Gradient Methods of Training and Generalization Ability of a Biological Neural Network Model

Aleksandra Świetlicka, Krzysztof Kolanowski
Poznan University of Technology
aleksandra.swietlicka@put.poznan.pl, krzysztof.kolanowski@put.poznan.pl

Abstract

In this research we focus on the generalization ability of a model of a biological neuron with dendritic structure (which can be understood as a simple biological neural network). Our previous work covers issues like development of the biological neural network model based on Markov kinetic schemes [6] (discretization, implementation and simulation of the model) and introduction of training in this model [4]. We also performed a test of generalization ability of a biological neuron model with a point-like structure in [5]. A natural continuation of this research is to examine the generalization ability of the kinetic model of a biological neural network, while it is one of the most important features of neural networks in general.

Commonly used method of testing the generalization ability of any neural network is by examining the Vapnik-Chervonenkis dimension. A huge disadvantage of this method is a requirement of a sufficiently large training set. Another approach, which is more and more often a subject of research in this area, assumes modifying the error function of a neural network by adding a performance index (regularizer). This performance index can be in one of many forms, e.g. Tikhonov functional [2], penalty function [3] or a square norm of the network curvature [1].

In this work, we examine three different forms of a regularizer to test the generalization ability of the stochastic kinetic model of a biological neural network. Additionally, we consider an improved training procedure, where - besides the gradient descent algorithm - we use conjugate gradient, stochastic gradient and Newton's methods. As an application we use a problem of noise reduction in an image.

References

1. M. Galicki and L. Leistritz and E.B. Zwick and H. Witte. Improving Generalization Capabilities of Dynamic Neural Networks. Neural Computation (2014), vol. 16, no. 6, pp. 1253–1282, doi: 10.1162/089976604773717603.
2. S. Haykin. Neural Networks: A Comprehensive Foundation. second edition, Pearson Education, 1998.
3. A. Krogh and J. Hertz. A simple weight decay can improve generalization. in: Advances in Neural Information Processing Systems 4, Morgan Kaufmann, 1992, pp. 950-957.
4. A. Świetlicka. Trained stochastic model of biological neural network used in image processing task. Applied Mathematics and Computation (2015), vol. 267, pp. 716-726, doi: 10.1016/j.amc.2014.12.082.
5. A. Świetlicka and K. Kolanowski and R. Kapela and M. Galicki and A. Rybarczyk. Investigation of generalization ability of a trained stochastic kinetic model of neuron. Applied Mathematics and Computation (2017), doi: 10.1016/j.amc.2017.01.058.
6. A. Świetlicka and K. Gugała and W. Pedrycz and A. Rybarczyk. Development of the Deterministic and Stochastic Markovian Model of a Dendritic Neuron. Biocybernetics and Biomedical Engineering (2017), vol. 37, issue 1, pp. 201-216, DOI: 10.1016/j.bbe.2016.10.002.