

УДК 518.9 + 517.9

ББК 65.050.2

ПРИНЦИПЫ УСТОЙЧИВОЙ КООПЕРАЦИИ

Л.А. ПЕТРОСЯН

Факультет прикладной математики –
процессов управления

Санкт-Петербургский государственный университет

Санкт-Петербург

e-mail: spbuoasis7@peterlink.ru

Н.А. ЗЕНКЕВИЧ

Высшая школа менеджмента

Санкт-Петербургский государственный университет

Санкт-Петербург

e-mail: zenkevich@gsom.pu.ru

Имеются три важных условия, которые должны быть исследованы, если рассматривается проблема устойчивости долгосрочного кооперативного соглашения: временная состоятельность (динамическая устойчивость) кооперативного соглашения, стратегическая устойчивость и защита от иррационального поведения такого соглашения. В работе получены математические результаты, основанные на использовании процедуры распределения дележа (ПРД), которые развивают разработанные ранее аспекты динамически устойчивой кооперации. В работе доказано для специального класса дифференциальных игр, что динамически устойчивое кооперативное соглашение может быть стратегически поддержано равновесием по Нэшу. Также приведен пример, в котором выполняются все три условия.

Ключевые слова: дифференциальная игра, кооперативное решение, временная состоятельность кооперативных соглашений, процедура распределения выигрыша (ПРВ), процедура распределения дележа (ПРД), стратегическая устойчивость, защита от иррационального поведения.

1. Введение

Кооперация представляет собой одну из основных форм человеческого поведения. Поэтому по многим практическим причинам важно, чтобы такая кооперация была устойчивой на всем временном промежутке ее реализации. Мы выделим три на наш взгляд основных условия такой устойчивости при рассмотрении проблемы устойчивости долгосрочных кооперативных соглашений.

1. *Состоятельность во времени (динамическая устойчивость) кооперативных соглашений.* Временная состоятельность представляет собой свойство кооперативного соглашения, когда следуя кооперативной траектории участники соглашения придерживаются одного и того же принципа оптимальности в каждый текущий момент времени, а поэтому не имеют объективных мотивов отклоняться от ранее выбранного решения о кооперации.
2. *Стратегическая устойчивость.* Предположим, что никакое индивидуальное отклонение от кооперации каждого участника не приносит выгоды отклонившемуся участнику. Это означает, что исход такого кооперативного соглашения достигается при некотором равновесии по Нэшу, которое и будет гарантировать стратегическую поддержку такой кооперации.
3. *Защита от иррационального поведения.* Это свойство кооперации должно рассматриваться, поскольку нет уверенности в том, что все участники кооперации будут вести себя рационально на всем продолжительном промежутке реализации кооперативного соглашения. Участники кооперации должны быть уверены, что даже в случае реализации наихудшего сценария (например, аннулирования кооперативного соглашения) их выигрыш будет не меньше, чем при изначальном некооперативном поведении.

В работе развит математический аппарат, основанный на применении процедуры распределения выигрыша (ПРВ) или процедуры распределения дележа (ПРД) применительно к анонсированным выше аспектам кооперации [1].

2. Случай непрерывного времени

Рассмотрим дифференциальную игру n - лиц $\Gamma(x_0, T - t_0)$ с предписанной продолжительностью и независимыми движениями на временном промежутке $[t_0, T]$. Уравнения движения имеют вид:

$$\begin{aligned} \dot{x}_i &= f_i(x_i, u_i), \quad u_i \in U_i \subset R^\ell, x_i \in R^m, \\ x_i(t_0) &= x_i^0, \quad i = 1, \dots, n. \end{aligned} \quad (2.1)$$

Предполагается, что система дифференциальных уравнений (2.1) удовлетворяет всем условиям существования, единственности и продолжимости решения для любого n - набора измеримых управлений $u_1(t), \dots, u_n(t)$.

Выигрыш игрока i определяется следующим образом:

$$H_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) = \int_{t_0}^T h_i(x(\tau)) d\tau,$$

где $h_i(x)$ представляет собой непрерывную функцию и $x(\tau) = \{x_1(\tau), \dots, x_n(\tau)\}$ решение системы (2.1) при допустимом программном управлении

$$u_1(\tau), \dots, u_n(\tau)$$

и начальных условиях

$$x(t_0) = \{x_1(t_0), \dots, x_n(t_0)\} = \{x_1^0, \dots, x_n^0\} = x_0.$$

Предположим, что существует n - набор программных управлений $\bar{u}(\tau) = (\bar{u}_1(\tau), \dots, \bar{u}_n(\tau))$ и траектория $\bar{x}(\tau)$, $\tau \in [t_0, T]$, такие что

$$\begin{aligned} & \max_{u_1(\tau), \dots, u_n(\tau)} \sum_{i=1}^n H_i(x_0, T - t_0; u_1(\tau), \dots, u_n(\tau)) = \\ & = \sum_{i=1}^n H_i(x_0, T - t_0; \bar{u}_1(\tau), \dots, \bar{u}_n(\tau)) = \sum_{i=1}^n \int_{t_0}^T h_i(\bar{x}(\tau)) d\tau. \end{aligned} \quad (2.2)$$

Траекторию $\bar{x}(\tau) = (\bar{x}_1(\tau), \dots, \bar{x}_n(\tau))$, удовлетворяющую (2.2), будем называть "оптимальной кооперативной траекторией".

Обозначим через $N = \{1, \dots, n\}$ - множество игроков и определим в игре $\Gamma(x_0, T - t_0)$ - классическим образом характеристическую функцию:

$$\begin{aligned} V(x_0, T - t_0; N) &= \sum_{i=1}^n \int_{t_0}^T h_i(\bar{x}(\tau)) d\tau, \\ V(x_0, T - t_0; \emptyset) &= 0, \\ V(x_0, T - t_0; S) &= Val \Gamma_{S, N \setminus S}(x_0, T - t_0), \end{aligned} \quad (2.3)$$

где $Val \Gamma_{S, N \setminus S}(x_0, T - t_0)$ обозначает значение антагонистической игры между коалицией S , действующей как игрок 1, и коалицией $N \setminus S$, действующей как игрок 2, при этом выигрыш игрока S равен:

$$\sum_{i \in S} H_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)).$$

Определим также $L(x_0, T - t_0)$ как множество дележей в игре $\Gamma(x_0, T - t_0)$ (см. [4]):

$$\begin{aligned} L(x_0, T - t_0) &= \{ \alpha = (\alpha_1, \dots, \alpha_n) : \\ &\alpha_i \geq V(x_0, T - t_0; \{i\}), \sum_{i \in N} \alpha_i = V(x_0, T - t_0; N) \} \end{aligned} \quad (2.4)$$

Регуляризованная игра $\Gamma_\alpha(x_0, T - t_0)$. Для каждого дележа $\alpha \in L(x_0, T - t_0)$ определим некооперативную игру $\Gamma_\alpha(x_0, T - t_0)$, которая отличается от игры $\Gamma(x_0, T - t_0)$ только выигрышами вдоль оптимальной кооперативной траектории $\bar{x}(\tau)$, $\tau \in [t_0, T]$.

Пусть $\alpha \in L(x_0, T - t_0)$. Определим процедуру распределения дележа (ПРД) ([5]) как функцию $\beta(\tau) = (\beta_1(\tau), \dots, \beta_n(\tau))$, $\tau \in [t_0, T]$, такую что

$$\alpha_i = \int_{t_0}^T \beta_i(\tau) d\tau. \quad (2.5)$$

Определим через $H_i^\alpha(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot))$ функцию выигрыша в игре $\Gamma_\alpha(x_0, T - t_0)$ и через $x(\tau)$ соответствующую траекторию. Тогда

$$H_i^\alpha(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) = H_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot))$$

если не существует такого $t \in (t_0, T]$, что $x(\tau) = \bar{x}(\tau)$ для $\tau \in (t_0, t]$. Пусть $t = \sup\{t' : x(\tau) = \bar{x}(\tau), \tau \in [t_0, t']\}$ и $t > t_0$. Тогда

$$\begin{aligned} H_i^\alpha(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) &= \\ &= \int_{t_0}^t \beta_i(\tau) d\tau + H_i(\bar{x}(t), T - t; u_1(\cdot), \dots, u_n(\cdot)) = \\ &= \int_{t_0}^t \beta_i(\tau) d\tau + \int_t^T h_i(x(\tau)) d\tau. \end{aligned}$$

В частном случае, когда $x(\tau) = \bar{x}(\tau)$, $\tau \in [t_0, T]$ (если $x(\tau)$ представляет собой кооперативную траекторию), имеем

$$H_i^\alpha(x_0, T - t_0; \bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot)) = \int_{t_0}^T \beta_i(\tau) d\tau = \alpha_i.$$

По определению функции выигрыша в игре $\Gamma_\alpha(x_0, T - t_0)$ получим, что вдоль оптимальной кооперативной траектории эти выигрыши равны компонентам дележа $\alpha = (\alpha_1, \dots, \alpha_n)$.

Рассмотрим текущие подыгры ([4]) $\Gamma(\bar{x}(t), T - t)$ вдоль $\bar{x}(t)$ и текущие множества дележей $L(\bar{x}(t), T - t)$. Пусть $\alpha(t) \in L(\bar{x}(t), T - t)$. Предположим, что $\alpha(t)$ может быть выбран как дифференцируемая по t , $t \in [t_0, T]$ функция.

Определение 2.1. *Игра $\Gamma_\alpha(x_0, T - t_0)$ называется регуляризацией игры $\Gamma(x_0, T - t_0)$ (α - регуляризация), если ПРД β определяется таким образом, что*

$$\alpha_i(t) = \int_t^T \beta_i(\tau) d\tau$$

или

$$\beta_i(t) = -\alpha_i'(t). \quad (2.6)$$

Из (2.6) получаем

$$\alpha_i = \int_{t_0}^t \beta_i(\tau) d\tau + \alpha_i(t), \quad (2.7)$$

где $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in L(x_0, T - t_0)$, и $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)) \in L(\bar{x}(t), T - t)$.

Пусть $M(x_0, T - t_0) \subset L(x_0, T - t_0)$ представляет собой некоторый принцип оптимальности для кооперативной версии игры $\Gamma(x_0, T - t_0)$, а $M(\bar{x}(t), T - t) \subset L(\bar{x}(t), T - t)$ - тот же принцип оптимальности, но определенный для подыгр $\Gamma(\bar{x}(t), T - t)$ с начальными условиями на кооперативной траектории. В качестве M может быть выбрано c -ядро, HM -решение, вектор Шепли, ядро и другие принципы оптимальности, используемые в кооперативной теории игр. Если $\alpha \in M(x_0, T - t_0)$ и $\alpha(t) \in M(\bar{x}(t), T - t)$, то условие (2.7) дает нам временную состоятельность выбранного дележа α , или выбранного принципа оптимальности, поскольку в этом случае условие (2.7) означает, что ожидаемый к получению выигрыш в текущей подыгре $\alpha(t)$ при всех t принадлежит одному и тому же принципу оптимальности $M(\bar{x}(t), T - t)$. В таком случае будем говорить, что имеет место *временная состоятельность (динамическая устойчивость) выбранного кооперативного соглашения*.

Рассмотрим теперь проблему *стратегической устойчивости* кооперативных соглашений. Основываясь на процедуре распределения дележа β , удовлетворяющей (2.5), можно доказать следующую основную теорему.

Теорема 2.1 *В регуляризованной игре $\Gamma_\alpha(x_0, T - t_0)$ для каждого $\varepsilon > 0$ существует ε -равновесие по Нэшу (Nash (1951)) с выигрышами $\alpha = (\alpha_1, \dots, \alpha_i, \dots, \alpha_n)$.*

Доказательство основано на конструктивном построении ε -равновесия по Нэшу в кусочно-программных стратегиях (КПС) с памятью.

Напомним определение КПС стратегий с памятью в дифференциальной игре. Обозначим через $\hat{x}(t)$ произвольную допустимую траекторию системы (2.1) на временном промежутке $[t_0, t]$, $t \in [t_0, T]$. Стратегия $u_i(\cdot)$ игрока i называется КПС, если она определяется парой (a, σ) , где σ представляет собой разбиение промежутка $[t_0, T]$, $t_0 < t_1 < \dots < t_l = T$, $(t_{k+1} - t_k = \delta > 0)$ и a отображение, кото-

рое каждой точке разбиения $(\hat{x}(t_k), t_k)$, $t_k \in \sigma$, ставит в соответствие программное управление $u_i(t)$, $t \in [t_k, t_{k+1})$.

Рассмотрим семейство ассоциированных с $\Gamma(x, T - t)$ (но не с $\Gamma_\alpha(x, T - t)$) антагонистических игр $\Gamma_{\{i\}, N \setminus \{i\}}(x, T - t)$ из начального состояния x продолжительности $T - t$ между коалицией S , состоящей из одного игрока i и дополнительной коалицией $N \setminus \{i\}$ с выигрышем игрока i равным

$$H_i(x, T - t; u_1(\cdot) \dots, u_n(\cdot)).$$

Выигрыш игрока $N \setminus \{i\}$ в игре $\Gamma_{\{i\}, N \setminus \{i\}}(x, T - t)$ равен $(-H_i)$. Пусть $\hat{u}(x, t; \cdot)$ есть ε -оптимальная КПС стратегия игрока $N \setminus \{i\}$ в игре $\Gamma_{\{i\}, N \setminus \{i\}}(x, T - t)$. Заметим, что $\hat{u}(x, t; \cdot) = \{\hat{u}_j(x, t; \cdot)\}$, $j \in N \setminus \{i\}$.

Пусть $\hat{x}(\tau) = \{\hat{x}_1(\tau), \dots, \hat{x}_n(\tau)\}$ - отрезок допустимой траектории (2.1), определенной на временном промежутке $[t_0, t]$, $t \in [t_0, T]$. Для каждого $i \in \{1, \dots, n\}$ определим $\bar{t}(i) = \sup\{t_i : \hat{x}_i(t_i) = \bar{x}_i(t_i)\}$ и $\bar{t}(j) = \min_i \{\bar{t}(i) = \bar{t}(j)\}$. $\bar{t}(j)$ принадлежит одному из промежутков $[t_k, t_{k+1})$, $k = 0, 1, \dots, l - 1$. Таким образом, $\bar{t}(i) - t_0$ представляет собой длину временного промежутка, начинающегося в t_0 , на котором $\hat{x}_i(t)$ совпадает с $\bar{x}_i(t)$ - i -ая компонента кооперативной траектории $\bar{x}(t)$. В свою очередь, $\bar{t}(j) - t_0$ представляет собой длину временного промежутка, начинающегося в t_0 , на котором $\hat{x}(t)$ совпадает с кооперативной траекторией $\bar{x}(t)$.

Определим следующие стратегии игрока $i \in N$:

$$u_i^*(\cdot) = \begin{cases} \bar{u}_i(t) & \text{для } (\hat{x}(t_k), t_k) \text{ на оптимальной} \\ & \text{кооперативной траектории} \\ & \bar{x}(t) \text{ } (\hat{x}(\tau) = \bar{x}(\tau), \tau \in [t_0, t_k]); \\ \hat{u}_i(\hat{x}(t_{k+1}), t_{k+1}; \cdot) & i\text{-ая компонента } \varepsilon/2 \text{ оптимальной КПС} \\ & \text{стратегии игрока } N \setminus \{j\} \text{ в игре} \\ & \Gamma_{\{j\}, N \setminus \{j\}}(x(t_{k+1}), T - t_{k+1}), \\ & \text{если } t_k \leq \bar{t}(j) < t_{k+1}; \\ \text{произвольно} & \end{cases}$$

Покажем, что ситуация $u^*(\cdot) = (u_1^*(\cdot), \dots, u_n^*(\cdot))$ и есть ε -равновесие по Нэшу в игре $\Gamma_\alpha(x_0, T - t_0)$.

Действительно, имеют место следующие равенства

$$H_i(x_0, T - t_0; u^*(\cdot)) = H_i(x_0, T - t_0; u_1^*(\cdot), \dots, u_n^*(\cdot)) = \int_{t_0}^T \beta_i(t) dt = \alpha_i. \quad (2.8)$$

Рассмотрим набор стратегий $(u^*(\cdot) || u_i(\cdot))$, где игрок i изменяет свою стратегию с $u_i^*(\cdot)$ на $u_i(\cdot)$. Нужно показать, что

$$H_i(x_0, T - t_0; u^*(\cdot)) \geq H_i(x_0, T - t_0; u^*(\cdot) || u_i(\cdot)) - \varepsilon. \quad (2.9)$$

для всех $i \in N$ и любой КПС игрока i .

Легко видеть, что когда разыгрывается ситуация $u^*(\cdot)$, игра развивается вдоль оптимальной траектории $\bar{x}(t)$. Если при $(u^*(\cdot) || u_i(\cdot))$ также реализуется траектория $\bar{x}(t)$, то (2.9) выполняется как равенство и поэтому утверждение верно. Предположим теперь, что при $(u^*(\cdot) || u_i(\cdot))$ реализовавшаяся траектория $x(t)$ отличается от $\bar{x}(t)$. Тогда пусть

$$\bar{t} = \inf\{t : \bar{x}(t) \neq x(t)\}.$$

и $\bar{t} \in [t_{k-1}, t_k)$. Поскольку движения игроков независимы, то сразу можем сказать, что $\bar{x}_m(t_k) = x_m(t_k)$ для $m \in N \setminus \{i\}$ и $\bar{x}_i(t_k) \neq x_i(t_k)$ (но $\bar{x}_j(t_{k-1}) = x_j(t_{k-1})$ для $j \in N$).

Из определения стратегии $u^*(\cdot)$ следует, что игроки $m \in N \setminus \{i\}$ будут использовать свои стратегии, которые являются $\varepsilon/2$ -оптимальными в антагонистической игре $\Gamma_{\{i\}, N \setminus \{i\}}(x(t_k), T - t_k)$ против игрока i , который отклонился от оптимальной траектории на временном промежутке $[t_{k-1}, t_k)$.

Если игроки из множества $N \setminus \{i\}$ будут использовать свои стратегии $\hat{u}_m(\hat{x}(t_k), t_k; \cdot)$, то игрок i , начиная из состояния $x(t_k), t_k$, получит не больше, чем

$$V(x(t_k), T - t_k; \{i\}) + \frac{\varepsilon}{2},$$

где $V(x(t_k), T - t_k; \{i\})$ представляет собой значение игры $\Gamma_{\{i\}, N \setminus \{i\}}(x(t_k), T - t_k)$. Тогда общий выигрыш игрока i в игре $\Gamma_\alpha(x_0, T - t_0)$ при реализации ситуации $(u^*(\cdot) || u_i(\cdot))$ не превышает величину

$$\int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) + \frac{\varepsilon}{2} + \int_{t_{k-1}}^{t_k} h_i(x(\tau)) d\tau. \quad (2.10)$$

Но выигрыш игрока i в ситуации $u^*(\cdot)$ равен

$$\alpha_i = \int_{t_0}^T \beta_i(\tau) d\tau = \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + \int_{t_{k-1}}^T \beta_i(\tau) d\tau = \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + \alpha_i(t_{k-1}). \quad (2.11)$$

По определению ПРД (см. (2.5), (2.6)), $\alpha_i(t_{k-1}) \in L(\bar{x}(t_{k-1}), T - t_{k-1})$,

$$\int_{t_{k-1}}^T \beta_i(\tau) d\tau = \alpha_i(t_{k-1}) \geq V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}). \quad (2.12)$$

Из непрерывности функции V и непрерывности траектории $x(t)$ при соответствующем выборе $\delta > 0$ ($t_{k+1} - t_k = \delta$) справедливо выполнение следующих неравенств:

$$|V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}) - V(x(t_k), T - t_k; \{i\})| < \frac{\varepsilon}{4},$$

$$\int_{t_{k-1}}^T \beta_i(\tau) d\tau = \alpha_i(t_{k-1}) \geq V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4}.$$

Сравним $\alpha_i(t_{k-1})$ и $V(x(t_k), T - t_k; \{i\}) + \frac{\varepsilon}{2} + \int_{t_{k-1}}^{t_k} h_i(x_i(\tau)) d\tau$. Выбирая величину $\delta = t_{k+1} - t_k$ достаточно малой, можно показать, что интеграл $\int_{t_{k-1}}^{t_k} h_i(x_i(\tau)) d\tau$ будет также мал (меньше, чем $\varepsilon/4$).

Добавляя в обе части (2.12) величину $\int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau$ и используя предыдущее неравенство получаем, что

$$\begin{aligned} \alpha_i &= \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + \alpha_i(t_{k-1}) \geq \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(\bar{x}(t_{k-1}), T - t_{k-1}; \{i\}) \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4} \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) - \frac{\varepsilon}{4} + \int_{t_{k-1}}^{t_k} h_i(\tau) d\tau - \frac{\varepsilon}{4} \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau) d\tau - \frac{\varepsilon}{2} \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau) d\tau + \\ &\quad + \frac{\varepsilon}{2} - \frac{\varepsilon}{2} - \frac{\varepsilon}{2}. \end{aligned} \quad (2.13)$$

Здесь первые четыре слагаемых в правой части неравенства составляют верхнюю границу выигрыша игрока i в ситуации $(u^*(\cdot)||u_i^*(\cdot))$.

Однако α_i представляет собой выигрыш игрока i в ситуации $u^*(\cdot)$, откуда

$$\begin{aligned} H_i(x_0, T - t_0; u^*(\cdot)) &= \alpha_i \geq \\ &\geq \int_{t_0}^{t_{k-1}} \beta_i(\tau) d\tau + V(x(t_k), T - t_k; \{i\}) + \int_{t_{k-1}}^{t_k} h_i(\tau) d\tau + \frac{\varepsilon}{2} - \varepsilon \geq \\ &\geq H_i(x_0, T - t_0; u^*(\cdot)||u_i(\cdot)) - \varepsilon \end{aligned} \quad (2.14)$$

И тем самым мы получаем (2.9). Теорема доказана. \square

Содержательно, утверждение теоремы означает, что кооперативное решение (некоторый дележ) может быть стратегически поддержано в регуляризованной игре $\Gamma_\alpha(x_0, T - t_0)$ (реализуемо в специально сконструированном равновесии по Нэшу) равновесием по Нэшу $u^*(\cdot)$, определенным в теореме 2.1

Условия защиты от иррационального поведения. Предположим теперь, что в некоторый момент времени иррациональное поведение игрока (или группы игроков) приводит к распаду большой коалиции, а тем самым и прекращению действия кооперативного соглашения. В этом случае *условие защиты от иррационального поведения* (см.[7]) требует, чтобы выполнялось следующее неравенство:

$$V(x_0, T - t_0; \{i\}) \leq \int_{t_0}^t \beta_i(\tau) d\tau + V(\bar{x}(t), T - t; \{i\}), \quad i \in N. \quad (2.15)$$

Если ПРД $\beta(t)$ может быть выбран так, что удовлетворяются и условие временной состоятельности и условие защиты от иррационального поведения (условие стратегической устойчивости следует из временной состоятельности в соответствии с теоремой 2.1), тогда кооперативное соглашение о выборе дележа $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ будем называть *устойчивым кооперативным соглашением*.

Если предположить дифференцируемость функции $V(x, T; \{i\})$, то для выполнения (2.15) достаточно, чтобы ПРД $\beta(\tau) = (\beta_1(\tau), \beta_2(\tau), \dots, \beta_n(\tau))$ удовлетворяло условию:

$$\beta_i(\tau) \geq -\frac{d}{d\tau} V(\bar{x}(\tau), T - \tau; \{i\}), \quad i = 1, \dots, n. \quad (2.16)$$

Другими словами, для устойчивого соглашения все три условия устойчивости должны выполняться. Не всегда это именно так. Однако следующий пример показывает, что в некоторых случаях в задаче с дискретным временем все три условия выполняются.

В (2.16) величина $V(\bar{x}(\tau), T - \tau; \{i\})$ представляет собой значение антагонистической игры между коалицией $N \setminus \{i\}$, действующей как один игрок и игроком $\{i\}$ с коалиционным выигрышем равным $[-H_i(\bar{x}(\tau), T - \tau; u_1, \dots, u_n)]$.

Предположим, что $y(t), t \in [\tau, T]$ есть траектория этой антагонистической игры при реализации оптимальных стратегий (стратегий из седловой точки). Предположим, что для любых начальных условий $\bar{x}(\tau), T - \tau, \tau \in [t_0, T]$ такая седловая точка существует (если нет, то можно рассмотреть ε -седловую точку в кусочно-программных стратегиях, которая всегда существует для каждого заданного $\varepsilon \geq 0$, но следующие формулы будут верны с точностью до ε).

Тогда можно написать

$$V(\bar{x}(\tau), T - \tau; \{i\}) = \int_{\tau}^T h_i(\bar{x}(\tau); y(t)) dt,$$

где $y(\tau) = \bar{x}(\tau)$. Из (2.16) имеем

$$\begin{aligned} \beta_i(\tau) &\geq -\frac{d}{d\tau} \int_{\tau}^T h_i(\bar{x}(\tau); y(t)) dt = \\ &= -[-h_i(\bar{x}(\tau); y(\tau)) + \int_{\tau}^T \sum_{l=1}^n \sum_{k=1}^m \frac{\partial h_i(\bar{x}(\tau); y(t))}{\partial x_{lk}} f_{lk}(\bar{x}(\tau), \bar{u}(\tau)) dt] = \\ &= h_i(\bar{x}(\tau); \bar{x}(\tau)) - \int_{\tau}^T \sum_{l=1}^n \sum_{k=1}^m \frac{\partial h_i(\bar{x}(\tau); y(t))}{\partial x_{lk}} f_{lk}(\bar{x}(\tau), \bar{u}(\tau)) dt \end{aligned}$$

или

$$\beta_i(\tau) \geq h_i(\bar{x}(\tau); \bar{x}(\tau)) - \int_{\tau}^T \sum_{l=1}^n \sum_{k=1}^m \frac{\partial h_i(\bar{x}(\tau); y(t))}{\partial x_{lk}} f_{lk}(\bar{x}(\tau), \bar{u}(\tau)) dt.$$

3. Случай дискретного времени

В качестве основной модели для случая дискретного времени рассмотрим игру в развернутой форме с полной (совершенной) информацией.

Определение 3.1 Дерево игры представляет собой конечный древовидный граф K с корневой вершиной (корнем) x_0 .

В дальнейшем будем использовать следующие обозначения. Пусть x некоторая вершина (позиция). Обозначим через $K(x)$ поддерево K с корнем x . Обозначим через $Z(x)$ - множество вершин, непосредственно следующих за x . Вершины y , непосредственно следующие за x , называются альтернативами в вершине x ($y \in Z(x)$).

Игрока, который делает ход в x (который выбирает следующую альтернативу в позиции x), будем обозначать через $i(x)$. Выбор игрока $i(x)$ в позиции x будем обозначать через $\bar{x} \in Z(x)$.

Пусть $N = \{1, \dots, n\}$ - множество всех игроков в игре.

Определение 3.2. Игра в развернутой форме с полной информацией (см. [2]) $G(x_0)$ представляет собой древовидный граф $K(x_0)$, который обладает следующими дополнительными свойствами:

- Множество вершин (позиций) разлагается на $n + 1$ подмножество

$$P_1, P_2, \dots, P_{n+1},$$

которые образуют разбиение множества всех вершин графа K . Вершины (позиции) $x \in P_i$ называются личными позициями игрока i , $i = 1, \dots, n$; вершины (позиции) $x \in P_{n+1}$ называются терминальными позициями.

- В каждой вершине (позиции) x задан набор действительных чисел $h(x) = (h_1(x), \dots, h_n(x))$, где $h_i(x)$ интерпретируется как выигрыш игрока i в вершине (позиции) x .

Определение 3.3. Стратегия игрока i представляет собой отображение $U_i(\cdot)$, которое каждой $x \in P_i$ ставит в соответствие единственную альтернативу $y \in Z(x)$.

Как и в предыдущем случае обозначим через $H_i(x; u_1(\cdot), \dots, u_n(\cdot))$ функцию выигрыша игрока $i \in N$ в подыгре $G(x)$ из позиции x . Будем предполагать, что

$$H_i(x; u_1(\cdot), \dots, u_n(\cdot)) = \sum_{i=1}^l h_i(x'_i)$$

где $x' = (x'_1, x'_2, \dots, x'_l)$ представляет собой путь, реализовавшийся в подыгре $G(x)$ в ситуации $(u_1(\cdot), \dots, u_n(\cdot))$, $x'_1 = x$.

Обозначим через $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot))$ ситуацию и соответствующую траекторию (путь) $\bar{x} = (\bar{x}_0, \bar{x}_1, \dots, \bar{x}_m)$, $\bar{x}_m \in P_{n+1}$, такие что

$$\begin{aligned} \max_{u_1(\cdot), \dots, u_n(\cdot)} \sum_{i=1}^n H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)) &= \\ &= \sum_{i=1}^n H_i(x_0; \bar{u}_1(\cdot), \dots, \bar{u}_n(\cdot)) = \sum_{i=1}^n h_i(\bar{x}_m). \end{aligned} \quad (3.1)$$

Путь $\bar{x} = (\bar{x}_0, \dots, \bar{x}_m)$, удовлетворяющий уравнению (3.1), будем называть "кооперативной траекторией" или "кооперативным путем".

Определим в $G(x_0)$ характеристическую функцию классическим образом:

$$\begin{aligned} V(x_0; N) &= \sum_{i=1}^n h_i(\bar{x}_m), \\ V(x_0; \emptyset) &= 0, \\ V(x_0; S) &= Val \Gamma_{S, N \setminus S}(x_0), \end{aligned}$$

где $Val \Gamma_{S, N \setminus S}(x_0)$ есть значение антагонистической игры между коалицией S , действующей как первый игрок, и дополнительной коалицией $N \setminus S$, действующей как игрок 2, с выигрышем игрока S равным

$$\sum_{i \in S} H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)).$$

Определим $L(x_0)$ как множество дележей в игре $G(x_0)$.

$$L(x_0) = \left\{ \alpha = (\alpha_1, \dots, \alpha_n) : \alpha_i \geq V(x_0; \{i\}), \sum_{i \in N} \alpha_i = V(x_0; N) \right\}.$$

Регуляризованная игра $G_\alpha(x_0)$. Для каждого $\alpha \in L(x_0)$ определим некооперативную игру $G_\alpha(x_0)$, которая отличается от игры $G(x_0)$ только выигрышами, определенными в вершинах (позициях) вдоль оптимального кооперативного пути $\bar{x} = (\bar{x}_0, \dots, \bar{x}_m)$. Пусть $\alpha \in L(x_0)$. Определим процедуру распределения дележа (ПРД) как функцию $\beta_k = (\beta_1(k), \dots, \beta_n(k))$, $k = 0, 1, \dots, m$, такую что

$$\alpha_i = \sum_{k=0}^m \beta_i(k). \quad (3.2)$$

Определим через $H_i^\alpha(x_0; u_1(\cdot), \dots, u_n(\cdot))$ функцию выигрыша в игре $G_\alpha(x_0)$ и через $\bar{x} = \{\bar{x}_0, \dots, \bar{x}_m\}$ кооперативный путь.

Положим, что при развитии игры вдоль кооперативной траектории

$$H_i^\alpha(x_0; u_1(\cdot), \dots, u_n(\cdot)) = H_i(x_0; u_1(\cdot), \dots, u_n(\cdot))$$

И пусть $\bar{x}_k = x_k$, при $0 \leq k \leq l$. Тогда

$$H_i^\alpha(x_0; u_1(\cdot), \dots, u_n(\cdot)) = H_i(x_0; u_1(\cdot), \dots, u_n(\cdot)) = \sum_{k=0}^l \beta_i(k) + \sum_{k=l+1}^m h_i(x_k)$$

Таким образом, по определению функции выигрыша в игре $G_\alpha(x_0)$ выигрыши вдоль оптимальной кооперативной траектории равны компонентам дележа $\alpha = (\alpha_1, \dots, \alpha_n)$.

Рассмотрим текущие подыгры $G(\bar{x}_k)$ вдоль оптимального пути \bar{x} и текущие множества дележей $L(\bar{x}_k)$. Пусть $\alpha^k \in L(\bar{x}_k)$.

Определение 3.4. Игра $G_\alpha(x_0)$ называется регуляризацией игры $G(x_0)$ (α -регуляризация), если ПРД β определена таким образом, что

$$\alpha_i^k = \sum_{j=k}^m \beta_i(j)$$

или $\beta_i(k) = \alpha_i^k - \alpha_i^{k+1}$, $i \in N$, $k = 0, 1, \dots, m-1$, $\beta_i(m) = \alpha_i^m$, $\alpha_i^0 = \alpha_i$.

Теорема 3.1. В регуляризации $G_\alpha(x_0)$ исходной игры существует равновесие по Нэшу с равновесными выигрышами $\alpha = (\alpha_1, \dots, \alpha_n)$.

Доказательство. Вдоль кооперативного пути имеем

$$\alpha_i^k \geq V(\bar{x}_k; \{i\}), \quad i \in N, k = 0, 1, \dots, m.$$

поскольку $\alpha^k = (\alpha_1^k, \dots, \alpha_n^k) \in L(\bar{x}_k)$ представляет собой дележ в $G(\bar{x}_k)$ (заметим, что здесь $V(\bar{x}_k; \{i\})$ вычисляется в подыгре $G(\bar{x}_k)$, а не в $G_\alpha(\bar{x}_k)$). В тоже время

$$\alpha_i^k = \sum_{j=k}^m \beta_i(j)$$

и мы получаем

$$\sum_{j=k}^m \beta_i(j) \geq V(\bar{x}_k; \{i\}), \quad i \in N, k = 0, 1, \dots, m. \quad (3.3)$$

Однако $\sum_{j=k}^m \beta_i(j)$ есть выигрыш игрока i в подыгре $G_\alpha(\bar{x}_k)$ вдоль кооперативного пути. Тогда из (3.3), используя аргументы аналогичные таковым в теореме 2.1, можно построить равновесие по Нэшу с выигрышами $\alpha = (\alpha_1, \dots, \alpha_n)$ и результирующим кооперативным путем $\bar{x} = (\bar{x}_0, \dots, \bar{x}_m)$. Условие защиты от иррационального поведения в случае дискретного времени примет вид

$$\sum_{j=0}^l \beta_i(j) + V_i(\bar{x}_{l+1}; \{i\}) \geq V_i(x_0; \{i\}), \quad 0 \leq l \leq m, \quad i \in N. \quad (3.4)$$

Пример.

В этом примере в качестве дележа рассмотрим вектор Шепли (Shapley value) [6]. Используя приведенную выше регуляризацию игры, покажем, что здесь существует равновесие по Нэшу с равновесными выигрышами равными компонентам вектора Шепли.

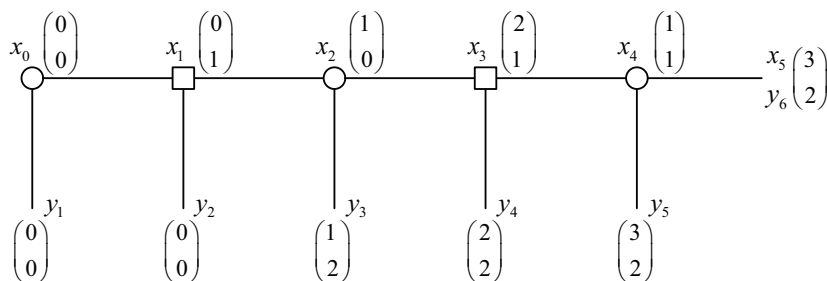


Рисунок 1. Игра $G(x_0)$

В игре $G(x_0) : N = \{1, 2\}$, $P_1 = \{x_0, x_2, x_4\}$, $P_2 = \{x_1, x_3\}$, $P_3 = \{y_1, y_2, y_3, y_4, y_5, y_6\}$. $h(x_0) = (1, 0)$, $h(x_1) = (0, 1)$, $h(x_2) = (1, 0)$, $h(x_3) = (2, 1)$, $h(x_4) = (1, 1)$, $h(x_5) = (4, 4)$, $h(y_1) = (0, 0)$, $h(y_2) = (0, 0)$, $h(y_3) = (1, 2)$, $h(y_4) = (2, 2)$, $h(y_5) = (3, 2)$, $h(y_6) = h(x_5) = (4, 4)$. Оптимальная кооперативная траектория (кооперативный путь) $\bar{x} = \{x_0, x_1, x_2, x_3, x_4, x_5\}$.

	x_0	x_1	x_2	x_3	x_4	x_5
$V(x; \{1\})$	0	0	2	4	5	4
$V(x; \{2\})$	0	3	5	4	5	4
$V(x; \{1, 2\})$	15	15	14	13	10	8
$Sh(x; \{1\})$	7,5	6	5,5	6,5	6	4
$Sh(x; \{2\})$	7,5	9	8,5	6,5	4	4
$\beta_1(x) = \beta_1(j)$	1,5	0,5	-1	0,5	2	4
$\beta_2(x) = \beta_2(j)$	-1,5	0,5	2	2,5	0	4

Легко увидеть, что в этом случае имеет место неравенство (3.3)

$$\sum_{j=k}^m \beta_i(j) \geq V(\bar{x}_k; \{i\})$$

для $i \in N$, и свойство защиты от иррационального поведения (3.4) также выполняется:

$$\sum_{j=0}^l \beta_i(j) + V_i(\bar{x}_{l+1}; \{i\}) \geq V(x_0; \{i\}), \quad i = 1, 2, \quad 1 \leq l \leq 4.$$

СПИСОК ЛИТЕРАТУРЫ

1. Петросян Л.А., Данилов Н.А. *Устойчивые решения неантагонистических дифференциальных игр с транзитивными выигрышами*// Вестник ЛГУ. 1979. № 1. С. 46-54.
2. Kuhn H.W. *Extensive games and the problem of imputation*. Contributions to the Theory of Games II (in H.W. Kuhn and A.W. Tucker (eds.). Princeton: Princeton University Press, 1953. P. 193-216.
3. Nash J. *Non-cooperative games* // Ann. Mathematics. 1951. V. 54. P. 286-295.
4. Neumann J., Morgenstern O. *Theory of Games and Economic Behavior*. Princeton, 1947.
5. Petrosjan L.A. *Differential Games of Pursuit*. World Scientific, Singapore, 1993.
6. Shapley L.S. *A Value for n-Person Games*. Contributions to the Theory of Games (in H.W. Kuhn and A.W. Tucker (eds.). Princeton: Princeton University Press, 1953, P. 307-315.
7. Yeung D.W.K. *An irrational-behavior-proofness condition in cooperative differential games*// Int. J. of Game Theory Review. 2007. V. 9. N 1. P. 256-273.

PRINCIPLES OF DYNAMIC STABILITY

Leon Petrosyan, Faculty of Applied Mathematics and Control Processes, St. Petersburg State University, Saint Petersburg, Doctor of Science, professor (spbuoasis7@peterlink.ru).

Nickolay Zenkevich, School of Management, St. Petersburg State University, Saint Petersburg, Cand.Sc. (zenkevich@gsom.pu.ru).

Abstract: There are three important aspects which must be taken into account when the problem of stability of long-range cooperative agreements is investigated: time-consistency of the cooperative agreements, strategic stability and irrational behavior proofness. The mathematical results based on imputation distribution procedures (IDP) are developed to deal with the above mentioned aspects of cooperation. We proved that for a special class of differential games time-consistent cooperative agreement can be strategically supported by Nash equilibrium. We also consider an example where all three conditions are satisfied.

Keywords: differential game, cooperative solution, time-consistency of the cooperative agreements, payoff distribution procedures (PDP), imputation distribution procedures (IDP), strategic stability, irrational behavior proofness.