The Optimal Stopping of Markov Chain and its Application to other Probability Problems

Isaac M. Sonin

Dept. of Mathematics and Statistics
Univ. of North Carolina at Charlotte

Central Economics and Mathematical Institute,
Russian Academy of Science

website: google Isaac Sonin

Chiba University, Japan /Petrozavodsk, Russia/ May, June, 2008

Outline

- (Pure) Probability Theory
- Many-armed Bandit problems (with dependent arms)
- Economic Theory. Games
- Applied Probability: secretary problem, optimal stopping (OS) problems, sufficiency of Markov, stationary strategies in MDP, replacement models, inventory models, *State Elimination Algorithm, Generalized Gittins index*
    - Optimal Stopping of Markov chains
    - State Reduction Approach
    - Generalized Gittins index. Four indices
    - Abstract Optimization Problems
    - Open Problems. Plans

Markov model $M = (X, P), \;\; P = \{p(x, y)\}, \;\; (Z_n), \;\; U_0 = \{(Z_n)\}$

family of all finite Markov chains

The behavior of U_0 is a classical result of Probability Theory derived in the 30's by A. Kolmogorov and W. Doeblin.

If P is replaced by (P_n) then ... U... family of *nonhomogeneous* MCs.

Is it possible to say something about the behavior of the family U if there are *no assumptions* about sequence (P_n) ?

If P is replaced by (P_n) then ... U... family of *nonhomogeneous* MCs.

Is it possible to say something about the behavior of the family U if there are *no assumptions* about sequence (P_n) ?

The surprising answer to this question is Yes. Such behavior is described by a theorem which we call a Decomposition-Separation (DS) Theorem. Simple deterministic interpretation plus martingales.

Net Present value (NPV) and **Internal rate of return (IRR)**

IRR = effective interest rate = yield to maturity

Income stream = *cash flow* = project $a = (a_0, a_1, ..., a_T)$.
If $a_i < 0$ (is negative) it means that an investor pays money, if its positive she receives money. For this cash flow we can calculate NPV for a particular interest rate r. If we set NPV equal to zero, the roots of this equation will give the IRsR

$$NPV(r) = a_0 + \frac{a_1}{1+r} + \frac{a_2}{(1+r)^2} + \dots \frac{a_T}{(1+r)^T} = 0.$$

Investment polynomial $\sum_{0 \le i \le T} a_i z^i = 0,\ z = 1/(1+r)$. Simplest case: you invest a_0 and after moment T you receive a_T, equation $|a_0|(1+r)^T = a_T$. In this case there is exactly one root; in general there may be many roots. The interpretation of multiple roots of IRR is given in my paper *Economic Theory*, 5, (1995). Presman, Sonin (200X)

To stop or not to stop, that is the question !

OS: many probability models, and now the pricing of American options

there are two approaches - " martingale theory of OS " and "Markovian approach"

classical monographs Chow, Robbins and Sigmund (1971)

A. Shiryayev (1969, 1978)

T. Ferguson (website)

G. Peskir, A. Shiryayev (2006)

T. Ferguson: "Most problems of optimal stopping without some form of Markovian structure are essentially untractable...".

Optimal Stopping Problem of Markov Chain (MC)

$M = (X, P, c, g, \beta)$, $P = \{p(x,y)\}$, (Z_n), P_x , E_x, $Pf(x) = \sum_y p(x,y) f(y)$, X finite (countable), $c(x)$ *one step cost* function, $g(x)$ *terminal reward* function, β discount factor, $\beta \leq 1$,

$v(x) = \sup_{\tau \geq 0} E_x[\sum_{i=0}^{\tau-1} \beta^i c(Z_i) + \beta^\tau g(Z_\tau)]$ *value function*

Remark: absorbing state e, $p(x,y) \longrightarrow \beta p(x,y)$, $p(x,e) = 1 - \beta$, $\beta \longrightarrow \beta(x) = P_x(Z_1 \neq e)$ probability of "survival"

Bellman (optimality) equation $v = \max(g, c + Pv) = \max(g, Fv)$

$S = \{x : g(x) = v(x)\}$ optimal stopping set

Theorem 1. (Shiryaev 1969). (a) *The value function* $v(x)$ *is the minimal solution of Bellman equation ...*

(e) *if state space* X *is finite then set* S *is not empty and* $\tau_0 = \min\{n \geq 0 : Z_n \in S\}$ *is an optimal stopping time.*

Basic methods of solving OS of MC

- the direct solution of the Bellman equation
- the value iteration method : one considers the sequence of functions $v_n(x) = \sup_{0 \le \tau \le n} E_x ...,\ v_{n+1}(x) = \max(g(x), Fv_n(x)),$ $v_0(x) = g(x)$. Then $v_0(x) \le v_1(x) \le ...; \quad v_n(x)$ converges to $v(x)$.
- the linear programming approach $(\ |X| < \infty),\ \ \min \sum_{y \in X} v(y),$ $v(x) \ge \sum_y p(x,y)v(y), v(x) \ge g(x), x \in X.$
- Davis and Karatzas (1994), interesting interpretation of the Doob-Meyer decomposition of the Snell‘s envelope
- The State Elimination Algorithm (EA) Sonin (1995, 1999, 2005)

Goal: a unified recursive approach to

Problem 1 To find a minimal solution v of a *Bellman equation*

$$v = \max(g, c + Pv) = \max(g, Fv)$$

Free boundary problem

Problem 2 To find the solution f of the discrete *Poisson equation*

$$f = c + Pf$$

Problem 3 To calculate *Gittins index (GI)* $\gamma(x)$ and the Generalized Gittins index $\alpha(x)$,

$$\gamma(x) = \sup_{\tau > 0} \frac{E_x \sum_{n=o}^{\tau-1} \beta^n c(Z_n)}{E_x \sum_{n=o}^{\tau-1} \beta^n},$$

i.e. $\gamma(x)$ is the maximum of a *expected discounted reward* per *expected discounted unit of time*, $0 < \beta < 1$, τ is a stopping time.

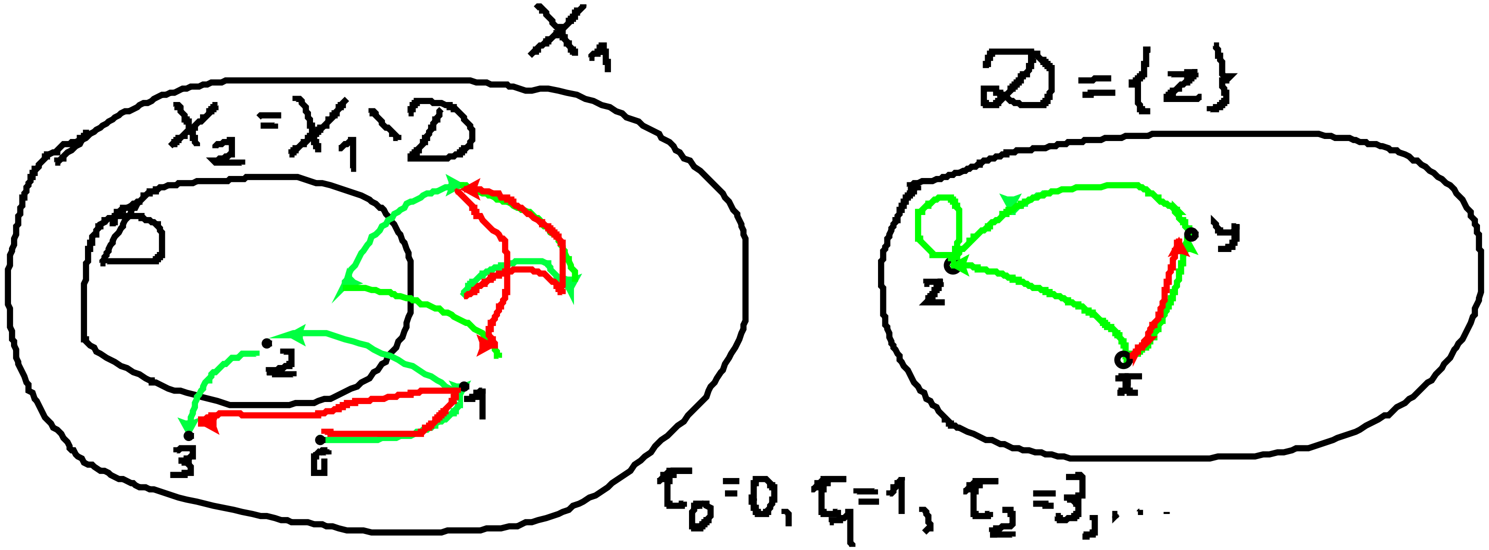
State Reduction (SR) approach

$M_1 = (X_1,\, P_1),\;\; D \subset X_1,\, X_2 = X_1 \setminus D,\, (Z_n), n = 0, 1, 2, \ldots$

$\tau_0, \tau_1, \ldots, \tau_n, \ldots,$ the moments of zero, first, and so on, visits of (Z_n) to the set X_2 : $\tau_0 = 0,\, \ldots\; \tau_{n+1} = \min\{k > \tau_n, Z_k \in X_2\}$,

$(Y_n),\quad Y_n = Z_{\tau_n},\quad n = 0, 1, 2, \ldots$.

$U_1 = \{u_1(x, y)\},\, u_1(x, y) = P_x(Z_{\tau_1} = y), x \in D,\; y \in X_2.$



Lemma 1. (Kolmogorov, Doeblin)

(a) (Y_n) *is a Markov chain in a model* $M_2 = (X_2, P_2)$, *where*

(b) *the transition matrix* $P_2 = \{p_2(x,y)\}$ *is given by the formula*

$$p_2(x,y) = p_1(x,y) + \sum_{z \in D} p_1(x,z) u_1(z,y), \quad (x, y \in X_2).$$

$$P_1 = \begin{bmatrix} Q_1 & T_1 \\ R_1 & P_1' \end{bmatrix}, \quad P_2 = P_1' + R_1 U_1 = P_1' + R_1 N_1 T_1, \tag{1}$$

N_1 is a (transient) *fundamental matrix*, i.e. $N_1 = (I - Q_1)^{-1}$,

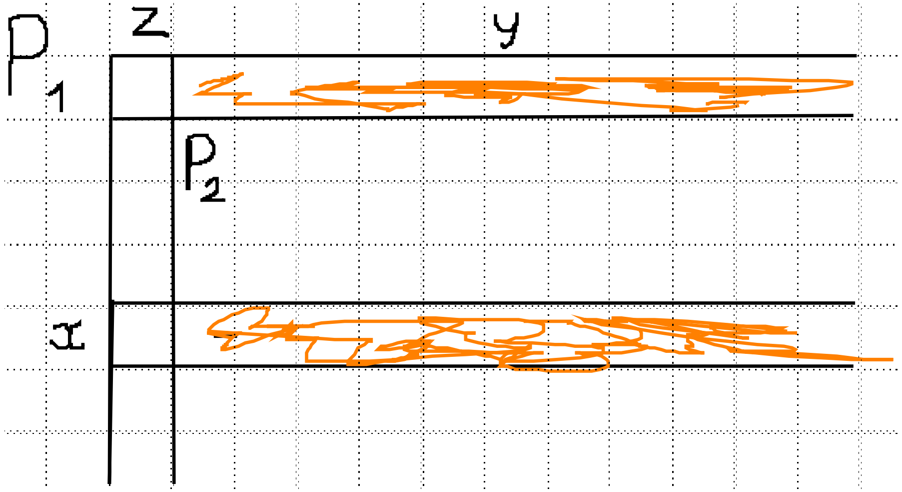
$N = I + Q + Q^2 + \ldots = (I - Q)^{-1}, U_G = \{u(x,s) = P_x(Z_{\tau_G} = s\},$

$U = NT$

If $D = \{z\}$ then

$$p_2(x,\cdot) = p_1(x,\cdot) + p_1(x,z)n_1(z)p_1(z,\cdot)\ , \tag{2}$$

where $n_1(z) = 1 + p_1(z,z) + (p_1(z,z))^2 + \ldots = 1/(1 - p_1(z,z))$.



Lemma 2. *Let* M_1 *and* M_2 *be two adjacent models with state spaces* X_1 *and* $X_2 = X_1 \setminus D$*, where* $D \subseteq X_1, G \subset X_2, N_1, N_2, U_1, U_2$*. Then* $u_1(x,y) = u_2(x,y), \quad y \in G$, $n_1(x,v) = n_2(x,v)$.

$$N_1 = N_2, \qquad\qquad U_1 = U_2$$

$M_1 = (X_1,\, P_1, c_1, g), \quad c_1(x)$ one step cost function transformation

If $D = \{z\}$ then (T.J. Sheskin (1999))

$$c_2(x) = c_1(x) + p_1(x,z)n_1(z)c_1(z), \ \ x \in X_2.$$

For the general set D, in matrix form

$$c_2 = c_{1,X_2} + R_1 N_1 c_{1,D} \tag{3}$$

OS $=$ Bellman equation $v(x) = \max(g(x), c(x) + Pv(x))$

The State Elimination algorithm (SEA): three facts:

1. It may be *difficult* to find the states where it is optimal *to stop*, but it is *easy* to find a state (states) where it is optimal *not to stop.*

stop if $g(x) \geq c(x) + Pv(x)$ but v is unknown, so...

do not stop if $g(z) < c(z) + Pg(z) \leq c(z) + Pv(z))$ or...

2. Then eliminate z, (set D) and recalculate $P_1 \longrightarrow P_2$ and $c_1 \longrightarrow c_2$, g. Lemma 2 implies *Elimination theorem:* $S_1 = S_2, v_1 = v_2$ for $x \in X_2$. After that repeat the first step and so on.

3. Finally, there is a well known

Proposition 1. *Let* $M = (X, P, g)$ *be an optimal stopping problem, and* $g(x) \geq c(x) + Pg(x)$ *for* all $x \in X$. *Then* X *is the optimal stopping set in the problem* M, *and* $v(x) = g(x)$ *for all* $x \in X$.

After the elimination... the stopping sets and the value functions in the remaining points will coincide

SE algorithm

$$\begin{tikzcd}
 & g(x)-(c_1(x)+P_1g(x)) = g - F_1g \arrow[dl] \arrow[dr] & & \\
g(x)-F_1g(x)\geq 0 \ \text{for } \textit{all } x \arrow[d, Rightarrow] & & \text{there is } z : g(z)-F_1g(z)<0 \arrow[d, Rightarrow] & \\
X_1 = S & & M_1 \longrightarrow M_2 : g(x)-F_2g(x) \arrow[dl] \arrow[dr] & \\
 & {} & & {} \\
 & & \dots &
\end{tikzcd}$$

On the 2nd stage we recursively calculate $v(x)$

Sonin (1995); Sonin, I.M.: The Elimination Algorithm for the Problem of Optimal Stopping. *Math. Meth. of Oper. Res.*, **49**, no. 1, (1999)

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. *Advances in Mathematics* **145**, no. 2, (1999)

SR Algorithms 1985, GTH/S

Grassman, Taksar, Heyman; Sheskin

$$M_1 = (X_1, P_1), |X_1| = k,$$

$$P_1 \longrightarrow P_2 \longrightarrow ... \longrightarrow P_k = \{1\}$$

$$\pi_1 \longleftarrow \pi_2 \longleftarrow ... \longleftarrow \pi_k = \{1\}$$

GTH Algorithm to calculate the invariant distribution of MC
GTH/S

1. T. J. Sheskin *Oper. Res.* **33** (1985).

2. W. K. Grassmann, M. I. Taksar, and D. P. Heyman, *Oper. Res.* **33** (1985).

other SR algorithms: FUND Heyman (1995), REFUND Sonin and Thornton (1999, 2001), Kohlas (1986),...

Countable State Space

- Random walk on a line
- Optimal Stopping of ”Seasonal observations”

There are m different ”dice” but which die to observe at moment n is specified by a position of an underlying Markov chain (U_n) independent of observations on dice

Sonin (2001), Presman and Sonin (2008), submitted

- Game: Optimal Stopping of ”Seasonal observations”

  There are m different ”dice” and $N, N \leq m$ players, player i can stop the game when her dice is observed

Recursive Solution of Poisson equation

$$f = c + Pf$$

Lemma 3. *Let* M_1 *and* M_2 *be two adjacent models with state spaces* X_1 *and* $X_2 = X_1 \setminus D$*, where* $D \subseteq X_1$*,* P_i *and* $F_i, i = 1, 2$ *be the corresponding averaging and reward operators, where functions* c_1 *and* c_2 *are related by* (3)*, matrices* R_1, T_1 *are as in* (1) *and matrix* N_1 *is a fundamental matrix for* Q_1*. Let* f *be a function defined on* X_1*. Then*

$$f_{X_2} - F_2 f_{X_2} = (f - F_1 f)_{X_2} + R_1 N_1 (f - F_1 f)_D.$$

$$f_D = N_1[(f - P_1 f)_D + T_1 f_{X_2}]$$

Theorem 2. *Let* M_1 *and* M_2 *be two adjacent models with state spaces* X_1 and $X_2 = X_1 \setminus z$, *where* P_1 *and* P_2, *functions* c_1 *and* c_2 *are related by* (2) and (3).

Then a function f *satisfies an equation*

$$f = c_1 + P_1 f$$

if and only if its restriction to X_2 *satisfies an equation*

$$f = c_2 + P_2 f$$

and $$f(z) = n_1(z)[\textstyle\sum_{y \in X_2} p_1(z,y) f(y) + c_1(z)] \qquad (4)$$

A new recursive algorithm:

$|X| = k, \quad \sum c_1(y)\pi_1(y) = 0 \Rightarrow \sum c_2\pi_2 = 0;$

$$P_1, c_1 \longrightarrow P_2, c_2 \longrightarrow \ldots \longrightarrow P_k = \{1\}, c_k = 0$$

$$f_1 \quad \longleftarrow f_2 \quad \longleftarrow \ldots \qquad \ldots \longleftarrow \quad f_k = 0$$

Example 1. $X_1 = \{1, 2, 3\}$, P_1, $\pi(1) = \frac{12}{35}$, $\pi(2) = \frac{14}{35}$, $\pi(2) = \frac{9}{35}, (c_1, \pi) = 0, c_1(1) = 3, c_1(2) = 2,$ and $c_1(3) = -\frac{64}{9}$.

$$P_1 = \begin{array}{|c|c|c|} \hline 1/3 & 1/3 & 1/3 \\ \hline 1/4 & 1/2 & 1/4 \\ \hline 1/2 & 1/3 & 1/6 \\ \hline \end{array}, \; P_2 = \begin{array}{|c|c|} \hline 5/8 & 3/8 \\ \hline 7/12 & 5/12 \\ \hline \end{array}, P_3 = \{1\}$$

By eliminating state 1 we obtain transition matrix P_2, and function $c_2, c_2(2) = 2 + \frac{1}{4}\frac{3}{2}3 = \frac{25}{8}$, and $c_2(3) = -\frac{64}{9} + \frac{1}{2}\frac{3}{2}3 = -\frac{175}{36}$. Now eliminate state 2. Then $P_3 = \{1\}, c_3(3) = 0$. Take e.g. $f(3) = 0$. Then $f(2) = \frac{8}{3}\frac{25}{8} = \frac{25}{3}$. Then for $n = 1$, we obtain $f(1) = \frac{3}{2}[\frac{1}{3}\frac{25}{3} + 3] = \frac{26}{3}$. If $(f, \pi) = 0$, we can set $f(1) = \frac{26}{3} + k, f(2) = \frac{25}{3} + k, f(3) = k$ and to find $k = -\frac{662}{105}$. Then $f(1) = \frac{248}{105}, f(2) = \frac{213}{105}$, and $f(3) = -\frac{662}{105}$.

Recursive Solution of Bellman equation $v = \max(g, c + Pv)$.

Theorem 3. (For $D = \{z\}$). *Let* M_1 *and* M_2 *be two adjacent models with state spaces* X_1 and $X_2 = X_1 \setminus z$, *where* $g(z) \leq c_1(z) + P_1 g(z)$, P_1 *and* P_2, *functions* c_1 *and* c_2 *are related by* (2) *and* (3).

Function f *is a (minimal) solution of Bellman equation*

$$f = \max(g, c_1 + P_1 f) \textit{ on } X_1$$

if and only if $f = c_1 + P_1 f$ *on* D, *its restriction to*X_2 *is a (minimal) solution of Bellman equation*

$$f = \max(g, c_2 + P_2 f) \textit{ on } X_2,$$

and the restrictions f_{X_2} *and* f_D *are related by formula*

$$f(z) = n_1(z)[\sum_{y \in X_2} p_1(z, y) f(y) + c_1(z)]$$

Example 2. Let $X_1 = \{1, 2, 3\}$, transition probabilities are given by a matrix P_0, $P_0 = P_1$ from an Example 1, cost function $c(x)$ is: $c_1(1) = 1, c_1(2) = -.5, c_1(3) = .5$, the terminal reward function $g(x)$ is: $g(1) = -1, g(2) = 2, g(3) = 3.5$, and discount factor $\beta = .9$.

$$P_0 = \begin{array}{|l|l|l|}\hline .33 & .33 & .34 \\ \hline .25 & .5 & .25 \\ \hline .5 & .33 & .17 \\ \hline \end{array}, \qquad \mathrm{P}_1 = \begin{array}{|l|l|l|l|}\hline .3 & .3 & .3 & .1 \\ \hline .225 & .45 & .225 & .1 \\ \hline .45 & .3 & .15 & .1 \\ \hline 0 & 0 & 0 & 1 \\ \hline \end{array}.$$

The first step: consider $g(x) - T_1 g(x) \equiv g(x) - (c_1(x) + P_1 g(x))$ and obtain $g(1) - T_1 g(1) = -3.35 < 0$. Therefore state 1 can be eliminated. Then P_2, $c_2(2) = -.18$, and $c_2(3) = 1.14$. At the second step $g(2) - T_2 g(2) = -.04 < 0$ and state 2 can be eliminated.

$$P_2 = \begin{array}{|l|l|l|}\hline .54 & .32 & .13 \\ \hline .50 & .34 & .16 \\ \hline 0 & 0 & 1 \\ \hline \end{array}, \qquad \mathrm{P}_3 = \begin{array}{|l|l|}\hline .69 & .31 \\ \hline 0 & 1 \\ \hline \end{array}$$

P_3, $c_3(3) = .95$.. Then $g(3) - T_3 g(3) = .13 > 0$ and therefore $S = \{3, x_*\}$, and $v(3) = g(3) = 3.5$. Applying formula (4) for $n = 2$, we obtain $v(2) = \frac{1}{.46}[.32(3.5) - .18] = 2.043$. Applying formula (4) again for $n = 1$, we obtain $v(1) = \frac{1}{.7}[.3(2.043) + .3(3.5) + 1] = 3.804$.

Gittins Index (GI) $\gamma(x)$ and Generalized GI $\alpha(x)$

$$\gamma(x) = \sup_{\tau>0} \frac{E_x \sum_{n=0}^{\tau-1} \beta^n c(Z_n)}{E_x \sum_{n=0}^{\tau-1} \beta^n}$$

Multi-armed bandit (MAB) problems with *independent* arms
Gittins (1979), Varaiya et al. (1985), Mandelbaum (1987)

El Karoui and Karatzas (1993), Kaspi and Mandelbaum (1998).

$(1-\beta)E_x \sum_{n=0}^{\tau-1} \beta^n = 1 - E_x\beta^\tau = P_x(Z_\tau = e) \equiv Q^\tau(x)$

GI $\gamma(x) = (1-\beta)\alpha(x)$, where

$$\alpha(x) = \frac{E_x \sum_{n=0}^{\tau-1} \beta^n c(Z_n)}{E_x \sum_{n=0}^{\tau-1} \beta^n} = \sup_{\tau>0} \frac{R^\tau(x)}{Q^\tau(x)},$$

$Q^\tau(x)$ is a probability of *termination*.

Four indices : GI $\gamma(x) = (1-\beta)\alpha(x)$, GGI (Mitten 1960); Kathehakis and Veinot (1987) $h(x)$; Whittle $w(x)$

Theorem 4. *If* $\beta = const$ *then* $\frac{\gamma(x)}{1-\beta} = \alpha(x) = h(x) = w(x)$.

But if $\beta(x)$ is not a constant then the *proportionality* of $\gamma(x)$ and $\alpha(x)$ as functions of x *completely disappears.*

Mitten (1960); Granot and Zuckerman (1991); Denardo, Rothblum, and Van der Heyden, (2004)

Presman and Sonin (2005) *A (Gittins) Index Theorem for Randomly Evolving Graphs* in: *The Shiryaev Festschrift*, Kabanov, Y; Lipster, R; Stoyanov, J (Eds.), Springer, 2006,

Sonin *A Generalized Gittins Index for Markov Chain and its Recursive Calculation.* appear in *Statistics and Probability Letters*, (2008).

Elementary (Mitten) model of independent trials

$$\begin{array}{cccccc} \uparrow & \uparrow & \cdot & \cdot & \cdot & \uparrow \\ e_1 & e_2 & \cdot & \cdot & \cdot & e_n \end{array}$$

Two possible outcomes in each trial, **"continuation"** with probability p_i in i-th trial, and **"termination"** with probability q_i.

A decision maker (DM) can choose an order in which to conduct (test) the trials. Each trial can be tested *only once*. The test of i-th trial brings a **reward** r_i. In the case of "termination" the testing has to be *terminated*. In the case of "continuation" the DM may continue testing or quit. The goal of DM is to select the optimal order to **maximize the expected total reward**.

The formulation is equivalent to a formulation where DM has to pay an amount c_i in advance, obtains a_i with probability p_i, and b_i with probability q_i, and $r_i = -c_i + a_i p_i + b_i q_i$.

A rather simple proof shows that the optimal strategy has a remarkable simple structure and is based on an **index** α calculated for *each trial* e_i, $\alpha(e_i)$ equal to **expected profit divided by probability of termination**, i.e.

$$\alpha(e_i) = \frac{r_i}{q_i}.$$

The optimal strategy has the following form: test the trials with positive index in the order of decreasing. If all trials must be tested then all they should be tested in the above order.

This problem is a reformulation of a “least cost testing sequencing” problem solved independently by a few authors in 1960, in particular by Mitten.

Indices $h(x)$ and $w(x)$.

Kathehakis and Veinot (1987) Markov Decision model

$M = (X, x_1, A(x), P, c(x), \beta(x))$

$A(x) = \{ c = \textit{continue}, \mathrm{r} = \textit{return to } x_1 \}$

$h(x) = \sup_\pi E_x^\pi \sum_{n=0}^\infty r(Z_n, A_n)$,

where $r(x, a) = c(x)$ if $a = c$, $r(x, a) = 0$ if $a = r$

$$h(x_1) = \sup_{\tau > 0} E_x [\sum_{n=0}^{\tau - 1} c(Z_n) + h(x_1)].$$

Whittle Retirement Process

family of OS models $M(k) = (X, P, c(x), k, \beta(x))$,

$g(x|k) = k$ if $x \neq e, g(x|k) = 0$ if $x = e$,

value function $v(x, k) = \sup_{\tau \geq 0} E_x[\sum_{n=0}^{\tau-1} c(Z_n) + g(x|k)]$

$w(x) = \inf\{k : v(x, k) = k\}$

Theorem 5. *The three indices coincide, i.e.* $\alpha(x) = h(x) = w(x)$.

Apply SE algorithm:
$g(x) - Fg(x) = k - (c(x) + \beta(x)k) = (1 - \beta(x))(k - d(x))$, where function $d_1(x) = \frac{c(x)}{1-\beta(x)}$, number $d = \max_{x \in X} d(x)$, set $D = \{x : d(x) = d\}$.

Lemma 4 *Let* $M(k)$ *be the corresponding Whittle OS model. Then* $\alpha(x) = d$ *for* $x \in D$, and $\alpha(x) < d$ *for all* $x \in X \setminus D$.

Algorithm to calculate $\alpha(x)$

model $M_1 = (X_1, P_1, c_1(x), \beta_1(x))$, function $d_1(x) = \frac{c_1(x)}{1-\beta_1(x)}$,

number d_1, set $D_1 : \alpha(x) = d_1$ for all $x \in D_1$;

eliminate D_1, recalculate $P_1 \to P_2, c_1(x) \to c_2, \beta_1(x) \to \beta_2$,

obtain $d_2(x)$, d_2, set D_2. Then by lemma 4 $\alpha(x) = d_2$ for all $x \in D_2$ and so on.

Example 3. Let our reward model has $X_1 = \{1, 2, 3, x_*\}$, with $c_1(1) = 3, c_1(2) = 2, c_1(1) = 1$, and $\beta(x) = .9$ for all $x \neq x_*$ and corresponding transition matrix $P = P_1$ is

$$P_1 = \begin{array}{|c|c|c|c|}\hline .3 & .3 & .3 & .1 \\ \hline .45 & .3 & .15 & .1 \\ \hline .1 & .5 & .3 & .1 \\ \hline 0 & 0 & 0 & 1 \\ \hline \end{array}\,, \; P_2 = \begin{array}{|c|c|c|}\hline .5 & .34 & .16 \\ \hline .54 & .34 & .12 \\ \hline 0 & 0 & 1 \\ \hline \end{array}\,, \; P_3 = \begin{array}{|c|c|}\hline .71 & .29 \\ \hline 0 & 1 \\ \hline \end{array}$$

Then $d_1 = c_1(1)/(1 - \beta) = 30, D_1 = \{1\}$, and by Lemma 2 $\alpha(1) = 30.$ Therefore, eliminate state 1 on a first step and, applying formulas (2) and (3), obtain new transition matrix P_2 and function $c_2(x)$ for a state space $X_2 = \{2, 3, x_*\}$; $c_2(2) = 3.93, c_2(3) = 1.43.$ Therefore $d_2 = c_2(2)/(1 - \beta_2(2)) = 3.93/0.16 = 24.56 = \alpha(2), D_2 = \{2\}$ and on the second step state 2 is eliminated, and we obtain matrix P_3, and $c_3(3) = 5.63.$ Therefore $\alpha(3) = c_3(3)/(1 - \beta_3(3)) = 5.63/0.29 = 19.41.$

Note that though we started in this example from a constant survival function $\beta(x)$, after the first step we deal with variable $\beta_i(x)$ for $i > 1$. The classical GI for this model $\gamma(x) = (1 - \beta)\alpha(x) = .1\alpha(x)$.

Three abstract optimization problems: There is an index set U, and $A = \{a_u\}$ and $B = \{b_u\}$ be two sets indexed by the elements of U. Suppose that an assumption AB holds,

$$a_u \le a < \infty, \quad 0 < b \le b_u \le 1 < \infty.$$

Problem 1 A. Restart pr-m. Find solution(s) of the equation

$$h = H(h) = \sup_{u \in U}[a_u + (1 - b_u)h]) \tag{1}$$

The interpretation of this problem is as follows. There is a set of "buttons" U. A DM can select one of them and push. She obtains a reward a_u and with probability b_u the game is over, with complimentary probability $1 - b_u$ she can select any button again. Her goal is to maximize the total (undiscounted) reward. It can be easily proved that her value satisfies the equation above.

[in Mitten problem each button can be used only once.]

Problem 2 A. Ratio (cycle) pr-m. Find

$$\alpha = \sup_{u \in U} \frac{a_u}{b_u} \tag{2}$$

Problem 3 A. A parametric family of Retirement pr-ms.

Find w, defined as follows: given parameter $k, -\infty < k < \infty$, let

$$v(k) = k \vee H(k), \quad w = \inf\{k : v(k) = k\}. \tag{3}$$

Theorem 1 A. a) $\alpha < \infty$ *and is a (unique) solution of equation* (1)

b) $\alpha = h = w$.

c) *The optimal index* u, *if exists, or an optimizing sequence* u_n *coincide for all three problems.*

Open Problems. Plans

- more examples when SEA is efficient
- analog of SEA in continuous time/space
- games
- Multiple restart problems
- Multidimensional Abstract Optimization ?

Bonus Slides

Let us consider the following simple random experiment; we flip a *coin*, we toss a *die*. Then our sample space consists of 12 outcomes each having a probability 1/12. This experiment is used in many textbook as an illustration of a concept of independent events.

Question 1. How many different *pairs* (A, B) *of independent events* are there ?

Question 2. How many different *tuples of independent events* $(A_1, A_2, ..., A_k)$ are there ?

The answers for both questions, numbers N_1 and N_2 are a bit puzzling.

Answers

The first number, $N_1 = 888,888.$

The second, $N_2 = 30,826,488.$

Thank you for your attention !