

解非线性方程组的神经网络方法

赵华敏, 陈开周

(西安电子科技大学理学院, 陕西西安 710071)

摘 要: 本文提出了一种求解非线性方程组的神经网络方法, 该方法对非线性方程组的任意给定的初始点, 都能稳定地收敛到它的一个实根. 文中首先严格地证明了该方法的稳定性、收敛性及可行性, 然后给出了一个模拟算法及其应用. 最后的数值试验结果表明该方法是有用的.

关键词: 非线性方程组; 神经网络; 能量函数; 稳定性

中图分类号: O241.7 **文献标识码:** A **文章编号:** 0372-2112 (2002) 04-0601-04

Neural Network for Solving Systems of Nonlinear Equations

ZHAO Hua-min, CHEN Kai-zhou

(School of Science, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract: A neural network for solving systems of nonlinear equations is proposed in this paper. For an arbitrarily given initial point of systems of nonlinear equations, the network converges to a real solution stably. First, the stability, convergence and feasibility of the network are strictly proved. Then a simulation algorithm and its applications are provided. Finally, some numerical tests are given to show the effectiveness of the network.

Key words: systems of nonlinear equations; neural network; energy function; stability

1 引言

自然生活和工程技术中的许多问题常常归结为解非线性方程组, 因此, 对非线性方程组的解法的研究具有重要的意义. 已有许多学者对此做了较好的研究, 例如, 文[1]针对一元非线性方程的求根问题, 提出了二阶收敛指数迭代法, 文[2]提出了求解非线性方程组的连续极小化方法, 但该文得到的结果还不太完善.

另外, 由于神经网络具有许多优点, 如大规模并行处理、分布式存储以及网络的计算时间复杂度几乎为零等^[3], 因此, 神经网络的应用已越来越广泛, 例如, 文[4]把神经网络用于解多目标动态规划问题, 文[5]则提出了用于解优化问题和信号处理的各种神经网络系统.

由此, 本文把神经网络用于非线性方程组的求解, 并得出了一种解非线性方程组的新方法, 该方法对非线性方程组的任意给定的初始点, 都能稳定地收敛到它的一个实根. 文中严格地证明了该方法的稳定性、收敛性及可行性, 得到了完善的结论, 并说明了该方法也可用于解多元非线性方程及线性方程组. 另外, 该方法还可用于全局优化问题的求解等许多实际问题之中.

2 神经网络模型

考虑非线性方程组

$$F(x) = 0 \quad (1)$$

其中 $F(x) = (f_1(x), f_2(x), \dots, f_m(x))^T \in R^m$, $x \in R^n$, $m \leq n$, $f_i: R^n \rightarrow R^1$, $i = 1, 2, \dots, m$ 为二阶连续可微函数. 设方程组(1)的全部实根的集合为

$$M = \{x \in R^n | F(x) = 0\}$$

又设 $M \neq \emptyset$, 且在 $R^n - M$ 上,

$$F(x)^T F(x) > 0 \quad (2)$$

定义能量函数

$$E(x) = \frac{1}{2} F(x)^T F(x)$$

显然, x^* 为非线性方程组(1)的根等价于 x^* 为能量函数 $E(x)$ 的极小点. 因此, 为求方程组(1)的根, 可以用降能的方法求出能量函数 $E(x)$ 的极小点, 从而可建立如下的神经网络模型

$$\frac{dx}{dt} = -W P(x) F(x) \quad (3)$$

其中 $W = (w_{ij})_{nm}$ 为对称正定矩阵 (通常, 取 W 为对角元素均为正数的对角矩阵), 函数 $P(x) = (p_{ij}(x))_{nm}$ 通常依赖于 $F(x)$, $P(x)$ 连续可微, 且满足

$$F(x)^T F(x) W P(x) F(x) > 0, \forall x \in R^n - M$$

令 $G(x) = P(x) F(x)$, 且

$$G(x) = (g_1(x), g_2(x), \dots, g_n(x))^T$$

我们可作出神经网络(3)的框图^[5] (如图1所示), 其输入为 W, x , 输出为 x .

特别地,若取

$$P(x) = F(x)^T$$

则神经网络(3)化为

$$\frac{dx}{dt} = -WF(x)^TF(x) \quad (4)$$

显然,常微分方程组(3)与(4)的初值问题的解存在且惟一.

下面,我们将对神经网络(4)的稳定性及应用进行分析与探讨.

3 稳定性分析

设 $x = x(t, x^0)$ 为神经网络(4)的以 x^0 为初始点的积分曲线.

定理 1 x^* 为非线性方程组(1)的根的充要条件为:
 $\nabla E(x^*) = 0$.

证明 必要性:若 x^* 为非线性方程组(1)的根,则 $F(x^*) = 0$,从而

$$\nabla E(x^*) = F(x^*)^TF(x^*) = 0$$

充分性:若 $\nabla E(x^*) = 0$,则

$$F(x^*)^TF(x^*) = 0 \quad (5)$$

又由式(2)知, $x^* \in M$, 即 x^* 为方程组(1)的根.

因此,该定理的结论成立.

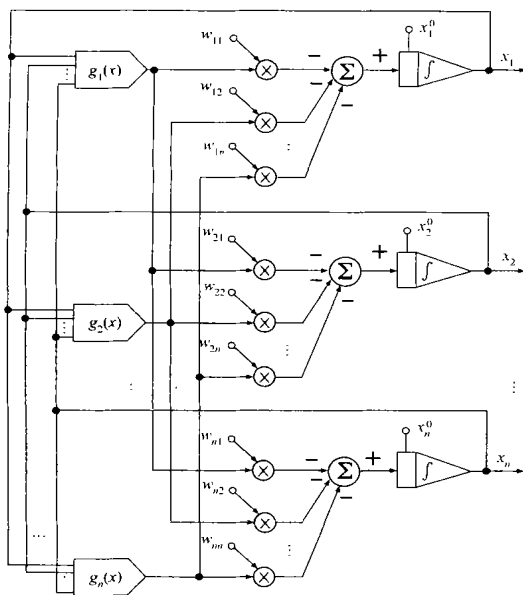


图 1 神经网络(3)的框图

由定理 1 可得如下推论:

推论 1 设神经网络(4)的平衡点集合为

$$E = \{x \in R^n | \nabla E(x) = 0\}$$

则 $E = M$.

对 x^* 的任一邻域

$$N(x^*) = \{x \in R^n | \|x - x^*\|_2 < \delta\}$$

其中 $\delta > 0$, 记其闭邻域为

$$\bar{N}(x^*) = \{x \in R^n | \|x - x^*\|_2 \leq \delta\}$$

定理 2 设非线性方程组(1)在 R^n 内有有限个实根,则

神经网络(4)一致渐近稳定收敛于方程组(1)的一个实根.

证明 沿神经网络(4)的轨线 $x = x(t, x^0)$, 有

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt} E(x(t, x^0)) = \nabla E(x)^T \frac{dx}{dt} \\ &= \nabla E(x)^T [-WF(x)^TF(x)] \end{aligned}$$

即

$$\frac{dE}{dt} = -\nabla E(x)^TW\nabla E(x) \quad (6)$$

又由式(2)可知,对任意 $x \in R^n - M$, 有

$$\nabla E(x) \neq 0$$

当 $x \in M$ 时,有 $\nabla E(x) = 0$. 故由式(6)得

$$\begin{cases} \frac{dE}{dt} < 0, & \text{对任意 } x \in R^n - M, \\ \frac{dE}{dt} = 0, & \text{当 } x \in M \end{cases} \quad (7)$$

若非线性方程组(1)在 R^n 内有有限个实根,设 x^* 为其任一实根,那么,必存在 x^* 的一个闭邻域 $\bar{N}(x^*)$, 其中 $\delta > 0$, 使得在 $\bar{N}(x^*)$ 上, $E(x)$ 为正定函数,并且,由式(7)知,沿轨线 $x = x(t, x^0)$ 的导数 $\dot{E}(x(t))$ 负定. 因此,由 Lyapunov 稳定性定理及推论 1 知,方程组(1)的根 x^* 是一致渐近稳定的,再由 x^* 的任意性知,该方程组的任一实根都是一致渐近稳定的,即结论成立.

推论 2 设非线性方程组(1)在 R^n 内有惟一实根 x^* , 则神经网络(4)一致渐近稳定收敛于方程组(1)的根 x^* .

定义 1^[6] 若集合 M 的导集为空集,则称集合 M 为离散集合.

定理 3 设非线性方程组(1)在 R^n 内有无穷多个实根,且集合 M 为离散集合,则神经网络(4)一致渐近稳定收敛于方程组(1)的一个实根.

证明 类似于定理 2.

对任意 $x^0 \in R^n$, 记能量函数 $E(x)$ 的水平集为

$$L(x^0) = \{x \in R^n | E(x) \leq E(x^0)\}$$

引理 1 假设水平集 $L(x^0)$ 为有界集,则任给初始点 x^0 , 当 $t \rightarrow +\infty$ 时,神经网络(4)的轨线 $x = x(t, x^0)$ 趋于它的一个平衡点,即

$$\lim_{t \rightarrow +\infty} x(t, x^0) = \bar{x}, \nabla E(\bar{x}) = 0$$

证明 (a) 首先,沿正半轨线 $x = x(t, x^0) (t \geq 0)$, 有

$$\dot{E}(x(t)) = -\nabla E(x)^TW\nabla E(x) \leq 0$$

因此,沿轨线 $x = x(t, x^0) (t \geq 0)$, $E(x(t, x^0))$ 是单调不增的.

(b) 证明正半轨线

$$^+(x^0) = \{x(t, x^0) | t \geq 0\}$$

是有界的.

事实上, $E(x) \geq 0$ 为有下界的连续函数,因此,由(a)知,

$$^+(x^0) = \{x(t, x^0) | t \geq 0\} \subseteq L(x^0)$$

且因 $L(x^0)$ 有界,所以 $^+(x^0)$ 是有界的.

(c) 证明存在 $\bar{x} \in L(x^0)$ 和严格单调递增数列 $\{t_n\}$, $t_n \rightarrow +\infty$, 使

$$\lim_{n \rightarrow +\infty} x(t_n, x^0) = \bar{x}$$

事实上,因 $\bar{x}^+(x^0)$ 是有界点集,取严格单调递增数列 $\{\bar{t}_n\}$:

$$0 \leq \bar{t}_1 < \bar{t}_2 < \dots < \bar{t}_n < \bar{t}_{n+1} < \dots, \bar{t}_n \rightarrow +\infty$$

则 $\{x(\bar{t}_n, x^0)\}$ 为有界点集,因而存在极限点 \bar{x} ,即存在严格单调子数列 $\{t_n\} \subseteq \{\bar{t}_n\}$, $t_n \geq 0$, $t_n \rightarrow +\infty$, 使

$$\lim_{n \rightarrow +\infty} x(t_n, x^0) = \bar{x} \quad L(x^0)$$

即 \bar{x} 为正半轨线 $\bar{x}^+(x^0)$ 的极限点.

(d) 最后证明 $\nabla E(\bar{x}) = 0$.

由 (a) 的证明知, $E(x)$ 是 R^n 上的 Lyapunov 函数, 且

$$\dot{E}(x(t)) = 0 \Leftrightarrow \nabla E(x) = 0$$

若令 S 为神经网络(4)的平衡点集中的最大不变集,那么,由 LaSalle 不变原理知,当 $t \rightarrow +\infty$ 时,有 $x(t, x^0) \in S$, 故

$$\lim_{t \rightarrow +\infty} x(t, x^0) = \lim_{n \rightarrow +\infty} x(t_n, x^0) = \bar{x}, \quad \nabla E(\bar{x}) = 0$$

因此,任给初始点 x^0 , 网络(4)的轨线 $x = x(t, x^0)$ 趋于它的一个平衡点 \bar{x} .

定理 4 假设水平集 $L(x^0)$ 为有界集,则任给初始点 x^0 , 当 $t \rightarrow +\infty$ 时,神经网络(4)的轨线 $x = x(t, x^0)$ 趋于非线性方程组(1)的一个实根,即

$$\lim_{t \rightarrow +\infty} x(t, x^0) = x^*, \quad F(x^*) = 0 \quad (8)$$

证明 由引理 1 与推论 1 显然可得.

由定理 4 可得如下推论:

推论 3 若非线性方程组(1)在 R^n 内有无穷多个实根,且水平集 $L(x^0)$ 为有界集,则任给初始点 x^0 , 当 $t \rightarrow +\infty$ 时,神经网络(4)的轨线 $x = x(t, x^0)$ 趋于方程组(1)的一个实根.

定理 5 若非线性方程组(1)在 R^n 内有惟一实根 x^* , 且水平集 $L(x^0)$ 为有界集,则神经网络(4)全局一致渐近稳定收敛于方程组(1)的根 x^* .

证明 由推论 2 与定理 4 显然可得.

定理 4 中的公式(8)给出了非线性方程组(1)的根与神经网络(4)的轨线 $x = x(t, x^0)$ 之间的关系,因此,可以通过对网络(4)进行求解,从而得到方程组(1)的一个实根.

4 模拟算法及其应用

根据上节中的理论分析,可以给出解非线性方程组(1)的神经网络算法.

步骤 1 初始化:令 $t=0$, 任选初始点 $x^0 = x(0) \in R^n$, 取 $t > 0$, 矩阵 $W > 0$, 精度 $\epsilon > 0$, 并令 $x = x^0$;

步骤 2 计算能量函数 $E(x)$ 的梯度:

$$\nabla E(x) = F(x)^T F(x);$$

步骤 3 进行网络状态更新:

$$x(t + \Delta t) = x - \Delta t W \nabla E(x);$$

步骤 4 计算 $s = \|F(x(t + \Delta t))\|_2$;

步骤 5 终止条件判断:若 $s < \epsilon$, 则停,令非线性方程组(1)的根

$$x^* \doteq x(t + \Delta t)$$

否则,令 $x = x(t + \Delta t)$, $t = t + \Delta t$, 转步骤 2.

我们可以用该算法解下列问题:

问题 1 解非线性方程组 $F(x) = 0$, 其中 $F: R^n \rightarrow R^n$, 且

对任意 $x \in R^n - M$, $F(x)$ 非奇异.

问题 2 解多元非线性方程

$$F(x) = ax_k + H(y) = 0$$

其中 $a \in R$, $y = (x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_n)^T \in R^{n-1}$, $H: R^{n-1} \rightarrow R^1$, $F: R^n \rightarrow R^1$, $1 \leq k \leq n$.

问题 3 解线性方程组

$$F(x) = Ax + b = 0$$

其中 $F: R^n \rightarrow R^n$, A 为 $n \times n$ 非奇异矩阵, $x \in R^n$, $b \in R^n$.

另外,对非线性方程组(1),若 $m = n$, 且对任意 $x \in R^n$, $F(x)$ 对称正定,则方程组(1)在 R^n 内有惟一实根 x^* ,并可建立如下的神经网络

$$\frac{dx}{dt} = -F(x) \quad (9)$$

显然,该网络比网络(3)与(4)简单.类似于前面对神经网络(4)的稳定性的分析,对网络(9),可得到如下定理:

定理 6 对非线性方程组(1),若 $m = n$, 且对任意 $x \in R^n$, $F(x)$ 对称正定,则网络(9)一致渐近稳定收敛于方程组(1)的根 x^* .

同样,若 $m = n$, 且对任意 $x \in R^n$, $F(x)$ 对称负定,则方程组(1)在 R^n 内有惟一实根 x^* ,并可建立如下的神经网络

$$\frac{dx}{dt} = F(x)$$

该网络一致渐近稳定收敛于方程组(1)的根 x^* .

5 数值试验

下面,我们给出利用上节的算法解非线性方程组、多元非线性方程及线性方程组的一些数值试验结果(在算法中,取 $t = 0.0001$, 矩阵 W 为单位矩阵,精度 $\epsilon = 1.0 \times 10^{-10}$).

表 1 例 1 的计算结果

初始点	近似根
$(1, 1)^T$	$(-0.451123, 0.445178)^T$
$(-1, 1)^T$	$(-0.451123, 0.445178)^T$
$(-1, -1)^T$	$(-0.451123, 0.445178)^T$
$(1, -1)^T$	$(-0.451123, 0.445178)^T$

例 1 求二元非线性方程组

$$\begin{cases} f_1(x) = 3x_1 + x_1^3 + x_2 + 1 = 0 \\ f_2(x) = x_1 + 2x_2 + e^{x_2} - 2 = 0 \end{cases}$$

在 R^2 中的实根.

令 $F(x) = (f_1(x), f_2(x))^T$, 分别取不同的初始点,由算法可得该方程组的近似根如表 1 所示.

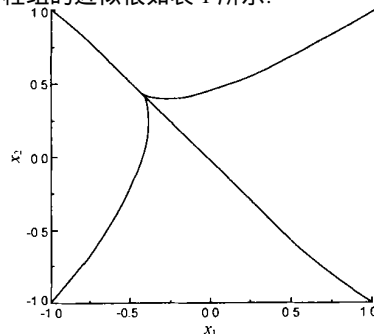


图 2 神经网络(9)的轨线

在图2中,我们画出了神经网络(9)的轨线,它们分别从四个不同的初始点收敛到该方程组的根.

例2 求三元非线性方程

$$F(x) = \sin x_1 - x_2^2 + x_3 - 2 = 0$$

在 R^3 中的实根.

分别取不同的初始点,由算法可得该方程的近似根如表2所示.

表2 例2的计算结果

初始点	近似根
$(1, 1, 1)^T$	$(1.159080, 0.506496, 1.340103)^T$
$(-1, 1, 1)^T$	$(-0.107423, 0.106727, 2.118607)^T$
$(-1, -1, -1)^T$	$(0.871353, -0.011438, 1.234930)^T$
$(1, -1, -1)^T$	$(1.491703, -0.018181, 1.003457)^T$

在图3中,我们画出了神经网络(4)的轨线(初始点 $x^0 = (-1, -1, -1)^T$).

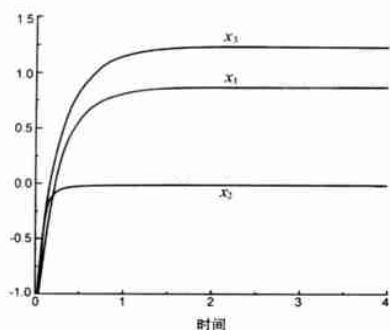


图3 神经网络(4)的轨线

例3 求线性方程组

$$F(x) = Ax + b = 0$$

在 R^n 中的实根,其中

$$A = \begin{pmatrix} 2 & 1 & 1 & \dots & 1 \\ 1 & 2 & 1 & \dots & 1 \\ 1 & 1 & 2 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 2 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 2 \\ 3 \\ \dots \\ n \end{pmatrix}$$

取 $n=20$, 初始点 $x^0 = (x_1^0, x_2^0, \dots, x_{20}^0)^T$, $x_i^0 = 0, i=1, 2, \dots, 20$, 由算法可得该方程组的根

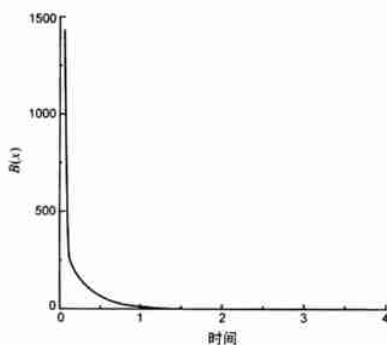


图4 神经网络(9)的能量函数 $E(x)$ 的动态变化过程

$$x^* = (9, 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -3, -4, -5, -6, -7, -8, -9, -10)^T.$$

图4显示了神经网络(9)的能量函数 $E(x)$ 的动态变化过程.

6 结论

本文提出了一种解非线性方程组的神经网络方法,该方法也可用于多元非线性方程及线性方程组的求解.此外,在模拟算法的步骤3中,采用不同的常微分方程组初值问题的数值解法就会得到不同的神经网络状态更新的计算公式,文中所提出的仅是网络状态更新的计算公式中的最简单的一种.数值试验结果表明,该方法对非线性方程组的任意给定的初始点 x^0 ,都能稳定地收敛到它的一个实根,因而,该方法是有效的.

参考文献:

- [1] 吴新元. 解非线性方程的二阶收敛指数迭代法[J]. 计算数学, 1998, 20(4): 367 - 370.
- [2] 侍乐媛. 求解非线性方程组的连续极小化方法[J]. 计算数学, 1987, 9(4): 438 - 445.
- [3] 焦李成. 神经网络系统理论[M]. 西安: 西安电子科技大学出版社, 1996. 1 - 25.
- [4] 胡铁松, 郭元裕. 多目标动态规划的神经网络方法[J]. 电子学报, 1999, 27(10): 70 - 73.
- [5] A Gichocki, R Unbehauen. Neural Networks for Optimization and Signal Processing [M]. New York: John Wiley & Sons, 1993. 88 - 101.
- [6] 江泽坚, 吴智泉. 实变函数论[M]. 北京: 高等教育出版社, 1989. 18 - 22.

作者简介:



赵华敏 女, 1976年8月生于河南省, 1997年毕业于西安电子科技大学应用数学系, 同年免试进入该系直接攻读博士学位. 主要研究方向为神经网络优化方法, 最优化理论及应用. e-mail: zhaohuamin @263.net



陈开周 男, 1930年10月生于四川省, 西安电子科技大学应用数学系教授, 博士生导师, 现任全国经济数学与管理数学学会理事长, 中国优选法、统筹法与经济数学研究会理事等职, 研究领域为神经网络优化方法, 最优化理论、算法与应用等.