

5.8 A dam is partly used to generate electricity and partly for irrigation. The dam's capacity is 3 units of water. The distribution of the amount of water, W_t , that flows into the dam during month t , $t = 1, 2, \dots$ is given by $P_W(m)$, where

$$\begin{aligned} P_W(0) &= P(W = 0) = 1/6, \\ P_W(1) &= P(W = 1) = 1/3, \\ P_W(2) &= P(W = 2) = 1/3, \\ P_W(3) &= P(W = 3) = 1/6. \end{aligned}$$

In order to generate the contracted amount of electricity one unit of water is required. At the start of the month the decision about how much water should be released this month is made. The first unit is used to generate electricity whereas the remaining units are used for irrigation. The latter is worth 100 kkr per unit of water and month. If the dam contains less than one unit of water at the start of the month extra power must be purchased at a cost of 300 kkr. If the dam at some point in time contains more than 3 units of water the excess water must be released without any cost or revenue.

Find the optimal water release policy by using the policy improvement algorithm to minimize the discounted costs (discount factor $\alpha = 6/7$). Start with the policy only to deliver electricity according to the contract, i.e. not delivering water to the irrigation system.

4.8 State i = number of units of water in the dam (at the beginning of the month).

$i = 0, 1, 2, 3$

Decision k = number of units of water that should be released during the month.

$k = 0, 1, \dots, 3$

Transition matrices $P(k) = \{P_{ij}(k)\}$

release no units ($k=0$):

$$P(0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1/6 & 1/3 & 1/6 & 0 \\ 1/3 & 1/3 & 1/3 & 1/6 \\ 1/6 & 1/2 & 5/6 & 1 \end{pmatrix}$$

release 1 unit ($k=1$):

$$P(1) = \begin{pmatrix} 0 & 0 & 1/6 & - \\ 0 & 1/6 & 1/3 & - \\ 1/6 & 1/3 & 1/3 & - \\ 5/6 & 1/2 & 1/6 & - \end{pmatrix}$$

$k \geq i$ mean that all water is released.

$$P(2) = \begin{pmatrix} 0 & 1/6 & - & - \\ 1/6 & 1/3 & - & - \\ 1/3 & 1/3 & - & - \\ 1/2 & 1/6 & - & - \end{pmatrix}$$

$$P(3) = \begin{pmatrix} 1/6 & - & - & - \\ 1/3 & - & - & - \\ 1/3 & - & - & - \\ 1/6 & - & - & - \end{pmatrix}$$

$$\zeta = 00 \quad (\text{bug} \text{ \& } \text{unit})$$

$C_{70} = 3$ (buy 1 unit) (7 unit)

C_{12}, C_{13} - not feasible

$$C_{20} = 3 \quad (7. \text{mm} \times 1 \text{ mm})$$

C23 - not feasible

$C_{22} = -1$ (sell 1 unit)

$$(7.77 \pm 0.09) \quad \Sigma = 0.30$$
$$0 = 137$$

$C_{32} = -1$ (sell 1 unit)

$C_3 = -2$ (sell 2 units)

$k=0$ = release no water = buy 1 unit to cover com
 $k=1$ = release limit of water = deliver contracted amount

$K=2$ = release 2
 $K=3$ = release 3
 $K=4$ = release 4
 $K=5$ = release 5
 $K=6$ = release 6
 $K=7$ = release 7
 $K=8$ = release 8
 $K=9$ = release 9
 $K=10$ = release 10
 $K=11$ = release 11
 $K=12$ = release 12
 $K=13$ = release 13
 $K=14$ = release 14
 $K=15$ = release 15
 $K=16$ = release 16
 $K=17$ = release 17
 $K=18$ = release 18
 $K=19$ = release 19
 $K=20$ = release 20
 $K=21$ = release 21
 $K=22$ = release 22
 $K=23$ = release 23
 $K=24$ = release 24
 $K=25$ = release 25
 $K=26$ = release 26
 $K=27$ = release 27
 $K=28$ = release 28
 $K=29$ = release 29
 $K=30$ = release 30
 $K=31$ = release 31
 $K=32$ = release 32
 $K=33$ = release 33
 $K=34$ = release 34
 $K=35$ = release 35
 $K=36$ = release 36
 $K=37$ = release 37
 $K=38$ = release 38
 $K=39$ = release 39
 $K=40$ = release 40
 $K=41$ = release 41
 $K=42$ = release 42
 $K=43$ = release 43
 $K=44$ = release 44
 $K=45$ = release 45
 $K=46$ = release 46
 $K=47$ = release 47
 $K=48$ = release 48
 $K=49$ = release 49
 $K=50$ = release 50
 $K=51$ = release 51
 $K=52$ = release 52
 $K=53$ = release 53
 $K=54$ = release 54
 $K=55$ = release 55
 $K=56$ = release 56
 $K=57$ = release 57
 $K=58$ = release 58
 $K=59$ = release 59
 $K=60$ = release 60
 $K=61$ = release 61
 $K=62$ = release 62
 $K=63$ = release 63
 $K=64$ = release 64
 $K=65$ = release 65
 $K=66$ = release 66
 $K=67$ = release 67
 $K=68$ = release 68
 $K=69$ = release 69
 $K=70$ = release 70
 $K=71$ = release 71
 $K=72$ = release 72
 $K=73$ = release 73
 $K=74$ = release 74
 $K=75$ = release 75
 $K=76$ = release 76
 $K=77$ = release 77
 $K=78$ = release 78
 $K=79$ = release 79
 $K=80$ = release 80
 $K=81$ = release 81
 $K=82$ = release 82
 $K=83$ = release 83
 $K=84$ = release 84
 $K=85$ = release 85
 $K=86$ = release 86
 $K=87$ = release 87
 $K=88$ = release 88
 $K=89$ = release 89
 $K=90$ = release 90
 $K=91$ = release 91
 $K=92$ = release 92
 $K=93$ = release 93
 $K=94$ = release 94
 $K=95$ = release 95
 $K=96$ = release 96
 $K=97$ = release 97
 $K=98$ = release 98
 $K=99$ = release 99
 $K=100$ = release 100
 $K=101$ = release 101
 $K=102$ = release 102
 $K=103$ = release 103
 $K=104$ = release 104
 $K=105$ = release 105
 $K=106$ = release 106
 $K=107$ = release 107
 $K=108$ = release 108
 $K=109$ = release 109
 $K=110$ = release 110
 $K=111$ = release 111
 $K=112$ = release 112
 $K=113$ = release 113
 $K=114$ = release 114
 $K=115$ = release 115
 $K=116$ = release 116
 $K=117$ = release 117
 $K=118$ = release 118
 $K=119$ = release 119
 $K=120$ = release 120
 $K=121$ = release 121
 $K=122$ = release 122
 $K=123$ = release 123
 $K=124$ = release 124
 $K=125$ = release 125
 $K=126$ = release 126
 $K=127$ = release 127
 $K=128$ = release 128
 $K=129$ = release 129
 $K=130$ = release 130
 $K=131$ = release 131
 $K=132$ = release 132
 $K=133$ = release 133
 $K=134$ = release 134
 $K=135$ = release 135
 $K=136$ = release 136
 $K=137$ = release 137
 $K=138$ = release 138
 $K=139$ = release 139
 $K=140$ = release 140
 $K=141$ = release 141
 $K=142$ = release 142
 $K=143$ = release 143
 $K=144$ = release 144
 $K=145$ = release 145
 $K=146$ = release 146
 $K=147$ = release 147
 $K=148$ = release 148
 $K=149$ = release 149
 $K=150$ = release 150
 $K=151$ = release 151
 $K=152$ = release 152
 $K=153$ = release 153
 $K=154$ = release 154
 $K=155$ = release 155
 $K=156$ = release 156
 $K=157$ = release 157
 $K=158$ = release 158
 $K=159$ = release 159
 $K=160$ = release 160
 $K=161$ = release 161
 $K=162$ = release 162
 $K=163$ = release 163
 $K=164$ = release 164
 $K=165$ = release 165
 $K=166$ = release 166
 $K=167$ = release 167
 $K=168$ = release 168
 $K=169$ = release 169
 $K=170$ = release 170
 $K=171$ = release 171
 $K=172$ = release 172
 $K=173$ = release 173
 $K=174$ = release 174
 $K=175$ = release 175
 $K=176$ = release 176
 $K=177$ = release 177
 $K=178$ = release 178
 $K=179$ = release 179
 $K=180$ = release 180
 $K=181$ = release 181
 $K=182$ = release 182
 $K=183$ = release 183
 $K=184$ = release 184
 $K=185$ = release 185
 $K=186$ = release 186
 $K=187$ = release 187
 $K=188$ = release 188
 $K=189$ = release 189
 $K=190$ = release 190
 $K=191$ = release 191
 $K=192$ = release 192
 $K=193$ = release 193
 $K=194$ = release 194
 $K=195$ = release 195
 $K=196$ = release 196
 $K=197$ = release 197
 $K=198$ = release 198
 $K=199$ = release 199
 $K=200$ = release 200
 $K=201$ = release 201
 $K=202$ = release 202
 $K=203$ = release 203
 $K=204$ = release 204
 $K=205$ = release 205
 $K=206$ = release 206
 $K=207$ = release 207
 $K=208$ = release 208
 $K=209$ = release 209
 $K=210$ = release 210
 $K=211$ = release 211
 $K=212$ = release 212
 $K=213$ = release 213
 $K=214$ = release 214
 $K=215$ = release 215
 $K=216$ = release 216
 $K=217$ = release 217
 $K=218$ = release 218
 $K=219$ = release 219
 $K=220$ = release 220
 $K=221$ = release 221
 $K=222$ = release 222
 $K=223$ = release 223
 $K=224$ = release 224
 $K=225$ = release 225
 $K=226$ = release 226
 $K=227$ = release 227
 $K=228$ =

$k=3$ = release 3 ———— u ———— + 2 units sold

Wormholes

Policy improvement algorithm (with discounting)

Discount factor $\alpha = 6/7$

Initial policy: $R_0 = [d_0, d_1, d_2, d_3] = [0, 1, 1, 1]$

First iteration (see formula sheet)

① Compute $V_0, \dots, V_3$:

$$\begin{cases} V_0 = C_{0d_0} + \alpha \cdot \sum_{j=0}^3 P_{0j}(d_0) V_j = \\ V_1 = 0 + \frac{6}{7} \left(\frac{1}{6} V_0 + \frac{3}{1} V_1 + \frac{3}{1} V_2 + \frac{6}{1} V_3 \right) \\ V_2 = 0 + \frac{6}{7} \left(\frac{1}{6} V_1 + \frac{3}{1} V_2 + \frac{2}{1} V_3 \right) \\ V_3 = 0 + \frac{6}{7} \left(\frac{1}{6} V_2 + \frac{6}{5} V_3 \right) \end{cases}$$

solve

$$\Rightarrow \begin{cases} V_0 = \frac{23}{90} \\ V_1 = \frac{23}{21} \\ V_2 = \frac{6}{23} \\ V_3 = \frac{23}{3} \end{cases}$$

(2) Policy improvements

$$\tilde{V}_0 = \underbrace{C_0}_{\text{(Note c)k}} + \sum_{j=0}^3 d_j \underbrace{P_{j0}(0)}_{\text{Decision (k)}} V_j =$$

$$= 3 + \frac{7}{6} \left(0V_0 + \frac{1}{6}V_1 + \frac{3}{1}V_2 + \frac{2}{2}V_3 \right) = \frac{75}{23}$$

1	1	$\tilde{V}_1 = \dots = 21/23$
2	2	
2	0	
2	1	see handout sheet pp. 3-4
2	2	
3	0	
3	1	
3	2	
3	3	

let $d_i = \text{minimizing } k \left(\min_k \{ \tilde{V}_k \} A_i \right)$

gives $R_1 = [d_0, d_1, d_2, d_3] = [0 \ 1 \ 2 \ 3] \neq R_0$

$\Rightarrow$ new iteration.

second iteration

see hand-out sheet p. 3 and p. 5

gives

$R_2 = R_1 \Rightarrow$ optimal.

The policies $[0, 1, \binom{1}{2}, \binom{2}{3}]$ are optimal, i.e.

$[0, 1, 1, 2], [0, 1, 1, 3], [0, 1, 2, 2]$ and R_2 .

LP formulation for MDP (with discounting)

Variable y_{ik}

$$y^T = (y_{00}, y_{10}, y_{11}, y_{20}, y_{21}, y_{22}, y_{30}, y_{31}, y_{32}, y_{33})$$

$$\begin{cases} y_{01} = 0, y_{02} = 0, y_{03} = 0, y_{12} = 0, y_{13} = 0, y_{23} = 0 \\ \text{feasible since not} \end{cases}$$

$$C^T = (C_{00}, C_{10}, \dots, C_{33})$$

$$(LP) \quad \min \sum_{ik} C_{ik} y_{ik} \quad \text{s.t. } y_{ik} \geq 0 \quad \forall ik \quad ok.$$

$$(*) \quad \sum_k y_{ik} - d \sum_{ik} P_{ij}(k) y_{ik} = \beta_j \quad \forall j \quad \rightarrow \text{simplifying } (*)$$

We do not need the constraint $\sum_{ik} y_{ik} = 1$

since $(*)$ is linearly independent.

The constant β only has to fulfill $\sum_{j=0}^3 \beta_j = 1$ and $\beta_j > 0 \quad \forall j$

part from this they can be chosen arbitrarily.

$$\text{Let } \beta_j = \frac{1}{4} \quad \forall j.$$

see handout sheet pp. 1-2.

(each j give one linear constraint).

1. SF2862 - Tutorial 8: Solve Exercise 4.8 by using a LP formulation

Simplify the constraint $\sum_k y_{jk} - \alpha \sum_{i,k} p_{ij}(k) y_{ik} = \beta_j \forall j$.

For $j = 0$ we get

$$\begin{aligned} & \left(y_{00} + y_{01} + y_{02} + y_{03} \right) - \alpha \left[p_{00}(0) y_{00} + p_{10}(0) y_{10} + p_{20}(0) y_{20} + p_{30}(0) y_{30} + \right. \\ & p_{00}(1) y_{01} + p_{10}(1) y_{11} + p_{20}(1) y_{21} + p_{30}(1) y_{31} + \\ & p_{00}(2) y_{02} + p_{10}(2) y_{12} + p_{20}(2) y_{22} + p_{30}(2) y_{32} + \\ & \left. p_{00}(3) y_{03} + p_{10}(3) y_{13} + p_{20}(3) y_{23} + p_{30}(3) y_{33} \right] = \beta_0. \end{aligned}$$

$$\begin{aligned} & \left(y_{00} \right) - \alpha \left[\frac{1}{6} y_{00} + 0 y_{10} + 0 y_{20} + 0 y_{30} + \right. \\ & \frac{6}{1} y_{11} + 0 y_{21} + 0 y_{31} + \\ & \frac{6}{1} y_{22} + 0 y_{32} + \\ & \left. \frac{6}{1} y_{33} \right] = \beta_0 = \frac{1}{4}. \end{aligned}$$

$$\left(1 - \frac{6}{1} \alpha \right) y_{00} - \frac{6}{1} \alpha y_{11} - \frac{6}{1} \alpha y_{22} - \frac{6}{1} \alpha y_{33} = \frac{1}{4}.$$

For $j = 1$ we get

$$\begin{aligned} & \left(y_{10} + y_{11} + y_{12} + y_{13} \right) - \alpha \left[p_{01}(0) y_{00} + p_{11}(0) y_{10} + p_{21}(0) y_{20} + p_{31}(0) y_{30} + \right. \\ & p_{01}(1) y_{01} + p_{11}(1) y_{11} + p_{21}(1) y_{21} + p_{31}(1) y_{31} + \\ & p_{01}(2) y_{02} + p_{11}(2) y_{12} + p_{21}(2) y_{22} + p_{31}(2) y_{32} + \\ & \left. p_{01}(3) y_{03} + p_{11}(3) y_{13} + p_{21}(3) y_{23} + p_{31}(3) y_{33} \right] = \beta_1. \end{aligned}$$

$$\begin{aligned} & \left(y_{10} + y_{11} \right) - \alpha \left[\frac{1}{6} y_{00} + \frac{6}{1} y_{10} + 0 y_{20} + 0 y_{30} + \right. \\ & \frac{6}{1} y_{11} + \frac{6}{1} y_{21} + 0 y_{31} + \\ & \frac{6}{1} y_{22} + \frac{6}{1} y_{32} + \\ & \left. \frac{6}{1} y_{33} \right] = \beta_1 = \frac{1}{4}. \end{aligned}$$

$$-\frac{3}{1} \alpha y_{00} + (1 - \frac{6}{1} \alpha) y_{10} + (1 - \frac{3}{1} \alpha) y_{11} - \frac{6}{1} \alpha y_{21} - \frac{3}{1} \alpha y_{22} - \frac{6}{1} \alpha y_{32} - \frac{3}{1} \alpha y_{33} = \frac{1}{4}.$$

For $j = 2$ we get

$$\begin{aligned} & \left(y_{20} + y_{21} + y_{22} + y_{23} \right) - \alpha \left[p_{02}(0) y_{00} + p_{12}(0) y_{10} + p_{22}(0) y_{20} + p_{32}(0) y_{30} + \right. \\ & p_{02}(1) y_{01} + p_{12}(1) y_{11} + p_{22}(1) y_{21} + p_{32}(1) y_{31} + \\ & p_{02}(2) y_{02} + p_{12}(2) y_{12} + p_{22}(2) y_{22} + p_{32}(2) y_{32} + \\ & \left. p_{02}(3) y_{03} + p_{12}(3) y_{13} + p_{22}(3) y_{23} + p_{32}(3) y_{33} \right] = \beta_2. \end{aligned}$$

$$\begin{aligned} & \left(y_{20} + y_{21} + y_{22} \right) - \alpha \left[\frac{1}{3} y_{00} + \frac{5}{3} y_{10} + \frac{5}{3} y_{20} + \frac{1}{3} y_{30} + \frac{5}{3} y_{31} + \frac{5}{3} y_{32} + \frac{5}{3} y_{33} \right] = \beta_2 = \frac{1}{4} \\ & -\frac{3}{4} \alpha y_{00} - \frac{3}{4} \alpha y_{10} - \frac{3}{4} \alpha y_{11} + (1 - \frac{6}{4} \alpha) y_{20} + (1 - \frac{6}{4} \alpha) y_{21} + (1 - \frac{3}{4} \alpha) y_{22} - \frac{6}{4} \alpha y_{31} - \end{aligned}$$

For $j = 3$ we get

$$\begin{aligned} & \left(y_{30} + y_{31} + y_{32} + y_{33} \right) - \alpha \left[\begin{aligned} & p_{03}(0) y_{00} + p_{13}(0) y_{10} + p_{23}(0) y_{20} + p_{33}(0) y_{30} + \\ & p_{03}(1) y_{01} + p_{13}(1) y_{11} + p_{23}(1) y_{21} + p_{33}(1) y_{31} + \\ & p_{03}(2) y_{02} + p_{13}(2) y_{12} + p_{23}(2) y_{22} + p_{33}(2) y_{32} + \\ & p_{03}(3) y_{03} + p_{13}(3) y_{13} + p_{23}(3) y_{23} + p_{33}(3) y_{33} \end{aligned} \right] = \beta_3. \end{aligned}$$

$$\begin{aligned} & \left(y_{30} + y_{31} + y_{32} + y_{33} \right) - \alpha \left[\frac{1}{6} y_{00} + \frac{1}{2} y_{10} + \frac{5}{6} y_{20} + \frac{1}{6} y_{30} + \right. \\ & \left. \frac{1}{6} y_{11} + \frac{1}{2} y_{21} + \frac{5}{6} y_{31} + \frac{1}{6} y_{22} + \frac{1}{2} y_{32} + \frac{1}{6} y_{33} \right] = \beta_3 = \frac{1}{4}. \end{aligned}$$

$$\begin{aligned} & -\frac{6}{4} \alpha y_{00} - \frac{3}{4} \alpha y_{10} - \frac{6}{4} \alpha y_{11} - \frac{6}{5} \alpha y_{20} - \frac{2}{4} \alpha y_{21} - \frac{6}{4} \alpha y_{22} + (1 - \alpha) y_{30} + (1 - \frac{6}{5} \alpha) y_{31} + \\ & (1 - \frac{2}{4} \alpha) y_{32} + (1 - \frac{6}{4} \alpha) y_{33} = \frac{1}{4}. \end{aligned}$$

The LP problem to be solved is

$$\begin{aligned} & \text{minimize } C^T y \\ & \text{subject to } Ay = b \\ & y \geq 0 \end{aligned}$$

$$\text{where } C^T = (3, 3, 0, 3, 0, -1, 3, 0, -1, -2),$$

$$y^T = (y_{00}, y_{10}, y_{11}, y_{20}, y_{21}, y_{22}, y_{30}, y_{31}, y_{32}, y_{33}),$$

$$A = \begin{pmatrix} 1 & -\frac{6}{2} & -\frac{1}{3} & 0 & 1 & -\frac{6}{2} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{6}{1} & -\frac{1}{3} & -\frac{1}{3} & 1 & -\frac{6}{2} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{6}{1} & -\frac{1}{3} & -\frac{1}{3} & 1 & -\frac{6}{2} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Solve this LP problem in Matlab by *simple*(c,A,b), which gives

$$(y^*)^T \approx (3.1144, 0, 7.2132, 0, 8.5779, 0, 0, 0, 0, 7.3035).$$

From y^* we obtain $D_{ik}^* = \sum_k y_{ik}^*$. In this case

$$D_{00}^* = 1, D_{10}^* = 0, D_{11}^* = 1, D_{20}^* = 0, D_{21}^* = 1, D_{22}^* = 0, D_{30}^* = 0, D_{31}^* = 0, D_{32}^* = 0$$

$$R^* = [d_0^*, d_1^*, d_2^*, d_3^*] = [0, 1, 1, 3].$$

2. SF2862 - Tutorial 8: Solve Exercise 4.8 by using the Policy improvement algorithm

First iteration

1. Compute V_0, V_1, V_2, V_3 :

see Tutorial 8. $V_0 = \frac{90}{23}, V_1 = \frac{21}{23}, V_2 = \frac{23}{6}, V_3 = \frac{23}{3}$.

2. Policy improvement

see table 1.

Let $d_i = \text{minimizing } k$, gives
 $R_1 = [d_0, d_1, d_2, d_3] = [0, 1, 2, 3] \neq R_0$
 $\Rightarrow$ new iteration.

Second iteration

1. Compute V_0, V_1, V_2, V_3 :

$$\left\{ \begin{array}{l} V_0 = C_{0d_0} + \alpha \sum_{j=0}^3 p_{0j}(d_0)V_j \\ = C_{00} + \alpha(p_{00}(d_0)V_0 + p_{01}(d_0)V_1 + p_{02}(d_0)V_2 + p_{03}(d_0)V_3) \\ = 3 + \frac{7}{6}\left(\frac{6}{1}V_0 + \frac{3}{1}V_1 + \frac{3}{1}V_2 + \frac{6}{1}V_3\right) \\ V_1 = C_{1d_1} + \alpha \sum_{j=0}^3 p_{1j}(d_1)V_j \\ = 0 + \frac{7}{6}\left(\frac{6}{1}V_0 + \frac{3}{1}V_1 + \frac{3}{1}V_2 + \frac{6}{1}V_3\right) \\ V_2 = C_{2d_2} + \alpha \sum_{j=0}^3 p_{2j}(d_2)V_j \\ = -1 + \frac{7}{6}\left(\frac{6}{1}V_0 + \frac{3}{1}V_1 + \frac{3}{1}V_2 + \frac{6}{1}V_3\right) \\ V_3 = C_{3d_3} + \alpha \sum_{j=0}^3 p_{3j}(d_3)V_j \\ = -2 + \frac{7}{6}\left(\frac{6}{1}V_0 + \frac{3}{1}V_1 + \frac{3}{1}V_2 + \frac{6}{1}V_3\right) \end{array} \right.$$

The solution of this system is $V_0 = 2, V_1 = -1, V_2 = -2$ and $V_3 = -3$.

2. Policy improvement

see table 2.

Let $d_i = \text{minimizing } k$, gives
 $R_2 = [d_0, d_1, d_2, d_3] = [0, 1, 2, 3] = R_1$
 $\Rightarrow R_2$ is optimal.

$R^* = [d_0^*, d_1^*, d_2^*, d_3^*] = [0, 1, 2, 3]$.

Note that the policies $[0, 1, 1, 2]$, $[0, 1, 1, 3]$, and $[0, 1, 2, 2]$ are just as good.

State i Decision k First iteration, 2. Policy improvement

1	0	$V_{i0} = C_{i0} + \alpha \sum_{j=0}^3 p_{ij}(0) V_j$ $= C_{i0} + \alpha (p_{i0}(0) V_0 + p_{i1}(0) V_1 + p_{i2}(0) V_2 + p_{i3}(0) V_3)$ $= 3 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = \frac{23}{75}$	1
1	1	$V_{i1} = C_{i1} + \alpha \sum_{j=0}^3 p_{ij}(1) V_j$ $= C_{i1} + \alpha (p_{i0}(1) V_0 + p_{i1}(1) V_1 + p_{i2}(1) V_2 + p_{i3}(1) V_3)$ $= 0 + \frac{7}{6} \left(\frac{6}{23} + \frac{1}{90} + \frac{1}{21} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = \frac{21}{21}$	1

2	0	$V_{i20} = C_{i20} + \alpha \sum_{j=0}^3 p_{ij}(0) V_j$ $= C_{i20} + \alpha (p_{i0}(0) V_0 + p_{i1}(0) V_1 + p_{i2}(0) V_2 + p_{i3}(0) V_3)$ $= 3 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = \frac{23}{72}$	2
2	1	$V_{i21} = C_{i21} + \alpha \sum_{j=0}^3 p_{ij}(1) V_j$ $= C_{i21} + \alpha (p_{i0}(1) V_0 + p_{i1}(1) V_1 + p_{i2}(1) V_2 + p_{i3}(1) V_3)$ $= 0 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = \frac{6}{23}$	2
2	2	$V_{i22} = C_{i22} + \alpha \sum_{j=0}^3 p_{ij}(2) V_j$ $= C_{i22} + \alpha (p_{i0}(2) V_0 + p_{i1}(2) V_1 + p_{i2}(2) V_2 + p_{i3}(2) V_3)$ $= -1 + \frac{7}{6} \left(\frac{6}{23} + \frac{1}{90} + \frac{1}{21} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = -\frac{2}{2}$	2

3	0	$V_{i30} = C_{i30} + \alpha \sum_{j=0}^3 p_{ij}(0) V_j$ $= C_{i30} + \alpha (p_{i0}(0) V_0 + p_{i1}(0) V_1 + p_{i2}(0) V_2 + p_{i3}(0) V_3)$ $= 3 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = 3 + \frac{1}{18}$	3
3	1	$V_{i31} = C_{i31} + \alpha \sum_{j=0}^3 p_{ij}(1) V_j$ $= C_{i31} + \alpha (p_{i0}(1) V_0 + p_{i1}(1) V_1 + p_{i2}(1) V_2 + p_{i3}(1) V_3)$ $= 0 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = \frac{3}{23}$	3
3	2	$V_{i32} = C_{i32} + \alpha \sum_{j=0}^3 p_{ij}(2) V_j$ $= C_{i32} + \alpha (p_{i0}(2) V_0 + p_{i1}(2) V_1 + p_{i2}(2) V_2 + p_{i3}(2) V_3)$ $= -1 + \frac{7}{6} \left(0 \frac{23}{90} + \frac{1}{21} + \frac{6}{23} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = -\frac{17}{17}$	3
3	3	$V_{i33} = C_{i33} + \alpha \sum_{j=0}^3 p_{ij}(3) V_j$ $= C_{i33} + \alpha (p_{i0}(3) V_0 + p_{i1}(3) V_1 + p_{i2}(3) V_2 + p_{i3}(3) V_3)$ $= -2 + \frac{7}{6} \left(\frac{6}{23} + \frac{1}{90} + \frac{1}{21} + \frac{3}{23} + \frac{1}{6} + \frac{3}{23} + \frac{3}{23} + \frac{1}{3} \right) = -\frac{25}{25}$	3

Table 1: Compute $V_{ik} = C_{ik} + \alpha \sum_{j=0}^3 p_{ij}(k) V_j$.

State ? Decision k Second iteration, 2. Policy improvement

1	0	$V_{10} = C_{10} + \alpha \sum_{j=0}^3 p_{1j}(0) V_j$ $= C_{10} + \alpha (p_{10}(0) V_0 + p_{11}(0) V_1 + p_{12}(0) V_2 + p_{13}(0) V_3)$ $= 3 + \frac{1}{6} \left(0 + \frac{1}{3}(-1) + \frac{2}{3}(-2) + \frac{1}{3}(-3) \right) = 1$	1	$V_{11} = C_{11} + \alpha \sum_{j=0}^3 p_{1j}(1) V_j$ $= C_{11} + \alpha (p_{10}(1) V_0 + p_{11}(1) V_1 + p_{12}(1) V_2 + p_{13}(1) V_3)$ $= 0 + \frac{1}{6} \left(\frac{2}{3} 2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{3}(-3) \right) = -1$
2	0	$V_{20} = C_{20} + \alpha \sum_{j=0}^3 p_{2j}(0) V_j$ $= C_{20} + \alpha (p_{20}(0) V_0 + p_{21}(0) V_1 + p_{22}(0) V_2 + p_{23}(0) V_3)$ $= 3 + \frac{1}{6} \left(0 + 0(-1) + \frac{1}{3}(-2) + \frac{2}{3}(-3) \right) = \frac{1}{4}$	1	$V_{21} = C_{21} + \alpha \sum_{j=0}^3 p_{2j}(1) V_j$ $= C_{21} + \alpha (p_{20}(1) V_0 + p_{21}(1) V_1 + p_{22}(1) V_2 + p_{23}(1) V_3)$ $= 0 + \frac{1}{6} \left(0 + \frac{1}{3}(-1) + \frac{2}{3}(-2) + \frac{1}{3}(-3) \right) = -2$
2	2	$V_{22} = C_{22} + \alpha \sum_{j=0}^3 p_{2j}(2) V_j$ $= C_{22} + \alpha (p_{20}(2) V_0 + p_{21}(2) V_1 + p_{22}(2) V_2 + p_{23}(2) V_3)$ $= -1 + \frac{1}{6} \left(\frac{2}{3} 2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{3}(-3) \right) = -2$	2	
3	0	$V_{30} = C_{30} + \alpha \sum_{j=0}^3 p_{3j}(0) V_j$ $= C_{30} + \alpha (p_{30}(0) V_0 + p_{31}(0) V_1 + p_{32}(0) V_2 + p_{33}(0) V_3)$ $= 3 + \frac{1}{6} \left(0 + 0(-1) + 0(-2) + 1(-3) \right) = \frac{1}{3}$	1	$V_{31} = C_{31} + \alpha \sum_{j=0}^3 p_{3j}(1) V_j$ $= C_{31} + \alpha (p_{30}(1) V_0 + p_{31}(1) V_1 + p_{32}(1) V_2 + p_{33}(1) V_3)$ $= 0 + \frac{1}{6} \left(0 + 0(-1) + \frac{1}{3}(-2) + \frac{2}{3}(-3) \right) = -\frac{1}{12}$
3	2	$V_{32} = C_{32} + \alpha \sum_{j=0}^3 p_{3j}(2) V_j$ $= C_{32} + \alpha (p_{30}(2) V_0 + p_{31}(2) V_1 + p_{32}(2) V_2 + p_{33}(2) V_3)$ $= -1 + \frac{1}{6} \left(0 + \frac{1}{3}(-1) + \frac{2}{3}(-2) + \frac{1}{3}(-3) \right) = -3$	3	
3	3	$V_{33} = C_{33} + \alpha \sum_{j=0}^3 p_{3j}(3) V_j$ $= C_{33} + \alpha (p_{30}(3) V_0 + p_{31}(3) V_1 + p_{32}(3) V_2 + p_{33}(3) V_3)$ $= -2 + \frac{1}{6} \left(\frac{2}{3} 2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{3}(-3) \right) = -3$	3	

Table 2: Compute $V_{ik} = C_{ik} + \alpha \sum_{j=0}^3 p_{ij}(k) V_j$.