УДК 519.21

Ал.А. ВОЙНА, Ан.А. ВОЙНА

ДИНАМИЧЕСКОЕ УПРАВЛЕНИЕ РИСКОМ В МНОГОМЕРНЫХ МАРКОВСКИХ МОДЕЛЯХ

Аннотация. Рассмотрена типичная для многих областей прикладной математики модель выбора оптимальных решений в стохастических системах. Управляемый на промежутке случайной длины многомерный процесс гибели, который изучается в статье, может описывать явления, возникающие в экономических, биологических, медицинских, технических и других приложениях. Предложено несколько подходов к определению критерия оптимальности и изучены свойства соответствующих функций риска, а также разработан конструктивный вычислительный алгоритм построения оптимальных стационарных стратегий.

Ключевые слова: оптимальные стационарные стратегии, управляемые процессы, марковские процессы принятия решений, функция риска, процессы гибели и рождения, итерационные алгоритмы.

В статье изучается типичная для многих экономических моделей ситуация относительно выбора оптимальных стратегий, описанная, в частности, в маркетинговых моделях [1, 2], задачах дистрибуции товаров [2], управления запасами [4], моделях страхования и управления риском [5, 6]. Подобные ситуации часто можно встретить также в технических приложениях, связанных, например, с контролем качества, в биологических моделях и медицинских исследованиях, касающихся профилактики заболеваний, борьбы с эпидемиями. Характерной особенностью многих упомянутых выше моделей есть то, что эргодические свойства соответствующих им случайных процессов не позволяют строить оптимальные стратегии с помощью «классических» подходов. В марковских моделях, как правило, выбор оптимальных решений основан на использовании их стационарных распределений. Важная особенность задачи, рассмотренной в настоящей статье, заключается в том, что все состояния управляемого многомерного процесса являются несущественными, за исключением одного. Стационарное распределение такого процесса сосредоточено в этом выделенном состоянии, которое является «оптимальным» с точки зрения значения целевой функции. Следующая существенная особенность заключается в выборе оптимальных управлений на промежутке времени T случайной длины.

Рассмотрим целочисленный многомерный марковский процесс $\xi(\pi, t)$ с дискретным временем $t = 0, 1, 2, \ldots$ и конечным множеством состояний $\xi(\pi, t) \in X(M)$, который управляется согласно стратегии π. При этом

$$\xi(\pi, t) = \{\xi_1(\pi, t), \xi_2(\pi, t), \ldots, \xi_m(\pi, t)\},$$

$$X(M) = \{(x_1, x_2, \ldots, x_m): 0 \le x_i \le M_i, i = 1, 2, \ldots, m\},$$

где $M = \{M_1, M_2, \ldots, M_m\}$ — известный вектор с целочисленными положительными координатами. Стратегия управления π определяется следующим образом. Предположим, что задано множество A, состоящее из r векторов:

$$A = \{a^{(1)}, a^{(2)}, \ldots, a^{(r)}\}, \; a^{(u)} = \{a_{u1}, a_{u2}, \ldots, a_{um}\}, \; u \in U = \{1, 2, \ldots, r\},$$

при этом для каждого $u \in U$ имеем

$$0 \le a_{uj} \le 1, \;\; j = 1, 2, \ldots, m.$$

Назовем $U = \{1, 2, \ldots, r\}$ множеством управлений. Для каждого управления $u \in U$ определена однородная цепь Маркова $x^{(u)}(t)$ с дискретным временем

$t = 1, 2, 3, \ldots$ и значениями в множестве $X(M)$. Переходные вероятности за один шаг марковской цепи представляются следующим образом.

• Если $x = (x_1, x_2, \ldots, x_m) \in X(M)$ и $y = (y_1, y_2, \ldots, y_m) \in X(M)$ — два состояния цепи $x^{(u)}(t)$ и при этом $x \geq y$, т.е. для произвольного числа $i = 1, 2, \ldots, m$ выполняется неравенство $x_i \geq y_i$, то

$$p^{(u)}(x, y) = P\{x^{(u)}(t+1) = (y_1, y_2, \ldots, y_m) / x^{(u)}(t) = (x_1, x_2, \ldots, x_m)\} =$$

$$= C_{x_1}^{x_1 - y_1} (a_{u1})^{x_1 - y_1} (1 - a_{u1})^{y_1} \times C_{x_2}^{x_2 - y_2} (a_{u2})^{x_2 - y_2} (1 - a_{u2})^{y_2} \times \ldots$$

$$\ldots \times C_{x_m}^{x_m - y_m} (a_{um})^{x_m - y_m} (1 - a_{um})^{y_m}.$$

• Если условие $x \geq y$ не выполняется, т.е. хотя бы для одной координаты i выполняется неравенство $x_i < y_i$, то

$$p^{(u)}(x, y) = P\{x^{(u)}(t+1) = (y_1, y_2, \ldots, y_m) / x^{(u)}(t) = (x_1, x_2, \ldots, x_m)\} = 0.$$

Введем множества

$$X_0 = X_1 = X_2 = \ldots = X_n = X(M); \; U_1 = U_2 = \ldots = U_n = U.$$

Для каждого $t \in \{0, 1, 2, \ldots, n\}$ определим множество

$$I(t) = X_0 \times U_1 \times X_1 \times U_2 \times X_2 \times \ldots \times U_t \times X_t.$$

Каждую последовательность

$$i(t) = \{x^{(0)}, u^{(1)}, x^{(1)}, u^{(2)}, x^{(2)}, .., u^{(t)}, x^{(t)}\} \in I(t)$$

назовем историей длины $t \in \{0, 1, 2, \ldots, n\}$, а символом $\nu_t(i(s))$, $s \geq t$, обозначим отображение

$$\nu_t : I(s) \to I(s - t), \; s \geq t,$$

где $i(s) = \{\xi(0), u(1), \xi(1), u(2), \xi(2), .., u(s), \xi(s)\} \in I(s)$, а

$$\nu_t(i(s)) = \{\xi(t), u(t+1), \xi(t+1), u(t+2), \xi(t+2), .., u(s), \xi(s)\} \in I(s - t).$$

Стратегией управления для управляемой марковской модели будем называть функцию $\pi = \{\pi_t(\cdot / i(t)), i(t) \in I(t), t \in \{0, 1, \ldots, n-1\}\}$, которая ставит в соответствие произвольной истории $i(t) \in I(t)$, $t \in \{0, 1, 2, \ldots, n\}$, некоторое распределение вероятностей $\pi_t(\cdot / i(t))$, сосредоточенное на множестве U_{t+1}.

Стратегия π называется простой, если для произвольного $t \in \{0, 1, 2, \ldots, n-1\}$ и произвольной истории $i(t) \in I(t)$ распределение вероятностей $\pi_t(\cdot / i(t))$ сосредоточено в одной точке множества U_{t+1}.

Если для произвольного момента $t \in \{0, 1, 2, \ldots, n-1\}$ выполняется равенство

$$\pi_t(\cdot / i(t)) = \pi_t(\cdot / \nu_t(i(t))),$$

то стратегию $\pi = \{\pi_t(\cdot / i(t)), i(t) \in I(t), \; t \in \{0, 1, \ldots, n-1\}\}$ назовем марковской. Множество всех марковских стратегий обозначим Π_{M}.

Пусть $\pi = \{\pi_t(\cdot / i(t)), i(t) \in I(t), t \in \{0, 1, \ldots, n-1\}\}$ — некоторая стратегия управления, а $\mu(\xi)$, $\xi \in X_0$, — мера, определяющая начальное распределение управляемого согласно стратегии π многомерного марковского процесса $\xi_\pi = \xi(\pi, t)$, $t = 0, 1, 2, \ldots$ Обозначим $P_n(\cdot)$ соответствующее ему распределение вероятностей на множестве $I(n)$ всех историй длины n и для произвольной функции $\zeta(l)$, $l \in I(n)$, определим математическое ожидание $M_n(\zeta) = E(\zeta(\xi_\pi))$:

$$M_n(\zeta) = \sum_{l \in I(n)} \zeta(l) \cdot P_n(l).$$

Для того чтобы ввести понятие оптимальной стратегии, зададим вектор коэффициентов $c = (c_1, c_2, \ldots, c_m)$ и определим на множестве $X(M)$ функцию $r(\xi)$,

$\xi \in X(M)$, следующим образом. Если $\xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X(M)$, то

$$r(\xi) = \sum_{i=1}^{m} c_i \cdot \xi_i .$$

Предположим теперь, что определена некоторая стратегия управления $\pi = \{\pi_t(\cdot / i(t)),\, i(t) \in I(t),\, t \in \{0, 1, \ldots, n-1\}\}$ и выбрано начальное состояние $\xi \in X_0$ управляемого многомерного марковского процесса. Оценкой состояния $\xi \in X_0$ для стратегии $\pi = \{\pi_t(\cdot / i(t)),\, i(t) \in I(t),\, t \in \{0, 1, \ldots, n-1\}\}$ назовем величину $S(\xi, \pi)$, определенную равенством

$$S(\xi, \pi) = \sum_{l \in I(n)} r(\nu_n(l)) \cdot P_n(l).$$

Соответственно величину

$$V(\xi) = \inf_{\pi \in \Pi_{\text{M}}} \{S(\xi, \pi)\}$$

назовем оценкой начального состояния $\xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_0$ в множестве Π_{M} всех марковских стратегий. Будем считать, что марковская стратегия

$$\pi^* = \{\pi_t^*(\cdot / i(t)),\, i(t) \in I(t),\, t \in \{0, 1, \ldots, n-1\}\}$$

является оптимальной, если выполняется равенство

$$V(\xi) = S(\xi, \pi^*).$$

Ниже в теореме предложен конструктивный алгоритм построения оптимальной стратегии в случае, когда временной интервал управления $t \in T = \{0, 1, 2, \ldots, n\}$ не случайный, а его длина n однозначно определена.

Теорема 1. Предположим, что начальное распределение $\mu(\xi)$, $\xi \in X_0$, процесса $\xi(t)$, $t = 0, 1, 2, \ldots$, сосредоточено в одной точке $M = \{M_1, M_2, \ldots, M_m\} \in X_0$, т.е.

$$P\{\xi(0) = \{M_1, M_2, \ldots, M_m\}\} = 1.$$

Определим семейство функций

$$[V_t(\xi),\ \xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_t,\ t \in \{n, n-1, \ldots, 2, 1, 0\}]$$

следующим образом. Для $t = n$ положим по определению

$$V_n(\xi) = r(\xi) = r(\xi_1, \xi_2, \ldots, \xi_m) = \sum_{i=1}^{m} c_i \cdot \xi_i ,\ \ \xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_n .$$

Для остальных $0 \le t < n$ соответствующие функции $V_t(\xi)$ определяем рекуррентно:

$$V_t(\xi) = \min_{u \in A} \left[\sum_{\zeta \in X_{t+1}} p^{(u)}(\xi, \zeta) \cdot V_{t+1}(\zeta) \right],\ \ t \in \{n-1, n-2, \ldots, 1, 0\},$$

$$\xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_t ,\ \ \zeta = \{\zeta_1, \zeta_2, \ldots, \zeta_m\} \in X_{t+1}.$$

Тогда оптимальной для управляемого многомерного марковского процесса $\xi_\pi = \xi(\pi, t)$, $t = 0, 1, 2, \ldots$, будет простая марковская стратегия

$$\pi = \{\pi_t(\cdot / i(t)) = \gamma_t,\, i(t) \in I(t),\ t \in \{0, 1, \ldots, n-1\}\},$$

которая строится следующим образом:

$$\pi_t = \pi_t(u / \xi) = \delta(u, u_t(\xi)) = \begin{cases} 1, \text{если } u = u_t(\xi), \\ 0, \text{если } u \ne u_t(\xi), \end{cases}$$

$$u_t(\xi) = \arg\min_{u \in U} \left[\sum_{\zeta \in X_{t+1}} p^{(u)}(\xi, \zeta) \cdot V_{t+1}(\zeta) \right].$$

Доказательство теоремы базируется на использовании идей динамического управления, исходя из общей теории управляемых марковских процессов [8, 7] с учетом особенностей рассматриваемой модели. Согласно постановке задачи необходимо выбрать в точности n управлений $u_{t+1} \in U_{t+1}$, $t = 0, 1, 2, \ldots, n-1$, причем каждое из них зависит от состояния $\xi(\pi, t) = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_t$ управляемого марковского процесса в момент времени $t = 0, 1, 2, \ldots, n-1$. Поэтому практическая реализация процедуры построения описанной в теореме 1 простой марковской стратегии $\pi = \{\pi_t(\cdot / i(t)) = \pi_t^*,\ i(t) \in I(t), t \in \{0, 1, \ldots, n-1\}\}$ представляется следующим образом.

• Процесс определения оптимальных решений начинаем с выбора управления $u_n \in U_n$ на n-м шаге в момент времени $t = n-1 \in T$. Если $\xi(\pi, t-1) = \{\xi_1, \xi_2, \ldots, \xi_m\}$, то

$$u_n = u_n(\xi) = \underset{u \in U}{\arg\min} \left[\sum_{\zeta \in X_1} p^{(u)}(\xi, \zeta) \cdot V_{(n)}(\zeta) \right] = \underset{u \in U}{\arg\min} \left[\sum_{\zeta \in X_1} p^{(u)}(\xi, \zeta) \cdot r(\zeta) \right]$$

или

$$\pi_{n-1}^* = \delta(u, u_n(\xi)).$$

• С использованием функции $[V_n(\xi), \xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_n]$ строим очередное семейство $[V_{n-1}(\xi), \xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_{n-1}]$, где

$$V_{n-1}(\xi) = \min_{u \in A} \left[\sum_{\zeta \in X_{t+1}} p^{(u)}(\xi, \zeta) \cdot V_n(\zeta) \right].$$

Оптимальным решением на $(n-1)$-м шаге при условии, что управляемый марковский процесс находится в состоянии $\xi = \xi_1, \xi_2, \ldots, \xi_m \in X_{n-2}$, является управление $u_{n-1} = u_{n-1}(\xi) \in U_{n-1}$, удовлетворяющее условию

$$u_{n-1} = u_{n-1}(\xi) = \underset{u \in U}{\arg\min} \left[\sum_{\zeta \in X_1} p^{(u)}(\xi, \zeta) \cdot V_{(n-1)}(\zeta) \right]$$

или

$$\pi_{n-2}^* = \delta(u, u_{n-1}(\xi)).$$

Продолжая описанную процедуру, построим простую марковскую стратегию $\pi^{(0)*} = \pi_0^* \cdot \pi_1^* \cdot \ldots \cdot \pi_{n-1}^*$, которая на основании теоремы 1 является оптимальной для управляемого многомерного марковского процесса $\xi_\pi = \xi(\pi, t)$, $t = 0, 1, 2, \ldots$

Как следует из определения простой марковской стратегии, для произвольного $t \in \{0, 1, \ldots, n-1\}$ функция $\pi_t(\cdot / \xi^{(t)})$, $\xi^{(t)} \in X_t$, является простой функцией и порождает разбиение множества X_t на r подмножеств:

$$\{Y_t(j), j \in U_{t+1} = \{1, 2, \ldots, r\}\}.$$

При этом

$$Y_t(j) = \{\xi \in X_t, u_{t+1}(\xi) = j \in U_{t+1} = \{1, 2, \ldots, r\}\}.$$

Очевидно, что

$$Y_t(j) \cap Y_t(l) = \varnothing,\ j \neq l, \quad j, l \in U_{t+1} = \{1, 2, \ldots, r\},$$

$$Y_t(1) \cup Y_t(2) \cup \ldots \cup Y_t(r) = X_t.$$

Для многомерного управляемого марковского процесса $\xi(\pi, t) = \{\xi_1(\pi, t), \xi_2(\pi, t), \ldots, \xi_m(\pi, t)\}$ соответствующие разбиения строятся следующим образом. На первом шаге алгоритма для всех $\xi = \{\xi_1, \xi_2, \ldots, \xi_m\} \in X_n$ вычисляем $V_{(n)}(\xi)$:

$$V_{(n)}(\xi) = r(\xi) = \sum_{i=1}^{m} c_i \cdot \xi_i.$$

Предположим, что u является некоторым управлением из множества $U = \{1, 2, \dots, r\}$, а $\xi = \{\xi_1, \xi_2, \dots, \xi_m\} \in X_{n-1}$. Учитывая вид переходных функций $p^{(u)}(x, y), x \in X(M), \ y \in X(M), \ u \in U$, семейства цепей Маркова $\{x^{(u)}(t)$, $u = 1, 2, \dots, r\}$, для вычисления $V_{(n-1)}^{(u)}(\xi)$ получаем следующее равенство:

$$V_{(n-1)}^{(u)}(\xi) = V_{(n-1)}^{(u)}(\xi_1, \xi_2, \dots, \xi_n) = \sum_{h \in X_n} p^{(u)}(\xi, h) \cdot V_{(n)}(h) = \sum_{i=1}^{m} c_i \cdot \xi_i \cdot (1 - a_{ui}).$$

Определяем оценку $V_{(n-1)}(\xi)$ состояния $\xi \in X_{n-1}$, и если окажется, что

$$V_{(n-1)}(\xi) = V_{(n-1)}^{(j)}(\xi),$$

то состояние ξ включаем в множество $Y_{(n-1)}(j)$, где

$$j = \min_{j \in U} \{V_{(n-1)}(\xi_1, \xi_2, \dots, \xi_m) = V_{(n-1)}^{(j)}(\xi_1, \xi_2, \dots, \xi_m)\}.$$

Продолжаем описанную процедуру до тех пор, пока ни вычислим для каждого состояния $\xi = \{\xi_1, \xi_2, \dots, \xi_m\} \in X_0$ его оценку $V_{(0)}(\xi)$ и ни построим разбиение $[Y_0(j), j = 1, 2, \dots, r\}$ множества X_0.

Предположим, что в каждый момент времени t независимо от номера очередного шага и истории $i(t) = \{x^{(0)}, u^{(1)}, x^{(1)}, u^{(2)}, x^{(2)}, \dots, u^{(t)}, x^{(t)}\} \in I(t)$ процесс управления может быть прерван. Другими словами, пусть длина τ временного интервала управления T_τ — случайная и при этом τ имеет геометрическое распределение с параметром p:

$$t \in T_\tau = \{0, 1, 2, \dots, \tau\}; \ \tau \in \{1, 2, 3, \dots\}, \ P\{\tau = n\} = (1-p)^{n-1} \cdot p, \ n = 1, 2, 3, \dots$$

В этом случае предложенное в теореме 1 решение нельзя непосредственно использовать, необходима соответствующая его модификация. Вначале скорректируем постановку оптимизационной задачи. Зададим вектор $d = (d_1, d_2, \dots, d_m)$ и для каждого $t \in T_\tau = \{0, 1, 2, \dots, \tau\}$ положим

$$L_\pi(\xi(t), \xi(t+1)) = \sum_{i=1}^{m} d_i \cdot (\xi_i(t) - \xi_i(t+1)).$$

Целевую функцию для стратегии управления $\pi = \{\pi_t(\cdot / i(t)), i(t) \in I(t)\}$ определим следующим образом:

$$L_\pi = \sum_{t=1}^{\tau} (L_\pi(\xi(t-1), \xi(t))).$$

Проведем также соответствующую модификацию формальной модели. Добавим к множеству $X(M)$ некоторое специальное состояние $x^* = (x_1^*, x_2^*, \dots, x_m^*)$ и по определению установим

$$x_i^* - x_i^* = 0, \ i \in \{1, 2, \dots, m\}; \ c - x_i^* = 0, \ i \in \{1, 2, \dots, m\},$$

где c — произвольное целое неотрицательное число. Определим новое семейство цепей Маркова

$$\{x^{*(u)}(t), u = 1, 2, \dots, r\} = (x^{*(1)}(t), x^{*(2)}(t), \dots, x^{*(r)}(t)\}$$

со значениями в множестве состояний $X^*(M) = \{x^*\} \cup X(M)$. Переходные функции за один шаг цепи $x^{*(u)}(t)$, $t = 0, 1, 2, \dots$, приобретают следующий вид:

$$p^{*(u)}(x, y) =$$

$$= P\{x^{*(u)}(t+1) = (y_1, y_2, \dots, y_m) / x^{*(u)}(t) = (x_1, x_2, \dots, x_m)\} = p^{(u)}(x, y),$$

если $x = (x_1, x_2, \dots, x_m) \in X(M)$ и $y = (y_1, y_2, \dots, y_m) \in X(M)$.

Для произвольного $x=(x_1, x_2, \dots, x_m) \in X(M)$ независимо от момента времени t и от решения $u \in U_{t+1}$

$$p^{*(u)}(x, x^*)=p, \qquad p^{*(u)}(x^*, x^*)=1.$$

Пусть $f=\{f_x, x \in X(M)\}$ — вектор, размерность (N) которого совпадает с размерностью пространства состояний $X(M)$ семейства цепей Маркова $\{x^{(u)}(t), u=1,2,\dots,r\}$, а координаты f_x, $x \in X(M)$, принимают свои значения в множестве U возможных решений: $f_x \in U=\{1,2,\dots,r\}$, $x \in X(M)$. Вектор f назовем вектором решений для управляемого марковского процесса. Произвольную последовательность $\{f(t), t=1,2,3,\dots\}$, где для каждого $t=1,2,3,\dots$ $f(t)=\{f_x(t), x \in X(M)\}$ — вектор решений, назовем стратегией и обозначим, как и ранее, символом π:

$$\pi=\{f(1), f(2), f(3), \dots\}.$$

Назовем π стратегией стационарной, если вектор решений $f(t)=f=\{f_x, x \in X(M)\}$ не зависит от $t \in T$.

Как и в случае процесса $\xi(\pi, t)$, $t=0,1,2,\dots$, семейства цепей Маркова $\{x^{*(u)}(t), u=1,2,\dots,r\}$ вместе со стратегией управления π определяют на множестве $I^*(t)$ всех историй $i^*(t) \in I^*(t)$ длины t распределение вероятностей $P^*(i^*(t))$ и соответственно некоторый управляемый случайный процесс $\xi_\pi^*=\xi^*(\pi, t)$, $t=0,1,2,\dots$ Чтобы определить целевую функцию $L_\pi^*=L^*(\{\xi^*(\pi, t), t \in T\})$ процесса $\xi_\pi^*=\xi^*(\pi, t)$, $t=0,1,2,\dots$, введем функционал

$$L_\pi^*(\xi^*(\pi, t), \xi^*(\pi, t+1))=\sum_{i=1}^m d_i \cdot (\xi_i^*(\pi, t)-\xi_i^*(\pi, t+1)),$$

который можно интерпретировать, как выгоду («прибыль») при переходе процесса $\xi_\pi^*=\xi^*(\pi, t)$, $t=0,1,2,\dots$, из состояния $\xi^*(\pi, t)=\xi(t) \in X(M)$ в состояние $\xi^*(\pi, t+1)=\xi(t+1) \in X(M)$. В свою очередь,

$$L_\pi^*=\sum_{t=0}^\infty L_\pi^*(\xi^*(\pi, t), \xi^*(\pi, t+1))$$

соответственно можно воспринимать как выгоду, полученную от управления случайным процессом $\xi_\pi^*=\xi^*(\pi, t)$, $t=0,1,2,\dots$, на бесконечном промежутке времени. Учитывая предположения, принятые относительно состояния $x^*=(x_1^*, x_2^*, \dots, x_m^*)$ процесса ξ_π^*, покажем, что

$$L_\pi^*=\sum_{t=0}^\infty L_\pi^*(\xi^*(\pi, t), \xi^*(\pi, t+1))=\sum_{t=1}^\tau (L_\pi(\xi(t-1), \xi(t)))=L_\pi.$$

Обозначим символом $P_{(f(t))}^*$ матрицу, размерность которой совпадает с размерностью пространства $X^*(M)=\{x^*\} \cup X(M)$, при этом

$$p_{(f(t))}^*(x, y)=p^{*(f_x(t))}(x, y), \; x \in X^*(M), \; y \in X^*(M).$$

Для определенности будем считать, что первая строка и первый столбец матрицы $P_{(f(t))}^*$ соответствуют состоянию $x^*=(x_1^*, x_2^*, \dots, x_m^*)$. Тогда совокупность матриц

$$P^{*(u)}=\| p^{*(u)}(x, y) \|, \quad x \in X^*(M), \; y \in X^*(M), \; u \in U,$$

определяет переходные вероятности за один шаг управляемого марковского процесса при условии, что после каждого очередного шага, независимо от его номера t, а также от истории управления, процесс управления может быть прерван. Если $\pi = \{f(1), f(2), f(3), \dots\}$ — некоторая стратегия, то для произвольного $t > 0$ символ $P_t^*(\pi)$ обозначает матрицу, определенную равенством

$$P_t^*(\pi) = P^*_{(f(1))} \times P^*_{(f(2))} \times \dots \times P^*_{(f(t))}.$$

В случае $t = 0$ для произвольной стратегии $\pi = \{f(1), f(2), f(3), \dots\}$ положим по определению

$$P_0^*(\pi) = E,$$

где E — единичная матрица соответствующей размерности. Введем вектор-столбец $m_\pi = \{m_\pi(x), x \in X(M)\}$, размерность которого совпадает с размерностью пространства состояний $X(M)$, а координаты определяются равенством

$$m_\pi(x) = E\{L_\pi / \xi^*(\pi, 0) = x\}, \; x = \{x_1, x_2, \dots, x_m\} \in X(M).$$

Если интерпретировать L_π^* как выгоду (прибыль), полученную от управления процессом ξ_π^*, то координаты $m_\pi(x)$ вектора m_π представляют среднюю прибыль при условии, что $x \in X(M)$ является начальным состоянием для управляемого случайного процесса $\xi_\pi^* = \xi^*(\pi, t)$, $t = 0, 1, 2, \dots$

Стационарную стратегию $\hat{\pi} = \hat{\pi}(\hat{f}) = \{\hat{f}, \hat{f}, \hat{f}, \dots\}$, где $\hat{f} = \{\hat{f}_x, x \in X(M)\}$, назовем оптимальной, если для произвольной допустимой стратегии $\pi(f) = \{f, f, f, \dots\}$, где $f = \{f_x, x \in X(M)\}$, выполняется неравенство

$$m_{\hat{\pi}} \ge m_\pi.$$

Справедлив следующий результат.

Теорема 2. Пусть $x = \{x_1, x_2, \dots, x_m\} \in X(M)$, а координаты $d(x)$ вектора $D = \{d(x), x \in X(M)\}$ определяются равенством

$$d(x) = \sum_{i=1}^{m} d_i \cdot x_i.$$

Тогда выполняется следующее соотношение:

$$m_\pi = \left(E - p \cdot \sum_{t=0}^{\infty} P_t(\pi) \right) \cdot D,$$

где $P_t(\pi)$ — матрица, которая получается из матрицы $P_t^*(\pi)$ исключением первой строки и первого столбца, а $P_0(\pi)$ — единичная матрица E, размерность которой совпадает с размерностью пространства состояний $X(M)$ управляемого марковского процесса $\xi_\pi = \xi(\pi, t)$, $t = 0, 1, 2, \dots$

Определим вектор $w_\pi = \{w_\pi(x), x \in X(M)\}$ следующим образом:

$$w_\pi = p \cdot \sum_{t=0}^{\infty} P_t(\pi) \cdot D.$$

Используя результат теоремы 2, легко убедиться, что стационарная стратегия $\hat{\pi} = \hat{\pi}(\hat{f}) = \{\hat{f}, \hat{f}, \hat{f}, \dots\}$, где $\hat{f} = \{\hat{f}_x, x \in X(M)\}$, будет оптимальной тогда и только тогда, когда для произвольной допустимой стратегии $\pi(f) = \{f, f, f, \dots\}$ выполняется неравенство $w_{\hat{\pi}} \le w_\pi$.

Пусть $\pi = \{f(1), f(2), f(3), \dots\}$ — некоторая стратегия, $g = \{g_x, x \in X(M)\}$ — некоторый вектор решений. Тогда символ (g, π) будет обозначать следующую стратегию: $\{g, f(1), f(2), f(3), \dots\}$.

Теорема 3. Если для произвольного вектора решений $g = \{g_x, x \in X(M)\}$ выполняется неравенство

$$w_{\hat{\pi}} \le w_{(g, \hat{\pi})},$$

то стратегия $\hat{\pi} = \hat{\pi}(\hat{f}) = \{\hat{f}, \hat{f}, \hat{f}, \dots\}$, где $\hat{f} = \{\hat{f}_x, x \in X(M)\}$, будет оптимальной.

Поскольку согласно теореме 2 выполняется равенство

$$m_{\pi} = D - w_{\pi},$$

то, используя теорему 3, можем сформулировать следующий практический критерий оптимальности стратегии $\hat{\pi}$: если для произвольного вектора решений $g = \{g_x, x \in X(M)\}$ выполняется неравенство

$$m_{\hat{\pi}} \geq m_{(g, \hat{\pi})},$$

то стационарная стратегия $\hat{\pi} = \hat{\pi}(\hat{f}) = \{\hat{f}, \hat{f}, \hat{f}, \dots\}$, где $\hat{f} = \{\hat{f}_x, x \in X(M)\}$ — некоторый вектор решений, будет оптимальной.

Сформулируем теперь алгоритм построения вектора решений $\hat{f} = \{\hat{f}_x, x \in X(M)\}$ и соответственно оптимальной стратегии $\hat{\pi}$.

Теорема 4. Пусть $\pi = \{f(1), f(2), \dots, f(t), \dots\}$ — некоторая стратегия, а $f = \{f_x, x \in X(M)\}$ — некоторый вектор решений, $\pi(f) = \{f, f, f, \dots\}$ — стационарная стратегия, построенная на основании вектора решений $f = \{f_x, x \in X(M)\}$. Если выполнено условие

$$m_{\{f\pi\}} \geq m_{\pi}, \text{ то } m_{\pi(f)} \geq m_{\pi}.$$

Теорема 5. Пусть $f = \{f_x, x \in X(M)\}$ — некоторый вектор решений. Определим для каждого вектора $x \in X(M)$ множество $Q(x, f)$ таких решений $u \in U = \{1, 2, \dots, r\}$, для которых выполняется соотношение

$$p \cdot d(x) + p \cdot \sum_{y \in X(M)} p^{*(u)}(x, y) \cdot w_{\pi(f)}(y) < w_{\pi(f)}(x).$$

Построенная на основании вектора решений $f = \{f_x, x \in X(M)\}$ стационарная стратегия $\pi(f) = \{f, f, f, \dots\}$ будет оптимальной тогда, когда для всех $x \in X(M)$ множество $Q(x, f)$ пусто.

Если же существуют такие $x \in X(M)$, что $Q(x, f)$ не пусто, то для каждого вектора решений $g = \{g_x, x \in X(M)\}$, где $g_x = f_x$ в случае, когда $g_x \notin Q(x, f)$, и (возможно) $g_x \neq f_x$ в случае, когда $g_x \in Q(x, f)$, выполняется неравенство

$$m_{\pi(g)} \geq m_{\pi(f)}.$$

Теорема 6. Для управляемого марковского процесса $\{x^{*(u)}(t), u = 1, 2, \dots, r\}$ с множеством состояний $X^*(M) = \{x^*\} \cup X(M)$, конечным множеством решений $u \in U = \{1, 2, \dots, r\}$ и семейством переходных вероятностей

$$\{p^{*(u)}(x, y), u = 1, 2, \dots, r\},\ x \in X^*(M),\ y \in X^*(M),\ X^*(M) = \{x^*\} \cup X(M)$$

существует оптимальная стационарная стратегия $\pi(f^*) = \{f^*, f^*, f^*, \dots\}$.

Вектор решений $f^* = \{f_x^*, x \in X(M)\}$, на основании которого строится стратегия $\pi(f^*) = \{f^*, f^*, f^*, \dots\}$, можно определить, используя следующую итерационную процедуру.

1. Выбираем произвольный вектор решений $f = \{f_x, x \in X(M)\}$ из условия

$$\sum_{y \in X(M)} p^{*(f_x)}(x, y) \cdot d(y) = \min_{u = 1, \dots, r} \sum_{y \in X(M)} p^{*(u)}(x, y) \cdot d(y).$$

2. Решаем систему уравнений

$$\begin{cases} V(x) = p \cdot d(x) + p \cdot \sum_{y \in X(M)} p^{*(f_x)}(x, y) \cdot V(y), \\ x \in X(M) \end{cases}$$

и находим вектор $V = \{V(x), x \in M\}$.

3. Для каждого состояния $x \in X(M)$ определяем множество $Q(x, f)$ таких решений $u \in U = \{1, 2, \ldots, r\}$, для которых выполняется неравенство

$$p \cdot d(x) + p \cdot \sum_{y \in X(M)} p^{(u)}(x, y) \cdot V(y) < V(x).$$

Если окажется, что для всех $x \in X(M)$ множество $Q(x, f)$ пусто, то:

• стационарная стратегия $\pi(f) = \{f, f, f, \ldots\}$, построенная на основании вектора решений $f = \{f_x, x \in X(M)\}$, будет оптимальной стационарной стратегией;

• вектор $m_{\pi(f^*)} = \{m_{\pi(f^*)}(x), x \in X(M)\}$ средних выгод (прибылей), соответствующий оптимальной стационарной стратегии $\pi(f^*)$, определяется равенством

$$m_{\pi(f^*)}(x) = d(x) - V(x), x \in X(M).$$

Если окажется, что по крайней мере для одного состояния $x \in X(M)$ множество $Q(x, f)$ не будет пустым, то можно построить более оптимальный по сравнению с $f = \{f_x, x \in X(M)\}$ вектор решений $g = \{g_x, x \in X(M)\}$, используя следующий алгоритм.

Пусть

$$Q^* = \{x \in X(M) \colon Q(x, f) \neq \varnothing\}.$$

• Если $x \notin Q^*$, то положим $g_x = f_x$.

• Если $x \in Q^*$, то для вектора решений $g = \{g_x, x \in X(M)\}$ в состоянии $x \in X(M)$ выберем произвольное решение из множества $Q(x, f)$, например решение g_x, удовлетворяющее соотношение

$$\sum_{y \in X(M)} p^{*(g_x)}(x, y) \cdot d(y) = \min_{u=1, \ldots, r} \sum_{y \in X(M)} p^{*(u)}(x, y) \cdot d(y).$$

• На основании полученного таким образом вектора решений $g = \{g_x, x \in X(M)\}$ строим стационарную стратегию $\pi(g) = \{g, g, g, \ldots\}$ и затем возвращаемся к п. 2.

Изложенные подходы являются актуальными в условиях увеличения вычислительных мощностей и развития распределенных облачных вычислений, позволяющих оперировать большими объемами исторических данных. При этом важной особенностью предложенных в статье моделей является применимость в условиях неопределенности сроков реализации, характерных для многих социально-экономических и технологических процессов, которые формализуются марковскими процессами гибели и рождения.

Сфера возможного применения результатов широка и включает, в частности, реализацию в рамках систем принятия решений и бизнес-планирования в экономике, оптимизацию технологических процессов, поиск эффективных шагов повышения экологической безопасности, построение эффективных процедур медицинского обслуживания, повышение производительности труда.

СПИСОК ЛИТЕРАТУРЫ

1. Война А.О. Про вибір оптимальних стратегій в стохастичних маркетингових моделях. *Доповіді НАН України.* 2002. № 8. С. 60–64.

2. Война А.А. Оптимальные динамические стратегии для стохастических моделей маркетингового комплекса продвижения. *Кибернетика и системный анализ.* 2004. № 2. С. 145–152.

3. Война А.О. Стохастичні математичні моделі та функція ризику при побудові оптимальних маркетингових стратегій. *Журнал обчислювальної та прикладної математики.* 2004, № 89, вип. 2. С. 2–5.

4. Война А.А., Клодзинська А. Функция риска в многомерных моделях управления запасами, функционирующих в случайной марковской среде. *Кибернетика и системный анализ.* 2004. № 4. С. 150–155.

5. Война О.А. Керування ризиком в багатовимірних моделях страхування. *Журнал обчислювальної та прикладної математики.* 2007. № 2(95). С. 13–23.
6. Wojna A. Ryzyko w procesach finansowych oraz metody badań koniunktury. Politechnika Koszalińska (Polska), 2009. 446 s.
7. Гихман И. И., Скороход А.В. Управляемые случайные процессы. Киев: Наук. думка, 1977. 252 с.
8. Майн Х., Осаки С. Марковские процессы принятия решений. Москва: Наука, 1977. 176 с.

Надійшла до редакції 06.07.2017

О.А. Война, А.О. Война

ДИНАМІЧНЕ КЕРУВАННЯ РИЗИКОМ У БАГАТОВИМІРНИХ МАРКОВСЬКИХ МОДЕЛЯХ

Анотація. Розглянуто типову для багатьох галузей прикладної математики модель для вибору оптимальних рішень у стохастичних системах. Керований на проміжку випадкової довжини багатовимірний процес загибелі, що вивчається у статті, може бути використано для аналізу явищ, які виникають в економічних, біологічних, медичних, технічних та інших застосунках. Запропоновано декілька підходів до визначення критерію оптимальності і вивчено властивості відповідних функцій ризику, а також розроблено конструктивний обчислювальний алгоритм побудови оптимальних стаціонарних стратегій.

Ключові слова: оптимальні стаціонарні стратегії, керовані процеси, марковські процеси прийняття рішень, функція ризику, процеси загибелі та народження, ітераційні алгоритми.

O.A. Voina, A.O. Voyna

DYNAMIC RISK CONTROL IN MULTIDIMENSIONAL MARKOV MODELS

Abstract. Typical model used in many fields of applied mathematics regarding selection of optimal decisions in stochastic systems is considered. The multidimensional death process controlled on an interval of random length, which is analyzed in the paper, can be used to describe phenomena that appear in economical, biological, medical, engineering, and other applications. Several approached are proposed to determine the of optimality criteria, the properties of corresponding risk function are investigated, and a computing algorithm is developed for construction of optimal stationary strategies.

Keywords: optimal stationary strategies, controlled processes, Markov decision processes, risk function, birth-death processes, iterative algorithms.

Война Александр Андреевич,
доктор физ.-мат. наук, профессор факультета экономических наук Кошалинского политехнического института, Кошалин, Республика Польша, e-mail: avoina@hotmail.com.

Война Андрей Александрович,
магистр социальной информатики, заместитель директора Планово-финансового департамента ПАО «Проминвестбанк», Киев, e-mail: andrey.voyna@gmail.com.