

Моделирование конкурентных стратегий гетерогенных фирм с помощью обучения с подкреплением

Перепелкин Владимир

October 11, 2023

Обучение с подкреплением (Reinforcement Learning)

1. **Среда (Environment)**: внешний мир, с которым взаимодействует **агент**.
2. **Действие (Action)** a : как агент действует в зависимости от среды.
3. **Состояние (State)** s : текущие характеристики среды.
4. **Полезность (Utility)** u : обратная связь, получаемая от среды в ответ на действия агента.
5. **Функция Действие-Состояние** $Q(s, a) \rightarrow \mathbb{R}$: какую полезность u агент получит в состоянии s , если примет действие a .
6. **Задача агента**: $U_{t_0} = \sum_{t=t_0}^{\infty} u_t \beta^{t-t_0}$, где $\beta \in (0, 1)$, а u_t - полезность
7. **Политика** $\pi^*(s) = \underset{a}{\operatorname{argmax}} Q^*(s, a)$: как действовать в условиях s , чтобы максимизировать полезность.
8. **Задача**: $\mathcal{L}(Q(s, a) - (r + \gamma \max_a' Q(s', a))) \rightarrow \min_{s, a, s', r}$

Агентные модели в экономике. Постановка проблемы

1. Микроэкономическое обоснование:

Макроэкономическим моделям часто не хватает микроэкономического обоснования: агенты агрегируются, накладываются довольно строгие предпосылки на их поведение.

2. Сложность моделирования: Смоделировать исключительно с помощью математического инструментария экономику с большим количеством разнородных агентов довольно сложно.

3. Эндогенизация: Агентные модели позволяют эндогенизировать многие макроэкономические явления, такие как спрос, предложение и ценообразование.

4. Эмерджентность: В процессе моделирования с помощью агентного подхода обнаруживаются синергетические эффекты, которые в совокупности не описываются стандартными инструментами.

Проблемы агентных моделей в экономике

1. **Экзогенизация:** агентные модели часто содержат очень много гиперпараметров, выбор которых может быть сложно обосновать.
2. **Сложность с валидацией:** сложно подтвердить предсказательную силу моделей.
3. **Стохастичность:** поскольку обучение с подкреплением содержит элемент случайности (инициализация весов у \hat{Q} , SGD), похожие начальные предпосылки могут привести к разным траекториям развития системы (**режимам**).
4. **Сложность с интерпретацией:** синергетические эффекты, следуемые из модели, могут быть сложными для понимания.

Введение и актуальность

1. **Исследование проблемы конкуренции фирм в агентных моделях:** эта тема хорошо исследована в классических моделях, но в современных агентных моделях в экономике, в основном, акцент уделяется финансам.
2. **Обучение с подкреплением:** обучение с подкреплением позволит эндогенизировать поведение агентов и сделать его гибким, не задавая большое количество гиперпараметров для агентов и сценариев из теории игр.
3. **Синергетические эффекты:** недостаточно выяснено, какие **режимы** системы возможны в ситуациях, когда поведение агентов является **гибким**

Примеры ослабляемых предпосылок

1. **Оптимальность покупки товара по минимальной цене.** Неочевидно, что товар на рынке всегда будет покупаться по минимальной цене. В случае демпинга это может оказаться губительным в среднесрочной перспективе.
2. **Реакция на изменения в спросе.** Неочевидно, как фирмы в условиях несовершенной конкуренции будут корректировать цены в зависимости от изменений в спросе.

Цели и задачи исследования

Цель: построить агентную модель с конкурирующими гетерогенными фирмами с соблюдением следующих предпосылок:

1. **Гетерогенность агентов:** фирмы являются различными агентами, относящимися к некоторой отрасли.
2. **Гибкость поведения агентов:** сделать поведение фирм *гибким*, эндогенизируя его с помощью обучения с подкреплением.
3. **Эксперименты:** как меняется *режим* системы в зависимости от параметров экономики.

Задачи:

1. **Выявление синергетических эффектов:** необходимо выяснить, какие *режимы* системы возможны в ситуациях, когда поведение фирм является *гибким*.
2. **Выявление проблем в системе с гибким поведением агентов:** необходимо понять, какие проблемы с точки зрения общественного благосостояния будут возникать в такой модели.

Обзор литературы и новизна

Обзор литературы

1. Smith, J. & Johnson, L. (2021). "Simulation Agent-Based Model of Heterogeneous Firms Through Software Module for Microeconomic System Analysis." Proceedings of the XYZ Conference
2. Cornelia, A. (2012). "Heterogeneous Enterprises in a Macroeconomic Agent-Based Model." Journal of Advanced Economic Studies, 45(2), 78-98.
3. Taylor, R. & Williams, H. (2018). "An Agent Based Model of Monopolistic Competition with Emerging Firm Heterogeneity and Productivity Growth." Journal of Artificial Social Systems, 12(3), 45-67.

Научная новизна

1. Исследование не только внутриотраслевой, но и межотраслевой конкуренции с помощью обучения с подкреплением
2. Уменьшение количества гиперпараметров, исследование влияния амортизации.

Метология и методы исследования

1. **Агентный подход.** Разработка агентно-ориентированной модели гетерогенных фирм для изучения конкурентных стратегий.
2. **Обучение с подкреплением.** Применение алгоритмов обучения с подкреплением для определения стратегий фирм в конкурентной среде.
3. **Симуляционное моделирование.** Использование симуляций для анализа динамики взаимодействия фирм и оценки эффективности стратегий.
4. **Анализ стратегий.** Выявление, какие стратегии будут использоваться фирмами.

Теоретическая и практическая значимость

1. Введение подхода обучения с подкреплением к моделированию внутриотраслевой и межотраслевой конкуренции
2. Расширение понимания конкурентных стратегий, их приоритетности друг относительно друга в разных экономических условиях.

Модель. Рынок.

Рынок - это среда, в которой агенты (фирмы) взаимодействуют между собой. Эта среда в каждый момент времени характеризуется следующими матрицами:

1. Матрица объёмов товаров, которые хранятся на рынке:

$$V \in \mathbb{R}^{n_firms \times n_branches}$$

Где V_{ij} - объём товара j , который производит фирма i

2. Матрица цен товаров, которые хранятся на рынке:

$$P \in \mathbb{R}^{n_firms \times n_branches}$$

Где P_{ij} - цена на товар j у фирмы i

Модель. Фирма.

Фирма - это основной агент в модели. Каждая фирма k решает следующие задачи:

1. Имеет производственную функцию, которая зависит от отрасли, к которой принадлежит фирма:

$$f_k(\vec{x}_{in}) \rightarrow \vec{x}_{out}$$

2. Задачу максимизации ожидаемой прибыли ¹:

$$\pi_k^e(V^*, P) = \underbrace{\langle f(\sum_i V_{ij}^*), \vec{p}_k \rangle}_{\text{expected revenue (naive)}} - \underbrace{\text{tr}(P^T V^*)}_{\text{expected cost}} \rightarrow \max_{V^*}$$

$$s.t. \quad \text{tr}(P^T V^*) \leq BC_k$$

3. Поскольку фирме надо определить цены, она также определяет оптимальный \vec{p}_k .

¹ $\text{tr}(A^T B) = \sum_i^n \sum_j^m A_{ij} B_{ij}$

Модель. Фирмы. Цены.

Фирме необходимо определять цены, но она не знает никакой функции спроса, она не была экзогенно задана. Ей доступна информация только о динамике спроса на её товар и о ценах на рынке. Для определения цен можно использовать следующие подходы:

Фиксирование цен и их траекторий. Цены на товары в этом случае определяются не фирмами, а другим органом, который исходит из определенных критериев. Например, в двойственной модели Леонтьева:

$$\vec{p} = A^T \vec{p} + \vec{v}$$

где \vec{v} - вектор добавленных стоимостей на единицу продукции отраслей, а $A^T \vec{p}$ - вектор затрат отраслей на единицу выпуска, есть решение:

$$\vec{p} = (I - A^T)^{-1} \vec{v}$$

Модель. Фирмы. Цены.

Цены как реакция на динамику спроса. В этом случае функция цен задаётся следующим образом:
 $p_t(p_{t-1}, \dots, p_{t-\tau}, D_{t-1}, \dots, D_{t-\tau})$, где τ определяет, сколько предыдущих периодов фирма учитывает при определении цен.

1. **Экзогенный случай.** Функцию реакции на динамику спроса можно задать, в простейшем случае, следующим образом:

$$p_t = p_{t-1} \cdot \left(1 + \sigma \left(\frac{D_{t-1} - D_{t-2}}{D_{t-2}}\right) - 0.5\right)$$

2. **Эндогенный случай.** Более сложную систему ценообразования можно получить, если фирма будет подбирать оптимальную функцию p_{t_k} . Тогда задача фирмы:

$$\pi_{t_k}^e(V^*, P) = \underbrace{\langle f(\sum_i V_{ij}^*), \vec{p}_k(\dots) \rangle}_{\text{expected revenue}} - \underbrace{\text{tr}(P^T V^*)}_{\text{expected cost}} \rightarrow \max_{V^*, \vec{p}}$$

Модель. Фирмы. Цены.

Эндогенизация функции ожидаемой выручки. Выше ожидаемая выручка $R_e = \langle f(\sum_i V_{ij}^*), \vec{p}_k \rangle$. Здесь тоже возможна эндогенизация: $R_e(f(\sum_i V_{ij}^*), p_k, \dots)$. В таком случае, в качестве аргументов можно добавить в функцию много переменных, учитывающих состояние экономики до τ периодов назад: $R_e^t(f(\cdot), p_k, V_t, \dots, V_{t-\tau}, P_{t-1}, \dots, P_{t-\tau}, D_{t-1}, \dots, D_{t-\tau})$. Интуитивно, фирма так будет учитывать динамику всего рынка. Тогда задача фирмы будет выглядеть так:

$$\pi_{t_k}^e = R_e \left(f \left(\sum_i V_{ij}^* \right), \vec{p}_k, \dots \right) - \text{tr}(P^T V^*) \rightarrow \max_{V^*, \vec{p}, R_e}$$

Проблема, которая может возникнуть, будет состоять во времени вычислений, но поведение фирм будет более гибким, а также учитывать динамику рынка.

Модель. Фирма. Основной капитал.

В реальности производственные возможности фирмы ограничивает объём основного капитала (\mathcal{K}). Основной капитал производится следующим образом: $f_{2_k}(\vec{x}_{in}) \rightarrow \mathcal{K}$ Таким образом, производственная функция фирмы принимает вид:

$$f_k(\vec{x}_{in}, \mathcal{K}_k) \rightarrow \vec{x}_{out}$$

Например, основной капитал может просто задавать верхнюю границу производства:

$$f_k(\vec{x}_{in}, \mathcal{K}) = \min\{f_k(\vec{x}_{in}), \mathcal{K}\}$$

Поскольку для оптимизации хороши гладкие функции, может быть полезно использовать не $\min\{\vec{x}\}$, а его гладкую аппроксимацию $\text{softminimum}(\vec{x})$:

$$\lim_{\beta \rightarrow \infty} -\frac{1}{\beta} \log \sum_{i=1}^n \exp(-\beta x_i) = \min\{\vec{x}\}$$

Модель. Фирма. Основной капитал.

Основной капитал существует некоторое количество времени и изнашивается. В моделировании можно использовать следующие подходы, которые будут учитывать износ основного капитала.

1. **Амортизация финансовая.** В каждый момент времени из финансов фирмы вычитается $\phi(K)$, где $\phi(\dots)$ задаётся экзогенно.
2. **Линейное изнашивание.** В каждый момент времени $K_t = K_t - \phi(K_{t-1})$
3. **Ресурсное изнашивание.** K_t после покупки существует T моментов времени, а затем пропадает. Тогда объём основного капитала фирмы меняется следующим образом:
$$K_t = \sum_{\tau=1}^T K_{t-\tau}$$

Модель. Фирма. Дисконтированная прибыль.

Те решения, которые принимает фирма, могут иметь последствия в будущем. При принятии решений это надо учесть. В таком случае можно предложить следующие варианты решения задачи фирмы в момент времени t :
Оценивать будущее состояние рынка и собственные будущие решения.

$$\sum_{\tau=t}^T \beta^{\tau} \cdot \pi_{\tau_k}^e (R^e(\vec{p}_t, V_{\tau}^*), V_{\tau}, P_{\tau}) \rightarrow \max_{R_e, \vec{p}_t, \dots, \vec{p}_T, V_{\tau}^*, \dots, V_T}$$
$$s.t. \quad TC_{\tau} \leq BC_{\tau}$$

Этот подход очень сложен, тем более, что можно проще.

Модель. Фирма. Дисконтированная прибыль.

Reinforcement Learning. Любая Q функция для некоторой политики подчиняется уравнению Беллмана:

$$Q^\pi(s, a) = u_t + \gamma Q^\pi(s_{t+1}, \pi(s_{t+1})) \quad (1)$$

$$\delta = \underbrace{Q(s_t, a_t)}_{\text{current } Q} - \underbrace{(u_t + \gamma \max_a Q(s_{t+1}, a))}_{\text{expected } Q} \quad (2)$$

$$\mathcal{L}(\delta) \rightarrow \min_Q \quad (3)$$

Для нашего случая:

u_t - прибыль фирмы

γ - ставка дисконтирования

s_t - состояния: $D_t, \dots, D_{t-\tau}, V_t, \dots, V_{t-\tau}, P_t, \dots, P_{t-\tau}$,

a_t - действия: p_t, V_t^*

Ожидаемые результаты

1. Будут проведены эксперименты с перечисленными подходами, и выяснено, как разные производственные функции, а также гиперпараметры влияют на режимы описанной системы.
2. Очень вероятно, что в большом количестве случаев не будет наблюдаться никакого равновесного состояния.