УДК 519.21

Ал.А. ВОЙНА, Ан.А. ВОЙНА

МАТЕМАТИЧЕСКИЕ МОДЕЛИ УПРАВЛЕНИЯ РИСКОМ ДЛЯ МАРКОВСКИХ ПРОЦЕССОВ РЕГЕНЕРИРУЮЩЕГО ТИПА

Аннотация. В рамках математической модели, условно названной параллельной марковской структурой, формализован ряд практических постановок задач оптимального управления. Изучены свойства моделей и использованных для их построения случайных процессов. Разработаны конструктивные алгоритмы вычисления значений соответствующих функций риска и построения оптимальных стратегий управления.

Ключевые слова: регенерирующий процесс, параллельная структура, оптимальные стационарные стратегии, марковские процессы принятия решений, функция риска, итерационные алгоритмы.

Настоящая статья представляет методы построения оптимальных стратегий в стохастических моделях и является продолжением работы [1]. Возможности практического применения этих методов в различных сферах достаточно разнообразны. Прежде всего это экономические приложения: модели дистрибуции и построения маркетинговых стратегий, управления запасами, страхования, управления риском и др. Подобные задачи возникают также в технических приложениях, связанных, например, с контролем качества, в социологии для построения эффективных моделей социального взаимодействия, в медицине, биотехнологии и других областях.

Характерной особенностью большинства рассматриваемых авторами прикладных моделей является многократное повторение одних и тех же процессов. Последние можно представить с помощью понятия «параллельная структура». В отличие от [1], посвященной вопросам принятия оптимальных решений для каждого отдельно взятого повтора процесса, построенная в настоящей статье формальная стохастическая модель описывает все повторяемые процессы и анализирует явление в целом. В результате возникает необходимость выбора существенно иных методов для построения оптимальных решений. В [1] все сводилось к анализу соответствующих управляемых марковских процессов на конечном промежутке времени, изложенных, например, в [2–4]. В случае параллельной структуры при определении оптимальной стратегии управления используются методы анализа управляемых случайных процессов на бесконечном промежутке времени [3]. Это связано с тем, что большое число повторений требует значительного увеличения длины промежутка, в котором анализируется значение функции риска.

Рассмотрим вначале обобщение на случай параллельной структуры результатов в [5]. Характерной особенностью исследуемой в [5] модели является то, что состояние управляемого процесса становится известным только по окончании всего интервала управления. В отличие от модели, приведенной в [1], после каждого принятого решения не имелось возможности наблюдать его последствия.

Опишем формальное определение модели. Рассмотрим множество векторов $X(M)=\{(x_1,x_2,\dots,x_m): 0\le x_i \le M_i,\ i=1,2,\dots,m\}$, где $M=\{M_1,M_2,\dots,M_m\}$ — известный вектор с целочисленными положительными координатами. Предположим, что задано множество A, состоящее из r векторов:

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

при этом для каждого $u \in U$ справедливо $0 \le a_{ui} \le 1$, $i = 1, 2, \ldots, m$. Множество $U = \{1, 2, \ldots, r\}$ будем называть множеством управлений. Для каждого управления $u \in U$ определена однородная цепь Маркова $x^{(u)}(t)$ с дискретным временем $t = 1, 2, 3, \ldots$, со значениями во множестве $X(M)$ и следующими переходными вероятностями за один шаг:

$$p^{(u)}(x, y) = P\{x^{(u)}(t+1) = (y_1, y_2, \ldots, y_m) / x^{(u)}(t) = (x_1, x_2, \ldots, x_m)\} =$$
$$= C_{x_1}^{x_1 - y_1} (a_{u1})^{x_1 - y_1} (1 - a_{u1})^{y_1} \times$$
$$\times C_{x_2}^{x_2 - y_2} (a_{u2})^{x_2 - y_2} (1 - a_{u2})^{y_2} \times \ldots \times C_{x_m}^{x_m - y_m} (a_{um})^{x_m - y_m} (1 - a_{um})^{y_m},$$

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 $x \ge y$, т.е. для произвольного числа $i = 1, 2, \ldots, m$ выполняется неравенство $x_i \ge y_i$, и

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

если условие $x \ge y$ не выполняется, т.е. хотя бы для одной координаты i выполняется неравенство $x_i < y_i$.

Предположим, что выбор решений $u \in U$ по управлению системой случайный. Стратегией управления будем называть множество векторов

$$\Gamma(T) = \{\gamma(t) = \{\gamma_1(t), \gamma_2(t), \ldots, \gamma_r(t)\}, t \in \Delta\},$$

в которых координата $\gamma_u(t)$ является вероятностью выбора управления u из множества управлений $U = \{1, 2, \ldots, r\}$ в момент времени t из интервала управления Δ.

Если вероятности $\gamma_u(t) = \gamma_u$, $u \in \{1, 2, \ldots, r\}$, не зависят от момента t, то вектор $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_r\}$ назовем стационарной стратегией, а множество всех стационарных стратегий γ обозначим Γ.

Очевидно, что координаты стационарной стратегии γ имеют следующие свойства:

- $\gamma_1 \ge 0, \gamma_2 \ge 0, \ldots, \gamma_r \ge 0$; $\gamma_1 + \gamma_2 + \ldots + \gamma_r = 1$;
- если принятое в момент времени $t = 0, 1, 2, \ldots$ решение обозначить $\eta(t)$, то последовательность $\{\eta(t), t = 0, 1, 2, \ldots\}$ будет независимым от семейства цепей Маркова $\{x^{(u)}(t), u \in \{1, 2, \ldots, r\}\}$ случайным процессом с независимыми значениями. При этом для каждого $u \in U$ имеем

$$P\{\eta(t+1) = u / \eta(0) = u_0, \eta(1) = u_1, \ldots, \eta(t) = u_t,\} = \gamma_u, \ t = 0, 1, 2, \ldots$$

Пусть вектор $q = \{q(x), x \in X(M)\}$ определяет некоторое распределение на множестве $X(M)$, т.е. $q(x) \ge 0$ для любого $x \in X(M)$ и $\sum_{x \in X(M)} q(x) = 1.$

Предположим при этом, что выполнено условие $q(M) > 0$.

Случайным процессом, управляемым в соответствии со стационарной стратегией γ, назовем неоднородную цепь Маркова $\xi(\gamma, t) = \{\xi_1(\gamma, t), \xi_2(\gamma, t), \ldots$ $\ldots, \xi_m(\gamma, t)\}$ с дискретным временем $t = 1, 2, 3, \ldots$ и значениями во множестве $X(M)$, которая определяется следующим образом:

- начальное распределение процесса $\xi(\gamma, t)$, $t = 0, 1, 2, \ldots$, определяется вектором $q = \{q(x), x \in X(M)\}$, т.е. $P\{\xi(\gamma, 0) = x\} = q(x)$, $x \in X(M)$;
- переходная вероятность процесса $\xi(\gamma, t)$ в момент времени t определяется

$$p_{(\xi(\gamma,t))}(x, y) =$$
$$= P\{\xi(\gamma, t+1) = (y_1, y_2, \dots, y_m) / \xi(\gamma, t) = (x_1, x_2, \dots, x_m)\} = p^{(\eta(t))}(x, y).$$

В работе [5] рассмотрено задачу нахождения оптимальной стационарной стратегии $\gamma^* \in \Gamma$ при условии, что последствия принимаемых решений будут известны только после завершения интервала управления $\Delta_\tau = \{0, 1, 2, \dots, \tau\}$ случайной длины τ. Оптимальная стационарная стратегия γ^* находилась как решение соответствующей задачи нелинейного программирования и определялась из условия минимизации $L(\gamma^*) = \inf\limits_{\gamma\in\Gamma}[L(\gamma)]$ значения следующей функции риска: $L(\gamma) = E\left(\sum\limits_{i=1}^{m} c_i \cdot \xi_i(\gamma, \tau)\right)$, где $c = (c_1, c_2, \dots, c_m)$ — заданный неотрицательный числовой вектор весовых коэффициентов, т.е. $c_i \geq 0$ для любого $i = 1, 2, \dots, m$.

Предположим теперь, что тот же «цикл» управления случайной длины τ многократно повторяется, причем эти повторения происходят независимо одно от другого. Пусть число таких повторений равно S. Тогда каждое повторение можно описать процессом

$$\xi^{(s)}(\gamma, t) = \{\xi_1^{(s)}(\gamma, t), \xi_2^{(s)}(\gamma, t), \dots, \xi_m^{(s)}(\gamma, t)\}, \; t = 0, 1, 2, \dots, \; s = 1, 2, \dots, S,$$

который строится подобно процессу $\xi(\gamma, t) = \{\xi_1(\gamma, t), \xi_2(\gamma, t), \dots, \xi_m(\gamma, t)\}$, $t = 0, 1, 2, \dots$

Построенный на основании семейства случайных процессов $\{\xi^{(s)}(\gamma, t)$, $s = 1, 2, \dots, S\}$ процесс

$$\zeta(\gamma, t) = \{\zeta_1(\gamma, t), \zeta_2(\gamma, t), \dots, \zeta_m(\gamma, t)\}, \; t = 0, 1, 2, \dots,$$

описываюший одновременно всю параллельную структуру, характеризуется тем, что в случайные моменты времени он возвращается в первоначальное состояние, распределение которого совпадает с распределением $\xi(\gamma, 0)$.

Если предположить, что момент окончания цикла управления с номером s является одновременно моментом начала цикла управления с номером $s+1$, то эти моменты можно назвать моментами регенерации. Поэтому рассматриваемую задачу построения оптимальной стратегии в параллельной структуре можно условно назвать управлением многомерным марковским процессом регенерирующего типа [6].

Пусть $\tau(1), \tau(2), \dots, \tau(n), \dots$ обозначает последовательность независящих от семейства процессов $\{\xi^{(s)}(\gamma, t), s = 1, 2, \dots, S\}$ и независимых между собой случайных величин, каждая из которых имеет такое же распределение, как и случайная длина τ интервала управления $\Delta_\tau = \{0, 1, 2, \dots, \tau\}$. Построим точечный случайный процесс $\theta_n, n = 0, 1, 2, \dots$, следующим образом:

$$\theta_0 = 0; \; \theta_1 = \tau(1); \; \theta_2 = \tau(1) + \tau(2); \; \dots; \; \theta_n = \tau(1) + \tau(2) + \dots + \tau(n); \; \dots$$

В интервале $t \in [0, 1, 2, \dots, \tau(1) - 1]$ положим по определению $\zeta(\gamma, t) = \xi(\gamma, t)$. Случайная величина $\tau(s)$ определяет продолжительность цикла управления с номером s, а построенный на основании последовательности случайных величин $\tau(1), \tau(2), \dots, \tau(n), \dots$ точечный случайный процесс θ_n, $n = 0, 1, 2, \dots$, является процессом восстановления и определяет случайные моменты времени, в которые процесс $\zeta(\gamma, t)$, $t = 0, 1, 2, \dots$, возвращается в начальное состояние $\xi(\gamma, 0)$.

Обозначим $\nu(t)$ количество восстановлений процесса θ_n, $n = 0, 1, 2, \dots$, на отрезке времени $[0, t]$, т.е. $\nu(t) = \max\{n \geq 0, \theta_n \leq t\}$, и предположим, что случайная величина τ, которая определяет длину интервала управления $\Delta_\tau = \{0, 1, 2, \ldots, \tau\}$, такова, что $\nu(T) \underset{T \to \infty}{\to} \infty$.

Случайный процесс $\zeta(\gamma, t)$, $t = 0, 1, 2, \ldots$, описывающий одновременно всю параллельную структуру на бесконечном промежутке времени ее функционирования, можно определить равенством

$$\zeta(\gamma, t) = \xi(\gamma, t - \theta_{\nu(t)}), \ \ t = 0, 1, 2, \ldots$$

По аналогии с [5] в качестве целевой функции $L_S(\gamma)$ для определения оптимальной стационарной стратегии можно выбрать следующее математическое ожидание:

$$L_S(\gamma) = E\left\{\sum_{s=1}^{S} \sum_{i=1}^{m} c_i \cdot \zeta_i(\gamma, \theta_s - 1)\right\}.$$

Когда количество циклов управления S параллельной структуры возрастет, то, с одной стороны, будет расти величина $L_S(\gamma)$, а с другой — длина $T_s = \theta_s = \tau(1) + \tau(2) + \ldots + \tau(S)$ случайного интервала управления $\Delta_{T_s} = \{0, 1, 2, \ldots, T_s\}$. Поэтому в случае, когда S достаточно велико, вместо анализа управляемого процесса на конечном промежутке можно использовать один из методов анализа на бесконечном промежутке [3]. Предположим, что процесс $\zeta(\gamma, t)$ управляется на промежутке $\Delta_T = \{0, 1, 2, \ldots, T\}$ длины T. Поскольку согласно предположениям вектор весовых коэффициентов $c = (c_1, c_2, \ldots, c_m)$ неотрицательный и общий для всех циклов управления, для определения целевой функции $\widetilde{L}(\gamma)$ введем вначале случайную величину

$$L(\gamma, T) = \sum_{s=1}^{\nu(T)} \sum_{i=1}^{m} c_i \cdot \zeta_i(\gamma, \theta_s - 1),$$

затем определим «средний прирост» $\dfrac{L(\gamma, T)}{\nu(T)}$ значения $L(\gamma, T)$ в течение одного цикла управления. Поскольку согласно предположению $\nu(T) \underset{T \to \infty}{\to} \infty$, в качестве целевой функции $\widetilde{L}(\gamma)$ целесообразно выбрать предельное значение «среднего прироста» при $T \to \infty$. Использовать такой подход можно только при условии эргодичности марковского процесса $\zeta(\gamma, t)$, $t = 0, 1, 2, \ldots$, т.е. при условии, что для процесса, описывающего одновременно всю параллельную структуру, существует стационарный режим. Тогда на основании усиленного закона больших чисел для эргодических конечных цепей Маркова [8] можно утверждать, что с вероятностью 1 существует следующий предел:

$$\widetilde{L}(\gamma) = \lim_{T \to \infty} \frac{L(\gamma, T)}{\nu(T)} = \lim_{T \to \infty}\left(\frac{1}{\nu(T)} \cdot \sum_{s=1}^{\nu(T)} \sum_{i=1}^{m} c_i \cdot \zeta_i(\gamma, \theta_s - 1)\right).$$

Зная стационарное распределение случайного процесса $\zeta(\gamma, t)$, $t = 0, 1, 2, \ldots$, можно определить значение $\widetilde{L}(\gamma)$ целевой функции в стационарном режиме, и поиск оптимальной стационарной стратегии γ^* сведется к выбору такого управления, при котором значение $\widetilde{L}(\gamma^*)$ целевой функции будет оптимальным, т.е. $\widetilde{L}(\gamma^*) = \inf_{\gamma \in \Gamma} [\widetilde{L}(\gamma)]$.

Анализируя для каждого управления $u \in U$ вид переходных вероятностей $p^{(u)}(x, y)$ многомерных цепей Маркова $x^{(u)}(t) = \{x_1^{(u)}(t), x_2^{(u)}(t), \ldots, x_m^{(u)}(t)\}$, приходим к выводу, что его компоненты $x_i^{(u)}(t)$ — независимые случайные процессы. Согласно принятым предположениям точечный случайный процесс θ_n, $n=0,1,2,\dots$, который используется при определении случайного процесса $\zeta(\gamma,t)$, $t=0,1,2,\dots$, не зависит от семейства цепей Маркова $x^{(u)}(t)$, $u\in U=\{1,2,\dots,r\}$. Поэтому изучение эргодичности марковского процесса $\zeta(\gamma,t)$, $t=0,1,2,\dots$, фактически сводится к исследованию свойств отдельных координат $\zeta_i(\gamma,t)$, $i=1,2,\dots,m$, и условий их эргодичности.

Для любого $i=1,2,\dots,m$ координата $\xi_i(\gamma,t)$ многомерного марковского процесса $\xi(\gamma,t)=\{\xi_1(\gamma,t),\xi_2(\gamma,t),\dots,\xi_m(\gamma,t)\}$ является однородной цепью Маркова с конечным множеством состояний $\{0,1,2,\dots,M_i\}$. Определим переходную вероятность $p_{(i)}(x_i,y_i)$, $0\le x_i\le M_i$, $0\le y_i\le M_i$, этой цепи за один шаг.

Очевидно, что если $x_i<y_i$, то $p_{(i)}(x_i,y_i)=0$.

Пусть теперь $x_i\ge y_i$. Введем полную группу событий $\{H_1,H_2,\dots,H_r\}$ для фиксированного значения t следующим образом: $H_u=\{\eta(t)=u\}$, $u\in\{1,2,\dots,r\}$. Тогда, используя формулу полной вероятности, получаем

$$p_{(i)}(x_i,y_i)=P\{\xi_i(t+1)=y_i\,/\,\xi_i(t)=x_i\}=$$

$$=\sum_{u=1}^{r}P\{H_u\}\times P\{\xi_i(t+1)=y_i\,/\,\xi_i(t)=x_i,H_u\}=$$

$$\sum_{u=1}^{r}\gamma_u\times P\{\xi_i(t+1)=y_i\,/\,\xi_i(t)=x_i,H_u\}=\sum_{u=1}^{r}\gamma_u\times C_{x_i}^{x_{1i}-y_i}(a_{ui})^{x_i-y_i}(1-a_{ui})^{y_i}.$$

Пусть $P^{(i)}=\|p_{(i)}(x_i,y_i)\|$, $0\le x_i\le M_i$, $0\le y_i\le M_i$, обозначает матрицу переходных вероятностей за один шаг цепи $\xi_i(\gamma,t)$, $i=1,2,\dots,m$. Обозначим $Q^{(i)}$ матрицу размерности, аналогичной $P^{(i)}$, которая состоит из одинаковых строк $q^{(i)}=\{q^{(i)}(x_i),0\le x_i\le M_i\}$, при этом вектор $q^{(i)}$ определяет начальное распределение координаты $\xi_i(\gamma,t)$, $i=1,2,\dots,m$, многомерного марковского процесса $\xi(\gamma,t)$, т.е.

$$q^{(i)}(x_i)=P\{\xi_i(\gamma,0)=x_i\},\ 0\le x_i\le M_i;\ q^{(i)}(x_i)\ge 0;\ \sum_{x_i=0}^{M_i}q^{(i)}(x_i)=1.$$

Введем случайный процесс $\gamma_t^+=\theta_{\nu(t)+1}-t$, определяющий время, прошедшее от текущего момента t до очередного момента восстановления. Очевидно, что $\gamma_t^+\ge 1$, $\gamma_t^+\in\{1,2,3,\dots\}$. В теории восстановления процесс γ_t^+ называется перескоком процесса восстановления $\theta_n, n=0,1,2,\dots$, через уровень t.

Теорема 1. Для произвольного $i=1,2,\dots,m$ координата $\zeta_i(\gamma,t)$ случайного процесса $\zeta(\gamma,t)=\{\zeta_1(\gamma,t),\zeta_2(\gamma,t),\dots,\zeta_m(\gamma,t)\}$, $t=0,1,2,\dots$, который описывает одновременно всю параллельную структуру, является неоднородной цепью Маркова с дискретным временем $t=0,1,2,\dots$ и множеством состояний $\{0,1,2,\dots,M_i\}$.

Начальное распределение процесса $\zeta_i(\gamma,t)$ определяется вектором $q^{(i)}=\{q^{(i)}(x_i),0\le x_i\le M_i\}$, $q^{(i)}(x_i)=P\{\zeta_i(\gamma,0)=x_i\}$, $0\le x_i\le M_i$.

Матрица $\widetilde{P}^{(i)}$ переходных вероятностей процесса $\zeta_i(\gamma,t)$ за один шаг в момент времени t:

$$\widetilde{P}^{(i)}=\|\widetilde{p}_{(i)}(x_i,y_i)\|,\ 0\le x_i\le M_i,\ 0\le y_i\le M_i,$$

где

$$\widetilde{p}_{(i)}(x_i,y_i)=P\{\zeta_i(t+1)=y_i\,/\,\zeta_i(t)=x_i\},$$

определяется равенством $\widetilde{P}^{(i)}=P^{(i)}\cdot P\{\gamma_t^+>1\}+Q^{(i)}\cdot P\{\gamma_t^+=1\}$.

Если предположить, что с положительной вероятностью ρ в каждый момент t процесс управления может прекратиться (т.е. длина τ интервала управления $T = \{0, 1, 2, \ldots, \tau\}$ определяется следующим образом: каждый очередной момент времени $t = 0, 1, 2, \ldots$ с положительной вероятностью ρ, которая не зависит ни от состояния управляемого процесса $\xi(\gamma, t)$ в этот момент, ни от принимаемых ранее решений, может быть последним), то справедливыми будут следующие утверждения:

• процесс $\zeta_i(\gamma, t)$ является однородной эргодической цепью Маркова, т.е. для произвольного значения $k \in \{0, 1, 2, \ldots, M_i\}$ существует предел

$$\tilde{\pi}^{(i)}(k) = \lim_{t\to\infty} \tilde{p}_{(i)}^{(t)}(k, x_i) = \lim_{t\to\infty} P\{\zeta_i(t) = k / \zeta_i(0) = x_i\},\ 0 \le x_i \le M_i,\ 0 \le k \le M_i;$$

• стационарное распределение $\tilde{\pi}^{(i)} = \{\tilde{\pi}^{(i)}(k), 0 \le k \le M_i\}$ процесса $\zeta_i(\gamma, t)$ удовлетворяет уравнению $\tilde{\pi}^{(i)} \cdot [E^{(i)} - (1-\rho) \cdot P^{(i)}] = \rho \cdot q^{(i)}$, в котором $E^{(i)}$ обозначает единичную матрицу такой же размерности, как и $P^{(i)}$.

Доказательство. Как следует из определения случайного процесса $\zeta(\gamma, t)$, $t = 0, 1, 2, \ldots$, независимо от его траектории $\{\zeta_i(\gamma, 0), \zeta_i(\gamma, 1), \ldots, \zeta_i(\gamma, \theta_k - 1)\}$ справедливо равенство

$$q^{(i)}(x_i) = P\{\zeta_i(\gamma, 0) = x_i\} = P\{\zeta_i(\gamma, \theta_k) = x_i\},\ 0 \le x_i \le M_i.$$

Предположим, что на отрезке времени $[0, t]$ произошло $\nu(t)$ восстановлений случайного процесса θ_n, $n = 0, 1, 2, \ldots$ Тогда, принимая во внимания строгую марковость случайных процессов $\xi_i(\gamma, t)$, $i = 1, 2, \ldots, m$, получаем

$$P\{\zeta_i(\gamma, t) \in A_i / \zeta_i(\gamma, 0), \zeta_i(\gamma, 1), \ldots, \zeta_i(\gamma, t-1), \zeta_i(\gamma, t)\} =$$
$$= P\{\zeta_i(\gamma, t) \in A_i / \zeta_i(\gamma, \theta_{\nu(t)}), \zeta_i(\gamma, \theta_{\nu(t)} + 1), \ldots, \zeta_i(\gamma, t-1), \zeta_i(\gamma, t)\}.$$

Для перескока $\gamma_t^+ = \theta_{\nu(t)+1} - t$ два события: $\{\gamma_t^+ = 1\}$ и $\{\gamma_t^+ > 1\}$, составляют полную группу. При этом случай, когда $\{\gamma_t^+ = 1\}$, означает, что $\nu(t+1) = \nu(t) + 1$. Соответственно событие $\{\gamma_t^+ > 1\}$ влечет за собой равенство $\nu(t+1) = \nu(t)$. Поэтому, используя предположение, что процесс $\nu(t)$ не зависит от $\xi_i(\gamma, t)$, после несложных преобразований получаем

$$P\{\zeta_i(\gamma, t+1) = y_i / \zeta_i(\gamma, 0), \zeta_i(\gamma, 1), \ldots, \zeta_i(\gamma, t-1), \zeta_i(\gamma, t)\} =$$
$$= P\{\gamma_t^+ > 1\} \times P\{\xi_i(\gamma, t+1) = y_i / \xi_i(\gamma, t)\} + P\{\gamma_t^+ = 1\} \times P\{\xi_i(\gamma, 0) = y_i\}.$$

В то же время с помощью подобных рассуждений выводим соотношение

$$P\{\zeta_i(\gamma, t+1) = y_i / \zeta_i(\gamma, t)\} =$$
$$= P\{\gamma_t^+ > 1\} \times P\{\xi_i(\gamma, t+1) = y_i / \xi_i(\gamma, t)\} + P\{\gamma_t^+ = 1\} \times P\{\xi_i(\gamma, 0) = y_i\},$$

т.е.

$$P\{\zeta_i(\gamma, t+1) = y_i / \zeta_i(\gamma, 0), \zeta_i(\gamma, 1), \ldots, \zeta_i(\gamma, t-1), \zeta_i(\gamma, t)\} =$$
$$= P\{\zeta_i(\gamma, t+1) = y_i / \zeta_i(\gamma, t)\},$$

что и доказывает марковость процесса $\zeta_i(\gamma, t)$, $t = 0, 1, 2, \ldots$

Зафиксируем $\xi_i(\gamma, t) = x_i$ значение процесса $\xi_i(\gamma, t)$ в момент времени t. Используя определение процесса $\zeta_i(\gamma, t)$, получаем

$$P\{\zeta_i(\gamma, t+1) = y_i / \zeta_i(\gamma, t) = x_i\} =$$
$$= P\{\gamma_t^+ > 1\} \times P\{\xi_i(\gamma, t+1) = y_i / \xi_i(\gamma, t) = x_i\} + P\{\gamma_t^+ = 1\} \times P\{\xi_i(\gamma, 0) = y_i\} =$$
$$= P\{\gamma_t^+ > 1\} \times p_{(i)}(x_i, y_i) + P\{\gamma_t^+ = 1\} \times q^{(i)}(y_i).$$

Принимая во внимание тот факт, что вероятности $P\{\gamma_t^+ > 1\}$ и $P\{\gamma_t^+ = 1\}$ не зависят от значений x_i и y_i, записываем последнее соотношение в матричном виде

$$\widetilde{P}^{(i)} = P^{(i)} \cdot P\{\gamma_t^+ > 1\} + Q^{(i)} \cdot P\{\gamma_t^+ = 1\},$$

что и доказывает приведенное в теореме 1 равенство.

Пусть теперь выполнено предположение теоремы 1, т.е. в каждый момент времени t с положительной вероятностью ρ процесс управления может прекратиться. Это означает, что длина τ интервала управления $T = \{0, 1, 2, \ldots, \tau\}$ является дискретной случайной величиной, которая не зависит от семейства цепей Маркова $x^{(u)}(t)$, $u \in \{1, 2, \ldots, r\}$, и имеет геометрическое распределение с параметром ρ, а именно

$$\tau \in \{1, 2, 3, \ldots\}, \; P\{\tau = n\} = (1-\rho)^{n-1} \cdot \rho, \; n = 1, 2, 3, \ldots$$

Для распределения $h(t, k) = P\{\gamma_t^+ = k\}$, $t = 0, 1, 2, \ldots$, $k = 1, 2, \ldots$, перескока γ_t^+ можно записать соотношение

$$h(t, k) = P\{\tau(1) = t + k\} + \sum_{l=1}^{t} P\{\tau(1) = l\} \cdot h(t-l, k), \; k = 1, 2, \ldots$$

Если ввести производящие функции

$$H_k(s) = E(s^{\gamma_t^+}) = \sum_{t=0}^{\infty} s^t \cdot h(t, k), \; k = 1, 2, \ldots, \; F(s) = E[s^{\tau}] = \sum_{l=1}^{\infty} s^l \cdot P\{\tau = l\},$$

то это соотношение примет вид

$$H_k(s) = (1-\rho)^{k-1} \cdot \rho + (1-\rho)^k \cdot F(s) + F(s) \cdot H_k(s).$$

Использовав свойства геометрического распределения, легко получим равенство $\dfrac{F(s)}{1-F(s)} = \dfrac{s \cdot \rho}{1-s}$. Поэтому

$$H_k(s) = (1-\rho)^{k-1} \times \left[\rho + \frac{F(s)}{1-F(s)}\right] = \sum_{t=0}^{\infty} s^t \cdot [\rho \cdot (1-\rho)^{k-1}], \; k = 1, 2, \ldots$$

Другими словами, $\widetilde{P}^{(i)} = (1-\rho) \cdot P^{(i)} + \rho \cdot Q^{(i)}$. Из этого следует, что переходные вероятности $\widetilde{p}_{(i)}(x_i, y_i)$ не зависят от момента времени t, т.е. $\zeta_i(\gamma, t)$, $t = 0, 1, 2, \ldots$, — однородная цепь Маркова. Поскольку $\widetilde{p}_{(i)}(M_i, M_i) = (1-\rho) \cdot p_{(i)}(M_i, M_i) + \rho \cdot q^{(i)}(M_i)$ и на основании принятых условий $q(M) > 0$, то $\widetilde{p}_{(i)}(M_i, M_i) > 0$. Поэтому случайный процесс $\zeta_i(\gamma, t)$ является непериодической, неприводимой конечной цепью Маркова. Эти условия гарантируют ее эргодичность. При этом стационарное распределение $\widetilde{\pi}^{(i)} = \{\widetilde{\pi}^{(i)}(k), 0 \le k \le M_i\}$ цепи $\zeta_i(\gamma, t)$ является единственным решением системы уравнений

$$\begin{cases} \widetilde{\pi}^{(i)} = \widetilde{\pi}^{(i)} \cdot \widetilde{P}^{(i)}, \\ \widetilde{\pi}^{(i)} \cdot I^{(i)} = 1, \end{cases}$$

где $I^{(i)}$ обозначает состоящий из единиц вектор-столбец такой же размерности, что и вектор $\widetilde{\pi}^{(i)} = \{\widetilde{\pi}^{(i)}(k), 0 \le k \le M_i\}$.

Поскольку координаты $\widetilde{\pi}^{(i)}(k)$, $0 \le k \le M_i$, вектора $\widetilde{\pi}^{(i)}$ имеют следующие свойства: $\widetilde{\pi}^{(i)}(k) \ge 0$; $\sum_{k=0}^{M_i} \widetilde{\pi}^{(i)}(k) = 1$, а все строки матрицы $Q^{(i)}$ совпадают с вектором $q^{(i)} = \{q^{(i)}(x_i), 0 \le x_i \le M_i\}$ начального распределения, очевидно, справедливо равенство $\widetilde{\pi}^{(i)} \cdot Q^{(i)} = q^{(i)}$. Поэтому $\widetilde{\pi}^{(i)} = (1-\rho) \cdot \widetilde{\pi}^{(i)} \cdot P^{(i)} + \rho \cdot q^{(i)}$ или окончательно $\widetilde{\pi}^{(i)} \cdot [E - (1-\rho) \cdot P^{(i)}] = \rho \cdot q^{(i)}$, что и завершает доказательство теоремы 1.

Решим задачу динамического управления в параллельной структуре, которая изучалась в [1, 7]. Для этого изменим рассмотренную в теореме 1 модель следующим образом.

Предположим, что после каждого принятого решения можно проконтролировать его результат, т.е. наблюдать состояние, в которое переходит управляемый процесс.

Случайные процессы $\{x^{(u)}(t) = (x_1^{(u)}(t), x_2^{(u)}(t), \dots, x_m^{(u)}(t)), u \in U = \{1, 2, \dots, r\}\}$ характеризуются тем, что их траектории монотонные. Поэтому, если предположить, что для каждого $u \in U$ справедливо $x^{(u)}(0) = M$, то пространством состояний всех цепей Маркова $x^{(u)}(t)$, $u \in U$, будет множество $X(M) = \{(x_1, x_2, \dots, x_m): 0 \le x_i \le M_i, i = 1, 2, \dots, m\}$. Как и в [1], определим множество историй $i(t) = \{x^{(0)}, u^{(1)}, x^{(1)}, u^{(2)}, x^{(2)}, \dots, u^{(t)}, x^{(t)}\} \in I(t)$ длины $t \in \{0, 1, 2, \dots\}$, а также стратегию $\pi = \{f(1), f(2), f(3), \dots\}$, построенную на основании семейства векторов решений $f(t) = \{f_x(t), x \in X(M)\}$, $t \in \{1, 2, \dots\}$. Стратегия $\pi(f) = \{f, f, f, \dots\}$, в которой вектор решений $f(t) = f = \{f_x, x \in X(M)\}$ не зависит от $t \in T$, называется стационарной. Оптимальное управление будем искать во множестве Π всех стационарных стратегий.

Обозначим $\zeta^{(f)}(t) = \{\zeta_1^{(f)}(t), \zeta_2^{(f)}(t), \dots, \zeta_m^{(f)}(t)\}$, $t = 0, 1, 2, \dots$, управляемый случайный процесс, который описывает одновременно всю параллельную структуру на бесконечном промежутке времени ее функционирования. Положим по определению, что с вероятностью 1 $\zeta^{(f)}(0) = (M_1, M_2, \dots, M_m)$. Пусть $P_{(f)} = \| p_{(f)}(x, y) \|$, $x = (x_1, x_2, \dots, x_m) \in X(M)$, $y = (y_1, y_2, \dots, y_m) \in X(M)$, обозначает матрицу переходных функций процесса $\zeta^{(f)}(t)$ за один шаг. Тогда вероятность $p_{(f)}(x, y)$ определяется следующим образом:

- для произвольного состояния $x = (x_1, x_2, \dots, x_m) \in X(M)$ справедливо $p_{(f)}(x, M) = \rho$;
- если для состояния $y \in X(M)$, $y = (y_1, y_2, \dots, y_m) \ne M$, существует такое i, что $x_i < y_i$, то $p_{(f)}(x, y) = 0$;
- если состояние $y \in X(M)$, $y = (y_1, y_2, \dots, y_m) \ne M$, такое, что для произвольного $i = 1, 2, \dots, m$ выполняется $x_i - y_i \ge 0$, то $p_{(f)}(x, y) = (1-\rho) \times \prod_{i=1}^{m} C_{x_1}^{x_1 - y_1} (a_{f_{x_i}})^{x_1 - y_1} \times (1 - a_{f_{x_i}})^{y_1}$.

Определим последовательность случайных моментов времени $s_0, s_1, s_2, \dots$ следующим образом: $s_0 = 0$, $s_k = \min\{n : n > s_{k-1};\ \zeta^{(f)}(n) = (M_1, M_2, \dots, M_m)\}$, $k = 1, 2, \dots$ Другими словами, $s_0, s_1, s_2, \dots$ — последовательные моменты возвращения процесса $\zeta^{(f)}(t)$ в состояние M. Поскольку согласно определению процесса $\zeta^{(f)}(t)$ возвращение его в состояние M фактически означает окончание очередного цикла управления случайной длины τ_ρ, траектория $\{\zeta^{(f)}(s_k), \zeta^{(f)}(s_k + 1), \dots, \zeta^{(f)}(s_{k+1} - 1)\}$ процесса $\zeta^{(f)}(t)$ в интервале $s_k \le t < s_{k+1}$ описывает развитие $(k+1)$-го по порядку цикла управления. Предположим, что задан вектор $d = (d_1, d_2, \dots, d_m)$ весовых коэффициентов, и определим случайный функционал $L_{\{\pi(f)\}}(T)$ в интервале управления $0 \le t < T$ следующим равенством:

$$L_{\{\pi(f)\}}(T) = \sum_{t=0}^{T} \left(\sum_{i=1}^{m} d_i \cdot (\zeta_i^{(f)}(t) - \zeta_i^{(f)}(t+1)) \cdot \chi[\zeta^{(f)}(t+1) \ne M] \right).$$

Здесь $\chi[H]$ обозначает индикатор события H, т.е. $\chi[H]=1$, если оно происходит, и $\chi[H]=0$ в противном случае. Функционал $L_{\{\pi(f)\}}(T)$ можно интерпретировать, как выгоду (прибыль) от управления процессом $\zeta^{(f)}(t)$ в соответствии со стационарной стратегией $\pi(f)$ в интервале $0 \le t < T$. Очевидно, что увеличение количества S повторений цикла управления влечет за собой возрастание промежутка $0 \le t < T$, необходимого для их реализации. Поэтому качество стационарной стратегии $\pi(f)$ можно оценивать средней выгодой за один шаг управления. Другими словами, в качестве критерия оптимальности стационарной стратегии $\pi(f)$ можно выбрать функционал

$$\hat{L}_{\{\pi(f)\}} = \lim_{T\to\infty} \frac{1}{T} L_{\{\pi(f)\}}(T).$$

Стационарную стратегию $\pi^* = \{f^*, f^*, f^*, \dots\}$ будем называть оптимальной, если

$$\hat{L}_{\{\pi^*\}} = \max_{\pi(f)\in\Pi} \{\hat{L}_{\{\pi(f)\}}\}.$$

Теорема 2. Функция риска для стационарной стратегии $\pi(f)$ вычисляется следующим образом:

$$\hat{L}_{\{\pi(f)\}} = \sum_{x\in X(M)} \sum_{\substack{y\in X(M)\\ y\ne M}} q_{\{f\}}(x) \cdot p_{\{f\}}(x, y) \times \left[\sum_{i=1}^{m} d_i \cdot (x_i - y_i)\right],$$

при этом $x=(x_1, x_2, \dots, x_m) \in X(M)$, $y=(y_1, y_2, \dots, y_m) \in X(M)$, а вектор $\hat{q}_{(f)} = \{\hat{q}_{(f)}(x), x \in X(M)\}$ является единственным решением системы уравнений

$$\begin{cases} \hat{q}_{\{f\}} \cdot P_{\{f\}} = \hat{q}_{\{f\}}, \\ \hat{q}_{\{f\}} \cdot I_{\{f\}} = 1, \end{cases}$$

где $I_{\{f\}}$ — единичный вектор-столбец, размерность которого совпадает с размерностью вектора $q_{(f)}$.

Доказательство. Процесс $\zeta^{(f)}(t)$ является цепью Маркова з дискретным временем и множеством состояний $X(M)$. Анализируя его переходную функцию за один шаг $p_{(f)}(x, y)$, легко приходим к выводу, что цепь $\zeta^{(f)}(t)$ непериодическая и неприводимая и при этом все ее состояния из множества $X(M)$ существенны. Поскольку множество $X(M)$ конечно, на основании эргодической теоремы для конечных цепей Маркова существует эргодическое распределение $\hat{q}_{(f)} = \{\hat{q}_{(f)}(x), x \in X(M)\}$ процесса $\zeta^{(f)}(t)$. Это означает, что для произвольного состояния $y=(y_1, y_2, \dots, y_m) \in X(M)$ существует независящий от начального состояния $x=(x_1, x_2, \dots, x_m) \in X(M)$ предел

$$\lim_{n\to\infty} P\{\zeta^{(f)}(n) = (y_1, y_2, \dots, y_m) / \zeta^{(f)}(0) = (x_1, x_2, \dots, x_m)\} = \hat{q}_{(f)}(y).$$

При этом вектор эргодического распределения $\hat{q}_{(f)} = \{\hat{q}_{(f)}(x), x \in X(M)\}$ процесса $\zeta^{(f)}(t)$ одновременно является его стационарным распределением и его можно найти как единственное решение приведенной в теореме 2 системы уравнений. Окончательное доказательство теоремы 2 следует из усиленного закона больших чисел для эргодических конечных цепей Маркова [8].

Оптимальную стационарную стратегию $\pi^* = \{f^*, f^*, f^*, \dots\}$ можно найти простым перебором. Согласно принятым условиям необходимо определить максимальное значение функционала качества $\hat{L}_{\{\pi(f)\}}$ на конечном множестве всех стационарных стратегий. При этом теорема 2 дает практический алгоритм определения этого значения для произвольной конкретной стационарной стратегии $\pi = \{f, f, f, \dots\}$.

Такой подход можно использовать в ситуации, когда количество возможных решений r, размерность m процесса $\zeta^{(f)}(t)$ и значения координат $M_1, M_2, \dots, M_m$ вектора M не очень велики. Если, кроме того, предметный анализ исследуемой ситуации позволяет ограничить множество возможных и разумных з практической точки зрения стратегий до небольшого их количества, то предложенный алгоритм прямого перебора может быть вполне приемлемым. Очевидно, однако, что в большинстве случаев такой подход нельзя реализовать на практике, поскольку количество стационарных стратегий очень быстро растет при увеличении значений параметров модели. Поэтому, чтобы построить эффективный вычислительный алгоритм для поиска оптимальной стационарной стратегии, найдем вначале критерий ее оптимальности.

Обозначим $M_{\{f\}}(T) = E[L_{\{\pi(f)\}}(T)]$ среднюю выгоду, полученную за T шагов управления процессом $\zeta^{(f)}(t)$ в интервале $0 \le t < T$. В свою очередь $m_{\{f\}}(T) = \{m_{\{f\}}(T, x), x \in X(M)\}$ обозначает вектор-столбец, размерность которого совпадает с размерностью пространства $X(M)$, а координаты $m_{\{f\}}(T, x)$ определяют условную среднюю выгоду $m_{\{f\}}(T, x) = E[L_{\{\pi(f)\}}(T) / \zeta^{(f)}(0) = x]$.

Если вектор $q = \{q(x), x \in X(M)\}$ определяет начальное распределение случайного процесса $\zeta^{(f)}(t)$, т.е. $q(x) \ge 0$, $\sum\limits_{x \in X(M)} q(x) = 1$, $q(x) = P\{\xi(\gamma, 0) = x\}$, $x \in X(M)$, то очевидно $M_{\{f\}}(T) = \sum\limits_{x \in X(M)} g(x) \cdot m_{\{f\}}(T, x)$.

Введем вектор-столбец $d_{\{f\}} = \{d_{\{f\}}(x), x \in X(M)\}$, размерность которого совпадает с размерностью пространства $X(M)$, а координаты $d_{\{f\}}(x)$ определяются равенством

$$d_{\{f\}}(x) = \sum_{\substack{y \in X(M) \\ y \ne M}} p_{\{f\}}(x, y) \left(\sum_{i=1}^{m} d_i \cdot (x_i - y_i) \right).$$

Обозначим $\hat{Q}_{\{f\}} = || \hat{q}_{\{f\}} ||$ квадратную матрицу, сложенную из одинаковых строк, которые совпадают со стационарным распределением $\hat{q}_{(f)} = \{\hat{q}_{(f)}(x), x \in X(M)\}$, и введем матрицу $H_{\{f\}} = (E_{\{f\}} - P_{\{f\}} + Q_{\{f\}})^{-1} - Q_{\{f\}}$, где $E_{\{f\}}$ — единичная матрица такой же размерности, как $Q_{\{f\}}$.

Теорема 3. Пусть $f = \{f_x, x \in X(M)\}$ и $g = \{g_x, x \in X(M)\}$ — два вектора решений, $\pi(f) = \{f, f, f, \dots\}$ и $\pi(g) = \{g, g, g, \dots\}$ — соответствующие им стационарные стратегии.

Если выполняется неравенство $d_{(g)} + P_{(g)} \cdot H_{(f)} \cdot d_{(f)} > d_{(f)} + P_{(f)} \cdot H_{(f)} \cdot d_{(f)}$, то $\hat{L}_{\{\pi(g)\}} > \hat{L}_{\{\pi(f)\}}$.

Доказательство. Предположим, что $\zeta^{(f)}(k) = x$, $x \in X(M)$, т.е. управляемый в соответствии со стратегией $\pi(f)$ процесс $\zeta^{(f)}(t)$ в результате проведенных k шагов находится в состоянии $x = (x_1, x_2, \dots, x_m) \in X(M)$. Если после k-го шага не будет принято решения о прекращении процесса управления, а осуществится очередной $(k+1)$-й шаг, то согласно условию с вероятностью $p_{(f)}(x, y)$ процесс

$\zeta^{(f)}(t)$ перейдет в состояние $y=(y_1, y_2, \ldots, y_m) \in X(M)$. При этом выгода от управления возрастет на величину $\sum_{i=1}^{m} d_i \cdot (x_i - y_i)$. Если будет принято решение о прекращении процесса управления, то процесс $\zeta^{(f)}(t)$ перейдет в состояние $M=(M_1, M_2, \ldots, M_m) \in X(M)$ и значение функционала $L_{\{\pi(f)\}}(k)$ не изменится. Вероятность ρ такого решения постоянна и не зависит от шага k, когда оно принимается. Поэтому вектор $d_{\{f\}}$ является вектором средних выгод управляемого случайного процесса $\zeta^{(f)}(t)=\{\zeta_1^{(f)}(t), \zeta_2^{(f)}(t), \ldots, \zeta_m^{(f)}(t)\}$, $t=0,1,2,\ldots$, за один шаг. Если ввести переходные вероятности процесса $\zeta^{(f)}(t)$ за t шагов

$$p_{\{f\}}^{(t)}(x, y) = P\{\zeta^{(f)}(t) = y / \zeta^{(f)}(0) = x\},$$

то, использовав свойство марковости процесса $\zeta^{(f)}(t)$, после несложных преобразований получим равенство

$$m_{\{f\}}(T, x) = E[L_{\{\pi(f)\}}(T) / \zeta^{(f)}(0) = x] = \sum_{t=0}^{T-1} \left(\sum_{y \notin X(M)} p_{\{f\}}^{(t)}(x, y) \cdot d_{\{f\}}(y) \right),$$

что в матричном виде представим следующим образом: $m_{\{f\}}(T) = \sum_{t=0}^{T-1} P_{\{f\}}^{(t)} \cdot d_{\{f\}}$,

где $P_{\{f\}}^{(t)}$ — матрица переходных вероятностей процесса $\zeta^{(f)}(t)$ за t шагов. Поскольку для однородной цепи Маркова справедливы равенства $P_{\{f\}}^{(t)} = [P_{\{f\}}]^t$, $t=0,1,2,\ldots$, то

$$m_{\{f\}}(T) = \sum_{t=0}^{T-1} (P_{\{f\}})^t \cdot d_{\{f\}}.$$

Принимая во внимание результат теоремы 2 и свойства матриц $P_{(f)}$ и $\hat{Q}_{\{f\}}$, убеждаемся в справедливости для произвольного натурального числа k следующих равенств:

$$P_{(f)} \cdot \hat{Q}_{(f)} = \hat{Q}_{(f)} \cdot P_{(f)} = \hat{Q}_{(f)} \cdot \hat{Q}_{(f)} = \hat{Q}_{(f)}; \ (P_{(f)})^k \cdot \hat{Q}_{(f)} = \hat{Q}_{(f)}.$$

Кроме того, по индукции легко проверить справедливость соотношения $(P_{(f)} - \hat{Q}_{(f)})^k = (P_{(f)})^k - \hat{Q}_{(f)}$.

Проводя ряд технических преобразований, приводим выражение для вектора $m_{\{f\}}(T)$ к виду

$$m_{\{f\}}(T) = T \cdot \hat{Q}_{(f)} \cdot d_{(f)} + \left[(E - P_{\{f\}} + \hat{Q}_{\{f\}})^{-1} - \sum_{t=T}^{\infty} (P_{\{f\}} - \hat{Q}_{\{f\}})^t - \hat{Q}_{\{f\}} \right] \cdot d_{\{f\}} =$$

$$= T \cdot \hat{Q}_{(f)} \cdot d_{(f)} + \left[H_{\{f\}} - \sum_{t=T}^{\infty} (P_{\{f\}} - \hat{Q}_{\{f\}})^t \right] \cdot d_{\{f\}}.$$

Поскольку процесс $\zeta^{(f)}(t)$ является эргодическим, то $\lim_{t \to \infty} (P_{\{f\}})^t = \hat{Q}_{(f)}$. Поэтому $(P_{(f)})^t - \hat{Q}_{(f)} = (P_{(f)} - \hat{Q}_{(f)})^t \underset{t \to \infty}{\to} 0$. Тогда полученное равенство принимает вид $m_{\{f\}}(T) = T \cdot \hat{Q}_{(f)} \cdot d_{(f)} + H_{\{f\}} \cdot d_{\{f\}} + \varepsilon_T$, где ε_T обозначает следующее выражение:

$$\varepsilon_T = -\left[\sum_{t=T}^{\infty}(P_{\{f\}} - \hat{Q}_{\{f\}})^t\right] \cdot d_{\{f\}}.$$

Поскольку $\sum_{t=0}^{\infty}(P_{\{f\}} - \hat{Q}_{\{f\}})^t = (E - P_{\{f\}} + \hat{Q}_{\{f\}})^{-1}$, приходим к выводу, что $\lim_{T\to\infty}(\varepsilon_T) = 0.$

Введем единичный вектор-столбец $I_{\{f\}}$, размерность которого совпадает с размерностью вектора $d_{\{f\}}$. Тогда $m_{\{f\}}(T) = T \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + H_{\{f\}} \cdot d_{\{f\}} + \varepsilon_T$.

В то же время $m_{\{f\}}(T) = \sum_{t=0}^{T-1}(P_{\{f\}})^t \cdot d_{\{f\}} = d_{\{f\}} + P_{\{f\}} \cdot m_{\{f\}}(T-1).$

Сопоставляя эти два равенства для $m_{\{f\}}(T)$, получаем

$$m_{\{f\}}(T) = d_{\{f\}} + P_{\{f\}} \cdot [(T-1) \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + H_{\{f\}} \cdot d_{\{f\}} + \varepsilon_{T-1}] =$$
$$= d_{\{f\}} + (T-1) \cdot P_{\{f\}} \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + P_{\{f\}} \cdot \nu_{\{f\}} + P_{\{f\}} \cdot \varepsilon_{T-1} =$$
$$= d_{\{f\}} + (T-1) \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + P_{\{f\}} \cdot \nu_{\{f\}} + \varepsilon'_{T-1},$$

где $\nu_{\{f\}} = H_{\{f\}} \cdot d_{\{f\}}$, а $\varepsilon'_{T-1} = P_{\{f\}} \cdot \varepsilon_{T-1} \underset{T\to\infty}{\to} 0$. Но $Q_{(f)} \cdot d_{(f)} = \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}}$. Поэтому $m_{\{f\}}(T) = T \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + \nu_{\{f\}} + \varepsilon_T = d_{\{f\}} + (T-1) \cdot \hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + P_{\{f\}} \cdot \nu_{\{f\}} + \varepsilon'_{T-1}$. Или $\hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + \nu_{\{f\}} = d_{\{f\}} + P_{\{f\}} \cdot \nu_{\{f\}} + \varepsilon''_T$, где $\varepsilon''_T = \varepsilon'_{T-1} - \varepsilon_T \underset{T\to\infty}{\to} 0.$

Пусть теперь $g = \{g_x, x \in X(M)\}$ — некоторый вектор решений, отличный от $f = \{f_x, x \in X(M)\}$, $\zeta^{(g)}(t)$ — управляемый в соответствии со стационарной стратегией $\pi(g) = \{g, g, g, \dots\}$ процесс, а $\hat{L}_{\{\pi(g)\}}$ — средняя выгода за один шаг. Предположим, что выполняется условие теоремы $d_{(g)} + P_{(g)} \cdot \nu_{(f)} > d_{(f)} + P_{(f)} \cdot \nu_{(f)}$. Используя полученное ранее равенство соответственно для стационарных стратегий $\pi(f) = \{f, f, f, \dots\}$ и $\pi(g) = \{g, g, g, \dots\}$, получаем следующее соотношение:

$$(\hat{L}_{\{\pi(g)\}} - \hat{L}_{\{\pi(f)\}}) \cdot I + (\nu_{\{g\}} - \nu_{\{f\}}) = \delta + P_{\{g\}} \cdot (\nu_{\{g\}} - \nu_{\{f\}}) + (\varepsilon''_{\{g\}T} - \varepsilon''_{\{f\}T}),$$

где $\delta = (d_{\{g\}} - d_{\{f\}}) + (P_{\{g\}} \cdot \nu_{\{f\}} - P_{\{f\}} \cdot \nu_{\{f\}})$.

На основании предположения теоремы 3 можем утверждать, что все координаты этого вектора неотрицательны и по крайней мере одна среди них строго положительна. Умножим обе части этого равенства на вектор-строку $\hat{q}_{(g)} = \{\hat{q}_{(g)}(x), x \in X(M)\}$ стационарного распределения цепи Маркова с матрицей переходных вероятностей $P_{\{g\}}$. После выполнения соответствующих упрощений получим $\hat{L}_{\{\pi(g)\}} - \hat{L}_{\{\pi(f)\}} = q_{(g)} \cdot \delta + q_{(g)} \cdot (\varepsilon''_{\{g\}T} - \varepsilon''_{\{f\}T})$.

Поскольку $q_{(g)} \cdot \delta > 0$, а $\lim_{T\to\infty}(\varepsilon''_{\{g\}T} - \varepsilon''_{\{f\}T}) = 0$, то всегда можно выбрать T так, чтобы $q_{(g)} \cdot \delta + q_{(g)} \cdot (\varepsilon''_{\{g\}T} - \varepsilon''_{\{f\}T}) > 0$. Следовательно, $\hat{L}_{\{\pi(g)\}} - \hat{L}_{\{\pi(f)\}} > 0$, что и требовалось доказать.

Главная трудность использования доказанной теоремы для решения практических задач заключается в проверке ее условия. Поэтому используем для построения соответствующего конструктивного алгоритма полученное при доказательстве теоремы 3 соотношение $\hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + \nu_{\{f\}} = d_{\{f\}} + P_{\{f\}} \cdot \nu_{\{f\}}$, в котором опущено несущественное слагаемое ε''_T. Изучаемая модель построена таким образом, что выбор вектора решений $f = \{f_x, x \in X(M)\}$ однозначно определяет величину $\hat{L}_{\{\pi(f)\}}$ и вектор $d_{\{f\}}$. Предположим, что вектор $V_{(f)} = \{V_{(f)}(x), x \in X(M)\}$ удовлетворяет системе уравнений

$$\hat{L}_{\{\pi(f)\}} \cdot I_{\{f\}} + V_{\{f\}} = d_{\{f\}} + P_{\{f\}} \cdot V_{\{f\}}.$$

Легко убедиться, что тогда аналогичное свойство имеет вектор $V^*_{\{f\}} = V_{\{f\}} + C \cdot I$, где C — произвольное действительное число. Это означает, что рассматриваемая система определяет вектор $V_{(f)}$ с точностью до произвольной действительной постоянной. Поэтому одну из координат вектора $V_{(f)}$ можно выбрать произвольно и положить, например, равной нулю. Положим по определению $V_{(f)}(M) = 0$, $M = \{M_1, M_2, \dots, M_m\}$. Остальные координаты $V_{(f)}(x)$, $x \in X(M)$, $x \neq M$, вектора $V_{(f)}$ определим как решение системы уравнений

$$\begin{cases} \hat{L}_{\{\pi(f)\}} + V_{\{f\}}(x) = d_{\{f\}}(x) + \sum\limits_{y \neq M} p_{\{f\}}(x, y) \cdot V_{\{f\}}(y), \\ x \in X(M), \end{cases}$$

в которой вместо известной координаты $V_{(f)}(M)$ имеется в качестве неизвестной величина $\hat{L}_{\{\pi(f)\}}$. Тогда для произвольного вектора решений $g = \{g_x, x \in X(M)\}$ соотношение $d_{(g)} + P_{(g)} \cdot \nu_{(f)} > d_{(f)} + P_{(f)} \cdot \nu_{(f)}$ эквивалентно неравенству $d_{(g)} + P_{(g)} \cdot V_{(f)} > d_{(f)} + P_{(f)} \cdot V_{(f)}$, где $V_{(f)}$ — определенный ранее вектор. Используя это, можем сформулировать критерий оптимальности для произвольной стационарной стратегии $\pi(f) = \{f, f, f, \dots\}$. Пусть $x = (x_1, x_2, \dots$ $\dots, x_m) \in X(M)$ — некоторое состояние процесса $\zeta^{(f)}(t)$, а $u \in U$ — некоторое решение из множества управлений $U = \{1, 2, \dots, r\}$. Построим на основании вектора решений $f = \{f_x, x \in X(M)\}$ новый вектор решений $f_u^x = \{f_u^x(y), y = (y_1, y_2, \dots, y_m) \in X(M)\}$ следующим образом:

$$f_u^x(y) = \begin{cases} u, & \text{если } y = x, \\ f_y, & \text{если } y \neq x. \end{cases}$$

Определим для каждого $x = (x_1, x_2, \dots, x_m) \in X(M)$ множество $G(x, f) \subseteq$ $\subseteq U = \{1, 2, \dots, r\}$. При этом $u \in G(x, f)$ тогда и только тогда, когда

$$d_{\{f_u^x\}}(x) + \sum_{y \neq M} p_{\{f_u^x\}}(x, y) \cdot V_{\{f\}}(y) > \hat{L}_{\{\pi(f)\}} + V_{\{f\}}(x).$$

Тогда будет справедливо следующее утверждение.

Теорема 4. Пусть $\pi(f) = \{f, f, f, \dots\}$ — стационарная стратегия, построенная на основании вектора решений $f = \{f_x, x \in X(M)\}$. Для того чтобы стационарная стратегия $\pi(f) = \{f, f, f, \dots\}$ была оптимальной, необходимо и достаточно, чтобы для произвольного состояния $x = (x_1, x_2, \dots, x_m) \in X(M)$ множество $G(x, f)$ было пусто.

Если стационарная стратегия $\pi(f) = \{f, f, f, \dots\}$ не является оптимальной, то построим новый вектор решений $g = \{g_x, x \in X(M)\}$ следующим образом:

- в случае, если $G(x, f) = \varnothing$, положим $g_x = f_x$;

• если $G(x, f) \neq \varnothing$, то $g_x = u$, где $u \in G(x, f)$ — произвольное решение из множества $G(x, f)$.

Тогда $\hat{L}_{\{\pi(g)\}} > \hat{L}_{\{\pi(f)\}}$ и стационарная стратегия $\pi(g) = \{g, g, g, \ldots\}$, построенная на основании вектора решений $g = \{g_x, x \in X(M)\}$, будет лучше, чем стационарная стратегия $\pi(f) = \{f, f, f, \ldots\}$.

Теорема 4 позволяет предложить итерационный численный алгоритм, на каждом шаге которого значение средней стационарной выгоды строго монотонно возрастает, что исключает возможность возникновения циклов при его использовании, а также с учетом количества стационарных стратегий гарантирует возможность построения оптимальной стационарной стратегии в результате проведения всех шагов. Это применимо, в частности, к моделям, рассмотренным в [1, 9, 10].

Предложенные в данной статье конструктивные подходы могут являться важными составляющими в реализациях программных комплексов поддержки принятия наиболее приемлемых решений во многих отраслях: в экономике и финансах, в сфере управления сложными социальными явлениями, при построении эффективных верификационных процедур для различных технологических процессов и в других направления деятельности. Реализации данных подходов на практике способствует несколько объективных факторов. Прежде всего рассмотренные в работе модели базируются на стохастическом, т.е. наиболее приближенном к реальности, описании исследуемых процессов. В настоящее время, что также важно, существенно снижается стоимость получения и обработки информации, необходимой для их построения и применения [11]. Происходит это, в первую очередь, благодаря методам сбора статистической информации с помощью современных интернет-технологий [12], а также возможности использовать уже накопленные массивы статистических данных из открытых и предоставляемых на коммерческой основе источников. Постоянный рост вычислительных мощностей и практика применения систем машинного обучения (искусственного интеллекта) позволяют стандартизировать оценки качества и выбирать оптимальные решения на сложных полирегиональных рынках. Применение на практике предложенных методов дает возможность сосредоточиться на наиболее значимых направлениях деятельности, получить необходимые конкурентные преимущества и связанные с этим выгоды.

СПИСОК ЛИТЕРАТУРЫ

1. Война Ал.А., Война Ан.А. Динамическое управление риском в многомерных марковских моделях. *Кибернетика и системный анализ*. 2018. Т. 54, № 2. С. 45–54.
2. Гихман И.И., Скороход А.В. Управляемые случайные процессы. Киев: Наук. думка, 1977. 252 с.
3. Майн Х., Осаки С. Марковские процессы принятия решений. Москва: Наука, 1977. 176 с.
4. Wojna A. Ryzyko w procesach finansowych oraz metody badan koniunktury. Politechnika Koszalinska (Polska), 2009. 446 s.
5. Война А.О. Стохастичні математичні моделі та функція ризику при побудові оптимальних маркетингових стратегій. *Журнал обчислювальної та прикладної математики*. 2004. Вип. 2, № 89. С. 2–5.
6. Война О.А. Оптимальне регулювання регенеративного типу в марковських моделях. *Дослідження операцій та АСУ*. 1992. Вып. 39. С. 3–10.
7. Война А.А. Оптимальные динамические стратегии для стохастических моделей маркетингового комплекса продвижения. *Кибернетика и системный анализ*. 2004. Т. 40, № 2. С. 145–152.

8. Гихман И.И., Скороход А.В. Теория случайных процессов. Т. 1. Москва: Наука, 1971. 664 с. Т. 2, 1973. 640 с.
9. Война А.А. Одна модель оптимального регулирования запасами. *Докл. АН УССР. Сер. А*. 1983. № 1. С. 11–14.
10. Война А.О. Про вибір оптимальних стратегій в стохастичних маркетингових моделях. *Доп. НАН України*. 2002. № 8. С. 60–64.
11. FLOPS. Performance records. Cost of computing. URL: https://en.wikipedia.org/wiki/FLOPS.
12. Internet Usage Statistics. World Internet Users and 2018 Population Stats. URL: https://www.internetworldstats.com/stats.htm.

Надійшла до редакції 10.12.2018

О.А. Война, А.О. Война

МАТЕМАТИЧНІ МОДЕЛІ КЕРУВАННЯ РИЗИКОМ ДЛЯ МАРКОВСЬКИХ ПРОЦЕСІВ РЕҐЕНЕРАЦІЙНОГО ТИПУ

Анотація. В рамках математичної моделі, умовно названої паралельною марковською структурою, формалізовано низку практичних постановок задач оптимального керування. Вивчено властивості моделей і використаних до їхньої побудови випадкових процесів. Розроблено алгоритми обчислення вартості відповідних функцій ризику та побудови оптимальних стратегій керуванння.

Ключові слова: реґенераційний процесс, паралельна структура, оптимальні стаціонарні стратегії, марковські процеси прийняття рішень, функція ризику, ітераційні алгоритми.

O.A. Voina, A.O. Voyna

DYNAMIC RISK CONTROL IN MULTIDIMENSIONAL MARKOV MODELS

Abstract. Within the framework of a mathematical model called conventionally “parallel Markov structure,” a number of practical set ups of optimal control problems are formalized. The properties of the models and of the random processes used to build them are examined. In addition, the algorithms on calculating the cost of the corresponding risk functions and on constructing optimal control strategies are developed.

Keywords: regenerating process, parallel structure, optimal stationary strategies, Markov decision processes, risk function, iterative algorithms.

Война Александр Андреевич,
доктор физ.-мат. наук, профессор кафедры Киевского национального университета имени Тараса Шевченко, e-mail: avoina@hotmail.com.

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