

Homework 2

due Wednesday, February 19, 2020

Problem 1. Assume that you play a chess match with a friend. If you play timid your probability of making a draw is $p = 0.9$, the probability to win is 0 and the probability to lose is 0.1. If you play bold you either win with probability $q = 0.45$, or you lose. Each win brings one point to the score of the winner. The match consists of 5 games. If the score is a tie after the fifth game, then a “sudden death” rule is adopted; that is, whoever wins the next game is a winner of the match; if it is a draw, then the game is repeated with the same rule.

Formulate a Markov decision problem to determine the optimal strategy of your play (to maximize the probability of winning the match) and solve it. Clearly describe the state space, control space, transition probabilities, and the reward function.

Problem 2. A software manufacturer can be in one of two states. In state 1 their software sells well, and in state 2, the product sells poorly. While in state 1, the company can invest in development of upgraded version of the software, in which case the one-stage reward is 4 units, and the probability of degrading to state 2 is 0.2. If no investment in new development occurs, then the reward is 6 units, but the probability of transition to state 2 is 0.5. While in state 2, if the company invests in software development, then the reward is -2 units, but the probability of transition to state 1 is 0.7. Without special efforts to improve, the reward is 1 and the probability of upgrading to state 1 is 0.

Formulate a dynamic programming problem to determine an optimal research and development policy. Solve the problem for a time horizon of 12 time intervals.

Problem 3. Consider the equipment replacement problem discussed in class, with the following data:

- operating cost per period $c_0 + c_1x$, $x = 0, 1, 2, \dots$;
- revenue per period R ;
- replacement cost K ;
- salvage value $\gamma K e^{-\mu x}$;

where $c_0 > 0$, $c_1 > 0$, $c_2 > 0$, $0 < \gamma < 1$, and $\mu > 0$. Assume that the salvage value can be collected whenever the item is replaced. The probabilities of deterioration by j steps in one period are given by the Poisson distribution

$$p_j = \frac{\lambda^j}{j!} e^{-\lambda}, \quad j = 0, 1, \dots$$

(3.1) Formulate the corresponding Markov decision problem. Clearly define the state space, action space, transition probabilities, and the reward function.

- (3.2) Solve the problem numerically for $c_0 = 1$, $c_1 = 1$, $R = 5$, $K = 10$, $\gamma = 0.8$, $\mu = 0.2$, $\lambda = 1$, and time horizon $T = 20$. To this end, argue that a constant value \bar{x} exists such that for all $x \geq \bar{x}$ replacement is always profitable. Then you will know that the value function for all $x \geq \bar{x}$ is the same as at \bar{x} . This will allow you have finite tables of the value function for each time t .