

递归多层感知器的稳定性分析——LMI 方法

刘妹琴, 颜钢锋

(浙江大学 电气工程学院 系统科学与工程学系, 浙江 杭州 310027)

摘要: 递归多层感知器(RMLP)在工程上应用比较多,但对其稳定性的研究还比较少.本文提出一种新的神经网络模型——标准神经网络模型(SNNM),通过状态空间扩展法,将 RMLP 转化为 SNNM,而 SNNM 的稳定性分析可转化为一组线性矩阵不等式(LMI)的求解,利用 Matlab/LMI Toolbox 求解 LMI,从而判定 RMLP 的 Lyapunov 稳定性,并考虑非零阈值对稳定性的影响.该方法也适用于其他类型的递归神经网络(RNN)的稳定性分析.

关键词: 标准神经网络模型;递归多层感知器;状态空间扩展法;线性矩阵不等式

中图分类号: TP183

文献标识码: A

Stability analysis of recurrent multilayer perceptrons: LMI approach

LIU Mei-qin, YAN Gang-feng

(Department of System Science & Engineering, College of Electrical Engineering, Zhejiang University, Zhejiang Hangzhou 310027, China)

Abstract: Recurrent multilayer perceptrons (RMLPs) were widely applied to the industrial processes, but stability analysis of RMLPs was seldom researched at present. A novel neural network model named as standard neural network model (SNNM) was advanced. By applying the state space extension method, RMLPs were converted to the SNNMs. Stability conditions of the SNNMs were transformed into some linear matrix inequalities (LMIs). LMIs were solved by Matlab/LMI Toolbox to determine whether RMLPs were Lyapunov stable or not. And the effect of nonzero biased in RMLPs on stability was taken into account. The proposed approach can also be applied to other forms of recurrent neural networks (RNNs).

Key words: standard neural network model (SNNM); recurrent multilayer perceptron (RMLP); state space extension method; linear matrix inequality (LMI)

1 引言(Introduction)

递归神经网络(RNN)在许多重要的应用中显示出巨大的优势^[1,2],而 RNN 是动力学系统,在应用中首要考虑的是其稳定性,如果稳定性无法保证,系统就无法正常工作,更谈不上其他的性能指标.目前,有许多学者在这方面进行了研究^[1,3],但是其证明过程较复杂,结论保守性较大,很难在工程上应用,而且所讨论的 RNN 都不考虑阈值,这使得 RNN 的逼近能力降低,无法充分发挥 RNN 的优点.本文仅考虑目前在工程上应用比较多的一种离散 RNN——递归多层感知器(RMLP)^[2,4]的 Lyapunov 稳定性,而其他类型的 RNN 稳定性分析可以借鉴.文献[4]通过状态空间扩展法将 RMLP 转化为适合稳定性分析的形式,利用绝对稳定性理论的一些结论(定理 1),将 RMLP 稳定性分析转化为一组线性

矩阵不等式(LMI)求解问题,通过求解 LMI 来判定 RMLP 的稳定性,并考虑非零阈值对稳定性的影响.但是要满足文献[4]中的定理 1 的三个条件是比较困难的,要引入一些较强的假设条件,从而使稳定性定理的保守性增加.本文在文献[4]的基础上,提出标准神经网络模型(SNNM)的概念,并利用 Lyapunov 方法推导出基于 LMI 的 SNNM 全局渐近稳定性定理,通过状态空间扩展法,将 RMLP 转化为 SNNM 形式,从而可以利用有关 SNNM 的一些结论判定 RMLP 的稳定性,避开讨论假设条件是否成立等问题,另一方面,考虑非零阈值对 RMLP 稳定性的影响,以便提高 RMLP 的性能,降低稳定性定理的保守性.

2 问题描述(Statement of problems)

带有全局反馈的离散 RMLP 的动态行为可用如下方程描述^[4]:

$$\begin{cases} x_1(k+1) = \tanh(W_1 x_1(k) + V_1 x_n(k) + b_1), \\ x_2(k+1) = \tanh(W_2 x_2(k) + V_1 x_1(k+1) + b_2), \\ x_3(k+1) = \tanh(W_3 x_3(k) + V_2 x_2(k+1) + b_3), \\ \vdots \\ x_n(k+1) = \tanh(W_n x_n(k) + V_{n-1} x_{n-1}(k+1) + b_n). \end{cases} \quad (1)$$

其中, $x_j(k)$ ($1 \leq j \leq n$) 表示在第 k 时刻第 j 层的状态向量, W_j, V_j ($1 \leq j \leq n$) 为权值矩阵, b_j ($1 \leq j \leq n$) 为阈值向量. 式(1)中的双曲正切函数 $\tanh$ 可以被任何单调递增的函数 $\phi(\cdot)$ 所代替. 本文中, 假定 RMLP 的训练事先已完成, 因此权值在稳定性分析过程中保持不变. 可以借助基于扩展卡尔曼滤波 (EKF) 算法和随时间演化反向传播 (BPTT) 算法来训练 RMLP^[2]. 在这里, 本文仅分析不受外力作用 (或无扰动) 的 RMLP 在平衡点的 Lyapunov 稳定性, 所以各层的外部输入为零. 本文所要解决的问题为: 权值和阈值满足什么条件, 系统(1)的所有轨迹才能收敛于唯一平衡点?

3 标准神经网络模型 (Standard neural network model)

标准神经网络模型 (SNNM) 是由一个线性动力学系统和有界激励函数构成的静态非线性算子连接而成, 这里仅讨论离散 SNNM 情形, 连续情形具有类似的结构和结论. 离散的 SNNM 结构如图 1 所示. 方框 Φ 由非线性激励函数 $\phi_i(\xi_i(k))$ 构成的块对角算子, $\phi_i(\xi_i(k))$ 一般为连续、可微、单调递增、斜率有界的有界输出函数. 矩阵 N 表示时滞 $z^{-1}I$ (连

续情形为积分器) 和算子 Φ 的输入和输出之间的线性映射. 向量 $\xi(k)$ 和 $\phi(\xi(k))$ 分别表示非线性算子 Φ 的输入和输出.

图 1 中矩阵 N 写成如下分块形式:

$$N = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

其中, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times L}$, $C \in \mathbb{R}^{L \times n}$, $D \in \mathbb{R}^{L \times L}$, $x \in \mathbb{R}^n$, $\phi \in \mathbb{R}^L$, $L \in \mathbb{R}$ 为非线性激励函数的个数 (即神经网络隐含层和输出层的神经元个数总和). 那么 SNNM 可以用下列线性微分包含 (LDI) 表示:

$$\begin{cases} x(k+1) = Ax(k) + B\phi(\xi(k)), \\ \xi(k) = Cx(k) + D\phi(\xi(k)), \\ \phi(\xi(k)) = \Phi(\xi(k)), \end{cases} \quad (2)$$

如果系统(2)中的非线性激励函数还满足扇区条件 $\phi_i(\xi_i(k))/\xi_i(k) \in [q_i, u_i]$, 即

$$[\phi_i(\xi_i(k)) - q_i \xi_i(k)] \cdot [\phi_i(\xi_i(k)) - u_i \xi_i(k)] \leq 0,$$

有如下稳定性定理成立.

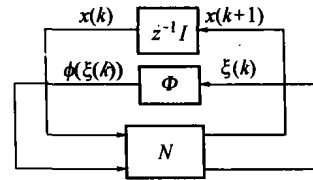


图 1 离散标准神经网络模型 (SNNM)

Fig. 1 Discrete standard neural network model (SNNM)

定理 1 离散系统(2)全局渐近稳定的充分条件是: 存在对称正定矩阵 P 和对角半正定矩阵 Λ, T 使得下列 LMI 成立

$$\begin{bmatrix} A^T P A - P - 2C^T T Q U C & A^T P B + C^T \Lambda - 2C^T T Q U D + C^T (Q + U) T \\ B^T P A + \Lambda C - 2D^T T Q U C + T (Q + U) C & B^T P B + 2\Lambda D - 2D^T T Q U D - 2T + D^T (Q + U) T + T (Q + U) D \end{bmatrix} < 0,$$

其中

$$Q = \text{diag}(q_1, q_2, \dots, q_L),$$

$$U = \text{diag}(u_1, u_2, \dots, u_L).$$

证 为了简化下面的证明过程, 将 $x(k)$ 简记为 x_k ; $\xi_i(k)$ 简记为 $\xi_{k,i}$; $\phi_i(\xi_i(k))$ 简记为 $\phi_{k,i}$; $\phi(\xi(k))$ 简记为 ϕ_k . 对于离散系统(2), 采用如下 Lyapunov 函数:

$$V(x_k) = x_k^T P x_k + 2 \sum_{i=1}^L \lambda_i \sum_{j=0}^{k-1} \phi_i(\xi_i(j)) \xi_i(j),$$

$P > 0, \lambda_i \geq 0$, 因此, $\forall x_k \neq 0, V(x_k) > 0$, 当且仅当 $x_k = 0$ 时, $V(x_k) = 0$. $V(x_k)$ 作系统(2)的解的差分为:

$$\Delta V(x_k) =$$

$$x_{k+1}^T P x_{k+1} - x_k^T P x_k + 2 \sum_{i=1}^L \lambda_i \phi_{k,i} \xi_{k,i} =$$

$$(Ax_k + B\phi_k)^T P (Ax_k + B\phi_k) - x_k^T P x_k +$$

$$2 \sum_{i=1}^L \lambda_i \phi_{k,i} (C_i x_k + D_i \phi_k) =$$

$$x_k^T (A^T P A - P) x_k + x_k^T (A^T P B + C^T \Lambda) \phi_k +$$

$$\phi_k^T (B^T P A + \Lambda C) x_k + \phi_k^T (B^T P B + 2\Lambda D) \phi_k$$

$$\begin{cases} < 0, \forall x_k \neq 0, \\ = 0, x_k = 0. \end{cases} \quad (3)$$

又因为 $(\phi_{k,i} - q_i \xi_{k,i})(\phi_{k,i} - u_i \xi_{k,i}) \leq 0$, 则

$$(\phi_{k,i} - q_i C_i x_k - q_i D_i \phi_k)(\phi_{k,i} - u_i C_i x_k - u_i D_i \phi_k) \leq 0 \Leftrightarrow$$

$$2\phi_{k,i}^2 - 2\phi_{k,i}((q_i + u_i)C_i x_k - 2\phi_{k,i}((q_i + u_i)D_i \phi_k +$$

$$2x_k^T C_i^T q_i u_i C_i x_k + 2\phi_k^T D_i^T q_i u_i D_i \phi_k +$$

$$2x_k^T C_i^T q_i u_i D_i \phi_k + 2\phi_k^T D_i^T q_i u_i C_i x_k \leq 0, \quad (4)$$

其中, C_i 为矩阵 C 的第 i 行, D_i 为矩阵 D 的第 i 行,

将式(3)和(4)写成矩阵形式:

$$\begin{bmatrix} x_k \\ \phi_k \end{bmatrix}^T \underbrace{\begin{bmatrix} A^T P A - P & A^T P B + C^T \Delta \\ B^T P A + \Delta C & B^T P B + 2\Delta D \end{bmatrix}}_{\tau_0} \begin{bmatrix} x_k \\ \phi_k \end{bmatrix} < 0, \quad \forall x_k \neq 0,$$

$$\begin{bmatrix} x_k \\ \phi_{k,1} \\ \vdots \\ \phi_{k,i-1} \\ \phi_{k,i} \\ \phi_{k,i+1} \\ \vdots \\ \phi_{k,L} \end{bmatrix}^T \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 & -C_i^T(q_i + u_i) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -d_{i,1}(q_i + u_i) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & -d_{i,i-1}(q_i + u_i) & 0 & \cdots & 0 \\ -(q_i + u_i)C_i & -(q_i + u_i)d_{i,1} & \cdots & -(q_i + u_i)d_{i,i-1} & 2-2(q_i + u_i)d_{i,i} & -(q_i + u_i)d_{i,i+1} & \cdots & -(q_i + u_i)d_{i,L} \\ 0 & 0 & \cdots & 0 & -d_{i,i+1}(q_i + u_i) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & -d_{i,L}(q_i + u_i) & 0 & \cdots & 0 \end{bmatrix}}_{\tau_i^1} \begin{bmatrix} x_k \\ \phi_{k,1} \\ \vdots \\ \phi_{k,i-1} \\ \phi_{k,i} \\ \phi_{k,i+1} \\ \vdots \\ \phi_{k,L} \end{bmatrix} +$$

$$\begin{bmatrix} x_k \\ \phi_k \end{bmatrix}^T \underbrace{\begin{bmatrix} 2C_i^T q_i u_i C_i & 2C_i^T q_i u_i D_i \\ 2D_i^T q_i u_i C_i & 2D_i^T q_i u_i D_i \end{bmatrix}}_{\tau_i^2} \begin{bmatrix} x_k \\ \phi_k \end{bmatrix} \leq 0,$$

式中, $d_{i,j}$ 为矩阵 D 的第 i 行, 第 j 列元素. 应用 S 方法^[5], 得:

$$T_0 - \sum_{i=1}^L \tau_i (T_i^1 + T_i^2) =$$

$$\begin{bmatrix} A^T P A - P & A^T P B + C^T \Delta \\ B^T P A + \Delta C & B^T P B + 2\Delta D \end{bmatrix} - \begin{bmatrix} 2C^T T Q U C & 2C^T T Q U D - C^T (Q + U) T \\ 2D^T T Q U C - T (Q + U) C & 2D^T T Q U D + 2T - D^T (Q + U) T - T (Q + U) D \end{bmatrix} =$$

$$\begin{bmatrix} A^T P A - P - 2C^T T Q U C & A^T P B + C^T \Delta - 2C^T T Q U D + C^T (Q + U) T \\ B^T P A + \Delta C - 2D^T T Q U C + T (Q + U) C & B^T P B + 2\Delta D - 2D^T T Q U D - 2T + D^T (Q + U) T + T (Q + U) D \end{bmatrix} < 0.$$

其中, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_L)$, $T = \text{diag}(\tau_1, \tau_2, \dots, \tau_L)$ 且 $\Lambda \geq 0, T \geq 0$. 证毕.

推论 1 对于定理 1 中的系统, 当 $D = 0$ 时, 系统(2) 全局渐近稳定的充分条件是: 存在对称正定矩阵 P 和对角半正定矩阵 Λ, T 使得下列 LMI 成立

$$\begin{bmatrix} A^T P A - P - 2C^T T Q U C & A^T P B + C^T \Delta + C^T (Q + U) T \\ B^T P A + \Delta C + T (Q + U) C & B^T P B - 2T \end{bmatrix} < 0. \quad (5)$$

4 系统转化(System transformation)

为了应用定理 1 进行 RMLP 的稳定性分析, 必须将 RMLP(1) 转化为系统(2), 通过以下叙述的状态空间扩展法^[4]可将 RMLP(1) 转化为系统(2).

对第 l 层的状态向量 x_l , 引入 n 个辅助变量:

$$\begin{cases} x_{l1}(k+1) = \tanh(W_l x_{ln}(k) + V_{l-1} x_{l-1,1}(k) + b_l), \\ x_{l2}(k+1) = x_{l1}(k), \\ x_{l3}(k+1) = x_{l2}(k), \\ \vdots \\ x_{ln}(k+1) = x_{l,n-1}(k), \\ l = 1, \dots, n \pmod{n}. \end{cases}$$

可以看出, 式(6)具有式(2)的形式, 但是式(6)中的阈值不为零. 式(6)描述了 n 个独立的过程, 每个过程都可以转化为原始系统(1), 但各个过程代表的系统互不相同. 例如: 状态变量为 $\text{col}(x_{1n}, x_{2,n-1}, x_{3,n-2}, \dots, x_{n1})$ 的系统可以写成如下形式:

$$\begin{aligned} x_{1n}(k+n) &= \\ \phi_1(x_{1n}(k), x_{2,n-1}(k), x_{3,n-2}(k), \dots, x_{n1}(k)) &= \\ \tanh(W_1 x_{1n}(k) + V_n x_{n1}(k) + b_1), \\ x_{2,n-1}(k+n) &= \\ \phi_2(x_{1n}(k), x_{2,n-1}(k), x_{3,n-2}(k), \dots, x_{n1}(k)) &= \\ \tanh(W_2 x_{2,n-1}(k) + V_1 x_{1n}(k+n) + b_2) &= \\ \tanh(W_2 x_{2,n-1}(k) + V_1 \tanh(W_1 x_{1n}(k) + & \\ V_n x_{n1}(k) + b_1) + b_2), \\ \vdots \\ x_{n,1}(k+n) &= \\ \phi_n(x_{1n}(k), x_{2,n-1}(k), x_{3,n-2}(k), \dots, x_{n1}(k)) &= \\ \tanh(W_n x_{n1}(k) + V_{n-1} x_{n-1,2}(k+n) + b_n) &= \\ \tanh(W_n x_{n1}(k) + V_{n-1} \phi_{n-1}(x_{1n}(k), & \\ x_{2,n-1}(k), x_{3,n-2}(k), \dots, x_{n1}(k)) + b_n). \end{aligned}$$

(6)

如果用 $k+1$ 代替上式中的 $k+n$, 并设 $x_{1n} = x_1$, $x_{2,n-1} = x_2, x_{3,n-2} = x_3, \dots, x_{n1} = x_n$, 上式就完全与式(1)相同. 对系统(6)的平衡点的全局渐近稳定性进行分析, 选择哪个过程无关紧要, 因为所有的过程都是等价的.

5 稳定性分析(Stability analysis)

文献[1,2]中将 RMLP 的阈值设为零, 不但降低了 RMLP 的映射能力, 而且得出的稳定性结论具有很大的保守性, 所以在 RMLP 稳定性分析中考虑非零阈值是很自然的, 也很有必要^[4].

将非零阈值的系统(6)写成矩阵表示形式:

$$\begin{cases} y(k+1) = Ay(k) + B\phi(\xi(k)), \\ \xi(k) = Cy(k) + b, \\ \phi(\xi(k)) = \tanh(\xi(k)). \end{cases} \quad (7)$$

其中

$$y = \text{col}(x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots,$$

$$x_{l1}, x_{l2}, \dots, x_{ln}, \dots, x_{n1}, x_{n2}, \dots, x_{nn}),$$

$$l = 1, 2, \dots, n(\text{mod } n),$$

$$A = \text{diag}(A_1, A_2, \dots, A_l, \dots, A_n),$$

$$A_l = \begin{bmatrix} 0_{m_l \times (n-1)m_l} & 0_{m_l \times m_l} \\ I_{(n-1)m_l \times (n-1)m_l} & 0_{(n-1)m_l \times m_l} \end{bmatrix},$$

m_l 为 x_{l1} 的维数,

$$B = \text{diag}(B_1, B_2, \dots, B_l, \dots, B_n),$$

$$B_l = \begin{bmatrix} I_{m_l \times m_l} \\ 0_{(n-1)m_l \times m_l} \end{bmatrix},$$

$$C = [C_1, C_2, \dots, C_l, \dots, C_n],$$

$$C_l = \begin{bmatrix} 0_{(l-1)m_l \times m_l} & 0_{(l-1)m_l \times (n-2)m_l} & 0_{(l-1)m_l \times m_l} \\ 0_{m_l \times m_l} & 0_{m_l \times (n-2)m_l} & W_l \\ V_l & 0_{m_l \times (n-2)m_l} & 0_{m_l \times m_l} \\ 0_{(n-l-1)m_l \times m_l} & 0_{(n-l-1)m_l \times (n-2)m_l} & 0_{(n-l-1)m_l \times m_l} \end{bmatrix},$$

$$b = \text{col}(b_1, b_2, \dots, b_l, \dots, b_n).$$

设 y_{eq} 为系统(7)的平衡点, 即 y_{eq} 满足下式:

$$y_{\text{eq}} = Ay_{\text{eq}} + B \tanh(Cy_{\text{eq}} + b).$$

对系统(7)作线性变换 $z = y - y_{\text{eq}}$, 并设 $a = Cy_{\text{eq}} + b$, 则

$$\begin{cases} z(k+1) = Az(k) + B\eta(\sigma(k)), \\ \eta(\sigma(k)) = \tanh(\sigma(k) + a) - \tanh(a), \\ \sigma(k) = Cz(k). \end{cases} \quad (8)$$

系统(8)与系统(7)具有相近的形式, 但是系统(8)的平衡点为原点, 其非线性函数 η 分量为:

$$\eta_i(\sigma_i) = \tanh(\sigma_i + a_i) - \tanh(a_i).$$

如果 a_i 不同, η_i 也不同, 但是 η_i 仍是单调的. 系统(8)中, 如果其阈值为 0 (即 $a = 0$), 则其扇区界为 $[0, 1]$, 当阈值不为 0 (即 $a \neq 0$), 其扇区界要小的多, 是前者的子集.

设 $\varphi_i(s) = \tanh(s + a_i) - \tanh(a_i)$, 对于系统(8)的扇区上界可由下式求得:

$$u_i = \max \{ \varphi_i(s)/s : s \neq 0 \}, \quad U = \text{diag} \{ u_i \}.$$

根据式(7)及 A 和 B 的特殊形式 ($\|A\| \leq 1, \|B\| \leq 1$), 状态向量 $y(k)$ 的每个分量的绝对值小于或等于 1, 因此

$$|\xi_i| \leq \sum_{j=1}^n |C_{ij}| + |b_i| = r_i \quad (i = 1, \dots, L),$$

可以看出 $|s + a_i| \leq r_i, |a_i| < r_i$.

系统(8)的扇区下界 q_i 的求解可以通过文献[4]中的引理 1 求得:

引理 1^[4] 如果 $|s + a_i| \leq r_i$, 则

$$\varphi_i(s)/s \geq q_i =$$

$$(\tanh(r_i) - \tanh(|a_i|))/(r_i - |a_i|),$$

如果 $|a_i| = r_i$, 则 $q_i = d(\tanh(s))/ds(s = r_i)$.

对于 RMLP(1)进行稳定性分析的步骤如下:

1) 利用状态空间扩展法, 将 RMLP(1)转化为系统(7)的形式.

2) 找出系统(7)的平衡点. 如果 RMLP(1)在训练过程中就可确定不动点, 该不动点就作为系统(7)的平衡点; 否则, 可以计算系统的一些状态轨迹, 直到系统趋于平衡点, 该过程可以利用 Matlab 进行运算.

3) 通过状态变量的平移, 将系统(7)的平衡点移到原点, 这样系统(7)就转化为系统(8), 非线性函数也相应发生变化.

4) 计算每个转换函数 $\varphi(s) = \tanh(s + a) - \tanh(a)$ 的扇区界. 利用 Matlab 函数 `fminbnd` 可求得 $-\varphi(s)/s$ 的极小值, $\varphi(s)/s$ 的上界为负的极小值; 根据引理 1 计算出 $\varphi(s)/s$ 的下界, 从而求得 Q 和 U .

5) 利用 Matlab 的 LMI 控制工具箱^[6]求解式(5), 判断系统是否稳定, 值得注意的是: 如果式(5)没有可行解, 并不说明系统不稳定, 可以通过其他复杂的方法判定系统的稳定性.

6 例子(Example)

下面分析一个具体的 RMLP 的稳定性, 对于两层的 RMLP, 每层具有两个节点, 其结构如图 2 所示, 动态行为如下方程所示:

$$\begin{cases} \mathbf{x}_1(k+1) = \tanh(\mathbf{W}_1 \mathbf{x}_1(k) + \mathbf{V}_2 \mathbf{x}_2(k) + \mathbf{b}_1), \\ \mathbf{x}_2(k+1) = \tanh(\mathbf{W}_2 \mathbf{x}_2(k) + \mathbf{V}_1 \mathbf{x}_1(k+1) + \mathbf{b}_2), \end{cases} \quad (9)$$

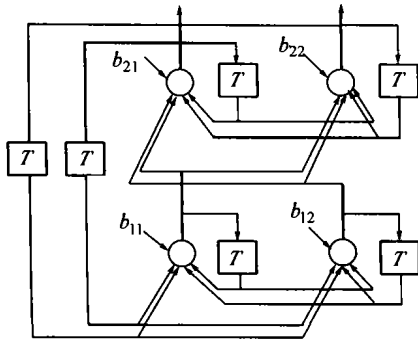


图2 两层的RMLP结构

Fig. 2 Structure of two-layer RMLP

其中权值矩阵和阈值向量如下:

$$\begin{aligned} \mathbf{W}_1 &= \begin{pmatrix} -0.2494 & -0.3483 \\ -0.1381 & -0.1054 \end{pmatrix}, \\ \mathbf{W}_2 &= \begin{pmatrix} 0.0981 & -0.1520 \\ -0.2939 & -0.2922 \end{pmatrix}, \\ \mathbf{V}_1 &= \begin{pmatrix} 0.9200 & 0.0920 \\ -0.2300 & 0.6440 \end{pmatrix}, \\ \mathbf{V}_2 &= \begin{pmatrix} 0.7500 & 0.2500 \\ 0.6000 & -0.0500 \end{pmatrix}, \\ \mathbf{b}_1 &= \text{col}(0.3000, -0.5000), \end{aligned}$$

$$\mathbf{b}_2 = \text{col}(-1.0000, 1.0000).$$

根据步骤1,将系统(9)转化为:

$$\begin{cases} \mathbf{x}_{11}(k+1) = \tanh(\mathbf{W}_1 \mathbf{x}_{12}(k) + \mathbf{V}_2 \mathbf{x}_{21}(k) + \mathbf{b}_1), \\ \mathbf{x}_{12}(k+1) = \mathbf{x}_{11}(k), \\ \mathbf{x}_{21}(k+1) = \tanh(\mathbf{W}_2 \mathbf{x}_{22}(k) + \mathbf{V}_1 \mathbf{x}_{11}(k) + \mathbf{b}_2), \\ \mathbf{x}_{22}(k+1) = \mathbf{x}_{21}(k). \end{cases} \quad (10)$$

这样,系统(10)可以写成式(7)的形式,其中, $\mathbf{y} = \text{col}(\mathbf{x}_{11}, \mathbf{x}_{12}, \mathbf{x}_{21}, \mathbf{x}_{22})$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ 和 $\mathbf{b}$ 取值如下所示:

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{I}_{2 \times 2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}_{2 \times 2} & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{I}_{2 \times 2} & 0 \\ 0 & 0 \\ 0 & \mathbf{I}_{2 \times 2} \\ 0 & 0 \end{pmatrix}, \\ \mathbf{C} &= \begin{pmatrix} 0 & \mathbf{W}_1 & \mathbf{V}_2 & 0 \\ \mathbf{V}_1 & 0 & 0 & \mathbf{W}_2 \end{pmatrix}, \quad \mathbf{b} = \text{col}(\mathbf{b}_1, \mathbf{b}_2). \end{aligned}$$

假如系统稳定,其平衡点为

$$\mathbf{y}_{\text{eq}} = \text{col}(0.0573, -0.7410, 0.0573, -0.7410, -0.8268, 0.5345, -0.8268, 0.5345).$$

计算出扇区界

$$\begin{aligned} \mathbf{U} &= \text{diag}\{0.9992, 0.8291, 0.7683, 0.9213\}, \\ \mathbf{Q} &= \text{diag}\{0.4884, 0.3243, 0.1400, 0.2420\}. \end{aligned}$$

利用 Matlab/LMI 控制工具箱求解式(5),可得:

$$\mathbf{P} = \begin{bmatrix} 153.9902 & -76.0801 & 0 & 0 & 0 & 0 & -33.8377 & -41.6013 \\ -76.0801 & 547.8634 & 0 & 0 & 0 & 0 & -117.0840 & 15.4305 \\ 0 & 0 & 213.3289 & -89.3822 & -287.3975 & 67.5761 & 0 & 0 \\ 0 & 0 & -89.3822 & 476.1148 & -65.8291 & -106.6357 & 0 & 0 \\ 0 & 0 & -287.3975 & -65.8291 & 615.3706 & -0.3046 & 0 & 0 \\ 0 & 0 & 67.5761 & -106.6357 & -0.3046 & 259.6717 & 0 & 0 \\ -33.8377 & -117.0840 & 0 & 0 & 0 & 0 & 172.7951 & 60.9458 \\ -41.6013 & 15.4305 & 0 & 0 & 0 & 0 & 60.9458 & 230.9313 \end{bmatrix},$$

$$\mathbf{\Lambda} = \text{diag}\{0.0182, 0.0834, 0.0239, 0.1588\}; \quad \mathbf{T} = \text{diag}\{121.7026, 308.3922, 352.4518, 195.5336\}.$$

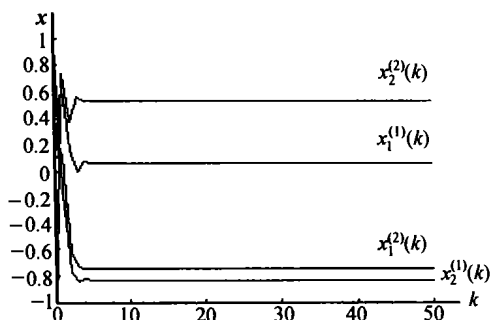


图3 两层的RMLP状态曲线

Fig. 3 States of two-layer RMLP

$\mathbf{P}$ 为正定矩阵, $\mathbf{\Lambda}$ 和 $\mathbf{T}$ 都是对角正定矩阵, 由推论 1 可知系统(9) 为全局渐近稳定的, 其状态曲线如图 3 所示. 由图 3 可知, 系统(9) 在平衡点 $\mathbf{y}_{\text{eq}}$ 是稳定的, 与推论 1 判别一致.

7 结论(Conclusion)

本文提供了一种基于 LMI 技术的新的 RMLP 稳定性分析方法. 标准神经网络模型(SNNM)的提出为 RMLP 的稳定性分析提供一种新的途径, 而且该方法还突破以往 RMLP 稳定性分析的局限, 将非零阈值引入稳定性分析, 不仅可以提高 RMLP 的逼近能力, 而且也可增强 RMLP 的稳定性. 如果能将其

他类型的网络(如时延神经网络)表示为 SNNM 的形式,也可用该方法分析它们的稳定性.值得说明的是:定理 1 和推论 1 是稳定的充分条件,而非充要条件,所以得不到 LMI 的可行解,并不说明系统不稳定;如果将扇区缩小(即增大阈值,缩小权值),可以获得可行解,但是这种作法会大大降低 RMLP 的性能.所以如何在提高性能的同时,增强 RMLP 的稳定性是将来研究的一个方向.另一方面,该方法对非线性函数的要求仅限于满足扇区条件,对于特定的非线性函数(如 $\tanh$),如果能利用其他的特性(如斜率有界等),就可以进一步降低稳定性定理的保守性,这也是未来研究的方向之一.

参考文献(References):

- [1] SUYKENS J A K, VANDEWALLE J P L, de MOOR B L R. *Artificial Neural Networks for Modeling and Control of Non-linear Systems* [M]. Boston, Dordrecht, London: Cluwer Academic Publishers, 1996.
 - [2] FELDKAMP L A, PUSKORIUS G V. A signal processing framework based on dynamic neural networks with application to problems in adaptation, filtering and classification [J]. *Proc of the IEEE*, 1998, 86(11): 2259 - 2277.
 - [3] TANAKA K. An approach to stability criteria of neural network control systems [J]. *IEEE Trans on Neural Networks*, 1996, 7(3): 629 - 642.
 - [4] BARABANOV N E, PROKHOROV D V. Stability analysis of discrete-time recurrent neural networks [J]. *IEEE Trans on Neural Networks*, 2002, 13(2): 292 - 303.
 - [5] BOYD S P, GHAOUI L E, FERON E, et al. *Linear Matrix Inequalities in System and Control Theory* [M]. Philadelphia, PA: SIAM, 1994.
 - [6] GAHINET P, NEMIROVSKI A, LAUB A J, et al. *LMI Control Toolbox* [M]. Natick, MA: The Math Works Inc, 1995.
- 作者简介:
- 刘妹琴 (1972 —), 女, 副教授, 1999 年毕业于中南工业大学, 获工学博士学位, 1999 年至 2001 年在华中科技大学从事博士后研究工作, 现为浙江大学系统科学与工程系教师, 在国内外知名刊物发表论文 30 余篇, 研究方向为混杂系统, 神经网络和信号处理等. E-mail: liumeiqin@cee.zju.edu.cn;
- 颜钢锋 (1959 —), 男, 教授, 1984 年毕业于浙江大学, 获硕士学位, 从事 CAD/CAM 智能控制, 非线性理论等方面的研究工作. E-mail: ygf@cee.zju.edu.cn.
-
- (上接第 896 页)
- MAGHRABY S E, ed. New York: Van Nostrand Reinhold, 1978: 127 - 169.
- [2] HOLT C C, MODIGLIANI F, MUTH J, et al. *Planning Production Inventories and Workforce* [M]. New Jersey: Prentice-Hall, Inc, 1960.
 - [3] WANG D, FANG Shu-cheng. A genetics-based approach for aggregate production planning in fuzzy environment [J]. *IEEE Trans on Systems, Man, and Cybernetics (Part A)*, 1997, 12(5): 636 - 645.
 - [4] BERGSTROM G L, SMITH B E. Multi-item production planning—an extension of the HMMS rules [J]. *Management Science*, 1970, 16(4): 614 - 629.
 - [5] HAUSMAN W H, McCLAIN J D. A note on the Bergstrom-Smith multi-bibitem production planning model [J]. *Management Science*, 1971, 17(5): 783 - 785.
 - [6] BITRAN G R, YANASSEE H H. Deterministic approximations to stochastic production problems [J]. *Operations Research*, 1984, 32(5): 999 - 1018.
 - [7] RINKS D B. The performance of fuzzy algorithm models for aggregate planning and differing cost structures [A]. *Approximate Reasoning in Decision Analysis* [M]. GUPTA M M, SACHEZ E, ed. Amsterdam, North-Holland, New York: [s.n.], 1982: 267 - 278.
 - [8] TANG J, FUNG RYK, WANG D. A fuzzy approach to modelling production and inventory planning [A]. *Proc of the 14th IFAC Congress* [C]. Beijing: [s.n.], 1999: 261 - 265.
 - [9] TANG J, WANG D, FUNG R. Fuzzy formulation for multi-product aggregate production planning [J]. *Production Planning and Control*, 2000, 11(7): 670 - 676.
 - [10] 唐加福, 汪定伟, 许宝栋. 多品种集约生产计划问题的模糊方法 [J]. *管理科学学报*, 2003, 6(1): 44 - 50.
(TANG J, WANG D, LIU S, et al. Fuzzy multiple product aggregate production planning [J]. *J of Management Sciences in China*, 2003, 6(1): 44 - 50.)
 - [11] ZIMMERMANN H J. *Fuzzy Set Theory and Its Applications* [M]. Hingham: Kluwer-Nijhoff, 1985.
 - [12] 唐加福, 汪定伟. 模糊优化理论与方法的研究综述 [J]. *控制理论与应用*, 2000, 17(2): 159 - 164.
(TANG J, WANG D. Fuzzy optimization theory and methodology: A survey [J]. *Control Theory & Applications*, 2000, 17(2): 159 - 164.)
- 作者简介:
- 唐加福 (1965 —), 男, 1999 年毕业于东北大学控制理论与控制工程专业, 获工学博士学位; 1998 和 2000 年任香港城市大学高级访问研究员和博士后研究员; 2002 年任香港理工大学客座研究员 (Research Fellow). 现任东北大学系统工程研究所教授. 感兴趣的领域是 fuzzy 优化理论与方法, 供应链计划与物流管理, 新产品质量优化设计. E-mail: jftang@mail.neu.edu.cn;
- 董颖 (1971 —), 女, 博士研究生. 感兴趣的领域包括模糊建模与优化, 信息管理与决策分析. E-mail: dy_neu@sina.com;
- 刘士新 (1968 —), 男, 博士, 副教授. 感兴趣的领域包括智能优化方法, 生产计划与调度. E-mail: sxliu@mail.neu.edu.cn;
- 汪定伟 (1948 —), 男, 教授, 博士生导师. 感兴趣的领域包括模糊建模与智能优化方法, 供应链管理, 电子商务的建模与优化. E-mail: dwwang@mail.neu.edu.cn.