

THE REAL-TIME SOLUTION OF LARGE-SCALE SIMULTANEOUS LINEAR EQUATIONS BASED ON NEURAL NETWORK

ZHAO Yan¹, ZHAO Wei-feng¹, LIAO Wu-dai²

(1. *Department of Basic, Henan Light Industry School, Zhengzhou 450000, China*)

(2. *School of Electrical and Information Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China*)

Abstract: The paper studies the disturbed Hopfield neural network of the large-scale simultaneous linear equations, and points out the criterion of neural network convergence. Under certain conditions, the steady-state errors of this network are consistently bounded or convergent to zero. That is to say, the network system has good robustness. We give a simulation example, and verify the authenticity of the conclusion.

Keywords: the large-scale simultaneous linear equation; disturbed gradient neural network; robustness

2010 MR Subject Classification: 92B20

Document code: A **Article ID:** 0255-7797(2014)02-0287-08

1 Introduction

Real-time problems of solving the large-scale simultaneous linear equations, which is in the sense of the minimal norm, often occur in the areas such as optimization, system identification, intelligent control, optimal control, robotics and control [1–4]. For many large-scale simultaneous linear equations, the number of unknowns involved is very large. These large-scale problems often need to be solved in real time.

Neural networks were proposed for solving real-time problems involved simultaneous equations, including linear cases [5–10]. We develop an electronic neural network called the gradient Hopfield neural network for solving the simultaneous linear equations. Meanwhile, we consider the case which will be errors such as circuit noise, the opamp finite gains and frequent correlations. Assume the errors are bounded and uncertain, then this kind of errors can be understood as system parameter perturbations caused by the disturbed effects or the integration effects from the internal and the external of the entire system. So, as to such a Hopfield neural network whose disturbance gradient is bounded and uncertain, the first

* **Received date:** 2013-11-04

Accepted date: 2014-01-10

Foundation item: Supported by the Teaching Reform of Higher Vocational Education and Research Project of Zhengzhou University of Light Industry.

Biography: Zhao Yan(1982–), female, born at Jiaozuo, Henan, lecturer, major in differential geometry.

thing should be considering its robustness. That is to say, under bounded conditions, for any value of the disturbance, the errors between the effective solution and the exact solution of the model are consistent or convergent [4].

The paper will study the robustness of disturbance gradient Hopfield neural network from the theory and will offer numerical simulation of some instances to check the validity of the theory. As follows: In Section 2, the disturbance gradient neural network. In Section 3, the robustness analysis. In Section 4, Matlab simulation of robustness. In Section 5, conclusion.

2 Disturbance Gradient Described by Hopfield Neural Network

Based on LMS $AX = b$ errors minimum problems, the negative gradient neural network model (theoretical model) [4] with non-linear activation function can be described as follows:

$$\dot{x}(t) = -\alpha A^T F(AX(t) - b), \quad (2.1)$$

where $A = (a_{ij})_{m \times n} \in R^{m \times n}$, $x = (x_1, x_2, \dots, x_n)^T \in R^n$, $b = (b_1, b_2, \dots, b_m)^T \in R^m$. And the matrix A is full row rank, namely $\text{rank}(A) = m$. Parameter $\alpha > 0$ determines the convergence of the network.

Non-linear activation function $f(\cdot)$ usually takes the following types.

- (1) The linear activation function $f(u) = u$.
- (2) Bipolar sigmoid activation function

$$f(u) = \frac{1 - e^{-\xi u}}{1 + e^{-\xi u}}, \xi \geq 2.$$

- (3) Power activation function

$$f(u) = u^p, p \geq 3$$

and odd.

- (4) Power-sigmoid activation function

$$f(u) = \begin{cases} u^p, & |u| \geq 1, \\ \frac{1 + e^{-\xi} - e^{-\xi u}}{1 - e^{-\xi} + e^{-\xi u}}, & \text{others,} \end{cases} \quad \xi \geq 2, \rho \geq 3$$

and odd.

Remark 1 For the power-sigmoid and bipolar sigmoid activation function, at any fixed point $y = (y_1, y_2, \dots, y_m)^T \in R^m$, Jacobi matrix $\frac{\partial F(y)}{\partial y} = \text{diag}(f'(y_1), f'(y_2), \dots, f'(y_m))$ is a diagonal positive definite matrix.

In neural network (2.1), the negative gradient neural network model (actual model) with parameter disturbance can be described as

$$\dot{X}(t) = -\alpha(A + \Delta A)^T F[(A + \Delta A)X(t) - (b + \Delta b)]. \quad (2.2)$$

Or the negative gradient neural network model (actual model) with overall disturbance can be described as

$$\dot{x}(t) = -\alpha A^T F(Ax(t) - b) + \Delta c, \tag{2.3}$$

where $\Delta A(t) \in R^{m \times n}$ and $\Delta b(t) \in R^m$ are circuit implementation errors or parameter disturbance of the matrix A and b , respectively, $\Delta c(t) \in R^n$ is the error for the non-accurate implementation of the overall model and there exists a positive number to make the consistently bounded disturbance conditions meet

$$\max\{\|\Delta A(t)\|, \|\Delta b(t)\|, \|\Delta c(t)\|\} \leq \varepsilon, \forall t \geq 0. \tag{2.4}$$

3 Robustness Analysis

First, we give two theorems to character convergence rate properties and convergence properties of equilibrium point of the disturbed Hopfield neural network.

Theorem 2 [4] The convergence rate of non-linear neural network system (2.1) depends on parameter and activation functions. The greater α is, the faster network converge. And

(1) When using the linear activation function, the neural network (2.1) globally exponential converges.

(2) When using the Bipolar-sigmoid activation function, in the interval $[-c, c] \subset [-1, 1]$, the neural network (2.1) converges faster than the linear activation function, where c is the intersection abscissa value of the bipolar sigmoid curve and the linear $f(u) = u$.

(3) When using the power activation function, within the interval $(-\infty, -1]$ and $[1, +\infty)$, neural network converges faster than the linear activation function.

(4) When using the power-sigmoid activation function, in the interval $(-\infty, +\infty)$, neural network converges faster than the linear activation function.

The following content will focus on the robustness of system (2.3).

Theorem 3 For the disturbed network system (2.3), as long as the parameter α is chosen large enough, it's easy to get that the steady-state error of the network model $\|x(t) - x^*\| \rightarrow 0, t \rightarrow \infty$. In other words, when the parameter is chosen appropriately, even if the implementation errors are large, the network still converges to the exact solution of the network system (2.1).

Proof Note x^* as the exact solution of the network system (2.1), $x(t)$ as the disturbance solution of network (2.3) at any initial value. Then two equations can be founded

$$\dot{x}^* = -\alpha A^T F(Ax^* - b) \dot{x}(t) = -\alpha A^T F(Ax(t) - b) + \Delta c.$$

Subtracting the two formulas above, note $z = x - x^*$ as dynamic error, $y^* = Ax^* - b$ as residuals, then

$$\dot{z}(t) = -\alpha A^T (F(y^* + Az(t)) - F(y^*)) + \Delta c(t). \tag{3.1}$$

When $\|Az\|$ is small enough, $F(y^* + Az) - F(y^*) \approx \frac{\partial F(y^*)}{\partial y} \cdot Az$ can be achieved. So the partial linear system of formula (3.1) at the point y^* is

$$\dot{z}(t) = -\alpha(A^T \frac{\partial F(y^*)}{\partial y} A)z + \Delta c(t). \quad (3.2)$$

By Remark 1, we have $\frac{\partial F(y^*)}{\partial y} = \text{diag}(f'(y_1^*), f'(y_2^*), \dots, f'(y_n^*))$ is a positive definite matrix. Since A is a full row rank matrix, it's not difficult to prove that $A^T \frac{\partial F(y^*)}{\partial y} A$ also is a positive definite matrix. This is because $\forall x \in R^n$, there always has

$$x^T (A^T \frac{\partial F(y^*)}{\partial y} A) x = x^T f'_i((y_1^*) A_i^T A_i x = f'_i((y_1^*) (A_i x)^2) \geq 0.$$

When and only when $A_i x = 0$, namely, $Ax = 0$, the equation can be achieved as follow:

$$x^T (A^T \frac{\partial F(y^*)}{\partial y} A) x = 0,$$

where A_i stands for the i row of A . Since A is a full row matrix, so it can be equivalent to $x = 0$.

Also because the disturbance $\Delta c(t)$ satisfies the bounded condition (2.4), by the linear system theory, when the parameter α is large enough, the eigenvalues of system (3.2) will stay in the left half complex plane and away from the imaginary axis. Thus system (3.2) converges, that is $z = x - x^* \rightarrow 0$. Theorem 3 is proved.

Remark 2 By the proof of theorem, we know the key condition of theorem is that the matrix $A^T \frac{\partial F(y^*)}{\partial y} A$ is a positive definite matrix, which needs

$$\frac{\partial F(y^*)}{\partial y} = \text{diag}(f'(y_1^*), f'(y_2^*), \dots, f'(y_m^*))$$

is a positive definite diagonal matrix. The power-sigmoid and sigmoid activation functions both can ensure the condition are met. Therefore, these two activation functions will have better performance than the linear activation function.

4 The Matlab Simulation of Robustness

To test the robustness of the gradient-based neural network, we consider the following LMS problem

$$A = \begin{pmatrix} 1 & 0 & 1 & 2 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 0 & 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 2 & 2 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 2 & 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Corresponding to the power-sigmoid activation function of neural network system (2.1) when $\xi = 4, p = 3$. We have the following model implementation errors of sinusoidal form

$$\Delta c(t) = \varepsilon(\cos 3t \ 0 \ 0 \ -\sin 3t \ 0 \ \sin 2t \ 0 \ \sin 3t \ 0)^T.$$

Using Matlab internal function ode45, plot the solution curve of the network in figures 1–3, respectively, when $\alpha = 1, \alpha = 10$ and $\alpha = 100$. Just as follows:

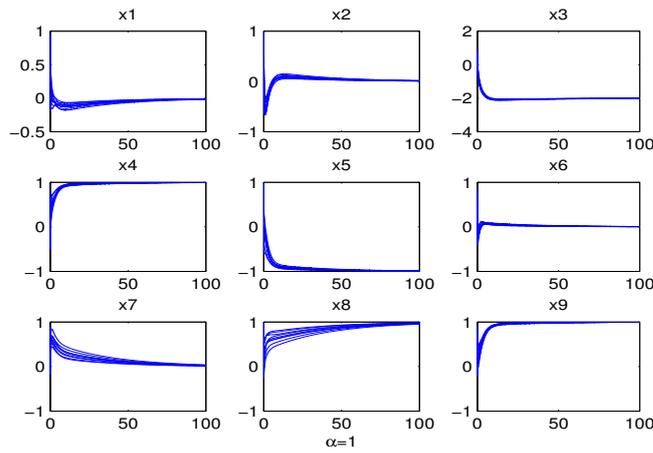


Figure 1

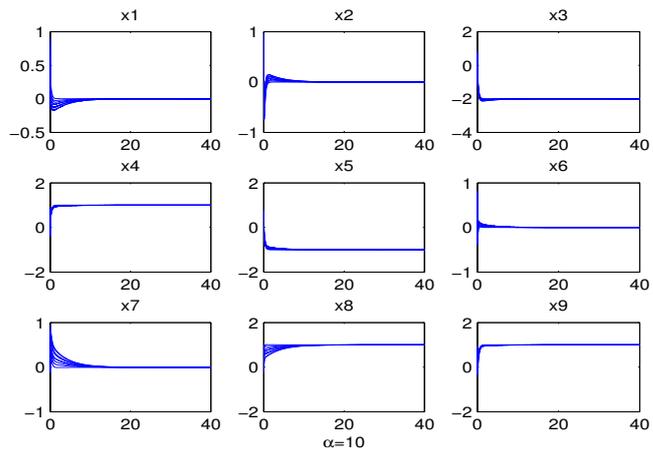


Figure 2

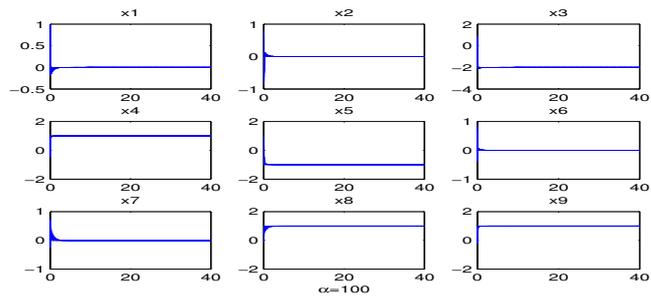


Figure 3

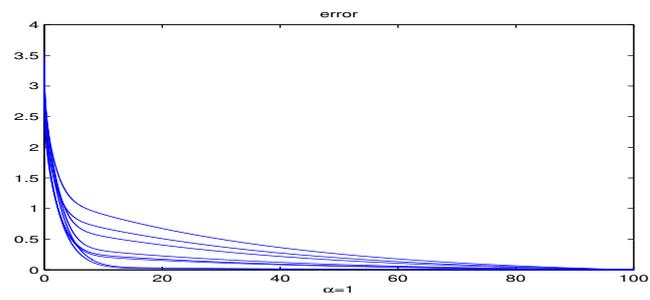


Figure 4

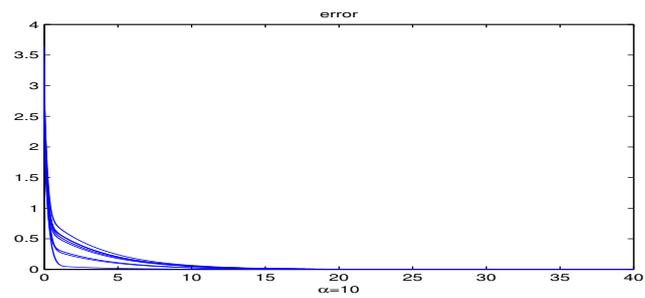


Figure 5

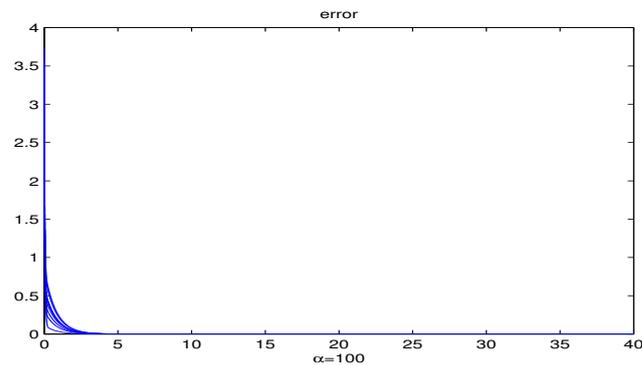


Figure 6

Figures 4–6 show us clearly that even when there are great errors, network (2.3) can run well and the error norm $\|x(t) - x^*\|$ is consistently bounded. At the same time, when α changes from 1 to 100, convergence time becomes faster and the errors fall down to a very small value when steady-state comes. The simulation result again confirms the correctness and effectiveness of the theoretical analysis in the paper.

5 Conclusion

Gradient-based Hopfield neural network (2.1) provides an effective real-time parallel computing to solve minimal norm of least mean square problems. By using different activation functions, the paper focuses on the disturbed neural network and analysis its robustness. The results show that when the activation function is power-sigmoid or bipolar-sigmoid form and the parameters are large enough, the disturbed network with implementation errors can operate well. Matlab simulation results reveal that neural network has the advantages of effectiveness and efficiency when it's used to solve minimal norm of the least mean square problems.

References

- [1] Kennedy M, Chua L O. Neural networks for nonlinear programming [J]. IEEE Trans., 1988, 35(5): 554–562 .
- [2] Atencia M, Joya G, Sandoval F. Hopfield neural networks for parametric identification of dynamical systems[J]. Neural Processing Letters, 2005, 21: 143–152.
- [3] Liao Wudai, Wang Jiangfeng, Wang Junyan. A discrete-time recurrent neural network for solving systems of complex-valued linear equations[J]. Advances in Swarm Intelligence, Lecture Notes in Computer Science, 2010, LNCS 6146: 315–320.
- [4] Zhang Yunong, Zhang Yuhang, Chen Ke. The gradiet neural network solution of linear matrix equation[J]. Sun Yat-sen University (Natural Science), 2008, 47(3): 26–32.
- [5] Wang Jun. Electronic realisation of recurrent neural networks for solving simultaneous linear equations[J]. Electroics Letters, 1992, 28(5): 493–495.

- [6] Deepak M, Prem K. Kalra. Modified hopfield neural network approach for solving nonlinear algebraic equations[J]. Engineering Letters, 2007, 14(1): 1–23.
- [7] Park J H, Kim Y S, Eom I K, Lee K Y. Economic load dispatch for piecewise quadratic cost function using hopfield neural network[J]. IEEE Transaction on Power System, 1993, 8(3): 1030–1038.
- [8] Zhang Y, Wang J. Global exponential stability of recurrent neural networks for synthesizing linear feedback control systems via pole assignment [J]. IEEE Transactions on Neural Networks, 2002, 13(3): 633–644.
- [9] Wang J. Recurrent neural network for solving quadratic programming problems with equality constraints[J]. Electronics Letter, 1992, 28(14): 1345–1347.
- [10] Liao Xiaoxin. Advances in robust stability for interval dynamical systems[J]. Advances Mathematics, 1992, 21(2): 168–184.

基于神经网络的大规模联立线性方程的实时解

赵 艳¹, 赵伟峰¹, 廖伍代²

(1. 河南省轻工业学校基础教学部, 河南 郑州 450000)

(2. 中原工学院电气与信息工程学院, 河南 郑州 450007)

摘要: 本文研究了大规模线性方程问题的扰动Hopfield神经网络, 并给出了网络收敛的判别准则. 在一定条件下, 网络的稳态误差一致有界或者收敛于0, 网络具有较好的鲁棒性. 最后数值仿真验证了方法的有效性.

关键词: 大规模联立线性方程; 扰动梯度神经网络; 系统的鲁棒性

MR(2010)主题分类号: 92B20 中图分类号: O233