ОБ ОДНОЙ ЗАДАЧЕ УПРАВЛЕНИЯ СЛУЧАЙНЫМИ БЛУЖДАНИЯМИ НА ПЛОСКОСТИ

Э. О. Рапопорт

Институт математики им. С. Л. Соболева СО РАН,
630090, Новосибирск, Россия

УДК 519.87

Изучается динамическая экономическая система с дискретным временем, состояния которой в каждый момент времени характеризуются целыми неотрицательными точками двумерного векторного пространства. Имеются два продукта и несколько различных производств, в каждом из которых состояние системы может изменяться на некоторый случайный вектор с целыми компонентами с различными наборами вероятностей. Под управлением понимается выбор в каждый момент времени одного из имеющихся производств. Цель управления — минимизация вероятности выхода из положительного квадранта. Исследуются вопросы существования цен на продукты, согласованные с оптимальным управлением.

Ключевые слова: оптимальное управление, случайные блуждания, цены, марковская цепь.

We study the dynamic economic system with discrete time, the state of which at any given time are characterized by non-negative integer points of the two-dimensional vector space. There are two products and several different manufactures, in each of which the state of the system may vary to some random vector with integer components with different sets of probabilities. Governance refers to the choice at every moment of one of the existing facilities. Management goal — to minimize the probability of leaving the positive quadrant. Study the existence of food prices, consistent with the optimal control.

Key words: control, optimal prices, random walk, Markov chain.

Введение. Одна из стандартных задач динамического управления экономикой — распределение в каждый момент времени ограниченного ресурса между конкурирующими производствами. Обычно рассматриваются ситуации, когда ресурс можно разбивать для такого распределения на несколько частей. Здесь будет изучаться случай, когда ресурс неделим, т. е. когда его необходимо целиком отдавать только одному производству. Ясно, что нехватка ресурса в тех производствах, для которых ресурс не выделяется, должна приводить к ухудшению состояния, характеризующего результаты этого производства, в то время как выделение ресурса — к улучшению состояния. Такие задачи обычно решаются методами целочисленного программирования, методами динамического программирования и т. п. Наличие стохастических факторов, влияющих на выпуск продукции в каждом из производств, существенно усложняет задачу.

Интересный подход был предложен Р. Раднером и М. Ротшильдом в [1]. В рассматриваемых авторами моделях необходимо распределить единичный неделимый ресурс между k различными производствами, каждое из которых производит k видов продуктов. Выделение ресурса какому-либо производству приводит к увеличению запаса одних продуктов и уменьшению запаса других (увеличение и уменьшение зависят от случайных факторов).

Считается, что система функционирует (живет) до тех пор, пока количество каждого продукта неотрицательно и хотя бы для одного — строго положительно. При этом ищется управление, которое максимизирует вероятность выживания, т. е. минимизирует вероятность выхода из положительного квадранта.

Подобные модели при $k = 2$ изучались автором в [2], [3], [4]. При разных предположениях о природе стохастических объектов и разных классах управлений удалось описать асимптотическое поведение оптимальной траектории и вероятности ее вырождения.

В [4] исследованы вопросы ценообразования, построена система цен, согласованная с оптимальным управлением.

Такая система цен оказалась полезной для исследования случая, когда число производств превышает число продуктов. Число продуктов по-прежнему предполагается равным двум.

1. Основная модель. Имеются два продукта и два способа производства этих продуктов. Состояние системы отождествляется с целочисленными точками неотрицательного квадранта плоскости N_+^2. В каждый дискретный момент времени инвестор должен вкладывать неделимый единичный ресурс в одно из двух производств. Опишем процесс функционирования такого вложения.

Пусть задан набор целочисленных векторов $(a_i, b_i),\ 1 \leqslant i \leqslant k$. При вложении в первое производство система переходит из точки (x, y) в одну из точек $(x + a_i, y + b_i)$ с вероятностями p_i, при вложении во второе производство — с вероятностями q_i. Некоторые из вероятностей могут быть равны нулю, $\sum_i p_i = \sum_i q_i = 1$. Через ξ будем обозначать двумерную случайную величину, определяемую набором вероятностей $\{p_i\}$, а через η — двумерную случайную величину, определяемую набором вероятностей $\{q_i\}$.

Под вырождением системы мы будем понимать выход ее из первого квадранта.

Под управлением системой будем понимать выбор в каждый момент времени вложения в первое или второе производство, т. е. выбор одной из двух случайных величин ξ или η, определяющих следующее состояние системы. Цель управления — минимизация вероятности вырождения системы, т. е. минимизация вероятности выхода из первого квадранта.

Естественно потребовать, чтобы первое управление (вложение ресурса в первое производство) было „лучше“ для первого продукта, а второе управление — „лучше“ для второго продукта. Формально это предположение будем записывать так:

$$\begin{array}{lclclcl} \sum_i p_i b_i & < & 0, & \sum_i p_i(a_i + b_i) & > & 0, \\ \sum_i q_i a_i & < & 0, & \sum_i q_i(a_i + b_i) & > & 0. \end{array} \tag{1}$$

Заметим, что из этих неравенств следует выполнение условий из [1], при которых там была доказана теорема существования управления с ненулевой вероятностью выживания. В принятых обозначениях эти условия сводятся к неравенству

$$\sum_i p_i b_i \cdot \sum_i q_i a_i < \sum_i p_i a_i \cdot \sum_i q_i b_i.$$

Управлением в точке (x, y) будем называть функцию $\gamma(x, y)$, принимающую значения 0 или 1. Через γ будем обозначать функцию, определенную на N_+^2, в каждой точке (x, y) принимающую значение $\gamma(x, y), \gamma = \{\gamma(x, y)\}$.

Множество всех допустимых управлений будем обозначать через Γ:

$$\Gamma = \{\gamma : \ \gamma(x,y) \in \{0,1\} \ \forall (x,y) \in N_+^2\}.$$

Тогда при фиксированном управлении γ получаем марковский процесс $Z(t) = (X(t), Y(t))$, матрица переходов которого определяется следующим образом: из точки (x,y) мы можем попасть в точки $(x+a_i, y+b_i)$ с вероятностями $s_i = \gamma(x,y)p_i + (1-\gamma(x,y))q_i$ соответственно. Таким образом, марковский процесс в каждый момент t определяется рекуррентными соотношениями

$$Z(t+1) = Z(t) + \gamma(X(t), Y(t))\xi + (1-\gamma(X(t), Y(t)))\eta,$$

причем $Z(0) = (x,y)$.

Оператор переходов этого марковского процесса обозначим через Φ_γ. В этих обозначениях $Z(t+1) = \Phi_\gamma(Z(t))$.

Как уже говорилось, для каждой точки $(x,y) \in N_+^2$ нужно выбрать политику $\gamma = \{\gamma(x,y)\}$, принимающую в каждой точке значения 0 или 1. При такой политике вероятность вырождения $g_\gamma(x,y)$ (выхода из первого квадранта) должна быть минимальным решением следующей системы уравнений ([6], с. 100)

$$g_\gamma(x,y) = \gamma \sum_i p_i g_\gamma(x+a_i, y+b_i) + (1-\gamma) \sum_i q_i g_\gamma(x+a_i, y+b_i)$$

с граничными условиями: при $(x,y) \notin N_+^2$ выполняется равенство $g_\gamma(x,y) = 1$. Положим,

$$g(x,y) = \inf_{\gamma \in \Gamma} \{g_\gamma(x,)\}.$$

Можно показать,(см. [5]) что существует такое управление $\hat{\gamma}$, что

$$g(x,y) = g_{\hat{\gamma}}(x,y).$$

Управление $\hat{\gamma}$ будем называть оптимальным управлением. Отметим, что $g(x,y)$ является решением системы

$$g(x,y) = \min(\sum_i p_i g(x+a_i, y+b_i), \sum_i q_i g(x+a_i, y+b_i)), (x,y) \in N_+^2$$

с граничными условиями $g(x,y) = 1$, если $(x,y) \notin N_+^2$.

Для исследования свойств марковской цепи, возникающей при оптимальном управлении, оказывается полезным рассмотрение системы уравнений (введенной в [4]), которую здесь запишем в следующем виде

$$\left\{ \begin{array}{lcl} \sum\limits_i p_i e^{-(\alpha a_i + \beta b_i)} & = & 1 \\ \sum\limits_i q_i e^{-(\alpha a_i + \beta b_i)} & = & 1. \end{array} \right. \tag{2}$$

Эту систему будем называть системой, ассоциированной с рассматриваемой марковской цепью, или просто ассоциированной системой. Очевидно, что пара $(0,0)$ является решением этой системы.

Было показано, что при выполнении условий (1) система (2) может иметь еще только положительные решения (β^*, α^*). В [5] были приведены примеры, когда имеется несколько положительных решений и когда положительных решений нет.

Всюду в этой работе будем предполагать, что **ассоциированная система имеет только одно положительное решение**.

Пусть, как и выше, $g(x, y)$ — вероятность вырождения при оптимальной политике, если система находится в точке (x, y). Тогда

$$g(x, y) = \min(\sum_i p_i g(x + a_i, y + b_i), \sum_i q_i g(x + a_i, y + b_i))$$

Эти равенства должны выполняться для всех точек (x, y), таких, что при любых i точка $(x + a_i, y + b_i)$ лежит в неотрицательном квадранте. Если же для некоторого i либо $x + a_i < 0$, либо $y + b_i < 0$, то $g(x + a_i, y + b_i) = 1$.

Покажем, что введенная функция g является субрегулярной $(g(x, y) \leqslant Mg(\Phi_\gamma(x, y)))$ при любом управлении $\gamma \in \Gamma$ для всех целочисленных точек первого квадранта.

Действительно, $Mg(x, y) = \sum\limits_i p_i g(x + a_i, y + b_i)$, если $\gamma(x, y) = 1$ и $Mg(x, y) = \sum\limits_i q_i g(x + a_i, y + b_i)$, если $\gamma(x, y) = 0$. Поэтому $g(x, y) \leqslant Mg(\Phi(x, y))$.

Отметим, что $1 - g$ является суперрегулярной и неотрицательной. По теореме о представлении неотрицательной суперрегулярной функции (см. [6], с. 99) функция $1 - g$ однозначно представима в виде

$$1 - g = r + Nf,$$

где r — регулярная неотрицательная функция, $f = (I - P)h$, Nf — потенциал некоторой меры, носитель которой сосредоточен в точках (x, y), для которых существуют i, такие, что либо $x + a_i < 0$, либо $y + b_i < 0$. Поэтому $g = 1 - r - Nf$, и это представление единственно.

Пусть (α, β) — какое-либо решение ассоциированной системы, отличное от (0,0). В [5] была доказана следующая нижняя оценка для функции g: существует такая постоянная c_1, что

$$c_1 e^{-(\alpha x + \beta y)} \leqslant g(x, y).$$

При дополнительном предположении

Условие А: $p_i = 0$ при $b_i > 0$ и $q_i = 0$ при $a_i > 0$,

была получена аналогичная верхняя оценка: существует константа c_2, такая, что

$$g(x, y) \leqslant c_2 e^{-(\alpha x + \beta y)}.$$

Описанную выше задачу о выборе управления, минимизирующего вероятность выхода из первого квадранта, будем называть **задачей 1**.

2. Цены и асимптотическое поведение суммарной стоимости. При управлении рассматриваемой выше системой возникает вопрос о разумных ценах, которые нужно назначать на производимые продукты. Цены являются естественным экономическим инструментом, влияющим на выбор того или иного управления опосредованно. Естественно считать, что разумным образом выбранные цены должны быть как-то согласованы с оптимальным управлением.

В этом разделе рассмотрим вопросы такого согласования.

Основной принцип согласования — цены должны быть такими, чтобы политика, являющаяся оптимальной для задачи 1, была „хорошей“ для задачи, возникающей при управлении системой с помощью цен. Формально это требование опишем позднее.

При введении системы цен (p_1, p_2) необходимо некоторое условие их нормировки — выбор масштаба. В литературе обычно принимается нормировка, порожденная лебеговой нормой: $\|(p_1, p_2)\| = |p_1| + |p_2|$. В рассматриваемой задаче более удобной является нормировка, порожденная эвклидовой нормой: $\|(p_1, p_2)\| = \sqrt{|p_1|^2 + |p_2|^2}$. При такой нормировке оказалось, что „хорошие“ цены должны быть пропорциональны решению ассоциированной системы.

Считая, что эвклидова норма вектора цен равна 1, цены (p_1, p_2) на производимые продукты можно задавать в виде пары $(\cos\varphi, \sin\varphi)$. При таких ценах стоимость s произведенного набора составляет величину $s = x\cos\varphi + y\sin\varphi$, $s \geqslant 0$. Мы будем теперь рассматривать одномерную ситуацию — управляемое блуждание системы на прямой. Под вырождением системы будем теперь понимать выход на отрицательную полуось, т. е. попадание в состояние, когда общее количество денег становится отрицательным. Цель управления (выбора одного из двух случайных блужданий), как и в задаче 1, — минимизация вероятности вырождения системы.

Обозначим $\delta_i = a_i\cos\varphi + b_i\sin\varphi$.

Предположим, что в каждом состоянии s выбрана некоторая политика γ – выбор в каждом состоянии одного из блужданий с вероятностями $\{p_i\}$ или $\{q_i\}$. В итоге при фиксированном управлении получаем марковскую цепь на множестве состояний S_0, порожденном некоторым начальным состоянием s_0 и всевозможными суммами величин δ_i. Отметим, что множество S_0 счетное. Поэтому все функции, определенные на этом множестве, можно рассматривать как последовательности.

Пусть, как и выше, $g_\gamma(s)$ — вероятность вырождения системы из состояния $(x, y), s = x\cos\varphi + y\sin\varphi$ при политике γ.

Тогда для каждого $s \in S_0$ имеем

$g_\gamma(s) = 1$, если $s < 0$ и

$$g_\gamma(s) = \gamma(s)\sum p_i g_\gamma(s + \delta_i) + (1 - \gamma(s))\sum q_i g_\gamma(s + \delta_i), s \geqslant 0.$$

Пусть Γ — множество всех допустимых политик: $\Gamma = \{\gamma(s) \in \{0, 1\} : s \in S_0\}$.

Положим,

$$g(s) = \inf_{\gamma\in\Gamma}\{g_\gamma(s)\}.$$

Таким образом, рассматривается **задача 2**: при фиксированных ценах (определяемых параметром φ) найти политику $\hat\gamma$, при которой для каждого $s \in S_z$

$$g_{\hat\gamma}(s) = g(s).$$

Политику, являющуюся решением **задачи 2**, естественно называть оптимальной политикой. Ясно, что оптимальная политика $\tilde\gamma$ должна удовлетворять равенству

$$g(s) = g_{\tilde\gamma}(s).$$

Теорема 1. Оптимальная политика существует.

Доказательство. Для каждого натурального n и для каждого $s \in S_0$ существует управление γ_n, такое, что

$$0 \leqslant g_{\gamma_n}(s) - g(s) \leqslant \frac{1}{n}.$$

Последовательность $\gamma_n = \{\gamma_n(s), s \in S_0\}$ есть последовательность в l_∞ — пространстве ограниченных последовательностей, которое является сопряженным к пространству абсолютно суммируемых последовательностей l_1. Но единичный шар в l_∞ слабо компактен, поэтому из последовательности γ_n можно выбрать подпоследовательность, слабо сходящуюся к некоторому пределу $\hat{\gamma}$. Этот слабый предел и является искомой оптимальной политикой. Теорема доказана.

Отметим здесь, что может существовать много оптимальных политик.

Очевидно, что функция $g(s)$ при $s \geqslant 0$ должна удовлетворять системе

$$g(s) = \min(\sum p_i g(s+\delta_i), \sum q_i g(s+\delta_i)),$$

если же $s < 0$, то $g(s) = 1$.

Отметим, что g — субрегулярная функция для этой марковской цепи (и при любой выбранной политике).

Удобно ввести множество $I(s) = \{i : s + \delta_i < 0\}$. Тогда уравнения для $g(s)$ принимают вид

$$g(s) = \min\left(\sum_{i \notin I(s)} p_i g(s+\delta_i) + \sum_{i \in I(s)} p_i, \sum_{i \notin I(s)} q_i g(s+\delta_i)) + \sum_{i \in I(s)} q_i\right).$$

Теперь мы можем сформулировать основное требование на согласование цен:

Условие B: цены должны быть такими, чтобы политика, являющаяся оптимальной для **задачи 1**, являлась оптимальной и для **задачи 2**.

Для дальнейшего оказывается полезным рассмотреть два уравнения относительно параметра α, связанных с рассматриваемой марковской цепью:

$$\sum p_i e^{-\alpha\delta_i} = 1, \quad \sum q_i e^{-\alpha\delta_i} = 1. \tag{3}$$

Один из корней каждого из этих уравнений обязательно 0. Поскольку функции $\phi_p(\alpha) = \sum p_i e^{-\alpha\delta_i}$ и $\phi_q(\alpha) = \sum p_i e^{-\alpha\delta_i}$ выпуклые, у каждого из этих уравнений может быть (а может и не быть) еще только один корень.

Всюду в дальнейшем будем предполагать, что эти корни существуют. Эти отличные от 0 корни будем обозначать через α_p и α_q соответственно.

Теорема 2. Для выполнения **условия B** необходимо, чтобы корни α_p и α_q были положительны.

Доказательство. Предположим сначала, что корень $\alpha_p \leqslant 0$, а корень $\alpha_q \geqslant 0$. Поскольку эти корни являются отличными от нуля решениями уравнений (4), то учитывая знаки производных функций ϕ_p и ϕ_q при $\alpha = 0$, получаем неравенства $\sum p_i\delta_i \leqslant 0$ и $\sum q_i\delta_i \geqslant 0$.

Рассмотрим случайные величины ξ и η, принимающие значения δ_i с вероятностями $\{p_i\}$ и $\{q_i\}$ соответственно. Определим случайную цепь $Z(t)$ в каждый момент времени t соотношением $Z(t+1) = Z(t) + \min(\xi, \eta)$.

Пусть $W(t+1) = W(t) + \kappa\xi + (1-\kappa)\eta$, где κ — некоторый числовой параметр между 0 и 1.

Очевидно, что $Z(t) \leqslant W(t)$.

Обозначим через Λ прирост математического ожидания случайной величины $W(t)$ за один шаг:

$$\Lambda = M(W(t+1) - W(t)) = \sum q_i\delta_i - \kappa \sum (q_i - p_i)\delta_i.$$

Поскольку $\sum p_i\delta_i \leqslant 0$, а $\sum q_i\delta_i \geqslant 0$, то можно выбрать параметр $\kappa \in (0,1)$, чтобы Λ стало отрицательным. Например,

$$\kappa = \frac{\sum q_i\delta_i - \frac{1}{2}\sum p_i\delta_i}{\sum (q_i - p_i)\delta_i}.$$

При выбранном таким образом κ получаем, что

$$\Lambda = \frac{1}{2}\sum p_i\delta_i < 0.$$

Воспользуемся теперь усиленным законом больших чисел (теоремой Колмогорова) (см., например, [8], с. 379).

Пусть $\{R_t\}$ — последовательность независимых, одинаково распределенных случайных величин с конечным математическим ожиданием h, $S(t) = \sum\limits_{j=0}^{t} R_j$. Тогда $\dfrac{S_t}{t} \to h$ почти всюду.

Положим, $MR_t = M(W(t) - W(t-1))$. Тогда $MR_t = \Lambda, S_t = W(t)$, и по усиленному закону больших чисел $W(t) \to -\infty$ почти всюду. Поэтому $Z(t) \to -\infty$ почти всюду.

Следовательно, если бесконечно много раз используется управление с набором вероятностей $\{p_i\}$, то система вырождается с вероятностью 1. Но при использовании только управления с набором вероятностей $\{q_i\}$ решение задачи 1 приводит к вырождению с вероятностью 1 за счет выхода из положительного квадранта по одной из координат (поскольку $\sum a_i q_i < 0$).

Если же оба корня отрицательны, то в качестве κ можно взять $\frac{1}{2}$, при этом величина Λ тоже отрицательна.

Использование же только случайного блуждания с набором вероятностей $\{q_i\}$ приводит, как и выше, с вероятностью 1 к выходу из первого квадранта.

Теорема доказана.

Легко получить необходимые и достаточные условия на цены $(\cos\varphi, \sin\varphi)$, гарантирующие положительность корней α_p и α_q. Если в точке $\alpha = 0$ функции $\phi_p(\alpha)$ и $\phi_q(\alpha)$ убывающие, то производные этих функций в точке $\alpha = 0$ отрицательны, а в точках $\alpha = \alpha_p$ и $\alpha = \alpha_q$ эти функции возрастающие. Поэтому

$$\sum p_i\delta_i > 0 \text{ или } \cos\varphi \sum p_i a_i > -\sin\varphi \sum p_i b_i,$$

$$cos\varphi \sum q_i a_i > -\sin\varphi \sum q_i b_i.$$

Учитывая условия (1) ($\sum p_i b_i < 0, \sum q_i a_i < 0$), получаем непустой допустимый диапазон для соотношения цен:

$$\frac{\sum q_i b_i}{-\sum q_i a_i} > tg\varphi > \frac{-\sum p_i b_i}{\sum p_i a_i}, \tag{4}$$

при котором соответствующие этим ценам корни α_p и α_q положительны.

Если же в точке $\alpha = 0$ производная одной из функций $\phi_p(\alpha)$ и $\phi_q(\alpha)$ положительна, то соответствующий этой функции второй корень должен быть отрицательным.

В качестве верхней оценки для вероятности вырождения $g(s)$ можно взять решение одной из двух систем (без учета минимума):

$$g_1(s) = \sum_{i \notin I(s)} p_i g_1(s+\delta_i) + \sum_{i \in I(s)} p_i,$$

$$g_2(s) = \sum_{i \notin I(s)} q_i g_2(s+\delta_i) + \sum_{i \in I(s)} q_i,$$

поскольку для каждого s справедливы неравенства $g(s) \leqslant g_i(s), i = 1, 2.$

Асимптотика решения таких систем хорошо известна. Она приведена, например, в [7, с. 479]. Из этой асимптотики следует существование постоянной k, такой, что

$$g_1(s) \leqslant ke^{-s\alpha_p}, \quad g_2(s) \leqslant ke^{-s\alpha_q}.$$

Не ограничивая общности, можно считать, что $0 < \alpha_q < \alpha_p$. Тогда для $g(s)$ справедлива верхняя оценка

$$g(s) \leqslant g_1(s) \leqslant ke^{-s\alpha_p}.$$

Теорема 3. Существуют постоянные k_1 и k_2, такие, что при $s \to \infty$ справедливы неравенства

$$k_1 e^{-s\alpha_p} \leqslant g(s) \leqslant k_2 e^{-s\alpha_p}.$$

Оценка сверху уже есть. Для завершения доказательства построим оценку снизу.

Заметим сначала, что если $\alpha_q < \alpha_p$, то

$$\sum q_i e^{-\alpha_p \delta_i} > 1.$$

Действительно, функция $\phi(\alpha) = \sum q_i e^{-\alpha\delta_i}$ выпуклая и равна единице в двух точках 0 и α_q, причем $\alpha_q > 0$. Поэтому эта функция принимает значения меньше единицы только в интервале $(0, \alpha_q)$. Поскольку $\alpha_q < \alpha_p$, то $\phi(\alpha_p) > 1$.

Положим, $d(s) = g(s)e^{s\alpha_p}$.

Тогда, если $I(s) = \{i : s + \delta_i < 0\}$, то

$$d(s) = \min\left(\sum_{i \notin I(s)} p_i^* d(s+\delta_i) + \sum_{i \in I(s)} p_i^* e^{\alpha_p(s+\delta_i)}, \sum_{i \notin I(s)} q_i d(s+\delta_i) e^{-\alpha_p\delta_i} + \sum_{i \in I(s)} q_i e^{\alpha_p(s+\delta_i)}\right),$$

Как было показано выше, $B = \sum_i q_i e^{-\delta_i\alpha_p} > 1$.

Пусть $r_i = \frac{q_i e^{-\delta_i\alpha_p}}{B}$. Ясно, что $\sum_i r_i = 1$. Тогда

$$d(s) = \min\left(\sum_{i \notin I(s)} p_i^* d(s+\delta_i) + \sum_{i \in I(s)} p_i^* e^{\alpha_p(s+\delta_i)}, B\sum_{i \notin I(s)} r_i d(s+\delta_i) + B\sum_{i \in I(s)} r_i e^{\alpha_p(s+\delta_i)}\right).$$

Уменьшив B (заменив на 1), получаем

$$d(s) > \min\left(\sum_{i \notin I(s)} p_i^* d(s+\delta_i) + \sum_{i \in I(s)} p_i^* e^{\alpha_p(s+\delta_i)}, \sum_{i \notin I(s)} r_i d(s+\delta_i) + \sum_{i \in I(s)} r_i e^{\alpha_p(s+\delta_i)}\right).$$

Наборы $\{p_i^*\}$ и $\{r_i\}$ — вероятностные.

Можно показать, что марковская цепь, порожденная случайными блужданиями с этими вероятностями, приводит к вырождению системы с вероятностью 1. Это замечание приводит к построению оценки снизу для соответствующей вероятности вырождения.

Рассмотрим $f(s)$ — решение системы

$$f(s) = \min\left(\sum_{i \notin I(s)} p_i^* f(s+\delta_i) + \sum_{i \in I(s)} p_i^* e^{\alpha_p(s+\delta_i)}, \sum_{i \notin I(s)} r_i f(s+\delta_i) + \sum_{i \in I(s)} r_i e^{\alpha_p(s+\delta_i)}\right).$$

Покажем, что $d(s) \geqslant f(s)$.

Действительно, пусть

$$f_1(s) = \min\left(\sum_{i \notin I(s)} p_i^* d(s+\delta_i) + \sum_{i \in I(s)} p_i^* e^{\alpha_p(s+\delta_i)}, \sum_{i \notin I(s)} r_i d(s+\delta_i) + \sum_{i \in I(s)} r_i e^{\alpha_p(s+\delta_i)}\right).$$

Ясно, что $d(s) \geqslant f_1(s)$. Это неравенство справедливо для любого s. Заменив в системе для f_1 все $d(s+\delta_i)$ на $f_1(s+\delta_i)$, получим

$$f_1(s) \geqslant f_2(s) = \min\left(\sum_{i \notin I(s)} p_i^* f_1(s+\delta_i) + \sum_{i \in I(t)} p_i^* e^{\alpha_p(s+\delta_i)}, \sum_{i \notin I(s)} r_i f_1(s+\delta_i) + \sum_{i \in I(s)} r_i e^{\alpha_p(s+\delta_i)}\right).$$

Повторяя этот процесс, для каждого s получаем монотонно убывающую последовательность $f_n(s)$, сходящуюся к $f(s)$, которая и является решением соответствующей системы.

Пусть

$$c = \min_{t: I(s) \neq \emptyset} \min_{i \in I(s)} e^{\alpha_p(s+\delta_i)}.$$

Заметим, что минимумы в этих формулах достигаются и являются строго положительными, поскольку набор возможных значений i при каждом s конечен, а замыкание множества s, таких, что $I(s) \neq \emptyset$, является компактом.

Тогда решение системы

$$v(s) = \min\left(\sum_{i \notin I(s)} p_i^* v(s+\delta_i) + c \sum_{i \in I(s)} p_i^*, \sum_{i \notin I(s)} r_i v(s+\delta_i) + c \sum_{i \in I(s)} r_i\right)$$

меньше, чем для предыдущей. По теореме единственности решение $v(s) = c$ и справедливо неравенство $d(t) \geqslant f(t) \geqslant c$.

Это неравенство и завершает доказательство теоремы 3.

3. О согласованных ценах. При оптимальном управлении в **задаче 2** (с ценами) минимизируется вероятность попадания случайной величины из точки (x, y) в область $\Omega = \{(z, v) : z \cos\varphi + v \sin\varphi < 0\}$. При этом было показано, что при достаточно больших x и y эту вероятность можно оценить с двух сторон

$$c_2 e^{-\lambda_p(x\cos\varphi + y\sin\varphi)} \leqslant g(x\cos\varphi + y\sin\varphi) =\leqslant c_1 e^{-\lambda_p(x\cos\varphi + y\sin\varphi)}.$$

При оптимальном управлении в **задаче 1** минимизируется вероятность g(x,y) попадания из точки (x, y) в область $\Delta = \{(z, v) : z < 0\} \cup \{(z, v) : v < 0\}$, при этом

$$c_4 e^{-(\alpha x + \beta y)} \leqslant g(x, y) \leqslant c_3 e^{-(\alpha x + \beta y)}.$$

Здесь c_1, c_2, c_3, c_4 — некоторые положительные постоянные.

Но $\Omega \subset \Delta$. Отсюда следует, что при любой фиксированной политике (в частности, при политике, оптимальной для **задачи 1**) вероятность попадания из любой точки (x, y) в множество Ω не больше вероятности попадания из этой точки в множество Δ. Поэтому

$$\lambda_p \cos\varphi \leqslant \alpha; \lambda_p \sin\varphi \leqslant \beta.$$

Пусть при некоторой цене $(\cos\varphi, \sin\varphi)$ выполняются строгие неравенства.

Это означает, при таком наборе цен любая оптимальная „ценовая“ политика приводит к вырождению исходной системы (к выходу из первого квадранта по одной из координат) с вероятностью 1. Набор цен, при котором существует „ценовая“ политика, не приводящая к автоматическому вырождению исходной задачи, будем называть согласованными ценами.

Легко видеть, что для таких цен справедливы формулы

$$\cos\varphi = \frac{\alpha}{\sqrt{\alpha^2+\beta^2}},\ \sin\varphi = \frac{\beta}{\sqrt{\alpha^2+\beta^2}}, \lambda_p = \lambda_q = \sqrt{\alpha^2+\beta^2},$$

где пара (α, β) — решение ассоциированной системы.

Вернемся к **задаче 1**. Для исследования поведения траектории в этой задаче в работе [1] было введено понятие пожарного управления, обобщение которого — **c**-пожарное управление — было предложено в [3]. Опишем это управление.

Пусть $\mathbf{c} = (c_1, c_2)$ — произвольный вектор с положительными компонентами. Управление γ будем называть **c**-пожарным, если в каждой целочисленной точке (x, y) первого квадранта

$$\gamma(x,y) = 1, \text{если} \frac{x}{c_1} \geqslant \frac{y}{c_2},$$

$$\gamma(x,y) = 0, \text{если} \frac{x}{c_1} < \frac{y}{c_2}.$$

Отметим, что, не сильно ограничивая общность, можно считать величины c_1 и c_2 несоизмеримыми, т. е. на прямой $\frac{x}{c_1} = \frac{y}{c_2}$ есть только одна целочисленная точка (0,0).

Пожарное управление — частный случай **c**-пожарного управления при $\mathbf{c} = (1, 1)$.

Обозначим через (X_n, Y_n) траекторию, возникающую в **задаче 1** при **c**-пожарном управлении.

Траектории, возникающие при таком управлении, изучались в [3].

Оказалось, что если процесс не вырождается, то $\frac{Y_n}{X_n} \to \frac{c_2}{c_1}$. Переходя к направляющим косинусам, получаем, что

$$\frac{Y_n}{X_n} \to tg\varphi,$$

где $\cos\varphi = \frac{c_1}{\sqrt{c_1^2+c_2^2}}, \quad \sin\varphi = \frac{c_2}{\sqrt{c_1^2+c_2^2}}$.

Если положить $X_n = t_n \cos\varphi_n, Y_n = t_n \sin\varphi_n$, то $tg\varphi_n \to tg\varphi$.

Было показано, что вероятность вырождения $R(X_n, Y_n)$ при **c**-пожарном управлении имеет вид

$$R(X_n, Y_n) = O(e^{-t_n(\alpha\cos\varphi + \beta\sin\varphi)}).$$

Тем самым, для минимизации асимптотического поведения этой вероятности следует выбирать угол φ (и связанное с этим углом **c**-пожарное управление) так, чтобы максимизировать выражение $\alpha\cos\varphi + \beta\sin\varphi$.

Поэтому

$$\cos\varphi = \frac{\alpha}{\sqrt{\alpha^2+\beta^2}},\ \sin\varphi = \frac{\beta}{\sqrt{\alpha^2+\beta^2}}.$$

Таким образом, согласованные цены $(\cos\varphi, \sin\varphi)$ и порождают оптимальное **c**-пожарное управление.

О выборе из нескольких блужданий. В предыдущих разделах изучался случай, когда число случайных блужданий совпадало с числом продуктов.

Пусть теперь число случайных блужданий превышает число продуктов, т. е. $k > 2$. Имеется k производств, каждое из которых характеризуется своей переходной матрицей: для производства j переходную матрицу будем обозначать через $\{p_{i,j}\}, i = 1, \cdots, n,\ j = 1, \cdots, k$. Будем предполагать, что каждое из производств является более предпочтительным к одному из продуктов, то есть для каждого j выполняется одно из пар неравенств

$$\sum_i b_i p_{i,j} < 0,\ \sum_i (a_i + b_i) p_{i,j} > 0$$

(„лучшее“ для первого продукта), или

$$\sum_i a_i p_{i,j} < 0,\ \sum_i (a_i + b_i) p_{i,j} > 0$$

(„лучшее“для второго продукта).

1. Рассмотрим сначала один простой случай: блуждания, являющиеся „хорошими“для одного продукта, таковы, что вероятность ухудшения по этому продукту равна нулю, а по второму продукту ухудшение с ненулевой вероятностью может произойти только на единицу. Это означает, что в качестве векторов (a_i, b_i) рассматриваются такие $(i, 0), (i, -1), (0, i), (-1, i)$. Для блужданий, „хороших“ для первого продукта, ненулевыми являются лишь вероятности перехода из точки (x, y) в точки $(x + i, y)$ (вероятности перехода p_i^1) или $(x+i, y-1)$ (вероятности p_i^2). Для блужданий, „хороших“для второго продукта, ненулевыми являются лишь вероятности перехода из точки (x, y) в точки $(x, y + i)$ (вероятности q_i^1) или $(x-1, y+i)$ (вероятности q_i^2). Тем самым условия предпочтительности соответственно принимают вид

$$\sum_i i p_i^1 + \sum_i (i-1) p_i^2 > 0$$

или

$$\sum_i i q_i^1 + \sum_i (i-1) q_i^2 > 0.$$

Для дальнейшего нам потребуется следующее утверждение:

Лемма 1.

Пусть $\sum (s_i + r_i) = 1, r_i \geqslant 0, s_i \geqslant 0$ и $\sum i s_i + \sum (i-1) r_i > 0$

Тогда функция $\mu(\lambda)$, определяемая из уравнения

$$\sum_{i=0}^{n} s_i e^{-i\lambda} + e^{\mu} \sum_{i=0}^{n} r_i e^{-i\lambda} = 1,$$

является возрастающей и выпуклой, причем $\mu'(0) > 1$.

Доказательство. Опуская для простоты пределы суммирования, легко видеть, что

$$\mu = ln(1 - \sum s_i e^{-i\lambda}) - ln(\sum r_i e^{-i\lambda})$$

$$\mu' = \frac{\sum i s_i e^{-i\lambda}}{1 - \sum s_i e^{-i\lambda}} + \frac{\sum i r_i e^{-i\lambda}}{\sum r_i e^{-i\lambda}}.$$

Очевидно, что $\mu' > 0$, т. е. функция $\mu(\lambda)$ — возрастающая. Заметим, кроме того, что $\mu'(\lambda) \to 0$ при $\lambda \to \infty$.

Покажем, что ее производная убывает. Действительно, с ростом λ числитель первого слагаемого убывает, а знаменатель возрастает, поэтому первое слагаемое убывает. Рассмотрим подробнее второе слагаемое. Докажем, что с ростом λ оно тоже убывает. Взяв его производную по λ, покажем, что она отрицательна.

Это утверждение сводится к неравенству

$$-\sum i^2 r_i e^{-i\lambda} \sum r_i e^{-i\lambda} + (\sum i r_i e^{-i\lambda})^2 < 0.$$

Докажем его по индукции. Обозначив для удобства $t_i = r_i e^{-i\lambda}(t_i > 0)$, сведем неравенство к следующему:

$$(\sum i t_i)^2 < \sum i^2 t_i \sum t_i.$$

Для двух слагаемых получаем очевидное неравенство

$$(t_1 + 4t_2)(t_1 + t_2) > (t_1 + 2t_2)^2.$$

Требуется доказать индукционный переход — следующее неравенство

$$(\sum i^2 t_i + (n+1)^2 t_{n+1})(\sum t_i + t_{n+1}) > (\sum i t_i + (n+1) t_{n+1})^2.$$

Раскрывая скобки, приводим наше неравенство к следующему

$$\sum i^2 t_i \sum t_i + \sum i^2 t_i t_{n+1} + (n+1)^2 t_{n+1} \sum t_i + (n+1)^2 t_{n+1}^2 >$$
$$> (\sum i t_i)^2 + 2(\sum i t_i)(n+1) t_{n+1} + (n+1)^2 t_{n+1}^2.$$

По индукционному предположению первое слагаемое слева больше, чем первое слагаемое справа. Поэтому остается доказать, что

$$\sum i^2 t_i t_{n+1} + (n+1)^2 t_{n+1} \sum t_i > 2(\sum i t_i)(n+1) t_{n+1}.$$

Разделив это неравенство на t_{n+1}, получаем, что оно сводится к очевидному неравенству

$$\sum t_i (n+1-i)^2 > 0.$$

Значит, оба слагаемых в формуле для $\mu'(\lambda)$ монотонно убывают, следовательно, функция $\mu(\lambda)$ выпуклая.

Покажем теперь, что $\mu'(0) > 1$. Действительно,

$$\mu'(0) = \frac{\sum i s_i}{1 - \sum s_i} + \frac{\sum i r_i}{\sum r_i}.$$

Но, по предположению, $\sum i r_i + \sum i s_i > \sum r_i$. Заменяя числитель второй дроби на меньшее число, получаем

$$\mu'(0) > \frac{\sum is_i}{1-\sum s_i} + \frac{\sum r_i - \sum is_i}{\sum r_i} = \frac{\sum is_i}{1-\sum s_i} + 1 - \frac{\sum is_i}{\sum r_i} = \sum is_i\Big(\frac{1}{1-\sum s_i} - \frac{1}{\sum r_i}\Big) + 1 = 1,$$

поскольку $\sum(s_i + r_i) = 1$.

Лемма доказана.

Если рассмотреть ассоциированную систему для пары блужданий из разных групп, то она имеет вид

$$\sum_i p^1_{i,j} e^{-i\lambda} + e^{\mu} \sum_i p^2_{i,j} e^{-i\lambda} = 1$$

$$\sum_i q^1_{i,j} e^{-i\mu} + e^{\lambda} \sum_i q^2_{i,j} e^{-i\mu} = 1.$$

Здесь наборы вероятностей $\{p^1_i\}$ и $\{p^2_i\}$ являются „хорошими“ для первого продукта, а $\{q_i\}$ и $\{q^2_i\}$ — для второго продукта. В этом случае условие, что блуждание является более предпочтительным к первому продукту, принимает вид

$$\sum_i (i-1) p^2_{i,j} > 0,$$

поскольку векторы (a_i, b_i) имеют вид $(i, 0)$ или $(i, -1)$.

Покажем, что такая система (кроме точки (0,0)) имеет только одно дополнительное решение.

По лемме 1 из первого уравнения этой системы получаем возрастающую и выпуклую функцию $\mu(\lambda)$, а из второго — возрастающую и выпуклую функцию $\lambda(\mu)$. Функция $\nu(\lambda)$, обратная к функции $\mu(\lambda)$, является возрастающей и вогнутой, причем $\nu'(0) < 1$. Заметим, что при $\mu = 0$ справедливы неравенства

$$-\sum_i i(q^1_{i,j} + q^2_{i,j}) + \lambda'(0) \sum_i q^2_{i,j} = 0.$$

Поскольку это уравнение соответствует блужданию, „хорошему“ для второго продукта, то $\lambda'(0) > 0$.

Аналогично, при $\lambda = 0$ справедливы неравенства

$$-\sum_i i(p^1_{i,j} + p^2_{i,j}) + \mu'(0) \sum_i p^2_{i,j} = 0.$$

Поскольку это уравнение соответствует блужданию, „хорошему“ для первого продукта, то $\lambda'(0) > 1$.

$$\mu'(0) = \frac{\sum_i i(p^1_{i,j} + p^2_{i,j})}{\sum_i p^2_{i,j}}.$$

Действительно, рассмотрим функции $\mu(\lambda)$ и $\nu(\lambda)$ — обратную к функции $\lambda(\mu)$. Обе эти функции возрастающие, и в точке (0,0) равны нулю. Но по лемме 1 $\mu(\lambda)$ выпуклая, $\mu'(0) > 1$, а $\nu(\lambda)$ — вогнутая, $\nu'(0) < 1$. Покажем теперь, что имеется еще только одна положительная точка пересечения графиков этих функций. Так как $\mu'(0) > 1$, а $\nu'(0) < 1$, то в некоторой окрестности точки 0 справедливо неравенство

$$\mu(\lambda) > \nu(\lambda).$$

Предположим, что это неравенство справедливо для всех λ. Тогда хорда, соединяющая точки (0,0) и $(\lambda, \mu(\lambda))$, лежит выше хорды, соединяющей точки (0,0). Но при $\lambda \to \infty$ имеем $\mu'(\lambda) \to 0$, а $\nu'(\lambda) \to \infty$. Поэтому угловой коэффициент первой хорды стремится к 0, а второй — растет. Это противоречие и показывает, что кривые пересекаются, при этом, вследствие выпуклости одной функции и вогнутости другой, точка пересечения единственная.

Тем самым задача сведена в этом случае к варианту с двумя случайными блужданиями, и асимптотика получается из предыдущих результатов.

Пусть число случайных блужданий, как и выше, превышает число продуктов (по-прежнему предполагаем, что продуктов только два), но при этом блуждания уже общего вида.

Здесь монотонность не сохраняется, поэтому приходится использовать другой подход, связанный с ценами на продукты.

Как и выше, все блуждания разбиваются на две группы, „хорошие“ по отношению к одному из продуктов. Целесообразно рассматривать только такие пары, в которых имеется по представителю из каждой группы.

Пусть $(\cos\varphi, \sin\varphi)$ — цены на продукты. Для каждой пары блужданий (по одной из каждой группы) мы можем рассмотреть блуждание на прямой, порожденное этими ценами.

Будем для каждой пары (i, j) рассматривать только согласованные цены, поскольку только они связаны с оптимальным управлением.

Асимптотика вырождения при с-политике, порожденной согласованными ценами и парой блужданий (i, j), определяется одномерными параметрами $\lambda_{i,j}$, которые можно упорядочить.

Есть пара блужданий из разных групп. Имеются два корня $\alpha_p(\varphi)$ и $\alpha_q(\varphi)$. Эти корни могут быть разными, но если цены согласованы с данной парой, то корни одинаковы.

Предположим противное: $\alpha_p(\varphi) < \alpha_q(\varphi)$. Тогда поведение Q выгоднее поведения P, оптимальное решение — принимать только поведение, связанное с Q, что приводит к вырождению по одной из компонент.

В [5] было показано, что для согласованных цен угол φ должен выбираться так, чтобы выполнялись равенства

$$\alpha_p = \alpha_q = \sqrt{\mu^2 + \lambda^2}.$$

Здесь пара (λ, μ) — решение системы, возникающей для соответствующей пары блужданий.

Тем самым, среди множества пар следует выбирать такое, что $\sqrt{\mu^2 + \lambda^2}$ максимально.

Список литературы

1. Radner R., Rothschild M. On the allocation of effort // J. Economic Theory. 1975. V. 10. N 3. P. 358–376.

2. Рапопорт Э. О. Об одной стохастической модели распределения неделимого ресурса // Труды Сибирской конференции по прикладной и индустриальной математике. Новосибирск: Ин-т математики СО РАН, 1997. С. 197–206.

3. Рапопорт Э. О. Магистральные стратегии при распределении неделимого ресурса // Дискретный анализ и исследование операций. Серия 1. Т. 4, № 1, 1997, С. 33–45.

4. Рапопорт Э. О. Об одной модели распределения неделимого ресурса // Дискретный анализ и исследование операций. Серия 2. Т. 12. № 1. 2005, С. 55–73.

5. Рапопорт Э. О. Распределение неделимого ресурса: оптимальное управление и цены // Сибирский журнал индустриальной математики. Том XII. № 3(39). 2009. С. 75–84.

6. Кемени Дж., Снелл Дж., Кнепп А. Счетные цепи Маркова. М.: Наука, 1987. С. 3–398.

7. Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, 1964. Т. 2. С. 3–734.

8. Ширяев А.Н. Вероятность. М.: Наука, 1980. С. 10–561.

Рапопорт Эрнест Ошерович — канд. физ.-мат. наук,
доцент Института математики им. С. Л. Соболева СО РАН,
раб. тел. 363-46-06, e-mail: rapoport@math.nsc.ru.

Дата поступления — 27.08.2014