Global exponential stability of discrete-time neural networks for constrained quadratic optimization

K.C. Tan\textsuperscript{a,*}, H.J. Tang\textsuperscript{a}, Z. Yi\textsuperscript{b}

\textsuperscript{a}Department of Electrical and Computer Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 117576, Singapore
\textsuperscript{b}School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610064, People’s Republic of China

Abstract

A class of discrete-time recurrent neural networks for solving quadratic optimization problems over bound constraints is studied. The regularity and completeness of the network are discussed. The network is proven to be globally exponentially stable (GES) under some mild conditions. The analysis of GES extends the existing stability results for discrete-time recurrent networks. A simulation example is included to validate the theoretical results obtained in this letter.

\textsuperscript{c}\textsuperscript{c} © 2003 Elsevier B.V. All rights reserved.

Keywords: Discrete-time neural networks; Global exponential stability; Quadratic optimization

1. Introduction

Since the early work of Tank and Hopfield [10] and Kennedy and Chua [5], the construction of a recurrent neural network (RNN) for solving linear and nonlinear programming has become an active research topic in the field of neural networks [1,3,7–9,11]. The study of nonlinear optimization is also valuable in the sense that it allows one to apply its results to variational inequality problems, since there is a two-way bridge connecting both issues [2,4]. In the recent literature, there exist a few RNN models for solving nonlinear optimization problems over convex constraints (see, e.g., [8,6]). In [8], a discrete-time RNN model was proposed to solve the strictly

\textsuperscript{*} Corresponding author. Tel.: +65-874-2127; fax: +65-779-91103.
\textsuperscript{E-mail address: eletankc@nus.edu.sg (K.C. Tan).}

0925-2312/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0925-2312(03)00442-9
convex quadratic optimization problem with bound constraints. Sufficient conditions for
the GES of the model and several corresponding neuron updating rules were presented.
In [6], a continuous-time RNN was presented for solving bound-constrained nonlinear
differentiable optimization problems.

Quadratic optimization is an important case of nonlinear optimization problems. We
restrict our discussion to bound-constrained quadratic optimization problems, where the
objective function \( E(x) : \mathbb{R}^n \to \mathbb{R} \) is assumed to be differentiable but not necessarily
convex, which is to be minimized over the vector \( x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n \) subject to
bound constraints as described by \( \min \{ E(x) | x \in \Omega \} \), or in an explicit form
\[
\min \left\{ \frac{1}{2} x^T A x + x^T b \mid x \in \Omega \right\},
\]
where \( A \) is a matrix, \( b \) is a vector and the superscript \( T \) denotes the transpose operator.
\( \Omega \) is defined by \( \Omega = \{ x \in \mathbb{R}^n | x_i \in [c_i, d_i], i = 1, \ldots, n \} \), where \( c_i \) and \( d_i \) are constants
such that \( c_i \leq d_i (i = 1, \ldots, n) \).

In this letter, we present new theoretical results for a class of discrete-time RNN for
solving quadratic optimization,
\[
(N1) \quad x(k) = f(x(k - 1) - \alpha(Ax(k - 1) + b))
\]
for all \( k \geq 1 \), where \( \alpha > 0 \) is a constant.

The vector-valued activation function \( f(x) = (f_1(x_1), \ldots, f_n(x_n))^T : \mathbb{R}^n \to \mathbb{R}^n \) is
defined as
\[
f_i(x_i) = \begin{cases} 
    c_i & \text{if } x_i < c_i, \\
    d_i & \text{if } x_i > d_i, \\
    x_i & \text{otherwise},
\end{cases}
\]
for \( i = 1, \ldots, n \). It should be noted that all the \( f(x) \) in this letter will take the form
of (3).

Since a neural network with global stability implies that the neural network has
a unique equilibrium and all solutions of the network converge to the equilibrium,
the analysis of global stability is important for solving optimization problems [12].
For strictly convex quadratic optimization problems with bound constraints, the model
(N1) is proven to be GES in this letter, which does not require additional conditions
to be imposed on the matrix. In addition, the lower bound of the global exponential
convergence rate is also obtained in this letter. The advantages of the proposed
model lie in its simplicity in numerical simulation and digital implementation over its
continuous-time counterpart.

2. Preliminaries

**Definition 1.** A vector \( x^e \in \mathbb{R}^n \) is said to be an equilibrium point of (2) if and only if
\( x^e = f(x^e - \alpha(Ax^e + b)) \).

For each \( x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n \), we denote \( \|x\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{x^T x} \).
**Definition 2.** Network (2) is said to be globally exponentially stable (GES) if it possesses a unique equilibrium \( x^e \) and there exist constants \( \eta > 0 \) and \( \mu \geq 1 \) such that 

\[
\| x(k) - x^e \| \leq \mu \| x(0) - x^e \| \exp(-\eta k), \text{ for all } k \geq 1.
\]

The constant \( \eta \) is the lower bound of the convergence rate for the network.

**Lemma 1.** \([f(x) - f(y)]^T[f(x) - f(y)] \leq [x - y]^T[x - y], \text{ for all } x, y \in \mathbb{R}^n.\)

The proof of this lemma can be derived directly from the definition of the activation function \( f(x) \).

### 3. Regularity and completeness

We define the solution set \( \Omega^* = \{ x^* \in \Omega | E(x) \geq E(x^*), \forall x \in \Omega \} \) for the minimization problem (1) and the equilibrium set \( \Omega^e = \{ x \in \mathbb{R}^n | x = f(x - \nabla E(x)) \} \) for the RNN model (2). It is assumed \( \Omega^* \neq \emptyset \). It can be seen that \( \Omega^* \) and \( \Omega^e \) are both closed. If \( \Omega^* \subseteq \Omega^e \), then it is said that the network is regular. If \( \Omega^* = \Omega^e \), then the network is said to be complete. Obviously, regular and complete properties indicate the capability of the network for solving constrained optimization problems [6]. In [6], a continuous-time recurrent network was proven to be regular or even complete when the objective function is convex on \( \mathbb{R}^n \). The analysis procedure and proof for the discrete-time recurrent neural network is analogous to the continuous-time network.

**Theorem 1.** The discrete-time RNN model (2) is regular, i.e., \( \Omega^* \subseteq \Omega^e \). Furthermore, if the objective function \( E(x) \) is convex on \( \mathbb{R}^n \), then the network is complete, i.e., \( \Omega^* = \Omega^e \).

**Proof.** Let \( x^* \in \mathbb{R}^n \) be a minimum of the minimization problem (1), if and only if it satisfies the first-order necessary optimum condition,

\[
\nabla E(x^*)^T(y - x^*) \geq 0, \quad \forall y \in \Omega.
\]

This inequality coincides with the variational inequality problem. According to the fundamental results in [2] (see also [4]), any solution to the variational inequality problem is equivalent to be an equilibrium of the recurrent system (2). This proves that \( \Omega^* \subseteq \Omega^e \).

In order to prove that the RNN model is complete, it remains to show \( \Omega^e \subseteq \Omega^* \). Let \( x^e \) be an equilibrium of the system (2), i.e., \( x^e = f(x^e - \alpha \nabla E(x^e)) \), which is equivalent to \( (x - x^e)^T z \nabla E(x^e) \geq 0, \forall x \in \Omega \). Since \( \alpha > 0 \), we obtain the first-order optimum condition as (4) and thus \( x^e \in \Omega^e \). This completes the proof.

### 4. Global exponential stability

For the strictly convex quadratic optimization problem (1), where \( A \) is symmetric positive definite, it has a unique minimum \( x^* \in \Omega \) and \( \Omega^* = \Omega^e = \{ x^* \} \). The following
Lemma about the existence and uniqueness of the equilibrium point was proven in [8], which can also be derived from the completeness of the network (2) by Theorem 1.

**Lemma 2.** For each \( \alpha > 0 \), the network (2) has a unique equilibrium point and this equilibrium point is the minimum of the quadratic optimization problem (1).

In the following, we present some new results on the analysis of GES. Under some mild conditions, the network (2) is guaranteed to be globally exponentially convergent.

Let \( \lambda_i > 0 (i = 1, \ldots, n) \) be the eigenvalues of \( A \), which are all positive values since \( A \) is a symmetric\(^1\) and positive definite matrix. Denote \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) as the smallest and the largest eigenvalue of \( A \), respectively. Define a continuous function \( \gamma(\alpha) = \max_{1 \leq i \leq n} |1 - \alpha \lambda_i| \) for \( \alpha > 0 \). It is clear that \( \gamma(\alpha) > 0 \) for all \( \alpha > 0 \). This function plays an important role in the estimation of the convergence rate for the network. Denote \( \alpha^* = 2/(\lambda_{\text{max}} + \lambda_{\text{min}}) \), obviously, \( \alpha^* > 0 \).

**Theorem 2.** For all \( \alpha > 0 \), we have

(i) \[ r(\alpha) = \begin{cases} 1 - \lambda_{\text{min}} \alpha, & 0 < \alpha \leq \alpha^*, \\ \lambda_{\text{max}} \alpha - 1, & \alpha^* < \alpha < +\infty. \end{cases} \]  

(ii) \( r(\alpha) < 1 \) if and only if \( \alpha \in \left( 0, \frac{2}{\lambda_{\text{max}}} \right) \).

(iii) \[ \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\lambda_{\text{max}} + \lambda_{\text{min}}} = r(\alpha^*) \leq r(\alpha). \]

**Proof.** Based on the definition of \( r(\alpha) \), it is easy to see that

\[ r(\alpha) = \max \{|1 - \alpha \lambda_{\text{min}}|, |1 - \alpha \lambda_{\text{max}}|\}. \]

Since

\[ |1 - \alpha^* \lambda_{\text{min}}| = \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\lambda_{\text{max}} + \lambda_{\text{min}}} = |1 - \alpha^* \lambda_{\text{max}}|, \]

it is obvious that (5) can be obtained via a simple calculation. Using the expression (8), the rest of the theorem can be easily derived (details are omitted due to the page limit). An intuitive explanation for Theorem 2 is shown in Fig. 1. It can be seen that \( \alpha^* \) gives the minimum of \( r(\alpha) \), and \( r(\alpha) < 1 \) if and only if \( 0 < \alpha < 2/\lambda_{\text{max}} \). \( \square \)

**Theorem 3.** For each \( \alpha \) in the range of \( 0 < \alpha < 2/\lambda_{\text{max}} \). Network (2) is globally exponentially stable with a lower bound of convergence rate \( \eta(\alpha) = -\ln r(\alpha) > 0 \).

\(^1\) If \( A \) is not symmetric, it can be decomposed by its symmetric and antisymmetric components as \( A = [(A + A^T)/2] + [(A - A^T)/2] \). Since \( \forall x, x^T A x = 0 \) for any antisymmetric matrix, i.e., the antisymmetric component contributes zero to the quadratic form, (1) can simply be reformulated by replacing \( A \) with its symmetric component as stated in [1].
Proof. Since \( \lambda_i(i=1,\ldots,n) \) are all the eigenvalues of \( A \), it is easy to see that \( (1-\alpha \lambda_i)^2(i=1,\ldots,n) \) are all the eigenvalues of the matrix \( (I-\alpha A)^2 \).

For any \( x(0) \in \mathbb{R}^n \), let \( x(k) \) be the solution of (2) starting from \( x(0) \). By Lemma 2, the network has a unique equilibrium \( x^* \). Using Lemma 1, it follows that for all \( k \geq 1 \),

\[
\|x(k) - x^*\|^2 \\
=\|f(x(k-1) - \alpha(Ax(k-1) + b)) - f(x^* - \alpha(Ax^* + b))\|^2 \\
\leq \|(I - \alpha A)(x(k - 1) - x^*)\|^2 \\
=\|x(k - 1) - x^*\|^2 \|x(k - 1) - x^*\|^2 \\
= r(\alpha)^2 \|x(k - 1) - x^*\|^2 
\]

That is \( \|x(k) - x^*\| \leq r(\alpha)\|x(k - 1) - x^*\| \leq r(\alpha)^k \|x(0)\| - x^*\|. \) It implies that \( \|x(k) - x^*\| \leq \exp(-\eta(\alpha)k)\|x(0) - x^*\| \), where \( \eta(\alpha) \) is given as \( \eta(\alpha) = -\ln r(\alpha) \). By Theorem 2, we have \( \eta(\alpha) > 0 \) for each \( \alpha \), since \( 0 < \alpha < 2/\lambda_{\max} \). This completes the proof that the network is GES and \( \eta(\alpha) > 0 \) is a lower bound of the global exponential convergence rate. \( \square \)

5. Discussions and simulations

In order to guarantee that the network has a large convergence rate, one needs to choose the parameter \( \alpha \) in the network (N1) such that \( r(\alpha) \) is as small as possible. It can be observed from Theorem 2 that \( \alpha^* \) gives the minimum of \( r(\alpha) \), i.e., we can set \( \alpha = \alpha^* \) in (N1). In this way, the network (N1) will have a large lower bound of global exponential convergence rate. However, this approach requires the calculation of the largest and the smallest eigenvalues of the matrix \( A \) in advance, which introduces additional computational effort, especially when the dimension of the matrix \( A \) is very
large. Since \( \text{tra}(A) = \sum_{i=1}^{n} a_{ii} \geq \lambda_{\text{max}} \) when \( n \geq 2 \), we have \( 2/\text{tra}(A) \in (0, 2/\lambda_{\text{max}}) \), for \( n \geq 2 \). As \( n = 1 \) is a trivial and special case, only the situation of \( n \geq 2 \) is considered.

Therefore, we can choose the parameter \( x = 2/\text{tra}(A) \) in the network \((N1)\) as given below

\[
(N2) \quad x(k) = f \left( x(k-1) - \frac{2}{\text{tra}(A)} (Ax(k-1) + b) \right), \tag{11}
\]

for all \( k \geq 1 \). Obviously, the network has a lower bound convergence rate \( \eta = -\ln r \) 

\( (2/\text{tra}(A)) > 0 \).

Network \((N2)\) is relatively easy to be implemented and is recommended for solving problem \((1)\). In order to further improve the numerical stability and convergence speed of the network, the preconditioning technique in [8] can be used. The aim of the preconditioning technique is to minimize the difference between the largest and the smallest eigenvalues of the matrix \( A \). The method is given as follows: A diagonal matrix \( P \) with the diagonal elements \( p_{ii} = 1/\sqrt{a_{ii}}(i = 1, \ldots, n) \) is first defined. Some transformations are then performed such that \( \tilde{A} = PAP, \ \tilde{b} = Pb, \ \tilde{c} = P^{-1}c, \ \tilde{d} = P^{-1}d \).

Since \( \text{tra}(\tilde{A}) = n \), the network \((N2)\) becomes

\[
(N3) \quad \left\{ \begin{array}{l}
\tilde{x}(k) = f(\tilde{x}(k-1) - \frac{2}{n} (\tilde{A}\tilde{x}(k-1) + \tilde{b})), \\
x(k) = P\tilde{x}(k).
\end{array} \right. \tag{12}
\]

**Example.** Consider the quadratic optimization problem \((1)\) with

\[
A = \begin{bmatrix}
0.180 & 0.648 & 0.288 \\
0.648 & 2.880 & 0.720 \\
0.288 & 0.720 & 0.720
\end{bmatrix},
\]

\( b = [0.4 \ 0.2 \ 0.3]^T \) and \( c = -d = [20 \ -20 \ -20]^T \).

The author in [8] illustrated that the recurrent network \((N1)\) is globally convergent for this problem but indicated it is not known whether the network is globally exponentially convergent in a priori. Unlike [8], the network is guaranteed to be globally exponentially convergent by the conditions obtained in this letter.

By applying the precondition technique, we have

\[
\tilde{A} = \begin{bmatrix}
1.0000 & 0.9000 & 0.8000 \\
0.9000 & 1.0000 & 0.5000 \\
0.8000 & 0.5000 & 1.0000
\end{bmatrix}, \quad \tilde{b} = \begin{bmatrix}
0.9428 \\
0.1179 \\
0.3536
\end{bmatrix}, \quad \tilde{c} = \begin{bmatrix}
-8.4853 \\
-33.9411 \\
-16.9706
\end{bmatrix}
\]

and \( \tilde{d} = \begin{bmatrix}
8.4853 \\
33.9411 \\
16.9706
\end{bmatrix} \).
Fig. 2. The convergence trace for each component of the trajectory starting from $x(0)$.

Fig. 3. The global exponential convergence for various trajectories in $\mathbb{R}^3$ space.

We set an error tolerance $\varepsilon = 0.0001$ and $x(0) = [15; 10; -10]$ as the starting point for the simulation. It takes 46 iterations to converge exponentially to the minimum $x^* = [-20.0000; 3.3796; 4.2038]$ for the above error tolerance. The convergence trace for each of the optimization variables in $x$ starting from $x(0)$ is shown in Fig. 2. We then perform the simulation by randomly selecting 30 points in $\mathbb{R}^3$ as the initial points of the network (N3). As can be seen from Fig. 3, all the trajectories have globally exponentially converged to the minimum $x^*$, as desired.
6. Conclusions

We have studied a class of discrete-time recurrent neural networks for constrained quadratic optimization problems. The regularity and completeness of the network have been discussed. The analysis of global exponential stability (GES) has presented new and mild conditions for the strictly convex quadratic optimization problems. Simulation results illustrated the applicability of the proposed theory.

Acknowledgements

The authors wish to thank Professor R. Newcomb and the reviewers for their valuable comments and helpful suggestions which greatly improved the letter’s quality.

References