

Note de synthèse - Reinforcement Learning appliqué à la gestion de portefeuille chez Aequam Capital

Dorian LAGADEC

Année scolaire 2018-2019

L'organisme : Aequam Capital

Aequam Capital est une société de gestion quantitative employant une dizaine de personnes et spécialisée dans la gestion multi-facteurs. Le fonds Aequam Dynamic Premia Equity (ADPE), ouvert il y a un an, a une allocation d'actifs pilotée par un algorithme propriétaire.

Dans le cadre de la recherche d'Aequam Capital, il a été décidé d'explorer des pistes jusqu'alors inconnues dans le but de générer une performance supérieure à celle connue actuellement. Une piste envisagée est celle du Reinforcement Learning - ou apprentissage par renforcement - car c'est un cadre théorique qui permet de répondre à des problèmes séquentiels et d'avoir une dynamique double, d'exploration et d'exploitation. C'est la raison pour laquelle Aequam Capital a fait appel à la société Avisia, qui m'a accompagné au cours de mon alternance, pour proposer mon travail : ayant un pied dans le monde académique et de la recherche avec l'ENSAE et un autre dans la finance de marché avec mes précédents stages, j'étais susceptible de les aider dans leur processus de découverte. Cette mission d'exploration a été menée de la manière suivante : j'étais responsable de lire la littérature scientifique, de proposer une démarche de *Proof-Of-Concept* (POC) et de la mettre en oeuvre ; les gérants d'Aequam Capital Thierry Béchu et Pierre Colonna m'apportaient leur éclairage sur des questions spécifiques de finance de marché, de données et de considérations pratiques sur le caractère faisable ou non des allocations algorithmiques proposées ; le chercheur avec lequel travaille Aequam Capital, Florian Berg, était chargé de m'aider sur l'aspect théorique et fonctionnel des algorithmes utilisés.

Un problème précis et une démarche incrémentale

Le problème était le suivant : étant donné 6 facteurs de risques (un facteur de risque étant un panier pondéré d'actions exposées à un même risque donné), est-il possible

de choisir dynamiquement les poids de ces facteurs pour que la valeur finale du portefeuille soit maximisée (sous contrôle de critères de volatilité et de *turnover*), en prenant en compte des coûts de transaction ? Si oui, l'allocation dynamique proposée est-elle interprétable, est-elle vendable ?

Le problème étant relativement ouvert, il a été décidé de commencer par un problème plus simple : l'indice de référence étant un portefeuille équi-pondéré de ces 6 facteurs, est-il possible de choisir des points d'entrée et de sortie optimaux dans ce portefeuille équi-pondéré pour battre ce portefeuille ? Ce problème est plus simple car plus contraint : les poids sont tous égaux à $\frac{1}{6}$ ou à 0.

Etant néophyte dans le domaine du Reinforcement Learning, la démarche a donc été la suivante : familiarisation avec les algorithmes de Reinforcement Learning, puis lecture d'articles qui portent sur l'application du Reinforcement Learning à des problématiques financières, enfin exploration de bibliothèques en ligne mettant en oeuvre lesdits algorithmes. Après cette phase de découverte, j'ai codé en Python mes propres classes pour adapter un algorithme de Reinforcement Learning au problème d'Aequam Capital. A l'aide d'outils de visualisation, j'ai pu observer le comportement de l'algorithme entraîné sur le jeu de données et, à plusieurs reprises, j'ai modifié tel ou tel bout de code pour que l'algorithme modifie son comportement afin que ce dernier s'approche au plus près d'une allocation dynamique "logique", du goût des gérants.

Etat des lieux à date et pistes d'amélioration

Après de nombreux *backtests*, aucune instance de mon algorithme n'a pu significativement faire mieux que l'indice de référence et j'ai tenté d'analyser les causes de cet échec. D'abord, il est possible, de nombreux articles le disent, qu'il soit impossible de prédire les cours boursiers, donc qu'aucun signal ne puisse être détecté dans les données. Si l'on écarte cette hypothèse qui ruine la raison d'être de ce projet, on peut soupçonner les données choisies et/ou leur traitement de ne pas donner lieu à la détection d'un signal fort pour notre algorithme. Autrement, il est possible que l'algorithme que j'ai codé soit largement sous-optimal, c'est pourquoi j'ai décidé de revenir aux algorithmes éprouvés (ceux d'Open AI Gym par exemple) pour voir si le problème venait de mon algorithme ou de mes données. Enfin, la question de la fonction de "récompense", clef dans tout algorithme de Reinforcement Learning car c'est elle qui indique à l'algorithme si chaque décision a été bonne ou non et à quel point, est sensible et il a été décidé d'aborder la question de l'Inverse Reinforcement Learning, une forme de *reverse engineering* du Reinforcement Learning, pour comprendre plus finement les paramètres de l'allocation que recherche réellement Aequam Capital, plus complexe qu'un arbitrage entre rendement et risque.