

П.С. КНОПОВ

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: knopov1@yahoo.com.

Т.В. ПЕПЕЛЯЄВА

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: pepelaev@yahoo.com.

КЕРОВАНІ СТОХАСТИЧНІ СИСТЕМИ

Анотація. Розглянуто керовані марковські та напівмарковські процеси і системи. Наведено огляд прикладних задач керування запасами. Розглянуто одно- та багатомасштабні моделі з різними типами функції витрат та критеріями оптимальності. Досліджено умови оптимальності та вигляд оптимальних стратегій у цих задачах.

Ключові слова: марковські процеси, напівмарковські процеси, керування запасами, (s, S) -стратегія, критерій оптимальності, оптимальна стратегія.

Керовані марковські та напівмарковські процеси і системи широко використовуються під час розв'язання багатьох задач оптимального керування в різних галузях техніки й економіки. У статті розглянуто методи й алгоритми, які істотно використовують результати, що стосуються існування оптимальних розв'язків для керованих випадкових процесів і полів. Досліджено методи й алгоритми знаходження оптимальних стратегій для низки задач теорії запасів.

1. КЕРОВАНІ ВИПАДКОВІ ПРОЦЕСИ

Наведемо деякі відомості з теорії керування випадковими процесами, які знадобляться надалі. Розглянемо керовану стохастичну систему з дискретним часом (система з випадковими впливами у випадкові моменти часу), що керується у деякий спосіб задля мінімізації витрат, пов'язаних із системою керування.

Нехай X і A — множини станів і керувань, X — фазовий простір (простір станів) стохастичного процесу $X = (X_n: n \in \mathbb{N})$, який описує розвиток системи в часі, A — простір прийняття рішень. Простори X і A є повними сепарабельними метричними просторами з борелівськими σ -алгебрами $\mathfrak{X}$ і $\mathfrak{A}$ відповідно. Якщо в момент часу $n \in \mathbb{N}$ система перебуває у стані $x \in X$, то приймають рішення $D_n = a$, $a \in A_x \in \mathfrak{A}$, де A_x — набір допустимих дій у стані x . Позначимо $A: X \rightarrow \mathfrak{A}$, $x \rightarrow A_x$ відображення, що пов'язує набір допустимих дій із відповідним станом системи. Вважатимемо, що $\Delta = \{(x, a), x \in X, a \in A_x\} \in (\mathfrak{X} \otimes \mathfrak{A})$ — борелівські підмножини простору $X \times A$.

1.1. Керовані марковські процеси. Випадкова еволюція системи керується множиною перехідних імовірностей

$$P\{B|x_n, a_n\} = P\{X_{n+1} \in B | X_0 = x_0, D_0 = a_0, \dots, X_n = x_n, D_n = a_n\},$$

де $B \in \mathfrak{X}$, $(x_k, a_k) \in \Delta$ і x_k — стан системи в момент часу k , a_k — вибране керування в момент часу k , $k \leq n$.

Позначимо $r(x, a)$ витрати за один період, якщо система перебуває у стані x на початку періоду і приймається рішення $a \in A_x$. Будемо вважати, що функція $r(x, a)$ — обмежена вимірна функція на Δ , $|r(x, a)| \leq C < \infty$, $(x, a) \in \Delta$, для деякого $0 < C < \infty$.

Загальною допустимою стратегією керування системою є послідовність $\delta = \{\delta_0, \delta_1, \dots, \delta_n, \dots\}$ ймовірнісних розподілів така, що ймовірнісна міра $\delta_n(\cdot | x_0, a_0, \dots, x_n)$ на $(A, \mathfrak{A})$ зосереджена на A_{x_n} .

Стратегія δ називається стаціонарною марковською, якщо $\delta_n(\cdot|x_0, a_0, \dots, x_n) = \delta(\cdot|x_n)$, $n = 0, 1, \dots$. Стаціонарна марковська стратегія є детермінованою, якщо міра $\delta(\cdot|x)$ зосереджена в точці для будь-якого $x \in X$. Позначимо $\delta(x)$ точку концентрації маси $\delta(\cdot|x)$.

Нехай $\mathfrak{R}$ — клас допустимих стратегій і $\mathfrak{R}_1$ — клас стаціонарних нерандомізованих стратегій.

Розглянемо два критерії оптимальності стратегій:

1. Середня вартість для вибраної стратегії δ

$$\varphi(x, \delta) = \lim_{n \rightarrow \infty} \sup \frac{1}{n+1} E_x^\delta \sum_{k=0}^n r(X_k, D_k);$$

2. Очікувана (дисконтована) загальна вартість із $\beta \in (0, 1)$ і стратегією δ

$$\psi_\beta(x, \delta) = E_x^\delta \sum_{k=0}^{\infty} \beta^k r(x_k, D_k),$$

де $E_x^\delta(\cdot)$ — математичне сподівання за ймовірнісною мірою, що індукована стратегією керування δ .

Стратегія δ^* є оптимальною за цими критеріями, якщо $\varphi(x, \delta^*) = \inf_{\delta \in \mathfrak{R}} \varphi(x, \delta)$ (φ -оптимальною) або $\psi_\beta(x, \delta^*) = \inf_{\delta \in \mathfrak{R}} \psi_\beta(x, \delta)$ (ψ_β -оптимальною).

Позначимо $\Xi(X)$ банахів простір обмежених вимірних за Борелем функцій на X з нормою $\|u\| = \sup_{x \in X} |u(x)|$.

Першою задачею, що виникає, є знаходження умов існування φ -оптимальних і ψ_β -оптимальних стратегій у тому чи іншому класі множини керувань. Тут слід розрізняти, які умови накладено на множини станів і керувань для доведення цих теорем. Найпростішим є випадок, коли X і A є скінченними. Для цього випадку теорему про існування оптимальної стратегії у класі детермінованих стратегій доведено в [1]. Доведення здійснено шляхом безпосередньої перевірки оптимальності деякого рішення з використанням рівняння оптимальності, побудованого на основі методу динамічного програмування.

Значно більш складною виявилася задача доведення оптимального керування для випадку, коли хоча б одна з множин X і A не буде скінченною. У разі компактності або зліченності множин X і A твердження про існування оптимальних стратегій у класі стаціонарних нерандомізованих стратегій вперше отримано у [2–4] і потім досліджено багатьма математиками. Детальну бібліографію наведено, наприклад, у [5].

Теорема 1 [2, 3]. Нехай простір керувань A є компактим і відображення $A: X \rightarrow \mathfrak{A}$, $x \rightarrow A_x$, є напівнеперервним зверху. Нехай існує міра μ на $(X, \mathfrak{N})$, $\mu(X) > 0$, така, що для перехідної ймовірності $P(\cdot|x)$

$$P\{B|x\} \geq \mu(B), \quad (x, a) \in \Delta, \quad B \in \mathfrak{N}.$$

До того ж нехай на множині Δ також виконано такі умови:

- 1) функція вартості $r(x, a)$ неперервна (напівнеперервна зверху) за (x, a) ;
- 2) ймовірність переходу $P(\cdot|x, a)$ слабко неперервна за (x, a) .

Тоді у класі $\mathfrak{R}_1$ існує стаціонарна нерандомізована φ -оптимальна стратегія з мінімальною вартістю

$$W = \int V(x) \mu(dx),$$

де $V(x)$ — єдина в $\Xi(X)$ і визначається розв'язком рівняння оптимальності

$$V(x) = \inf_{a \in A_x} \{r(x, a) + \int V(y) P'(dy|x, a)\}, \quad x \in X,$$

де $P'(B|x, a) = P(B|x, a) - \mu(B)$, $B \in \mathfrak{N}$.

Зауваження 1. Умови на множини X і A можуть бути й іншими, водночас можна одержувати різні варіанти теорем існування оптимальних стратегій.

Наприклад, можна припускати, що X є компактною, а A — скінченною або X є зліченною, а A — компактною. У цьому випадку також справджується твердження про існування оптимальної стратегії у класі стаціонарних нерандомізованих стратегій. Є й інші варіанти умов існування оптимальних стратегій. Аналогічні результати отримано для ψ_β -оптимальних стратегій.

1.2. Керовані напівмарковські процеси. Якщо у стані $x \in X$ прийнято рішення $a \in A_x$, то:

1) такий стан системи вибирається за допомогою перехідної ймовірності $P(\cdot / x, a)$;

2) за умови, що такий стан системи є $y \in X$, час перебування у стані x є випадковою величиною із функцією розподілу $\Phi(\cdot / x, a, y)$.

Будемо припускати, що $P(\cdot / x, a)$ і $\Phi(\cdot / x, a, y)$ є борелівськими функціями на Δ і $\Delta \times X$ відповідно.

Нехай x_n — стан системи після n -го переходу, a_n — прийняте рішення, а τ_n — час перебування у цьому стані ($n = 0, 1, 2, \dots$).

Допустима стратегія δ для керованої систем визначається як послідовність $\delta = \{\delta_0, \delta_1, \dots, \delta_n, \dots\}$ ядер переходу така, що ймовірнісна міра $\delta_n(\cdot / h_n)$ на $(A, \mathfrak{F})$ зосереджена на A_{X_n} і така, що вимірним чином залежить від $h_n = (x_0, a_0, \tau_0, \dots, x_{n-1}, a_{n-1}, \tau_{n-1}, x_n)$ — історії керованої системи до моменту n -го переходу. Стратегію δ називають марковською, якщо $\delta_n(\cdot / h_n) = \delta_n(\cdot / x_n)$, ($n = 0, 1, 2, \dots$). Марковську стратегію δ називають стаціонарною, якщо $\delta_n(\cdot / x_n) = \delta(\cdot / x_n)$, ($n = 0, 1, 2, \dots$), і стаціонарною нерандомізованою (детермінованою), якщо міра $\delta(\cdot / x_n)$ вироджена і зосереджена в точці для будь-якого $x \in X$. У цьому випадку $\delta(x)$ будемо позначати точку зосередження маси $\delta(\cdot / x)$. Нехай (як і в п. 1.1) $\mathfrak{R}$ — клас усіх допустимих стратегій, $\mathfrak{R}_1$ — клас стаціонарних нерандомізованих стратегій.

Вибором стратегії δ визначаємо випадковий процес, який називатимемо процесом, керованим стратегією δ . Якщо стратегія δ — марковська стаціонарна, то керований процес є напівмарковським процесом.

Якщо у стані $x \in X$ прийнято рішення $a \in A_x$ і час, проведений у стані x , дорівнює t , то очікувані витрати за час s ($s \leq t$) становлять $r(s / x, a)$. Функцію $r(s / x, a)$ вважаємо вимірною за Борелем на $[0; +\infty) \times \Delta$.

Розглянемо такий критерій оптимальності вибраної стратегії.

Середня очікувана вартість стратегії δ має вигляд

$$\varphi(x, \delta) = \limsup_{n \rightarrow \infty} \frac{E_x^\delta \sum_{k=0}^n r(\tau_k / x_k, a_k)}{E_x^\delta \sum_{k=0}^n \tau_k}, \quad (1)$$

де $\xi_0 = x$, а E_x^δ — математичне сподівання, що відповідає процесу, керованому стратегією δ за умови, що $\xi_0 = x$.

Стратегія δ^* є оптимальною за цим критерієм, якщо

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

Нехай

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

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

Припустимо, що $\tau(x, a)$ і $r(x, a)$ існують і є скінченними для всіх $(x, a) \in \Delta$ та $|r(x, a)| \leq C < \infty$, $(x, a) \in \Delta$. Оскільки критерій (1) залежить тільки від $P(\cdot / x, a)$ і усереднених характеристик $\tau(x, a)$ і $r(x, a)$, обмежимося розглядом

керуваних процесів, для яких

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

$$r(t/x, a) = \begin{cases} 0, & t < \tau(x, a), \\ r(x, a), & t \geq \tau(x, a). \end{cases} \quad (2)$$

Нехай $\Xi(X)$ — банахів простір обмежених вимірних за Борелем функцій на X із нормою $\|u\| = \sup_{x \in X} u(x)$.

Теорема 2 [3, 4]. Нехай простір керувальних впливів A є компактним, відображення $F: X \rightarrow (2)_{set}^A$, $x \rightarrow A_x$, є напівнеперервним зверху. Крім того, нехай виконуються такі припущення:

- 1) $0 < l < \tau(x, a) \leq L < \infty$, $(x, a) \in \Delta$;
- 2) існує невід’ємна міра μ на $(X, \mathfrak{N})$ така, що виконуються нерівності:
 - а) $\mu(B) \leq P(B/x, a)$, $(x, a) \in \Delta$, $B \in \mathfrak{N}$,
 - б) $\mu(X) > 0$.

Тоді, якщо

а) функція $r(x, a)$ напівнеперервна зверху, а $\tau(x, a)$ неперервна за x , a , $(x, a) \in \Delta$;

б) перехідна ймовірність $P(\cdot/x, a)$ слабко неперервна за x , a , $(x, a) \in \Delta$, то у класі $\mathfrak{R}_1$ існує стаціонарна нерандомізована оптимальна стратегія δ^* з мінімальною вартістю

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

де функція $v(x)$ — єдина у просторі $\Xi(X)$ і визначається розв’язком рівняння оптимальності

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

де

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

Зауваження 2. Ця теорема справедлива для функції витрат зі значеннями в $[0; +\infty)$, яку потрібно мінімізувати. У [2, 4] наведено умови максимізації винагороди (прибутку) $r(x, a)$ за один період, якщо система перебуває у стані $x \in X$ і прийнято рішення $a \in A_x$.

2. КЕРОВАНІ СТОХАСТИЧНІ СИСТЕМИ ІЗ СЕПАРАБЕЛЬНОЮ ФУНКЦІЄЮ ВИТРАТ (ІЗ БАГАТОВИМІРНИМИ ФАЗОВИМ ПРОСТОРОМ І ПРОСТОРОМ ПРИЙНЯТТЯ РІШЕНЬ)

2.1. Марковські керовані стохастичні системи. Розглянемо модель керування системою з багатовимірними фазовим простором і простором прийняття рішень.

Нехай простір станів є декартовим добутком m множин, тобто $X = X_1 \times X_2 \times \dots \times X_m$. Простір прийнятих рішень $A = A_1 \times A_2 \times \dots \times A_m$.

Для кожної пари $x_i \in X_i$, $a_i \in A_i$, позначимо $r_i(x_i, a_i)$ очікувані витрати (витрати) за один період, якщо i -та підсистема перебуває у стані x_i на початку періоду і приймається рішення $a_i \in A_i$.

Нехай очікувані витрати всієї системи за один період становлять $r(x, a)$, де $x = (x_1, x_2, \dots, x_m)$, $a = (a_1, a_2, \dots, a_m)$, і є сепарабельною функцією, тобто має

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

Далі вважатимемо, що простори X_i , A_i , $i=1, \dots, m$, і функції $r_i(x_i, a_i)$ задовольняють відповідні умови, наведені в п 1.1 для $r(x, a)$ загального вигляду.

Тоді критерій φ -оптимальності цієї стратегії запишемо так:

$$\varphi(x, \delta) = \limsup_{n \rightarrow \infty} \frac{1}{n+1} E_x^\delta \sum_{k=0}^n r(X_k, D_k) = \limsup_{n \rightarrow \infty} \frac{1}{n+1} E_x^\delta \sum_{k=0}^n \sum_{i=1}^m r_i(x_i^k, d_i^k),$$

де $X_k = (x_1^k, x_2^k, \dots, x_m^k)$ — стан системи в момент часу k , $D_k = (d_1^k, d_2^k, \dots, d_m^k)$ — вибране керування в момент часу k .

Позначимо $\Xi_1(X)$ банахів простір обмежених вимірних за Борелем функцій на X з нормою $v(x) = \sum_{i=1}^m \sup_{x_i \in X_i} |v_i(x_i)|$.

У [6] доведено таке твердження.

Теорема 3. Нехай A — компактний простір і відображення $A: X \rightarrow 2^A$ є напівнеперервним зверху, та нехай існує $\mu_i(X_i) > 0$ на $(X, \mathfrak{N})$, $i = \overline{1, m}$:

$$\mu_i(X_i) \leq Q_i(B_i / x_i, a_i), B_i \in \mathfrak{N}, i = \overline{1, m}.$$

Нехай також виконано такі умови:

- 1) функції $r_i(x_i, a_i)$ напівнеперервні знизу на (x_i, a_i) ;
- 2) перехідні ймовірності $Q_i(B_i / x_i, a_i)$ слабко неперервні на (x_i, a_i) .

Тоді у класі стаціонарних нерандомізованих стратегій існує оптимальна стратегія з мінімальною вартістю

$$W = \int V(x) \mu(dx),$$

де

$$V = \inf_{a \in A} \left\{ r(x, a) + \int_X V(y) Q'(dy / x, a) \right\} =$$

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

2.2. Напівмарковські керовані стохастичні системи. Розглянемо модель керування системою, у якої простір станів є декартовим добутком m множин, тобто $X = X_1 \times X_2 \times \dots \times X_m$. Простір прийняття рішень $A = A_1 \times A_2 \times \dots \times A_m$. Позначимо: x_i^k — стан i -ї підсистеми після k -го переходу, a_i^k — прийняте рішення, τ_i^k — час перебування i -ї підсистеми у цьому стані, $k = 0, 1, 2, \dots$.

Якщо у стані $x_i \in X_i$ прийнято рішення $a_i \in A_{x_i}$ і час, проведений у стані x_i , дорівнює t_i , то очікувані витрати i -ї підсистеми за час s_i ($s_i \leq t_i$) становлять $r_i(s_i / x_i, a_i)$. Функції $r_i(s_i / x_i, a_i)$ вважаємо вимірними за Борелем на $[0; +\infty) \times \Delta$.

Нехай очікувані витрати $r(s / x, a)$ всієї системи за час s , де $x = (x_1, x_2, \dots, x_m)$, $a = (a_1, a_2, \dots, a_m)$, є сепарабельною функцією, тобто має вигляд $r(s / x, a) = \sum_{i=1}^m r_i(s_i / x_i, a_i)$.

Далі будемо вважати, що простори X_i , A_i , і функції $r_i(s_i / x_i, a_i)$, $i = \overline{1, m}$, задовольняють відповідні умови.

Для вибраної стратегії розглянемо такий критерій оптимальності.

Середня очікувана вартість стратегії δ визначається величиною

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

Стратегія δ^* є оптимальною за цим критерієм, якщо $\varphi(x, \delta^*) = \inf_{\delta \in \mathfrak{H}} \varphi(x, \delta)$, $x \in X$.

Позначимо

$$\begin{aligned} \tau_i(x_i, a_i) &= \int_{X_i} \int_0^\infty t d\Phi_i(t / x_i, a_i, y) P(dy / x_i, a_i), \\ r_i(x_i, a_i) &= \int_{X_i} \int_0^\infty r_i(t / x_i, a_i) d\Phi_i(t, a_i, y) P(dy / x_i, a_i), \\ \Phi_i(t / x_i, a_i, y) &= \begin{cases} 1, & t \geq \tau_i(x_i, a_i), \\ 0, & t < \tau_i(x_i, a_i), \end{cases} \\ r_i(t / x_i, a_i) &= \begin{cases} 0, & t < \tau_i(x_i, a_i), \\ r_i(x_i, a_i), & t \geq \tau_i(x_i, a_i), \end{cases} \\ r(x, a) &= \sum_{i=1}^m r_i(x_i, a_i). \end{aligned}$$

Нехай $\Xi_1(X)$ — банахів простір обмежених вимірних за Борелем функцій на X з нормою $v(x) = \sum_{i=1}^m \sup_{x_i \in X_i} |v_i(x_i)|$.

У [7, 8] доведено таку теорему.

Теорема 4. Нехай A — компактний простір і відображення $F: X \rightarrow (2)_{set}^A$, $x \rightarrow A_x$, є напівнеперервним зверху. До того ж нехай виконуються такі припущення та умови:

- 1) $0 < l < \tau_i(x_i, a_i) \leq L < \infty$, $(x_i, a_i) \in \Delta_i$, $i = \overline{1, m}$;
- 2) для кожного $i = \overline{1, m}$ існує невід’ємна міра μ_i на $(X_i, \mathfrak{N}_i)$ така, що виконуються нерівності:
 - а) $\mu_i(X_i) \leq Q_i(B_i / x_i, a_i)$, $B_i \in \mathfrak{N}_i$, $i = \overline{1, m}$,
 - б) $\mu_i(X_i) > 0$.
- 3) функції $r_i(x_i, a_i)$ напівнеперервні зверху на (x_i, a_i) , а $\tau_i(x_i, a_i)$ — неперервні за x_i, a_i , $(x_i, a_i) \in \Delta_i$;
- 4) перехідні ймовірності $Q_i(B_i / x_i, a_i)$ слабко неперервні за (x_i, a_i) .

Тоді у класі стаціонарних нерандомізованих стратегій $\mathfrak{H}_1$ існує оптимальна стратегія δ^* з мінімальною вартістю

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

де

$$\begin{aligned} V &= \inf_{a \in A} \{r(x, a) + \int_X V(y) Q'(dy / x, a)\} = \sum_{i=1}^m \inf_{a_i \in A_i} \left\{ r_i(x_i, a_i) + \right. \\ &\left. + \int_{X_i} V_i(y_i) \left[Q_i(dy_i / x_i, a_i) - \frac{1}{L} \mu_i(dy_i) \tau_i(x_i, a_i) \prod_{j=1, j \neq i}^m \mu_j(X_j) \right] \right\}. \end{aligned}$$

3.1. Марковські моделі керування запасами

3.1.1. Однономенклатурні моделі керування запасами для марковських систем. Модель керування запасами є такою. Розглянемо систему керування запасами одного продукту, що може неперервно поповнюватись. Максимальний рівень запасу дорівнює Q , так що запас продукту набуває значень на $[0, Q]$.

Періодично в дискретні моменти часу виконують перевірку і залежно від фактичного запасу продукту приймають рішення про поповнення складів у такий спосіб: якщо рівень запасів продукту X_n у момент часу $n \in \mathbb{N}$ дорівнює $x \in [0, Q]$, то здійснюють замовлення продукту $D_n \in A_x$, $A_x := [0, Q - x]$.

У момент перевірки ($n+$) надходить випадкова вимога ξ_n і $\xi = (\xi_n : n \in \mathbb{N})$ — послідовність незалежних випадкових величин із функцією розподілу $G(x)$, $x \geq 0$.

Вважатимемо, що ξ_n не залежать від історії системи до моменту часу n включно, $G(x) < 1$, та $G(\cdot)$ є неперервною.

Рівняння еволюції процесу запасу системи має такий вигляд:

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

де $(a)_+ = \max(a, 0)$ — додатна частина $a \in \mathbb{R}_+$.

Надалі вважатимемо, що усі випадкові величини визначені на загальному основному ймовірнісному просторі моделі (Ω, F, P) .

Нехай $X = [0, Q]$, $X = (X^n : n \in \mathbb{N})$ — простір станів системи, що описує розвиток системи в часі, $A = \{d_a, a \in [0, Q]\}$ — простір прийняття рішень, d_a — рішення про дозамовлення в розмірі a . У стані x множина допустимих дій є такою: $A_x = \{d_a, a \in [0, Q - x]\}$.

Розглянемо задачу визначення структури оптимальної стратегії за умов теореми про існування оптимальної стратегії, які далі вважатимемо виконаними. Відомо, що для багатьох систем керування запасами оптимальною є стратегія, для якої існує такий основний рівень запасів S , що після замовлення оптимальний рівень запасу наближається до S . Оскільки є можливість замовити будь-яку кількість продукту, то рівень S в точності досягається. Далі зазвичай доводять оптимальність (s, S) -стратегії: замовлення на поповнення продукту здійснюють тільки тоді, коли рівень запасу менше за s і замовлення дорівнює $S - x$.

Далі розглянемо один із варіантів умов, за яких оптимальною буде (s, S) -стратегія. Припустимо, що функція витрат поповнення запасу має вигляд

$$C(x) = A + Cx, \quad A > 0, \quad C > 0.$$

Нехай $f(x)$ — функція втрат, що пов'язана зі зберіганням та дефіцитом товару в розмірі x , яку будемо припускати опуклою та неперервною.

Отже, функція витрат $r(x, d_a)$ має такий вигляд:

$$r(x, d_0) = f(x),$$

$$r(x, d_a) = f(x + a) + C(a).$$

Ймовірності переходу для будь-якої борелівської підмножини $[0, Q]$ задаються таким співвідношенням:

$$P\{[y_1, y_2] / x, d_a\} = G(x + a - y_1) - G(x + a - y_2),$$

$$a \in [0, Q - x], \quad 0 < y_1 \leq y_2 \leq x, \quad 0 < x \leq y_1 \leq y_2 \leq x + a,$$

$$P\{\{0\} / x, d_a\} = 1 - G(x + a), \quad a \in [0, Q - x].$$

Теорема 5 [9]. Нехай функція $c \cdot x + f(x)$, $x \in [0, Q]$, $i = \overline{1, m}$, монотонно спадає.

Тоді для описаної вище моделі існує поріг $x^* \in [0, Q]$ такий, що

$$\delta^* = \begin{cases} d_{Q-x}, & x < x^*; \\ d_0, & x \geq x^*. \end{cases}$$

Зауваження 3. Однономенклатурні моделі керування запасами для критерію з переоцінкою детально розглянуто в [10].

Зауваження 4. Однономенклатурні моделі керування запасами для марковських систем із загальними функціями загальних витрат детально розглянуто в [9].

3.1.2. Багатономенклатурні моделі керування запасами для марковських систем зі спадними функціями загальних витрат. Розглянемо систему керування запасами m продуктів, кожний із яких можна неперервно поповнювати. Припустимо, що Q_i — максимальний рівень запасу i -го продукту, тоді запас i -го продукту набуває значень на $[0, Q_i]$. Стан запасів кожного продукту перевіряють у дискретні моменти часу N , після чого приймають відповідні рішення щодо поповнення складів у наведений нижче спосіб.

Якщо рівень запасів i -го продукту в момент часу $n \in \mathbb{N}$ становить $X_i^n = x_i \in [0, Q_i]$, то здійснюють замовлення цього продукту $D_i^n \in A_i^x$, $A_i^x := [0, Q_i - x_i]$. Тоді простір станів системи, що описує розвиток системи в часі, має вигляд $X = X_1 \times X_2 \times \dots \times X_m$, $X_i = [0, Q_i]$, $X = (X^n : n \in \mathbb{N})$, а простір прийнятих рішень — $A = A_1^x \times A_2^x \times \dots \times A_m^x$.

У момент часу $(n+)$ за кожним i -м продуктом надходить випадкова вимога ξ_i^n і $\xi_i = (\xi_i^n : n \in \mathbb{N})$, $i = \overline{1, m}$, — послідовність незалежних однаково розподілених випадкових величин із функціями розподілу $G_i(x)$, $x \geq 0$, $i = \overline{1, m}$.

Нехай $\xi = (\xi_1, \xi_2, \dots, \xi_m)$, і $\xi_n = (\xi_1^n, \xi_2^n, \dots, \xi_m^n)$, $n \in \mathbb{N}$, не залежать від історії системи до моменту часу n включно, $G_i(x) < 1$, $G(x) = G_1(Q_1) \times \dots \times G_m(Q_m)$, $Q = (Q_1, \dots, Q_m)$, і $G_i(\cdot) \in$ неперервними.

Вимога за i -м товаром, що надійшла в момент часу $(n+)$, задовольняється (якщо це можливо) із запасу цього продукту $X_i^n + D_i^n$, який є наприкінці інтервалу часу $[n, n+1)$. Система є такою, що дефіцит або частковий дефіцит якогось із m продуктів призводить до втрати, а не до відкладання вимоги.

Еволюцію процесу запасу системи описує таке рівняння:

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

де $(a)_+ = \max(a, 0)$ — додатна частина $a \in \mathbb{R}_+$, або за кожним i -м продуктом ($i = \overline{1, m}$)

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

Будемо припускати, що вартість замовлення (яка може включати витрати виробництва) складається з фіксованої вартості і лінійної функції відносно розміру замовлення. Тоді у разі замовлення $x_i > 0$ товарів очікувані витрати замовлення становитимуть $C_i(x_i) = A_i + c_i \cdot x_i$, $A_i \geq 0$, $c_i \geq 0$.

Далі припустимо, що загальні витрати збереження і дефіциту за період часу $f_i(x_i)$ залежать від рівня запасів $x_i \geq 0$, де $f_i(\cdot) \in$ напівнеперервними знизу.

Для системи, що перебуває у стані x на початку періоду, у разі прийняття рішення $a \in A$ очікувані витрати за один період становлять $r(x, a) = \sum_{i=1}^m r_i(x_i, a_i)$,

де $r_i(x_i, a_i)$ — очікувані витрати за i -м продуктом за один період часу, якщо стан відповідного продукту дорівнює x_i і на початку періоду прийнято рішення d^{a_i} :

$$r_i(x_i, d^0) = f_i(x_i), \quad i = \overline{1, m}, \quad (3)$$

і для $a_i > 0$

$$r_i(x_i, d^{a_i}) = f_i(x_i + a_i) + C_i(a_i), \quad i = \overline{1, m}. \quad (4)$$

Ймовірності переходу на X_i для будь-якої борелівської підмножини $[0, Q_i]$ задають так:

$$P([y_i^1, y_i^2] / x_i, d^{a_i}) = G_i(x_i + a_i - y_i^1) - G_i(x_i + a_i - y_i^2), \quad (5)$$

$$a_i \in [0, Q_i - x_i], \quad 0 \leq y_i^1 \leq y_i^2 \leq x_i + a_i,$$

$$P(\{0\} / x_i, d^{a_i}) = 1 - G_i(x_i + a_i -), \quad x_i \in [0, Q_i]. \quad (6)$$

Ймовірності переходу системи мають вигляд $P(B / x, d^a) = \prod_{i=1}^m P(B_i / x_i, d^{a_i})$, де B_i — борелівські підмножини $[0, Q_i]$.

Наступна теорема вказує на існування оптимальної стратегії для цієї моделі керування запасами.

Теорема 6 [7, 11]. Для розглядуваної марковської моделі керування у класі $\mathfrak{R}_1$ існує φ -оптимальна стратегія з мінімальною вартістю $W = \int V(x) \mu(dx)$.

Тут $\mu(\cdot) = \mu_1(\cdot) \dots \mu_m(\cdot)$, $\mu_i(\cdot)$ — міра, сконцентрована в точці 0 із вагою $\overline{G_i} = 1 - G(x)$, $i = \overline{1, m}$, $\overline{G} = \overline{G_1} \dots \overline{G_m}$, а $V(x)$ задовольняє рівняння оптимальності

$$V(x) = LV(x) = \sum_{i=1}^m \min_{a \in A} \left\{ f_i(x_i) + f_i(x_i + a_i) + C_i(a_i) + \right. \\ \left. + \int_{X_i} V_i(y_i) \left[P(dy_i / x_i, a_i) - \mu_i(dy_i) \prod_{j=1, j \neq i}^m \mu_j(x_j) \right] \right\}.$$

Наступна теорема дає змогу визначити структуру оптимальної стратегії для багатоміноміклатурної системи запасів.

Теорема 7 [7, 11]. Нехай $c_i \cdot x_i + f_i(x_i)$, $x_i \in [0, Q_i]$, $i = \overline{1, m}$, монотонно спадає.

Тоді φ -оптимальна стратегія $\delta^* \in \mathfrak{R}_1$ для задачі керування запасами має такий вигляд: існує поріг $x^* = (x_1^*, \dots, x_m^*) \in [0, Q_1] \times \dots \times [0, Q_m]$ такий, що $\delta^* = (\delta_1^*, \dots, \delta_m^*)$ і для $i = \overline{1, m}$

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

Зауваження 5. Багатоміноміклатурні моделі керування запасами для марковських систем із загальними функціями загальних витрат детально розглянуто в [6, 7].

3.1.3. Багатоміноміклатурні моделі керування запасами для критерію з переоцінкою. Розглянемо таку саму систему керування запасами, як у п. 3.1.2.

Критерій ψ_β -оптимальності стратегії запишемо в такий спосіб:

$$\begin{aligned}\psi_\beta(x, \delta) &= E_x^\delta \sum_{k=0}^{\infty} \beta^k r(x_k, D_k) = E_x^\delta \sum_{k=0}^{\infty} \beta^k \sum_{i=1}^m r_i(x_i^k, d_i^k) = \\ &= E_x^\delta \sum_{k=0}^{\infty} \sum_{i=1}^m \beta^k r_i(x_i^k, d_i^k),\end{aligned}$$

де $X_k = (x_1^k, x_2^k, \dots, x_m^k)$ — стан системи в момент часу k , $D_k = (d_1^k, d_2^k, \dots, d_m^k)$ — вибране керування в момент часу k .

Еволюція процесу запасу системи та вартість замовлення такі самі, як у п. 3.1.2.

Теорема 8 [12]. Нехай $r(x, a)$ визначається співвідношеннями (3) і (4), а $P(B/x, a)$ — співвідношеннями (5) і (6). Тоді у класі $\mathfrak{R}_1$ існує ψ_β -оптимальна стратегія, яку можна отримати методом послідовних наближень.

Знайдемо вигляд оптимальної стратегії.

Нехай $F_i(x_i)$, $i = \overline{1, m}$, визначаються такими співвідношеннями:

$$\begin{aligned}F_i(x_i) &= c_i x_i + f_i(x_i) + \int_0^{x_i} \psi_\beta(y_i) dG_i(x_i - y_i), \\ F_i^n(x_i) &= c_i x_i + f_i(x_i) + \int_0^{x_i} \psi_\beta^{(n-1)}(y_i) dG_i(x_i - y_i), \\ \psi_\beta^n(y_i) &= L \psi_\beta^{n-1}(y_i),\end{aligned}$$

де $LF_i(x_i) = \min_{x_i \in A_{x_i}} \{c_i x_i + f_i(x_i) + \int F_i(y_i) P(dy_i / x_i, a_i)\}$.

Нехай $S_i \in [0, Q_i]$ — точка локального мінімуму функції $F_i(x_i)$, $i = \overline{1, m}$, $A_i > 0$ з означення A_i -опуклої функції, $i = \overline{1, m}$.

Теорема 9 [12]. Нехай $C_i(x_i) = \overline{A_i} + C_i x_i$, $f_i(x_i)$ — опукла диференційовна на $[0, S_i]$ функція для кожного $i = \overline{1, m}$, та існує $g_i(x_i) = G'_i(x_i)$, $i = \overline{1, m}$. Тоді у класі $\mathfrak{R}_1$ існує ψ_β -оптимальна стратегія $\delta^* = (\delta_1^*, \dots, \delta_m^*)$ вигляду

$$\begin{aligned}\delta_i^* &= \begin{cases} d_{S_i^* - x_i}, & x_i < x_i^*; \\ d_0, & x_i \geq x_i^*, \end{cases} \text{ якщо } A_i > 0, \\ \delta_i^* &= d_{S_i^* - x_i}, \quad A_i = 0,\end{aligned}$$

де $S_i^* = \arg \min_{x_i \in [0, Q_i]} F_i(x_i)$, $x_i^* \in [0, S_i^*]$ для $i = \overline{1, m}$, $x^* = (x_1^*, \dots, x_m^*)$.

Для розв'язування практичних задач важливим є випадок, коли $S^* = Q$.

Теорема 10 [12]. Нехай функція $C_i x_i + f_i(x_i)$ монотонно спадає для кожного $i = \overline{1, m}$. Тоді у класі $\mathfrak{R}_1$ існує оптимальна стратегія $\delta^* = (\delta_1^*, \dots, \delta_m^*)$ вигляду

$$\begin{aligned}\delta_i^* &= \begin{cases} d_{Q_i - x_i}, & x_i < x_i^*; \\ d_0, & x_i \geq x_i^*, \end{cases} \text{ якщо } A_i > 0, \\ \delta_i^* &= d_{Q_i - x_i}, \quad A_i = 0.\end{aligned}$$

3.2. Напівмарковські моделі керування запасами.

3.2.1. Однономенклатурні моделі керування запасами для напівмарковських систем із загальними функціями витрат. Дослідимо задачу знаходження умови, що гарантує існування оптимальної стратегії (s, S) -структури для напівмарковської моделі.

Розглянемо систему керування запасами одного продукту, що може неперервно поповнюватися з еволюцією процесу запасу системи, як описано в п. 3.1.1.

Припустимо, що час τ_n ($\tau_n > 0$) перебування системи у стані x_n є випадковою величиною із функцією розподілу $\Phi(\cdot / X_n, D_n, X_{n+1})$.

Вважатимемо, що функція $\Phi(\cdot / X_n, D_n, X_{n+1})$ вимірна за Борелем на $[0; Q] \times [0; Q] \times [0; Q]$.

Модель керування запасами враховує вартість замовлення (яка може включати витрати виробництва), вартість зберігання і дефіциту. Витрати зберігання рівня запасу x за час s , якщо у стані x система перебувала протягом часу t , $s \leq t$, становлять $c_1(s/x)$, $c_1: R_+ \times [0; Q] \rightarrow R_+$, вартість замовлення продукції в розмірі x становить $c_2(x)$, $c_2: (0; Q] \rightarrow R_+$. Витрати, спричинені дефіцитом, становлять $c_3(x)$, $c_3: [0; +\infty) \rightarrow [0; +\infty)$, якщо вимоги x не можуть бути виконані.

Припустимо, що:

- 1) функція $c_1(s/x)$ невід'ємна монотонно неспадна за s та x , а функції $c_2(x)$, $c_3(x)$ невід'ємні монотонно неспадні за x ;
- 2) функція $c_3(x)$, $x \in [0; +\infty)$, задовольняє такі умови:

$$c_3(0) = 0, \quad \int_0^{+\infty} c_3(x) dG(x) < \infty.$$

Якщо система перебуває у стані $x \in X$, прийнято рішення d_a про дозамовлення в розмірі a і час, проведений до моменту такого дозамовлення, дорівнює t , то очікувані витрати $r(s/x, d_a)$ за час s ($s \leq t$) є такими:

$$r(s/x, d_0) = c_1(s/x) + \int_x^{+\infty} c_3(y-x) dG(y),$$

а для $a > 0$

$$\begin{aligned} r(s/x, d_a) &= p c_1(s/x + a) + (1-p) c_1(s/x) + c_2(x) + \\ &+ p \int_{x+a}^{+\infty} c_3(y - (x+a)) dG(y) + (1-p) \int_x^{+\infty} c_3(y-x) dG(y). \end{aligned}$$

Ймовірності $P(\cdot / x, a)$ на X задають для будь-якої борелівської підмножини відрізка $[0; Q]$ у такий спосіб:

$$P\{[y_1; y_2] / x, d_0\} = G(x - y_1) - G(x - y_2), \quad x \in [0; Q], \quad 0 \leq y_1 < y_2 \leq x;$$

$$P\{[y_1; y_2] / x, d_a\} =$$

$$= p(G(x + a - y_1) - G(x + a - y_2)) + (1-p)(G(x - y_1) - G(x - y_2)),$$

$$a \in [0; Q - x], \quad 0 \leq y_1 < y_2 \leq x;$$

$$P\{[y_1; y_2] / x, d_a\} = p(G(x + a - y_1) - G(x + a - y_2)), \quad 0 \leq x \leq y_1 < y_2 \leq x + a;$$

$$P\{\{0\} / x, d_0\} = 1 - G(x-), \quad x \in [0; Q];$$

$$P\{\{0\} / x, d_0\} = p(1 - G(x + a-)) + (1-p)(1 - G(x-)), \quad x \in [0; Q],$$

оскільки $G(Q) < 1$.

Нехай

$$c_1(x, d_a) = \int_0^Q \int_0^\infty c_1(s/x) d\Phi(t/x, d_a, y) P(dy/x, d_a).$$

Тоді для моделі керування запасами умова (2) набуває вигляду

$$r(t/x, d_x) = \begin{cases} 0, & t < \tau(x, d_a), \\ r(x, d_a), & t \geq \tau(x, d_a), \end{cases}$$

де

$$\begin{aligned} r(x, d_0) &= c_1(x, d_0) + \int_x^\infty c_3(y-x) dG(y), \\ r(x, d_a) &= p c_1(x+a, d_a) + (1-p) c_1(x, d_a) + c_2(a) + \\ &+ p \int_{x+a}^\infty c_3(y-(x+a)) dG(y) + (1-p) \int_x^\infty c_3(y-x) dG(y) \end{aligned}$$

для $a > 0$.

Теорема 11 [13]. Нехай функції $c_i(x)$, $i=1,2,3$, є напівнеперервними знизу і виконано припущення 1) теореми 2. Тоді для моделі керування запасами у класі $\mathfrak{R}$ усіх допустимих стратегій існує оптимальна нерандомізована стратегія, для якої досягається мінімум витрат $W = \frac{1}{L} \int V(x) \mu(dx)$. Тут $\mu(\cdot)$ — міра, зосереджена в точці 0 із вагою $\bar{G}(x) := 1 - G(x)$, а $V(x)$ задовольняє рівняння оптимальності

$$\begin{aligned} V(x) &= KV(x) = \\ &= \min \left\{ c_1(x, d_0) + \int_x^\infty c_3(y-x) dG(y) + \int_0^x V(y) d\bar{G}(x-y) + \right. \\ &\quad \left. + V(0) \left[\bar{G}(x) - \frac{1}{L} \bar{G}(Q) \tau(x, d_0) \right] \right\}, \\ &\min_{a \in (0; Q-x]} \left[p c_1(x+a, d_a) + (1-p) c_1(x, d_a) + c_2(a) + \right. \\ &\quad + p \int_{x+a}^\infty c_3(y-(x+a)) dG(y) + (1-p) \int_x^\infty c_3(y-x) dG(y) + \\ &\quad + p \int_0^{x+a} V(y) d\bar{G}(x+a-y) + (1-p) \int_0^x V(y) d\bar{G}(x-y) + \\ &\quad \left. + V(0) \left[p \bar{G}(x+a) + (1-p) \bar{G}(x) - \frac{1}{L} \bar{G}(Q) \tau(x, d_a) \right] \right] \Bigg\}, \end{aligned}$$

$x \in X$.

Теорема 12 [13]. Нехай виконані умови теореми 11 і для будь-якого $x \in X$

$$L(x, a) = p c_1(x+a, d_a) + p \int_{x+a}^\infty c_3(y-(x+a)) dG(y) + c_2(a)$$

монотонно спадає за $a \in (0; Q - x]$. Крім того, нехай функція $\tau(x, d_a)$ монотонно зростає за $x \in X$ і $a \in (0; Q - x]$, а функція $c_1(x, d_x) - c_1(x, d_0)$ монотонно зростає за $x \in X$. Нехай також $\tau(x, d_0) - \tau(x, d_{Q-x})$ монотонно зростає за $x \in [0; Q)$, а

$$c_1(x, d_0) - p c_1(Q, d_{Q-x}) - (1 - p) c_1(x, d_{Q-x}) + p \int_x^{\infty} c_3(y - x) dG(y) - c_2(Q - x)$$

монотонно спадає за $x \in [0; Q)$.

Тоді оптимальна стратегія δ^* для задачі керування запасами має такий вигляд: існує поріг $x^* \in [0; Q]$ такий, що

$$\delta^* = \begin{cases} d_{Q-x}, & x < x^*, \\ d_0, & x \geq x^*. \end{cases}$$

Зауваження 6. Однономенклатурні моделі керування запасами для напівмарковських систем зі спадними функціями загальних витрат детально розглянуто у [13].

3.2.2. Багатономенклатурна модель керування запасами для напівмарковських систем зі спадною функцією витрат. Розглянемо систему керування запасами m продуктів, кожний із яких можна неперервно поповнювати. Припустимо, що Q_i — максимальний рівень запасу i -го продукту, тоді запас i -го продукту набуває значень на $[0, Q_i]$.

У дискретні моменти часу N перевіряють стан запасів кожного продукту і приймають відповідні рішення щодо поповнення складів у такий спосіб.

Якщо рівень запасів i -го продукту в момент часу $n \in N$ дорівнює $X_i^n = x_i \in [0, Q_i]$, то здійснюють замовлення цього продукту $D_i^n \in A_i^x$, $A_i^x := [0, Q_i - x_i]$.

Простір станів системи, яка описує розвиток системи в часі, позначимо $X = X_1 \times X_2 \times \dots \times X_m$, $X_i = [0, Q_i]$, $X = (X^n: n \in N)$, а простір прийнятих рішень — $A = A_1^x \times A_2^x \times \dots \times A_m^x$.

Припустимо, що вартість замовлення (яка може включати витрати виробництва) складається з фіксованої вартості та лінійної функції. Тоді замовлення $x_i > 0$ товарів призводить до очікуваних витрат замовлення $C_i(x_i) = A_i + c_i \cdot x_i$, $C_i(0) = 0$, $A_i \geq 0$, $c_i \geq 0$.

Далі припустимо, що загальні витрати збереження та дефіциту за період часу $f_i(x_i, d^{a_i})$ залежать від рівня запасів $x_i \geq 0$ і замовлення $a_i \in [0, Q_i - x_i]$ на початку періоду, де $f_i(\cdot, \cdot)$ — функції, напівнеперервні знизу за сукупністю змінних.

Розглянемо систему із середніми витратами (φ -критерієм). Для системи, що перебуває у стані x на початку періоду, у разі прийняття рішення $a \in A$ очікувані витрати за один період становлять $r(x, a) = \sum_{i=1}^m r_i(x_i, a_i)$, де $r_i(x_i, a_i)$ — очікувані витрати за i -м продуктом за один період часу, якщо стан цього продукту дорівнює x_i , і на початку періоду прийнято рішення d^{a_i} :

$$r_i(x_i, d^0) = f_i(x_i, d^0), \quad i = \overline{1, m},$$

і для $a_i > 0$

$$r_i(x_i, d^{a_i}) = f_i(x_i + a_i, d^{a_i}) + C_i(a_i), \quad i = \overline{1, m}.$$

Ймовірності переходу на X_i для будь-якої борелівської підмножини $[0, Q_i]$ задають так:

$$\begin{aligned} P([y_i^1, y_i^2] / x_i, d^{a_i}) &= G_i(x_i + a_i - y_i^1) - G_i(x_i + a_i - y_i^2), \\ a_i &\in [0, Q_i - x_i], \quad 0 \leq y_i^1 \leq y_i^2 \leq x_i + a_i, \\ P(\{0\} / x_i, d^{a_i}) &= 1 - G_i(x_i + a_i -), \quad x_i \in [0, Q_i]. \end{aligned}$$

Ймовірності переходу системи мають вигляд $P(B / x, d^a) = \prod_{i=1}^m P(B_i / x_i, d^{a_i})$,

де B_i — борелівські підмножини $[0, Q_i]$. Вважатимемо, що $G(Q) < 1$.

Наступна теорема дає умови існування оптимальної стратегії багатоміноменклатурної моделі керування запасами.

Теорема 13 [14]. Для розглядуваної напівмарковської моделі керування у класі $\mathfrak{R}_1$ існує φ -оптимальна стратегія з мінімальною вартістю $W = \frac{1}{L} \int V(x) \mu(dx)$.

Тут $\mu(\cdot) = \mu_1(\cdot) \dots \mu_m(\cdot)$, $\mu_i(\cdot)$ — міра, сконцентрована в точці 0 із вагою $\overline{G_i} = 1 - G(x)$, $i = \overline{1, m}$, $\overline{G} = \overline{G_1} \dots \overline{G_m}$, а $V(x)$ задовольняє рівняння оптимальності

$$\begin{aligned} V(x) = LV(x) &= \sum_{i=1}^m \min_{a \in A} \{f_i(x_i + a_i, d_{a_i}) + C_i(a_i) \cdot 1_{a_i > 0} + \\ &+ \int_{X_i} V_i(y_i) [P(dy_i / x_i, a_i) - \frac{1}{L} \mu_i(dy_i) \tau_i(x_i, a_i) \prod_{j=1, j \neq i}^m \mu_j(x_j)]\}. \end{aligned}$$

Наступна теорема дає змогу визначити форму оптимальної стратегії для багатоміноменклатурної системи запасів.

Теорема 14 [14]. Нехай $c_i \cdot x_i + f_i(x_i, d^{a_i})$, $i = \overline{1, m}$, монотонно спадає за $x_i \in [0, Q_i]$ і $a_i \in (0, Q_i - x_i]$. Крім того, нехай функція $\tau_i(x_i, d^{a_i})$ монотонно зростає за $x_i \in [0, Q_i]$ і $a_i \in (0, Q_i - x_i]$.

Тоді φ -оптимальна стратегія $\delta^* \in \mathfrak{R}_1$ для задачі керування запасів має такий вигляд: існує поріг $x^* = (x_1^*, \dots, x_m^*) \in [0, Q_1] \times \dots \times [0, Q_m]$ такий, що $\delta^* = (\delta_1^*, \dots, \delta_m^*)$ і для $i = \overline{1, m}$

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

3.2.3. Багатоміноменклатурна модель керування запасами для напівмарковських систем із загальною функцією витрат. Розглянемо систему керування запасами m продуктів, кожний з яких можна неперервно поповнювати (таку саму, як у п. 3.2.2).

Припустимо, що час τ_i^n ($\tau_i^n > 0$) перебування i -ї системи у стані x_i^n є випадковою величиною із функцією розподілу $\Phi_i(\cdot / X_i^n, D_i^n, X_i^{n+1})$, $i = \overline{1, m}$. Будемо вважати, що функції $\Phi_i(\cdot / X_i^n, D_i^n, X_i^{n+1})$ вимірні за Борелем на $[0, Q_i] \times [0, Q_i] \times [0, Q_i]$.

Ці припущення гарантують, що для описаної марковської стратегії послідовності $X_i = (X_i^n : n \in \mathbb{N})$ є стаціонарними напівмарковськими процесами, $i = \overline{1, m}$, і модель керування запасами відповідає структурі напівмарковського динамічного програмування.

Далі вважатимемо, що всі випадкові величини визначені на загальному основному ймовірнісному просторі моделі (Ω, F, P) .

Модель керування запасами системи враховує вартість запасу (яка може включати витрати виробництва), вартість зберігання і дефіцит для кожного i -го продукту. Витрати на зберігання рівня запасу x_i i -го продукту за час s_i , якщо у стані x_i i -та підсистема перебувала протягом часу t_i , $s_i \leq t_i$, становлять $C_i^1(s_i / x_i)$, $C_i^1: [0, Q_i] \rightarrow R_+$, а вартість замовлення продукції в розмірі x_i для i -го товару становить $C_i^2(x_i)$, $C_i^2: [0, Q_i] \rightarrow R_+$, при цьому витрати, спричинені дефіцитом, становлять $C_i^3(x_i)$, $C_i^3: [0, +\infty) \rightarrow [0, +\infty)$, якщо вимоги x_i не можуть бути виконані.

Припустимо, що:

1) функції $C_i^1(s_i / x_i)$ — невід’ємні монотонно неспадні за s_i та x_i , а функції $C_i^2(x_i)$, $C_i^3(x_i)$, $i = \overline{1, m}$, — монотонні неспадні невід’ємні функції;

2) функції $C_i^3(x_i)$, $x_i \in [0, \infty)$, задовольняють умову

$$C_i^3(0) = 0, \quad \int_0^\infty C_i^3(y) dG_i(y) < \infty, \quad i = \overline{1, m}.$$

Якщо i -та підсистема перебуває у стані x_i на початку періоду, прийнято рішення d_{a_i} про дозамовлення в розмірі $a_i \in A_i$ і час, що минув до моменту такого дозамовлення, дорівнює t_i , то очікувані витрати за час s_i , $s_i \leq t_i$, становлять:

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

і для $a_i > 0$

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

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

Ймовірності переходу $P(\cdot / x_i, a_i)$ на X_i для будь-якої борелівської підмножини $[0, Q_i]$ задають у такий спосіб:

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

$$a_i \in [0, Q_i - x_i], \quad 0 \leq y_i^1 < y_i^2 \leq x_i + a_i,$$

$$P(\{0\} / x_i, d^{a_i}) = 1 - G_i(x_i + a_i -), \quad x_i \in [0, Q_i].$$

Ймовірності переходу системи мають вигляд $P(B / x, d^a) = \prod_{i=1}^m P(B_i / x_i, d^{a_i})$, де B_i — борелівські підмножини $[0, Q_i]$.

Нехай

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

Тоді для моделі керування запасами маємо:

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

де

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

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

для $a_i > 0$.

Теорема 15 [8]. Нехай функції $C_i^1, C_i^2, C_i^3, i=\overline{1, m}$, — напівнеперервні знизу і виконано припущення 1) теореми 4. Тоді для моделі керування у класі $\mathfrak{R}$ всіх допустимих стратегій існує φ -оптимальна стратегія з мінімальною вартістю $W = \frac{1}{L} \int V(x) \mu(dx)$.

Тут $\mu(\cdot) = \mu_1(\cdot) \dots \mu_m(\cdot)$, $\mu_i(\cdot)$ — міра, сконцентрована в точці 0 із вагою $\overline{G} = \overline{G_1} \dots \overline{G_m}$, $\overline{G_i} = 1 - G(x)$, $i = \overline{1, m}$, а $V(x)$ задовольняє рівняння оптимальності

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

У [8] розглянуто задачу визначення структури оптимальної стратегії за умов теореми 15.

Теорема 16 [8]. Нехай виконуються умови теореми 15, а також такі умови:

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

монотонно спадають за $a_i \in (0, Q_i - x_i]$;

2) функції $\tau_i(x_i, d^{a_i})$ монотонно зростають за $x_i \in X_i$ та $a_i \in (0; Q_i - x_i]$, а функції $C_i^1(x_i, d^{x_i}) - C_i^1(x_i, d^0)$ монотонно зростають за $x_i \in X_i$;

3) $\tau_i(x_i, d^0) - \tau_i(x_i, d^{Q_i - x_i})$ монотонно зростають за $x_i^* \in [0; Q_i]$, а

$$\tilde{L}_i(x_i) = C_i^1(x_i, d^0) - C_i^1(Q_i, d^{Q_i - x_i}) + \int_{x_i}^{\infty} C_i^3(y - x_i) dG_i(y) - C_i^2(Q_i - x_i),$$

$x_i \in [0, Q_i)$, монотонно спадає за $x_i \in [0, Q_i)$.

Тоді φ -оптимальна стратегія $\delta^* \in \mathfrak{R}_1$ для задачі керування запасами має такий вигляд: існує поріг $x^* = (x_1^*, \dots, x_m^*) \in [0, Q_1] \times \dots \times [0, Q_m]$ такий, що $\delta^* = (\delta_1^*, \dots, \delta_m^*)$ і для $i = \overline{1, m}$

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

Підсумовуючи наведені результати, зазначимо, що основна ідея отримання оптимальних розв'язків для деяких моделей керування запасами полягала у відшуканні умов, за яких задача знаходження оптимальних керувань зводиться до більш простої задачі теорії оптимізації, що має велике значення для конкретних прикладних застосувань. Обговорення деяких проблем, які стосуються реалізації цього підходу до знаходження оптимальних розв'язків, представлено також у роботах [15–19].

СПИСОК ЛІТЕРАТУРИ

1. Висков О.В., Ширяев А.Н. Об управлениях, приходящих к оптимальным стационарным режимам. *Труды МИАН им. Стеклова*. 1964. № 71. С. 35–45.
2. Gubenko L.G., Shtatland E.S. On discrete-time Markov decision processes. *Theory of Probability and Mathematical Statistics*. 1975. N 7. P. 47–61.

3. Губенко Л.Г., Штатланд Э.С. Об управляемых марковских и полумарковских моделях и некоторых прикладных проблемах оптимизации стохастических систем. *Труды Конференции по управляемым системам*. Киев. 1972. С. 87–119.
4. Губенко Л.Г., Штатланд Э.С. Об управляемых полумарковских процессах. *Кибернетика*. 1972. № 2. С. 26–29.
5. Feinberg E.A., Shwartz A. Handbook of Markov decision processes: methods and applications. New York: Springer Verlag, 2002. 566 p.
6. Пепеляева Т.В., Вовк Л.Б., Демченко И.Ю. Об одной задаче многономенклатурной модели теории запасов. *Управляющие системы и машины*. 2015. № 2. С. 32–38.
7. Knopov P.S., Pepelyaeva T.V. Some multidimensional stochastic models of inventory control with a separable cost function. *Cybernetics and Systems Analysis*. 2022. Vol. 58, N 4. P. 523–529. <https://doi.org/10.1007/s10559-022-00487-6>.
8. Пепеляева Т.В., Демченко И.Ю. Об одной многономенклатурной модели для полумарковской системы запасов. *Компьютерная математика*. 2015. № 2. С. 150–162.
9. Daduna H., Knopov P.S., Tur L.P. Optimal strategies for an inventory system with cost functions of general form. *Cybernetics and Systems Analysis*. 1999. Vol. 35, N 4. P. 602–618. <https://doi.org/10.1007/BF02835856>.
10. Demchenko S.S., Knopov A.P., Pepelyaev V.A. Optimal strategies for inventory control systems with a convex cost function. *Cybernetics and Systems Analysis*. 2000. Vol. 36, N 6. P. 891–897. <https://doi.org/10.1023/A:1009413511883>.
11. Pepelyaeva T.V., Vovk L.B., Demchenko I.Yu. Optimal strategies for the multi-task inventory control model. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N 1. P. 107–112. <https://doi.org/10.1007/s10559-016-9805-6>.
12. Knopov P.S., Pepelyaeva T.V. Some multi-task inventory control models for a criterion with revaluation. *Cybernetics and Systems Analysis*. 2023. Vol. 59, N 6. P. 928–933. <https://doi.org/10.1007/s10559-023-00628-5>.
13. Демченко С.С., Кнопов П.С., Чорней Р.К. Оптимальные стратегии для полумарковской системы запасов. *Кибернетика и системный анализ*. 2002. Т. 38, № 1. С. 146–160.
14. Knopov P.S., Pepelyaeva T.V., Demchenko I.Yu. A semi-Markov inventory control model. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N 5. P. 730–736. <https://doi.org/10.1007/s10559-016-9874-6>.
15. Кнопов А.П., Пепеляев В.А. Про оптимальні стратегії для систем керування запасами з опуклими функціями вартості. *Компьютерная математика*. 2005. № 3. С. 116–121.
16. Knopov A.P., Pepelyaev V.A. Some continuous models of inventory control. *Cybernetics and Systems Analysis*. 2005. Vol. 41, N 3. P. 465–467. <https://doi.org/10.1007/s10559-005-0080-1>.
17. Knopov P.S., Pepelyaev V.A. On one model of controlled semi-Markov process. *Tavrisheskij Vestn. Inform. Mat.* 2005. N 2. P. 5–13.
18. Knopov P., Norkin V. Stochastic optimization methods for the stochastic storage process control. In: Intelligent Control and Smart Energy Management. Blondin M.J., Fernandes Trovão J.P., Chaoui H., Pardalos P.M. (Eds.). *Springer Optimization and Its Applications*. 2022. Vol. 181. P. 79–111. https://doi.org/10.1007/978-3-030-84474-5_3.
19. Дериева Е.Н., Пепеляева Т.В. Об одной задаче управления случайными процессами. *Кибернетика и системный анализ*. 2004. Т. 40, № 1. С. 116–121.

P.S. Knopov, T.V. Pepelyaeva
CONTROLLED STOCHASTIC SYSTEMS

Abstract. Controlled Markov and semi-Markov processes and systems are considered. A review of inventory control problems is provided. One- and multiproduct models with different types of cost function and optimality criteria are considered. The optimality conditions and structures of optimal strategies in the problems are investigated.

Keywords: Markov processes, semi-Markov processes, inventory control, (s, S) -strategy, optimality criterion, optimal strategy.

Надійшла до редакції 28.03.2024