

## 暂态混沌神经网络中的模拟退火策略优化

李薪宇, 吕炳朝

(电子科技大学 自动化工程学院, 四川 成都 610054)

(lixu@uestc.edu.cn)

**摘要:**分析了暂态混沌神经网络(TCNN)模型的动力学特性对自反馈连接权值的敏感性,研究了退火函数对优化过程中的准确性和计算速度的影响。给出一种对模拟退火策略的优化算法,在保证优化算法准确性的基础上,加快收敛速度,并利用两个典型的函数优化的例子验证了这种优化策略的有效性。

**关键词:**暂态混沌神经网络;自反馈连接权;函数优化

**中图分类号:** TP183 **文献标识码:** A

## Optimal strategy for simulated annealing mechanics in transiently chaotic neural networks

LI Xin-Yu, LÜ Bing-chao

(School of Automation Engineering, University of Electronic Science & Technology of China, Chengdu Sichuan 610054, China)

**Abstract:** This paper analyzed that the dynamic characteristics of transiently chaotic networks (TCNN) quite sensitively depend on value of the self-feedback connection weights, and researched the annealing function that intensively influences the veracity and search speed of TCNN module. It is proposed an optimal strategy for value of the self-feedback connection weights that can accelerate the search speed and guarantee the assurance of the veracity of the optimal arithmetic. To demonstrate the validity of this optimal strategy, two examples of function optimization are given.

**Key words:** transiently chaotic network; self-feedback connection weights; function optimization

### 0 引言

随着神经网络的发展,神经网络的寻求稳定平衡点的大规模集体计算能力,已经被广泛地应用于求解大量困难的优化问题,取得了显著的效果。但是由于其利用梯度下降的动力学,因此这种网络在求解许多实际优化问题时所遇到的最大困难是极易陷入局部极小点。为了避免这一问题,人们将混沌动力学的全局搜索特性引入神经网络之中,提出了多种混沌神经网络模型,并取得了良好的效果<sup>[1~4]</sup>。

目前,用于优化搜索的混沌神经网络模型,大多数是在Hopfield网络基础上加上自反馈项,通过其产生混沌现象,故网络的动态特性很敏感地依赖于自反馈连接权值 $z_i(t)$ 。 $z_i(t)$ 类似于随机模拟退火中的温度,其一般按指数退火函数动态变化,对网络的优化性能和收敛速度有很大的影响。有关学者针对这一问题,提出了一些改进方法,并取得了好的效果<sup>[5~7]</sup>。本文利用一种改进的变指数退火函数,提出对自反馈权值 $z_i(t)$ 的一种新优化策略,既充分利用混沌的动态特性进行搜索,又克服其带来的速度问题,减少收敛时间。

### 1 具有暂态混沌和时变增益的混沌神经网络

混沌动力学特性与收敛动力学特性结合起来可以更好地求解非线性问题的全局最优解,综合以前各种神经网络模型,采用一种具有暂态混沌和时变增益的混沌神经网络模型,其动力学特性定义为:

$$x_i(t) = \frac{1}{1 + e^{-\mu y_i(t)(1 + \varepsilon_i(t))}} \quad (1)$$

$$y_i(t+1) = ky_i(t) + \alpha \left( \sum_{j=1, j \neq i}^n w_{ij} x_j(t) + I_i \right) - z_i(t)(x_i(t) - I_i) \quad (2)$$

$$z_i(t+1) = (1 - \beta)z_i(t) \quad (3)$$

$$\varepsilon_i(t+1) = (1 - \gamma)\varepsilon_i(t), i = 1, 2, \dots, n \quad (4)$$

式中 $x_i, y_i$ 和 $I_i$ 为第 $i$ 个神经元的输出,内部状态和输入偏置; $w_{ij}$ 为从第 $j$ 个神经元到第 $i$ 个神经元的连接权; $\alpha$ 为比例参数; $k(0 \leq k \leq 1)$ 为神经元的衰减因子; $z_i(t)(z_i(t) \geq 0)$ 为自反馈连接权; $\beta(0 \leq \beta \leq 1)$ 为时变参量 $z_i(t)$ 的衰减因子; $\mu$ 为 $x_i(t)$ 的陡度参数; $\varepsilon_i(t)(\varepsilon_i(t) \geq 0)$ 为输出函数的时变增益参量; $\gamma(0 \leq \gamma \leq 1)$ 为时变参量的衰减因子。随着时间变化,当自反馈连接权 $z_i(t)$ 以指数方式(即: $z_i(t) = z_i(0)e^{-\beta t}$ )趋于零时,此网络逐渐退化为一个HNN。故此网络用来求解非线性优化问题的过程可分为两个阶段:混沌搜索阶段和梯度收敛阶段。在第一个阶段,由自反馈项来产生一个混沌过程以避免陷入网络的局部最小问题。第一阶段结束后,可以为第二阶段提供一个全局最优解附近的初始值。可见如何控制第一阶段产生的混沌搜索过程,是利用混沌神经网络解决非线性优化问题的关键。所以 $z_i(t)$ 的演变策略对优化的性能和时

### 2 模拟退火策略的优化算法

为了理解上述网络模型的运行机理,以单个神经元为例

收稿日期:2005-04-14;修订日期:2005-07-12

作者简介:李薪宇(1979-),男,四川营山人,助教,硕士研究生,主要研究方向:模式识别与智能系统;吕炳朝(1938-),男,云南人,教授,硕士,主要研究方向:控制系统理论。

来检验该网络的动力行为:

单个神经网络模型( $\alpha = 0$ )为:

$$\begin{cases} y(t+1) = ky(t) - z(t)(x(t) - I_0) \\ x(t) = \frac{1}{1 + e^{-\mu y(t)(1+\varepsilon(t))}} \\ z(t+1) = (1-\beta)z(t) \\ \varepsilon(t+1) = (1-\gamma)\varepsilon(t) \end{cases} \quad (5)$$

式(5)中的参数值为:

$k = 1, \beta = 0.001, \mu = 250, \gamma = 0.001, I_0 = 0.65, y(0) = 0.5, z(0) = 0.008, \varepsilon(0) = 0.6$ 。

图 1 是网络参数为以上值时的网络输出  $x(t)$ , 自反馈连接权  $z(t)$  的时间演化图。

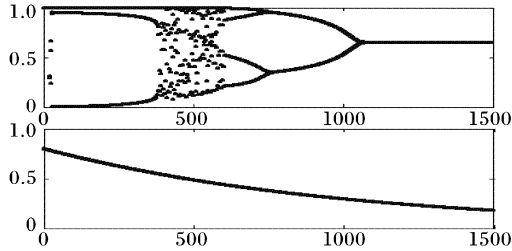


图 1 时间演化图( $\beta = 0.001$ )

从图 1 可以看出,该网络具有暂态混沌动力学的行为:随着  $z(t)$  在时间上的不断衰减,通过一个连续的混沌搜索过程,网络逐渐趋于稳定的平衡点。

下面通过改变网络参数  $\beta$  而保持其余参数不变来研究其对网络混沌动态行为的影响。

图 2 是网络参数  $\beta = 0.004$  时的网络输出  $x(t)$ , 自反馈连接权  $z(t)$  的时间演化图。

从图 2 中可以看出,随着参数  $\beta$  的增大,网络输出  $x(t)$  的混沌动态消失得更快;而且因为自反馈连接权  $z_i(t)$  以指数方式(即:  $z_i(t) = z_i(0)e^{-\beta t}$ ) 衰减,所以其趋于 0 的收敛速度也随着参数  $\beta$  的增大而加快了。

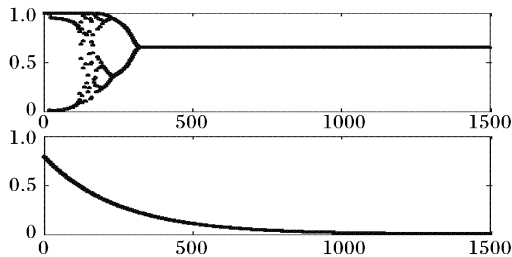


图 2 时间演化图( $\beta = 0.004$ )

由图 1、2 可知,网络的混沌动态特性很敏感地依赖于自反馈连接权值  $z_i(t)$ ,  $z_i(t)$  的下降速度直接影响到判断优化算法的两个重要指标:准确性和速度。当  $z_i(t)$  下降速度很快时,将通过短暂的搜索阶段直接进入收敛过程,虽然算法的整体速度很快,但因为没有充分利用混沌的丰富动态特征,容易陷入局部最小值,准确性大大降低;相反地,如果  $z_i(t)$  变化小,可以提高准确性而牺牲了优化速度。

通过研究大量利用混沌神经网络进行函数优化的文献[3]、[4],发现其虽然都能得到全局最优解,但收敛速度太慢。究其原因主要有两个:1) 在整个优化过程中只采用单一的参数  $\beta$ ,使得  $z_i(t)$  的动态特性变化过于单一,造成网络退火策略无法同时满足准确性和速度两方面的要求;2) 当网络的输出趋于稳态后,即得到了一个全局最优解附近的值,其自反馈项  $z_i(t)(x_i(t) - I_0)$  还是存在较小的数值,每次叠代对网络

第二阶段的梯度收敛过程有扰动,使得网络不得不用较长的时间来收敛到全局最优解。下面针对以上情况提出一个对自反馈项的优化策略:

文献[5]针对  $z_i(t)$  的动态特性变化过于单一的问题,采用分段指数退火函数(6)来代替原网络模型中的式(3):

$$z_i(t+1) = \begin{cases} (1-\beta_1)z_i(t), & z_i(t) > z_i(0)/2 \\ (1-\beta_2)z_i(t), & \text{others} \end{cases} \quad (6)$$

式中  $\beta_1, \beta_2$  为常数,  $\beta_1 < \beta_2$ 。

开始时  $z_i(t)$  的值较大,而参数  $\beta$  值较小可以充分利用混沌的丰富动态特征,使得网络输出值能在大范围内进行遍历搜索,使算法可以从局部最优值中跳出,而得到全局最优解;随着  $z_i(t)$  值的减小,网络输出逐渐收敛于分岔点,所以可以采用比较大的指数衰减,减小收敛时间。

但是我们发现当网络的输出趋于一个稳态值后,网络就转入梯度收敛阶段,此时希望网络能迅速收敛到全局最优解,然而自反馈项对此阶段的收敛过程有扰动,上述(6)式不能解决此问题,所以我们对(6)式进行改进:当网络的输出趋于一个稳态值后,可以将  $z_i(t)$  的值直接令为零,来消除其对梯度收敛阶段的扰动。

而网络混沌动力特性的结束可以用网络输出值  $x_i(t)$  的变化来确定,在  $z_i(t) \leq z_i(0)/2$  后当  $|x_i(t+1) - x_i(t)|$  小于一个给定的足够小的正值,就认为网络输出由混沌状态逐渐趋于稳定的平衡状态,即可令  $z_i(t)$  等于零。

用函数表达式(7)来描述此改进优化策略:

$$z_i(t+1) = \begin{cases} (1-\beta_1)z_i(t), & \text{if } z_i(t) > z_i(0)/2 \\ (1-\beta_2)z_i(t), & \text{if } z_i(t) \leq z_i(0)/2 \\ & \text{and } |x_i(t+1) - x_i(t)| > \delta \\ 0, & \text{if } z_i(t) \leq z_i(0)/2 \\ & \text{and } |x_i(t+1) - x_i(t)| < \delta \end{cases} \quad (7)$$

式中  $\beta_1, \beta_2$  为常数,  $\beta_1 < \beta_2$ ,  $\delta$  为给定的足够小的正值(一般取  $10^{-3}$  量级)。

### 3 函数优化的仿真分析

以下两个函数优化的例子来检验此自反馈项的优化策略运用于具有暂态混沌和时变增益的混沌神经网络模型后带来的优化效果。

例 1 求解下述非线性优化问题

$$\min f_1(x_1, x_2) = (x_1 - 0.7)((x_2 + 0.6)^2 + 0.1) + (x_2 - 0.5)^2((x_1 + 0.4)^2 + 0.15) \quad (8)$$

目标函数(8)的全局最优解为(0.7, 0.5),局部极小值有 3 个,分别为(0.6, 0.4), (0.6, 0.5) 和 (0.7, 0.4)。

运用上述网络模型和自反馈项的优化策略来求解此函数的优化时,令:

$$\sum_{j=1, j \neq i}^n w_{ij}x_j(t) + I_i = -\frac{\partial f_1}{\partial x_i}$$

网络参数取值:  $k = 1, \beta_1 = 0.02, \beta_2 = 0.1, \alpha = 0.005, \mu = 250, \gamma = 0.001, I_0 = 0.65, \delta = 0.001$ 。

采用文献[4]中的网络初始值为:

$$y(0) = [-0.283, -0.047], z(0) = [0.065, 0.065], \varepsilon(0) = [0.08, 0.08]$$

图 3(a)、(b) 分别给出了采用文献[4]中的网络和采用本文自反馈项优化策略的网络计算出的神经元 1 的输出  $x_1(t)$  随时间  $t$  的演化图。从图 3(a) 中可以看出采用文献[4]

的网络,神经元 1 要经过大约 300 次叠代后才稳定于 0.7 值处;而从图 3(b)中可以看出采用本文自反馈项优化策略的网络,神经元 1 只需要经过大约 60 次叠代后就稳定于 0.7 值处,大约节约了 4/5 的叠代次数。对比图 3(a) 和图 3(b),无论从第一阶段混沌搜索还是第二阶段的梯度收敛,采用本文自反馈项优化策略的网络的收敛速度均大大优于文献[4] 的网络,而且这种优化策略并不以减小准确性为代价。

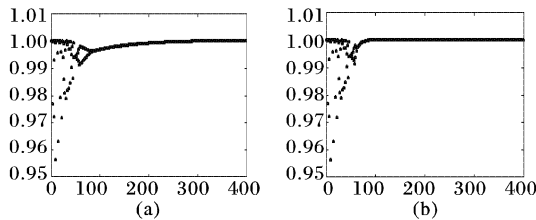


图 3 神经元 1 的输出

例 2 Rosenbrock 函数(又称香蕉函数)的优化问题<sup>[8]</sup>

$$\min f_2(x_1, x_2) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2 \quad (9)$$

该函数的全局最小点为(1,1),相应的最小值为 0。在 Rosenbrock 函数曲面山谷中的点的最速下降方向几乎与到函数最小值的最佳方向垂直,因而许多传统算法(如最速下降法)较难求解此问题。采用文献[4] 中的网络参数初始值为:

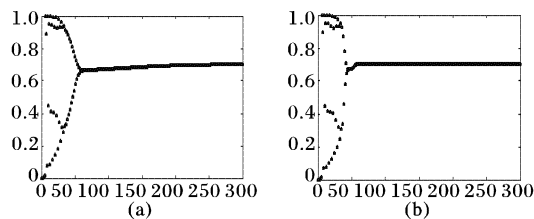


图 4 神经元 1 的输出

$$k = 1.015, \alpha = 0.0017, I_0 = 0.495, \mu = 250, \beta_1 = 0.01,$$

(上接第 2389 页)

数:  $k_1 = 1, k_2 = 1, k_3 = 0.5, k_4 = 0.5, k_5 = 0.06, k_6 = 0.1$ , 排挤操作的参数  $N_t = 20, \delta_F = 0.01, \delta_t$  按如下方式随进化代数  $t$  变化:

$$\delta_t = 0.12 - (t - 1) \frac{0.11}{N_{gen}} \quad (8)$$

其中  $N_{gen}$  是预先设置的最大进化代数,使用欧几里德距离作为个体之间的度量。通过多次对比实验,对 Musk1 数据集来说,基于反向传播训练的局部搜索算子中训练周期数目  $T_t$  取 5 左右的值比较合适,故本文实验时取了 5。停机条件中  $T_c$  和  $N_{gen}$  分别取为 0.05 和 90 000。本文算法和 Srinivas 算法的种群规模都取 220,染色体初始化使用均匀分布。

#### 4 结语

1) 多示例神经网络的反向传播训练中由于不可微函数需要使用近似函数,而且反向传播训练方法容易陷入局部极值,因此多示例神经网络的预测准确性不高。为了提高预测准确性,本文构造了一类优化多示例神经网络参数的改进遗传算法。

2) 由于多示例神经网络参数组成的染色体的长度相当长,优化难度比较大,通常的进化算法因局部搜索能力较弱,难以较快地求出多示例神经网络参数的近似全局最优值,提出了借助基于反向传播训练的局部搜索算子、排挤操作和适应性操作概率计算方式来提高收敛速度和防止早熟收敛。

3) 通过公开的多示例测试集上实验结果的分析 and 对比,

$$\beta_2 = 0.1, \gamma = 0.001, \delta = 0.001.$$

网络的初始值为:  $y(0) = [0.034 \ 6, 0.85], z(0) = [0.03, 0.031], \varepsilon(0) = [0.353, 0.008]$ 。

同例 1 图 4(a)、(b) 分别给出了采用文献[4] 中的网络和采用本文自反馈项优化策略的网络计算出的神经元 1 的输出  $x_1(t)$  随时间  $t$  的演化图。网络的输出均能最终达到 Rosenbrock 函数的全局最小点(1,1),但从图 4(a)、(b) 的对比中,可以明显看出采用本文自反馈项优化策略的网络的优化性能优于文献[4] 中的网络。

#### 参考文献:

- [1] TANK DW, JHOFIELD J. Simple "Neural" Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit[J]. IEEE Trans Circuit Syst, 1986, 33(4): 533-541.
- [2] CHEN L, AIHARA K. Chaotic simulated annealing by a neural network model with transient chaos[J]. Neural networks, 1995, 8(6): 915-930.
- [3] 谭营, 王保云, 何振亚. 一种具有暂态混沌和时变增益的神经网络及其在优化计算中的应用[J]. 电子学报, 1998, 25(7): 818-822.
- [4] 张强, 马润年, 许进. 一种混沌神经网络模型及其在优化中的应用[J]. 系统工程与电子技术, 2002, 24(2): 48-50.
- [5] 谢传泉, 何晨. 混沌神经网络模型中的模拟退火策略[J]. 上海交通大学学报, 2003, 37(3): 36-39.
- [6] 康波. 混沌控制与混沌优化及其在组合电路测试生成中的应用研究[D]. 电子科技大学博士学位论文, 2002, 10.
- [7] ZHOU C, CHEN T. Chaotic annealing for optimization[J]. Physical Review E, 1997, 55(3): 2580-2587.
- [8] 梁艳春, 周春光, 李寿范. 基于遗传算法的 Rosenbrock 函数优化问题的研究[J]. 软件学报, 1997, 8(9): 701-708.

证实了改进型遗传算法能够明显地提高多示例神经网络的预测准确性,而且具有较快的收敛速度。

#### 参考文献:

- [1] DIETTERICH TG, LATHROP RH, LOZANO - PEREZ T. Solving the multiple-instance problem with axis - parallel rectangles [J]. Artificial Intelligence, 1997, 89(1-2): 31-71.
- [2] RAMON J, RAEDT LD. Multi-instance neural networks [A]. Proceedings of ICML-2000 workshop on Attribute - Value and Relational Learning[C], 2000.
- [3] 李敏强, 徐博艺, 寇纪松. 遗传算法与神经网络的结合[J]. 系统工程理论与实践. 1999, 19(2): 107-112.
- [4] SIDDIQUE MNH, TOKHI MO. Training neural networks: back-propagation vs. genetic algorithms [A]. In: Proceedings of the IEEE International Joint Conference on Neural Networks[C], 2001, 4: 2673-2678.
- [5] SRINIVAS M, PATNAI LM. Adaptive probabilities of crossover and mutation in genetic algorithms [J]. IEEE Trans. Syst., Man, and Cybern, 1994, 24(4): 656-666.
- [6] BLAKE C, KEOGH E, MERZ CJ. UCI Repository of Machine Learning Databases. <http://www.ics.uci.edu/~mlearn/MLRepository.html>, Department of Information and Computer Science, University of California, Irvine, CA, 1998.
- [7] TOM M. 机器学习[M]. 曾华军, 张银奎, 译. 北京: 机械工业出版社, 2003.