F1 Python

Jaemin Park

2021-01-28

차 례

p.9 Preparation

```
states = np.arange(0,80,10)
P_normal = np.matrix([[0,1,0,0,0,0,0,0],[0,0,1,0,0,0,0,0],
[0,0,0,1,0,0,0,0],[0,0,0,0,1,0,0,0],[0,0,0,0,0,1,0,0],[0,0,0,0,0,0,1,0],
[0,0,0,0,0,0,0,1],[0,0,0,0,0,0,0,1]])
P_speed = np.matrix([[0.1,0,0.9,0,0,0,0,0],[0.1,0,0,0.9,0,0,0,0],
[0,0.1,0,0,0.9,0,0,0],[0,0,0.1,0,0,0.9,0,0],[0,0,0,0.1,0,0,0.9,0],
[0,0,0,0,0.1,0,0,0.9],[0,0,0,0,0,0.1,0,0.9],[0,0,0,0,0,0,0,1]])
P_normal = pd.DataFrame(P_normal,states,states)
P_speed = pd.DataFrame(P_speed,states,states)
R_s_a = np.matrix([[-1,-1,-1,-1,0,-1,-1,0],[-1.5,-1.5,-1.5,-1.5,-0.5,-1.5,-1.5,0]]).T
R_s_a = pd.DataFrame(R_s_a,states,["n","s"])

pi_speed = np.hstack((np.repeat(0,len(states)).reshape(8,1),np.repeat(1,len(states)).reshape(8,1)))
pi_speed = pd.DataFrame(pi_speed,states,["n","s"])

pi_50 = np.hstack((np.repeat(0.5,len(states)).reshape(8,1),np.repeat(0.5,len(states)).reshape(8,1)))
pi_50 = pd.DataFrame(pi_50,states,["n","s"])
pi_speed.T
```

```
##    0   10  20  30  40  50  60  70
## n   0   0   0   0   0   0   0   0
## s   1   1   1   1   1   1   1   1
```

```
pi_50.T
```

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

p.11 Simulator π^{speed}

```
pi = pi_speed
history = []
MC_N = 10000
for MC_i in range(MC_N):
    s_now = 0
    history_i = []
    while(s_now!=70):
        if(np.random.uniform(0,1)<pi.loc[s_now]["n"]):
            a_now="n"
            P=P_normal
        else:
            a_now="s"
            P=P_speed
        r_now = R_s_a.loc[s_now][a_now]
        s_next = states[np.argmin(P.loc[s_now].cumsum()<np.random.uniform(0,1))].item()
        history_i.extend([s_now,a_now,r_now])
        s_now = s_next
    history.append(history_i)

history_speed = history
print(pd.DataFrame(np.matrix(history_speed[:20])).T)
```

```
##                                                    0
## 0   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 1   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 2   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 3   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 4   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 5   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 6   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 7   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 8   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 9   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 10  [0, s, -1.5, 20, s, -1.5, 10, s, -1.5, 30, s, ...
## 11  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 12  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 13  [0, s, -1.5, 20, s, -1.5, 10, s, -1.5, 30, s, ...
## 14  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 30, s, ...
```

```
## 15  [0, s, -1.5, 20, s, -1.5, 10, s, -1.5, 30, s, ...
## 16  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 17  [0, s, -1.5, 0, s, -1.5, 0, s, -1.5, 20, s, -1...
## 18  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 19  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
##
## D:\miniconda3\envs\r-reticulate\lib\site-packages\numpy\matrixlib\defmatrix.py:145: VisibleDeprec
##   arr = N.array(data, dtype=dtype, copy=copy)
```

p.13 Simulator π^{50}

```
pi = pi_50
history = []
MC_N = 10000
for MC_i in range(MC_N):
    s_now = 0
    history_i = []
    while(s_now!=70):
        if(np.random.uniform(0,1)<pi.loc[s_now]["n"]):
            a_now="n"
            P=P_normal
        else:
            a_now="s"
            P=P_speed
        r_now = R_s_a.loc[s_now][a_now]
        s_next = states[np.argmin(P.loc[s_now].cumsum()<np.random.uniform(0,1))].item()
        history_i.extend([s_now,a_now,r_now])
        s_now = s_next
    history.append(history_i)

history_50 = history
print(pd.DataFrame(np.matrix(history_50[:20])).T)
```

```
##                                                    0
## 0   [0, n, -1.0, 10, s, -1.5, 0, s, -1.5, 20, n, -...
## 1   [0, n, -1.0, 10, n, -1.0, 20, n, -1.0, 30, n, ...
## 2   [0, s, -1.5, 20, n, -1.0, 30, n, -1.0, 40, n, ...
## 3   [0, n, -1.0, 10, s, -1.5, 30, n, -1.0, 40, s, ...
## 4   [0, s, -1.5, 20, n, -1.0, 30, n, -1.0, 40, n, ...
## 5   [0, n, -1.0, 10, s, -1.5, 30, n, -1.0, 40, s, ...
## 6   [0, n, -1.0, 10, n, -1.0, 20, s, -1.5, 40, s, ...
## 7   [0, n, -1.0, 10, s, -1.5, 30, n, -1.0, 40, s, ...
## 8   [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 9   [0, n, -1.0, 10, s, -1.5, 30, s, -1.5, 20, n, ...
## 10  [0, s, -1.5, 20, s, -1.5, 40, s, -0.5, 60, s, ...
## 11  [0, s, -1.5, 20, n, -1.0, 30, n, -1.0, 40, s, ...
## 12  [0, s, -1.5, 20, s, -1.5, 10, n, -1.0, 20, n, ...
## 13  [0, n, -1.0, 10, n, -1.0, 20, n, -1.0, 30, n, ...
## 14  [0, n, -1.0, 10, n, -1.0, 20, s, -1.5, 40, n, ...
```

```
## 15  [0, s, -1.5, 20, n, -1.0, 30, s, -1.5, 50, n, ...
## 16  [0, s, -1.5, 20, n, -1.0, 30, n, -1.0, 40, s, ...
## 17  [0, s, -1.5, 20, s, -1.5, 40, n, 0.0, 50, s, -...
## 18  [0, s, -1.5, 20, s, -1.5, 40, n, 0.0, 50, s, -...
## 19  [0, n, -1.0, 10, n, -1.0, 20, n, -1.0, 30, s, ...
##
## D:\miniconda3\envs\r-reticulate\lib\site-packages\numpy\matrixlib\defmatrix.py:145: VisibleDeprec
##   arr = N.array(data, dtype=dtype, copy=copy)
```

p.17 Implementation 1 π^{speed} (vectorized)

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), index=states, col
pol_eval.T
```

```
##         0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## sum    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_speed[MC_i]

    for j in range(0,len(history_i),3):
        pol_eval.loc[history_i[j]]['count']+=1

        if j < len(history_i) :
            pol_eval.loc[history_i[j]]['sum']+=pd.Series(history_i)[range(j+2,len(history_i)-1,3)].a

        else:
            pol_eval.loc[history_i[j]]['sum']+=0
print(pol_eval.T)
```

```
##             0       10       20      30      40      50      60   70
## count  11272.0  1066.0  10304.0  1905.0  9491.0  2538.0  8546.0  0.0
## sum   -48708.5 -3930.5 -27212.0 -3760.0 -8044.0  -598.0 -1425.5  0.0
```

```
pol_cal=pd.DataFrame(pol_eval['sum']/pol_eval['count'])
print(pol_cal.T)
```

```
##          0         10        20        30       40        50        60  70
## 0 -4.321194 -3.687148 -2.640916 -1.973753 -0.84754 -0.235619 -0.166803 NaN
```

p.19 Implementation 2 π^{speed} (running estimate)

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), index=states, col
pol_eval.T
```

```
##          0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## est    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_speed[MC_i]

    for j in range(0,len(history_i),3):
        # update count
        pol_eval.loc[history_i[j]]['count']+=1
        current_cnt=pol_eval.loc[history_i[j]]['count']

        # return is the new info
        if j < len(history_i):
            new_info=pd.Series(history_i)[range(j+2,len(history_i)-1,3)].astype('float').sum()

        else:
            new_info=0

        # update the last estimate with new info
        alpha=1/current_cnt
        pol_eval.loc[history_i[j]]['est']+=alpha*(new_info-pol_eval.loc[history_i[j]]['est'])


np.round(pol_eval.T,2)
```

```
##              0        10        20       30       40       50       60   70
## count  11272.00  1066.00  10304.00  1905.00  9491.00  2538.00  8546.00  0.0
## est       -4.32    -3.69     -2.64    -1.97    -0.85    -0.24    -0.17  0.0
```

p.21 Implementation 3 π^{50} (vectorized)

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), index=states, col
pol_eval.T
```

```
##         0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## sum    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_50[MC_i]

    for j in range(0,len(history_i),3):
        pol_eval.loc[history_i[j]]['count']+=1

        if j < len(history_i) :
            pol_eval.loc[history_i[j]]['sum']+=pd.Series(history_i)[range(j+2,len(history_i)-1,3)].a

        else:
            pol_eval.loc[history_i[j]]['sum']+=0
pol_eval.T
```

```
##             0        10       20       30      40      50      60   70
## count  10845.0   5912.0   8058.0   7058.0  7546.0  7324.0  7074.0  0.0
## sum   -50262.0 -22547.5 -22345.0 -14418.0 -5489.0 -4871.0  -714.0  0.0
```

```
pol_cal=pd.DataFrame(pol_eval['sum']/pol_eval['count'])
print(pol_cal.T)
```

```
##          0         10        20        30        40        50        60  70
## 0 -4.634578 -3.813853 -2.773021 -2.042788 -0.727405 -0.665074 -0.100933 NaN
```

p.23 Implementation 4 π^{50} (running estimate)

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), index=states, col
pol_eval.T
```

```
##          0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## est    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_50[MC_i]

    for j in range(0,len(history_i),3):
        # increment count
        pol_eval.loc[history_i[j]]['count']+=1
        current_cnt=pol_eval.loc[history_i[j]]['count']

        # return is the new info
        if j < len(history_i):
            new_info=pd.Series(history_i)[range(j+2,len(history_i)-1,3)].astype('float').sum()

        else:
            new_info=0

        # update the last estimate with new info
        alpha=1/current_cnt
        pol_eval.loc[history_i[j]]['est']+=alpha*(new_info-pol_eval.loc[history_i[j]]['est'])


np.round(pol_eval.T,2)
```

```
##               0        10       20       30       40       50       60   70
## count  10845.00  5912.00  8058.00  7058.00  7546.00  7324.00  7074.0  0.0
## est       -4.63    -3.81    -2.77    -2.04    -0.73    -0.67    -0.1  0.0
```

p.35 Implementation 5 π^{speed}

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), states, ['count',
pol_eval.T
```

```
##          0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## est    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_speed[MC_i]

    for j in range(0,len(history_i),3):
        pol_eval.loc[history_i[j]]['count']+=1
        current_cnt=pol_eval.loc[history_i[j]]['count']

        if j < len(history_i)-3 :
            new_info = history_i[j+2]+pol_eval.loc[history_i[j+3]]['est']

        else:
            new_info =0

        alpha=1/current_cnt
        pol_eval.loc[history_i[j]]['est']+=alpha*(new_info-pol_eval.loc[history_i[j]]['est'])

print(np.round(pol_eval.T,2))
```

```
##               0        10        20       30       40       50       60   70
## count  11272.00  1066.00  10304.00  1905.00  9491.00  2538.00  8546.00  0.0
## est       -4.28    -3.71     -2.64    -1.98    -0.85    -0.23    -0.17  0.0
```

p.37 Implementation 5 π^{50}

```
pol_eval=pd.DataFrame(np.matrix(np.zeros((len(states)*2))).reshape(len(states),2), states, ['count',
pol_eval.T
```

```
##         0    10   20   30   40   50   60   70
## count  0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
## est    0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
```

```
for MC_i in range(MC_N):
    history_i=history_50[MC_i]

    for j in range(0,len(history_i),3):
        pol_eval.loc[history_i[j]]['count']+=1
        current_cnt=pol_eval.loc[history_i[j]]['count']

        if j < len(history_i)-3 :
            new_info = history_i[j+2]+pol_eval.loc[history_i[j+3]]['est']

        else:
            new_info =0

        alpha=1/current_cnt
        pol_eval.loc[history_i[j]]['est']+=alpha*(new_info-pol_eval.loc[history_i[j]]['est'])

print(np.round(pol_eval.T,2))
```

```
##              0        10       20       30       40       50       60   70
## count  10845.00  5912.00  8058.00  7058.00  7546.00  7324.00  7074.00  0.0
## est       -4.62    -3.79    -2.77    -2.03    -0.71    -0.66    -0.11  0.0
```