

УДК: 518.9

ДИНАМИЧЕСКИЕ МАТРИЧНЫЕ ИГРЫ. ОБОСНОВАНИЕ ПРИМЕНЕНИЯ ПРИНЦИПА МИНИМАКСА В КЛАССЕ ЧИСТЫХ КОМБИНИРОВАННЫХ СТРАТЕГИЙ

Н. Н. Данилов

DYNAMIC MATRIX GAMES. SUBSTANTIATION OF APPLICATION OF MINIMAX PRINCIPLE IN THE CLASS OF PURE COMBINED STRATEGIES

N. N. Danilov

Матричные игры имеют достаточно обширное практическое применение. Поскольку процессы принятия решений в социально-экономических и других системах имеют динамический характер, возникает необходимость обобщения класса матричных игр так, чтобы учитывались временные факторы. Динамические матричные игры построены в форме многошаговой задачи конфликтного управления. Введен новый класс стратегий (k-стратегий). Определены понятия допустимых, оптимальных и чистых k-стратегий, понятие оптимальной траектории и цены игры. Обосновано применение принципа минимакса. Найдено достаточное условие существования единственной оптимальной стратегии. Установлена динамическая устойчивость оптимальной траектории.

Matrix games have rather extensive practical application. As decision-making processes in social and economic and other systems have dynamic nature, there is a necessity of generalization of matrix games class so that time factors were considered. Dynamic matrix games are built in the form of a multistage problem of conflict control. New class of strategies (k-strategies) are introduced. Concepts of optimal pure k-strategies, optimal trajectory and value of game are defined. The use of minimax principle in dynamical matrix games are substantiated. Sufficient condition of existence of unique optimal trajectory is found. Dynamical stability of optimal trajectory is ascertained.

Ключевые слова: матричные игры, принцип минимакса, седловая точка, цена игры, оптимальные стратегии и траектория, динамическая устойчивость.

Keywords: matrix games, minimax principle, saddle point, value of game, optimal strategy and trajectory, dynamical stability.

Пусть переменная t , которую будем называть «временем», принимает лишь дискретные значения: $t = 0, 1, \dots, T$. Рассмотрим некоторую управляемую систему, изменение состояния которой происходит в дискретные моменты времени t и описывается уравнением

$$x(t) = f^t(x(t-1), u_1(t), u_2(t)), \quad t = 1, \dots, T, \quad (1)$$

из заданного начального состояния

$$x^0 = x(0). \quad (2)$$

В (1) – (2) $x(t) \in R^k$ – вектор состояния; $u_1(t) \in R^{l_1}, u_2(t) \in R^{l_2}$ – векторы управления в момент t ; $f^1: R^k \times R^{l_1} \times R^{l_2} \rightarrow R^k$ – вектор-функция, характеризующая динамические возможности системы.

Предполагается, что управляющие параметры удовлетворяют условиям

$$u_1(t) \in U_1^t, u_2(t) \in U_2^t, \quad t = 1, \dots, T, \quad (3)$$

где $U_1^t \subset R^{l_1}, U_2^t \subset R^{l_2}$ – заданные непустые конечные множества.

Соотношения (1) – (3) определяют дискретную систему с двумя управлениями. Будем считать, что выбором управления $u_1(\cdot) = \{u_1(1), \dots, u_1(T)\}$ распоряжается игрок I, а выбором управления

$$u_2(\cdot) = \{u_2(1), \dots, u_2(T)\} \text{ – игрок II.}$$

Множества допустимых управлений игроков есть соответственно

$$U_1 = U_1^1 \times \dots \times U_1^T, \quad U_2 = U_2^1 \times \dots \times U_2^T.$$

Предполагается, что каждой паре управлений $u_1(\cdot), u_2(\cdot) \in U_1 \times U_2$ соответствует единственная последовательность

$$x(\cdot) = x(\cdot, x^0, u_1(\cdot), u_2(\cdot)) = \{x^0, x(1), \dots, x(T)\}$$

решений системы (1) – (2), которую будем называть траекторией. Множество всех траекторий (1) – (2) обозначим символом $X(x^0, T)$.

Введем в рассмотрение множество:

$$G(x(t-1), t) = \{x(t) = x(t, x(t-1), u_1(t), u_2(t)) \mid$$

$$x(t-1) \in G(x(t-2), t-1), u_1(t) \in U_1^t, u_2(t) \in U_2^t,$$

которое будем называть множеством достижимости t -го уровня, $t = 1, \dots, T$ (будем считать, что $G(x(-1), 0) = x^0$). Множество

$$G(x^0, T) = \bigcup_{t=1}^T G(x(t-1), t)$$

будем называть множеством достижимости системы (1) – (3). Считается, что $x^0 \in G(x^0, T)$ по определению.

Пусть в каждой точке $x = x(t)$ ($t = 1, \dots, T$) множества достижимости $G(x^0, T)$ системы (1) – (3) определена матрица выигрышей игрока I:

$$h(x(t)) = \begin{pmatrix} a_{11}(x(t)) & a_{12}(x(t)) & \dots & a_{1n}(x(t)) \\ a_{21}(x(t)) & a_{22}(x(t)) & \dots & a_{2n}(x(t)) \\ \dots & \dots & \dots & \dots \\ a_{m1}(x(t)) & a_{m2}(x(t)) & \dots & a_{mn}(x(t)) \end{pmatrix}. \quad (4)$$

$(h(x^0) = \|0\|_{m \times n}$ – $m \times n$ -матрица с нулевыми элементами, где m и n конечные числа). Целью игрока I является максимизация значения функции

$$H(x(\cdot), I, J) = \sum_{t=1}^T a_{i_t j_t}(x(t)), \quad (5)$$

где $I = \{i_t, t = 1, \dots, T\}$, $J = \{j_t, t = 1, \dots, T\}$ – последовательность чистых стратегий игроков, выбираемых ими в матричных играх

$h(x(t))$, $t = 1, \dots, T$. Целью игрока II является минимизация значения функции (5).

Определение 1. Чистой комбинированной стратегией (чистой k -стратегией) игрока I (II) будем называть отображение $\varphi_1(\varphi_2)$, которое каждому шагу t и реализуемому на этом шаге состоянию $x(t)$ ставит в соответствие номер $i_t(j_t)$ выбранной строки (столбца) матрицы $h(x(t))$ некоторое допустимое на этом шаге управление $u_1(t+1) \in U_1^{t+1}(u_2(t+1) \in U_2^{t+1})$.

k -Стратегии игроков можно представить так:

$$\begin{aligned} \varphi_1(\cdot) &= \{(u_1(1), i_1), \dots, (u_1(T), i_T)\}, \\ \varphi_2(\cdot) &= \{(u_2(1), j_1), \dots, (u_2(T), j_T)\}, \end{aligned} \quad (6)$$

где i_t и j_t – номера выбранных в момент t (в состоянии $x(t)$) строки и столбца матрицы $h(x(t))$ ($t = 1, \dots, T$). Множества стратегий игроков обозначим соответственно Φ_1 и Φ_2 .

Совокупность

$$\Gamma(x^0, T) = \left\langle \sum(x^0, T); \Phi_1, \Phi_2; H \right\rangle, \quad (7)$$

где $\sum(x^0, T)$ – символическое обозначение системы (1) – (2), назовем многошаговой конечной антагонистической игрой (или коротко – динамической матричной игрой).

Протекание игры $\Gamma(x^0, T)$ можно представить следующим образом. Игроки выбирают некоторые стратегии $\varphi_1(\cdot) \in \Phi_1$, $\varphi_2(\cdot) \in \Phi_2$. Далее вычисляется траектория

$$x(\cdot) = x(\cdot, x^0, u_1(\cdot), u_2(\cdot)) = \{x^0, x(1), \dots, x(T)\}$$

как решение системы (1) – (2), соответствующее ситуации $(u_1(\cdot), u_2(\cdot))$. Затем по формуле (5) вычисляется выигрыш игрока I вдоль этой траектории $x(\cdot)$:

$$K(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) = H(x(\cdot), I, J) = \sum_{t=1}^T a_{i_t j_t}(x(t)),$$

где i_t, j_t – номера строк и столбцов матрицы $h(x(t))$, предписанные k -стратегиями $\varphi_1(\cdot), \varphi_2(\cdot)$. Выигрыш игрока II вдоль той же траектории равен числу $-\sum_{t=1}^T a_{i_t j_t}(x(t))$.

Классическая матричная игра (см. [1] – [3]) – это статическая математическая модель, с помощью которой можно определить оптимальное поведение двух конфликтующих сторон, имеющих противоположные интересы и конечное число стратегий (решений, альтернатив). Матричные игры имеют достаточно широкое применение (см., напр., [2], [3]). Поскольку процессы принятия решения в социально-экономических и других системах имеют динамический характер, то возникает необходимость обобщения класса матричных игр так, чтобы учитывались временные факторы. Игра вида (7) и представляет собой такое обобщение.

Пусть игроками выбраны управления $\bar{u}_1(\cdot)$, $\bar{u}_2(\cdot)$ и вычислена соответствующая им траектория $\bar{x}(\cdot)$ системы (1) – (2). k -Стратегии соответствующие этим управлениям имеют вид:

$$\varphi_1(\cdot) = \{(\bar{u}_1(t), i_t), t = 1, \dots, T\},$$

$$\varphi_2(\cdot) = \{(\bar{u}_2(t), j_t), t = 1, \dots, T\}.$$

Введем следующие обозначения:

$$M = \{1, \dots, m\}, N = \{1, \dots, n\};$$

$$\hat{M} = \underbrace{M \times \dots \times M}_T, \quad \hat{N} = \underbrace{N \times \dots \times N}_T.$$

Определение 2. Последовательность $\{(\bar{i}_t, \bar{j}_t), t = 1, \dots, T\}$ назовем седловой точкой вдоль траектории $\bar{x}(\cdot)$, если для любых $\{i_1, \dots, i_T\} \in \hat{M}$, $\{j_1, \dots, j_T\} \in \hat{N}$ выполняются условия:

$$a_{i_t \bar{j}_t}^t(\bar{x}(t)) \leq a_{i_t \bar{j}_t}^t(\bar{x}(t)) \leq a_{i_t j_t}^t(\bar{x}(t)), \quad t = 1, \dots, T. \quad (8)$$

Из этого определения следует, что если вдоль траектории $\bar{x}(\cdot) = \{x^0, \bar{x}(1), \dots, \bar{x}(T)\}$ в какой-то из моментов $t = 1, \dots, T$ неравенство (8) не выполняется, т. е. в статической матричной игре $h(x(t))$ нет седловой точки, то седловая точка вдоль траектории $\bar{x}(\cdot)$ не существует.

Множество всех траекторий, вдоль которых существует седловая точка, обозначим символом $\bar{X}(x^0, T)$ ($\bar{X}(x^0, T) \subset X(x_0, T)$).

Для фиксированного t неравенство (8) определяет седловую точку (\bar{i}_t, \bar{j}_t) в статической матричной игре $h(\bar{x}(t))$. Из неравенства (8) получаем: для всех $\{i_1, \dots, i_T\} \in \hat{M}$, $\{j_1, \dots, j_T\} \in \hat{N}$

$$\sum_{t=1}^T a_{i_t \bar{j}_t}(\bar{x}(t)) \leq \sum_{t=1}^T a_{\bar{i}_t \bar{j}_t}(\bar{x}(t)) \leq \sum_{t=1}^T a_{\bar{i}_t \bar{j}_t}(\bar{x}(t)).$$

Для $\bar{x}(\cdot) \in \bar{X}(x_0, T)$, число

$$v(x^0, \bar{\varphi}_1(\cdot), \bar{\varphi}_2(\cdot)) = \sum_{t=1}^T a_{\bar{i}_t \bar{j}_t}(\bar{x}(t))$$

назовем ценой вдоль траектории $\bar{x}(\cdot)$.

Предположим, что $\bar{X}(x^0, T) \neq \emptyset$. Каждой траектории $\bar{x}(\cdot) \in \bar{X}(x_0, T)$ соответствует своя седловая точка $\{(\bar{i}_t, \bar{j}_t), t = 1, \dots, T\}$. В отличие от статической игры вида $h(x(t))$, седловые точки, соответствующие разным траекториям из $\bar{X}(x_0, T)$, вообще говоря, не обладают свойствами взаимозаменяемости и эквивалентности. Действительно, пусть $\tilde{x}(\cdot), \hat{x}(\cdot) \in \bar{X}(x_0, T)$, а $(\tilde{\varphi}_1(\cdot), \tilde{\varphi}_2(\cdot)), (\hat{\varphi}_1(\cdot), \hat{\varphi}_2(\cdot))$ – соответствующие этим траекториям стратегии. Нельзя утверждать, что пары $\{(\tilde{i}_t, \tilde{j}_t)\}$ и $\{(\hat{i}_t, \hat{j}_t)\}$ образуют седловую точку (невозможна взаимозаменяемость), так как необязательно траектории $x'(\cdot) = x(\cdot, x^0, \tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ и $x''(\cdot) = x(\cdot, x^0, \hat{u}_1(\cdot), \hat{u}_2(\cdot))$ принадлежат множеству $\bar{X}(x_0, T)$ и нельзя утверждать, что

$$\sum_{t=1}^T a_{\tilde{i}_t \tilde{j}_t}(x'(t)) = \sum_{t=1}^T a_{\hat{i}_t \hat{j}_t}(x''(t))$$

(неэквивалентность).

В общем случае факт «непрямоугольности» множества

$$\bar{\Phi}_1 \times \bar{\Phi}_2 = \{(\bar{\varphi}_1(\cdot), \bar{\varphi}_2(\cdot)) \mid x(\cdot, x^0, \bar{u}_1(\cdot), \bar{u}_2(\cdot)) \in \bar{X}(x^0, T)\}$$

не исключает выгоды одностороннего отклонения (любого из двух игроков) от ситуации $(\bar{\varphi}_1(\cdot), \bar{\varphi}_2(\cdot)) \in \bar{\Phi}_1 \times \bar{\Phi}_2$ (причем отклонение понимается как от $(\bar{u}_1(\cdot), \bar{u}_2(\cdot))$, так и от (\bar{I}, \bar{J})) (см. (6)). Это делает бессмысленным применение правила максимина (минимакса) в игре (7) как принципа оптимальности.

Множество $\bar{\Phi}_1 \times \bar{\Phi}_2$ будет «прямоугольным», а применение принципа максимина содержательным в том случае, если множество $\bar{X}(x^0, T)$ будет состоять только из одного элемента, т. е. в игре $\Gamma(x^0, T)$ будет реализована лишь одна траектория, вдоль которой существует седловая точка (см. определение 2).

В данной работе построена специфическая процедура применения правила минимакса в игре $\Gamma(x^0, T)$, вытекающая из ее особенностей как

многошаговой матричной игры в классе к-стратегий.

Установлено, что предложенная процедура порождает единственную реализуемую траекторию из множества $\bar{X}(x^0, T)$, которая соответствует единственной седловой точке (в классе к-стратегий) в игре $\Gamma(x^0, T)$. Таким образом, в данной работе строго обоснована состоятельность (правомочность) применения правила минимакса (максимина) как принципа оптимального поведения в динамической матричной игре.

Цену игры вдоль траектории можно представить как функцию $v(x^0, \cdot) : \bar{\Phi}_1 \times \bar{\Phi}_2 \rightarrow R^1$ (или $v(x^0, \cdot) : \bar{X}(x^0, T) \rightarrow R^1$).

Введем в рассмотрение два числа:

$$\underline{v}(x^0) = \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) \quad (9)$$

– нижняя цена игры $\Gamma(x^0, T)$,

$$\bar{v}(x^0) = \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) \quad (10)$$

– верхняя цена игры $\Gamma(x^0, T)$.

Так как множества $\bar{\Phi}_1$ и $\bar{\Phi}_2$ конечные, то экстремумы в (9) и (10) существуют.

к-Стратегии, на которых достигаются внешние экстремумы в (9) и (10), назовем максиминной и минимаксной к-стратегиями первого и второго игроков соответственно.

Можно показать, что в игре $\Gamma(x^0, T)$

$$\begin{aligned} \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) &\leq \\ &\leq \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) \end{aligned} \quad (11)$$

и, в случае строгого неравенства в (11), максиминная и минимаксная пара к-стратегий не являются устойчивыми против индивидуальных отклонений игроков от нее. Напротив, равенство

$$\begin{aligned} \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) &= \\ = \min_{\varphi_2(\cdot) \in \bar{\Phi}_2} \max_{\varphi_1(\cdot) \in \bar{\Phi}_1} v(x^0, \varphi_1(\cdot), \varphi_2(\cdot)) \end{aligned} \quad (12)$$

порождает устойчивую пару к-стратегий.

Определение 3. Если выполнено равенство (12), то соответствующие ему максиминную и минимаксную к-стратегии назовем оптимальными к-стратегиями (будем обозначать их $\varphi_1^*(\cdot)$ и $\varphi_2^*(\cdot)$) первого и второго игроков, число

$$\begin{aligned} v(x^0) &= v(x^0, \varphi_1^*(\cdot), \varphi_2^*(\cdot)), \\ \varphi_2^*(\cdot) &= \underline{v}(x^0) = \bar{v}(x^0) \end{aligned} \quad (13)$$

– ценой игры $\Gamma(x^0, T)$, а пару $(\varphi_1^*(\cdot), \varphi_2^*(\cdot))$ будем называть седловой точкой игры $\Gamma(x^0, T)$.

Пусть $x^*(\cdot) = x(\cdot, x^0, u_1^*(\cdot), u_2^*(\cdot))$ – траектория системы (1) – (2), соответствующая седловой

точке $(\varphi_1^*(\cdot), \varphi_2^*(\cdot))$. Тогда цена (13) игры $\Gamma(x^0, T)$ вычисляется так:

$$v(x^0) = \sum_{t=1}^T a_{i_t^* j_t^*}^t(x^*(t)),$$

где (i_t^*, j_t^*) – седловая точка статической матричной игры $h(x^*(t))$, $t = 1, \dots, T$.

Для упрощения дальнейших выкладок примем следующие условия:

все матрицы $h(x)$, $x \in G(x^0, T)$, имеют одну и ту же размерность $m \times n$;

для любых $t = 1, \dots, T$ мощности множеств $U_1^t(U_2^t)$ одни и те же и равны соответственно $q(r)$ (т. е. эти множества содержат q и r допустимых значений управляющих параметров).

Любой точке $x(t) \in G(x(t-1), t)$ и паре множеств U_1^{t+1}, U_2^{t+1} соответствуют (в силу системы (1)) $q \cdot r$ точек $x(t+1) \in G(x(t), t+1)$. Поэтому мощность множества $G(x(t), t+1)$ равна $(q \cdot r)^{t+1}$.

Пусть игроками выбраны управления:

$$u_1(\cdot) = \{u_1^{k_1}(1), u_1^{k_2}(2), \dots, u_1^{k_T}(T)\},$$

$$u_2(\cdot) = \{u_2^{l_1}(1), u_2^{l_2}(2), \dots, u_2^{l_T}(T)\},$$

где $u_1^{k_i}(t) \in U_1^t$, $u_2^{l_i}(t) \in U_2^t$.

Паре управлений

$$(u_1(\cdot), u_2(\cdot)) = \{(u_1^{k_1}(1), u_2^{l_1}(1)), \dots, (u_1^{k_T}(T), u_2^{l_T}(T))\}$$
 соответствует траектория

$$x(\cdot) = x(\cdot, x^0, u_1(\cdot), u_2(\cdot)) = \{x^0, x^{k_1 l_1}(1), \dots, x^{k_T l_T}(T)\}, \quad (14)$$

где

$$x^{k_t l_t}(t) = x(t, x^{k_{t-1} l_{t-1}}(t-1), u_1^{k_t}(t), u_2^{l_t}(t)) \in G(x(t-1), t).$$

Для траектории (14) последовательность $\{x^{k_{t-1} l_{t-1}}(t-1), \dots, x^{k_1 l_1}(1)\}$ будем называть предысторией точки $x^{k_t l_t}(t)$ и будем писать:

$$x^{k_t l_t}(t) = x^{k_t l_t}(t) \mid x^{k_{t-1} l_{t-1}}(t-1), \dots, x^{k_1 l_1}(1).$$

В дальнейшем, для удобства, цену игры A будем обозначать символом $val A$.

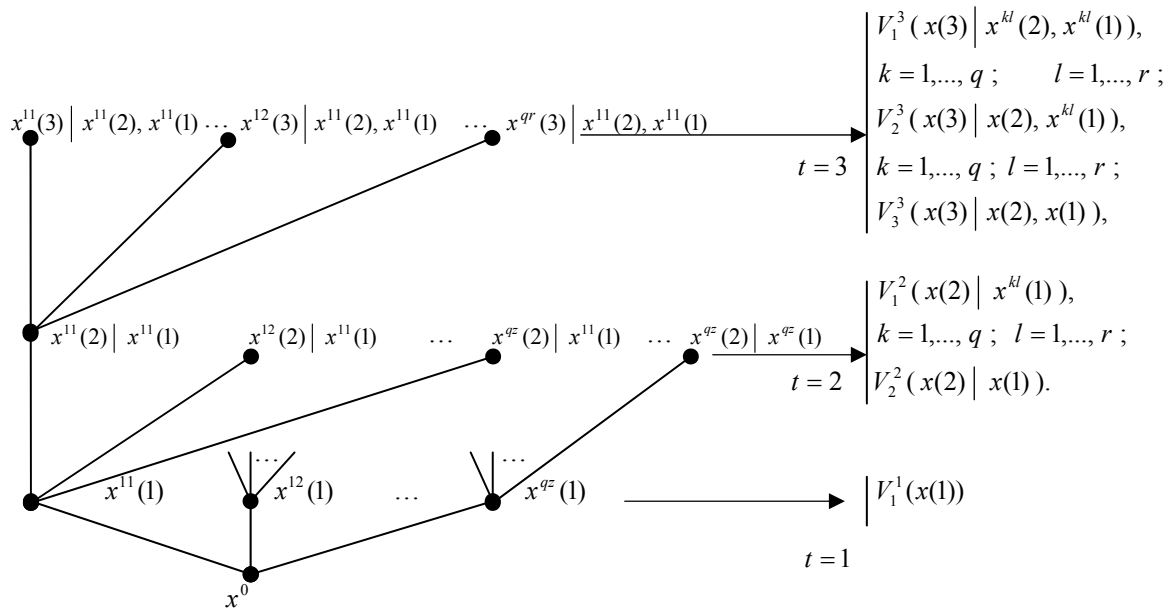


Рис. 1. Соответствие уровневых матриц фазовому дереву игры (фрагмент)

Для доказательства существования оптимальных k -стратегий в игре $\Gamma(x^0, T)$ построим на каждом шаге t ($t = 1, \dots, T$) рекуррентное семейство (уровневых) матричных игр и определим их цены:

$$V_1^t(x(t) \mid x^{kl}(t-1), \dots, x^{kl}(1)) = val \left\| val h(x^{kl}(t) \mid x^{kl}(t-1), \dots, x^{kl}(1)) \right\|_{q \times r}, \quad (15)$$

$$V_2^t(x(t) \mid x(t-1), x^{kl}(t-2), \dots, x^{kl}(1)) = val \left\| V_1^t(x(t) \mid x^{kl}(t-1), \dots, x^{kl}(1)) \right\|_{q \times r}, \quad (16)$$

$$V_{t-1}^t(x(t) \mid x(t-1), \dots, x(2), x^{kl}(1)) = val \left\| V_{t-2}^t(x(t) \mid x(t-1), \dots, x(3), x^{kl}(2), x^{kl}(1)) \right\|_{q \times r}, \quad (17)$$

$$V_t^t(x(t) | x(t-1), \dots, x(1)) = \text{val} \left\| V_{t-1}^t(x(t) | x(t-1), \dots, x(2), x^{kl}(1)) \right\|_{q \times r}, \quad (18)$$

В соответствии с мощностью множества допустимости t -го уровня $G(x(t-1), t)$, существует $(q \cdot r)^t$ чисел вида (15), $(q \cdot r)^{t-1}$ чисел вида (16) и т. д., $q \cdot r$ чисел вида (17) и одно число вида (18).

Последовательность (15) – (18) для каждого фиксированного t ($t = 1, \dots, T$) представляет собой систему упорядоченных чисел (цен уровней матриц), предписанных к вершинам дерева, соответствующего развернутой в фазовом пространстве форме игры. Фрагмент такого дерева (при $t = 1, 2, 3$) приведен на рис. 1.

Лемма 1. Пусть $B^{kl} = \|b_{ij}^{kl}\|_{m \times n}$ – система из $q \cdot r$ $m \times n$ – матричных игр, имеющих седловые точки в чистых стратегиях, $k = 1, \dots, g$, $l = 1, \dots, r$. Тогда в $q \times r$ – матричной игре $\|val B^{kl}\|_{q \times r}$ существует седловая точка в чистых стратегиях.

Доказательство. Нижнюю и верхнюю цены игры $B = \|val B^{kl}\|_{q \times r}$ можно представить в виде:

$$\begin{aligned} \underline{v}(B) &= \max_{k=1, \dots, q} \min_{l=1, \dots, r} (val B^{kl}) = \max_{k=1, \dots, q} \min_{l=1, \dots, r} \\ & \quad (\max_{i=1, \dots, m} \min_{j=1, \dots, n} b_{ij}^{kl} = \min_{j=1, \dots, n} \max_{i=1, \dots, m} b_{ij}^{kl}), \\ \bar{v}(B) &= \min_{l=1, \dots, r} \max_{k=1, \dots, q} (val B^{kl}) = \min_{l=1, \dots, r} \max_{k=1, \dots, q} \\ & \quad (\max_{i=1, \dots, m} \min_{j=1, \dots, n} b_{ij}^{kl} = \min_{j=1, \dots, n} \max_{i=1, \dots, m} b_{ij}^{kl}) \end{aligned}$$

Предположим, что $\underline{v}(B) \neq \bar{v}(B)$. Так как $\underline{v}(B) \leq \bar{v}(B)$ всегда, то наше предположение сводится к неравенству $\underline{v}(B) < \bar{v}(B)$.

Из неравенства

$$\max_k \min_l (val B^{kl}) < \min_k \max_l (val B^{kl})$$

следует:

$$\begin{aligned} \min_l (val B^{kl}) &< \min_l \min_l (val B^{kl}), \\ k &= 1, \dots, q, \end{aligned}$$

и далее:

$$\begin{aligned} \min_l (val B^{kl}) &< \max_k (val B^{kl}), \quad k = 1, \dots, q, \\ l &= 1, \dots, r. \end{aligned}$$

Следовательно, для любой фиксированной пары k и l выполняется неравенство:

$$\min_l (\max_i \min_j b_{ij}^{kl}) < \max_k (\max_i \min_j b_{ij}^{kl}). \quad (19)$$

Так как игра B^{kl} имеет седловую точку, то (19) можно переписать в виде:

$$\min_l (\max_j \min_i b_{ij}^{kl}) < \max_k (\max_i \min_j b_{ij}^{kl}). \quad (20)$$

Предположим, что внешние экстремумы в (20) достигаются при $l = \bar{l}$ и $k = k^*$. Тогда для всех $k = 1, \dots, q$ и $l = 1, \dots, r$

$$\min_j \max_i b_{ij}^{k\bar{l}} < \max_i \min_j a^{k^*\bar{l}}$$

и, в частности,

$$\min_j \max_i b_{ij}^{k\bar{l}} < \max_i \min_j a_{ij}^{k^*\bar{l}}.$$

Последнее неравенство означает, что

$$\bar{v}(B^{k\bar{l}}) < \underline{v}(B^{k^*\bar{l}}).$$

Полученное противоречие говорит о несостоятельности нашего предположения и на самом деле справедливо равенство $\underline{v}(B) = \bar{v}(B)$. Лемма доказана.

Лемма 2. Если во всех играх $h(x)$, $x \in G(x^0, T)$ существуют седловые точки в чистых стратегиях, то существуют все цены (15) – (18) и соответствующие им седловые точки в чистых стратегиях.

Справедливость этого утверждения непосредственно вытекает из леммы 1.

Теорема 1. Если во всех играх $h(x)$, $x \in G(x^0, T)$ существуют седловые точки в чистых стратегиях, то в динамической матричной игре $\Gamma(x^0, T)$ существует единственная седловая точка $(\varphi_1^*(\cdot), \varphi_2^*(\cdot))$, единственная оптимальная (в смысле принципа минимакса) траектория $x^*(\cdot)$ и цена игры в чистых k -стратегиях, вычисляемая по формуле:

$$\text{val } \Gamma(x^0, T) = \sum_{t=1}^T a_{i_t^* j_t^*} (x^*(t)),$$

где $\{(i_t^*, j_t^*), t = 1, \dots, T\}$ – седловая точка вдоль траектории $x^*(\cdot)$.

Доказательство. На каждом шаге t ($t = 1, \dots, T$), последовательно применяя в матричных играх

$$h(x), \quad x \in G(x(t-1), t),$$

$$\left\| V_1^t(x(t) | x^{kl}(t-1), \dots, x^{kl}(1)) \right\|_{q \times r},$$

$$k = 1, \dots, q, \quad l = 1, \dots, r;$$

$$\left\| V_{t-1}^t(x(t) | x(t-1), \dots, x(2), x^{kl}(1)) \right\|_{q \times r}$$

$$k = 1, \dots, q, \quad l = 1, \dots, r,$$

принцип минимакса, получаем число $V_t^t(x(t) | x(t-1), \dots, x(1))$ – итоговую цену семей-

ства матричных игр t -уровня дерева игры (в фазовом пространстве (см. рис. 1)), которая, согласно лемме 2, существует.

Предысторию точки $x(t) = x^*(t)$, соответствующую числу $V_t^t(x(t) | x(t-1), \dots, x(1))$, обозначим $x^*(t-1), \dots, x^*(1), x_0$, т. е.

$$x^*(t) = x^*(t) | x^*(t-1), \dots, x^*(1).$$

Седловую точку $m \times n$ – матричной игры $h(x^*(t))$, обозначим (i_t^*, j_t^*) (по условию теоремы она существует).

По построению последовательности (15) – (18) и исходя из леммы 2, последовательность $\{(i_t^*, j_t^*), t = 1, \dots, T\}$ является седловой точкой вдоль траектории $x^*(\cdot) = \{x^0, x^*(1), \dots, x^*(T)\}$, а цена игры $\Gamma(x^0, T)$ равна:

$$val \Gamma(x^0, T) = \sum_{t=1}^T a_{i_t^*, j_t^*}(x^*(t)).$$

Допустимые управления, соответствующие траектории $x^*(\cdot)$, обозначим $(u_1^*(\cdot), u_2^*(\cdot)) = \{(u_1^*(t), u_2^*(t)), t = 1, \dots, T\}$. Оптимальные стратегии игроков в игре $\Gamma(x^0, T)$ имеют вид:

$$\varphi_1^*(\cdot) = \{(u_1^*(t), i_t^*), t = 1, \dots, T\}, \varphi_2^*(\cdot) = \{(u_2^*(t), i_t^*), t = 1, \dots, T\},$$

а траектория $x^*(\cdot) = x(\cdot, x^0, u_1^*(\cdot), \dots, u_2^*(\cdot))$ оптимальна в смысле принципа минимакса. Теорема доказана.

Теорема 1 является конструктивной, так как из ее доказательства следует алгоритм построения оптимальных стратегий, оптимальной траектории и цены игры $\Gamma(x^0, T)$.

В теории игр достаточно примеров, когда принципы оптимальности, определяющие наилучшие исходы (решения) статических игр, будучи примененными в динамических играх, оказываются несостоятельными во времени, т. е. динамически неустойчивыми (см., напр., [4] – [6]).

Для определения принципа динамической устойчивости в игре $\Gamma(x^0, T)$, как обобщения принципа оптимальности Р. Беллмана из теории оптимального управления, погрузим ее в семейство $\{\Gamma(x^*(t), T-t), t = 0, 1, \dots, T\}$ аналогичных игр, определенных вдоль оптимальной траектории $x^*(\cdot)$. Текущая игра

$$\Gamma(x^*(t), T-t) = \left\langle \sum(x^*(t), T-t); \Phi_1[t, T], \Phi_2[t, T], H[t, T] \right\rangle$$

отличается от исходной игры $\Gamma(x^0, T)$ только начальным состоянием и продолжительностью

$(\sum(x^*(t), T-t)$ – система (1), определенная на отрезке $[t, T]$, $\Phi_1[t, T]$ и $\Phi_2[t, T]$ – сужение множеств Φ_1 и Φ_2 на отрезке $[t, T]$,

$$H[t, T] = \sum_{\tau=t}^T a_{i_\tau j_\tau}(x(\tau)).$$

Определение 4. Будем говорить, что оптимальная (в смысле принципа минимакса) траектория $x^*(\cdot)$ игры $\Gamma(x^0, T)$ динамически устойчива, если выполнены следующие два условия:

1) во всех играх $\Gamma(x^*(t), T-t)$, $t = 0, 1, \dots, T$ существуют цены в чистых k -стратегиях;

$$2) \sum_{\tau=1}^{t-1} a_{i_\tau^* j_\tau^*}(x^*(\tau)) + val \Gamma(x^*(t), T-t) = val \Gamma(x^0, T). \quad (21)$$

Теорема 2. Пусть во всех играх $h(x)$, $x \in G(x^0, T)$ существуют седловые точки в чистых стратегиях. Тогда оптимальная (в смысле принципа минимакса) траектория игры $\Gamma(x^0, T)$ динамически устойчива.

Доказательство. Выполнение первого условия определения 4 следует из теоремы 1. Покажем выполнения равенства (21). Применяя соотношения (9) – (10) и утверждение теоремы 1, можно записать:

$$\sum_{\tau=1}^{t-1} a_{i_\tau^* j_\tau^*}(x^*(\tau)) + \max_{\varphi_1 \in \Phi_1[t, T]} \min_{\varphi_2 \in \Phi_2[t, T]} v(x^0, \varphi_1, \varphi_2) = \sum_{t=1}^T a_{i_t^* j_t^*}(x^*(t)). \quad (22)$$

Так как условие 1 определения 4 выполнено, то

$$\max_{\varphi_1 \in \Phi_1[t, T]} \min_{\varphi_2 \in \Phi_2[t, T]} v(x^0, \varphi_1, \varphi_2) = \min_{\varphi_2 \in \Phi_2[t, T]} \max_{\varphi_1 \in \Phi_1[t, T]} v(x^0, \varphi_1, \varphi_2).$$

Следовательно, равенство (22) можно переписать в виде:

$$\sum_{\tau=1}^{t-1} a_{i_\tau^* j_\tau^*}(x^*(\tau)) + \max_{(i_t, \dots, i_T) \in \hat{M}[t, T]} \min_{(j_1, \dots, j_T) \in \hat{N}[t, T]} \sum_{\tau=t}^T a_{i_\tau j_\tau}(x(\tau)) = \sum_{t=1}^T a_{i_t^* j_t^*}(x^*(t)),$$

где $\hat{M}[t, T] = \underbrace{M \times \dots \times M}_{T-t}$, $\hat{N}[t, T] = \underbrace{N \times \dots \times N}_{T-t}$.

Отсюда, с учетом справедливости условия 1 определения 4, получаем:

$$\left[\begin{array}{l} \max_{(i_1, \dots, i_T) \in \tilde{M}[t, T]} \min_{(j_1, \dots, j_T) \in \tilde{N}[t, T]} \\ \sum_{\tau=t}^T a_{i_\tau j_\tau}(x(\tau)) = \min_{(j_1, \dots, j_T) \in \tilde{N}[t, T]} \max_{(i_1, \dots, i_T) \in \tilde{M}[t, T]} \\ \sum_{\tau=t}^T a_{i_\tau j_\tau}(x(\tau)) \end{array} \right] = \sum_{\tau=t}^T a_{i_\tau^* j_\tau^*}(x^*(t)).$$

Следовательно,

$$val \Gamma(x^*(t), T-t) = \sum_{\tau=t}^T a_{i_\tau^* j_\tau^*}(x(\tau)).$$

Подставляя выражения для $val \Gamma(x^*(t), T-t)$ и $val \Gamma(x^0, T)$ в (21), получаем тождество. Теорема доказана.

Теорема 2 показывает, что принцип минимакса относится к тем принципам оптимальности, динамическая устойчивость которых гарантируется фактом существования решения игры.

Литература

1. Данилов, Н. Н. Представление динамической матричной игры в форме задачи конфликтного управления / Н. Н. Данилов // Вестник КемГУ. – 2009. – Вып. 2 (38).
2. Оуэн, Г. Теория игр / Г. Оуэн. – М.: Мир, 1971.
3. Воробьев, Н. Н. Теория игр. Лекции для экономистов-кибернетиков / Н. Н. Воробьев. – Л.: Изд-во ЛГУ, 1974.
4. Петросян, Л. А. Устойчивость решений в дифференциальных играх со многими участниками / Л. А. Петросян // Вестник ЛГУ. – 1977. – № 19.
5. Петросян, Л. А. Кооперативные дифференциальные игры и их приложения / Л. А. Петросян, Н. Н. Данилов. – Томск: Изд-во ТГУ, 1985.
6. Данилов, Н. Н. Кооперативное поведение в динамических системах со многими управлениями / Н. Н. Данилов. – Томск: Изд-во ТГПУ, 2008. – 232 с.