Управляемые системы, ИМ,ИК СО АН СССР 1971 г. выпуск 9.
УДК.519.95.

ВОПРОСЫ ОПТИМИЗАЦИИ ДИСКРЕТНЫХ СТОХАСТИЧЕСКИХ СИСТЕМ

И.З.Фельдман /Москва/

Доказываются необходимые и достаточные условия оптимальности для линейных дискретных стохастических систем с фазовыми ограничениями. Предлагается метод решения задач, основанный на доказанных условиях. Метод иллюстрируется примером.

Для эффективного использования математических методов в экономических исследованиях очень важно строить и анализировать модели, описывающие изучаемые явления, с учетом вероятностного характера данных. Вопросы теории оптимальных стохастических систем применительно к экономическим и организационным задачам разработаны еще недостаточно полно. Весьма существенными в этих задачах оказываются ограничения, накладываемые на управление и на фазовые координаты, описывающие динамику экономического процесса. Кроме того, при разработке математического аппарата нужно учитывать большую размерность, сопутствующую, как правило, экономическим и организационным задачам.

Применение известных методов математической теории оптимальных процессов [1 - 3] для изучения вероятностных экстремальных задач требует построения стохастического аналога принципа максимума Понтрягина. Вероятностные системы, для которых этот аналог был получен, не включали в себя фазовые ограничения [4 - 5]. На возможность применения итеративных методов для решения задач с фазовыми ограничениями указывалось в работах [6] и [7]. Важность рассмотрения дискретных систем показана в [6].

В данной работе рассматриваются вопросы оптимизации дискретных линейных стохастических систем, включающих в себя фазовые ограничения. Доказаны необходимые и достаточные условия оптимальности, и на основе этих условий разработан метод решения задач рассматриваемого класса, учитывающий вероятностный характер данных. Для построения метода проводится анализ вероятностных ограничений и исследуются многогранники ограничений. В качестве иллюстрации решается задача об оптимальной работе склада со случайным спросом.

1. Постановка задачи. Опишем состояние системы в любой дискретный момент времени $t \in \Theta(T)$ $(\Theta(T)=\{0,1,\ldots,T\})$ фазовым вектором $x(t,\omega)$, представляющим собой набор r вещественных случайных величин $x_i(t,\omega)$ $(i=0,1,\ldots,r-1)$, заданных на вероятностном пространстве (Ω, σ, P), где элементарное событие ω принадлежит пространству элементарных событий Ω. Случайные величины $x_i(t,\omega)$ в любой момент времени t предполагаются измеримыми относительно вероятностной меры P и имеющими математическое ожидание $Mx_i(t,\omega)$.

Соответственно $Mx(t,\omega)$ составлено из набора $Mx_i(t,\omega)$, так что $Mx(t,\omega)=(Mx_0(t,\omega), Mx_1(t,\omega), \ldots Mx_{r-1}(t,\omega)$.

Детерминированный ограниченный вектор $u(t)$, принадлежащий k - мерному векторному пространству E^k при любом $t\in\theta(T-1)$, назовем управлением. Будем рассматривать случай, при котором в любой момент времени $t_0, t_0\in\theta(T-1)$, решение о выборе управления в этот момент $u(t_0)$ принимается после того, как становится известной реализация $x(t_0)$ случайного вектора $x(t_0,\omega)$. Это дает возможность использовать знание текущего состояния системы и искать управление в виде

$$u(t)=(u(x(t,\omega))/x(t,\omega)=x(t))=u(x(t)).$$

Введем две последовательности статистически независимых случайных векторов $\xi(t,\omega)$ и $\eta(t,\omega)(\omega\in\Omega, t\in\theta(T-1))$. Функции распределения координат этих векторов-случайных величин $\xi_i(t,\omega)$ и $\eta_i(t,\omega)$ в любой фиксированный момент времени будем считать известными. Изменение фазового вектора $x(t,\omega)$ опишем системой конечно - разностных уравнений

$$x(t+1,\omega)=Ax(t,\omega)+Bu(x(t))+\xi(t,\omega) \qquad /I.I/$$

при начальном условии

$$x(0)=C_0. \qquad /I.2/$$

На управление и фазовый вектор наложены вероятностные ограничения

$$P((Q^i x(t,\omega)+R^i u(x(t,\omega))+b_i-\eta_i(t,\omega)\geq 0)/x(t))\geq$$
$$\geq\alpha_i(t);\ 0\leq\alpha_i(T)\leq 1;\quad t\in\theta(T-1);\ i=0,1,\ldots,\ell, \qquad /I.3/$$

где $\alpha_i(t)$ - заданные детерминированные функции, а на управление $u(x(t))$ ограничения

$$K^i u(x(t))-d_i\geq 0;\ i=0,1,\ldots,\ell_2. \qquad /I.4/$$

Соотношения вида $P((f(t,\omega)\geq 0/x(t))$ и $M((g(t,\omega)/x(t))$ означают вероятность и математическое ожидание при условии, что в момент времени t $x(t,\omega)$ принимает значение $x(t)$.

Условие /I.3/ удобно записать в иной форме

$$(Q^i x(t,\omega)+R^i u(x(t,\omega))+b_i-F_i^{-1}(\alpha_i)\geq 0/x(t)),$$

где F_i^{-1} -функция, обратная функции распределения случайной величины $\eta_i(t,\omega)$ в момент времени t.

Критерием качества системы будем считать функционал

$$J=M(cx(T,\omega)). \qquad /I.5/$$

В условиях /I.I/ - /I.5/ векторы $\xi(t,\omega)$ и $\eta(t,\omega)$ представляют собой набор r и ℓ_1 случайных величин $\xi_i(t,\omega)$ и $\eta_i(t,\omega)$; A, B, Q, R, K - постоянные матрицы размерностей соответственно $(r\times r), (r\times k), (\ell_1\times r), (\ell_1\times k), (\ell_2\times k)$; $\ell_1+\ell_2=\ell$; b, d, c' - постоянные векторы размерности ℓ_1, ℓ_2 и r /штрих здесь и в дальнейшем означает транспонирование/, Q^i, R^i, K^i - строки соответствующих матриц.

Случайный вектор $x(t,\omega)$ и ограниченный измеримый вектор $u(x(t,\omega))$, удовлетворяющие условиям /1.1/ - /1.4/, назовем допустимыми фазовым вектором и управлением. Пусть $u^1(x(t))$ и $u^2(x(t))$ - два допустимых управления, соответствующие одному и тому же значению фазового вектора. Разность $u^1(x(t)) - u^2(x(t)) = \delta u(x)$ назовем допустимой вариацией. Сформулируем задачу оптимизации.

З а д а ч а А. Определить последовательность допустимых управлений $u(x(t))$, которая обеспечивала бы

$$\min J = \min M(cx(T,\omega)).$$

Решение задачи "А", т.е. допустимое управление, минимизирующее /1.5/, назовем оптимальным и обозначим

$$u^0(x^0(t)).$$

Сформулированная задача характерна для экономических проблем. На общность такой постановки указывалось, в частности, в работах [6 - 7], где проводились некоторые возможные экономические интерпретации задачи. Заметим, что ограничения вида /1.3/ соответствуют нежесткой постановке задач линейного стохастического программирования при $\alpha_i < 1$ и жесткой постановке при $\alpha_i = 1$. В соответствии со смыслом ограничений /1.3/ функции $\beta_i(t)$ такие, что

$$\beta_i(t) = 1 - \alpha_i(t) \quad (i = 0, 1, \ldots, \ell)$$

будем называть функциями риска.

2. Необходимые и достаточные условия оптимальности.

Получим условия оптимальности, аналогичные условиям принципа максимума. Введем сопряженную случайную последовательность n -мерных вектор-строк $\psi(t,\omega)$. Введем, кроме того, ℓ_1 -мерную последовательность вектор-строк $\mu(t,\omega)$ такую, что $\mu_i(t,\omega) \geqslant 0$, $i = 0, 1, \ldots, \ell_1$. Изменение последовательности $\psi(t,\omega)$ опишем разностным векторным уравнением

$$\psi(t,\omega) = \psi(t+1,\omega)A - \mu(t+1,\omega)Q. \qquad /2.1/$$

при выполнении условия

$$\psi(T-1,\omega) = c. \qquad /2.2/$$

Л е м м а 1. Положим в задаче "А", что $\xi(t,\omega)$ и $\eta(t,\omega)$ принимают только детерминированные значения $\xi(t)$ и $\eta(t)$; $\alpha_i(t) \equiv 1$, $T = 1$.

Для того, чтобы допустимое управление $u^0(x(0))$ было оптимальным, необходимо и достаточно существование вектор-строк $\mu(t)$ и $v(t)$ размерности ℓ_1 и ℓ_2 таких, что при

$$t = 0 \qquad /2.3/$$

выполняются условия

$$\mu_i(t) \geqslant 0, \quad i = 0, 1, \ldots, \ell_1, \qquad /2.4/$$

$$v_i(t) \geqslant 0, \quad i = 0, 1, \ldots, \ell_2, \qquad /2.5/$$

$$\mu(t)(Qx(t) + Ru(x(t)) + b - \eta(t)) = 0, \qquad /2.6/$$

$$v(t)(Ku(x(t)) - d) = 0 \qquad /2.7/$$

и справедливо равенство

$$\varphi(t)B = \mu(t)R + v(t)K. \qquad /2.8/$$

Доказательство. Определим в векторном пространстве E^K норму с помощью равенства

$$\|u\| = (u', u)^{1/2}.$$

Получим, что E^K есть пространство типа В. Тогда из существования линейных непрерывных неотрицательных функционалов $<\mu>$ и $<v>$ [8] и в силу самосопряженности конечномерного векторного пространства E^K и общего вида линейных функционалов на нем [9] следует, что в любой фиксированный момент времени t допустимое $u(x(t))$, удовлетворяющее условиям /2.4/ -/2.8/ обеспечивает $\min(\varphi(t)Bu(x(t)))$. Такое управление будем обозначать $u^*(x(t))$. Так как $x^0(0)=x(0)$, то при $t=0$ для любого допустимого $u(x^0(0))$ выполняется условие

$$\varphi(0)(Ax^0(0) + Bu^*(x^0(0)) + \xi(0)) = \min_u \varphi(0)(Ax^0(0) + Bu(x^0(0)) + \xi(0)).$$

Утверждение леммы доказано, $u^*(x(0)) = u^0(x^0(0))$ /здесь $\varphi(0)=C$/. Заметим, что утверждение леммы I следует и из общей теории линейного программирования, в частности, /2.6/ и /2.7/ соответствуют условиям "дополняющей нежесткости".

Лемма 2. Положим в задаче "А", что $\xi(t,\omega)$ и $\eta(t,\omega)$ принимают только детерминированные значения $\xi(t)$ и $\eta(t)$; $\alpha_i(t) \equiv 1$. Рассмотрим допустимые $x^0(t)$, $u^0(x^0(t))$, $x(t)$, $u(x(t))$. Если $\mu(t), v(t), x^0(t), u^0(x^0(t))$ удовлетворяют условиям /2.4/ - /2.8/, $t = 0, 1, \ldots, T-1$, то

$$\varphi(t)(x(t+1) - x^0(t+1)) \geqslant \varphi(t-1)(x(t) - x^0(t)). \qquad /2.9/$$

Доказательство. Согласно доказательству леммы I управление $u^0(x^0(t))$ в любой фиксированный момент времени минимизирует линейную форму $\varphi(t)Bu(x^0(t))$. Кроме того, существует допустимое управление $u^*(x(t))$, удовлетворяющее условиям /2.4/-/2.8/ и минимизирующие линейную форму $\varphi(t)Bu(x(t))$. Вектор $x(t+1)$, соответсвующий $x(t)$ и $u^*(x(t))$, обозначим $x^*(t+1)$. Тогда

$$\begin{aligned}
&\Psi(t)(x(t+1) - x^0(t+1)) \geqslant \Psi(t)(x^*(t+1) - x^0(t+1)) = \\
&= \Psi(t)(Ax^*(t) - Ax^0(t)) + \Psi(t)(Bu^*(x(t)) - Bu^0(x^0(t))) = \\
&= \Psi(t)(Ax^*(t) - Ax^0(t)) + \mu(t)(Ru^*(x(t)) - Ru^0(x^0(t))) + \\
&+ v(t)(Ku^*(x(t)) - Ku^0(x^0(t))) = \Psi(t)(Ax^*(t) - \\
&- Ax^0(t)) - \mu(t)(Qx(t) - Qx^0(t)).
\end{aligned}$$

Так как

$$\mu(t)(Qx(t) - Qx^0(t)) + \mu(t)(Ru^*(x(t)) - Ru^*(x^0(t))) = 0,$$

то имеем

$$\Psi(t)(x(t+1)-x^0(t+1)) \geqslant \Psi(t)(x^*(t+1)-x^0(t+1)) =$$
$$= \Psi(t-1)(x(t)-x^0(t)),$$

что и требовалось доказать.

Следствие I. В условиях леммы 2 при выполнении дополнительного условия

$$QBu(x^0(t-1)) \geqslant QBu^0(x^0(t-1)).$$

Оптимальное управление $u^0(x^0(t))$ допустимо совместно с $x(t)$, т.е. $Qx(t)+Ru^0(x^0(t))+b-\eta(t) \geqslant 0$.

При оговоренных выше условиях $Qx(t) \leqslant Qx^0(t)$, откуда легко следует утверждение следствия I.

Следствие 2. В условиях следствия I существует допустимое управление $u^*(x(t))$ такое, что

$$\Psi(t)Bu^*(x(t)) \leqslant \Psi(t)Bu^0(x^0(t)) \qquad /2.10/$$

Обозначим $U^k(x(t))$ и $U^k(x^0(t))$ множества, к которым принадлежат допустимые управления $u^*(x(t))$ и $u(x^0(t))$. Так как на основании леммы I и следствия I $U^k(x^0(t)) \subset U^k(x(t))$, то

$$\min_{u\in U^k(x(t))} \Psi(t)Bu(t) \leqslant \min_{u\in U^k(x^0(t))} \Psi(t)Bu(t),$$

что и доказывает утверждение следствия 2.

Заметим, что неравенство /2.10/ характерно именно для задачи с фазовыми ограничениями и противоречит условиям оптимальности в отсутствие их.

Лемма 3. Пусть $x(t,\omega)$ и $u(x(t,\omega))$ допустимы а $u^0(x^0(t_0))$-оптимально в задаче "А", и допустимое $u(x(t))$ удовлетворяет условиям

$$u(x(t)) = u^0(x^0(t)), \qquad t \neq t_0 \in \Theta(T-2),$$
$$u(x(t_0)) = u^0(x^0(t_0)) + \delta u(x(t_0)).$$

Тогда

$$M(cx(T,\omega) - cx^0(T,\omega)/x^0(t_0,\omega)) \geqslant 0. \qquad /2.11/$$

Доказательство. Пусть неравенство /2.11/ не выполняется. Тогда существует множество X такое, что

$$P(x^0(t_0,\omega)) \in X) = P_X > 0,$$

и существует допустимое управление $u(x(t,\omega))$ и определяемая им траектория $x(t,\omega)$ такие, что

$$M(cx(T,\omega) - cx^0(T,\omega)/x^0(t_0) = x(t_0)) < 0.$$

Построим управление $u^*(x(t))$ и соответсвующий ему фазовый вектор следующим образом

$$u^*(x(t)) = \begin{cases} u^0(x^0(t)); & t < t_0 \\ u(x(t)); & t \geqslant t_0;\ x^0(t_0) \in X \\ u^0(x^0(t)); & t \geqslant t_0;\ x^0(t_0) \notin X \end{cases}$$

Тогда

$$M(cx(T,\omega)-cx^{\circ}(T,\omega)) = P_x M(cx(T,\omega)-cx^{\circ}(T,\omega))/ x^{\circ}(t_0,\omega)\in X) + (1-P_x) M(cx(T,\omega)-cx^{\circ}(T,\omega))/ x^{\circ}(t_0,\omega)\notin X) .$$

Так как

$$M(cx(T,\omega)/x^{\circ}(t_0,\omega)\notin X) = M(cx(T,\omega)/x(t_0,\omega) = x^{\circ}(t_0,\omega)\notin X),$$

то из предположения от противного следует отрицательность левой части /2.12/, что противоречит оптимальности управления $u^{\circ}(x^{\circ}(t))$. Полученное противоречие доказывает лемму. Утверждение леммы 3 аналогично принципу оптимальности Беллмана.

Т е о р е м а I. /Необходимые и достаточные условия оптимальности/. Для того,чтобы допустимое управление $u^{\circ}(x^{\circ}(t))$ было оптимальным, необходимо и достаточно выполнение условия

$$M(\Psi(t,\omega)Bu^{\circ}(x^{\circ}(t,\omega))/x^{\circ}(t,\omega)) \leq \leq M(\Psi(t,\omega)Bu(x^{\circ}(t,\omega))/x^{\circ}(t,\omega)) \qquad /2.13/$$

для всех допустимых $u(x^{\circ}(t))$ в любой момент времени $t\in\Theta(T-1)$, причем $\mu(t,\omega)$ таково, что справедливо равенство

$$\Psi(t,\omega)B = \mu(t,\omega)R + \nu(t,\omega)K .$$

Д о с т а т о ч н о с т ь. Пусть $u(x(t))$ - произвольное допустимое управление, не совпадающее с $u^{\circ}(x^{\circ}(t))$.Рассмотрим выражение

$$H(t) = M(\Psi(t,\omega)(x(t+1,\omega)-x^{\circ}(t+1,\omega)/x(t),x^{\circ}(t)).$$

Так как при любом значении t выполняется равенство

$$M(x(t+1,\omega)-x^{\circ}(t+1,\omega)/x(t),x^{\circ}(t)) = = (Ax(t,\omega)-Ax^{\circ}(t,\omega)+Bu(x(t,\omega))- - Bu^{\circ}(x^{\circ}(t,\omega)/x(t),x^{\circ}(t)),$$

то, учитывая /2.8/, имеем

$$H(t) = (\Psi(t-1,\omega)(x(t,\omega)-x^{\circ}(t,\omega)+\mu(t,\omega)\times \times Qx(t,\omega)-Qx^{\circ}(t,\omega))+\Psi(t,\omega)(Bu(x(t,\omega))- - Bu^{\circ}(x^{\circ}(t,\omega))/x(t),x^{\circ}(t)).$$

Рассмотрев выражение

$$H(0)+M(H(1))+M(H(2))+M(H(3))+\ldots+M(H(t))+\ldots+ + M(H(T-1)/x(T-2),x^{\circ}(T-2))/x(T-3)\ldots/x(1),x^{\circ}(1)/C_0),$$

получим

$$M(cx(T,\omega)-cx^{\circ}(T,\omega)/x(0)) = M\Big(\sum_{t=1}^{T-1}(\Psi(t,\omega)(Bu(x(t,\omega))- \qquad /2.14/$$

$$-Bu^\circ(x^\circ(t,\omega))+\mu(t,\omega)(Qx(t,\omega)-Qx^\circ(t,\omega)/x(t),$$
$$x^\circ(t))/x(0))+\psi(0,\omega)(Bu(0)-Bu^\circ(0)).$$

Для любых реализаций $\mu(t)$ и $\Psi(t)$ случайных вектор - строк $\mu(t,\omega)$ и $\Psi(t,\omega)$ на основании леммы 2 и условия /2.13/ выполняется неравенство

$$(Bu(x(t,\omega)-Bu^\circ(x^\circ(t,\omega))+\mu(t)(Qx(t,\omega)-$$
$$-Qx^\circ(t,\omega))/x(t),x^\circ(t))\geqslant 0.$$

Следовательно правая часть /2.14/ неотрицательна, и для произвольного допустимого $x(T,\omega)$

$$M((cx(T,\omega)-cx^\circ(T,\omega))/x(0))\geqslant 0,$$

что и требовалось доказать.

Необходимость. Утверждение теоремы для $t=T-1$ непосредственно следует из леммы I и леммы 3. Докажем необходимые условия для произвольного момента времени $t_0\in\Theta(T-1)$. Предположим противное. Пусть существует допустимое управление $u(x(t_0))$ такое, что

$$M((\psi(t,\omega)(Bu(x^\circ(t_0))-Bu^\circ(x^\circ(t_0)))/x^\circ(t_0))<0.$$

Для любых реализаций допустимых $x(t,\omega),x^\circ(t,\omega)$ и произвольной реализации $\Psi(t)$ случайной вектор - строки $\Psi(t,\omega)$ введем минимизирующие управления $u^*(x(t))$ и $u^*(x^\circ(t))$. Согласно лемме 2 и в силу произвольности выбора реализаций

$$M(cx(T,\omega)-cx^\circ(T,\omega)/x(t_0))\leqslant M(cx(T,u^*(x(T-1))-$$
$$-cx(T,u^*(x^\circ(T-1))/x(T-1))/x(t_0))\leqslant M(\Psi(t_0,\omega)\times$$
$$\times(x(t_0+1,\omega)-x^\circ(t+1,\omega)/x(t_0))=M(\Psi(t_0,\omega)\times$$
$$\times(Bu(x^\circ(t_0))-Bu^\circ(x^\circ(t_0))/x^\circ(t_0)).$$

Следовательно, на основании предположения от противного,

$$M((cx(T,\omega)-cx^\circ(T,\omega))/x^\circ(t_0))\leqslant 0.$$

Полученное неравенство противоречит оптимальности $x^\circ(T,\omega)$ ввиду леммы 3. Таким образом, необходимые и достаточные условия оптимальности доказаны.

3. Метод дополнительных ограничений. Аналогично условиям принципа максимума начальные условия для фазового вектора $x(t,\omega)$ и сопряженной вектор - строки $\Psi(t,\omega)$ заданы в различные моменты времени. Для построения последовательности оптимальных управлений на основе полученных условий оптимальности рассмотрим некоторый метод, учитывающий вероятностный характер условий и использующий линейность задачи.

На допустимые управления накладываются ограничения

$$((Qx(t,\omega)+Ru(x^\circ(t,\omega))+\tilde{b})/x(t))\geqslant 0, \qquad /3.1/$$

$$((Ku^\circ(x^\circ(t,\omega))+d\ )/x(t,\omega))=x(t))\geq 0, \qquad /3.2/$$

где $\widehat{b}_i = b_i - F^{-1}(\alpha_i)$. Каждую строку этих неравенств будем обозначать символом z_i, где $i\in\mathcal{L};\mathcal{L}=(1,2,\ldots,\ell)$. Пересечение конечного числа замкнутых полупространств $z_i\geq 0$ задает в пространстве E^k замкнутый выпуклый многогранник $U^k(x(t))$ [8]. Пусть совокупность I из k строк z_i такова, что существует единственное решение системы

$$z_i=0;\ i\in\mathcal{L}_I\ ;\ \mathcal{L}_I\subset\mathcal{L} \qquad /3.3/$$

относительно $u(x(t))$. Тогда совокупность I определяет некоторую вершину многогранника $U^k(x(t))$, которую обозначим $A(t)$. Подставив решение системы /3.3/ в условия

$$z_i\geq 0;\ i\in(\mathcal{L}\,|\,\mathcal{L}_I\ ), \qquad /3.4/$$

получим систему равенств - неравенств относительно координат фазового вектора $x(t,\omega)$. Если условия /3.4/ совместны, то вершину $A(t)$ назовем квазидопустимой, в противном случае - запрещенной. Если вершина $A(t)$ квазидопустима, то система /3.4/ имеет решение.

$$M(x(t))+m\geq 0, \qquad /3.5/$$

где матрица M имеет размерность $(\ell_3\times r);\ell_3\leq\ell-k$, а вектор m - размерность ℓ_3. Неравенства /3.5/ - решение системы /3.4/ - будем называть условиями реализации вершины $A(t)$. Будем говорить, что квазидопустимая вершина $A(t)$ допустима в момент времени t, если $x(t)$ удовлетворяет условиям ее реализации. В противном случае вершину $A(t)$ будем называть недопустимой. Каждой квазидопустимой вершине многогранника $U^k(x(t))$ в любой момент времени t в пространстве управлений E^k сопоставим присоединенный конус Ω, образованный нормалями к гиперплоскостям $(z_i=0;\ i\in\mathcal{L}_I)$, причем нормали направлены в сторону $(z_i=0;\ i\in\mathcal{L}_I)$. Если оптимальное значение управления $u(x(t))$ достигается в вершине $A(x(t))$, то, согласно пункту 2, вектор $(\varphi(t,\omega))'/x(t,\omega)$ принадлежит присоединенному конусу этой вершины. В этом случае вершину $A(t)$ будем называть оптимальной.

Согласно условиям оптимальности для определения управления $u^\circ(x^\circ(t))$ в каждый момент времени $t\in\theta(T-1)$ решается локальная оптимизационная задача "В"(t) : определить допустимое управление, минимизирующее $M(\psi(t,\omega)Bu(x^\circ(t,\omega))/x^\circ(t,\omega))$. Значение управления $u^\circ(x^\circ(t))$ зависит от реализации того или иного из возможных видов многогранника $U^k(x(t))$ /геометрически подобные многогранники принадлежат одному и тому же виду/. Реализация определенного вида многогранника зависит от реализации соответствующей вершины, т.е. от выполнения условий /3.5/. Условия реализации некоторого вида многогранника $U^k(x(t))$ в момент времени t накладывают дополнительные ограничения на управление. Так как эти условия зависят от $\xi(t-1,\omega)$, то дополнительные ограничения имеют вероятностный характер. Обозначим индексом $P(D^j)$ условную вероятность выполнения дополнительных ограничений, соответствующих j - му виду многогранника:

$$P(D^j)=P(D_i^j \geqslant 0 \;/x(t,\omega)),\; i=1,2,\ldots .$$

С помощью функции распределения $F(\xi)$ случайной величины $\xi(t)$ вероятность $P(D^j)$ может быть представлена как функция $x(t,\omega)$ и $u(x(t,\omega))$

$$P(D^j)=P(x(t),u(x(t))).$$

Введение вероятностей выполнения дополнительных ограничений позволяет определить условное математическое ожидание сопряженной вектор-строки $\Psi(t,\omega)$. Пусть в момент времени $t+1$ вектор - строка $\Psi(t+1,\omega)$ принимает некоторое значение $\Psi^i(t+1)$. При фиксированном значении $\Psi^i(t+1)$ каждому j - му типу многогранника однозначно соответствуют значения $\mu^{ji}(t+1)$ и $\Psi^{ji}(t)$, определяемые с помощью условия

$$\Psi^{ji}(t)=\Psi^i(t+1)A-\mu^{ji}(t+1)Q \qquad /3.2/$$

Определим, кроме того, вероятность

$$P(\Psi(t+1,\omega)=\Psi^i(t+1,\omega))=P^i(t+1)/x(t+1,\omega)).$$

Учитывая введенные вероятности, получим

$$M(\Psi(t,\omega)/x(t,\omega))=\sum_j\sum_i P^iP^{ji}(\Psi^i(t+1)A-\mu^{ij}(t+1)Q).$$

Следовательно, задача "В" $/(t)$, с учетом соотношения

$$M(\Psi(t,\omega)Bu(x^0(t,\omega))/x^0(t,\omega))=M(\Psi(t,\omega)/$$
$$/x^0(t,\omega))Bu(x^0(t,\omega)),$$

запишется в детерминированной форме:

Задача "С". Определить допустимое управление $u^0(x^0(t))$ из условия минимизации

$$\sum_i\sum_j P^i(D^j)P^i(\Psi^i(t+1)A-\mu^{ij}(t+1)Q).$$

Таким образом, на основе полученных условий оптимальности и с помощью введения вероятностей реализации типа многогранника, оптимальная стохастическая задача "А" свелась к последовательному решению локальных детерминированных задач "С"

4. Задача о складе со случайным спросом. Рассмотрим модель распределительного склада [8], [10]. В дискретные моменты времени на склад поступают заявки, величина которых носит стохастический характер. Поступившая заявка удовлетворяется или ставится в очередь. Свои запасы склад пополняет за счет поступлений с базы, соответствующих запросам склада, но размеры поставляемой партии ограничены. Критерием оптимальности служит минимум суммарных издержек при работе склада, возникающих за счет штрафов при невыполнении за-

казов и платы за хранение на складе запасов. Последовательность опрераций на складе следующая: отсылается запрос на базу, заказанная партия поступает с базы, фиксируется заявка, определяется величина задолженности или запасов. Так как решение о величине запроса на базу принимается до того, как становится известной величина заявки, а решение о величине задолженности после, разделим моменты принятия этих решений фиктивным промежутком времени Δ . Сделаем следующие предположения:

1/ случайные величины, описывающие спрос, статистически независимы, а их функции распределения известны;

2/ величины штрафа и платы за хранение пропорциональны задолженности и величине запасов соответственно;

3/ заказанная на базе партия поступает мгновенно.

Примем следующие обозначения:

$x_1(t,\omega)$ и $x_2(t,\omega)$ - величины запасов на складе и задолжности в момент времени t до удовлетворения заявок; $x_1(t+\Delta,\omega)$ и $x_2(t+\Delta,\omega)$ - те же величины после удовлетворения заявок;

$\xi(t+\Delta,\omega)$ - величина заявок в момент времени t ;

$F(t+\Delta)$ - функция распределения случайной величины $\xi(t+\Delta)$;

$u_1(t)$ - величина поступлений с базы в момент t ;

$u_2(t)$ - величина задолженности, погашаемой в момент t ;

$u_3(t+\Delta)$ - величина задолженности, образующейся в момент t ;

a - плата за хранение единичного количества запасов;

b - штраф за невыполнение единичной заявки.

Введем переменную $x_0(t,\omega)$ пс правилу

$$x_0(t+\Delta)=x_0(t,\omega)+x_1(t+\Delta,\omega)+\frac{b}{a}x_2(t+\Delta,\omega);\ x(0)=0.$$

Значение $x_0(t+\Delta,\omega)$ в момент времени $T-\Delta$ с точностью до постоянного положительного множителя a совпадает с величиной суммарных издержек при работе склада за период времени $(0,T)$.

При сделанных предположениях задача минимизации суммарных издержек эквивалентна следующей задаче линейного динамического стохастического программирования: Определить

$$u_1(t),u_2(t),u_3(t),u_1(t+\Delta),u_2(t+\Delta),u_3(t+\Delta)$$

из условия

$$\min M(x_0(T,\omega)) \Big/$$

$$x_0(t+\Delta)=x_0(t,\omega)+x_1(t,\omega)+rx_2(t,\omega)+u_1(t)-(r+1)u_2(t),$$
$$x_1(t+\Delta,\omega)=x_1(t,\omega)+u_1(t)-u_2(t),$$
$$x_2(t+\Delta,\omega)=x_2(t,\omega)-u_2(t),$$
$$x_0(t+1,\omega)=x_0(t+\Delta,\omega)+x_1(t+\Delta,\omega)+rx_2(t+\Delta,\omega)-$$
$$-\xi(t+\Delta,\omega)+(r+1)u_3(t+\Delta),$$
$$x_1(t+1,\omega)=x_2(t+\Delta,\omega)+u_3(t+\Delta),$$

$$x_2(t+1,\omega)=x_2(t+\Delta,\omega)+u_3(t+\Delta),$$
$$x(t+\Delta)\geqslant 0,\ x(t+1)=0,\ u(t)=0,\ u(t+\Delta)\geqslant 0,\ u_1(t)\leqslant U(t),$$
$$x_0(0)=0,\ x_1(0)=C,\ x_2(0)=0,\ t=0,1,\ldots,T-1;\ r=\frac{b}{a}=2).$$

Будем решать задачу описанным выше методом. Положим $T=2$, так как рассмотрение этого значения T достаточно для выяснения особенностей предлагаемого метода и характера получаемого с его помощью решения.

Запишем локальную задачу "В"$(1+\Delta)$. Определить $u^0(1+\Delta)$ из условия $\min M(\Psi(1+\Delta)B_1u/u(1+\Delta)\geqslant 0,\ u_3(1+\Delta)\geqslant\xi(1+\Delta,\omega)-x_1(1+\Delta;\omega);$ $x_1(1+\Delta,\omega)$, где $\Psi(1+\Delta)=(1,0,0)$,

а
$$B_1=\begin{pmatrix}0&0&r+1\\0&0&1\\0&0&1\end{pmatrix}$$

Обозначим разность $\xi(1+\Delta,\omega)-x_1(1+\Delta,\omega)=D_1(1+\Delta)$. Решение задачи "В"$(1+\Delta)$ запишется как
$$u_1^0(1+\Delta)=0,$$
$$u_2^0(1+\Delta)=0,$$
$$u_3^0(1+\Delta)=\begin{cases}\xi(1+\Delta,\omega)-x_1(1+\Delta,\omega); & D_1\geqslant 0;\\ 0; & D_1<0.\end{cases}$$

Вектор строка $\mu(1+\Delta,\omega)$, определяемая из условия
$$\Psi(1+\Delta,\omega)B=\mu(1+\Delta,\omega)R+\nu(1+\Delta,\omega)K$$
равна
$$\mu(1+\Delta,\omega)=\begin{cases}(0;\ 0;\ r+1); & D_1(1+\Delta)\geqslant 0;\\ (0;\ 0;\ 0); & D_1(1+\Delta)<0.\end{cases}$$

Соответсвенно сопряженная вектор - строка $\Psi(1,\omega)$
$$\Psi(1;\omega)=\begin{cases}\Psi^1=(1;-r;\ r); & D_1(1+\Delta)\geqslant 0\\ \Psi^2=(1;\ 1;\ r); & D_1(1+\Delta)<0.\end{cases}$$

Дополнительное ограничение $D_1(1+\Delta)$, рассматриваемое в момент времени $t=1$ как функция $x(1,\omega)$ и $u(1)$ имеет вероятностный характер, причем
$$P(D_1(1+\Delta)\geqslant 0)=1-F_1(x_1(1,\omega)+u_1(1)-u_2(1)),$$
где F_1 - функция распределения случайной величины $\xi(1+\Delta,\omega)$ Таким образом,

$$M(\Psi(1,\omega)/x(1,\omega))=(1-F_1(x_1(1,\omega)+u_1(1)-u_2(1)))\times$$
$$\times(1;-r;\ r)+F_1(x_1(1,\omega)+u_1(1)-u_2(1))(1;\ 1;\ r).$$

Оптимальное управление $u^0(\mathrm{I})$ находится при решении задачи "В"(1) Определить $u(1)$ из условия:

$$\min M(\Psi(1,\omega)/x_1(1,\omega))B_2u$$
$$x_2(1,\omega)-u_2(1)\geqslant 0,\ x_1(1)+u_1(1)-u_2(1)\geqslant 0,$$

где

$$B = \begin{pmatrix} 1 & -(r+1) & \bar{0} \\ 1 & -1 & 0 \\ 0 & -1 & 0 \end{pmatrix} .$$

Многогранник ограничений на управление в момент времени $t=1$ изображен на рис.I.

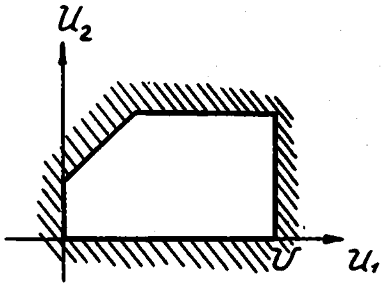

Рис. I

Положим, что случайные величины $\xi(t,\omega)$ распределены равномерно и

$$F_0(y) = F_1(y) = \begin{cases} y/20 \; ; & 0 \le y \le 20 \; ; \\ 0 \; ; & y < 0 \; ; \\ 1 \; ; & y > 20 \; . \end{cases}$$

Тогда

$$u_1^0 = \begin{cases} \frac{10}{3} + x_2(1) - \frac{x_1(1)}{2} \; ; & D_2(1) \geq 0 \\ 0 \; ; & D_2(1) < 0 \end{cases}$$

$$u_2^0 = x_2(1) \; ,$$

где

$$D_2(1) = \frac{10}{3} - \frac{x_3(\Delta,\omega)}{2} + \frac{x_2(\Delta,\omega)}{2} + \frac{\xi(\Delta,\omega)}{2} + \frac{u_3(\Delta)}{2}$$

Заметим, что локальная задача "В" (1) является нелинейной, для определения же $\mu(1,\omega)$ решается линейная задача. Вектор-строка $\mu(1,\omega)$ равна

$$\mu(1,\omega) = \begin{cases} (0; \; 3; \; 0) \; ; & \psi(1,\omega) = \psi^1(1) \\ (2; \; 4; \; 0) & \psi(1,\omega) = \psi^2(1) \end{cases}$$

В момент времени $t=2$ $\psi(\Delta,\omega)$ равно

$$\psi(\Delta,\omega) = \begin{cases} \psi^1(\Delta,\omega) = (1; -1; \; 1) \; ; D_1(\Delta) \geq 0 \; , \\ \psi^2(\Delta,\omega) = (1; \; 0; \; 0) \; ; D_1(\Delta) < 0 \; , \end{cases}$$

Причем $P(\psi(\Delta,\omega) = \psi^1(\Delta)) = P(D_1(\Delta) \geq 0 / D_2(\Delta) \geq 0) \times$

$$\times P(D_2(\Delta) \geq 0) + P(D_1(\Delta) \geq 0 / D_2(\Delta) < 0) P(D_2(\Delta) < 0) =$$
$$= 1 - P(\xi(1+\Delta) + \tfrac{1}{2}\xi(\Delta) < \tfrac{1}{2} x_1(\Delta,\omega) + \tfrac{1}{2} u_3(\Delta) + \tfrac{10}{3}) \times$$
$$\times P(\xi(\Delta) < x_1(\Delta) - 2x_2(\Delta) - u_3(\Delta) - \tfrac{20}{3}) = P_1(\Delta) .$$

Аналогично определяется и

$$P(\psi(\Delta,\omega)=\psi^2(\Delta))=P_2(\Delta)$$

При решении нелинейной задачи "В"(Δ) оптимальное управление равно

$$u_1^0=0$$

$$u_2^0=0 \quad u_3^0=\begin{cases}\xi(\Delta)-x_1(\Delta); & D_3(\Delta)\geq 0\\ 0; & D_3(\Delta)<0;\ D_3(\Delta)=\xi(\Delta)-x_1(\Delta)\end{cases}$$

Аналогично предыдущему определяется $\psi(0,\omega)$

$$\psi(0,\omega)=\begin{cases}\psi^1(0)=(1;\ 0;\ 0\ ); & D_1\geq 0;\ D_3\geq 0\\ \psi^2(0)=(1;\ 0;\ 3); & D_1\geq 0;\ D_3<0\\ \psi^3(0)=(1;\ 1;-1); & D_1<0;\ D_3\geq 0\\ \psi^4(0)=(1;\ 1;\ 2); & D_1<0;\ D_3<0\end{cases}$$

и вероятности $P(\psi(0,\omega)=\psi^j(0))$. Оптимальное управление $u^0(0)$ находится при решении задачи "В"(0) : определить $u^0(0)$ из условия

$$\min(P_1(0)(1;0;0)+P_2(0)(1;0;3)+P_3(0)(1;1;-1)+P_4(0)(1;1;2))\times$$
$$\times B_2u(0);\ u(0)\geq 0,\ u_2(0)\leq x_2(0),\ u_1(0)-u_2(0)+x_1(0)\geq 0.$$

Оптимальное значение управления $u^0(0)$ равно

$$u^0(0)=\begin{cases}(\bar u;0;\ 0)'; & \bar u\leq U(0),\\ (U,\ 0;\ 0)'; & \bar u>U(0),\end{cases}$$

где $\bar u$ корень уравнения

$$\frac{37}{12}-\frac{(c+u_1)}{90}+\frac{(c+u_1)^2}{40}-\frac{3(c+u_1)^3}{32000}=0$$

Полученное решение изображено на рис. 2.

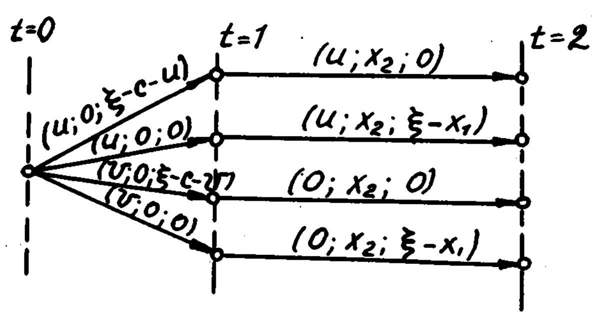

Рис. 2

При оптимальной работе склада и постоянной величине штрафа и платы за хранение запасы на складе в любой момент времени допустимы только при отсутствии долгов; оптимальная величина запросов на базу определяется последовательностью $u^0(t)$. При переменных ценах и при больших значениях T метод решения не изменяется.

В заключение считаю своим долгом выразить благодарность Ю.П. Кривенкову за помощь и внимание к работе и профессору Ю.В.Гермейеру за полезные советы и обсуждение.

Поступила в редакцию 20.12.1970 г.

Литература

1. Л.С.Понтрягин,В.Г.Болтянский, Р.В.Гамкрелидзе,Е.Ф.Мищенко, Математическая теория оптимальных процессов. М.,ГИФМЛ, 1961 г.

2. Л.И.Розоноэр.Принцип максимума Л.С.Понтрягина в теории оптимальных систем. Автоматика и телемеханика, 1959, т.XX,№10,11,12.

3. В.П.Аноров. Принцип максимума для процессов с ограничениями общего вида. Автоматика и телемеханика, 1967 г., т.XX, №11,№3,4.

4. Kushner H.I. Schweppe F.S. A maxmin principle for stochastik control systems. G. Math. and Appl., 1964,v.8, N 2.

5. Svorcher D.P. Optimal Control of Discretetime Stochastic Systems. G, Math. and Appl.,1969,v.15,N2.

6. Ю.М.Ермольев. Об оптимальном управлении случайным процессом, "Кибернетика" 1970 г. № 2.

7. В.М.Ефимов. Исследование стохастических экстремальных задач при помощи функционального анализа. "Кибернетика", 1970 г. № 3.

8. Ю.П.Кривенков. Математические и вычислительные вопросы линейного динамического программирования, Изд.Вычисл.Центра АН СССР, М. 1969.

9. Л.А.Люстерник, В.И.Соболев. Элементы функционального анализа, "Наука", 1965 г.

10. С.Н.Кащеев. Оптимальное управление запасами в динамических системах со случайным спросом. Автореферат диссертации М., ЦЭМИ АН СССР, 1967 г.