STA 623 HW9

Lingyun Shao

2018/12/06

Homework Assignment 9

Part I

The assumption of the probability 0.5 may not be accurate. Instead, we now assume that the probability of the machine being repaired is now θ, where $\theta \sim beta(2,3)$.

Additionally, we have data from the previous month that in the last $n = 8$ malfunctions, the number of days the machine could not be repaired are

$$D = (x_1, ..., x_8) = (2, 0, 0, 0, 4, 0, 1, 1)$$

1. Given θ, what is the distribution of the number of days the machine remains down before being fully operational? What is the likelihood of the data?

Given θ, the number of days the machine remains down before being fully repaired is like the number of failures we have before having an success in a Bernoulli experiment, so the distribution is Geometric distribution with parameter θ.

We can regard each malfunction as independent, so the likelihood would be

$$L(X|\theta) = \prod_{i=1}^{8} L(x_i|\theta) = \prod_{i=1}^{8} (1-\theta)^{x_i}\theta = (1-\theta)^8 \theta^8$$

2. What is the posterior distribution of the probability θ?

$$
\begin{aligned}
p(\theta|D) &\propto L(X|\theta)p(\theta) \\
&\propto (1-\theta)^8 \theta^8 \theta^{2-1}(1-\theta)^{3-1} \\
&= \theta^9 (1-\theta)^{10}
\end{aligned}
$$

From the kernel, we know the posterior is $Beta(10, 11)$.

3. Obtain the optimal speed a now, by using a Monte-Carlo approximation.

From the hint, we know

$$
\begin{aligned}
&E(R_a(s_{t-1}, s_t)|D) \\
&= \int_\Theta E(R_a(s_{t-1}, s_t)|\theta)p(\theta|D)d\theta \\
&\approx \sum_{t=1}^{T} E(R_a(s_{t-1}, s_t)|\theta_t)
\end{aligned}
$$

where $\theta_t's$ are samples from the posterior of θ.

Referring to the code in `Machine.R`, we can now get the optimal a as the following.

```r
# Problem parameters
M <- 1   # Idle-repair cost / day
gamma <- 0.8   # Discount factor

# Transition matrix
P <- function(a, theta) {
    matrix(c(a, 1 - a, theta, 1 - theta), 2, 2)
}

# Daily cost of going from state i to state o under decision a.
R <- function(i, o, a) {
    if (i == 1 & o == 1) {
        a
    } else if (i == 1 & o == 2) {
        M
    } else if (i == 2 & o == 2) {
        M
    } else {
        0
    }
}

# Set some values of 'a', and an array of expected costs
a_values <- seq(0.05, 0.95, 0.05)
n_sam = 10000
theta_sam = rbeta(n_sam, 10, 11)
c_values <- array(0, length(a_values))

# This loop evaluates each 'a', it may take a while.  Using N chains of
# length TIME
N <- 2000
TIME <- 30

# There exists an analytical approximation by using the long-term or
# stationary distribution.
LR <- function(a, theta) {
    lr = array(0, length(theta))
    for (i in 1:length(theta)) {
        pi <- c(theta[i], 1 - a)/(theta[i] + 1 - a)
        lr[i] = R(1, 1, a) * pi[1] * a + R(1, 2, a) * pi[1] * (1 - a) + R(2,
            1, a) * pi[2] * theta[i] + R(2, 2, a) * pi[2] * (1 - theta[i])
    }
    return(mean(lr))
}

# The approximation values
app.c_values <- (gamma - gamma^(TIME + 1))/(1 - gamma) * sapply(a_values, function(a) LR(a,
    theta = theta_sam))

matplot(a_values, app.c_values, lty = 2, type = "l", xlab = "Speed (a)", ylab = "Expected cost")
```
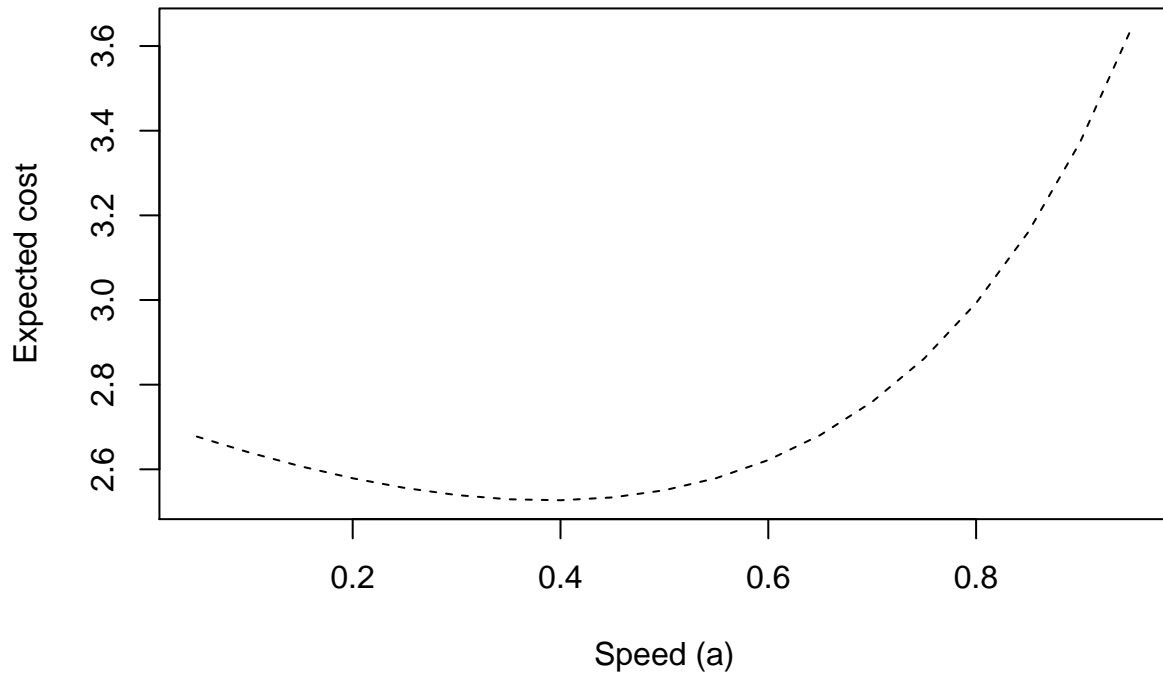
```r
optimize(LR, theta = theta_sam, lower = 0, upper = 1)
```

```
## $minimum
## [1] 0.3885463
##
## $objective
## [1] 0.6324667
```

We can sample θ from its posterior and use Mo nte Carlo Approximation and redefine the functions. Using the analytical approximation by stationary distribution, we can get the optimal results as above.

By introducing some variation in θ, we can see the optimal minimun is slightly bigger than using a fixed $\theta = 0.5$

Part II

Let's focus on the transition matrix Λ_0. Suppose each cell $\lambda_{0,ij}$ has values,

$$\lambda_{0,ij}(\theta) = \binom{j}{i}(1 - \theta)^i \theta^{j-i}$$

for all $i \leq j$, and 0 elsewhere. This is interpreted as the number of sold items each day are independent and are Binomial(j, θ).

4. Write the likelihood, discarding the first step. Assuming a uniform prior for θ, what is the form of the posterior distribution? How would you sample $\theta^{(q)}$ in order to construct draws $\Lambda_0^{(q)}$

3

The likelihood can be written as the following

$$L(D|\theta) = \prod_{j=0}^{5} \prod_{i=0}^{5} \lambda_{0,ij}^{n_{ij}}$$

Since the prior for θ is uniform, we know the posterior has the form

$$p(\theta|D) \propto L(D|\theta)p(\theta)$$
$$= \prod_{j=0}^{5} \prod_{i=0}^{5} \lambda_{0,ij}^{n_{ji}}$$
$$\propto \prod_{j=0}^{5} \prod_{i=0}^{j} [(1-\theta)^i \theta^{j-i}]^{n_{ji}}$$

where n_{ji} is the total number from state j to i in data.

From the form, we know the posterior for θ is $Beta(\sum_{j=0}^{5} \sum_{i=0}^{j} i n_{ji}, \sum_{j=0}^{5} \sum_{i=0}^{j} (j-i) n_{ji})$.

To construct $\Lambda_0^{(q)}$, we can sample $\theta^{(q)}$ from its posterior and use the posterior samples to get $\lambda_{0,ij}^{(q)}$, thus constructing a $\Lambda_0^{(q)}$ draw.

5. From Λ_0, what is a reasonable way to construct Λ_1 and Λ_2?

We can construct Λ_1 in a similar way. Let

$$\lambda_{1,ij} = \begin{cases} \binom{j}{i-1}(1-\theta)^{i-1}\theta^{j-(i-1)} & 0 \le i-1 \le 4, 0 \le j \le 4, i \le j \\ \binom{j}{i-1}(1-\theta)^{i-1}\theta^{j-(i-1)} & 0 \le i-1 \le 3, j = 5 \\ \sum_{i=5}^{6} \binom{j}{i-1}(1-\theta)^{i-1}\theta^{j-(i-1)} & i = 5, j = 5 \\ 0 & o.w. \end{cases}$$

$$\lambda_{2,ij} = \begin{cases} \binom{j}{i-2}(1-\theta)^{i-2}\theta^{j-(i-2)} & 0 \le i-2 \le 3, 0 \le j \le 3, i \le j \\ \binom{j}{i-2}(1-\theta)^{i-2}\theta^{j-(i-2)} & 0 \le i-2 \le 2, j = 4, 5 \\ \sum_{i=5}^{6} \binom{j}{i-1}(1-\theta)^{i-1}\theta^{j-(i-1)} & i = 5, j = 4 \\ \sum_{i=5}^{7} \binom{j}{i-1}(1-\theta)^{i-1}\theta^{j-(i-1)} & i = 5, j = 5 \\ 0 & o.w. \end{cases}$$