ID de Contribution: **107** Type: **Contributed talk**

Lagrange Duality Gap of Decomposed Dynamic Programming: The Case of Bandits

mercredi 30 juillet 2025 10:45 (30 minutes)

This study explores **multi-armed bandit** problems under the premise that the decision-maker possesses prior knowledge of the arms' distributions and knows the finite time horizon. These conditions render the problems suitable for **stochastic multistage optimization decomposition techniques**. On the one hand, multi-armed bandit algorithms are integral to reinforcement learning and are supported by extensive theoretical analysis, particularly regarding regret minimization. Meanwhile, stochastic multistage optimization emphasizes examining the **performance gap between** a given policy and the optimal, adapted policy, which can be estimated numerically through dual methods. Within this framework, decomposition strategies have been demonstrated to efficiently tackle large-scale stochastic multistage optimization challenges. A fact in **dire need of better theoretical understanding**. On the empirical side, our experiments corroborate **the approach's efficacy on multi-armed bandit.** On the theoretical side, **we reinterpret the duality gap by relating it to quantities defined along the trajectory of the decision-making process**: the Lagrangian multiplier, the instantaneous reward and the value of information. This lays the groundwork for subsequent investigations into the performance of decomposition methods.

Authors: HEYMANN, Benjamin; Prof. CHANCELIER, Jean-Philippe (ENPC); DE LARA, Michel (Ecole des Ponts ParisTech)

Orateur: HEYMANN, Benjamin

Classification de Session: Sequential decision-making under uncertainty

Classification de thématique: Sequential decision making under uncertainty