Course Description

The aim of this course is to cover modern tools for data-driven decision making. Most decision making tasks involve uncertainty that is directly impacted by the amount and complexity of data at hand. Classical decision models rely on strong distributional assumptions about the uncertain events. But in recent years, and due to growing availability of rich data, there has been a rapid adoption of models from machine learning and statistics that provide more accurate and personalized picture of the uncertainty which in turn leads to better decisions. The interplay between the multiple objectives of modeling the data, personalization, and decision optimization has created a number mathematical models that the course aims to cover.

Topics[1]

1. The "Data $\rightarrow$ prediction $\rightarrow$ decision" paradigm.
   Methods: Parametric estimation and regularization
   Reference: Bayati et al. [7].

2. Indirect and direct methods for learning and decision making in presence of covariates
   Methods: Non-parametric estimation, random forest
   References: Ban and Rudin [3], Bertsimas and Kallus [8].

3. Indirect and direct methods for learning and decision making for multi-outcome settings
   Methods: Matrix estimation, low-rank methods
   References: Kao and Van Roy [14].

4. Multi-armed bandit, a toy model for dynamic learning and decision making Methods: ϵ-Greedy, Upper Confidence Bound (UCB), Thompson Sampling
   References: Scott [17], Chapelle and Li [10].

5. Multi-armed bandits with covariates I
   Methods: Linear bandits, UCB and Thompson sampling for linear bandits
   References: Li et al. [15], Abbasi-Yadkori et al. [1], Agrawal and Goyal [2], Russo and Van Roy [16], Hamidi and Bayati [13].

6. Multi-armed bandits with covariates II
   Methods: Two estimator bandits, lasso bandit for high-dimensional problems
   References: Goldenshluger and Zeevi [11], Bastani and Bayati [4].

7. Free exploration in contextual bandits
   Methods: Matrix tail bounds, covariate diversity, Greedy-First algorithm
   References: Bastani et al. [5], Bietti et al. [9].

8. Bayesian learning and decision-making
   Methods: Factor graphs and message-passing algorithms
   References: Graepel and Herbrich [12], Bayati and Montanari [6].

[1] There may be some modifications to this list as the quarter unfolds

Logistics

Time and Location. Thursdays 3 PM - 5:50 PM over zoom (links will be provided on Canvas).

Prerequisites. Knowledge of probability and linear algebra is required. Some assignments may need programming in one language of your preference (e.g. Matlab, Python, R, . . .)

Course Calendar. Contains the schedule and due dates, this link.

Grading Criteria and Policies

1. Class participation (20%)
   - Scribing lecture notes, we provide you with a LaTeX template (make your selections on (course calendar, contains the schedule and due dates))
2. Two Problem set (20%)
3. Projects (50%), you can work with teams of size up to three (we strongly encourage at least two)
   - Proposal, one page (5%)
   - Progress report, two pages (10%)
   - Final presentation (15-20 minutes, depending on the number of projects), during last lecture (15%)
   - Final report (at most 5 pages) (10%)

References

[1] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *NIPS*, pages 2312–2320, 2011.

[2] Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *ICML*, pages 127–135, 2013.

[3] Gah-Yi Ban and Cynthia Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1):90–108, 2019. doi: 10.1287/opre.2018.1757. URL https://doi.org/10.1287/opre.2018.1757.

[4] H. Bastani and M. Bayati. Online decision-making with high-dimensional covariates. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2661896, 2015.

[5] H. Bastani, M. Bayati, and K. Khosravi. Mostly exploration-free algorithms for contextual bandits. https://arxiv.org/abs/1704.09011.

[6] M. Bayati and A. Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Transactions on Information Theory*, 57: 764–785, 2011.

[7] M. Bayati, M. Braverman, M. Gillam, K. Mack, G. Ruiz, M. Smith, and E. Horvitz. Data-driven decisions for reducing readmissions for heart failure: General methodology and case study. *PLOS ONE*, 9, 2014. doi: 10.1371/journal.pone.0109264.

[8] Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. https://arxiv.org/abs/1402.5481, 2014.

[9] Alberto Bietti, Alekh Agarwal, and John Langford. A Contextual Bandit Bake-off. *arXiv e-prints*, art. arXiv:1802.04064, Feb 2018.

[10] Olivier Chapelle and Lihong Li. An empirical evaluation of thompson sampling. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 2249–2257. Curran Associates, Inc., 2011. URL http://papers.nips.cc/paper/4321-an-empirical-evaluation-of-thompson-sampling.pdf.

[11] Alexander Goldenshluger and Assaf Zeevi. A linear response bandit problem. *Stochastic Systems*, 3(1):230–261, 2013.

[12] Thore Graepel and Ralf Herbrich. Large Scale Data Analysis and Modelling in Online Services and Advertising. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 2, Las Vegas, 2008. URL http://www.herbrich.me/papers/kdd2008.pdf.

[13] Nima Hamidi and Mohsen Bayati. A General Theory of the Stochastic Linear Bandit and Its Applications. 2020. doi: 10.48550/ARXIV.2002.05152. URL https://arxiv.org/abs/2002.05152.

[14] Yi-Hao Kao and Benjamin Van Roy. Directed principal component analysis. *Operations Research*, 62(4):957–972, 2014. doi: 10.1287/opre.2014.1290.

[15] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 661–670, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. doi: 10.1145/1772690.1772758. URL http://doi.acm.org/10.1145/1772690.1772758.

[16] Daniel Russo and Benjamin Van Roy. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.

[17] Steven L. Scott. A modern bayesian look at the multi-armed bandit. *Appl. Stoch. Model. Bus. Ind.*, 26(6):639–658, November 2010. ISSN 1524-1904. doi: 10.1002/asmb.874. URL http://dx.doi.org/10.1002/asmb.874.