Повторяющаяся игра как модель для анализа соглашений об охране окружающей среды (The repeated game as a model for analysis of environmental agreements)*

А. А. Васин (A. A. Vasin), А. Г. Дивцова (A. G. Divtsova)

Московский государственный университет им. М.В. Ломоносова, Москва, Россия

vasin@cs.msu.su, nastyakislaeva@gmail.com

Рассматривается повторяющаяся игра, в которой взаимодействие в каждый период происходит в два этапа. Первый этап представляет собой обобщение известной модели “трагедия общин” [3], в которой индивидуальная оптимизация производственной деятельности каждым игроком приводит к “плохому” равновесию с высоким уровнем загрязнения окружающей среды. На втором этапе игроки перераспределяют выигрыши с помощью побочных платежей. В каждый период отдельного участника интересует суммарный выигрыш за несколько предстоящих повторений до его горизонта планирования. Целью работы является поиск условий существования совершенного подыгрового равновесия (СПР) повторяющейся игры, реализующего парето-оптимальный исход, максимизирующий суммарный выигрыш в однократной игре. Подобные математические задачи рассматривались в работе [2]. Данная проблема представляет интерес в контексте изучения международных соглашений об ограничении загрязнений окружающей среды (см. [1, 4]). Существование СПР означает возможность стабильного и эффективного соглашения такого рода.

В однократной игре, моделирующей взаимодействие в данный период времени, стратегией i-го игрока является величина $z_i \in Z_i = [0, z_{\max,i}]$ выбрасываемого им загрязнения, прямо связанная с интенсивностью его производственной деятельности. Функция выигрыша $F_i(z) = v_i(z_i) - H_i\big(\sum_{j\in I} \pi_{ji} z_j\big)$, $z = (z_1, \dots, z_n)$, где $v_i(z_i)$ — полезность, которую получает страна i в результате своей деятельности, а $H_i(\hat{z}_i)$ — ущерб стране i от загрязнения, $\hat{z}_i = \sum_{j\in I} \pi_{ji} z_j$ — суммарное загрязнение, получаемое страной i от всех стран, π_{ji} обозначает переносимую на страну i долю загрязнения, выбрасываемого страной j. Пусть в игре существует

*Работа выполнена при финансовой поддержке РФФИ, проект 16-01-00353/16.

равновесие Нэша z^* в доминирующих стратегиях (например, $\pi_{ii} = 0$ или H_i — линейная функция в практически интересном диапазоне). Ситуация $\bar{z}$ называется оптимальной по Парето, если в ней реализуется максимум суммарной функции выигрыша $\max_{\vec{z}} \sum_{i=1}^{n} F_i(\vec{z})$.

В реальности рассматриваемое взаимодействие многократно повторяется в примерно одинаковых условиях. Формальной моделью служит повторяющаяся игра с полной информацией и скользящими горизонтами планирования, каждый период взаимодействия $t = 1, 2, \ldots$ включает в себя два этапа: на этапе $t1$ происходит выбор объемов загрязнения z_i^t, $i \in I$, которые определяют выигрыши $F_i^t = F_i(z^t)$, $i \in I$, с учетом потребления, затрат на очистку и ущерба от загрязнения. На этапе $t2$ игроки осуществляют побочные платежи y_i^t, $i \in I$, $\sum_i y_i^t = 0$, и определяют итоговые выигрыши $\bar{F}_i^t(z^t, y^t) = F_i(z^t) + y_i^t$, $i \in I$, за период t. Каждый игрок принимает решение на текущем этапе, исходя из полной информации о действиях всех игроков в предыдущее время. При этом игрок стремится максимизировать свой суммарный выигрыш за $T_i + 1$ периодов начиная с текущего. Последовательность действий игроков $h^{t-1} = (z^\tau, y^\tau)_{\tau=1}^{t-1}$ называется историей игры до этапа $t1$. К этапу $t2$ к истории добавляется значение z^t. Стратегия игрока i в каждый период t определяет выбор z_i^t, а затем y_i^t в зависимости от истории. Формально стратегия задается функциями $z_i^t = \sigma_i^1(h^{t-1})$, $y_i^t = \sigma_i^2(h^{t-1}, z^t)$. Ситуация σ является СПР, если для любых h и t имеем $\sigma_i^* = \arg\max_{\sigma_i} \sum_{\tau=t}^{t+T_i} \bar{F}_i(z^\tau(h^{t-1}, \sigma^* \| \sigma_i), y^\tau(h^{t-1}, \sigma^* \| \sigma_i))$, т.е. какова бы ни была предшествующая история до периода $t-1$, для игрока i оптимально следовать стратегии σ_i^*, если остальные игроки придерживаются своих стратегий σ_j^*, $j \in I \setminus i$.

Нас интересует существование СПР, для которого в каждом повторении реализуется парето-оптимальная ситуация $\bar{z}$. Рассмотрим конструкцию такого равновесия, основанную на следующих идеях. При отклонении на некотором этапе какого-то игрока остальные страны в простейшем случае (далее случай (а), см. также [1]) со следующего этапа переходят к реализации равновесия Нэша z^*, при этом побочные платежи, естественно, прекращаются. Во втором случае (случай (б)) после отклонения остальные страны продолжают сотрудничество без отклонившегося игрока. Размеры побочных платежей определяются заново. При отклонении второй страны все игроки переходят к равновесию Нэша.

Для любой вектор-функции $(f_i(z),\ i \in I)$ введем обозначения $f_i^* :=$ $:= f_i(z^*)$, $\bar{f}_i := f_i(\bar{z})$, $f_\Sigma(z) := \sum_{i \in I} f_i(z)$, $f_{\Sigma(I \setminus i)}(z) := \sum_{j \in I \setminus i} f_j(z)$.

Теорема 1. *СПР для случая* (а) *существует, если и только если* $\sum_{i\in I}\big(\big(H_i^* - H_i\big(\sum_{j\in I\setminus i}\pi_{ji}\bar z_j + \pi_{ii}z_i^*\big)\big)/(1+T_i)\big) \le \bar F_\Sigma - F_\Sigma^*$. *При этом условии побочные платежи можно определить как* $y_i = F_i^* - \bar F_i - \big(H_i^* - H_i\big(\sum_{j\in I\setminus i}\pi_{ji}\bar z_j + \pi_{ii}z_i^*\big)\big)/(1+\lambda T_i)$, *где* $\lambda \le 1$ — *корень уравнения* $\sum_{i\in I}\big(\big(H_i^* - H_i\big(\sum_{j\in I\setminus i}\pi_{ji}\bar z_j + \pi_{ii}z_i^*\big)\big)/(1+\lambda T_i)\big) = \bar F_\Sigma - F_\Sigma^*$.

В случае (б) положим $\bar{\bar z}_{I\setminus i}(i) = \arg\max_{(z_j, j\in I\setminus i)} \sum_{j\in I\setminus i} F_j(z_{I\setminus i}, z_i{}^*)$, $\bar{\bar z}(i) = (\bar{\bar z}_{I\setminus i}, z_i^*)$, $\bar{\bar F}_j(i) := F_j(\bar{\bar z}(i))$ при всех $i, j \in I$.

Теорема 2. *СПР для случая* (б) *существует, если и только если выполняются следующие условия*:

1) $\bar{\bar F}_\Sigma \le \bar F_\Sigma$;
2) $\bar{\bar F}_{\Sigma(I\setminus i)} \ge F^*_{\Sigma(I\setminus i)} \quad \forall i \in I$;
3) $\sum_{i\in I}\big(\big(H_i\big(\sum_{j\in I}\pi_{ji}\bar{\bar z}_j(i)\big) - H_i\big(\sum_{j\in I\setminus i}\pi_{ji}\bar z_j + \pi_{ii}z_i^*\big)\big)/(1+T_i)\big) \le \le \bar F_\Sigma - \bar{\bar F}_\Sigma$;
4) $\sum_{j\in I\setminus i}\big(\big(H_j\big(\sum_{k\in I}\pi_{kj}z_k^*\big) - H_j\big(\sum_{k\in I\setminus\{i,j\}}\pi_{kj}\bar{\bar z}_k + \pi_{ij}z_i^* + \pi_{jj}z_j^*\big)\big)/ /(1+T_j)\big) \le \bar{\bar F}_{\Sigma(I\setminus i)}(i) - F^*_{\Sigma(I\setminus i)}$, $i \in I$.

Замечание 1. Случай (б) описывает поведение игроков, более рациональное с точки зрения общего выигрыша, поскольку в нем предпринимается попытка сохранить кооперацию при отклонении одной страны. Однако условия существования СПР в этом случае не всегда выполнены даже при неограниченных горизонтах планирования. В то же время в случае (а) при достаточно длинных горизонтах планирования они выполнены всегда.

Замечание 2. Теоремы 1 и 2 допускают следующие обобщения для модели, в которой горизонты планирования $T_i(t)$ зависят от времени. Для существования указанных СПР достаточно, чтобы условия утверждений выполнялись для $T_i = \min_t T_i(t)$. Однако эти условия не являются необходимыми в общем случае. Другое обобщение связано с повторяющимися играми с переменным дисконтированием, в которых игрок i в период t стремится максимизировать приведенный будущий выигрыш $\sum_{\tau=t}^{\infty} d^i_{\tau t} W_i(\tau)$, где $W_i(\tau)$ — его выигрыш в период τ, $d^i_{\tau t}$ — коэффициент приведения к текущему периоду, $d^i_{tt} = 1$. Указанные утверждения остаются справедливыми как достаточные условия, если положить $T_i(t) = \sum_{\tau=t+1}^{\infty} d^i_{\tau t}$.

Список литературы

1. *Васин А.А., Дивцова А.Г.* Теоретико-игровая модель соглашения об ограничении загрязнения атмосферы // Математическая теория игр и приложения. 2017. Т. 9, № 1. С. 27–44.
2. *Kleimenov A.F., Kryazhimskiy A.V.* Minimum-noncooperative trajectories in repeated games // Complex dynamical systems with incomplete information / Ed. by E. Reithmeier, G. Leitmann. 1999. V. 1. P. 94–107.
3. *Hardin G.* The tragedy of the commons // Science, New Series. 1968. V. 162, N 3859.
4. *Sacco A., Zaccour G.* Impact of social externalities on the formation of an international environmental agreement: An exploratory analysis. 2016.

О снижении зашумленности при моделировании параметров методом динамической регуляризации (On the reduction of noisiness in the modeling of parameters by the method of dynamic regularization)

А. Ю. Вдовин (A. Yu. Vdovin),
С. С. Рублева (S. S. Rubleva)

Уральский государственный лесотехнический университет, Екатеринбург, Россия

vdovin@usfeu.ru, rublevas@mail.ru

Рассматривается обратная задача для динамической системы, которая описывается обыкновенным дифференциальным уравнением

$$\begin{gathered}\dot{x}(t) = f(t, x(t), v(t)) = f_1(t, x(t), v(t)) + f_2(t, x(t), v(t))v(t),\\ t \in [t_0, \vartheta] = T, \qquad x(t_0) = x_0.\end{gathered} \tag{1}$$

Здесь $f_1(\cdot)$, $f_2(\cdot)$ — непрерывные отображения $T \times \mathbb{R}^m$ в пространство $\mathbb{R}^m$ с евклидовой нормой $|\cdot|$ и в пространство $\mathbb{R}^{m \times q}$ матриц со спектральной нормой. Решение $x(\cdot)$ по Каратеодори задачи (1) станем называть движением, а измеримую по Лебегу функцию $v(\cdot)\colon T \to Q \subset \mathbb{R}^q$ (Q — выпуклый компакт) — воздействием, порождающим это движение. Пусть