

О КОЛЬЦЕ ЭНДОМОРФИЗМОВ КОНЕЧНОПОРОЖДЕННОГО МОДУЛЯ НАД КОЛЬЦОМ ГЛАВНЫХ ИДЕАЛОВ

В настоящей работе изучаются эндоморфизмы конечнопорожденного модуля над кольцом главных идеалов. Устанавливается необходимое и достаточное условие аддитивной порождаемости кольца эндоморфизмов модуля рассматриваемого типа его автоморфизмами. Устанавливается также необходимое и достаточное условие совпадения множества характеристических подмодулей заданного модуля с множеством его вполне характеристических подмодулей. В случае конечных абелевых групп соответствующие результаты содержатся в работе Шода¹.

В дальнейшем будут использоваться следующие обозначения и понятия;

1) $\langle S \rangle$ — подмодуль модуля M , порожденный системой $S \subset M$;
2) (α) — главный идеал кольца R , порожденный элементом $\alpha \in R$;

3) $\text{Ann } S$ — аннулятор множества S , состоящего из элементов модуля M над кольцом R :

$$\text{Ann } S = \{r \in R \mid (\forall x \in S) rx = 0\};$$

4) $\text{Aut } M$ — группа автоморфизмов модуля M ;

5) $\text{End } M$ — группа эндоморфизмов модуля M ;

6) модуль M будет называться однородным, если M разлагается в прямую сумму изоморфных, между собой неразложимых циклических подмодулей. Число последних будет называться рангом ($\text{rg } M$) модуля². Очевидно, однородный модуль является либо свободным, либо примарным (т. е. $\text{Ann } M = (p^e)$, где p простой элемент кольца). Произвольный конечнопорожденный модуль M над кольцом главных идеалов может быть представлен в виде прямой суммы однородных компонент H_i ($i = 1, \dots, n$):

$$M = H_1 \oplus H_2 \oplus \dots \oplus H_n,$$

$$H_i = \langle b_i^1 \rangle \oplus \dots \oplus \langle b_i^{r_i} \rangle,$$

где

$$\text{Ann } b_i^j = (p_i^{e_{ij}}) \text{ при } i \neq 1,$$

и

$$\text{Ann } b_i^1 = \begin{cases} p_i^{e_i}, & \text{если модуль } M \text{ периодичен,} \\ 0 & \text{в противном случае;} \end{cases}$$

¹ Shoda K. Über die charakteristischen Untergruppen einer endlichen Abelschen Gruppe. — «Mathematische Zeitschrift», 1930, № 31, p. 27—39.

² В случае свободного однородного модуля введенное понятие совпадает с общепринятым.

7) произвольное кольцо A аддитивно порождается множеством $S \subset A$, если S является системой образующих аддитивной группы кольца A .

Пусть M — конечнопорожденный модуль над кольцом главных идеалов R . Согласно (1) $M = H_1 \oplus \dots \oplus H_n$, где $H_i = \langle b_i^1 \rangle \oplus \dots \oplus \langle b_i^r \rangle$ — однородные компоненты модуля M .

Теорема 1а. Если кольцо R порождается своими единицами, то кольцо $\text{End } M$ тогда и только тогда аддитивно порождается группой $\text{Aut } M$, когда модуль M не имеет двух различных однородных компонент, удовлетворяющих условию

$$\left. \begin{aligned} \text{rg } H_i &= 1; \\ \text{Ann } H_i &\text{ равен } 0 \text{ или } (q^*), \text{ где } q \end{aligned} \right\} \quad (2)$$

такой простой элемент кольца R , что $|R/(q)| = 2$.

Замечание. В случае конечнопорожденной аддитивной абелевой группы таким элементом q является число 2. В общем случае $2 = 1 + 1 \equiv 0 \pmod{q}$, так как $|R/(q)| = 2$. Если указанный элемент q существует, то сумма $\epsilon_1 + \dots + \epsilon_m$, где ϵ_i ($i = 1, \dots, m$) — единицы кольца R , сравнима с 1 по $\text{mod } q$ при нечетном m и сравнима с 0 по $\text{mod } q$ при четном m . Поэтому в разложении элемента из R в сумму единиц число слагаемых имеет определенную четность. Отсюда следует, что такой элемент q единственен с точностью до ассоциированности. Действительно, пусть q_1 и q_2 — два таких элемента. Так как $q_1 \equiv 0 \pmod{q_1}$, то q_1 разлагается в сумму четного числа единиц, и, следовательно, $q_1 \equiv 0 \pmod{q_2}$. Ввиду неприводимости элементов q_1 и q_2 отсюда вытекает, что q_1 ассоциировано с q_2 .

Доказательство теоремы 1.

1. *Достаточность.*

1. Сначала рассмотрим случай однородного модуля M . Пусть $\{b^1, \dots, b^r\}$ — базис модуля M и S — матрица эндоморфизма $\sigma \in \text{End } M$ в этом базисе. В силу однородности M S является матрицей над кольцом $\tilde{R} = R/\text{Ann } M$. Как известно, существуют такие унимодулярные матрицы P и Q над соответствующим кольцом, что матрица PSQ диагональна. Так как унимодулярные матрицы соответствуют автоморфизмам модуля M , то достаточно доказать, что любая диагональная матрица $\text{diag}(s_1, \dots, s_r)$ представима в виде суммы унимодулярных.

Если $r > 1$ либо $r = 1$, но $\text{Ann } M = (p^*)$, где $|R/(p)| \neq 2$, то матрица $\text{diag}(s_1, \dots, s_r)$ над кольцом $\tilde{R}$ представима в виде суммы как двух, так и трех унимодулярных матриц. Так как в случае необходимости всегда может быть прибавлено нужное число раз $E - E$, где E — единичная матрица, то в рассматриваемой ситуации матрица $\text{diag}(s_1, \dots, s_r)$ представима в виде суммы любого числа $k \geq 2$ унимодулярных матриц. Очевидно, достаточно рас-

смотреть случаи, когда $1 \leq r \leq 3$. При $r = 2$ утверждение вытекает из равенств

$$\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} = \begin{pmatrix} s_1 & -1 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & s_2 \end{pmatrix} = \begin{pmatrix} s_1 - 1 & -1 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & s_2 - 1 \end{pmatrix} + E.$$

При $r = 3$ утверждение вытекает из равенств

$$\begin{aligned} \begin{pmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & s_2 - 1 & 1 \\ 0 & 1 & s_3 - 1 \end{pmatrix} + \begin{pmatrix} s_1 & 0 & -1 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & s_2 - 2 & 1 \\ 0 & 1 & s_3 - 2 \end{pmatrix} + \begin{pmatrix} s_2 - 1 & 0 & -1 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} + E. \end{aligned}$$

Если $r = 1$, $\tilde{R} = R/(p^s)$, $|R/(p)| \neq 2$, то утверждение вытекает из того, что любой элемент $s \in R$ допускает следующее разложение:

при $s \equiv 0 \pmod{p}$ $s = (s + 1) - 1 = (s - 1) + \gamma + (1 - \gamma)$,

при $s \equiv 1 \pmod{p}$ $s = (s + \gamma) - \gamma = s + s - s$,

при $s \not\equiv 0 \pmod{p}$, $s \not\equiv 1 \pmod{p}$ $s = (s - 1) + 1 = s + s - s$,

где γ такой элемент из R , что $\gamma \neq 0$, $\gamma \neq 1 \pmod{p}$.

В общем случае при $r = 1$ можно лишь утверждать, что матрица $S = (s)$ представима в виде суммы некоторого числа унимодулярных матриц (это следует из аддитивной порождаемости кольца R своими единицами).

2. Пусть M — модуль, не имеющий однородных компонент, удовлетворяющих условию (2). Тогда каждый эндоморфизм модуля M представим в виде суммы любого числа $k \geq 2$ автоморфизмов.

Действительно, если число n однородных компонент модуля M равно 1, то утверждение уже доказано; допуская его справедливость при $n \leq m$, докажем ее для $n = m + 1$.

Пусть $N_1 = H_1$, $N_2 = H_2 \oplus \dots \oplus H_{m+1}$, $\{\pi_1, \pi_2\}$ — пара проекторов, отвечающая прямому разложению $M = N_1 \oplus N_2$. Из $\sigma \in \text{End } M$ следует, что $\sigma = \pi_1 \sigma \pi_1 + \pi_1 \sigma \pi_2 + \pi_2 \sigma \pi_1 + \pi_2 \sigma \pi_2$. В силу индуктивного предположения для любого числа $k \geq 2$ $\pi_i \sigma \pi_i|_{N_i} = \sum_{l=1}^k \sigma_l^i$, где $\sigma_l^i \in \text{Aut } N_i$. Очевидно, $\sigma_l = \sigma_l^1 + \sigma_l^2$ при $l \geq 3$ — авто-

морфизмы модуля M . Докажем, что $\sigma_1 = (\sigma_1^1 + \sigma_2^1) + \pi_1 \sigma \pi_2$ и $\sigma_2 = (\sigma_1^2 + \sigma_2^2) + \pi_2 \sigma \pi_1$ также автоморфизмы модуля M . Действительно, пусть $x = x_1 + x_2$ ($x_1 \in N_1$, $x_2 \in N_2$) и $\sigma_1(x) = \sigma_1(x_1 + x_2) = 0$. Тогда

$$\sigma_1^1(x_1) + \sigma_2^1(x_2) + \pi_1 \sigma \pi_2(x_2) = 0 \Rightarrow [\sigma_2^1(x_2) = 0,$$

$$\sigma_1^1(x_1) + \pi_1 \sigma \pi_2(x_2) = 0, (\sigma_2^1(x_2) \in N_2, \sigma_1^1(x_1) + \pi_1 \sigma \pi_2(x_2) \in N_1)] \Rightarrow$$

$$\Rightarrow x_2 = 0 (\sigma_2^1 \in \text{Aut } N_2) \Rightarrow x_1 = 0 (\sigma_1^1 \in \text{Aut } N_1).$$

Следовательно, σ_1 — мономорфизм. Пусть $x = x_1 + x_2 \in M$ ($x_1 \in N_1$, $x_2 \in N_2$). Полагая $y_1 = (\sigma_1^1)^{-1}(x_1)$ и $y_2 = (\sigma_2^1)^{-1}(x_2 - \pi_1 \sigma \pi_2(y_1))$, получим $\sigma_1(y_1 + y_2) = x_1 + x_2 = x$. Таким образом, σ_1 — эпиморфизм, и, следовательно, $\sigma_1 \in \text{Aut } M$. Аналогично $\sigma_2 \in \text{Aut } M$. В результате получаем $\sigma = \sigma_1 + \sigma_2 + \dots + \sigma_k$, где $\sigma_i \in \text{Aut } M$.

3. Пусть модуль M имеет однородную компоненту H_i , удовлетворяющую условию (2). В этом случае $M = N_1 + N_2$, где $N_1 = H_i$, $N_2 = \bigoplus_{j \neq i} H_j$ и N_2 уже не имеет однородных компонент, удовлетворяющих условию (2). Применяя предыдущие соображения, получим, что из возможности разложения $\pi_1 \sigma \pi_1$ в сумму k автоморфизмов модуля N_1 следует возможность разложения σ в сумму k автоморфизмов модуля M . Тем самым достаточность условия теоремы доказана.

II. *Необходимость.* Допустим, что модуль M имеет две различные однородные компоненты, удовлетворяющие условию (2). Покажем, что тогда существует эндоморфизм, не представимый в виде суммы автоморфизмов. Предварительно докажем следующие леммы.

Лемма 1. Если подмодуль M' периодического конечнопорожденного модуля M над кольцом главных идеалов R изоморфен M , то $M' = M$.

Доказательство. Так как периодический конечнопорожденный модуль разлагается в прямую сумму примарных, то лемму достаточно доказать для примарных модулей. Очевидно, что утверждение леммы верно для однородных модулей (M' в этом случае выделяется прямым слагаемым). Допустим, что его справедливость уже доказана для всех примарных модулей, число однородных компонент у которых в разложении (1) меньше, чем у модуля M . Докажем ее для модуля M . Пусть φ — такой изоморфизм модуля M в себя, что $\varphi(M) = M' \subseteq M$, и H_1 — однородная компонента модуля M с наименьшим аннулятором, т. е. $\text{Ann } H_1 = \text{Ann } M$. Тогда $\varphi(H_1) = H'_1 \cong H_1$. Можно считать, что $H'_1 = H_1$. Действительно, так как $\text{Ann } H'_1 = \text{Ann } M$, то H'_1 можно дополнить до однородной компоненты H''_1 модуля M так, что $H''_1 = H'_1 \oplus H$, где H — подмодуль модуля M . Ввиду того что $\text{rg } H'_1 = \text{rg } H_1 = \text{rg } H''_1$, $H'_1 = H''_1$ и, следовательно, изоморфизм $\alpha \varphi$ модуля M в себя, где $\alpha \in \text{Aut } M$ и $\alpha H''_1 = H_1$, уже обладает указанным свойством. Итак, пусть $M = H_1 \oplus N$ и $M' = H_1 \oplus (N \cap M')$. Тогда $N \supset N \cap M' \cong M'/H_1 = \varphi(M)/\varphi(H_1) \cong M/H_1 \cong N$. В силу индуктивного предположения $N = N \cap M'$ и, следовательно, $M' = M$. Лемма доказана.

Лемма 2. Пусть $M = H_1 \oplus \dots \oplus H_n$, где H_i ($i = 1, \dots, n$) — однородные компоненты модуля M . Тогда, если $\varphi \in \text{Aut } M$, то $\pi_i \varphi \pi_i|_{H_i} \in \text{Aut } H_i$, где π_i — проектор на H_i .

Доказательство. При $n = 1$ утверждение справедливо. Допуская доказанной его справедливость при $n < m$, докажем ее для $n = m$. Пусть H_1 — такая однородная компонента модуля M , что либо а) $\text{Ann } H_1 = 0$ либо б) модуль M периодичен, $\text{Ann } H_1 = (p^e)$ и из того, что $\text{Ann } H_j = (p^e)$ ($j \neq 1$), следует, что $\delta < e$. Из $x \in H_1$ и $\pi_1 \varphi \pi_1(x) = 0$ следует, что $\varphi(x) \in H_2 \oplus \dots \oplus H_m$. Так как H_1 — однородный модуль, то существует такой элемент $y \in H_1$, что $\text{Ann } y = \text{Ann } H_1$ и $x = \xi y$ ($\xi \in R$). Отсюда $\xi \pi_1 \varphi(y) = 0$. Так как $\varphi \in \text{Aut } M$, то $\text{Ann } \varphi(y) = \text{Ann } y = \text{Ann } H_1$. Из выбора однородной компоненты H_1 следует, что $\text{Ann } \pi_1 \varphi(y) = \text{Ann } H_1$, откуда $\xi \in \text{Ann } H_1$ и $x = \xi y = 0$. Таким образом, $\pi_1 \varphi \pi_1|_{H_1}$ — мономорфизм. Покажем, что $\pi_1 \varphi \pi_1|_{H_1} \in \text{Aut } H_1$. Если $\text{Ann } H_1 \neq 0$, то это вытекает из леммы 1. Допустим, что $\text{Ann } H_1 = 0$. Так как $\varphi \in \text{Aut } M$, то для любого $x \in H_1$ существует такой элемент $z = z_1 + \dots + z_m$ ($z_i \in H_i$, $i = 1, \dots, m$), что $x = \varphi(z)$. Из периодичности H_i при $i \geq 2$ следует, что $\text{Ann } \varphi(z_i) \neq 0$. Поэтому $\pi_1 \varphi(z_i) = 0$ при $i \geq 2$ и, следовательно, $x = \pi_1 \varphi(z) = \pi_1 \varphi(z_1) = \pi_1 \varphi \pi_1(z)$. Таким образом, $\pi_1 \varphi \pi_1(H_1) = H_1$, откуда и вытекает, что $\pi_1 \varphi \pi_1|_{H_1} \in \text{Aut } H_1$.

Покажем теперь, что отображение

$$\psi = \left(\sum_{i \geq 2} \pi_i \right) \varphi \left(\sum_{i \geq 2} \pi_i \right) \Big|_{H_2 \oplus \dots \oplus H_m}$$

является автоморфизмом модуля $H_2 \oplus \dots \oplus H_m$. Если $x \in H_2 \oplus \dots \oplus H_m$ и $\psi(x) = 0$, то, очевидно, $\varphi(x) \in H_1$ и, следовательно, $\varphi(x) = \alpha y$, где $y \in H_1$, $\text{Ann } y = \text{Ann } H_1$, $\alpha \in R$. Так как $\varphi \in \text{Aut } M$, то $y = \varphi(z)$, где $z \in M$. Отсюда $\varphi(x) = \alpha y = \alpha \varphi(z) = \varphi(\alpha z)$ и, следовательно, $x = \alpha z$.

Поэтому $\alpha \pi_1(z) = \pi_1(x) = 0$, т.е. $\alpha \in \text{Ann } \pi_1(z)$. Так как $\varphi \in \text{Aut } M$, то $\text{Ann } z = \text{Ann } \varphi(z) = \text{Ann } H_1$. Поэтому $\text{Ann } \pi_1(z) = \text{Ann } H_1$ и, таким образом, $\alpha \in \text{Ann } H_1$. Следовательно, $\varphi(x) = \alpha y = 0$ и $x = 0$. Тем самым доказано, что ψ — мономорфизм. В силу леммы 1 $\psi \in \text{Aut}(H_2 \oplus \dots \oplus H_m)$. Применяя индуктивное предположение к модулю $H_2 \oplus \dots \oplus H_m$, получаем: $\pi_i \varphi \pi_i|_{H_i} \in \text{Aut } H_i$ ($i \geq 2$). Лемма доказана.

Пусть $H_1 = \langle b_1 \rangle$ и $H_2 = \langle b_2 \rangle$ — две однородные компоненты модуля M , удовлетворяющие условию (2). Это означает, что $\text{Ann } H_1$ равен нулю или (q^{e_1}) , $\text{Ann } H_2 = (q^{e_2})$, где q — такой простой элемент кольца R , что $|R/(q)| = 2$. Покажем, что эндоморфизм φ модуля M , определяемый условиями: $\varphi(b_1) = qb_1$, $\varphi(b_2) = b_2$, $\varphi(H_i) = 0$ ($i > 2$), не представим в виде суммы автоморфизмов. Пусть $\{\sigma_1, \dots, \sigma_k\}$ — произвольный набор автоморфизмов модуля M и

$$\sigma_i(b_1) = \xi_{i1}b_1 + \xi_{i2}b_2 + a_{i1},$$

$$\sigma_i(b_2) = \eta_{i1}b_1 + \eta_{i2}b_2 + a_{i2},$$

$$(\xi_{i1}, \eta_{ij} \in R; a_{ij} \in H_3 \oplus \dots \oplus H_n; i = 1, \dots, k).$$

В силу леммы $2\xi_{i1}b_1$ является базисным элементом модуля H_1 . Следовательно, ξ_{i1} является единицей кольца R , если $\text{Ann } H_1 = 0$,

и $(\xi_{i1}, q) = 1$, если $\text{Ann } H_1 = (q^{\varepsilon_1})$. В обоих случаях, так как $|R/(q)| = 2$, $\xi_{i1} \equiv 1 \pmod{q}$. Аналогично $\eta_{i2} \equiv 1 \pmod{q}$. Поэтому, если

$$\sigma_1(b_1) + \dots + \sigma_k(b_1) = \alpha b_1 + c_1,$$

$$\sigma_1(b_2) + \dots + \sigma_k(b_2) = \beta b_2 + c_2,$$

где $\alpha, \beta \in R$, $c_1 \in H_2 \oplus \dots \oplus H_n$, $c_2 \in H_1 \oplus H_3 \oplus \dots \oplus H_n$, то $\alpha \equiv \beta \equiv k \pmod{q}$. Отсюда следует, что указанный выше эндоморфизм φ не может совпадать с $\sigma_1 + \dots + \sigma_k$. Теорема доказана.

Теорема 16. Если кольцо R не порождается своими единицами, то кольцо $\text{End } M$ тогда и только тогда порождается группой $\text{Aut } M$, когда ранг свободной компоненты модуля M отличен от единицы, и модуль M не имеет двух однородных компонент, удовлетворяющих условию (2).

Доказательство. Для доказательства достаточности условия теоремы 16, очевидно, достаточно заметить, что каждый эндоморфизм φ периодического циклического модуля $\langle b \rangle$, удовлетворяющего условию (2), представим в виде суммы автоморфизмов (если $\varphi(b) = \rho b$ и $\varphi \notin \text{Aut } \langle b \rangle$, то $\rho \equiv 0 \pmod{q}$ и $\varphi = \varphi_1 + \varphi_2$, где $\varphi_1, \varphi_2 \in \text{Aut } \langle b \rangle$, $\varphi_1(b) = b$, $\varphi_2(b) = (\rho - 1)b$), и затем воспользоваться доказательством достаточности условия теоремы 1а.

Пусть модуль M имеет свободную компоненту $\langle b \rangle$ ранга один. Эндоморфизм, переводящий b в ρb , где ρ — элемент кольца R , не разложимый в сумму единиц, не представим в виде суммы автоморфизмов. Если модуль M имеет две периодические однородные компоненты, удовлетворяющие условию (2), то такой эндоморфизм построен при доказательстве теоремы 1а. Тем самым доказана необходимость условия теоремы 16. Теорема доказана.

Следствие. Если модуль M удовлетворяет условию теоремы 1а или 1б, то его любой характеристический подмодуль является вполне характеристическим.

Эти условия, как оказывается, не являются необходимыми для совпадения множества характеристических подмодулей модуля M с множеством его вполне характеристических подмодулей. Справедлива

Теорема 2. Для того, чтобы все характеристические подмодули модуля M являлись вполне характеристическими, необходимо и достаточно, чтобы модуль M не имел таких двух однородных компонент H_1 и H_2 , удовлетворяющих условию (2), что а) $\text{Ann } H_1 = (q^{\varepsilon_1})$, $\text{Ann } H_2 = (q^{\varepsilon_2})$ и $\varepsilon_1 - \varepsilon_2 > 1$ или б) $\text{Ann } H_1 = 0$, $\text{Ann } H_2 = (q^{\varepsilon_2})$.

Доказательство. I. 1. Пусть модуль M имеет не более одной однородной компоненты, удовлетворяющей условию (2). Если ранг свободной компоненты отличен от единицы, то справедливость утверждения теоремы 2 вытекает из предыдущих теорем. Пусть ранг свободной компоненты H_1 равен единице и L — характеристический подмодуль модуля M . Элемент $x \in L$ имеет

вид $x = \xi b + a$, где $\xi \in R$, $a \in \sum_{i>1} H_i$, $H_1 = \langle b \rangle$. Из характеристичности L следует, что $\xi(\sum_{i>1} H_i) \subset L$. Выберем в модуле M такой базис $B = \{b_1^1, \dots, b_n^1\}$, что $a = \sum_{i>1} \eta_i a_i$, где $a_i \in H_i \cap B$ и $\eta_i \in R$.

Пусть ε — тождественный автоморфизм модуля M . Если $\text{rg } H_i > 1$, то обозначим через φ_i автоморфизм модуля M , переводящий a_i в $a_i + a'_i$, где $a'_i \in H_i \cap B$, $a'_i \neq a_i$, и оставляющий остальные базисные элементы неподвижными. Через φ' обозначим автоморфизм модуля M , переводящий a'_i в $a_i + a'_i$. Тогда $(\varphi_i - \varepsilon)(\xi b + a) = \eta_i a'_i \in L$, $\varphi'_i(\eta_i a'_i) = \eta_i(a_i + a'_i) \in L$ и, следовательно, $\eta_i a_i \in L$.

Если $\text{rg } H_i = 1$ ($i \neq 1$), $\text{Ann } H_i = (p_i^{\text{ст}})$, то существует такой элемент $\gamma \in R$, что $\gamma \not\equiv 0 \pmod{p_i}$ и $\gamma \not\equiv 1 \pmod{p_i}$, и, следовательно, эндоморфизм φ_i , переводящий a_i в $a_i - \gamma a_i$ и оставляющий остальные базисные элементы неподвижными, является автоморфизмом модуля M .

Поэтому $(\varepsilon - \varphi_i)(\xi b + a) = \eta_i \gamma a_i \in L$, откуда $\eta_i a_i \in L$. Таким образом, $a \in L$ и $\xi b \in L$. Следовательно, $L = \zeta b + N$ ($\zeta \in R$), где N — подмодуль модуля M , содержащийся в $\sum_{i>1} H_i$. Обозначив через π_1 проектор на H_1 , через π_2 — проектор на $\sum_{i>1} H_i$, для произвольного $\psi \in \text{End } M$ будем иметь:

$$\psi = \pi_1 \psi \pi_1 + \pi_2 \psi \pi_1 + \pi_1 \psi \pi_2 + \pi_2 \psi \pi_2.$$

Если $x \in L$, то, очевидно, $\pi_1 \psi \pi_1(x) \in L$, $\pi_1 \psi \pi_2(x) = 0$ и $\pi_2 \psi \pi_1(x) \in \zeta(\sum_{i>1} H_i) \subset L$.

Так как для модуля $H_2 + \dots + H_n$ справедливо утверждение теорем 1а и 1б, то $\pi_2 \psi \pi_2|_{H_2 + \dots + H_n} = \sum_{j=1}^n \sigma_j$, где $\sigma_j \in \text{Aut}(\sum_{i>1} H_i)$. Из этого вытекает, что $\psi(x) \in L$, и следовательно, подмодуль L является вполне характеристическим.

2. Допустим, что у модуля M , удовлетворяющего условию теоремы, имеются периодические однородные компоненты $N_1, \dots, N_k$ ($k \geq 2$), удовлетворяющие условию (2), $\text{Ann } N_i = (q_i^{\text{ст}})$, и L — характеристический подмодуль модуля M . Обозначим через K_i ($i = 1, \dots, k$) подмодуль модуля M , состоящий из всех таких элементов подмодуля L , что их аннуляторы равны степени простого идеала (q_i) . Любой элемент из M с нулевым аннулятором имеет вид: $\xi b + a$, где $\xi \in R$, $a, b \in M$, $\text{Ann } a \neq 0$ и b — базисный элемент модуля M . Если $\xi b + a \in L$, то $a, \xi b \in L$. Для доказательства вполне характеристичности подмодуля L , следовательно, достаточно доказать, что любой эндоморфизм модуля M переводит K_i в K_i , т. е. задача сводится к примарному случаю.

Пусть M — примарный модуль, удовлетворяющий условию теоремы 2, $H_1 = \langle b_1 \rangle$ и $H_2 = \langle b_2 \rangle$ — его однородные компо-

ненты, удовлетворяющие условию (2), и L — его характеристический подмодуль. Возьмем произвольные $\varphi \in \text{End } M$ и $x \in L$. Элемент x равен $\xi_1 b_1 + \xi_2 b_2 + h$, где $\xi_1, \xi_2 \in R$, $h \in H_3 \oplus \dots \oplus H_n$. Если $\xi_1 \neq 0$, $\xi_2 \neq 0$, то можно считать, что $\xi_1 = q^{\alpha_1}$, $\xi_2 = q^{\alpha_2}$, т. е. $x = q^{\alpha_1} b_1 + q^{\alpha_2} b_2 + h$. Если $\alpha_1 \leq \alpha_2$, то можно в базисе заменить b_1 на $b_1 + q^{\alpha_2 - \alpha_1} b_2$. ($\text{Ann } b_1 = (q^{\varepsilon_1}) \subset \text{Ann } b_2 = (q^{\varepsilon_2})$). Поэтому можно считать, что $x = q^{\alpha_1} b_1 + h$. Обозначив через π_i проектор на H_i , получаем

$$\varphi(x) = \pi_2 \varphi(x) + \sum_{i \neq 2} \pi_i \varphi(x).$$

Пусть $v \in \text{Aut } H_2$. Подмодуль $N = \bigoplus_{i \neq 2} H_i$ удовлетворяет условию теоремы 1а или 1б, и поэтому

$$\left(\sum_{i \neq 2} \pi_i \right) \varphi \left(\sum_{i \neq 2} \pi_i \right) |_N = \sum_{i=1}^k \sigma_i,$$

где $\sigma_i \in \text{Aut } N$.

Очевидно, что $\sigma_i + v \in \text{Aut } M$.

В силу характеристичности подмодуля L

$$\sum_{i=1}^k \sigma_i(x) = \sum_{i=1}^k (\sigma_i + v)(x) \in L.$$

Кроме того, если $\sigma \in \text{Aut } N$, то $\sigma + v \in \text{Aut } M$ и $(\sigma + v) + \pi_2 \varphi \times \left(\sum_{i \neq 2} \pi_i \right) \in \text{Aut } M$ (это доказано в пункте 1, 2 доказательства теоремы 1а). Отсюда вытекает, что $\sigma(x) + v(x) \in L$ и $\sigma(x) + v(x) + \pi_2 \varphi \left(\sum_{i \neq 2} \pi_i \right)(x) \in L$. Следовательно, $\pi_2 \varphi(x) \in L$, откуда $\varphi(x) \in L$.

Если $\alpha_1 > \alpha_2$, то в базисе можно заменить b_2 на $q^{\alpha_1 - \alpha_2} b_1 + b_2$. Это возможно благодаря тому, что $q^{\varepsilon_2} q^{\alpha_1 - \alpha_2} b_1 = 0$. Таким образом, можно считать, что $x = q^{\alpha_2} b_2 + h$. Как и в первом случае получаем: $\varphi(x) \in L$ и, следовательно, подмодуль L вполне характеристичен. Тем самым доказана достаточность условия теоремы 2.

Докажем необходимость условия теоремы.

Допустим, что модуль M имеет такие однородные компоненты $H_1 = \langle b_1 \rangle$ и $H_2 = \langle b_2 \rangle$, удовлетворяющие условию (2), что а) $\text{Ann } b_1 = (q^{\varepsilon_1})$, $\text{Ann } b_2 = (q^{\varepsilon_2})$, $\varepsilon_1 - \varepsilon_2 > 1$ или б) $\text{Ann } b_1 = 0$, $\text{Ann } b_2 = (q^{\varepsilon_1})$.

Пусть $H_3, \dots, H_k$ — все такие однородные компоненты модуля M , что $\text{Ann } H_i = (q^{\varepsilon_i})$ ($i = 3, \dots, k$), и пусть $\varphi \in \text{Aut } M$. В случае а) рассмотрим подмодуль

$$L = \langle q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2 \rangle \oplus \sum_{i=3}^k q^{\max(\varepsilon_i - \varepsilon_1, 0) + \varepsilon_2} H_i.$$

Если $x \in L$, то

$$x = \xi_1 (q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) + \sum_{i=3}^k \xi_i q^{\max(\varepsilon_i - \varepsilon_1, 0) + \varepsilon_2} h_i,$$

где $h_i \in H_i$, $\text{Ann } h_i = \text{Ann } H_i$, $\xi_i \in R$ ($i = 1, \dots, k$). Так как при $j > k$ $\text{Ann } H_j$ равен нулю или (p^a) , где $(p, q) = 1$, то $\pi_j \varphi \pi_i(x) = 0$ при $j > k$, $i \leq k$. Так как H_j — однородные компоненты, то при $j \leq k$ $\pi_j \varphi(h_i) = \alpha_{ij} q^{\max(\varepsilon_j - \varepsilon_i, 0)} h_j^i$, где $h_j^i \in H_j$, $h_1 = h_1^1 = b_1$, $h_2 = h_2^1 = b_2$, $\text{Ann } h_j^i = \text{Ann } H_j$,

$$\pi_i \varphi \pi_i(x) = \begin{cases} \text{при } j \geq 3, i \neq 2 \\ \alpha_{ij} \xi_i q^{\max(\varepsilon_j - \varepsilon_i, 0) + \max(\varepsilon_i - \varepsilon_1, 0) + \varepsilon_2} h_j^i \in L, \\ \text{при } j \geq 3, i = 2 \\ \alpha_{2j} \xi_1 q^{\max(\varepsilon_j - \varepsilon_2, 0) + \varepsilon_2 - 1} h_j^2 \in L, \\ \text{при } j = 2, i \neq 2 \\ \alpha_{i2} \xi_i q^{\max(\varepsilon_2 - \varepsilon_i, 0) + \max(\varepsilon_i - \varepsilon_1, 0) + \varepsilon_2} b_2 = 0, \\ \text{при } j = 1, i \geq 3 \\ \alpha_{i1} \xi_i q^{\max(\varepsilon_1 - \varepsilon_i, 0) + \max(\varepsilon_i - \varepsilon_1, 0) + \varepsilon_2} b_1 = \\ = \alpha_{i1} \xi_i q^{|\varepsilon_i - \varepsilon_1|} (q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) \in L, \\ \text{при } j = 1, i = 2 \\ \alpha_{21} \xi_1 q^{\varepsilon_1 - \varepsilon_2 + \varepsilon_2 - 1} b_2 = \alpha_{21} \xi_1 q^{\varepsilon_1 - \varepsilon_2 - 1} (q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) \in L. \end{cases}$$

Согласно лемме 2 $\pi_1 \varphi \pi_1|_{H_1} \in \text{Aut } H_1$ и $\pi_2 \varphi \pi_2|_{H_2} \in \text{Aut } H_2$. Поэтому $\pi_1 \varphi \pi_1(x) = \alpha \xi_1 q^{\varepsilon_2} b_1$ и $\pi_2 \varphi \pi_2(x) = \beta \xi_1 q^{\varepsilon_2 - 1} b_2$, где $\alpha, \beta \in R$, $\alpha \not\equiv 0 \pmod{q}$, $\beta \not\equiv 0 \pmod{q}$. Так как $\alpha - \beta \equiv 0 \pmod{q}$, т. е. $\alpha - \beta = \delta q$, где $\delta \in R$, то

$$\begin{aligned} \pi_1 \varphi \pi_1(x) + \pi_2 \varphi \pi_2(x) &= (\alpha - \beta) \xi_1 q^{\varepsilon_2} b_1 + \beta \xi_1 (q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) = \\ &= (\delta q + \beta) \xi_1 (q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) \in L. \end{aligned}$$

Тем самым доказано, что $\varphi(x) \in L$, и, следовательно, подмодуль L характеристичен.

В случае б рассматриваем подмодуль

$$L = \langle q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2 \rangle \oplus \sum_{i=3}^k q^{\varepsilon_i} H_i$$

и аналогично случаю а получаем, что L — характеристический подмодуль.

Рассмотрим такой эндоморфизм $\psi \in \text{End } M$, что $\psi(b_1) = q b_1$, $\psi(b_2) = b_2$ и $\psi(H_i) = 0$ при $i > 2$. Тогда $\psi(q^{\varepsilon_2} b_1 + q^{\varepsilon_2 - 1} b_2) = q^{\varepsilon_2 + 1} b_1 + q^{\varepsilon_2 - 1} b_2 \notin L$, откуда следует, что подмодуль L не является вполне характеристическим. Теорема доказана.

К ПРОБЛЕМЕ СИНТЕЗА ОПТИМАЛЬНОЙ СТРАТЕГИИ В УПРАВЛЯЕМЫХ ПРОЦЕССАХ

1. Постановка задачи. Рассмотрим обобщенный управляемый процесс, введенный в [1]. Пространства состояний E и управлений Y будем полагать непустыми борелевскими множествами.

Пусть $v(x, \pi)$ — ожидаемый выигрыш, полученный при стратегии π и начальном состоянии $x \in E$. Функцию $v^*(x) = \sup_{\pi} v(x, \pi)$ назовем ценой.

Стратегию $\pi^* : v(x, \pi^*) = v^*(x) \geq v(x, \pi)$ для всех π назовем оптимальной.

Свойства функций $v(x, \pi)$ и $v^*(x)$ достаточно полно исследованы в [2]. Наша задача в данной работе состоит в исследовании существования оптимальной стратегии и построении алгоритма ее поиска.

2. Проблема существования. В [2] было установлено, что если существует оптимальная стратегия, то она содержится в классе нерандомизированных марковских стратегий Π_D . Поэтому мы будем сразу исследовать класс Π_D . При этом будем использовать без дополнительных разъяснений определения и обозначения из [2]. В [2] были введены операторы T и F , связанные с нерандомизированными стратегиями:

$$Tv(x) = T(f)v(x) = \int v(z) dq(z/x, f(x)),$$

где $q(\cdot/x, f(x))$ — переходная функция управляемого процесса; f — элемент стратегии $\pi = (f, f, \dots)$;

$$F(f)v(x, \pi) = T[w(x) + v(x, \pi)].$$

Здесь $w = w(x, y, z)$ — непосредственный выигрыш за переход из $x \in E$ в $z \in E$ под влиянием управления $y = f(x) \in Y$.

Следующая теорема, которую мы приводим без доказательства, обосновывает процесс усовершенствования стратегии, рассмотренный Ховардом [3].

Теорема 2. 1. Если $F(f)v(x, \pi) \geq v(x, \pi)$, то $v(x, f^\infty) \geq v(x, \pi)$, где $f^\infty = (f, f, \dots)$ — стационарная стратегия.

Пусть Π_D — класс нерандомизированных марковских стратегий и $f \in P(Y/E)$, где P — множество вырожденных условных распределений, осуществляющих отображение множества состояний E в множество управлений Y .

Функции $v(x, \pi) \in L$ сопоставим оператор $Q = Q(f)$, $f \in P$, определенный следующим образом:

$$Q(f)v(x, \pi) = \max\{v(x, \pi), F(f)v(x, \pi)\}. \quad (2.2)$$

С результатом применения оператора Q к функции $v(x, \pi)$ свяжем стратегию σ :

$$\sigma = \begin{cases} \pi, & \text{если } v(x, \pi) > F(f)v(x, \pi), \\ f^\infty, & \text{если } F(f)v(x, \pi) \geq v(x, \pi). \end{cases} \quad (2.3)$$

Определим N -ю степень оператора Q . Пусть $\pi^{(1)} = \sigma$ — стратегия, полученная в результате применения оператора Q к $v(x, \pi)$. Тогда положим

$$Q^2v(x, \pi) = Q(f_2)Q(f_1)v(x, \pi) = \max\{v(x, \pi^{(1)}), F(f_2)v(x, \pi^{(1)})\}.$$

Пусть теперь $\pi^{(N-1)}$ — стратегия, полученная в результате $(N-1)$ -кратного применения оператора Q к функции $v(x, \pi)$ и $v(x, \pi^{(N-1)})$ — выигрыш, соответствующий стратегии $\pi^{(N-1)}$. Тогда

$$Q^N v(x, \pi) = Q(f_N) \dots Q(f_1) v(x, \pi) = \\ = \max \{v(x, \pi^{(N-1)}), F(f_N) v(x, \pi^{(N-1)})\}.$$

Теорема 2.2. Пусть $Q = Q(f)$ — оператор, определенный в (2.2), σ — стратегия, определенная в $(\mathcal{E}, \mathcal{Z})$, и $v(x, \sigma)$ — выигрыш, соответствующий стратегии σ . Тогда $v(x, \sigma) \geq Qv(x, \pi)$.

Доказательство. Обозначим $u = \max \{v(x, \pi), F(f) v(x, \pi)\}$. Пусть σ_1 — первый элемент стратегии σ , π_1 — элемент стационарной стратегии π и f — элемент стратегии f^∞ .

Для доказательства нам потребуется следующий факт: если $\pi = (f, f, \dots)$ — стационарная стратегия, то $F(f) v(x, \pi) = v(x, \pi)$.

С учетом этого факта, теоремы 2.1 и монотонности оператора F имеем:

а) если $v > Fv$, то $\sigma_1 = \pi_1$ и

$$F(\sigma_1) u = F(\pi_1) u \geq F(\pi_1) v(x, \pi) = v(x, \pi) = u;$$

б) если $Fv \geq v$, то $\sigma_1 = f$ и

$$F(\sigma_1) u = F(f) u \geq F(f) [F(f) v(x, \pi)] \geq F(f) v(x, \pi) = u.$$

Из а и б следует, что $F(\sigma) u \geq u$. Но, если $F(\sigma) u \geq u$, то в силу теоремы 2.1 $v(x, \sigma) \geq u$. Поэтому

$$v(x, \sigma) \geq u = \max \{v(x, \pi), F(f) v(x, \pi)\} = Qv(x, \pi).$$

Следствие 2.1. Пусть $\pi^{(N)}$ — стратегия, полученная в результате N -кратного применения оператора Q к $v(x, \pi)$. Тогда $v(x, \pi^{(N)}) \geq Q^N v(x, \pi)$ для любого $N = 1, 2, \dots$

Доказательство. Воспользуемся методом индукции. Для $N = 1$ это неравенство доказано в теореме 2.2. Покажем теперь, что

$$v(x, \pi^{(N-1)}) \geq Q^{N-1} v(x, \pi) \Rightarrow v(x, \pi^{(N)}) \geq Q^N v(x, \pi).$$

Действительно пусть $v(x, \pi^{(N-1)}) \geq Q^{N-1} v(x, \pi) = Q(f_{N-1}) \dots Q(f_1) v(x, \pi)$. Стратегию $\pi^{(N)}$ получаем в результате применения оператора $Q(f_N)$ к $v(x, \pi^{(N-1)})$. Тогда, применяя теорему 2.2, получаем:

$$v(x, \pi^{(N)}) \geq Q(f_N) v(x, \pi^{(N-1)}) \geq Q(f_N) [Q^{N-1} v(x, \pi)] = \\ = Q(f_N) Q(f_{N-1}) \dots Q(f_1) v(x, \pi) = Q^N v(x, \pi).$$

Следствие 2.2. $v(x, \pi^{(N+1)}) \geq v(x, \pi^{(N)})$, $N = 1, 2, \dots$

Доказательство. $v(x, \pi^{(N+1)}) \geq Qv(x, \pi^{(N)}) = \max \{v(x, \pi^{(N)}), F(f) v(x, \pi^{(N)})\} \geq v(x, \pi^{(N)})$.

Обозначим $g(\omega_n)$ непосредственный выигрыш, получаемый за n шагов, где

$$\omega_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}, x \in F, y \in Y.$$

Пусть $g^-(\omega_n) = -\min [0, g(\omega_n)]$ и $g^+(\omega_n) = \max [0, g(\omega_n)]$.

Обозначим далее $v_n(x, \pi) = M_x^\pi g(\omega_n)$ — математическое ожидание выигрыша, получаемое за n шагов при стратегии π и начальном состоянии $x \in E$. Пусть $L(A^-)$ и $L(A^+)$ — множество функций из L , удовлетворяющих дополнительно условиям

$$A^+ : M_x^\pi [\sup_n g^+(\omega_n)] < \infty,$$

$$A^- : M_x^\pi [\sup_n g^-(\omega_n)] < \infty.$$

Кроме того, положим $L(A^-, A^+) = L(A^-) \cap L(A^+)$.

Определение. Функция $v(x, \pi) \in L$ эксцессивна относительно оператора F , (F_a) , если $Fv \leq v$. Эксцессивная функция (э. ф.) v называется эксцессивной мажорантой (э. м.) $g = g(x, \pi)$, если $v(x, \pi) \geq g(x, \pi)$ для всех $x \in E$. Э. м. v называется наименьшей э. м. (н. э. м.) функции g , если v меньше или равна другой э. м.

Теорема 2.3. Пусть $v(x, \pi) \in L(A^-, A^+)$ для всех $\pi \in \Pi$, $\pi^{(N)}$ — стратегия, порожденная оператором Q^N , $\pi^* = \lim_{N \rightarrow \infty} \pi^{(N)}$ и $v(x, \pi^*)$ — выигрыш, соответствующий стратегии π^* . Обозначим $s(x) = \lim_{n \rightarrow \infty} Q^n v(x, \pi)$. Тогда

$$s(x) = v^*(x) = v(x, \pi^*) \sup v(x, \pi);$$

$s(x) \in L(A^-, A^+)$ и является наименьшей F_0 -эксцессивной мажорантой выигрышей $v(x, \pi)$.

Доказательство. Заметим прежде всего, что в силу определения Q^N и следствия 2.1 $Q^{N+1}v(x, \pi) \geq Q^N v(x, \pi)$. Тогда существует предел $\lim_{N \rightarrow \infty} Q^N v(x, \pi) = s(x)$.

Кроме того, в силу следствия 2.2 $v(x, \pi^{(N+1)}) \geq v(x, \pi^{(N)})$, т. е. $v(x, \pi^{(N)})$, $N=1, 2, \dots$ образует монотонно неубывающую ограниченную последовательность. Отсюда существует предел

$$\lim_{N \rightarrow \infty} v(x, \pi^{(N)}) = v(x, \lim_{N \rightarrow \infty} \pi^{(N)}) = v(x, \pi^*)$$

и тем самым существует предел $\lim_{N \rightarrow \infty} \pi^{(N)} = \pi^*$.

В силу неравенства $v(x, \pi^{(N)}) \geq Q^N v(x, \pi)$, которое согласно следствию 2.1. верно для любых $N=1, 2, \dots$ справедливо также неравенство

$$\begin{aligned} v^*(x) &= \sup_\pi v(x, \pi) \geq v(x, \pi^*) = \lim_{N \rightarrow \infty} v(x, \pi^{(N)}) \geq \\ &\geq \lim_{N \rightarrow \infty} Q^N v(x, \pi) = s(x). \end{aligned} \quad (2.4)$$

По условию $v(x, \pi) \in L(A^-, A^+)$ для всех π . Отсюда (см. [2]) $v^*(x) \in L(A^-, A^+)$. Тогда и $s(x) \in L(A^-, A^+)$.

Ясно также, что $s(x) = \lim_{N \rightarrow \infty} Q^N v(x, \pi) \geq Q^N v(x, \pi)$ для всех $N=1, 2, \dots$. Тогда $s(x) \geq v(x, \pi^{(N)})$ для всех $N=1, 2, \dots$. Отсюда

$$s(x) \geq \lim_{N \rightarrow \infty} v(x, \pi^{(N)}) = v(x, \pi^*).$$

Тем самым $s(x) \geq v(x, \pi)$ для всех $\pi \in \Pi$, откуда

$$s(x) \geq \sup_{\pi} v(x, \pi) = v^*(x). \quad (2.5)$$

Таким образом, $s(x)$ является мажорантой всех выигрышей $v(x, \pi)$.

Далее из неравенств (2.4) и (2.5) следуют равенства

$$s(x) = v^*(x) = v(x, \pi^*) = \sup_{\pi} v(x, \pi). \quad (2.6)$$

Покажем теперь, что $s(x)$ эксцессивна относительно оператора F_a . Действительно, в силу неравенства $Q^N v(x, \pi) \geq F_a(f)(Q^{N-1}v(x, \pi))$ и теоремы Лебега о монотонной сходимости имеем

$$\begin{aligned} s(x) &= \lim_{N \rightarrow \infty} Q^N v(x, \pi) \geq \lim_{N \rightarrow \infty} F_a(f) Q^{N-1} v(x, \pi) = \\ &= \lim_{N \rightarrow \infty} T(f) \beta_a(x) [w(x) + Q^{N-1}v(x, \pi)] = \\ &= T(f) \beta_a(x) [w(x) + \lim_{N \rightarrow \infty} Q^{N-1}v(x, \pi)] = \\ &= F_a(f) [\lim_{N \rightarrow \infty} Q^{N-1}v(x, \pi)] = F_a(f) s(x). \end{aligned} \quad (2.7)$$

Покажем, наконец, что $s(x)$ — наименьшая F_a -эксцессивная мажоранта выигрышей $v(x, \pi)$. Пусть $g(x)$ — также F -э.м. $v(x, \pi)$. Тогда $g(x) \geq F_a(f)g(x)$ и значит,

$$Qg(x) = \max \{g(x)\} = F_a(f)g(x)g(x),$$

$$g(x) = Q^N g(x) \geq Q^N v(x, \pi),$$

откуда $g(x) \geq \lim_{N \rightarrow \infty} Q^N v(x, \pi) = s(x)$. Таким образом, $s(x)$ -н. F_a -э.м. $v(x, \pi)$.

Следствие 2.3 $s(x)$ эксцессивна относительно оператора $U_a = \sup_{\pi} F_{\sigma}(f)$.

Доказательство. Из цепочки равенств (2.6) следует, что $s(x)$ является ценой. Тогда (см. [2]) $s(x)$ эксцессивна относительно оператора U_a .

Теорема 2.4. Пусть $v(x, \pi) \in L(A^-, A^+)$. Тогда существует оптимальная стационарная нерандомизированная марковская стратегия π^* такая, что $v(x, \pi^*) = \sup_{\pi} v(x, \pi) = v^*(x)$. Причем $v^*(x) = \lim_{N \rightarrow \infty} Q^N v(x, \pi) = s(x) \in L(A^-, A^+)$.

Доказательство. Существование оптимальной нерандомизированной марковской стратегии непосредственно вытекает из существования пределов $\lim_{N \rightarrow \infty} Q^N v(x, \pi) = s(x) \in L(A^-A^+)$ и $\lim_{N \rightarrow \infty} v(x, \pi^{(N)}) = v(x, \lim_{N \rightarrow \infty} \pi^{(N)}) = v(x, \pi^*)$ и цепочки равенств (2.6).

Покажем теперь, что оптимальная стратегия π^* стационарна. Пусть f^* — первый элемент стратегии π^* , а $\pi = (f, f, \dots)$ — произвольная стационарная стратегия. Очевидно, что

$$F(f^*)v(x, \pi) = \sup_i F(f_i)v(x, \pi) \geq F(f)v(x, \pi) = v(x, \pi).$$

Отсюда в силу теоремы 2.1 получаем $v(x, j^{*\infty}) \geq v(x, \pi)$ для всех $\pi \in \Pi$, что и доказывает теорему.

Теорема 2.5. Пусть $v(x, \pi) \in L(A^-)$. Тогда существует оптимальная стационарная нерандомизированная марковская стратегия π^* такая, что

$$v(x, \pi^*) = \sup_{\pi} v(x, \pi) = v^*(x) = s(x) = \lim_{N \rightarrow \infty} Q^N v(x, \pi),$$

причем цена $v^*(x) \in L(A^-)$.

Доказательство. Пусть $v(x, \pi) \in L(A^-)$. Рассмотрим функцию $v^b(x, \pi) = \min(b, v(x, \pi))$, $b \geq 0$. Очевидно, что $v^b(x, \pi) \in L(A^-, A^+)$. Тогда $\lim_{N \rightarrow \infty} Q^N v^b(x, \pi) = s^b(x) \in L(A^-, A^+)$, где $s^b(x)$ — эксцессивна относительно оператора U_a . Отсюда в силу теоремы 2.4. при каждом фиксированном $b \geq 0$ существует оптимальная стационарная стратегия, удовлетворяющая (2.7). Очевидно, что при $b \rightarrow \infty$ получим некоторую последовательность оптимальных стратегий. Причем, поскольку последовательность $\lim_{N \rightarrow \infty} Q^N v^b(x, \pi)$ является монотонно неубывающей при $b \rightarrow \infty$, то существует предел $\lim_{b \rightarrow \infty} (\lim_{N \rightarrow \infty} Q^N v^b(x, \pi))$. Отсюда существует предел

$$\begin{aligned} & \lim_{b \rightarrow \infty} (\lim_{N \rightarrow \infty} v^b(x, \pi^{(N)})) = \\ & = \lim_{b \rightarrow \infty} v^b(x, \lim_{N \rightarrow \infty} \pi^{(N)}) = \lim_{b \rightarrow \infty} v^b(x, \pi^*) = v(x, \pi^*). \end{aligned} \quad (2.8)$$

Таким образом, существует предел последовательности стратегий $\lim_{N \rightarrow \infty} \pi^{(N)} = \pi^*$, который, как видно из (2.8), не зависит от предельного перехода при $b \rightarrow \infty$.

Отсюда следует, что последовательность оптимальных стратегий, порождаемая предельным переходом при $b \rightarrow \infty$, составлена из одной и той же стратегии $\pi^* = \lim_{N \rightarrow \infty} \pi^{(N)}$, причем, в силу теоремы 2.4, имеют место равенства

$$s^b(x) = v^b(x, \pi^*) = (v^b)^*(x) = \sup_{\pi} v^b(x, \pi),$$

которые справедливы для любого $b \geq 0$. Тогда, переходя к пределу при $b \rightarrow \infty$, получаем

$$\begin{aligned} \lim_{b \rightarrow \infty} s^b(x) &= \lim_{b \rightarrow \infty} v^b(x, \pi^*) = \lim_{b \rightarrow \infty} (v^b)^*(x) = \\ &= \lim_{b \rightarrow \infty} \sup_{\pi} v^b(x, \pi) \end{aligned}$$

или $s(x) = v(x, \pi^*) = v^*(x) = \sup_{\pi} v(x, \pi)$, что и устанавливает существование оптимальной стратегии в рассматриваемом случае.

Кроме того, поскольку оптимальная стратегия π^* порождается оператором $\lim_{N \rightarrow \infty} Q^N v^b(x, \pi)$ и согласно (2.8) не зависит от предельного перехода при $b \rightarrow \infty$, то

$$\begin{aligned} v(x, \pi^*) &= \lim_{b \rightarrow \infty} v^b(x, \pi^*) = \lim_{b \rightarrow \infty} \lim_{N \rightarrow \infty} Q^N v^b(x, \pi) = \\ &= \lim_{N \rightarrow \infty} \lim_{b \rightarrow \infty} Q^N v^b(x, \pi) = \lim_{N \rightarrow \infty} Q^N v(x, \pi). \end{aligned}$$

Теперь, поскольку $v(x, \pi) \in L(A^-)$, то тем более $\sup_{\pi} v(x, \pi) = v^*(x) \in L(A^-)$, что и завершает доказательство теоремы.

Замечание. Аналогичным образом можно показать существование оптимальной стратегии, если $v(x, \pi) \in L(A^+)$. Однако в этом случае цена $v^*(x)$ не обязательно принадлежит классу $L(A^+)$.

3. Оптимизационный алгоритм. Построим теперь алгоритм поиска оптимальной стратегии. При этом для простоты будем полагать множества состояний E и управлений Y конечными, а управляемый процесс образует марковскую цепь с дискретным временем. В качестве критерия оптимальности используем ожидаемый средний выигрыш

$$\varphi(x, \pi) = \liminf_{N \rightarrow \infty} \frac{1}{N} M_x^\pi \sum_{k=1}^N w(x_k, y_k, x_{k+1}), \quad (3.1)$$

где $w(x_k, y_k, x_{k+1})$ — непосредственный выигрыш.

Положим $|w(x_k, y_k, x_{k+1})| \leq c < \infty$. Тогда, обозначив $R = \sup_{x, y, z} |w(x, y, z)|$, получаем $\varphi(x, \pi) \leq R < \infty$, откуда следует $\varphi(x, \pi) \in L(A^-, A^+)$. Но тогда в силу теоремы 2.4 существует оптимальная нерандомизированная марковская стратегия π^* такая, что $\varphi(x, \pi^*) = \sup_{\pi} \varphi(x, \pi)$.

Рассмотрим ожидаемый выигрыш $v_n(x, \pi)$ за n шагов при стратегии π и начальном состоянии $x \in E$:

$$v_n(x, \pi) = M_x^\pi \sum_{k=1}^n w(x_k, y_k, x_{k+1}) = \sum_{k=1}^n T^k w(x).$$

Как показано в [2], для $v_n(x, \pi)$ справедливо представление $v_n(x, \pi) = [F^n 0](x)$.

Теорема 3.1. Пусть для любого $n \geq 1$ справедливо неравенство

$$F(f) v_{n-1}(x, \pi) \geq v_n(x, \pi).$$

тогда

$$v_n(x, f) \geq v_n(x, \pi),$$

где π и f — стационарные стратегии.

Доказательство. По условию имеем

$$v_n(x, \pi) = F(\pi) v_{n-1}(x, \pi) \leq F(f) v_{n-1}(x, \pi).$$

Поскольку $n \geq 1$ произвольно, то это неравенство справедливо для любого $n \geq 1$. Тогда, воспользовавшись соотношением $v_n(x, \pi) = [F^n(\pi) 0](x)$ и полагая $F^0 \equiv 1$, получаем

$$\begin{aligned} v_n(x, \pi) &= F(\pi) [F^{n-1}(\pi) 0](x) \leq F(f) [F^{n-1}(\pi) 0](x) \leq \\ &\leq F(f) [F(f) F^{n-2}(\pi) 0](x) = F^2(f) [F^{n-2}(\pi) 0](x) \leq \\ &\leq \dots \leq F^n(f) [F^0(\pi) 0](x) = [F^n(f) 0](x) = v_n(x, f). \end{aligned}$$

Введем теперь оператор Q :

$$Qv_n(x, \pi) = \max \{v_n(x, \pi), F(f)v_{n-1}(x, \pi)\}. \quad (3.2)$$

С результатом применения оператора Q к $v_n(x, \pi)$ свяжем стратегию σ :

$$\sigma = \begin{cases} \pi, & \text{если } v_n(x, \pi) > F(f)v_{n-1}(x, \pi), \\ f, & \text{если } F(f)v_{n-1}(x, \pi) \geq v_n(x, \pi). \end{cases} \quad (3.3)$$

Теорема 3.2. Пусть π — произвольная стратегия, σ — стратегия, порожденная оператором Q , и $v_n(x, \sigma)$ — выигрыш, соответствующий стратегии σ . Тогда

$$v_n(x, \sigma) \geq Qv_n(x, \pi), \quad n = 1, 2, \dots$$

Доказательство. Аналогично доказательству теоремы 2.2.

Следствие 3.1. Пусть π — произвольная стратегия, σ — стратегия, порожденная оператором Q , и пусть существуют пределы

$$\lim_{N \rightarrow \infty} \frac{1}{N} v_N(x, \sigma) = \varphi(x, \sigma) \quad \text{и} \quad \lim_{n \rightarrow \infty} \frac{1}{n} v_n(x, \pi) = \varphi(x, \pi). \quad \text{Тогда}$$

$$\varphi(x, \sigma) \geq \varphi(x, \pi).$$

Доказательство. Из неравенства $v_n(x, \sigma) \geq v_n(x, \pi)$, $n = 1, 2, \dots$ следует $\frac{1}{n} v_n(x, \sigma) \geq \frac{1}{n} v_n(x, \pi)$, $n = 1, 2, \dots$. Поскольку это неравенство справедливо при любом $n = 1, 2, \dots$, то, переходя к пределу при $n \rightarrow \infty$, получаем требуемое неравенство.

Пусть P — переходная матрица управляемого процесса, образующего эргодическую марковскую цепь.

Относительно матрицы P справедливы утверждения, содержащиеся в лемме 3.1, доказательство которых можно найти в [4, 5].

Лемма 3.1. Пусть $P = P(\pi)$ — переходная матрица однородной эргодической цепи Маркова. Тогда а) P суммируема по Чезаро

$$\text{к матрице } A, \text{ т. е. } A = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P^k; \quad (3.4)$$

в) матрица $A = A(\pi)$ обладает свойством

$$PA = AP = AA = A \quad (3.5)$$

и является матрицей финальных вероятностей состояний, по строкам которой стоит один и тот же вектор $\alpha(\pi) = \{\alpha_j(\pi), j = 1, \dots, N\}$, т. е. $A(\pi) = e\alpha(\pi)$, где e — N -мерный вектор-столбец, все элементы которого равны 1, N — число состояний в множестве E .

Вектор $\alpha(\pi)$ не зависит от начального распределения вероятностей состояний и является единственным решением системы

$$\begin{aligned}\alpha &= \alpha P, \\ \alpha_j &\geq 0, \quad j \in E \\ \alpha e &= 1.\end{aligned}\tag{3.6}$$

Определим теперь N -мерный вектор непосредственных выигрышей $\omega = \omega(\pi)$, компонентами которого являются

$$\omega(i, \pi(i)) = \sum_j p_{ij}(\pi(i)) \omega_{ij}(\pi(i)), \tag{3.7}$$

где p_{ij} — элементы матрицы $P(\pi)$, $\pi(i) = y \in Y$. Кроме того, пусть $\varphi(\pi)$ — N -мерный вектор, компонентами которого являются средние выигрыши $\varphi(i, \pi)$, определенные в (3.1).

С учетом свойства (3.4) имеем

$$\begin{aligned}\varphi(\pi) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P^k \omega(\pi) = \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P^k \omega(\pi) = A(\pi) \omega(\pi).\end{aligned}\tag{3.8}$$

Компоненты вектора $\varphi(\pi)$ имеют вид $\varphi(i, \pi) = \alpha(\pi) \omega(\pi)$. Но вектор $\alpha(\pi)$ не зависит от начальных состояний. Поэтому $\varphi(i, \pi)$ также не зависит от $i \in E$, т. е. $\varphi(i, \pi) = g$, где $g = g(\pi)$ — константа. Тогда вектор $\varphi(\pi)$ может быть записан так: $\varphi(\pi) = ge$.

Лемма 3.2. Пусть π — некоторая стратегия и σ — стратегия, порожденная оператором Q . Тогда

$$g(\sigma) e \geq g(\pi) e. \tag{3.9}$$

Доказательство. Замечая, что $\varphi(x, \sigma) = g(\sigma)$ и $\varphi(x, \pi) = g(\pi)$, из следствия 3.1 получаем $g(\sigma) \geq g(\pi)$ или в векторном виде

$$g(\sigma) e \geq g(\pi) e. \tag{3.10}$$

Теорема 3.3. Пусть π — произвольная стратегия и σ — стратегия, порожденная оператором Q . Тогда необходимым и достаточным условием выполнения неравенства (3.9) является выполнение неравенства

$$P(\sigma) \omega(\sigma) \geq P(\pi) \omega(\pi). \tag{3.11}$$

Доказательство. Покажем вначале, что для выполнения неравенства (3.9) необходимо выполнение неравенства (3.10). Для этого достаточно показать, что для выполнения неравенства (3.9) необходимо выполнение неравенства

$$P^n(\sigma) \omega(\sigma) \geq P^n(\pi) \omega(\pi) \text{ для всех } n = 1, 2, \dots, \tag{3.12}$$

откуда при $n = 1$ следует неравенство (3.10).

Предположим противное, т.е. что существует некоторое $n = N$ такое, что

$$P^N(\sigma) \omega(\sigma) \geq P^N(\pi) \omega(\pi), \quad (3.13)$$

но

$$P^{N+m}(\sigma) \omega(\sigma) \leq P^{N+m}(\pi) \omega(\pi), \quad m = 1, 2, \dots \quad (3.14)$$

или

$$P^k(\sigma) \omega(\sigma) \leq P^k(\pi) \omega(\pi) \text{ для всех } k > N.$$

Тогда

$$\sum_{k=N+1}^{N+m} P^k(\sigma) \omega(\sigma) \leq \sum_{k=N+1}^{N+m} P^k(\pi) \omega(\pi), \quad m = 1, 2, \dots$$

или

$$P^N(\sigma) \sum_{k=1}^m P^k(\sigma) \omega(\sigma) \leq P^N(\pi) \sum_{k=1}^m P^k(\pi) \omega(\pi).$$

Неравенство не нарушится, если его обе части разделить на m , т. е.

$$\frac{1}{m} P^N(\sigma) \sum_{k=1}^m P^k(\sigma) \omega(\sigma) \leq \frac{1}{m} P^N(\pi) \sum_{k=1}^m P^k(\pi) \omega(\pi).$$

Отсюда

$$\lim_{m \rightarrow \infty} \frac{1}{m} P^N(\sigma) \sum_{k=1}^m P^k(\sigma) \omega(\sigma) \leq \lim_{m \rightarrow \infty} \frac{1}{m} P^N(\pi) \sum_{k=1}^m P^k(\pi) \omega(\pi).$$

С учетом свойств (3.4) и (3.5) получаем

$$P^N(\sigma) A(\sigma) \omega(\sigma) \leq P^N(\pi) A(\pi) \omega(\pi), \\ A(\sigma) \omega(\sigma) \leq A(\pi) \omega(\pi)$$

или

$$g(\sigma) \omega \leq g(\pi) \omega,$$

что противоречит неравенству (3.9).

Таким образом, если предположить неравенства (3.13) и (3.14), то неравенство (3.9) невозможно. Отсюда следует, что для выполнения неравенства (3.9) необходимо выполнение неравенства (3.12) для всех $n = 1, 2, \dots$ и тем самым неравенства (3.10). Пусть теперь для стратегий σ и π имеет место неравенство (3.9), т. е.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P^k(\sigma) \omega(\sigma) \geq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n P^k(\pi) \omega(\pi).$$

Обозначая

$$v_n = \sum_{k=1}^n P^k \omega,$$

последнее неравенство запишется в виде

$$\lim_{n \rightarrow \infty} \frac{1}{n} v_n(\sigma) \geq \lim_{n \rightarrow \infty} \frac{1}{n} v_n(\pi).$$

Отсюда и с учетом теоремы 3.2 получаем

$$v_n(\sigma) \geq v_n(\pi) \text{ для всех } n = 1, 2, \dots,$$

т. е.

$$\sum_{k=1}^n P^k(\sigma) \omega(\sigma) \geq \sum_{k=1}^n P^k(\pi) \omega(\pi)$$

для всех $n = 1, 2, \dots$. Отсюда

$$P(\sigma) \omega(\sigma) \geq P(\pi) \omega(\pi),$$

что и требовалось показать.

Следствие 3.2. Пусть π — произвольная стратегия и σ — стратегия, порожденная оператором Q . Тогда

$$P(\sigma) \omega(\sigma) \geq P(\pi) \omega(\pi).$$

Доказательство. Непосредственно вытекает из леммы 3.2 и теоремы 3.3.

Последнее следствие позволяет построить простой итеративный алгоритм поиска оптимальной стратегии.

Обозначим Y_i : множество допустимых управлений в состоянии $i \in E$:

Пусть дана некоторая стратегия π . Введем тогда множество управлений $G_i(\pi)$, определенное следующим образом:

$$G_i(\pi) = \{y \in Y_i : \sum_j [p_{ij}(y) - p_{ij}(y(\pi))] \omega_j(y(\pi)) > 0\},$$

где

$$y(\pi) = \pi(i) \in Y_i.$$

Лемма 3.3. Если $G_i(\pi) \neq \emptyset$, то стратегия π может быть улучшена выбором управлений из множества $G_i(\pi)$.

Доказательство. Пусть σ — стратегия, составленная из управлений $y \in G_i(\pi)$. В силу условия

$$\sum_j [p_{ij}(y) - p_{ij}(y(\pi))] \omega_j(y(\pi)) > 0. \quad (3.15)$$

переходные вероятности $p_{ij}(y(\sigma))$ при стратегии σ «лучше», чем $p_{ij}(y(\pi))$ при стратегии π . Если при этом выполняется условие

$$\sum_j p_{ij}(y(\sigma)) \omega_j(y(\sigma)) > \sum_j p_{ij}(y(\pi)) \omega_j(y(\pi)), \quad i \in F, \quad (3.16)$$

то в силу теоремы 3.3 выполняется неравенство $g(\sigma)e \geq g(\pi)e$ и тем самым стратегия σ является улучшенной, что и доказывает лемму.

Пусть $\sigma^{(n)}$ — последовательность стратегий, которые составлены из управлений, выбираемых из соответствующих множеств

$G_i(\sigma^{n-1})$ и удовлетворяющих неравенству (3.16). Тогда в силу теоремы 3.3 и следствия 3.2 стратегию $\sigma^{(n)}$ можно полагать порожденной n -й степенью оператора Q .

Теорема 3.4. Если для некоторой стратегии σ^* из последовательности $\sigma^{(n)}$, $n = 1, 2, \dots$ множество $G_i(\sigma^*) = \emptyset$ для всех $i \in E$ либо если $G_i(\sigma^*) = \emptyset$, но

$$\sum_i p_{i1}(y^*) w_1(y^*) < \sum_i p_{i1}(\sigma^*) w_1(\sigma^*)$$

для всех

$$y^* \in G_i(\sigma^*), \quad i \in E.$$

то стратегия σ^* оптимальна.

Доказательство. Выберем в качестве начальной стратегии для последовательности $\sigma^{(n)}$ стратегию π такую, что

$$w_i(y(\pi)) = \max_{y \in Y_i} w(y). \quad (3.17)$$

Очевидно, эта стратегия может быть улучшена только лишь за счет улучшения переходных вероятностей, т. е. в том случае, если $G_i(\pi) \neq \emptyset$. Если же $G_i(\pi) = \emptyset$ для всех $i \in E$, то стратегия π не может быть улучшена, следовательно, она оптимальна.

Кроме того, возможен случай такой, что $G_i(\pi) \neq \emptyset$, но

$$\sum_i p_{i1}(y) w_1(y) < \sum_i p_{i1}(\pi) w_1(\pi) \quad \forall y \in G_i(\pi), \quad i \in E. \quad (3.18)$$

Это означает, что условие (3.10) теоремы 3.3 не может быть выполнено. Тогда стратегия π не может быть улучшена, следовательно, она оптимальна.

Таким образом, если в качестве начальной стратегии выбрана стратегия π , удовлетворяющая условию (3.17), и производится последовательное улучшение стратегии за счет выбора управлений из множеств $G_i(\sigma^{(n)})$, то на каждой итерации улучшение стратегии возможно только за счет улучшения переходных вероятностей.

Если теперь на некоторой итерации улучшения окажется, что либо $G_i(\sigma^*) = \emptyset$ для всех $i \in E$, либо $G_i(\sigma^*) \neq \emptyset$, но выполнено неравенство (3.18) для всех $y \in G_i(\sigma^*)$, то стратегия σ^* оптимальна. Теорема доказана.

Из доказанной теоремы вытекает следующий итеративный алгоритм поиска оптимальной стратегии.

1. Выбрать в качестве начального приближения стратегию π такую, что для всех $i \in E$ выполняется условие (3.17), т. е.

$$w_i(y(\pi)) = \max_{y \in Y_i} w(y).$$

2. Для каждого $i \in E$ выбрать управление

$$y^* \in G_i(\pi) : \sum_i [p_{ij}(y^*) - p_{ij}(y(\pi))] w_j(y(\pi)) > 0.$$

3. Из выбранных управлений $y^* \in G_i(\pi)$ образовать стратегию $\sigma: \{i \rightarrow y_i^*, i \in E\}$.

4. Проверить условие (3.16)

$$\sum_I p_{ij}(\sigma) w_j(\sigma) > \sum_I p_{ij}(\pi) w_j(\pi), i \in E.$$

Если да, то σ использовать в качестве нового приближения и перейти к п. 2.

Если нет, то начальное приближение оставить прежним и перейти снова к п. 2. При этом управления $y_i^* \in G_i(\pi)$, не удовлетворяющие условию (3.16), из множества Y_i исключить, если $G_i(\pi) \neq \emptyset$.

5. Если $G_i(\pi) = \emptyset$ для всех $i \in E$, то останов.

Теорема 5. *Описанный алгоритм за конечное число шагов сходится к оптимальной стратегии.*

Доказательство. В силу условий (3.15) и (3.16) на каждой итерации имеет место строгое улучшение стратегии. Поскольку множества состояний E и управлений Y конечны, то множество стратегий также конечно. Тогда за конечное число итераций будет достигнута оптимальная стратегия.

Пример. Пусть имеется некоторая техническая система, которая может находиться в двух состояниях: исправном (1) и в состоянии отказа (2). Пусть установлен некоторый период времени, в течение которого исправная система работает и приносит доход 3 руб. При этом в конце периода система остается в исправном состоянии с вероятностью 0,7 и вероятностью 0,3 переходит в состояние отказа.

Отказавшая система может быть отремонтирована двумя способами. Первый способ (1) переводит в исправное состояние с вероятностью 0,6, при этом начисляется штраф 2 руб. (доход равен (-2) руб.). Второй способ (2) переводит в исправное состояние с вероятностью 0,4, при этом начисляется штраф 1 руб. (доход (-1) руб.). Требуется отыскать оптимальную стратегию.

В условиях примера имеется две стратегии f_1 и f_2 , которые отличаются способами ремонта. Для этих стратегий переходные матрицы и векторы выигрышей имеют вид

$$Q(f_1) = \begin{bmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{bmatrix} \quad w(f_1) = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

$$Q(f_2) = \begin{bmatrix} 0,7 & 0,3 \\ 0,4 & 0,6 \end{bmatrix} \quad w(f_2) = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Согласно описанному алгоритму в качестве начального приближения, удовлетворяющего условию (18), выберем стратегию f_2 .

Выполняя второй шаг алгоритма, находим, что $G_1(f_2) = \emptyset$, а $G_2(f_2) \neq \emptyset$ и содержит способ (1), при котором

$$\sum_I [p_{2j}(1) - p_{2j}(2)] w_j(f_2) = (0,6 - 0,4) 3 + (0,4 - 0,6) (-1) > 0.$$

Выполняя третий шаг алгоритма, получаем стратегию $\sigma = f_1$.
Выполняя четвертый шаг алгоритма, находим, что для состояния 1 условие (17) не выполняется, т. е.

$$0,7 \cdot 3 + 0,3(-2) < 0,7 \cdot 3 + 0,3 \cdot (-1).$$

Исключая тогда первый способ из множества допустимых действий Y_2 , получаем $G_i(f_2) = \emptyset$ для всех $i = 1, 2$. Тогда стратегия f_2 является оптимальной.

ЛИТЕРАТУРА

1. Баранов В. В. Обобщенные управляемые процессы.— «Тезисы докладов на III Всесоюзном совещании по статистическим методам в процессах управления», 1973, с. 141—143.
2. Баранов В. В. О свойствах выигрыша в управляемых процессах.— См статью в настоящем вестнике.
3. Ховард Р. А. Динамическое программирование и марковские процессы. М., «Сов. радио», 1964. 95 с.
4. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М., «Наука», 1970, с. 93—104.
5. Чжун Кай-Лай. Однородные цепи Маркова. М., «Мир», 1964, с. 50—59.
6. Manne A. S. Linear programming and sequential decisions.— «Management Science», 1960, № 6, p. 259—267.

УДК 519.21 + 621.391

В. З. СОКОЛОВСКИЙ

ОБ ОПТИМАЛЬНЫХ ОЦЕНКАХ СЛУЧАЙНЫХ ПОЛЕЙ

1. Многие важные задачи теории управления носят стохастический характер. Рассмотрение случайных уравнений вместо детерминированных вызвано целым рядом причин.

Одна из них состоит в том, что поведение реального объекта, функционирующего в условиях естественных, промышленных и других «шумов», характеризуется некоторой неопределенностью. Кроме того, участие человека в системе управления также вносит элемент неоднозначности. Описание таких систем с помощью детерминированных уравнений не всегда адекватно и часто не отражает действительной картины функционирования объекта.

Существенной частью теории управления являются задачи фильтрации, представляющие предмет изучения данной работы.

Теория оценивания случайных сигналов, возникающих в системах с сосредоточенными параметрами, интенсивно развивается. Однако производственные процессы обычно описываются системами, имеющими распределенные в пространстве параметры.

Кроме того, некоторые сложные системы с сосредоточенными параметрами описываются обыкновенными дифференциальными уравнениями столь высокого порядка, что их исследование вызывает большие затруднения. Иногда такие системы целесообразно аппроксимировать системами с распределенными параметрами.

Поэтому возникла необходимость в дальнейшем развитии теории оптимальных систем и ее обобщении на случай систем с распределенными параметрами. Основные результаты в этой новой области приведены в книгах Бутковского [1] и Лионса [2].

Между тем теория фильтрации для систем с распределенными параметрами развивалась слабо. В данной статье развиваются и уточняются общие положения и теоремы работы Бен-Суссана [3].

2. Задачи прогноза и фильтрации стационарных случайных процессов впервые были рассмотрены Колмогоровым [4] (см. также [5]) для систем с дискретным временем и Винером [6] для систем с непрерывным временем.

Теория Колмогорова — Винера основывается на сведении задач оценивания стационарных случайных процессов и последовательностей к так называемому интегральному уравнению Винера — Хопфа. Существуют различные приближенные и точные методы решения этого уравнения.

В дальнейшем теория Колмогорова — Винера интенсивно развивалась и обобщалась. Отметим, в частности, работы [7—10]. Общей для всех этих работ была конечная цель: определить линейную динамическую систему (фильтр), которая осуществляла бы прогноз, фильтрацию или детектирование случайного сигнала.

В 1960 году Калман указал на ряд недостатков этого подхода, которые существенно ограничивают возможности его практического применения. К этим недостаткам относятся следующие:

1) оптимальный фильтр определяется своей передаточной функцией, а восстановить по ней фильтр достаточно сложно;

2) численное решение возникающих задач мало приспособлено для решения на ЭВМ. При этом трудность решения резко возрастает с увеличением сложности задачи;

3) методы, развитые Колмогоровым и Винером, не применимы к задачам нестационарной фильтрации.

В той же работе [11] Калман предложил новый подход к задачам линейной фильтрации и прогноза, частично свободный от указанных недостатков. При этом рассматривались, главным образом, системы с дискретным временем. Годом позже Калман и Бьюси [12] перенесли полученные результаты на системы с непрерывным временем.

Одним из центральных пунктов теории Калмана — Бьюси явилось установление двойственности между задачами оценивания для стохастических систем и задачами управления детерминированными системами.

Фильтры Калмана интенсивно изучаются до настоящего времени. Подробно исследованы условия устойчивости решений матричного уравнения Риккати. Получен и ряд обобщений теории Калмана.

Так, Бьюси [14] распространил теорию на случай входных шумов, образуемых в результате прохождения белого шума через линейную нестационарную систему. Фалб [15] перенес результаты Калмана — Бьюси, полученные для конечномерного случая, на гильбертовы пространства.

Дальнейшие ссылки можно найти в работах [16—17].

3. Рассмотрим задачи фильтрации для случайных полей, описываемых уравнениями в частных производных первого порядка. На системы такого вида теория Калмана переносится наиболее полным образом.

Уравнениями в частных производных первого порядка описываются различные технологические процессы, например, теплообмен в химическом реакторе или в проходной печи в случае, когда нагреваемый материал можно считать тонким телом в тепло-техническом отношении.

Так, при определенных условиях, температура $Q(x, t)$ в циклическом реакторе с неподвижным слоем катализатора описывается уравнением

$$\frac{\partial Q}{\partial t} + v \frac{\partial Q}{\partial x} = -h \exp(\alpha + \beta Q), \quad (3.1)$$

где x — координата точки, в которой измеряется температура; t — момент времени; h — теплота реакции; α , β и v — некие параметры [1]. Для построения системы автоматического управления процессом в реакторе необходимо знать текущие значения температуры. В случае, когда в уравнении (1) в правой части добавляются случайные воздействия и измерения температуры сопровождаются случайными ошибками, естественным образом возникает задача фильтрации.

Другой пример. Если предположить, что материал, который нагревается в проходной печи, является «тонким телом», то распределение температуры материала будет описываться функцией

$$Q(x, t), \quad 0 \leq x \leq L, \quad 0 \leq t \leq T,$$

которая удовлетворяет уравнению

$$b \frac{\partial Q}{\partial t} + bv \frac{\partial Q}{\partial x} + Q = u(x, t)$$

и граничному условию при входе в греющую зону

$$Q(0, t) = Q_1(t).$$

Здесь $u(x, t)$ — распределение температуры в неподвижной греющей среде; $v(t)$ — скорость движения материала; $Q_1(t)$ — функция, характеризующая изменение во времени температуры материала, входящего в зону обработки. Функция $b(x, t)$ характеризует теплофизические параметры материала. Задача состоит в выборе такого управляющего воздействия $u(x, t)$, чтобы материал на выходе из печи имел требуемую температуру [1]. И в этом

случае при переходе к стохастической модели естественно возникают задачи фильтрации.

Задача оценивания для полей, описываемых уравнениями в частных производных первого порядка, ставится следующим образом.

Пусть искомое случайное поле $u(x, t)$ генерируется системой, описываемой уравнением

$$\frac{\partial u}{\partial t} = a(x, t) \frac{\partial u}{\partial x} + b(x, t) u + c(x, t) + m(x, t) \quad (3.2a)$$

с начальным условием

$$u(x, 0) = f(x). \quad (3.2b)$$

Вследствие наличия помех в измерительном устройстве случайное поле $u(x, t)$ определяется с ошибкой, и на выходе измерительного прибора наблюдается величина $w(x, t)$, связанная с полем $u(x, t)$ соотношением

$$w(x, t) = u(x, t) + n(x, t). \quad (3.3)$$

Необходимо по наблюдаемым значениям $w(x, t)$ построить несмещенную и эффективную оценку $\hat{u}(x, t)$ поля $u(x, t)$.

Относительно функций, входящих в соотношения (2) и (3), сделаем следующие предположения. Функции a , b и c — детерминированы, причем $a \neq 0$. Случайные поля m , n и f предполагаются взаимно не коррелированными

$$\langle m(x, t) n(\xi, \tau) \rangle = \langle m(x, t) f(\xi) \rangle = \langle n(x, t) f(\xi) \rangle = 0.$$

Кроме того, потребуем, чтобы

$$\langle m(x, t) \rangle = \langle n(x, t) \rangle = 0,$$

$$\langle m(x, t) m(\xi, \tau) \rangle = M(x, \xi, t) \delta(t - \tau),$$

$$\langle n(x, t) n(\xi, \tau) \rangle = N(x, \xi, t) \delta(t - \tau),$$

$$\langle f^2(x) \rangle < \infty, \quad \langle f(x) \rangle = F(x),$$

где $M(x, \xi, t)$, $N(x, \xi, t)$, $F(x)$ — непрерывно дифференцируемые функции, причем M — неотрицательна, а N — всюду положительна.

Следуя идее Калмана, представим оценку $\hat{u}(x, t)$ случайного поля $u(x, t)$ как решение уравнения

$$\frac{\partial \hat{u}}{\partial t} = a(x, t) \frac{\partial \hat{u}}{\partial x} + a(x, t) \hat{u} + \beta(x, t) + \gamma(x, t) w(x, t) \quad (3.4a)$$

при начальном условии

$$\hat{u}(x, 0) = f(x). \quad (3.4b)$$

Неизвестные детерминированные функции α , β и γ нужно определить таким образом, чтобы ошибка оценивания

$$\varepsilon(x, t) = u(x, t) - \hat{u}(x, t) \quad (3.5)$$

удовлетворяла условиям несмещенности

$$\langle \varepsilon(x, t) \rangle = 0$$

и эффективности

$$\langle \varepsilon^2(x, t) \rangle = \min.$$

Вычитая из уравнения (2а) уравнение (4а), получим

$$\begin{aligned} \frac{\partial \varepsilon}{\partial t} = & a(x, t) \frac{\partial \varepsilon}{\partial x} + [b(x, t) - \alpha(x, t) - \gamma(x, t)] u + \alpha(x, t) \varepsilon + \\ & + c(x, t) - \beta(x, t) + m(x, t) - \gamma(x, t) n(x, t). \end{aligned}$$

Возьмем математические ожидания от обеих частей этого равенства. Поскольку шумы имеют нулевые средние, получим

$$\begin{aligned} \frac{\partial \langle \varepsilon \rangle}{\partial t} = & a(x, t) \frac{\partial \langle \varepsilon \rangle}{\partial x} + [b(x, t) - \alpha(x, t) - \gamma(x, t)] \times \\ & \times \langle u \rangle + \alpha(x, t) \langle \varepsilon \rangle + c(x, t) - \beta(x, t). \end{aligned} \quad (3.6)$$

Отсюда видно, что для выполнения условия несмещенности оценки достаточно потребовать, чтобы

$$\alpha(x, t) = b(x, t) - \gamma(x, t)$$

и

$$\beta(x, t) = c(x, t).$$

Легко видеть, что в предположении о том, что

$$\langle u(x, t) \rangle \neq 0,$$

условия (5) и (6) не только достаточны для несмещенности, но и необходимы.

В силу (5) и (6) уравнение для ошибки можно переписать следующим образом:

$$\frac{\partial \varepsilon}{\partial t} = a(x, t) \frac{\partial \varepsilon}{\partial x} + [b(x, t) - \gamma(x, t)] \varepsilon + m(x, t) - \gamma(x, t) n(x, t). \quad (3.7)$$

Применим оператор

$$\frac{\partial}{\partial t} - a(x, t) \frac{\partial}{\partial x}$$

к функции $\varepsilon^2(x, t)$. В силу (7) получим, что

$$\frac{\partial \varepsilon^2}{\partial t} - a(x, t) \frac{\partial \varepsilon^2}{\partial x} = 2\varepsilon \left[\frac{\partial \varepsilon}{\partial t} - a(x, t) \frac{\partial \varepsilon}{\partial x} \right] = \quad (3.8)$$

$$= 2[b(x, t) - \gamma(x, t)] \varepsilon^2 + 2m(x, t) \varepsilon - 2\gamma(x, t) n(x, t) \varepsilon.$$

Введем обозначение

$$r(x, t) = \langle \varepsilon^2(x, t) \rangle$$

перейдем в равенстве (8) к математическим ожиданиям. Полу-

$$\frac{\partial r}{\partial t} - a(x, t) \frac{\partial r}{\partial x} = 2 [b(x, t) - \gamma(x, t)] r + \\ + 2 \langle m(x, t) \varepsilon(x, t) \rangle - 2 \gamma(x, t) \langle n(x, t) \varepsilon(x, t) \rangle. \quad (3.9)$$

Введем понятие об интегральном ядре («функции Грина») для уравнения

$$\frac{\partial y}{\partial t} = a(x, t) \frac{\partial y}{\partial x} + d(x, t) y.$$

то такая функция $G(x, t; \xi, \tau)$, что решение задачи

$$\frac{\partial y}{\partial t} = a(x, t) \frac{\partial y}{\partial x} + d(x, t) y + h(x, t), \quad y(x, 0) = f(x)$$

представимо в виде

$$y(x, t) = \int_0^t \int_{-\infty}^{\infty} G(x, t; \xi, \tau) h(\xi, \tau) d\xi d\tau + \int_{-\infty}^{\infty} G(x, t; \xi, 0) f(\xi) d\xi. \quad (3.10)$$

Возможность такого представления устанавливается следующим рассуждением. Рассмотрим сперва однородное уравнение

$$\frac{\partial \tilde{y}(x, t, \tau)}{\partial t} - a(x, t) \frac{\partial \tilde{y}(x, t, \tau)}{\partial x} - d(x, t) \tilde{y}(x, t, \tau) = 0$$

при условии

$$\tilde{y}(x, t, \tau)|_{t=\tau} = h(x, \tau).$$

Если ввести формальное представление

$$y(x, t, \tau) = \int_{-\infty}^{\infty} G(x, t; \xi, \tau) h(\xi, \tau) d\xi,$$

то функция $G(x, t; \xi, \tau)$ должна удовлетворять уравнению

$$\frac{\partial G(x, t; \xi, \tau)}{\partial t} = a(x, t) \frac{\partial G(x, t; \xi, \tau)}{\partial x} + d(x, t) G(x, t; \xi, \tau) \quad (3.11)$$

условию

$$G(x, t; \xi, \tau)|_{t=\tau} = \delta(x - \xi). \quad (3.12)$$

Решение задачи (11) — (12) существует при достаточной гладкости коэффициентов уравнения.

Легко видеть, что решение неоднородного уравнения

$$\frac{\partial y(x, t)}{\partial t} - a(x, t) \frac{\partial y(x, t)}{\partial x} - d(x, t) y(x, t) = h(x, t)$$

с нулевым начальным условием

$$y|_{t=0} = 0$$

будет таким:

$$y(x, t) = \int_0^t \tilde{y}(x, t, \tau) d\tau = \int_0^t \int_{-\infty}^{\infty} G(x, t; \xi, \tau) h(\xi, \tau) d\xi d\tau.$$

То, что $y(x, t)$ удовлетворяет начальному условию, очевидно. Подстановка в правую часть уравнения дает

$$\int_0^t \int_{-\infty}^{\infty} \left\{ \frac{\partial G(x, t; \xi, \tau)}{\partial t} - a(x, t) \frac{\partial G(x, t; \xi, \tau)}{\partial x} - d(x, t) G(x, t; \xi, \tau) \right\} \times \\ \times h(\xi, \tau) d\xi d\tau + \int_{-\infty}^{\infty} G(x, t; \xi, t) h(\xi, t) d\xi.$$

Но первый член равен нулю в силу (11), а второй равен $h(x, t)$ в силу (12).

Следовательно, представление (10) возможно при достаточной гладкости коэффициентов уравнения.

Очевидно, что начальное условие для уравнения (7) будет нулевым. Поэтому решение можно представить в виде

$$\varepsilon(x, t) = \int_0^t \int_{-\infty}^{\infty} G_1(x, t; \xi, \tau) [m(\xi, \tau) - \gamma(\xi, \tau) n(\xi, \tau)] d\xi d\tau, \quad (3.13)$$

где $G_1(x, t; \xi, \tau)$ — интегральное ядро для уравнения

$$\frac{\partial \varepsilon}{\partial t} = a(x, t) \frac{\partial \varepsilon}{\partial x} + [b(x, t) - \gamma(x, t)] \varepsilon.$$

С помощью представления (13) легко вычислить корреляционные функции, входящие в (9). При этом используются корреляционные свойства полей m и n , а также свойство (12).

Имеем

$$\langle \varepsilon(x, t) m(x, t) \rangle = \int_0^t \int_{-\infty}^{\infty} G_1(x, t; \xi, \tau) \langle m(x, t) m(\xi, \tau) \rangle d\xi d\tau = \\ = \int_0^t \int_{-\infty}^{\infty} G_1(x, t; \xi, \tau) M(x, \xi, t) \delta(t - \tau) d\xi d\tau = \\ = \frac{1}{2} \int_{-\infty}^{\infty} G_1(x, t; \xi, t) M(x, \xi, t) d\xi = \\ = \frac{1}{2} \int_{-\infty}^{\infty} \delta(x - \xi) M(x, \xi, t) d\xi = \frac{1}{2} M(x, x, t).$$

Аналогично находим, что

$$\langle n(x, t) \varepsilon(x, t) \rangle = -\frac{1}{2} \gamma(x, t) N(x, x, t).$$

Подставляя полученные значения корреляционных функций в (9), получим следующее уравнение:

$$\frac{\partial r}{\partial t} - a \frac{\partial r}{\partial x} = 2[b - \gamma]r + M(x, x, t) + \gamma^2 N(x, x, t). \quad (3.14)$$

В это уравнение входят две неизвестные функции $r(x, t)$ и $\gamma(x, t)$, и следующим этапом будет установление еще одной связи между этими функциями. Для этого используется условие эффективности оценки и вытекающее из него уравнение Винера — Хопфа.

Случайное поле $\hat{u}(x, t)$ в силу (4) представимо в виде

$$\hat{u}(x, t) = \int_0^t \int_{-\infty}^{\infty} G_2(x, t; \xi, \tau) \gamma(\xi, \tau) \omega(\xi, \tau) d\xi d\tau + \int_{-\infty}^{\infty} G_2(x, t; \xi, 0) f(\xi) d\xi, \quad (3.15)$$

где $G_2(x, t; \xi, \tau)$ — интегральное ядро для уравнения

$$\frac{\partial \hat{u}}{\partial t} = a(x, t) \frac{\partial \hat{u}}{\partial x} + [b(x, t) - \gamma(x, t)] \hat{u} + c(x, t). \quad (3.16)$$

Если же выбрать не оптимальный коэффициент $\gamma(x, t)$, а произвольный $\tilde{\gamma}(x, t)$, то получим оценку

$$z(x, t) = \int_0^t \int_{-\infty}^{\infty} \tilde{G}(x, t; \xi, \tau) \tilde{\gamma}(\xi, \tau) \omega(\xi, \tau) d\xi d\tau + \int_{-\infty}^{\infty} \tilde{G}(x, t; \xi, 0) f(\xi) d\xi \quad (3.17)$$

(интегральное ядро также изменится, поскольку оно зависит от $\gamma(x, t)$).

Для вывода уравнение Винера — Хопфа воспользуемся теоремой об ортогональном проектировании [18].

Введем гильбертово пространство H случайных полей $u(x, t)$, для которых

$$\langle u^2(x, t) \rangle < \infty.$$

Скалярное произведение в этом пространстве задается так:

$$(u(x, t), v(x, t)) = \langle u(x, t) v(x, t) \rangle.$$

Метрика при этом будет среднеквадратической. Поэтому можно утверждать, что оценка $\hat{u}(x, t)$ будет эффективной тогда и только тогда, когда ошибка оценивания ортогональна к любой оценке

$$(u - \hat{u}, z) = 0.$$

Поскольку $u - \hat{u} = \varepsilon$, это соотношение примет вид

$$\int_0^t \int_{-\infty}^{\infty} \tilde{G}_2(x, t; \xi, \tau) \tilde{\gamma}(\xi, \tau) \langle \varepsilon(x, t) \omega(\xi, \tau) \rangle d\xi d\tau + \int_{-\infty}^{\infty} \tilde{G}_2(x, t; \xi, 0) \langle \varepsilon(x, t) f(\xi) \rangle d\xi = 0. \quad (3.18)$$

В силу произвольности $\tilde{\gamma}(x, t)$ функция $\tilde{G}_2$ тоже может меняться произвольным образом (в пределах определенного класса). Поэтому из соотношения (16) следует, что

$$\langle \varepsilon(x, t) \omega(\xi, \tau) \rangle = 0,$$

$$\langle \varepsilon(x, t) f(\xi) \rangle = 0.$$

Эти соотношения являются аналогами уравнениями Винера — Хопфа в классической теории.

Условие (17) перепишем в виде

$$\langle [u(x, t) - \hat{u}(x, t)][u(\xi, \tau) + n(\xi, \tau)] \rangle = 0. \quad (3.19)$$

Легко видеть, что

$$\langle u(x, t) n(\xi, \tau) \rangle = 0. \quad (3.20)$$

В самом деле,

$$\begin{aligned} u(x, t) = & \int_0^t \int_{-\infty}^{\infty} G_3(x, t; \xi, \tau) m(\xi, \tau) d\xi d\tau + \\ & + \int_{-\infty}^{\infty} G_3(x, t; \xi, 0) f(\xi) d\xi, \end{aligned}$$

где $G_3(x, t; \xi, \tau)$ — интегральное ядро для уравнения

$$\frac{\partial u}{\partial t} = a(x, t) \frac{\partial u}{\partial x} + b(x, t) u + c(x, t).$$

Утверждение (20) вытекает из полученного представления функции $u(x, t)$ и предположенной ортогональности n с m и f .

В силу (20) соотношение (19) можно записать так:

$$\langle \varepsilon(x, t) u(\xi, \tau) \rangle = \langle \hat{u}(x, t) n(\xi, \tau) \rangle.$$

Заменяя в этом выражении $\hat{u}$ его представлением (15), получаем

$$\begin{aligned} \langle \varepsilon(x, t) u(\xi, \tau) \rangle = & \int_{-\infty}^{\infty} G_2(x, t; \eta, 0) \langle f(\eta) n(\xi, \tau) \rangle d\eta + \\ & + \int_0^t \int_{-\infty}^{\infty} G_2(x, t; \eta, \theta) \gamma(\eta, \theta) \langle \omega(\eta, \theta) n(\xi, \tau) \rangle d\eta d\theta. \end{aligned}$$

Учитывая (20) и условия ортогональности, можно написать

$$\begin{aligned} \langle \varepsilon(x, t) u(\xi, \tau) \rangle = & \int_0^t \int_{-\infty}^{\infty} G_2(x, t; \eta, \theta) \gamma(\eta, \theta) \langle n(\eta, \theta) n(\xi, \\ & \tau) \rangle d\eta d\theta = \int_{-\infty}^{\infty} G_2(x, t; \eta, \tau) \gamma(\eta, \tau) N(\xi, \eta, \tau) d\eta. \end{aligned}$$

Перейдем в последнем соотношении к пределу при $\tau \rightarrow t$ и $\xi \rightarrow x$. Получим, что

$$\langle \varepsilon(x, t) u(x, t) \rangle = \int_{-\infty}^{\infty} G_2(x, t; \eta, t) \gamma(\eta, t) N(x, \eta, t) d\eta = \\ = \gamma(x, t) N(x, x, t). \quad (3.21)$$

Покажем теперь, что

$$\langle \varepsilon(x, t) \hat{u}(x, t) \rangle = 0. \quad (3.22)$$

Для этого достаточно умножить обе части (15) на $\varepsilon(x, t)$, перейти к математическим ожиданиям и воспользоваться равенствами (17)–(18).

Вычитая из (21) соотношение (20), найдем, что

$$\langle \varepsilon^2(x, t) \rangle = \gamma(x, t) N(x, x, t)$$

или с учетом ранее введенного обозначения

$$\gamma(x, t) = \frac{r(x, t)}{N(x, x, t)}. \quad (3.23)$$

Подставляя это соотношение в (14), получим следующее дифференциальное уравнение для определения r :

$$\frac{\partial r}{\partial t} - a(x, t) \frac{\partial r}{\partial x} = M(x, x, t) + 2b(x, t) r - \frac{r^2}{N(x, x, t)}. \quad (3.24a)$$

Уравнение это содержит квадратичную нелинейность точно так же, как и уравнение Риккати, которое возникает при определении коэффициентов фильтра Калмана.

Для решения уравнения (24) нужно задать начальное условие. Очевидно, что оно будет нулевым

$$r|_{t=0} = 0. \quad (3.24b)$$

Полученный результат можно сформулировать в виде следующей теоремы.

Теорема 1. *Случайное поле $\hat{u}(x, t)$, представляющее собой решение задачи (4), будет несмещенной и эффективной оценкой поля $u(x, t)$, которое является решением задачи (2), если коэффициенты α , β , и γ в уравнении (4a) определяются соотношениями (5), (6) и (23), в котором $r(x, t)$ есть решение задачи (24).*

Численное решение задачи (24) легко может быть осуществлено методом сеток.

4. Решение задач фильтрации для случайных полей, описываемых гиперболическими уравнениями, важно в связи с тем, что многие процессы, для которых создаются системы автоматического управления, описываются именно такими уравнениями. О связи между задачами управления и фильтрации речь шла выше.

Теорию Калмана не удается полностью перенести на системы, описываемые уравнениями в частных производных второго порядка. Удастся, однако, свести задачу оценивания к двойственной задаче детерминированного управления.

Для простоты изложения ограничимся обычным уравнением малых колебаний

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + m(x, t) \quad (4.1a)$$

с нулевыми краевыми условиями

$$u|_{x=a} = u|_{x=b} = 0 \quad (4.1b)$$

и начальными условиями

$$u|_{t=t_0} = f(x), \quad \left. \frac{\partial u}{\partial t} \right|_{t=t_0} = g(x). \quad (4.1c)$$

Наблюдаемое поле $w(x, t)$ связано с исследуемым $u(x, t)$ соотношением

$$w(x, t) = u(x, t) + n(x, t). \quad (4.2)$$

Требования на случайные поля $m(x, t)$, $n(x, t)$, $f(x)$ и $g(x)$ налагаются следующие:

$$\begin{aligned} \langle f(x) \rangle &= F(x), \quad \langle g(x) \rangle = G(x), \\ \langle [f(x) - F(x)][f(\xi) - F(\xi)] \rangle &= \Phi(x) \delta(x - \xi), \\ \langle m(x, t) \rangle &= \langle n(x, t) \rangle = 0, \\ \langle [g(x) - G(x)][g(\xi) - G(\xi)] \rangle &= \Psi(x) \delta(x - \xi), \\ \langle m(x, t)f(\xi) \rangle &= \langle m(x, t)g(\xi) \rangle = \langle n(x, t)f(\xi) \rangle = \\ &= \langle n(x, t)g(\xi) \rangle = \langle m(x, t)n(\xi, \tau) \rangle = \langle f(x)g(\xi) \rangle = 0, \\ \langle m(x, t)m(\xi, \tau) \rangle &= \mu(x, t) \delta(x - \xi) \delta(t - \tau), \\ \langle n(x, t)n(\xi, \tau) \rangle &= \nu(x, t) \delta(x - \xi) \delta(t - \tau), \end{aligned}$$

где $F(x)$, $G(x)$, $\Phi(x)$, $\Psi(x)$, $\mu(x, t)$ и $\nu(x, t)$ — непрерывно дифференцируемые функции, причем

$$\Phi(x) \geq 0, \quad \Psi(x) \geq 0, \quad \mu(x, t) \geq 0$$

и

$$\nu(x, t) > 0$$

во всей области определения.

Будем оценивать следующий функционал:

$$J[\alpha, \beta] = \int_a^b u(x, t_1) \alpha(x) dx - \int_a^b \frac{\partial u(x, t_1)}{\partial t} \beta(x) dx, \quad (4.3)$$

где $\alpha(x)$ и $\beta(x)$ — произвольные детермированные функции, заданные на сегменте $[a, b]$. Если нужно оценить значение $u(x_1, t_1)$, то достаточно выбрать $\alpha(x) = \delta(x - x_1)$, $\beta(x) = 0$. Если же оценивается величина

$$\frac{\partial u(x_1, t_1)}{\partial t},$$

выбираем $\alpha(x) = 0$, $\beta(x) = -\delta(x - x_1)$.

Наиболее общий вид линейных оценок следующий:

$$\begin{aligned} \hat{J}[\alpha, \beta] = & \int_{t_0}^{t_1} \int_a^b w(x, t) \lambda(x, t) dx dt + \\ & + \int_a^b \varphi(x) F(x) dx + \int_a^b x(x) G(x) dx. \end{aligned} \quad (4.4)$$

Задача состоит в том, чтобы выбрать детерминированные функции $\lambda(x, t)$, $\varphi(x)$ и $x(x)$ таким образом, чтобы оценка была несмещенной

$$\langle J[\alpha, \beta] - \hat{J}[\alpha, \beta] \rangle = 0 \quad (4.5)$$

и эффективной

$$\langle (J[\alpha, \beta] - \hat{J}[\alpha, \beta])^2 \rangle = \min. \quad (4.6)$$

Рассмотрим двойственную краевую задачу

$$\frac{\partial^2 z}{\partial t^2} = \frac{\partial^2 z}{\partial x^2} + \lambda(x, t), \quad (4.7a)$$

$$z|_{x=a} = z|_{x=b} = 0, \quad (4.7b)$$

$$z|_{t=t_1} = \beta(x), \quad \left. \frac{\partial z}{\partial t} \right|_{t=t_1} = \alpha(x). \quad (4.7c)$$

В силу (7c) функционал $J[\alpha, \beta]$ можно представить в виде

$$J[\alpha, \beta] = \int_a^b u(x, t_1) \frac{\partial z(x, t)}{\partial t} dx - \int_a^b \frac{\partial u(x, t_1)}{\partial t} z(x, t_1) dx.$$

Проведем ряд несложных преобразований:

$$\begin{aligned} J[\alpha, \beta] = & \int_a^b u(x, t_0) \frac{\partial z(x, t_0)}{\partial t} dx - \int_a^b \frac{\partial u(x, t_0)}{\partial t} z(x, t_0) dx + \\ & + \int_{t_0}^{t_1} d \left[\int_a^b u(x, t) \frac{\partial z(x, t)}{\partial t} dx \right] - \int_{t_0}^{t_1} d \left[\int_a^b \frac{\partial u(x, t)}{\partial x} z(x, t) dx \right]; \\ & d \left[\int_a^b u \frac{\partial z}{\partial t} dx \right] = \int_a^b u \frac{\partial^2 z}{\partial t^2} dx dt + \int_a^b \frac{\partial u}{\partial t} \frac{\partial z}{\partial t} dx dt, \\ & d \left[\int_a^b z \frac{\partial u}{\partial t} dx \right] = \int_a^b z \frac{\partial^2 u}{\partial t^2} dx dt + \int_a^b \frac{\partial z}{\partial t} \frac{\partial u}{\partial t} dx dt, \\ & d \left[\int_a^b \left(u \frac{\partial z}{\partial t} - z \frac{\partial u}{\partial t} \right) dx \right] = \int_a^b \left(u \frac{\partial^2 z}{\partial t^2} - z \frac{\partial^2 u}{\partial t^2} \right) dx dt = \end{aligned}$$

(в силу (7а) и (1а))

$$= \int_a^b \left\{ u \left[\frac{\partial^2 z}{\partial x^2} + \lambda(x, t) \right] - z \left[\frac{\partial^2 u}{\partial x^2} + m(x, t) \right] \right\} dx dt =$$

(после интегрирования по частям с учетом (7в) и (1в))

$$= \int_a^b [u(x, t) \lambda(x, t) - z(x, t) m(x, t)] dx dt.$$

Следовательно, учитывая (1с), получаем

$$\begin{aligned} J[\alpha, \beta] &= \int_a^b f(x) \frac{\partial z(x, t_0)}{\partial t} dx - \int_a^b g(x) z(x, t_0) dx + \\ &+ \int_{t_0}^{t_1} \int_a^b [u(x, t) \lambda(x, t) - z(x, t) m(x, t)] dx dt. \end{aligned} \quad (4.8)$$

Вычтем из (8) равенство (4). Получим с учетом (2)

$$\begin{aligned} J[\alpha, \beta] - \hat{J}[\alpha, \beta] &= \int_a^b \left[f(x) \frac{\partial z(x, t_0)}{\partial t} - \varphi(x) F(x) \right] dx - \\ &- \int_a^b [g(x) z(x, t_0) + \kappa(x) G(x)] dx - \\ &- \int_{t_0}^{t_1} \int_a^b [n(x, t) \lambda(x, t) + z(x, t) m(x, t)] dx dt. \end{aligned} \quad (4.9)$$

Возьмем математические ожидания от обеих частей этого равенства. Нетрудно видеть, что условие (5) эквивалентно равенству

$$\int_a^b \left[\frac{\partial z(x, t_0)}{\partial t} - \varphi(x) \right] F(x) dx - \int_a^b [z(x, t_0) + \kappa(x)] G(x) dx = 0.$$

Следовательно, для того, чтобы выполнялось условие несмещенности, достаточно положить

$$\varphi(x) = \frac{\partial z(x, t_0)}{\partial t}, \quad (4.10)$$

$$\kappa(x) = -z(x, t_0). \quad (4.11)$$

Если возвести равенство (9) в квадрат и взять математическое ожидание, то с учетом корреляционных свойств случайных полей и соотношений (10), (11) получим, что

$$\begin{aligned} \langle (J[\alpha, \beta] - \hat{J}[\alpha, \beta])^2 \rangle &= \\ &= \int_a^b \int_a^b \frac{\partial z(x, t_0)}{\partial t} \frac{\partial z(\xi, t_0)}{\partial t} \langle [f(x) - F(x)][f(\xi) - F(\xi)] \rangle dx d\xi + \end{aligned}$$

$$\begin{aligned}
& + \int_a^b \int_a^b z(x, t_0) z(\xi, t_0) \langle [g(x) - G(x)][g(\xi) - G(\xi)] \rangle dx d\xi + \\
& + 2 \int_a^b \int_a^b z(x, t_0) \frac{\partial z(\xi, t_0)}{\partial t} G(x) F(\xi) dx d\xi + \\
& + \int_{t_0}^{t_1} \int_a^b \int_{t_0}^{t_1} \int_a^b [z(x, t) z(\xi, \tau) \langle m(x, t) m(\xi, \tau) \rangle + \\
& + \lambda(x, t) \lambda(\xi, \tau) \langle n(x, t) n(\xi, \tau) \rangle] dx d\xi dt d\tau = \\
& = \int_a^b \left[\frac{\partial z(x, t_0)}{\partial t} \right]^2 \Phi(x) dx + \int_a^b z^2(x, t_0) \psi(x) dx + \\
& + 2 \int_a^b z(x, t_0) G(x) dx \int_a^b \frac{\partial z(\xi, t_0)}{\partial t} F(\xi) d\xi + \\
& + \int_{t_0}^{t_1} \int_a^b [z^2(x, t) \mu(x, t) + \lambda^2(x, t) \nu(x, t)] dx dt.
\end{aligned}$$

Отсюда вытекает, что справедлива следующая

Теорема 2. Для того чтобы функционал (4) давал оптимальную (несмещенную и эффективную) оценку функционала (3), достаточно, чтобы коэффициент $\lambda(x, t)$ представлял собой оптимальное управление системой (7) по критерию качества

$$\begin{aligned}
I[\lambda] = & \int_a^b \left\{ \left[\frac{\partial z(x, t_0)}{\partial t} \right]^2 \Phi(x) + z^2(x, t_0) \psi(x) \right\} dx + \\
& + 2 \int_a^b z(x, t_0) G(x) dx \cdot \int_a^b \frac{\partial z(\xi, t_0)}{\partial t} F(\xi) d\xi + \\
& + \int_{t_0}^{t_1} \int_a^b [z^2(x, t) \mu(x, t) + \lambda^2(x, t) \nu(x, t)] dx dt,
\end{aligned}$$

а коэффициенты $\varphi(x)$ и $\chi(x)$ определялись соотношениями (10), (11), в которых $z(x, t)$ — решение задачи (7), соответствующее оптимальному управлению.

Автор пользуется случаем выразить глубокую признательность А. А. Янцевичу за постановку задачи и полезные дискуссии.

ЛИТЕРАТУРА

1. Бутковский А. Г. Теория оптимального управления системами с распределенными параметрами. М., «Наука», 1965. 474 с.
2. Лионс Ж.-Л. Оптимальное управление системами, описываемыми уравнениями с частными производными. М., «Мир», 1972. 414 с.
3. Bensoussan A. Sur l'identification et le filtrage de systèmes gouvernés par des équations aux dérivées partielles. Paris, AGIP, 1969. 233 p.

4. Колмогоров А. Н. Интерполирование и экстраполирование случайных последовательностей. — «Изв. АН СССР, сер. мат.», 1941, 5, с. 3—14.
5. Яелом А. М. Введение в теорию стационарных случайных функций. — «Усп. мат. наук», 1952, 7:5, с. 3—168.
6. Wiener N. The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. N. U., Wiley, 1949. 163 p.
7. Zadeh L. A., Ragazzini J. R. An Extension of Wiener's Theory of Prediction. — J. of Appl. Physics, 1950, 21, p. 645—655.
8. Shinbrot M. Optimization of Time — Varying Linear Systems with Nonstationary Inputs. — «Trans. ASME», 1958, 80 D, p. 457—465.
9. Пугачев В. С. Применение канонических разложений случайных функций к определению оптимальной линейной системы. — «Автом. и телемех.», 1956, 17, с. 489—499.
10. Солодовников В. В., Батков А. М. К теории самооптимизирующихся систем. — «Proc. Heidelberg Conference on Automatic Controls», 1956, p. 308—323.
11. Kalman R. E. A New Approach to Linear Filtering and Prediction Problems. — «Trans. ASME», 1960, 82 D, p. 35—45.
12. Kalman R. E., Bucy R. S. New Results in Linear Filtering and Prediction Theory. — «Trans ASME», 1961, 83 D, p. 95—108.
13. Ройтенберг Я. Н. Автоматическое управление. М., «Наука», 1971. 395 с.
14. Bucy R. S. Optimal Filtering for Correlated Noise. — «J. of Math. Analysis and Appl.», 1967, 20, p. 1—8.
15. Falb P. L. Infinite — dimensional Filtering: the Kalman — Bucy Filter in Hilbert Space. «Information and Control», 1967, 11, p. 102—137.
16. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. М., «Мир», 1971. 400 с.
17. Острем К. Ю. Введение в стохастическую теорию управления. М., «Мир», 1973. 322 с.
18. Ахиезер Н. И., Глазман И. М. Теория линейных операторов в гильбертовом пространстве. М., «Наука», 1966. 543 с.

УДК 530

В. Г. КОЛОДЯЖНЫЙ

О СЛУЧАЙНЫХ ЭЛЕКТРОМАГНИТНЫХ ПОЛЯХ

Реальная среда, в которой распространяется электромагнитная волна, искажает эту волну. Представляет интерес изучение статистических свойств флуктуаций волновых полей, обусловленных случайным характером неоднородностей среды.

Рассмотрим распространение плоской монохроматической электромагнитной волны в статистически однородной и изотропной среде, тензор диэлектрической проницаемости которой представляется в виде

$$\epsilon_{ik} = \epsilon_0 \delta_{ik} + \epsilon'_{ik}. \quad (1)$$

Здесь $\epsilon'_{ik} = \epsilon''_{ik}(A)$ — случайное тензорное поле в точке A среды с заданной прямоугольной декартовой системой координат x_1, x_2, x_3 , вариации которого относительно ϵ_0 малы, со средним значением $\overline{\epsilon'_{ik}} = 0$ и корреляционной функцией $\tau_{iklm}(A_1, A_2)$, являющейся в силу предполагаемых свойств среды двухточечным изотропным тензором четвертого ранга [1]:

$$\tau_{iklm}(A_1, A_2) = \overline{\epsilon'_{ik}(A_1) \epsilon'_{lm}(A_2)} = \tau_{iklm}(r), \quad (2)$$

где $r = |\vec{r}|$ и $\vec{r}$ — радиус-вектор между точками A_2 и A_1 . Исходными при рассмотрении являются уравнения Максвелла [2]

$$\begin{cases} e_{ijk} \frac{\partial E_k}{\partial x_j} + i \frac{\omega}{c} H_i = 0, & \frac{\partial}{\partial x_i} (\epsilon_{ik} E_k) = 0, \\ e_{ijk} \frac{\partial H_k}{\partial x_j} - i \frac{\omega}{c} \epsilon_{ik} E_k = 0, & \frac{\partial H_i}{\partial x_i} = 0, \end{cases} \quad (3)$$

e_{ijk} — символ Леви-Чивитта, $\mu_{ik} = \delta_{ik}$.

Поле в среде ищется в виде

$$E_i = [E^0 \delta_{ii} + E'_i(A)] \exp(ik_3 x_3), \quad (4)$$

где первое слагаемое представляет собой плоскую волну, распространяющуюся вдоль оси x_3 с волновым вектором $\vec{k} = (0, 0, k_3)$, второе слагаемое — флуктуирующее электрическое поле с нулевым средним значением, имеющее тот же порядок, что и флуктуации тензора диэлектрической проницаемости.

Компоненты эффективного тензора диэлектрической проницаемости ϵ_{ik}^e определяются из выражений для вектора электрической индукции и

$$\epsilon_{ik}^e = \epsilon_0 \delta_{ik} + \frac{1}{E_m^0} \overline{\epsilon'_{ig}(A) E'_g(A)}, \quad (5)$$

где E_m^0 — величина вектора напряженности электрического поля волны, плоскость поляризации которой направлена по оси x_m .

Корреляционные функции

$$S_{jq} = \overline{E'_j(A_1) E'^*_{q}(A_2)}, \quad (6)$$

$$L_{jlm} = \overline{E'_j(A_1) \epsilon'^*_{lm}(A_2)} \quad (7)$$

можно считать двухточечными изотропными тензорами с зависимостью от двух пространственных направлений — направления распространения волны x_3 и вектора напряженности электрического поля x'_1 . S_{jq} , например, представляется в виде

$$\begin{aligned} S_{jq} = & S_1 r_j r_q + S_2 \delta_{jq} + (S_3^+ + i S_3^-) r_j \delta_{q1} + (S_3^+ - i S_3^-) r_q \delta_{j1} + \\ & + (S_4^+ + i S_4^-) r_j \delta_{q3} + (S_4^+ - i S_4^-) r_q \delta_{j3} + S_5 \delta_{j1} \delta_{q1} + \\ & + (S_6^+ + i S_6^-) \delta_{j1} \delta_{q3} + (S_6^+ - i S_6^-) \delta_{j3} \delta_{q1} + S_7 \delta_{j3} \delta_{q3}, \end{aligned} \quad (8)$$

где $S_R = S_R(r, r_1, r_3)$ — неизвестные скалярные функции. Тензор L_{jlm} строится аналогично.

Исключая H_i из уравнений системы (3), с учетом того, что для среды с невозмущенными параметрами эти уравнения для

поля $\vec{E}^0$ выполняются тождественно, в первом приближении по величинам порядка ε'_{ik} получаем

$$\begin{aligned} \frac{\partial^2 E'_j}{\partial x_n^2} + p^2 E'_j + 2ik_3 \delta_{n3} \frac{\partial E'_j}{\partial x_n} = -\frac{E^0}{\varepsilon_0} \delta_{1k} \left[\frac{\partial^2 \varepsilon'_{nk}}{\partial x_j \partial x_n} + \right. \\ \left. + \frac{\omega^2}{c^2} \varepsilon_0 \varepsilon'_{jk} - k_3^2 \delta_{j3} \delta_{n3} \varepsilon'_{nk} + ik_3 \left(\delta_{n3} \frac{\partial \varepsilon'_{nk}}{\partial x_j} + \delta_{j3} \frac{\partial \varepsilon'_{nk}}{\partial x_n} \right) \right], \end{aligned} \quad (9)$$

где

$$p^2 = \frac{\omega^2}{c^2} \varepsilon_0 - k_3^2. \quad (10)$$

Уравнения для нахождения S_R и $L_M^\pm$ -функций, определяющих тензоры S_{jq} и L_{ilm} , получаются с помощью уравнения (9). Так, умножая комплексно сопряженное уравнение (9), рассматриваемое в точке A_2 , на $E'_q(A_1)$ и усредняя, с помощью несложных преобразований получаем

$$\begin{aligned} \left(1 - \frac{r_3^2}{r^2} \right) \frac{\partial^2 S_R}{\partial r^2} + 2 \frac{r_1}{r} \frac{\partial^2 S_R}{\partial r_1 \partial r} + \frac{\partial^2 S_R}{\partial r_1^2} + \\ + \left(2 \frac{r_3^2}{r^3} + \frac{h-2}{r} \right) \frac{\partial S_R}{\partial r} + p^2 S_R = F_R, \end{aligned} \quad (11)$$

здесь и далее h — известное целое число. Для функций $L_M^\pm$ с помощью уравнения (9) получаем систему уравнений

$$\begin{aligned} \frac{\partial^2 L_M^+}{\partial r^2} + 2 \frac{r_1}{r} \frac{\partial^2 L_M^+}{\partial r_1 \partial r} + 2 \frac{r_3}{r} \frac{\partial^2 L_M^+}{\partial r_3 \partial r} + \frac{\partial^2 L_M^+}{\partial r_1^2} + \frac{\partial^2 L_M^+}{\partial r_3^2} + \\ + \frac{h \partial L_M^+}{r \partial r} + p^2 L_M^+ - 2k_3 \left(\frac{r_3}{r} \frac{\partial L_M^-}{\partial r} + \frac{\partial L_M^-}{\partial r_3} \right) = F_M, \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{\partial^2 L_M^-}{\partial r^2} + 2 \frac{r_1}{r} \frac{\partial^2 L_M^-}{\partial r_1 \partial r} + 2 \frac{r_3}{r} \frac{\partial^2 L_M^-}{\partial r_3 \partial r} + \frac{\partial^2 L_M^-}{\partial r_1^2} + \frac{\partial^2 L_M^-}{\partial r_3^2} + \\ + \frac{h \partial L_M^-}{r \partial r} + p^2 L_M^- + 2k_3 \left(\frac{r_3}{r} \frac{\partial L_M^+}{\partial r} + \frac{\partial L_M^+}{\partial r_3} \right) = \Phi_M. \end{aligned} \quad (13)$$

В этих уравнениях функции F_M и Φ_M являются линейными комбинациями частных производных функций τ_N , определяющих корреляционный тензор τ_{iklm} , и функций $L^\pm$, номер которых меньше M . Функции F_R в уравнении (11) являются линейными комбинациями частных производных функций $L_M^\pm$ и функций S , номер которых меньше R . Таким образом, при заданном тензоре τ_{iklm} и находимых последовательно L_1^+ , $L_2^\pm$ и т. д., затем S_1 , S_2 , $S_3^\pm$ и т. д. в правых частях уравнений (11), (12) и (13) все время имеем известные функции.

В общем случае нахождение решений этих уравнений довольно затруднительно, однако при рассмотрении корреляций вдоль осей прямоугольной декартовой системы координат, связанной с исходной волной описанным образом, легко отыскать фундаментальные системы решений соответствующих однородных уравнений. Тем самым задача в указанном случае решается до конца.

ЛИТЕРАТУРА

1. Бэтчелор Дж. Теория однородной турбулентности. М., ИЛ, 1955. 87 с.
2. Ландау Л. Д., Лифшиц Е. М. Электродинамика сплошных сред. ГИФМЛ, М., 1959. 148 с.

УДК 681.3:519.21

В. В. БАРАНОВ

О СВОЙСТВАХ ВЫИГРЫША В ОБОБЩЕННЫХ УПРАВЛЯЕМЫХ ПРОЦЕССАХ

1. Постановка задачи. Рассмотрим обобщенный управляемый процесс, определенный в [1] набором $\{x_{\sigma_k}(\pi), \tau, F_\zeta(t), F_{\sigma_k}, P_x\}$, где $x_{\sigma_k}(\pi)$ — траектории процесса, управляемого стратегией π в моменты σ_k , $k = 1, 2, \dots$; $\sigma_k = \sigma_{k-1} + \zeta$; $\zeta = \min(\xi, \tau)$; $\xi = \xi(\omega)$ — случайная величина, соответствующая длительности времени до обрыва наблюдаемого процесса; τ — неслучайный параметр, имеющий смысл периода наблюдения; $F_\zeta(t)$ — функция распределения величины ζ ; $F_{\sigma_k} - \sigma$ — алгебра, порожденная траекториями $x_{\sigma_k}(\pi)$; P_x — мера на σ -алгебре $\sigma\left(\bigcup_k F_{\sigma_k}\right)$.

Пространства состояний E и управлений Y будем полагать непустыми борелевскими множествами.

Введем следующие обозначения. Пусть $v(x, \pi)$ — ожидаемый выигрыш, полученный при стратегии π и начальном состоянии $x \in E$. Функцию $v^*(x) = \sup_{\pi} v(x, \pi)$ назовем ценой. Стратегию $\pi^*: v(x, \pi^*) = v^*(x) \geq v(x, \pi)$ для всех π назовем оптимальной. Наша задача состоит в исследовании наиболее важных свойств ожидаемого выигрыша и цены.

2. Свойства выигрыша $v(x, \pi)$.

Условимся обозначать прямое произведение множеств $A \times B = AB$. Положим тогда $H = EY$ и k -ю степень произведения множеств H обозначим $\Omega_k = H^k$, $k = 1, 2, \dots$. Пусть $\Omega = \lim_{k \rightarrow \infty} \Omega_k$ σ -алгебру на Ω_k обозначим F_k .

Пусть P — множество условных вероятностных мер, определенных для элементов из $\Omega_k E$ и сосредоточенных на Y , т. е. $P = P(Y/\Omega_k E)$. Бесконечное произведение множеств P обозначим $\Pi = P \times P \times \dots$. Элементы из $P(Y/\Omega_{k-1} E)$ будем обозначать $\pi_k = \pi_k(y/h_{k-1}, x)$, $h_{k-1} \in \Omega_{k-1}$, $x \in E$.

Определение 2.1. Последовательность элементов $(\pi_1, \pi_2, \dots) = \pi \in \Pi$ назовем стратегией.

Если для каждого $k = 1, 2, \dots$ $\pi_k(y/h_{k-1}, x) = \pi_k(y/x)$, то стратегию π назовем марковской.

Если $\pi_k(y/x) = \pi(y/x)$, $k = 1, 2, \dots$, то стратегию π назовем стационарной.

Если мера $\pi_k(y/x)$ вырождена, то π будем называть нерандомизированной. Очевидно, что нерандомизированная стратегия осуществляет отображение $X \rightarrow Y$.

Пусть далее L_n — класс F_n -измеримых функций g таких, что $|g(\omega_n)| \leq K < \infty$, где

$$\omega_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \in \Omega_n.$$

Обозначим $w(\zeta/x, y, z)$ — непосредственный выигрыш, полученный на промежутке ζ за переход из состояния $x \in E$ в $z \in E$ при управлении $y \in Y$. Пусть $F_\zeta(t)$ — функция распределения случайной величины ζ , где ζ является промежутком времени между соседними моментами выбора управлений. Положим

$$w(x, y, z) = \int w(\zeta/x, y, z) dF_\zeta(t). \quad (2.1)$$

Определим тогда $g(\omega_1) \equiv w(x, y, z)$ и $g(\omega_n) \equiv \sum_{k=1}^n w(x_k, y_k, x_{k+1})$.

Функцию $g(\omega_n)$ назовем непосредственным выигрышем за n шагов.

Положим далее $g(\omega) = \liminf_{n \rightarrow \infty} g(\omega_n)$. Будем допускать, что $|g(\omega)| \leq \infty$. Класс функций $g(\omega)$ обозначим L .

Определим теперь $v_n(x, \pi)$ — ожидаемый выигрыш за n шагов при условии, что используется стратегия π и начальным состоянием является $x \in E$:

$$v_n(x, \pi) = M_x^\pi g(\omega_n) = M_x^\pi \sum_{k=1}^n w(x_k, y_k, x_{k+1}). \quad (2.2)$$

Предел

$$\liminf_{n \rightarrow \infty} v_n(x, \pi) = \liminf_{n \rightarrow \infty} M_x^\pi g(\omega_n) = v(x, \pi) \quad (2.3)$$

будем называть просто ожидаемым выигрышем. Очевидно, что $v_n(x, \pi) \in L_n$, а $v(x, \pi) \in L$.

Введем интенсивность переоценки (дисконтирования) $\alpha > 0$, означающую, что непосредственный выигрыш $w = w(\zeta/x, y, z)$ имеет на самом деле значение $w(\zeta/\cdot) e^{-\alpha \zeta}$. Поскольку в (2.1) рассматривается математическое ожидание $M_\zeta w(\zeta/x, y, z)$, то определим величину $0 \leq \beta(\alpha) = \beta(\alpha; x, y, z) < 1$:

$$\beta(\alpha) = M_\zeta e^{-\alpha \zeta(x, y, z)} = \int_0^\infty e^{-\alpha t} dF_\zeta(t/x, y, z).$$

Тогда непосредственный дисконтированный выигрыш есть $\beta(\alpha) w(x, y, z)$, а ожидаемый дисконтированный выигрыш

$$v_n(\alpha; x, \pi) = M_x^\pi \sum_{k=1}^n w(x_k, y_k, x_{k+1}) \prod_{m=1}^k \beta_m(\alpha)$$

$$v_a(x, \pi) = \lim_{n \rightarrow \infty} v_n(a; x, \pi). \quad (2.4)$$

Обозначим далее $g^- = -\min(g, 0)$ и $g^+ = \max(g, 0)$. Пусть $L(A^-)$ и $L(A^+)$ — совокупность функций из L , удовлетворяющих дополнительно условиям

$$A^-: M_x^\pi \left[\sup_n g^-(\omega_n) \right] < \infty, \quad x \in E,$$

$$A^+: M_x^\pi \left[\sup_n g^+(\omega_n) \right] < \infty, \quad x \in E.$$

Положим также $L(A^-, A^+) = L(A^-) \cap L(A^+)$.

Лемма 2.1. Пусть $v_n(x, \pi) \in L(A^-, A^+)$. Тогда найдутся $\alpha > 0$ и стратегия π' такие, что

$$v_n(x, \pi) = v_\alpha(x, \pi') \in L(A^-, A^+).$$

Доказательство. Если $v_n(x, \pi) \in L(A^-, A^+)$, то это означает, что

$$v_n(x, \pi) = M_x^\pi \left| \sum_{k=1}^n w(x_k, y_k, x_{k+1}) \right| = C < \infty.$$

Очевидно, что в силу сходимости ряда в (2.4) при любом $\alpha > 0$ всегда можно подобрать $\alpha > 0$ и стратегию π' такие, что

$$v_\alpha(x, \pi') = \lim_{n \rightarrow \infty} M_x^{\pi'} \sum_{k=1}^n w(x_k, y_k, x_{k+1}) \prod_{m=1}^k \beta_m(\alpha) = C.$$

Следствие 2.1. Пусть функция выигрыша $v(x, \pi) \in L(A^-, A^+)$. Тогда она совпадает с некоторым дисконтированным выигрышем и

$$\sup_\pi v(x, \pi) = v^*(x) \in L(A^-, A^+).$$

Доказательство. В силу того, что ожидаемый выигрыш $v(x, \pi)$ представляется рядом (2.3) либо (2.4), то включение $v(x, \pi) \in L(A^-, A^+)$ возможно, если ряд (2.3) либо содержит конечное число членов, т. е. $v(x, \pi) = v_n(x, \pi)$, либо является дисконтированным выигрышем. Тогда с учетом леммы (2.1) $v(x, \pi)$ можно отождествить с некоторым дисконтированным выигрышем $v_\alpha(x, \pi')$. Но для дисконтированного выигрыша ряд (2.4) сходится при любой стратегии $\pi \in \Pi$. Тогда

$$\sup_\pi M_x^\pi \left[\sup_n |g_\alpha(\omega_n)| \right] < \infty.$$

Отсюда

$$\sup_\pi v_\alpha(x, \pi) = v_\alpha^*(x) \in L(A^-, A^+).$$

Введем теперь следующие обозначения:

$$M_x^\pi w(x, y, z) = \pi_1 q w(x) = \iint w(x, y, z) q(dz/x, y) \pi_1(dy/x). \quad (2.5)$$

$$\begin{aligned}
M_x^\pi \omega(x_k, y_k, x_{k+1}) &= \pi_1 q \dots \pi_k q \omega(x) = \\
&= \int \dots \int \omega(x, y, z) q(dz/x_k, y_k) \pi_k(dy/x_k) \dots \\
&\quad \dots q(dz/x_1, y_1) \pi_1(dy/x_1).
\end{aligned} \tag{2.6}$$

Очевидно тогда, что

$$v_n(x, \pi) = M_x^\pi \sum_{k=1}^n \omega(x_k, y_k, x_{k+1}) = \sum_{k=1}^n \pi_1 q \dots \pi_k q \omega(x).$$

Докажем теперь одно вспомогательное утверждение, которое для процессов с дискретным временем содержится в [2]. Однако наше доказательство несколько отличается от него и распространено на более общий случай обобщенных процессов.

Лемма 2.2. Для любой стратегии $\pi \in \Pi$ и любого $n = 1, 2, \dots$ найдется марковская стратегия $\hat{\pi}$ такая, что $v_n(x, \pi) = v_n(x, \hat{\pi})$.

Доказательство. Лемма очевидным образом справедлива для $n = 1$, так как в этом случае $\pi_1 = \pi_1(\cdot/x_1) = \hat{\pi}_1(\cdot/x_1)$ и, следовательно, с учетом обозначений (2.5) получаем

$$\begin{aligned}
v_1(x, \pi) &= M_x^\pi \omega(x_1, y_1, x_2) = \pi_1 q \omega(x) = \\
&= \hat{\pi}_1 q \omega(x) = M_x^{\hat{\pi}} \omega(x_1, y_1, x_2) = v_1(x, \hat{\pi}).
\end{aligned}$$

Предположим теперь, что лемма справедлива для $n = N$. Рассмотрим $M_x^\pi \omega(x_{N+1}, y_{N+1}, x_{N+2})$. Из свойств матожиданий следует

$$\begin{aligned}
M_x^\pi \omega(x_{N+1}, y_{N+1}, x_{N+2}) &= M_x^\pi \{M \omega(x_{N+1}, y_{N+1}, x_{N+2})/x_{N+1}\} = \\
&= M_x^\pi \{M_{x_{N+1}}^\pi \omega(x_{N+1}, y_{N+1}, x_{N+2})\} = M_x^\pi u(x_{N+1}),
\end{aligned}$$

где $u(x_{N+1}) = M_{x_{N+1}}^\pi \omega(x_{N+1}, y_{N+1}, x_{N+2})$. Очевидно, что всегда найдется марковское $\hat{\pi}_{N+1}$ такое, что (2.5)

$$u(x_{N+1}) = \hat{\pi}_{N+1} q \omega(x_{N+1}).$$

Тогда с учетом предположения индукции имеем

$$\begin{aligned}
M_x^\pi \omega(x_{N+1}, y_{N+1}, x_{N+2}) &= M_x^\pi u(x_{N+1}) = \pi_1 q \dots \pi_N q u(x_{N+1}) = \\
&= \hat{\pi}_1 q \dots \hat{\pi}_N q u(x_{N+1}) = \hat{\pi}_1 q \dots \hat{\pi}_N q \hat{\pi}_{N+1} q \omega(x_{N+1}) = \\
&= M_x^{\hat{\pi}} \omega(x_{N+1}, y_{N+1}, x_{N+2}).
\end{aligned}$$

Отсюда и с учетом предположения индукции получаем:

$$v_{N+1}(x, \pi) = M_x^\pi \sum_{k=1}^{N+1} \omega(x_k, y_k, x_{k+1}) = \sum_{k=1}^{N+1} M_x^\pi \omega(x_k, y_k, x_{k+1}) =$$

$$\begin{aligned}
&= \sum_{k=1}^N M_x^{\pi} \omega(x_k, y_k, x_{k+1}) + M_x^{\pi} \omega(x_{N+1}, y_{N+1}, x_{N+2}) = \\
&= \sum_{k=1}^N \hat{M}_x^{\pi} \omega(x_k, y_k, x_{k+1}) + \hat{M}_x^{\pi} \omega(x_{N+1}, y_{N+1}, x_{N+2}) = \\
&= \sum_{k=1}^{N+1} \hat{M}_x^{\pi} \omega(x_k, y_k, x_{k+1}) = \hat{M}_x^{\pi} \sum_{k=1}^{N+1} \omega(x_k, y_k, x_{k+1}) = v_{N+1}(x, \hat{\pi}),
\end{aligned}$$

что и доказывает лемму.

Из доказанной леммы вытекает

Теорема 2.1. Для любой стратегии $\pi \in \Pi$ найдется марковская стратегия $\hat{\pi}$ такая, что $v(x, \pi) = v(x, \hat{\pi})$.

Доказательство.

$$v(x, \pi) = \liminf_{n \rightarrow \infty} v_n(x, \pi) = \liminf_{n \rightarrow \infty} v_n(x, \hat{\pi}) = v(x, \hat{\pi}).$$

Доказанная теорема позволяет при исследовании свойств функции выигрыша $v(x, \pi)$ ограничиться классом марковских стратегий.

Рассмотрим класс ограниченных функций $\omega \in L_1$. Докажем еще одно вспомогательное утверждение, сформулированное в [3], но не доказанное.

Лемма 2.3. Для любых $\pi \in P(Y/E)$ и $\omega \in L_1$ найдется вырожденное $f \in P(Y/E)$ такое, что

$$\int_Y \omega(x, y, z) df(y/x) \geq \int \omega(x, y, z) d\pi(y/x).$$

Доказательство. Поскольку L_1 является множеством ограниченных функций, то для каждого $x, z \in E$ существует элемент $y^* \in Y$ такой, что

$$\omega(x, y^*, z) \geq \sup_y \omega(x, y, z) - \varepsilon, \quad \varepsilon > 0.$$

Множество $P(Y/E)$, являясь множеством вероятностных мер на Y , содержит вырожденные меры. Тогда можно выбрать элемент $f \in P(Y/E)$ такой, что для каждого $x \in E$

$$f(y/x) = \begin{cases} 1, & \text{если } y = y^* \\ 0, & \text{если } y \neq y^*. \end{cases}$$

Тогда для фиксированных $x, z \in E$

$$\begin{aligned}
&\int_Y \omega(x, y, z) df(y/x) = \omega(x, y^*, z) \geq \\
&\geq \sup_y \omega(x, y, z) - \varepsilon \geq \int_Y \omega(x, y, z) d\pi(y/x) - \varepsilon
\end{aligned}$$

для любых $\pi \in P(Y/E)$ и $\varepsilon > 0$. Отсюда в силу произвольности $\varepsilon > 0$

$$\int_Y \omega(x, y, z) df(y/x) \geq \int_Y \omega(x, y, z) d\pi(y/x).$$

Из этой леммы и теоремы 2.1 очевидным образом вытекает следующая теорема, которую мы приводим без доказательства.

Теорема 2.2. Для любой марковской стратегии $\pi = (\pi_1, \pi_2, \dots)$ найдется мажорирующая ее нерандомизированная марковская стратегия $f = (f_1, f_2, \dots)$ такая, что

$$v(x, f) \geq v(x, \pi) \text{ и } v_\alpha(x, f) \geq v_\alpha(x, \pi).$$

Теперь мы можем доказать следующую важную теорему.

Теорема 2.3. Если существует оптимальная стратегия π^* такая, что

$$v(x, \pi^*) = \sup_{\pi} v(x, \pi),$$

то она содержится в классе нерандомизированных марковских стратегий.

Доказательство. Обозначим $\{f\} = \Pi_D$ класс нерандомизированных марковских стратегий. Тогда очевидно включение $\Pi_D \subset \Pi$. Отсюда

$$\sup_{\pi \in \Pi} v(x, \pi) \geq \sup_{f \in \Pi_D} v(x, f). \quad (2.7)$$

С другой стороны, в силу теоремы 2.2 для любой стратегии $\pi \in \Pi$ найдется нерандомизированная марковская стратегия $f \in \Pi_D$ такая, что $v(x, f) \geq v(x, \pi)$. Отсюда

$$\sup_{f \in \Pi_D} v(x, f) \geq \sup_{\pi \in \Pi} v(x, \pi). \quad (2.8)$$

Неравенства (2.7) и (2.8) дают равенство

$$\sup_{f \in \Pi_D} v(x, f) = \sup_{\pi \in \Pi} v(x, \pi),$$

которое и доказывает теорему.

Доказанная теорема позволяет при исследовании управляемых процессов ограничиться классом нерандомизированных марковских стратегий.

Введем теперь операторы, связанные с нерандомизированными стратегиями $f \in \Pi_D$. С этой целью любой функции $g \in L$ сопоставим оператор $T = T(f) : L \rightarrow L$:

$$Tg(x) = \int_E \int_Y g(z) dq(z/x, y) df(y/x) = f q g(x).$$

Если $g \in L_1$, то

$$Tg(x) = \int_E \int_Y g(x, y, z) dq(z/x, y) df(y/x) = f q g(x).$$

Теперь функции $v(x, \pi) \in L$ поставим в соответствие оператор $F_\alpha = F_\alpha(f) : L \rightarrow L$, определенный следующим образом:

$$F_\alpha v(x, \pi) = T \beta_\alpha(x) [\omega(x) + v(x, \pi)],$$

где $\omega(x) \in L_1$.

Если $\alpha = 0$, то $F_\alpha = F$:

$$Fv(x, \pi) = T[\omega(x) + v(x, \pi)].$$

Нам потребуются следующие очевидные свойства операторов F_α и F , которые мы сформулируем без доказательства.

Лемма 2.4. а) F_α , (F) монотонен, т. е. $u \geq v \Rightarrow F_\alpha u \geq F_\alpha v$, б) $F_\alpha(f) v_\alpha(x, \pi) = v_\alpha(x, (f\pi))$, где $(f\pi) = (f, \pi_1, \pi_2, \dots)$. Последнее свойство, естественно, имеет место и для оператора F .

Лемма 2.5. Пусть π некоторая стратегия. Тогда $F^n v(x, \pi) \rightarrow v(x, f^\infty)$ при $n \rightarrow \infty$, где $f^\infty = (f, f, \dots)$ — стационарная стратегия.

Доказательство. В силу свойства б) оператора F имеем

$$Fv(x, \pi) = v(x, (f\pi)),$$

откуда

$$F^n v(x, \pi) = v(x, (f^n \pi)),$$

где

$$(f^n \pi) = (f, \dots, f, \pi_1, \pi_2, \dots).$$

Тогда при

$$\underbrace{\quad\quad\quad}_n$$

$$n \rightarrow \infty \quad F^n v(x, \pi) \rightarrow v(x, f^\infty).$$

Следствие 2.2. Пусть $\pi = (f, f, \dots)$ — стационарная стратегия. Тогда

$$F(f)v(x, \pi) = v(x, \pi).$$

Доказательство. В силу свойства б) оператора F $Fv(x, \pi) = v(x, (f\pi))$. Но $(f\pi) = (f, f, \dots) = \pi$, что и доказывает следствие.

Заметим, что из определения оператора T и ожидаемого выигрыша имеем

$$v_n(x, \pi) = M_x^n \sum_{k=1}^n \omega(x_k, y_k, x_{k+1}) = \sum_{k=1}^n f_1 q \dots f_k q \omega(x) =$$

$$= \sum_{k=1}^n T(f_1) T(f_2) \dots T(f_k) \omega(x).$$

Очевидно, если стратегия π стационарна, то

$$v_n(x, \pi) = \sum_{k=1}^n T^k(f) \omega(x).$$

Используя теперь оператор F , нетрудно получить следующее рекуррентное соотношение:

$$v_n(x, \pi) = F(f) v_{n-1}(x, \pi). \quad (2.9)$$

Действительно,

$$v_n(x, \pi) = \sum_{k=1}^n T^k \omega(x) = T \omega(x) + T \sum_{k=2}^n T^{k-1} \omega(x) =$$

$$= T \omega(x) + T \sum_{k=1}^{n-1} T^k \omega(x) = T[\omega(x) + v_{n-1}(x, \pi)] = F v_{n-1}(x, \pi).$$

Лемма 2.6. Пусть $\pi = (f, f, \dots)$ — стационарная стратегия. Положим $v_0(x, \pi) \equiv 0$. Тогда

$$v(x, \pi) = \liminf_{n \rightarrow \infty} [F^n(f) 0](x).$$

Доказательство. По определению (2.3)

$$v(x, \pi) = \liminf_{n \rightarrow \infty} v_n(x, \pi),$$

где $v_n(x, \pi)$ удовлетворяет соотношению (2.9). Из (2.9) рекуррентным способом получаем

$$v_n(x, \pi) = F^n(f) v_0(x, \pi).$$

Но

$$v_0(x, \pi) \equiv 0.$$

Отсюда

$$v_n(x, \pi) = [F^n(f) 0](x).$$

Теперь

$$v(x, \pi) = \liminf_{n \rightarrow \infty} [F^n(f) 0](x).$$

Определение 2.2. Будем говорить, что оператор F сохраняет $v = v(x, \pi) \in L$, если $Fv \geq v$.

Определение 2.3. Функцию $v = v(x, \pi) \in L$ будем называть эксцессивной относительно оператора F (F -эксцессивной), если $Fv \leq v$.

Эксцессивная функция (э. ф.) v называется эксцессивной мажорантой (э. м.) $g = g(x, \pi)$, если $v(x, \pi) \geq g(x, \pi)$ для всех $x \in E$.

Э. м. v называется наименьшей э. м. (н. э. м.) функции g , если v меньше или равна любой другой э. м. функции g .

Замечание 2.1. Если некоторая функция $v(x, \pi)$ F -эксцессивна, то она также и F_α -эксцессивна.

Доказательство. Из определения F -эксцессивности v следует

$$v(x, \pi) \geq T[\omega(x) + v(x, \pi)].$$

Очевидно, что это неравенство не нарушится, если его правую часть в квадратных скобках умножить на некоторую величину $\beta_\alpha(x) \leq 1$ для каждого $x \in E$.

Тогда

$$v(x, \pi) \geq T\beta_\alpha(x)[\omega(x) + v(x, \pi)] = F_\alpha v(x, \pi).$$

Теорема 2.4. Пусть $\pi = (f, f, \dots)$ — стационарная стратегия. Тогда а) F сохраняет $v(x, \pi)$ в том и только в том случае, если выигрыш $\omega = \omega(x, y, z)$ и оператор $T = T(f)$ таковы, что существует предел

$$\lim_{k \rightarrow \infty} T^k = T^*, \quad T\omega(x) \geq 0 \quad \text{и} \quad \lim_{k \rightarrow \infty} T^k \omega(x) = T^* \omega(x) \geq 0;$$

б) если $T\omega(x) \leq 0$ и $T^*\omega(x) \leq 0$, то функция $v(x, \pi)$ F -экссивна;

в) если $T\omega(x) \geq 0$, а $\lim_{k \rightarrow \infty} T^k\omega(x) = T^*\omega(x) < 0$, то $v_n(x, \pi)$ меняет знак при $n \rightarrow \infty$.

Доказательство. а) Пусть $T\omega(x) \geq 0$ и $T^*\omega(x) \geq 0$. Покажем, что F сохраняет v . Действительно, если $T\omega(x) \geq 0$ и $T^*\omega(x) \geq 0$, то $T^k\omega(x) \geq 0$ для всех $k = 1, 2, \dots$. Очевидно, что тогда для любого $n = 1, 2, \dots$ выполняется неравенство

$$\sum_{k=1}^n T^k\omega(x) \geq \sum_{k=1}^{n-1} T^k\omega(x) \quad (2.10)$$

или

$$\begin{aligned} \sum_{k=1}^n T^k\omega(x) &= T\omega(x) + T \sum_{k=2}^n T^{k-1}\omega(x) = \\ &= T[\omega(x) + v_{n-1}(x, \pi)] \geq v_{n-1}(x, \pi). \end{aligned}$$

Поскольку (2.10) справедливо при всех $n = 1, 2, \dots$, то в последнем неравенстве, переходя к пределу при $n \rightarrow \infty$ и используя теорему Лебега о монотонной сходимости, получаем

$$T[\omega(x) + v(x, \pi)] \geq v(x, \pi).$$

Но левая часть этого неравенства определяет оператор F , откуда $Fv(x, \pi) \geq v(x, \pi)$, т. е. оператор F сохраняет $v(x, \pi)$.

Наоборот, пусть $Fv(x, \pi) \geq v(x, \pi)$, где

$$v(x, \pi) = \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k\omega(x).$$

Тогда

$$T\left[\omega(x) + \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k\omega(x)\right] \geq \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k\omega(x).$$

В силу леммы Фату имеем

$$\begin{aligned} &\liminf_{n \rightarrow \infty} T\left[\omega(x) + \sum_{k=1}^n T^k\omega(x)\right] \geq \\ &\geq T\left[\omega(x) + \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k\omega(x)\right] \geq \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k\omega(x). \end{aligned}$$

Откуда

$$\begin{aligned} &\liminf_{n \rightarrow \infty} T\left[\omega(x) + \sum_{k=1}^n T^k\omega(x)\right] = \\ &= \liminf_{n \rightarrow \infty} \left[T\omega(x) + T \sum_{k=1}^n T^k\omega(x)\right] = \\ &= \liminf_{n \rightarrow \infty} \left[T\omega(x) + \sum_{k=1}^n T^{k+1}\omega(x)\right] = \end{aligned}$$

$$\begin{aligned}
&= \liminf_{n \rightarrow \infty} \left[T w(x) + \sum_{k=2}^{n+1} T^k w(x) \right] = \\
&= \liminf_{n \rightarrow \infty} \sum_{k=1}^{n+1} T^k w(x) \geq \liminf_{n \rightarrow \infty} \sum_{k=1}^n T^k w(x).
\end{aligned}$$

Очевидно, что последнее неравенство возможно в том случае, если для всех $n = 1, 2, \dots$

$$\sum_{k=1}^{n+1} T^k w(x) \geq \sum_{k=1}^n T^k w(x)$$

или

$$\sum_{k=1}^n T^k w(x) \geq \sum_{k=1}^{n-1} T^k w(x),$$

т. е.

$$\sum_{k=1}^{n-1} T^k w(x) + T^n w(x) \geq \sum_{k=1}^{n-1} T^k w(x).$$

Отсюда следует, что $T^n w(x) \geq 0$ для всех $n = 1, 2, \dots$. Тогда $T w(x) \geq 0$ и $\lim_{n \rightarrow \infty} T^n w(x) = T^* w(x) \geq 0$, что и доказывает утверждение а теоремы; б доказывается аналогичным образом с заменой неравенств на противоположные.

Докажем теперь последнее утверждение теоремы, в. Из существования предела $\lim_{k \rightarrow \infty} T^k = T^*$ следует, что оператор T^k монотонен.

По условию теоремы $T w(x) \geq 0$, а $\lim_{k \rightarrow \infty} T^k w(x) = T^* w(x) < 0$. Тогда существует некоторый номер $m < \infty$ такой, что $T^{m-1} w(x) \geq 0$, а $T^m w(x) < 0$ и в силу монотонности оператора T^k имеем $T^k w(x) < 0$ для всех $k \geq m$.

Возьмем $n > m$ и рассмотрим

$$v_n(x, \pi) = \sum_{k=0}^n T^{k+1} w(x).$$

Очевидно, равенство

$$\sum_{k=1}^n T^k w(x) = \sum_{k=1}^{m-1} T^k w(x) + \sum_{k=m}^n T^k w(x). \quad (2.11)$$

Поскольку $T^k w(x) \geq 0$ для $k \leq m-1$ и $T^k w(x) < 0$ для $k \geq m$, то первая сумма в (2.11) неотрицательна, а вторая отрицательна, причем в силу монотонности оператора T^k получаем

$\sum_{k=m}^n T^k w(x) \downarrow -\infty$ при $n \rightarrow \infty$. Отсюда следует, что существует

такой номер $N \geq m$, что $v_{N-1}(x, \pi) = \sum_{k=1}^{N-1} T^k w(x) \geq 0$, а $v_N(x, \pi) =$

$= \sum_{k=1}^{\infty} T^k \omega(x) < 0$. Таким образом получаем, что $v_n(x, \pi)$ меняет знак при $n \rightarrow \infty$. Теорема доказана.

Замечание 2.2. Теорема очевидным образом справедлива для оператора F_a .

Следствие 2.3. а) Если оператор F сохраняет $v(x, \pi)$, то $v_n(x, \pi) = [F^n 0](x) \uparrow \infty$ при $n \rightarrow \infty$;

б) если F не сохраняет $v(x, \pi)$, то $v_n(x, \pi) = [F^n 0](x) \rightarrow -\infty$ при $n \rightarrow \infty$.

Доказательство. а) Если F сохраняет $v(x, \pi)$, то в силу теоремы 2.4 $T^k \omega(x)$ для всех $k = 1, 2, \dots$. При этом выполняется неравенство

$$v_n(x, \pi) = \sum_{k=1}^n T^k \omega(x) \geq \sum_{k=1}^{n-1} T^k \omega(x) = v_{n-1}(x, \pi).$$

Отсюда $v_n(x, \pi) \uparrow \infty$ при $n \rightarrow \infty$.

Замечая, что в силу леммы 2.6 $v_n(x, \pi) = [F^n 0](x)$, получаем утверждение а).

б) Если F не сохраняет $v(x, \pi)$, то, как и в теореме 2.4,

$$v_n(x, \pi) = \sum_{k=1}^{m-1} T^k \omega(x) + \sum_{k=m}^n T^k \omega(x),$$

где

$$0 \leq \sum_{k=1}^{m-1} T^k \omega(x) \leq C < \infty, \text{ а } \sum_{k=m}^n T^k \omega(x) \downarrow -\infty$$

при $n \rightarrow \infty$. Отсюда

$$v_n(x, \pi) \rightarrow -\infty$$

при $n \rightarrow \infty$.

Учитывая снова лемму 2.6, получаем утверждение б).

Теорема 2.5. а) Если оператор F сохраняет $v(x, \pi)$, то

$$v(x, \pi) \in L(A^-);$$

б) если F не сохраняет $v(x, \pi)$, то $v(x, \pi) \in L(A^+)$;

с) если $F = F_a$, то $v_a(x, \pi) \in L(A^-, A^+)$.

Доказательство. а) Пусть F сохраняет $v(x, \pi)$. Тогда, так же как в теореме 2.4, легко проверить, что F сохраняет $v_n(x, \pi)$ для всех $n = 1, 2, \dots$, т. е.

$$Fv_n(x, \pi) \geq v_n(x, \pi), \quad n = 1, 2, \dots$$

Но $v_n(x, \pi)$ удовлетворяет рекуррентному соотношению

$$Fv_n(x, \pi) = v_{n+1}(x, \pi).$$

Отсюда

$$v_{n+1}(x, \pi) \geq v_n(x, \pi), \quad n = 1, 2, \dots$$

или с учетом обозначения (2.2)

$$M_x^\pi g(\omega_{n+1}) = v_{n+1}(x, \pi) \geq v_n(x, \pi) = M_x^\pi g(\omega_n), \quad n = 1, 2, \dots \quad (2.12)$$

Заметим теперь, что $g(\omega_1) \in L_1$, где L_1 — класс ограниченных функций. Следовательно, $g(\omega_1) > -\infty$. Из (2.12) теперь получаем

$$M_x^\pi g(\omega_{n+1}) \geq M_x^\pi g(\omega_n) \geq \dots \geq M_x^\pi g(\omega_1) > -\infty.$$

Откуда $g(\omega_n) > -\infty$ для всех $n = 1, 2, \dots$, или

$$g^-(\omega_n) = -\min[0, g(\omega_n)] < \infty$$

для всех $n = 1, 2, \dots$. Отсюда немедленно следует

$$M_x^\pi[\sup_n g(\omega_n)] > -\infty \quad \text{или} \quad M_x^\pi[\sup_n g^-(\omega_n)] < \infty.$$

Тем самым

$$v(x, \pi) = \liminf_{n \rightarrow \infty} v_n(x, \pi) = \liminf_{n \rightarrow \infty} M_x^\pi g(\omega_n) \in L(A^-).$$

в) Пусть F не сохраняет $v(x, \pi)$. Тогда, как и в теореме 2.4, для достаточно большого $n < \infty$ существует $m \leq n$ такое, что

$$v_n(x, \pi) = M_x^\pi \sum_{k=1}^{m-1} w(x_k, y_k, x_{k+1}) + M_x^\pi \sum_{k=m}^n w(x_k, y_k, x_{k+1}),$$

где

$$0 \leq M_x^\pi \sum_{k=1}^{m-1} w(x_k, y_k, x_{k+1}) \leq C < \infty,$$

а

$$M_x^\pi \sum_{k=m}^n w(x_k, y_k, x_{k+1}) \downarrow -\infty \quad \text{при} \quad n \rightarrow \infty.$$

Отсюда

$$M_x^\pi[\sup_n g^+(\omega_n)] = M_x^\pi[\sup_m g(\omega_{m-1})] \leq C < \infty.$$

Таким образом,

$$v(x, \pi) = \liminf_{n \rightarrow \infty} v_n(x, \pi) = \liminf_{n \rightarrow \infty} M_x^\pi g(\omega_n) \in L(A^+).$$

с) Пусть $F = F_\alpha$. Тогда с учетом леммы 2.6 имеем

$$\begin{aligned} v_\alpha(x, \pi) &= \liminf_{n \rightarrow \infty} M_x^\pi \sum_{k=1}^n \prod_{m=1}^k \beta_m(\alpha) w(x_k, y_k, x_{k+1}) = \\ &= \liminf_{n \rightarrow \infty} \sum_{k=1}^n M_x^\pi \prod_{m=1}^k \beta_m(\alpha) w(x_k, y_k, x_{k+1}) = \\ &= \liminf_{n \rightarrow \infty} [F_\alpha^n 0](x). \end{aligned} \quad (2.13)$$

Обозначим $\prod_{m=1}^k \beta_m(a) \omega(x_k, y_k, x_{k+1}) = \omega_k(a)$. Из сходимости

ряда в (2.13) следует $\lim_{n \rightarrow \infty} M_x^\pi \sum_{k=1}^n |\omega_k(a)| < \infty$, причем знак inf можно опустить. Но $|\omega(a)| = \omega^-(a) + \omega^+(a)$. Тогда получаем

$$\lim_{n \rightarrow \infty} M_x^\pi \sum_{k=1}^n \omega_k^-(a) < \infty \text{ и}$$

$$\lim_{n \rightarrow \infty} M_x^\pi \sum_{k=1}^n \omega_k^+(a) < \infty. \quad (2.14)$$

Обозначая $\sum_{k=1}^n \omega_k^-(a) = g_a^-(\omega_n)$ и $\sum_{k=1}^n \omega_k^+(a) = g_a^+(\omega_n)$, из (2.14)

получаем

$$M_x^\pi \left[\sup_n g_a^-(\omega_n) \right] < \infty, \quad M_x^\pi \left[\sup_n g_a^+(\omega_n) \right] < \infty.$$

Отсюда

$$v_a(x, \pi) = \lim_{n \rightarrow \infty} M_x^\pi \sum_{k=1}^n \omega_k(a) \in L(A^-, A^+).$$

3. Свойства цены $v^*(x)$. Для исследования свойств цены $v^*(x)$ введем операторы U и U_a :

$$Uv(x, \pi) = \sup_f F(f)v(x, \pi)$$

и

$$U_av(x, \pi) = \sup_f F_a(f)v(x, \pi).$$

Заметим прежде всего, что из монотонности оператора F следует монотонность оператора U , а из определения оператора U вытекает, что, если F сохраняет v , то U также сохраняет v .

В дальнейшем при исследовании свойств цены будем предполагать, что оптимальная стратегия $\pi^*: v(x, \pi^*) = \sup_{\pi} v(x, \pi) = v^*(x)$ существует.

Заметим также, что из определения оператора U следует, что существует элемент $f^s = f^s(x) \in P(Y|E)$ такой, что

$$F(f^s)v(x, \pi) \geq Uv(x, \pi) - \varepsilon$$

и соответственно

$$[F(f^s)0](x) \geq [U0](x) - \varepsilon.$$

С учетом этих замечаний справедлива

Теорема 3.1.

$$\lim_{n \rightarrow \infty} [U^n 0](x) = \sup_{\pi} v(x, \pi) = v(x, \pi^*) = v^*(x).$$

Доказательство. Из определения оператора U следует, что $[U0](x) \geq [F(f)0](x)$ для всех $f \in P(Y|E)$. Отсюда и в силу монотонности оператора U получаем $[U^n 0](x) \geq [F^n(f)0](x) = v_n(x, \pi)$ для всех $\pi = (f, f, \dots)$ и $n = 1, 2, \dots$

Следовательно,

$\lim_{n \rightarrow \infty} [U^n 0](x) \geq \liminf_{n \rightarrow \infty} [F^n 0](x, \bar{v} = (x, \pi))$ для всех π . Отсюда

$$\lim_{n \rightarrow \infty} [U^n 0](x) \geq \sup_{\pi} v(x, \pi) = v(x, \pi^*) = v^*(x). \quad (3.1)$$

С другой стороны, существует элемент $f^{\varepsilon'} \in P(Y|E)$ такой, что $[F(f^{\varepsilon'}) 0](x) \geq [U 0](x) - \varepsilon'$ или $[U 0](x) \leq [F(f^{\varepsilon'}) 0](x) + \varepsilon'$. Выберем ε'_k , $k = 1, 2$ такие, что $\sum_{k=1}^{\infty} \varepsilon'_k = \varepsilon$.

Тогда

$$[U^n 0](x) \leq [F^n(f^{\varepsilon'}) 0](x) + \sum_{k=1}^n \varepsilon'_k.$$

Отсюда

$$\begin{aligned} \lim_{n \rightarrow \infty} [U^n 0](x) &\leq \liminf_{n \rightarrow \infty} \left\{ [F^n(f^{\varepsilon'}) 0](x) + \sum_{k=1}^n \varepsilon'_k \right\} = \\ &= v(x, \pi^{\varepsilon'}) + \sum_{k=1}^{\infty} \varepsilon'_k = v(x, \pi^{\varepsilon'}) + \varepsilon \leq \\ &\leq \sup_{\pi} v(x, \pi) + \varepsilon = v^*(x) + \varepsilon. \end{aligned}$$

В силу произвольности $\varepsilon > 0$ получаем

$$\lim_{n \rightarrow \infty} [U^n 0](x) \leq \sup_{\pi} v(x, \pi) = v^*(x). \quad (3.2)$$

Из неравенств (3.1) и (3.2) следует равенство

$$\lim_{n \rightarrow \infty} [U^n 0](x) = v^*(x) = \sup_{\pi} v(x, \pi) = v(x, \pi^*).$$

Теорема доказана.

Используя оператор U_α , аналогичным образом можно доказать эту теорему для случая с дисконтированием.

Теорема 3.2. а) Если $v(x, \pi) \in L(A^-)$ или $v(x, \pi) \in L(A^-, A^+)$, то цена $v^*(x) = \sup_{\pi} v(x, \pi)$ принадлежит соответствующему классу $L(A^-)$ или $L(A^-, A^+)$; б) если $v(x, \pi) \in L(A^+)$, то цена принадлежит либо классу $L(A^+)$, либо $L(A^-)$.

Доказательство. а) Если $v(x, \pi) \in L(A^-)$, то

$$M_x^\pi \left[\sup_n g(\omega_n) \right] > -\infty.$$

Отсюда

$$\sup_{\pi} M_x^\pi \left[\sup_n g(\omega_n) \right] > -\infty,$$

т.е.

$$\sup_{\pi} M_x^\pi \left[\sup_n g^-(\omega_n) \right] < \infty.$$

Тогда

$$v^*(x) = \sup_{\pi} v(x, \pi) \in L(A^-).$$

Если $v(x, \pi) \in L(A^-, A^+)$, то в силу следствия 2.1 $v^*(x) = \sup_{\pi} v(x, \pi) \in L(A^-, A^+)$.

б) Пусть $v(x, \pi) \in L(A^+)$. Тогда $v_n(x, \pi) = M_x^n[g(\omega_n)]$ либо меняет знак при $n \rightarrow \infty$, либо $v_n(x, \pi) < 0$ для всех $n = 1, 2, \dots$. В обоих случаях оператор F не сохраняет $v_n(x, \pi)$, а тем самым и $v(x, \pi)$.

Обозначим f^* элемент $f \in P(Y|E)$ такой, что

$$F(f^*)v(x, \pi) = \sup_i F(f)v(x, \pi) = Uv(x, \pi).$$

Рассмотрим теперь оператор $F(f^*)$ и стратегию $\pi^* = (f^*, f^*, \dots)$. Очевидно, что относительно элемента f^* теперь уже имеется две возможности: либо $F(f^*)$ так же как $F(f)$ не сохраняет $v(x, \pi)$, либо $F(f^*)$ сохраняет $v(x, \pi)$. Тогда в силу теоремы 2.5, учитывая, что

$$\liminf_{n \rightarrow \infty} [F^n(f^*)0](x) = \liminf_{n \rightarrow \infty} [U^n 0](x) = v(x, \pi^*) = v^*(x),$$

получаем в первом случае $v^*(x) \in L(A^+)$ и во втором случае $v^*(x) \in L(A^-)$, что и завершает доказательство теоремы.

Теорема 3.3. Пусть некоторая функция выигрыша $u(x) \in L(A^-)$ либо $u(x) \in L(A^+)$ эксцессивна относительно оператора F . Тогда она является мажорантой всех выигрышей, т.е. $u(x) \geq v(x, \pi)$ для всех π .

Для доказательства теоремы вначале предположим, что $u(x) \in L(A^-, A^+)$. Тогда справедлива

Лемма 3.1. Пусть $u(x) \in L(A^-, A^+)$ эксцессивна относительно оператора F_a . Тогда она является мажорантой всех выигрышей из класса $L(A^-, A^+)$ т.е. $u(x) \geq v_a(x, \pi)$ для всех π .

Доказательство. Обозначим $\beta_a = \sup_{x, y} \beta_a(x, y)$ и $R = \sup_{x, y, z} \omega(x, y, z) > 0$. Так как $\omega(x, y, z) \in L_1$, то $R \leq K < \infty$ и, если $\alpha > 0$, то $\beta_a < 1$. Поскольку $u(x) \in L(A^-, A^+)$, то в силу следствия 2.1 $u(x)$ совпадает с дисконтированным выигрышем при некоторой стратегии π , т.е.

$$u(x) = v_a(x, \pi) = M_x^n \sum_{k=1}^n \omega(x_k, y_k, x_{k+1}) \prod_{m=1}^k \beta_m(\alpha).$$

Отсюда получаем, что

$$u(x) \leq R \left(\sum_{k=0}^{\infty} \beta_a^k - 1 \right) = \frac{R\beta_a}{1 - \beta_a}.$$

Обозначим $\frac{R\beta_a}{1 - \beta_a} = b_a$ и $\varphi^{b_a}(x) = \min(b_a, u(x))$. Покажем, что $\varphi^{b_a}(x)$ F_a -эксцессивна.

Действительно, поскольку $u(x) \geq \varphi^{b_a}(x)$, то в силу монотонности F_a

$$u(x) \geq F_a u(x) \geq F_a \varphi^{b_a}(x). \quad (3.3)$$

Далее, учитывая, что $b_\alpha = R\beta_\alpha + b_\alpha\beta_\alpha$, имеем

$$b_\alpha = R\beta_\alpha + b_\alpha\beta_\alpha = T\beta_\alpha(x) [\omega(x) + \varphi^{b_\alpha}(x)] = F_\alpha \varphi^{b_\alpha}(x). \quad (3.4)$$

Из неравенств (3.3) и (3.4) теперь получаем

$$\varphi^{b_\alpha}(x) = \min(\beta_\alpha, u(x)) \geq F_\alpha \varphi^{b_\alpha}(x). \quad (3.5)$$

Отсюда и в силу монотонности оператора F_α

$$\varphi^{b_\alpha}(x) \geq F_\alpha \varphi^{b_\alpha}(x) \geq \dots \geq F_\alpha^n \varphi^{b_\alpha}(x).$$

Тогда в силу леммы 2.5

$$\varphi^{b_\alpha}(x) \geq F_\alpha^n \varphi^{b_\alpha}(x) \downarrow v_\alpha(x, f^\infty),$$

где $f^\infty = (f, f, \dots)$ стационарная нерандомизированная стратегия. Так как элемент $f \in P(Y|E)$, соответствующий оператору $F(f)$, выбран произвольно, то неравенство $\varphi^{b_\alpha}(x) \geq v_\alpha(x, f^\infty)$ справедливо для всех f^∞ , в том числе и для стратегии π^* :

$$v_\alpha(x, \pi^*) \geq v_\alpha(x, \pi) \text{ для всех } \pi.$$

Отсюда

$$u(x) \geq \varphi^{b_\alpha}(x) \geq v_\alpha(x, \pi) \text{ для всех } \pi. \quad (3.6)$$

т.е. $u(x)$ является мажорантой всех выигрышей $v_\alpha(x, \pi) \in L(A^-, A^+)$.

Доказательство теоремы 3.3. Пусть $u(x) \in L(A^-)$ эксцессивна относительно оператора F . Рассмотрим функцию $u^b(x) = \min(b, u(x))$, которая при каждом фиксированном $b > 0$ очевидным образом принадлежит классу $L(A^-, A^+)$. Выберем $\alpha = \alpha(b)$ такое, чтобы $b = \frac{R\beta_\alpha}{1 - \beta_\alpha}$. Покажем, что $u^b(x) F_{\alpha(b)} -$ эксцессивна. Действительно, поскольку, $u(x)$ F -эксцессивна, то в силу замечания 2.1 она и F_α -эксцессивна. Тогда

$$u(x) \geq F_\alpha u(x) \geq F_\alpha u^b(x). \quad (3.7)$$

Далее, также как в (3.4), имеем

$$b \geq F_{\alpha(b)} u^b(x). \quad (3.8)$$

Теперь из (3.7) и (3.8) следует

$$u^b(x) = \min(b, u(x)) \geq F_{\alpha(b)} u^b(x).$$

Тогда для $u^b(x)$ справедливо неравенство (3.6) при каждом фиксированном $b > 0$, т.е.

$$u(x) \geq u^b(x) \geq v_{\alpha(b)}(x, \pi) \text{ для всех } \pi. \quad (3.9)$$

Переходя к пределу при $b \rightarrow \infty$, получаем

$$u(x) = \lim_{b \rightarrow \infty} u^b(x) \geq \lim_{\alpha(b) \rightarrow 0} v_{\alpha(b)}(x, \pi) = v(x, \pi)$$

для всех π , что и доказывает утверждение теоремы для $u(x) \in L(A^-)$.

Пусть теперь F — эксцессивная функция $u(x) \in L(A^+)$. Положим $u_a(x) = \max(a, u(x))$, $a \leq 0$. Очевидно, что $u_a(x) \in (A^-, A^+)$. Аналогичным образом выберем $\alpha = \alpha(a)$ такое, чтобы

$a = \frac{R\beta_{a(a)}}{1 - \beta_{a(a)}}$. Покажем, что $u_a(x)$ F_a -эксцессивна. Действительно, поскольку $u_a(x) \geq u(x)$ и в силу замечаний 2.1 $u_a(x) \geq u(x) \geq F_a u(x)$. Отсюда

$$u(x) = \lim_{a \rightarrow -\infty} u_a(x) \geq \lim_{a(a) \rightarrow 0} v_a(x, \pi) = v(x, \pi) \text{ для всех } \pi,$$

что и завершает доказательство теоремы.

Следствие 3.1. Теорема 3.2 справедлива, если функция $u(x)$ эксцессивна относительно оператора $U = \sup F(f)$.

Доказательство. Обозначим $f^* \in P(Y|E)$: $F(f^*) = \sup F(f)$.

Теперь $U = F(f^*)$, откуда немедленно получаем требуемое утверждение.

Теорема 3.4. Цена $v^*(x) = \sup_{\pi} v(x, \pi)$ эксцессивна относительно оператора U .

Доказательство. При исследовании цены мы предположили, что существует оптимальная стратегия. Тогда существует стратегия $\pi^\varepsilon = (f^\varepsilon, f^\varepsilon, \dots)$ и оператор $F(f^\varepsilon)$ такие, что $v(x, \pi^\varepsilon) \geq v^*(x) - \varepsilon$ и $F(f^\varepsilon)v(x, \pi) \geq Uv(x, \pi) - \varepsilon$, $\varepsilon > 0$. Тогда с учетом следствия 2.2 имеем

$$\begin{aligned} v^*(x) &\geq v(x, \pi^\varepsilon) = F(f^\varepsilon)v(x, \pi^\varepsilon) \geq F(f^\varepsilon)[v^*(x) - \varepsilon] = \\ &= F(f^\varepsilon)v^*(x) - \varepsilon \geq Uv^*(x) - \varepsilon - \varepsilon. \end{aligned}$$

В силу произвольности $\varepsilon > 0$ получаем

$$v^*(x) \geq Uv^*(x),$$

что и доказывает теорему.

Следствие 3.2. Цена $v^*(x)$ является U -эксцессивной мажорантой (U — э. м.) выигрышей $v(x, \pi)$.

Доказательство. Непосредственно вытекает из теорем 3.3 и 3.4.

Теорема 3.5. Цена $v^*(x)$ является н. U -э. м. выигрышей $v(x, \pi)$.

Доказательство. С учетом следствия 3.2 необходимо только показать, что $v^*(x)$ является наименьшей мажорантой $v(x, \pi)$.

Предположим, что имеется некоторая другая U -э. м. $g(x)$. Тогда $g(x) \geq v(x, \pi)$ для всех π и тем самым $g(x) \geq \sup_{\pi} v(x, \pi) = v^*(x)$. Отсюда следует, что цена $v^*(x)$ н. U -э. м. $v(x, \pi)$.

Лемма 3.2. Если $g(x)$ — н. U -э. м. $v(x, \pi)$, то

$$g(x) = \max \{v(x, \pi), Ug(x)\}.$$

Доказательство. Очевидно, если $g(x)$ U -э.м. $v(x, \pi)$, то

$$g(x) \geq \max \{v(x, \pi), Ug(x)\}.$$

Рассмотрим

$$\varphi(x) = \max \{v(x, \pi), Ug(x)\}.$$

Очевидно, что $\varphi(x)$ — U -э.м. $v(x, \pi)$. Действительно, поскольку $\varphi(x) \geq v(x, \pi)$ и в силу $\varphi(x) \leq g(x)$ получаем

$$U\varphi(x) \leq Ug(x) \leq \max \{v(x, \pi), Ug(x)\} = \varphi(x).$$

С другой стороны, $g(x)$ — н. U -э.м. $v(x, \pi)$. Поэтому $\varphi(x) \geq g(x)$. Следовательно, $\varphi(x) = g(x)$.

Теорема 3.6. Цена $v^*(x)$ удовлетворяет рекуррентному уравнению

$$v^*(x) = \max \{v(x, \pi), Uv^*(x)\}.$$

Доказательство. Непосредственно вытекает из теоремы 3.5 и леммы 3.2.

Теорема 3.7. Цена $v^*(x)$ является единственным решением уравнения оптимальности

$$v^*(x) = Uv^*(x).$$

Если $v(x, \pi) \in L(A^-, A^+)$, то это решение ограничено.

Доказательство. Поскольку цена U -эксцессивна, то

$$v^*(x) \geq Uv^*(x). \quad (3.10)$$

С другой стороны, цена $v^*(x)$ достигается на некоторой стратегии $\pi^* = (f^*, f^*, \dots)$, т.е. $v^*(x) = v(x, \pi^*)$. Тогда с учетом следствия 2.2.

$$\begin{aligned} v^*(x) &= v(x, \pi^*) = F(f^*)v(x, \pi^*) \leq \\ &\leq \sup_f F(f)v(x, \pi) = Uv^*(x). \end{aligned} \quad (3.11)$$

Из неравенств (3.10) и (3.11) следует равенство $v^*(x) = Uv^*(x)$.

Единственность следует из того, что $v^*(x)$ — н. U -э.м. $v(x, \pi)$.

Наконец, если $v(x, \pi) \in L(A^-, A^+)$, то в силу следствия 2.1 $v^*(x, \pi) \in L(A^-, A^+)$, откуда следует ограниченность $v^*(x)$. Но $v^*(x)$ является решением уравнения оптимальности. Следовательно, это решение с необходимостью должно быть ограниченным.

4. Заключение. В данной работе исследованы свойства функции выигрыша и цены для обобщенных управляемых процессов, которые необходимы в дальнейшем при исследовании проблемы существования оптимальной стратегии, построении алгоритма ее поиска и решении задачи об оптимальной остановке управляемых процессов.

ЛИТЕРАТУРА

1. Баранов В. В. Обобщенные управляемые процессы. Тезисы докладов на III Всесоюзном совещании по статистическим методам в процессах управления. Вильнюс, 1973. с. 141—143.
2. Strauch R. E. Negative dynamic programming. The Annals of Mathematical Statistics, 1966, N 37, p. 871—890.
3. Blackwell D. Discounted dynamic programming. The Annals of Mathematical Statistics, 1965, N 36, p. 226—235.

СОДЕРЖАНИЕ

Львов В. А., Хруслов Е. Я. Краевые задачи в областях с мелкозернистой границей для линеаризованной системы уравнений Навье — Стокса	3
Львов В. А., Хруслов Е. Я. Краевые задачи в областях с мелкозернистой границей для общей системы уравнений Навье — Стокса	23
Маринич А. П. Каноническая форма идеально наблюдаемых систем	36
Геворкян Ю. Л., Пашкова Т. В. О полноте и базисности некоторых систем в L_p	40
Куриной Г. Ч. Условие изоморфизма конечных модулярных решеток	45
Куриной Г. Ч. О расширениях конечных структур	47
Керер Е. Ш. Некоторые теоремы о строении конечнопорожденных модулей над кольцами главных идеалов	55
Керер Е. Ш. О кольце эндоморфизмов конечнопорожденного модуля над кольцом главных идеалов	67
Баранов В. В. К проблеме синтеза оптимальной стратегии в управляемых процессах	75
Соколовский В. З. Об оптимальных оценках случайных полей	88
Колодяжный Б. Г. О случайных электромагнитных полях	102
Баранов В. В. О свойствах выигрыша в обобщенных управляемых процессах	105

ВЕСТНИК ХАРЬКОВСКОГО УНИВЕРСИТЕТА

№ 119

Математика и механика

Выпуск 40

Редактор А. П. Гужва

Технический редактор Г. П. Александрова

Корректоры Л. К. Бердута, А. В. Евлахова

Сдано в набор 26/XI 1974 г. Подписано в печать 20/III 1975 г. Формат 60×90^{1/16}. Бумага типографская № 3. Усл.-печ. л. 8. Уч.-изд. л. 7,3. Тираж 1000. Заказ 4-2425. БЦ 50084. Цена 49 коп.

Издательство издательского объединения «Вища школа» при Харьковском государственном университете. 310003, Харьков, 3, ул. Университетская, 16.

Отпечатано с матриц книжной фабрики «Коммунист» Республиканского производственного объединения «Полиграфкнига» Госкомиздата УССР в Харьковской городской типографии № 16 Областного управления по делам издательств, полиграфии и книжной торговли, Харьков-3, Университетская, 16. Зак. 1473.

РЕФЕРАТЫ

УДК 517.9

Краевые задачи в областях с мелкозернистой границей для линеаризованной системы уравнений Навье — Стокса. Львов В. А., Хруслов Е. Я. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 3—22.

Рассматривается первая краевая задача для линеаризованной системы уравнений Навье — Стокса в области, являющейся дополнением к большому числу мелких связанных множеств (зерен). Изучается асимптотическое поведение решения задачи, когда число зерен неограниченно растет, а диаметры их уменьшаются. Выведены дифференциальные уравнения и граничные условия для описания первого числа асимптотики.

Ил. 2. Библиогр. 3.

УДК 517.9

Краевые задачи в областях с мелкозернистой границей для общей системы уравнений Навье — Стокса. Львов В. А., Хруслов Е. Я. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 23—36.

Рассматривается первая краевая задача (стационарная и нестационарная) для нелинейной системы уравнений Навье — Стокса в области, являющейся дополнением к большому числу непересекающихся связанных множеств (зерен). Параметры задачи предполагаются такими, чтобы существовало единственное решение. Изучается асимптотическое поведение решения, когда диаметры зерен стремятся к нулю, а число их неограниченно растет. Выведены уравнения и граничные условия, описывающие первые члены асимптотики решений этих задач.

Библиогр. 6.

УДК 517.91

Каноническая форма идеально наблюдаемых систем. Маринич А. П. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 36—40.

Для линейной наблюдаемой системы

$$\dot{x} = Ax + Bu, y = hx$$

с одним выходом получена каноническая форма идеально наблюдаемых систем, а также необходимые и достаточные условия идеально наблюдаемых систем.

Библиогр. 2.

УДК 513.8

О полноте и базисности некоторых систем в L_p . Геворкян Ю. Л., Пашкова Т. В. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 40—45.

Рассматриваются условия, при которых полная система неотрицательных функций в L_p будет сверхполной. Система называется сверхполной, если она остается полной после удаления любого количества конечного числа элемен-

тов. Вводится понятие обобщенных шаудеровых систем в пространствах C и L_1 и выясняются условия полноты этих систем в L_p . Доказывается, что широкий класс некоторых шаудеровых систем не может быть базисом в L_p .

Библиогр. 2.

УДК 512.9

Услови́е изоморфизма конечных модулярных решеток. Куринной Г. Ч. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 45—47.

Доказывается следующая теорема: две конечные модулярные решетки изоморфны тогда и только тогда, когда изоморфны их частично упорядоченные подмножества, состоящие из тех элементов, которые представимы в виде объединения неразложимых элементов.

Библиогр. 4.

УДК 512.9

О расширениях конечных модулярных структур. Куринной Г. Ч. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 47—55.

По двум заданным конечным модулярным структурам, одна из которых предполагается подпрямой неразложимой, в статье строятся все конечные модулярные структуры, являющиеся их подпрямым произведением.

Библиогр. 3.

УДК 519.48

Некоторые теоремы о строении конечнопорожденных модулей над кольцами главных идеалов. Керер Е. Ш. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 55—67.

Доказывается, что каждый конечнопорожденный модуль над кольцом главных идеалов является расширением циклического модуля с помощью модуля симметрического типа. В случае периодических модулей доказывается также двойственное утверждение: каждый периодический конечнопорожденный модуль над кольцом главных идеалов является расширением модуля симметрического типа с помощью циклического модуля. Затем приводится применение этих результатов к теории матриц.

Библиогр. 2

УДК 519.48

О кольце эндоморфизмов конечнопорожденного модуля над кольцом главных идеалов. Керер Е. Ш. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 67—75.

Изучаются эндоморфизмы конечнопорожденного модуля над кольцом главных идеалов. Устанавливается необходимое и достаточное условие аддитивной порождаемости кольца эндоморфизмов модуля рассматриваемого типа его автоморфизмами. Устанавливается также необходимое и достаточное условие совпадения множества характеристических подмодулей заданного модуля с множеством его вполне характеристических подмодулей.

УДК 681.3:519.21

К проблеме синтеза оптимальной стратегии в управляемых процессах. Баранов В. В. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 75—88.

Доказано существование оптимальной стратегии в обобщенных управляемых процессах и построен итеративный алгоритм ее поиска. Этот алгоритм имеет важное преимущество по сравнению с алгоритмами Ховарда и сведения к задаче линейного программирования в том, что в нем отсутствует необходимость решения систем линейных уравнений. Алгоритм построен для управляемых процессов с конечным множеством состояний и управлений. Однако

он легко может быть распространен на более сложные ситуации с компактным либо счетным множеством состояний и управлений.

Библиогр. 6.

УДК 519. 21+621.391

Об оптимальных оценках случайных полей. Соколовский В. З. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 88—102.

Некоторые важные результаты теории фильтрации Калмана — Бьюси перенесены на случай полей, описываемых уравнениями в частных производных. Для уравнений первого порядка определение коэффициентов оптимального фильтра сводится к решению дифференциального уравнения в частных производных первого порядка с квадратичной нелинейностью. Для уравнений гиперболического типа доказана теорема двойственности задач фильтрации и детерминированного управления.

Библиогр. 18.

УДК 530

О случайных электромагнитных полях. Колодяжный Б. Г. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 102—105.

Находятся уравнения для отыскания тензорных корреляционных функций электрического поля и тензора диэлектрической проницаемости при распространении плоской монохроматической электромагнитной волны в среде с малыми случайными вариациями тензора диэлектрической проницаемости.

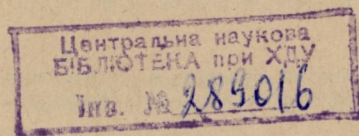
Библиогр. 2.

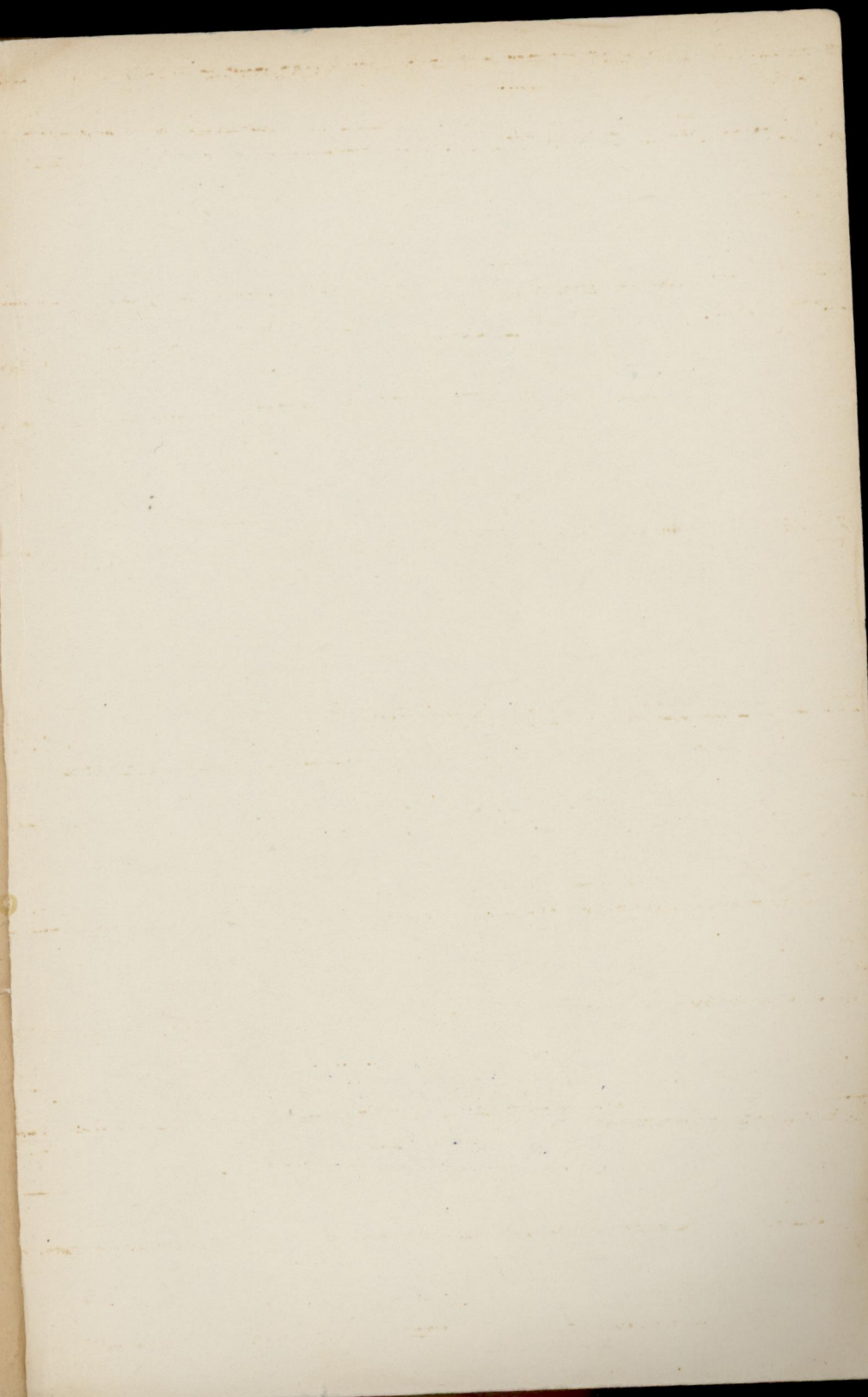
УДК 681.3:519:21

О свойствах выигрыша в обобщенных процессах. Баранов В. В. «Вестник Харьковского университета. Математика и механика», вып. 40, 1975, с. 105—122.

Исследованы свойства функции выигрыша и цены в обобщенных управляемых процессах. Показано, что оптимальная стратегия содержится в классе нерандомизированных марковских стратегий. Доказано, что цена является наименьшей эксцессивной мажорантой выигрышей. Установлено рекуррентное соотношение, которому удовлетворяет цена.

Библиогр. 3.





49 коп.

V.N. Karazin Kharkiv National University



00280558

6