References

1. Приемы объектно-ориентированного проектирования. Паттерны проектирования / Э. Гамма, Р. Хелм, Д. Джонсон, Дж. Влиссидес ; пер. с англ. – СПб : Питер, 2006. – 366 с.
2. Дидманидзе И. Ш. Использование тренажеров в учебном процессе. Материалы II Международной научно-практической интернет конференций / И. Ш. Дидманидзе, Н. О. Худжадзе, Д. З. Дидманидзе // Материалы II Международной научно-практической интернет-конференций «Сучасні тенденції розвитку математики та її прикладні аспекти». – Донецк : ДонНУЕТ, 2013. – С. 205–206.
3. Софиев А. Э. Компьютерные обучающие системы / А. Э. Софиев, Е. А. Черткова – М. : ДеЛи принт, 2006. — 296 с.

Дідманідзе І. Ш., Худжадзе Н. О., Дідманідзе Д. З.

ТРЕНАЖЕРИ В УЧБОВОМУ ПРОЦЕСІ

У статті описано віртуальні тренажери в учбовому процесі. Розглянуто їх використання у віртуальній лабораторії.

Ключові слова: тренажер, учбові системи, візуальний тренажер, віртуальний тренажер, віртуальна лабораторія.

Матеріал надійшов 31.07.2013

УДК 519.21

Ковальчук Н. Д., Чорней Р. К.

НАПІВМАРКОВСЬКІ ПРОЦЕСИ ПРИЙНЯТТЯ РІШЕНЬ З РАНДОМІЗОВАНИМ ДИСКОНТОМ

У роботі розглянуто керовані напівмарковські процеси з випадковим дисконтуючим фактором. Знаходяться достатні умови існування та єдиності оптимальної стаціонарної нерандомізованої стратегії керування зазначеними процесами на скінченному та нескінченному горизонті.

Ключові слова: процеси прийняття рішень, напівмарковські процеси, рандомізований дисконт, рівняння оптимальності.

Вступ

Чимало наукових робіт присвячено вивченню випадкових керованих процесів з дисконтованим критерієм. Вибір критерію пояснюють, в основному, його можливим застосуванням в економічних задачах. Дисконт розуміють як величину, обернену до відсоткової ставки (рівня інфляції тощо). Однак навіть неглибокий аналіз показує, що ці економічні показники не є сталими в часі (рис. 1–3).

Питання зміни дисконту в часі розглядалося в роботах [1; 4; 6]. У даній роботі досліджуються напівмарковські процеси прийняття рішень з випадковим дисконтуючим множником. Досі такі процеси не розглядалися. Автори узагальнюють результати, отримані в [3] для марковських процесів прийняття рішень з рандомізованим дисконтом.

Постановка задачі

Розглянемо напівмарковський процес прийняття рішень для випадку, коли дисконт-фактор $\beta=(1+r)^{-1}$ є випадковою величиною (r – відсоткова ставка). Даний випадок є узагальненням керованого марковського процесу з рандомізованим дисконтом [3]. Подамо дисконт-фактор у вигляді $e^{-\alpha}=(1+r)^{-1}$, звідки $\alpha=\ln(1+r)$.

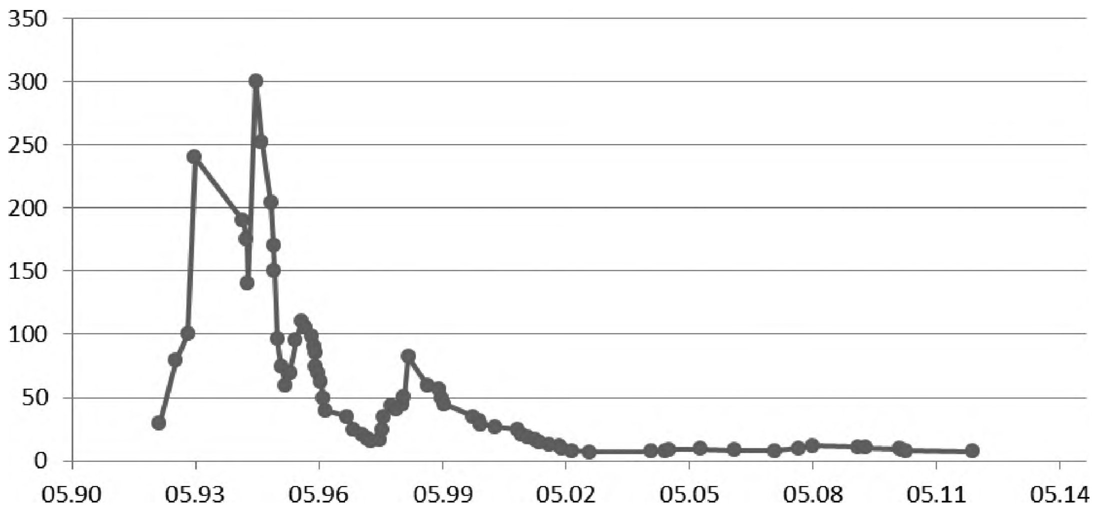

Рис. 1. Облікова ставка Національного банку України за 1991–2012 роки

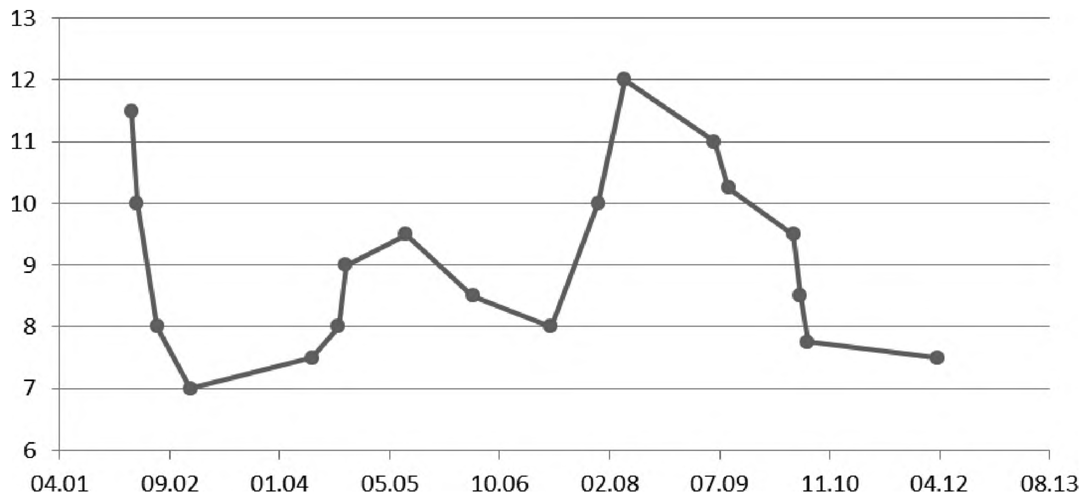

Рис. 2. Облікова ставка Національного банку України за 2002–2012 роки

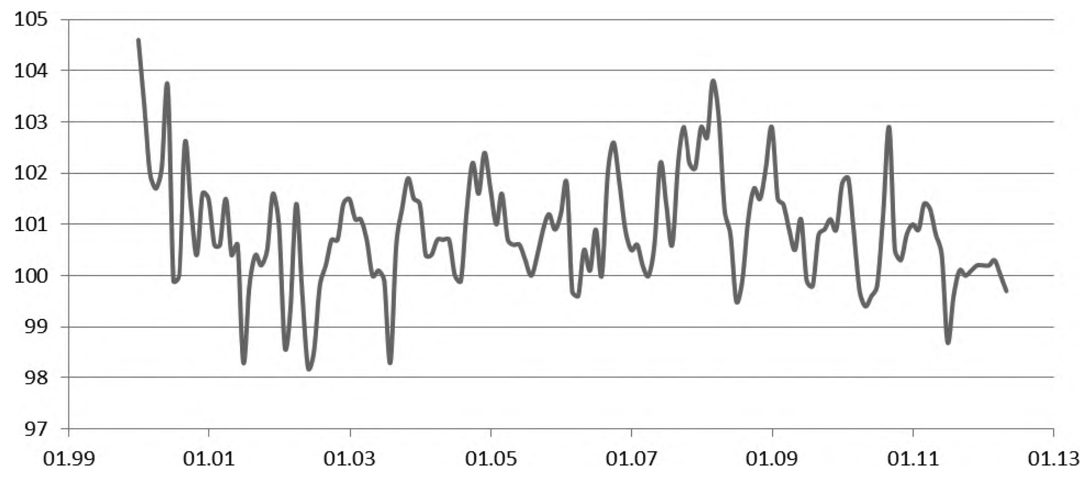

Рис. 3. Індекс інфляції в Україні за 1991–2012 роки

Оскільки $r \in (0;+\infty)$, то і $\alpha \in (0;+\infty)$. Надалі простором станів вважатимемо X', де X' – простір станів випадкового процесу, а R^+ – множина значень випадкової величини α.

Означення 1. Напівмарковським процесом прийняття рішень з рандомізованим дисконтом (або керованим процесом) називатимемо сукупність об'єктів: $\mathbf{P} = \langle (X,\Xi),(A,\mathrm{A}),F,\Delta,\mathrm{Q},\Phi,c\rangle$,

де:

- $X = X' \times R^+$ – простір станів системи, Ξ – σ-алгебра борелівських підмножин простору X;
- A – простір керуючих впливів, A – σ-алгебра борелівських підмножин простору F;
- F – вимірне відображення, що ставить у відповідність кожному $(x,\alpha) \in X$ деяку непорожню замкнену підмножину $A_{(x,\alpha)}$ множини A, $A_{(x,\alpha)}$ – множина допустимих керуючих впливів системи в стані (x,α);
- $\Delta = \left\{ \left(x,\alpha,a \,\middle|\, (x,\alpha) \in X, a \in A_{(x,\alpha)}\right) \right\}$ – множина можливих наборів станів і керуючих впливів, вимірна підмножина $X \times A$;
- $\mathrm{Q}(\cdot \mid x,\alpha,a)$ – стохастичне ядро, що задане на X, породжене Δ, називається законом переходу;
- $\Phi(\cdot \mid x',\alpha',a,x'',\alpha'')$ – функція розподілу додатної випадкової величини τ, яка є часом перебування системи в стані (x',α') за умови, що прийнято рішення $a \in A_{(x,\alpha)}$ і наступним станом системи є $(x'',\alpha'') \in X$;
- $c(x,\alpha,a)$ – вимірна дійснозначна функція на Δ, що позначає очікувану вартість одиниці часу перебування системи в стані $(x,\alpha) \in X$ за умови прийняття рішення $a \in A_{(x,\alpha)}$.

Означення 2. Для кожного моменту $n \in \mathbf{Z}^+$ переходу керованого процесу $\mathbf{P}$ визначимо множину H_n допустимих історій до n-го моменту переходу включно таким чином:

$H_0 = X$ (множина початкових станів);

$$H_n = \left(\prod_{i=0}^{n-1} \left(\Delta \times \mathbf{R}^+\right) \right) \times X, \; n \in \mathbf{N}.$$

Елементи множини H_n, $n \in Z^+$, називатимемо допустимими n-історіями керованого процесу, що мають вигляд:

$$h_n = \left(x_0,\alpha_0,a_0,t_0,\ldots,x_{n-1},\alpha_{n-1},a_{n-1},t_{n-1},x_n,\alpha_n\right),$$

де $(x_i,\alpha_i,a_i) \in \Delta$, $i = \overline{0,n}$, t_i – реалізація випадкової величини τ, t_i – час перебування системи в стані (x_i,α_i).

При цьому для кожного $n \in \mathbf{N}$ має місце рівність

$$H_n \subseteq \bar{H}_n = \left(\prod_{i=0}^{n-1} \left(X \times A \times \mathbf{R}^+\right) \right) \times X, \; H_0 \subseteq \bar{H}_0 = X.$$

Означення 3. Допустиму рандомізовану стратегію π визначимо як послідовність $\{\pi_n, n \in \mathbf{Z}^+\}$, де $\pi_n(\cdot \mid h_n)$ – ймовірнісна міра на (A,A), зосереджена на $A_{(x_n,\alpha_n)}$, тобто:

$$\pi_n\left(A_{(x_n,\alpha_n)} \mid h_n\right) = 1 \text{ для всіх } h_n \in \mathbf{Z}^+, \; n \in \mathbf{Z}^+$$

і вимірно залежить від n-історії керованого процесу $\mathbf{P}$.

Стратегія π називається *марковською*, якщо

$$\pi_n(\cdot \mid h_n) = \pi_n(\cdot \mid x_n,\alpha_n) \text{ для всіх } n \in \mathbf{Z}^+.$$

Марковська стратегія π називатиметься *стаціонарною,* якщо

$$\pi_n(\cdot \mid x_n,\alpha_n) = \pi(\cdot \mid x_n,\alpha_n) \text{ для всіх } n \in \mathbf{Z}^+.$$

та *стаціонарною нерандомізованою*, якщо міра $\pi(\cdot \mid x,\alpha)$ вироджена для всіх $(x,\alpha) \in X$, тобто стаціонарну нерандомізовану стратегію π можна ототожнити з деякою вимірною функцією $f: X \to A$ такою, що $f(x,\alpha) \in A_{(x,\alpha)}$ для всіх $(x,\alpha) \in X$. Дана функція називатиметься *селектором* багатозначної функції $(x,\alpha) \mapsto A_{(x,\alpha)}$. Множину всіх селекторів позначатимемо F.

Клас всіх допустимих рандомізованих стратегій позначимо S_R, марковських – S_M, стаціонарних – S_S, стаціонарних нерандомізованих – S_N. Для даних класів має місце включення: $S_N \subset S_S \subset S_M \subset S_R$.

Означення 4. Як критерій керування розглянемо такий функціонал:

$$\Re(\pi,x,\alpha) := E^{\pi}_{(x,\alpha)} \left[\sum_{n=0}^{\infty} e^{-s_n} c(x_n,\alpha_n,a_n) \right], \; \pi \in S_R,$$

$$(x,\alpha) \in X_0,$$

де $E^{\pi}_{(x,\alpha)}$ – математичне сподівання, що відповідає процесу $\mathbf{P}$, керованому стратегією π, з початковим станом $(x,\alpha) \in X$;

$$s_0 = 0, \; s_n = \sum_{i=0}^{n-1} \alpha_i \tau_i, \; n \in \mathbf{N}.$$

Стратегію $\pi^* \in S_R$ називатимемо оптимальною, якщо

$$\Re(\pi^*,x,\alpha) = \inf_{\pi \in S_R} \Re(\pi,x,\alpha) =: \Re^*(x,\alpha)$$

$$\text{для всіх } (x,\alpha) \in X_0,$$

а функцію $\Re^*$ – функцією вартості.

Припустимо, що функція

$$\tau(x,\alpha,a)=\int_X\int_0^\infty t d\Phi(t\mid x,\alpha,a,y,\beta)Q(d(y,\beta)\mid x,\alpha,a),$$

що позначає очікувану тривалість часу перебування в стані (x,α) за умови прийняття рішення a, існує та скінченна для всіх $(x,\alpha,a)\in\Delta$. Тоді, внаслідок того, що критерій оптимальності залежить лише від стохастичного ядра $Q(\cdot\mid x,\alpha,a)$, очікуваної ціни $c(x,\alpha,a)$ та усередненої характеристики $\tau(x,\alpha,a)$, не втрачаючи загальності, розглядатимемо тільки такі керовані процеси, для яких функція розподілу випадкової величини τ має вигляд:

$$\Phi(t\mid x,\alpha,a,y,\beta)=\begin{cases}0, & t\le\tau(x,\alpha,a);\\ 1, & t>\tau(x,\alpha,a).\end{cases}$$

Керування на скінченному горизонті

Розглянемо керовану напівмарковську модель

$$\mathbf{P}=\langle(X,\Xi),(A,\mathrm{A}),F,\Delta,\mathrm{Q},\Phi,c\rangle$$

на скінченному горизонті, тобто вважатимемо, що процес відбувається протягом N кроків:

$$e^{-S_m}c(x_m,\alpha_m,a_m)=0 \text{ для всіх } m>N.$$

Крім того, на останньому кроці рішення не приймається, тобто вартість перебування на ньому залежатиме лише від стану: $c=c(x_N,\alpha_N)$.

Отже, потрібно мінімізувати критерій

$$\Re(\pi,x,\alpha):=E^{\pi}_{(x,\alpha)}\left[\sum_{n=0}^{N-1}e^{-S_n}c(x_n,\alpha_n,a_n)+e^{-S_N}c(x_N,\alpha_N)\right],$$

$$\pi\in S_R,\ (x,\alpha)\in X_0, \tag{1}$$

$$\Re^*(x,\alpha)=\inf_{\pi\in S_R}\Re(\pi,x,\alpha) \text{ для всіх } (x,\alpha)\in X_0. \tag{2}$$

Необхідно знайти стратегію $\pi^*\in S_R$ таку, що

$$\Re(\pi^*,x,\alpha)=\Re^*(x,\alpha). \tag{3}$$

Основним результатом цього розділу є наступна теорема динамічного програмування, що визначає цінову функцію та оптимальну стратегію.

Теорема 5. Для $n=\overline{0;N}$ визначимо функцію V_n, задану на X, починаючи з останнього кроку $n=N$:

$$V_N(x,\alpha):=c(x,\alpha); \tag{4}$$

а для $n=\overline{N-1;0}$

$$V_n(x,\alpha):=\min_{a\in A_{(x,\alpha)}}\left\{e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,a)+\right.\right.$$

$$\left.\left.+\int_X V_{n+1}(y,\beta)Q(d(y,\beta)|x,\alpha,a)\right]\right\}. \tag{5}$$

Припустимо, що ця функція вимірна і для кожного $n=\overline{0;N-1}$ існує селектор $f_n\in F: f_n(x,\alpha)\in A_{(x,\alpha)}$, який досягає мінімуму в (5) для всіх $(x,\alpha)\in X$; тобто для всіх $(x,\alpha)\in X$ і $n=\overline{0;N-1}$

$$V_n(x,\alpha)=e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,f_n)+\right.$$

$$\left.+\int_X V_{n+1}(y,\beta)Q(d(y,\beta)|x,\alpha,f_n)\right]. \tag{6}$$

Тоді стратегія $\pi^*=\{f_0,\ldots,f_{N-1}\}$ – оптимальна і $\Re^*=V_0$, тобто

$$\Re^*(x,\alpha)=V_0(x,\alpha)=\Re(\pi^*,x,\alpha)$$

$$\text{для всіх } (x,\alpha)\in X. \tag{7}$$

Доведення. Нехай $\pi=\{\pi_n\}$ – довільна стратегія, а $C_n(\pi,x,\alpha)$ відповідна очікувана загальна вартість за період часу від кроку n до останнього кроку N, починаючи зі стану $(x_n,\alpha_n)=(x,\alpha)$ на кроці n, тобто

$$C_n(\pi,x,\alpha):=E^{\pi}\left[\sum_{k=n}^{N-1}e^{S_n-S_k}c(x_n,\alpha_n,a_n)+\right.$$

$$\left.+e^{S_n-S_N}c(x_N,\alpha_N)\middle|x_n=x,\alpha_n=\alpha\right] \tag{8}$$

для $n=\overline{0;N-1}$ і

$$C_N(\pi,x,\alpha):=E^{\pi}\left[c(x_N,\alpha_N)|x_N=x,\alpha_N=\alpha\right]=c(x,\alpha)$$

.

$C_n(\pi,x,\alpha)$ – очікувана вартість за період з моменту n і далі за умови використання стратегії π та $(x_n,\alpha_n)=(x,\alpha)$. У такому разі з (1) і (8) випливає, що

$$\Re(\pi,x,\alpha)=C_0(\pi,x,\alpha). \tag{9}$$

Для доведення теореми покажемо, що для всіх $(x,\alpha)\in X$ і $n=\overline{0;N}$

$$C_n(\pi,x,\alpha)\ge V_n(x,\alpha) \tag{10}$$

з рівністю при $\pi=\pi^*$, тобто

$$C_n(\pi^*,x,\alpha)=V_n(x,\alpha). \tag{11}$$

Зокрема, при $n=0$

$$\Re(\pi,x,\alpha)\ge V_0(x,\alpha) \text{ і}$$

$$\Re(\pi^*,x,\alpha)=V_0(x,\alpha) \text{ для всіх } (x,\alpha)\in X,$$

що доводить твердження (7), оскільки з $\Re(\pi,\cdot,\cdot)\geq V_0(\cdot,\cdot)$ для довільної стратегії π випливає, що $\Re^*(\cdot,\cdot)\geq V_0(\cdot,\cdot)$.

Доведення тверджень (10), (11) проведемо за допомогою зворотньої індукції. Для $n=N$ істинність (10), (11) очевидна, оскільки з (9) і (3) слідує $C_N(\pi,x,\alpha)=V_N(x,\alpha)=c(x,\alpha)$.

Далі, за припущенням індукції, для деякого $n=N-1;0$

$$C_{n+1}(\pi,x,\alpha)\geq V_{n+1}(x,\alpha) \text{ для всіх } (x,\alpha)\in X. \quad (12)$$

Тоді матимемо

$$C_n(\pi,x,\alpha)=E^{\pi}\left[\sum_{k=n}^{N-1}e^{s_k-s_n}c(x_k,\alpha_k,a_k)+e^{s_n-s_N}c(x_N,\alpha_N)\middle|x_n=x,\alpha_n=\alpha\right]=$$

$$=\int_A e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,a)+\int_X C_{n+1}(\pi,y,\beta)Q(d(y,\beta)|x,\alpha,a)\right]\pi_n(da|x,\alpha).$$

Отже,

$$C_n(\pi,x,\alpha)\geq\min_{a\in A_{(x,\alpha)}}\left\{e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,a)+\right.\right.$$

$$\left.\left.+\int_X V_{n+1}(y,\beta)Q(d(y,\beta)|x,\alpha,a)\right]\right\}=V_n(x,\alpha),$$

що доводить (10).

З іншого боку, якщо в (12) рівність виконується при $\pi=\pi^*$ так, що $\pi_n(\cdot|h_n)$ – міра Дірака, зосереджена в $f(x_n,\alpha_n)$, то рівність виконується також і в попередніх обрахунках, що і доводить (11).

Вимірні умови вибору

У попередній теоремі припускалося існування вимірного селектора, що задовольняє певні співвідношення. Визначимо умови, що забезпечують існування селектора для напівмарковських процесів.

Припущення 6. Для відомої вимірної функції $u:X\to\mathbf{R}$ функцію $u^*:X\to\mathbf{R}$ визначено як:

$$u^*(x,\alpha):=\inf_{a\in A_{(x,\alpha)}}\left\{e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,a)+\right.\right.$$

$$\left.\left.+\int_X u(y,\beta)Q(d(y,\beta)|x,\alpha,a)\right]\right\}, \quad (13)$$

є вимірною і при цьому існує селектор $f\in F$ такий, що функція в дужках досягає свого мінімуму в $f(x,\alpha)\in A_{(x,\alpha)}$ для будь-якого $(x,\alpha)\in X$, тобто:

$$u^*(x,\alpha):=e^{-\alpha\tau(x,\alpha,f)}\left[c(x,\alpha,f)+\int_X u(y,\beta)Q(d(y,\beta)|x,\alpha,f)\right]$$

для всіх $(x,\alpha)\in X$.

Перед представленням умов сформулюємо деякі визначення в термінах даної роботи.

Нехай Y – метричний простір, а v – функція, що відображає Y в $\mathbf{R}\cup\{\infty\}$ так, що $v(y)<\infty$ принаймні для однієї точки $y\in Y$. Кажуть, що функція v *напівнеперервна знизу* в точці $y\in Y$, якщо $\liminf_{y_n\to y}v(y_n)\geq v(y)$ для всіх послідовностей $\{y_n\}$ з Y, що збігаються до y. Функцію v називають *напівнеперервною знизу*, якщо вона напівнеперервна знизу в усіх точках Y. Функцію $v:\Delta\to\mathbf{R}$ називатимемо *інфімум-компактною* на Δ, якщо для будь-яких $(x,\alpha)\in X$ і $r\in\mathbf{R}$ множина $\{a\in A_{(x,\alpha)}|v(x,\alpha,a)\leq r\}$ компактна. Багатозначна функція ψ з X в A називається напівнеперервною зверху, якщо $\psi^{-1}[F]$ замкнена для будь-якої замкненої множини $F\subset A$.

Позначимо як $\mathbf{B}(X)$ сімейство вимірних обмежених функцій на X, а $\mathbf{C}(X)\subset\mathbf{B}(X)$ – підсімейство неперервних функцій.

Умова 7. 1) Множина станів $A_{(x,\alpha)}$ компактна для всіх $(x,\alpha)\in X$;

2) функція вартості $c(x,\cdot)$ напівнеперервна знизу на $A_{(x,\alpha)}$ для всіх $(x,\alpha)\in X$;

3) функція

$$v'(x,\alpha,a):=\int_X v(y,\beta)Q(d(y,\beta)|x,\alpha,a) \quad (14)$$

на Δ задовольняє одну з наступних умов:

3.1) $v'(x,\alpha,\cdot)$ – напівнеперервна знизу на $A_{(x,\alpha)}$ для всіх $(x,\alpha)\in X$ і будь-якої $v\in\mathbf{C}(X)$;

3.2) $v'(x,\alpha,\cdot)$ – напівнеперервна знизу на $A_{(x,\alpha)}$ для всіх $(x,\alpha)\in X$ і будь-якої $v\in\mathbf{B}(X)$.

Умова 8. 1) $A_{(x,\alpha)}$ – компакт для всіх $(x,\alpha)\in X$, а багатозначна функція $(x,\alpha)\mapsto A_{(x,\alpha)}$ – напівнеперервна зверху;

2) функція вартості c напівнеперервна та обмежена знизу;

3) перехідне ядро Q задовольняє принаймні одну з умов:

3.1) Q – слабо неперервне, тобто для довільної функції $v\in\mathbf{C}(X)$ функція v', визначена співвідношенням (14), неперервна і обмежена на Δ;

3.2) Q – строго неперервне, тобто v' неперервна і обмежена на Δ для кожного $v\in\mathbf{B}(X)$.

Умова 9. 1) функція вартості c напівнеперервна та обмежена знизу та інфімум–компактна на Δ;

2) перехідне ядро Q або:

2.1) слабо неперервне;

2.2) строго неперервне.

Наступна теорема вказує на зв'язок між припущенням 6 та трьома умовами, що були наведені вище. Її доведення аналогічне доведенню теореми 7 в роботі [3], тому опускається.

Теорема 10. 1) Обидві умови, 7 і 8, тягнуть за собою виконання припущення 6 для будь-якої невід'ємної вимірної функції $u: X \to \mathbf{R}$. Крім того, якщо виконуються умови 7 (3.1) та 8 (3.1), то достатньо, щоб u була невід'ємною та напівнеперервною знизу, а при виконанні умови 8 (1, 2, 3.1) u в (13) – напівнеперервна знизу;

2) Умова 9 тягне за собою виконання припущення 6, якщо при виконанні пункту 2.1 u – невід'ємна та напівнеперервна знизу, а пункту 2.2 u – невід'ємна вимірна функція. Якщо, крім того, багатозначна функція $(x,\alpha) \mapsto A^*_{(x,\alpha)}$, де $A^*_{(x,\alpha)}$ дорівнює

$$\left\{ a \in A_{(x,\alpha)} \middle| u^*(x,\alpha) = e^{-\alpha\tau(x,\alpha,a)} \left[c(x,\alpha,a) + \int_X u(y,\beta) Q(d(y,\beta)|x,\alpha,a) \right] \right\},$$

напівнеперервна знизу, тоді u^* також напівнеперервна знизу.

Керування на нескінченному горизонті

У цьому розділі розглянемо керовану напівмарковську модель

$$\mathbf{P} = \langle (X,\Xi),(A,\mathrm{A}),F,\Delta,\mathrm{Q},\Phi,c \rangle$$

на нескінченному горизонті, тобто вважатимемо, що процес відбувається протягом зліченної кількості кроків.

Надалі припускатимемо, що функція вартості c невід'ємна. Крім того, для визначення поточного значення критерію якості керування в момент n-го переходу використовуватимемо позначення

$$\mathfrak{R}_n(\pi,x,\alpha) := E^{\pi}_{(x,\alpha)} \left[\sum_{k=0}^{n-1} e^{-S_k} c(x_k,\alpha_k,a_k) \right],\ \pi \in S_R,\quad (x,\alpha) \in X_0. \tag{15}$$

Можна переписати вираз для визначення критерію якості керування, застосувавши теорему про монотонну збіжність:

$$\mathfrak{R}(\pi,x,\alpha) = \lim_{n\to\infty} \mathfrak{R}_n(\pi,x,\alpha) \tag{16}$$

Вимірну функцію $v: X \to \mathbf{R}$ називатимемо розв'язком рівняння оптимальності, якщо вона задовольняє таку рівність:

$$v(x,\alpha) = \min_{a \in A_{(x,\alpha)}} \left\{ e^{-\alpha\tau(x,\alpha,a)} \left[c(x,\alpha,a) + \int_X v(y,\beta) Q(d(y,\beta)|x,\alpha,a) \right] \right\}. \tag{17}$$

Перш ніж сформулювати основні результати цього розділу, введемо до розгляду деякі припущення.

Припущення 11. 1) Функція вартості c напівнеперервна знизу та інфімум-компактна на Δ;

2) ядро переходу Q строго неперервне.

Припущення 12. Існує така стратегія π, що $\mathfrak{R}(\pi,x,\alpha) < \infty$ для будь-якого $(x,\alpha) \in X$.

Сформулюємо допоміжні твердження.

Лема 13. Нехай u і u_n, $n \in \mathbf{N}$, – напівнеперервні та обмежені знизу функції, інфімум-компактні на Δ. Тоді, якщо $u_n \uparrow u$, то:

$$\lim_{n\to\infty} \min_{a \in A_{(x,\alpha)}} u_n(x,\alpha,a) = \min_{a \in A_{(x,\alpha)}} u(x,\alpha,a) \quad \text{для будь-якого } (x,\alpha) \in X. \tag{18}$$

Доведення цього твердження цілком аналогічне доведенню леми 4.2.4 в роботі [5, с. 47].

Для аргументування існування вимірного селектора використаємо теорему 5 та наступне означення.

Означення 14. Як $M(X)^+$ позначимо пучок невід'ємних вимірних функцій на X, а для кожної функції $u \in M(X)^+$ введемо оператор T так, що Tu – функція, визначена на X таким співвідношенням:

$$Tu(x,\alpha) := \min_{a \in A_{(x,\alpha)}} \left\{ e^{-\alpha\tau(x,\alpha,a)} \left[c(x,\alpha,a) + \int_X u(y,\beta) Q(d(y,\beta)|x,\alpha,a) \right] \right\}. \tag{19}$$

Лема 15. За припущенням 11 T відображає $M(X)^+$ в себе, тобто для кожної функції $u \in M(X)^+$ функція Tu також належить $M(X)^+$; крім того, існує такий селектор $f \in F$, що

$$Tu := e^{-\alpha\tau(x,\alpha,f)} \left[c(x,\alpha,f) + \int_X u(y,\beta) Q(d(y,\beta)|x,\alpha,f) \right]$$

для будь-якого $(x,\alpha) \in X$.

Лема 16. Якщо припущення 2 і 3 виконуються, тоді:

1) якщо функція $u \in M(X)^+$ така, що $u \geq Tu$, то $u \geq \mathfrak{R}^*$;

2) якщо $u: X \to \mathbf{R}$ вимірна функція така, що Tu повністю визначена і, крім того, $u \geq Tu$ і

$$\lim_{n\to\infty} E^{\pi}_{(x,\alpha)} \left[e^{-S_n} u(x_n,\alpha_n) \right] = 0 \text{ для всіх } \pi \in S_R \text{ і } (x,\alpha) \in X, \tag{20}$$

то $u \leq \mathfrak{R}^*$.

Доведення. 1) Нехай функція $u \in M(X)^+$ така, що $u \geq Tu$. Тоді з леми 15 слідує, що можна вибрати селектор $f \in F$ такий, що

$$u(x,\alpha)\ge e^{-\alpha\tau(x,\alpha,f)}\left[c(x,\alpha,f)+\int\limits_X u(y,\beta)Q\big(d(y,\beta)\big|x,\alpha,f\big)\right].$$

Після N ітерацій останньої нерівності матимемо:

$$u(x,\alpha)\ge E^{\pi}_{(x,\alpha)}\left[\sum_{n=0}^{N-1}e^{-s_n}c(x_n,\alpha_n,a_n)\right]+E^{\pi}_{(x,\alpha)}\left[e^{-s_N}u(x_N,\alpha_N,a_N)\right], \quad (21)$$

де $\pi=(f,f,\ldots)=:f^{\infty}$ і

$$E^{\pi}_{(x,\alpha)}\left[e^{-s_N}u(x_N,\alpha_N,a_N)\right]=\int\limits_X u(y,\beta)Q^N\big(d(y,\beta)\big|x,\alpha,f\big).$$

З того, що $u\ge 0$, маємо:

$$u(x,\alpha)\ge E^{\pi}_{(x,\alpha)}\left[\sum_{n=0}^{N-1}e^{-s_n}c(x_n,\alpha_n,a_n)\right].$$

Граничний перехід при $N\to\infty$ дає $u(x,\alpha)\ge\Re(\pi,x,\alpha)\ge\Re^*(x,\alpha)$ для всіх $(x,\alpha)\in X$, що і завершує доведення пункту 1).

2) Нехай $\pi\in S_R$ і $(x,\alpha)\in X$. З властивості марковості та припущення $u\ge Tu$ отримуємо:

$$E^{\pi}_{(x,\alpha)}\left[e^{-s_{n+1}}u(x_{n+1},\alpha_{n+1})\big|h_n,a_n\right]=$$

$$=E^{\pi}_{(x,\alpha)}\left[e^{-s_{n+1}}u(x_{n+1},\alpha_{n+1})\big|x_n,\alpha_n,a_n\right]=$$

$$=e^{-s_{n+1}}\int\limits_X u(y,\beta)Q\big(d(y,\beta)\big|x_n,\alpha_n,a_n\big)=$$

$$=e^{-s_n}\left[c(x_n,\alpha_n,a_n)+\int\limits_X u(y,\beta)Q\big(d(y,\beta)\big|x_n,\alpha_n,a_n\big)-c(x_n,\alpha_n,a_n)\right]\ge$$

$$\ge e^{-s_n}\left[u(x_n,\alpha_n)-c(x_n,\alpha_n,a_n)\right].$$

Звідси

$$e^{-s_n}c(x_n,\alpha_n,a_n)\ge -E^{\pi}_{(x,\alpha)}\left[e^{-s_{n+1}}u(x_{n+1},\alpha_{n+1})\big|h_n,a_n\right]+e^{-s_n}u(x_n,\alpha_n).$$

Тепер, беручи математичне сподівання $E^{\pi}_{(x,\alpha)}$ і просумувавши по $n=\overline{0;N-1}$, отримаємо

$$E^{\pi}_{(x,\alpha)}\left[\sum_{n=0}^{N-1}e^{-s_n}c(x_n,\alpha_n,a_n)\right]\ge$$

$$\sum_{n=0}^{N-1}E^{\pi}_{(x,\alpha)}\left[-e^{-s_n}u(x_n,\alpha_n)+e^{-s_{n-1}}u(x_{n-1},\alpha_{n-1})\right]=$$

$$-E^{\pi}_{(x,\alpha)}\left[e^{-s_N}u(x_N,\alpha_N)\right]+u(x,\alpha) \text{ для будь-якого } N.$$

При $N\to\infty$ і з гіпотези (20) отримуємо $\Re(\pi,x,\alpha)\ge u(x,\alpha)$, що зводиться до $\Re^*\ge u$ при довільних $\pi\in S_R$ і $(x,\alpha)\in X$.

Лема 17. Якщо припущення 11 і 12 виконуються, тоді $v_n\uparrow V^*$ і V^* задовольняє рівняння оптимальності.

Доведення останньої леми цілком аналогічне доведенню леми 4 з [3].

Теорема 18. Нехай виконуються припущення 11 і 12. Тоді:

1) функція вартості $\Re^*$ – мінімальний розв'язок рівняння оптимальності, тобто

$$\Re^*(x,\alpha)=\min_{a\in A_{(x,\alpha)}}\left\{e^{-\alpha\tau(x,\alpha,a)}\left[c(x,\alpha,a)+\right.\right.$$

$$\left.\left.+\int\limits_X\Re^*(y,\beta)Q\big(d(y,\beta)\big|x,\alpha,a\big)\right]\right\} \quad (22)$$

для всіх $(x,\alpha)\in X$, і для будь-якого іншого розв'язку u рівняння оптимальності виконуватиметься нерівність $u(\cdot,\cdot)\ge\Re^*(\cdot,\cdot)$;

2) існує селектор $f_*\in F$ такий, що при $f_*(x,\alpha)\in A_{(x,\alpha)}$ досягається мінімум в (22), тобто

$$\Re^*(x,\alpha)=e^{-\alpha\tau(x,\alpha,f_*)}\left[c(x,\alpha,f_*)+\right.$$

$$\left.+\int\limits_X\Re^*(y,\beta)Q\big(d(y,\beta)\big|x,\alpha,f_*\big)\right] \quad (23)$$

для всіх $(x,\alpha)\in X$ і стаціонарна нерандомізована стратегія f_*^{∞} – оптимальна. І навпаки, якщо f_*^{∞} – стаціонарна нерандомізована оптимальна стратегія, то вона задовольняє (23);

3) якщо π^* – стратегія, така, що $\Re(\pi^*,\cdot,\cdot)$ – розв'язок оптимального рівняння і задовольняє умову

$$\lim_{n\to\infty}E^{\pi}_{(x,\alpha)}\left[e^{-s_n}\Re(\pi^*,x_n,\alpha_n)\right]=0 \text{ при довільних}$$

$$\pi\in S_R \text{ і } (x,\alpha)\in X, \quad (24)$$

то $\Re(\pi^*,\cdot,\cdot)=\Re^*(\cdot,\cdot)$ і π^* – оптимальна. Іншими словами, якщо виконується (24), то π^* оптимальна тоді і тільки тоді, коли $\Re(\pi^*,\cdot,\cdot)$ задовольняє (23);

4) якщо оптимальна стратегія існує, то існує оптимальна стаціонарна нерандомізована стратегія.

Доведення теореми 18 аналогічне доведенню теореми 10 з [3], тому його опускаємо.

Висновки

У роботі розглянуто керовані напівмарковські процеси з рандомізованим дисконтом. Знайдено достатні умови існування та єдиності оптимальних стратегій для задачі на скінченному та нескінченному горизонті. Запропоновано ітераційний метод знаходження оптимальної стратегії керування серед стаціонарних нерандомізованих стратегій. Їх вибір зумовлюється простотою практичної реалізації.

Одним з можливих напрямків подальшого дослідження таких процесів є вивчення багатовимірних моделей взаємодії економічних агентів на макроекономічному рівні. Зокрема застосування результатів роботи [2].

Список літератури

1. Berument H. The effects of different inflation risk prepius on interest rate spreads / Hakan Berument, Zubeyir Kilinc, Umit Ozlale // Physica A. – 2004. – Vol. 333. – P. 317–324.
2. Chornei R. K. Control of Spatially Structured Random Processes and Random Fields with Applications / Ruslan K. Chornei, Hans Daduna, Pavel S. Knopov. – New York : Springer Science + Business Media, Inc., 2006. – 261 p.
3. González-Hernández J. Markov control processes with randomized discounted cost / Juan González-Hernández, Raquiel R. López-Martínez, J. Rubén Pérez-Hernández // Math. Meth. Oper. Res. – 2007. – Vol. 65, № 1. – P. 27–44.
4. Gil-Alana L. A. Modelling the U.S. interest rate in terms of I(d) statistical models / Luis Alberiko Gil-Alana // Quarterly Review of Economics & Finance. – 2004. – Vol. 44, № 4. – P. 475–486.
5. Hernández-Lerma O. Discrete-time Markov control processes: basic optimality criteria / O. Hernández-Lerma, J. B. Lasserre. – Berlin Heidelberg New York : Springer, 1996. – 216 p.
6. Newell R. G. Discounting the distant future: how much do uncertain rates increase valuations? / Richard G. Newell, William A. Pizer // Journal of Environmental Economics and Management. – 2003. – Vol. 46, № 1 – P. 52–71.

N. Kovalchuk., R. Chornei

SEMI-MARKOV DECISION-MAKING PROCESSES WITH RANDOMIZED DISCOUNT

In this paper we consider semi-Markov decision processes with randomized discounted factors. We find sufficient conditions for the existence and uniqueness of optimal stationary nonrandomized strategies for control these processes on finite and infinite horizon.

Keywords: Decision-making processes, semi-Markov processes, randomized discount, optimality equation.

Матеріал надійшов 07.08.2013

УДК 519.7

Галкін О. А.

АВТОМАТИЧНИЙ ВИБІР ПАРАМЕТРІВ ЯДРА ОПОРНО-ВЕКТОРНИХ МАШИН

Досліджено методологію автоматичного вибору параметрів ядра опорно-векторних машин. Використано метод градієнтного спуску для мінімізації оцінок тестової похибки. Отримано гладку апроксимацію тестової помилки з використанням оцінки апостеріорних ймовірностей. Запропоновано та реалізовано алгоритм, де градієнтним кроком є напрямок градієнта в просторі параметрів. Проведено експериментальні дослідження для оцінки ефективності запропонованого методу.

Ключові слова: ядро, метод градієнтного спуску, згладжування оцінок, гладка апроксимація тестової помилки, множина перевірки.

Вступ

Алгоритм опорно-векторних машин (ОВМ) залежить від кількох параметрів. Один із них, що позначається через C, контролює відповідність між максимізацією поля та мінімізацією помилки. Інші параметри знаходяться в нелінійному відображенні у простір характеристик. Вони називаються *параметрами ядра*.

Зазвичай, встановлення параметрів здійснюється з використанням мінімаксного підходу, який полягає в максимізації поля по коефіцієнтах гіперплощини та зведенні до мінімуму оцінки помилки узагальнення по множині параметрів ядра. Однак, коли є множина параметрів, класична стратегія, що ґрунтується на методі перебору в просторі параметрів, стає нерозв'язною,