D1 Python Code

Kang, Eui Hyeon

2021-01-13

차 례

Monte-Carlo Simulation (10p)

```
# 10p
def soda_simul(this_state):
    u=np.random.uniform(0,1,size=1)

    if this_state=='c':
        if u<=0.7:
            next_state='c'
        else:
            next_state='p'
    else:
        if u<=0.5:
            next_state='c'
        else:
            next_state='p'

    return next_state

def cost_eval(path):
    cost_one_path=path.count('c')*1.5+path.count('p')*1
```

```
    return cost_one_path


MC_N=10000

spending_records=np.zeros((MC_N,))

for i in range(MC_N):
    path='c'

    for t in range(9):
        this_state=path[-1]
        next_state=soda_simul(this_state)
        path+=next_state

    spending_records[i]=cost_eval(path)
```

Monte-Carlo Simulation (11p)

```
# 11p
def state_value_function(num_episode, time_horizon):
    episode_i=0
    cum_sum_G_i=0

    while episode_i<num_episode:
        path='c'

        for i in range(time_horizon):
            this_state=path[-1]
            next_state=soda_simul(this_state)
            path+=next_state

        G_i=cost_eval(path)
        cum_sum_G_i+=G_i
        episode_i+=1

    V_t=cum_sum_G_i/MC_N

    return V_t

result=state_value_function(10000,9)
result
```

```
## 13.34815
```

Exercise 17p

For general t

$$
\begin{aligned}
V_t(S) &= \mathbb{E}[G_t|S_t = s] \\
&= \mathbb{E}[\sum_{i=t}^{n} r_t|S_t = s] \\
&= \mathbb{E}[r_t + r_{t+1} + r_{t+2}...r_n|S_t = s] \\
&= R(s) + \mathbb{E}[r_{t+1} + r_{t+2}...r_n|S_t = s] \\
&= R(s) + \mathbb{E}[G_{t+1}|S_t = s] \\
&= R(s) + \sum_{s^{'} \in S} P_{ss'} \mathbb{E}[G_{t+1}|S_t = s, S_{t+1} = s^{'}] \\
&= R(s) + \sum_{s^{'} \in S} P_{ss'} \mathbb{E}[G_{t+1}|S_{t+1} = s^{'}] \quad (MarkovProperty) \\
&= R(s) + \sum_{s^{'} \in S} P_{ss'} V_{t+1}(s^{'})
\end{aligned}
$$

Implementation strategy (20p)

```
# 20p

P=np.array([0.7,0.5,0.3,0.5]).reshape(2,2, order='F')
R=np.array([1.5,1.0]).reshape(2,1)

H=10

v_t1=np.array([0,0]).reshape(2,1)
t=H-1

while t>=0 :
    v_t=R+np.dot(P,v_t1)
    t=t-1
    v_t1=v_t

v_t
```

```
## array([[13.35937498],
##        [12.73437504]])
```