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, $t = 1, 2, \ldots$ 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 <u>State</u> i = number of units of water in the dam (at the beginning of the month).

$i = 0, 1, 2, 3$

<u>Decision</u> k = number of units of water that should be released during the month.

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

<u>Transition matrices</u> $P(k) = \{ p_{ij}(k) \}$

release no units ($k=0$):

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

$$\quad 0 \quad 1 \quad 2 \quad 3 \quad j$$

release 1 unit ($k=1$):

$$P(1) = \begin{pmatrix} - & - & - & - \\ 1/6 & 1/3 & 1/3 & 1/6 \\ 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 1/6 & 5/6 \end{pmatrix} \leftarrow k \geq i \text{ mean that all water is released.}$$

$$P(2) = \begin{pmatrix} - & - & - & - \\ - & - & - & - \\ 1/6 & 1/3 & 1/3 & 1/6 \\ 0 & 1/6 & 1/3 & 1/2 \end{pmatrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} k \geq i$$

$$P(3) = \begin{pmatrix} - & - & - & - \\ - & - & - & - \\ - & - & - & - \\ 1/6 & 1/3 & 1/3 & 1/6 \end{pmatrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} k \geq i$$

Expected cost for the coming month:

$C_{00} = 3$ (buy 1 unit)

C_{01}, C_{02}, C_{03} – correspond to infeasible decisions

$C_{10} = 3$ (buy 1 unit)

$C_{11} = 0$

C_{12}, C_{13} – not feasible (may not be possible)

$C_{20} = 3$ (buy 1 unit)

$C_{21} = 0$

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

C_{23} – not feasible

$C_{30} = 3$ (buy 1 unit)

$C_{31} = 0$

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

$C_{33} = -2$ (sell 2 units)

Decisions

$k=0$ = release no water = buy 1 unit to cover contracted distribution
$k=1$ = release 1 unit of water = deliver contracted amount
$k=2$ = release 2 — " — = — " — + 1 unit sold
$k=3$ = release 3 — " — = — " — + 2 units sold

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 $V_0, \ldots, V_3$:

$$\begin{cases} V_0 = C_{0d_0} + \alpha \cdot \sum_{j=0}^{3} P_{0j}(d_0) V_j = \\ \quad = 3 + \frac{6}{7}\left(\frac{1}{6} V_0 + \frac{1}{3} V_1 + \frac{1}{3} V_2 + \frac{1}{6} V_3\right) & \{P(0), i=0\} \\ V_1 = \overset{C_{1d_1}}{0} + \frac{6}{7}\left(\frac{1}{6} V_0 + \frac{1}{3} V_1 + \frac{1}{3} V_2 + \frac{1}{6} V_3\right) & \{P(1), i=1\} \\ V_2 = \overset{C_{2d_2}}{0} + \frac{6}{7}\left(1/6\, V_1 + \frac{1}{3} V_2 + \frac{1}{2} V_3\right) & \{P(1), i=2\} \\ V_3 = \overset{C_{3d_3}}{0} + \frac{6}{7}\left(\frac{1}{6} V_2 + \frac{5}{6} V_3\right) & \{P(1), i=3\} \end{cases}$$

solve

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

(2) Policy improvement

State (i)	Decision (k)	
1	0	$\tilde{V}_{10} = \underbrace{C_{10}}_{\text{(Note } C_{ik})} + \alpha \sum_{j=0}^{3} P_{1j}(0)\, V_j = 3 + \frac{6}{7}\left(0 V_0 + \frac{1}{6} V_1 + \frac{1}{3} V_2 + \frac{1}{2} V_3\right) = \frac{75}{23}$
1	1	$\tilde{V}_{11} = \dots = 21/23$
2	0	see handout sheet pp. 3–4
2	1	
2	2	
3	0	
3	1	
3	2	
3	3	

let d_i = minimizing k $\left(\min_k \{ \tilde{V}_{ik} \} \quad \forall i \right)$

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

$\Rightarrow$ new iteration.

<u>Second iteration</u>

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

gives

$R_2 = R_1 \Rightarrow$ optimal.

The policies $\left[0, 1, \binom{1}{2}, \binom{2}{3}\right]$ 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};

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

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

$$C^T = (C_{00},\ C_{10},\ \text{—} \quad \text{—} \quad \text{—} \quad ,\ C_{33})$$

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

- $(*) \quad \sum_{k} y_{jk} - \alpha \sum_{i,k} P_{ij}(k)\, y_{ik} = \beta_j \quad \forall j$ ← simplify (*

We do not need the constraint

$$\sum_{i,k} y_{ik} = 1$$

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.

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

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

$$\Big(y_{00}\Big) - \alpha\Big[\\ \tfrac{1}{6}y_{00} + 0y_{10} + 0y_{20} + 0y_{30} + \\ \tfrac{1}{6}y_{11} + 0y_{21} + 0y_{31} + \\ \tfrac{1}{6}y_{22} + 0y_{32} + \\ \tfrac{1}{6}y_{33}\Big] = \beta_0 = \tfrac{1}{4}.$$

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

For $j = 1$ we get

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

$$\Big(y_{10} + y_{11}\Big) - \alpha\Big[\tfrac{1}{3}y_{00} + \tfrac{1}{6}y_{10} + 0y_{20} + 0y_{30} + \\ \tfrac{1}{3}y_{11} + \tfrac{1}{6}y_{21} + 0y_{31} + \\ \tfrac{1}{3}y_{22} + \tfrac{1}{6}y_{32} + \\ \tfrac{1}{3}y_{33}\Big] = \beta_1 = \tfrac{1}{4}.$$

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

For $j = 2$ we get

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

$\Big(y_{20}+y_{21}+y_{22}\Big)-\alpha\Big[\frac{1}{3}y_{00}+\frac{1}{3}y_{10}+\frac{1}{6}y_{20}+0y_{30}+$
$\frac{1}{3}y_{11}+\frac{1}{6}y_{21}+\frac{1}{6}y_{31}+\frac{1}{3}y_{22}+\frac{1}{3}y_{32}+\frac{1}{3}y_{33}\Big]=\beta_2=\frac{1}{4}.$

$-\frac{1}{3}\alpha y_{00}-\frac{1}{3}\alpha y_{10}-\frac{1}{3}\alpha y_{11}+(1-\frac{1}{6}\alpha)y_{20}+(1-\frac{1}{6}\alpha)y_{21}+(1-\frac{1}{3}\alpha)y_{22}-\frac{1}{6}\alpha y_{31}-\frac{1}{3}\alpha y_{32}-\frac{1}{3}\alpha y_{33}=\frac{1}{4}.$

For $j=3$ we get
$\Big(y_{30}+y_{31}+y_{32}+y_{33}\Big)-\alpha\Big[$
$p_{03}(0)y_{00}+p_{13}(0)y_{10}+p_{23}(0)y_{20}+p_{33}(0)y_{30}+$
$p_{03}(1)y_{01}+p_{13}(1)y_{11}+p_{23}(1)y_{21}+p_{33}(1)y_{31}+$
$p_{03}(2)y_{02}+p_{13}(2)y_{12}+p_{23}(2)y_{22}+p_{33}(2)y_{32}+$
$p_{03}(3)y_{03}+p_{13}(3)y_{13}+p_{23}(3)y_{23}+p_{33}(3)y_{33}\Big]=\beta_3.$

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

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

The LP problem to be solved is

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

where $C^T=(3,\ 3,\ 0,\ 3,\ 0,\ -1,\ 3,\ 0,\ -1,\ -2)$,
$y^T=(y_{00},\ y_{10},\ y_{11},\ y_{20},\ y_{21},\ y_{22},\ y_{23},\ y_{31},\ y_{32},\ y_{33})$,

$$A=\begin{pmatrix} 1-\frac{\alpha}{6} & 0 & -\frac{1}{6} & 0 & 0 & -\frac{1}{6} & 0 & 0 & 0 & -\frac{1}{6} \\ -\frac{1}{3} & 1-\frac{\alpha}{6} & 1-\frac{\alpha}{3} & 0 & -\frac{1}{6} & -\frac{1}{3} & 0 & 0 & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 1-\frac{\alpha}{6} & 1-\frac{\alpha}{3} & 1-\frac{\alpha}{3} & 0 & -\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{2} & -\frac{1}{6} & -\frac{5}{6} & -\frac{1}{2} & -\frac{1}{6} & 0 & 1-\frac{5\alpha}{6} & 1-\frac{\alpha}{2} & 1-\frac{\alpha}{6} \end{pmatrix} \text{ and } b=\begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}.$$

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 $D^*_{ik}=\frac{y^*_{ik}}{\sum_k y^*_{ik}}$. In this case
$D^*_{00}=1$, $D^*_{10}=0$, $D^*_{11}=1$, $D^*_{20}=0$, $D^*_{21}=1$, $D^*_{22}=0$, $D^*_{30}=0$, $D^*_{31}=0$, $D^*_{32}=0$ and $D^*_{33}=1$, i.e. $d^*_0=0$, $d^*_1=1$, $d^*_2=1$ and $d^*_3=3$.

$R^*=[d^*_0,\ d^*_1,\ d^*_2,\ d^*_3]=[0,\ 1,\ 1,\ 3]$.

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

2. Policy improvment
see table 1.

Let $d_i =$ 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:

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

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

2. Policy improvment
see table 2.

Let $d_i =$ 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 improvment**
1	0	$\tilde{V}_{10} = C_{10} + \alpha \sum_{j=0}^{3} p_{1j}(0)V_j$ $= C_{10} + \alpha\left(p_{10}(0)V_0 + p_{11}(0)V_1 + p_{12}(0)V_2 + p_{13}(0)V_3\right)$ $= 3 + \frac{6}{7}\left(0\frac{90}{23} + \frac{1}{6}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{2}\frac{3}{23}\right) = \frac{75}{23}$
1	1	$\tilde{V}_{11} = C_{11} + \alpha \sum_{j=0}^{3} p_{1j}(1)V_j$ $= C_{11} + \alpha\left(p_{10}(1)V_0 + p_{11}(1)V_1 + p_{12}(1)V_2 + p_{13}(1)V_3\right)$ $= 0 + \frac{6}{7}\left(\frac{1}{6}\frac{90}{23} + \frac{1}{3}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{6}\frac{3}{23}\right) = \mathbf{\frac{21}{23}}$
2	0	$\tilde{V}_{20} = C_{20} + \alpha \sum_{j=0}^{3} p_{2j}(0)V_j$ $= C_{20} + \alpha\left(p_{20}(0)V_0 + p_{21}(0)V_1 + p_{22}(0)V_2 + p_{23}(0)V_3\right)$ $= 3 + \frac{6}{7}\left(0\frac{90}{23} + 0\frac{21}{23} + \frac{1}{6}\frac{6}{23} + \frac{5}{6}\frac{3}{23}\right) = \frac{72}{23}$
2	1	$\tilde{V}_{21} = C_{21} + \alpha \sum_{j=0}^{3} p_{2j}(1)V_j$ $= C_{21} + \alpha\left(p_{20}(1)V_0 + p_{21}(1)V_1 + p_{22}(1)V_2 + p_{23}(1)V_3\right)$ $= 0 + \frac{6}{7}\left(0\frac{90}{23} + \frac{1}{6}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{2}\frac{3}{23}\right) = \frac{6}{23}$
2	2	$\tilde{V}_{22} = C_{22} + \alpha \sum_{j=0}^{3} p_{2j}(2)V_j$ $= C_{22} + \alpha\left(p_{20}(2)V_0 + p_{21}(2)V_1 + p_{22}(2)V_2 + p_{23}(2)V_3\right)$ $= -1 + \frac{6}{7}\left(\frac{1}{6}\frac{90}{23} + \frac{1}{3}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{6}\frac{3}{23}\right) = -\mathbf{\frac{2}{23}}$
3	0	$\tilde{V}_{30} = C_{30} + \alpha \sum_{j=0}^{3} p_{3j}(0)V_j$ $= C_{30} + \alpha\left(p_{30}(0)V_0 + p_{31}(0)V_1 + p_{32}(0)V_2 + p_{33}(0)V_3\right)$ $= 3 + \frac{6}{7}\left(0\frac{90}{23} + 0\frac{21}{23} + 0\frac{6}{23} + \frac{1}{6}\frac{3}{23}\right) = 3 + \frac{18}{161}$
3	1	$\tilde{V}_{31} = C_{31} + \alpha \sum_{j=0}^{3} p_{3j}(1)V_j$ $= C_{31} + \alpha\left(p_{30}(1)V_0 + p_{31}(1)V_1 + p_{32}(1)V_2 + p_{33}(1)V_3\right)$ $= 0 + \frac{6}{7}\left(0\frac{90}{23} + 0\frac{21}{23} + \frac{1}{6}\frac{6}{23} + \frac{5}{6}\frac{3}{23}\right) = \frac{3}{23}$
3	2	$\tilde{V}_{32} = C_{32} + \alpha \sum_{j=0}^{3} p_{3j}(2)V_j$ $= C_{32} + \alpha\left(p_{30}(2)V_0 + p_{31}(2)V_1 + p_{32}(2)V_2 + p_{33}(2)V_3\right)$ $= -1 + \frac{6}{7}\left(0\frac{90}{23} + \frac{1}{6}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{2}\frac{3}{23}\right) = -\frac{17}{23}$
3	3	$\tilde{V}_{33} = C_{33} + \alpha \sum_{j=0}^{3} p_{3j}(3)V_j$ $= C_{33} + \alpha\left(p_{30}(3)V_0 + p_{31}(3)V_1 + p_{32}(3)V_2 + p_{33}(3)V_3\right)$ $= -2 + \frac{6}{7}\left(\frac{1}{6}\frac{90}{23} + \frac{1}{3}\frac{21}{23} + \frac{1}{3}\frac{6}{23} + \frac{1}{6}\frac{3}{23}\right) = -\mathbf{\frac{25}{23}}$

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

State i	Decision k	Second iteration, 2. Policy improvment
1	0	$\tilde{V}_{10} = C_{10} + \alpha \sum_{j=0}^{3} p_{1j}(0)V_j = C_{10} + \alpha\Big(p_{10}(0)V_0 + p_{11}(0)V_1 + p_{12}(0)V_2 + p_{13}(0)V_3\Big) = 3 + \frac{6}{7}\Big(0 + \frac{1}{6}(-1) + \frac{1}{3}(-2) + \frac{1}{2}(-3)\Big) = 1$
1	1	$\tilde{V}_{11} = C_{11} + \alpha \sum_{j=0}^{3} p_{1j}(1)V_j = C_{11} + \alpha\Big(p_{10}(1)V_0 + p_{11}(1)V_1 + p_{12}(1)V_2 + p_{13}(1)V_3\Big) = 0 + \frac{6}{7}\Big(\frac{1}{6}2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{6}(-3)\Big) = \mathbf{-1}$
2	0	$\tilde{V}_{20} = C_{20} + \alpha \sum_{j=0}^{3} p_{2j}(0)V_j = C_{20} + \alpha\Big(p_{20}(0)V_0 + p_{21}(0)V_1 + p_{22}(0)V_2 + p_{23}(0)V_3\Big) = 3 + \frac{6}{7}\Big(0 + 0(-1) + \frac{1}{6}(-2) + \frac{5}{6}(-3)\Big) = \frac{4}{7}$
2	1	$\tilde{V}_{21} = C_{21} + \alpha \sum_{j=0}^{3} p_{2j}(1)V_j = C_{21} + \alpha\Big(p_{20}(1)V_0 + p_{21}(1)V_1 + p_{22}(1)V_2 + p_{23}(1)V_3\Big) = 0 + \frac{6}{7}\Big(0 + \frac{1}{6}(-1) + \frac{1}{3}(-2) + \frac{1}{2}(-3)\Big) = \mathbf{-2}$
2	2	$\tilde{V}_{22} = C_{22} + \alpha \sum_{j=0}^{3} p_{2j}(2)V_j = C_{22} + \alpha\Big(p_{20}(2)V_0 + p_{21}(2)V_1 + p_{22}(2)V_2 + p_{23}(2)V_3\Big) = -1 + \frac{6}{7}\Big(\frac{1}{6}2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{6}(-3)\Big) = \mathbf{-2}$
3	0	$\tilde{V}_{30} = C_{30} + \alpha \sum_{j=0}^{3} p_{3j}(0)V_j = C_{30} + \alpha\Big(p_{30}(0)V_0 + p_{31}(0)V_1 + p_{32}(0)V_2 + p_{33}(0)V_3\Big) = 3 + \frac{6}{7}\Big(0 + 0(-1) + 0(-2) + 1(-3)\Big) = \frac{3}{7}$
3	1	$\tilde{V}_{31} = C_{31} + \alpha \sum_{j=0}^{3} p_{3j}(1)V_j = C_{31} + \alpha\Big(p_{30}(1)V_0 + p_{31}(1)V_1 + p_{32}(1)V_2 + p_{33}(1)V_3\Big) = 0 + \frac{6}{7}\Big(0 + 0(-1) + \frac{1}{6}(-2) + \frac{5}{6}(-3)\Big) = -\frac{17}{7}$
3	2	$\tilde{V}_{32} = C_{32} + \alpha \sum_{j=0}^{3} p_{3j}(2)V_j = C_{32} + \alpha\Big(p_{30}(2)V_0 + p_{31}(2)V_1 + p_{32}(2)V_2 + p_{33}(2)V_3\Big) = -1 + \frac{6}{7}\Big(0 + \frac{1}{6}(-1) + \frac{1}{3}(-2) + \frac{1}{2}(-3)\Big) = \mathbf{-3}$
3	3	$\tilde{V}_{33} = C_{33} + \alpha \sum_{j=0}^{3} p_{3j}(3)V_j = C_{33} + \alpha\Big(p_{30}(3)V_0 + p_{31}(3)V_1 + p_{32}(3)V_2 + p_{33}(3)V_3\Big) = -2 + \frac{6}{7}\Big(\frac{1}{6}2 + \frac{1}{3}(-1) + \frac{1}{3}(-2) + \frac{1}{6}(-3)\Big) = \mathbf{-3}$

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