

# RLD – Rapport TP 1 : Bandits multi bras

Maxime DARRIN

10 janvier 2020

## 1 Baselines

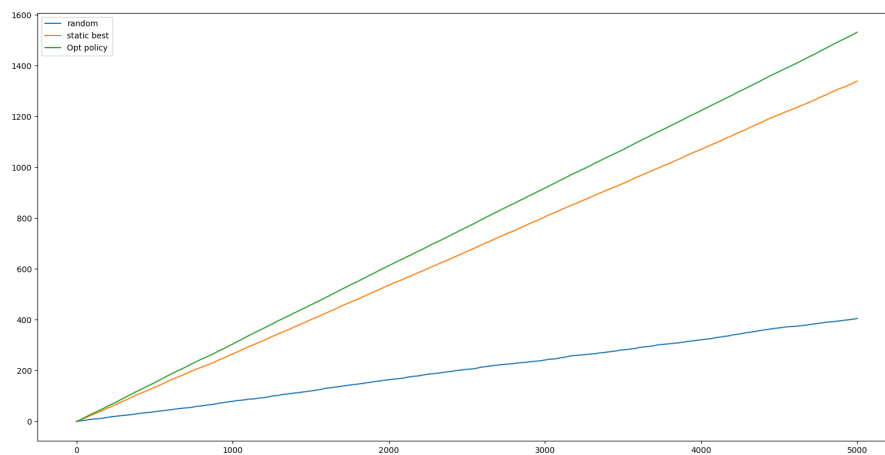


FIGURE 1 – Baselines – agents omniscients

## 2 UCB

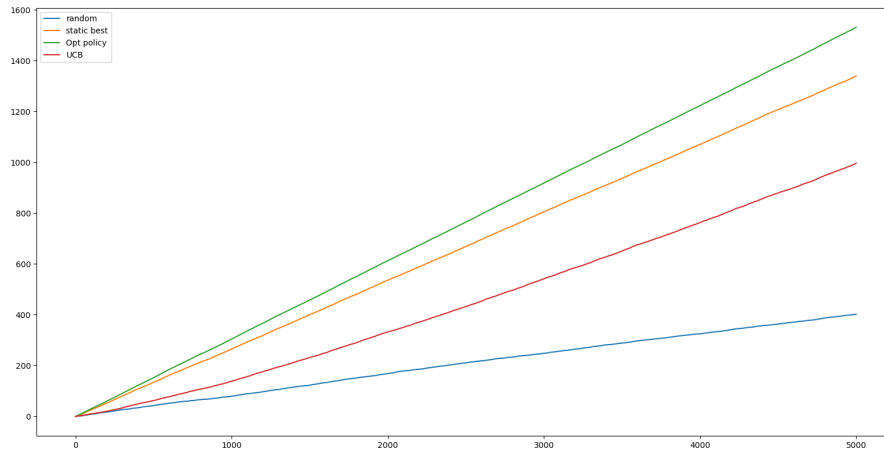


FIGURE 2 – Algorithm UCB VS Baselines

## 3 LinUCB

Dans figure suivante on compare *Lin UCB* aux *baselines* en faisant varier  $\alpha$ . En particulier on vérifie bien qu'un  $\alpha$  bas correspond à une forte exploration (ici  $\alpha = 0.01$  donne des résultats proches de l'aléatoire) tandis que des valeurs plus importantes permettent un meilleur compromis exploration-exploitation.

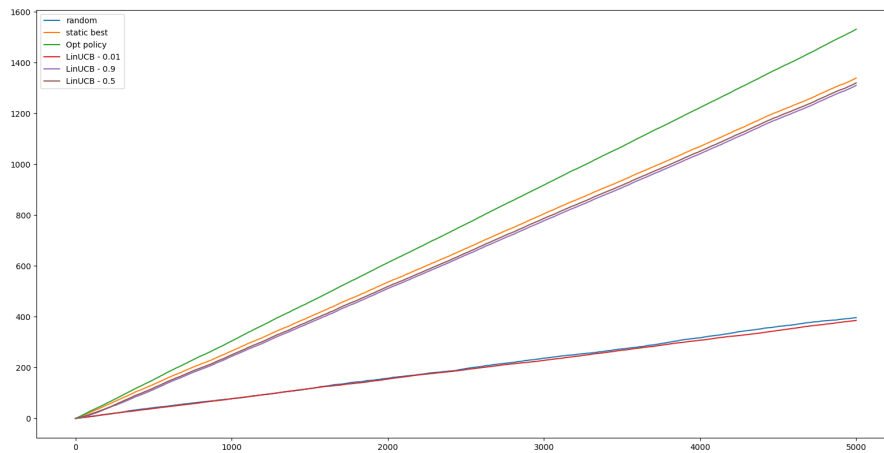


FIGURE 3 – Algorithm LinUCB VS Baselines

Par ailleurs, on observe bien que *lin UCB* est clairement meilleur que *UCB*. En effet, en utilisant le contexte pour prendre des décisions plus averties.

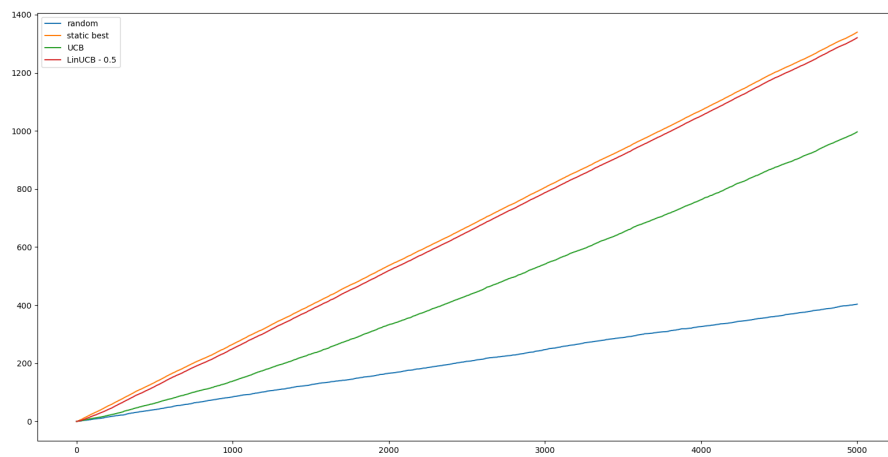


FIGURE 4 – Algorithmme LinUCB VS UCB