Vitesse minimax du regret interne en prédiction de suites individuelles

Sebastien Gerchinovitz

Vitesse minimax du regret interne en prédiction de suites individuelles

Sébastien Gerchinovitz

DMA, Ecole Normale Supérieure, 45 rue d'Ulm, 75005 Paris.

sebastien.gerchinovitz@ens.fr

Résumé Le problème de la prédiction séquentielle avec avis d'experts consiste à prédire tour après tour les valeurs d'une certaine suite (températures, pics d'ozone journaliers), à l'aide de prédictions de base qu'on peut combiner pour former une seule prédiction. Dans cette communication, nous nous intéressons à une formalisation générique de ce problème de décision séquentielle, et étudions la vitesse minimax d'un critère de performance, le *regret interne*. D'après les travaux de Stoltz (2005), Stoltz et Lugosi (2005), ainsi que Blum et Mansour (2007), cette vitesse est comprise entre $\Omega(\sqrt{n})$ et $\mathcal{O}(\sqrt{n \ln N})$, où n désigne le nombre de tours de prédiction et N le nombre d'actions. Nous montrons que le terme $\sqrt{\ln N}$ est absent dans deux quantités maximin et minimax associées, où les pertes (stochastiques) sont supposées indépendantes et i.i.d. respectivement.

Mots-clés : apprentissage séquentiel, prédiction de suites individuelles, vitesses minimax et maximin.

Abstract Prediction with expert advice consists in forecasting, in a sequential fashion, the values of an unknown sequence (temperatures, daily ozone peaks) from some base forecasts available to the forecaster. In this work, we look at a decision-theoretic formulation of this problem, and study the minimax rate of one of its performance criterion, the *internal regret*. Stoltz (2005), Stoltz and Lugosi (2005), and Blum and Mansour (2007) proved that this rate lies between $\Omega(\sqrt{n})$ and $\mathcal{O}(\sqrt{n \ln N})$, where n is the number of prediction rounds and N the number of actions. We show that the $\sqrt{\ln N}$ term is unnecessary in two related maximin and minimax quantities, where the (stochastic) losses are assumed to be independent and i.i.d. respectively.

Index terms: on-line learning, prediction of individual sequences, maximin and minimax rates.

1 Un modèle générique de décision séquentielle

1.1 Introduction

Nous considérerons le problème de décision (ou prédiction) séquentielle suivant, qui puise sa formalisation dans les travaux de Blackwell et Hannan dans les années 1950, et a récemment été précisé, entre autres, par Freund et Shapire (1997), ainsi que Foster et Vohra (1999).

Un statisticien doit choisir à chaque date $t \in \mathbb{N}^*$ une combinaison convexe $p_t \in \mathcal{X}_N$ de N actions distinctes, où

$$\mathcal{X}_N = \{x \in \mathbb{R}_+^N, \quad \sum_{i=1}^{N} x_i = 1\} \ .$$

A chaque action i correspond une perte $\ell_{i,t} \in [0,1]$, et le statisticien encoure la perte moyenne $\widehat{\ell_t} = p_t \cdot \ell_t = \sum_{i=1}^N p_{i,t}\,\ell_{i,t}$, où $\ell_t = (\ell_{i,t})_{1 \le i \le N}$. Après n tours de prédiction, sa perte cumulée est alors $\sum_{t=1}^n \widehat{\ell_t}$. L'objectif du statisticien est de minimiser cette perte cumulée.

1.2 Formalisation à l'aide d'un jeu répété

Le problème de décision séquentielle précédent peut être décrit sous la forme d'un jeu répété entre le statisticien et l'environnement.

A chaque date $t \in \mathbb{N}^*$,

1. le statisticien choisit et révèle le vecteur de poids $p_t \in \mathcal{X}_N$;
2. l'environnement choisit et révèle le vecteur de pertes $\ell_t = (\ell_{i,t})_i \in [0,1]^N$;
3. chaque action i encoure la perte $\ell_{i,t}$, et le statisticien encoure la perte moyenne $\widehat{\ell_t} = p_t \cdot \ell_t$.

Cette description est commode pour synthétiser les dépendances entre les différentes quantités introduites, même si celles-ci seront évidemment omises. Par exemple, p_t est fonction (mesurable) de tous les vecteurs de pertes passés ℓ_s, $1 \le s \le t-1$.

On appelle alors *stratégie* toute suite de fonctions boréliennes $(p_t)_{t \in \mathbb{N}^*}$, où $p_t : \left([0,1]^N\right)^{t-1} \ \rightarrow \ \mathcal{X}_N$ associe aux vecteurs de pertes $\ell_1, \ldots, \ell_{t-1}$ la prédiction au tour t, encore notée p_t.

1.3 Un critère de performance : le regret interne

Il existe plusieurs critères pour mesurer la qualité d'une stratégie. L'objectif à terme étant de minimiser la perte cumulée $\sum_{t=1}^n \widehat{\ell_t}$ encourue, la stratégie de décision retenue pourra être considérée comme performante lorsque cette perte cumulée est presque aussi petite que la perte cumulée encourue par la meilleure des stratégies d'une classe de référence.

Lorsque cette classe est composée des stratégies constantes $\delta_1, \ldots, \delta_N$, le critère de performance retenu s'appelle *regret externe.* Nous nous intéresserons pour notre part à un autre critère de performance, le *regret interne*, introduit par Foster et Vohra (1999), et qui joue un rôle important en théorie des jeux répétés[1]. Les stratégies composant ici la classe de référence sont des modifications légères de la stratégie maître. Plus précisément, le regret interne sur la suite $(\ell_t)_{1\le t\le n}$ est défini par la différence

$$\sum_{t=1}^{n} p_t \cdot \ell_t - \min_{1\le i\neq j\le N} \sum_{t=1}^{n} p_t^{i\to j} \cdot \ell_t \ , \tag{1}$$

où le poids modifié $p_t^{i\to j}$ est obtenu à partir de p_t selon la transformation définie par

$$(p_t^{i\to j})_k = \begin{cases} p_{k,t} & \text{si } k \notin \{i,j\} \\ 0 & \text{si } k = i \\ p_{i,t} + p_{j,t} & \text{si } k = j \end{cases} \quad , \qquad 1 \le k \le N.$$

Ainsi, le regret interne mesure le regret qu'encoure le statisticien à n'avoir pas choisi l'action j à chaque fois qu'il a choisi l'action i, et ce pour tous les couples (i,j) possibles, $i \neq j$.

Comme les poids $p_t^{i\to j}$ et p_t ne diffèrent qu'en deux coordonnées au plus, beaucoup de termes se simplifient dans la différence (1), et le regret interne prend la forme plus simple

$$\max_{1\le i\neq j\le N} \sum_{t=1}^{n} p_{i,t}(\ell_{i,t} - \ell_{j,t}) \ . \tag{2}$$

2 Regret interne minimax

L'objet de cette section est l'étude du *regret interne minimax*, qui est défini en fonction du nombre n de tours de prédiction et du nombre N d'actions par

$$\overline{\mathrm{RI}}(n,N) = \inf_{S} \sup_{\ell_1,\ldots,\ell_n\in[0,1]^N} \max_{1\le i\neq j\le N} \sum_{t=1}^{n} p_{i,t}(\ell_{i,t} - \ell_{j,t}) \ , \tag{3}$$

où l'infimum est pris sur toutes les stratégies $S = (p_t)_{1\le t\le n}$. Il s'agit donc du meilleur des pires cas qu'on puisse espérer (arbitrairement) atteindre.

[1]Comme l'ont montré Foster et Vohra (1997; 1999), dans un jeu répété (randomisé) fini, si tous les joueurs suivent une stratégie dont le regret interne est asymptotiquement négligeable devant n, leur probabilité empirique (jointe) d'actions converge vers l'ensemble des équilibres corrélés du jeu.

2.1 Résultats existants et question ouverte

Nous nous sommes intéressés à l'ordre de grandeur de $\overline{\mathrm{RI}}(n, N)$ en n et N, dont on connaît pour l'instant un encadrement à un $\sqrt{\ln N}$ près. Plus précisément, Stoltz (2005) a montré qu'il existe deux constantes $C_1, c > 0$ telles que, pour tout $N \geq 1$ et $n \geq cN^2$,

$$C_1 \sqrt{n} \leq \overline{\mathrm{RI}}(n, N) \ .$$

Quant à la borne supérieure, Stoltz et Lugosi (2005) ainsi que Blum et Mansour (2007) ont montré qu'il existe $C_2 > 0$ telle que, pour tous $n \geq 1$ et $N \geq 1$,

$$\overline{\mathrm{RI}}(n, N) \leq C_2 \sqrt{n \ln N} \ .$$

Nous nous sommes donc attachés à la précision de cet encadrement. Nous avons ainsi montré que le $\sqrt{\ln N}$ n'était pas nécessaire pour deux quantités minimax ou maximin qui minorent $\overline{\mathrm{RI}}(n, N)$; la quantité plus génerale $\overline{\mathrm{RI}}(n, N)$ est quant à elle en cours d'étude.

2.2 Dualité maximin/minimax

Comme le font par exemple Cesa-Bianchi et Lugosi (2003), ou bien Abernethy et al. (2009), le regret interne minimax peut être réinterprété comme une quantité (stochastique) maximin. Plus précisément, appliquant pour la deuxième égalité le lemme de Sion, on a

$$\begin{aligned}
\inf_S \sup_{\ell_1,\ldots,\ell_n \in [0,1]^N} \max_{1 \leq i \neq j \leq N} \sum_{t=1}^n p_{i,t}(\ell_{i,t} - \ell_{j,t}) &= \inf_S \sup_{\mathbb{Q} \in \mathcal{M}_1^+([0,1]^{Nn})} E_{\mathbb{Q}} \left[\max_{1 \leq i \neq j \leq N} \sum_{t=1}^n p_{i,t}(\ell_{i,t} - \ell_{j,t}) \right] \\
&= \sup_{\mathbb{Q} \in \mathcal{M}_1^+([0,1]^{Nn})} \inf_S E_{\mathbb{Q}} \left[\max_{1 \leq i \neq j \leq N} \sum_{t=1}^n p_{i,t}(\ell_{i,t} - \ell_{j,t}) \right] .
\end{aligned}$$

Les deux quantités de droite ci-dessus sont donc égales à $\overline{\mathrm{RI}}(n, N)$, et nous avons pour l'instant étudié un minorant de chacune d'entre elles, à savoir

$$\underline{\mathrm{RI}}_{\mathrm{indep}}(n, N) = \sup_{\mathbb{Q} \in \mathcal{Q}_{\mathrm{indep}}(n,N)} \inf_S E_{\mathbb{Q}} \left[\max_{1 \leq i \neq j \leq N} \sum_{t=1}^n p_{i,t}(\ell_{i,t} - \ell_{j,t}) \right] , \tag{4}$$

pour le regret interne maximin indépendant, et, pour le regret interne minimax i.i.d.,

$$\overline{\mathrm{RI}}_{\mathrm{iid}}(n, N) = \inf_S \sup_{\mathbb{Q} \in \mathcal{Q}_{\mathrm{iid}}(n,N)} E_{\mathbb{Q}} \left[\max_{1 \leq i \neq j \leq N} \sum_{t=1}^n p_{i,t}(\ell_{i,t} - \ell_{j,t}) \right] , \tag{5}$$

où les deux parties de $\mathcal{M}_1^+([0,1]^{Nn})$ auxquelles on a restreint le $\sup_{\mathbb{Q}}$ sont définies par

$$\begin{aligned}
\mathcal{Q}_{\mathrm{indep}}(n, N) &= \left\{ \otimes_{t=1}^n \mu_t, \quad \mu_1, \ldots, \mu_n \in \mathcal{M}_1^+([0,1]^N) \right\} \\
\mathcal{Q}_{\mathrm{iid}}(n, N) &= \left\{ \mu^{\otimes n} ; \mu \in \mathcal{M}_1^+([0,1]^N) \right\} \ .
\end{aligned}$$

2.3 Résultats obtenus

La borne inférieure de Stoltz (2005) valant en particulier dans les cas maximin indépendant et minimax i.i.d., les deux théorèmes ci-dessous montrent que les deux quantités $\underline{\mathrm{RI}}_{\mathrm{indep}}(n,N)$ et $\overline{\mathrm{RI}}_{\mathrm{iid}}(n,N)$ ont pour vitesse $\Theta(\sqrt{n})$, sans dépendance en la dimension N du problème donc.

Pour majorer $\underline{\mathrm{RI}}_{\mathrm{indep}}(n,N)$, nous avons considéré une variante d'un algorithme très étudié pour le regret externe (Weighted Majority Algorithm ou Exponentially Weighted Average Forecaster), et dont nous montrons qu'il est maximin optimal au sens du regret interne, pour des pertes indépendantes. Notez qu'il dépend de la probabilité $\mathbb{Q}$ sous-jascente, ce qui est licite dans le cas maximin. La majoration $\mathcal{O}(\sqrt{n})$ est obtenue à l'aide de simples outils de concentration de martingales (comme les inégalités d'Hoeffding-Azuma et de Mc Diarmid).

Théorème 1 (Regret interne maximin indépendant) *Etant donnés $N \geq 1$, $n \geq 1$ et $\mathbb{Q} \in \mathcal{Q}_{\mathrm{indep}}(n,N)$, la stratégie constante (p_t), sans dépendance en les pertes ni en t, et définie par*

$$p_{i,t} = p_{i,1} = \frac{e^{-n^{-1/2}\sum_{s=1}^{n} E_{\mathbb{Q}}[\ell_{i,s}]}}{\sum_{j=1}^{N} e^{-n^{-1/2}\sum_{s=1}^{n} E_{\mathbb{Q}}[\ell_{j,s}]}} \ , \qquad 1 \leq i \leq N, \quad 1 \leq t \leq n \ .$$

vérifie, pour une constante absolue $C_3 > 0$,

$$E_{\mathbb{Q}}\left[\max_{1\leq i,j\leq N} \sum_{t=1}^{n} p_i(\ell_{i,t} - \ell_{j,t})\right] \leq C_3\sqrt{n} \ .$$

Quant au cas minimax, la stratégie considérée ne pourra plus dépendre de la probabilité $\mathbb{Q}$ sous-jascente. La stratégie retenue est la version plug-in de la stratégie précédente, où les espérances sont estimées par leurs moyennes empiriques.

Une stratégie minimax optimale pour des lois i.i.d.
Le temps est partitionné en régimes de taille exponentiellement croissante $\mathcal{T}_r = \{t_r, \ldots, t_{r+1}-1\}$, $r \geq 0$, où $t_0 = 1$, $t_1 = 2$ et, par récurrence, $t_{r+1} = t_r + 2^{r-1}$ pour tout $r \geq 1$. Dès lors, pour tout $r \geq 0$ et tout $t \in \mathcal{T}_r$, le poids choisi $\widehat{p}_t = \widehat{p}^{(r)}$ est défini par

$$\widehat{p}_i^{(r)} = \frac{e^{-\sqrt{t_{r+1}-t_r}\,\hat{m}_i^{(r)}}}{\sum_{j=1}^{N} e^{-\sqrt{t_{r+1}-t_r}\,\hat{m}_j^{(r)}}} \ , \quad 1 \leq i \leq N \ ,$$

où $\widehat{m}_i^{(r)} = \frac{1}{t_r-1}\sum_{t=1}^{t_r-1} \ell_{i,t}$.

Toujours au moyen d'outils simples de concentration de martingales, on montre que les poids de la stratégie précédente approchent avec grande probabilité une variante de la loi de Gibbs $(p_{i,1})_{1 \le i \le N}$ définie dans le cas maximin indépendant, et ce suffisamment rapidement sur tout régime r pour permettre une borne distribution-free en $\sqrt{n}$.

Théorème 2 (Regret interne minimax i.i.d.) *Il existe une constante absolue $C_4 > 0$ telle que, pour tous $N \ge 1$, $n \ge 1$ et $\mathbb{Q} \in \mathcal{Q}_{\mathrm{iid}}(n, N)$, la stratégie $(\widehat{p}_t)$ précédente vérifie*

$$E_{\mathbb{Q}} \left[\max_{1 \le i,j \le N} \sum_{t=1}^{n} \widehat{p}_{i,t} \left(\ell_{i,t} - \ell_{j,t} \right) \right] \le C_4 \sqrt{n} \ .$$

Ce résultat pointe une différence essentielle entre les regrets externe et interne : dans le cas de pertes i.i.d., si plusieurs actions i sont quasi-optimales au sens où elles quasi-minimisent $E_{\mathbb{Q}}[\ell_{i,1}]$, une répartition uniforme de la masse p_t sur ces actions n'aura aucun effet bénéfique sur le regret externe, alors que c'est précisément cette procédure qui permet de supprimer la dépendance du regret interne en le nombre N d'actions.

Bibliographie

[1] Abernethy, J., Agarwal, A., Bartlett, P. and Rakhlin, A. (2009) *A Stochastic View of Optimal Regret through Minimax Duality*, Proceedings of the 22nd Annual Conference on Learning Theory.

[2] Blum, A. and Mansour, Y. (2007) *From External to Internal Regret*, Journal of Machine Learning Research 8, 1307-1324.

[3] Cesa-Bianchi, N. and Lugosi, G. (2006) *Prediction, learning, and games*, Cambridge University Press.

[4] Foster, D. and Vohra, R. (1997) *Calibrated learning and correlated equilibrium*, Games and Economic Behavior, 21:40-55.

[5] Foster, D. and Vohra, R. (1999) *Regret in the on-line decision problem*, Games and Economic Behavior, 29:7-36.

[6] Freund, Y. and Schapire, R.E. (1997) *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of Computer and System Sciences, 55(1):119-139.

[7] Massart, P. (2007) *Concentration inequalities and model selection*, Saint-Flour summer school lecture notes, Springer, New-York.

[8] Stoltz, G. and Lugosi, G. (2005) *Internal regret in on-line portfolio selection*, Proceedings of the 16th Annual Conference on Computational Learning Theory and 7th Kernel Workshop, 403-417, Springer.

[9] Stoltz, G. (2005) *Incomplete information and internal regret in prediction of individual sequences*, PhD dissertation, Université Paris-Sud.