SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN-POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

STOPPING PROBLEMS AS SPECIAL STATISTICAL DECISION PROBLEMS

PETER NEUMANN

Section of Mathematics, Technical University of Dresden, Dresden, G.D.R

0. Introduction

The problem of the optimal stopping of stochastic processes is well known and often described in the literature. In [2], problems for discrete-time processes with arbitrary dependences are dealt with extensively. Shiryayev ([12]) obtained structural results for discrete and continuous-time Markov processes and applied them to the sequential statistical analysis. Recently, stopping problems in several respects have been generalized: consideration of costs which undertake a valuation of the process before stopping, optimal control and stopping [5], [10], [11], multiple optimal stopping [9] as well as stopping games [13].

In this paper, the basic stopping problem for arbitrary discrete-time processes is described as a statistical decision problem. For this purpose we refer very closely to [7], give a relevant dynamic system (DS) and formulate the corresponding Bellman optimality equation. Finally, we only briefly refer to the solution methods of this equation if the components of DS are Markovian.

1. A stopping system

As in the description of a statistical decision problem by a DS (see [7], 1) we start from the so-called *stopping system* StS: $(N, \mathfrak{S}, \mathfrak{p}, \mathfrak{g}, \mathfrak{R})$ with the following components:

– end-time point or horizon N, $N \leqslant \infty$;

– state space $\mathfrak{S}$, countable (for the purpose of coupling with [7], $\mathfrak{S} \neq \emptyset$);

– law of motion $\mathfrak{p} = (p_n)$: sequence of probability distributions

$p_n(\cdot|\mathfrak{s}_{n-1})$, $\mathfrak{s}_{n-1} = (s_0, \ldots, s_{n-1}) \in \mathfrak{S}^n$, $n = 1, 2, \ldots$ In the case of non-stopping, the transition into $s' \in \mathfrak{S}$ takes place with probability $p_n(s'|\mathfrak{s}_{n-1})$;

– reward structure $\mathfrak{g} = (g_n)$: sequence of reward functions $g_n|\mathfrak{S}^{n+1} \to \boldsymbol{R}$, $n = 0, 1, \ldots$ If it was stopped at time n, dependent on $\mathfrak{s}_n$, the reward $g_n(\mathfrak{s}_n)$ results;

– set $\mathfrak{N}$ of admissible stopping times, i.e., functions $\tau|\mathfrak{S}^{N+1} \to \{0, 1, \ldots, N\}$ such that for arbitrary $\mathfrak{s}_N = (s_0, s_1, \ldots)$ and $n = 0, 1, \ldots$ $\tau(\mathfrak{s}_N) = n$ implies $\tau(\mathfrak{s}'_N) = n$ for all $\mathfrak{s}'_N = (s'_0, s'_1, \ldots)$ with $s'_\nu = s_\nu$, $\nu = 0, \ldots, n$. Whether for a concrete state sequence $\mathfrak{s}_N$ the stopping time τ stops at time n depends only on the first n-part $\mathfrak{s}_n = (s_0, \ldots, s_n)$.

2. Description by a dynamic system

In the following, we give a possibility for fitting a stopping system to a DS from [7], 1, in a suitable manner. The purpose is to apply immediately the optimality equation given in [7], 14, to stopping problems. On the one hand, the components of DS will prove to be simple. Since, however, the process to be stopped runs a random time, but in a DS only nonrandom horizons are provided, on the other hand there will be some particularities.

Now we have a DS: $(N, \mathfrak{S}, \mathfrak{D}, p_0, \mathfrak{p}, \mathfrak{k}, \Theta)$ with

– horizon N and

– state space $\mathfrak{S}$ as above;

– decision space $\mathfrak{D} = \{0, 1\}$. Here $d = 1$ denotes stopping, $d = 0$ non-stopping;

– initial distribution p_0: There exists an initial state $s = \mathfrak{S}$ with $p_0(s) = 1$;

– law of motion $\mathfrak{p} = (p_n)$, $p_n = p_n(\cdot|h_{n-1}, d_{n-1})$, $h_{n-1} = (s_0, d_0, s_1, \ldots \ldots, s_{n-1})$, $d_{n-1} \in \mathfrak{D}$. For pairs (h_{n-1}, d_{n-1}) with $d_0 = \ldots = d_{n-1} = 0$, at which no stopping occurred, let $p_n(s'|h_{n-1}, d_{n-1}) := p_n(s'|\mathfrak{s}_{n-1})$; otherwise we put

$$p_n(s'|h_{n-1}, d_{n-1}) := \delta_{s_{n-1}, s'}. \tag{1}$$

(1) means that the DS preserves that state until the end-time which it has been reached at the stopping time;

– cost structure $\mathfrak{k} = (k_n)$, $k_n|\mathfrak{H}_n \to \boldsymbol{R}$. In particular we put

$$k_n(s_0, 0, s_1, 0, \ldots, 0, s_n) = 0, \qquad n < N, \tag{2}$$

$$\begin{aligned} &k_m(s_0, 0, \ldots, 0, s_n, 1, s_{n+1}, d_{n+1}, \ldots, d_{m-1}, s_m) \\ &\quad = k_{n+1}(s_0, 0, \ldots, 0, s_n, 1, s_{n+1}) = -g_n(s_0, \ldots, s_n), \qquad n < m \leqslant N. \end{aligned} \tag{3}$$

(2) means that before the first stopping decision $d = 1$ nothing is paid, (3) means that after the first $d = 1$ nothing more is paid. Thus, (-1) times the quantity in (3) represents a reward resulting once. Moreover, in the case N

$< \infty$ the quantity $-k_N(s_0, 0, \ldots, 0, s_N)$ is the reward paid at the end-time if the process has never been stopped before;

– set Θ of admissible strategies: One can find (not only in one manner) strategies $\vartheta = (\delta_n)$, $\delta_n | \mathfrak{H}_n \to \mathfrak{D}$, the decisions of which correspond to a stopping time τ. For the purpose of an appropriate mathematical description, starting from τ we choose that strategy $\vartheta^\tau = (\delta_n^\tau)$ which makes a stopping decision anew at all the following time-points: For $\mathfrak{s}_N$ with $\tau(\mathfrak{s}_N) = k$ and $h_n = h_n^{\vartheta^\tau}(a_n\, \mathfrak{s}_N)$ (see [7], 2) we have

$$\delta_n^\tau(h_n) = \begin{cases} 0 & \text{if } n = 0, \ldots, k-1,\ k < N, \\ 1 & \text{if } n = k,\ k+1, \ldots,\ k < N \end{cases}$$

or

$$\delta_n^\tau(h_n) = 0 \qquad \text{if } n = 0,\ 1, \ldots,\ k = N.$$

Finally we define $\Theta := \{\vartheta\colon\ \vartheta = \vartheta^\tau,\ \tau \in \mathfrak{N}\}$.

3. The process to be stopped and the optimal value function

We consider the stopping time $\bar{\tau} := N$. For the corresponding strategy $\bar{\vartheta} := \vartheta^{\bar{\tau}}$, by the theorem of Ionescu-Tulcea (see [7], 3) there exists a probability space $[\Omega, \mathfrak{F}, P_{s,\bar{\vartheta}}]$, $\Omega = \mathfrak{S}^{N+1}$, $\mathfrak{F} = (\mathfrak{P}(\mathfrak{S}))^{(N+1)}$, $P_{s,\bar{\vartheta}} =: P_s$, s being the initial state of the DS, as well as a stochastic process $X := (X_n)$, $n = 0, 1, \ldots$ over $[\Omega, \mathfrak{F}, P_s]$ describing the state of the DS. We call this process the *process to be stopped* or the *running process*.

It is worth while to note that the corresponding measure $P_{s,\vartheta}$ for every other strategy $\vartheta \in \Theta$ can be described uniquely by P_s. Furthermore, for every stopping time τ one can define by

$$Y_n(\omega) := X_{\min(\tau(\omega), n)}(\omega), \qquad n = 0, 1, \ldots$$

a stochastic process $X^\tau := (Y_n)$ over the same probability space $[\Omega, \mathfrak{F}, P_s]$. We call it the *process stopped by* τ. We denote $\mathfrak{F}_n := \sigma(\mathfrak{S}_n)$, $\mathfrak{S}_n = (X_0, \ldots, X_n)$. The stopping time τ, first defined purely algebraically, now proves to be a random variable $\tau | [\Omega, \mathfrak{F}, P_s] \to \{0, 1, \ldots, N\}$ for which $\{\omega\colon\ \tau(\omega) = n\} \in \mathfrak{F}_n$, $n = 0, 1, \ldots$ holds. Thus, τ becomes a stopping time in the sense of measure theory [1]. If we denote further $G_n := g_n(X_0, \ldots, X_n)$, $n = 0, 1, \ldots$, G_τ is defined as

$$G_\tau(\omega) := \begin{cases} G_n(\omega) & \text{on } \{\omega\colon\ \tau(\omega) = n\},\ n = 0, 1, \ldots, \\ 0 & \text{on } \{\omega\colon\ \tau(\omega) = \infty\}. \end{cases}$$

If we postulate $E_s g_n^-(X_0, \ldots, X_n) < \infty$, $n = 0, 1, \ldots$, there exists $E_s G_\tau$, eventually being infinity.

In future, we shall consider the class $\mathfrak{M}_n := \{\tau\colon\ P_s(n \leqslant \tau < \infty) = 1,\ s \in \mathfrak{S}\}$ of stopping times. Let $(\mathfrak{S}^{n+1})_s^+ := \{\mathfrak{s}_n\colon\ P_s(\mathfrak{S}_n = \mathfrak{s}_n) > 0\}$ be the set of all P_s-essential state-n-histories. (For such $\mathfrak{s}_n$, we must necessarily have $s_0 = s$.) Then, the expected reward

$$v_n^\tau(\mathfrak{s}_n) := E_s(G_\tau | \mathfrak{S}_n = \mathfrak{s}_n),$$

given $\mathfrak{s}_n \in (\mathfrak{S}^{n+1})_s^+$ caused by $\tau \in \mathfrak{M}_n$, is called the value relative to τ at stage n, $n = 0, 1, \ldots$ The function v_n^* defined by

$$v_n^*(\mathfrak{s}_n) := \sup_{\tau \in \mathfrak{M}_n} v_n^\tau(\mathfrak{s}_n), \qquad \mathfrak{s}_n \in (\mathfrak{S}^{n+1})_s^+$$

is called the *optimal value function*. Determining this function is the first object in solving a stopping problem.

We note that $v_n^*(\mathfrak{s}_n)$ is a realization of the so-called *random optimal value function*

$$V_n^* := \sup_{\tau \in \mathfrak{M}_n} E_s(G_\tau | \mathfrak{S}_n), \quad n = 0, 1, \ldots \tag{4}$$

It is an interesting fact here that although, in general, $\mathfrak{M}_n$ is a more than countable set, because of the countability assumption of $\mathfrak{S}$ V_n^* is a measurable function, i.e., a random variable on $[\Omega, \mathfrak{F}, P_s]$. Finally we remark that V_0^* is a constant.

4. Reactivity of stopping times

DEFINITION. A stopping time $\tau \in \mathfrak{M}_n$ is called *reactive* if for all $\omega \in \Omega$ with $\mathfrak{s}_m = a_m \omega \in (\mathfrak{S}^{m+1})_s^+$ the inequality $\tau(\omega) > m$ implies $E_s(G_\tau | \mathfrak{S}_m)(\omega) > G_m(\omega)$, $m = n, n+1, \ldots$

Remark. If $\tau \in \mathfrak{M}_m$ is reactive, for some $\omega \in \Omega$ the first m-part $\mathfrak{s}_m = a_m \omega$ of which is P_s-essential and for which

$$E_s g_\tau(\mathfrak{s}_m, X_{m+1}, \ldots, X_\tau) \leqslant g_m(\mathfrak{s}_m) \tag{5}$$

holds, the equality $\tau(\omega) = m$ follows. If an essential state-m-history would result a reward at least as large as the one would expect by continuing the process, a reactive stopping time stops at time m. The only important case here is that of equality of (5), in which it seems irrelevant whether we stop or continue. A reactive stopping time does not make the latter decision.

Let us denote $\mathfrak{M}_n' := \{\tau \in \mathfrak{M}_n, \tau \text{ reactive}\}$, $n = 0, 1, \ldots$ We now cite a theorem and a lemma which are due to [2].

THEOREM. *From a stopping time $\tau \in \mathfrak{M}_n$ we construct a stopping time $\varrho(\tau)$. Then we have*:

(a) $\varrho(\tau) \leqslant \tau$,

(b) $\varrho(\tau) \in \mathfrak{M}_n'$,

(c) $E_s(G_{\varrho(\tau)} | \mathfrak{S}_n) \geqslant E_s(G_\tau | \mathfrak{S}_n)$,

(d) $\varrho(\varrho(\tau)) = \varrho(\tau)$.

LEMMA. *For $\tau_1, \tau_2 \in \mathfrak{M}'_n$ the stopping time $\tau := \max(\tau_1, \tau_2)$ also belongs to $\mathfrak{M}'_n$, and we have*

$$E_s(G_\tau | \mathfrak{S}_n) \geq \max_{i=1,2} E_s(G_{\tau_i} | \mathfrak{S}_n), \qquad n = 0, 1, \ldots$$

Now we can state the following

THEOREM. *The subset $\Theta' \subset \Theta$ of strategies of the* DS *considered above, the relative stopping times of which form the set $\mathfrak{M}'_0$, has the property of completeness* (see [7], 6a), *with regard to the stopping problem formulation, $\leq$ is to be replaced here by $\geq$ and* min *by* max.

Proof. Let $n = 0, 1, \ldots, \vartheta \in \Theta'$ and $\mathfrak{s}_n \in (\mathfrak{S}^{n+1})^+(\vartheta)$ be given. For every $\vartheta_1, \vartheta_2 \in \Theta_n(\vartheta)$ we have to show the existence of a $\vartheta' \in \Theta_n(\vartheta)$ such that

$$v_{n+1,\vartheta'}(\mathfrak{s}_n, s') \geq \max\big(v_{n+1,\vartheta_1}(\mathfrak{s}_n, s'), v_{n+1,\vartheta_2}(\mathfrak{s}_n, s')\big)$$

holds for all $s' \in \mathfrak{S}$ with $p_{n+1,\vartheta}(s' | \mathfrak{s}_n) > 0$. (For all notation, see [7], 5.) Obviously we can restrict ourselves to such $\mathfrak{s}_n$ and ϑ, relative to which until time n no stopping decision is made. For the stopping times τ_i corresponding to ϑ_i, by assumption we have $\tau_i \in \mathfrak{M}'_{n+1}$, $i = 1, 2$, and, because of

$$v_{n+1,\vartheta_i}(\mathfrak{s}_{n+1}) = v^{\tau_i}_{n+1}(\mathfrak{s}_{n+1}) = E_s(G_{\tau_i} | \mathfrak{S}_{n+1} = \mathfrak{s}_{n+1}),$$

$\mathfrak{s}_{n+1} = (\mathfrak{s}_n, s')$, the conclusion follows by the lemma. ■

5. The optimality equation

THEOREM. *For the optimal expected reward V^*_n* (4) *the Bellman optimality equation*

$$V^*_n = \max\big(G_n, E_s(V^*_{n+1} | \mathfrak{S}_n)\big) \tag{6}$$

holds; for every $\mathfrak{s}_n \in (\mathfrak{S}^{n+1})^+_s$ therefore we have

$$v^*_n(\mathfrak{s}_n) = \max\Big(g_n(\mathfrak{s}_n), \sum_{s' \in \mathfrak{S}} p_{n+1}(s' | \mathfrak{s}_n) v^*_{n+1}(\mathfrak{s}_n, s')\Big) \tag{6'}$$

where $n = 0, 1, \ldots$

Moreover, in the case $N < \infty$ we have

$$V^*_N = G_N; \tag{7}$$

for every $\mathfrak{s}_N \in (\mathfrak{S}^{N+1})^+_s$ therefore

$$v^*_N(\mathfrak{s}_N) = g_N(\mathfrak{s}_N). \tag{7'}$$

Proof. Without loss of generality, in considering (4) we can restrict ourselves to $\mathfrak{M}'_n \subset \mathfrak{M}_n$. Furthermore, we have seen that Θ' is complete, and because of

$$\big(E_s(G_\tau | \mathfrak{S}_n)\big)^- \leq E_s(G^-_\tau | \mathfrak{S}_n) < \infty, \qquad \tau \in \mathfrak{M}'_n,$$

a condition analogous to condition (V) of [7], 14, holds. Then, by [7], 14, we have the optimality equation

$$v^*_{\vartheta|n}(\mathfrak{s}_n) = \sup_{\eta \in \Theta_{n-1}(\vartheta)} \sum_{s' \in \mathfrak{S}} p_{n+1,\eta}(s'|\mathfrak{s}_n)\big(c_{n+1,\eta}(\mathfrak{s}_n, s') + v^*_{\eta|n+1}(\mathfrak{s}_n, s')\big)$$

for every $\vartheta \in \Theta$ for which $\Theta_{n-1}(\vartheta) = \Theta_{n-1}(\bar{\vartheta})$ holds. However, the decision function δ_n of every $\eta \in \Theta_{n-1}(\vartheta)$ has the values $d = 1$ (stopping) and $d = 0$ (non stopping) only, and $\sup\limits_{\eta \in \Theta_{n-1}(\vartheta)}$ reduces to a maximization over two terms. Finally, for $d = 1$ we have

$$p_{n+1,\eta}(s'|\mathfrak{s}_n) = \delta_{s_n, s'}, \qquad c_{n+1,\eta}(\mathfrak{s}_n, s') = g_n(\mathfrak{s}_n) \qquad \text{and} \qquad v^*_{\eta|n+1}(\mathfrak{s}_n, s') = 0,$$

whereas for $d = 0$ we have $c_{n+1,\eta}(\mathfrak{s}_n, s') = 0$, so that (6), (6′) follow. (7), (7′) are trivial. ∎

A numerical calculation of the optimal value $v_0^*(s)$ of a stopping problem by means of the optimality equation succeeds – if at all – at most for $N < \infty$. In this case Bellman's well-known backward induction technique is applicable. Under additional assumptions the optimal value function of an infinite horizon stopping problem is approximated by a sequence of optimal value functions of finite ones. A greater chance for the practical solvability of a stopping problem, however, can be found at most in the so-called Markov case.

6. The Markov case

We shall say that we have the *Markov case for the stopping system StS* if

(a) $p_n(\cdot|\mathfrak{s}_{n-1})$ for all $\mathfrak{s}_{n-1} = (s_0, \ldots, s_{n-1}) \in (\mathfrak{S}^n)_s^+$ depends on s_{n-1} only,

(b) $g_n(\mathfrak{s}_n)$ for all $\mathfrak{s}_n = (s_0, \ldots, s_n) \in (\mathfrak{S}^{n+1})_s^+$ depends on s_n only, (we shall write shortly $g_n(s_n)$),

(c) every stopping time $\tau \in \mathfrak{N}$ has the property that $\tau(\mathfrak{s}_N) = n$ implies $\tau(\mathfrak{s}'_N) = n$ for all $\mathfrak{s}'_N = (s'_0, s'_1, \ldots)$ with $s'_n = s_n$.

We note that in the Markov case the stochastic process X to be stopped as well as the process X^τ stopped by a stopping time $\tau \in \mathfrak{N}$ are Markov chains.

Now we want to fit the StS to a DS with Markov components. In particular, we suppose that the set Θ of admissible strategies consists of all Markov strategies, i.e., the decision functions depend on the last state only.

We have the following particularity here: Whereas in a DS rewards (or costs) will occur at every step, in a stopping problem a payment is provided only once. If we do not register whether a stopping decision is already made or not, an immediate description of a stopping system as a DS is not possible.

The first step in overcoming this difficulty consists in marking every state $s \in \mathfrak{S}$ according to whether before reaching $s \in \mathfrak{S}$ a stopping decision d

$= 1$ is made or not. This leads to a definition of new states of the form (s, d), $s \in \mathfrak{S}$, $d \in \{0, 1\}$.

However, because of the fact that after a stopping decision the DS must remain in the state reached and that in every state $(s, 1)$ the same reward, namely 0, is paid at every step, the description can further be simplified by clustering all states $(s, 1)$, $s \in \mathfrak{S}$, into a state σ, the so-called *cemetery*. The states $(s, 0)$, $s \in \mathfrak{S}$, will be denoted by s, as before.

Now, a stopping system in the Markov case can be described as a DS: $(N, \tilde{\mathfrak{S}}, \mathfrak{D}, p_0, (p_n), (c_n), \Theta)$ with step costs c_n instead of cumulative costs k_n (see [7], 5), where in detail we have the following:

– N, $\mathfrak{D}$, p_0 as before;

– $\tilde{\mathfrak{S}} = \mathfrak{S} \cup \{\sigma\}$, $\sigma \notin \mathfrak{S}$;

– $p_n(s'|s, 0) = p_n(s'|s)$, $s, s' \in \mathfrak{S}$; $p_n(\sigma|s, 1) = 1$, $s \in \mathfrak{S}$; $p_n(\sigma|\sigma, d) = 1$, $d \in \mathfrak{D}$, $n = 1, 2, \ldots$;

– $c_n(s, 0) = 0$, $s \in \mathfrak{S}$; $c_n(s, 1) = -g_n(s)$, $s \in \mathfrak{S}$; $c_n(\sigma, d) = 0$, $d \in \mathfrak{D}$, $n = 0, 1, \ldots$;

– Θ, set of all Markov strategies.

We note that v_n^* depends on the last reached state s only. Since $v_n^*(\sigma) = 0$ is known, the state σ does not disturb practical calculations.

For a finite state space $\mathfrak{S}$, in the Markov case we have the following procedures for practical calculations:

– value iteration procedure (for instance, see [12]);

– linear programming (for instance, see [4]);

– stopping set enlargement procedure (for instance, see [3], [6]).

A comparison of all three methods is outlined in [8]. There is a reference to an ALGOL program package which by numerical experience automatically switches to the most favourable solving routine.

References

[1] H. Bauer, *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie*, Walter de Gruyter, Berlin 1974.

[2] Y. S. Chow, H. Robbins and D. Siegmund, *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston 1971.

[3] O. Emrich, *Optimales Stoppen von endlichen Markov-Ketten*, Proc. Oper. Res. **1** (1972), 604–617.

[4] W. R. Heilmann, *Solving stochastic dynamic programming problems by linear programming*, Z. Oper. Res. Ser. A-B **22** (1978), 33–53.

[5] S. Iwamoto, *Stopped decision processes on compact metric spaces*, Bull. Math. Statist. **14** (1970), 51–60.

[6] H. Kogelschatz, *Über ein Verfahren zur Bestimmung optimaler Stoppregeln*, Proc. Oper. Res. **2** (1973), 267–277.

[7] P. H. Müller, *An optimality-equation for discrete stochastic decision problems with general sets of admissible strategies*, this volume.

[8] P. Neumann, *Zur numerischen Lösung von Stopproblemen bei Markov-Ketten*, Tagung *Mathematische Statistik in der Technik*, Bergakademie Freiberg 1980, 107–112.
[9] M. L. Nikolaev, *Obobščennyje posledovateľnyje procedury*, Litovsk. Mat. Sb. **19** (1979), 35–44.
[10] U. Rieder, *On stopped decision processes with discrete time parameter*, Stochastic Process. Appl. **3** (1975), 365–383.
[11] H. Sieler, *Zur Theorie der optimalen Steuerung von Markowschen steuerbaren Prozessen mit diskreter Zeit*, Dissertation, Universität Jena, Jena 1978.
[12] A. N. Shiryayev, *Statističeskij posledovateľnyj analiz*, Nauka, Moskva 1976.
[13] J. Van der Wal, *The method of successive approximations for the discounted Markov game*, Internat. J. Game Theory **6** (1977), 11–22.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981