



Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет прикладної математики
Кафедра прикладної математики

Реферат

З дисципліни «Практичний курс англійської мови
для наукової комунікації. Частина 1»

«Hidden Markov Models and it's applications»

Роботу виконав:
Студент групи КМ-31мн,
Цибульник Антон Владиславович

Викладач:
Муханова Олена Миколаївна

Київ 2023

Contents

Abstract	2
Abstract (translation)	3
Introduction.	4
1 Overview of Markov chains	5
1.1 States: the building blocks of dynamics	5
1.2 Observations: insight into the system's dynamics	5
1.3 Transitions: navigating the dynamics	6
1.4 Conclusions to section 1	6
2 Applications of Hidden Markov Models	7
2.1 Speech Recognition: deciphering phonemes and beyond	7
2.2 Bioinformatics: decoding the genetic blueprint.	7
2.3 Financial Modeling: navigating market dynamics	8
2.4 Natural Language Processing: parsing linguistic complexity.	8
2.5 Conclusions to section 2	8
3 Markov Chain Monte Carlo in application to HMMs.	9
3.1 Introduction to Markov Chain Monte Carlo (MCMC).	9
3.2 Recall of Hidden Markov Models	9
3.3 Markov Chain Monte Carlo in HMMs.	10
3.4 Conclusions to section 3	11
References	12

Abstract

This research delves into the intricate realm of Hidden Markov Models (HMMs) and explores their wide-ranging applications in various domains. Commencing with a concise overview of the foundational Markov chain concept, we establish the groundwork for understanding the evolution of HMMs.

Markov chains, characterized by states, observations, and transitions, form the backbone upon which HMMs extend, introducing hidden states and observations as well as emission and transition probabilities. The study emphasizes the crucial role of HMMs in capturing dynamic processes with unobservable states, highlighting their versatility and applicability.

The core focus of this research lies in the applications of HMMs within Applied Mathematics, with a spotlight on their pivotal role in fields such as Speech Recognition, Bioinformatics, Financial Modeling, and Natural Language Processing. From phoneme identification and gene prediction to stock price forecasting and part-of-speech tagging, HMMs emerge as indispensable tools for modeling and understanding complex systems.

Furthermore, the research investigates the integration of Markov Chain Monte Carlo (MCMC) methods within the context of HMMs. By employing MCMC techniques, we explore how HMMs can be refined and adapted to better suit real-world scenarios, offering a more robust and accurate representation of dynamic systems.

In conclusion, this research not only illuminates the fundamental principles of Hidden Markov Models but also underscores their significance across diverse applications. By incorporating MCMC methods, we aim to enhance the precision and adaptability of HMMs, paving the way for continued advancements in understanding complex, hidden processes within Applied Mathematics.

Abstract (translation)

Реферат заглиблюється у складну сферу прихованих марковських моделей (ПММ) і вивчає їхнє широке застосування в різних галузях. Починаючи з короткого огляду фундаментальної концепції ланцюгів Маркова, ми закладаємо основу для розуміння еволюції ПММ.

Ланцюги Маркова, що характеризуються станами, спостереженнями та переходами, утворюють основу, на якій розвиваються ПММ, вводячи приховані стани та спостереження, а також імовірності переходів. Дослідження підкреслює вирішальну роль ПММ у відображенні динамічних процесів з неспостережуваними станами, підкреслюючи їхню універсальність у застосуваннях.

Основна увага у рефераті зосереджена на застосуванні ПММ у прикладній математиці з акцентом на їхній ключовій ролі в таких галузях, як розпізнавання мовлення, біоінформатика, фінансове моделювання та обробка природної мови. Від ідентифікації фонем і прогнозування генів до прогнозування цін на акції та виділення частин мови, ПММ стають незамінними інструментами для моделювання та розуміння складних систем.

Крім того, у роботі розглянуто інтеграційний метод марковських ланцюгів Монте-Карло (МСМС) в контексті ПММ. Використовуючи методи МСМС проводять дослідження, як ПММ можуть бути вдосконалені та адаптовані, щоб краще відповідати реальним сценаріям, пропонуючи більш надійне і точне представлення динамічних систем.

Отже, реферат не лише висвітлює фундаментальні принципи прихованих марковських моделей, але й підкреслює їхню важливість у різних сферах застосування. При цьому, впроваджуючи методи МСМС дослідники прагнуть підвищити точність і адаптивність ПММ, прокладаючи шлях до подальшого прогресу в розумінні складних, прихованих процесів у прикладній математиці.

Introduction

Markov models have a wide and effective arsenal of tools for analyzing the dynamics of systems, whose behavior at each subsequent moment in time is determined only by the current state of the system and does not depend on the nature of the evolution at previous moments in time.

At the same time, when direct observation of the Markov chain evolution is impossible or limited, hidden Markov chain models (HMMs) are used. In this case, the analysis of the process behavior is based on some indirect information about the «hidden», true states of the chain.

For example, in bioinformatics (Koski, 2001, chapter 9) the Markov chain apparatus is used to study the evolution of DNA molecules over time, considering the state of the system to be a bound sequence of so-called nucleotides that are formed over the alphabet of four nitrogenous bases $\{T, C, A, G\}$.

The existence of statistical dependencies in the alternation of phonemes or words in natural languages determines the effectiveness of using hidden Markov Models for such tasks as creating voice commands, transcription services, and voice assistants (Rabiner, 1989).

The tasks of sign language recognition are no exception (Chaaaraoui et al., 2013): for example, by representing gestures as sequences of hidden states, HMMs can recognize the dynamics and variations of hand movements.

1 Overview of Markov chains

Markov Chains, a fundamental concept in probability theory, serve as a cornerstone in understanding the dynamic evolution of systems over time. They are stochastic processes that exhibit the Markov property, which posits that the future state of the system depends solely on its current state, independent of how it arrived at that state. This section delves into the key components of Markov Chains, namely states, observations, and transitions, shedding light on their collective significance in modeling a diverse array of phenomena.

1.1 States: the building blocks of dynamics

At the heart of Markov Chains lie the concept of states, representing the distinct, identifiable conditions or situations that a system can occupy at any given time. These states encapsulate the essential characteristics defining the system's behavior, serving as the foundational elements upon which the entire Markov process unfolds. States can be discrete or continuous, depending on the nature of the system under consideration.

In a discrete state space, the system transitions between distinct, separate conditions. For example, consider a weather model with states such as «sunny», «cloudy», and «rainy». Conversely, continuous state spaces involve an uncountable range of possible conditions. A stock price, for instance, can be considered as a continuous state variable, with the system evolving smoothly through an infinite set of potential values.

1.2 Observations: insight into the system's dynamics

Observations in a Markov Chain represent the information available to an external observer about the system's state at a given time. These observations may or may not fully reveal the underlying state, introducing an element of uncertainty. The relationship between states and observations is encapsulated by probability distributions, providing a

statistical framework for understanding the likelihood of particular observations given the current state.

The connection between states and observations is crucial in scenarios where not all aspects of the system's state are directly observable. This characteristic becomes particularly prominent in Hidden Markov Models (HMMs), where certain states remain unobservable, adding an extra layer of complexity to the modeling process.

1.3 Transitions: navigating the dynamics

Transitions constitute the essence of Markov Chains, embodying the shifts between different states over time. The transitions are guided by transition probabilities, representing the likelihood of moving from one state to another. These probabilities encapsulate the dynamics of the system, offering insights into the tendencies and patterns governing its evolution.

The transition matrix, a fundamental component of Markov Chains, succinctly captures the probabilities of moving from one state to another in a single time step. It encapsulates the inherent memorylessness of Markov Chains, where the probability of transitioning to a new state depends solely on the current state, irrespective of the past sequence of events.

1.4 Conclusions to section 1

In summary, Markov Chains provide a powerful framework for modeling dynamic systems through the interplay of states, observations, and transitions. States define the possible conditions of the system, observations offer insights into its current state, and transitions navigate the evolving dynamics, guided by probabilities. This foundational understanding sets the stage for the extension of Markov Chains into more complex models, such as Hidden Markov Models, as we explore in subsequent sections.

2 Applications of Hidden Markov Models

Hidden Markov Models (HMMs) stand as a versatile and powerful tool in the arsenal of Applied Mathematics, finding applications across diverse domains. Their inherent ability to model systems with unobservable states makes them particularly adept at capturing complex real-world dynamics. This section navigates through the applications of HMMs in four key fields within Applied Mathematics: Speech Recognition, Bioinformatics, Financial Modeling, and Natural Language Processing.

2.1 Speech Recognition: deciphering phonemes and beyond

In the realm of Speech Recognition, HMMs have revolutionized the way machines interpret and understand spoken language. Phoneme identification, a foundational task in this domain, involves categorizing and distinguishing individual speech sounds. HMMs excel in modeling the temporal dependencies and variability inherent in spoken language, enabling accurate identification of phonetic units. Moreover, HMMs are instrumental in language modeling, allowing systems to predict and comprehend the sequential nature of spoken words, thus enhancing the overall accuracy and efficiency of speech recognition technologies.

2.2 Bioinformatics: decoding the genetic blueprint

In Bioinformatics, Hidden Markov Models play a pivotal role in unraveling the complexities of biological sequences. Gene prediction, a critical task in genomics, involves identifying the locations of genes within DNA sequences. HMMs, with their capacity to model the hidden states representing gene structures, contribute significantly to accurate gene prediction. Additionally, HMMs find application in predicting protein structures, aiding researchers in understanding the functional implications of different amino acid arrangements.

2.3 Financial Modeling: navigating market dynamics

Financial markets, characterized by dynamic and often unpredictable behavior, benefit from the application of HMMs in modeling stock prices and assessing risk. HMMs prove instrumental in stock price prediction by capturing the underlying patterns and transitions in market trends. Traders and investors leverage these models to make informed decisions based on the evolving dynamics of financial markets. Furthermore, HMMs contribute to risk assessment by modeling the latent factors influencing market fluctuations, providing a valuable tool for anticipating and managing financial risks.

2.4 Natural Language Processing: parsing linguistic complexity

In the domain of Natural Language Processing, HMMs offer a robust framework for tackling intricate linguistic tasks. Part-of-speech tagging, a fundamental task in syntactic analysis, involves assigning grammatical categories to individual words in a sentence. HMMs excel in modeling the sequential dependencies between words, facilitating accurate and context-aware part-of-speech tagging. Additionally, HMMs are employed in Named Entity Recognition (NER), extracting and categorizing entities such as names, locations, and organizations from text. The ability to capture hidden states enables HMMs to discern subtle contextual cues, enhancing the accuracy of language processing tasks.

2.5 Conclusions to section 2

In conclusion, the applications of Hidden Markov Models in Applied Mathematics extend across a spectrum of disciplines, showcasing their adaptability and effectiveness in modeling complex systems. From deciphering spoken language to decoding the genetic blueprint, navigating financial markets, and parsing linguistic complexity, HMMs continue to be at the forefront of innovation, contributing to advancements in diverse fields within Applied Mathematics.

3 Markov Chain Monte Carlo in application to HMMs

3.1 Introduction to Markov Chain Monte Carlo (MCMC)

Before delving into the intricate relationship between Markov Chain Monte Carlo (MCMC) (Andrieu and Thoms, 2008) and Hidden Markov Models (HMMs), it is imperative to illuminate the significance of MCMC in the realm of Applied Mathematics. MCMC is a powerful computational technique designed to explore and sample from complex probability distributions, especially when analytical solutions are elusive or computationally infeasible.

At its core, MCMC operates on the principles of Markov Chains, where the generation of a sequence of dependent samples converges towards the desired distribution. This sequential sampling technique has become a cornerstone in Bayesian statistics, enabling practitioners to make inferences about parameters, model structures, and uncertainties. The elegance of MCMC lies in its ability to transform complex problems into a series of simpler, interconnected steps, allowing for efficient exploration of high-dimensional parameter spaces.

3.2 Recall of Hidden Markov Models

Hidden Markov Models, celebrated for their prowess in modeling systems with unobservable states, gain additional depth and flexibility through the integration of MCMC techniques. The marriage of MCMC and HMMs unlocks new avenues for refining model parameters, accommodating uncertainties, and improving the overall accuracy of the model.

In the context of HMMs, the hidden states represent the unobservable variables that govern the system's dynamics. Traditionally, parameter estimation in HMMs relies on algorithms like the Baum-Welch algorithm or the Expectation-Maximization (EM) algorithm. However, these methods may have limitations, especially in scenarios with complex, high-dimensional state spaces or when dealing with incomplete and noisy data.

By leveraging the sequential, Markovian nature of MCMC, we can traverse the parameter space of HMMs with greater flexibility. This is particularly valuable when dealing with intricate models where hidden states are not explicitly observable or when the model structure is uncertain. MCMC allows for the exploration of a wide range of potential parameter values, providing a more comprehensive understanding of the hidden dynamics within the system.

3.3 Markov Chain Monte Carlo in HMMs

One of the primary applications of MCMC in the realm of HMMs is in model refinement. Traditional algorithms may struggle with finding optimal parameter values in highly complex or nonlinear systems. MCMC steps in by iteratively proposing new parameter values, evaluating their likelihood given the observed data, and accepting or rejecting the proposals based on a carefully crafted acceptance criterion. This iterative process allows the model to adapt and refine its parameters, gradually converging towards a more accurate representation of the underlying dynamics.

Furthermore, MCMC provides a natural framework for handling uncertainties associated with hidden states. In scenarios where the true state is ambiguous or challenging to observe directly, MCMC offers a probabilistic approach to characterize the uncertainty inherent in the system. Bayesian inference, facilitated by MCMC, enables the incorporation of prior knowledge and observed data to update the model's beliefs about the hidden states.

While the synergy between MCMC and HMMs opens new doors in modeling and inference, it is not without challenges. The computational demands of MCMC, especially in high-dimensional spaces, can be substantial. Careful consideration must be given to algorithmic efficiency, convergence diagnostics, and the choice of appropriate priors to ensure the reliability and accuracy of the results.

3.4 Conclusions to section 3

In conclusion, the connection between MCMC and Hidden Markov Models is a symbiotic relationship that enhances the capabilities of both. MCMC, with its ability to navigate complex parameter spaces, provides a powerful tool for refining and adapting HMMs to real-world complexities. This integration not only advances our understanding of hidden dynamics but also offers a robust framework for tackling complex problems in Applied Mathematics.

References

- Andrieu, C., & Thoms, J. (2008). A tutorial on adaptive mcmc. *Statistics and Computing*, 18(4), 343–373. <https://doi.org/10.1007/s11222-008-9110-y>
- Charaoui, A. A., Climent-Pérez, P., & Flórez-Revuelta, F. (2013). One shot learning for gesture recognition using hmms and hand appearance features. *Pattern Recognition Letters*, 34(9), 1009–1017.
- Koski, T. (2001). *Hidden markov models for bioinformatics* (2002nd ed.). Springer.
- Rabiner, L. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286. <https://doi.org/10.1109/5.18626>