

微粒群优化动态神经网络模型结构分析

范剑超, 韩敏

(大连理工大学 电子信息与电气工程学部, 辽宁 大连 116023)

摘要: 微粒群算法由于进化机制中的随机不确定性, 其稳定性很难进行分析, 所以对微粒群的研究多是根据经验的实际优化模型求解. 针对该问题, 利用鲁棒不确定性理论, 将算法分解为时不变和不确定时变的结构, 减少原有参数固定的假设条件, 从而对引入动态惯性权重的微粒群算法的渐近稳定性进行分析. 在此基础上, 采用李雅普诺夫方法, 得到基于微粒群参数优化的动态神经网络收敛的充分条件, 自适应调整微粒速度的上下限, 为组合模型的实际应用提供参数选择的理论基础. 最后, 通过仿真实例验证了所给出微粒群算法稳定性条件和基于微粒群优化的动态神经网络收敛条件的有效性.

关键词: 微粒群优化; 动态神经网络; 鲁棒不确定性; 稳定性

中图分类号: TP18 **文献标识码:** A

Model-structure analysis of dynamic neural networks with particle-swarm optimization

FAN Jian-chao, HAN Min

(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian Liaoning 116023, China)

Abstract: Due to the random uncertainty in the particle-swarm optimization algorithm(PSO), the analysis of stability is difficult to be performed. Most of the researchers on PSO solve this problem based on the practical model obtained by experience. Being different, we employ the robust uncertainty theory to decompose the original algorithm into the time-invariant part and the uncertain time-variant part for reducing the original fixed constraints on parameters, and perform the asymptotic stability analysis by using the PSO algorithm with dynamic inertia weight. By using the Lyapunov method, we obtain the sufficient conditions of stability for the dynamic neural networks based on PSO and the upper and lower bounds of the parameters to be adjusted, providing the theoretical basis for parameter selection. Finally, simulation examples validate the stability conditions and the effectiveness of the proposed dynamic neural networks based on PSO algorithm.

Key words: particle-swarm optimization; dynamic neural networks; robust uncertain; stability

1 引言(Introduction)

微粒群算法(particle swarm optimization, PSO)自 1995 年由社会心理学家 Kennedy 和电气工程师 Eberhart 提出之后, 在无约束和多目标优化等问题中得到了广泛地应用^[1,2].

到目前为止, 更多的文献主要集中于 PSO 的实际应用中, 没有对微粒群的运动轨迹进行详细的分析^[3], 有关 PSO 的参数选择, 稳定性的工作主要集中在以下 3 个方面: 1) 基于实验仿真的统计结果. Ozcan 等^[4]指出了粒子的速度和位置可以近似看成正弦波状, 瞬间不同的参数决定了波的频率和振幅, 而以 Clerc 等^[5]为代表的学者建立了微粒群算法的收敛模型, 给出了全局收敛系数 χ ; 2) 线性系统理论. 将 PSO 速度和位置进化公式转化为状态空间描述, 从而利用线性系统的稳定性理论进行求解. Tre-

lea^[6]将 D 维空间的 PSO 方法简化为 1 维, 然后利用离散时不变线性系统方法进行分析, 在 Bergh 等^[7]的文献里虽然加入了惯性权重的分析, 但是仍然假设所有时变的参数为固定的. Kadiramanathan 等^[8]将微粒动态轨迹转化为非线性反馈控制系统, 克服仅对线性模型的分析, 放松对随机变量 $\varphi(k)$ 的假设条件, 但主模型仍然是在时不变的情况下进行研究. 潘峰和陈杰等^[9]同样利用时变离散系统的方法, 在谱半径小于 1 的情况下得到加速系数的取值范围. 然而, 目前为了平衡 PSO 算法的 Exploration 和 Exploitation 的能力, 惯性权重多采用非线性动态递减的方法, 所以完整的微粒群算法稳定性分析应该建立在充分考虑惯性权重和随机参数的时变非线性特性的基础上. 3) 随机过程. 将微粒的轨迹作为序列分析其极值, Chen 等^[10]在速度更新公式中引入随机向量, 从

而保证高的扩展搜索能力, Jiang等^[11]在Bergh等^[7]的研究基础上, 利用随机过程的理论对参数的选择从另外一个角度进行了重新的理解. 此外, 金欣磊等^[12]提出系数矩阵 $A(k)$ 的谱半径在单位圆内是系统稳定的既非充分也非必要条件, 所以笔者给出了一个系统依概率收敛的充分条件. 上述方法多是基于标准微粒群算法中随机参数 $\varphi(k)$, $w(k)$ 为固定常数的假设条件, 因此很难得到标准微粒群算法中粒子真正的运动轨迹.

而神经网络作为一种有效的智能方法被广泛应用于非线性动态系统的建模过程中^[13]. 利用PSO优化神经网络参数的方法, 可以克服梯度下降求导计算. 其结构简单, 寻优高效的特点使其成功应用于电力分配、图像处理和预测控制等众多领域中^[14~16]. 然而两者结合的文献中未见有关分析组合模型稳定性的讨论, 是否由于微粒群算法参数的选择不当从而导致整个模型的不收敛仍然是一个需要解决的问题.

因此, 本文的工作主要集中在两个方面: 一是从一个新的角度利用鲁棒不确定性的方法, 将微粒群动态轨迹分析分解为一个时不变与不确定参数变化的系统, 利用李雅普诺夫理论给出PSO算法稳定性的条件, 其中包含动态变化的惯性权重 $w(k)$ 和随机变量 $\varphi(k)$, 放松原有研究中的假设条件, 解决其中存在的局限性; 二是在此基础上, 利用PSO算法更新网络的权值、阈值和动态延迟参数, 避免Jacobian矩阵的求取运算, 得到了基于PSO优化的动态神经网络收敛条件、自适应调整微粒幅值的上下限, 减少人工参数设定. 在保证精度基本不变的情况下, 提高了组合模型的收敛速度, 为两种方法更好地结合和应用给出理论保证.

2 PSO收敛性分析(Analysis of PSO convergence)

标准微粒群算法主要有速度和位置的进化公式, 如下所示^[1]:

$$\begin{cases} v_{id}(k+1) = \\ w(k)v_{id}(k) + c_1\text{rand}_1[p_{id}(k) - x_{id}(k)] + \\ c_2\text{rand}_2[p_{gd}(k) - x_{id}(k)], \\ x_{id}(k+1) = x_{id}(k) + v_{id}(k+1), \end{cases} \quad (1)$$

其中: $x_{id}(k)$ 和 $v_{id}(k)$ 分别是第 i 个微粒在第 d 维的速度和位置, 且 $v_{id}(k) \in [v_{\min}, v_{\max}]$ 具有上下边界, 其边界通常是通过人工经验进行设定; $w(k)$ 为动态惯性权重, 可以动态减小微粒自身惯性运动; $p_{id}(k)$ 是第 i 个微粒的历史最优位置, $p_{gd}(k)$ 是整个群体的历史最优位置, $c_1 = c_2 = 2$ 为加速系数. 将微粒群算法的进化公式写成如下的矩阵形式, 可以得到如下时

变离散系统的描述:

$$\begin{bmatrix} x(k+1) \\ v(k+1) \end{bmatrix} = \begin{bmatrix} 1 & w(k) \\ 0 & w(k) \end{bmatrix} \begin{bmatrix} x(k) \\ v(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(k), \quad (2)$$

$$y(k) = [1 \ 0] \begin{bmatrix} x(k) \\ v(k) \end{bmatrix}, \quad (3)$$

$$u(k) = -\varphi(k)(y(k) - p), \quad (4)$$

这里 $u(k)$ 可以看作是一个控制输入信号, 随机参数为 $\varphi(k) = c_1\text{rand}_1 + c_2\text{rand}_2$, 此外, 参数 p 代表着系统的平衡点 $x_* = p$, $v_* = 0$, 其表达式如下所示:

$$p = \frac{c_1\text{rand}_1 p_{id}(k) + c_2\text{rand}_2 p_{gd}(k)}{\varphi(k)}, \quad (5)$$

可以注意到随着迭代过程的进行, 当微粒达到最佳位置时 $p_{id}(k) = p_{gd}(k) = p$. 因此, p 可以表示为微粒寻优过程的极值点. 系统状态的描述图如图1所示.

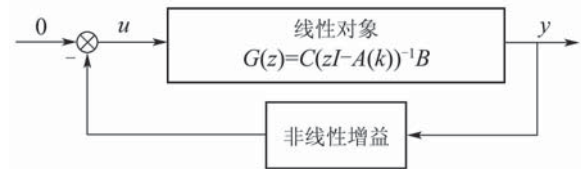


图1 微粒轨迹的系统状态空间描述

Fig. 1 The system state description for particles trajectory

在控制系统的前向通道中为一个线性时变被控对象, 将非线性增益放置到反馈通道中. 为了简化整个动态系统方程, 引入如下向量:

$$\zeta(k) = \begin{bmatrix} x(k) - p \\ v(k) \end{bmatrix}.$$

这样可以使系统更加具有实际物理意义, 即所有变量达到平衡点时, $x(k) = p$, $v(k) = 0$, 意味着所有微粒到达极值点, 而且微粒停止搜索. 因此, 系统状态方程可以转变为如下形式:

$$\zeta(k+1) = A(k)\zeta(k) + Bu(k), \quad (6)$$

$$y(k) = C\zeta(k), \quad (7)$$

$$u(k) = -\varphi(k)y(k), \quad (8)$$

其中: 状态矩阵 $A(k) = \begin{bmatrix} 1 & w(k) \\ 0 & w(k) \end{bmatrix}$, 输入矩阵为

$B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, 输出矩阵为 $C = [1 \ 0]$. 因此, 状态方程

描述可以重新写成如下形式:

$$\zeta(k+1) = (A(k) - \varphi(k)BC)\zeta(k). \quad (9)$$

为了充分考虑参数 $w(k)$, $\varphi(k)$ 引入的不确定性, 将状态矩阵 $[A(k) - \varphi(k)BC]\zeta(k)$ 拆成两项: 时不变项 A_0 和不确定时变项 $\Delta A_0(k)$, 因此可以得到式(10):

$$A(k) - \varphi(k)BC =$$

$$\begin{bmatrix} 1 - \varphi(k) & w(k) \\ -\varphi(k) & w(k) \end{bmatrix} = A_0 + \Delta A_0(k), \quad (10)$$

其中:

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \Delta A_0(k) = \begin{bmatrix} -\varphi(k) & w(k) \\ -\varphi(k) & w(k) \end{bmatrix}. \quad (11)$$

矩阵 $\Delta A_0(k)$ 包含了微粒群算法的随机不确定进化因素. 如果 $w(k) \neq 0$, 矩阵 $(A(k) - \varphi(k)BC)$ 为非奇异矩阵, 当且仅当 $\zeta_* = 0$, 方程 $\zeta_* = (A(k) - \varphi(k)BC)\zeta_*$ 满足. 因此, 上述状态空间按描述在原点处有唯一的平衡点.

至此, 微粒群动态轨迹问题转变为如下的鲁棒不确定性系统的稳定性分析, 对式(9)重新表述为如下形式:

$$\zeta(k+1) = (A_0 + \Delta A_0(k))\zeta(k). \quad (12)$$

下面给出相关的引理和微粒群算法稳定性证明过程.

引理 1 如果矩阵 $S = S^T \in \mathbb{R}^{(n+m) \times (n+m)}$ 可以表述为如下形式:

$$S = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix},$$

$A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{m \times m}$, 那么对于 $S > 0$ 的充分必要条件为^[17]

$$C > 0, A - B^T C^{-1} B > 0. \quad (13)$$

引理 2 假设参数 $D, E, F(x, t)$ 是具有合适维数的实值矩阵, 且 $F^T(x, t)F(x, t) \leq I$, 对于任意一个向量 $\varepsilon_0 > 0$, 在如下不等式^[17]:

$$DF(x, t)E + E^T F^T D^T \leq \varepsilon DD^T + \varepsilon^{-1} E^T E. \quad (14)$$

定理 1 完整微粒群算法表示为如式(12)所示的不确定系统, 如果存在一个正定矩阵 P , 相关参数 $\varepsilon_0, \alpha > 0$, 并且满足下述的矩阵不等式, 则微粒群算法在原点处渐近稳定:

$$S = \begin{bmatrix} \bar{A}_0^T P + P \bar{A}_0 + 2\alpha P + \varepsilon_0 \bar{G}_0^T \bar{G}_0 & P \bar{D}_0 \\ \bar{D}_0^T P & -\varepsilon_0 I \end{bmatrix} < 0, \quad (15)$$

其中:

$$\begin{aligned} \bar{A}_0 &= [A_0 + (c+r)I]^{-1}[A_0 + (c-r)I], \\ \Delta \bar{A}_0(t) &= 2r[A_0 + (c+r)I]^{-1}\Delta A_0(k)[A_0 + \\ &\quad (c+r)I + \Delta A_0(k)]^{-1}. \end{aligned}$$

证 首先对不确定系统(12)做线性分式变换, 如式(16)所示:

$$w = f(s) = (s + c - r)/(s + c + r). \quad (16)$$

它将 s 平面的圆域 $D(c, r)$ 变为 w 平面以虚轴为界的左半平面, 且利用矩阵和的逆的反演公式, 将离散不确定系统(12)转变为连续时变不确定系统, 如式(17)所示:

$$\dot{\bar{\zeta}}(t) = (\bar{A}_0 + \Delta \bar{A}_0(t))\bar{\zeta}(t), \quad (17)$$

其中:

$$\begin{aligned} \bar{A}_0 &= [A_0 + (c+r)I]^{-1}[A_0 + (c-r)I], \\ \Delta \bar{A}_0(t) &= 2r[A_0 + (c+r)I]^{-1}\Delta A_0(k)[A_0 + \\ &\quad (c+r)I + \Delta A_0(k)]^{-1}. \end{aligned}$$

针对连续系统(17)引入一个新的状态平移向量 $z(t) = e^{\alpha t}\bar{\zeta}(t) > 0$, 因此得到

$$\begin{aligned} \dot{z}(t) &= \alpha e^{\alpha t}\bar{\zeta}(t) + e^{\alpha t}\dot{\bar{\zeta}}(t) = \\ &\alpha z(t) + e^{\alpha t}(\bar{A}_0 + \Delta \bar{A}_0(t))\bar{\zeta}(t). \end{aligned} \quad (18)$$

这里的李雅普诺夫函数可以定义为如下形式:

$$V(z(t)) = z^T(t)Pz(t). \quad (19)$$

对上述公式求导数, 得到

$$\begin{aligned} \dot{V}(z(t)) &= \dot{z}^T(t)Pz(t) + z^T(t)P\dot{z}(t) = \\ &[(\bar{A}_0 + \Delta \bar{A}_0(t) + \alpha I)z(t)]^T Pz(t) + \\ &z^T(t)P[(\bar{A}_0 + \Delta \bar{A}_0(t) + \alpha I)z(t)] = \\ &2\alpha z^T(t)Pz(t) + z^T(t)(\Delta \bar{A}_0^T(t)P + \\ &P\Delta \bar{A}_0(t))z(t) + z^T(t)(\bar{A}_0^T P + P\bar{A}_0)z(t). \end{aligned} \quad (20)$$

在标准微粒群算法中, 因为 $\varphi(t) \in \text{rand}(0, 1)$, $w(t) \in [0, 1]$, 所以利用范数有界不确定模型, 不确定时变项 $\Delta \bar{A}_0(t)$ 可以表述为 $\Delta \bar{A}_0(t) = \bar{D}_0 \bar{F}_0(t) \bar{G}_0$, \bar{D}_0, \bar{G}_0 分别代表具有合适维数的实值矩阵, $\bar{F}_0(t)$ 是未知的实值时变矩阵, 满足Lebesgue条件, 如式(21)所示:

$$\bar{F}_0^T(t)\bar{F}_0(t) \leq I. \quad (21)$$

因此, 根据引理1和引理2可以得到如下形式:

$$\begin{aligned} \dot{V}(z(t)) &= \\ &2\alpha z^T(t)Pz(t) + z^T(t)(\bar{G}_0^T \bar{F}_0^T(t) \bar{D}_0^T P + \\ &P \bar{D}_0 \bar{F}_0(t) \bar{G}_0)z(t) + z^T(t)(\bar{A}_0^T P + P \bar{A}_0)z(t) \leq \\ &z^T(t)(\bar{A}_0^T P + P \bar{A}_0 + 2\alpha P + \\ &\varepsilon_0^{-1} P \bar{D}_0 \bar{D}_0^T P + \varepsilon_0 \bar{G}_0^T \bar{G}_0)z(t) = \\ &z^T(t)S z(t) < 0, \end{aligned}$$

这里

$$S = \begin{bmatrix} \bar{A}_0^T P + P \bar{A}_0 + 2\alpha P + \varepsilon_0 \bar{G}_0^T \bar{G}_0 & P \bar{D}_0 \\ \bar{D}_0^T P & -\varepsilon_0 I \end{bmatrix} < 0.$$

因此定理1得证. 证毕.

至此,没有任何限制条件下的标准PSO算法的稳定性得到了证明.随着时间的推移,可以发现

$$\lim_{k \rightarrow \infty} x(k) = p_{id}(k) = p_{gd}(k) = p,$$

每个微粒的历史最优值和群体的最优值重合,这样利用鲁棒不确定性原理即可保证 $x_0 \in \mathbb{R}^D \in [x_{\min}, x_{\max}]$,算法可以收敛到唯一的平衡点 p .

3 动态前向神经网络(Dynamic feedforward neural networks)

为了提高前向神经网络的动态特性,本文提出一种新的动态前向神经网络(dynamic feedforward neural networks, DFNN),神经元之间只有一个连接,网络结构如图2所示.

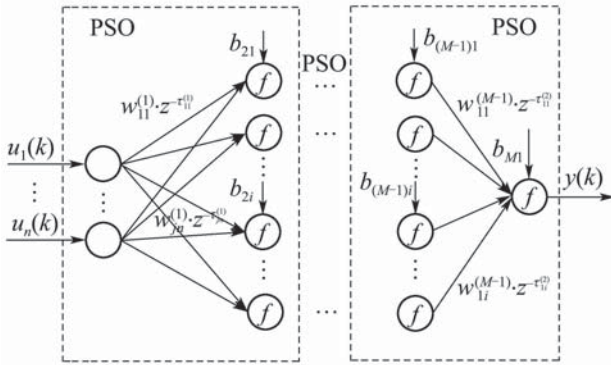


图2 动态前向神经网络结构图

Fig. 2 Basic structure of dynamic feedforward neural networks

在输入输出层与中间隐层的连接处,分别引入动态延迟算子 τ ,神经网络中动态关系如式(22)所示:

$$\text{net}_j(k) = \sum_{i=1}^N \omega_{ji} o_i(k - \tau_{ji}) + b_j, \quad (22)$$

其中: b_j 为阈值, ω_{ji} , τ_{ji} 分别表示第 i 个神经元到下一层第 j 个神经元之间的连接权值和动态延迟. 权值、阈值和动态延迟参数都通过PSO算法进行离线更新,从而构成了组合模型PSO+DFNN.

4 PSO优化的动态神经网络收敛性分析 (Convergence analysis of PSO+DFNN)

目前已有微粒群算法与多种不同的神经网络模型进行结合,本文以图2所示的前向动态神经网络为例,对组合模型的收敛性进行分析,从而提供PSO参数选择范围,减少人工设定,为PSO+DFNN组合模型提供理论保障.

定理2 假设满足定理1,惯性权重 $|w(k)| < 1$, $w(k) \neq 0$. 则基于PSO优化的动态神经网络在平衡点处渐近稳定的条件为

$$-2\sqrt{\frac{1}{3}} < w(k)A(k) + \varphi(k)B(k) < 2\sqrt{\frac{1}{3}}. \quad (23)$$

为了便于表达,定义 $A(k) = v(k)/e(k)$ 和 $B(k) = d(k)/e(k)$, 其中: $v(k)$ 为微粒的速度, $e(k)$ 为系统误差, $d(k) = x(k) - p$ 为微粒到平衡点的距离.

证 定义李雅普诺夫方程如下所示:

$$V(k) = \frac{1}{2}e^2(k). \quad (24)$$

根据神经网络输出值 $y(k)$ 和期望输出 $y_d(k)$ 的误差,可以将其定义为

$$e(k) = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_d(k) - y(k))^2} = \sqrt{J}.$$

因此李雅普诺夫函数每步的变化为

$$\Delta V(k) = V(k+1) - V(k) = \frac{1}{2}[e^2(k+1) - e^2(k)]. \quad (25)$$

此外,在离散形式下系统误差的变化为

$$e(k+1) = e(k) + \Delta e(k). \quad (26)$$

定义1 系统误差的全微分形式如式(27)所示:

$$\Delta e(k) = \frac{\partial e(k)}{\partial \omega} \Delta \omega + \frac{\partial e(k)}{\partial b} \Delta b + \frac{\partial e(k)}{\partial \tau} \Delta \tau, \quad (27)$$

这里 ω, b, τ 是动态神经网络中3个自适应调整的参数. 根据微粒群算法的进化公式,得到参数的更新公式如下:

$$\begin{cases} \Delta \omega = -\frac{\partial J}{\partial \omega} = w(k)v^\omega(k) + \varphi(k)d^\omega(k), \\ \Delta b = -\frac{\partial J}{\partial b} = w(k)v^b(k) + \varphi(k)d^b(k), \\ \Delta \tau = -\frac{\partial J}{\partial \tau} = w(k)v^\tau(k) + \varphi(k)d^\tau(k), \end{cases} \quad (28)$$

这里 $w(k)$, $\varphi(k)$ 分别是微粒群算法中的惯性权重和随机参数; 而 v^* , d^* 分别是调整 ω, b, τ 的速度和距离.

因此,李雅普诺夫函数的变化如下所示:

$$\begin{aligned} \Delta V(k) &= \frac{1}{2}(e^2(k+1) - e^2(k)) = \\ &= \frac{1}{2}((e(k) + \Delta e(k))^2 - e^2(k)) = \\ &= \frac{1}{2}\Delta e^2(k) + e(k)\Delta e(k) = \\ &= \frac{1}{2}\left(\frac{\partial e(k)}{\partial \omega} \Delta \omega\right)^2 + \frac{1}{2}\left(\frac{\partial e(k)}{\partial b} \Delta b\right)^2 + \\ &= \frac{1}{2}\left(\frac{\partial e(k)}{\partial \tau} \Delta \tau\right)^2 + \frac{\partial^2 e(k)}{\partial \omega \partial b} \Delta \omega \Delta b + \\ &= \frac{\partial^2 e(k)}{\partial \omega \partial \tau} \Delta \omega \Delta \tau + \frac{\partial^2 e(k)}{\partial b \partial \tau} \Delta b \Delta \tau + \\ &= e(k)\left(\frac{\partial e(k)}{\partial \omega} \Delta \omega + \frac{\partial e(k)}{\partial b} \Delta b + \frac{\partial e(k)}{\partial \tau} \Delta \tau\right), \end{aligned}$$

将式(28)代入到上式中可以到如下的表达式:

$$\begin{aligned} \Delta V(k) = & 2e^2(k)[(\frac{\partial e(k)}{\partial \omega})^2((\frac{\partial e(k)}{\partial \omega})^2 + 2(\frac{\partial e(k)}{\partial b})^2 - 1) + \\ & (\frac{\partial e(k)}{\partial b})^2((\frac{\partial e(k)}{\partial b})^2 + 2(\frac{\partial e(k)}{\partial \tau})^2 - 1) + \\ & (\frac{\partial e(k)}{\partial \tau})^2((\frac{\partial e(k)}{\partial \tau})^2 + 2(\frac{\partial e(k)}{\partial \omega})^2 - 1)]. \end{aligned} \quad (29)$$

通过观察式(29), 可以发现是由 $(\frac{\partial e(k)}{\partial \omega})^2$, $(\frac{\partial e(k)}{\partial b})^2$ 和 $(\frac{\partial e(k)}{\partial \tau})^2$ 3项组成的轮换对称多项式. 根据因式分解定理, 对其进行进一步简化可得

$$\begin{aligned} \Delta V(k) = & 2e^2(k)[(\frac{\partial e(k)}{\partial \omega})^2 + (\frac{\partial e(k)}{\partial b})^2 + (\frac{\partial e(k)}{\partial \tau})^2] \cdot \\ & [(\frac{\partial e(k)}{\partial \omega})^2 + (\frac{\partial e(k)}{\partial b})^2 + (\frac{\partial e(k)}{\partial \tau})^2 - 1], \end{aligned} \quad (30)$$

因此, 可以很明显的发现, 如果

$$(\frac{\partial e(k)}{\partial \omega})^2 + (\frac{\partial e(k)}{\partial b})^2 + (\frac{\partial e(k)}{\partial \tau})^2 < 1, \quad (31)$$

那么 $\Delta V(k) < 0$, 组合模型的误差逐步减小, 最终区域稳定. 进一步, 根据轮换对称的特性, 对其充分条件进行放缩, 可以得到如下的条件仍然符合公式(32):

$$\begin{cases} 0 \leq (\frac{\partial e(k)}{\partial \omega})^2 < \frac{1}{3}, \\ 0 \leq (\frac{\partial e(k)}{\partial b})^2 < \frac{1}{3}, \\ 0 \leq (\frac{\partial e(k)}{\partial \tau})^2 < \frac{1}{3}. \end{cases} \quad (32)$$

根据式(28)和式(32), 可以得到如下形式:

$$\begin{cases} 0 \leq (w(k)v^\omega(k) + \varphi(k)d^\omega(k))^2 < \frac{1}{3}e^2(k), \\ 0 \leq (w(k)v^b(k) + \varphi(k)d^b(k))^2 < \frac{1}{3}e^2(k), \\ 0 \leq (w(k)v^\tau(k) + \varphi(k)d^\tau(k))^2 < \frac{1}{3}e^2(k). \end{cases} \quad (33)$$

由于目标函数 $J = e^2(k)$, 且 $A(k) = v(k)/e(k)$, $B(k) = d(k)/e(k)$, 上述的不等式可以转变为如下的形式:

$$-2\sqrt{\frac{1}{3}} < w(k)A(k) + \varphi(k)B(k) < 2\sqrt{\frac{1}{3}}.$$

定理2得证. 证毕.

为了更好地将所得到的结论更直观应用于基于微粒群优化的动态神经网络组合模型, 对定理2进行扩展, 可以得到推论1.

推论 1 在基于微粒群优化的动态神经网络组合模型中, 微粒速度的上下需要动态调整为:

$$v_{\min}(k) = -2\sqrt{\frac{1}{3}}e(k), \quad v_{\max}(k) = 2\sqrt{\frac{1}{3}}e(k), \quad (34)$$

组合模型在平衡点处渐近稳定.

证 根据定理2可以得到神经网络参数的调整范围, 进行变换形式:

$$\begin{cases} -2\sqrt{\frac{1}{3}} < w(k)A(k) + \varphi(k)B(k) < 2\sqrt{\frac{1}{3}}, \\ -2\sqrt{\frac{1}{3}} < \frac{w(k)v(k) + \varphi(k)(x(k)-p)}{e(k)} < 2\sqrt{\frac{1}{3}}, \\ -2\sqrt{\frac{1}{3}}e(k) < v(k+1) < 2\sqrt{\frac{1}{3}}e(k). \end{cases} \quad (35)$$

而 $v(t) \in [v_{\min}, v_{\max}]$, 所以推论1得证. 证毕.

从结果中可以看出原先需要人工设定的速度上下边界, 可以进行动态的自适应调整, 从而保证PSO+DFNN的稳定性. 可以观察不等式两边的形式, 随着神经网络的训练, $e(k)$ 逐渐减小, 微粒速度的范围也动态的减小, 所以微粒最终可以趋于停止.

5 仿真实验(Simulations)

为了验证本文提出微粒群算法稳定性条件和基于微粒群算法优化的动态神经网络模型的有效性. 区别于已有文献的惯性权重和随机参数固定的方法, 本文设定惯性权重为 $w(k) = \cos(k\pi/100)$, 随着迭代次数的进行按余弦函数非线性递减, 有利于PSO算法的后期持续逼近效果, $\phi(k) = \text{rand}(0, 1)$, 加速系数为 $c_1 = c_2 = 2$.

因此, 可以得到系统(10)中的 $A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, 将

其进行变换到 $D(0, 1)$ 的 w -平面内, 可以得到 $\bar{A}_0 = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$, 取 $\bar{D}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\bar{G}_0 = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$, $\alpha =$

0.33, 可以得到正定矩阵 $P = \begin{bmatrix} 0.0047 & 0 \\ 0 & 0.0363 \end{bmatrix}$, 且

$$\begin{aligned} \bar{A}_0^T P + P \bar{A}_0 + 2\alpha P + \varepsilon_0^{-1} P \bar{D}_0 \bar{D}_0^T P + \varepsilon_0 \bar{G}_0^T \bar{G}_0 = \\ \begin{bmatrix} -0.0009 & 0 \\ 0 & -0.2486 \end{bmatrix} < 0. \end{aligned}$$

满足定理1, 微粒群算法稳定. 选取速度和位置的初始点为 $x(0) = -10$, $v(0) = 2$, 得到的满足稳定条件的粒子运动轨迹和相平面分别如图3和图4所示.

从图3和图4中可以看出, 粒子可以很快达到稳定状态. 进而, 通过非线性动态延迟系统仿真实验, 验证在保证收敛条件下的PSO+DFNN模型, 可以充分表征动态系统的非线性关系, 提高辨识精度. 具体的非线性延迟系统方程描述如下:

$$y(k) = \frac{y(k-1)}{1+y(k-1)^2} + u(k-10)^3. \quad (36)$$

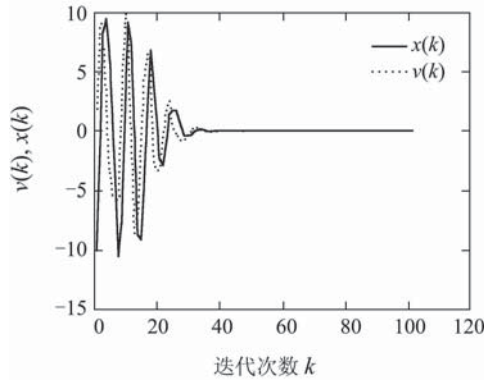


图3 粒子运动轨迹示意图

Fig. 3 Schematic diagram of particle trajectories

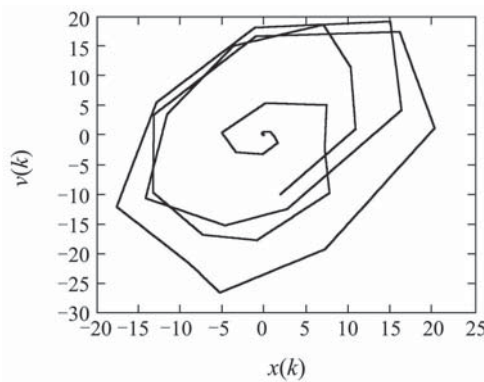


图4 粒子运动轨迹相平面图

Fig. 4 Phase planet of particle trajectories

从式(36)中可以看出,当前输出与输入有10个延迟时间.采用PID控制器进行控制,得到系统输入输出2000组数据作为训练和测试数据,训练次数为10000次,网络结构选为2-20-1,种群规模为20,维数 $D = 141$.在神经网络的学习过程中,每次迭代都根据推论1,对更新的参数进行限制,如果超出范围就设置为相应的边界 $\pm 2\sqrt{\frac{1}{3}}e(k)$,整个算法MATLAB伪代码如下:

Start

Initialize the population

Evaluate the error: $e(1), k = 1$

Update P_i and P_g

While ($k > Iter_{max}$)

Do

Update $v(k)$

If $v(k+1) > +2\sqrt{\frac{1}{3}}e(k), v(k+1) = +2\sqrt{\frac{1}{3}}e(k)$

If $v(k+1) < -2\sqrt{\frac{1}{3}}e(k), v(k+1) = -2\sqrt{\frac{1}{3}}e(k)$

Update $x(k)$

Evaluate the error: $e(k+1)$

Update P_i and P_g

Update inertia weight $w(k) = \cos(k\pi/100)$

End do

End

图5为经过训练后网络的辨识结果图,实线为被控对象的实际输出,虚线为神经网络输出.可以发现,PSO+DFNN具有较好的泛化效果,可以很好地模拟被控对象输入输出特性.

本文将PSO+DFNN与其他神经网络模型进行仿真比较,分别是原始BPNN, PSO+BPNN, Genic Algorithm(GA)+BPNN, 动态递归Elman神经网络,采用MATLAB工具箱提供的3种经典的改进BP神经网络优化方法, Variable Learning Rate, Quasi-Newton, Levenberg-Marquardt.采用相同的网络结构2-20-1,进行网络学习,其中GA算法的初始规模为50,交叉概率为0.95,结果如表1所示.

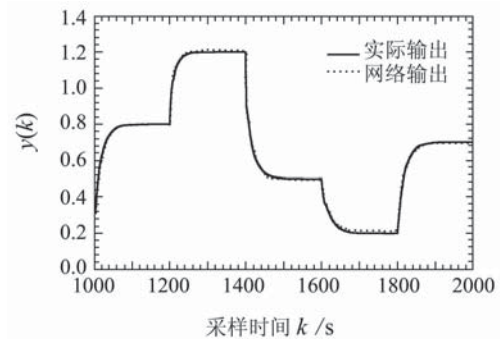


图5 PSO+DFNN的系统辨识结果

Fig. 5 System identification result of PSO+DFNN

表1 神经网络学习结果

Table 1 Learning results of neural networks

神经网络	训练误差	泛化误差	时间消耗/min
BPNN	0.79191	0.018030	14.55
PSO+BPNN ^[16]	0.065247	0.004765	4.13
GA+BPNN ^[18]	0.069623	0.003773	31.22
VLR+BPNN	0.00162	0.006620	5.6
QN+BPNN	0.00501	0.000964	9.1
LM+BPNN	0.00443	0.000902	13.25
Elman	0.00221	0.000150	850.1
本文PSO+DFNN	0.00241	0.000240	4.97

从式(35)中可以发现随着 $e(k)$ 减小,算法自适应减小微粒随机更新的幅度,既保证了组合模型前期的全局寻优,又保证了算法寻优后期神经网络后期的持续逼近能力.从表中可以得出,PSO+DFNN的训练和泛化效果都得到了显著地提升.所以推论1是在保证较高的寻优精度下,充分保证算法的稳定性.GA+BPNN方法的效果要比原始的BP神经网络好,但仍然不如微粒群算法与神经网络结合的效果好.

虽然PSO+DFNN不如Elman网络精度高,但是时间消耗却减小为原来的1/171.此外,与改进的梯度下降方法相比,精度和时间都达到了更好的效果.在实际工程应用中,重点要求的是稳定性、精确度和实时有效性.本文提出的方法,在非线形系统建模过程中,提供了理论保证,可以充分保证系统的稳定性,减小模型人工参数的设定,防止了算法不收敛导致设备的损坏.另外,在保证较高精度的同时,所需要的计算时间较短,更容易满足工艺要求.

6 结论与展望(Conclusion and prospect)

本文在微粒群算法分析基础之上,充分考虑了动态惯性权重和随机参数的影响,减少了已有方法中利用线性时不变系统理论的条件约束,根据鲁棒不确定性理论,将算法分解成为一个时不变和不确定参数变化的结构,从而给出了完整微粒群算法的粒子运动轨迹分析,得到算法渐近稳定的充分条件.此外,并在此基础之上,将微粒群算法应用于动态神经网络结构参数更新,进而得到了基于微粒群算法优化的动态神经网络的收敛的充分条件,自适应调整微粒速度的上下边界参数,在系统辨识的过程中,对参数优化的幅值进行自适应可以调整,既保证了神经网络后期的持续逼近能力,同样也保证了组合模型的稳定性,仿真实验中发现PSO+DFNN提高了对非线性动态延迟系统的辨识精度和计算速度.然而,推论1仅是通过误差进行自适应调整,还可以从非线性动力学的流形架构理念出发,分析误差曲面复杂地形的深沟豁和缓坡度的无规则分布,以误差趋势预测设计自适应策略.

参考文献(References):

- [1] KENNEDY J, EBERHART R C. Particle swarm optimization[C] // *Proceedings of IEEE International Conference on Neural Networks*. Piscataway, NJ: IEEE, 1995: 1942 – 1948.
- [2] LIANG J J, QIN A K, SUGANTHAN P N, et al. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions[J]. *IEEE Transactions on Evolutionary Computation*, 2006, 10(3): 281 – 295.
- [3] 潘峰, 陈杰, 辛斌, 等. 粒子群优化方法若干特性分析[J]. *自动化学报*, 2009, 35(7): 1010 – 1016.
(PAN Feng, CHEN Jie, XIN Bin, et al. Several characteristics analysis of particle swarm optimizer[J]. *Acta Automatica Sinica*, 2009, 35(7): 1010 – 1016.)
- [4] OZCAN E, MOHAN MOHAN C K. Particle swarm optimization: surfing the waves[C] // *Proceedings of the 1999 Congress on Evolutionary Computation*. Washington D.C, USA: IEEE, 1999: 1939 – 1944.
- [5] CLERC M, KENNEDY J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space[J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(1): 58 – 73.
- [6] TRELEA I C. The particle swarm optimization algorithm: convergence analysis and parameter selection[J]. *Information Processing Letters*, 2003, 85(6): 317 – 325.
- [7] BERGH F, ENGELBRECHT A P. A study of particle swarm optimization particle trajectories[J]. *Information Sciences*, 2006, 176(8): 937 – 971.
- [8] KADIRKAMANATHAN V, SELVARAJAH K, FLEMING P J. Stability analysis of the particle dynamics in particle swarm optimizer[J]. *IEEE Transactions on Evolutionary Computation*, 2006, 10(3): 245 – 255.
- [9] 潘峰, 陈杰, 甘明刚, 等. 粒子群优化算法模型分析[J]. *自动化学报*, 2006, 32(3): 368 – 377.
(PAN Feng, CHEN Jie, GAN Minggang, et al. Model analysis of particle swarm optimizer[J]. *Acta Automatica Sinica*, 2006, 32(3): 368 – 377.)
- [10] CHEN X, LI Y M. A modified PSO structure resulting in high exploration ability with convergence guaranteed[J]. *IEEE Transactions on Systems Man, and Cybernetics, Part B: Cybernetics*, 2007, 37(5): 1271 – 1289.
- [11] JIANG M, LUO Y P, YAGN S Y. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm[J]. *Information Processing Letters*, 2007, 102(1): 8 – 16.
- [12] 金欣磊, 马龙华, 吴铁军, 等. 基于随机过程的PSO收敛性分析[J]. *自动化学报*, 2007, 33(12): 1263 – 1268.
(JIN Xinlei, MA Longhua, WU Tiejun, et al. Convergence analysis of the particle swarm optimization based on stochastic processes[J]. *Acta Automatica Sinica*, 2007, 33(12): 1263 – 1268.)
- [13] HAN M, FAN J C, HAN B. An adaptive dynamic evolution feed-forward neural network on modified particle swarm optimization[C] // *Proceedings of the 2009 International Joint Conference on Neural Networks*. Atlanta, USA: IEEE, 2009: 1083 – 1089.
- [14] ZIELINSKI K, WEITKEMPER P, LAUR R, et al. Optimization of power allocation for interference cancellation with particle swarm optimization[J]. *IEEE Transactions on Evolutionary Computation*, 2009, 13(1): 128 – 150.
- [15] HAN M, FAN J C. Delay nonlinear system predictive control on MPSO+DNN[C] // *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*. San Antonio, USA: IEEE, 2009: 4309 – 4314.
- [16] SONG Y, CHEN Z Q, YUAN Z Z. New chaotic PSO-based neural network predictive control for nonlinear process[J]. *IEEE Transactions on Neural Networks*, 2007, 18(2): 595 – 600.
- [17] BOYD S, GHAOUI L E, FERON E, et al. *Linear Matrix Inequalities in System and Control Theory*[M]. Philadelphia: SIAM, 1994.
- [18] 刘威, 李小平, 柴天佑, 等. 基于实数编码遗传算法的神经网络成本预测模型及其应用[J]. *控制理论与应用*, 2004, 21(3): 423 – 426, 431.
(LIU Wei, LI Xiaoping, CHAI Tianyou, et al. Neural network cost prediction model based on real-coded genetic algorithm and its application[J]. *Control Theory & Applications*, 2004, 21(3): 423 – 426, 431.)

作者简介:

范剑超 (1985—), 男, 博士研究生, 主要研究方向为计算智能、神经网络和预测控制, E-mail: fjchao@mail.dlut.edu.cn;

韩敏 (1959—), 女, 教授, 博士生导师, 主要研究方向为神经网络、混沌时间序列和智能控制和建模, E-mail: minhan@dlut.edu.cn, 通信作者.