JEUX À DEUX JOUEURS À SOMME NULLE

DYLAN LAPLACE MERMOUD

Brèves notes de cours adaptées du cours de Tristan Garrec et du livre intitulé "Bases mathématiques de la théorie des jeux" de Rida Laraki, Jérôme Renault et Sylvain Sorin.

1. Premières définitions

Dans ce cours, nous allons étudier une classe de jeux pour laquelle l'équilibre de Nash a une grande importance. Nous avons fréquemment rencontré des exemples de jeux symétriques lors des deux cours précédents, c'est-à-dire des jeux pour lesquels les fonctions de paiements sont les mêmes pour chaque joueurs. Ici, nous sommes intéressés par des que nous pouvons considérer comme antisymétriques.

Définition 1. Un jeu à deux joueurs à somme nulle est un jeu sous forme normale $\Gamma = (N, S, g)$ tel que $|N| = 2$ et $g_1 = -g_2$.

Dans un jeu à deux joueurs à somme nulle, les intérêts des deux joueurs sont diamétralement opposés, ce que l'un gagne, l'autre le perd. De plus, les données d'un jeu à deux joueurs à somme nulle sont plus compactes: le jeu est complètement déterminé par la donnée de l'ensemble de stratégie du premier joueur S, l'ensemble de stratégies du second joueur T et de la fonction de paiement du premier joueur $g : S \times T \to \mathbb{R}$. Le premier joueur va alors chercher à maximiser la fonction de paiement g, tandis que le second joueur va chercher à la minimiser.

Exemple 1. Reprenons encore une fois l'exemple du "Tir de penalty", donné par

	G	D
G	$1, -1$	$-1, 1$
D	$-1, 1$	$1, -1$

On remarque que dans chaque cellule, la somme des gains est bien égale à 0. Ainsi, il suffit simplement de connaitre le paiement du premier joueur pour chaque profil. On peut alors représenter le jeu de la manière suivante :

	G	D
G	1	-1
D	-1	1

$\Diamond$

Les deux joueurs s'affrontant frontalement, l'analyse des équilibres y est légèrement simplifié, dû à la grande correlation des deux fonctions de paiements.

Définition 2. Le *supinf* du jeu Γ, noté $\underline{v}(\Gamma)$ (ou simplement $\underline{v}$), est la quantité

$$\underline{v} \coloneqq \sup_{s \in S} \inf_{t \in T} g(s,t).$$

Lorsque le sup et l'inf sont atteints, on parle également de *maxmin*.

De manière générale, pas abus de langage, on parle de *maxmin* même lorsque le sup ou l'inf ne sont pas atteint. Le maxmin d'un jeu représente la quantité que le joueur j_1 peut s'assurer peu importe ce que le joueur j_2 fait.

Proposition 1. *Le maxmin $\underline{v}$ est caractérisé par*

- ▷ *Le joueur j_1* garantit *$\underline{v}$ à ε près, c'est-à-dire, pour tout $\varepsilon > 0$, il existe une stratégie pour le premier joueur $s \in S$ telle que, pour toute stratégie du second joueur $t \in T$, on a $g(s,t) \geq \underline{v} - \varepsilon$;*
- ▷ *Le joueur j_2* défend *$\underline{v}$) ε près, c'est-à-dire, pour tout $\varepsilon > 0$ et pour toute stratégie du premier joueur $s \in S$, il existe une stratégie pour le second joueur $t \in T$ telle que $g(s,t) \leq \underline{v} + \varepsilon$.*

Posons $f = \inf_{t \in T} g(\cdot, t)$ une fonction définie sur l'ensemble de stratégie S du joueur j_1. La fonction f peut être interprétée comme une fonction de meilleure réponse pour le joueur j_2, qui prend en entrée une stratégie $s \in S$ et retourne la valeur minimale de g possible pour ce choix de stratégie. Dans cette situation, le joueur j_2 a un avantage important, car il sait ce que le joueur j_1 joue, et choisit sa stratégie en fonction. On a alors $\underline{v} = \sup_{s \in S} f(s)$, et le maxmin du jeu représente la valeur que le joueur j_1 peut obtenir dans le cas où le joueur j_2 joue nécessairement la meilleure réponse.

Il s'agit d'un calcul *prudent*, ce nombre étant déterminé dans le pire cas pour le joueur j_1, afin de déterminer une borne sous laquelle son paiement ne doit rationnellement pas se trouver. C'est ce comportement que nous retrouvons dans la proposition ci-dessus.

Il existe ce même calcul prudent pour donner une borne inférieure au paiement du joueur j_2, appelé le *minmax*, noté $\overline{v}$ et défini par

$$\overline{v} \coloneqq \inf_{t \in T} \sup_{s \in S} g(s,t).$$

Il s'agit de la même idée, mais cette fois-ci les rôles des joueurs j_1 et j_2 sont inversés.

Exemple 2. Dans l'exemple du "Tir de penalty", le maxmin du jeu est -1, car l'adversaire connait notre stratégie. À l'inverse, le minmax est 1, car nous sommes capable de jouer là où l'adversaire ne se trouve pas. ♢

Proposition 2. *Pour tout jeu, on a $\underline{v} \leq \overline{v}$.*

La différence $\overline{v} - \underline{v} \geq 0$ entre le minmax et le maxmin est appelé le *saut de dualité*.

Définition 3. Si $\overline{v} - \underline{v} = 0$, le jeu a une *valeur* (en stratégies pures), notée $v = \underline{v} = \overline{v}$.

Lorsqu'un jeu admet une valeur, il s'agit du seul paiement auquel on peut s'attendre si les joueurs jouent de manière optimale.

Exemple 3. Le jeu représenté par la matrice de gain suivante

	G	D
G	1	-1
D	1	1

a une valeur $v = 1$. En effet, même le maxmin, qui est censé être à l'avantage du joueur j_2 n'est pas négatif, car la stratégie D du joueur j_1 lui assurera toujours 1. ♢

On s'intéresse à présent aux stratégies atteignant cette valeur.

2. Stratégies optimales

Définition 4. Considérons un jeu Γ admettant une valeur v. Soit $\varepsilon > 0$.

- ▹ Une stratégie $s \in S$ est *ε-optimale* pour j_1 si elle lui garantit $v - \varepsilon$, c'est-à-dire, si pour toute stratégie adverse $t \in T$, on a $g(s,t) \geq v - \varepsilon$;
- ▹ Une stratégie $t \in T$ est *ε-optimale* pour j_2 si elle lui garantit $v + \varepsilon$, c'est-à-dire, si pour toute stratégie adverse $s \in S$, on a $g(s,t) \leq v + \varepsilon$.

Les stratégies 0-optimales sont simplement dites *optimales*.

Définition 5. Le profil (s^*, t^*) est un *point-selle* si pour tout profil (s,t), on a

$$g(s, t^*) \leq g(s^*, t^*) \leq g(s^*, t).$$

On voit que chaque déviation unilatérale d'un point-selle réduit le paiement du joueur qui dévie, ainsi c'est un équilibre de Nash.

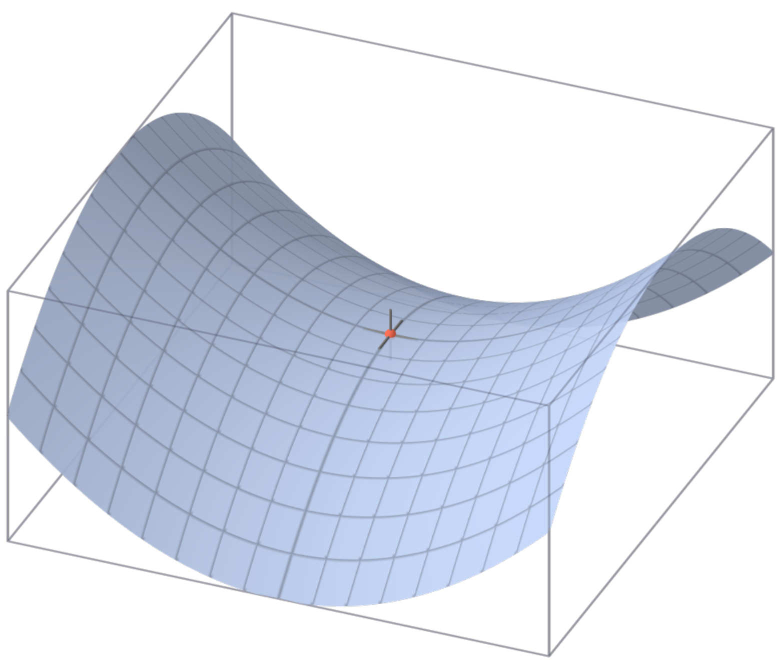

FIGURE 1. Un point selle est un point critique qui n'est pas un optimum.

Proposition 3. *S'il existe un équilibre de Nash en stratégies pures* (s^*, t^*)*, alors le jeu a une valeur* $g(s^*, t^*)$*. De plus,* s^* *et* t^* *sont des stratégies optimales. Réciproquement, si un jeu a une valeur* v *et des stratégies optimales* (s^*, t^*)*, alors* (s^*, t^*) *est un équilibre de Nash et* $v = g(s^*, t^*)$*.*

Cette proposition propose une caractérisation des équilibres de Nash en stratégies pures grâce à des propriétés seulement unilatérales des stratégies qui les composent. Au lieu de devoir chercher des profils pour lesquels chacune des stratégies est une meilleure réponde au profil partiel correspondant, il suffit, dans les jeux à deux joueurs à somme nulle, de trouver des stratégies optimales.

Bien que les stratégies optimales ressemblent aux stratégies dominantes, il peut exister des stratégies optimales qui ne sont pas des stratégies dominantes.

On s'intéresse maintenant aux propriétés de l'opérateur "valeur", que l'on peut voir comme une fonction définie sur l'espace des jeux.

Proposition 4. *Soient* $c \in \mathbb{R}$ *un réel,* $\Gamma = (S, T, g)$ *et* $\Gamma' = (S, T, g')$ *deux jeux définis sur les mêmes ensembles de stratégies, et* $\Gamma'' = (S, T, g + c)$ *étant défini comme* Γ*, mais en translatant chaque paiement par* c*. Alors on a*

- ▷ monotonie*: si* $g \leq g'$*, alors* $\underline{v}(\Gamma) \leq \underline{v}(\Gamma')$ *;*
- ▷ translation des constantes*: on a* $\underline{v}(\Gamma'') = \underline{v}(\Gamma) + c$ *;*
- ▷ contraction*: on a* $|\underline{v}(\Gamma) - \underline{v}(\Gamma')| \leq \|g - g'\|_\infty$*.*

La dernière propriété découle des deux précédentes. En effet, on a nécessairement

$$g \leq g' + \|g - g'\|_\infty,$$

on peut alors appliquer l'opérateur maxmin de chaque côté de l'inégalité par monotonie, et enfin la translation des constantes.

La proposition ci-dessus s'applique de la même manière à l'opérateur minmax, et à l'opérateur valeur si celui-ci est bien défini.

On rappelle les définitions de stratégies strictement dominées et faiblement dominées. Une stratégie $s' \in S$ est *strictement dominée* par $s \in S$ si pour toute stratégie $t \in T$, on a $g(s', t) < g(s, t)$. Une stratégie $s' \in S$ est *faiblement dominée* par $s \in S$ si pour toute stratégie $t \in T$ on a $g(s', t) \leq g(s, t)$ et il existe une stratégie $t' \in T$ telle que $g(s', t') < g(s, t')$.

Proposition 5. *Soient* $\Gamma = (S, T, g)$ *un jeu à deux joueurs à somme nulle,* $s' \in S$ *une stratégie faiblement dominée et* $\Gamma' = (S \setminus \{s'\}, T, g)$*. Alors* $\underline{v}(\Gamma) = \underline{v}(\Gamma')$ *et* $\overline{v}(\Gamma) = \overline{v}(\Gamma')$*. De plus, si* s' *est strictement dominée et que le maxmin est atteint, alors* s' *n'est pas une stratégie optimale dans* Γ*.*

Ce résultat ressemble aux résultats du cours précédent à propos des équilibres de Nash et de l'élimination itérée de stratégies dominées. En effet, il affirme que, lorsque l'on cherche la valeur d'un jeu, on peut ignorer les stratégies faiblement dominées, et lorsque l'on cherche les stratégies optimales, on peut ignorer les stratégies strictement dominées.

Exemple 4. Par éliminations itérées des stratégies strictement dominées, donnons la valeur et les stratégies optimales du jeu suivant :

	t_1	t_2	t_3	t_4
s_1	0	3	3	9
s_2	6	1	4	0
s_3	7	5	6	8
s_4	8	4	1	9

On obtient alors la succession :

	t_1	t_2	t_3	t_4
s_1	0	3	3	9
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	7	5	6	8
s_4	8	4	1	9

	t_1	t_2	t_3	~~t_4~~
s_1	0	3	3	~~9~~
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	7	5	6	~~8~~
s_4	8	4	1	~~9~~

	t_1	t_2	t_3	~~t_4~~
~~s_1~~	~~0~~	~~3~~	~~3~~	~~9~~
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	7	5	6	~~8~~
s_4	8	4	1	~~9~~

	~~t_1~~	t_2	t_3	~~t_4~~
~~s_1~~	~~0~~	~~3~~	~~3~~	~~9~~
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	~~7~~	5	6	~~8~~
s_4	~~8~~	4	1	~~9~~

	~~t_1~~	t_2	t_3	~~t_4~~
~~s_1~~	~~0~~	~~3~~	~~3~~	~~9~~
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	~~7~~	5	6	~~8~~
~~s_4~~	~~8~~	~~4~~	~~1~~	~~9~~

	~~t_1~~	t_2	~~t_3~~	~~t_4~~
~~s_1~~	~~0~~	~~3~~	~~3~~	~~9~~
~~s_2~~	~~6~~	~~1~~	~~4~~	~~0~~
s_3	~~7~~	5	~~6~~	~~8~~
~~s_4~~	~~8~~	~~4~~	~~1~~	~~9~~

$\Diamond$

3. Valeurs en stratégies mixtes

Dans la suite de ce cours, on considère que les deux ensembles de stratégies S et T sont finis. On rappelle que les ensembles de stratégies mixtes sont donnés par

$$\Delta(S) = \left\{ x \in \mathbb{R}_+^{|S|} \,\middle|\, \sum_{i \in S} x_i = 1 \right\}, \qquad \text{et} \qquad \Delta(T) = \left\{ y \in \mathbb{R}_+^{|T|}, \,\middle|\, \sum_{i \in T} y_i = 1 \right\}.$$

La fonction de paiement g est étendue multilinéairement, par

$$\begin{array}{rccl} g: & \Delta(S)\times\Delta(T) & \longrightarrow & \mathbb{R} \\ & (\sigma,\tau) & \longmapsto & g(\sigma,\tau) \coloneqq \mathbb{E}_{\sigma,\tau}[g] = \displaystyle\sum_{s\in S}\sum_{t\in T}\sigma(s)\tau(t)g(s,t). \end{array}$$

Comme les ensembles de stratégies S et T sont finis, on peut représenter la fonction de paiement par une matrice

$$A = \big[g(s,t)\big]_{(s,t)\in S\times T}.$$

Ainsi, pour chaque profil de stratégies mixtes (σ,τ), le gain est donné par

$$g(\sigma,\tau) = \sigma^\top A\tau.$$

Exemple 5. Soient $\sigma = (\sigma_1,\sigma_2)$ une stratégie mixte pour le gardien et $\tau = (\tau_1,\tau_2)$ une stratégie mixte pour le tireur dans le jeu "Tir de penalty". Le gain du gardien est donné par

$$\sigma^\top A\tau = \begin{pmatrix}\sigma_1 & \sigma_2\end{pmatrix}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}\tau_1\\ \tau_2\end{pmatrix} = \begin{pmatrix}\sigma_1 & \sigma_2\end{pmatrix}\begin{pmatrix}\tau_1-\tau_2\\ \tau_2-\tau_1\end{pmatrix} = \sigma_1(\tau_1-\tau_2)+\sigma_2(\tau_2-\tau_1)$$

$\diamondsuit$

Pour la théorie générale des jeux finis sous forme normale, prendre l'extension mixte d'un jeu impliquait l'existence d'au moins un équilibre de Nash. On a un résultat similaire dans la théorie des jeux à deux joueurs à somme nulle.

Théorème 1 (du minmax, von Neumann (1928))**.** *Soit A une matrice réelle de taille $|S|\times|T|$. Il existe $(\sigma^*,\tau^*,v)\in\Delta(S)\times\Delta(T)\times\mathbb{R}$ tel que, pour tout profil (σ,τ), on a*

$$\sigma^\top A\tau^* \leq v \leq (\sigma^*)^\top A\tau.$$

L'extension mixte du jeu matriciel a une valeur et les joueurs ont des stratégies optimales. Le réel v est unique, c'est la valeur *(en stratégies mixtes) de la matrice A:*

$$v = \max_{\sigma\in\Delta(S)}\min_{\tau\in\Delta(T)}\sigma^\top A\tau = \min_{\tau\in\Delta(T)}\max_{\sigma\in\Delta(S)}\sigma^\top A\tau.$$

Ce résultat est généralement considéré comme le point de départ de la théorie des jeux moderne. Il a été généralisé plus de 20 ans plus tard, par Nash.

En stratégies pures, le maxmin $\underline{v}$ donne une borne inférieure quant au paiement attendu par le joueur j_1, et le minmax $\overline{v}$ donne une borne supérieure. Il est alors peu surprenant de voir que la valeur d'un jeu en stratégies mixtes est comprise entre le maxmin et la minmax, c'est-à-dire,

$$\underline{v}(\Gamma) \leq v(\Gamma^\Delta) \leq \overline{v}(\Gamma).$$

On a également le résultat suivant.

Lemme 2. *Le joueur qui joue en second peut jouer en stratégies pures, c'est-à-dire,*

$$\min_{\tau\in\Delta(T)}\max_{\sigma\in\Delta(S)} g(\sigma,\tau) = \min_{\tau\in\Delta(T)}\max_{s\in S} g(s,\tau).$$

Cette propriété est due au fait que la fonction de paiement g est étendu multilinéairement aux stratégies mixtes. De façon similaire, on a le résultat suivant.

Proposition 6. *Si (σ^*, τ^*) est un profil de stratégies optimales, et si $\sigma^*(s) > 0$, alors*

$$g(s, \tau^*) = v.$$

Exemple 6. Considérons le jeu donné par la matrice

	G	D
H	3	-1
M	0	0
B	-2	1

La stratégie M du joueur j_1 est strictement dominée par la stratégie mixte $0.49H + 0.51B$, on l'élimine donc sans que cela change la valeur du jeu ainsi que les stratégies optimales. On a alors la matrice

	G	D
H	3	-1
B	-2	1

Soit (σ, τ) un profil de stratégies optimales. On voit que σ ne peut pas être une stratégie pure. Notons $\sigma = p_1 H + (1 - p_1)B$ avec $p_1 \in (0, 1)$ et $\tau = p_2 G + (1 - p_2)D$. Étudions les meilleures réponses du premier joueur en fonction de p_2. On a

$$g(H, p_2) = 3p_2 - (1 - p_2) = 4p_2 - 1, \quad \text{et} \quad g(B, p_2) = -2p_2 + (1 - p_2) = 1 - 3p_2.$$

Le joueur j_1 est indifférent entre ses deux stratégies pures lorsque $g(H, p_2) = g(B, p_2)$, c'est-à-dire,

$$4p_2 - 1 = 1 - 3p_2 \qquad \Longleftrightarrow \qquad p_2 = \frac{2}{7}.$$

On répète le même processus pour le second joueur. On a alors

$$g(p_1, G) = -3p_1 + 2(1 - p_1) = 2 - 5p_1, \quad \text{et} \quad g(p_1, D) = p_1 - (1 - p_1) = 2p_1 - 1.$$

Le joueur j_2 est indifférent entre ses deux stratégies pures lorsque $g(p_1, G) = g(p_1, D)$, c'est-à-dire,

$$2 - 5p_1 = 2p_1 - 1 \qquad \Longleftrightarrow \qquad p_1 = \frac{3}{7}.$$

Ainsi, les stratégies optimales sont

$$\sigma = \frac{3}{7}H + \frac{4}{7}B, \qquad \text{et} \qquad \tau = \frac{2}{7}G + \frac{5}{7}D,$$

et la valeur est

$$g(\sigma, \tau) = \begin{pmatrix} 3/7 & 4/7 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2/7 \\ 5/7 \end{pmatrix} = \begin{pmatrix} 3/7 & 4/7 \end{pmatrix} \begin{pmatrix} 1/7 \\ 1/7 \end{pmatrix} = \frac{7}{49} = \frac{1}{7}.$$

$\Diamond$

4. Lien avec la programmation linéaire

Considérons le maxmin d'un jeu matriciel en stratégies mixtes :

$$\max_{\sigma\in\Delta(S)} \min_{\tau\in\Delta(T)} \sigma^\top A\tau.$$

Grâce au Lemme 2, on peut réécrire ce problème comme

$$\max_{\sigma\in\Delta(S)} \min_{t\in T} \left[\sigma^\top A\right]_t,$$

où le fait de prendre la t-ième coordonnées du vecteur $\sigma^\top A$ correspond à prendre son produit scalaire avec le vecteur indicatrice de $t \in T$. On peut réécrire ce problème comme un programme linéaire :

$$\max_{\sigma\in\Delta(S)} z, \qquad \text{tel que} \qquad z \leq \left[\sigma^\top A\right]_t, \ \forall t \in T,$$

que l'on réécrit à nouveau comme

$$\max_{\sigma\in\mathbb{R}^{|S|}} z$$
$$\text{s.c.} \begin{cases} z\mathbf{1} - \sigma^\top A \leq \mathbf{0}, \\ \sigma^\top \mathbf{1} = 1, \\ \sigma \geq \mathbf{0}. \end{cases}$$

Le problème dual s'écrit

$$\min_{\tau\in\mathbb{R}^{|T|}} w$$
$$\text{s.c.} \begin{cases} \mathbf{1}w - A\tau \leq \mathbf{0}, \\ \mathbf{1}^\top \tau = 1, \\ \tau \geq \mathbf{0}, \end{cases}$$

qui lui-même se réécrit comme un problème de minmax

$$\min_{\tau\in\Delta(T)} \max_{\sigma\in\Delta(S)} \sigma^\top A\tau.$$

La théorie des jeux à deux joueurs à somme nulle et la théorie de la programmation linéaire sont en réalité encore plus similaires qu'il n'y parait sur cet exemple, car tout programme linéaire correspond également à un jeu à deux joueurs à somme nulle.