Prédiction de suites individuelles et théorie des jeux

Nathan Levy et Samuel Ronsin

3 avril 2006

Table des matières

1 Préliminaires

Le problème de la prédiction de suites individuelles peut être formulé de différentes façons, le principe général reste toujours le même : un joueur doit choisir une action parmi un ensemble de choix, il écope d'une perte qu'il doit minimiser. Plus précisement, il reçoit avant de jouer les conseils de différents experts et doit faire à peu près aussi bien que le meilleur d'entre eux.

1.1 Problème général

On considère un ensemble de sortie, $\mathcal{Y}$ et un ensemble de décision, $\mathcal{X}$ (on prendra un sous espace convexe d'un espace vectoriel). On considère également une fonction de perte $\ell : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. On se dote de plus d'une classe finie d'experts : $\mathcal{E}$.

A chaque instant t,

1. L'environnement choisit $y_t \in \mathcal{Y}$ et ne le révèle pas
2. Chaque expert $e \in \mathcal{E}$ choisit $f_{e,t} \in \mathcal{X}$ et le rend public
3. Le joueur choisit $p_t \in \mathcal{X}$
4. y_t est rendu publique
5. On calcule la perte du joueur : $\ell(p_t, y_t)$ et celle des experts : $\ell(f_{e,t}, y_t)$, $e \in \mathcal{E}$

Typiquement, il faut voir la fonction de perte comme une distance entre la prédiction et la réalité : le joueur doit être le plus près possible de y_t, d'où le terme de *prédiction*.

Soulignons le fait que le joueur prédit de façon *séquentielle* : il choisit p_t à chaque tour en ne connaissant que le passé jusqu'à $t-1$ et sans tenir compte du nombre de coup total (potentiellement infini). De plus, on cherchera à obtenir des résultat valables *quelle que soit la suite* y_n, on ne cherche pas à modéliser les probabilités de sortie.

Notation : On notera $y_{1..n}$ la suite $(y_1, \ldots, y_n)$ et $N = \operatorname{card} \mathcal{E}$.

Définition 1.1.1 (Perte cumulée)**.** *On définit la* perte cumulée *jusqu'au temps n :*
- *pour le joueur :* $L(y_{1..n}) = \sum_{t=1}^{n} \ell(p_t, y_t)$
- *pour l'expert e :* $L_e(y_{1..n}) = \sum_{t=1}^{n} \ell(f_{e,t}, y_t)$.

Définition 1.1.2 (Regret cumulé)**.** *De même, le* regret cumulé *par rapport à l'expert e est :*

$$R_{e,n} = L(y_{1..n}) - L_e(y_{1..n}) = \sum_{t=1}^{n} (\ell(p_t, y_t) - \ell(f_{e,t}, y_t)) \ .$$

Le *regret* exprime l'intérêt qu'aurait eu le joueur à écouter l'expert e. On cherche à avoir un regret négligeable par rapport à n pour tous les experts, cela quelle que soit la suite $(y_t)_{t \in \mathbb{N}}$:

$$\max_{e \in \mathcal{E}} R_{e,n} = L(y_{1..n}) - \min_{e \in \mathcal{E}} L_e(y_{1..n}) = o(n) \ .$$

Cela peut être obtenu sous de faibles hypothèses par l'algorithme à pondération exponentielle

1.2 L'algorithme à pondération exponentielle

L'idée, assez naturelle, est de faire plus confiance à l'expert qui a subi la moindre perte, cela pour deux raisons : tout d'abord parce qu'il y a de fortes chances qu'il soit bon, mais aussi et surtout parce que c'est lui qui réalise $\min_{e \in \mathcal{E}} L_e(y_{1..n})$ et donc en jouant comme lui, on fait moins augmenter le regret. On prendra comme coefficients de pondération les regrets cumulés auxquels on applique une fonction exponentielle avec un paramètre η à choisir.

Définition 1.2.1. (Algorithme à pondération exponentielle).

On définit l'algorithme à pondération exponentielle de façon séquentielle
- $p_1 = \sum_{e\in\mathcal{E}} f_{e,1}/N$
- *pour* $t \geq 2$

$$p_t = \frac{\sum_{e\in\mathcal{E}} e^{\eta R_{e,t-1}} f_{e,t}}{\sum_{e\in\mathcal{E}} e^{\eta R_{e,t-1}}} = \frac{\sum_{e\in\mathcal{E}} e^{-\eta L_e(y_{1..t-1})} f_{e,t}}{\sum_{e\in\mathcal{E}} e^{-\eta L_e(y_{1..t-1})}} \quad (\in \mathcal{X} \text{ par convexité})$$

Théorème 1.2.1. *On se donne une fonction de perte $\ell : \mathcal{X} \times \mathcal{Y} \to [0,1]$ où $\mathcal{X}$ est un sous espace convexe d'un espace vectoriel. Supposons que la fonction de perte ℓ est convexe en son premier argument. Alors pour tout $\eta > 0$, pour tout n et pour toute suite $y_1, \ldots, y_n \in \mathcal{Y}$, le regret de l'algorithme pondération exponentielle vérifie :*

$$\max_{e\in\mathcal{E}} R_{e,n} \leq \frac{\ln N}{\eta} + \frac{n\eta}{8} \ ,$$

où $N = \operatorname{card} \mathcal{E}$. En particulier, avec $\eta = \sqrt{8(\ln N)/n}$ la borne supérieure devient $\sqrt{(n/2)\ln N}$.

Remarque : Le fait que η dépende du nombre de coup est un problème, cela contredit la contrainte de *prédiction séquentielle* : on ne le connaît pas à l'avance, il peut même être infini. Il y a plusieurs façon d'y remédier, par exemple en prenant un $\eta_t = \sqrt{8(\ln N)/t}$ qui dépend du temps, on obtient

$$\max_{e\in\mathcal{E}} R_{e,n} \leq 2\sqrt{\frac{n}{2}\ln N} + \sqrt{\frac{\ln N}{8}} \ . \tag{1.2.1}$$

Démonstration. On note

$$p_{e,t} = \frac{e^{\eta R_{e,t-1}}}{\sum_{e\in\mathcal{E}} e^{\eta R_{e,t-1}}} = \frac{e^{-\eta L_e(y_{1..t-1})}}{\sum_{e\in\mathcal{E}} e^{-\eta L_e(y_{1..t-1})}} \ ,$$

de sorte que

$$p_t = \sum_{e\in\mathcal{E}} p_{e,t} f_{e,t} \ .$$

On peut considérer les $(p_{e,t})_{e\in\mathcal{E}}$ comme une distribution de probabilité sur $\mathcal{E}$ Alors,

$$\begin{aligned}
\max_{e\in\mathcal{E}} R_{e,n} &= \sum_{t=1}^{n} \ell(p_t, y_t) - \min_{e\in\mathcal{E}} L_e(y_{1..n}) \\
&= \sum_{t=1}^{n} \ell\Big(\sum_{e\in\mathcal{E}} p_{e,t} f_{e,t}, y_t\Big) - \min_{e\in\mathcal{E}} L_e(y_{1..n}) \\
&\leq \sum_{t=1}^{n} \sum_{e\in\mathcal{E}} p_{e,t} \ell(f_{e,t}, y_t) - \min_{e\in\mathcal{E}} L_e(y_{1..n}) \qquad (1.2.2)
\end{aligned}$$

(ℓ est convexe en son premier argument : c'est l'inégalité de Jensen).

Pour $t = 1$ et $e \in \mathcal{E}$, on note

$$w_{e,t} = e^{-\eta L_e(y_{1..t-1})} \text{ et } W_t = \sum_{e \in \mathcal{E}} w_{e,t} \text{ tel que } p_{e,t} = \frac{w_{e,t}}{W_t} \ .$$

Alors, d'une part,

$$\begin{aligned} \ln \frac{W_{n+1}}{W_1} &= \ln \left(\sum_{e \in \mathcal{E}} e^{-\eta L_e(y_{1..n})} \right) - \ln N \geq \ln \left(\max_{e \in \mathcal{E}} e^{-\eta L_e(y_{1..n})} \right) - \ln N \\ &\geq -\eta \min_{e \in \mathcal{E}} L_e(y_{1..n}) - \ln N \ , \end{aligned} \tag{1.2.3}$$

et d'autre part, pour tout $t = 1, \ldots, n$,

$$\begin{aligned} \ln \frac{W_{t+1}}{W_t} &= \ln \frac{\sum_{e \in \mathcal{E}} e^{-\eta \ell_{e,t}} e^{-\eta L_e(y_{1..t-1})}}{\sum_{e \in \mathcal{E}} e^{-\eta L_e(y_{1..t-1})}} = \ln \left(\sum_{e \in \mathcal{E}} p_{e,t} e^{-\eta \ell_{e,t}} \right) \\ &= \ln \mathbb{E} \left[e^{-\eta \ell_{e,t}} \right] \\ &= \ln \mathbb{E} \left[e^{\eta(\mathbb{E}[\ell_{e,t}] - \ell_{e,t})} \right] - \eta \sum_{e \in \mathcal{E}} p_{e,t} \ell_{e,t} \\ &\leq -\eta \sum_{e \in \mathcal{E}} p_{e,t} \ell_{e,t} + \frac{\eta^2}{8} \quad \text{inégalité de Hoeffding (voir en appendice).} \end{aligned}$$

On somme sur $t = 1, \ldots, n$:

$$\ln \frac{W_{n+1}}{W_1} \leq \frac{n\eta^2}{8} - \eta \sum_{t=1}^{n} \sum_{e \in \mathcal{E}} p_{e,t} \ell(f_{e,t}, y_t) \ . \tag{1.2.4}$$

En combinant les deux inégalités (1.2.3) et (1.2.4) on obtient

$$-\eta \min_{e \in \mathcal{E}} L_e(y_{1..n}) - \ln N \leq \frac{n\eta^2}{8} - \eta \sum_{t=1}^{n} \sum_{e \in \mathcal{E}} p_{e,t} \ell(f_{e,t}, y_t) \ .$$

On applique l'inégalité (1.2.2) et on divise par η,

$$\max_{e \in \mathcal{E}} R_{e,n} \leq \frac{\ln N}{\eta} + \frac{n\eta}{8} \ .$$

$\square$

1.3 Stratégies pondérées

Nous allons maintenant changer légèrement de point de vue : on prend comme ensemble de décision $\mathcal{X}$ un ensemble fini $\{1, \ldots, N\}$. De plus, à l'instant t, au lieu de choisir une action dans $\mathcal{X}$, le joueur choisit une distribution de probabilité sur $\mathcal{X}$: $\boldsymbol{p}_t = (p_{1,t}, \ldots, p_{N,t})$ et joue l'action i avec la probabilité $p_{i,t}$. Formellement, le joueur dispose de $(U_t)_{t \in \mathbb{N}}$ une suite de variable aléatoires indépendantes uniformément distribuées sur $[0, 1[$ et I_t, l'action du joueur à l'instant t est définie comme il suit :

$$I_t = i \text{ si et seulement si } U_t \in \left[\sum_{j=1}^{i-1} p_{j,t}, \sum_{j=1}^{i} p_{j,t} \right[\ .$$

Cela permet notamment d'avoir des tirages indépendants.

On considère comme au paragraphe précédent une fonction de perte à valeurs dans $[0,1]$ et on cherche à rendre la perte cumulée la plus proche possible de celle du meilleur expert. Nous nous limiterons dans cet exposé à une classe d'experts particulière : les experts constants qui conseillent toujours la même action. Comme auparavant, on autorise le tirage à l'instant t à dépendre des coups du joueurs au instants précedents (et de la façon dont ils sont tirés au hasard), ainsi, le coup de l'environnement, Y_t est une fonction mesurable des variables $U_1, \ldots, U_{t-1}$. Formellement, on cherche à atteindre

$$R_n = L_n - \min_{i=1,\ldots,N} L_{i,n} = \sum_{t=1}^{n} \ell(I_t, Y_t) - \min_{i=1,\ldots,N} \sum_{t=1}^{n} \ell(i, Y_t) = o(n) \qquad \text{p.s} \ .$$

Cette propriété est la *consistance de Hannan*

Définition 1.3.1 (Consistance de Hannan, consistant au sens de Hannan)**.** *On appelle consistance de Hannan la propriété suivante :*

$$\limsup_{n\to\infty} \frac{1}{n} \left(\sum_{t=1}^{n} \ell(I_t, Y_t) - \min_{i=1,\ldots,N} \sum_{t=1}^{n} \ell(i, Y_t) \right) \leq 0 \qquad \text{p.s} \ ,$$

un algorithme qui vérifie cette propriété est dit consistant au sens de Hannan.

L'existance d'algorithmes consistants au sens de Hannan sera essentielle dans la suite du développement. L'algorithme à pondération exponentielle est consistant au sens de Hannan, en voici la démonstration.

On introduit la notion *espérance de perte*, c'est l'espérance conditionnelle de $\ell(I_t, Y_t)$ par rapport à la tribu engendrée par $U_1, \ldots, U_{t-1}$.

Définition 1.3.2 (Espérance de perte)**.**

$$\overline{\ell}(\boldsymbol{p}_t, Y_t) = \mathbb{E}\left[\ell(I_t, Y_t) | U_1, \ldots, U_{t-1}\right] = \sum_{i=1}^{N} p_{i,t} \ell(i, Y_t) \ .$$

Cela permet de tuer la composante aléatoire qui dépend de l'instant t c'est-à-dire le tirage au hasard de I_t ($\boldsymbol{p}_t$ et y_t ne dépendent pas de U_t), ainsi $\overline{\ell}(\boldsymbol{p}_t, Y_t)$ ne dépend que du passé jusqu'à l'instant $t-1$.

Remarque : La fonction d'espérance de perte prend son premier argument dans le simplexe sur $\mathcal{X}$ ce qui permet de se ramener à une partie convexe d'un espace vectoriel et donc d'utiliser les resultats précédents.

Définition 1.3.3 (Espérance de perte cumulée, espérance de regret)**.** *On aura de même les notions d'espérance de perte cumulée et d'espérance de regret notés* $\overline{L}_n = \sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p}_t, Y_t)$ *et* $\overline{R}_n = \overline{L}_n - \min_{i=1,\ldots,N} L_{i,n}$.

L'idée est de raisonner en deux temps, on borne d'abord l'espérance de regret puis la différence entre l'espérance de regret et le regret réel. L'espérance de perte sur le simplexe sur $\mathcal{X}$ vérifie les hypothèses du Théorème 1.2.1 (elle est même linéaire en son premier argument). On retrouve ainsi l'inégalité (1.2.1),

Lemme 1.3.1. *Dans le cadre des stratégies pondérées, pour tout* n *et pour toute suite* $Y_1, \dots, Y_n \in \mathcal{Y}$*, le regret de l'algorithme à pondération exponentielle vérifie :*

$$\overline{R}_n \leq 2\sqrt{\frac{n}{2}\ln N} + \sqrt{\frac{\ln N}{8}} \ .$$

On peut ensuite majorer la différence entre L_n et $\overline{L}_n$

Lemme 1.3.2. *Pour tout* $\delta > 0$*, avec une probabilité au moins* $1 - \delta$

$$L_n - \overline{L}_n \leq \sqrt{\frac{n}{2}\ln\frac{1}{\delta}} \ .$$

Remarque : la probabilité de référence est celle associée à la suite des $(U_t)_{t\in\mathbb{N}}$.

Démonstration. Ce lemme découle de l'inégalité d'Hoeffding-Azuma (voir en appendice). On considère la filtration canonique associée aux $(U_t)_{t\in\mathbb{N}}$. Alors, $\big(\ell(I_t, Y_t) - \overline{\ell}(\boldsymbol{p}_t, Y_t)\big)_{t\in\mathbb{N}}$ est un acroissement de martingale. De plus, $\overline{\ell}(\boldsymbol{p}_t, Y_t)$ est $\mathcal{F}_{t-1}$–mesurable et $\ell(I_t, Y_t) \in [0, 1]$. On peut donc appliquer l'inégalité avec $x = \sqrt{\frac{n}{2}\ln\frac{1}{\delta}}$ et $c_t = 1 (\forall t \in \mathbb{N})$. □

En combinant les lemmes 1.3.1 et 1.3.2, on obtient

Lemme 1.3.3. *Pour tout* $\delta > 0$*, avec une probabilité au moins* $1 - \delta$

$$R_n \leq 2\sqrt{\frac{n}{2}\ln N} + \sqrt{\frac{\ln N}{8}} + \sqrt{\frac{n}{2}\ln\frac{1}{\delta}} \ .$$

Théorème 1.3.1. *Dans le cadre des stratégies pondérées, l'algorithme à pondération exponentielle est Hannan-consistant.*

La preuve donnera même une convergence en $O(\sqrt{\frac{\ln n}{n}})$.

Démonstration. On applique le lemme 1.3.3 avec $\delta = \frac{1}{n^2}$: avec une probabilité au moins $1 - \frac{1}{n^2}$

$$R_n \leq u_n \quad \text{où} \quad u_n = 2\sqrt{\frac{n}{2}\ln N} + \sqrt{\frac{\ln N}{8}} + \sqrt{n\ln n}$$

$$\mathbb{P}\left[\frac{R_n}{u_n} > 1\right] \leq \frac{1}{n^2} \quad \text{donc par le lemme de Borel-Cantelli}$$

$$\mathbb{P}\left[\limsup_{n\to\infty}\left\{\frac{R_n}{u_n} > 1\right\}\right] = 0 \ .$$

Avec une probabilité 1, il existe seulement un nombre fini de rangs tels que $\frac{R_n}{u_n} > 1$ donc avec une probabilité 1

$$\limsup_{n\to\infty}\left(\frac{R_n}{u_n}\right) \leq 1 \ .$$

Or $\lim_{n\to\infty}\frac{1}{n}u_n = 0$. Cela prouve le théorème et donne la convergence en $O(\sqrt{\frac{\ln n}{n}})$ □

Remarque : L'algorithme à pondération exponentielle vérifie aussi

$$\lim_{n\to\infty} \frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, Y_t) - \min_{i=1,\dots,N} \sum_{t=1}^{n} \ell(i, Y_t)\right) = 0 \qquad \text{p.s} \ .$$

(Dans la définition de la consistance de Hannan, on a seulement une limsup négative). En effet, on peut appliquer la même démonstration à $-R_n$ avec à la place du lemme 1.3.1 l'inégalité $-\overline{R}_n \leq 0$ obtenue ainsi,

D'une part :

$$\begin{aligned}\ln \frac{W_{n+1}}{W_1} &= \ln\left(\sum_{i=1}^{n} e^{-\eta L_{i,n}}\right) - \ln N \leq \ln\left(N \max_{i=1,\dots,n} e^{-\eta L_{i,n}}\right) - \ln N \\ &\leq -\eta \min_{i=1,\dots,n} L_{i,n} \ ,\end{aligned}$$

d'autre part, pour $t = 1, \dots, n$:

$$\begin{aligned}\ln \frac{W_{t+1}}{W_t} &= \ln\left(\sum_{i=1}^{n} p_{j,t} e^{-\eta \ell_{j,t}}\right) \\ &\geq -\eta \sum_{i=1}^{n} p_{i,t}\ell_{i,t} \quad \text{(inégalité de Jensen)} \\ &\geq -\eta\, \overline{\ell}_t(\boldsymbol{p}_t) \ .\end{aligned}$$

On somme sur $t = 1, \dots, n$:

$$\ln \frac{W_{n+1}}{W_1} \geq -\eta\, \overline{L}_n \quad \text{on obtient} \quad \min_{i=1,\dots,n} L_{i,n} \leq \overline{L}_n \quad \text{puis} \quad -\overline{R}_n \leq 0 \ .$$

Le deuxième point de l'inégalité d'Hoeffding-Azuma donne une version symétrique du lemme 1.3.2.

1.4 Cadre pour l'exposé

Nous alons à présent définir un cadre plus restreint pour la suite de l'exposé. Nous allons appliquer les résultats rappelés plus haut à des problèmes de théorie des jeux. On se donne un jeu à K joueurs, chaque joueur k ($k \in \{1, \dots, K\}$) a le choix à chaque tour entre N_k actions possibles (on appelle ces actions des stratégies pures). Si chaque joueur k choisit l'action i_k, on appelle $\boldsymbol{i} = (i_1, \dots, i_K) \in \prod_{k=1}^{K}\{1, \dots, N_K\}$ le K-uplet des actions de tous les joueurs et dans ce cas, le joueur k subit une perte $\ell^{(k)}(\boldsymbol{i})$ qu'on supposera comprise entre 0 et 1.

Une stratégie mixte pour le joueur k est une distribution de probabilité $\boldsymbol{p}^{(k)} = \big(p_1^{(k)}, \dots, p_{N_k}^{(k)}\big)$ sur l'ensemble $\{1, \dots, N_k\}$. Quand des stratégies mixtes sont utilisées, les joueurs choisissent une action au hasard selon la distribution donnée par la stratégie mixte. On note alors I_k l'action jouée par le joueur k, c'est une variable aléatoire qui prend ses valeurs dans $\{1, \dots, N_k\}$ avec la loi $\boldsymbol{p}^{(k)}$. Soit $\boldsymbol{I} = (I_1, \dots, I_K)$ le *profil d'action*. On note π la loi de $\boldsymbol{I}$, on appelle π le *profil de stratégies mixtes*. Si les variables aléatoires $I_1, \dots, I_K$ sont indépendantes (c'est-à-dire si les joueurs tirent leurs actions indépendamment les uns des autres), π est la probabilité produit $p^{(1)} \otimes \cdots \otimes p^{(K)}$ sur l'ensemble $\prod_{k=1}^{K}\{1,\ldots,N_k\}$. Ainsi, pour $\boldsymbol{i} = (i_1,\ldots,i_K) \in \prod_{k=1}^{K}\{1,\ldots,N_k\}$

$$\pi(\boldsymbol{i}) = \mathbb{P}[\boldsymbol{I} = \boldsymbol{i}] \quad \left(= p^{(1)}_{i_1} \cdots p^{(K)}_{i_K} \text{ si } \pi \text{ est une probabilité produit} \right) .$$

Définition 1.4.1 (Espérance de perte)**.** *L'espérance de perte pour le joueur k est définie ainsi :*

$$\begin{aligned}
\mathbb{E}_\pi \left[\ell^{(k)}\right] &= \mathbb{E}\left[\ell^{(k)}(\boldsymbol{I})\right] \\
&= \sum_{\boldsymbol{i} \in \prod_{k=1}^{K}\{1,\ldots,N_k\}} \pi(\boldsymbol{i})\ell^{(k)}(\boldsymbol{i}) \\
\Bigg(&= \sum_{i_1=1}^{N_1} \cdots \sum_{i_K=1}^{N_K} p^{(1)}_{i_1} \cdots p^{(K)}_{i_K} \ell^{(k)}(i_1,\ldots,i_K) \\
&\text{si } \pi \text{ est une probabilité produit}\Bigg) .
\end{aligned}$$

Le problème étant défini, nous pouvons maintenant utiliser ce cadre pour obtenir des résultats de théorie des jeux.

2 Jeux à deux joueurs

Nous allons aborder, dans ce chapitre, une première illustration des méthodes développées dans la théorie de la prédiction de suites individuelles. Les jeux à deux joueurs présentent un intérêt particulier en théorie des jeux, en raison de la simplicité de la notion d'équilibre. Un autre point essentiel, dans le contexte de prédiction de suites, sera la convergence de la fréquence empirique des profils d'actions vers certains ensembles d'équilibres.

2.1 Théorèmes minimax

Un concept fondamental de théorie des jeux est celui d'équilibre. Considérons deux joueurs qui jouent l'un contre l'autre. Leur jeu est formalisé de la manière suivante. On se donne une fonction $\ell : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, qui représente la perte algébrique du premier joueur : lorsque le premier joueur choisit l'action $x \in \mathcal{X}$ et le second $y \in \mathcal{Y}$, alors, le premier joueur perd $\ell(x, y)$. On suppose le jeu à somme nulle, c'est-à-dire que ce qui est perdu (resp. gagné) par l'un est gagné (resp. perdu) par l'autre. Une concept d'équilibre naturel peut alors être explicité :

Définition 2.1.1 (équilibre à deux joueurs)**.** *Un équilibre est un couple d'actions (x^*, y^*) tel que pour tout $x \in \mathcal{X}$, $\ell(x^*, y^*) \leq \ell(x, y^*)$ et pour tout $y \in \mathcal{Y}$, $\ell(x^*, y^*) \geq \ell(x^*, y)$*

Dans une situation d'équilibre, aucun des deux joueurs n'aura intérêt à modifier son action si son adversaire ne change pas la sienne. L'aspect naturel de ce concept est étayé par l'hypothèse d'indépendance des joueurs : on suppose qu'ils ne communiquent pas leur intention et qu'ils dévoilent simultanément leur action, sans rien savoir de ce que jouera l'autre.

Remarque : On observe que si un tel équilibre (x^*, y^*) existe, alors :

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) \le \sup_{y\in\mathcal{Y}} \ell(x^*,y) \le \ell(x^*,y^*) \le \inf_{x\in\mathcal{X}} \ell(x,y^*) \le \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y)$$

Or, on remarque facilement que $\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) \ge \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y)$. L'existence d'un équilibre implique donc que

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) = \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) \ .$$

Définition 2.1.2. *On appelle valeur du jeu et l'on note V la quantité*

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) = \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) \ ,$$

si elle existe.

Cette notion de valeur est particulièrement importante en théorie des jeux, puisqu'elle représente l'unique valeur possible de la perte du permier joueur (et donc du gain du second) en situation d'équilibre.

Remarque : Sous l'hypothèse de l'existence de la valeur V du jeu, une condition nécessaire et suffisante à l'existence d'un équilibre est : $\exists (x^*, y^*), \ell(x^*, y^*) = V$. Autrement dit, toujours sous cette hypothèse, il est nécessaire et suffisant que la valeur du jeu soit atteinte.

Il est aussi à noter que si le joueur ligne (resp. colonne) choisit l'action x^* (resp. y^*), il est certain de réaliser au pire la perte V (resp. au mieux le gain $-V$).

Les théorèmes de minimax sont des conditions suffisantes sur les données du jeu (ensembles d'actions $\mathcal{X}$ et $\mathcal{Y}$, propriétés de la fonction ℓ ...) pour que la valeur du jeu soit bien définie.

Nous allons énoncer le théorème dans la version établie par Von Neumann en 1928, considéré comme un théorème fondamental en théorie des jeux. Avant cela, rappelons le concept de stratégie mixte. Considérons un jeu $\ell : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. On peut alors, au lieu de considérer les actions $x \in \mathcal{X}$ et $y \in \mathcal{Y}$, manipuler les distributions de probabilités $\boldsymbol{p} \in \mathcal{P}$ et $\boldsymbol{q} \in \mathcal{Q}$ respectivement sur $\mathcal{X}$ et $\mathcal{Y}$. La nouvelle fonction de perte sera définie par $\overline{\ell}(\boldsymbol{p}, \boldsymbol{q}) = \mathbb{E}_{\boldsymbol{p}\otimes\boldsymbol{q}}[\ell]$. Le théorème de Von Neumann assure qu'un jeu fini à deux personnes admet, en stratégies mixtes, au moins un équilibre.

Théorème 2.1.1 (Minimax de Von Neumann, 1928)**.** *Soit $\mathcal{X}$, $\mathcal{Y}$ et ℓ un jeu fini à deux joueurs. On note, comme précédemment, $\mathcal{P}$ et $\mathcal{Q}$ les ensembles des mesures de probabilité sur $\mathcal{X}$ et $\mathcal{Y}$ respectivement. Alors :*

$$\min_{\boldsymbol{p}\in\mathcal{P}} \max_{\boldsymbol{q}\in\mathcal{Q}} \overline{\ell}(\boldsymbol{p}, \boldsymbol{q}) = \max_{\boldsymbol{q}\in\mathcal{Q}} \min_{\boldsymbol{p}\in\mathcal{P}} \overline{\ell}(\boldsymbol{p}, \boldsymbol{q}) \ .$$

Un raffinement essentiel, qui nous permettra d'illustrer les méthodes de prédiction, est le théorème énoncé par Sion en 1958, dont nous fournissons une preuve originale, fondée sur l'inégalité du théorème 1.2.1.

Théorème 2.1.2 (Minimax de Sion, 1958)**.** *Soit $\ell : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ une fonction réelle bornée. On suppose $\mathcal{X}$ et $\mathcal{Y}$ convexes, $\mathcal{X}$ compact. On suppose de plus que pour tout $y \in \mathcal{Y}$, $\ell(.,y)$ est une fonction convexe, semi-continue supérieurement, et que pour tout $x \in \mathcal{X}$, $\ell(x,.)$ est concave.Alors,*

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) = \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) \ .$$

Démonstration. L'inégalité $\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) \geq \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y)$ est claire. D'autre part, comme ℓ est bornée, on peut supposer, à une multiplication et une translation près, que f prend ses valeurs dans $[0,1]$. Soit $\delta > 0$. Par compacité de $\mathcal{X}$, il existe un entier $N > 0$, et des points $x^{(1)}, \ldots, x^{(N)} \in \mathcal{X}$ tels que $(B(x^{(i)}, \delta))_{1\leq i\leq N}$ soit un recouvrement ouvert de $\mathcal{X}$. Soit maintenant n un entier strictement positif. On définit les suites $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ par récurrence :

$$x_t = \frac{\sum_{i=1}^N x^{(i)} e^{-\eta \sum_{s=0}^{t-1} \ell(x^{(i)}, y_s)}}{\sum_{i=1}^N e^{-\eta \sum_{s=0}^{t-1} \ell(x^{(i)}, y_s)}} \ ,$$

avec $\eta = \sqrt{8\frac{lnN}{n}}$ et y_t choisi de sorte que $\ell(x_t, y_t) \geq \sup_{y\in\mathcal{Y}} \ell(x_t, y) - \frac{1}{n}$. On peut alors traduire cela en terme de prédiction de suites : la fonction de perte est ℓ, les experts constants $x^{(1)}, \ldots, x^{(N)}$, les prévisions $x_1, \ldots, x_t$ et les résultats $y_1, \ldots, y_t$. Le théorème 1.2.1 assure alors que :

$$\frac{1}{n}\sum_{t=1}^n \ell(x_t, y_t) \leq \min_{i=1,\ldots,N} \sum_{i=1}^n \ell(x^{(i)}, y_t) + \sqrt{\frac{\ln N}{2n}} \ .$$

On en déduit alors :

$$\begin{aligned}
\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) &\leq \sup_{y\in\mathcal{Y}} \ell(\frac{1}{n}\sum_{t=1}^n x_t, y) \\
&\leq \sup_{y\in\mathcal{Y}} \frac{1}{n}\sum_{t=1}^n \ell(x_t, y) \qquad \text{par convexité en le premier argument} \\
&\leq \frac{1}{n}\sum_{t=1}^n \sup_{y\in\mathcal{Y}} \ell(x_t, y) \\
&\leq \frac{1}{n}\sum_{t=1}^n \ell(x_t, y_t) + \frac{1}{n} \qquad \text{par définition de } y_t \\
&\leq \min_{1\leq i\leq N} \frac{1}{n}\sum_{t=1}^n \ell(x^{(i)}, y_t) + \frac{1}{n} + \sqrt{\frac{\ln N}{2n}} \\
&\leq \min_{1\leq i\leq N} \ell(x^{(i)}, \frac{1}{n}\sum_{t=1}^n y_t) + \frac{1}{n} + \sqrt{\frac{\ln N}{2n}}
\end{aligned}$$

par concavité en le second argument

$$\leq \sup_{y\in\mathcal{Y}} \min_{1\leq i\leq N} \ell(x^{(i)}, y) + \frac{1}{n} + \sqrt{\frac{\ln N}{2n}} \ .$$

Or cette inégalité est valable pour tout n, et ce indépendamment de N (qui ne dépend que de δ). En faisant tendre n vers $+\infty$, on obtient :

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \ell(x,y) \leq \sup_{y\in\mathcal{Y}} \min_{1\leq i\leq N} \ell(x^{(i)},y) \ .$$

Pour conclure, il reste à voir

$$\sup_{y\in\mathcal{Y}} \min_{1\leq i\leq N_\delta} \ell(x^{(i)},y) \xrightarrow[\delta\to 0]{} \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) \ .$$

Pour cela, donnons-nous un réel $\varepsilon > 0$. On considère alors y_ε tel que :

$$|\inf_{x\in\mathcal{X}} \ell(x,y_\varepsilon) - \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{Y}} \ell(x,y)| < \varepsilon \ .$$

Choisissons maintenant x_ε tel que :

$$|\inf_{x\in\mathcal{X}} \ell(x,y_\varepsilon) - \ell(x_\varepsilon,y_\varepsilon)| < \varepsilon \ .$$

En fait,

$$\sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) - \varepsilon < \ell(x_\varepsilon,y_\varepsilon) < \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{Y}} \ell(x,y) + \varepsilon \ .$$

Par semi-continuité supérieure, il existe $\delta > 0$ tel que pour tout x tel que $\|(x - x_\varepsilon)\| < \delta$,

$$f(x,y_\varepsilon) < \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{Y}} \ell(x,y) + \varepsilon \ .$$

Prenons alors un recouvrement de $\mathcal{X}$ par des $B(x^i,\delta)$, $(1 \leq i \leq N)$. On remarque alors,

$$\begin{aligned}\sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) - \varepsilon < \inf_{x\in\mathcal{X}} \ell(x,y_\varepsilon) &\leq \min_{1\leq i\leq N} \ell(x^i,y_\varepsilon) \\ &\leq \sup_{y\in\mathcal{Y}} \min_{1\leq i\leq N} \ell(x^i,y) \leq \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y) + \varepsilon \ .\end{aligned}$$

En particulier, pour δ suffisamment petit,

$$|\sup_{y\in\mathcal{Y}} \min_{1\leq i\leq N_\delta} \ell(x^{(i)},y) - \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \ell(x,y)| < \varepsilon$$

ce qui conclut la preuve.

□

Remarque : La notion d'équilibre définie précédemment se généralise naturellement au cas de jeux à K joueurs, où $K \geq 2$. Un jeu à K joueurs sera la donnée de $\mathcal{X}_1 \dots \mathcal{X}_K$ (ensembles des actions possibles pour chacun des joueurs), et de $\ell_i : \mathcal{X}_1 \times \cdots \times \mathcal{X}_K \to \mathbb{R}$ $(\forall i \in \{1,\dots,K\})$ les fonctions de gain des joueurs. La définition 2.1.1 s'étend comme suit :

Définition 2.1.3 (Équilibre de Nash)**.** *Un équilibre de Nash est un K-uplet $(x_1^*,\dots,x_K^*) \in \mathcal{X}_1 \dots \mathcal{X}_K$ tel que :*

$$\forall i \in \{1,\dots,K\}, \forall x_i \in \mathcal{X}_i, \ell_i(x_1^*,\dots,x_i,\dots,x_K^*) \leq \ell_i(x_1^*,\dots,x_i,\dots,x_K^*)$$

Autrement dit, un équilibre de Nash est un choix fixé de chacun des K joueurs, pour lequel, aucun joueur n'a intérêt à revenir sur sa décision individuellement. Une telle situation sera donc stable au sens où si les joueurs ne communiquent pas entre eux, personne n'aura envie de bouger de cette situation.

On conclut cette section en énonçant le Théorème de Nash :

Théorème 2.1.3 (Théorème de Nash, 1951)**.** *Soient $\mathcal{X}_1, \dots, \mathcal{X}_K$ des parties convexes compactes d'un espace vectoriel normé. On suppose pour tout $i \in \{1, \dots, K\}$, ℓ_i continue sur $\prod_{i=1}^K \mathcal{X}_i$ et pour tout $(x_1, \dots, x_K)$ la fonction $y_i \to \ell_i(x_1, \dots, y_i, \dots, x_K)$ convexe. Alors, le jeu admet un équilibre de Nash.*

Remarque : Une conséquence immédiate de ce théorème est l'existence d'un équilibre de Nash en stratégies mixtes pour tout jeu fini.

2.2 Stratégies Hannan-consistantes

Dans cette partie, on étudie les conséquences d'une stratégie Hannan-consistante sur l'évolution du jeu. On verra qu'on s'approche, en un certain sens, de la valeur du jeu.

Notation : Le jeu est caractérisé par une matrice de perte ℓ. Dans le cadre du jeu à deux joueurs, on notera $\boldsymbol{p}$ (resp. $\boldsymbol{q}$) pour $\boldsymbol{p}^{(1)}$ (resp. $\boldsymbol{p}^{(2)}$) et N (resp. M) pour N_1 (resp. N_2). On rappelle que la valeur du jeu est définie par

$$V = \max_{\boldsymbol{q}} \min_{\boldsymbol{p}} \overline{\ell}(\boldsymbol{p}, \boldsymbol{q})$$

où le maximum est pris sur tous les vecteurs $\boldsymbol{q} = (q_1, \dots, q_M)$ et le minimum sur les vecteurs $\boldsymbol{p} = (p_1, \dots, p_N)$ et

$$\overline{\ell}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^{N} \sum_{j=1}^{M} p_i q_j \ell(i, j)$$

Théorème 2.2.1. *On considère un jeu à deux joueurs à somme nulle. Supposons que le joueur-ligne joue en suivant une stratégie Hannan-consistante. Alors*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \ell(I_t, J_t) \le V \qquad \text{p.s}\ .$$

Démonstration. La stratégie étant Hannan consistante, il suffit de prouver que pour $n \in \mathbb{N}$

$$\min_{i=1,\dots,N} \frac{1}{n} \sum_{t=1}^{n} \ell(i, J_t) \le V$$

Pour voir cela, il faut considérer $f : \boldsymbol{p} \to \frac{1}{n} \sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p}, J_t)$. C'est une forme linéaire sur le simplexe, elle atteint son minimum dans un coin. Ainsi

$$\min_{i=1,\dots,N} \frac{1}{n}\sum_{t=1}^{n} \ell(i,J_t) = \min_{\boldsymbol{p}} \frac{1}{n}\sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p},J_t)$$

$$= \min_{\boldsymbol{p}} \sum_{j=1}^{M} \widehat{q}_{j,n}\overline{\ell}(\boldsymbol{p},j)$$

(où $\widehat{q}_{j,n} = (1/n)\sum_{t=1}^{n} \mathbb{I}_{\{J_t=j\}}$ est la probabilité empirique que le coup du joueur-ligne soit j)

$$= \min_{\boldsymbol{p}} \overline{\ell}(\boldsymbol{p},\widehat{\boldsymbol{q}}_n)$$

(où $\widehat{\boldsymbol{q}}_n = (\widehat{q}_{1,n},\dots,\widehat{q}_{M,n})$)

$$\le \max_{\boldsymbol{q}} \min_{\boldsymbol{p}} \overline{\ell}(\boldsymbol{p},\boldsymbol{q}) = V$$

□

Le théorème 2.2.1 montre que, quels que soient les coups de l'adversaire, si le joueur-ligne utilise une stratégie Hannan-consistante, alors il peut être sûr que sa perte moyenne est asymptotiquement majorée par V. Par raison de symetrie, si les deux joueurs utilisent une stratégie Hannan-consistante, alors la perte moyenne du joueur-ligne converge vers V.

Corollaire 2.2.1. *On considère un jeu à deux joueurs à somme nulle. Supposons que les deux joueurs jouent en suivant une stratégie Hannan-consistante. Alors*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \ell(I_t,J_t) = V \qquad \text{p. s} \ .$$

Démonstration. Par le Théoreme 2.2.1

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \ell(I_t,J_t) \le V \qquad \text{p. s} \ .$$

Le même théorème appliqué au joueur-colonne implique, sachant que la perte du joueur-colonnes est l'opposé de celle du joueur ligne,

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \ell(I_t,J_t) \ge \min_{p} \max_{q} \overline{\ell}(p,q) \qquad \text{p. s} \ .$$

Par le théorème du minimax de von Neumann, la quantité de droite est égale à V. □

Remarque Si les deux joueurs suivent une stratégie consistante au sens de Hannan, alors la distribution produit $\widehat{p}_n \otimes \widehat{q}_n$, formée par les distributions empiriques marginales, converge, p. s, vers l'ensemble des équilibres de Nash du jeu.

$$\widehat{p}_{i,n} = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}_{\{I_t=i\}} \text{ et } \widehat{q}_{j,n} = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}_{\{J_t=j\}}$$

Cependant, il est important de noter que la fréquence empirique des profils d'actions : $\widehat{\pi}_n(i,j) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}_{\{(I_t,J_t)=(i,j)\}}$ ne converge pas vers l'ensemble des équilibres de Nash. La partie suivante traite des résultats qu'on peut obtenir sur cette convergence.

3 Convergence de la mesure empirique

Nous allons étudier dans cette partie les différents résultats obtenus sur la fréquence empirique des profils d'actions. Dans une première partie, nous étudierons un outil, le regret interne, puis nous étudierons deux types d'équilibres, les équilibres de Hannan et les équilibres corrélés. Enfin, nous verrons une autre approche du problème : le calibrage.

3.1 Regret interne

Avant de donner les résultats sur la convergence de la fréquence empirique des profils d'actions, voici un outil qui sera utile dans cette partie : *le regret interne.* L'idée est la suivante : un joueur a un faible regret interne si pour chaque couple d'experts (i,j) il ne regrette pas d'avoir suivi i au lieu de j chaque fois qu'il a suivi i.

Définition 3.1.1 (Regret interne)**.** *Le regret interne cumulé de l'algorithme à pondération $\boldsymbol{p}_t$ est donné par la formule suivante :*

$$\max_{\substack{i,j=1,\ldots,N\\ i\neq j}} R_{(i,j),n} = \max_{\substack{i,j=1,\ldots,N\\ i\neq j}} \sum_{t=1}^{n} p_{i,t}\big(\ell(i,Y_t) - \ell(j,Y_t)\big) \ .$$

Ainsi, $r_{(i,j),t} = p_{i,t}\big(\ell(i,Y_t) - \ell(j,Y_t)\big)$ est le regret que le joueur ressent d'avoir, au temps t, mis la masse $p_{i,t}$ sur le i–ème expert au lieu du j–ème.

Remarque : Un faible regret interne implique un faible regret externe (en espérance). En effet le regret externe de l'algorithme à pondération $\boldsymbol{p}_t$ vaut

$$\max_{j=1,\ldots,N} \sum_{i=1}^{N} R_{(i,j),n} \leq N \max_{\substack{i,j=1,\ldots,N\\ i\neq j}} R_{(i,j),n}$$

En revanche, la réciproque n'est pas vraie, ainsi, l'algorithme a pondération exponentielle peut avoir un regret interne qui croit linéairement. Voici une méthode qui permet de convertir un algorithme qui minimise le regret externe en un algorithme qui minimise le regret interne.

Notation : Pour $\boldsymbol{p}$ une distribution de probabilité sur l'ensemble des actions et i et j deux actions différentes, on définit $\boldsymbol{p}^{i\to j}$ la probabilité obtenu à partir de $\boldsymbol{p}$ en déplaçant la masse de i en j. Formellement, $p_i^{i\to j} = 0$, $p_j^{i\to j} = p_i + p_j$ et $p_k^{i\to j} = p_k$ pour $k \neq i,j$.

Algorithme :

On définit l'algorithme de façon séquentielle :

1. Au temps $t = 1$ on prend $\boldsymbol{p}_1 = (1/N, \ldots, 1/N)$
2. Pour $t > 1$, on considère une stratégie de minimisation du regret *externe* sur les experts fictifs de pertes $\ell(\boldsymbol{p}_{t-1}^{i\to j}, y_t)$ pour $i \neq j$ (on note $\boldsymbol{\Delta}_t$ la distribution de probabilité sur les paires $\{(i,j), i \neq j\}$)
3. Avec ce méta-algorithme, on définit $\boldsymbol{p}_t$ par l'équation de point fixe suivante :
$$\boldsymbol{p}_t = \sum_{(i,j):i\neq j} \Delta_{(i,j),t} \boldsymbol{p}_t^{i\to j} \ .$$

Concrètement, une stratégie de minisation du regret *externe* sur les experts $\boldsymbol{p}_{t-1}^{i\to j}$ est $\boldsymbol{\Delta}_t$, une distribution de probabilité sur les paires $\{(i,j), i \neq j\}$ telle que l'espérence de perte cumulée est bornée par celle de la meilleure stratégie modifiée (avec un terme corréctif qui tend vers 0). On peut par exemple utiliser l'algorithme à pondération exponentielle :

$$\Delta_{(i,j),t} = \frac{\exp\big(-\eta \sum_{s=1}^{t-1} \overline{\ell}(\boldsymbol{p}_s^{i\to j}, Y_s)\big)}{\sum_{(k,l):k\neq l} \exp\big(-\eta \sum_{s=1}^{t-1} \overline{\ell}(\boldsymbol{p}_s^{k\to l}, Y_s)\big)} \ .$$

Théorème 3.1.1. *L'algorithme ainsi défini permet d'obtenir un regret interne négligeable devant n.*

Démonstration. Notons avant tout que dans le point 3, l'existence d'une solution est assurée par le théorème du point fixe de Brouwer. Pour tout t

$$\overline{\ell}(\boldsymbol{p}_t, Y_t) = \sum_{\substack{i,j=1,\ldots,N \\ i\neq j}} \Delta_{(i,j),t} \overline{\ell}(\boldsymbol{p}_t^{i\to j}, Y_t) \ .$$

Ce qui permet de borner la perte cumulée :

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p}_t, Y_t) &= \frac{1}{n}\sum_{t=1}^{n} \sum_{\substack{i,j=1,\ldots,N \\ i\neq j}} \Delta_{(i,j),t} \overline{\ell}(\boldsymbol{p}_t^{i\to j}, Y_t) \\ &\leq \min_{\substack{i,j=1,\ldots,N \\ i\neq j}} \frac{1}{n}\sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p}_t^{i\to j}, Y_t) + \varepsilon_n \ ,\end{aligned}$$

avec $\varepsilon_n = 2\sqrt{\frac{n}{2}\ln\frac{N(N-1)}{2}} + \sqrt{\frac{\ln N(N-1)/2}{8}}$ (par l'inégalité (1.2.1)) ,

puis le regret interne :

$$\begin{aligned}\max_{\substack{i,j=1,\ldots,N \\ i\neq j}} \frac{1}{n}\sum_{t=1}^{n} \overline{\ell}(\boldsymbol{p}_t, Y_t) - \overline{\ell}(\boldsymbol{p}_t^{i\to j}, Y_t) &\leq \varepsilon_n \\ \frac{1}{n}\max_{\substack{i,j=1,\ldots,N \\ i\neq j}} R_{(i,j),n} &\leq \varepsilon_n \ .\end{aligned}$$

□

Remarque : La notion de regret interne définie ici est en fait à rapprocher du regret externe *en espérance*. On définira dans le paragraphe suivant une notion à rapprocher du regret externe : le *regret interne inconditionnel*. Le regret interne est l'espérance du regret interne inconditionnel par rapport au passé jusqu'à l'instant $t-1$.

3.2 Équilibres de Hannan et équilibres corrélés

On commence, dans cette section, par définir deux sous-sensembles de l'ensemble des mesures de probabilité sur $\prod_{k=1}^{K}\{1,\ldots,N_k\}$, qui vont jouer un rôle particulier dans l'analyse de la convergence des fréquences empiriques des profils d'actions.

Définition 3.2.1 (Ensemble de Hannan)**.** *Une distribution de probabilité π sur l'ensemble $\prod_{k=1}^{K}\{1,\ldots,N_k\}$ est dans l'ensemble de Hannan si pour tout $k=1,\ldots,K$ et $j\in\{1,\ldots,N_k\}$:*

$$\mathbb{E}\left[\ell^{(k)}(\boldsymbol{I})\right] \le \mathbb{E}\left[\ell^{(k)}(I^{-},j)\right] \ ,$$

avec les notations suivantes : $\boldsymbol{I}=(I_1,\ldots,I_K)$ est une variable aléatoire de loi π et $(\boldsymbol{I}^{-},j)=(I_1,\ldots,I_{k-1},j,I_{k+1},\ldots,I_K)$. On note $\mathcal{H}$ l'ensemble de Hannan.

Définition 3.2.2 (Équilibres corrélés)**.** *Une distribution de probabilité π sur l'ensemble $\prod_{k=1}^{K}\{1,\ldots,N_k\}$ est un équilibre corrélé si pour tout $k=1,\ldots,K$:*

$$\mathbb{E}\left[\ell^{(k)}(\boldsymbol{i})\right] \le \mathbb{E}\left[\ell^{(k)}(I^{-},I'_k)\right] \ ,$$

avec les notations suivantes : $\boldsymbol{I}=(I_1,\ldots,I_K)$ est une variable aléatoire de loi π et $(I^{-},I'_k)=(I_1,\ldots,I_{k-1},I'_k,I_{k+1},\ldots,I_K)$, I'_k étant une variable aléatoire à valeurs dans $\{1,\ldots,N_k\}$ quelconque mesurable par rapport à I_k (c'est à dire qu'il existe une fonction $\psi_k:\{1,\ldots,N_k\}\to\{1,\ldots,N_k\}$ telle que $I'_k=\psi_k(I_k)$. On note $\mathcal{C}$ l'ensemble des équilibres corrélés.

On introduit par ailleurs, pour simplifier les notations, la notion de regret interne inconditionnel.

Définition 3.2.3. *Le regret interne inconditionnel à l'instant n est la variable aléatoire définie par :*

$$\widehat{r}^{(k)}_{(j,j'),t} = \mathbb{I}_{\{I_t^{(k)}=j\}}\left(\ell^{(k)}(I_t)-\ell^{(k)}(I_t^{-},j')\right)$$

Remarques :

1. Le point clé est que contrairement aux équilibres de Nash, les équilibres corrélés ne sont pas nécessairement des équilibres produit (d'où l'adjectif "corrélés"). Un équilibre corrélé qui est une mesure produit est en fait un équilibre de Nash. L'existence d'équilibres de Nash nous assure l'existence d'équilibres corrélés. On peut interpréter les équilibres corrélés ainsi : avant chaque tour, les joueurs reçoivent une recommendation I_k telle que $I=(I_1,\ldots,I_K)$ a pour loi π ; aucun joueur n'a intérêt à dévier de sa recommandation tant que les autres ne dévient pas de la leur.

2. Contrairement aux équilibres de Hannan, les modifications considérées ne sont plus seulement les stratégies constantes : on considère aussi les stratégie mixtes mesurables par rapport à la stratégie de base.
3. On remarque aisément que $\mathcal{C} \subset \mathcal{H}$. L'inclusion inverse est fausse en général.

Rappelons la définition essentielle à la compréhension de ce paragraphe, celle des fréquences empiriques des profils d'actions.

Définition 3.2.4. *On appelle fréquence empirique des profils d'actions à l'instant n, la fonction définie par :*

$$\widehat{\pi}_n(\boldsymbol{i}) = \frac{1}{n} \sum_{t=1}^{n} \mathbb{I}_{\{\boldsymbol{I}_t = \boldsymbol{i}\}} \qquad \forall \boldsymbol{i} \in \prod_{k=1}^{K} \{1, \ldots, N_k\} \ .$$

Un premier résultat concerne les stratégies de prédiction à minimisation du regret externe :

Théorème 3.2.1. *Considérons un jeu à K personnes. On suppose que chacun des joueurs suit une stratégie de prédiction consistante au sens de Hannan. Alors,*

$$dist(\widehat{\pi}_n, \mathcal{H}) \xrightarrow[n \to \infty]{} 0 \quad i.e. \quad \widehat{\pi}_n \to \mathcal{H}$$

Nous ne démontrons pas ce résultat dont un preuve se calque immédiatement sur celle du prochain résultat en remplaçant $\mathcal{C}$ par $\mathcal{H}$ et "regret interne" par "regret externe".

On va maintenant démontrer que, moyennant un comportement raisonnable des joueurs (ici, on supposera qu'ils agissent selon une une stratégie de minimisation du regret interne), la mesure empirique du jeu (à l'instant n), atteint asymptotiquement l'ensemble des équilibres corrélés quand n tend vers l'infini. Le théorème suivant formalise cette intuition.

Théorème 3.2.2. *Considérons un jeu à K personnes. On suppose que chacun des joueurs suit une stratégie de minimisation de son regret interne. Alors,*

$$dist(\widehat{\pi}_n, \mathcal{C}) \xrightarrow[n \to \infty]{} 0 \quad (\widehat{\pi}_n \to \mathcal{C}) \ ,$$

c'est à dire :

$$\inf_{\pi \in \mathcal{C}} \sum_{i \in \prod_{k=1}^{K} \{1, \ldots, N_k\}} |\pi(i) - \widehat{\pi}_n(i)| \xrightarrow[n \to \infty]{} 0$$

Démonstration. Notons

$$r^{(k)}_{(j,j'),t} = p^{(k)}_{j,t} \big(\ell^{(k)}(I_t) - \ell^{(k)}(I_t^-, j') \big)$$

l'espérance conditionnelle de $\widehat{r}^{(k)}_{(j,j'),t}$ connaissant le passé avant t et les actions des autres joueurs. D'une part, l'analyse de la section précédente assure que $\max_{j,j'} \sum_{t=1}^{n} r^{(k)}_{(j,j'),t} = o(n)$. D'autre part pour tout j et j', $\big(r^{(k)}_{(j,j'),t} - \widehat{r}^{(k)}_{(j,j'),t}\big)_{t \geq 0}$ est une suite d'accroissements de martingales bornée. Alors, une démonstration semblable à celles du lemme 1.3.2 montre que pour tout $k \in \{1, \ldots, N\}$ et tous $j, j' \in \{1, \ldots, N_k\}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \big(r^{(k)}_{(j,j'),t} - \widehat{r}^{(k)}_{(j,j'),t} \big) = 0 \qquad \text{p. s}$$

$$\text{D'où} \qquad \limsup_{n\to\infty} \max_{j,j'} \frac{1}{n} \sum_{t=1}^{n} \widehat{r}^{(k)}_{(j,j'),t} \leq 0 \ .$$

Supposons maintenant que $\widehat{\pi}_n$ n'atteigne pas $\mathcal{C}$ asymptotiquement. Par compacité de l'ensemble des mesures de probabilité sur $\prod_{k=1}^{K} \{1, \ldots, N_k\}$, il existe une sous-suite $(\widehat{\pi}_{n_k})_{k \leq 0}$ et une probabilité π^*, tels que $lim_{n\to\infty} \widehat{\pi}_{n_k} = \pi^* \notin \mathcal{C}$. Par définition d'un équilibre corrélé, il existe un joueur $k \in \{1, \ldots, K\}$ et des actions $j, j' \in \{1, \ldots, N_k\}$ tels que

$$\sum_{i \in \{i_k = j\}} \pi^*(\boldsymbol{i}) \big(\ell^{(k)}(\boldsymbol{i}) - \ell^{(k)}(\boldsymbol{i}^-, j') \big) > 0$$

ce qui est absurde.

□

3.3 Calibrage

On cherche un critère différent du regret pour évaluer la qualité de la prédiction. Dans cette section, on va développer le concept de calibrage pour une stratégie de prédiction, puis montrer comment cette notion intervient dans la convergence des fréquences empiriques vers l'ensemble des équilibres corrélés. On considère ici que $\mathcal{Y}$ est un ensemble fini, de cardinal M et $\mathcal{X}$ est l'ensemble des probabilités sur $\mathcal{Y}$.

Définition 3.3.1 (ε-calibré, bien calibré)**.** *Posons, pour A partie de $\mathcal{X}$,*

$$A_\varepsilon = \{ x \in A \mid \exists y \in A \quad \|x - y\| < \varepsilon \} \qquad \textit{et}$$

$$\rho^\varepsilon_n(A) = \begin{cases} \frac{\sum_{t=1}^{n} y_t \mathbb{I}_{q_t \in A_\varepsilon}}{\sum_{t=1}^{n} \mathbb{I}_{q_t \in A_\varepsilon}} & \textit{si } \sum_{t=1}^{n} \mathbb{I}_{q_t \in A_\varepsilon} > 0 \\ 0 & \textit{sinon.} \end{cases}$$

On dit que la prédiction est ε-calibrée si

$$\limsup_{n\to\infty} \left| \rho^\varepsilon_n(A) - \frac{\int_{A_\varepsilon} x dx}{\lambda(A_\varepsilon)} \right| < \varepsilon$$

où λ désigne la mesure de Lebesgue. On dit que la prédiction est bien calibrée si elle est ε-calibrée pour tout $\varepsilon > 0$

Le calibrage est un moyen naturel pour évaluer une prédiction : si l'on trace, à chaque tour, le point ayant pour abscisse la valeur de la prédiction et pour ordonnée, la valeur du résultat. On veut que ce nuage de points tende à se concentrer sur la droite d'équation $y = x$.

Remarque : Si les joueurs sont autorisés à tirer au sort, une stratégie bien calibrée existe toujours. C'est en fait un corollaire non trivial de l'existence d'une stratégie de minimisation du regret interne, que nous ne prouverons pas ici.

On considère jeu à deux personnes dans lequel chaque joueur tente de prédire l'action de l'autre selon une méthode bien calibrée. On va maintenant démontrer que la suite des fréquences empiriques des profils d'actions converge vers l'ensemble des équilibres corrélés.

Théorème 3.3.1. *On considère un jeu à deux personnes, et on suppose que chaque joueur suit une stratégie bien calibrée de prédiction des actions de son adversaire. Alors, la fréquence empirique des profils d'actions définie par*

$$\widehat{\pi}_n(i,j) = \frac{1}{n}\sum_{t=1}^{n} \mathbb{I}_{\{(I_t,J_t)=(i,j)\}}$$

vérifie :

$$dist(\widehat{\pi}_n, \mathcal{C}) \xrightarrow[n\to\infty]{} 0 \quad (\widehat{\pi}_n \to \mathcal{C}) \ ,$$

c'est à dire :

$$\inf_{\pi\in\mathcal{C}} \sum_{(i,j)\in\{1,\ldots,N\}\times\{1,\ldots,M\}} |\pi(i,j) - \widehat{\pi}_n| \xrightarrow[n\to\infty]{} 0 \qquad \text{p. s}$$

Démonstration. Comme l'ensemble des mesures de probabilité sur un ensemble fini est compact, il suffit de voir que toutes les valeurs d'adhérence de la suite $(\widehat{\pi}_n)_{n\leq 0}$ sont dans $\mathcal{C}$. Soit $(\widehat{\pi}_{n_k})_{k\leq 0}$ une telle sous-suite, et notons π sa limite. Pour tout $i \in \{1,\ldots,N\}$, notons $B_i = \{\boldsymbol{q} \mid \overline{\ell}^{(1)}(i,\boldsymbol{q}) = \min_{i'\in\{1,\ldots,N\}} \overline{\ell}^{(1)}(i,\boldsymbol{q})\}$ et $\widehat{B}_i = \{\boldsymbol{q}$ tels que le joueur ligne joue i si $\boldsymbol{q} = \boldsymbol{q}_t\}$. Alors, π est dans $\mathcal{C}$ si et seulement si :

$$\begin{aligned}\forall i \in \{1,\ldots,N\}, \boldsymbol{q}(.|i) &= (q(1|i),\ldots,q(M|i)) \\ &= \left(\frac{\pi(i,1)}{\sum_{j'=1}^{M}\pi(i,j')},\ldots,\frac{\pi(i,1)}{\sum_{j'=1}^{M}\pi(i,j')}\right) \in B_i \ .\end{aligned}$$

Pour cela, on remarque que, en notant pour tout $i \in \{1,\ldots,N\}$ et $n \leq 0$,

$$\boldsymbol{q}_n(.|i) = \left(\frac{\pi_n(i,1)}{\sum_{j'=1}^{M}\pi_n(i,j')},\ldots,\frac{\pi_n(i,1)}{\sum_{j'=1}^{M}\pi_n(i,j')}\right) \ ,$$

on a :

$$\boldsymbol{q}_n(.|i) = \rho_{n_k}(\widehat{B}_i) \ .$$

Or $(\widehat{\pi}_{n_k})_{k\leq 0}$ est convergente. Soit $\overline{x}$ sa limite. En utilisant les notations de la définition 3.3.1, prenons $\varepsilon > 0$, et, par définition d'une stratégie bien calibrée, il vient que :

$$\limsup_{k\to\infty} |\rho^{\varepsilon}_{n_k}(\widehat{B_i}) - \frac{\int_{(\widehat{B}_i)_\varepsilon} x dx}{\lambda((\widehat{B}_i)_\varepsilon)}| < \varepsilon \ .$$

En outre, par continuité de la mesure de Lebesgue, $\frac{\int_{(\widehat{B}_i)_\varepsilon} xdx}{\lambda((\widehat{B}_i)_\varepsilon)}$ converge quand $\varepsilon > 0$ tend vers 0. Notons $\overline{x}'$ sa limite. Or, clairement,

$$\forall k \leq 0, \lim_{k\to\infty} \rho_{n_k}((\widehat{B}_i)_\varepsilon) = \rho_{n_k}(\widehat{B}_i) \ .$$

Il s'ensuit donc que $|\overline{x} - \overline{x}'| < \varepsilon$ pour tout $\varepsilon > 0$ et donc $\overline{x} = \overline{x}'$. Or $\widehat{B}_i \subset B_i$ qui est convexe, et donc $(\widehat{B}_i)_\varepsilon \subset (B_i)_\varepsilon$, convexe lui aussi. On en déduit que :

$$\frac{\int_{(\widehat{B}_i)_\varepsilon} xdx}{\lambda((\widehat{B}_i)_\varepsilon)} \in (\widehat{B}_i)_\varepsilon \ .$$

Finalement, $\overline{x} \in B_i$. Cela montre que $dist(\pi_{n_k}, \mathcal{C}) \xrightarrow[n\to\infty]{} 0$. □

Appendice

Inégalité d'Hoeffding

Lemme. *Soit $V_1, V_2, \ldots$ une suite d'acroissements de martingale par rapport à une filtration $\mathcal{F}_1, \mathcal{F}_2, \ldots$, telle que $V_t \in [A_t, A_t + c_t]$ pour une variable aléatoire A_t mesurable par rapport à $\mathcal{F}_{t-1}$ et une constante positive c_t. On pose $S_n = \sum_{t=1}^n V_t$, alors pour $s > 0$,*

$$\ln \mathbb{E}\left[e^{sS_n}\right] \leq \frac{s^2}{8} \sum_{t=1}^n c_t^2 \ .$$

Inégalité d'Hoeffding-Azuma

Lemme. *Soit $V_1, V_2, \ldots$ une suite d'acroissements de martingale par rapport à une filtration $\mathcal{F}_1, \mathcal{F}_2, \ldots$, telle que $V_t \in [A_t, A_t + c_t]$ pour une variable aléatoire A_t mesurable par rapport à $\mathcal{F}_{t-1}$ et une constante positive c_t. On pose $S_n = \sum_{t=1}^n V_t$, alors, pour tout $x > 0$,*

$$\mathbb{P}[S_n > x] \leq \exp\left(\frac{-2x^2}{\sum_{t=1}^n c_t^2}\right)$$

et

$$\mathbb{P}[S_n < -x] \leq \exp\left(\frac{-2x^2}{\sum_{t=1}^n c_t^2}\right) \ .$$