



UNIwersytet
Warszawski

Wydział Matematyki, Informatyki i Mechaniki
Instytut Matematyki Stosowanej i Mechaniki

prof. dr hab. Urszula Forýś
kierownik Zakładu Biomatematyki i Teorii Gier
Instytut Matematyki Stosowanej i Mechaniki
Wydział Matematyki, Informatyki i Mechaniki
Uniwersytetu Warszawskiego
Banacha 2, 02-097 Warszawa

Święcice, 13.11.2024

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy: Zastosowanie kawałkami deterministycznych procesów Markowa w modelach stochastycznej ekspresji genów

Autor rozprawy: mgr Agnieszka Kozdęba

Promotor rozprawy: prof. dr hab. Ryszard Rudnicki

Recenzja została sporządzona w związku z powołaniem mnie przez Radę Dyscypliny Naukowej Matematyka Uniwersytetu Jagiellońskiego w dniu 26.09.2024 roku do pełnienia funkcji recenzenta w postępowaniu o nadanie stopnia naukowego doktora nauk ścisłych i przyrodniczych w dyscyplinie MATEMATYKA pani mgr Agnieszce Kozdębie. Celem recenzji jest ocena spełnienia przez rozprawę doktorską mgr Agnieszki Kozdęby warunków określonych w art. 187. pkt 2. Ustawy z dn. 20 lipca 2018 Prawo o szkolnictwie wyższym i nauce, jak również w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 65, poz. 595, z późniejszymi zmianami).

Rozprawa doktorska „Zastosowanie kawałkami deterministycznych procesów Markowa w modelach stochastycznej ekspresji genów” dotyczy matematycznego modelowania ekspresji genów przy użyciu – jak wskazuje temat – kawałkami deterministycznych procesów Markowa. Część wyników prezentowanych w rozprawie została opublikowana w artykule:

- A. Kozdęba and A. Tamski, Application of the Goodwin model to autoregulatory feedback for stochastic gene expression, *Mathematical Biosciences* **327** (2020): 108413;

natomiast kod użyty do symulacji jest dostępny online:

- A. Kozdęba and A. Tamski, Goodwin model simulations, https://github.com/alnilam6/NEG_feedback, 2019.

1. Tematyka rozprawy doktorskiej

Tematyka rozprawy doktorskiej została wyraźnie uwidoczniiona w tytule tej rozprawy. Autorka połączyła idee deterministycznego modelowania ekspresji genów (model Goodwina) z modelowaniem stochastycznym. W rezultacie ekspresja genów jest opisana za pomocą PDMC, czyli kawałka-

mi deterministycznych procesów Markowa. Oba podejścia są dobrze znane, natomiast połączenie obu typów modelowania (w kontekście nieliniowej dynamiki deterministycznej) jest podejściem względnie nowym – praca doktorantki (z A. Tomskim) ukazała się w 2020 roku. Sam proces ekspresji genów jest jednym z podstawowych procesów zachodzących we wszystkich organizmach, stąd nie da się przecenić badań prowadzonych w tym kierunku. Nowe podejście pozwoliło na wskazanie możliwości całkowitej inhibicji tego procesu w przypadku dodatniego sprzężenia zwrotnego, podczas gdy poprzednie prace wskazywały na istnienie rozkładu niezmienniczego.

2. Teza i cele badawcze

Niestety, choć rozprawa doktorska dotyczy konkretnych zastosowań matematyki i w takim przypadku istotne wydaje się postawienie konkretnej hipotezy badawczej na początku pracy oraz konsekwentne dążenie do jej zweryfikowania w poszczególnych rozdziałach, to Autorka takiej hipotezy nie postawiła. Zasadnicze cele pracy są realizowane w dwóch głównych rozdziałach, gdzie zostały zbadane dwie wersje modelu PDMP z różnego typu sprzężeniem zwrotnym. Można zatem stwierdzić, że celem pracy było wykazanie, że połączenie modelu Goodwina ze stochastyczną ekspresją genów może prowadzić do różnego typu asymptotyki w powstałym PDMP, w zależności od sprzężenia zwrotnego, tzn. w zależności od tego, czy jest to sprzężenie dodatnie, czy ujemne.

W mojej ocenie, choć cele badawcze nie zostały wyraźnie zaakcentowane w pracy, to są one ważne, rozważana tematyka jest interesująca i aktualna, podobnie jak zastosowana metodologia, tym samym więc spełniają wymagania stawiane przed pracami doktorskimi.

3. Organizacja i strona redakcyjna rozprawy

Recenzowana rozprawa doktorska liczy łącznie 64 strony, wraz ze spisem treści, rysunków oraz streszczeniem i została podzielona na 4 rozdziały, oprócz wstępu i zakończenia. W pracy zawarto bibliografię, która liczy 75 pozycji. Wśród pozycji przedstawionych w tym rozdziale 2 stanowią publikacje Autorki (artykuły i kod do symulacji).

„Wstęp” zawiera nakreślenie tematyki, zarówno w kontekście PDMP, jak i ekspresji genów, jak również prezentuje zawartość poszczególnych rozdziałów.

Rozdział 1, „Kawałkami deterministyczne procesy Markowa”, składa się z czterech podrozdziałów. Pierwszy stanowi wprowadzenie w świat procesów Markowa, a drugi – w świat procesów kawałkami deterministycznych. W następnym pojawiają się pojęcia i szereg twierdzeń dotyczących półgrup stochastycznych, a na koniec poznajemy podstawy zastosowania PDMP do opisu stochastycznej ekspresji genów. Omówione zostały dwa modele samoregulującej ekspresji genów z zastosowaniem PDMP, gdzie układ dynamiczny w części deterministycznej jest liniowy.

W rozdziale 2, „Model stochastycznej ekspresji genów z ujemnym sprzężeniem zwrotnym”, mamy ponownie cztery podrozdziały. W pierwszym z nich opisany został proces biologiczny wraz ze schematem obrazującym pętlę sprzężenia zwrotnego oraz wykresem malejącej funkcji Hilla, która za to sprzężenie odpowiada. Kolejny podrozdział zawiera opis modelu, analizę dwóch podmodeli deterministycznych, następnie powiązanie tej analizy w całość jako PDMP i zbadanie asymptotyki modelu PDMP, tzn. udowodnienie istnienia gęstości niezmienniczej z wykorzystaniem twierdzeń zaprezentowanych w poprzednim rozdziale. Tutaj Autorka wykazała się pomysłowością w konstruowaniu odpowiednich zbiorów niezmienniczych. Podrozdział trzeci dotyczy badania tzw. granicy adiabatycznej i zbadania możliwości istnienia cyklu granicznego. Okazało się, że w tym przypadku zachowania cykliczne nie są możliwe. W ostatnim podrozdziale widzimy ilustracje numeryczne omawianego modelu z ujemnym sprzężeniem zwrotnym.

Rozdział 3, „Model stochastycznej ekspresji genów z dodatnim sprzężeniem zwrotnym”, zawiera trzy podrozdziały. W pierwszym z nich ponownie mamy opis biologiczny z wykresem rosnącej funkcji Hilla. W kolejnym podrozdziale, analogicznie jak poprzednio, wprowadzony został model i omówiona jego asymptotyka, która różni się zasadniczo od tej dla sprzężenia ujemnego. Badany PDMP jest tym razem wymiatający. W trzecim podrozdziale Autorka przedstawiła symulacje nume-

ryczne dla modelu z dodatnim sprzężeniem oraz porównanie modeli z omawianymi dwoma typami sprzężeń.

Rozdział 4, „Sieci regulatorowe”, zawierający dwa podrozdziały – jeden na temat interakcji pomiędzy dwoma genami z pominięciem transkrypcji, a drugi z jej uwzględnieniem, stanowi przyczynek do zastosowania PDMP do opisu sytuacji bardziej skomplikowanej, gdzie bierzemy pod uwagę co najmniej dwa geny stanowiące sieć regulatorową.

W „Zakończeniu” autorka krótko podsumowała otrzymane wyniki oraz wskazała kolejną ścieżkę dalszego uogólniania analizowanego podejścia, np. dla modeli na bazie trójwymiarowych układów deterministycznych.

Rozprawa została w całości przygotowana po polsku, ale zawiera także „Abstract” w języku angielskim. Od strony edytorskiej mam kilka uwag. Po pierwsze, można znaleźć drobne błędy edytorskie, jak np. pozostawienie pojedynczej litery (o, w itp.) na końcu linii, zastosowanie kreski zamiast myślnika (myślnik w języku polskim tworzymy w kodach Tex-a za pomocą trzech kresek) czy zapis postaci rysunek (3), gdy numery rysunków nie są objęte nawiasami, więc powinno być rysunek 3, itp. Znalazłam także kilka błędów interpunkcyjnych. Mam też wątpliwości odnośnie użytej nazwy „malejąca funkcja Hilla” w przypadku modelu z ujemnym sprzężeniem zwrotnym (raczej zapisałabym tę funkcję jako $1 - H(x)$, gdzie H to funkcja Hilla), jak też nazwy „semipotok” (raczej półpotok) czy też „macierz linearyzacji” (klasycznie mówimy o macierzy Jacobiego). W niektórych fragmentach pracy brakuje także dostatecznych odnośników do bibliografii, przez co trudno się zorientować, z których podręczników pochodzą poszczególne pojęcia, twierdzenia, w jakich artykułach coś się pojawiło najwcześniej itp.

4. Osiągnięcia naukowe

Główne osiągnięcia naukowe, na podstawie których wnioskuję poniżej o dopuszczenie do obrotu, to:

- Zbudowanie modelu ekspresji genów w oparciu o PDMP, łączącego podejście deterministyczne (model Goodwina) z przełączeniami stochastycznymi.
- Udowodnienie twierdzenia dotyczącego asymptotycznej stabilności dla modelu z ujemnym sprzężeniem zwrotnym.
- Udowodnienie twierdzenia dotyczącego wymiatania dla modelu z dodatnim twierdzeniem zwrotnym.
- Zaproponowanie kodu do numerycznego rozwiązywania modeli PDMP, w oparciu o który wyniki analityczne zostały zilustrowane numerycznie.
- Wskazanie możliwości wykorzystania PDMP do opisu sieci regulatorowych.

5. Uwagi krytyczne i dyskusyjne

Powyżej przedstawiłam już uwagi krytyczne dotyczące braku wyrażnie sformułowanej hipotezy, jak również redakcyjnej strony rozprawy. Poniżej przedstawiam bardziej szczegółowe uwagi dotyczące zawartości pracy.

Uwagi szczegółowe:

- Na końcu podrozdziału 1.1 przedstawione zostało pojęcie mocnego procesu Markowa, które dalej w pracy się nie pojawia, więc nie bardzo rozumiem, po co zostało przedstawione.
- Na dole str. 11 pojawia się tekst: „będący jednorodnym, kawałkami deterministycznym procesem Markowa.”, jednak nie zostało powiedziane, co to jest proces jednorodny, ani czy są niejednorodne PDMP.
- Na str. 12 po definicji 1.7 czytamy: „Dla półgrup generowanych przez procesy stochastyczne asymptotyczna stabilność będzie oznaczać, że gęstości jednowymiarowych rozkładów danego procesu zbiegają do jedynej gęstości niezmienniczej”, ale stwierdzenie to nie zostało dokładniej umotywowane. W szczególności – skąd wiadomo, że istnieje taka półgrupa, jak również z czego wynika, że jest jedyna gęstość niezmiennicza?

- Na str. 14 po wzorze (1.17) czytamy: „W szczególności zbiory X_n są niezmiennicze względem półgrupy $\{P(t)\}$ ” – nie bardzo wiadomo, o jaką półgrupę chodzi. Wydaje się, że ta półgrupa nie została wcześniej wprowadzona. Dalej mamy „w przypadku gdy półgrupa $\{P(t)\}$ powiązana jest z układem dynamicznym z losowymi przełączeniami.” – w jaki sposób powiązana?
- Nie rozumiem zapisu definiującego τ w definicji 1.9.
- W dolnej części str. 16 mamy: „Można więc ograniczyć wartości $x(t)$ do przedziału $[0,1]$.” Tutaj przydałoby się w sposób jawny wskazać, że $[0,1]$ jest zbiorem niezmienniczym i udowodnić ten fakt.
- Na str. 17 czytamy: „Wówczas funkcje $u_i(t,x)$ spełniają następujący układ Fokkera-Planka” – nie bardzo wiadomo, co to jest ten układ Fokkera-Planka, ani dlaczego te funkcje go spełniają.
- Przejście str. 17/18: „Dzięki temu założeniu mamy pewność, że trajektorie procesu startujące z dowolnego punktu trafią do zbioru niezmienniczego” – dlaczego tak będzie? Czy q_0, q_1 występujące w tym założeniu mają na zbiór niezmienniczy jakiś wpływ?
- Na str. 21 pojawia się wzór malejącej funkcji Hilla i założenie, że parametr Hilla $n > 0$. Byłabym ostrożna z wyborem wartości n z przedziału $(0,1)$, gdyż wtedy funkcja Hilla nie ma w zerze nieskończoną pochodną i są kłopoty z jednoznacznością rozwiązań.
- Po wzorze (2.6) czytamy: „Jak łatwo zauważyć, jest on globalnie asymptotycznie stabilny” – należało to wykazać (tym bardziej, że faktycznie łatwo to zrobić...)
- Na str. 23 (i później dla kolejnych układów) obliczane są wartości własne macierzy Jacobiego w celu określenia stabilności stanu stacjonarnego. Jest to jednak całkiem zbędne, gdyż w przypadku modelu dwuwymiarowego mamy wygodniejsze narzędzia – są nim wyznacznik $\det(J)$ i ślad macierzy $\text{tr}(J)$ Jacobiego. Ponieważ są to jednocześnie współczynniki wielomianu charakterystycznego (w przypadku śladu z dokładnością do znaku), to łatwo podać warunku asymptotycznej stabilności – mianowicie $\det(J) > 0$ jest warunkiem koniecznym, a w połączeniu z $\text{tr}(J) < 0$ daje warunek wystarczający.
- Poczynając od rysunku 6, Autorka przedstawia szereg przykładowych portretów fazowych. Nie zostało jednak wyjaśnione, jak te portrety zostały przygotowane.
- Na str. 26 pod rysunkiem 6 czytamy: „trajektorie... będą wchodzić do zbioru...” – należało to udowodnić.
- Na dole str. 26 mamy stwierdzenie, że „Ponieważ forma h jest ujemnie określona”, to można przyjąć maksimum na danym zbiorze równe $-c_1$. Zdecydowanie jest tu potrzebny komentarz, dlaczego tak można, bo w ogólnym przypadku ujemnie określona forma kwadratowa nie jest oddzielona od zera.
- Na górze str. 28 czytamy: „Skoro intensywności przejścia będą dążyć do nieskończoności” – nie jest dla mnie jasne, dlaczego tak ma być.
- Nad wzorem (2.17) mamy „proces (x_i, i_t) , gdzie ξ_i ma wartości...” – rozumiem, że powinno być nie ξ_i tylko x_i .
- Skąd wynika ostatnia równość we wzorze (2.18). Dodatkowo – wcześniej mieliśmy q_{01}, q_{10} , a przy wprowadzaniu tego wzoru pojawiły się q_0 i q_1 . Czy to te same wartości?
- Na górze str. 39 mamy implikację dotyczącą zbiorów niezmienniczych i półgrupy wymiatającej. Czemu z istnienia tych zbiorów niezmienniczych wynika wymiatanie?
- Na górze str. 40 mamy „dla dostatecznie małych y spełniony jest warunek” – można łatwo określić, dla jakich y ten warunek będzie spełniony.

6. Konkluzja

W mojej ocenie rozprawa doktorska mgr Agnieszki Kozdęby, mimo pewnych braków, zawiera wartościowe wyniki, a ponadto wykazuje potencjał rozwojowy, gdyż stosując analogiczne metody można badać znacznie bardziej skomplikowane procesy, w szczególności sieci regulatorowe, co sama Autorka pokreśliła w ostatnim rozdziale rozprawy.

Autorka opublikowała część wyników w renomowanym czasopiśmie, więc spełniła warunek konieczny w tym kontekście.

W związku z powyższym stwierdzam, że rozprawa ta spełnia warunki określone w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 65, poz. 595, z późniejszymi zmianami) i może być przedmiotem publicznej obrony w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie matematyka.

A handwritten signature in blue ink, appearing to read 'U. Forys', with a stylized flourish above it.

Urszula Forys