A Fast Algorithm for Solving Large Scale Nonlinear Optimization Problems Using RNN

Xiaoliang Xu
Institute of Software & Intelligent Technology
Hangzhou Dianzi University
Hangzhou, China 310018
E-mail: xxl@hdu.edu.cn

Huajin Tang
Institute for Infocomm Research
Agency for Science, Technology and Research
Singapore 119613
E-mail: htang@i2r.a-star.edu.sg

Xiaoxin Shi
Department of Mathematics
Sichuan University
Chengdu, China 610064
E-mail: sxx_2006@163.com

***Abstract*—This paper presents a discrete-time recurrent neural network (RNN) model for solving nonlinear differentiable constrained optimization problems, which contain the special case of convex optimizations over constrained sets and variational inequality problem. The qualitative analysis results about the regularity and completeness of the proposed network have been obtained. It is shown that all trajectories starting from any initial point in $\Re^n$ converge to the equilibrium set of the recurrent system. This RNN model shows its great simplicity in contrast to other existing neural network solvers. Simulations for a classs of large scale linear complementarity problems illustrate the fast convergence and features of the proposed RNN model.**

***Index Terms*—Discrete time recurrent neural networks, convergence, nonlinear optimization, quadratic optimization.**

I. Introduction

Constrained nonlinear differentiable optimization (NDO), especially the quadratic optimization, is an important problem in mathematical programming which has numerous applications in many fields of science and engineering. Since the early work of [1] and [2], the construction of recurrent neural network (RNN) for solving linear and nonlinear programming has been an active topic in the field of neural networks ([3]–[11]). Since there is a two-way bridge connecting the nonlinear optimization and variational inequality problem (VIP) ([12] and [13]), solving nonlinear optimization is also valuable in the sense that it allows one to apply the results from NDO to VIP. Thus, finding solutions of NDO has importance not only in theory but also in applications. The optimization problem over a convex subset can be described by

$$\text{minimize} \quad E(x) \quad \text{subject to} \quad x \in \Omega. \tag{1}$$

In the recent neural network literature, there exist a few RNN models for solving nonlinear optimization and linear variational inequality problem over convex constraints (see, e.g., [9], [14], [15]). In [9], a discrete-time RNN model is proposed to solve strictly convex quadratic programs with bound constraints. Sufficient conditions for the global exponential convergence of the RNN model and several corresponding neuron updating rules are resented in [9]. In [14], a continuous time RNN is presented for solving bound-constrained nonlinear optimization. In [15] a continuous time neural network model based on projection-contraction (PC) method to solve monotone linear asymmetric variational inequality problem (LVIP) is presented.

In this paper, we propose a discrete-time RNN model for solving nonlinear optimization with any continuously differentiable objective function. Therefore, quadratic optimization problem as the special cases of nonlinear optimization, as well as LVIP, can also be solved by the proposed RNN model. Furthermore, the RNN model is regular in the sense that any constrained optimum of the objective function is also an equilibrium point of the RNN. If the minimized function is convex, then the RNN is complete in the sense that the set of optima of the objective function with bound constraints is equal to the set of equilibria of the RNN. It has the quasiconvergence and attractivity property that all trajectories starting from the $\Re^n$, including the feasible region will converge to the set of equilibria of the RNN. For strictly convex quadratic optimization problems with bound constraints, the RNN model is global exponential stable which does not require additional conditions against the matrix. The most attracting advantages of the proposed RNN model lie in its simplicity and its efficiency to handle large scale optimization problems. In addition, the discrete-time RNN model has some advantages over the continuous-time counterpart in numerical simulation and digital implementation.

The organization of this paper is as follows: Section II gives some preliminaries and the problem formulation. Section III describes the discrete time RNN model for solving nonlinear differentiable optimization problems. Global exponential stability (GES) analysis of the strictly convex quadratic optimization problem is given in Section IV. In section V the performances of the network are illustrated by two illustrative simulations. Conclusions are drawn in Section VI.

II. Preliminaries and Problem Formulation

For each vector $x = (x_1, \cdots, x_n)^T \in \Re^n$, we denote $\|x\| = \sqrt{x^T x}$.

Definition 1: Let Ω be a nonempty, closed and convex subset of $\Re^n$, the *projection* of a point $y \in \Re^n$ onto the set Ω, denoted by $Proj_\Omega(y)$, is defined as the unique solution to the mathematical program: $\min \|y - x\|$, where x is a vector in Ω.

Property 1: The projector operator $Proj(\cdot)$ has the following properties:

(1)$Proj_\Omega(x) = x \Leftrightarrow x \in \Omega$.
(2)$(Proj_\Omega(y) - y)^T(Proj_\Omega(y) - x) \le 0 \quad \forall y \in \Re^n, \forall x \in \Omega$.
(3)$\|Proj_\Omega(y) - Proj_\Omega(x)\| \le \|y - x\|$, $\forall y, x \in \Re^n$, i.e., it is nonexpansive in $\Re^n$.

Definition 2: Let Ω be a nonempty, closed and convex subset of $\Re^n$, the distance from the point y to the set Ω is defined by

$$dist(y, \Omega) = \min_{x \in \Omega} \|y - x\| = \|y - Proj_\Omega(y)\|.$$

And we define that a sequence $\{y(k)\}$ converges to Ω if $dist(y(k), \Omega) \to 0$ as $k \to \infty$.

Lemma 1: Let $F(x) : \Re^n \to \Re^n$ be a continuous mapping. If there exist a bounded sequence $x(k) \in \Re^n$ for $k \ge 0$ such that $F(x(k)) \to 0$ as $k \to \infty$, then

$$\Omega_0 = \{x \in \Re^n | F(x) = 0\}$$

is a closed and nonempty set and

$$dist(x(k), \Omega_0) \to 0 \qquad \text{as} \quad k \to \infty.$$

Proof: The boundedness of $x(k)$ implies that it has a convergent subsequence whose limit is obviously in Ω_0 from the continuity of $F(x)$ and the assumption that $F(x(k)) \to 0$ as $k \to \infty$, i.e., the set Ω_0 is nonempty.

In the following, we will show $dist(x(k), \Omega_0) \to 0$ as $k \to \infty$ by the contradiction method. It is assumed that there exists a positive constant $\epsilon \ge 0$ such that for any positive number K_0 there exists a corresponding positive number $K \ge K_0$ satisfying $dist(x(K), \Omega_0) \ge \epsilon$. Let $x(k)$ has a convergent subsequence $x(k_j)$ such that

$$dist(x(k_j), \Omega_0) \ge \epsilon > 0 \quad \forall j \ge 1 \tag{2}$$

where $K_0 \le k_1 < k_2 < \cdots < k_{j-1} < k_j < \cdots < \infty$ and $k_j \to \infty$ as $j \to \infty$. Let the vector $\bar{x}$ be the limit of $x(k_j)$, i.e., $\lim_{j \to \infty}(x(k_j)) = \bar{x}$, which satisfies $F(\bar{x}) = 0$, i.e., $\bar{x} \in \Omega_0$. Thus, it follows that

$$dist(x(k_j), \Omega_0) \le \|x(k_j) - \bar{x}\|$$

for all $j \ge 1$. The right side of this inequality approaches to zero if taking the limit as $j \to \infty$. Therefore we obtain $dist(x(k_j), \Omega_0) \to 0$ as $j \to \infty$, which contradicts the initial assumption. Thus the lemma is proved. ■

III. The Discrete Time RNN Model for Nonlinear Differentiable Optimization

In this section, we will analyze the qualitative property of the proposed discrete time RNN model to solve the nonlinear differentiable minimization problems over convex constraints. The discrete time neural network model is described as

$$x_i(k+1) = f\Big(x_i(k) - \alpha_i \frac{\partial E(x)}{\partial x_i(k)}\Big) \tag{3}$$

for all $k \ge 0, i = 1, 2, \cdots, n$, where α_i is any positive constant. The activation function f is a projection operator as defined by Definition 1 and $f(x) = (f_1(x_1), f_2(x_2), \cdots, f_n(x_n))^T$.

The projection operator is assumed to be easily implemented. For bounded constraint set, $\Omega = [c, d] = \{x \in \Re^n | c_i \le x_i \le d_i, i = 1, 2, \cdots, n\}$, it takes the form of, in component wise,

$$f_i(\theta) = max(c_i, \min(\theta, d_i)).$$

If $\Omega = [0, \infty) = \{x \in \Re^n | x_i \ge 0, i = 1, 2, \cdots, n\}$, then

$$f_i(\theta) = max(0, \theta).$$

If we define a positive diagonal matrix $\Lambda = diag(\alpha_1, \alpha_2, \cdots, \alpha_n)$, then the RNN model can be represented by the following compact matrix form:

$$\text{(N)} \quad x(k+1) = f(x(k) - \Lambda \nabla E(x(k))). \tag{4}$$

We define the solution set

$$\Omega^* = \{x^* \in \Omega | E(x) \ge E(x^*), \forall x \in \Omega\}$$

for the minimization problem (8) and the equilibrium set

$$\Omega^e = \{x \in \Re^n | x = f(x - \Lambda \nabla E(x))\}$$

for the RNN model (N). It is assumed $\Omega^* \ne \emptyset$. It can be seen that Ω^* and Ω^e are both closed.

Theorem 1: The discrete time RNN model (N) is *regular*, i.e., $\Omega^* \subseteq \Omega^e$. Furthermore, if the objective function $E(x)$ is convex on R^n, then the RNN model is *complete*, i.e., $\Omega^* = \Omega^e$.

Proof: Let $x^* \in \Re^n$ be a minimizer of the minimization problem (8), if and only if it satisfies the first-order necessary optimum condition

$$\nabla E(x^*)^T(y - x^*) \ge 0, \forall y \in \Omega. \tag{5}$$

This inequality coincides with the variational inequality problem. By the fundamental results in [12] (see also [13]), any solution to variational inequality problem is equivalent to be an equilibrium of the recurrent system (N). This proves that $\Omega^* \subseteq \Omega^e$.

In order to prove that the RNN model is complete, it remains to show that $\Omega^e \subseteq \Omega^*$. Let x^e be an equilibrium of the system (N), i.e.,

$$x^e = f(x^e - \Lambda \nabla E(x^e)),$$

which is equivalent to

$$(x - x^e)^T \Lambda \nabla E(x^e) \ge 0, \forall x \in \Omega.$$

Since $\Lambda > 0$, we obtain the first-order optimum condition as (5). Thus $x^e \in \Omega^e$. The proof is completed. ■

In contrast to the continuous time dynamical systems (see, e.g., [14]), the solution trajectory of the discrete time system (N) starting from any initial point in $\Re^n$ will be mapped into the constraint set after the first iteration, and it will remain in this set, i.e., $dist(x(k+1), \Omega) = 0, \forall k \ge 0$. Clearly, Ω is a *positive invariant* and *attracting* set ([19]) of the system (N).

Theorem 2: Any solution trajectory of the discrete time system (N) starting from the inside of Ω converges to Ω^e.

Proof: Without loss of generality, we establish the proof by showing each component of any solution trajectory of the discrete time system (N) from the inside of Ω converges to Ω^e.

Let $h_i(k) = \alpha_i E(x(k))$ and

$$y_i(k) = x_i(k) - \alpha_i E(x(k)), \quad \alpha_i > 0,$$

for $i = 1, 2, \cdots, n, \forall k \geq 0$, we have

$$\begin{aligned}
\Delta h_i(k) &= h_i(k+1) - h_i(k) \\
&= (\alpha_i \nabla E(x_i(k)))^T (x_i(k+1) - x_i(k)) \\
&= [(x_i(k) - f_i(y_i(k))) + (f_i(y_i(k)) - y_i(k))]^T \\
&\quad\ [f_i(y_i(k))) - x_i(k)] \\
&\quad\ \text{[use Property 1]} \\
&\leq -\|x_i(k) - f_i(x_i(k) - \alpha_i \nabla E(x_i(k)))\|^2 \\
&= -\|e_i(k)\|^2 \\
&\leq 0.
\end{aligned}$$

Here $\|e\| = \sqrt{\sum_{i=1}^n \|e_i\|^2}$ is called an *error bound* which measures how much $x(k)$ fails to be in Ω^e. Since

$$h_i(k+1) = \Delta h_i(k) + h_i(k)$$

Repeatedly, we have

$$\begin{aligned}
h_i(k+1) &= \sum_{j=0}^{k} \Delta h_i(j) + h_i(0) \\
&\leq -\sum_{j=0}^{k} \|e_i(j)\|^2 + h_i(0).
\end{aligned} \tag{6}$$

Consider the boundedness of $h_i(k)$, $\|e_i(k)\| \to 0$ as $k \to \infty$, i.e.,

$$x_i(k) - f_i[x_i(k) - \alpha_i \nabla E(x_i(k))] \to 0 \quad \text{as} \quad k \to \infty.$$

Therefore, from Lemma 1, we obtain

$$dist(x_i(k), \Omega^e) \to 0$$

as $k \to \infty, i = 1, 2, \cdots, n$,

In other words, every component of any solution trajectory converges to Ω^e. This completes the proof. ■

Let Ω be a nempty subset of $\Re^n$ and let F be a mapping from $\Re^n$ onto itself. The variational inequality problem, denoted by VIP(Ω, F), is to find a vector $x^* \in \Omega$ such that

$$F(x^*)^T (y - x^*) \geq 0 \quad \forall y \in \Omega. \tag{7}$$

When $F(x)$ is an affine function of x, i.e., $F(x) = Mx + q$ for some given vector $q \in \Re^n$ and matrix $M \in \Re^{n \times n}$, the problem VI reduces to the linear variational inequality problem (LVIP). If $F(x)$ be the gradient of a real-valued differentiable function $E : \Re^n \to \Re$, and E is a convex function, then solving VI(Ω, F) is equivalent to find the solution of the following optimization problem over a convex subset [13]:

$$\text{minimize} \quad E(x) \quad \text{subject to} \quad x \in \Omega. \tag{8}$$

For the case of LVIP(Ω, M, q), the objective function $E(x)$ is not known a-priori, in order to employ the network (N), we should construct an equivalent differentiable function. Let $F(x) = Mx + q$, then the objective function is given as

$$E(x) = -F(x)^T (f(x) - x) - \frac{1}{2}(f(x) - x)^T (f(x) - x). \tag{9}$$

Clearly, $E(x)$ is continuously differentiable and its gradient is given by

$$\nabla E(x) = F(x) - (\nabla F(x) - I)(f(x) - x), \tag{10}$$

where I is an identify matrix (see Theorem 3.2, [16]).

IV. GES Analysis for Strictly Convex Quadratic Optimization Over Bound Constraints

Quadratic optimization is an important case of nonlinear optimization problem. In this section, we consider the strictly convex quadratic optimization over bound constraints. This class of optimization problem can be described by

$$\min \left\{ \frac{1}{2} x^T A x + x^T b \,\middle|\, x \in \Omega \right\}, \tag{11}$$

where A is positive definite.

We will use the following neural network to solve this problem:

$$\text{(NQ)} \quad x(k+1) = f\Big(x(k) - \alpha(Ax(k) + b)\Big), \tag{12}$$

where $\alpha > 0$ is some constant.

Definition 3: The neural network (NQ) is said to be globally exponentially converge, if the network exists a unique equilibrium x^e and there exist constants $\eta > 0$ and $\mu \geq 1$ such that

$$\|x(k+1) - x^e\| \leq \mu \|x(0) - x^e\| \exp(-\eta k)$$

for all $k \geq 0$. The constant η is called a lower bound of the convergence rate of the network (NQ).

Since A is a positive definite matrix, then all of its eigenvalues are positive. Let $\lambda_i > 0 (i = 1, \cdots, n)$ be all the eigenvalues of A. Denote λ_{min} and λ_{max} the smallest and largest eigenvalues of A, respectively.

Since the constraint region is a convex set, the strict convex quadratic optimization problem has a unique minimizer $x^* \in \Omega$ and $\Omega^* = \Omega^e = \{x^*\}$. The following lemma about the existence and uniqueness of the equilibrium point was proved in [9], which can also be derived from the completeness of the network (NQ) (Theorem 1 in the last section).

Lemma 2: For each $\alpha > 0$, the network (NQ) has a unique equilibrium point and this equilibrium point is the minimizer of the quadratic optimization problem (11).

Exponential stability is an important dynamical property for recurrent neural networks, and the exponential convergence rate can be calculated explicitly as reported in [17]. In the previous work [18], the following theorem has been proved.

Theorem 3: For each α, if

$$0 < \alpha < \frac{2}{\lambda_{max}}$$

then the network (NQ) is globally exponentially converge with a lower bound of convergence rate

$$\eta(\alpha) = -\ln r(\alpha) > 0.$$

Remark 1: To further improve the numerical stability and convergence speed of the network, we can use the preconditioning technique in [3] to reduce the conditioner number of the matrix A. We first define a diagonal matrix P with the diagonal elements $p_{ii} = 1/\sqrt{|a_{ii}|}(i = 1, \cdots, n)$, then some transformations are performed such that

$$\tilde{A} = PAP, \quad \tilde{b} = Pb, \quad \tilde{c} = P^{-1}c, \quad \tilde{d} = P^{-1}d.$$

Since $tra(\tilde{A}) = n$, the network (NQ1) becomes

$$\text{(NQ2)} \quad \begin{cases} \tilde{x}(k+1) = f\left(\tilde{x}(k) - \frac{2}{n}(\tilde{A}\tilde{x}(k) + \tilde{b})\right) \\ \\ x(k+1) = P\tilde{x}(k+1) \end{cases} \tag{13}$$

for $k \geq 0$.

V. Numerical Simulations

In this section, we apply the proposed discrete time RNN model to a special class of linear variational problems, namely the linear complementarity problem: Find a vector $x^* \in \Omega$ such that

$$(Mx^* + q)^T(y - x^*) \geq 0 \quad \forall y \in \Omega \tag{14}$$

where Ω is the nonnegative orthant of R^n, i.e., $\Omega = [0, +\infty)$.

To apply the proposed recurrent neural network model (4), an equivalent optimization function is firstly to be defined. It follows from equation (9) that

$$E(x) = -(Mx+q)^T(f(x)-x) - \frac{1}{2}(f(x)-x)^T(f(x)-x). \tag{15}$$

Then its gradient is derived from (10):

$$\nabla E(x) = Mx + q - (M - I)(f(x) - x). \tag{16}$$

Hence, the general RNN model (4) is reduced to such an explicit dynamic system: $k \geq 1$

$$\begin{aligned} g(k) &= Mx(k) + q - (M-I)(f(x(k)) - x(k)), \\ x(k+1) &= f(x(k) - \Lambda g(k)). \end{aligned} \tag{17}$$

In the following simulations, the error estimation is $\|e\|_\infty \leq 10^{-6}$ for the stopping criterion.

Example 1: Consider a classical linear complementarity problem LCP(Ω, M, q), where $q = (-1, \cdots, -1)^T \in \Re^n$ and M is an $n \times n$ upper triangular matrix

$$M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

which is positive definite but asymmetric. This problem has a unique solution $x^* = (0, \cdots, 0, 1)^T$ for any n.

This example is the same as the second example in [15] where a continuous time neural network model based on projection-contraction (PC) method is employed to handle this problem. The network based on PC method, which involves several steps to calculate a descent direction of a distance function, demonstrates its advantages in faster convergence speed in a linear programming example compared with three classes of neural network models investigated in [7] and in the above asymmetric LCP compared with Damped-Newton method.

Since M is positive definite, the LCP can be solved by the discrete time RNN model (N) for general nonlinear differential objective functions presented in Section III. By running the dynamical equation (17), we simulate the problem of dimensions from $n = 8$ up to $n = 4000$, although it still works well for larger dimensions for this problem. Let the scaling constants $\alpha_i = 0.6, i = 1, 2, \cdots, n$. All elements of the starting vectors are randomly uniformly distributed in the range $(0, 1)$. Table I gives the iteration numbers of the solution trajectories converging to the unique solution with respect to different dimensions.

TABLE I
Results by the proposed RNN model

n	8	16	32	64	128	256	512	1000	2000	4000
k	11	11	11	11	10	11	11	11	11	10

n=Dimension of the problem, k=Iteration number.

We depict the last five components of the trajectory starting from a randomly initial point in $\Re^n$ for the problem of 40, 400 and 4000 dimensions in Figs. 1, 2 and 3, respectively. For all different dimensionalities of the problem, tt is found that the network converges to the unique solution after about 11 iterations. This simulation example illustrates that the proposed network has a very fast convergence behavior, and the computation cost does not scale up when the scale of the simulated problem increased.

The searching trajectories are significantly different for the proposed network behavior and the gradient projection method. For example, when the current point lies within the constraint region, the gradient projection searches along the minus gradient direction in the next iteration; while this behavior only happens with the network when all the α_i take the same value [9]. Such behaviors imply that parallel implementation for the proposed discrete time network is direct. In addition, from the programming point of view, the implementation of the presented network model is simpler than that of neural network based on gradient projection method which involves several steps to compute a descent direction.

VI. Conclusions

In this paper we have proposed a general discrete time recurrent neural network model for nonlinear differentiable constrained optimization and obtained the qualitative analysis results about the regularity, completeness, and attractivity of

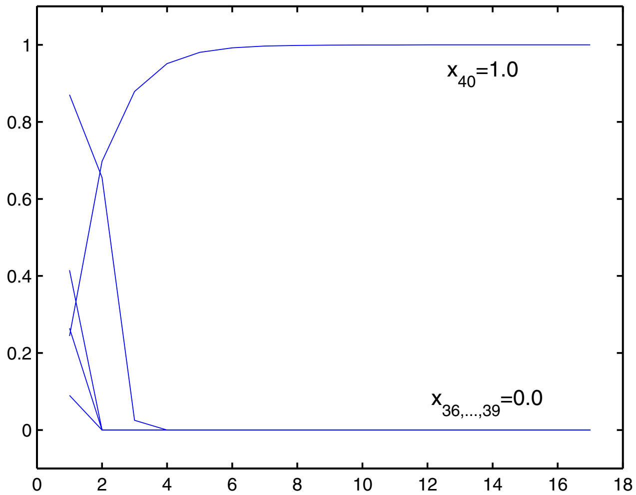

Fig. 1. The trajectories of the last five components of x for $n = 40, \alpha = 0.6$.

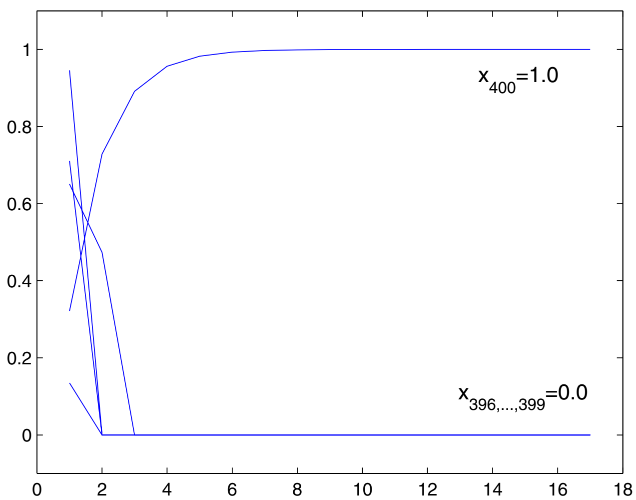

Fig. 2. The trajectories of the last five components of x for for $n = 400, \alpha = 0.6$.

the RNN model. All the solution trajectories starting from any point in $\Re^n$ converge to the equilibrium set of the network. The network is globally exponentially convergent for strictly convex quadratic optimization problem and it does not impose additional requirements over the matrix. Simple and practical conditions are derived to guarantee this property. The network demonstrates fast convergence for a quadratic optimization and a classical LCP and great insensitivity to different dimensions of the latter hard problem. Its technical simplicity suggests it to be suitable for digital implementation and a good choice for real-time solver.

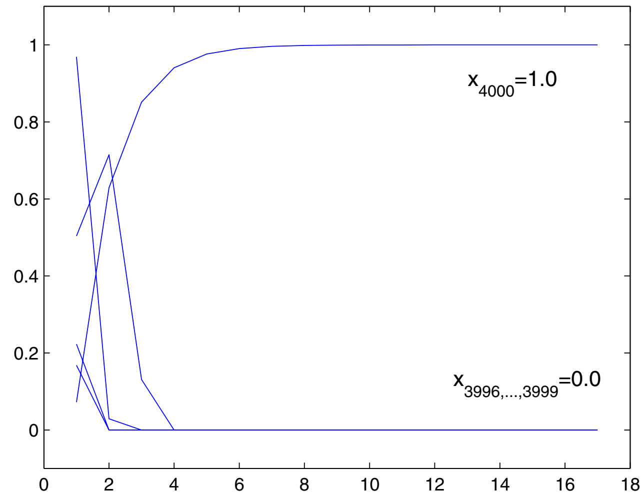

Fig. 3. The trajectories of the last five components of x for for $n = 4000, \alpha = 0.6$.

References

[1] J. J. Hopfiled and D. W. Tank, Neural computation of decision in optimization problem, *Biol. Cybern.*, vol. 52, pp. 141–152, 1985.

[2] M. P. Kennedy and L. O. Chua, Neural networks for nonlinear programming, *IEEE Trans. Circuits Syst.*, vol. 35, pp. 554–562, 1988.

[3] A. Bouzerdoum and T. R. Pattison, Neural network for quadratic optimization with bound constraints, *IEEE Trans. Neural Networks*, vol. 4, no. 2, pp. 293–303, 1993.

[4] C. Y. Maa and M. Shanblatt, Linear and quadratic programming neural network analysis, *IEEE Trans. Neural Networks*, vol. 3, pp. 580–594, 1992.

[5] S. Sudharsanan and M. Sundareshan, Exponential stability and systematic synthesis of a neural network for quadratic minimization, *Neural Networks*, vol. 4, no. 5, pp. 599–613, 1991.

[6] M. Forti and A. Tesi, New conditions for global stability of neural networks with applications to linear and quadratic programming problems, *IEEE Trans. Circuits Syst.*, vol. 42, pp. 354–366, 1995.

[7] S. H. Ż, V. Upatising and S. Hui, Solving linear programming problems with neural networks: A comparative study, *IEEE Trans. Neural Networks*, vol. 6, pp. 94–104, 1995.

[8] J. Wang, Analysis and design of a recurrent neural network for linear programming, *IEEE Trans. Circuits Syst. I*, vol. 40, no. 9, pp. 613–618, 1993.

[9] M. J. Pérez-Ilzarbe, Convergence analysis of a discrete-time recurrent neural networks to perform quadratic real optimization with bound constraints, *IEEE Trans. Neural Networks*, vol. 9, no. 6, pp. 1344–1351, 1998.

[10] Y. Xia and J. Wang, Global exponential stability of recurrent neural networks for solving optimization and related problems, *IEEE Trans. Neural Networks*, vol. 11, pp. 1017–1022, 2000.

[11] X.-B. Gao, L.-Z. Liao and L. Qi, A novel neural network for variational inequalities with linear and nonlinear constraints, *IEEE Trans. Neural Networks*, vol. 16, no. 6, pp. 1305–1317, 2005.

[12] B. C. Eaves, On the basic theorem of complementarity, *Math. Programming*, vol. 1, pp. 68-75, 1971.

[13] P. T. Harker and J. S. Pang, Finite-dimensional variational inequality and nonlinear complementary problems: A survey of theory, algorithms and applications, *Math. Programming*, vol. 48, pp. 161–220, 1990.

[14] X. B. Liang and J. Wang, A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints, *IEEE Trans. Neural Networks*, vol. 11, no. 6, pp. 1251–1262, 2000.

[15] B. He and H. Yang, A neural-network model for monotone linear asymmetric variational inequalities, *IEEE Trans. Neural Networks*, vol. 11, no. 1, pp. 3-16, 2000.

[16] M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Math. Programming*, vol. 53, pp. 99–110, 1992.
[17] Z. Yi, P.A. Heng and A.W.C. Fu, Estimate of exponential convergence rate and exponential stability for neural networks, *IEEE Trans. Neural Networks*, vol. 10, no. 6, pp. 1487–1493, 1999.
[18] K. C. Tan, H. Tang and Z. Yi, Global exponential stability of discrete-time neural networks for constrained quadratic optimization, *Neurocomputing*, vol. 56, pp. 399–406, 2004.
[19] J. K. Hale, Ordinary Differential Equations, John Wiley & Sons Inc., 1969.