Міністерствоосвіти і науки України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Кафедра математики

РІВНОВАГА ЗА НЕШЕМ У СТОХАСТИЧНИХ ІГРАХ З ДИСКОНТОВНИМ КРИТЕРІЄМ

Текстовачастина до курсовоїроботи за спеціальністю „Прикладна математика”6.040301

Керівниккурсовоїроботи

____

____

“____” _____________2020 р.

Виконав студент___________

____

“____” _____________2020 р.

Київ 2020

ЗМІСТ

Вступ

Теорія ігор як наука була заснована відносно нещодавно – вона бере свій початок у 1950 роках. Концепцію рівноваги у своїй праці у 1950 році визначив Джон Неш[1]. Пізніше, у 1953 році ЛлойдШеплі опублікував статтю, де визначив термін стохастичної гри [2]. Взагалі кажучи, стохастична гра відрізняється від Марківського процесу прийняття рішень кількістю гравців – більше, ніж одиниця. Найбільш дослідженими є повторювані ігри в нормальній формі, для них існує певна кількість алгоритмів вирішення, але стохастичні ігри відрізняються від повторюваних тим, що мають певну скінченну або нескінченну множину станів, в яких в певні моменти часу знаходиться гра. Можна сказати, що стохастична гра – це певний набір ігор в нормальній формі, і коли змінюється стан стохастичної гри, гравці просто грають в іншу гру. В цій роботі розглядатиметься гра зі скінченною множиною станів.

Об’єктом дослідження виступає дисконтована стохастична гра, яка є моделлю конкурентної ситуації між декількома особами для отримання вигоди з певної ситуації. Предметом є рекомендації щодо прийняття рішень людьми у різних життєвих ситуаціях, моделлю яких якраз і виступає стохастична гра. Метою роботи є розробка застосунку, що прийматиме формальну модель стохастичної гри з визначеною кількістю гравців, станів та дій і повертатиме профіль стратегій, що є рівновагою за Нешем для даної гри. Для розробки було обрано мову програмування Python.

Перший розділ присвячено аналізу та формулюванню моделі стохастичної гри з усіма ключовими елементами, визначення поняття стратегій для гравців та який профіль стратегій є рівновагою за Нешем.

В другому розділі наведений аналіз алгоритмів стохастичної гри з урахуванням дисконтованого критерію, порівняння рівноваги за Нешем з

Марківською рівновагою, та обрання шляху для обчислення рівноваги у застосунку.

Третій розділ присвячено огляду алгоритму, його реалізації на мові Python. Також наведені результати роботи програми і їх аналіз.

РОЗДІЛ 1: Аналіз задачі пошуку рівноваги за Нешем

1.1 Моделювання стохастичної гри з дисконтованим критерієм

Гра з точки зору теорії ігор – це будь-яке протистояння або конфлікт двох або більше сторін, які називаються гравцями. Гравці мають задану скінченну множину дій, доступних для них. Оскільки розглядається стохастична гра, визначена також скінченна множина станів, у яких може перебувати гра. За свої дії гравці отримують виплату – вона може бути від’ємною, додатною, або ж нульовою. Тому також визначена функція, яка у відповідь на дію гравця у певному стані гри повертатиме дійсне число, яке є виплатою цього гравця. Також має бути визначений закон руху між станами, який визначатиме розподіл вірогідностей, за якими гра переходить в інший стан, або ж залишається в поточному. Коли постає питання щодо визначення вигоди гравця за дуже тривалий період, недостатньо просто просумувати всі виплати і поділити їх на кількість зіграних ігор. В результаті, якщо виплати, отримані гравцем на початкових етапах гри будуть кардинально відрізнятися від тих, які він буде отримувати впродовж наступних етапів, то середнє значення не буде враховувати цю різницю за рахунок того, що гравець набагато частіше у майбутньому отримуватиме інші виплати.Щоб оцінити виплати гравців за тривалий період часу, або у далекому майбутньому, визначається дисконтований критерій – величина, яка вказує на те, наскільки гравців «хвилює ситуація в майбутньому». Саме цей критерій враховується при обрахуванні суми виплат гравця за тривалий проміжок часу, зберігаючи дані за увесь час.

Отже, модель дисконтованої стохастичної гри наступна :

$$(N, S, A^{*}, (A^{j}, r^{j}) : j \in N, q, \beta),$$

де N –скінченна множина гравців;

S – скінченнамножинастанів;

$A*$ – скінченнамножинаусіхдій, доступнихгравцям;

$A^j : S \to A*$- функція, що визначає набір дій для кожного гравця $j \in N$ в стані $s \in S$;

$r^j : \{(s,a): s \in S, a \in A*\} \to \mathrm{R}$ - функція виплат;

$q: \{(s,a): q \in S, a \in A*\} \to S$ – закон руху;

β – дисконтований критерій, $\beta \in (0;1)$.

Гра розглядається в чистих стратегіях, що означає вибір гравцем певної дії свідомо, на відміну від змішаних стратегій, де гравці виконують дії з певною ймовірністю.

Сама гра проходить наступним чином : гра знаходиться в початковому стані $s_1 \in S$. У стані m гравці мають доступ до історії гри

$$(s_1, a_1, s_2, a_2, ..., a_{m-1}, s_m),$$

де s_i – стан гри на i-му етапі;

a_i – комбінація дій, зіграних гравцями на відповідному етапі.

Кожен гравець i обирає незалежно від інших гравців та одночасно з ними дію $a_m^i \in A^{\mathrm{i}}(s_m)$, отримує виплату $r^{\mathrm{i}}(s_m, a_m)$, і гра переходить до наступного стану s_{m+1} відповідно до закону руху $q(\cdot | s_m, a_m)$.

1.2 Визначення стратегії гравців

Нехай H^m – множина усіх скінченних історій довжиною m. Тоді нехай множина усіх історій буде об’єднанням :

$$H = U_{m \in N} H^m$$

Для кожної скінченної історії

$$h = (s_1, a_1, s_2, a_2, ..., a_{m-1}, s_m)$$

позначатимемо довжину як

$$L(h) = m,$$

та останній стан, в якому перебувала гра як

$$s_L(h) = s_m.$$

Стратегією для гравця *i* є функція

$$\tau^{\mathrm{i}} : H \to A^*,$$

де $\tau^{\mathrm{i}}(h) \in A^{\mathrm{i}}(s_L(h)), \forall h \in H$.

Тоді $\tau = (\tau_1, \tau_2, ..., \tau_n)$ – профіль стратегій, що складається зі стратегій τ_i для кожного гравця i.

1.3 Визначення рівноваги за Нешем

Для профілю стратегій τ β-дисконтовнаочікуванавиплатагравця i

$$\gamma_i(\tau)(s) = E_s^\tau\left(\sum_{n=1}^{\infty}(1-\beta)\beta^{n-1}r_i(s_n, a_n)\right),$$

де $\beta \in (0;1)$ – фіксований дисконтований критерій;

E_s^τ – очікуваний оператор відносно ймовірнісного розподілу.

Нехай $\tau^* = (\tau_1^*, \tau_2^*, ..., \tau_m^*)$ – фіксований профіль стратегій. Запис (τ_{-i}^*, τ_i^*) позначає, що гравець i обирає інакшу стратегію, за умови того, що всі інші гравці грають ті ж самі стратегії.[3]

Тоді рівновага за Нешем у стохастичній грі з дисконтованим критерієм означає, що ні одному з гравців не вигідно відхилятись від обраної стратегії, за умови того, що інші гравці гратимуть ті ж самі стратегії, і має вигляд

$$\gamma_i(\tau^*)(s) \geq \gamma_i(\tau_{-i}^*, \tau_i^*)(s).$$

РОЗДІЛ 2: Пошук алгоритму знаходження рівноваги за Нешем

2.1 Марківська рівновага та її зв'язок з рівновагою за Нешем

В переважній більшості досліджень на тему стохастичних ігор розглядаються чисті стратегії Маркова. Тобто, стратегії гравців залежать лише від поточного стану, в якому знаходиться гра[3]. Такі стратегії мають перевагу в тому, що вони просто обчислюються.

Відповідно, рівновага, яка базується на таких стратегіях називається Марківською рівновагою, і її також асоціюють з рівновагою за Нешем. Алгоритм для знаходження саме такої рівноваги буде розроблений в цій роботі.

Залишається вирішити проблему існування такої рівноваги у стохастичних іграх. Починаючи ще з робіт Шеплі[2], ця проблема залишається вкрай важливою. Але спочатку було доведення існування такої рівноваги для ігор з двома гравцями, та скінченними множинами станів та дій [4], а згодом це було узагальнено в роботах [5], [6]для загального випадку зі скінченною кількістю гравців m та, відповідно, скінченною кількістю станів та скінченною кількістю дій. Саме існування такої рівноваги потрібно для того, щоб алгоритм працював для будь-якого загального випадку, адже в роботі розглядаються лише скінченні множини.

Якщо всі гравці гратимуть лише Марківські стратегії, то для кожного гравця i Марківська стратегія у поточному стані s задаватиметься

$$f_{is} : S \to A^{i},$$

де f_{is} - обмежена функція, задана у просторі станів.

У такому разі запис

$$f = (f_1, f_2, ..., f_n),$$

буде комбінацією Марківських стратегій у чистому вигляді, коли гравці самі обирають дію. Тоді Марківська рівновага в чистих стратегіях для будь-якого стану s та для всіх гравців i матиме вигляд

$$\gamma_i(f^*)(s) \geq \gamma_i(f_{-i}^*, f_i^*)(s).$$

Така гра, якщо вона має скінченну кількість гравців та скінченну кількість станів матиме рівновагу за Нешем[7].

2.2 Алгоритм знаходження рівноваги

По суті, знаходження рівноваги у стохастичній грі, де ми маємо справу з декількома станами системи, відрізняється від алгоритму для повторюваних ігор, де система не змінює стан. Але, оскільки множина станів скінченна, і гравці грають Марківські стратегії, гра розбивається на під-ігри, кожна з яких відрізняється одна від одної функцією виплат, остільки вона залежить від поточного стану, у якому знаходиться гра. Отже, якщо для кожного стану знайти рівновагу за Нешем, то в результаті буде отримана рівновага для гри, залежно від стану, в якому гра знаходиться.

Для цього потрібно перенести абстрактну модель стохастичної гри в більш конкретну, не втрачаючи загальності. Отже, модель стохастичної гри буде наступною:

а) Потужність множини гравців буде задаватися користувачем та має бути цілим числом. Сама множина гравців пронумерована відповідно від нуля до m, де m-кількість гравців;

б) Потужність множини станів буде задаватися користувачем та має бути цілим числом. Після цього користувач задає кожен стан як ціле додатне число;

в)Потужність множини доступних дій гравців задаватиметься користувачем та має бути цілим числом. Множина дій містить в собі номери дій, які можуть зіграти гравці від нуля до n, де n- потужність множини дій;

г) Функція для визначення набору дій гравця у стані s задана аналітично та є функцією від двох аргументів

$$A^{i} = f(s,i),$$

де s- поточний стан системи;

i-гравець, для якого визначається множина дій;

A^{i}- скінченна множина дій для гравця i, що є підмножиною $A*$-множини всіх доступних дій;

д) Функція виплат задана аналітично та є обмеженою функцією від двох аргументів

$$R = r\left(s,a\right),$$

де R- виплата, отримана гравцем у результаті дій на цьому кроці;

s- поточний стан системи;

a- набір дій усіх гравців на цьому кроці;

ж) Закон руху системи є функцією від одного аргументу s, яка задана аналітично, та повертає новий стан гри з множини S;

з) Дисконтований критерій β є дійсним числом, яке вводиться користувачем та є спільним для усіх гравців, $\beta \in (0;1)$;

и) Очікувана виплата гравця i є функцією від двох аргументів

$$g = \gamma_i(a, s),$$

де a – профіль стратегій гравців у стані s;

s - поточний стан, у якому знаходиться гра;

g - дійсне число, що позначає очікувану виплату гравця, якщо всі гравці дотримуються стратегій з профілю a.

і) Користувач вводить кількість ітерацій гри, тобто у скільки етапів будуть змінюватись стани впродовж усієї гри. Від етапу гри залежатиме очікувана виплата гравців у стані, в якому гра знаходиться на даному етапі.

Алгоритм полягає в знаходженні максимальної вигоди g для кожного з гравців, що і є рівновагою за Нешем. Знаходження такої рівноваги для кожного стану гри s і буде результатом виконання програми.

Для кожного стану матимемо допоміжну структуру словник, яка у відповідність кожному стану ставитиме множини доступних дій кожного гравця у цьому стані за допомогою функції $f(s,i)$.

На кожній ітерації фіксуватимемо кожен стан s, і після цього фіксуватимемо кожну з можливих дій гравця i, знаходячи для кожної з них таку комбінацію дій інших гравців, за якої очікувана виплата буде максимальною.

Після цього, щоб визначити профіль стратегій, за якого гравцям не вигідно відхилятись від своїх дій, тобто їх очікувана вигода не буде нижчою за ту, яку вони вже мають, потрібно знайти профіль, вигода якого буде максимальною серед тих, які були відібрані на попередньому кроці.

В результаті матимемо для кожної ітерації гри рівновагу за Нешем, в залежності від стану гри s, при чому за рахунок дисконтованого критерія вони будуть відрізнятись одне від одної в залежності від ітерації, навіть якщо система перейде в стан, у якому уже знаходилась до цього.

РОЗДІЛ 3: Впровадження алгоритму пошукурівноваги за Нешем

Отже, сам алгоритм :

for n := 1 **to** *iterations* **do**

for all комбінації стратегій гравців *combinations* зафіксувати

по черзі стратегії гравця, і записати max(*combinations*) для кожної

стратегії в структуру *firstmax*

for all $t \in$*firstmax* **do**

$g(s,t,n)$, записуючи в структуру *gtempmax*

max(gtempmax) і буде результатом, тепер за індексом відновити

профіль стратегій з *firstmax*

Return index_of(max(gtempmax)) in *firstmax*

Чим більша кількість гравців, тим менша швидкість роботи алгоритму. Більшість існуючих алгоритмів розроблені для вирішення ігор з двома гравцями. Алгоритм можна покращити, прибравши деякі проміжні структури, використовуючи іншу мову програмування, ніж Python.

Висновки

В результаті виконаної роботи було створено програму для обчислення рівноваги за Нешем у стохастичних іграх з дисконтованим критерієм на основі теоретичного матеріалу з досліджень як першовідкривачів науки теорії ігор, так і більш сучасних авторів. Застосунок можна покращити, змінивши мову програмування з Python на, наприклад, Ruby, де реалізація має бути більш простою та швидкою. Оскільки теорія ігор відносно молода наука, існуватимуть кращі реалізації алгоритмів для стохастичних ігор, будуть виникати нові підтипи стохастичних ігор, які, можливо, більш точно характеризуватимуть поведінки людей.

Список літератури

1. NashJF (1950) Equilibrium points in n-person games. Proc Natl Acad Sci USA36(1):48–49.

2. Shapley LS (1953) Stochastic games. Proc Natl Acad Sci USA, 39(10):1095–1100.

3.Discounted Stochastic Games Eilon Solan October 26, 1998, 5-6.

4.M. Sobel, Non-Cooperative Stochastic Games, Annals of Mathematical Statistics 42 (1971), 1930–1935.

5.A. Fink, Equilibriumin a Stochastic n-PersonGame, JournalofScienceof

theHiroshimaUniversity 28 (1964), 89–93.

6.P. Rogers, Non-ZeroSumStochasticGames, OperationsResearchCenter

Report 69–8, UniversityofCalifornia, Berkeley, 1969.

7. Pure Strategy Markov Equilibrium in Stochastic Games with Concave Transition Probabilities Subir.K.Chakrabarti Department of Economics Indiana University Purdue University Indianapolis425 University Blvd. Indianapolis, IN 46202, 6-7.