

F3_Exercises

Kwon do yun

2021-02-23

차 례

Policy Iteration 1 - MC Control	2
Policy Iteration 2 - TD Control (a.k.a. sarsa)	2

Policy Iteration 1 - MC Control

```
num_ep = 10**4
beg_time =time.time()
q_s_a = q_s_a_init
pi = pi_50

for epi_i in range(1,num_ep) :
    sample_path_i = simul_path(pi, P_normal, P_speed, R_s_a)
    q_s_a = pol_eval_MC(sample_path_i, q_s_a, alpha = 0.1)
    pi = pol_imp(pi, q_s_a, 0.1)

end_time =time.time()
result_q = pd.DataFrame(q_s_a, columns =['n','s'], index= states)
result_pi = pd.DataFrame(pi, columns =['n','s'], index= states)
print("Time difference of {} sec".format(end_time- beg_time))
```

```
## Time difference of 42.380091190338135 sec
```

```
print(result_pi.T)
```

```
##      0   10   20   30   40   50   60   70
## n  1.0  1.0  1.0  0.5  1.0  0.0  1.0  1.0
## s  0.0  0.0  0.0  0.5  0.0  1.0  0.0  0.0
```

```
print(result_q.T)
```

```
##      0      10      20      30      40      50      60      70
## n -5.959966 -4.984162 -3.928099 -2.895617 -1.766032 -2.024560 -1.000000  0.0
## s -6.629561 -5.248498 -4.498140 -3.251987 -1.933679 -1.713061 -1.656527  0.0
```

Policy Iteration 2 - TD Control (a.k.a. sarsa)

```
num_ep = 10**4
beg_time =time.time()
q_s_a = q_s_a_init
pi = pi_50
exploration_rate = 1
for epi_i in range(1,num_ep) :
    s_now = "0"
    while s_now != "70":
```

```

sample_step = simul_step(pi,s_now, P_normal, P_speed, R_s_a)
q_s_a = pol_eval_TD(sample_step, q_s_a, alpha = 0.01)
pi = pol_imp(pi, q_s_a, epsilon= exploration_rate)
s_now = sample_step[3]
exploration_rate *=0.995
end_time =time.time()
result_q = pd.DataFrame(q_s_a, columns =['n','s'], index= states)
result_pi = pd.DataFrame(pi, columns =['n','s'], index= states)
print("Time difference of {} sec".format(end_time- beg_time))

```

```
## Time difference of 129.33292317390442 sec
```

```
print(result_pi.T)
```

```
##      0   10   20   30   40   50   60   70
## n  0.0  1.0  0.0  1.0  1.0  0.0  1.0  1.0
## s  1.0  0.0  1.0  0.0  0.0  1.0  0.0  0.0
```

```
print(result_q.T)
```

```
##           0           10           20           30           40           50           60   70
## n -5.224649 -4.437470 -3.581390 -2.673233 -1.634618 -1.784836 -0.999789  0.0
## s -5.081633 -4.471394 -3.370718 -2.677920 -1.726274 -1.642682 -1.001449  0.0
```

```
"F3_Exercises"
```