F3 Python

Jaemin Park

2021-02-22

차 례

MC control

```
#F3 MC
num_ep = 1000
beg_time = time.time()
q_s_a = q_s_a_init
pi = pi_50
exp_rate = 1
for i in range(1000):
    sample_path_i = simul_path(pi,P_normal,P_speed, R_s_a,True)
    q_s_a = pol_eval_MC(sample_path_i,q_s_a,alpha = 1/(i+1))
print(q_s_a.T)
```

```
##          0         10        20        30        40        50         60   70
## n -4.568462 -2.880586 -2.893204 -1.541354 -0.580992 -0.975313  0.000000  0.0
## s -4.433742 -3.127271 -2.497822 -2.122910 -0.786928 -0.150241 -0.186155  0.0
```

```
for i in range(num_ep):
    if i%1000 == 0:
        print(i)
    sample_path_i = simul_path(pi,P_normal,P_speed, R_s_a,False)
    q_s_a = pol_eval_MC(sample_path_i,q_s_a,alpha = max(1/(i+1),0.1))
    pi = pol_imp(pi,q_s_a, exp_rate)
    exp_rate = exp_rate * 0.9997
```

```
## 0
```

```
end_time = time.time()
print(end_time-beg_time)
```

```
## 27.199814319610596
```

```
print(pi.T)
```

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

```
print(q_s_a.T)
```

```
##          0         10        20        30        40        50         60   70
## n -4.426932 -3.546458 -2.732696 -1.597353 -0.668141 -1.233602  0.000000  0.0
## s -3.687434 -3.631142 -2.204898 -2.207810 -0.560296 -0.104494 -0.269952  0.0
```

TD control

```
num_ep = 1000
beg_time = time.time()
q_s_a = q_s_a_init
pi = pi_50
exp_rate = 1
for i in range(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 = max(1/(i+1),0.005))
        pi = pol_imp(pi,q_s_a, exp_rate)
        s_now = sample_step[3]
        exp_rate = exp_rate*0.998

end_time = time.time()
print(end_time-beg_time)
```

```
## 28.912672519683838
```

```
print(pi.T)
```

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

```
print(q_s_a.T)
```

```
##          0         10        20        30        40        50        60   70
## n -4.757975 -3.908347 -3.188800 -2.443763 -1.565966 -1.621335 -0.975013  0.0
## s -4.752970 -3.909867 -3.193081 -2.730802 -1.561854 -1.602753 -1.710272  0.0
```