Chebyshev Neural network model with linear and nonlinear active functions

Shahriar Seddighi Chaharborj[1*] **and Yaghoub Mahmoudi**[1]

[1]*Department of Mathematics, Tabriz Branch, Islamic Azad University, Tabriz, Iran*
[*]*Corresponding author E-mail:sseddighi2014@yahoo.com.my*

Abstract

In this paper the second order non-linear ordinary differential equations of Lane-Emden type as singular initial value problems using Chebyshev Neural Network (ChNN) with linear and nonlinear active functions has been studied. Active functions as, F(z)=z, sinh(x), tanh(z) are considered to find the numerical results with high accuracy. Numerical results from Chebyshev Neural Network shows that linear active function has more accuracy and is more convenient compare to other functions.

Keywords: Lane-Emden equation; Chebyshev Neural Network; Error back propagation algorithm; Feed forward neural network.

1. Introduction

The Lane-Emden type equation describes the variety of phenomena in physics and astrophysics, such as the aspects of stellar structure, the thermal history of a spherical cloud of gas and isothermal gas spheres [1]. Lane-Emden type equation is named after astrophysicists Jonathan Hamer Lane and Robert Emden which is a singular second order different equation as follows:

$$\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + f(x,y) = g(x) \qquad (1)$$

with initial condition : $y(0) = y_0, y'(0) = 0$.
The $f(x,y)$ in Eq. (1) is a nonlinear function of x and y and $g(x)$ is a function of x. Here we shown two popular type of $f(x,y)$ and $g(x)$.
Type I: Describes the thermal behavior of a spherical cloud of gas acting under the mutual attraction of its molecules and subject to the classical laws of the thermodynamics.

$$
\begin{aligned}
f(x,y) &= y^m, \ y(0) = 1, \ y'(0) = 0 \\
g(x) &= 0.
\end{aligned}
\qquad (2)
$$

where m is the polytrophic index which is a constant.
Type II: Describes the isothermal gas spheres where the temperature remains constant.

$$
\begin{aligned}
f(x,y) &= e^y, \ y(0) = 1, \ y'(0) = 0 \\
g(x) &= 0.
\end{aligned}
\qquad (3)
$$

It has been reported by several researcher such as Chandrasekharr [1] and Davis [2] that only for $m = 0, 1$ and 5 of type I Lane-Emden equation has exact solutions. For the type II Lane-Emden equation, an approximate implicit solution has been obtained by Momoniat and Harley [11] based on power series. Besides that, the exact solution of the Lane-Emden type equation also has been proposed by Khalique and Ntsime [6] with respect to the standard Lagrangian

according to the Noether point symmetries approach.
Recently, an analytical and numerical solution to the Lane-Emden equation presented by Gorder and Vajravelu [14] by using the traditional power series approach and the homotopy anlaysis method. The collocation method has been used by Mechee and Senu [10] to obtain the numerical solution of the Lane-Emden type. Recently, Chebyshev Neural Network based model has been used by Mall and Chakraverty [7] to solve the Lane-Emden type equations where the artificial neutral network used to solve the singularity in Lane-Emden type equations. A numerical method based on hybrid of Chebyshev wavelet and finite difference method has been proposed by Nasab *et al.* [12] to solve the Lane-Emden equation which is reduce the computation of the problem to a set of nonlinear algebraic equations. In this paper, we are concerned for solving Lane-Emden type equations using Chebyshev neural network method with different active functions include a single layer and eliminated hidden layer by expanding the input pattern by Chebyshev polynomials [8, 9]. To minimize the computed error function a feed forward neural network model with error back propagation principle is used.

2. Learning algorithm of Chebyshev neural network

Fig.(1) shows the structure of Chebyshev Neural Network (ChNN) which consists of single input unit, one output unit and a functional expansion block based on Chebyshev polynomials. Chebyshev Neural Network model is a single layer neural model where each input data is expanded to several terms using Chebyshev polynomials. For updating the network parameters and minimizing the error function can use the learning algorithm. To update weights of the ChNN; error back propagation algorithm is used [8, 9]. Functions F(z)=z; sinh(z); tanh(z) are considered as the activation functions. The network output with input data x and weights p

may be computed as,

$$N(x, p) = F(z), \qquad (4)$$

where z a weighted sum of expanded input data is as,

$$z = \sum_{j=1}^{M} w_j T_{j-1}(x), \qquad (5)$$

where x is the input data, $T_{j-1}(x)$ and w_j with $j = 1, 2, ..., M$ denote the expanded input data and the weight vector, respectively. The first two Chebyshev polynomials are as,

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x. \end{aligned}$$

The higher order Chebyshev polynomials can be evaluate by,

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad (6)$$

where $T_n(x)$ denotes nth order Chebyshev polynomial. Here n dimensional input pattern is expanded to m dimensional enhanced Chebyshev polynomials.

Now, the weights of ChNN may be modified by using principle of back propagation [8, 9],

$$w_j^{k+1} = w_j^k + \Delta w_j^k = w_j^k + \left(-\eta \left[\frac{\partial E(x, p)}{\partial w_j} \right]^k \right), \qquad (7)$$

where η is learning parameter, k is iteration step which is used to update the weights as usual in Artificial Neural Network (ANN) and $E(x, p)$ is the error function.
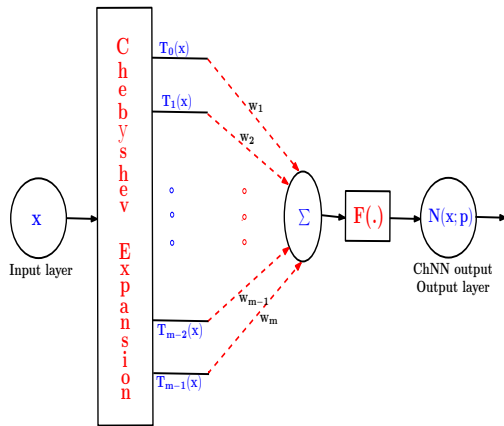

Figure 1: Structure of single layer Chebyshev Neural Network.

3. Formulation of the method

In this section, general formulation of differential equations using Neural Network have described. In particular the formulations of ordinary differential equations (ODEs) are incorporated in detail with computation of the gradient of the network parameters with respect to its inputs [8, 9].

3.1. Chebyshev neural network formulation for differential equations

A general differential equation which represents ordinary or partial differential equations is as,

$$\Psi\left[x, y(x), \nabla y(x), \nabla^2 y(x), ..., \nabla^n y(x)\right] = 0, \qquad x \in \bar{D} \subseteq R^n, \qquad (8)$$

where Ψ is the function which defines the structure of differential equation, $y(x)$ and ∇ denote the solution and differential operator, respectively. Let $y_t(x, p)$ denotes the trial solution with adjustable parameters p and then the above general differential equation changes to the form,

$$\Psi\left[x, y_t(x, p), \nabla y_t(x, p), \nabla^2 y_t(x, p), ..., \nabla^n y_t(x, p)\right] = 0, \qquad (9)$$

The problem is transformed into the following minimization problem [8, 9, 5],

$$\min_p \frac{1}{2}\left[\sum_{x \in \bar{D}}\left[\Psi\left(x, y_t(x, p), \nabla y_t(x, p), \nabla^2 y_t(x, p), ..., \nabla^n y_t(x, p)\right)\right]^2\right]. \qquad (10)$$

Now, we consider the second order ordinary differential equation (ODE) as follows,

$$\frac{d^2y(x)}{dx^2} = f(x, y, y'); \quad x \in [a, b] \qquad (11)$$

with the initial conditions, $y(a) = A$ and $y'(a) = A'$, the trial solution $y_t(x, p)$ of feed forward neural network with input x and parameters p is written as,

$$y_t(x, p) = A + A'(x-a) + (x-a)^2 N(x, p), \qquad (12)$$

where, $\texttt{N(x, p)=z; sinh(z); tanh(z)}$.

General form of corresponding error function for the ODE's may be formulated as [8, 9],

$$E(x, p) = \sum_{i=1}^{n} \frac{1}{2}\left[\frac{d^2 y_t(x_i, p)}{dx^2} - f(x_i, y_t(x, p), y_t'(x, p))\right]^2. \qquad (13)$$

To minimize the error function $E(x, p)$ corresponding to every entry x, we differentiate $E(x, p)$ with respect to the parameters. Then the gradient of network output with respect to their inputs is computed as below.

3.2. Computation of gradient for optimizing values of weight

The error computation involves both output and derivatives of the network output with respect to the corresponding inputs [8, 9]. Then the gradient of network output with respect to their inputs for $\texttt{N(x, p)=z; sinh(z); tanh(z)}$ is computed as below.

- **If $N(\mathbf{x}, \mathbf{p}) = \mathbf{z}$:**
  Derivatives of $N(x, p) = z$ with respect to input x is as follows,

$$\frac{dN}{dx} = \sum_{j=1}^{M} w_j T_{j-1}'(x), \qquad (14)$$

  and

$$\frac{d^2N}{dx^2} = \sum_{j=1}^{M} w_j T_{j-1}''(x). \qquad (15)$$

- **If $N(\mathbf{x}, \mathbf{p}) = \sinh(\mathbf{z})$:**
  Derivatives of $N(x, p) = sinh(z)$ with respect to input x is as follows,

$$\frac{dN}{dx} = \frac{1}{2} \sum_{j=1}^{M} w_j T_{j-1}'(x)\left(e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} + e^{-\sum_{j=1}^{M} w_j T_{j-1}(x)}\right), \qquad (16)$$

  and

$$\frac{d^2N}{dx^2} = \frac{1}{2}\left[\begin{array}{l} \sum_{j=1}^{M} w_j T_{j-1}''(x) e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} \\ + \left(\sum_{j=1}^{M} w_j T_{j-1}'(x)\right)^2 e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} \\ + \frac{\sum_{j=1}^{M} w_j T_{j-1}''(x)}{e^{\sum_{j=1}^{M} w_j T_{j-1}(x)}} - \frac{\left(\sum_{j=1}^{M} w_j T_{j-1}'(x)\right)^2}{e^{\sum_{j=1}^{M} w_j T_{j-1}(x)}} \end{array}\right]. \qquad (17)$$

- **If $N(\mathbf{x}, \mathbf{p}) = \tanh(\mathbf{z})$:**
  Derivatives of $N(x,p) = tanh(z)$ with respect to input x is as follows,

$$\frac{dN}{dx} = \frac{4\sum_{j=1}^{M} w_j T'_{j-1}(x)}{\left(e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} + e^{-\sum_{j=1}^{M} w_j T_{j-1}(x)}\right)^2}, \qquad (18)$$

and

$$\frac{d^2 N}{dx^2} = \frac{4\begin{bmatrix} \sum_{j=1}^{M} w_j T''_{j-1}(x) e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} \\ +\sum_{j=1}^{M} w_j T''_{j-1}(x) e^{-\sum_{j=1}^{M} w_j T_{j-1}(x)}, \\ -2\left(\sum_{j=1}^{M} w_j T'_{j-1}(x)\right)^2 e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} \\ +2\left(\sum_{j=1}^{M} w_j T'_{j-1}(x)\right)^2 e^{-\sum_{j=1}^{M} w_j T_{j-1}(x)} \end{bmatrix}}{\left(e^{\sum_{j=1}^{M} w_j T_{j-1}(x)} + e^{-\sum_{j=1}^{M} w_j T_{j-1}(x)}\right)^3}. \quad (19)$$

where w_j denote parameters of network and $T'_{j-1}(x)$, $T''_{j-1}(x)$ denote first and second derivatives of Chebyshev polynomials. From Eq. (12) we have,

$$\frac{dy_t(x,p)}{dx} = A' + 2(x-a)N(x,p) + (x-a)^2\frac{dN}{dx}, \qquad (20)$$

and

$$\frac{d^2 y_t(x,p)}{dx^2} = 2N(x,p) + 4(x-a)\frac{dN}{dx} + (x-a)^2\frac{d^2 N}{dx^2}. \qquad (21)$$

Also, from Eq. (13) we have,

$$\frac{\partial E(x,p)}{\partial w_j} = \frac{\partial}{\partial w_j}\left[\sum_{i=1}^{n}\frac{1}{2}\left(\frac{d^2 y_t(x_i,p)}{dx^2} - f(x_i, y_t(x,p), y'_t(x,p))\right)\right]. \qquad (22)$$

Finally, we can use the converged ChNN results in Eq. (12) to obtain the approximate solutions.

4. Numerical results

In this section, we consider Lane-Emden equation to show the powerfulness of the proposed method [8]. Active functions updated as, `F(z)=z; sinh(x); tanh(z)` are considered to find the numerical results with high accuracy. In this paper also we used first five Chebyshev polynomials (`m=5`).

4.1. Example 1

The For $n = 0$, the equation becomes linear ordinary differential equation,

$$\frac{d^2 y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + 1 = 0, \qquad (23)$$

with initial conditions $y(0) = 1$, $y'(0) = 0$. The exact solution of the above equation is given as,

$$y(x) = -\frac{1}{6}x^2. \qquad (24)$$

As discussed above we can write the ChNN trail solution as,

$$y_t(x,p) = 1 + x^2 N(x,p). \qquad (25)$$

The network is trained for ten equidistant points in [0, 1] with first five Chebyshev polynomials (`n=5`). Fig. (2.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (2.b). As we can see error between analytical and ChNN solutions with active function `F(z)=z` is more better than active functions `tanh(z); sinh(z)`. Numerical solution with active function `F(z)=z` has good agreement with analytical solution.
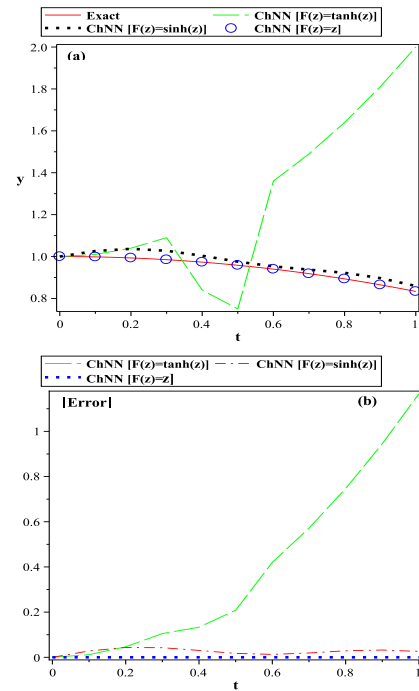

Figure 2: Numerical results and absolute errors for Example 1, (a): numerical results, (b): absolute errors.

4.2. Example 2

Let us consider LaneEmden equation for $n = 1$,

$$\frac{d^2 y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + y = 0, \qquad (26)$$

with initial conditions $y(0) = 1$, $y'(0) = 0$. The exact solution of the above equation is given as,

$$y(x) = \frac{sin(x)}{x}. \qquad (27)$$

The ChNN trial solution in this case is represented as,

$$y_t(x,p) = 1 + x^2 N(x,p). \qquad (28)$$

The network have been trained for ten equidistant points in interval [0, 1] for computing the results. Fig. (3.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (3.b). Obtained errors show good agreement between analytical and ChNN solutions for all three active functions `F(z)=z; sinh(x); tanh(z)`.

4.3. Example 3

Let us consider LaneEmden equation for $n = 5$,

$$\frac{d^2 y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + y^5 = 0, \qquad (29)$$

with initial conditions $y(0) = 1$, $y'(0) = 0$.
The exact solution of the above equation is given in [3, 4] as,

$$y(x) = (1 + \frac{x^2}{3})^{(-1/2)} \qquad (30)$$

The ChNN trial solution may be expressed as,

$$y_t(x,p) = 1 + x^2 N(x,p). \qquad (31)$$

Figure 3: Numerical results and absolute errors for Example 2, (a): numerical results, (b): absolute errors.
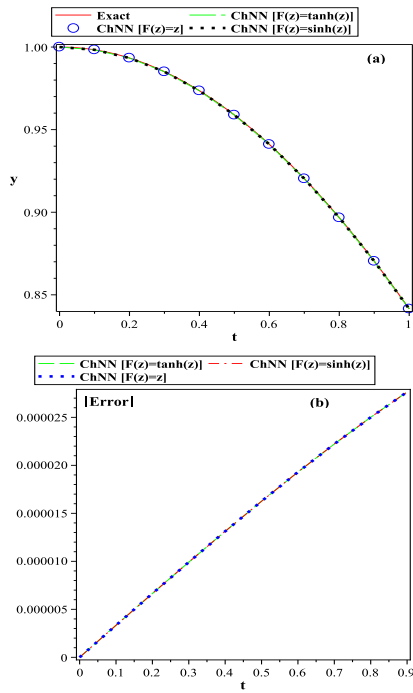

Figure 4: Numerical results and absolute errors for Example 3, (a): numerical results, (b): absolute errors.

The network have been trained for ten equidistant points in interval $[0, 1]$ for computing the results. Fig. (4.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (4.b). Obtained errors show good agreement between analytical and ChNN solutions for all three active functions `F(z)=z; sinh(x); tanh(z)`.

4.4. Example 4

Let us consider LaneEmden equation for $n = 0.5$,

$$\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + y^{0.5} = 0, \tag{32}$$

with initial conditions $y(0) = 1$, $y'(0) = 0$.
The ChNN trial solution is written as,

$$y_t(x, p) = 1 + x^2 N(x, p). \tag{33}$$

Ten equidistant points in interval $[0, 1]$ are considered here to train the model. Fig. (5.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (5.b). Obtained errors show good agreement between analytical and ChNN solutions for all three active functions `F(z)=z; sinh(x); tanh(z)`.

4.5. Example 5

Here we take Lane-Emden equation for $n = 2.5$,

$$\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + y^{2.5} = 0, \tag{34}$$

with initial conditions $y(0) = 1$, $y'(0) = 0$.
ChNN trail solution may be written as,

$$y_t(x, p) = 1 + x^2 N(x, p). \tag{35}$$

The network have been trained for ten equidistant points in interval $[0, 1]$ for computing the results. Fig. (6.a) shows comparison
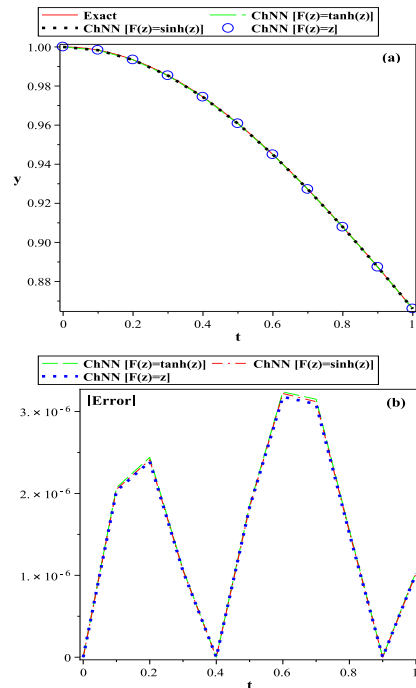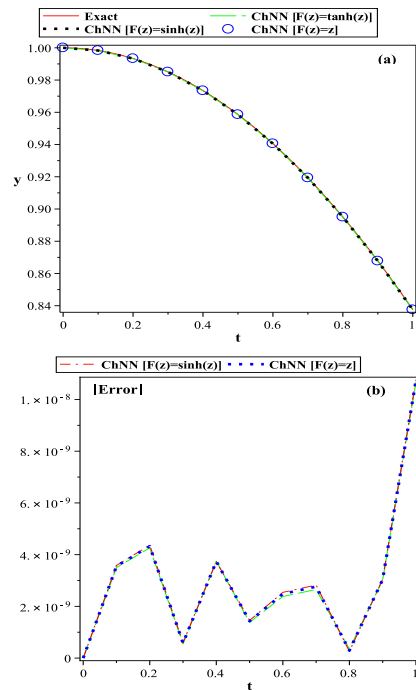

Figure 5: Numerical results and absolute errors for Example 4, (a): numerical results, (b): absolute errors.

between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (6.b). Obtained errors show good agreement between analytical and ChNN solutions for all three active functions `F(z)=z; sinh(x); tanh(z).`
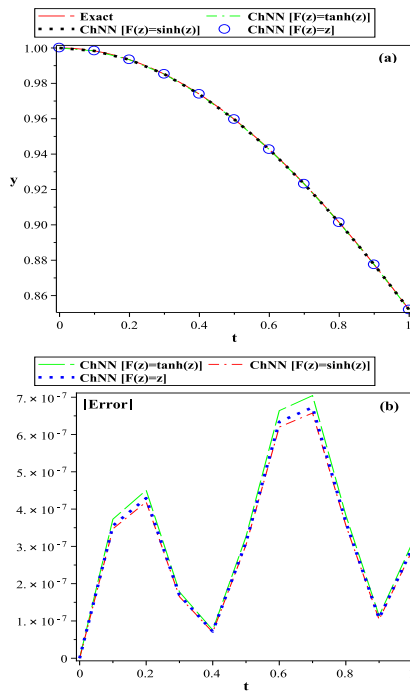

Figure 6: Numerical results and absolute errors for Example 5, (a): numerical results, (b): absolute errors.

4.6. Example 6

Below we now consider an example of Lane-Emden equation with $f(x,y) = -2(2x^2+3)y$. As such second order homogeneous Lane-Emden equation will be,

$$\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} - 2(2x^2+3)y = 0, \tag{36}$$

with initial conditions $y(0) = 1$, $y'(0) = 0$. The exact solution of the above equation is given as,

$$y(x) = e^{x^2}. \tag{37}$$

As discussed above we can write the ChNN trail solution as,

$$y_t(x,p) = 1 + x^2 N(x,p). \tag{38}$$

The network have been trained for ten equidistant points in interval $[0,1]$ for computing the results. Fig. (7.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (7.b). Obtained errors show good agreement between analytical and ChNN solutions for two active functions `F(z)=z; sinh(x).` As we can see ChNN results with active function `F(z)=tanh(z)` is not in good agreement with analytical solution.

5. The non homogeneous Lane-Emden equation

Following non homogeneous Lane-Emden equations have been solved by Wazwaz [15] and Singh et al. [13] using Adomian decomposition and modified homotopy analysis method. Here the same problem is solved using Chebyshev Neural Network.
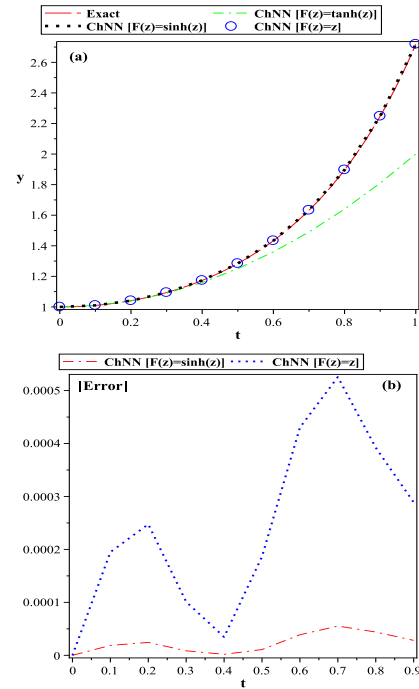

Figure 7: Numerical results and absolute errors for Example 6, (a): numerical results, (b): absolute errors.

5.1. Example 7

The non homogeneous Lane-Emden equation is written as,

$$\frac{d^2y}{dx^2} + \frac{2}{x}\frac{dy}{dx} + y = 6 + 12x + 12x^2 + x^3, \tag{39}$$

with initial conditions $y(0) = 0$, $y'(0) = 0$. This equation has the exact solution for $x \geq 0$ [13] as,

$$y(x) = x^2 + x^3. \tag{40}$$

Here, we can write the related ChNN trial solution as,

$$y_t(x,p) = x^2 N(x,p). \tag{41}$$

The network have been trained for ten equidistant points in interval $[0,1]$ for computing the results. Fig. (8.a) shows comparison between analytical and Chebyshev neural results. The error plot between analytical and ChNN results with different active functions are shown in Fig. (8.b). Obtained errors show good agreement between analytical and ChNN solutions with two active functions `F(z)=z; sinh(x).` As we can see form Fig. (8.a) ChNN results with active function `F(z)=tanh(z)` is not in good agreement with analytical solution.

6. Conclusion

The second order non-linear ordinary differential equations of Lane-Emden type as singular initial value problems using Chebyshev Neural Network (ChNN) has been studied with different kind active functions as, `F(z)=z; sinh(x); tanh(z).` Numerical results from ChNN method have been compared with analytical solutions. Comparison of obtained results with analytical results shows that results from proposed method with active function `F(z)=z` has very good agreement with analytical results. Can conclude the linear active function can applied for solvating any type of linear and non-linear ordinary differential equations.
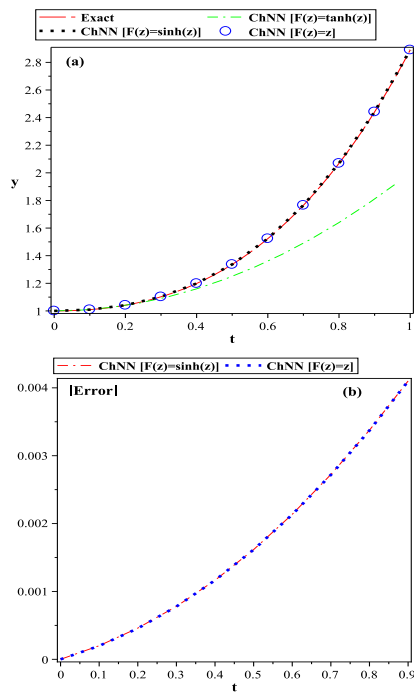
Figure 8: Numerical results and absolute errors for Example 7, (a): numerical results, (b): absolute errors.

References

[1] S. Chandrasekhar and S. Chandrasekhar. *An introduction to the study of stellar structure*, volume 2. Courier Dover Publications, 2010.

[2] H. T. Davis. *Introduction to nonlinear differential and integral equations*. Courier Dover Publications, 1962.

[3] H. T. Davis. *Introduction to nonlinear differential and integral equations*. Courier Dover Publications, 1962.

[4] L. Dresner. *Similarity solutions of nonlinear partial differential equations*, volume 88. Pitman London, 1983.

[5] S. Hoda and H. Nagla. Neural network methods for mixed boundary value problems. *International Journal of Nonlinear Science*, 11(3):312–316, 2011.

[6] C. Khalique and P. Ntsime. Exact solutions of the lane–emden-type equation. *New Astronomy*, 13(7):476–480, 2008.

[7] S. Mall and S. Chakraverty. Chebyshev neural network based model for solving lane–emden type equations. *Applied Mathematics and Computation*, 247:100–114, 2014.

[8] S. Mall and S. Chakraverty. Chebyshev neural network based model for solving lane–emden type equations. *Applied Mathematics and Computation*, 247:100–114, 2014.

[9] S. Mall and S. Chakraverty. Numerical solution of nonlinear singular initial value problems of emden–fowler type using chebyshev neural network method. *Neurocomputing*, 149:975–982, 2015.

[10] M. S. Mechee and N. Senu. Numerical study of fractional differential equations of lane-emden type by method of collocation. *Applied Mathematics*, 3:851, 2012.

[11] E. Momoniat and C. Harley. Approximate implicit solution of a lane-emden equation. *New Astronomy*, 11(7):520–526, 2006.

[12] A. K. Nasab, A. Kılıçman, Z. P. Atabakan, and W. Leong. A numerical approach for solving singular nonlinear lane–emden type equations arising in astrophysics. *New Astronomy*, 34:178–186, 2015.

[13] O. P. Singh, R. K. Pandey, and V. K. Singh. An analytic algorithm of lane–emden type equations arising in astrophysics using modified homotopy analysis method. *Computer Physics Communications*, 180(7):1116–1124, 2009.

[14] R. A. Van Gorder and K. Vajravelu. Analytic and numerical solutions to the lane–emden equation. *Physics Letters A*, 372(39):6060–6065, 2008.

[15] A.-M. Wazwaz. A new algorithm for solving differential equations of lane–emden type. *Applied Mathematics and Computation*, 118(2):287–310, 2001.