Gilles Stoltz

Chargé de recherche CNRS
Professeur affilié à HEC Paris

Journée de formation IPR 2013 / Théorie des jeux

Jeux répétés

Convergence vers des ensembles d'équilibres

Contenu

- Équilibres corrélés (jeux en un coup)
- Notion de stratégie dans un jeu répété
- Un outil fondamental : le théorème d'approchabilité de Blackwell
- Minimisation du regret et conséquences : convergences vers des ensembles d'équilibres

Coordonnées : stoltz@math.ens.fr --- http://www.math.ens.fr/~stoltz/

Notion d'équilibre corrélé (illustration sur deux jeux : guerre des sexes / poule mouillée)

Déf mathématique : $\pi \in \Delta(A \times B)$ est un équilibre corrélé si

$$\forall \varphi : A \to A \qquad E_\pi[r_1(\varphi(I), J)] \leq E_\pi[r_1(I,J)]$$
$$\forall \psi : B \to B \qquad E_\pi[r_2(I, \psi(J))] \leq E_\pi[r_2(I,J)]$$

Interprétation : Un « médiateur » tire un couple $(I, J) \in A \times B$ selon la loi jointe π et informe $\left\{ \begin{array}{l} J1 \text{ de } I \\ J2 \text{ de } J \end{array} \right\}$ en privé.

Aucun joueur n'a intérêt à dévier unilatéralement de l'information / de la recommandation qui lui a été faite.

Rq : Si $\pi = x \otimes y$ où (x,y) est un équilibre de Nash alors π est également un équilibre corrélé

(Nash $\longleftrightarrow$ absence de médiateur / ou : médiateur tire I et J indépendamment selon x et y)

Rq : Eq. corrélé $\leftrightarrow$ (id, id) est un éq. de Nash pour un certain jeu.

Csq : Les paiements obtenus par un équilibre corrélé peuvent être plus satisfaisants que ceux d'un équilibre de Nash, grâce à la coordination rendue possible par l'existence d'un médiateur.

Ex 1 [guerre des sexes] :

$\left\{ \begin{array}{l} 1 \text{ point si son activité préférée} \\ 2 \text{ points si conjoint présent} \end{array} \right.$

	Opera	Foot	J_2 ♂
Opera	(3,2)	(1,1)	
Foot	(0,0)	(2,3)	
J_1 ♀			

Equilibres (purs) de Nash sont inégalitaires : (Opera, Opera) et (Foot, Foot)

Equilibre mixte : cf page suivante

Nash mixte : $x = (3/4, 1/4)$ et $y = (1/4, 3/4)$ [on peut vérifier que c'est l'unique]

paiements : $3 \times \frac{3}{4} \times \frac{1}{4} + 1 \times \frac{3}{4} \times \frac{3}{4} + 2 \times \frac{1}{4} \times \frac{3}{4}$

$= \frac{24}{16} = 3/2$

↳ $(3/2, 3/2)$: c'est peu !

Equilibres corrélés : Il y en a plein ! en sont ; $\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ et $\begin{pmatrix} 1/3 & 1/3 \\ 0 & 1/3 \end{pmatrix}$

paiements respectifs : $(5/2, 5/2)$ et $(2,2)$

Note : parmi les équilibres corrélés de la forme $\begin{pmatrix} \pi_{O/O} & \pi_{O/F} \\ 0 & \pi_{F/F} \end{pmatrix}$ c'est $\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ celui qui mène aux meilleurs paiements.

Note : il y a d'autres équilibres corrélés que ceux-ci ; ex : le Nash mixte !

Ccl : $(5/2, 5/2)$ rend les choses plus égalitaires

ici, ⟷ Signal public : on lance une pièce et on suit sa décision.

Autre ex. de signal public : feux rouges (pour jeu de la poule mouillée)

Jeux répétés, notion de stratégies (randomisées)

- On part d'un jeu en un coup, de fonctions de paiement
$$r_1 : A \times B \to \mathbb{R} \quad \text{et} \quad r_2 : A \times B \to \mathbb{R}$$

- Choix simultanés de $I_t \in A$ et $J_t \in B$, éventuellement au hasard selon $x_t \in \Delta(A)$ et $y_t \in \Delta(B)$ et en se fondant sur les actions passées : (I_s, J_s) $s \leq t-1$.

  Une stratégie du joueur 1 est donc une fonction
$$\varphi : \bigcup_{s \in \mathbb{N}} (A \times B)^s \longrightarrow \Delta(A)$$
  qui à tout passé associe une action mixte, selon laquelle est tirée une action pure. → La stratégie, c'est l'essence de la théorie des jeux !

- On s'intéressera aux paiements moyens lorsque $T \to +\infty$:
$$\frac{1}{T} \sum_{t=1}^{T} r_1(I_t, J_t) \quad \text{et} \quad \frac{1}{T} \sum_{t=1}^{T} r_2(I_t, J_t)$$

OBJECTIF DE CET EXPOSÉ :

Montrer que des interactions répétées entre joueurs jouant de « bonnes » stratégies ($\leftrightarrow$ processus évolutionnaires) permettent de converger vers des équilibres (du jeu en 1 coup)

↳ Ça justifie qu'on observe des équilibres dans le monde qui nous entoure... (en économie p.ex.)

Théorie du contrôle / théorie des jeux :

Théorème d'approchabilité de Blackwell (1956)

Fonction vectorielle $m : A \times B \to \mathbb{R}^d$ (on verra plus tard le lien avec r_1 et r_2)

étendue linéairement à $\Delta(A) \times \Delta(B)$:

$$m(x,y) = \sum_{a \in A} \sum_{b \in B} x_a \, y_b \, m(a,b)$$

Objectif joueur 1 : faire en sorte que $\frac{1}{T} \sum_{t=1}^{T} m(I_t, J_t) \to \mathcal{C}$ ps

(où $\mathcal{C}$ est un ensemble)

vs

Objectif joueur 2 : tout faire pour que cette convergence n'ait pas lieu !

↳ Qui va gagner ?

On peut montrer (cf. convergences de martingale ou inégalité de Hoeffding)

que $\left| \frac{1}{T} \sum_t m(I_t, J_t) - \frac{1}{T} \sum_t m(x_t, y_t) \right| \to 0$ ps

Alors : si $(*)$ $\exists y^* \in \Delta(B) \mid \forall x \in \Delta(A),\ m(x,y^*) \notin \mathcal{C}$

le joueur 2, en choisissant $y_t \equiv y^*$ à chaque tour assure

que pour toute stratégie du joueur 1, $\frac{1}{T} \sum_t m(x_t, y^*) \not\to \mathcal{C}$

et même : $d\left(\frac{1}{T} \sum_t m(x_t, y^*), \mathcal{C} \right) \geq \delta$

où $\delta = \min_{x \in \Delta(A)} d(\underbrace{m(x,y^*), \mathcal{C}}_{\text{fonction continue en } x \text{ et } > 0 \ \forall x})$

En revanche, si $(**)$ $\forall y \in \Delta(B),\ \exists x \in \Delta(A) \mid m(x,y) \in \mathcal{C}$

alors il y a un espoir : si on voyait y_t avant de choisir notre x_t, on prendrait x_t tq. $m(x_t, y_t) \in \mathcal{C}$ et emballé c'est pesé !

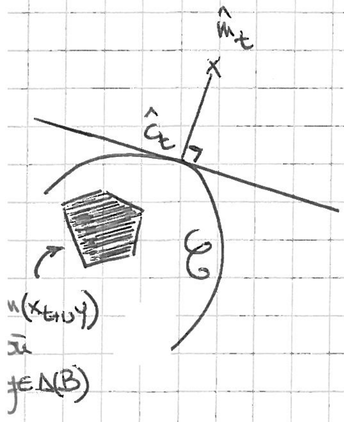
Le choix simultané de x_t et y_t empêche cela, mais la répétition du jeu va nous sauver → dans le cas où $\mathcal{C}$ est CONVEXE.

<u>Stratégie</u> : $t \geq 1$, $\hat{m}_t = \frac{1}{t} \sum_{s \leq t} m(x_s, J_s)$

$\hat{c}_t$ = projection euclidienne de $\hat{m}_t$ sur $\mathcal{C}$ convexe

$\exists x_{t+1} \in \Delta(A) \mid \forall y \in \Delta(B), \quad \langle \hat{m}_t - \hat{c}_t, m(x_{t+1}, y) - \hat{c}_t \rangle \leq 0$

↳ jouer un tel x_{t+1}



<u>Pourquoi ?</u> Montrer que

$$\min_x \max_y \langle \hat{m}_t - \hat{c}_t, m(x,y) - \hat{c}_t \rangle \overset{?}{\leq} 0$$

cf. valeur d'1 certain jeu ...

$$= \max_y \underbrace{\min_x \langle \hat{m}_t - \hat{c}_t, m(x,y) - \hat{c}_t \rangle}_{\text{par } (**) \ \exists x \mid m(x,y) \in \mathcal{C}}$$

et cf. proj. orthogonale :

$\forall c \in \mathcal{C}, \quad \langle \hat{m}_t - \hat{c}_t, c - \hat{c}_t \rangle \leq 0$

<u>Théorème</u> : Si $\mathcal{C}$ est convexe et sous l'hypothèse $(**)$ la stratégie ci-dessus assure que $\frac{1}{T} \sum_t m(x_t, J_t) \longrightarrow \mathcal{C}$

<u>Cor</u> : Par convergence de martingales / inégalité de Hoeffding, on a également $\frac{1}{T} \sum_t m(I_t, J_t) \rightarrow \mathcal{C}$ ps

<u>Eléments de preuve</u> :

$$d_{T+1}^2 \triangleq d\Big(\underbrace{\frac{1}{T+1} \sum_{t \leq T+1} m(x_t, J_t)}_{\hat{m}_{T+1}}, \mathcal{C}\Big)^2 \leq \| \overbrace{\hat{m}_{T+1} - \hat{m}_T}^{\frac{1}{T+1}(m(x_{T+1}, J_{T+1}) - \hat{m}_T)} + \hat{m}_T - \hat{c}_T \|^2$$

$$= \underbrace{\| \hat{m}_{T+1} - \hat{m}_T \|^2}_{\leq 4M^2/(T+1)^2} + \underbrace{\| \hat{m}_T - \hat{c}_T \|^2}_{= d_T^2}$$

$$+ \frac{2}{T+1} \underbrace{\Big\langle \substack{m(x_{T+1}, J_{T+1}) - \hat{m}_T \\ - \hat{c}_T + \hat{c}_T}, \hat{m}_T - \hat{c}_T \Big\rangle}_{\leq 0 \text{ (cf. pourquoi) } - d_T^2}$$

Au final :

$$d_{T+1}^2 \leq \left(1 - \frac{2}{T+1}\right) d_T^2 + \frac{4M^2}{(T+1)^2}$$

et par récurrence : $d_T^2 \leq 4M^2/T^2$.

Applications du théorème d'approchabilité : Minimisation du regret (interne) & Convergence vers les équilibres corrélés

oui, je sais, ça fait "philo", surtout le "interne"

Le joueur 1 a une stratégie telle que

$$\limsup \max_{k,l} \frac{1}{T} \sum_{t=1}^{T} \left(r_1(l, J_t) - r_1(k, J_t) \right) \mathbb{1}_{\{I_t = k\}} \leq 0 \quad \text{p.s.}$$

et donc également

(S1) $$\limsup \max_{\varphi : A \to A} \frac{1}{T} \sum_{t=1}^{T} r_1(\varphi(I_t), J_t) - \frac{1}{T} \sum_{t=1}^{T} r_1(I_t, J_t) \leq 0 \quad \text{ps}$$

Pourquoi ? / Conséquence :

Pourquoi ? Th. d'approchabilité, $m : A \times B \to \mathbb{R}^{A \times A}$

avec $m(i,y) = \left[\left(r_1(l,y) - r_1(k,y) \right) \mathbb{1}_{\{i=k\}} \right]_{(k,l)}$

↳ $\forall y$, en prenant $i^*(y) \in \arg\max_k r_1(k,y)$, on a :

$$m(i^*(y), y) \in \underbrace{(\mathbb{R}_-)^{A \times A}}_{= \mathcal{C}}$$

La CNS de Blackwell est vérifiée pour x = Dirac en $i^*(y)$.

Conséquence ? Symétriquement le joueur 2 a une stratégie telle que

(S2) $$\limsup \max_{\psi : B \to B} \frac{1}{T} \sum_{t=1}^{T} r_2(I_t, \psi(J_t)) - \frac{1}{T} \sum_{t=1}^{T} r_2(I_t, J_t) \leq 0 \quad \text{ps}$$

On note $\pi_m = \frac{1}{T} \sum_{t=1}^{T} \delta_{(I_t, J_t)}$ la loi empirique des coups d'actions joués.

Th : (S1) et (S2) entraînent la convergence de $(\pi_n)_n$ vers l'ensemble des équilibres corrélés :

$$\inf_{\pi \in \mathrm{CORR}} \| \pi_n - \pi \| \longrightarrow 0 \quad \text{ps}$$

Rq : Pas de convergence, en général, vers un équilibre corrélé donné.

Preuve : Cf. raisonnement par l'absurde et le fait que $\Delta(A\times B)$ soit compact : il suffit de montrer que toute valeur d'adhérence de (π_n) est dans Corr.

Soit $\pi_{k_n} \to \pi$:

(S1) se réécrit comme $\underbrace{\limsup_{k_n\to+\infty}}_{\text{vraie limite}} \mathbb{E}_{\pi_{k_n}}\left[r_1(\varphi(I),J)\right] - \mathbb{E}_{\pi_{k_n}}\left[r_1(I,J)\right] \leq 0$ ps

$$= \mathbb{E}_{\pi}\left[r_1(\varphi(I),J)\right] - \mathbb{E}_{\pi}\left[r_1(I,J)\right]$$

et de même par (S2) :

$$\mathbb{E}_{\pi}\left[r_2(I,\psi(J))\right] - \mathbb{E}_{\pi}\left[r_2(I,J)\right] \leq 0$$

On a bien $\pi \in$ Corr.

Peut-on converger vers l'ensemble des équilibres de Nash ?

* Algorithmes computationnellement inefficaces (et encore : souvent convergence vers ε-Nash), preuves longues et douloureuses...

* Sauf (bien sûr !) le cas particulier des jeux à somme nulle, $r_1 = -r_2$, où les équilibres de Nash correspondent aux équilibres minmax : on peut montrer que (S1) et (S2) entraînent

$$(\hat{x}_T, \hat{y}_T) = \left(\frac{1}{T}\sum_{t=1}^{T}\delta_{I_t}, \frac{1}{T}\sum_{t=1}^{T}\delta_{J_t}\right) \longrightarrow \mathcal{V}$$

convergence du produit des marginals empiriques qui est une quantité moins naturelle...