HETEROGENEITY, REINFORCEMENT LEARNING AND CHAOS IN POPULATION GAMES SUPPLEMENTARY INFORMATION

JAKUB BIELAWSKI, THIPARAT CHOTIBUT, FRYDERYK FALNIOWSKI, MICHAŁ MISIUREWICZ, AND GEORGIOS PILIOURAS

1. Multiplicative Weights Update

Multiplicative Weights Update (MWU) is an algorithm discovered and used in economics, computer science and AI, with close connection to models of theoretical biology [2, 3]. Multiplicity of fields of applications is a consequence both of its simplicity and of its intuitive property — it is a no regret algorithm. No regret algorithms play fundamental role in applications, as they can be used when agents lack information, because they guarantee that the sequence of chosen actions is asymptotically as good as the best fixed action in hindsight.

Let Γ be a set of strategies. At time $t = 1, 2 \ldots, T$ a player chooses a mixed strategy x^t, that is a probability distribution over the action set Γ, and then receives a cost c^t. The (time-averaged) regret of the action sequence $\gamma^1, \gamma^2, \ldots, \gamma^T \in \Gamma$ with respect to the fixed strategy $\gamma \in \Gamma$ is

$$\frac{1}{T}\left[\sum_{t=1}^{T} c^t(\gamma^t) - \sum_{t=1}^{T} c^t(\gamma)\right]. \tag{1}$$

An algorithm is no regret when for all $\gamma \in \Gamma$ the expression in (1) with $T \to \infty$ converges to zero. MWU is a canonical example of no regret algorithm. In addition, it gives optimal bounds for regret [6].

Multiplicative Weights Update maintains weights of the strategies. At each step the algorithm chooses a strategy with probability proportional to its current weight. Initially we set equal weights $w^1(\gamma) = 1$ for every $\gamma \in \Gamma$. Then at step $t = 1, 2, \ldots, T$ the action $\hat{\gamma}$ is chosen by a player with probability

$$x^t(\hat{\gamma}) = \frac{w^t(\hat{\gamma})}{\sum_{\gamma \in \Gamma} w^t(\gamma)}. \tag{2}$$

The weight of a strategy $\gamma \in \Gamma$ is updated as follows

$$w^{t+1}(\gamma) = w^1(\gamma)(1-\varepsilon)^{\mathcal{C}^t(\gamma)},$$

where $\mathcal{C}^t(\gamma) = \sum_{\tau=1}^{t} c^\tau(\gamma)$ is the cumulative cost of the play of strategy γ up to step t and $\varepsilon \in (0, 1)$ is a common learning rate of the agents. We express the update rule of the weight of the action γ in terms of the previous-step weight

$$\begin{aligned} w^{t+1}(\gamma) &= w^1(\gamma)(1-\varepsilon)^{\mathcal{C}^{t-1}(\gamma)}(1-\varepsilon)^{c^t(\gamma)} \\ &= w^t(\gamma)(1-\varepsilon)^{c^t(\gamma)}. \end{aligned}$$

The weight $w^t(\gamma)$ decreases with time and the rate of its decrease depends of the cumulative cost of the previous play of strategy γ.

By using (2) and (1) we can express the probabilities x^{t+1} in terms of of the previous-step probabilities

$$\begin{aligned} x^{t+1}(\hat{\gamma}) &= \frac{w^t(\hat{\gamma})(1-\varepsilon)^{c^t(\hat{\gamma})}}{\sum_{\gamma\in\Gamma} w^t(\gamma)(1-\varepsilon)^{c^t(\gamma)}} \\ &= \frac{\frac{w^t(\hat{\gamma})}{\sum_{\gamma\in\Gamma} w^t(\gamma)}(1-\varepsilon)^{c^t(\hat{\gamma})}}{\sum_{\gamma\in\Gamma} \frac{w^t(\gamma)}{\sum_{\gamma\in\Gamma} w^t(\gamma)}(1-\varepsilon)^{c^t(\gamma)}} \\ &= \frac{x^t(\hat{\gamma})(1-\varepsilon)^{c^t(\hat{\gamma})}}{\sum_{\gamma\in\Gamma} x^t(\gamma)(1-\varepsilon)^{c^t(\gamma)}}. \end{aligned}$$

Thus,

$$x^{t+1}(\hat{\gamma}) = \frac{x^t(\hat{\gamma})(1-\varepsilon)^{c^t(\hat{\gamma})}}{\sum_{\gamma\in\Gamma} x^t(\gamma)(1-\varepsilon)^{c^t(\gamma)}}$$

2. Background in Dynamical Systems

2.1. **Invariant measures and the ergodic theorem.** We can also discuss a discrete dynamical system in terms of a measure preserving transformation defined on a probability space. This approach can handle not only purely mathematical concepts but also physical phenomena in nature. This subsection is devoted to invariant measures, absolutely continuous measures and the most fundamental idea in ergodic theory — the Birkhoff Ergodic Theorem, which states that with probability one the average of a function along an orbit of an ergodic transformation is equal to the integral of the given function.

Let $(X, \mathcal{B}, \mu)$ be a probability space and $f\colon X \mapsto X$ be a measurable map. The measure μ is f-invariant (a map f is μ-invariant) if $\mu(f^{-1}E) = \mu(E)$ for every $E \in \mathcal{B}$. For an f-invariant measure μ we say that μ is ergodic (f is ergodic) if for all $E \in \mathcal{B}$ if $f^{-1}E = E$ then $\mu(E) = 0$ or 1. If X is an interval, a measure μ is absolutely continuous with respect to Lebesgue measure if and only if for every set $E \in \mathcal{B}$ of zero Lebesgue measure $\mu(E) = 0$.

We can now state the ergodic theorem.

Theorem (Birkhoff Ergodic Theorem)**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space. If f is μ-invariant and g is integrable, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(f^k(x)) = g^*(x)$$

for some $g^* \in L^1(X, \mu)$ with $g^*(f(x)) = g^*(x)$ for almost every x. Furthermore if f is ergodic, then g^* is constant almost everywhere and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(f^k(x)) = \int_X g \, d\mu$$

for almost every x.

Lastly, why do absolutely continuous invariant measures matter? Computer-based investigations are widely used to gain insights into the dynamics of chaotic phenomena. However, one must exercise caution in the interpretation of computer simulations. Usually, chaotic systems possess multiple ergodic probability invariant measures, but if the absolutely continuous measure with respect to Lebesgue measure exists, the averages of a given observable (function) along the orbits obtained from the computer simulations will be equal to the integral of this observable with respect to our measure [1]. Thus, the theoretical measure and the computational measure coincide in this situation.

3. Proofs

3.1. **Proof of Lemma 3.1.**

Proof of Lemma 3.1. We have

$$F(\xi_s)(\omega) = \frac{\xi_s(\omega)}{\xi_s(\omega) + (1 - \xi_s(\omega)) \exp\left(a(\omega)\left(\int \xi_s \, d\mu - b\right)\right)}.$$

By applying the formula for ξ_s we get

$$\begin{aligned}
&F(\xi_s)(\omega) = \\
&= \frac{\xi_s(\omega)}{\xi_s(\omega) + (1 - \xi_s(\omega)) \exp\left(a(\omega)\left(\int \xi_s \, d\mu - b\right)\right)} \\
&= \frac{\xi(\omega)}{\xi(\omega) + (1 - \xi(\omega)) \exp(sa(\omega)) \exp\left(a(\omega)\left(\int \xi_s \, d\mu - b\right)\right)} \\
&= \frac{\xi(\omega)}{\xi(\omega) + (1 - \xi(\omega)) \exp\left(\left(s + \int \xi_s \, d\mu - b\right) a(\omega)\right)} \\
&= \xi_{G(s)}(\omega),
\end{aligned}$$

where

$$G(s) = s + \int \xi_s \, d\mu - b.$$

□

This means that ξ can be embedded in a one-parameter family, invariant for F, on which F is topologically conjugate to G.

3.2. **Proof of Theorem 3.2.**

Proof of Theorem 3.2. Observe that if $s \geq \varepsilon$, then $\xi_{a_n,s} \leq \xi_{a_n,\varepsilon}$. The functions $\xi_{a_n,\varepsilon}$ are commonly bounded by 1 on a space of finite measure and converge pointwise to 0 as $n \to \infty$. Therefore, if we fix a and let s go to ∞ (respectively $-\infty$), then $\int \xi_{a,s} \, d\mu$ goes to 0 (respectively 1). Therefore, if s is sufficiently big, then $G_a(s) < s$, and if s is sufficiently small then $G_a(s) > s$. Hence, G_a has a fixed point. Since $\xi_{a,s}$ is a strictly decreasing function of s, so is $G_a(s) - s$, and therefore the fixed point is unique. □

3.3. **Proof of Theorem 3.5.** Let us now check what happens if the function a is small.

Lemma 1. *If* $\sup_{\omega\in\Omega} a(\omega) \le A$ *then* $1 - \frac{A}{4} \le G_a'(s) < 1$ *for all* $s \in \mathbb{R}$.

Proof. We have

$$G_a'(s) = 1 + \frac{d}{ds}\int \xi_{a,s}(\omega)\, d\mu(\omega).$$

By differentiating the integrand, we get

$$\frac{d}{ds}\xi_{a,s}(\omega) = -a(\omega)\frac{\xi(\omega)(1-\xi(\omega))\exp(sa(\omega))}{\big(\xi(\omega)+(1-\xi(\omega))\exp(sa(\omega))\big)^2}.$$

Since $\xi(\omega) \in (0,1)$ for all $\omega \in \Omega$, this derivative is negative. On the other hand, if $t, u \in \mathbb{R}$, then $(t+u)^2 - 4tu = (t-u)^2 \ge 0$, so $\frac{tu}{(t+u)^2} \le \frac{1}{4}$. Applying this to $t = \xi(\omega)$ and $u = (1-\xi(\omega))\exp(sa(\omega))$, we get

$$-a(\omega)\frac{d}{ds}\xi_{a,s}(\omega) \ge -\frac{a(\omega)}{4} \ge -\frac{A}{4}.$$

Therefore,

$$\left|\frac{d}{ds}\xi_{a,s}(\omega)\right| \le \frac{A}{4},$$

and since constant functions are integrable for μ, we can apply the Leibniz integral rule for differentiation under the integral sign. In such way we get $1 - \frac{A}{4} \le G_a'(s) < 1$. □

Proof of Theorem 3.5. If $\sup_{\omega\in\Omega} a(\omega) < 8$, then by Lemma 1 we have $-1 < G_a'(s) < 1$ for all s. Therefore, this fixed point is globally attracting. □

3.4. **Proof of Theorem 3.7.** Define functions G_- and G_+ by $G_-(s) = s + 1 - b$ and $G_+(s) = s - b$.

Lemma 2. *If the sequence* $(a_n)_{n=1}^\infty$ *of measurable functions from* Ω *to* $(0,\infty)$ *converges pointwise to infinity, then for every* $\varepsilon > 0$ *functions* G_{a_n} *converge uniformly to* G_- *on* $(-\infty, -\varepsilon]$ *and to* G_+ *on* $[\varepsilon, \infty)$. *Moreover, for every* n *we have* $G_+ < G_{a_n} < G_-$.

Proof. Observe that if $s \ge \varepsilon$, then $\xi_{a_n,s} \le \xi_{a_n,\varepsilon}$. The functions $\xi_{a_n,\varepsilon}$ are commonly bounded by 1 on a space of finite measure and converge pointwise to 0 as $n \to \infty$. Therefore, their integrals converge to 0, so the functions G_{a_n} converge to G_+ uniformly with respect to $s \in [\varepsilon, \infty)$. Similarly, they converge to G_- uniformly with respect to $s \in (-\infty, -\varepsilon]$. The last inequality is obvious. □

Proposition 3. *If* $b \ne 1/2$ *and the sequence* $(a_n)_{n=1}^\infty$ *of measurable functions from* Ω *to* $(0,\infty)$ *converges pointwise to infinity, then there exists* N *such that for every* $n \ge N$ *the map* G_{a_n} *has a periodic orbit of period* 3.

Proof. Assume first that $b > 1/2$. Fix a positive number $\varepsilon < \min\left(\frac{1-b}{3}, \frac{2b-1}{3}\right)$. By Lemma 2 there exists N such that for every $n \ge N$, if $s \le -\varepsilon$ then

$$s + 1 - b - \varepsilon < G_{a_n}(s) < s + 1 - b,$$

and if $s \ge \varepsilon$ then

$$s - b < G_{a_n}(s) < s - b + \varepsilon.$$

We have $G_{a_n}(-\varepsilon) > 1 - b - 2\varepsilon > \varepsilon$ and $G_{a_n}(\varepsilon) < -b + 2\varepsilon < -\varepsilon$. Therefore, there exists $s_0 \in (-\varepsilon, \varepsilon)$ such that $G_{a_n}(s_0) = \varepsilon$. We have $G^2_{a_n}(\varepsilon) < G_{a_n}(-b + 2\varepsilon) < 1 - 2b + 2\varepsilon < -\varepsilon$. Thus, $s_0 < \varepsilon = G_{a_n}(s_0)$ and $G^3_{a_n}(s_0) < -\varepsilon < s_0$. By [4], this implies the existence of a periodic point of period 3 for G_{a_n}.

The case $b < 1/2$ can be reduced to the case $b > 1/2$ by replacing b by $1 - b$ and conjugating G_{a_n}, G_-, G_+ via $s \mapsto 1 - s$ (this switches G_- and G_+). □

Proof of Theorem 3.7. By the Sharkovsky Theorem [7], existence of a periodic orbit of period 3 implies existence of periodic orbits of all periods, and by the result of [5], it implies that the map is Li-Yorke chaotic. Thus, use of Proposition 3 completes the proof of Theorem 3.7. □

3.5. **Proof of Theorem 3.9.** Now we look at the averages. For a given $\zeta \in M(\Omega, I)$, we can consider its space average, $\int \zeta \, d\mu$. It can be interpreted as the expected value of ζ. We will prove that the sequence of averages of the expected values of $F^n(\zeta)$ (which is the same as the sequence of expected values of averages of $F^n(\zeta)$) converges to b.

Proof of Theorem 3.9. We started our construction by fixing a function ξ. It was arbitrary, so we can take $\xi = \zeta$. If $s_n = G^n_a(0)$, then

$$s_{n+1} = s_n + \int \xi_{a,s_n} \, d\mu - b = s_n + \int \zeta_n \, d\mu - b.$$

In such a way we get

$$\sum_{k=0}^{n-1} \int \zeta_k \, d\mu - nb = s_n.$$

From Lemma 2 it follows that $G_-(s) - G(s) \to 0$ when $s \to -\infty$ and $G(s) - G_+(s) \to 0$ as $s \to \infty$. Thus, for s close to $-\infty$ we have $G(s) > s$ and for s close to ∞ we have $G(s) < s$. Therefore, the trajectory of 0 under the iterates of G_a is bounded, so the left-hand side above is bounded uniformly in n. □

References

[1] Abraham Boyarsky and Pawel Gora. *Laws of chaos: invariant measures and dynamical systems in one dimension*. Springer Science & Business Media, 2012.
[2] Erick Chastain, Adi Livnat, Christos Papadimitriou, and Umesh Vazirani. Algorithms, games, and evolution. *Proceedings of the National Academy of Sciences*, 111(29):10620–10623, 2014.
[3] Tobias Galla and J Doyne Farmer. Complex dynamics in learning complicated games. *Proceedings of the National Academy of Sciences*, 110(4):1232–1236, 2013.
[4] T.-Y. Li, M. Misiurewicz, G. Pianigiani, and J.A. Yorke. Odd chaos. *Physics Letters A*, 87, 1982.
[5] Tien-Yien Li and James A. Yorke. Period three implies chaos. *The American Mathematical Monthly*, 82(10):985–992, 1975.
[6] Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.
[7] Aleksandr N. Sharkovsky. Coexistence of the cycles of a continuous mapping of the line into itself. *Ukrain. Math. Zh.*, 16:61–71, 1964.

(J. Bielawski) Department of Mathematics, Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland
E-mail address: jakub.bielawski@uek.krakow.pl

(T. Chotibut) Chula Intelligent and Complex Systems, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.
E-mail address: Thiparat.C@chula.ac.th, thiparatc@gmail.com

(F. Falniowski) Department of Mathematics, Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland
E-mail address: falniowf@uek.krakow.pl

(M. Misiurewicz) Department of Mathematical Sciences, Indiana University-Purdue University Indianapolis, 402 N. Blackford Street, Indianapolis, IN 46202, USA
E-mail address: mmisiure@math.iupui.edu

(G. Piliouras) DeepMind, 14-18 Handyside Street, London N1C 4DN, UK
E-mail address: gpil@deepmind.com