

Recenzja rozprawy doktorskiej Andrzeja Tomskiego „Kawałkami deterministyczne procesy Markowa w wybranych modelach biologicznych”

Rozprawa poświęcona jest matematycznym modelom procesów biochemicznych zachodzących w komórkach, związanych z inhibicją enzymów oraz ekspresją i regulacją genów. Procesy te, składające się z szeregu reakcji biochemicznych, są zazwyczaj modelowane za pomocą układów równań różniczkowych zwyczajnych (równań kinetyki chemicznej) opisujących ewolucję w czasie koncentracji substratów i produktów. W przypadku małych komórek i małej liczby substratów i produktów należy przejść od opisu deterministycznego do opisu stochastycznego, na przykład przy pomocy skokowych procesów stochastycznych opisujących zmianę w czasie liczby bio-molekuł danego typu. Dla pewnego zakresu parametrów reakcji odpowiednie jest podejście hybrydowe: produkcja i degradacja cząsteczek opisywana jest równaniami różniczkowymi, natomiast zmiana stanu genu lub inhibitora (z aktywnego na nieaktywny lub odwrotnie) skokowym w czasie procesem typu Poissona. Mamy tutaj do czynienia z kawałkami deterministyczną ewolucją układu, który w losowych momentach czasu zmienia swój stan i ewoluuje zgodnie ze zmienionym układem równań różniczkowych zwyczajnych.

Rozprawa Andrzeja Tomskiego jest rozprawą interdyscyplinarną. Do opisu inhibicji enzymów oraz regulacji genów autor proponuje nowe modele kawałkami deterministycznych procesów Markowa. Bada ich asymptotykę czasową. Udowadnia, że odpowiednia półgrupa Markowa jest asymptotycznie stabilna, to znaczy istnieje gęstość niezmiennicza, do której zbiega w czasie rozkład prawdopodobieństwa oraz wyznacza nośnik rozkładu stacjonarnego. Pod względem matematycznym, analizowane przez autora procesy Markowa różnią się od dotychczasowo badanych w dwóch aspektach: w modelach inhibicji, część deterministyczna ewolucji nie jest różniczkowalna, w modelach regulacji genów rozważany jest układ trzech równań różniczkowych (dotychczas badano układy dwuwymiarowe). Dowody standardowych twierdzeń wymagały od autora nowego podejścia i technik dowodzenia. Twierdzenia o asymptotycznej stabilności półgrup Markowa (Twierdzenia 4-6) są głównymi wynikami matematycznymi rozprawy.

Następnie autor przeprowadza symulacje procesów stochastycznych algorytmem Gillespiego. To jest część interesująca biologów. Możemy zobaczyć czasowe trajektorie poziomu inhibitorów i enzymów oraz cząsteczek pre-mRNA, mRNA i białek, oraz estymację rozkładów stacjonarnych. W trójfazowym (trójwymiarowym) modelu ekspresji genów z auto-aktywacją autor pokazał, że dla odpowiedniego doboru parametrów reakcji biochemicznych, układ staje się bistabilny (stacjonarny rozkład prawdopodobieństwa jest bimodalny) a dla modelu z auto-represją pojawia się cykl graniczny. Jak wiemy, cykl graniczny nie może się pojawić w układach dwuwymiarowych. Wykazanie jego istnienia w modelu trójwymiarowym jest bardzo ważnym wynikiem, niezwykle interesującym z punktu widzenia biologii.

Dodatkowo, w celu zbadania wpływu czynników transkrypcyjnych (czynników pod wpływem których gen przełącza się pomiędzy stanem aktywnym i nieaktywnym) autor rozprawy (ze współpracownikami) stworzył przydatne narzędzia bioinformatyczne.

Tak więc widzimy, że rozprawa doktorska Andrzeja Tomskiego zawiera ściśle powiązane ze sobą części matematyczne, biologiczne i informatyczne. Każda z nich z osobna zasługuje na uznanie a razem tworzą rozprawę, która zdecydowanie spełnia wymogi stawiane rozprawom doktorskim z matematyki.

Swoje uwagi krytyczne chciałbym podzielić na te odnoszące się do metodologii i konstrukcji modeli i na uwagi redakcyjne.

Uwagi metodologiczne

1. Kawałkami deterministyczne procesy Markowa (PDMP) są modelami przybliżonymi całkowicie stochastycznych modeli, gdzie wszystkie reakcje opisuje się skokowymi procesami Markowa. W PDMP przyjmuje się, że dla układów z odpowiednio dużą liczbą cząstek, można zaniedbać efekty stochastyczne związane z ich produkcją i degradacją. Stosowalność PDMP w opisie konkretnych procesów biologicznych zależy od parametrów modelu. Autor nigdzie o tym nie pisze. W tym kontekście ciekawe byłoby sprawdzenie czy w pełnych stochastycznych układach trójwymiarowych nadal mogą wystąpić bistabilności i cykle graniczne.
2. Nie możemy sprawdzać poprawności teoretycznych wyników (jak to w wielu miejscach pisze autor) czyli prawdziwości udowodnionych twierdzeń symulacjami komputerowymi. Możemy natomiast przeprowadzić symulację komputerową pełnego układu stochastycznego i sprawdzić w ten sposób poprawność przybliżenia PDMP.
3. W opisie modelu inhibicji enzymów, autor pisze na stronie 19-tej, że inhibitor przyłącza się do enzymu (patrz również Rys. 4) a następnie na stronie 20-tej, że stężenie substancji zewnętrznej gra rolę inhibitora i wpływa negatywnie na tempo produkcji enzymu. To są dwa różne mechanizmy biologiczne i dwa różne modele matematyczne. Jest to mylące i nie wiemy co tak naprawdę modeluje autor. Następnie autor pisze, że inhibitor przechodzi w sposób losowy między stanem aktywności i nieaktywności. Autor nie objaśnia jakie są biologiczne podstawy takiego przełączania i dlaczego intensywności przełączeń zależą od koncentracji zarówno inhibitorów jak i enzymów. W (20) składnik degradacyjny enzymu zależy od koncentracji inhibitorów ale już nie od koncentracji enzymów, dlaczego w (20) nie mamy wyrażenia typu $x_1 * x_2$? Nawiasem mówiąc nie mamy tu inhibitorów aktywnych lub nieaktywnych tylko gen produkujący inhibitory jest aktywny albo nieaktywny. Model (19) jest bardzo niejasno opisany.
4. W (25) autor wprowadza ale nie wyprowadza równania, które nazywa równaniami Fokkera-Plancka. Równania FP zawierają tradycyjnie drugie pochodne. Przydałaby się tutaj jakaś dyskusja.
5. Symulacje w III.4 przeprowadzone są tylko dla jednego zestawu stałych intensywności przełączania. To jest bardzo specyficzny przykład modelu bez jakiegokolwiek samoregulacji, w którym nota bene można policzyć wszystkie momenty

rozkładu stacjonarnego dla pełnego modelu stochastycznego. Szkoda, że autor nie wykonał bardziej systematycznych symulacji.


6. W modelu ekspresji genu, dla stałych intensywności przełączania genu można policzyć ściśle wszystkie momenty rozkładu stacjonarnego, nie tylko wartości oczekiwane, o których na stronie 41 pisze autor, nie precyzuje jednak jak to można zrobić. Autor twierdzi na stronie 42, że wartości oczekiwane oscylują w czasie, ale chyba nie dla stałych intensywności przełączania?
7. Cykl graniczny pojawia się dla modelu z represją (Rys. 27). Czy dla jego pojawienia się konieczna jest 10-ta potęga koncentracji białka w funkcji intensywności wyłączania genu?

Uwagi redakcyjne

1. Na stronie 8, punkt w przestrzeni d -wymiarowej ma $d+1$ współrzędnych, warunek początkowy został tam wklejony.
2. Na stronie 10 mamy „funkcja intensywności przejścia i funkcja prawdopodobieństwa przejścia”, czy w tym drugim przypadku chodzi o rozkład czasów oczekiwania?
3. W (14-15) zamiast μ powinno być h .
4. W Twierdzeniu 3 pojawia się tajemnicze C .
5. Na stronie 16 (i w innych miejscach też) x_1 i x_2 zamieniają się nagle w x i y .
6. Na Rys. 2 występują nieopisane E^+ i E^- .
7. Na stronie 19, metody z [8] nie mogą być zastosowane w [10], chyba że odwrócimy strzałkę czasu.
8. W (19) zamiast γ pojawia się β .
9. Na Rys. 15, rozkład brzegowy to chyba raczej rozkład stacjonarny i wtedy nie powinniśmy mieć t w argumentach ?
10. Na Rys. 16, strzałka autoregulacyjna oznacza aktywację, brak strzałki represyjnej.

Uważam, że wyniki uzyskane w rozprawie są bardzo ważne, zawiera ona oryginalne rozwiązanie kilku matematycznych problemów. Wyniki są bardzo interesujące z biologicznego punktu widzenia. Zostały zamieszczone w 4 publikacjach. Rozprawa wypełnia wszelkie wymogi stawiane rozprawom doktorskim z matematyki. Wnioskuje o dopuszczenie mgra Andrzeja Tomskiego do dalszych etapów przewodu doktorskiego.

prof. dr hab. Jacek Miękisz
Instytut Matematyki Stosowanej i Mechaniki
Uniwersytet Warszawski


Warszawa, 2 grudnia 2015

