A COMPARISON BETWEEN REAL AND COMPLEX VALUED NEURAL NETWORKS IN COMMUNICATION APPLICATIONS

N. Benvenuto(*) , M. Marchesi(**) , F. Piazza, A. Uncini

Dip. di Elettronica e Automatica, Università di Ancona
Via Brecce Bianche, 60131 Ancona - Italy
(*) Dip. di Elettronica ed Informatica, Università di Padova
Via Gradenigo 6/a, 35131 Padova - Italy
(**) Dip. di Ingegneria Biof. ed Elettronica, Università di Genova
Via Opera Pia 11/a, 16145 Genova - Italy

Abstract

Neural networks are becoming a very useful tool for the processing of data in several communication applications, where it is of great interest to have complex computing structures. In this paper, after a brief review of existing techniques, a new complex Multi-Layer Perceptron is proposed. In particular, after defining a proper neuron structure (i.e. using a non-analytic activation function), a novel learning algorithm is derived.

On equalizing an 8-PSK complex signal distorted by a nonlinear communication channel, a comparison between this new complex network and the real-valued Multi-Layer Perceptron is also reported in this paper. The results show that the proposed network has a lower computation complexity and fewer parameters, for the same performance, than the classical real-valued network.

1. Introduction

In the last few years, due to their capabilities of classification, generalization, and learning-from-examples, neural network models have been used in many DSP applications. Among the various models, the Multi-Layer Perceptron (MLP) can be considered a very general scheme for the nonlinear processing of real-valued digital signals [1]. It has been successfully used, by means of the Back Propagation (BP) learning algorithm, as a predictor [2], as a filter of biological data [3], and as an equalizer of signals distorted by nonlinear communication channels [4].

However, in some DSP problems the available signals and desired output are in complex form. A typical example is in communication systems where a complex envelope representation is used instead of the more cumbersome real-valued bandpass. In these cases, the classical MLP cannot by directly applied because it requires both input and activation signals be real-valued. In order to extend the MLP model to complex signals, two approaches can be considered:

1) one can use a classical real-valued MLP where the complex input and output signals of the network are replaced by pairs of independent real-valued signals;
2) another approach is to use a MLP structure extended to work with complex signals. In this case we need to redefine both the complex neuron and the learning algorithm.

In this paper, based on the second approach, a new complex MLP algorithm is proposed. Moreover, using a realistic communication problem as a test case, a comparison between the two approaches, in terms of performance and computation complexity, is presented.
2. Complex-Valued Multi-Layer Perceptron

Although complex-valued neural networks have already been proposed [6,7], they presents some constraints for our applications. Therefore a novel complex MLP is presented here.

Firstly, we considered the problem of extending the nonlinear activation function to complex values. Let us note that if the classical real-valued sigmoidal function [1] is extended through a standard process of analytic continuation, the obtained complex function becomes unbounded in the complex plane, i.e. it presents some discontinuities [6,7]. This problem can be alleviated by choosing different, more suitable, analytic activation functions [6]. However, it cannot be eliminated because of the Liouville's theorem.

On the other hand, the analytic property of the activation function is necessary in order to extend the Back Propagation algorithm to complex signals. To solve this apparent dilemma, let $f(.)$ be the classical real-valued sigmoidal function [1], and $z$ a complex variable, we propose the following complex activation function:

$$F(z) = f(\text{Re}(z)) + j f(\text{Im}(z))$$

where $j$ is the imaginary unit.

Now the complex function $F(z)$ is always bounded for any $z$ and behaves in a way very close to its real-valued counterpart.

Unfortunately, $F(z)$ is not analytic and therefore the standard Back Propagation algorithm cannot be used. For MLP networks employing neurons with above activation functions and complex inputs, a novel BP recursive algorithm must be derived. Here we only report the final formulation.

Let "s" be the layer index, $s=0$ is the input layer, $s=1,...,M-1$, are the hidden layers, and $s=M$ is the output layer], "k" be the neuron index $k=0$ refers to bias neurons with output set to 1, while $k=1,...,N_s$, refer to neurons of layer (s), $o_k^{(s)}$ be the activation output of $k$-th neuron of $s$-th layer, and $w_{kj}^{(s)}$ be the synaptic weight of $k$-th neuron of $s$-th layer, relative to the $j$-th neuron of (s-1)-th layer. Hence, remembering that both signals and network parameters are complex, it can be shown that the following expressions hold true:

**Forward Phase:**

$$s = 1, ..., M$$

$$\text{net}_k^{(s)} = \sum_{j=0}^{N_{s+1}} w_{kj}^{(s)} o_j^{(s-1)} , \quad o_k^{(s)} = F(\text{net}_k^{(s)}) , \quad k = 1, ..., N_s$$

**Learning Phase:**

$$s = M, ..., 1$$

$$\sigma_k^{(s)} = \begin{cases} t_k \cdot o_k^{(M)} , & \text{for } s = M \\ \sum_{j=1}^{N_{s+1}} w_{kj}^{*(s+1)} \delta_j^{(s+1)} , & \text{for } s = M-1, ..., 1 \end{cases}$$

with

$$\delta_k^{(s)} = \text{Re}[\sigma_k^{(s)}] f'(\text{Re}[\text{net}_k^{(s)}]) + j \text{ Im}[\sigma_k^{(s)}] f'(\text{Im}[\text{net}_k^{(s)}]) , \quad k = 1, ..., N_s$$

$$\Delta w_{kj}^{(s)} = -\eta \delta_k^{(s)} o_j^{*(s-1)} , \quad k = 1, ..., N_s , \quad j = 0, ..., N_{s-1}$$

where $(.)^*$ denotes complex conjugate, $t_k, k=1,...,N_M$, is the target output, $f'(.)$ is the derivative of $f(.)$, $\eta$ is a suitable positive real constant (learning rate), and $\Delta w_{kj}^{(s)}$ is the computed synaptic
weight variation. It is interesting to note that for $M=1$ the above recursions yield the well known LMS algorithm for complex signals. Moreover, in the real case, they collapse to the classical real-valued BP algorithm.

3. Experimental Results

There are no canonical problems for testing a complex MLP; also the extension of some classical problems, like the XOR problem, to the complex plane [7] seems to be actually useless. Since the scope of application of these complex networks is mainly restricted to DSP problems within the communication area, we use as test case the problem of equalizing a data signal distorted by nonlinear channels as those found in typical PSK/TDMA satellite transmission systems [8]. Here, the channel nonlinearities are modeled by a complex Volterra expansion up to the $5$-th order. We just mention that real-valued MLP has already been proven to be a good equalizer for similar real-valued channels [5].

The neural equalizer is composed of a set of complex inputs, one hidden layer, and an output layer. In order to correctly reproduce the signal level range at the output, the output neurons have a linear activation function, i.e. they act as linear combiners.

To compare performance of the two approaches introduced in section 1, the two structures have been applied to the equalization of the output of a nonlinear 8-PSK channel described by the same Volterra coefficients of [8]. The performance criterion is the mean square error between the output at time $n$ and the complex transmitted data sequence $a_{in}$:

$$\text{MSE} = \begin{cases} E \left[ (o_{1,n}(M) - \text{Re}(a_{in}))^2 + (o_{2,n}(M) - \text{Im}(a_{in}))^2 \right], & \text{real-valued MLP} \\ E \left[ |o_{1,n}(M) - a_{in}|^2 \right], & \text{complex-valued MLP}. \end{cases}$$

Some simulations are reported in Tables I and II for real and complex-valued neural networks, respectively. Performance is computed for various networks whose configurations are denoted by $i-j-k$ (i units in the input layer, j units in the hidden layer and k outputs). The epochs processed were 10000 and each epoch is formed of 256 complex samples. In all cases, the evaluated MSE is relative to the last epoch, i.e. after 2 560 000 iterations.

For each network configuration, we report both the number of real parameters (i.e. the number of independent real coefficients) and the number of the real multiplications the network implies, when working in forward mode. Here each complex multiplication is considered to be equivalent to four real multiplications, although with proper definition the number of real multiplications could be reduced to three.

A comparison between the two approaches indicates that for the same computation complexity, the two networks yield almost the same performance. Actually, even if the real-valued network includes the complex-valued configuration, its performance, after 10000 epochs, is slightly lower than the equivalent complex-valued network. We believe this is due to the fact that the real-valued network has many more parameters (i.e. degrees of freedom) and therefore the adaptation algorithm is prone to find a local minimum rather than the optimum configuration.

We can conclude that our definition of a complex neuron and related Back Propagation algorithm is successful both in terms of computation complexity and number of parameters. Other areas of application of this tool are under investigation.

4. References


TABLE I
MEAN SQUARE ERROR ACHieved WITH REAL-VALUED NEURAL NETWORKS OVER THE COMPLEX NONLINEAR CHANNELS [8]

<table>
<thead>
<tr>
<th>real-valued network configuration</th>
<th>number of real multiplications</th>
<th>number of real parameters</th>
<th>MSE (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 - 4 - 2</td>
<td>48</td>
<td>54</td>
<td>-26.4</td>
</tr>
<tr>
<td>10 - 6 - 2</td>
<td>72</td>
<td>80</td>
<td>-28.8</td>
</tr>
<tr>
<td>10 - 8 - 2</td>
<td>96</td>
<td>106</td>
<td>-28.6</td>
</tr>
<tr>
<td>14 - 6 - 2</td>
<td>96</td>
<td>104</td>
<td>-27.9</td>
</tr>
<tr>
<td>14 - 8 - 2</td>
<td>128</td>
<td>138</td>
<td>-27.8</td>
</tr>
<tr>
<td>14 - 10 - 2</td>
<td>160</td>
<td>172</td>
<td>-27.7</td>
</tr>
<tr>
<td>14 - 14 - 2</td>
<td>224</td>
<td>240</td>
<td>-28.4</td>
</tr>
</tbody>
</table>

TABLE II
MEAN SQUARE ERROR ACHieved WITH COMPLEX-VALUED NEURAL NETWORKS OVER THE COMPLEX NONLINEAR CHANNELS [8]

<table>
<thead>
<tr>
<th>complex-valued network configurations</th>
<th>number of real multiplications</th>
<th>number of real parameters</th>
<th>MSE (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 - 2 - 1</td>
<td>48</td>
<td>30</td>
<td>-27.0</td>
</tr>
<tr>
<td>5 - 3 - 1</td>
<td>72</td>
<td>44</td>
<td>-28.3</td>
</tr>
<tr>
<td>5 - 4 - 1</td>
<td>96</td>
<td>58</td>
<td>-29.6</td>
</tr>
<tr>
<td>7 - 3 - 1</td>
<td>96</td>
<td>56</td>
<td>-29.0</td>
</tr>
<tr>
<td>7 - 4 - 1</td>
<td>128</td>
<td>74</td>
<td>-28.4</td>
</tr>
<tr>
<td>7 - 5 - 1</td>
<td>160</td>
<td>92</td>
<td>-29.5</td>
</tr>
<tr>
<td>7 - 7 - 1</td>
<td>224</td>
<td>128</td>
<td>-30.6</td>
</tr>
</tbody>
</table>

This work was supported in part by the Consiglio Nazionale delle Ricerche of Italy under the project "Rei Neurali" and in part by the Ministero dell'Università e della Ricerca Scientifica e Tecnologica of Italy.