VE558: 随机控制及强化学习
Stochastic Control and Reinforcement Learning

Instructor:
金力 Li Jin
Assistant Professor of Electrical and Computer Engineering, Shanghai Jiao Tong University
li.jin@sjtu.edu.cn, Longbin Lou 424

Course description

Control and optimization of discrete-time and continuous-time Markov processes. Probability model, convergence of random variables. Countable-state Markov chains, continuous-state Markov chains, Foster-Lyapunov stability theory, Markov decision processes, dynamic programming, Monte-Carlo method, temporal-difference method, approximate dynamic programming. Continuous-time Markov processes, Poisson processes, queuing theory, infinitesimal generator, piecewise-deterministic Markov processes. Applications include connected and autonomous vehicles, intelligent transportation systems, computer and communication systems, social networks, epidemics, and finance.

Textbook:

1. Lecture notes.
2. Optional references (free online):
   a) Robert G. Gallager. *Stochastic Processes: Theory for Applications*. Cambridge University Press, 2013. https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-262-discrete-stochastic-processes-spring-2011/course-notes/
   b) Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT Press, 2018. http://incompleteideas.net/book/RLbook2020.pdf
   c) Meyn, S. P., & Tweedie, R. L. (2012). *Markov Chains and Stochastic Stability*. Springer Science & Business Media. http://www.lib.ysu.am/disciplines_bk/faab27c775970f84b2e87f848a9ffb79.pdf

Grading

Homework: 5×6=30: Independent work.
Quizzes: 25×2=50: 90 minutes in class, open-book.
Project: 20: Independent work based on stochastic control and/or reinforcement learning; Topics related to your own research is encouraged. 10-min presentation & 5-page report.
Honor code: https://www.ji.sjtu.edu.cn/academics/academic-integrity/honor-code/

Schedule

No.	Date	Topic	Note
1	9.13	Probability basics: Probability model, convergence of random variables, law of total probability/expectation	
2	9.15	Finite-state Markov chains: Classes, transient behavior, steady-state equations, matrix representation	

3	9.18	Countable-state Markov chains: Irreducibility, recurrence, ergodicity, birth-death chains, branching processes	HW1 due
4	9.22	Foster's theorem: Lyapunov function, Foster's theorem, comparison theorem.	
5	9.27	Continuous-state Markov chains: Basics of measure theory, transition probability kernel, stability theory, large deviation theory.	
6	9.29	Markov decision processes & dynamic programming: SAR formulation, feasible solution, optimal solution; Value function, Bellman equation, Bellman optimality equation, value iteration, policy evaluation/ improvement	HW 2 due
7	10.11	Tabular reinforcement learning methods: Multi-arm bandit, Monte-Carlo method, temporal-difference method	
8	10.13	Approximate algorithms I: on-policy methods	
9	10.18	Approximate algorithms II: off-policy methods	HW 3 due
10	10.20	Quiz 1	
11	10.25	Countable-state Markov processes: Definition, Poisson processes, infinitesimal generator	
12	10.27	Queuing systems: Formulation, Kendall's notation, Little's theorem, M/M/1, M/D/1, M/G/1, Foster-Lyapunov theorem	
13	11.1	Queuing networks: Jackson networks, Bernoulli routing, dynamic routing.	HW4 due
14	11.3	Markovian jump linear systems	
15	11.8	Piecewise-deterministic Markov processes	
16	11.10	Stochastic optimal control	HW5 due
17	11.15	Approximate dynamic programming I	
18	11.17	Approximate dynamic programming II	
19	11.22	Learning-based adaptive control	HW6 due
20	11.24	Quiz 2	
21	11.29	Engineering applications: Connected and autonomous vehicles, intelligent transportation systems, communication networks	
22	12.1	Socio-economical applications: Finance, social networks, epidemics	
23	12.6	Project presentation I	Project proposal due
24	12.8	Project presentation II	
	12.15		Project report due