Advanced Dynamic Programming

Prof. Xie

Summer Semester 2022

Language: English

Occurrence: Summer semester

Scheduling: Kick-off April 28, 2022, 9:00 pm – 10:30 am via Zoom. The course will be taught as a series of seminars every Thursday morning. Dates could be coordinated with participants. If possible, the course will be held in person; otherwise, via Zoom.

Registration: Until April 21, 2022, via Moodle.

Description of Achievement and Assessment Methods: Participants will be assessed based on their seminar presentation (70%) and oral contributions to the course (30%). Each participant will present a topic and lead the discussion. Everyone else should read the relevant literature before class and participate actively in the discussion. Students who show up unprepared and/or unable to contribute actively may be considered absent. The presentation of 90 min to 120 min shall introduce and explain the respective method as well as applications. Presenters will suggest an article in which this method is applied. The course is pass/fail, not graded. In order to pass the course, participants must take part in all classes. In case of excused absence due to illness they need to hand in a written assignment about the content of the class they have missed.

Prerequisites: The seminar requires solid knowledge in advanced mathematics, especially the knowledge of probabilities, Markov chains, and algorithms etc. Mathematical maturity and the ability to write down precise and rigorous arguments and proofs are also important. Computer programming skills are required as well.

Content:
Dynamic programming is an optimization approach that has been widely applied in operations management. In this course, the students will study the most recent and advanced models in dynamic programming.

The seminar will focus on one for more of the following topics:
1. Dynamic programming
2. Approximate dynamic programming
3. Markov decision process
4. Partially observable Markov decision process
5. Deep reinforcement learning
6. Robust dynamic programming
7. Other topics related to advanced dynamic programming

Intended learning outcomes: Participants shall be able to understand the mathematical derivation of dynamic programming; to model the sequential decision-making problems; and to find the optimal solution efficiently.

Teaching and learning methods: Learning methods are a mix of seminar presentations by the participants and group discussions. There will be an intense 180 minutes session presentation and discussion each week. During this time, we are going to discuss a specific topic in dynamic programming in the greatest depth.

Topics and readings:

Topic 1: Dynamic programming

- Dimitri P. Bertsekas. Dynamic Programming and Optimal Control (Vol. I and II, 4th Edition), Athena Scientific (2012)

Topic 2: Approximate dynamic programming

- Warren B. Powell. Approximate Dynamic Programming: Solving the Curses of Dimensionality, (2nd Edition), Willy (2011)

Topic 3: Markov decision process

- Martin L. Puterman. Markov Decision Processes. Wiley-Interscience (2005)
- Sennott, L. I. (2009). Stochastic dynamic programming and the control of queueing systems (Vol. 504). John Wiley & Sons.

Topic 4: Partially observable Markov decision process

- Vikram Krishnamurthy. Partially Observed Markov Decision Processes. Cambridge University Press (2016)

Topic 5: Reinforcement learning

- Sutton R.S., Barton A.G. Reinforcement learning: An introduction, The MIT Press (2018)
- Das, T.K., Gosavi, A., Mahadevan, S. and Marchalleck, N., 1999. Solving semi-Markov decision problems using average reward reinforcement learning. Management Science, 45(4), pp.560-574.
- Gosavi, A., 2009. Reinforcement learning: A tutorial survey and recent advances. INFORMS Journal on Computing, 21(2), pp.178-192.
- Wen, Z. and Van Roy, B., 2017. Efficient reinforcement learning in deterministic systems with value function generalization. Mathematics of Operations Research, 42(3), pp.762-782.
- Cheung, W.C., Simchi-Levi, D. and Zhu, R., 2020, November. Reinforcement learning for non-stationary markov decision processes: The blessing of (more) optimism. In International Conference on Machine Learning (pp. 1843-1854). PMLR.
- Li, J., Luo, Y. and Zhang, X., 2021. Causal Reinforcement Learning: An Instrumental Variable Approach. Available at SSRN 3792824.
- Dai, J.G. and Gluzman, M., 2021. Queueing network controls via deep reinforcement learning. Stochastic Systems.
- Oroojlooyjadid, A., Nazari, M., Snyder, L.V. and Takáč, M., 2021. A Deep Q-Network for the Beer Game: Deep Reinforcement Learning for Inventory Optimization. Manufacturing & Service Operations Management.

Topic 6: Robust dynamic programming

- Iyengar, G. N. (2005). Robust dynamic programming. Mathematics of Operations Research, 30(2), 257-280.
- Nilim, A. and El Ghaoui, L., 2005. Robust control of Markov decision processes with uncertain transition matrices. Operations Research, 53(5), pp.780-798.
- Xu, H. and Mannor, S., 2012. Distributionally robust Markov decision processes. Mathematics of Operations Research, 37(2), pp.288-300.
- Wiesemann, W., Kuhn, D. and Rustem, B., 2013. Robust Markov decision processes. Mathematics of Operations Research, 38(1), pp.153-183.
- Lim, S.H., Xu, H. and Mannor, S., 2016. Reinforcement learning in robust markov decision processes. Mathematics of Operations Research, 41(4), pp.1325-1353.
- Goyal, V. and Grand-Clement, J., 2018. Robust Markov decision process: Beyond rectangularity. arXiv preprint arXiv:1811.00215.
- Bäuerle, N. and Glauner, A., 2021. Distributionally robust Markov decision processes and their connection to risk measures. Mathematics of Operations Research.

Responsible for module: Prof. Dr. Jingui Xie (jingui.xie@tum.de).