BANDIT CONVEX OPTIMIZATION

SÉBASTIEN BUBECK
MICROSOFT RESEARCH

ABSTRACT. The multi-armed bandit and its variants have been around for more than 80 years, with applications ranging from medial trials in the 1930s to ad placement in the 2010s. In this mini-course I will focus on a groundbreaking model introduced in the 1990s which gets rid of the unrealistic i.i.d. assumption that is standard in statistics and learning theory. This paradigm shift leads to exciting new mathematical and algorithmic challenges. I will focus the lectures on the foundational results of this burgeoning field, as well as their connections with classical problems in mathematics such as the geometry of martingales and high-dimensional phenomena.

Tentative schedule:

Lecture 1: Introduction to regret. Game theoretic viewpoint and derivation of the minimax regret via geometry of martingales.

Lecture 2: Introduction to the mirror descent algorithm. Connections with competitive analysis in online computations will also be discussed.

Lecture 3: Bandit Linear Optimization. Two proofs of optimal regret: one via low-rank decomposition in the information theoretic argument, and the other via mirror descent with self-concordant barriers.

Lecture 4: Bandit Convex optimization. Kernel methods for online learning, Bernoulli convolution based kernel, Gaussian approximation of Bernoulli convolutions, and restart type strategies.