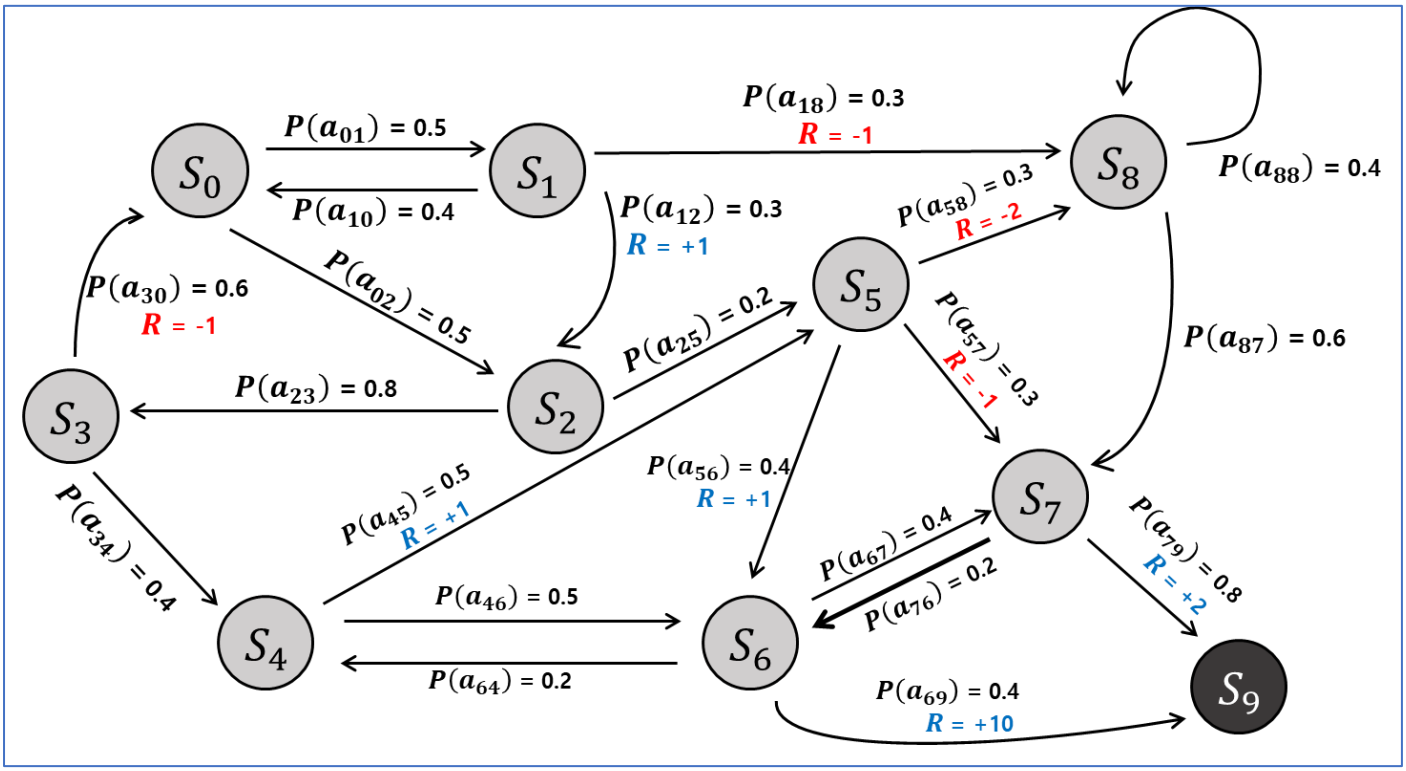
지능시스템

과제1 - 벨만 기대방정식

2019305059

이현수

● 다음의 State diagram으로 나타난 MDP에 대해 벨만 기대방정식을 세우고 이를 행렬식을 이용하여 푸시오.



벨만 기대 방정식

V(S0) = 0.5 X 0.9 X V(S1) + 0.5 X 0.9 X V(S2)

V(S1) = 0.4 X 0.9 X V(S0) + 0.3 X (1+0.9xV(S2)) + 0.3 X (-1+0.9 x V(S8))

V(S2) = 0.8 X 0.9 X V(S3) + 0.2 X 0.9 X V(S5)

V(S3) = 0.6 X (-1+0.9 x V(S0)) + 0.4 X 0.9 X V(S4)

V(S4) = 0.5 X (1+0.9 X V(S5)) + 0.5 X 0.9 X V(S6)

V(S5) = 0.4 X (1+0.9 X V(S6)) + 0.3 X (-1+0.9 X V(S7)) + 0.3 X (-2+0.9 X V(S8))

V(S6) = 0.2 X 0.9 X V(S4) + 0.4 X 0.9 X V(S7) +0.4 X (10 + 0.9 X V(S9))

V(S7) = 0.2 X 0.9 X V(S6) + 0.8 X (2+0.9 X V(S9))

V(S8) = 0.6 X 0.9 X V(S7) + 0.4 X 0.9 X V(S8)

행렬식

$$\begin{bmatrix} V(S0) \\ V(S1) \\ V(S2) \\ V(S3) \\ V(S4) \\ V(S5) \\ V(S6) \\ V(S7) \\ V(S8) \end{bmatrix} = \begin{bmatrix} 0 & 0.45 & 0.45 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.36 & 0 & 0.27 & 0 & 0 & 0 & 0 & 0 & 0.27 \\ 0 & 0 & 0 & 0.72 & 0 & 0.18 & 0 & 0 & 0 \\ 0.54 & 0 & 0 & 0 & 0.36 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.45 & 0.45 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.36 & 0.27 & 0.27 \\ 0 & 0 & 0 & 0 & 0.18 & 0 & 0 & 0.36 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.18 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.54 & 0.36 \end{bmatrix} \begin{bmatrix} V(S0) \\ V(S1) \\ V(S2) \\ V(S3) \\ V(S4) \\ V(S5) \\ V(S6) \\ V(S7) \\ V(S8) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -0.6 \\ 0.5 \\ -0.5 \\ 4 \\ 1.6 \\ 0 \end{bmatrix}$$

V = AV + B

V − AV = B

(E − A)V = B

V = $(\boldsymbol{E}-\boldsymbol{A})^{-1}$B

E 행렬은 아래와 같다.

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \mathbf{0} & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{bmatrix}$$

E-A 행렬은

$$\begin{bmatrix} \mathbf{1} & \mathbf{-0.45} & \mathbf{-0.45} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{-0.36} & \mathbf{1} & \mathbf{-0.27} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.27} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-0.72} & \mathbf{0} & \mathbf{-0.18} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{-0.54} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-0.36} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-0.45} & \mathbf{-0.45} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-0.36} & \mathbf{-0.27} & \mathbf{-0.27} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.18} & \mathbf{0} & \mathbf{1} & \mathbf{-0.36} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.18} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.54} & \mathbf{0.64} \end{bmatrix}$$

$(\boldsymbol{E}-\boldsymbol{A})^{-1}$은 파이썬으로 구해보면

```
import numpy as np

x = np.array([[1,-0.45,-0.45,0,0,0,0,0,0],
              [-0.36,1,-0.27,0,0,0,0,0,-0.27],
              [0,0,1,-0.72,0,-0.18,0,0,0],
              [-0.54,0,0,1,-0.36,0,0,0,0],
              [0,0,0,0,1,-0.45,-0.45,0,0],
              [0,0,0,0,0,1,-0.36,-0.27,-0.27],
              [0,0,0,0,-0.18,0,1,-0.36,0],
              [0,0,0,0,0,0,-0.18,1,0],
              [0,0,0,0,0,0,0,-0.54,0.64]])

y=np.linalg.inv(x)
print(y)
```

```
[[1.62390176 0.73075579 0.92805985 0.66820309 0.2982193  0.30124946
  0.32036771 0.43177342 0.43537721]
 [0.75507534 1.3397839  0.70152556 0.5050984  0.23253253 0.23091424
  0.2816506  0.52156573 0.66263828]
 [0.631373   0.28411785 1.36082967 0.97979736 0.43017703 0.438529
  0.43027764 0.43793077 0.30486664]
 [0.87690695 0.39460813 0.50115232 1.36082967 0.57272765 0.34793486
  0.46016096 0.42876093 0.31326032]
 [0.         0.         0.         0.         1.1435812  0.51461154
  0.79767334 0.54334246 0.21710174]
 [0.         0.         0.         0.         0.09896175 1.04453279
  0.54978748 0.71790497 0.44066227]
 [0.         0.         0.         0.         0.22010759 0.09904841
  1.22281993 0.48952271 0.04178605]
 [0.         0.         0.         0.         0.03961937 0.01782871
  0.22010759 1.08811409 0.00752149]
 [0.         0.         0.         0.         0.03342884 0.01504298
  0.18571578 0.91809626 1.56884626]]
```

$(E-A)^{-1}$ X B는

$$\begin{bmatrix} 1.623902 & 0.730756 & 0.928060 & 0.668203 & 0.298219 & 0.301249 & 0.320368 & 0.431773 & 0.435377 \\ 0.755075 & 1.339784 & 0.701526 & 0.505098 & 0.232533 & 0.230914 & 0.281651 & 0.521566 & 0.662638 \\ 0.631373 & 0.284118 & 1.360830 & 0.979797 & 0.430177 & 0.438529 & 0.430278 & 0.437931 & 0.304867 \\ 0.876907 & 0.394608 & 0.501152 & 1.360830 & 0.572728 & 0.347935 & 0.460161 & 0.428761 & 0.313260 \\ 0 & 0 & 0 & 0 & 1.143581 & 0.514612 & 0.797673 & 0.543342 & 0.217102 \\ 0 & 0 & 0 & 0 & 0.098962 & 1.044533 & 0.549787 & 0.717905 & 0.440662 \\ 0 & 0 & 0 & 0 & 0.220108 & 0.099048 & 1.222820 & 0.489823 & 0.041786 \\ 0 & 0 & 0 & 0 & 0.039619 & 0.017829 & 0.220108 & 1.088114 & 0.007521 \\ 0 & 0 & 0 & 0 & 0.033429 & 0.015043 & 0.185716 & 0.918096 & 1.568846 \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 0 \\ -0.6 \\ 0.5 \\ -0.5 \\ 4 \\ 1.6 \\ 0 \end{bmatrix}$$

파이썬으로 계산하면

```
b=np.array([[0],[0],[0],[-0.6],[0.5],[-0.5],[4],[1.6],[0]])
result=np.dot(y,b)
print(result)
```

```
[[1.56987137]
 [1.65885766]
 [1.82974539]
 [1.82255994]
 [4.37452611]
 [2.87501237]
 [5.73504566]
 [2.63230822]
 [2.22101006]]
```

$$(E-A)^{-1}\text{X B} = \begin{bmatrix} 1.569871 \\ 1.658858 \\ 1.829745 \\ 1.822560 \\ 4.374526 \\ 2.875012 \\ 5.735046 \\ 2.632308 \\ 2.221010 \end{bmatrix}$$

결과 :

V(S0) = 1.569871

V(S1) = 1.658858

V(S2) = 1.829745

V(S3) = 1.822560

V(S4) = 4.374526

V(S5) = 2.875012

V(S6) = 5.735046

V(S7) = 2.632308

V(S8) = 2.221010

V(S9) = 0