Optimal Stopping of Seasonal Observations and Projection of a Markov Chain

Isaac M. Sonin

Abstract

We consider the recently solved problem of Optimal stopping of Seasonal Observations and its more general version. Informally, there is a finite number of dice, each for a state of "underlying" finite MC. If this MC is in a state k, then k-th die is tossed. A Decision Maker (DM) observes both MC and the value of a die, and at each moment of discrete time can either continue observations or to stop and obtain a discounted reward. The goal of a DM is to maximize the total expected discounted reward. This problem belongs to an important class of stochastic optimization problems - the problem of optimal stopping of Markov chains (MCs). The solution was obtained via an algorithm which is based on the general, so called, State Elimination algorithm developed by the author earlier. An important role in the solution is played by the relationship between the fundamental matrix of a transient MC in the "large" state space and the fundamental matrix for the modified underlying transient MC. In this paper such relationship is presented in a transparent way using the general concept of a projection of a Markov model. The general relationship between two fundamental matrices is obtained and used to clarify the solution of the optimal stopping problem.

Keywords Markov chain. Optimal Stopping. Elimination algorithm. Seasonal observations

Mathematical Subject Classification (2010) 60G42, 60J10, 82B35

1 Introduction

The problem described below was formulated in [7] and dubbed as Optimal stopping of *Seasonal Observations*. The solution was published recently in [5]. The goal of this note is to introduce the notion of a *projection* of a Markov chain (MC), which is of interest in its own right, and using this concept to obtain one of the key equalities in [5] in a more general form.

Seasonal Observations. Suppose that (U_n), $n \geq 0$ is a MC with values in a finite set $B = \{1, 2, ..., m\}$ and known transition matrix $U = \{u(s, k), s, k \in B\}$. Suppose that there are m different "dice", each die for a state in B, and the probability that k-th die takes value $j \in Z = \{1, 2, ...\}$ is $f(j|k), k \in B, j \in Z$. If at the moment n the MC (U_n) takes value k, then the k-th die is tossed and a Decision Maker (DM) observes both U and the value j obtained. At each moment $n = 0, 1, 2, ...$ a DM can either continue observations or to stop and obtain a discounted reward $\beta^n g(k, j)$, where β is a discount factor, $0 < \beta \leq 1$, and $g(k, j)$ is the *terminal reward function.* The goal of a DM is to maximize the total expected discounted reward. This problem can be generalized if one introduces a *one step cost function* $c(k)$, but for simplicity we assume that $c(k) = 0$ for all k. Formally, we assume that a DM observes MC (Z_n) with values in $X = B \times Z$ and with transition probabilities $p(x, y) \equiv p(s, i; k, j) = u(s, k)f(j|k), s, k \in B, i, j \in Z$. Thus, these probabilities depend only on the first "horizontal" coordinate of a state $x = (s, i)$. We can represent this relationship symbolically by the "factorization equality"

$$P = U \times F, \tag{1}$$

where U is $m \times m$ stochastic matrix and $F = \{f(\cdot|k), k \in B\}$ is a vector of distributions on Z.

2 Optimal Stopping of MC

The problem described above belongs to an important class of stochastic optimization problems - the *problem of optimal stopping (OS) of MC*, where a DM observing a MC, has two possible actions at each moment of discrete time: to continue observations or to stop, and then to obtain a terminal reward. Formally, such a problem is specified by a tuple $M = (X, P, c, g, \beta)$, where X is a state space, $P = \{p(x, y)\}$ is a transition matrix, $c(x)$ is a *one step cost function*, $g(x)$ is a *terminal reward function*, and β is a discount factor, $0 < \beta \leq 1$. We call such a model *OS model* and a tuple $M = (X, P)$, we call a *Markov model.* The *value function* $v(x)$ for OS model is defined as $v(x) = \sup_{\tau \geq 0} E_x[\sum_{i=0}^{\tau-1} \beta^i c(Z_i) + \beta^\tau g(Z_\tau)]$, where the sup is taken over all stopping times $\tau \leq \infty$. To simplify our presentation we will assume that $c(x) = 0$ and $v(x) < \infty$ for all x.

It is well-known that in stochastic optimization problems the discounted case can be treated as undiscounted if an absorbing point e is introduced and the transition probabilities are modified as follows, $p^\beta(x, y) = \beta p(x, y)$ for $x, y \in X$, $p^\beta(x, e) = 1 - \beta, p^\beta(e, e) = 1$. In other words, with probability β the Markov chain "survives" and with complimentary probability it transits to an absorbing state e. More than that, for our method it is convenient and important to consider a more general situation when the constant β can be replaced by the probability of "survival", function $\beta(x) = P_x(Z_1 \neq e), 0 \leq \beta(x) \leq 1$. Further we will assume that this transformation is made and we skip the superscript β, using again notation P_x and E_x.

Let $Pf(x)$ be the *averaging operator,* $Pf(x) = \sum_y p(x,y)f(y)$. It is well-known that the value function v is a minimal solution of a corresponding *Bellman (optimality) equation* $v = \max(g, c + Pv)$. Let $A \subset B \times Z$, i.e. $A = \{A(k)\}, A(k) \subset Z, k \in B$ and let us denote by $F(A(k)|k) = \sum_{j \in A(k)} f(j|k)$ and by $F_d(A)$ the $m \times m$ diagonal matrix $F_d(A) = (\delta_{sk} F(A(k)|k)),\ s,k \in B$. The complement of a set D_* is denoted by S_*. The following theorem was proved in [5].

Theorem 1. *There is a vector* $d^* = (d_1^*, \ldots, d_m^*)$, *such that*

a) an optimal stopping time τ^* *is the moment of first visit of the Markov chain* Z *to the set* $\{e\} \cup S^*$, *where*

$$S^* = \{z = (k,j) : k \in B,\ j \in S^*(k)\},\ S^*(k) = \{j : g(k,j) \geq d_k^*\};$$

b) the value function satisfies the equation

$$v(x) = g(x), x \in S^*,\ \ v(x) = d_k^* > g(k,j), x = (k,j) \in D^* = X \diagdown S^*, \tag{2}$$

and d^* *satisfies the equation*

$$d_s^* = \sum_{k \in B} l^*(s,k) \sum_{j \in D^*(k)} g(k,j)\, f(j|k), \tag{3}$$

where the matrix $L^* = \{l^*(s,k),\ s,k \in B\}$ *is defined by the equality*

$$L^* = [I - UF_d(D^*)]^{-1} U. \tag{4}$$

The proof of Theorem 1 is obtained via an algorithm which allows one to find the vector d^*, and, therefore, to construct the value function and the optimal stopping set in a finite number of steps. This algorithm is based on the general, so called, State Elimination (SE) algorithm developed by the author earlier and described in [8] (see also [9]). This algorithm has some features in common with the so called State Reduction (SR) approach used in computational MCs and which is exemplified by works of Grassmann, Taksar, Heyman [1] and Sheskin [6], who independently developed GTH/S algorithm to calculate the invariant distribution for an ergodic MC. The explanation of this approach is given in [9]. We first briefly describe this approach and afterwards we explain the SE algorithm. Our notations in these sections are slightly different than those used in the original author's papers.

3 Recursive Calculation of Characteristics of MC and the State Reduction (SR) Approach

Let us assume that a *Markov model* $M = (X, P)$ is given and let $D \subset X$, $S = X \setminus D$. Then the matrix $P = \{p(x, y)\}$ can be decomposed as the first matrix below

$$P = \begin{bmatrix} Q & T \\ R & P_0 \end{bmatrix}, \qquad P'_S = \begin{bmatrix} 0 & NT \\ 0 & P_S \end{bmatrix} \tag{5}$$

where the substochastic matrix Q describes the transitions inside of D, P_0 describes the transitions inside of S and so on. Let us introduce the sequence of Markov times $\tau_0, \tau_1, ..., \tau_n, ...,$ the moments of zero, first, and so on, return of (Z_n) to the set S, i.e., $\tau_0 = 0$, $\tau_{n+1} = \min\{k > \tau_n, Z_k \in S\}$. Let us consider the random sequence $Y_n = Z_{\tau_n}, n = 0, 1, 2, ..., Z_0 \in S$. The strong Markov property and standard probabilistic reasoning imply the following basic lemma of the SR approach which probably should be credited to Kolmogorov and Doeblin.

Lemma 1. *(a) The random sequence* (Y_n) *is a Markov chain in a model* $M_S = (S, P_S)$*, where* $S = X \setminus D$ *and*

(b) the transition matrix $P_S = \{p_S(x, y), x, y \in S\}$ *is given by the formula*

$$P_S = P_0 + RV = P_0 + RN_D T, \tag{6}$$

where $V = N_D T$ *is a matrix of the distribution of the MC at the moment of first return to* S*, and* $N_D = N$ *is the fundamental matrix for the substochastic matrix* $Q = \{p(x, y), x, y \in D\}$.

We remind that $N = \sum_{n=0}^{\infty} Q^n = (I - Q)^{-1}$, where I is the $|D| \times |D|$ identity matrix. This representation is proved, for example, in the classical text of Kemeny and Snell, [3]. This matrix N satisfies also the equality

$$N = I + QN = I + NQ. \tag{7}$$

An important case is when the set D consists of one nonabsorbing point z. In this case formula (6) takes the form

$$p_S(x, \cdot) = p(x, \cdot) + p(x, z)n(z)p(z, \cdot), \tag{8}$$

where $n(z) = 1/(1 - p(z,z))$. According to this formula, each row-vector of the new stochastic matrix P_S is a linear combination of two rows of P (with the z-column deleted). This transformation corresponds formally to one step of the Gaussian elimination method. This matrix P_S describes the behaviour of MC with values in a set S, or we can extend this matrix to the full size $X \times X$ matrix P'_S, see the second matrix in (5), assuming that MC (Y_n) can have an initial point in set D also. But in both cases, to obtain the matrix P_S, we need to study the behaviour of the related transient MC with values in D.

The matrix N, a fundamental matrix for this transient MC with transition matrix Q, has the following well known probabilistic interpretation, $N = \{n(x,y), x, y \in D\}, n(x,y) = E_x \sum_{n=0}^{\tau_S} I_y(Z_n)$, where τ_S is the moment of the first visit to S, i.e. $\tau_S = \min(n \geq 0 : x_n \in S)$ (moment of first exit from D), i.e. the expected number of visits to y starting from x till τ_S. In this case, i.e. when the transition matrix P is changed in such a way that S become an absorbing set, we shall say that MC (Z_n) is *stopped* at $S = X \setminus D$, and we shall denote this new MC as (Z_n^D).

The recursive calculation of the second fundamental matrix, for the ergodic MC was described in [10].

If an initial Markov model $M_1 = (X_1, P_1)$, is finite, $|X_1| = k$, and only one point is eliminated each time, then a sequence of stochastic matrices $(P_n), n = 2, ..., k$, can be calculated recursively on the basis of formula (8). Generally, a set of points D can be eliminated using formula (6). In both cases such sequence of stochastic matrices provides an opportunity to calculate many characteristics of the initial Markov model M_1 recursively starting from some reduced model $M_s, 1 < s \leq k$.

4 State Elimination (SE) Algorithm

In this section we describe briefly the SE algorithm (for the case of $c(x) = 0$). Let an OS model $M = (X, P, g)$, be given, and suppose that an optimal stopping set $S^* = \{x : v(x) = g(x)\}$ does exists. Let a subset $D \subset \{x : g(x) < Pg(x)\}$. Since $g(x) \leq v(x)$, and $Pg(x) \leq Pv(x)$ the optimality equation implies that $D \cap S^* = \emptyset$. It was proved in [8] that the optimal stopping set in the reduced OS model $M_S = (X_S = X \setminus D, P_S, g)$ will be the same as in the initial OS model and the value functions will be the same for all points in X_S. After that we can repeat the process by eliminating points in a set $D' \subset \{x : g(x) - P_S g(x) < 0\}$ and so on. If at some stage after k steps, with $D_1 = D, D_2 = D' \cup D_1$ and so on, we obtain that $g(x) - P_{S_k} g(x) \geq 0$ for all remaining points, then $S^* = S_k = X \setminus D_k$. For the finite space X this algorithm solves the OS problem in no more than $|X|$ steps, and allows us also to find the distribution of the MC at the moment of stopping in an optimal stopping set S^*. Recently E. Presman modified this idea and applied to the case of OS in continuous time, see [4].

5 Projection of MC and Seasonal Observations

Note that the matrix $[I - UF_d(D^*)]^{-1}$ from formula (4) is the *fundamental* matrix for the transient MC obtained from the underlying MC (U_n) by modifying its transition matrix U. An important role in the proof of Theorem 1 is played by the relationship between the fundamental matrix of a transient MC in the state space X and the fundamental matrix for the modified transient MC in the state space B. This relationship can be presented in a transparent way using the concept of *projection of a Markov model,* and, correspondingly of projection of a MC.

Let $M_i = (X_i,\ P_i)$ be two Markov models, $i = 1, 2$ and let $h : X_1 \longrightarrow X_2$ be a mapping. If (Z_n) is a MC in model M_1 then generally random sequence $(U_n), U_n = h(Z_n)$ is not a MC in model M_2. In [2] Howard introduced a notion of a "mergeable" Markov chain when the random sequence (U_n) is a MC. In terms of two models, model M_1 is mergeable if the transitional probabilities in these models for any $x, x' \in h^{-1}(s) \subset X_1$ and any $s, k \in X_2$ satisfy the following equality: $\sum_{y \in h^{-1}(k)} p_1(x, y) = \sum_{y \in h^{-1}(k)} p_1(x', y)$. If these two Markov models have terminal reward functions $g_1(x), x \in X_1$, $g_2(k), k \in X_2$ and terminal reward function g_1 is also "mergeable", i.e. if $g_1(x) = g_2(h(x))$ for all $x \in h^{-1}(k), k \in X_2$ then of course the solution of the OS in M_1 can be reduced to the solution in M_2, but this is a trivial situation. To be able to consider the OS problem for the seasonal observations we need a stronger assumption.

We say that model M_2 is a *projection* of a model M_1 (under h) if the transitional probabilities in these models satisfy the following property for all $x, y \in X_1$,

$$p_1(x, y) = p_2(h(x), h(y)) f_1(y|h(y)), \tag{9}$$

where $f_1(y|t)$ is a probability distribution on a set $h^{-1}(t) = \{y \in X_1 : h(y) = t\}$, defined for each $t \in X_2$. In other words, the state space X_1 is partitioned into classes $T_t = h^{-1}(t)$, $t \in X_2$ and transitions from state x in class T_s to state y in T_k depend only on s, k and y but not on x. A reader may think about model M_1 as a "large", basic model and about model M_2 as a "small", more manageable model. It is easy to check that if (Z_n) is a MC in model M_1 then the random sequence $(U_n), U_n = h(Z_n)$ is a MC in model M_2.

To simplify our presentation we will assume that sets X_1 and X_2 are discrete and that Markov model M_2 has an absorbing state e. Let $|X_2| = m + 1$, where $m, m \leq \infty$ is the number of proper states, i.e. $x \neq e$. Let the set $D \subset X_1$, $S = X_1 \setminus D$. We consider MC (Z_n^D) stopped at $S = X \setminus D$. According to the SE algorithm if a set D should be eliminated then in order to find the matrix P_S by formula (6), we have to find the fundamental matrix $N_{1,D} = \{n_{1,D}(x, y), x, y \in D\}$.

To accomplish this goal we will introduce MC (U_n^D) in model M_2, the "projection"of MC (Z_n^D), defined by the equality $U_n^D = h_D(Z_n^D)$, where function $h_D(x) = h(x)$ if $x \in D$ and $h_D(X) = e$ if $x \in X_1 \setminus D$. In Theorem 2 we will relate the fundamental matrix $N_{2,D}$ for this MC with the matrix $N_{1,D}$.

If P is an $m \times m$ stochastic matrix, $D \subset X_1$ and $F_d(D)$ is the $m \times m$ diagonal matrix with elements $F(D(k)) = \sum_{j \in D(k)} f_1(j|k)$, $D(k) = D \cap h^{-1}(k)$, then we denote substochastic matrix $P_D^F = PF_d(D)$ and we denote the fundamental matrix for P_D^F as $(I - P_D^F)^{-1} = \sum_{n=0}^{\infty} (P_D^F)^n$.

We denote $M_{2,D} = (X_2, P^*_{2,D})$ a Markov model obtained from model M_2 as follows. The state space is the same, X_2 and the transition matrix $P^*_{2,D} = P_2 F^*(D)$, where $F^*(D)$ is the $(m + 1) \times (m + 1)$ matrix, which has in the upper left corner the $m \times m$ diagonal matrix $F_d(D)$ described above, and the last column of matrix $F^*(D)$ contains entries $f^*(s, e) = 1 - \sum_k p_2(s, k)F(D(k))$, $s \neq e$, $f^*(e, e) = 1$. In other words, in this model the transitional probabilities are: $p_{2,D}(s, k) = p_2(s, k)F(D(k))$, for $k \neq e$, and $p_{2,D}(s, e) = p_2(s, e) + \sum_k p_2(s, k)F(S(k))$. We denote the $m \times m$ upper left corner of matrix $P^*_{2,D}$ by $P_{2,D}$. According to the definition of $P^*_{2,D}$, we have $P_{2,D} = P_2 F_d(D) \equiv P_D^F$. This is a substochastic matrix for the transient MC in model $M_{2,D}$ with absorption in e.

Let us consider $N_{2,D} = \{n_{2,D}(s, k), s, k \in X_2, s, k \neq e\}$, the fundamental matrix for $P_{2,D}$. The following theorem holds.

Theorem 2. *If (Z_n) is a Markov chain in model M_1 and $D \subset X_1$ then*

(a) *the random sequence $(U_n), U_n = h(Z_n)$ is a MC in model M_2 with the transition matrix P_2; the random sequence $(U_n^D), U_n^D = h_D(Z_n^D)$ is a MC in model $M_{2,D}$ with the transition matrix (for the proper states) $P_{2,D}^F$ described above;*

(b) *the fundamental matrices in the original and the projected models, $N_{1,D}$ and $N_{2,D}$ are related by the equalities valid for all $x, y \in D \subset X_1, s, k \in X_2, s, k \neq e$,*

$$n_{1,D}(x, y) = n_{2,D}(s, k) f_1(y|k)/F(D(k)), s = h(x), k = h(y); \tag{10}$$

(c) *stochastic matrix $P_{1,S}$ has factorization*

$$P_{1,S} = P_{2,S} \times F_S, \tag{11}$$

where $F_S = \{f_{1,S}(y|k) = f_1(y|k)/F(S(k))\}, k \in X_2$ *and*

$$P_{2,S} = P^F_{2,S} + P^F_{2,D}(I - P^F_{2,D})^{-1}P^F_{2,S} = (I - P^F_{2,D})^{-1}P^F_{2,S}. \qquad (12)$$

Proof. We omit the proof of point (a) which can be obtained using standard probability reasoning. To prove (b) note that by the definition of a fundamental matrix for a MC (Z_n^D) stopped at $S = X_1 \setminus D$, we have $n_{1,D}(x,y) = E_{1,x}\sum_{n=0}^{\infty} I_y(Z_n^D) = \sum_{n=0}^{\infty} P_{1,x}(Z_n^D = y)$. According to (9) we have $P_{1,x}(Z_n^D = y) = P_{2,s}(U_n^D = k)P_{1,x}(Z_n^D = y|h(Z_n^D) = k) = P_{2,s}(U_n^D = k)f_1(y|k)/F(D(k))$. Using the equality $n_{2,D}(s,k) = \sum_{n=0}^{\infty} P_{2,s}(U_n^D = k)$, we obtain (10).

Point (c). Using (6), factorization (1), (9) and notations $x = (s,x'), y = (k,y'), z = (l,z')$ and $v = (t,v')$, we have

$$\begin{aligned} p_{1,S}(x,y) &= p_1(x,y) + \sum_z p_1(x,z)\sum_v n_{1,D}(z,v)p_1(v,y) = \\ &= p_2(s,k)f_1(y'|k) + \sum_l p_2(s,l)\sum_{z'\in D(l)} f_1(z'|l)\sum_v n_{1,D}(z,v)p_2(t,k)f_1(y'|k). \end{aligned}$$

Using point (b), i.e. replacing $n_{1,D}(z,v)$ by $n_{2,D}(l,t)f_1(y'|t)/F_1(D(t))$, and using the equality $\sum_{z'\in D(t)} f_1(z'|t) = F(D(t)), t \in X_2$, we have

$$\sum_{v=(t,v')} n_{1,D}(z,v)p_2(t,k) = \sum_t n_{2,D}(l,t)\sum_{v'\in D(t)}\frac{f_1(v'|t)}{F(D(t))}p_2(t,k) = \sum_t n_{2,D}(l,t)p_2(t,k). \qquad (13)$$

Using the equalities $\sum_{z'\in D(l)} f_1(z'|l) = F(D(l)), l \in X_2$, $P_{2,D} = P_2F_{1,d}(D)$, and (13), we obtain finally

$$p_{1,S}(x,y) = [p_2(s,k)F(S(k)) + \sum_l p_{2,D}(s,l)\sum_t n_{2,D}(l,t)p_2(t,k)F(S(k))]f_1(y'|k)/F(S(k)).$$

The expression in square brackets in matrix notation is $P_{2,S} + P_{2,D}(I - P_{2,D})^{-1}P_{2,S}$, which equals the last term in (12) by the first equality in (7). The expression outside of square brackets corresponds to the term F_S. Theorem 2 is proved.

6 Open problem

Let $M_i = (X_i,\ P_i)$ be two Markov models, $i = 1,2$ and let $h : X_1 \longrightarrow X_2$ be a mapping. An open problem is to find all relationships between the transitional probabilities in these two models such that the solution of the OS problem for the "large" model M_1 can be simplified using the projection model M_2. For example, a potential candidate is the case when the transition probabilities for all $x, y \in X_1$ satisfy

$$p_1(x,y) = p_2(s,k) \sum_{i=1}^{N} \alpha_i(s,k) f_1(y|k,i), \tag{14}$$

where $s = h(x), k = h(y), \alpha_i(s,k) \geq 0, \sum_{i=1}^{N} \alpha_i(s,k) = 1, s, k \in X_2$. In other words, instead of one die for each state of $k \in X_2$, there are sets of N dice, and transitions are defined using randomization over these sets.

The author would like to thank Joe Quinn, Ernst Presman, and an anonymous referee for valuable comments.

References

[1] Grassmann, W. K., Taksar, M., Heyman, D., 1985. Regenerative analysis and steady state distributions for Markov chains. *Oper. Res.* **33**, 5, 1107-1116.

[2] Howard, R., 1971. Dynamic probabilistic systems. Markov Models., John Wiley & Sons, NY-London.

[3] Kemeny, J., Snell L.,(1960, 1983), Finite Markov Chains. Springer-Verlag.

[4] Presman E, 2010. The solution of optimal stopping problem based on a modification of a payoff function, this volume.

[5] Presman, E., Sonin I., 2010. On optimal stopping of random sequences modulated by Markov chain, *Theory of Probability and Its Applications,* **54**, 3, 534-542.

[6] Sheskin, T. 1985. A Markov chain partitioning algorithm for computing steady state probabilities. *Oper. Res.* **33**, 1, 228–235.

[7] Sonin, I., 2001. The optimal stopping of "Seasonal observations". In: *Proc.11th INFORMS Appl. Prob. Conf., N*ew York, 18.

[8] Sonin, I., 1999. The Elimination Algorithm for the Problem of Optimal Stopping, *Math. Meth. of Oper. Res.,* **4**, 1, 111-123.

[9] Sonin, I., 1999. The State Reduction and related algorithms and their applications to the study of Markov chains, graph theory and the Optimal Stopping problem, *Advances in Mathematics*, **145**, 2, 159-188.

[10] Sonin I, Thornton J., 2001. Recursive Algorithm for the Fundamental/Group Inverse Matrix of a Markov Chain from an Explicit Formula, *SIAM J. on Matrix Analysis and Appl.* **23**, 1, 209 - 224.