Optimizing Multilayer Perceptrons to Approximate Nonlinear Quaternion Functions

Arturo Buscarino[**], Luigi Fortuna [**] and Gabriele Puglisi[*]

[*]*Dipartimento di Ingegneria Elettrica Elettronica e Informatica, University of Catania, Italy*

[**]*IASI, Consiglio Nazionale delle Ricerche (CNR), Roma, Italy*

Abstract. In this contribution, a novel approach to optimize Multilayer Perceptron Artificial Neural Networks (MLP-ANN) devoted to approximate quaternion valued functions is presented. The approach is based on the definition of proper auxiliary networks devoted to predict the trends of the main network weights during the learning phase, thus reducing the number of epochs needed to reach a suitable abstraction level.

Introduction

In the last decade the use of artificial neural networks has received a growing interest in a wide variety of scientific fields. Our interest in this contribution focuses on multilayer perception based structures, able to deal with quaternion algebra. Quaternions have been introduced in order to generalize complex number properties to a three-dimensional space and have been used in a wide variety of practical applications, including the reformulation of Maxwell equations, the analysis of stresses in three-dimensional objects, and in digital filter design. Moreover, applications can be found also in robotics, due to the possibility of representing rotations in the three-dimensional space with quaternions instead of matrices. From the large variety of applications of quaternion algebra, the idea of developing a MLP-ANN structure able to deal directly with quaternions to approach these problems more efficiently has been discussed in [1]. In this contribution, we aim at optimizing the MLP-ANN structure for quaternions in order to improve the approximation of complex valued nonlinear functions and the efficiency of the learning algorithm. To this latter aim, we will focus on a recent implementation based on parallel computing of generic MLP-ANNs [2].

The main idea is to exploit the parallel computing capabilities of modern microprocessors in embedded devices to instantiate a set of auxiliary networks devoted to the prediction of the learning algorithm output, thus sensibly reducing the time to get a robust and efficient learning error. MLP-ANNs, in fact, are based on a learning phase during which inputs are presented to the network in order to favour its plasticity in the direction of abstracting the relationship among the inputs with the corresponding outputs. The variables of the problem are the so-called weights connecting the neurons of the ANN. The proposed approach introduces a series of auxiliary networks running in parallel with the main network, each devoted to predict the trend of the weights.

Results and discussion

Thanks to the parallel computing properties of current hardware and software technical solutions it is possible to reconsider algorithms and approaches presented in the literature before the diffusion of modern computers. This consideration led us to the implementation of the auxiliary networks paradigm to improve the efficiency of MLP-ANN for quaternion valued nonlinear functions.

A software improving the existing learning procedures for neural networks has been proposed and programmed through ANSI C language. It is possible to show practically the improvements, in terms of epochs needed to reach the convergence of the learning phase, by using datasets of different nature coming from technical high-precision circuits simulations, from highly complex experimental scenarios, such as nuclear fusion experimentations, and from planning of trajectory of robot manipulators. The adoption of auxiliary networks for the computing of the optimal weights increases remarkably the convergence velocity of a very simple kind of neural network, especially when the dimensions of the net in terms of layers and neurons are very high. In all the case studies taken into consideration, the desired approximation is reached with a reduced number of learning epochs in the order of 8000, as shown in Fig. 1.

References

[1] Arena, P., Fortuna, L., Muscato, G., and Xibilia, M. G. (1997). Multilayer Perceptions to Approximate Functions Quaternion Valued. Neural Networks, 10(2), 335-342.

[2] Bucolo, M., Buscarino, A., Fortuna, L., and Puglisi, G. (2022). Learning-on-learning approach for modeling. 48th Annual Conference of the IEEE Industrial Electronics Society, IECON 2022.

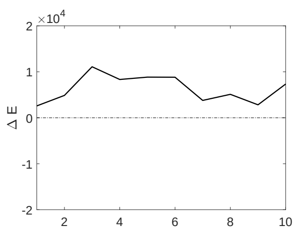

Figure 1: Experimental results: reduction in the number of learning epochs.