Iranian Journal of Numerical Analysis and Optimization
Vol. 12, No. 3 (Special Issue), 2022, pp 497–512
DOI:10.22067/ijnao.2022.73876.1079
https://ijnao.um.ac.ir/



How to cite this article
**Research Article**

# Estimation of the regression function by Legendre wavelets

M. Hamzehnejad* , M.M. Hosseini and A. Salemi

**Abstract**

We estimate a function $f$ with $N$ independent observations by using Legendre wavelets operational matrices. The function $f$ is approximated with the solution of a special minimization problem. We introduce an explicit expression for the penalty term by Legendre wavelets operational matrices. Also, we obtain a new upper bound on the approximation error of a differentiable function $f$ using the partial sums of the Legendre wavelets. The validity and ability of these operational matrices are shown by several examples of real-world problems with some constraints. An accurate approximation of the regression function is obtained by the Legendre wavelets estimator. Furthermore, the proposed estimation is compared with a non-parametric regression algorithm and the capability of this estimation is illustrated.

**AMS subject classifications (2020):** 65T60; 41A30; 65D10; 62G08.

**Keywords:** Legendre wavelet; Operational matrix; Wavelet approximation; Regression function; Error analysis.

---

* Corresponding author

Received 29 November 2021; revised 5 March 2022; accepted 9 March 2022

Mehdi Hamzehnejad
Department of Mathematic, Graduate University of Advanced Technology, Kerman, Iran. e-mail: mhdhamzehnejad@gmail.com

Mohammad Mehdi Hosseini
Department of Applied Mathematics and Mahani Mathematical Research Center, Shahid Bahonar University of Kerman, Kerman, Iran. e-mail: Hossem25@gmail.com

Abbas Salemi
Department of Applied Mathematics and Mahani Mathematical Research Center, Shahid Bahonar University of Kerman, Kerman, Iran. e-mail: Salemi@uk.ac.ir

## 1 Introduction

Let $f : [a, b] \longrightarrow \mathbb{R}$ with independent observations $\{(x_i, y_i), i = 1, \ldots, N\}$. Consider the following nonparametric regression model to provide an estimate for $f$:

$$y_i = f(x_i) + \epsilon_i, \tag{1}$$

where $x_i \in [a, b]$ and $\epsilon_i$ have Gaussian noise. It is well known that the following optimization problem approximate the regression function $f$ [7, 13]:

$$\min_{f \in \mathbf{S}} \frac{1}{N} \sum_{i=1}^{N} (y_i - f(x_i))^2 + \frac{\lambda}{N} \int_a^b (f^{(r)}(x))^2 dx, \tag{2}$$

where $\mathbf{S}$ denotes the set of functions $f$ satisfying the constraints and the constant $\lambda$ is called smoothing parameter. The first term measures closeness to the data, while the second term penalizes curvature in the function. This optimization problem appears in many branches of applied mathematics including economics, stochastic processes, statistics, machine learning, and control theory, and several studies have been conducted to determine the function $f$ [7, 9, 18, 5, 13].

Using linear combinations of basis functions, such as orthogonal polynomials, wavelets, and splines is a popular approach to estimating the function $f$ [7, 18, 5, 17, 11, 3, 6, 16]. This kind of method can be expressed as a matrix equation that contains a penalty term. Although it is not possible to get a clear and accurate answer to this problem, it is necessary to use approximate methods to solve it. Calculating the penalty term is an important issue for the authors. Wand and Ormerod [18] obtained an exact explicit expression for each entry of the penalty matrix by solving numerical integrals.

It is well known that a single method cannot work for all functions without any restrictions. Some of these restrictions include monotonicity, convexity, unimodality, or combinations of several types of constraints. For example, Mammen et al. [8] considered the regression function under the monotonicity constraint and Meyer [9] considered the regression function under constraints of convexity and monotone. Also in [1, 12], the authors considered the regression function under combinations of several types of restrictions.

In this paper, by using properties of the Legendre wavelets, we provide an exact explicit expression for the penalty term only by matrix multiplications, which reduce the complexity of the problem. Also, an accurate approximation of differentiable functions is obtained by Legendre wavelets. For this purpose, we provide an upper bound for the first term of (2). Moreover, by using the examples that have been mentioned in [9, 1, 4], we show that the Legendre wavelets are a good candidate for the estimation of regression functions under various constraints.

The rest of this paper is organized as follows. In Section 2, we state some definitions and properties of the Legendre wavelets. Furthermore, we recall the operational matrix of derivatives, and by using this operational matrix, we provide an exact explicit expression for the penalty matrix. In Section 2, a new upper bound on the approximation error of the partial sums of the Legendre wavelets is presented. In Section 3, the performance of the proposed estimation is compared with a nonparametric regression method, by numerical examples.

## 2 Legendre polynomials and wavelets

In this section, we study Legendre polynomials and wavelets by presenting some necessary definitions and theorems. The well-known Legendre polynomials are defined on the interval $[-1, 1]$ and can be determined by the following recurrence formulas [15].

$$(m+1)L_{m+1}(x) = (2m+1)xL_m(x) - mL_{m-1}(x), \quad m = 1, 2, 3, \dots,$$

where $L_0(t) = 1$ and $L_1(x) = x$. The following relation is hold for Legendre polynomials [15, eq. 3.176a]

$$L_m(x) = \frac{1}{2m+1}\left(L'_{m+1}(x) - L'_{m-1}(x)\right). \tag{3}$$

Moreover, we have the following uniform bound for Legendre polynomials [15]

$$|L_m(x)| \leq 1, \quad x \in [-1, 1],\ m \geq 0. \tag{4}$$

Legendre wavelets are defined on the interval $[0, 1]$ as follows:

$$\psi_{n,m}(t) = \begin{cases} \sqrt{(m+\frac{1}{2})}2^{\frac{k+1}{2}} L_m(2^{k+1}t - (2n+1)), & \frac{n}{2^k} \leq t < \frac{n+1}{2^k}, \\ 0, & \text{otherwise}, \end{cases}$$

where $k \in \mathbb{N}$, $m = 0, 1, \dots, M-1$, and $n = 0, 1, \dots, 2^k - 1$. The Legendre wavelets are an orthonormal basis for $L^2[0, 1]$ and the following orthogonality holds:

$$\int_0^1 \psi_{m,n}(t)\psi_{r,s}(t)dt = \delta_{mr}\delta_{ns}.$$

Let $f(t) \in L^2[0, 1]$. Then

$$f(t) \simeq \sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(t) = C^T\Psi(t),$$

where $c_{n,m} = \int_0^1 f(t)\psi_{n,m}(t)dt$. The vectors $C$ and $\Psi(t)$ are $2^kM \times 1$ vectors given by

$$C = \left[c_{0,0}, \dots, c_{0,M-1}, c_{1,0}, \dots, c_{1,M-1}, \dots, c_{2^k-1,0}, \dots, c_{2^k-1,M-1}\right]^T,$$

$$\begin{aligned}\Psi(t) = &[\psi_{0,0}(t), \dots, \psi_{0,M-1}(t), \psi_{1,0}(t), \dots, \psi_{1,M-1}(t),\\ &\dots, \psi_{2^k-1,0}(t), \dots, \psi_{2^k-1,M-1}(t)]^T.\end{aligned}$$

The Legendre wavelets approximation finds a shape constrained $f$ to the minimization problem (2). In the minimization problem (2), we set

$$f(t) \simeq \sum_{n=0}^{2^k-1} \sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(t).$$

For simplicity, we can set$\psi_{((i-1)\times M)+j+1}(t) := \psi_{i,j}(t)$ and $c_{((i-1)\times M)+j+1} := c_{ij}$ for $i = 1, \dots, 2^k$ and $j = 0, \dots, M-1$. Hence the following vectors are obtained:

$$\Psi(t) = [\psi_1(t), \dots, \psi_{2^kM}(t)]^T, \quad C = [c_1, c_2, \dots, c_{2^kM}]^T. \tag{5}$$

Therefore, we have

$$f(t) = \sum_{j=1}^{2^kM} c_j\psi_j(t),$$

where $\psi_j(t)$ are the Legendre wavelets. Therefore the objective function to minimize (2) is the following penalized least square:

$$\min_{c_j} \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \sum_{j=1}^{2^kM} c_j\psi_j(x_i)\right)^2 + \frac{\lambda}{N}\int_0^1 \left(\sum_{j=1}^{2^kM} c_j\psi_j^{(r)}(t)\right)^2 dt,$$

where

$$\int_0^1 \left(\sum_{j=1}^{2^kM} c_j\psi_j^{(r)}(t)\right)^2 dt = \sum_{i=1}^{2^kM}\sum_{j=1}^{2^kM} c_ic_j \int_0^1 \psi_i^{(r)}(t)\psi_j^{(r)}(t)dt.$$

Suppose that $V$ is a matrix by elements of the form $V_{ij} := \frac{1}{N}\sum_{l=1}^{N}\psi_i(x_l)\psi_j(x_l)$, that $P$ is a matrix by elements $P_{ij} = \int_0^1 \psi_i^{(r)}(t)\psi_j^{(r)}(t)dt$, and that the elements of vector $b$ are defined by $b_i = \frac{1}{N}\sum_{l=1}^{N}\psi_i(x_l)y_l,\ i,j = 1, \dots, 2^{k-1}M$, so the minimization problem (2) has the following quadratic form of minimization [5]:

$$\min_{C\in\mathbb{R}^{2^k M}} \frac{1}{2}C^T VC - bC + \lambda(\frac{1}{2}C^T PC). \tag{6}$$

By taking the derivative of (6) in terms of $C$ and put it equal zero, we obtain the following equation:

$$(V + \lambda P)C = b. \tag{7}$$

Now focus on the second term, to determine an appropriate operator matrix to solve the problem (2). An important issue is to calculate the elements of the matrix $P$. We use Legendre wavelets operational matrix of derivative, to get the new structure of the matrix $P$. The following theorems determine the Legendre wavelet operational matrices of derivatives, which are used to solve differential equations.

**Theorem 1.** [10, Theorem 1] Let $\Psi(t)$ be the Legendre wavelets vector as in (5). Then the derivative of the vector $\Psi(t)$ can be expressed by

$$\frac{d\Psi(t)}{dt} = D\Psi(t),$$

where $D$ is the $2^k M$ operational matrix

$$D = \begin{bmatrix} F & 0 & \cdots & 0 \\ 0 & F & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F \end{bmatrix},$$

where $F$ is an $M \times M$ matrix such that $(r, s)$th entry of $F$ is defined as follows:

$$F_{r,s} = \begin{cases} 2^{k+1}\sqrt{(2r-1)(2s-1)}, \begin{cases} r = 2, \dots, M, \\ s = 1, \dots, r-1, \end{cases} & (r+s) \text{ odd}, \\ 0, & \text{otherwise.} \end{cases}$$

**Theorem 2.** [10, Theorem 2] By using Theorem 1, the operational matrix for $n$th derivative can be derived as

$$\frac{d^n\Psi(t)}{dt^n} = D^n\Psi(t),$$

where $D^n$ is the $n$th power of the matrix $D$.

Therefore, using these operational matrices, the elements of the matrix $P$ in (7) are introduced in the next theorem.

**Theorem 3.** Let $\Psi(t)$ be the Legendre wavelets vector defined in (5). Assume that $r$ is a nonnegative integer and that the elements of the matrix $P = [P_{ij}]$ are $P_{ij} = \int_0^1 \psi_i^{(r)}(t)\psi_j^{(r)}(t)dt$. Then $P_{ij}$ has the following exact explicit expression

$$P_{ij} = (D_i^r)(D_j^r)^T, \quad i,j = 1,\ldots,2^k M, \tag{8}$$

where $D_i^r$ is the $i$th row of the operational matrix $D^r$ as in Theorem 2.

*Proof.* By using Theorem 2, the elements of the matrix $P$ are as follows:

$$P_{ij} = \int_0^1 \psi_i^{(r)}(t)\psi_j^{(r)}(t)dt = \int_0^1 (D_i^r \Psi(t))(D_j^r \Psi(t))dt, \quad i,j = 1,\ldots,2^k M. \tag{9}$$

Let $D_i^r \Psi(t) = \sum_{s=1}^{2^k M} d_{is}^{(r)} \psi_s(t)$. Then

$$\begin{aligned} P_{ij} &= \int_0^1 \left(d_{i1}^{(r)}\psi_1(t) + \cdots + d_{i2^kM}^{(r)}\psi_{2^kM}(t)\right)\left(d_{j1}^{(r)}\psi_1(t) + \cdots + d_{j2^kM}^{(r)}\psi_{2^kM}(t)\right) dt \\ &= \int_0^1 \sum_{s=1}^{2^kM}\sum_{l=1}^{2^kM} d_{is}^{(r)} d_{jl}^{(r)} \psi_s(t)\psi_l(t)dt = \sum_{s=1}^{2^kM}\sum_{l=1}^{2^kM} d_{is}^{(r)} d_{jl}^{(r)} \int_0^1 \psi_s(t)\psi_l(t)dt. \end{aligned}$$

According to the property of orthogonality, we have

$$\int_0^1 \psi_s(t)\psi_l(t)dt = \delta_{sl}. \tag{10}$$

By using (10), $P_{ij} = \sum_{s=1}^{2^kM} d_{is}^{(r)} d_{js}^{(r)} = (D_i^r)(D_j^r)^T$. □

Therefore, we can calculate the elements of the matrix $P$ only by a matrix multiplication. By solving system (7), the appropriate weight coefficients are obtained to approximate the function $f$.

## 3 Error analysis

In this section, we present an error estimate of the partial sums of Legendre wavelets to the regression function $f$. For this purpose, we benefit from the well-known mean-square error (MSE). By using the MSE [16], we measure the performance of the estimator $\hat{f}$ as follows:

$$MSE(\hat{f}, f) = \frac{1}{N}\sum_{i=1}^{N} E\left[\hat{f}(x_i) - f(x_i)\right]^2.$$

The Legendre wavelets estimator $\hat{f}$ can be written as

$$\hat{f} = (\hat{f}(x_1),\ldots,\hat{f}(x_N)) = \left(\sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(x_1),\ldots,\sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(x_N)\right).$$

We present a new approximation error of the function $f$, using the partial sums of Legendre wavelets. We know that

$$
\begin{aligned}
f(t) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_{n,m}\psi_{n,m}(t) \\
&= \sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(t) + \sum_{n=0}^{2^k-1}\sum_{m=M}^{\infty} c_{n,m}\psi_{n,m}(t) + \sum_{n=2^k}^{\infty}\sum_{m=0}^{\infty} c_{n,m}\psi_{n,m}(t).
\end{aligned}
\tag{11}
$$

The last part in (11), $\sum_{n=2^k}^{\infty}\sum_{m=0}^{\infty} c_{n,m}\psi_{n,m}(t) = 0$, because the Legendre wavelets $\psi_{n,m}(t)$ are zero outside of the interval $[0,1]$. Then

$$
\begin{aligned}
\left\| f(t) - \sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(t) \right\|^2 &= \left\| \sum_{n=0}^{2^k-1}\sum_{m=M}^{\infty} c_{n,m}\psi_{n,m}(t) \right\|^2 \\
&\leq \sum_{n=0}^{2^k-1}\sum_{m=M}^{\infty} \left|c_{n,m}\right|^2 \left\|\psi_{n,m}(t)\right\|^2 .
\end{aligned}
$$

We know that $\left\|\psi_{n,m}(t)\right\|^2 = 1$. Therefore

$$
\left\| f(t) - \hat{f}(t) \right\|^2 = \left\| f(t) - \sum_{n=0}^{2^k-1}\sum_{m=0}^{M-1} c_{n,m}\psi_{n,m}(t) \right\|^2 \leq \sum_{n=0}^{2^k-1}\sum_{m=M}^{\infty} \left|c_{n,m}\right|^2 . \tag{12}
$$

Hence, the approximation error of the truncated series of Legendre wavelets depends on the Legendre wavelets coefficients $c_{n,m}$. Now, we obtain an upper bound for Legendre wavelets coefficients.

**Theorem 4.** Suppose that $k \in \mathbb{N}$ and that $f, f', \dots, f^{(r)}$ are absolutely continuous on $[0,1]$. Suppose that $V = \max\left\{V_n,\ n = 0,\dots,2^k-1\right\}$, where

$$
V_n = \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} \left| f^{(r+1)}(t) \right| dt, \qquad n = 0,1,\dots,2^k-1.
$$

Then for $m \geq r+1$,

$$
\left|c_{n,m}\right| \leq \begin{cases} \dfrac{V}{2^{rk}(2m-2r+3)\cdots(2m-1)(2m+3)\cdots(2m+2r-1)\sqrt{2^k(2m-2r+1)}}, & r \text{ odd}, \\[2ex] \dfrac{V}{2^{rk}(2m-2r+3)\cdots(2m+1)(2m+5)\cdots(2m+2r-1)\sqrt{2^k(2m-2r+1)}}, & r \text{ even}. \end{cases} \tag{13}
$$

*Proof.* For each $0 \leq i \leq r$, define the following sequence

$$\begin{aligned} c^{(i)}_{n,m} &= \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} f^{(i)}(t)\psi_{n,m}(t)dt \\ &= \sqrt{(m+\frac{1}{2})}2^{\frac{k+1}{2}} \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} f^{(i)}(t)L_m(2^{k+1}t-(2n+1))dt, \end{aligned} \tag{14}$$

where $c^{(0)}_{n,m} = c_{n,m}$. Let $x = 2^{k+1}t - (2n+1)$. Then

$$\begin{aligned} c^{(r+1)}_{n,m} &= \sqrt{(m+\frac{1}{2})}2^{\frac{k+1}{2}} \int_{-1}^{1} f^{(r+1)}\left(\frac{x+2n+1}{2^{k+1}}\right) L_m(x)\frac{dx}{2^{k+1}} \\ &= \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}} \int_{-1}^{1} f^{(r+1)}\left(\frac{x+2n+1}{2^{k+1}}\right) L_m(x)dx. \end{aligned} \tag{15}$$

By using the equation (3), we have

$$c^{(r)}_{n,m} = \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}(2m+1)} \int_{-1}^{1} f^{(r)}\left(\frac{x+2n+1}{2^{k+1}}\right)\left(L'_{m+1}(x) - L'_{m-1}(x)\right)dx. \tag{16}$$

Using integration by parts, we have

$$\begin{aligned} c^{(r)}_{n,m} &= \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}(2m+1)} \left[f^{(r)}\left(\frac{x+2n+1}{2^{k+1}}\right)(L_{m+1}(x) - L_{m-1}(x))\right]_{-1}^{1} \\ &+ \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}2^{k+1}(2m+1)} \int_{-1}^{1} f^{(r+1)}\left(\frac{x+2n+1}{2^{k+1}}\right)(L_{m+1}(x) - L_{m-1}(x))\,dx. \end{aligned} \tag{17}$$

Using the properties $L_m(1) = 1^m$ and $L_m(-1) = (-1)^m$ for $m \geq 0$, easy computations shows that the first term of (17) vanishes. Thus we have

$$c^{(r)}_{n,m} = \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}2^{k+1}(2m+1)} \int_{-1}^{1} f^{(r+1)}\left(\frac{x+2n+1}{2^{k+1}}\right)(L_{m+1}(x) - L_{m-1}(x))\,dx. \tag{18}$$

From (14) and (18), we obtain the following relation between the coefficients

$$c^{(r)}_{n,m} = \frac{1}{2^{k+1}(2m+1)}\left(c^{(r+1)}_{n,m-1} - c^{(r+1)}_{n,m+1}\right). \tag{19}$$

Now, we obtain an upper bound for $c^{(r+1)}_{n,m}$. We can see that

$$c_{n,m}^{(r+1)} = \frac{\sqrt{(m+\frac{1}{2})}}{2^{\frac{k+1}{2}}} \int_{-1}^{1} f^{(r+1)}\left(\frac{x+2n+1}{2^{k+1}}\right) L_m(x)dx$$
$$= \sqrt{(m+\frac{1}{2})2^{k+1}} \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} f^{(r+1)}(t) L_m(2^{k+1}t-(2n+1))dt.$$

From (9) and by easy computation, we obtain

$$\left|c_{n,m}^{(r+1)}\right| = \sqrt{(m+\frac{1}{2})2^{k+1}} \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} \left|f^{(r+1)}(t)\right| \left|L_m(2^{k+1}t-(2n+1))\right| dt$$
$$\leq \sqrt{(2m+1)2^k} \int_{\frac{n}{2^k}}^{\frac{n+1}{2^k}} \left|f^{(r+1)}(t)\right| dt \leq V\sqrt{2^k(2m+1)}. \tag{20}$$

Applying (20) in (19), we have

$$\left|c_{n,m}^{(r)}\right| \leq \frac{1}{2^{k+1}(2m+1)} \left(\left|c_{n,m-1}^{(r+1)}\right| + \left|c_{n,m+1}^{(r+1)}\right|\right)$$
$$\leq \frac{V\sqrt{2^k(2m-1)} + V\sqrt{2^k(2m+3)}}{2^{k+1}(2m+1)}. \tag{21}$$

Since

$$\sqrt{2m-1} + \sqrt{2m+3} \leq 2\sqrt{2m+1},$$

(21) becomes to

$$\left|c_{n,m}^{(r)}\right| \leq \frac{2V\sqrt{2^k(2m+1)}}{2^{(k+1)}(2m+1)} = \frac{V}{\sqrt{2^k(2m+1)}}. \tag{22}$$

Also, by using (22) in (19), we obtain the following upper bound for $c_{n,m}^{(r-1)}$:

$$\left|c_{n,m}^{(r-1)}\right| \leq \frac{1}{2^{k+1}(2m+1)} \left(\left|c_{n,m-1}^{(r)}\right| + \left|c_{n,m+1}^{(r)}\right|\right)$$
$$\leq \frac{1}{2^{k+1}(2m+1)} \left(\frac{V}{\sqrt{2^k(2m-1)}} + \frac{V}{\sqrt{2^k(2m+3)}}\right)$$
$$= \frac{V}{2^{k+1}(2m+1)\sqrt{2^k}} \left(\frac{\sqrt{(2m+3)} + \sqrt{(2m-1)}}{\sqrt{(2m-1)(2m+3)}}\right)$$
$$\leq \frac{2V\sqrt{(2m+3)}}{2^{k+1}(2m+1)\sqrt{2^k(2m-1)(2m+3)}}$$
$$= \frac{V}{2^k(2m+1)\sqrt{2^k(2m-1)}}.$$

If we continue the above process, then by easy computation for an integer $s \geq 2$, we obtain the following upper bound for $c_{n,m}^{(r-s-1)}$:

$$|c_{n,m}^{(r-s)}| \leq \begin{cases} \frac{V}{2^{(s-1)k}(2m-2s+5)\cdots(2m-1)(2m+3)\cdots(2m+2s-3)\sqrt{2^k(2m-2s+3)}}, & s \text{ odd}, \\ \frac{V}{2^{(s-1)k}(2m-2s+5)\cdots(2m+1)(2m+5)\cdots(2m+2s-3)\sqrt{2^k(2m-2s+3)}}, & s \text{ even}. \end{cases}$$

Then (13) holds when $s+1=r$. □

Now, we are ready to provide an approximation error of the partial sums of Legendre wavelets. We show that if the regression function $f$ is smooth, then the partial sums of Legendre wavelets converge to it rapidly.

**Theorem 5.** Suppose that $k \in \mathbb{N}$ and that $f, f', \dots, f^{(r)}$ are absolutely continuous on $[0,1]$. Moreover, suppose that $E_{k,M}(f(t)) = \left\| f(t) - \hat{f}(t) \right\|$. Then for $M \geq r+1$ and $r \geq 1$,

$$E_{k,M}(f(t)) \leq \begin{cases} \frac{V}{r2^{(r-1)k}(2M-2r+1)\cdots(2M-1)(2M+3)\cdots(2M+2r-7)\sqrt{2^k(2M-2r+1)}}, & r \text{ odd}, \\ \frac{V}{r2^{(r-1)k}(2M-2r+1)\cdots(2M+1)(2M+5)\cdots(2M+2r-7)\sqrt{2^k(2M-2r+1)}}, & r \text{ even}. \end{cases}$$

*Proof.* Let $r$ be an odd integer. Applying (13) in (12), we obtain

$$\begin{aligned}
&E_{k,M}(f(t)) \\
&\leq \sum_{n=0}^{2^k-1} \sum_{m=M}^{\infty} \frac{V}{2^{rk}(2m-2r+3)\cdots(2m-1)(2m+3)\cdots(2m+2r-1)\sqrt{2^k(2m-2r+1)}} \\
&\leq \frac{V}{2^{rk}\sqrt{2^k(2M-2r+1)}} \sum_{n=0}^{2^k-1} \sum_{m=M}^{\infty} \frac{1}{(2m-2r+3)\cdots(2m-1)(2m+3)\cdots(2m+2r-1)} \\
&= \frac{V}{2^{rk}\sqrt{2^k(2M-2r+1)}} \sum_{n=0}^{2^k-1} \sum_{m=M}^{\infty} \frac{1}{(2m+2r-1)^{r-1}(1-\frac{4r-4}{(2m+2r-1)})\cdots(1-\frac{4}{(2m+2r-1)})} \\
&\leq \frac{V}{2^{rk}\sqrt{2^k(2M-2r+1)}(1-\frac{4r-4}{(2M+2r-3)})\cdots(1-\frac{4}{(2M+2r-3)})} \\
&\quad \sum_{n=0}^{2^k-1} \int_{M-1}^{\infty} \frac{1}{(2x+2r-1)^{r-1}} dx \\
&= \frac{2^k V}{r2^{rk}(2M-2r+1)\cdots(2M-1)(2M+3)\cdots(2M+2r-7)\sqrt{2^k(2M-2r+1)}} \\
&= \frac{V}{r2^{(r-1)k}(2M-2r+1)\cdots(2M-1)(2M+3)\cdots(2M+2r-7)\sqrt{2^k(2M-2r+1)}}
\end{aligned} \tag{23}$$

By a similar approach, the results hold for an even integer $r$ and complete the proof. □

**Remark 1.** The aim of this remark is to draw an approximation error for a function $f(x)$, using the partial sums of the Legendre wavelets. Consider two functions $f(x) = 1 + x - 0.45\exp[-5(x-0.5)^2]$ and $f(x) = \frac{1}{6}x^2|x|$. The function $f(x) = 1 + x - 0.45\exp[-5(x-0.5)^2]$ is infinitely differentiable. In Table 1, numerical results are shown for this function for some values of $M, k$, and $r$. The numerical results obtained from this table indicate that by increasing $M$, $k$, and $r$, the partial sums of Legendre wavelets converge to the function $f(x)$ rapidly. Also, consider the function $f(x) = \frac{1}{6}x^2|x|$

Table 1: Approximation errors of the function $f(x) = 1 + x - 0.45\exp[-5(x-0.5)^2]$ evaluated by Theorem 5.

| $M$ | $k$ | $r$ | $E_{k,M}(f(x))$ | $M$ | $k$ | $r$ | $E_{k,M}(f(x))$ |
|---|---|---|---|---|---|---|---|
| 10 | 1 | 3 | $1.920 \times 10^{-3}$ | 10 | 1 | 5 | $6.977 \times 10^{-5}$ |
| 15 | 2 | 3 | $5.669 \times 10^{-5}$ | 15 | 2 | 5 | $1.686 \times 10^{-7}$ |
| 20 | 3 | 3 | $3.373 \times 10^{-6}$ | 20 | 3 | 5 | $8.670 \times 10^{-10}$ |

[19]. This function and its derivatives are absolutely continuous on $[0,1]$ and $f^{(2)}(x) = |x|$. Also, $f^{(3)}(x) = 2H(x) - 1$, where $H(x)$ is the Heaviside step function, which is of bounded variation and $f^{(4)}(x) = 2\delta(x)$, where $\delta(x)$ is the Dirac delta function. In Table 55, the numerical results are listed for some values of $M, k$, and $r$. Moreover, the logarithm of absolute errors is displayed

Table 2: Approximation errors of the function $f(x) = \frac{1}{6}x^2|x|$ evaluated by Theorem 5.

| $M$ | $k$ | $r$ | $E_{k,M}(f(x))$ | $M$ | $k$ | $r$ | $E_{k,M}(f(x))$ |
|---|---|---|---|---|---|---|---|
| 10 | 1 | 3 | $1.067 \times 10^{-4}$ | 10 | 2 | 3 | $1.887 \times 10^{-5}$ |
| 15 | 1 | 3 | $3.251 \times 10^{-5}$ | 15 | 2 | 3 | $5.747 \times 10^{-6}$ |
| 20 | 1 | 3 | $1.459 \times 10^{-5}$ | 20 | 2 | 3 | $2.579 \times 10^{-6}$ |

in Figure 1.

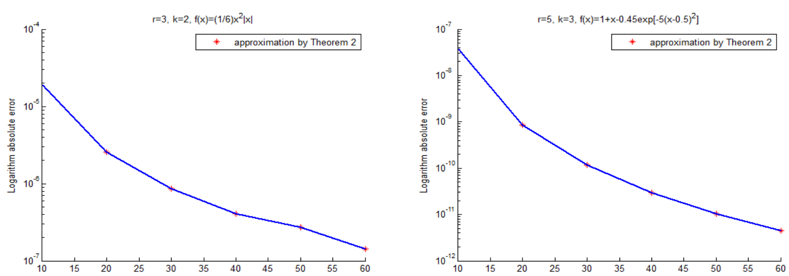


Figure 1: Approximation error of the functions $f(x) = \frac{1}{6}x^2|x|$ and $f(x) = 1 + x - 0.45\exp[-5(x-0.5)^2]$.

## 4 Numerical results

In this section, we present some examples to illustrate the validity and ability of the Legendre wavelets. For this purpose, we use some real-world test functions. Suppose that $(x_i, y_i), i = 1, \ldots, N$ are $N$ independent data with the same distribution such that $X_i, i = 1, \ldots, N$ have normal distribution, that is, $x_i \sim N(\mu, \sigma)$. Let $y_i = f(x_i) + \epsilon_i$ and let $x_i, \epsilon_i, f$ be independent with penalization order $r = 2$. We consider different kinds of regression functions, which have different constraints on interval $[0, 1]$.

**Remark 2.** Choosing the suitable smoothing parameter $\lambda$ is also an important issue in solving the minimization problem (2). Corlay [5] showed that $\lambda = \sigma_{x_i}^{2r-1}$ is a suitable smoothing parameter, where the quantity $\sigma_{x_i}$ is the standard deviation, which scales proportionally with $x_i$. Hence, in all examples, the coefficient of the penalty term $\frac{\lambda}{N} = \frac{\sigma_{x_i}^{2r-1}}{N}$ is used.

**Example 1.** Consider two real regression functions $f_1(x) = 15(x - 0.25)^2$ [9] and $f_2(x) = 1 + x - 0.45 \exp[-5(x - 0.5)^2]$ [4]. Then $f_1(x)$ is convex over $[0, 1]$ and $f_2(x)$ is strictly monotone over $[0, 1]$. Penalized Legendre wavelets regression of samples are plotted in Figure 2.

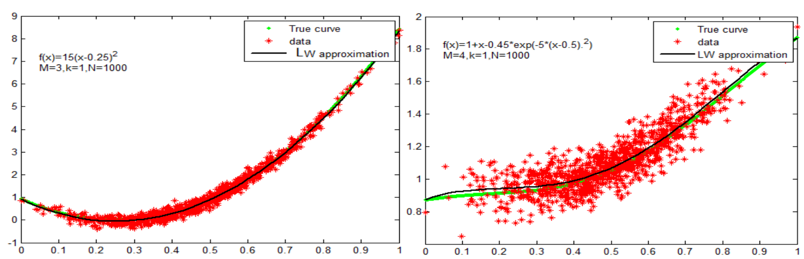


Figure 2: Approximate solution for the regression functions $f_1(x)$ and $f_2(x)$ in Example 1

**Example 2.** Consider the real regression function $f_3(x) = 15x^2 \sin(3.7x) + \frac{2}{\sigma\sqrt{2\pi}} \exp[(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)]$ [7, 1], where $\sigma = 0.1$ and $\mu = 0.3$. This function is unimodal (first increasing and then decreasing), concave on $[0.55, 1]$, and twice differentiable. We approximate the minimization problem (2) for $N = 1000$ samples of $(x_i, y_i)$. In Figure 3, the numerical results are shown.

**Example 3.** Consider the real regression function $f_4(x) = 10(x - 0.5)^3 - \exp[-100(x - 0.25)^2]$[4]. In Figure 4, the numerical results are shown.

In the following example, we compare our method by a nonparametric Regression (NR) method. NR methods are very sensitive to parameters such as

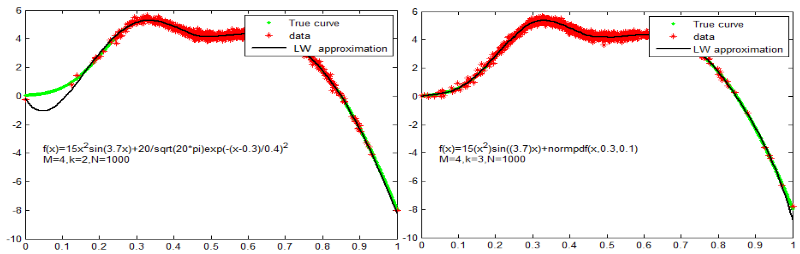


Figure 3: Approximate solution for the regression function $f_3(x)$.

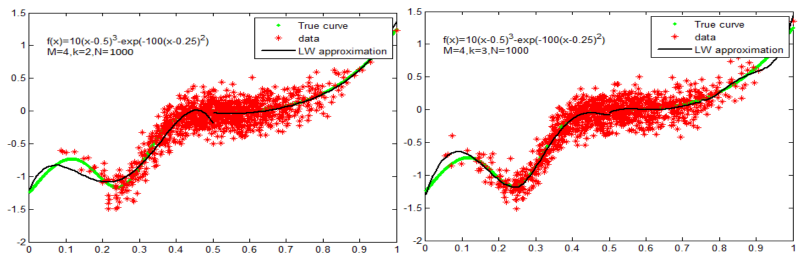


Figure 4: Approximate solution for the regression function $f_4(x)$.

the bandwidth selection, the regression order, and the shape of the smoothing kernel. In these methods, the choice of order and especially the bandwidth parameter can be a hassle [14]. In the previous example, we observed that the Legendre wavelets regression (LWR) method provides a good estimate for $N$ samples $(x_i, y_i)$, which does not depend on any parameter except the choice of $k$ and $M$, where $k$ specifies the level of resolution, $2^k$ sub-intervals on $[0, 1]$, and $M$ determines the degree of wavelets. Note that the selection of these two parameters is easy.

**Example 4.** Consider the functions $f_1(x) = 1 + x - 0.45 \exp[-5(x - 0.5)^2]$, $f_2(x) = 15x^2 \sin(3.7x) + \frac{2}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$ and $f_3(x) = -x^3 - x^2$. In Figure 5, we approximate the minimization problem (2) for $N = 250$ samples of $(x_i, y_i)$ and compare this method by a nonparametric regression method, which are shown in Figure 5.

## 5 Conclusion

In this paper, Legendre wavelets were used to approximate the regression function. A new operational matrix was introduced to simplify the minimization problem in (2), which is useful for new research in financial mathematics and numerical analysis. Moreover, a new approximation error of

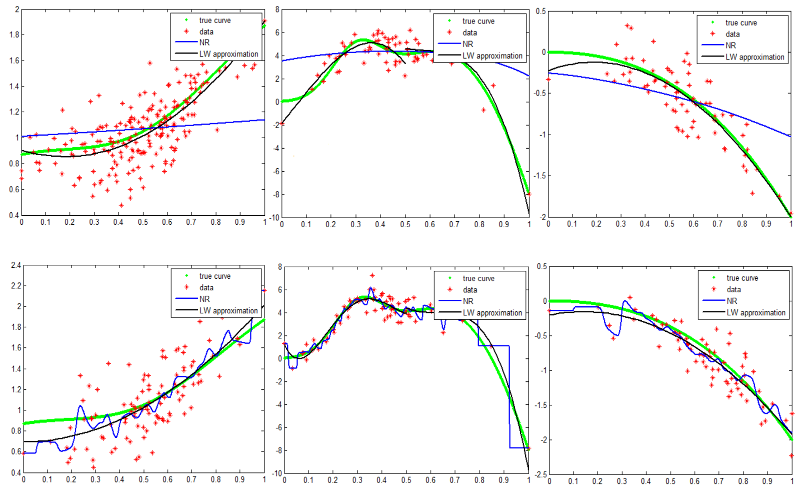


Figure 5: Comparing the Legendre wavelets estimation (black curve) with the nonparametric regression (blue curve). Due to nonoptimal choices of $h$, under-fitting occurred in the first row and over-fitting occurred in the second row for nonparametric regression for the functions mentioned in Example 3.

a differentiable function $f$ using the partial sums of the Legendre wavelets was provided. Numerical experiments were performed for a variety of real regression functions (see [9, 1, 4]). The proposed method was executed on some popular functions, and the numerical results were compared with the nonparametric regression method. Finally, the capability of the proposed method was illustrated.

## References


1. Abraham, C. *Bayesian regression under combinations of constraints,* J. Statist. Plann. Inference, 142 (2012) 2672–2687.

2. Abraham, C. and Khadraoui, K. *Bayesian regression with B-splines under combinations of shape constraints and smoothness properties*,Stat. Neerl. 69 (2015), 150–170.

3. Angelini, C., Canditiis, D.D. and Leblanc, F. *Wavelet regression estimation in nonparametric mixed effect models*, J. Multivariate Anal. 85 (2003) 267–291.

4. Bowman, W., Jones, M.C. and Gijbels, I. *Testing monotonicity of regression*, J. Comput. Graph Stat. 7 (1998) 489–500.

5. Corlay, S. *B-spline techniques for volatility modeling,* J. Comput. Finance, 19 (2016) 97–135.

6. Hamzehnejad, M., Hosseini, M.M. and Salemi, A. *An improved upper bound for ultraspherical coefficients*, Journal of Mathematical Modeling, 10 (2022), 1–11.

7. Khadraoui, K. *A smoothing stochastic simulated annealing method for localized shapes approximation*, JJ. Math. Anal. Appl. 446 (2017), 1018–1029.

8. Mammen, E., Marron, J., Turlach, B. and Wand, M. *A general projection framework for constrained smoothing*, Stat. Sci., 16 (2001) 232–248.

9. Meyer, M.C. *Inference using shape-restricted regression splines,* Ann. Appl. Stat. 2 (2008), 1013–1033.

10. Mohammadi, F. and Hosseini, M.M. *A new Legendre wavelet operational matrix of derivative and its applications in solving the singular ordinary differential equations,* J. Franklin Inst. 348 (2011) 1787–1796.

11. Mohammadi, M. and Bahrkazemi, M. *Bases for polynomial-based spaces*,J. Math. Model. 7 (2019) 21–34.

12. Polpo, A., Louzada, F.,Rifo, L.L.R., Stern, J.M. and Lauretto, M. *Interdisciplinary Bayesian Statistics*, Proceedings of the 12th Brazilian Meeting on Bayesian Statistics (EBEB 2014) held in Atibaia, March 10–14, 2014. Springer Proceedings in Mathematics & Statistics, 118. Springer, Cham, 2015.

13. Rasmussen, C.E. and Williams, C.K.I. *Gaussian processes for machine learning,* Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2006.

14. Raykar, C. and Duraiswami, R. *Fast optimal bandwidth selection for kernel density estimation,* Proceedings of the Sixth SIAM International Conference on Data Mining, 524–528, SIAM, Philadelphia, PA, 2006.

15. Shen, J., Tang, T. and Wang, L.L. *Spectral methods: Algorithms, analysis and applications*, Vol. 41. Springer Science & Business Media, 2011.

16. Vidakovic, B. *Statistical modeling by wavelets*, Wiley Series in Probability and Statistics: Applied Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999.

17. Wahba, G. *Spline models for observational data,* CBMS-NSF Regional Conference Series in Applied Mathematics, 59. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

18. Wand, M. and Ormerod, J. *On semiparametric regression with O'Sullivan penalized splines,* Aust. N. Z. J. Stat., 50 (2008) 179–198.

19. Wang, H. *A new and sharper bound for Legendre expansion of differentiable functions,* Appl. Math. Lett. 85 (2018) 95–102.

**How to cite this article**

M. Hamzehnejad, M.M. Hosseini and A. Salemi Estimation of the regression function by Legendre wavelets. *Iranian Journal of Numerical Analysis and Optimization*, 2022; 12(3 (Special Issue), 2022): 497-512. doi: 10.22067/ijnao.2022.73876.1079.