1.5em 0pt

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED BOUDIAF-M'SILA
Faculté des Mathématiques et d'Informatique
Département d'Informatique

Thèse

En vue de l'obtention du diplôme de Doctorat LMD en Informatique

---

Sur la Stabilité des modèles stochastiques

Présentée par :

SALAH EDDINE SEMATI

Les membres du jury:

Prof. ALLAOUA HEMMAK	Univ.M'sila	Président
Prof. ABDELKADER GASMI	Univ.M'sila	Encadreur
Prof. SALIM BOUAMAMA	Univ.Setif1	Examinateur
M.C.A. MAKHLOUF BENAZI	Univ.M'sila	Examinateur
M.C.A. MOUSSA SEMCHEDINE	Univ.Setif1	Examinateur

ANNÉE UNIVERSITAIRE : 2023-2024

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

1985

جامعة محمد بوضياف - المسيلة
Université Mohamed Boudiaf - M'sila

University of Mohamed Boudiaf-M'sila
Faculty of mathematics and computer science
Department of computer science

Thesis

For obtaining the diploma of Doctorate LMD in Computer Science

On the Stability of stochastic models

Presented By:

Salah Eddine Semati

The jury members:

Prof. Allaoua Hemmak	Univ.M'sila	President
Prof. Abdelkader Gasmi	Univ.M'sila	Supervisor
Prof. Salim Bouamama	Univ.Setif1	Examiner
M.C.A. Makhlouf Benazi	Univ.M'sila	Examiner
M.C.A. Moussa Semchedine	Univ.Setif1	Examiner

Academic year: 2023-2024

بسم الله الرحمن الرحيم

وقل
رب زدني علما

صدق الله العظيم

ABSTRACT

In the ever-evolving landscape of digital systems, the stability of stochastic models is a critical aspect influencing decision-making processes and algorithmic applications. This thesis delves into the examination of stability properties within various models, with a primary focus on Markov Interval Chains (IMC). The study aims to contribute to the broader understanding of stability in stochastic systems.

Commencing with an extensive literature review, the research surveys existing models and methodologies pertaining to stability. Subsequently, it outlines specific research objectives and defines the scope, employing mathematical analysis to explore stability in-depth.

Through case studies and practical applications, the thesis demonstrates the relevance of stability analysis in diverse contexts.

The presentation and discussion of results emphasize the implications for theoretical frameworks and practical implementations. The conclusion consolidates key findings and suggests potential directions for future research, contributing substantively to the field of stability analysis for stochastic models.

Keywords: *Interval Markov Chain, Stability, Stochastic Process*

ملخص

في المشهد المتطور باستمرار للأنظمة الرقمية، يعد استقرار النماذج العشوائية جانبًا حاسمًا يؤثر على عمليات صنع القرار والتطبيقات الخوارزمية. تتعمق هذه الأطروحة في فحص خصائص الاستقرار في نماذج مختلفة، مع التركيز بشكل أساسي على سلاسل ماركوف الفاصلة (MIC). تهدف الدراسة إلى المساهمة في فهم أوسع للاستقرار في الأنظمة العشوائية، بشكل مستقل عن المنشورات السابقة. بدءاً بمراجعة واسعة النطاق للمنشورات السابقة، يقوم البحث بمسح النماذج الموجودة والنماذج الحالية و منهجيات الاستقرار. وبعد ذلك، يحدد أهدافًا بحثية محددة ويحدد النطاق، ويستخدم التحليل الرياضي والمحاكاة لاستكشاف الاستقرار بشكل متعمق.

من خلال دراسات الحالة والتطبيقات العملية، توضح الأطروحة أهمية تحليل الاستقرار في سياقات متنوعة. ويؤكد عرض النتائج ومناقشتها على الآثار المترتبة على الأطر النظرية والتطبيقات العملية. ويعزز الاستنتاج النتائج الرئيسية ويقترح الاتجاهات المحتملة للبحث المستقبلي، مما يساهم بشكل كبير في مجال تحليل الاستقرار للنماذج العشوائية.

الكلمات المفتاحية: سلاسل ماركوف على مجالات، الاستقرار، النماذج العشوائية.

RÉSUMÉ

Dans le paysage en constante évolution des systèmes numériques, la stabilité des modèles stochastiques est un aspect essentiel qui influence les processus décisionnels et les applications algorithmiques. Cette thèse se penche sur l'examen des propriétés de stabilité au sein de divers modèles, en mettant l'accent principalement sur les chaînes d'intervalles de Markov (IMC). L'étude vise à contribuer à une compréhension plus large de la stabilité dans les systèmes stochastiques, indépendamment des publications antérieures.

En commençant par une revue approfondie de la littérature, la recherche examine les modèles et méthodologies existants relatifs à la stabilité. Par la suite, il décrit les objectifs de recherche spécifiques et définit la portée, en utilisant des analyses mathématiques et des simulations pour explorer en profondeur la stabilité.

À travers des études de cas et des applications pratiques, la thèse démontre la pertinence de l'analyse de stabilité dans divers contextes.

La présentation et la discussion des résultats mettent l'accent sur les implications pour les cadres théoriques et les mises en œuvre pratiques. La conclusion consolide les principales conclusions et suggère des orientations potentielles pour de futures recherches, contribuant ainsi de manière substantielle au domaine de l'analyse de la stabilité des modèles stochastiques.

Mots-clés : *Chaîne De Markov, Stabilité, Modèles Stochastiques*

ACKNOWLEDGEMENTS

Writing a PhD thesis is a long and arduous journey and I am deeply grateful to all those who have supported me along the way. .First of all, I am deeply grateful to my GOD for empowering me in all my life stations, including this journey.

I would like to express my sincere gratitude to my advisor, Prof. Abdelkader Gasmi, for its invaluable guidance, support, and encouragement throughout my research.

I am so thankful to the jury members: Prof.Allaoua Hemmak, Prof Salim Bouamama, Dr.Makhlouf Benazi , and Dr.Moussa Semchedine accepting the request to join us and contribute to evaluating the presented work based on their rich experiences.

I would like to express my heartfelt gratitude to my family, who have always been there for me, providing love, support, and encouragement throughout my academic journey.

Finally, I am also thankful to my colleagues and friends who have provided me with support, encouragement, and motivation during my PhD journey. Their friendship have made this experience truly unforgettable.

This work would not have been possible without the support and help of all of these individuals, and I am deeply grateful to each and every one of them.

TABLE OF CONTENTS

LIST OF TABLES

TABLE **Page**

LIST OF FIGURES

LISTINGS

LIST OF ABBREVIATIONS

IMC	Interval Markov Chain
PR	Page Rank
DTSP	Discrete-Time Stochastic Process
CTSP	Continuous-Time Stochastic Process
M(P$_i$)	Set of Pages that Link to Page P_i
L(P$_j$)	Number of Outbound Links on Page P_j
d	Damping Factor (typically set to 0.85 in Page Rank)
CDF	Cumulative Distribution Function
PDF	Probability Density Function
MC	Markov Chain
MDP	Markov Decision Process
QL	Q-Learning
SP	Stochastic Process
HMM	Hidden Markov Model

CHAPTER 1

INTRODUCTION

"The validation of a model is not that it is 'true' but that it generates good testable hypotheses relevant to important problems"

Richard Levins, "The Strategy of Model Building in Population Biology", 1966.

1.1 Problem Statement

Stochastic models serve as indispensable tools in various domains of science and engineering, enabling the analysis and prediction of complex systems under uncertainty. In the realm of computer science, where decisions are made based on data and algorithms, the stability of stochastic models holds paramount importance.

The stability of stochastic models pertains to their ability to maintain predictable behavior over time and under varying conditions. Stability ensures that the modeled system remains bounded and does not exhibit erratic or uncontrolled behavior, which is critical for reliable decision-making and system performance in computer science. Understanding and assessing the stability of stochastic models are therefore essential endeavors that drive the advancement and application of these models in real-world scenarios.

In this context, the Markov interval chain (IMC) method emerges as a promising approach to address stability concerns in stochastic models, particularly in decision-making processes involving dynamic and uncertain environments. The IMC method extends traditional Markov chain models to accommodate interval uncertainty, offering a more comprehensive framework for analyzing and predicting system behavior.

The research background encompasses previous studies and developments in the field of stochastic modeling and stability analysis. Previous works have investigated various aspects of stability, including criteria for stability assessment, methods for stability analysis, and applications of stable stochastic models in diverse domains. However, there remains a need for further research to deepen our understanding of stability in stochastic models, particularly in the context of computer science.

The problem statement addressed in this thesis revolves around the investigation of stability in stochastic models, with a primary focus on the IMC approach. The objectives of this research include analyzing stability properties, developing stability criteria, applying stable models to real-world decision problems, and evaluating their performance.

By addressing these objectives, this thesis aims to contribute to the advancement of knowledge in the field of stability analysis in stochastic models, particularly within the domain of computer science. The subsequent chapters will delve into a comprehensive review of literature, theoretical frameworks, stability analysis methodologies, applications of stable models, and discussion of findings, paving the way for a deeper understanding of this critical aspect of stochastic modeling.

1.2 Thesis Objective and Contribution

1.2.1 Objective

The primary objective of this thesis is to investigate the stability of stochastic models, with a focus on the Markov interval chain (IMC) approach proposed in the published article titled "Markov Interval Chain (IMC) for Solving a Decision Problem." The specific objectives are as follows:

I. To Analyze Stability Properties: The thesis aims to analyze the stability properties of stochastic models, including traditional Markov chains and the IMC approach, under various conditions and scenarios.

II. To Develop Stability Criteria: Another objective is to develop stability criteria and methodologies for assessing the stability of stochastic models, particularly those incorporating interval uncertainty.

III. To Apply Stable Models to Real-World Problems: The thesis seeks to apply stable stochastic models, including the IMC method, to real-world decision problems . This involves demonstrating the practical utility and effectiveness of stable models in addressing complex decision-making challenges.

IV. To Evaluate Performance and Effectiveness: Lastly, the thesis aims to evaluate the performance and effectiveness of stable stochastic models in comparison to traditional approaches.

1.2.2 Contribution

This thesis contributes to the field of stochastic modeling and computer science in several ways:

I. Advancement of Stability Analysis: By analyzing the stability properties of stochastic models, the thesis contributes to advancing our understanding of stability in dynamic and uncertain systems. This includes identifying conditions under which stability can be ensured and developing criteria for stability assessment

II. Methodological Development: The development of stability criteria and methodologies enhances the toolbox available for stability analysis in stochastic models. The thesis contributes new approaches for assessing stability, particularly in models incorporating interval uncertainty.

III. Practical Applications: The application of stable stochastic models, such as the IMC method, to real-world decision problems demonstrates their practical utility and effectiveness in computer science. This contributes to bridging the gap between theoretical developments and practical applications.

IV. Performance Evaluation: By evaluating the performance and effectiveness of stable models, the thesis provides insights into the impact of stability on system performance, accuracy, and robustness.

Overall, this thesis makes significant contributions to both theoretical understanding and practical applications of stable stochastic models in computer science, paving the way for improved decision-making and system performance in dynamic and uncertain environments.

1.3 Thesis organization

The rest of the thesis has been structured in two parts: the first one (Chapters 2, 3) is devoted to the state-of-the-art in the Stochastic models field . The second one (Chapters 4, and 5) presented our contributions to the Stochastic Models. Below, we briefly show the content of each Chapter:

- Chapter 2 "Background": In this chapter we talk about Overview of Mathematical Modeling and we start talking about types of models and focus on stochastic models as well as Importance of Stability in Stochastic Models, then we talk about Markov Interval Chain and we explain Page Rank Algorithm, then we talk about the most of the concepts related to probability and then addresses queuing theory and its relationship to random models and their stability .
- Chapter 3 " Matrix and Vector Calculations": This chapter covers everything related to the matrix and vectors.
- Chapter 4 "Markov interval Chain (IMC) for solving a decision problem": In this chapter we address the problem of decision using interval Markov chains, by giving a real example based on a article we published
- Chapter 5 "Application of interval Markov chain in the page rank algorithm":In this chapter we talk about the use of interval Markov chains in the Page rank algorithm that Google uses to adjust the ranking of pages on the web after the crawling and indexing process.

1.4 Publications list

This research has resulted in two publications in international journals and conference papers.

1.4.1 Journal paper:

- **Semati, Salah eddine, and Abdelkader Gasmi**, “Markov interval chain (IMC) for solving a decision problem”, *OPSEARCH*, pp. 802-811, 2023. doi: `https://doi.org/10.1007/s12597-023-00632-5`

1.4.2 Conference paper:

- **Semati, Salah eddine, and Abdelkader Gasmi**, “Application of interval Markov chain in the page rank algorithm”, *In:The First International Conference on Nonlinear Mathematical Analysis and Its Applications,Bordj-Bou-Arreridj, Algeria from May 14-15 2024.*

Part I

State-of-the-art

CHAPTER 2

BACKGROUND STUDY AND LITERATURE REVIEW

> "Probability theory is nothing but common sense reduced to calculation."
>
> *Pierre-Simon Laplace, French scholar and polymath.*

2.1 Introduction

The chapter titled "Background" provides a comprehensive overview of the foundational concepts and methodologies essential for understanding the various scientific and engineering disciplines discussed in the thesis. It begins by delving into mathematical modeling, which is a critical tool for researchers to translate complex systems and processes into manageable and predictive mathematical frameworks. The section outlines different classifications of mathematical models such as static vs. dynamic, linear vs. non-linear, discrete vs. continuous, mechanistic vs. data-driven, and deterministic vs. stochastic models. These classifications highlight the diverse nature and applications of mathematical modeling in various fields. Following this, the chapter explores deterministic and stochastic models in detail, emphasizing their roles in providing systematic understanding and predictive insights into systems characterized by both certainty and randomness. The importance of stability in stochastic models is also discussed, underscoring its relevance in ensuring the reliability and applicability of these models in real-world scenarios. Additionally, the chapter covers the Page Rank algorithm, a cornerstone of modern search engine technology, elucidating its conceptual foundation and mathematical formulation. Lastly, the chapter touches on queuing theory, which applies mathematical principles to analyze and optimize the behavior of waiting lines in various systems, particularly in informatics. Through this diverse yet interconnected set of topics, the chapter sets the stage for a deeper exploration of the specific applications and advancements in these areas as discussed in the subsequent sections of the thesis

2.2 Mathematical Modelling

Scientific researchers reveal the nature of systems and processes that appear in various fields of computer science by expressing their ideas and understanding of systems through an efficient and effective way known as mathematical modelling [1]. The translation of ideas into mathematical models provides many advantages, including process prediction [2], quality control [3], discovery of information [4], measurement optimisation [5], and system optimisation [6]. A tremendously increasing number of mathematical models have been developed in computer science to gain better insights into realistic systems. Mathematical models, depending on their features and structures, can be classified into the following groups:

- Static vs dynamic: A static model usually describes the steady state of a system, or a process which is considered in equilibrium, while a dynamic model concerns the time-dependent change of variables in a system. Both static and dynamic models can exist in one system, even described by the same set of mathematical expressions, and characterise the system behaviour at different time periods. For example, static models in computer networks can analyze the equilibrium state of network traffic, while dynamic models can describe the time-varying behavior of network congestion control algorithms.

- Linear vs non-linear: In mathematical programming, the linearity/non-linearity of expressions (i.e. objective functions, constraints) significantly affects the model solvability and precision of optimisation. In machine learning, linear models like linear regression are easier to solve and interpret, whereas non-linear models such as neural networks can capture more complex patterns in data but require more computational resources and advanced techniques to optimize .

- Discrete vs continuous: A discrete model in time usually describes a system whose time-evolution can be considered as a result of individual and separate events occurring in the system. Examples include the execution of discrete algorithms, such as sorting algorithms or the analysis of packet-switching in networks. A continuous model describes a system in which some variables hold continuous change in time domain, and is usually applied to characterise the time-dependent evolution of physical properties, such as the fluid dynamics in cooling systems of data centers or the continuous updates in real-time systems by differential equations .
- Mechanistic vs data-driven: Mechanistic models, also known as knowledge-driven models, are formulated based on well-understood processes and hypotheses, such as the modeling of computer architectures and their performance characteristics. Data-driven models, on the other hand, become very useful to characterize the relationship between system conditions and outputs using statistical approaches, such as machine learning algorithms for classification and regression, where the model is trained on empirical data to make predictions or decisions.
- Deterministic vs stochastic: Deterministic models are implemented with explicit relations among several (non-random) model variables, such as deterministic finite automata in theoretical computer science, where the same input will always produce the same output. Stochastic models, however, incorporate randomness and can be used to model systems with inherent uncertainty, such as probabilistic algorithms and Markov chains in queuing theory, where outcomes are described by probability distributions.

Nevertheless, the majority of mathematical models are composite and composed of several features rather than only one listed above. The combination of these features contributes to great diversity and flexibility of mathematical modelling on systems of interest. Particularly, the classification of models into deterministic and stochastic ones allows one to study the same system from different perspectives. Deterministic models present the mean field evolution of system states from a macroscopic perspective, while stochastic models take into account the uncertainties in system states and therefore provide additional predictive information about system evolutions.

2.2.1 Deterministic models

A deterministic model is represented by a set of mathematical expressions, usually differential and algebraic equations (DAEs), for a system or process, establishing a logical relationship between variables whose physical quantities are obtainable from the system or to be predicted and parameters whose values are unknown or unmeasurable in the system. In the engineering community, deterministic models are extensively used to study a variety of systems, including heat, mass and momentum transfers Patankar and Spalding[7], fluid dynamics Ferziger et al.[8], and chemical reaction processes Schmidt and Ray[9]. These models allow one to identify the evolution of system states from a macroscopic perspective. The general form of a deterministic model can be described by the following expression:

$y(t) = f(x(t), u(t), \theta)$ where

y(t): Output or response at time t

x(t): State variables or system states at time t

u(t): Control inputs or external influences at time t.

θ: Parameters of the model.

Example:Consider a simple deterministic model of a car's motion:

s(t)=s(0)+v× t

Where:

s(t) is the position of the car at time t.

s(0)is the initial position of the car.

v is the constant speed of the car.

t is time.

If you know the initial position s(0) and the speed v, you can predict the position s(t) at any time t.

For example, if s(0)=0 (the car starts at the origin) and v=60 km/h, then after 1 hour (t=1), the car's position will be

$$s(1) = 0 + 60 \times 1 = 60km.$$

In this model, if you run the scenario again with the same initial conditions, the car will always be at 60 km after 1 hour, showing that the model is deterministic. Detailed investigations on general deterministic models are not included in this Thesis. More information about the identification and improvement of deterministic models can be found in Quaglio [10].

2.2.2 Stochastic models

stochastic models, which are mathematical models used to describe systems subject to randomness. Stochastic models are widely applied in various fields such as finance, engineering, biology, and informatics due to their ability to capture uncertainty and variability in real-world phenomena [11]. Stochastic processes form the foundation of stochastic modeling. A stochastic process is a collection of random variables indexed by time or another parameter. These processes can be discrete or continuous and are often used to model dynamic systems evolving over time [12].

2.2.2.1 Discrete-Time Stochastic Processes

In discrete-time processes, the state of the system is observed at discrete points in time. Examples include Markov chains, which describe systems with memoryless transitions, and branching processes, which model population growth and extinction.

2.2.2.2 Continuous-Time Stochastic Processes

Continuous-time processes are characterized by continuous evolution over time. Stochastic differential equations (SDEs) are commonly used to model continuous-time stochastic processes. SDEs describe systems where randomness affects the rate of change of a quantity .

2.2.3 Some Types of Stochastic Models

Stochastic models come in various forms, each suited to different types of systems and applications.

2.2.3.1 Markov Models

Markov models are widely used in both discrete and continuous settings. These models assume that the future behavior of a system depends only on its current state, not on its past history. Markov chains and Markov processes are common examples of such models.

2.2.3.2 Stochastic Differential Equations (SDEs)

SDEs are differential equations with stochastic terms, often used to describe systems subject to continuous random fluctuations. They find applications in physics, biology, finance, and engineering, among other fields.

2.2.3.3 Monte Carlo Methods

Monte Carlo methods are computational techniques based on random sampling. These methods are used to solve problems by simulating random processes and averaging the results. They are particularly useful for estimating quantities and solving optimization problems.

2.2.3.4 Agent-Based Models

Agent-based models simulate the interactions of individual agents within a system to understand emergent behavior. These models are used in sociology, ecology, and economics, among other disciplines, to study complex systems.

2.2.4 Applications of Stochastic Models

Stochastic models find applications in diverse fields, addressing a wide range of problems: In finance, stochastic models are used to model stock prices, interest rates, and option pricing. Geometric Brownian motion and jump-diffusion processes are common models used in this domain [13]. Engineering applications of stochastic models include reliability analysis, signal processing, and control systems. Stochastic differential equations are used to model dynamic systems subject to noise and uncertainty . Stochastic models are used in biology to study population dynamics, gene expression, and biochemical reactions. Gillespie's algorithm is a popular method used to simulate biochemical systems. In informatics, stochastic models are applied in machine learning, optimization, and algorithm design. Stochastic gradient descent and Markov chain Monte Carlo methods are widely used in these applications

2.2.5 Importance of Stability in Stochastic Models

Stability is a fundamental property of stochastic models, crucial for their reliability and applicability across various fields. Understanding and ensuring stability in stochastic models is essential for developing robust and predictive models that can be effectively utilized in real-world applications. Before delving into the importance of stability, let's briefly review the key stability concepts in stochastic models:

1. **Asymptotic Stability:**A stochastic model is asymptotically stable if its behavior converges to a stable state as time progresses, even in the presence of random fluctuations.

2. **Lyapunov Stability:**Lyapunov stability ensures that a small perturbation from an equilibrium state does not lead to unbounded growth, indicating the system's robustness.

2.2.6 Stability in Probability:

Stability in probability ensures that the probability of the system remaining within a certain region converges to unity as time goes to infinity, indicating the system's long-term behavior.
Stability plays a crucial role in numerous fields where stochastic models are applied:

1. **Finance:**In finance, stability is paramount for risk management and portfolio optimization. Stable financial models ensure that investment strategies remain viable over time, even in volatile market conditions
   Example: Stable models for stock price dynamics help investors make informed decisions and hedge against market risks.

2. **Engineering:**In engineering, stability is essential for designing robust control systems, ensuring stable operation under varying conditions and disturbances. Example: Stable models for feedback control systems ensure the stability of aircraft autopilots and robotic systems .

3. **Biology:**In biology, stability is critical for understanding the dynamics of ecological systems and population growth. Example: Stable models for predator-prey interactions help ecologists predict the long-term stability of ecosystems.

4. **Informatics:**In informatics, stability is vital for designing reliable algorithms and machine learning models. Example: Stable optimization algorithms ensure the convergence of machine learning models, improving their accuracy and generalization.

2.2.7 Previous Work on Stability of Stochastic Models

This subsection provides an extensive review of previous studies on the stability analysis of stochastic models. Stability analysis is fundamental in understanding the behavior of stochastic systems, crucial for various applications in computer science, machine

learning, and optimization.
Stability analysis of stochastic processes has been extensively studied across various disciplines. Here, we review several key works in this area: C. Obropta and Josef S. Kardos [14] reviewed deterministic, stochastic, and hybrid modeling approaches in flood management in their research paper, highlighting six deterministic models, three stochastic models, and three hybrid models. The hybrid methods demonstrated promising potential in reducing prediction errors and uncertainties

The book by Rafail Khasminskii[15] is an essential reference for graduate students in mathematics and its applications. It is also beneficial for researchers looking to explore this field or study mechanical systems affected by random disturbances.

A. Lahrouz ey al [16] aimed to incorporate random noise into the deterministic smoking model in their research. The developed stochastic model demonstrated non-negative solutions. They also noted that the stochastic stability of the equilibrium state in smoking depends on the intensity of the noise and the parameters involved in the model system.

A simplified stochastic model of the predator-prey system is presented in the work of Debjit Pal et al[17].

In his study, Cheng-Shang Chang [18] presented two types of stability problems: the first relates to queueing networks that ensure limited waiting times and bounded delays for customers, while the second concerns queueing networks where the distribution of queue lengths does not follow a normal distribution.

This research provides both Otso Ovaskainen and Baruch Meerson [19] with effective tools to determine extinction times and describe the pathways leading to extinction, offering deep insights into extinction processes and showing significant potential for other applications in theoretical biology

In their article, Lloyd Demetrius et al [20] explored the relationship between population stability, defined by the rate of decay of fluctuations due to demographic randomness, and the variation in birth and death rates across different age groups.

V. G. Kulkarni [21] provides a comprehensive examination of how stochastic modeling techniques can be applied to various aspects of health care. This book discusses the role of randomness and uncertainty in health care systems and presents models that help analyze patient flow, resource allocation, and service delivery. By integrating mathematical and statistical methods, Kulkarni's work offers valuable insights for improving health care efficiency and effectiveness. The application of these models can lead to better decision-making and policy formulation in health care settings.

In their paper, Păuna, B. et al.[22] aimed to develop a comprehensive index for the stability of the Romanian financial system, with the goal of enhancing the analyses used by authorities to assess this stability

In this research, J. A. ADEBISI [23] developed a stochastic model for several selected stocks on the Nigerian Stock Exchange (NSE). The model included four different stocks with their market prices and made short-term predictions for stock prices. By analyzing stock prices over a period of forty months, investors can identify which stock is likely to yield the best returns

In their article, Mohammed El Khomssi et al [24] presented an innovative approach to predicting gold prices over an extended period, utilizing an advanced stochastic process based on traditional models

In the work of Andrew Omame et al [25], a new stochastic model for two strains of SARS-CoV-2 was developed. Conditions for the existence and uniqueness of the global solution to the model were derived, along with numerical simulations to validate the theoretical results.

The article by Shah Hussain et al [26] presents a stochastic version of the MERS-CoV epidemic model, initially formulating the mathematical dynamics of the disease while incorporating unexpected factors. The research by Yongsheng Ding [27] presents a stochastic model for the transmission of HIV/AIDS

In conclusion, previous works on stability analysis of stochastic models provide a rich foundation for understanding their behavior and applications in informatics. The diverse range of methodologies, applications, and challenges discussed in this chapter highlights the importance of stability analysis in ensuring the reliability and robustness of computational models.

2.3 Markov Chains

Markov chains are a fundamental concept in probability theory and have extensive applications in various fields such as physics, economics, biology, and computer science. Named after the Russian mathematician Andrey Markov, a Markov chain is a stochastic process that undergoes transitions from one state to another in a state space. The defining property of a Markov chain is that the future state depends only on the present state and not on the sequence of states that preceded it. We will explores in This chapter the basics of Markov chains, their mathematical formulation, types, and applications.

2.3.1 Definition and Basic Properties

A Markov chain is defined by a set of states and transition probabilities between these states. Let $S = \{s_1, s_2, \ldots, s_n\}$ be a finite set of states. The process moves through these states at discrete time steps $t = 0, 1, 2, \ldots$. The probability of moving from state s_i to state s_j in one time step is denoted by P_{ij}, forming a transition matrix P:

$$P = \begin{pmatrix} P_1^1 & P_1^2 & \cdots & P_1^n \\ P_2^1 & P_2^2 & \cdots & P_2^n \\ \vdots & \vdots & \ddots & \vdots \\ P_n^1 & P_n^2 & \cdots & P_n^n \end{pmatrix}$$

The transition probabilities satisfy the following conditions:

1. $P_{ij} \geq 0$ for all i, j.
2. $\sum_{j=1}^{n} P_{ij} = 1$ for all i.

2.3.2 Types of Markov Chains

Markov chains can be classified based on various properties:

2.3.2.1 Discrete-time and Continuous-time Markov Chains

- **Discrete-time Markov Chains (DTMC):** Transitions occur at fixed discrete time intervals.
- **Continuous-time Markov Chains (CTMC):**Continuous-time Markov processes extend the concept of Markov chains to continuous time. These models are used when transitions between states occur at random times, making them suitable for modeling systems with continuous dynamics. Continuous-time Markov processes find applications in areas like population dynamics, chemical reaction kinetics, and reliability analysis [28].

2.3.2.2 Homogeneous and Non-homogeneous Markov Chains

- **Homogeneous Markov Chains:** Transition probabilities P_{ij} are independent of time.
- **Non-homogeneous Markov Chains:** Transition probabilities $P_{ij}(t)$ can vary with time.

2.3.2.3 Absorbing Markov Chains

A state s_i is called absorbing if $P_{ii} = 1$ and $P_{ij} = 0$ for all $j \neq i$. A Markov chain is called an absorbing Markov chain if it contains at least one absorbing state, and it is possible to go from any state to an absorbing state.

2.3.3 Transition Matrix and State Distribution

The state of the Markov chain at time t can be represented by a probability distribution vector $\pi^{(t)}$, where $\pi_i^{(t)}$ is the probability of being in state s_i at time t. The evolution of the state distribution over time is given by:

$$\pi^{(t+1)} = \pi^{(t)} \times P$$

For a long-term behavior analysis, we are often interested in the stationary distribution π, which satisfies:

$$\pi = \pi \times P$$

and

$$\sum_{i=1}^{n} \pi_i = 1$$

2.3.4 Classification of States

States in a Markov chain can be classified based on their properties:

2.3.4.1 Recurrent and Transient States

- **Recurrent (Persistent) State:** A state s_i is recurrent if, starting from s_i, there is a non-zero probability of returning to s_i eventually.
- **Transient State:** A state s_i is transient if, starting from s_i, the probability of never returning to s_i is non-zero.

2.3.4.2 Periodic and Aperiodic States

A state s_i has period d if any return to s_i must occur in multiples of d steps. If $d = 1$, the state is aperiodic.

2.3.4.3 Ergodic States

A state s_i is ergodic if it is aperiodic and positive recurrent (i.e., it is recurrent and the expected return time is finite). A Markov chain is ergodic if all states are ergodic.

2.3.5 Applications of Markov Chains

Markov chains have a wide range of applications across different fields:

1. **Physics and Chemistry:**In physics and chemistry, Markov chains are used to model the behavior of particles and molecules in systems such as gases, liquids, and solids.
2. **Economics and Finance:**In economics and finance, Markov chains are applied to model market behavior, stock prices, and decision-making processes.
3. **Biology and Medicine:**Markov chains are used in biology to model population dynamics, genetic sequences, and the spread of diseases.
4. **Computer Science:**In computer science, Markov chains are fundamental in algorithms for search engines (e.g., Google's Page Rank algorithm), machine learning, and natural language processing.

2.3.6 recent work on Markov chain

There are many studies in which Markov chains have been used primarily, including the following:

- Xiafei Ding et al[29] talk about Markov chain models in COVID-19 forecasting: latest developments and future prospects.

- A systematic review of hidden Markov models and their applications was conducted by Musa et al[30].

- Schweitzer[31] studied the assembly and disassembly of large Markov chains.

- In their paper, Juan et al[32]. present a prediction methodology to deal with complex deterioration processes for which condition monitoring data constitute the only available source of physical information. They test it using a case study on fatigue crack propagation in steel structures.

- Parametric Markov chains were introduced in the paper by Christel et al[33]. as a model for families of stochastic systems that are based on the same graph structure, but differ in their concrete transition probabilities.

- Joakim et al[34] used this study of a Markov Chain Mixture Distribution Model (MCM) to forecast short-term load of household electricity consumption. The model is used to forecast household electricity consumption data with half-hour accuracy ahead.

- Daryl et al[35] solved the constituency problem, which is the problem of dividing a set of geographic units into a fixed number of subsets called constituencies, subject to a list of rules and priorities.

Markov chains provide a powerful mathematical framework for modeling stochastic processes with the Markov property. Their applications span numerous disciplines, reflecting their versatility and utility in both theoretical and practical contexts. Understanding Markov chains involves grasping the concepts of state transitions, classification of states, and the behavior of the chain over time. As we continue to encounter complex systems and datasets, Markov chains will remain an essential tool in the arsenal of probabilistic and statistical methods. Markov chains are a fundamental type of stochastic model where the probability of transitioning from one state to another depends only on the current state, following the Markov property. They are extensively used in modeling systems with discrete states and probabilistic transitions, such as queueing systems, genetic algorithms, and weather prediction models[36, 37]

2.3.7 Interval Markov Chains (IMCs)

2.3.7.1 Definition and Basic Properties

An Interval Markov Chain is defined by a set of states and transition probability intervals between these states. Let $S = \{s_1, s_2, \ldots, s_n\}$ be a finite set of states. The transition probability from state s_i to state s_j is given by an interval $[P_{ij}^{\min}, P_{ij}^{\max}]$, where $P_{ij}^{\min}$ and $P_{ij}^{\max}$ are the lower and upper bounds of the transition probability, respectively[38]. The transition interval matrix P is represented as:

$$P = \begin{pmatrix} \left[P_{1,1}^{(min)}, P_{1,1}^{(max)}\right] & \left[P_{1,2}^{(min)}, P_{1,2}^{(max)}\right] & \cdots & \left[P_{1,n}^{(min)}, P_{1,n}^{(max)}\right] \\ \left[P_{2,1}^{(min)}, P_{2,1}^{(max)}\right] & \left[P_{2,2}^{(min)}, P_{2,2}^{(max)}\right] & \cdots & \left[P_{2,n}^{(min)}, P_{2,n}^{(max)}\right] \\ \vdots & \vdots & \ddots & \vdots \\ \left[P_{n,1}^{(min)}, P_{n,1}^{(max)}\right] & \left[P_{n,2}^{(min)}, P_{n,2}^{(max)}\right] & \cdots & \left[P_{n,n}^{(min)}, P_{n,n}^{(max)}\right] \end{pmatrix}$$

The intervals must satisfy the following conditions for each state s_i:

1. $0 \leq P_{ij}^{(min)} \leq P_{ij}^{(max)} \leq 1$ for all j.
2. $\sum_{j=1}^{n} P_{ij}^{(min)} \leq 1$ and $\sum_{j=1}^{n} P_{ij}^{(max)} \geq 1$.

Markov Interval Chain (IMC) provides a valuable framework for modeling stochastic processes with uncertain transition probabilities. By representing transition probabilities as intervals, IMC enhances the realism and applicability of traditional Markov chains, making it a powerful tool for decision-making, reliability analysis, and fault-tolerant system design.Markov Interval Chain (IMC) provides a valuable framework for modeling stochastic processes with uncertain transition probabilities. By representing transition probabilities as intervals, IMC enhances the realism and applicability of traditional Markov chains, making it a powerful tool for decision-making, reliability analysis, and fault-tolerant system design. The Interval Markov Chain (IMC) is a stochastic model that extends traditional Markov chains to incorporate uncertainty in transition probabilities. While traditional Markov chains assume fixed transition probabilities between states, IMC allows representing these probabilities as intervals, reflecting the uncertainty inherent in real-world systems [39].

2.3.7.2 Transition Diagram of Interval Markov Chain

The transition diagram for an IMC visually represents the states and the intervals of transition probabilities between them. Each node in the diagram corresponds to a state s_i, and each directed edge between nodes s_i and s_j is labeled with the interval $\left(\left[P_{i,j}^{(min)}, P_{i,j}^{(max)}\right]\right)$. This diagram is useful for understanding the range of possible transitions in the system and for analyzing the behavior of the Markov chain under uncertainty. The figure below illustrates a simple IMC with three states s_1, s_2, and s_3. The intervals on the edges indicate the possible range of transition probabilities between these states.

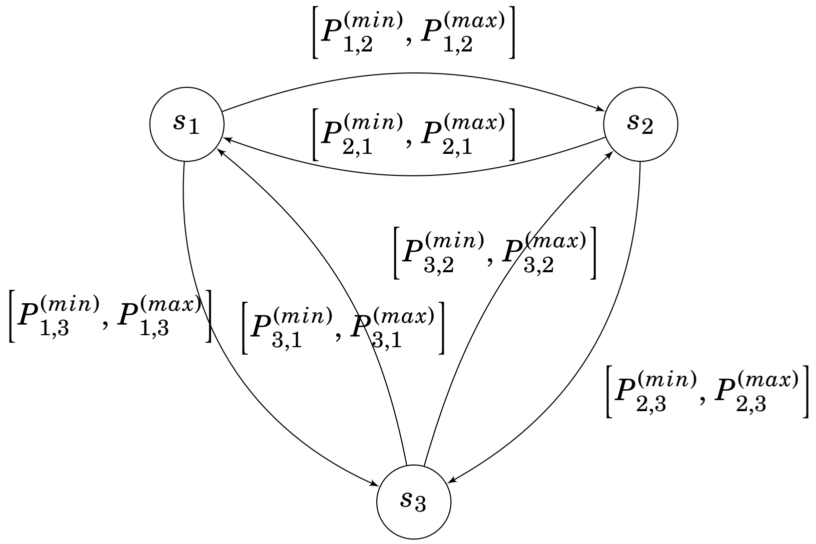

Figure 2.1: Transition Diagram of Interval Markov Chain

2.3.7.3 Characteristics of IMC

1. **Uncertainty Representation:**IMC captures the uncertainty in transition probabilities, allowing for more realistic modeling of stochastic processes
2. **Flexibility:**By representing transition probabilities as intervals, IMC provides flexibility in modeling systems with imprecisely known or varying transition probabilities

2.3.8 Stability Analysis in IMC

Stability analysis in IMC is crucial for understanding the long-term behavior of the system. Stability conditions for IMC depend on the intervals of transition probabilities and the structure of the underlying Markov chain. Various methods, such as eigenvalue analysis, Lyapunov functions, and interval arithmetic, are employed for stability analysis

.

2.3.8.1 Applications of IMC

1. **Decision Making Under Uncertainty:**IMC provides a framework for decision-making in uncertain environments, where the consequences of decisions may vary due to uncertain transition probabilities.
2. **Reliability Analysis:**IMC is used to analyze the reliability of systems where component failure probabilities are uncertain or varying.
3. **Fault-Tolerant Systems:** IMC aids in designing fault-tolerant systems by modeling uncertain behaviors of components .

2.3.8.2 Case Studies and Examples

1. **Robust System Design:**In engineering, IMCs can model systems with components that have probabilistic failure rates within known bounds, helping to design robust and reliable systems.
2. **Modeling Biological Systems:**In biology, IMCs can represent uncertain transitions between states in biological processes, such as gene expression levels or the spread of diseases.
3. **Finance and Economics:**In finance, IMCs can model market behavior under uncertainty, providing a range of possible outcomes based on varying transition probabilities.
4. **Computer Science:**In computer science, IMCs can be used in verification and validation of systems with probabilistic behaviors, such as randomized algorithms and protocols.

They have applications in reliability analysis, fault-tolerant systems, and decision-making under uncertainty. Several case studies and examples demonstrate the applicability of IMC in various domains.

2.4 Page Rank Algorithm

In the digital age, the explosion of information available on the internet has necessitated sophisticated methods for retrieving relevant data. One of the most groundbreaking contributions to this field is the Page Rank algorithm, developed by Larry Page and

Sergey Brin in 1996 while they were PhD students at Stanford University. This algorithm underpins the search engine Google, fundamentally transforming the way we navigate and utilize web content.

2.4.1 The Conceptual Foundation of Page Rank

At its core, Page Rank is a method for ranking web pages in search engine results. The primary insight behind Page Rank is that the importance of a web page can be inferred from the links that point to it. In essence, a page is considered important if it is linked to by other important pages. This recursive idea reflects the notion that high-quality pages are more likely to be linked by other pages. Page Rank operates on the principle of democratic voting, where each link from one page to another represents a vote of confidence. However, not all votes are equal; votes from highly ranked pages carry more weight than those from less significant pages. This iterative process of evaluating and re-evaluating the importance of pages continues until the rankings stabilize.

2.4.2 Mathematical Formulation of Page Rank

The Page Rank of a page p_i is defined by a formula that considers both the quantity and quality of incoming links. The basic Page Rank algorithm as described by Brin et al [40] can be expressed as follows::

$$PR(p_i) = \frac{1-d}{N} + d\left(\sum_{P_j \in M(p_i)} \frac{PR(P_j)}{L(p_j)}\right) \tag{2.1}$$

Where:

- $PR(p_i)$:is the Page Rank of page p_i.
- d: is the damping factor, typically set to 0.85. This factor accounts for the probability that a user will continue clicking on links rather than starting a new search.
- N: is the total number of pages in the web graph.
- $M(p_i)$: is the set of pages that link to p_i.
- $L(p_j)$: is the number of outbound links on page p_j.

The equation consists of two main components:

1. **Teleportation Term:** $\frac{1-d}{N}$

  - This term ensures that the algorithm can handle disconnected web graphs by introducing a probability that a random surfer will jump to any page in the web.

2. **Linking Term:** $d\left(\sum_{P_j \in M(P_i)} \frac{PR(P_j)}{L(P_j)}\right)$
   - This term captures the probability that the random surfer follows links, distributing the rank of the current page among the pages it links to.

2.4.3 Iterative Computation of Page Rank

The computation of Page Rank is an iterative process. Initially, each page is assigned an equal rank. The algorithm then repeatedly updates the ranks using the formula until the values converge to a steady state. This convergence is typically fast, often requiring only a few iterations.

2.4.4 The Damping Factor

The damping factor d plays a crucial role in the Page Rank algorithm. It models the behavior of a "random surfer" who, with probability $1-d$, jumps to a random page rather than following links. This aspect helps in managing the issue of rank sinks;pages or groups of pages that link exclusively to each other, trapping rank. The typical value of $d = 0.85$ balances the likelihood of random jumps and link-following behavior. This value was determined experimentally to provide good performance across a wide range of web structures.

2.4.5 Practical Implications and Extensions

The Page Rank algorithm has profound implications for search engine optimization (SEO). Understanding how Page Rank works can help webmasters improve their page rankings by focusing on acquiring high-quality backlinks from reputable sources.
However, this also led to manipulative practices like link farms and paid links, prompting Google to continuously refine its algorithm to combat such tactics. Over time, Page Rank has been extended and modified to address various challenges. For example:

- **Personalized Page Rank:** Tailors search results based on user preferences and browsing history.

- **Topic-Sensitive Page Rank:** Weighs links differently based on the topic or category of pages.
- **TrustRank:** Combats web spam by using a subset of trusted pages as seeds.

The Page Rank algorithm represents a pioneering approach to information retrieval, emphasizing the interconnected nature of web content. By leveraging the structure of links, Page Rank effectively ranks web pages, making it easier for users to find relevant information. While the algorithm has evolved and been augmented by numerous other factors in modern search engines, its core principles remain a testament to the power of leveraging collective intelligence through link analysis. Understanding Page Rank is not only essential for grasping the mechanics of search engines but also offers insights into the broader field of network theory and the dynamics of interconnected systems. As the digital landscape continues to grow, the principles underlying Page Rank will undoubtedly continue to influence the evolution of information retrieval and management technologies. The Page Rank algorithm is a stochastic model used to rank web pages in search engine results. It models the web as a directed graph, where the importance of a page is determined by the number and quality of inbound links. Page Rank has revolutionized information retrieval and plays a crucial role in web ranking algorithms [41, 42]

2.4.6 Limitations

While Page Rank is effective, it has some limitations. It may not accurately represent the importance of pages in dynamically changing web environments. Additionally, it can be manipulated by webmasters through techniques such as link farming.

2.5 Connection between Markov Interval Chain (IMC) and Page Rank Algorithm

The connection between the Markov Interval Chain (IMC) and the Page Rank Algorithm presents an intriguing intersection between stochastic modeling and web search.
As we mentioned earlier:
IMC, an extension of traditional Markov chains, is a powerful tool for modeling stochastic processes under uncertainty. It represents transition probabilities as intervals, offering a

more realistic approach to modeling complex systems . and Page Rank, initially developed by Larry Page and Sergey Brin, is the cornerstone of Google's search algorithm. It assigns a numerical value to each web page in a network based on the importance of incoming links, effectively ranking web pages in search results [43]. So, the connection between Markov Interval Chain (IMC) and Page Rank Algorithm presents a rich area for exploration, with potential applications in modeling dynamic systems, enhancing web search algorithms, and improving the robustness of stochastic models.

2.6 Probability

The aim of this section is to offer a comprehensive overview of probability principles, random variables, and distributions. Probability is linked with the execution of random experiments or trials and evaluating the ensuing outcomes.

Definition 2.6.1 (Outcome)**.** An outcome is any possible observation of a random experiment.

For example, the random experiment might be flipping a coin and the outcome would be heads or tails depending on whether the coin lands face up or down. The possible outcomes of a random experiment could be discrete or continuous. An example of a random experiment with discrete outcomes is rolling a die since the possible outcomes would be the numbers 1, 2,..., 6. An example of a random experiment with continuous outcomes is spinning a pointer and measuring its angle relative to some reference direction. The angle we measure could be anything between 0°and 360°.

Definition 2.6.2 (Sample space)**.** The collection of all possible, mutually exclusive outcomes is called the sample space S.

For the random experiment of rolling a die, the sample space will be the set S={1,2,3,4,5,6} This sample space is discrete since the possible outcomes were discrete. For the example of spinning a pointer, the sample space is specified by the equation S= $\{\theta \mid 0^\circ \leq \theta < 360^\circ\}$

Definition 2.6.3 (Event)**.** An event is a set of outcomes sharing a common characteristic

Definition 2.6.4 (Event space)**.** The event space is the collection of all possible,mutually exclusive events of a random experiment.

In that sense, the event space is the partitioning of S into subsets that do not share their elements. For the case of die rolling experiment, we might be interested in the event E that the outcome is an even number. Thus we can specify E: the number is even or $\overline{E}$:the number is odd. In that case, there are six outcomes and two events. The event space will be composed of two events S={E,$\overline{E}$}

Definition 2.6.5 (Complementary event)**.** Given an event A,the complementary event $\overline{A}$ is a set of all outcomes that do not belong to the set A

2.6.1 Applying Set Theory to Probability

sets play a crucial role in defining various aspects of random experiments, including the sample space, outcomes, and events. Table 2.1 illustrates the relationship between set theory terminology and probability definitions. Let's consider two events, denoted as A and B, in a random experiment. These events can be visually represented using the Venn diagram . A or B is referred to as the union of A and B. This union is symbolically represented by the expression A∪ B The event defined as any outcome that belongs to both A and B is called the intersection of A and B and is represented by the expression A∩B Figure 2.2 shows the union operation as the Purple area . Figure 2.3 shows the intersection operation as the red area.

Probability	Set theory
Outcome	Element of a set
Sample space S	Universal set ∪
Event	Set
Impossible event	Null set ∅

Table 2.1: Correspondence between set theory and probability definitions

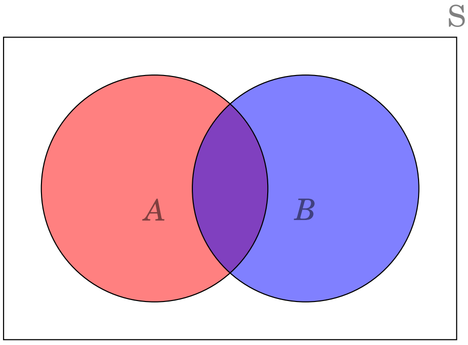

Figure 2.2: Venn diagram for two events A and B in a sample space S. The intersection operation $A \cap B$.

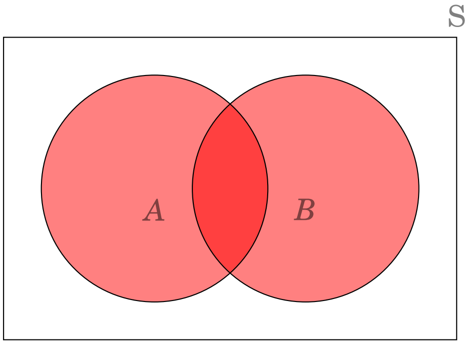

Figure 2.3: Venn diagram for two events A and B in a sample space S. The union operation $A \cup B$.

Figure 2.4 shows that the universal set S is partitioned into two sets A and $\overline{A}$. These two sets are not overlapping in the sense that there is not a single outcome that belongs to both A and $\overline{A}$ simultaneously.

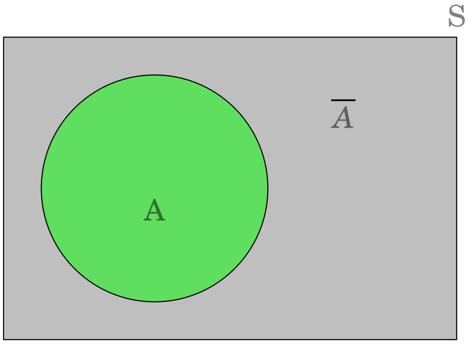

Figure 2.4: Venn diagram for Event A and its complement $\overline{A}$.

Exemple: Let S={a,b,c,d,e,f,g,h,i,j,k,l,m},A={a,c,e,h,j}, B={c,d,e,f,k} We have: A∪B={a,c,d,e,f,j,h,k} A∩B={c,e} $\overline{A}$={b,d,f,g,i,k,l,m}

Definition 2.6.6 (Mutually exclusive events)**.** Events A and B are said to be mutually exclusive or disjoint events if they have no elements in common. From that definition, we can write A∩B=∅

2.6.2 Axioms of Probability

We defined our sample space S as the set of all possible outcomes of an experiment.An impossible outcome defines the empty set or null event ∅ . Based on this we can state four basic axioms for the probability.

- The probability P(A) of an event A is a non negative fraction in the range 0≤ P(A)≤ 1
- The probability of the null event ∅ is zero,P(∅)=0.
- The probability of all possible events S is unity,P(S)=1.
- If A and B are mutually exclusive events (cannot happen at the same time), then the probability that event A or event B occurs is P(A ∪ B)=P(A)+P(B)

Event E and its complement E^c are mutually exclusive. By applying the above axioms of probability, we can write
P(E)+P(E^c)=1.
P(E ∩E^c)=0.
If A and B are two events (they need not be mutually exclusive), then the probability that event A or event B occurring is:
P(A∪B)=P(A)+P(B)-P(A∩B)
The probability that event A occurs given that event B occurred is denoted by P(A|B) and is sometimes referred to as the probability of A conditioned by B.This is given by P(A|B)=$\frac{P(A\cap B)}{P(B)}$
Now, if A and B are two independent events, then we can write P(A|B)=P(A) because the probability of event A taking place will not change whether event B occurs or not. From the above equation we can now write the probability that event A and event B occurs is
P(A∩ B)=P(A)× P(B)
Provided that the two events are independent.
P(A^c)=1-P(A)

2.6.3 Random Variables

A random variable is a variable in a statistical experiment or scenario whose possible values are outcomes of a random phenomenon. In other words, it's a variable that can take on different values as a result of chance. Random variables are used to model uncertainty in various situations, such as in probability theory, statistics, and applied mathematics. There are two main types of random variables: discrete and continuous.

- Discrete random variables: These are random variables that can only take on a countable number of distinct values. Examples include the number of heads obtained when flipping a coin multiple times, the number of customers arriving at a store in a given hour.
- Continuous random variables: These are random variables that can take on any value within a certain range. Examples include the height of a person, the time it takes for a car to travel from one point to another, or the temperature in a given location at a specific time.

Random variables are fundamental in probability theory and are often used to describe the uncertainty associated with real world phenomena, making them essential in fields such as statistics, finance, engineering, and many others.

2.6.3.1 Poisson RV

A Poisson random variable (RV) is a discrete probability distribution that represents the number of events occurring in a fixed interval of time or space, given a constant average rate of occurrence, and where the events happen independently of the time since the last event. It's often used to model rare events like the number of phone calls at a call center within a given time period, the number of accidents at an intersection in a day, or the number of goals scored in a soccer match. The probability mass function (PMF) of a Poisson random variable is given by:

$$P(X=k)=\frac{e^{-\lambda}\lambda^{k}}{k!}$$

Where:

- P(X=k):is the probability of observing k events.
- e:is the base of the natural logarithm (approximately equal to 2.71828).

- λ:is the average rate of occurrence (also called the rate parameter).
- k:is the number of events.
- k!:denotes the factorial of k.

The mean and variance of a Poisson random variable are both equal to λ, i.e.,E(X)=λ and Var(X)=λ

2.6.3.2 Exponential RV

An exponential random variable (RV) is a continuous probability distribution that represents the time between events in a Poisson process, where events occur continuously and independently at a constant average rate. It's commonly used to model the time until the next event occurs, such as the time between arrivals of customers at a service point, the time until a radioactive atom decays, or the lifetime of certain electronic components. The probability density function (PDF) of an exponential random variable is given by:

$$f(x|\lambda) = \lambda e^{-\lambda x}$$

Where:

- f(x | λ):is the probability density function at x given the rate parameter λ
- e:is the base of the natural logarithm (approximately equal to 2.71828).
- λ:is the rate parameter, which is the average number of events per unit time.
- x:is the value of the random variable.

The mean and variance of an exponential random variable are both equal to $\frac{1}{\lambda}$, i.e.,E(X)=$\frac{1}{\lambda}$ and Var(X)=$\frac{1}{\lambda^2}$

2.6.4 Cumulative Distribution Function (CDF)

The Cumulative Distribution Function (CDF) is a concept used in probability theory and statistics. It describes the probability that a random variable takes on a value less than or equal to a given value. Mathematically, for a random variable X, the CDF, denoted as ,is defined as:F(x)=P(X≤x) Where

- F(x) is the cumulative distribution function.

- $P(X \leq x)$ is the probability that the random variable X takes on a value less than or equal to x.

The CDF has several important properties:

- It is a non-decreasing function: As x increases,F(x) either remains constant or increases.
- It ranges between 0 and 1: As x approaches negative infinity, F(x) approaches 0, and as x approaches positive infinity, F(x) approaches 1.
- It is right-continuous: The function has no jumps,meaning that the value of the CDF at each point x is equal to the limit of the function as x approaches that point from the right.

The CDF is a fundamental concept in probability and statistics, and it is often used to describe and analyze random variables and their distributions.

2.6.5 Probability Mass Function(PMF)

The Probability Mass Function (PMF) is a concept used in probability theory and statistics, particularly in the context of discrete random variables. It gives the probability that a discrete random variable is exactly equal to some value. Mathematically, for a discrete random variable X, the PMF is denoted by p(x), where x represents the possible values that X can take on. The PMF p(x) is defined as:$p(x)=P(X=x)$ Where

- p(x) is the probability mass function.
- $P(X=x)$ is the probability that the random variable X takes on the specific value x.

The PMF has the following properties:

- Non-negativity: $p(x) \geq 0$ for all x.
- Summing to 1: The sum of probabilities over all possible values of X equals 1

The PMF provides a complete description of the probability distribution of a discrete random variable, specifying the probabilities associated with each possible outcome. It is commonly used to characterize the behavior of discrete random variables in various applications, such as in counting problems, games, and queueing systems.

2.6.6 Probability Density Function (PDF)

The Probability Density Function (PDF) is a concept used in probability theory and statistics, particularly in the context of continuous random variables. It describes the relative likelihood of a continuous random variable taking on a particular value within a given interval. Mathematically, for a continuous random variable X, the PDF is denoted by f(x), where x represents the possible values that X can take on. The PDF f(x) is defined such that the probability of X falling within a specific interval [a,b] is given by the integral of f(x) over that interval:

$$P(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{d}x$$

Where

- f(x) is the probability density function.
- P(a ≤ X ≤ b) is the probability that the random variable X falls within the interval[a,b].
- dx represents an infinitesimally small interval around x, indicating integration over all possible values of x within the interval.

The PDF has the following properties:

- Non-negativity: f(x) ≥0 for all x.
- Area under the curve: The total area under the PDF curve over the entire range of possible values of X equals 1, i.e.,

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$$

Since the PDF represents a density rather than a probability, the probability of X being equal to any specific value is technically zero. Instead, the PDF is used to calculate probabilities for intervals of values. The PDF is a fundamental concept in continuous probability distributions, such as the normal distribution, exponential distribution, and uniform distribution. It provides essential information about the shape and characteristics of the distribution of a continuous random variable.

2.6.6.1 Expected Value and Variance

The pdf and pmf we obtained above help us find the expected value E[X] of a random variable X. For the continuous case, the expected value is given by

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$$

For the discrete case, the expected value is given by the weighted sum

$$E[X] = \sum_{i} x_i p(x_i)$$

2.7 Queuing Theory

Queuing theory is the study of queues or waiting lines. It applies mathematical principles to understand and predict the behavior of queues, which helps in designing systems to manage them efficiently. Originating in the early 20th century with the work of Agner Krarup Erlang in telephony, queuing theory has become essential in various fields, including informatics.

2.7.1 Fundamentals of Queuing Theory

- Key Concepts and Definitions
  - Arrival Process: Describes how items (customers, data packets) arrive at the queue, typically modeled using a Poisson process.
  - Service Process: Describes how items are processed or serviced, often modeled using exponential distributions.
  - Queue Discipline: The rule determining the order in which items are serviced (e.g., FIFO - First In, First Out).
- Basic Queuing Models
  - M/M/1 Queue: A single server with exponential inter-arrival and service times.
  - M/M/c Queue: Multiple servers with exponential inter-arrival and service times.
  - M/G/1 Queue: A single server with a general service time distribution.

2.7.2 Relationship to Stochastic Models

Queuing theory heavily relies on stochastic processes to model the randomness inherent in arrival and service times. Key stochastic tools used include:

- Poisson Processes: For modeling random arrival times.
- Markov Chains: For analyzing systems where the future state depends only on the current state (memoryless property).

2.7.3 Applications in Informatics

- Computer Networks
  - Packet Switching: Queuing models help in designing efficient packet-switched networks by optimizing router queue management. Example: In a network router, packets arrive randomly and must be forwarded to their destinations. Using an M/M/1 model helps in predicting the average queue length and waiting time, which informs buffer size design and congestion control mechanisms [44]
  - Traffic Engineering: Managing data traffic to avoid congestion and ensure smooth data flow. Example: Internet service providers use queuing models to balance traffic loads across their networks, improving service quality [45]
- Operating Systems
  - Process Scheduling: Operating systems use queuing theory to manage CPU scheduling, ensuring efficient use of processing resources and minimizing process wait times. Example: In a multi-programmed operating system, processes arrive at the CPU queue randomly. Using an M/M/c model (multiple CPUs), the system can predict queue lengths and waiting times to optimize the scheduling algorithm [46]
  - Resource Management: Queuing models help in allocating and managing resources like memory and I/O devices, ensuring that the system remains responsive under varying loads. Example: Disk I/O requests are often modeled using queuing theory to improve disk scheduling algorithms, reducing average wait times and enhancing system throughput [47]
- Distributed Systems

  - Load Balancing: Queuing models assist in distributing workloads across multiple servers or nodes, ensuring no single node becomes a bottleneck. Example: In a cloud computing environment, tasks are distributed across a pool of servers. By modeling the task arrival and service rates using an M/M/c queue, the system can dynamically adjust load balancing to maintain performance [48]
  - Concurrency Control: Ensuring data consistency and system performance in databases and distributed systems often involves queuing theory to manage concurrent access to shared resources. Example: Database systems use queuing models to manage transactions, optimizing locking mechanisms to balance throughput and consistency [49]

- Performance Evaluation
  - Queuing models are used to predict and evaluate the performance of computer systems, from single-server setups to large-scale distributed systems. Example: Performance modeling of web servers to predict response times and optimize server configurations based on expected traffic patterns [50]
  - Simulation: Queuing theory often underpins simulation tools that model and analyze the behavior of computer systems under various scenarios and workloads. Example: Network simulators like NS-3 use queuing theory to model packet queues and simulate the performance of networking protocols [51]

2.7.4 Examples:

2.7.4.1 Web Server Performance Optimization

- Problem: A web server experiences high variability in request rates, leading to performance degradation.

- Solution: Using an M/M/1 queuing model, the server's request handling process is analyzed. By understanding the relationship between arrival rates, service rates, and queue lengths, the server's configuration is optimized to handle peak loads more effectively.

- Outcome: Reduced average response time and improved user experience during high traffic periods.

2.7.4.2 Cloud Resource Allocation

- Problem: A cloud service provider needs to allocate resources dynamically to handle fluctuating workloads.
- Solution: Using an M/M/c queuing model, the provider models the arrival of tasks and the service capacity of its servers. This analysis informs an elastic resource allocation strategy that adjusts the number of active servers based on current demand.
- Outcome: Efficient use of resources, cost savings, and maintained performance levels under varying workloads.

Queuing theory provides a robust mathematical framework for analyzing and optimizing various systems in informatics. By understanding the stochastic nature of these systems, queuing models help improve performance, manage resources efficiently, and ensure the stability and reliability of computing environments.

2.8 Conclusion

The "Background" chapter serves as a crucial foundation for the thesis by systematically presenting the essential concepts and frameworks that underpin the research and analysis in the subsequent chapters. By covering mathematical modeling and its various classifications, the chapter provides a clear understanding of how complex systems can be translated into predictive models. The detailed examination of deterministic and stochastic models, along with the discussion on stability, highlights the dual perspectives from which systems can be studied and understood. The inclusion of the Page Rank algorithm exemplifies the practical application of these mathematical concepts in developing sophisticated tools that drive modern technology. Furthermore, the exploration of queuing theory underscores its significance in optimizing systems that deal with variability and congestion, particularly in the field of informatics. Overall, this chapter not only lays the groundwork for the technical discussions to follow but also emphasizes the importance of these foundational concepts in addressing real-world problems across various scientific and engineering domains.

CHAPTER 3

MATRIX AND VECTOR CALCULATIONS

> "Mathematics is the music of reason."
>
> *James Joseph Sylvester, English mathematician*

3.1 Introduction

The study of matrices and vectors is fundamental in linear algebra and has wide-ranging applications in various fields, particularly in the analysis and stability of stochastic models. This chapter delves into the essential arithmetic and logical operations involving matrices and vectors, which form the backbone of many mathematical and computational methods. By understanding these operations, one can effectively represent and manipulate multidimensional data, which is crucial for modeling complex systems. The chapter also includes practical examples using the Python programming language to demonstrate how these operations can be implemented in real-world scenarios.

3.1.1 Vectors

A vector in $\mathbb{R}^n$ is an ordered list of n real numbers. It can be represented as:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Vectors are used to represent state variables in stochastic models, enabling the analysis of system dynamics over time[52].

3.2 Matrices

A matrix $\mathbf{A}$ of size $m \times n$ is an ordered array of $m \times n$ real numbers. It can be represented as:

$$\mathbf{A} = \begin{pmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^1 & a_m^2 & \cdots & a_m^n \end{pmatrix}$$

Matrices are fundamental in representing linear transformations and systems of linear equations, which are prevalent in the analysis of stochastic models[53].

3.3 Arithmetic Operations

3.3.1 Vector Addition and Subtraction

Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, their addition and subtraction are defined element-wise:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}, \quad \mathbf{u} - \mathbf{v} = \begin{pmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{pmatrix}$$

We can calculate Vector Addition and Subtraction using this code in the Python programming language:

```python
import numpy as np
# Example usage
u = np.array([1, 2, 3])
v = np.array([4, 5, 6])
add_result = u + v
sub_result = u - v
print("Vector Addition:", add_result)
print("Vector Subtraction:", sub_result)
```

Listing 3.1: Python program to calculate vector addition and subtraction

3.3.2 Complexity Analysis of vector addition and subtraction algorithm

3.3.2.1 Vector Addition (`u + v`)

The operation `u + v` performs element-wise addition of two arrays `u` and `v`. Let n be the length of the vectors `u` and `v`. The addition operation iterates over each element of the arrays once, performing a constant-time addition for each pair of elements. Therefore, the time complexity for vector addition is $O(n)$.

3.3.2.2 Vector Subtraction (`u - v`)

The operation `u - v` performs element-wise subtraction of two arrays `u` and `v`. Similar to the addition operation, the subtraction operation iterates over each element of the arrays once, performing a constant-time subtraction for each pair of elements. Therefore, the time complexity for vector subtraction is $O(n)$.

3.3.2.3 Overall Complexity

Both operations, vector addition and vector subtraction, are performed separately and independently. Each operation has a time complexity of $O(n)$. Therefore, the overall complexity for performing both operations is $O(n)$. This complexity is linear, indicating that the time required to perform the operations scales linearly with the size of the input vectors.

3.3.3 Scalar Multiplication

For a scalar $\alpha \in \mathbb{R}$ and a vector $\mathbf{v} \in \mathbb{R}^n$, scalar multiplication is:

$$\alpha \mathbf{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{pmatrix}$$

Scalar multiplication is used to scale vectors, which is essential in various transformation processes in stochastic models[54].

3.3.4 Matrix Addition and Subtraction

Given two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, their addition and subtraction are:

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} a_1^1 + b_1^1 & \cdots & a_1^n + b_1^n \\ \vdots & \ddots & \vdots \\ a_m^1 + b_m^1 & \cdots & a_m^n + b_m^n \end{pmatrix}, \quad \mathbf{A}-\mathbf{B} = \begin{pmatrix} a_1^1 - b_1^1 & \cdots & a_1^n - b_1^n \\ \vdots & \ddots & \vdots \\ a_m^1 - b_m^1 & \cdots & a_m^n - b_m^n \end{pmatrix}$$

We can calculate Matrix Addition and Subtraction using this code in the Python programming language:

```python
import numpy as np
# Example usage
A = np.array([[1, 2], [3, 4]])
B = np.array([[5, 6], [7, 8]])
add_result = A + B
sub_result = A - B
print("Matrix Addition:")
print(add_result)
print("Matrix Subtraction:")
print(sub_result)
```

Listing 3.2: Python program to calculate matrix addition and subtraction

3.3.5 Complexity Analysis of matrix addition and subtraction algorithm

3.3.5.1 Matrix Addition (`A + B`)

The operation `A + B` performs element-wise addition of two matrices `A` and `B`. Let m and n be the dimensions of the matrices `A` and `B` (i.e., each matrix has m rows and n columns). The addition operation iterates over each element of the matrices once, performing a constant-time addition for each pair of elements. Therefore, the time complexity for matrix addition is $O(m \times n)$.

3.3.5.2 Matrix Subtraction (`A - B`)

The operation `A - B` performs element-wise subtraction of two matrices `A` and `B`. Similar to the addition operation, the subtraction operation iterates over each element of the matrices once, performing a constant-time subtraction for each pair of elements. Therefore, the time complexity for matrix subtraction is $O(m \times n)$.

3.3.5.3 Overall Complexity

Both operations, matrix addition and matrix subtraction, are performed separately and independently. Each operation has a time complexity of $O(m \times n)$. Therefore, the overall complexity for performing both operations is $O(m \times n)$. This complexity is linear with respect to the number of elements in the matrices, indicating that the time required to perform the operations scales linearly with the size of the input matrices.

3.3.6 Scalar Multiplication of Matrices

For a scalar $\alpha \in \mathbb{R}$ and a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, scalar multiplication is:

$$\alpha \mathbf{A} = \begin{pmatrix} \alpha a_1^1 & \cdots & \alpha a_1^n \\ \vdots & \ddots & \vdots \\ \alpha a_m^1 & \cdots & \alpha a_m^n \end{pmatrix}$$

3.3.7 Matrix Multiplication

Given two matrices $\mathbf{A} \in \mathbb{R}^{m \times p}$ and $\mathbf{B} \in \mathbb{R}^{p \times n}$, their product $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times n}$ is defined as:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

Matrix multiplication is fundamental in transforming coordinate systems and combining linear equations, which is extensively used in stochastic model analysis[55]. We can calculate Matrix Multiplication using this code in the Python programming language:

```
import numpy as np
# Example usage
A = np.array([[1, 2], [3, 4]])
B = np.array([[5, 6], [7, 8]])
result = np.dot(A, B)
print("Matrix Multiplication:")
print(result)
```

Listing 3.3: Python program to calculate matrix multiplication

3.3.7.1 Complexity Analysis of matrix multiplication algorithm

Matrix Multiplication (`np.dot(A, B)`) The `np.dot` function performs matrix multiplication of two matrices `A` and `B`. Let A be of dimensions $m \times p$ and B be of dimensions $p \times n$. The matrix multiplication operation iterates over each element of the resulting matrix C of dimensions $m \times n$, performing p multiplications and $p-1$ additions for each element. Therefore, the time complexity for matrix multiplication is $O(m \times p \times n)$.

So the time complexity of this algorithm is $O(m \times p \times n)$, where m and n are the dimensions of the resulting matrix and p is the shared dimension of the input matrices `A` and `B`. This complexity indicates that the time required to perform matrix multiplication scales with the product of the dimensions of the matrices involved.

Transpose of a Matrix

The transpose of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is denoted by $\mathbf{A}^T \in \mathbb{R}^{n \times m}$ and is defined as:

$$(\mathbf{A}^T)_{ij} = a_{ji}$$

We can calculate Transpose of a Matrix using this code in the Python programming language:

```
import numpy as np
def matrix_transpose(A):
return np.transpose(A)
```

```
# Example usage
A = np.array([[1, 2], [3, 4], [5, 6]])
result = matrix_transpose(A)
print("Transpose of Matrix:")
print(result)
```

Listing 3.4: Python program to calculate transpose of a matrix

3.3.7.2 Complexity Analysis of transpose of a matrix algorithm

Matrix Transpose (`np.transpose(A)`**)**

The `np.transpose` function performs the transpose operation on the matrix `A`. Let A be of dimensions $m \times n$. The transpose operation creates a new matrix A^T of dimensions $n \times m$ by swapping the rows and columns of the original matrix.

The `np.transpose` operation iterates over each element of the matrix, swapping its row and column indices. Therefore, the time complexity for the matrix transpose operation is $O(m \times n)$.

So the time complexity of this algorithm is $O(m \times n)$, where m and n are the dimensions of the input matrix `A`. This complexity indicates that the time required to perform the transpose operation scales linearly with the number of elements in the matrix.

3.3.8 Determinant and Inverse of a Matrix

For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, the determinant is a scalar value that provides important information about the matrix, including whether it is invertible. The inverse of $\mathbf{A}$, denoted by $\mathbf{A}^{-1}$, is defined such that:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix. The matrix $\mathbf{A}$ is invertible if and only if its determinant is non-zero[56]. We can calculate the Determinant and Inverse of a Matrix using this code in the Python programming language:

```
import numpy as np
# Example usage
A = np.array([[1, 2], [3, 4]])
# Step 1: Calculate Determinant
```

```python
det_A = np.linalg.det(A)
# Step 2: Check if matrix is invertible
if det_A != 0:
    # Step 3: Calculate Inverse
    A_inv = np.linalg.inv(A)
else:
    A_inv = None
print("Determinant:", det_A)
print("Inverse:\n", A_inv)
```

Listing 3.5: Python program to calculate determinant and inverse of a matrix

3.3.8.1 Complexity Analysis of determinant and inverse of a matrix algorithm

Determinant Calculation (`np.linalg.det(A)`) The `np.linalg.det` function calculates the determinant of the matrix `A`. Let `A` be a square matrix of dimensions $n \times n$. The time complexity of determinant calculation for a square matrix using standard algorithms (such as LU decomposition) is $O(n^3)$.

Inverse Calculation (`np.linalg.inv(A)`) The `np.linalg.inv` function calculates the inverse of the matrix `A`. For a square matrix of dimensions $n \times n$, the time complexity of matrix inversion is also $O(n^3)$, typically achieved through methods like Gaussian elimination or LU decomposition.

Overall Complexity The calculations for both the determinant and inverse are performed sequentially. Since both operations have a time complexity of $O(n^3)$, the overall complexity for this process is also $O(n^3)$. So the time complexity of this algorithm is $O(n^3)$, where n is the dimension of the input square matrix `A`. This complexity indicates that the time required to calculate the determinant and inverse of the matrix scales cubically with the size of the matrix.

3.4 Logical Operations

3.4.1 Element-wise Comparison

Given two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, the element-wise comparison is defined as:

$$(\mathbf{A} == \mathbf{B})_{ij} = \begin{cases} 1 & \text{if } a_{ij} = b_{ij} \\ 0 & \text{if } a_{ij} \neq b_{ij} \end{cases}$$

Similar operations can be defined for other comparisons like $<, \leq, >$, and $\geq$.

3.4.2 Logical Operations on Matrices

Logical operations such as AND, OR, and NOT can be performed element-wise on matrices with binary elements (0 and 1). These operations are particularly useful in probabilistic models and decision-making processes in informatics[53].

3.5 Eigenvalues and Eigenvectors

Consider a square matrix A of order $n \times n$. An eigenvector of A is a non-zero vector v in R^n, satisfying the condition $Av = \lambda v$, where λ is a scalar. An eigenvalue of A is a scalar λ for which the equation $Av = \lambda v$ has a nontrivial solution.
If $Av = \lambda v$ for $v \neq 0$, we denote that λ is the eigenvalue associated with v, and v is termed an eigenvector corresponding to λ. Eigenvalues and eigenvectors are only for square matrices If someone hands you a matrix A and a vector v, it is easy to check if v is an eigenvector of A: simply multiply v by A and see if Av is a scalar multiple of v. On the other hand, given just the matrix A, it is not obvious at all how to find the Eigenvectors.
Example (Verifying eigenvectors).
Consider the matrix $A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix}$
and vectors $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
Which are eigenvectors?What are their eigenvalues?
solution. We have $Av = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4v$
As shown in the following figure(The vector is in red):

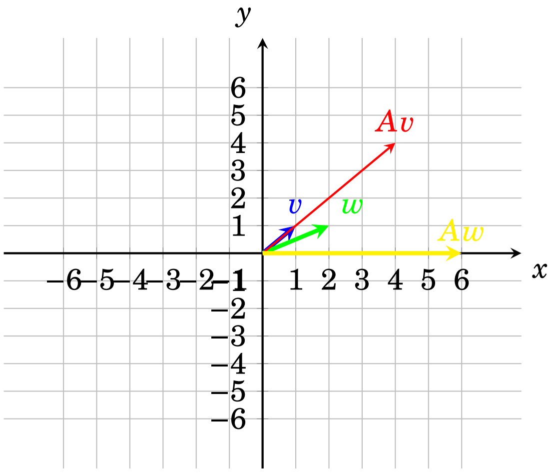

Figure 3.1: Vector Transformation Example

Hence,v is an eigenvector of A, with eigenvalue $\lambda = 4$, on the other hand,

$$Aw = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$$

which is not a scalar multiple of w. Hence w is not an eigenvector of A. To say that $Av = \lambda v$ means that Av and λv are collinear with the origin(lie on the same line). So, an eigenvector of A is a nonzero vector v such that Av and v lie on the same line through the origin. In this case, Av is a scalar multiple of v ; the eigenvalue is the scaling factor.

3.6 Arithmetic Operations Between Matrix and Vector

In linear algebra, arithmetic operations between a matrix and a vector are fundamental. These operations include matrix-vector multiplication and element-wise operations. This section will detail these operations and provide illustrative examples.

3.6.1 Element-wise Operations

Element-wise operations between a matrix and a vector involve performing the operation on each corresponding element of the matrix and the vector. For instance, if we have a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{u} \in \mathbb{R}^{n}$, an element-wise addition can be performed if the vector is broadcasted to match the dimensions of the matrix.

```
import numpy as np
def element_wise_addition(B, u):
```

```python
    return B + u
def element_wise_subtraction(B, u):
    return B - u
# Example usage
B = np.array([[1, 2, 3], [4, 5, 6], [7, 8, 9]])
u = np.array([1, 1, 1])
add_result = element_wise_addition(B, u)
sub_result = element_wise_subtraction(B, u)
print("Element-wise Addition Result:")
print(add_result)
print("Element-wise Subtraction Result:")
print(sub_result)
```

Listing 3.6: Python program to perform element-wise operations

3.6.1.1 Complexity Analysis of the algorithm performing element-wise operations

Element-wise Addition (`B + u`)

The element-wise addition operation adds each element of the matrix `B` with the corresponding element of the vector `u`. Let `B` be a matrix of dimensions $m \times n$ and `u` be a vector of length n. The element-wise addition operation involves adding $m \times n$ pairs of elements.

Therefore, the time complexity for element-wise addition is $O(m \times n)$.

Element-wise Subtraction (`B - u`)

The element-wise subtraction operation subtracts each element of the vector `u` from the corresponding element of the matrix `B`. Similar to the addition operation, the element-wise subtraction operation involves subtracting $m \times n$ pairs of elements.
Therefore, the time complexity for element-wise subtraction is $O(m \times n)$.

Overall Complexity

Both `element_wise_addition` and `element_wise_subtraction` functions are called separately and independently. Each function has a time complexity of $O(m \times n)$. Therefore, the overall complexity for performing both element-wise addition and element-wise subtraction is $O(m \times n)$.
So the time complexity of the algorithm provided in the Python code snippet is $O(m \times n)$, where m and n are the dimensions of the input matrix `B`. This complexity indicates that the time required to perform the element-wise operations scales linearly with the number of elements in the matrix.

Example

Consider the matrix **B** and the vector **u**:

$$\mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

To perform element-wise addition, **u** is broadcasted to match the dimensions of **B**:

$$\mathbf{B} + \mathbf{u} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 5 & 6 \end{pmatrix}$$

Note that this example assumes the vector **u** is applied to each column of the matrix **B**. If we wanted to apply **u** to each row instead, the vector would need to be appropriately shaped or transposed. Arithmetic operations between matrices and vectors are essential in many areas of mathematics, physics, and engineering. Understanding how to perform these operations and their implications is crucial for anyone working with linear algebra.

3.7 Interval Vector

An interval vector is a mathematical construct where each element of the vector is an interval. This concept is particularly useful in numerical analysis and computational applications where uncertainty or range constraints are involved.

Definition 3.7.1. An interval vector is a vector whose components are intervals instead of single numerical values. Each interval represents a range of possible values. Formally, an interval vector **v** can be written as:

$$\mathbf{v} = \begin{pmatrix} \left[a_1, b_1\right] \\ \left[a_2, b_2\right] \\ \vdots \\ \left[a_n, b_n\right] \end{pmatrix}$$

where $\left[a_i, b_i\right]$ denotes the interval for the i-th component of the vector, with $a_i \leq b_i$.

3.7.1 Arithmetic Operations

Arithmetic operations involving interval vectors follow specific rules to account for the range of values in each interval.

3.7.1.1 Addition

The addition of two interval vectors **v** and **w**, where

$$\mathbf{v} = \begin{pmatrix} \left[a_1, b_1\right] \\ \left[a_2, b_2\right] \\ \vdots \\ \left[a_n, b_n\right] \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} \left[c_1, d_1\right] \\ \left[c_2, d_2\right] \\ \vdots \\ \left[c_n, d_n\right] \end{pmatrix}$$

is defined as:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} \left[a_1 + c_1, b_1 + d_1\right] \\ \left[a_2 + c_2, b_2 + d_2\right] \\ \vdots \\ \left[a_n + c_n, b_n + d_n\right] \end{pmatrix}$$

3.7.1.2 Subtraction:

The subtraction of two interval vectors **v** and **w** is defined as:

$$\mathbf{v} - \mathbf{w} = \begin{pmatrix} [a_1 - d_1, b_1 - c_1\right] \\ [a_2 - d_2, b_2 - c_2\right] \\ \vdots \\ [a_n - d_n, b_n - c_n\right] \end{pmatrix}$$

3.7.1.3 Scalar Multiplication

The scalar multiplication of an interval vector **v** by a scalar k is defined as:

$$k \cdot \mathbf{v} = \begin{pmatrix} \left[\min\left(k \cdot a_1, k \cdot b_1\right), \max\left(k \cdot a_1, k \cdot b_1\right)\right] \\ \left[\min\left(k \cdot a_2, k \cdot b_2\right), \max\left(k \cdot a_2, k \cdot b_2\right)\right] \\ \vdots \\ \left[\min\left(k \cdot a_n, k \cdot b_n\right), \max\left(k \cdot a_n, k \cdot b_n\right)\right] \end{pmatrix}$$

3.7.2 Applications

Interval vectors are widely used in fields that require robust computations under uncertainty, such as numerical analysis, optimization, and control theory. They provide a means to model and manage the uncertainty and variability in system parameters and data.

Example

Consider an interval vector representing measurement uncertainties:

$$\mathbf{v} = \begin{pmatrix} [1.0, 1.1] \\ [2.0, 2.2] \\ [3.0, 3.3] \end{pmatrix}$$

If we add another interval vector representing additional uncertainties:

$$\mathbf{w} = \begin{pmatrix} [0.1, 0.2] \\ [0.2, 0.3] \\ [0.3, 0.4] \end{pmatrix}$$

The resulting interval vector is:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} [1.0 + 0.1, 1.1 + 0.2] \\ [2.0 + 0.2, 2.2 + 0.3] \\ [3.0 + 0.3, 3.3 + 0.4] \end{pmatrix} = \begin{pmatrix} [1.1, 1.3] \\ [2.2, 2.5] \\ [3.3, 3.7] \end{pmatrix}$$

Interval vectors extend the concept of traditional vectors to encompass ranges of values, making them a powerful tool in various computational fields. By accounting for uncertainty and variability, interval vectors enable more robust and reliable analysis and decision-making.

3.7.3 Arithmetic Operations on Interval Matrices

An interval matrix is a matrix in which each element is an interval rather than a single number. This concept extends the idea of interval arithmetic to matrices, allowing for the representation of uncertainty and variability in matrix elements. Arithmetic operations on interval matrices are defined in such a way that the result of the operation captures all possible outcomes given the ranges of the input intervals.

Definition 3.7.2. An interval matrix **A** can be represented as:

$$\mathbf{A}=\begin{bmatrix} \left[a_1^1,b_1^1\right] & \left[a_1^2,b_1^2\right] & \cdots & \left[a_1^n,b_1^n\right] \\ \left[a_2^1,b_2^1\right] & \left[a_2^2,b_2^2\right] & \cdots & \left[a_2^n,b_2^n\right] \\ \vdots & \vdots & \ddots & \vdots \\ \left[a_m^1,b_m^1\right] & \left[a_m^2,b_m^2\right] & \cdots & \left[a_m^n,b_m^n\right] \end{bmatrix}$$

where each $\left[a_i^j,b_i^j\right]$ is an interval representing all possible values between a_i^j and b_i^j.

3.7.3.1 Addition

The sum of two interval matrices **A** and **B**(has elements $\left[c_i^j,d_i^j\right]$) of the same dimensions is defined element-wise. For $\mathbf{A}+\mathbf{B}$, the resulting matrix **C** has elements $\left[e_i^j,f_i^j\right]$ where:

$$\left[e_i^j,f_i^j\right]=\left[a_i^j+c_i^j,b_i^j+d_i^j\right]$$

Thus, the addition of two intervals produces a new interval whose lower bound is the sum of the lower bounds of the original intervals, and whose upper bound is the sum of the upper bounds of the original intervals.

3.7.3.2 Subtraction

The difference between two interval matrices **A** and **B** of the same dimensions is also defined element-wise. For $\mathbf{A}-\mathbf{B}$, the resulting matrix **C** has elements $\left[e_i^j,f_i^j\right]$ where:

$$\left[e_i^j,f_i^j\right]=\left[a_i^j-d_i^j,b_i^j-c_i^j\right]$$

Here, the subtraction of two intervals results in a new interval where the lower bound is the difference between the lower bound of the first interval and the upper bound of the second interval, and the upper bound is the difference between the upper bound of the first interval and the lower bound of the second interval.

3.7.3.3 Multiplication

The product of two interval matrices **A** and **B** is more complex due to the nature of matrix multiplication. For interval matrices, the element (i,j) of the product matrix **C** is given by:

$$\left[e_i^j, f_i^j\right] = \sum_{k=1}^{n} \left[a_i^k \cdot c_k^j, b_i^k \cdot d_k^j\right]$$

where each term in the summation is the interval product, defined as the interval that contains all possible products of elements from $\left[a_i^k, b_i^k\right]$ and $\left[c_k^j, d_k^j\right]$. The interval product $\left[p,q\right]$ for intervals $\left[a,b\right]$ and $\left[c,d\right]$ is calculated as:

$$\left[p,q\right] = \left[\min(ac,ad,bc,bd), \max(ac,ad,bc,bd)\right]$$

This ensures that the resulting interval contains all possible products of the values within the given intervals.

3.7.3.4 Scalar Multiplication

Multiplying an interval matrix **A** by a scalar k is straightforward. Each interval element $\left[a_i^j, b_i^j\right]$ is scaled by k:

$$\left[e_i^j, f_i^j\right] = \begin{cases} \left[k \cdot a_i^j, k \cdot b_i^j\right] & \text{if } k \geq 0 \\ \left[k \cdot b_i^j, k \cdot a_i^j\right] & \text{if } k < 0 \end{cases}$$

This operation scales the bounds of each interval, preserving the direction of the interval's range.

3.7.3.5 Applications

Interval matrices are particularly useful in fields where uncertainty and variability must be accounted for, such as robust control systems, optimization problems, and economic modeling. By capturing the range of possible values, interval matrices provide a way to analyze and compute outcomes that consider all possible variations in the input data. We used it in the Page Rank algorithm as well as in solving a decision problem.

3.7.3.6 Some symbols and their meanings

we have:

P:is the transition matrix

P_c:is the Centre of the transition matrix

Δ = Spread of the transition matrix

$P^I = \left[\underline{P}, \overline{P}\right]$

$P_c = \frac{1}{2} \times \left[\underline{P} + \overline{P}\right]$

$\Delta = \frac{1}{2} \times \left[\overline{P} - \underline{P}\right]$

Also, any $P^I = \left[P_c - \Delta, P_c + \Delta\right]$

exemple:

$$P^I = \begin{bmatrix} [0.1, 0.2] & [0.8, 0.9] \\ [0.3, 0.6] & [0.4, 0.7] \end{bmatrix} \tag{3.1}$$

Here:

$$\underline{P} = \begin{bmatrix} 0.1 & 0.8 \\ 0.3, & 0.4 \end{bmatrix} \tag{3.2}$$

$$\overline{P} = \begin{bmatrix} 0.2 & 0.9 \\ 0.6, & 0.7 \end{bmatrix} \tag{3.3}$$

Centre: $P_c = \frac{1}{2} \times \left[\underline{P} + \overline{P}\right] = \frac{1}{2} \times \left[\begin{bmatrix} 0.1 & 0.8 \\ 0.3, & 0.4 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.9 \\ 0.6, & 0.7 \end{bmatrix}\right] = \begin{bmatrix} 0.15 & 0.85 \\ 0.45, & 0.65 \end{bmatrix}$

Spread: $\Delta = \frac{1}{2} \times \left[\overline{P} - \underline{P}\right] = \frac{1}{2} \times \left[\begin{bmatrix} 0.2 & 0.9 \\ 0.6, & 0.7 \end{bmatrix} - \begin{bmatrix} 0.1 & 0.8 \\ 0.3, & 0.4 \end{bmatrix}\right] = \begin{bmatrix} 0.05 & 0.05 \\ 0.15, & 0.15 \end{bmatrix}$

Also, any $P^I = \left[P_c - \Delta, P_c + \Delta\right]$, it can be easily shown.

3.7.4 Determinant of an Interval Matrix

$$det\, P^I = \left\{ detP \mid P \in P^I \right\} = \left[min\, \forall\, P \in P^I\, detP\,,\, max\, \forall\, P \in P^I\, detP \right]$$

Exemple:Take: $P^I = \begin{bmatrix} [0.1, 0.2] & [0.8, 0.9] \\ [0.3, 0.6] & [0.4, 0.7] \end{bmatrix}$

$letS = \left\{ det \begin{bmatrix} a & b \\ c & d \end{bmatrix} : 0.1 \le a \le 0.2, 0.8 \le b \le 0.9, 0.3 \le c \le 0.6, 0.4 \le d \le 0.7 \right\}$

$= \{a \times d - b \times c : 0.1 \le a \le 0.2, 0.8 \le b \le 0.9, 0.3 \le c \le 0.6, 0.4 \le d \le 0.7\}$

So $det\, P^I = \left[minS, maxS\right] = \left[-0.5, -0.1\right]$

3.7.5 Regular Interval Matrix

Definition 3.7.3. A square interval matrix P^I is called regular if each $P \in P^I$ is non singular that is $detP \neq 0$.
Exemple:$S = \{a \times d - b \times c : 0.1 \le a \le 0.2, 0.8 \le b \le 0.9, 0.3 \le c \le 0.6, 0.4 \le d \le 0.7\}\, if\ 0 \notin S$,then each P is non singular.

$$det\, P^I = det \begin{bmatrix} [0.1, 0.2] & [0.8, 0.9] \\ [0.3, 0.6] & [0.4, 0.7] \end{bmatrix} = \left[-0.5, -0.1\right]$$

Since $0 \notin \left[-0.5, -0.1\right]$;$P^I$will be regular

3.7.6 Singularity of an Interval Matrix

Definition 3.7.4. An interval matrix P^I is called singular interval matrix if it contains a singular matrix that is for any $P \in P^I$, det P = 0.
Exemple:

$$\det P^I = \begin{bmatrix} [0.3, 0.5] & [0.5, 0.7] \\ [0.2, 0.6] & [0.4, 0.8] \end{bmatrix} = \left[-0.3, 0.3\right]$$

Since $0 \in \left[-0.3, 0.3\right]$, So P^I contains atleast one singular matrix.

3.7.7 Theoretical Result on Singular Interval Matrix

Theorem 3.7.1: An interval matrix P^I is singular if and only if it satisfies $\left|P^I x\right| \le \Delta |x|$, for some non zero vector x.

3.8 Eigen Values of an Interval Matrix

Definition 3.8.1. Let $P \in M_{n \times m}(R)$.A nonzero vector $v \in R^n$is called an eigen vector of P if $P \times v = \lambda v$ for some scalar λ. The scalar λ is called the eigenvalue of P.

Theorem 3.8.1: $\lambda \in L$ if and only if the interval matrix $\left[(P_c - \lambda I) - \Delta, (P_c - \lambda I) + \Delta\right]$ is singular matrix. Where

$$L = \left\{\lambda \in \mathbb{R} \mid Px = \lambda x \text{ for some } P \in P^I, x \neq 0\right\}$$

3.8.1 Applications in Markov Chains

In Markov Chains, matrices are used to represent transition probabilities between states. The stability and long-term behavior of the chain are analyzed using matrix operations and eigenvalues[55].

3.9 Conclusion

This chapter has provided a comprehensive overview of the basic arithmetic and logical operations involving matrices and vectors, along with their applications in stochastic models. These operations are fundamental in the study of stochastic models and their stability. Understanding these concepts will enable us to delve deeper into the more complex aspects of stochastic model stability in subsequent chapters.

Part II

Contributions

CHAPTER 4

MARKOV INTERVAL CHAIN (IMC) FOR SOLVING A DECISION PROBLEM

> "The quality of decision is like the well-timed swoop of a falcon which enables it to strike and destroy its victim."
>
> *Sun Tzu, Chinese - Philosopher.*

4.1 Introduction

As living standards improve, insurance options diversify, making it essential for the company to draw in more customers to expand its market share. Gaining a clear grasp of market trends and securing a competitive edge are key to achieving this. A company's primary objectives include profit, customer satisfaction, market dominance, and continuity. Among these, continuity is crucial and implies long-term stability and balance. Without stability, continuity is unattainable, and focusing solely on profit might jeopardize the company's survival. A company must align with societal needs to maintain balance and ensure its ongoing presence. The key question is: How can a company forecast its future state and monitor its equilibrium effectively? In this chapter of the graduation thesis, we will talk about Application of Markov prediction method in the decision of insurance company by using Markov chains based on Mr Lu et al [57] . Alongside, we explain our contribution by replacing traditional Markov chains with interval Markov chains and applying our ideas to the same example applied in Mr Lu et al[57]. Then, we present our conclusions and analyses based on the results we obtain. To address this, we propose using an effective model based on Markov interval chains, as outlined by Jonsson and Larsen [58]. This model helps companies make informed decisions and enhance their services, ultimately aiming to attract more customers, ensure continuity, and boost profits. The Markov model serves as a robust tool for evaluating the present state and predicting future outcomes. Since the Russian mathematician Markov introduced the Markov chain, its theory and methods have been extensively utilized across various domains. Markov's predictive approach is particularly prevalent in market economics.

4.2 Study of the insurance market situation using traditional Markov chains:

In a scenario similar to the one discussed by Lu et al. [57], there are three insurance companies A, B, and C serving a total of N customers (with N_A, N_B, and N_C representing the number of customers for each company, respectively). Over time, the customer distribution among these companies fluctuates

There are several possibilities for the movement of customers entering the insurance market or wishing to do so, according to the following cases:

I. **There are no new uninsured customers in one of the insurance companies we are studying entering the insurance market, and no old insured customers in these companies are leaving the insurance market**.

Thus, the total number of customers remains unchanged.With the possibility of customers moving between the three companies under consideration. Therefore, we can derive a one-step transition probability matrix, represented as follows:

$$P = \begin{pmatrix} \frac{N_A^A}{N_A} & \frac{N_A^B}{N_A} & \frac{N_A^C}{N_A} \\ \frac{N_B^A}{N_B} & \frac{N_B^B}{N_B} & \frac{N_B^C}{N_B} \\ \frac{N_C^A}{N_C} & \frac{N_C^B}{N_C} & \frac{N_C^C}{N_C} \end{pmatrix} \tag{4.1}$$

The process of reading the data of this matrix is very simple, for example: $\frac{N_A^A}{N_A}$ represents the retention rate of company A, $\frac{N_A^B}{N_A}$ represents the probability of customers transferring from company A to company B, and $\frac{N_B^A}{N_B}$ represents the probability of company A gaining customers from company B. After one step by the customers and their transition between these three companies, the market share for each company can be expressed as:

$$\pi(0)=\left(\frac{N_A+N_B^A+N_C^A-N_A^B-N_A^C}{N},\frac{N_B+N_A^B+N_C^B-N_B^A-N_B^C}{N},\frac{N_C+N_A^C+N_B^C-N_C^A-N_C^B}{N}\right)$$

Depending on the transition matrix as well as the initial market share for each company, we can predict the market share after (k) periods using the initial probability matrix and the one-step transition probability matrix as follows:

$$\pi(k)=\pi(0)\times P^{(k-1)}$$

After a certain period, the market distribution stabilizes, then p_A, p_B, p_C represent the market shares of each company in order, and can be determined using the following formula:

$$\begin{cases} (p_A,p_B,p_C)=(p_A,p_B,p_C)\times P \\ p_A+p_B+p_C=1 \end{cases} \tag{4.2}$$

II. **There are new customers insuring with these three companies (with no old customers leaving these companies)**

This case can be divided into two parts:

- **New customers are evenly distributed between the three companies**

  If 3n new customers enter the insurance market, each company increases the number of its customers by n. Hence, the one-step transition probability is:

  $$P=\begin{pmatrix} \frac{N_A^A+n}{N_A+n} & \frac{N_A^B}{N_A+n} & \frac{N_A^C}{N_A+n} \\ \frac{N_B^A}{N_B+n} & \frac{N_B^B+n}{N_B+n} & \frac{N_B^C}{N_B+n} \\ \frac{N_C^A}{N_C+n} & \frac{N_C^B}{N_C+n} & \frac{N_C^C+n}{N_C+n} \end{pmatrix} \tag{4.3}$$

  Each company's market share is:

  $$\pi(0)=\left(\frac{N_A+N_B^A+N_C^A-N_A^B-N_A^C+n}{N+3n},\frac{N_B+N_A^B+N_C^B-N_B^A-N_B^C+n}{N+3n},\frac{N_C+N_A^C+N_B^C-N_C^A-N_C^B+n}{N+3n}\right)$$

  The market share after a period of k can be estimated using the equation:

  $$\pi(k)=\pi(0)\times P^{(k-1)}$$

- **New customers are not evenly distributed among the three companies.**

In this case, if we assume that the number of new customers entering the insurance market is n1 for company A, n2 for company B, and n3 for company C, the one-step transition probability matrix can be represented as follows:

$$P = \begin{pmatrix} \frac{N_A^A+n1}{N_A+n1} & \frac{N_A^B}{N_A+n1} & \frac{N_A^C}{N_A+n1} \\ \frac{N_B^A}{N_B+n2} & \frac{N_B^B+n2}{N_B+n2} & \frac{N_B^C}{N_B+n2} \\ \frac{N_C^A}{N_C+n3} & \frac{N_C^B}{N_C+n3} & \frac{N_C^C+n3}{N_C+n3} \end{pmatrix} \tag{4.4}$$

The market share is:

$$\pi(0) = \left(\frac{N_A+N_B^A+N_C^A-N_A^B-N_A^C+n1}{N+n1+n2+n3}, \frac{N_B+N_A^B+N_C^B-N_B^A-N_B^C+n2}{N+n1+n2+n3}, \frac{N_C+N_A^C+N_B^C-N_C^A-N_C^B+n3}{N+n1+n2+n3} \right)$$

The two cases discussed provide a broad overview of the general situation by using three companies as examples. Nevertheless, this approach is not restricted to just three examples and can be applied to additional instances as well.

4.2.1 Application in example

Let's now look at the example Mr Lu et al [57], The one-step transition probability can then be represented by:

$$P = \begin{pmatrix} \frac{192}{240} & \frac{24}{240} & \frac{24}{240} \\ \frac{42}{600} & \frac{540}{600} & \frac{18}{600} \\ \frac{30}{360} & \frac{24}{360} & \frac{306}{360} \end{pmatrix} \tag{4.5}$$

And through the formula:

$$\pi(k) = \pi(0) \times P^{(k-1)}$$

$$\pi(k) = (0.22, 0.49, 0.29) \times \begin{pmatrix} \frac{192}{240} & \frac{24}{240} & \frac{24}{240} \\ \frac{42}{600} & \frac{540}{600} & \frac{18}{600} \\ \frac{30}{360} & \frac{24}{360} & \frac{306}{360} \end{pmatrix}^{(k-1)}$$

We can obtain the values of

$$p_A, p_B, p_C$$

As follows:

$$p_A = \frac{250}{981}, \quad p_B = \frac{793}{1667}, \quad p_C = \frac{260}{981}$$

4.3 Study of the insurance market situation using Interval Markov chains:

4.3.1 Some definitions

Definition 4.3.1. Following de Campos et al [59], the probability interval is defined by:

$$P = \left[\underline{p}_{ij},\ \ \overline{p}_{ij}\right] \tag{4.6}$$

In which $\underline{p}_{ij}$ and $\overline{p}_{ij}$ verifying:

$$\begin{cases} 0 \leq \underline{p}_{ij} \leq \overline{p}_{ij} \leq 1, \forall 1 \leq i,j \leq n; \\ \sum_i \underline{p}_{ij} \leq 1 \leq \sum_i \overline{p}_{ij}, \forall 1 \leq j \leq n \end{cases} \tag{4.7}$$

Definition 4.3.2. The transition probability of any system from an actual state i to another future state j by one step is noted by p_{ij}, in which $\underline{p}_{ij} \leq p_{ij} \leq \overline{p}_{ij}$, then the one-step transition interval probability matrix can be expressed as:

$$P = \begin{pmatrix} \left[\underline{p}_{11}, \overline{p}_{11}\right] & . & . & . & \left[\underline{p}_{1n}, \overline{p}_{1n}\right] \\ . & & . & & . \\ . & & & . & . \\ . & & & . & . \\ \left[\underline{p}_{n1}, \overline{p}_{n1}\right] & . & . & . & \left[\underline{p}_{nn}, \overline{p}_{nn}\right]. \end{pmatrix} \tag{4.8}$$

Definition 4.3.3. Let $\pi(0) = \left(\left[\underline{p}_1^{(0)}, \overline{p}_1^{(0)}\right], \ldots, \left[\underline{p}_n^{(0)}, \overline{p}_n^{(0)}\right]\right)$ to be the initial interval probability distribution. Then, the one-step distribution $\pi(1)$ could be predicted as:

$$\pi(1) = \pi(0) \times P. \tag{4.9}$$

The distribution after k period should be expressed as:

$$\pi(k) = \pi(k-1) \times P. \tag{4.10}$$

From the formula, we can predict future distributions if we know the one-step transition probability matrix P and the initial distribution $\pi(0)$. With the transition matrix, we can calculate the distribution for any future period using this formula.

$$\begin{cases} \pi = \pi \times P, \pi = \left(\left[\underline{p}_1, \overline{p}_1\right], \ldots, \left[\underline{p}_n, \overline{p}_n\right]\right); \\ \sum_i^n \underline{p}_i \leq 1 \leq \sum_i^n \overline{p}_i. \end{cases} \tag{4.11}$$

In the same way that Lu et al.[57] studied the insurance market using traditional Markov chains as mentioned above. Considering A, B and C respectively as three insurance firms on the market, and the number of customers attached to each company is given as follows N_A, N_B and N_C. $N = N_A + N_B + N_C$ denotes the total number of the customers in the three firms. These values are considered likely to increase and decrease by a certain amount, and therefore,The idea we will contribute that instead of considering these values, we consider arbitrary intervals that contain these values as centers. After one transition step, the interval which contains the number of customers of each firm is changed.
We will study the market according to the following cases:

I. **There are no new uninsured customers in one of the insurance companies we are studying entering the insurance market, and no old insured customers in these companies are leaving the insurance market**.
by considering the flow of the customers as a midpoint of a proposed interval, then the one-step interval transition probability matrix can be calculated and presented by

$$P = \begin{pmatrix} \left[\frac{N_A^A - M_A}{N_A}, \frac{N_A^A + M_A}{N_A}\right] & \left[\frac{N_A^B - M_A^B}{N_A}, \frac{N_A^B + M_A^B}{N_A}\right] & \left[\frac{N_A^C - M_A^C}{N_A}, \frac{N_A^C + M_A^C}{N_A}\right] \\ \left[\frac{N_B^A - M_B^A}{N_B}, \frac{N_B^A + M_B^A}{N_B}\right] & \left[\frac{N_B^B - M_B}{N_B}, \frac{N_B^B + M_B}{N_B}\right] & \left[\frac{N_B^C - M_B^C}{N_B}, \frac{N_B^C + M_B^C}{N_B}\right] \\ \left[\frac{N_C^A - M_C^A}{N_C}, \frac{N_C^A + M_C^A}{N_C}\right] & \left[\frac{N_C^B - M_C^B}{N_C}, \frac{N_C^B + M_C^B}{N_C}\right] & \left[\frac{N_C^C - M_C}{N_C}, \frac{N_C^C + M_C}{N_C}\right] \end{pmatrix} \tag{4.12}$$

where M_A, M_B and M_A are arbitrary natural numbers, the lower index indicates the original firm and the upper index indicates the firm to which it is transferred, for example N_A^A indicate the number of customers of A who will stay in A and for N_A^C the number of customers of A which will transform to C. The rows of the interval matrix P a pointing to the interval probability of a firms losing customers to other firms, the columns of the interval matrix can be indicate to the interval probability of a firm getting customers from other firms. For example, $\left[\frac{N_A^B - M_A^B}{N_A}, \frac{N_A^B + M_A^B}{N_A}\right]$ describe the interval probability of customers moving from firm A to B, $\left[\frac{N_B^A - M_B^A}{N_B}, \frac{N_B^A + M_B^A}{N_B}\right]$ represent the interval probability of firm A obtaining customers form firm B. The interval probability distribution of the initial state can be expressed as

$$\pi(0) = \left(\left[\frac{S_A}{N}, \frac{S_A}{N}\right], \left[\frac{S_B}{N}, \frac{S_B}{N}\right], \left[\frac{S_C}{N}, \frac{S_C}{N}\right]\right) \tag{4.13}$$

where S_A, S_B and S_C are expressed respectively by

$$\begin{aligned} S_A &= N_A^A - M_A + N_B^A + M_B^A + N_C^A + M_C^A - (N_A^B + M_A^B + N_A^C + M_A^C), \\ S_B &= N_B^B - M_B + N_A^B + M_A^B + N_C^B + M_C^B - (N_B^A + M_B^A + N_B^C + M_B^C), \\ S_C &= N_C^C - M_C + N_A^C + M_A^C + N_B^C + M_B^C - (N_C^A + M_C^A - N_C^B + M_C^B). \end{aligned} \tag{4.14}$$

After one transition step, each interval share of firm can be given by

$$\pi(1) = \pi(0) \times P \tag{4.15}$$

After k period the market share can be predicted through the initial interval probability matrix and one-step transition probability matrix namely

$$\pi(k) = \pi(k-1) \times P. \tag{4.16}$$

Considering that after a certain period, the distribution of the market reached a stable state, $p_A = [\underline{p}_A, \overline{p}_A]$, $p_B = [\underline{p}_B, \overline{p}_B]$, $p_C = [\underline{p}_C, \overline{p}_C]$ is respectively the market share interval of each firm, and p_A, p_B, p_C can be obtained by the formula

$$\begin{cases} \pi(*) = \pi(*) \times P, \qquad \pi(*) = (p_A, p_B, p_C); \\ \sum_i^n \underline{p}_i \leq 1 \leq \sum_i^n \overline{p}_i \quad . \end{cases} \tag{4.17}$$

II. **New customers are evenly distributed between the three companies**

Considering that $3n$ customers have access to insurance markets A, B and C respectively, and that they are shared by the three companies equally, namely, each company increase n customers, then the one-step transition interval matrix probability is

$$P = \begin{pmatrix} \left[\frac{N_A^A+n-M_A}{N_A+n}, \frac{N_A^A+n+M_A}{N_A+n}\right] & \left[\frac{N_A^B-M_A^B}{N_A+n}, \frac{N_A^B+M_A^B}{N_A+n}\right] & \left[\frac{N_A^C-M_A^C}{N_A+n}, \frac{N_A^C+M_A^C}{N_A+n}\right] \\ \left[\frac{N_B^A-M_B^A}{N_B+n}, \frac{N_B^A+M_B^A}{N_B+n}\right] & \left[\frac{N_B^B+n-M_B}{N_B+n}, \frac{N_B^B+n+M_B}{N_B+n}\right] & \left[\frac{N_B^C-M_B^C}{N_B+n}, \frac{N_B^C+M_B^C}{N_B+n}\right] \\ \left[\frac{N_C^A-M_C^A}{N_C+n}, \frac{N_C^A+M_C^A}{N_C+n}\right] & \left[\frac{N_C^B-M_C^B}{N_C+n}, \frac{N_C^B+M_C^B}{N_C+n}\right] & \left[\frac{N_C^C+n-M_C}{N_C+n}, \frac{N_C^C+n+M_C}{N_C+n}\right] \end{pmatrix} \tag{4.18}$$

Each company's market share is

$$\pi(0) = \left(\left[\frac{S_A+n}{N+3n}, \frac{S_A+n}{N+3n}\right], \left[\frac{S_B+n}{N+3n}, \frac{S_B+n}{N+3n}\right], \left[\frac{S_C+n}{N+3n}, \frac{S_C+n}{N+3n}\right]\right) \tag{4.19}$$

If the situation remains the same, then the market share after period k can be predicted by

$$\pi(k) = \pi(k-1) \times P. \tag{4.20}$$

The formula can also be used to predict the market share at a time when the market has been in a stable state for some time

$$\begin{cases} \pi(*) = \pi(*) \times P, \qquad \pi(*) = (p_A, p_B, p_C); \\ \sum_i^n \underline{p}_i \leq 1 \leq \sum_i^n \overline{p}_i \quad . \end{cases} \tag{4.21}$$

If we assume the number of increasing customers is respectively n_1, n_2, n_3, the one-step transition probability matrix is calculated as follows:

$$P = \begin{pmatrix} \left[\frac{N_A^A+n_1-M_A}{N_A+n_1}, \frac{N_A^A+n_1+M_A}{N_A+n_1}\right] & \left[\frac{N_A^B-M_A^B}{N_A+n_1}, \frac{N_A^B+M_A^B}{N_A+n_1}\right] & \left[\frac{N_A^C-M_A^C}{N_A+n_1}, \frac{N_A^C+M_A^C}{N_A+n_1}\right] \\ \left[\frac{N_B^A-M_B^A}{N_B+n_2}, \frac{N_B^A+M_B^A}{N_B+n_2}\right] & \left[\frac{N_B^B+n_2-M_B}{N_B+n_2}, \frac{N_B^B+n_2+M_B}{N_B+n_2}\right] & \left[\frac{N_B^C-M_B^C}{N_B+n_2}, \frac{N_B^C+M_B^C}{N_B+n_2}\right] \\ \left[\frac{N_C^A-M_C^A}{N_C+n_3}, \frac{N_C^A+M_C^A}{N_C+n_3}\right] & \left[\frac{N_C^B-M_C^B}{N_C+n_3}, \frac{N_C^B+M_C^B}{N_C+n_3}\right] & \left[\frac{N_C^C+n_3-M_C}{N_C+n_3}, \frac{N_C^C+n_3+M_C}{N_C+n_3}\right] \end{pmatrix} \tag{4.22}$$

Each company's market share is

$$\pi(0) = \left(\left[\frac{S_A+n_1}{N+n_1+n_2+n_3}, \frac{S_A+n_1}{N+n_1+n_2+n_3}\right], \left[\frac{S_B+n_2}{N+n_1+n_2+n_3}, \frac{S_B+n_2}{N+n_1+n_2+n_3}\right], \left[\frac{S_C+n_3}{N+n_1+n_2+n_3}, \frac{S_C+n_3}{N+n_1+n_2+n_3}\right]\right) \tag{4.23}$$

4.3.2 Algorithm for computing the interval probability vector after k transitions

by using the interval arithmetic operation given in Nirmala et al. [60] by extending Kaucher [61] interval arithmetic. This arithmetic operation satisfying group properties with respect to addition and multiplication operations and satisfying the distributive relations between intervals, while maintaining the inclusion monotonicity. We give an iterative algorithm (4.1) to find the prediction for any given future period.

```
P = ($\underline{P}$ + $\overline{P}$) / 2;
for s = 1:k do
    $\pi(s-1)$ = ($\underline{\pi}(s-1)$ + $\overline{\pi}(s-1)$) / 2;
    for i = 1:n do
        for j = 1:n do
            Alpha(i,j) = min(($\underline{\pi}_i(s-1)$ * $\underline{P}_{ij}$), ($\overline{\pi}_i(s-1)$ * $\underline{P}_{ij}$), ($\underline{\pi}_i(s-1)$ * $\overline{P}_{ij}$), ($\overline{\pi}_i(s-1)$ * $\overline{P}_{ij}$));
        end for
    end for
```

```

    for i = 1:n do
        for j = 1:n do
            Beta(i,j) = max(($\underline{\pi}_i(s-1)$ * $\underline{P}_{ij}$), ($\overline{\pi}_i(s-1)$ * $\underline{P}_{ij}$), ($\underline{\pi}_i(s-1)$ * $\overline{P}_{ij}$),
                ($\overline{\pi}_i(s-1)$ * $\overline{P}_{ij}$));
        end for
    end for

    for i = 1:n do
        for j = 1:n do
            D(i,j) = min(($\pi_i(s-1)$ * $P_{ij}$ - Alpha(i,j)), (Beta(i,j) -
                $\pi_i(s-1)$ * $P_{ij}$));
        end for
    end for

    for i = 1:n do
        for j = 1:n do
            C1(i,j) = $\pi_i(s-1)$ * $P_{ij}$ - D(i,j);
            C2(i,j) = $\pi_i(s-1)$ * $P_{ij}$ + D(i,j);
        end for
    end for

    for j = 1:n do
        $\underline{\pi}_j(s)$ = 0; $\overline{\pi}_j(s)$ = 0;
        for i = 1:n do
            $\underline{\pi}_j(s)$ = $\underline{\pi}_j(s)$ + C1(i,j);
            $\overline{\pi}_j(s)$ = $\overline{\pi}_j(s)$ + C2(i,j);
        end for
    end for

    return $\underline{\pi}(s)$, $\overline{\pi}(s)$;
end for
```

Listing 4.1: Calculation of the interval probability vector after k transitions

4.3.2.1 Calculate the time complexity of this Matlab code

To calculate the time complexity of this algorithm, we will analyze each part of the algorithm step by step and then combine the results.

The algorithm consists of several nested loops, and we will focus on the operations within these loops:

1. **Outer loop** (`for s = 1:k`):

- This loop runs k times.

2. **First set of nested loops** (`for i = 1:n` and `for j = 1:n`):

- Inside the outer loop, there are two nested loops that iterate n times each.
- The operations inside these loops (calculating Alpha and Beta) are constant time operations.
- Time complexity: $O(n^2)$.

3. **Second set of nested loops** (`for i = 1:n` and `for j = 1:n`):

- These loops also iterate n times each.
- The operations inside these loops (calculating D) are constant time operations.
- Time complexity: $O(n^2)$.

4. **Third set of nested loops** (`for i = 1:n` and `for j = 1:n`):

- These loops iterate n times each.
- The operations inside these loops (calculating C1 and C2) are constant time operations.
- Time complexity: $O(n^2)$.

5. **Fourth set of nested loops** (`for j = 1:n` and `for i = 1:n`):

- These loops iterate n times each.
- The operations inside these loops (updating $\underline{\pi}$ and $\overline{\pi}$) are constant time operations.
- Time complexity: $O(n^2)$.

Since each of these sets of nested loops is inside the outer loop that runs k times, we multiply the time complexity of each set of nested loops by k.

Combining all these results:

- First set of nested loops: $O(k \cdot n^2)$
- Second set of nested loops: $O(k \cdot n^2)$

- Third set of nested loops: $O(k \cdot n^2)$
- Fourth set of nested loops: $O(k \cdot n^2)$

Adding these together, the total time complexity of the algorithm is:

$$O(k \cdot n^2) + O(k \cdot n^2) + O(k \cdot n^2) + O(k \cdot n^2) = O(k \cdot n^2)$$

Thus, the overall time complexity of the given algorithm is $\mathbf{O(k \cdot n^2)}$.

4.3.3 Illustrative example

We consider the same example treated in Lu et al. [57], which analyzes the current situation of three companies based on the Markov theory and predict the company's marked share in the future, but we consider the number of customer's declared by each company as center of symmetrical interval, then the flow of the costumers from the first year to the second year is shown in the following Table 4.1.

Table 4.1: Estimation ranges of the customers numbers (Unit: 10 thousand)

Firm		A	B	C
First year		[240,240]	[600,600]	[360,360]
From other firms	From A	[0,0]	[24,30]	[24,30]
	From B	[42,48]	[0,0]	[18,24]
	From C	[30,36]	[24,30]	[0,0]
To other firms	To A	[0,0]	[42,48]	[30,36]
	To B	[24,30]	[0,0]	[24,30]
	To C	[24,30]	[18,24]	[0,0]
Second year		[252,276]	[576,600]	[336,360]

And we can obtain the reserve rate interval and share interval of each firm in Table4.2.

Table 4.2: Estimation ranges of the rate and share

Firm	A	B	C
First year	[240,240]	[600,600]	[360,360]
Customers going to other firms	[48,60]	[60,72]	[54,66]
Customers coming from other firms	[72,84]	[48,60]	[42,54]
reserved customers range of the second year	[0.75,0.8]	[0.88,0.9]	[0.82,0.85]
share	[0.21,0.23]	[0.48,0.5]	[0.28,0.3]

It follows that the one-step transition interval matrix can be expressed by:

$$P = \begin{pmatrix} \left[\frac{180}{240}, \frac{192}{240}\right] & \left[\frac{24}{240}, \frac{30}{240}\right] & \left[\frac{24}{240}, \frac{30}{240}\right] \\ \left[\frac{42}{600}, \frac{48}{600}\right] & \left[\frac{528}{600}, \frac{540}{600}\right] & \left[\frac{18}{600}, \frac{24}{600}\right] \\ \left[\frac{30}{360}, \frac{36}{360}\right] & \left[\frac{24}{360}, \frac{30}{360}\right] & \left[\frac{294}{360}, \frac{306}{360}\right] \end{pmatrix} \tag{4.24}$$

Then the share interval after k steps can be presented by the formula:

$$\pi(k) = ([0.21, 0.23], [0.48, 0.5], [0.28, 0.3]) \times \begin{pmatrix} \left[\frac{180}{240}, \frac{192}{240}\right] & \left[\frac{24}{240}, \frac{30}{240}\right] & \left[\frac{24}{240}, \frac{30}{240}\right] \\ \left[\frac{42}{600}, \frac{48}{600}\right] & \left[\frac{528}{600}, \frac{540}{600}\right] & \left[\frac{18}{600}, \frac{24}{600}\right] \\ \left[\frac{30}{360}, \frac{36}{360}\right] & \left[\frac{24}{360}, \frac{30}{360}\right] & \left[\frac{294}{360}, \frac{306}{360}\right] \end{pmatrix}^{k-1}$$

After k period, the interval share can be calculated by applying the proposed algorithm. Some obtained results for various periods are illustrated in the Table 4.3

Table 4.3: Some results for different periods

Period	Iteration K	Interval probability distribution$\pi(k)$
First year	1	([0.21,0.23],[0.48,0.5],[0.28,0.3])
Third year	3	([0.2152,0.2716],[0.4457,0.5085],[0.2510,0.3080])
Fifth year	5	([0.2102,0.2991],[0.4161,0.5235],[0.2301,0.3209])
Seventh year	7	([0.2011,0.3190],[0.3896,0.5417],[0.2135,0.3350])
Ninth year	9	([0.1905,0.3349],[0.3653,0.5611],[0.1993,0.3489])

From these results the interval probability distribution and share can be predicted in any market period. We note that if we take the midpoints of each of these intervals for example that after 9 years we have the probability distribution vector of the customers $(0.2627, 0.4632, 0.2741)$, we also note that company B remains the leader in the number of customers compared to companies A and C. Decision makers can predict the expected number of customers for future years in several seconds by knowing the movement of customers from one year to the next, as shown in the Table4.4 below:

Table 4.4: execution time for different transition steps

The time period (years)	Total Time (second)
10	$09.114s$
20	$19.247s$
30	$29.410s$
40	$39.574s$
50	$49.684s$

The first beneficiary is the company, which will be able to attract a larger number of customers in order to maintain its continuity and increase profit, which will have a positive impact on customer service.

4.4 Conclusion

We demonstrated how to apply interval Markov chains (IMC), an extension of discrete Markov chains, to model and predict customer movements between companies. The model treats customer numbers as discrete values within symmetric intervals, addressing fluctuations in customer counts. The results show that the IMC model effectively analyzes current conditions and forecasts future trends, aiding businesses in making timely decisions, enhancing services, and attracting more customers to ensure continuity and boost profits.

The use of interval Markov chains extends beyond this study and is applicable in various other contexts.

CHAPTER 5

APPLICATION OF INTERVAL MARKOV CHAIN IN THE PAGE RANK ALGORITHM

> "The next Bill Gates will not start an operating system. The next Larry Page won't start a search engine. The next Mark Zuckerberg won't start a social network company. If you are copying these people, you are not learning from them."
>
> *Peter Thiel, American Businessman*

5.1 Introduction

The World Wide Web consists of billions of web-pages and a huge amounts of information available in the form of pages. To retrieve required information from the World Wide Web, search engines perform a number of tasks based on their respective architecture. These can be complicated and time consuming. Every search engine process goes from crawling, to indexing, to searching, and then to a ranking of information. A crawler visits and downloads all the web-pages of a given website in order to retrieve the required information, the information provided by crawler has to be stored in a particular order to be accessed by the search engine. The information is then indexed in order to decrease the time needed to look into it. A web search engine represents the interface needed for allowing users to put a query for the required information. This represents the connection between the user and the information repository. So, when a user sends a query to a search engine, there exists an incredible number of web pages related to a given query. However, only a small number of web pages is really useful to the user. Still this number can be very large (in the millions). On the other hand, web users do not have the time and patience to go through all of them. It has actually been documented at [62, 63], that most Web users do not look beyond the first page of results. Therefore, it is important for the ranking function to output the desired results within the top few pages. Otherwise, the search engine would be rendered useless, there are various ranking algorithms [64, 65] including Page Rank, HITS, SALSA, RANDOMZE HITS, SUBSPACE HITS, SIMRANK, etc. In this chapter, we will focus on Page Rank and a proposed improvement for it.

5.2 Google Search Mechanism

Google's search mechanism is a sophisticated process that involves several stages to provide the most relevant results to the user. This section describes the key stages involved in this process, illustrated in Figure 5.1.

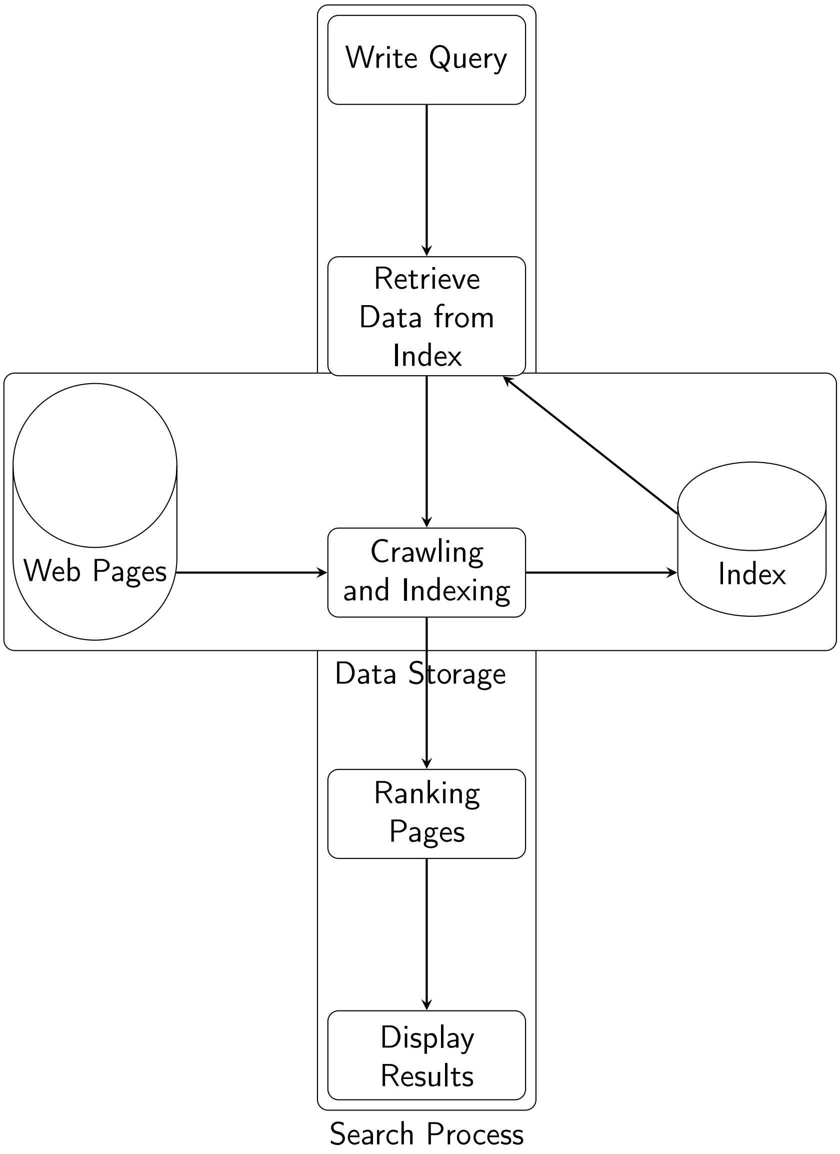

Figure 5.1: Google Search Mechanism

This mechanism is divided into several stages:

5.2.1 Writing the Query

The process starts when the user writes a query in the search bar. This query is the input that drives the entire search mechanism. Google's algorithms analyze the query to understand the user's intent and to identify the relevant keywords.

5.2.2 Retrieving Data from the Index

Once the query is written, Google retrieves data from its pre-built index. The index is a vast database containing information about numerous web pages. This retrieval process involves matching the query keywords with the indexed data to find relevant pages.

5.2.3 Crawling and Indexing

The crawling and indexing stage is crucial for keeping the index up to date. Google's web crawlers, often referred to as spiders or bots, continuously scour the web for new and updated pages. These pages are then indexed, which involves processing and storing the content in a way that makes it easily retrievable during a search.

5.2.4 Ranking Pages

After retrieving data from the index, Google ranks the pages based on relevance and quality. This ranking is determined by various algorithms, including the Page Rank algorithm, which evaluates the importance of pages based on the number and quality of links pointing to them. Other factors, such as content relevance, user engagement, and freshness, also play a role in determining the rank.

5.2.5 Displaying Results

Finally, the ranked pages are displayed to the user as search results. This display includes snippets of information to help the user quickly identify the most relevant results. The goal is to provide a seamless and efficient search experience, ensuring that the user finds the most pertinent information as quickly as possible.

In summary, the Google search mechanism is a complex, multi-stage process designed to deliver the most relevant search results efficiently. Figure 5.1 provides a visual representation of this process, highlighting the flow from query writing to displaying the search results.

5.3 Page Rank algorithm

Page Rank is an algorithm used by Google Search to rank web pages in their search engine results. It was developed by Page and Brin in 1996 as part of their research project at Stanford University. This algorithm assigns a numerical value, called a Page

Rank score, to each web page based on the number and quality of links pointing to that page. The idea behind the algorithm is that a web page is considered more important if it has many links from other important pages. The Page Rank score of a web page is calculated recursively as follows:

1. Each web page is assigned an initial Page Rank score of 1/N, where N is the total number of web pages in the network.
2. The Page Rank score of each web page is updated iteratively based on the Page Rank scores of the web pages linking to it.
3. The Page Rank score of a page is increased by the sum of the Page Rank scores of the pages linking to it, weighted by the importance of each linking page, for more details see Brin and Page [41].
4. The process continues until the Page Rank scores converge to a stable value.

The importance of a linking page is determined by its own Page Rank score and the number of outgoing links it has. A page with a high Page Rank score and few outgoing links will pass more Page Rank value to the pages it links to than a page with a low Page Rank score and many outgoing links. The Page Rank algorithm is one of the most widely used algorithms for ranking web pages, and it has been instrumental in making Google the dominant search engine it is today.

5.4 The Role of Markov Chains in the Page Rank Algorithm

Markov chains form the mathematical backbone of the Page Rank algorithm, providing a robust framework for assessing the relative importance of web pages. By leveraging the properties of Markov chains and incorporating a damping factor, Page Rank effectively handles the complexity of the web's hyperlink structure, ensuring convergence to a stable ranking of pages. The Page Rank algorithm, crucial to Google's search engine, employs Markov chains to evaluate the importance of web pages. Understanding the stability of random models in this context involves recognizing how Markov chains form the foundation of Page Rank's design and functionality.

5.4.1 Markov Chains and Page Rank

At its core, the Page Rank algorithm is a Markov chain model used to determine the relative importance of web pages. Below is a detailed breakdown of how Markov chains are applied:

5.4.1.1 Web page Representation as a Markov Chain

The web is represented as a directed graph where nodes (vertices) correspond to web pages, and directed edges (arcs) represent hyperlinks between these pages. In this graph, a Markov chain is defined where each state corresponds to a web page, and the transition probabilities between states are determined by the likelihood of moving from one page to another via hyperlinks.

5.4.1.2 Transition Matrix

Construct the transition matrix P, where each entry p_{ij} represents the probability of transitioning from web page i to web page j. This matrix is derived from the hyperlink structure of the web. Specifically, p_{ij} is the fraction of links on page i that point to page j. For a web page with no outbound links (a *dangling node*), the algorithm typically redistributes its probability mass uniformly across all pages to ensure a well-defined stochastic matrix.

5.4.1.3 The Page Rank distribution

The Page Rank distribution π is a probability vector where each entry π_i represents the importance or rank of web page i. The goal is to find the steady-state distribution of the Markov chain, where the probability distribution over the pages stabilizes and does not change with further transitions.

5.4.1.4 Irreducibility and Aperiodicity

To ensure that the Markov chain represented by the transition matrix is ergodic (i.e., it has a unique steady-state distribution), the matrix is adjusted to guarantee irreducibility (every state can be reached from every other state) and aperiodicity (the chain does not get trapped in cycles). In practice, a damping factor is introduced to address these properties. The transition matrix is modified to incorporate a damping factor α (Many researchers [41, 66, 67] say the value of α used by the Page Rank algorithm of Google is

0.85.), which accounts for the probability of randomly jumping to any page rather than following hyperlinks. Brin and Page [41] forced all the entries in the transition matrix to satisfy $0 < P_{i,j} < 1$ to make it regular. This adjusted matrix is given by:

$$P' = \alpha \times P + (1-\alpha) \times \frac{E}{N}. \tag{5.1}$$

where N is the number of pages, and E is a matrix with all entries equal to 1.

5.4.1.5 Computing the Page Rank

The steady-state distribution of the adjusted Markov chain P' is computed by solving the eigenvalue problem:

$$\pi = P'\pi, \tag{5.2}$$

where π is the principal eigenvector corresponding to the eigenvalue 1. This steady-state distribution represents the Page Rank scores of the web pages.

5.4.2 Stability and Convergence

The stability of the Page Rank algorithm relies on the convergence properties of the Markov chain. Given the adjustments to the transition matrix to ensure irreducibility and aperiodicity, the Markov chain defined by the Page Rank algorithm converges to a unique stationary distribution.

5.5 Application of Markov chain models:

Our example will be drawn from the study by K. G. A. P. R. Kumar and all [68].
Figure 5.2 illustrates a sample web graph sourced from a university site, featuring seven pages: Home, Admin, Staff, Student, Library, Dept, and Alumni

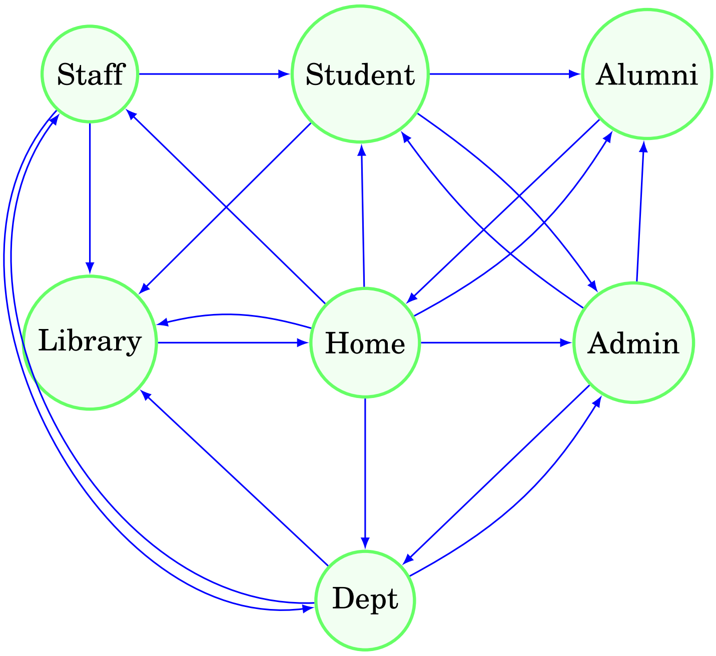

Figure 5.2: Web graph obtained from a university website

The connections between pages are represented by a graph. A node represents a web page and an arrow from page A to page B means that there is a link from page A to page B. The number of out-going links is an important parameter. We use the notation "out-degree of a node" to stand for the number of out-going links contained in a page. This graph is usually referred to as "the web graph", each node in the graph is identified with a page. We will use the term "node and page" interchangeably, and let $L(p)$ be the number of out-going links in a page p. We have: $L(Staff) = 3$, $L(Student) = 3$ and $L(Alumni) = 1$, $L(Library) = 1$, $L(Home) = 6$, $L(Admin) = 3$, $L(Dept) = 3$. So p_{ij} is defined as(5.3):

$$p_{ij} = \begin{cases} \frac{1}{|L_j|} & \text{if there is a link from j to i;} \\ 0 & \text{otherwise,} \end{cases} \tag{5.3}$$

5.5.1 Calculating The Page rank distribution

To create the transition matrix P, we apply Equation (5.3) to the web graph illustrated in Figure 5.2.

$$p = \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{pmatrix}. \tag{5.4}$$

We calculate the Google Matrix in equation 5.1 using the sample Web graph.

$$P' = 0.85 \times \begin{pmatrix} 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{pmatrix} + 0.15 \times \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{pmatrix}$$

$$P' = \begin{pmatrix} 0.021 & 0.305 & 0.021 & 0.305 & 0.021 & 0.021 & 0.305 \\ 0.021 & 0.021 & 0.305 & 0.305 & 0.021 & 0.305 & 0.021 \\ 0.021 & 0.021 & 0.021 & 0.021 & 0.871 & 0.021 & 0.021 \\ 0.021 & 0.021 & 0.021 & 0.021 & 0.871 & 0.021 & 0.021 \\ 0.1631 & 0.1631 & 0.1631 & 0.1631 & 0.021 & 0.1631 & 0.1631 \\ 0.021 & 0.305 & 0.305 & 0.021 & 0.021 & 0.021 & 0.305 \\ 0.305 & 0.021 & 0.021 & 0.305 & 0.021 & 0.305 & 0.021 \end{pmatrix}$$

We repeat the process shown in the equation 5.2 until we obtain the steady state. After the 63th iteration, the following are the stationary vector for our sample 7-page Web graph.

$$s = [0.171, 0.110, 0.117, 0.139, 0.238, 0.115, 0.110]$$

5.6 Interval Markov Chain Model's Application

Now we will change this web graph by placing two types of links (green links and red ones), the green links represent the largest expected links to move from one page to the rest of the pages, and the red links represent the smallest expected links to move from one page to the rest of them.

Therefore, moving from one page to another will be through an interval and not a fixed link. This interval will be according to the links in red and green. This change is in order to embody and clarify an idea, the idea is to use Markov Interval Chain models instead of Markov chains, because after completing the crawling and indexing phase, the links between pages may change before performing a new crawling process. The user initiates his query, and the search results will be among the old pages, before changing the links between them, and here lies the issue that we seek to solve then the links between pages can be changed over time. So, the results obtained through Markov chains will be inaccurate. Overall, the proposed model involves considering the number of links between pages that are discrete values as centers of symmetric intervals. Through said model, it would be possible to avoid the problem of dangling node. as shown in the following figure 5.3.

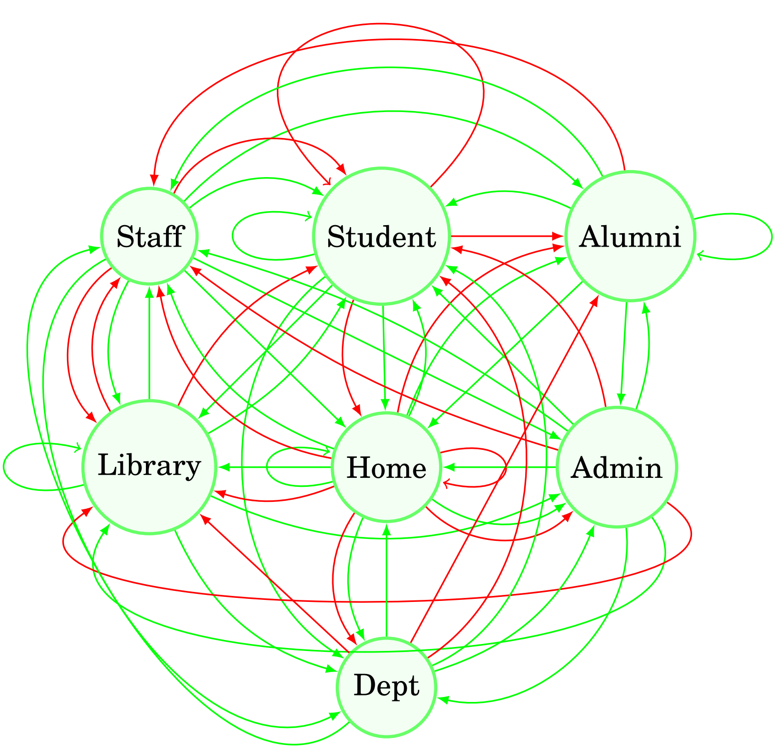

Figure 5.3: Web graph site after modifying the links

Let $\overline{L}_{(P)}$, be the largest number of out-going links in a page p, and $\underline{L}_{(P)}$, be the smallest number of out-going links in a page p, and we have:
$\overline{L}_{(Staff)} = 6$, $\underline{L}_{(Staff)} = 2$
$\overline{L}_{(Student)} = 4$, $\underline{L}_{(Student)} = 3$
$\overline{L}_{(Alumni)} = 5$, $\underline{L}_{(Alumni)} = 1$
$\overline{L}_{(Library)} = 5$, $\underline{L}_{(Library)} = 2$
$\overline{L}_{(Home)} = 7$, $\underline{L}_{(Home)} = 6$
$\overline{L}_{(Admin)} = 6$, $\underline{L}_{(Admin)} = 3$
$\overline{L}_{(Dept)} = 4$, $\underline{L}_{(Dept)} = 3$
By comparing the two figures above, we find that the links between pages can change (whether by adding new links or deleting old links), and therefore we suggest replacing the fixed values (which represent the probability of moving from one page to another) by intervals, and this is what provides us with an interval Markov chain.

5.6.1 Discussing the case of adding new links or deleting old ones

These links are considered likely to increase and decrease by a certain amount like what we did in our previous article [69], and therefore, the idea behind implementing this work that instead of considering these links, we need to consider arbitrary intervals that contain these links as centers. After one transition step, the interval which represents the possibility of moving from one page to another. Let N be the total number of pages, and we create an $N \times N$ matrix A by defining the (i,j)-entry as:

$$\underline{p}_{ij} = \begin{cases} \frac{1}{\overline{L}_{(j)}} & \text{if there is a link from i to j;} \\ 0 & \text{otherwise,} \end{cases} \tag{5.5}$$

and

$$\overline{p}_{ij} = \begin{cases} \frac{1}{\underline{L}_{(j)}} & \text{if there is a link from i to j;} \\ 0 & \text{otherwise,} \end{cases} \tag{5.6}$$

where $\underline{p}_{ij}$ represents the smallest probability that go from page i to page j, while $\overline{p}_{i,j}$ representing the largest probability that goes from page i to page j. Then the one-step interval transition probability matrix can be calculated and presented by:

$$p = \begin{pmatrix} \left[\underline{p}_{1,1},\overline{p}_{1,1}\right] & \left[\underline{p}_{1,2},\overline{p}_{1,2}\right] & \left[\underline{p}_{1,3},\overline{p}_{1,3}\right] & \left[\underline{p}_{1,4},\overline{p}_{1,4}\right] & \left[\underline{p}_{1,5},\overline{p}_{1,5}\right] & \left[\underline{p}_{1,6},\overline{p}_{1,6}\right] & \left[\underline{p}_{1,7},\overline{p}_{1,7}\right] \\ \left[\underline{p}_{2,1},\overline{p}_{2,1}\right] & \left[\underline{p}_{2,2},\overline{p}_{2,2}\right] & \left[\underline{p}_{2,3},\overline{p}_{2,3}\right] & \left[\underline{p}_{2,4},\overline{p}_{2,4}\right] & \left[\underline{p}_{2,5},\overline{p}_{2,5}\right] & \left[\underline{p}_{2,6},\overline{p}_{2,6}\right] & \left[\underline{p}_{2,7},\overline{p}_{2,7}\right] \\ \left[\underline{p}_{3,1},\overline{p}_{3,1}\right] & \left[\underline{p}_{3,2},\overline{p}_{3,2}\right] & \left[\underline{p}_{3,3},\overline{p}_{3,3}\right] & \left[\underline{p}_{3,4},\overline{p}_{3,4}\right] & \left[\underline{p}_{3,5},\overline{p}_{3,5}\right] & \left[\underline{p}_{3,6},\overline{p}_{3,6}\right] & \left[\underline{p}_{3,7},\overline{p}_{3,7}\right] \\ \left[\underline{p}_{4,1},\overline{p}_{4,1}\right] & \left[\underline{p}_{4,2},\overline{p}_{4,2}\right] & \left[\underline{p}_{4,3},\overline{p}_{4,3}\right] & \left[\underline{p}_{4,4},\overline{p}_{4,4}\right] & \left[\underline{p}_{4,5},\overline{p}_{4,5}\right] & \left[\underline{p}_{4,6},\overline{p}_{4,6}\right] & \left[\underline{p}_{4,7},\overline{p}_{4,7}\right] \\ \left[\underline{p}_{5,1},\overline{p}_{5,1}\right] & \left[\underline{p}_{5,2},\overline{p}_{5,2}\right] & \left[\underline{p}_{5,3},\overline{p}_{5,3}\right] & \left[\underline{p}_{5,4},\overline{p}_{5,4}\right] & \left[\underline{p}_{5,5},\overline{p}_{5,5}\right] & \left[\underline{p}_{5,6},\overline{p}_{5,6}\right] & \left[\underline{p}_{5,7},\overline{p}_{5,7}\right] \\ \left[\underline{p}_{6,1},\overline{p}_{6,1}\right] & \left[\underline{p}_{6,2},\overline{p}_{6,2}\right] & \left[\underline{p}_{6,3},\overline{p}_{6,3}\right] & \left[\underline{p}_{6,4},\overline{p}_{6,4}\right] & \left[\underline{p}_{6,5},\overline{p}_{6,5}\right] & \left[\underline{p}_{6,6},\overline{p}_{6,6}\right] & \left[\underline{p}_{6,7},\overline{p}_{6,7}\right] \\ \left[\underline{p}_{7,1},\overline{p}_{7,1}\right] & \left[\underline{p}_{7,2},\overline{p}_{7,2}\right] & \left[\underline{p}_{7,3},\overline{p}_{7,3}\right] & \left[\underline{p}_{7,4},\overline{p}_{7,4}\right] & \left[\underline{p}_{7,5},\overline{p}_{7,5}\right] & \left[\underline{p}_{7,6},\overline{p}_{7,6}\right] & \left[\underline{p}_{7,7},\overline{p}_{7,7}\right] \end{pmatrix}. \tag{5.7}$$

where:

1. the page with the number 1 is the Staff page,
2. the page with the number 2 is the Student page,
3. the page with the number 3 is the Alumni page,
4. the page with the number 4 is the Library page,

5. the page with the number 5 is the Home page,
6. the page with the number 6 is the Admin page,
7. the page with the number 7 is the Dept page.

The interval probability distribution of the initial state can be expressed as:

$$\pi(0)=\left(\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right],\left[\frac{1}{N},\frac{1}{N}\right]\right) \tag{5.8}$$

5.6.2 Google's matrix

Google's matrix for this type is of the following form:5.1

$$P'=0.85\times\begin{pmatrix}
\left[0,0\right] & \left[\frac{1}{6},\frac{1}{2}\right] & \left[0,\frac{1}{6}\right] & \left[\frac{1}{6},\frac{1}{2}\right] & \left[0,\frac{1}{6}\right] & \left[0,\frac{1}{6}\right] & \left[0,\frac{1}{6}\right] \\
\left[0,0\right] & \left[\frac{1}{4},\frac{1}{3}\right] & \left[0,\frac{1}{3}\right] & \left[0,\frac{1}{4}\right] & \left[\frac{1}{4},\frac{1}{3}\right] & \left[0,0\right] & \left[0,\frac{1}{4}\right] \\
\left[\frac{1}{5},1\right] & \left[0,\frac{1}{5}\right] & \left[0,\frac{1}{5}\right] & \left[0,0\right] & \left[0,\frac{1}{5}\right] & \left[0,\frac{1}{5}\right] & \left[0,0\right] \\
\left[\frac{1}{5},\frac{1}{2}\right] & \left[\frac{1}{5},\frac{1}{2}\right] & \left[0,0\right] & \left[0,\frac{1}{5}\right] & \left[0,0\right] & \left[0,\frac{1}{5}\right] & \left[0,\frac{1}{5}\right] \\
\left[\frac{1}{7},\frac{1}{6}\right] & \left[0,\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{6}\right] & \left[\frac{1}{7},\frac{1}{6}\right] & \left[\frac{1}{7},\frac{1}{6}\right] & \left[\frac{1}{7},\frac{1}{6}\right] & \left[\frac{1}{7},\frac{1}{6}\right] \\
\left[\frac{1}{6},\frac{1}{3}\right] & \left[\frac{1}{6},\frac{1}{3}\right] & \left[0,\frac{1}{6}\right] & \left[\frac{1}{6},\frac{1}{3}\right] & \left[0,\frac{1}{6}\right] & \left[0,0\right] & \left[0,\frac{1}{6}\right] \\
\left[0,\frac{1}{4}\right] & \left[\frac{1}{4},\frac{1}{3}\right] & \left[0,\frac{1}{3}\right] & \left[0,\frac{1}{3}\right] & \left[0,\frac{1}{4}\right] & \left[0,\frac{1}{4}\right] & \left[0,0\right]
\end{pmatrix}$$

$$+0.15\times\begin{pmatrix}
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] \\
\left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right] & \left[\frac{1}{7},\frac{1}{7}\right]
\end{pmatrix}$$

$$=\begin{pmatrix}
[0.0214,0.0214] & [0.1631,0.4464] & [0.0214,0.1631] & [0.1631,0.4464] & [0.0214,0.1631] & [0.0214,0.1631] & [0.214,0.1631] \\
[0.0214,0.0214] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.0214] & [0.0214,0.2339] \\
[0.1914,0.8714] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] \\
[0.1914,0.4464] & [0.1914,0.4464] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] \\
[0.1429,0.1631] & [0.0214,0.1429] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] \\
[0.1631,0.3048] & [0.1631,0.3048] & [0.0214,0.1631] & [0.1631,0.3048] & [0.0214,0.1631] & [0.0214,0.0214] & [0.0214,0.1631] \\
[0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.0214,0.2339] & [0.0214,0.0214]
\end{pmatrix}$$

After one transition step, each interval page rank can be given According to Equation 4.9. Then after k period, the rank of pages can be predicted through the initial interval probability matrix and one-step transition probability matrix. According to Equation 5.3 considering that after a certain period, the page rank reached a stable state:

$$P_1 = \left[\underline{p}_1, \overline{p}_1\right], P_2 = \left[\underline{p}_2, \overline{p}_2\right], P_3 = \left[\underline{p}_3, \overline{p}_3\right], P_4 = \left[\underline{p}_4, \overline{p}_4\right], P_5 = \left[\underline{p}_5, \overline{p}_5\right], P_6 = \left[\underline{p}_6, \overline{p}_6\right], P_7 = \left[\underline{p}_7, \overline{p}_7\right]. \tag{5.9}$$

Is respectively the page rank interval of each page, and $P_1, P_2, P_3, P_4, P_5, P_6$, and P_7 can be obtained by the formula:

$$\begin{cases} \pi(*) = \pi(*) \times P & \pi(*) = (P_1, P_2, P_3, P_4, P_5, P_6, P_7) \\ \sum_i^n \underline{P} \leq 1 \leq \sum_i^n \overline{P}. \end{cases} \tag{5.10}$$

5.6.3 The Algorithm for computing the interval probability vector after m transitions

We use the algorithm that we presented in the previous chapter(Algorithm 4.1)

5.6.4 Illustrative example

We're now going to apply all of this to our example of seven pages see figure 5.3.

Table 5.1: Estimation ranges of pages

pages	Staff	Student	Alumni	Library	Home	Admin	Dept
first Iteration	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$	$\left[\frac{1}{7}, \frac{1}{7}\right]$
to Staff	$[0,0]$	$[2,6]$	$[0,6]$	$[2,6]$	$[0,6]$	$[0,6]$	$[0,6]$
to Student	$[0,0]$	$[3,4]$	$[0,3]$	$[0,4]$	$[3,4]$	$[0,0]$	$[0,4]$
to Alumni	$[1,5]$	$[0,5]$	$[0,5]$	$[0,0]$	$[0,5]$	$[0,5]$	$[0,0]$
to Library	$[2,5]$	$[2,5]$	$[0,0]$	$[0,5]$	$[0,0]$	$[0,5]$	$[0,5]$
to Home	$[6,7]$	$[0,6]$	$[6,7]$	$[6,7]$	$[6,7]$	$[6,7]$	$[6,7]$
to Admin	$[3,6]$	$[3,6]$	$[0,6]$	$[0,6]$	$[0,6]$	$[0,0]$	$[0,6]$
to Dept	$[0,4]$	$[3,4]$	$[0,3]$	$[0,3]$	$[0,4]$	$[0,4]$	$[0,0]$
second Iteration	$[0.1076, 0.2946]$	$[0.1469, 0.3059]$	$[0.0388, 0.1874]$	$[0.0793, 0.2380]$	$[0.0691, 0.1773]$	$[0.0388, 0.1408]$	$[0.0388, 0.1368]$

It follows that the one-step transition interval matrix can be expressed by:

$$\text{p}=\begin{pmatrix} [0.0214,0.0214] & [0.1631,0.4464] & [0.0214,0.1631] & [0.1631,0.4464] & [0.0214,0.1631] & [0.0214,0.1631] & [0.214,0.1631] \\ [0.0214,0.0214] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.0214] & [0.0214,0.2339] \\ [0.1914,0.8714] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] \\ [0.1914,0.4464] & [0.1914,0.4464] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] \\ [0.1429,0.1631] & [0.0214,0.1429] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] \\ [0.1631,0.3048] & [0.1631,0.3048] & [0.0214,0.1631] & [0.1631,0.3048] & [0.0214,0.1631] & [0.0214,0.0214] & [0.0214,0.1631] \\ [0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.0214,0.2339] & [0.0214,0.0214] \end{pmatrix}. \tag{5.11}$$

Then the interval rank pages after k steps can be presented by the formula:

$$\pi(k)=\pi(0)\times P^{k-1}$$

$$\pi(k)=\left(\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right],\left[\frac{1}{7},\frac{1}{7}\right]\right)\times$$

$$\begin{pmatrix} [0.0214,0.0214] & [0.1631,0.4464] & [0.0214,0.1631] & [0.1631,0.4464] & [0.0214,0.1631] & [0.0214,0.1631] & [0.214,0.1631] \\ [0.0214,0.0214] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.0214] & [0.0214,0.2339] \\ [0.1914,0.8714] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] & [0.0214,0.0214] \\ [0.1914,0.4464] & [0.1914,0.4464] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.0214] & [0.0214,0.1914] & [0.0214,0.1914] \\ [0.1429,0.1631] & [0.0214,0.1429] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] & [0.1429,0.1631] \\ [0.1631,0.3048] & [0.1631,0.3048] & [0.0214,0.1631] & [0.1631,0.3048] & [0.0214,0.1631] & [0.0214,0.0214] & [0.0214,0.1631] \\ [0.0214,0.2339] & [0.2339,0.3048] & [0.0214,0.3048] & [0.0214,0.3048] & [0.0214,0.2339] & [0.0214,0.2339] & [0.0214,0.0214] \end{pmatrix}^{k-1}$$

After k period, the interval share can be calculated by applying the proposed algorithm. Some obtained results for various periods are illustrated in Table2

Table 5.2: Some results for different iterations

Iteration K	Interval probability distribution ($\pi(k)$)
1	([0.1076,0.2946], [0.1469,0.3059], [0.0388,0.1874], [0.0793,0.2380], [0.0691,0.1773], [0.0388,0.1408], [0.0388,0.1368])
2	([0.0451,0.2967], [0.0848, 0.3945], [0.0195,0.2052], [0.0403,0.2871], [0.0507,0.2159], [0.0195,0.1491], [0.0195,0.1720])
4	([0.0144,0.3313], [0.0242,0.4482], [0.0070,0.2221], [0.0123,0.3051], [0.0164,0.2563], [0.0070,0.1630], [0.0070,0.1855])
8	([0.0014,0.3450], [0.0024,0.4700], [0.0007,0.2282], [0.0012,0.3162], [0.0016,0.2707], [0.0007,0.1694], [0.0007,0.1918])
9	([0.0008,0.3457], [0.0013,0.4711], [0.0004,0.2285], [0.0007,0.3167], [0.0009,0.2714], [0.0004,0.1697], [0.0004,0.1921])

5.6.5 Important findings:

1. After 66th iteratins, we reach a state of stability,where: $\pi(66) = \pi(67)$
2. If we take the center of each intervals among the intrval as the rank of these pages, we have the probability distribution vector of page as follows:

$$\left(\left[0,1732\right], \left[0,2362\right], \left[0,1145\right], \left[0,1587\right], \left[0,1362\right], \left[0,0851\right], \left[0,0962\right]\right)$$

3. If we sum the ranking results of all pages, we get 1.
4. We can give each page its final ranking after the state of stability according to the following table:

Table 5.3: pages Rank after the stability state

Page	Staff	Student	Alumni	Library	Home	Admin	Dept
The Rank	0,1732	0,2362	0,1145	0,1587	0,1362	0,0851	0,0962

5.7 Conclusion

This chapter explores the application of interval Markov chains (IMC) as an extension of discrete Markov chains, to model page rank algorithm. This proposed model is based on considering the number of links between pages as an interval, and this can be more realistic, so that the minimum value of the interval is the smallest possible number of links between pages and the largest value is the largest possible number of links, and this is what embodies reality because the links between pages are constantly exposed to change, whether by adding new links between pages or deleting existing ones, we conclude from this work that if we consider the number of links between pages as a fixed number that does not change, then the rank of pages will be inaccurate, in addition to that it is possible to get rid of the problem of dangling node (and this is if adding New links between these pages and the rest of them). After repeating the process of calculating the pages rank, we will eventually get a stationary interval that represents the final pages rank.Therefore,we adopt the middle of the interval as the final pages rank.

CHAPTER 6

CONCLUSION AND FUTURE WORK

> "Little by little, a little becomes a lot."
>
> *Tanzanian proverb*

In this thesis, we have delved into the stability of stochastic models with a specific focus on Markov interval chains (IMC) as a tool for solving decision problems. Our primary objective was to establish a comprehensive understanding of the stability characteristics of these models and to elucidate the practical applications of IMC in decision-making scenarios.

We began by exploring the theoretical underpinnings of stochastic models, emphasizing the critical role that stability plays in ensuring the reliability and predictability of these models. Through a rigorous examination of the mathematical foundations, we developed a framework for assessing the stability of various stochastic processes. This framework was subsequently applied to IMC, enabling us to derive significant insights into their behavior under different conditions.

The Markov interval chain, a novel concept introduced in our research, was shown to be a robust method for addressing decision problems characterized by uncertainty and variability. By discretizing the state space and incorporating interval-based transitions, IMC offers a flexible and adaptive approach to model stochastic processes. Our experimental results demonstrated the effectiveness of IMC in providing accurate and stable solutions across a range of decision-making scenarios.

One of the key contributions of this thesis is the development of a set of stability criteria for IMC. These criteria, derived from both theoretical analysis and empirical validation, provide a valuable tool for practitioners seeking to ensure the reliability of their models. We have shown that by adhering to these criteria, it is possible to mitigate the risks associated with instability and enhance the overall performance of the decision-making process.

Furthermore, our research has highlighted the versatility of IMC in handling diverse types of decision problems. From resource allocation and scheduling to risk assessment and policy planning, the applications of IMC are vast and varied. By demonstrating the practical utility of IMC in these domains, we have paved the way for future research and application in both academic and industrial settings.

However, it is important to acknowledge the limitations of our study. One area where our research could be extended is in the scalability of IMC to very large decision spaces. While our experiments have shown promising results in moderate-sized scenarios, applying IMC to extremely large-scale problems may require further algorithmic optimizations and computational resources.

The dissertation's future work can be summarized as follows:

- Theoretical Advancements:
  - Extension of Stability Criteria: The stability criteria developed in this thesis can be further refined and extended to accommodate more complex and dynamic systems. Future research could focus on identifying additional factors that influence stability and developing more comprehensive criteria that account for these factors.
  - Advanced Analytical Techniques: Incorporating advanced mathematical techniques, such as stochastic calculus and dynamic programming, could provide deeper insights into the behavior of IMC and enhance our ability to predict their stability under various conditions
  - Comparative Analysis: Conducting comparative studies between IMC and other stochastic modeling approaches, such as Markov decision processes (MDPs) and Bayesian networks, could highlight the relative strengths and weaknesses of each method. This analysis could inform the selection of the most appropriate tool for specific decision problems.
- Practical Implementations:

  - Algorithm Optimization: While IMC has proven to be effective, there is room for optimization in terms of computational efficiency and scalability. Developing more efficient algorithms for interval calculations and state transitions could significantly enhance the practical utility of IMC in large-scale applications.
  - Software Development: Creating user-friendly software tools and libraries that implement IMC would facilitate its adoption by practitioners and researchers. These tools could include visualization capabilities, simulation environments, and integration with existing decision-support systems.
  - Case Studies and Real-World Applications: Conducting detailed case studies in various domains, such as finance, healthcare, and logistics, could provide valuable insights into the practical challenges and benefits of using IMC. Collaborating with industry partners to implement and evaluate IMC in real-world settings would further validate its effectiveness

- Cross-Disciplinary Applications:
  - Interdisciplinary Research: Exploring the application of IMC in fields beyond traditional decision science, such as biology, environmental science, and social sciences, could uncover new opportunities and challenges. Interdisciplinary research efforts could lead to innovative solutions and broaden the impact of IMC.
  - Integration with Machine Learning: Combining IMC with machine learning techniques, particularly in the context of reinforcement learning and predictive modeling, could enhance the ability to handle complex and dynamic environments. This integration could result in more adaptive and intelligent decision-making systems.

In conclusion, the exploration of stability in stochastic models and the development of Markov interval chains represent significant contributions to the field of decision science. The insights gained from this research provide a solid foundation for future work, offering numerous opportunities for theoretical, practical, and interdisciplinary advancements. By continuing to build on this foundation, we can unlock the full potential of IMC and other stochastic modeling approaches, ultimately leading to more robust and effective decision-making processes in an increasingly uncertain world.

BIBLIOGRAPHY

[1] N. A. Gershenfeld, *The Nature of Mathematical Modeling*. Cambridge University Press, 1999.

[2] H. Peng, W. Zhao, H. Tan, Z. Ji, K. Li, and X. Zhou, "Prediction of treatment efficacy for prostate cancer using a mathematical model," *Scientific Reports*, pp. 1–13, 2016.

[3] V. Boekel, "Kinetic modeling of food quality: a critical review," *Comprehensive Reviews in Food Science and Food Safety*, pp. 144–158, 2008.

[4] J. Krumm, "A markov model for driver turn prediction," in *Society of Automotive Engineers (SAE) 2008 World Congress*, 2008.

[5] F. Galvanin, A. Boschiero, M. Barolo, and F. Bezzo, "Model-based design of experiments in the presence of continuous measurement systems," *Industrial & Engineering Chemistry Research*, vol. 50, no. 4, pp. 2167–2175, 2011.

[6] L. Papageorgiou, G. Rotstein, and N. Shah, "Strategic supply chain optimization for the pharmaceutical industries," *Industrial & Engineering Chemistry Research*, vol. 40, no. 1, pp. 275–286, 2001.

[7] S. V. Patankar and D. B. Spalding, "A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows," in *Numerical Prediction of Flow*, pp. 54–73, Elsevier, 1983.

[8] P. M. Ferziger, M. Perić, and R. L. Street, *Computational Methods for Fluid Dynamics*. Springer, 2002.

[9] A. D. Schmidt and W. H. Ray, "The dynamic behavior of continuous polymerization reactors—i: Isothermal solution polymerization in a cstr," *Chemical Engineering Science*, vol. 36, no. 8, pp. 1401–1410, 1981.

[10] M. Quaglio, *Novel techniques for kinetic model identification and improvement*. PhD thesis, University College London, 2020.

[11] S. Ross, *Introduction to Probability Models*. Academic Press, 2014.

[12] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Springer, 1992.

[13] J. C. Hull, *Options, Futures, and Other Derivatives*. Prentice Hall, 1989.

[14] C. Obropta and J. S. Kardos, “Review of urban stormwater quality models: Deterministic, stochastic, and hybrid approaches,” *American Water Resources Association*, 01 November 2007.

[15] R. Khasminskii, “Stochastic stability of differential equations,” *Springer Berlin, Heidelberg*, pp. 1–342, 20 September 2011.

[16] A. Lahrouz, L. Omari, D. Kiouach, and A. Belmaâti, “Deterministic and stochastic stability of a mathematical model of smoking,” *Statistics Probability Letters*, pp. 1276–1284, 25 March 2011.

[17] D. Pal, D. Kesh, and D. Mukherjee, “Pattern dynamics in a predator–prey model with smith growth function and prey refuge in predator poisoned environment,” *Chinese Journal of Physics*, pp. 99–108, 19 September 2024.

[18] C. Cheng-Shang, “Stability, queue length, and delay of deterministic and stochastic queueing networks,” *IEEE Transactions on Automatic Control*, pp. 913–931, 1 May 1994.

[19] O. Ovaskainen and B. Meerson, “Stochastic models of population extinction,” *Trends in Ecology Evolution (TREE)*, pp. 643–652, November 2010.

[20] L. Demetrius, V. M. Gundlach, and G. Ochs, “Complexity and demographic stability in population models,” *Theoretical Population Biology*, pp. 211–225, May 2004.

[21] V. G. Kulkarni, “Stochastic models in health care,” pp. 1–129, 15 November 2016.

[22] C. T. Albulescu, “forecasting the romanian financial system stability using a stochastic simulation model,” *Romanian Journal of Economic Forecasting*, 2010.

[23] A. Ofomata, S. Inyama, A. Omame, and R. Umaña, “A stochastic model of the dynamics of stock price for forecasting,” *Journal of Advances in Mathematics and Computer Science*, pp. 1–24, January 2018.

[24] M.E.Khomssi, B.D.Coulibaly, and Ch.Ghizlane, “An approach to stochastic differential equations for long-term forecasting,” *Mathematical Modelling and Numerical Simulation with Applications*, pp. 165–192, 30.06.2024.

[25] A. Omame, M. Abbas, and A. Din, “Global asymptotic stability, extinction and ergodic stationary distribution in a stochastic model for dual variants of sars-cov-2,” *Mathematics and Computers in Simulation*, pp. 302–336, February 2023.

[26] S. Hussain, O. Tunç, H. K. G. Rahman, and E. Nadia, “Mathematical analysis of stochastic epidemic model of mers-corona application of ergodic theory,” *Mathematics and Computers in Simulation*, pp. 130–150, May 2023.

[27] Y. Ding, M. Xu, and L. Hu, “Asymptotic behavior and stability of a stochastic model for aids transmission,” *Applied Mathematics and Computation*, pp. 99–108, 1 October 2008.

[28] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*. John Wiley & Sons, 1983.

[29] X.Ding, Y. Ma, and J. Tang, “Markov chain models in covid-19 prediction: State-of-the art and future perspectives,” *Highlights in Science, Engineering and Technology*, 2023.

[30] B. Mor, S. Garhwal, and A. Kumar, “A systematic review of hidden markov models and their applications,” *Springer*, 2021.

[31] P. J. Schweitzer, “A survey of aggregation-disaggregation in large markov chains,” *Taylor Francis Online*, 2021.

[32] J.Chiachio, M. L. Jalon, M.Chiachio, and A.Kolios, “A markov chains prognostics framework for complex degradation processes,” *Elsevier*, 2020.

[33] C.Baier, C. Hensel, L. Hutschenreiter, S.Junges, and J. Klein, “Parametric markov chains: Pctl complexity and fraction-free gaussian elimination,” *Elsevier*, 2020.

[34] J. W. J. Munkhammar, D.Meer, “Very short term load forecasting of residential electricity consumption using the markov-chain mixture distribution (mcm) model,” *Elsevier*, 2021.

[35] D.DeFord, M. Duchin, and J. Solomon, “Very short term load forecasting of residential electricity consumption using the markov-chain mixture distribution (mcm) model,” *Harvard Data Science Review*, 2021.

[36] J. R. Norris, “Markov chains,” *Cambridge University Press*, 1997.

[37] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*. Athena Scientific, 2008.

[38] R. Howard, *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[39] K. G. Murty, *Linear and Combinatorial Programming*. Wiley-Interscience, 2015.

[40] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web,” *Proceedings of the 7th International World Wide Web Conference (WWW7)*, 1996.

[41] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web,” Tech. Rep. SIDL-WP-1999-0120, Stanford Digital Libraries, 1999.

[42] A, Ntoulas and Junghoo,C and Christopher,O, “What’s new on the web? the evolution of the web from a search engine perspective,” 2004.

[43] S. Fortunato, *Community Detection in Graphs*. Cambridge University Press, 2009.

[44] D. P. Bertsekas and R. G. Gallager, *Data Networks*. Prentice Hall, 1992.

[45] D. Medhi and K. Ramasamy, *Network Routing: Algorithms, Protocols, and Architectures*. Elsevier, 2007.

[46] A. Silberschatz, P. B. Galvin, and G. Gagne, *Operating System Concepts*. Wiley, 2018.

[47] W. Stallings, *Operating Systems: Internals and Design Principles*. Pearson, 2018.

[48] A. S. Tanenbaum and M. Van Steen, *Distributed Systems: Principles and Paradigms*. Pearson, 2016.

[49] E. N. Philip A. Bernstein, *Principles of Transaction Processing*. Morgan Kaufmann, 2009.

[50] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991.

[51] G. F. Riley and T. R. Henderson, "The ns-3 network simulator," in *Modeling and Tools for Network Simulation*, pp. 15–34, Springer, 2010.

[52] G. Strang, *Linear Algebra and Its Applications*. Cengage Learning, 2006.

[53] D. C. Lay, *Linear Algebra and Its Applications*. Addison-Wesley, 2021.

[54] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.

[55] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[56] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. SIAM, 1997.

[57] Y. Lu, M. Zhang, T. Yu, and M. Qu, "Application of markov prediction method in the decision of insurance company," in *Proceedings of the Conference on Insurance Decision Making*, pp. 294–297, 2014.

[58] B. Jonsson and K. G. Larsen, "Specification and refinement of probabilistic processes," *IEEE Computer*, pp. 266–277, 1991.

[59] L. M. de Campos, J. F. Huete, S. Mora, and S. Moral, "Probability intervals: a tool for uncertain reasoning," pp. 167—-196, 1994.

[60] T. Nirmala, D. Datta, H. Kushwaha, and K. Ganesan, "On the solution of system of interval linear equations," in *Proceedings of the Conference on Interval Linear Equations*, pp. 15777–15797, 2015.

[61] E. Kaucher, "Interval analysis in the extended interval space ir, fundamentals of numerical computation," in *Proceedings of the Conference on Numerical Computation*, pp. 33–49, 1980.

[62] Y. Jiang and K. C. Sinha, "Bridge service life prediction model using the markov chain," in *Proceedings of the Conference on Bridge Service Life Prediction*, pp. 24–30, 1989.

[63] C. Silverstein, M. Henzinger, H. Marais, and M. Moricz, "Analysis of a very large altavista query log," Tech. Rep. 1998-014, Digital SRC, 1998.

[64] J. Kleinberg, “Authoritative sources in a hyper-linked environment,” *Journal of the ACM*, vol. 46, no. 5, pp. 604–632, 1999.

[65] R. Lempel and S. Moran, “The stochastic approach for link-structure analysis (salsa) and the tkc effect,” in *Proceedings of the 9th World Wide Web Conference (WWW9)*, pp. 387–401, Elsevier Science, 2000.

[66] A. Langville and C. D. Meyer, “Survey of eigenvector methods of web information retrieval,” in *Proceedings of the Conference on Web Search and Data Mining*, pp. 135–161, 2005.

[67] M. Bianchini, M. Gori, and F. Scarselli, “Inside pagerank,” in *Proceedings of the Conference on Web Search and Data Mining*, pp. 92–128, 2005.

[68] R.Kumar, K.S.Ashutosh, and K.L.Goh, “Application of markov chain in the pagerank algorithm,” *Pertanika Journal of Science and Technology*, pp. 541–554, 2013.

[69] S. E. Semati and A. Gasmi, “Markov interval chain (imc) for solving a decision problem,” *OPSEARCH*, pp. 802–811, 2023.