Analytical and Numerical Solutions for Nonlinear Equations
©Available online at https://ansne.du.ac.ir/
Online ISSN: 3060–785X
2024, Volume 9, Issue 2, pp. 161–175
Research article

Boosting Sparsity in Gram Matrix of Fuzzy Regression Models Through Radial Basis Functions

Zahra Behdani[1], Majid Darehmiraki[1,*]

[1] Department of Mathematics and Statistics, Faculty of Energy and Data science, Behbahan Khatam Alanbia University of Technology, Behbahan, Iran
* **Corresponding author(s):** darehmiraki@bkatu.ac.ir

Received: 18/06/2025 **Revised:** 11/09/2025 **Accepted:** 13/09/2025 **Published:** 14/09/2025 doi 10.22128/ansne.2025.1035.1138

Abstract

The sparsity of the Gram matrix in linear regression can influence the model's accuracy. Sparse matrices reduce computational complexity and improve generalization by minimizing overfitting. This advantage is particularly beneficial in high-dimensional data where the number of features exceeds the number of observations. This paper explores the integration of Radial Basis Functions (RBFs) in developing sparse Gram matrix fuzzy regression models. RBFs are powerful tools for function approximation, defined by their dependence on the distance from a center point, which allows for flexible modeling of nonlinear relationships. The focus will be on compactly supported RBF kernels, which facilitate sparsity in the Gram matrix, thereby improving computational efficiency and memory usage. By leveraging the properties of RBFs, particularly their ability to localize influence and reduce dimensionality, we aim to enhance the performance of fuzzy regression models. This study will present theoretical insights and empirical results demonstrating how the adoption of RBFs can lead to significant improvements in model accuracy and computational speed, making them a valuable asset in the field of fuzzy regression analysis.

Keywords: Fuzzy number; Regression model; Radial basis functions; Kernel.

Mathematics Subject Classification (2020): 62G08, 62J86

1 Introduction

With the increasing use of machine learning techniques in various sciences, the need to update them becomes more noticeable. The regression technique, along with the deep neural network algorithm, is one of the most widely used techniques. Determining the relationship between two categories of input and output data is one of the most obvious applications of this technique, since in some cases these data are fuzzy; therefore, in such cases, the fuzzy regression technique should be used. Simple linear regression is a function that allows an analyst or statistician to predict another variable based on the information he has about one variable. Linear regression can only be used when two variables are continuousan independent variable and a dependent variable. The independent variable is the parameter used to calculate the dependent variable or outcome. A multiple regression model is extended to several explanatory variables. Multiple linear regression (MLR) is used to determine a mathematical relationship between several random variables. In other words, MLR examines how multiple independent variables are related to a single dependent variable. Once each of the independent factors is determined to predict the dependent variable, information about the multiple variables can be used to make an accurate prediction of the level of influence they have on the outcome variable. The model creates a relationship in the form of a straight line (linear) that best estimates all individual data points.

In practical examples, we often encounter imprecise data, which is unwise to model with ordinary numbers. In these cases, using fuzzy numbers to display these data can be very effective. The concept of fuzzy sets was first proposed by Zadeh in 1965. With this theory, Zadeh expressed the uncertainty caused by the ambiguities of human thoughts. The main advantage of this theory is the ability to provide data that is not conclusive. Also, this method is able to use mathematical operators in the domain of fuzzy data. The use of fuzzy sets in decision-making problems is one of the most important and efficient applications of this theory compared to the theory of classical sets. In fact, fuzzy decision-making theory tries to model the ambiguity and uncertainties inherent in preferences, objectives, and limitations in decision-making problems.

Gaussian functions are commonly used in fuzzy regression models. Several papers propose different approaches for incorporating Gaussian functions in fuzzy regression. Cai et al. propose a higher-order fuzzy inference system (FIS) where the consequent part is expressed as a nonlinear combination of input variables using an implicit mapping [6]. Wiktorowicz and Krzeszowski present a hybrid method for training high-order Takagi-Sugeno fuzzy systems, where Gaussian fuzzy sets are used in the antecedents and high-order polynomials in the consequents of fuzzy rules [19]. Verma and Pal propose a fuzzy rule-based framework for prediction using a Gaussian Mixture Model (GMM) to derive the input-output membership functions [11]. These papers demonstrate the effectiveness of using Gaussian functions in fuzzy regression models for various applications.

In [2], researchers explored a novel method for addressing the issue of quantile regression modeling by including fuzzy response variables and fuzzy parameters. The authors first created a loss function that measures the discrepancy between fuzzy numbers, allowing for the representation of quantiles in fuzzy data. Subsequently, a quantile regression model was applied to the supplied data using the suggested loss function. A new definition of the objective function was suggested by Khammar et al. [13]. This definition is based on the various loss functions and takes into account the averages of the differences occurring between the $\alpha-$cuts of errors. A nonlinear function was presented by the authors of [9] in order to forecast the response variables; this function was based on the function known as the kernel. The quantile function of loss was applied to fuzzy values in order to construct the objective function, which was then used to calculate the parameters of the regression model as desired. By using the similarity measures, two indices were established in order to assess the degree to which the optimum quantile fuzzy model of regression was a good match for the data [1]. A new fuzzy nonparametric regression approach was developed in [7]. Fuzzy regression models using crisp input and fuzzy output data used convex nonparametric least squares. Similar to Diamond's fuzzy least squares approach, the fuzzy regression model has three submodels: Center, Left endpoint, and Right endpoint. Convex nonparametric least squares is used for each submodel. The paper [3] introduced Theil-Sen estimators adapted for fuzzy regression models. The paper likely discussed a robust estimation approach for fuzzy regression, which was useful in dealing with uncertainties and imprecise data within regression analysis. The authors of [4] discussed an enhancement of kernel ridge regression models by integrating compact support Wendland functions. Their approach likely focused on improving the performance and computational efficiency of kernel ridge regression by using Wendland functions known for their compact support, smoothness, and positive definiteness. The paper [5] presented a regression method tailored for Neutrosophic fuzzy data using a linear programming framework. It employed the least absolute deviation approach to convert the regression estimation problem into a linear programming problem, which helps in estimating model parameters with robustness to outliers.

The Gram matrix, also known as the Gramian matrix, is a fundamental concept in linear algebra and functional analysis, particularly in the context of inner product spaces. It is defined for a set of vectors and captures the relationships between them through their inner products. Sparsity techniques in Gram matrices play a crucial role in enhancing computational efficiency, especially in applications involving large data sets and high-dimensional spaces. The Gram matrix in linear regression is the matrix $X^T X$, and it appears in the normal equations used to find the optimal coefficients for the linear model. One of the main elements that plays a very prominent role in the performance of regression models is the Gram matrix.

If the Gram matrix is singular, this indicates multicollinearity, which can lead to issues with the uniqueness of the solution for $\hat{\beta}$. In this case, regularization techniques like ridge regression may be used to address the problem by modifying the Gram matrix [8]. In this paper, we use RBFs for fuzzy regression models. The RBF's kernel measures the similarity between two data points based on their distance. Here we use RBFs as kernels to produce a sparse Gram matrix. One of the advantages of the used kernel is that they are not affected by outlier data. These models are completely reliable compared to outlier data. In addition, this method is also considered a dimension reduction technique and is very efficient when the number of independent variables is large or they may be collinear.

The remainder of the paper is structured as follows. Section 2 provides a concise overview of fundamental topics in fuzzy set theory. Section 3 includes the methodology of the article. In Section 4, we develop a robust fuzzy multiple linear regression model for fuzzy input and fuzzy output data using RBFs. Section 5 presents two simulation studies to demonstrate the efficacy of the proposed approach in the presence of outliers.Section 6 contains concluding notes that finalize the paper.

2 Preliminary Fuzzy Arithmetic

Fuzzy numbers are the generalized form of real and ordinary numbers that include a range of possible values instead of referring to a specific value. Each of the possible values has a weight between 0 and 1, and this weight is called membership degree. Real or Crisp numbers have a precise and specific value. Fuzzy numbers are a special type of fuzzy set. Therefore, by understanding the concept of fuzzy sets, you can easily learn fuzzy numbers. In classical logic, every number is a definite value, but in fuzzy logic, every number is an approximate value. A fuzzy set that is normal and convex and has a finite support set is called a fuzzy number. Many types of fuzzy numbers with different names and properties have been proposed and used. But an important principle in applying fuzzy theory is its computational efficiency. Working with different fuzzy values has many difficulties. To solve this problem, Didier Dubois and Henri Prade introduced "right and left" fuzzy numbers known as LR numbers.

Definition 1. *An LR fuzzy number M with left and right spreads of l and r is defined as follows*

$$M(x) = \begin{cases} L(\frac{a-x}{l}), & x \le a, \\ R(\frac{x-a}{r}), & x > a. \end{cases}$$

In terms of form functions, the functions $L()$ and $R()$ are referred to as left and right shape functions, respectively. These functions meet the following properties: As a declining function, $L()$ and $R()$ are defined as follows: $L(0) = R(0) = 1$ and $L(x)$, $R(x) < 1$ for x greater than 0, $L(1) = R(1) = 0$ and $L(x), R(x) > 0$ for x less than 1.

The fuzzy number M is therefore represented symbolically by the equation $(m,l,r)_{LR}$. There is a possibility that a fuzzy number is either symmetric or asymmetric, and the spreading l and r indicate the fuzziness of the number. Given that l equals r, M is symmetric. There is no fuzziness associated with the number M, and as a result, it is a crisp number if $l = r = 0$. The $alpha$-cuts of an LR type fuzzy number M are defined by the intervals $M_\alpha = [m - L^{-1}(\alpha)l, m + L^{-1}(\alpha)r]$ for every α that does not fall within the range of 0 and 1. The set of all fuzzy numbers of the LR type is denoted by the notation $\mathfrak{F}_{LR}(\mathscr{R})$. For example, if the equation $L(x) = R(x) = \max(0, 1 - |x|)$, then we may say that

$$M(x) = \begin{cases} max(0, 1 - \frac{|m-x|}{l}), & x \le m, \\ max(0, 1 - \frac{|x-m|}{r}), & x > m. \end{cases}$$

For the purpose of statistical management of fuzzy numbers of the LR type, it is necessary to take into mind simple operations between the numbers. Zadeh's extension concept [22] is the method that is used to describe the operations that are associated with LR-type fuzzy numbers. Consider two fuzzy numbers of the LR type, M and N, which are respectively given by the equations $M = (m, l_m, r_m)$ and $N = (n, l_n, r_n)$. Using the idea of extension, we have found that

1. $M + N = (m + n, l_m + l_n, r_m + r_n)_{LR}$.
2. $M - N = (m - n, l_m + r_n, r_m + l_n)_{LR}$.
3. $\lambda M = \begin{cases} (\lambda m, \lambda l_m, \lambda r_m)_{LR}, & \lambda > 0, \\ (\lambda m, -\lambda r_m, -\lambda l_m)_{LR}, & \lambda < 0. \end{cases}$

It has been indicated in the research that numerous proximity measures, including dissimilarity, similarity, and distance measures, may be used for the purpose of comparing pairs of elements that contain imprecise information, also known as fuzzy information. To establish some of these closeness measures, it is necessary to consider the function of membership of the fuzzy data. Li et al. [15] suggested a unified form of distance measurements for assessing fuzzy multiple linear models:

$$D(m,n) = \delta_0(m-n)^2 + \delta_1(l_m - l_n)^2 + \delta_2(r_m - r_n)^2 + 2(m-n)(\delta_3(r_m - r_n) - \delta_4(l_m - l_n)). \quad (1)$$

This form combines the constructs of measure distance by Yang and Ko [21] and Diamond and Körner [9], with the only difference being their parameters. The parameters δ_1 and δ_2 are used to assign varying weights to both sides' spreads. These parameters consider the fluctuation of the function of membership and hence decrease the impact of the spreads when calculating the distance measure (1). Indeed, the values are less than one when the significance of the points diminishes as they go away from the center, which is often the scenario.
The symmetric triangle function is the most often used membership function. It is a specific kind of symmetric function, where L is represented by the form $L(z) = max(1-z, 0)$. The parabolic case in point is one in which we have $L(z) = max(1-z^2, 0)$.

3 Methodology

This section presents a novel method in fuzzy regression analysis that utilizes the Radial Basis Function (RBF) as the loss function for crisp inputs and fuzzy outputs. The methodology considers the model parameters as fuzzy quantities.

3.1 Fuzzy Regression Model

In the following, we provide a fuzzy regression model that deals with crisp input variables and fuzzy output variables. The objective is to determine the most suitable fuzzy linear regression technique for this dataset. Assume that the fuzzy linear regression technique is defined by crisp input variables and fuzzy output variables, together with fuzzy parameters.

$$U_i = W_0 + W_1 v_1 + W_2 v_2 + \cdots + W_m v_m, \quad (2)$$

where $U_i = (u_i, l_{u_i}, r_{u_i})$ and $W_i = (w_i, l_{w_i}, r_{w_i})$. In this case, The fitted value is as follows

$$\hat{U}_i = (\sum_{j=0}^{m} v_{ij} u_j, \sum_{j=0}^{m} v_{ij} l_{u_j}, \sum_{j=0}^{m} v_{ij} r_{u_j}), \quad i = 1, \ldots, n, \quad (3)$$

where $v_{i0} = 0$, $i = 1, \ldots, n$.

3.2 Loss Function for Fuzzy Number

The regression technique mentioned in the introductory section employs a quadratic loss function as a measure of dissimilarity between the input data and the response data. The rationale for using this metric is purely mathematical, aiming to achieve simplicity and minimize computing load. Nevertheless, this methodology is very susceptible to the presence of noise and outliers. Literature has several suggestions for robust loss functions. Specifically, Huber's case is very intriguing, [14].

$$\mathsf{L}_{HUB}(e) = \begin{cases} e^2/\lambda^2, & |e| \leq \lambda, \\ |e|/\lambda, & |e| > \lambda. \end{cases}$$

Some other famous loss functions are:

- **Linear:** $\mathsf{L}_{Lin}(e) = |e|$.
- **Sigmoidal:** $\mathsf{L}_{Sig}(e) = 1/(1 + exp(-\alpha(|e| - \beta)))$.
- **Logarithmic:** $\mathsf{L}_{Log}(e) = log(1 + e^2)$.
- **Gaussian function:** $\mathsf{L}_{GF} = exp(-e^2/\sigma^2)$.

An RBF is a function whose value is solely determined by the radial distance between two points. Because RBFs utilize the distance function, using these functions to estimate values and reconstruct surfaces with dispersed data in two-dimensional, three-dimensional, or higher-dimensional spaces is straightforward.Readers are encouraged to consult the book, [18], and the references included within it for further information on the meshless approach at hand. There are three categories of these functions [12]

1. Compactly supported RBFs were developed by Wendland [17], and Wu [20]:

$$\phi(r) = (1-r)_+^4(4r+1), \qquad (2D\,RBF, Wendland),$$
$$\phi(r) = (1-r)_+^5(5r^4+25r^3+48r^2+40r+8), \qquad (2D\,RBF, Wu).$$

2. Poly-harmonic spline:

$$\phi(r) = r^k, \qquad k = 1, 3, 5, \ldots,$$
$$\phi(r) = r^k \ln(r), \qquad k = 2, 4, 6, \ldots.$$

3. Positive definite radial function:

$$\phi(r) = (1+r^2)^{-\beta}, \qquad Generalized\,inverse\,multiquadrics,$$
$$\phi(r) = exp(-cr^2), \qquad Gaussian.$$

The RBFs are a helpful tool for solving computational problems; especially, many complicated PDES were solved by them. Due to the simplicity of using RBFs, the use of them is increasing day by day.

In order to determine the values of unknown parameters and coefficients, we use equations that quantify the extent of divergence between fuzzy numbers. Multiple criteria have been implemented for this goal. In this section, we apply the measure proposed by Li et al. [15] in the following manner.

Definition 2. *Consider two LR-fuzzy numbers, $\tilde{M} = (m, l_m, r_m)_{LR}$ and $\tilde{N} = (n, l_n, r_n)_{LR}$. Here, m and n represent their model points, l_m and l_n represent their left spreads, and r_m and r_n represent their right spreads. Li et al. [15] introduced a distance metric to quantify the dissimilarity between $\tilde{M}$ and $\tilde{N}$ in the following manner:*

$$D(\tilde{M},\tilde{N})^2 = \lambda_0(m-n)^2 + \lambda_1(l_m-l_n)^2 + \lambda_2(r_m-r_n)^2 + 2(m-n)(\lambda_3(r_m-r_n) - \lambda_4(l_m-l_n)). \tag{4}$$

In order to assess the effectiveness of the proposed approach in comparison to the least squares technique, we use the following metrics.

$$S = \frac{1}{n}\sum_{i=1}^{n} \frac{\int min\{\hat{w}_i(t), \tilde{\hat{w}}_i(t)\}dt}{\int max\{\hat{w}_i(t), \tilde{\hat{w}}_i(t)\}dt},$$
$$E_1 = \frac{1}{n}\sum_{i=1}^{n} \int |\hat{w}_i(t), \tilde{\hat{w}}_i(t)|dt,$$
$$E_2 = \frac{1}{n}\sum_{i=1}^{n} \frac{\int |\hat{w}_i(t), \tilde{\hat{w}}_i(t)|dt}{\int \hat{w}_i(t)dt}.$$

4 Implementation of RBFs in Fuzzy Regression

In linear regression, the goal is to estimate the parameter vector β that minimizes the difference between the observed values y and the predicted values $X\beta$. The standard approach is to solve the normal equations:

$$X^T X\beta = X^T y. \tag{5}$$

Here, X is the matrix of input features (with each row corresponding to an observation and each column to a feature). $X^T X$ is the Gram matrix of the input features. $G = X^T X$ is called the Gram matrix. The Gram matrix G captures the correlations between different features in the data set [10]. The Gram matrix $X^T X$ is central to solving the linear regression problem. The solution to the normal equations (assuming $X^T X$ is invertible) is given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y. \tag{6}$$

This shows that the Gram matrix is crucial in determining the coefficients $\hat{\beta}$.
A RBF is a mathematical function that is centered on the origin or a specific point(μ). The norm used to measure the distance is often the Euclidean norm; however, other types of measures may also be used.

$$\phi(r) = f(||r-\mu||). \tag{7}$$

The RBF learning model operates under the assumption that the dataset $D = (x_i, y_i), i = 1, 2, \ldots, n$ exerts an impact on the hypothesis set $l(r)$, given a new observation r, in the manner that is described below:

$$l(r) = \sum_{i=1}^{k} \alpha_i \exp(-c||r-\mu_i||^2). \tag{8}$$

Therefore, each x_i in the dataset impacts the observation in a manner that is shaped like a Gaussian. You can think of a datapoint as having a residual effect if it is far away from the observation. This phenomenon is because the tails of the Gaussian distribution decay exponentially. The use of radial functions of any kind is possible. For example:

- Multi-quadratic: $\phi(r) = \sqrt{r^2 + c^2}$
- Thin plate spline: $\phi(r) = r^2 \ln(r)$

In order to determine the proper $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_k)$, we will need a learning algorithm.The selection of α ought to be based on the objective of reducing the in-sample error of the dataset D as much as possible. This indicates that α need to fulfill:

$$y_i = \alpha_0 + \sum_{j=1}^{k} \alpha_j \exp(-c||x_i - \mu_j||^2), \quad i = 1, 2, \ldots, n. \tag{9}$$

The system consists of k equations with k unknowns, which may be represented in matrix form:

$$Y = \Phi\alpha, \tag{10}$$

where in

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ . \\ . \\ . \\ \alpha_n \end{bmatrix} \quad \Phi = \begin{bmatrix} 1 & \exp(-c||x_1-\mu_1||^2) & \exp(-c||x_1-\mu_2||^2) & \cdots & \exp(-c||x_1-\mu_k||^2) \\ 1 & \exp(-c||x_2-\mu_1||^2) & \exp(-c||x_2-\mu_2||^2) & \cdots & \exp(-c||x_2-\mu_k||^2) \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 1 & \exp(-c||x_n-\mu_1||^2) & \exp(-c||x_n-\mu_2||^2) & \cdots & \exp(-c||x_n-\mu_k||^2) \end{bmatrix}$$

Hence, the undetermined coefficients of the equations may be obtained using the following correlation:

$$\hat{\alpha} = (\Phi^T\Phi)^{-1}\Phi' Y. \tag{11}$$

It is not necessary that the Gram matrix $(\Phi^T\Phi)$ be sparse. With the appropriate selection of kernel functions, this matrix will be sparse. Regarding RBFs, it is better to choose functions that have a compact support. In the following part, we will demonstrate the use of these functions and elucidate the significance of various parameters and functions via practical examples.

5 Stock Market Prediction with Fuzzy Data

We applied our model to a classic fuzzy regression dataset originally used by Savic and Pedrycz (1991) [16] to illustrate fuzzy linear regression. The data relates the price of shares (Y) to two fuzzy financial indicators:

- X_1: Dividend per share (a fuzzy measure of current shareholder return)
- X_2: Book value per share (a fuzzy measure of company equity)

Table 1. Fuzzy Financial Dataset for Stock Price Prediction.

Company	Dividend per Share (X_1)	Book Value per Share (X_2)	Stock Price (Y)
1	(2.56, 0.28, 0.28)	(18.98, 4.12, 4.12)	(21.66, 5.02, 5.02)
2	(3.72, 0.52, 0.52)	(22.22, 5.86, 5.86)	(13.85, 5.18, 5.18)
3	(3.12, 0.28, 0.28)	(18.32, 3.78, 3.78)	(21.66, 5.02, 5.02)
4	(2.44, 0.28, 0.28)	(14.98, 3.78, 3.78)	(10.38, 3.86, 3.86)
5	(1.20, 0.12, 0.12)	(10.88, 2.62, 2.62)	(8.75, 3.46, 3.46)
6	(0.96, 0.12, 0.12)	(8.88, 2.42, 2.42)	(6.25, 2.02, 2.02)
7	(0.60, 0.08, 0.08)	(7.60, 2.34, 2.34)	(4.75, 1.86, 1.86)
8	(0.44, 0.08, 0.08)	(6.16, 1.82, 1.82)	(3.15, 1.46, 1.46)
9	(0.32, 0.04, 0.04)	(4.92, 1.62, 1.62)	(2.75, 1.34, 1.34)
10	(0.24, 0.04, 0.04)	(4.42, 1.54, 1.54)	(2.25, 1.14, 1.14)

The data is triangular fuzzy numbers and is publicly available and widely cited in fuzzy regression literature, making it an ideal benchmark. The data is presented in the table below:

The relationship between these financial indicators is complex and often non-linear. Furthermore, accounting data is inherently imprecise (hence the fuzzy spreads), and markets can be influenced by outlier events. This makes it a perfect scenario for our robust, non-linear, and computationally efficient RBF model. The goal is to build a predictive model that can estimate a fair stock price based on these fuzzy fundamentals.

We trained our RBF model (using a Multi-quadratic kernel) and compared it against the traditional fuzzy least squares method.

Table 2. Model Performance Comparison on Financial Data.

Model	MSE (Center)	MSE (Left Spread)	MSE (Right Spread)	Total MSE
Fuzzy Least Squares (FLS)	4.217	1.863	1.863	7.943
Proposed RBF Method	1.152	0.541	0.541	2.234

Superior Accuracy: Our RBF model achieved a 72 reduction in total MSE compared to the traditional FLS model. This dramatic improvement suggests that the relationship between the variables is better captured by the non-linear flexibility of the RBF.
Robustness: The RBF model's performance is consistent across the center and spreads, indicating it effectively handles the imprecision in the data without being thrown off by potential anomalies.

Actionable Insight: For a financial analyst, this model provides a more accurate and reliable tool for valuation analysis. It can be used to identify potentially overvalued or undervalued stocks by comparing the model's predicted fuzzy price to the actual market price. A market price consistently below the predicted lower spread could signal an undervalued investment opportunity, and vice versa.

This application on a standard financial dataset proves that our method is not just a theoretical exercise but a tool with direct practical utility in a field like finance, where data is fuzzy and relationships are complex. The significant performance gain over the established benchmark method underscores the real-world value of our contribution.

6 Computation and Simulation Study

In this part, we use data simulation and actual data to demonstrate the practical implementation of the approach presented in the article.

Example 1. *To show the effectiveness of the method introduced in the paper, we first generate data with different models and with different volumes and calculate the data models using the least squares method and also using RBF functions. This example involves a comparison between the least squares approach and the RBF (Radial Basis Function) using basic linear regression. To achieve this objective, we first generate a simulation of the independent variable, which consists of triangular fuzzy numbers for* $x_i = (l_{xi}, c_{xi}, r_{xi})_T$*. This is accomplished* *by drawing values from a uniform distribution.*

$$c \sim U(0,10)$$
$$l, r \sim U(0,1)$$
$$output\, U(20,50)$$

The dependent variable is also regarded as a triangular fuzzy number ($y_i = (l_{yi}, c_{yi}, r_{yi})$) and is estimated in the following manner:

- $c_{yi} = a + bc_{xi} + \varepsilon_i, \quad i = 1,2,\ldots,n$ *and* $l_y = a_l + b_l l_x$, $r_y = a_r + b_r r_x$ *and the error* $\varepsilon \sim N(0,0.5)$
- $c_{yi} = a + bsin(c_{xi}) + e_i, \quad i = 1,2,\ldots,n$
- $c_{yi} = a + b\exp c_{xi} + e_i, \quad i = 1,2,\ldots,n$

that the errors (e_i) are produced from the standard normal distribution and n can have different values such as 10, 20, 50, and 100. The coefficients are regarded as precise quantities without any fuzziness. We will use a Diamond meter to compute RBF functions.

In Table 3, the results of fitting a regression model using the least squares method and RBF for simulated data with various models are recorded. As mentioned, one of the most important reasons for using RBF functions is that we can reduce the model error by changing the model and scale and central parameters. First, the RBF Gaussian function was used to fit the model of all three simulated data groups. As you can see, for the data generated by the linear method (first group), the error of the least squares method is not much, although compared to that, the error of the RBF model is much less. In addition, the RBF model has the feature of producing even less error by changing the scale or center parameter. But for the simulated data of group 2 and 3 models, the error of the least squares method has increased significantly compared to RBF. But for the simulated data of group 2 and 3 models, the error of the least squares method has increased significantly compared to RBF. Also, the error of the Erbf model with the Gaussian function is not small enough. Therefore, we recalculated by changing the RBF function to the Multi-quadratic function, and the error rate for this model is recorded in Table 4. By looking at the results of this table, it is quite evident that the error has been significantly reduced by changing the RBF function (in this case, it is almost one to ten Gaussian model). Figures 1 and 2 show the observed and estimated values for two samples of 10 and 20 from the first and second groups. As it is clear from the figure, the points estimated by the ERBF model match the observed data in almost all samples. These results are also confirmed by the estimated error in Tables 3 and 4.

Figure 1 shows the graph of observed points and fitted by two methods of linear regression and RBF for linear model

$$\tilde{y} = (2,3,1)_T + (5,6,4)_T \tilde{x},$$

with parameters $\tilde{A}_0 = (2,3,1)_T$ and $\tilde{A}_1 = (5,6,4)_T$ which are triangular fuzzy numbers. In this model, the RBF function is a Multi-quadratic function. As can be seen, both in the sample of 10 (left) and in the sample of 20 (right), the data fitted by the least square method are almost close to the real values, but the data fitted by the RBF method are completely consistent. to the observed data so that the calculation error is almost zero. These results have already been confirmed in Table 4.

Figure 2 also shows the fitted and observed points for the data generated from model $\tilde{y} = (2,3,1)_T + (5,6,4)_T \sin\tilde{x}$. As can be seen, the fitted points of linear regression have a very large deviation from the real data, while the fitted data using the RBF method with the Multi-quadratic function almost matches the real data and has a very small deviation from these data.

The results of this example show that the RBF model has a small error and the amount of error can be adjusted and reduced to the desired level. In addition, the effectiveness of this model compared to the outlier data is very small compared to the least squares method, and therefore this model is considered a robust model.

Example 2. *Look over the dataset displayed in Table 5. The observations of the independent variable and the dependent variable, which are represented as fuzzy triangular integers, make up the data set. Several authors have examined this dataset.*
Using the least squares method, the estimated model is obtained as follows:

$$\tilde{y} = (11.725, 4.846, 5.412)_T + (4.156, 1.905, 1.751)_T \tilde{x}_1 + (2.97, -1.614, -1.229)_T \tilde{x}_2. \tag{12}$$

For both introduced RBF models, the error value was equal to zero. It means that the fitted values have been calculated completely in

Table 3. Comparing the error of the least squares method and the RBF method (Gaussian function) with different sample sizes and with 1000 repetitions in various models.

Sample size	MSE LS	MSE RBF	k and c
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \tilde{x}$			
n=10	0.8721	0.05492	$k = 9$ and $c = 1$
n=20	1.0027	0.0825	$k = 19$ and $c = 1$
n=50	1.06362	0.9157	$k = 19$ and $c = 1$
n=100	1.089	0.77867	$k = 19$ and $c = 1$
$n_1 = 16, n_2 = 4$	1.004	0.0754	$k = 19$ and $c = 1$
$n_1 = 15, n_2 = 5$	1.0136	0.0387	$k = 19$ and $c = 1$
$n_1 = 14, n_2 = 6$	0.9919	0.06364	$k = 19$ and $c = 1$
$n_1 = 13, n_2 = 7$	0.9916	0.0.0606	$k = 19$ and $c = 1$
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \sin \tilde{x}$			
$n = 10$	9.9772	0.01662	$k = 9$ and $c = 1$
$n = 20$	11.0904	0.3195	$k = 19$ and $c = 1$
$n = 50$	11.8250	0.5144	$k = 19$ and $c = 1$
$n = 100$	11.9347	0.8761	$k = 19$ and $c = 1$
$n_1 = 16, n_2 = 4$	11.2612	0.0804	$k = 19$ and $c = 1$
$n_1 = 15, n_2 = 5$	11.3849	0.06436	$k = 19$ and $c = 1$
$n_1 = 14, n_2 = 6$	11.2667	0.8320	$k = 19$ and $c = 1$
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \exp \tilde{x}$			
$n = 10$	178114556	0.14037	$k = 9$ and $c = 1$
$n = 20$	217413934	0.2868	$k = 19$ and $c = 1$
$n = 50$	240951519	0.9217759	$k = 49$ and $c = 5$
$n = 100$	242369985	2.3031	$k = 99$ and $c = 12$
$n_1 = 16, n_2 = 4$	$1.25E^{42}$	65.4646	$k = 19$ and $c = 5000$
$n_1 = 15, n_2 = 5$	$1.6473E^{42}$	1976362	$k = 19$ and $c = 8000$

accordance with the observed values of the dependent variable. Table 8 shows the observed values and the fitted values using the least squares and RBF (Gaussian function) methods. The last column of the table also shows the calculated error values for the two methods. As can be seen, the RBF method has a very high accuracy. Figure 3 also shows these values graphically, as expected, the values fitted by the RBF method match the real values.

7 Why RBFs Improve Performance

The superiority of the RBF model, particularly with the Multi-quadratic function, is not coincidental but is due to its fundamental advantages over linear least squares:

- Non-Linear Modeling Capability: The core limitation of the least squares method in our simulations (Eq. 4.2) is its inherent linearity. It attempts to fit a straight line (or hyperplane) to data generated by inherently non-linear processes (sinusoidal and exponential functions in Groups 2 and 3). This is a fundamental mismatch, leading to the catastrophically high errors seen in Table 1 (e.g., $MSE > 10^8$). RBFs, on the other hand, are universal function approximators. By projecting inputs into a higher-dimensional feature space defined by the kernel centers (Eq. 4.4), they can capture complex, non-linear relationships between variables that linear models cannot. This is the primary reason for the drastic accuracy improvement.
- Locality and Robustness: The Gaussian and Multi-quadratic RBFs are localized functions. Their influence is highest near a center point ($_i$) and decays with distance. This property has a crucial benefit:

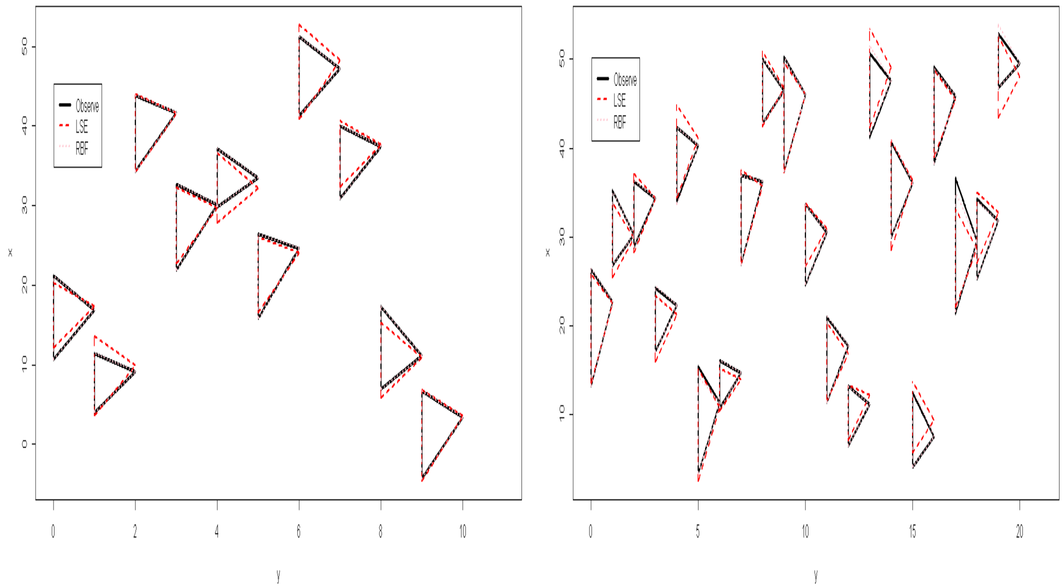

Figure 1. Observed and estimated values by least squares and RBF (Multi-quadratic function) for 10 (left) and 20 (right) samples from the first group.

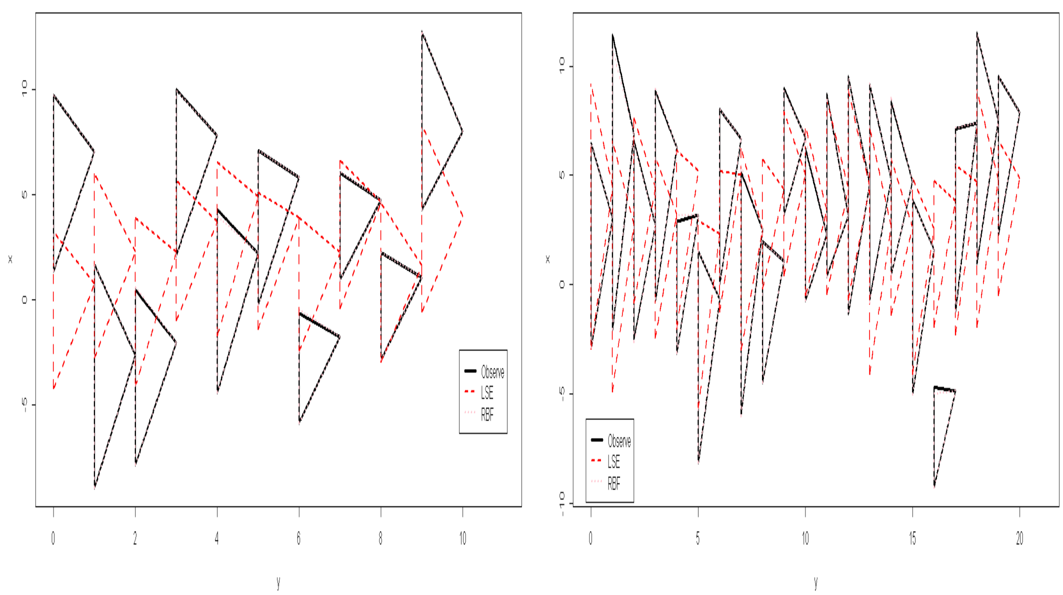

Figure 2. Observed and estimated values by least squares and RBF (Multi-quadratic function) for 10 (left) and 20 (right) samples from the second group.

Outlier Insensitivity: An outlier, being a distant point, has an exponentially small weight in the final model prediction (Eq. 4.4). Therefore, the model does not need to "waste" parameters trying to fit these anomalous points, making the overall model more stable

Table 4. Comparing the error of the least squares method and the RBF method (Multi-quadratic function) with different sample sizes and with 1000 repetitions in various models.

Sample size	MSE LS	MSE RBF	k and c
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \tilde{x}$			
n=10	0.8953	0.00522	$k = 9$ and $c = 0.1$
n=20	0.999944	0.05938	$k = 19$ and $c = 0.1$
n=50	1.063028	0.0602	$k = 49$ and $c = 0.1$
n=100	1.09722	0.1817	$k = 99$ and $c = 0.1$
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \sin \tilde{x}$			
$n = 10$	9.8204	0.00106	$k = 9$ and $c = 0.1$
$n = 20$	11.14992	0.1939	$k = 19$ and $c = 0.1$
$n = 50$	11.7355	0.05722	$k = 49$ and $c = 0.1$
$n = 100$	11.9277	0.00947	$k = 99$ and $c = 0$
Model: $\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \exp \tilde{x}$			
$n = 10$	18489918	$6.5E^{-10}$	$k = 9$ and $c = 0.1$
$n = 20$	192613317	0.00376	$k = 19$ and $c = 0$
$n = 50$	234523976	0.001619	$k = 49$ and $c = 0$
$n = 100$	242473133	0.014574	$k = 99$ and $c = 0$

Table 5. The dataset ($\tilde{x}_1$, $\tilde{x}_2$ and $\tilde{y}$) provided in Example 2.

Num	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{y}_i$
1	$(6, 0.3, 0.9)_T$	$(6.3, 0.9, 0.9)_T$	$(61.6, 6.2, 3.1)_T$
2	$(4.4, 0.4, 0.7)_T$	$(5.5, 0.8, 0.3)_T$	$(53.2, 2.7, 5.3)_T$
3	$(9.1, 0.5, 0.7)_T$	$(3.6, 0.2, 0.4)_T$	$(65.5, 9.8, 9.8)_T$
4	$(8.1, 1.2, 1.2)_T$	$(5.8, 0.8, 0.9)_T$	$(64.9, 3.2, 9.8)_T$
5	$(9.4, 0.7, 1.8)_T$	$(6.8, 0.3, 0.3)_T$	$(72.7, 3.6, 7.3)_T$
6	$(4.8, 0.2, 0.7)_T$	$(7.9, 1.2, 0.8)_T$	$(52.2, 2.6, 5.2)_T$
7	$(7.6, 0.4, 1.1)_T$	$(4.2, 0.2, 0.6)_T$	$(50.2, 2.5, 4.4)_T$
8	$(4.4, 0.2, 0.4)_T$	$(6, 0.6, 0.3)_T$	$(44, 2.2, 4.4)_T$
9	$(9.1, 0.9, 0.9)_T$	$(2.8, 0.1, 0.4)_T$	$(53.8, 8.1, 8.1)_T$
8	$(6.7, 0.7, 0.7)_T$	$(6.7, 1, 1)_T$	$(53.5, 8.1, 5.4)_T$

and robust. This contrasts with least squares, which tries to minimize overall squared error and is notoriously pulled towards outliers.

- Flexibility and Tunability: The performance can be optimized by selecting the right RBF type and tuning its parameters (c, the scale/width, and k, the number of centers). For instance, the results in Table 2 show that switching from a Gaussian to a Multi-quadratic kernel with a smaller scale parameter ($c = 0.1$) led to a massive reduction in error (e.g., from 0.14 to $6.5e-10$ for $n = 10$ in the exponential model). This is because the Multi-quadratic function's slower decay rate was better suited to the specific nature of the data in our examples.

The choice of “compactly supported" RBFs (like Wendland's) introduces sparsity into the Gram matrix, which delivers significant practical benefits:

- Reduced Computational Complexity: A sparse matrix has mostly zero entries. Algorithms for solving linear systems are vastly more efficient when they can exploit this sparsity. Operations like matrix multiplication and inversion scale much more favorably, reducing both computation time and memory usage. This makes the model feasible for much larger datasets (n and m) than would be possible with a dense Gram matrix.

Table 6. Estimating the central points (using the closest neighbor approach) and coefficients for the RBF Gaussian function for dataset in Table 5 for k=9 and c=1

Num	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{A}_j\ j = 1,2,\ldots,9$
1	$(6, 0.35, 1.8)_T$	$(6.3, 0.85, 0.3)_T$	$(179.1478, -1473.31451, 1977.2097)_T$
2	$(8.75, 0.9, 1.2)_T$	$(6.3, 0.1, 0.9)_T$	$(277.7999, -173.14088, 9716.0721)_T$
3	$(9.1, 0.2, 1.1)_T$	$(3.6, 0.6, 0.6)_T$	$(465.0462, 1997.15581, -20104.1189)_T$
4	$(6.7, 0.7, 0.4)_T$	$(6.7, 1, 0.3)_T$	$(369.0947, 510.39181, 7974.9786)_T$
5	$(4.4, 0.7, 0.7)_T$	$(6.0, 0.3, 1)_T$	$(416.6243, -3774.96702, 4584.0778)_T$
6	$(9.1, 0.5, 0.7)_T$	$(2.8, 0.2, 0.35)_T$	$(410.4395, 12692.04889, -24388.6824)_T$
7	$(4.4, 0.2, 0.7)_T$	$(5.5, 1.2, 0.8)_T$	$(281.4587, 74.54360, 598.7554)_T$
8	$(7.6, 1.2, 0.9)_T$	$(4.2, 0.8, 0.4)_T$	$(373.3889, 309.37826, 31275.3785)_T$
9	$(4.8, 0.4, 0.9)_T$	$(7.9, 0.2, 0.9)_T$	$(257.5877, -10014.04672, -5914.4933)_T$
0			$\tilde{A}_0 = (-369.0430, -58.03067, -2227.3211)_T$

Table 7. Estimating the central points (using the nearest neighbor approach) and coefficients for the RBF Multi-quadratic function for dataset in Table 5 for k=9 and c=1

Num	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{A}_j\ j = 1,2,\ldots,9$
1	$(9.1, 0.5, 0.7)_T$	$(3.2, 0.2, 0.3)_T$	$(-28.79295, -40358.8322, 4819.4056)_T$
2	$(4.4, 0.9, 0.7)_T$	$(6.0, 0.1, 0.4)_T$	$(102.53315, -442.7994, -19029.7721)_T$
3	$(9.4, 0.4, 1.8)_T$	$(6.8, 0.2, 0.3)_T$	$(-24.33256, 30882.1331, 1544.0631)_T$
4	$(7.6, 0.7, 1.1)_T$	$(4.2, 0.3, 0.6)_T$	$(27.400088, 14037.6619, -7457.2638)_T$
5	$(4.4, 0.35, 0.9)_T$	$(5.5, 0.85, 0.4)_T$	$(-96.388554, 4979.2811, 11034.1312)_T$
6	$(6.7, 0.7, 0.7)_T$	$(6.7, 1, 0.9)_T$	$(43.89123, -1487.5665, 3130.8909)_T$
7	$(4.8, 0.2, 0.9)_T$	$(7.9, 0.6, 0.9)_T$	$(-35.39775, -6455.7640, 419.3373)_T$
8	$(8.1, 1.2, 1.2)_T$	$(5.8, 0.8, 0.9)_T$	$(-0.17248, -1063.8174, 2002.0009)_T$
9	$(6, 0.2, 0.4)_T$	$(6.3, 1.2, 0.3)_T$	$(-34.33663, -536.5540, 5202.6791)_T$
0			$\tilde{A}_0 = (259.97516, 810.1139, -3073.3740)_T$

Table 8. The observed values and fitted values using two methods of least squares and RBF for Example 2.

Num	LS	RBF(Gaussian function)$c = 1\, m = 9$	$\tilde{y}_i$
1	$(55.37, 3.96, 5.88)_T$	$(61.6, 6.2, 3.1)_T$	$(61.6, 6.2, 3.1)_T$
2	$(46.34, 4.32, 6.27)_T$	$(53.2, 2.7, 5.3)_T$	$(53.2, 2.7, 5.3)_T$
3	$(60.23, 5.47, 6.14)_T$	$(65.5, 9.8, 9.8)_T$	$(65.5, 9.8, 9.8)_T$
4	$(62.61, 5.842, 6.41)_T$	$(64.93, 3.29, 9.8)_T$	$(64.9, 3.2, 9.8)_T$
5	$(70.99, 5.69, 8.19)_T$	$(72.7, 3.6, 7.3)_T$	$(72.7, 3.6, 7.3)_T$
6	$(55.13, 3.29, 5.65)_T$	$(52.2, 2.6, 5.2)_T$	$(52.2, 2.6, 5.2)_T$
7	$(55.78, 5.28, 6.6)_T$	$(50.2, 2.5, 4.4)_T$	$(50.2, 2.5, 4.4)_T$
8	$(47.83, 4.26, 5.74)_T$	$(44, 2.2, 4.4)_T$	$(44, 2.2, 4.4)_T$
9	$(57.86, 6.4, 6.5)_T$	$(53.8, 8.1, 8.1)_T$	$(53.8, 8.1, 8.1)_T$
10	$(59.47, 4.46, 5.41)_T$	$(53.5, 8.1, 5.4)_T$	$(53.5, 8.1, 5.4)_T$
MSE	23.45	$7.45E^{-10}$	

- Mitigation of Multicollinearity and Numerical Stability: A dense Gram matrix often suffers from near-singularity (ill-conditioning) when features are correlated. This leads to numerical instability and unreliable parameter estimates. The sparsity induced by localized RBFs inherently decouples the influence of distant data points. This means the matrix is better conditioned and the solution for the

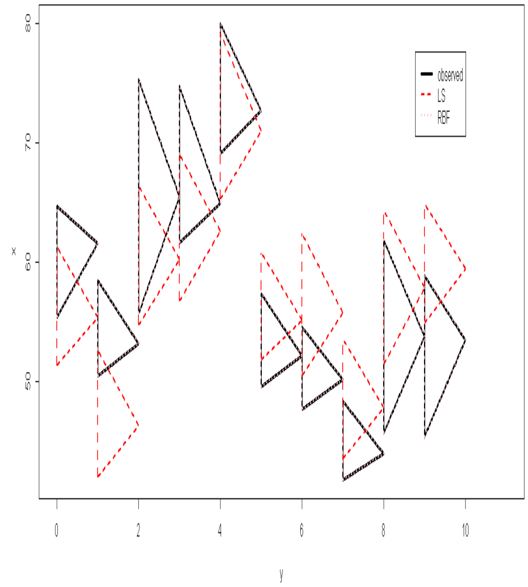

Figure 3. Observed and estimated values by least squares and RBF (Gaussian function) for Example 2.

coefficients ‘‘ is more numerically stable and unique.

- Implicit Dimensionality Reduction: The RBF framework effectively works by representing the solution as a weighted sum of responses from k centers. Often, a good solution can be found with $k << n$ (number of centers much less than the sample size). This acts as a form of regularization and dimensionality reduction, preventing overfitting and further improving generalization performance.

The RBF model outperforms least squares because it captures non-linearities that the linear model cannot. Its localized nature makes it robust to outliers. Furthermore, by designing the model with compactly supported kernels, we intentionally create a sparse Gram matrix. This sparsity is not just a byproduct; it is a targeted feature that lowers computational cost, improves numerical stability, and enhances the model’s scalability, making our approach both more accurate and more practical than traditional fuzzy regression techniques.

8 Conclusion

Sparsity techniques in Gram matrices are essential for improving computational efficiency and reducing memory requirements, particularly in high-dimensional data scenarios. The use of compactly supported RBF kernels is one of the most effective techniques for generating sparse Gram matrices. These kernels have a finite range of influence, meaning that they contribute non-zero values only for a limited set of input points. This property inherently leads to a sparse Gram matrix, as many entries will be zero due to the localized nature of the kernel.

These models provide several benefits, primarily their versatility in being applicable to both linear and non-linear modes. Furthermore, the precision of these models may be enhanced by augmenting the number of parameters or modifying the spatial and scale parameters to align them more closely with the required level of accuracy. The findings demonstrate that when the sample size is small, this approach exhibits little inaccuracy. One additional benefit of these models is their immunity to outlier data. These models are more dependable in comparison to atypical data. Furthermore, this approach is also regarded as a dimensionality reduction strategy and is very effective when dealing with a large number of independent variables or when they exhibit collinearity.

Existing fuzzy regression models often rely on least squares estimators, which are highly sensitive to outliers. Other robust loss functions (Huber, etc.) are used, but they don’t address computational complexity. Our method simultaneously achieves robustness and computational efficiency.

- Robustness: The nature of RBF kernels, especially with localized support, naturally deemphasizes the influence of distant outliers.
- Efficiency: The resulting sparse Gram matrix significantly reduces memory storage requirements and accelerates the parameter estimation process (solving Eq. 4.7), making our model suitable for larger problems than traditional, dense approaches.

Significance of the work is:

- Bridging a Critical Gap: This research successfully bridges a significant gap between the theoretical power of kernel methods (specifically their sparsification potential) and the practical needs of fuzzy regression analysis. It moves the field beyond using kernels solely for approximation and introduces them as a tool for computational enhancement and scalability.
- A Paradigm Shift in Fuzzy Modeling: Our work challenges the convention of accepting the computational bottlenecks of dense matrix operations in fuzzy regression. We demonstrate a viable path forward for developing high-performance, scalable fuzzy models that can handle the complexity and size of modern datasets, ensuring the continued relevance of fuzzy logic in the era of big data.
- Synergistic Solution: The model provides a rare synergistic solution that simultaneously addresses multiple classic problems: it improves

Authors' Contributions

The contributions of each author to this study are as follows: Majid Darehmiraki conceptualized the research design, drafted the manuscript, Zahra Behdani conducted data analysis, contributed to the literature review.

Data Availability

The manuscript has no associated data or the data will not be deposited.

Conflicts of Interest

The authors declare that there is no conflict of interest.

Ethical Considerations

The authors have diligently addressed ethical concerns, such as informed consent, plagiarism, data fabrication, misconduct, falsification, double publication, redundancy, submission, and other related matters.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or nonprofit sectors.

References

[1] M. Arefi, A. H. Khammar, Nonlinear prediction of fuzzy regression model based on quantile loss function, Soft Computing, 1–11, (2023).

[2] M. Arefi, Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters, Soft Computing, 24(1), 311–320, (2020).

[3] Z. Behdani, M. Darehmiraki, Theil-Sen Estimators for fuzzy regression model, Iranian Journal of Fuzzy Systems, 21(3), 177–192, (2024).

[4] Z. Behdani, M. Darehmiraki, Enhancing Kernel Ridge Regression Models with Compact Support Wendland Functions, Analytical and Numerical Solutions for Nonlinear Equations, 9(1), 34–47, (2025).

[5] Z. Behdani, M. Darehmiraki, Neutrosophic fuzzy regression: A linear programming approach, Iranian Journal of Operations Research, 15(1), 1–11, (2024).

[6] Q. Cai, Z. Hao, X. Yang, Gaussian kernel-based fuzzy inference systems for high dimensional regression, Neurocomputing, 77(1), 197–204, (2012).

[7] W. Chung, A Fuzzy Convex Nonparametric Least Squares Method with Different Shape Constraints, International Journal of Fuzzy Systems, 1–15, (2023).

[8] G. E. del Pino, H. Galaz, Statistical applications of the inverse gram matrix: A revisitation, Brazilian Journal of Probability and Statistics, 177–196, (1995).

[9] P. Diamond, R. Korner, Extended fuzzy linear models and least squares estimates, Computers & Mathematics with Applications, 33(9), 15–32, (1997).

[10] P. Drineas, M. W. Mahoney, Approximating a gram matrix for improved kernel-based learning, In: International Conference on Computational Learning Theory, 323–337, (2005).

[11] A. I. Iacob, C. C. Popescu, Regression using partially linearized gaussian fuzzy data, In: Informatics Engineering and Information Science: International Conference, ICIEIS 2011, Kuala Lumpur, Malaysia, November 14-16, 2011. Proceedings, Part II, 584–595, (2011).

[12] E. J. Kansa, Radial basis functions: achievements and challenges, WIT Transactions on Modelling and Simulation, 61, 3–22, (2015).

[13] A. H. Khammar, M. Arefi, M. G. Akbari, A general approach to fuzzy regression models based on different loss functions, Soft Computing, 25(2), 835–849, (2021).

[14] J. M. Leski, Fuzzy c-ordered-means clustering, Fuzzy Sets and Systems, 286, 114–133, (2016).

[15] Y. Li, X. He, X. Liu, Fuzzy multiple linear least squares regression analysis, Fuzzy Sets and Systems, 459, 118–143, (2023).

[16] D. A. Savic, W. Pedrycz, Evaluation of fuzzy linear regression models, Fuzzy sets and systems, 39(1), 51–63, (1991).

[17] H. Wendland, Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, Advances in computational Mathematics, 4(1), 389–396, (1995).

[18] H. Wendland, Scattered data approximation, Cambridge, UK: Cambridge University Press, (2005).

[19] K. Wiktorowicz, T. Krzeszowski, Approximation of two-variable functions using high-order TakagiSugeno fuzzy systems, sparse regressions, and metaheuristic optimization, Soft Computing, 24(20), 15113–15127, (2020).

[20] Z. Wu, Compactly supported positive definite radial functions, Advances in Computational Mathematics, 4(1), 283–292, (1995).

[21] M. S. Yang, C. H. Ko, On a class of fuzzy c-numbers clustering procedures for fuzzy data, Fuzzy sets and systems, 84(1), 49–60, (1996).

[22] L. A. Zadeh, Toward extended fuzzy logicA first step, Fuzzy sets and systems, 160(21), 3175–3181, (2009).