Turk J Math
25 (2001) , 345 – 354.

Global Convergence for Discrete Dynamical Systems and Forward Neural Networks

Asuman G. Aksoy, Mario Martelli

Abstract

Using a theoretical result regarding the global stability of discrete dynamical systems of lower triangular form, we establish convergence properties of forward neural networks when the neuron response functions fail to be continuous.

1. Introduction

Global convergence to a stationary state is a fundamental property of a neural network. When an input $\vec{x}_I$ is received and regardless of the initial state $\vec{x}_0$ of the net, it is expected that a stationary state $\vec{x}_s$ is reached which depends only on $\vec{x}_I$, i.e. $\vec{x}_s = \vec{x}_s(\vec{x}_I)$. The purpose of this note is to prove a theorem regarding discrete dynamical systems of triangular type and to show that certain types of neural networks do have, as a consequence of the theorem, the required global convergence, even in these cases in which the neuron response functions fail to be continuous. The established result can be viewed as a generalization of [5] where analogous theorems were obtained under the more restrictive assumption of differentiability of the various neuron response functions. The paper is divided into four parts. Part 2 contains some notations and definitions. In part 3 we prove a global convergence result on triangular maps and in the last part we apply it to a family of neural networks.

2. Notation and Definitions

Let X be a region in R^q. A function $K : X \to X$ defines a discrete dynamical system as follows. Assume that the state at time $n, \vec{x}_n$ is known. Then the state of the system at time $n+1, \vec{x}_{n+1}$ is given by $\vec{x}_{n+1} = K(\vec{x}_n)$. The *orbit* or *trajectory* of a point $\vec{x}_0 \in X$ is the sequence of states

$$O(\vec{x}_0) = \{\vec{x}_0, \vec{x}_1 = K(\vec{x}_0), \ldots, \vec{x}_{n+1} = K(\vec{x}_n), \ldots\}.$$

A point $\vec{x}_s \in X$ is called a *fixed point* or *stationary state* of the dynamical system governed by K if $O(\vec{x}_s) = \{\vec{x}_s\}$. A point $\vec{a} \in X$ is called a *limit point* of $O(\vec{x}_0)$ if there is a subsequence of the orbit, which converges to $\vec{a}$. We will use $L(\vec{x}_0)$ to denote the limit points of $O(\vec{x}_0)$. When K is continuous and the orbit $O(\vec{x}_0)$ is bounded, it can be easily verified that $L(\vec{x}_0)$ is non-empty, compact, and invariant under the action of K, i.e. [4]

$$K(L(\vec{x}_0)) = L(\vec{x}_0). \tag{1}$$

A continuous map $F : R^q \to R^q$ of the form

$$F(x_1, x_2, \ldots, x_q) = (f_1(x_1), f_2(x_1, x_2), \ldots, f_q(x_1, x_2, \ldots, x_q)) \tag{2}$$

is called *lower triangular*. Upper triangular maps are defined in a similar manner. A discrete model of a *neural network* (Hopfield model [2]) with q neurons is a dynamical system in R^q of the form

$$\vec{x}_{n+1} = \vec{x}_n - C\vec{x}_n + TF(\vec{x}_n) + \vec{x}_I. \tag{3}$$

The components of the vector $\vec{x}_{n+1}$ represent the energies of the various neurons of the net at time $n+1$. C is a $q \times q$ diagonal matrix (leakage matrix) whose diagonal entries are numbers in $(0, 1)$ and it represents the energy lost by the system at each iteration. T is the *connectivity matrix*, with its entries $t_{i,j}$ representing the strength of the connection between neuron j and neuron i. $t_{i,j} > 0$ means that neuron j acts in an excitatory manner on neuron i while $t_{i,j} < 0$ implies that its action is inhibitory.

$$F(\vec{x}) = F(x_1, x_2, \ldots, x_q) = (f_1(x_1), f_2(x_2), \ldots, f_q(x_q))$$

is the neuron response function. $F(\vec{x})$ provides a threshold level below which the neurons are inactive. Each f_i is frequently taken in the form of a unit step function (0 for $x < 0$, 1 for $x \geq 0$). Finally, x_I represents the q-dimensional input vector, namely the signal received by the net. A network modeled as in (3) is called forward if $t_{i,j} = 0$ for $j \geq i$. The map associated with the system (3) is

$$K(\vec{x}) = (I - C)\vec{x} + TF(\vec{x}) + \vec{x}_I, \tag{4}$$

and the k^{th} equation, $k = 1, 2, \ldots q$ of (3) assumes the form

$$x_{k,n+1} = (1 - c_k)x_{k,n} + t_{k,1}f_1(x_{1,n}) + \cdots + t_{k,q}f_q(x_{q,n}) + x_{kI}. \tag{5}$$

When a signal $\vec{x}_I$ arrives it finds the net in a certain state $\vec{x}_0$ and the dynamical process starts. The problem is to have convergence of all orbits of (3) to a unique fixed point, depending only on $\vec{x}_I$, no matter what $\vec{x}_0$ is. This is obviously the case if K is a contraction, but that may be too much to ask for. In [5] global convergence was established when K is triangular and under the assumptions that all f_i's are differentiable with continuous derivatives. Later in [3] it was proved that, for a continuous K which may fail to be differentiable at the points of a linearly denumerable set one has global convergence of orbits if the norm of the derivative satisfies certain inequalities. Notice that in [3] K is neither a triangular map nor a contraction. In this paper we do not assume the differentiability of $f_i, i = 1, 2, \dots, q$ at every point and the continuity of the derivatives and show that with suitable assumptions a map of the form (2) admits a global attractor. This result is applied to forward neural networks, for which K (see(4)) has the form (2). It is proved that such systems converge to a fixed point which is independent of the initial state.

3. Results

Three preliminary lemmas will simplify the proof of our main result. In the first and second lemma we use a version of the Mean Value Inequality which can be found in [3] (see also [1], pg. 158). It states that given a continuous function $f : [a, b] \to \mathbb{R}$ which is differentiable except possibly on a denumerable set A, there exists $c \in [a, b]$ such that

$$|f(b) - f(a)| \leq |f'(c)|(b - a) \tag{6}$$

Lemma 1 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and differentiable except possibly on a denumerable set A. Assume that $|f'(x)| \leq 1$ and there is a point c such that $|f'(c)| < 1$. Then for every x, y such that $(x - c)(y - c) < 0$ we have $|f(x) - f(y)| < |x - y|$.*

Proof. Clearly $|f(x)-f(y)| \leq |x-y|$ for every $x, y \in \mathbb{R}$ (see [3]). Assume that $f'(c) > 0$ and consider the function $g(x) = x - f(x)$. Then g is strictly increasing at c i.e., there is an open interval $I, c \in I$, such that $w < c < z, w, z \in I$ implies $g(w) < g(c) < g(z)$. Hence $f(z) - f(w) < z - w$. Similarly, setting $h(x) = x + f(x)$, we obtain $f(w) - f(z) < z - w$. Thus $|f(z) - f(w)| < z - w = |z - w|$. Now, with $x < w$ and $z < y$, we have

$$|f(x) - f(w) + f(w) - f(z) + f(z) - f(y)| < |x - w| + |w - z| + |z - y| \leq \quad |x - y|.$$

□

Lemma 2 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and differentiable except possibly on a denumerable set A. Assume that*

1. $f(0) = 0$
2. $|f'(x)| \leq 1$ *for all* $x \in A^c$.
3. *There are two sequences* c_n *and* d_n *such that* $c_n \uparrow 0$ *and* $d_n \downarrow 0$ *and* $|f'(c_n)| < 1$, $|f'(d_n)| < 1$.

Then 0 is a global attractor.

Proof. We first show that 0 is the only fixed point of f. Define $g(x) = x - f(x)$. Then $g'(x) \geq 0$ and g is non-decreasing (see [3]). Let $a < 0$ and assume that $g(a) = 0$. Notice that g cannot be constant in $[a, 0]$ since there are points $c_n \in (a, 0)$ such that $g'(c_n) > 0$. Hence, there exists $r \in (a, 0)$ such that either $g(r) < g(a) = 0$ or $0 = g(0) < g(r)$, contradicting the non-decreasing character of g. The proof is similar when $a > 0$. Consequently, $g(x)x > 0$ for every $x \neq 0$ and 0 is the only fixed point of f.

Next we show that f does not have any periodic orbit of period 2. In fact, assume that $a < b$ are such that $f(a) = b$ and $f(b) = a$. Then f has a fixed point in $[a, b]$. Since the only fixed point of f is $x = 0$ we must have $a < 0 < b$. But, by Lemma 1 $|f(b) - f(a)| < b - a$. Hence, we cannot have $f(a) = b$ and $f(b) = a$.

Finally, we prove that every orbit converges to 0. From $g(x)x > 0$ we derive $x_0 < x_1 = f(x_0)$ when $x_0 < 0$ and $x_0 > x_1 = f(x_0)$ when $x_0 > 0$. Moreover, by the Mean Value Inequality (see [3]), $|x_1| = |f(x_0) - f(0)| \leq |x_0|$. It follows that the sequence $\{|x_n| : n = 0, 1, \ldots\}$ is non-increasing. Thus $\{|x_n| : n = 0, 1, \ldots\}$ converges to some $r \geq 0$. The limit set of $\{x_n : n = 0, 1, \ldots\}$ is either a singleton, $\{r\}$ or $\{-r\}$, or it coincides with the set $\{-r, r\}$. In the first case r or $-r$ is a fixed point. Hence $r = 0$ by the first part of this proof. In the second case $\{-r, r\}$ is a periodic orbit of period 2. Hence $r = 0$ by the second part of this proof. □

Lemma 3 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Assume that f has a bounded orbit. Then f has a fixed point.*

Proof. Let $O(x_0)$ be bounded. Then $L(x_0)$ is non-empty and compact. Moreover, (see (1)) $f(L(x_0)) = L(x_0)$. Let $z_1, z_2 \in L(x_0)$ be such that $z_1 \leq z \leq z_2$ for every $z \in L(x_0)$. Then $f(z_1), f(z_2) \in [z_1, z_2]$. By the Intermediate Value Theorem f has a fixed point in $[z_1, z_2]$. □

We are now ready to state and prove the main result of this paper. As a consequence of Lemma 3 the assumption $F(\vec{0}) = \vec{0}$ can be replaced by the condition that for every $i = 1, 2, ...q$ the function f_i has a bounded orbit when its first $i - 1$ coordinates remain fixed.

Theorem 1 *Let $F : \mathbb{R}^q \to \mathbb{R}^q$ be continuous and such that*

$$F(\vec{x}) = F(x_1, x_2, \dots, x_q) = (f_1(x_1), f_2(x_1, x_2), \dots, f_q(x_1, x_2, \dots, x_q)).$$

Assume that f_1 satisfies the conditions of Lemma 2. Moreover assume that for $i = 2, \dots, q$ the following conditions hold.

1. $f_i(0, 0, \dots, 0) = 0$.
2. *There exist constants r_i and subsets $A_i \subset \mathbb{R}^i$ and $P_i = \{(x_1, x_2, \dots, x_i) : |x_j| \leq r_i \text{ for } j = 1, 2, \dots, i - 1\}$ such that*
   (a) *$A_i \subset P_i$, and every line $\ell = \{(x_1, \dots, x_{i-1}, t) \mid t \in \mathbb{R}\} \subset P_i$ intersects A_i in at most a denumerable set of points;*
   (b) *$\left|\frac{\partial f_i}{\partial x_i}\right| \leq 1$ in $P_i \setminus A_i$ and there are sequences $c_{in} \uparrow 0$ and $d_{in} \downarrow 0$ such that $\left|\frac{\partial f_i}{\partial x_i}(x_1, x_2, \dots, c_{in})\right| < 1$ and $\left|\frac{\partial f_i}{\partial x_i}(x_1, x_2, \dots, d_{in})\right| < 1$ in $P_i \setminus A_i$.*

Then $\vec{0}$ is the only fixed point of F and every orbit converges to $\vec{0}$.

Remark. To clarify condition 2 of the above theorem, suppose $A_3 \subset \mathbb{R}^3$ and $P_3 = \{(x, y, z) : |x| < r_3, |y| < r_3\}$. Then $|\frac{\partial f_3}{\partial z}| \leq 1$ in $P_3 \setminus A_3$. Moreover, there are two sequences $c_{3n} \uparrow 0$ and $d_{3n} \downarrow 0$ such that $|\frac{\partial f_3}{\partial z}(x, y, c_{3n})| < 1$ and $|\frac{\partial f_3}{\partial z}(x, y, d_{3n})| < 1$ in $P_3 \setminus A_3$.

Proof. The result is true for $q = 1$ (see Lemma 2). Assume it true for $k = 1, 2, \dots, i-1$ and let us show that the same conclusion holds for $k = i$. By Lemma 2, for every $|x_k| \leq \frac{r_i}{2}, k = 1, 2, \dots, i - 1$ we have

$$|f_i(x_1, x_2, \dots, x_{i-1}, c_{i1}) - f_i(x_1, x_2, \dots, x_{i-1}, 0)| < |c_{i1}|.$$

Likewise

$$|f_i(x_1, x_2, \dots, x_{i-1}, d_{i1}) - f_i(x_1, x_2, \dots, x_{i-1}, 0)| < d_{i1}.$$

Since the set $|x_k| \le \frac{r_i}{2}, k = 1, \dots, i-1, x_i = c_{i1}$ or $x_i = d_{i1}$ is compact, there exists $s_i > 0$ such that

$$|f_i(x_1, \dots, x_{i-1}, c_{i1}) - f_i(x_1, \dots, x_{i-1}, 0)| \le |c_{i1}| - s_i$$

and

$$|f_i(x_1, \dots, x_{i-1}, d_{i1}) - f_i(x_1, \dots, x_{i-1}, 0)| < d_{i1} - s_i.$$

It follows that for $x_i < c_{i1}$ and $|x_k| \le \frac{r_i}{2}, k = 1, 2, \dots, i-1$ we have

$$\begin{aligned} & |f_i(x_1, \dots, x_{i-1}, x_i) - f_i(x_1, \dots, x_{i-1}, 0)| \\ \le\ & |f_i(x_1, \dots, x_{i-1}, x_i) - f_i(x_1, \dots, x_{i-1}, c_{i1})| \\ & + |f_i(x_1, \dots, x_{i-1}, c_{i1}) - f_i(x_1, \dots, x_{i-1}, 0)| \\ \le\ & |x_i - c_{i1}| + |c_{i1}| - s_i \qquad (7) \\ =\ & c_{i1} - x_i - c_{i1} - s_i \\ =\ & |x_i| - s_i \end{aligned}$$

Similarly for $x_i > d_{i1}$ we have

$$|f_i(x_1, \dots, x_{i-1}, x_i) - f_i(x_1, \dots, x_{i-1}, 0)| \le x_i - s_i \qquad (8)$$

Since f_i is continuous there is $0 < \delta \le \frac{r_i}{2}$ such that $|f_i(x_1, \dots, x_{i-1}, 0)| \le \frac{s_i}{2}$ whenever $|x_k| \le \delta$ for $k = 1, 2, \dots, i-1$. By the induction argument, we know that the sequences $x_{jn} \to 0, j = 1, 2, \dots, i-1$. Hence, without loss of generality, we can assume that $|x_{j0}| \le \delta$ for every $j = 1, 2, \dots, i-1$. Now, let $x_{i0} < c_{i1}$. An easy computation, based on the inequalities (7) and (8) shows that $|x_{i1}| \le |x_{i0}| - \frac{s_i}{2}$ and a similar result holds if $x_{i0} > d_{i1}$. Thus, as long as $x_{in} < c_{i1}$ or $x_{in} > d_{i1}$ we have $|x_{i,n+1}| \le |x_{in}| - \frac{s_i}{2}$. Therefore, the sequence $\{x_{in}, n = 1, 2, \dots\}$ is bounded and the set L of its limit points is non-empty and compact. Moreover, an easy adaptation of (1), together with the property $x_{jn} \to 0, j = 1, 2, \dots, i-1$, imply that for every $z \in L$ we have that $f_i(0, 0, \dots, z) \in L$ and there exists $w \in L$ such that $f_i(0, 0, \dots, w) = z$. By applying Lemma 2 to $f_i(0,0,\ldots,s)$ we obtain that every orbit of f_i goes to 0. Therefore, we conclude that $L=\{0\}$. □

4. Application to Forward Neural Networks

Consider the dynamical system (Hopfield Model)

$$\vec{x}_{n+1} = \vec{x}_n - C\vec{x}_n + TF(\vec{x}_n) + \vec{x}_I$$

The neuron response function F is frequently of the form $F(\vec{x}) = (f(x_1),\ldots,f(x_q))$ with

$$f(x) = \begin{cases} 0 & x \le 0 \\ x & 0 \le x \le 1 \\ 1 & 1 \le x \end{cases} \tag{9}$$

However in many cases, it may be useful to assume that f is not a continuous function, for example

$$f(x) = \begin{cases} 0 & x \le a \\ 1 & x > a \end{cases} \tag{10}$$

where $a \in [0,1]$. Another possibility could be

$$f(x) = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x. \end{cases} \tag{11}$$

It may happen that the neuron response function is different for each neuron. The vector x_I is the input vector and the matrix T is lower triangular (forward network). Hence the non-zero entries of T are all below the main diagonal. Under these circumstances we have the following result.

Theorem 2 *Let $\vec{x}_{n+1} = (I-C)\vec{x}_n + TF(\vec{x}_n) + \vec{x}_I$ be a discrete dynamical system representing a forward neural network. Assume that $c_i \in (0,1]$ for every $i = 1,2,\ldots,q$. Then given an initial condition $\vec{x}_0$ the iteration scheme converges to a point $\vec{x}_s$ which depends only on $\vec{x}_I$.*

Proof. Let $K(\vec{x}) = (I - C)\vec{x} + TF(\vec{x}) + \vec{x}_I$. Notice that under the assumptions of the theorem we have $\frac{\partial k_i}{\partial x_i}(\vec{x}) = 1 - c_i \in [0, 1)$. Hence, Theorem 1 can be applied whenever the functions $f_i, i = 1, \ldots, q$ are either of the form (9) or (11). In the case where at least one of the functions is of the form (10) we cannot apply Theorem 1 directly since F is not continuous. However, observe that in the case when all neuron response functions are of the form (10) we have

$$x_{1,(n+1)} = k_1(x_{1,n}) = (1 - c_1)x_{1,n} + x_{1,I}.$$

Consequently, $x_{1,n} \to \frac{x_{1,I}}{c_1}$ and $f_1(x_{1,n})$ is eventually a constant k_1 ($k_1 = 0$ or $k_1 = 1$). Now let us look at what is happening to $x_{2,n}$ as $n \to \infty$. Notice that for n large enough,

$$x_{2,(n+1)} = k_2(x_{2,n}) = (1 - c_2)x_{2,n} + k_1 + x_{2I}.$$

It follows that $x_{2,n} \to (k_1 + x_{2I})/c_2$. Therefore, also $f_2(x_{2,n})$ is eventually a constant k_2. An induction argument shows that $O(\vec{x}_0)$ is convergent to a fixed point $\vec{x}_s$ which depends only on $\vec{x}_I$.

When some of the neuron response functions are of the form (10) and the others are of the form (9) or (11) the proof of the global convergence is obtained by combining the above strategy with the result established in Theorem 1. □

The following two examples show that when $t_{i,i} \neq 0$ for some $i = 1, 2, \ldots, q$ then a neuron response function of type (10) or (11) does not guarantee global convergence to a unique fixed point.

Example 1 *Let*

$$(x_{n+1}, y_{n+1}, z_{n+1}) = \begin{bmatrix} .7 & 0 & 0 \\ 0 & .4 & 0 \\ 0 & 0 & .4 \end{bmatrix} (x_n, y_n, z_n) + \begin{bmatrix} .2 & 0 & 0 \\ -.6 & .5 & 0 \\ .2 & -.8 & .5 \end{bmatrix} (f(x_n), f(y_n), f(z_n)) + (0, 1, 0)$$

where f is as in (10) with $a = 0$. With initial condition $\vec{x}_0 = (1, 0, 1)$ the system converges to $(\frac{2}{3}, 1.5, -1)$, while with initial condition $\vec{x}_0 = (0, 0, 1)$ the system converges to $(0, 2.5, -\frac{4}{3})$.

Example 2 *With the same matrices as above, consider the system*

$$(x_{n+1}, y_{n+1}, z_{n+1}) = \begin{bmatrix} .7 & 0 & 0 \\ 0 & .4 & 0 \\ 0 & 0 & .4 \end{bmatrix} (x_n, y_n, z_n) + \begin{bmatrix} .2 & 0 & 0 \\ -.6 & .5 & 0 \\ .2 & -.8 & .5 \end{bmatrix} (g(x_n), g(y_n), g(z_n)) + (.1, 0, 0)$$

where

$$g(x) = \begin{cases} 0 & x \leq .4 \\ 5/2(x - .4) & .4 \leq x \leq .8 \\ 1 & .8 \leq x \end{cases}$$

With initial condition $\vec{x}_0 = (.1, 0, .1)$ the system converges to $(1/3, 0, 0)$ while with $\vec{x}_0 = (1, 1, 1)$ the system converges to $(1, -1, 7/6)$.

We would like to express our gratitude to the referee for reading our paper so carefully and pointing out the pauls where changes were needed.

References

[1] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York and London, 1969.

[2] J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proc. Nat. Acad. Sci. USA* **79**, 2254-2258, 1982.

[3] M. Martelli, "Global Stability of Stationary States of Discrete Dynamical Systems", *Annales des Sciences Mathématiques du Québec*, **22,** 2, 201-212, 1998.

[4] M. Martelli, *Introduction to Discrete Dynamical Systems and Chaos*, Wiley, Interscience Series in Discrete Mathematics and Optimization, 1999.

[5] M. Martelli and B. Johnston "Global Attractivity and Forward Neural Networks," *Appl. Math. Lett.* **9,** 4, 77-83, 1996.

Asuman G. AKSOY, Mario MARTELLI
Department of Mathematics,
Claremont McKenna College,
Claremont, CA 91711, USA

Received 12.01.2001