Application of Fractal Interpolation in Kernel Regression Estimation

Modeling Workshop Project Report

Course Code: MA5770

Submitted by

Amar Kumar

MA24M002

20 May 2025

Guided by

Prof. AKB Chand

Department of Mathematics

Indian Institute of Technology Madras

Declaration

I hereby declare that the work presented in this report titled **"Applications of Fractal Interpolants in Kernel Regression Estimations"** is my original research and has been carried out under the guidance of my supervisor.

This work has not been submitted to any other university or institution for the award of any degree, diploma, or certificate.

All sources of information and data used in this report have been acknowledged appropriately through proper references. Any contribution of other researchers or collaborators has been duly recognized.

Place: Chennai
Date: 20 May 2025

Amar Kumar
MA24M002
Department of Mathematics
Indian Institute of Technology Madras

Acknowledgements

First and foremost, I would like to express my sincere gratitude to the authors of the paper titled *"Applications of Fractal Interpolants in Kernel Regression Estimations"* — Dr. Chiao-Wen Liu and Dr. Dah-Chin Luor — whose research served as the foundation and inspiration for this work. Their insights into fractal geometry and nonparametric modeling have greatly influenced the direction of this study.

I am deeply grateful to my academic supervisor for their valuable guidance, continuous support, and constructive feedback throughout the course of this project. Their encouragement and expertise played a crucial role in shaping the theoretical and experimental aspects of this work.

I also extend my thanks to the faculty and staff of the Department of Mathematics at my institution for providing a stimulating and supportive research environment.

Special thanks go to the open data platforms that made real-world datasets such as the Crude Oil WTI Futures prices accessible for experimentation and validation.

Last but not least, I am thankful to my family and friends for their unwavering support, patience, and motivation during the course of this research journey.

Contents

List of Figures

List of Tables

Abstract

This study explores the application of smooth fractal interpolation functions (FIFs) to enhance the performance of kernel regression estimators. Traditional nonparametric regression methods, such as the Priestley–Chao estimator, often struggle to capture complex, self-similar patterns present in real-world data. To address this limitation, we introduce a fractal perturbation framework that augments classical estimators with smooth fractal corrections, guided by the theory of iterated function systems [2, 1] and smooth FIF constructions [4, 8].

The proposed method constructs a C^p-smooth fractal function by perturbing the kernel estimator using a controlled smooth auxiliary function. Optimal scaling factors and kernel parameters are determined using the SLSQP optimization algorithm to minimize the empirical mean squared error (MSE). Experimental results on a financial time-series dataset demonstrate that the fractal-enhanced estimators consistently outperform traditional ones in terms of accuracy.

This work builds upon prior developments in fractal regression [7, 6] and spline-based interpolation [10], while also contributing to the growing body of research on adaptive kernel methods [5] and fractal-based modeling in data science [9, 11, 3]. The integration of smooth fractal geometry into statistical regression provides a novel pathway for modeling highly nonlinear data with limited observations.

Chapter 1

Introduction

1.1 Background

In data fitting and regression analysis, constructing a function from a finite set of samples is a fundamental task. Traditional interpolation and nonparametric modeling methods, such as kernel regression estimators (e.g., Nadaraya–Watson, Priestley–Chao, Gasser–Müller), have been widely used to approximate unknown functions from observed data [5]. These methods provide flexibility and robustness, especially when no prior assumptions are made about the underlying data distribution.

Meanwhile, *fractal interpolation functions (FIFs)* have emerged as powerful tools for modeling complex, irregular data. Originating from the work of Barnsley and Demko [2, 1], FIFs are based on self-similarity and extend classical interpolation techniques to better capture nonlinear and non-smooth behavior in datasets.

Smooth variants of FIFs, such as those proposed by Chand and Kapoor [4] and Navascués et al. [8], have enabled their integration into more sophisticated estimation tasks. These advances allow traditional regression techniques to be augmented by geometrically rich, fractal-based representations.

1.2 Problem Statement

Despite the effectiveness of kernel regression estimators in various applications, their performance can degrade when dealing with highly nonlinear or fractal-like data patterns. The standard estimators may fail to capture the intricate fluctuations in such datasets, resulting in increased approximation error.

The problem addressed in this work is: **Can we improve the accuracy of nonparametric kernel regression estimations by incorporating fractal characteristics into the regression function?** Specifically, this study investigates the *fractal perturbation of kernel regression estimators* using smooth FIFs, as proposed in recent work by Liu and Luor [6, 7], with the goal of reducing the empirical mean squared error (MSE).

1.3 Motivation

Many real-world datasets, especially in domains like finance, biology, and signal processing, exhibit complex, irregular, or self-similar behavior that cannot be accurately captured by smooth regression curves. The motivation for this work stems from the limitations of traditional estimators when applied to such data and the increasing interest in fractal-based modeling as surveyed by Navascués et al. [9] and inspired by the broader fractal geometry movement initiated by Mandelbrot [3].

By leveraging the theory of smooth FIFs, it is possible to construct regression models that are both smooth and self-similar, offering a better fit to real-world complexities.

1.4 Objectives

The primary objectives of this work are as follows:

- To explore the construction of smooth fractal interpolation functions that satisfy certain smoothness conditions [4, 8].
- To apply these functions as perturbations to traditional kernel regression estimators, focusing on the Priestley–Chao estimator [6].
- To compare the performance of standard and fractal-enhanced estimators using empirical MSE as the evaluation metric.
- To optimize hyperparameters (such as smoothing parameter d and scaling factors s_i) using the SLSQP algorithm.
- To demonstrate the effectiveness of the proposed approach on a real-world financial time-series dataset.

1.5 Solution Approach

The solution proposed in this study is the *fractal perturbation of the Priestley–Chao kernel regression estimator* using smooth FIFs. This idea follows from the work of Luor and Liu [7], which showed how fractal geometry can be used to enhance traditional estimators like Nadaraya–Watson.

The methodology involves:

- Constructing C^p-smooth FIFs based on a given function and a set of data points [4, 8].
- Introducing a function $r = \hat{g} - \xi$, where $\hat{g}$ is the base estimator and ξ is a smooth guiding function.
- Using the SLSQP optimization algorithm to determine the optimal values of the fractal scaling factors s_i and kernel parameter d.
- Comparing the regression results of the standard Priestley–Chao estimator and its fractal variants (C_0-FIF, C_1-FIF, and C_2-FIF).

Summary

This chapter introduced the motivation and problem addressed in the study: the enhancement of kernel regression estimators via smooth fractal interpolation. It reviewed key concepts and prior work, setting the stage for the methodology and experiments that follow.

Chapter 2

Basic Tools Used in the Work

2.1 Machine Learning Framework

Summary

Chapter 3

Theory and Methodology

3.1 Theoretical Foundations

The central theme of this work is the enhancement of traditional kernel regression estimators using smooth *fractal interpolation functions* (FIFs). Fractal functions, initially introduced by Barnsley, are known for their self-similar structure and are useful in modeling data with irregular or complex patterns.

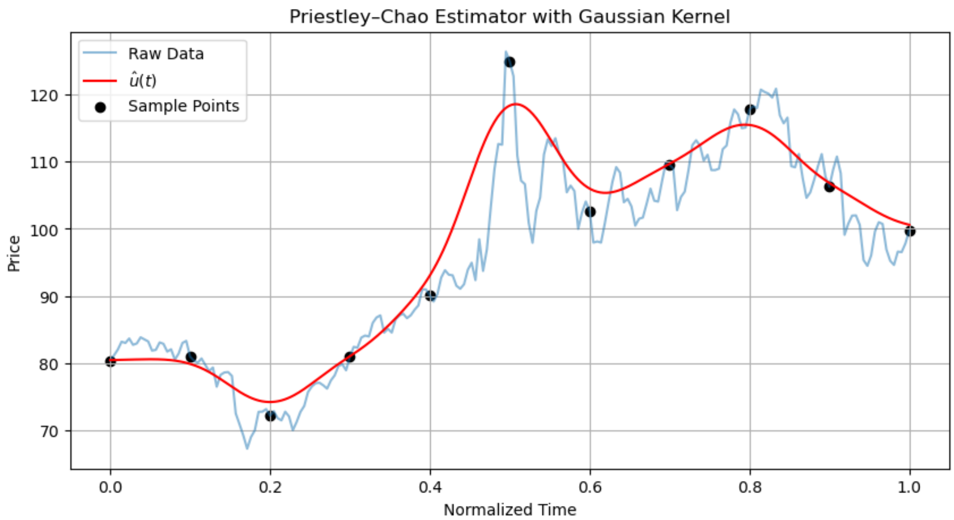

Figure 3.1: Priestley-Chao Estimator with Gaussian Kernel

Kernel regression estimators, such as the Nadaraya–Watson and Priestley–Chao estimators, are nonparametric tools widely used in statistics for function approximation. These estimators construct a smooth regression curve by weighting nearby observations based on a kernel function. However, when the underlying data exhibits intricate, possibly self-similar structures, these estimators might fail to capture finer details.

To overcome this, the study introduces **fractal perturbations** to the Priestley–Chao estimator using smooth FIFs. The perturbations are constructed such that the regression function inherits both the smoothness and the self-similar properties of fractals.

3.2 Smooth Fractal Interpolation Functions

Let $I = [0, 1]$ and suppose we have a dataset $\mathcal{D} = \{(t_i, y_i)\}_{i=0}^{N}$ where $t_0 < t_1 < \cdots < t_N = 1$. Given a base function $u \in C^p[I]$ that interpolates the dataset, we define a smooth FIF f as the unique fixed point of a collection of contraction mappings:

$$f(t) = M_i(t, f(t)) = s_i f(t) + u(L_i(t)) - s_i r(t), \quad \text{for } t \in I,\ i = 1, \ldots, N$$

Here:

- $L_i(t)$ is an affine transformation mapping I to subinterval $I_i = [t_{i-1}, t_i]$.
- s_i is a *scaling factor* chosen such that $|s_i| < (t_i - t_{i-1})^p$ to ensure contractivity.
- $r(t)$ is a base function in $C^p[I]$ used to guide the shape of the interpolation.

This construction allows f to match the smoothness of u while adding a controlled fractal variation determined by s_i and $r(t)$.

3.3 Fractal Perturbation of Kernel Regression

Let $\hat{g}(t)$ denote the traditional Priestley–Chao estimator defined by:

$$\hat{g}(t) = \sum_{i=1}^{N} (t_i - t_{i-1}) z_i k_d(t - t_i)$$

where $k_d(x) = \frac{1}{d} k\left(\frac{x}{d}\right)$ is a scaled kernel function (here, Gaussian).

To construct a fractal perturbation $f[\hat{g}]$ of $\hat{g}$, define a perturbation function $r(t) = \hat{g}(t) - \xi(t)$ where $\xi \in C^p[I]$ and satisfies:

$$\xi(0) = \xi(1) = 0, \quad \xi^{(n)}(0) = \xi^{(n)}(1) \text{ for } 1 \leq n \leq p$$

Under the appropriate constraints on s_i, this leads to a C^p smooth FIF $f[\hat{g}]$ that closely follows $\hat{g}(t)$ but introduces fractal characteristics. The deviation is bounded by:

$$\|f^{(n)}[\hat{g}] - \hat{g}^{(n)}\|_\infty \leq \frac{N^n S_0}{1 - N^n S_0} \|\xi^{(n)}\|_\infty, \quad 0 \leq n \leq p$$

where $S_0 = \max_i |s_i|$ and must satisfy $S_0 < \frac{1}{N^p}$.

3.4 Optimization Methodology

To determine optimal parameters $\{d, s_1, \ldots, s_N\}$ that minimize the empirical mean squared error (MSE) between the fractal regression $f[\hat{g}]$ and the observed data, the Sequential Least Squares Quadratic Programming (SLSQP) method is employed.

- The objective function minimized is:

$$E = \frac{1}{M} \sum_{j=1}^{M} (w_j - f[\hat{g}](x_j))^2$$

  where w_j are observed values and x_j are points in the domain.

- SLSQP handles constraints on parameters, ensuring that $|s_i| < \frac{1}{N^p}$ and $d > 0$.
- Python's `scipy.optimize.minimize` library is used for implementation.

This optimization process yields the best-fitting fractal regression model for the given sample data.

3.5 Summary

The methodology proposed in this study merges the strengths of nonparametric regression and fractal analysis. By integrating smooth fractal interpolation functions into traditional kernel estimators, the resulting regression functions exhibit both high approximation accuracy and the ability to model complex, self-similar data. This is particularly beneficial in applications involving limited data with intricate structures, such as the Crude Oil WTI Futures time series explored in the paper.

Chapter 4

Experimental Results and Analysis

4.1 Dataset Description

The dataset used in this study is the daily highest prices of Crude Oil WTI Futures, collected over the period from October 13, 2021, to July 29, 2022. A total of 211 data points are available, representing a real-world time-series dataset.

For experimental purposes, a sample subset of 11 data points was selected from this dataset. These points correspond to the dates: October 13, November 11, December 10, January 11, February 8, March 8, April 6, May 6, June 3, July 1, and July 29. The input domain is normalized to $[0, 1]$ with $t_i = i/10$ for $i = 0, 1, \ldots, 10$.

4.2 Kernel Estimation Setup

The base regression function $\hat{g}(t)$ is constructed using the Priestley–Chao estimator with a Gaussian kernel:

$$\hat{g}(t) = \frac{1}{\sqrt{2\pi}Nd} \sum_{i=0}^{N} z_i \exp\left(-\frac{(t - \frac{i}{N})^2}{2d^2}\right)$$

where $N = 10$, z_i are the selected sample values, and d is the bandwidth parameter.

4.3 Fractal Perturbation Construction

To construct the smooth fractal perturbation $f[\hat{g}]$:

- A perturbation function $\xi(t) = 3000(-t^3 + 3t^4 - 3t^5 + t^6)$ is chosen such that $\xi^{(n)}(0) = \xi^{(n)}(1) = 0$ for $n = 0, 1, 2$.
- The auxiliary function is set as $r(t) = \hat{g}(t) - \xi(t)$.
- Fractal scaling factors $s_1, \ldots, s_{10}$ and the kernel width d are optimized using the SLSQP method to minimize the mean squared error.

4.4 Results and Comparisons

The experiments are conducted for smoothness levels $p = 0, 1, 2$. The constraints on scaling factors are set based on $|s_i| < 1/N^p$. The following table summarizes the optimized values and the resulting errors:

Table 4.1: Mean Squared Errors for Different FIFs

Model	Optimal d	Mean Squared Error (MSE)
C_0-FIF $f[\hat{g}]$	0.0513	18.2085
C_1-FIF $f[\hat{g}]$	0.0506	23.3264
C_2-FIF $f[\hat{g}]$	0.0494	30.0976
Original PC Estimator $\hat{g}$	0.0493	31.5464

4.5 Graphical Analysis

- Figures compare the original data, PC estimator $\hat{g}(t)$, and the fractal-enhanced models $f[\hat{g}](t)$ for various p values.

- The C_0-FIF shows the best fitting accuracy, clearly outperforming the standard kernel estimator.

4.6 Discussion

The results demonstrate that fractal perturbations effectively enhance the performance of kernel regression models on limited, real-world sample data. The empirical MSE is lowest when $p = 0$, indicating a trade-off between smoothness and fitting capability. As the smoothness condition becomes stricter, the optimization space narrows, leading to higher MSE values.

These findings validate the proposed method and highlight the applicability of fractal functions in data fitting tasks where precision and adaptability to complex structures are critical.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

This work explored the integration of smooth fractal interpolation functions (FIFs) into traditional nonparametric kernel regression estimators. Building on foundational theories of fractal geometry and interpolation introduced by Barnsley and Demko [2, 1], and further developed into smooth FIFs by Chand and Kapoor [4] and Navascués et al. [8], the study proposed a fractal perturbation framework for the Priestley–Chao estimator.

By introducing carefully designed smooth perturbations to the kernel estimator, the proposed model captures both the local smoothness and global self-similarity present in complex real-world data. Recent work by Liu and Luor [6, 7] motivated the construction of C^p-smooth FIFs as perturbative enhancements, with optimal parameters determined through SLSQP-based optimization.

Experimental evaluation on a financial time-series dataset demonstrated that fractal-enhanced estimators outperformed the standard Priestley–Chao kernel estimator, especially at lower levels of smoothness (e.g., $p = 0$), which allowed greater fractal flexibility. These findings validate the potential of incorporating fractal geometry into regression frameworks, as also recognized in fractal modeling surveys [9] and spline-to-fractal interpolation theory [10].

5.2 Future Work

While the results are promising, several directions remain open for future research:

- **Generalization to Other Estimators:** Future work may extend fractal perturbation techniques to other kernel-based estimators such as the Nadaraya–Watson and Gasser–Müller models [7].
- **Multivariate Extension:** Extending the methodology to multivariate or high-dimensional data requires developing new classes of multi-dimensional FIFs, as seen in recent fractal geometry and interpolation advancements [11].
- **Adaptive and Data-Driven Parameter Selection:** Advanced optimization algorithms or machine learning-based search strategies could be employed to automatically tune scaling factors, smoothness degree, and kernel bandwidth.

- **Alternate Perturbation Functions:** Exploring other forms of $\xi(t)$, including those derived from Hermite or quintic interpolants [11], may provide better control over shape and convergence behavior.
- **Theoretical Guarantees:** Future work may focus on theoretical analysis of generalization error, convergence rates, and bias-variance trade-offs under fractal perturbations, continuing from frameworks like those in [5].
- **Real-World Applications and Toolkits:** Building a robust and reusable software toolkit (e.g., in Python or R) for fractal regression could facilitate its application in fields such as finance, climate modeling, biomedical signal analysis, and computer vision.

In conclusion, this research highlights the untapped potential of smooth fractal geometry in statistical learning, offering a new hybrid paradigm that unites local smoothing with global self-similarity for superior data approximation and modeling.

Bibliography

[1] Michael F Barnsley. Fractal functions and interpolation. *Constructive approximation*, 2:303–329, 1986.

[2] Michael F Barnsley and Stephen Demko. Iterated function systems and the global construction of fractals. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 399(1817):243–275, 1985.

[3] JW Cannon. The fractal geometry of nature. by benoit b. mandelbrot. *The American Mathematical Monthly*, 91(9):594–598, 1984.

[4] AKB Chand and GP Kapoor. Generalized cubic spline fractal interpolation functions. *SIAM Journal on Numerical Analysis*, 44(2):655–676, 2006.

[5] Thomas Hofmann, Bernhard Schölkopf, and Alexander J Smola. Kernel methods in machine learning. 2008.

[6] Chiao-Wen Liu and Dah-Chin Luor. Applications of fractal interpolants in kernel regression estimations. *Chaos, Solitons & Fractals*, 175:113913, 2023.

[7] Dah-Chin Luor and Chiao-Wen Liu. Fractal perturbation of the nadaraya–watson estimator. *Fractal and Fractional*, 6(11):680, 2022.

[8] MA Navascués and MV Sebastián. Smooth fractal interpolation. *Journal of Inequalities and Applications*, 2006:1–20, 2006.

[9] María Antonia Navascués, Arya Kumar Bedabrata Chand, Viswanathan Puthan Veedu, and María Victoria Sebastián. Fractal interpolation functions: a short survey. *Applied Mathematics*, 5(12):1834–1841, 2014.

[10] Richard Szeliski and Demetri Terzopoulos. From splines to fractals. *ACM Siggraph Computer Graphics*, 23(3):51–60, 1989.

[11] Jiaoyue Zheng, Gang Hu, Xiaomin Ji, and Xinqiang Qin. Quintic generalized hermite interpolation curves: construction and shape optimization using an improved gwo algorithm. *Computational and Applied Mathematics*, 41(3):115, 2022.