

Théorème d'approchabilité de Blackwell (1956)

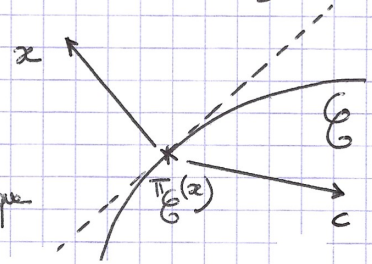
= CVS (+ algorithme associé) pour qu'un problème séquentiel admette une solution

Deux utilisations récentes au LJO $\rightarrow$ Apprentissage séquentiel équilibré (Christophe + Éric + Gilles)
(et jamais 2 sans 3?) $\rightarrow$ Apprentissage conforme pour séries temporelles (Guillaume + Gilles)

I. Deux briques pour commencer

1) Projection euclidienne sur un convexe fermé E
(notée Π_E)

L'angle entre les vecteurs $x - \Pi_E(x)$ et $c - \Pi_E(x)$ est obtus lorsque $x \notin E$
 $\forall x \notin E, \quad \langle x - \Pi_E(x), c - \Pi_E(x) \rangle \leq 0$



2) Théorème minmax de von Neumann: A, B finis, matrice de pertes $(L_{ab})_{a \in A, b \in B}$,
 p et q ds probs sur A et B

$$\text{On a: } \min_p \max_q p^T L q = \inf_p \sup_q p^T L q = \sup_q \inf_p p^T L q = \max_q \min_p p^T L q$$

Preuve: $\inf$ et $\sup$ atteints par ds arguments de continuité et compacité, comme $p \mapsto \max_q p^T L q$ est Lipschitz

L'inégalité $\geq$ est directe (vaut pour toute fonction $f(p, q)$ même non régulière)

Pour $\leq$: couverture de $\mathcal{P}(A)$ par des ε -bails de centres $p_1, \dots, p_{|\mathcal{A}|}$, ie

$$\forall p, q \exists j \mid |p^T L q - p_j^T L q| \leq \varepsilon \|L\|_1$$

puis $\varepsilon \|L\|_1 + \max_q \min_p p^T L q \geq \max_q \min_j p_j^T L q = \min_j \max_q p_j^T L q \geq \min_j \max_q p_j^T L q \geq \max_q \min_p p^T L q$
cf. min fini et max atteint

II. Cadre et énoncé du théorème de Blackwell

Toujours ensembles d'actions A et B finis pour les deux joueurs et fonction $w: A \times B \rightarrow \mathbb{R}^d$ de paiements vectoriels pour le joueur 1, connue
Par abus de notation, $w(p, q) = \sum_a \sum_b p_a q_b w(a, b)$ extension bilinéaire

Protocole: à chaque tour $t \geq 1$,

- les joueurs choisissent simultanément $a_t \in A$ et $b_t \in B$, éventuellement de manière randomisée selon p_t et q_t
- le joueur 1 reçoit et observe $w(a_t, b_t)$, ie, monitoring bandits pour lui, tandis que le joueur 2 peut observer a_t, p_t et w par exemple

Objectifs: $\rightarrow$ pour joueur 1, approcher E ie, assurer que $\bar{m}_t = \frac{1}{t} \sum_{s=1}^t w(a_s, b_s) \rightarrow E$ ps
i.e., $\text{dist}(\bar{m}_t, E) = \inf_{c \in E} \|\bar{m}_t - c\| \rightarrow 0$ ps

$\rightarrow$ pour joueur 2 empêcher cette convergence (et on verra qu'en fait, le joueur 2 peut approcher le complémentaire d'un voisinage de E si le joueur 1 ne peut pas approcher E)

Théorème: CNS d'approchabilité: $\forall q, \exists p \mid m(p, q) \in \mathcal{G}$ pour $\mathcal{G}$ convexe fermé

Algorithme associé: pour tout t tel que $\bar{m}_t \notin \mathcal{G}$, jouer p_{t+1} (il en existe) tel que

$$\max_q \langle \bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t), m(p_{t+1}, q) - \pi_{\mathcal{G}}(\bar{m}_t) \rangle \leq 0$$
 (et sinon, p_{t+1} quelconque, y compris pour p_1)

Interprétation: Ce qu'on peut réaliser avec des coups successifs en actions mixtes dans le jeu sur un tour, on peut le réaliser avec des coups simultanés dans le jeu répété (i.e. sans que le joueur 1 connaisse la stratégie du joueur 2)

Remarque: En général, l'algorithme de Blackwell n'obtient pas la meilleure réponse de convergence possible ni n'est la stratégie la plus efficace computationnellement (cf. calcul de la projection).

III. Preuve du théorème

1) Remarque préliminaire: Hoeffding-Azuma + Borel-Cantelli impliquent que

$$\lim \left| \frac{1}{t} \sum_{\tau=1}^t m(a_{\tau}, b_{\tau}) - \frac{1}{t} \sum_{\tau=1}^t m(p_{\tau}, q_{\tau}) \right| = 0 \text{ ps}$$

2) CN: supposons $\exists q^* \mid \forall p \ m(p, q^*) \notin \mathcal{G}$ i.e. ($\mathcal{G}$ fermé) $\exists \gamma > 0 \ \forall p, \frac{m(p, q^*)}{m(p, q)} \in \overline{\mathcal{G}^{\gamma}}$
 $\overline{\mathcal{G}^{\gamma}}$ = complémentaire du γ -voisinage de $\mathcal{G}$

Lorsque joueur 2 joue q^* à tous les tours: $\forall t, \frac{1}{t} \sum_{\tau=1}^t m(p_{\tau}, q^*) \in \overline{\mathcal{G}^{\gamma}}$
 Non seulement joueur 1 n'approche pas $\mathcal{G}$, mais c'est joueur 2 qui approche $\overline{\mathcal{G}^{\gamma}}$ et donc $\frac{1}{t} \sum_{\tau=1}^t m(a_{\tau}, b_{\tau}) \rightarrow \overline{\mathcal{G}^{\gamma}}$ ps

3) CS: Voyons d'abord l'existence de p_{t+1} lorsque $\bar{m}_t \notin \mathcal{G}$:

$$\min_p \max_q \langle \bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t), m(p, q) - \pi_{\mathcal{G}}(\bar{m}_t) \rangle = \max_q \min_p \langle \bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t), m(p, q) - \pi_{\mathcal{G}}(\bar{m}_t) \rangle$$

donc c'est ≤ 0 et $\forall q \exists p \mid m(p, q) \in \mathcal{G}$
donc que $\langle \cdot, \cdot \rangle \leq 0$

Puis, récurrence sur d_t^2 avec $d_t = \text{dist}(\bar{m}_t, \mathcal{G}) = \|\bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t)\|$

$$\begin{aligned} d_{t+1}^2 &= \|\bar{m}_{t+1} - \pi_{\mathcal{G}}(\bar{m}_{t+1})\|^2 \leq \|\bar{m}_{t+1} - \pi_{\mathcal{G}}(\bar{m}_t)\|^2 = \|\bar{m}_{t+1} - \bar{m}_t + \bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t)\|^2 \\ &= \underbrace{\|\bar{m}_{t+1} - \bar{m}_t\|^2}_{= d_t^2} + \underbrace{\|\bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t)\|^2}_{= d_t^2} + 2 \underbrace{\langle \bar{m}_{t+1} - \bar{m}_t, \bar{m}_t - \pi_{\mathcal{G}}(\bar{m}_t) \rangle}_{\text{voir décomposition page suivante}} \end{aligned}$$

$$\begin{aligned} \bar{m}_{t+1} - \bar{m}_t &= \frac{1}{t(t+1)} \left(t m(a_{t+1}, b_{t+1}) - \sum_{\tau=1}^t m(a_{\tau}, b_{\tau}) \right) \\ &= \frac{1}{t+1} \left(m(a_{t+1}, b_{t+1}) - \bar{m}_t \right) \text{ et de norme au carré} \leq 4 \|m\|_{\infty}^2 / (t+1)^2 \end{aligned}$$

$$\begin{aligned}
2 \langle \bar{m}_{t+1} - \bar{m}_t, \bar{m}_t - \Pi_G(\bar{m}_t) \rangle &= \frac{2}{t+1} \langle m(a_{t+1} b_{t+1}) - \bar{m}_t, \bar{m}_t - \Pi_G(\bar{m}_t) \rangle \\
&= \frac{2}{t+1} \left(\langle m(a_{t+1} b_{t+1}) - m(p_{t+1}, b_{t+1}), \bar{m}_t - \Pi_G(\bar{m}_t) \rangle \quad E[\cdot | \mathcal{F}_t] = 0 \\
&\quad + \langle m(p_{t+1}, b_{t+1}) - \Pi_G(\bar{m}_t), \bar{m}_t - \Pi_G(\bar{m}_t) \rangle \leq 0 \quad \text{si } \bar{m}_t \notin G \\
&\quad + \langle \Pi_G(\bar{m}_t) - \bar{m}_t, \bar{m}_t - \Pi_G(\bar{m}_t) \rangle \right) = -d_t^2 \quad \text{par projection convexe}
\end{aligned}$$

pour la filtration $\mathcal{F}_t = \sigma((a_s b_s)_{s \leq t} \text{ et } b_{t+1})$: $b_{t+1}, p_{t+1}, d_t, \bar{m}_t$ sont $\mathcal{F}_t$ -mesurables
conditionnées par rapport à $\mathcal{F}_t$, c'est intégrer par rapport à $a_{t+1} \sim p_{t+1}$

3.1) Cas simple de la convergence
 ℓ^2 donc en proba:

On a donc prouvé que si $\bar{m}_t \notin G$

$$E[d_{t+1}^2 | \mathcal{F}_t] \leq \frac{4 \|m\|_{\infty}^2}{(t+1)^2} + \left(1 - \frac{2}{t+1}\right) d_t^2 \quad (**)$$

et si $\bar{m}_t \in G$,

$$d_{t+1}^2 \leq \text{directement} \|\bar{m}_{t+1} - \bar{m}_t\|^2 \leq \frac{4 \|m\|_{\infty}^2}{(t+1)^2} \quad \text{d'où la même borne}$$

En particulier, $u_{t+1} \triangleq E[d_{t+1}^2] \leq \frac{4 \|m\|_{\infty}^2}{(t+1)^2} + \left(1 - \frac{2}{t+1}\right) u_t$ et par récurrence,
 $u_t = E[d_t^2] \leq \frac{4 \|m\|_{\infty}^2}{t}$

cf. ok $t=1$
et $\frac{1}{(t+1)^2} + \left(1 - \frac{2}{t+1}\right) \frac{1}{t} = \frac{1}{t(t+1)^2} (t + (t+1)^2 - 2(t+1)) \leq \frac{1}{t+1}$
 $= t^2 + t - 1 \leq t(t+1)$

3.2) Pour ceux que ça intéresse: prouve que $d_t \rightarrow 0$ p.s. avec même une vitesse uniforme

$$Z_t = d_t^2 + \sum_{G \geq t+1}^{\infty} \frac{1}{G^2} E[\|m(a_{G+1} b_{G+1}) - \bar{m}_G\|^2 | \mathcal{F}_t] \geq d_t^2$$

Seule limite possible (cf $d_t \xrightarrow{P} 0$ et reste série convergente): 0
Il suffit donc de mentionner que Z_t converge et pour cela, montrons que c'est une surmartingale (positive)

$$\begin{aligned}
E[Z_{t+1} | \mathcal{F}_t] &= E[d_{t+1}^2 | \mathcal{F}_t] + \sum_{G \geq t+2} \frac{1}{G^2} E[E[\cdot | \mathcal{F}_{t+1}] | \mathcal{F}_t] \leq Z_t \\
&\leq d_t^2 + \frac{1}{(t+1)^2} E[\|m(a_{t+1} b_{t+1}) - \bar{m}_t\|^2 | \mathcal{F}_t] = E[\cdot | \mathcal{F}_t] \\
&\quad \text{avec } (**) \text{ en conservant l'expression complète du premier terme}
\end{aligned}$$

Inégalité maximale de Doob
pour une vitesse uniforme:

$$P\left(\sup_{G \geq T} Z_G \geq \varepsilon\right) \leq E[Z_T] / \varepsilon$$

qu'on applique ainsi:

$$E[Z_T] \leq E[d_T^2] + \sum_{G \geq T+1} \frac{4\|m\|_\infty^2}{G^2} \leq \left(\frac{1}{T} + \int_T^{+\infty} \frac{1}{G^2} dG \right) \times 4\|m\|_\infty^2$$

puis $P(\sup_{G \geq T} d_G \leq \sqrt{\varepsilon}) \leq P(\sup_{G \geq T} Z_G \geq \varepsilon) = 8\|m\|_\infty^2 / T\varepsilon$

ce qui conclut à la convergence par lemmes de Borel-Cantelli $\limsup d_G / \sqrt{T \ln T} \leq \text{constante} \sqrt{\delta T}$

Remarque sur la preuve:

(Mais on perd a priori la C.N.)

Elle montre que la CS s'étend au cas où les fonctions de paiement m_t changent au cours du temps en étant révélées au début du tour, et en étant uniformément bornées.

Et pour le regret externe classique, ou la stratégie induite, pas besoin de connaître m_t au début du tour

IV. Exemples d'application

Ex1: Regret (externe classique)

Fonction de perte $\ell: A \times B \rightarrow [m, M]$ pour le joueur 1

$$R_T = \sum_{t=1}^T \ell(a_t, b_t) - \min_{i \in A} \sum_{t=1}^T \ell(i, b_t) = \max_{i \in A} \sum_{t=1}^T (\ell(a_t, b_t) - \ell(i, b_t))$$

Objectif: $\limsup R_T \leq 0$ p.s. i.e., approcher $\mathcal{E} = (-\infty, 0]^A$ pour $m(a, b) = (\ell(a, b) - \ell(i, b))_{i \in A}$

CNS Blackwell: $\forall q$, prendre p sur $\arg \min_{i \in A} \ell(i, q)$

Stratégie associée: $P_{t+1} \propto q$. $P_{t+1} = (1/|A|, \dots, 1/|A|)$ si $\bar{m}_t \in \mathcal{E}$, i.e., $\forall i, \bar{m}_{t,i} \leq 0$
et sinon, $P_{t+1,j} = (\bar{m}_{t,j})^+ / \sum_{i \in A} (\bar{m}_{t,i})^+$ poids polynomiaux sur le regret avec $p=2$ (cf. Gha-Bianchi-Lugosi 2003)

En effet,

$$\begin{aligned} \max_q \langle \bar{m}_t - \Pi_{\mathcal{E}}(\bar{m}_t), m(P_{t+1}, q) - \Pi_{\mathcal{E}}(\bar{m}_t) \rangle \\ = \langle (\bar{m}_t)^+, m(P_{t+1}, q) - (\bar{m}_t)^- \rangle = \left(\sum_{i \in A} (\bar{m}_{t,i})^+ \right) \times \underbrace{\langle P_{t+1}, m(P_{t+1}, q) \rangle}_{=0} \\ \text{cf. } \forall p, q, \langle p, m(p, q) \rangle = \sum_i p_i (\ell(p, q) - \ell(i, q)) = 0 \end{aligned}$$

Ex2: ε -calibration binaire $B = \{0, 1\}$ et $A = \{0, 1, 2, \dots, L/\varepsilon\}$

À chaque tour t , joueur 1 joue $k_t \in A$, ce qui donne une loi proba $\text{Ber}(k_t \varepsilon)$ sur la réalisation b_t choisie par joueur 2

Objectif: qu'en moyenne, les prévisions du joueur 1 soient bien calibrées:

soit, pour $k \in A$: $N_k(T) = \sum_{t=1}^T \mathbb{1}_{\{k_t = k\}}$

et $\hat{q}_T(k) = \begin{cases} \text{arbitraire si } N_k(T) = 0 \\ \frac{1}{N_k(T)} \sum_{t=1}^T \mathbb{1}_{\{k_t = k\}} \mathbb{1}_{\{b_t = 1\}} \end{cases}$ sinon

la fréquence empirique de 1 de joueur 2 lorsque joueur 1 prévoit $k\varepsilon$

On veut que $\hat{q}_T(k)$ et $k\varepsilon$ soient proches $\forall k$:

$\text{Cal}_T = \sum_{k=0}^{L/\varepsilon} \frac{N_k(T)}{T} |\hat{q}_T(k) - k\varepsilon|$ et désir que $\limsup \text{Cal}_T \leq \varepsilon$ ps

$$= \sum_{k=0}^{L/\varepsilon} \left| \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{\{k_t = k\}} (k\varepsilon - \mathbb{1}_{\{b_t = 1\}}) \right|$$

ie, approcher

(au b_t tout simplement)

$\mathcal{G} = \{ (u_k)_{k=0, \dots, L/\varepsilon} : \|u\|_1 \leq \varepsilon \}$ convexe fermé

pour $u(k, b) = (0, 0, \dots, 0, k\varepsilon - b, 0, \dots, 0)$
 ↑ position 0 ↑ position k ↑ position L/ε

CNS Blackwell: $\forall q \in \mathcal{G} \quad \exists k \in A \quad |k\varepsilon - q(1)| \leq \varepsilon/2$ ie $u(k, q) \in \mathcal{G}$
 ↑ b sur $\{0, 1\}$ ↑ k sera un k sur A