On the convergence of no-regret learning in selfish routing
Supplementary material, ICML 2014

Walid Krichene Benjamin Drighès Alexandre Bayen

1 Proof of Lemma 1

Lemma 1. *Let $(\gamma_\tau)_{\tau\in\mathbb{N}}$ be a non-summable sequence of positive weights. If a real sequence $(u^{(\tau)})_{\tau\in\mathbb{N}}$ converges absolutely to u in the sense of Cesàro means w.r.t. $(\gamma_\tau)_\tau$, that is $\lim_{T\to\infty}\frac{\sum_{\tau\le T}\gamma_\tau|u^{(\tau)}-u|}{\sum_{\tau\le T}\gamma_\tau}=0$, then there exists a subset of indexes $\mathcal{T}$ of density one such that the subsequence $(u^{(\tau)})_{\tau\in\mathcal{T}}$ converges to u.*

Proof. Let $\epsilon>0$, and define the set $\mathcal{I}_\epsilon=\{\tau\in\mathbb{N}\colon \|u^{(\tau)}-u\|\ge\epsilon\}$. First, we show that for all $\epsilon>0$, $\mathcal{I}_\epsilon$ has zero density. Indeed, we have for all $\tau\in\mathcal{I}_\epsilon$, $0\le\epsilon\le\|u^{(\tau)}-u\|$, thus

$$0\le\frac{\sum_{\tau\in\mathcal{I}_\epsilon\colon\tau\le T}\gamma_\tau\epsilon}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}\le\frac{\sum_{\tau\in\mathbb{N}:\tau\le T}\gamma_\tau\|u^{(\tau)}-u\|}{\sum_{\tau\in\mathbb{N}:\tau\le T}\gamma_\tau}$$

But the right-hand side converges to 0 by assumption, therefore $\lim_{T\to\infty}\frac{\sum_{\tau\in\mathcal{I}_\epsilon:\tau\le T}\gamma_\tau}{\sum_{\tau\in\mathbb{N}:\tau\le T}\gamma_\tau}=0$, and $\mathcal{I}_\epsilon$ has zero density w.r.t. (γ_τ).

Next, we will construct a set $\mathcal{I}\subset\mathbb{N}$ of zero density, such that the subsequence $(u_\tau)_{\tau\in\mathbb{N}\setminus\mathcal{I}}$ converges. For all $k>0$, let

$$p_k(T)=\sum_{\tau\in\mathcal{I}_{\frac{1}{k}}\colon\tau\le T}\gamma_\tau$$

Since $\lim_{T\to\infty}\frac{p_k(T)}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}=0$, $\exists T_k>0$ such that for all $T\ge T_k$, $\frac{p_k(T)}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}\le\frac{1}{k}$. Without loss of generality, we can assume that $(T_k)_{k\in\mathbb{N}^*}$ is increasing. Now, let

$$\mathcal{I}=\bigcup_{k\in\mathbb{N}^*}\left(\mathcal{I}_{\frac{1}{k}}\cap\{T_k,\dots,T_{k+1}-1\}\right).$$

Then we have for all $k\in\mathbb{N}^*$, $\mathcal{I}\cap\{0,\dots,T_{k+1}-1\}=\left(\cup_{j=1}^k\mathcal{I}_{\frac{1}{j}}\right)\cap\{0,\dots,T_{k+1}-1\}$. But since $\mathcal{I}_1\subset\mathcal{I}_{\frac{1}{2}}\subset\dots\subset\mathcal{I}_{\frac{1}{k}}$, we have $\mathcal{I}\cap\{0,\dots,T_{k+1}-1\}\subset\mathcal{I}_{\frac{1}{k}}\cap\{0,\dots,T_{k+1}-1\}$, thus for all T such that $T_k\le T<T_{k+1}$, we have

$$\frac{\sum_{\tau\in\mathcal{I}\colon\tau\le T}\gamma_\tau}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}\le\frac{\sum_{\tau\in\mathcal{I}_{\frac{1}{k}}\colon\tau\le T}\gamma_\tau}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}=\frac{p_k(T)}{\sum_{\tau\in\mathbb{N}\colon\tau\le T}\gamma_\tau}\le\frac{1}{k}$$

which proves that $\mathcal{I}$ has zero density.

Let $\mathcal{T}=\mathbb{N}\setminus\mathcal{I}$. We have that $\mathcal{T}$ has density one, and it remains to prove that the subsequence $(u^{(\tau)})_{\tau\in\mathcal{T}}$ converges to u. For all k, there exists $S_k\in\mathcal{T}$ such that $S_k\ge T_k$. For all $\tau\in\mathcal{T}$ with $\tau\ge S_k$, there exists $k'\ge k$ such that $T_{k'}\le\tau<T_{k'+1}$. Since $\tau\notin\mathcal{I}$, we must have $\tau\notin\mathcal{I}_{\frac{1}{k'}}$, therefore

$$\|u^{(\tau)}-u\|<\frac{1}{k'}\le\frac{1}{k}.$$

This proves that $(u^{(\tau)})_{\tau\in\mathcal{T}}$ converges to u. □

2 Proof of Lemma 3

Lemma 3 (Convergence of potentials under AREP algorithms)**.** *Let Γ be a compact invariant set for the replicator ODE $\dot{\mu} = F(\mu)$, V a Lyapunov function for Γ, and assume $V(\Gamma)$ has empty interior. Assume that the sequence of distributions $(\mu^{(\tau)})_{\tau\in\mathbb{N}}$ obey an AREP update rule. Then the sequence of potentials $(V(\mu^{(\tau)}))_\tau$ converges.*

Proof. Let M be the affine interpolated process of the sequence $(\mu^{(\tau)})_\tau$ with time steps $(\gamma_\tau)_\tau$. That is, M is the continuous function defined on $\mathbb{R}_+$ and with values in Δ, such that for all τ and all $s \in [0, \gamma_\tau)$

$$M(T_\tau + s) = \mu^{(\tau)} + s\frac{\mu^{(\tau+1)} - \mu^{(\tau)}}{\gamma_\tau}$$

where $T_\tau = \sum_{t\leq\tau} \gamma_t$.

First, by definition of the AREP class, we have $(\mu^{(\tau)})$ satisfies an update equation of the form

$$\mu^{(\tau+1)} - \mu^{(\tau)} = \gamma_\tau \left(F(\mu^{(\tau)}) + U^{(\tau+1)}\right)$$

where $(U^{(\tau)})$ is a bounded sequence of perturbations that satisfies the condition

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

Furthermore, the sequence $(\mu^{(\tau)})$ is bounded since Δ is compact. by Proposition 4.1 in [1], the affine interpolated process M is an asymptotic pseudo trajectory of the vector field F. It follows by Theorem 5.7 in [1] that the set $L(M)$ of limit points of M is internally chain transitive.

Since V is a Lyapunov function for the compact invariant set Γ, and $V(\Gamma)$ has empty interior, by Proposition 6.4 in [1], Γ contains every internally chain transitive set and V is constant on Γ. In particular, we have $L(M) \subset \Gamma$ and there exists $v \in \mathbb{R}$ such that $V(\mu) = v$ for all $\mu \in L(M)$.

Now consider the sequence of potential values $(V(\mu^{(\tau)}))_{\tau\in\mathbb{N}}$. Since V is continuous and Δ is compact, $V(\Delta)$ is compact, and $(V(\mu^{(\tau)}))_{\tau\in\mathbb{N}}$ has at least one limit point. Let $\hat{v}$ be such a limit point, that is, $\hat{v}$ is the limit of a subsequence $(V(\mu^{(\tau)}))_{\tau\in\mathcal{T}}$ where $\mathcal{T} \subseteq \mathbb{N}$. The subsequence $(\mu^{(\tau)})_{\tau\in\mathcal{T}}$ lives in the compact set Δ, thus we can extract a further subsequence which converges, that is, there exists $\mathcal{T}' \subset \mathcal{T}$ and $\hat{\mu} \in \Delta$ such that $(\mu^{(\tau)})_{\tau\in\mathcal{T}'}$ converges to $\hat{\mu}$. By continuity of V, $V(\hat{\mu}) = \hat{v}$. But since $\hat{\mu} \in L(M)$, we also have $V(\hat{\mu}) = v$, therefore $\hat{v} = v$. This proves that $V(\mu^{(\tau)})_\tau$ has a unique limit point equal to v, thus it converges to v. □

References

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