KTH Matematik

Home assignment 3, February 2009, in SF2862 Stochastic decision support models

Examiner: Krister Svanberg, *krille@math.kth.se*.

This home assignment should also be carried out in groups of at most two students. The problems may be discussed with other groups, but each group should, on their own and with their own words, write a *short* report where the obtained results are presented, including the rate diagrams and balance equations. The code used for the numerical calculations (Matlab is recommended) should be attached to the report.

A paper version of your report should be handed in to Mikael Fallgren or Krister Svanberg, not later than **March 4, 2009 at 16.00**.

Write your name, "personnummer" and e-mail address on the front page of the report.

Some groups may (partly by random) be seleced to give an oral presentation to the teachers. Read your e-mail to check if you are selected.

If the solutions and presentation are adequate, you get 2 bonus points to the final exam.

Consider the machine problem delt with in chapter 19 on Markov decision processes in Hillier Lieberman (and also delt with on some lectures), but with the following modifications: In each of the states $i \in \{1,2,3,4\}$, each of the three actions $k \in \{1,2,3\}$ is possible. The transition matrices $\mathbf{P}(k)$ for the three actions are given by

$$\mathbf{P}(1) = \begin{bmatrix} 0.3 & 0.4 & 0.2 & 0.1 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}, \quad \mathbf{P}(2) = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.2 & 0.6 & 0.2 & 0 \\ 0 & 0.5 & 0.3 & 0.2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{P}(3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

while the matrix with immediate expected costs C_{ik} is given by $\mathbf{C} = \begin{bmatrix} 0 & 1000 & 6000 \\ 1000 & 2200 & 6000 \\ 2000 & 3400 & 6000 \\ 6000 & 6000 & 6000 \end{bmatrix}$.

1. Use the linear programming approach to find a stationary policy which minimizes the long-run expected average cost per time period.

2. Use the dynamic programming approach (called Method of Successive Approximation in the book) to minimize the expected cost for $N = 10$ time periods. Compare the obtained (time dependent) policies with the stationary policy obtained above.

3. Now assume that there is a discount factor α. Use the linear programming approach to find a stationary policy which minimizes the expected total discounted cost (for infinite time horizon). Do this for both $\alpha = 0.95$ and $\alpha = 0.9$.

4. Assume again that there is a discount factor α. Use the dynamic programming approach to minimizes the expected total discounted cost for $N = 10$ time periods. Do this for both $\alpha = 0.95$ and $\alpha = 0.9$. Compare the obtained (time dependent) policies with the stationary policy obtained above.