Ma 661 Dynamic Programming and Reinforcement Learning

Darinka Dentcheva

darinka.dentcheva@stevens.edu

Homework 6

due Wednesday, April 22, 2020

- **Problem 1.** Prove Lemma 2.4 (ii) of the note on "Infinite-Horizon Average Cost Problem", that is, show that $HP^* = 0$.
- **Problem 2.** In Lemma 8.1 of the note on "Infinite-Horizon Average Cost Problem", show that the equation $(I \alpha P)^{-1} = I + \sum_{k=1}^{\infty} \alpha^k P^k$ holds.
- **Problem 3.** Each quarter the marketing manager of a retail store divides the customers into two groups based on their purchase behavior in the previous quarter. The classes are denoted by L and H. The manager wishes to determine to which group of customers he should sent a catalog. The cost of sending a catalog is \$ 15 per customer. If a customer from group L receives a catalog, then the expected purchase in the current quarter is \$ 20, otherwise it is \$ 10. If a customer from group H receives a catalog, then the expected purchase in the current quarter is \$ 50, otherwise it is \$ 25. Furthermore, if a customer from group L receives a catalog, then the probability that he will stay in group L for the next quarter is 0.3, otherwise, it is 0.5. If a customer from group H receives a catalog, then he probability that he will stay in group H for the next quarter is 0.8, otherwise, it is 0.4.
 - (a) Formulate an average reward problem to help the manager.
 - (b) Determine an optimal policy using policy iteration method.
 - (c) Solve the problem using linear programming.
 - (d) For what discount factor is the discounted infinite-horizon problem equivalent to the average reward problem in this context?