BUSF 40902-50: Online Optimization and Decision Making under Uncertainty

Spring 2018

Syllabus (March 18, 2018)

Instructor: **Varun Gupta**

Office:	331 Harper Center
e-mail:	guptav@uchicago.edu
Phone:	773-702-7315
Office hours:	by appointment

Class Times: Thu – 8:30-11:30 am – Harper Center (C04)

Course Website: Canvas

Course Objectives

In many (almost all?) operations problems, decisions have to be made over time without perfect knowledge of future input. Portfolio are balanced periodically with uncertain stock returns; jobs have to be scheduled on servers without knowledge of future arrivals; power generation has to be committed with imperfect information about demand and solar/wind availability; ads have to be placed on webpages, and items have to be priced without information about demand.

In this course we will look at some of the models and algorithmic frameworks developed in various communities for dealing with problems involving uncertainty. The goals are two-fold:

1. Develop an understanding of the strenghts and limitations of each approach. Decision making under uncertainty is an active area of research, and the hope is that by the end of the course students will gain enough maturity to start contributing to this field.
2. Expose students to the rich set of mathematical techniques used: concentration inequalities, information theory, convex optimization, duality, potential function analysis.

Tentative list of topics:

- Competitive Analysis - Dual fitting and resource augmentation analysis; Primal-dual algorithms. Potential function analysis.
- Online Learning - Prediction with Experts; Online Mirror Descent/Online Linear Optimization. Shifting experts; Strongly Adaptive Regret. Multi-armed Bandits (stochastic - UCB1, adversarial - EXP3, Thompson sampling); bandit linear optimization; contextual bandits.
- Experts via Primal-dual analysis.
- Secretary problems and Random order problems.
- Prophet inequalities; stochastic knapsack.

- Introduction to Reinforcement Learning.
- k-server problem; work-function analysis; Markov k-server problem.

Time permitting and depending on interest, we may touch upon some of the following:

- Multistage optimization with recourse; online network design.
- Information relaxation bounds and links to Approximate Dynamic Programming.
- Models between random and worst-case.
- Markov Decision Processes; Bandits with Markov rewards - Gittins index policies; Restless bandits.
- Prior-free mechanism design.

Prerequisites

Exposure to undergraduate probability (random variables, discrete and continuous probability distributions) and calculus. Linear Progamming.

Required Course Material

There is no required textbook. I will post links to relevant online references and research papers.

Supplemental Texts/Readings

- *Prediction, Learning, and Games* by Nicolò Cesa-Bianchi and Gábor Lugosi. Cambridge University Press. https://homes.di.unimi.it/~cesabian/predbook/
- *Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems* by Sébastien Bubeck and Nicolò Cesa-Bianchi. https://arxiv.org/abs/1204.5721
- *Introduction to Online Convex Optimization* by Elad Hazan. http://ocobook.cs.princeton.edu/OCObook.pdf
- *The Design of Competitive Online Algorithms via a Primal-Dual Approach* by Niv Buchbinder and Joseph (Seffi) Naor. http://www.tau.ac.il/~nivb/download/pd-survey.pdf
- *Online Computation and Competitive Analysis* by Allan Borodin and Ran El-Yaniv. Cambridge University Press. http://www.cs.technion.ac.il/~rani/book.html
- *Statistical Learning and Sequential Prediction* by Sasha Rakhlin and Karthik Sridharan. http://stat.wharton.upenn.edu/~rakhlin/book_draft.pdf (Working book draft)

Grades

- Based on lecture scribes, assignments, a final paper presentation, and class participation.

- **Scribe** : Each student is expected to scribe at least one lecture. The student who scribes lecture n is also required to proofread the scribe of lecture $n-1$ for correctness, and to maintain continuity. I will give my hand-written notes to help.
- This class can not be taken pass/fail or audited.
- Homework is **due at the start** of the lecture on the homework due date.
- **No late homework will be accepted**. If you can not attend the lecture for some reason, you can scan/email or submit the homework in person to me prior to the lecture.
- You may only discuss the homework assignment with others enrolled in the course. You must write and submit your own solution, and declare who you discussed the homework with (if any).

TENTATIVE SCHEDULE - Subject to change

PART 0: Introduction (0.5 lectures)

- Ski Rental, Graham's load balancing heuristic
- Models of adversaries
- Taxonomy of Online Decision Making Models

PART I : Competitive Analysis-1 (1.5 lecture).

- Primal-Dual heuristic for design of online algorithms
- Ski-rental analysis via dual-fitting.
- Online Matching, Online Set Covering
- Dual-fitting and resource augmentation analysis of Online Scheduling on Unrelated machines

PART II: Regret Minimization under Full information (1.5 lectures)

- Randomized Weighted Majority and Multiplicative Weight Update algorithms
- Online Mirror Descent view of MWU; Online Linear Optimization
  - Hannan-Kalai-Vempala algorithms (Follow the Perturbed Leader, Follow the Regularized Leader)
  - Zinkevich's Online Convex Optimization algorithm
- Shifting regret and strongly adaptive regret

PART III: Regret Minimization under Bandit feedback (1.5 lectures)

- Multi-armed Bandits
  - Stochastic rewards – Lai-Robbins; UCB1 algorithm; Thompson sampling
  - Adversarial rewards – EXP3 algorithm
  - Information theoretic lower bound
- Bandit Linear Optimization

PART IV: Secretary Problems and Random Order Models (1 lectures).

- Vanilla secretary problem
- Analysis of secretary problem via LP
- Incentive compatible secretary via LP
- Policy design via LPs for stochastic box selection
- Online edge-weighted matching under random arrival model
- Multiple secretary problem

PART V: Prophet Inequalities (1 lectures).

- Application to online auctions
- Sequential search with exploration cost

PART VI: Introduction to Reinforcement Learning (1.5 lectures).

- On-policy vs. Off-policy RL
- ϵ-GREEDY and REINFORCE

PART VII: Competitive Analysis - 2 (1 lectures).

- Metrical Task System, k-server problems
- Work function algorithms for MTS, k-server
- Markov k-server problem
- k-server using Online Convex Optimization