An Effective Solution to Regression Problem by RBF Neuron Network

Dang Thi Thu Hien, Faculty of Information Technology, University of Transport and Communications, Hanoi, Vietnam

Hoang Xuan Huan, Faculty of Information Technology, Vietnam National University, Hanoi, Vietnam

Le Xuan Minh Hoang, Faculty of Information Technology, Vietnam National University, Hanoi, Vietnam

ABSTRACT

Radial Basis Function (RBF) neuron network is being applied widely in multivariate function regression. However, selection of neuron number for hidden layer and definition of suitable centre in order to produce a good regression network are still open problems which have been researched by many people. This article proposes to apply grid equally space nodes as the centre of hidden layer. Then, the authors use k-nearest neighbour method to define the value of regression function at the center and an interpolation RBF network training algorithm with equally spaced nodes to train the network. The experiments show the outstanding efficiency of regression function when the training data has Gauss white noise.

Keywords: Artificial Neural Network, Multivariate Interpolation and Approximation Function, Radial Basis Functions, Regression k-Nearest Neighbors

1. INTRODUCTION

Multivariable function regression is a traditional and very important problem in numerical analysis which is widely applied (Alpaydin, 2010; Bartels, Beatty, & Barsky, 1987; Bromhead & Lowe, 1988; Collatz, 1966; Park & Sandberg, 1993; Tomohiro, Sadanori, & Seiya, 2008). When no concerning on noise distribution characteristics, this problem is stated and applied through interpolation and function approximation problems. 1-D case was researched and solved by Lagrange and Chebyshev by using polynomial as regression function. Since mid of 20th century till now, with research development and application of machine learning, image processing, computer graphic and technical problems, the regression problem has been also attracting many people to study. Among of them, selection of regression function form and good definition method is still

DOI: 10.4018/IJORIS.2015100104
an interested prime research topic for researchers (Blanzieri, 2003; Huan, Hien, & Huu-Tue, 2007; Schwenker, Kesler, & Gunther, 2001; Tomohiro et al., 2008).

During 3 recent decades, MLP (Multiple-Layered Perceptron), RBF (Radial Basis Function) neuron networks are effective tools to solve this problem in applications (Blanzieri, 2003; Haykin, 1999; Huan, Hien, & Huu-Tue, 2011; Looney, 1997; Rudenko & Bezsonov, 2011).

RBF regression method was proposed by Powell, introduced by Broomhead and Lowe as a neuron network (Powell, 1988; Bromhead et al., 1988). In comparison to MLP neuron network, RBF neuron network (hereinafter called RBF network) has short training period and is suitable for Regression problems. Training process of RBF network includes: 1) Defining the number of neurons in hidden layer and corresponding centre; 2) defining radius parameters of hidden neurons and the weight of output layer, in which definition of suitable neuron number in hidden layer, and the centre and radius parameters to produce a good Regression function is still an open problem (Powell, 1988; Blanzieri, 2003; Fasshauer, 2007; Pérez-Godoy, Rivera, Carmona, & del Jesus, 2014; Schwenker et al., 2001; Tomohiro et al., 2008; Weruaga & Via, 2014). The authors usually base on interpolation nodes’ distribution characteristics o determine center and radius parameter (Guang-Bin Huang, Saratchandran, & Sundararajan, 2004; Pérez-Godoy et al., 2014; Tomohiro et al., 2008; Sum, Chi-Sing Leung, & Ho, 2009).

Another approach that is to use interpolation nodes as centers for Gauss norm radius functions of interpolation network when number of interpolation node is not many. Based on mathematical analysis, Huan et al. proposed an iterative two-phase HDH to build interpolation RBF network (Huan et al., 2007). This HDH training algorithm has two phases: in the first, it iteratively computes the RBF width parameters; and in the second, the weights of the output layer are determined by the simple iterative method. HDH has very fast training time and being used for the problems with number of nodes much larger than the other ones, which are having the same approach. When, equally spaced nodes by using an appropriate metric to pre-determine radius widths in the first phase. By that time, HDH becomes a quick one-phase algorithm QHDH (Huan et al., 2011). This allows to train interpolation network very fast and generality is also better than HDH.

In this article, we explore the advantages of interpolation network training method by the above algorithm to build a regression network. Firstly, generating an equally spaced nodes grid on a specific domain and use them as the centre of hidden neurons; and use nearest neighbour \( k \) algorithm basing on training data set to define the approximated value of Regression function at the centres. Then use fast training RBF network algorithm proposed by Huan et al. to train the Regression network (Huan et al., 2011). This method may be applied when training data contains white noises. Experiments of the article comparing the methods of Tomohiro and Guang-Bin Huang on real data set show outstanding efficiency of the newly proposed method (Tomohiro et al., 2008; Guang-Bin Huang et al., 2004).

Next, in section 2, we will make a short introduction about regression and multivariable interpolation problems, fast training algorithm for RBF network with equally spaced nodes proposed by Huan et al. as the base for new method (Huan et al., 2011). Section 3 explains newly proposed solution for Regression problem. Section 4 states testing results and compares with GIC method of Tomohiro, method GGAP of Guang-Bin Huang on real data set (Guang-Bin Huang et al., 2004; Tomohiro et al., 2008). Then conclusion will be the last.