Математическая Теория Игр и её Приложения, т.6, в.1, с. 56–72

УДК 519.83
ББК 22.18

СТРАТЕГИЧЕСКАЯ УСТОЙЧИВОСТЬ ОДНОТОЧЕЧНЫХ ПРИНЦИПОВ ОПТИМАЛЬНОСТИ В КООПЕРАТИВНЫХ СТОХАСТИЧЕСКИХ ИГРАХ

Елена М. Парилина
Санкт-Петербургский государственный университет
Факультет прикладной математики –
процессов управления
198504, Санкт-Петербург, Университетский пр., 35
e-mail: e.parilina@spbu.ru

В работе рассматриваются кооперативные стохастические игры, для которых актуальной является задача построения устойчивой кооперации. Одним из условий устойчивой кооперации является стратегическая устойчивость выбранного игроками кооперативного решения или кооперативного принципа оптимальности, гарантирующая, что исход кооперативного поведения игроков достигается при некотором равновесии по Нэшу. В работе получены достаточные условия для того, чтобы одноточечный кооперативный принцип оптимальности обладал свойством стратегической устойчивости.

Ключевые слова: кооперативная стохастическая игра, равновесие по Нэшу, стратегическая устойчивость, процедура распределения дележа, позиционная состоятельность, динамическая устойчивость.

1. Введение

В работе рассматривается класс динамических игр – стохастические игры, в которых допускается кооперация игроков. Построение кооперативного варианта игры, а именно, определение характеристической функции, задающей кооперативную игру, было предложено в работе [12], но стохастическая игра в этой работе была определена на конечном древовидном графе. В настоящей работе рассматривается стохастическая игра, определение которой для случая двух игроков введено в статье [14], в этом случае множество состояний или одновременных игр, которые могут реализоваться на каждом шаге игры, конечно и не меняется на протяжении всей игры, продолжительность игры бесконечна, и выигрыши игроков дисконтируются с некоторым коэффициентом из интервала $(0, 1)$, называемым дисконтирующим фактором. Построение характеристической функции, и, как следствие, кооперативной формы стохастической игры в классе стационарных стратегий предложено в работе [6]. В качестве решения кооперативной игры может быть выбран любой кооперативный принцип оптимальности (например, вектор Шепли, C-ядро, N-ядро).

Поскольку стохастическая игра является динамической, то возникает вопрос о позиционной состоятельности (динамической устойчивости) выбранного игроками кооперативного принципа оптимальности [2]. В общем случае, кооперативный принцип оптимальности не является позиционно состоятельным [3], но перераспределение выигрышей игроков на каждом шаге, проведенное с помощью кооперативной процедуры распределения дележа, впервые предложенной в работе [4], позволяет получить позиционно состоятельный принцип оптимальности. Помимо динамической устойчивости, дополнительными условиями устойчивой кооперации игроков являются условия защиты от иррационального поведения, а также стратегической устойчивости [5]. Условия устойчивой кооперации для класса стохастических игр были получены в [1].

Настоящая работа посвящена нахождению достаточных условий стратегической устойчивости одноточечных кооперативных принципов оптимальности в стохастической игре. Это условие позволяет гарантировать, что исход кооперативного поведения достигается при некотором равновесии по Нэшу, т.е. ни один из игроков не может увеличить свой выигрыш, отклонившись индивидуально от кооперативного поведения. В данной работе рассматриваются различные походы к построению равновесия по Нэшу, гарантирующего стратегическую устойчивость кооперативного принципа оптимальности.

Работа организована следующим образом: в разделе 2 вводится определение стохастической игры и ее кооперативной формы. В разделе 3 содержатся основные результаты о стратегической устойчивости принципов оптимальности в кооперативной стохастической игре, в разделе 4 представлен численный пример нахождения кооперативного принципа оптимальности и проверки условий, достаточных для его стратегической устойчивости. Выводы о результатах и открытые вопросы приводятся в Заключении.

2. Кооперативная стохастическая игра

Определение 2.1. *Стохастической игрой G назовем набор*

$$G = \left\langle N, \Gamma, \pi, p^0, \delta \right\rangle, \tag{2.1}$$

где

- $N = \{1, \ldots, n\}$ – *конечное множество игроков.*
- $\Gamma = \{\Gamma^j\}_{j=1}^{t}$ – *конечное множество состояний, каждое из которых – одновременная игра n лиц $\Gamma^j = \langle N, X_1^j, \ldots, X_n^j, K_1^j, \ldots, K_n^j \rangle$, множество игроков N одинаково для всех Γ^j, X_i^j – конечное множество стратегий игрока i, $K_i^j(x^j)$ – функция выигрыша игрока i в состоянии Γ^j, $x^j \in \prod\limits_{i \in N} X_i^j$.*
- $\pi : T \to \Delta(\Gamma)$ *определяет вероятности перехода, $T = \{(\Gamma^j, x^j) | j = 1, \ldots, t; x^j \in X^j\}$ и $\Delta(\Gamma)$ – вероятностное распределение на множестве состояний Γ, т. е. $\pi(\Gamma^j, x^j) = (\pi(1|j; x^j), \ldots, \pi(t|j; x^j))$.*
- $p^0 = (p_1^0, \ldots, p_t^0)$ – *вектор начального распределения по состояниям множества Γ.*
- $\delta \in (0, 1)$ – *дисконтирующий фактор.*

Стохастическая игра происходит следующим образом. До первого шага игры выбирается начальное состояние из множества Γ в соответствии с вектором начального распределения p^0. Допустим, было выбрано состояние Γ^j, т.е. игроки на первом шаге игры G играют в игру Γ^j, и реализуют ситуацию $x^j = (x_1^j, \dots, x_n^j) \in \prod\limits_{i \in N} X_i^j$, и игрок i получает выигрыш $K_i^j(x^j)$, $i \in N$. Далее стохастическая игра переходит на следующий шаг, на котором реализуется состояние Γ^k с вероятностью $\pi(k|j; x^j)$ и т.д.

Определение 2.2. *Стохастической подыгрой G^j, $j = 1, \dots, t$ назовем стохастическую игру* (2.1) *с вектором $p^0 = (0, \dots, 0, 1, 0, \dots, 0)$* (*с единицей на j-ом месте*), *т. е. стохастическую игру, начинающуюся из состояния Γ^j.*

Будем предполагать, что игроки используют чистые стационарные стратегии в стохастической игре. Обозначим через $\eta_i : \Gamma \to X_i$ чистую стационарную стратегию игрока i, т.е. для любого $j = 1, \dots, t$ стратегия $\eta_i(\Gamma^j)$ определяет чистую стратегию $x_i^j \in X_i^j$. Тогда $\eta = (\eta_1, \dots, \eta_n)$ – ситуация в стохастической игре G. Очевидно, что стратегия η_i в игре G является также стратегией в любой подыгре G^j. Множество чистых стационарных стратегий игрока i в стохастической игре G обозначим через H_i.

В качестве выигрыша игрока i в стохастической игре G (подыгре G^j) будем рассматривать математическое ожидание его дисконтированного выигрыша, который зависит от ситуации η, и будем обозначать его через $E_i(\eta)$ $(E_i^j(\eta))$. Если стратегии игроков в стохастической игре – стационарные, то возможно только t подыгр. В дальнейшем, будем использовать вектор $\mathbf{E}_i(\eta) = (E_i^1(\eta), \dots, E_i^t(\eta))^T$ для обозначения вектора математических ожиданий дисконтированных выигрышей игрока i в подыграх $G^1, \dots, G^t$ соответственно. Очевидно, что $E_i(\eta) = p^0 \mathbf{E}_i(\eta)$. Вектор $\mathbf{E}_i(\eta)$ можно вычислить по формуле [9, 15]:

$$\mathbf{E}_i(\eta) = (\mathbf{I}_t - \delta \Pi(\eta))^{-1} \mathbf{K}_i(\eta), \tag{2.2}$$

где $\mathbf{K}_i(\eta) = (K_i^1(x^1), \dots, K_i^t(x^t))^T$, ситуация η такая, что $\eta(\Gamma^j) = x^j$, $j = 1, \dots, t$, $\mathbf{I}_t$ – единичная матрица размерности $t \times t$, и матрица

$\Pi(\eta)$ определяется следующим образом:

$$\Pi(\eta) = \begin{pmatrix} \pi(1|1;x^1) & \ldots & \pi(t|1;x^1) \\ \pi(1|2;x^2) & \ldots & \pi(t|2;x^2) \\ \ldots & \ldots & \ldots \\ \pi(1|t;x^t) & \ldots & \pi(t|t;x^t) \end{pmatrix}. \tag{2.3}$$

Для определения кооперативной стохастической игры необходимо задать характеристическую функцию как для всей стохастической игры G, так и для любой ее подыгры G^j. Предположим, что игроки из множества N решили объединиться с целью получения максимального суммарного выигрыша. Найдем ситуацию в чистых стационарных стратегиях, максимизирующую сумму математических ожиданий дисконтированных выигрышей игроков в игре G:

$$\bar{\eta} = \arg \max_{\eta \in \prod\limits_{i \in N} H_i} \sum_{i \in N} E_i(\eta).$$

Ситуацию $\bar{\eta}$ будем называть кооперативным поведением в стохастической игре G.

Определим характеристическую функцию для подыгры G^j, $j = 1, \ldots, t$ следующим образом:

$$V^j(S) = \begin{cases} E_i^j(\bar{\eta}), & \text{если } S = N\,, \\ \max\limits_{\eta_S \in \prod\limits_{i \in S} H_i} \min\limits_{\eta_{N \setminus S} \in \prod\limits_{i \in N \setminus S} H_i} \sum\limits_{i \in S} E_i^j(\eta_S, \eta_{N \setminus S}), & \text{если } S \subset N\,, \\ 0, & \text{если } S = \emptyset\,. \end{cases} \tag{2.4}$$

Если $\mathbf{V}(S) = (V^1(S), \ldots, V^t(S))$, то (2.4) можно переписать следующим образом:

$$\mathbf{V}(S) = \begin{cases} \mathbf{E}_i(\bar{\eta}) = (\mathbf{I}_t - \delta \Pi(\bar{\eta}))^{-1} \sum\limits_{i \in N} K_i(\bar{\eta}), & \text{если } S = N\,, \\ \max\limits_{\eta_S \in \prod\limits_{i \in S} H_i} \min\limits_{\eta_{N \setminus S} \in \prod\limits_{i \in N \setminus S} H_i} \sum\limits_{i \in S} \mathbf{E}_i(\eta_S, \eta_{N \setminus S}), & \text{если } S \subset N\,, \\ \mathbf{0}, & \text{если } S = \emptyset\,. \end{cases} \tag{2.5}$$

Определим характеристическую функцию $V(S)$ в игре G следующим образом:

$$V(S) = p^0 \mathbf{V}(S), \quad S \subseteq N. \tag{2.6}$$

Определение 2.3. *Кооперативной стохастической игрой* (*подыгрой, начинающейся из состояния* Γ^j) *назовем набор* $\langle N, V(S)\rangle$, *где* $V(S) : 2^N \longrightarrow R$ – *характеристическая функция, определяемая формулой* (2.6) ($\langle N, V^j(S)\rangle$, *где* $V^j(S) : 2^N \longrightarrow \mathbb{R}$ – *характеристическая функция, определяемая формулой* (2.4)).

Характеристические функции $V(S)$ и $V^j(S)$ являются супераддитивными по построению. Определение множества дележей в подыгре $G^j, j = 1, \ldots, t$ вводится классическим для кооперативной теории игр способом, а именно, это множество $A^j = \Big\{\alpha^j = (\alpha_1^j, \ldots, \alpha_n^j) : \sum\limits_{i \in N} \alpha_i^j = = V^j(N), \alpha_i^j \geqslant V^j(\{i\}), \forall i \in N\Big\}$. Множеством дележей в игре G будем называть множество $A = \{\alpha = (\alpha_1, \ldots, \alpha_n) : \alpha_i = \sum_{j=1}^t p_j^0 \alpha_i^j, \quad \forall i \in \in N$ и $\alpha^j = (\alpha_1^j, \ldots, \alpha_n^j) \in A^j, \quad \forall j = 1, \ldots, t\Big\}$. Нетрудно показать, что при таком определении дележа α в игре G, он также обладает свойствами индивидуальной и коллективной рациональности, это следует из определения характеристической функции (2.6). Под кооперативным принципом оптимальности будем понимать подмножество множества дележей. Предположим, что кооперативный принцип оптимальности является непустым множеством для игры G и всех подыгр G^j.

3. Стратегическая устойчивость кооперативного принципа оптимальности

Предположим, что игроки объединились в коалицию N и выбрали в качестве дележа в каждой подыгре G^j вектор $\alpha^j = (\alpha_1^j, \ldots, \alpha_n^j)$ из некоторого принципа оптимальности $C(G^j)$, одинакового для всех подыгр G^j, $j = 1, \ldots, t$. Тогда дележ в игре G есть вектор $\alpha = = (\alpha_1, \ldots, \alpha_n)$, где $\alpha_i = \sum\limits_{j=1}^{t} p_j^0 \alpha_i^j$ для любого $i \in N$.

В результате реализации кооперативного поведения, или ситуации $\bar{\eta}$, игроки с течением времени попадают в подыгры (точнее, для стохастической игры, определенной ранее, они могут попасть в t различных подыгр), в которых их индивидуальные ожидаемые дисконтированные выигрыши могут не совпадать с компонентами дележа, принадлежащего тому же принципу оптимальности $C(G^j)$, который был выбран ими в качестве принципа распределения суммарного ожидаемого дисконтированного выигрыша. То есть может не выполняться следующее условие: для всех подыгр G^j, $j = 1,\dots,t$ и для любой компоненты α_i^j дележа из принципа оптимальности $C(G^j)$ существуют $\tilde{\alpha}_i^l$ такие, что $\tilde{\alpha}^l = (\tilde{\alpha}_1^l,\dots,\tilde{\alpha}_n^l) \in C(G^l)$, что α_i^j представима в виде

$$\alpha_i^j = K_i^j(\bar{x}^j) + \delta \sum_{l=1}^{t} \pi(l|j;\bar{x}^j)\tilde{\alpha}_i^l, \quad i \in N. \tag{3.1}$$

В случае, когда выбранный игроками принцип оптимальности – одноточечный, условие (3.1) может быть записано в векторном виде следующим образом:

$$(\mathbf{I}_t - \delta\Pi(\bar{\eta}))\alpha_i = \mathbf{K}_i(\bar{\eta}), \tag{3.2}$$

где $\alpha_i = (\alpha_i^1,\dots,\alpha_i^t)^T$ – вектор, состоящий из i-ых компонент дележей $\alpha^1,\dots,\alpha^t$, каждый из которых определяется единственным образом, $\alpha^j = C(G^j)$, $j = 1,\dots,t$. Таким образом, производя выплаты игрокам в течение игры в соответствии с их функциями выигрыша, определенными стохастической игрой G, а именно, $K_i^j(\bar{x}^j)$ игроку i в состоянии Γ^j, принцип оптимальности может оказаться позиционно несостоятельный (динамически неустойчивым). В связи в этим, в работе [4] была предложена процедура регуляризации (определения процедуры перераспределения выплат игрокам на каждом шаге игры) динамически неустойчивого принципа оптимальности. Ее реализация для класса стохастических игр учитывает то, что в игре G при использовании стационарных стратегий возможно лишь конечное число подыгр. В связи с этим, для каждого $\alpha^j \in C(G^j)$ переопределим функции выигрыша игроков в игре G по правилу

$$K_i^{\alpha,j}(x^j) = \begin{cases} \beta_i^j, & \text{если } x^j = \bar{x}^j, \\ K_i^j(x^j), & \text{если } x^j \neq \bar{x}^j, \end{cases} \tag{3.3}$$

где $\beta = (\beta_1,\dots,\beta_n)$, $\beta_i = (\beta_i^1,\dots,\beta_i^t)^T$ называется процедурой распределения дележа и находится по формуле

$$\beta_i^j = \alpha_i^j - \delta \sum_{l=1}^{t} \pi(l|j;\bar{x}^j)\tilde{\alpha}_i^l, \tag{3.4}$$

где $\tilde{\alpha}^l \in C(G^l)$, или в случае одноточечного принципа оптимальности по формуле

$$\beta_i = (\mathbf{I}_t - \delta\Pi(\bar{\eta}))\alpha_i. \tag{3.5}$$

Будем называть стохастическую игру с функциями выигрыша в состояниях, определяемыми формулой (3.3), α-регуляризацией игры G (подыгр G^j).

Рассмотрим условие стратегической устойчивости принципа оптимальности при условии, что проведена α-регуляризация игры G. Предположим, что игроки перед переходом игры на следующий шаг знают, какая ситуация реализовалась на текущем шаге. Это предположение существенно для построения стратегически устойчивого принципа оптимальности. В работе [1] доказана следующая теорема, содержащая достаточные условия стратегической устойчивости.

Теорема 3.1. *Если $C(\cdot)$ – одноточечный принцип оптимальности, и процедура распределения дележа β_i, определяемая формулой* (3.5)*, удовлетворяет следующему неравенству для любого игрока $i \in N$:*

$$\beta_i \geqslant (\mathbf{I}_t - \delta\Pi(\bar{\eta}))\mathbf{W}_i, \tag{3.6}$$

где $\mathbf{W}_i = (W_i^1, \ldots, W_i^t)^T$,

$$W_i^j = \max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \bar{x}_i^j}} \left\{ K_i^{\alpha,j}(\bar{x}^j \parallel x_i^j) + \delta \sum_{l=1}^{t} \pi(l|j; \bar{x}^j \parallel x_i^j) V^l(\{i\}) \right\},$$

тогда в α-регуляризации игры G существует ситуация равновесия по Нэшу с выигрышами $(\alpha_1, \ldots, \alpha_n)$.

Доказательство теоремы 3.1 схоже по построению с доказательствами «народных теорем» с использованием специальным образом построенных стратегий наказания. Если на текущем шаге игры кто-то из игроков отклоняется от кооперативного поведения $\bar{\eta}$ в индивидуальном порядке, то перед началом следующего шага все остальные игроки узнают об этом и, начиная со следующего шага и до конца игры, используют стратегии наказания, т. е. начинают играть в антагонистическую игру против отклонившегося игрока.

Однако, может быть предложено отличное от вышеописанного поведение игроков в случае отклонения игрока i. Пусть игроки перед началом подыгры G^k идентифицировали отклонение игрока i от кооперативного решения $\bar{\eta}$, тогда они могут, начиная со следующего шага, реализовать стратегии равновесные по Нэшу в подыгре G^k. Существование равновесия по Нэшу в стохастической игре n лиц в классе смешанных стационарных стратегий было доказано в работах

[9, 15]. Для построения стратегии наказания в стохастической игре, рассматриваемой в классе чистых стационарных стратегий, где наказанием отклонившегося игрока является реализация равновесных по Нэшу стратегий, необходимо предположить существование ситуации равновесия по Нэшу в этом классе стратегий. В случае, если в игре существует несколько ситуаций равновесия по Нэшу, предположим, что игроки договариваются об использовании одной из них для построения стратегий наказания.

Теорема 3.2. *Пусть в стохастической игре G существует ситуация равновесия по Нэшу η^{NE} в чистых стационарных стратегиях. Если $C(\cdot)$ – одноточечный принцип оптимальности, и β_i – процедура распределения дележа, определяемая формулой* (3.5), *для любого $i \in N$ удовлетворяет неравенству:*

$$\beta_i \geqslant (\mathbf{I}_t - \delta\Pi(\bar{\eta}))\mathbf{U}_i, \tag{3.7}$$

где $\mathbf{U}_i = (U_i^1, \ldots, U_i^t)^T$,
$$U_i^j = \max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \bar{x}_i^j}} \left\{ K_i^{\alpha,j}(\bar{x}^j \parallel x_i^j) + \delta \sum_{l=1}^{t} \pi(l|j; \bar{x}^j \parallel x_i^j) E_i^l\left(\eta^{NE}\right) \right\},$$
тогда в α-регуляризации игры G существует ситуация равновесия по Нэшу с выигрышами $(\alpha_1, \ldots, \alpha_n)$.

Доказательство. Обозначим через $\Gamma(k)$ состояние в игре G, реализовавшееся на шаге k, $\Gamma(k) \in \{\Gamma^1, \ldots, \Gamma^t\}$. Через $x(k)$ будем обозначать ситуацию, реализовавшуюся в состоянии $\Gamma(k)$. Подыгру игры G, начинающуюся из состояния $\Gamma(k)$, обозначим через $G^{\Gamma(k)}$. Предысторией шага k будем называть последовательность $((\Gamma(1), x(1)), (\Gamma(2), x(2)), \ldots, (\Gamma(k-1), x(k-1)))$, которую обозначим через $h(k)$. Обозначим через T набор $\{(\Gamma^1, \bar{x}^1), (\Gamma^2, \bar{x}^2), \ldots, (\Gamma^t, \bar{x}^t)\}$.

Рассмотрим ситуацию $\widehat{\varphi}(h) = (\widehat{\varphi}_1(h), \ldots, \widehat{\varphi}_n(h))$ в стратегиях поведения как функцию предыстории в α-регуляризации игры G:

$$\widehat{\varphi}_i(h(k)) = \begin{cases} \bar{x}_i^j, & \text{если } \Gamma(k) = \Gamma^j, j = \overline{1,t}, h(k) \subset T, \\ x_i^{NE,j}, & \text{если } \Gamma(k) = \Gamma^j,\ j = \overline{1,t},\ \exists\ l \in [1, k-1] \\ & \text{и } p \in N,\ p \neq i\text{: } h(l) \subset T, \\ & \text{а } (\Gamma(l), x(l)) \notin T, \text{но } (\Gamma(l), (x(l) \,\|\, \bar{x}_p(l)) \in T, \\ \text{произвольна} & \text{в других случаях,} \end{cases} \tag{3.8}$$

где $x_i^{NE,j}$ – чистая стратегия i-го игрока в состоянии Γ^j такая, что $\eta_i^{NE}(\Gamma^j) = x_i^{NE,j}$, $i \in N$.

Используя структуру стратегии $\widehat{\varphi}_i(h)$, $i \in N$, докажем, что $\widehat{\varphi}(h) = (\widehat{\varphi}_1(h), \ldots, \widehat{\varphi}_n(h))$, определенная в (3.8), является равновесием по Нэшу в α-регуляризации игры G.

Из (3.8) следует, что ожидаемый дисконтированный выигрыш игрока i в α-регуляризации подыгры G^j, $j = 1, \ldots, t$, при условии, что все игроки придерживаются кооперативного поведения $\bar{\eta}$, равен

$$E_i^j(\widehat{\varphi}) = E_i^j(\bar{\eta}(\cdot)).$$

Пусть $\mathbf{E}_i(\widehat{\varphi}) = (E_i^1(\widehat{\varphi}), \ldots, E_i^t(\widehat{\varphi}))^T$, тогда для любого игрока $i \in N$ имеет место равенство:

$$\mathbf{E}_i(\widehat{\varphi}) = (\mathbf{I}_t - \delta\Pi(\bar{\eta}))^{-1}\beta_i. \tag{3.9}$$

Рассмотрим ситуацию $(\widehat{\varphi} \,\|\, \varphi_i)$, $i \in N$, когда некоторый игрок i отклоняется от своей стратегии $\widehat{\varphi}_i$. Пусть шаг k такой, что существует номер $l \in [1, k-1]$ такой, что предыстория $h(l) \subset T$, а элемент $(\Gamma(l), x(l)) \notin T$, но $(\Gamma(l), (x(l) \,\|\, \bar{x}_i(l))) \in T$. Не умоляя общности, предположим, что $\Gamma(k) = \Gamma^j$. Вычислим ожидаемый выигрыш игрока i в α-регуляризации игры G в ситуации $(\widehat{\varphi} \,\|\, \varphi_i)$ по формуле $E_i(\widehat{\varphi} \,\|\, \varphi_i) = p^0\mathbf{E}_i(\widehat{\varphi} \,\|\, \varphi_i)$, где

$$\mathbf{E}_i(\widehat{\varphi} \,\|\, \varphi_i) = \mathbf{E}_i^{[1,k-1]}(\widehat{\varphi} \,\|\, \varphi_i) + \delta^{k-1}\Pi^{k-1}(\widehat{\varphi} \,\|\, \varphi_i)\mathbf{E}_i^{[k,\infty)}(\widehat{\varphi} \,\|\, \varphi_i), \tag{3.10}$$

где первое слагаемое в правой части – ожидаемый дисконтированный выигрыш игрока i на первых $k-1$ шагах α-регуляризации игры G, во втором слагаемом $\mathbf{E}_i^{[k,\infty)}(\widehat{\varphi} \,\|\, \varphi_i)$ – ожидаемый дисконтированный выигрыш игрока i в подыгре игры G, начинающейся с шага k. Так как до шага $k-1$ включительно никакой из игроков не отклонялся от кооперативного поведения $\bar{\eta}$, то, как было показано ранее, для элементов из правой части (3.10) имеют место равенства

$$\mathbf{E}_i^{[1,k-1]}(\widehat{\varphi} \parallel \varphi_i) = \mathbf{E}_i^{[1,k-1]}(\bar{\eta}),$$

$$\Pi^{k-1}(\widehat{\varphi} \parallel \varphi_i) = \Pi^{k-1}(\bar{\eta}).$$

Во втором слагаемом в правой части (3.10) под $\mathbf{E}_i^{[k,\infty)}(\widehat{\varphi} \parallel \varphi_i)$ понимается вектор $(E_i^1(\widehat{\varphi} \parallel \varphi_i), \ldots, E_i^t(\widehat{\varphi} \parallel \varphi_i))^T$, где $E_i^j(\widehat{\varphi} \parallel \varphi_i)$ – ожидаемый дисконтированный выигрыш игрока $i \in N$ в α-регуляризации подыгры G^j, начинающейся из состояния Γ^j.

Вычислим ожидаемый дисконтированный выигрыш игрока i в α-регуляризации подыгры G^j, начинающейся с шага k, и $\Gamma(k) = \Gamma^j$. Имеет место равенство

$$E_i^j(\widehat{\varphi} \parallel \varphi_i) = K_i^{\alpha,j}(\bar{x}^j \parallel x_i^j) + \delta \sum_{l=1}^{t} \pi(l|j; \bar{x}^j \parallel x_i^j) E_i^j(\eta^{NE}), \tag{3.11}$$

поскольку, согласно определению ситуации $\widehat{\varphi}$, все игроки кроме i будут играть чистые стационарные стратегии η_i^{NE}, образующими равновесие по Нэшу в α-регуляризации G, поэтому игрок i может ожидать получение выигрыша $E_i^j(\eta^{NE})$ в подыгре G^j.

Так как ожидаемые выигрыши игрока i в ситуациях $\widehat{\varphi}$ и $(\widehat{\varphi} \parallel \varphi_i)$ до шага $k-1$ совпадают, то в результате отклонения игрок i может гарантировать себе увеличение выигрыша только за счет части игры G, начинающейся с шага k, т. е. за счет ожидаемого выигрыша в подыгре G^j, $j = 1, \ldots, t$. Игрок i в ситуации $(\widehat{\varphi} \parallel \varphi_i)$ может гарантировать себе с шага k следующий ожидаемый выигрыш:

$$\max_{\substack{x_i^j \in X_i^j \\ x_i^j \neq \bar{x}_i^j}} \left\{ K_i^{\alpha,j}(\bar{x}^j \parallel x_i^j) + \delta \sum_{l=1}^{t} \pi(l|j; \bar{x}^j \parallel x_i^j) E_i^j(\eta^{NE}) \right\}. \tag{3.12}$$

Ожидаемый выигрыш игрока i в α-регуляризации G^j в ситуации $\widehat{\varphi}$ согласно определению процедуры распределения дележа (3.5) может быть найден из уравнения

$$\mathbf{E}_i(\widehat{\varphi}) = (\mathbf{I}_t - \delta \Pi(\bar{\eta}))^{-1} \beta_i, \tag{3.13}$$

где $\mathbf{E}_i(\widehat{\varphi}) = (E_i^1(\widehat{\varphi}), \ldots, E_i^t(\widehat{\varphi}))^T$. Из (3.12), (3.13), учитывая неравенство (3.7) и рассуждения, приведенные выше, получаем справедливость неравенства

$$\mathbf{E}_i(\widehat{\varphi}) \geqslant \mathbf{E}_i(\widehat{\varphi} \parallel \varphi_i).$$

Следовательно, ситуация $\widehat{\varphi}$ в стратегиях поведения (3.8) является равновесием по Нэшу в α-регуляризации игры G. Причем, ожидаемый выигрыш i-го игрока в α-регуляризации игры G в ситуации $\widehat{\varphi}$ равен $\alpha_i = \sum\limits_{j=1}^{t} p_j^0 \alpha_i^t$, вектор $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^t)$ состоит из i-ых компонент дележей α^1, …, α^t, рассчитанных для кооперативных подыгр G^1, …, G^t соответственно. Утверждение теоремы доказано. $\square$

Замечание 3.1. Доказательство теоремы 3.1 можно обобщить на случай, когда отклоняется несколько игроков [1], т.е. показать, что при выполнении условия, аналогичного неравенству (3.6), в α-регуляризации игры G существует сильное трансферабельное равновесие с выигрышами $(\bar{\alpha}_1, \ldots, \bar{\alpha}_n)$. В этом случае игроки могут реализовать специальным образом построенную ситуацию в стратегиях наказания, где в качестве наказания отклонившейся коалиции не отклонившиеся игроки будут реализовывать стационарные стратегии, минимизирующие суммарный выигрыш отклонившейся коалиции во всей оставшейся игре. Аналогичный результат может быть доказан для случая наказания в виде реализации равновесных по Нэшу стратегий.

4. Пример

Рассмотрим стохастическую игру 3 лиц с 2 состояниями, каждое из которых – одновременная игра типа «дилемма заключенного»[1]. В первом состоянии Γ^1 у каждого игрока – 2 стратегии: A и B. Во втором состоянии Γ^2 у каждого игрока – 2 стратегии: A' и B'. Состояние Γ^1 задается матрицами выигрышей (игрок 1 выбирает строки, игрок 2 – столбцы, игрок 3 – матрицы):

$$\left[\begin{pmatrix} (6,6,6) & (3,10,3) \\ (10,3,3) & (6,6,0) \end{pmatrix} \qquad \begin{pmatrix} (3,3,10) & (0,6,6) \\ (6,0,6) & (1,1,1) \end{pmatrix}\right],$$

[1]Все вычисления этого раздела производились с программы, написанной в системе Maxima.

и состояние Γ^2 задается матрицами:

$$\left[\begin{pmatrix}(10,10,10) & (3,12,3)\\ (12,3,3) & (10,10,0)\end{pmatrix} \qquad \begin{pmatrix}(3,3,12) & (0,10,10)\\ (10,0,10) & (1,1,1)\end{pmatrix}\right].$$

Вероятности перехода из состояния Γ^1 в состояния Γ^1 и Γ^2 равны $1/3$ и $2/3$ соответственно, если два или три игрока выбрали одновременно стратегию A в состоянии Γ^1, и равны $2/3$ и $1/3$ соответственно, если никто из игроков или только один игрок выбрали стратегию A в состоянии Γ^1. Вероятности перехода из состояния Γ^2 в состояния Γ^1 и Γ^2 равны $1/4$ и $3/4$ соответственно, если два или три игрока выбрали одновременно стратегию A' в состоянии Γ^2, и равны $3/4$ и $1/4$ соответственно, если никто из игроков или только один игрок выбрали стратегию A' в состоянии Γ^2.

Пусть также вектор начального распределения по состояниям p^0 есть $(0.5; 0.5)$, и дисконтирующий фактор $\delta = 0.95$.

Кооперативным поведением в игре является ситуация $\bar\eta = (\bar\eta_1, \bar\eta_2)$, где $\bar\eta_i(\Gamma^1) = A$, $\bar\eta_i(\Gamma^2) = A'$, $i \in N = \{1, 2, 3\}$. Ожидаемые выигрыши игроков при кооперативном решении:

$$\begin{aligned}\mathbf{E}_i(\bar\eta) =& (175.02; 179.37)^T,\\ E_i(\bar\eta) =& p^0\mathbf{E}_i(\bar\eta) = 177.19\end{aligned}$$

для любого $i \in N$. Ожидаемый суммарный выигрыш коалиции N равен 531.58. Характеристическая функция для подыгр G^1 и G^2 определяется по формулам (2.4) и (2.5):

$$\begin{aligned}&\mathbf{V}(\{1\}) = \mathbf{V}(\{2\}) = \mathbf{V}(\{3\}) = (60.00; 60.00)^T,\\ &\mathbf{V}(\{1,2\}) = \mathbf{V}(\{1,3\}) = \mathbf{V}(\{2,3\}) = (156.30; 159.16)^T,\\ &\mathbf{V}(\{1,2,3\}) = (525.06; 538.10)^T.\end{aligned}$$

Значения характеристической функции для всей игры G определяем по формуле (2.6):

$$\begin{aligned}&V(\{1\}) = V(\{2\}) = V(\{3\}) = 60.00,\\ &V(\{1,2\}) = V(\{1,3\}) = V(\{2,3\}) = 157.73,\\ &V(\{1,2,3\}) = 531.58.\end{aligned}$$

В качестве принципа оптимальности рассмотрим N-ядро [13]. Известно, что этот принцип оптимальности является одноточечным. Вычислим его для подыгр G^1 и G^2:

$$\alpha^1 = (175.02; 175.02; 175.02)^T,$$
$$\alpha^2 = (179.367; 179.367; 179.367)^T,$$

и для всей игры N-ядро равно

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (177.19; 177.19; 177.19).$$

Выбранный принцип оптимальности (N-ядро) является позиционно состоятельным в рассматриваемой игре, так как

$$(\mathbf{I}_2 - \delta\Pi(\bar{\eta}))\alpha_i = \mathbf{K}_i(\bar{\eta}),$$

для любого $i \in N$, поэтому процедура распределения дележа, определяемая по формуле (3.5):

$$\beta_i = (\beta_i^1, \beta_i^2) = (6.00; 10.00)^T, \quad i \in N,$$

совпадает с вектором выигрышей

$$\mathbf{K}_i(\bar{\eta}) = (K_i^1(A, A, A), K_i^2(A', A', A'))^T$$

в состояниях Γ^1 и Γ^2 соответственно.

Однако, следует заметить, что кооперативное поведение $\bar{\eta}$, реализация которого позволяет игрокам достигнуть N-ядра, не является равновесием по Нэшу в стохастической игре. В игре G существует ситуация равновесия по Нэшу в чистых стратегиях, а именно, это ситуация $\eta^{NE} = (\eta_1^{NE}, \eta_2^{NE}, \eta_3^{NE})$, где $\eta_i^{NE}(\Gamma^1) = B$, $\eta_i^{NE}(\Gamma^2) = B'$ для всех $i \in N$. Предположим, что игроки будут использовать эту ситуацию для построения ситуации в стратегиях наказания, определенных в доказательстве теоремы 3.2. Проверим выполнение условия (3.7), оно может быть записано для рассматриваемой игры в следующем виде:

$$\begin{pmatrix} 6 \\ 10 \end{pmatrix} \geqslant \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 0.95 \begin{pmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{pmatrix} \right] \mathbf{U}_i \tag{4.1}$$

для любого $i \in N$, где

$$\mathbf{U}_i = \begin{pmatrix} 10 \\ 12 \end{pmatrix} + 0.95 \begin{pmatrix} 1/3 & 2/3 \\ 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} 20 \\ 20 \end{pmatrix}.$$

Очевидно, что условие (4.1) выполняется для всех игроков. Из теоремы 3.2 можно сделать вывод о том, что N-ядро является стратегически устойчивым принципом оптимальности в данной стохастической игре.

В примере найдено кооперативное поведение игроков, построена кооперативная форма стохастической игры, найдено N-ядро, показано, что оно является позиционно состоятельным и стратегически устойчивым принципом оптимальности в данной игре.

5. Заключение

В работе предложен способ проверки принципа оптимальности, выбранного в качестве решения кооперативной стохастической игры, на стратегическую устойчивость. Доказанные теоремы 3.1 и 3.2 позволяют проверить условия, достаточные для стратегической устойчивости одноточечного принципа оптимальности. В работе существенным является предположение о существовании ситуации равновесия по Нэшу в чистых стационарных стратегиях. Доказательство теоремы 3.2 основано на построении ситуации в стратегиях наказания с использованием равновесных по Нэшу чистых стационарных стратегий. В этом случае не возникает вопрос нахождения ситуации равновесия по Нэшу. Если рассматривать стохастическую игру в классе смешанных стационарных стратегий, то существование ситуации равновесия доказано в работах [9, 15], но открытым остается вопрос нахождения равновесия по Нэшу. Существует несколько алгоритмов, позволяющих найти равновесие по Нэшу в стационарных стратегиях в стохастической игре, на параметры которой наложены ограничения [7, 11]. Для общего случая, для нахождения равновесия по Нэшу в смешанных стационарных стратегиях обычно используют приближенные численные методы.

СПИСОК ЛИТЕРАТУРЫ

1. Парилина Е.М. *Устойчивая кооперация в стохастических играх* // Математическая теория игр и ее приложения. 2010. Т. 2. № 3. С. 21–40.

2. Петросян Л.А. *Устойчивость решений в дифференциальных играх со многими участниками* // Л.: Вестник Ленинградского университета. Сер. 1. 1977. Вып. 19. С. 46–52.

3. Петросян Л.А., Баранова Е.М., Шевкопляс Е.В. *Многошаговые кооперативные игры со случайной продолжительностью* // Оптимальное управление и дифференциальные игры. Сборник статей. Труды института математики и механики. 2004. Т. 10. № 2. С. 116–130.

4. Петросян Л.А., Данилов Н.А. *Устойчивые решения неантагонистических дифференциальных игр с транзитивными выигрышами* // Вестник ЛГУ. 1979. №1. С. 46–54.

5. Петросян Л.А., Зенкевич Н.А. *Принципы устойчивой кооперации* // Математическая теория игр и ее приложения. 2009. Т.1. Вып. 1. С. 102–117.

6. Baranova E.M., Petrosyan L.A. *Cooperative Stochastic Games in Stationary Strategies* // Game Theory and Applications. 2006. V. XI. P. 7–17.

7. Breton M. *Algorithms for stochastic games.* In: T.E.S. Raghavan, T.S. Ferguson, T. Parthasarathy, O.J. Vrieze (Eds.), Stochastic Games and Related Topics: In Honor of Professor L.S. Shapley, Theory and Decision Library C: Game Theory, Mathematical Programming and Operations Research. V. 7. Kluwer Academic Publishers, Dordrecht, 1991. P. 45—57.

8. Dutta P. *A Folk Theorem for Stochastic Games* // Journal of Economic Theory. 1995. V. 66. P. 1–32.

9. Fink A.M. *Equilibrium in a stochastic n-person game* // Journal of Science of the Hiroshima University. Series A-I. 1964. V. 28. N. 1. P. 89–93 .

10. Grauer L.V., Petrosjan L.A. *Strong Nash Equilibrium in Multistage Games* // International Game Theory Review. 2002. V. 4(3). P. 255-264.

11. Herings P.J.-J., Peeters R. J. A. P. *Stationary Equlibria in Stochastic Games: Structure, Selection, and Computation* // Journal of Economic Theory. 2004. Vol. 118. N. 1. P. 32–60.

12. Petrosjan L.A. *Cooperative Stochastic Games* // Advances in Dynamic Games. Annals of the ISDG. Application to Economics, Engineering and Environmental Management, ed. by A. Haurie, S. Muto, L. A. Petrosjan, T.E.S. Raghavan. 2006. P. 139–146.

13. Schmeidler D. *The nucleolus of a characteristic function game* // SIAM Journal of Applied Mathematics. 1969. V. 17 (6). P. 1163–1170.

14. Shapley L.S. *Stochastic Games* // Proceedings of National Academy of Sciences of the USA. 1953. V. 39. P. 1095–1100.

15. Takahashi M. *Equilibrium points of stochastic non-cooperative n-person games* // Journal of Science of the Hiroshima University. Series A-I. 1964. V. 28. N. 1. P. 95–99.

STRATEGIC STABILITY OF ONE-POINT OPTIMALITY PRINCIPLES IN COOPERATIVE STOCHASTIC GAMES

Elena M. Parilina, Saint Petersburg State University, Cand.Sc, assoc. prof. (e.parilina@spbu.ru).

Abstract: Cooperative stochastic games are investigated in the paper. The construction of stable cooperation is an actual problem for this class of dynamic games. One of stable cooperation principles is the strategic stability of the cooperative decision or the cooperative optimality principle chosen by the players. Strategic stability guarantees that the players' cooperative payoffs can be obtained if players realize a Nash equilibrium. Sufficient conditions of the strategic stability of the single-point optimality principle are obtained.

Keywords: stochastic game, Nash equilibrium, strategic stability, payoff distribution procedure, subgame consistency, time consistency.