Автоматика и телемеханика, № 6, 2025

Управление в социально-экономических системах

© 2025 г. **М.И. ГЕРАСЬКИН**, д-р экон. наук (innovation@ssau.ru)
(Самарский национальный исследовательский университет
имени академика С.П. Королева)

ИНФОРМАЦИОННОЕ УПРАВЛЕНИЕ СТРАТЕГИЕЙ ИГРОКА В ИГРЕ ОЛИГОПОЛИИ n ЛИЦ ПРИ РЕФЛЕКСИВНОМ ПОВЕДЕНИИ ИГРОКОВ

Рассматривается игра олигополии n лиц при объемной конкуренции для случая функций спроса и издержек общего вида. Игроки считаются рефлексирующими, т.е. выдвигают предположения о стратегиях других игроков, в результате чего в игре образуются подмножества игроков с различными уровнями лидерства по Штакельбергу, т.е. рассматривается игра с многоуровневым лидерством. Рефлексия игроков формализована в виде предположительных вариаций, т.е. предположений игроков о влиянии их действий на действие контрагента. Исследуется проблема управления стратегией одного игрока (управляемого игрока) со стороны остальных $n-1$ игроков (Центра), в результате чего устанавливается равновесие Нэша, оптимальное по функциям полезности игроков Центра. Предложена модель взаимодействий игроков в виде иерархической игры, для которой найдена зависимость максимума функции полезности Центра от вектора сумм предположительных вариаций (СПВ) всех игроков, позволяющая вычислить значение СПВ управляемого игрока, оптимизирующее функцию полезности Центра. Разработан способ информационного управления действиями игрока, осуществляемого Центром путем управляющего воздействия, побуждающего игрока выбрать оптимальную для Центра функцию реакции.

Ключевые слова: олигополия, предположительная вариация, лидерство по Штакельбергу, иерархическая игра.

DOI: 10.31857/S0005231025060061, **EDN:** IKHLEQ

1. Введение

Игра олигополии – агрегативная игра [1], т.е. игра, в которой выигрыш каждого игрока зависит от суммы (агрегата) действий всех игроков. Решением этой игры является равновесие Курно–Нэша [2,3]. Г. Штакельберг впервые определил [4] на основе предположительной вариации стратегию лидера в игре в отличие от стратегии ведомого.

Предпосылкой появления в агрегативной игре понятия «предположительная вариация» послужило осознание того факта, что игроки при выборе своих оптимальных действий неизбежно будут предполагать оптимальное поведение соперников, т.е. рефлексировать. Следовательно, предположительная вариация представляет собой математическую формализацию ментального процесса рефлексии [5], которая в данном случае трактуется как выполняемая некоторым игроком мыслительная операция вычисления оптимальной реакции (или наилучшего ответа) другого игрока на действие первого. Как правило, рассматривается объемная предположительная вариация, характеризующая предполагаемое игроком ответное изменение действия (объема предложения продукта) контрагента, оптимизирующее функцию полезности последнего при выбранном действии первого. В современных исследованиях предположительная вариация широко используется для анализа лидерства по Штакельбергу в рамках двух направлений: во-первых, увеличение числа рефлексирующих игроков приводит к появлению в игре множества лидеров по Штакельбергу [6]; во-вторых, углубление рефлексии влечет возникновение лидеров более высоких уровней, т.е. формируется многоуровневое лидерство [7]. Второй аспект выражается в иерархии предположений игроков, приводящих к следующей иерархии их ментальных типов: 1) ведомый, не выдвигающий предположений о стратегиях окружения, вследствие чего его предположительная вариация равна нулю; 2) лидер по Штакельбергу (первого уровня), предполагающий, что его окружают ведомые; 3) лидер по Штакельбергу второго уровня (или более высоких уровней), предполагающий, что его окружают лидеры первого уровня (или иных низших уровней). Иерархия ментальных типов определяет ранг рефлексии игрока r как номер в описанной последовательности ментальных типов. Отметим, что описанная иерархия выстраивается только в представлениях игроков (в этом случае говорят об игре с фантомными игроками), а фактически имеет место неиерархическая игра с равноправными игроками, которая изучается в большинстве исследований проблемы олигополии. Как исключение, иерархическая агрегативная игра с управлением со стороны Центра исследовалась в [8] как задача стимулирования при условии, что игроки (университеты) институционально зависят от Центра (государства).

С учетом описанной стратификации лидеров в игре олигополии могут сосуществовать игроки различных ментальных типов, зависящих от характера информированности каждого из них, причем игрок определенного ментального типа будет совершать предсказуемое действие в игре согласно его предположительной вариации. Поэтому появляется возможность целенаправленного изменения действия некоторого игрока посредством формирования для него определенного информационного поля, что приводит к известной концепции информационного управления. Идея информационного управления [9–11] базируется на формировании целенаправленной последовательности мнений в социальной группе в зависимости от мнений так называемых агентов влияния. Формально говоря, перед информационным управлением стоит задача целенаправленного индуцирования заданного органом управления образа мышления одного или нескольких игроков.

В контексте игры олигополии концепция информационного управления конструируется следующим образом. Пусть группа, состоящая из $n-1$ игроков, которую обозначим символом j, стремится добиться выгодного для себя действия не входящего в группу игрока i. Для этого группа совершает действия, из которых игрок i заключает, что группа имеет некоторый ранг рефлексии r. Поэтому для игрока i оптимальным является ранг рефлексии $r+1$, которому соответствуют определенные значения его предположительных вариаций, что, в свою очередь, предопределяет желательное для группы действие этого игрока. Для конкретной реализации такого процесса управления необходимо найти сумму предположительных вариаций (СПВ) игрока i, которые оптимальны с позиций группы, а также определить зависимости равновесных действий всех игроков от параметров их ментального типа.

Далее в статье будут рассмотрены процедура вычисления оптимальной по критерию окружения величины СПВ некоторого игрока, способ оценки соответствующего этой СПВ ментального типа игрока или его функции реакции, а также алгоритм вычисления действий группы, индуцирующий требуемую реакцию игрока.

2. Базовая модель игры олигополии

Теоретико-игровая модель описывает взаимодействия n игроков, представляющих собой фирмы на рынке олигополии. Традиционно считается [6], что эти фирмы предлагают на рынок идентичный продукт с единой для всех убывающей обратной функцией спроса и в случае объемной конкуренции фирмы-игроки выбирают действия в виде объемов предложения. Игроки рациональны, т.е. максимизируют индивидуальные, вогнутые по действиям, функции полезности $\pi_i(Q, Q_i) = P(Q)Q_i - C_i(Q_i)$, информированы о функциях полезности окружения и выбирают действия одновременно, однократно и независимо. Тогда базовая модель выбора действия игрока представляется в следующем виде:

$$\max_{Q_i \geqslant 0} \pi_i(Q, Q_i) = \max_{Q_i \geqslant 0} \left[P(Q)Q_i - C_i(Q_i)\right], \quad i \in N = \{1, \dots, n\}, \tag{1}$$

$$Q = \sum_{i \in N} Q_i, \tag{2}$$

где Q_i, π_i – действие и функция полезности i-го игрока; Q – агрегат действий; N – множество игроков; n – количество игроков; $P(Q)$ – обратная функция спроса, $P'_Q < 0$; $C_i(Q_i)$ – функция издержек i-го игрока, $C'_{Q_i} > 0$.

Равновесие Нэша в игре $\Gamma = \langle N, \{Q_i, i \in N\}, \{\pi_i, i \in N\} \rangle$ при некотором известном векторе предположительных вариаций определяется путем решения системы уравнений реакций следующего типа:

$$\frac{\partial \pi_i(Q_i^*, \rho_{ij})}{\partial Q_i} = 0, \quad i, j \in N, \tag{3}$$

где $\rho_{ij} = Q'_{jQ_i}$ – объемная предположительная вариация i-го игрока, т.е. предполагаемое им изменение выпуска j-го игрока в ответ на единичный прирост выпуска i-го игрока; Q_i^* – равновесное значение.

Оптимальная (в зарубежной литературе consistent – совместимая) предположительная вариация вычисляется из уравнения (3) j-го игрока, т.е. соответствует его наилучшему ответу. При этом система (3) для функции полезности (1) имеет следующий вид:

$$P(Q) + (1 + S_i^r)Q_i P'_Q - C'_{iQ_i} = 0, \quad i \in N, \quad S_i^r = \sum_{j \in N \setminus i} \rho_{ij}^r, \tag{4}$$

где S_i^r – сумма предположительных вариаций i-го игрока на r-м ранге рефлексии.

Рекуррентная формула СПВ на произвольном ранге рефлексии для независящих от действий игроков предположительных вариаций, т.е. при $\rho'_{ijQ_i} = 0$, имеет вид [6]:

$$S_i^r = \left(\frac{1}{\sum\limits_{j \in N \setminus i} \dfrac{1}{u_j - S_j^{r-1} + 1}} - 1 \right)^{-1}, \tag{5}$$

где $u_i = -1 + \dfrac{P'_{Q_i} + (1 + S_i^{r-1})Q_i P''_{QQ_i} - C''_{iQ_iQ_i}}{|P'_Q|}$ – коэффициент нелинейности, выражающий влияние нелинейности функций спроса и издержек на унимодальность функции полезности i-го игрока.

Поскольку из (4) следует, что вектор равновесия Нэша $\mathbf{Q}^* = \{Q_i^*, i \in N\}$ в игре олигополии n лиц зависит от вектора СПВ всех игроков $\mathbf{S^r} = \{S_i^r, i \in N\}$, то существует и обратная зависимость. А именно, при известном векторе действий $\mathbf{Q}^* = \{Q_i^*, i \in N\}$ можно установить вектор $\mathbf{S^r} = \{S_i^r, i \in N\}$, порождающий эти действия игроков. На этой основе рассмотрим инструменты управления (манипулирования) поведением игрока со стороны окружения.

3. Модель оптимального управления поведением игрока

Рассмотрим следующую модификацию базовой модели игры олигополии в виде иерархической игры. Игрок i является объектом управления, а окружение этого игрока (т.е. остальные игроки) выступает субъектом управления или Центром. Поэтому рассматривается иерархическая игра типа «Центр-агент» (рис. 1). Окружение, которое для упрощения обозначим индексом j (т.е. $j = \{N \setminus i\}$), имеет общую цель – побудить игрока i выбирать действие $\overline{Q}_i$, оптимальное по функциям полезности окружения. Поэтому сформулируем

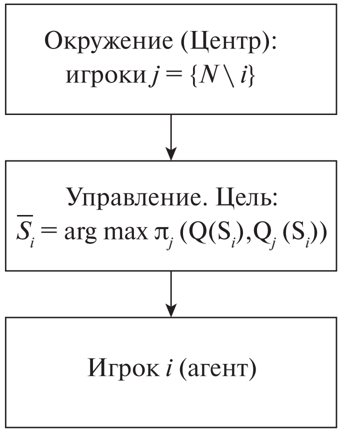

Рис. 1. Схема иерархической игры.

целевую функцию центра как вектор функций полезности игроков окружения, который, поскольку эти функции предполагаются идентичными, можно представить как одну функцию следующим образом:

$$\pi^{(i)} = \pi_j, \quad j \in \{1, \ldots, i-1, i+1, \ldots, n\},$$

и для упрощения записи обозначим целевую функцию центра как

$$\pi = \pi^{(i)}.$$

Опишем основные предположения, принятые при анализе иерархической игры.

1) Предполагается идентичность функций полезности всех игроков окружения, т.е. эти игроки имеют однотипные функции издержек с одинаковыми значениями коэффициентов:

$$C_j(Q_j) = C_k(Q_k), \quad \pi_j(Q_j) = \pi_k(Q_k) \forall j, \quad k = \{N \backslash i\}.$$

2) Информированность игроков описывается следующими множествами информированности I_i, I_j:

– множество информированности управляемого игрока включает в себя множество игроков, действия и функции полезности всех игроков:

$$I_i = \{N, Q_k, \pi_k, k \in N\};$$

– множество информированности окружения охватывает множество игроков, действия и функции полезности всех игроков, СПВ игроков окружения и целевую функцию Центра:

$$I_j = \{N, Q_k, \pi_k, S_j, \pi, j \in N \backslash i, k \in N\},$$

где $\pi = \pi_j$, $j \in N \backslash i$ – целевая функция окружения (Центра), идентичная функциям полезности игроков окружения.

3) Предположительные вариации всех игроков вычисляются ими на основе множеств I_i, I_j, причем могут быть оптимальными (т.е. совместимыми с функциями полезности игроков), или могут быть определены игроками исходя из действий других игроков. Если действия игроков не совместимы с их функциями полезности, то игроки ориентируются на действия. В каждый момент игры либо СПВ управляемого игрока, либо СПВ окружения имеет постоянное значение:

$$S_i = \mathrm{const} \vee S_j = \mathrm{const},$$

поскольку для изменения СПВ игроки оценивают действия других игроков в предыдущий момент.

Управление поведением игрока i окружение осуществляет через его рефлексию в следующем порядке:

– окружение вычисляет оптимальное по своим функциям полезности целевое значение $\overline{S}_i$ СПВ игрока i:

$$\overline{S}_i = \arg\max_{S_i} \pi_j(Q(S_i), Q_j(Si)),$$

– окружение совершает действия $\overline{Q}_j$, которые приведут к выбору i-м игроком СПВ, равной $\overline{S}_i$, вследствие чего управляемый игрок выбирает оптимальное для окружения действие $\overline{Q}_i$.

Второй этап описанного порядка управления поведением игрока объясняет смысл предположения 3 в контексте двойственности подхода игроков к оценке предположительных вариаций. Поскольку окружение для управления совершает действия $\overline{Q}_j$, определенные не из максимума его функции полезности, а из условия побуждения игрока выбрать $\overline{S}_i$, то игрок, предсказав СПВ по функциям полезности окружения, придет к противоречию. Следовательно, игрок при этом противоречии отдаст предпочтение способу оценки СПВ по действиям окружения. Поэтому при оценке СПВ принимаются во внимание две альтернативы: если действия игроков согласуются с их функциями полезности, то другие игроки оценивают оптимальные СПВ; если же СПВ, найденные по действиям игроков, не совпадают с СПВ по функциям полезности, то приоритет отдается первому способу оценки как более реалистичному.

4. Методы расчета оптимального управления

Управление поведением игрока i основано на зависимости функции полезности каждого игрока от СПВ всех игроков, предопределенной системой (4). Поэтому вначале выведем формулу максимума функции полезности окружения в зависимости от вектора СПВ игроков.

У т в е р ж д е н и е 1. Максимальное значение функции полезности окружения вычисляется по формуле

$$(6)\quad \pi_j^*(\mathbf{S}^r) = P\left[Q^*(\mathbf{S}^r)\right] Q_j^*(\mathbf{S}^r) - \int\limits_0^{Q_j^*(\mathbf{S}^r)} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0)$$

где $\mathbf{S}^r = \{S_k^r, k \in N\}$ – вектор СПВ всех игроков.

Д о к а з а т е л ь с т в о у т в е р ж д е н и я 1.

Выразим из уравнения (4), записанного для окружения, предельные издержки

$$C'_{jQ_j} = P(Q^*) + (1+S_j^r)Q_j^* P_Q',$$

проинтегрируем это выражение по Q_j

$$C_j(Q_j^*) = \int\limits_0^{Q_j^*} C'_{jQ_j} dQ_j + C_j(0) = \int\limits_0^{Q_j^*} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0)$$

и подставим в функцию полезности окружения

$$\pi_j^* = P(Q^*)Q_j^* - C_j(Q_j^*) = P(Q^*)Q_j^* - \int\limits_0^{Q_j^*} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0),$$

где $C_j(0)$ – постоянные издержки. Заметим, что в этом выражении равновесное действие игрока Q_j^*, равновесная цена $P(Q^*)$ и равновесное агрегированное действие Q^* являются функцией от вектора СПВ всех игроков $\mathbf{S}^r = \{S_i^r, i \in N\}$. Таким образом, максимальная полезность окружения также зависит от этого вектора, откуда следует (6). ■

Выведем выражение для СПВ игрока i, оптимальной по критерию полезности окружения:

$$\overline{S}_i = \arg\max_{S_i} \pi_j^*(\mathbf{S}^r).$$

У т в е р ж д е н и е 2. Оптимальное по функции полезности окружения значение СПВ $\overline{S}_i$ управляемого игрока вычисляется из уравнения

$$(7)\qquad 2(1+S_j^r)Q_j^* Q_{jS_i}^{*'} P_Q' + \left((1+S_j^r)P_{QS_i}'' + P_Q' \frac{\partial S_j^r}{\partial S_i}\right) Q_j^{*2} = 0$$

при условии

$$(7a)\qquad 2\frac{\partial S_j^r}{\partial S_i} Q_{jS_i}^{*'} + (1+S_j^r) Q_{jS_i}^{*'} < 0,$$

в случае слабого влияния изменения СПВ на сдвиг равновесия и относительно малого значения второй производной СПВ окружения по СПВ игрока по сравнению с первой производной.

Доказательство утверждения 2.

Базируясь на формуле Лейбница для производной интеграла, пределы которого есть функции параметра S_i, запишем необходимое условие максимума первого порядка функции (6):

$$\text{(7б)}\quad \pi^{*'}_{jS_i} = P'_{S_i}Q^*_j + PQ^{*'}_{jS_i} - \Bigg\{\left[P + (1+S^r_j)Q^*_jP'_Q\right]Q^{*'}_{jS_i} + \\ + \int\limits_0^{Q^*_j}\left[P(Q^*) + (1+S^r_j)Q^*_jP'_Q\right]'_{S_i} dQ_j\Bigg\} = 0.$$

Преобразуем в этом уравнении интеграл:

$$I = \int\limits_0^{Q^*_j}\left[P + (1+S^r_j)Q^*_jP'_Q\right]'_{S_i} dQ_j =$$

$$= \int\limits_0^{Q^*_j}\left[P'_{S_i} + (1+S^r_j)(Q^{*'}_{jS_i}P'_Q + Q^*_jP''_{QS_i}) + Q^*_jP'_Q\frac{\partial S^r_j}{\partial S_i}\right] dQ_j.$$

В подынтегральном выражении параметры $Q^*_j, Q^*, Q^{*'}_{jS_i}, P'_{S_i}(Q^*), P'_Q(Q^*)$ являются характеристиками равновесия всех игроков, поэтому не зависят от действия игрока Q_j; параметр S^r_j и, следовательно, $\frac{\partial S^r_j}{\partial S_i}$ слабо зависят от действия Q_j, что было доказано ранее [6]. Поэтому считаем, что не зависят от Q_j следующие переменные:

$$Q^*_j, Q^*, P'_{S_i}, P'_Q, \frac{\partial S^r_j}{\partial S_i}, Q^{*'}_{jS_i}, S^r_j.$$

В этом случае интеграл преобразуется к виду:

$$I = (P'_{S_i} + (1+S^r_j)Q^{*'}_{jS_i}P'_Q)Q^*_j + \left((1+S^r_j)P''_{QS_i} + P'_Q\frac{\partial S^r_j}{\partial S_i}\right)Q^{*2}_j.$$

При подстановке этого интеграла в (7а) получим выражение:

$$\pi^{*'}_{jS_i} = P'_{S_i}Q^*_j + PQ^{*'}_{jS_i} - \left[P + (1+S^r_j)Q^*_jP'_Q\right]Q^{*'}_{jS_i} -$$

$$-(P'_{S_i} + (1+S^r_j)Q^{*'}_{jS_i}P'_Q)Q^*_j - \left((1+S^r_j)P''_{QS_i} + P'_Q\frac{\partial S^r_j}{\partial S_i}\right)Q^{*2}_j =$$

$$= -2(1+S^r_j)Q^*_jQ^{*'}_{jS_i}P'_Q - \left((1+S^r_j)P''_{QS_i} + P'_Q\frac{\partial S^r_j}{\partial S_i}\right)Q^{*2}_j.$$

Следовательно, уравнение (7б), из которого вычисляется оптимальная для окружения СПВ игрока i, имеет вид (7).

Условие максимума второго порядка функции (6) имеет вид:

$$\pi^{*''}_{jS_iS_i} = -\left\{2\frac{\partial S^r_j}{\partial S_i}Q^*_j Q^{*'}_{jS_i}P'_Q + \right.$$
$$+ 2(1+S^r_j)\left[Q^*_j Q^{*''}_{jS_iS_i}P'_Q + Q^*_j(Q^{*''}_{jS_iS_i}P'_Q + Q^{*'}_{jS_i}P''_{QS_i})\right] +$$
$$+ \frac{\partial S^r_j}{\partial S_i}P''_{QS_i}Q^{*2}_j + (1+S^r_j)(P'''_{QS_iS_i}Q^{*2}_j + 2Q^*_j Q^{*'}_{jS_i}P''_{QS_i}) +$$
$$\left. + P''_{QS_i}\frac{\partial S^r_j}{\partial S_i}Q^{*2}_j + P'_Q\left(\frac{\partial^2 S^r_j}{\partial S^2_i}Q^{*2}_j + 2Q^*_j Q^{*'}_{jS_i}\frac{\partial S^r_j}{\partial S_i}\right)\right\} < 0.$$

Преобразование этого условия приводит к виду:

$$\frac{\partial S^r_j}{\partial S_i}Q^*_j(2P''_{QS_i}Q^*_j + 4Q^{*'}_{jS_i}P'_Q) +$$
$$+ (1+S^r_j)\left(2(Q^{*'}_{jS_i})^2P'_Q + 2Q^*_j Q^{*''}_{jS_iS_i}P'_Q + P'''_{QS_iS_i}Q^{*2}_j + 4Q^*_j Q^{*'}_{jS_i}P''_{QS_i}\right) +$$
$$+P'_Q\frac{\partial^2 S^r_j}{\partial S^2_i}Q^{*2}_j > 0.$$

С учетом предположения о слабом влиянии изменения СПВ на сдвиг равновесия имеем:

$$P''_{QS_i} = 0, \quad P'''_{QS_iS_i} = 0, \quad Q^{*''}_{jS_iS_i} = 0.$$

С учетом предположения о малом значении второй производной СПВ окружения по СПВ игрока по сравнению с первой производной имеем:

$$\left|\frac{\partial^2 S^r_j}{\partial S^2_i}\right| \ll \left|\frac{\partial S^r_j}{\partial S_i}\right| \Rightarrow \frac{\partial^2 S^r_j}{\partial S^2_i} \approx 0.$$

При этих предположениях условие примет вид:

$$4\frac{\partial S^r_j}{\partial S_i}Q^*_j Q^{*'}_{jS_i}P'_Q + 2(1+S^r_j)(Q^{*'}_{jS_i})^2P'_Q > 0.$$

Отсюда, учитывая, что $P'_Q < 0$ в силу свойства обратной функции спроса и $Q^{*'}_{jS_i} > 0$ из свойства лидерства по Штакельбергу, получим достаточное условие максимума для решения (7) в виде (7а). ■

Определим методику вычисления производных $Q^{*'}_{jS_i}$, входящих в уравнение (7).

Утверждение 3. Производные $Q^{*'}_{jS_i}$ являются корнями следующей системы линейных уравнений:

$$\sum_{k\in N} a_{jk} Q^{*'}_{kS_i} = b_j, \quad j \in N, \tag{8}$$

где $b_j = -\left(\dfrac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q + (1+S^r_j) Q^*_j P''_{QS_i} C''_{jQ_jS_i}\right)$,

$$a_{jk} = \begin{cases} \gamma_{jk} + P'_Q, & \text{при } j \neq k, \\ \gamma_{jk} + P'_Q + (1+S^r_j)P'_Q, & \text{при } j = k, \end{cases}$$

$$\gamma_{jk} = P'_{Q_k}(Q^*) + (1+S^r_j)\left\{Q^*_j P''_{QQ_k} + P'_Q Q'_{jQ_jQ_k}\right\} - C''_{jQ_jQ_k}.$$

Доказательство утверждения 3.

Полагая, что оптимальные действия всех игроков окружения, заданные системой (4), зависят от S_i, рассмотрим n неявно заданных функций:

$$F_j(Q^*, S_i) = P(Q^*) + (1+S^r_j)Q^*_j P'_Q - C'_{jQ_j} = 0, \quad j \in N.$$

В этом случае производные $Q^{*'}_{jS_i}$ для неявно заданных функций, зависящих от нескольких независимых переменных, вычисляются из следующей системы [12]:

$$\sum_{k\in N} \frac{\partial F_j}{\partial Q_k}\frac{\partial Q_k}{\partial S_i} + \frac{\partial F_j}{\partial S_i} = 0, \quad j \in N, \tag{8a}$$

где $\dfrac{\partial F_j}{\partial Q_k} = P'_Q(Q^*) + (1+S^r_j)\left\{Q^*_j P''_{QQ_k} + P'_Q Q^{*'}_{jQ_k}\right\} - C''_{jQ_jQ_k} = \gamma_{jk}$,

$$\frac{\partial F_j}{\partial S_i} = P'_Q(Q^*)Q^{*'}_{S_i} + \frac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q + $$
$$+ (1+S^r_j)\left\{Q^*_j P''_{QS_i} + P'_Q Q^{*'}_{jS_i}\right\} - C_{jQ_jS_i}, \quad Q^{*'}_{S_i} = \sum_{k\in N} Q^{*'}_{S_i},$$

обозначение γ_{jk} введено для компонента, явно не зависящего от искомых параметров $Q^{*'}_{jS_i}$, которые обозначим $x_j = Q^{*'}_{jS_i}$. В этом случае система (8a) имеет вид:

$$\sum_{k\in N} \gamma_{jk} x_k + P'_Q \sum_{k\in N} x_k + (1+S^r_j)P'_Q x_j + \frac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q +$$
$$+ (1+S^r_j)Q^*_j P''_{QS_i} - C''_{jQ_jS_i} = 0.$$

Обозначив $b_j = -\left(\dfrac{\partial S_j^r}{\partial S_i} Q_j^* P_Q' + (1+S_j^r) Q_j^* P_{QS_i}'' - C_{jQ_jS_i}''\right)$,

$$a_{jk} = \begin{cases} \gamma_{jk} + P_Q', & \text{при } j \neq k, \\ \gamma_{jk} + P_Q' + (1+S_j^r)P_Q', & \text{при } j = k, \end{cases}$$ можно записать следующую систему линейных алгебраических уравнений относительно неизвестных $x_k : \sum_{k \in N} a_{jk} x_k = b_j, j \in N$, что соответствует (9). ■

Система (9) позволяет определить производные $Q_{jS_i}^{*'}$ в виде функций от СПВ S_j^r всех игроков, в том числе от искомой величины $\overline{S}_i$. Таким образом, определен метод вычисления целевого значения СПВ управляемого игрока, состоящий в решении уравнения (7) с учетом выражений производных $Q_{jS_i}^{*'}$ через S_i, подставленных из решения системы (8).

5. Механизм расчета оптимального управления

Рассмотрим способ, который может использовать окружение для побуждения управляемого игрока к выбору целевого значения СПВ, равного $\overline{S}_i$. Для наглядной иллюстрации будем интерпретировать рассуждения на примере

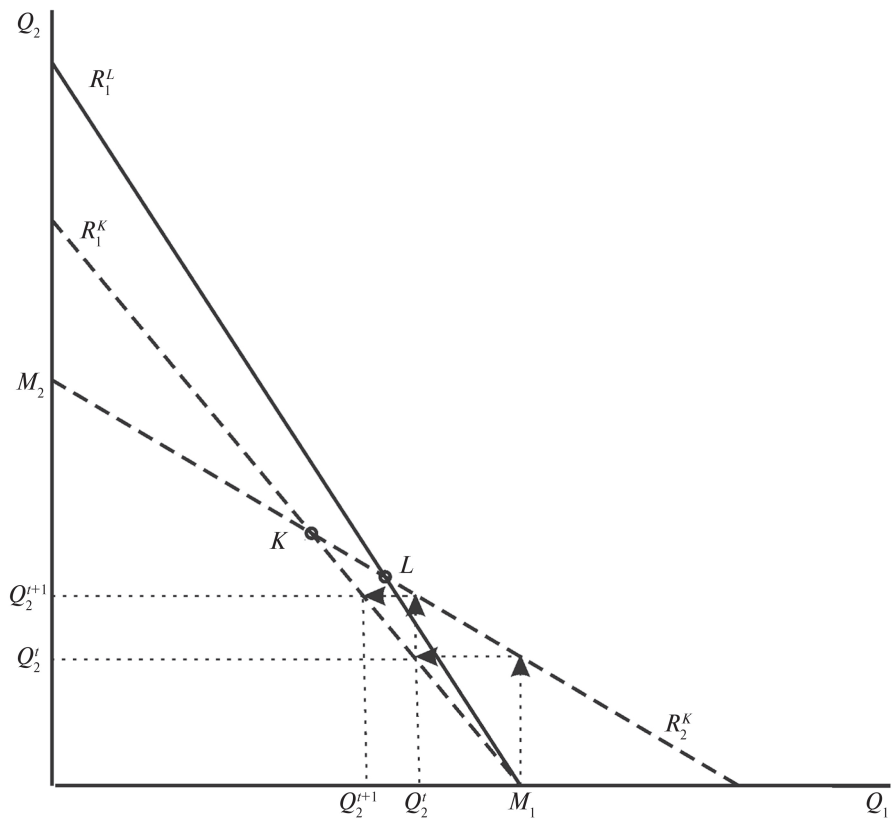

Рис. 2. Графическая иллюстрация возникновения лидера по Штакельбергу.

дуополии и начнем с описания классического принципа возникновения лидера по Штакельбергу в игре изначально равноправных участников (рис. 2), т.е. из ситуации реакций Курно (равновесие и реакции Курно обозначены символом «K»). Как известно [13], в линейной модели дуополии Курно функции оптимальных реакций игроков имеют вид

$$Q_1 = \frac{\alpha_1 - Q_2}{2}, \quad Q_2 = \frac{\alpha_2 - Q_1}{2},$$

где $\alpha_1 = \frac{a - B_i}{b}$, a, b – максимальная цена и темп снижения цены в обратной функции спроса, B_i – предельные издержки i-го игрока. Однако если бы функции реакций не были известны игрокам, они могли бы восстановить эти функции исходя из наблюдений за действиями друг друга. Если рассматривать первого игрока как потенциального лидера, то он наблюдает в динамике игры реакцию второго игрока: на действие M_1 второй игрок совершает действие Q_2^t, а на ответное действие первого Q_1^t совершает действие $Q_2^{(t+1)}$. На основе этих наблюдений первый игрок (его функция реакции R_1^K) определяет функцию реакции второго игрока R_2^K, из которой вычисляет свою предположительную вариацию (в дуополии она равна СПВ), следующим образом:

$$S_1 = Q'_{2Q_1} = -\frac{1}{2}.$$

В результате функция реакции первого игрока трансформируется к виду

$$Q_1 = \frac{\alpha_1 - Q_2}{2 + S_2} = \frac{\alpha_1 - Q_2}{2 - \frac{1}{2}},$$

после чего этот игрок становится лидером по Штакельбергу, реакция которого и соответствующее равновесие обозначено на рис. 2 символом «L». Другими словами, первый игрок, наблюдая реакцию второго игрока $Q_2 = \frac{\alpha_2 - Q_1}{2+0}$, пересмотрел свою СПВ, т.е. СПВ второго игрока $S_2 = 0$ побудила второго игрока к установлению СПВ на уровне $S_1 = -\frac{1}{2}$.

Если распространить эту процедуру на многоуровневое лидерство, то можно сформулировать следующую закономерность: для того чтобы некоторый игрок изменил свою предположительную вариацию до некоторой заданной величины, соответствующей лидеру по Штакельбергу определенного уровня, он должен наблюдать такое действие другого игрока, которое соответствует реакции лидера по Штакельбергу предшествующего уровня. Следовательно, другой игрок должен создать так называемого фантомного агента, действующего не по его истинной функции реакции, поэтому будем называть эту реакцию фантомной и обозначать символом «f». Формально это означает, что для побуждения i-го игрока установить СПВ, равную $\overline{S}_i$, окружение должно совершить действие по следующей фантомной функции реакции:

$$Q_j^f = \frac{\alpha_j - Q_i}{2 + S_j^f}$$

при условии

$$S_i = Q^{f'}_{jQ_i} = -\frac{1}{2+S^f_j} = \overline{S}_i.$$

Следовательно, СПВ окружения для этого действия вычисляется по формуле:

$$S^f_j = -\frac{1}{\overline{S}_i} - 2.$$

В общем виде этот принцип доказан ранее [6] в виде формулы (5), на основе которой можно сформулировать метод расчета фантомной реакции в общем случае нелинейных функций издержек, когда функции реакции невозможно выразить в явном виде. Обобщим представленные рассуждения в виде следующего утверждения.

Утверждение 4. Фантомная функция реакции окружения, побуждающая управляемого игрока установить целевое значение СПВ $\overline{S}_i$, соответствует СПВ окружения S^f_j, вычисленной из решения следующего уравнения:

$$\overline{S}_i = \left(\frac{1}{\sum\limits_{j \in N \setminus i} \frac{1}{u_j - S^f_j + 1}} - 1 \right)^{-1}. \tag{9}$$

На основе этого принципа опишем процедуру процесса управления в рассмотренном выше примере дуополии, предположив, что целевое значение СПВ управляемого (второго) игрока равно $\overline{S}_2 = -\frac{3}{4}$, т.е. первый игрок стремится перевести второго игрока на реакциию лидера по Штакельбергу третьего уровня (напомним, что в линейной дуополии лидерам первого, второго и третьего уровней соответствуют СПВ $-\frac{1}{2}, -\frac{2}{3}, -\frac{3}{4}$). Исходное состояние равновесия обозначим индексом «0», т.е. равновесные действия при этом будут Q^0_1, Q^0_2, СПВ игроков S^0_1, S^0_2, функции реакции R^0_1, R^0_2. Процесс управления проиллюстрирован на рис. 3.

В момент t первый игрок вычисляет свое действие по фантомной функции реакции R^f_1 лидера второго уровня $Q^{f(t)}_1 = \frac{\alpha_1 - Q^0_2}{2-\frac{2}{3}}$. Поэтому состояние игры в этот момент, обозначенное точкой E^t, описывается вектором действий $(Q^{f(t)}_1, Q^0_2)$.

В момент $t+1$ второй игрок по этому действию вычисляет $\overline{S}_2 = -\frac{3}{4}$ и переходит на целевую функцию реакции $\overline{R}_2$, заданную уравнением $Q^{t+1}_2 = \frac{\alpha_2 - Q^{f(t)}_1}{2-\frac{3}{4}}$, по которой реагирует на действие $Q^{f(t)}_1$ действием $Q^{t+1}_2 = \frac{\alpha_2 - Q^{f(t)}_1}{2-\frac{3}{4}}$. В этот момент состояние игры обозначено точкой E^{t+1}.

В момент $t+2$ первый игрок совершает действие по своей истинной функции реакции $Q^{t+2}_1 = \frac{\alpha_1 - Q^{t+1}_2}{2+S^0_2}$, поскольку он максимизировал свою полезность

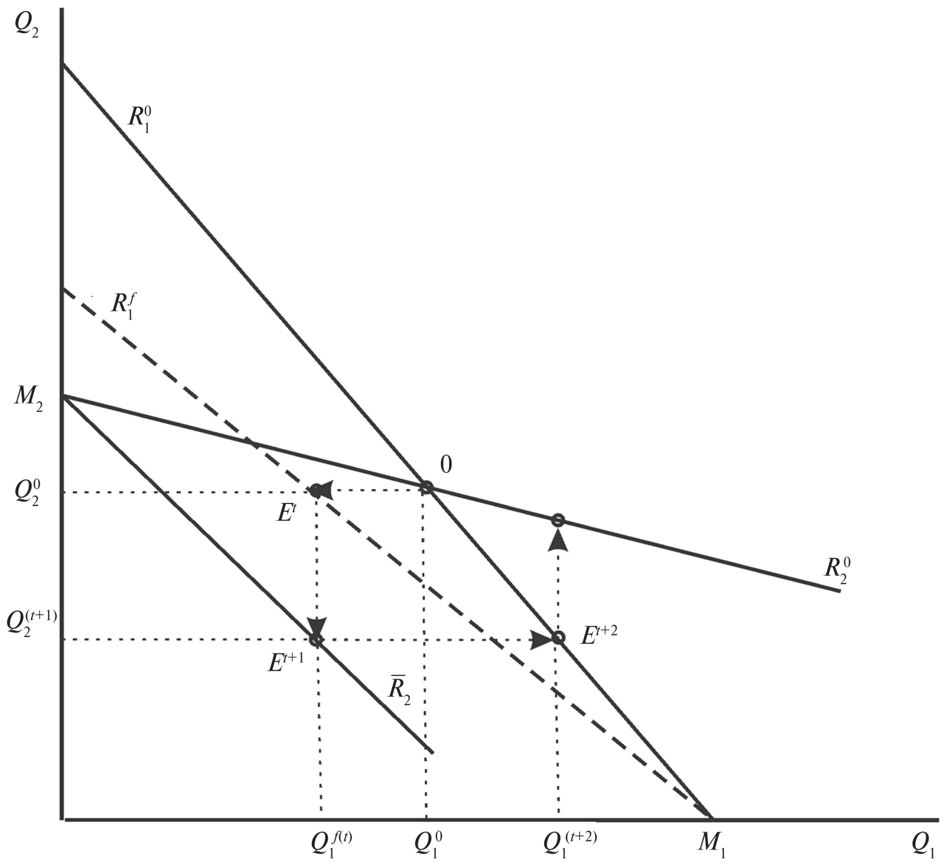

Рис. 3. Графическая иллюстрация процесса информационного управления.

для исходного равновесия. Состояние игры при сочетании (Q_1^{t+2}, Q_2^{t+1}) обозначено точкой E^{t+2}.

В последующие моменты игры исходное равновесие восстанавливается согласно процедуре, показанной выше на рис. 2. В результате описанной процедуры первый игрок получает дополнительный выигрыш полезности в моменты $t+1$ и $t+2$, поскольку второй игрок при этом совершает действия согласно целевой величине СПВ $\overline{S}_2$.

Поскольку состояние игры возвращается в исходное равновесие за бесконечное число шагов, то оценить эффективность управления для Центра можно по следующему условию:

$$\sum_{\tau=t}^{\infty} \pi^{*(\tau)} e^{-\rho\tau} - \pi^{*(0)} \sum_{\tau=t}^{\infty} e^{-\rho\tau} \geqslant 0,$$

где $\pi^{*(0)}$ – максимум полезности Центра при исходном равновесии; $\pi^{*(\tau)}$ – максимум полезности Центра в момент τ; ρ – ставка дисконтирования.

6. Заключение

В статье разработан способ информационного управления действиями игрока в теоретико-игровой модели олигополии, которое осуществляется другими игроками путем выполнения управляющего действия, побуждающего игрока сделать оптимальный для окружения ответный ход. Основаниями управления являются, во-первых, зависимость действий игроков от их предположений об ожидаемых действиях контрагентов и, во-вторых, априорная неинформированность игроков относительно предположений друг друга, связанная с дуальной природой предположительных вариаций игроков. Эти вариации, с одной стороны, базируются на анализе функций полезности окружения, а, с другой стороны, игрок не может игнорировать характер реагирующих действий окружения. Поэтому в статье принята гипотеза о том, что в случае противоречия между этими двумя подходами игроки оценивают предположительные вариации по действиям друг друга, которые являются более достоверной информацией. Благодаря этому окружение может, не изменяя своих предположительных вариаций, совершить действие как бы от лица фантомного игрока, побуждающего управляемого игрока к выгодной для окружения реакции, а последний воспринимает это действие как сигнал об изменении истинной реакции окружения и совершает желательное для него действие.

Основные результаты исследования можно обобщить следующим образом. Сформулирована модель взаимодействий игроков в ситуации олигополии в виде иерархической игры, в которой окружение рассматривается как Центр, а некоторый игрок как объект управления. Выведено выражение максимума функции полезности окружения в зависимости от вектора СПВ всех игроков, позволяющее найти значение СПВ управляемого игрока, оптимизирующее функцию полезности окружения. Определена методика вычисления целевого с позиций окружения значения СПВ управляемого игрока. Разработана итерационная процедура, побуждающая управляемого игрока выбрать функцию реакции, соответствующую целевому значению СПВ, в результате которой окружение максимизирует свою полезность.

Поскольку задача оптимального управления сформулирована для игры n лиц и для функций полезности общего вида, то в разработанных методиках и процедурах невозможно получить решения в явном виде, позволяющие проанализировать результаты игры. Поэтому дальнейший этап исследования предусматривает приложение этих общих инструментов к конкретным видам функций полезности и проведение численных экспериментов.

СПИСОК ЛИТЕРАТУРЫ

1. *Anderson S.P., Erkal N., Piccinin D.* Aggregative games and oligopoly theory: short-run and long-run analysis // RAND J. Econ. 2020. 51(2), P. 470–495.
2. *Cournot A.A.* Researches into the Mathematical Principles of the Theory of Wealth. London: Hafner, 1960. (Original 1838).

3. *Nash J.* Non-cooperative Games // Ann. Math. 1951. No. 54. P. 286–295.
4. *Stackelberg H.* Market Structure and Equilibrium: 1st Edition. Translation into English, Bazin, Urch & Hill, Springer, 2011. (Original 1934).
5. *Novikov D.A., Chkhartishvili A.G.* Reflexion and Control: Mathematical Models. London: CRC Press, 2014.
6. *Julien L.A.* On noncooperative oligopoly equilibrium in the multiple leader-follower game // Eur. J. Oper. Res. 2017. V. 256(2). P. 650–662.
7. *Гераськин М.И.* Свойства предположительных вариаций в нелинейной модели олигополии Штакельберга// АиТ. 2020. № 6. С. 105–130.
Geraskin M.I. The Properties of Conjectural Variations in the Nonlinear Stackelberg Oligopoly Model // Autom. Remote Control. 2020. V. 81. No. 6. P. 1051–1072.
8. *Malsagov M., Ougolnitsky G., Usov A.* A differential Stackelberg game theoretic model of the promotion of innovations in universities // Advances Syst. Sci. Appl. 2020. V. 20. No. 3. P. 166–177.
9. *Губанов Д.А., Новиков Д.А., Чхартишвили А.Г.* Модели информационного влияния и информационного управления в социальных сетях // Проблемы управления. 2009. № 5. С. 28–35.
10. *Губанов Д.А., Новиков Д.А., Чхартишвили А.Г.* Социальные сети: модели информационного влияния, управления и противоборства. М.: Изд-во физ-мат. лит-ры, 2010.
11. *Губанов Д.А., Петров И.В., Чхартишвили А.Г.* Многомерная модель динамики мнений в социальных сетях: индексы поляризации // Проблемы управления. 2020. № 3. С. 26–33.
Gubanov D.A., Petrov I.V., Chkhartishvili A.G. Multidimensional Model of Opinion Dynamics in Social Networks: Polarization Indices // Autom. Remote Control. 2021. V. 82. No. 10. P. 1802–1811.
12. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. М.: Наука, 1974.
13. *Алгазин Г.И., Алгазина Д.Г.* Моделирование динамики коллективного поведения в рефлексивной игре с произвольным числом лидеров // Информатика и автоматизация. 2022. Т. 21. № 2. С. 339–375.

Статья представлена к публикации членом редколлегии А.Г. Чхартишвили.

Поступила в редакцию 29.11.2024

После доработки 21.03.2025

Принята к публикации 25.03.2025