Convergence of online learning dynamics In routing games

Walid Krichene, Benjamin Drighes, Alex Bayen

Dept. of Electrical Engineering & Computer Sciences,
UC Berkeley, CA, USA

Motivation

The online learning model

- Is a realistic model for population dynamics (weak information assumptions)
- Has convergence guarantees

Can be used

- As a model of population dynamics for optimal control

$$\begin{aligned} \text{minimize}_{u \in \mathcal{U}} \quad & \sum_t J^{(t)}(u^{(t)}, \mu^{(t)}) \\ \text{subject to} \quad & \mu^{(t+1)} = h^{(t)}(u^{(t)}, \mu^{(t)}) \end{aligned}$$

- As an algorithm for distributed load balancing.
- Fast convergence = fast recovery.

Motivation

The online learning model

- Is a realistic model for population dynamics (weak information assumptions)
- Has convergence guarantees

Can be used

- As a model of population dynamics for optimal control

$$\begin{aligned} \text{minimize}_{u\in\mathcal{U}} \quad & \sum_t J^{(t)}(u^{(t)},\mu^{(t)}) \\ \text{subject to} \quad & \mu^{(t+1)} = h^{(t)}(u^{(t)},\mu^{(t)}) \end{aligned}$$

- As an algorithm for distributed load balancing.
- Fast convergence = fast recovery.

Routing game

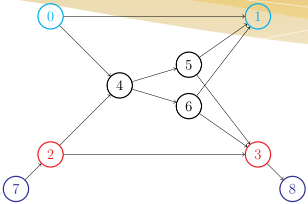


Figure : Example network

- Graph (V, E)
- Source-sink pairs, (s_k, t_k): paths $\mathcal{P}_k$
- Population distribution $\mu^k \in \Delta^{\mathcal{P}_k}$,
- Loss on path p: $\ell_p^k(\mu)$
- Players want to minimize personal loss

Routing game

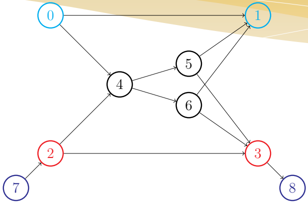


Figure : Example network

- Graph (V, E)
- Source-sink pairs, (s_k, t_k): paths $\mathcal{P}_k$
- Population distribution $\mu^k \in \Delta^{\mathcal{P}_k}$,
- Loss on path p: $\ell_p^k(\mu)$
- Players want to minimize personal loss

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Routing game

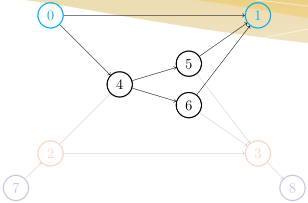


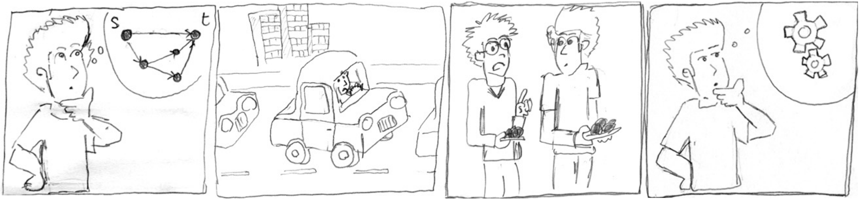
Figure : Example network

- Graph (V, E)
- Source-sink pairs, (s_k, t_k): paths $\mathcal{P}_k$
- Population distribution $\mu^k \in \Delta^{\mathcal{P}_k}$,
- Loss on path p: $\ell_p^k(\mu)$
- Players want to minimize personal loss

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Online learning model

$\pi^{(t)}$ | Sample $p \sim \pi^{(t)}$ | Discover $\ell^{(t)}$ | Update $\pi^{(t+1)}$



Convergence to Nash equilibria

Nash equilibria

$$\mathcal{N} = \arg\min_{\mu \in \Delta} V(\mu)$$

Average strategies

$$\bar{\mu}^{(T)} = \sum_{t \le T} \eta_t \mu^{(t)} / \sum_{t \le T} \eta_t$$

Convergence of averages to Nash equilibria

If an update has sublinear regret, then

$$\bar{\mu}^{(T)} \to \mathcal{N}$$

Proof: show

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

Convergence to Nash equilibria

Nash equilibria

$$\mathcal{N} = \arg\min_{\mu \in \Delta} V(\mu)$$

Average strategies

$$\bar{\mu}^{(T)} = \sum_{t \le T} \eta_t \mu^{(t)} / \sum_{t \le T} \eta_t$$

Convergence of averages to Nash equilibria

If an update has sublinear regret, then

$$\bar{\mu}^{(T)} \to \mathcal{N}$$

Proof: show

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

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Convergence to Nash equilibria

Nash equilibria

$$\mathcal{N} = \arg\min_{\mu \in \Delta} V(\mu)$$

Average strategies

$$\bar{\mu}^{(T)} = \sum_{t \le T} \eta_t \mu^{(t)} / \sum_{t \le T} \eta_t$$

Convergence of averages to Nash equilibria

If an update has sublinear regret, then

$$\bar{\mu}^{(T)} \to \mathcal{N}$$

Proof: show

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

Example dynamics

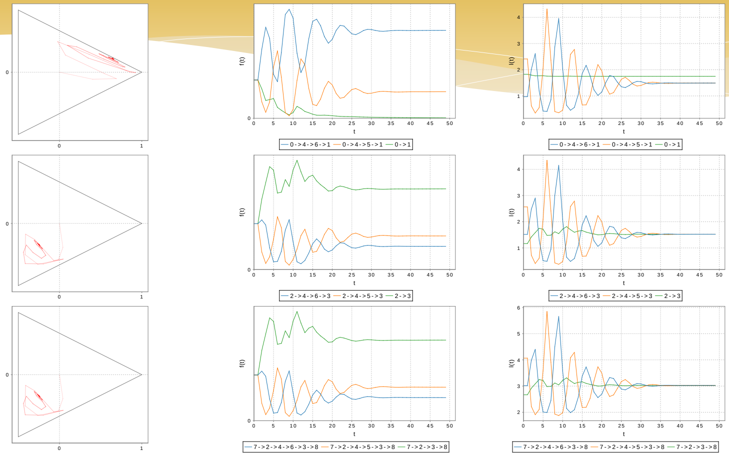
Hedge algorithm

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

REP algorithm

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

Simulations



(a) Trajectories $(\mu^{k(t)})_t$.

(b) Path flows $\left(\mu_p^{k(t)}\right)_{p \in \mathcal{P}_k}$

(c) Path losses $\ell_p^k(\mu^{(t)})$

Figure : Population dynamics under Hedge updates with $\eta_t \downarrow 0$ and $\sum_t \eta_t = \infty$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Convergence of $(\mu^{(t)})_t$

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

Sufficient condition 1

If $V(\mu^{(t)})$ converges, then $\mu^{(t)} \to \mathcal{N}$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Convergence of $(\mu^{(t)})_t$

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

Sufficient condition 1

If $V(\mu^{(t)})$ converges, then $\mu^{(t)} \to \mathcal{N}$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Approximate Replicator algorithms

Underlying continuous time. Updates happen at $\eta_1, \eta_1 + \eta_2, \ldots$

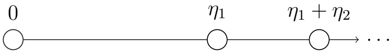


Figure : Underlying continuous time

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

Replicator equation

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

Approximate Replicator algorithms

Underlying continuous time. Updates happen at $\eta_1, \eta_1 + \eta_2, \ldots$

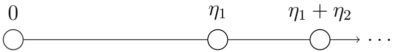


Figure : Underlying continuous time

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

Replicator equation

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

Approximate Replicator algorithms

Discretization of the continuous-time replicator dynamics

Approximate REP algorithm

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

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

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

Theorem

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

Uses sufficient condition 1.

Approximate Replicator algorithms

Discretization of the continuous-time replicator dynamics

Approximate REP algorithm

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

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

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

Theorem

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

Uses sufficient condition 1.

Summary

Convergence of $\mu^{(t)}$ under

- No-regret and AREP algorithms

Current work

- Optimal control under online-learning dynamics
- Robustness of convergence (perturbations in losses)
- Distributed tolling and load balancing

FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

Summary

Convergence of $\mu^{(t)}$ under

- No-regret and AREP algorithms

Current work

- Optimal control under online-learning dynamics
- Robustness of convergence (perturbations in losses)
- Distributed tolling and load balancing

Thank you.