International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Boundedness of a Batch Gradient Method with Smoothing $L_{1/2}$ Regularization for Pi-sigma Neural Networks

Kh. Sh. Mohamed[1,2], Y. Sh. Mohammed[3,4,5], Abd Elmoniem A. Elzain[3,6], Mohamed El-Hafiz M. N. [3,7], and Elnoor. A. A. Noh[3,8]

[1] School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, PR China

[2] Mathematical Department, College of Science, Dalanj University, Dalanj, Sudan

[3] Department of Physics, College of Science & Art, Qassim University, Oklat Al- Skoor, P.O.Box: 111, Saudi Arabia.

[4] Physics Department, College of Education, Dalanj University, Dalanj, Sudan

[5] Africa City for Technology, Khartoum, Sudan

[6] Department of Physics, University of Kassala, Kassala, P.O.Box: 266, Sudan

[7] Mathematical Department, College of Education, Zalingy University, Zalingy, Sudan

[8] Department of Chemistry, College of Art & Science, Albaha University, Baljurashi, Saudi Arabia.

Abstract: ***This paper considers a batch gradient method with $L_{1/2}$ regularization for Pi –sigma neural networks. In origin, by introducing an $L_{1/2}$ regularization term involves absolute value and is not differentiable into the error function. A key point of this paper, specifically, the smoothing $L_{1/2}$ regularization is a term proportional to the norm of the weights. The role of the smoothing $L_{1/2}$ regularization term is to control the magnitude of the weights and to improve the generalization performance of the networks. The weights are proved to be bounded during the training process, thus the conditions that are required for convergence analysis of batch gradient method in literature are simplified.***

Keywords: ***Batch gradient method, Pi-sigma neural network, L½ regularization, Boundedness.***

1. Introduction

In fact, higher order neural networks (HONN) have been widely applied in many applications such as intention to enhance the nonlinear descriptive capacity of the feed forward multilayer perceptron networks [1 - 6]. Pi-sigma neural network (PSNN) is a class of higher-order feed forward polynomial neural network and is known to provide inherently more powerful mapping abilities than traditional feed forward neural networks. The (PSNN) modules are widely used for pattern classification and approximation problems [7 - 9]. By adding a penalty term of the error function [10 - 16]. The penalty has become a common practice to make the network weights keeping bounded during the training process. The boundedness of the weights is an obvious fact when a convergence training method ‘’ such as the quadratic programming used in support vector machines in [17] is used to minimize the cost function with penalty term. When using online gradient method to minimize the cost of error function with regularization penalty term, the boundedness of weights is not obvious because the decrease of the cost function and convergence of the method during the learning process are usually obtained by first condition the network weights are bounded. Recently, most of the studies have been focused on the $L_{1\backslash 2}$ regularization penalty term adding to the error function usual is not smooth at the origin, which causes difficulty in the convergence analysis to speed this drawback, we use the modified $L_{1\backslash 2}$ regularization term is proposed by the usual one at the origin [18, 19], in [18] the $L_{1\backslash 2}$ regularization term is introduced into the batch gradient learning algorithm for the pruning of FNN. Some convergence analyses of the online gradient method (OGM for short) with fixed order inputs (OGM_F) and with special stochastic inputs (OGM_SS) for PSNN were respectively presented in [20,21]. Especially, convergence analysis of the online gradient learning algorithm with $L_{1\backslash 2}$ regularization term for the pruning of FNN [19]. However, in [19 - 21] obtain both the weak and strong convergence results. The main purpose of this paper , in doing so, by prove that the weights are indeed bounded deterministically in the batch gradient learning algorithm process by adding a smoothing $L_{1\backslash 2}$ regularization, a term proportional to the norm of the weights. That the weights of the network will keep bounded in the training process.

The rest of this paper is organized as follows. The network model and the batch gradient method with smoothing $L_{1\backslash 2}$ regularization are described in the section. The convergence of this algorithm is discussed and a convergence theorem is established in section 3. and conclusion in section 4.

PSNN and Randomized Batch gradient with smoothing $L_{1\backslash 2}$ regularization

Structure of PSNN:

Consider a three- layer network consisting of P input node, N hidden nodes, and 1-output nodes. Suppose that by $\omega_k = (\omega_{k1}, \dots, \omega_{kp})^T \in R^p$ be the weight vector between the input units and the hidden unit$(k = 1,2,\dots,N)$. $\xi^j = (\xi_1^j, \xi_2^j, \dots \xi_p^j) \in \mathbb{R}^p$, stands for input vector. To simplify the presentation, we write all the weight parameters in a compact form $= (\omega_1^T, \dots, \omega_N^T) \in R^{NP}$. The weights on the connections between the product node and the summation node are fixed to one. We have included a special input unit $\xi_P = -1$, corresponding to the biases ω_{np} . Let $g: \mathbb{R} \to \mathbb{R}$ be a transfer function for the hidden and output node, which is typically, but not necessarily, a sigmoid function. For any given input ξ and weight ω , the output of the network is

$$y = g\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right) \quad (1)$$

Randomized Batch gradient method for PSNN

In general, the batch gradient method is a simple and efficient learning method for feed-forward neural networks. Usually PSNN and the networks with pi-sigma building blocks are also trained by it but with randomized modification.

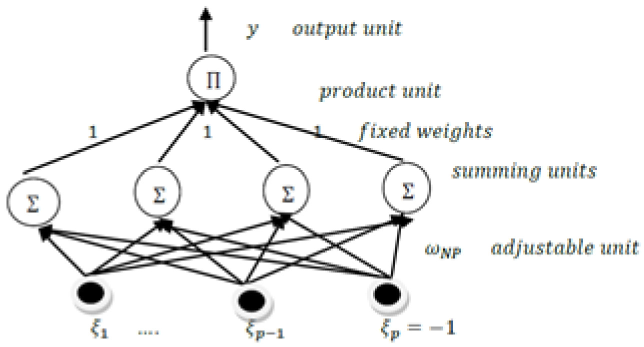

Figure 1 : A pi − sigms neural network

Let $\{x^j, o^j\}_{j=1}^{J} \subset R^p \times R$ is a given of training samples. The usual mean square error function for the network is

$$\hat{E}(\omega) = \frac{1}{2}\sum_{j=1}^{J}\left(o^j - g\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right)\right)^2 = \sum_{j=1}^{J} g_j\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right) \quad (2)$$

Where $g_j(t) = \frac{1}{2}\left(o^j - g(t)\right)^2 \quad (1 \le j \le J, t \in \mathbb{R})$.

Batch gradient with $L_{1/2}$regularization (BG $L_{½}$)

We denote the error function with $L_{1/2}$ regularization penalty term is

$$E(\omega) = \frac{1}{2}\sum_{j=1}^{J}\left(o^j - g\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right)\right)^2 + \lambda\sum_{k=1}^{N}|\omega_k|^{1/2}$$
$$= \sum_{j=1}^{J} g_j\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right) + \lambda\sum_{k=1}^{N}|\omega_k|^{1/2} \quad (3)$$

The gradient of error function with $L_{1/2}$ regularization respect to the weight vector ω_k $(k = 1,2,\dots N)$ is

$$E_{\omega_k}(\omega) = \sum_{j=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i.\xi^j)\,\xi^j + \frac{\lambda\, sgn\,(\omega_k)}{2\,|\omega_k|^{1/2}} \quad (4)$$

Starting from an arbitrary initial weightW^0, the batch gradient method with $L_{1\backslash 2}$ regularization update the weights iteratively by:

$$\omega_k^{mJ+j} = \omega_k^{mJ+j-1} - \eta_m \Delta_j\, \omega_k^{mJ+j-1}, m = 0,1,2,\dots \quad (5)$$

and

$$\Delta_j\,\omega_k^{mJ+j-1} = g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\,\xi^j + \frac{\lambda\, sgn\,(\omega_k^{mJ+j-1})}{2\,J\,|\omega_k^{mJ+j-1}|^{1/2}} \quad (6)$$

Where $k = 1,2,,\dots N$; and $\eta_m > 0$ represents the learning rate.

Smoothing $L_{1\backslash 2}$ regularization (BGS$L_{1\backslash 2}$)

A modified $L_{1\backslash 2}$ regularization term is proposed by smoothing the usual one at the origin, resulting in the following error function with a smoothing $L_{1\backslash 2}$ regularization penalty term:

$$E(\omega) = \sum_{\ell=1}^{J} g_\ell\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right) + +\lambda\sum_{k=1}^{N} f(\omega_k)^{1/2} \quad (7)$$

Where $f(x)$ is a smooth function that approximates $|x|$. for definiteness and simplicity, we choose $f(x)$ as a piecewise polynomial function:

$$f(x) = \begin{cases} |x| \; if \; |x| \ge a \\ -\frac{1}{8a^3}x^4 + \frac{3}{4a}x^2 + \frac{3}{8}, if \; |x| < a \end{cases} \quad (8)$$

Where a is a small positive constant. Then it is easy to get

$$f(x) \in \left[\frac{3}{8}a, +\infty\right), f'^{(x)} \in [-1,1], f''^{(x)} \in \left[0, \frac{3}{2a}\right] \quad (9)$$

The gradient of the error function can be written as (4) with

$$E_{\omega_k}(\omega) = \sum_{j=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i.\xi^j)\,\xi^j + \frac{\lambda\, f'(\omega_{nk})}{2J\, f\,(\omega_k)^{1/2}} \quad (10)$$

Where $\lambda > 0$ is a penalty parameter and $k = 1,2,3,\dots N$.

Starting from an arbitrary initial weightW^0, the batch gradient method with $L_{1\backslash 2}$ regularization update the weights iteratively by

$$\omega_k^{mJ+j} = \omega_k^{mJ+j-1} - \eta_m \Delta_j\, \omega_k^{mJ+j-1}, m = 0,1,2,\dots \quad (11)$$

and

$$\Delta_j\,\omega_k^{mJ+j-1} = g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\,\xi^j + \frac{\lambda\, f'(\omega_k^m)}{2\,J\, f(\omega_k^m)^{1/2}} \quad (12)$$

Where $k = 1,2,3,\dots N$; and $\eta_m > 0$ represents the learning rate

2. Main Results

Suppose that K is any positive integer and consider the Euclidean space $\mathbb{R}^K$. For $x = (x_1, x_2, \dots, x_K)^T$ and $y = (y_1, y_2, \dots, y_K)^T$, we define $x.y = \sum_{i=1}^{K} x_i y_i$, and $\|x\|^2 = (x\,.x)^2$, the following assumptions in this paper are described below:

Assumption (A1)

$|g_j(t)|, |g_j'(t)|, |g_j''(t)|$ $(j = 1,2,\dots,J)$ are uniformly bounded for $t \in \mathbb{R}$.

Assumption (A2)

$0 < \eta_m < 1$, i.e., $\sum_{m=0}^{\infty}\eta_m < \infty$,

Assumption (A3)

η and λ are chosen to satisfy $0 < \eta < \frac{1}{\lambda + C_1}$, where

$$C_1 = \frac{\lambda M}{J}(1 + C_3\eta_0)^2 + C_4 N + \frac{1}{2}CC_5^2(1 + NC_3\eta_m)^2 + \frac{1}{2}JC^{(N+1)}(N-1)C_4^2$$
$$C_2 = max\{C_7 + C_7 C_{2,j-1}\, N(j-1)\eta_0\, , C_{2,j-1}\} \quad (13)$$

Assumption (A4)

The set $\Omega_0 \in \{w \in \Omega: E_w(w) = 0\}$ Contains finite points, where Ω is closed bounded region such that $\{\omega^m\} \subset \Omega$.

Theorem 3.1 (boundedness Theorem).

Suppose that the weight sequence $\{\omega^m\}$ is generated by the algorithm (11) for any initial valueω^0, that (A1) is valid, and then $\{\omega^m\}$ is uniformly bounded.

Theorem 3.2 (convergence Theorem).

Suppose that the error function is given by (7), that the weight sequence $\{\omega^m\}$ is generated by the algorithm (11) for any initial valueω^0, and Assumption (A1) is valid.
Then we have

(a) $E(\omega^{(m+1)J}) \le E(\omega^{mJ})$,
(b) There is $E^* \ge 0$ such that $\lim_{m\to\infty} E(\omega^{mJ}) = E^*$;
(c) $\lim_{m\to\infty} \|\Delta_j^m \omega_i^{mJ}\| = 0$, $\lim_{m\to\infty} \|E_\omega(\omega^{mJ})\| = 0$.

Moreover, if Assumption (A4) is valid, then we have the strong convergence:

(d) There exists $\omega^* \in \Omega_0$ such that $\lim_{m\to\infty} \omega^m = \omega^*$.

Proofs

The next two lemmas will be used to prove our convergence result. Their proofs are omitted since they are quite similar to those of *lemma 3.5* in [22] and Theorem 3.5.10 in [23], respectively.

Lemma 4.1

Suppose that the learning rate η_m satisfies (A2) and that the sequence $\{a_m\}(m \in \mathbb{N})$ satisfies $a_m \ge 0$ $\sum_{m=0}^{\infty} \eta_m a_m^\beta < \infty$ and $|a_{m+1} - a_m| \le \mu\eta_m$ for some constants β *and* μ . Then we have $\lim_{m\to\infty} a_m = 0$.

Lemma 4.2

Let $F: \Phi \subset R^p \to \mathrm{R}$ $(\mathrm{p} \ge 1)$ be continuous for a bounded closed region Φ. if the set $\Phi_0 = \{x \in \Phi: \mathrm{F_x(x)} = 0\}$ has finite points and the sequence $\{x_n\} \in \Phi$ satisfy: $\lim_{n\to\infty} \|F_x(x_n)\| = 0$ and $\lim_{n\to\infty} \|x_{n-1} - x_n\| = 0$.
Then, there exists $x^* \in \Phi_0$ such that $\lim_{n\to\infty} x_n = x^*$
In this work, by choosing an initial $\eta_0 \in (0, 1]$ and positive constantβ, we inductively, determine η_m in (16) by (cf. [22])

$$\frac{1}{\eta_{m+1}} = \frac{1}{\eta_m} + \beta\ , m = 0,1,2 \ldots\ldots (14)$$

First, we define $r_k^{m,j} = \Delta_j^m \omega_k^{mJ+j-1} - \Delta_j^m \omega_k^{mJ}$

$$r_k^{m,j} = 0\ , 1 \le k \le N, m = 0,1,2, \ldots\ (15)$$

and $d_k^{m,j} = \omega_k^{mJ+j-1} - \omega_k^{mJ}$ (16)
Then, we have

$$d_k^{m,j} = \eta_m \sum_{t=1}^{j} \Delta_t^m \omega_k^{mJ+j} = \eta_m \sum_{t=1}^{j} (\Delta_t^m \omega_k^{mJ} + r_k^{t,m})\ (17)$$

$$1 \le j \le J, 1 \le k \le N, m = 0,1,2 \ldots$$

Then, by the error function (7), we have

$$E(\omega^{(m+1)J}) = \sum_{j=1}^{J} g_j\left(\prod_{i=1}^{N}\left(\omega_i^{(m+1)J}.\xi^j\right)\right) + \lambda \sum_{k=1}^{N} f(\omega_k^{(m+1)J})^{\frac{1}{2}}\ (18)$$

$$E(\omega^{mJ}) = \sum_{j=1}^{J} g_j\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right) + \lambda \sum_{k=1}^{N} f(\omega_k^{mJ})^{\frac{1}{2}}\ (19)$$

Lemma 4.3

Let t $\{\eta_m\}$ *be given by* (15)*.There hold*

$$0 < \eta_m < \eta_{m+1} \le 1, m = 1,2, \ldots\ldots\ (20)$$

$$\frac{\tau}{m} < \eta_m < \frac{\rho}{m}, \tau = \frac{\eta_0}{1+\eta_0\beta}, \rho = \frac{1}{\beta}, m = 1,2, \ldots\ldots (21)$$

Proof.

This lemma is easy to validate by virtue of (15) and $\eta_0 \in (0, 1]$, see Lemma 4 in [24] and Lemma 2.1. in [25]

Lemma 4.*4*

Suppose that (A1) and (A2) are satisfied. Then, there exists constants$C_2, C_3, C_4 > 0$, such that for any $m = 0,1, \ldots$

$$\|r_k^{j,m}\| \le C_2 \sum_{i=1}^{N}\sum_{t=1}^{j-1} \|\Delta_t^m \omega_i^{mJ}\|\ , 2 \le j \le J, 1 \le k \le N\ (22)$$

$$\sum_{t=1}^{j} \|r_k^{j,m}\| \le C_3 \sum_{i=1}^{N}\sum_{t=1}^{j} \|\Delta_t^m \omega_i^{mJ}\|\ , 1 \le j \le J, 1 \le k \le N\ (23)$$

$$\|d_k^{m,j}\| \le C_4 \sum_{i=1}^{N}\sum_{t=1}^{j} \|\Delta_t^m \omega_i^{mJ}\|\ ,1 \le j \le J, 1 \le k \le N\ (24)$$

Proof.

By Assumption (A2), (18) and Cauchy- Schwartz inequality, we have

$$\left|\prod_{i=1}^{N}(\omega_i^{mJ+j}.\xi^j) - \prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right|$$

$$\le \left|\prod_{i=1}^{N-1}(\omega_i^{mJ+j}.\xi^j)\right| |(\omega_N^{mJ+j} - \omega_N^{mJ})\xi^j|$$

$$+ \left|\prod_{i=1}^{N-2}(\omega_i^{mJ+j}.\xi^j)(\omega_N^{mJ}.\xi^j)\right| |(\omega_{N-1}^{mJ+j} - \omega_{N-1}^{mJ})\xi^j| \ldots$$

$$+ \left|\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right| |(\omega_1^{mJ+j} - \omega_1^{mJ})\xi^j|$$

$$\le C_5\left(\sum_{i=1}^{N}\sum_{t=1}^{j} \|\Delta_t^m \omega_i^{mJ}\| + \sum_{i=1}^{N}\sum_{t=1}^{j} \|r_i^{t,m}\|\right)\ (25)$$

Where $C_5 = C^N$ $(1 \le j \le J, m = 0,1,2, \ldots)$
Similarly, easy to get

$$\left|\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ+j}.\xi^j) - \prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ}.\xi^j)\right| \le C_6\left(\sum_{i=1}^{N}\sum_{t=1}^{j} \|\Delta_t^m \omega_i^{mJ}\| + \sum_{i=1}^{N}\sum_{t=1}^{j} \|r_i^{t,m}\|\right)\ (26)$$

Where $C_6 = C^{N-1}$ $(1 \le j \le J, m = 0,1,2, \ldots)$
By Assumption (A1), (A2), (12), (16), (26), (27) and differential mean value theorem, for $1 \le j \le J, 1 \le k \le N, m = 0,1,2, \ldots$, we have

$$\|r_k^{j,m}\| = \left\| \eta_m g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\xi^j - \eta_m g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ}.\xi^j)\xi^j + \frac{\lambda}{2J}\left(\frac{f'(\omega_k^{mJ+j-1})}{f(\omega_k^{mJ+j-1})^{1/2}} - \frac{f'(\omega_k^m)}{f(\omega_k^m)^{1/2}}\right)\right\|$$

$$\le \left\|\eta_m g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\left(\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j) - \prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ}.\xi^j\right)\xi^j\right\|$$

$$+ \left\|\eta_m g_j''(t_{j,m})\left(\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j) - \prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\xi^j\right\| + \frac{\lambda}{2J}\eta_m F''(t_{k,j})|d_k^{m,j}|$$

$$\leq C_7\eta_m\left(\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|+\sum_{i=1}^{N}\sum_{t=1}^{j}\|r_i^{t,m}\|\right) \quad (27)$$

Where $t_{i,m}\in\mathbb{R}$ is on the line segment between $\omega_i^m.\xi^j$ and $\omega_i^{m+1}.\xi^j$ and $C_7=C_6C^2+C_5C^{N+1}+\lambda M/J$.

By mathematical induction to prove the following formula

$$\|r_k^{j,m}\|\leq C_{2,j}\eta_m\sum_{i=1}^{N}\sum_{j=1}^{j-1}\|\Delta_t^m\omega_i^{mJ}\|, 2\leq j\leq J, 1\leq k\leq N, m=0,1,2,\ldots (28)$$

Where $C_{2,j}$ constant.

By (16) and (28), for $j=2$ the (29) is clearly established. Suppose that $j<J\ (2<j\leq J)$,(27) establish. Then, proof for (29) also founded. By (28) and (21), we have

$$\|r_k^{j,m}\|\leq C_7\eta_m\sum_{i=1}^{N}\sum_{j=1}^{J-1}\|\Delta_t^m\omega_i^{mJ}\|+C_7C_{2,j-1}\eta_m^2\sum_{i=1}^{N}\sum_{t=1}^{J-1}\sum_{i_1=1}^{N}\sum_{t_1=1}^{J-1}\|\Delta_{t_1}\omega_{i_1}^{mJ}\|$$

$$\leq C_{2,j}\,\eta_m\sum_{i=1}^{N}\sum_{j=1}^{J-1}\|\Delta_t^m\omega_i^{mJ}\|\quad 1\leq k\leq N, m=0,1,2,\ldots$$

Where $C_{2,j}=max\{C_7+C_7C_{2,j-1}N(j-1)\eta_0, C_{2,j-1}\}$.

Therefore, $j=J$, (29) established. By the mathematical induction for $2\leq j\leq J$, then (29) it is also establish. Suppose $C_2=C_{2,J}$ in (23) easily to get (29).

Next, by (16) and (23), we have

$$\sum_{t=1}^{j}\|r_k^{j,m}\|=\sum_{t=2}^{j}\|r_k^{j,m}\|\leq\sum_{t=2}^{j}C_3\eta_m\sum_{i=1}^{N}\sum_{t_1=1}^{t-1}\|\Delta_t^m\omega_i^{mJ}\|$$

$$\leq C_2(j-1)\eta_m\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|$$

$$\leq C_3\eta_m\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\| \quad (29)$$

Where $C_3=C_2(J-1)$ and $2\leq j\leq J, 1\leq k\leq N, m=0,1,2,\ldots$

Finally, (24) established on the basis of proof (25). By Lemma 4.3 for $2\leq j\leq J, 1\leq k\leq N, m=0,1,2,\ldots$,

We have

$$\|\omega_k^{mJ+j}-\omega_k^{mJ}\|\leq\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|+\sum_{t=1}^{j}\|r_i^{t,m}\|$$

$$\leq\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|+C_3\eta_m\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|$$

$$=(1+C_3\eta_m)\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\|$$

$$\leq C_4\sum_{i=1}^{N}\sum_{t=1}^{j}\|\Delta_t^m\omega_i^{mJ}\| \quad (30)$$

Where $C_4=1+C_3\eta_0$, the proof it is completed

Proof (*of Theorem* 1).

See **[19],** and By the Assumption (A2), i.e., $\sum_{m=0}^{\infty}\eta_m<\infty$, we can easily get that the sequence $S_m=\eta_0+\eta_1+\cdots+\eta_{m-1}$ is convergence sequence. By the Cauchy's test for convergence, for $\forall\mathcal{E}>0$, there exists a positive integer $N_1\in\mathbb{N}$, for$\forall m>N_1$, $\forall p\in\mathbb{N}$, we have

$$|S_{m+p}-S_m|=\eta_m+\eta_{m+1}+\cdots+\eta_{m+p-1}<\varepsilon$$

$$|S_{m+p+1}-S_{m+1}|=\eta_{m+1}+\eta_{m+2}+\cdots+\eta_{m+p}<\varepsilon$$

By (11), (12) and Assumption (A2) result in

$$|\omega_k^{mJ+j-1}-\omega_k^{mJ}|=\eta_m|\Delta_j^m\omega_k^{mJ}|$$

$$\leq\eta_m\left(\sum_{j=1}^{J}g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\xi^j\right) \quad (31)$$

By Assumption (A1), there is a constant $C_6>0$ such that all $(m\in\mathbb{N}.; j=1,2,..,J)$

$$C_7=sup\left\{\sum_{j=1}^{J}g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\|\xi^j\|\right\} \quad (32)$$

In addition, for all $f(x)\in\left[\frac{3}{8}a,+\infty\right), f'^{(x)}\in[-1,1]$ holds.

By the updating (31), we have

$$|\omega_k^{mJ+j}-\omega_k^{mJ+j-1}|\leq\eta_m\left(\left|\sum_{j=1}^{J}g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ+j-1}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j-1}.\xi^j)\xi^j+\frac{\lambda f'(\omega_k^{mJ+j-1})}{2J\,f(\omega_k^{mJ+j-1})^{1/2}}\right|\right)$$

$$\leq\eta_m\left(C_7+\frac{\lambda}{3a}\sqrt{6a}\right)\leq\eta_m C_6 \quad (33)$$

Where $C_6=C_7+(\lambda/3a)\sqrt{6\,a}$. Then

$$\begin{aligned}|\omega_k^{(m+1)J+j}-\omega_k^{mJ+j}|&\leq|\omega_k^{(m+1)J+j}-\omega_k^{(m+1)J+j-1}|\\&+|\omega_k^{(m+1)J+j-1}-\omega_k^{(m+1)J+j-2}|+\cdots\\&+|\omega_k^{(m+1)J+1}-\omega_k^{(m+1)J}|+|\omega_k^{mJ+J}-\omega_k^{mJ+j-1}|\\&+|\omega_k^{mJ+j-1}-\omega_k^{mJ+j-2}|+\cdots+|\omega_k^{mJ+j+1}-\omega_k^{mJ+j}|\\&\leq(j\eta_{m+1}+(J-j)\eta_m)C_6 \quad (34)\end{aligned}$$

Since

$$\begin{aligned}|\omega_k^{(m+p)J+j}-\omega_k^{mJ+j}|&\leq|\omega_k^{(m+p)J+j}-\omega_k^{(m+p-1)J+j}|\\&+|\omega_k^{(m+p-1)J+j}-\omega_k^{(m+p-2)J+j}|+\cdots\\&+|\omega_k^{(m+1)J+j}-\omega_k^{mJ+j}|\\&\leq C_6j(\eta_{m+p}+\eta_{m+p-1}+\cdots.+\eta_{m+1})\\&+C_2(J-j)(\eta_{m+p-1}+\eta_{m+p-2}+\cdots.+\eta_m)\\&\leq JC_6\varepsilon \quad (35)\end{aligned}$$

Therefore, the weight sequence $\{\omega_k^{mJ+j}\}$ is a convergence sequence.

By the properties of convergence sequence, $\{\omega_k^{mJ+j}\}$ must be a bounded sequence, so we get $\|\omega_k^{mJ+j}\|$ $(m=0,1,\ldots,k=$ *1,2,..,P, j=1,2,..,J*) is also bounded. Then we obtain the uniform boundedness of the weight sequence $\{\omega_k^{mJ+j}\}$. Namely, there exists a constant $M>0$ such that

$$\|\omega_k^{mJ+j}\|\leq M,, m=0,1,2,\ldots; 1\leq j\leq J; 1\leq k\leq P \quad (36)$$

Naturally, there also exists a constant $\bar{M}>0$ such that

$$\|\Delta_k^m\omega_k^{mJ+k}\|\leq\bar{M},\quad k=1,2,\ldots\ldots,J. \quad (37)$$

This proof is completed.

Proof (*of Theorem* 2).

Using Taylor expansion to first and second orders, we have

$$\prod_{i=1}^{N}(\omega_i^{(m+1)J}.\xi^j)=\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)+\sum_{k=1}^{N}\left(\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{(m+1)J}.\xi^j)\left(\omega_k^{(m+1)J}-\omega_k^{mJ}\right)\xi^j\right)$$

$$+\frac{1}{2}\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{N}\left(\prod_{\substack{i=1\\i\neq k_1,k_2}}^{N}t_{i,m,j}\right)[(\omega_{k_1}^{(m+1)J}-\omega_{k_1}^{mJ})][(\omega_{k_2}^{(m+1)J}-\omega_{k_2}^{mJ})]\,\xi^j \quad (38)$$

Where $t_{i,m,j}\in\mathbb{R}$ is on the line segment between $\omega_i^m.\xi^j$ and $\omega_i^{m+1}.\xi^j$. Again applying the Taylor expansion and noting (11) and (38), we have

$$g_j\left(\prod_{i=1}^{N}(\omega_i^{mJ+j}.\xi^j)\right)=g_j\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)$$

$$+g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\sum_{k=1}^{N}\left(\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ+j}.\xi^\ell)(\omega_k^{mJ+j}-\omega_k^{mJ})\xi^j\right)$$

$$+\frac{1}{2}\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{N}\left(\prod_{\substack{i=1\\i\neq k_1,k_2}}^{N}t_{i,m,j}\right)[(\omega_{k_1}^{mJ+j}-\omega_{k_1}^{mJ})][(\omega_{k_2}^{mJ+j}-\omega_{k_2}^{mJ})]\,\xi^j$$

$$+\frac{1}{2}g_j''(t_{i,m})\left(\prod_{i=1}^{N}(\omega_i^{mJ+j}.\xi^j)-\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)^2 \quad (39)$$

Where $t_{i,m} \in \mathbb{R}$ is on the line segment between $\omega_i^m.\xi^j$ and $\omega_i^{m+1}.\xi^j$, by combination (7), (11), and (12) and (39), we have

$$E(\omega^{(m+1)J}) - E(\omega^{mJ}) \leq -\frac{1}{\eta_m}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m \omega_i^{mJ}\|^2 + \frac{\lambda}{2J}\sum_{k=1}^{N}\sum_{\ell=1}^{J}\left(\frac{f'(\omega_k^{mJ})}{f(\omega_k^{mJ})^{1/2}} + F''(t_{n,k,m})d_k^{m,j}\right)(\omega_k^{(m+1)J} - \omega_k^{mJ})$$

$+\delta_1 + \delta_2 + \delta_3$ (40)

Where

$$\delta_1 = \sum_{\ell=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\sum_{k=1}^{N}\left(\prod_{\substack{i=1\\ i\neq n}}^{N}(\omega_i^{mJ}.\xi^j)(\omega_k^{(m+1)J} - \omega_k^{mJ})\xi^j\right)$$

$$\delta_2 = \sum_{\ell=1}^{J}\frac{1}{2}g_j''(t_{i,m}')\left(\prod_{i=1}^{N}(\omega_i^{(m+1)J}.\xi^j) - \prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)^2$$

$$\delta_3 = \frac{1}{2}\sum_{\ell=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\sum_{\substack{k_1,k_2=1\\ k_1\neq k_2}}^{N}\left(\prod_{\substack{i=1\\ i\neq k_1,k_2}}^{N} t_{i,m,j}\right)[(\omega_{k_1}^{(m+1)J} - \omega_{k_1}^{mJ})][(\omega_{k_2}^{(m+1)J} - \omega_{k_2}^{mJ})]\,\xi^j$$

Where $t_{i,m}'$ *and* $t_{n,k,m,j}$ lies in between $\omega_i^{mJ}.x^j$ and $\omega_i^{(m+1)J}.x^j$, and from (23), (24) and (45), $M = \frac{\sqrt{6}}{\sqrt{a^3}}$, *and* $F(x) \equiv (f(x))^{\frac{1}{2}}$. Note that

$$F'(x) = \frac{f'(x)}{2\sqrt{f(x)}}$$

$$F''(x) = \frac{2f''(x).f(x) - [f'(x)]^2}{4[f(x)]^{\frac{3}{2}}} \leq \frac{f''(x)}{2\sqrt{f(x)}} \leq \frac{\sqrt{6}}{2\sqrt{a^3}}$$

By *(25), (30) and Lemma 4.3 for* $1 \leq j \leq J, 1 \leq k \leq N, m = 0,1,2,\ldots$, *and Cauchy- Schwartz Theorem, we have*

$$\frac{\lambda}{2J}F''(t_{n,k,m})\left|\left(\sum_{i=1}^{N}\sum_{j=1}^{J}(d_k^{m,j}).(\omega_k^{(m+1)J} - \omega_k^{mJ})\right)\right| \leq \lambda M\sum_{i=1}^{N}\sum_{j=1}^{J}|d_k^{m,j}|_\circ \|\omega_k^{(m+1)J} - \omega_k^{mJ}\|$$

$$\leq \frac{\lambda M}{J}C_4^2\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \leq C_{10}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (41)$$

Where $C_{10} = \lambda M(1 + C_3\eta_0)^2/J$ and $t_{n,k,m}$ lies in between $\omega_i^{mJ}.x^j$ and $\omega_i^{(m+1)J}.x^j$

By Assumption (A1), (A2), (12) and (25), we have

$$|\delta_1| \leq \frac{1}{\eta_m}\left|\sum_{i=1}^{N}\sum_{j=1}^{J}\left(g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\prod_{\substack{i=1\\ i\neq k}}^{N}(\omega_i^{mJ}.\xi^j)\xi^j + \frac{\lambda}{2J}\frac{f'(\omega_k^{mJ})}{f(\omega_k^{mJ})^{1/2}}\right)\right|\|\omega_k^{(m+1)J} - \omega_k^{mJ}\|$$

$$\leq C_{11}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (42)$$

Where $C_{11} = C_4NJ$.

By Assumption (A1), (21), (24), and (26) for $m = 0,1,2,\ldots$, we have

$$|\delta_2| \leq \frac{1}{2}CC_5^2(1 + NC_3\eta_m)^2\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \leq C_{12}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (43)$$

Where $C_{12} = \frac{1}{2}CC_5^2(1 + NC_3\eta_m)^2$.

Using Assumption (A1), (A2), (25) and Cauchy- Schwartz Theorem, we get

$$|\delta_2| \leq \frac{1}{2}C^{N-1}\left|\sum_{j=1}^{J}\sum_{\substack{k_1,k_2=1\\ k_1\neq k_2}}^{N}\left((\omega_{k_1}^{(m+1)J} - \omega_{k_1}^{mJ})(\omega_{k_2}^{(m+1)J} - \omega_{k_2}^{mJ})\xi^j\right)\right|$$

$$\leq \frac{1}{2}JC^{N+1}\sum_{\ell=1}^{J}\sum_{\substack{k_1,k_2=1\\ k_1\neq k_2}}^{N}|d_{k_1}^{m,j}|.|d_{k_2}^{m,j}|$$

$$\leq \frac{1}{2}JC^{(N+1)}(N-1)C_4^2\sum_{k=1}^{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|\right)^2$$

$$\leq C_{13}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2, m = 0,1,2,\ldots \quad (44)$$

Where $C_{13} = JC^{(N+1)}(N-1)C_4^2/2$.

Substituting (41) - (44) into (40), then, we have

$$E(\omega^{(m+1)J}) - E(\omega^{mJ}) \leq \left(-\frac{1}{\eta_m} + C_{10} + C_{11} + C_{12} + C_{13}\right)\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \leq -\left(\frac{1}{\eta_m} - C_1\right)\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2$$

$\leq 0.$ (45)

This completes the proof to statement (i) of theorem 3.2.

Proof to (ii) of theorem 3.2.

From the conclusion of (i), we know that the nonnegative sequence $\{E(W^m)\}$ is monotone. However, it is also bounded below. Hence there must exist $E^* \geq 0$ such that $\lim_{k\to\infty} E(W^m) = E^*$. The proof to (ii) it thus completed.

Proof to (iii) of theorem 3.2.

It is follows from Assumption (A4) that $\beta > 0$. Taking $\beta = \frac{1}{\eta_m} - C_1$ and using (45), we suppose that M is positive integer, we have

$$E(W^{(M+1J)}) \leq E(W^{MJ}) - \beta\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{MJ}\|^2 \leq \cdots \leq E(W^0) - \beta\sum_{m=0}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2.$$

Since $E(W^{m+1}) \geq 0$, we have

$$\beta\sum_{m=0}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \leq E(\omega^0) \leq \infty.$$

Let $M \to \infty$, then

$$\sum_{m=0}^{\infty}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 < \frac{1}{\beta}E(W^0) < \infty.$$

Thus results in

$$\lim_{m\to\infty}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 = 0.$$

From (10) - (12) and (A1) it is easily get

$$\lim_{m\to\infty}\|\Delta_j^m\omega_i^{mJ}\| = 0, \lim_{m\to\infty}\|E_{\omega_k}(\omega^{mJ})\| = 0 \quad (46)$$

The proof to (iii) is thus completed.

Proof to (iv) of theorem 3.2.

Note that the error function $E(W)$ defined in (7) is continuous and differentiable. According to (46), (A5) and

Lemma 4.2, we can easily get the desired results, i.e., there exists a point $\omega^* \in \Omega_0$ such that

$$\lim_{m\to\infty} (\omega_i^{mJ}) = \omega_i^*$$

This completes the proof to (iv)

3. Conclusions

In this paper, we investigate a Batch Gradient Method with Smoothing $L_{1/2}$ Regularization for Pi-sigma Neural Networks. The Smoothing $L_{1/2}$ Regularization is a term proportional to the magnitude of the weights. We prove under moderate conditions that the weights of the networks are keeping bounded in the learning process. The both weak and convergence results require the boundedness of the weights is precondition.

4. Acknowledgment

We gratefully acknowledge Dalanj University and Dalian University of Technology for supporting this research. Special thanks to Prof. Dr. Wei Wu and Dr. Yan Liu for their kind helps during the period of the study.

References

[1] M. Sinha, K. Kumar, and P. K. Kalra, "Some new neural network architectures with improved learning schemes", Soft computing, 4, pp. 214 – 223, 2000.
[2] P. Estevez, Y. Okabe, "Training the piecewise linear-high order neural network through error back propagation", International Joint Conference on Neural Networks, 1, pp. 711 -716, 1991.
[3] Y. Shin, J. Ghosh, "The pi-sigma network: an efficient higher-order neural network for pattern classification and function approximation", International Joint Conference on Neural Networks, 1, pp. 13 -18, 1991.
[4] A. J. Hussain, P. Liatsis, "A new recurrent polynomial neural network for predictive image coding", Image Processing and Its Applications, 1, pp. 82- 86, 1999.
[5] L. Villalobos, F. L. Merat, "Learning capability assessment and feature space optimization for higher-order neural networks", IEEE Transaction on Neural Networks,, 6, pp. 267 – 272, 1995.
[6] Y. Shin, J. Ghosh, D. Samani, "Computationally efficient invariant pattern classification with higher-order pi-sigma networks", Intelligent Engineering Systems through Artificial Neural Networks, 2, pp. 379 – 384, 1992.
[7] J. L. Jing, F. Xu, R. S. Piao, "Application of pi-sigma neural network to real- time classification of seafloor sediments", Applied Acoustics, 24, pp. 346 – 350, 2005.
[8] J. A. Hussaina, P. Liatsisb, "Recurrent pi- sigma networks DPCM image coding", Neurocomputing, 55, pp. 363- 382, 2002.
[9] M. Sinha, K. Kumar, P. K. Kalra, "Some new neural network architures with improved learning schemes", Soft Convergence, 4, pp. 214 – 223, 2004.
[10] S. Haykin, "Neural Networks: A Comprehensive Foundation", 2nd edition. Beijing: *Tsinghua University Press* and Prentice Hall, 2001.
[11] M. Ishikawa, "Structural learning with forgetting", Neural Networks, 9(3), pp. 509 – 521, 1996.
[12] V. Setiono, "A penalty – function approach for pruning feed forward neural networks", Neural Networks, 9, pp. 185 -204, 1997.
[13] E. D. Karnin, "A simple Procedure for pruning back – propagation trained neural networks", IEEE Trans. Neural Network, 1(2), pp. 239 – 242, 1990.
[14] Wu Wei, "Hongmei Shao and Zhengxue Li, Convergence of batch BP algorithm with penalty for FNN training" Lecture Notes in Computer Science, 4232, pp. 562-569, 2006.
[15] Shao Hongmei, Zheng Gaofeng, "Boundedness and convergence of online gradient method with penalty and momentum", Neurocomputing, 74, pp. 765-770, 2011.
[16] Shao Hongmei, Wu Wei, Li Feng, "Convergence of online gradient method with a penalty term for feed forward neural networks with stochastic inputs". Numerical Mathematics, 1(14), pp. 87-96, 2005.
[17] T. Evgeniou, M. Pontil, T. Poggio, "Regularization networks and support vector machine"s, Adv. Comput. Math., 13, pp. 1 – 50, 2000.
[18] Wu Wei, Fan Qinwei, Jacek M. Zurada, Jian Wang, Dakun Yang, Yan Liu, "Batch gradient method with smoothing $L_{1/2}$ regularization for training of feed forward neural networks", Neural Networks, 50, pp. 72-78, 2014.
[19] Fan Qinwei, Jacek M. Zurada and Wei Wu, "Convergence of online gradient method for feed forward neural networks with smoothing $L_{1/2}$ regularization penalty", Neurocmputing, 131, pp. 208 – 216, 2014.
[20] P. Billingsley, "Probability and Measure", New York: John Wiley, 1979.
[21] X. D. Kang, Y. Xiong, C. Zhang, W. Wu, "Convergence of online gradient algorithm with stochastic inputs for Pi-sigma neural networks", Proc. IEEE Symp. Foundations of Comput. Intelligence, IEEE Press, pp. 564 – 569, 2007.
[22] W. Wu, G. R. Feng, X. Li, "Training multilayer perceptrons via minimization of sum of ridge functions", Advances in Computational Mathematics, 17, pp. 335 – 347, 2002.
[23] Y. X. Yuan, W. Y. Sun, "Optimization Theory and Methods", Science Press, Beijing, 2011.
[24] Shao Hongmei, Zheng Gaofeng, "Boundedness and convergence of online gradient method with penalty and momentum", Neurocomputing, 74, pp. 765 – 770, 2011.
[25] Z. X. Li, W. Wu, Y. L. Tian, "Convergence of an online gradient method for feed forward neural networks with stochastic inputs", Journal of Computational and Applied Mathematics, 163(1), pp. 165 –176, 2004

Authors Profiles

Khidir Shaib Mohamed (PhD student in Computational Mathematics) received the B.S. in Mathematics from Dalanj University – Dalanj – Sudan (2006) and M.S. in Applied Mathematics from Jilin University – Changchun – China (2011). He works as a lecturer of mathematics at College of Science – Dalanj University since (2011). Now he is a PhD student in Computational Mathematics at School of University of Technology, Dalian.

Yousif Shoaib Mohammed (Assistant Professor of Computational Physics) received the B.S. in Physics from Khartoum University – Oudurman – Sudan (1994) and High Diploma in Solar Physics from Sudan University of Science and Technology – Khartoum – Sudan (1997) and M.S. in Computational Physics (Solid – Magnetism) from Jordan University – Amman – Jordan and PhD in Computational Physics (Solid – Magnetism – Semi Conductors) from Jilin University – Changchun – China (2010) and worked as Researcher at Africa City of Technology – Khartoum – Sudan since 2012. He worked at Dalanj University since 1994 up to 2013 then from 2013 up to now at Qassim University – Kingdom of Saudi Arabia.

Abd Elmoniem Ahmed Elzain is Associate Professor in Physics and Researcher received the B.Sc. in Physics from Kassala University – Kassala - Sudan (1996) and High Diploma of Physics from Gezira University – Madani – Sudan (1997) and M.Sc. in Physics from Yarmouk University – Erbid – Jordan (2000) and PhD in Applied Radiation Physics from Kassala University – Kassala – Sudan (2006). He worked at Kassala University since 1996 up to 2010 then from 2010 up to now at Qassim University – Kingdom of Saudi Arabia.

Mohamed El-Hafiz M. Noor is Assistant Professor in Mathematice, received the B.Sc. in Mathematics from Wadi Elneel University – Atbara - Sudan (1996) and M.Sc. of Mathematics from Sudan University of Science and Technology – Khartoum – Sudan (1999) and PhD in Mathematics from Alneelain University – Khartoum – Sudan (2010). He worked at Zalingi University since 1997 up to 2013 then from 2012 up to now at Qassim University – Kingdom of Saudi Arabia.

Ellnoor Abaker Abdrhman Noh (Associate Professor of Physical Chemistry) received the B.S. in Physics from Khartoum University – Oudurman – Sudan (1991) and M.S. in Physical Chemistry (Corrosion) from Yarmouk University – Erbid – Jordan (1998) and PhD in Physical Chemistry (Computational) from North East Normal University – Changchun – China (2005). He worked at Dalanj University since 1992 up to 2011 then from 2011 up to now at Albaha University – Kingdom of Saudi Arabia.