ISSN 2078-3744. Вісник Львів. ун-ту. Серія мех.-мат. 2026. Випуск 98. С. 101–107
Visnyk of the Lviv Univ. Series Mech. Math. 2026. Issue 98. P. 101–107
http://publications.lnu.edu.ua/bulletins/index.php/mmf
doi: http://dx.doi.org/10.30970/vmm.2026.98.101-107

УДК 519.217

КЕРОВАНІ МАРКІВСЬКІ ПРОЦЕСИ З ОБРИВОМ У ФІНАНСОВИХ ЗАДАЧАХ

Ярослав ЄЛЕЙКО, Антон НОС

Львівський національний університет імені Івана Франка,
вул. Університетська,1 м. Львів, 79000
e-mail: yaroslav.yeleyko@lnu.edu.ua, ORCID: 0000-0002-2200-8447
e-mail: anton.nos@lnu.edu.ua, ORCID: 0009-0003-5233-1805

У роботі розглянуто дискретну модель керованого марківського процесу з обривом на скінченному горизонті часу. Обрив інтерпретується як банкрутство фінансової системи, а керування впливає як на поточний дохід, так і на ймовірність переходу в обривний стан. Для моделі введено функціонал якості, побудовано оператори динамічного програмування, виведено рівняння Беллмана, доведено існування оптимальної марківської стратегії та наведено умову побудови рівномірно ε-оптимальних стратегій. Наприкінці подано фінансовий приклад для моделі капіталу з двома режимами ризику.

Ключові слова: керований марківський процес, обривний стан, функціонал якості, рівняння Беллмана, оптимальна стратегія, банкрутство.

Рівняння Беллмана для скінченногоризонтних керованих марківських процесів та існування оптимальної марківської стратегії є класичними результатами теорії динамічного програмування і марківських процесів керування ([1, 2, 3, 4]). У цій роботі ці результати адаптовано до моделі з обривним станом та фінансовою інтерпретацією.

Нехай на проміжку дискретного часу $t = 0, 1, \dots, N$ задано випадковий процес станів

$$X_t \in \mathsf{X}_t, \qquad t = 0, 1, \dots, N,$$

де X_t — множина допустимих фінансових станів системи в момент t. До кожного моменту $t = 1, \dots, N$ відповідає множина керувань

$$A_t \in \mathsf{A}_t,$$

2020 Mathematics Subject Classification: 60J25

причому після вибору керування $A_t = a$ система переходить зі стану $x \in \mathsf{X}_{t-1}$ у випадковий стан $y \in \mathsf{X}_t$. Введемо спеціальний обривний стан $x_t^* \in \mathsf{X}_t$, який інтерпретується як банкрутство, і нехай

$$\mathsf{X}_t = \widehat{\mathsf{X}}_t \cup \{x_t^*\},$$

де $\widehat{\mathsf{X}}_t$ — множина не обривних станів. Припускаємо, що якщо процес потрапив у x_t^*, то подальше функціонування системи втрачає економічний зміст. Аналогічно до позначень у роботах з обривних керованих процесів, модель запишемо у вигляді

$$Z_\mu^* = (\mathsf{X}, \mathsf{A}, j, p, q, r, c, \mu),$$

де: $X = (X_0, \ldots, X_N)$ простір станів; $\mathsf{A} = (\mathsf{A}_1, \ldots, \mathsf{A}_N)$ — простір керувань; $j_t : \mathsf{A}_t \to \widehat{\mathsf{X}}_{t-1}$ — відображення допустимості, тобто $j_t(a) = x$ означає, що керування a допустиме в стані x; $p_t(\cdot \mid a)$ — розподіл переходу на X_t при керуванні $a \in \mathsf{A}_t$; $q_t : \mathsf{A}_t \to \mathbb{R}$ — миттєва винагорода; $r : \widehat{\mathsf{X}}_N \to \mathbb{R}$ — термінальний виграш у фінальний момент; $c_t(x_t^*) < 0$ — штраф у випадку обриву; μ — початковий розподіл на X_0.

Означення 1. Шляхом процесу називається послідовність

$$l = (x_0, a_1, x_1, \ldots, a_t, x_t, \ldots, a_k, x_k),$$

де $x_s \in \widehat{\mathsf{X}}_s$ для $s < k$, а $x_k \in \mathsf{X}_k$; якщо $x_k = x_k^*$, то шлях обривається в момент k.

Нехай $\pi = (\pi_1, \ldots, \pi_N)$ — стратегія керування, де для кожного t і кожної історії

$$h_{t-1} = (x_0, a_1, x_1, \ldots, a_{t-1}, x_{t-1})$$

функція $\pi_t(\cdot \mid h_{t-1})$ є розподілом імовірностей на множині допустимих керувань

$$\mathsf{A}_t(x_{t-1}) = \{a \in \mathsf{A}_t \colon j_t(a) = x_{t-1}\}.$$

Тоді імовірність довільного не обривного шляху

$$l = (x_0, a_1, x_1, \ldots, a_N, x_N), \qquad x_t \in \widehat{\mathsf{X}}_t,$$

визначається за формулою

$$P_\mu^\pi(l) = \mu(x_0) \prod_{t=1}^{N} \pi_t(a_t \mid h_{t-1}) p_t(x_t \mid a_t). \tag{1}$$

Якщо ж обрив відбувся в момент k, тобто $x_k = x_k^*$, то

$$P_\mu^\pi(l) = \mu(x_0) \prod_{t=1}^{k} \pi_t(a_t \mid h_{t-1}) p_t(x_t \mid a_t). \tag{2}$$

Визначимо оцінку шляху l. Якщо до моменту N обрив не відбувся, то покладемо

$$I(l) = \sum_{t=1}^{N} q_t(a_t) + r(x_N). \tag{3}$$

Якщо ж обрив відбувся в момент $k \le N$, тобто $x_k = x_k^*$, то

$$I(l) = \sum_{t=1}^{k} q_t(a_t) + c_k(x_k^*). \tag{4}$$

Отже, математичне сподівання функціонала якості дорівнює

$$\omega(\mu, \pi) = E_\mu^\pi I(l) = \sum_l I(l) P_\mu^\pi(l). \tag{5}$$

Для виродженого початкового розподілу в точці $x \in \mathsf{X}_0$ використовуватимемо скорочення

$$\omega(x, \pi) = \omega(\delta_x, \pi).$$

Оптимальне значення моделі визначимо як

$$\upsilon(\mu) = \sup_\pi \omega(\mu, \pi), \qquad \upsilon_t(x) = \sup_\pi \omega_t(x, \pi). \tag{6}$$

Для підмоделі з початковим станом $x \in \widehat{\mathsf{X}}_{t-1}$ і керуванням $a \in \mathsf{A}_t(x)$ маємо рекурентне співвідношення

$$\omega_{t-1}(x, \pi) = \sum_{a \in \mathsf{A}_t(x)} \pi_t(a \mid x) \left[q_t(a) + \sum_{y \in \mathsf{X}_t} p_t(y \mid a) \omega_t(y, \pi^a) \right], \tag{7}$$

де π^a — хвіст стратегії після вибору керування a. Оскільки

$$\omega_t(x_t^*, \pi) = c_t(x_t^*),$$

то з (7) випливає

$$\omega_{t-1}(x, \pi) = \sum_{a \in \mathsf{A}_t(x)} \pi_t(a \mid x) \left[q_t(a) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid a) \omega_t(y, \pi^a) + p_t(x_t^* \mid a) c_t(x_t^*) \right]. \tag{8}$$

Введемо оператори

$$(U_t f)(a) = q_t(a) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid a) f(y) + p_t(x_t^* \mid a) c_t(x_t^*), \qquad a \in \mathsf{A}_t, \tag{9}$$

та

$$(V_t g)(x) = \sup_{a \in \mathsf{A}_t(x)} g(a), \qquad x \in \widehat{\mathsf{X}}_{t-1}. \tag{10}$$

Тоді композитний оператор

$$T_t = V_t U_t \tag{11}$$

породжує рівняння Беллмана

$$\upsilon_{t-1} = T_t \upsilon_t, \qquad t = 1, 2, \ldots, N, \tag{12}$$

з кінцевими умовами

$$\upsilon_N(x) = r(x), \qquad x \in \widehat{\mathsf{X}}_N, \tag{13}$$

та

$$\upsilon_t(x_t^*) = c_t(x_t^*), \qquad t = 1, \ldots, N. \tag{14}$$

Класичне рівняння Беллмана для задачі керування на скінченному горизонті, адаптоване до наявності обривного стану, має такий вигляд ([1, 2, 3]).

Теорема 1. *Для кожного $t = 1, \ldots, N$ та кожного $x \in \widehat{\mathsf{X}}_{t-1}$ справджується рівність*

$$\upsilon_{t-1}(x) = \sup_{a \in \mathsf{A}_t(x)} \left[q_t(a) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid a)\upsilon_t(y) + p_t(x_t^* \mid a)c_t(x_t^*) \right]. \tag{15}$$

Доведення. З формули (8) для довільної стратегії π маємо

$$\omega_{t-1}(x, \pi) \leq \sup_{a \in \mathsf{A}_t(x)} \left[q_t(a) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid a)\upsilon_t(y) + p_t(x_t^* \mid a)c_t(x_t^*) \right].$$

Переходячи до супремуму за всіма стратегіями, дістаємо нерівність "$\leq$". Обернена нерівність отримується фіксацією керування $a \in \mathsf{A}_t(x)$ та подальшим використанням у підмоделі $[t, N]$ оптимальної стратегії; після переходу до границі по $\varepsilon \downarrow 0$ одержуємо рівність. $\square$

Теорема 2. *Для довільного початкового розподілу μ виконується*

$$\upsilon(\mu) = \sum_{x \in \mathsf{X}_0} \mu(x)\upsilon_0(x). \tag{16}$$

Доведення. З лінійності математичного сподівання маємо

$$\omega(\mu, \pi) = \sum_{x \in \mathsf{X}_0} \mu(x)\omega(x, \pi) \leq \sum_{x \in \mathsf{X}_0} \mu(x)\upsilon_0(x).$$

Далі використовуємо комбінування ε-оптимальних стратегій для кожного початкового стану x і переходимо до границі $\varepsilon \downarrow 0$. $\square$

Нехай для кожного t і $x \in \widehat{\mathsf{X}}_{t-1}$ множина $\mathsf{A}_t(x)$ скінченна. Тоді супремум у (15) досягається. Позначимо через $\psi_t(x)$ селектор, який задовольняє

$$\psi_t(x) \in \operatorname*{argmax}_{a \in \mathsf{A}_t(x)} \left[q_t(a) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid a)\upsilon_t(y) + p_t(x_t^* \mid a)c_t(x_t^*) \right]. \tag{17}$$

Наступне твердження про існування оптимальної марківської стратегії є класичним для скінченногоризонтних моделей керування; тут його записано в позначеннях моделі Z_μ^* ([2, 3, 4]).

Теорема 3. *Стратегія $\psi = (\psi_1, \ldots, \psi_N)$, побудована за правилом* (17)*, є оптимальною марківською стратегією, тобто*

$$\omega(x, \psi) = \upsilon_0(x), \qquad x \in \mathsf{X}_0, \tag{18}$$

і

$$\omega(\mu, \psi) = \upsilon(\mu) \tag{19}$$

для довільного початкового розподілу μ.

Доведення. Проведемо зворотну індукцію за t. Для $t = N$ з (15) та (13) випливає оптимальність ψ_N. Припустимо, що стратегія $(\psi_t, \dots, \psi_N)$ оптимальна в підмоделі $[t, N]$. Тоді для $x \in \widehat{\mathsf{X}}_{t-1}$, використовуючи (8), маємо

$$\omega_{t-1}(x, \psi_t, \dots, \psi_N) = q_t(\psi_t(x)) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid \psi_t(x))\upsilon_t(y) + p_t(x_t^* \mid \psi_t(x))c_t(x_t^*) = \\ = \upsilon_{t-1}(x),$$

оскільки $\psi_t(x)$ максимізує праву частину (15). Індукцією дістаємо (18); формула (19) випливає з (16). □

Стратегія π називається ε-оптимальною для моделі Z_μ^*, якщо

$$\omega(\mu, \pi) \geq \upsilon(\mu) - \varepsilon.$$

Вона називається рівномірно ε-оптимальною, якщо ця нерівність виконується для всіх початкових розподілів μ. Нехай для кожного t вибрано селектор $\psi_t^{(\varepsilon)}$ такий, що

$$q_t(\psi_t^{(\varepsilon)}(x)) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t(y \mid \psi_t^{(\varepsilon)}(x))\upsilon_t(y) + p_t(x_t^* \mid \psi_t^{(\varepsilon)}(x))c_t(x_t^*) \geq \upsilon_{t-1}(x) - \frac{\varepsilon}{N} \tag{20}$$

для всіх $x \in \widehat{\mathsf{X}}_{t-1}$.

Твердження 1. *Стратегія $\psi^{(\varepsilon)} = (\psi_1^{(\varepsilon)}, \dots, \psi_N^{(\varepsilon)})$ є рівномірно ε-оптимальною.*

Доведення. Застосуємо зворотну індукцію. Для останнього кроку $t = N$ з (20) випливає

$$\omega_{N-1}(x, \psi_N^{(\varepsilon)}) \geq \upsilon_{N-1}(x) - \frac{\varepsilon}{N}.$$

Припустимо, що в підмоделі $[t, N]$ вже маємо оцінку

$$\omega_t(x, \psi_t^{(\varepsilon)}, \dots, \psi_N^{(\varepsilon)}) \geq \upsilon_t(x) - \frac{N-t}{N}\varepsilon.$$

Тоді

$$\omega_{t-1}\big(x, \psi_{t-1}^{(\varepsilon)}, \dots, \psi_N^{(\varepsilon)}\big) \geq q_t\big(\psi_t^{(\varepsilon)}(x)\big) + \sum_{y \in \widehat{\mathsf{X}}_t} p_t\big(y \mid \psi_t^{(\varepsilon)}(x)\big)\left(\upsilon_t(y) - \frac{N-t}{N}\varepsilon\right) + \\ + p_t\big(x_t^* \mid \psi_t^{(\varepsilon)}(x)\big)c_t(x_t^*) \geq \upsilon_{t-1}(x) - \frac{N-t+1}{N}\varepsilon.$$

Поклавши $t = 1$, дістаємо $\omega_0(x, \psi^{(\varepsilon)}) \geq \upsilon_0(x) - \varepsilon$, а далі за формулою (16) — рівномірну ε-оптимальність. □

Розглянемо капітал інвестора на горизонті N періодів. Нехай

$$\widehat{\mathsf{X}}_t = \{1, 2, \dots, M\}, \qquad x_t^* = 0,$$

де значення x означає рівень капіталу, а стан 0 — банкрутство. На кожному кроці доступні два керування

$$\mathsf{A}_t(x) = \{a_L, a_H\},$$

де a_L — консервативна стратегія, a_H — агресивна стратегія. Нехай перехідні ймовірності задаються формулами

$$p_t(x+1 \mid a_L) = \alpha_L, \qquad p_t(x \mid a_L) = 1 - \alpha_L - \beta_L, \qquad p_t(0 \mid a_L) = \beta_L, \tag{21}$$

$$p_t(x+2 \mid a_H) = \alpha_H, \quad p_t(x-1 \mid a_H) = 1 - \alpha_H - \beta_H, \quad p_t(0 \mid a_H) = \beta_H, \tag{22}$$

де $0 < \beta_L < \beta_H < 1$, $0 < \alpha_L < \alpha_H < 1$, а вихід за межі множини $\{1, \ldots, M\}$ замінюється найближчим допустимим станом. Миттєві виграші покладемо рівними

$$q_t(a_L) = u_L, \qquad q_t(a_H) = u_H, \qquad u_H > u_L > 0, \tag{23}$$

а штраф за банкрутство візьмемо у вигляді

$$c_t(0) = -K, \qquad K > 0. \tag{24}$$

Кінцевий виграш нехай дорівнює поточному капіталу

$$r(x) = \gamma x, \qquad \gamma > 0. \tag{25}$$

Тоді рівняння Беллмана набуває вигляду

$$\begin{aligned} \upsilon_{t-1}(x) = \max\Big\{ & u_L + \alpha_L \upsilon_t(x+1) + (1 - \alpha_L - \beta_L)\upsilon_t(x) + \beta_L c_t(0), \\ & u_H + \alpha_H \upsilon_t(x+2) + (1 - \alpha_H - \beta_H)\upsilon_t(x-1) + \beta_H c_t(0)\Big\}. \end{aligned} \tag{26}$$

Позначимо різницю між агресивною та консервативною стратегіями через

$$\begin{gathered} \Delta_t(x) = (u_H - u_L) + \alpha_H \upsilon_t(x+2) - \alpha_L \upsilon_t(x+1) + \\ +(1 - \alpha_H - \beta_H)\upsilon_t(x-1) - (1 - \alpha_L - \beta_L)\upsilon_t(x) + (\beta_H - \beta_L)c_t(0). \end{gathered} \tag{27}$$

Тоді оптимальний вибір визначається правилом

$$\psi_t(x) = \begin{cases} a_H, & \text{якщо } \Delta_t(x) \geq 0; \\ a_L, & \text{якщо } \Delta_t(x) < 0. \end{cases} \tag{28}$$

Оскільки $c_t(0) = -K < 0$, доданок $(\beta_H - \beta_L)c_t(0)$ є від'ємним і зменшує привабливість агресивного керування. Отже для малих рівнів капіталу, коли ризик банкрутства суттєвий, оптимальною зазвичай є консервативна стратегія, а для великих значень x можлива перевага агресивної стратегії. Таким чином у моделі виникає порогова структура оптимального керування. Побудовано математичну модель керованого марківського процесу з обривом для фінансових задач на скінченному горизонті. Для введеного функціонала якості одержано фундаментальне рекурентне співвідношення, побудовано оператори U_t, V_t, T_t та виведено рівняння Беллмана. Доведено існування оптимальної марківської стратегії та наведено спосіб побудови рівномірно ε-оптимальних стратегій. Фінансовий приклад показує, що врахування банкрутства як обривного стану приводить до суттєво іншої структури оптимального керування порівняно з класичними моделями без поглинання.

Список використаної літератури

1. Bellman R. Dynamic programming, Princeton University Press, 1957.
2. Bertsekas D.P., Shreve S.E. Stochastic Optimal control: the discrete-time case, Academic Press, 1978.

3. Puterman M.L. Markov decision processes: discrete stochastic dynamic programming, John Wiley & Sons, 1994.
4. Hernández-Lerma O., Lasserre J.B. Discrete-time Markov control processes: basic optimality criteria, Springer, 1996.

Стаття: вперше надійшла 18.11.2025
прийнята до друку 15.04.2026
опублікована 29.04.2026

CONTROLLED MARKOV PROCESSES WITH TERMINATION IN FINANCIAL PROBLEMS

Yaroslav YELEYKO, Anton NOS

Ivan Franko Lviv National University,
Universytetska str., 1, L'viv, 79000
e-mail: yaroslav.yeleyko@lnu.edu.ua, ORCID: 0000-0002-2200-8447
e-mail: anton.nos@lnu.edu.ua, ORCID: 0009-0003-5233-1805

This article considers a discrete-time controlled Markov process with termination over a finite time horizon. The termination is interpreted as the bankruptcy of a financial system, while control actions affect both the current payoff and the probability of transition to the termination state. For the model, a performance functional is introduced, dynamic programming operators are constructed, and the Bellman equation is derived. The existence of an optimal Markov strategy is proven, and a condition for constructing uniformly ε-optimal strategies is established. Finally, a financial example is presented for a capital model with two risk regimes.

Key words: Controlled Markov process, termination state, performance functional, Bellman equation, optimal strategy, bankruptcy.