THEORY PROBAB. APPL.
Vol. 54, No. 3, pp. 1–19

Translated from Russian Journal

ON OPTIMAL STOPPING OF RANDOM SEQUENCES MODULATED BY MARKOV CHAIN*

PRESMAN E.L.† AND SONIN I. M.‡

(*Translated by the authors*)

Abstract.
This paper has two main goals: first, to describe a new class of optimal stopping problems for which the solutions can be found either in an explicit form, or in a finite number of steps, and second, to demonstrate the potential of the State Elimination algorithm developed by one of the authors earlier, for the problem of optimal stopping of a finite or countable Markov chain.

Key words. Markov chain, optimal stopping, State Elimination algorithm, seasonal observations.

DOI. 10.1137/S0040585X979XXXX

1. Introduction and problem setting. Let us consider the following three problems which have a common structure and each of them being a generalization of the previous one. In each problem a statistician observes a random sequence $Z = (Z_n)_{n\geqq 0}$, and her goal is to maximize $\mathbf{E}\,\beta^\tau g(Z_\tau)$ over all possible stopping times τ, where β is a discount factor, $0 < \beta \leqq 1$, and g is a reward function.

Problem 1: A classical problem of optimal stopping of independent trials, see e.g. [6]. A sequence of i.i.d. random variables $(Z_n)_{n\geqq 0}$ with known distribution F is observed and $\beta < 1$. For example, a six side die is rolled repeatedly. It is well known that the solution for this problem is given by a threshold value $c^* = c^*(\beta)$, such that the moment of a first visit to a set $\{z\colon\ g(z) \geqq c^*\}$ is optimal.

Problem 2: "Seasonal observations". In this problem a "die" is replaced by m different "dice" and they are rolled sequentially, i.e. first, second, ..., mth, again first, second, and so on. Formally, let $(Y_n^s)_{n\geq 0}$ be m independent i.i.d. sequences with known distributions $F^s = \{q^s(i), i = 1, 2, ...\}$. A sequence of observations (Z_n) has a form $Z_n = Y_n^s$, where $s \equiv n \pmod m$, $s \in B = \{1, 2, ..., m\}$, $\beta < 1$. Problem 1 is a special case of Problem 2 when $m = 1$.

Problem 3: Sequence of observations modulated by Markov chain (generalized seasonal observations). In this problem, as in Problem 2, there are m different "dice" but which die to observe at moment n is specified by the position of an underlying Markov chain $(U_n)_{n\geq 0}$ which take values in a set B, independent of observations on dice. Formally, as in Problem 2, m independent i.i.d. sequences $(Y_n^s)_{n\geqq 0}$ with known distributions $F^s = \{q^s(i), i = 1, 2, ...\}$ are given. A sequence of observations (Z_n) now has a form $Z_n = (U_n, Y_n^{U_n})$, where (U_n) is a finite Markov chain with values in B and known transitional probabilities $p(s, k)$. Problem 2 is a special case of Problem 3 when Markov chain $(\widetilde{U}_n)_{n\geqq 0}$ is a deterministic cyclical movement along set B.

Let us remind that in the applied probability models the case $\beta < 1$ can be reduced to the case $\beta = 1$ if an additional *absorbing state* is introduced. So we introduce an absorbing state $e = (\widetilde{e}, 0)$, and instead of $(\widetilde{U}_n)_{n\geqq 0}$ we consider new Markov chain

*The work of the first author was supported by RFBR (project 07-01-00541).
†CEMI RAS, Nakhimovski prospect, 47, 117428 Moscow, Russia; e-mail: presman@cemi.rssi.ru
‡Dept. of Mathematics, Univ. of North Carolina at Charlotte, Charlotte, NC 28223, USA; CEMI RAS, Moscow, Russia; e-mail: imsonin@uncc.edu

$(U_n)_{n\geqq 0}$ with values in $\widetilde{e}\cup B$, and transition probabilities $p_{s,k} = \beta\widetilde{p}_{s,k},\ s,k\in B$; $p_{s,\tilde{e}} = 1-\beta$, $s\in B$, and we define $Y_n^{\tilde{e}}\equiv 0$, $g(e)=0$. Then for any stopping time τ

$$\mathbf{E}\big[\beta^\tau g(\widetilde{U}_\tau, Y_\tau^{\widetilde{U}_\tau})\big] = \mathbf{E}\big[g(U_\tau, Y_\tau^{U_\tau})\big].$$

In this paper we consider even more general problem than Problem 3, when the transition probability into absorbing state may depend on the state of Markov chain $(\widetilde{U}_n)_{n\geqq 0}$. We assume that a sequence of random variables $Z=(Z_n)_{n\geqq 0}$ with values in $X=\{e\}\cup(B\otimes\mathbf{R})$, is defined on a probability space $(\Omega,\mathcal{F},\mathbf{P})$, such that $Z_n=(U_n,Y_n), n\geqq 0$, and

$$\begin{aligned} &\mathbf{P}[Z_{n+1}=e\,|\,Z_0,\dots,Z_{n-1},Z_n=e]=1,\\ &\mathbf{P}[U_{n+1}=j,Y_{n+1}\in A\,|\,Z_0,\dots,Z_{n-1},Z_n=z=(i,y)]=p_{ij}F(j,A), \qquad (1)\\ &\mathbf{P}[Z_{n+1}=e\,|\,Z_0,\dots,Z_{n-1},Z_n=z=(i,y)]=1-\sum_{j\in B}p_{ij}\end{aligned}$$

for $i,j\in B$, $y\in\mathbf{R}$, where $F(j,\cdot)$ are probability measures on a Borel space $\mathbf{R}$, $F(j,\mathbf{R})=1$, $j\in B$, a $P=(p_{ij},\ i,j\in B)$ is a strictly substochastic matrix, i.e. a matrix with nonnegative values such that

$$\sum_{j\in B}p_{ij}<1 \text{ for all } i\in B. \qquad (2)$$

Obviously, the random sequences $Z=(Z_n)_{n\geqq 0}$ and $(U_n)_{n\geqq 0}$ form Markov chains and Markov chain $(U_n)_{n\geqq 0}$, with transitions for all states except an absorbing state defined by matrix P, is in a sense a *master* chain “modulating” sequence $(Y_n)_{n\geqq 0}$.

Such situation is described sometimes by saying that *random sequence* $(Y_n)_{n\geqq 0}$ *is defined on a Markov chain* $(U_n)_{n\geqq 0}$, or that *Markov chain Z is defined by matrix P and measures $F(j,\cdot)$, $j\in B$.*

If $\sum_{j\in B}p_{ij}=\beta$ for all $i\in B$, $0<\beta<1$, we obtain a problem equivalent to Problem 3, with transition probabilities $\widetilde{p}_{ij}=p_{ij}/\beta$, $i,j\in B$. Note that the assumption (2) can be relaxed, e.g. assuming that only some power of matrix P satisfies this condition.

We assume that a measurable function $g(\cdot)$ is defined on X, such that $g(e)=0$, $\int_{\mathbf{R}}|g(i,v)|\,F(i,dv)<\infty$ for all $i\in B$, and we consider the problem of optimal stopping (OS) of Markov chain Z, defined by a matrix P, measures $F(i,\cdot)$, with terminal reward function $g(\cdot)$ and $\beta=1$. The *value function* in this problem is denoted by $V(z)=\sup_\tau\mathbf{E}[g(Z_\tau)\,|\,Z_0=z]$.

The main result of our paper is the following theorem which describes the optimal stopping set and value function for this OS problem. Before formulating this theorem let us introduce some basic notation. For any $A=\{(j,y)\colon\ j\in B,\ y\in A_j\}$, where A_j, $j\in B$, are some (possibly empty) sets from $\mathbf{R}$, we denote by $F_d(A)$ the diagonal matrix

$$F_d(A)=(\delta_{ij}F(j,A_j),\ i,j\in B). \qquad (3)$$

Theorem. *There is a vector $d^*=(d_1^*,\dots,d_m^*)$, such that*

a) *an optimal stopping time τ^* is a moment of the first visit of Markov chain Z into set $\{e\}\cup D^*$, where $D^*=\{z=(i,y)\colon\ i\in B,\ y\in D_i^*\}$, $D_i^*=\{y\colon\ g(i,y)\geqq d_i^*\}$;*

b) *the value function satisfies the equation*

$$V(z)=g(z),\quad z\in D^*,\qquad V(i,y)=d_i^*>g(i,y),\quad z=(i,y)\notin D^*, \qquad (4)$$

and d^* satisfies the equation

$$d_i^* = \sum_{j\in B} p_{ij}^* \int_{D_j^*} g(j,v)\, F(j,dv), \tag{5}$$

where matrix $P^* = (p_{ij}^*,\ i,j \in B)$ is defined by the equality

$$P^* = [I + PF_d(D^*) - P]^{-1}P. \tag{6}$$

For any initial state i, function $F(j, D_j^*)p_{ij}^* = \mathbf{P}[U_{\widetilde{\tau}} = j \,|\, U_0 = i]$ gives a distribution of the first coordinate of Markov chain Z at the moment $\widetilde{\tau}$ of the first, after moment zero, visit of Z into set D^*;

c) *there is an algorithm to find vector d^*, and therefore to construct the value function and the optimal stopping set. In the case of discrete distributions with support sets without finite limit points, this construction requires only a finite number of steps.*

The proof of this theorem is given in Section 3. Not only do we prove this theorem, but we also present in there an algorithm mentioned in point c) (see (22) – (24) and Lemma 3). This algorithm is based on a general, so-called the State Elimination Algorithm (SEA) developed by I. Sonin for the problem of OS of discrete Markov chain. In Section 2 we recall the main facts from the general theory of optimal stopping and discuss this algorithm. Some preliminary results of this paper were obtained in [12] for the discrete case.

2. The problem of optimal stopping of Markov chain and the State Elimination algorithm. There are two different approaches to the problem of OS, usually called the "martingale" and the "Markovian". Both were developed intensively in the 1960's. The first one is represented by the classical monographs Chow, Robbins and Sigmund (1971) [3] (see also the book of T. Fergusson [6]). The second approach is represented by books A. N. Shiryayev (1969, 1976) [4] and E. B. Dynkin and A. A. Yushkevich (1969) [1]. We mention only two other monographs, [5] and [8], from the numerous lists of books with chapters or sections discussing this subject. The modern presentation of both approaches can be found in the monograph by G. Peskir and A. N. Shiryayev [7]. Our presentation follows the Markovian approach.

Let random sequence $(Z_n)_{n\geqq 0}$ with values in state space $(X, \mathcal{B})$, defined on a measurable space $(\Omega, \mathcal{F})$ with probability measures $\mathbf{P}_z$, measurable with respect to $z \in X$, be a Markov chain with respect to the filtration $\mathcal{F}_n$, where $\mathcal{F}_n$ is a σ-algebra generated by $Z_0, \dots, Z_n$ and $\mathbf{P}_z[Z_0 = z] = 1$. Note that in this section Markov chain Z is a general one and does not coincide with a Markov chain from the previous section. Let $\mathcal{P}$ be a *transition operator* defined by Z, i.e., $\mathcal{P}f(z) = \mathbf{E}_z[f(Z_1)]$. Operator $\mathcal{P}$ maps any measurable function f defined on X with finite mathematical expectation to a measurable function on X. We assume that a number β, $0 < \beta \leqq 1$, and measurable functions $g(z)$ and $c(z)$, are given. The stopping times are defined with respect to the sequence of σ-algebras $\mathcal{F}_n$, $n \geqq 0$. We assume that for any $z \in X$ and any stopping time τ a function $V_\tau(z) = \mathbf{E}_z[g(Z_\tau)\beta^\tau - \sum_{k=0}^{\tau-1} c(Z_k)\beta^k]$ is well defined. Function $g(z)$ is a *terminal reward* for stopping at state z, and $c(z)$ is an *observation fee* (reward) to make one more observation (both functions can be of any sign). To solve the problem of optimal stopping means to find the value function

$$V(z) = \sup_\tau V_\tau(z) = \sup_\tau \mathbf{E}_z\left[g(Z_\tau)\beta^\tau - \sum_{k=0}^{\tau-1} c(Z_k)\beta^k\right], \tag{7}$$

where supremum is taken over all stopping times, and to find an optimal stopping time, i.e., a stopping time where this supremum is attained. As in the previous section, the case when $0 < \beta < 1$ can be reduced to the case $\beta = 1$ by introduction of an absorbing state. Thus, in what follows, without restriction of generality, we assume that $\beta = 1$.

Let us define a *reevaluation operator* $\mathcal{T}$ as follows:

$$\mathcal{T}f(z) = -c(z) + \mathcal{P}f(z). \tag{8}$$

It is well known that under standard conditions the function $V(z)$ is finite and the following theorem holds (see. [7, theorem 1.11, corollary 1.12 and section 11 of chapter 1]).

Statement 1. a) *Function* $V(z)$ *is a minimal solution of optimality equation* (*Bellman equation*)

$$V(z) = \max[g(z), \mathcal{T}V(z)]; \tag{9}$$

b) *if* $\mathbf{P}_z[\tau^* < \infty] = 1$ *for all* $z \in X$, *where* $\tau^* = \inf\{n \geqq 0\colon Z_n \in D^*\}$, $D^* = \{z\colon V(z) = g(z)\}$, *then a stopping time* τ^* *is optimal and* $\tau^* \leqq \tau'$ $\mathbf{P}_z$*-a.s. for all* z *and any stopping time* τ'*;*

c) *the sequence* $\widetilde{V}^{(0)}(z) = g(z)$, $\widetilde{V}^{(k+1)}(z) = \max[g(z), \mathcal{T}\widetilde{V}^{(k)}(z)]$ *satisfies* $\widetilde{V}^{(k)} \uparrow V$.

The set D^* is called a *stopping set*, and the set $C^* = X \backslash D^* = \{z\colon V(z) > g(z)\}$ is called a *continuation set*.

It is sometimes stated that point c) provides a constructive method for calculating function $V(z)$. Note, however, that as a rule $\widetilde{V}^{(k+1)} \neq \widetilde{V}^{(k)} \neq V$ for all k. This maybe true even if Z_n is a Markov chain with only two values, z_1 and z_2, and $g(z_1) \neq g(z_2)$. If Z_n takes values in a finite set, then the equation (9) can be solved via linear programming, but under this approach a probabilistic interpretation is missing and it is not clear how to generalize this approach to the countable state space, for example. Sonin in [9] – [11] proposed an algorithm to solve the OS problem for the finite and sometimes countable state space. This algorithm calculates function $V(z)$ by sequential elimination of some states, and guarantees that if the state space has m, $m < \infty$, points, then the set D^* can be obtained in no more than $(m-1)$ steps and after that the function $V(z)$ can be obtained for all z in the same number of steps. It was mentioned that sometimes $V(z)$ and D^* can be found in a finite number of steps even if the state space is countable.

The Elimination algorithm for the OSP of a Markov chain is based on the following three considerations (see [10], [11]).

1. Although in the OS problem it may be *difficult* to find the states where it is optimal *to stop*, it is relatively *easy* to find a state (states) where it is optimal *not to stop*. Indeed, it is optimal to stop at z if $g(z) \geq c(z) + Pv(z) \equiv Fv(z)$, but v is unknown until the problem is solved. On the other hand, it is optimal not to stop at z if $g(z) < \mathcal{T}g(z)$, i.e., when the expected reward of taking *one more step* is larger than the reward from stopping. (Generally, it is optimal not to stop at any state where the expected reward of taking some, perhaps random number of steps, is larger than the reward from stopping).

2. After we have found a set of states C, which are not in the optimal stopping set, we can eliminate them and recalculate the transition matrix for all remaining states $D = X \setminus C$, i. e., to consider an embedded Markov chain. The current fee

function must also be recalculated with the new fee equal to the expectation of the sum of the original fees between two subsequent visit to set D. It is almost obvious, and we prove this below, that an optimal stopping set for the new Markov chain is the same as for the original Markov chain and the original value function coincides with the new one on a set D, and on the eliminated set C, equals the expected sum of the observation fees, plus the value function at the moment of the first visit to set D.

3. It may happen that after a finite number of such steps, we obtain a situation where for the new Markov chain the following inequality will hold for all states z from the remaining set $\widetilde{D}$

$$g(z) \geqq \widetilde{\mathcal{T}} g(z), \qquad z \in \widetilde{D}, \tag{10}$$

where $\widetilde{c}(z)$ is an observation fee, and $\widetilde{\mathcal{T}}$ is a reevaluation operator for the new Markov chain. If $Z_0 \in \widetilde{D}$, then the inequality (10) will hold for the entire state space of the new Markov chain. Note that if an initial state space has m states, then such situation occurs in no more than $m-1$ steps. But Statement 1 implies that if (10) holds for the entire state space of the new Markov chain, then the state space coincides with the optimal stopping set and the value function is equal to the terminal reward. It remains to calculate value function for states z from C.

A computational procedure for this algorithm for the finite state space was described in [10], and it was a mention of a possibility to use it in the countable case, when after a finite number of steps the situation (10) occurs. The following modification was proposed in [2], that allows also to consider a general state space. The main point is to consider the new Markov chain with the same state space as the initial one, see Lemma 1 and Corollary below. Note that a version of this Lemma on the reduced state space was presented for the countable case in [13]. For simplicity we consider a case where $c(z) \equiv 0$.

For any $A \subset X$, in addition to operator $\mathcal{P}$, let us consider also an operator $\mathcal{P}_A$, defined by an equality

$$\mathcal{P}_A f = \mathcal{P} I_A f, \tag{11}$$

where I_A is an operator of multiplication by a characteristic function of set A.

Let a subset $D \subset X$ be given. Let us denote by C the complement of set D, i.e., $C = X \setminus D$. In the following, a complement of any set D (with possible subindexes and/or indication to dependence on some other parameters) will be denoted as C with the same subindexes and parameters.

Let $\tau_0 = 0$, and $\tau_n > 0$, $n \geqq 1$, be the moments of subsequent visits of Markov chain Z into set D (if $Z_0 = z$ and $z \in D$, then τ_1 is the moment of the first return). Suppose that $\mathbf{P}_z[\tau_1 < \infty] = 1$ for any $z \in X$. Let us introduce a Markov chain $Z' = (Z'_n)_{n \geqq 0}$, where $Z'_n = Z_{\tau_n}$, $n \geqq 0$. Let us denote by $\mathcal{P}'$ the transition operator of Z', and by I, the identity operator, i.e. $I = I_X$.

Lemma 1. *The following equalities are true*

$$\mathcal{P}' f = \sum_{l=0}^{\infty} (\mathcal{P}_C)^l \mathcal{P}_D f = (I - \mathcal{P}_C)^{-1} \mathcal{P}_D f, \tag{12}$$

$$\mathcal{P}' f = \mathcal{P} f + (I - \mathcal{P}_C)^{-1} \mathcal{P}_C (\mathcal{P} f - f). \tag{13}$$

Proof. The second equality in (12) follows from the definition of operator $(I - \mathcal{P}_C)^{-1}$, and the first one is a formula of total probability for the partition $\{\tau_1 = l+1\}$, $l \geqq 0$, since $(\mathcal{P}_C)^l \mathcal{P}_D f(z)$ is $f(Z_{l+1})$, averaged over all trajectories which start in state z, then spend l moments in C, and finally go to D.

Substituting the equality $\mathcal{P}_D = \mathcal{P} - \mathcal{P}_C$ into the right side of equality (12), and using also the equality $(I - \mathcal{P}_C)^{-1} = I + (I - \mathcal{P}_C)^{-1}\mathcal{P}_C$, we obtain

$$
\begin{aligned}
\mathcal{P}' &= (I - \mathcal{P}_C)^{-1}\mathcal{P} - (I - \mathcal{P}_C)^{-1}\mathcal{P}_C \\
&= (I + (I - \mathcal{P}_C)^{-1}\mathcal{P}_C)\mathcal{P} - (I - \mathcal{P}_C)^{-1}\mathcal{P}_C \\
&= \mathcal{P} + (I - \mathcal{P}_C)^{-1}\mathcal{P}_C\mathcal{P} - (I - \mathcal{P}_C)^{-1}\mathcal{P}_C,
\end{aligned}
\tag{14}
$$

which is equivalent to (13). This completes the proof of Lemma 1.

Note that in the case of finite number of states an operator $(I - \mathcal{P}_C)^{-1}$ is represented by a matrix, which is called the *fundamental* matrix for matrix $\mathcal{P}_C$. The entries of this matrix are equal to the expected number of visits to a particular state before an exit from set C.

Formula (13) implies the following statement.

Corollary 1. *If $C \subseteq \{z\colon\ \mathcal{P}f(z) > f(z)\}$, then $\mathcal{P}'f(z) \geqq \mathcal{P}f(z)$ for any $z \in X$.*

In the next section we show how Lemma 1 can be used to construct an optimal stopping set and a value function for the problem in Section 1.

3. Random sequences defined on Markov chain. Let Markov chain Z be defined by a strictly substochastic matrix P and probability measures $F(i,\cdot)$, $i \in B$, defined on the real line. Let us consider the transition operator $\mathcal{P}$ of this Markov chain on a set of functions, such that $f(e) = 0$. Since e is an absorbing state to describe this operator it suffices to consider its action on a set of functions defined only on set $B \otimes \mathbf{R}$. We denote by $\mathcal{P}$ this reduced operator. Functions defined on $B \otimes \mathbf{R}$ can be considered as vector-functions $f(y) = (f(1,y), \dots, f(m,y))$, defined on $\mathbf{R}$. Let us define the operator on a set of such functions by

$$
\mathcal{L}f(i,y) = \int_{\mathbf{R}} f(i,v)\,F(i,dv), \qquad (i,y) \in B \otimes \mathbf{R}. \tag{15}
$$

The definition of operator $\mathcal{L}$ implies that it maps any vector-function to a vector-function with constant coordinates, and therefore the operator $\mathcal{P}$ has the same property and then

$$
\mathcal{P}f = P\mathcal{L}f. \tag{16}
$$

In the following we say that a Markov chain Z is defined by a strictly substochastic matrix P and operator $\mathcal{L}$ в (15).

For any $A = \{(j,y)\colon\ j \in B,\ y \in A_j\}$ in addition to operator $\mathcal{L}$ we shall consider also an operator $\mathcal{L}_A$, defined by the equality

$$
\mathcal{L}_A f = \mathcal{L} I_A f. \tag{17}
$$

This equality, (11) and (15) imply that

$$
\mathcal{P}_A f = P\mathcal{L}_A f. \tag{18}
$$

Before we prove our main theorem let us present some auxiliary statements.

Let a set $D = \{(j, y)\colon\ j \in B,\ y \in D_j\}$ be given. Let $\tau_0 = 0$, and let $\tau_n > 0$, $n \geqq 1$, be the moments of subsequent visits of Markov chain Z into set $\{e\} \cup D$ (if $Z_0 = z$ and $z \in \{e\} \cup D$, then τ_1 is the moment of the first return). Obviously, $\tau_1 \leqq \tau_e$, where τ_e is a moment of the first visit to state e. Since P is a strictly substochastic matrix, then $\mathbf{P}_z[\tau_e < \infty] = 1$, and therefore $\mathbf{P}_z[\tau_1 < \infty] = 1$ for any $z \in B \otimes \mathbf{R}$, and we can apply Lemma 1.

Let us consider a Markov chain $Z' = (Z'_n)_{n \geqq 0}$, where $Z'_n = (U'_n, Y'_n)$, $U'_n = U_{\tau_n}$, $Y'_n = Y_{\tau_n}$, $n \geqq 0$. If $Z_0 = z$, then for any $z \in X$ with $n \geqq 1$, the values of Z'_n belong to $\{e\} \cup D$. Let us denote by $\mathcal{P}'$ the transition operator of Markov chain Z'.

Lemma 1 implies the following key lemma which shows how the factorization (16) is changed when set $C = (B \otimes \mathbf{R}) \setminus D$ is "eliminated" .

LEMMA 2. *Markov chain Z' is defined by a strictly substochastic matrix P' and an operator $\mathcal{L}'$, where*

$$P' = [I - PF_d(C)]^{-1} PF_d(D), \tag{19}$$

$$\mathcal{L}' f(i, y) = (F(i, D_i))^{-1} \int_{D_i} f(i, v)\, F(i, dv), \quad \textit{если } F(i, D_i) \neq 0, \tag{20}$$

and if $F(i, D_i) = 0$, then $\mathcal{L}' f(i, y)$ can be defined arbitrarily with matrix $F_d(A)$ defined in (3) *for any set A.*

Proof. From now on, the vector functions with constant coordinates will be denoted by bold font. Formula (15) then implies that $\mathcal{L}\mathbf{f} = \mathbf{f}$. Formulas (15), (18) and (3) imply that $\mathcal{L}_A \mathbf{f} = F_d(A)\mathbf{f}$. Hence, using also (18) we obtain that $\mathcal{P}_A \mathbf{f} = PF_d(A)\mathbf{f}$, and therefore $(\mathcal{P}_A)^l \mathbf{f} = (PF_d(A))^l \mathbf{f}$ for any $l \geqq 1$. From the last equality with $A = C$ and equality (18) with $A = D$, taking into account that $\mathcal{P}_D f$ is a vector function with constant coordinates, we obtain that $(\mathcal{P}_C)^l \mathcal{P}_D f = (PF_d(C))^l P \mathcal{L}_D f$. Substituting this equation into the right side of the first equality in (12), we obtain

$$\mathcal{P}' f = \widetilde{P}' \mathcal{L}_D f, \quad \text{где } \widetilde{P}' = [I - PF_d(C)]^{-1} P. \tag{21}$$

Now, normalizing measures, defining operator $\mathcal{L}_D$, we obtain (19) и (20). If $F(i, D_i) = 0$ for some i, then corresponding measure $F'(i, \cdot)$ can be defined arbitrarily. Lemma 2 is proved.

Now let us prove the main theorem. As in [2], we construct a sequence of Markov chains Z^k and a nondecreasing sequence of sets D^k, $k \geqq 0$, such that an initial state of each Markov chain is the same as the initial state for the initial Markov chain, and the state of Markov chain Z^k (for $k \geqq 1$) at any moment $r > 0$ coincides with the position of the initial Markov chain at the moment of its r-th visit into set D^k. Under this construction, the sequence of functions $V^k(z)$, $k \geqq 0$, where $V^k(z)$ is an expected reward of stopping at the moment of the first visit of Markov chain Z^k into set D^k, will converge to the value function of the initial Markov chain, and a sequence of sets D^k will converge to an optimal stopping set of the initial Markov chain.

Let $Z^0 = Z$ and $\mathcal{P}^0 = \mathcal{P}$ be a transition operator of Markov chain Z^0. Let $d^1 = \mathcal{P}^0 g$. Note that d^1 is a vector function defined for all $z \in B \otimes \mathbf{R}$, with constant coordinates. We may consider d^1 simply as a real valued vector with corresponding coordinates. Let $D^1 = \{g(z) \geqq d^1\}$. If $F_d(D^1) = I$, we define $Z^1 = Z^0$.

If $F_d(D^1) \neq I$, then it is better to continue at the states of set $C^1 = (B \otimes \mathbf{R}) \setminus D^1$, since one more step brings higher expected reward than immediate stopping. As a result, we "eliminate" this set and consider Markov chain Z^1, similar to Markov chain Z' from Lemma 2 with $D = D^1$. After that, applying the same procedure to Markov chain Z^1 as was applied to Markov chain Z^0, we obtain Markov chain Z^2, and so on. As a result, we obtain a sequence of Markov chains $(Z^k)_{k\geqq 0}$, $Z_0^k = Z_0^0$, corresponding sequence of transition operators $(\mathcal{P}^k)_{k\geqq 0}$, and a sequence of vectors $(d^k)_{k\geqq 1}$, where

$$d^k = \mathcal{P}^{k-1} g, \qquad k \geqq 1, \tag{22}$$

and a sequence of sets $(D^k)_{k\geqq 1}$, where

$$D^k = \{z\colon\ g(z) \geqq d^k\}, \qquad k \geqq 1. \tag{23}$$

For $k \geqq 1$, let us consider a seuence of vector functions

$$V^k(z) = g(z), \quad z \in D^k, \qquad V^k(z) = d^k, \quad z \in C^k = (B \otimes \mathbf{R}) \setminus D^k; \tag{24}$$

Lemma 3. a) *sequence* $(V^k(z))_{k\geqq 1}$ *is nondecreasing and converges to function* $V(z)$, *which is the value function in the OS problem for any Markov chain* Z^k, $k \geqq 0$*;*

b) *sequence of vectors* $(d^k)_{k\geqq 1}$ *is nondecreasing, bounded and converges to vector* d^*. *Sequence of sets* $(D^k)_{k\geqq 1}$ *is nondecreasing and converges to set* D^*, *such that* $\{e\} \cup D^*$ *is an optimal stopping set for any Markov chain* Z^k, $k \geqq 0$, *and Markov chain* Z^*, *corresponding to the visits of Markov chain* Z^0 *into set* D^*. *If* $d^{l+1} = d^l$ *for some* l, *then* $d^k = d^*$, $D^k = D^*$ *for* $k \geqq l$*;*

c) *transition operator of Markov chain* Z^k *for* $k \geqq 1$ *and for* $k = *$ *has the form*

$$\mathcal{P}^k f = \widetilde{P}^k \mathcal{L}_{D^k} f, \quad \textit{where } \widetilde{P}^k = [I - PF_d(C^k)]^{-1} P, \tag{25}$$

so that Markov chain Z^k *is defined by matrix* $P^k = \widetilde{P}^k F(D^k)$ *and by operator* $\mathcal{L}^k$, *which, if* $[F(D^k)]^{-1}$ *exists, has the form* $\mathcal{L}^k = [F(D^k)]^{-1} \mathcal{L}_{D^k}$, *and in general case is defined according to Lemma* 2*;*

d) *let* $d^* = (d_1^*, \dots, d_m^*)$, *and let* $G(i, \cdot)$, $i \in B$, *be the distributions of* $g(i, \widehat{Y}_i)$, *where* $\widehat{Y}_i$ *has distribution* $F(i, \cdot)$. *If there is an* $\varepsilon > 0$, *such that* $G(i, \{(-\varepsilon + d_i^*, d_i^*)\}) = 0$ *for all* $i \in B$, *then there is a* k^*, *such that* $Z^s = Z^{k^*}$ *for* $s \geqq k^*$.

Proof. From (22), (23) and Corollary 1 we obtain that

$$d^{k+1} \geqq d^k, \qquad k \geqq 1, \tag{26}$$

and therefore (23) implies that

$$D^{k+1} \subseteq D^k, \qquad k \geqq 1. \tag{27}$$

Note that Markov chain Z^k was obtained from Markov chain Z^{k-1} by "elimination" of set C^k, i.e. the values of Z^k at moment $n \geqq 1$ coincide with the values of Z^{k-1} at the moment of n-th visit of Markov chain Z^{k-1} into set D^k. Formula (27) implies that the same Markov chain will be obtained from Markov chain Z^0, if set C^k is "eliminated" at once, and then in correspondence with (21)

$$\mathcal{P}^k f = [I - PF_d(C^k)]^{-1} P \mathcal{L}_{D^k} f. \tag{28}$$

The strict substochasticity of matrix P implies that the elements of matrix $[I - PF_d(C^k)]^{-1}$ are bounded. According to our initial assumptions, $\int_{\mathbf{R}} g(i, v)\, F(i, dv)$ is well defined and finite for every $i \in B$. Then, using also (28) and (22), we obtain that the sequence of vectors $(d^k)_{k\geqq 1}$ is bounded. By (26) this sequence will converge to some vector d^*. Therefore, by (23) и (27), the sequence of sets $(D^k)_{k\geqq 1}$ will converge to some set D^*, and then, by (23), the sequence of matrices $(P^k)_{k\geqq 1}$ will converge to the matrix $P^* = [I - PF_d(C^*)]^{-1}P$, where $C^* = (B \otimes \mathbf{R}) \setminus D^*$. Markov chain, corresponding to $\mathcal{P}^*$, is obtained by the "elimination" of C^* for any of Markov chains Z^k, $k \geqq 0$. If $d^{l+1} = d^l$ for some l, then formulas (23), (22), and (25) and the convergence we obtained, imply that $d^k = d^*$, $D^k = D^*$ for $k \geqq l$.

Note now that $V^k(z)$, $k \geqq 1$, is the expected reward from stopping at the moment of the first visit of any of Markov chains Z^s, $0 \leqq s \leqq k$, into set D^k. Hence $V^k(z) \leqq V(z)$, $k \geqq 1$. This inequality and formulas (24), (22) and (26) imply that $V^k(z)$ converges to some function $V^*(z)$, for which $V^*(z) \leqq V(z)$. Formulas (23) and (24) imply that $V^k(z) = g(z) \geqq d^k = \mathcal{P}^{k-1}g(z)$ for $z \in D^k$ and $V^k(z) = d^k = \mathcal{P}^{k-1}g(z) > g(z)$ for $z \in C^k$. Passing to the limit in these relations, we obtain that function $V^*(z)$ satisfies optimality equation (9). But according to Statement 1, $V(z)$ is a minimal solution of the optimality equation. Then $V^*(z) \leqq V(z)$ implies that $V^*(z) = V(z)$. The points a), b) and c) of the theorem are proved. Let us prove point d). The assumption of this point and convergence of d^k to d^* imply that there exists a k^*, such that $d^k > -\varepsilon I + d^*$ for all $k \geqq k^*$. Then $\mathcal{L}_{D^k} = \mathcal{L}_{D^*}$ for all $k \geqq k^*$.Therefore $d^k = d^*$, $D^k = D^*$ for all $k \geqq k^*$. Lemma 3 is proved.

Lemma 3 directly implies all points of the Theorem.

The relations (22)–(25) provide an effective algorithm to calculate the value function and an optimal stopping set in a finite number of steps if the distributions of all random variables $g(i, \widehat{Y}_i)$ do not have finite limit points. We demonstrate that using a specific example. Note that at each step from k to $k+1$ we can "eliminate" not all the states where $d^k \leqq g(z) < d^{k+1}$, but only some of them.

Example. Let us consider Problem 3, where set B consists of only two elements. If Markov chain $\widetilde{U}$ is in state 1, a fair six-side die is rolled, so the distribution $F(1, \cdot)$ is uniform on the set $\{1, 2, 3, 4, 5, 6\}$, and if $\widetilde{U}$ is in a state 2, then a regular tetrahedron is rolled, so the distribution $F(2, \cdot)$ is uniform on the set $\{1, 2, 3, 4\}$. The discount factor β and probabilities p_i that Markov chain $\widetilde{U}$ remains in a state i, are given. Correspondingly, the probabilities of transition from state i to state $3 - i$, are equal to $q_i = 1 - p_i$, $i = 1, 2$. Thus, Markov chain Z is defined by the matrix $P = \begin{bmatrix} p_1\beta, & q_1\beta \\ q_2\beta, & p_2\beta \end{bmatrix} = \beta\widetilde{P}$. Let us consider an OS problem with $g(1, y) = g(2, y) = y$. For integers k_1, k_2, $0 \leqq k_1 \leqq 6$, $0 \leqq k_2 \leqq 4$, let us define $D(k_1, k_2) = \{z = (i, j)\colon i = 1,\ k_1 < j \leqq 6;\ i = 2,\ k_2 < j \leqq 4\}$. To describe operator $\mathcal{P}'$, specified according to Lemma 2 by set $D = D(k_1, k_2)$, we denote $d(k_1, k_2) = \begin{bmatrix} d_1(k_1, k_2) \\ d_2(k_1, k_2) \end{bmatrix} = \mathcal{P}'g$. In this case, formula (21) has the form $P = \beta \begin{bmatrix} p_1, & q_1 \\ q_2, & p_2 \end{bmatrix}$, $F_d(C) = \begin{bmatrix} k_1/6, & 0 \\ 0, & k_2/4 \end{bmatrix}$, $\mathcal{L}_D g = \begin{bmatrix} 3.5 - a_{k_1} \\ 2.5 - a_{k_2} \end{bmatrix}$, where $a_{k_1} = k_1(k_1+1)/12$, $a_{k_2} = k_2(k_2+1)/8$.

Using the explicit formula to invert the 2×2 matrix from formula (21), and using simple algebraic calculations, we obtain

$$[I - PF_d(C)]^{-1} = \frac{1}{\gamma} \begin{bmatrix} 1 - \beta p_2 k_2/4, & \beta q_1 k_2/4 \\ \beta q_2 k_1/6, & 1 - \beta p_1 k_1/6 \end{bmatrix},$$

where $\gamma \equiv \gamma(k_1, k_2) = 1 - \beta(p_1 k_1/6 + p_2 k_2/4) + \beta^2(p_1 - q_2)k_1 k_2/24$ is the determinant of matrix $I - PF_d(C)$. Then, using (21) and the definition of vector $d(k_1, k_2)$, and by multiplying the matrices, we obtain that

$$d_1(k_1,k_2)=\frac{\beta}{\gamma}\Big(2.5-a_{k_2}+p_1(1+a_{k_2}-a_{k_1})+\beta(q_2-p_1)\,\frac{k_2}{4}\,(3.5-a_{k_1})\Big),$$

$$d_2(k_1,k_2)=\frac{\beta}{\gamma}\Big(2.5-a_{k_2}+q_2(1+a_{k_2}-a_{k_1})+\beta(q_2-p_1)\,\frac{k_1}{6}\,(2.5-a_{k_2})\Big).$$

Let $[a]$ denote the integer part of a number a. Then, by formulas (22)–(24), we have: $d^1\equiv(d_1^1,d_2^1)=d(0,0)$, $d^{i+1}\equiv(d_1^{i+1},d_2^{i+1})=d(k_1^i,k_2^i)$, where $k_1^i=[d_1^i]$, $k_2^i=[d_2^i]$. We present below the calculations for certain particular values of the parameters.

N	β	p_1	p_2	d^1	d^2	d^3	d^4	d^5
1	0.7	0.60	0.30	2.17, 2.24	2.44, 2.50	2.44, 2.50		
2	0.8	0.60	0.30	2.48, 2.56	2.96, 3.01	2.957, 3.014	2.957, 3.014	
3	0.8	0.64	0.20	2.51, 2.64	2.99, 3.09	3.005, 3.092	3.006, 3.093	3.006, 3.093
4	0.9	0.80	0.05	2.97, 3.11	3.76, 3.81	3.938, 4.001	3.938, 4.001	3.938, 4.001

The last statement of point b) of Lemma 3 implies the following. For case 1, the optimal stopping set is found in one step, and in both states it is optimal to stop if number three or greater is observed. For case 2, the optimal stopping set is found in two steps, and if a die is rolled, it is optimal to stop if number three or greater is observed, whereas if a tetrahedron is rolled, then it is optimal to stop if the number four is observed. For case 3 the optimal stopping set is found in three steps and if a die or a tetrahedron is rolled, then it is optimal to stop if number four or greater is observed. For case 4, the optimal stopping set is found in three steps, and if a die is rolled, it is optimal to stop if the number four or greater is observed, whereas if a tetrahedron is rolled then it is optimal to roll no matter what number is observed.

REFERENCES

[1] *Dynkin E. B.*, *Yushkevich A. A.* Markov processes: Theorems and problems. New York: Plenum Press, 1969.

[2] *Presman E. L.* On Sonin's algorithm for solution of the optimal stopping problem, Proceedings of the Fourth international conference on control problems (January 26—30, 2009). — M.: Institute of Control Sciences, 2009, pp. 300–309.

[3] *Chow Y. S.*, *Robbins H.*, *Sigmund D.* Great Expectations: The Theory of Optimal Stopping. Boston: Houghton Mifflin, 1971, 141 p.

[4] *Shiryayev A. N.* Optimal Stopping Rules, New York – Heidelberg: Springer-Verlag, 1978. (Russian editions published by "Nauka": 1969 (first ed.), 1976 (second ed.).)

[5] *Çinlar E.* Introduction to Stochastic Processes. Englewood Cliffs: Prentice-Hall, 1975, 402 p.

[6] *Ferguson T. S.* Optimal Stopping and Applications. http://www.math.ucla.edu/~tom/

[7] *Peskir G.*, *Shiryaev A. N.* Optimal Stopping and Free-Boundary Problems. Basel: Birkhäuser, 2006, 500 p.

[8] *Puterman M. L.* Markov Decision Processes: Discrete Stochastic Dynamic Programming. New York: Wiley, 1994, 649 p.

[9] *Sonin I. M.* Two simple theorems in the problems of optimal stopping. — Proceedings of the 8th INFORMS Appl. Probab. Conf., Atlanta, Georgia, 1995, p. 27.

[10] *Sonin I. M.* The elimination algorithm for the problem of optimal stopping. — Math. Methods Oper. Res., 1999, v. 49, N 1, p. 111-123.

[11] *Sonin I. M.* The state reduction and related algorithms and their applications to the study of Markov chains, graph theory, and the optimal stopping problem. — Adv. Math., 1999, v. 145, N 2, p. 159–188.

[12] *Sonin I. M.* The optimal stopping of "seasonal" observations. — Proceedings of the 11th INFORMS Appl. Probab. Conf., New York, 2001, p. 28.

[13] *Sonin I. M.* The optimal stopping of a Markov chain and recursive solution of Poisson and Bellman equations. — From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift. Berlin: Springer-Verlag, 2006, p. 609-621.