НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
Кафедра математичного моделювання та аналізу даних

«До захисту допущено»
В.о. завідувача кафедри
__________ І. М. Терещенко
«__» __________ 2023 р.

Дипломна робота

на здобуття ступеня бакалавра

зі спеціальності: 113 Прикладна математика
на тему: **«Оцінювання параметрів частково спостережуваного ланцюга Маркова на двійкових послідовностях»**

Виконав: студент 4 курсу, групи ФІ-91
Цибульник Антон Владиславович

Керівник: ст.викл. Наказной П. О. __________

Консультант: к.ф.-м.н., доцент Ніщенко І. І. __________

Рецензент: (згодом) __________

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

Студент __________

Київ — 2023

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені Ігоря СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
Кафедра математичного моделювання та аналізу даних

Рівень вищої освіти — перший (бакалаврський)
Спеціальність (освітня програма) — 113 Прикладна математика, ОПП «Математичні методи моделювання, розпізнавання образів та комп'ютерного зору»

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
__________ І. М. Терещенко
«__» __________ 2023 р.

ЗАВДАННЯ
на дипломну роботу

Студент: Цибульник Антон Владиславович

1. Тема роботи: «*Оцінювання параметрів частково спостережуваного ланцюга Маркова на двійкових послідовностях*»,

керівник: ст.викл. Наказной П. О.,

затверджені наказом по університету №__ від «__» __________ 2023 р.

2. Термін подання студентом роботи: «__» __________ 2023 р.

3. Вихідні дані до роботи: *опубліковані джерела за тематикою дослідження.*

4. Зміст роботи: *спостерігаючи часткову інформацію про динаміку бінарних послідовностей, за допомогою математичного апарату прихованих марковських моделей та із використанням методів математичної статистики дослідити деякі характреристики заданої моделі, зокрема: оцінити керуючий параметр динаміки системи; оцінити керуючий параметр динаміки системи у випадку додаткового зашумлення спостережуваних даних.*

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо): *презентація доповіді.*

6. Дата видачі завдання: 9 жовтня 2022 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми роботи із науковим керівником	09-14 жовтня 2022 р.	Виконано
2	Огляд опублікованих джерел за тематикою дослідження	15-30 жовтня 2022 р.	Виконано
3	Моделювання еволюції прихованої Марковської моделі	Листопад 2022 р.	Виконано
4	Побудова оцінки для керуючого параметра системи	Грудень 2022 р.	Виконано
5	Побудова оцінки для керуючого параметра системи у випадку додаткового зашумлення спостережуваних даних	Січень 2023 р.	Виконано
6	Дослідження інших харакреристик побудованої прихованої Марковської моделі	Лютий 2023 р.	Виконано
7	Побудова статистичних оцінок невідомих параметрів моделі	Березень 2023 р.	Виконано
8	Програмна реалізація алгоритмів	Квітень-травень 2023 р.	Виконано

Студент _______________ Цибульник А. В.

Керівник _______________ Наказной П. О.

РЕФЕРАТ

Кваліфікаційна робота містить: ??? стор., ??? рисунки, ??? таблиць, ??? джерел.

Об'єктом дослідження є ланцюг Маркова зі значеннями в множині двійкових послідовностей фіксованої довжини. Динаміка ланцюга задається як випадкове блукання вершинами одиничного куба, розмірність якого збігається з довжиною двійкової послідовності.

Стани заданого ланцюга є неспостережуваними (прихованими). Спостережуваними величинами в кожен момент часу є набір значень певного функціонала від фіксованих підмножин двійкової послідовності, яка описує поточний стан прихованого ланцюга.

Метою дослідження є побудова оцінок невідомих параметрів заданої Марковської моделі за допомогою математичного апарату прихованих Марковських моделей та із використанням методів математичної статистики.

Результати чисельного експерименту продемонстрували ефективність використаних методів, зокрема збіжність побудованих оцінок до істинних значень параметрів при збільшенні кількості спостережень.

ЛАНЦЮГ МАРКОВА, ПРИХОВАНА МАРКОВСЬКА МОДЕЛЬ, АЛГОРИТМ БАУМА-ВЕЛША, АЛГОРИТМ ВІТЕРБІ

ABSTRACT

The English abstract must be the exact translation of the Ukrainian "annotation" (including statistical data and keywords).

ЗМІСТ

Перелік умовних позначень, скорочень і термінів 7

Вступ 8

1 Методи дослідження прихованих марковських моделей 10

1.1 Основні поняття і властивості ланцюгів Маркова 10

1.2 Поняття прихованого ланцюга Маркова 13

1.3 Ітераційний алгоритм Баума-Велша 14

1.4 Алгоритм Вітербі 17

1.5 Статистика на ланцюгах Маркова 18

Висновки до розділу 1 19

2 (Назва другого розділу) 20

2.1 (Якийсь підрозділ) 20

2.2 (Якийсь наступний підрозділ з дуже-дуже довгою назвою, загальна кількість слів в якій, однак, не повинна перевищувати 12 слів) 21

Висновки до розділу 2 22

3 (Назва третього розділу) 23

3.1 (якийсь підрозділ) 23

Висновки до розділу 3 24

Висновки 25

Перелік посилань 26

Додаток А Тексти програм 27

А.1 Програма 1 27

Додаток Б Великі рисунки та таблиці 28

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

$\mathbb{1}$ — індикаторна функція;

$\mathbb{N}$ — множина натуральних чисел;

$\forall$ — квантор загальності: будь-який або для усіх;

Вип. вел. / в. в. — випадкова величина;

ЗВЧ — Закон великиз чисел

ПММ — прихована Марковська модель.

ВСТУП

Актуальність дослідження. Марковські моделі мають широкий та ефективний арсенал інструментів для аналізу динаміки систем, поведінка яких у кожен наступний момент часу зумовлюється лише поточним станом системи та не залежить від характеру еволюції у попередні моменти часу.

Водночас, у випадку, коли безпосереднє спостереження еволюції ланцюга Маркова є неможливим чи обмеженим, застосовують моделі прихованих ланцюгів Маркова (ПММ). У такому випадку аналіз поведінки процесу відбувається за деякою опосередкованою інформацією про «приховані», справжні стани ланцюга.

Наприклад, в біоінформатиці [1, глава 9] апарат ланцюгів Маркова застосовують при дослідженні еволюції молекул ДНК протягом певного часу, вважаючи при цьому за стан системи зв'язану послідовність так званих нуклеотидів, які формуються над алфавітом чотирьох азотистих основ $\{T, C, A, G\}$.

За рахунок використання ймовірнісних переходів між фонемами чи словами, приховані Марковські моделі є ефективним інструментом для таких завдань, як голосові команди, служби транскрипції та голосові помічники [2].

Не винятком стають і задачі розпізнавання мови жестів [3]: наприклад, представляючи жести як послідовності прихованих станів, ПММ можуть фіксувати динаміку та варіації рухів рук.

Відтак, враховуючи актуальність вивчення еволюції систем, стани яких є послідовностями, наборами символів певної довжини, у роботі розглядається дослідження динаміки бінарних послідовностей, вважаючи, що спостерігаються не безпосередньо самі послідовності, а лише деяка опосередкована, часткова інформація про них.

Метою дослідження є оцінка характеристик частково спостережуваної Марковської моделі на бінарних послідовностях. Для досягнення мети необхідно розв'язати **задачу дослідження**, яка полягає у вирішенні таких завдань:

1) провести огляд опублікованих джерел за тематикою дослідження;

2) перевірити, чи задана модель відповідає необхідним умовам для використання апарату прихованих Марковських ланцюгів;

3) побудувати оцінки згідно обраних методів;

4) експериментально перевірити ефективність отриманих теоретичних оцінок.

Об'єктом дослідження є процеси, які описуються моделями частково спостережуваних ланцюгів Маркова.

Предметом дослідження є моделі та методи оцінки параметрів частково спостережуваного ланцюга Маркова на бінарних послідовностях.

При розв'язанні поставлених завдань використовувались такі *методи дослідження*: методи лінійної алгебри, теорії імовірностей, математичної статистики, методи комп'ютерного та статистичного моделювання, методи та алгоритми дослідження ланцюгів Маркова.

Наукова новизна отриманих результатів полягає у ...

Практичне значення результатів полягає у ...

Апробація результатів та публікації. Частина даної роботи була представлена на XXI Науково-практичній конференції студентів, аспірантів та молодих вчених «Теоретичні і прикладні проблеми фізики, математики та інформатики» (11-12 травня 2023 р., м. Київ).

1 МЕТОДИ ДОСЛІДЖЕННЯ ПРИХОВАНИХ МАРКОВСЬКИХ МОДЕЛЕЙ

У цьому розділі будуть окреслені основні поняття та методи, які використовуватимуться в подальших викладках при розв'язуванні поставлених задач.

1.1 Основні поняття і властивості ланцюгів Маркова

Нехай $\{X^t\}_{t\geqslant 1}$ — послідовність випадкових величин зі значеннями в скінченній або зліченній множині $E = \{e_1, e_2, \ldots\}$.

Означення 1.1. Послідовність $\{X^t\}_{t\geqslant 1}$ утворює ланцюг Маркова, якщо:

$$\forall t \geqslant 2 \quad \forall i^1, i^2, \ldots, i^{t+1} \in E:$$
$$P\left(X^{t+1} = i^{t+1} \,|\, X^t = i^t, \ldots, X^1 = i^1\right) = P\left(X^{t+1} = i^{t+1} \,|\, X^t = i^t\right)$$

Цю умову називають «умовою марковості».

Множина E називається множиною станів ланцюга, а випадкова величина X^t трактується як стан системи в момент часу t. Надалі у ході дослідження по замовчуванню розглядатимуться так звані однорідні ланцюги Маркова, для яких ймовірності переходу з одного стану $i \in E$ в інший $j \in E$

$$p_{ij} = P\left(X^{t+1} = j \,|\, X^t = i\right)$$

не залежать від t, тобто

$$P\left(X^2 = j \,|\, X^1 = i\right) = P\left(X^3 = j \,|\, X^2 = i\right) = \ldots = P\left(X^{10} = j \,|\, X^9 = i\right) = \ldots$$

Ймовірність $p_{ij} = P\left(X^{t+1} = j \,|\, X^t = i\right)$ називається перехідною ймовірністю однорідного ланцюга Маркова. Матриця A, складена із цих імовірностей, називається матрицею перехідних імовірностей:

$$A = \Big(p_{ij}\Big)_{i,j \in E} = \Big(P\left(X^{t+1} = j \mid X^t = i\right)\Big)_{i,j \in E}$$

Ця матриця є стохастичною, тобто

$$\forall i,j \in E : p_{ij} \geqslant 0 \ \text{ та } \ \forall i \in E : \sum_{j \in E} p_{ij} = 1$$

Окрім матриці A, для характеристики ланцюга Маркова слід задати початковий розподіл

$$\pi = \Big(\pi_i\Big)_{i \in E} = \Big(P\left(X^1 = i\right)\Big)_{i \in E}$$

Іншими словами, задається ймовірнісний розподіл, згідно якого відбуватиметься початкова ініціалізація ланцюга на множині станів E. При цьому, якщо ініціалізація ланцюга на кожному наступному кроці також відбувається згідно заданого розподілу, то π називають інваріантним. Умова інваріантності ланцюга записується таким чином:

Означення 1.2. Ланцюг Маркова $\{X^t\}_{t \geqslant 1}$ з матрицею перехідних імовірностей A та початковим розподілом $\pi = \Big(\pi_i\Big)_{i \in E}$ є інваріантним, якщо:

$$\pi = \pi A$$

Наступна лема [4] встановлює умови для такого роду перевірки: чи утворює ланцюг Маркова деяка задана послідовність випадкових величин.

Лема 1.1. *Нехай $\{\xi^t\}_{t \geqslant 1}$ — послідовність незалежних випадкових величин, і нехай X^1 — незалежна від цієї послідовності випадкова величина. Означимо для деякої функції $f : \mathbb{N} \times E \times E \longrightarrow E$ та при $t \geqslant 1$*

$$X^{t+1} = f(t, X^t, \xi^{t+1})$$

Тоді $\{X^t\}_{t \geqslant 1}$ утворює ланцюг Маркова на множині станів E.

Надалі позначатимемо Марковську модель з матрицею перехідних імовірностей A та початковим розподілом π як $\lambda = (\pi, A)$.

Сформулюємо у теоремі нижче [5] ще одну важливу ознаку ланцюгів Маркова — так звану «марковську властивість».

Теорема 1.1 («марковська властивість»)**.** *Нехай* $\{X^t\}_{t\geqslant 1}$ *— ланцюг Маркова* $\lambda = (\pi, A)$*. Тоді для довільного* $s \geqslant 2$ *та для довідного стану* $j \in E$ *при умові, що* $X^s = j$*, послідовність* $\{X^{s+t}\}_{t\geqslant 1}$ *є ланцюгом Маркова з початковим розподілом*

$$\delta = (0, 0, \ldots, \delta_j, 0, \ldots, 0), \text{ де } \delta_j = 1$$

та матрицею перехідних імовірностей A*. Більш того, розподіл випадкових величин* $X^{s+1}, X^{s+2}, \ldots$ *при* $X^s = j$ *не залежить від* $X^1, X^2, \ldots, X^{s-1}$ *:*

$$\begin{aligned}
&\forall t \geqslant 1 \quad \forall s \geqslant 2 \quad \forall i^1, i^2, \ldots, i^{t+1}, j \in E : \\
&\quad P\left(X^{t+1} = i^{t+1}, X^t = i^t, \ldots, X^{s+1} = i^{s+1} \,|\, X^s = j, \ldots, X^1 = i^1\right) = \\
&\qquad\qquad = P\left(X^{t+1} = i^{t+1}, X^t = i^t, \ldots, X^{s+1} = i^{s+1} \,|\, X^s = j\right)
\end{aligned}$$

Наостанок, зазначимо вигляд скінченновимірних розподілів [5] ланцюга Маркова $\lambda = (\pi, A)$. Інакше кажучи, визначимо для довільного t сумісну ймовірність перебування системи в станах $i^1, i^2, \ldots, i^t \in E$ у послідовні моменти часу:

$$P\left(X^1 = i^1, X^2 = i^2, \ldots, X^{t-1} = i^{t-1}, X^t = i^t\right) = \pi_{i^1} \cdot p_{i^1 i^2} \cdot \ldots \cdot p_{i^{t-1} i^t}$$

Таким чином, вказана ймовірність повністю визначається через компонени матриці перехідних імовірностей A та елементи вектора початкового розподілу π.

1.2 Поняття прихованого ланцюга Маркова

Розглянемо послідовність випадкових величин $\{X^t\}_{t\geqslant 1}$ на скінченній або зліченній множині станів $E = \{e_1,e_2,\ldots\}$. Крім того, нехай $\{Y^t\}_{t\geqslant 1}$ — послідовність випадкових величин на скінченній або зліченній множині $F = \{f_1,f_2,\ldots\}$.

Означення 1.3. Пара $\{(X^t,Y^t)\}_{t\geqslant 1}$, задана на декартовому добутку $E \times F$, є прихованою Марковською моделлю за виконання таких умов:

1) послідовність $\{X^t\}_{t\geqslant 1}$ утворює ланцюг Маркова з початковим розподілом π та матрицею перехідних імовірностей A;

2) послідовність $\{(X^t,Y^t)\}_{t\geqslant 1}$ є ланцюгом Маркова;

3) випадкові величини $Y^1,Y^2,\ldots,Y^t$ є умовно незалежними при заданому наборі величин $X^1,X^2,\ldots,X^t$:

$$\forall t \geqslant 2 \quad \forall j^1,j^2,\ldots,j^t \in F \quad \forall i^1,i^2,\ldots,i^t \in E :$$
$$P\left(Y^1 = j^1,\ldots,Y^t = j^t \mid X^1 = i^1,\ldots,X^t = i^t\right) = \\ = \prod_{k=1}^{t} P\left(Y^k = j^k \mid X^1 = i^1,\ldots,X^t = i^t\right)$$

4) умовний розподіл випадкової величини Y^k в момент часу k при заданих $X^1,X^2,\ldots,X^t$ залежить лише від X^k :

$$\forall t \geqslant 2 \quad \forall k = \overline{1,t} \quad \forall j^k \in F \quad \forall i^1,i^2,\ldots,i^t \in E :$$
$$P\left(Y^k = j^k \mid X^1 = i^1,\ldots,X^t = i^t\right) = P\left(Y^k = j^k \mid X^k = i^k\right)$$

У парі $\{(X^t,Y^t)\}_{t\geqslant 1}$ послідовність $\{X^t\}_{t\geqslant 1}$ називають «прихованою», а послідовність $\{Y^t\}_{t\geqslant 1}$ — «спостережуваною». Фактично, в таких термінах в Озн. 1.3 умови 3) та 4) в сукупності вказують на взаємозв'язок спостережень та прихованих станів виключно через поточний момент часу.

Окрім матриці A та вектора π, властивих ланцюгу Маркова

$\{X^t\}_{t\geqslant 1}$, прихована Марковська модель визначається матрицею умовних імовірностей спостережень $j \in F$ при заданих прихованих станах $i \in E$:

$$B = \left(B_{ij}\right)_{i,j \in E \times F} = \left(P\left(Y^t = j \,|\, X^t = i\right)\right)_{i,j \in E \times F}$$

Позначатимемо приховану Марковську модель з початковим розподілом π, матрицею перехідних імовірностей A та матрицею умовних ймовірностей спостережень при заданих прихованих станах B таким чином: $\lambda = (\pi, A, B)$.

1.3 Ітераційний алгоритм Баума-Велша

Нехай протягом деякого часу $t = \overline{1,T}$ спостерігається послідовність випадкових величин

$$\left(Y^1 = y^1, \ldots, Y^T = y^T\right) \Longleftrightarrow Y = y,$$

деякої прихованої Марковської моделі, ознаки $\lambda = (\pi, A, B)$ якої є невідомими. Постає питання так званої задачі навчання: як за набором наявних даних віднайти оптимальні параметри моделі?

Скористаємося методом максимальної правдоподібності, шукаючи оцінку $\lambda^* = (\pi^*, A^*, B^*)$ шляхом максимізації ймовірності вигляду:

$$\lambda^* = \underset{\lambda}{\operatorname{argmax}}\, P\left(Y = y \,|\, \lambda\right) \tag{1.1}$$

Інакше кажучи, шукатимемо такі параметри моделі, які найкраще пояснюють отримані спостереження y.

Ймовірність $P\left(Y = y \,|\, \lambda\right)$ називається функцією правдоподібності. І у випадку прихованих Марковських моделей безпосередня максимізація цієї функції через, наприклад, диференціювання по відповідним невідомим параметрам часто є складною аналітично чи неможливою в цілому задачею в силу громіздкості отриманого виразу.

Однак, для Марковських моделей можна застосувати інший підхід: модифікацію ЕМ-алгоритму [1, розділ 4] для дослідження прихованих ланцюгів Маркова — ітераційний алгоритм Баума-Велша [1, розділ 15].

Задавши деяке наближення невідомої моделі $\lambda^{(0)} = (\pi^{(0)}, A,^{(0)}, B^{(0)})$, покладемо для наступної ітерації $n+1$

$$\lambda^{(n+1)} = \underset{\lambda}{\operatorname{argmax}}\, Q\left(\lambda^{(n)}, \lambda\right),$$

де

$$Q\left(\lambda^{(n)}, \lambda\right) = \sum_{x \in E^T} L_{\lambda^{(n)}} \cdot \ln L_\lambda \tag{1.2}$$

є так званою функцією квазі-log правдоподібності, а вираз

$$L_\lambda \equiv P\left(X = x, Y = y \,|\, \lambda\right)$$

називається повною функцією правдоподібності.

Доведено [1, розділ 4], що така ітераційна процедура є збіжною і приводить до точки локального максимуму логарифму функції правдоподібності (1.1). А оскільки у довільної функції та у відповідної логарифмічної функції від неї точки екстремумів є однаковими, ця процедура повністю задовільняє поставлену задачу.

В залежності від специфіки системи (характер динаміки; особливість прихованих станів ланцюга; повноцінність, частковість чи зашумленість наявних спостережень), явний аналітичний вигляд результату максимізації функції (1.2) для кожної окремо заданої прихованої Марковської моделі є різним.

Тим не менш, серед загальних особливостей можна виокремити наявність в отриманих ітераційних формулах переоцінки невідомих параметрів так званих коефіцієнтів прямого (1.3) та зворотного ходу (1.4).

Вказані коефіцієнти визначаються наступним чином:

$$\forall x \in E:$$
$$\alpha_t(x) = P\left(Y^1 = y^1, \ldots, Y^t = y^t, X^t = x \mid \lambda^{(n)}\right) \tag{1.3}$$
$$\beta_t(x) = P\left(Y^{t+1} = y^{t+1}, \ldots, Y^T = y^T \mid X^t = x, \lambda^{(n)}\right) \tag{1.4}$$

Крім того, вказані ймовірності можна визначити рекурентно [6, розділ 5]. Відповідні рекурентні співвідношення наведені нижче для коефіцієнтів прямого ходу

$$t = 1 \qquad \forall x \in E : \alpha_1(x) = \pi_x \, B_{x,y^1}$$
$$t = \overline{2,T} \qquad \forall x \in E : \alpha_t(x) = \sum_{x' \in E} \alpha_{t-1}(x') \, A_{x'x} \, B_{xy^t}$$

та коефіцієнтів зворотного ходу

$$t = T \qquad \forall x \in E : \beta_T(x) = 1$$
$$t = \overline{T-1,1} \qquad \forall x \in E : \beta_t(x) = \sum_{x' \in E} \beta_{t+1}(x') \, A_{xx'} \, B_{x'y^{t+1}}$$

Зауваження. При великих значеннях довжини ланцюга виникає потреба у шкалюванні [6, розділ 5] коефіцієнтів прямого та зворотного ходу, адже їхні значення стають нерозрізнювано малими для обчислювальних ресурсів. Процедура шкалювання полягає в наступному: на кожному кроці t після обчислення істинних (1.3) слід виконати відповідне нормування

$$\forall x \in E : \qquad \widehat{\alpha}_t(x) = \frac{\alpha_t(x)}{C_t}, \text{ де } C_t = \sum_{x' \in E} \alpha_t(x'),$$

а тоді коефіцієнти (1.4) переоцінюватимуться так:

$$\forall x \in E : \qquad \widehat{\beta}_t(x) = \frac{\beta_t(x)}{C_t}$$

1.4 Алгоритм Вітербі

Отримавши за наявними спостереженнями оптимальну модель $\lambda^* = (\pi^*, A^*, B^*)$, перейдемо до так званої задачі декодування: віднайдемо ланцюжок прихованих станів системи. Алгоритм, який дозволяє ефективно розв'язати задачу декодування, називається алгоритмом Вітербі [6, розділ 6].

Отже, шукатимемо таку послідовність прихованих станів $\widehat{X}^1, \widehat{X}^2, \ldots, \widehat{X}^T$, яка найкращим чином описує наявні спостереження:

$$\widehat{X} = \underset{x \in E^T}{\operatorname{argmax}} \, P\left(X = x, Y = y \mid \lambda^*\right)$$

Введемо величини $\delta_t(x)$ — ймовірності спостереження ланцюжка довжини t, використовуючи найкращий шлях, що закінчується станом $x \in E$ в момент часу t :

$$\delta_t(x) = \max_{x^1, \ldots, x^{t-1}} P\left(X^1 = x^1, \ldots, X^t = x, Y^1 = y^1, \ldots, Y^t = y^t \mid \lambda^*\right)$$

Вказані ймовірності можна визначити рекурентно:

$$\begin{aligned}
&t = 1 && \forall x \in E : \delta_1(x) = \pi_x \, B_{xy^1} \\
&t = \overline{2,T} && \forall x \in E : \delta_t(x) = B_{xy^t} \cdot \max_{x' \in E} \{\delta_{t-1}(x') \, A_{x'x}\}
\end{aligned}$$

При цьому, щоб знайти оптимальний ланцюжок прихованих станів, необхідно відстежувати аргумент, при якому досягається максимум $\delta_t(x)$ для кожного t та x. Таким чином, алгоритм Вітербі знаходження найбільш ймовірного ланцюжка прихованих станів є таким:

1) ініціалізація:

$$\forall x \in E : \qquad \delta_1(x) = \pi_x B_{xy^1}, \;\; \psi_1(x) = 1$$

2) обчислити коефіцієнти $\delta_t(x)$ та відповідні аргументи $\psi_t(x)$:

$$\forall t = \overline{1,T},\ \forall x \in E : \qquad \delta_t(x) = B_{xy^t} \cdot \max_{x' \in E}\{\delta_{t-1}(x')\, A_{x'x}\}$$
$$\forall t = \overline{1,T},\ \forall x \in E : \qquad \psi_t(x) = \operatorname*{argmax}_{x' \in E}\{\delta_{t-1}(x')\, A_{x'x}\}$$

3) покласти зворотну точку відліку:

$$\widehat{\delta} = \max_{x \in E}\{\delta_T(x)\}$$
$$\widehat{\psi} = \operatorname*{argmax}_{x \in E}\{\delta_T(x)\}$$

4) визначити оптимальний ланцюжок станів (у зворотному порядку), починаючи з останнього $\widehat{x}^T = \widehat{\psi}$:

$$\forall t = \overline{T-1, 1} : \qquad \widehat{x}^t = \psi_{t+1}(\widehat{x}^{t+1})$$

1.5 Статистика на ланцюгах Маркова

Окреслимо основні інструменти математисної статистики, які будуть використані при побудові статистичних оцінок невідомих параметрів.

Вектор незалежних однаково розподілених випадкових величин $X^1, \ldots, X^T$ з деякого розподілу $F(\theta)$ називають вибіркою об'єму T. При цьому, параметр розподілу θ може бути невідомим. Функція від вибірки $S_T = S_T(\overrightarrow{X})$ називається статистикою.

Якщо значення статистики $S_T(\overrightarrow{X})$ при заданій реалізації вибірки приймає наближене значення невідомого параметра розподілу θ, тоді S_T називають точковою оцінкою θ.

Означення 1.4. Статистика S_T називається змістовною оцінкою θ, якщо вона збігається за ймовірністю до істинного значення оцінюваного параметра, тобто

$$S_T \xrightarrow[T \longrightarrow \infty]{P} \theta \quad \Longleftrightarrow \quad \forall \varepsilon > 0 : \lim_{T \longrightarrow \infty} P\big(|S_T - \theta| \geqslant \varepsilon\big) = 0$$

Означення 1.5. Статистика S_T називається незміщеною оцінкою θ, якщо її математичне сподівання дорівнює істинному значенню оцінюваного параметра, тобто

$$MS_T = \theta$$

Також зазначимо для набору незалежних однаково розподілених випадкових величин так званий Закон великих чисел [7]:

Теорема 1.2. *Нехай $X^1, \ldots, X^T$ — незалежні однаково розподілені випадкові величини зі скінченними математичними сподіваннями $\forall k = \overline{1,T} : MX^k = m < \infty$ та дисперсіями $\forall k = \overline{1,T} : DX^k = \sigma^2 < \infty$, тоді вибіркова дисперсія $\overline{X}$ збігається за ймовірністю до значення математичного сподівання заданої вибірки:*

$$\overline{X} \xrightarrow[T\longrightarrow\infty]{P} m \quad \Longleftrightarrow \quad \forall\varepsilon > 0 : \lim_{T\longrightarrow\infty} P\left(\left|\frac{1}{T}\sum_{k=1}^{T} X^k - m\right| \geqslant \varepsilon\right) = 0$$

Висновки до розділу 1

Розглядаючи динаміку бінарних послідовностей протягом певного часу при наявних опосередкованих даних про неї, у наступному розділі будуть розв'зані задачі оцінки деяких характеристик зазначної моделі. При цьому, важливо перконатися, що модель відповідатиме умовам використання окреслених у цьому розділі алгоритмів та методів.

2 (НАЗВА ДРУГОГО РОЗДІЛУ)

До другого розділу також краще написати малесенький вступ. Зокрема, це збільшує загальний об'єм роботи та покращує її читабельність.

2.1 (Якийсь підрозділ)

У другому розділі необхідно наводити розв'язання поставленої перед вами задачі у теоретичному або аналітичному сенсі (хоча, звісно, все залежить від того, яка саме задача перед вами поставлена).

Для подання матеріалів можна використовувати таблиці (наприклад, Таблицю 2.1). Розмір шрифту у таблиці може бути меншим за 14 pt (наприклад, 12 pt, або навіть 10 pt, якщо так таблиця виглядає зрозуміліше та компактніше).

Таблиця 2.1 – Розрахунок якоїсь фантастичної дичини у декілька кроків

Параметр x_i	Параметр x_j				Перший крок		Другий крок	
	X_1	X_2	X_3	X_4	w_i	$K_{\text{в}i}$	w_i	$K_{\text{в}i}$
X_1	1	1	1.5	1.5	5	0.31	19	0.32
X_2	1	1	1.5	1.5	5	0.31	19	0.32
X_3	0.5	0.5	1	0.5	2.5	0.16	9.25	0.16
X_4	0.5	0.5	1.5	1	3.5	0.22	12.25	0.20
Разом:					16	1	59.5	1

Бажано, щоб кожен пункт завдань, окреслених у вступі, відповідав певному розділу або підрозділу у дипломній роботі.

Теорема 2.1. *Нумерація у наступних розділах також* *проставляється автоматично та коректно.*

2.2 (Якийсь наступний підрозділ з дуже-дуже довгою назвою, загальна кількість слів в якій, однак, не повинна перевищувати 12 слів)

Для подання матеріалів також дуже зручними є рисунки (наприклад, рисунки 2.1 або 2.2).

(а)

(б)

Рисунок 2.1 – Різні види риб: (а) судак, (б) тудак.

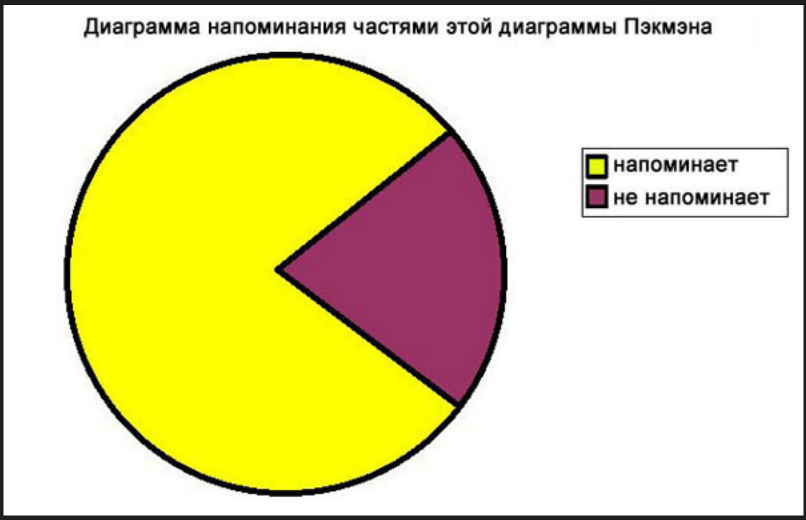

Рисунок 2.2 – Частка кругових діаграм, які схожі на Пекмена

Висновки до розділу 2

Наприкінці розділу знову наводяться коротенькі підсумки.

3 (НАЗВА ТРЕТЬОГО РОЗДІЛУ)

3.1 (якийсь підрозділ)

Подивіться, як нераціонально використовується простір, якщо не писати вступи до розділів. :)

Зазвичай третій розділ присвячено опису практичного застосування або експериментальної перевірки аналітичних результатів, одержаних у другому розділі роботи. Втім, це не обов'язкова вимога, і структура основної частини диплому більш суттєво залежить від характеру поставлених завдань. Навіть якщо у вас є певне експериментальне дослідження, але його загальний опис займає дві сторінки, то краще приєднайте його підрозділом у попередній розділ.

При описі експериментальних досліджень необхідно:

– наводити повний опис експериментів, які проводились, параметрів обчислювальних середовищ, засобів програмування тощо;

– наводити повний перелік одержаних результатів у чисельному вигляді для їх можливої перевірки іншими особами;

– представляти одержані результати у вигляді таблиць та графіків, зрозумілих людському оку;

– інтерпретувати одержані результати з точки зору поставленої задачі та загальної проблематики ваших досліджень.

У жодному разі не потрібно вставляти у даний розділ тексти інструментальних програм та засобів (окрім того рідкісного випадку, коли саме тексти програм і є результатом проведення експериментів). За необхідності тексти програм наводяться у додатках.

Висновки до розділу 3

Висновки до останнього розділу є, фактично, підсумковими під усім дослідженням; однак вони повинні стостуватись саме того, що розглядалось у розділі.

ВИСНОВКИ

Загальні висновки до роботи повинні підсумовувати усі ваші досягнення у даному напрямку досліджень.

За кожним пунктом завдань, поставлених у вступі, у висновках повинен міститись звіт про виконання: виконано, не виконано, виконано частково (І чому саме так). Наприклад, якщо першим поставленим завданням у вас іде «огляд літератури за тематикою досліджень», то на початку висновків ви повинні зазначити, що «у ході даної роботи був проведений аналіз опублікованих джерел за тематикою (...), який показав, що (...)». Окрім простої констатації про виконання ви повинні навести, які саме результати ви одержали та проінтерпретувати їх з точки зору поставленої задачі, мети та загальної проблематики.

В ідеалі загальні висновки повинні збиратись з висновків до кожного розділу, але ідеал недосяжний. :) Однак висновки не повинні містити формул, таблиць та рисунків. Дозволяється (та навіть вітається) використовувати числа (на кшталт «розроблена методика дозволяє підвищити ефективність пустопорожньої балаканини на 2.71%»).

Наприкінці висновків необхідно зазначити напрямки подальших досліджень: куди саме, як вам вважається, необхідно прямувати наступним дослідникам у даній тематиці.

ПЕРЕЛІК ПОСИЛАНЬ

1. *Koski T.* Hidden Markov models for bioinformatics. — 2002-е вид. — New York, NY : Springer, 11.2001. — (Computational Biology).
2. *Kumar L. A.*, *Renuka D. K.* Fundamentals of Speech Recognition // Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision. — CRC Press, 02.2023. — С. 99—125. — DOI: 10.1201/9781003348689-5. — URL: https://doi.org/10.1201/9781003348689-5.
3. *Chaaraoui A. A.*, *Climent-Pérez P.*, *Flórez-Revuelta F.* One shot learning for gesture recognition using HMMs and hand appearance features // Pattern Recognition Letters. — 2013. — Т. 34, № 9. — С. 1009—1017.
4. *Serfozo R.* Basics of Applied Stochastic Processes. — Springer Berlin Heidelberg, 2009. — DOI: 10.1007/978-3-540-89332-5. — URL: https://doi.org/10.1007/978-3-540-89332-5.
5. *Norris J. R.* Markov Chains. — Cambridge University Press, 02.1997. — DOI: 10.1017/cbo9780511810633. — URL: https://doi.org/10.1017/cbo9780511810633.
6. *Nilsson M.* First Order Hidden Markov Model: Theory and Implementation Issues : тех. звіт. / Blekinge Institute of Technology, School of Engineering, Department of Signal Processing. — 2005.
7. *Larsen R. J.*, *Marx M. L.* An introduction to mathematical statistics and its applications. — 6-е вид. — "Upper Saddle River, NJ" : "Pearson", 01.2017.

ДОДАТОК А ТЕКСТИ ПРОГРАМ

Тексти інструментальних програм для проведення експериментальних досліджень необхідно виносити у додатки.

А.1 Програма 1

Зауважте, як змінилась нумерація.

ДОДАТОК Б ВЕЛИКІ РИСУНКИ ТА ТАБЛИЦІ

Якщо результати вашої роботи описуються величезними рисунками і таблицями (один аркуш та більше) у незліченній кількості, їх також необхідно виносити у додатки.