The Neural-Network Approaches to Solve Nonlinear Equation

Xiangde GUO
College of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha, China
Email: guoxiangde@126.com

Zhezhao ZENG
College of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha, China
Email: hncs6699@yahoo.com.cn

Abstract—In this paper, we proposed two neural-network approaches for solving nonlinear equations or polynomials. The first method is suitable for finding simple roots of nonlinear equation or polynomial, and the second approach is fit to finding both the multiple and simple roots of nonlinear equations or polynomial, which were not well solved by the other methods. The convergence of algorithm proposed was researched. The convergence theorem provides the theory criterion selecting learning rate of neural network. The specific examples showed that the proposed method can find the simple or multiple roots of nonlinear equations or polynomials at a very rapid convergence and very high accuracy with less computation.

Index Terms—neural-network, nonlinear equations, polynomials, multiple and simple roots

I. INTRODUCTION

Finding rapidly and accurately the roots of nonlinear equations is an important problem in various areas of control and communication systems engineering, signal processing and in many other areas of science and technology. The problem of finding the zeros of nonlinear equations has fascinated mathematicians for centuries, and the literature is full of ingenious methods, analysis of these methods, and discussions of their merits [1-3]. Over the last decades, there exist a large number of different methods for finding nonlinear equations roots either iteratively or simultaneously. Most of them yield accurate results only for small degree or can treat only special polynomials, e.g., polynomials with simple real or complex roots [4].

So far, some better modified methods of finding roots of nonlinear equations or polynomials cover mainly the Jenkins/Traub method [5], the Markus/Frenzel method [4], the Laguerre method [6], the Routh method [7], the Truong, Jeng and Reed method [8], the Fedorenko method [9], the Halley method[10], and some modified Newton’s methods[11-13], etc. Although the Laguerre method is faster convergent than all other methods mentioned above, it has more computation. Among other methods, some have low accuracy, and some have more computation, especially to say is the modified Newton’s methods must have a good initial value near solution. Furthermore, it is very difficult for the all methods mentioned above to find multiple real or complex roots of nonlinear equations or polynomials.

In this paper, we consider a neural network algorithm to find a simple root \( \alpha \), i.e., \( f(\alpha) = 0 \) and \( f'(\alpha) \neq 0 \), of a nonlinear equation \( f(x) = 0 \).

Newton’s method is the well-known iterative method for finding \( \alpha \) by using

\[
x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}
\]

that converges quadratically in some neighborhood of \( \alpha \) [14-15].

In recent years, some iterative methods have been proposed and analyzed for solving nonlinear equations that improve order of convergence of the classical methods such as Jarratt’s method, Euler-Chebyshev methods, Ostrowski’s method [14] given by

\[
y^k = x^k - \frac{f(x^k)}{f'(x^k)}
\]

\[
x^{k+1} = y^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)}
\]

which are improvements of Newton’s method; the order increases by at least two at the expense additional function evaluation at another point iterated by the classical methods [15].

Grau et al. [16] developed the sixth-order variant of Ostrowski’s method which is given by

\[
y^k = x^k - \frac{f(x^k)}{f'(x^k)}
\]

\[
z^k = y^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)}
\]

\[
x^{k+1} = z^k - \frac{f(x^k)}{f'(x^k)} \frac{f(z^k)}{f'(z^k)}
\]

Supported by the key project of the Education Department of Hunan Province (08A006)
Another sixth-order family of modified Ostrowski’s method was considered by Sharma et al. [17].

\[ y^k = x^k - \frac{f(x^k)}{f'(x^k)} \]  
\[ z^k = y^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)} \frac{f(z^k)}{f'(z^k)} \]  
\[ x^{k+1} = z^k - H(u^k) \frac{f(z^k)}{f'(z^k)} \]  

where \( \beta \in \mathbb{R} \)

Changbum Chun and YoonMee Ham proposed some sixth-order variant of Ostrowski’s method which is given by [15]

\[ y^k = x^k - \frac{f(x^k)}{f'(x^k)} \]  
\[ z^k = y^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)} \]  
\[ x^{k+1} = z^k - H(u^k) \frac{f(z^k)}{f'(z^k)} \]

where \( u^k = f(y^k) / f(x^k) \) and \( H(t) \) represents a real-valued function, it is observed that if we take
\[ H(t) = 1/(1 - 2t) \]

or
\[ H(t) = [1 + (\beta + 2)t] / (1 + \beta t) \], then the stepsub defined by (12) reduces to (6) and (9), respectively.

In order to solve the problems finding the multiple roots of nonlinear equations or polynomial, Dong [18], Neta and Johnson [19], Neta [20], etc. developed some effective methods.

In this paper, we proposed two neural-network methods for finding simple or multiple roots of nonlinear equations or polynomials. In section 2 and 3 we constructed a neural network algorithm and analyzed its convergence, respectively. We give the algorithm steps in section 4. The efficiency of the considered method was illustrated by numerical examples in section 5.

II. CONSTRUCTION OF THE NEURAL NETWORK ALGORITHM

A. For the function \( f(x) \) with only simple roots

We suppose that \( x^0 \) is close initial approximations of the root \( \alpha \) and \( x^k \) is a weight of nonlinear equation \( f(x) = 0 \). The algorithm of neural network is as follows

**Algorithm 2.1.** Define the error function and the objective function as follows

\[ e(k) = 0 - f(x^k) = -f(x^k) \]  
\[ J = \frac{1}{2} e^2(k) \]

To minimize the objective function \( J \), the weight \( x^k \) is recursively calculated via using a simple gradient descent rule:

\[ y^k = x^k - \mu(k) \frac{dJ}{dx} \]  

where \( \eta(k) \) is learning rate and \( 0 < \mu(k) < 1 \), and \( \Delta x^k = -\mu(k) \frac{dJ}{dx} \).

Differentiating (14) with respect to \( x^k \), it can be obtained that

\[ \Delta x^k = -\mu(k) \frac{dJ}{dx} = \mu(k)e(k) f'(x^k) \]

Substituting (16) into (15), we have

\[ y^k = x^k + \mu(k)e(k) f'(x^k) \]

According to [3], let

\[ z^k = y^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)} \frac{f(z^k)}{f'(z^k)} \]

Then

\[ x^{k+1} = z^k - \frac{f(x^k)}{f'(x^k)} \frac{f(y^k)}{f'(y^k)} \frac{f(z^k)}{f'(z^k)} \]

B. For the function \( f(x) \) with multiple roots

Consider the nonlinear equation or polynomial of the type

\[ f(x) = 0 \]

We assume that \( x \) is a multiple root of the Eq. (20) and \( x_0 \) is an initial guess sufficiently close to \( x \). Using the Taylor’s series expansion of the function \( f(x) \), we have

\[ f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) = 0 \]  

This alternative equivalence formulation has been used to develop a class of iterative methods for solving nonlinear equation or polynomial \( f(x) = 0 \). We can rewrite (21) in the following form:

\[ f''(x_0)(x - x_0)^2 + 2f'(x_0)(x - x_0) + 2f(x_0) = 0 \]

This is a quadratic equation. From which we have

\[ x - x_0 = \frac{-2f'(x_0) \pm \sqrt{[f'(x_0)]^2 - 8f''(x_0)f(x_0)}}{2f''(x_0)} \]  

or

\[ x - x_0 = \frac{-f'(x_0) \pm \sqrt{[f'(x_0)]^2 - 2f''(x_0)f(x_0)}}{f''(x_0)} \]

This formula gives two possibilities for \( x \), depending on the sign proceeding the radical term. Thus,
\[ x = x_0 - \frac{f'(x_0) + \sqrt{[f'(x_0)]^2 - 2f(x_0)f''(x_0)}}{f''(x_0)} \]  
\[ x = x_0 - \frac{f'(x_0) - \sqrt{[f'(x_0)]^2 - 2f(x_0)f''(x_0)}}{f''(x_0)} \]

This fixed point formulation allows us to suggest the following method.

**Algorithm 2.2.** For a given \( x_0 \), compute approximate solution \( x_{k+1} \) by the iterative schemes

For the smaller root, using the following iterative formula

\[ x_{k+1} = x_k - \frac{f'(x_k) + \sqrt{[f'(x_k)]^2 - 2f(x_k)f''(x_k)}}{f''(x_k)} \]

and for the larger root, using the following iterative formula

\[ x_{k+1} = x_k - \frac{f'(x_k) - \sqrt{[f'(x_k)]^2 - 2f(x_k)f''(x_k)}}{f''(x_k)} \]

Since the denominator of both formula (27) and (28) does not include the term \( f'(x_k) \), they are specially fit to finding double roots of nonlinear equation or polynomial.

### III. ANALYSIS OF CONVERGENCE

In this section we analyzed the convergence property of the Algorithm 2.1. The following theorem deals with the choice of the variable parameters \( \eta(k) \).

**Theorem 1.** Let \( f(x) = 0 \) be a nonlinear equation or polynomial, with simple root \( \alpha \). If \( x^0 \) is the initial approximations of the root \( \alpha \), only when the parameter satisfies

\[ 0 < \mu(k) < 2/\left| f'(x_k^k) \right|^2 \]

The Algorithm 2.1 is convergent.

**Proof.** Define an error function and a Lyapunov function respectively as follows

\[ V(k) = \frac{1}{2} e^2(k) \]

Then we have

\[ \Delta V(k) = \frac{1}{2} e^2(k+1) - \frac{1}{2} e^2(k) \]

Since

\[ e(k+1) = e(k) + \Delta e(k) = e(k) + \frac{de(k)}{dx^k} \Delta x^k \]

and

\[ \Delta e(k) = \frac{de(k)}{dx^k} \Delta x^k = -f'(x^k) \Delta x^k \]

According to formula (31), (32), and (33), we have

\[ \Delta V(k) = \Delta e(k)e(k) + \frac{1}{2} [\Delta e(k)]^2 \]

\[ = -e(k)f'(x^k) \Delta x^k + \frac{1}{2} [f'(x^k) \Delta x^k]^2 \]

Known from the (16) that

\[ \Delta x^k = \mu(k)e(k)f'(x^k) \]

Substituting formula (35) into the formula (34) gives

\[ \Delta V(k) = -[e(k)f'(x^k)]^2 \mu(k) \]

\[ + \frac{1}{2} \left[ \mu(k)e(k)f'(x^k) \right]^2 \]

If the formula (15) is convergent, i.e. \( \Delta V(k) < 0 \), then it is easy to see from (36) that

\[ 0 < 1/2 \left[ \mu(k)e(k)f'(x^k) \right]^2 < [e(k)f'(x^k)]^2 \mu(k) \]

Since \( \mu(k) > 0 \), hence we have from formula (37)

\[ 0 < \mu(k) < \frac{2}{\left| f'(x^k) \right|^2} \]

which proves the theorem.

### IV. ALGORITHM 2.1 STEPS

To find zero of nonlinear equation or polynomial \( f(x) = 0 \) given one approximation \( x^0 \):

**INPUT:** \( x^0 \); tolerance \( Tol \); maximum number of iterations \( N \); let \( k = 0 \);

**OUTPUT:** approximate solution \( x^k \) or message of failure.

**Step 1:** While \( k \leq N \) do Steps 2-5

**Step 2:** compute: \( e(k) = -f'(x^k) \), \( f'(x^k) \), and

\[ \mu_{opt}(k) = \begin{cases} 
0.5, & \left| f'(x^k) \right| < 1 \\
(1.0 - 1.6) \left| f'(x^k) \right|^2, & \left| f'(x^k) \right| \geq 1 
\end{cases} \]

**Step 3:** Estimate: \( y^k = x^k + \mu(k)e(k)f'(x^k) \), \( \hat{e}(k) = -f'(y^k) \), and

\[ z^k = y^k - \mu(k)e(k)\hat{e}(k)f'(x^k)/[2\hat{e}(k) - f'(x^k)] \]

Update weight:

\[ x^{k+1} = z^k + \mu(k)e(k) f'(y^k) f(z^k) / [2\hat{e}(k) - f'(x^k)] \]

**Step 4:** If \( \left| f(x^{k+1}) \right| \leq Tol \) then

**OUTPUT** \( (x^{k+1}) \); (The procedure was successful.)

**STOP**

**Step 5:** Set \( k = k + 1 \)

Go back to step 2
Step 6: OUTPUT ('the method failed after N iterations, 
\( n = k \));
(The procedure was unsuccessful.)

STOP

V. NUMERICAL EXAMPLES

A. For the function \( f(x) \) with only simple roots

To confirm the validity of the algorithm discussed, we gave two examples to evaluate the polynomial or nonlinear equation at the initial values, and compared results with other methods.

Example 1. Let
\[
f(z) = z^7 + z^5 - 10z^4 - z^3 - z + 10
\]
with zeros \( \xi_1 = 2 , \xi_2 = 1 , \xi_3 = -1 , \xi_4 = i , \xi_5 = -i , \xi_6 = -1 + 2i , \xi_7 = -1 - 2i \). As initial approximations, the following complex numbers have been chosen [21]:
\[
z_1^0 = 1.66 + 0.23i , \quad z_2^0 = 1.36 - 0.3i , \quad z_3^0 = -0.76 + 0.18i , \quad z_4^0 = -0.35 + 1.17i , \quad z_5^0 = 0.29 - 1.37i , \quad z_6^0 = -0.75 + 2.36i , \quad z_7^0 = -1.27 - 1.62i .
\]
The Table 1 display the absolute errors:
\[
e_i^k = |z_i^k - \xi_i| \quad (i = 1,2,3,4,5,6,7)
\]
using the neural network algorithm proposed and the formula (11) in [21]. Where, the formula (11) in [21] is as follows
\[
x_{i+1} = x_i - \frac{2W_i^k}{1 + s_i^k + p_i^k}
\]
\[
p_i^k = \sqrt{1 + s_i^k + 4W_i^k/\sum j \neq i (x_i^k - x_j^k)(x_i^k - W_j^k - x_j^k)}
\]
\[
s_i^k = \sum j \neq i W_j^k (x_i^k - x_j^k), \quad u_j^k = f(y_j^k) / f(x_j^k).
\]

Example 2. Let
\[
f(z) = z^4 - 1
\]
with zeros \( \xi_1 = 1 , \xi_2 = -1 , \xi_3 = i , \xi_4 = -i \). As initial approximations, the following complex numbers have been chosen [21]:
\[
z_1^0 = 0.5 + 0.5i , \quad z_2^0 = -1.36 + 0.42i , \quad z_3^0 = -0.25 + 1.28i , \quad z_4^0 = 0.46 - 1.37i
\]
The Table 2 displays the absolute errors:
\[
e_i^k = |z_i^k - \xi_i| \quad (i = 1,2,3,4)
\]
using the neural network algorithm proposed and the formula (11) in [21].

<table>
<thead>
<tr>
<th>( \text{Method proposed} )</th>
<th>Iteration formula (11) [21]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_1^k )</td>
<td>( k=4 )</td>
</tr>
<tr>
<td>( e_1^k )</td>
<td>exact</td>
</tr>
<tr>
<td>( e_2^k )</td>
<td>exact</td>
</tr>
<tr>
<td>( e_3^k )</td>
<td>exact</td>
</tr>
<tr>
<td>( e_4^k )</td>
<td>exact</td>
</tr>
<tr>
<td>( e_5^k )</td>
<td>0.674948 e-18</td>
</tr>
<tr>
<td>( e_6^k )</td>
<td>0.370194 e-18</td>
</tr>
<tr>
<td>( e_7^k )</td>
<td>exact</td>
</tr>
</tbody>
</table>

(2) Chebyshev’s method
\[
x_{n+1} = x_n + \frac{m(m-3)}{m} u_n \left[ 1 - \frac{m-3}{m} f'(x_n) \right] f(x_n)
\]
where \( u_n = f(x_n) / f'(x_n) \).

(3) The modified Chebyshev’s method
\[
x_{n+1} = x_n - u_n \left[ \beta + \gamma \frac{f'(y_n)}{f(x_n)} \right]
\]
where

© 2010 ACADEMY PUBLISHER
\[
y_n = x_n - \alpha u_n, \quad \alpha = -\frac{m(m - 3)}{2}, \quad \beta = -\frac{m}{(m - 3)}
\]
\[
\gamma = \frac{m(m - \alpha)}{\rho \alpha^2}, \quad \rho = \left(\frac{m - \alpha}{m}\right)^m.
\]

(4) Neta’s method
\[
x_{n+1} = x_n + \frac{m(m - 3)}{2} u_n \left[1 - \frac{m}{m - 3} u_n w_n \right]
\]
where
\[
w_n = \frac{6[f(x_{n-1}) - f(x_n)] + 2hf'(x_{n-1}) + 4hf''(x_n)}{h^2 f'(x_n)},
\]
\[h = x_n - x_{n-1}.
\]

(5) The modified Osada method
\[
x_{n+1} = x_n - \frac{m(m + 1)}{2} u_n + \frac{(m - 1)^2}{2w_n}
\]

Example 3. In third example we took a quadratic polynomial having a double root at \(\xi = 1\) [20]
\[f(x) = x^2 - 2x + 1\]
The iteration starts with \(x_0 = 0\) and the results are summarized in Table 3.

| Method | Initial value \(x_0 = 0\) | Iterations: \(n\) | \(|f(x_n)|\) |
|--------|-----------------|-----------------|---------|
| Proposed method (27) | 1 | 0 |
| Proposed method (28) | 1 | 0 |
| Halley’s method (2) in reference [20] | 1 | 0 |
| Chebyshev’s method (29) in ref. [20] | 1 | 0 |
| The modified Chebyshev’s method (39) in ref. [20] | 1 | 0 |
| Neta’s method (49) in reference [20] | 2 | 0 |
| The modified Osaka method (51) in ref. [20] | 2 | 0 |

Example 4. In fourth example we took a polynomial having two double roots at \(\xi = \pm 1\) [20]
\[f(x) = x^4 - 2x^2 + 1\]
The iteration starts respectively with \(x_0 = 0.8\) or \(x_0 = 0.6\). The results are summarized in Table 4.

| Method | Initial value \(x_0 = 0\) | Iterations: \(n\) | \(|f(x_n)|\) |
|--------|-----------------|-----------------|---------|
| Proposed method (27) | 1 | 0 |
| Proposed method (28) | 22 | 0 |
| Halley’s method (2) in ref. [20] | 3 | 1.0e-18 |
| Chebyshev’s method (29) in ref. [20] | 3 | 1.0e-15 |
| The modified Chebyshev’s method (39) in ref. [20] | 3 | 0 |
| Neta’s method (49) in ref. [20] | 7 | 1.0e-14 |
| The modified Osaka method (51) in ref. [20] | 9 | 0 |

Example 5. The fifth example is a polynomial with triple root at \(\xi = 1\) [20]
\[f(x) = x^5 - 8x^4 + 24x^3 - 34x^2 + 23x - 6\]
The iteration starts with \(x_0 = 0\) and the results are summarized in Table 5.

Example 6. The sixth example is a nonlinear equation with double root at \(\xi = 0\) [20]
\[f(x) = x^2 e^x\]
Starting at \(x_0 = 0.1\) or at \(x_0 = 0.2\), the results are given in Table 6.

Example 7. The seventh example having a double root at \(\xi = 1\) is [20]
\[f(x) = 3x^4 + 8x^3 - 6x^2 - 24x + 19\]
We started with $x_0 = 0$ and $x_0 = 0.5$ and the results are summarized in Table 7.

Starting at $x_0 = 1$ or at $x_0 = 4$, the results are given in Table 8.

<table>
<thead>
<tr>
<th>TABLE 6</th>
<th>COMPARISON OF 7 METHODS FOR EXAMPLE 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>$x_0 = 0.1$</td>
</tr>
<tr>
<td></td>
<td>Iters. $n$</td>
</tr>
<tr>
<td>Proposed method (27)</td>
<td>7</td>
</tr>
<tr>
<td>Proposed method (28)</td>
<td>7</td>
</tr>
<tr>
<td>Halley's method (2) in ref. [20]</td>
<td>2</td>
</tr>
<tr>
<td>Chebyshev's method (29) in ref. [20]</td>
<td>2</td>
</tr>
<tr>
<td>The modified Chebyshev's method (39) in ref. [20]</td>
<td>2</td>
</tr>
<tr>
<td>Neta's method (49) in ref. [20]</td>
<td>4</td>
</tr>
<tr>
<td>The modified Osada method (51) in ref. [20]</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 7</th>
<th>COMPARISON OF 7 METHODS FOR EXAMPLE 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>$x_0 = 0$</td>
</tr>
<tr>
<td></td>
<td>Iteration $n$</td>
</tr>
<tr>
<td>Proposed method (27)</td>
<td>10</td>
</tr>
<tr>
<td>Proposed method (28)</td>
<td>20</td>
</tr>
<tr>
<td>Halley’s method (2) in ref. [20]</td>
<td>3</td>
</tr>
<tr>
<td>Chebyshev’s method (29) in ref. [20]</td>
<td>5</td>
</tr>
<tr>
<td>The modified Chebyshev’s method (39) in ref. [20]</td>
<td>4</td>
</tr>
<tr>
<td>Neta’s method (49) in ref. [20]</td>
<td>5</td>
</tr>
<tr>
<td>The modified Osada method (51) in ref. [20]</td>
<td>5</td>
</tr>
</tbody>
</table>

**Example 8.** The sixth example having two simple roots at $\xi = 2$ and $\xi = 3$ is

$$f(x) = x^2 - 5x + 6$$

VI. CONCLUDING REMARKS

We can see from the table 1 to table 2 that the method proposed can simultaneously find the real or complex roots of polynomials with fast convergence and very high accuracy. Although the iterative numbers using the method proposed are four times as same as the iterative formula (11) in [21], see Table 1, the iterative formula (11) in [21] has more computation than the method proposed at each iterative. The results in example 2 show that the method proposed has much higher accuracy and faster convergence than the iterative formula (11) in [21]. Furthermore, we have developed two new methods to obtain not only multiple roots but also simple roots. We can know from the table 3 to table 8 that the numerical experiments show the rapid convergence and high accuracy for finding the multiple or simple roots of polynomials or nonlinear equations compared with other methods. It is specially to say that the methods proposed are more efficient than other methods in finding roots of quadratic polynomial or nonlinear equation with double or simple roots (see Table 3 and Table 8). Hence, the algorithm 2.2 proposed will play a very important role in the many fields of science and engineering practice.

REFERENCES


ZENG Zhe-zhao (1963-9) received the B.S., M.S. and D.S. degree from Xiangtan University in 1987, Tsinghua University in 1989 and Hunan University in 2008 respectively. From 1990 to 1993, he worked in Dongguang, Guangdong, China, where he was a technical manager from 1991 to 1993. From 1993 to 1999, he worked in Xiangtan University of Science & Technology, China. From 2000 to now, he has worked in Changsha University of Science & Technology, China. Since 2002, he has been a professor. He is engaging in theory and novel technology of the electronics and information engineering, principally in research on signal processing , theory and application of neural networks, and PID intelligent control.

Email: hncs6699@yahoo.com.cn; hncsdlxdz@vip.sina.com