

非线性方程组的耦合神经网络算法

赵启林 卓家寿

(河海大学土木工程学院 ,江苏 南京 210098)

摘要 :提出一种权值耦合的前馈神经网络模型 ,并探讨利用它进行非线性方程组求解的可能性 ,给出了具体的算法 ,最后求解了四个数值例子 .算例表明 ,用神经网络算法求解非线性方程组时可获得与解析法、牛顿法相同的结果 .
关键词 :耦合神经网络 ;非线性方程 ;BP 算法
中图分类号 :O241 文献标识码 :A 文章编号 :1000-1980(2000)05-0038-03

神经网络现在已成为国际上十分活跃的研究课题 ,它已被广泛地应用于各个领域 ,并且获得了极大的成功 .文献 [1] 利用线性方程组与神经网络在数学表达上的相似性 ,讨论了求解超定线性方程组的可能性 ,并且取得了满意的结果 .这种相似性对于非线性方程组已经不复存在 .基于此 ,本文提出了权值发生关联的耦合神经网络 .实践证明 ,这一网络模型可用于非线性方程组的求解 .

1 权值关联的耦合前馈神经网络

在传统的多层前馈神经网络中 ,各层连接权值与同层连接权值之间都是独立变化的 ,即在基于梯度法求解目标函数极小值时 ,权值修正是互不影响的 .这一模型在实际应用中不仅能较好地解决一部分问题 ,同时也能在一定程度上反映神经元的自学习特性 [2] .但现代脑科学研究发现 ,同层神经元之间也会相互影响 [3] .为了反映这一特性 ,我们通过权值耦合的方法 ,提出了一种新的网络模型——耦合前馈神经网络模型 .为了说明这一模型 ,我们首先给出基权和关联权的概念 .

a. 基权 W_j .权值本身是独立的 ,即通过网络的学习实现自身的调整 .其学习表达式为 $W_j(t+1) = W_j(t) - \eta \Delta W_j(t)$,其中 η 为修正系数 .

b. 关联权 W_k .权值本身是某一基权或几个基权的函数 : $W_k = F(W_j)$.它的调整通过基权的调整来实现 ,可用下式表达 :

$$W_k(t+1) = F_k(W_j(t+1)) = F_k(W_j(t) - \eta \Delta W_j(t)) \quad (1)$$

一个网络模型中 ,可以没有关联权 ,但至少应有一个基权 .没有关联权只有基权的耦合神经网络模型则可转化为传统的网络模型 .一个典型的二层耦合前馈神经网络模型如图 1 所示 .这是一个多输入、单输出的网络模型 ,有 n 个基权 ,即 $W_{ji} (i = 1, \dots, n)$, m 个关联权 ,即 $W_{ki} = F_{ki}(W_{j1}, \dots, W_{jn})$. f 为输出单元的处理函数 , $f =$

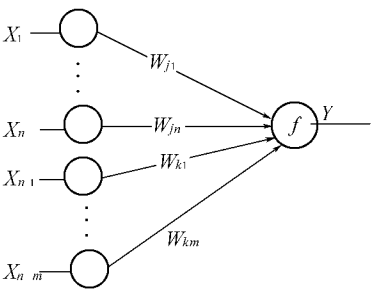


图 1 耦合神经网络模型

Fig.1 Model of coupling neural network

$f(\sum_{i=1}^n W_{ji} X_i + \sum_{i=n+1}^{m+n} W_{ki} X_i)$.它与传统的神经网络模型相似 .其学习与训练的基本思路是 :将输入加到网络上 (图 1) ,对每个输入向量 $X_p = (x_{p1}, x_{p2}, \dots, x_{p(n+m)})$ 计算其输出 Y_p ,而后按式

$$W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji}(t) \tag{2}$$

$$\Delta W_{ji}(t) = \eta \sum_{p=1}^l (T_p - Y_p) \cdot X_{pi} \tag{3}$$

修正每个基权 W_{ji} 的值 ,按

$$W_{ki}(t+1) = F_{ki}(W_{si}(t+1), \dots, W_{ji}(t+1)) \tag{4}$$

修正每个关联权 W_{ki} 的值。

式(3)中： η ——训练速率系数，一般取 $\eta \in (0, 1)$ ； T_p ——样本 p 的理想输出， Y_p ——实际输出； l ——总样本数。

其算法实施的步骤如下：

a. 将所有基权随机置值，并根据(4)式计算关联权的权值。

b. 加入每个输入向量 X_p ，计算相应的每个输出 Y_p ：

$$Y_p = f\left(\sum_{i=1}^n X_i W_{ji} + \sum_{i=1}^m X_{n+i} W_{ki}\right) \quad W_{ki} = F_{ki}(W_{j1}, \dots, W_{jn}) \quad p = 1, \dots, l \tag{5}$$

c. 计算基权的调整值
$$\Delta W_{ji}(t) = \eta \sum_{p=1}^l (T_p - Y_p) \cdot X_{pi} \tag{6}$$

d. 调整基权
$$W_{ji}(t+1) = W_{ji}(t) + \Delta W_{ji}(t) \tag{7}$$

e. 调整关联权
$$W_{ki}(t+1) = F_{ki}(W_{j1}(t+1), W_{j2}(t+1), \dots, W_{jn}(t+1)) \tag{8}$$

f. 判断是否达到误差要求，达到停止，否则回到 b。

2 非线性方程的神经网络算法

设有一非线性方程组为

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + a_{1(n+1)}F_0(x_1, \dots, x_n) + \dots + a_{1(n+m)}F_m(x_1, \dots, x_n) &= T_1 \\ a_{n1}x_1 + \dots + a_{nm}x_n + a_{n(n+1)}F_0(x_1, \dots, x_n) + \dots + a_{n(n+m)}F_m(x_1, \dots, x_n) &= T_n \end{aligned} \right\} \tag{9}$$

式中： $x_1, \dots, x_n$ ——基本变量； $F_i(x_1, \dots, x_n)$ ——基本变量的函数项。

分析这一非线性方程组可知，如果以系数向量 $A_i = (a_{i1}, \dots, a_{in}, \dots, a_{i(n+m)})$ 作为输入向量，对一有 n 个基权、 m 个相关权的二层耦合前馈神经网络进行学习，则其输出 Y_i 应为

$$Y_i = f\left(\sum_{l=1}^n a_{il}W_{jl} + \sum_{l=n+1}^{n+m} a_{il}W_{kl}\right) = f\left(\sum_{l=1}^n a_{il}W_{jl} + \sum_{l=n+1}^{n+m} a_{il}F_{kl}(W_{j1}, \dots, W_{jn})\right) \tag{10}$$

若取 $f(x) = x$ ，则式(10)变为
$$Y_i = \sum_{l=1}^n a_{il}W_{jl} + \sum_{l=n+1}^{n+m} a_{il}F_{kl}(W_{j1}, \dots, W_{jn}) \tag{11}$$

比较式(9)(11)可以看出，这两式在数学表达上完全相似，因此完全可以用这一耦合神经网络模型求解非线性方程组。其设计思路是：用 n 个基权对应方程组中的 n 个基本变量一次项或单变量项、 m 个关联权对应 m 个基本变量的函数项，然后构成一个有 $n+m$ 个输入、一个输出的耦合前馈神经网络；以方程组中的系数向量 $A_i = (a_{i1}, a_{i2}, \dots, a_{i(n+m)})$ 与数值 T_i 作为神经网络中的输入和输出的训练对，通过这些训练对反复对网络进行学习，直到权向量 $W = (W_{i1}, \dots, W_{in})$ 满足理想输出 T_i 为止。

3 数值算例

为验证算法求解非线性方程组的可行性，现给出几个算例。

例 1 非线性方程组为

$$\begin{cases} 3x_1 + 4x_2 + 5x_1x_2 = 1 \\ 5x_1 + x_2 + 4x_1x_2 = 7 \end{cases}$$

这里 η 取 0.0008，输出整体误差 E_0 取 0.00005，则基权为 $W_1 = x_1, W_2 = x_2$ ，关联权为 $W_3 = x_1x_2$ ，训练的输入样本 $x_1 = (3, 4, 5), x_2 = (5, 1, 4)$ ，输出样本 $y_1 = 1, y_2 = 7$ 。按神经算法求得的解如表 1 所示。

例 2 非线性方程组为

$$\begin{cases} 4x_1^2 + x_2^2 + 2x_1x_2 - x_2 = 2 \\ 2x_1^2 + x_2^2 + 3x_1x_2 = 1 \end{cases}$$

η 取 0.0004，输出整体误差 $E_0 = 0.00005$ ，则基权为 $W_1 = x_2, W_2 = x_1x_2$ ，关联权为 $W_3 = W_1^2, W_4 = (\frac{W_2}{W_1})^2$ ，训练的输入样本 $x_1 = (-1, 2, 1, 4), x_2 = (0, 3, 1, 2)$ ，输出样本 $y_1 = 2, y_2 = 1$ 。按神经算法求得的解如表 2 所

示.

表 1 例 1 的解值			表 2 例 2 的解值		
Table 1 The outcome of the first example			Table 2 The outcome of the second example		
方法	x_1	x_2	方法	x_1	x_2
神经算法	2.036 13	- 0.357 738	神经算法	0	- 0.997 265
解析法	2.076 92	- 0.363 636	解析法	0	- 1

例 3 非线性方程组为

$$\begin{cases} x_2 + \sin x_1 = 2.2 \\ 4x_1 + x_2^2 = 7 \end{cases}$$

η 取 0.000 8 输出整体误差 $E_0 = 0.000\,05$ 则基权为 $W_1 = x_1, W_2 = x_2$,关联权为 $W_3 = \sin W_1, W_4 = W_2^2$,训练的输入样本 $x_1 = (0\,1\,1\,0)$, $x_2 = (4\,0\,0\,1)$ 输出样本 $y_1 = 2.2, y_2 = 7$.按神经算法求得的解如表 3 所示.

例 4 非线性方程组为

$$\begin{cases} 4x_1^2 + x_2^2 + 2x_1x_2 - x_2 = 2 \\ 2x_1^2 + 3x_1x_2 + x_2^2 = 3 \end{cases}$$

η 取 0.008 整体误差 $E_0 = 0.000\,5$ 则基权为 $W_1 = x_2, W_2 = x_1x_2$,关联权为 $W_3 = (\frac{W_2}{W_1})^2, W_4 = W_1^2$,训练的输入样本 $x_1 = (-1\,2\,4\,1)$, $x_2 = (0\,3\,2\,1)$ 输出样本 $y_1 = 2, y_2 = 3$.神经算法与牛顿法的对比见表 4.

表 3 例 3 的解			表 4 例 4 的解		
Table 3 The outcome of the third example			Table 4 The outcome of the fourth example		
方法	x_1	x_2	方法	x_1	x_2
神经算法	1.446 22	1.108 87	神经算法	0.499 53	1.004 30
解析法	1.378 90	1.218 30	牛顿法 ^[4]	0.5	1.0

因目前还没有正确确定 η 值的理论依据^[1] ,所举算例的 η 值都是根据经验和在反复试验的基础上确定的.

4 结 束 语

本文提出的耦合神经网络模型是传统的前馈神经网络模型的一种推广 ,为解决非线性方程组之类的问题开辟了一种新的途径 .但它自身与传统神经网络模型的异同点及其收敛性、多值问题有待进一步研究.

参考文献 :

[1] 任伟波 ,于万明 ,云大真 .求解超定线性方程组及其相关问题的神经网络算法[J].大连理工大学学报 ,1996 ,36(4) :419 ~ 423.

[2] 董聪 .多层前向网络的逼近与泛化机制[J].控制与决策 ,1998 ,13(增) :413 ~ 417.

[3] 杨雄里 .脑科学的现代进展[M].上海 :上海科技教育出版社 ,1998.16 ~ 20.

[4] 王淑云 ,方保 ,王如云 .数值分析方法[M].南京 :河海大学出版社 ,1996.69 ~ 70.

Algorithm of Coupling Neural Network for Solving Non-linear Equations

ZHAO Qi-lin ,ZHUO Jia-shou

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

Abstract : A Weights-Coupling neural network is proposed , and the possibility of solving non-linear equations with it is discussed. A specific algorithm is given. Finally ,four numerical examples are provided which prove that the present method may yield the same outcome as that from the Newton method.

Key words : coupling neural network ; non-linear equation ; BP algorithm