Summer 2017

Finite Element Maximum Entropy Method for Approximating Absolutely Continuous Invariant Measures

Tulsi Upadhyay
University of Southern Mississippi

The University of Southern Mississippi

FINITE ELEMENT MAXIMUM ENTROPY METHOD FOR APPROXIMATING ABSOLUTELY CONTINUOUS INVARIANT MEASURES

by

Tulsi Upadhyay

Dissertation
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

August 2017

ABSTRACT

FINITE ELEMENT MAXIMUM ENTROPY METHOD FOR APPROXIMATING ABSOLUTELY CONTINUOUS INVARIANT MEASURES

by Tulsi Upadhyay

August 2017

In a chaotic dynamical system, the eventual behavior of iterates of initial points of a map is unpredictable even though the map is deterministic. A system which is chaotic in a deterministic point of view may be regular in a statistical viewpoint. The statistical viewpoint requires the study of absolutely continuous invariant measure (ACIM) of a map with respect to the Lebesgue measure. An invariant density of the Frobenius-Perron (F-P) operator associated with a nonsingular map is employed to evaluate an ACIM of the map. The ACIM is a key factor for studying the eventual behavior of iterates of almost all initial points of the map. It is difficult to obtain an invariant density of the F-P operator in an exact mathematical form except for some simple maps. Different numerical schemes have been developed to approximate such densities.

The maximum entropy principle gives a criterion to select a least-biased density among all densities satisfying a system of moment equations. In this principle, a least-biased density maximizes the Boltzmann entropy. In this dissertation, piecewise quadratic functions and quadratic splines are used in the maximum entropy method to calculate the L^1 errors between the exact and the approximate invariant densities of the F-P operator associated with nonsingular maps defined from $[0,1]$ to itself. The numerical results are supported by rigorous mathematical proofs.

The L^1 errors between the exact and approximate invariant densities of the Markov operator associated with Markov type position dependent random maps, defined from $[0,1]$ to itself, are calculated by using the piecewise linear polynomials maximum entropy method.

FINITE ELEMENT MAXIMUM ENTROPY METHOD FOR APPROXIMATING ABSOLUTELY CONTINUOUS INVARIANT MEASURES

by

Tulsi Upadhyay

A Dissertation
Submitted to the Graduate School
and the Department of Mathematics
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved:

Dr. Jiu Ding, Committee Chair
Professor, Mathematics

Dr. Ching-Shyang Chen, Committee Member
Professor, Mathematics

Dr. James V. Lambers, Committee Member
Associate Professor, Mathematics

Dr. Haiyan Tian, Committee Member
Associate Professor, Mathematics

Dr. Bernd Schroeder
Chair, Department of Mathematics

Dr. Karen S. Coats
Dean of the Graduate School

August 2017

ACKNOWLEDGMENTS

I would like to extend my sincere thanks to my supervisor, Dr. Jiu Ding, whose patient guidance and insightful comments left an indelible mark on this dissertation. The completion of this dissertation is possible due to his help and academic guidance. I have had opportunities to talk with him on this research and on possible future research. I am thankful for his patient mentorship that has enhanced my ability to develop insightful thinking and essential concepts to complete the current research. I am sure his influence on my professional development will last throughout my career.

I also sincerely thanks to my dissertation committee members Dr. Ching-Shyang Chen, Dr. James Lambers, and Dr. Haiyan Tian. Their needful instructions have helped me to carry out the dissertation through to this form. I would like to thank department chair Dr. Bernd Schroeder for his kind and influential support during the study period. His readiness to help in any circumstances is greatly appreciated.

I am humbly thankful to my mother, Pabitra Upadhyay, and my late father, Sadakhar Upadhyay, who have influenced me to pursue a higher study. My wife, Shova Sharma Upadhyay, and son, Bishal Saptusho Upadhyay, have always supported and encouraged me during my study. Without their support, it would not be possible to complete my research and write this dissertation.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

ACIM	-	Absolutely continuous invariant measure
BV	-	Bounded variation
F-P	-	Frobenius-Perron
HMEM	-	Homogeneous maximum entropy method
MEM	-	Maximum entropy method
PLMEM	-	Piecewise linear maximum entropy method
PQMEM	-	Piecewise quadratic maximum entropy method
PQMFA	-	Piecewise quadratic Markov finite approximation

NOTATION AND GLOSSARY

General Usage and Terminology

Standard mathematical symbols are used in this dissertation. The blackboard fonts $\mathbb{R}$ and $\mathbb{C}$ are used to denote sets of real and complex numbers respectively, and $\mathbb{R}^n$ is used for an n-dimensional Euclidean space. The symbols X and Y are used to denote measure spaces, upper case letters with subscripts $S_1,\ S_2$, etc are used to denote maps and the lowercase letters $f,\ g$, etc are used to write functions.

Chapter 1

Introduction

Studying the asymptotic behavior of trajectories defined by a transformation is the main concern of discrete dynamical systems. However, it is not always possible to describe the asymptotic behavior of trajectories except for certain simple dynamical systems. Especially, for a chaotic dynamical system, the behavior of trajectories is unpredictable for most initial points. Therefore, it is natural to try to find ways to describe the asymptotic behavior of trajectories.

Statistical methods have been developed to describe the behavior of a system as a whole. In brief, these methods have been applied to prove the existence of invariant measures that are absolutely continuous with respect to the Lebesgue measure. Absolutely continuous invariant measures (abbreviated as ACIMs) are important physical measures in the statistical study of dynamical systems because the computer simulation of trajectories of a system reveals only the measures which are absolutely continuous with respect to the Lebesgue measure [36].To know about the ACIM, one needs definitions of some basic mathematical terms from measure theory.

Definition 1.0.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, where X is a set, $\mathcal{A}$ is a σ-algebra of subsets of X, and μ is a measure defined on $\mathcal{A}$. Let $S : X \to X$ be a measurable transformation. The measure μ is said to be invariant under the transformation S or S is invariant with respect to the measure μ if $\mu(S^{-1}(A)) = \mu(A)$ for every $A \in \mathcal{A}$. In this case, the transformation S is also called a measure preserving transformation. When $\mu(X) = 1$, the measure space $(X, \mathcal{A}, \mu)$ is called a probability space.

The following are two examples of measure preserving transformations.

Example 1.0.2. Let μ be the Lebesgue measure defined on $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$. For any fixed $\beta \in \mathbb{R}$, the transformation $T : \mathbb{S}^1 \to \mathbb{S}^1$ defined by $T(z) = e^{2\pi i \beta} z$ is measure preserving.

Example 1.0.3. (Baker's transformation) Let μ be the Lebesgue measure defined on the space $[0,1] \times [0,1]$. Then the transformation

$$T(x,y) = \begin{cases} (2x, \frac{1}{2}y), & \text{if } x \leq \frac{1}{2}, \\ (2x-1, \frac{1}{2}y + \frac{1}{2}), & \text{if } x > \frac{1}{2} \end{cases}$$

preserves μ.

Lemma 1.0.4. For a measure preserving transformation $T : X \to X$, the following are equivalent:
a) T is measure preserving;
b) for every $f \in L^1(X, \mathcal{A}, \mu)$,

$$\int_X f d\mu = \int_X f \circ T d\mu.$$

If one does not impose additional requirements for an invariant measure μ of S, the measure μ might be trivial and has no physical importance in applications. For instance, $x = 0$ is a fixed point of the logistic map $S : [0,1] \to [0,1]$ defined by $S(x) = 4x(1-x)$. It can be shown that the Dirac measure δ_0, defined by

$$\delta_0(A) = \begin{cases} 1, & \text{if } 0 \in A, \\ 0, & \text{if } 0 \notin A \end{cases},$$

concentrated at 0 and is an invariant measure. A Dirac measure concentrates on a fixed point $a \in [0,1]$ of the logistic map is an invariant measure, but it is not absolutely continuous with respect to the Lebesgue measure. Because of the lack of absolute continuity, this measure can not be represented by the integral of an integrable function.

The example of the Dirac measure shows that an invariant measure might not be physically significant. For that reason, it is necessary to find an invariant measure having the absolute continuity property with respect to the Lebesgue measure. An absolutely continuous measure can be defined in a number of alternative ways. A definition is given below.

Definition 1.0.5. Let μ and ν be two measures on a measurable space $(X, \mathcal{A})$. The measure μ is said to be absolutely continuous with respect to the measure ν if every ν-null set is also a μ-null set, in other words, for every measurable set A, $\mu(A) = 0$ whenever $\nu(A) = 0$. Generally, it is written as $\mu \ll \nu$ to denote that μ is absolutely continuous with respect to ν.

Note 1.0.6. If a measure on a real line is called an ACIM, it means the measure is absolutely continuous with respect to the Lebesgue measure.

Definition 1.0.7. Let $(X, \mathcal{A}, \mu)$ be a measure space. A measurable transformation $S : X \to X$ is a nonsingular transformation if $\mu(S^{-1}(A)) = 0$ for all $A \in \mathcal{A}$ such that $\mu(A) = 0$.

In this case, the measure μ is called a quasi-invariant measure for S. This assumption on S in Definition 1.0.7 is equivalent to the statement of the measure $\mu \circ S^{-1}$ being absolutely continuous with respect to the measure μ, where $\mu \circ S^{-1}(A) = \mu(S^{-1}(A))$ for all $A \in \mathcal{A}$.

An example of an absolutely continuous measure is:

Example 1.0.8. Let $(X, \mathcal{A}, \mu)$ be a measure space and f be a nonnegative integrable function. The measure defined by

$$\lambda(A) = \int_A f d\mu, \; A \in \mathcal{A}$$

satisfies

$$\mu(A) = 0 \Rightarrow \lambda(A) = 0.$$

Thus, λ is absolutely continuous with respect to the measure μ.

The importance of an invariant measure can be found in Poincaré's Recurrence Theorem.

Theorem 1.0.9. *(Poincaré's Recurrence Theorem) Let T be a measure preserving transformation of a probability space $(X, \mathcal{A}, \mu)$ and $A \subset X$ be a measurable set. For any $N \in \mathbb{N}$*

$$\mu\left(\left\{x \in X : \{T^n(x)\}_{n \geq N} \subset X \backslash A\right\}\right) = 0.$$

The theorem states that, for any $A \in \mathcal{A}$ in a probability space $(X, \mathcal{A}, \mu)$ with $\mu(A) > 0$, almost all points in A return to A infinitely many times under the iteration of T. The phrase "almost all points of A" refers to all points of A except for the set of points in A of μ-measure zero. The result of the theorem is possible because of the existence of an invariant measure.

Birkhoff's Ergodic Theorem [50] is an illustration which provides importance of invariant measures.

Theorem 1.0.10. *(Birkhoff's Ergodic Theorem) Let T be a measure preserving transformation on a σ-finite measure space $(X, \mathcal{A}, \mu)$, and let f be a μ-integrable function. Then there exists a μ-integrable function f^* such that*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = f^*(x)$$

for almost all $x \in X$, and $f^* \circ T = f^*$. Furthermore, if $\mu(X) < \infty$, then $\int f d\mu = \int f^* d\mu$.

When the measure μ is ergodic, that is, $T^{-1}(A) = A \Rightarrow$ either $\mu(A) = 0$, or $\mu(A^c) = 0$, then

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_A f d\mu$$

for almost all points x in X and all density functions f, namely f is nonnegative with integral 1. Theorem 1.0.10 gives the statistical nature of the trajectories of almost all points of X. Let μ be an invariant ergodic measure for the transformation $T : X \to X$, where X is a probability space. The asymptotic behavior of almost all points of X under the iteration of T is given by the asymptotic frequency of the visits of the trajectories of $x \in X$ in $A \in \mathcal{A}$. The relative frequency of the first n iterates of $x \in X$,

$$\{x, T(x), T^2(x), \ldots, T^{n-1}(x)\},$$

in A, is

$$\frac{1}{n}\sum_{i=0}^{n-1}\chi_A(T^i(x)),$$

where the characteristic function, χ_A, is defined by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}.$$

By Theorem 1.0.10

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\chi_A(T^i(x)) = \int_X \chi_A d\mu = \int_A d\mu = \mu(A).$$

The left side is called the time average, and the right side is called the space average. The result indicates that the trajectories of almost all points $x \in X$ lie in A with the probability $\mu(A)$. The theorem establishes a relation between two entirely different quantities; however, it does not give any information about the existence of the invariant measures.

The existence criterion of invariant measures for a certain class of transformations is given by the Krylov-Bogolubov Theorem:

Theorem 1.0.11. *[43] Any continuous transformation on a metrizable compact space has an invariant probability measure.*

Theorem 1.0.11 gives an existence criterion of an invariant measure of a certain class of transformations, but it does not give a way to find such a measure. The study of dynamical systems from the statistical viewpoint provides an approach to find the invariant measures of the systems. In this viewpoint, the study focuses on iterates of state space points by introducing a measure theoretical concept. The existence of an ACIM of a nonsingular transformation can be studied by using the Frobenius-Perron (abbreviated as F-P) operator, which connects an invariant density of the operator associated with a transformation with an ACIM of the transformation. More discussion of the F-P operator can be found in Chapter 2.

For a nonsingular transformation S, a function f is a density of an ACIM (say μ) if and only if f is an invariant density of P_S, that is

$$P_S f = f.$$

This equation establishes an existence criterion of ACIMs of nonsingular transformations. The concern of this study is to find such densities f. Unfortunately, solving the equation $P_S f = f$ for finding an invariant density f is difficult except in some simple cases.

To demonstrate how difficult it is to find an invariant density of the F-P operator even for a simple transformation, one can take the logistic map $S(x) = 4x(1-x)$. Let $A = [0,x] \subset [0,1]$.

As will be defined in the next chapter, the F-P operator corresponding to S maps an integrable function f to an integrable function $P_S f$ that satisfies

$$\int_0^x P_S f(t)\,dt = \int_{S^{-1}([0,x])} f(t)\,dt. \tag{1.1}$$

Differentiating (1.1) with respect to x gives

$$P_S f(x) = \frac{d}{dx}\int_{S^{-1}([0,x])} f(t)\,dt, \tag{1.2}$$

where,

$$S^{-1}[0,x] = \left[0, \frac{1}{2} - \frac{1}{2}\sqrt{1-x}\right] \cup \left[\frac{1}{2} + \frac{1}{2}\sqrt{1-x}\right].$$

By using the chain rule, (1.2) becomes

$$P_S f(x) = \frac{1}{4\sqrt{1-x}}\left[f\left(\frac{1}{2} - \frac{1}{2}\sqrt{1-x}\right) + f\left(\frac{1}{2} + \frac{1}{2}\sqrt{1-x}\right)\right].$$

This is the formula to find the F-P operator associated with the map S. This formula demonstrates how the operator changes a density function f into a new density function $P_S f$. For a uniformly distributed initial points, the density function is $f(x) \equiv 1$ and the density function $P_S f$ is

$$P_S f(x) = \frac{1}{2\sqrt{1-x}}.$$

The function sequence $\{P_S^n f\}$ then converges to the density

$$f^*(x) = \frac{1}{\pi\sqrt{x(1-x)}} \tag{1.3}$$

as $n \to \infty$. The graphs of the density functions $P_S f$ and f^* are presented respectively in Figure 1.1 and in Figure 1.2. The above example shows the difficulty in obtaining an invariant density of the F-P operator.

1.1 Literature Review

Numerical schemes have been developed to overcome the difficulties of obtaining invariant densities of the F-P operator. The main concern of a numerical scheme is to make a suitable conversion of the infinite dimensional F-P operator into a finite dimensional matrix. In this scheme, a nonnegative fixed point of the operator equation $P_S f = f$, which is called an invariant density, can be approximated by a fixed point of a corresponding finite dimensional matrix operator. Ulam [69] suggested a numerical scheme for approximating invariant densities of the F-P operator corresponding to one-dimensional nonsingular maps defined

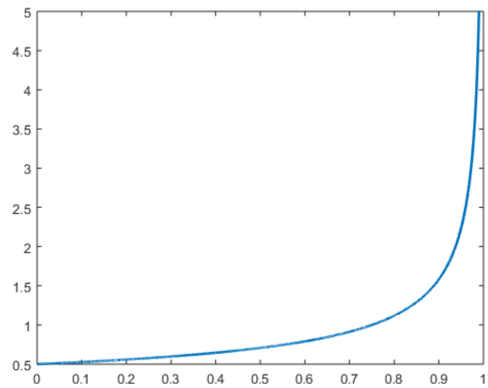

Figure 1.1: The Density Function $P_S f(x)$

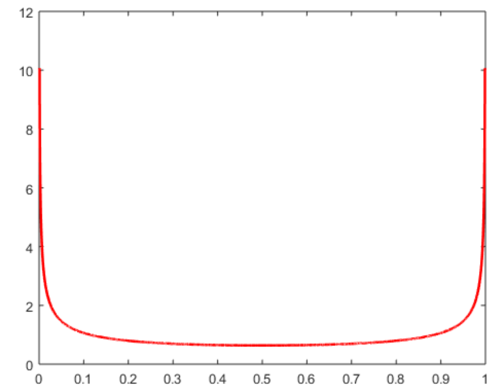

Figure 1.2: The Invariant Density $f^*(x)$

from $[0,1]$ to itself. The scheme was based on piecewise constant functions. Although the convergence of the scheme was slow, it served as a basis for the development of the more advanced numerical schemes. In the general case, the convergence of Ulam's numerical scheme is still an open question.

Li [51] proved Ulam's conjecture for a class of C^2 and stretching maps $S:[0,1]\to[0,1]$. Li's proof was based on the Lasota and Yorke [49] result for the existence of ACIMs for a class of piecewise C^2 and stretching maps. In the paper, Li gave a numerical procedure which can be implemented on a computer with negligible round-off errors. The paper showed how the original infinite-dimensional operator can be approximated by a finite dimensional operator (even though the original operator is not compact). It also gave a solution to Ulam's conjecture concerning a finite approximation for the F-P operator.

The result in [49] was generalized into higher dimensions, but the generalization took a long time due to the difficulty in defining bounded variation in higher dimensions. Analogous to Lasota and Yorke's proof, Jablonski proved the existence of ACIMs for piecewise C^2-transformations of the n-dimensional cube using the Tonelli definition of the functions of bounded variation [40]. The results for the existence of ACIMs of smooth maps on boundary-less domains were investigated in [45].

Rychlik [63] proved the existence of ACIMs in a more general setting for the monotonic piecewise functions, S, satisfying three specific conditions. The proof is considered to be a general proof because it did not depend on the principle of bounded variation. Under the given conditions, the sequence $\{P_S^n\mathbf{1}\}_{n=1}^{\infty}$ is bounded in L^∞ and the boundedness of this sequence shows a weak compactness of the sequence in L^1. Due to the weak compactness condition, any of the weak limit points of the sequence, according to the Kakutani-Yosida theorem, is a P_S invariant density.

Ding and Zhou [21] proved the convergence of Ulam's piecewise constant approximation algorithm for the computation of an absolutely continuous invariant measure associated with a piecewise C^2 expanding transformation or a Jablonski transformation $S : [0,1]^N \subset \mathbb{R}^N \to [0,1]^N$. It was an extension of Ulam's conjecture from a one dimensional space to a higher dimensional space.

Kohda and Murao [54] proposed a piecewise polynomial Galerkin approximation to invariant densities of the F-P operator associated with one-dimensional maps from $[0,1]$ to itself. The piecewise polynomial functions are linear combinations of $N \times K$ piecewise polynomial bases of degree $K-1$, where N is the number of subintervals of $[0,1]$. The approximate solution for $K = 1$ was identical to the Ulam-Li solution. The sufficient condition for the L^1 convergence of the approximate solution for $(K = 2,3)$ and a large N was given in the paper. The numerical results of the paper showed that the approximate solution for $K = 3$ is a significant improvement on the Ulam-Li solution, but the results were not supported by a rigorous convergence analysis.

Ulam's scheme was based on piecewise constant functions so its convergence rate was low. To improve the convergence rate, a piecewise linear Markov finite approximation method was proposed in [22]. The method has a better convergence rate than Ulam's method had. In the paper, Ulam's piecewise constant approximation was extended to a higher dimensional space and it was established that the method is a first-order method.

A piecewise linear least squares method was proposed in [27] for the F-P operator and the norm convergence of the method was also proved. Linear algebra arguments were used to prove L^1-norm and BV-norm convergence for computing invariant densities of the F-P operator associated with piecewise C^2 and stretching maps defined from $[0,1]$ to itself. The paper also gave a theoretical proof of the convergence rate under the BV-norm. A standard hat function was used to define a basis for the vector space of continuous piecewise linear functions. The proof of the BV-norm convergence was based on L^1-norm convergence, on the Lasota-Yorke inequality, and on the local convergence of the BV-norm. According to the numerical results, the L^1-norm convergence rate was close to the order 2.

In the paper [14] Bose and Murray investigated how the exact rate of convergence in Ulam's method can not be better than $O(\frac{\log n}{n})$, where n is the number of subintervals in the discretization. The results of the paper showed that the conjectured rate of $O(\frac{1}{n})$ cannot be obtained even for extremely regular maps.

The maximum entropy method (abbreviated as MEM) was first introduced in [20] for solving the F-P operator equation. The paper developed a foundation to use the MEM in the numerical approximation to find an invariant density of the F-P operator. Especially, the author of the paper used fundamental relations between the F-P operator and the Koopman operator to develop a general MEM. The rigorous theoretical work, which linked the MEM with this type of approximation scheme, was the main outcome of the paper. The proof of the convergence rate of the proposed method was based on the general convergence analysis of the best entropy estimate. The results were better in comparison to the results from Ulam's method. The results of the paper, obtained by the implementation of monomials, can be improved by the use of higher order piecewise polynomials, orthogonal polynomials, or splines. Newton's iteration method with three-node Gaussian quadrature was used to obtain the numerical results. The paper also showed that the sequence $\{f_N\}$ of the maximum entropy solution converges, at least weakly, to an invariant density f^*.

Ding and Mead [24] described how to compute both the invariant densities of the Markov operator in stochastic analysis and the F-P operator in ergodic theory of chaotic dynamics using the MEM. The authors used a high precision Gaussian quadrature to approximate an invariant density f^* of the Markov operator and they also applied the method to the Markov operator $P : L^1(0,1) \to L^1(0,1)$ with a stochastic kernel

$$Pf(x) = \int_0^1 \frac{ye^{xy}}{e^y - 1} f(y)\, dy.$$

The unique invariant density f^* of P is

$$f^*(x) = \frac{e^x - 1}{ax},$$

where $a \equiv \int_0^1 \frac{e^x-1}{x}\, dx \simeq 1.317902151$. In the paper, the authors defined the dual operator P^* of the integral operator P to find P^*x^n explicitly. The dual operator

$$P^*g(y) = \int_0^1 \frac{ye^{xy}}{e^y - 1} g(x)\, dx = \frac{y}{e^y - 1} \int_0^1 e^{xy} g(x)\, dx$$

gave the explicit expression for P^*x^n as

$$P^*x^n = \frac{x}{e^x - 1} \int_0^1 e^{xt} t^n\, dt = \frac{e^x \sum_{j=0}^{n} (-1)^j \frac{n!}{(n-j)} x^{n-j} + (-1)^{n+1} n!}{x^n (e^x - 1)}.$$

The paper also provided approximations of the invariant density of the F-P operator P_S, where S is a nonsingular transformation $S : [0,1]$ to itself, and the Lyapunov exponent corresponding to S. Newton's iteration method, to a high precision (up to 15 significant digits), was used to solve the nonlinear system,

$$\int_0^1 [I - P^*] \exp \sum_{n=1}^{N} \lambda_n (I - P^*) x^n\, dx = 0,\ i = 1, 2, \ldots, N,$$

to obtain the unique solution,

$$f_N(x) = \frac{\exp\sum_{n=1}^{N}\lambda_n(I-P^*)x^n}{\int_0^1 \exp\sum_{n=1}^{N}\lambda_n(I-P^*)x^n\,dx},$$

of the maximum entropy problem

$$\max\left\{H(f) : \int_0^1 f(x)(I-P^*)x^n\,dx = 0,\ 1\le n\le N\right\}.$$

To compute the numerical value of the integration, a 50-node Gaussian quadrature was used.

Ding, Jin, Rhee, and Zhou [25] developed the MEM based on piecewise linear polynomials for the recovery of an invariant density of the F-P operator. A set of piecewise linear polynomials, $\{\phi_i, i = 0,1,2,\ldots,n\}$, are used as moment functions. The set forms a canonical basis for the space of all continuous piecewise linear functions and satisfies the partition of unity property, i.e.,

$$\sum_{i=0}^{n}\phi_i(x) = 1,\ \forall x\in[0,1].$$

The partition of unity property of the moment functions plays an important role in defining a modified MEM. The modified MEM removed the singularity problem presented in the traditional MEM. The modified method also simplified the computational problem by converting the Jacobian matrix of the nonlinear system

$$\int_0^1 \phi_i(x)e^{\sum_{k=0}^{n}\lambda_k\phi_k(x)}\,dx = m_i,\ i = 0,1,\ldots,n,$$

into a positive definite and tri-diagonal matrix. The modified method was more efficient than the traditional MEM and produced more accurate results than the results obtained from the Markov finite approximation method and traditional maximum. This paper [25] is the first paper to lay a foundation for the possible extension of the MEM by using different moment functions such as higher order piecewise polynomials and spline functions to approximate invariant densities of the F-P operator.

1.2 Organization of the Dissertation

The dissertation extends the results of the paper [25] by applying piecewise quadratic polynomials and quadratic splines in the MEM for approximating invariant densities of the F-P operator associated with nonsingular transformations defined from $[0,1]$ to itself. The L^1 errors between the approximate and exact invariant densities are calculated and supported by rigorous theoretical proofs. The efficiency of the methods can be judged by the time taken by the Matlab program, used for numerical calculation, to get the numerical results. Similarly, the accuracy of the methods can be found by comparing the errors given by this method with the errors produced by contemporary methods. The numerical results are given along with sufficient theoretical discussion. A brief study of position dependent random maps is included in this dissertation. A study to compare the errors between the exact and approximate invariant densities of the F-P operator associated with some Markov type random maps with position dependent probabilities has been done.

Chapter 2 contains definitions and introduction of entropy, maximum entropy. In the chapter, some examples are presented to show how a density function can be chosen, when a set of density functions satisfy same constraints, by maximizing Boltzmann entropy. The chapter also gives the definition and some properties of the F-P operator.

Chapter 3 consists of an approximation of invariant densities of the F-P operator associated with some nonsingular transformations defined from $[0,1]$ to itself. A piecewise quadratic MEM is presented using finite element method by dividing the interval $[0,1]$ into finite number of subintervals. How the scheme overcomes the singularity problems of classical MEMs is discussed there. The numerical results of the L^1 errors between the exact and approximate invariant densities are presented in tabular forms. A rigorous theory is developed to study the convergence analysis of the method and the numerical results follow the theoretical convergence rate which is of $O(h^3)$. The numerical results of the study are compared with the results from both the Piecewise Linear Maximum Entropy Method (PLMEM) [25] and of the Piecewise Quadratic Markov Finite Approximation Method (PQMFA) [18].

Chapter 4 presents the background required to understand splines in general and quadratic splines in particular and their applicability in the MEM. Theoretical detail of the convergence of the method is followed by the tabular presentations of L^1 errors between the exact and approximate invariant densities of the F-P operator associated with some nonsingular maps defined from $[0,1]$ to itself. In the tabular presentations, errors from this method are compared with erros from PQMFA [18], PLMEM [25], and Piecewise Quadratic Maximum Entropy Method (PQMEM) [70].

In Chapter 5 a brief introduction of position dependent random maps defined from $[0,1]$ to itself is given. A Piecewise Linear MEM is used to approximate invariant densities of the F-P operator associated with Markov random maps defined from $[0,1]$ to itself. The L^1 errors between exact and approximate invariant densities of the F-P operator are presented in tabular forms.

Chapter 6 contains both concluding comments and suggestions for future research.

Chapter 2

Terminology and Definition

2.1 Entropy

The concept of entropy was first introduced by Clausius. It gained its importance when L. Boltzmann used it in his pioneering work on the kinetic theory of gasses in 1866, and Shannon used it in his work on information theory in 1948.

In a physical system, entropy provides a measure of the amount of thermal energy that cannot be used to do work. In some other definitions of entropy, it is a measure of how evenly energy (or some analogous property) is distributed in a system.

In probability theory, the entropy of a random variable measures the uncertainty about the value that might be assumed by the variable. If a random variable X takes different values $x_1,\ldots,x_n$, and the probability of x_i is p_i, for $1 \leq i \leq n$, then the equation

$$S = -K\sum_i \left[P_i \log(P_i)\right], \tag{2.1}$$

gives a way of using entropy in probability theory. Here $\log u = \log_e u$.

Claude Shannon is recognized as the father of modern communication and information theory. In his work [64], he defined entropy as follows:

Suppose there is a set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$. From these known probabilities, the concern is to know which event will occur. Can a measure be found to quantify how much "choice" is involved in the selection of the event or how much uncertainty would exist in the outcome? The uncertainty is quantified by using a probability measure $H(p_1,\ldots,p_n)$.

The measure, $H(p_1,\ldots,p_n)$, should have the following properties:

1) H is continuous at p_i.

2) If all p_i are equal, that is $p_i = \frac{1}{n}, i = 1,2,\ldots,n$, then H is a monotonically increasing function of n. With equally likely events, there is more choice or uncertainty when there is a great number of events.

3) If a choice is broken down into two successive choices, the original H is the weighted sum of the individual values of H.

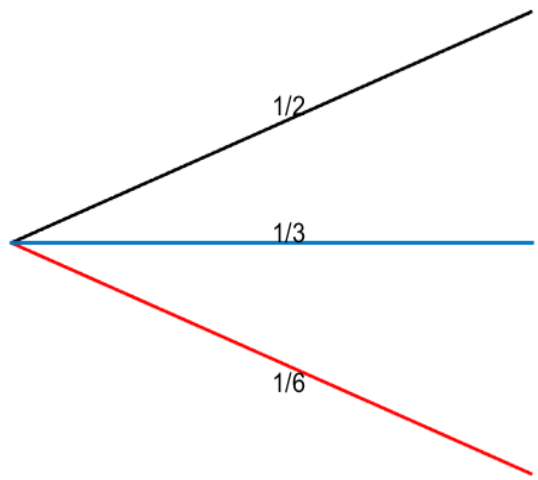

Figure 2.1: Three Possibilities

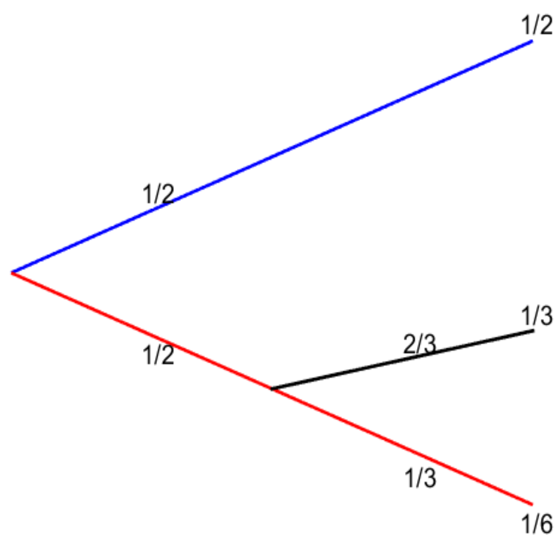

Figure 2.2: Decomposition of a Choice From Three Possibilities

In Figure 2.1, there are three probabilities: $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, and $p_3 = \frac{1}{6}$. In Figure 2.2, the first event between the two events, each with the probability $\frac{1}{2}$, is chosen. If the second event is chosen, the next choice would have the probabilities $\frac{2}{3}$ or $\frac{1}{3}$. The final result has the same probability as before. In this case,

$$H\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) = H\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2} H\left(\frac{2}{3}, \frac{1}{3}\right).$$

Shannon derived the information entropy in Appendix 2 of [64]. The derivation was based on the three facts given above and the entropy was the same as in (2.1) with $K = 1$. The entropy defined in this way is the entropy of a discrete set of probabilities $p_1, \dots, p_n$.

The mathematical expression, $-\sum p_i \log p_i$, which appears both in statistical mechanics and in information theory, does not establish any connection between these two fields [41].

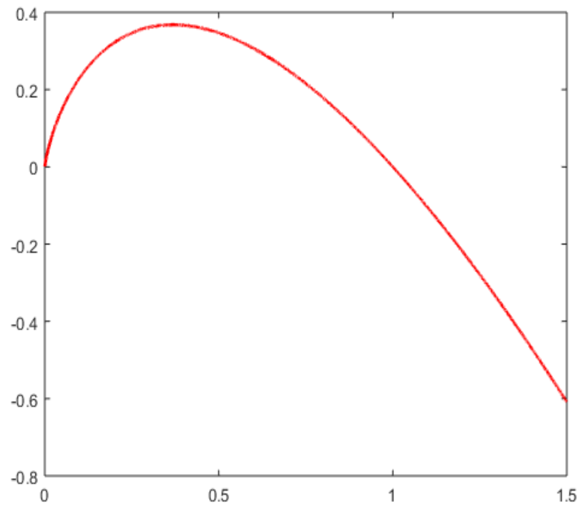

Figure 2.3: Graph of $\eta(u) = -u\log(u)$

Let η be the function defined by

$$\eta(u) = -u\log u, \; \eta(0) = 0.$$

Definition 2.1.1. If $f \geq 0$ and $\eta \circ f \in L^1$, then the entropy of f is defined as

$$H(f) = \int_X \eta(f(x))\,\mu(dx). \tag{2.2}$$

Remark 2.1.1. The integral (2.2) is always well defined for a function $f \geq 0$ when $\mu(X) < \infty$ and the integral is either finite or $-\infty$.

The function η is continuous for all $u \geq 0$. Since,

$$\eta''(u) = -\frac{1}{u},$$

the function $\eta(u)$ is concave for $u > 0$.

A line tangent to the graph of the function always lies above the graph, so

$$\begin{aligned}\eta(u) &\leq \eta(u_1) + \eta'(u_1)(u - u_1) \\ &= \eta(u_1) - (1 - \log u_1)(u - u_1) \\ &= -u + u_1 - u\log u_1 + u_1.\end{aligned}$$

By the definition of η

$$\begin{aligned} -u\log u &\leq -u+u_1-u\log u_1 \\ u-u\log u &\leq u_1-u\log u_1, \quad u,u_1>0. \end{aligned} \tag{2.3}$$

The inequality (2.3) is called the Gibbs inequality. The inequality is used to prove some important theorems and inequalities which are necessary for the later use.

Let f and g be two densities. If $\eta \circ f$ and $f \log g$ are integrable functions, then the Gibbs inequality (2.3) gives that

$$\begin{aligned} \int_X f(x)\,\mu(dx) - \int_X f(x)\log(f(x))\,\mu(dx) &\leq \int_X g(x)\,\mu(dx) - \int_X f(x)\log(g(x))\,\mu(dx) \\ -\int_X f(x)\log(f(x))\,\mu(dx) &\leq -\int_X f(x)\log(g(x))\,\mu(dx). \end{aligned} \tag{2.4}$$

The equality in (2.4) holds if and only if $f=g$.

The following theorem is an application of the inequality (2.4).

Theorem 2.1.1. *Let $(X,\mathcal{A},\mu)$ be a finite measure space and f be a density defined on X. The maximal entropy, $H(f)$ of f, attains for the constant density*

$$f_c(x) = \frac{1}{\mu(X)}.$$

Proof. Let f be a density in X. The entropy of f is

$$H(f) = -\int_X f(x)\log(f(x))\,\mu(dx).$$

The inequality (2.4) gives

$$\begin{aligned} H(f) &= -\int_X f(x)\log(f(x))\,\mu(dx) \\ &\leq -\int_X f(x)\log(f_c(x))\,\mu(dx) \\ &= -\log\left(\frac{1}{\mu(x)}\right). \end{aligned}$$

The entropy of $f_c(x)$ is

$$\begin{aligned} H(f_c(x)) &= -\int_X f_c(x)\log(f_c(x))\,\mu(dx) \\ &= -\int_X \frac{1}{\mu(x)}\log\left(\frac{1}{\mu(x)}\right)\mu(dx) \\ &= -\log\left(\frac{1}{\mu(x)}\right). \end{aligned}$$

Thus, $H(f) \leq H(f_c(x))$ for all densities f defined on X. $\square$

Remark 2.1.2. The theorem fails if $\mu(X) = \infty$ because there does not exist any constant density under this condition.

The following examples on the maximum entropy are from [50].

Example 2.1.2. Let $X = [0, \infty)$ and consider all possible densities f such that the first moment of f is given by

$$\int_0^\infty x f(x)\, dx = \frac{1}{\lambda}. \tag{2.5}$$

Then the density

$$f_\lambda(x) = \lambda e^{-\lambda x} \tag{2.6}$$

maximizes the entropy. From the inequality (2.4), for any density f that satisfies (2.5) holds

$$\begin{aligned} H(f) &= -\int_0^\infty f(x)\log(f(x))\, dx \\ &\leq -\int_0^\infty f(x)\log\left(\lambda e^{-\lambda x}\right)\, dx \\ &= -\log\lambda \int_0^\infty f(x)\, dx + \int_0^\infty \lambda x f(x)\, dx \\ &= -\log\lambda + 1. \end{aligned}$$

The entropy of the density function f_λ defined in (2.6) is

$$\begin{aligned} H(f_\lambda) &= -\int_0^\infty \lambda e^{-\lambda x}\log\left(\lambda e^{-\lambda x}\right)\, dx \\ &= -\log\lambda + 1. \end{aligned}$$

Thus $H(f) \leq H(f_\lambda)$ for all densities defined on X that satisfy (2.5).

In the following example, the second moment is used as a constraint for a density function defined on X.

Example 2.1.3. Consider all densities f defined on $X = (-\infty, \infty)$ such that the second moment of f is

$$\int_{-\infty}^\infty x^2 f(x)\, dx = \sigma^2. \tag{2.7}$$

Then the Gaussian density

$$f_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{x^2}{2\sigma^2}\right)}, \tag{2.8}$$

has maximum entropy. The entropy of a density function f defined on X is

$$\begin{aligned} H(f) &= \int_{-\infty}^{\infty} f(x)\log(f(x))\,dx \\ &\leq \int_{-\infty}^{\infty} f(x)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{\left(-\frac{x^2}{2\sigma^2}\right)}\right)dx \\ &= -\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\int_{-\infty}^{\infty} f(x)dx + \frac{1}{2\sigma^2}\int_{-\infty}^{\infty} x^2 f(x)\,dx \\ &= \frac{1}{2} - \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right). \end{aligned}$$

On the other hand,

$$\begin{aligned} H(f_\sigma(x)) &= -\int_{\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{\left(-\frac{x^2}{2\sigma^2}\right)}\right)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}e^{\left(-\frac{x^2}{2\sigma^2}\right)}\right)dx \\ &= \frac{1}{2} - \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right). \end{aligned}$$

It shows that $H(f) \leq H(f_\sigma)$ for all densities f satisfy (2.7). Thus the entropy is maximum for the Gaussian density (2.8).

2.2 Markov Operators

Markov operators are used in the study of dynamical systems and dynamical systems with stochastic perturbations [62]. In the statistical observation of dynamical systems, the evolution of a probability measure describing the distribution of points in a phase space is generally observed. This kind of observation is related to a transformation defined on the space of probability measures. The transformation defined in this way is linear.

Definition 2.2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. A linear operator

$$M : L^1(X, \mathcal{A}, \mu) \to L^1(X, \mathcal{A}, \mu)$$

is called a Markov operator if:

1) $Mf \geq 0$, for all $f \geq 0$;

2) $\int_X Mf\,d\mu = \int_X f\,d\mu$, for all $f \geq 0$.

Some properties of a Markov operator are given below.

Proposition 2.2.2. [50] Let $(X, \mathcal{A}, \mu)$ be a measure space and M be a Markov operator. Then for every $f \in L^1$,

1) $(Mf)^+ \leq Mf^+$,

2) $(Mf)^- \leq Mf^-$,

3) $|Mf| \leq M|f|$,

4) $\|Mf\| \leq \|f\|$,

5) If $f \leq g$, then $Mf \leq Mg$, for all $f, g \in L^1$.

The property 4) is called the contraction property of the operator.

2.3 Frobenius-Perron Operator

The F-P operator is a transfer operator named after Ferdinand Georg Frobenius and Oskar Perron, however; they did not propose the operator. Ulam gave the name because the operator shares some properties of nonnegative matrices. Those properties are studied by the classic Perron-Frobenius theorem for irreducible nonnegative matrices. In linear algebra, the F-P theorem says that a real square irreducible nonnegative matrix has a positive eigenvalue (spectral radius) which is greater than or equal to absolute value of other eigenvalues and the eigenvector corresponding to the eigenvalue has nonnegative components. This theorem has applications in probability theory, ergodic theory, economics [57] and others. In mathematics, the transfer operator generally provides information about the density evolution for an iterated map, so it is mostly used to study dynamical systems, statistical mechanics, chaos, and fractals.

According to [10], the transfer operator approach to the study of a dynamical system allows:

- the exploration of global dynamics and the characterization of global attractors,
- an estimation of invariant manifolds,
- the partitioning of the phase space into invariant regions, almost invariant regions, and coherent sets,
- an estimation of the rates of transport between these partitioned regions,
- the calculation of the decay of correlation.

The F-P operator associated with a transformation transfers a density function into another density function. A graphical representation of this transformation is given here. The density

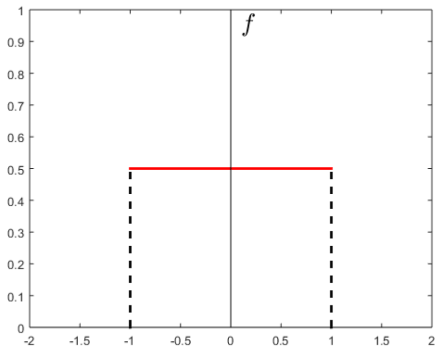

Figure 2.4: Density $f(x) = \frac{1}{2}\chi_{[-1,1]}(x)$

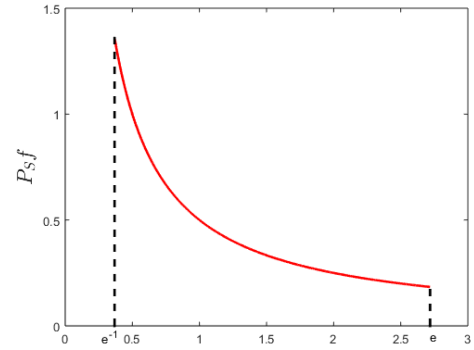

Figure 2.5: $P_S f(x) = (2x)^{-1}\chi_{[e^{-1},e^1]}(x)$

function in Figure 2.4 is transferred to the new density function in Figure 2.5 by the action of the F-P operator. Here the transformation is $S = e^x$, $x \in \mathbb{R}$ [47].

The use of this operator, in the measure-theoretic approach, can be described as [15]: Let $\{x_i\}$ be a collection of initial points in a phase space X. Consider a transformation $S : X \to X$ that describes the dynamics of a system and transfers each point $x_i \in X$ into a new point $S(x_i) \in X$. If the starting points are distributed according to a probability distribution function f, then the set of points $\{S(x_i)\}$ is distributed according to another probability distribution function, denoted by $P_S f$. The operator P_S is a linear operator known as the F-P operator. Now, instead of studying orbits of points in the phase space, it is easy and suitable to study the evolution (orbit) $\{f, P_S f, P_S^2 f, \ldots, P_S^n f, \ldots\}$ of the initial density f. If such a sequences converges, it generally converges to the density of a measure that is invariant under S and supported on the strange attractor. The advantage of studying P_S comes from the fact that, P_S is linear as well as bounded on $L^1(X)$ even if S is discontinuous on the phase space. Due to this fact, one can apply functional analysis tools to study the asymptotic behavior of dynamical systems.

Before giving a more general set up for the existence of the F-P operator, the following intuitive idea, from [50], is given to describe the operator.

Let $S : [0,1] \to [0,1]$ be a map. Let

$$\{x_1^0, x_2^0, x_3^0, \ldots, x_n^0\}$$

be a set of initial points of S. The set of initial points will change to a new set of points by an application of S. The new set of points is

$$\{x_1^1, x_2^1, x_3^1, \ldots, x_n^1\},$$

where $x_i^1 = S(x_i^0)$ for $i = 1, 2, \ldots, n$. The density function f_0 of the initial points $\{x_1^0, x_2^0, x_3^0, \ldots, x_n^0\}$

roughly says that, for any interval $I_0 \subset [0,1]$ (not too small),

$$\int_{I_0} f_0(u)\,du \simeq \frac{1}{n}\sum_{i=1}^{n} \chi_{I_0}(x_i^0). \tag{2.9}$$

Similarly, the new density function $f_1(x)$ for the set of points $\{x_1^1, x_2^1, x_3^1, \ldots, x_n^1\} \subset [0,1]$, gives

$$\int_I f_1(u)\,du \simeq \frac{1}{n}\sum_{i=1}^{n} \chi_{\Delta}(x_i^1). \tag{2.10}$$

The inverse image of an interval $I \subset [0,1]$ under S is

$$S^{-1}(I) = \{x \in [0,1] : S(x) \in I\}.$$

For any $I \subset [0,1]$,

$$x_i^1 \in I \text{ if and only if } x_i^0 \in S^{-1}(I).$$

It gives,

$$\chi_I(S(x)) = \chi_{S^{-1}(I)}(x). \tag{2.11}$$

Rewriting (2.11) using (2.10) gives

$$\int_I f_1(u)\,du \simeq \frac{1}{n}\sum_{i=1}^{n} \chi_{S^{-1}(I)}(x_i^0). \tag{2.12}$$

For a arbitrary subinterval I, define $I_0 = S^{-1}(I)$. This definition makes the right-hand side of (2.9) and (2.12) equal and

$$\int_I f_1(u)\,du = \int_{S^{-1}(I)} f_0(u)\,du. \tag{2.13}$$

The equation (2.13) shows a relation between f_0 and f_1 and says how a density function f_0 of initial states transfers into a density function f_1 of new states.

Taking $I = [a,x]$ and using the fact in (2.13) gives

$$\int_a^x f_1(u)\,du = \int_{S^{-1}([a,x])} f_0(u)\,du,$$

differentiating it with respect to x,

$$f_1(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f_0(u)\,du. \tag{2.14}$$

We write $f_1 = P_S f_0$, so (2.14) becomes

$$P_S f_0(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f_0(u)\,du.$$

The change of the arbitrary function f_0 into f_1 in (2.14) yields

$$Pf(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f(u)\,du. \tag{2.15}$$

The equation (2.15) explicitly defines the F-P operator.

The study of absolutely continuous invariant measures is an important component of study in the physical sciences. Many problems in the physical sciences are related to the study of existence and computation of densities of absolutely continuous invariant measures of nonsingular transformations. The study of absolutely continuous invariant measure in statistical physics, especially in the Boltzmann ergodic hypothesis in statistical physics, is related to the study of measure preserving transformations on measure spaces [7].

The Radon-Nikodym Theorem provides a rigorous basis for the existence of the F-P operator. Nonsingular transformations and absolutely continuous invariant measures are the main ingredients for defining the operator.

Theorem 2.3.1. *(Radon-Nikodym Theorem) Let λ be a finite measure and μ be a σ-finite measures on $(X,\mathcal{A})$. If $\lambda \ll \mu$, then there exists a μ-integrable function f such that*

$$\lambda(A) = \int_A f(x)\,\mu(dx),\ A \in \mathcal{A}. \tag{2.16}$$

The function f is unique μ-a.e. It means that if $\lambda(A) = \int_A g d\mu$ for all $A \in \mathcal{A}$, then $f(x) = g(x)$ μ-a.e.

Let $S : X \to X$ be a nonsingular transformation. For a given $f \in L^1(X)$ define

$$\mu_f(A) = \int_{S^{-1}(A)} f(x)\,\mu(dx),\ A \in \mathcal{A}.$$

S is nonsingular, so $\mu(A) = 0$ implies that $\mu_f(A) = 0$. By the Radon-Nikodyn Theorem 2.3.1 there exists a function $P_S f \in L^1(X)$ such that

$$\mu_f(A) = \int_A P_S f(x)\,\mu(dx),\ A \in \mathcal{A}.$$

The operator $P_S : L^1 \to L^1$ defined by

$$\int_A P_S f(x)\,\mu(dx) = \int_{S^{-1}(A)} f(x)\,\mu(dx),\ A \in \mathcal{A}, \tag{2.17}$$

is called the F-P operator associated with S.

From the above definition, it is clear that the F-P operator is a linear operator with norm 1.

The F-P operator P_S has the following properties:

Proposition 2.3.2.

(i) $P_S(\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n) = \lambda_1 P_S f_1 + \lambda_2 P_S f_2 + \cdots + \lambda_n P_S f_n$, for all $f_1, f_2, \ldots, f_n \in L^1$, and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$.

(ii) $P_S f \geq 0$ if $f \geq 0$; and

(iii) $\int_X P_S f(x)\, \mu(dx) = \int_X f(x)\, \mu(dx)$

(iv) If $S_n = S \circ S \circ \cdots \circ S$ and P_{S_n} is the F-P operator associated with S_n, then $P_{Sn} = P_S^n$.

(v) Let f be a density function. Then the measure μ_f defined by

$$\mu_f(A) = \int_A f\, dm,$$

for all measurable subsets A, is absolutely continuous with respect to the Lebesgue measure m. The measure μ_f is invariant under S if and only if f is a fixed point of P_S, that is $P_S f = f$.

Proof. The proofs of $(i) - (iii)$ are straightforward from the definition. The proof of (iv) is given below. Let S and T be two nonsingular transformations. By the definition of the F-P operator, we have

$$\begin{aligned} \int_A P_{T \circ S} f(x)\, \mu(dx) &= \int_{(T \circ S)^{-1}(A)} f(x)\, \mu(dx) \\ &= \int_{T^{-1}(A)} P_S f(x)\, \mu(dx) \\ &= \int_A P_T P_S f(x)\, \mu(dx). \end{aligned}$$

The required result is obtained by generalizing this process to a composite of n nonsingular transformations S. The proof of the first part of (v) comes from the definition of absolutely continuous invariant measure. For the proof of the second part, we proceed as follows:
If the measure μ is invariant under S, that is $\mu(S^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{A}$, then

$$\begin{aligned} \int_A f dm = \mu_f(A) &= \int_{S^{-1}(A)} f\, dm \\ &= \int_A P_S f\, dm, \end{aligned}$$

since $A \in \mathcal{A}$ is arbitrary, thus $f = P_S f$.

Conversely, let $P_S f = f$.

$$\begin{aligned} \mu_f(A) &= \int_A f\, dm \\ &= \int_A P_S f\, dm, \end{aligned}$$

which is true for every $A \in \mathcal{A}$, hence $\mu_f(A) = \mu_f(S^{-1}(A))$. $\square$

Remark 2.3.1. The F-P operator is not one-to-one in general. For example, let S be the symmetric triangle transformation on $[0,1]$. Take

$$f(x) = \begin{cases} 1, & \text{if } 0 \leq x < \frac{1}{2}, \\ -1, & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Then

$$\begin{aligned} P_S f(x) &= \frac{d}{dx}\left[\int_0^{\frac{1}{2}} \mu(dx) + \int_{\frac{1}{2}}^{1} -1\,\mu(dx)\right] \\ &= 0, \end{aligned}$$

which shows that P_S is not one-to-one.

A large amount of literature in dynamical systems concerns the study of absolutely continuous invariant measures. The invariant densities of the F-P operators associated with the nonsingular transformations are used to find the absolutely continuous invariant measures. The existence criteria of invariant densities of the F-P operators are considered important topics in the literature. The iterates of the F-P operator P_S starting from an initial density function f_0 produce a sequence $\{f_n = P_S^n f_0\}$. If the sequence $\{f_n\}$ converges to a density f^*, then f^* is an invariant density of P_S. In this formulation, the convergence of $\{f_n = P_S^n f_0\}$ is in question. When does the sequence $\{f_n = P_S^n f_0\}$ converge?

To answer the above question, the question can be divided into two parts. In the first part P_S is a finite dimensional, and in the second part, P_S is an infinite dimensional. For the finite and infinite dimensional cases, we use completely different approach to prove the convergence.

When P_S is a finite dimensional and has an irreducible matrix representation under a density basis, if 1 is the unique eigenvalue with absolute value 1, then the sequence $\{f_n = P_S^n f_0\}$ converges to f^*.

Bounded and closed sets of the L^1 space may not be compact or weakly compact.When P_S is an infinite dimensional, the convergence of the sequence $\{f_n = P_S^n f_0\}$ can be guaranteed by the Kakutani-Yosida Theorem, under certain conditions.

Theorem 2.3.3. *[15] (Kakutani-Yosida Theorem)*

Let $P : L^1 \to L^1$ be a Markov operator, and let f be a density function. If the Césaro sum sequence from the iterates $P^k f$,

$$\left\{ \frac{1}{n} \sum_{k=0}^{n-1} P^k f \right\}$$

is weakly precompact, i.e., it contains a weakly convergent subsequence, then

$$\frac{1}{n}\sum_{k=0}^{n-1} P^k f \to f^*$$

under the L^1- norm, and $Pf^* = f^*$, i.e., f^* is a fixed density of P.

2.3.1 The Frobenius-Perron Operators for a Certain Class of Transformations

When the transformation $S : [0,1] \to [0,1]$ is bijective, the F-P operator associated with S is defined by

$$P_S f(x) = f(S^{-1}(x))|(S^{-1})(x)|', \tag{2.18}$$

which is obtained from (2.17) by using the fundamental theorem of integral calculus If there exists a partition $0 = a_0 < a_1 < a_2 < \ldots < a_n = 1$ of $[0,1]$ such that S is bijective and differentiable in each subinterval $[a_{i-1}, a_i],\ i = 1, 2, \ldots, n$, one-sided limits being counted at the end points, then the F-P operator associated with S is

$$P_S f(x) = \sum_{i=1}^{n} f(S_i^{-1}(x))|(S_i^{-1})(x)|' \ f \in L^1(0,1). \tag{2.19}$$

To compute the F-P operator numerically, sometimes it is convenient to write the operator using the delta function

$$\delta(x) = \begin{cases} \infty, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

It satisfies the property

$$\int_a^b f(x)\delta(x)\,dx = \begin{cases} 1, & \text{if } 0 \in [a,b] \\ 0, & \text{if } 0 \notin [a,b] \end{cases} .$$

Using the basic property

$$\int_{-\infty}^{\infty} \delta(x-y) f(x)\,dx = f(y),$$

where f is any function with compact support and y any real number, the F-P operator associated to S can be written in the form

$$P_S f(x) = \int_0^1 \delta(x - S(y)) f(y)\,dy,\ f \in L^1(0,1). \tag{2.20}$$

An equivalence relation can be established between the formulas (2.17) and (2.20) when S is bijective and differentiable on the interval $[0,1]$ (differentiable at the end points means the existence of one-sided derivative at those points).

$$\begin{aligned} P_S f(x) &= \int_0^1 \delta(x-S(y))f(y)\,dy \\ &= \int_0^1 \delta(x-z)f(S^{-1}(z))(S^{-1})'(z)\,dz \\ &= f(S^{-1}(x))|(S^{-1})'(x)|, \end{aligned}$$

which is defined in (2.18).

When S is piecewise differentiable and bijective on each subinterval from the given partition, as described above, then

$$\begin{aligned} P_S f(x) &= \sum_{i=1}^{n} \int_0^1 \delta(x-S_i(y))f(y)\,dy \\ &= \sum_{i=1}^{n} \int_0^1 \delta(x-z)f(S_i^{-1}(z))(S_i^{-1})'(z)\,dz \\ &= \sum_{i=1}^{n} f(S^{-1}(x))|(S^{-1})'(x)|, \end{aligned}$$

which is the equation (2.19).

It is not so easy to find an expression for the F-P operator for general maps; however, for a certain class of maps, it is possible to express the F-P operator in a precise form. The F-P. operators corresponding to a class of piecewise monotonic maps, having certain properties, have nice representations in matrix forms. The definition of a class of piecewise monotonic and expanding maps is given before discussing the expressions of the F-P operators associated with those maps.

Definition 2.3.4. [15] Let $I = [a,b]$ be an interval. A transformation S defined from I to itself is called a piecewise monotonic transformation if there exists a partition

$$\mathcal{P} = \{I_i = (a_{i-1}, a_i) : i = 1, \ldots, q\}$$

of I and a number p such that
i) $S|_{I_i}$ is a C^p function, $i = 1, 2, \ldots, q$, and
ii) $|S'(x)| > 0$ on $I_i,\ i = 1, 2, \ldots, q$.

If S is a piecewise monotonic on I and $f \in L^1$, then the representation of the F-P operator associated with S is

$$P_S f(x) = \sum_{z \in \{S^{-1}(x)\}} \frac{f(z)}{|S'(z)|}. \tag{2.21}$$

For any x, the set $\{S^{-1}(x)\}$ consists of at most q points.

The representation (2.21) is obtained by the following way:
The definition of the F-P operator gives

$$\int_A P_S f(x) dx = \int_{S^{-1}(A)} f(x)\, dx, \tag{2.22}$$

for any measurable $A \subset I$.

As S is monotonic on each $I_i,\ i = 1,2,\ldots,q$, it is possible to define inverse function for each $S|_{I_i}$. Let $J_i = S(\overline{I_i}),\ i = 1,2,\ldots,q$. Define

$$\phi_i = S^{-1}|_{J_i}.$$

Now, $\phi_i : J_i \to \overline{I_j}$ and

$$S^{-1}(A) = \cup_{i=1}^{q} \phi_i(J_i \cap A), \tag{2.23}$$

so the sets $\phi_i(J_i \cap A)$ are mutually disjoint. Substituting (2.23) in ((2.22)) yields

$$\begin{aligned} \int_A P_S f(x)\, dx &= \sum_{i=1}^{q} \int_{\phi_i(J_i \cap A)} f(x)\, dx \\ &= \sum_{i=1}^{q} \int_{J_i \cap A} f(\phi_i(x)) |\phi_i'(x)|\, dx, \end{aligned}$$

by the change of variable formula.

$$\begin{aligned} \int_A P_S f(x)\, dx &= \sum_{i=1}^{q} \int_A f(\phi_i(x)) |\phi_i'(x)| \chi_{J_i}(x)\, dx \\ &= \int_A \sum_{i=1}^{q} \frac{f(S_i^{-1}(x))}{|S'(S_i^{-1}(x))|} \chi_{S(I_i)}(x)\, dx,\ S_i = S|_{I_i}. \end{aligned}$$

Since A is arbitrary,

$$P_S f(x) = \sum_{i=1}^{q} \frac{f(S_i^{-1}(x))}{|S'(S_i^{-1}(x))|} \chi_{S(I_i)}(x).$$

The more compact form is (2.21).

Example 2.3.5. Let $S : [0,1] \to [0,1]$ be defined by

$$S(x) = \begin{cases} -2x+1,\ x \in [0, \frac{1}{2}], \\ \\ 2x-1,\ x \in (\frac{1}{2}, 1]. \end{cases}$$

We find the representation of the F-P operator associated with S as follows.
The transformation S is monotonic on the intervals $I_1 = (0, \frac{1}{2}),\ I_2 = (\frac{1}{2}, 1)$. As above

$$P_S f(x) = \sum_{i=1}^{2} f(\phi_i(x)) |\phi_i'(x)| \chi_{S(I_i)}(x).$$

Here,

$$\begin{aligned}\phi_1(x) &= \frac{1-x}{2},\ \phi_1'(x) = -\frac{1}{2},\ \text{and}\\ \phi_2(x) &= \frac{1+x}{2},\ \phi_2'(x) = \frac{1}{2}.\end{aligned}$$

Hence,

$$P_S f(x) = \frac{1}{2}\left(f\left(\frac{1-x}{2}\right) + f\left(\frac{1+x}{2}\right)\right),$$

where $f \in L^1(0,1)$.

2.3.2 Matrix Representation of the Frobenius-Perron Operator

The F-P operator associated with a transformation S has a matrix representation when S is a piecewise linear Markov transformation.

Definition 2.3.6. Let $\mathcal{P} = \left\{I_i = (a_{i-1}, a_i) : i = 1, \ldots, q,\ a_0 = a,\ a_q = b\right\}$ be a partition of $I = [a,b]$. A transformation

$$S : I \to I$$

is said to be a Markov transformation if $S_i \equiv S|_{I_i}, i = 1, 2, \ldots, q$, is a homeomorphism from I_i onto some connected union of intervals of $\mathcal{P}$. The partition is called a Markov partition with respect to S.

If $|S'(x)| > 0$ on each I_i, then S is called piecewise monotonic. If each S_i is linear on I_i, then S is called a piecewise linear Markov transformation.

Definition 2.3.7. Let $S : I \to I$ be a piecewise monotonic transformation and $\mathcal{P} = \{I_i\}_{i=1}^q$ be a partition of I. The incidence matrix $A_S = (a_{ij})_{1 \le i,j \le q}$ induced by S and $\mathcal{P}$ is defined by

$$a_{ij} = \begin{cases} 1, \text{ if } I_j \subset S(I_i), \\ \\ 0, \text{ otherwise.} \end{cases}$$

The matrix is called the $0-1$ matrix induced by S if the partition $\mathcal{P}$ is clear, and the matrix is called $0-1$ matrix induced by $\mathcal{P}$ is S is clear.

The representation of P_S turns into a matrix when S is a piecewise linear Markov transformation. If g is a piecewise constant function on the partition $\mathcal{P}$, then the representation of g is

$$g(x) = \sum_{i=1}^{q} c_i \chi_{I_i}(x),$$

where $c_1, c_2, \ldots, c_q$ are constants. In such a case the function can be represented by

$$g(x) = c^g = (c_1, c_2, \ldots, c_q)^T,$$

where T denotes the transpose.

Theorem 2.3.8. *[15] Let $S : I \to I$ be a piecewise linear Markov transformation on the partition $\mathcal{P} = \{I_i\}_{i=1}^{q}$. Then there exists an $q \times q$ matrix M_S such that*

$$P_S f = M_S^T c^g$$

for every piecewise constant function g and c^g is the column vector obtained from g.

The matrix M_S is of the form

$$M_S = (m_{ij})_{1 \le i,j \le q},$$

where

$$m_{ij} = \frac{a_{ij}}{|S'|} = \frac{\lambda(I_i \cap S^{-1}(I_j))}{\lambda(I_i)},\ 1 \le i,j \le q,$$

where $\lambda(B)$ is the measure of B and $A_S = (a_{ij})_{1 \le i,j \le q}$ is the incidence matrix induced by S and $\mathcal{P}$.

Remark 2.3.9. For a given S, the matrix M_S is unique, but the converse is not true. There may be piecewise Markov transformations, other than S, which induce M_S. For instance, the transformation S_i on I_i can be replaced by a linear transformation which has the same domain and range as S_i has but having slope $-S_i'$. In this way, we can find 2^n piecewise linear Markov transformations which induce the same matrix M_S.

Chapter 3

A Continuous Piecewise Quadratic Maximum Entropy Method

3.1 Piecewise Quadratic Polynomials in the Maximum Entropy Method

Moment problems can be found in spectra estimation, geophysics, radio astronomy, sonar and radar (see [44, 46] and references therein). They also arise in theoretical physics such as quantum spin systems, Ising models, and the divergent series obtained from Stieltjes transformations [56]. The moment problem has appeared in pure mathematics since Hausdorff [38].

In 1866, L. Boltzmann used the concept of entropy for the kinetic theory of gasses. The second law of thermodynamics states that in an isolated system, the entropy never decreases. Jaynes introduced the principle of maximum entropy in 1957. In his paper [41] he formulated the maximum entropy problem to recover a least biased density among all densities satisfying a finite number of given constraints. This concept has been widely used in different fields of science and engineering [56, 72].

The author of [20] developed a method for solving the fixed point equation of the F-P operators based on Jaynes' maximum entropy principle and it was extended to find the Lyapunov exponents of chaotic maps in [23]. A discrete version of the MEM for computing invariant densities of the F-P operator and the Lyapunov exponent using Boltzmann entropy functional can be found in [9], whereas its continuous version, based on orthogonal polynomials is found in [26].

The use of polynomials in the continuous version of MEM to the recovery of an invariant density leads to solving systems of nonlinear equations. The difficulty arises in solving such systems when they are in ill-condition. The difficulty can be resolved by taking piecewise polynomial functions instead of taking polynomial functions over the whole domain.

Piecewise constant polynomials to approximate the F-P operator in the MEM were used in [28], which was the first publication on MEM based on piecewise polynomials to recover invariant densities. The authors of [25] considered piecewise linear polynomials for the approximation of an invariant density of the F-P operator.

This research uses piecewise quadratic polynomials for solving the maximum entropy problem in order to numerically recover an invariant density of the F-P operator associated with a nonsingular transformation defined from $[0,1]$ to itself. The use of piecewise quadratic polynomials as the moment functions gives a system of nonlinear equations. The Jacobian matrix of the system is five-diagonal, nonsingular, and positive definite.

Gaussian quadrature of order 3 is used to find the numerical evaluations of the integrals. The three weights and three nodes are respectively $\frac{5}{9}$, $\frac{8}{9}$, and $\frac{5}{9}$; and $-\sqrt{\frac{3}{5}}$, 0, and $\sqrt{\frac{3}{5}}$. Newton's method is implemented to solve the system of nonlinear equations.

3.2 The Maximum Entropy Problem

Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. As defined in the previous chapter, the Boltzmann entropy of f is

$$H(f) = -\int_X f(x) \log f(x)\, \mu(dx). \tag{3.1}$$

Some properties of the Boltzmann entropy $H(f)$ described below can be found in [12, 50].
i) $H(f)$ is either finite or $-\infty$.
ii) H is a proper, upper semi continuous, concave functional and strictly concave on the set

$$\{f \in L(X,\mu), f \geq 0 : H(f) > -\infty\}.$$

iii) The level sets

$$\{f \in L^1(X,\mu) : f \geq 0, H(f) \geq \alpha\}$$

for all $\alpha \in \mathbb{R}$ are weakly compact.

In what follows, the set of all densities is denoted by D.

As proved in the previous chapter, for every $f, g \in D$, the Gibbs inequality implies that

$$-\int_X f(x) \ln f(x)\, \mu(dx) \leq -\int_X f(x) \ln g(x)\, \mu(dx). \tag{3.2}$$

Here, MEM is related to solving the following constrained optimization problem:

$$\max \left\{ H(f) : f \in D, \int_X f(x) g_i(x)\, \mu(dx) = m_i,\ 1 \leq i \leq r \right\}, \tag{3.3}$$

where $g_i \in L^\infty(X,\mu)$ are known moment functions and m_i are given moments for all $1 \leq i \leq r$. The inequality (3.2) is helpful to prove the following proposition.

Proposition 3.2.1. The solution of the constrained optimization problem (3.3) is

$$f_r(x) = \frac{e^{\sum_{i=1}^r \lambda_i g_i(x)}}{\int_X e^{\sum_{i=1}^r \lambda_i g_i(x)}\, \mu(dx)}, \tag{3.4}$$

where the numbers $\lambda_1, \lambda_2, \cdots, \lambda_r$ satisfy the relations

$$\int_X e^{\sum_{j=1}^r \lambda_j g_j(x)} g_i(x)\, \mu(dx) = m_i \int_X e^{\sum_{j=1}^r \lambda_j g_j(x)}\, \mu(dx). \tag{3.5}$$

Proof. See [50]. □

When the moment functions g_i are $\{1, x, x^2, \cdots, x^r\}$, the maximum entropy problem (3.3) with $X = [0,1]$ is the traditional maximum entropy problem where the moments m_i are given by

$$m_i = \int_0^1 x^i p(x)\, dx,$$

where p is a probability density.

3.3 Piecewise Quadratic Polynomials

A piecewise quadratic MEM for the numerical recovery of an invariant density of the F-P operator is proposed in this research. For this purpose, two types of piecewise quadratic polynomials as the moment functions are considered. The mathematical formulation of the piecewise polynomials is given in the following way:

Let $X = [0,1]$ be partitioned into n subintervals $I_j = [x_{j-1}, x_j],\ j = 1,2,\ldots,n$, of uniform length $h_j = x_j - x_{j-1} = h = 1/n$. On the set X, define the space

$$\mathcal{P} = \left\{\phi : \phi \in C^0(X), \phi|_{I_j} \in P_2(I_j), j = 1,2,\ldots,n\right\}, \tag{3.6}$$

where $C^0(X)$ denotes the space of continuous functions defined on X, and $P_2(I_j)$ the space of quadratic polynomials defined on I_j. The piecewise quadratic polynomials $\{\phi_k\}_{k=0}^{2n}$ are defined as

$$\phi_{2j}(x) = \tau\left(\frac{x - x_j}{h}\right),\ j = 0,1,\ldots,n,$$

and

$$\phi_{2j-1}(x) = \rho\left(\frac{x - x_{j-1}}{h}\right),\ j = 1,2,\ldots,n,$$

where

$$\tau(x) = \begin{cases} (x+1)^2, & -1 \le x \le 0, \\ (x-1)^2, & 0 \le x \le 1, \\ 0, & \text{otherwise,} \end{cases} \tag{3.7}$$

$$\rho(x) = \begin{cases} 2x(1-x), & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.8}$$

The set $\{\phi_0, \phi_1, \cdots, \phi_{2n}\}$ of polynomials forms a basis for the space $\mathcal{P}$ of continuous piecewise quadratic polynomials. It is a $2n+1$ dimensional vector subspace of $C^0(X)$. The support of ϕ_k is the set $\text{supp}(\phi_k) = \left\{x \in [0,1] : \phi_k(x) \neq 0\right\}$. So $\text{supp}(\phi_0) =$

$(0,h)$, $\text{supp}(\phi_{2n}) = (1-h,1)$, $\text{supp}(\phi_{2j}) = ((j-1)h,(j+1)h)$ for $j = 1,2,,\ldots,n-1$, and $\text{supp}(\phi_{2j-1}) = ((j-1)h, jh)$ for $j = 1,2,\ldots,n$. These polynomials satisfy the partition of unity property

$$\sum_{k=0}^{2n} \phi_k(x) \equiv 1, \; \forall\, x \in [0,1].$$

The partition of unity property plays an important role in modifying the maximum entropy problem in a relatively new form which leads to an easier numerical computation.

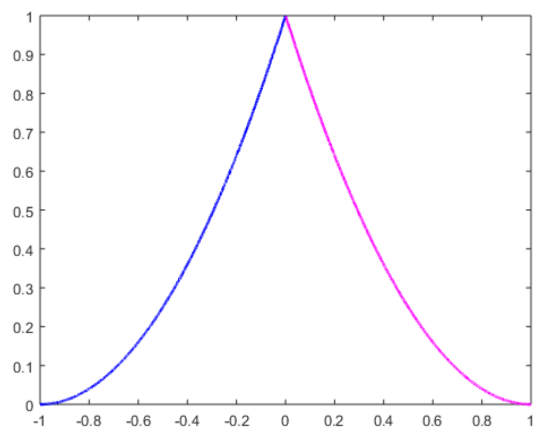

Figure 3.1: Graph of $\tau(x)$

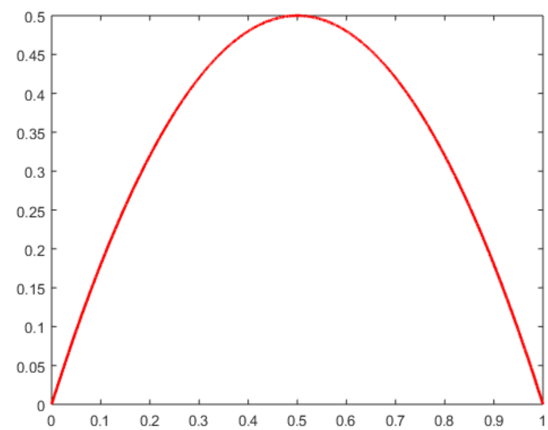

Figure 3.2: Graph of $\rho(x)$

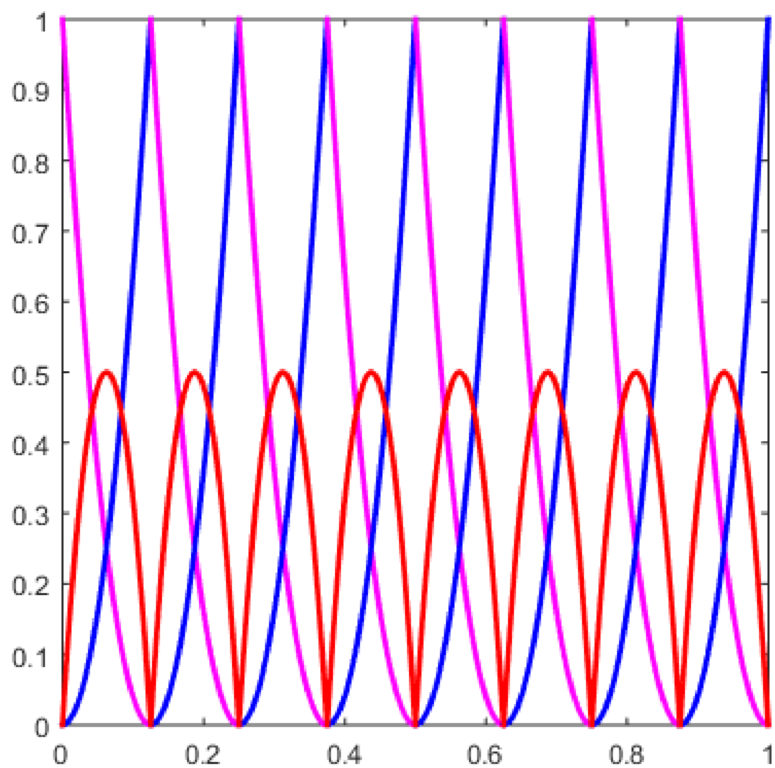

Figure 3.3: Piecewise Quadratic polynomials

For each $k = 0, 1, \ldots, 2n$, we define the kth moment, m_k, of an invariant density f^* with respect to the basis functions ϕ_k as

$$m_k = \int_0^1 f^*(x)\phi_k(x)\,dx. \tag{3.9}$$

Proposition 3.3.1. The sum of the moments m_k in (3.9) is unity.

Proof.

$$\text{Since } \sum_{k=0}^{2n} \phi_k(x) = 1 \text{ and } f^* \text{ is a density,}$$

$$\sum_{k=0}^{2n} m_k = \sum_{k=0}^{2n} \int_0^1 f^*(x)\phi_k(x)\,dx = \int_0^1 f^*(x) \sum_{k=0}^{2n} \phi_k(x)\,dx = \int_0^1 f^*(x)\,dx = 1.$$

□

3.4 Solution of the Maximum Entropy Problem by Piecewise Quadratic Polynomials

The Birkhoff individual ergodic theorem in [50] gives,

$$m_i = \lim_{M\to\infty} \frac{1}{M}\sum_{j=0}^{M-1} \phi_i(S^j(x)),\ \forall\, x \in [0,1]\ \mu_{f^*} - a.e.,\ i = 0,1,\cdots,2n, \tag{3.10}$$

so, for a large natural number N, the moments m_i are

$$m_i \simeq \frac{1}{N}\sum_{j=0}^{N-1} \phi_i(S^j(x)),\ \forall\, x \in [0,1]\ \mu_{f^*} - a.e.,\ i = 0,1,\cdots,2n.$$

Solving the maximum entropy problem (3.3), using the piecewise quadratic polynomials ϕ_i, requires solving the following problem:

$$\max\Big\{H(f) : f \in D, \int_0^1 f(x)\phi_i(x)\,dx = m_i, 0 \le i \le 2n\Big\}. \tag{3.11}$$

The solution of the problem is

$$f(x) = \frac{e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}}{\int_0^1 e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\,dx}, \tag{3.12}$$

where the Lagrange multipliers $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$ satisfy the relations

$$\int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\,dx = m_i \int_0^1 e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\,dx,\ i = 0,1,\ldots,2n. \tag{3.13}$$

Due to the partition of unity property of the basis functions ϕ_i, the Jacobian matrix $J(\lambda_0, \lambda_1, \ldots \lambda_{2n})$ of the equation (5.10) has left eigenvector $e_1 = (1,1,\ldots,1)$ corresponding to the eigenvalue 0, so it is a singular matrix.

To overcome this singularity problem, a modified method was developed in [25] by defining a new form of (5.10) based on the following Proposition:

Proposition 3.4.1. If $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$ satisfy

$$\int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\,dx = m_i,\ i = 0,1,\ldots,2n, \tag{3.14}$$

then $e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}$ is a density.

Proof. Since By Proposition 3.3.1, $\sum_{i=0}^{2n} m_i = 1$. Thus,

$$\begin{aligned}
\int_0^1 e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\,dx &= \int_0^1 e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)} \sum_{i=0}^{2n}\phi_i(x)\,dx \\
&= \sum_{i=0}^{2n}\int_0^1 e^{\sum_{k=0}^{2n}\lambda_k\phi_k(x)}\phi_i(x)\,dx \\
&= \sum_{i=0}^{2n} m_i = 1.
\end{aligned}$$

□

It can be easily seen that the numbers $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$, which are the solutions of the nonlinear system (3.14), are also solutions of (5.10).

Remark 3.4.2. Let λ=$(\lambda_0, \lambda_1, \ldots, \lambda_{2n}) \in \mathbb{R}^{2n+1}$ and $\mathbf{F} = (F_0, F_1, \ldots, F_{2n})^T$, where $F_i(\lambda)$=$\int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)}\, dx - m_i,\ i = 0, 1, \ldots, 2n$. The system (3.14) can be written as $\mathbf{F}(\lambda) = \mathbf{0}$.

Proposition 3.4.3. The Jacobian matrix $J(\lambda_0, \lambda_1, \ldots, \lambda_{2n})$ of the nonlinear system (3.14) is symmetric, nonsingular, five-diagonal, and positive definite.

Proof. The $(i, j)^{th}$ entry of the Jacobian matrix J at $(\lambda_0, \lambda_1, \ldots, \lambda_{2n})$ is

$$\begin{aligned} J_{i,j} &= \frac{\partial F_i}{\partial \lambda_j} = \frac{\partial}{\partial \lambda_j}\Big(\int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)}\, dx - m_i\Big) \\ &= \int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)} \phi_j(x)\, dx. \end{aligned}$$

The symmetry of the matrix is clear from the fact that $\phi_i(x)\phi_j(x) = \phi_j(x)\phi_i(x)$ for all $i, j = 0, 1, \ldots, 2n$. The supports of ϕ_i and ϕ_j are disjoint for $\mid i - j \mid > 2$ so the matrix J is five-diagonal. Furthermore, the $(i, j)^{th}$ entries of J_{ij} are zero when i and j are both odd and $\mid i - j \mid = 2$. $J_{ij} \neq 0$ for at least one pair of (i, j), otherwise $\phi_i(x) = 0$ for every $i = 0, 1, \ldots, 2n$ which is impossible.

Let $y^T = (y_0, y_1, \ldots, y_{2n})$ be a nonzero vector in $\mathbb{R}^{2n+1}$. Then

$$\begin{aligned} y^T J y &= \sum_{j=0}^{2n} \sum_{i=0}^{2n} y_j y_i J_{i,j} \\ &= \sum_{j=0}^{2n} \sum_{i=0}^{2n} y_j y_i \int_0^1 \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)} \phi_j(x)\, dx \\ &= \int_0^1 \sum_{j=0}^{2n} \sum_{i=0}^{2n} y_j y_i \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)} \phi_j(x)\, dx \\ &= \int_0^1 \sum_{i=0}^{2n} y_i \phi_i(x) e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)} \sum_{j=0}^{2n} y_j \phi_j(x)\, dx \\ &= \int_0^1 \Big(\sum_{i=0}^{2n} y_i \phi_i(x)\Big)^2 e^{\sum_{k=0}^{2n} \lambda_k \phi_k(x)}\, dx > 0. \end{aligned}$$

Thus J is positive definite and it completes the proof. □

The system (3.14) has a unique solution which can be shown exactly the same way as shown in [25].

In our algorithm, we calculate m_i using the formula (3.9) and use them in equation (3.14) to get $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$. By Proposition 3.4.1, the solution of the maximum entropy problem (3.11) is given by

$$f_n(x) = e^{\sum_{j=0}^{2n} \lambda_j \phi_j(x)}. \tag{3.15}$$

3.5 Convergence Analysis

In this section, we analyze the L^1 error of the proposed MEM. We provide the convergence rate and approximate error bounds of our method that concentrates on the numerical recovery of the unique invariant density f^* of the F-P operator P_S of a nonsingular transformation $S : [0,1] \to [0,1]$. The convergence analysis of the proposed numerical method is based on the convergence theory for the moment problem developed in [12, 13] and used in [25] .

Let X be a locally convex topological vector space with a nested sequence of closed subsets $\{G_n\}$ and let $K : X \to [-\infty, \infty)$ be a functional with compact level sets.

Lemma 3.5.1. Let f_n be an optimal solution of

$$\max\{K(f) : f \in G_n\},$$

and f_∞ the unique optimal solution of the limiting problem

$$\max\left\{K(f) : f \in \bigcap_{n=1}^{\infty} G_n\right\}$$

with $K(f_\infty) > -\infty$. Then $\lim_{n\to\infty} f_n = f_\infty$ under the topology of X and $\lim_{n\to\infty} K(f_n) = K(f_\infty)$.

Proof. See [13] (Propositions 1.4 and 1.5). □

To fit the result of Lemma 3.5.1 in our scheme, we use the following method developed in [25]. Divide the interval $[0,1]$ into two equal subintervals,divide one of those intervals into two subintervals, and keep this procedure for each step for each subinterval until getting n subintervals. On these subintervals, define a set of piecewise quadratic polynomials $\{\phi_0, \phi_1, \ldots, \phi_{2n}\}$. Taking these functions as basis functions, we have a sequence of subspaces Δ_n, of $C^0[0,1]$, of dimension $2n+1$, of piecewise quadratic polynomial. Since Δ_n is a nested increasing sequence of subspaces, the set $\{G_n\}$ of feasible solutions of the maximum entropy problem (3.11) is a monotonically decreasing sequence of sets of $L^1(0,1)$. The entropy functional H defined in the previous section has weakly compact level sets; it implies that the set of feasible solutions is weakly compact in $L^1(0,1)$. The space $L^1(0,1)$ is a locally convex topological vector space under the weak topology, so we have the following weak convergence result of our method:

Proposition 3.5.2. If the entropy $H(f^*) = -\int_0^1 f^* \ln f^* \, dx$ of f^* is finite, then the solution f_n given by (3.15) associated with Δ_n has the following properties
i) f_n converge weakly to f^*, that is
$\lim_{n\to\infty} \int_0^1 f_n(x)g(x)\,dx = \int_0^1 f^*(x)g(x)\,dx, \;\; \forall\, g \in L^\infty(0,1)$.
ii) $\lim_{n\to\infty} H(f_n) = H(f^*)$.

Proof. The result follows from Lemma 3.5.1 and Theorem 3.1 of [12]. □

Since the Boltzmann entropy functional H is Kadec, from Theorem 2.7 of [12], the weak convergence of Proposition 3.5.2 implies the strong convergence of the entropy solutions. Hence, Proposition 3.5.2 implies that $\lim_{n\to\infty} \|f_n - f^*\| = 0$.

Now the convergence rate analysis of the piecewise quadratic MEM is given. By the general convergence theory for the moment problem from [12, 13], it is enough to estimate the minimal distance of a continuous function g to the subspace of $C^0[0,1]$ spanned by $\phi_0, \phi_1, \ldots, \phi_{2n}$ under the maximum norm $\|g\|_\infty = \max\{|g(x)| : x \in [0,1]\}$. In other words, we want to estimate the quantity

$$\min_{(\lambda_0,\lambda_1,\ldots,\lambda_{2n})} \max_{x\in[0,1]} |g(x) - \sum_{k=0}^{2n} \lambda_k \phi_k(x)|.$$

For this purpose, note that the continuous piecewise quadratic polynomials $\phi_0, \phi_1, \ldots, \phi_{2n}$ have the expressions

$$\begin{aligned}
\phi_{2j}(x) &= \Big(\frac{x - jh}{h} - 1\Big)^2 = \Big(\frac{x - (j+1)h}{h}\Big)^2 \\
\phi_{2j+1}(x) &= 2\Big(\frac{x - jh}{h}\Big)\Big(1 - \frac{x - jh}{h}\Big) = 2\Big(\frac{x - jh}{h}\Big)\Big(\frac{(j+1)h - x}{h}\Big) \\
\phi_{2(j+1)}(x) &= \Big(\frac{x - (j+1)h}{h} + 1\Big)^2 = \Big(\frac{x - jh}{h}\Big)^2
\end{aligned}$$

on each subinterval $[x_j, x_{j+1}] = [jh, (j+1)h]$ of $[0,1]$. Our next lemma shows that if we set $\lambda_{2j} = g(x_j)$ for $j = 0,1,2,\ldots,n$, then the difference $|g(x) - \sum_{k=0}^{2n} \lambda_k \phi_k(x)|$ is of order $O(h^3)$ for a suitable choice of $\lambda_{2j+1},\; j = 0,1,2,\ldots,n-1$.

Lemma 3.5.3. Let $g \in C^3[0,1]$ and $p(x) = g(x_j)\phi_{2j}(x) + \mu_{2j+1}\phi_{2j+1}(x) + g(x_{j+1})\phi_{2(j+1)}(x)$, where $x \in [x_j, x_{j+1}] \subseteq [0,1], j = 0,1,\ldots,n-1$. If $\mu_{2j+1} = g(jh) + \frac{h}{2} g'(jh)$, then $g(x) - p(x) = \frac{1}{3!}\Big[g'''(\xi)(x - jh) - g'''(\eta)h\Big](x - jh)^2 = O(h^3)$, where $\xi \in (jh, x)$ and $\eta \in (jh, (j+1)h)$.

Proof. Using the expressions of ϕ_{2j}, ϕ_{2j+1}, and $\phi_{2(j+1)}$, we have

$$
\begin{aligned}
p(x) &= g(jh)\Big(\frac{x-(j+1)h}{h}\Big)^2 + 2\mu_{2j+1}\Big(\frac{x-jh}{h}\Big)\Big(\frac{(j+1)h-x}{h}\Big) + g((j+1)h)\Big(\frac{x-jh}{h}\Big)^2 \\
&= \frac{1}{h^2}\Big[g(jh)(x-(j+1)h)^2 + 2\mu_{2j+1}(x-jh)(jh-x+h) + g((j+1)h)(x-jh)^2\Big] \\
&= \frac{1}{h^2}\Big[g(jh)(x-jh-h)^2 + 2\mu_{2j+1}(x-jh)((j+1)h-x) + g((j+1)h)(x-jh)^2\Big] \\
&= \frac{1}{h^2}\Big[g(jh)\big\{(x-jh)^2 - 2h(x-jh) + h^2\big\} - 2\mu_{2j+1}\big\{(x-jh)^2 - h(x-jh)\big\} \\
&+ g((j+1)h)(x-jh)^2\Big] \\
&= \frac{1}{h^2}\Big[\big(g(jh) - 2\mu_{2j+1} + g((j+1)h)\big)(x-jh)^2 - 2h\big(g(jh)-\mu_{2j+1}\big)(x-jh) + h^2 g(jh)\Big] \\
&= \frac{1}{h^2}\Big(g(jh) - 2\mu_{2j+1} + g((j+1)h)(x-jh)^2\Big) - \frac{2}{h}\Big(g(jh)-\mu_{2j+1}\Big)(x-jh) + g(jh) \quad (3.16)
\end{aligned}
$$

Expanding $g(x)$ near $x = jh$ and using (3.16),

$$
\begin{aligned}
g(x)-p(x) &= g(jh) + g'(jh)(x-jh) + \frac{g''(jh)}{2!}(x-jh)^2 + \frac{g'''(\xi)}{3!}(x-jh)^3 - \frac{1}{h^2}\big(g(jh) - 2\mu_{2j+1} \\
&- g((j+1)h)(x-jh)^2\big) + \frac{2}{h}\big(g(jh)-\mu_{2j+1}\big)(x-jh) - g(jh), \text{where } \xi \in (jh,x) \\
&= \Big(g'(jh) + \frac{2}{h}(g(jh)-\mu_{2j+1})\Big)(x-jh) + \Big(\frac{g''(jh)}{2!} - \frac{g(jh)-2\mu_{2j+1}+g((j+1)h)}{h^2}\Big) \\
&(x-jh)^2 + \frac{g'''(\xi)}{3!}(x-jh)^3.
\end{aligned}
$$

Choose μ_{2j+1} such that $g'(jh) + \frac{2}{h}(g(jh)-\mu_{2j+1}) = 0$, that is $\mu_{2j+1} = \frac{h}{2}g'(jh) + g(jh)$. Then

$$
\begin{aligned}
g(x)-p(x) &= \Big(\frac{g''(jh)}{2!} - \frac{g(jh) - hg'(jh) - 2g(jh) + g((j+1)h)}{h^2}\Big)(x-jh)^2 + \frac{g'''(\xi)}{3!}(x-jh)^3 \\
&= \frac{g'''(\xi)}{3!}(x-jh)^3 - \frac{g'''(\eta)}{3!}h(x-jh)^2, \\
&= \frac{1}{3!}\Big[g'''(\xi)(x-jh) - g'''(\eta)h\Big](x-jh)^2 \\
&= O(h^3)
\end{aligned}
$$

where $\xi \in (jh,x)$, and $\eta \in (jh,(j+1)h)$. $\square$

Theorem 3.5.4. *Let $g \in C^3[0,1]$ and $\{\phi_k\}_{k=0}^{2n}$ be the basis of the space of continuous piecewise quadratic functions. Then*

$$\min_{\lambda_0,\lambda_1,\ldots,\lambda_{2n}} \max_{x\in[0,1]} |g(x) - \sum_{k=0}^{2n} \lambda_k \phi_k(x)| = O(h^3).$$

Proof. Apply Lemma 3.5.3 on each subinterval $[x_j, x_{j+1}],\ j = 0, 1, \ldots, n-1$. □

The combination of the above result with $g = \ln f^* \in C^3[0,1]$ and Theorem 4.7 of [12] gives the L^1 norm estimation as expressed in the following theorem:

Theorem 3.5.5. *Let $g = \ln f^* \in C^3[0,1]$. If f_n is defined as above, then $\|f^* - f_n\| = O(h^3)$.*

Proof. Let $E_n = \min\{\|\sum_{k=0}^{2n} \lambda_k \phi_k - \ln f^*\|_\infty : \lambda_k \in \mathbb{R}\}$. From Theorem 4.7 of [12], $\|f_n - f^*\| \le E_n e^{\frac{E_n}{2}}$, so $E_n = O(h^3)$ by Theorem 3.5.4. Since $e^{\frac{E_n}{2}} = O(1)$, it follows that $\|f_n - f^*\| = O(h^3)$. □

3.6 Numerical Results

In this section we present the numerical results of error analysis corresponding to the following five transformations from $[0,1]$ to itself.

$$
\begin{aligned}
S_1(x) &= \begin{cases} \frac{2x}{1-x^2}, & 0 \le x \le \sqrt{2}-1 \\ \frac{1-x^2}{2x}, & \sqrt{2}-1 \le x \le 1 \end{cases}, \\
S_2(x) &= \begin{cases} \frac{2x}{1-x}, & 0 \le x \le \frac{1}{3} \\ \frac{1-x}{2x}, & \frac{1}{3} \le x \le 1 \end{cases}, \\
S_3(x) &= 4x(1-x), \\
S_4(x) &= 1 - \sqrt{|2x-1|}, \\
S_5(x) &= \left(\frac{1}{8} - 2|x - \frac{1}{2}|^3\right)^{1/3}.
\end{aligned}
$$

The unique invariant densities of the F-P operator corresponding to the above transformations, $S_i,\ i = 1, \ldots, 5$, are respectively:

$$
\begin{aligned}
f_1^*(x) &= \frac{4}{\pi(1+x^2)}, \\
f_2^*(x) &= \frac{2}{(1+x)^2}, \\
f_3^*(x) &= \frac{1}{\pi\sqrt{x(1-x)}}, \\
f_4^*(x) &= 2(1-x), \\
f_5^*(x) &= 12\left(x - \frac{1}{2}\right)^2.
\end{aligned}
$$

The graphs of the nonsingular transformations $S_1, \ldots, S_5$, and the invariant densities of the F-P operator associated with those transformations are given below:

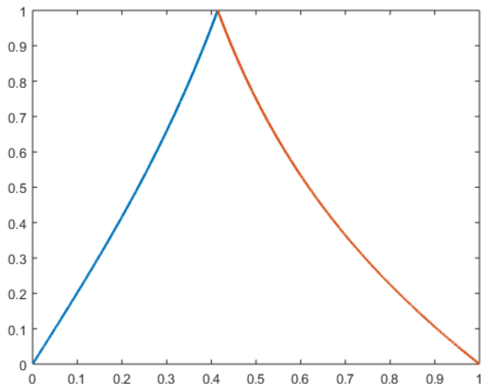

Figure 3.4: Graph of $S_1(x)$

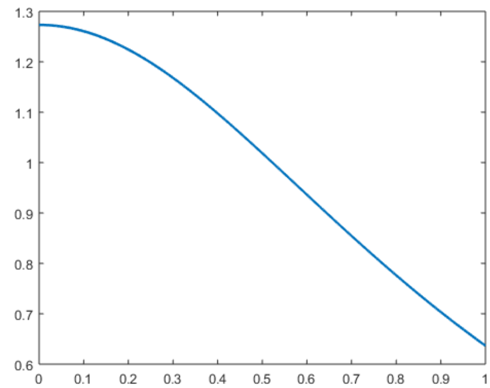

Figure 3.5: Invariant Density

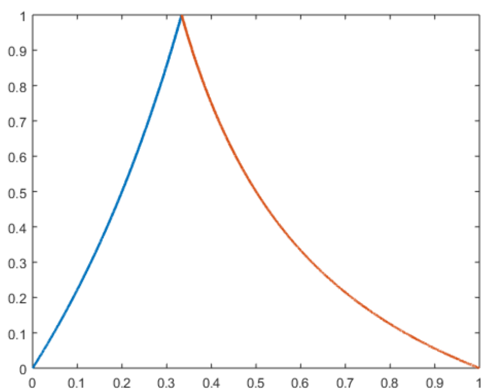

Figure 3.6: Graph of $S_2(x)$

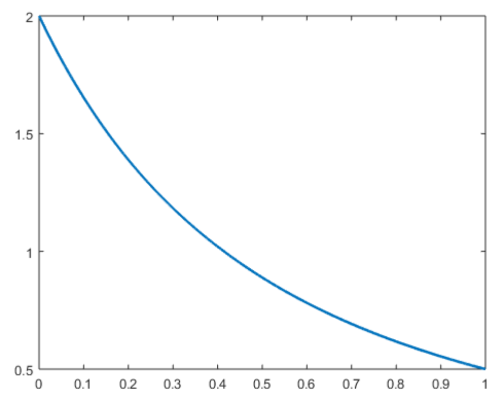

Figure 3.7: Invariant Density

To calculate the error, we divide the interval $[0,1]$ into $n = 2^r,\ r = 2,3,\ldots,8$ subintervals $I_i = [x_{i-1}, x_i],\ i = 1,2,\ldots,n$ of equal length $h = \frac{1}{n}$. The $m_i,\ i = 0,1,\ldots,2n$, are calculated using the actual invariant densities f_1^*, f_2^*, f_3^*, f_4^*, and f_5^* in (3.9); those m_i are used in (3.14) to get $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$. Finally, the $\lambda_0, \lambda_1, \ldots, \lambda_{2n}$ are used in (3.15) to obtain the approximate invariant densities. The errors, e_n, are calculated using the formula

$$e_n = \|f_n - f^*\|_{L^1} = \int_0^1 |f_n - f^*|\, dx.$$

The errors are presented in Table 3.1, Table 3.2, Table 3.3, Table 3.4, and Table 3.5.

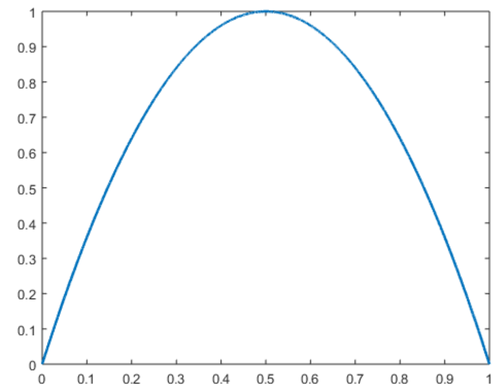

Figure 3.8: Graph of $S_3(x)$

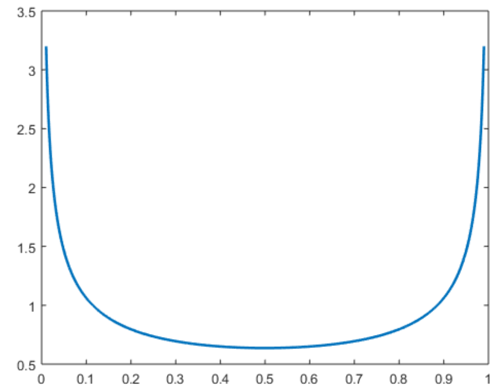

Figure 3.9: Invariant Density

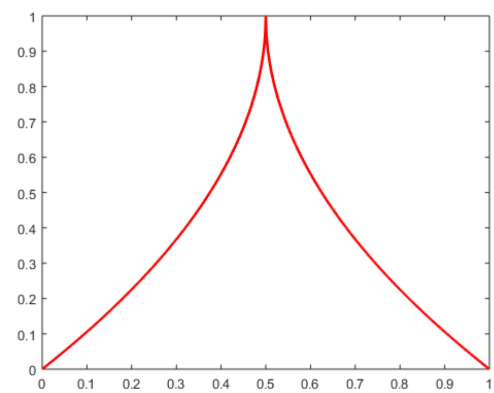

Figure 3.10: Graph of $S_4(x)$

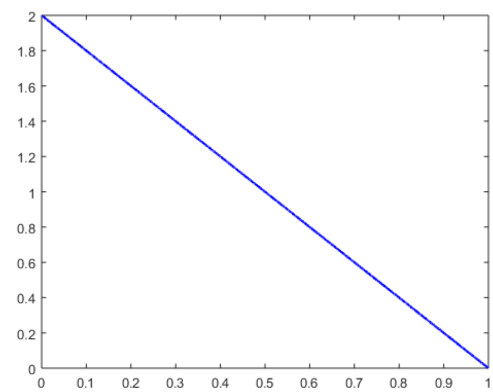

Figure 3.11: Invariant Density

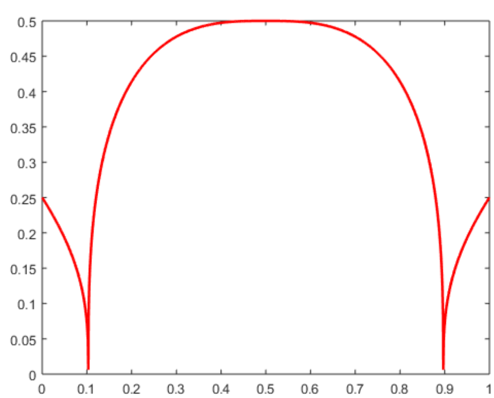

Figure 3.12: Graph of $S_5(x)$

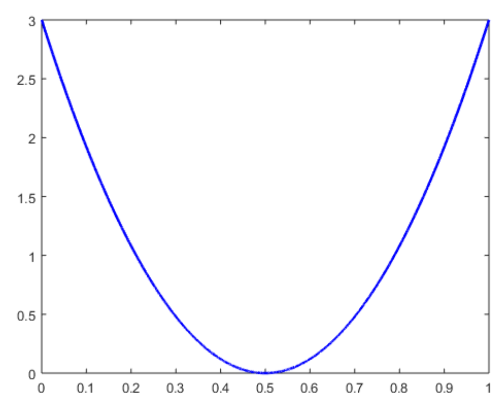

Figure 3.13: Invariant Density

n	PQMFA	PLMEM	PQMEM	CPU Time (Sec)
4	2.096419×10^{-2}	2.3×10^{-3}	1.5314×10^{-4}	0.774722
8	7.047357×10^{-3}	5.7×10^{-4}	2.0188×10^{-5}	0.500334
16	2.159652×10^{-3}	1.4×10^{-4}	2.5498×10^{-6}	0.554427
32	6.358728×10^{-4}	3.6×10^{-5}	3.1965×10^{-7}	0.796904
64	1.836483×10^{-4}	8.9×10^{-6}	4.0000×10^{-8}	1.563933
128	5.215950×10^{-5}	2.2×10^{-6}	5.0034×10^{-9}	3.2084285
256	$*$	5.6×10^{-7}	6.2562×10^{-10}	6.443589
512	$*$	$*$	7.8216×10^{-11}	13.487430
1024	$*$	$*$	9.7778×10^{-12}	28.769752
2048	$*$	$*$	1.2223×10^{-12}	69.269342

Table 3.1: Comparison of Errors, e_n, Associated with S_1

n	PQMFA	PLMEM	PQMEM	CPU Time (Sec)
4	4.366420×10^{-2}	2.4×10^{-3}	1.1855×10^{-4}	0.846421
8	1.920112×10^{-2}	5.9×10^{-4}	1.6623×10^{-5}	0.589024
16	8.297540×10^{-3}	1.5×10^{-4}	2.2120×10^{-6}	0.702382
32	3.117798×10^{-3}	3.7×10^{-5}	2.8645×10^{-7}	0.890773
64	1.034574×10^{-3}	9.2×10^{-6}	3.6522×10^{-8}	1.638339
128	3.202538×10^{-4}	2.3×10^{-6}	4.6107×10^{-9}	3.228516
256	$*$	5.7×10^{-7}	5.7920×10^{-10}	6.522812
512	$*$	$*$	7.2580×10^{-11}	13.318838
1024	$*$	$*$	9.0837×10^{-12}	28.698942
2048	$*$	$*$	1.1362×10^{-12}	69.510323

Table 3.2: Comparison of Errors, e_n, Associated with S_2

n	PQMFA	PLMEM	PQMEM	CPU Time (Sec)
4	3.681119×10^{-1}	2.4×10^{-1}	1.6916×10^{-1}	26.191738
8	3.057296×10^{-1}	1.7×10^{-1}	1.1931×10^{-1}	30.665695
16	2.453954×10^{-1}	1.2×10^{-1}	8.3465×10^{-2}	29.392794
32	1.801104×10^{-1}	8.4×10^{-2}	5.8523×10^{-2}	35.060111
64	1.346728×10^{-1}	5.9×10^{-2}	4.1834×10^{-2}	38.502450
128	9.663345×10^{-2}	4.2×10^{-2}	2.8717×10^{-2}	1093.743123
256	*	3.0×10^{-2}	2.0026×10^{-2}	2379.527132

Table 3.3: Comparison of Errors, e_n, Associated with S_3

n	PQMEM	CPU Time (Sec)
4	6.2000×10^{-3}	0.418198
8	1.6000×10^{-3}	0.333783
16	4.0657×10^{-4}	0.583267
32	1.0230×10^{-4}	1.277854
64	2.5671×10^{-5}	2.758252
128	6.4278×10^{-6}	5.528695
256	1.6065×10^{-6}	12.291537
512	4.0093×10^{-7}	27.524598
1024	1.0025×10^{-7}	60.307657
2048	2.5066×10^{-8}	151.421663

Table 3.4: Errors, e_n, Associated with S_4

n	PQMEM	CPU Time (Sec)
4	1.0126×10^{-2}	0.597442
8	2.2500×10^{-3}	0.409689
16	4.0313×10^{-4}	0.828435
32	6.1283×10^{-5}	1.647523
64	9.0157×10^{-6}	3.495872
128	1.2920×10^{-6}	8.132633
256	1.8222×10^{-7}	17.538370
512	2.5364×10^{-8}	40.068020
1024	3.4935×10^{-9}	92.235110
2048	4.7705×10^{-9}	215.851727

Table 3.5: Errors, e_n, Associated with S_5

The errors of the Piecewise Linear MEM (PLMEM) are taken from the paper [25] and the errors of the Piecewise Quadratic Markov Finite Approximation Method (PQMFA) are obtained from [18]. The numerical results, except the last two rows of the Table 5.3, show that the proposed Piecewise Quadratic MEM (PQMEM) is the best among the three. The invariant densities which are taken in Table 3.1, Table 3.2, Table 3.4, and Table 3.5 are smooth functions on $[0,1]$ and the order of convergence of errors is close to $O(h^3)$ for S_1, S_2 and S_5 and close to $O(h^2)$ for S_4. The function f^* in Table 3.3 is unbounded on $[0,1]$ so the order of convergence of errors doesn't follow any definite rule. The tabulated errors for this function are obtained by keeping λ_0 and λ_{2n} fixed.

Chapter 4

A Quadratic Spline Maximum Entropy Method

4.1 Introduction

When two or more curves are connected to form a curve that satisfies given conditions, the curve is called a spline. These constraints may consist of parametric and geometric continuity. Splines have been used to approximate complex curves. The use of splines was increased after the invention of computers.

Before the use of splines, one generally used a polynomial for approximation problems. In computational works, step functions and polynomials have been applied due to their simplicity in writing and easy-to-execute nature.

The following is a way to get a spline of degree n. Let $\{(x_i, y_i)_{i=0}^{n}\}$ be a set of $n+1$ points. The ordered numbers $x_1 < x_2 < \ldots < x_n$ are called knots. A piecewise polynomial function $f(x)$ defined by

$$f(x) = \begin{cases} f_0(x), & x_0 \le x < x_1 \\ f_1(x), & x_1 \le x < x_2 \\ \vdots & \vdots \\ f_{n-1}(x), & x_{n-1} \le x \le x_n \end{cases}$$

is called a spline of degree n if

1) each f_i is a polynomial of degree n,

2) $f(x)$ is $n-1$ times differentiable,

3) for each $j = 1, 2, \cdots, n-1$,

$$\begin{aligned} f_{(j-1)}(x_j) &= f_j(x_j) \\ f'_{(j-1)}(x_j) &= f'_j(x_j) \\ &\vdots \\ f^{(n-1)}_{(j-1)}(x_j) &= f^{(n-1)}_j(x_j). \end{aligned}$$

Linear, quadratic and cubic splines are common in applications. Among these three, cubic is the most used spline function.

A basis of a spline is known as B-splines and a spline function formed from basis functions. A spline function is given in the form

$$\sum_{i=0}^{n} B_i^k(t) P_i,$$

where $B_i^k(t)$ is a basis function that can be defined by using the Cox-de Boor recursion formula and $P_0, P_1, \ldots, P_n$ are control points.

In the previous chapter, a piecewise quadratic maximum entropy method was developed to approximate an invariant density of the F-P operator associated with nonsingular transformations S, $S : [0,1] \to [0,1]$. The scheme was based on the partition of the interval $[0,1]$ into n uniform subintervals and application of $2n+1$ continuous piecewise quadratic polynomials which satisfy the partition of unity property. This method improved the numerical results by producing the smaller errors than the errors produced by the piecewise linear maximum entropy method [25] and the piecewise quadratic Markov finite approximation method [18]. The scheme in [70] also opened a way to implement continuous higher order piecewise polynomials that satisfy the partition of unity property. Although this scheme produced the best results in comparison to two existing methods, it needed a long time and a large amount of computing memory for the numerical accomplishment due to the implementation of a large number, $2n+1$, of piecewise quadratic polynomials.

A new numerical scheme based on finite element MEM which will overcome the drawbacks of the PQMEM will be presented in this chapter. In that scheme, quadratic splines are used as moment functions and this new scheme reduces the number of nonlinear equations to $n+2$ which is nearly one half in comparison to the number of equations, $2n+1$, required in [70]. Since the number of moment functions reduces by nearly one half, this numerical scheme is faster and more efficient than PQMEM and other methods.

4.2 Quadratic Spline and B-Spline

The theory of splines and their applications are relatively new but splines are widely used in applications due to their approximation powers and other characteristics [65].

Definition 4.2.1. Let $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$, and we write $\Delta = \{x_i\}_{i=0}^{n}$. The set Δ partitions the interval $[a,b]$ into n subintervals $I_i = [x_{i-1}, x_i]$, $i = 1,2,\ldots,n$. Let d be a positive integer. Define

$$\mathcal{S}_d(\Delta) = \mathcal{P}_d(\Delta) \cap C^{d-1}[a,b],$$

where $\mathcal{P}_d(\Delta)$ is the space of piecewise polynomial functions of degree at most d with respect to Δ and $C^{d-1}[a,b]$ is the space of functions having continuous $d-1$ derivatives on $[a,b]$.

Thus, $\mathcal{S}_d(\Delta)$ is the set of the piecewise polynomials of degree at most d that are $d-1$ times continuously differentiable at all nodes $x_0, x_1, \ldots, x_n$, and each function in this set is called a degree-d *spline function* associated with the partition.

From the above definition, $\mathcal{S}_2(\Delta)$ consists of all continuously differentiable functions f such that f is a quadratic polynomial on each subinterval $[x_{i-1}, x_i]$, $i = 1, 2, \ldots, n$. It is well known from [65] that the space $\mathcal{S}_d(\Delta)$ of degree d spline functions associated with Δ is a vector space of dimension $n+d$. A special basis of $\mathcal{S}_d(\Delta)$, consisting of the so-called B-splines, can be constructed via a recurrence relation with respect to the lower degree B-splines. This can be done as follows.

It is easier to develop the recurrence relations for the B-splines if we assume that there are nodes $\cdots < x_{-2} < x_{-1}$ to the left of a and $x_{n+1} < x_{n+2} < \cdots$ to the right of b; in other words,it can be written as a bi-infinite sequence

$$\cdots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 < \cdots$$

of nodes on the real line.

The B-splines of degree 0, denoted as B_i^0, are just piecewise constant functions and defined as

$$B_i^0(x) = \begin{cases} 1, & x_i \le x < x_{i+1}, \\ 0, & \text{otherwise} \end{cases}$$

on the ith subinterval for $i = 1, 2, \ldots, n$, which are nothing but the characteristic functions of the subintervals. Starting from B_i^0, the recursive formula of degree-d B-splines with $d = 1, 2, \ldots$ are defined as

$$B_i^d(x) = \frac{x - x_i}{x_{i+d} - x_i}\, B_i^{d-1}(x) + \frac{x_{i+d+1} - x}{x_{i+d+1} - x_{i+1}}\, B_{i+1}^{d-1}(x)$$

for each i. It is obvious to see that each B_i^d is a piecewise polynomial of degree d and belongs to $\mathcal{S}_d(\Delta)$, so it is called a B-spline of degree d. Some basic properties of the B-splines are summarized in the following theorem; more properties can be referred to [11, 65].

Theorem 4.2.2. *(i) If $x \in (x_i, x_{i+d+1})$, then $B_i^d(x) > 0$.*

(ii) If $x \notin [x_i, x_{i+d+1})$, then $B_i^d(x) = 0$.

(iii) $\sum_i B_i^d(x) \equiv 1$ for all x.

(iv) The B-splines of degree-d

$$B_{-d}^d, B_{-d+1}^d, \ldots, B_{n-2}^d, B_{n-1}^d$$

constitute a basis for the space $\mathcal{S}_d(\Delta)$ of all functions in $C^{d-1}[0,1]$ which are polynomials of degree at most $\le d$ on each of the n subintervals

$$[x_0, x_1],\ [x_1, x_2],\ \ldots,\ [x_{n-1}, x_n].$$

The properties (i) and (ii) mean the *local positive support* of the B-splines, which shows that B-splines are positive on their support which is the union of only $d+1$ consecutive subintervals of the partition. The property (iii) gives the partition of unity and plays an important role in our numerical scheme when the B-splines are used as the moment functions. The property (iv) is the continuity property of the derivatives of the degree-d B-splines up to the order $d-1$, and it also gives the dimension of the space $\mathcal{S}_d(\Delta)$, which is $n+d$. In practical computations with the splines, the nodes x_i are usually evenly distributed so the resulting subintervals have the same length $h=\frac{b-a}{n}$. The resulting B-splines can be obtained from a "mother spline" followed by scaling and translation techniques. For example, the degree-1 B-splines used in [29], can be expressed as

$$B_i^1(x) = l\left(\frac{x-x_i}{h}\right),$$

where

$$l(x) = \begin{cases} x, & 0 \le x \le 1, \\ 2-x, & 1 < x \le 2, \\ 0, & x \notin [0,2] \end{cases}$$

is the standard tent function.

The degree-2 B-splines can be expressed as

$$B_i^2(x) = q\left(\frac{x-x_i}{h}\right),$$

where

$$q(x) = \begin{cases} \frac{1}{2}x^2, & 0 \le x \le 1, \\ \frac{3}{4} - \left(x-\frac{3}{2}\right)^2, & 1 < x \le 2, \\ \frac{1}{2}(x-3)^2, & 2 < x \le 3, \\ 0, & x \notin [0,3]. \end{cases}.$$

4.3 B-spline Functions in Maximum Entropy

B-splines are employed in the numerical estimation of an invariant density of the F-P operator associated with a nonsingular transformation $S:[0,1]\to[0,1]$. To define the B-splines properly, we divide the interval $[0,1]$ into the n subintervals $I_j=[x_{j-1},x_j],\ j=1,2,\ldots,n$, of equal width $h=1/n$. As mentioned previously, $2d$ external knots will be needed for the B-spline of degree d. The knots are defined as $x_{-i}=-ih$ and $x_{n+i}=(n+i)h$ for $i=1,2,\ldots,d$.

The maximization problem in term of B-spline functions has the form:

$$\text{maximize}\left\{H(f): f\in D, \int_0^1 f(x)B_i^d(x)\,dx = m_i,\ i=-d,-d+1,\ldots,n-1\right\}. \tag{4.1}$$

Since the B-splines B_i^d of degree d satisfy the partition of unity property, the solution of the maximum entropy problem (4.1) is:

$$f_n(x) = e^{\sum_{i=-d}^{n-1} \lambda_i B_i^d(x)}, \tag{4.2}$$

where the Lagrange multipliers λ_i satisfy the constraint conditions

$$\int_0^1 e^{\sum_{j=-d}^{n-1} \lambda_j B_j^d(x)} B_i^d(x)\, dx = m_i,\ i = -d, -d+1, \ldots, n-1. \tag{4.3}$$

Quadratic splines are used in MEM, so $d = 2$ in what follows. The figures of a single quadratic B-spline basis function, with the simple nodes at $x = 0, 1, 2, 3$, and the corresponding B-splines are given in Figure 4.1 and in Figure 4.2, respectively. When $[0, 1]$ is divided into n subintervals with the uniform length $h = 1/n$, then there will be $n+2$ B-splines $B_{-2}, B_{-1}, \ldots, B_{(n-1)}$, where the superscript "2" are removed for simplicity.

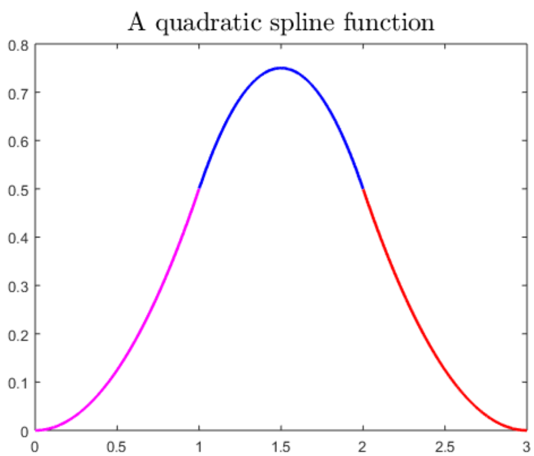

Figure 4.1: A Quadratic Spline Function

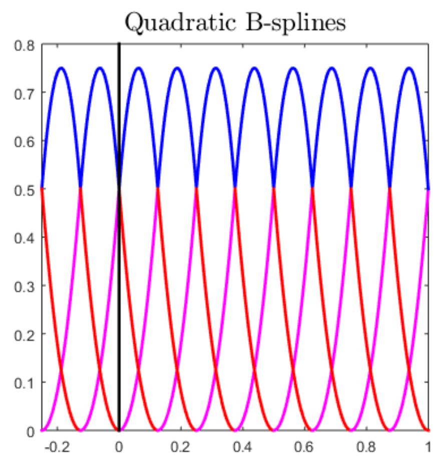

Figure 4.2: Quadratic B-Splines

4.4 Convergence Analysis

In this section, the error bounds for the proposed quadratic spline maximum entropy method are presented with respective mathematical proofs. If the F-P operator P_S associated with a nonsingular transformation $S : [0, 1] \to [0, 1]$ has an invariant density f^*, a useful result in the maximum entropy method will be obtained, which gives the weak convergence of the method when the Boltzmann entropy of f^* is finite.

Theorem 4.4.1. *(Theorem 3.1 [12]) If the entropy $H(f^*) = -\int_0^1 f^* \log f^*\, dx$ of f^* is finite, then the solutions f_n for all n given by (4.2) associated with $\mathcal{S}_2(\Delta)$ has the following*

properties:
i) f_n *converge weakly to* f^**, that is*
$\lim_{n\to\infty}\int_0^1 f_n(x)g(x)\,dx = \int_0^1 f^*(x)g(x)\,dx,\ \ \forall\, g\in L^\infty(0,1)$.
ii) $\lim_{n\to\infty} H(f_n) = H(f^*)$.

According to the theory developed in the papers [12, 13], it is essential to estimate the minimal distance of a function $f\in C[0,1]$ to the subspace spanned by the functions B_{-2}, $B_{-1},\dots,$ B_{n-1} under the infinity norm $\|f\|_\infty = \max\{|f(x)| : x\in[0,1]\}$. The estimation of the following quantity fulfills the above requirement:

$$\min_{(\lambda_{-2},\lambda_{-1},\dots,\lambda_{n-1})}\ \max_{x\in[0,1]}\left|f(x)-\sum_{k=-2}^{n-1}\lambda_k B_k(x)\right|.$$

Define a linear operator $Q : C[0,1]\to S_2(\Delta_n)$ by

$$(Qf)(x) = \sum_{i=-2}^{n-1}\mu_i(f)B_i(x),\ \ \forall\, x\in[0,1], \tag{4.4}$$

where

$$\mu_i(f) = \begin{cases} f(x_0), & i=-2,\\ \frac{1}{2}\left[4f(x_{i+3/2})-f(x_{i+1})-f(x_{i+2})\right], & i=-1,0,\dots,n-2,\\ f(x_n), & i=n-1. \end{cases}$$

Here for convenience, we write $x_{k+1/2} = (x_k+x_{k+1})/2$ for any k. A direct computation shows that Q satisfies that $Qp = p$ for $p(x) = 1,\ x$, and x^2.

Let $f\in C^3[0,1]$. To estimate $\max_{x\in[0,1]}|f(x)-(Qf)(x)|$, it is enough to estimate $|f(x)-(Qf)(x)|$ on a particular subinterval $[x_i,x_{i+1}]$ for some $1\le i\le n-2$. The error analysis on the subinterval $[x_0,x_1]$ and $[x_{n-1},x_n]$ can be done similarly. So we focus on the subinterval $[x_i,x_{i+1}]$ with a fixed $i\in\{1,\dots,n-2\}$. On this interval, there are only three nonzero terms in the expression (4.4) of Qf, which are associated with the B-splines

$$B_{i-2}(x) = \frac{1}{2}\left(\frac{x-x_{i-2}}{h}-3\right)^2,$$

$$B_{i-1}(x) = \frac{3}{4}-\left(\frac{x-x_{i-1}}{h}-\frac{3}{2}\right)^2,$$

$$B_i(x) = \frac{1}{2}\left(\frac{x-x_i}{h}\right)^2.$$

Then for $x \in [x_i, x_{i+1}]$, from the definition of Qf, we have

$$
\begin{aligned}
(Qf)(x) &= \mu_{i-2}(f)B_{i-2}(x) + \mu_{i-1}(f)B_{i-1}(x) + \mu_i(f)B_i(x) \\
&= \frac{1}{2}\left[4f(x_{i-1/2}) - f(x_{i-1}) - f(x_i)\right] B_{i-2}(x) \\
&\quad + \frac{1}{2}\left[4f(x_{i+1/2}) - f(x_i) - f(x_{i+1})\right] B_{i-1}(x) \\
&\quad + \frac{1}{2}\left[4f(x_{i+3/2}) - f(x_{i+1}) - f(x_{i+2})\right] B_i(x) \\
&= \frac{1}{4}\left[4f(x_{i-1/2}) - f(x_{i-1}) - f(x_i)\right] \left(\frac{x - x_{i-2}}{h} - 3\right)^2 \\
&\quad + \frac{1}{2}\left[4f(x_{i+1/2}) - f(x_i) - f(x_{i+1})\right] \left(\frac{3}{4} - \left(\frac{x - x_{i-1}}{h} - \frac{3}{2}\right)^2\right) \\
&\quad + \frac{1}{4}\left[4f(x_{i+3/2}) - f(x_{i+1}) - f(x_{i+2})\right] \left(\frac{x - x_i}{h}\right)^2.
\end{aligned}
$$

Let $t = (x - x_i)/h$. Then

$$
\begin{aligned}
(Qf)(x) &= \frac{1}{4}\left[4f(x_{i-1/2}) - f(x_{i-1}) - f(x_i)\right](t-1)^2 + \frac{1}{2}\left[4f(x_{i+1/2}) - f(x_i) - f(x_{i+1})\right] \\
&\quad \left(\frac{3}{4} - \left(t - \frac{1}{2}\right)^2\right) + \frac{1}{4}\left[4f(x_{i+3/2}) - f(x_{i+1}) - f(x_{i+2})\right] t^2 \\
&= \left[f(x_{i-1/2}) - \frac{1}{4}f(x_{i-1}) - \frac{1}{4}f(x_i)\right]\left(t^2 - 2t + 1\right) + \left[2f(x_{i+1/2}) - \frac{1}{2}f(x_i) - \frac{1}{2}f(x_{i+1})\right] \\
&\quad \left(-t^2 + t + 1/2\right) + \left[f(x_{i+3/2}) - \frac{1}{4}f(x_{i+1}) - \frac{1}{4}f(x_{i+2})\right] t^2 \\
&= \left[f(x_{i-1/2}) - \frac{1}{4}f(x_{i-1}) - \frac{1}{4}f(x_i) - 2f(x_{i+1/2}) + \frac{1}{2}f(x_i) + \frac{1}{2}f(x_{i+1}) + f(x_{i+3/2})\right. \\
&\quad \left. - \frac{1}{4}f(x_{i+1}) - \frac{1}{4}f(x_{i+2})\right] t^2 + \left[-2f(x_{i-1/2}) + \frac{1}{2}f(x_{i-1}) + \frac{1}{2}f(x_i) + 2f(x_{i+1/2})\right. \\
&\quad \left. - \frac{1}{2}f(x_i) - \frac{1}{2}f(x_{i+1})\right] t + f(x_{i-1/2}) - \frac{1}{4}f(x_{i-1}) - \frac{1}{4}f(x_i) + f(x_{i+1/2}) - \frac{1}{4}f(x_i) - \frac{1}{4}f(x_{i+1}) \\
&= \left[f(x_{i-1/2}) - \frac{1}{4}f(x_{i-1}) + \frac{1}{4}f(x_i) - 2f(x_{i+1/2}) + \frac{1}{4}f(x_{i+1}) + f(x_{i+3/2}) - \frac{1}{4}f(x_{i+2})\right] t^2 \\
&\quad + \left[-2f(x_{i-1/2}) + \frac{1}{2}f(x_{i-1}) + 2f(x_{i+1/2}) - \frac{1}{2}f(x_{i+1})\right] t \\
&\quad + \left[f(x_{i-1/2}) - \frac{1}{4}f(x_{i-1}) - \frac{1}{2}f(x_i) + f(x_{i+1/2}) - \frac{1}{4}f(x_{i+1})\right].
\end{aligned}
$$

Expanding each term inside the brackets of the above expression at the point x_i gives

$$
\begin{aligned}
(Qf)(x) \;=\; & \Big[f(x_i)+f'(x_i)(-h/2)+\frac{f''(x_i)}{2!}(-h/2)^2+O(h^3) \\
& -\frac{1}{4}\Big(f(x_i)+f'(x_i)(-h)+\frac{f''(x_i)}{2!}(-h)^2+O(h^3)\Big)+\frac{1}{4}f(x_i) \\
& -2\Big(f(x_i)+f'(x_i)(h/2)+\frac{f''(x_i)}{2!}(h/2)^2+O(h^3)\Big) \\
& +\frac{1}{4}\Big(f(x_i)+f'(x_i)h+\frac{f''(x_i)}{2!}h^2+O(h^3)\Big) \\
& +f(x_i)+f'(x_i)(3h/2)+\frac{f''(x_i)}{2!}(3h/2)^2+O(h^3) \\
& -\frac{1}{4}\Big(f(x_i)+f'(x_i)(2h)+\frac{f''(x_i)}{2!}(2h)^2+O(h^3)\Big)\Big]t^2 \\
& +\Big[-2\Big(f(x_i)+f'(x_i)(-h/2)+\frac{f''(x_i)}{2!}(-h/2)^2+O(h^3)\Big) \\
& +\frac{1}{2}\Big(f(x_i)+f'(x_i)(-h)+\frac{f''(x_i)}{2!}(-h)^2+O(h^3)\Big) \\
& +2\Big(f(x_i)+f'(x_i)(h/2)+\frac{f''(x_i)}{2!}(h/2)^2+O(h^3)\Big) \\
& -\frac{1}{2}\Big(f(x_i)+f'(x_i)h+\frac{f''(x_i)}{2!}h^2+O(h^3)\Big)\Big]t \\
& +\Big[(f(x_i)+f'(x_i)(-h/2)+\frac{f''(x_i)}{2!}(-h/2)^2+O(h^3) \\
& -\frac{1}{4}\Big(f(x_i)+f'(x_i)(-h)+\frac{f''(x_i)}{2!}(-h)^2+O(h^3)\Big) \\
& -\frac{1}{2}(f(x_i)+f(x_i)+f'(x_i)(h/2)+\frac{f''(x_i)}{2!}(h/2)^2+O(h^3) \\
& -\frac{1}{4}\Big(f(x_i)+f'(x_i)h+\frac{f''(x_i)}{2!}h^2+O(h^3)\Big)\Big].
\end{aligned}
$$

After simplification and substituting t back to $x = x_i + ht$, we have

$$(Qf)(x) = \left(\frac{1}{2}f''(x_i)+O(h)\right)(x-x_i)^2+\big(f'(x_i)-O(h^2)\big)(x-x_i)+f(x_i)+O(h^3).$$

Thus, by expanding $f(x)$ at x_i and noting that $x - x_i = O(h)$, we obtain

$$
\begin{aligned}
(Qf-f)(x) \;=\; & \left(\frac{1}{2}f''(x_i)+O(h)\right)(x-x_i)^2+\big(f'(x_i)-O(h^2)\big)(x-x_i)+f(x_i)-f(x_i) \\
& -f'(x_i)(x-x_i)-\frac{f''(x_i)}{2!}(x-x_i)^2-O(h^3)=O(h^3).
\end{aligned}
$$

This proves the following approximation result.

Theorem 4.4.2. *For a given function $f \in C^3[0,1]$,*

$$\min_{(\lambda_{-2},\lambda_{-1},\ldots,\lambda_{n-1})} \max_{x\in[0,1]} \left| f(x) - \sum_{k=-2}^{n-1} \lambda_k B_k^2(x) \right| = O(h^3),$$

where h is the length of the uniform subintervals.

The combination of the above result with the assumption $\log f^* \in C^3[0,1]$ and Theorem 4.7 of [12] gives the L^1-norm estimation as the following convergence rate result shows.

Theorem 4.4.3. *Suppose f^* is a unique invariant density of the F-P operator such that $\log f^* \in C^3[0,1]$. If f_n is the sequence of the degree-2 spline maximum entropy approximations of f^*, then $\|f^* - f_n\| = O(h^3)$.*

Proof. Let $E_n = \min\left\{ \|\sum_{k=-2}^{n-1} \lambda_k B_k - \log f^*\|_\infty : \lambda_k \in \mathbb{R} \right\}$. Then $E_n = O(h^3)$ by Theorem 4.4.2. From Theorem 4.7 of [12], $\|f_n - f^*\| \le E_n e^{\frac{E_n}{2}}$. Since, $e^{\frac{E_n}{2}} = O(1)$, it follows that $\|f_n - f^*\| = O(h^3)$. □

4.5 Numerical Results

This section presents the numerical results from the degree-2 spline MEM implemented to the same set of the five transformations given in Section 3.6.

To calculate the errors, divide the interval $[0,1]$ into $n = 2^s,\ s = 2,3,\ldots,8$ subintervals $I_i = [x_{i-1}, x_i],\ i = 1,2,\ldots,n$ of equal length $h = \frac{1}{n}$. The moments $m_i,\ i = -2,-1,\ldots,n-1$, are calculated using the actual invariant densities f_1^*, f_2^*, and f_3^*, so

$$m_i = \int_0^1 B_i(x) f_j^*(x)\,dx,\ i = -2,-1,\ldots n-1.$$

Those m_i are used in (4.3) to get $\lambda_{-2}, \lambda_{-1}, \ldots, \lambda_{n-1}$. Finally, the $\lambda_{-2}, \lambda_{-1}, \ldots, \lambda_{n-1}$ are used in (4.2) to obtain the approximate invariant densities. The errors e_n are calculated using the formula

$$e_n = \|f_n - f^*\| = \int_0^1 |f_n(x) - f^*(x)|\,dx$$

and presented in Table 4.1, Table 5.2, Table 5.3, Table 5.1, and Table 4.5. The errors of the PQMFA are taken from [18], the errors for the PLMEM are based on [25], and the errors in the PQMEM are obtained from [70].

From the comparison between the performance of our degree-2 spline method and the piecewise quadratic method from the last chapter, it is clear that, although both of them have the same order of convergence with respect to the same partition of the interval, the involved numerical work of the current method is much reduced due to the fact that the number of the nonlinear equations of the new system for the Lagrange multipliers from the maximum entropy method using the continuously differentiable B-splines of degree 2, is about half of that from the method that uses only continuous piecewise quadratic moment functions in [70].

The convergence and errors estimation are shown when the moments are exactly known. The moments $m_{-2}, m_{-1}, \ldots, m_{n-1}$ can't be calculated when the corresponding invariant density f^* is unknown. In this situation, the way to estimate those moments, via iteration, is to use the Birkhoff individual ergodic theorem [15]. The approximated moments, no matter how large N may be, have errors in comparison to the exact moments. The theoretical analysis of the effects, of the errors in moments that influence the approximate maximum entropy solution, was presented in [70]. The analysis established that, under certain conditions, the estimated moments influence the maximum entropy estimation by $O(\|\delta\|_\infty)$, where $\delta = (\delta_{-2}, \delta_{-1}, \ldots, \delta_{n-1})^T$, and δ_i is the difference between an ith exact (m_i) and estimated ($\hat{m}_i$) moment. Though the ergodic theorem very often exhibits a slow convergence, we think that the choice of a large N in the ergodic theorem can produce a small difference between two moments m_i and $\hat{m}_i$ as we desire. When $\|\delta\|_\infty$ is small enough, the approximated invariant density is not much affected by the choice of the moments.

n	**PQMFA**	**PLMEM**	**PQMEM**	**QSMEM**	**CPU Time(Sec)**
4	2.096419×10^{-2}	2.3×10^{-3}	1.5314×10^{-4}	1.6895×10^{-4}	0.654131
8	7.047357×10^{-3}	5.7×10^{-4}	2.0188×10^{-5}	2.0848×10^{-5}	0.498099
16	2.159652×10^{-3}	1.4×10^{-4}	2.5498×10^{-6}	2.5718×10^{-6}	0.577895
32	6.358728×10^{-4}	3.6×10^{-5}	3.1965×10^{-7}	3.2038×10^{-7}	0.786626
64	1.836483×10^{-4}	8.9×10^{-6}	4.0000×10^{-8}	4.0027×10^{-8}	1.705700
128	5.215950×10^{-5}	2.2×10^{-6}	5.0034×10^{-9}	5.0043×10^{-9}	3.234171
256	$*$	5.6×10^{-7}	6.2562×10^{-10}	6.2566×10^{-10}	6.460725
512	$*$	$*$	7.8218×10^{-11}	7.8218×10^{-11}	12.816098
1024	$*$	$*$	9.7778×10^{-12}	9.7779×10^{-12}	26.251176
2048	$*$	$*$	1.2223×10^{-12}	1.2223×10^{-12}	53.800612

Table 4.1: Comparison of Errors, e_n, Associated with S_1

n	PQMFA	PLMEM	PQMEM	QSMEM	CPU Time (Sec)
4	4.366420×10^{-2}	2.4×10^{-3}	1.1855×10^{-4}	1.2210×10^{-4}	0.559815
8	1.920112×10^{-2}	5.9×10^{-4}	1.6623×10^{-5}	1.6783×10^{-5}	0.529327
16	8.297540×10^{-3}	1.5×10^{-4}	2.2120×10^{-6}	2.2244×10^{-6}	0.628871
32	3.117798×10^{-3}	3.7×10^{-5}	2.8645×10^{-7}	2.8745×10^{-7}	0.888102
64	1.034574×10^{-3}	9.2×10^{-6}	3.6522×10^{-8}	3.6581×10^{-8}	1.631717
128	3.202538×10^{-4}	2.3×10^{-6}	4.6107×10^{-9}	4.6144×10^{-9}	3.247069
256	*	5.7×10^{-7}	5.7920×10^{-10}	5.7944×10^{-10}	6.542880
512	*	*	7.2580×10^{-11}	7.2594×10^{-11}	13.552166
1024	*	*	9.0837×10^{-12}	9.0846×10^{-12}	27.303408
2048	*	*	1.1362×10^{-12}	1.1362×10^{-12}	55.576483

Table 4.2: Comparison of Errors, e_n, Associated with S_2

n	PQMFA	PLMEM	PQMEM	QSMEM	CPU Time (Sec)
4	3.681119×10^{-1}	2.4×10^{-1}	1.6916×10^{-1}	1.8881×10^{-1}	23.740532
8	3.057296×10^{-1}	1.7×10^{-1}	1.1931×10^{-1}	1.3768×10^{-1}	24.584100
16	2.453954×10^{-1}	1.2×10^{-1}	8.3465×10^{-2}	9.7225×10^{-2}	27.577844
32	1.801104×10^{-1}	8.4×10^{-2}	5.8523×10^{-2}	6.8380×10^{-2}	30.930162
64	1.346728×10^{-1}	5.9×10^{-2}	4.1834×10^{-2}	4.8221×10^{-2}	35.086297
128	9.663345×10^{-2}	4.2×10^{-2}	2.8717×10^{-2}	3.4053×10^{-2}	65.791768
256	*	3.0×10^{-2}	2.0026×10^{-2}	1.5171×10^{-1}	354.794894

Table 4.3: Comparison of Errors, e_n, Associated with S_3

n	PQMEM	QSMEM	CPU Time (Sec)
4	6.2000×10^{-3}	7.4000×10^{-3}	0.855025
8	1.6000×10^{-3}	2.0000×10^{-3}	0.283095
16	4.0657×10^{-4}	4.9250×10^{-4}	0.578824
32	1.0230×10^{-4}	1.2378×10^{-4}	1.282375
64	2.5671×10^{-5}	3.1023×10^{-5}	2.747500
128	6.4278×10^{-6}	7.7658×10^{-6}	5.504208
256	1.6065×10^{-6}	1.9417×10^{-6}	12.064600
512	4.0093×10^{-7}	4.8622×10^{-7}	26.236218
1024	1.0025×10^{-7}	1.2157×10^{-7}	52.520748
2048	2.5066×10^{-8}	3.0396×10^{-8}	112.108948

Table 4.4: Comparison of Errors, e_n, Associated with S_4

n	**PQMEM**	**QSMEM**	**CPU Time (Sec)**
4	1.0126×10^{-2}	1.0640×10^{-1}	0.491448
8	2.2500×10^{-3}	2.2700×10^{-2}	0.466884
16	4.0313×10^{-4}	3.200×10^{-3}	0.626806
32	6.1283×10^{-5}	4.1322×10^{-4}	1.386648
64	9.0157×10^{-6}	5.3001×10^{-5}	2.999914
128	1.2920×10^{-6}	6.7911×10^{-6}	6.969542
256	1.8222×10^{-7}	8.6961×10^{-7}	14.974448
512	2.5364×10^{-8}	1.1129×10^{-7}	32.576887
1024	3.4935×10^{-9}	1.4234×10^{-8}	71.906818
2048	4.7705×10^{-9}	1.8196×10^{-9}	151.107587

Table 4.5: Comparison of Errors, e_n, Associated with S_5

Chapter 5

Invariant Densities of the Frobenius-Perron operator Associated with Random Maps

5.1 Introduction

A random dynamical system is a family of maps defined from a state space into itself in which iterates of a point are determined by a given probability distribution. For a random map $T = \{T_1, T_2, \ldots, T_k; p_1, p_2, \ldots, p_k\}$ a sequence of iterates of an initial point, x_0, is defined in the following way:

Suppose a member of the family of the maps, say T_{k_1}, is chosen randomly with a probability p_{k_1}. When the map acts on x_0 with the probability p_{k_1}, the first iterate is $x_1 = T_{k_1}(x_0)$. If a map T_{k_2} is chosen randomly, independent with T_{k_1}, with a probability p_{k_2}, then the second iterate is $x_2 = T_{k_2}(x_1) = T_{k_2}(T_{k_1}(x_0))$. In this way the nth iterate will be

$$x_n = T_{k_n}(x_{n-1}) = T_{k_n} \circ T_{k_{n-1}} \circ \cdots \circ T_{k_1}(x_0).$$

The sequence $\{x_n\}$ is called a Markov process. A Markov process which is stable in distribution is studied in [8].

Let $T_1, T_2, \ldots, T_k$ be a family of maps defined from $[0,1]$ to itself, and let $p_1, p_2, \ldots, p_k$ be the probabilities, where $p_i \geq 0, \forall i$ and $\sum_{i=1}^{k} p_i = 1$. A random map T is a random dynamical system written as $T = \{T_1, T_2, \ldots, T_k; p_1, p_2, \ldots, p_k\}$. The iterates of a point in the state space are generated according to the rule defined above. Bhattacharya and Majumdar [8] described Solow's growth model [66] using a random dynamical system with state space $\mathbb{R}_+$, and taking a continuous map depending on saving rate and net output function in per capita. Random maps are used to construct a generalized binomial model [3] for the study of existence of an invariant asymptotic density. Random maps with contraction property have been used to study fractals [6]. In the position dependent random map, the probability of switching from one transformation to another transformation depends on the position of the point in the iteration process. In a more precise way, a position dependent random map $T = \{T_1, T_2, \ldots, T_k : p_1, p_2, \ldots, p_k\}$ is a collection of maps defined on state space and the selection of such maps depend on a probability distribution function $p = \{p_i(x)\}_{i=1}^{k}$ where each p_i depends on the position x. The probability functions, $p_i(x)$, are all point functions.

Pelikan [58] described a sufficient condition for the existence of invariant measures for a random map with constant probabilities and also discussed the number of ergodic components of the map. The study focused on finding a class of random maps for which it is possible to describe the distribution of almost every trajectory. In the paper, Pelikan used the concept of pseudo skew product so every random map may be realized as a transformation of the square $[0,1] \times [0,1]$ to itself.

A deterministic representation of random maps is presented in [5] that discussed the structure of sets which have invariant measures. The deterministic skew-type representation for random maps with constant probabilities was provided. This representation coincides with the skew product for random maps with constant probabilities. It also established a one-to-one correspondence between eigenfunctions of the F-P operator associate with the deterministic skew-type maps and the eigenfunctions of the transform operator associated with the random maps for which the skew-type products are defined.

If random maps have position dependent probabilities, then the maps can not be taken as a skew product. A sufficient condition for the existence of an invariant density of the F-P operator associated with position dependent random maps was studied in [37]. The geometrical and topological properties of the sets of absolutely continuous invariant measures were also discussed there. Few attempts have be done for numerical approximations of invariant densities of the Markov operator associated with random maps and error analysis of such approximations. In this research, the invariant densities of the F-P operator associated with different random maps are approximated by using the piecewise linear MEM. The L^1 errors between the exact and approximate invariant densities of the Markov operator are calculated and presented in tabular form.

5.2 Markov Operator of Position Dependent Random Maps

Let $([0,1], \mathcal{A}, \mu)$ be a normalized measure space. Let $T = \{T_1, T_2, \dots, T_k; p_1, p_2, \dots, p_k\}$ be a position dependent random map, where $T_i : [0,1] \to [0,1],\ i = 1,2,\dots,k,$ are one-to-one and onto functions defined on common partitions $I_j = [x_{j-1}, x_j], j = 1,2,\dots,n,$ and each T_i is chosen by the probability function $p_i(x) \geq 0,\ i = 1,2,\dots,k$ with $\sum_{i=1}^{k} p_i(x) = 1$, for all $x \in [0,1]$. The restriction of T_i on the partition I_j is given by

$$T_{i,j} = T_{i|_j},\ \forall i = 1,2\dots,k; j = 1,2,\dots,n.$$

A measure ν is called an invariant measure of T if

$$\nu(A) = \sum_{i=1}^{n} \int_{T_i^{-1}(A)} p_i(x)\, d\nu(x),\ \forall A \in \mathcal{A}.$$

Under the transition probability

$$\mathbf{P}(x,A) = \sum_{i=1}^{n} p_i(x)\chi_A(T_i(x)),$$

where χ_A denotes the characteristic function of $A \in \mathcal{A}$, and the random map T is a Markov process.

Definition 5.2.1. [37] The Markov operator corresponding to T, $\mathcal{P} : L^1[0,1] \to L^1[0,1]$, is defined by

$$P_T f(x) = \sum_{i=1}^{k} P_{T_i}\left(p_i(x)f(x)\right), \tag{5.1}$$

where P_{T_i} is the F-P operator corresponding to T_i and

$$P_{T_i} f(x) = \sum_{j=1}^{n} \frac{f(T_{i,j}^{-1}(x))}{|T_i'(T_{i,j}^{-1}(x))|}\chi_{T_{I_j}(x)}.$$

Let $T = \{T_1, T_2, \ldots, T_k; p_1, p_2, \ldots, p_k\}$ be a random map, the Nth iteration of T is taken as

$$T^N(x) = (T_{r_N} \circ T_{r_{N-1}} \circ \cdots \circ T_{r_1}(x)),$$

with the probability

$p_{r_N}(T_{r_{N-1}} \circ \cdots \circ T_{r_1}(x))p_{r_{N-1}}(T_{r_{N-2}} \circ \cdots \circ T_{r_1}(x)) \cdots p_{r_1}(x)$, where $r_N, r_{N-1}, \cdots, r_1 \in \{1, 2, \ldots, k\}$.

Let $(I, \mathcal{A}, \lambda)$ be a normalized measure space with $I = [a,b]$. Let a random map $T = \{T_1, T_2, \ldots, T_k; p_1, p_2, \ldots, p_k\}$ defined on I. Let $T_r : I \to I,\ r = 1, 2, \ldots, k$ be one-to-one and differentiable on a partition P. Let BV(I) be the set of functions of bounded variations defined on I equipped with the norm $\|\cdot\|_{BV} = V(\cdot) + \|\cdot\|_1$, where $V(\cdot)$ is the variation of a function. Define

$$h_r(x) = \frac{p_r(x)}{|T_k'(x)|}, r = 1, 2, \ldots, k. \tag{5.2}$$

If

$$\sum_{r=1}^{k} h_r(x) < \alpha < 1,\ x \in I, \tag{5.3}$$

and

$$h_r \in BV(I),\ r = 1, 2, \ldots, k, \tag{5.4}$$

then

$$V_I \mathcal{P}_T^n f \leq AV_I f + B\|f\|_1, \tag{5.5}$$

for some $n \geq 1$, $0 < A < 1$ and $B > 0$. The inequality (5.5) is a criterion for the existence of an absolutely continuous invariant measure for the position dependent random map T [4].

5.3 Piecewise Linear Polynomials

Let P_n be a partition of the interval $[0,1]$ consisting of n uniform subintervals $I_j = [x_{j-1}, x_j],\ j = 1,2,\ldots,n$ of length $h = 1/n$. Piecewise linear polynomials are defined by:

$$\phi_k(x) = \ r(\tfrac{x-x_k}{h}), \quad k = 0,1,\ldots,n, \tag{5.6}$$

where the function $r(x)$ is

$$r(x) = \begin{cases} 1+x, & x \in [-1,0] \\ 1-x, & x \in [0,1] \\ 0, & \text{otherwise} \end{cases}. \tag{5.7}$$

The function $r(x)$ is called a Hat function. From definitions (5.6) and (5.7),

$$\phi_k(x) = r\left(\frac{x-x_k}{h}\right) = \begin{cases} 1+\frac{x-x_k}{h}, & -1 \le \frac{x-x_k}{h} \le 0 \\ \\ 1-\frac{x-x_k}{h}, & 0 \le \frac{x-x_k}{h} \le 1 \end{cases},$$

for $k = 0,1,\ldots,n$.

The support of ϕ_0 is $[0,h]$, the support of ϕ_k is $[x_{k-1}, x_{k+1}]$ for $k = 1,2,\ldots,n-1$, and the support of ϕ_n is $[1-h,1]$. The piecewise functions satisfy the partition of unity property

$$\sum_{k=0}^{p} \phi_k(x) = 1,\ \forall x \in [0,1].$$

The partition of unity property will be utilized to find the numerical approximation of invariant densities of the F-P operator. Let Δ_n be a, $n+1$ dimensional, subspace of $L^1[0,1]$ generated by the piecewise linear polynomials defined on the partition P_n of $[0,1]$. That is $\Delta_n = \text{span}\{\phi_k(x) : \phi_k(x) \in C^0[0,1], k = 0,1,\ldots,n\}$.

5.4 Piecewise Linear Maximum Entropy

The piecewise linear maximum entropy problem has the form

$$\text{maximize } \left\{ H(f) : f \in D, \int_0^1 f(x)\phi_k(x)\,dx = m_k, 0 \le k \le n \right\}. \tag{5.8}$$

The unique solution of the above problem is

$$f_n(x) = \frac{e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)}}{\int_0^1 e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)}\,dx}, \tag{5.9}$$

where the coefficients $\lambda_0, \lambda_2, \ldots, \lambda_n$ satisfy the system of equations

$$\int_0^1 \phi_i(x) e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)}\,dx = m_i \int_0^1 e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)}\,dx,\ i = 0,1,\ldots,n. \tag{5.10}$$

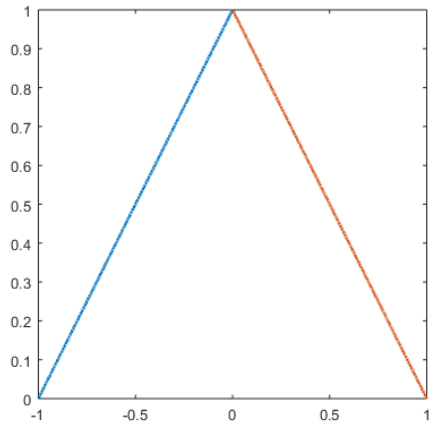

Figure 5.1: A Hat Function

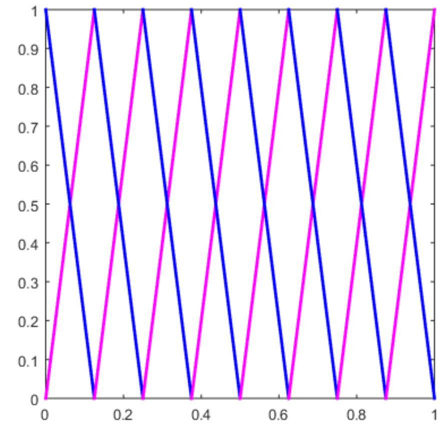

Figure 5.2: Piecewise Linear Polynomials

If the F-P operator, P_T, associated to a nonsingular transformation $T : [0,1] \to [0,1]$ has an invariant density f^*, then the kth moment, m_k, is defined by

$$m_k = \int_0^1 f^*(x)\phi_k(x)\,dx.$$

It can be shown that the sum, $\sum_{k=0}^{n} m_k$, of the moments m_k is unity. Using this fact, it can be shown that the function

$$f_n(x) = e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)} \tag{5.11}$$

is a density function [70] that maximizes (5.8) under the given constraints.

5.5 Numerical Results

Example 5.5.1. [37] Let T_1 and T_2 be two maps defined by:

$$T_1(x) = \begin{cases} 2x, & 0 \le x \le 0.2 \\ 3x - 0.2, & 0.2 \le x \le 0.4 \\ 2x - 0.6, & 0.4 < x \le 0.6 \\ 2x - 1.2, & 0.6 < x \le 0.8 \\ 2x - 1, & 0.8 < x \le 1 \end{cases}, \tag{5.12}$$

$$T_2(x) = \begin{cases} 2x + 0.4, & 0 \le x \le 0,2 \\ 2x - 0.4, & 0.2 < x \le 0.4 \\ 2x - 0.2, & 0.4 < x \le 0.6 \\ 2x - 0.6, & 0.6 < x \le 0.8 \\ 3x - 2.4, & 0.8 < x \le 1 \end{cases}.$$

Let the probability be defined by,

$$p_1(x) = \begin{cases} \frac{3}{4}, & 0 \le x \le 0.2 \\ \frac{3}{4}, & 0.2 < x \le 0.4 \\ \frac{1}{4}, & 0.4 < x \le 0.6 \\ \frac{1}{4}, & 0.6 < x \le 0.8 \\ \frac{1}{4}, & 0.8 < x \le 1 \end{cases},$$

$$p_2(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 0,2 \\ \frac{1}{4}, & 0.2 < x \le 0.4 \\ \frac{3}{4}, & 0.4 < x \le 0.6 \\ \frac{3}{4}, & 0.6 < x \le 0.8 \\ \frac{3}{4}, & 0.8 < x \le 1 \end{cases},$$

and $\tau_1 = \{T_1, T_2; p_1, p_2\}$ be the random map.

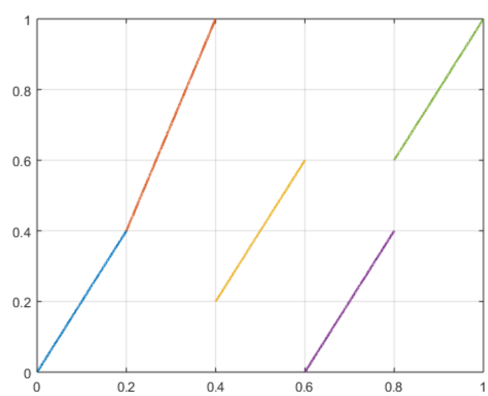

Figure 5.3: Graph of $T_1(x)$

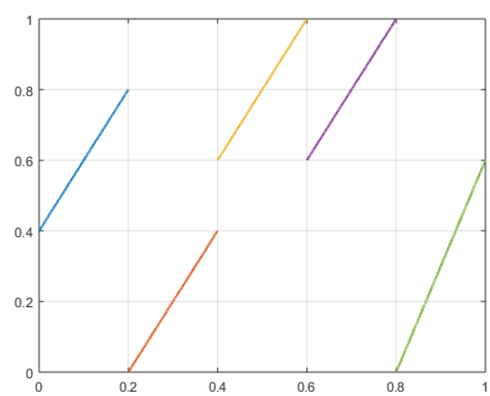

Figure 5.4: Graph of $T_2(x)$

Since the maps $T_1(x)$ and $T_2(x)$ are Markov, the F-P operator associated with each of these maps has a matrix representation. The matrix representation of the F-P operator associated with $T_1(x)$ can be derive in the following way:

Let $T_{11}(x) = 2x,\ T_{12}(x) = 3x - 0.2,\ T_{13}(x) = 2x - 0.6,\ T_{14} = 2x - 1.2,\ T_{15}(x) = 2x - 1,$ and $I_1 = [0, 0.2],\ I_2 = (0.2, 0.2],\ I_3 = (0.4, 0.6],\ I_4 = (0.6, 0.8],\ I_5 = (0.8, 1].$

Let $f = [f_1, f_2, f_3, f_4, f_5]$ be a piecewise constant function on the given partitions I_1, I_2, I_3, I_4, I_5. The F-P operator associated with $T_1(x)$ is

$$
\begin{aligned}
P_{T_1} f(x) &= \sum_{i=1}^{5} f_i |T'_{1\mathrm{i}}(x)|^{-1} \chi_{T_{1\mathrm{i}(I_i)}}(x) \\
&= f_1 |T'_{11}(x)|^{-1} \chi_{T_{11}(I_1)}(x) + f_2 |T'_{12}(x)|^{-1} \chi_{T_{12}(I_2)}(x) + f_3 |T'_{13}(x)|^{-1} \chi_{T_{13}(I_3)}(x) \\
&+ f_4 |T'_{14}(x)|^{-1} \chi_{T_{14}(I_4)}(x) + f_5 |T'_{15}(x)|^{-1} \chi_{T_{15}(I_5)}(x) \\
&= \frac{1}{2} f_1 \chi_{[0,0.4]}(x) + \frac{1}{3} f_2 \chi_{(0.4,1]}(x) + \frac{1}{2} f_3 \chi_{(0.2,0.6]} \\
&+ \frac{1}{2} f_4 \chi_{(0,0.4]} + \frac{1}{2} f_5 \chi_{(0.6,1]}.
\end{aligned} \tag{5.13}
$$

Consider the following cases in (5.13):

1) when $x \in [0, 0.2]$

$$P_{T_1} f(x) = \frac{1}{2} f_1 + 0 + 0 + \frac{1}{2} f_4 + 0. \tag{5.14}$$

2) when $x \in (0.2, 0.4]$

$$P_{T_1} f(x) = \frac{1}{2} f_1 + 0 + \frac{1}{2} f_3 + \frac{1}{2} f_4 + 0. \tag{5.15}$$

3) when $x \in (0.4, 0.6]$

$$P_{T_1} f(x) = 0 + \frac{1}{3} f_2 + \frac{1}{2} f_3 + 0 + 0. \tag{5.16}$$

4) when $x \in (0.6, 0.8]$

$$P_{T_1} f(x) = 0 + \frac{1}{3} f_2 + +0 + \frac{1}{2} f_5. \tag{5.17}$$

5) when $x \in (0.8, 1]$

$$P_{T_1} f(x) = 0 + \frac{1}{3} f_2 + +0 + \frac{1}{2} f_5. \tag{5.18}$$

The matrix representation of the F-P operator, using (5.14), (5.15), (5.16), (5.17), and (5.18), is

$$
P_{T_1} f(x) = [f_1 \ f_2 \ f_3 \ f_4 \ f_5] \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{5.19}
$$

The F-P operator matrix associated with T_1 is

$$M_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{5.20}$$

Similarly, the F-P operator matrix associated with T_2 is

$$M_2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \end{bmatrix}. \tag{5.21}$$

From (5.2),

$$\begin{aligned} \sum_{k=1}^{2} h_k(x) &= \sum_{k=1}^{2} \left(\sum_{i=1}^{5} \frac{p_{ki}}{|T'_{ki}(x)|} \right) \\ &= \frac{p_{11}(x)}{|T'_{11}(x)|} + \frac{p_{12}(x)}{|T'_{12}(x)|} + \frac{p_{13}(x)}{|T'_{1}(x)|} + \frac{p_{14}(x)}{|T'_{14}(x)|} + \frac{p_{15}(x)}{|T'_{15}(x)|} \\ &+ \frac{p_{21}(x)}{|T'_{21}(x)|} + \frac{p_{22}(x)}{|T'_{22}(x)|} + \frac{p_{23}(x)}{|T'_{23}(x)|} + \frac{p_{24}(x)}{|T'_{24}(x)|} + \frac{p_{25}(x)}{|T'_{25}(x)|}, \end{aligned}$$

1) when $x \in [0, 0.2]$,

$$\sum_{k=1}^{2} h_k(x) = \frac{p_{11}(x)}{|T'_{11}(x)|} + \frac{p_{21}(x)}{|T'_{21}(x)|} = \frac{\frac{3}{4}}{2} + \frac{\frac{1}{4}}{2} = \frac{4}{8} < 1,$$

2) when $x \in (0.2, 0.4]$

$$\sum_{k=1}^{2} h_k(x) = \frac{3}{8} < 1,$$

3) when $x \in (0.4, 0.6]$

$$\sum_{k=1}^{2} h_k(x) = \frac{4}{8} < 1,$$

4) when $x \in (0.6, 0.8]$

$$\sum_{k=1}^{2} h_k(x) = \frac{4}{8} < 1,$$

5) when $x \in (0.8, 1]$

$$\sum_{k=1}^{2} h_k(x) = \frac{3}{8} < 1.$$

Since $\sum_{k=1}^{2} h_k(x) < 1$ for all $x \in [0,1]$, τ_1 satisfies the condition (5.3) and (5.4). By theorem 4.5 of [4], τ_1 has an invariant measure absolutely continuous with respect to Lebesgue measure.

An invariant density $g = [g_1\ g_2\ g_3\ g_4\ g_5]$ of the F-P operator associated with the map $T_1(x)$ is obtained by solving

$$[g_1\ g_2\ g_3\ g_4\ g_5]M_1 = [g_1\ g_2\ g_3\ g_4\ g_5]$$

$$[g_1\ g_2\ g_3\ g_4\ g_5]\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [g_1\ g_2\ g_3\ g_4\ g_5] \tag{5.22}$$

Solving the system (5.22) gives:

$$g = g_1\left[1\ \ \frac{3}{2}\ \ 1\ \ \frac{5}{6}\ \ \frac{2}{3}\right].$$

Normalizing g

$$g_1\left(\int_0^{0.2} dx + \frac{3}{2}\int_{0.2}^{0.4} dx + \int_{0.4}^{0.6} dx + \frac{5}{6}\int_{0.6}^{0.8} dx + \frac{2}{3}\int_{0.8}^{1} dx\right) = 1$$

$$g = \frac{1}{5}\left[1\ \ \frac{3}{2}\ \ 1\ \ \frac{5}{6}\ \ \frac{2}{3}\right].$$

The normalized invariant density of the F-P operator associated with the map $T_2(x)$ is

$$\frac{10}{13}\left[1\ \ 1\ \ 1\ \ 2\ \ \frac{3}{2}\right].$$

The F-P operator of the random map $\tau_1 = \{T_1, T_2; p_1, p_2\}$ [4] is

$$M_{\tau_1} = \begin{bmatrix} \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{8} & \frac{1}{8} & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{8} & \frac{1}{8} & 0 & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}.$$

Let $f^* = [f_1^* \ f_2^* \ f_3^* \ f_4^* \ f_5^*]$ be an invariant density of the F-P operator associated with the random map $\tau_1 = \{T_1, T_2; p_1, p_2\}$ satisfying $f^* M_{\tau_1} = f^*$. The invariant density is in the form:

$$f^* = \frac{1}{1608} f_3^* \, [1968 \ \ 2169 \ \ 1608 \ \ 2721 \ \ 2475].$$

After normalization,

$$f^* = [0.8994 \ \ 0.9912 \ \ 0.7349 \ \ 1.2435 \ \ 1.1310].$$

Example 5.5.2. [3] Consider the maps $T_1, T_2 : [0,1] \to [0,1]$ defined by

$$T_1(x) = \begin{cases} 2x, & 0 \le x < 0.5 \\ 2x - 0.3, & 0.5 \le x < 0.6 \\ x + 0.1, & 0.6 \le x < 0.7 \\ x & 0.7 \le x \le 1 \\ & \end{cases}, \ T_2(x) = \begin{cases} x, & 0 \le x < 0.1 \\ x - 0.1, & 0.1 \le x < 0.3 \\ x - 0.2, & 0.3 \le x < 0.5 \\ 2x - 0.9 & 0.7 \le x \le 0.9 \\ 2x - 0.1 & 0.9 \le x \le 1 \end{cases}.$$

Let p_1 and p_2 be the probabilities corresponding to T_1 and T_2 respectively,

$$p_1(x) = \begin{cases} 0.8, & \le x < 0.5 \\ 0.725, & 0.5 \le x < 0.7 \\ 0.4, & 0.7 \le x < 1 \end{cases}, \ p_2(x) = \begin{cases} 0.2, & 0 \le x < 0.5 \\ 0.275, & 0.5 \le x < 0.7 \\ 0.6, & 0.7 \le x < 1 \end{cases}.$$

The normalized invariant density, $f^* = [f_1^* \ f_2^* \ f_3^* \ f_4^* \ f_5^* \ f_6^* \ f_7^* \ f_8^* \ f_9^* \ f_{10}^*]$, of the F-P operator associated with the random map $\tau_2 = \{T_1, T_2; p_1, p_2\}$ is

$f^* = [0.11591 \ 0.23183 \ 0.48548 \ 0.44184 \ 0.19419 \ 1.28694 \ 1.26949 \ 3.64250 \ 2.07290 \ 0.25892]$.

Theorem 5.5.3. *Let B, D be two row stochastic matrices of order n. If A, C are diagonal matrices of order n with the corresponding entries sum to unity, then*

$$AB + CD$$

is a row stochastic matrix.

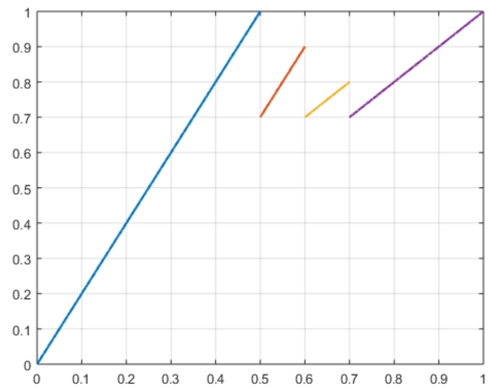

Figure 5.5: Graph of $T_1(x)$

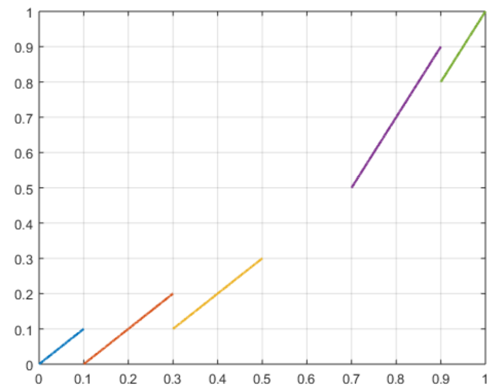

Figure 5.6: Graph of $T_2(x)$

Let $A = \{a_{ij}\}_{i,j=1}^{n}, a_{ij} = 0, \;\; i \neq j$, $B = \{b_{ij}\}_{i,j=1}^{n}$, $C = \{c_{ij}\}_{i,j=1}^{n}, c_{ij} = 0, \;\; i \neq j$, and $D = \{d_{ij}\}_{i,j=1}^{n}$ such that $\sum_{j=1}^{n} b_{ij} = \sum_{j=1}^{n} d_{ij} = 1$, for all $i = 1, 2, \ldots, n$. The i, jth entry of the matrix AB is $\{a_{ii}b_{ij}\}$ for all $i, j = 1, 2, \ldots, n$ and the i, jth entry of the matrix CD is $\{c_{ii}d_{ij}\}$ for all $i, j = 1, 2, \ldots, n$. Consequently, the i, jth entry of the matrix $AB + CD$ is the $\{a_{ii}b_{ij} + c_{ii}d_{ij}\}$ for all $i, j = 1, 2, \ldots, n$. The sum of entries in the ith row of the matrix $AB + CD$ is

$$\begin{aligned}
\sum_{j=1}^{n} a_{ii}b_{ij} + c_{ii}d_{ij} &= a_{ii}b_{i1} + a_{ii}b_{i2} + \cdots + a_{ii}b_{in} + c_{ii}d_{i1} + c_{ii}d_{i2} + \cdots + c_{ii}d_{in} \\
&= a_{ii}(b_{i1} + b_{i2} + \cdots + b_{in}) + c_{ii}(d_{i1} + d_{i2} + \cdots + d_{in}) \\
&= a_{ii} + c_{ii} \\
&= 1, \; \forall i = 1, 2, \ldots, n.
\end{aligned}$$

Hence $AB + CD$ is a row stochastic matrix.

Corollary 5.5.4. *Let B, D be two column stochastic matrices of order n. If E, F are diagonal matrices of order n with the corresponding entries sum to unity, then*

$$BE + DF$$

is a column stochastic matrix.

Example 5.5.5. [39] Take a position dependent random map $\tau_3 = \{T_1, T_2; p_1, p_2\}$, where

$T_i : [0,1] \to [0,1], i = 1,2$ defined by

$$T_1(x) = \begin{cases} 3x + \frac{1}{4}, & 0 \leq x < \frac{1}{4}, \\ 3x - \frac{3}{4}, & \frac{1}{4} \leq x < \frac{1}{2}, \\ 4x - 2, & \frac{1}{2} \leq x < \frac{3}{4}, \\ 4x - 3, & \frac{3}{4} \leq x \leq 1 \end{cases},$$

$$T_2(x) = \begin{cases} 4x, & 0 \leq x < \frac{1}{4}, \\ 4x - 1, & \frac{1}{4} \leq x < \frac{1}{2}, \\ 3x - \frac{3}{2}, & \frac{1}{2} \leq x < \frac{3}{4}, \\ 3x - \frac{9}{4}, & \frac{3}{4} \leq x \leq 1 \end{cases}.$$

The probability functions are defined by $p_i : [0,1] \to [0,1], i = 1,2,$

$$p_1(x) = \begin{cases} \frac{1}{4}, & 0 \leq x < \frac{1}{4} \\ \frac{1}{4}, & \frac{1}{4} \leq x < \frac{1}{2} \\ \frac{3}{4}, & \frac{1}{2} \leq x < \frac{3}{4} \\ \frac{3}{4}, & \frac{3}{4} \leq x \leq 1 \end{cases},$$

$$p_2(x) = \begin{cases} \frac{3}{4}, & 0 \leq x < \frac{1}{4} \\ \frac{3}{4}, & \frac{1}{4} \leq x < \frac{1}{2} \\ \frac{1}{4}, & \frac{1}{2} \leq x < \frac{3}{4} \\ \frac{1}{4}, & \frac{3}{4} \leq x \leq 1 \end{cases}.$$

The F-P operator matrix of the random map $\tau_3 = \{T_1, T_2; p_1, p_2\}$ is

$$M_{\tau_3} = \begin{bmatrix} \frac{3}{16} & \frac{13}{48} & \frac{13}{48} & \frac{13}{48} \\ \frac{13}{48} & \frac{13}{48} & \frac{13}{48} & \frac{3}{16} \\ \frac{13}{48} & \frac{13}{48} & \frac{13}{48} & \frac{3}{16} \\ \frac{13}{48} & \frac{13}{48} & \frac{13}{48} & \frac{3}{16} \end{bmatrix}.$$

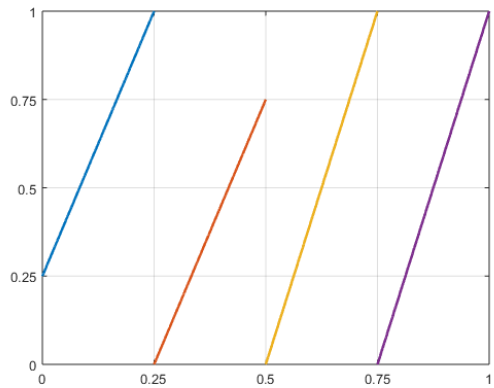

Figure 5.7: Graph of $T_1(x)$

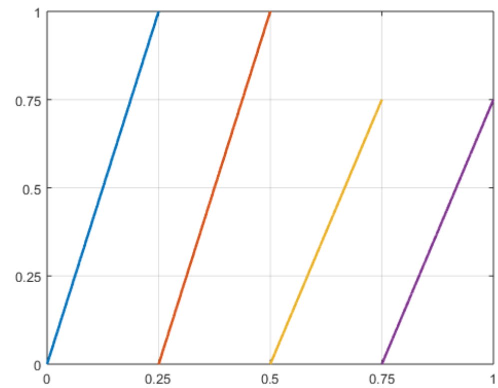

Figure 5.8: Graph of $T_2(x)$

The invariant density, $f = [f_1\ f_2\ f_3\ f_4]$, of τ_3 satisfying $fM_{\tau_3} = f$, after normalization, is

$$f = \begin{bmatrix} 1 & \frac{13}{12} & \frac{13}{12} & \frac{5}{6} \end{bmatrix}.$$

5.5.1 Error Calculation

In MEM, an invariant density of the F-P operator associated with a random map is approximated by the formula

$$f_n(x) = e^{\sum_{k=0}^{n} \lambda_k \phi_k(x)},$$

given in (5.11). The numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ satisfy the relations (3.5). The L^1 errors between the exact, f^*, and approximate, f_n, invariant densities of the F-P operator associated with a random map $\tau_i,\ i = 1, 2, 3$ is defined by

$$e_n = \|f_n - f^*\|_{L^1(0,1)} = \int_0^1 |f_n(x) - f^*(x)|\, dx$$

. The L^1 errors associated to the random map are given below. The numerical work has been done by MATLAB© on a 64-bit OS and 16 GB memory with Intel(R) Core(TM) i7-4770K CPU @ 3.50GHz processor.

n	Error	CPU Time (Sec)
4	1.7700×10^{-02}	14.682806
8	7.5000×10^{-03}	14.627364
16	7.600×10^{-03}	14.777599
32	2.600×10^{-03}	15.020649
64	1.300×10^{-03}	15.253307
128	4.7164×10^{-04}	15.507617
256	4.5027×10^{-04}	16.802797
512	1.6036×10^{-04}	19.223928
1024	7.9284×10^{-05}	25.052764

Table 5.1: Errors, e_n, Associated with τ_1

n	Error	CPU Time (Sec)
4	2.4300×10^{-02}	14.342839
8	6.810×10^{-02}	28.820306
16	4.990×10^{-02}	28.784264
32	2.700×10^{-02}	30.018860
64	9.000×10^{-03}	29.466120
128	5.800×10^{-03}	30.209280
256	3.200×10^{-03}	31.907281
512	1.700×10^{-03}	35.226260
1024	5.5179×10^{-04}	41.971688

Table 5.2: Errors, e_n, Associated with τ_2

n	Error	CPU Time (Sec)
4	3.1000×10^{-03}	0.044880
8	3.2000×10^{-03}	0.067648
16	1.7000×10^{-03}	0.126008
32	8.4117×10^{-04}	0.236896
64	4.1393×10^{-04}	0.500686
128	2.0534×10^{-04}	0.980884
256	1.0227×10^{-04}	1.917636
512	5.1033×10^{-05}	3.408522
1024	2.5491×10^{-05}	8.764972

Table 5.3: Errors, e_n, Associated with τ_3

Chapter 6

Conclusion and Future Research

6.1 Conclusion

The dissertation focuses on approximating the invariant densities of the F-P operator associated with various nonsingular maps defined from $[0,1]$ to itself and the Markov operator associated with random maps, by using the maximum entropy method combined with the finite elements idea. The $L^1(0,1)$ errors between the exact and approximate invariant densities of the F-P operator are presented with necessary theoretical works. The finite elements are the uniform divisions of the interval $[0,1]$ into the subintervals. Piecewise polynomial functions having partition of unity property are defined on such subintervals. The MEMs based on piecewise quadratic functions, quadratic splines and piecewise linear polynomials are applied on the finite elements of $[0,1]$. The estimated errors between exact and approximate invariant densities based on piecewise quadratic functions are presented in Chapter 3. Similar results based on quadratic splines are presented in Chapter 4. Chapter 5 introduces position dependent random maps defined from the unit interval $[0,1]$ to itself. It also gives an overview of how to use the MEM based on piecewise linear functions in error estimation.

When a unique invariant density f^* has a continuous third order derivative on $[0,1]$, the numerical and theoretical results of PQMEM show that the order of convergence of the L^1 errors is $O(h^3)$, but for an invariant density which is unbounded on $[0,1]$ the proposed method does not work well. In this case, the order of L^1 error is not of order $O(h^3)$. The method gives better results in comparison to the results from PQMFA and PLMEM, but due to the lack of differentiability of the moment functions $\{\phi_k\}_{k=0}^{2n}$, a large number of moment functions are needed. This is a drawback of this method.

Quadratic B-splines are used in MEM, which needs only $n+1$ moments. The number is nearly half of the moment functions, $2n+1$, needed in PQMEM. QSMEM is used in error estimation between the exact and approximate invariant densities of the F-P operator associated with different nonsingular maps defined from $[0,1]$ to itself. A theoretical discussion is presented to show the L^1 norm convergence rate of this method.

When a unique invariant density f^* exists and it has a continuous third derivative on $[0,1]$, the convergence rate of the L^1 errors is $O(h^3)$ unless the invariant density is unbounded on $[0,1]$. The slow convergence for such an invariant density is due to the singularity of the density at the end points of the interval $[0,1]$.

Both methods have the same convergence rate. Since QSMEM needs a fewer number of piecewise quadratic functions than PQMEM does, the former method is faster than the latter.

Piecewise linear functions have been utilized in the MEM for the error estimation between exact and approximate invariant densities of the F-P operator associated with random maps defined from the unit interval $[0,1]$ to itself. The L^1 errors between the exact invariant densities and approximated densities are presented in tabular forms. The numerical results show that the convergence rate is of $O(h)$. The low convergence rate can be improved by applying higher order piecewise polynomials or splines or other suitable functions.

6.2 Future Work

The dissertation shows that MEM based on piecewise linear and quadratic polynomials can be used to approximate invariant densities of the F-P operator associated to various nonsingular maps defined from $[0,1]$ to itself. The following research can be done in the future:

- Using numerical methods other than Newton-Raphson method to solve nonlinear systems. Applying different integration techniques in place of the Gaussian 3-point Quadrature method.
- Applying higher degree polynomials and higher order splines in the MEM to develop a fast convergence method.
- Developing a fast numerical scheme, based on a rigorous theory, to calculate moments from given maps.
- Investigating a homogeneous maximum entropy method (HMEM) based on piecewise polynomials with a faster convergence rate than in the original HMEM [20] which used the monomials $1, x, x^2, \cdots$ for the moment functions and which had a slow convergence rate.
- Extending MEM to higher dimensional spaces.
- Investigating new criteria for the existence of invariant densities of position dependent random maps.

- Developing fast and efficient numerical methods to calculate moments of position dependent random maps.

BIBLIOGRAPHY

[1] Aston P. J., Dellnitz, M.: *The computation of Lypunov exponents via spatial integration with application to blowout bifurcation*. Comput. Methods in Appl. Mech. Eng. **170** (1999) 223-237.

[2] Aston, P. J., Junge, O.: *Computing the invariant measure and the Lyapunov exponent for one-dimensional maps using a measure-preserving polynomial basis*, Math. Comp., **83** (288) (2014) 1869-1902.

[3] Bahsoun, W., Góra, P., Mayoral, S., Morales, M.: *Random dynamics and finance: Constructing implied binomial trees from a predetermined stationary density*, Appl. Stochastic Models Bus. Ind., **23** (2007) 181-212.

[4] Bahsoun, W., Góra, P.: *Position dependent random maps in one and higher dimensions*, Studia Math., **166** (3) (2005) 271-286.

[5] Bahsoun, W., Bose, C., and Quas, A.: *Deterministic representation for position dependent random maps* , DCDS., **22** (3) (2008) 529-540.

[6] Barnsley, M.: *Fractals Everywhere*, 2nd Ed. Academic Press, London (1993).

[7] Beck, C., Schlögl, F.: *Thermodynamics of Chaotic Systems*. Cambridge University Press (1993).

[8] Bhattacharya, R., Majumdar, M. : *On a Class of Stable Random Dynamical Systems: Theory and Applications*, J. Econ. Theory, **96** (2001) 2008-229.

[9] Biswas, P., Shimoyama, H., and Mead, L. R.: *Lyapunov exponent and natural invariant density determination of chaotic maps: An iterative entropy ansatz,* J. Phys. A **43**(12) (2010) 103-125.

[10] Bolt, E. M., Santitissadeekorn, N.: *Applied and Computational Measurable Dynamics*, SIAM (2013).

[11] de Boor, C.: *A practical Guide to Splines*, Revised Edition, Springer, (2001).

[12] Borwein, J. M., Lewis, A. S.: *Convergence of the best entropy estimates*, SIAM J. Optim. **1** (2) (1991) 191-205.

[13] Borwein, J. M., Lewis, A. S.: *On the convergence of moment problems*, Trans. Amer. Math. Soc. **325** (1) (1991) 249-271.

[14] Boss, C., Murray, R.: *The exact rate of approximation in Ulam's method*, Discrete Contin Dyn Syst, **7**(1)(2001) 219-235

[15] Boyarsky, A., Góra, P.: *Laws of Chaos*, Birkhaüser, (1997).

[16] Choe, G. H.: *Computational Ergodic Theory, Algorithms and Computation in Mathematics* , Springer-Verlag, Berlin, vol.13, (2005).

[17] Devaney, R. L.: *An Introduction to Chaotic Dynamical Systems*, Westview Press Boulder, Colorado, USA, 2nd Ed. (2003).

[18] Ding, J., Li, T. Y.: *Markov finite approximation of Frobenius-Perron operator*, Nonlinear Anal., **17** (8) (1991) 759-772.

[19] Ding, J., Zhou, A.: *Finite approximations of Frobenius-Perron operators: a solution of Ulam's conjecture to multi-dimensional transformations*, Physica D. **92** (1-2) (1996) 61-68.

[20] Ding, J.: *A maximum entropy method for solving Frobenius-Perron operator equations*, Appl. Math. Comput., **93** (1998) 155-168.

[21] Ding, J., Li, T. Y.: *A convergence rate analysis for Markov finite approximations to a class of Frobenius-Perron operators*, Nonlinear Anal-Theor, **31**(5/6)(1998) 765-776.

[22] Ding, J, Li, T. Y., and Zhou, A.: *Finite approximations of Markov operators*, J. Comput. Appl. Math., **147** (2002) 137-152.

[23] Ding, J., Mead, L. R.: *Maximum entropy approximation for the Lyapunov exponents of chaotic maps*, J. Math. Phys. **43** (2002) 2518-2522.

[24] Ding, J., Mead, L. R.: *The maximum entropy method applied to stationary density computation*, Appl. Math. Comput., **185** (2007) 658-666.

[25] Ding, J., Jin, C., Rhee, N. H. and Zhou, A.: *A maximum entropy method based on piecewise linear functions for the recovery of a stationary density of interval mapping*, âĂŐJ. Stat. Phys. **145**(2011) 1620-1639.

[26] Ding, J., Rhee, N.: *A maximum entropy method based on orthogonal polynomials for Frobenius-Perron operators*, Adv. Appl. Math. Mech. **3** (2)(2011) 204-218.

[27] Ding, J., Rhee, N. H.: *On the norm convergence of a piecewise linear least squares method for Frobenius-Perron operators*, J. Math. Anal. Appl., **366** (1) (2012) 91-102.

[28] Ding, J., Rhee, N.: *Birkhoff's ergodic theorem and the piecewise maximum entropy method for Frobenius-Perron operators*, Int. J. Comput. Math., **89** (8) (2012) 1083-1091.

[29] Ding, J., Rhee, N.: *A unified maximum entropy method via spline functions for Frobenius-Perron operators*, Numer. Algebra Contr. Optim, **3** (2)(2013) 235-245.

[30] Edward, O.: *Chaos in Dynamical Systems*, Cambridge University Press, 2nd Ed. (2002).

[31] Eric, L.: *Mathematics for 3D Game Programming and Computer Graphics*, Charles River Media, Inc., Massachusetts, 2nd Ed. (2004).

[32] Feigenbaum, M.J.: J. Statist. Phys,**19** (1978) 25-52 .

[33] Froyland, G.: *Finite approximation of Sinai-Bowen-Ruelle measures for Anosov systems in two dimensions*, Random and Computational Dynamics. **3** (4) (1995) 251-264.

[34] Gerasoulis, A.: *The use of piecewise quadratic polynomials for the solution of singular equations of Cauchy type*, Computers & Mathematics with Applications, **8** (1982) 15-22.

[35] Giffin, A.: *Maximum entropy: The universal method for inference (PhD dissertation)*, University at Albany, State University of New York, (2008).

[36] Góra, P., Boyarsky, A.: *Why computers like Lebesgue measure*, Comput Math Appl., **16** (1988) 321-329.

[37] Góra, P., Boyarsky, A.: *Absolutely continuous invariant measures for random maps with position dependent probabilities*, J. Math. Anal. Appl. **278** (2003) 225-242.

[38] Hausdorff, F.: *Summationsmethoden und momentfolgen*, I, Math. Z., **9** (1921) 74-109.

[39] Islam, S. M: *Invariant measures for random maps via interpolation*, Int. J. Bifurc. Chaos, **23** (2) (2013) 1350025 (1-8).

[40] Jablonski, M.: *On invariant measures for piecewise C^2- transformations of the n-dimensional cube*, Annales Polonici Mathematici, **43** 2 (1983) 185-195.

[41] Jaynes, E. T.: *Information theory and statistical mechanics*, Phys. Rev.**106** (1957) 620-630.

[42] Johansson, M.: *Piecewise Linear Control systems*, Ph. D. thesis, Department of Automatic Control. Lund Institute of Technology, Sweden (1999).

[43] Katok, A., Hasselblatt, B.: *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, New York (1996).

[44] Kay, S. M., Marple,S. L.: *Spectrum analysis - a modern perspective*, Proc. IEEE. **69** (1981) 1380-1419.

[45] Krzyzewski, K., Szlenk, W.: *On invariant measures for expanding differentiable mappings* , Studia Mathematica , **33** (1969) 82-92.

[46] Lang, S. W., McCellan, J. H.: *Spectral estimation for sensor arrays*, IEEE Trans. Acoust. Speech Signal Process. **31** (1983) 349-358.

[47] Lasota, A., Macke. M.: Probabilistic Properties of Deterministic Systems, Cambridge University Press, New York (1985).

[48] Lasota, A., Macke. M.: *Chaos, Fractals and Noise*, 2nd Ed. Springer, New York, (1994).

[49] Lasota. A., Yorke. J. A.: *On the existence of invariant measure for piecewise monotonic transformations*, Trans. Amer. Math. Soc., **186** (1973) 481-488.

[50] Lasota, A., Macke, M.: *Chaos, Fractals and Noise, 2nd Ed.*, Springer, New York (1994).

[51] Li. T. Y.: *Finite approximation for the Frobenius-Perron operator, a solution to Ulam's conjecture*, J. Approx. Theory, **17** :2 (1976) 177-186.

[52] Li, T.Y., Yorke, J. A.: *Ergodic transformations from an interval into itself*. Trans. Amer. Math. Soc.**235** (1978) 177-186.

[53] Lorenz, E. N.: *Deterministic nonperiodic flow J. Atmospheric Sci.*, **20** (1963) 130-141.

[54] Kohda, G., Murao, K.: *Piecewise polynomial Galerkin approximation to invariant densities of one-dimensional difference equations*, Electron. Comm. Jpn. 1, **65** :6 (1982) 1-11.

[55] Rasmussen, M.: *Attractivity and Bifurcation for Nonautonomous Dynamical System*, Springer-Verlag, Berlin Heidelberg, (2007).

[56] Mead, L., Papanicolaou, N.: *Maximum entropy in the problem of moments*, J. Math. Phys. **25** (1984) 2404-2417.

[57] Meyer, C. D.: *Matrix analysis and applied linear algebra*, SIAM (2000).

[58] Pelikan, S.: *Invariant densities for random maps of the interval*, Transactions of the AMS. **281** (2) (1984) 813-825.

[59] Poincaré, J. H.: Les mé thodes nouvelles de la mécanique céleste *Gauthiers-Villars*, volume 1-3 (1892), (1893), (1899).

[60] Richard, H. A.: *A First Course in Discrete Dynamical Systems*. Springer-Verlag, 1st Ed. (1994).

[61] Robinson, C.: *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, CRC Press, (1994).

[62] Rudnicki, R.: *Markov operators: Applications to diffusion processes and population dynamics*, Appl. Math., **27** (1)(2000) 67-79.

[63] Rychlik, M.: *Invariant measures and the variational principle for Lozi mappings*, Ph. D. Thesis, University of California, Berkeley, (1983).

[64] Shannon, C. E.: *A mathematical theory of communication Bell System Technical Journal*, **27** (1948) 379-423 and 623-656, July and October.

[65] Schumaker, L. L.: *Spline functions basic theory, Pure and applied Mathematics*-A Wiley interscience series of texts, monographs & tracts, John wiley & sons New York, (1980).

[66] Solow, R. M.: *A Contribution to the Theory of Economic Growth* Q. J. Econ. **70**(1)(1956) 65-94.
Published by: The MIT Press Stable URL: http://www.jstor.org/stable/1884513

[67] Janke, S. J.: *Mathematical structures for computer graphics*, Hoboken, New Jersey : John Wiley & Sons, Inc. (2015).

[68] Uffink, J.: *Boltzmann's work in statistical Physics*, Stanford Encyclopedia of Philosophy, (2004).

[69] Ulam, S. M.: *A Collection of Mathematical Problems*, Interscience Tracts in Pure and Applied Mathematics Vol. 8, Interscience, New York, (1960).

[70] Upadhyay, T. , Ding, J., and Rhee, N.: *A piecewise quadratic maximum entropy method for the statistical study of chaos*, J. Math. Anal. Appl. **421** (2015) 1487-1501.

[71] Wong, S.: *Some metric properties of piecewise monotonic mappings of the unit interval*. Trans. Amer. Math. Soc. **246** (1978) 493-500.

[72] Wu, N.: *The Maximum Entropy Method*, Springer (1997).