

Efficient Training of RBF Networks Via the BYY Automated Model Selection Learning Algorithms

Jing Wang¹ and Jun Wang² *

¹Department of Mathematics, Tsinghua University, Beijing 100084, P. R. China
²Department of Mathematics, Tsinghua University, Beijing 100084, P. R. China
jwma@math.pku.edu.cn

Abstract. In this paper, we propose a novel algorithm for training radial basis function (RBF) networks. The algorithm is based on the BYY (Bregman-Young-Yang) automated model selection learning algorithm. The algorithm is efficient and robust. It can be applied to training RBF networks for function approximation, pattern classification, and other applications. The algorithm is based on the BYY (Bregman-Young-Yang) automated model selection learning algorithm. The algorithm is efficient and robust. It can be applied to training RBF networks for function approximation, pattern classification, and other applications.

1 Introduction

The radial basis function (RBF) network is a type of artificial neural network. It is used for function approximation, pattern classification, and other applications. The RBF network is composed of three layers: an input layer, a hidden layer, and an output layer. The input layer consists of input nodes. The hidden layer consists of hidden nodes. The output layer consists of output nodes. The RBF network is trained by minimizing the error function. The error function is defined as the sum of the squared errors between the target values and the output values. The RBF network is trained by minimizing the error function. The error function is defined as the sum of the squared errors between the target values and the output values. The RBF network is trained by minimizing the error function. The error function is defined as the sum of the squared errors between the target values and the output values.

*The corresponding author.

$$x_i, y_i = (x_i, y_i) \in \mathcal{X} \times \mathcal{Y} \quad \forall i = 1, \dots, n$$

Let \mathcal{H} be a hypothesis space. A hypothesis $h \in \mathcal{H}$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h that minimizes the expected loss over the joint distribution \mathcal{D} . The loss function $\ell(h, (x, y))$ measures the discrepancy between the predicted output $h(x)$ and the target output y . The expected loss is defined as $\mathbb{E}_{(x, y) \sim \mathcal{D}} [\ell(h, (x, y))]$. The learning algorithm aims to find a hypothesis h that minimizes this expected loss.

Let \mathcal{H}_k be a hypothesis space. A hypothesis $h_k \in \mathcal{H}_k$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h_k(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h_k that minimizes the expected loss over the joint distribution \mathcal{D} .

Let \mathcal{H}_k be a hypothesis space. A hypothesis $h_k \in \mathcal{H}_k$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h_k(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h_k that minimizes the expected loss over the joint distribution \mathcal{D} . The learning algorithm aims to find a hypothesis h_k that minimizes this expected loss.

Let \mathcal{H}_k be a hypothesis space. A hypothesis $h_k \in \mathcal{H}_k$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h_k(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h_k that minimizes the expected loss over the joint distribution \mathcal{D} . The learning algorithm aims to find a hypothesis h_k that minimizes this expected loss.

Let \mathcal{H}_k be a hypothesis space. A hypothesis $h_k \in \mathcal{H}_k$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h_k(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h_k that minimizes the expected loss over the joint distribution \mathcal{D} . The learning algorithm aims to find a hypothesis h_k that minimizes this expected loss.

2 BYY-AMS Adaptive Gradient Learning Algorithm

Let \mathcal{H}_k be a hypothesis space. A hypothesis $h_k \in \mathcal{H}_k$ is a function that maps each input $x \in \mathcal{X}$ to a predicted output $h_k(x) \in \mathcal{Y}$. The goal of the learning algorithm is to find a hypothesis h_k that minimizes the expected loss over the joint distribution \mathcal{D} . The learning algorithm aims to find a hypothesis h_k that minimizes this expected loss.

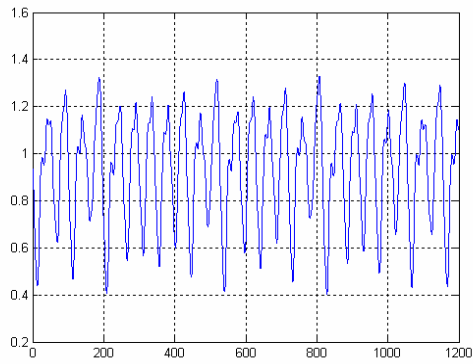


Fig. 3. Time evolution of the temperature $T_{\alpha\alpha}(t)$ for a 1D chain.

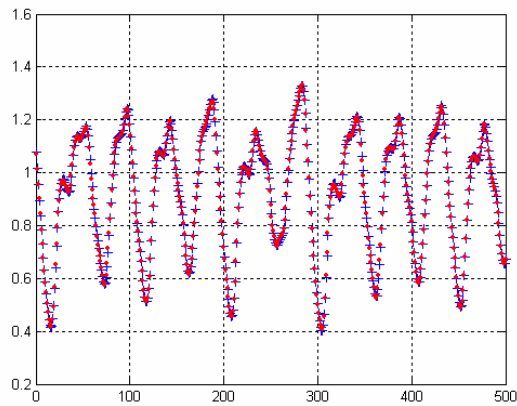


Fig. 4. Time evolution of the temperature $T_{\alpha\alpha}(t)$ for a 1D chain, showing the periodicity of the temperature fluctuations.

5 Conclusions

In this paper, we have presented a new method for calculating the time evolution of the temperature $T_{\alpha\alpha}(t)$ for a 1D chain, showing the periodicity of the temperature fluctuations.

