Sequential learning, sequential optimisation

Gilles Stoltz

Laboratoire de mathématiques d'Orsay

Note: as I will do in class—lecture notes in English but comments in French

Two models of online learning with much structure → RL not covered

1. Stochastic bandits

Simplest case: K unknown distributions $\nu_1, \ldots, \nu_K$ over $[0,1]$

At each round, pick an arm A_t and get a reward $Y_t \sim \nu_{A_t}$

Exploration – exploitation trade-off

2. Individual sequences – Robust aggregation of predictors

Meta-statistical framework: expert predictions $f_{1,t}, \ldots, f_{K,t}$ are available

Aggregate them into $\widehat{y}_t = \sum_{j \in [K]} p_{j,t} f_{j,t}$

In both cases: control the regret

Computationally efficient algorithms with regret upper bounds

Exhibit lower bounds holding for all algorithms

Focus

Strong emphasis on neat proofs

Challenges ($\to$ bonus points) dedicated to improving current writing of notes

100% theoretical course, no use of a computer

Mathematical tools at hand

No need of a statistics background, but measure/integration theory required

Martingale theory   Hoeffding-Azuma, Doob's max. inequalities [both!]

Information theory   KL divergences, data-processing inequality, Fano's lemma

$\to$ Links with P. Massart's course   but different view on the same objects

Probability students: opportunity to study IA/machine learning!

Practical details

7 or 8 sessions, of 2h30 when not at a distance (shorter otherwise)

One homework due after the winter break

One final exam or a second homework   depending on sanitary situation

Webpage with all lecture notes

http://stoltz.perso.math.cnrs.fr/enseignements.html

You can already go there and register

Will use BBB for the 1st session, may change afterwards