JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume **29**, No. 4, November 2016
http://dx.doi.org/10.14403/jcms.2016.29.4.631

# APPROXIMATE SOLUTIONS TO ONE-DIMENSIONAL BACKWARD HEAT CONDUCTION PROBLEM USING LEAST SQUARES SUPPORT VECTOR MACHINES

Ziku Wu*, Fule Li**, and Do Young Kwak***

Abstract. This article deals with one-dimension backward heat conduction problem (BHCP). A new approach based on least squares support vector machines (LS-SVM) is proposed for obtaining their approximate solutions. The approximate solution is presented in closed form by means of LS-SVM, whose parameters are adjusted to minimize an appropriate error function. The approximate solution consists of two parts. The first part is a known function that satisfies initial and boundary conditions. The other is a product of two terms. One term is known function which has zero boundary and initial conditions, another term is unknown which is related to kernel functions. This method has been successfully tested on practical examples and has yielded higher accuracy and stable solutions.

## 1. Introduction

Heat conduction problems usually fall into two classes: the direct heat conduction problems and the inverse heat conduction problems. Solutions of an inverse heat conduction problem entail determining unknown model parameters based on some observed values [1]. These model parameters include initial and boundary conditions, heat conduction coefficient, source function. Recently, the inverse heat conduction problems had attracted much attention from the researchers because

Received January 13, 2016; Accepted October 13, 2016.
2010 Mathematics Subject Classification: Primary 12A34, 56B34; Secondary 78C34.
Key words and phrases: backward heat conduction problem, least squares support vector machines, approximate solutions.
Correspondence should be addressed to Do Young Kwak, kdy@kaist.ac.kr.
This work is supported by international cooperation for excellent lectures by Shandong provincial education department and the NNSF of China (61403233).

they have powerful benefits in engineering and science [2]-[3]. Traditionally, analytical and numerical methods are the main methods which are employed to handle the inverse problems, such as adjoint assimilation method, integral transformation method, finite difference method, finite element method and exact methods. Due to their ill-posedness, almost all methods concern regularization technique. The backward heat conduction problems are known as determination of the initial conditions from the known distribution of the final time. There has been a lot of research works related to the inverse problem [4]-[6]. It is all known that the problems are ill posed. It is difficult to obtain stable solutions, so some regularization strategies should be considered. During the last two decades, some researches employed artificial neural networks (ANN) methods to solve the problems [7]-[10]. And to some extent the ANN method is very successful which improved the stability of the solutions. Although the ANN method has so many properties, it has two drawbacks. The first is the existence of many local minima solutions. The second is how to choose the number of hidden units. Support vector machines (SVM) proposed by Vapnik et al. 1995[11] has been successfully applied in many aspects for its high generalization ability and global optimization property. The simplicity of least square support vector machines (LS-SVM [12]-[13]) promotes the applications of SVM. Over the last decade, many pattern recognition and function approximation problems have successfully been tackled with LS-SVM method. Recently, Mehrkanoon et al. [14]-[16] proposed a new approach based on LS-SVM to solve ODEs and to estimate the unknown parameters in ODEs.

In this work, we employ a method based on LS-SVM to solve the one-dimensional backward heat conduction problems. In contrast to the traditional approach, the advantage of applying this method in solving the backward heat conduction problems is that no prior information is needed on the functional form of the unknown quantities. In addition, no initial guess is required and the solution can be computed by solving linear equations. This paper is organized as follows. The next section we introduce LS-SVM theory briefly. In Section 3, we formulate the LS-SVM method for the solution of the problems. In Section 4, we illustrate the method by means of examples and compare our results to analytic solutions. Finally, in Section 5 we give a conclusion of this work.

## 2. Brief introduction to the LS-SVM approach

Let us consider a given training set $\{V_i, Y_i\}_{i=1}^N$ with input data $V_i \in \mathbb{R}^k$ and output data $Y_i \in \mathbb{R}$. Our goal is to estimate a model of the following form using the regression:

$$Y(V) = \sum_{j=1}^{N} \alpha_j K(V, V_j) + b. \tag{2.1}$$

The LS-SVM model for regression can be written as the following quadratic programming problem:

$$\min_{a,b,e} \frac{1}{2}\alpha^T \alpha + \frac{\gamma}{2} e^T e \tag{2.2}$$

$$\text{such that } Y_i = \sum_{j=1}^{N} \alpha_j K(V_i, V_j) + b + e_i, \ \ i = 1, \cdots, N, \tag{2.3}$$

where $\gamma \in \mathbb{R}^+$ is a penalty factor, $e_i \in \mathbb{R}$ are bias terms, $K(p, q) = \exp(-\frac{\|p-q\|^2}{2\sigma^2})$ is the Gaussian RBF kernel, and $\alpha, b$ are regression parameters. With $\vec{Y} = (Y_1, Y_2, \cdots, Y_N)^T \in \mathbb{R}^N$ and $1_N = (1, 1, \cdots, 1)^T \in \mathbb{R}^N$, the solutions is given by

$$\begin{bmatrix} \Omega + \frac{I_N}{\gamma} & I_N \\ I_N^T & 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} \hat{Y} \\ 0 \end{bmatrix}, \tag{2.4}$$

where $\Omega_{i,j} = K(V_i, V_j)$ is the $i, j$-th entry of the positive definite kernel matrix and $I_N$ is the identity matrix of order $N$.

## 3. Formulation of the method for one dimensional heat conduction problem

In this paper, we consider the following equations

$$\begin{cases} \dfrac{\partial T}{\partial t} = k\dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,a) \times (0,b) \\ T(0,t) = g_1(t), & T(a,t) = g_2(t), \ t \in [0,b] \\ T(x,0) = f(x), & x \in [0,a] \end{cases} \tag{3.1}$$

$$\begin{cases} \dfrac{\partial T}{\partial t} = k\dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,a) \times (0,b) \\ T(0,t) = g_1(t), & T(a,t) = g_2(t), \ t \in [0,b] \\ T(x,0) = F(x), & x \in [0,a] \end{cases} \tag{3.2}$$

Equations (3.1) and (3.2) are the direct problem and the inverse problem, respectively. The aim of the inverse problem is to find $T(x, 0) = f(x)$.

In order to get approximate solutions of (3.1) and (3.2) , we consider a mesh of $M$ uniformly distributed interior points as training set of points. We reshape these points as a vector $V = [V_1, V_2, \cdots, V_M]^T, V_i = (x_i, t_i)$ is the i-th interior point. Assume the approximate solution of the equation has the following expression:

$$\hat{T}(V) = A(V) + B(V)\big(\sum_{j=1}^{M} \alpha_j K(V, V_j) + b\big), \tag{3.3}$$

where $V = (x, t),\ V_j = (x_j, t_j), K(U, V) = \exp(-\frac{\|U-V\|^2}{2\sigma^2})$ is the Gaussian RBF kernel function, and $\sigma$ is the kernel bandwidth. $\alpha$ and $b$ are the regression parameters that have to be determined. In this expression, $A(V) = A(x, t)$ is a known function, which satisfies the initial conditions and boundary conditions. On the other hand, the function $B(V) = B(x, t)$ takes zero on boundary and at the given moment. Neither functions includes any adjustable parameter.

For the direct problem (2.4), we have

$$\begin{aligned} A(V) &= (1 - \frac{x}{a})g_1(t) + \frac{x}{a}g_2(t) + (1 - \frac{t}{b}) \\ &\quad \cdot\big[f(x) - \big(1 - \frac{x}{a}\big)f(0) - \frac{x}{a}f(a)\big] &(3.4a) \\ B(V) &= x(a - x)t. &(3.4b) \end{aligned}$$

For the inverse problem (3.1), we have

$$\begin{aligned} A(V) &= (1 - \frac{x}{a})g_1(t) + \frac{x}{a}g_2(t) + \frac{t}{b} \\ &\quad \cdot[F(x) - (1 - \frac{x}{a})F(0) - \frac{x}{a}F(a)] &(3.5a) \\ B(V) &= x(a - x)(b - t). &(3.5b) \end{aligned}$$

It is easy to check that $\hat{T}(V)$ satisfies the initial and the boundary conditions. Inserting (3.3) into original equations (3.1) and (3.2), we obtain the following equation:

$$\sum_{j=1}^{M} \alpha_j G(V, V_j) + Q(V)b + W(V) = 0, \tag{3.6}$$

where

$$\begin{aligned} Q(V) &= B_t(V) - kB_{xx}(V),\ W(V) = A_t(V) - kA_{xx}(V) \\ G(V, V_j) &= G_1(V, V_j) + G_2(V, V_j) + G_3(V, V_j) + G_4(V, V_j) \end{aligned}$$

$$\begin{aligned} G_1(V,V_j) &= Q(V)K(V,V_j),\ G_2(V,V_j) = B(V)K_t(V,V_j) \\ G_3(V,V_j) &= -2kB_x(V)K_x(V,V_j),\ G_4(V,V_j) = -kB(V)K_{xx}(V,V_j). \end{aligned}$$

To obtain the optimal values of $\alpha$ and $b$, we solve the following quadratic optimization problem:

$$\min_{\alpha,e,b} \frac{1}{2}\alpha^T\alpha + \frac{\gamma}{2}e^Te \tag{3.7}$$

(3.8)

$$\text{such that } \sum_{j=1}^{M} \alpha_j G(V_i,V_j) + Q(V_i)b + W(V_i) + e_i = 0,\ i = 1,2,\cdots,M,$$

where $\gamma \in \mathbb{R}^+$ is regularization parameter, and $e_i$ is bias terms. The Lagrangian function of the constrained optimization problem (3.7) and (3.8) becomes

(3.9)

$$\begin{aligned} L(\alpha,e,b,\eta) = &\ \tfrac{1}{2}\alpha^T\alpha + \tfrac{\gamma}{2}e^Te \\ &-\sum_{i=1}^{M}\eta_i\Big(\sum_{j=1}^{M}\alpha_j G(V_i,V_j) + Q(V_i)b + W(V_i) + e_i\Big). \end{aligned}$$

Then the Karush-Kuhn-Tucher (KKT) optimality conditions as follows:

$$\begin{cases} \dfrac{\partial L}{\partial \alpha_i} &= \alpha_i - \displaystyle\sum_{j=1}^{M}\eta_j G(V_j,V_i) = 0 \\ \dfrac{\partial L}{\partial e_i} &= \gamma e_i - \eta_i = 0 \\ \dfrac{\partial L}{\partial \eta_i} &= \displaystyle\sum_{j=1}^{M}\alpha_j G(V_i,V_j) + Q(V_i)b + W(V_i) + e_i = 0 \\ \dfrac{\partial L}{\partial b} &= -\displaystyle\sum_{i=1}^{M}\eta_i Q(V_i) = 0 \end{cases} \tag{3.10}$$

After elimination of the primal variable $e_i$, the solution is given by

$$\begin{bmatrix} -I_M & KM^T & Z \\ KM & \frac{I_M}{\gamma} & LQ \\ Z^T & LQ^T & 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ \eta \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ rh \\ 0 \end{bmatrix}, \tag{3.11}$$

where $rh = -[W(V_1), W(V_2), \cdots, W(V_M)]^T$, $Z = [0,\cdots,0]^T$ and $LQ = [Q(V_1), Q(V_2), \cdots, Q(V_M)]^T$ are $M$ dimensional vectors, $I_M$ is a unit matrix of order $M$, $KM = [G(V_i,V_j)]$ is the kernel matrix of order $M$.

Equation (3.11) is linear, which can be solved easily. Finally, we can obtain the approximate function as (3.3). The performance of the LS-SVM model depends on the selection of the tuning parameters. In this paper, the Gaussian RBF kernel is employed. Therefore, the results are determined by two parameters, namely, the regularization $\gamma$ and the kernel bandwidth $\sigma$. In order to achieve approximate solutions accurate enough, we should take $\gamma$ quite large. In all examples, the chosen value for $\gamma$ was $10^8$. So, the only parameter that needs to be tuned is $\sigma$. In this paper, an approach is proposed which is based on Crank-Nicolson difference method. The parameter $\sigma$ is chosen according to the following rule:

$$\min_{\sigma} \|E_i^k\|_\infty \tag{3.12}$$

$$\begin{aligned} E_i^k &= \frac{T_i^{k+1} - T_i^k}{\tau} - \frac{1}{2h^2}[T_{i+1}^{k+1} - 2T_i^{k+1} + T_{i-1}^{k+1} - (T_i^k - 2T_i^k + T_{i-1}^k)] \\ T_i^k &= A(x_i, t_k) + B(x_x, t_k)(\sum_{j=1}^{M} \alpha_j \Phi(x_i, t_k, V_j) + b), \end{aligned}$$

where $h$ and $\tau$ are the step size of space and time, respectively. Theoretically, $\|E_i^k\|_\infty$ satisfies $\|E_i^k\|_\infty = O(\tau^2 + h^2)$. Therefore, the smaller $\|E_i^k\|_\infty$, the better $\sigma$ is. So it can be an optimal criterion of $\sigma$.

## 4. Examples

In this section, we test the performance of the method on two problems. In order to show the approximation capability of the method, we compare the computed approximate solutions with the analytic solutions. We divide the domain $\Omega$ by equal steps on $x$- axis $h$ and $t$ axis $\tau$. We illustrate the results on two cases.

Case I: $h = \pi/20, \tau = 0.1$ and $M = 171$. Case II: $h = \pi/40, \tau = 0.05$ and $M = 741$. Here $T^A$ and $T^L$ stand for the analytic solutions and the LS-SVM solutions, respectively.

PROBLEM 4.1. Consider the following equations. The direct problem is

$$\begin{cases} \dfrac{\partial T}{\partial t} = \dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,\pi) \times (0,1) \\ T(0,t) = T(\pi,t) = 0, & t \in [0,1] \\ T(x,0) = \sin x + \sin 2x, & x \in [0,\pi] \end{cases} \tag{4.1}$$

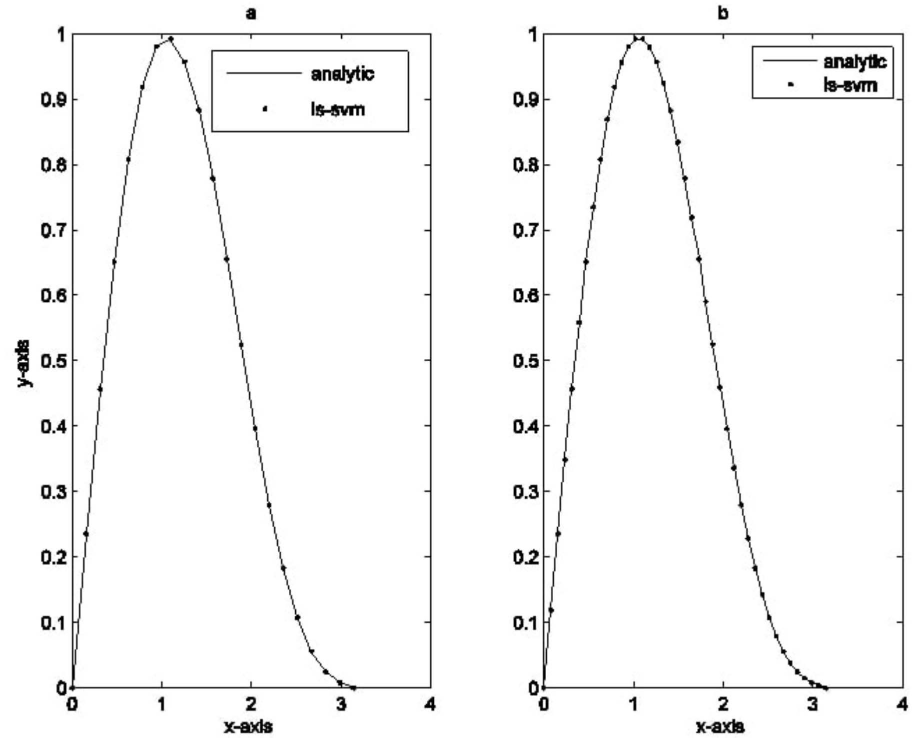


FIGURE 1. Numerical results for equation (3.12) at $t = 0$.
a: Case I, b: Case II.

And the inverse problem is

$$(4.2) \qquad \begin{cases} \dfrac{\partial T}{\partial t} = \dfrac{1}{4}\dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,\pi)\times(0,1) \\ T(0,t) = T(\pi,t) = 0, & t \in [0,b] \\ T(x,1) = e^{-\frac{1}{4}}\sin x + e^{-1}\sin 2x, & x \in [0,\pi] \end{cases}$$

The analytic solution is $T(x,t) = \exp(-\frac{t}{4})\sin x + \exp(-t)\sin 2x$. The approximate solutions obtained by our method are compared with the analytic solutions and results are depicted in Fig. 1-2. The related parameters information is tabulated in Table 1. It is clear that the approximate solutions are quite acceptable, despite the fact that fewer training points are employed.

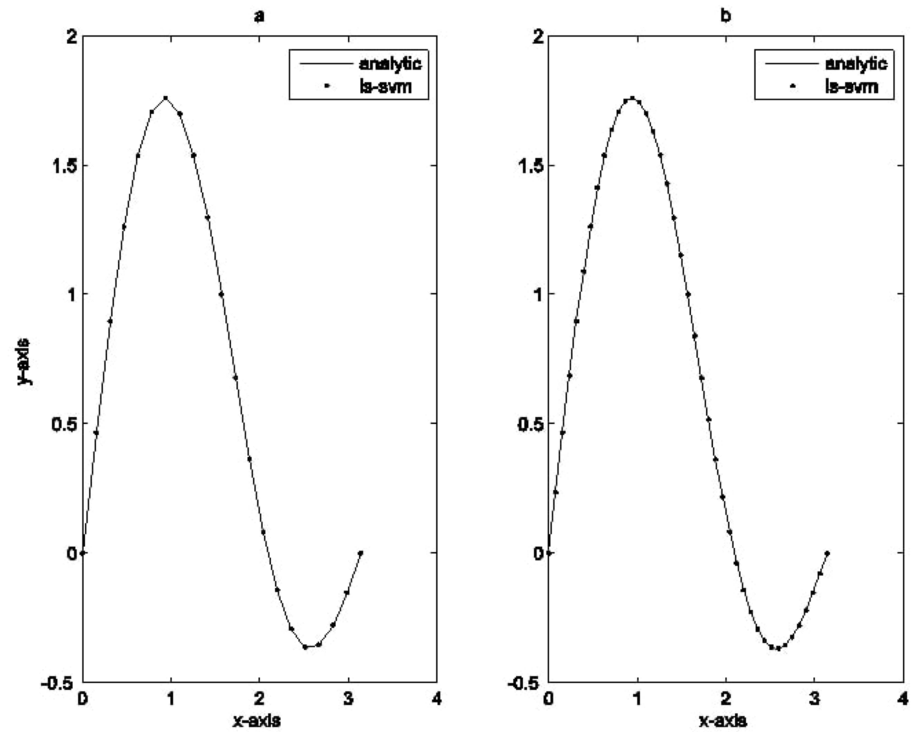


FIGURE 2. Numerical results for equation (3.12) at $t = 0$.
a: Case I, b: Case II.

| | Direct | | | Inverse | | |
|---|---|---|---|---|---|---|
| case | $\|T^A - T^L\|_\infty$ | aver err | $2\sigma^2$ | $\|T^A - T^L\|_\infty$ | aver err | $2\sigma^2$ |
| case I | 1.66 e-4 | 6.01e-5 | 1.215 | 8.79e-4 | 4.26e-4 | 1.475 |
| case II | 4.47e-5 | 1.37e-5 | 1.345 | 2.66e-4 | 8.91e-5 | 1.675 |

TABLE 1. Numerical results of the proposed method for solving example I

PROBLEM 4.2. Consider the following equations. The direct problem is

$$\begin{cases} \dfrac{\partial T}{\partial t} = \dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,\pi) \times (0,1) \\ T(0,t) = e^{-t}, T(\pi,t) = -e^{-t}, & t \in [0,1] \\ T(x,0) = \cos x, & x \in [0,\pi] \end{cases} \tag{4.3}$$

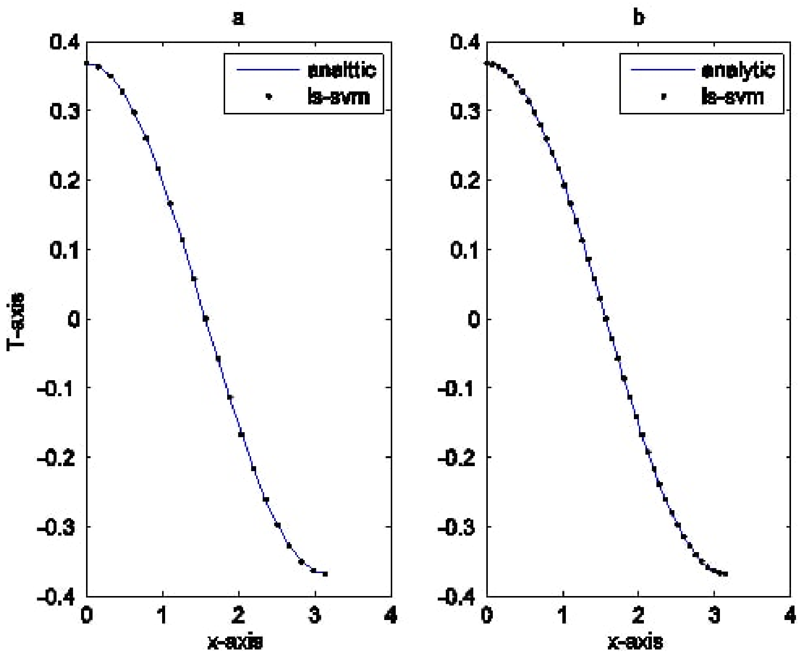


FIGURE 3. Numerical results for equation (4.2) at $t = 1$.
a: Case I, b: Case II.

And the inverse problem is

$$\begin{cases} \dfrac{\partial T}{\partial t} = \dfrac{\partial^2 T}{\partial x^2}, & (x,t) \in (0,\pi) \times (0,1) \\ T(0,t) = e^{-t},\ T(\pi,t) = -e^{-t}x, & t \in [0,1] \\ T(x,1) = e^{-1}\cos x, & x \in [0,\pi] \end{cases} \tag{4.4}$$

The analytic solution is $T(x,t) = \exp(-t)\cos x$. The approximate solutions obtained by the method are compared with the analytic solutions and results are depicted in Fig. 3-4. The related parameters information is tabulated in Table 2. It is clear that the approximate solutions of Case II are superior to Case I. It is due to more training points.

It can be seen that as the number of the training points increase, the kernel bandwidth $\sigma$ becomes larger and the error decrease. In all cases, we always have $\|T^A - T^L\|_\infty = O(\tau^2 + h^2)$, and the proposed method has higher accuracy and stability. In the future, this method may be employed to solve more complex inverse problems.

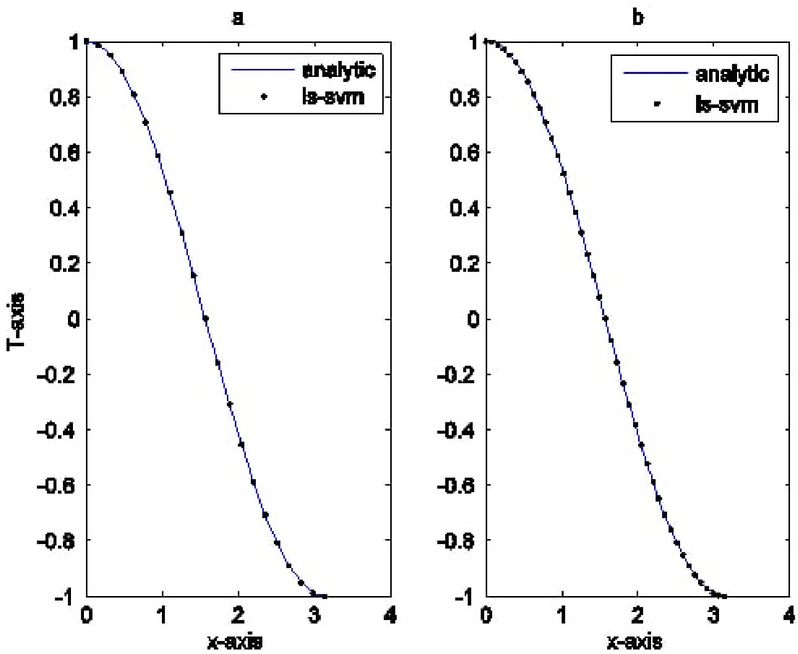


FIGURE 4. Numerical results for equation (4.3) at $t = 1$.
a: Case I, b: Case II.

| | Direct | | | Inverse | | |
|---|---|---|---|---|---|---|
| case | $\|T^A - T^L\|_\infty$ | aver err | $2\sigma^2$ | $\|T^A - T^L\|_\infty$ | aver err | $2\sigma^2$ |
| case I | 2.16e-5 | 5.53e-6 | 0.985 | 1.27e-4 | 2.84e-5 | 0.525 |
| case II | 7.80e-6 | 2.11e-6 | 1.385 | 2.93e-5 | 4.72e-6 | 1.300 |

TABLE 2. Numerical results of the proposed method for solving example II

## 5. Conclusion

In this paper, we introduced a method based on LS-SVM for one-dimension backward heat conduction problems. We have shown our method can solve the problem successfully. Although the method based on ANN can solve the problems with higher accuracy, it has some obvious drawbacks. Theoretically, PDEs can be solved by ANN can also be solved by LS-SVM. Because of complex boundary conditions and nonlinearity, the method based on LS-SVM may have some trouble to solve PDEs. That is why we focus on one-dimension heat conduction problem in this paper. Taking into account complexity, we assume approximate solutions directly which does not need dual form. This is different from previous LS-SVM based ones. On the tested problems, the method proposed in this paper is successful with accuracy and stability. Consequently, this method can be used for backward heat conduction problems with complex boundary conditions. We believe this method can be used for a wide class of inverse problems.

## References

[1] O. M. Alifanov, *Inverse Hear Transfer Problems. International series in heat and mass transfer*, Spingger-Verlag, 1994.

[2] Ghosh, D. K. Pratihar, B. Maiti, and P. K. Das, *Inverse estimation of location of internal heat source in conduction*, Inverse problem in science and engineering **19** (2011), no. 3, 337-361.

[3] X. T. Xiong, X. H. Liu, Y. M. Yan, and H. B. Guo, *A numerical method for identifying heat transfer coefficient*, Applied Mathematical Modelling **34** (2010), 1930-1938.

[4] C. W. Chang and C. S. Liu, *A backward group preserving scheme for multidimensional backward heat conduction problems*, CMES, Tech Science Press **59** (2010), no. 3, 239-274.

[5] D. Colton, *The approximation of solutions to the backwards heat equation in a nonhomogeneous medium.* J. Math. Anal. Appl. **72** (1979), 419-429.

[6] N. S. Mera, L. Elliott, D. B. Ingham, and D. Lesnic, *An iterative boundary element for solving one-dimensional backward heat conduction problem*, Int. J. Heat Mass Transfer. **44** (2001), 1937-1946.

[7] I. E. Lagaris, A. Likas, and D. I. Fotiadis, *Artificial neural networks for solving ordinary and partial differential equations*, IEEE Trans. Neural Networks **9** (1998), no. 5, 987-1000.

[8] I. E. Lagaris, A. Likas, and D. G. Papageorgio, *Neural-network methods for boundary value problems with irregular boundaries*, IEEE Trans. Neural Networks **11** no. 5 (2000), 1041-1049.

[9] M. Baymani, A. Kerayechian, and S. Effati, *Artificial neural networks approach for solving Stokes problem*, Applied Mathematics 2010, no. 1, 288-292.

[10] H. Alli, A. Ucar, and Y. Demir, *The solutions of vibration control problems using artificial neural networks*, Journal of the Franklin Institute **340** (2003), 307-325.

[11] V. N. Vapnik, *The nature of statistical learning theory*, New York, Springer-Verlag, 1995.

[12] J. A. K. Suykens and J. Vandewalle, *Least squares support vector machine classifiers*, Neural Process Letter **9** (1999), no. 3, 293-300.

[13] J. A. K. Suykens, T. V. Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least squares support vector machines*, Singapore, World Scientific, 2002.
[14] S. Mehrkanoon and J. A. K. Suykens, *LS-SVM approximate solution to linear time varying descriptor systems*, Automatica **48** (2012), no. 10, 2502-2511.
[15] S. Mehrkanoon, T. Falck, and J. A. K. Suykens, *Approximate solution to ordinary differential equations using least squares support vector machines*, IEEE Transactions on Neural Networks and Learning Systems **23** (2012), no. 9, 1356-1367.
[16] S. Mehrkanoon, S. Mehrkanoon, and J. A. K. Suykens, *Parameter estimation of delay differential equations: An integration-free LS-SVM approach*, Commun. Nonlinear Sci. Numer. Simulat. **19** (2014), 830-841.

*
Department of Mathematics
Qingdao Agricultural University
Qingdao, China
*E-mail*: zkwu1968@126.com

**
Department of Mathematics
Qingdao Agricultural University
Qingdao, China
*E-mail*: lfl2004666@126.com

***
Department of Mathematical Sciences
Korea Advanced Institute of Science and Technology
Daejeon 305-701, Republic of Korea
*E-mail*: kdy@kaist.ac.kr