Comments and Corrections

Correction to "Convergence and Rate Analysis of Neural Networks for Sparse Approximation"

Aurèle Balavoine, *Student Member, IEEE*, Justin Romberg, *Senior Member, IEEE*, and Christopher J. Rozell, *Senior Member, IEEE*

***Abstract*— This document provides a correction to the proof of the theorem establishing the exponential speed of convergence of the Locally Competitive Algorithm (LCA) in the paper "Convergence and Rate Analysis of Neural Networks for Sparse Approximation."**

***Index Terms*— Exponential convergence, global stability, locally competitive algorithm, Lyapunov function, nonsmooth objective, sparse approximation.**

I. Introduction

The paper [1] studies the convergence properties of a neural network designed to solve a class of nonsmooth and nonlinear optimization programs of the form

$$V\,(a(t)) = \frac{1}{2}\,\|y - \Phi a\|_2^2 + \lambda \sum_{n=1}^{N} C(a_n). \tag{1}$$

The network proposed is the Locally Competitive Algorithm (LCA) defined by the following differential equation:

$$\begin{aligned} \tau\dot{u}(t) &= -u(t) - (\Phi^T\Phi - I)\,a(t) + \Phi^T y \\ a(t) &= T_\lambda(u(t)). \end{aligned} \tag{3}$$

Note that we keep the same notations and equation numbers as in [1].

The two main contributions of [1] are: first to provide a proof of global asymptotic convergence for the Locally Competitive Algorithm (LCA) to the solution a^* that minimizes (1) for a wide class of cost functions $C(\cdot)$ and second, to provide a proof for the exponential rate of convergence of the network. Despite extensive work on neural network convergence in the literature, the analysis of the LCA network proves challenging due to a nonsmooth, unbounded and not strictly increasing activation function, as well as a potentially singular interconnection matrix.

It has come to the authors' attention that the proof of Theorem 3 (regarding the exponential convergence rate) contains an error. This correction provides a new statement of the theorem with a proof that corrects the error.

II. Exponential Convergence

Theorem 3 in [1] guarantees an exponential rate of convergence of the LCA neural network under some condition on the interconnection matrix that involves two constants. The constant α bounds the derivative of $f(\cdot)$ in (5)

$$|f'(u_n(t))| \leq \alpha \qquad \forall t \geq 0, \forall n = 1, \ldots, N. \tag{14}$$

Letting Γ_* be the active set of the solution a^* to (1), we define $\widetilde{\Gamma} = \Gamma \cup \Gamma_*$. Then, the second constant δ is the smallest positive constant such that for any vector x in $\mathbb{R}^N$ with active set $\widetilde{\Gamma}$, we have

$$(1-\delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1+\delta)\|x\|_2^2. \tag{15}$$

The constant δ depends on the singular values of the matrix $\Phi_{\widetilde{\Gamma}}$ and on the sequence of active sets visited by the system. The mistake in the proof of Theorem 3 in [1] happens when splitting the vector $\Phi(\widetilde{u} - \widetilde{a})$ into its components on the active set $\widetilde{\Gamma}$ and its complement $\widetilde{\Gamma}^c$. Since those sets index the columns of Φ and not its rows, the resulting vectors $\Phi_{\widetilde{\Gamma}}(\widetilde{u}_{\widetilde{\Gamma}} - \widetilde{a})$ and $\Phi_{\widetilde{\Gamma}^c}\widetilde{u}_{\widetilde{\Gamma}^c}$ are not orthogonal and Pythagoras' theorem does not apply. The following theorem is a revised statement of Theorem 3 in [1], with a simpler condition between the constants α and δ and a more favorable convergence speed.

Revised Theorem 3: Under conditions (6) on the activation function in (5), and provided that the constants α and δ defined in (14) and (15) satisfy

$$\alpha\delta < 1 \tag{16}$$

the LCA network defined in (3) is globally exponentially convergent to a unique equilibrium with convergence speed

$$c = \frac{1-\alpha\delta}{\tau}.$$

III. Proof of Theorem II

Contrary to the approach taken in [1] that made use of a Lure-Postnikov-type Lyapunov function, we directly analyze the energy function

$$E(t) = \frac{1}{2}\left\|u(t) - u^*\right\|_2^2 \tag{E1}$$

for an arbitrary fixed point u^* of (3), and show that it converges exponentially fast to zero. We begin by analyzing $E(t)$ on the set of indices $\widetilde{\Gamma}$ and obtain a convergence result for the outputs. We then use this result to prove convergence of the entire state vector.

Proof: We again define new output and state variables in terms of the distance from the fixed point u^* as follows:

$$\begin{aligned} \widetilde{u}_n(t) &= u_n(t) - u_n^*, \\ \widetilde{a}_n(t) &= a_n(t) - a_n^* = T_\lambda(\widetilde{u}_n(t) + u_n^*) - T_\lambda(u_n^*). \end{aligned} \tag{18}$$

Using the fact that u^* is a fixed point, we set $\dot{u}^*(t) = 0$ in (3) and rewrite the dynamics in terms of the new variables

$$\tau\dot{\widetilde{u}}(t) = -\widetilde{u}(t) - (\Phi^T\Phi - I)\widetilde{a}(t). \tag{19}$$

Manuscript received April 22, 2013; revised September 3, 2013; accepted November 21, 2013. Date of publication December 6, 2013; date of current version July 14, 2014. This work was supported by the National Science Foundation under Grant CCF-0905346.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: faurele.balavoine@gatech.edu; crozell@gatech.edu; jromg@gatech.edu).

Digital Object Identifier 10.1109/TNNLS.2013.2292700

We start by showing that the partial energy function $E_{\widetilde{\Gamma}}(t) = 1/2 \left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2^2$ converges exponentially fast, where $\widetilde{\Gamma}$ is the support of $\widetilde{a}(t)$. Using the chain rule, the time derivative of $E_{\widetilde{\Gamma}}(t)$ along the network trajectory is

$$\begin{aligned} \tau \dot{E}_{\widetilde{\Gamma}}(t) &= \tau \dot{\widetilde{u}}_{\widetilde{\Gamma}}(t)^T \widetilde{u}_{\widetilde{\Gamma}}(t) \\ &= -\widetilde{u}_{\widetilde{\Gamma}}(t)^T \left(\widetilde{u}_{\widetilde{\Gamma}}(t) + \left(\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right) \widetilde{a}_{\widetilde{\Gamma}}(t) \right) \\ &= -\left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2^2 - \widetilde{u}_{\widetilde{\Gamma}}(t)^T \left(\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right) \widetilde{a}_{\widetilde{\Gamma}}(t). \end{aligned}$$

Since $\widetilde{a}_{\widetilde{\Gamma}}$ is supported on $\widetilde{\Gamma}$, assumption (15) implies that the eigenvalues of $\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}}$ lie between $(1-\delta)$ and $(1+\delta)$ and so

$$\begin{aligned} \left\| \left(\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right) \widetilde{a}_{\widetilde{\Gamma}} \right\|_2 &\le \left\| \Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right\|_2 \|\widetilde{a}_{\widetilde{\Gamma}}\|_2 \\ &\le \max\{(1+\delta)-1,\ 1-(1-\delta)\} \|\widetilde{a}_{\widetilde{\Gamma}}\|_2 \\ &= \delta \|\widetilde{a}_{\widetilde{\Gamma}}\|_2. \end{aligned}$$

Finally, property (iii) of Lemma 1 in [1] states that for any set $\mathcal{T}$, $\left\| \widetilde{a}_{\mathcal{T}} \right\|_2^2 \le \alpha^2 \left\| \widetilde{u}_{\mathcal{T}} \right\|_2^2$. Using the Cauchy-Schwartz inequality and putting everything together, we obtain

$$\begin{aligned} \left| \widetilde{u}_{\widetilde{\Gamma}}^T \left(\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right) \widetilde{a}_{\widetilde{\Gamma}} \right| &\le \left\| \widetilde{u}_{\widetilde{\Gamma}} \right\|_2 \left\| \left(\Phi_{\widetilde{\Gamma}}^T \Phi_{\widetilde{\Gamma}} - I_{\widetilde{\Gamma}} \right) \widetilde{a}_{\widetilde{\Gamma}} \right\|_2 \\ &\le \left\| \widetilde{u}_{\widetilde{\Gamma}} \right\|_2 \delta \left\| \widetilde{a}_{\widetilde{\Gamma}} \right\|_2 \\ &\le \alpha\delta \left\| \widetilde{u}_{\widetilde{\Gamma}} \right\|_2^2. \end{aligned}$$

As a consequence, the time derivative of the partial energy function satisfies

$$\begin{aligned} \tau \dot{E}_{\widetilde{\Gamma}}(t) &\le -\left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2^2 + \alpha\delta \left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2^2 \\ &\le -2(1-\alpha\delta) E_{\widetilde{\Gamma}}(t). \end{aligned}$$

Using Gronwall's inequality on the interval $[t_k, t_{k+1}]$ where $\widetilde{\Gamma}$ is constant yields

$$E_{\widetilde{\Gamma}}(t) = \frac{1}{2} \left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2^2 \le \frac{1}{2} \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2^2 e^{-2(1-\alpha\delta)(t-t_k)/\tau}.$$

Since $\|\widetilde{a}(t)\|_2 \le \alpha \left\| \widetilde{u}_{\widetilde{\Gamma}}(t) \right\|_2$, $\forall t \in [t_k, t_{k+1}]$

$$\|\widetilde{a}(t)\|_2 \le \alpha \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau}. \tag{E2}$$

Using this result on the output, we now prove that the state $u(t)$ converges exponentially fast. For this we write the solution to (19) as follows $\forall t \in [t_k, t_{k+1}]$:

$$\begin{aligned} \widetilde{u}(t) &= e^{-(t-t_k)/\tau} \widetilde{u}(t_k) \\ &\quad + e^{-(t-t_k)/\tau} \int_{t_k}^{t} e^{(\nu-t_k)/\tau} \left(I - \Phi^T \Phi \right) \widetilde{a}(\nu) d\nu. \end{aligned}$$

Denoting by $h(t)$ the second term in the right-hand side, and plugging in (E2), we bound the norm of $h(t)$ by

$$\begin{aligned} \|h(t)\|_2 &\le e^{-(t-t_k)/\tau} \int_{t_k}^{t} e^{(\nu-t_k)/\tau} \left\| \left(\Phi^T \Phi - I \right) \widetilde{a}(\nu) \right\|_2 d\nu \\ &\le e^{-(t-t_k)/\tau} \int_{t_k}^{t} \underbrace{\left\| \Phi^T \Phi - I \right\|_2}_{=C_1} e^{(\nu-t_k)/\tau} \left\| \widetilde{a}_{\widetilde{\Gamma}}(\nu) \right\|_2 d\nu \\ &\le e^{-(t-t_k)/\tau} \int_{t_k}^{t} C_1 \alpha \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2 e^{\alpha\delta(\nu-t_k)/\tau} d\nu \\ &= \frac{C_1 \tau}{\delta} \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2 e^{-(t-t_k)/\tau} \left[e^{\alpha\delta(t-t_k)/\tau} - 1 \right] \\ &= C_2 \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2 e^{-(t-t_k)/\tau} \left[e^{\alpha\delta(t-t_k)/\tau} - 1 \right] \\ &\le C_2 \left\| \widetilde{u}_{\widetilde{\Gamma}}(t_k) \right\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau} \\ &\le C_2 \|\widetilde{u}(t_k)\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau} \end{aligned}$$

where $C_2 = (\|\Phi^T \Phi - I\|_2 \tau/\delta)$. We plug this bound in the expression for $\widetilde{u}(t)$ to get $\forall t \in [t_k, t_{k+1}]$

$$\begin{aligned} \|\widetilde{u}(t)\|_2 &= \left\| e^{-(t-t_k)/\tau} \widetilde{u}(t_k) + h(t) \right\|_2 \\ &\le \|\widetilde{u}(t_k)\|_2 e^{-(t-t_k)/\tau} + \|h(t)\|_2 \\ &\le \|\widetilde{u}(t_k)\|_2 e^{-(t-t_k)/\tau} \\ &\quad + C_2 \|\widetilde{u}(t_k)\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau} \\ &\le (1 + C_2) \|\widetilde{u}(t_k)\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau} \\ &= C_3 \|\widetilde{u}(t_k)\|_2 e^{-(1-\alpha\delta)(t-t_k)/\tau} \end{aligned}$$

where $C_3 = 1 + C_2$. Since $\|\widetilde{u}(t)\|_2$ is continuous for all time t, it is easy to show (by induction on t_k) that

$$\|\widetilde{u}(t)\|_2 \le e^{-(1-\alpha\delta)t/\tau} C_3 \|\widetilde{u}(0)\|_2.$$

This last inequality shows that the state variable converges exponentially fast to a unique fixed point u^* with convergence speed $(1-\alpha\delta)/\tau$. ∎

References

[1] A. Balavoine, J. Romberg, and C. J. Rozell, "Convergence and rate analysis of neural networks for sparse approximation," *IEEE Trans. Neural Netw.*, vol. 23, no. 9, pp. 1377–1389, Sep. 2012.