ISAE-SUPAERO
Jérôme Renault

12 Février 2021

Introduction à la Théorie des Jeux

Introduction

Une **interaction stratégique** est une situation où :

a) il y a plusieurs entités (au sens large : personnes physiques, personnes morales, animaux, automates, logiciels,...) appelées joueurs,

b) chaque joueur peut ou doit faire quelque chose (choix d'actions ou stratégies),

c) l'utilité (le but, bonheur, argent, etc...) de chacun ne dépend pas seulement de ce qu'il fait, mais aussi de ce que font les autres joueurs.

La théorie des jeux étudie les interactions stratégiques, qu'on appelle des jeux.

Examples :

1. Enchères pour un bien indivisible (Google Ads)
2. Concurrence entre des entreprises sur un même marché qui choisissent leur prix de vente
3. Elections (par exemple élection présidentielle en France)
4. Jeux ludiques (échecs, poker, ...)
5. etc...

Situations extrêmement fréquentes. Les jeux sont omniprésents dans la société, ce qui rend difficile d'énoncer des propriétés à la fois très générales et utiles dans un cas précis. Etudiés en maths, économie, informatique, biologie, sociologie,...

Quelques questions :

1. Comment modéliser une interaction stratégique ?

2. Peut-on déterminer ce que des joueurs "rationnels" devraient faire ? Que veut dire bien jouer ?

3. Dans quels jeux jouer rationnellement aboutit-il à quelque chose de socialement optimal ? Comment construire des mécanismes qui amènent à jouer de tels jeux, et pas les autres ?

Et aussi : quid des jeux hybrides entre des agents stratégiques et des algorithmes connus ? Comment utiliser des modèles de jeux en IA ?...

Il y a **plusieurs types de théorie de jeux**. On va aborder ici les jeux stratégiques (ou non coopératifs, les règles du jeu sont explicites), mais il y a aussi :

• les jeux coalitionnels (on dit aussi coopératifs) et le choix social (comment partager une note de taxi, comment voter, comment former une coalition majoritaire dans un parlement),

• les jeux évolutionnaires en théorie de l'évolution (comment le patrimoine génétique évolue en fonction de la confrontation avec d'autres espèces),

• la théorie des jeux algorithmique ("prix de l'anarchie") • ....

Plan :

1. Jeux à somme nulle

1.a. Jeux matriciels

1.b. Jeux à information parfaite

1.c. Un poker simplifié

1.d. Fictitious play

2. Apprentissage sans regret et calibration

2.a. Absence de regret

2.b. Calibration

3. Jeux à somme non nulle

3.a. Définitions

3.b. Equilibres en stratégies dominantes

3.c. Equilibres de Nash

3.d. Equilibre corrélé (exemple)

Annexes

1. Jeux à somme nulle

Un jeu à somme nulle est un jeu avec 2 joueurs ayant des intérêts opposés.

1.a Jeux matriciels

Définition 1 : Un jeu fini à somme nulle, ou jeu matriciel, est donné par une matrice $A = (a_{i,j})$ in $\mathbb{R}^{n \times p}$.

Interprétation : il y a 2 joueurs, appelés joueur 1 et joueur 2 . Simultanément, le joueur 1 choisit une ligne i dans $I = \{1, ..., n\}$ et le joueur 2 choisit une ligne j dans $J = \{1, ..., p\}$. Puis le joueur 2 paie $a_{i,j}$ au joueur 1. Les deux joueurs connaissent la matrice A et les règles du jeu.

Examples : Que doit jouer le joueur 1 ?

$$A_1 = \begin{pmatrix} 7 & -2 \\ -1 & -1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \text{(MatchingPennies)},$$

$$A_3 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

On suppose que chaque joueur peut choisir son action aléatoirement, et veut maximiser son paiement *espéré*. Une probabilité sur I (resp. J) est appelée une stratégie mixte du joueur 1 (resp. joueur 2). **L'ensemble des stratégies mixtes du joueur 1 est noté :**

$$\Delta(I) = \{x = (x_1, ..., x_n) \in \mathbb{R}^n_+, \sum_{i=1}^{n} x_i = 1\},$$

et l'ensemble des stratégies mixtes du joueur 2 est $\Delta(J) = \{y = (y_1, ..., y_p) \in \mathbb{R}^p_+, \sum_{j=1}^{p} y_j = 1\}$.

Une ligne i (colonne j) est maintenant appelée une action pure du joueur 1 (joueur 2), et assimilée à la mesure de Dirac sur i (sur j).

On note aussi $g : I \times J \to I\!R$ la fonction de paiement du jeu, vérifiant $g(i,j) = a_{i,j}$ pour tout (i,j).

Le paiement espéré du joueur 1 si le couple de stratégies mixtes (x,y) est joué, est par définition :

$$g(x,y) := I\!E_{x\otimes y}(g) = \sum_{i=1}^{n}\sum_{j=1}^{p} x_i y_j g(i,j) = xAy,$$

où dans la dernière expression, x est vu comme un vecteur ligne et y comme un vecteur colonne. La fonction g est donc étendue à $\Delta(I) \times \Delta(J)$.

Example de calcul de paiement espéré :

Remarque : Pour tout y dans $\Delta(J)$, $\max_{x\in\Delta(I)} g(x,y) = \max_{i\in I} g(i,y)$. Et pour tout x de $\Delta(I)$, $\min_{y\in\Delta(J)} g(x,y) = \min_{j\in J} g(x,j)$. *Les joueurs ont toujours des meilleures réponses en stratégies pures.*

Théorème 1 : du Minmax de Von Neumann

Il existe un unique réél v pouvant être garanti par chacun des joueurs, i.e. tel que :

$$(\exists x \in \Delta(I)\, \forall y \in \Delta(J), g(x,y) \geq v), \quad \text{et} \quad (\exists y \in \Delta(J)\, \forall x \in \Delta(I), g(x,y) \leq v).$$

$$\text{Et on a } v = \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} xAy = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} xAy.$$

On note $v = \texttt{val}(A)$ et on dit que v est la *valeur* de la matrice A. C'est la somme équitable que le joueur 1 devrait avoir à payer au joueur 2 pour avoir le droit de jouer le jeu matriciel A.

Une stratégie x qui atteint le max dans $\max_{x \in \Delta(I)} \min_{y \in \Delta(J)} xAy$ est appelée stratégie optimale du joueur 1, et une stratégie y qui atteint le min dans $\min_{y \in \Delta(J)} \max_{x \in \Delta(I)} xAy$ est appelée stratégie optimale du joueur 2.

Remarques :

1) $x \in \Delta(I)$ est optimale pour le joueur 1 si et seulement si : $\forall j \in J$, $g(x,j) \geq v$. Et $y \in \Delta(J)$ est optimale pour le joueur 2 ssi : $\forall i \in I$, $g(i,y) \leq v$.

2) Si $\min_{j \in J} \max_{i \in I} a_{i,j} = \max_{i \in I} \min_{j \in J} a_{i,j}$, alors c'est la valeur du jeu et les joueurs ont des stratégies optimales pures.

3) En général il n'existe pas forcément de stratégie optimale pure.

4) Si on ajoute à A une ligne inférieure (coordonnée par coordonnée) à une ligne existante, ou une colonne supérieure (coordonnée par coordonnée) à une colonne existante, la valeur ne change pas.

Examples :

Preuve du Théorème du Minmax :

Notons $\alpha = \sup_{x\in\Delta(I)} \inf_{y\in\Delta(J)} g(x,y)$ et $\beta = \inf_{y\in\Delta(J)} \sup_{x\in\Delta(I)} g(x,y)$.
Par continuité et compacité,
$\alpha = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} g(x,y)$ et $\beta = \min_{y\in\Delta(J)} \max_{x\in\Delta(I)} g(x,y)$.

Il est facile de voir que $\alpha \leq \beta$:

Il reste à montrer $\alpha \geq \beta$.

Pour x dans $\Delta(I)$, on définit $l(x) = (g(x,1), ..., g(x,p)) \in I\!R^p$. l est affine, donc l'ensemble $E := \{l(x), x \in \Delta(I)\}$ est compact convexe dans $I\!R^p$. On introduit :

$$C = E - I\!R^p_+ = \{l(x) - u,\ x \in \Delta(I), u \in I\!R^p_+\},$$

$$\text{et} \ \ \bar{\beta} = (\beta, \beta, ..., \beta) \in I\!R^p.$$

E est convexe compact et $I\!R^p_+$ est convexe fermé, donc C est convexe et fermé dans $I\!R^p$.

1) Supposons $\overline{\beta} \in C$. Alors il existe x in $\Delta(I)$ tel que $l(x) \geq \overline{\beta}$ (coordonnée par coordonnée). Donc $\alpha \geq \beta$, OK ☺.

2) Sinon $\overline{\beta} \notin C$. Dans $I\!R^p$, on peut séparer strictement $\overline{\beta}$ du convexe fermé C. Donc il existe $y = (y_1, ..., y_p)$ dans $I\!R^p$, $\varepsilon > 0$ t.q.

$$< \overline{\beta}, y > \geq \sup_{z \in C} < z, y > +\varepsilon.$$

On a :

$$< \overline{\beta}, y > \geq \sup_{z \in C} < z, y > +\varepsilon.$$

Alors $y \neq 0$, et

$$\beta \sum_{j=1}^{p} y_j \geq \sum_{j=1}^{p} y_j z_j + \varepsilon, \forall z \in C.$$

La forme de C implique $y_j \geq 0$ pour tout j. Et en particulier,

$$\forall x \in \Delta(I), \;\; \beta \sum_{j=1}^{p} y_j \geq \sum_{j=1}^{p} y_j g(x, j) + \varepsilon.$$

Soit $S = \sum_{j \in J} y_j > 0$, et $y^* = \frac{y}{S}$. Alors $y^* \in \Delta(J)$ peut être vue comme une stratégie mixte du joueur 2 dans le jeu matriciel ! On divise par S et on obtient :

$$\forall x \in \Delta(I), \; g(x, y^*) < \beta - \frac{\varepsilon}{S}.$$

Contradiction avec la définition de β. Le cas 2) est impossible. ☺.

Proposition 1 : Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in I\!R^{2\text{x}2}$ un jeu matriciel. On a :

$$\mathtt{val}(A) \in \{a, b, c, d, \frac{ad - bc}{a + d - b - c}\}.$$

Remarque : Prise de décision dans le pire cas.

Les jeux à somme nulle sont souvent utilisés pour modéliser des situations à un seul joueur voulant maximiser son paiement dans le pire cas (optimisation robuste).

Example : Soit un agent pouvant investir aujourd'hui dans 3 actifs i_1, i_2 et i_3, tous de prix 1 (l'unité est par exemple le millier d'euros). L'état du monde demain est inconnu, il y a 3 possibilités j_1, j_2 et j_3. La matrice A suivante indique le rendement de chaque actif en fonction de l'état du monde :

$$\begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} & \begin{pmatrix} 0 & 2 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \end{array}$$

Par exemple si l'état demain est j_2, alors l'actif i_2 rapportera 0. L'actif i_3 est sans risque, il rapporte 1 quel que soit l'état du monde demain. L'agent veut investir 5 (milliers d'euros) aujourd'hui et peut repartir son investissement sur plusieurs actifs. Il veut maximiser sa richesse demain dans le pire des cas, quel est son investissement optimal aujourd'hui ?

$$\begin{array}{cc} & \begin{array}{ccc} j_1 & j_2 & j_3 \end{array} \\ \begin{array}{c} i_1 \\ i_2 \\ i_3 \end{array} & \begin{pmatrix} 0 & 2 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \end{array}$$

Réponse : résoudre $\max_{x \in \Delta(I)} \min_j \sum_i x_i a_{i,j}$.

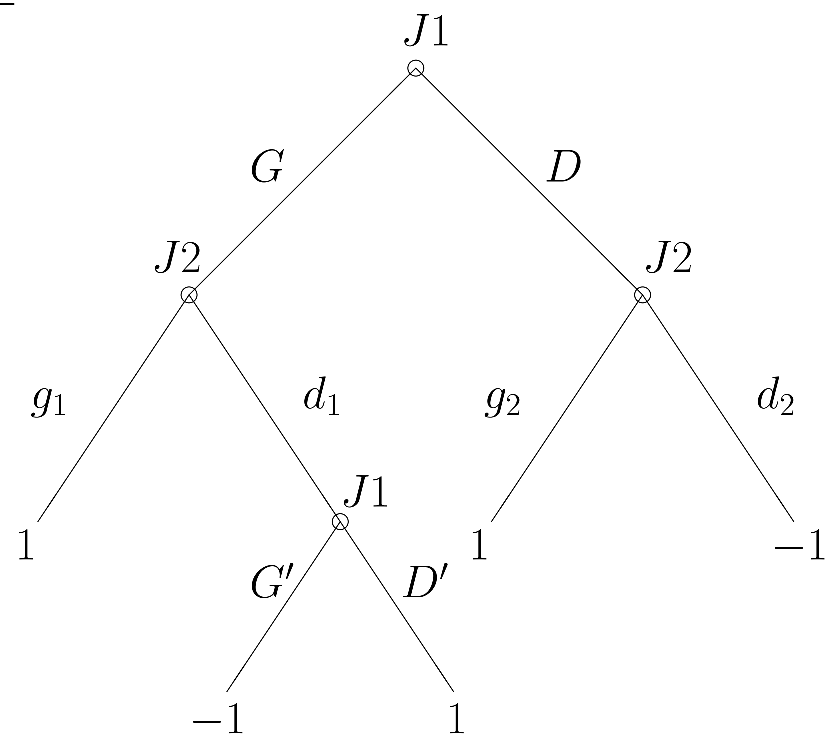
On trouve $\texttt{val}(A) = 6/5$, investir $5\frac{3}{5} = 3$ sur i_1 et $5\frac{2}{5} = 2$ sur i_2, 0 sur i_3.

Et les échecs ?

1.b Jeux à information parfaite : quand un joueur joue, il sait exactement ce qui s'est passé avant et il n'y a pas de choix simultanés.

Ils se représentent par des arbres. Par convention, les feuilles contiennent les paiements du J1.

Exemple :



Préférez-vous être le joueur 1 ou le joueur 2 ? Que doit jouer le joueur 1 au premier coup ?

Paiements généraux : le paiement d'un noeud terminal est un nombre réel (somme que doit payer le J2 au joueur J1).

Une stratégie s_i d'un joueur $i = 1, 2$ spécifie, à chaque noeud de ce joueur, un des successeurs immédiats de ce noeud. Un couple de stratégies induit un paiement (en suivant la partie induite sur l'arbre), ce qui définit un jeu matriciel !

Théorème 2 : (Zermelo) Dans les jeux à somme nulle à information parfaite (avec un arbre fini), il existe un unique réel v satisfaisant :

(il existe s_1^* t.q. $\forall s_2, g(s_1^*, s_2) \geq v$, et il existe s_2^* t.q. $\forall s_1, g(s_1, s_2^*) \leq v$)

Ce nombre s'appelle la *valeur* du jeu. C'est la somme équitable que le joueur 1 devrait payer pour avoir le droit de jouer. s_1^* et s_2^* sont des stratégies *optimales*.

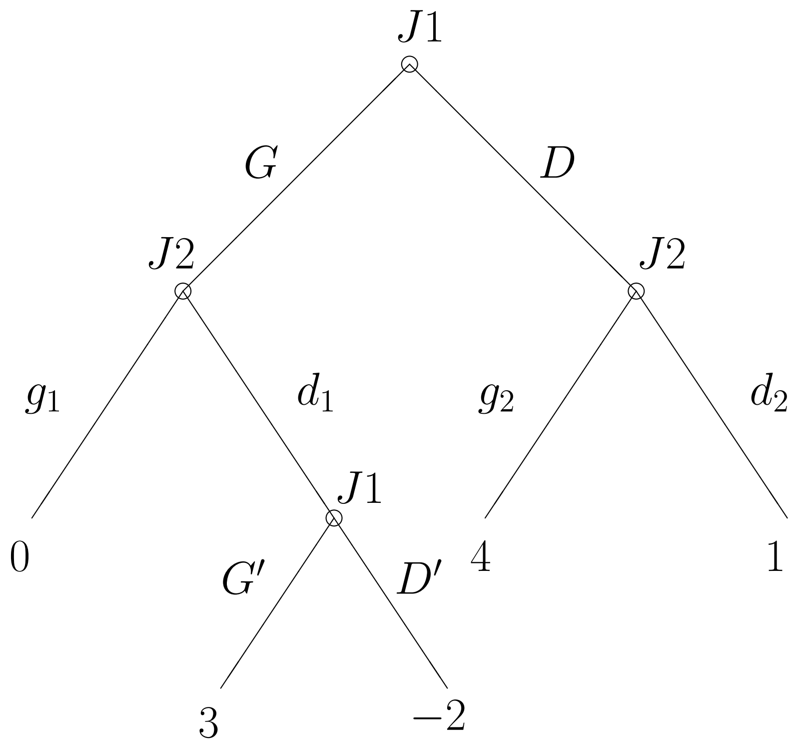
Remarque : mêmes notions que pour les jeux matriciels. Plus simple ici, car les stratégies pures suffisent.

Corollaire (Zermelo) Si en plus le jeu est du type gagnant-perdant (tous les paiements sont 1 ou -1, comme au Go), alors :
soit le joueur 1 a une stratégie gagnante quoi que fasse le joueur 2,
soit le joueur 2 a une stratégie gagnante quoi que fasse le joueur 1.

Et aux échecs où il peut aussi y avoir des matchs nuls, de 3 choses l'une :
soit les blancs ont une stratégie gagnante,
soit les noirs ont une stratégie gagnante,
soit chacun des joueurs a une stratégie qui lui assure au moins le match nul.

(personne ne sait laquelle est vraie)

Exemple : Quelle est la valeur de ce jeu ?



1.c Un poker simplifié : le bluff

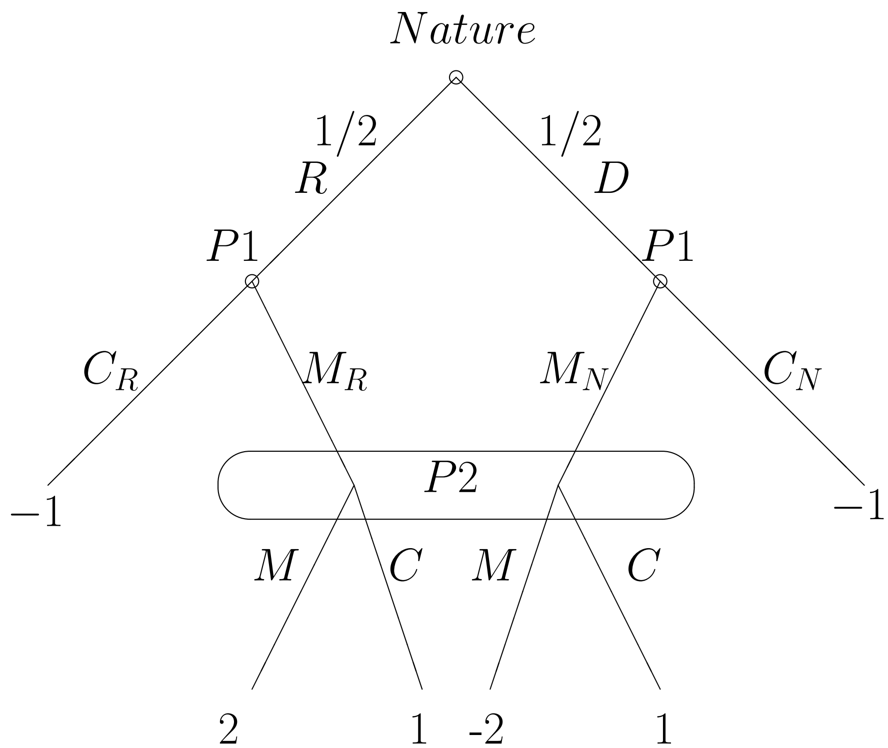
2 joueurs jouent au jeu suivant. Ils commencent par miser chacun 1 euro, c'est-à-dire qu'ils mettent chacun 1 euro au centre de la table.

• Puis un jeu de 52 cartes est battu uniformément, une des cartes est tirée et observée uniquement par le joueur 1

• Le joueur 1 peut alors quitter le jeu (il a alors perdu l'euro qu'il a misé), ou bien rester dans le jeu en rajoutant un euro supplémentaire au centre de la table.

• Si le joueur 1 n'a pas quitté le jeu, le joueur 2 a le choix entre se coucher (c'est-à-dire quitter le jeu, le jeu est alors fini et au final le joueur 1 aura gagné un euro) ou bien rester dans le jeu en rajoutant également 1 euro supplémentaire au centre de la table. Dans ce dernier cas, la carte sélectionnée est montrée aux 2 joueurs : si elle est rouge, le joueur 1 gagne et empoche l'argent qui est au milieu de la table (il aura alors gagné 2 euros), si la carte est noire c'est le joueur 2 qui gagne et empoche l'argent situé au centre de la table.

En tant que joueur 1, combien êtes-vous prêt à payer pour participer ? Avec quelle fréquence le joueur 1 doit-il bluffer ? Avec quel fréquence le joueur 2 doit-il rester dans le jeu et aller voir la carte du joueur 1 ?



$$\begin{array}{cc} & \begin{array}{cc} M & C \end{array} \\ \begin{array}{c} C_R, C_N \\ C_R, M_N \\ M_R, C_N \\ M_R, M_N \end{array} & \begin{pmatrix} -1 & -1 \\ -3/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

(remarque : paiements espérés)

Valeur =1/3, J1 doit miser si R et bluffer 1/3 du temps si N, J2 doit miser 2/3 du temps.

Tous les jeux sous forme d'arbre peuvent se ramener à un jeu matriciel ! et vice-versa.

1.d Fictitious play

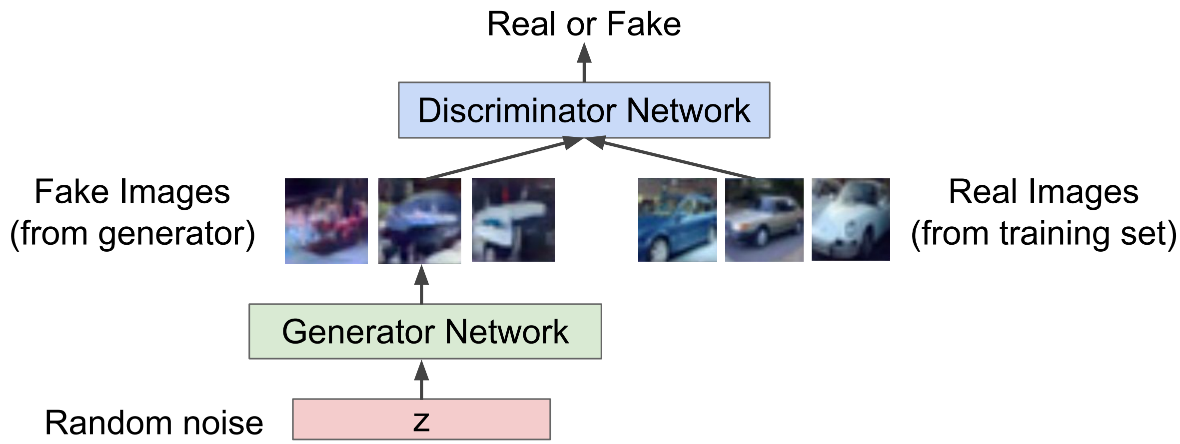
On répète le même jeu matriciel A à chaque date $n = 1, ...,$. Supposons qu'à chaque date $n \geq 2$, le joueur 1 joue i_n qui est une meilleure réponse contre la moyenne empirique $y_{n-1} := \frac{1}{n-1} \sum_{t=1}^{n-1} j_t$ des actions passées du joueur 2, et le joueur 2 joue j_n qui est une meilleure réponse contre la moyenne empirique $x_{n-1} := \frac{1}{n-1} \sum_{t=1}^{n-1} i_t$ des actions passées du joueur 1.

Théorème 3 (Robinson) Fictitious Play :

Quand n tend vers $+\infty$, la distance de x_n à l'ensemble des stratégies mixtes optimales du joueur 1 converge vers 0, et la distance de y_n à l'ensemble des stratégies optimales du joueur 2 converge vers 0. Et le paiement $\frac{1}{n} \sum_{t=1}^{n} A_{i_t, j_t}$ converge vers la valeur $\texttt{val}(A)$ de la matrice A.

Remarque : GANs (Generative Adversarial Networks)

Inventés en 2014 par I. Goodfellow, “the coolest idea in the last 20 years in machine learning” selon Yann Le Cun 2016. Comment simuler une distribution inconnue à l'aide d'un jeu à somme nulle entre réseaux de neurones (un faussaire-générateur contre un policier-discriminateur) ?



2. Apprentissage sans regret et calibration

2.a. Absence de regret

On se donne comme avant une matrice de gains $A = (a_{i,j})$ in $I\!R^{I \times J}$. Un joueur doit choisir, à chaque étape $n = 1, 2, ...$, une action i_n dans I. L'environnement (nature, adversaire, autres joueurs poursuivant leurs buts propres) induit un processus stochastique $(j_n)_{n \geq 1}$ à valeurs dans J, et à chaque étape n le joueur gagne a_{i_n,j_n}.

Exemple : Le joueur est un météorologue qui doit annoncer chaque soir à minuit s'il pleuvra demain à midi sur le campus. $I = J = \{0, 1\}$, $a_{i,j} = 1$ si $i = j$, et $a_{i,j} = 0$ si $i \neq j$. Etape= jour, l'action 0 signifie pas de pluie aujourd'hui, et l'action 1 signifie pluie aujourd'hui. $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

Problème : le joueur ne sachant rien a priori sur la loi de $(j_n)_{n \geq 1}$, comment doit-il jouer ?

Remarque : si l'environnement était un adversaire qui voulait minimiser le paiement d'étape a_{i_n,j_n} pour tout n, il suffirait de jouer à chaque étape une stratégie optimale du joueur 1 pour la matrice A. Ici, on veut *apprendre de façon automatique* (Machine Learning) comment bien jouer en fonction de la suite aléatoire $(j_n)_{n\geq 1}$ dont on ne sait rien a priori.

Que veut dire bien jouer ici ?

Si par exemple on devine correctement 90% des fois mais que $j_n = 1$ pour tout n, est-ce satisfaisant ?

Définitions 2 :

Une stratégie (ou algorithme) du joueur est un élément $\sigma = (\sigma_n)_{n\geq 1}$, avec pour tout n, $\sigma_n : (I \times J)^{n-1} \to \Delta(I)$.

Une stratégie de la nature (environnement) est un élément $\tau = (\tau_n)_{n\geq 1}$, avec pour tout n, $\tau_n : (I \times J)^{n-1} \to \Delta(J)$.

Un couple de stratégies (σ, τ) induit une probabilité $\mathbf{P}_{\sigma,\tau}$ sur l'ensemble des parties possibles $\{(i_1, j_1), (i_2, j_2), ...,\} = (I \times J)^{\infty}$ (muni de la tribu produit).

Pour tout $n \geq 1$, $g_n = g(i_n, j_n) = a_{i_n,j_n}$ est la variable aléatoire du paiement d'étape n, et $\overline{g}_n = \frac{1}{n}\sum_{t=1}^{n} g_t$ est le paiement moyen du joueur jusqu'à l'étape n.

Supposons qu'à la fin de la date n, $(i_1, j_1, ..., i_n, j_n)$ a été joué. Le paiement moyen du joueur est $\overline{g}_n$, et il peut comparer ce paiement avec ce qu'il aurait obtenu s'il avait joué constamment autre chose, ou plus généralement s'il avait, à chaque fois qu'il a joué une certaine action i, remplacé i par une autre action $l \in I$.

Définition 3 : Le regret interne du joueur de ne pas avoir joué l à la place de i est

$$\overline{R}_n(i, l) = \frac{1}{n} \sum_{t \in \{1,..,n\}, i_t = i} (g(l, j_t) - g(i_t, j_t)).$$

Examples :

Définition 4 : Une stratégie σ du joueur est sans regret interne si pour chaque stratégie τ de la nature,

$$\max_{i \in I, l \in I} \overline{R}_n(i, l) \xrightarrow[n \to \infty]{} 0 \quad \mathbf{P}_{\sigma,\tau} \ p.s.$$

Remarque : notion forte car valable pour toute stratégie τ, et ces stratégies peuvent s'adapter aux actions passées.

Théorème 4 : Il existe une stratégie σ sans regret, vérifiant pour toute stratégie τ de la nature :

$$\max_{i\in I, l\in I} \overline{R}_n(i,l) \xrightarrow[n\to\infty]{} 0 \quad \text{et} \quad \forall n\geq 1, \forall (i,l)\in I^2, \mathbb{E}_{\sigma,\tau}\left(\overline{R}_n(i,l)\right) \leq \frac{4\,\|A\|_\infty}{\sqrt{n}}.$$

Construction de σ sur l'exemple : $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

Jouer arbitrairement à l'étape 1. A la fin de chaque étape $n \geq 1$, calculer les regrets :

$$\overline{R}_{n,1,2} = \frac{1}{n} \sum_{t\in\{1,\ldots,n\}, i_t=1} (\mathbf{1}_{j_t=2} - \mathbf{1}_{j_t=1}), \text{ et } \overline{R}_{n,2,1} = \frac{1}{n} \sum_{t\in\{1,\ldots,n\}, i_t=2} (\mathbf{1}_{j_t=1} - \mathbf{1}_{j_t=2}).$$

Et à l'étape $n+1$, jouer $i_{n+1}=1$ avec une (la) probabilité $x\in[0,1]$ telle que :

$$x\,\overline{R}^{+}_{n,1,2} = (1-x)\,\overline{R}^{+}_{n,2,1}.$$

Pourquoi ça marche ? *Preuve basée sur la* **théorie de l'approchabilité de Blackwell** *pour les jeux à paiements vectoriels.* (pour les plus motivés, voir par exemple les notes de cours Strategic Optimization à l'adresse s
https ://sites.google.com/site/jrenaultsite/lecturenotes).

Principe géométrique qui intervient dans la preuve : Soit $(x_n)_{n\geq 1}$ une suite bornée de $\mathbb{R}^K$, et C un convexe fermé non vide de $\mathbb{R}^K$. On note $\overline{x}_n = \frac{1}{n}\sum_{t=1}^n x_t$ pour tout n.

Si pour tout $n \geq 1, < x_{n+1} - \pi_C(\overline{x}_n), \overline{x}_n - \pi_C(\overline{x}_n) > \leq 0,$ alors $d(\overline{x}_n, C) \xrightarrow[n\to\infty]{} 0$).

Exemple en dimension 1, avec $C = \{0\}$. Soit (x_n) une suite bornée de réels telle que pour tout n, $x_{n+1}\overline{x}_n \leq 0$. Alors $\overline{x}_n \xrightarrow[n\to\infty]{} 0$.

Autre stratégie sans regret : Multiplicative Weight Algorithm.

Pour la matrice $\begin{array}{cc} & \begin{array}{cc} G & D \end{array} \\ \begin{array}{c} H \\ B \end{array} & \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{array}$ et N fixé. On se donne un paramètre $\varepsilon > 0$ petit, et on définit l'algo σ à partir de poids dynamiques $(\alpha_n, \beta_n)_{n \geq 1}$ t.q. $\alpha_1 = \beta_1 = 1$, et pour $n = 1, ..., N$:

Si $j_n = G$ alors $\alpha_{n+1} = \alpha_n$ et $\beta_{n+1} = (1 - \varepsilon)\beta_n$.

Si $j_n = D$ alors $\alpha_{n+1} = (1 - \varepsilon)\alpha_n$ et $\beta_{n+1} = \beta_n$.

Et à chaque date $n = 1, ..., N$, σ joue H avec proba $\frac{\alpha_n}{\alpha_n + \beta_n}$ et B avec proba $\frac{\beta_n}{\alpha_n + \beta_n}$.

Soit une suite j_1,...,j_N déterministe jouée par la nature, on note $OPT = \max\{\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{j_n = G}, \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{j_n = R}\}$. On prouve alors (avec un peu de travail) :

$$\mathbb{E}(\overline{g_N}) \geq OPT - \varepsilon - \frac{\ln 2}{N\varepsilon}.$$

Et pour $\varepsilon = \sqrt{\frac{\ln 2}{N}}$, on a donc $\mathbb{E}(\overline{g_N}) \geq OPT - 2\sqrt{\frac{\ln 2}{N}}$.

2.b Calibration : Prédire la météo sans connaître le temps.

On étudie les précipitations chaque jour $t = 1, ..., T$ à midi à un endroit donné : $j_t = 1$ s'il pleut, $j_t = 0$ sinon.

Un prévisionniste doit décider chaque soir s'il va pleuvoir le lendemain à midi. Il donne des prévisions probabilistes dans une grille A fixée :

$$p_t \in A = \{0, \frac{1}{N}, \frac{2}{N}, ..., \frac{N-1}{N}, 1\}.$$

(N paramètre donné, par exemple $N = 100$).

A la fin de la période T, quand doit-on considérer que le prévisionniste a été bon ?

On introduit, pour chaque p de la grille, l'écart :

$$K(p) = \left| p - \frac{\sum_{s=1}^{T} \mathbf{1}_{p_s=p} j_s}{\sum_{s=1}^{T} \mathbf{1}_{p_s=p}} \right| \in [0, 1].$$

Les prévisions sont considérées bonnes si l'écart moyen $K = \frac{1}{T} \sum_{t=1}^{T} K_T(p_t)$ est petite.

Théorème 5 : Pour $T \geq N^2(N+1)$, il existe une stratégie aléatoire du prévisionniste telle que quelque soit $(j_1, ..., j_T)$ dans $\{0, 1\}^T$, l'espérance de K est au plus $\frac{1}{N}$.

Idée de la preuve :

a) Supposons que la nature choisisse de faire pleuvoir chaque jour de façon indépendamment avec une probabilité p dans A, et que le prévisionniste le sache.

Alors la prévision est facile : en prédisant p chaque jour, l'espérance de K sera $\mathbb{E}\left(\left|p-\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{j_t=1}\right|\right)$. Et

$$\begin{aligned}
\mathbb{E}\left(\left|p-\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{j_t=1}\right|\right) &\leq \sqrt{\mathbb{E}\left(p-\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{j_t=1}\right)^2} \\
&= \sqrt{V(p-\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{j_t=1})} \\
&= \sqrt{\frac{p(1-p)}{T}} \\
&\leq \frac{1}{2\sqrt{T}}
\end{aligned}$$

b) idem mais $p \notin A$. En prédisant le point de la grille le plus proche de p, on fait une erreur d'au plus $1/(2N)$, et donc l'espérance de K sera au plus $\frac{1}{2\sqrt{T}}+\frac{1}{2N} \leq \frac{1}{N}$ pour N grand.

c) oui mais p n'est pas forcément le même chaque jour, et peut dépendre de l'histoire passé. Si on suppose que chaque soir le prévisionniste connaît la probabilité employée par la nature le lendemain, il peut prédire à chaque fois le popint de la grille le plus proche, et pour T assez grand on montre que l'on a toujours $\mathbb{E}(K) \leq \frac{1}{N}$.

d) enfin, le prévisionniste ne connaît pas chaque soir la probabilité employée par la nature le lendemain. Rappelons-nous le théorème du minmax : pour toute matrice A,

$\texttt{val}(A) = \min_{y \in \Delta(J)} \max_{x \in \Delta(I)} xAy = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} xAy.$

On applique cela à la matrice où les lignes sont les stratégies pures de la nature, les colonnes sont les stratégies pures du prévisionnistes, et le paiement du joueur 1 est K.

On a vu que : pour toute stratégie x de la nature, il existe une stratégie y du prévisionniste telle que $\mathbb{E}_{x,y}(K) \leq \frac{1}{N}$.

Donc $\texttt{val}(A) \leq \frac{1}{N}$.

On considère une stratégie optimale y du prévisionniste : pour toute stratégie x de la nature, $\mathbb{E}_{x,y}(K) \leq \frac{1}{N}$. ☺

3. Jeux à somme non nulle

3.a Définition 5 : Un jeu sous forme stratégique (ou forme normale) est donné par :

- un ensemble de joueurs N,

- pour chaque joueur i de N, un ensemble d'actions A_i et une fonction de paiement $g_i : \prod_{j \in N} A_j \to \mathbb{R}$.

On suppose N et les ensembles d'actions non vides, et on note le jeu $G = (N, (A_i)_{i \in N}, (g_i)_{i \in N}$. On note $A = \prod_{j \in N} A_j$, un élément de A s'appelle un profil d'actions.

Interprétation : Simultanément, chaque joueur i de N choisit une action i dans A_i. Puis chaque joueur i reçoit le paiement $g_i(a)$, où $a = (a_j)_{j \in N} \in A$. Les joueurs connaissent G et les règles du jeu.

Exemple : Jeu matriciel

Exemple : $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (10,0) & (0,1) \\ (0,0) & (1,1) \end{pmatrix} \end{array}$ Que doit jouer le joueur 1 ?

Exemple : **Enchères au second prix.**

$N = \{1, ..., n\}$ joueurs, v_1,...,v_n sont des paramètres fixés ("valuations" des joueurs) dans $I\!R_+$. Pour tout i de N, $A_i = I\!R_+$ et pour tout profil d'actions $(p_1, ..., p_n)$ dans A

$$g_i(p_1, ..., p_n) = \begin{cases} 0 & \text{if } \; p_i < \max\{p_j, j \in N\} \\ \frac{v_i - p_i}{|\{j \in N, p_j = p_i\}|} & \text{if } \; p_i = \max\{p_j, j \in N\} \end{cases}$$

Retour au modèle général. Soit i un joueur. On note $A_{-i} = \prod_{j \in N, j \neq i} A_j$ l'ensemble des profils d'actions des autres joueurs que i. Pour a dans A, on note $a = (a_i, a_{-i})$.

Problème du joueur i : il doit choisir a_i de façon à maximiser $g_i(a_i, a_{-i})$ mais il ne connait pas a_i.

3.b Equilibres en stratégies dominantes

Le cas “facile” du point de vue stratégique. Souvent unique quand il existe, mais n’existe pas souvent.

Définition 6 : Pour i dans N, $a_i \in A_i$ est une stratégie dominante du joueur i si :

$$\forall a_{-i} \in A_{-i}, \forall b_i \in A_i, g_i(a_i, a_{-i}) \geq g_i(b_i, a_{-i}).$$

Un équilibre en stratégies dominantes est un profil d’actions $a = (a_i)_{i \in N}$ tel que pour tout i, a_i est une stratégie dominante du joueur i.

Exemple : $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (1,1) & (3,3) \\ (2,0) & (5,2) \end{pmatrix} \end{array}$

Exemple : **le dilemme du prisonnier** $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (1,1) & (-1,2) \\ (2,-1) & (0,0) \end{pmatrix} \end{array}$

Aurore a un abricot, Barnabé une banane. Ils préfèrent le fruit de l'autre, sont gourmands et égoïstes, ne se reverront jamais. Doivent décider simultanément d'offrir leur fruit (C_1 ou C_2) ou le garder (D_1 ou D_2).

Comment jouer *rationnellement* ? Garder son fruit ! (Jouer D_1, D_2)

Un drame de l'humanité : Rationalité stratégique et optimum social sont deux notions distinctes, "on peut bien jouer et arriver à un mauvais résultat."

Exercice : Au enchères au second prix, il est dominant de jouer sa propre valuation.

Les équilibres en stratégies dominantes n'existent pas "souvent" : $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (10,0) & (0,1) \\ (0,0) & (1,1) \end{pmatrix} \end{array}$

3.c Equilibres de Nash

Définition 7 : Un équilibre de Nash du jeu G est un profil de stratégies $a = (a_i)_{i \in N}$ tel que :

$$\forall i \in N, \forall b_i \in A_i,\ g_i(b_i, a_{-i}) \leq g_i(a).$$

Si les autres joueurs jouent a_i, le joueur i a intérêt à jouer a_i. Un équilibre de Nash est un profil d'actions tel que chaque joueur joue une meilleure réponse contre la stratégie des autres joueurs.

Exemple : $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (10,0) & (0,1) \\ (0,0) & (1,1) \end{pmatrix} \end{array}$

Exemple : la **"Bataille des Sexes"** $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (2,1) & (0,0) \\ (0,0) & (1,2) \end{pmatrix} \end{array}$

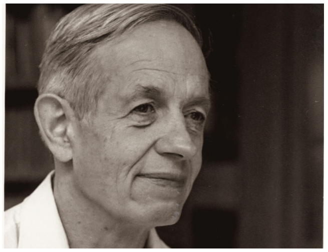
Il peut y avoir plusieurs équilibres de Nash, avec des paiements différents.

Considérons un jeu $G = (N, (A_i)_{i\in N}, (g_i))_{i\in N}$ fini : $N = \{1, ..., n\}$ et A_i est fini pour chaque i. Comme pour les jeux matriciels, on autorise les joueurs à jouer des stratégies mixtes et on considère qu'ils veulent maximiser leur paiement espéré :

$$g_i(x_1, ..., x_n) = \mathbb{E}_{x_1\otimes...\otimes x_n}(g_i) = \sum_{(a_1,...,a_n)\in A} x_1(a_1)...x_n(a_n)\ g_i(a_1, ..., a_n).$$

Théorème 6 : Existence d'équilibres de Nash (J. Nash) Dans un jeu fini, il existe toujours un équilibre de Nash en stratégies mixtes.

Formellement, il existe $x = (x_i)_{i\in N} \in \prod_{i\in N} \Delta(A_i)$ tel que pour tout joueur i et tout action b_i dans A_i : $g_i(x) \geq g_i(b_i, x_{-i})$.



Bataille des sexes : 3 équilibres de Nash en stratégies mixtes

On montre que si G est un jeu matriciel, les équilibres de Nash sont les couples (x, y), où x est une stratégie optimale du joueur 1 et y une stratégie optimale du joueur 2. Et il n'y a qu'un paiement d'équilibre de Nash $(v, -v)$, où v est la valeur du jeu.

Preuve du théorème de Nash : utilise le Théorème de point fixe de Brouwer.

Exemple : Jeu de congestion

Un exemple où la “valeur de l’information” est négative

Jeu G :

- Un arbitre tire à Pile ou Face et montre le résultat au joueur 1 seulement.
- Puis le joueur 1 doit annoncer publiquement P_1 ou F_1.
- Puis le joueur 2 annonce P_2 ou F_2.

Paiements du jeu : les joueurs essaient d’annoncer la pièce. Un joueur qui annonce la mauvaise face a un paiement de 0, un joueur qui annonce la bonne face reçoit 1 si l’autre a aussi bien annoncé, et 100 sinon.

Cette description est publique.

Il y a un seul équilibre de Nash, où les 2 joueurs annoncent la bonne face. Paiement $(1,1)$.

Jeu G' : exactement comme G sauf que personne ne voit le résultat de la pièce lancée par l’arbitre. La description du jeu est publique.

En équilibre de Nash, le joueur 2 va annoncer la face non annoncée par le joueur 1. Paiement d’équilibre unique $\frac{1}{2}(100,0)+\frac{1}{2}(0,100)=(50,50)$.

Conclusion : les 2 joueurs préfèrent jouer G' que G. Si l’arbitre commence à demander publiquement au joueur 1 “voulez-vous voir la pièce ?”, il doit répondre négativement.

3.d Equilibre corrélé (exemple)

Jeu du *chicken* (poule mouillée) : $\begin{array}{ccc} & L & R \\ T & \multicolumn{2}{c}{\multirow{2}{*}{$\begin{pmatrix} (6,6) & (2,7) \\ (7,2) & (0,0) \end{pmatrix}$}} \\ B & & \end{array}$

$$\begin{array}{c|c} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (6,6) & (2,7) \\ (7,2) & (0,0) \end{pmatrix} \end{array}$$

3 paiements d'équilibre de Nash : $(7,2)$, $(2,7)$, and $(14/3, 14/3)$.

Supposons qu'un médiateur choisisse une entrée de la matrice avec les probabilités : $\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & 0 \end{pmatrix} \end{array}$, puis annonce de façon privée la ligne choisie au joueur 1 et la colonne choisie au joueur 2.

Dans le jeu étendu, jouer ce qu'on vous a annoncé est un équilibre de Nash, de paiement $\frac{1}{3}(6,6) + \frac{1}{3}(7,2) + \frac{1}{3}(2,7) = (5,5)$.

Un équilibre de Nash d'un jeu étendu où un médiateur a envoyé des messages au joueur avant de jouer s'appelle un équilibre corrélé du jeu de départ.

On montre que l'ensemble des paiements d'équilibres corrélés d'un jeu fini est un polytôpe qui contient les paiements d'équilibres de Nash.

Annexe 1 : Retour sur le dilemme du prisonnier On peut tous “bien” jouer et arriver à un résultat très mauvais.

$$\begin{array}{cc} & \begin{array}{cc} C_2 & D_2 \end{array} \\ \begin{array}{c} C_1 \\ D_1 \end{array} & \begin{pmatrix} (1,1) & (-1,2) \\ (2,-1) & (0,0) \end{pmatrix} \end{array}$$

- Sport : 2 cyclistes dans une échappée. C : rouler, D : en garder sous la pédale
- Ecologie : entreprises ou pays : C : Ne pas polluer, D : polluer
- Economie : entreprises sur le même marché.

  C : publicité modérée, D : beaucoup de publicité.

  C : maintenir les prix, D : baisser les prix.

  C : pêcher modérément, D : pêcher intensivement.
- Politique Internationale : 2 pays avec des désaccords

  C : financer la culture, D : course aux armements

  C : coopérer, dialoguer, D : être agressif, attaquer (pour être réélu)
- Politique : C : taxer certaines multinationales (éventuellement), D : taxer moins que les autres pays pour les attirer.
- Jeu TV *Golden Balls* : *Split or Steal ?* $\begin{array}{cc} & \begin{array}{cc} C_2 & D_2 \end{array} \\ \begin{array}{c} C_1 \\ D_1 \end{array} & \begin{pmatrix} (50,50) & (0,100) \\ (100,0) & (0,0) \end{pmatrix} \end{array}$
- ...

Comment y échapper ?

1) *Changer les règles du jeu.*

S'échanger les fruits naturellement. Interdire la vente à perte. Introduire des quotas de pêche. Politique Internationale : avoir recours à un médiateur extérieur (ONU, Cour internationale). Signer un traité, faire des accords *vérifiables* sur le climat, la pollution.

C'est possible en sciences sociales. → Conception de mécanismes d'incitation (*mechanism design*).

2) *Se placer dans un contexte dynamique, avec une mémorisation des comportements.*

Cyclisme : il y a des courses toutes les semaines. Celui qui ne roule pas aura une mauvaise réputation, et personne ne voudra faire une échappée avec lui.

Idée : Personne n'a interêt de commencer à jouer agressif (D) si cela implique une escalade d'agressivité par la suite. → Si on répète le jeu indéfiniment, on peut soutenir à l'équilibre la répétition de (C_1, C_2) (bon paiements), chaque joueur menaçant l'autre de représailles en cas déviation.

Les jeux à somme nulle permettent de quantifier les niveaux de menaces pour chaque joueur. On aboutit au "Folk Théorème" : les paiements d'équilibres du jeu répété sont les paiements réalisables où chaque joueur a au moins son niveau de menace.

Annexe 2 : Des roulettes et le paradoxe de Condorcet

Il y a 2 joueurs et 3 roulettes indépendantes A,B, et C. La roulette A sort 1,6 ou 8 avec probas égales, la roulette B sort 3,5 ou 7 avec proba égales et la roulette C sort 2,4 ou 9 avec probas égales.

Jeu à 2 joueurs : d'abord, le joueur 1 choisit une des roulettes, puis le joueur 2 choisit une roulette restante. Puis chacun fait rouler sa roulette, et le joueur qui sort le plus grand numéro gagne un euro de la part de son opposant.

Si vous êtes le joueur 1, combien êtes-vous prêt à payer pour pouvoir jouer ce jeu une fois ? Rapport avec le paradoxe de Condorcet ?

Annexe 3 : Chomp ! ou le mange-savon. *(David Gale)*

On fixe 2 entiers $n \geq 1$ et $m \geq 1$, on pose $P(n,m) = \{0,...,n\} \times \{0,...,m\}$. On définit le jeu $\Gamma(n,m)$ suivant, où les joueurs jouent à tour de rôle.

Au début du jeu, une pierre est placée sur chaque point de la grille $P(n,m)$. Le joueur 1 commence et choisit une pierre. Il enlève cette pierre ainsi que toutes les pierres ayant l'abscisse et l'ordonnée au moins égales à celles de la pierre choisie. Puis c'est au joueur 2 de jouer : il choisit une des pierres restantes, enlève cette pierre ainsi que toutes les pierres ayant l'abscisse et l'ordonnée au moins égales à celles de la pierre qu'il vient de choisir. Etc... Le jeu continue et le joueur qui prend la dernière pierre (qui est forcément $(0,0)$) a perdu.

1. Pour $n \geq 1$, résoudre le jeu $\Gamma(n,n)$.

2. Pour n et m quelconques, déterminer quel joueur a une stratégie gagnante dans $\Gamma(n,m)$

3. Problème ouvert : trouver un premier coup gagnant dans $\Gamma(n,m)$.

Annexe 4 : Exercice : duopole à la Cournot

2 firmes produisent le même bien avec coût marginal constant $c > 0$. La demande des consommateurs est $D(p) = 1 - p$. Simultanément chaque firme i produit une quantité q_i, puis vend toute sa production au prix $p = 1 - (q_1 + q_2)$. Les firmes veulent maximiser leur profit. Montrer qu'il y a un unique équilibre de Nash et le calculer.