LYAPUNOV - LIKE APPROACH FOR STUDYING GLOBAL CONVERGENCE OF NEURAL NETWORKS WITH DISCONTINUOUS OR NON-LIPSCHITZ ACTIVATIONS

C.REVATHI[1], S.RAMADEVI[2]

[1]*Research Scholar, Department of Mathematics*

[2]*Assistant Professor, Department of Mathematics*

Vivekanandha College of Arts and Sciences for Women (Autonomous), Tamilnadu, India.

ABSTRACT

The paper considers a class of additive neural networks where the neuron activations are modelled by discontinuous functions or by continuous non-lipschitz functions. Some tools are developed which enable us to apply a Lyapunov-like approach to differential equations with discontinuous right-hand side modelling the neural network dynamics. The tools include a chain rule for computing the time derivative along the neural network solutions of a non differentiable Lyapunov function, and a comparison principal for this time derivative, which yields conditions for exponential convergence or convergence in finite time. By means of Lyapunov-like approach, a general result is proved on global exponential convergence toward a unique equilibrium point of the neural network solutions, Moreover, new results on global convergence in finite time are established, which are applicable to neuron activations with jump discontinuities, or neuron activations modelled by means of continuous (non-Lipschitz) Holder functions.

Keywords: *Discontinuous neural networks; Global exponential stability; convergence in finite time*

---

1. GLOBAL EXPONENTIAL CONVERGENCE

INTRODUCTION

We address global exponential convergence toward a unique equilibrium point of the state solution of $\dot{x} = Bx + Tg(x) + I$ .... (1.1)

Let ξ be an equilibrium point of (1.1), with corresponding output equilibrium point η. We find it useful to consider the change of variables $z = x - \xi$, which transform (1.1) into the differential equation

$\dot{z} = Bz + TG(z)$ ....(1.2)

Where $G(z) = g(z + \xi) - \eta$. If $g \in \mathbb{D}$, then we also have $G \in \mathbb{D}$.

Note that (1.2) has an equilibrium point , and a corresponding output equilibrium point, which are both located at the origin. If $z(t), t \geq 0$, is a solution of (1.2), then we denote by

$\gamma^0(t) = T^{-1}(\dot{z}(t) - Bz(t)) \in \overline{co}\,[g(z(t))]$ ....(1.3)

the output solution of (1.2) corresponding to $z(t)$, which is defined for almost all $t \geq 0$.

DEFINITION: 1

We say that matrix $A \in \mathbb{R}^{n\times n}$ is Lyapunov Diagonally Stable (LDS), if there exists a positive definite diagonal matrix α = diag(α_1, . . . , α_n), such that (1/2)(αA +A^T α) is positive definite.

Suppose that −T ∈LDS and, as in [1] consider for (1.2) the (candidate) Lyapunov function

$v(z) = \sum_{i=1}^{n}\frac{1}{b_i} z_i^2 + 2c\sum_{i=n}^{n} \alpha_i \int_0^{z_i} G_i(\rho)d\,\rho$ ....(1.4)

where c > 0 is a constant, and α = diag(α_1, . . . , α_n) is a positive definite diagonal matrix such that (1/2)(α(−T) + (−T)$^T\alpha$) is positive definite.

THEOREM :1

Suppose that g ∈ 𝔻and that $-T \in LDS$. Let $z(t)$, t ≥ 0, be any solution of (1.2), and $v(t) = V(z(t))$, t ≥0. Then,

we have

$\dot{v}(t) \leq -b_m v(t)$, for almost all. t ≥0

where $b_m = min_{i=1,2,...,m}$ {bi } > 0,

and hence

$0 \leq v(t) \leq v(0)\, e^{-b_m t}, t \geq 0$

i.e., $v(t)$ converges exponentially to 0 with convergence rate b_m.

Furthermore,

we have

$\|z(t)\| \leq \sqrt{b_M v(0)}\, e^{-\frac{b_m}{2}t}\ \ t \geq 0$ ....(1.5)

where $b_M = max_{i=1,...,m}$ {bi } > 0, i.e., z(t) is exponentially convergent to 0 with convergence rate b_m/2.

Proof

We start by observing that, since each G_i is monotone non-decreasing and 0 ϵ $\overline{co}[G_i(0)]$, it easily follows that for any $z_i \epsilon\ \mathbb{R}$ we have

$0 \leq \int_0^{z_i} G_i(\rho)\, d\rho \leq z_i\, \xi_i \quad \forall \xi_i \in \overline{co}\,[G_i(z_i)]$................................ (1.6)

Let $z(t)$, t ≥ 0, be any solution of (1.2) hence z(t) is absolutely continuous on any compact interval of [0,+∞).Since V is regular at any z(t), it is possible to apply chain rule property

[Suppose that V satisfies Function $V(x): \mathbb{R}^n \to \mathbb{R}$ is:

(i) Regular in $\mathbb{R}^n$;

(ii) positive definite, i.e, we have $V(x) > 0$ for $x \neq 0$, and $V(0) = 0$;

(iii) radially unbounded, i.e., $V(x) \to +\infty$ as $\|x\| \to +\infty$

and that $x(t) : [0, +\infty) \to \mathbb{R}^n$ is absolutely continuous on any compact interval of $[0, +\infty)$.

Then, $x(t)$ and V $x(t)$: $[0, +\infty) \to \mathbb{R}$ are differentiable for almost all $t \in [0, +\infty)$, and we have

$$\frac{d}{dt} V\big(x(t)\big) = \langle \zeta, \dot{x}(t) \rangle \quad \forall\, \zeta \epsilon\, \partial V(x(t))]$$

From above we get $\dot{v}(t) = \langle \zeta, \dot{z}(t) \rangle \quad \forall\, \zeta \epsilon\, \partial V(z(t))$ ....(1.7)

Below, we extend the argument used in the proof of [1, Theorem 2] in order to prove by means of [2, property 3] that $z(t)$ is exponentially convergent to 0.

By evaluating the scalar product in (1.7), it has been proved in [1,App. IV] that there exists $\lambda > 0$ such that for almost all $t \geq 0$

We have

$$\dot{v}(t) \leq -\|z(t)\|^2 = \|B^{-1}\dot{z}(t)\|^2 - \lambda \|\gamma^0(t)\|^2 + 2cz^T(t)\alpha B \gamma^0(t)$$

$$\leq -\|z(t)\|^2 + 2cz^T(t)\alpha B \gamma^0(t) \text{ .... (1.8)}$$

Now, note from (16) it follows that

$$2cz^T(t)\alpha B \gamma^0(t) = -2c\sum_{i=1}^{n} \alpha_i b_i z_i \gamma_i^0(t)$$
$$\leq -2c\sum_{i=1}^{n} \alpha_i b_i \int_0^{z_i(t)} G_i(t)\, d\rho \leq 0. \text{ ....(1.9)}$$

Hence, we obtain

$$\dot{v}(t) \leq -\sum_{i=1}^{n} z_i^2(t) - 2c\sum_{i=1}^{n} \alpha_i b_i z_i \gamma_i^0(t)$$

$$\leq -b_m(\sum_{i=1}^{n} \frac{1}{b_i} z_i^2(t) + 2c\sum_{i=1}^{n} \alpha_i \int_0^{z_i(t)} G_i(t)\, d\rho \leq 0.$$

For almost all. $t \geq 0$. By applying [2,property 3], we conclude that

$v(t) \leq v(0)e^{-b_m t}$ for all $t \geq 0$.

Since $\alpha_i \geq 0$, and the graph of G_i is contained in the first and third quadrant, we have

$$\alpha_i \int_0^{z_i(t)} G_i(t)\, d\rho \geq 0.$$

For all $z_i \in R$. Therefore,

$$V(z) = \sum_{i=1}^{n} \frac{1}{b_i} z_i^2 + 2c\sum_{i=1}^{n} \alpha_i \int_0^{z_i} G_i(t)\, d\rho$$
$$\geq \sum_{i=1}^{n} \frac{1}{b_i} z_i^2 \geq \frac{1}{b_M} \|z\|^2$$

The result on exponential convergence of $x(t)$ to 0 thus follows from (4) of [2, property 3]

2. GLOBAL CONVERGENCE IN FINITE TIME

2.1 $LDS - matrices$

Suppose that $-T \in$ LDS, and consider for (1.2) the Lyapunov function V defined in (1.4). Let
$\theta_D = \{i \epsilon \{1, \ldots \ldots n\}: G_i$ is discontinuous at $z_i = 0\}$
And $\theta_c = \{1, \ldots \ldots n\} / \theta_D$.

In the next theorem we establish a result on global convergence in finite time of the state and output solutions of (1.2).

THEOREM 2

Suppose that $g \in \mathbb{D}$, and that $-T \in LDS$. Moreover, suppose that for any $i \in \theta_D$ we have $G_i(0^+) > 0$ and $G_i(0^-) < 0$, while for any $i \in \theta_c$ there exist $\delta_i, k_i, k_i^+ > 0, \mu_i \in (0,1)$ and $\mu_i^+ \in (0,1)$ such that

$$K_i|\rho|^{\mu_i} \le |G_i(\rho)| \le K_i^+|\rho|^{\mu_i^+}, \quad |\rho| < \delta_i \quad \text{......................................(2.0)}$$

Furthermore , suppose that

$$\mu_M = \max_{i \in \theta_c}\{\frac{2\mu_i}{1+\mu_i^+}\} < 1. \quad \text{......................................(2.1)}$$

Let $z(t), t \ge 0$, be any solution of (1.2), and $v(t) = V(z(t)), t \ge 0$.
Moreover ,let γ^0(t), for almost all, t≥0, be the output solution of (1.2) corresponding to z(t). Then , there exists $t_\delta > 0$ such that we have

$$\dot{v}(t) \le -Qv^{\mu}(t), \text{ for almost all. } t > t_\delta$$

Where $\mu \in [\mu_M, 1)$ and Q>0.as a consequence, we have

$$V(t) = 0, \quad \forall t \ge t_\phi$$

$$Z(t) = 0, \quad \forall t \ge t_\phi$$

$$\gamma^0(t) = 0, \quad \text{for almost all. } t \ge t_\phi$$

Where

$$t_\phi = t_\delta + \frac{v^{1-\mu}(t_\delta)}{Q(1-\mu)}$$

i.e.,v(t),z(t), and γ^0(t) converge to zero in finite time t_ϕ

proof

We need the following additional notations.
Since $-T \in LDS$, there exists $\alpha = \text{diag}(\alpha_1, \dots \dots \dots \dots, \alpha_n)$, where $\alpha_i > 0$ for $i = 1, \dots \dots \dots, n$, such that $(1/2)(\alpha(-T) + (-T)^T\alpha)$ is positive definite.
Let

$$\alpha_M = \max_{i \in \theta_c}\{\alpha_i\} > 0$$

We also define

$$K_m = \min_{i \in \theta_c}\{K_i\} > 0, \quad \mu_m^+ = \min_{i \in \theta_c}\{\mu_i^+\} \ge 0 \;.$$

For any $i \in \theta_D$, we let

$$m_i = min\{-G_i(0^-), G_i(0^+)\} > 0.$$

Since G_i has a finite number of discontinuities in any compact interval of $\mathbb{R}$, for any $i \in \theta_D$ there exist $\delta_i \in (0,1]$ such that G_i is a continuous function in $[-\delta_i, 0) \cup (0, \delta_i]$. For any $i \in \theta_D$ we define

$$K_i^+ = sup_{\rho \in [-\delta_i, \delta_i]}\{|G_i(\rho)|\} = max\{-G_i(-\delta_i), G_i(\delta_i)\} > 0, \text{ and}$$

we let
$K_M^+ = max\{K_1^+, \dots, K_n^+\} > 0.$

Finally, we get

$\delta = min\left\{1, \min\{\delta_1, \delta_2, \dots \dots \dots \delta_n\}, \min_{i\in\theta_D}\left\{\left(\frac{m_i}{K_M}\right)^{\frac{2}{\mu M}}\right\}\right\} > 0.$

We are now in a position to address the theorem proof, It is seen from (1.5) that for

$t > t_\delta = \frac{2}{b_m}\ln\left(\frac{\sqrt{b_M v(0)}}{g}\right)$ .... (2.2)

We have $z(t) \in [-\delta, \delta]^n$. Since by definition $\delta \in (0,1]$, then for $t > t_\delta$ we also have $|z_i(t)|^{p_2} \le |z_i(t)|^{p_1} \le 1$, for any $p_1, p_2 > 0$ such that $p_1 \le p_2$.

For $t > t_\delta$, let

$P(t) = \{i \in 1, \dots . n: z_i(t) = 0\}, \quad \theta_C^{\#}(t) = \theta_C \setminus P(t), \quad \theta_D^{\#} = \theta_C \setminus P(t).$

For any $i \in p(t)$, we have $|\gamma_i^0(t)| \ge 0 = z_i(t)$, while for any $i \in \theta_C^{\#}$ we have

$|\gamma_i^0(t)| = |G_i(z_i(t))| \ge K_i|z_i(t)|^{\mu_i}$

Moreover, for any $i \in \theta_D^{\#}(t)$

We obtain $m_i \le |\gamma_i^0(t)| = |G_i(z_i(t))| \le K_i^+$.

Suppose that $v(t)$ is differential at $t > t_\delta$.

Then, from (1.8) and (1.9) we have

$$\dot{v}(t) \le -\lambda\|\gamma^0(t)\|^2 = -\lambda\sum_{i=1}^{n}|\gamma_i^0(t)|^2$$

$$\le -\lambda\left(\sum_{i\in\theta_C^{\#}(t)} K_i^2|z_i(t)|^{2\mu_i} + \sum_{i\in\theta_C^{\#}(t)} m_i^2\right)$$

$$\le -\lambda(K_m^2\sum_{i\in\theta_C^{\#}(t)}|z_i(t)|^{2\mu_i} + \sum_{i\in\theta_D^{\#}(t)} m_i^2)$$

$$= -\lambda K_m^2(\sum_{i\in\theta_C^{\#}(t)}|z_i(t)|^{2\mu_i} + \sum_{i\in\theta_D^{\#}(t)}\frac{m_i^2}{K_m^2}) \qquad (2.3)$$

Let $\mu\epsilon(0,1)$. Since $z_i(t) \le \delta_i$, we obtain

$$v^\mu(t) = (-z^T(t)B^{-1}z(t) + 2c\sum_{i=1}^{n}\alpha_i\int_0^{z_i(t)} G_i(\rho)d\rho)^\mu$$

$$= (\sum_{i\notin P(t)}\frac{|z_i(t)|^2}{b_i} + 2c\sum_{i\notin P(t)}\alpha_i\int_0^{z_i(t)} G_i(\rho)\,d\rho)^\mu$$

$$\le (\sum_{i\notin P(t)}\frac{|z_i(t)|^2}{b_i} + 2c\sum_{i\in\theta_C^{\#}(t)}\alpha_i\int_0^{|z_i(t)|} K_i^+\rho^{\mu_i^+}d\rho + 2c\sum_{i\in\theta_D^{\#}(t)}\alpha_i\int_0^{|z_i(t)|} K_i^+ d\rho)^\mu$$

$$=(\sum_{i\notin P(t)}\frac{|z_i(t)|^2}{b_i} + 2c\sum_{i\in\theta_C^{\#}(t)}\alpha_i K_i^+\frac{|z_i(t)|^{1+\mu_i^+}}{1+\mu_i^+} + 2c\sum_{i\in\theta_D^{\#}(t)}\alpha_i K_i^+|z_i(t)|)^\mu$$

Since $(a+b)^\mu \le a^\mu + b^\mu$ for $a, b \ge 0$ and $\mu \in (0,1)$,

we have

$$v^\mu(t) \le \sum_{i\notin P(t)}\frac{|z_i(t)|^{2\mu}}{b_m^\mu} + (2c)^\mu(\sum_{i\in\theta_C^{\#}(t)}\left(\frac{\alpha_i K_i^+}{1+\mu_i^+}\right)^\mu|z_i(t)|^{\mu(1+\mu_i^+)} + \sum_{i\in\theta_D^{\#}(t)}(\alpha_i K_i^+)^\mu|z_i(t)|^\mu).$$

Recall that $\mu_i^+ > 1$, hence $2\mu > \mu(1+\mu_i^+)$ and being $|z_i(t)| \le \delta \le 1$,

it follows that $|z_i(t)|^{2\mu} \le |z_i(t)|^{2\mu} \le |z_i(t)|^{\mu(1+\mu_i^+)}$ and $|z_i(t)|^{2\mu} \le |z_i(t)|^\mu$.

Therefore,

$$v^{\mu}(t) \leq \sum_{i\epsilon\,\theta_C^{\#}(t)} \frac{|z_i(t)|^{\mu(1+\mu_i^+)}}{b_m^{\mu}} + \sum_{i\epsilon\,\theta_D^{\#}(t)} \frac{|z_i(t)|^{\mu}}{b_m^{\mu}} + \left(\frac{2c\alpha_M K_M^+}{1+\mu_m^+}\right)^{\mu} \sum_{i\epsilon\,\theta_C^{\#}(t)} |z_i(t)|^{\mu(1+\mu_i^+} + (2c\alpha_M K_M^+)^{\mu} \sum_{i\epsilon\,\theta_D^{\#}(t)} |z_i(t)|^{\mu}$$

$$\leq \left(\frac{1}{b_m^{\mu}} + \left(\frac{2c\,\alpha_M K_M^+}{1+\mu_m^+}\right)^{\mu}\right) \sum_{i\epsilon\,\theta_C^{\#}(t)} |z_i(t)|^{\mu(1+\mu_i^+)} + \left(\frac{1}{b_m^{\mu}} + (2c\,\alpha_M K_M^+)^{\mu}\right) \sum_{i\epsilon\,\theta_D^{\#}(t)} |z_i(t)|^{\mu}$$

If $\mu \in [\mu_m, 1)$, then from (2.1)
it follows that $\mu(1+\mu_i^+) \geq 2\mu_i$ and hence

$$v^{\mu}(t) \leq \left(\frac{1}{b_m^{\mu}} + \left(\frac{2c\alpha_M K_M^+}{1+\mu_m^+}\right)^{\mu}\right) \sum_{i\epsilon\,\theta_C^{\#}(t)} |z_i(t)|^{2\mu_i} + \left(\frac{1}{b_m^{\mu}} + (2c\,\alpha_M K_M^+)^{\mu}\right) \sum_{i\epsilon\,\theta_D^{\#}(t)} |z_i(t)|^{\mu}$$

$$\leq \left(\frac{1}{b_m^{\mu}} + (2c\,\alpha_M K_M^+)^{\mu}\right) \times \left(\sum_{i\epsilon\,\theta_C^{\#}(t)} |z_i(t)|^{2\mu_i} + \sum_{i\epsilon\,\theta_D^{\#}(t)} |z_i(t)|^{\mu}\right) \text{........................ (2.4)}$$

Where we have taken into account that $(2c\alpha_M K_M^+) \,/\, (1+\mu_m^+) \leq (2c\alpha_M K_M^+)$.

Now let us show that for any $i \in \theta_D$ we have $|z_i(t)|^{\mu} \leq m_i^2 / K_m^2$.

There are the following two possibilities.

(a) $m_i^2 / K_m^2 \geq 1$. In this case, since $|z_i(t)| \leq \delta \leq 1$ we have $|z_i(t)|^{\mu} \leq 1 \leq \frac{m_i^2}{K_m^2}$.

(b) $m_i^2 / K_m^2 < 1$. Then by the definition of δ we have $|z_i(t)|^{\mu} \leq \delta^{\mu} \leq \left(\frac{m_i}{K_m}\right)^{\frac{2\mu}{\mu_M}} \leq \frac{m_i^2}{K_m^2}$

Where we have considered that $\mu / \mu_M \geq 1$.

Therefore, from (2.4) we have

$$v^{\mu}(t) \leq \left(\frac{1}{b_m^{\mu}} + (2c\alpha_M K_M^+)^{\mu}\right) \times \left(\sum_{i\epsilon\,\theta_C^{\#}(t)} |z_i(t)|^{2\mu_i} + \sum_{i\epsilon\,\theta_D^{\#}(t)} \frac{m_i^2}{K_m^2}\right) \text{........................ (2.5)}$$

Eqs. (2.3) and (2.5) thus yield

$$\dot{v}(t) \leq -\lambda K_m^2 \left(\sum_{i\epsilon\theta_C^{\#}(t)} |z_i(t)|^{2\mu_i} + \sum_{i\epsilon\,\theta_D^{\#}(t)} \frac{m_i^2}{K_m^2}\right)$$

$$= -\frac{\lambda K_m^2}{\frac{1}{b_m^{\mu}} + (2c\alpha_M K_M^+)^{\mu}} \left(\frac{1}{b_m^{\mu}} + (2c\,\alpha_M K_M^+)^{\mu}\right) \times \left(\sum_{i\epsilon\theta_C^{\#}(t)} |z_i(t)|^{2\mu_i} + \sum_{i\epsilon\,\theta_D^{\#}(t)} \frac{m_i^2}{K_m^2}\right)$$

$$\leq -\frac{\lambda K_m^2}{\frac{1}{b_m^{\mu}} + (2c\alpha_M K_M^+)^{\mu}}\, v^{\mu}(t). \qquad \text{...(2.6)}$$

In conclusion, we have shown that a.a $t > t_g$ we have $\dot{v}(t) \leq -Q v^{\mu}(t)$

Where

$$Q = \frac{\lambda K_m^2}{\frac{1}{b_m^\mu} + (2c\,\alpha_M K_M^+)^\mu} > 0$$

By applying the result in point (b) of[2, property 4], we obtain that $v(t) = 0$ and $z(t) = 0$ for $t \geq t_\emptyset$, where

$$t_\emptyset = t_\delta + \frac{v^{1-\mu}(t_\delta)}{Q\,(1-\mu)} \qquad \text{......................} \qquad (2.7)$$

And t_δ is given in (2.2). Finally, (1.3) implies that $\gamma^\circ(t) = 0$ for almost all $t \geq t_\emptyset$.

CONCLUSION

The paper has proved results on global exponential convergence toward a unique equilibrium point, and global convergence in finite time, for a class of additive neural networks possessing discontinuous neuron activations or continuous non-Lipschitz neuron activations. The results are of potential interest in view of the neural network applications for solving global optimization problems in real time, where global convergence toward an equilibrium point, fast convergence speed and the ability to quantitatively estimate the convergence time, are of crucial importance. The results have been proved by means of a generalized Lyapunov-like approach, which has been developed in the paper, and is suitable for addressing convergence of nonsmooth dynamical systems described by differential equations with discontinuous right-hand side. An important open question is whether the results on global convergence here obtained may be extended to more general neural network models incorporating the presence of a delay in the neuron interconnections. This topic goes beyond the scope of the present paper and will constitute a challenging issue for future investigations.

REFERENCES

[1] M. Forti, P. Nistri, Global convergence of neural networks with discontinuous neuron actiations, IEEE Trans. Circuits Syst.I 50 (11) (2003) 1421-1435
[2] M. Forti, M. Grazzini, P. Nistri, L. Pancioni, Generalized Lyapunov for convergence of neural networks with discontinuous or non-Lipschitz activations Physica D 214 (2006) 88-99
[3] M. Hirsch, Convergent activation dynamics in continuous time networks, Neural Netw. 2 (1989) 331–349.
[4] C.M. Marcus, R.M. Westervelt, Stability of analog neural network with delay, Phys. Rev. A 39 (1989) 347–359.
[5] E. Kaszkurewicz, A. Bhaya, On a class of globally stable neural circuits, IEEE Trans. Circuits Syst. I 41 (1994) 171–174.