Modified least squares method and a review of its applications in machine learning and fractional differential/integral equations

Abhishek Kumar Singh[1,2], Mani Mehra*[1], and Anatoly A. Alikhanov[3]

[1]Deptartment of Mathematics, Indian Institute of Technology Delhi, India
[2]Institute of Mathematics and Computer Science, Universität Greifswald, Walther-Rathenau-Straße 47, 17489 Greifswald, Germany
[3]North-Caucasus Center for Mathematical Research, North-Caucasus Federal University, Russia
assinghabhi@gmail.com, mmehra@maths.iitd.ac.in, aaalikhanov@gmail.com

May 2, 2024

Abstract

The least squares method provides the best-fit curve by minimizing the total squares error. In this work, we provide the modified least squares method based on the fractional orthogonal polynomials that belong to the space $M_n^\lambda := \text{span}\{1, x^\lambda, x^{2\lambda}, \dots, x^{n\lambda}\},\ \lambda \in (0, 2]$. Numerical experiments demonstrate how to solve different problems using the modified least squares method. Moreover, the results show the advantage of the modified least squares method compared to the classical least squares method. Furthermore, we discuss the various applications of the modified least squares method in the fields like fractional differential/integral equations and machine learning.

Keywords. Modified least squares method; Müntz-Legendre polynomials; Machine learning; Fractional differential/integral equations.

1 Introduction

The least squares method is one of the oldest methods of modern statistics used to obtain the physical parameters from the experimental data. The first use of the least squares method is generally attributed to Gauss in 1795, although Legendre concurrently and independently used it [12]. Gauss invented the least squares method to estimate planets' orbital motion from telescopic measurements. In modern statistics, Galton [2] was the first to use the least squares method in his work on the heritability of size, which laid down the foundations of correlation and regression analysis. Nowadays, the least squares method is widely used to find the best-fit curve while finding the parameter involved in the curve. There are many versions of the least squares method available in the literature. The simpler version is called the ordinary least squares method, and the more advanced one is the weighted least squares method, which performs better than the ordinary least squares method. The recent version of the least squares method is the moving least squares method [6], and the partial least squares method [27].

One of the areas where the least squares method is frequently used is machine learning, where we analyze data for regression analysis and classification [24]. Machine learning is a field of artificial intelligence that allows computer systems to learn using available data. Recently machine learning algorithms (regression analysis and classification) have become very popular for analyzing data and making predictions. Another application of the least squares method is solving fractional differential/integral equations. Fractional differential/integral equations give an excellent way to deal with complex phenomena in nature, such as biological systems, control theory, finance, signal and image processing, sub-diffusion and

*Corresponding author

super-diffusion process, viscoelastic fluid, electrochemical processes, and so on [3, 13, 28, 26, 18]. The fractional differential equations are equivalent to the Hammerstein form of Volterra's second kind integral equations for the specific choice of kernel (for more details see [9]). Due to the importance of fractional differential/integral equations, people are interested in solving them numerically because of the non-availability of exact solutions. Many numerical methods are available in the literature to solve fractional differential/integral equations, such as finite difference [17, 19], compact finite difference [22, 4], finite element [8, 15] , and wavelet methods [16, 25]. Recently, the least squares method based on classical polynomials is also used to solve fractional differential/integral equations [23].

The aforementioned discussion concludes with the importance of the least squares method in various disciplines. In this work, we proposed the modified least squares method based on fractional polynomials. The main contribution of this paper consists of the following aspects:

- A modified least squares method based on fractional orthogonal polynomials belongs to the space $M_n^\lambda := \text{span}\{1, x^\lambda, x^{2\lambda}, \ldots, x^{n\lambda}\},\ \lambda \in (0, 2]$ has been proposed.
- Applications of the modified least squares method have been discussed in detail, especially in the field of fractional differential/integral equations and machine learning.

The paper has been arranged in the following pattern: Section 2 describes some basic concepts of approximation theory and orthogonal polynomials like Müntz-Legendre polynomials. Section 3 deals with the review of a least squares method. In Section 4, we develop a modified least squares method based on fractional orthogonal polynomials. Section 5 provides the numerical results to validate the modified least squares method. Section 6 summarizes the applications part of the method. Finally, Section 7 gives the brief conclusion.

2 Preliminaries

In this Section, some necessary results from approximation theory and orthogonal polynomials are summarized. Let define the space

$$M^\lambda := \text{span}\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}, \ldots\}, \tag{2.1}$$

and for its subspaces we define

$$M_n^\lambda := \text{span}\{x^{\lambda_0}, x^{\lambda_1}, \ldots, x^{\lambda_n}\}, \tag{2.2}$$

with $0 \le \lambda_0 < \lambda_1 < \ldots \to \infty$. The space M^λ is known as Müntz space. Now, we recall one of the fundamental theorems of approximation theory called Müntz-Szász theorem, which is related to the denseness of polynomial belonging to space M^λ.

Theorem 2.1. *(Müntz-Szász Theorem) The Müntz polynomials of the forms $\sum_{i=0}^n a_i x^{\lambda_i} \in M_n^\lambda$ with real coefficients are dense in $L^2[0,1]$ if and only if*

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i} = +\infty. \tag{2.3}$$

Moreover, if $\lambda_0 = 0$, then Müntz polynomials are dense in $C[0,1]$.

Proof. See [7].

In this work, we assume that $\lambda_i = i\lambda,\ i = 0, 1, \ldots, n, \ldots$, where λ is a real constant and $\lambda \in (0, 2]$. Now, we review one of the general forms of classical orthogonal polynomials called Jacobi polynomials. Moreover, we also review the Müntz-Legendre polynomials, and because of some advantage in terms of computational accuracy, we define the Müntz-Legendre polynomials with the help of Jacobi polynomials.

2.1 Jacobi polynomials

The Jacobi polynomials with parameters $a, b > -1$, denoted by $P_n^{a,b}$, is defined in the interval $-1 \le x \le 1$ as

$$P_n^{a,b}(x) = \sum_{i=0}^{n} \frac{(-1)^{n-i}(1+b)_n(1+a+b)_{i+n}}{i!(n-i)!(1+b)_i(1+a+b)_n} \left(\frac{1+x}{2}\right)^i, \tag{2.4}$$

where $(1+b)_i = (1+b)(2+b)\dots(i+b)$ and $(1+b)_0 = 1$.

In practice, one can compute the Jacobi polynomials using the following recurrence relation

$$
\begin{aligned}
P_0^{a,b}(x) = 1,\ \ P_1^{a,b}(x) = \frac{1}{2}[(a-b)+(a+b+2)x],\\
c_{1,n}^{a,b} P_n^{a,b}(x) = c_{2,n}^{a,b}(x) P_n^{a,b}(x) - c_{3,n}^{a,b} P_{n-1}^{a,b}(x),
\end{aligned} \tag{2.5}
$$

where

$$
\begin{aligned}
c_{1,n}^{a,b} &= 2(n+1)(n+a+b+1)(2n+a+b),\\
c_{2,n}^{a,b}(x) &= (2n+a+b+1)[(2n+a+b)(2n+a+b+2)x + a^2 - b^2],\\
c_{3,n}^{a,b} &= 2(n+a)(n+b)(2n+a+b+2).
\end{aligned} \tag{2.6}
$$

2.2 Müntz Legendre polynomials

One can defined Müntz Legendre polynomials on the interval $[0,1]$ as follows

$$
L_n(x) = \sum_{i=0}^{n} \eta_{n,i} x^{\lambda_i}, \quad \eta_{n,i} = \frac{\prod_{k=0}^{n-1}(\lambda_i + \bar{\lambda}_k + 1)}{\prod_{k=0,k\neq i}^{n}(\lambda_i - \lambda_k)}. \tag{2.7}
$$

These polynomials satisfy the following orthogonality condition with respect to weight function $W(x) = 1$

$$
\int_0^1 L_n(x) L_m(x) dx = \frac{\delta_{n,m}}{\lambda_n + \bar{\lambda}_n + 1}. \tag{2.8}
$$

Since we assumed that $\lambda_i = i\lambda$, where λ is a real constant, then the Müntz-Legendre polynomials on the interval $[0,1]$ are represented by the formula

$$
L_n(x;\lambda) := \sum_{i=0}^{n} \eta_{n,i} x^{i\lambda}, \quad \eta_{n,i} = \frac{(-1)^{n-i}}{\lambda^n i!(n-i)!} \prod_{k=0}^{n-1}((i+k)\lambda + 1). \tag{2.9}
$$

From Equation (2.7), one can observe that evaluating Müntz-Legendre polynomials, mainly when n is vast, and x is closed to 1, is problematic in finite arithmetic. Milovanovic has addressed these problems [20]. We will use a method for evaluating Müntz-Legendre polynomials, which is based on three-term recurrence relation induced from the accompanying theorem with the help of Jacobi polynomials.

Theorem 2.2. *Let $\lambda > 0$ be a real number and $x \in [0,1]$. Then the representation*

$$
L_n(x;\lambda) = P_n^{(0,1/\lambda-1)}\left(2x^\lambda - 1\right)
$$

holds true.

Proof. See [10]. □

So in view of Theorem 2.2 and Equations (2.5) and (2.7), the Müntz-Legendre polynomials $L_n(x;\lambda)$ can be evaluated by means of three-terms recurrence relation

$$
\begin{aligned}
L_0(x;\lambda) = 1, \quad L_1(x;\lambda) = \left(\frac{1}{\lambda}+1\right) x^\lambda - \frac{1}{\lambda},\\
d_{1,n} L_{n+1} = d_{2,n}(t) L_n(x;\lambda) - d_{3,n} L_{n-1}(x;\lambda),
\end{aligned} \tag{2.10}
$$

where

$$
d_{1,n} = c_{1,n}^{0,1/\lambda-1}, \quad d_{2,n}(x) = c_{2,n}^{0,1/\lambda-1}\left(2x^\lambda - 1\right), \quad d_{3,n} = c_{3,n}^{0,1/\lambda-1}.
$$