Lecture 9: Exercises

October 2023

In the following exercises, we investigate a system consisting of 4 identical components, which can either be in functioning or failed state. The lifetimes of the respective components are i.i.d. with distribution:

$$X_i \sim exp(\beta)$$

for $1 \leq i \leq 4$ and $\beta \in \mathbb{R}^+$. Failure is irreversible. The system is inspected at regular intervals of τ time units. The state of the system is given as the number of functioning components at the time of inspection. Thus, the state space is given by:

$$S = \{0, ..., 4\}$$

Without maintenance, the system will deteriorate according to a DTMC.

Exercise 1: Show that the transition probabilities between two inspections are given by:

$$p(t|s) = \binom{s}{t} e^{-\beta\tau t}(1 - e^{-\beta\tau})^{s-t} \tag{1}$$

for $0 \leq t \leq s \leq 4$. *Hint:* Find expressions for $P(X_i \leq \tau) = F(\tau)$ and $P(X_i > \tau) = R(\tau)$. ■

At each inspection, we have the opportunity to replace any number of failed components. Thus, in state $s \in S$, the action set is given by:

$$A_s = \{a_0, ..., a_{4-s}\}$$

where a_i is the action of replacing i components. Failed components can be replaced immediately and are substituted by statistically equivalent (functioning) components.

Exercise 2: Find a general expression for the transition probabilities

$$p(t|s, a_i)$$

for $0 \leq t, s \leq 4$ and $a_i \in A_s$. *Hint:* Use the expression in (1). ■

Assume now, that the cost function is given by:

$$c(s, a_i, t) = \mathbb{1}_{i>0} \cdot \alpha + i \cdot \beta + \mathbb{1}_{t=0} \cdot \gamma$$

for $\alpha, \beta, \gamma \in \mathbb{R}^+$ where $\mathbb{1}_{(\cdot)}$ is the indicator function and t is the realisation of the random variable T representing the successor state (at the next inspection). Here, α and β may represent the fixed and variable cost of replacing components, while γ may be the breakdown/downtime cost.

Exercise 3: Show that the expected cost when taking action a_0 in state $s \in S$ is given by:

$$\bar{c}(s, a_0) := \mathbb{E}\{c(s, a_0, T)\} = \gamma(1 - e^{-\beta\tau})^s$$

What does this say about the cost of systematically leaving the system unmaintained in the long run? ∎

We are now given a (stationary deterministic) policy $\pi : S \to A$. Consider the sequence of state-action pairs $S_0, \pi(S_0), S_1, \pi(S_1), S_2, \pi(S_2), \ldots$ and let

$$c(S_0, \pi(S_0), S_1),\ \ c(S_1, \pi(S_1), S_2),\ \ c(S_2, \pi(S_2), S_3), \ldots$$

be the sequential costs incurred in the first, second, third etc. decision epochs. To reduce the weight of future costs, we introduce the discount factor $0 < \lambda < 1$ and consider the discounted costs

$$c(S_0, \pi(S_0), S_1),\ \ \lambda c(S_1, \pi(S_1), S_2),\ \ \lambda^2 c(S_2, \pi(S_2), S_3), \ldots$$

as an alternative cost measure when evaluating long-term plans.

Exercise 4: Show that the *total discounted cost*:

$$\sum_{i=0}^{\infty} \lambda^i c(S_i, \pi(S_i), S_{i+1})$$

over an *infinite horizon* is well-defined. ∎

It is now given that the system, when unmaintained, deteriorates according to the DTMC:

$$
\begin{array}{ccccc}
4 & 3 & 2 & 1 & 0
\end{array}
$$
$$
M_{\varnothing} = \left[\begin{array}{ccccc}
.0625 & .2500 & .3750 & .2500 & .0625 \\
\cdot & .1250 & .3750 & .3750 & .1250 \\
\cdot & \cdot & .2500 & .5000 & .2500 \\
\cdot & \cdot & \cdot & .5000 & .5000 \\
\cdot & \cdot & \cdot & \cdot & 1
\end{array}\right]
\begin{array}{c} 4 \\ 3 \\ 2 \\ 1 \\ 0 \end{array}
$$

Assume, that we adopt the (stationary deterministic) maintenance policy $\pi : S \to A$ given by:

$$\pi(4) = \pi(3) = a_0,\ \pi(2) = a_2,\ \pi(1) = a_3,\ \pi(0) = a_4$$

That is, if there are 2 or fewer functioning components, π prescribes replacing all failed components. Else, π prescribes doing nothing (a_0). Implementing a policy reduces an MDP to a DTMC (why?).

Exercise 5: Write up the Markov chain:

$$M_\pi \in \mathbb{R}^{5\times 5}$$

induced by the policy π. *Hint:* all the rows in M_π can be read off $M_\emptyset$. ■

Assume now, that the expected cost function $\bar{c}$ takes the following values:

Table 1: Expected cost of state action pairs under π

s	4	3	2	1	0
$\pi(s)$	0	0	2	3	4
$\bar{c}(s, \pi(s))$	625	1,250	2,825	3,425	4,025

Exercise 6: Determine the long run expected cost per epoch under the maintenance policy π. *Hint:* Under the policy π the MDP is reduced to a DTMC which has a limiting distribution. ■

Solutions:

Exercise 1: The probability that a single functioning component survives another period (between inspections) is given by:

$$R(\tau) = \int_{\tau}^{\infty} \beta e^{-\beta t} dt = e^{-\beta\tau}$$

Also, the probability that a single functioning component fails before the next inspection is $F(\tau) = 1 - e^{-\beta\tau}$. Starting with s functioning components, the number of components surviving a period is binomially distributed with:

$$\binom{s}{t} R(\tau)^t F(\tau)^{s-t}$$

The expression in (1) follows directly.

Exercise 2: As noted, replacement is immediate. Thus, after replacement, we have:

$$p(t|s, a_i) = p(t|s+i)$$

Referring to the result from Exercise 1, we conclude that:

$$p(t|s, a_i) = \binom{s+i}{t} e^{-\beta\tau t}(1 - e^{-\beta\tau})^{s+i-t}$$

for $t \leq s+i$. Because failure is irreversible, we have:

$$p(t|s, a_i) = 0$$

for $t > s+i$.

Exercise 3: Note that the cost function reduces to:

$$c(s, a_0, T) = \mathbb{1}_{0>0} \cdot \alpha + 0 \cdot \beta + \mathbb{1}_{T=0} \cdot \gamma = \mathbb{1}_{T=0} \cdot \gamma$$

under the action a_0. Therefore the expected cost conditioned on s and a_0 is given by:

$$\mathbb{E}\{c(s, a_0, T)\} = \mathbb{E}\{\mathbb{1}_{T=0} \cdot \gamma \mid s, a_0\} = \gamma P(T = 0 \mid s, a_0) = \gamma F(\tau)^{s+0} = \gamma(1 - e^{-\beta\tau})^s$$

as claimed. If maintenance is systematically neglected, the system will tend to the state $s = 0$ and the cost per period will eventually be γ.

Exercise 4: Note that the cost function is bounded by:

$$c(s, a_i, t) \leq \alpha + 4 \cdot \beta + \gamma := c^+$$

Since the discount factor satisfies: $|\lambda| < 1$ we have:

$$\sum_{i=0}^{\infty} \lambda^i c(S_i, \pi(S_i), S_{i+1}) \leq \sum_{i=0}^{\infty} \lambda^i c^+ = \frac{c^+}{1-\gamma}$$

Thus, the discounted total cost is increasing an bounded above, which implies that it converges.

Exercise 5: Note, that:

$$\pi(4) = \pi(3) = a_0$$

which implies that the system under policy π is left unmaintained when observed in state $s \geq 3$. Therefore:

$$M_\pi[3:4,] = M_\emptyset[3:4,]$$

(using computer-inspired notation). Furthermore, when $0 \leq s \leq 2$, all failed components are immediately replaced, which means that:

$$M_\pi[0,] = M_\pi[1,] = M_\pi[2,] = M_\emptyset[4,]$$

The resulting matrix is given as:

$$M_\pi = \begin{array}{c} \\ \end{array} \begin{array}{ccccc} 4 & 3 & 2 & 1 & 0 \end{array}$$

$$M_\pi = \begin{bmatrix} .0625 & .2500 & .3750 & .2500 & .0625 \\ \cdot & .1250 & .3750 & .3750 & .1250 \\ .0625 & .2500 & .3750 & .2500 & .0625 \\ .0625 & .2500 & .3750 & .2500 & .0625 \\ .0625 & .2500 & .3750 & .2500 & .0625 \end{bmatrix} \begin{array}{c} 4 \\ 3 \\ 2 \\ 1 \\ 0 \end{array}$$

Exercise 6: Since M_π is regular, the limiting probabilities exist and can be found with first-step analysis:

$$\rho = (\rho_4, ..., \rho_0) \approx (0.0486,\ 0.2222,\ 0.3750,\ 0.2778,\ 0.0764)$$

The long run expected cost per time period is given by:

$$\sum_{s=0}^{4} \rho_s \cdot \bar{c}(s, \pi(s)) \approx 2626.39$$

which is the expected cost of the state-action pairs $(s, \pi(s))_{s\in S}$ weighted by the long-run fraction of time that the process is expected to be in the respective states.