The policy iteration method for the optimal stopping of a Markov chain and applications to a free boundary problem for random walks

TECHNOLOGICAL UNIVERSITY EINDHOVEN

Department of Mathematics

STATISTICS AND OPERATIONS RESEARCH GROUP

Memorandum COSOR 74-12

The policy iteration method for the optimal stopping of a Markov chain and applications to a free boundary problem for random walks

by

K.M. van Hee

Eindhoven, november 1974

The policy iteration method for the optimal stopping of a Markov chain and applications to a free boundary problem for random walks

by

K.M. van Hee

0. Summary

In this paper we study the problem of the optimal stopping of a Markov chain with a countable state space. In each state i the controller receives a reward r(i) if he stops the process and he must pay the cost c(i) otherwise. We show that, under some conditions, the policy iteration method, introduced by Howard, gives the optimal stopping rule in a finite number of iterations. For random walks with a special reward and cost structure the policy iteration method gives the solution of a free boundary problem. Using this property we shall derive a simple algorithm for the determination of the optimal stopping time of such random walks.

1. Introduction

Consider a Markov chain $\{X_n \mid n = 0,1,2,\dots\}$ defined on the probability space $(\Omega, F, \mathbb{P})$. The state space S is countable. We suppose that $\mathbb{P}[X_0 = i] > 0$ for all $i \in S$. For all $A \in F$ $\mathbb{P}_i[A]$ is the conditional probability of A given $X_0 = i$. On S real functions r and c are defined, where r(i) is the reward if the process is stopped in state i and c(i) is the cost if the process goes on. We consider stopping times T (for a definition see [7]). The expected reward at time T, given $X_0 = i$, is defined by

$$\mathbb{E}_i[r(X_T)] = \int_{\{T<\infty\}} r(X_T)\, d\mathbb{P}_i \ .$$

We restrict our attention to reward functions r with

$$\mathbb{E}_i[\,|r(X_T)|\,] < \infty$$

for all $i \in S$ and all stopping times T.

Let P be the transition matrix of the Markov chain, with components $P(i,j)$ for $i,j \in S$. If c is a function on S, we define the function Pc by

$$Pc(i) := \sum_{j \in S} P(i,j)c(j)$$

and the function $P^n c$ by $P^n c := P(P^{n-1}c)$.
We call a function c on S a charge (see [3]) if

$$\sum_{n=0}^{\infty} P^n |c| < \infty .$$

(Note that the function $v \le w$ if $v(i) \le w(i)$ for all $i \in S$ and $v < w$ if $v \le w$ and for at least one $i \in S$ $v(i) < w(i)$.)
We suppose the cost function c to be either a charge or a nonnegative function.
For a stopping time T the expected return $v_T(i)$, given the starting state i, is defined by

$$v_T(i) := \mathbb{E}_i[r(X_T) - \sum_{n=0}^{T-1} c(X_n)].$$

The existence of the expected return $v_T(i)$ is guaranteed for all T since $|\mathbb{E}_i[r(X_T)]| < \infty$ for all i and c is either a charge or a nonnegative function.
Note that $v_T(i) = -\infty$ is permitted.
The value function $v(i)$ is the supremum over all the stopping times T

$$v(i) := \sup_T v_T(i) .$$

In the rest of this section we summarize some properties of stopping problems.

1.1. The value function $v(i)$ satisfies the functional equation

$$v(i) = \max\{r(i), -c(i) + \sum_{j \in S} P(i,j)v(j)\}$$

(see [3] and [7]).

1.2. The value function $v(i)$ is the smallest solution of this functional equation under each of the following conditions.
a) c is a charge
b) $c \ge 0$ and $r \ge 0$
(for c is a charge this is proved in [3], for the other case the proof proceeds analogously).

1.3. If an optimal stopping time exists the entrance time T_Γ in the set

$$\Gamma := \{i \mid r(i) = v(i)\}$$

is optimal. (See [6].)

1.4. The entrance time T_Γ is optimal under each of the following conditions

a) r is bounded and $c \geq \delta > 0$ (δ is a constant vector)

b) r is bounded, $\mathbb{P}_i[T_\Gamma < \infty] = 1$ and either c is a charge or $c \geq 0$

(for the proof of a) see [7], for the case b) see [2] and [3]).

In this paper we shall prove, as a by-product, that T_Γ is optimal under the two conditions

a) $S\backslash\Gamma$ is finite and the Markov chain is irreducible

b) either c is a charge or both r and c are nonnegative.

2. Some preparations and notations

A stopping rule f is a mapping from S to $\{0,1\}$ where $f(i) = 0$ means that the process is stopped in i and $f(i) = 1$ means that the process goes on in state i. The stopping rule f is equivalent with the entrance time T_f in the set $\{i \mid f(i) = 0\}$.

The expected return under a stopping rule f is indicated by $v_f(i)$.

For a stopping rule f we define

2.1. $D_f := \{i \in S \mid f(i) = 1\}$, the go-ahead set.

$\Gamma_f := S\backslash D_f$, the stopping set.

2.2. P_f is the matrix with components

$$P_f(i,j) := \begin{cases} P(i,j) & \text{if } i \in D_f \\ 0 & \text{otherwise .} \end{cases}$$

2.3. d_f is a function on S with

$$d_f(i) := \begin{cases} r(i) & \text{if } i \in \Gamma_f \\ -c(i) & \text{otherwise .} \end{cases}$$

Sometimes we shall suppose for a stopping rule f that the corresponding entrance time T_f in the set Γ_f satisfies the condition

2.4. There exists a k and an ε, $0 < \varepsilon \leq 1$ such that

$$\mathbb{P}_i[T_f > k] \leq 1 - \varepsilon, \ \forall_{i \in D_f} .$$

<u>Lemma 1</u>. For a stopping rule f, satisfying condition 2.4 it holds that

a) $(I - P_f)$ is invertible (I is the matrix on S with components

$$I(i,j) := \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{otherwise .} \end{cases}$$

b) $v_f = (I - P_f)^{-1} d_f$.

c) $\mathbb{P}_i[T_f < \infty] = 1, \ \forall_{i \in S}$.

<u>Proof</u>. Define, for a natural number k, the vector a_k on S by

$$a_k(i) := \begin{matrix} \mathbb{P}_i[T_f \geq k] & \text{if } i \in D_f \\ 0 & \text{otherwise .} \end{matrix}$$

It is easily verified that

$$a_k = P_f^k . 1$$

(where 1 is the vector on S with all components equal to one).
Condition 2.4 implies

$$\| P_f^{k+1} \| \leq 1 - \varepsilon$$

(where $\| A \|$ is the supremum of the row sums of A).
Hence $(I - P_f)$ is invertible. As a consequence of

$$\| a_{(k+1)m} \| = \| P_f^{(k+1)m} \| \leq \| P_f^{k+1} \|^m$$

we get

$$\lim_{n \to \infty} \mathbb{P}_i[T_f > n] = 0$$

for all $i \in S$. Hence

$$\mathbb{P}_i[T_f < \infty] = 1$$

for all $i \in S$. For $i \in D_f$ we have

$$v_f(i) = -c(i) + \sum_{j \in D_f} P(i,j) v_f(j) + \sum_{j \in \Gamma_f} P(i,j) r(j)$$

and for $i \in \Gamma_f$

$$v_f(i) = r(i) .$$

Hence $v_f = P_f v_f + d_f$ which implies $v_f = (I - P_f)^{-1} d_f$. □

The next lemma gives a sufficient condition for a stopping problem to satisfy condition 2.4.

Lemma 2. Let r be bounded and $c(i) \geq \delta > 0$ for all $i \in S$. For the optimal stopping rule f condition 2.4 holds.

Proof. The existence of an optimal stopping rule follows from 1.4. Suppose $m \leq r(i) \leq M$ for all $i \in S$, $M > 0$. Choose ε and k such that $(1-\varepsilon)k > \frac{M-m}{\delta}$. Suppose $\mathbb{P}_{i_0}[T_f \leq k] < \varepsilon$ for some $i_0 \in S$. Then

$$v_f(i_0) \leq M - \delta . \mathbb{E}_{i_0}[T_f] \leq M - (1-\varepsilon)\delta k < m$$

so the stopping rule f^* defined by

$$f_*(i) := \begin{cases} f(i) & \text{for } i \neq i_0 \\ 0 & \text{for } i = i_0 \end{cases}$$

is better than f in at least one state which produces a contradiction. □

3. The policy iteration method

Let f be a stopping rule. For f we define the improved stopping rule f^* by

3.1. $$f_*(i) := \begin{cases} 0 & \text{if } r(i) \geq -c(i) + \sum_{j \in S} P(i,j) v_f(j) \\ 1 & \text{otherwise .} \end{cases}$$

Lemma 3. For a stopping rule f and its improved stopping rule f_* it holds that

$$v_f \leq d_{f*} + P_{f*} v_f .$$

Proof. Let $i \in D_{f*}$ then $f_*(i) = 1$, $d_{f*}(i) = -c(i)$, $P_{f*}(i,.) = P(i,.)$ and

$$r(i) < -c(i) + \sum_{j \in S} P(i,j)v_f(j) = d_{f*}(i) + \sum_{j \in S} P_{f*}(i,j)v_f(j) .$$

Since

$$v_f(i) = -c(i) + \sum_{j \in S} P(i,j)v_f(j)$$

or $v_f(i) = r(i)$, the statement is true for $i \in D_{f*}$. If $i \in \Gamma_{f*}$ then $f_*(i) = 0$, $d_{f*}(i) = r(i)$, $P_{f*}(i,.) = 0$ and

$$r(i) \geq -c(i) + \sum_{j \in S} P(i,j)v_f(j) ,$$

which completes the proof. □

Lemma 4. If the improved stopping rule f_*, derived from f, satisfies condition 2.4 then

$$v_{f*} \geq v_f .$$

Proof. From lemma 1 it follows that $(I - P_{f*})^{-1}$ exists and that $v_{f*} = (I - P_{f*})^{-1} d_{f*}$. Since $(I - P_{f*})^{-1} = \sum_{n=0}^{\infty} P_{f*}^{\ n}$ we know that $(I - P_{f*})^{-1}$ has nonnegative components only. We conclude from lemma 3 that $(I - P_{f*})v_f \leq d_{f*}$, hence $v_f \leq (I - P_{f*})^{-1} d_{f*} = v_{f*}$. □

Now we are ready to derive a method to determine an optimal stopping rule. This method is called the policy iteration method. The method determines a sequence of stopping rules $f_0, f_1, f_2, \ldots$ where f_{n+1} is the improved stopping rule of f_n.

3.2. The policy iteration method:

a) $f_0(i) = 0$ for all i S.

b) determine $v_{f_n} = (I - P_{f_n})^{-1} d_{f_n}$.

c) define f_{n+1} by 3.1.

d) repeat b and c.

Since the value function v satisfies $v \geq v_{f_n} \geq r$ for all n, the following will be true

$$\Gamma = \{i \in S \mid v(i) = r(i)\} \subset \{i \in S \mid v_{f_n}(i) = r(i)\} = \Gamma_{f_n} .$$

(The last equality will follow from theorem 1.)
Hence, for the entrance time T_{f_n} in the set Γ_{f_n}, we have

$$T_{f_n} \leq T_\Gamma .$$

If T_Γ satisfies condition 2.4 the same is true for T_{f_n} for all n.

Theorem 1. Under the conditions

a) either c is a charge or $c \geq 0$ and $r \geq 0$, both

b) T_{f_n} satisfies condition 2.4 for all natural numbers n

the following holds

1) $f_n(i)$ and $v_{f_n}(i)$ are nondecreasing in n

2) if $f_n(i) < f_{n+1}(i)$ for at least $i = i_0$ then $v_{f_n}(i_0) < v_{f_{n+1}}(i_0)$

3) if $f_n(i) = f_{n+1}(i)$ for all $i \in S$ then $f_n(i)$ is optimal.

Proof.

Assertion 1. From lemma 4 it follows that $v_{f_n} \geq v_{f_{n-1}}$ for $n \geq 1$. If $f_{n-1}(i) = 1$ then

$$r(i) < -c(i) + \sum_{j\in S} P(i,j)v_{f_{n-2}}(j) \leq -c(i) + \sum_{j\in S} P(i,j)v_{f_{n-1}}(j) \quad (n \geq 2).$$

Hence $f_n(i) = 1$ for $n \geq 2$. For $n = 1$ the assertion is trivial.

Assertion 2. Suppose $f_n(i_0) = 0$ and $f_{n+1}(i_0) = 1$.

$$v_{f_n}(i_0) = r(i_0) < -c(i_0) + \sum_{j\in S} P(i_0,j)v_{f_n}(j) \leq -c(i_0) +$$

$$+ \sum_{j\in S} P(i_0,j)v_{f_{n+1}}(j) = v_{f_{n+1}}(i_0) .$$

Assertion 3. Let $f_n(i) = f_{n+1}(i)$ for all $i \in S$. Then $v_{f_n}(i) = v_{f_{n+1}}(i)$ and if $f_n(i) = 0$ then

$$v_{f_n}(i) = r(i) \geq -c(i) + \sum_{j \in S} P(i,j) v_{f_n}(j)$$

and if $f_n(i) = 1$ then

$$v_{f_n}(i) = -c(i) + \sum_{j \in S} P(i,j) v_{f_n}(j) > r(i) \ .$$

Hence v_{f_n} satisfies the functional equation 1.1. Condition a guarantees that the value function v is the smallest solution of 2.1. Since $v_{f_n} \leq v$ it must hold that $v = v_{f_n}$. □

Note that $v_{f_n}(i) = r(i)$ implies $f_n(i) = 0$, hence

$$\Gamma_{f_n} = \{i \in S \mid v_{f_n}(i) = r(i)\} \ .$$

Lemma 5. Let $S \backslash \Gamma$ be finite. Under each of the following conditions 2.4 is satisfied for T_Γ.

a) The Markov chain is irreducible.

b) $c \geq 0$ and $r \geq 0$.

c) c is a charge and $r \geq 0$.

Proof. If a holds the proof is straightforward. For b and c suppose the contrary of the statement. Then there is at least a state i_0 with $\mathbb{P}_{i_0}[T_\Gamma > k] = 1$ for all k. Hence i_0 does not communicate with the states of Γ. Let

$$\Delta := \{i \in S \backslash \Gamma \mid \mathbb{P}_i[T_\Gamma = \infty] = 1\} \ .$$

Then Δ is a distinct Markov chain with a finite state space, hence there must be a recurrent class B. (The subscript B indicates the restriction to B of vectors and matrices.) Hence $v_B = -c_B + P_B . v_B$, with P_B a Markov matrix. Suppose that b holds. Then $v_B \leq P_B v_B$ and therefore $v_B \leq P_B^* v_B$, where P_B^* is the Cesaro-limit of P_B^n for $n \to \infty$. The vector $d := P_B^* c_B$ is constant and so v_B is constant. Hence $c_B = 0$. It is easy to verify that there exists an optimal stopping rule for the chain B and that $v(i) = \max_{j \in B} r(j)$ for $i \in B$. But for all $i \in B$ $v(i) > r(i)$. This produces a contradiction. Suppose that c holds. Since c is a charge, $\lim_{n\to\infty} P_B^n|c_B| = 0$. Let e be the period of the class B. Then P_B^{ne} has a limit, different from 0. Hence $c_B = 0$. Again there exists an optimal stopping rule for the chain B and $v(i) = \max_{j\in B} r(j)$ for $i \in B$. Which also produces a contradiction. □

Corollary.

1) If $S\backslash\Gamma$ is finite and condition 2.4 holds for T_Γ then condition b of theorem 1 is satisfied. Hence if in addition either c is a charge or both $r \geq 0$ and $c \geq 0$, we proved the existence of an optimal stopping rule.
2) In this case the policy iteration method leads in a finite number of steps to the optimal stopping rule. If in addition, from each state $i \in S\backslash\Gamma$ only a finite number of states is within reach, it is easy to derive an algorithm for the policy iteration method.

4. Free boundary problems for random walks

We shall study the optimal stopping problem of a random walk as an application of the theory develloped in the foregoing sections. For simplicity we restrict our attention to one dimensional random walks.

Problem formulation

Consider a random walk with state space the set of integers ($\mathbb{Z}$) and transition matrix P defined by

$$
\text{4.1.}\qquad
\begin{aligned}
P(i,i+1) &:= p_i\\
P(i,i-1) &:= q_i\\
P(i,i) &:= s_i
\end{aligned}
$$

where p_i, q_i, $s_i \geq 0$ and $p_i + q_i + s_i = 1$ for all $i \in \mathbb{Z}$.
Let the reward function r and the cost function c satisfy one of the conditions 1.2. Define the set c by

$$C := \{i \in \mathbb{Z} \mid r(i) < -c(i) + p_i r(i+1) + q_i r(i-1) + s_i r(i)\}\ .$$

Suppose that

4.2. $d,e \in \mathbb{Z}$ exist such that $C = \{i \in \mathbb{Z} \mid d \leq i \leq e\}$.

Further suppose that

4.3. either $p_i > 0$ and $q_i > 0$ for all i or $r \geq 0$.

Condition 4.2 says that for $i \in \mathbb{Z}\backslash C$ immediately stopping is more profitable than making one more transition. In statistical sequential analysis this condition is satisfied in a natural way (see for example [5]). Condition 4.3 guarantees that for each go-ahead set D_{f_n} the entrance time in the set Γ_{f_n}, T_{f_n}, satisfies condition 2.4.

Theorem 2. Consider the stopping problem described above. The policy iteration method applied to this problem has the following properties:

a) for the sequence of sets D_{f_n}, $n = 0,1,2,...$ it holds that for each $n \in \mathbb{N}$ $\exists_{k,\ell\in\mathbb{Z}}$ such that $D_{f_n} = \{i \in \mathbb{Z} \mid k \le i \le \ell\}$

b) the entrance time T_{f_n} in the set Γ_{f_n} satisfies condition 2.4

c) if $D_{f_n} = \{i \in \mathbb{Z} \mid k \le i \le \ell\}$ then

$$D_{f_n} \subset D_{f_{n+1}} \subset \{i \in \mathbb{Z} \mid k-1 \le i \le \ell+1\},\ n \ge 1 .$$

The set Γ has the form

$$\Gamma = \{i \in \mathbb{Z} \mid i \ge \ell \vee i \le k\},\ k,\ell \in \mathbb{Z} \cup \{-\infty,+\infty\} .$$

Proof. Use induction $f_0 = 0$. It is easy to verify that $f_1(i) = 1$ for all $i \in C$ and $f_1(i) = 0$ for all $i \in \mathbb{Z}\backslash C$, hence $D_{f_1} = C$. We shall prove that T_{f_1} satisfies 2.4 in a way analogously to lemma 5. If the random walk is irreducible the proof is straightforward. Suppose now $r \ge 0$. Suppose T_{f_1} does not satisfy 2.4. Then there exist a recurrent class $B \subset D_{f_1}$ (the subscript B indicates the restriction to B of the matrix P and the vectors v_{f_n}, r and c). Let $c \ge 0$. Then $v_B \le P_B v_B$ and so $v_B \le P_B^* v_B$ where P_B^* is the Cesaro-limit of P_B^n for $n \to \infty$. Hence v_B is constant and therefore $c_B = 0$. If c is a charge $P_B^n|c_B| \to 0$ and therefore $c_B = 0$ too. Hence

$$v_B = -\sum_{n=0}^{\infty} P_B^n c_B = 0 .$$

We know from theorem 1 that $v_{f_1}(i) > r(i)$ for $i \in D_{f_1}$. But $r_B \geq 0$. This produces a contradiction. Hence T_{f_1} must satisfy condition 2.4. Suppose that $D_{f_n} = \{i \in \mathbb{Z} \mid k \leq i \leq \ell\}$ with $C \subset D_{f_n}$. In the same way as above we can prove that T_{f_n} satisfies condition 2.4. From the proof of theorem 1 it follows that $f_{n+1}(i) = 1$ if $i \in D_{f_n}$. For $i \geq \ell+2$ and $i \leq k-2$ it holds that $f_{n+1}(i) = 0$ because $v_{f_n}(i) = r(i)$ for $i \geq \ell+1$ and $i \leq k-1$ and since $i \in \mathbb{Z}\backslash C$. Therefore it only can happen in the points $i = k-1$ and $i = \ell+1$ that $f_{n+1}(i) > f_n(i)$. This proves the assertions a, b and c. If C is empty it holds that $f_1(i) = 0$ for all $i \in \mathbb{Z}$ such that $r = v$ and $\Gamma = \mathbb{Z}$. If for some n $f_n = f_{n+1}$ then it follows from theorem 1 that f_n is optimal and that $\Gamma = D_{f_n}$. If there does not exist a n for which $f_n = f_{n+1}$ then D_{f_n} is ascending to a (half) open interval. Γ is the complement of this set. □

From now on we shall suppose that $s_i = 0$ for all $i \in \mathbb{Z}$. This is allowed since we may define, if $s(i) \neq 0$ for all $i \in \mathbb{Z}$,

$$\bar{c}(i) := \frac{c(i)}{1 - s(i)} \quad , \quad \bar{p}_i := \frac{p_i}{1 - s_i} \quad \text{and} \quad \bar{q}_i := \frac{q_i}{1 - s_i} \; .$$

For this process $\bar{s}_i = 0$.

We shall discuss the connection between a free boundary problem and the policy iteration method.

Consider the function $\{(i,x(i)) \mid i \in \mathbb{Z}\}$. The difference operator Δ for this function is defined by

$$\Delta x(i) := x(i + 1) - x(i) \; .$$

For the difference equation

4.4. $\quad p_i \Delta x(i) - q_i \Delta x(i - 1) = c(i)$

we define the free boundary problem in the following way.

4.5. Find the function x, satisfying 4.4, and the smallest interval $[k,\ell]$, such that $C \subset [k,\ell]$ with the properties

a) $x(i) \geq r(i)$ for $k \leq i \leq \ell$, $x(k) = r(i)$ if $i = k-1$ and if $i = \ell+1$.

b) $p_{k-1}\Delta x(k-1) - q_{k-1}\Delta r(k-2) \leq c(k-1)$.

c) $p_{\ell+1}\Delta r(\ell+1) - q_{\ell+1}\Delta x(\ell) \leq c(\ell+1)$.

Theorem 3. The policy iteration method determines the solution of the free boundary problem 4.5, if the solution exists. If $[k,\ell]$ belongs to the solution of 4.5, then the entrance time in the set $\{i \in \mathbb{Z} \mid i < k \vee i > \ell\}$ is the optimal stopping time of the stopping problem specified by 4.1, 4.2 and 4.3.

Proof. Let $[k,\ell]$ be the solution of 4.5 and let $x(i)$ be the unique function determined by 4.4 and 4.5. Define $w(i)$ for $i \in \mathbb{Z}$ by

$$w(i) := \begin{cases} x(i) & \text{if } k \leq i \leq \ell \\ r(i) & \text{otherwise .} \end{cases}$$

Then $w(i)$ satisfies the functional equation 1.1, hence $w(i) \geq v(i)$ (where $v(i)$ is the value function of the stopping problem). So

$$\{i \in \mathbb{Z} \mid v(i) > r(i)\} \subset \{i \in \mathbb{Z} \mid w(i) > r(i)\} = \{i \in \mathbb{Z} \mid k \leq i \leq \ell\}.$$

According to theorem 2 $\{i \in \mathbb{Z} \mid v(i) > r(i)\} = D_{f_n}$ for some n, hence $D_{f_n} \subset [k,\ell]$. On the other hand $v_{f_n}(i)$ is a solution of 4.4, 4.5 a, b and c, so $\{i \in \mathbb{Z} \mid k \leq i \leq \ell\} \subset D_{f_n}$. $\square$

The correspondence between the solution of boundary value problems and optimal stopping of the Wiener process, the continuous analog of our process is treated in [5].

Example. Let $p_i = q_i = \frac{1}{2}$, $c(i) = c$ with $0 < c < 1$ and $r(i) = |i|$. The difference equation 4.4 has the form

4.6. $\Delta x(i) - \Delta x(i-1) = 2c$.

Hence

$$\Delta^2 x(i) = 2c \quad \text{and so} \quad \Delta^3 x(i) = 0 \ .$$

By the elementary theory of linear difference equations the solution is

$$x(i) = \alpha + i\beta + i^2\gamma$$

where α, β and γ are constants. Substituting in 4.6 gives $\gamma = c$, and from $x(k - 1) = r(k - 1)$ and $x(\ell + 1) = r(\ell + 1)$ it follows that

$$\beta = \frac{r(\ell + 1) - r(k - 1)}{\ell - k + 2} - (k + \ell)c$$

$$\alpha = \frac{(\ell + 1)r(k - 1) - (k - 1)r(\ell + 1)}{\ell - k + 2} + c(\ell + 1)(k - 1) .$$

By induction it follows from the symmetry of $r(i)$ and the transition probabilities that D_{f_n} is symmetric around $i = 0$. Hence $k = -\ell$. Therefore $\beta = 0$ and $\alpha = (\ell + 1)\{1 - c(\ell + 1)\}$, which implies that $x(i)$ is also symmetric around $i = 0$. According to 4.6 $\Delta x(i) = (2i + 1)c$ and from the conditions 4.5 b and c it follows that

$$\ell \geq \max\{0, \frac{1 - 3c}{2c}\} .$$

Hence

$$\Gamma = \{i \in Z \mid |i| \geq 1 + \max(0, \frac{1 - 3c}{2c})\} .$$

If $c = 0$ $x(i)$ is constant and 4.5 b and c will not be satisfied for all ℓ, hence never stopping is optimal.

Consider the free boundary problem 4.5. Now we shall derive an algorithm to determine the solution. Call

$$z_i := \Delta x(i) , \quad a_i := \frac{q_i}{p_i} \quad \text{and} \quad b_i := \frac{c(i)}{p_i} .$$

The difference equation 4.4 becomes

$$z_i - a_i z_{i-1} = b_i .$$

With induction on m it is easy to verify that for $k \leq m$

4.7. $$z_m = z_{k-1} \prod_{i=k}^{m} a_i + \sum_{i=k}^{m} \{b_i \prod_{j=i+1}^{m} a_j\}$$

(an empty product has the value 1, an empty sum the value 0). Because $x(\ell + 1) = r(\ell + 1)$ and $x(k - 1) = r(k - 1)$ it holds that

$$r(\ell + 1) - r(k - 1) = \sum_{m=k-1}^{\ell} z_m$$

hence

$$\text{4.8.} \qquad z_{k-1} = \frac{r(\ell + 1) - r(k - 1) - \sum\limits_{m=k-1}^{\ell} \sum\limits_{i=k}^{m} \{b_i \prod\limits_{j=i+1}^{m} a_j\}}{\sum\limits_{m=k-1}^{\ell} \prod\limits_{i=k}^{m} a_i} .$$

From 4.7 and 4.8 one can compute $z_{\ell+1}$.
Now we shall give the algorithm. A stop criterion is not included, so that in case Γ is half open or empty the algorithm is not finite.

<u>The algorithm</u>

1. $k := d$, $\ell := e$, $i := 0$
2. compute z_{k-1} and $z_{\ell+1}$
3. if $z_{k-1} > b_{k-1} + a_{k-1}\{r(k - 1) - r(k - 2)\}$ then $k := k - 1$ and $i := 1$
4. if $z_{\ell+1} > b_{\ell+1} + a_{\ell+1}\{r(\ell + 2) - r(\ell + 1)\}$ then $\ell := \ell + 1$ and $i := 1$
5. if $i = 0$ then stop, otherwise goto 2.

If the algorithm stops then

$$\Gamma = \{i \in Z \mid i \geq \ell + 1 \text{ or } i \leq k - 1\} .$$

It is easy to verify that the sums and products in 4.7 and 4.8 can be computed recursively. The algorithm uses the fact that for checking the boundary conditions 4.5 b anc c. It is only necessary to know the differences of x(i).

Literature

[1] Dynkin, E.B., Juschkewitsch, A.A.; Sätze und Aufgaben über Markoffsche Prozesse. Springer-Verlag (1969).

[2] Hordijk, A., Potharst, R., Runnenburg, J.Th.; Optimaal stoppen van Markov ketens. MC-rapport, SD 104/72 (1972).

[3] Hordijk, A.; Dynamic programming and Markov potential theory. MC tract (1974).

[4] Howard, R.A.; Dynamic programming and Markov processes. Technology Press, Cambridge Massachusetts (1960).

[5] van Hee, K.M., Hordijk, A.; A sequential sampling problem solved by optimal stopping. MC-rapport SW 25/73 (1973).

[6] van Hee, K.M.; Note on memoryless stopping rules. COSOR-notitie R-74-12, T.H. Eindhoven (1974).

[7] Ross, S.; Applied probability models with optimization applications. Holden-Day (1970).