Answer 2

```
[[ 3.30903373  8.78932925  4.42765654  5.32240493  1.49221608]
 [ 1.52162547  2.99235524  2.25017731  1.90760904  0.54744003]
 [ 0.05085989  0.73820797  0.67315062  0.35822355 -0.40310382]
 [-0.97355491 -0.43545805 -0.35484491 -0.58556775 -1.18303775]
 [-1.85766316 -1.34519388 -1.2292299  -1.42288081 -1.97514172]]
```

Answer 4

```
[[21.97744338 24.41938153 21.97744338 19.41938153 17.47744338]
 [19.77969904 21.97744338 19.77969904 17.8017056  16.02153504]
 [17.8017056  19.77969904 17.8017056  16.02153504 14.41938153]
 [16.02153504 17.8017056  16.02153504 14.41938153 12.97744338]
 [14.41938153 16.02153504 14.41938153 12.97744338 11.67969904]]
```

```
[0, 0, 24.41938153311321, 19.7796990418217]
Optimal Policy for State: (0, 0) --- [2]
[16.021535036792454, 16.021535036792454, 16.021535036792454, 16.021535036792454]
Optimal Policy for State: (0, 1) --- [0, 1, 2, 3]
[24.41938153311321, 0, 19.41938153311321, 19.7796990418217]
Optimal Policy for State: (0, 2) --- [0]
[16.021535036792454, 16.021535036792454, 16.021535036792454, 16.021535036792454]
Optimal Policy for State: (0, 3) --- [0, 1, 2, 3]
[19.41938153311321, 0, 0, 16.021535036792454]
Optimal Policy for State: (0, 4) --- [0]
[0, 21.97744337980189, 21.97744337980189, 17.80172913763953]
Optimal Policy for State: (1, 0) --- [1, 2]
[19.7796990418217, 24.41938153311321, 19.7796990418217, 19.7796990418217]
Optimal Policy for State: (1, 1) --- [1]
[21.97744337980189, 21.97744337980189, 17.80172913763953, 17.80172913763953]
Optimal Policy for State: (1, 2) --- [0, 1]
[19.7796990418217, 19.41938153311321, 16.021535036792454, 16.021535036792454]
Optimal Policy for State: (1, 3) --- [0]
[17.80172913763953, 17.47744337980189, 0, 14.419381533113208]
Optimal Policy for State: (1, 4) --- [0]
[0, 19.7796990418217, 19.7796990418217, 16.021535036792454]
Optimal Policy for State: (2, 0) --- [1, 2]
[17.80172913763953, 21.97744337980189, 17.80172913763953, 17.80172913763953]
Optimal Policy for State: (2, 1) --- [1]
[19.7796990418217, 19.7796990418217, 16.021535036792454, 16.021535036792454]
```

Optimal Policy for State: (2, 2) --- [0, 1]
[17.80172913763953, 17.80172913763953, 14.419381533113208, 14.419381533113208]
Optimal Policy for State: (2, 3) --- [0, 1]
[16.021535036792454, 16.021535036792454, 0, 12.977443379801887]
Optimal Policy for State: (2, 4) --- [0, 1]
[0, 17.80172913763953, 17.80172913763953, 14.419381533113208]
Optimal Policy for State: (3, 0) --- [1, 2]
[16.021535036792454, 19.7796990418217, 16.021535036792454, 16.021535036792454]
Optimal Policy for State: (3, 1) --- [1]
[17.80172913763953, 17.80172913763953, 14.419381533113208, 14.419381533113208]
Optimal Policy for State: (3, 2) --- [0, 1]
[16.021535036792454, 16.021535036792454, 12.977443379801887, 12.977443379801887]
Optimal Policy for State: (3, 3) --- [0, 1]
[14.419381533113208, 14.419381533113208, 0, 11.6796990418217]
Optimal Policy for State: (3, 4) --- [0, 1]
[0, 16.021535036792454, 16.021535036792454, 0]
Optimal Policy for State: (4, 0) --- [1, 2]
[14.419381533113208, 17.80172913763953, 14.419381533113208, 0]
Optimal Policy for State: (4, 1) --- [1]
[16.021535036792454, 16.021535036792454, 12.977443379801887, 0]
Optimal Policy for State: (4, 2) --- [0, 1]
[14.419381533113208, 14.419381533113208, 11.6796990418217, 0]
Optimal Policy for State: (4, 3) --- [0, 1]
[12.977443379801887, 12.977443379801887, 0, 0]
Optimal Policy for State: (4, 4) --- [0, 1]

Answer 6

Value iteration and Policy iteration

Iteration - 1
[[0. -1. -1. -1.]
 [-1. -1. -1. -1.]
 [-1. -1. -1. -1.]
 [-1. -1. -1. 0.]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0, 1, 2, 3]
Optimal Policy for State: (0, 3) --- [0, 1, 2, 3]
Optimal Policy for State: (1, 0) --- [1]

```
Optimal Policy for State: (1, 1) --- [0, 1, 2, 3]
Optimal Policy for State: (1, 2) --- [0, 1, 2, 3]
Optimal Policy for State: (1, 3) --- [0, 1, 2, 3]
Optimal Policy for State: (2, 0) --- [0, 1, 2, 3]
Optimal Policy for State: (2, 1) --- [0, 1, 2, 3]
Optimal Policy for State: (2, 2) --- [0, 1, 2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [0, 1, 2, 3]
Optimal Policy for State: (3, 1) --- [0, 1, 2, 3]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  5
[[ 0.   -1.75 -2.   -2.  ]
 [-1.75 -2.   -2.   -2.  ]
 [-2.   -2.   -2.   -1.75]
 [-2.   -2.   -1.75  0.  ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 1, 2, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 1, 2, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [0, 1, 2, 3]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [0, 1, 2, 3]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  9
[[ 0.     -2.4375 -2.9375 -3.    ]
 [-2.4375 -2.875  -3.     -2.9375]
 [-2.9375 -3.     -2.875  -2.4375]
 [-3.     -2.9375 -2.4375  0.    ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
```

```
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  13
[[ 0.      -3.0625  -3.84375 -3.96875]
 [-3.0625  -3.71875 -3.90625 -3.84375]
 [-3.84375 -3.90625 -3.71875 -3.0625 ]
 [-3.96875 -3.84375 -3.0625   0.     ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  17
[[ 0.        -3.65625   -4.6953125 -4.90625  ]
 [-3.65625   -4.484375  -4.78125   -4.6953125]
 [-4.6953125 -4.78125   -4.484375  -3.65625  ]
 [-4.90625   -4.6953125 -3.65625    0.       ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
```

```
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  21
[[ 0.         -4.20898438 -5.50976562 -5.80078125]
 [-4.20898438 -5.21875    -5.58984375 -5.50976562]
 [-5.50976562 -5.58984375 -5.21875    -4.20898438]
 [-5.80078125 -5.50976562 -4.20898438  0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  25
[[ 0.         -4.734375   -6.27734375 -6.65527344]
 [-4.734375   -5.89941406 -6.36425781 -6.27734375]
 [-6.27734375 -6.36425781 -5.89941406 -4.734375  ]
 [-6.65527344 -6.27734375 -4.734375    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
```

```
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  29
[[ 0.         -5.2277832  -7.0078125  -7.46630859]
 [-5.2277832  -6.54931641 -7.08837891 -7.0078125 ]
 [-7.0078125  -7.08837891 -6.54931641 -5.2277832 ]
 [-7.46630859 -7.0078125  -5.2277832   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  33
[[ 0.         -5.69622803 -7.6975708  -8.23706055]
 [-5.69622803 -7.15808105 -7.77856445 -7.6975708 ]
 [-7.6975708  -7.77856445 -7.15808105 -5.69622803]
 [-8.23706055 -7.6975708  -5.69622803  0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
```

```
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  37
[[ 0.         -6.13796997 -8.35235596 -8.96731567]
 [-6.13796997 -7.73739624 -8.42782593 -8.35235596]
 [-8.35235596 -8.42782593 -7.73739624 -6.13796997]
 [-8.96731567 -8.35235596 -6.13796997  0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  41
[[ 0.         -6.55693054 -8.97136688 -9.65983582]
 [-6.55693054 -8.28289795 -9.0448761  -8.97136688]
 [-8.97136688 -9.0448761  -8.28289795 -6.55693054]
 [-9.65983582 -8.97136688 -6.55693054  0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
```

```
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  45
[[  0.          -6.95279884  -9.55825233 -10.31560135]
 [ -6.95279884  -8.80090332  -9.62713242  -9.55825233]
 [ -9.55825233  -9.62713242  -8.80090332  -6.95279884]
 [-10.31560135  -9.55825233  -6.95279884   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  49
[[  0.          -7.32798862 -10.11344624 -10.93692684]
 [ -7.32798862  -9.28996563 -10.17957783 -10.11344624]
 [-10.11344624 -10.17957783  -9.28996563  -7.32798862]
 [-10.93692684 -10.11344624  -7.32798862   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
```

```
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  53
[[  0.          -7.68285012 -10.63948488 -11.52518654]
 [ -7.68285012  -9.75378323 -10.70170593 -10.63948488]
 [-10.63948488 -10.70170593  -9.75378323  -7.68285012]
 [-11.52518654 -10.63948488  -7.68285012   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  57
[[  0.          -8.01902956 -11.13730687 -12.08233571]
 [ -8.01902956 -10.19227803 -11.19663405 -11.13730687]
 [-11.13730687 -11.19663405 -10.19227803  -8.01902956]
 [-12.08233571 -11.13730687  -8.01902956   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
```

```
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  61
[[  0.          -8.33715361 -11.60882655 -12.60982129]
 [ -8.33715361 -10.60783181 -11.66479245 -11.60882655]
 [-11.60882655 -11.66479245 -10.60783181  -8.33715361]
 [-12.60982129 -11.60882655  -8.33715361   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  65
[[  0.          -8.63845299 -12.05514847 -13.10932392]
 [ -8.63845299 -11.00097303 -12.10832918 -12.05514847]
 [-12.05514847 -12.10832918 -11.00097303  -8.63845299]
 [-13.10932392 -12.05514847  -8.63845299   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
```

```
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  69
[[  0.          -8.92364362 -12.47781364 -13.5822362 ]
 [ -8.92364362 -11.37339108 -12.52806075 -12.47781364]
 [-12.47781364 -12.52806075 -11.37339108  -8.92364362]
 [-13.5822362  -12.47781364  -8.92364362   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  73
[[  0.          -9.19371209 -12.87793855 -14.03002492]
 [ -9.19371209 -11.72585219 -12.92560236 -12.87793855]
 [-12.87793855 -12.92560236 -11.72585219  -9.19371209]
 [-14.03002492 -12.87793855  -9.19371209   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
```

```
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  77
[[  0.          -9.44937571 -13.25681948 -14.45398174]
 [ -9.44937571 -12.05965722 -13.30189537 -13.25681948]
 [-13.25681948 -13.30189537 -12.05965722  -9.44937571]
 [-14.45398174 -13.25681948  -9.44937571   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  81
[[  0.          -9.6914631  -13.61551807 -14.85540061]
 [ -9.6914631  -12.37563554 -13.65823835 -13.61551807]
 [-13.61551807 -13.65823835 -12.37563554  -9.6914631 ]
 [-14.85540061 -13.61551807  -9.6914631    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
```

```
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  85
[[  0.          -9.92065418 -13.95515503 -15.23545934]
 [ -9.92065418 -12.67485073 -13.99557681 -13.95515503]
 [-13.95515503 -13.99557681 -12.67485073  -9.92065418]
 [-15.23545934 -13.95515503  -9.92065418   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  89
[[  0.         -10.13766499 -14.27671134 -15.59530719]
 [-10.13766499 -12.95811549 -14.31500288 -14.27671134]
 [-14.27671134 -14.31500288 -12.95811549 -10.13766499]
 [-15.59530719 -14.27671134 -10.13766499   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
```

```
Iteration -  93
[[  0.         -10.34312295 -14.5811716  -15.93600926]
 [-10.34312295 -13.22633393 -14.61741342 -14.5811716 ]
 [-14.5811716  -14.61741342 -13.22633393 -10.34312295]
 [-15.93600926 -14.5811716  -10.34312295   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  97
[[  0.         -10.53765712 -14.86942931 -16.25859043]
 [-10.53765712 -13.48026819 -14.90375277 -14.86942931]
 [-14.86942931 -14.90375277 -13.48026819 -10.53765712]
 [-16.25859043 -14.86942931 -10.53765712   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  101
```

```
[[  0.         -10.72183865 -15.14235741 -16.56400987]
 [-10.72183865 -13.72070494 -15.17484875 -15.14235741]
 [-15.14235741 -15.17484875 -13.72070494 -10.72183865]
 [-16.56400987 -15.14235741 -10.72183865   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  105
[[  0.         -10.89622525 -15.40076367 -16.85318364]
 [-10.89622525 -13.9483437  -15.43153118 -15.40076367]
 [-15.40076367 -15.43153118 -13.9483437  -10.89622525]
 [-16.85318364 -15.40076367 -10.89622525   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  109
[[  0.         -11.06133316 -15.64542593 -17.12697365]
```

```
 [-11.06133316 -14.16387821 -15.67455368 -15.64542593]
 [-15.64542593 -15.67455368 -14.16387821 -11.06133316]
 [-17.12697365 -15.64542593 -11.06133316   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  113
[[  0.         -11.21765933 -15.87707161 -17.38619979]
 [-11.21765933 -14.36794342 -15.90465207 -15.87707161]
 [-15.87707161 -15.90465207 -14.36794342 -11.21765933]
 [-17.38619979 -15.87707161 -11.21765933   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  117
[[  0.         -11.36566859 -16.0963957  -17.6316357 ]
 [-11.36566859 -14.5611557  -16.12250751 -16.0963957 ]
```

```
 [-16.0963957  -16.12250751 -14.5611557  -11.36566859]
 [-17.6316357  -16.0963957  -11.36566859   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  121
[[  0.         -11.505805   -16.30405188 -17.8640157 ]
 [-11.505805   -14.74408805 -16.3287757  -16.30405188]
 [-16.30405188 -16.3287757  -14.74408805 -11.505805  ]
 [-17.8640157  -16.30405188 -11.505805     0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  125
[[  0.         -11.63848623 -16.50066207 -18.08403379]
 [-11.63848623 -14.91729035 -16.52406996 -16.50066207]
 [-16.50066207 -16.52406996 -14.91729035 -11.63848623]
```

```
 [-18.08403379 -16.50066207 -11.63848623   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  129
[[  0.         -11.76410966 -16.68681301 -18.29234793]
 [-11.76410966 -15.0812781  -16.70897621 -16.68681301]
 [-16.68681301 -16.70897621 -15.0812781  -11.76410966]
 [-18.29234793 -16.68681301 -11.76410966   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  133
[[  0.         -11.88305019 -16.8630617  -18.48958047]
 [-11.88305019 -15.23654293 -16.88404555 -16.8630617 ]
 [-16.8630617  -16.88404555 -15.23654293 -11.88305019]
 [-18.48958047 -16.8630617  -11.88305019   0.        ]]
```

```
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  137
[[  0.         -11.99566371 -17.02993448 -18.67632109]
 [-11.99566371 -15.38354787 -17.04980232 -17.02993448]
 [-17.02993448 -17.04980232 -15.38354787 -11.99566371]
 [-18.67632109 -17.02993448 -11.99566371   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  141
[[  0.         -12.10228652 -17.1879304  -18.85312778]
 [-12.10228652 -15.52273301 -17.20674118 -17.1879304 ]
 [-17.1879304  -17.20674118 -15.52273301 -12.10228652]
 [-18.85312778 -17.1879304  -12.10228652   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
```

```
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  145
[[  0.         -12.20323748 -17.33752147 -19.02052909]
 [-12.20323748 -15.65451385 -17.35533171 -17.33752147]
 [-17.33752147 -17.35533171 -15.65451385 -12.20323748]
 [-19.02052909 -17.33752147 -12.20323748   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  149
[[  0.         -12.2988182  -17.47915494 -19.17902528]
 [-12.2988182  -15.77928459 -17.49601766 -17.47915494]
 [-17.47915494 -17.49601766 -15.77928459 -12.2988182 ]
 [-19.17902528 -17.47915494 -12.2988182    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
```

```
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  153
[[  0.         -12.38931443 -17.61325402 -19.32909011]
 [-12.38931443 -15.89741793 -17.62921977 -17.61325402]
 [-17.61325402 -17.62921977 -15.89741793 -12.38931443]
 [-19.32909011 -17.61325402 -12.38931443   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  157
[[  0.         -12.47499659 -17.74021958 -19.47117206]
 [-12.47499659 -16.0092671  -17.75533597 -17.74021958]
 [-17.74021958 -17.75533597 -16.0092671  -12.47499659]
 [-19.47117206 -17.74021958 -12.47499659   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
```

```
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  161
[[  0.         -12.55612082 -17.86043105 -19.60569582]
 [-12.55612082 -16.11516628 -17.87474334 -17.86043105]
 [-17.86043105 -17.87474334 -16.11516628 -12.55612082]
 [-19.60569582 -17.86043105 -12.55612082   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  165
[[  0.         -12.63292954 -17.97424776 -19.73306344]
 [-12.63292954 -16.21543208 -17.98779867 -17.97424776]
 [-17.97424776 -17.98779867 -16.21543208 -12.63292954]
 [-19.73306344 -17.97424776 -12.63292954   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
```

```
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  169
[[  0.         -12.70565234 -18.08200985 -19.8536556 ]
 [-12.70565234 -16.3103641  -18.09483992 -18.08200985]
 [-18.08200985 -18.09483992 -16.3103641  -12.70565234]
 [-19.8536556  -18.08200985 -12.70565234   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  173
[[  0.         -12.77450657 -18.18403943 -19.96783272]
 [-12.77450657 -16.40024613 -18.19618698 -18.18403943]
 [-18.18403943 -18.19618698 -16.40024613 -12.77450657]
 [-19.96783272 -18.18403943 -12.77450657   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
```

```
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  177
[[  0.         -12.83969803 -18.28064143 -20.07593608]
 [-12.83969803 -16.48534678 -18.29214278 -18.28064143]
 [-18.28064143 -18.29214278 -16.48534678 -12.83969803]
 [-20.07593608 -18.28064143 -12.83969803   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  181
[[  0.         -12.90142156 -18.37210458 -20.17828875]
 [-12.90142156 -16.56592041 -18.3829941  -18.37210458]
 [-18.37210458 -18.3829941  -16.56592041 -12.90142156]
 [-20.17828875 -18.37210458 -12.90142156   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
```

```
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  185
[[  0.         -12.95986164 -18.45870225 -20.27519666]
 [-12.95986164 -16.64220783 -18.46901249 -18.45870225]
 [-18.45870225 -18.46901249 -16.64220783 -12.95986164]
 [-20.27519666 -18.45870225 -12.95986164   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  189
[[  0.         -13.01519293 -18.54069326 -20.36694946]
 [-13.01519293 -16.71443706 -18.55045504 -18.54069326]
 [-18.54069326 -18.55045504 -16.71443706 -13.01519293]
 [-20.36694946 -18.54069326 -13.01519293   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
```

```
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  193
[[  0.         -13.06758081 -18.61832267 -20.45382136]
 [-13.06758081 -16.78282398 -18.62756516 -18.61832267]
 [-18.61832267 -18.62756516 -16.78282398 -13.06758081]
 [-20.45382136 -18.61832267 -13.06758081   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  197
[[  0.         -13.11718187 -18.6918225  -20.53607201]
 [-13.11718187 -16.84757299 -18.70057333 -18.6918225 ]
 [-18.6918225  -18.70057333 -16.84757299 -13.11718187]
 [-20.53607201 -18.6918225  -13.11718187   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
```

```
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  201
[[  0.         -13.16414434 -18.76141243 -20.61394726]
 [-13.16414434 -16.9088776  -18.76969774 -18.76141243]
 [-18.76141243 -18.76969774 -16.9088776  -13.16414434]
 [-20.61394726 -18.76141243 -13.16414434   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  205
[[  0.         -13.20860859 -18.82730044 -20.68767984]
 [-13.20860859 -16.96692104 -18.83514501 -18.82730044]
 [-18.82730044 -18.83514501 -16.96692104 -13.20860859]
 [-20.68767984 -18.82730044 -13.20860859   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
```

```
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  209
[[  0.         -13.25070752 -18.88968347 -20.75749014]
 [-13.25070752 -17.0218768  -18.89711074 -18.88968347]
 [-18.88968347 -18.89711074 -17.0218768  -13.25070752]
 [-20.75749014 -18.88968347 -13.25070752   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  213
[[  0.         -13.29056695 -18.94874797 -20.82358681]
 [-13.29056695 -17.07390913 -18.95578014 -18.94874797]
 [-18.94874797 -18.95578014 -17.07390913 -13.29056695]
 [-20.82358681 -18.94874797 -13.29056695   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
```

```
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  217
[[  0.         -13.32830601 -19.00467047 -20.88616739]
 [-13.32830601 -17.12317354 -19.01132855 -19.00467047]
 [-19.00467047 -19.01132855 -17.12317354 -13.32830601]
 [-20.88616739 -19.00467047 -13.32830601   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  221
[[  0.         -13.3640375  -19.0576181  -20.94541893]
 [-13.3640375  -17.16981728 -19.063922   -19.0576181 ]
 [-19.0576181  -19.063922   -17.16981728 -13.3640375 ]
 [-20.94541893 -19.0576181  -13.3640375    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
```

```
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  225
[[  0.         -13.39786822 -19.10774913 -21.00151852]
 [-13.39786822 -17.21397975 -19.11371769 -19.10774913]
 [-19.10774913 -19.11371769 -17.21397975 -13.39786822]
 [-21.00151852 -19.10774913 -13.39786822   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  229
[[  0.         -13.42989928 -19.15521339 -21.05463382]
 [-13.42989928 -17.25579296 -19.16086444 -19.15521339]
 [-19.15521339 -19.16086444 -17.25579296 -13.42989928]
 [-21.05463382 -19.15521339 -13.42989928   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
```

```
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  233
[[  0.         -13.46022641 -19.20015273 -21.10492361]
 [-13.46022641 -17.29538186 -19.20550317 -19.20015273]
 [-19.20015273 -19.20550317 -17.29538186 -13.46022641]
 [-21.10492361 -19.20015273 -13.46022641   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  237
[[  0.         -13.48894025 -19.24270148 -21.15253817]
 [-13.48894025 -17.33286479 -19.2477673  -19.24270148]
 [-19.24270148 -19.2477673  -17.33286479 -13.48894025]
 [-21.15253817 -19.24270148 -13.48894025   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
```

```
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  241
[[  0.         -13.51612663 -19.2829868  -21.19761983]
 [-13.51612663 -17.36835377 -19.28778314 -19.2829868 ]
 [-19.2829868  -19.28778314 -17.36835377 -13.51612663]
 [-21.19761983 -19.2829868  -13.51612663   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  245
[[  0.         -13.5418668  -19.3211291  -21.24030331]
 [-13.5418668  -17.40195488 -19.32567029 -19.3211291 ]
 [-19.3211291  -19.32567029 -17.40195488 -13.5418668 ]
 [-21.24030331 -19.3211291  -13.5418668    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
```

```
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  249
[[  0.         -13.5662377  -19.35724237 -21.28071621]
 [-13.5662377  -17.43376854 -19.36154199 -19.35724237]
 [-19.35724237 -19.36154199 -17.43376854 -13.5662377 ]
 [-21.28071621 -19.35724237 -13.5662377    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  253
[[  0.         -13.58931215 -19.39143457 -21.31897929]
 [-13.58931215 -17.46388984 -19.39550546 -19.39143457]
 [-19.39143457 -19.39550546 -17.46388984 -13.58931215]
 [-21.31897929 -19.39143457 -13.58931215   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
```

```
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  257
[[  0.         -13.61115914 -19.42380787 -21.35520693]
 [-13.61115914 -17.49240881 -19.4276622  -19.42380787]
 [-19.42380787 -19.4276622  -17.49240881 -13.61115914]
 [-21.35520693 -19.42380787 -13.61115914   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  261
[[  0.         -13.63184395 -19.45445904 -21.3895074 ]
 [-13.63184395 -17.51941067 -19.45810834 -19.45445904]
 [-19.45445904 -19.45810834 -17.51941067 -13.63184395]
 [-21.3895074  -19.45445904 -13.63184395   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
```

```
Iteration -  265
[[  0.         -13.65142842 -19.48347968 -21.42198322]
 [-13.65142842 -17.54497615 -19.48693485 -19.48347968]
 [-19.48347968 -19.48693485 -17.54497615 -13.65142842]
 [-21.42198322 -19.48347968 -13.65142842   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  269
[[  0.         -13.66997106 -19.51095654 -21.45273145]
 [-13.66997106 -17.56918163 -19.51422791 -19.51095654]
 [-19.51095654 -19.51422791 -17.56918163 -13.66997106]
 [-21.45273145 -19.51095654 -13.66997106   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  273
```

```
[[  0.         -13.68752731 -19.53697174 -21.481844  ]
 [-13.68752731 -17.59209949 -19.54006909 -19.53697174]
 [-19.53697174 -19.54006909 -17.59209949 -13.68752731]
 [-21.481844   -19.53697174 -13.68752731   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  277
[[  0.         -13.70414963 -19.56160303 -21.50940787]
 [-13.70414963 -17.6137982  -19.56453561 -19.56160303]
 [-19.56160303 -19.56453561 -17.6137982  -13.70414963]
 [-21.50940787 -19.56160303 -13.70414963   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  281
[[  0.         -13.71988772 -19.58492404 -21.53550545]
```

```
 [-13.71988772 -17.63434262 -19.58770062 -19.58492404]
 [-19.58492404 -19.58770062 -17.63434262 -13.71988772]
 [-21.53550545 -19.58492404 -13.71988772   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  285
[[  0.         -13.73478859 -19.60700446 -21.56021474]
 [-13.73478859 -17.65379417 -19.60963333 -19.60700446]
 [-19.60700446 -19.60963333 -17.65379417 -13.73478859]
 [-21.56021474 -19.60700446 -13.73478859   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  289
[[  0.         -13.7488968  -19.62791028 -21.5836096 ]
 [-13.7488968  -17.67221096 -19.63039931 -19.62791028]
```

```
 [-19.62791028 -19.63039931 -17.67221096 -13.7488968 ]
 [-21.5836096  -19.62791028 -13.7488968    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  293
[[  0.         -13.76225451 -19.647704   -21.60575994]
 [-13.76225451 -17.68964806 -19.65006062 -19.647704  ]
 [-19.647704   -19.65006062 -17.68964806 -13.76225451]
 [-21.60575994 -19.647704   -13.76225451   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  297
[[  0.         -13.77490164 -19.66644477 -21.62673197]
 [-13.77490164 -17.70615757 -19.66867603 -19.66644477]
 [-19.66644477 -19.66867603 -17.70615757 -13.77490164]
```

```
 [-21.62673197 -19.66644477 -13.77490164   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  301
[[  0.         -13.78687599 -19.6841886  -21.64658837]
 [-13.78687599 -17.72178884 -19.68630117 -19.6841886 ]
 [-19.6841886  -19.68630117 -17.72178884 -13.78687599]
 [-21.64658837 -19.6841886  -13.78687599   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  305
[[  0.         -13.79821336 -19.70098853 -21.66538849]
 [-13.79821336 -17.73658858 -19.70298872 -19.70098853]
 [-19.70098853 -19.70298872 -17.73658858 -13.79821336]
 [-21.66538849 -19.70098853 -13.79821336   0.        ]]
```

```
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  309
[[  0.         -13.80894762 -19.71689477 -21.68318851]
 [-13.80894762 -17.75060104 -19.71878856 -19.71689477]
 [-19.71689477 -19.71878856 -17.75060104 -13.80894762]
 [-21.68318851 -19.71689477 -13.80894762   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  313
[[  0.         -13.81911086 -19.73195486 -21.70004164]
 [-13.81911086 -17.76386809 -19.73374791 -19.73195486]
 [-19.73195486 -19.73374791 -17.76386809 -13.81911086]
 [-21.70004164 -19.73195486 -13.81911086   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
```

```
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  317
[[  0.         -13.82873345 -19.74621382 -21.71599825]
 [-13.82873345 -17.77642938 -19.74791148 -19.74621382]
 [-19.74621382 -19.74791148 -17.77642938 -13.82873345]
 [-21.71599825 -19.74621382 -13.82873345   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  321
[[  0.         -13.83784416 -19.75971425 -21.73110604]
 [-13.83784416 -17.78832246 -19.7613216  -19.75971425]
 [-19.75971425 -19.7613216  -17.78832246 -13.83784416]
 [-21.73110604 -19.75971425 -13.83784416   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
```

```
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  325
[[  0.         -13.84647022 -19.77249651 -21.74541014]
 [-13.84647022 -17.79958288 -19.77401836 -19.77249651]
 [-19.77249651 -19.77401836 -17.79958288 -13.84647022]
 [-21.74541014 -19.77249651 -13.84647022   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  329
[[  0.         -13.8546374  -19.78459881 -21.75895333]
 [-13.8546374  -17.81024429 -19.7860397  -19.78459881]
 [-19.78459881 -19.7860397  -17.81024429 -13.8546374 ]
 [-21.75895333 -19.78459881 -13.8546374    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
```

```
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  333
[[  0.         -13.86237012 -19.79605731 -21.77177607]
 [-13.86237012 -17.82033855 -19.79742155 -19.79605731]
 [-19.79605731 -19.79742155 -17.82033855 -13.86237012]
 [-21.77177607 -19.79605731 -13.86237012   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  337
[[  0.         -13.8696915  -19.80690626 -21.78391669]
 [-13.8696915  -17.82989584 -19.80819793 -19.80690626]
 [-19.80690626 -19.80819793 -17.82989584 -13.8696915 ]
 [-21.78391669 -19.80690626 -13.8696915    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
```

```
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  341
[[  0.         -13.8766234  -19.81717809 -21.79541148]
 [-13.8766234  -17.83894471 -19.81840105 -19.81717809]
 [-19.81717809 -19.81840105 -17.83894471 -13.8766234 ]
 [-21.79541148 -19.81717809 -13.8766234    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  345
[[  0.         -13.88318655 -19.8269035  -21.80629478]
 [-13.88318655 -17.84751222 -19.8280614  -19.8269035 ]
 [-19.8269035  -19.8280614  -17.84751222 -13.88318655]
 [-21.80629478 -19.8269035  -13.88318655   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
```

```
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  349
[[  0.         -13.88940057 -19.83611156 -21.81659914]
 [-13.88940057 -17.85562398 -19.83720786 -19.83611156]
 [-19.83611156 -19.83720786 -17.85562398 -13.88940057]
 [-21.81659914 -19.83611156 -13.88940057   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  353
[[  0.         -13.89528403 -19.84482978 -21.82635535]
 [-13.89528403 -17.86330422 -19.84586777 -19.84482978]
 [-19.84482978 -19.84586777 -17.86330422 -13.89528403]
 [-21.82635535 -19.84482978 -13.89528403   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
```

```
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  357
[[  0.         -13.90085451 -19.85308423 -21.83559257]
 [-13.90085451 -17.8705759  -19.854067   -19.85308423]
 [-19.85308423 -19.854067   -17.8705759  -13.90085451]
 [-21.83559257 -19.85308423 -13.90085451   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  361
[[  0.         -13.90612866 -19.86089958 -21.8443384 ]
 [-13.90612866 -17.87746075 -19.86183007 -19.86089958]
 [-19.86089958 -19.86183007 -17.87746075 -13.90612866]
 [-21.8443384  -19.86089958 -13.90612866   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
```

```
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  365
[[  0.         -13.91112225 -19.86829918 -21.85261899]
 [-13.91112225 -17.88397936 -19.86918017 -19.86829918]
 [-19.86829918 -19.86918017 -17.88397936 -13.91112225]
 [-21.85261899 -19.86829918 -13.91112225   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  369
[[  0.         -13.9158502  -19.87530514 -21.86045908]
 [-13.9158502  -17.89015121 -19.87613927 -19.87530514]
 [-19.87530514 -19.87613927 -17.89015121 -13.9158502 ]
 [-21.86045908 -19.87530514 -13.9158502    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
```

```
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  373
[[  0.         -13.92032664 -19.88193842 -21.86788211]
 [-13.92032664 -17.89599473 -19.88272818 -19.88193842]
 [-19.88193842 -19.88272818 -17.89599473 -13.92032664]
 [-21.86788211 -19.88193842 -13.92032664   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  377
[[  0.         -13.92456495 -19.88821884 -21.87491027]
 [-13.92456495 -17.90152741 -19.88896658 -19.88821884]
 [-19.88821884 -19.88896658 -17.90152741 -13.92456495]
 [-21.87491027 -19.88821884 -13.92456495   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
```

```
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  381
[[  0.         -13.9285778  -19.89416516 -21.88156455]
 [-13.9285778  -17.90676576 -19.89487312 -19.89416516]
 [-19.89416516 -19.89487312 -17.90676576 -13.9285778 ]
 [-21.88156455 -19.89416516 -13.9285778    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  385
[[  0.         -13.93237718 -19.89979516 -21.88786486]
 [-13.93237718 -17.91172546 -19.90046546 -19.89979516]
 [-19.89979516 -19.90046546 -17.91172546 -13.93237718]
 [-21.88786486 -19.89979516 -13.93237718   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
```

```
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  389
[[  0.         -13.93597445 -19.90512566 -21.89383001]
 [-13.93597445 -17.91642132 -19.90576031 -19.90512566]
 [-19.90512566 -19.90576031 -17.91642132 -13.93597445]
 [-21.89383001 -19.90512566 -13.93597445   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  393
[[  0.         -13.93938036 -19.91017261 -21.89947784]
 [-13.93938036 -17.92086738 -19.91077349 -19.91017261]
 [-19.91017261 -19.91077349 -17.92086738 -13.93938036]
 [-21.89947784 -19.91017261 -13.93938036   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
```

```
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  397
[[  0.         -13.94260509 -19.91495107 -21.90482522]
 [-13.94260509 -17.92507693 -19.91551999 -19.91495107]
 [-19.91495107 -19.91551999 -17.92507693 -13.94260509]
 [-21.90482522 -19.91495107 -13.94260509   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  401
[[  0.         -13.94565827 -19.91947534 -21.90988815]
 [-13.94565827 -17.92906254 -19.920014   -19.91947534]
 [-19.91947534 -19.920014   -17.92906254 -13.94565827]
 [-21.90988815 -19.91947534 -13.94565827   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
```

```
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  405
[[  0.         -13.94854904 -19.92375894 -21.91468175]
 [-13.94854904 -17.93283614 -19.92426894 -19.92375894]
 [-19.92375894 -19.92426894 -17.93283614 -13.94854904]
 [-21.91468175 -19.92375894 -13.94854904   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  409
[[  0.         -13.95128603 -19.92781467 -21.91922034]
 [-13.95128603 -17.93640899 -19.92829754 -19.92781467]
 [-19.92781467 -19.92829754 -17.93640899 -13.95128603]
 [-21.91922034 -19.92781467 -13.95128603   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
```

```
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  413
[[  0.         -13.95387742 -19.93165464 -21.9235175 ]
 [-13.95387742 -17.93979178 -19.93211183 -19.93165464]
 [-19.93165464 -19.93211183 -17.93979178 -13.95387742]
 [-21.9235175  -19.93165464 -13.95387742   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  417
[[  0.         -13.95633096 -19.93529035 -21.92758607]
 [-13.95633096 -17.94299462 -19.93572321 -19.93529035]
 [-19.93529035 -19.93572321 -17.94299462 -13.95633096]
 [-21.92758607 -19.93529035 -13.95633096   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
```

```
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  421
[[  0.         -13.95865398 -19.93873265 -21.93143821]
 [-13.95865398 -17.94602709 -19.93914249 -19.93873265]
 [-19.93873265 -19.93914249 -17.94602709 -13.95865398]
 [-21.93143821 -19.93873265 -13.95865398   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  425
[[  0.         -13.96085343 -19.94199183 -21.93508543]
 [-13.96085343 -17.94889824 -19.94237987 -19.94199183]
 [-19.94199183 -19.94237987 -17.94889824 -13.96085343]
 [-21.93508543 -19.94199183 -13.96085343   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
```

```
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  429
[[  0.         -13.96293587 -19.94507764 -21.93853863]
 [-13.96293587 -17.95161665 -19.94544503 -19.94507764]
 [-19.94507764 -19.94544503 -17.95161665 -13.96293587]
 [-21.93853863 -19.94507764 -13.96293587   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  433
[[  0.         -13.96490754 -19.9479993  -21.94180814]
 [-13.96490754 -17.95419045 -19.94834715 -19.9479993 ]
 [-19.9479993  -19.94834715 -17.95419045 -13.96490754]
 [-21.94180814 -19.9479993  -13.96490754   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
```

```
Iteration -  437
[[  0.         -13.96677432 -19.95076553 -21.94490372]
 [-13.96677432 -17.95662734 -19.95109487 -19.95076553]
 [-19.95076553 -19.95109487 -17.95662734 -13.96677432]
 [-21.94490372 -19.95076553 -13.96677432   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  441
[[  0.         -13.9685418  -19.95338461 -21.94783462]
 [-13.9685418  -17.9589346  -19.95369644 -19.95338461]
 [-19.95338461 -19.95369644 -17.9589346  -13.9685418 ]
 [-21.94783462 -19.95338461 -13.9685418    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  445
```

```
[[  0.         -13.97021525 -19.95586437 -21.95060962]
 [-13.97021525 -17.96111912 -19.9561596  -19.95586437]
 [-19.95586437 -19.9561596  -17.96111912 -13.97021525]
 [-21.95060962 -19.95586437 -13.97021525   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  449
[[  0.         -13.97179968 -19.95821221 -21.95323699]
 [-13.97179968 -17.96318743 -19.95849174 -19.95821221]
 [-19.95821221 -19.95849174 -17.96318743 -13.97179968]
 [-21.95323699 -19.95821221 -13.97179968   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  453
[[  0.         -13.97329983 -19.96043516 -21.9557246 ]
```

```
 [-13.97329983 -17.96514571 -19.96069982 -19.96043516]
 [-19.96043516 -19.96069982 -17.96514571 -13.97329983]
 [-21.9557246  -19.96043516 -13.97329983   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  457
[[  0.         -13.97472018 -19.96253985 -21.95807988]
 [-13.97472018 -17.96699983 -19.96279044 -19.96253985]
 [-19.96253985 -19.96279044 -17.96699983 -13.97472018]
 [-21.95807988 -19.96253985 -13.97472018   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  461
[[  0.         -13.97606496 -19.96453259 -21.96030987]
 [-13.97606496 -17.96875531 -19.96476984 -19.96453259]
```

```
 [-19.96453259 -19.96476984 -17.96875531 -13.97606496]
 [-21.96030987 -19.96453259 -13.97606496   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  465
[[  0.         -13.97733821 -19.96641931 -21.96242123]
 [-13.97733821 -17.9704174  -19.96664395 -19.96641931]
 [-19.96641931 -19.96664395 -17.9704174  -13.97733821]
 [-21.96242123 -19.96641931 -13.97733821   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  469
[[  0.         -13.97854373 -19.96820567 -21.96442027]
 [-13.97854373 -17.97199108 -19.96841836 -19.96820567]
 [-19.96820567 -19.96841836 -17.97199108 -13.97854373]
```

```
 [-21.96442027 -19.96820567 -13.97854373   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  473
[[  0.         -13.97968512 -19.96989701 -21.96631297]
 [-13.97968512 -17.97348104 -19.97009838 -19.96989701]
 [-19.96989701 -19.97009838 -17.97348104 -13.97968512]
 [-21.96631297 -19.96989701 -13.97968512   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  477
[[  0.         -13.98076579 -19.97149837 -21.96810499]
 [-13.98076579 -17.97489175 -19.97168903 -19.97149837]
 [-19.97149837 -19.97168903 -17.97489175 -13.98076579]
 [-21.96810499 -19.97149837 -13.98076579   0.        ]]
```

```
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  481
[[  0.         -13.98178898 -19.97301454 -21.96980168]
 [-13.98178898 -17.97622741 -19.97319506 -19.97301454]
 [-19.97301454 -19.97319506 -17.97622741 -13.98178898]
 [-21.96980168 -19.97301454 -13.98178898   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  485
[[  0.         -13.98275773 -19.97445007 -21.97140811]
 [-13.98275773 -17.97749202 -19.97462098 -19.97445007]
 [-19.97445007 -19.97462098 -17.97749202 -13.98275773]
 [-21.97140811 -19.97445007 -13.98275773   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
```

```
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  489
[[  0.         -13.98367495 -19.97580922 -21.97292909]
 [-13.98367495 -17.97868936 -19.97597104 -19.97580922]
 [-19.97580922 -19.97597104 -17.97868936 -13.98367495]
 [-21.97292909 -19.97580922 -13.98367495   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  493
[[  0.         -13.98454338 -19.97709608 -21.97436916]
 [-13.98454338 -17.979823   -19.97724929 -19.97709608]
 [-19.97709608 -19.97724929 -17.979823   -13.98454338]
 [-21.97436916 -19.97709608 -13.98454338   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
```

```
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  497
[[  0.         -13.98536561 -19.97831448 -21.97573262]
 [-13.98536561 -17.98089634 -19.97845954 -19.97831448]
 [-19.97831448 -19.97845954 -17.98089634 -13.98536561]
 [-21.97573262 -19.97831448 -13.98536561   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  501
[[  0.         -13.98614411 -19.97946806 -21.97702355]
 [-13.98614411 -17.98191258 -19.97960541 -19.97946806]
 [-19.97946806 -19.97960541 -17.98191258 -13.98614411]
 [-21.97702355 -19.97946806 -13.98614411   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
```

```
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  505
[[  0.         -13.98688119 -19.98056028 -21.9782458 ]
 [-13.98688119 -17.98287476 -19.98069032 -19.98056028]
 [-19.98056028 -19.98069032 -17.98287476 -13.98688119]
 [-21.9782458  -19.98056028 -13.98688119   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  509
[[  0.         -13.98757906 -19.9815944  -21.97940304]
 [-13.98757906 -17.98378575 -19.98171752 -19.9815944 ]
 [-19.9815944  -19.98171752 -17.98378575 -13.98757906]
 [-21.97940304 -19.9815944  -13.98757906   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
```

```
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  513
[[  0.         -13.9882398  -19.9825735  -21.98049872]
 [-13.9882398  -17.98464829 -19.98269007 -19.9825735 ]
 [-19.9825735  -19.98269007 -17.98464829 -13.9882398 ]
 [-21.98049872 -19.9825735  -13.9882398    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  517
[[  0.         -13.9888654  -19.98350052 -21.98153611]
 [-13.9888654  -17.98546494 -19.98361089 -19.98350052]
 [-19.98350052 -19.98361089 -17.98546494 -13.9888654 ]
 [-21.98153611 -19.98350052 -13.9888654    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
```

```
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  521
[[  0.         -13.98945772 -19.98437823 -21.98251832]
 [-13.98945772 -17.98623815 -19.98448273 -19.98437823]
 [-19.98437823 -19.98448273 -17.98623815 -13.98945772]
 [-21.98251832 -19.98437823 -13.98945772   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  525
[[  0.         -13.99001852 -19.98520925 -21.98344828]
 [-13.99001852 -17.98697022 -19.98530819 -19.98520925]
 [-19.98520925 -19.98530819 -17.98697022 -13.99001852]
 [-21.98344828 -19.98520925 -13.99001852   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
```

```
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  529
[[  0.         -13.9905495  -19.98599606 -21.98432876]
 [-13.9905495  -17.98766336 -19.98608974 -19.98599606]
 [-19.98599606 -19.98608974 -17.98766336 -13.9905495 ]
 [-21.98432876 -19.98599606 -13.9905495    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  533
[[  0.         -13.99105223 -19.98674101 -21.98516241]
 [-13.99105223 -17.98831962 -19.98682971 -19.98674101]
 [-19.98674101 -19.98682971 -17.98831962 -13.99105223]
 [-21.98516241 -19.98674101 -13.99105223   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
```

```
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  537
[[  0.         -13.99152822 -19.98744634 -21.98595171]
 [-13.99152822 -17.98894097 -19.98753032 -19.98744634]
 [-19.98744634 -19.98753032 -17.98894097 -13.99152822]
 [-21.98595171 -19.98744634 -13.99152822   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  541
[[  0.         -13.99197888 -19.98811415 -21.98669903]
 [-13.99197888 -17.98952927 -19.98819365 -19.98811415]
 [-19.98811415 -19.98819365 -17.98952927 -13.99197888]
 [-21.98669903 -19.98811415 -13.99197888   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
```

```
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  545
[[  0.         -13.99240557 -19.98874643 -21.98740659]
 [-13.99240557 -17.99008627 -19.98882171 -19.98874643]
 [-19.98874643 -19.98882171 -17.99008627 -13.99240557]
 [-21.98740659 -19.98874643 -13.99240557   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  549
[[  0.         -13.99280957 -19.98934507 -21.98807651]
 [-13.99280957 -17.99061364 -19.98941635 -19.98934507]
 [-19.98934507 -19.98941635 -17.99061364 -13.99280957]
 [-21.98807651 -19.98934507 -13.99280957   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
```

```
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  553
[[  0.         -13.99319207 -19.98991187 -21.98871079]
 [-13.99319207 -17.99111296 -19.98997936 -19.98991187]
 [-19.98991187 -19.98997936 -17.99111296 -13.99319207]
 [-21.98871079 -19.98991187 -13.99319207   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  557
[[  0.         -13.99355423 -19.99044852 -21.98931133]
 [-13.99355423 -17.99158571 -19.99051242 -19.99044852]
 [-19.99044852 -19.99051242 -17.99158571 -13.99355423]
 [-21.98931133 -19.99044852 -13.99355423   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
```

```
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  561
[[  0.         -13.99389712 -19.99095662 -21.98987993]
 [-13.99389712 -17.99203332 -19.99101712 -19.99095662]
 [-19.99095662 -19.99101712 -17.99203332 -13.99389712]
 [-21.98987993 -19.99095662 -13.99389712   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  565
[[  0.         -13.99422176 -19.9914377  -21.99041828]
 [-13.99422176 -17.99245712 -19.99149497 -19.9914377 ]
 [-19.9914377  -19.99149497 -17.99245712 -13.99422176]
 [-21.99041828 -19.9914377  -13.99422176   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
```

```
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  569
[[  0.         -13.99452914 -19.99189318 -21.99092799]
 [-13.99452914 -17.99285837 -19.99194741 -19.99189318]
 [-19.99189318 -19.99194741 -17.99285837 -13.99452914]
 [-21.99092799 -19.99189318 -13.99452914   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  573
[[  0.         -13.99482017 -19.99232443 -21.99141058]
 [-13.99482017 -17.99323828 -19.99237577 -19.99232443]
 [-19.99232443 -19.99237577 -17.99323828 -13.99482017]
 [-21.99141058 -19.99232443 -13.99482017   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
```

```
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  577
[[  0.         -13.99509572 -19.99273274 -21.9918675 ]
 [-13.99509572 -17.99359797 -19.99278135 -19.99273274]
 [-19.99273274 -19.99278135 -17.99359797 -13.99509572]
 [-21.9918675  -19.99273274 -13.99509572   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  581
[[  0.         -13.99535661 -19.99311933 -21.99230012]
 [-13.99535661 -17.99393854 -19.99316536 -19.99311933]
 [-19.99311933 -19.99316536 -17.99393854 -13.99535661]
 [-21.99230012 -19.99311933 -13.99535661   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
```

```
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  585
[[  0.         -13.99560362 -19.99348535 -21.99270973]
 [-13.99560362 -17.99426098 -19.99352893 -19.99348535]
 [-19.99348535 -19.99352893 -17.99426098 -13.99560362]
 [-21.99270973 -19.99348535 -13.99560362   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  589
[[  0.         -13.99583749 -19.99383191 -21.99309754]
 [-13.99583749 -17.99456627 -19.99387317 -19.99383191]
 [-19.99383191 -19.99387317 -17.99456627 -13.99583749]
 [-21.99309754 -19.99383191 -13.99583749   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
```

```
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  593
[[  0.         -13.99605892 -19.99416003 -21.99346472]
 [-13.99605892 -17.99485533 -19.99419909 -19.99416003]
 [-19.99416003 -19.99419909 -17.99485533 -13.99605892]
 [-21.99346472 -19.99416003 -13.99605892   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  597
[[  0.         -13.99626857 -19.99447069 -21.99381237]
 [-13.99626857 -17.995129   -19.99450768 -19.99447069]
 [-19.99447069 -19.99450768 -17.995129   -13.99626857]
 [-21.99381237 -19.99447069 -13.99626857   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
```

```
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  601
[[  0.         -13.99646707 -19.99476483 -21.99414153]
 [-13.99646707 -17.99538812 -19.99479985 -19.99476483]
 [-19.99476483 -19.99479985 -17.99538812 -13.99646707]
 [-21.99414153 -19.99476483 -13.99646707   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  605
[[  0.         -13.996655   -19.99504332 -21.99445318]
 [-13.996655   -17.99563346 -19.99507647 -19.99504332]
 [-19.99504332 -19.99507647 -17.99563346 -13.996655  ]
 [-21.99445318 -19.99504332 -13.996655     0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
```

```
Iteration -  609
[[  0.         -13.99683294 -19.99530699 -21.99474825]
 [-13.99683294 -17.99586574 -19.99533839 -19.99530699]
 [-19.99530699 -19.99533839 -17.99586574 -13.99683294]
 [-21.99474825 -19.99530699 -13.99683294   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  613
[[  0.         -13.99700142 -19.99555664 -21.99502762]
 [-13.99700142 -17.99608567 -19.99558637 -19.99555664]
 [-19.99555664 -19.99558637 -17.99608567 -13.99700142]
 [-21.99502762 -19.99555664 -13.99700142   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  617
```

```
[[  0.         -13.99716093 -19.99579301 -21.99529213]
 [-13.99716093 -17.99629389 -19.99582115 -19.99579301]
 [-19.99579301 -19.99582115 -17.99629389 -13.99716093]
 [-21.99529213 -19.99579301 -13.99716093   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  621
[[  0.         -13.99731196 -19.99601681 -21.99554257]
 [-13.99731196 -17.99649104 -19.99604345 -19.99601681]
 [-19.99601681 -19.99604345 -17.99649104 -13.99731196]
 [-21.99554257 -19.99601681 -13.99731196   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  625
[[  0.         -13.99745495 -19.9962287  -21.99577969]
```

```
 [-13.99745495 -17.99667771 -19.99625393 -19.9962287 ]
 [-19.9962287  -19.99625393 -17.99667771 -13.99745495]
 [-21.99577969 -19.9962287  -13.99745495   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  629
[[  0.         -13.99759034 -19.99642932 -21.99600419]
 [-13.99759034 -17.99685444 -19.9964532  -19.99642932]
 [-19.99642932 -19.9964532  -17.99685444 -13.99759034]
 [-21.99600419 -19.99642932 -13.99759034   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  633
[[  0.         -13.99771852 -19.99661926 -21.99621676]
 [-13.99771852 -17.99702177 -19.99664188 -19.99661926]
```

```
 [-19.99661926 -19.99664188 -17.99702177 -13.99771852]
 [-21.99621676 -19.99661926 -13.99771852   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  637
[[  0.         -13.99783989 -19.9967991  -21.99641801]
 [-13.99783989 -17.9971802  -19.99682052 -19.9967991 ]
 [-19.9967991  -19.99682052 -17.9971802  -13.99783989]
 [-21.99641801 -19.9967991  -13.99783989   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  641
[[  0.         -13.9979548  -19.99696938 -21.99660856]
 [-13.9979548  -17.9973302  -19.99698965 -19.99696938]
 [-19.99696938 -19.99698965 -17.9973302  -13.9979548 ]
```

```
 [-21.99660856 -19.99696938 -13.9979548    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  645
[[  0.         -13.9980636  -19.9971306  -21.99678897]
 [-13.9980636  -17.99747223 -19.99714979 -19.9971306 ]
 [-19.9971306  -19.99714979 -17.99747223 -13.9980636 ]
 [-21.99678897 -19.9971306  -13.9980636    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  649
[[  0.         -13.9981666  -19.99728324 -21.99695978]
 [-13.9981666  -17.99760669 -19.99730141 -19.99728324]
 [-19.99728324 -19.99730141 -17.99760669 -13.9981666 ]
 [-21.99695978 -19.99728324 -13.9981666    0.        ]]
```

```
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  653
[[  0.         -13.99826413 -19.99742776 -21.99712151]
 [-13.99826413 -17.99773401 -19.99744497 -19.99742776]
 [-19.99742776 -19.99744497 -17.99773401 -13.99826413]
 [-21.99712151 -19.99742776 -13.99826413   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  657
[[  0.         -13.99835648 -19.99756459 -21.99727463]
 [-13.99835648 -17.99785455 -19.99758088 -19.99756459]
 [-19.99756459 -19.99758088 -17.99785455 -13.99835648]
 [-21.99727463 -19.99756459 -13.99835648   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
```

```
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  661
[[  0.         -13.9984439  -19.99769415 -21.99741961]
 [-13.9984439  -17.99796868 -19.99770957 -19.99769415]
 [-19.99769415 -19.99770957 -17.99796868 -13.9984439 ]
 [-21.99741961 -19.99769415 -13.9984439    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  665
[[  0.         -13.99852668 -19.99781681 -21.99755688]
 [-13.99852668 -17.99807674 -19.99783141 -19.99781681]
 [-19.99781681 -19.99783141 -17.99807674 -13.99852668]
 [-21.99755688 -19.99781681 -13.99852668   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
```

```
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  669
[[  0.         -13.99860506 -19.99793295 -21.99768684]
 [-13.99860506 -17.99817905 -19.99794677 -19.99793295]
 [-19.99793295 -19.99794677 -17.99817905 -13.99860506]
 [-21.99768684 -19.99793295 -13.99860506   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  673
[[  0.         -13.99867926 -19.99804291 -21.9978099 ]
 [-13.99867926 -17.99827592 -19.998056   -19.99804291]
 [-19.99804291 -19.998056   -17.99827592 -13.99867926]
 [-21.9978099  -19.99804291 -13.99867926   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
```

```
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  677
[[  0.         -13.99874952 -19.99814702 -21.9979264 ]
 [-13.99874952 -17.99836763 -19.99815941 -19.99814702]
 [-19.99814702 -19.99815941 -17.99836763 -13.99874952]
 [-21.9979264  -19.99814702 -13.99874952   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  681
[[  0.         -13.99881604 -19.99824559 -21.99803671]
 [-13.99881604 -17.99845447 -19.99825732 -19.99824559]
 [-19.99824559 -19.99825732 -17.99845447 -13.99881604]
 [-21.99803671 -19.99824559 -13.99881604   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
```

```
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  685
[[  0.         -13.99887902 -19.99833891 -21.99814115]
 [-13.99887902 -17.99853668 -19.99835003 -19.99833891]
 [-19.99833891 -19.99835003 -17.99853668 -13.99887902]
 [-21.99814115 -19.99833891 -13.99887902   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  689
[[  0.         -13.99893866 -19.99842728 -21.99824003]
 [-13.99893866 -17.99861452 -19.9984378  -19.99842728]
 [-19.99842728 -19.9984378  -17.99861452 -13.99893866]
 [-21.99824003 -19.99842728 -13.99893866   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
```

```
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  693
[[  0.         -13.99899511 -19.99851094 -21.99833365]
 [-13.99899511 -17.99868823 -19.9985209  -19.99851094]
 [-19.99851094 -19.9985209  -17.99868823 -13.99899511]
 [-21.99833365 -19.99851094 -13.99899511   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  697
[[  0.         -13.99904857 -19.99859015 -21.9984223 ]
 [-13.99904857 -17.99875801 -19.99859958 -19.99859015]
 [-19.99859015 -19.99859958 -17.99875801 -13.99904857]
 [-21.9984223  -19.99859015 -13.99904857   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
```

```
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  701
[[  0.         -13.99909918 -19.99866515 -21.99850623]
 [-13.99909918 -17.99882408 -19.99867408 -19.99866515]
 [-19.99866515 -19.99867408 -17.99882408 -13.99909918]
 [-21.99850623 -19.99866515 -13.99909918   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  705
[[  0.         -13.9991471  -19.99873616 -21.99858569]
 [-13.9991471  -17.99888663 -19.99874461 -19.99873616]
 [-19.99873616 -19.99874461 -17.99888663 -13.9991471 ]
 [-21.99858569 -19.99873616 -13.9991471    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [3]
```

```
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  709
[[  0.         -13.99919247 -19.99880339 -21.99866092]
 [-13.99919247 -17.99894586 -19.9988114  -19.99880339]
 [-19.99880339 -19.9988114  -17.99894586 -13.99919247]
 [-21.99866092 -19.99880339 -13.99919247   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  713
[[  0.         -13.99923543 -19.99886705 -21.99873216]
 [-13.99923543 -17.99900193 -19.99887462 -19.99886705]
 [-19.99886705 -19.99887462 -17.99900193 -13.99923543]
 [-21.99873216 -19.99886705 -13.99923543   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
```

```
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  717
[[  0.         -13.9992761  -19.99892731 -21.9987996 ]
 [-13.9992761  -17.99905503 -19.99893449 -19.99892731]
 [-19.99892731 -19.99893449 -17.99905503 -13.9992761 ]
 [-21.9987996  -19.99892731 -13.9992761    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  721
[[  0.         -13.99931461 -19.99898438 -21.99886346]
 [-13.99931461 -17.9991053  -19.99899117 -19.99898438]
 [-19.99898438 -19.99899117 -17.9991053  -13.99931461]
 [-21.99886346 -19.99898438 -13.99931461   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
```

```
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  725
[[  0.         -13.99935107 -19.9990384  -21.99892392]
 [-13.99935107 -17.99915289 -19.99904484 -19.9990384 ]
 [-19.9990384  -19.99904484 -17.99915289 -13.99935107]
 [-21.99892392 -19.9990384  -13.99935107   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  729
[[  0.         -13.99938559 -19.99908956 -21.99898116]
 [-13.99938559 -17.99919795 -19.99909565 -19.99908956]
 [-19.99908956 -19.99909565 -17.99919795 -13.99938559]
 [-21.99898116 -19.99908956 -13.99938559   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
```

```
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  733
[[  0.         -13.99941828 -19.99913799 -21.99903536]
 [-13.99941828 -17.99924062 -19.99914376 -19.99913799]
 [-19.99913799 -19.99914376 -17.99924062 -13.99941828]
 [-21.99903536 -19.99913799 -13.99941828   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  737
[[  0.         -13.99944922 -19.99918385 -21.99908667]
 [-13.99944922 -17.99928102 -19.9991893  -19.99918385]
 [-19.99918385 -19.9991893  -17.99928102 -13.99944922]
 [-21.99908667 -19.99918385 -13.99944922   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2]
```

```
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  741
[[  0.         -13.99947852 -19.99922726 -21.99913526]
 [-13.99947852 -17.99931926 -19.99923243 -19.99922726]
 [-19.99922726 -19.99923243 -17.99931926 -13.99947852]
 [-21.99913526 -19.99922726 -13.99947852   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  745
[[  0.         -13.99950626 -19.99926837 -21.99918126]
 [-13.99950626 -17.99935548 -19.99927326 -19.99926837]
 [-19.99926837 -19.99927326 -17.99935548 -13.99950626]
 [-21.99918126 -19.99926837 -13.99950626   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
```

```
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  749
[[  0.         -13.99953253 -19.99930729 -21.99922481]
 [-13.99953253 -17.99938976 -19.99931192 -19.99930729]
 [-19.99930729 -19.99931192 -17.99938976 -13.99953253]
 [-21.99922481 -19.99930729 -13.99953253   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  753
[[  0.         -13.99955739 -19.99934414 -21.99926605]
 [-13.99955739 -17.99942222 -19.99934852 -19.99934414]
 [-19.99934414 -19.99934852 -17.99942222 -13.99955739]
 [-21.99926605 -19.99934414 -13.99955739   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
```

```
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  757
[[  0.         -13.99958094 -19.99937903 -21.99930509]
 [-13.99958094 -17.99945296 -19.99938318 -19.99937903]
 [-19.99937903 -19.99938318 -17.99945296 -13.99958094]
 [-21.99930509 -19.99937903 -13.99958094   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  761
[[  0.         -13.99960323 -19.99941206 -21.99934206]
 [-13.99960323 -17.99948206 -19.99941599 -19.99941206]
 [-19.99941206 -19.99941599 -17.99948206 -13.99960323]
 [-21.99934206 -19.99941206 -13.99960323   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
```

```
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  765
[[  0.         -13.99962434 -19.99944334 -21.99937706]
 [-13.99962434 -17.99950961 -19.99944706 -19.99944334]
 [-19.99944334 -19.99944706 -17.99950961 -13.99962434]
 [-21.99937706 -19.99944334 -13.99962434   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  769
[[  0.         -13.99964432 -19.99947295 -21.9994102 ]
 [-13.99964432 -17.9995357  -19.99947647 -19.99947295]
 [-19.99947295 -19.99947647 -17.9995357  -13.99964432]
 [-21.9994102  -19.99947295 -13.99964432   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
```

```
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  773
[[  0.         -13.99966324 -19.99950099 -21.99944157]
 [-13.99966324 -17.9995604  -19.99950432 -19.99950099]
 [-19.99950099 -19.99950432 -17.9995604  -13.99966324]
 [-21.99944157 -19.99950099 -13.99966324   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  777
[[  0.         -13.99968116 -19.99952753 -21.99947128]
 [-13.99968116 -17.99958378 -19.99953069 -19.99952753]
 [-19.99952753 -19.99953069 -17.99958378 -13.99968116]
 [-21.99947128 -19.99952753 -13.99968116   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
```

```
Iteration -  781
[[  0.         -13.99969812 -19.99955266 -21.99949941]
 [-13.99969812 -17.99960592 -19.99955566 -19.99955266]
 [-19.99955266 -19.99955566 -17.99960592 -13.99969812]
 [-21.99949941 -19.99955266 -13.99969812   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  785
[[  0.         -13.99971418 -19.99957646 -21.99952604]
 [-13.99971418 -17.99962689 -19.99957929 -19.99957646]
 [-19.99957646 -19.99957929 -17.99962689 -13.99971418]
 [-21.99952604 -19.99957646 -13.99971418   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  789
```

```
[[  0.         -13.99972938 -19.99959899 -21.99955125]
 [-13.99972938 -17.99964674 -19.99960167 -19.99959899]
 [-19.99959899 -19.99960167 -17.99964674 -13.99972938]
 [-21.99955125 -19.99959899 -13.99972938   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  793
[[  0.         -13.99974378 -19.99962032 -21.99957512]
 [-13.99974378 -17.99966553 -19.99962286 -19.99962032]
 [-19.99962032 -19.99962286 -17.99966553 -13.99974378]
 [-21.99957512 -19.99962032 -13.99974378   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  797
[[  0.         -13.99975741 -19.99964052 -21.99959772]
```

```
 [-13.99975741 -17.99968332 -19.99964293 -19.99964052]
 [-19.99964052 -19.99964293 -17.99968332 -13.99975741]
 [-21.99959772 -19.99964052 -13.99975741   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  801
[[  0.         -13.99977031 -19.99965964 -21.99961912]
 [-13.99977031 -17.99970017 -19.99966192 -19.99965964]
 [-19.99965964 -19.99966192 -17.99970017 -13.99977031]
 [-21.99961912 -19.99965964 -13.99977031   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  805
[[  0.         -13.99978253 -19.99967775 -21.99963938]
 [-13.99978253 -17.99971612 -19.99967991 -19.99967775]
```

```
 [-19.99967775 -19.99967991 -17.99971612 -13.99978253]
 [-21.99963938 -19.99967775 -13.99978253   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  809
[[  0.         -13.9997941  -19.99969489 -21.99965857]
 [-13.9997941  -17.99973122 -19.99969693 -19.99969489]
 [-19.99969489 -19.99969693 -17.99973122 -13.9997941 ]
 [-21.99965857 -19.99969489 -13.9997941    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  813
[[  0.         -13.99980505 -19.99971112 -21.99967673]
 [-13.99980505 -17.99974552 -19.99971306 -19.99971112]
 [-19.99971112 -19.99971306 -17.99974552 -13.99980505]
```

```
 [-21.99967673 -19.99971112 -13.99980505   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  817
[[  0.         -13.99981542 -19.99972649 -21.99969393]
 [-13.99981542 -17.99975905 -19.99972832 -19.99972649]
 [-19.99972649 -19.99972832 -17.99975905 -13.99981542]
 [-21.99969393 -19.99972649 -13.99981542   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  821
[[  0.         -13.99982524 -19.99974104 -21.99971021]
 [-13.99982524 -17.99977187 -19.99974277 -19.99974104]
 [-19.99974104 -19.99974277 -17.99977187 -13.99982524]
 [-21.99971021 -19.99974104 -13.99982524   0.        ]]
```

```
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  825
[[  0.        -13.99983454 -19.99975482 -21.99972562]
 [-13.99983454 -17.99978401 -19.99975646 -19.99975482]
 [-19.99975482 -19.99975646 -17.99978401 -13.99983454]
 [-21.99972562 -19.99975482 -13.99983454   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  829
[[  0.        -13.99984334 -19.99976786 -21.99974022]
 [-13.99984334 -17.9997955  -19.99976941 -19.99976786]
 [-19.99976786 -19.99976941 -17.9997955  -13.99984334]
 [-21.99974022 -19.99976786 -13.99984334   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
```

```
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  833
[[  0.         -13.99985167 -19.99978021 -21.99975404]
 [-13.99985167 -17.99980638 -19.99978168 -19.99978021]
 [-19.99978021 -19.99978168 -17.99980638 -13.99985167]
 [-21.99975404 -19.99978021 -13.99985167   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  837
[[  0.         -13.99985956 -19.9997919  -21.99976712]
 [-13.99985956 -17.99981668 -19.99979329 -19.9997919 ]
 [-19.9997919  -19.99979329 -17.99981668 -13.99985956]
 [-21.99976712 -19.9997919  -13.99985956   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
```

```
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  841
[[  0.         -13.99986703 -19.99980297 -21.99977951]
 [-13.99986703 -17.99982643 -19.99980429 -19.99980297]
 [-19.99980297 -19.99980429 -17.99982643 -13.99986703]
 [-21.99977951 -19.99980297 -13.99986703   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  845
[[  0.         -13.99987411 -19.99981345 -21.99979124]
 [-13.99987411 -17.99983566 -19.9998147  -19.99981345]
 [-19.99981345 -19.9998147  -17.99983566 -13.99987411]
 [-21.99979124 -19.99981345 -13.99987411   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
```

```
Optimal Policy for State: (0, 3) --- [0]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0, 1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  849
[[  0.         -13.9998808  -19.99982337 -21.99980235]
 [-13.9998808  -17.9998444  -19.99982456 -19.99982337]
 [-19.99982337 -19.99982456 -17.9998444  -13.9998808 ]
 [-21.99980235 -19.99982337 -13.9998808    0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  853
[[  0.         -13.99988715 -19.99983277 -21.99981286]
 [-13.99988715 -17.99985268 -19.99983389 -19.99983277]
 [-19.99983277 -19.99983389 -17.99985268 -13.99988715]
 [-21.99981286 -19.99983277 -13.99988715   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
```

```
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  857
[[  0.         -13.99989315 -19.99984167 -21.99982282]
 [-13.99989315 -17.99986052 -19.99984273 -19.99984167]
 [-19.99984167 -19.99984273 -17.99986052 -13.99989315]
 [-21.99982282 -19.99984167 -13.99989315   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [1]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration -  861
[[  0.         -13.99989883 -19.99985009 -21.99983224]
 [-13.99989883 -17.99986794 -19.99985109 -19.99985009]
 [-19.99985009 -19.99985109 -17.99986794 -13.99989883]
 [-21.99983224 -19.99985009 -13.99989883   0.        ]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
```

Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [0, 3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [1, 2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]
Iteration - 865
[[0. -13.99990421 -19.99985806 -21.99984116]
[-13.99990421 -17.99987496 -19.99985901 -19.99985806]
[-19.99985806 -19.99985901 -17.99987496 -13.99990421]
[-21.99984116 -19.99985806 -13.99990421 0.]]
Optimal Policy for State: (0, 0) --- [0, 1]
Optimal Policy for State: (0, 1) --- [0]
Optimal Policy for State: (0, 2) --- [0]
Optimal Policy for State: (0, 3) --- [0, 3]
Optimal Policy for State: (1, 0) --- [1]
Optimal Policy for State: (1, 1) --- [0]
Optimal Policy for State: (1, 2) --- [3]
Optimal Policy for State: (1, 3) --- [3]
Optimal Policy for State: (2, 0) --- [1]
Optimal Policy for State: (2, 1) --- [2]
Optimal Policy for State: (2, 2) --- [2, 3]
Optimal Policy for State: (2, 3) --- [3]
Optimal Policy for State: (3, 0) --- [1, 2]
Optimal Policy for State: (3, 1) --- [2]
Optimal Policy for State: (3, 2) --- [2]
Optimal Policy for State: (3, 3) --- [2, 3]

Policy Improvement

New State Values
[[0. -1. -1. -1.]
 [-1. -1. -1. -1.]
 [-1. -1. -1. -1.]
 [-1. -1. -1. 0.]]
Updated Policy with policy improvement
[[0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.   0.   0.   0.]
 [0.   0.   0.   0.]]
New State Values
[[0.   -1.75 -2.   -2.]
 [-1.75 -2.   -2.   -2.]
 [-2.   -2.   -2.   -1.75]
 [-2.   -2.   -1.75 0.]]
Updated Policy with policy improvement
[[1.   0.   0.   0.]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]

```
 [0.   1.   0.   0.  ]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.   0.   0.   1.  ]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.   0.   1.   0.  ]
 [0.   0.   0.   0.  ]
 [0.   0.   0.   0.  ]]
New State Values
[[ 0.   -1.   -2.75 -3.  ]
 [-2.75 -2.75 -3.   -3.  ]
 [-3.   -3.   -3.   -3.  ]
 [-3.   -3.   -3.    0.  ]]
Updated Policy with policy improvement
[[1.   0.   0.   0.  ]
 [1.   0.   0.   0.  ]
 [0.25 0.25 0.25 0.25]
 [0.   1.   0.   0.  ]
 [0.   1.   0.   0.  ]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.   0.   0.   1.  ]
 [0.25 0.25 0.25 0.25]
 [0.25 0.25 0.25 0.25]
 [0.   0.   1.   0.  ]
 [0.   0.   0.   0.  ]
 [0.   0.   0.   0.  ]]
New State Values
[[ 0.   -1.   -2.   -3.75]
 [-3.75 -3.75 -3.75 -4.  ]
 [-4.   -4.   -4.   -4.  ]
 [-4.   -4.   -4.    0.  ]]
Updated Policy with policy improvement
[[1.   0.   0.   0.  ]
 [1.   0.   0.   0.  ]
```

[1. 0. 0. 0.]
[0. 1. 0. 0.]
[0. 1. 0. 0.]
[0. 1. 0. 0.]
[0.25 0.25 0.25 0.25]
[0.25 0.25 0.25 0.25]
[0.25 0.25 0.25 0.25]
[0.25 0.25 0.25 0.25]
[0. 0. 0. 1.]
[0.25 0.25 0.25 0.25]
[0.25 0.25 0.25 0.25]
[0. 0. 1. 0.]
[0. 0. 0. 0.]
[0. 0. 0. 0.]]

Answer 7