ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ КОНЕЧНОЙ МАРКОВСКОЙ ЦЕПЬЮ С НЕЧЁТКО ЗАДАННОЙ ВЕРОЯТНОСТНОЙ МЕРОЙ

Карманов Анатолий Вячеславович
Д.ф.-м.н., профессор, РГУ нефти и газа (НИУ) имени И. М. Губкина
abkar2007@yandex.ru
Попадько Владимир Ефимович
К.т.н., профессор, РГУ нефти и газа (НИУ) имени И. М. Губкина
pve@gubkin.ru

Аннотация. В статье ставится задача управления однородной Марковской цепью, часто встречающаяся в приложениях, в условиях, когда вероятностная мера на пространстве траекторий этой цепи задана не чётко, т.е. мера является неизвестным элементом некоторого определённым образом заданного множества вероятностных мер. При этом формулируется задача оптимального управления такой Марковской цепью с функционалом — гарантированный доход, и критерием оптимальности — максимум гарантированного дохода. Приводится итерационная процедура решения поставленной задачи.

Ключевые слова: однородная конечная Марковская цепь с доходом, управляемая конечная Марковская цепь, предел по Чезаро стохастической матрицы, оптимальное управление Марковской цепью с доходом.

OPTIMAL CONTROL WITH A FINITE MARKOV CHAIN AND A FUZZY GIVEN PROBABILITY MEASURE

A. Karmanov
V. Popadko

Summary. The article poses the problem of controlling a homogeneous Markov chain, which is often found in applications, under conditions when the probability measure on the trajectory space of this chain is not clearly defined, i.e. the measure is an unknown element of some certain set of probability measures. In this case, the problem of optimal control with such Markov chain with a guaranteed income functional and maximum guaranteed income optimality criterion is formulated. An iterative procedure for solving the problem is given.

Keywords: homogeneous finite Markov chain with income controlled finite Markov chain, Cesiaro limit of the stochastic matrix, the Markov chain optimal control with income.

1. Постановка задачи

Рассмотрим управляемую конечную Марковскую цепь (УКМЦ), которая задаётся следующей совокупностью данных:

$$(J, p_0, \{U_i,\ iJ\}, \{h_i(s),\ s\ U_i\}), \tag{1}$$

где $J = \{1,., n\}$ — конечное множество состояний, p_0 — n-мерный стохастический вектор-строка, являющийся начальным распределением для УКМЦ, $U_i = \{1,\dots, m(i)\}$ — конечное множество управлений в состоянии i, $i = 1,., n$, $h_i(s) = [h_{i,1}(s),\dots, h_{i,n}(s)$,

$h_{i,n+1}(s)]$, $[h_{i,1}(s),\dots, h_{i,n}(s)]$- n-мерный стохастический вектор-строка, являющийся при фиксированным i, вектором переходных вероятностей из состояния i в состояние j при управлении $s\ U_i$, $j = 1,\dots, n$, $h_{i,n+1}(s)$- средний доход в состоянии i при управлении $s\ U_i$. При этом средний доход определяется соотношением:

$$h_{i,n+1}(s) = h_{i,1}(s)\cdot c_{i,1}(s) + \ldots + h_{i,n}(s)\cdot c_{i,n}(s), \tag{2}$$

где $c_{i,j}(s)$ — доход, получаемый от перехода УКМЦ из состояния i в состояние j при управлении s, $j = 1,\dots, n$. В дальнейшем этот доход будет считаться ограниченной функцией, т.е. $c_{i,j}(s)\ I_1$ для любых iJ, J, $s\ U_i$, I_1 — ограниченный числовой интервал вида:

$$[-\beta_0, \beta_0],\ \text{где}\ \beta_0 > 0. \tag{3}$$

Для формулирования стандартной задачи оптимального управления УКМЦ введём следующие обозначения:

1. $U = U_1\, x \dots x\, U_n$ — множество управлений, являющееся прямым

произведением множеств U_i, $i = 1,\ldots, n$. Произвольный элемент u U будем представлять (в случае необходимости) в виде: $u = (u_1,\ldots, u_n)$, где u_i U_i, $i = 1,\ldots, n$.

2. $h(u) = [h_1(u_1),\ldots, h_n(u_n)]^T$ — матрица размерности $n \times (n+1)$, где

i-ой строкой является $h_i(u_i) = [h_{i,1}(u_i),\ldots, h_{i,n}(u_i), h_{i,n+1}(u_i)]$, u_i U_i. Каждому элементу $h(u)$, где u U, соответствует 1) стохастическая матрица $P(h(u))$ порядка n, i-ая строка которой — $[h_{i,1}(u_i),\ldots, h_{i,n}(u_i)]$, $i = 1,\ldots, n$, 2) вектор дохода $q(h(u)) = [h_{1,n+1}(u_i),\ldots, h_{n,n+1}(u_i)]^T$.

Стандартная задача оптимального управления УКМЦ состоит в нахождении такого управления u^* U, которое удовлетворяет соотношению:

$$G(u^*) = max\{p_0{\cdot}\Pi(u){\cdot}q(u) \mid u\ U\}, \qquad (4)$$

где $\Pi(u)$- предел по Чезаро матрицы $P(h(u))$, т.е. $\Pi(u) = lim\ n^{-1}{\cdot}[E + P(h(u)) +\ldots+ P^{n-1}(h(u))]$, $n \to \infty$, $q(u) = q(h(u))$.

Решение задачи (4) существует и определяется по процедуре Р. Ховарда, изложенной, например, в [1]. Однако в приложениях часто возникает ситуация, когда вероятности $h_{i,j}(u_i)$, где $i = 1,\ldots, n$, $j = 1,\ldots, n$, известны с точностью до некоторого интервала, т.е. при каждом фиксированном i выполняются соотношения:

$$h_{i,j}(u_i)\ I_j(u_i),\ h_{i,1}(u_i) +\ldots+ h_{i\,n}(u_i) = 1, \qquad (5)$$

$I_j(u_i)$ — числовой интервал $[h^{(1)}_{i,j}(u_i), h^{(2)}_{i,j}(u_i)]$, включенный в интервал [0, 1].

Из соотношений (5), (2) и (3) следует, что каждому управлению u_i U_i, где $i = 1,\ldots, n$, соответствует замкнутый линейный многогранник (симплекс) $D(u_i)$ следующего вида:

$$D(u_i) = \{h_i = [h_{i,1},\ldots, h_{i,n+1}] \mid$$
$$h_{i,1}+\ldots+ h_{i,n} = 1,\ h_{i,j} \geq 0,\ h_{i,j}I_j(u_i), j = 1,\ldots, n\}, \qquad (6)$$

где

$$h_{i,n+1} = h_{i,1}{\cdot}c_{i,1} +\ldots+ h_{i,n}\cdot c_{i,n},\ c_{i,j}\ [-\beta_0,\beta_0],\ \beta_0 > 0. \qquad (7)$$

Замечание *1*. Из соотношения (5) следует, что вероятностная мера на пространстве траекторий УКМЦ [2, *с*.11] определяется нечётко, т.е. при выбранном управлении u U она является элементом некоторого множества $S(u)$ вероятностных мер, каждая из которых может в равной степени определять решение стандартной задачи.

Если ввести обозначение $D(u) = D(u_1)\ x\ldots\ x\ D(u_n)$, то в условиях замечания 1 стандартную задачу можно трансформировать в следующую задачу 1: найти такое управление u^*U, для которого выполняется соотношение:

$$g(u^*) = max\{g(u) \mid uU\}, \qquad (8)$$

где

$$g(u) = inf\{p_0{\cdot}\Pi(h){\cdot}q(h) \mid h\ D(u)\}, \qquad (9)$$

$\Pi(h)$ — предел по Чезаро стохастической матрицы P(h), в которой i-ая строка равна $[h_{i,1},\ldots, h_{i,n}]$, $i = 1,\ldots, n$, а вектор дохода $q(h) = [h_{1,n+1},\ldots, h_{n,n+1}]^T$, $h_{i,n+1}$ соотношением (7).

Замечание 2. Из (9) и (8) следует, что $g(u)$ является гарантированным доходом при управлении u U, а $g(u^*)$ является максимальным гарантированным доходом на УКМЦ. При этом, именно управление u^*U в задаче 1 объявляется оптимальным управлением УКМЦ с нечётко заданной вероятностной мерой на пространстве траекторий этой цепи.

В [2] доказано, что задача 1 эквивалентна задачи 2, т.е. управление u^*U, для которого выполняются соотношения (8), (9), удовлетворяет следующим соотношениям:

$$g_1(u^*) = max\{r(t(u)) \mid uU\},\ r(t(u)) = inf\{r(h) \mid h\ D(u)\}, \qquad (10)$$

где

$$r(h) = \Pi(h){\cdot}q(h). \qquad (11)$$

Отметим, что из (10) и (11) следует выполнение следующих векторных соотношений:

$$r(t(u)) \leq r(h) \text{ для любого } h\ D(u), \qquad (12)$$

$$g_1(u^*) \geq r(t(u)) \text{ для любого } uU, \qquad (13)$$

где неравенства для рассматриваемых векторов понимаются как покомпонентные неравенства. Таким образом, $g_1(u^*)$ является максимальным вектором гарантированного дохода. При этом управление u^* не зависит от начального распределения p_0.

Ниже излагается процедура, которая решает за конечное число итераций задачу 2 для случая, когда характеристики $D(u_i)$, где u_i U_i, $i = 1,\ldots, n$, являются замкнутыми линейными многогранниками — симплексами, задаваемыми выражением (6).

2. Решение вспомогательных задач

Итерационная процедура решения задачи 2 основывается на алгоритмах решения ряда вспомогательных задач.

2.1.Вспомогательная задача 2.1

Введём следующее обозначение для множества всех стохастических векторов размерности n:

$P_{1,\,n} = \{z = (z_1,\ldots, z_n) \mid z_i \geq 0, z_1+\ldots+ z_n = 1\}$, (14)

и сформулируем первую вспомогательную задачу 2.1.

Задача *2.1.* Пусть имеется симплекс D:

$$D = \{z = [z_1, \ldots, z_n] \in P_{1,n} \mid 0 \leq p_i^{(1)} \leq z_i \leq p_i^{(2)},$$

$$p_i^{(1)} < p_i^{(2)} i = 1, \ldots, n\}, \tag{15}$$

и пусть имеются две линейные формы, имеющие вид

$$F_k(z) = a_0^{(k)} + a_1^{(k)} z_1 + \ \ldots \ + a_n^{(k)} z_n, k = 1, 2. \tag{16}$$

Тогда необходимо найти один из следующих двух элементов, принадлежащих симплексу D:

1. Элемент $z^{(1)}$, который удовлетворяет условию:

$F_1(z^{(1)}) > 0$, (17)

2. В том случае, если $F_1(z^{(1)}) = 0$, найти элемент $z^{(2)}$, который удовлетворяет условию:

$$F_2(z^{(2)}) = min \{F_2(z) \mid z, F_1(z) = 0\} > 0. \tag{18}$$

В случае, когда элементы $z^{(1)}$ и $z^{(2)}$, удовлетворяющие условиям (17), (18), не существуют, то будем считать, что решение задачи 2.1 отсутствует.

Задача 2.1 решается в соответствии с общим алгоритмом симплекс-метода [3].

2.2. Вспомогательная задача 2.2

Задача 2.*2*. Пусть имеются симплекс D, определяемый выражением (15), и три линейные формы, имеющие вид:

$$F_k(z) = a_0^{(k)} + a_1^{(k)} z_1 + \ \ldots \ + a_n^{(k)} z_n, k = 1, 2, 3, \tag{19}$$

где z. При этом линейные формы обладают следующим свойством: существует такой элемент, для которого выполняются равенства

$$F_k(t) = 0, k = 1, 2, 3. \tag{20}$$

Тогда необходимо найти такие элементы $z^{(1)}$, $z^{(2)}$ и $z^{(3)}$ из симплекса D, которые удовлетворяют одному из следующих трех условий:

1. $F_1(z^{(1)}) = min \{F_1(p_1)\} < 0,$ (21)

2. Если $F_1(z^{(1)}) = 0$, то $F_2(z^{(2)}) =$
$= min \{F_2(z) F_1(z) = 0\} < 0,$ (22)

3. Если $F_1(z^{(1)}) = F_2(z^{(2)}) = 0$, то

$$F_3(z^{(3)}) = min \{F_3(z), F_1(z) = 0, F_2(z) = 0\} \leq 0. \tag{23}$$

Отметим, что из соотношения (20) следует, что задача 2.2 всегда имеет решение, которое может быть найдено по общему алгоритму симплекс-метода [3].

2.3. Вспомогательная задача 2.3

Ниже ставится и решается вспомогательная задача 2.3 по определению основных стационарных характеристик однородной конечной Марковской цепи с доходом $\xi(p_0, h)$, задаваемой начальным распределением p_0, матрицей переходных вероятностей $P(h)$ и вектором дохода $q(h)$, где $h = (h_1,\ldots, h_n)\ H^n\ h_i\ H$, i=1,…, n, $H = x\ [-\beta_0 \beta_0]$. Отметим, что $D(u)$ является подмножеством множества H^n для любого uU.

Задача *2.3.* Пусть имеется однородная Марковская цепь с доходом $\xi(p_0, h)$, которая имеет $n(h)$ эргодических классов состояний, где $n(h)\{1,\ldots, n\}$. Необходимо определить стационарные характеристики этой цепи где = $\Pi(h)\cdot q(h)$, $\Pi(h)$- предел по Чезаро матрицы $P(h)$,— вектор первого «веса»,— фундаментальная матрица Марковской цепи $\xi(p_0, h)$,— вектор второго «веса».

Ниже приводится алгоритм 2.3 решения задачи 2.3, основанный на решении рекуррентных систем линейных уравнений, приведенных в разделе 2.3.2 [2].

Алгоритм 2.3 определения стационарных характеристик представляет собой следующую двухшаговую вычислительную процедуру.

Шаг 1. Для каждого m-го эргодического класса состояний однородной Марковской цепи $\xi(p_0, h)$, где определяются (рекуррентно по k = 1, 2, 3) величины: c_m^kи $w_{k,i}^0$ как единственное решение системы линейных уравнений

$$c_m^k + w_{k,i}^0 = b_i^{k-1} + \sum_{j=1}^n p_{i,j} \cdot w_{k,j}^0, i \in J_m(h) \tag{24}$$

где $J_m(h)$ — множество состояний, образующее m-й эргодический класс, $b_i^0 = q_i(h)$— i-я компонента вектора $q(h)$, $b_i^1 = -w_{1,i}^0$, $b_i^2 = -w_{2,i}^0$, $p_{i,j}$ — соответствующий элемент матрицы $P(h)$ и одно любое равно нулю.

Далее, положив $r_i = c_m^1$, $w_{1,i} = w_{1,i}^0 + c_m^2$, $w_{2,i} = w_{2,i}^0 + c_m^3$, если $i \in J_m(h)$, перейдем к реализации второго шага алгоритма.

Шаг 2. Для класса невозвратных состояний Марковской цепи $\xi(p_0, h)$ определяются величины:

2.1. r_i и $w_{1,i}$, где $i \in J_0(h)$ как единственное решение системы линейных уравнений:

$$r_i = \sum_{j=1}^{n} p_{i,j} \cdot r_j, r_i + w_{1,i} =$$

$$= q_i + \sum_{j=1}^{n} p_{i,j} \cdot w_{1,j}, i \in J_0(h), \quad (25)$$

где $w_{1,j}$, r_j — величины, определенные на шаге 1,

$$j \in [J_1(h) + \ldots + J_{n(h)}(h)]$$

2.2. $w_{2,i}$, $i \in J_0(h)$ как единственное решение системы линейных уравнений:

$$w_{2,i} = -w_{1,i} + \sum_{j=1}^{n} p_{i,j} \cdot w_{2,j}, i \in J_0(h), \quad (26)$$

где величина $w_{2,j}$ при $j \in [J_1(h) + \ldots + J_{n(h)}(h)]$ определяется на шаге 1 алгоритма 2.3. Действие алгоритма завершается.

Таким образом, по алгоритму 2.3 определяются три вектора

$$r(h) = [r_1, \ldots, r_n]^T, w_1(h) = [w_{1,1}, \ldots, w_{1,n}]^T,$$

$$w_2(h) = [w_{2,1}, \ldots, w_{2,n}]^T.$$

1.4. Вспомогательная задача 2.4

Ниже ставится задача 2.4, решение которой представляет собой основу итерационной процедуры решения исходной задачи 2 оптимального управления, определяемой выражением (10).

Задача *2.4.* Пусть имеется управление

$$u = [u_1, \ldots, u_n] \in U.$$

Тогда в множестве $D(u) = D_1(u_1) \times \ldots \times D_n(u_n)$, $D_1(u_1)$, где — симплекс, определяемый выражением (6), $i = 1, \ldots, n$ необходимо найти такой элемент $t(u)$, для которого справедливо выражение

$$r(t(u)) \leq r(h), \forall h \in D(u) \quad (27)$$

В разделе 4.3 [2] доказывается, что $t(u)$ существует в каждом множестве, где, и является 1-минимальным элементом в множестве $D(u)$, т.е. для $t(u)$ выполняется соотношение:

$$r(t(u)) = \min \{r(h): h \in D(u)\} \quad (28)$$

Следующий алгоритм 2.4, основанный на решении задач 2.2 и 2.3, позволяет решить задачу 2.4. Алгоритм 2.4 представляет собой итерационную пошаговую процедуру, в которой на каждом m-ом шаге (m = 0, 1, 2…, m_0) производятся следующие вычислительные действия.

1. Для имеющегося элемента $h^{(m)} \in D(u)$, где $h^{(0)}$ — любой элемент из множества $D(u)$, по алгоритму 2.3 определяются векторы $r(h^{(m)})$, $w_1(h^{(m)})$, $w_2(h^{(m)})$

2. Для каждого i, где $i = 1, \ldots, n$, по алгоритму 2.2 решается задача 2.2, где $D = D_i(u_i)$ — симплекс, определяемый выражением (6), $t = h^{(m)}$ — элемент, для которого выполняются равенства (20), $z = p_i(h_i)$,

$$F_1(p_i(h_i)) = [p_i(h_i) - e_i] r(h^{(m)}),$$

$$F_2(p_i(h_i)) = q_i(h_i) + [p_i(h_i) - e_i] w_1(h^{(m)}) - r_i(h^{(m)}),$$

$$F_3(p_i(h_i)) = [p_i(h_i) - e_i] w_2(h^{(m)}) - w_{1,i}(h^{(m)}), \quad (29)$$

e_i — n-мерный вектор-строка, в котором i-я компонента равна единице, а остальные — нулю.

Возможные варианты решения задачи 2.2 имеют вид:
2.1. Если $F_1(z^{(2)}) < 0$, то положить: $h_i^{(m+1)} = z^{(1)}$.
2.2. Если $F_1(z^{(2)}) = 0$, $F_2(z^{(2)}) < 0$, то положить: $h_i^{(m+1)} = z^{(2)}$.
2.3. Если $F_1(z^{(3)}) = F_2(z^{(3)}) = 0$, $F_3(z^{(3)}) < 0$ то положить: $h_i^{(m+1)} = z^{(3)}$.
2.4. Если $F_1(z^{(1)}) = F_2(z^{(2)}) = F_3(z^{(3)}) = 0$, то положить: $h_i^{(m+1)} = z^{(m)}$.

В результате выполнения действий пункта 2 возможны следующие исходы:
- $h_i^{(m+1)} = h^{(m)}$, тогда полагаем и действие алгоритма 2.4 прекращается;
- $h_i^{(m+1)} \neq h^{(m)}$, тогда делается переход к следующему (m+1)-му шагу алгоритма 2.4.

С помощью алгоритма 2.4 в множестве

$$D(u) = D_1(u_1) \times \ldots \times D_n(u_n),$$

где $u \in U$, за конечное число m_0 итераций определяется минимальный элемент $t(u)$, т.е. элемент, для которого справедливо выражения (27), (28).

3. Решение задач 1 и 2

Ниже приводится итерационная процедура, позволяющая отыскать за конечное число итераций оптимальное управление u^*, для которого выполняется соотношение (10) и, следовательно (7), (8), т.е. эта процедура решает как задачу 2, так и задачу 1.

Итерационная процедура представляет собой пошаговую процедуру, в которой на каждом k-ом шаге, где $k = 0, 1,\ldots, \omega$, производятся следующие действия.

1. Для управления выбирается произвольно из множества U, в множестве D(с помощью алгоритма 2.4 вычисляется элемент и по алгоритму 2.3. вектора $r(t^{(k)})$, $w_1(t^{(k)}), w_2(t^{(k)})$. Далее осуществляется переход к пункту 2 этой процедуры;

2. Для каждого i, где $i = 1, \ldots, n$, и для каждого u_i применяется алгоритм решения задачи 2.1, где — симплекс, определяемый выражением (6), $z = p_i(h_i)$, n-мерный вектор-строка, в котором i-я компонента равна единице, а остальные — нулю. При этом возможны следующие исходы:

2.1. Если отыскивается такое управление u_i для которого решение задачи 2.1 существует, то:

2.1.1. Если $F_1(z^{(1)}) = \{F_1(z) \mid z \in D\} > 0$, то $u_i^{(k+1)} = z^{(1)}$.

2.1.2. Если $F_1(z^{(2)}) = 0, F_2(z^{(2)}) = min\{F_2(z) \mid z, F_1(z) = 0\} > 0$, $u_i^{(k+1)} = z^{(2)}$.

2.2. Если для управления $u_i \in U_i$ решение задачи 2.1 не существует, то полагаем $u_i^{(k+1)} = u_i^{(k)}$.

В результате вычислений по пункту 2 возможны следующие исходы:

А. Если $u_i^{(k+1)} = u_i^{(k)}$ для каждого i, где $i = 1, \ldots, n$, то полагаем $u^* = (u_1^{(k)},\ldots, u_n^{(k)}) \in U$, и действие процедуры прекращается.

Б. Если $u_i^{(k+1)} \neq u_i^{(k)}$ хотя бы для одного i, где $i = 1, \ldots, n$, то делается переход к следующему (k+1)–му шагу п. 1 настоящей процедуры.

В результате осуществления изложенной итерационной процедуры строится конечная последовательность управлений {, k=1,…, ω} из множества U, для которой выполняются соотношения

$$r(t^{(k-1)}) < r(t^{(k)})\ k = 1,\ldots, \omega, \tag{30}$$

где $t^{(v)} = t(u^{(v)})$ — минимальный элемент в множестве $D(u^{(v)})$, $v = 0,\ldots, \omega$.

При этом из теоремы 1 (п. 2.2) раздела 8.2 и процедуры 1 раздела 5.2 [2] непосредственно следует, что полученное управление $u^* = u^{(\omega)}$ является решением задачи 2 и, следовательно, задачи 1.

ЛИТЕРАТУРА

1. Ховард Р. А. Динамическое программирование и марковские процессы — М.: Советское радио, 1964. 189 с.
2. Карманов А. В. Исследование управляемых конечных марковских цепей с неполной информацией. — М.: Физматлит, 2002. 173 с.
3. Карманов В. Г. Математическое программирование. — М.: Физматлит, 2000. 256 с.

Журнал «Современная наука: актуальные проблемы теории и практики»