

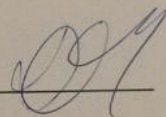
**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики
Кафедра обчислювальної математики

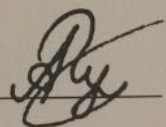
Кваліфікаційна робота
на здобуття ступеня магістра
за спеціальністю 113 Прикладна математика
на тему:

**Об'єднання якісного та кількісного моделювання складної
динаміки за допомогою штучного інтелекту на прикладі автоматичного
моделювання патернів будови органів**

Виконав студент 2-го курсу магістратури
Чергикало Денис Олександрович

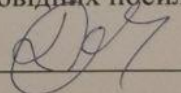


Науковий керівник:
кандидат фізико-математичних наук
Гуляницький Андрій Леонідович



Засвідчую, що в цій роботі немає запозичень з
праць інших авторів без відповідних посилань.

Студент

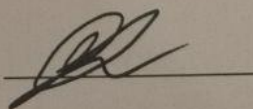


Роботу розглянуто й допущено до захисту на
засіданні кафедри обчислювальної математики

5 травня 2023 р., протокол № 7

Завідувач кафедри

проф. Сергій Ляшко



Київ - 2023

РЕФЕРАТ

Обсяг роботи 62 сторінки, 14 джерел посилання, 8 ілюстрацій.

Ключові слова:

ДИФЕРЕНЦІЙОВАНІ СИСТЕМИ З САМООРГАНІЗАЦІЄЮ, СТОХАСТИЧНІ ДИФЕРЕНЦІЙНІ РІВНЯННЯ, НЕЛІНІЙНІ ДИФЕРЕНЦІЙНІ РІВНЯННЯ З ЧАСТИННИМИ ПОХІДНИМИ, МУЛЬТИКОМПОНЕНТНІ ДИФЕРЕНЦІЙНІ РІВНЯННЯ З ЧАСТИННИМИ ПОХІДНИМИ, КЕРУВАННЯ РОЗПОДІЛЕНОЮ СИСТЕМОЮ, РІВНЯННЯ РЕАКЦІЇ-ДИФУЗІЇ, ТЕОРІЯ СТІЙКОСТІ, БАГАТО-МАСШТАБНЕ МОДЕЛЮВАННЯ.

Об'єктом роботи є задача оптимізації параметрів стохастичного нелінійного мультикомпонентного диференційного рівняння з частковими похідними.

Предмет застосування є багато-масштабне моделювання в біології, автоматизований пошук моделей утворення складних біологічних патернів.

Мета роботи. Побудова нових алгоритмів, методів та загальної методології моделювання та оптимізації складних біологічних систем.

Результати роботи. Була побудована модифікація моделі реакції-дифузії що є більш оптимальною ніж інші. На її основі було побудована модель розвитку патернів шкіри *Ictidomys tridecemlineatus*, що є, на думку автора, найбільш якісною моделлю для цього патерну з тих що є.

Апробація результатів. Результати роботи було представлено на міжнародній науково-практичній конференції "Сучасні проблеми математики та її застосування у природничих науках та інформаційних технологіях"(12-13 травня 2023 р.).

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАЧЕННЯ.....	5
ВСТУП.....	6
РОЗДІЛ I Теоретичні відомості.....	9
1.1 Стохастичне нелінійне мультикомпонентне ДРЧП	9
1.2 Диференційована система з самоорганізацією.....	14
1.3 Основні підходи математичної біології на даний час.....	15
РОЗДІЛ II Постановка задачі	18
2.1 Формалізація задачі пошуку систем з самоорганізацією.....	18
2.1.1 Ріст домену.....	18
2.1.2 Обмеження на функцію u	18
2.1.3 Використання фільтрів “уваги”.....	19
2.1.4 Оцінка стійкості організму.....	19
2.2 Зв'язок з теорією оптимального керування.....	20
2.3 Найважливіші підзадачі та способи їх розв'язку.....	22
2.4 Пошук області стійких рішень.....	24
2.5 Обмеження на оператор та їх імплементація.....	28
2.6 Обмеження на функцію та їх імплементація.....	29
2.7 Модифікації градієнту.....	29
2.8 Нестрогі та строгі обмеження на коефіцієнти.....	32
2.9 Проблема неопуклості цільової функції та малої кількості початкових даних.....	34
2.9.1 Отримання якісних характеристик.....	34
2.9.2 Вирішення проблеми неоднозначності рішень.....	37
РОЗДІЛ III Опис методів та алгоритмів.....	39

3.1 Теорія QQISA(Кількісно-Якісного інтегрованого аналізу стійкості)	39
3.2 Можливості застосування теорії QQISA як загальної теорії що відповідає на потреби багато-масштабного моделювання в біології та інших складних систем.....	42
3.2.1 Управління неправильно поставленими проблемами.....	42
3.2.2 Виявлення відсутньої інформації, створення сурогатних моделей.....	44
3.2.3 Дискретизація простору та часу.....	44
3.2.4 Перекриття масштабів, розуміння виникнення функції.....	44
3.2.5 Кількісна невизначеність.....	45
3.2.6 Вивчення складних просторів структур, підвищення точності та відтворюваності.....	46
3.2.7 З'ясування механізмів.....	46
3.2.8 Використання біологічно натхненного навчання.....	47
3.2.9 Запобігання перенавчанню, мінімізація зміщення даних.....	47
3.3 Отримання одночасно більш простих і більш якісних моделей.....	47
РОЗДІЛ IV Практична частина	49
4.1 Дослідження дії оператора.....	49
4.2 Моделювання патернів шкіри.....	54
ВИСНОВОК	62
ВИКОРИСТАНА ЛІТЕРАТУРА	63

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАЧЕННЯ

ДРЧП - диференційне рівняння в частинних похідних

СНМДРЧП - стохастичне нелінійне мультикомпонентне диференційне рівняння в частинних похідних

DSS - Differentiable Self-organizing Systems

QQISA - quantitative and qualitative integrated stability analysis

ML-MSM - Symposium on Integrating Machine Learning with Multiscale Modeling for Biological, Biomedical, and Behavioral Systems

IMAG - Interagency Modeling and Analysis Group

FNN - fully connected neural network

CNN - convolutional neural network

RNN - recurrent neural network

RCNN - recurrent convolutional neural network

Adam - Adaptive Moment Estimation

ВСТУП

В 21 столітті, все більше з'являється досліджень і методів які показують важливість тісного взаємозв'язку між елементами на одному рівні масштабу(наприклад між різними органами) та між самими рівнями системи(хімія, клітина, тканина, орган, організм, ніша в середовищі і т.д.)[13][12][3][14]. Взаємодію між компонентами цієї системи не звести до декількох параметрів, більш того це величина має масштабуватися в випадку деталізації процесів та взаємозв'язків що є в організмі.

В організмі є багато внутрішньоклітинних процесів стан яких складно виміряти через неінвазивні методи. Більш того, не зважаючи на можливість отримання ДНК клітин в кожній з клітин організму можуть бути деякі відмінності в епігенетичних конфігураціях які складно описати лише специфікацією клітин. Та досить складно дати фіксовану кількість параметрів які б дали повний опис того які значення приймає цей розподіл в різних частинах тіла, органу або навіть однієї тканини[25].

Те що кількість параметрів які потрібні для приблизного опису системи росте до нескінченності при спробах моделювати систему все точніше і точніше показує те що теоретичний опис стану системи не можна спростити до скінченної кількості характеристик. Іншими словами простір який описує стани нашої системи є нескінченновимірним. Більш того, неможливо розбити цей простір на прості незалежні компоненти.

Звичайно, є й задачі фізики в яких стан системи може описуватися через нескінченновимірний простір - це задачі де є диференційні рівняння в частинних похідних[1][4]. Але не для опису стану організма в одній точці не достатньо локальної поведінки розподілу однієї хімічної речовини чи одного коефіцієнта що параметризує конфігурації епігенетики. Із-за великого шуму додаткова інформація про стан в точці що дає одна просторово-часова змінна обмежена, так як похідні вищих порядків не дадуть ніякої інформації. Саме

тому для параметризації стану клітин потрібно декілька значень а це приводить нас до того що однокомпонентним рівнянням(навіть з досить високим порядком похідної) нам не описати стан нашого організму. Це приводить нас до багатоконпонентного диференційного рівняння. Також, у випадках коли в клітинах може бути регуляція вмикання та вимикання автокаталітичних реакцій то кількість параметрів що потрібно мати може стати ще більше, а явища які виникають можуть стати ще складнішими. В практичній частині роботи для параметризації просторово-часового розподілу стану ми буде використовувати рівняння хімічної кінетики для ключових хімічних речовин що відповідають за морфогенез та регенерацію. Але в теорії та методах буде наданий опис того якими компонентами можна природньо розширити це диференційне рівняння.

На даний час не є можливим повністю зрозуміти механізми експресії генів а тому зазвичай модель для цих механізмів пробують шукати з самих даних. В цій роботі ми не будемо дивитися на проміжні хімічні речовини та механізми які на локальному рівні сприяють хімічній динаміці основних морфогенів. Ми будемо вважати що подібні механізми проявляються в диференціальних рівняннях з невідомими функціями які їх описують. В роботі буде наданий опис того як приблизно можуть виглядати ці функції та буде надано деяке підґрунтя для цього.

Задачами цієї роботи є:

1. Пошук коефіцієнтів що параметризують простір функцій які використовуються для опису просторово-часової динаміки що описує локальні та глобальний стан органу.
2. Обґрунтування необхідності використання стохастичного нелінійного мультикомпонентного ДРЧП(СНМДРЧП) для пошуку диференціального рівняння при одному знімку компоненти яку можна

споглядати. Створення методології яка необхідна для перевірки необхідності використання цієї задачі.

3. Опис методу розв'язання задачі ідентифікації параметрів реакції-дифузії при декількох проміжних знімках однієї компоненти яку можна споглядати.
4. Опис методу розв'язання задачі ідентифікації параметрів СНМДРЧП при одному знімку однієї компоненти яку можна споглядати.

Під знімком мається на увазі значення якоїсь компоненти в області в якій розвивається патерн, для фіксованого часу. Це може бути рентген, мамографія, фото або інші неінвазивні методи які дозволяють отримати просторовий розподіл значень цієї компоненти.

На сучасний стан для отримання коефіцієнтів моделі потрібно мати дані на протязі усієї динаміки від початку і до кінця.

Існують різні роботи по виявленню коефіцієнтів в диференціальних рівняннях наприклад коли задача нелінійна, стохастична та відомо лише частина спостережуваних компонентів[24]. Або пошук коефіцієнтів нелінійного диференціального рівняння з частинними похідними[19][21].

Але автору цієї роботи не відомо щоб хтось до нього давав математичний опис задачі ідентифікації коефіцієнтів СНМДРЧП.

Більш того, на відміну від інших робіт, в нашій ми не маємо відносно неперервних спостережень по часу - в нас є всього декілька точок часу в яких ми можемо спостерігати значення нашої динаміки. Крім цього, в роботі ми можемо спостерігати значення лише однієї компоненти і тільки по ній маємо ідентифікувати коефіцієнти.

РОЗДІЛ I Теоретичні відомості

1.1 Стохастичне нелінійне мультикомпонентне ДРЧП

Зі вступу стає зрозуміло що для опису стану мультимасштабної біологічної системи необхідне мультикомпонентне нелінійне рівняння в частинних похідних. А із-за того що на реальні біологічні системи впливає шум який може в них викликати якісно різні та процеси які потрібні для стійкості нам не завжди може бути достатньо стійкості лінеаризованого ДРЧП. Це із-за того що стійкість лінеаризованого ДРЧП не гарантує те що випадкові впливи в сукупності не виведуть систему із ладу. А так як наявність стійкості яка працює для випадкових впливів сильно звужує простір можливих моделей нам потрібно поставити стохастичну задачу. Таким чином, щоб отримувати коефіцієнти нашого рівняння із невеликої кількості початкових даних нам потрібно поставити задачу ідентифікації коефіцієнтів стохастичного нелінійного мультикомпонентного ДРЧП(СНМДРЧП). Але перед цим ми маємо розібратися з нестохастичним випадком нелінійного мультикомпонентного ДРЧП.

Найвідомішим з таких є рівняння реакції-дифузії:

$$\frac{\partial \mathbf{u}}{\partial t} - (f(\mathbf{u}) + \Delta \mathbf{u}) = 0,$$

де $\mathbf{u}(t, x, y) = (u_1(t, x, y), u_2(t, x, y), \dots, u_n(t, x, y))$.

Якщо записувати для кожної компоненти, то маємо:

$$\frac{\partial u_i}{\partial t} - (f_i(u_1, u_2, \dots, u_n) + \Delta u_i) = 0, i \in \overline{1, n}$$

Однією з найважливіших характеристик цих мультикомпонентних ДРЧП є те що ці рівняння не змінюються від поворотів системи координат, що дуже важливо в моделюванні хімічної динаміки біологічних систем, оскільки вона інваріантна відносно повороту.

Одним з найзагальніших рівнянь для яких виконується ця властивість є:

$$\frac{\partial \mathbf{u}}{\partial t} - A(\mathbf{u}, \Delta \mathbf{u}, \begin{pmatrix} \nabla u_1 \nabla u_1 & \cdots & \nabla u_1 \nabla u_n \\ \vdots & \ddots & \vdots \\ \nabla u_n \nabla u_1 & \cdots & \nabla u_n \nabla u_n \end{pmatrix}) = 0.$$

де:

$$\nabla u_i = \left(\frac{\partial u_i}{\partial x}, \frac{\partial u_i}{\partial y} \right)$$

$$\Delta \mathbf{u} = (\Delta u_1, \Delta u_2, \dots, \Delta u_n), \Delta u_i = \frac{\partial^2 u_i}{\partial x^2} + \frac{\partial^2 u_i}{\partial y^2},$$

$A \in (L_2(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times n}))^n$ - вектор-функція від 2-ух векторів та матриці.

Але існує багато прикладів[1][7] які показують що патерни органів можна описати системою реакції-дифузії або їхніми модифікаціями які ми узагальнимо наступним чином:

$$\frac{\partial \mathbf{u}}{\partial t} - A(\mathbf{u}, \Delta \mathbf{u}) = 0.$$

Надалі ми будемо відштовхуватися від цього ДРЧП як від природного методу створення патернів органів. Але без крайових умов це все ще не повноцінне ДРЧП, крім того, як ми покажемо в “3.1 Дослідження дії оператора” такої постановки нам не достатньо для отримання нової динаміки.

Якщо б нас цікавило різноманіття самої динаміки то ми могли б використовувати мультикомпонентне нелінійне ДРЧП та дивитися як воно діє при різних нелінійних впливів та спробувати дати деяку інтерпретацію. Але так як ми хочемо оптимізувати ці процеси, які в тому числі включають в себе природне різноманіття біологічних систем, то ми вже не можемо дивитися на це як на зіткнення детермінованої системи з випадковим зовнішнім середовищем. Це змінює характер системи на стохастичний, а простір в якому розглядаються стани тепер описує не одного “індивіда” з його поточним станом а цілій випадковий простір “індивідів”, які можуть

утворюватися деякою динамікою та, при бажанні обмежені деякими відомими спостереженнями. На практиці ми не можемо спростити цей опис до нестохастичного тому що зазвичай ми можемо виміряти лише деякі компоненти - будову тканин та кісток, можемо дати деяке наближення для розподілу хімічних речовин в тілі. Але повністю ми не можемо отримати всю інформацію, із найбільших проблем в цьому напрямку це невідомі конфігурації епігенетики що відповідає за те наскільки активні різноманітні гени та процеси в клітині, вони можуть відрізнятись в різних органах, тканинах та навіть у клітинах в рамках однієї тканини може існувати сотні різних подібних конфігурацій[25]. Можливо основні компоненти що впливають на динаміку клітин і вдасться повністю просканувати в майбутньому, але навіть тоді це, скоріше за всього, буде занадто дорого, довго і складно.

Бачимо що для можливості оптимізації глибинної адаптації організму ми маємо як мінімум розглядати СНМДРЧП(стохастичне нелінійне мультикомпонентне ДРЧП), але основне питання залишається: “Чи достатньо цього мінімального опису для нашої задачі?”. Останнім часом складними взаємодіями в організмі займалася системна біологія але опису організму як взаємодіючої системи органів та їх відділів було недостатньо, а тому з’явилося так зване “Мультимасштабне моделювання” в рамках якого біологічна система розглядається як взаємодія між різними системами на різному масштабі. Критерії та методи того які процеси дають більш ефективний розвиток мають вибратися окремо. Ми будемо вважати що для першої спроби втілити оптимізацію глибинної адаптації тіла, нам буде досить СНМДРЧП. А далі, при необхідності, можна для оптимуму задачі оптимізації СНМДРЧП, поступово розширювати список можливих впливів та хімічних речовин, щоб знайти стійкий оптимум з достатньо гарним значенням цільової

функції. Для цього також потрібні також спеціальні методи градієнтного спуску які будуть описані в наступних підрозділах.

Якщо не робити модифікацій для пошуку оптимума, який підходить для цілей описаних вище, та вважати що ми можемо спроектувати взаємодію між різними системами та органами організму на 2-вимірну площину, то маємо наступну, найбільш мінімальну задачу:

Маємо $\mathbf{u} \in (L_2([0, T_{\max}] \times B))^n$ - вектор функція визначена на множині $[0, T_{\max}] \times B$ ($B \subset \mathbb{R}^2$), яка є розв'язком задачі Коші:

$$L(\mathbf{u}, A) = \frac{\partial \mathbf{u}}{\partial t} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta \mathbf{u}) = f,$$

з початковими умовами:

$$\mathbf{u}(0, x, y) = s_0(x, y),$$

та крайовими умовами:

$$\frac{\partial \mathbf{u}}{\partial n} = b = 0, (x, y) \in \partial B,$$

де n - нормаль до множини B .

Пояснення:

$L \in \mathbb{R}, L > 0$ - зазвичай невелика константа, тому функція A діє схоже в областях з невеликими похідними. Гарна функція A будує структуру з границь яка має може відновлювати себе при пошкодженнях та в якої цільові компоненти прямують до цільової функції.

$k * g(t, x, y)$ - моделює шум в сигналах який пришвидшує або загальмовує їх передачу. $g(t, x, y)$ - знаходиться в межах від $[-1, 1]$ - для моделювання ми будемо брати проекцію гаусівського випадкового поля з дуже маленьким радіусом на якому є кореляції - $G(t, x, y)$, та проектуємо його на $[-1, 1]$. Коефіцієнт $k \in (0, 1)$.

$f(t, x, y)$ - моделює зовнішні подразники - такі як шум в початкових умовах, та випадковий сильний вплив на деяку підобласть (поранення організму).

$s_0 \in (L_2(B))^n$ - при моделюванні ембріогенезу це перша клітина, зазвичай це маленьке коло посередині множини B з однаковим значенням вектору $\mathbf{u}(t, x, y)$. При моделюванні патернів це деякий початковий шум.

$s_T \in (L_2(B))^n$ - вектор-функція яка моделює будову органу/організму (компоненти що відповідають за різні підсистеми організму) та динамічні характеристики (компоненти що відповідають важливим векторним полям що описують транспортування та переробку речовин в органі/організмі). Структурні компоненти обмежені значеннями від 0 до 1 а векторні від -1 до 1.

$u \in (L_2([0, T_{\max}] \times B))^n$ - вектор-функція що моделює будову організму і його найголовніші векторні поля транспортування речовин + додає деякі додаткові компоненти які відповідають за підтримання цього стану при зовнішніх чинниках.

Введемо оператор проєкції вектор функції u на підпростір з тих m компонентів нашої динаміки які ми можемо споглядати:

$$P(\mathbf{u}) = P((u_1, u_2, \dots, u_n)) = (u_1, u_2, \dots, u_m)$$

Тоді похибкою розв'язку будемо вважати відхил цієї проєкції від цільового стану в час T :

$$\delta \mathbf{u} = P(\mathbf{u}(T, x, y)) - s_T(x, y)$$

Нам потрібно знайти неперервну за Ліпшицем функцію A , з фіксованою константою Ліпшиця L , таку що мінімізує *матсподівання квадрату похибки*:

$$F(\mathbf{u}, A) = M[F(\mathbf{u}, A, f, g)] = M[\|\delta \mathbf{u}\|_{L_2(B)}^2] \rightarrow \min$$

При деякому випадковому білому шумі $g(t)$ та випадкових подій зовнішнього середовища що моделюється $f(t)$.

Надалі кроками ми будемо називати дискретизацію процесу розвитку організму а ітераціями поступову зміну функції A - так щоб вона все краще і краще описувала розвиток і будову патерну органу.

1.2 Диференційована система з самоорганізацією

Недавно Google Research також помітили нові можливості в моделюванні біологічних систем, і організували групу людей які б займалися DSS(Differentiable Self-organizing Systems). Їхня робота, яка написана в 2020[5] стала першопрохідцем в цій області.

Команда яка займається DSS не ставила за мету опис подібних систем чи використання їх для пошуку нових моделей чи підходів в біології, а тому в них немає чіткої теорії і методології того як потрібно оптимізувати біологічні системи.

Ціль зведення алгоритмів до диференційованих - отримати універсальний набір базових блоків, використовуючи які можна не тільки вирішити задачі що можна описати через ці блоки а ще й оптимізувати параметри в цих базових блоках. І це стосується не тільки оптимізації FNN(повнозв'язних нейромереж) і CNN(згорткових нейромереж), також є можливість робити такі операції як згортку, обрахування відстані Вассерштейна або Сінкхорна між двома багатокomпонентними випадковими розподілами, а також використовувати алгоритми сортування. Все це поступово стає диференційованими функціями які також можуть брати участь в оберненому розповсюдженні градієнту. А тому їх можна використовувати в бібліотеках штучного інтелекту так ж просто як і CNN та FNN.

З цих нових методів ми візьмемо саме відстань Сінкхорна (Sinkhorn divergence)[2] так як на практиці він швидко дає гарні результати та гарно працює в різних випадках[2].

Відстань Вассерштейна між випадковими величинами α, β :

$$OT(\alpha, \beta) = \min_{\pi_1=\alpha, \pi_2=\beta} M[C(\pi_1, \pi_2)],$$

де $C(x, y)$ - функція від 2 змінних для якої виконується:

$$C(x, y) \geq 0, C(x, y) = C(y, x), C(x, x) = 0.$$

Але може існувати декілька випадкових величин $\pi = (\pi_1, \pi_2)$ для яких виконується цей оптимум, та також це не дуже гладка задача, тож зазвичай використовують ентропійну регуляризацію:

$$OT_\varepsilon(\alpha, \beta) = \min_{\pi_1=\alpha, \pi_2=\beta} M[C(\pi_1, \pi_2)] - \varepsilon H((\pi_1, \pi_2) | (\alpha \otimes \beta)),$$

де H – це умовна інформаційна ентропія.

Але із-за регуляризації ми можемо отримати що:

$$OT_\varepsilon(\alpha, \alpha) < 0$$

Тому відстань Сінкхорна визначають як:

$$S_\varepsilon(\alpha, \beta) = OT_\varepsilon(\alpha, \beta) - \frac{1}{2}(OT_\varepsilon(\alpha, \alpha) + OT_\varepsilon(\beta, \beta))$$

Найбільша користь для нас в тому що після обрахування розподілу π ми можемо спробувати отримати градієнт вже для розподілу α .

В підрозділі “2.3 Найважливіші підзадачі та способи їх розв'язку” буде показано як з градієнтів, які відповідають за цільову функцію, отримувати градієнти для інших задач.

1.3 Основні підходи математичної біології на даний час

Для того щоб патерни були стійкими було досліджено[7][11] деякі узагальнення звичайних рівнянь реакції-дифузії та проведено аналіз стійкості через лінеаризацію.

В цій роботі найбільш загальним є наступне рівняння:

$$\frac{\partial \mathbf{u}}{\partial t} - (f(\mathbf{u}) + D\Delta \mathbf{u}) = 0,$$

Узагальнюючи результат з розділу “2.3 General Conditions for Diffusion-Driven Instability: Linear Stability Analysis and Evolution of Spatial Pattern” в роботі [7], маємо необхідні умови лінійної стійкості:

$$\forall k \in \mathbb{R}: \operatorname{Re}(\Lambda(-f'_{\mathbf{u}} + k^2 D)) > 0$$

Де $f'_{\mathbf{u}} = J_f$ - яacobіан функції f . Для матриць під $\Lambda(A)$ ми будемо мати на увазі вектор з власних значень матриці, де для вектору: $v > c$ -позначає покомпонентне порівняння $v_j > c, j \in \overline{1, n}$.

Узагальнюючи результат з [7], маємо, що для $n \leq 2$ для лінійної стійкості є наступні необхідні і достатні умови:

$$P_{-f'_{\mathbf{u}}+k^2 D}(\lambda) = \det(\lambda E - f'_{\mathbf{u}} + k^2 D) = \sum_{i=0}^n p_{-f'_{\mathbf{u}}+k^2 D}^i \lambda^i$$

повинен мати додатні коефіцієнти.

Це також є необхідною і достатньою умовою для випадку коли усі власні числа дійсні.

У випадку з $\mathbf{u} = (u_1, u_2)$ нам необхідно і достатньо щоб:

$$\det(-f'_{\mathbf{u}} + k^2 D) > 0$$

$$\operatorname{tr}(-f'_{\mathbf{u}} + k^2 D) > 0$$

З чого маємо необхідні та достатні умови:

$$\min_{k \geq 0} (\det(-f'_{\mathbf{u}} + kD)) > 0$$

$$\operatorname{tr}(-f'_{\mathbf{u}}) > 0, \operatorname{tr}(D) \geq 0$$

в випадку $n > 2$ ці умови також мають виконуватися але вони не є достатніми.

Маємо для поліному:

$$Q_{-f'_{\mathbf{u}}, D}(k) = \det(-f'_{\mathbf{u}} + kD)$$

Повинні виконуватися наступні умови

$$Q_{-f'_{u,D}}(0) > 0$$

$Q_{-f'_{u,D}}(k)$ – не має коренів для дійсних $k > 0$

Для $n = 2$ отримуємо:

$$\det(-f'_u) > 0, \det(D) \geq 0$$

$$Q_{-f'_{u,D}}(k) = \det(D)k^2 + q_{-f'_{u,D}}k + \det(-f'_u)$$

–або немає дійсних коренів або має тільки додатні коефіцієнти

$$q_{-f'_{u,D}} \geq 0 \text{ або } q_{-f'_{u,D}}^2 < 4 * \det(D) * \det(-f'_u)$$

що еквівалентно:

$$\det(-f'_u) > 0, \det(D) \geq 0$$

$$q_{-f'_{u,D}} + 2\sqrt{\det(D) * \det(-f'_u)} > 0$$

Звичайно рівняння в біології мають лінійну стійкість, в тій ж роботі[7] є багато-масштабний аналіз невеликих впливів на загальну динаміку.

В випадку стійких біологічних систем вони зазвичай мають динаміку яка близька до лінійно-стійкої але не є такою - наприклад в деяких моделях серцебиття[18] моделюються подібним чином.

Багатомасштабність дозволяє мати сили що регулюють динаміку так щоб вона не виходила з досить складної множини, а вже в рамках цієї множини можуть бути різноманітні нелінійні складні ефекти.

Багатомасштабність є одним з найбільш просунутих методів в системній біології. Багато науковців покладають великі надії на нього як способу повного опису роботи організму[3].

Але, на даний час немає гарних інструментів які б могли гарно поєднати різні масштаби та диференційні рівняння. Для цього не було не прийнятних функцій для оцінки якості моделювання ні методів які б змогли

охопити різні ефекти. В цій роботі ми створимо ці інструменти та методологію як їх використовувати.

РОЗДІЛ II Постановка задачі

2.1 Формалізація задачі пошуку систем з самоорганізацією

Щоб наша система була оптимальною для різних можливих розширень СНМДРЧП який ми використовуємо для розв'язання нашої задачі ми маємо провести деякі модифікації.

2.1.1 Ріст домену

Так як при ембріогенезі тканини та органи постійно ростуть часто виникає проблема масштабу - від десятих міліметру до десятків сантиметрів тому в літературі зустрічається таке поняття як “domain growth”[8] тобто враховується постійний ріст організму а з ним і зменшення масштабу хімічних реакцій в порівнянні з розміром всього організму:

$$\frac{d\mathbf{u}}{dt} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta\mathbf{u} * \frac{1}{w(t)^2}) = \mathbf{f}$$

де $w(t)$ - коефіцієнт масштабу для організму.

2.1.2 Обмеження на функцію u

Необхідно щоб динаміка обмежувала функцію u - так як зазвичай на організм діють фізичні і хімічні обмеження, але ці обмеження набагато ширші за ті яким має відповідати кінцевий організм - їх можна моделювати тим що похідна буде мати напрямок(у середньому) лише в гіперкубі $C = [-R, R]^n$:

$$\frac{\partial \mathbf{u}}{\partial t} - (I_C(x, y) + k * g(t, x, y))A(\mathbf{u}, \Delta\mathbf{u} * \frac{1}{w(t)^2}) = \mathbf{f}$$

$$\text{Де } I_C(x, y) = \begin{cases} 1, & (x, y) \in C \\ 0, & (x, y) \in R^n \setminus C \end{cases}$$

2.1.3 Використання фільтрів “уваги”

Можна на кожній ітерації алгоритму використовувати деякі фільтри які мають “знайти і підкреслити” найбільш базову топологію патернів а далі поступово підкреслювати все більше і більше важливих деталей розвиваючи структуру організму.

Фільтри можуть мати як завгодно складний алгоритм пошуку так як вони даються для ітеративного алгоритму як зовнішній фактор. А тому послідовність областей можна шукати як за допомогою інших нейромереж так і аналізуючи структуру помилок алгоритму.

Після отримання фільтру G_i він використовується як частина формули для отримання похибки:

$$\delta \mathbf{u} = G_i(\mathbf{P}(\mathbf{u}(T, x, y)) - s_T(x, y))$$

Тому градієнт, що обраховується на i -тому кроці, кожен раз може бути обрахований для формули з новим фільтром.

2.1.4 Оцінка стійкості організму

Для того щоб оцінити не тільки можливість диференційного рівняння не зважаючи на шум отримати потрібну картинку а ще й оцінити стійкість отриманого організму, потрібно оцінювати стан не тільки в час T а стани по проміжку $[T, T + T_2]$.

Для цього потрібно змінити функцію що мінімізується:

$$\mathbf{F}(\mathbf{u}, A) = M[F(\mathbf{u}, A, f, g)] = M\left[\int_T^{T+T_2} (\delta \mathbf{u})^2 dt\right]$$

Але часто в такому вигляді вона сприяє створенню локальних оптимумів, тому краще на кожній ітерації використовувати стару оцінку, але час в якому оцінювати стан вибирати випадково з інтервалу $t_i \in [T, T + T_2]$:

$$\mathbf{F}_i(\mathbf{u}, A) = M[F(\mathbf{u}, A, f, g)] = M[(\delta \mathbf{u})^2, t = t_i]$$

Хоч це складно на практиці але гіпотетично за допомогою різноманітного аналізу(в тому числі перетворення Karhunen–Loeve(детальніше в пункті “3.2.5 Кількісна невизначеність”) та фільтрів які виділять процеси що призводять до нестійкості) знайти оптимальний час і фільтр який би допоміг системі за одну чи дві ітерації адаптуватися до відхилень які є.

2.2 Зв'язок з теорією оптимального керування

Базовий алгоритм для вирішення нелінійного рівняння з частковими похідними.

$$F(\mathbf{u}, A) = \|\delta\mathbf{u}\|_2^2 = \int_T^{T+T_2} \|\delta\mathbf{u}\|_{L_2(B)}^2 dt = \int_T^{T+T_2} \|\mathbf{u}(t, x, y) - s_T(x, y)\|_{L_2(B)}^2 dt \Rightarrow \min$$

Лінеризуємо в точці \bar{A} - маємо:

$$\begin{aligned} F(\bar{\mathbf{u}} + \varepsilon\bar{\bar{\mathbf{u}}}, \bar{A} + \varepsilon\bar{\bar{A}}) - F(\bar{\mathbf{u}}, \bar{A}) &= \varepsilon\delta F + O(\varepsilon^2) = \\ &= \varepsilon((\mathbf{u}(t, x, y) - s_T(x, y)), \bar{\bar{\mathbf{u}}})_2 + O(\varepsilon^2), \\ L(\bar{\mathbf{u}} + \varepsilon\bar{\bar{\mathbf{u}}}, \bar{A} + \varepsilon\bar{\bar{A}}) - L(\bar{\mathbf{u}}, \bar{A}) &= \\ &= \varepsilon \frac{\partial \bar{\bar{\mathbf{u}}}}{\partial t} - \varepsilon(1 + k * g(t, x, y))(A_1' \bar{\bar{\mathbf{u}}} + \frac{1}{w(t)^2} A_2' \Delta \bar{\bar{\mathbf{u}}} - \bar{\bar{A}}) + O(\varepsilon^2) = 0 \end{aligned}$$

Лінеаризоване PDE:

$$\frac{\partial \bar{\bar{\mathbf{u}}}}{\partial t} - (1 + k * g(t, x, y))(A_1' \bar{\bar{\mathbf{u}}} + \frac{1}{w(t)^2} A_2' \Delta \bar{\bar{\mathbf{u}}}) = (1 + k * g(t, x, y))\bar{\bar{A}}$$

Замінімо $(1 + k * g(t, x, y))\bar{\bar{A}}$ на r .

Підставимо:

$$\bar{\bar{\mathbf{u}}} = (K, r)_1 = (K, r)_{L_2([0, T+T_2] * B)},$$

де K - це ядро згортки.

$$\begin{aligned}\delta F &= 2((\mathbf{u}(t, x, y) - s_T(x, y)), (K, r)_1)_2 = \\ &= 2((\mathbf{u}(t, x, y) - s_T(x, y), K)_2, r)_1 = (c, r)_1\end{aligned}$$

Таким чином якщо рахувати градієнт F по r то отримаємо c .

Для знаходження c , використаємо формулу для ядра:

$$\begin{aligned}\frac{\partial K}{\partial t} - (1 + k * g(t, x, y))(A'_1 K + \frac{1}{w(t)^2} A'_2 \Delta K) = \\ \delta(\bar{t} - t, \bar{x} - x, \bar{y} - y)\end{aligned}$$

Та помножимо її на $(\mathbf{u}(\bar{t}, \bar{x}, \bar{y}) - s_T(\bar{x}, \bar{y}))$ використовуючи скалярний добуток $(u, v)_2 = (I_{[T, T+T_2]} u, v)_1$ для змінних $\bar{t}, \bar{x}, \bar{y}$, отримуємо спряжене до лінеаризованої задачі ДРЧП:

$$\begin{aligned}\frac{\partial c}{\partial t} - (1 + k * g(t, x, y))(A'_1 c + \frac{1}{w(t)^2} A'_2 \Delta c) = \\ I_{[T, T+T_2]}(t, x, y) 2(\mathbf{u}(t, x, y) - s_T(x, y)),\end{aligned}$$

$$\text{де } I_{[T, T+T_2]}(t, x, y) = \begin{cases} 1, & t \in [T, T + T_2] \\ 0, & t \notin [T, T + T_2] \end{cases}$$

$$\text{Також маємо обмеження: } c(T + T_2, x, y) = 0, \frac{\partial c}{\partial n} = 0.$$

Таким чином маємо лінійне ДРЧП для c для змінних t, x, y .

Можна легко помітити що c - не що інше як градієнт задачі керування системою:

$$\frac{d\mathbf{u}}{dt} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta \mathbf{u} * \frac{1}{w(t)^2}) = r_{\text{control}}$$

Тобто для того щоб знайти потрібний нам градієнт ми паралельно розв'язали більш просту задачу пошуку градієнту для керування.

Надалі ми маємо зрозуміти які є основні підзадачі.

2.3 Найважливіші підзадачі та способи їх розв'язку

Найбільш базовою є підзадача прямого керування яку ми розв'язали в попередньому розділі.

За допомоги розв'язку цієї задачі ми можемо розв'язати нестохастичний випадок нашої задачі (f, g - фіксовані). Для цього ми маємо отримати градієнт від оператора.

Тепер згадаємо що таке r в минулому розділі:

$$\delta F = (c, (1 + k * g(t, x, y))\bar{A})_1 = ((1 + k * g(t, x, y))c, \bar{A})_1,$$

$(1 + k * g(t, x, y))c$ - градієнт по A .

Так як у нас досить багато обмежень на A у нас є багато обмежень на те як ми можемо його змінювати а тому при ми маємо ще спроектувати цей градієнт на обмеження яким задовольняє A , в найпростішому випадку:

$$A_{i+1} = P_A(A_i - \lambda_i(1 + k * g(t, x, y))c).$$

В нашому випадку ми користуємось алгоритмом градієнтного спуску який ми опишемо в “ та алгоритмом який ми опишемо в “2.7 Модифікації градієнту””.

Цей базовий алгоритм можна легко модифікувати, розглянувши вже випадкові функції.

Для цього лише потрібно переробити скалярний добуток на:

$$(u, v)_1 = M[(u, v)_{L_2([0, T+T_2]*B)}]$$

Для цієї задачі ми також маємо отримати градієнт - позначимо його як **grad**. Іншими словами **grad** - градієнт матсподівання похибки.

Для того щоб його обрахувати необхідно буде зробити достатню кількість симуляцій з випадковими f та g . А потім для кожної симуляції порахувати свій градієнт c_j і потім порахувати середнє:

$$\mathbf{grad} = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{grad}_j = \frac{1}{N_i} \sum_{j=1}^{N_i} (1 + k * \mathbf{g}_{i,j}(t, x, y))c_j,$$

Де $\mathbf{g}_{i,j}$ - випадкові реалізації випадкового просторового-часового процесу \mathbf{g} .

Тобто для пошуку градієнту стохастичної задачі ми маємо знайти N_i градієнтів для нестохастичних задач.

При досить великій кількості симуляцій цей градієнт наблизатиметься до справжнього градієнту мат-сподівання похибки.

А далі потрібно спроектувати його на обмеження яким задовольняє A :

$$A_{i+1} = P_A(A_i - \lambda_i \mathbf{grad}).$$

В підрозділі “2.5 Обмеження на оператор та їх імплементація” ми покажемо скінченновимірне та реалістичне наближення функції A , а тому оператор проєкції буде проектувати з простору функцій в скінченновимірний простір. Також, в підрозділі “2.7 Модифікації градієнту” ми покажемо як трохи модифікувати цей алгоритм так щоб він гарно працював для нечастих “катастрофічних градієнтів”.

Але це не всі підзадачі які в нас є. Якщо функція оцінки стану складна то задачу з минулого розділу можна розглядати як дві підзадачі:

- 1) Пошук градієнта для самої функції оцінки $F(\mathbf{u}, A)$. Це потрібно щоб зрозуміти як саме змінювати наш стан \mathbf{u} в області цільових рішень $t \in [T, T + T_2]$.
- 2) Використання градієнту функції оцінки для отримання градієнту задачі керування.

Щоб узагальнити нашу задачу нам не потрібно багато чого змінювати в минулому розв’язку, в нас лише:

$$\delta F = 2((\mathbf{u}(t, \mathbf{x}, y) - s_T(\mathbf{x}, y)), (K, \mathbf{r})_1)_2$$

на:

$$\delta F = (GF(\mathbf{u}), (K, \mathbf{r})_1)_2$$

де $GF(\mathbf{u})$ - градієнт функції оцінки $F(\mathbf{u}, A)$ по \mathbf{u} .

А тому, аналогічними до попереднього розділу міркуваннями, легко отримати те як цей градієнт утворює градієнт задачі керування, а саме:

$$\frac{\partial c}{\partial t} - (1 + k * g(t, x, y))(A'_1 c + \frac{1}{w(t)^2} A'_2 \Delta c) =$$

$$I_{[T, T+T_2]}(t, x, y) GF(\mathbf{u}),$$

$$c(T + T_2, x, y) = 0, \frac{\partial c}{\partial n} = 0.$$

2.4 Пошук області стійких розв'язків

Для цього потрібні модифікації градієнту які дадуть поступове отримання безпечної зони в якій випадкові невеликі зміни параметрів диференціального рівняння від мутацій геному не руйнують організм.

Також потрібні обмеження на значення функції а в рамках цих обмежень має працювати обмеження на оператор що дадуть стійкість розв'язку для відомих стійких станів та опише можливі зони нестійкості що можуть бути характерні для подібних організмів.

І завдяки цим обмеженням можна буде тримати систему в її області стійкості, але для цього потрібно також зрозуміти як саме може змінюватись оператор і як зменшити розмірність задачі не втративши ключові властивості системи.

Надалі ми пояснимо та введемо всі ці обмеження, а також наведемо основні ідеї для їх імплементації.

2.5 Обмеження на оператор та їх імплементація

Крім звичайних обмежень такі як константа Ліпшиця L , необхідні також обмеження на оператор для запобігання великим помилкам -

наприклад зробити так щоб оператор диференційного рівняння не робив сильного експоненційного росту.

Лінеаризуємо задачу в якійсь точці $\bar{\mathbf{u}}$, маємо:

$$L(\mathbf{u}, \mathbf{f}) = \frac{\partial \mathbf{u}}{\partial t} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta \mathbf{u} * \frac{1}{w(t)^2}) - \mathbf{f} = 0,$$

$$L(\bar{\mathbf{u}} + \varepsilon \bar{\bar{\mathbf{u}}}, \bar{\mathbf{f}} + \varepsilon \bar{\bar{\mathbf{f}}}) - L(\bar{\mathbf{u}}) = \varepsilon \frac{\partial \bar{\bar{\mathbf{u}}}}{\partial t} - \varepsilon (1 + k * g(t, x, y)) (A'_1 \bar{\bar{\mathbf{u}}} + \frac{1}{w(t)^2} A'_2 \Delta \bar{\bar{\mathbf{u}}}) + O(\varepsilon^2) - \varepsilon \bar{\bar{\mathbf{f}}}(t, x, y) = 0.$$

Маємо лінеаризоване рівняння:

$$\frac{\partial \bar{\bar{\mathbf{u}}}}{\partial t} - (1 + k * g(t, x, y)) (A'_1 \bar{\bar{\mathbf{u}}} + \frac{1}{w(t)^2} A'_2 \Delta \bar{\bar{\mathbf{u}}}) = \bar{\bar{\mathbf{f}}}(t, x, y),$$

$$\bar{\bar{\mathbf{u}}}(0, x, y) = 0, \quad \frac{\partial \bar{\bar{\mathbf{u}}}}{\partial n} = 0.$$

Маємо що для того щоб $\bar{\mathbf{u}}$ була стійкою траєкторією достатньо щоб для кожної точки власні числа:

$$\forall l \in \mathbb{R}: \operatorname{Re}(\Lambda((1 + k * g(t, x, y)) (A'_1 \bar{\bar{\mathbf{u}}} - l^2 \frac{1}{w(t)^2} A'_2 \Delta \bar{\bar{\mathbf{u}}})) \leq -\alpha.$$

Обмежимо:

$$\frac{1}{2} < (1 + k * g(t, x, y)).$$

Маємо що нам для цього достатньо довести що:

$$\forall l \in \mathbb{R}: \operatorname{Re}(\Lambda((A'_1 \bar{\bar{\mathbf{u}}} - l^2 \frac{1}{w(t)^2} A'_2 \Delta \bar{\bar{\mathbf{u}}})) \leq -2\alpha.$$

Провівши схожі обрахунки з підрозділом “1.3 Основні підходи математичної біології на даний час” матимемо умову стійкості:

$$\operatorname{tr}(-A'_1 - 2\alpha E) \geq 0$$

$$\operatorname{tr}(A'_2) \geq 0$$

$$\det(-A'_1 - 2\alpha E) \geq 0$$

$$\det(A'_2) \geq 0$$

$$q_{-A'_1-2\alpha E, A'_2} + 2\sqrt{\det(-A'_1 - 2\alpha E) * \det(A'_2)} \geq 0$$

Або, якщо ввести вектор-функцію:

$$\text{cond}_{A'_1, A'_2} = \left(\text{tr}(-A'_1 - 2\alpha E), \text{tr}(A'_2), \det(-A'_1 - 2\alpha E), \det(A'_2), q_{-A'_1-2\alpha E, A'_2} + 2\sqrt{\det(-A'_1 - 2\alpha E) * \det(A'_2)} \right)$$

$$\text{cond}_{A'_1, A'_2}(\mathbf{u}, \Delta \mathbf{u}) \geq 0$$

Для нас потрібно робити для цієї функції досить широку область стійкості, на практиці ми будемо зазвичай дивитися на стійкість при $u_i > 0, i = 1..n$ так як в цих межах будуть знаходитися найголовніші процеси.

Але ця умова є достатньою лише для $n \leq 2$.

Для загального випадку має виконуватися умова:

$$\text{COND}_{A'_1, A'_2} = \left(\min_{l \geq 0} \text{Re}(\lambda_{-A'_1-2\alpha E + lA'_2}^i) \right) = \left(-2\alpha + \min_{l \geq 0} \text{Re}(\lambda_{-A'_1 + lA'_2}^i) \right),$$

$$\text{COND}_{A'_1, A'_2}(\mathbf{u}, \Delta \mathbf{u}) \geq 0$$

де λ_M^i - власні значення M , які ми беремо як корені поліному $P_{-M}(\lambda) = \det(\lambda - M) = 0$.

Звичайно, існують і більш прості підмножини області стійкості, наприклад тоді коли:

$$-A'_1 - 2\alpha E \geq 0, A'_2 \geq 0$$

Для матриць під “ $A \geq B$ ” ми будемо мати на увазі те що $A - B$ - додатно-визначена матриця.

На відміну від $\text{cond}_{A'_1, A'_2}(\mathbf{u}, \Delta \mathbf{u}) \geq 0$ для $n > 2$ це є достатньою умовою для стійкості, але вона сильно звужує область пошуку.

Для простих підмножин легко вивести формули для проектування на них. Це можна зробити напряду в ітераційному алгоритмі(наприклад

створити формулу для проекції на підмножину подібних функцій), або можна отримати через суму більш простих функцій що задовольняють цим умовам:

$$A(\mathbf{u}, \Delta\mathbf{u}) = -\alpha\mathbf{u} + \sum_{i=1}^K c_i(\mathbf{u}) * d_i(\Delta\mathbf{u})$$

$$c_i' \leq 0, d_i' \geq 0$$

Щоб спростити можна взяти:

$$A(\mathbf{u}, \Delta\mathbf{u}) = c(\mathbf{u}) + d(\Delta\mathbf{u}),$$

А щоб ще спростити можна подивитися як диференціальне рівняння буде вести себе при $d(\Delta\mathbf{u}) = D\Delta\mathbf{u}$, D - додатно визначена матриця.

$c(\mathbf{u})$ - можна взяти як поліноми від n до 3 порядку включно.

Ми хочемо щоб $c_i' < 0$ для компонентів \mathbf{u} що $u_i > 0$.

Використовуючи критерій Сільвестра для матриці $-(c_i' + (c_i')^T)$ маємо що нам потрібно щоб усі кутові мінори матриці $m_k(\mathbf{u})$ (детермінанти матриці з перших k стовпчиків та k рядків) були додатними.

Ми можемо брати такі функції як початкові. А потім, під час оптимізації, розширити простір допустимих функцій та поступово будувати їх для стійкої динаміки вже в більш загальних умовах.

Для розширення нам можна навіть залишити функцію як поліном, бо це спрощення дає нам деяке якісне наближення.

Часто поліноми використовуються в рівняннях реакції-дифузії як деяка абстракція від того як саме реалізована функція на мікрмасштабі. Вважається що цього наближення достатньо для моделювання складної взаємодії між хімічними речовинами.

В рівняннях хімічної кінетики зазвичай добуток 2 або більше величин позначають те які саме компоненти одночасно потрібні для якоїсь хімічної реакції що дає нові хімічні речовини. Всі ці прирости сумуються і якраз і виходять поліноми.

Звичайно є різні ефекти на мікромасштабі але вони зазвичай нейтралізують одна одну при переході на макромасштаб а тому поліноми є досить гарним способом описати вплив хімічних реакцій на макромасштаб.

Цю логіку ми можемо розширити не тільки на взаємовідносини між хімічними речовинами а й на взаємовідносини між нормами векторів в різних векторних полях(наприклад кровотоку) та важливими ефектами на мікромасштабі що впливають на весь організм. Тут також наявність деякого фактору може підсилити або загальмувати процеси - це також можна моделювати через поліном.

Тому хоч хімічна кінетика і не може описати динаміку всього організму з усіма його глибокими ефектами, але використавши додаткові компоненти що не є хімічними концентраціями в нашій задачі ми можемо розширити це.

Тому ми можемо користуватися поліномом як гарним способом поєднання мікромасштабу з макромасштабом та зменшення розмірності простору потрібних операторів до скінченновимірною.

2.6 Обмеження на функцію та їх імплементація

Ми вводим деякі м'які обмеження на функцію \mathbf{u} : $\mathbf{u} \in [-1, 1]$. Для цього ми маємо зробити так щоб вихід за рамки нерівності $|\mathbf{u}| - 1 \leq 0$ отримав свій штраф який врахується в функції що ми мінімізуємо.

Тому ми беремо і додаємо до нашої цільової функції додатковий штраф:

$$\|\text{relu}(|\mathbf{u}(T, \cdot, \cdot)| - 1)\|_{L_2(B)}^2 = \iint_B (\text{relu}(|\mathbf{u}(T, x, y)| - 1))^2 dx dy,$$

де $\text{relu}(x) = \max(x, 0)$.

Також в нас є фізичні обмеження які реалізуються строго:

$$|\mathbf{u}| \leq R_{\text{phys}},$$

Ці обмеження ми будемо застосовувати вже на рівні диференційного рівняння і будемо вважати що для $|\mathbf{u}|^2 = R_{\text{phys}}^2$:

$$\frac{\partial}{\partial t} |\mathbf{u}|^2 = 2 \left(\frac{\partial}{\partial t} \mathbf{u}, \mathbf{u} \right) = (A(\mathbf{u}, \Delta \mathbf{u}), \mathbf{u}) < 0.$$

Але так як на практиці таке втілення обмежень не дуже корисне так як ці межі алгоритм може випадково “переступити” а робити там сильний градієнт що тягне його назад може привести до великих похідних в ДРЧП.

А тому ці фізичні обмеження ми реалізуємо на практиці постійним проектуванням функції на множину $E = [-R_{\text{phys}}, R_{\text{phys}}]^n$:

$$\text{Pr}_E(\mathbf{u}) = \mathbf{u} * \left(\frac{\min(R_{\text{phys}} + \varepsilon, |\mathbf{u}| + \varepsilon)}{|\mathbf{u}| + \varepsilon} \right),$$

Де $\varepsilon > 0$ - деяка дуже маленька константа, яка потрібна щоб в формулі не було невизначеності типу $\frac{0}{0}$.

2.7 Модифікації градієнту

Найвідомішим алгоритмом для зашумленого градієнта є алгоритм Adam, для якого:

$$x_{n+1} = x_n - \alpha \frac{\bar{m}_n}{\sqrt{\bar{v}_n + \varepsilon}},$$

Де \bar{m}_n, \bar{v}_n - експоненційне згладжування для градієнту g_i та квадрату його норми $|g_i|^2$ з коефіцієнтами β_1, β_2 відповідно.

Але цей алгоритм гарно підходить коли є якийсь гаусівський шум або якийсь інший який легко згладжується за допомогою експоненційного згладжування:

$$\bar{m}_n = \frac{\beta_1 g_n + \beta_1^2 g_{n-1} + \dots + \beta_1^n g_1}{1 - \beta_1^{n+1}}$$

$$\bar{v}_n = \frac{\beta_2 |g_n|^2 + \beta_2^2 |g_{n-1}|^2 + \dots + \beta_2^n |g_1|^2}{1 - \beta_2^{n+1}}$$

І-за того що наші диференційні рівняння можуть при деяких конфігураціях бути хаотичними слідує те що в цих конфігураціях є висока чутливість до параметрів що експоненційно росте з часом. Тобто ми маємо що інколи наш в нас виникають “катастрофічні” градієнти які можуть впливати на \bar{m}_n в алгоритмі Adam сотні кроків.

При такому розподілі випадкових відхилень ми не можемо згладжувати самі градієнти, а також ми маємо не дати сильним викидам довго впливати на процес градієнтного спуску.

Для цього ми пропонуємо наступний крок для градієнтного спуску:

$$x_{n+1} = x_n - \alpha \bar{s}_n,$$

де \bar{s}_n - експоненційне згладжування для $\frac{g_i}{|g|+\epsilon}$, з коефіцієнтом β_1 :

$$\bar{s}_n = \frac{\beta_1 \frac{g_n}{|g_n| + \epsilon} + \beta_1^2 \frac{g_{n-1}}{|g_{n-1}| + \epsilon} + \dots + \beta_1^n \frac{g_1}{|g_1| + \epsilon}}{1 - \beta_1^{n+1}}$$

Або через алгоритм:

Input: $x_0, S_0 = \frac{g_0}{|g_0| + \varepsilon}$

for i **in** $1, \dots, N$ **do**

$$S_i \leftarrow \beta_1 \frac{g_i}{|g_i| + \varepsilon} + (1 - \beta_1) S_{i-1}$$

$$\bar{s}_i \leftarrow \frac{S_i}{1 - \beta_1^{i+1}}$$

$$x_{i+1} \leftarrow x_i - \alpha_i \bar{s}_i$$

end for

Output: x_N

Цей алгоритм має деякі властивості ослаблювати та пригнічувати як і звичайний шум так і нечастий але катастрофічний шум. Також він подібний до алгоритму Adam в тому що він також оминає локальні оптимуми та знаходить локальний оптимум з досить високою областю стійкості.

В рамках цієї статті ми будемо називати цей метод - метод усереднення напрямків. Тому що $\frac{g_i}{|g_i|}$ - одиничний вектор що показує напрямок зростання.

Останнє дуже важливо для того щоб алгоритм був стійким до випадкових “генетичних змін” та для того щоб потім можна було дискретизувати неперервне рівняння його хімічною реалізацією.

В статті[6] був використаний метод градієнтного спуску який можна спростити до того що ми описали вище. І як показує практика, як мінімум стиск 4-ьох байтових величин до одного байту не особливо скажуться на точності відтворення необхідної динаміки.

2.8 Нестрогі та строгі обмеження на коефіцієнти

Чим більше обмежень на коефіцієнти оператора то тим вужче простір можливих стохастичних процесів які можна отримати через ці коефіцієнти, тим менша ймовірність перенавчання.

Тому якщо в нас багато коефіцієнтів - позначимо їх W , то ми маємо ввести деякий додатковий штраф для цільової функції:

$$\varepsilon|W|^2$$

який допоможе мінімізувати перенавчання.

Але при маленькій кількості коефіцієнтів цього робити не обов'язково.

Також в нас можуть бути деякі додаткові обмеження на те які можуть бути коефіцієнти поліному.

Наприклад в нас немає коефіцієнтів для таких доданків як u_i^2, u_i^3 тому-що це означає реакцію речовини з самою собою.

Також ми вже отримали деяку область стійкості від якої ми відштовхуємося. Але потрібно дати деяке зауваження що не завжди є можливість повністю стабілізувати рівняння. Це часто проявляються в невеличких хаотичних атраторах. Для нелінійної динаміки що стійка на великому масштабі мати хаотичність на маленькому масштабі є досить звичним явищем. І ми мажемо при моделюванні також зберегли це явище не жертвуючи лінійною стійкістю рівняння на більшому масштабі. Тобто, ми відштовхуємося від того що є функція яка має лінійну стійкість \bar{A} та невелика за нормою функція $\varepsilon\bar{A}$ яка б при додаванні утворювали A :

$$A(\mathbf{u}, \Delta\mathbf{u}) = \bar{A}(\mathbf{u}, \Delta\mathbf{u}) + \varepsilon\bar{A}(\mathbf{u}, \Delta\mathbf{u})$$

Це значить що ми шукаємо розв'язок не тільки в самій області лінійної стійкості а ще й в її околі.

Позначимо область стійких рішень як St . В випадку навчання дуже важливо тримати функцію біля St , тому що якщо почати з області де навкруги тільки нестійкі рішення то градієнти будуть мати дуже сильний

шум і це буде ламати роботу звичайних алгоритмів градієнтного спуску. Навіть метод усереднення напрямків який ми використовуємо та строгі і гладкі обмеження на u якщо і дадуть локальний оптимум то досить поганий.

Тому ми на кожній ітерації нашого алгоритму яка оновлює коефіцієнти ми проектуємо їх на область в рамках якої ми шукаємо рішення.

Для того щоб отримати оператор проектування потрібно розглянути який вигляд мають область стійких рішень:

$$\text{COND}_{A'_1, A'_2} = \left(-2\alpha + \min_{l \geq 0} \text{Re}(\lambda_{-A'_1 + lA'_2}^l) \right) \geq 0$$

Де $A'_1: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $A'_2: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ - якобіан функції $A(\mathbf{u}_1, \mathbf{u}_2)$ по вектору \mathbf{u}_1 та \mathbf{u}_2 відповідно.

Для $n \leq 4$ можна буде отримати явні формули для $\lambda_{-A'_1 - 2\alpha E + lA'_2}^l$.

Для того щоб спростити вираз $\min_{l \geq 0} \text{Re}(\lambda_{-A'_1 + lA'_2}^l)$ на практиці пропонується використовувати для обрахунку та спрощення формул для цих коефіцієнтів поліномів символні обчислення. Також символні обчислення можна використати для того щоб знайти функцію що проектує точки на цю область.

При великих значень n , потрібно зробити гладку апроксимацію цієї проєкції, для цього можна використати сплайни або нейромережу.

Позначимо коефіцієнти що визначають нашу функцію A як a . Тобто у нас є система S така що при різних коефіцієнтах ми отримуємо різні значення. $S(a, \mathbf{u}, \Delta \mathbf{u}) = A(\mathbf{u}, \Delta \mathbf{u})$ Таким чином ми в будь-якому випадку зможемо побудувати проєкцію або її наближення, що обмежує наші коефіцієнти які визначають нашу систему:

$$\text{Pr}_{St}(a) = \underset{A(\bar{a}) \in St}{\text{argmin}} |\bar{a} - a|$$

Також часто буває що початкова динаміка не стійка але для деякого часу $t > T_{st}^\varepsilon$ для всіх початкових умов, крім множини з мірою що менше ε , динаміка розділяється по областям в яких вже можна довести стійкість.

Також є складні стійкі стани які в деяких своїх підобластях ведуть себе як більш прості стійкі стани такий випадок ми розглянемо в “3.1. Теорія QQISA”.

2.9 Проблема неопуклості цільової функції та малої кількості початкових даних

2.9.1 Отримання якісних характеристик

Хоч я і використовував для прикладу обрахування градієнту лінійний фільтр але він все одно може не дати гарний функціонал який легко оптимізувати та часто можна отримати не те що потрібно. Так як суть паттерна не змінюється при зсуві чи невеликих поворотах - нам не потрібні звичайні лінійні простори функцій для опису відстані між паттернами. Більш того нам потрібен простір не самих функцій а саме простір класів еквівалентності цих функцій який не змінюються при поворотах зсувах та невеликому випадковому шумі.

Найбільш простим рішенням є побудова деякого простору ознак що відповідають за паттерни та які не змінюються при вищенаведених перетвореннях.

Є дослідження де для цього простору використовуються характеристики що виловлюють нейромережі. Найчастіше для цього використовують характеристики на внутрішніх шарах нейромережі VGG.

Але виловлення характеристик через нейромережу це досить складна річ. При спробі виловлювати характеристики нелінійним чином може вийти так, що отримується багато локальних оптимумів які можуть домалювати до паттерна якісь інші об'єкти. В найгіршому випадку - коли нейромережа не

“бачила” якійсь патерн під час навчання, її характеристики по ньому не будуть стійкими до шумів і вона буде бачити один і той же об’єкт з різним шумом як різний. Інколи буває і навпаки - бачити в ньому деяке невелике відхилення від відомого патерну. А тому, із-за наших методів пошуку стійкого оптимуму, вона завжди буде зупинятися на стійкому патерні що був для неї найбільш схожий.

Звичайно її можна довчити але вона тоді стане занадто вузькою в своїх оцінках і не зможе гарно дати інші патерни.

Але є методи що досить близькі до лінійних але при цьому також дають гарний результат.

Наприклад, можна розглянути гаусівське згладжування:

$$v_{\sigma^2} = \langle \mathbf{u}, G(\sigma^2, \mathbf{x}, \mathbf{y}) \rangle$$

$$v_0 = \mathbf{u}$$

де $G(\sigma^2, \mathbf{x}, \mathbf{y})$ - ядро фільтру гауса, $\langle \cdot, \cdot \rangle$ - згортка.

Побудуємо наступну послідовність з m елементами:

$$w_0 = v_0$$

$$w_i = \langle w_{i-1}, G(2^{i-1}\sigma^2, \mathbf{x}, \mathbf{y}) \rangle$$

$$w_0, \dots, w_{m-1} = V_{(2^0-1)\sigma^2}, V_{(2^1-1)\sigma^2}, \dots, V_{(2^{m-1}-1)\sigma^2}$$

Ми будемо називати цю техніку “Проста ієрархія згорток з m рівнями”.

Але для неї ми бачимо що сусідні згортки сильно корелюють одна з одною.

Будемо вважати що в середньому:

$$\text{cor}(\langle w, G(2^{i-1}\sigma^2, \mathbf{x}, \mathbf{y}) \rangle, w) = k_{\text{coef}}$$

Тому ми трохи модифікуємо наш метод отримання ієрархії згорток так, щоб, перед отриманням нового рівня ми позбулися кореляції з попереднім:

$$w_0 = v_0$$

$$\bar{w}_{i-1} = \langle w_{i-1}, G(2^{i-1}\sigma^2, x, y) \rangle + \frac{k_{\text{coef}}}{(1 - k_{\text{coef}})} (\langle w_{i-1}, G(2^{i-1}\sigma^2, x, y) \rangle - w_{i-1})$$

$$w_i = \langle \bar{w}_{i-1}, G(2^{i-1}\sigma^2, x, y) \rangle$$

Також цю додаткову операцію можна вважати виділенням границь на малюнку з попереднього рівня так щоб виділити границі малюнка які присутні на новому рівні.

На практиці ми будемо брати $k_{\text{coef}} = \frac{2}{3}$.

Цю техніка також використовувалася в [6] але вона не була формалізована для різних коефіцієнтів кореляції. Ця техніка допомагає пришвидшити та покращити збіжність алгоритму.

Ми назвемо її “модифікована ієрархія згорток з m рівнями”.

Границі високого рівня відповідають за взаємозв'язки між різними об'єктами та підобластями в \mathbf{u} , а більш меншого за різні процеси між підобластями та в самих підобластях.

Для того щоб відрізнити один патерн від іншого цих розподілів нам достатньо. Більш того при розгляданні сумісних розподілів ми можемо аналізувати таким чином ієрархію розподілів.

Але для нашого випадку нам достатньо близькості розподілів на одному рівні, побудуємо їх.

Якщо ми візьмемо рівномірний випадковий розподіл для $(\xi_1, \xi_2) = U(B)$, то ми можемо ввести наступну випадкову величину:

$$\zeta_i(\mathbf{u}(t, \cdot, \cdot)) = W_i(t, \xi_1, \xi_2)$$

Аналогічну модифіковану ієрархію згорток ми можемо побудувати вже не для самого \mathbf{u} а для його проекції на ті компоненти що ми можемо споглядати, та отримати новий спосіб описати близькість патернів на i -тому масштабі:

$$S_\varepsilon(\zeta_i(P(\mathbf{u})(T, \cdot, \cdot)), \zeta_i(P(s_T)))$$

А для того щоб весь патерн на різних його масштабах був описаний, ми просумуємо цю відстань на всіх масштабах

$$F(\mathbf{u}, A) = PD(P(\mathbf{u})(T, \cdot, \cdot), s_T) = \sum_{i=0}^{m-1} S_{\varepsilon}(\zeta_i(P(\mathbf{u})(T, \cdot, \cdot)), \zeta_i(s_T))$$

Де PD - скорочення від pattern divergence.

Для цього $F(\mathbf{u}, A)$ позначимо градієнт $PD(P(\mathbf{u})(T, \cdot, \cdot), s_T)$ від \mathbf{u} як $GPD(P(\mathbf{u})(T, \cdot, \cdot), s_T)$.

Модифікована ієрархія згладжень - це досить універсальна структура, та яка при своїй універсальності більш ефективна за VGG що було показано в роботі[6].

2.9.2 Вирішення проблеми неоднозначності рішень

Так як PD -це відстань в просторі класів еквівалентностей, ми маємо деяку складність з тим коли патерн у тварини має якусь фіксовану направленість. Як досліджується в роботах[8] це демонструється тим щоб є деякий початковий розподіл клітин та хімічних речовин що йдуть від нервового гребня та розподіляються по всьому організму. Цей початковий розподіл часто може формувати деякі прості базові структури в яких є направленість - наприклад періодичні лінії.

Для отримання цих ліній а також і цільового патерну в роботі[8] було просимульований постійний вплив який починається від хребта. Ми прибираємо етап отримання ліній вважаючи його лінійним і не цікавим. Будемо вважати що початкові хімічні речовини що утворилися на ранніх етапах постійно впливають на організацію видимих змін в організмі.

Тому ми будемо моделювати вплив не однієї лінії що виступає зовнішнім полем що стимулює і направляє ріст патерну, а відразу періодичні лінії.

Для отримання цих основних ліній рекомендується зробити перетворення Фур'є і вибрати найсильніший пік по модулю - це і буде головна хвиля яка стимулює рівняння.

Щоб не ламати динаміку розвитку патерну ми зробимо цей вплив таким що може бути тільки при високих концентраціях хімічної речовини в цій хвилі.

Тому щоб патерн знав в якому напрямку розвиватися ми використаємо цю просту нестохастичну та фіксовану зовнішню стимуляцію:

$$\frac{d\mathbf{u}}{dt} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta\mathbf{u} * \frac{1}{w(t)^2}) = \mathbf{f} + fl_{\varepsilon}$$

$$fl_{\varepsilon} = (I_{1-\sin(c_2x+c_1y+c_0)<\varepsilon}^{\sigma}, 0),$$

де I_M^{σ} - згладжування I_M гаусовим ядром з невеликим σ .

Так як ця хвиля вже моделює вплив початкових та самих активних етапів росту організму. А для більш пізніх ріст не такий великий $\frac{1}{w(t)^2} \approx \text{constant}$ тому, при такому методі ми можемо спростити наше рівняння до:

$$\frac{d\mathbf{u}}{dt} - (1 + k * g(t, x, y))A(\mathbf{u}, \Delta\mathbf{u}) = \mathbf{f} + fl_{\varepsilon}.$$

РОЗДІЛ III Опис методів та алгоритмів

3.1 Теорія QQISA(Кількісно-Якісного інтегрованого аналізу стійкості)

Звичайно використання СНМДРЧП є крайнім випадком - коли є потреба в отриманні якісних моделей але майже немає даних.

Є й більш прості способи отримувати і оптимізувати складні моделі.

А так як прості способи моделювання легше інтерпретувати перед використанням основного функціоналу теорії QQISA ми маємо використати більш прості - щоб зрозуміти яку інформацію ми можемо легко отримати з різних стійких зон в просторовій структурі u та їх флуктуацій.

А тому вводиться наступна методологія:

- 1) Використання символічних обчислень для того щоб отримати модель та обмеження для системи.
- 2) Досліджень структури стійких станів та їх флуктуацій
- 3) Розв'язання задачі оптимізації СНМДРЧП через основний функціонал наших методів - градієнтний спуск для СНМДРЧП
- 4) Інтерпретація

Хоч і використання диференційних рівнянь та фізичних обмежень в кількісному моделюванні і можуть назвати інтеграцією якісного та кількісного моделювання але це тільки перші кроки в інтеграції.

Більш просунутим методом є не сліпе використання фізичних або інших обмежень будови а розуміння звідки вони беруться. Зазвичай це впливає з деяких властивостей оптимумів. Так як ми хочемо отримати можливість покращити глибоку адаптацію організму ми маємо навчитися розуміти властивості організму лише по тому як саме він підтримує свій

стійкий стан. Або в ще більш складній постановці - розуміти динаміку організму лише по одній картинці його флуктуацій в оптимумі.

Часто розподіл значень компонентів u_k біологічної динаміки \mathbf{u} може бути не неперервним. Майже завжди для будь-якої компоненти є деяка константа для якої є досить велика область в якій значення цієї компоненти знаходяться біля цієї константи. Це явище також і характерне для великого класу нелінійних ДРЧП. Для того щоб пояснити цей феномен потрібно розглянути динаміку в неперервній підобласті з динамікою близькою до вектора c_1 .

Назвемо цю область M_{c_1} , для неї маємо:

$$\begin{aligned}\mathbf{u}(t, x, y) &= c_1 + \boldsymbol{\varepsilon}(t, x, y) \cong c_1, \\ \Delta \mathbf{u}(t, x, y) &= \Delta \boldsymbol{\varepsilon}(t, x, y) \cong 0 \\ A(c_1, 0) &= 0\end{aligned}$$

Використавши наші умови стійкості маємо що:

$$\begin{aligned}\text{COND}_{A_1', A_2'} &= \left(-2\alpha + \min_{l \geq 0} \text{Re}(\lambda_{-A_1' + lA_2'}^l) \right) \\ \text{COND}_{A_1', A_2'}(c_1 + \boldsymbol{\varepsilon}, \Delta \boldsymbol{\varepsilon}) &\geq 0\end{aligned}$$

Ці умови стійкості в області M_{c_1} можна наблизити:

$$\begin{aligned}\text{COND}_{A_1', A_2'}^{\alpha_1, \alpha_2} &= \text{COND}_{A_1' + \alpha_1, A_2' - \alpha_2} = \left(-\alpha_1 - 2\alpha + \min_{l \geq 0} \text{Re}(\lambda_{-A_1' + l(A_2' - \alpha_2)}^l) \right) \\ \text{COND}_{A_1', A_2'}^{\alpha_1, \alpha_2}(c_1, 0) &\geq 0\end{aligned}$$

де $\alpha_1, \alpha_2 > 0$ - невеликі константи що більші за $\|\boldsymbol{\varepsilon}\|_{C(M_{c_1})}$ та $\|\Delta \boldsymbol{\varepsilon}\|_{C(M_{c_1})}$.

Таким чином маємо деяку ієрархію областей в стійкому оптимумі, а на нижчому рівні є деякі підобласті які ведуть себе як самостійні прості стійкі точки та під-області з більш складною динамікою які їх “зв’язують”.

Підобласті по типу стійкої точки найбільш гарно підходять так як їх зазвичай багато і в них можна дослідити як і самі ці точки так і розподіл флуктуацій що йдуть від цих точок.

При маленькій кількості вхідної інформації це досить важливо і дає можливість зрозуміти вплив тих компонент динаміки які ми на даний момент не можемо виміряти.

Такий спосіб аналізу ми назвемо “аналізом флуктуацій стійких під-областей”. Хоч це і набагато більш примітивний спосіб ніж дослідження через відстані Сінкхорна, що є наступним етапом, але ми можемо працювати з задачами лінійної регресії що дозволяє нам зрозуміти які дані ми можемо отримати без використання складних методів а також отримати зрозумілу інтерпретацію методу збору цих даних.

Також області по типу стійкої точки можуть мати трохи різну динаміку - деякі мають постійні флуктуації пов’язані з зовнішніми впливами, а деякі мають прошарок що захищає їх від зовнішніх впливів. Але при цьому значення \mathbf{u} на цьому прошарку все одно близькі до c_1 .

Найбільш загальним алгоритмом для 2 етапу будемо вважати:

- 1) Аналіз областей в яких при відомому \mathbf{u} можна буде знайти цільові параметри.
- 2) Аналіз флуктуації в точках та вибір корисних.
- 3) Отримання деяких параметрів по флуктуаціям.
- 4) Отримання наближення невідомих компонентів в області описаній в пункті 1).
- 5) Отримання цільових параметрів.

Також ми будемо вважати “Аналіз флуктуацій стійких під-областей” самостійним методом ідентифікації коефіцієнтів в СНМДРЧП.

3.2 Можливості застосування теорії QQISA як загальної теорії що відповідає на потреби багатомасштабного моделювання в біології та інших складних систем

Тепер потрібно описати те як теорія QQISA відповідає на найбільш важливі питання в дисциплінах які займалися схожими задачами

Для цього є дуже релевантна стаття[3]. Вона є підсумком симпозиуму 2019 року з інтеграції машинного навчання з багатомасштабним моделюванням для біологічних, біомедичних і поведінкових систем (ML-MSM) у рамках міжвідомчої групи моделювання та аналізу (IMAG).

І в ній задається найбільші проблеми та виклики що є в задачах по багато-масштабній біології. Опишемо їх усі та те як теорія QQISA відповідає на них.

3.2.1 Управління неправильно поставленими проблемами

Для початку потрібно визначити основні випадки проблеми.

Органи є тканинами що згорнуті одна на іншій в єдину функціональну систему, вони часто можуть мати крайові умови близькі к циклічним(тор) або розвиватися на якомусь іншому обмеженому топологічному об'єкті(куля, петлі і т.п.). А тому в нормальному для організмі стану проблеми некоректності крайових умов не виникає(вплив сусідніх органів можна моделювати як впливи зовнішніх сил що регулюються патернами більш високого масштабу).

Але при пошкодженнях або при описі всього організму це має сенс.

Для моделювання розвитку організму потрібно використовувати як крайові умови, для деякої множини В:

$$\frac{\partial \mathbf{u}}{\partial \mathbf{n}} = 0$$

Для моделювання розвитку патернів краще всього використовувати як область прямокутник $[0, a] * [0, b]$ і задати на ньому циклічні умови:

$$\mathbf{u}(t, 0, y) = \mathbf{u}(t, a, y)$$

$$\mathbf{u}(t, x, 0) = \mathbf{u}(t, x, b)$$

У випадку невизначених умов ми можемо почати з нульових а розв'язок так щоб вони дали мінімізацію для:

$$\mathbf{F}(\mathbf{u}, A) + K * ||\mathbf{b}||^2,$$

де $K > 1$.

Також скоріше за всього потрібні самі обмеження на можливі крайові функції \mathbf{b} так щоб вони були узгоджені з зовнішнім середовищем.

Також в майбутньому можна буде взяти як крайові умови:

$$\frac{\partial \mathbf{u}}{\partial n} = \mathbf{b} + \mathbf{e},$$

де \mathbf{e} - випадкове гаусове поле з невеликою дисперсією. Таким чином можна буде зрозуміти як тіло оптимізує зовнішні умови в рамках своїх можливостей так щоб мінімізувати проблеми які виникають із-за шуму.

3.2.2 Виявлення відсутньої інформації, створення сурогатних моделей

Ми поступово розширюємо кількість додаткових компонентів і таким чином можемо прикинути які мінімальні додаткові описи ми маємо додати щоб отримати модель що відповідає реальності.

В практичній частині ми покажемо що створити сурогатні моделі можна всього лише по одній реальній картинці патерна.

3.2.3 Дискретизація простору та часу

Дискретизація простору та часу може мати декілька негативних наслідків - відсутність опису на мікромасштабі та проблеми з описом на макромасштабі.

Для першого випадку пропонується параметризувати стан мікромасштабу - а тому відносно незалежні ефекти мікромасштабу, що можуть впливати на макромасштаб, будуть описані через додаткові компоненти.

Для другого випадку - при великих значеннях дійсної частини власних чисел матриці крос-дифузії маємо, що високочастотні хвилі легко пригнічуються, а тому дискретна сітка буде мати майже всю інформацію про просторову будову.

3.2.4 Перекриття масштабів, розуміння виникнення функції

1. Через теорію або обрахунки ми розуміємо в яких межах ми можемо керувати мікроскопічним масштабом
2. Використовуємо ці обмеження при моделюванні макромасштабу отримання цільової функції для мікромасштабу
3. Підгонка мікромасштабу під цю функцію

4. Моделювання для перевірки того що підгонка спрацювала
5. якщо ні, то потрібно ще більше уточнити обмеження та повторити пункти 2-4.
6. Доповнення навчальних даних

Можливо, якщо мікрорівень буде досить складний, то не буде простого способу отримати формулу обмеження на можливу динаміку.

Але швидше за все можна буде звужити область.

Також можна буде навчити класифікатор який намагатиметься відокремлювати динаміку згенеровану через випадкові зміни мікрорівня від випадкових функцій що всього лише задовольняють формулі.

Значення цього класифікатора можна використовувати для створення додаткового штрафу на “неправдоподібний” оператор що описує динаміку в мікромасштабі.

3.2.5 Кількісна невизначеність

Ми працюємо з стохастичними процесами а тому в нас натуральним чином є кількісна невизначеність яку ми можемо контролювати та оптимізувати ще на етапі навчання моделі.

За допомоги відстаней Сінкхорна $S_\varepsilon(\alpha, \beta)$ ми можемо не просто ідентифікувати деякі кореляції а й ідентифікувати та оптимізувати будь-який розподіл невизначеностей.

А вже після цієї оптимізації можна й подивитися на більш прості елементи поведінки між початковими на кінцевими невизначеностями.

Щоб побачити цю невизначеність можна використати розклад Кархун-Лоева одночасно для початкового і кінцевого значення. Цей розклад є нескінченномірним аналогом методу головних компонент, з нього ми можемо отримати вектори що роблять основний вклад в дисперсію. А з цього

ми маємо і будову взаємозв'язків між початковими та кінцевими невизначеностями.

Її можна використати для відхилень в якісному та звичайному просторі що залишилися після спроби організму стабілізуватися - зазвичай простір таких відхилень не такий великий і цю множину відхилень легко апроксимувати функціями що описані через 10 компонент[10].

3.2.6 Вивчення складних просторів структур, підвищення точності та відтворюваності

Завдяки додатковим якісним просторам які ми використовуємо ми можемо спростити опис того що нам потрібно. Отримати якісні характеристики з правил динаміки системи не так складно, а отримання правил динаміки системи з якісних характеристик - це і є розв'язання нашої основної задачі.

Тому простори якісних характеристик в сукупності в якомусь сенсі кодують різні стохастичні процеси.

3.2.7 З'ясування механізмів

Моделювання це лише один із можливих способів використання цієї системи. Після створення моделі ми можемо відразу перейти до їх оптимізації.

Звичайно деякі поверхові механізми та кореляції ми можемо вивести для інтерпретації результатів, але для справжнього лікування майбутнього цього буде мало і нам буде потрібно розв'язувати задачу оптимального керування для зміни довгострокових параметрів системи що зможуть не тількивилікувати організм але й дати організму довгострокову стійкість та якість життя.

3.2.8 Використання біологічно натхненого навчання

Як я писав у статті [9] створюючи блоки(FNN, RNN, CNN, RCNN і т.д.) та архітектури нейромережі ми часто не тільки копіюємо у природи а іноді і вгадуємо деякі глибокі характеристики нашого мозку намагаючись зробити ШІ більш вправним. Можливо розуміння того як робити стійку динаміку так щоб вона розпізнавала зовнішні загрози ми зможемо краще зрозуміти ту базову нейромережу яка вкладається в нас з народження і яку ми вже під час свого існування донавчаємо.

3.2.9 Запобігання перенавчанню, мінімізація зміщення даних

Для того щоб в нас не було перенавчання ми беремо рівняння редакції-дифузії та обмежуємо наш оператор і даємо невелику кількість параметрів які мають описати систему але так щоб вони могли описати функції з різними властивостями і будовою що можуть допомогти описати систему.

3.3 Отримання одночасно більш простих і більш якісних моделей

Хоч і ми отримали метод щоб будувати моделі різної складності але є також підстави що він може знаходити гарні та прості моделі що ще не відкрили люди.

Це пов'язано з тим що зараз в рамках цих простих моделей найкращі конфігурації шукаються ручним перебором що часто може пропустити багато цікавих фаз які мають в просторі параметрів порівняно з іншими не такий великий об'єм але достатній щоб мати в собі кулю з таким радіусом який буде запобігати повній поломці системи від невеликих впливів на параметри.

Умовно можна поділити сучасні системи біологічного моделювання на якісні та кількісні.

Якісні намагаються дати інтерпретацію коефіцієнтам диференційного рівняння щоб дати пояснення на які характеристики системи вони впливають а також це допомагає людям зменшити розмірність задачі пошуку і звести до такої в якій вони можуть спробувати ручний перебір.

Кількісні намагаються отримати модель поведінки системи по даним з відсутньою або мінімально присутньою теорією походження цих даних.

По рівню використання теорії можна описати такі моделі:

- 1) звичайні RNN або CRNN нейромережі що прогнозують наступний стан по попереднім
- 2) Отримання диференційних рівнянь по даним
- 3) Отримання диференційних рівнянь по даним так щоб дотриматися певних фізичних обмежень.[19][20][21][22][23][24]

Наші методи є продовженням цих і розраховані на випадки коли ці методи не працюють або їх не можна використати.

Тому ми отримуємо додаткових 2 рівні які раніше не використовувалися:

- 4) Аналіз зон в будові *u* де за допомогою методів 1-3 рівня можна дістати якусь інформацію
- 5) Автоматизація та якісне покращення 4 рівня через аналіз структури розподілу взаємозв'язків та флуктуацій в *u* та того яка обґрунтована стохастична фізична модель може їх утворити.

В нашій методології ці рівні пов'язані і є основою другого та третього етапу методології QQISA.

РОЗДІЛ IV Практична частина

4.1 Дослідження дії оператора

Для перевірки методології ми виберемо досить прості моделі, а тому більша частина їх опису вже є в різних роботах[7][8], завдяки чому ми зможемо порівняти якість моделей та теорій.

Покладемо:

$$\mathbf{u} = (u_1, u_2)$$

$$\frac{\partial \mathbf{u}}{\partial t} - (1 + k * g)A(\mathbf{u}, \Delta \mathbf{u}) = f$$

$$A(\mathbf{u}, \Delta \mathbf{u}) = (K, l(\mathbf{u})),$$

Де $(K, l(\mathbf{u}))$ -множення матриці коефіцієнтів K на вектор $l(\mathbf{u})$.

Ми розглянемо дві задачі:

$$1) \quad l_1(\mathbf{u}, \Delta \mathbf{u}) = (u_1, u_2, \frac{1}{w(t)^2} \Delta u_1, \frac{1}{w(t)^2} \Delta u_2, u_1 * u_2, 1)$$

$$2) \quad l_2(\mathbf{u}, \Delta \mathbf{u}) = (u_1, u_2, \frac{1}{w(t)^2} \Delta u_1, \frac{1}{w(t)^2} \Delta u_2, u_1 * u_2, u_1^2 * u_2, 1)$$

Є дослідження які виводять формули для області стійкості, тому ми можемо не обраховувати їх самостійно. З чого маємо що в цьому спрощеному випадку ми можемо пропустити перший етап нашої методології.

Якщо динаміка \mathbf{u} веде функцію до її стійкого стану, то в ньому:

$$\frac{\partial \mathbf{u}}{\partial t} \cong 0.$$

Тому, якщо в нас є переріз динаміки $\mathbf{u}(T, \cdot, \cdot)$ для якої ми впевнені що динаміка вже дуже близько до стійкої точки, то для знаходження коефіцієнтів нашого диференційного рівняння ми могли би просто розв'язати задачу лінійної регресії.

$$\|A_K(\mathbf{u}, \Delta \mathbf{u})\|_{L_2(B)}^2 = \|(K, l((u_1, u_2), (\Delta u_1, \Delta u_2)))\|_{L_2(B)}^2 \rightarrow \min$$

Але основна проблема в тому що в нас не завжди є вся інформація про стан об'єкта, наприклад коли ми бачимо візерунок шкіри тварини ми не бачимо які саме хімічні речовини підтримують цей візерунок, але їх динаміка також є важливою.

Але проблема в тому що в нас немає одного з компонентів, а тому задача пошуку коефіцієнтів попереднім способом стає некоректною.

Звичайно якщо знати деякі характеристики цієї множини то можна б було розв'язати таку задачу:

$$|| (K, l((u_1, v), (\Delta u_1, \Delta v))) || \rightarrow \min, \text{ при } F(v) = h$$

Але так як в нас мало знань про невідомі компоненти ми використаємо метод для 2-ого етапу що ми описали в “3.1. Теорія QQISA”.

Маємо малюнок патерну:



Рисунок 1 - Фото патерну

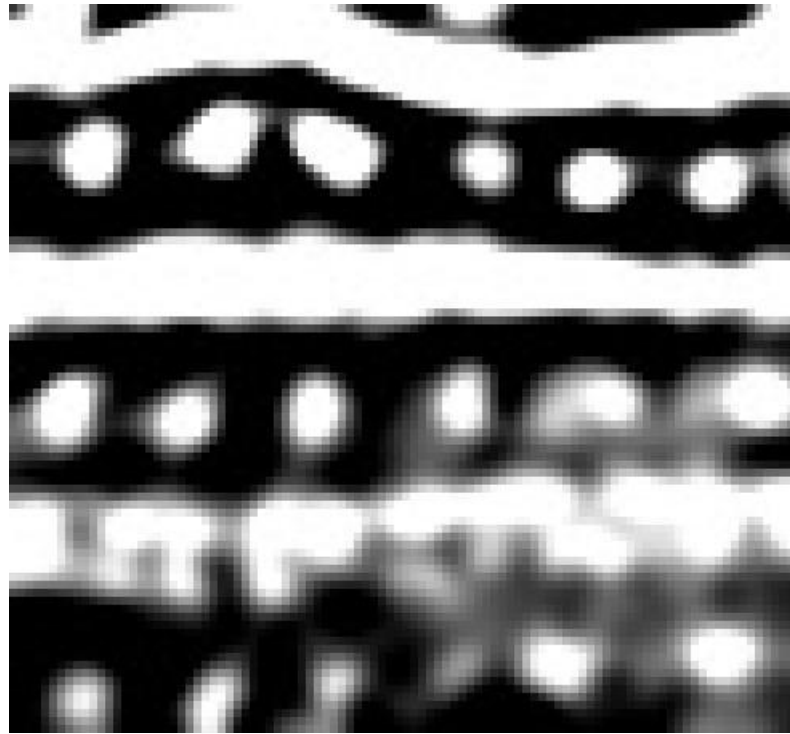


Рисунок 2 - Фото патерну після фільтрації від тіней

По малюнку ми відразу бачимо області з однаковими значеннями - саме вона відображає в собі нелінійну просторову динаміку та явища дифузії. Саме в цій області ми маємо знайти значення невідомих компонентів.

Для того щоб перевірити наш метод ми вибираємо відому модель для цього патерну яка описана в [8]:

$$\frac{\partial u_1}{\partial t} - (16 - u_1 * u_2 + d_1 \frac{1}{w(t)^2} \Delta u_1) = 0$$

$$\frac{\partial u_2}{\partial t} - (u_1 * u_2 - u_2 - 12 + d_2 \frac{1}{w(t)^2} \Delta u_2) = 0$$

З обмеженнями $u_1, u_2 \in [0, 7.2]$

Також в цій роботі було показано як багато складних патернів можна описати для випадків з різним ростом області.

Так як ріст області дуже повільний $\frac{1}{w(t)^2}$ можна розглядати як фіксований параметр під час якого система приходить до стійкого стану.

Тобто, достатньо мати по одному такому стійкому стану на кожен етап росту і якщо розв'язати для них задачу пошуку дифузійних параметрів то можна буде відновити весь процес генерації патерну.

Тобто задача пошуку динаміки росту області та значення дифузійних параметрів зводиться до декількох задач пошуку d_1, d_2 :

$$\begin{aligned} 16 - u_1 * u_2 + d_1 \Delta u_1 &\cong 0 \\ u_1 * u_2 - u_2 - 12 + d_2 \Delta u_2 &\cong 0 \end{aligned}$$

Ми будемо вважати що максимум та мінімум на нашій картиці відповідають $u_{\min} = 0, u_{\max} = 7.2$. Також ми вважаємо що картинка відповідає за значення u_2 , значення u_1 для нас є невідомим.

Наша початкова ціль знайти значення u_1 в області $u_{\min} + \varepsilon < u_2 < u_{\max} - \varepsilon$, назвемо її M_I .

Для цього ми розглянемо область прошарку для $u_2 = 0 : u_{\min} + \frac{\varepsilon}{10} < u_2 < u_{\min} + \varepsilon$, назвемо її M_S .

В цій області ми маємо що $u_2 \ll \Delta u_2$, з чого отримуємо задачу лінійної регресії для d_2 :

$$\| -12 + d_2 \Delta u_2 \|_{L_2(M_S)}^2 \rightarrow \min$$

Якщо нам відоме середнє значення $M[u_1] = m_1$ то можна його використати: і відповідь стане краща:

$$\| (m_1 - 1)u_2 - 12 + d_2 \Delta u_2 \|_{L_2(M_S)}^2 \rightarrow \min$$

Але так як ми вибрали варіант без попередніх знань про u_1 то ми використаємо результат попередньої формули. Тепер ми маємо деяке наближення для коефіцієнту дифузії 2-ого компоненту \bar{d}_2 .

Ми вже можемо спробувати знайти u_1 в області M_I , для цього ми мінімізуємо наступний вираз:

$$\|(u_1 - 1)u_2 - (12 - d_2\Delta u_2)\|_{L_2(M_I)}^2 + \varepsilon\|\nabla u_1\|_{L_2(M_I)}^2 \rightarrow \min$$

Після того як ми його мінімізували і отримали \bar{u}_1 ми обріжемо значення які вийшли за межі $[u_{\min}, u_{\max}]$, після чого згладимо отриманий результат гаусівським ядром з маленьким σ :

$$\bar{u}_1 = \langle \min(\max(\bar{u}_1, u_{\min}), u_{\max}), G_{\sigma^2} \rangle$$

Основну проблему з пошуком значень u_1 в області M_I ми виконали. А тому ми можемо використати ці значення для отримання остаточної відповіді, для цього вирішемо наступні задачі регресії для коефіцієнтів d_1 та d_2 :

$$\|16 - u_1 * u_2 + d_1\Delta\bar{u}_1\|_{L_2(M_I)}^2 \rightarrow \min$$

$$\|\bar{u}_1 * u_2 - u_2 - 12 + d_2\Delta u_2\|_{L_2(M_I)}^2 \rightarrow \min$$

З чого ми отримаємо наші коефіцієнти дифузії:

$$d_1, d_2 = 35.85, 3.396$$

Використаємо лінійне наближення для функції росту домена $w(t) = 0.5 + 0.5\frac{t}{T}$

Та виведемо її для часу $\frac{T}{8}, \frac{T}{4}, \frac{T}{2}, T$:

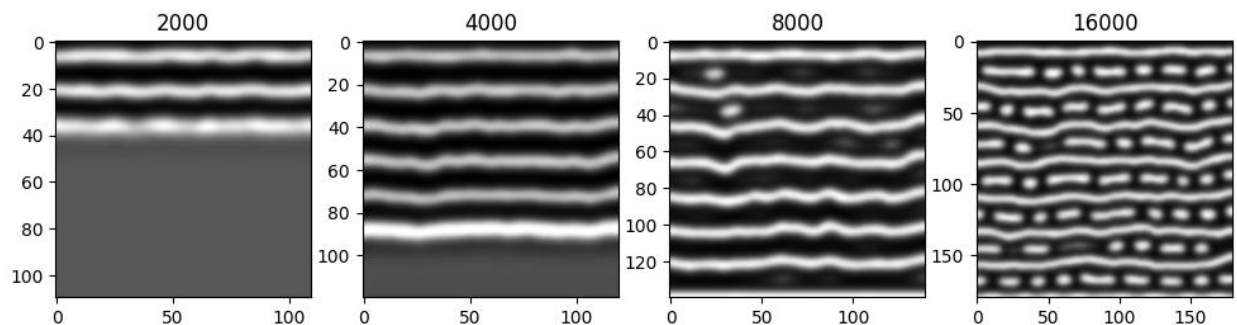


Рисунок 3 - Ріст патерну при зростаючому домені

Для порівняння покажемо результат в статті де ці параметри підбирали вручну[8]:

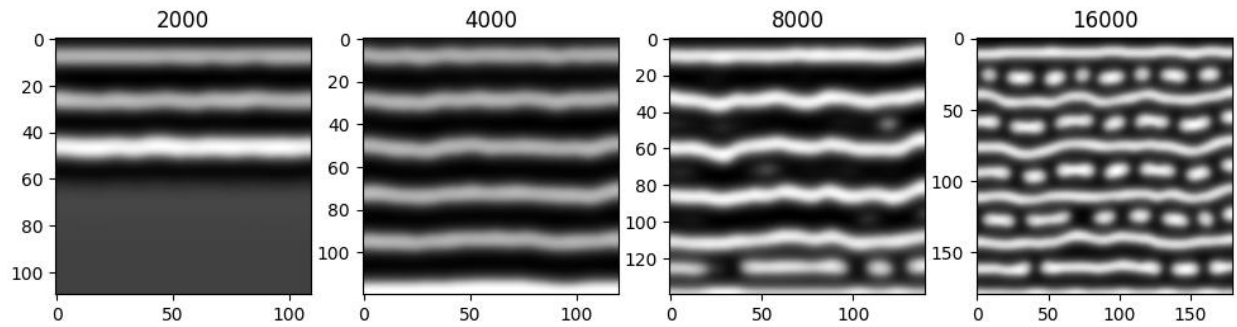


Рисунок 4 - Ріст патерну *Ictidomys tridecemlineatus* в [8]

Бачимо що наш варіант приблизно схожої якості.

Як видно 2-ий етап методології QQISA дав нам можливість краще зрозуміти будову u та дав нам алгоритм який ми можемо використовувати для того щоб ми могли, всього завдяки декільком знімкам всього одного компонента нашої динаміки, вибудувати весь процес росту та утворення патерна.

Але цей алгоритм погано масштабується у складності можливих патернів, а також його складно перенести в область багатомасштабної біології для якої потрібно описувати різні рівні патернів та структур разом з їх взаємодіями.

Крім цього всього, отриманий патерн, насправді, досить не оптимальний і завдяки оптимізації СНМДРЧП ми можемо отримати навіть більш кращий патерн при цьому використовуючи лише один малюнок одного компонента нашого патерну.

4.2 Моделювання патернів шкіри

Як вже було сказано в “2.9.2 Вирішення проблеми неоднозначності рішень” щоб мати можливість прибрати неоднозначність та для того щоб спростити модель росту до такої де не потрібно моделювати ріст домену ми

маємо ввести невеликий “вплив етапів з маленьким масштабом” а саме постійний маленький вплив однією хвилиною.

Крім того для моделювання A ми виберемо модель з невеликою кількістю параметрів:

$$A(\mathbf{u}, \Delta \mathbf{u}) = (K, l_2(\mathbf{u}, \Delta \mathbf{u})),$$

а тому ми можемо не робити додатковий штраф на розмір коефіцієнтів K .

Це дає нам змогу спростити нашу задачу до наступної:

$$\frac{d\mathbf{u}}{dt} - (1 + k * \mathbf{g}(t, x, y))(K, l_2(\mathbf{u}, \Delta \mathbf{u})) = \mathbf{f} + fl_\varepsilon$$

$$M \left[\int_T^{T+T_2} PD(P(\mathbf{u})(T, \cdot, \cdot), s_T) dt + \|\text{relu}(|\mathbf{u}(T, \cdot, \cdot)| - 1)\|_{L_2(B)}^2 \right] \rightarrow \min$$

Де PD - це відстань в просторі патернів яка введена в розділі “2.9.1. Отримання якісних характеристик”, а GPD - градієнт від цієї відстані.

З “2.7. Модифікації градієнту” отримаємо алгоритм для градієнтного спуску. Отримуючи з нього послідовність коефіцієнтів: K_1, K_2, K_3, \dots

В нього на i -тому кроці ми будемо підставляти наступний градієнт:

$$\begin{aligned} \mathbf{grad}_K &= (l_2(\mathbf{u}, \Delta \mathbf{u}), \mathbf{grad}_A) \\ \mathbf{grad}_A &= \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{grad}_{A_j} = \frac{1}{N_i} \sum_{j=1}^{N_i} (1 + k * \mathbf{g}_{i,j}(t, x, y)) c_j, \end{aligned}$$

де c_j обраховується через моделювання на сітці наступного лінійного ДРЧП:

$$\begin{aligned} \frac{\partial c}{\partial t} - (1 + k * \mathbf{g}(t, x, y)) (A'_1 c + \frac{1}{w(t)^2} A'_2 \Delta c) \\ = I_{[T, T+T_2]}(t, x, y) GPD(P(\mathbf{u})(t, \cdot, \cdot), s_T) \\ c(T + T_2, x, y) = 0, \frac{\partial c}{\partial n} = 0 \end{aligned}$$

Де A'_1, A'_2 обраховуються наступним чином:

$$A'_1 = (K_i, l_{2_1}') = (K_i, \begin{pmatrix} 1 & 0 & 0 & 0 & u_2 & 2u_1u_2 & 0 \\ 0 & 1 & 0 & 0 & u_1 & u_1^2 & 0 \end{pmatrix})$$

$$A'_2 = (K_i, l_{2_2}') = (K_i, \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}) = D$$

Ці похідні обраховуються в програмі автоматично так що для запуску алгоритму потрібно тільки ввести формулу для A та картинку з цільовим патерном. Як видно A'_2 - у нас фіксована матриця 2×2 що відповідає за коефіцієнти дифузії.

А \mathbf{u} обраховується як рішення на сітці нашого рівняння для K_i :

$$\frac{d\mathbf{u}}{dt} - (1 + k * \mathbf{g}(t, x, y))(K_i, l_2(\mathbf{u}, \Delta\mathbf{u})) = \mathbf{f} + \mathbf{f}_\varepsilon$$

$$\mathbf{u}(0, x, y) = s_0(x, y), \frac{\partial \mathbf{u}}{\partial n} = \mathbf{b} = 0, (x, y) \in \partial B$$

Для обрахування значень в обох диференційних рівняннях для s та u ми використаємо одну й ту ж сітку 100 на 100, $B = [0,10] \times [0,10]$, просторові кроки $h_1, h_2 = \frac{1}{10}$. Тепер підберемо часовий крок.

Проведемо аналіз стійкості для диференційних схем для лінеаризованого випадку, маємо:

$$-1 + O(\tau) < \text{Re} \left(\Lambda \left(E - \left(\frac{4\tau}{h_1^2} + \frac{4\tau}{h_2^2} \right) A'_2 \right) \right)$$

З чого маємо:

$$\max(\text{Re}(\lambda_i)) \left(\frac{2\tau}{h_1^2} + \frac{2\tau}{h_2^2} \right) < 1$$

$$\max(\text{Re}(\lambda_i)) < \frac{1}{2\tau \left(\frac{1}{h_1^2} + \frac{1}{h_2^2} \right)},$$

де λ_i -власні числа матриці A'_2 .

В нашому спрощеному випадку A_2' - це фіксована матриця D і відповідає за коефіцієнти дифузії.

Підберемо τ так щоб $2\tau(\frac{1}{h_1^2} + \frac{1}{h_2^2}) = 1$, маємо:

$$\tau = \frac{1}{2(\frac{1}{h_1^2} + \frac{1}{h_2^2})} = \frac{h^2}{4} = \frac{1}{400}.$$

$$0 \leq \text{Re}(\lambda_i) < 1$$

З чого маємо обмеження на матрицю D :

$$0 \leq \text{Re}(\lambda_D^i) < 1$$

Також ми візьмемо $[T, T + T_2] = [\frac{32}{400}, \frac{256}{400}]$.

На цьому інтервалі ми будемо навчати наш випадковий процес прямувати до цільового стану. Але нам не обов'язково щоб він стабілізувався на цьому етапі. Додатковими запусками на попередніх станах так щоб він “підтримував” та не “руйнував” стан для якого ми знаємо що він в реальності стійкий. Це дає нам можливість паралельно “донавчати” алгоритм стійкості.

Таким чином ми виконаємо 10000 ітерацій для нашого алгоритму. Також для випадку патерну ми зробимо початкові умови також випадковим шумом. Для симуляції Монте-Карло на кожному кроці ми будуватимемо $N_i = 8$ симуляцій для f , g та s_0 . Із-за цього наша функція оцінки також буде стохастичною.

Навчивши наш алгоритм маємо наступні значення для $F(\mathbf{u}, A)$:

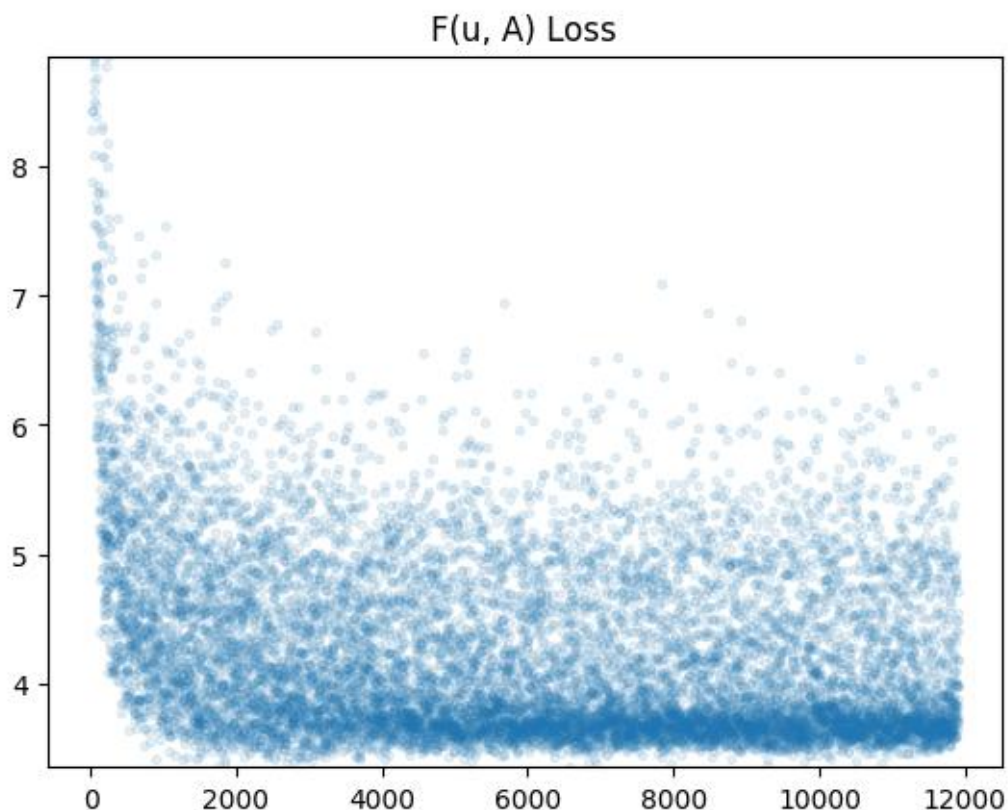


Рисунок 5 - Розкид $F(\mathbf{u}, A)$ на кожній ітерації

Ми отримуємо матрицю коефіцієнтів:

$$K = \begin{pmatrix} -1.1 & 0.6 & 0.6 & -1. & 6.6 & -11.7 & -1.1 \\ 13.6 & -2.2 & 0.2 & -0.2 & -6.8 & -69.3 & -0.6 \end{pmatrix}^T$$

Бачимо, що:

$$D = \begin{pmatrix} 0.6 & -1. \\ 0.2 & -0.2 \end{pmatrix},$$

що має комплексні власні значення $\lambda_1 = 0.2 \pm 0.2i$.

Якщо прибрати випадкові функції, то ми отримає наступне диференціальне рівняння в частинних похідних:

$$\frac{\partial u_1}{\partial t} = (-1.1 - 1.1u_1 + u_2(0.6 + 6.6u_1 - 11.7u_1^2) + 0.6\Delta u_1 - 1. \Delta u_2)$$

$$\frac{\partial u_2}{\partial t} = (-0.6 + 13.6u_1 + u_2(-2.2 - 6.8u_1 - 69.3u_1^2) + 0.2\Delta u_1 - 0.2\Delta u_2)$$

Промодельюємо нашу модель та проінтрапретуємо її етапи:

1) Спочатку вона організовує шум в маленькі деталі структур які відповідають динаміці патерну на маленькому масштабі:

$$t_{1,2,3} = \frac{32}{400}, \frac{64}{400}, \frac{128}{400}$$

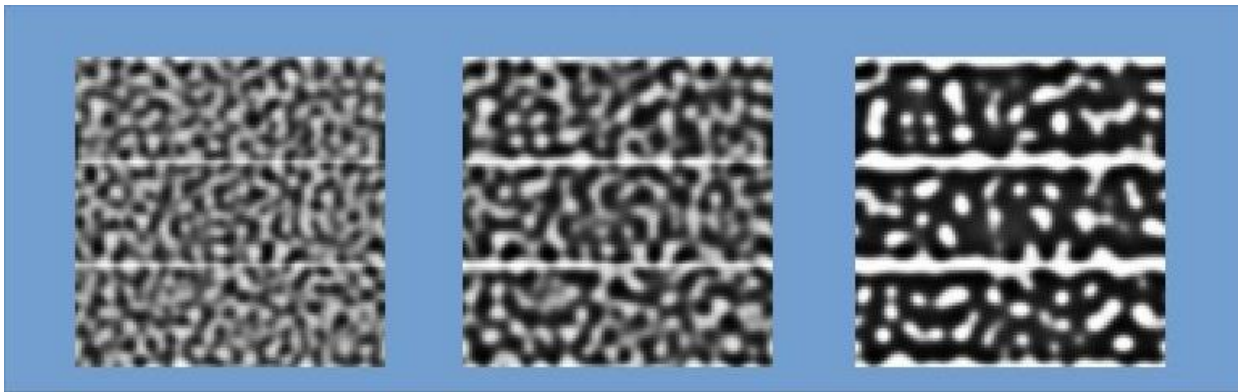


Рисунок 6 - Стадія формування локальних патернів

2) Далі вона будує з цих патернів на маленькому масштабі патерни на більших масштабах:

$$t_{4,5,6} = \frac{256}{800}, \frac{512}{800}, \frac{1024}{800}$$

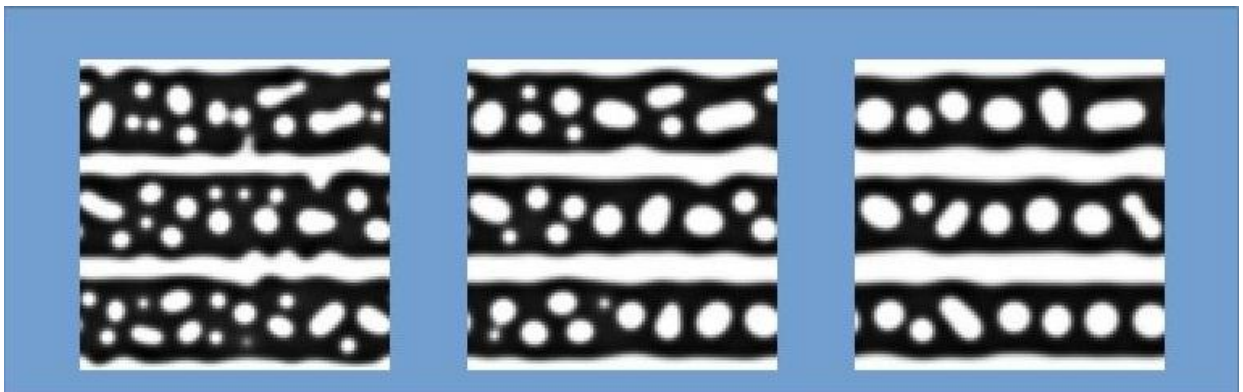


Рисунок 7 - Стадія формування глобального патерна

3) Якщо не зупиняти модель то вона стабілізує ці паттерни:

$$t_{7,8} = \frac{2048}{400}, \frac{4096}{400}$$

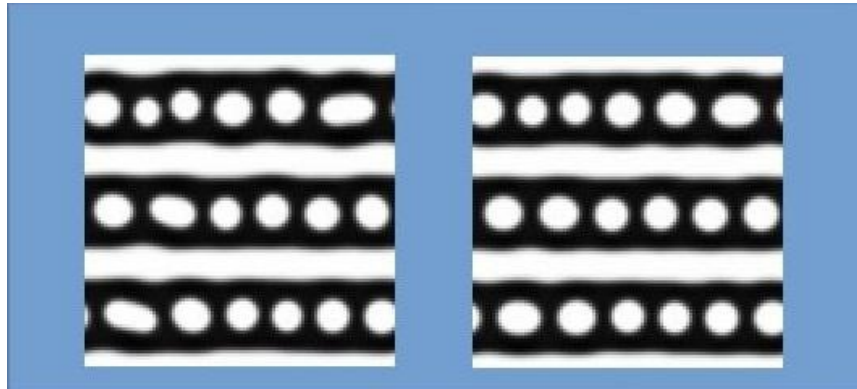


Рисунок 8 - Стадія стабілізації

Як видно найпростіші методи нашої теорії вже автоматизують створення простих патернів, яке раніше було ручною працею. Але більш просунуті створили більш кращий патерн навіть при тому що використовували менше різних параметрів таких наприклад як моделювання функції росту області $w(t)$ через ламану лінію або інші додаткові маніпуляції.

На відміну від минулих моделей[8] тут набагато краще видніше округлість плям на відміну від смужок що були раніше.

На даний час це найкраща модель по якості отриманого малюнка з тих що є для моделювання патерну шкіри суслика *Ictidomys tridecemlineatus*.

Наша модель описується кросс-дифузєю, причому такою в якій лінеаризація для коефіцієнтів при дифузійних компонентах дає матрицю з комплексними значеннями. Для того щоб отримати таку модель потрібно використовувати негативну кросс-дифузью яка мало досліджена. Досить мало робіт які досліджують саме можливості формування більш складних патернів за допомоги такого опису[15][16]. І більшість робіт які є на даний час фокусуються в основному на пошуку області стійкості. Є зрозумілим

причини чому подібне моделювання дає можливість більшої адаптації для патернів та можливість одночасного утримання декількох мод[17].

Але на даний час майже немає робіт які б використовували матриці коефіцієнтів для дифузійних компонентів які б при лінеаризації мали комплексні власні значення, та цим питанням не особливо переймалися дослідники. Але як я вважаю це може бути одним із ключових областей для опису того як моди коливань впливають одна на іншу, та інших більш глибоких ефектів які ще потрібно відкрити. В наступних роботах я дам більше пояснень чому це так.

Для того щоб пришвидшити процес сумісного дослідження подібних патернів ми дамо назву цій області моделювання - “комплексне моделювання патернів”.

ВИСНОВОК

Була вперше описана задача ідентифікації коефіцієнтів СНМДРЧП. Для неї був наданий математичний опис її розв'язку.

Було пояснено значимість цієї задачі для багатомасштабного моделювання в біології та для отримання моделі розвитку патернів органів та організму.

Було надано декілька методів для розв'язку цієї задачі:

- 1) Аналіз флуктуацій стійких під-областей
- 2) Градієнтний спуск для СНМДРЧП

Було показано що “Аналіз флуктуацій стійких під-областей” шукає параметри для моделі не гірше ніж це робить людина. Але все одно обмежена деякими спрощеннями які використовуються в існуючих теорія по моделюванню патернів.

Було показано що “Градієнтний спуск для СНМДРЧП” створює більш якісні моделі для патернів ніж це роблять люди вручну в існуючих роботах по моделюванню патернів шкіри.

Таким чином було створена нова теорія QQISA(quantitative and qualitative integrated stability analysis) яка має допомогти пришвидшити отримання більш якісних моделей, що в свою чергу дасть нам нове розумінні морфогенезу та множини тих лікувальних впливів на організм які не порушують існуючі процеси в органах та організмі.

ВИКОРИСТАНА ЛІТЕРАТУРА

- [1] Murray, J. D. (1990). Discussion: Turing's theory of morphogenesis—its influence on modelling biological pattern and form. *Bulletin of Mathematical Biology*, 52(1-2), 119-152.
- [2] Feydy, J., Séjourné, T., Vialard, F. X., Amari, S. I., Trounev, A., & Peyré, G. (2019, April). Interpolating between optimal transport and mmd using sinkhorn divergences. In *The 22nd International Conference on Artificial Intelligence and Statistics* (pp. 2681-2690). PMLR.
- [3] Alber, M., Buganza Tepole, A., Cannon, W. R., De, S., Dura-Bernal, S., Garikipati, K., ... & Kuhl, E. (2019). Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ digital medicine*, 2(1), 115.
- [4] Levin, S. (2003). Complex adaptive systems: exploring the known, the unknown and the unknowable. *Bulletin of the American mathematical Society*, 40(1), 3-19.
- [5] Mordvintsev, A., Randazzo, E., Niklasson, E., & Levin, M. (2020). Growing neural cellular automata. *Distill*, 5(2), e23.
- [6] Mordvintsev, A., & Niklasson, E. (2021). μ NCA: Texture Generation with Ultra-Compact Neural Cellular Automata. arXiv preprint arXiv:2111.13545.
- [7] Murray, J. D. (2001). *Mathematical biology II: Spatial models and biomedical applications* (Vol. 3). New York: Springer.
- [8] de Gomensoro Malheiros, M., Fensterseifer, H., & Walter, M. (2020). The leopard never changes its spots: realistic pigmentation pattern formation by coupling tissue growth with reaction-diffusion. *ACM Transactions on Graphics (TOG)*, 39(4), 63-1.
- [9] Chergykalo, D. O., and D. A. Klyushin. "Biomorphic artificial intelligence: achievements and challenges." *Enabling AI Applications in Data Science* (2021): 537-556.

- [10] Unser, M., Hugo Van hamme, De Muynck, P., Van Denhaute, E., & Cornelis, J. (1988, January). Karhunen-Loeve analysis of dynamic sequences of thermographic images for early breast cancer detection. In CVPR (pp. 592-596).
- [11] Madzvamuse, A., Barreira, R., & Gerisch, A. (2017). Cross-diffusion in reaction-diffusion models: Analysis, numerics, and applications. In Progress in Industrial Mathematics at ECMI 2016 19 (pp. 385-392). Springer International Publishing.
- [12] Enkavi, G., Javanainen, M., Kulig, W., Róg, T., & Vattulainen, I. (2019). Multiscale simulations of biological membranes: the challenge to understand biological phenomena in a living substance. *Chemical reviews*, 119(9), 5607-5774.
- [13] Weinan, E. *Principles of Multiscale Modeling*; Cambridge University Press: Cambridge, UK, 2011; ISBN 9781107096547
- [14] Benes, B., Guan, K., Lang, M., Long, S. P., Lynch, J. P., Marshall-Colón, A., ... & Turk, M. J. (2020). Multiscale computational models can guide experimentation and targeted measurements for crop improvement. *The Plant Journal*, 103(1), 21-31.
- [15] Jorné, J. (1975). Negative ionic cross diffusion coefficients in electrolytic solutions. *Journal of theoretical Biology*, 55(2), 529-532.
- [16] Satulovsky, J. E. (1996). Lattice Lotka–Volterra models and negative cross-diffusion. *Journal of theoretical biology*, 183(4), 381-389.
- [17] Iqbal, N., & Wu, R. (2019). Turing patterns induced by cross-diffusion in a 2D domain with strong Allee effect. *Comptes Rendus Mathematique*, 357(11-12), 863-877.
- [18] Valenza, G., Allegrini, P., Lanatà, A. & Scilingo, E. P. Dominant Lyapunov exponent and approximate entropy in heart rate variability during emotional visual elicitation. *Front. Neuroeng* 5, 3 (2012).
- [19] Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems

involving nonlinear partial differential equations. *J. Comput. Phys.* 378, 686–707 (2019).

[20] Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci.* 113, 3932–3937 (2016).

[21] Wang, Z., Huan, X. & Garikipati, K. Variational system identification of the partial differential equations governing the physics of pattern-formation: inference under varying fidelity and noise. *Comput. Methods Appl. Mech. Eng.* (2019). in press.

[22] Tartakovsky, A. M., Marrero, C. O., Perdikaris, P., Tartakovsky, G. D., & Barajas-Solano, D. Learning parameters and constitutive relationships with physics informed deep neural networks. Preprint at <http://arxiv.org/abs/1808.03398> (2018).

[23] Sanchez-Lengeling, B. & Aspuru-Guzik, A. Inverse molecular design using machine learning: Generative models for matter engineering. *Science* 361, 360–365 (2018).

[24] Yang, L., Zhang, D. & Karniadakis, G.E. Physics-informed generative adversarial networks for stochastic differential equations. Preprint at <https://arxiv.org/abs/1811.02033> (2018).

[25] Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics*, 14(9), 618-630.

ВІДГУК

на випускню кваліфікаційну роботу магістра
Чергикала Дениса Олександровича
«Об'єднання якісного та кількісного моделювання складної динаміки за
допомогою штучного інтелекту на прикладі автоматичного моделювання
патернів будови органів»

Роботу Д.О. Чергикала присвячено розробленню нових методів ідентифікації параметрів математичних моделей біологічних систем. Основним математичним апаратом, використаним у роботі, є стохастичне нелінійне багатокomпонентне рівняння в частинних похідних (СНМДРЧП).

Перші розділи роботи мають вступний характер, у них наведено математичну модель ембріогенезу й сформульовано допоміжні підзадачі – пошуку топології патернів, оцінювання стійкості тощо.

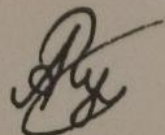
Третій розділ стосується запропонованої автором методології кількісно-якісного аналізу інтегрованого стану стійкості як математичного апарату багатомасштабного моделювання.

Четвертий розділ присвячено застосуванню зазначеної методології до моделювання патернів шкіри тварин. Результати не тільки не поступаються відомим із літератури, а й одержані в більш автоматизований спосіб. Загалом, це підтверджує наявність у роботі оригінального внеску в прикладну математику – способу пришвидшеної побудови якісних моделей патернів.

Результати, продемонстровані у практичній частині роботи, можна вважати такими, що становлять потенційний інтерес у науках про життя, передусім у розвитковій біології і її медичних застосуваннях.

На мою думку, Д.О. Чергикало виконав роботу на високому науковому рівні, вона відзначається творчим використанням апарату різних розділів математики, грамотним здійсненням обчислювальних експериментів і глибоким аналізом одержаних результатів. Вважаю, що робота заслуговує на оцінку «відмінно».

Асистент кафедри обчислювальної математики,
факультету комп'ютерних наук та кібернетики
Київського національного університету імені Тараса Шевченка,
канд. фіз-мат. наук



Андрій ГУЛЯНИЦЬКИЙ

Рецензія

на кваліфікаційну роботу магістра на тему:

«Об'єднання якісного та кількісного моделювання складної динаміки за допомогою штучного інтелекту на прикладі автоматичного моделювання патернів будови органів»

студента 6-го курсу факультету комп'ютерних наук та кібернетики

Київського національного університету імені Тараса Шевченка

Чергикало Дениса Олександровича

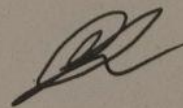
Метою роботи є побудова нових алгоритмів, методів та загальної методології моделювання та оптимізації складних біологічних систем. Опис області стійкості в мультикомпонентних нелінійних диференціальних рівнянь з частинними похідними та методів ідентифікації коефіцієнтів в стохастичних мультикомпонентних нелінійних диференціальних рівнянь з частинними похідними.

Була вперше детально описана задача ідентифікації коефіцієнтів в стохастичних мультикомпонентних нелінійних просторово-часових РЧП з компонентами які не можливо споглядати та всього декількома точками часу в яких можна споглядати стан системи, яку до цього часу розв'язували лише ручним підбором. Автор дає загальні формули для області лінійної стійкості нестохастичного випадку, для випадку 2-ух компонентів автор наводить явні формули. Також розглядається спряжена лінеаризоване РЧП для знаходження градієнту задачі ідентифікації коефіцієнтів мультикомпонентних нелінійних РЧП, і показано як за допомогою симуляцій Монте-Карло отримати з таких градієнтів градієнт стохастичної задачі. Було показано на прикладах що цей метод дає більш кращий результат ідентифікації коефіцієнтів для будови патерна ніж це роблять люди вручну, таким чином зробивши автоматизацію для цього процесу. На її основі було побудована модель розвитку патернів шкіри *Ictidomys tridecemlineatus*, що є, на думку автора роботи, найбільш якісною моделлю для цього патерну з тих що є.

За змістом роботи можна зробити зауваження. Автор каже що створив нову теорію яка б могла б бути корисною для багато-масштабного моделювання в біології але як обґрунтування ефективності своїх методів розглянув на практиці лише один орган - шкіру. Тому поки ще не можна казати що його методи будуть корисні для багато-масштабного моделювання в біології в загальному випадку. Але не вважаю це зауваження таким, що зменшує загальну позитивну оцінку роботи.

Вважаю, що кваліфікаційна робота студента Д. О. Чергикало відповідає вимогам, які висуваються до магістерських робіт, і заслуговує на оцінку «відмінно», а її автор заслуговує на присвоєння кваліфікації магістра.

Член-кореспондент НАН України,
доктор фізико-математичних наук, професор,
завідувач кафедри обчислювальної математики
факультету комп'ютерних наук та кібернетики
Київського національного університету
імені Тараса Шевченка



Сергій ЛЯШКО



Ім'я користувача:
Оноцький В'ячеслав ФКомпНаук

ID перевірки:
1015192467

Дата перевірки:
22.05.2023 22:41:12 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
22.05.2023 22:42:09 EEST

ID користувача:
100002816

Назва документа: ЧергикалоДенисОлександрович (1)

Кількість сторінок: 61 Кількість слів: 11582 Кількість символів: 78511 Розмір файлу: 1.54 MB ID файлу: 1014870623

1.15% Схожість

Найбільша схожість: 0.16% з Інтернет-джерелом (<https://kgu.kz/sites/default/files/documents/nauka/konf/5.pdf>)



0% Цитат

Вилучення цитат вимкнено

Вилучення списку бібліографічних посилань вимкнено

0% Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи 157