

反馈网络在优化问题应用中的理论研究

赵启林 卓家寿

(河海大学土木工程学院 , 江苏 南京 210098)

摘要 : 针对反馈网络在优化问题求解中的问题 , 探讨了反馈网络与能量函数之间的关系 , 证实了构造的神经网络只有满足一定条件时 , 网络平衡态才可能对应能量函数极小点 . 同时文中给出了 3 条用于优化求解网络设计的设计准则和方法 .

关键词 : 工程优化 , 平衡态 , 网络设计

中图分类号 : O224 文献标识码 : A 文章编号 : 1000-198X(2001)03-0073-03

美国物理学家 Hopfield 教授于 1984 年提出了著名的 Hopfield 网络^[1] , 并用一组微分方程表达如下 :

$$\begin{cases} C_i \frac{d\mu_i}{dt} = \sum_{j=1}^N T_{ij} V_j - \frac{\mu_i}{R_i} + I_i \\ V_i = g(\mu_i) \end{cases} \quad i = 1, 2, \dots, N \quad (1)$$

式中 : $T_{ij} = T_{ji}$, g ——可微的严格单调上升的函数 ; $C_i, R_i > 0$. 同时 Hopfield 构造了一个能量函数

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} V_i V_j - \sum_{i=1}^N V_i I_i + \sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i} g^{-1}(t) dt \quad (2)$$

随后 Hopfield 严格证明了

$$\frac{dE}{dt} \leq 0, \text{ 当且仅当 } \frac{dV_i}{dt} = 0 \text{ 时 } \frac{dE}{dt} = 0 \quad (3)$$

基于 Hopfield 网络总是朝着能量 E 减少的方向运行 , 而且网络的稳定平衡点就是 E 的极小点这一事实 , 许多学者提出了许多用于工程优化问题求解的广义的 Hopfield 网络^[2~5] . 这些网络的共同特点是将优化目标函数与约束条件利用罚函数法建立一个能量函数 , 然后再利用这一能量函数给出一个神经网络动态演化方程 $\dot{X} = f(X)$ 使式 (3) 成立 . 认为网络的解就是优化问题的一个解 , 但是许多数值模拟结果却是失败的 . TSP 问题就是一个典型的实例^[6] . Wilson 在利用 Hopfield 网络求解 TSP 问题时发现它在 100 次的重复试验中 , 只有 15 次能在 1000 次的迭代中收敛到有意义的路径 , 40 次在 1000 次迭代内不收敛 , 其他 45 次则毫无意义 . 本文则通过讨论分析说明 : 用神经网络进行优化问题的求解时 , 能量函数与网络动态演化方程仅仅满足式 (3) 不一定能够得到优化问题的解 .

1 用于优化问题的神经网络稳定性

当我们构造一个网络用于优化问题的求解时 , 首先考虑的是该网络的稳定性 , 即网络从任何初值都可以收敛到平衡态 . 下面的定理将说明仅满足式 (3) 是不能保证所构造的神经网络是稳定的 .

定理 1^[7] 连续自治动力系统的稳定性判据 : 对于由微分方程

$$\dot{X} = f(X) \quad (4)$$

所表示的系统 $f: R^n \rightarrow R^n$ 是连续的 , 它是完全稳定的条件 : 存在一标量函数 $V(X): R^n \rightarrow R^1$, $V(X)$ 有一阶偏连续导数并且 (a) $\frac{dV}{dt} \leq 0$ 且 $V(t) = 0$ 时 , 当且仅当 $\frac{dX}{dt} = 0$; (b) $X = f(X)$ 的所有解都是有界的 .

可见如果要使构造的网络是稳定的 , $X = f(X)$ 的解必须有界 . 而在实际应用中 , 许多网络都没有考虑到这一点 . 而 Hopfield 网络在满足式 (3) 的条件下是稳定的 , 是因为已经证明 Hopfield 网络在 Hopfield 严格的假设条

件下网络的解是有界的^[8]。一旦这些条件被破坏,网络就可能不稳定。下面取一个简单的算例来说明这个问题。
 g 不取连续可微 Sigmoid 函数,而取线性函数 $v = g(u) = u$ 。对简单的一维问题 $\frac{dX}{dt} = X$ 来说,方程的解是 $X(t) = X(0)e^t$,可知该解是无界的,因而网络是不稳定的。计算机模拟也不能够得到方程平衡态的解 0。

2 能量函数极小点与平衡态的关系

事实上即使根据优化问题的目标函数与约束条件而构造的神经网络是稳定的,网络的平衡态也不一定对应能量函数的极小点。下例可以说明。

取能量函数 $E(X) = -X$,网络的动态演化方程 $\frac{dX}{dt} = X^2$,可以证明这一网络是稳定的。

证 $\frac{dE}{dt} = -\frac{dX}{dt} = -X^2 \leq 0$,且 $\frac{dE}{dt} = 0$ 当且仅当 $\frac{dX}{dt} = 0$ 。由 $\frac{dX}{dt} = X^2$ 得 $-dX^{-1} = dt$,从而有 $X = (\frac{1}{X_0} - t)^{-1}$, $\lim_{t \rightarrow \infty} X(t) = 0$ 。可见网络的解是有界的,并且网络的平衡态 $X = 0$ 。但它明显不是 $E(X)$ 的极小点。这是因为在平衡态下 $\frac{dX}{dt} = 0$,根据式(3)有 $\frac{dE}{dt} = 0$,但不一定有 $\frac{dE}{dX} = 0$ 。那么平衡态就不一定满足能量函数 E 极小的一阶必要条件,故网络的平衡态也就不可能得到优化问题的极小解。上面所取例子就属于这种情况。

另一种情况下,如果 $\frac{dE}{dX}$ 与 $\frac{dX}{dt}$ 存在

$$\frac{dE}{dX} = F(X) \frac{dX}{dt} \text{ 或者由 } \frac{dX}{dt} = 0 \text{ 一定有 } \frac{dE}{dX} = 0 \tag{5}$$

则平衡态对应能量函数极小的一阶必要条件。因为在平衡态下 $\frac{dX}{dt} = 0$ 有 $\frac{dE}{dX} = 0$ 。但即使这样,由网络的解也不一定得到优化问题的解,因为式(5)仅满足了优化问题的一阶必要条件,并不是充分条件。

现用一个算例来说明这个问题。取网络的动态演化方程是 $\frac{dX}{dt} = 9X^2 - 4X^3$,能量函数 $E(X) = X^4 - 3X^3$ 。容易知道动态演化方程是有界的,而且是稳定的。 $X = 0, 2.25$ 是方程的平衡态。表 1 给出了不同开始点下 1000 次迭代后计算机模拟的收敛数值。可见不同的开始点将收敛到平衡态 0 与 2.25。但是 0 不是函数 E 的极小点,而是它的一个拐点。

但对于工程问题,当目标函数满足极小点的一阶必要条件时,一般可以得到一个极小点。因此,对于如下利用罚函数形成的优化问题: $\min E(X)$ 和它的网络动态演化方程 $\dot{X} = f(X)$ 应该满足如下几条准则 (a) $\frac{dE}{dt} \leq 0$ 且 $\frac{dE}{dt} = 0$ 时当且仅当 $\frac{dX}{dt} = 0$ (b) $\dot{X} = f(X)$ 的解有界 (c) $\frac{dE}{dX}$ 可以表达成 $\frac{dE}{dX} = F(X) \frac{dX}{dt}$, $F(X)$ 是任一函数。则满足以上 3 个准则的神经网络一般可以得到工程优化问题的极小解。

表 1 不同开始点的收敛值

Table 1 Convergent outcomes for different start-points

$X(0)$	$X(t)$	$E(X)$
-10.5	-0.0953	0.0027
-1.0	-0.0914	0.0024
0.5	-0.8520	0.0019
0.5	2.2500	-8.5430
10.5	2.2500	-8.5430

证明:由条件(a)(b)可知网络是稳定的,即最终网络的解满足 $\frac{dE}{dt} = 0$ 。由条件(c)得 $\frac{dE}{dX} = 0$,则网络的解对应 $E(X)$ 的极小的一阶必要条件。同时由条件(a)可知 $E(X)$ 是单调下降的,所以一般可以得到 $E(X)$ 的一个极小点。

3 优化问题的神经网络方法设计

考虑工程结构优化问题

$$\left. \begin{aligned} &\min f(X); X \in R^n \\ &\text{s.t. } g_k(X) \geq 0, k = 1, 2, \dots, K; h_p(X) = 0, p = 1, 2, \dots, P \end{aligned} \right\} \tag{6}$$

采用外罚函数法,构造其增广目标函数

$$F(X, \lambda) = f(X) + \lambda \sum_{k=1}^K \alpha(g_k(X)) + \lambda \sum_{p=1}^P H(h_p(X)) \tag{7}$$

这里 λ 是 Lagrange 乘子, $\alpha(y)$ 和 $H(y)$ 是罚函数:

$$\alpha(y) = \begin{cases} 0 & y \geq 0 \\ y^2/2 & y < 0 \end{cases} \quad H(y) = y^2/2 \tag{8}$$

以增广目标函数(7)作为网络的能量函数,首先取网络的动态演化方程满足条件(c):

$$\frac{dE}{dX} = -F(X) \quad F(X) \geq \alpha \text{ 对于任意定义域内的 } X \tag{9}$$

$F(X)$ 取 $1-X^2$ Sigmoid 函数等,则根据式(8)(9)有网络动态演化方程

$$\frac{dX}{dt} = \frac{1}{F(X)} \left(\frac{\partial f}{\partial X} + \lambda \sum_k \alpha(g_k) \frac{\partial g_k}{\partial X} + \lambda \sum_p \alpha(h_p) \frac{\partial h_p}{\partial X} \right) \tag{10}$$

很容易证明式(7)和(10)是满足准则(a)与(c)的,只要选取适当的 $F(X)$ 使式(10)的解有界,则一般情况下可得到问题(6)的极小解.

4 结 论

本文讨论了利用反馈神经网络求解工程优化问题时应注意的问题,并指出,根据优化目标函数所设计出来的神经网络仅仅能满足能量单调下降与平衡态对应 $\frac{dE}{dt} = 0$ 的条件,并不能保证网络稳定及得到能量函数的极小解,它还应该满足网络的解有界等具体的条件.这样就为反馈神经网络在工程结构优化设计中的应用提供了可靠的理论基础,同时还给出一个用于求解优化问题时网络设计的简单方法.这个方法中有界性的证明是网络设计成功的关键,但也是设计的难点,因为网络解的本身是待求的.

参考文献:

[1] Hopfield J J. Neural network and physical systems with emergent collective computational abilities[J]. Proc Matl Acad Sci, 1982, 79: 2554 ~ 2558.
[2] 吴剑国,赵丽萍.工程结构优化的神经网络方法[J].计算力学学报, 1998, 15(1): 70 ~ 74.
[3] 夏又生.求解绝对极小拟合问题的神经网络方法[J].东南大学学报, 1997, 27(2): 46 ~ 49.
[4] Vinod V V. Resulant projection neural networks for optimization under inequality constraint[J]. Neural Network, 1996, 2(4): 509 ~ 521.
[5] 周祖昊,郭宗楼.具有一类不可导多元的约束函数的非线性规划神经网络模型[J].武汉水利电力大学学报, 1999, 32(5): 102 ~ 105.
[6] 胡超,郑南宁,高岩.组合优化设计网络的胞元设计及稳定性分析[J].西安交通大学学报, 1996, 30(12): 27 ~ 33.
[7] 焦李成.神经网络系统理论[M].西安:西安电子科技大学, 1996. 60 ~ 67.
[8] 章毅,周明天,王平安.关于神经网络的能量函数[J].计算机研究与发展, 1999, 36(7): 794 ~ 799.

Theoretical Research of Application of Feedback
Neural Network to Optimization

ZHAO Qi-lin, ZHUO Jia-shou

(College of Civil Engineering, Hohai Univ., Nanjing 210098, China)

Abstract: In consideration of the problem of feedback neural network in solving optimization, the relation between the feedback neural network and energy function is discussed. It is proved that, only if the neural network satisfies some conditions, the equilibrium state is the least point of Energy function. At the same time, three design criteria for design of the neural network are put forward.

Key words: optimization; equilibrium state; network design