УДК 519.83

ЭВОЛЮЦИОННАЯ ИГРА ПРЕДУСМОТРИТЕЛЬНЫХ АГЕНТОВ С БИНАРНЫМ ВЫБОРОМ

Г.Т. Броницкий
МГУ им. М.В. Ломоносова
Россия, 119991, Москва, Ленинские горы, 1
E-mail: getibr@gmail.com

Е.Е. Серебрянникова
Физический институт им. П.Н. Лебедева РАН
Россия, 119991, Москва, Ленинский просп., 53, стр.4
E-mail: serebryannikova@lpi.ru

Ключевые слова: бинарный выбор, дальновидное поведение, близорукое поведение, полный граф, уравнение Беллмана.

Аннотация: В докладе[a] рассматривается эволюционная игра N агентов, находящихся в узлах полного графа в дискретном времени. Каждый агент может находиться в одном из двух состояний (±1). Ключевая особенность модели состоит в том, что при принятии решения агент руководствуется не мгновенной полезностью, а пытается максимизировать полезность на заданном фиксированном интервале, то есть является предусмотрительным. Поиск оптимальной стратегии для игроков моделируется при помощи построения и решения уравнения Бэллмана. Основная цель работы состоит в изучении характеристик решения при различной степени предусмотрительности агентов.

[a]Доклад основан на материалах работы Georgy Bronitsky, Andrey Leonidov, and Ekaterina Serebryannikova. Binary choice evolutionary game of farsighted agents. *Work in progress.*

1. Введение

В докладе рассматривается эволюционная игра N агентов, расположенных в узлах полного графа и совершающих бинарный выбор, формализуемый значением переменной $\sigma_i(t) = \pm1,\ i = 1, \ldots, N$, характеризующей состояние, в котором находится агент. Мгновенная (одпопериодная) функция выигрыша агента i при этом определяется следующим соотношением:

(1) $$v^i(s_i(t), s_{-i}(t)) = s_i(t)\frac{J}{N}\sum_{j=1,\ldots,N,\ j\neq i} s_j(t) + \varepsilon^i_{\sigma_i},$$

где $\varepsilon^i_{\sigma_i}$ - это некоторая случайная величина, $s_{-i}(t)$ - вектор значений стратегий всех игроков, кроме i. Или иначе, если ввести определение $N^+(t)$ ($N^-(t)$) как количества игроков в состоянии $+1$ (-1), а также

$$n_i(t) = \sum_{j=1,\dots,N,\ j\neq i} s_j(t) \equiv N^+(t) - N^-(t) - s_i(t),$$

то соотношение (1) примет вид

$$v^i(s_i(t), s_{-i}(t)) = \frac{J}{N} s_i(t) n_i(t) + \varepsilon^i_{\sigma_i} \tag{2}$$

Эволюционная игра такого рода рассматривалась в работе [1], где предполагалось, что в определенные моменты времени агентам предоставляется возможность выбрать оптимальное состояние, так чтобы максимизировать значение полезности (1), исходя из текущих состояний остальных игроков. Моменты выбора в модели [1] определялись «пуассоновскими часами» агентов. Такое поведение агентов называют близоруким, поскольку, принимая решение, агент стремится максимизировать значение текущей полезности, не пытаясь оптимизировать выбор относительно ожидаемой в будущем динамики. Важный результат работы [1] состоит в том, что в данной системе при определенных значениях параметров имеет место перестройка пространства решений («фазовый переход»).

В текущей работе предложена модель, аналогичная [1], но предполагающая наличие предусмотрительности у агентов. Под предусмотрительностью агентов понимается стратегическое поведение, при котром каждый агент в момент принятия решения t выбирает действие так, чтобы максимизировать приведенное значение ожидаемой полезности следующего вида:

$$\begin{aligned} W^i(t, s_i(t), s_i(t+1), \dots, s_i(T), s_{-i}(t), s_{-i}(t+1), \dots, s_{-i}(T)) = \\ v^i\left(s_i(t), s_{-i}(t)\right) + \mathbb{E}\left[\sum_{\tau=t+1}^{T} \gamma^{\tau-1} v^i(s_i(\tau), s_{-i}(\tau))\right], \end{aligned} \tag{3}$$

где γ - норма дисконтирования, характеризующая степень предусмотрительности агента. То есть ожидаемая полезность состоит из двух компонент: отвечающей близорукому поведению (одноопериодный выигрыш) и ожидаемая полезность от будущих действий. В результате, механизм близорукого выбора дополняется механизмом стратегического предусмотрительного поведения.

Рассматривается игра в дискретном времени на ограниченном временном интервале $t = 1, 2, \dots, T$.

В каждый момент времени происходят следующие события:

- Случайному игроку i выпадает возможность изменить свою стратегию
- Игрок i, которому выпала возможность изменить свое состояние, делает выбор ($s_i = \pm 1$) так чтобы максимизировать суммарную ожидаемую дисконтированную полезность (3) на горизонте T лет. Задача максимизации полезности (3) решается путем построения уравнения Беллмана.
- Все остальные игроки не меняют свои состояния и получают выигрыш согласно формуле (1).

2. Построение уравнения Беллмана

Цель каждого агента в момент принятия решения состоит в максимизации полезности вида (3). Пользуясь принципом оптимальности Беллмана, для каждого агента $i = 1, \ldots, N$ необходимо найти вид функции Беллмана $V^i(\Phi_i, t)$, где Φ_i - это некоторый вектор, характеризующий состояние системы (относительно игрока i), определяемой соотношением

$$V(\Phi_i, t) = \max_{s_i(t), s_i(t+1), \ldots, s_i(T)} W^i(t, s_i(t), \ldots, s_i(T), s_{-i}(t), \ldots, s_{-i}(T) | \Phi_i(t) = \Phi_i), \tag{4}$$

При этом состояние системы (относительно игрока i) характеризуется следующим вектором переменных:

$$\Phi_i(t) = (\mathbb{I}_i(t), s_i(t-1), n_i(t), \varepsilon^i_{+1}, \varepsilon^i_{-1}),$$

где переменная $\mathbb{I}_i(t)$ принимает значение 1, если в момент времени t играет игрок i, и 0 иначе.

Для того, чтобы выписать уравнение Беллмана, проанализируем сначала действия игрока i в последний момент времени T.

$$V^i(\Phi_i(T), T) = \begin{cases} \max_{s_i = \pm 1} v^i(s_i, s_{-i}(T-1)), & \text{если ходит игрок } i, \\ v^i(s_i(T-1), s_{-i}(T)), & \text{иначе} \end{cases} \tag{5}$$

Другими словами,

$$V^i(\Phi_i(T), T) = \mathbb{I}_i(T) \max_{s_i = \pm 1} v^i(s_i, s_{-i}(T-1)) + (1 - \mathbb{I}_i(T)) v^i(s_i(T-1), s_{-i}(T)). \tag{6}$$

Аналогично получаем, что для произвольного момента времени t уравнение Беллмана имеет вид:

$$\begin{aligned} &V^i(\Phi_i(t), t) = \\ &= \begin{cases} \max_{s_i = \pm 1}(v^i(s_i, s_{-i}(T-1)) + \gamma E V^i(\Phi_i(t+1), t+1)), & \text{если ходит } i, \\ v^i(s_i(T-1), s_{-i}(T)) + \gamma E V^i(\Phi_i(t+1), t+1), & \text{иначе,} \end{cases} \end{aligned} \tag{7}$$

где

$$EV^i(\Phi_i(t+1), t+1) = \int_{\Phi_i} p(\Phi_i | s_i(t)) V^i(\Phi_i, t+1) d\Phi_i,$$

$p(\Phi_i | s_i(t))$ - вероятность попасть в состояние Φ_i, если в момент времени t игрок i играет стратегию $s_i(t)$.

Решение уравнения Беллмана предлагается производить численно, методом обратной индукции. Результатом решения данного уравнения является вид функции Беллмана $V^i(\Phi_i, t)$, который затем используется в каждый момент времени для поиска оптимальной стратегии.

3. Заключение

В докладе сформулирована задача, обобщающая динамическескю модель бинарного выбора [1] на случай наличия предусмотрительных агентов. На основе анализа численных экспериментов в дальнейшем планируется проведение сравнения кинетических характеристик решения при различной степени предусмотрительности агентов. В частности, наибольший интерес представляет сравние со случаем игры близоруких агентов, описанным в [1].

Список литературы

1. Blume L., Durlauf S. Equilibrium concepts for social interaction models // International Game Theory Review, 2003. Vol. 5. No. 3. P. 193-209.