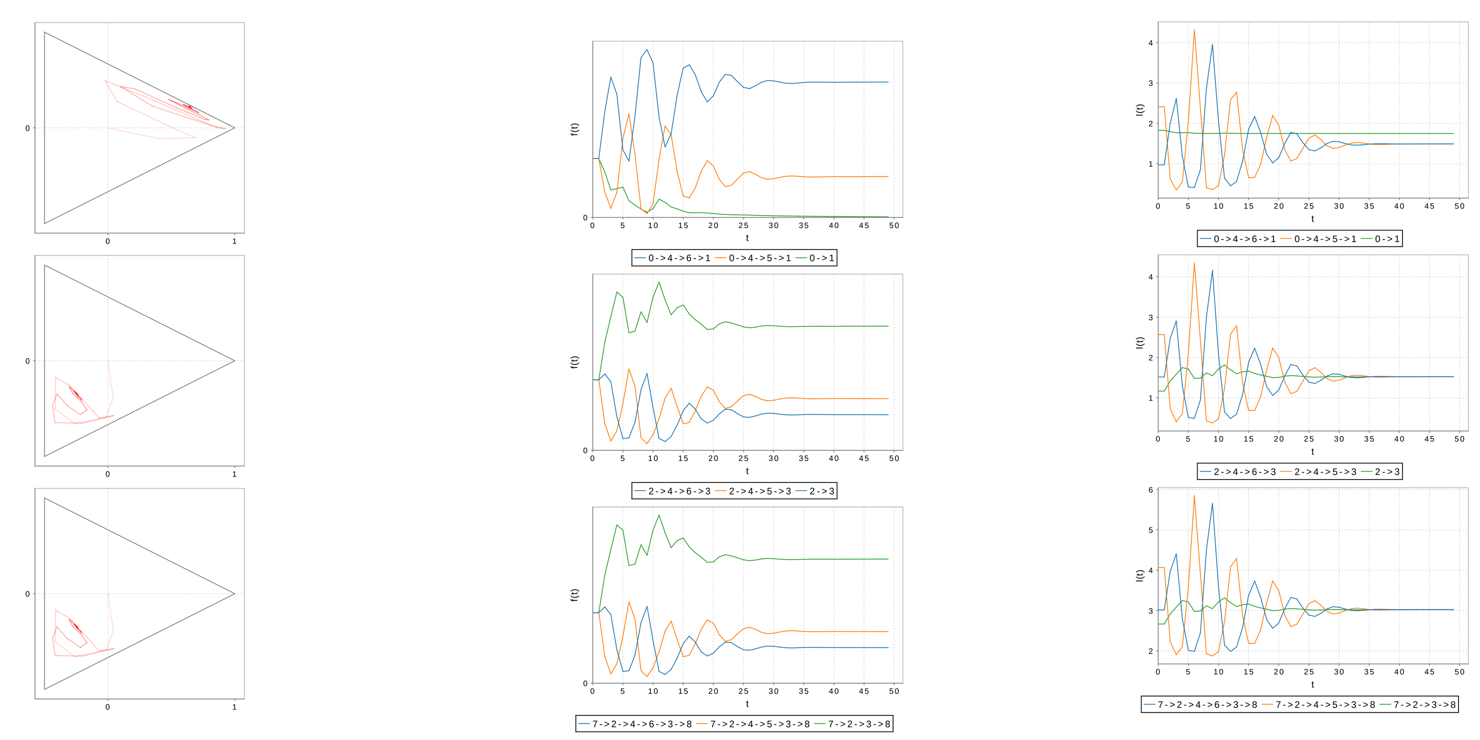
Simulation: convergence of $\bar{\mu}^{(t)}$



(a) Trajectories $(\mu^{k(t)})_t$ (b) Distributions $\mu_p^{k(t)}, p \in \mathcal{P}_k$ (c) Path losses $\ell^k_p(\mu^{(t)})$ (d) $\ell^k_p(\frac{1}{t}\sum_{\tau \le t} \mu^{(\tau)})$

Figure : Constant step size η. $\bar{\mu}^{(t)}$ converges, but $\mu^{(t)}$ does not.

Simulation: convergence of $(\mu^{(t)})$ with decreasing η_t



(a) Trajectories $(\mu^{k(t)})_t$. (b) Path flows $\mu^k_p, p \in \mathcal{P}_k$ (c) Path latencies $\ell^k_p(\mu^{(t)})$

Figure : harmonic step sizes. $\mu^{(t)}$ converges.

Replicator dynamics

Imagine an underlying continuous time. Updates happen at $\eta_1, \eta_1 + \eta_2, \ldots$

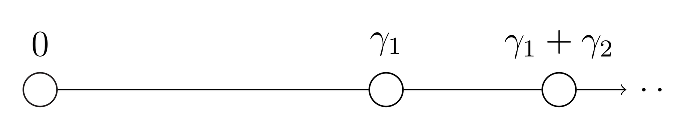


Figure : Underlying continuous time

In the MD solution $\mu_p^{(t+1)} \propto \mu_p^{(t)} \exp\left(-\eta \ell^k_p(\mu^{(t)})\right)$, take $\eta \to 0$

Replicator ODE

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

Convergence of the continuous-time dynamics

- V is a Lyapunov function for the ODE (1)

Theorem

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

Approximate REP algorithm

$$\mu_p^{k(t+1)} - \mu_p^{k(t)} = \eta_t \mu_p^{k(t)} \left(\left\langle \ell^k(\mu^{(t)}), \mu^{k(t)} \right\rangle - \ell^k_p(\mu^{(t)}) \right) / \rho + \eta_t U_p^{k(t+1)}$$

$(U^{(t)})_{t \ge 1}$ perturbations, with $\lim_{\tau_1 \to \infty} \max_{\tau_2 : \sum_{t=\tau_1}^{\tau_2} \eta_t < T} \left\| \sum_{t=\tau_1}^{\tau_2} \eta_t U^{(t+1)} \right\| = 0$

Theorem

Under any AREP algorithm with sublinear regret, $\mu^{(t)} \to \mathcal{N}$.

(1) Show that discrete trajectories approach the solution trajectories of the ODE.

(2) Consequence: $V(\mu^{(t)})$ converges.

Example

Hedge

$$\mu_p^{k(t+1)} \propto \mu_p^{k(t)} \exp\left(-\eta_t \ell^k_p(\mu^{(t)})/\rho\right)$$

REP

$$\mu_p^{k(t+1)} - \mu_p^{k(t)} = \eta_t \mu_p^{k(t)} \left(\left\langle \ell^k(\mu^{(t)}), \mu^{k(t)} \right\rangle - \ell^k_p(\mu^{(t)}) \right) / \rho$$

Both Mirror Descent algorithms $\mu^{(t+1)} \in \arg\min_{\mu \in \Delta} \left\langle \frac{\ell(\mu^{(t)})}{\rho}, \mu \right\rangle + \frac{1}{\eta_t} D(\mu \| \mu^{(t)})$

- Hedge: $D(\mu \| \nu) = \sum_k D_{KL}(\mu^k \| \nu^k)$
- REP: $D(\mu \| \nu) = \frac{1}{2} \sum_k \sum_{p \in \mathcal{P}_k} \nu^k_p \left(\frac{\mu^k_p}{\nu^k_p} - 1 \right)^2$

Mirror Descent

$\text{minimize}_{\mu \in \Delta} V(\mu)$

Algorithm 1 Mirror Descent Method

1: **for** $t \in \mathbb{N}$ **do**

2: $\mu^{(t+1)} = \arg\min_{\mu \in \Delta} \left\langle \nabla V(\mu^{(t)}), \mu \right\rangle + \frac{1}{\eta} D_\psi(\mu, \mu^{(t)})$

3: **end for**

Example:

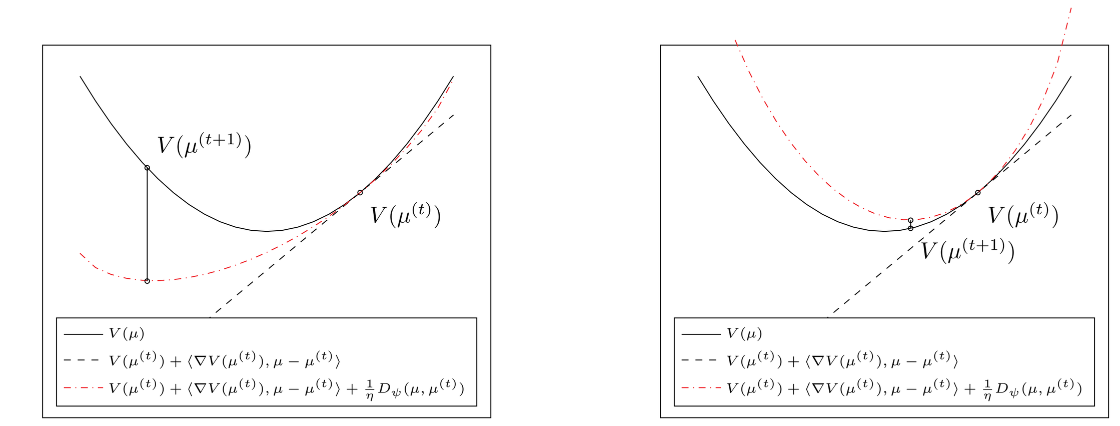
- Take $D_\psi(\mu, \nu) = \sum_k D_{KL}(\mu^k, \nu^k)$
- Objective decomposes, solution is exactly Hedge on each $\Delta^{\mathcal{P}_k}$

Then

$$V\left(\frac{\sum_{t \le T} \eta_t \mu^{(t)}}{\sum_{t \le T} \eta_t}\right) - V(\mu^*) \le \frac{1}{\sum_{t \le T} \eta_t} D_\psi(\mu^{(\mu^*)}, \mu^{(0)}) + \frac{\rho^2}{2} \frac{\sum_{t \le T} \eta_t^2}{\sum_{t \le T} \eta_t}$$

Strong convergence of Mirror Descent

If η_t small enough, MD update guarantees $V(\mu^{(t+1)}) \le V(\mu^{(k)})$. Then $\mu^{(t)} \to \mathcal{N}$.



(a) Large η (b) Small η

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

Summary

- $\bar{\mu}^{(t)} \to \mathcal{N}$ under no-regret updates.
- $\mu^{(t)} \to \mathcal{N}$ under no-regret AREP algorithms (includes Hedge and REP)

Open questions: can one guarantee convergence

- Under different learning rates per population?
- In the presence of perturbation on losses?

This work was supported in part by FORCES (Foundations Of Resilient CybEr-physical Systems), which receives support from the National Science Foundation (NSFaward numbers CNS-1238959, CNS-1238962, CNS-1239054, CNS-1239166).