УДК: 517.977.54

MSC2010: 91A10

ИСХОД И РИСК В МНОГОШАГОВОЙ ПОЗИЦИОННОЙ ЗАДАЧЕ ПРИ НЕОПРЕДЕЛЕННОСТИ

© В. И. Жуковский, А. С. Горбатов

Московский государственный университет имени М. В. Ломоносова
Факультет вычислительной математики и кибернетики
кафедра оптимального управления
Ленинские горы, МГУ, ВМК, ГСП-1, Москва, 119991, Российская Федерация
e-mail: *zhkvlad@yandex.ru, gorbatovanton@gmail.com*

© М. И. Высокос

Государственный гуманитарно-технологический университет
Физико-математический факультет
кафедра математики и физики
ул. Зеленая, 22, Орехово-Зуево, 142611, Российская Федерация
e-mail: *mvysokos@mail.ru*

Outcome and Risk in a Multi-step Positional Problem under Uncertainty.

Zhukovskiy V. I., Vysokos M. I. and Gorbatov A. S.

Abstract. In single-criteria problem under strategical uncertainty from the point of view of DM tasks of decision making are examined. DM tries to increase the guaranteed outcome with possible smaller guaranteed risk. We are based on the principle of minimax regret (Savage-Nichans) with the help of mathematical apparatus of the method of dynamic programming for discrete problems.

First, we examine single-criteria problem of two forms which differs by pairs: contrstrategy — pure uncertainty and pure strategy — strategical uncertainty.

In the problem of the first type, a hierarchical procedure of formation the counterstrategies of DM is used. The discrimination of uncertainty appears: the decision-maker learns the uncertainty which has realized and only then he chooses his strategy, namely for example for single-criteria problem under uncertainty

$$\Gamma = \langle \Sigma, \mathbf{U}_Z, \mathbf{Z}, J(U_Z, Z, x_0)\rangle$$

$$\Sigma \div x(k+1) = f(k, x(k), u(k, x(k), z(k, x(k)))),\ \ x(0) = x_0\ k = 0, ..., K,$$

$$U_Z = (U_Z(0), U_Z(1), ..., U_Z(K-1)) \div (u(0, x, z), u(1, x, z), ..., u(K-1, x, z)),$$

$$Z = (Z(0), Z(1), ..., Z(K-1)) \div (z(0, x), ..., z(K-1, x)),$$

$$\mathbf{U}_Z = \{U_Z\},\ \ \mathbf{Z} = \{Z\}.$$

$$J(U_Z, Z, x_0) = \Phi(x(K)) + \sum_{k=0}^{K-1} F_i(k, x(k), u[k], z[k]),$$

the functions of risk are defined

$$R(U_Z, Z, x_0) = \max_{U_Z \in \mathbf{U}_Z} J(U_Z, Z, x_0) - J(U_Z, Z, x_0).$$

In the problem of the second type strategic uncertainties are used. They are formed on the assumption of discrimination DM, who transmits them to the researcher selected by him pure strategy for the formation of strategic uncertainty.

In the problem about uncertainties only boundaries of changes are known (any probabilistic characteristics are absent).

In the first case the regret function is constructed, in the second case — guarantee of outcome and risk. Secondly, it is offered the initial single -criteria problem to set in accordance two-criteria discrete positional problem, where the first criteria — the guaranteed outcomes, the second one — "minus" guaranteed risk. For this two-criteria problem we construct the Pareto maximal pure strategy, which defines the value of guaranteed outcomes and guaranteed risk accompanying the realization of guaranteed outcome.

As the example, the explicit form of suggested solution for linear-quadratic one step variant of single-criteria problem is obtained.

Keywords: *multi-step problem, management problem strategy, multicriteria problem, guarantees, Pareto maximum.*

1. Постановка задачи

Рассматриваем однокритериальную многошаговую позиционную задачу при неопределенности.

Функция риска Севиджа–Ниханса.

Здесь ограничимся интервальными неопределенностями. Для построения функции риска рассматриваем однокритериальную многошаговую оптимизационную задачу

$$\Gamma_Z = \langle \mathbf{\Sigma}, \mathbf{U}_Z, \mathbf{Z}, J(U_Z, Z, x_0) \rangle,$$

изменение управляемой системы $\mathbf{\Sigma}$ описывается векторным разностным уравнением

$$\begin{aligned} x(k+1) = f(k, x(k), u(k, x(k), z(k, x(k))), z(k, x(k))), \\ x(0) = x_0, (k = 0, 1, ..., K-1) \end{aligned} \tag{1}$$

(момент окончания управления) постоянную $K > 0$ фиксируем; в (1) фазовый $n-$вектор (вектор состояния) в момент времени k есть $x(k) \in \mathbb{R}^n$, пара $(k, x(k))$ — позиция в момент k в задаче Γ_Z, $(0, x_0)$ — начальная позиция; стратегия ЛПР (лица, принимающего решения) в момент времени k, именно $U_Z(k)$ отождествляется с m-вектор-функцией $u(k, x, z) \in \mathbb{R}^m$, этот факт обозначается $U_Z(k) \div u(k, x, z)$. В теории дифференциальных игр $U_Z(k)$ называется контрстратегией в момент времени k и множество контрстратегий в момент времени k обозначим $\mathbf{U}_Z(k)$. Стратегией ЛПР в Γ_Z является набор

$$U_Z = (U_Z(0), U_Z(1), ..., U_Z(K-1)) \div (u(0,x,z), u(1,x,z), ..., u(K-1,x,z)),$$

множество таких контрстратегий U_Z обозначим $\mathbf{U}_Z$.

Согласно теории дифференциальных позиционных игр, задача Γ_Z относится к категории *минимаксных.* В ней применяем в момент $k = 0, ..., K-1$ *чистую неопределенность* $Z(k) \div z(k,x)$, $z \in \mathbb{R}^S$, множество таких неопределенностей в момент k обозначим $\mathbf{Z}(k) = \{Z(k)\}$, а сама неопределенность Z тогда определится набором

$$Z = (Z(0), Z(1), ..., Z(K-1)) \div (z(0,x), z(1,x), ..., z(K-1,x)),$$

множество которых обозначим $\mathbf{Z} = \prod\limits_{k=0}^{K-1} \mathbf{Z}(k)$.

При формировании контрстратегии $U_Z(k)$ в момент k используем *иерархическую* процедуру. Именно на k-м ходу неопределенность $Z(k) \div z(k,x)$ поступает к ЛПР. Лицо, принимающее решение, на этой основе и исходя из некоторых соображений (например, оптимизации своего критерия, об этом далее) выбирает в момент времени k свою стратегию $U_z(k) \div u(k,x,z)$; как раз эти вектор-функции $u(k,x,z)$ и $z(k,x)$ и фигурируют в (1).

В этом проявляется *дискриминация неопределенности*: при формировании своей стратегии ЛПР становится известной реализовавшаяся в момент k неопределенность $Z(k) \div z(k,x)$.

Перейдем к *динамике* задачи Γ_Z. В момент $k = 0$ ЛПР узнает неопределенность $Z(0) \div z(0,x)$ и формирует свою стратегию $U_Z(0) \div u(0, x_0, z(0,x_0))$. С помощью системы (1) при $k = 0$ находит

$$x(1) = f(0, x_0, u(0, x_0, z(0, x_0)), z(0, x_0)).$$

В следующий момент $k = 1$ ЛПР, получив информацию о реализовавшейся чистой неопределенности $Z(1) \div z(1,x)$, формирует свою стратегию $U(1) \div u(1,x,z(1,x))$ и опять-таки с помощью (1) при $k = 1$ находит

$$x(2) = f(1, x(1), u(1, x(1), z(1, x(1))), z(1, x(1))),$$

затем процедура с очевидными изменениями повторяется до момента $k = K-1$ и в этот момент K уже ЛПР находит на основе (1)

$$x(K) = f(K-1, x(K-1), u(K-1, x(K-1), z(K-1, x(K-1))), z((K-1), x(K-1))).$$

В результате получаем дискретную траекторию

$$x_0, x(1), ..., x(K),$$

порожденную ей реализацию выбранных ЛПР стратегий

$$u[0] = u(0, x_0, z(0, x_0)), u[1] = u(1, x(1), z(1, x(1))), ..., u[K-1] = \\ = u(K-1, x(K-1), z(K-1, x(K-1))),$$

и реализовавшихся, появившихся независимо от действий ЛПР, последовательности неопределенностей

$$z[0] = z(0, x_0),\ z[1] = z(1, x(1)), ..., z[K-1] = z(K-1, x(K-1)).$$

На полученных таким образом трех последовательностях:

$$\{x(k)\}_{k=0}^{K},\ \{u[k]\}_{k=0}^{K-1},\ \{z[k]\}_{k=0}^{K-1}$$

определен критерий, оценивающий качество поведения ЛПР (результат от выбранной им стратегии $U_Z = (U_Z(0), ..., U_Z(K-1)))$, который задается функционалом

$$J(U_Z, Z, x_0) = \Phi(x(K)) + \sum_{k=0}^{K-1} F(k, x(k), u[k], z[k]), \tag{2}$$

первое слагаемое в (2) называется *терминальным*, а второе — *интегральным.*

На содержательном уровне цель ЛПР состоит в выборе стратегии $U_Z^* \in \mathbf{U}_Z$, при которой критерий (2) принимает как можно большее (максимальное) значение; при этом ЛПР учитывает возможность реализации любой чистой неопределенности $Z \in \mathbf{Z}$.

Наконец перейдем к построению самой функции риска Сэвиджа–Ниханса [2, 3]. Для этого следует найти

$$\max_{U_Z \in \mathbf{U}_Z} J(U_Z, Z, x_0) = J(U_Z^*, Z, x_0) = J[Z, x_0]$$

при любых $Z \in \mathbf{Z}$, $x_0 \in \mathbb{R}^n$. Далее, формировать указанную функцию риска по формуле

$$R(U, Z, x_0) = J[Z, x_0] - J(U, Z, x_0). \tag{3}$$

Гарантии исхода и риска

При построении гарантий будем следовать подходу, предложенному в недавно опубликованной серии работ [5, 6]. Для этого в задаче приходится использовать так называемые *стратегические неопределенности* Z_U, а сама Γ_Z изменится на

$$\Gamma_U = \langle \mathbf{\Sigma}, \mathbf{U}, \mathbf{Z}_U, J(U, Z_U, x_0) \rangle. \tag{4}$$

Здесь употребляемая система $\mathbf{\Sigma}$ также имеет вид (1), но вместо $\mathbf{U}_Z$ используем множество чистых позиционных стратегий в момент k, именно вида $U(k) \div u(k,x)$, множество их есть $\mathbf{U}(k)$, а сама позиционная стратегия становится

$$U = (U(0), ..., U(K-1)) \div (u(0,x), ..., u(K-1,x)),$$

множество таких чистых стратегий в (4) обозначено через $\mathbf{U}$. Стратегические неопределенности в момент k уже будут $Z_U(k) \div z(k,x,u)$. Они формируются в предположении *дискриминации* ЛПР, которое на первом ходу передает для формирования неопределенности $Z_U(k)$ в момент k выбранную им стратегию $U(k) \div u(k,x)$. Сама стратегическая неопределенность Z для случая (4) представляется рядом

$$Z_U = (Z_U(0), ..., Z_U(k-1)) \div (z(0,x,u(0,x)), ..., z(K-1,x,u(K-1,x))).$$

Множество таких стратегических неопределенностей в задаче (4) обозначено через $\mathbf{Z}_U$. Далее гарантии исхода и риска строим для Γ_U из (4), а также для

$$\Gamma_U = \langle \mathbf{\Sigma}, \mathbf{U}, \mathbf{Z}_U, -R(U, Z_U, x_0) \rangle, \tag{5}$$

где $\mathbf{\Sigma}$, $\mathbf{U}$, $\mathbf{Z}_U$ те же, что в (4), а "минус" критерий взят из (3).

Перейдем к динамике задач (4) и (5). В момент $k = 0$ ЛПР передает выбранную им чистую стратегию $U(0) \div u(0,x)$ исследователю, формирующему стратегическую неопределенность при $k = 0$, именно $Z_U(0) \div z(0,x,u)$; тот определяет $z(0,x,u(0,x))$ и подставляет в систему (1), заменяя $z(0,x)$ на $z(0,x,u(0,x))$, а $u(0,x,z(0,x))$ на $u(0,x)$. В результате получаем из (1) при $k = 0$

$$x(1) = f(0, x_0, u(0,x_0), z(0, x_0, u(0,x_0))),$$

аналогично на втором шаге (при $k = 2$)

$$x(2) = f(1, x(1), u(1,x(1)), z(1, x(1), u(1,x(1)))).$$

Эта процедура продолжается до $k = K-1$ и на последнем шаге получаем

$$x(K) = f(K-1, x(K-1), u(K-1, x(K-1)), z(K-1, x(K-1), u(K-1, x(K-1)))).$$

Получаем три последовательности: векторов состояния $\{x(k)\}_{k=0}^{K}$, последовательность реализаций выбранных чистых стратегий $\{u[k] = u(k, x(k))\}_{k=0}^{K-1}$ и последовательность стратегических неопределенностей $\{z[k] = z(k, x(k), u(k, x(k)))\}_{k=0}^{K-1}$.

На них по формуле (2) определяется критерий $J(U, Z, x_0)$ (значение которого называется *исходом*) и по формуле (3) функция риска по Севиджу–Нихансу (значение которой называют *риском*).

Тогда гарантия по исходу, с учетом (4), будет

$$J[U, x_0] = \min_{Z_U \in \mathbf{Z}_U} J(U, Z_U, x_0), \tag{6}$$

а по риску, с учетом (3), будет

$$-R[U, x_0] = \min_{Z_U \in \mathbf{Z}_U} [-R(U, Z_U, x_0)]. \tag{7}$$

2. Формализация гарантированного решения

Наличие неопределенности в Г доставляет “большой простор” для определения гарантированных решений задачи Г.

Во-первых, применение общеиспользуемых принципов: максимина, минимаксного сожаления, Гурвица, Лапласа и т. д. (см. [1]).

Во-вторых, одновременно с указанными принципами учитывать и собственное значение критерия (2) и тем самым переходить к двух-, трех- и т. д. критериальным задачам.

В настоящей работе будем следовать второму подходу, заключающемуся в добавлении к (2) еще одного критерия — “минус” функции риска Севиджа–Ниханса [2, 3], и поиске максимальной по Парето стратегии в полученной в результате двухкритериальной задаче.

При этом основываемся на “трех китах”:

1. способе доказательства метода динамического программирования для дискретных систем, предложенном Болтянским В. Г. в книге [4];

2. двух способах построения гарантий из недавно опубликованной серии работ [5, 6];

3. математической теории оптимумов по Парето из книги [7].

Итак, прежде всего перейдем к формализации *сильно гарантированного по исходу и риску максимального по Парето решения* (СГИРП) задачи Г.

Следуя подходу из [2, 3], построим для критерия (2) функцию риска по Севиджу–Нихансу в два этапа.

Во-первых, найдем в задаче Γ_Z стратегию $U_Z^* \in \mathbf{U}_Z$, $U_Z^* = (U_Z^*(0), ..., U_Z^*(K-1))$, $U_Z^*(k) \in \mathbf{U}_Z(k)$, удовлетворяющую условию

$$\max_{U_Z \in \mathbf{U}_Z} J(U_Z, Z, x_0) = J[Z, x_0] \; \forall Z \in \mathbf{Z}, \; x_0 \in \mathbb{R}^n,$$

при этом используя класс чистых неопределенностей

$$Z = (Z(0), ..., Z(K-1)), \; Z(k) \div z(k, x), \; Z(k) \in \mathbf{Z}(k), \; \mathbf{Z} = \prod_{k=0}^{K-1} \mathbf{Z}(k).$$

Во-вторых, построим с учетом (3) саму функцию риска по Сэвиджу–Нихансу

$$R(U, Z, x_0) = J[Z, x_0] - J(U, Z, x_0),$$

значение которой для каждой пары $(U, Z) \in \mathbf{U} \times \mathbf{Z}$ как раз и определяет риск, сопровождающий использование чистой стратегии U. Естественно, что ЛПР желает за счет подходящего выбора своей стратегии $U \in \mathbf{U}$ уменьшить риск $R(U, Z, x_0)$ и одновременно *увеличить исход*, т. е. значение критерия $J(U, Z_U, x_0)$. В результате приходим к двухкритериальной многошаговой позиционной задаче при стратегической неопределенности

$$\langle \mathbf{\Sigma},\ \mathbf{U},\ \mathbf{Z}_U, \{J(U, Z_U, x_0), -R(U, Z_U, x_0)\}. \tag{8}$$

В (8), как и в Г, управляемая ситема $\mathbf{\Sigma}$ описывается разностным уравнением (1), $\mathbf{U}$ — множество чистых позиционных стратегий $U \div (u(0, x), ..., u(K-1, x))$, $\mathbf{Z}_U$ — множество стратегических позиционных неопределеннностей $Z_U \div (z(0, x, u), ..., z(K-1, x, u))$. Стремление увеличить оба критерия одновременно, благодаря "минусу" перед вторым критерием (функция риска по Сэвиджу), эквивалентно стремлению увеличить исход (первый критерий $J(U, Z_U, x_0)$) и одновременно уменьшить второй (риск) $R(U, Z_U, x_0)$. При этом ЛПР вынуждено ориентироваться на возможность реализации любой неопределенности $Z_U \in \mathbf{Z}_U$.

Подход, предложенный в [5, 6], позволяет учесть "действия" неопределенности в (8) за счет оценки действия гарантий обоих критериев. Для этого решим две вспомогательные задачи: найти $Z_U^{(l)} \in \mathbf{Z}_U$ $(l = 1, 2)$ такие, что

$$\begin{aligned} &\min_{Z_U \in \mathbf{Z}_U} J(U, Z_U, x_0) = J(U, Z_u^{(1)}, x_0) = J[U, x_0], \\ &\min_{Z_U \in \mathbf{Z}_U} [-R(U, Z_U, x_0)] = -R(U, Z_U^{(2)}, x_0) = -R[U, x_0] \end{aligned} \tag{9}$$

при $\forall\ U \in \mathbf{U}$, $x_0 \in \mathbb{R}^n$. Найденные согласно (9) критерии $J[U, x_0]$ и $-R[U, x_0]$ действительно будут гарантиями, ибо из (9) следует

$$\begin{aligned} J(U, Z_U, x_0) &\geq J[U, x_0], \\ -R(U, Z_U, x_0) &\geq -R[U, x_0] \end{aligned} \tag{10}$$

при $\forall\ U \in \mathbf{U}$, $x_0 \in \mathbb{R}^n$, и поэтому имеют место при $\forall\ Z_U \in \mathbf{Z}_U$.

Итак, от задачи (4) перейдем к двухкритериальной многошаговой позиционной "задаче гарантий"

$$\langle \mathbf{\Sigma},\ \mathbf{U},\ \{J[U, x_0], -R[U, x_0]\rangle, \tag{11}$$

но уже без неопределенности.

Тогда в качестве решения (СГИРП) задачи (1) предлагается тройка

$$(U^P, J^P[x_0], R^P[x_0]) \tag{12}$$

такая, что позиционная стратегия $U^P \in \mathbf{U}$ максимальна по Парето в (11), т. е. при $\forall\ x_0 \in \mathbb{R}^n$ и $\forall\ U \in \mathbf{U}$ несовместна система неравенств

$$J[U, x_0] \geq J[U^P, x_0] = J^P[x_0],$$
$$R[U, x_0] \leq R[U^P, x_0] = R^P[x_0],$$

из которых хотя бы одно строгое.

Замечание 1. Почему же тройка (12) выбрана в качестве гарантированного решения задачи Г?

Во-первых, (12) отвечает на вопрос: "Что делать?" Ответ — использовать указанную выше максимальную по Парето стратегию U^P.

Во-вторых, вследствие (10), стратегия U^P гарантирует исход $J(U^P, Z, x_0)$ не меньший гарантированного $J^P[x_0]$, а так же риск не больший $R^P[x_0]$ при $\forall\ x_0 \in \mathbb{R}^n$, любых неопределенностях Z_U и всяких стратегиях $U \in \mathbf{U}$, не совпадающих с U^P.

Более того, увеличение исхода (за счет использования $U \neq U^P$) автоматически приведет к увеличению риска и обратно, уменьшение риска вызовет уменьшение исхода (цель же ЛПР — увеличить исход при одновременном уменьшении риска).

Перейдем к строгому определению.

Определение 1. Тройку $(U^P, J^P[x_0], R^P[x_0])$ назовем сильно гарантированным по исходам и рискам Парето-максимальным решением (СГИРП) задачи Γ_U, если:
1^0. в задаче Γ_Z существует контрстратегия $U_Z^* \in \mathbf{U}_Z$, для которой

$$\max_{U_Z \in \mathbf{U}_Z} J(U_Z, Z, x_0) = J(U_Z^*, Z, x_0) = J[Z, x_0] \tag{13}$$

при $\forall Z \in \mathbf{Z}$, $x_0 \in \mathbb{R}^n$;
2^0. (далее используем функцию риска по Сэвиджу (3) и исход (2)) в задаче Γ_Z существуют две стратегические неопределенности $Z_U^{(l)}$ $(l = 1, 2)$ такие, что при $\forall\ U \in \mathbf{U}$ и $\forall\ x_0 \in \mathbb{R}^n$

$$\begin{aligned} &\min_{Z_U \in \mathbf{Z}_U} J(U, Z_U, x_0) = J(U, Z_U^{(1)}, x_0) = J[U, x_0], \\ &\min_{Z_U^{\in} \mathbf{Z}_U} [-R(U, Z_U, x_0)] = -R(U, Z_U^{(2)}, x_0) = -R[U, x_0]; \end{aligned} \tag{14}$$

3^0. чистая позиционная стратегия $U^P \in \mathbf{U}$ максимальна по Парето в двухкритериальной задаче

$$\langle \mathbf{\Sigma},\ \mathbf{U},\ \{J[U,x_0],\ -R[U,x_0]\}\rangle \tag{15}$$

при $\forall\ x_0 \in \mathbb{R}^n$. Тогда для $\forall x_0 \in \mathbb{R}^n$ имеем $J^P[x_0] = J[U^P, x_0]$, $R^P[x_0] = R(U^P, x_0)$.

3. Достаточные условия

В этом разделе статьи используем следующие обозначения:
$V_{\text{и}}^{(k)}(x,z)$ — функция Беллмана для исходов,
$V_p^{(k)}(x,z)$ — функция Беллмана для рисков,

$$W_{\text{и}}(k,x,z,u,V_{\text{и}}^{(k+1)}(f(k,x,z,u),z)) = F(k,x,z,u) + V_{\text{и}}^{(k+1)}(f(k,x,z,u),z) \tag{16}$$

для задачи Γ_Z построим

$$\begin{gathered} V_{\text{и}}^{(k)}(x,z) = W_{\text{и}}[k,x,z] = \max_u \{W_{\text{и}}(k,x,z,u,V_{\text{и}}^{(k+1)}(f(k,x,z,u),z))\} = \\ = Idem\{u \to u^*(k,x,z)\}\ \forall x \in \mathbb{R}^n,\ z \in \mathbb{R}^m\ (k = K-1,...,1,0); \end{gathered} \tag{17}$$

здесь и далее $Idem\{u \to u^*(k,x,z)\}$ означает, что в выражении в фигурных скобках заменяется u на $u^*(k,x,z)$, т.е. максимум достигается при $u = u^*(k,x,z)$ $\forall x \in \mathbb{R}^n$, $z \in \mathbb{R}^s$; составим

$$R(k,x,z,u) = W_{\text{и}}[k,x,z] - W_{\text{и}}(k,x,z,u,V_{\text{и}}^{(k+1)}(f(k,x,z,u),z)); \tag{18}$$

для задачи Γ_Z найдем

$$\begin{gathered} R[k,x,u] = \max_z \{R(k,x,z,u)\} = Idem\{z \to z^{(2)}(k,x,u)\}, \\ W_{\text{и}}[k,x,u] = \min_z \{W_{\text{и}}(k,x,z,u,V_{\text{и}}^{(k+1)}(f(k,x,z,u),z))\} = \\ = Idem\{z \to z^{(1)}(k,x,u)\}. \end{gathered} \tag{19}$$

Здесь опять-таки $Idem\{z \to z^{(l)}(k,x,u)\}$ означает замену z на $z_u^{(l)}$ соответственно $(l = 1,2)$.

Справедливость следующего положения теоретически обоснована утверждением и способом доказательства метода динамического программирования, примененными В. Г. Болтянским в монографии [4].

Утверждение 1. *Пусть для задачи* Γ_U *существует последовательность функций Беллмана* $\{V_u^{(k)}(x,z)\}_0^K$, *определенных на* $\mathbb{R}^{n+m}$ *и таких, что*
1^0. $V_u^{(K)}(x,z) = \Phi(x)\ \ \forall x \in \mathbb{R}^n$*;*
2^0. *при каждом* $k = K-1,...,0$ *последовательно существует контрстратегия в момент* k, *т.е.* $U_Z^*(k) \div u^*(k,x,z)$, $U_Z^*(k) \in \mathbf{U}_Z(k)$, *удовлетворяющая (17) при*

$\forall\ x \in \mathbb{R}^n,\ z \in \mathbb{R}^m$ в задаче Γ_U;
3^0. с учетом (16) и (18) для задачи Γ_Z при каждом $k = K-1, ..., 0$ существуют по две стратегические неопределенности $Z_U^{(l)}(k),\ Z_U^{(l)}(k) \div z^{(l)}(k, x, u),\ \ Z_U^{(l)}(k) \in \mathbf{Z}_U(k)$ $(l = 1, 2)$, для которых имеют место оба тождества (19) при $\forall x \in \mathbb{R}^n,\ u \in \mathbb{R}^m$;
4^0. в результате для $k = 0$ получаем две стратегические неопределенности $Z^{(l)} \in \mathbf{Z}_U$, определяющие две гарантии $J(U, Z^{(1)}, x_0) = J[U, x_0]$, $-R(U, Z^{(2)}, x_0) = -R[U, x_0]$, затем в задаче (1) строим максимальную по Парето чистую стратегию, например, найдя чистую стратегию $U^P \in \mathbf{U}$, реализующую максимум

$$\max_{U \in \mathbf{U}} \{J[U, x_0] - R[U, x_0]\} = Idem\{U \to U^P\}.$$

Находим гарантии $J^P[x_0] = J[U^P, x_0],\ R^P[x_0] = R[U^P, x_0]$.

Тогда при любом выборе начального состояния $x_0 \in \mathbb{R}^n$ СГИРП образует тройка $(U^P, J^P[x_0], R^P[x_0])$.

4. Линейно-квадратичная одношаговая задача при неопределенности

Предполагаем, что в Γ_Z, управляемая система (1) одношаговая, линейна и имеет вид

$$x(1) = Ax(0) + u + z,\ x(0) = x_0, \tag{20}$$

n-вектора $x,\ u,\ z \in \mathbb{R}^n$ и момент окончания $K = 1$, $n \times n$-матрица A постоянна (далее множество таких матриц обозначаем $\mathbb{R}^{n\times n}$). Чистую стратегию $U \div Px$, их множество $\mathbf{U}^n = \{U \div Px \mid \forall P \in \mathbb{R}^{n\times n}\}$, используем также множество контрстратегий $\mathbf{U}_Z^n = \{U_Z \div P_1x + P_2z \mid \forall P_i \in \mathbb{R}^{n\times n}\ (i = 1, 2)\}$, множество чистых неопределенностей $\mathbf{Z}^n = \{Z \div Qx \mid \forall Q \in \mathbb{R}^{n\times n}\}$, множество стратегических неопределенностей $Z_U^n = \{Z_U \div Q_1x + Q_2u \mid \forall Q_i \in \mathbb{R}^{n\times n}\ (i = 1, 2)\}$; штрих сверху означает операцию транспонирования.

Тогда Γ_Z превращается в

$$\Gamma_Z^n = \langle \mathbf{\Sigma} \div (20), \mathbf{U}_Z^n, \mathbf{Z}^n, J(U_Z, Z, x_0)\rangle,$$

а Γ_U в

$$\Gamma_U^n = \langle \mathbf{\Sigma} \div (20), \mathbf{U}^n, \mathbf{Z}_U^n, J(U, Z_U, x_0)\rangle,$$

где

$$J(U, Z_U, x_0) = -x^{'}(1)Cx(1) - u^{'}[1]Du[1] + z^{'}[1]Mz[1].$$

Причем предполагаем выполненным

Условие 1. *Матрицы* $C, D, M \in \mathbb{R}^{n\times n}$ *симметричны и квадратичные формы* $y^{'}Cy$, $y^{'}Dy$, $y^{'}My$ *определенно положительны (обозначаем этот факт,* $C > 0$, $D > 0$, $M > 0$*).*

Задача: построить СГИРП для Γ_U^n.

Утверждение 2. *Пусть* $n \times n$*-матрицы* $C > 0$, $M > 0$, $D > 0$*. Тогда при любом выборе начального ненулевого фазового вектора* x_0 *СГИРП будет* $(U^P, J^P[x_0], R^P) = (0_n, 0, 0)$.

Доказательство. В соответствии с утверждением 1 разделим построение СГИРП на четыре этапа. Ищем функции Беллмана в виде квадратичных форм $x^{'}\Theta_i(k)x$ $(i = ,)$.

1 этап: Используем тождества

$$V_i(1, x, z) = x^{'}\Theta_i(1)x = -x^{'}C^{'}x \quad \forall x \in \mathbb{R}^n,$$

откуда

$$\Theta_i(1) = -C \ (i = ,).$$

2 этап: для задачи Γ_Z^n, согласно (16), (17), построим $U_Z^* \div P_1^*x + P_2^*z$ исходя из условия

$$\max_u W_{\text{и}}(0, x, z, u, V_{\text{и}}^{(1)}(Ax + u + z, z)) = \max_u\{-u^{'}Du + z^{'}Mz-$$
$$-(Ax + u + z)^{'}C(Ax + u + z)\} = Idem\{u \to u^*(0, x, z) = P_1^*(0)x + P_2(0)z\}.$$

Достаточными условиями здесь будут

$$\begin{cases} -2Du^*(0, x, z) - 2C(Ax + u(0, x, z) + z) = 0_n, \\ -2(D + C) < 0. \end{cases}$$

Второе требование выполнено (ибо $D > 0$ и $C > 0$), а из первого тождества получаем

$$u^*(0, x, z) = -D^{-1}C(Ax + z).$$

Тогда, подставляя $u = u^*(0, x, z)$ в $W_{\text{и}}(...)$, приходим к

$$\max_u W_{\text{и}}(...) = -(Ax + z)^{'}\mathfrak{M}(Ax + z) + z^{'}Mz,$$

где симметричная постоянная $n \times n$ матрица

$$\mathfrak{M} = C(D + C)^{-1}D(D^{-1} + C^{-1})D(D + C)^{-1}C > 0. \tag{21}$$

Используя (18), имеем

$$\begin{aligned} R(0, x, u, z) = -(Ax + z)^{'}\mathfrak{M}(Ax + z) + z^{'}Mz + u^{'}Du - z^{'}Mz+ \\ +(Ax + z + u)C(Ax + z + u) = u^{'}(D + C)u+ \\ +(Ax + z)^{'}(C - \mathfrak{M})(Ax + z) + 2u^{'}C(Ax + z). \end{aligned} \tag{22}$$

3 этап: для задачи Γ_U^n построим гарантии по рискам и исходам. Используя при этом (19), для гарантии по риску имеем

$$\max_z \{R(0,x,u,z)\} = \{z \to z^*(0,x,u)\}.$$

Но это соотношение имеет место, если при $\forall x, u \in \mathbb{R}^n$ будет

$$\begin{array}{l} \frac{\partial R(0,x,u,z)}{\partial z}\Big|_{z^*(0,x,u)} = -2\mathfrak{M}[Ax + z(0,x,u)] + 2C(Ax + u + z(0,x,u)) = 0, \\ \frac{\partial^2 R(0,x,u,z)}{\partial z^2} = -2(C - \mathfrak{M}) < 0. \end{array} \tag{23}$$

Из (23) также

$$Ax + u + z^*(0,x,u) = C^{-1}\mathfrak{M}(Ax + z^*(0,x,u)),$$

и тогда для (22) получаем

$$R(0,x,u,z^*(0,x,u)) = u^{'}Du + (Ax + z^*(0,x,u))^{'}\mathfrak{M}(\mathfrak{M}^{-1} - C^{-1})\mathfrak{M}(Ax + z^*(0,x,u)).$$

Поэтому с учетом эквиваленции

$$[0 < C < \mathfrak{M}] \Leftrightarrow [0 < \mathfrak{M}^{-1} < C^{-1}]$$

приходим к

$$z^*(0,x,u) = -Ax$$

и тогда

$$\max_z R(0,x,u,z) = R[0,x,u] = u^{'}Du. \tag{24}$$

Перейдем к гарантии по исходам

$$\min_z \{-u^{'}Du + z^{'}Mz - (Ax + u + z)^{'}C(Ax + u + z)\} = Idem\{z \to z^0(0,x,u)\}.$$

Тождество имеет место, если

$$\begin{cases} 2Mz^0(0,x,u) - 2C(Ax + u + z^0(0,x_0,u)) = 0 \\ [2(M - C) > 0] \Leftrightarrow [M > C], \end{cases}$$

из первого тождества также имеем

$$Ax + u + z(0,x,u) = C^{-1}Mz(0,x,u).$$

Следовательно,

$$\begin{array}{l} W_{\text{и}}(0,x,z^0(0,x,n),u,V_{\text{и}}^{(1)}(Ax + z^0(0,x,u) + u)) = W_{\text{и}}[0,x,u] = \\ = -u^{'}Du + [z^0(0,x,u)]^{'}Mz^0(0,x,u) - z^0(0,x,u)MC^{-1}Mz^0(0,x,u) = \\ = -u^{'}Du + z^0(0,x,u)M(M^{-1} - C^{-1})Mz^0(0,x,u). \end{array} \tag{25}$$

4 этап: из (24) и (25) сразу следует, что в задаче (11) максимальной по Парето является чистая стратегия $U^P \div 0_n$.

□

Исследования выполнены в рамках проекта кафедры оптимального управления МГУ им. М. В. Ломоносова: «Методы решения динамических задач управления, оптимизации и идентификации».

Список литературы

1. Мушек, Э. Методы принятия технических решений / Э. Мушек, П. Миллер. — М.: Мир, 1990. — 208 с.

   MUSCHICK, E., MULLER, P. (1990) *Methods of technical decisions.* Moscow: Mir.

2. SAVAGE, L. Y. (1954) *The Foundation of Statistics.* New York: Wiley.

3. NICHANS, J. (1951) Zur Preisbuldung bei undewissen Erwartungen. *Schweizerische Zeitschrift Assotiation.* Vol.46 (No 253). p. 55–67.

4. Болтянский, В. Г. Оптимальное управление дискретными системами / В. Г. Болтянский. — М.: Наука, 1973. — 488 с.

   BOLTYANSKII, V. G. (1973) *Optimal control of discrete systems.* Moscow: Nauka.

5. Жуковский, В. И. Уравновешивание конфликтов при неопределенности. I. Аналог седловой точки // Математические основы теории игр и приложения / В. И. Жуковский, К. Н. Кудрявцев. — 2013. — Т. 5. — № 1. — С. 27–44.

   ZHYKOVSKY, V. I., KUDRYAVTSEV, K. N. (2013) Balancing conflicts uncertainty. I. An analogue of the saddle point. *Mathematical foundations of the theory of games and applications.* 5 (1). p. 27–44.

6. Жуковский, В. И. Уравновешивание конфликтов при неопределенности. II. Аналог максимина / В. И. Жуковский, К. Н. Кудрявцев // Математические основы теории игр и приложения. — 2013. — Т. 5. — № 2. — С. 3–45.

   ZHYKOVSKY, V. I., KUDRYAVTSEV, K. N. (2013) Balancing conflicts uncertainty. II. The analogue maximin. *Mathematical foundations of the theory of games and applications.* 5 (2). p. 3–45.

7. Подиновский, В. В. Парето-оптимальные решения многокритериальных задач / В. В. Подиновский, В. Д. Ногин. — М.: Наука, 1982. — 256 с.

   PODINOVSKII, V. V., NOGIN, V. D. (1982) *Pareto-optimal solutions for multiobjective problems.* Moscow: Nauka.