Une introduction aux jeux stochastiques à somme nulle

Guillaume Vigeral

CEREMADE Université Paris-Dauphine

$20-24$ Mai 2019, Talence
Colloque Inter'Actions en Mathématiques 2019

Table of contents

Table of contents

Modèle

Un jeu statique à somme nulle Γ est défini par :

- Un ensemble non vide I d'actions pures du joueur 1.
- Un ensemble non vide J d'actions pures du joueur 2.
- Une fonction de paiement bornée $g : I \times J \to [-M, M]$.

Déroulement du jeu : simultanément le joueur 1 (resp. 2) choisit $i \in I$ (resp. $j \in J$),et cherche à maximiser (resp. minimiser) le paiement résultant $g(i,j)$.

La description du jeu est connaissance commune.

Généralisation des problèmes d'optimisation à des situations à deux agents totalement antagonistes.

Exemples

J1 choisit une ligne, J2 une colonne.
Exemple 1 : "Tir au but"

	Saute à gauche	Saute à droite
Tir à gauche	0	1
Tir à droite	1	0

Exemple 2 : "Tir au but avec tireur borgne et mauvais gardien"

	Saute à gauche	Saute à droite
Tir à gauche	$2/3$	1
Tir à droite	0	0

Valeur en actions pures

$$\underline{V} = \sup_{i \in I} \left(\inf_{j \in J} g(i,j) \right)$$

$\underline{V}$ est la valeur du jeu où J1 joue en premier, puis J2, en ayant observé le choix de J1. Plus grande quantité que J1 peut garantir.

$$\overline{V} = \inf_{j \in J} \left(\sup_{i \in I} g(i,j) \right)$$

$\underline{V}$ est la valeur du jeu où J2 joue en premier, puis J1, en ayant observé le choix de J2. Plus petite quantité que J2 peut garantir.

Clairement $\underline{V} \leq \overline{V}$. Si égalité, la valeur commune V est la valeur du jeu en actions pures.

Retour aux exemples

Exemple 1

	Saute à gauche	Saute à droite
Tir à gauche	0	1
Tir à droite	1	0

$0 = \underline{V} < \overline{V} = 1$. Pas de valeur.

Exemple 2

	Saute à gauche	Saute à droite
Tir à gauche	$2/3$	1
Tir à droite	0	0

$2/3 = \underline{V} = \overline{V} = 2/3$. Valeur $V = 2/3$.

Jeu en actions mixtes

Supposons pour le moment I et J finis. Soit $X := \Delta(I)$ l'ensemble des probas sur I. De même $Y := \Delta(J)$ on étend bilinéairement g à $X \times Y$.

Maintenant J1 choisit $x \in X$, J2 choisit $y \in Y$, et le paiement est $g(x, y)$.

Interprétation : les joueurs peuvent jouer de manière aléatoire, et maximisent (ou minimisent) l'espérance de gain.

Valeur en actions mixtes

$$\underline{v} = \max_{x \in X} \left(\min_{y \in Y} g(x,y) \right) = \max_{x \in X} \left(\min_{j \in J} g(x,j) \right) \tag{1}$$

$$\overline{v} = \min_{y \in Y} \left(\max_{x \in X} g(x,y) \right) = \min_{y \in Y} \left(\max_{i \in I} g(i,y) \right) \tag{2}$$

Si égalité, le jeu admet une valeur v en actions mixtes. Une action qui réalise le $\max$ dans (1) (resp. le $\min$ dans (2)) est dite optimale pour J1 (resp. J2).

Théorème du minimax

Théorème (von Neumann, 1928)

Tout jeu fini a une valeur en stratégie mixte, et chaque joueur a (au moins) une action optimale.

Preuve originelle via le théorème du point fixe de Brouwer, preuves ultérieures via Hahn-Banach ou la dualité en programmation linéaire.

Remarques

- Si les entrées sont des rationnels, la valeur également.
- La valeur est calculable en temps polynomial.
- Le résultat s'étend dans des cas plus généraux (I et J infinis) avec des hypothèses de compacité et continuité. Mais dans ce cas les stratégies optimales peuvent être très complexes : jeu où $I = J = [0, 1]$, g fraction rationnelle et où le support de l'unique stratégie optimale de chaque joueur est un ensemble de Cantor.

Table of contents

Table of contents

Modèle

Un jeu stochastique fini à somme nulle Γ est 5-uplet (Ω, I, J, g, ρ) où :

- Ω est l'ensemble fini d'états.
- I (resp. J) est l'ensemble fini d'actions du joueur 1 (resp. Joueur 2).
- $g : I \times J \times \Omega \to [-M, M]$ est la fonction de paiement (que J1 maximise et J2 minimise).
- $\rho : I \times J \times \Omega \to \Delta(\Omega)$ est la probabilité de transition.

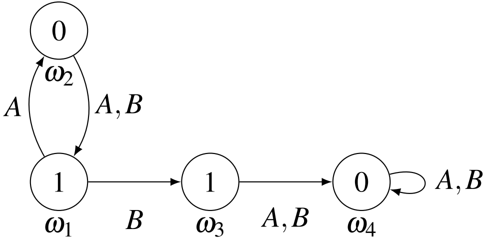
Déroulement du jeu

Un état initial ω_1 est donné et observé par chacun.
A chaque étape $t \in \mathbb{N}$:

- Les joueurs observent l'état courant ω_t et se rappellent de tout le passé.
- Simultanément J1 (resp. J2) choisit une action mixte x_t dans $X = \Delta(I)$ (resp. y_t dans $Y = \Delta(J)$).
- Une action i_t du joueur 1 (resp. j_t du joueur 2) est tirée en fonction de x_t (resp. y_t) et observée par tous.
- Paiement d'étape t : $g_t = g(i_t, j_t, \omega_t)$.
- Nouvel état ω_{t+1} tiré selon $\rho(i_t, j_t, \omega_t)$.

Exemple 1 : un processus de décision Markovien (MDP)

Un seul joueur (qui maximise).
2 actions A and B, 4 états, transitions déterministes, paiement ne dépend que de l'état courant.



Exemple 2 : le "Big Match"

3 états : 1^*, 0^* et ω_0. 2 actions par joueur.
1^* état absorbant de paiement 1 : quelques soient les actions le paiement est de 1 et l'état ne change pas .
0^* état absorbant de paiement 0 : quelques soient les actions le paiement est de 0 et l'état ne change pas .
Dans ω_0 :

	Left	Right
Top	0^*	1^*
Bottom	1	0

où une étoile représente une transition vers l'état absorbant correspondant et dans les autres entrées l'état reste en ω_0.

Histoires et stratégies

H_t est l'ensemble des histoires avant la date t :
$H_t := (\Omega \times I \times J)^{t-1} \times \Omega$.
Une stratégie (de comportement) du joueur 1 est une fonction

$$\sigma : \cup_t H_t \longrightarrow X$$

Une stratégie (de comportement) du joueur 2 est une fonction

$$\tau : \cup_t H_t \longrightarrow X$$

Un couple (σ, τ) et un état initial ω génèrent une unique probabilité $\mathbb{P}(\sigma, \tau, \omega)$ sur l'ensemble des parties $H_\infty = (\Omega \times I \times J)^\infty$. A partir de maintenant $\mathbb{E}(\sigma, \tau, \omega)$ désigne l'espérance par rapport à cette probabilité.

Jeu en n étapes

Soient $n \in \mathbb{N}$, et $\omega \in \Omega$, le jeu en n étapes $\Gamma_n(\omega)$ est le jeu de paiement

$$\gamma_n^{\omega}(\sigma, \tau) := \mathbb{E}_{\sigma,\tau,\omega}\left\{\frac{1}{n}\sum_{t=1}^{n} g_t\right\},$$

où J1 maximise et J2 minimize.

La valeur de $\Gamma_n(\omega)$ existe par le théorème du minimax de von Neumann et est notée $v_n(\omega)$.

Le jeu λ-escompté

Soient $\lambda \in]0,1[$, et $\omega \in \Omega$, le jeuλ-escompté est le jeu de paiement

$$\gamma_\lambda^\omega(\sigma,\tau) := \mathbb{E}_{\sigma,\tau,\omega}\left\{\lambda\sum_{t=1}^{+\infty}(1-\lambda)^{t-1}g_t\right\},$$

où J1 maximise et J2 minimize.

La valeur de $\Gamma_\lambda(\omega_1)$ existe par un théorème du minimax et est notée $v_\lambda(\omega)$.

Table of contents

Structure récursive

Shapley (1953) : les valeurs ont une structure récursive (principe de programmation dynamique)

$$
\begin{aligned}
v_n(\omega) &= \max_{x\in X}\min_{y\in Y}\left\{\frac{1}{n}g(x,y,\omega)+\frac{n-1}{n}E_{\rho(x,y,\omega)}(v_{n-1}(\cdot))\right\}\\
&= \min_{y\in Y}\max_{x\in X}\left\{\frac{1}{n}g(x,y,\omega)+\frac{n-1}{n}E_{\rho(x,y,\omega)}(v_{n-1}(\cdot))\right\}\\
v_\lambda(\omega) &= \max_{x\in X}\min_{y\in Y}\left\{\lambda g(x,y,\omega)+(1-\lambda)E_{\rho(x,y,\omega)}(v_\lambda(\cdot))\right\}\\
&= \min_{y\in Y}\max_{x\in X}\left\{\lambda g(x,y,\omega)+(1-\lambda)E_{\rho(x,y,\omega)}(v_\lambda(\cdot))\right\}.
\end{aligned}
$$

Conséquences

- In Γ_n, les joueurs ont des stratégies optimales Markoviennes : elles ne dépendent que de la date et de l'état courant.
- In Γ_λ ples joueurs ont des stratégies optimales stationnaires : elles ne dépendent que de l'état courant.
- Ainsi v_n et v_λ ne dépendent pas de l'observation des actions.

Modéle général

Tout ceci s'étend à des modèles plus généraux (I, J et Ω infinis) avec des hypothèses de compacité, continuité, mesurabilité,...

Table of contents

Questions ?

1. Convergence de v_n et v_λ quand n tend vers $+\infty$ ou quand λ tend vers 0 ?
2. Existence de stratégies optimales robustes ?
3. Egalité de $\lim v_n$ et $\lim v_\lambda$?
4. Caractérisation de la limite ?
5. Calcul de la limite ?

Table of contents

Table of contents

Preuve d'existence de la limite des valeurs escomptées

Pour chaque λ le joueur a une stratégie pure stationnaire optimale dans Γ_λ.
Etant donnée une stratégie pure stationnaire $h : \Omega \to I$, le paiement λ-escompté avec état inital ω vérifie

$$\gamma_\lambda^\omega(h) = \lambda g(\omega, h(\omega)) + (1-\lambda) \sum_{\omega'} \rho(\omega'|\omega, h(\omega)) \gamma_\lambda^w(h)$$

et le vecteur de paiement $\gamma_\lambda(h) := (\gamma_\lambda^\omega(h))_{\omega\in\Omega}$ vérifie donc

$$\gamma_\lambda(h) = \lambda U_h + (1-\lambda) A_h \gamma_\lambda(h)$$

pour une matrice stochastique A_h et un vecteur U_h. Donc

$$\gamma_\lambda(h) = (Id - (1-\lambda)A_h)^{-1}(\lambda U_h)$$

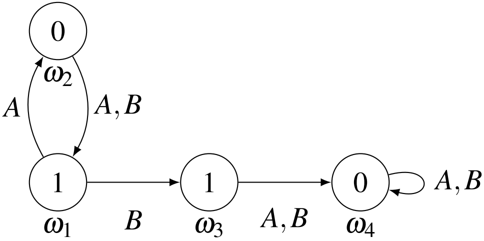
et $\gamma_\lambda^\omega(h)$ est donc une fonction rationnelle de λ.

Suite de la preuve

- A h et ω fixés, $\gamma_\lambda^\omega(h)$ est une fonction rationnelle de λ.
- Pour chaque h, h' and ω, $\gamma_\lambda^\omega(h) - \gamma_\lambda^\omega(h')$ a un signe fixe pour λ petit.
- Pour tout ω, il existe h optimal dans $\Gamma_\lambda(\omega)$ pour λ suffisamment petit (Blackwell optimalité).
- Donc $\lim_{\lambda\to 0} v_\lambda(\omega) = \lim_{\lambda\to 0} \gamma_\lambda^\omega(h)$
- Cette limite existe donc dans $[-\infty, +\infty]$, et donc dans $\mathbb{R}$ puisque paiement est borné.

Remarque 1 : Si le MDP est défini par des nombres rationnels, $v(\omega)$ est rationnel.
Remarque 2 : Calculable en temps polynomial.
Remarque 3 : On peut montrer que v_n converge vers la même limite.

Exemple 1



$$v_\lambda(\omega_1) = \begin{cases} \lambda(2-\lambda) & \text{if } \lambda \geq \frac{3-\sqrt{5}}{2} \\ \frac{1}{2-\lambda} & \text{if } \lambda \leq \frac{3-\sqrt{5}}{2} \end{cases}$$

$$v_n(\omega_1) = \begin{cases} \frac{n+2}{2n} & \text{si } n \text{ est pair} \\ \frac{n+1}{2n} & \text{si } n \text{ est impair} \end{cases}$$

Table of contents

Outil principal : théorème de Tarski-Seidenberg

Un ensemble $E \subset \mathbb{R}^d$ est dit semi-algébrique s'il peut s'écrire comme union et intersection d'un nombre finis d'ensemble définis par une inégalité (large ou stricte) polynomiale.
Exemple :

$$\begin{aligned} E &= \{(a,b,c,x) | ax^2+bx+c=0\} \\ &= \{(a,b,c,x) | ax^2+bx+c \geq 0\} \cap \{(a,b,c,x) | ax^2+bx+c \leq 0\}. \end{aligned}$$

Théorème (Tarski-Seidenberg)

Toute projection d'un ensemble semiagébrique reste semialgébrique.

Par exemple la projection de l'ensemble ci-dessus sur les trois premières coordonnées est

$$\{(a,b,c) | b^2-4ac \geq 0\} \cap \left(\{(a,b,c) | a^2+b^2>0\} \cup \{(a,b,c) | c=0\}\right)$$

Outil principal : théorème de Tarski-Seidenberg

Un ensemble $E \subset \mathbb{R}^d$ est dit semi-algébrique s'il peut s'écrire comme union et intersection d'un nombre finis d'ensemble définis par une inégalité (large ou stricte) polynomiale.
Exemple :

$$\begin{aligned} E &= \{(a,b,c,x) | ax^2+bx+c=0\} \\ &= \{(a,b,c,x) | ax^2+bx+c \geq 0\} \cap \{(a,b,c,x) | ax^2+bx+c \leq 0\}. \end{aligned}$$

Théorème (Tarski-Seidenberg)

Toute projection d'un ensemble semiagébrique reste semialgébrique.

Par exemple la projection de l'ensemble ci-dessus sur les trois premières coordonnées est

$$\{(a,b,c) | b^2-4ac \geq 0\} \cap \left(\{(a,b,c) | a^2+b^2>0\} \cup \{(a,b,c) | c=0\}\right)$$

Théorème de Bewley Kohlberg

Théorème (Bewley-Kohlberg 76)

Dans tout jeu stochastique fini, v_λ converge lorsque λ tend vers 0.

Idée de la preuve : on considère l'ensemble

$$E := \{(\lambda, v_\lambda(\cdot), x_\lambda(\cdot), y_\lambda(\cdot))\}$$

avec

- $\lambda \in]0,1]$
- $v_\lambda(\omega)$ la valeur du jeu λ escompté d'état initial ω
- $x_\lambda(\omega) \in \mathbb{R}^I$ une stratégie mixte optimale du joueur 1 dans l'équation de Shapley.
- $y_\lambda(\omega) \in \mathbb{R}^I$ une stratégie mixte optimale du joueur 2 dans l'équation de Shapley.

Théorème de Bewley Kohlberg

$E \subset \mathbb{R}^{(1+|\Omega|+|\Omega||I|+|\Omega||J|)}$ est semi algébrique car défini par des inégalités du genre

$$v_\lambda(\omega) \leq \lambda \sum_i x_\lambda(\omega)(i) g(i,j,\omega) + (1-\lambda) \sum_{i,\omega'} x_\lambda(\omega)(i) \rho(\omega'|i,j,\omega) v_\lambda(\omega')$$

pour tout ω et j ; et

$$\sum_i x_\lambda(\omega)(i) = 1$$

pour tout ω.
Fixons ω, l'ensemble $\{\lambda, v_\lambda(\omega)\}$, graphe de la fonction $\lambda \to v_\lambda(\omega)$, est donc semi-algébrique par Tarski-Seidenberg.

Théorème de Bewley Kohlberg

Cela entraîne que $v_\lambda(\omega)$ peut s'écrire comme une série de Puiseux :

$$v_\lambda(\omega) = \sum_{k \geq k_0} a_k \lambda^{\frac{k}{C}}$$

pour tout $\lambda \leq \lambda(0)$, avec $k_0 \in \mathbb{Z}$ et $C \in \mathbb{N}^*$.
Puisque v_λ est bornée $k_0 \geq 0$ et $v_\lambda(\omega)$ converge vers a_0.
Remarques :

- On montre que $v_n(\omega)$ converge vers la même limite.
- Pas de Blackwell optimalité.
- Si le jeu est défini par des nombres rationnels, la limite $v(\omega)$ est algébrique.
- Pas d'algorithme polynomial (pour le moment) .

Exemple 2

	Left	Right
Top	0^*	1^*
Bottom	1	0

Vlalues : $v_\lambda = v_n = 1/2$.
Stratégie optimale pour J^2 : $y_\lambda = y_n = 1/2L + 1/2R$.
Stratégie optimale pour J^1 : $x_\lambda = \frac{\lambda}{1+\lambda}\text{Top} + \frac{1}{1+\lambda}\text{Bottom}$.
Stratégie optimale pour J^1 : $x_n = \frac{1}{n+1}\text{Top} + \frac{n}{n+1}\text{Bottom}$.

Table of contents

Définition

Idée : valeur uniforme v si les joueurs peuvent garantir v dans tout jeu long sans connaître la longueur précise.
Joueur 1 (resp. 2) peut garantir que "$\lim_{n\to\infty}\mathbb{E}\{\frac{1}{n}\sum_{t=1}^{n} g_t(\cdot)\}$" est supérieur (resp. inférieur) à v.
Définition rigoureuse : $v(\omega)$ valeur uniforme du jeu avec état initial ω si

$$\forall \varepsilon > 0, \exists T, \exists \sigma, \forall \tau, \forall t > T, \gamma_n^{\omega}(\sigma,\tau) \geq v - \varepsilon$$

et une formule duale.
Si v valeur uniforme, alors v_n converges vers v (trivial) and v_λ également.

Valeur uniforme dans le Big Match : premier essai

	Left	Right
Top	0^*	1^*
Bottom	1	0

$v_n = v_\lambda = 1/2$ pour tout λ and et n. Stratégie optimale pour J^1 : jouer Top avec proba $\frac{1}{n+1}$ and $\frac{\lambda}{1+\lambda}$ respectivement.
Facile de montrer que $1/2L + 1/2R$ garantit $1/2$ au joueur 2 pour tout λ et n.
Que dois jouer J^1 ?
Première idée : jouer la limite des stratégies optimales stationnaires dans Γ_λ lorsque λ tend vers 0 : Bottom. Garantit seulement 0.
Problème $\lim_{\lambda\to 0} \gamma_\lambda^\omega(x_\lambda,\cdot) \neq \lim_{\lambda\to 0} \gamma_\lambda^\omega(\lim_\lambda x_\lambda,\cdot)$

Valeur uniforme dans le Big Match : deuxième essai

	Left	Right
Top	0^*	1^*
Bottom	1	0

Essayons des stratégies stationnaires. Bottom garantit seulement 0.

Et si J^1 joue $\alpha \mathrm{Top} + (1-\alpha)\mathrm{Bottom}$, $\alpha > 0$?

Alors J^2 joue L à chaque étape et le paiement est ε-près de 0 quand l'horizon est long.

Valeur uniforme dans le Big Match : troisième essai

	Left	Right
Top	0^*	1^*
Bottom	1	0

Essayons une stratégie Markovienne : J^1 joue Top avec proba α_t à l'étape t. Alors :

- Si $\Pi_{t=1}^{+\infty}(1-\alpha_t)=0$, la partie finira presque surement par être absorbée. J^2 joue L à chaque étape et le paiement est ε-près de 0 quand l'horizon est long.
- Si $\Pi_{t=1}^{+\infty}(1-\alpha_t)>0$ alors $\Pi_{t=T}^{+\infty}(1-\alpha_t)\geq 1-\varepsilon$ pour un certain T. J^2 joue L jusqu'à l'étape T puis R pour toujours et le paiement est 2ε-près de 0 quand l'horizon est long.

Valeur uniforme dans le Big Match : des idées ?

- La probabilité α_t avec laquelle J^1 joue Top doit dépendre des actions passées de J^2. !
- Si J^2 a beaucoup joué L dans le passé, le paiement courant est bon donc pas besoin de prendre de risque : on joue Top avec proba petite.
- Si J^2 a beaucoup joué R dans le passé, le paiement courant est mauvais donc il faut jouer Top avec proba plus grande.
- La proabilité de jouer Top doit être faible au début du jeu et varier lentement.
- Si J^2 joue L plus de la moitié du temps, il faudrait que $\Pi_{t=1}^{+\infty}(1-\alpha_t) > 0$; si J^2 joue L moins de la moitié du temps il faudrait que $\Pi_{t=1}^{+\infty}(1-\alpha_t) = 0$.

Valeur uniforme dans le Big Match : quatrième essai

	Left	Right
Top	0^*	1^*
Bottom	1	0

Soit K un entier grand, et notons L_t et R_t le nombres de L et R joués entre les dates 1 and T.

J^1 joue Top à l'étape $t+1$ avec proba $\frac{1}{(K+L_t-R_t)^2}$.

Calculs... Ca marche.

Theorem (Blackwell Ferguson '68)

Le Big Match a une valeur uniforme.

Cas général

Theorem (Mertens Neyman '81)

Tout jeu fini admet une valeur uniforme.

Idée : jouer une action optimale dans Γ_λ pour un λ qui dépend du passé et est d'autant plus petit que notre paiement courant est bon.
Ingrédient clé : v_λ varie lentement, conséquence de l'expression en série de Puiseux.

Table of contents

Questions ?

1. Convergence de v_n et v_λ quand n tend vers $+\infty$ ou quand λ tend vers 0 ?
2. Existence de stratégies optimales robustes ?
3. Egalité de $\lim v_n$ et $\lim v_\lambda$?
4. Caractérisation de la limite ?
5. Calcul de la limite ?

Valeur limite ?

Convergence de v_n et v_λ quand n tend vers $+\infty$ ou quand λ tend vers 0 ?

1. Oui dans le cas fini (Bewley et Kohlberg '76). Preuve "élémentaire" récente (Oliu-Barton 2014).
2. Oui pour un jeu absorbant sans hypothèse de finitude (Rosenberg Sorin '01)
3. Contre exemple avec Ω fini, I et J compacts, g et ρ continues (V. '13).
4. Problème ouvert : 4 états, I et J intervalles compacts, g et ρ fractions rationnelles. Difficulté : même dans ce cas simple v_λ n'est pas forcément semi-algébrique.

Valeur limite ?

Convergence de v_n et v_λ quand n tend vers $+\infty$ ou quand λ tend vers 0 ?

1. Oui dans le cas fini (Bewley et Kohlberg '76). Preuve "élémentaire" récente (Oliu-Barton 2014).
2. Oui pour un jeu absorbant sans hypothèse de finitude (Rosenberg Sorin '01)
3. Contre exemple avec Ω fini, I et J compacts, g et ρ continues (V. '13).
4. Problème ouvert : 4 états, I et J intervalles compacts, g et ρ fractions rationnelles. Difficulté : même dans ce cas simple v_λ n'est pas forcément semi-algébrique.

Théorème tauberiens

La convergence (uniforme) de v_λ implique t-elle celle de v_n (et vice-versa) ?

1. Oui à 0 joueurs (Hardy-Littlewood 1914)
2. Oui à 1 joueur (Lehrer-Sorin 1991)
3. Oui à 2 joueurs (Ziliotto 2016).
4. Problème ouvert : évaluations plus générales des paiements $\sum_{t=1}^{+\infty} \theta_t g_t$.

Théorème tauberiens

La convergence (uniforme) de v_λ implique t-elle celle de v_n (et vice-versa) ?

1. Oui à 0 joueurs (Hardy-Littlewood 1914)
2. Oui à 1 joueur (Lehrer-Sorin 1991)
3. Oui à 2 joueurs (Ziliotto 2016).
4. Problème ouvert : évaluations plus générales des paiements $\sum_{t=1}^{+\infty} \theta_t g_t$.

Formules

Formule “simple” pour la limite des valeurs ?
Oui dans le cas absorbant (Laraki 2010, Sorin-V. 2019). Ouvert en général, même pour des jeux finis à 4 états.

Complexité

Complexité du calcul de la limite des valeurs, dans le cas fini ?
Complexité du problème de décision : est-ce que $\lim v_n(\omega) \geq 0$ pour un ω donné ?
On sait le faire en temps exponentiel (Chatterjee Majumdar Heinzinger 2008) et le problème est même dans $NP \cap co(NP)$ dans le cas d'observation parfaite, mais aucun algorithme polynomial connu pour le moment.

Domaines connexes

- Jeux avec signaux, par exemple les joueurs n'observent pas l'état courant.
- Jeux avec objectifs non numériques : jeux de parité.
- Jeux en temps continu : jeux différentiels
- Jeux stochastiques à N joueurs
- ...

Définition formelle de la Zibeline