MACHINE LEARNING IN ROBOTICS

ASSIGNMENT 2

Submitted By

Bharti Munjal

Matriculation No:03658271

Exercise 1

Learned GMM Parameters

means =
  -0.0432   0.0446
  -0.0147 -0.0796
   0.0262   0.0617
  -0.0193 -0.0167

pis =
   0.2402
   0.2011
   0.2616
   0.2971

sigma1 =

  1.0e-03 *

   0.1750   0.2619
   0.2619   0.3981

sigma2 =

  1.0e-03 *

   0.3942   0.2165
   0.2165   0.1275

sigma3 =

   0.0011 -0.0004
  -0.0004   0.0002

sigma4 =

  1.0e-03 *

```
0.7436   -0.5916
-0.5916   0.6103
```

Exercise 2

1. Log-likelihood of A_Train_Binned.txt
   - -57.7141
   - -56.1732
   - -61.2662
   - -59.1239
   - -57.6957
   - -54.9307
   - -63.5531
   - -61.9493
   - -60.4662
   - -56.0470

2. Log-likelihood of A_Test_Binned.txt
   - -511.4069
   - -570.6697
   - -387.9167
   - -427.3069
   - -437.5989
   - -426.1784
   - -473.3031
   - -400.2880
   - -377.1776
   - -401.0614

3. Classification of A_Test_Binned.txt : All sequences (10) are classified as Test
4. Classification of A_Train_Binned.txt : All sequences (10) are classified as Train

Exercise 3

PolicyIteration

1. Reward Matrix:
   ```
   [7 -7  7 -7;
    -7 7 -7 -7;
   ```

```
       7 -7 -7 -7;
      -7 -7 7 -7;
      -7 -7 -7 7;
       7 -7 7 -7;
       7 -7  7 -7;
      -7 7 -7 -7;
      -7 -7 7 -7;
       7 -7 7  -7;
       7  7  7  7;
      -7 7 -7 7;
       7 -7 -7 -7;
      -7 -7 -7 7;
      -7 7 -7 7;
      -7 7 -7 7];
```

2. Effect of discount factor on no of steps required for convergence
   discount factor=0.9 ; No of steps=5 ;
   discount factor=0.8; No of steps=4 ;
   discount factor=0.7 No of steps=4;
   discount factor=0.6 No of steps=3;
   discount factor=0.5 No of steps=3;
   discount factor=0.4 No of steps=3;
   discount factor=0.2 No of steps=3;
   discount factor=0.1 No of steps=2;
   As the discount factor increases more steps are required for convergence
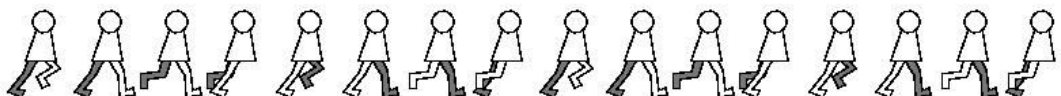3. Approx 4 to 5 iterations are required to converge
4. Outputs:
   a. For initial state 10

   b. For initial state 3

QWalkLearning

1. Reward Matrix

   [7 -7 7 -7;
    -7 7 -7 -7;
    7 -7 -7 -7;
    -7 -7 7 -7;
    -7 -7 -7 7;
    7 -7 7 -7;
    7 -7 7 -7;
    -7 7 -7 -7;
    -7 -7 7 -7;
    7 -7 7 -7;
    7 7 7 7;
    -7 7 -7 7;
    7 -7 -7 -7;
    -7 -7 -7 7;
    -7 7 -7 7;
    -7 7 -7 7];

2. alpha= 0.1, epsilon= 0.1, discount factor=0.9
   ❖ If pure greedy policy (epsilon=0) is used then less number of iterations are required (Approx. 45)
   ❖ Yes, it does matter what values of epsilon is used. Large value of epsilon means more iterations are required for convergence.

3. For epsilon=0   No Of Iterations= 45
   For epsilon=0.1 No Of Iterations= 70

4. Output
   a. initial state=5

   b. initial state=12