

Information incomplète et regret interne en prédiction de suites individuelles

Gilles Stoltz

Résumé

Le domaine de recherche dans lequel s'inscrit ce travail de thèse est la théorie de la prédiction des suites individuelles. Cette dernière considère les problèmes d'apprentissage séquentiel pour lesquels on ne peut ou ne veut pas modéliser le problème de manière stochastique, et fournit des stratégies de prédiction très robustes. Elle englobe aussi bien des problèmes issus de la communauté du *machine learning* que de celle de la théorie des jeux répétés, et ces derniers sont traités avec des méthodes statistiques, incluant par exemple les techniques de concentration de la mesure ou de l'estimation adaptative. Les résultats obtenus aboutissent, entre autres, à des stratégies de minimisation des regrets externe et interne dans les jeux à information incomplète, notamment les jeux répétés avec signaux. Ces stratégies s'appliquent au problème d'ajustement séquentiel des prix de vente (économie), ou d'allocation séquentielle de bande passante en télécommunications (informatique). Le regret dit interne a également été étudié plus spécifiquement, d'abord dans le cadre de l'investissement séquentiel dans le marché boursier, pour lequel des simulations sur des données historiques sont proposées, puis pour l'apprentissage des équilibres corrélés des jeux infinis à ensembles de stratégies convexes et compacts.

Cette thèse relève des mathématiques, avec pour domaine principal la statistique, pour rattachement secondaire, la théorie des jeux, et a donc des applications en économie. Elle porte sur des développements en théorie de l'apprentissage, plus précisément, en théorie de la prédiction des suites individuelles. L'exposé qui suit est structuré de la sorte :

-
1. **Contexte scientifique et historique**
 - 1.1 Présentation des fondements de la théorie des suites individuelles
 - 1.2 Au carrefour de la théorie des jeux, de l'informatique, de la statistique
 2. **Vue d'ensemble des résultats obtenus**
 - 2.1 Convergence vers les équilibres corrélés dans des jeux convexes et compacts
 - 2.2 Ordre de grandeur optimal du regret en information imparfaite
 3. **Application à venir : Apprentissage en prévision de pics d'ozone**
-

Contexte scientifique et historique

Contexte scientifique : Présentation des fondements de la théorie des suites individuelles

L'objet de la statistique mathématique est la prédiction de phénomènes ; en particulier elle fournit des prévisions de certaines valeurs quantitatives (par exemple, la concentration d'ozone du lendemain) grâce à des observations faites les jours, semaines, mois, années précédents.

Formellement, étant données n observations passées y_1, y_2, \dots, y_n , le statisticien a pour charge de fournir une prévision aussi précise que possible de toute valeur y à venir. On suppose classiquement qu'il existe un mécanisme, que l'on suppose stochastique car il serait souvent trop complexe de le décrire de manière déterministe, générant les y_t . On dit que les y_t sont les réalisations d'un certain processus stochastique sous-jacent ; pour peu que ce dernier soit régulier en un certain sens (stationnaire ou markovien par exemple), il est décrit par un nombre fini de caractéristiques. On en déduit une méthode de prédiction de y : estimer ces caractéristiques grâce aux observations y_1, y_2, \dots, y_n , étudier le processus qui leur est associé (et qui est sûrement proche du vrai processus inobservé), et en déduire une règle de prédiction efficace. L'hypothèse essentielle de cette approche est la génération des observations y_t , mais aussi y , selon un processus stochastique, que certes l'on ne connaît pas, mais qui a le mérite d'exister.

Que se passe-t-il si l'on abandonne ou doit abandonner cette hypothèse ? D'une part, on pourrait n'avoir aucune idée sur la forme d'un éventuel processus générateur. D'autre part, dans un cadre de théorie des jeux, il peut arriver que les observations y_t résultent du comportement d'autres agents, rationnels ou non, agissant pour leur seul compte ou cherchant intentionnellement à nuire au prévisionniste. Dans les deux cas, il s'agit de construire des méthodes de prédiction efficaces sur toutes les suites possibles (déterministes) d'observations ; elles seront alors très robustes.

Eclaircissons le propos en considérant un exemple, appelé ajustement séquentiel des prix de vente. Un vendeur dispose d'un grand nombre d'exemplaires d'un produit ; il ouvre un site Internet pour les proposer aux clients potentiels. Ces derniers se connectent l'un après l'autre au dit site, et le t -ième client se voit proposer le prix p_t . Or, il a en tête un prix maximal y_t qu'il est prêt à payer, mais comme aucun marchandage ni échange d'information n'est possible, ce prix restera inconnu au vendeur. Lorsque $y_t > p_t$, l'achat est réalisé et le vendeur subit une perte (en réalité, un manque à gagner) de $\ell(p_t, y_t) = y_t - p_t$; lorsque $y_t \leq p_t$, le client n'achète pas et la perte du vendeur est une constante c , $\ell(p_t, y_t) = c$, représentant les coûts fixes de gestion et de stockage. L'objectif du vendeur est de fixer ses prix de manière dynamique, de sorte que sa perte cumulée totale soit proche de celle du meilleur prix constant q ; formellement, on prouve dans la thèse que le vendeur a une stratégie telle que quelque soit la suite y_1, y_2, \dots , son regret croisse,

presque sûrement¹, plus lentement que le nombre n de clients,

$$\text{regret} = \sum_{t=1}^n \ell(p_t, y_t) - \min_{q \in \mathbb{R}^+} \sum_{t=1}^n \ell(q, y_t) \leq \square n^{4/5} \quad \text{p.s.} \quad (1)$$

(où \square est une constante dont on ne donne pas ici la valeur exacte). Ainsi, la différence entre la stratégie dynamique du vendeur et le meilleur prix constant devient négligeable lorsque le nombre n de clients augmente.

Une remarque importante est que le meilleur prix q face à la suite des y_t ne peut être au mieux déterminé que rétrospectivement, une fois les n clients servis; “au mieux”, parce que cela supposerait que les clients soient conscients de leur prix personnel y_t et acceptent de le donner pour étude. Or, on rappelle ici que le vendeur est assujéti à une contrainte séquentielle forte, la détermination de p_t ne repose que sur ce qui s’est passé avec les $t - 1$ premiers clients.

Cependant, on peut arguer qu’une étude de marché bien conduite donnerait le prix optimal q ; il suffirait alors de comparer le prix de l’étude de marché au regret; pour les grandes valeurs de n , il n’y aurait alors pas de différence, tout bien calculé, entre stratégie dynamique et stratégie procédant par étude de marché. Cependant, cela supposerait une certaine propriété de stationnarité dans la génération de la suite des y_t : que les comportements des clients interrogés pour l’étude de marché restent inchangés pour l’éternité. Or, les exemples récents de mise en concurrence dans le domaine de la téléphonie ou du transport aérien montrent qu’il n’en est rien, qu’une étude de marché ne vaut qu’un temps (court), que les prix que les clients ont en tête changent – pire, changent en fonction des prix que les vendeurs fixent. Nous sommes vraiment ici dans un cadre de théorie des jeux, où les actions du vendeur ont une influence sur les prix y_t à venir, de sorte que l’hypothèse de génération, disons stochastique stationnaire, ne peut tenir.

Des méthodes robustes de prédiction sont ainsi nécessaires, et un des deux objets de ma thèse a été de les introduire dans un cadre d’information imparfaite. Ici, le vendeur a peu d’information, en particulier il n’observe pas y_t , mais simplement la position du prix y_t du client par rapport au prix proposé p_t ; cependant, ce retour sur prédiction limité (binaire), correctement exploité, lui sera suffisant.

Avant de développer ce point, présentons rapidement le deuxième objet de la thèse, l’extension de la notion de regret interne. Le regret défini en (1) est appelé regret externe parce que l’on compare dans sa définition la perte cumulée du statisticien à celle de la meilleure action constante, l’ensemble des actions étant défini préalablement, extérieurement, à tout algorithme de prédiction. Le regret interne mesure quant à lui la différence à la perte cumulée de la meilleure modification simple de la stratégie maître. Une telle

¹ Puisque les observations ne sont pas supposées stochastiques, où se cache l’aléa? C’est dans les stratégies du statisticien qu’il se trouve: considérer toutes les suites possibles, c’est comme jouer contre le diable; si on a une stratégie déterministe en tête, il la connaîtra et anticipera nos actions; le seul moyen de le surprendre est de déterminer une probabilité sur les actions, et de s’en remettre à une randomisation auxiliaire pour tirer son action

modification propose à chaque tour, dans l'exemple précédent, $p'_t = \psi(p_t)$ en lieu et place de p_t , où ψ est une fonction réelle issue d'une classe de fonctions \mathcal{F} ; \mathcal{F} pourrait être par exemple l'ensemble des fonctions de seuil (le seuil ν est $\psi(x) = x\mathbb{I}_{x \leq \nu} + \nu\mathbb{I}_{x > \nu}$), on interdirait les prix trop élevés. Formellement, on définit

$$\text{regret interne par rapport à } \mathcal{F} = \sum_{t=1}^n \ell(p_t, y_t) - \min_{\psi \in \mathcal{F}} \sum_{t=1}^n \ell(\psi(p_t), y_t) \quad (2)$$

Cela correspond à nos expériences quotidiennes : chacun évalue ses actions toutes choses égales par ailleurs, la suite y_1, y_2, \dots est fixée (même si, naturellement, agir autrement aurait pu changer le futur) ; une évaluation rétrospective conduit à des sentiments de regret, plus ou moins forts. La définition (2) quantifie ce sentiment humain.

Le vocabulaire fondamental de la théorie des suites individuelles étant posé, il est maintenant possible de la replacer dans un contexte historique et de narrer la genèse de cette thèse.

Contexte historique : Au carrefour des jeux, de l'informatique, de la statistique

Les travaux les plus anciens dans le domaine sont issus de la théorie des jeux et remontent aux années 50 ; ils ont été effectués, indépendamment, par Hannan et Blackwell. En théorie de l'information, le problème considéré est non pas la prédiction, mais la compression de suites individuelles. Les pionniers sont Lempel et Ziv, à la fin des années 80 ; en particulier, ce sont eux les auteurs de la terminologie de "suites individuelles". Leurs travaux ont été approfondis par Feder, Merhav et Gutman, puis Willems, Shtarkov et Tjalkens, dans les années 90. Ces dernières voient la véritable éclosion de la théorie, dans le cadre de la théorie de l'apprentissage (*machine learning*). On peut notamment citer le prédicteur à pondération exponentielle introduit par Vovk d'une part, Littlestone et Warmuth d'autre part. Enfin, Cover a proposé une application de la théorie dans le cadre de l'investissement dans le marché boursier ; d'autres, comme Kivinen et Warmuth, ou Cesa-Bianchi, se sont intéressés au problème, plus fondamentalement statistique, de la régression linéaire séquentielle au sens des moindres carrés.

C'est dans ce contexte d'arrivée à l'âge mûr, au niveau international, de cette théorie que Gábor Lugosi, professeur des universités à Barcelone, a donné un cours dressant l'état de l'art, à l'institut Henri-Poincaré à Paris, dans le cadre d'un semestre international de statistique, en 2001. Ses qualités d'orateur, la nouveauté de cette théorie, non étudiée en France, ses perspectives d'application, m'ont convaincu alors de me placer sous sa direction de thèse et de tenter l'aventure franco-espagnole, durant la période entre septembre 2002 et mai 2005.

Vue d'ensemble des résultats obtenus

Convergence vers les équilibres corrélés dans des jeux convexes et compacts

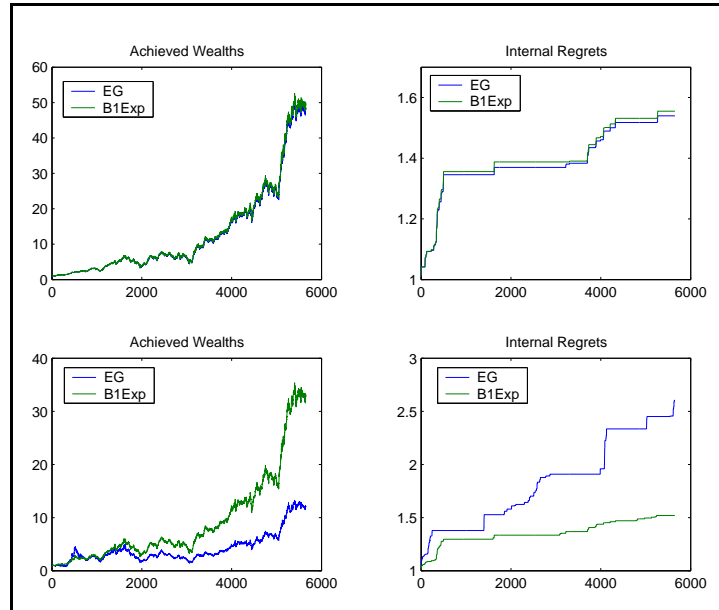


FIG. 1. Evolution des revenus (à gauche) et des regrets internes (à droite) de deux stratégies d'investissement, sur des données réelles de la bourse de New-York. En haut, on a utilisé les paramètres optimaux tels que prescrits par la théorie ; en bas, on a regardé ce qui se passait pour ces mêmes algorithmes lorsqu'ils étaient mal calibrés. EG est un algorithme antérieur à ma thèse, B1EXP est ma proposition.

La première moitié de mes trois ans de thèse a été consacrée à construire des algorithmes, de faible complexité d'implémentation et si possible optimaux en un sens minimax, permettant de contrôler le regret interne dans des jeux où les joueurs ont des ensembles de stratégies convexes et compacts.

C'est par exemple le cas de l'investissement dans le marché boursier : la répartition des capitaux dans N valeurs boursières fixées est modélisée par un élément du simplexe de \mathbb{R}^N . Une stratégie avec un faible regret interne est essentiellement satisfaisante pour le courtier, qui ne verra pas son travail critiqué par ses mandants. Cependant, j'ai pu exhiber des stratégies avec à la fois un faible regret interne et un faible regret externe (par rapport aux meilleures combinaisons linéaires fixes de valeurs boursières). La performance de ces dernières ne peut non seulement pas être améliorée substantiellement par des modifications simples, mais en plus, elles obtiennent également de bons résultats en pratique. (L'extremum local de performance est bien un maximum.)

La figure 1 révèle le sens profond du regret interne dans le cadre du marché boursier : un algorithme d'investissement avec un faible regret interne est plus stable en pratique lorsqu'il est mal calibré ; or, justement, le calibrage est une opération délicate qui demande quelques connaissances sur le marché, notamment des bornes limites d'évolution instantanée (par exemple, aucune valeur boursière ne peut prendre plus de $y\%$ d'un jour à l'autre, ni en perdre plus de $x\%$).

J'ai ensuite étudié² le cas d'un jeu général, dans lequel deux ou plusieurs joueurs ont chacun un ensemble de stratégies convexe et compact. J'ai introduit dans ce cadre la notion de regret interne et ai pu étendre les résultats du cas des jeux finis ; à savoir que si chaque joueur encourt un faible regret interne, alors, en moyenne, les profils de stratégies joués convergent vers l'ensemble des équilibres dits corrélés. Ces derniers ont été introduit par Aumann³ dans deux articles, le premier de 74, le second de 87.

Ordre de grandeur optimal du regret en information imparfaite

La seconde partie de ma thèse a été dévolue à la prédiction en information imparfaite, lorsqu'après avoir formé sa prévision p_t , le statisticien n'a pas accès à l'observation y_t , mais à une version dégradée de cette dernière. L'objectif était d'obtenir les ordres de grandeur optimaux du regret (disons externe) en le nombre n de tours de jeu. Pour un nombre fini d'actions, il est de $n^{2/3}$. La preuve est en deux temps ; la borne supérieure découle d'arguments de concentration de la mesure, elle améliore celle en $n^{3/4}$ obtenue en 2001 par Piccolboni et Schindelhauer. La borne inférieure est obtenue par des techniques de théorie de l'information, notamment l'inégalité de Pinsker. (De cette vitesse $n^{2/3}$ pour un nombre fini d'actions, on déduit la vitesse $n^{4/5}$ proposée dans le cadre de l'ajustement séquentiel des prix de vente, en relançant la stratégie optimale sur une grille de discrétisation de plus en plus fine des prix.)

Les deux parties de la thèse se sont alors rejointes quand j'ai étudié le regret interne dans des cadres de prédiction avec information imparfaite. Cela a été rendu facile par une remarque en apparence anodine : tout algorithme de prédiction avec un faible regret externe peut être converti facilement en un algorithme avec un faible regret interne. En un certain sens, le tout est de pouvoir, dans un cadre donné, minimiser le regret externe.

Perspectives d'application : Méta-apprentissage en prévision de pics d'ozone

Pour l'instant, hormis le cas particulier du marché boursier, les méthodes de suites individuelles n'ont été validées que sur des données artificielles issues de simulations. Or, dans le cas du marché, on peut modéliser les variations boursières, et même, on doit le faire pour augmenter le rendement de ses stratégies. Il ne faut donc pas lancer les stratégies de suites individuelles directement sur les valeurs boursières, mais sur un ensemble donné de stratégies d'investissement stochastiques, dont on pense que chacune obtient de bons résultats mais que l'on ne saurait départager a priori. Alors, les méthodes de suites individuelles permettent de combiner efficacement ces stratégies et d'obtenir des résultats

² Les cinq « bonnes pages » sont les premières pages du chapitre dédié à l'exposé de ce résultat

³ C'est, entre autres, pour cette notion fondamentale d'équilibre qu'il a eu le prix Nobel d'économie il y a quelques mois ; les équilibres corrélés sont des équilibres de Nash d'une extension du jeu originel ; ils expriment une forme de rationalité bayésienne ; leur succès vient de ce qu'ils sont, en général, plus faciles à calculer que les équilibres de Nash

presqu'aussi bons que la meilleure d'entre elles, alors que l'on ne savait pas à l'avance laquelle se révélerait la meilleure.

C'est ce que l'on appelle du méta-apprentissage : on dispose d'un nombre fini de modèles ; on pense que sur toute suite de données, au moins l'un d'entre eux sera efficace, mais ce ne sera pas toujours le même, et surtout, on ne le connaît pas à l'avance. Or les résultats de suites individuelles permettent de ne pas vraiment choisir et de combiner les prévisions, pour *in fine* prédire presqu'aussi bien que ce meilleur modèle inconnu.

En météorologie par exemple, pour prédire les pics d'ozone, on dispose de nombreux modèles, chacun devant spécifier les stations et le maillage utilisés pour recueillir les données passées, de même que plusieurs modèles physiques et chimiques (entre autres, un modèle physique de diffusion verticale, un modèle chimique d'interaction entre les composants de l'atmosphère, etc.). On considère toutes les combinaisons de modèles et maillages possibles ; chaque combinaison forme un méta-modèle, et on combine les prévisions des méta-modèles. Finalement, on évite une phase de calibration ; les méthodes de suites individuelles permettent d'exploiter les connaissances déjà acquises séparément en physique et chimie de l'atmosphère, avant qu'une théorie plus approfondie n'explique quel est le méta-modèle optimal – une sorte de travail en première approximation. Cette ligne de recherche sera mise en chantier bientôt avec le CEREAS, laboratoire mixte ENPC – EDF.