AN ENERGY FUNCTION APPROACH FOR FINDING ROOTS OF CHARACTERISTIC EQUATION

Deepak Mishra¹ and Prem K. Kalra²

¹Department of Avionics, Indian Institute of Space Science and Technology, Kerala, India
E-mail: dkmishra@gmail.com
²Department of Electrical Engineering, Indian Institute of Technology Rajasthan, India
E-mail: kalra@iitk.ac.in

Abstract
In this paper, an energy function approach for finding roots of a characteristic equation has been proposed. Finding the roots of a characteristic equation is considered as an optimization problem. We demonstrated that this problem can be solved with the application of feedback type neural network. The proposed approach is fast and robust against variation of parameter.

Keywords:
Hopfield Neural Network, Energy Function, Eigenvalues, Neural Networks

1. INTRODUCTION

Hopfield proposed a feedback neural network popularly known as Hopfield neural network [9], [10]. In his work, Hopfield designed a recurrent neural network as dynamical associative memory and showed that a single-layer of a fully connected network is capable of restoring a previously learned static pattern called a memory vector, ensuring its convergence from any initial condition [9]. The Hopfield network may be operated in a continuous mode or discrete mode, depending on the model adopted for describing the neurons. The continuous model of operation is based on an additive model [10]. On the other hand, the discrete mode of operation is based on McCulloch-Pitts model [9].

Hopfield network has been used for a variety of applications, such as associative memory [9], complex-valued associative memory [6], classification [31], and object recognition [27]. It is also used to solve optimization problems like Traveling salesman problem [11], finding solutions of linear and nonlinear equations [4], [23], Analog-to-Digital (ADC) converter [5], [29], Job Scheduling [30], Assignment problem [3] etc. However, the application of conventional Hopfield neural network in optimization was limited to quadratic cost functions only. To overcome this limitation, Samad and Harper [28] extended this network for higher order cost functions and proposed a higher order Hopfield neural network. Later on Miguel et al. applied it to several well known problems of optimization [2], [21].

The techniques of finding roots of a polynomial, found applications in diverse fields such as control system [25], digital signal processing [32], and image processing [1]. These problems can easily be solved by many traditional methods such as Laguerre method, Muller method, Hurwitz method, Newton-Raphson method [32], and Routh criterion [7]. Since last two decades, Neural network technique has been used successfully in many fields, e.g., in factorization of polynomials with two or more variables [12], [26], inversion of the non-singular matrices [19], solving linear/non-linear algebraic equations [4], [22], [24], etc. Similar to above mentioned applications application of neural networks in finding the roots of polynomial is a novel and important research topic [20]. Huang et al. contributed significant applications of feed-forward neural network for determining the roots of arbitrary order polynomials. The approach using backpropagation network to zero polynomial is to factorize the polynomial into subfactors on the hidden layer of the network, and use a suitable adaptive learning algorithm to train the connection weights, i.e. roots, from the input layer to the hidden layer until the defined output error between actual output and the desired output (the polynomial) converges to a predefined accuracy [16]. All these learning schemes are supervised learning schemes. In these schemes initialization of neural network parameters is one of the key issues. In their work, Huang et al. depicts that neural network based root finder schemes compete equally with conventional root finder schemes. However, the application of Hopfield neural network for extracting the roots of the characteristic equation is not explored even today. In this paper, an energy function approach using slightly modified Hopfield neural network is attempted to compute the roots of a given characteristic equation. Our study concludes that a fixed weight neural network such as Hopfield neural network with minor modifications can be used for root finding.

The paper is organized as follows. Section 2 presents the formulation of modified Hopfield neural network for finding the roots of a given characteristic polynomial. Stability analysis of the proposed approach is carried out in section 3. Section 4, demonstrates few numerical experiments with proposed approach. We concluded our work with a brief discussion in section 5.

2. MODIFIED HOPFIELD NEURAL NETWORK FOR DETERMINING NON-REPEATED ROOTS OF A CHARACTERISTIC EQUATION

Consider the characteristic equation,

\[ q(s) = a_0s^n + a_1s^{(n-1)} + \ldots + a_{(n-1)}s + a_n \]  \hspace{1cm} (1)

\[ a_0 > 0 \]

Eq.(1), can be written in the following form,

\[ q(s) = (s - s_1)(s - s_2)(s - s_3)\ldots(s - s_n) \]  \hspace{1cm} (2)

where, \( s_1, s_2, \ldots, s_n \) are the roots of Eq.(1). These roots can be real or in complex conjugate pair. Here, we assumed that all the
roots can be written in the form of \((x + iy)\). Variable \(s\) in Eq.(1) can be replaced by \((x + iy)\). That results in following,

\[
g(q) = a_0(x + iy)^n + a_1(x + iy)^{n-1} + \ldots + a_{n-1}(x + iy) + a_n > 0
\]  
(3)

Expanding the Eq.(3) gives us a complex algebraic equation of form \((\sigma + i\omega)\). That can be represented as follows,

\[
q(s) = \Re\{a_0(x + iy)^n + \ldots + a_{n-1}(x + iy) + a_n\} + i\Im\{a_0(x + iy)^n + \ldots + a_{n-1}(x + iy) + a_n\}
\]

or

\[
q(s) = \sigma + i\omega
\]  
(4)

Here, \(\Re(.)\) gives the real part of the characteristic equation \(q(s)\) after expansion, while \(\Im(.)\) gives the imaginary part. The individual terms of Eq.(4), i.e. \(\sigma\) and \(\omega\) are used for energy function formulation for the proposed model. This energy function will be used further to determine all possible non-repeated roots of the given characteristic equation. In our work, we consider this problem as an optimization problem and, in order to use a neural network model to solve an optimization problem, the problem is cast into the form of an energy function that the model minimizes. Hence, the energy function using Eq.(4) can be shown as follows,

\[
E(x, y) = \left[\Re\{a_0(x + iy)^n + \ldots + a_{n-1}(x + iy) + a_n\}\right]^2
\]

\[
+ \left[\Im\{a_0(x + iy)^n + \ldots + a_{n-1}(x + iy) + a_n\}\right]^2
\]

or

\[
E(x, y) = (\sigma^2 + \omega^2)
\]  
(5)

The energy function given in Eq.(6) is used to design a modified Hopfield neural network for finding the non-repeated roots of the characteristic equation Eq.(1). Calculation of partial derivatives of Eq.(6) with respect to variables \(x\) and \(y\), gives the dynamics of network, which can be written as,

\[
\frac{du}{dt} = -\frac{\partial}{\partial x} E(x, y)
\]

\[
\frac{dv}{dt} = -\frac{\partial}{\partial y} E(x, y)
\]

(7)

Equations (6) and (7) are used to determine weights and biases of the proposed neural network.

As evident from [9], that the conventional Hopfield neural network utilizes linear aggregation of weights, biases, and inputs at the particular node. The aggregated inputs are then passed through a nonlinearity to achieve a nonlinear input-output relation. The output of every neuron in the network is fed back to all the neurons of the network. Hence, it forms a fully connected recurrent neural network. However, in the proposed modified Hopfield neural network model, nonlinear combinations of inputs are aggregated at the summing node. The nonlinear combination includes linear as well as nonlinear terms. This is loosely inspired from the phenomenon of nonlinear dendritic interaction. This nonlinear combination of inputs is aggregated via unique weights. Other features are identical to conventional Hopfield model.

Fig.1 helps us to write dynamical equations for the proposed network. These dynamical equations are expressed as,

\[
\frac{du}{dt} = w_1 F_1(x, y) + \ldots + w_p F_p(x, y) + I_{bias}
\]
\[
\frac{dx}{dt} = v_1 F_1(x,y) + \ldots + v_i F_i(x,y) + \ldots + v_p F_p(x,y) + I_{bias1}
\]
\[
x = \varphi(u_x)
\]
\[
y = \varphi(u_y)
\]
\[
P = \frac{p(p+3)}{2}; p > 2
\] (8)

where \(w_1, \ldots, w_i, \ldots, w_p\) and \(v_1, \ldots, v_i, \ldots, v_p\) are the associated weights of the network. \(I_{bias1}\) and \(I_{bias2}\) are the constant bias values to the neurons \(x\) and \(y\), respectively. \(F_i(x,y), \ldots, F_i(x,y)\) are various combinations of inputs \(x\) and \(y\). The function \(\varphi(\cdot)\) is the input-output transfer function for the neurons. This function should be continually differentiable and monotonically increasing. Comparing Eq.(7) with Eq.(8) yields the weights \(w_j\), \(w_j, \ldots, w_j, w_p, v_1, \ldots, v_i, \ldots, v_p\) and biases \(I_{bias1}\) and \(I_{bias2}\).

3. STABILITY ANALYSIS OF THE MODIFIED HOPFIELD NEURAL NETWORK

In this section, the stability analysis for the proposed modified Hopfield neural network is carried out. We use the basic notions of Lyapunov stability theory [25], [34] in order to prove that the proposed energy function (Eq. 6) is a valid energy function.

Consider an autonomous system described by,
\[
\frac{dx}{dt} = f(x(t))
\] (9)

and let, \(x(t_0), t)\) be a solution. Let \(E(x)\) is the total energy associated with the system. If the derivative \(dE(x)/dt\) is negative for all \(x(t_0), t)\) except the equilibrium point, then it follows that energy of the system decreases as \(t\) increases and finally the system reaches the equilibrium point. This holds because energy is non negative function of system state which reaches a minimum only if the system motion stops.

In section 2, we formulated the energy function for the proposed model with the help of Eq.(1) through Eq.(4). This energy \(E(x,y)\) is the function of variables \(x, y\) and polynomial constant coefficients \(a_0, a_1, \ldots, a_n\). Eq.(5) can be rewritten as,
\[
E(x,y) = \sigma^2(x,y) + \sigma^2(x,y)
\] (10)

For the system to being stable, at any instant, the total energy \(E(x,y)\) in the system is positive unless the system is at rest at the equilibrium state, where the energy is zero. This can be written as,
\[
E(x,y) > 0; \text{when} x, y \neq 0
\]
\[
E(0,0) = 0; \text{when} x, y = 0
\] (11)

The rate of change of energy is given by,
\[
\frac{d}{dt} E(x,y) = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt}
\] (12)

The dynamics for the modified Hopfield neural network is given by,
\[
\frac{du_x}{dt} = -\frac{\partial}{\partial x} E(x,y)
\] (13)
\[
\frac{du_y}{dt} = -\frac{\partial}{\partial y} E(x,y)
\]

Using equations (12) and (13) we get,
\[
\frac{d}{dt} E(x,y) = -\frac{du_x}{dt} \frac{dx}{dt} - \frac{du_y}{dt} \frac{dy}{dt}
\] (14)

States \(u_x\) and \(u_y\) are obtained by,
\[
u_x = \varphi^{-1}(u_x)
\]
\[
u_y = \varphi^{-1}(u_y)
\] (15)

The activation function \(\varphi(\cdot)\) is chosen in such a way that, \(\varphi(\cdot)\) is continuously differentiable and strictly monotonically increasing, that is, \(\varphi(p) > \varphi(p_0)\) if \(p > p_0\). Considering the relation as shown in Eq.(14) and Eq.(15) results in,
\[
\frac{d}{dt} E(x,y) = -\left(\frac{d u_x}{dt}\right)^2 -\left(\frac{d u_y}{dt}\right)^2 \frac{\partial \varphi^{-1}(\cdot)}{\partial (x,y)}
\] (16)

Thus, \(\frac{d}{dt} E(x,y)\) is negative at all points except the equilibrium state. This proves the stability of the modified Hopfield like network in Lyapunov sense. The total energy of the system will decrease with time and the proposed network is stable.

The gradient descent approach for minimizing the energy function \(E(x,y)\) necessitates a positive slope for activation function \(\varphi(\cdot)\) with respect to states \(x\) and \(y\). In our experimentation, the input-output transfer function \(\varphi(\cdot)\) was linear. The input dynamics of network follows a gradient descent approach for minimizing the energy function \(E(x,y)\), the network inputs are guaranteed to converge at a local minimum of \(E(x,y)\), if \(E(x,y)\) has a one. Energy at a local minimum is zero, i.e. every minimum of \(E(x,y)\) is a global minimum and hence corresponds to a solution [4].

4. ILLUSTRATIVE EXAMPLES

Methodology for calculating the non-repeated roots of a characteristic equation by using the modified Hopfield neural network approach is discussed in previous section. In this section we demonstrate usefulness of proposed method by several numerical experiments. Section 4.1 will describe the proposed approach in details. However, for subsequent examples only results have been presented. Initial states of the network are randomly initialized in the range of [-1, 1] and network is simulated on a machine configured with Intel Pentium 4 CPU 3.20GHz 512 MB RAM.

4.1 EXAMPLE 1

Consider a characteristic equation of quadratic form, i.e. \(Ax^2 + Bx + C = 0\). The parameters \(A, B,\) and \(C\) of this equation are any arbitrary real valued constants.
4.1.1 Energy Function Formulation:

The given characteristic equation has two roots which can be either complex conjugate pair or real depending on the values of parameters $A$, $B$, and $C$. The parameters $A$, $B$, $C \in \mathbb{R}$ and variable $s \in \mathbb{C}$; here $\mathbb{R}$ is the set of real number and $\mathbb{C}$ represents the set of complex number. Replacing $s$ with $x + iy$ in the given characteristic equation results in,

$$A(x+iy)^2 + B(x+iy) + C = 0$$

which is further simplified as,

$$\left(A(x^2 - y^2) + Bx + C\right) + i\left(2Axy + By\right) = 0$$

The energy function for this problem is given by,

$$E(x, y) = (A(x^2 - y^2) + Bx + C)^2 + (2Axy + By)^2$$

(19)

Eq. (19) is used further to calculate the weights and biases for the proposed network.

4.1.2 Design Details of Proposed Network:

With the calculated weights and bias values, network attempts to minimize the energy function $E(x, y)$ for obtaining the solutions. We assumed that the proposed network does not have any inherent loss terms, and its dynamics is governed by Eq.(7). Using equations (7) and (19), we get,

$$\frac{du_x}{dt} = w_1x^3 + w_2y^3 + w_3x^2y + w_4xy^2 + w_5x^2 + w_6y^2$$

$$+ w_7xy + w_8x + w_9y + I_{bias1}$$

$$\frac{du_y}{dt} = v_1x^3 + v_2y^3 + v_3x^2y + v_4xy^2 + v_5x^2 + v_6y^2$$

$$+ v_7xy + v_8x + v_9y + I_{bias2}$$

$$x = \varphi(u_x)$$

$$y = \varphi(u_y)$$

(20)

Values of weights and biases are determined by comparing the RHS of Eq.(19) and Eq.(20). The architecture of the proposed network for quadratic characteristic equation is shown in Fig.2. Results for Example 1 are obtained by numerical simulation of Eq.(20) with $A = 1$, $B = 2$, and $C = 3$. In our experimentation, we noticed that in few cases the numerical method fails to converge if the integration step size is not appropriate and hence the selection of optimal integration step size is essential. Energy profile for the proposed network is drawn in Fig.3. It is evident from this figure that the energy of system minimizes with iterations and network requires small number of iterations to reach its minimum.

It is also noted here that all the non-repeated roots of the network are calculated sequentially. One complete simulation of the network results in one non-repeated root for the given polynomial. To evaluate all non-repeated roots the same network is simulated again and again with different initial states. In other words every time for a unique root for a characteristic equation the network is trained with unique random initial states. Each starting point in the basin of attraction must have different minima. If the starting point is same every time, we will get one root of the given characteristic equation. This is one of the limitations of proposed method because there is no mechanism for the selection of starting point. Because of this limitation the computational time is high as compared to conventional methods.

Table.1 depicts the comparison between proposed energy minimization based neural network technique (PM) and conventional methods used for root determination, i.e. Laguerre[32] (LM) and Eigenvalues of balanced companion matrix method [8] (EBCM). It is evident from this table that the roots obtained with all the methods are almost identical but the computation time required by the proposed technique is high. Thus, the incorporation of nonlinear dendritic interaction in the Hopfield neural network can be used for determination of non-repeated complex roots of polynomial. However, this method is not computationally efficient method.
### 4.2 EXAMPLE 2

The proposed methodology is used to calculate the non-repeated roots of a cubic characteristic equation of form, i.e. \( Ax^3 + Bx^2 + Cx + D = 0 \). The parameters \( A, B, C \) and \( D \) can have any arbitrary real values. Table 2 presents the simulation results and comparison with conventional numerical methods.

Table 2. The comparison of performance of the proposed method and conventional methods for the cubic characteristic equation at \( A = 1, B = 2, C = 3, \) and \( D = 5 \)

<table>
<thead>
<tr>
<th>Indices</th>
<th>PM</th>
<th>LM</th>
<th>EBCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated roots</td>
<td>-1.9081</td>
<td>-0.0781+1.644i</td>
<td>-0.0781+1.644i</td>
</tr>
<tr>
<td>CPU time(s)</td>
<td>0.39</td>
<td>0.0781-1.644i</td>
<td>0.0781-1.644i</td>
</tr>
</tbody>
</table>

### 4.3 EXAMPLE 3

In this section, we extended the proposed approach for finding all the non-repeated roots for a fifth order characteristic equation of form, i.e. \( Ax^5 + Bx^4 + Cx^3 + Dx^2 + Ex + F = 0 \). The parameters \( A, B, C, D, E, \) and \( F \) can have any arbitrary real values. Table 3 presents the simulation results and comparison with conventional numerical methods.

Table 3. The comparison of performance of the proposed method and conventional methods for the fifth order characteristic equation at \( A = 1, B = 2, C = 3, D = 6, E = 3, \) and \( F = 5 \)

<table>
<thead>
<tr>
<th>Indices</th>
<th>PM</th>
<th>LM</th>
<th>EBCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated roots</td>
<td>-0.2703-1.166i</td>
<td>-0.2703-1.166i</td>
<td>-0.2703-1.166i</td>
</tr>
<tr>
<td>CPU time(s)</td>
<td>1.34</td>
<td>0.031</td>
<td>0.062</td>
</tr>
</tbody>
</table>

### 4.4 EXAMPLE 4

Here we exhibit simulation results for the seventh order characteristic equation of form \( Ax^7 + Bx^6 + Cx^5 + Dx^4 + Ex^3 + Fx^2 + Gx + H = 0 \). The parameter \( A, B, C, D, E, F, G, \) and \( H \) can be any arbitrary real values. Table 4 presents the simulation results and comparison with conventional numerical methods.

Table 4. The comparison of performance of the proposed method and conventional methods for the seventh order characteristic equation at \( A = 1, B = 2, C = 3, D = 6, E = 3, F=5, G = 6 \) and \( H = 6 \)

<table>
<thead>
<tr>
<th>Indices</th>
<th>PM</th>
<th>LM</th>
<th>EBCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated roots</td>
<td>-0.7031+0.8971i</td>
<td>-0.7129+0.6506i</td>
<td>-0.03607+1.6113i</td>
</tr>
<tr>
<td>CPU time(s)</td>
<td>5.891</td>
<td>0.078</td>
<td>0.062</td>
</tr>
</tbody>
</table>

### 5. CONCLUSION AND DISCUSSION

In this paper, modified Hopfield neural network whose formulation is inspired from nonlinear dendritic aggregation is proposed. The modification lies in the application of linear as well as nonlinear terms in calculating the aggregation at each neuron unit.

We proposed a modified Hopfield neural network approach for determining the roots of a characteristic equation. We formulated an energy function that is used to determine the weights and biases for the proposed network. A generalized architecture of the proposed network is also given, which can be used to calculate non-repeated roots of any order of characteristic equation. Lyapunov stability analysis for the proposed network is carried out and it is found that proposed energy function for the network is stable in Lyapunov sense. Results for quadratic, cubic, fifth order, and seventh order characteristic equations are shown. These results are compared with the conventional Lagurre and Eigenvalues of balanced companion matrix methods. Computation time with each method is also compared. In our simulation the proposed method is not as fast as the conventional methods.

However, we believe that it is faster in comparison to feed forward type neural network. For determining the convergence and stability of the proposed network numerically, energy profile for the network is plotted. It has been noticed that the energy of the system minimizes with iterations. Thus, dynamics of the proposed network is stable.

Determination of proper weights and bias values, selection of suitable starting point, and higher computational time as compared to conventional numerical methods are some limitations of the proposed network. It is also observed that the proposed network should be trained with unique random initial states for approximately 5 times to the maximum number of roots in a given polynomial to evaluate non-repeated roots of given characteristic equation. It is also noted that each starting point in the basin of attraction must have different minima. If the starting point is same every time, we will get only one root. In this application, selection of starting point plays an important role. This can be rectified if some mechanism for the selection of starting point is developed.

This work can be taken as future research work. The proposed approach can also be extended for handling repeated roots of a characteristic polynomial. It is concluded that incorporating nonlinear dendritic aggregation in conventional Hopfield neural network and a proper formulation of energy function we can determine real as well as complex non-repeated roots of characteristic equations and set of nonlinear algebraic equations.
REFERENCES


