An Optimization Method for Numerically Solving Three-point BVPs of Linear Second-order ODEs with Variable Coefficients

Ming Hou, Liya Fan*
Liaocheng University
School of Mathematics Sciences
252059, Liaocheng
P.R. CHINA
houmingbb@126.com; fanliya63@126.com

Abstract: It is known that most numerical methods for solving differential equations are based on iterative methods or Taylor expansion methods. This paper tries to study a numerical method from a new perspective—optimization method. By means of the idea of kernel ε-SVR, the paper constructs an optimization model for a class of three-point boundary value problems (BVPs) of linear second-order ordinary differential equations (ODEs) with variable coefficients and proposes a novel numerical method for solving them. The proposed method has a certain versatility and can be used to solve some other kinds of differential equations and integral equations. In order to verify the effectiveness of the proposed method, comparative experiments with six specific linear second-order ODEs are performed. Experimental results show that the proposed method has a good approximation property.

Key–Words: Optimization modeling, numerical method, ordinary differential equation, kernel ε-support vector regression, Lagrange function

1 Introduction

Differential equations can be found in the mathematical formulation of physical phenomena in a wide variety of applications especially in science and engineering. Depending on the form of the boundary conditions to be satisfied by the solution, problems involving ordinary differential equations (ODEs) can be divided into two main categories, initial value problems (IVPs) and boundary value problems (BVPs). Exact solutions for these problems are not generally available and hence numerical methods must be applied.

Multi-point BVPs for ODEs can arise in solving linear partial differential equations by using the separation variable method [1]. Moreover, in the engineering problem to increase the stability of a rod, one also imposes a fixed interior point except for the ends of the rod [2]. Along this line, the solutions of multi-point BVPs for ODEs have great significance in mathematical theory and practical applications [3-4]. It is seen that the three-point BVPs of nonlinear second-order ODEs have attracted much attention [5-8] and it is significant to provide the solutions and Green's functions for these problems [9-10]. As shown in mentioned works, the existence and uniqueness of the solution are always focused on by using a fixed point theorem. However, from the viewpoint of practical applications, one should provide explicit solutions or approximate solutions for multi-point BVPs. In this paper, inspired by work of S. Mehrkanoon et. al [11], we try from a new perspective to study a class of three-point BVPs of linear second-order ODEs with variable coefficients

$$\begin{aligned} &y''(x) + p(x)y'(x) + q(x)y(x) = g(x), \\ &x \in [a, b], \\ &y(a) = p_0, \\ &y(b) + \lambda y(\mu) = q_0,\ \mu \in (a, b), \end{aligned} \tag{1}$$

where the known functions $p(x) \in C^1[a, b]$, $q(x), g(x) \in C[a, b]$ and λ, μ are the constants, and research how to obtain its approximal solutions by using an optimization method based on kernel support vector regression (KSVR).

Kernel support vector regressions (KSVRs) based on optimization modelings are a powerful methodology for solving pattern recognition and function estimation problems, which based on Vapnik and Chervonenkis structural risk minimization principle [12]. They show better generalization ability compared to other machine learning methods on a wide variety of real-world problems, such as optimal control [13], image segmentation [14], image denoising [15], time series prediction [16], cyber security problems [17], network traffic predictive [18] and so on.

*Corresponding author.

KSVRs are used to deal with nonlinear regression problems by means of kernel skill. Data can be mapped into a high dimensional feature space by using a kernel function and then linear regression problems can be solved, which results in solving quadratic programming problems (QPPs). The main challenge in developing a useful regression modeling is to capture accurately the underlying functional relationship between the given inputs and their output values. Once the resulting model is obtained, it can be used as a tool for analysis, simulation and prediction. KSVR aims at determining a regressor for a given set of data points and a kernel function. The basic idea of kernel ε-SVR with ε-insensitive loss function proposed by Vapnik [12] is to find a linear function $y(x)$ in the higher dimensional feature space such that, on the one hand, more mapped data samples locate in the ε-intensive tube between $y(x) - \varepsilon$ and $y(x) + \varepsilon$ and on the other hand, the function $y(x)$ is as flat as possible, which leads to introduce the regularization term. Thus, the structural risk minimization principle is implemented.

The paper mainly researches how to use kernel ε-SVR method to obtain numerical solutions of Eq.(1). It can be seen from the derivation process that some IVPs and BVPs of general linear ODEs may be studied by using the similar way. The paper does not involve nonlinear ODEs, which will be our next work. The remainder of this paper is organized as follows: Section 2 briefly reviews kernel ε-SVR and some basic concepts. Section 3 presents a new approximation method for the solution of Eq.(1). In order to verify the effectiveness of the presented method, a series of comparative experiments with six specific linear second-order ODEs are performed in Section 4 and some concluding remarks are given in Section 5.

2 Kernel ε-SVR

This section briefly recalls kernel ε-SVR, for details see [12]. Let $T = \{(x_i, y_i)\}_{i=1}^{l}$ be a set of sample data, where $x_i \in R^n$ and $y_i \in R$ are input and output of the ith sample data respectively. Let $k : R^n \times R^n \to R$ be a kernel function with the reproducing kernel Hilbert space (RKHS) H and the nonlinear feature mapping $\phi : R^n \to H$. The space H is a higher dimensional space and can be expressed as an expended space of the mapped inputs $\{\phi(x_i)\}_{i=1}^{l}$, that is, $H = \text{span}\{\phi(x_1), \cdots, \phi(x_l)\}$. It is known that $k(u, v) = \langle\phi(u), \phi(v)\rangle$ for all $u, v \in R^n$, where $\langle\cdot,\cdot\rangle$ denotes the inner product in RKBS H. Let $K = [k(x_i, x_j)] \in R^{l\times l}$ be the kernel matrix and K_i denote the ith column of K. ε-SVR is based on ε-insensitive loss function defined by

$$L_\varepsilon(y) = \begin{cases} 0 & \text{if } |f(x) - y| \le \varepsilon, \\ |f(x) - y| - \varepsilon & \text{otherwise,} \end{cases}$$

where $f(x)$ and y are the predicted value and true value of the input $x \in R^n$, respectively. The regression function obtained by using kernel ε-SVR has the form $f(x) = w^T\phi(x) + b$, where unknown normal vector $w \in H$ and bias $b \in R$ can be obtained by solving the following optimization problem:

$$\begin{aligned} &\min_{w,b} \tfrac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_{1i} + \xi_{2i}), \\ &s.t.\ w^T\phi(x_i) + b - y_i \le \varepsilon + \xi_{1i}, i = 1, \cdots, l, \\ &\quad y_i - (w^T\phi(x_i) + b) \le \varepsilon + \xi_{2i}, i = 1, \cdots, l, \\ &\quad \xi_{1i}, \xi_{2i} \ge 0, i = 1, \cdots, l, \end{aligned} \tag{2}$$

where $C > 0$ is a pre-specified value and ξ_{1i}, ξ_{2i}, $i = 1, \cdots, l$ are slack variables. Due to

$$H = \text{span}\{\phi(x_1), \cdots, \phi(x_l)\},$$

we can set $w = \sum_{i=1}^{l}\beta_i\phi(x_i)$. Put

$$\begin{aligned} &\beta = (\beta_1, \cdots, \beta_l)^T, \\ &y = (y_1, \cdots, y_l)^T, \\ &e_l = (1, \cdots, 1)^T \in R^l, \\ &\xi_1 = (\xi_{11}, \cdots, \xi_{1l})^T \in R^l, \\ &\xi_2 = (\xi_{21}, \cdots, \xi_{2l})^T \in R^l, \end{aligned}$$

then $w^T\phi(x_i) = K_i^T\beta$ and the problem (2) can be rewritten as the matrix form:

$$\begin{aligned} &\min_{w,b} \tfrac{1}{2}\beta^T K\beta + Ce_l^T(\xi_1 + \xi_2), \\ &s.t.\ K\beta + be_l - y \le \varepsilon e_l + \xi_1, \\ &\quad y - (K\beta + be_l) \le \varepsilon e_l + \xi_2, \\ &\quad \xi_1, \xi_2 \ge 0. \end{aligned} \tag{3}$$

By solving the Wolfe dual form of the problem (3):

$$\begin{aligned} &\min_{\alpha_1,\alpha_2} \tfrac{1}{2}(\alpha_1 - \alpha_2)^T K(\alpha_1 - \alpha_2) - y^T(\alpha_1 - \alpha_2) \\ &\quad +\varepsilon e_l^T(\alpha_1 + \alpha_2) \\ &s.t.\ e_l^T(\alpha_1 - \alpha_2) = 0, \\ &\quad 0 \le \alpha_1, \alpha_2 \le Ce_l, \end{aligned}$$

we can obtain the optimal solution (α_1^*, α_2^*) and then

$$\begin{aligned} &\beta^* = \alpha_1^* - \alpha_2^*, \\ &b^* = y_i - K_i^T\beta^* + \varepsilon \text{ for some } i : 0 < \alpha_{1i}^* < C, \end{aligned}$$

or

$$b^* = y_k - K_k^T\beta^* - \varepsilon \text{ for some } k : 0 < \alpha_{2k}^* < C.$$

Therefore, the optimal regression function is

$$f(x) = [k(x_1, x), \cdots, k(x_l, x)]\beta^* + b^*.$$

3 A numerical method based on Kernel ε-SVR

It is known that KSVRs based on optimization model are a powerful methodology for solving function estimation problems and show better generalization ability than other machine learning methods on a wide variety of real-world problems. But the main challenge in developing a useful regression model is to capture the underlying functional relationship between the given inputs and their output values accurately.

This section tries to study how to seek the numerical solutions of Eq.(1) by using kernel ε-SVR. Specifically, it assumes that numerical solutions of Eq.(1) have the form $\hat{y}(x) = w^T\phi(x) + b$ for a given kernel function $k : R \times R \to R$ with RKHS H and feature mapping $\phi : R \to H$ and then learns unknown $w \in H$ and $b \in R$ by kernel ε-SVR. For this end, the domain $[a, b]$ of Eq.(1) needs to be discretized into a set of collocation points $a = x_1 < x_2 < \cdots < x_l = b$ with the same stepsize $h = \frac{b-a}{l-1}$ such that $\mu = x_n \in (x_1, x_l)$, and then the points $\{x_i\}_{=1}^{l}$ are selected as input values. Since there are no available output values to learn (w, b) from Eq.(1), we have to substitute the function $\hat{y}(x) = w^T\phi(x) + b$ into Eq.(1). So, it needs to define the derivative of the kernel function. Making use of the Mercer's Theorem [19], derivatives of the feature map can be written in terms of derivatives of the kernel function [20]. Define the following differential operator which will be used in subsequent sections,

$$\nabla_n^m k(u, v) = \frac{\partial^{n+m} k(u, v)}{\partial u^n \partial v^m}, \forall u, v \in R, n, m = 1, 2, \cdots.$$

By the definition of kernel function, it has

$$\nabla_n^m k(u, v) = (\phi^{(n)}(u))^T \phi^{(m)}(v),$$

and then

$$\begin{aligned}
\hat{y}(x) &= w^T\phi(x) + b = \textstyle\sum_{j=1}^{l} \beta_j \phi(x_j)^T \phi(x) + b \\
&= \textstyle\sum_{j=1}^{l} \beta_j k(x_j, x) + b, \\
\hat{y}'(x) &= w^T\phi'(x) = \textstyle\sum_{j=1}^{l} \beta_j \phi(x_j)^T \phi'(x) \\
&= \textstyle\sum_{j=1}^{l} \beta_j \nabla_0^1 k(x_j, x), \\
\hat{y}''(x) &= w^T\phi''(x) = \textstyle\sum_{j=1}^{l} \beta_j \phi(x_j)^T \phi''(x) \\
&= \textstyle\sum_{j=1}^{l} \beta_j \nabla_0^2 k(x_j, x).
\end{aligned}$$

In the sequel, put

$$\begin{aligned}
K_0^0 &= [\nabla_0^0 k(x_i, x_j)] = [k(x_i, x_j)] = K \in R^{l\times l}, \\
K_n^m &= [\nabla_n^m k(x_i, x_j)] = [\nabla_n^m k(u, v)|_{u=x_i, v=x_j}] \in R^{l\times l},
\end{aligned}$$

and denote by $(K_n^m)_i$ and K_i the ith columns of the matrices K_n^m and K, respectively. Let

$$L(y(x)) = y''(x) + p(x)y'(x) + q(x)y(x).$$

If $\hat{y}(x)$ is an exact solution of Eq.(1), then $\hat{y}(a) = p_0, \hat{y}(b) + \lambda\hat{y}(\mu) = q_0$ and $L(\hat{y}(x_i)) = g(x_i)$ for all $i = 1, \cdots, l$. But in general, $\hat{y}(a), \hat{y}(b) + \lambda\hat{y}(\mu)$ and $L(\hat{y}(x_i))$ are not necessarily equal to p_0, q_0 and $g(x_i)$, respectively. So, we hope the smaller the better of the absolute values $|\hat{y}(a) - p_0|, |\hat{y}(b) + \lambda\hat{y}(\mu) - q_0|$ and $|L(\hat{y}(x_i)) - g(x_i)|$ for $i = 1, \cdots, l$. For this end, choose $\{g(x_i)\}_{i=1}^{l}$ as output values and construct the following optimization problem:

$$\begin{aligned}
&\min_{w,b} \tfrac{1}{2}\|w\|^2 + C\textstyle\sum_{i=1}^{l}(\xi_{1i} + \xi_{2i}), \\
&\quad s.t.\ L(\hat{y}(x_i)) - g(x_i) \le \varepsilon + \xi_{1i}, \\
&\qquad \xi_{1i} \ge 0,\ i = 1, \cdots, l, \\
&\qquad g(x_i) - L(\hat{y}(x_i)) \le \varepsilon + \xi_{2i}, \\
&\qquad \xi_{2i} \ge 0,\ i = 1, \cdots, l, \\
&\qquad \hat{y}(x_1) = p_0, \\
&\qquad \hat{y}(x_l) + \lambda\hat{y}(x_n) = q_0,
\end{aligned} \tag{4}$$

where $C > 0$ is a pre-specified value and $\xi_{1i}, \xi_{2i}, i = 1, \cdots, l$ are slack variables.

In the problem (4), it assumes the boundary conditions being satisfied and hopes the most values of $\{L(\hat{y}(x_i))\}_{i=1}^{l}$ satisfying $|L(\hat{y}(x_i)) - g(x_i)| < \varepsilon$ for some given sufficiently small $\varepsilon > 0$, which indicates that the errors between exact values and estimated values at most of discrete points $\{x_i\}_{i=1}^{l}$ are smaller than ε. By discussion in the previous section, it knows that the numerical solution $\hat{y}(x) = w^T\phi(x) + b$ is continuous if the kernel function $k : R \times R \to R$ is continuous. According to the symbol property of continuous functions, it can conclude that the error between exact solution and numerical solution in the whole domain $[a, b]$ is smaller than ε as long as the stepsize h is taken sufficiently small. In general, ε is often taken as $10^{-4} \sim 10^{-2}$. In this paper, take $\varepsilon = 10^{-2}$.

According to the discussion in Section 2, we can let $w = \sum_{j=1}^{l} \beta_j \phi(x_j)$ and then deduce that

$$\begin{aligned}
&\hat{y}(x_i) = \textstyle\sum_{j=1}^{l} \beta_j k(x_j, x_i) + b = K_i^T\beta + b, \\
&\hat{y}(x_l) + \lambda\hat{y}(x_n) = (K_l + \lambda K_n)^T\beta + (1 + \lambda)b, \\
&\hat{y}'(x_i) = (K_0^1)_i^T\beta, \\
&\hat{y}''(x_i) = (K_0^2)_i^T\beta, \\
&L(\hat{y}(x_i)) = [(K_0^2)_i + p(x_i)(K_0^1)_i + q(x_i)K_i]^T\beta + q(x_i)b, \\
&i = 1, \cdots, l.
\end{aligned}$$

Put

$$\begin{aligned}
&m_i = (K_0^2)_i + p(x_i)(K_0^1)_i + q(x_i)K_i \in R^l, i = 1, \cdots, l, \\
&M = [m_1, \cdots, m_l] \in R^{l\times l}, \\
&g = (g(x_1), \cdots, g(x_l))^T \in R^l, \\
&q = (q(x_1), \cdots, q(x_l))^T \in R^l,
\end{aligned}$$

then the problem (4) can be rewritten as the matrix

form:

$$\begin{aligned}\min_{\beta,b,\xi_1,\xi_2}\ & \tfrac{1}{2}\beta^T K\beta + Ce_l^T(\xi_1+\xi_2)\\ s.t.\ & M^T\beta + qb - g \le \varepsilon e_l + \xi_1,\ \xi_1 \ge 0,\\ & g - M^T\beta - qb \le \varepsilon e_l + \xi_2,\ \xi_2 \ge 0,\\ & K_1^T\beta + b = p_0,\\ & (K_l + \lambda K_n)^T\beta + (1+\lambda)b = q_0.\end{aligned} \tag{5}$$

Considering the Lagrange function of the problem (5)

$$\begin{aligned}&L(\beta,b,\xi_1,\xi_2,\eta_1,\eta_2,\varsigma_1,\varsigma_2,\alpha_1,\alpha_2)\\ &= \tfrac{1}{2}\beta^T K\beta + Ce_l^T(\xi_1+\xi_2)\\ &\quad +\eta_1^T(M^T\beta + qb - g - \varepsilon e_l - \xi_1)\\ &\quad +\eta_2^T(g - M^T\beta - qb - \varepsilon e_l - \xi_2)\\ &\quad -\varsigma_1{}^T\xi_1 - \varsigma_2{}^T\xi_2 + \alpha_1(K_1^T\beta + b - p_0)\\ &\quad +\alpha_2((K_l + \lambda K_n)^T\beta + (1+\lambda)b - q_0),\end{aligned}$$

where $\eta_1, \eta_2, \varsigma_1, \varsigma_2 \in R^l$ and $\alpha_1, \alpha_2 \in R$ are the vectors of Lagrange multipliers, and letting $\frac{\partial L}{\partial \beta} = \frac{\partial L}{\partial b} = \frac{\partial L}{\partial \xi_1} = \frac{\partial L}{\partial \xi_2} = 0$, it can deduce that

$$\begin{aligned}&K\beta + M(\eta_1 - \eta_2) + \alpha_1 K_1 + \alpha_2(K_l + \lambda K_n) = 0,\\ &q^T(\eta_1 - \eta_2) + \alpha_1 + (1+\lambda)\alpha_2 = 0,\\ &Ce_l - \eta_1 - \varsigma_1 = 0 \Rightarrow 0 \le \eta_1 \le Ce_l,\\ &Ce_l - \eta_2 - \varsigma_2 = 0 \Rightarrow 0 \le \eta_2 \le Ce_l.\end{aligned} \tag{6}$$

Since the kernel matrix K is symmetric and nonnegative definite, it can be assumed as a nonsingular matrix without loss of generality. Otherwise, to take care of problems due to possible ill-conditioning, a regularization term $K + \delta I_l$ can be introduced with a sufficiently small number $\delta > 0$ and a l order unit matrix I_l. Put

$$\begin{aligned}&D = [M, -M, K_1, K_l + \lambda K_n] \in R^{l\times(2l+2)},\\ &\gamma = [\eta_1^T, \eta_2^T, \alpha_1, \alpha_2]^T \in R^{2l+2},\\ &d = [q^T, -q^T, 1, (1+\lambda)]^T \in R^{2l+2}.\end{aligned}$$

It can get $\beta = -K^{-1}D\gamma$ from (6). Substituting (6) and $\beta = -K^{-1}D\gamma$ into the Lagrange function, it has

$$L(\gamma) = -\tfrac{1}{2}\gamma^T D^T K^{-1} D\gamma - z^T\gamma,$$

where

$$z = ((\varepsilon e_l + g)^T, (\varepsilon e_l - g)^T, p_0, q_0)^T \in R^{2l+2}.$$

Consequently, the Wolfe dual form of the problem (5) can be written as

$$\begin{aligned}\min_{\gamma}\ & \tfrac{1}{2}\gamma^T D^T K^{-1} D\gamma + z^T\gamma\\ s.t.\ & d^T\gamma = 0,\\ & 0 \le \gamma_j \le C,\ j = 1,\cdots,2l.\end{aligned} \tag{7}$$

After getting the optimal solution γ^* of the problem (7), it has

$$\begin{aligned}\beta^* &= -K^{-1}D\gamma^*,\\ b^* &= p_0 - K_1^T\beta^*,\end{aligned}$$

and then the numerical solution of Eq.(1) is

$$\hat{y}(x) = [k(x_1,x),\cdots,k(x_l,x)]\beta^* + b^*.$$

The specific procedure is as follows.

Algorithm 1.

Step 1. Discrete the domain $[a,b]$ by $a = x_1 < x_2 < \cdots < x_l = b$ with a sufficiently small same step-size $h = \frac{b-a}{l-1}$ such that $\mu = x_n \in (x_1, x_l)$.

Step 2. Select a proper kernel function.

Step 3. Choose proper kernel parameters and model parameters.

Step 4. Solve the problem (7) and obtain the optimal solution γ^*.

Step 5. Compute $\beta^* = -K^{-1}D\gamma^*$ and $b^* = p_0 - K_1^T\beta^*$.

Step 6. construct the numerical solution of Eq.(1) by $\hat{y}(x) = [k(x_1,x),\cdots,k(x_l,x)]\beta^* + b^*$.

4 Numerical examples

In order to demonstrate the effectiveness of Algorithm 1, in this section, a series of comparative experiments are performed between numerical solutions and exact solutions in the following six elaborated three-point BVPs of linear second-order ODEs with variable coefficients. The first five examples come from a the second kind Fredholm integral equation [10]

$$\begin{aligned}&\varphi(x) - \tfrac{1}{2}\int_0^1 (x+1)e^{-xy}\varphi(y)dy = e^{-x} - \tfrac{1}{2} + \tfrac{1}{2}e^{-(x+1)},\\ &0 \le x \le 1,\end{aligned}$$

and have the same exact solution $\varphi(x) = e^{-x}$. The sixth example is also taken from [10] and has the exact solution $\varphi(x) = e^x$.

Example 1. Consider a three-point BVP of linear second-order ODE with variable coefficients:

$$\begin{cases}\varphi''(x) + \varphi'(x) + x\varphi(x) = xe^{-x}, 0 \le x \le 1,\\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e^{-1} + e^{-1/2}.\end{cases}$$

Example 2. Consider a three-point BVP of linear second-order ODE with variable coefficients:

$$\begin{cases}\varphi''(x) - (1+\sin x)\varphi(x) = -e^{-x}\sin x, 0 \le x \le 1,\\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e^{-1} + e^{-1/2}.\end{cases}$$

Example 3. Consider a three-point BVP of linear second-order ODE with variable coefficients:

$$\begin{cases}\varphi''(x) + x^2\varphi'(x) - x\varphi(x) = (1 - x - x^2)e^{-x}, 0 \le x \le 1,\\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e^{-1} + e^{-1/2}.\end{cases}$$

Example 4. Consider a three-point BVP of linear second-order ODE with variable coefficients:

$$\begin{cases} \varphi''(x) + (1+x)\varphi'(x) = -xe^{-x}, 0 \le x \le 1, \\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e^{-1} + e^{-1/2}. \end{cases}$$

Example 5. Consider a three-point BVP of linear second-order ODE with variable coefficients:

$$\begin{cases} \varphi''(x) + \sin x\varphi'(x) + x^2\varphi(x) = (1 - \sin x + x^2)e^{-x}, \\ 0 \le x \le 1, \\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e^{-1} + e^{-1/2}. \end{cases}$$

Example 6. Consider a three-point BVP for a second-order ODE with variable coefficients:

$$\begin{cases} \varphi''(x) - (1 + \sin x)\varphi(x) = -e^x \sin x, x \in [0, 1], \\ \varphi(0) = 1, \varphi(1) + \varphi(1/2) = e + e^{1/2}. \end{cases}$$

All computations are implemented in Matlab 2010b on a PC with 2.5 GHz CPU and 4 G bytes memory. The Gaussian RBF kernel function

$$k(u, v) = \exp\{-\frac{(u-v)^2}{2\sigma^2}\}, \forall u, v \in R$$

is chosen in all experiments and the model parameter C and kernel parameter σ are taken as $C = 1$ and $\sigma = 2$. The experiment results are listed in Table 1 and the diagrams of comparisons are listed in Figure 1.

Because most of known methods for solving differential equations numerically are focus on Taylor expansion or discrete approaches, almost no use of optimization method, it can only compare the errors between numerical solutions and exact solutions in this paper. From Table 1, it can see that the proposed numerical method has good accuracy, which can be seen more obvious from Figure 1. According to the experiment results, we can conclude that the proposed method is feasible and effective for solving numerical three-point BVPs of linear second-order ODEs with variable coefficients.

5 Conclusions

This paper tries to seek the numerical solutions of three-point BVPs of linear second-order ODEs with variable coefficients by using optimization technology. Although kernel ε-SVR based on optimization model is a powerful methodology for solving function estimation problems, the main challenge in developing a useful regression model is to capture the underlying functional relationship between the given inputs and their output values accurately. In order to construct optimization model corresponding to kernel ε-SVR, this paper selects the discrete points of the domain of Eq.(1) as input values and selects values of the function $L(\hat{y}(x))$ at the discrete points as output values. By solving the Wolfe dual problem of the model, a numerical method is proposed. experiment results indicates that the proposed method has a good approximation property. From the derivation process, it can be seen that the proposed method has a certain versatility and can be used to study some other kinds of linear ODEs. The paper only involves linear ODEs and nonlinear ODEs problems will be our next work.

6 Acknowledgements

The research was supported by Science and Technology Project of Higher Education of Shandong Province, P.R. China (grant No.L13LI10).

References:

[1] F.V. Atkinson, Multiparameter Eigenvalue Problems, Academic Press, New York, 1972.

[2] A. Boucherif, Nonlinear three-point boundary value problems, J. Math. Anal. Appl. vol. 77, pp. 577-600, 1980.

[3] C.Z. Bai, J.X. Fang, Existence of multiple positive solutions for nonlinear m-point boundary-value problems, Appl. Math. Comput. vol. 140, pp. 297-305, 2003.

[4] X.J. Liu, J.Q. Liu, Three positive solutions for second-order m-point boundary value problems, Appl. Math. Comput. vol. 156, no. 3, pp. 733-742, 2004.

[5] X. Xu, Multiplicity results for positive solutions of some semi-positone three-point boundary value problems, J. Math. Anal. Appl. vol. 291, pp. 673-689, 2004.

[6] X. Xu, Three solutions for three-point boundary value problems, Nonlinear Anal. vol. 62, pp. 1053-1066, 2005.

[7] H. Chen, B. Li, Three-point boundary value problems for second-order ordinary differential equations in Banach spaces, Comput. Math. Appl. vol. 56, pp. 1852-1860, 2008.

[8] P.D. Phung, L.X. Truong, On the existence of a three point boundary value problem at resonance in Rn, J. Math. Anal. Appl. vol. 416, pp. 522-533, 2014.

Table 1: Comparison for six examples

(a) Comparison for Example 1.

x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	0.9048	0.9160	0.0112
0.2	0.8187	0.8350	0.0163
0.3	0.7408	0.7575	0.0167
0.4	0.6703	0.6842	0.0139
0.5	0.6065	0.6156	0.0091
0.6	0.5488	0.5521	0.0033
0.7	0.4966	0.4943	0.0022
0.8	0.4493	0.4426	0.0067
0.9	0.4066	0.3973	0.0092
1	0.3679	0.3588	0.0091

(b) Comparison for Example 2.

x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	0.9048	0.9136	0.0087
0.2	0.8187	0.8307	0.0120
0.3	0.7408	0.7522	0.0113
0.4	0.6703	0.6784	0.0081
0.5	0.6065	0.6100	0.0035
0.6	0.5488	0.5475	0.0014
0.7	0.4966	0.4913	0.0053
0.8	0.4493	0.4418	0.0075
0.9	0.4066	0.3994	0.0072
1	0.3679	0.3644	0.0035

(c) Comparison for Example 3.

x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	0.9048	0.9143	0.0094
0.2	0.8187	0.8320	0.0132
0.3	0.7408	0.7537	0.0129
0.4	0.6703	0.6801	0.0097
0.5	0.6065	0.6116	0.0051
0.6	0.5488	0.5488	0.0000
0.7	0.4966	0.4921	0.0044
0.8	0.4493	0.4420	0.0073
0.9	0.4066	0.3988	0.0078
1	0.3679	0.3628	0.0051

(d) Comparison for Example 4.

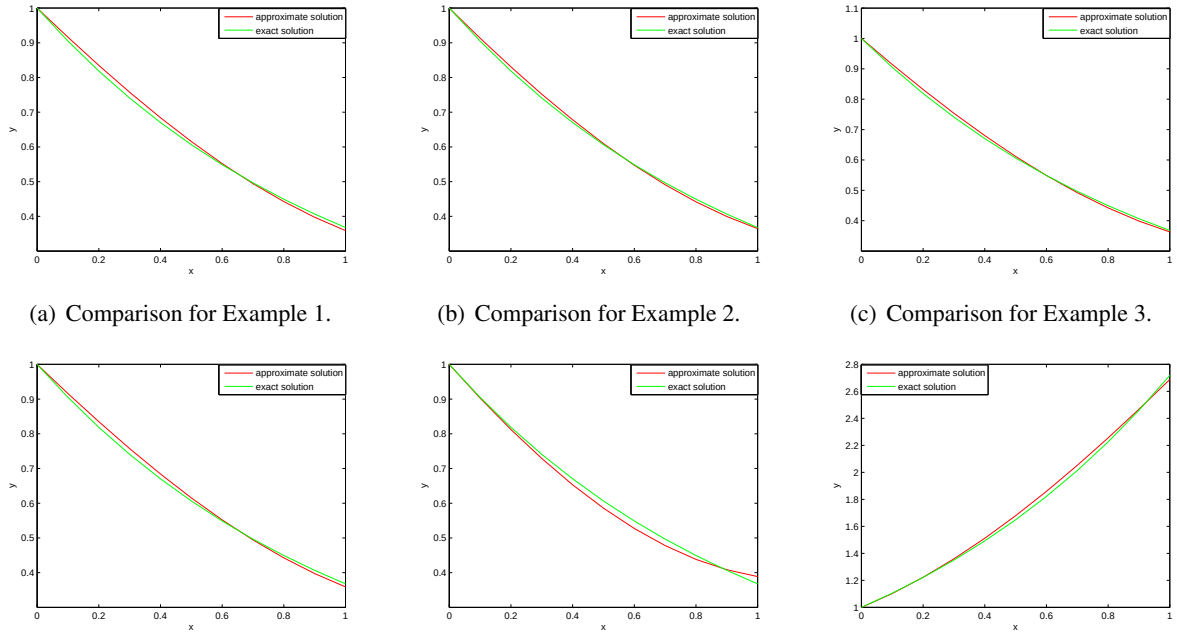
x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	0.9048	0.9161	0.0113
0.2	0.8187	0.8351	0.0164
0.3	0.7408	0.7577	0.0169
0.4	0.6703	0.6844	0.0141
0.5	0.6065	0.6158	0.0092
0.6	0.5488	0.5523	0.0035
0.7	0.4966	0.4944	0.0021
0.8	0.4493	0.4426	0.0067
0.9	0.4066	0.3973	0.0093
1	0.3679	0.3586	0.0092

(e) Comparison for Example 5.

x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	0.9048	0.9038	0.0011
0.2	0.8187	0.8138	0.0049
0.3	0.7408	0.7307	0.0101
0.4	0.6703	0.6552	0.0152
0.5	0.6065	0.5877	0.0188
0.6	0.5488	0.5288	0.0200
0.7	0.4966	0.4790	0.0176
0.8	0.4493	0.4385	0.0108
0.9	0.4066	0.4077	0.0011
1	0.3679	0.3867	0.0188

(f) Comparison for Example 6.

x	Exact solution	Numerical solution	Absolute error
0	1.0000	1.0000	0.0000
0.1	1.1052	1.1022	0.0030
0.2	1.2214	1.2221	0.0007
0.3	1.3499	1.3587	0.0089
0.4	1.4918	1.5112	0.0194
0.5	1.6487	1.6784	0.0297
0.6	1.8221	1.8591	0.0369
0.7	2.0138	2.0519	0.0381
0.8	2.2255	2.2554	0.0299
0.9	2.4596	2.4682	0.0086
1	2.7183	2.6886	0.0297



(a) Comparison for Example 1. (b) Comparison for Example 2. (c) Comparison for Example 3.

(d) Comparison for Example 4. (e) Comparison for Example 5. (f) Comparison for Example 6.

Figure 1: diagrams of comparisons for six examples

[9] Z. Zhao, Solutions and Greens functions for some linear second-order three-point boundary value problems, Comput. Math. Appl. vol. 56, pp. 104-113, 2008.

[10] Xian-Ci Zhong, Qiong-Ao Huang, Approximate solution of three-point boundary value problems for second-order ordinary differential equations with variable coefficients, Applied Mathematics and Computation, vol. 247, pp. 18-29, 2014.

[11] S. Mehrkanoon, T. Falck, J.A.K. Suykens, Approximate Solutions to Ordinary Differential Equations Using Least Squares Support Vector Machines, IEEE Transactions on Neural Networks and Learning Systems, vol. 23, no. 9, pp. 1356-1367, 2012.

[12] V. N.Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 2000.

[13] J. A. K. Suykens, J. Vandewalle, B. De Moor, Optimal control by least squares support vector machines, Neural Networks, vol. 14, no. 1, pp. 23-25, 2001.

[14] S. Chen, M. Wang, Seeking multi-thresholds directly from support vectors for image segmentation, Neurocomputing, vol. 67, pp. 335-344, 2005.

[15] Z. Liu, S. Xu, C. L. P. Chen, Y. Zhang, X. Chen, Y. Wang, A three-domain fuzzy support vector regression for image denoising and experimental studies. IEEE Transactions on Cybernetics, vol. 44, no. 4, pp. 516-525, 2014.

[16] K. W. Lau, Q. H. Wu, . Local prediction of non-linear time series using support vector regression, Pattern Recognition, vol. 41, no. 5, pp. 1539-1547, 2008.

[17] Y. Shang, Optimal control strategies for virus spreading in inhomogeneous epidemic dynamics. Canadian Mathematical Bulletin, vol. 56, no. 3, pp. 621-629, 2013.

[18] A. K. Sharma, P. S. Parihar, An effective DoS prevention system to analysis and prediction of network traffic using support vector machine learning. International Journal of Application or Innovation in Engineering and Management, vol. 2, no. 7, pp. 249-256, 2013.

[19] V. Vapnik, Statistical learning theory, New York: Wiley, 1998.

[20] M. L*á*zaro, I. Santamaria, F. P*é*rez-Cruz, A. Art*é*s-Rodriguez, Support vector regression for the simultaneous learning of a multivariate function and its derivative, Neurocomputing, vol. 69, pp. 42-61, 2005.