Chapter 1

The simple discounted problem

In this chapter, a simplified version of the problem is formulated and we attempt to solve it. In this simple problem, the time to live of the machine is a random variable and no data about the condition of the machine is observed before it breaks.

1.1 Problem Formulation

We consider a machine that is subject to deterioration over time. Calendar time is discretized in steps of size δ, i.e. the k'th decision stage is at time $t_k = k\delta$. We refer to the interval $(k\delta, (k+1)\delta]$ as the k'th time interval. At stage $k \in \mathbb{N} \cup \{0\}$, we denote the state of the machine by $x_k \in X = \mathbb{N} \cup \{0\}$ and initially $x_0 = 1$. When $x_k = x$, this means that after the k'th time interval, the machine has age $x\delta$. We now introduce some random disturbances $\omega_k := \omega_k(x_k)$ at decision stage k which only depend on x_k.

$$\omega_k(x_k) := \begin{cases} 1, & \text{if the machine lives longer than the } k\text{'th decision time given} \\ & \text{that it has been alive for } x_k \text{ decision times} \\ 0, & \text{else.} \end{cases} \tag{1.1}$$

At times step k, we shall choose an action u_k from the action set $U(x_k)$. Where

$$U(x_k) := \begin{cases} \{a_W, a_R\}, & \text{if } x_k > 0 \\ \{a_R\}, & \text{if } x_k = 0. \end{cases} \tag{1.2}$$

These actions are

- a_R: Repair (or replace) the machine.
- a_W: Wait and do not repair the machine.

The state evolves in the following way

$$x_{k+1} = f(x_k, u_k, \omega_k) := \begin{cases} 1, & \text{if } u_k = a_R \\ 0, & \text{if } u_k = a_W \text{ and } \omega_k = 0 \\ x_k + 1, & \text{if } u_k = a_W \text{ and } \omega_k = 1. \end{cases} \tag{1.3}$$

For convenience, we define the random variable $S(x_k) := f(x_k, a_W, \omega_k)$. Note that in the above definition, repairing the machine takes exactly one time interval. When the machine is in state x_k and the action u_k is chosen, the following cost is incurred

$$g(x_k, a_k) := \begin{cases} c + a, & \text{if } x_k = 0 \\ c, & \text{if } x_k > 0 \text{ and } u_k = a_R \\ 0, & \text{else.} \end{cases} \tag{1.4}$$

Furthermore, a discount α_δ is introduced such that costs n decision stages in the future are discounted by α_δ^n. We consider the total discounted cost

$$V(x_k) = \sum_{m=k}^{\infty} \alpha_\delta^m g(x_m, u_k) \tag{1.5}$$

We want to find a stationary policy $\mu : X \to \{a_W, a_R\}$ that chooses the action $u_k = \mu(x_k)$ that minimizes the expected total discounted cost $V_\mu(x_k)$ for each state. $V_\mu(x_k)$ is given by

$$V_\mu(x_k) = g(x_k, \mu(x_k)) + \alpha_\delta \mathbb{E}[V_\mu(f(x_k, \mu(x_k), \omega_k))]$$

The Bellman equations for the optimal cost V^* read

$$V^*(x_k) = \begin{cases} \min\{c + \alpha_\delta V^*(1), \alpha_\delta \mathbb{E}[V^*(S(x_k))]\}, & \text{if } x_k > 0 \\ c + a + \alpha_\delta V^*(1), & \text{else.} \end{cases} \tag{1.6}$$

μ is optimal if $V_\mu(x) = V^*(x)$ for all x.

1.1.1 Alternative models

Instantaneous repairs

If we would want to make repairs take zero time, we would have to change the definition of x_{k+1} to

$$x_{k+1} = f_2(x_k, u_k, \omega_k) := \begin{cases} 0, & \text{if } \omega_k = 0 \\ 2, & \text{if } u_k = a_R \text{ and } \omega_k = 1 \\ x_k + 1, & \text{if } u_k = a_W \text{ and } \omega_k = 1. \end{cases} \tag{1.7}$$

This would however, introduce the possibility of having to correctively repair the machine twice in a row.

Stochastic inter-decision times

Another possibility would be to have positive stochastic i.i.d. inter-decision times Δ_k and to use a continuous discount such that costs at time t are discounted by $e^{-\beta t}$ for $\beta > 0$. The state space could then be modeled as $X = \mathbb{R}^+ \cup \{0, x_f\}$ where x_f denotes that the machine is broken. The state evolution would be as follows

$$x_{k+1} = f(x_k, u_k, \omega_k, \Delta_k) := \begin{cases} \Delta_k, & \text{if } u_k = a_R \\ x_f, & \text{if } u_k = a_W \text{ and } \omega_k = 0 \\ x_k + \Delta_k, & \text{if } u_k = a_W \text{ and } \omega_k = 1. \end{cases} \tag{1.8}$$

The Bellman equations should then be changed to

$$V^*(x_k) = \begin{cases} \min\{c + \mathbb{E}[e^{-\beta\Delta} V^*(\Delta)], \mathbb{E}[e^{-\beta\Delta} V^*(f(x_k, a_W, \omega_k, \Delta))], & \text{if } x_k \neq x_f \\ c + a + \mathbb{E}[e^{-\beta\Delta} V^*(\Delta)], & \text{else.} \end{cases} \tag{1.9}$$

Where repair would again take one inter-decision time. Where Δ is of the same family of i.i.d. random variables as the Δ_i's.

Time to live in state information

The time-to-live could also be included in the state information. We would then denote the cost when the machine still has a time q to live by $V(q)$ and in each time step, this time-to-live would decrease by δ until the time-to-live is negative, when the machine must be repaired. The total discounted cost would then be $\mathbb{E}[V(Q)]$ and the Bellman equations would be:

$$V(q) = \min\{c + \alpha_\delta \mathbb{E}[V(Q)], \alpha_\delta V(q - \delta)\}$$

for $q > 0$ and

$$V(q) = c + a + \alpha_\delta \mathbb{E}[V(Q)],$$

else. These equations are easy to solve although the resulting policy is not very useful in a realistic setting: If $0 < q \leq \delta$, then

$$\begin{aligned} V(q) &= \min\{c + \alpha_\delta \mathbb{E}[V(Q)], \alpha_\delta V(q - \delta)\} \\ &= \min\{c + \alpha_\delta \mathbb{E}[V(Q)], \alpha_\delta (c + a + \alpha_\delta \mathbb{E}[V(Q)])\} \\ &= c + \alpha_\delta \mathbb{E}[V(Q)] \end{aligned} \tag{1.10}$$

(Assuming $a > \frac{1-\alpha_\delta}{\alpha_\delta}(c + \alpha_\delta \mathbb{E}[V(Q)])$, note that for sufficiently small discount factor α_δ, this does not hold.). And by induction, we can prove that for $(k-1)\delta < q \leq k\delta$ $(k > 0)$

$$\begin{aligned} V(q) &= \min\{c + \alpha_\delta \mathbb{E}[V(Q)], \alpha_\delta V(q - \delta)\} \\ &= \min\{c + \alpha_\delta \mathbb{E}[V(Q)], \alpha_\delta \alpha_\delta^{k-2} (c + \alpha_\delta \mathbb{E}[V(Q)])\} \\ &= \alpha_\delta^{k-1} (c + \alpha_\delta \mathbb{E}[V(Q)]). \end{aligned} \tag{1.11}$$

As can be seen, the cost does not depend on a as correctively repairing never occurs. This is because in this formulation, the time-to-live is known when making the decision (when choosing the minimum). Such that we will repair at the last opportunity before the machine breaks down. If, for instance, we choose $alpha_\delta = e^{-\beta\delta}$ for $\beta > 0$, we have for $(k-1)\delta < q \leq k\delta$

$$V(q) = e^{-\beta(k-1)\delta}(c + \alpha_\delta \mathbb{E}[V(Q)]).$$

If we let δ approach zero, the cost approaches

$$V(q) = e^{-\beta q}(c + \alpha_\delta \mathbb{E}[V(Q)]).$$

As can be seen, this is the cost of repairing the machine preventively at exactly the instant that it will break, which is indeed optimal but can unfortunately not be realized as the time-to-live cannot be directly observed.

1.2 Solving the simple discounted problem

In this section, the following Bellman equations will be solved

$$V(x) = \begin{cases} \min\{c + \alpha_\delta V(1), \alpha_\delta \mathbb{E}[V(S(x))]\}, & \text{if } x > 0 \\ c + a + \alpha_\delta V(1), & \text{else.} \end{cases} \tag{1.12}$$

Where $\mathbb{P}(S(x) = 0) = \mathbb{P}(Q \leq x+\delta | Q \geq x) = \delta h(x) + o(\delta^2)$ (for lifetime $Q \sim F(x)$ and corresponding hazard rate $h(x)$) and $\mathbb{P}(S(x) = x + 1) = 1 - \delta h(x) + o(\delta^2)$ and $\alpha_\delta = e^{-\beta\delta} = 1 - \beta\delta + o(\delta^2)$ for $\beta > 0$. We define $V_\delta(n\delta) := V(n)$ and for convenience, we define $V_\delta(0^+) := V_\delta(\delta)$. If we assume that

$$c + \alpha_\delta V_\delta(0^+) > \alpha_\delta \mathbb{E}[V_\delta(S(x)\delta)],$$

we can write

$$V_\delta(x) = \alpha_\delta \mathbb{P}(Q \leq x+\delta | Q \geq x)(c+a+\alpha_\delta V(0^+)) + \alpha_\delta \mathbb{P}(Q > x+\delta | Q \geq x) V_\delta(x+\delta) \tag{1.13}$$

We are now going to let δ approach zero.

$$\begin{aligned} \lim_{\delta\to 0} V_\delta(x) = \lim_{\delta \to 0} & (1 - \beta\delta + o(\delta^2))(\delta h(x) + o(\delta^2))(c + a + (1 - \beta\delta + o(\delta^2))V'(0^+)) \\ & + (1 - \beta\delta + o(\delta^2))(1 - \delta h(x) + o(\delta^2)) V_\delta(x + \delta). \end{aligned} \tag{1.14}$$

Gathering the terms of $o(\delta^2)$, we get

$$\lim_{\delta\to 0} V_\delta(x) = \lim_{\delta\to 0} \delta h(x)(c + a + V_\delta(0^+)) + (1 - \delta\beta - \delta h(x))V_\delta(x + \delta) + o(\delta^2). \tag{1.15}$$

And by moving one $V_\delta(x + \delta)$ to the left and dividing by $-\delta$, we get

$$\begin{aligned} \frac{d}{dx} V_0(x) &= \lim_{\delta\to 0} \frac{V_\delta(x+\delta) - V_\delta(x)}{\delta} \\ &= \lim_{\delta\to 0} -h(x)(c + a + V_\delta(0^+)) + (\beta + h(x))V_\delta(x + \delta) + o(\delta) \\ &= -h(x)(c + a + V_0(0^+)) + (\beta + h(x))V_0(x). \end{aligned} \tag{1.16}$$

Where

$$V_0(x) := \lim_{\delta \to 0} V_\delta(x).$$

(Note that $V_0(0^+) = V_0(0) - c - a$). This differential equation seems counterintuitive since for small β, $V_0(x)$ would be decreasing as $V_0(x) < c + e^{-\beta\delta}V_0(0^+) < c + a + V_0(x)$. We will try to solve this O.D.E. anyway. We use the method of the integrating factor. Our integrating factor will be

$$e^{\int_0^x (-\beta - h(q))dq} = e^{-\beta x - H(x)}.$$

Where $H(x)$ is the cumulative hazard function. We get

$$V_0(x) = e^{\beta x + H(x)}[C + \int_0^x e^{-\beta q - H(q)}(-h(q)(c + a + V_0(0^+)))dq]$$

$$= \frac{e^{\beta x}}{1 - F(x)}[C - (c + a + V_0(0^+)) \int_0^x e^{-\beta q} f(q)dq].$$

Using the identities $e^{H(x)} = (e^{-H(x)})^{-1} = \frac{1}{1-F(x)}$ and $h(x)e^{-H(x)} = f(x)$. The C is an integrating constant and since $\lim_{x \to 0} V_0(x) = V_0(0^+)$ should hold, we find $C = V_0(0^+)$. We can rewrite the expression to

$$V_0(x) = \frac{e^{\beta x}}{1 - F(x)}[V_0(0^+) - (c + a + V_0(0^+))\mathbb{P}(Q < x)\mathbb{E}[e^{-\beta Q}|Q < x]].$$

Concluding

$$\begin{aligned} V_0(x) = \min\{&c + V_0(0^+), \\ &\frac{e^{\beta x}}{1 - F(x)}[V_0(0^+) - (c + a + V_0(0^+))\mathbb{P}(Q < x)\mathbb{E}[e^{-\beta Q}|Q < x]]\} \end{aligned} \tag{1.17}$$

and preventive maintenance is chosen if and only if $V_0(x) = c + V_0(0^+)$. However, the value of $V_0(0^+)$ depends on the policy that is chosen and it seems difficult to solve $V_0(x) = c + V_0(0^+)$ analytically for x. In the rest of this text we will write $V(x)$ instead of V_0 to not clutter the notation too much. Let μ be the smallest positive x that satisfies $V(x) = c + V(0^+)$ if such x exist and $\mu = \infty$ else. The policy that we just derived, schedules preventive maintenance at time μ if the machine has not already failed by then. We denote the total discounted cost of this policy by $V(0^+, \mu)$. Distinguising these two cases (machine survives until μ and machine breaks before μ), we get the following expression for $V(0^+, \mu)$

$$V(0^+, \mu) = \mathbb{P}(Q > \mu)e^{-\beta\mu}(c + V(0^+, \mu)) + \mathbb{P}(Q \le \mu)\mathbb{E}[e^{-\beta Q}|Q \le \mu](c + a + V(0^+, \mu)). \tag{1.18}$$

While for any $0 < x < \mu$ we get a similar expression for $V(x)$

$$\begin{aligned} V(x, \mu) = &\mathbb{P}(Q > \mu|Q > x)e^{-\beta(\mu - x)}(c + V(0^+, \mu)) \\ &+ \mathbb{P}(Q \le \mu|x < Q)\mathbb{E}[e^{-\beta(Q - x)}|x < Q \le \mu](c + a + V(0^+, \mu)), \end{aligned} \tag{1.19}$$

which also adheres to (1.17).

Example 1.2.1. Let $Q \sim \text{Exp}(\lambda)$. Because of the memoryless property, we would expect $V(x)$ to be constant. Filling this in into (1.16), we get

$$\frac{d}{dx}V(x) = 0 = -\lambda(c+a+V(0^+))+(\beta+\lambda)V(x) = -\lambda(c+a+V(0^+))+(\beta+\lambda)V(0^+).$$

Which results in

$$V(0^+) = \frac{\lambda}{\beta}(c+a)$$

which equals exactly the total discounted cost for control limit ∞.

Instead of minimizing $V(0^+, \mu)$ for μ using the Bellman equations, we can also minimize it by looking for extreme values of this function. We take its derivative to μ

$$\begin{aligned}\frac{d}{d\mu}V(0^+,\mu) &= f(\mu)e^{-\beta\mu}(c+V(0^+,\mu)) - \beta\mathbb{P}(Q>\mu)e^{-\beta\mu}(c+V(0^+,\mu)) \\ &+ \frac{d}{d\mu}V(0^+,\mu)\mathbb{P}(Q>\mu|Q>x)e^{-\beta(\mu-x)} \\ &+ f(\mu)e^{-\beta\mu}(c+a+V(0^+,\mu)) \\ &+ \frac{d}{d\mu}V(0^+,\mu)\mathbb{P}(Q\le\mu|x<Q)\mathbb{E}[e^{-\beta(Q-x)}|x<Q\le\mu].\end{aligned} \tag{1.20}$$

We are interested in the zeroes of this derivative:

$$0 = -\beta\mathbb{P}(Q>\mu)e^{-\beta\mu}(c+V(0^+,\mu)) + af(\mu)e^{-\beta\mu}.$$

Which results in

$$\frac{f(\mu)}{\mathbb{P}(Q>\mu)} = h(\mu) = \beta\frac{c+V(0^+,\mu)}{a} \tag{1.21}$$

(1.18) can also attain its minimum at $\mu = \infty$. Note however, that $\mu = 0$ results in replacing the machine infinitely often each instant and would hence result in infinite cost so that the minimum is not attained there.

So currently, we know about the optimal control limit μ that if $\mu < \infty$, it holds that $h(\mu) = \beta\frac{c+V(0^+,\mu)}{a}$ and from the Bellman equation approach, we know that $c + V(0^+,\mu) \le V(\mu)$. It can be shown that the second implies the first. For this, we will briefly return to discretized time: If the control limit equals μ, then in the time step before μ, $c+V(0^+,\mu) \ge V(\mu-\delta)$. Using (1.15), we get

$$c+V_\delta(0^+) \ge V_\delta(\mu-\delta) = \delta h(\mu)(c+a+V_\delta(0^+))+(1-\delta\beta-\delta h(\mu))V_\delta(\mu+\delta)+o(\delta^2).$$

Since we repair in the next time step, we can write

$$c+V_\delta(0^+) \ge \delta h(\mu)(c+a+V_\delta(0^+)) + (1-\delta\beta-\delta h(\mu))(c+V_\delta(0^+)) + o(\delta^2).$$

Which simplifies to

$$0 \ge ah(\mu) - \beta(c+V_\delta(0^+)) + o(\delta^2)$$

and can be rewritten as

$$h(\mu) \leq \beta \frac{c + V_\delta(0^+)}{a} + o(\delta^2) \to \beta \frac{c + V_\delta(0^+)}{a}.$$

If we now look at control limit μ, the Bellman equations yield

$$c+V_\delta(0^+) \leq V_\delta(\mu-\delta) = \delta h(\mu)(c+a+V_\delta(0^+))+(1-\delta\beta-\delta h(\mu))V_\delta(\mu+\delta)+o(\delta^2)$$

And using the same steps, we get

$$h(\mu) \geq \beta \frac{c + V_\delta(0^+)}{a} + o(\delta^2) \to \beta \frac{c + V_\delta(0^+)}{a}$$

such that the result is proven when δ approaches zero. From the above, it also follows that the hazard rate is increasing at the control limit. Hence, if the hazard rate is monotonously decreasing, preventive repair will never be chosen.

1.3 Value iteration

(1.21) can also be challenging to solve analytically. Hence we will attempt a variation on value iteration. Consider a one state, continuous time, continuous decision space and infinite horizon problem where at each (discrete) decision time, a control limit is chosen. Let $\mu_i \in \mathbb{R}^+ \cup \{\infty\}$ be the i'th control limit. The decision times have stochastic time intervals, dependent on the chosen control limit. The i'th time interval equals $I_i = \min\{Q_i, \mu_i\}$, where Q_i is the i'th lifetime. The i'th decision time is hence $T_i = \sum_{j=0}^{i-1} I_j$. The cost that is incurred when a control μ_i is chosen equals

$$g(i, \mu_i) := \mathbb{P}(Q > \mu_i)e^{-\beta\mu_i}c + \mathbb{P}(Q_i \leq \mu_i)\mathbb{E}[e^{-\beta Q_i}|Q_i \leq \mu_i](c + a).$$

So that the total discounted cost for policy $\pi = \{\mu_0, \mu_1, ...\}$ equals

$$V_\pi = \sum_{i=0}^{\infty} e^{-\beta T_i} g(i, \mu_i). \tag{1.22}$$

Assuming a stationary policy, (a variant of) the Bellman equations can be constructed

$$V_{\mu^*} = \inf_{\mu>0} \mathbb{E}[\mathbb{P}(Q > \mu)e^{-\beta\mu}(c + V_{\mu^*}) + \mathbb{P}(Q \leq \mu)\mathbb{E}[e^{-\beta Q}|Q \leq \mu](c + a + V_{\mu^*})]. \tag{1.23}$$

This differs from the Bellman equations since it lacks the additive structure as the interdecision times depend on the chosen actions.

We can now attempt to employ a variant of value iteration on it. We iteratively compute

$$V_{n+1} = \inf_{\mu_{n+1}>0} \mathbb{E}[\mathbb{P}(Q > \mu_{n+1})e^{-\beta\mu_{n+1}}(c+V_n)+\mathbb{P}(Q \leq \mu_{n+1})\mathbb{E}[e^{-\beta Q}|Q \leq \mu_{n+1}](c+a+V_n)]$$

and μ_{n+1} is the control limit at which this infimum is achieved (or ∞). Similarly as (1.21), for Q with continuously differentiable cumulative distribution function, this infimum can either be attained at ∞ or at some μ where

$$h(\mu) = \beta \frac{c + V_n}{a}$$

holds. Since (1.23) is not a Bellman equation, the convergence of this value iteration is not yet proven. When the iteration is started from an initial value V_0 which equals the total discounted cost of some control limit μ_0, then

$$\begin{aligned} V_1 &= \inf_{\mu_1>0} \mathbb{E}[\mathbb{P}(Q > \mu_1)e^{-\beta\mu_1}(c + V_0) + \mathbb{P}(Q \leq \mu_1)\mathbb{E}[e^{-\beta Q}|Q \leq \mu_1](c + a + V_0)] \\ &\leq \mathbb{E}[\mathbb{P}(Q > \mu_0)e^{-\beta\mu_0}(c + V_0) + \mathbb{P}(Q \leq \mu_0)\mathbb{E}[e^{-\beta Q}|Q \leq \mu_0](c + a + V_0)] = V_0. \end{aligned} \tag{1.24}$$

And by induction

$$\begin{aligned} V_{n+1} &= \inf_{\mu_{n+1}>0} \mathbb{E}[\mathbb{P}(Q > \mu_{n+1})e^{-\beta\mu_{n+1}}(c + V_n) + \mathbb{P}(Q \leq \mu_{n+1})\mathbb{E}[e^{-\beta Q}|Q \leq \mu_{n+1}](c + a + V_n)] \\ &\leq \mathbb{E}[\mathbb{P}(Q > \mu_n)e^{-\beta\mu_n}(c + V_n) + \mathbb{P}(Q \leq \mu_n)\mathbb{E}[e^{-\beta Q}|Q \leq \mu_n](c + a + V_n)] = V_n. \end{aligned} \tag{1.25}$$

Such that this value iteration produces a monotonously decreasing sequence $V_0 \geq V_1 \geq ....$

1.3.1 Remarks

In many cases, it may be easier just to directly minimize (1.18), as this method requires finding an intersection of the hazard rate function in each iteration. Furthermore, there may be multiple intersections of this hazard rate and these may need to be compared which requires determining $\mathbb{P}(Q \leq \mu_n)\mathbb{E}[e^{-\beta Q}|Q \leq \mu_n]$ which may be difficult. However, in cases where the hazard rate function is monotonously increasing (such as the Weibull distribution), this approach may be easier.