Politechnika Wrocławska

Wydział Matematyki
Kierunek studiów: Matematyka stosowana
Specjalność: –

Praca dyplomowa – inżynierska

ZASTOSOWANIE STOCHASTYCZNEJ OPTYMALIZACJI DO GIER CZĘŚCIOWO OBSERWOWALNYCH

Aleksander Jakóbczyk

słowa kluczowe:
tutaj podajemy najważniejsze słowa kluczowe (łącznie nie powinny być dłuższe niż 150 znaków).

krótkie streszczenie:
Celem rozprawy jest wyznaczenie nieoczekiwanych strategii w grach częściowo obserwowalnych za pomocą metod stochastycznej optymalizacji. Praca będzie opierać się na wynikach Cauwet i Teytauda z 2018 roku, w których przedstawili nieoczekiwane strategie dla kilku klasycznych gier oraz kilku metod optymalizacji. Podjęta zostanie próba odtworzenia oraz rozszerzenia wyników na kolejną grę. Przeprowadzona zostanie analiza porównawcza dla różnych metod optymalizacji.

Opiekun pracy dyplomowej	dr inż. Andrzej Giniewicz	....	....
	Tytuł/stopień naukowy/imię i nazwisko	ocena	podpis

*Do celów archiwalnych pracę dyplomową zakwalifikowano do:**

a) kategorii A (akta wieczyste)

b) kategorii BE 50 (po 50 latach podlegające ekspertyzie)

** niepotrzebne skreślić*

pieczątka wydziałowa

Wrocław, rok 2022

Wrocław University
of Science and Technology

Faculty of Pure and Applied Mathematics
Field of study: Mathematics
Specialty: Theoretical Mathematics

Engineering Thesis

TYTUŁ PRACY DYPLOMOWEJ W JĘZYKU ANGIELSKIM

Aleksander Jakóbczyk

keywords:
tutaj podajemy najważniejsze słowa kluczowe w języku angielskim (łącznie nie powinny być dłuższe niż 150 znaków)

short summary:
Tutaj piszemy krótkie streszczenie pracy w języku angielskim (nie powinno być dłuższe niż 530 znaków).

Supervisor	dr inż. Andrzej Giniewicz	....	....
	Title/degree/name and surname	grade	signature

*For the purposes of archival thesis qualified to:**
- *a) category A (perpetual files)*
- *b) category BE 50 (subject to expertise after 50 years)*

** delete as appropriate*

stamp of the faculty

Wrocław, 2022

Spis treści

Wstęp

We wstępie zapowiadamy, o czym będzie praca. Próbujemy zachęcić czytelnika do dalszej lektury, np. krótko informując, dlaczego wybraliśmy właśnie ten temat i co nas w nim zainteresowało.

Rozdział 1

Rozdział pierwszy

Tabela 1.1 przedstawia przykładową tabelę. Do tworzenia tabeli służą m.in. środowiska `tabular` oraz `table`. Istnieje możliwość numeracji dwustopniowej, gdzie pierwsza cyfra oznacza numer rozdziału, a druga – kolejny numer tabeli w tym rozdziale. Tytuł powinien znajdować się centralnie nad tabelą, 12 pkt odstępu od tekstu zasadniczego nad i pod tabelą wraz z tytułem. Jeśli tabela jest cytowana – należy podać centralnie pod tabelą źródło jej pochodzenia, np. opracowanie własne, opracowano na podstawie danych z GUS.

Tabela 1.1: Podstawowa Tabela

Państwo	PKB (w milionach USD)	Stopa bezrobocia
Stany Zjednoczone	75 278 049	4,60%
Chiny	11 218 281	4,10%
Japonia	4 938 644	3,10%
Niemcy	3 466 639	6,00%
Wielka Brytania	2 629 188	4,60%

Źródło: opracowanie własne

Do cytowania używamy komendy `cite`. W nawiasie klamrowym podajemy klucz, którego użyliśmy w pliku *bibliografia.bib*. Przykład: [**?**] lub [**?**, chap. 2].

1.1 Podrozdział pierwszy

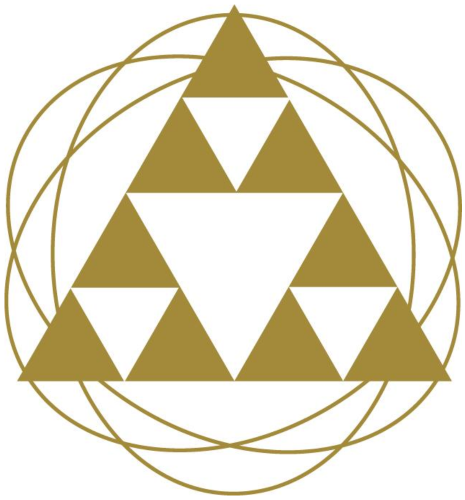
Tabela 1.2: Podstawowa Tabela

Państwo	PKB (w milionach USD)	Stopa bezrobocia
Stany Zjednoczone	75 278 049	4,60%
Chiny	11 218 281	4,10%
Japonia	4 938 644	3,10%
Niemcy	3 466 639	6,00%
Wielka Brytania	2 629 188	4,60%

Źródło: opracowanie własne

1.2 Podrozdział drugi

Rysunki do pracy dyplomowej należy wstawiać w sposób podobny do wstawiania tabel, z zasadniczą różnicą polegającą na tym, że podpis powinno umieszczać się centralnie pod rysunkiem, a nie powyżej niego. Numeracja i sposób cytowania pozostają bez zmian, przy czym tabele i rysunki nie mają numeracji wspólnej, np. po Tabeli 1.2 występuje Rysunek 1.1 (o ile jest to pierwszy rysunek rozdziału pierwszego), a nie Rysunek 1.3.



Rysunek 1.1: Podstawowy Rysunek

Rozdział 2

Definicje, lematy, twierdzenia, przykłady i wnioski

Celem algorytmów, których będziemy wykorzystywać, jest znalezienie optymalnej strategi w grach częściowo obserwowalnych. Zdefiniujmy zatem podstawowe pojęcia potrzebne nam do tego, aby matematycznie opisać czym jest gra i czym jest strategia optymalna. W tym celu wprowadzi kilka podstawowych definicji.

2.1 Typy gier

Podstawową kategorią, na jaką możemy podzielić gry, jest podział ze względu na czas, w którym gracze podejmują decyzje:

Definicja 2.1 (Gra w postaci strategicznej)**.** Jest to typ gry, w której gracze podejmują decyzje w tym samym momencie.

Definicja 2.2 (Gra w postaci ekstensywnej)**.** Jest to typ gry, w której gracze podejmują decyzje we wcześniej ustalonej kolejności.

Przykładami gier w postaci strategicznej są gry papier-kamień-nożyce, oszust czy też mora. Natomiast przykładami gier w postaci ekstensywnej są szachy, warcaby oraz go.

Gry możemy również dzielić ze względu na posiadaną wiedzę.

Definicja 2.3 (Gra z kompletną informacją)**.** Jest to typ gry, w której gracze mają informacje o możliwych przyszłych wynikach gry i o zbiorach możliwych strategii.

Definicja 2.4 (Gra częściowo obserwowalna)**.** Jest to przeciwnie gier z kompletną informacją.

Przykładami gier z kompletną informacją są szachy, warcaby oraz go. Natomiast przykładami gier częściowo obserwowanymi są wszelkie gry posiadające w rozgrywce pewien elementy losowe takie jak rzut kostką czy też dobieranie kart.

Istnieje jeszcze wiele innych podziałów gier ze wglądu na kategorie takie jak liczba graczy, ze względu na zbiory dostępnych akcji, ze względu na możliwość tworzenia koalicji i wiele innych.

2.2 Definicje i Oznaczenia

Wprowadźmy podstawowe oznaczenia potrzebne nam do tego, aby móc zdefiniować czym strategia optymalna.

2.2.1 Strategie proste

- $N = \{1, 2, \dots, n\}$ – zbiór graczy,
- $A_i,\ i \in N$ – niepusty zbiór strategii czystych gracza i,
- $m_i = |A_i|$ – liczba strategi gracza i,
- $A = \prod_{i \in N} A_i$ – zbiór wszystkich strategii gry,
- $u_i : A \to \mathbb{R}$ – funkcja wypłaty gracza i,
- $a = (a_1, a_2, \dots, a_n) = (a_i)_{i \in N},\ a_i \in A_i$ – profil gry w strategiach czystych,
- $u_i(a) = u_i(a, a_{-i})$ – wypłata gracza i z profilu a,
- $a_{-i} = (a_i)_{i \in N \setminus \{i\}},$ – profil wszystkich strategii poza strategią graca i.

Definicja 2.5 (Gra strategiczna)**.** Grą strategiczną nazywamy trójkę $GS = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$.

Definicja 2.6 (Równowaga Nasha w strategiach czystych gry strategicznej)**.** Równowaga Nasha w strategiach czystych gry strategicznej jest to taki profil gry $a^* = (a_1^*, a_2^*, \dots, a_N^*) \in A$, takim, że:

$$\underset{i \in N}{\forall} \underset{a_i \in A_i}{\forall} \quad u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*)$$

Zatem jest to profil gry, w którym istnieje strategia czystej dająca niegorszeń wyniki od dowolnej innej strategii czystej. Okazuje się jednak, że taki stan nie zawsze istnieje w strategiach czystych, np. w grze kamień papier nożyce strategia grania tylko kamienia daje gorsze rezultat przeciwko graniu tylko papieru. Podobnie ze strategią grania tylko nożyc i grania tylko papieru.

2.2.2 Strategie mieszane

Definicja 2.7 (Strategia mieszana)**.** Strategia mieszana σ_i graca i w grze strategicznej $GS = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$. Nazywamy rozkład prawdopodobieństwa na zbiorze strategi czystych A_i:

$$\sigma_i = (\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{im_i})$$

gdzie σ_{ik} oznacza prawdopodobieństwo, że gracz i zagra strategie czysta $k \in A_i$.

Wprowadźmy dodatkowe oznaczenia:

Fakt 2.8. *Strategia czysta jest szczególnym przypadkiem strategij mieszanej, w którym prawdopodobieństwo zagrania jednej z dostępnych strategii wynosi 1.*

Wprowadźmy dodatkowe oznaczenia:

- $\Sigma_i = \left\{ \sigma_i : A_i \to [0, 1], \sum_{k=1}^{n} \sigma_{ik} = 1, \sigma_{ki} \geq 0 \right\}$ – zbiór strategii mieszanych gracza i,
- $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ – profil gry,

- $u_i(\sigma) = u_i(\sigma_i, \sigma_{-i})$ – wypłata gracza i z profilu σ,
- $\sigma_{-i} = (\sigma_i)_{i \in N \setminus \{i\}}$, – profil wszystkich strategii poza strategią graca i.

Definicja 2.9 (Równowaga Nasha w strategiach mieszanej gry strategicznej)**.** Profil gry strategicznej σ_i^* jest Równowagą Nasha gdy:

$$\underset{i \in N}{\forall} \underset{\sigma_i \in \Sigma_i}{\forall} \quad u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*)$$

Równowaga Nasha interpretujemy jako taki profil gry, w którym żaden z graczy nie opłaca się zmieniać swojej strategi, ponieważ nie skutkuje to zwiększeniem swoich zysków.

2.3 Twierdzenia

Algorytmy wykorzystywane w poniższej pracy oparte są o dwa twierdzenia a dokładniej o szczególne przypadki wynikające z nierówności 2.10 i 2.12:

Twierdzenie 2.10 (Nierówność Hoffdinga)**.** *Niech $X_1, X_2, \ldots, X_t$ będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że $a_i \leq X_i \leq b_i$, wtedy:*

$$S_t = \sum_{i=1}^{t} X_i, \quad c_i = b_i - a_i,$$
$$P(|S_t - \mathbb{E}(S_t)| \geq \epsilon) \leq 2 \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^{n} c_i^2}\right)$$

Lemat 2.11. *Niech $X_1, X_2, \ldots, X_t$ będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że $0 \leq X_i \leq 1$, wtedy:*

$$\overline{S}_t = \frac{S_t}{t}, \quad \mu = \mathbb{E}(\overline{S}_t), \quad P(|\overline{S}_t - \mu| \leq \epsilon) = 1 - \delta,$$
$$\epsilon \leq \sqrt{\frac{\ln(2/\delta)}{2t}}$$

Twierdzenie 2.12 (Empiryczna nierówność Bernsteina)**.** *Niech $X_1, X_2, \ldots, X_t$ będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że $a \leq X_i \leq b$, wtedy:*

$$P(|S_n - \mathbb{E}(E_n)| \geq \epsilon) = \delta, \quad \overline{\sigma}_t^2 = \frac{1}{t} \sum_{i=i}^{t} (S_i - \overline{S}_t)^2,$$
$$|S_t - \mathbb{E}(E_t)| \leq \overline{\sigma}_t \sqrt{\frac{2\ln(3/\delta)}{t}} + \frac{3R\ln(3/\delta)}{t}$$

Lemat 2.13. *Niech $X_1, X_2, \ldots, X_t$ będzie ciągiem niezależnych zmiennych losowych (i.i.d.) takim, że $0 \leq X_i \leq 1$, wtedy:*

$$P(|\overline{S}_t - \mu| \leq \epsilon) = 1 - \delta,$$
$$\epsilon \leq \overline{\sigma}_t \sqrt{\frac{2\ln(3/\delta)}{t}} + \frac{3\ln(3/\delta)}{t}$$

2.4 Problem porównań wielokrotnych

Załóżmy, że z prawdopodobieństwem $1-\delta$ chcemy wiedzieć który z 2 graczy, p_1 i p_2 jest lepszy. W tym celu będziemy przeprowadzać testy statystyczne, dla których prawdopodobieństwo pomyłki k-tego testu wynosi δ_k, aż do momentu, gdy jeden z graczy wygra przeważającą ilość razy. Wtedy po przeprowadzeniu n takich testów

$$P(\text{Chociaż jeden z } n \text{ testów się pomylił}) \overset{(*)}{\leq} \sum_{k=1}^{n} \delta_k \implies$$

$$P(\text{Żaden test się nie pomylił}) \leq 1 - \sum_{k=1}^{n} \delta_k$$

$(*) \quad P(X+Y) \leq P(X) + P(Y)$

Oznacza to, że musimy wprowadzić pewną korektę, aby ostateczne prawdopodobieństwo popełnienia błędu było mniejsze niż δ. Możemy wprowadzić jedna z dwóch poprawek:

2.4.1 Niech n będzie maksymalną liczbą testów jaką pozwalamy wykonać, aby wyznaczyć lepszego gracza. Wtedy $\delta_k = \frac{\delta}{n}$.

2.4.2 Niech δ_k spełnia nierówność $\delta \geq \sum_{k=1}^{\infty} \delta_k$. Wtedy niezależnie od ilości przeprowadzonych testów, ostateczne prawdopodobieństwo pomyłki będzie nie większe niż δ.

2.5 Algorytmy wyścigowe

Jednymi z algorytmów stosującymi dane korekty są algorytmy wyścigowe (Racing Algorytms). Dwoma najpopularniejszymi typami algorytmów racingowych są "Hoeffding raceóraz "Bernstein race". Oparte są one odpowiednio o Twierdzenie 2.10 i Twierdzenie 2.13. Mają one jednak pewną wadę, mogą zajmować one dużą ilość czasu, a w momencie, gdy poziom umiejętności porównywanych graczy jest sobie równy (prawdopodobieństwo wygranej wynosi 50%), wtedy z prawdopodobieństwem równym $1-\delta$ algorytm się nigdy nie zatrzyma.

W celu poradzenia sobie z problemami, jakie wiążą się z klasycznymi algorytmami wyścigowymi, wprowadzimy tzw. Limited Racing algorytm. Załóżmy zatem dodatkowy warunek, który mówi, że przerywamy działanie algorytmu w momencie, gdy empiryczna wartość oczekiwana z prawdopodobieństwem większym bądź równym $1-\delta$ jest znana z dokładnością co do zadanego ϵ.

2.5.1 Bernstein race without maximum race length

Wykorzystując Lemat 2.13 oraz korektę 2.4.2 z $\delta_k = \frac{c\delta}{k^2}, c = \frac{6}{\pi^2}$ otrzymujemy, że dla ciągu $X_1, X_2, \ldots, X_t$ i.i.d. takim, że $0 \leq X_i \leq 1$:

$$\epsilon_{t,k} \leq \overline{\sigma}_t \sqrt{\frac{2\ln(3/\delta_k)}{t}} + \frac{3\ln(3/\delta_k)}{t} = \overline{\sigma}_t \sqrt{\frac{2\ln(\frac{k^2\pi^2}{2\delta})}{t}} + \frac{3\ln(\frac{k^2\pi^2}{2\delta})}{t}$$

Wtedy $e_{t,k}$ interpretujemy jako maksymalną różnicę między empiryczną a teoretyczną wartością oczekiwaną po przeprowadzeniu k testów i rozegraniu t gier, z prawdopodobieństwem pomyłki równym

Fakt 2.14. *Przeprowadzać testy po każdej rozegranej grze ($k = t$)* $\lim_{t\to\infty} e_{t,k} = 0$

Dowód Faktu 2.14.

$$0 \le X_i \le 1 \implies \overline{\sigma}_t^2 \le \frac{1}{2}$$
$$\epsilon_{t,t} = \epsilon_t \le \sqrt{\frac{\ln(\frac{t^2\pi^2}{2\delta})}{t}} + \frac{3\ln(\frac{t^2\pi^2}{2\delta})}{t} \le \frac{4\ln(\frac{t^2\pi^2}{2\delta})}{t}$$
$$\forall_{a>0} \lim_{t\to\infty} \frac{\ln(at^2)}{t} \overset{[\frac{\infty}{\infty}]}{=} \lim_{t\to\infty} \frac{2}{t} = 0 \implies \lim_{t\to\infty} e_t \le 0$$

Zatem $0 \le \lim_{t\to\infty} e_t \le 0 \implies \lim_{t\to\infty} e_t = 0$ $\square$

Podsumowanie

Podsumowanie w pracach matematycznych nie jest obligatoryjne. Warto jednak na zakończenie krótko napisać, co udało nam się zrobić w pracy, a czasem także o tym, czego nie udało się zrobić.

Dodatek

Dodatek w pracach matematycznych również nie jest wymagany. Można w nim przedstawić np. jakiś dłuższy dowód, który z pewnych przyczyn pominęliśmy we właściwej części pracy lub (np. w przypadku prac statystycznych) umieścić dane, które analizowaliśmy.