КОМПЬЮТЕРНАЯ
МАТЕМАТИКА

Рассмотрена задача управления полумарковской многономенклатурной системой запасов. Найдены условия существования оптимальной стратегии, а также определена структура оптимальной стратегии при выполненных условиях оптимальности.

УДК 519.21

Т.В. ПЕПЕЛЯЕВА, И.Ю. ДЕМЧЕНКО

ОБ ОДНОЙ МНОГОНОМЕНКЛАТУРНОЙ МОДЕЛИ ДЛЯ ПОЛУМАРКОВСКОЙ СИСТЕМЫ ЗАПАСОВ

Введение. Управляемые полумарковские процессы используются при решении многих задач оптимизации систем массового обслуживания, управления запасами, исследовании надежности сложных технических систем и др. В настоящей статье полумарковские процессы исследуются при рассмотрении задачи оптимизации многономенклатурной модели теории управления запасами. Запасы каждого продукта могут пополняться непрерывно, поэтому, уровни запасов и дозаказов каждого продукта принимают значения в R_+, поэтому множества состояний и действий непрерывны. Причем уровень запасов каждого продукта ограничен сверху.

Цель работы – нахождение условий оптимальности (s,S)-стратегии для многономенклатурной модели управления запасами с функцией стоимости, которая связана с хранением запаса, дозаказами и пополнением товаров, а также дефицитом. В работе [1] приведен общий обзор теории запасов. Теория динамического полумарковского программирования берет свое начало в [2 – 7], и продолжает развиваться в [8 – 17]. Проблема нахождения условий оптимальности (s,S)-стратегии для полумарковской однономенклатурной системы теории управления запасами рассматривалась в [18] и ряде других.

1. Предварительные сведения об управляемых полумарковских процессах

Приведем некоторые сведения из теории управления полумарковскими процессами, которые будут использоваться в данной работе. Рассматривается система со случайными воздействиями в случайные моменты времени, которая управляется некоторым образом для минимизации издержек, связанных с системой управления. Пространство состояний (фазовое пространство) стохастического процесса $X = (X_n : n \in \mathrm{N})$, описывающее развитие системы во времени, обозначим *X*, а пространство управляющих воздействий (решений или действий) обозначим *A*. Пусть *X*, *A* – некоторые полные сепарабельные метрические пространства с борелевскими σ-алгебрами $\aleph$ и $\Im$, соответственно. Пусть задано измеримое отображение, ставящее в соответствие каждому $x \in X$ некоторое непустое замкнутое множество $A_x \subseteq A$, т. е. отображение *F* связывает данное состояние системы с допустимым набором действий (решений). При этом множество $\Delta = \{(x,a) : x \in X, a \in A_x\}$ измеримо по Борелю в произведении пространств $X \times A$ [19].

Если в состоянии $x \in X$ принято решение $a \in A_x$, то:

1) следующее состояние системы выбирается с помощью переходной вероятности $P(\cdot / x,a)$;

2) при условии, что следующее состояние системы есть $y \in X$, время пребывания в x – случайная величина с функцией распределения $\Phi(\cdot / x,a,y)$.

Будем предполагать, что $\mathrm{P}(\cdot / x,a)$ и $\Phi(\cdot / x,a,y)$ – борелевские функции на Δ и $\Delta \times X$, соответственно.

Обозначим x_n состояние системы после *n*-го перехода, a_n – принятое решение, а τ_n – время пребывания в этом состоянии (n = 0, 1, 2, …).

Допустимая стратегия δ для управляемой систем определяется как последовательность $\delta = \{\delta_0, \delta_1, \dots, \delta_n, \dots\}$ ядер перехода такая, что вероятностная мера $\delta_n(\cdot / h_n)$ на $(A, \Im)$, сосредоточенная на A_{X_n} и измеримым образом зависящая от $h_n = (x_0, a_0, \tau_0, \dots, x_{n-1}, a_{n-1}, \tau_{n-1}, x_n)$ – истории управляемой системы к моменту *n*-го перехода. Стратегия δ называется марковской, если $\delta_n(\cdot / h_n) = \delta_n(\cdot / x_n)$, ($n$ = 0, 1, 2, …). Марковская стратегия δ называется стационарной, если $\delta_n(\cdot / x_n) = \delta(\cdot / x_n)$, ($n$ = 0, 1, 2, …) и стационарной нерандомизированной (детерминированной), если мера $\delta(\cdot / x_n)$ вырождена и сосредоточена в точке для любого $x \in X$. В этом случае $\delta(x)$ будем обозначать точку сосредоточения массы $\delta(\cdot / x)$. Обозначим $\Re$ – класс всех допустимых стратегий, $\Re_1$ – класс стационарных марковских нерандомизированных (детерминированных) стратегий. Непустоту класса $\Re_1$ гарантирует условие измеримости отображения *F* [3].

Выбором стратегии δ мы определяем случайный процесс, который будем называть процессом, управляемым стратегией δ. Если стратегия δ – марковская стационарная, то управляемый процесс является полумарковским процессом.

Введем понятие издержек, связанных с управляемым процессом. Если в состоянии $x \in X$ принято решение $a \in A_x$ и время, проведенное в состоянии *x*, равно *t*, то ожидаемые издержки за время s ($s \le t$) равны $r(s / x, a)$. Функция $r(s / x, a)$ предполагается измеримой по Борелю на $[0; +\infty) \times \Delta$.

Рассмотрим следующий критерий оптимальности выбранной стратегии.

Средняя ожидаемая стоимость стратегии δ

$$\varphi(x,\delta)=\lim_{n\to\infty}\sup\frac{E_x^\delta\sum_{k=0}^{n} r(\tau_k / x_k, a_k)}{E_x^\delta\sum_{k=0}^{n}\tau_k},$$

где $\xi_0 = x,$ а E_x^δ – математическое ожидание, соответствующее процессу, управляемому стратегией δ при условии, что $\xi_0 = x$.

Стратегия δ^* оптимальна относительно этого критерия, если

$$\varphi(x,\delta^*)=\inf_{\delta\in R}\varphi(x,\delta),\ \ x\in X.$$

Обозначим

$$\tau(x,a)=\int_X\int_0^\infty t\,d\Phi(t/x,a,y)\,P(dy/x,a),$$

$$r(x,a)=\int_X\int_0^\infty r(t/x,a)\,d\Phi(t,a,y)\,P(dy/x,a).$$

Будем предполагать, что $\tau(x, a)$ и $r(x, a)$ существуют и конечны для всех $(x, a) \in \Delta$ и $|r(x, a)| \le C < \infty$, $(x, a) \in \Delta$. Поскольку первый критерий зависит только от $P(\,\cdot / x, a)$ и усредненных характеристик $\tau(x, a)$ и $r(x, a)$, мы ограничимся рассмотрением управляемых процессов, для которых

$$\Phi(t/x,a,y)=\begin{cases}1, & t\ge\tau(x,a),\\ 0, & t<\tau(x,a),\end{cases}$$

$$r(t/x,a)=\begin{cases}0, & t<\tau(x,a),\\ r(x,a), & t\ge\tau(x,a).\end{cases}$$

Обозначим $\Xi(X)$ – банахово пространство ограниченных измеримых по Борелю функций на X с нормой $\|u\|=\sup_{x\in X}u(x)$. Далее будут использованы такие результаты [4].

Теорема 1. Пусть пространство A управляющих воздействий компактно, отображение $F:\ X\to(2)^A_{set}$, $x \to A_x$, полунепрерывно сверху. Кроме того, пусть выполняются следующие предположения:

1) $0<l<\tau(x,a)\le L<\infty$, $(x, a) \in \Delta$;

2) существует неотрицательная мера μ на $(X,\aleph)$, такая, что выполняются неравенства:

а) $\mu(B)\le P(B/x,a)$, $(x,a)\in\Delta$, $B\in\aleph$,

б) $\mu(X)>0$.

Тогда, если

а) функция $r(x, a)$ полунепрерывна сверху, а $\tau(x, a)$ непрерывна по x, a, $(x, a) \in \Delta$;

б) переходная вероятность $P(\,\cdot\,/\,x,a)$ слабо непрерывна по x, a; $(x, a) \in \Delta$, то в классе $\Re_1$ существует стационарная марковская детерминированная оптимальная стратегия δ^* с минимальной стоимостью

$$W = \frac{1}{L}\int_X v(x)\mu(dx),$$

где функция $v(x)$ – единственная в пространстве $\Xi(X)$ и определяется решением уравнения оптимальности

$$v(x) = \inf_{a\in A_x}\{r(x,a) + \int_X v(y)\,P'(dy\,/\,x,a)\},\quad x\in X,$$

где

$$P'(B\,/\,x,a) = P(B\,/\,x,a) - \frac{1}{L}\mu(B)\,\tau(x,a),\ \ B\in X.$$

Замечание 1. Данная теорема имеет место для функции издержек со значениями в [0; +∞), которые необходимо минимизировать. В работе [4] приведены условия максимизации вознаграждения (дохода) $r(x, a)$ за один период, если система находится в состоянии x и принято решение $a \in A_x$..

Здесь теорема 1 переформулирована с использованием отрицательной функции $r_1(x, a)$ = J $r(x, a)$.

1. Управление системой с многомерным фазовым пространством и пространством принятия решений.

Рассмотрим модель управления системой, у которой пространство состояний является декартовым произведением m множеств, т. е. $X = X_1 \times X_2 \times \ldots$ $\ldots \times X_m$. Пространство принимаемых решений $A = A_1 \times A_2 \times \ldots \times A_m$. Обозначим x_i^k – состояние i-й подсистемы после k-го перехода, a_i^k – принятое решение, τ_i^k – время пребывания i-й подсистемы в этом состоянии, $k = 0,1,2,\ldots$

Если в состоянии $x_i \in X_i$ принято решение $a_i \in A_{x_i}$ и время, проведенное в состоянии x_i равно t_i, то ожидаемые издержки i-й подсистемы за время s_i $(s_i \le t_i)$ равны $r_i(s_i\,/\,x_i, a_i)$. Функции $r_i(s_i\,/\,x_i, a_i)$ предполагаются измеримыми по Борелю на [0; + ∞) × Δ.

Пусть ожидаемые издержки всей системы за время s $r(s\,/\,x,a)$, где $x = (x_1, x_2, \ldots, x_m)$, $a = (a_1, a_2, \ldots, a_m)$ – сепарабельна, т. е. имеет вид $r(s\,/\,x,a) = \sum_{i=1}^{m} r_i(s_i\,/\,x_i, a_i)$.

Далее будем считать, что пространства X_i, A_i, и функции $r_i(s_i\,/\,x_i, a_i)$, $i = 1,\ldots,m$ удовлетворяют соответствующим условиям из п. 2.

Для выбранной стратегии рассмотрим следующий критерий оптимальности.

Средняя ожидаемая стоимость стратегии δ определяется величиной:

$$\varphi(x,\delta)=\sum_{i=1}^{m}\limsup_{n\to\infty}\frac{E_x^{\delta}\sum_{k=0}^{n}r_i(\tau_i^k/x_i^k,a_i^k)}{E_x^{\delta}\sum_{k=0}^{n}\tau_i^k}.$$

Стратегия δ^* оптимальна относительно этого критерия, если $\varphi(x,\delta^*)=$ $=\inf_{\delta\in R}\varphi(x,\delta),\ x\in X$.

Обозначим

$$\tau_i(x_i,a_i)=\int_{X_i}\int_0^{\infty}t\,d\Phi_i(t/x_i,a_i,y)\,P(dy/x_i,a_i)\ ,$$

$$r_i(x_i,a_i)=\int_{X_i}\int_0^{\infty}r_i(t/x_i,a_i)\,d\Phi_i(t,a_i,y)\,P(dy/x_i,a_i),$$

$$\Phi_i(t/x_i,a_i,y)=\begin{cases}1, & t\ge\tau_i(x_i,a_i),\\ 0, & t<\tau_i(x_i,a_i),\end{cases}$$

$$r_i(t/x_i,a_i)=\begin{cases}0, & t<\tau_i(x_i,a_i),\\ r_i(x_i,a_i), & t\ge\tau_i(x_i,a_i),\end{cases}$$

$$r(x,a)=\sum_{i=1}^{m}r_i(x_i,a_i).$$

Обозначим $\Xi_1(X)$ – банахово пространство ограниченных измеримых по Борелю функций на X с нормой $v(x)=\sum_{i=1}^{m}\sup_{x_i\in X_i}|v_i(x_i)|$.

Теорема 2. Пусть A – компактное пространство и отображение $F:X\to(2)_{set}^{A}$, $x\to A_x$, полунепрерывно сверху. Кроме того, пусть выполняются следующие предположения:

1) $0<l<\tau_i(x_i,a_i)\le L<\infty,\ (x_i,a_i)\in\Delta_i,\ i=\overline{1,m}$;

2) для каждого $i=\overline{1,m}$ существует неотрицательная мера μ_i на $(X_i,\aleph_i)$, такая, что выполняются неравенства:

а) $\mu_i(X_i)\le Q_i(B_i/x_i,a_i),\ \ B_i\in\aleph_i,\ \ i=1,m,$

б) $\mu_i(X_i)>0$.

Пусть также выполнены следующие условия:

3) функции $r_i(x_i,a_i)$ полунепрерывны сверху на (x_i,a_i), а $\tau_i(x_i,a_i)$ непрерывны по $x_i,a_i,\ (x_i,a_i)\in\Delta_i$;

4) переходные вероятности $Q_i(B_i/x_i,a_i)$ слабо непрерывны по (x_i,a_i).

Тогда в классе стационарных марковских детерминированных стратегий $\Re_0$ существует оптимальная стратегия δ^* с минимальной стоимостью

$$W = \frac{1}{L}\int V(x)\mu(dx),$$

где $V = \inf_{a\in A}\left\{ r(x,a) + \int_X V(y)Q'(dy/x,a)\right\} = \sum_{i=1}^{m}\inf_{a_i\in A_i}\{r_i(x_i,a_i) + \int_{X_i} V_i(y_i)[Q_i(dy_i/x_i,a_i) -$

$$-\frac{1}{L}\mu_i(dy_i)\tau_i(x_i,a_i)\prod_{j=1,\, j\neq i}^{m}\mu_j(X_j)]\}.$$

Доказательство. Из условия 3) теоремы вытекает, что $r(x,a)$ является полунепрерывной на (x,a) как сумма полунепрерывных сверху функций. $\tau(x,a) = (\tau_1(x_1,a_1),...,\tau_m(x_m,a_m))$ непрерывна по (x, a) как вектор непрерывных функций. Из условия 4) теоремы вытекает, что $Q = Q_1\times Q_2\times ...\times Q_m$ – слабо непрерывна.

Тогда условие теоремы 1 выполняется и φ-оптимальная стратегия существует с минимальной стоимостью

$$W = \frac{1}{L}\int V(x)\mu(dx),$$

где $V = \inf_{a\in A}\left\{ r(x,a) + \int_X V(y)Q'(dy/x,a)\right\}$, $x\in X$, где $Q'(B/x,a) = Q(B/x,a) - \frac{1}{L}\mu(B)\tau(x,a)$, $B\in\aleph$.

Условия 1) и 2) теоремы дают выполнение условий теоремы 1 для каждой i-й подсистемы, согласно которой существует оптимальная стратегия для каждой i-й подсистемы $(i = 1,...,m)$ с минимальной стоимостью $W_i = \frac{1}{L}\int V_i(x)\mu_i(dx)$, $i = \overline{1,m}$, где $V_i(x_i) = \inf_{a_i\in A_i}\left\{ r_i(x_i,a_i) + \int_{X_i} V_i(y)Q_i'(dy_i/x_i,a_i)\right\} =$

$$= \inf_{a_i\in A_i}\left\{ r_i(x_i,a_i) + \int_{X_i} V_i(y_i)\left[Q_i(dy_i/x_i,a_i) - \frac{1}{L}\mu_i(dy_i)\tau_i(x_i,a_i)\prod_{j=1,\, j\neq i}^{m}\mu_j(X_j)\right]\right\},$$

$$x_i\in X_i.$$

В силу сепарабельности функций $r(x,a)$, $x\in X$, $a\in A$ имеем

$$V(x) = \sum_{i=1}^{m}\sup V_i(x_i) =$$

$$= \sum_{i=1}^{m}\inf_{a_i\in A_i}\left\{ r_i(x_i,a_i) + \int_{X_i} V_i(y_i)\left[Q_i(dy_i/x_i,a_i) - \frac{1}{L}\mu_i(dy_i)\tau_i(x_i,a_i)\prod_{j=1,\, j\neq i}^{m}\mu_j(X_j)\right]\right\}.$$

Теорема доказана.

2. Многономенклатурные модели управления запасами

Рассмотрим систему управления запасами m продуктов, каждый из которых может непрерывно пополнятся. Предположим Q_i – максимальный уровень запаса i-го продукта, тогда запас i-го продукта принимает значение на $[0, Q_i]$.

В дискретные моменты времени N проверяется состояние запасов каждого продукта и принимаются соответствующие решения о пополнении складов следующим образом.

Если уровень запасов i-го продукта в момент времени $n \in N$ $X_i^n = x_i \in [0, Q_i]$, $x_i \in [0, Q_i]$, то производится заказ этого продукта $D_i^n \in A_i^x$, $A_i^x := [0, Q_i - x_i]$.

Таким образом пространство состояний системы, которое описывает развитие системы во времени, обозначим $X = X_1 \times X_2 \times ... \times X_m$, $X_i = [0, Q_i]$, $X = (X^n : n \in \mathrm{N})$. Пространство принимаемых решений обозначим $A = A_1^x \times A_2^x \times ... \times A_m^x$.

В момент времени $(n+)$ по каждому i-му продукту поступает случайное требование ξ_i^n и $\xi_i = (\xi_i^n : n \in \mathrm{N})$, $i = \overline{1, m}$ – последовательность независимых одинаково распределенных случайных величин с функциями распределения $G_i(x)$, $x \geq 0$, $i = \overline{1, m}$.

Обозначим $\xi = (\xi_1, \xi_2, ..., \xi_m)$, будем считать, что $\xi_n = (\xi_1^n, \xi_2^n, ..., \xi_m^n)$, $n \in \mathrm{N}$ не зависят от истории системы до момента времени n включительно и что $G_i(Q_i) < 1$, $G(Q) = G_1(Q_1) \times ... \times G_m(Q_m)$, $Q = (Q_1, ..., Q_m)$, а также $G_i(\cdot)$ непрерывны, $i = \overline{1, m}$.

Требование по i-му товару, которое поступило в момент времени $(n+)$, удовлетворяется (если это возможно) из запаса этого продукта $X_i^n + D_i^n$, который имеется в конце интервала времени $[n, n+1)$. Система такова, что дефицит или частичный дефицит какого-либо из m продуктов приводит к потери, а не откладыванию требования.

Итак, следующее уравнение описывает эволюцию процесса запаса системы

$$X^{n+1} = (X^n + D^n - \xi^n)_+, \; n \in \mathrm{N},$$

где $(a)_+ = \max(a, 0)$ – положительная часть $a \in \mathrm{R}_+$, или по каждому i-му продукту $(i = \overline{1, m})$.

$$X_i^{n+1} = (X_i^n + D_i^n - \xi_i^n)_+, \; n \in \mathrm{N}, \; i = \overline{1, m}.$$

Предположим, что время $\tau_i^n (\tau_i^n > 0)$ пребывания i-й системы в состоянии x_i^n – случайная величина с функцией распределения $\Phi_i(\cdot / X_i^n, D_i^n, X_i^{n+1})$, $i = \overline{1, m}$.

Будем считать, что функции $\Phi_i(\cdot / X_i^n, D_i^n, X_i^{n+1})$ измеримы по Борелю на $[0; Q_i] \times [0; Q_i] \times [0; Q_i]$.

Данные предположения гарантируют, что для описанной марковской стратегии последовательности $X_i = (X_i^n : n \in \mathrm{N})$ являются стационарными полумарковскими процессами, $i = \overline{1,m}$ и модель управления запасами удовлетворяет структуре полумарковского динамического программирования (см. разд. 2).

В дальнейшем предполагается, что все случайные величины определены на общем основном вероятностном пространстве модели (Ω, $\mathcal{F}$, P).

Модель управления запасами системы учитывает стоимость запаса (которая может включать издержки производства), стоимость хранения и дефицит для каждого i-го продукта. Издержки хранения уровня запаса x_i i-го продукта за время s_i, если в состоянии x_i i-ая подсистема пребывала время $t_i, s_i \le t_i$, составляет $C_i^1(s_i / x_i)$, $C_i^1 : [0, Q_i] \to \mathrm{R}_+$, а стоимость заказа продукции в размере x_i для i-го товара составляет $C_i^2(x_i)$, $C_i^2 : [0, Q_i] \to \mathrm{R}_+$, издержки, вызванные дефицитом, составляют $C_i^3(x_i)$, $C_i^3 : [0, +\infty) \to [0, +\infty)$, если требования x_i не могут быть выполнены.

Будем предполагать, что:

1) функции $C_i^1(s_i / x_i)$ – неотрицательные монотонно неубывающие по s_i и x_i, а функции $C_i^2(x_i)$, $C_i^3(x_i)$, $i = \overline{1,m}$ – монотонные неубывающие неотрицательные функции;

2) функции $C_i^3(x_i)$, $x_i \in [0, \infty)$ удовлетворяют $C_i^3(0) = 0$, $\int_0^\infty C_i^3(y) dG_i(y) < \infty$, $i = \overline{1,m}$.

Если i-я подсистема находится в состоянии x_i в начале периода, принято решение d_{a_i} о дозаказе в размере $a_i \in A_i$ и время проведенное до момента следующего дозаказа, равно t_i, то ожидаемые издержки за время s_i, $s_i \le t_i$ составляют:

$$r_i(s_i / x_i, d^0) = C_i^1(s_i / x_i) + \int_{x_i}^\infty C_i^3(y - x_i) dG_i(y),\ i = \overline{1,m}$$

и для $a_i > 0$

$$r_i(s_i / x_i, d^{a_i}) = C_i^1(s_i / x_i + a_i) + C_i^2(x_i) + \int_{x_i + a_i}^\infty C_i^3\left(y - (x_i + a_i)\right) dG_i(y),\ i = \overline{1,m}.$$

Причем $r(s / x, a) = \sum_{i=1}^m r_i(s_i / x_i, a_i)$.

Вероятности перехода $P\left(\cdot / x_i, a_i\right)$ на X_i для любого борелевского подмножества $[0, Q_i]$ задаются:

$$P\left(\left[y_i^1, y_i^2\right] / x_i, d^{a_i}\right) = G_i(x_i + a_i - y_i^1) - G_i(x_i + a_i - y_i^2),$$

$$a_i \in [0, Q_i - x_i],\ 0 \le y_i^1 < y_i^2 \le x_i + a_i,$$
$$P\left(\{0\} / x_i, d^{a_i}\right) = 1 - G_i(x_i + a_i -),\ x_i \in [0, Q_i].$$

Вероятности перехода системы $P\left(B / x, d^a\right) = \prod_{i=1}^{m} P(B_i / x_i, d^{a_i})$, где B_i – борелевские подмножества $[0, Q_i]$.

Обозначим

$$C_i^1(x_i, d^{a_i}) = \int_0^{Q_i}\int_0^{\infty} C_i^1(t / x_i)\, d\Phi_i(t / x_i, d^{a_i}, y)\, P_i(dy / x_i, d^{a_i}).$$

Тогда для модели управления запасами:

$$r_i(t_i / x_i, d^{x_i}) = \begin{cases} 0, & t_i < \tau_i(x_i, d^{a_i}), \\ r_i(x_i, d^{a_i}), & t_i \ge \tau_i(x_i, d^{a_i}), \end{cases}$$

где

$$r_i(x_i, d^0) = C_i^1(x_i, d^0) + \int_{x_i}^{\infty} C_i^3(y - x_i)\, dG_i(y),$$

$$r_i(x_i, d^{a_i}) = C_i^1(x_i + a_i, d^{a_i}) + C_i^2(a_i) + \int_{x_i + a_i}^{\infty} C_i^3\left(y - (x_i + a_i)\right) dG_i(y),\ i = \overline{1, m}.$$

для $a_i > 0$.

Теорема 3. Пусть функции C_i^1, C_i^2, C_i^3, $i = \overline{1, m}$ – полунепрерывны снизу и выполнено предположение 1 теоремы 2. Тогда для модели управления в классе $\mathfrak{R}$ всех допустимых стратегий существует φ-оптимальная стратегия с минимальной стоимостью $W = \frac{1}{L}\int V(x)\mu(dx)$.

Здесь $\mu(\cdot) = \mu_1(\cdot)...\mu_m(\cdot)$, $\mu_i(\cdot)$ – мера сконцентрированная в точке 0 с весом $\overline{G} = \overline{G_1}...\overline{G_m}$, $\overline{G_i} = 1 - G(x)$, $i = \overline{1, m}$, а $V(x)$ – удовлетворяет уравнению оптимальности

$$V(x) = KV(x) = \sum_{i=1}^{m} \min_{a \in A} \Big\{ \Big[C_i^1(x_i, d^{a_i}) + \int_{x_i}^{\infty} C_i^3\left(y - x_i\right) dG_i(y) + C_i^1(x_i + a_i, d^{a_i}) + C_i^2(a_i) +$$
$$+ \int_{x_i + a_i}^{\infty} C_i^3\left(y - (x_i + a_i)\right) dG_i(y) \Big] + \int_{X_i} V_i(y_i)[P(dy_i / x_i, a_i) -$$
$$- \frac{1}{L}\mu_i(dy_i)\tau_i(x_i, d^{a_i}) \prod_{j=1,\, j \ne i}^{m} \mu_j(x_j)\Big]\Big\}.$$

Доказательство. Применим теорему 2 о существовании φ-оптимальной стратегии, принадлежащей классу детерминированных (марковских) стратегий, для которых достигается минимальное значение издержек $W = \int V(x)\mu(dx)$. Проверим выполнение предположений этой теоремы.

1. Пусть $F:[0,Q_1]\times\ldots\times[0,Q_m]\to(2)^A_{set}, x\to A_x$ – отображение, которое связывает с каждым состоянием x набор допустимых решений A_x. Тогда F – полунепрерывно сверху. Действительно, если для $x,\ x^n\in X=[0,Q_1]\times\ldots\times[0,Q_m]$,

$$a^n=(a_1^n,\ldots,a_m^n)\in A^{x^n}=[0,Q_1-x_1^n]\times\ldots\times[0,Q_m-x_m^n],$$

$$\lim_{x\to\infty}x^n=x=(x_1,\ldots,x_m),\ \lim_{a\to\infty}a^n=a=(a_1,\ldots,a_m),$$

тогда в пределе $(0,\ldots,0)\le(a_1,\ldots,a_m)\le(Q_1-x_1,\ldots,Q_m-x_m)$, т. е. $a\in A^x$. Поэтому A – полунепрерывна сверху.

2. Докажем, что функции $r_i(x_i,d^{a_i})$ полунепрерывны снизу. В соответствии с предположением, функции $C_i^1(x_i,d^{a_i}),\ C_i^2(x_i),\ C_i^3(x_i),\ i=\overline{1,m}$ – полунепрерывны снизу. Из определения для $r(\cdot,\cdot)$ достаточно доказать полунепрерывность снизу

$$\int_{x_i+a_i}^{\infty}C_i^3\big(y-(x_i+a_i)\big)dG_i(y)\ \text{по}\ (x_i,a_i).$$

Введем случайный процесс $\eta_i(z_i)=\begin{cases}\xi_i-z_i, & \xi_i>z_i\\ 0, & \xi_i\le z_i\end{cases},\ \ z_i\in[0,Q_i].$

Тогда $u_i(x_i,a_i)=\int\limits_{x_i+a_i}^{\infty}C_i^3\big(y-(x_i+a_i)\big)dG_i(y)=EC_i^3\big(\eta_i(x_i+a_i)\big).$

Процесс $\eta_i(z_i)$ имеет непрерывную траекторию на $[0,Q_i]$. По лемме Фату получаем $\lim\limits_{(x_i',a_i')\to(x_i',a_i')}\inf Eu_i(x_i',a_i')\ge E\lim\limits_{(x_i',a_i')\to(x_i',a_i')}\inf Eu_i(x_i',a_i')\ge Eu_i(x_i,a_i)$ для любых $(x_i',a_i'),\ (x_i,a_i),$ то функция $\int\limits_{x_i+a_i}^{\infty}C_i^3\big(y-(x_i+a_i)\big)dG_i(y),$ и следовательно функции $r_i(x_i,d^{a_i}),\ i=\overline{1,m}$ полунепрерывны снизу.

3. Слабая непрерывность вероятностей перехода $P_i\big(B_i/x_i,d^{a_i}\big)$ следует из их определения.

4. Ограниченность выигрыша $r_i(x_i,d^{a_i}),\ i=\overline{1,m}$ следует из ограниченности $C_i^1,\ C_i^2,\ C_i^3,\ i=\overline{1,m}.$

Рассмотрим теперь задачу определения структуры оптимальной стратегии при условиях теоремы 3, которые далее считаются выполненными. Известно, что для многих систем управления запасами оптимальной является стратегия, для которой существует такой основной уровень запасов S, что после заказа оптимальный уровень запаса приближается к S. Так как есть возможность заказать любое количество продукта, то уровень S в точности достигается. Дальше обычно доказывается оптимальность (s,S) -стратегии: заказ на пополнение продукта производится только, когда уровень запаса меньше s.

Далее будем использовать следующие результаты, полученные в [18] для задачи управления запасами для системы с одним видом продукции.

Лемма 1. Пусть выполнены условия теоремы 3 и пусть для любого $x_i \in X_i$

$L_i(x_i, a_i) = C_i^1(x_i + a_i, d^{a_i}) + \int\limits_{x_i+a_i}^{\infty} C_i^3\left(y - (x_i + a_i)\right)dG_i(y) + C_i^2(a_i)$ монотонно убывают по $a_i \in (0, Q_i - x_i]$. Кроме того, пусть функции $\tau_i(x_i, d^{a_i})$ монотонно возрастают по $x_i \in X_i$ и $a_i \in (0; Q_i - x_i]$, а функции $C_i^1\left(x_i, d^{x_i}\right) - C_i^1\left(x_i, d^0\right)$ монотонно возрастают по $x_i \in X_i$. Тогда $V_i(x_i)$ монотонно убывает по $x_i \in X_i$.

Теорема 4. Пусть выполнены условия теоремы 3 и леммы 1, и при этом

$$\tau_i(x_i, d^0) - \tau_i(x_i, d^{Q_i - x_i})$$

монотонно возрастают по $x_i^* \in [0; Q_i]$, а

$$\tilde{L}_i(x_i) = C_i^1(x_i, d^0) - C_i^1(Q_i, d^{Q_i - x_i}) + \int\limits_{x_i}^{\infty} C_i^3\left(y - x_i\right)dG_i(y) - C_i^2(Q_i - x_i),\ x_i \in [0, Q_i),$$

монотонно убывают по $x_i \in [0, Q_i)$.

Тогда оптимальная стратегия $\delta_i^* \in \Re_i^1$ ($\Re_i^1$ – класс стационарных марковских детерминированных стратегий для *i*-го продукта) для задачи управления запасов имеет вид.

Существует порог $x_i^* \in [0, Q_i)$ такой, что

$$\delta_i^* = \begin{cases} d^{Q_i - x_i}, & x_i < x_i^*; \\ d^0, & x_i \geq x_i^*. \end{cases}$$

Замечание 2. Лемма 1 и теорема 4 приведены в обозначениях настоящей статьи и с учетом того, что в данной модели вероятность выполнения заказа для каждого продукта равна 1.

Докажем для многономенклатурной системы запасов следующий результат.

Теорема 5. Пусть выполняются условия теоремы 3, а также следующие условия:

1) $\forall x_i \in X_i \quad L_i(a_i, x_i) = C_i^1(x_i + a_i, d^{a_i}) + \int\limits_{x_i+a_i}^{\infty} C_i^3\left(y - (x_i + a_i)\right)dG_i(y) + C_i^2(a_i)$

монотонно убывают по $a_i \in (0, Q_i - x_i]$;

2) пусть функции $\tau_i(x_i, d^{a_i})$ монотонно возрастают по $x_i \in X_i$ и $a_i \in (0; Q_i - x_i]$, а функции $C_i^1\left(x_i, d^{x_i}\right) - C_i^1\left(x_i, d^0\right)$ монотонно возрастают по $x_i \in X_i$;

3) пусть $\tau_i(x_i, d^0) - \tau_i(x_i, d^{Q_i - x_i})$ монотонно возрастают по $x_i^* \in [0; Q_i]$,

а $\tilde{L}_i(x_i) = C_i^1(x_i, d^0) - C_i^1(Q_i, d^{Q_i - x_i}) + \int\limits_{x_i}^{\infty} C_i^3\left(y - x_i\right)dG_i(y) - C_i^2(Q_i - x_i), \quad x_i \in [0, Q_i),$

монотонно убывает по $x_i \in [0, Q_i)$.

Тогда φ-оптимальная стратегия $\delta^* \in \Re_1$ для задачи управления запасов имеет вид: существует порог $x^* = (x_1^*,...,x_m^*) \in [0,Q_1] \times ... \times [0,Q_m]$ такой, что $\delta^* = (\delta_1^*,...,\delta_m^*)$ и для $i = \overline{1,m}$

$$\delta_i^* = \begin{cases} d_{Q_i - x_i}, & x_i < x_i^*; \\ d_0 & , x_i \ge x_i^*. \end{cases}$$

Доказательство. Условия данной теоремы обеспечивают выполнение условий теоремы 4, которая задает структуру оптимальной стратегии δ_i^* $i = \overline{1,m}$ по каждому i-му товару, т. е. $\varphi_i(x_i, \delta_i^*) = \inf_{\delta_i \in \Re_i} \phi_i(x_i, \delta_i)$, $\varphi_i(x_i, \delta_i) =$

$= \lim_{n\to\infty}\sup \dfrac{\mathrm{E}_x^\delta \sum_{k=0}^n r_i(\tau_i^k / x_i^k, d_i^k)}{\mathrm{E}_x^\delta \sum_{k=0}^n \tau_i^k}$, где φ_i – средняя ожидаемая стоимость стратегии

δ_i, $\Re_i$ – класс допустимых стратегий для i-го товара, $i = \overline{1,m}$.

Поскольку $\varphi(x,\delta) = \sum_{i=1}^m \lim_{n\to\infty}\sup \dfrac{E_x^\delta \sum_{k=0}^n r_i(\tau_i^k / x_i^k, d_i^k)}{E_x^\delta \sum_{k=0}^n \tau_i^k} = \sum_{i=1}^m \varphi_i(x_i, \delta_i)$,

т. е. $\varphi(x,\delta^*) = \inf_{\delta\in\Re} \varphi_i(x,\delta) = \inf_{\delta\in\Re} \sum_{i=1}^m \varphi_i(x_i,\delta_i) = \sum_{i=1}^m \inf_{\delta_i\in\Re_i} \varphi_i(x_i,\delta_i) = \sum_{i=1}^m \varphi_i(x_i,\delta_i^*)$,
то утверждение теоремы выполнено.

Выводы. Таким образом, найдены условия существования оптимальной стратегии для полумарковской многономенклатурной системы. Рассмотрена полумарковская многономенклатурная модель управления запасами, которая учитывает стоимость заказа продуктов, стоимость хранения и издержки, вызванные дефицитом продуктов. Для данной модели найдены условия существования оптимальной стратегии, а также определена структура оптимальной стратегии при выполненных условиях оптимальности.

1. *Porteus E.L., Heyman Ed.D.P., Sobel M.J.* Stochastic Inventory Theory // Stochastic Models: Handbooks Oper. Res. and Manag. Sci. – Amsterdam: North Holland, 1990. – **2**, chap. 12. – P. 605 – 652.
2. *Lippman S.A.* Maximal Average-Reward Policies for Semi-Markov Decision Processes with Arbitrary State and Action Space // Ann. Math. Stat. – 1971. – **42**, N 5. – P. 1717 – 1726.
3. *Губенко Л.Г., Штатланд Э.С.* Об управляемых марковских процессах с дискретным временем // Теория вероятностей и математическая статистика. – 1972. – № 7. – С. 51 – 64.
4. *Губенко Л.Г., Штатланд Э.С.* Об управляемых полумарковских процессах // Кибернетика. – 1972. – № 2. – С. 26 – 29.
5. *Губенко Л.Г., Штатланд Э.С.* Управляемые марковские и полумарковские модели и некоторые конкретные задачи оптимизации стохастических систем // Управляемые случайные процессы и системы. – Киев, 1973. – С. 87 – 119.
6. *Yasuda Masami.* Semi-Markov Decision Processes with Countable State Space and Compact Action Space // Bull. Math. Statist. – 1978. – **18**, N 1 – 2. – P. 35 – 54.

7. *Federgruen A., Tijms H.S.* The Optimality Equation in Average Cost Denumerable State Semi-Markov Decision Problems, Recurrency Conditions and Algorithms // J. Appl. Probab. – 1978. – **15**, N 2. – P. 356 – 373.
8. *Kitaev M.* Elimination of Randomization in Semi-Markov Decision Models with Average Cost Criterion // Optimization. – 1987. – **18**, N 3. – P. 439 – 446.
9. *Ksir B.* Controle Optimal des Processus Semi-Markoviens sur des Espaces Compacts Metriques et Solution au Probleme de Replacement d'un Systeme Soumis a des Chocs Aleatoires Semi-Markoviens // Cah. Rech. – 1982. – N 17. – P. 23 – 52.
10. *Kurano Masami.* Semi-Markov Decision Processes and Their Applications in Replacement Models // J. Oper. Res. Soc. Jap. – 1985. – **28**, N 1. – P. 18 – 30.
11. *Wakuta K.* Arbitrary State Semi-Markov Decision Processes with Unbounded Rewards // Optimization. – 1987. – **18**, N 3. – P. 447 – 454.
12. *Виноградская Т.М., Генинсон Б.А., Рубчинский А.А.* Полумарковские процессы принятия решений с векторными доходами // Теория вероятностей и ее применения. – 1983. – **28**, № 1. – С. 182 – 184.
13. *Китаев М.Ю.* Полумарковские и скачкообразные марковские управляемые модели. Критерий средней цены // Там же. – 1985. – **30**, № 2. – С. 252 – 268.
14. *Юшкевич А.А.* О полумарковских управляемых моделях с критерием среднего дохода // Там же. – 1981. – **26**, № 4. – С. 808 – 815.
15. *Maitra A.P., Sudderth W.D.* Discrete Gambling and Stochastic Games. – New York, Berlin, Heidelberg: Springer-Verlag Inc., 1996. – 248 p.
16. *Vega-Amaya O.* Average Optimality in Semi-Markov Control Models on Borel Spaces: Unbounded Cost and Controls // Bol. Soc. Math. Mex. – 1993. – **38**, N 1 – 2. – P. 47 – 60.
17. *Xiao Guonueng, Guo Xianping, Liu Qingping.* The Optimizing Condition for Semi-Markov Decision Programming with Average Criterion // Hunan Ann. Math. – 1995. – **15**, N 1. – P. 6 – 13.
18. *Демченко С.С., Кнопов П.С., Чорней Р.К.* Оптимальные стратегии для полумарковской системы запасов // Кибернетика и системный анализ. – 2002. – № 1. – С. 146 – 160.
19. *Rockafellar R.T.* Measurable Dependence of Convex Sets and Functions on Parameter // J. Math. Anal. and Appl. – 1969. – **28**, N 1. – P. 45 – 57.

Т.В. Пепеляєва, І.Ю. Демченко

ПРО ОДНУ БАГАТОНОМЕНКЛАТУРНУ МОДЕЛЬ ДЛЯ НАПІВМАРКІВСЬКОЇ СИСТЕМИ ЗАПАСІВ

Розглянута задача керування напівмарківською багатономенклатурною системою запасів. Знайдені умови існування оптимальної стратегії, а також визначено структуру оптимальної стратегії при виконаних умовах оптимальності.

T.V. Pepelyaeva, S.Yu. Demchenko

ABOUT SOME MULTI-TASK MODEL FOR SEMI-MARKOV INVENTORY SYSTEM

We consider the control problem for semi-Markov inventory system. The existence conditions for the optimal strategy are found, and the optimal strategy structure is determined in case the existence conditions take place.

Получено 25.09.2015

Об авторах:

Пепеляева Татьяна Владимировна,
старший научный сотрудник Института кибернетики имени В.М. Глушкова НАН Украины,

Демченко Ирина Юриевна,
аспирантка Киевского Национального университета имени Тараса Шевченко.