Optimal Risk Probability for First Passage Models

— in Semi-Markov Decision Processes

Xianping Guo (Coauthor: Yonghui Huang)
(Zhongshan University, Guangzhou)

Peking University, July 2010

Outline

2 Motivation

2 Semi-Markov decision processes

2 Optimality problems

2 Optimality equation

2 Existence of optimal policies

2 Numerable examples

1. Motivation

Background: Reliability engineering, and risk analysis

Problem: $\sup_{\pi} P_i^{\pi}(\tau_B > \lambda)$,

- i an initial state
- π is a policy
- B is a given target set
- τ_B is a first passage time to B
- λ is a threshold value.

2. Semi-Markov Decision Processes

The model of SMDP:

$$\{S, B, (A(i), i \in S), Q(t, j|i, a)\}$$

where

- S : the state space, a denumerable set;
- B: a given target set, a subset of S;
- $A(i)$: finite set of actions available at $i \in S$;
- $Q(t, j|i, a)$: semi-Markov kernel, $a \in A(i), i, j \in S$;

Notation:

- **Policy π**: A sequence $\pi = \{\pi_n, n = 0, 1, \ldots\}$ of stochastic kernels π_n on the action space A given H_n satisfying

$$\pi_n(A(i_n) | (0, i_0, \lambda_0, a_0, \ldots, t_{n-1}, i_{n-1}, \lambda_{n-1}, a_{n-1}, t_n, i_n)) = 1$$

- **Stationary policy**: measurable f, $f(i, \lambda) \in A(i)$ for all (i, λ)
- $P^{\pi}_{(i,\lambda)}$: Probability measure on $(S \times [0, \infty) \times (\cup_{i \in S} A(i)))^{\infty}$
- S_n, J_n, A_n: n-th decision epoch, the state and action at the S_n, respectively.

Assumption A. There exist $\delta > 0$ and $\varepsilon > 0$ such that

$$\sum_{j \in S} Q(\delta, j \mid i, a) \leq 1 - \varepsilon, \text{ for all } (i, a) \in K.$$

Assumption A $\Rightarrow$ $P^{\pi}_{(i,\lambda)}(\{S_\infty = \infty\}) = 1$

Semi-Markov decision process $\{(Z(t), A(t), t \geq 0\}$:

$$Z(t) = J_n, A(t) = A_n, \quad \text{for } \; S_n \leq t < S_{n+1}$$

The first passage time into B, is defined by

$$\tau_B := \inf\{t \geq 0 \mid Z(t) \in B\}, \quad (\text{with } \inf \emptyset := \infty),$$

3. Optimality Problems

The risk probability:

$$F^{\pi}(i,\lambda) := P^{\pi}_{(i,\lambda)}(\tau_B \le \lambda)$$

The optimal value:

$$F_*(i,\lambda) := \inf_{\pi \in \Pi} F^{\pi}(i,\lambda),$$

Definition 1. A policy $\pi^* \in \Pi$ is called optimal if

$$F^{\pi^*}(i,\lambda) = F_*(i,\lambda) \quad \forall\, (i,\lambda) \in S \times R.$$

- Existence and computation of optimal policies ???

4. Optimality Equation

For $i \in B^c; a \in A(i)$, and $\lambda \geq 0$, let

$$T^a u(i;\lambda) := Q(\lambda; B|i;a) + \sum_{j \in B^c} \int_0^{\lambda} Q(dt; j|i;a) u(j; \lambda - t);$$

with $u \in F_{[0;1]}$ (the set of measurable functions $0 \leq u \leq 1$),

$$Q(\lambda; B|i;a) := \sum_{j \in B} Q(\lambda; j|i;a); \quad T^a u(i;\lambda) := 0 \text{ for } \lambda < 0.$$

Then, define operators T and T^f:

$$Tu(i;\lambda) := \min_{a \in A(i)} T^a u(i;\lambda); \quad T^f u(i;\lambda) := T^{f(i;\lambda)} u(i;\lambda);$$

for each stationary policy f.

Theorem 1. Let Under Assumption A, we have

(a) $F^f = \lim_{n \to \infty} u_n^f$, where $u_n^f := T^f u_{n-1}; u_{-1}^f := 1$;

(b) F^f satisfied the equation, $u = T^f u$, for all $f \in F$;

- Theorem 1 gives an approximation of risk probability F^f.

For each $(i; \lambda) \in B^c \times R_+$ and $\pi \in \Pi$, let

$$F_{-1}^{\pi}(i; \lambda) := 1;$$

$$F_n^{\pi}(i; \lambda) := 1 - \sum_{m=0}^{n} P_{(i;\lambda)}^{\pi}(S_m \le \lambda < S_{m+1}; J_k \in B^c; 0 \le k \le m)$$

Theorem 2. Let $F_n^*(i;\lambda) := \inf_\pi F_n^\pi(i;\lambda)$, then

(a) $F_{n+1}^* = TF_n^*$ for all $n \geq -1$, and $\lim_{n\to\infty} F_n^* = F_*$.

(b) F_* satisfies the optimality equation: $F_* = TF_*$.

(c) F_* is the maximal fixed point of T in $\mathcal{F}_{[0,\infty]}$.

Remark 1.

- Theorem 2(a) gives a value iteration algorithm for computing the optimal value function F_*.
- Theorem 2(b) establishes the optimality equation.

5. Existence of Optimality Policise

To ensure the existence of optimal policies, we introduce the following condition.

Assumption B. For every $(i; ¸)$ 2 B^c £ R and f,

$$P^{f}_{(i;¸)}(¿_B < 1) = 1:$$

To verify Assumption B, we have a fact below:

Theorem 3. If there exists a constant ® > 0 such that

$$\sum_{j2B} Q(1; jji; a) \; ¸ \; ® \quad \text{for all } i \; 2 \; B^c; a 2 A(i)$$

Theorem 4. Under Assumptions A and B, we have

(a) F^f and $F_{¤}$ are the unique solution in $F_{[0;1]}$ to equations $u = T^f u$ and $u = Tu$, respectively;

(b) any f, such that $F_{¤} = T^f F_{¤}$, is optimal;

(c) there exists a stationary policy $f^{¤}$ satisfying the optimality equation: $F_{¤} = TF_{¤} = T^{f^{¤}} F_{¤};$ and such policy $f^{¤}$ is optimal.

Remark 2.

[2] Theorem 4(c) shows the existence of an optimal poliy.

To give the existence of special optimal policies, let

$$A^*(i,\lambda) := \{a \in A(i) \mid F^*(i,\lambda) = T^a F^*(i,\lambda)\}.$$

$$A^*(i) := \bigcap_{\lambda \geq 0} A^*(i,\lambda)$$

Theorem 5. If $\sup_{i \in B^c} \sup_{a \in A(i)} Q(t, B^c \mid i, a) < 1$ for some $t > 0$, and Assumptions A and B hold, then,

(a) for any $g \in G := \{g \mid g(i) \in A(i) \forall i \in S\}$, F^g is the unique solution in $F_{[0,1]}$ to the equation: $u = T^g u$;

(b) there exists an optimal policy $f \in G$ if and only if $A^*(i) \neq \emptyset$ for all $i \in B^c$.

5. Numerable examples

Example 5.1. Let $S = \{1; 2; 3\}$, $B = \{3\}$, where

- state 1: the good state
- state 2: the medium state
- state 3: the failure state

Let $A(1) = \{a_{11}; a_{12}\}; A(2) = \{a_{21}; a_{22}\}; A(3) = \{a_{31}\}$.

The semi-Markov kernel is of the form:

$$Q(t; j \mid i; a) = H(t \mid i; a)p(j \mid i; a)$$

- $H(t \mid i; a)$: the distribution functions of the sojourn time
- $p(j \mid i; a)$: the transition probabilities.

$$H(t \mid 1; a_{11}) := \begin{cases} 1/25; & t \in [0; 25]; \\ 1; & t > 25; \end{cases}$$

$$H(t \mid 2; a_{21}) := \begin{cases} 1/20; & t \in [0; 20]; \\ 1; & t > 20; \end{cases}$$

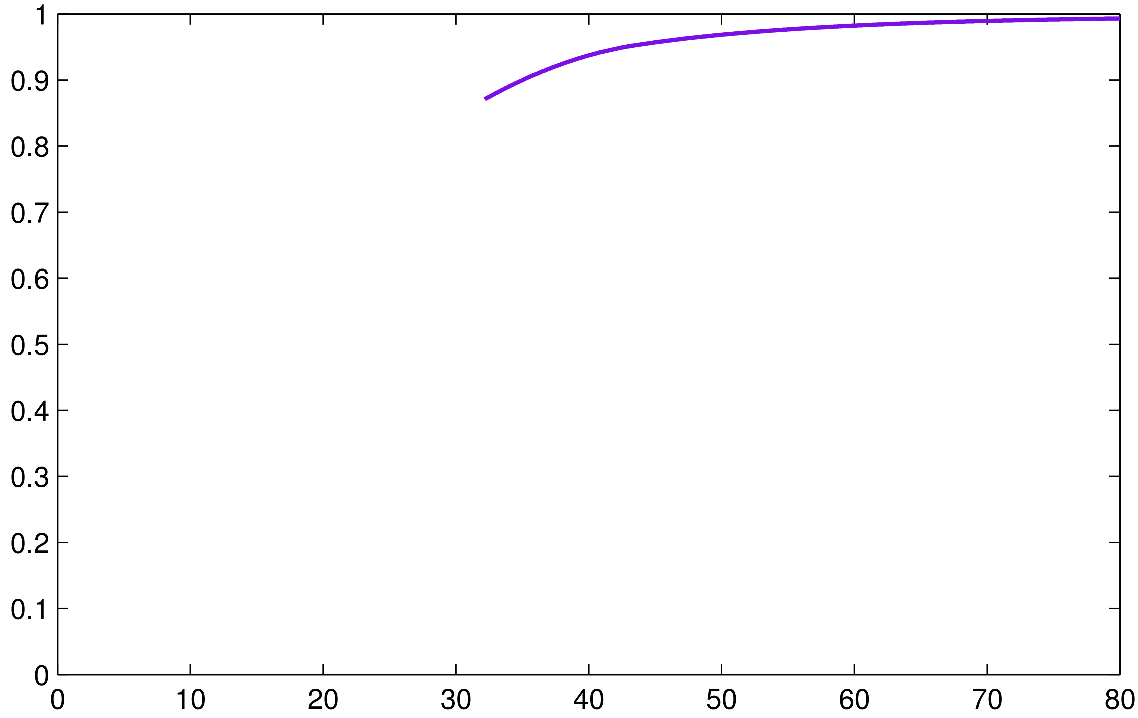
$$H(t \mid 3; a_{31}) := 1 - e^{-0.2t}.$$

$$H(t \mid 1; a_{12}) = 1 - e^{-0.08t};$$

$$H(t \mid 2; a_{22}) = 1 - e^{-0.15t};$$

$$p(1 \mid 1, a_{11}) = 0, \; p(2 \mid 1, a_{11}) = \frac{9}{20}, \; p(3 \mid 1, a_{11}) = \frac{11}{20},$$
$$p(1 \mid 1, a_{12}) = 0, \; p(2 \mid 1, a_{12}) = \frac{1}{2}, \; p(3 \mid 1, a_{12}) = \frac{1}{2},$$
$$p(1 \mid 2, a_{21}) = \frac{1}{5}, \; p(2 \mid 2, a_{21}) = 0, \; p(3 \mid 2, a_{21}) = \frac{4}{5},$$
$$p(1 \mid 2, a_{22}) = \frac{1}{4}, \; p(2 \mid 2, a_{22}) = 0, \; p(3 \mid 2, a_{22}) = \frac{3}{4},$$
$$p(3 \mid 3, a_{31}) = 1.$$

Using the value iteration algorithm in Theorem 2, we obtain some computational results as in Figure 1 and Figure 2.



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
80

Figure 1. The functions $T^{a}F^{¤}(i; ¸)$

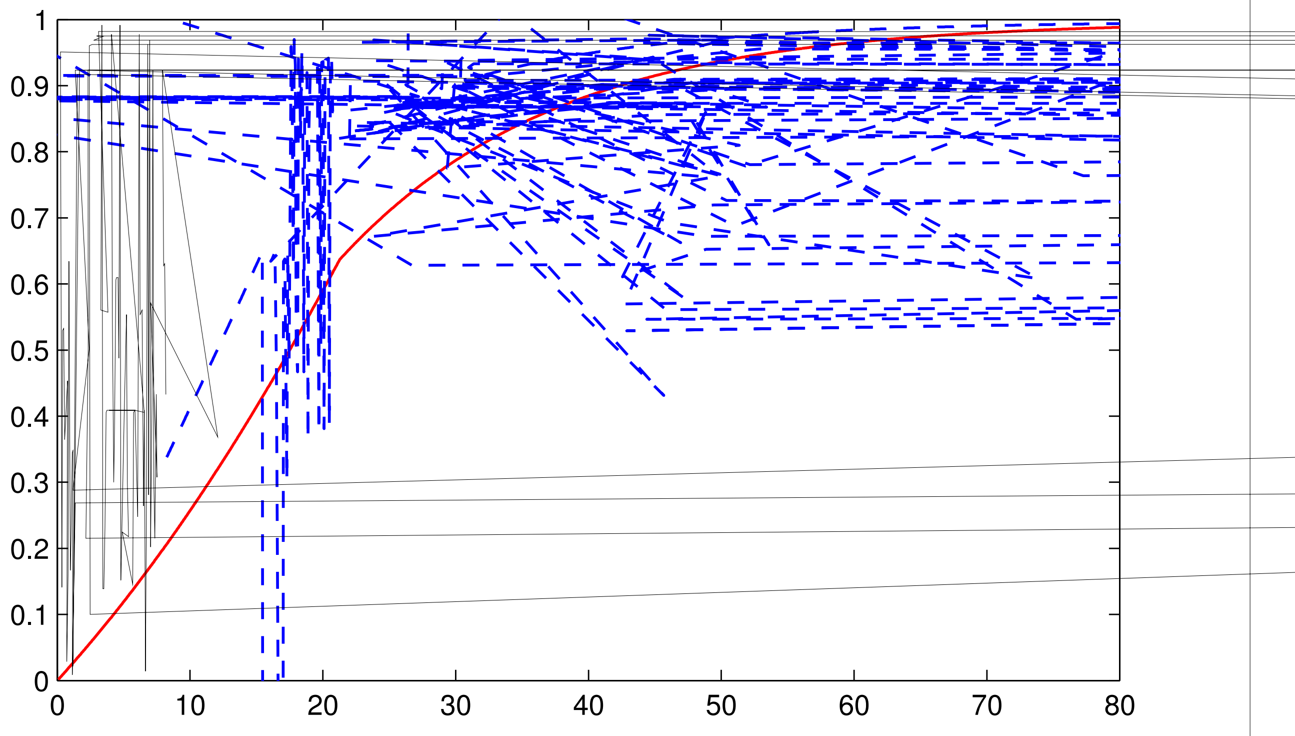


Figure 2. The value function $F^{¤}(i; ¸)$

More clearly, we have

$$F^*(1;\theta) = \begin{cases} T^{a_{11}}F^*(1;\theta); & 0 \le \theta < 21.36; \\ T^{a_{11}}F^*(1;\theta) = T^{a_{12}}F^*(1;\theta); & \theta = 21.36; \\ T^{a_{12}}F^*(1;\theta); & 21.36 < \theta < 29.3; \\ T^{a_{11}}F^*(1;\theta) = T^{a_{12}}F^*(1;\theta); & \theta = 29.3; \\ T^{a_{11}}F^*(1;\theta)(= 0.7742); & \theta > 29.3; \end{cases}$$

$$F^*(2;\theta) = \begin{cases} T^{a_{21}}F^*(2;\theta); & 0 \le \theta < 18.54; \\ T^{a_{21}}F^*(2;\theta) = T^{a_{22}}F^*(2;\theta); & \theta = 18.54; \\ T^{a_{22}}F^*(2;\theta); & 18.54 < \theta < 23.82; \\ T^{a_{21}}F^*(2;\theta) = T^{a_{22}}F^*(2;\theta); & \theta = 23.82; \\ T^{a_{21}}F^*(2;\theta)(= 0.8542); & \theta > 23.82. \end{cases}$$

Define a policy f^* by

$$f^*(1; \lambda) = \begin{cases} a_{11}; & 0 \le \lambda \le 21.36; \\ a_{12}; & 21.36 < \lambda \le 29.3; \\ a_{11}; & \lambda > 29.3; \end{cases}$$

$$f^*(2; \lambda) = \begin{cases} a_{21}; & 0 \le \lambda \le 18.54; \\ a_{22}; & 18.54 < \lambda \le 23.82; \\ a_{21}; & \lambda > 23.82; \end{cases}$$

Then, we have

- $F^*(i; \lambda) = T^{f^*} F^*(i; \lambda)$ for $i = 1, 2$ and all $\lambda \ge 0$,
- f^* is an optimal stationary policy.

$$A^*(1,\lambda) = \begin{cases} \{a_{11}\}, & 0 \le \lambda < 21.36, \\ \{a_{11}, a_{12}\}, & \lambda = 21.36, \\ \{a_{12}\}, & 21.36 < \lambda < 29.3, \\ \{a_{11}, a_{12}\}, & \lambda = 29.3, \\ \{a_{11}\}, & \lambda > 29.3, \end{cases}$$

$$A^*(2,\lambda) = \begin{cases} \{a_{21}\}, & 0 \le \lambda < 18.54, \\ \{a_{21}, a_{22}\}, & \lambda = 18.54, \\ \{a_{22}\}, & 18.54 < \lambda < 23.82, \\ \{a_{21}, a_{22}\}, & \lambda = 23.82, \\ \{a_{21}\}, & \lambda > 23.82, \end{cases}$$

Hence,

$$A^{¤}(1) = \bigcap_{¸\ ¸\ 0} A^{¤}(1; ¸) = ;; A^{¤}(2) = \bigcap_{¸\ ¸\ 0} A^{¤}(2; ¸) = ;;$$

which show there is no optimal policy in G.

Remark 3. This shows that the assumption in the previous literature is not satis¯ed for this example !!!

Example 5.2. Let $S = \{1, 2\}$, $B = \{2\}$,

$A(1) = \{a_{11}, a_{12}\}, A(2) = \{a_{21}\}$,

$Q(t, j \mid i, a)$ is given by

$$Q(t, j \mid 1, a_{11}) = \begin{cases} 1/2, & \text{if } t \geq 1, j = 1, 2, \\ 0, & \text{otherwise}, \end{cases}$$

$$Q(t, j \mid 1, a_{12}) = \begin{cases} 1, & \text{if } t \geq 2, j = 2, \\ 0, & \text{otherwise}, \end{cases}$$

$$Q(t, j \mid 2, a_{21}) = \begin{cases} 1 - e^{-t}, & \text{if } t \geq 0, j = 2, \\ 0, & \text{otherwise}. \end{cases}$$

Assumptions A and B holds in this example.

We now define a policy d as follows:

$$d(1;\theta) = \begin{cases} a_{12}; & 0 \le \theta \le 2; \\ a_{11}; & \theta > 2. \end{cases}$$

Then, by Theorem 1, we have $F^d(1;\theta) = \lim_{n \to \infty} F^d_n(1;\theta)$, which yields

$$F^d(1;\theta) = \begin{cases} 0; & 0 \le \theta < 2; \\ 1; & \theta = 2; \\ 1/2; & 2 < \theta < 3. \end{cases}$$

Hence, $F^d(1;\theta)$ is not a distribution function of θ.

Many Thanks !!!