Complex-Valued Neural Networks

Tohru Nitta
AIST, Japan

INTRODUCTION

The usual real-valued artificial neural networks have been applied to various fields such as telecommunications, robotics, bioinformatics, image processing and speech recognition, in which complex numbers (two dimensions) are often used with the Fourier transformation. This indicates the usefulness of complex-valued neural networks whose input and output signals and parameters such as weights and thresholds are all complex numbers, which are an extension of the usual real-valued neural networks. In addition, in the human brain, an action potential may have different pulse patterns, and the distance between pulses may be different. This suggests that it is appropriate to introduce complex numbers representing phase and amplitude into neural networks.

Aizenberg, Ivaskiv, Pospelov and Hudiakov (1971) (former Soviet Union) proposed a complex-valued neuron model for the first time, and although it was only available in Russian literature, their work can now be read in English (Aizenberg, Aizenberg & Vandewalle, 2000). Prior to that time, most researchers other than Russians had assumed that the first persons to propose a complex-valued neuron were Widrow, McCool and Ball (1975). Interest in the field of neural networks started to grow around 1990, and various types of complex-valued neural network models were subsequently proposed. Since then, their characteristics have been researched, making it possible to solve some problems which could not be solved with the real-valued neuron, and to solve many complicated problems more simply and efficiently.

BACKGROUND

The generic definition of a complex-valued neuron is as follows. The input signals, weights, thresholds and output signals are all complex numbers. The net input U_n to a complex-valued neuron n is defined as:

$$U_n = \sum_m W_{nm} X_m + V_n \tag{1}$$

where W_{nm} is the complex-valued weight connecting complex-valued neurons n and m, X_m is the complex-valued input signal from the complex-valued neuron m, and V_n is the complex-valued threshold of the neuron n. The output value of the neuron n is given by $f_C(U_n)$ where $f_C : \boldsymbol{C} \rightarrow \boldsymbol{C}$ is called *activation function* ($\boldsymbol{C}$ denotes the set of complex numbers). Various types of activation functions used in the complex-valued neuron have been proposed, which influence the properties of the complex-valued neuron, and a complex-valued neural network consists of such complex-valued neurons.

For example, the component-wise activation function or real-imaginary type activation function is often used (Nitta & Furuya, 1991; Benvenuto & Piazza, 1992; Nitta, 1997), which is defined as follows:

$$f_C(z) = f_R(x) + i f_R(y) \tag{2}$$

where $f_R(u) = 1/(1+\exp(-u))$, $u \in \boldsymbol{R}$ ($\boldsymbol{R}$ denotes the set of real numbers), i denotes $\sqrt{-1}$, and the net input U_n is converted into its real and imaginary parts as follows:

$$U_n = x + iy = z. \tag{3}$$

That is, the real and imaginary parts of an output of a neuron mean the sigmoid functions of the real part x and imaginary part y of the net input z to the neuron, respectively.

Note that the component-wise activation function (eqn (2)) is bounded but non-regular as a complex-valued function because the Cauchy-Riemann equations do not hold. Here, as several researchers have pointed out (Georgiou & Koutsougeras, 1992; Nitta, 1997) in the complex region, we should recall the Liouville's theorem, which states that if a function g is regular at all $z \in \boldsymbol{C}$ and bounded, then g is a constant function. That is, we need to choose either regularity or boundedness for an activation function of complex-valued neurons. In addition, it has been proved that the complex-valued neural network with the component-wise activation function (eqn (2)) can approximate any continuous complex-valued function, whereas a network with a regular activation function (for example, $f_C(z) = 1/(1+exp(-z))$ (Kim & Guest, 1990), and $f_C(z) = tanh(z)$ (Kim & Adali, 2003)) cannot approximate any non-regular complex-valued function (Arena, Fortuna, Re & Xibilia, 1993; Arena, Fortuna, Muscato & Xibilia, 1998). That is, the complex-valued neural network with the non-regular activation function (eqn (2)) is a universal approximator, but a network with a regular activation function is not. It should be noted here that the complex-valued neural network with a regular complex-valued activation function such as $f_C(z) = tanh(z)$ with the poles can be a universal approximator on the compact subsets of the deleted neighbourhood of the poles (Kim & Adali, 2003). This fact is very important theoretically, however, unfortunately the complex-valued neural network for the analysis is not usual, that is, the output of the hidden neuron is defined as the product of several activation functions. Thus, the statement seems to be insufficient to compare with the case of component-wise complex-valued activation function. Thus, the ability of complex-valued neural networks to approximate complex-valued functions depends heavily on the regularity of activation functions used.

On the other hand, several complex-valued activation functions based on polar coordinates have been proposed. For example, Hirose (1992) proposed the following amplitude-phase type activation function:

$$f_C(z) = \tanh\left(\frac{\alpha}{m}\right) \cdot \exp(i\beta), \; z = \alpha \cdot \exp(i\beta), \tag{4}$$

where m is a constant. Although this amplitude-phase activation function is not regular, Hirose noted that the non-regularity did not cause serious problems in real applications and that the amplitude-phase framework is suitable for applications in many engineering fields such as optical information processing systems, and amplitude modulation, phase modulation and frequency modulation in electromagnetic wave communications and radar. Aizenberg et al. (2000) proposed the following activation function:

$$f_C(z) = \exp\left(i\frac{2\pi j}{k}\right), \text{ if } \frac{2\pi j}{k} \le \beta < \frac{2\pi(j+1)}{k},$$
$$z = \alpha \cdot \exp(i\beta), \quad j = 0, 1, \ldots, k-1 \tag{5}$$

where k is a constant. Eqn (5) can be regarded as a type of amplitude-phase activation functions. Only phase information is used and the amplitude information is discarded, however, many successful applications show that the activation function is sufficient.

INHERENT PROPERTIES OF THE MULTI-LAYERED TYPE COMPLEX-VALUED NEURAL NETWORK

This article presents the essential differences between multi-layered type real-valued neural networks and multi-layered type complex-valued neural networks, which are very important because they expand the real application fields of the multi-layered type complex-valued neural networks. To the author's knowledge, the inherent properties of complex-valued neural networks with regular complex-valued activation functions have not been revealed except their learning performance so far. Thus, only the inherent properties of the complex-valued neural network with the non-regular complex-valued activation function (eqn (2)) are mainly described: (a) the learning performance, (b) the ability to transform geometric figures, and (c) the orthogonal decision boundary.

Learning Performance

In the applications of multi-layered type real-valued neural networks, the error back-propagation learning algorithm (called here, *Real-BP*) (Rumelhart, Hinton & Williams, 1986) has often been used. Naturally, the complex-valued version of the Real-BP (called here, *Complex-BP*) can be considered, and was actually proposed by several researchers (Kim & Guest, 1990; Nitta & Furuya, 1991; Benvenuto & Piazza, 1992; Georgiou & Koutsougeras, 1992; Nitta, 1993, 1997; Kim & Adali, 2003). This algorithm enables the network to learn complex-valued patterns naturally.