

AEQUAM Capital

—

Projet Reinforcement Learning

Réunion du 8 juillet 2019

1ers tests : mise en œuvre

- Les codes tournent et aucune erreur n'est apparue
- La complexité calculatoire est présente mais relativement bien mesurée : le code a tourné pendant 54h sur 60h prévues maximum (légère sous-estimation due partiellement à une marge d'erreur volontaire)
- 128 scénarios testés, sur 37 « épisodes », correspondant à :
 - 4 jeux de données (prix, prix + macros, macros de vol, macros de surprise éco)
 - 2 types de signal : signal de changement de portefeuille, signal de position
 - 4 fenêtres de mise à jour du prédicteur (5, 10, 30, 100 jours)
 - 2 métriques de récompense ($a * \text{Return} * \text{sharpe} - b * \text{coûts de transaction}$, $\text{Return} - c * \text{Semivariance} - d * \text{coûts de transaction}$)
 - 2 paramètres pour a et c

1ers tests : résultats

- **Une bonne moitié des résultats ne présente aucun intérêt**, le portefeuille est changé beaucoup trop souvent et le timing de ces changements ne présente aucune forme d' « intelligence »
- Les « **bons** » **résultats** sont une forme déguisée de **buy-and-hold** (quelques arbitrages à la marge, pas nécessairement les meilleurs possibles)
- **Bonne nouvelle** : en cas de position « unique », le portefeuille choisit toujours le marché par rapport au cash ; **mauvaise nouvelle** : très peu de positions cash tenues plusieurs jours consécutifs

1ers tests : leçons pour prochains tests

- Remise à l'échelle du jeu de données pour partir sur une base 1000
- Horizon de mise à jour autour de 20 – 40 jours (plus court, le signal est erratique ; plus long, le signal est figé)
- Pas de leçon particulière sur le choix des données, ni sur la métrique
- Résultats relatifs, car sur 37 épisodes : vrais tests plus longs (200-500)

→ Moins de paramètres, plus resserrés, tests plus longs

En parallèle

- Wiki et partage du code pour valider la démarche et les premiers éléments