Answer to Q2

July 15, 2019

In [1]:
```
import numpy as np
print("Jun Qing 1002088, and Shaun Toh 1002012, done in collaboration")
action0 = np.array([[0.1,0.5,0.4],[0.2,0.2,0.6],[0.8,0.1,0.1]])
action1 = np.array([[0.3,0.3,0.4],[0.3,0.4,0.3],[0.6,0.2,0.2]])
a0r=np.array([1,2,4])
a1r=np.array([0,5,2])
action0=action0.transpose()
action1=action1.transpose() # this is transposed to allow for
#correct multiplication
action0mat = action0*a0r # rewards for action0 and its various states.
action1mat = action1*a1r # blah
print("action 0:\n",action0)
print()
print("action 1:\n",action1)
print()
print("reward for action 0:\n",action0)
print()
print("rewards for action 1:\n",action1)
print()
print("rewards weighed via probabilities for action 0\n",action0mat)
print()
print("rewards weighed via probabilities for action 1\n",action1mat)
print()
eff_val_0 = np.sum(action0mat,axis=1)
eff_val_1 =np.sum(action1mat,axis=1)
print("In the order of state0,state1,state2...")
print("effective value of an action 0 given current state:\n",eff_val_0)
print()
print("effective value of action 1 given current state:\n",eff_val_1)
print()
print("however, we have a set policy already of")
print("doing action 1 in state 0, and state 2")
print("while doing action 0 in state 1 only.")
print("As a result we can ignore the rewards for irrelevant actions,")
print("obtaining the rewards obtained by the policy per movement.")
true_eff_val = np.array([eff_val_1[0],eff_val_0[1],eff_val_1[2]])
print(["s=0","s=1","s=2"])
print(true_eff_val)
```

```
Jun Qing 1002088, and Shaun Toh 1002012, done in collaboration
action 0:
 [[0.1 0.2 0.8]
 [0.5 0.2 0.1]
 [0.4 0.6 0.1]]

action 1:
 [[0.3 0.3 0.6]
 [0.3 0.4 0.2]
 [0.4 0.3 0.2]]

reward for action 0:
 [[0.1 0.2 0.8]
 [0.5 0.2 0.1]
 [0.4 0.6 0.1]]

rewards for action 1:
 [[0.3 0.3 0.6]
 [0.3 0.4 0.2]
 [0.4 0.3 0.2]]

rewards weighed via probabilities for action 0
 [[0.1 0.4 3.2]
 [0.5 0.4 0.4]
 [0.4 1.2 0.4]]

rewards weighed via probabilities for action 1
 [[0.  1.5 1.2]
 [0.  2.  0.4]
 [0.  1.5 0.4]]

In the order of state0,state1,state2...
effective value of an action 0 given current state:
 [3.7 1.3 2. ]

effective value of action 1 given current state:
 [2.7 2.4 1.9]

however, we have a set policy already of
doing action 1 in state 0, and state 2
while doing action 0 in state 1 only.
As a result we can ignore the rewards for irrelevant actions,
obtaining the rewards obtained by the policy per movement.
['s=0', 's=1', 's=2']
[2.7 1.3 1.9]
```

In [2]:
```
print("there are hence 3^3 possible policy values depending on the states.")
```

```
print("\n")
val_list1 = []
val_list2 = [] # tuple output... of real, gamma, gamma square
val_list3 = []
states_list =[]
for i in range(len(true_eff_val)):
    for j in range(len(true_eff_val)):
        for k in range(len(true_eff_val)):
            print("Landed in State",i, "then State",j,"then State", k)

            print("k=1 value")
            val_list1.append((true_eff_val[i],0,0))
            print("Real Value: ",true_eff_val[j],
                  " Gamma value:",0, "Gamma square value:",0)

            print("k=2 value")
            val_list2.append((true_eff_val[j],true_eff_val[i],0))
            print("Real Value: ",true_eff_val[j],
                  " Gamma value:",true_eff_val[i], "Gamma square value:",0)

            print("k=3 value")
            print("Real Value: ",true_eff_val[k],
                  " Gamma value:",true_eff_val[j], "Gamma square value:",true_eff_val[i]
            val_list3.append((true_eff_val[k],true_eff_val[j],true_eff_val[i]))
            states_list.append((i,j,k))
            print()
            print()
```

```
there are hence 3^3 possible policy values depending on the states.


Landed in State 0 then State 0 then State 0
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 2.7


Landed in State 0 then State 0 then State 1
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 2.7
```

```
Landed in State 0 then State 0 then State 2
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 2.7


Landed in State 0 then State 1 then State 0
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 2.7


Landed in State 0 then State 1 then State 1
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 2.7


Landed in State 0 then State 1 then State 2
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 2.7


Landed in State 0 then State 2 then State 0
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 2.7


Landed in State 0 then State 2 then State 1
k=1 value
```

```
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 2.7


Landed in State 0 then State 2 then State 2
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 2.7


Landed in State 1 then State 0 then State 0
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 1.3


Landed in State 1 then State 0 then State 1
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 1.3


Landed in State 1 then State 0 then State 2
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 1.3


Landed in State 1 then State 1 then State 0
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 0
```

```
k=3 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 1.3


Landed in State 1 then State 1 then State 1
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 1.3


Landed in State 1 then State 1 then State 2
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 1.3


Landed in State 1 then State 2 then State 0
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 1.3


Landed in State 1 then State 2 then State 1
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 1.3


Landed in State 1 then State 2 then State 2
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 1.3
```

```
Landed in State 2 then State 0 then State 0
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 2.7 Gamma square value: 1.9


Landed in State 2 then State 0 then State 1
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 2.7 Gamma square value: 1.9


Landed in State 2 then State 0 then State 2
k=1 value
Real Value:  2.7  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 2.7 Gamma square value: 1.9


Landed in State 2 then State 1 then State 0
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 1.3 Gamma square value: 1.9


Landed in State 2 then State 1 then State 1
k=1 value
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.3 Gamma square value: 1.9


Landed in State 2 then State 1 then State 2
k=1 value
```

```
Real Value:  1.3  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.3 Gamma square value: 1.9


Landed in State 2 then State 2 then State 0
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  2.7  Gamma value: 1.9 Gamma square value: 1.9


Landed in State 2 then State 2 then State 1
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.3  Gamma value: 1.9 Gamma square value: 1.9


Landed in State 2 then State 2 then State 2
k=1 value
Real Value:  1.9  Gamma value: 0 Gamma square value: 0
k=2 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 0
k=3 value
Real Value:  1.9  Gamma value: 1.9 Gamma square value: 1.9
```