Recurrent Neural Networks Training With Stable Bounding Ellipsoid Algorithm

Wen Yu, Senior Member, IEEE, and José de Jesús Rubio

Abstract—Bounding ellipsoid (BE) algorithms offer an attractive alternative to traditional training algorithms for neural networks, for example, backpropagation and least squares methods. The benefits include high computational efficiency and fast convergence speed. In this paper, we propose an ellipsoid propagation algorithm to train the weights of recurrent neural networks for nonlinear systems identification. Both hidden layers and output layers can be updated. The stability of the BE algorithm is proven.

Index Terms—Bounding ellipsoid (BE), identification, recurrent neural networks.

I. INTRODUCTION

RECENT results show that neural network techniques seem to be effective to identify a broad category of complex nonlinear systems, when complete model information cannot be obtained. Neural networks can be classified as feedforward and recurrent ones [8]. Feedforward networks, for example multilayer perceptrons, are implemented to approximate nonlinear functions. The main drawback of these neural networks is that the weights’ updating does not utilize information on the local data structure and the function approximation is sensitive to the training data [17]. Since recurrent networks incorporate feedback, they have powerful representation capabilities and can successfully overcome the disadvantages of feedforward networks [13]. Even though backpropagation has been widely used as a practical training method for neural networks, there are some limitations such as slow convergence, local minima, and sensitivity to noise.

In order to overcome these problems, many methods for neural identification, filtering, and training have been proposed, for example, Levenberg–Marquardt, momentum algorithms [15], extended Kalman filter [23], and least squares approaches [17], which can speed up the backpropagation training. Most of them use static structures. There are some special restrictions for recurrent structure. In [2], the output layer must be linear and the hidden-layer weights are chosen randomly. The extended Kalman filter with decoupling structure has fast convergence speed [22], however the computational complexity in each interaction is increased. Decoupled Kalman filter with diagonal matrix $P$ [19] is similar to the gradient algorithm, so the convergence speed cannot be increased. A main drawback of the Kalman filter training is that theoretical analysis requires the uncertainty of neural modeling to be Gaussian process.

In 1979, Khachiyan indicated how an ellipsoid method for linear programming can be implemented in polynomial time [1]. This result caused great excitement and stimulated a flood of technical papers. The ellipsoid technique is a helpful tool in state estimation of dynamic systems with bounded disturbances [5]. There are many potential applications to problems outside the domain of linear programming. Weyer and Campi [27] obtained confidence ellipsoids which are valid for a finite number of data points, whereas Ros et al. [20] presented an ellipsoid propagation such that the new ellipsoid satisfies an affine relation with another ellipsoid. In [3], the ellipsoid algorithm is used as an optimization technique that takes into account the constraints on cluster coefficients. Lorenz and Boyd [14] described in detail several methods that can be used to derive an appropriate uncertainty ellipsoid for the array response. In [16], the problem concerning asymptotic behavior of ellipsoid estimates is considered for linear discrete-time systems. There are few application of ellipsoid on neural networks. In [4], unsupervised and supervised learning laws in the form of ellipsoids are used to find and tune the fuzzy function rules. In [12], ellipsoid type of activation function is proposed for feedforward neural networks. In [10], multivariate optimization for bounding ellipsoid (BE) algorithms is introduced. In [6], a simple adaptive algorithm is proposed that estimates the magnitude of noise. They are based on two operations of ellipsoid calculus: summation and intersection which correspond to the prediction and correction phase of the recursive state estimation problem, respectively.

In [21], we used the BE algorithm to train recurrent neural networks. But the training algorithm does not have standard recurrent form, so theory analysis cannot be implemented. In this paper, we modify the above algorithm and analyze the stability of nonlinear system identification. To the best of our knowledge, neural network training and stability analysis with the ellipsoid or the BE algorithm has not yet been established in the literature, and this is the first paper to successfully apply the BE algorithm for stable training of recurrent neural networks.

In this paper, the BE algorithm is modified to train the weights of a recurrent neural network for nonlinear system identification. Both hidden layers and output layers can be updated. Stability analysis of identification error with the BE algorithm is given by a Lyapunov–like technique.