STABILITY OF A NEW RECURRENT QUADRATIC NEURAL NETWORK

Franco Rubio López[1*], Orlando Hernández Bracamonte[1]

[1]Department of Mathematics, University of Trujillo, Trujillo, Peru

Abstract. In this paper we construct a new recurrent discrete neural network from the fixed points of a quadratic function. Our goal is to provide a criterion for the assignment of synaptic weights to the network, and we prove the stability of the neural network, without using the energy function used in Hopfield networks. In addition, we give an application to recognition of a pattern.

Keywords: *discrete Neural network, recurrent neural network, stability, fixed point, Hopfield network.*

AMS Subject Classification: *37D05, 37D25, 37D40, 37D45.*

Corresponding Author: *Prof. Franco Rubio López, Department of Mathematics, National University of Trujillo, Avenue Juan Pablo II S/N , Trujillo, Peru, e-mail: frubio@unitru.edu.pe*

Manuscript received: *16 June 2017*

1. Introduction

Since 1943, when Warren McCulloch and Walter Pitts [2] developed the first mathematical model of an artificial neuron, today the theory of neural networks has developed remarkably, and has been applied to many areas of science.

In 1982 Jhon Hopfield [1] presented a new model of a recurrent discrete neural network, which constitutes an associative memory with many applications such as recognition of patterns, images, signals, etc.

In this paper we construct a new discrete neural network called quadratic, in the sense that we use the fixed points of a quadratic function constructed by Rubio and Hernández [3]. Our goal is to establish a criterion for the assignment of synaptic weights to the network, which will guarantee the stability of the network at the fixed point previously given. In addition we make an application of our network in the area of recognition of a pattern; and we compared the results with a Hopfield bipolar network, only in the case of a single fixed point.

2. Quadratic function

In this section we present some results obtained by Rubio and Hernández [3], in which two points x_0 , $x_1 \in R$, $x_0 < x_1$ are given as fixed points a priori, and the quadratic function is determined:

$$f(x) = Ax^2 + Bx + C \tag{1}$$

where:

$$\begin{cases} A = \dfrac{y_m - x_m}{(x_m - x_1)(x_m - x_0)} \\ B = \dfrac{y_m(x_0 + x_1) - x_0 x_1 - x_m^2}{(x_1 - x_m)(x_m - x_0)} \\ C = \dfrac{x_0 x_1 (y_m - x_m)}{(x_m - x_1)(x_m - x_0)} \end{cases} \tag{2}$$

The point (x_m, y_m) is given in such a way that $(x_0, x_0), (x_1, x_1)$ y (x_m, y_m) are non-collinear.

Using the theorem (5.1) of [3], with $x_m = x_0 - \varepsilon,\ y_m = x_0\ ,\ \varepsilon = 0.1$, we have:

a) x_0 is a fixed point attractor. (3)
b) x_1 is a fixed point repellent.

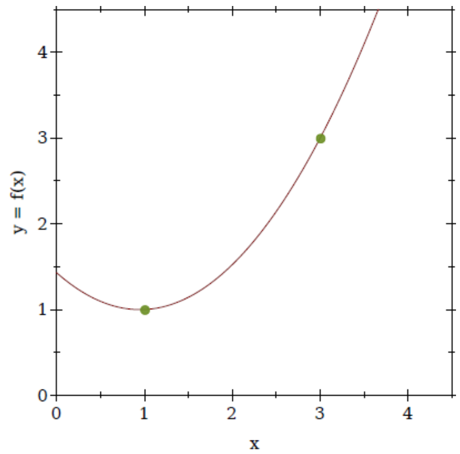

Fig. 1. Fixed point $x_0 = 1$, attractor

Using the theorem (5.4) of [3], with $x_m = x_1 + \varepsilon,\ y_m = x_1\ ,\ \varepsilon = 0.1$, we have:

a) x_0 is a fixed repellent point.
b) x_1 is an attractor fixed point. (4)

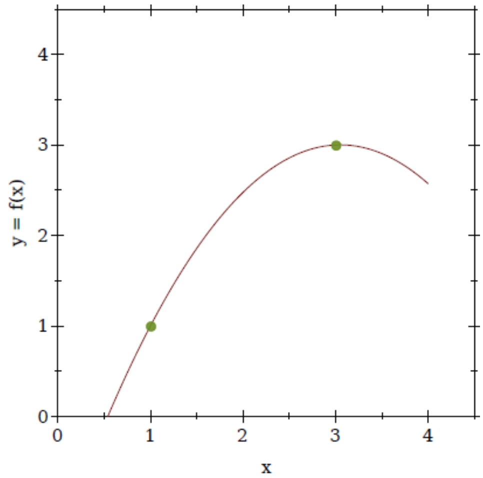

Fig. 2. Fixed point $x_1 = 3$, attractor.

3. N-dimensional case

Let $F: R^n \to R^n$ be an application defined by

$$F(x) = (F_1(x), F_2(x), \dots, F_n(x)),$$

where:

$$F_i(x) = A_i(\textstyle\sum_{j=1}^{n} w_{ij}x_j)^2 + B_i(\sum_{j=1}^{n} w_{ij}x_j) + C_i, \ \forall i = 1, \dots, n. \tag{5}$$

where $A_i, B_i\ y\ C_i$ are constants.

Now consider $X_p = (x_p^1, x_p^2, \dots, x_p^n) \in R^n$, such that:

$$f_i(x_p^i) = x_p^i, \forall i = 1, \dots, n. \tag{6}$$

where:

$$f_i(y) = A_i y^2 + B_i y + C_i, \ \forall i = 1, \dots, n, \text{ is given by (1).}$$

Theorem 1. Let $X_p = (x_p^1, x_p^2, \dots, x_p^n) \in R^n$, such that

$$f_i(x_p^i) = x_p^i, \ \forall i = 1, \dots, n.$$

X_p is a fixed point of F(x) if and only if $\sum_{j=1}^{n} w_{ij}x_p^j = x_p^i, \ \forall i = 1, \dots, n.$ (7)

Proof. If X_p is a fixed point of F(x), ie: $F(X_p) = X_p$, then:

$$A_i(\textstyle\sum_{j=1}^{n} w_{ij}x_p^j)^2 + B_i(\sum_{j=1}^{n} w_{ij}x_p^j) + C_i = x_p^i, \ \forall i = 1, \dots, n.$$
$$= f_i(x_p^i) = A_i(x_p^i)^2 + B_i(x_p^i) + C_i.$$

Therefore:

$$\textstyle\sum_{j=1}^{n} w_{ij}x_p^j = x_p^i, \ \forall i = 1, \dots, n.$$

Conversely, if $\sum_{j=1}^{n} w_{ij}x_p^j = x_p^i, \ \forall i = 1, \dots, n,$ then:

$$F_i(X_p) = A_i\left(\textstyle\sum_{j=1}^{n} w_{ij}x_p^j\right)^2 + B_i\left(\sum_{j=1}^{n} w_{ij}x_p^j\right) + C_i, \ \forall i = 1, \dots, n,$$
$$= A_i\left(x_p^i\right)^2 + B_i\left(x_p^i\right) + C_i = f_i\left(x_p^i\right) = x_p^i, \ \forall i = 1, \dots, n.$$

Therefore: $F(X_p) = X_p$.

Developing (7), gives:

$$\begin{cases} w_{11}x_p^1 + w_{12}x_p^2 + \cdots + w_{1n}x_p^n = x_p^1 \\ w_{21}x_p^1 + w_{22}x_p^2 + \cdots + w_{2n}x_p^n = x_p^2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ w_{n1}x_p^1 + w_{n2}x_p^2 + \cdots + w_{nn}x_p^n = x_p^n \end{cases}$$

The matrix associated with the previous system is:

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{pmatrix} \tag{8}$$

Furthermore, from (7) we obtain:

$$w_{ii} = 1 - \sum_{\substack{j=1 \\ j \neq i}}^{n} w_{ij}\frac{x_p^j}{x_p^i}, \ \ x_p^i \neq 0, \tag{9}$$

Now, for our study consider the space of Hamming $H^n = \{-1; 1\}^n$, and let $X_p \in H^n$.

Using (9), we in this paper, establish a criterion to assign values to the elements of matrix W, which will be the synaptic weights of the network, considering the more general criterion proposed by Rubio and Hernández [4].

Criterion

For assigning values to the elements of matrix W, follow the steps.

1. $h = \frac{1}{n}$.
2. If $\frac{x_p^j}{x_p^i} > 0$, then $w_{ij} = h$. (10)
3. If $\frac{x_p^j}{x_p^i} < 0$, then $w_{ij} = -h$. (11)

Theorem 2. The diagonal elements of the matrix W, are given by:

$$w_{ii} = h, \;\; \forall i = 1, \dots, n. \tag{12}$$

Proof . From (9):

$$\begin{aligned} w_{ii} &= 1 - \sum_{\substack{j=1 \\ j \neq i}}^{n} w_{ij} \frac{x_p^j}{x_p^i}, \quad x_p^i \neq 0, \\ &= 1 - \sum_{\substack{j=1 \\ j \neq i}}^{n} h, \\ &= 1 - h(n-1) = h \end{aligned}$$

Therefore:

$$w_{ii} = h, \;\; \forall i = 1, \dots, n.$$

Definition 1. Given $X_p \in H^n$, the point $X_a = -X_p$, is called the Antipode point of X_p.

Theorem 3. Let be $X_p = \left(x_p^1,\, x_p^2, \dots, x_p^n\right) \in H^n, y\; W_i = (w_{i1}, w_{i2}, \dots, w_{in})$ the nth row of W.

1. If $x_p^i = 1$, then $W_i = hX_p$. (13)
2. If $x_p^i = -1$, then $W_i = hX_a$. (14)

Proof.

1. If $x_p^i = 1$, then $\frac{x_p^j}{x_p^i} = x_p^j, \forall\, j = 1, \dots, n; j \neq i$. We have two cases:

   a) If $x_p^j = 1$, then $w_{ij} = h = h(1) = hx_p^j$.

   b) If $x_p^j = -1$, then $w_{ij} = -h = h(-1) = hx_p^j$.

Therefore:

$$W_i = \left(hx_p^1, hx_p^2, \dots, hx_p^n\right) = hX_p.$$

2. If $x_p^i = -1$, then $\frac{x_p^j}{x_p^i} = -x_p^j, \forall\, j = 1, \dots, n; j \neq i$. We have two cases:

   a) If $x_p^j = 1$, then $w_{ij} = -h = h(-1) = h(-x_p^j)$.

   b) If $x_p^j = -1$, then $w_{ij} = h = h(-1)(-1) = h(-x_p^j)$.

Therefore:

$$W_i = \left(h(-x_p^1), h(-x_p^2), \dots, h(-x_p^n)\right) = h(-X_p) = hX_a.$$

Theorem 4. Let $X_p \in H^n$, w_{ij} given by (10) or (11), then the matrix $W = (w_{ij})_{nxn}$ is symmetric.

Proof. Since $x_p^i,\, x_p^j \in H^1 = \{-1; 1\}$, then $\frac{x_p^j}{x_p^i} = \frac{x_p^i}{x_p^j}$, then: $w_{ij} = w_{ji}$. Therefore, W is symmetric.

Now, let us consider the functions $V_i: R^n \to R$, defined by:

$$V_i(x) = \sum_{j=1}^{n} w_{ij}x_j\,,\ \forall\, i = 1,\dots,n \tag{15}$$

Then:

$$V_i(x) =< (w_{i1},\ w_{i2},\dots,w_{in}),(x_1,x_2,\dots,x_n) >$$
$$V_i(x) =< W_i, x > \tag{16}$$

Theorem 5. Let $x_p \in H^n$. We then have:

1. If $W_i = hx_p$, then $V_i(x_p) = 1$. (17)
2. If $W_i = hx_a$, then $V_i(x_p) = -1$. (18)

Proof.

1. If $W_i = hx_p$, then $V_i(x_p) =< hx_p, x_p >= h < x_p, x_p >= hn = 1$.
2. If $W_i = hx_a$, then $V_i(x_p) =< hx_a, x_p >= -h < x_p, x_p >= -1$.

Theorem 6. Let $x_p \in H^n$, $x_a = -x_p$. We then have:

1. If $W_i = hx_p$, then $V_i(x_a) = -1$. (19)
2. If $W_i = hx_a$, then $V_i(x_a) = 1$. (20)

Proof.

1. If $W_i = hx_p$, then $V_i(x_a) =< hx_p, x_a >= h < x_p, -x_p >= -1$.
2. If $W_i = hx_a$, then $V_i(x_a) =< hx_a, x_a >= h < x_a, x_a >= 1$.

4. Stability

In this section we prove results, which are fundamental to establish the stability of the quadratic recurrent discrete neural network.

From (5) we have:

$$F_i(x) = A_i\Big(\sum_{j=1}^{n} w_{ij}x_j\Big)^2 + B_i\Big(\sum_{j=1}^{n} w_{ij}x_j\Big) + C_i\,,\ \forall i = 1,\dots,n.$$

Thus, the application $F(x) = (F_1(x),\dots,F_n(x))$ is differentiable of class $C^\infty(R^n)$.
Then:

$$\frac{\partial F_i(x)}{\partial x_k} = 2A_i\Big(\sum_{j=1}^{n} w_{ij}\, x_j\Big)w_{ik} + B_i w_{ik},\ \ \forall\, i,k = 1,\dots,n,$$
$$\frac{\partial F_i(x)}{\partial x_k} = \Big(2A_i\Big(\sum_{j=1}^{n} w_{ij}\, x_j\Big) + B_i\Big)w_{ik} \tag{21}$$

Therefore, from (21) the Jacobian matrix de F in $x_p \in H^n$, is:

$$\mathrm{JF}(\mathrm{x_p}) = \Big(\frac{\partial \mathrm{F_i}(\mathrm{x_p})}{\partial \mathrm{x_k}}\Big)_{\mathrm{nxn}} = \Big(\Big(2\mathrm{A_i}\Big(\sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{w_{ij}}\, \mathrm{x_p^j}\Big) + \mathrm{B_i}\Big)\mathrm{w_{ik}}\Big)_{\mathrm{nxn}} \tag{22}$$

Theorem 7. Let $x_p = (x_p^1,\dots,x_p^n) \in H^n$, x_p^i are attracting fixed points given by (3) or (4), and $W = (w_{ij})_{nxn}$ the matrix given by (8). We then have:

$$\|JF(x_p)\|_\infty < \|W\|_\infty \tag{23}$$

Proof. From (21) we have:

$$\sum_{k=1}^{n}\left|\frac{\partial F_i(x_p)}{\partial x_k}\right| = \sum_{k=1}^{n}\Big|\Big(2A_i\Big(\sum_{j=1}^{n} w_{ij}\, x_j\Big) + B_i\Big)w_{ik}\Big|$$
$$= \sum_{k=1}^{n}\Big|\Big(2A_i\Big(\sum_{j=1}^{n} w_{ij}\, x_j\Big) + B_i\Big)\Big||w_{ik}|$$
$$< \sum_{k=1}^{n}|w_{ik}| \le \|W\|_\infty.$$

Therefore: $\|JF(x_p)\|_\infty < \|W\|_\infty$.

Theorem 8. Let $W = (w_{ij})_{nxn}$ be the matrix given by (8).
Then

$$\|W\|_\infty = 1. \tag{24}$$

Proof. Since $h = \frac{1}{n}$, then:

$$\sum_{k=1}^{n} |w_{ik}| = \sum_{k=1}^{n} |\pm h| = hn = 1,\ \forall i = 1, \dots, n.$$

Therefore:

$$\|W\|_\infty = 1.$$

Theorem 9. If $x_p^i \in H = \{-1; 1\}$, are fixed points attracting of the functions $f_i(y) = A_i y^2 + B_i y + C_i, \forall\ i = 1, \dots, n$, then $x_p = \left(x_p^1, \dots, x_p^n\right)$ is a fixed point attractor of $F(x)$.
Proof. By theorem (7), we have $\left\|JF\left(x_p\right)\right\|_\infty < \|W\|_\infty$, and by theorem (8), we have $\|W\|_\infty = 1$, then $\left\|JF\left(x_p\right)\right\|_\infty < 1$. Therefore, x_p is an attractor point.
Theorem 10. If $x_p \in H^n$ is a fixed point attractor of F, then $x_a = -x_p$ it is a repellent fixed point of F.
Proof. Since $F(x_a) = F(-x_p)$, we have:

$$F_i(x_a) = A_i(\sum_{j=1}^{n} w_{ij}\, x_a^j)^2 + B_i\left(\sum_{j=1}^{n} w_{ij}\, x_a^j\right) + C_i\,, \qquad \forall i = 1, \dots, n.$$

$$F_i(x_a) = A_i(\sum_{j=1}^{n} w_{ij}\,(-x_p^j))^2 + B_i\left(\sum_{j=1}^{n} w_{ij}\,(-x_p^j)\right) + C_i\,, \forall i = 1, \dots, n.$$

$$= A_i(\textstyle\sum_{j=1}^{n} w_{ij}\, x_p^j)^2 - B_i\left(\sum_{j=1}^{n} w_{ij}\, x_p^j\right) + C_i\,,\ \forall i = 1, \dots, n.$$

$$= A_i(x_p^i)^2 - B_i\left(x_p^i\right) + C_i\,,\ \forall i = 1, \dots, n.$$

$$= A_i(-x_p^i)^2 + B_i\left(-x_p^i\right) + C_i\,,\ \forall i = 1, \dots, n.$$

$$= f_i\left(-x_p^i\right) = -x_p^i = -x_a^i\,,\ \forall i = 1, \dots, n.$$

Therefore:
$F(x_a) = x_a$, and as x_p^i is a fixed point attractor $\forall i = 1, \dots, n$, then $-x_p^i$ it is a repellent fixed point. Therefore x_a it is a repellent fixed point of F.

5. Application

In this section we give an application of our quadratic recurrent neural network to the recognition of a pattern.

Using (2), (3) and (4), with $x_0 = -1$, $x_1 = 1$, we construct the quadratic functions:
a) $f_+(x) = -0.4762x^2 + x + 0.4762$, with $x_1 = 1$, fixed point attractor.
b) $f_-(x) = 0.4762x^2 + x - 0.4762$, with $x_0 = -1$, fixed point attractor.
Example 1. The pattern given in Figure 3 is represented using 64 neurons. In addition, the disturbed pattern is considered, a tolerance equal to 0.01; and the information is processed using our quadratic recurrent neural network. As can be seen in the figure, our neural network allows to completely regenerate the pattern.

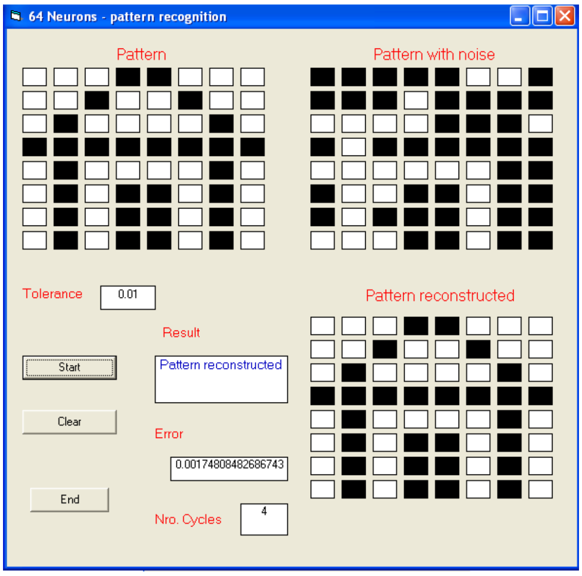

Fig. 3. Result of the quadratic neural network.

The following figure shows that the perturbation of the initial pattern is greater than the previous one; however, our quadratic recurrent neural network allows a complete regeneration of the pattern, which is not the case with a single-point bipolar Hopfield neural network.

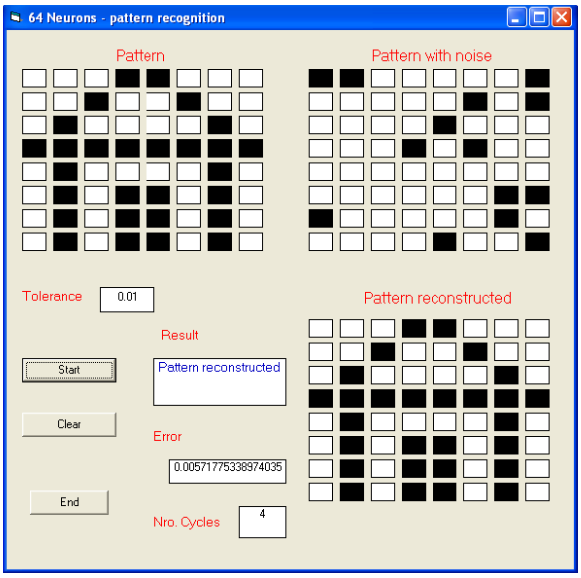

Fig. 4. Greater disturbance in the pattern.

Example 2. Consider the pattern given by the figure.

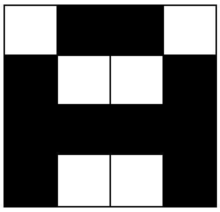

Fig. 5. Letter A.

Which is represented using 16 neurons:

$$P = [-1, 1, 1, -1, 1, -1, -1, 1, 1, 1, 1, 1, 1, -1, -1, 1].$$

Using a Hopfield bipolar network with a single pattern to be memorized, the matrix of synaptic weights is given by the table:

0	-1	-1	1	-1	1	1	-1	-1	-1	-1	-1	-1	1	1	-1
-1	0	1	-1	1	-1	-1	1	1	1	1	1	1	-1	-1	1
-1	1	0	-1	1	-1	-1	1	1	1	1	1	1	-1	-1	1
1	-1	-1	0	-1	1	1	-1	-1	-1	-1	-1	-1	1	1	-1
-1	1	1	-1	0	-1	-1	1	1	1	1	1	1	-1	-1	1
1	-1	-1	1	-1	0	1	-1	-1	-1	-1	-1	-1	1	1	-1
1	-1	-1	1	-1	1	0	-1	-1	-1	-1	-1	-1	1	1	-1
-1	1	1	-1	1	-1	-1	0	1	1	1	1	1	-1	-1	1
-1	1	1	-1	1	-1	-1	1	0	1	1	1	1	-1	-1	1
-1	1	1	-1	1	-1	-1	1	1	0	1	1	1	-1	-1	1
-1	1	1	-1	1	-1	-1	1	1	1	0	1	1	-1	-1	1
-1	1	1	-1	1	-1	-1	1	1	1	1	0	1	-1	-1	1
-1	1	1	-1	1	-1	-1	1	1	1	1	1	0	-1	-1	1
1	-1	-1	1	-1	1	1	-1	-1	-1	-1	-1	-1	0	-1	-1
1	-1	-1	1	-1	1	1	-1	-1	-1	-1	-1	-1	1	0	-1
-1	1	1	-1	1	-1	-1	1	1	1	1	1	1	-1	-1	0

Table 1. Table of the synaptic weights

Now, we enter the disturbed pattern:

$$Pert = [1, -1, 1, 1, 1, -1, 1, -1, -1, -1, -1, 1, -1, -1, 1, -1, 1].$$

Which goes to the fixed point:

$$Ps = [1, -1, -1, \ 1, -1, \ 1, \ 1, -1, -1, -1, -1, -1, -1, 1, 1, -1],$$

That corresponding to:

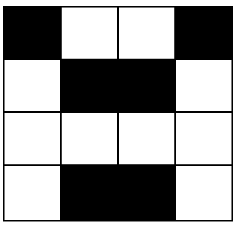

Fig. 6. Antipode fixed point.

As shown, when entering Pert, the network does not converge to the pattern P; but to the antipode -P.

Now, we use our quadratic recursive neural network with the pattern P, and the disturbed *Pert* pattern. After processing the information our network is stabilized obtaining the initial pattern; which is not the case with the Hopfield network.

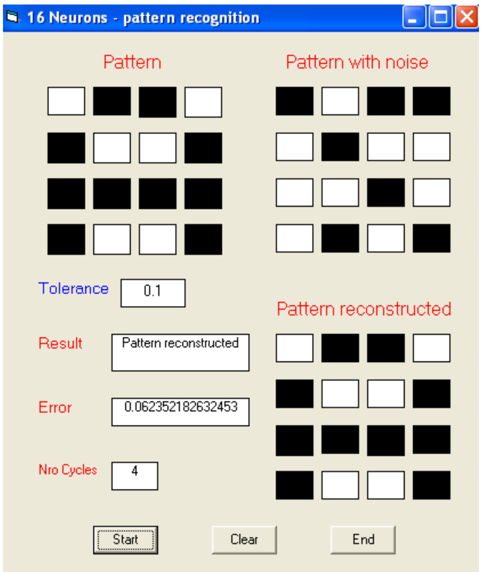

Fig. 7. Result of the quadratic recursive neural network.

In the following example, we further modify the pattern P; we process the information with our quadratic recurrent neural network; and we see that it stabilizes allowing to reconstruct the initial pattern; which is observed in the following figure.

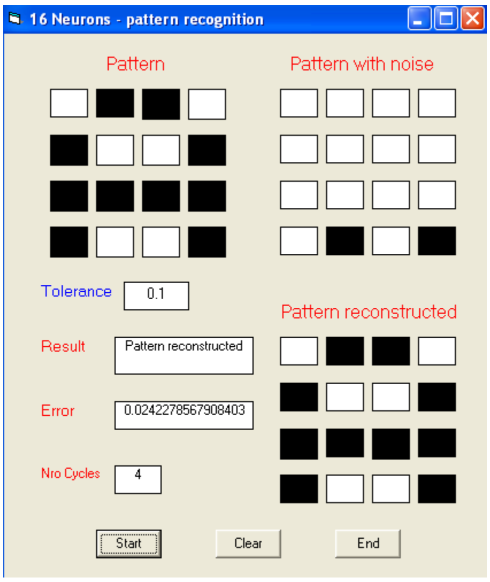

Fig. 8. Stability of the quadratic recurrent neural network.

This shows that our quadratic recursive neural network presents an improvement over the Hopfield network in the case of a single fixed point.

6. Conclusion

In this paper we construct a new quadratic discrete recurrent neural network with a fixed point attractor given previously, using the fixed points of quadratic functions given by (1) – (4).

Using the relation given in (9); we established a criterion for assigning values (synaptic weights) to the elements of matrix W; which allowed to prove that the $\|W\|_\infty = 1$.

In theorem 7 it is proved that the norm of the Jacobian matrix associated with the neural network at fixed point x_p is less than $\|W\|_\infty$; which guarantees the stability of the fixed point; methodology different from that used by Hopfield [1], which makes use of the energy function associated to the system.

This new quadratic discrete recurrent neural network behaves as auto-associative memory; allowing to reconstruct objects from certain information; as in the recognition of images, sounds; as in the application example to the recognition of a pattern.

References

1. Hopfield J., (1982) Neural Networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci.*, *USA*, 79, 2554-2558.
2. McCulloch W., Pitts W., (1943) A logical calculus of the ideas immanent in neurons activity, *Bulletin Mathematical Biophysics*, 5, 115-133.
3. Rubio López F., Hernández Bracamonte O., (2015) Construcción de una función polinómica a partir de los puntos fijos dados previamente, *Selecciones Matemáticas*, 2(01), 54-67.
4. Rubio López F., Hernández Bracamonte O., (2017) Construcción de una función Vectorial a partir de un punto fijo dado previamente, *Selecciones Matemáticas*, 4(01), 124-138.