DPRL Assignment 2: Markov Chains

Zhiyan Yao
Vrije Universiteit Amsterdam
z.yao2@student.vu.nl

Bo-Chian Chen
Vrije Universiteit Amsterdam
b.chen4@student.vu.nl

I. Problem Definition

Given a system that slowly deteriorates over time:

- When it is new is has a failure probability of 0.1
- The probability of failure increases every time unit linearly with 0.01
- Replacement costs are 1, after replacement the part is new

II. QUESTION A

A. *The state space $\mathcal{X}$*

The state space for this deteriorating system can be defined as the set of discrete states representing the level of deterioration. Let's denote the state space as $\mathcal{X}$:

$$\mathcal{X} : \{0, 1, 2, 3, \ldots, 90\}$$

The state space is the level of deterioration, which can range from 0 to 90. In addition, each state represents a certain level of deterioration of the system. The Markov Chain is depicted in figure 1, where tuple (p, r) represents transition probability, which is $0.1 + 0.01 * level\_of\_deterioration$, and replacement cost, which is always 1. Notice that the cost used in later questions is cost expectation $p \times r$.

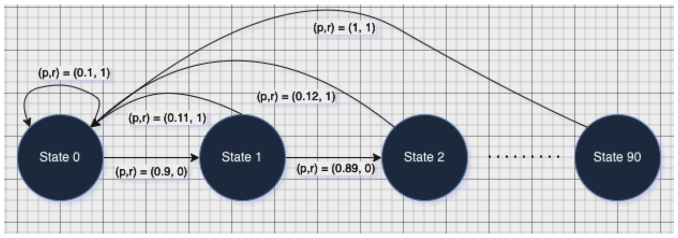

Fig. 1. System Markov Chain

B. *Markov property*

The Markov property is satisfied because the future state of the system (at time t+1) depends only on its current state (at time t) and not on the sequence of events that preceded it. In this case, the probability of failure at time t+1 depends only on the current state of deterioration at time t.

III. QUESTION B

A. *Compute the stationary distribution*

Our computation is based on the following assumptions:

Let's set $\pi_0 = \{1, 0, 0, 0, \ldots, 0\}$, all 0 except for the first one. After computing for 100000 times, we get the stationary distribution π_* (shown in programming b.py result).

B. *Long-run average replacement costs*

We can also get the long-run reward ϕ_*, by letting $r_*(i) = 0.1 + 0.01 * i$. As a result, we can get the long-run average replacement costs, which is 0.1461075, by computing $\pi_* \times r_*$.

IV. QUESTION C

A. *Simulation*

We use a random float number which is between 0 and 1. If it is equal to or smaller that the failure probability, the total cost will be added replacement cost 1. In addition, the state (the level of deterioration) will be reset to 0.

B. *Results*

The result is 0.1461297, which is close to our long-run average replacement costs 0.1461075.

V. QUESTION D

A. *The average-cost Poisson equation*

Firstly, based on Poisson equation

$$V_*(x) + \phi_* = r(x) + \sum_y p(y|x) V_*(y)$$

we can derive the general formula, where i is from 0 to 90:

$$\begin{aligned} V(i) + \phi &= (0.1 + 0.01 \times i) \times (1 + V(0)) \\ &\quad + (0.9 - 0.01 \times i) \times V((i+1) \mod 91) \end{aligned}$$

Secondly, let's list the formulas:

$$\begin{aligned} V(0) + \phi &= 0.1 + 0.9 \times V(1) + 0.1 \times V(0) \\ V(1) + \phi &= 0.11 + 0.89 \times V(2) + 0.11 \times V(0) \\ V(2) + \phi &= 0.12 + 0.88 \times V(3) + 0.12 \times V(0) \\ &\ldots \\ V(90) + \phi &= 1 + V(0) \end{aligned}$$

Thirdly, we rewrite the formulas and set V(0) = 0, listed as below:

$$
\begin{aligned}
V(0) - 0.9 \times V(1) - 0.1 \times V(0) + \phi &= 0.1 \\
V(1) - 0.89 \times V(2) - 0.11 \times V(0) + \phi &= 0.11 \\
V(2) - 0.88 \times V(3) - 0.12 \times V(0) + \phi &= 0.12 \\
&\ldots \\
V(90) - V(0) + \phi &= 1
\end{aligned}
$$

Lastly, we compute the $V(1), V(2), \ldots, V(90), \phi$.

B. *Results interpretation*

After compute the Poisson equation, we get the the average cost: 0.14609. And the result is nearly the same as the long-run average replacement cost in question (B), thus we do the right calculation and prove it correctly.

VI. QUESTION E

Given additional condition that preventive replacement is possible at every state with cost 0.6, any state can be renewed by preventive replacement and thus go back to the initial state 0 with cheaper cost 0.6. The new system Markov chain with preventive replacement is depicted in figure 2, where black lines represent action 1 (no action and may fail) and red lines represent action 2 (preventive replacement). This is a Markov decision chain and average optimal policy can be derived by policy iteration or value iteration.

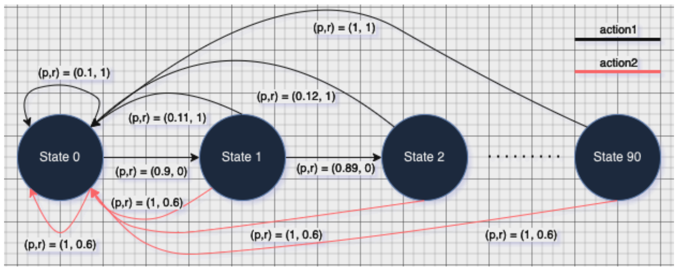

Fig. 2. System Markov Chain with Preventive Replacement

A. *Policy iteration*

We can do the following policy iterations to get average optimal policy:

1) Fix an initial policy α, in this case we take α with all action 1 as our initial policy.
2) Find a solution (V_α, ϕ_α) of $V + \phi e = r_\alpha + P_\alpha V$
3) Compute $\tilde{\alpha} = \underset{\alpha'}{\operatorname{argmin}}\{r_{\alpha'} + P_{\alpha'} V_\alpha\}$
4) If $\alpha = \tilde{\alpha}$, terminate with (V_α, ϕ_α), which is the average optimal policy, otherwise, let $\alpha = \tilde{\alpha}$ and go back to step 1 for a new iteration.

For more implementation details please refer to e.py.

B. *Value iteration*

We can do the following value iterations to get average optimal policy:

1) Start from some value function V_t, in this case we take V_t with all 0 as our initial value function.
2) Repeat $V_{t+1} = \underset{\alpha}{min}\{r_\alpha + P_\alpha V_t\}$, new policy $\alpha' = \underset{\alpha}{argmin}\{r_\alpha + P_\alpha V_t\}$
3) Stop when $span\{v_{t+1} - v_t\} \leq \epsilon$, where $\epsilon = 1e-5$ in our case. And when converged, $\phi e = V_{t+1} - V_t$

For more implementation details please refer to e.py.

C. *Result*

The average optimal policy α we get from both policy iteration and value iteration is

$$
\begin{gathered}
\alpha = \{1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1, \\
2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2, \\
2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2, \\
2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2\}
\end{gathered}
$$

And long-run average replacement cost $\phi = 0.14587$.

The first 17 actions in optimal policy is 1, which means in the early stage of deterioration, preventive replacement is not needed because fail rate is low. After fail rate is bigger than 0.26, preventive replacement jumps in to avoid failure with cost 1. Under the optimal policy, long-run average replacement cost decreases by a small margin from 0.1461075 to 0.14587.