Title: Conversion theorem and minimax optimality for continuum contextual bandits

Abstract:

We study the continuum contextual bandit problem, where the learner sequentially receives a side information vector (a context) and has to choose an action in a convex set, minimizing a function depending on the context. The goal is to minimize a dynamic (contextual) regret, which provides a stronger guarantee than the standard static regret. Considering a meta-algorithm that to any input non-contextual bandit algorithm associates an output contextual bandit algorithm, we prove a conversion theorem, which allows one to derive a bound on the contextual regret from the static regret of the input algorithm. We apply this strategy to obtain upper bounds on the contextual regret in several settings (losses that are Lipschitz, convex and Lipschitz, strongly convex and smooth with respect to the action variable). Inspired by the interior point method and employing self-concordant barriers, we propose an algorithm achieving a sub-linear contextual regret for strongly convex and smooth functions in noisy setting. We show that it achieves, up to a logarithmic factor, the minimax optimal rate of the contextual regret as a function of the number of queries.

Joint work with Arya Akhavan, Karim Lounici and Massi Pontil.