

文章编号: 1007- 2985(2008)02- 0030- 05

# 对参数不确定神经网络的鲁棒状态估计\*

黄 淦, 曹进德

(东南大学数学系, 江苏 南京 210096)

**摘 要:** 讨论了参数不确定神经网络的状态估计问题. 由于参数的不确定性, 无法设计观测器使观测器的状态与原来的神经网络达到完全同步, 只能将误差控制在一定范围之内. 对于给定的观测器中的增益矩阵, 给出了判据来估计观测器的状态与原神经网络的状态间最终的误差界; 同时, 通过使用线性矩阵不等式的技术, 给出了观测器中增益矩阵的设计方法. 此外, 对误差界的估计进行讨论, 说明了影响估计准确性的主要原因. 最后用 2 个例子来说明这些判据的有效性.

**关键词:** 神经网络; Lyapunov 函数; 线性矩阵不等式; 状态估计

**中图分类号:** TP18; O175. 13

**文献标识码:** A

近几十年里, 神经网络得到了广泛的研究, 同时它也成功地应用于模式识别、图像处理、优化计算等多个领域. 其中神经网络的状态估计问题也得到很多学者的关注<sup>[1-3]</sup>. 在很多实际问题中, 神经网络中各个神经元的状态并不能由网络输出完全确定. 所谓神经网络的状态估计, 就是根据神经网络的输出来设计观测器, 而对原神经网络中各个神经元的状态进行估计. 但在现实问题中, 网络的参数带有不确定性也可能随着时间的变化而发生变化, 大量的文献<sup>[4-7]</sup>对这类不确定系统进行了分析. 笔者讨论了这类参数不确定神经网络的状态估计问题, 利用稳定性理论以及线性矩阵不等式(LMI)技术, 得到了观测器中的增益矩阵的设计方法, 并对最后的误差界给出估计公式. 此外, 还讨论了误差界的估计准确性, 并在仿真中作了仔细比较.

文中  $A > 0$  表示  $A$  是正定矩阵,  $E_n$  表示  $n$  阶单位矩阵,  $\|x\| = (\sum x_i^2)^{1/2}$ ,  $\|A\| = [\lambda_{\max}(A^* A)]^{1/2}$ .

## 1 模型介绍

考虑参数不确定的神经网络:

$$\dot{u}_i(t) = - (c_i + \Delta c_i) u_i(t) + \sum_{j=1}^n (a_{ij} + \Delta a_{ij}) g_j(u_j(t)) + I_i + \Delta I_i \quad i = 1, 2, \dots, n. \quad (1)$$

用向量形式表示为

$$\dot{\mathbf{u}}(t) = - (\mathbf{C} + \Delta \mathbf{C}) \mathbf{u}(t) + (\mathbf{A} + \Delta \mathbf{A}) \mathbf{g}(\mathbf{u}(t)) + \mathbf{I} + \Delta \mathbf{I}. \quad (2)$$

这里:  $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))^T$  为神经网络的状态向量;  $\mathbf{C} = \text{diag}(c_1, \dots, c_n)$  是对角正定阵, 有  $c_i > 0$ ;  $\mathbf{A} = (a_{ij})_{n \times n}$  表示连接矩阵;  $\mathbf{I} = (I_1, \dots, I_n)$  为神经网络的外部输入; 激活函数  $\mathbf{g}(\mathbf{u}(t)) = (g_1(u_1), \dots, g_n(u_n))^T$ ; 而  $\Delta \mathbf{C} = \text{diag}(\Delta c_1, \dots, \Delta c_n)$ ,  $\Delta \mathbf{A} = (\Delta a_{ij})_{n \times n}$ ,  $\Delta \mathbf{I} = (\Delta I_1, \dots, \Delta I_n)^T$  表示系统中参数的不确定性.

假设激活函数  $G(\mathbf{u}(t))$  满足 Lipschitz 条件:

$$\|g(\mathbf{u}) - g(\mathbf{v})\| \leq \|G(\mathbf{u} - \mathbf{v})\|, \quad (3)$$

且有界  $\|g(\mathbf{u})\| \leq \alpha_1$ . 这里并不要求激活函数是单调的或者是光滑的. 同时假设  $\|\Delta \mathbf{C}\| \leq \delta_1$ ,  $\|\Delta \mathbf{A}\| \leq$

\* 收稿日期: 2007- 11- 05

基金项目: 国家自然科学基金资助项目(60574043); 江苏省自然科学基金资助项目(BK2006093)

作者简介: 黄 淦(1982-), 男, 江苏苏州人, 东南大学数学系 2005 级硕士研究生, 主要从事神经网络、复杂网络研究; 曹进德(1963-), 男, 安徽和县人, 东南大学特聘教授, 博士生导师, 主要从事复杂网络和神经网络理论及应用、网络控制与优化等研究.

$\delta_2, \|\Delta I\| \leq \delta_3$ . 另外由激活函数的有界性, 很容易知道到状态  $u(t)$  也是有界的, 不妨设  $\|u(t)\| \leq \alpha_2$ .

设神经网络的输出为

$$y(t) = Bu(t) + f(t, u(t)). \tag{4}$$

其中:  $y(t) \in \mathbf{R}^m$  表示网络的输出;  $B$  为适当维数的矩阵;  $f: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^m$  表示神经网络输出中的非线性扰动, 它满足 Lipschitz 条件

$$|f(t, u) - f(t, v)| \leq F|u - v|. \tag{5}$$

对于不确定的神经网络(2), 可以设计如下观测器:

$$\dot{v}(t) = -Cv(t) + Ag(v(t)) + I + K[y(t) - Bv(t) - f(t, v(t))], \tag{6}$$

其中增益矩阵  $K \in \mathbf{R}^{n \times m}$ . 设误差状态为  $e(t) = u(t) - v(t)$ , 则由方程(1), (4), (6), 得

$$\begin{aligned} \dot{e}(t) = & -(C + KB)e(t) + A[g(u(t)) - g(v(t))] - K[f(t, u(t)) - f(t, v(t))] - \\ & \Delta Cu + \Delta Ag(u(t)) + \Delta I. \end{aligned} \tag{7}$$

## 2 主要结论

引理 1 (Schur complement)<sup>[8]</sup> 对于给定的常数对称阵  $\Omega_1, \Omega_2, \Omega_3$ , 若  $\Omega_1 = \Omega_1^T, \Omega_2 = \Omega_2^T > 0$ , 则

$$\Omega_1 + \Omega_3 \Omega_2^{-1} \Omega_3 < 0,$$

当且仅当

$$\begin{pmatrix} \Omega_1 & \Omega_3^T \\ \Omega_3 & -\Omega_2 \end{pmatrix} < 0 \text{ 或 } \begin{pmatrix} -\Omega_2 & \Omega_3 \\ \Omega_3^T & \Omega_1 \end{pmatrix} < 0.$$

定理 1 对于给定增益矩阵  $K \in \mathbf{R}^{n \times m}$ , 若存在实对称正定矩阵  $P \in \mathbf{R}^{n \times n}$ , 以及正实数  $\varepsilon_1 > 0, \varepsilon_2 > 0,$

$\gamma > 0$ , 使得  $\Omega < 0$  成立, 则方程(6) 与方程(2) 将以误差界  $\frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}}$  渐进同步, 其中  $\beta = 2\lambda_{\max}(P)(\delta_1 \alpha_2 + \delta_2 \alpha_1 + \delta_3)$ ,

$$\Omega = -((C + KB)^T P + P(C + KB)) + P(\varepsilon_1^{-1} A A^T + \varepsilon_2^{-1} K K^T) P + \varepsilon_1 G^T G + \varepsilon_2 F^T F + \gamma I.$$

这里以误差界  $\frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}}$  渐进同步指的是, 对于方程(2) 与方程(6), 存在  $t_0 > 0$ , 当  $t > t_0$  时  $\|u(t) -$

$$v(t)\| = \|e(t)\| \leq \frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}}.$$

证明 令  $V(e(t)) = e^T(t) P e(t)$ , 则

$$\begin{aligned} D^+ V(e(t)) = & 2e^T(t) P \{- (C + KB)e(t) + A[g(u(t)) - g(v(t))] - K[f(t, u(t)) - f(t, v(t))] - \\ & \Delta Cu + \Delta Ag(u(t)) + \Delta I\} \leq e^T(t) \{- ((C + KB)^T P + P(C + KB)) + P(\varepsilon_1^{-1} A A^T + \\ & \varepsilon_2^{-1} K K^T)\} e(t) + \varepsilon_1 [g(u(t)) - g(v(t))]^T [g(u(t)) - g(v(t))] + \varepsilon_2 [f(t, u(t)) - \\ & f(t, v(t))]^T [f(t, u(t)) - f(t, v(t))] + 2\lambda_{\max}(P)(\delta_1 \alpha_2 + \delta_2 \alpha_1 + \delta_3) \|e(t)\| \leq \\ & e^T(t) \{- ((C + KB)^T P + P(C + KB)) + P(\varepsilon_1^{-1} A A^T + \varepsilon_2^{-1} K K^T) P + \varepsilon_1 G^T G + \\ & \varepsilon_2 F^T F\} e(t) + \beta \|e(t)\| \leq e^T(t) \Omega e(t) - \gamma \|e(t)\|^2 + \\ & \beta \|e(t)\| < (\beta - \gamma) \|e(t)\| \|e(t)\|. \end{aligned}$$

所以当  $\|e(t)\| \leq \frac{\beta}{\gamma}$  时,  $D^+ V(e(t)) < 0$ . 由  $V$  函数的定义, 得

$$\lambda_{\min}(P) \|e(t)\|^2 \leq V(e(t)) \leq \lambda_{\max}(P) \|e(t)\|^2.$$

因此,  $\lambda_{\min}(P) \|e(t)\|^2 \leq \frac{\beta^2}{\gamma^2} \lambda_{\max}(P)$ . 由此可以得到方程(6) 和(2) 最终同步时的误差界为  $\frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}}$ .

定理 2 若存在实对称正定矩阵  $P \in \mathbf{R}^{n \times n}$ , 正实数  $\varepsilon_1 > 0, \varepsilon_2 > 0, \gamma > 0$ , 以及矩阵  $R \in \mathbf{R}^{n \times m}$ , 使得不

等式(8) 成立, 则可以设计增益矩阵  $K = P^{-1} R$ , 使得方程(6) 与方程(2) 将以误差界  $\frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}}$  渐进同步,

$$\begin{pmatrix} \Pi & PA & R \\ * & -\varepsilon_1 E_n & 0 \\ * & * & -\varepsilon_2 E_m \end{pmatrix} < 0, \quad (8)$$

$$\Pi = -CP - PC - B^T R^T - RB + \varepsilon_1 G^T G + \varepsilon_2 F^T F + \forall E_n.$$

证明 在方程(8)两边同乘以对角阵  $\text{diag}(E_n, \varepsilon_1^{-\frac{1}{2}} E_n, \varepsilon_2^{-\frac{1}{2}} E_m)$ , 得

$$\begin{pmatrix} \Pi & \varepsilon_1^{-\frac{1}{2}} PA & \varepsilon_2^{-\frac{1}{2}} R \\ * & -E_n & 0 \\ * & * & -E_m \end{pmatrix} < 0.$$

由引理1得  $\Omega < 0$ . 因此由定理1得定理2的结论成立.

注 这里误差界的估计依赖于  $P$  的条件数  $\frac{\lambda_{\max} P}{\lambda_{\min} P}$ . 从大量计算机仿真的模拟中发现,  $P$  的条件数较大往往会造成这个误差界估计得不准确. 因此需要准确估计误差的时候, 人们总是希望  $P$  的条件数尽可能地小.

定理3 若存在正实数  $\varepsilon_1 > 0, \varepsilon_2 > 0, \forall > 0$ , 以及矩阵  $R \in \mathbf{R}^{n \times m}$ , 使得不等式(9)成立, 则系统(6)与系统(2)将以误差界  $\frac{\beta}{\forall}$  渐进同步, 其中  $\beta = 2(\delta_1 \alpha_2 + \delta_2 \alpha_1 + \delta_3)$ ,

$$\begin{pmatrix} \Pi' & A & K \\ * & -\varepsilon_1 E_n & 0 \\ * & * & -\varepsilon_2 E_m \end{pmatrix} < 0, \quad (9)$$

$$\Pi' = -2CC - B^T K^T - KB + \varepsilon_1 G^T G + \varepsilon_2 F^T F + \forall E_n.$$

证明 事实上只要在定理2中取  $P = E_n$ , 即可得到定理3.

注 在定理3中取  $P = E_n$  使得  $P$  的条件数为1, 可有效提高估计误差的准确性, 而且通过解下列线性优化问题(10), 可以设计控制器  $K$  使得误差的范围尽可能地小. 但对  $P$  的限制使得定理3的适用范围比定理2小得多, 这也带来了一定的局限性.

$$\begin{array}{l} \max \quad \forall \\ \text{s. t.} \quad \begin{pmatrix} \Pi' & A & K \\ * & -\varepsilon E_n & 0 \\ * & * & -\varepsilon E_m \end{pmatrix} < 0. \end{array} \quad (10)$$

### 3 数值仿真

例1 考虑三维神经网络(2):

$$\begin{aligned} \dot{\mathbf{u}}(t) &= -(\mathbf{C} + \Delta\mathbf{C})\mathbf{u}(t) + (\mathbf{A} + \Delta\mathbf{A})\mathbf{g}(\mathbf{u}(t)) + \mathbf{I} + \Delta\mathbf{I}, \\ \mathbf{y}(t) &= \mathbf{B}\mathbf{u}(t) + \mathbf{f}(t, \mathbf{u}(t)). \end{aligned}$$

$$\text{这里: } \mathbf{C} = \text{diag}(1, 1, 1); g_i(s) = \frac{1}{2}(|s+1| + |s-1|); \mathbf{I} = (0, 0, 0)^T; \mathbf{A} = \begin{pmatrix} 1.25 & -3.2 & -3.2 \\ -3.2 & 1.1 & -4.4 \\ -3.2 & 4.4 & 1 \end{pmatrix};$$

$|\Delta c_j| \leq 0.1, |\Delta a_{ij}| \leq 0.1, |\Delta I_i| \leq 0.1$ . 因此,  $|g_i(s)| < 1, \mathbf{G} = \text{diag}(1, 1, 1)$ , 可以推算  $\mathbf{u}(t)$  最终有界, 存在不变集  $\{\mathbf{u} \mid |u_1| < 8.95, |u_2| < 10.11, |u_3| < 10\}$ .

若输出函数中

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \mathbf{f}(t, \mathbf{u}(t)) = \begin{pmatrix} 0.1 \sin(t) u_1(t) \\ 0.1 \cos(\sqrt{2t}(u_3(t) + u_2(t))) \end{pmatrix},$$

则  $F = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0.1 \end{bmatrix}$ .

解 LMI(8), 得

$$P = \begin{bmatrix} 0.4886 & -0.2305 & 0.4101 \\ -0.2305 & 0.1447 & -0.2044 \\ 0.4101 & -0.2044 & 0.3774 \end{bmatrix}, R = \begin{bmatrix} 0.3452 & 0.4179 \\ 0.0157 & 0.6558 \\ -0.4441 & 0.3987 \end{bmatrix},$$

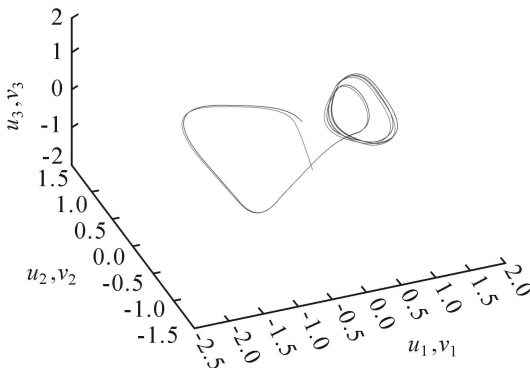
$$\gamma = 0.2271, \varepsilon_1 = 0.5366, \varepsilon_2 = 1.2862,$$

因此可以得到  $K = P^{-1}R = \begin{bmatrix} 18.8824 & 4.9019 \\ -1.9476 & 26.8832 \\ -22.7483 & 10.2911 \end{bmatrix}$ .

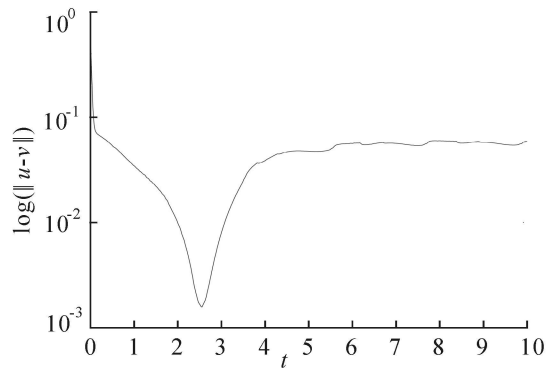
随机选取一组  $\Delta C, \Delta A, \Delta I$ , 并且随机选取方程(6),(2)的初值进行仿真. 如图 1 所示, 经过一段时间

后,  $\|e(t)\| \leq 0.1$ , 但由定理 2 估计出的误差界  $\frac{\beta}{\gamma} \sqrt{\frac{\lambda_{\max} P}{\lambda_{\min} P}} = 11.5724$ .

值得注意的是在这里定理 3 并不适用, 当  $P = E_n$  时 LMI(9) 无解.



a 例 1 中, 方程(2) (蓝色) 与方程(6) (红色) 的相图



b 例 1 中, 方程(2) 与方程(6) 的误差  $\|e(t)\|$

图 1 例 1 中方程(2) 与方程(6) 的动力学行为

例 2 若将例 1 的输出函数变为:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, f(t, u(t)) = \begin{bmatrix} 0.1 \sin(t) u_1(t) \\ 0.1 \cos \sqrt{2t} (u_2(t) + u_3(t)) \\ 0.1 \sin(t) u_2(t) \end{bmatrix},$$

则  $F = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0.1 \\ 0 & 0.1 & 0 \end{bmatrix}$ .

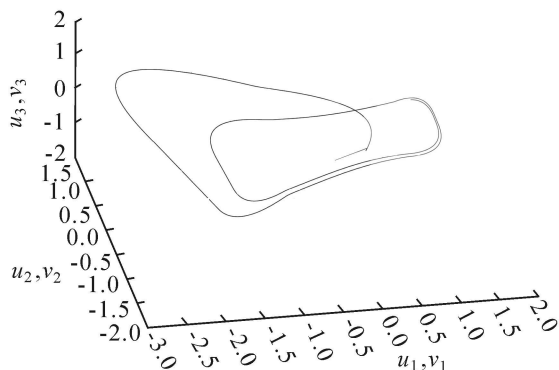
解 LMI(9), 得

$$K = \begin{bmatrix} 2364 & -27.7 & 48.3 \\ 4.8 & 2188.7 & 232.8 \\ 16.4 & -1957.1 & 2161.7 \end{bmatrix}, \gamma = 507.29, \varepsilon_1 = 1275.3, \varepsilon_2 = 6033.1.$$

随机选取一组  $\Delta C, \Delta A, \Delta I$ , 并且随机选取方程(6),(2)的初值进行仿真. 由定理 2 估计出的误差界  $\frac{\beta}{\gamma} = 0.0013 = 10^{-2.8832}$  (在图 2- b 中用直线标出), 对比例 1 这一估计比较准确. 另外, 这里增益矩阵中各个元素都比较大, 从图 2 可以发现 2 个系统很快就同步了, 而且同步后 2 个系统误差较小.

### 4 结语

讨论了参数不确定神经网络的状态估计问题. 由于参数的不确定性, 观测器的状态与原神经网络的状态



a 例2中,方程(2)(蓝色)与方程(6)(红色)的相图

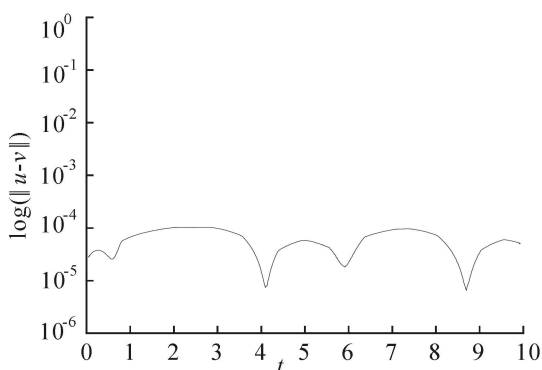
b 例2中方程(2)与方程(6)的误差  $\|e(t)\|$ 

图2 例2中方程(2)与方程(6)的动力学行为

态无法完全同步,而只能将误差控制在一定范围之内。通过推导得出这个误差界的估计,并给出方法来设计观测器中的增益矩阵。同时通过对问题的讨论与数值仿真,可以发现矩阵  $P$  的条件数对误差界估计的准确性有严重影响。为此,给出了减小误差界的方法,并对该方法的适用性作了进一步的讨论和比较。

与文献[1-3]不同,讨论的模型并不含有时滞。这是因为在处理参数不确定延时神经网络的状态估计问题中,对误差界进行有效的估计仍然存在一定的困难,这也是下一步工作的主要内容。

#### 参考文献:

- [1] LIU Y, WANG Z, LIU X. Design of Exponential State Estimators for Neural Networks with Mixed Time Delays. *Star, Open [J]. Physics Letters A*, 2007, 364(5): 401-412.
- [2] WANG Z, HO D W C, LIU X. State Estimation for Delayed Neural Networks [J]. *IEEE Trans. Neural Networks*, 2005, 16(1): 279-284.
- [3] HE Y, WANG Q G, WU M, et al. Delay-Dependent State Estimation for Delayed Neural Networks [J]. *IEEE Trans. Neural Networks*, 2006, 17(4): 1077-1081.
- [4] CHEN A, CAO J, HUANG L. Global Robust Stability of Interval Cellular Neural Networks with Time-Varying Delays [J]. *Chaos, Solitons & Fractals*, 2005, 23(3): 787-799.
- [5] LIAO X, YU J. Robust Stability for Interval Hopfield Neural Networks with Time Delay [J]. *IEEE Trans. Neural Networks*, 1998, 9(5): 1042-1045.
- [6] 张化光, 季策, 张铁岩. 多时滞 Hopfield 神经网络的鲁棒稳定性分析及吸引域的估计 [J]. *自动化学报*, 2006, 32(1): 84-90.
- [7] 张文安, 俞立, 张贵军. 离散多时滞系统的时滞相关鲁棒稳定性分析 [J]. *控制理论与应用*, 2006, 23(4): 636-639.
- [8] BOYD S P. *Linear Matrix Inequalities in System and Control Theory [M]*. Society for Industrial & Applied, 1994.

## Robust State Estimation for Uncertain Neural Networks

HUANG Gan, CAO Jia-de

(Department of Mathematics, Southeast University, Nanjing 210096, China)

**Abstract:** The robust state estimation problem for uncertain neural network is studied in this paper. As the uncertainty of the parameters, the states of the estimator can't be complete synchronous with the neural network, but asymptotically synchronous with errorbound is accessible. For given state estimator gain matrix, the error bound is derived. By using stable theory and linear matrix inequality approach, the design of the robust state estimator is also given in this paper. And the discussion of the estimate of the error bound is also presented. The simulation samples have proved the effectiveness of the conclusion.

**Key or s:** neural networks; Lyapunov function; linear matrix inequality; state estimation

(责任编辑 向阳洁)