
Optimism and randomness in linear multi-armed bandit

Alessandro Lazaric*¹ and Marc Abeille¹

¹SEQUEL (INRIA Lille - Nord Europe) – Université Lille III - Sciences humaines et sociales, Université Lille I - Sciences et technologies, CNRS : UMR8146, CNRS : UMR8022, INRIA, Ecole Centrale de Lille
– France

Abstract

In this talk I will review the linear multi-armed bandit problem and discuss two main algorithmic design approaches. The first one relies on the celebrated principle of "optimism in face of uncertainty", which prescribes to select at each step the action that maximizes the performance when the problem parameters are optimistically chosen within the current level of uncertainty. While the resulting algorithms are effective from a learning point of view, they are computationally expensive as the number of available actions increases. An alternative approach is to use a randomized algorithm, the so-called Thompson sampling, which drastically reduces the computational complexity. On the other hand, this poses theoretical challenges to "control" the negative effects of the stochastic choice of actions. I will indeed show that even in this case it is crucial to preserve a minimum level of optimism to guarantee an order-optimal performance. Finally, I will discuss possible extensions beyond the linear bandit problem.

*Speaker