भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

INDIAN INSTITUTE OF TECHNOLOGY HYDERABAD

AI1001
ASSIGNMENT 4

Introduction to Modern AI

Submitted To:
Prof. M. Vidyasagar
IIT Hyderabad

Submitted By :
Vijay Tadikamalla
CS17BTECH11040

1. Consider the Markov process with the state transition matrix

$$A = \begin{bmatrix} 0.1 & 0.4 & 0.3 & 0.2 & 0.0 \\ 0.1 & 0.0 & 0.4 & 0.3 & 0.2 \\ 0.0 & 0.3 & 0.0 & 0.5 & 0.2 \\ 0.0 & 0.1 & 0.3 & 0.2 & 0.4 \\ 0.0 & 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix}$$

Number the states consecutively from 1 to 5. What are the absorbing states, if any?

Solution:

Every state in A has transition to some other state. Hence there are no absorbing state in A.

2. Consider the Markov process with the state transition matrix

$$A = \begin{bmatrix} 0.1 & 0.4 & 0.3 & 0.2 & 0.0 \\ 0.1 & 0.0 & 0.4 & 0.3 & 0.2 \\ 0.0 & 0.3 & 0.0 & 0.5 & 0.2 \\ 0.0 & 0.1 & 0.3 & 0.2 & 0.4 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

Number the states consecutively from 1 to 5.

- Find the absorbing state(s) if any.
- Compute the average number of time steps needed to reach the absorbing state, starting from each of the other states.

Solution:

Only State 5 is absorbing state because $A[5, 5] = 1$.

$$Average\ time\ step = \begin{bmatrix} 4.8055 \\ 3.9253 \\ 3.7513 \\ 3.1474 \end{bmatrix}$$

The row i gives the average time needed to reach the absorbing state, starting from state i.

3. Now consider the Markov process with the state transition matrix

$$A = \begin{bmatrix} 0.1 & 0.4 & 0.3 & 0.2 & 0.0 & 0.0 \\ 0.1 & 0.0 & 0.4 & 0.3 & 0.1 & 0.1 \\ 0.0 & 0.3 & 0.0 & 0.4 & 0.1 & 0.2 \\ 0.0 & 0.1 & 0.3 & 0.3 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

- What are the absorbing states?
- For each of the nonabsorbing states, compute the probability of hitting each of the absorbing states.

Solution: State 5 and 6 are absorbing states ($A[5,5] = A[6,6] = 1$).

$$Hitting\ Probability = \begin{bmatrix} 0.5130 & 0.4870 \\ 0.5124 & 0.4876 \\ 0.4795 & 0.5205 \\ 0.5644 & 0.4356 \end{bmatrix}$$

Column 1 and 2 shows hitting probability for State 5 and 6 respectively.
Each row adds up to one, as expected.

4. Consider the following Markov reward process with state transition matrix

$$A = \begin{bmatrix} 0.1 & 0.4 & 0.3 & 0.2 & 0.0 \\ 0.1 & 0.0 & 0.4 & 0.3 & 0.2 \\ 0.0 & 0.3 & 0.0 & 0.5 & 0.2 \\ 0.0 & 0.1 & 0.3 & 0.2 & 0.4 \\ 0.0 & 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix}$$

and the reward vector

$$r = \begin{bmatrix} 3 & 1 & 4 & 5 & 2 \end{bmatrix}$$

With a discount factor of $\gamma = 0.8$, compute the value vector. Repeat with a discount factor of $\gamma = 0.5$.

Solution:
For $\gamma = 0.5$

$$Value\ Vector = \begin{bmatrix} 6.1266 & 4.540 & 7.281 & 8.231 & 5.462 \end{bmatrix}$$

For $\gamma = 0.8$

$$Value\ Vector = \begin{bmatrix} 16.032 & 14.667 & 17.286 & 18.208 & 15.560 \end{bmatrix}$$

5. Consider a Markov Decision Process transition matrices are given as with six states and two actions, where the two state transition matrices are given as

$$A^{(u1)} = \begin{bmatrix} 0.1 & 0.4 & 0.3 & 0.2 & 0.0 & 0.0 \\ 0.1 & 0.0 & 0.4 & 0.3 & 0.1 & 0.1 \\ 0.0 & 0.3 & 0.0 & 0.4 & 0.1 & 0.2 \\ 0.0 & 0.1 & 0.3 & 0.3 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.7 & 0.3 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.7 \end{bmatrix}, A^{(u2)} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 & 0.0 & 0.0 \\ 0.1 & 0.0 & 0.2 & 0.4 & 0.2 & 0.1 \\ 0.0 & 0.2 & 0.0 & 0.4 & 0.2 & 0.2 \\ 0.0 & 0.1 & 0.4 & 0.2 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.0 & 0.3 & 0.7 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.7 & 0.3 \end{bmatrix}$$

and corresponding reward matrix

$$R = \begin{bmatrix} 1 & 4 \\ 3 & 2 \\ 1 & 2 \\ 2 & 1 \\ 4 & 3 \\ 4 & 5 \end{bmatrix}$$

Note that the first column is the reward vector for the first action and the second column is the reward vector for the second action. Using polity iteration, compute the optimal value and optimal policy, when $\gamma = 0.5$, and when $\gamma = 0.8$.

Solution:

For $\gamma = 0.5$

$$Optimal\ Policy = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$
$$Optimal\ Value = \begin{bmatrix} 7.139 & 6.092 & 5.403 & 5.185 & 8.295 & 9.295 \end{bmatrix}$$

For $\gamma = 0.8$

$$Optimal\ Policy = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$
$$Optimal\ Value = \begin{bmatrix} 17.951 & 17.332 & 17.020 & 16.615 & 21.173 & 22.173 \end{bmatrix}$$