Apprentissage dans les Jeux Anonymous Non-atomique: Application aux Jeux Champs Moyens

Saeed HADIKHANLOO, CEREMADE

Nous présentons un modéle des Jeux Anonymes où les actions sont choisies parmi les ensembles dépendant des joueurs. Nous proposons des procédures d'apprentissage similaires aux *Ficititious Play* et *Online Mirror Descent*. Nous montrons la convergence de ces procédures vers l'Equilibre Nash dans le cas monotone. Les cas d'application principaux de notre travail sont les jeux champs moyens.
Soit I l'ensemble des joueurs et $\lambda \in \mathcal{P}(I)$ une distribution non-atomique qui représente la répartition des joueurs sur I. Pour tout joueur $i \in I$, soit $A_i \subseteq V$ l'ensemble des actions de i. Les profiles des actions admissibles sont

$$\mathcal{A} = \{\Psi : I \to V \text{ mesurable} \mid \Psi(i) \in A_i \text{ pour } \lambda\text{-presque tout } i \in I\}.$$

Le jeu est appellé *anonyme* si le coût des joueurs est sous forme de $J(a, \eta)$, où $a \in V$ est l'action choisi par le joueur et η la distribution des actions choisies par les adversaires, noté par $\Psi\sharp\lambda$, si le profil des action est Ψ. Le coût J est appellé monotone si pour tout $\eta, \eta' \in \mathcal{P}(V)$ on a $\int_V (J(a,\eta) - J(a,\eta'))\ (\eta - \eta')(a) \geq 0$. Nous verrons que la mononicité avec la condition de minimiseur unique nous donnera l'unicité d'équilibre. Ensuite, nous présenterons un procédure d'apprentissage similaire au *Fictitious Play* comme suivant:

$$\begin{array}{llcl} (i) & \Psi_{n+1}(i) & = & BR(i, \bar{\eta}_n), \quad \text{pour } \lambda\text{-presque tout } i \in I, \\ (ii) & \eta_{n+1} & = & \Psi_{n+1}\sharp\lambda, \\ (iii) & \bar{\eta}_{n+1} & = & \frac{1}{n+1}\sum_{i=1}^{n+1} \eta_i. \end{array} \tag{1}$$

Le deuxième procédure que nous étudierons sera le *Online Mirror Descente*:

$$\begin{array}{llcll} (i) & \Phi_{n+1}(i) & = & \Phi_n(i) - \frac{1}{n}\nabla J(\Psi_n(i), \eta_n), & \text{pour tous } i \in I \\ (ii) & \Psi_{n+1}(i) & = & \arg\max_{a \in A_i} \langle \Phi_{n+1}(i), a\rangle - h(a), & \text{pour tous } i \in I \\ (iii) & \eta_{n+1} & = & \Phi_{n+1}\sharp\lambda. & \end{array} \tag{2}$$

Dans tous les deux procédures, en considérant les conditions de monotonicité, régularité et convexité pour le cas OMD, nous montrons la convergence de η_n vers l'unique équilibre. Nous démontrons que le Jeu Champ Moyens en Premier l'Ordre est bien un exemple des jeux anonymes et les résultats des convergences des précédents procédures peuvent directement être appliqués.

Références

[1] Brown, G. W., *Iterative solution of games by Fictitious Play.* Activity analysis of production and allocation 13.1 (1951), 374-376.

[2] Cardaliaguet P., Hadikhanloo S., *Learning in Mean Field Games: the Fictitious Play.* ESAIM: Control, Optimisation and Calculus of Variations. 2015 Jul 30.

[3] Lasry, J.-M., and Lions, P.-L., *Jeux à champ moyen. II. Horizon fini et contrôle optimal.* C. R. Math. Acad. Sci. Paris 343 (2006), no. 10, 679-684.

[4] Mas-Colell, Andreu. *On a theorem of Schmeidler.* Journal of Mathematical Economics 13.3 (1984): 201-206.

[5] Shalev-Shwartz, Shai. *Online learning and online convex optimization.* Foundations and Trends in Machine Learning 4.2 (2011): 107-194.

[6] Schmeidler, David. *Equilibrium points of nonatomic games.* Journal of statistical Physics 7.4 (1973): 295-300.

Saeed HADIKHANLOO, Université Paris-Dauphine, PSL Research University, CNRS, UMR [7534], CEREMADE, 75016 PARIS, FRANCE
saeed.hadikhanloo@dauphine.fr