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
__________ Іван 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
__________ Іван 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 досліджено деякі характеристики заданої моделі, зокрема: оцінено керуючий параметр динаміки системи; оцінено 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

Qualification work contains: ??? pages, ??? figures, ??? tables, ??? sources.

The object of the study is the Markov chain with values in the set of binary sequences of fixed length. Chain dynamics is given as a random walk on the vertices of a unit cube whose dimension coincides with the length of the binary sequence.

The states of a given circuit are unobservable (hidden). The observed values at each time point are a set of values of a certain functional from fixed subsets of the binary sequence that describes the current state of the hidden chain.

The purpose of the study is to construct estimates of unknown parameters of a given Markov model using the mathematical apparatus of hidden Markov models and using methods of mathematical statistics.

The results of the numerical experiment demonstrated the effectiveness of the methods used, in particular the convergence of the constructed estimates to the true values of the parameters with an increase in the number of observations.

MARKOV CHAIN, HIDDEN MARKOV MODEL, BAUM-WELCH ALGORITHM, VITERBI ALGORITHM

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

$\mathbb{1}$ — індикаторна функція;

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

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

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

ЗВЧ — закон великих 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 $\{T, C, A, G\}$.

Існування статистичних залежностей в чергуванні фонем чи слів в природних мовах зумовлює ефективність використання прихованих марковських моделей до таких завдань, як створення голосових команди, служб транскрипції та голосових помічників [2].

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

Відтак, враховуючи актуальність вивчення еволюції систем, стани яких є послідовностями чи наборами символів певної довжини, у роботі розглядається ланцюг Маркова на множині двійкових послідовностей, динаміка якого відстежується за зміною в часі набору функціоналів від його станів.

Метою дослідження є побудова оцінок невідомих параметрів частково спостережуваного ланцюга Маркова 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 моделей;

3) побудувати оцінки згідно з обраними методами;

4) експериментально перевірити ефективність отриманих теоретичних оцінок.

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

Предметом дослідження є 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 прихованих марковських моделей.

Наукова новизна отриманих результатів полягає у ...(можливо, прибрати цей абзац шаблону)

Практичне значення результатів полягає у ...(можливо, прибрати цей абзац шаблону)

Апробація результатів та публікації. Частина даної роботи була представлена на XXI Всеукраїнській 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 марковською властивістю.

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 $P\left(X^{t+1} = j \,|\, X^t = i\right)$ з одного стану $i \in E$ в інший $j \in E$ не залежать від t.

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:

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

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 ланцюга Маркова слід задати вектор початкового розподілу ймовірностей

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

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

Скінченновимірні розподіли [4] ланцюга Маркова $\lambda = (\pi, A)$ повністю визначаються матрицею перехідних імовірностей A та вектором початкового розподілу ймовірностей π. А саме:

$$\begin{aligned}&\forall t \geqslant 2 \quad \forall i^1, i^2, \ldots, i^t \in E : \\ &P\left(X^1 = i^1, X^2 = i^2, \ldots, X^{t-1} = i^{t-1}, X^t = i^t\right) = \pi_{i^1} \cdot p_{i^1 i^2} \cdot \ldots \cdot p_{i^{t-1} i^t}\end{aligned} \tag{1.1}$$

1.2 Поняття прихованої марковської моделі

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

Означення 1.2. 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 $X^1, X^2, \dots, X^t$:

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

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

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

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».

Окрім матриці A та вектора π, які задають ланцюг Маркова $\{X^t\}_{t \geqslant 1}$, прихована марковська модель визначається матрицею умовних імовірностей спостережень $j \in F$ при заданих прихованих станах $i \in E$:

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

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

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

деякої прихованої марковської моделі $\lambda = (\pi, A, B)$, параметри π, A, B якої є невідомими. Постає питання: як за набором наявних даних віднайти оптимальні параметри моделі?

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

$$\lambda^* = \underset{\lambda}{\operatorname{argmax}} \, P\left(Y = y \mid \lambda\right) \tag{1.2}$$

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

Ймовірність $P\left(Y = y \mid \lambda\right)$ називається функцією правдоподібності. Задача максимізації цієї функції є складною чи неможливою в цілому через громіздкість отриманого виразу: зважаючи на вигляд (1.1) скінченновимірних розподілів ланцюга Маркова та враховуючи умови 3) й 4) означення прихованої марковської моделі (Озн. 1.2), функція правдоподібності набуває вигляду:

$$P\left(Y = y \mid \lambda\right) = \sum_{x \in E^T} P\left(X = x, Y = y \mid \lambda\right) = \sum_{x \in E^T} \pi_{x^1} \cdot \prod_{t=1}^{T-1} A_{x^t x^{t+1}} \cdot \prod_{t=1}^{T} B_{x^t y^t}$$

Безпосередня максимізація за цією формулою вимагає охоплення близько $T \cdot |E|^T$ множників, де T, як правило, є великим. Однак, для прихованих марковських моделей можна застосувати інший підхід: модифікацію EM-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

$$Q\left(\lambda^{(n)}, \lambda\right) = \sum_{x \in E^T} L_{\lambda^{(n)}} \cdot \ln L_\lambda \tag{1.3}$$

є так званою функцією квазі-log правдоподібності, а вираз

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

називається функцією повної правдоподібності.

Доведено [1, розділ 4], що така ітераційна процедура є збіжною і приводить до точки локального максимуму логарифма функції правдоподібності (1.2). А оскільки точки екстремумів довільної функції та її логарифму збігаються, ця процедура розв'язує поставлену задачу.

Особливістю алгоритму навчання Баума-Велша є використання так званих змінних прямого (1.4) та зворотного (1.5) ходу, за допомогою яких обчислення функції правдоподібності вимагає лише $T \cdot |E|^2$ добутків. Вказані коефіцієнти визначаються наступним чином:

$$\begin{aligned}
&\forall x \in E : \\
&\alpha_t(x) = P\left(Y^1 = y^1, \ldots, Y^t = y^t, X^t = x \,|\, \lambda^{(n)}\right) && (1.4)\\
&\beta_t(x) = P\left(Y^{t+1} = y^{t+1}, \ldots, Y^T = y^T \,|\, X^t = x, \lambda^{(n)}\right) && (1.5)
\end{aligned}$$

Перевага цих коефіцієнтів полягає у тому, що їх можна обчислити рекурентно згідно з відповідними співвідношеннями [5, розділ 5] для змінних прямого ходу

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

та змінних зворотного ходу

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

Зауваження. При великих значеннях довжини ланцюга виникає потреба у шкалюванні [5, 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 змінних (1.4) 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 (1.5) нормуються так:

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

1.4 Алгоритм Вітербі

Отримавши оптимальну модель $\lambda^* = (\pi^*, A^*, B^*)$ як розв'язок задачі навчання, перейдемо до так званої задачі декодування: віднайдемо ланцюжок прихованих станів системи. Алгоритм, який дозволяє ефективно розв'язати задачу декодування, називається алгоритмом Вітербі [5, розділ 6].

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

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

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

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

Вказані ймовірності можна визначити рекурентно:

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

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 Елементи математичної статистики

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

Вектор $\overrightarrow{X} = (X^1, \ldots, X^T)$ незалежних однаково розподілених випадкових величин з деякої параметричної сім'ї розподілів $\mathfrak{F} = \{F_\theta(x),\, \theta \in \Theta\}$ 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 приймають за наближене значення невідомого параметра θ розподілу $F_\theta(x)$, тоді S_T називають точковою оцінкою θ.

Означення 1.3. Статистика S_T називається змістовною оцінкою θ, якщо вона збігається за ймовірністю до істинного значення оцінюваного параметра, тобто

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

Означення 1.4. Статистика S_T називається незміщеною оцінкою θ, якщо її математичне сподівання дорівнює істинному значенню оцінюваного параметра, тобто

$$\forall \theta \in \Theta \ : \ M_\theta S_T = \theta$$

Також зазначимо для набору незалежних однаково розподілених випадкових величин так званий закон великих чисел [6]:

Теорема 1.1. *Нехай $\{X^t\}_{t \geqslant 1}$ — послідовність незалежних однаково розподілених випадкових величин зі скінченним математичним сподіванням $MX^1 = m$. Тоді вибіркове середнє $\overline{X}$ збігається за ймовірністю до значення математичного сподівання заданої вибірки:*

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

Висновки до розділу 1

У цьому розділі було розглянуто основний математичний апарат прихованих марковських моделей, необхідний для подальшого дослідження динаміки частково спостережуваного ланцюга Маркова на бінарних послідовностях.

Використовуючи викладки цього розділу, проведемо побудову теоретичних оцінок невідомих параметрів вказаної моделі.

2 ОЦІНКА ПАРАМЕТРІВ МОДЕЛІ

Структуруємо дослідження таким чином:

1) спершу формально опишемо досліджуваний об'єкт, визначимо ключові параметри системи та переконаємося, що утворена модель є марковською;

2) розв'яжемо задачу навчання: за наявними спостереженнями про динаміку набору функціоналів від станів прихованого ланцюга бінарних послідовностей оцінимо керуючий параметр системи, використовуючи математичний апарат прихованих марковських моделей та методи побудови статистичних оцінок;

3) розв'яжемо задачу декодування: за наявними спостереженнями та оцінкою керуючого параметра відновимо ланцюг прихованих станів;

4) розв'яжемо задачу локалізації: за відомими значеннями набору функціоналів від деякої невідомої підмножини стану прихованого ланцюга, оцінимо потужність та набір елементів цієї підмножини;

5) розв'яжемо задачу навчання, окреслену в пункті 2), враховуючи, що наявні спостереження є зашумленими, спотвореними.

Кожен з підрозділів цього розділу матиме відповідну лаконічну назву: «Моделювання об'єкта дослідження», «Задача навчання», «Задача декодування», «Задача локалізації» та «Задача навчання за спотвореними спостереженнями».

2.1 Моделювання об'єкта дослідження

2.2 Задача навчання

2.3 Задача декодування

2.4 Задача локалізації

2.5 Задача навчання за спотвореними спостереженнями

Висновки до розділу 2

Наприкінці розділу наводяться короткі підсумки.

3 РЕЗУЛЬТАТИ ЧИСЕЛЬНОГО ЕКСПЕРИМЕНТУ

3.1 Моделювання об'єкта дослідження

3.2 Задача навчання

3.3 Задача декодування

3.4 Задача локалізації

3.5 Задача навчання за спотвореними спостереженнями

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.

ПЕРЕЛІК ПОСИЛАНЬ

1. *Koski T.* Hidden Markov models for bioinformatics. — 2002-е вид. — New York, NY : Springer, 11.2001. — (Computational Biology).
2. *Rabiner L.* A tutorial on hidden Markov models and selected applications in speech recognition // Proceedings of the IEEE. — 1989. — Т. 77, № 2. — С. 257—286. — DOI: 10.1109/5.18626. — URL: https://doi.org/10.1109/5.18626.
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. *Norris J. R.* Markov Chains. — Cambridge University Press, 02.1997. — DOI: 10.1017/cbo9780511810633. — URL: https://doi.org/10.1017/cbo9780511810633.
5. *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.
6. *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.