

三层 BP网隐层节点数确定方法的研究

伍春香 刘琳 王葆元

(武汉测绘科技大学信息工程学院,武汉市珞喻路 129号,430079)

摘要 对确定三层 BP网络隐层节点数的理论依据和现有做法进行了研究,提出了一种三层 BP网络隐层节点数的双向确定法。

关键词 BP网络;隐层节点数;双向确定法

分类号 TP73 918

BP网络(back-propagation network)具有结构简单、工作状态稳定、易于硬件实现等优点。在众多的神经网络模型中,BP网络的应用最为广泛,尤其是在识别分类、非线性映射、复杂系统仿真等方面。然而,BP网络也存在着不足之处,它的训练速度慢,易陷入局部极小。要使其得到更好的应用,具有更高的功效,主要通过对训练算法的改进来提高训练速度。而用人工神经网络来解决实际问题,关键在于网络结构和训练算法的确定,其中网络结构的确定是首要的。如果网络的结构未选好,即使对训练算法作了很多改进,在网络特性的提高上也难有大的进展。对于三层BP网络来说,输入节点数和输出节点数是由实际问题本身决定的,确定网络结构实际上就是隐层节点个数的确定。本文提出的三层BP网络隐层节点数的双向确定方法也可以用于自适应神经网络的研究。

1 确定隐层节点数的理论依据

在大量BP网络的应用实例中,很少有人提及确定隐层节点个数的依据,有人认为隐层节点个数可以任意选择,实际上并非如此。隐层节点太少,网络的特性会降低;隐层节点太多,网络的权重矩阵太大,不便于推广。

依据Kolmogorov定理^[1],对于任意连续函数 H ,可以有一个三层网络来精确实现它,网络的第一层有 m 个单元,中间层有 $2m+1$ 个单元,第三层有 n 个单元。在这一定理中,提到了隐层节点的数目。Kolmogorov定理保证了三层网络精确地实现任意连续函数 H ^[3]。

关于三层BP网络隐层节点数的确定,文献

[4]从隐层节点与记忆样本的关系出发,证明了对于任意给定 k 个实数值样本, $k-1$ 个隐层单元的三层网络能够记忆它们。网络隐层节点的激发函数可以是任何渐近函数。采用阶跃激发函数时, $k+1$ 个随机给定的实数值样本能够被 $k-1$ 个隐层节点的网络记忆的概率为零^[4]。

2 隐层节点数现有确定方法

2.1 由Kolmogorov定理确定隐层节点数

直接依据Kolmogorov定理,对于 m 个输入的问题,采用 $2m+1$ 个隐层节点的BP网络来实现。

由于关于隐层节点数的理论尚在研究之中,而确定网络结构(隐层节点数)的目的就是为了找到一个经过样本训练后,网络输出与实际结果最接近、结构最简单(隐层节点数最少)的网络,这原则上是一个“搜索”过程^[9],实际上则是需要对一系列网络进行比较的过程。所以,在具体应用中,除了参考Kolmogorov定理确定隐层节点数外,大多数BP网络中隐层节点数的确定仍采用试凑的方式。

在理论上,要求求解某个问题BP网络的规模与实例规模相匹配。为了记住较多的样本,必须有相当数量的隐层节点。在试凑方面,也应找出一些规律来,以缩短神经网络的研制周期。

2.2 隐层节点单向渐变确定法

2.2.1 隐层节点渐增法

这种方法的思路是:先选取较少数目的隐层节点组成网络进行训练;然后根据训练达到的精度,适当增加一定数目的隐层节点,继续训练;重复这个工作,直至达到精度要求^[8]。

2.2.2 隐层节点渐减法

与上述方法正好相反,先选定较多隐层节点的网络进行训练,然后逐步对网络中那些对输入影响不大,而又干扰降低模型精度的节点进行删除^[7]。

3 三层 BP网络隐层节点数的双向确定法

在隐层节点单向渐变确定法中,存在如何选取初始隐层节点数问题。另外,假设实际应取隐层节点数为 p ,而初始选取的隐层节点数为 p_0 ,按上述渐变法来确定网络结构,不能排除这样的可能: $p < p_0$ 或 $p > p_0$ 。在这两种情况下,分别按§2.2中的两种做法,网络的结构会离目标越来越远。三层BP网络隐层节点数的双向确定法,可改进上述不足。

选取一个初始隐层节点数为 p_0 的网络进行训练。 p_0 的确定可参照Kolmogorov定理取 $2m+1$ (m 为输入节点数),或是其他有一定依据的估计值,如采用聚类分析确定样本总类别数作为隐层节点数^[10],采用实值样本数减1作为隐层节点数^[4]等,但以略多于可能的最终值为好(为了记住较多样本,必须有相当数量的隐层节点,以免隐层节点数太少,使输入的信息过份压缩,很多输入信息不能通过隐含层,导致网络特性下降)。

在初始网络训练的结果上,分别增加或减少(删去一些灵敏度小的节点)一定数目的隐层节点,重新进行训练,根据训练结果误差的增大或减小趋势,确定下一步隐层节点数变化的方向是继续增多还是减小,然后对改变了隐层节点数的网络重新进行训练。

按照每一批训练结果后(每批训练需对所有样本进行训练)误差变化的趋势重复“改变隐层节点数,重新训练”这个过程,即若上一批训练是对增加(或减少)了 k 个隐层节点的网络进行的,而训练结果是误差增大了,则将网络隐层节点减少(或增加) k' 个, $k' < k$,然后重新训练。

上述过程继续到到达一定数目的隐层节点后,网络的特性不再明显增加时为止。

具体步骤如下:

1) 选取初始的隐层节点数 p_0 (p_0 略多于隐层节点数的估计值)。

2) 采用选取的样本集对此网络进行训练,得到输出误差 Δy_0 。

3) 减少 k 个隐层节点,对 p 个隐层节点的网

络重新训练, $p = p_0 - k$,得到新的输出误差 Δy 。

4) 重复下述过程,直至网络的特性不再明显增加:

(1) 比较 Δy 与 Δy_0

a. 若 $\Delta y > \Delta y_0$,则:若 Δy 是减少 k 个隐层节点的网络的输出误差,则增加 k' ($k' < k$)个隐层节点;若 Δy 是增加 k 个隐层节点后的网络输出误差,则减少 k' ($k' < k$)个隐层节点。

b. 若 $\Delta y \leq \Delta y_0$,则:若 Δy 是减少 k 个隐层节点后的网络输出误差,则再继续减少 k' ($k' < k$)个隐层节点;

若 Δy 是增加 k 个隐层节点后的网络输出误差,则再继续增加 k' ($k' < k$)个隐层节点。

(2) 将原 Δy 作为新的 Δy_0 ,将 k' 作为新的 k 。

(3) 对改变了隐层节点数的网络重新训练,得到新的输出误差作为新的 Δy 。

这个方法,实际上是通过试验,从大于和小于最终节点数 p 这两个方向,依据误差变化的趋势,以减少或增加隐层节点数的方式不断向 p 靠近,以求得满意的 p 值。

笔者设计了一个试验系统,系统采用了三层BP网络来处理测量所得数据,网络输入节点为6个。表1是实验中关于 k 和 k' 取值及实验结果的部分数据。

按照Kolmogorov定理,隐层节点数为 $2m+1$ (m 为输入节点数),即13个。但按双向确定法试验得出的结论是:隐层节点为2时,网络输出误差最小;隐层节点为13时,误差是0.0346355,而隐层节点是2时,误差是0.0331830。

设计这个方法的最初目的主要是为了避免前述 $p > p_0$ 或 $p < p_0$ 时仍不断增加或减少隐层节点。试验结果表明,使用这一方法,还有助于找到与实际结果最接近的网络,缩短了寻找合适的隐层节点数的时间。

参 考 文 献

- 1 胡守仁,余少波,戴葵.神经网络导论.长沙:国防科技大学出版社,1993
- 2 沈清,胡德文,时春.神经网络应用技术.长沙:国防科技大学出版社,1993
- 3 焦李成.神经网络系统理论.西安:西安电子科技大学出版社,1990
- 4 应行仁.三层神经网络隐单元与样本记忆的关系.模式识别与人工智能.1990,3(1):29~34
- 5 鲍立威,何敏,沈平.关于BP模型的缺陷的讨论.模式识别与人工智能.1995,8(1):1~5

表 1 p_0 值不同时 k, k' 及最终误差 Δy
Tab. 1 k, k' and the Final Error Δy with Different p_0

	k	k'				
		$[k/2]$	$[k/3]$	$[k/4]$	$[k/5]$	$[k/6]$
$p_0 = 2m + 3$	$[p_0/2]$	0.034 064 5	0.034 064 5	0.033 895 6		
	$[p_0/3]$	0.033 895 0	0.033 895 6	0.033 895 6		
	$[p_0/4]$	0.034 635 5	0.034 635 5	0.034 197 9		
	$[p_0/6]$	0.034 272 4	0.034 635 5	0.034 635 5		
	$[p_0/8]$	0.034 272 4	0.034 272 4	0.034 272 4		
$p_0 = 2m + 5$	$[p_0/2]$	0.033 450 4	0.033 183 0	0.033 183 0	0.033 627 1	
	$[p_0/3]$	0.033 455 7	0.033 895 6	0.033 895 6	0.033 895 6	
	$[p_0/4]$	0.034 064 5	0.034 197 9	0.034 197 9	0.034 635 5	
	$[p_0/5]$	0.034 635 5	0.034 635 5	0.034 272 4	0.034 272 4	
	$[p_0/6]$	0.034 272 4	0.033 553 1	0.033 553 1	0.033 553 1	
	$[p_0/9]$	0.035 740 1	0.035 740 1	0.035 740 1	0.035 740 1	
$p_0 = 2m + 7$	$[p_0/2]$	0.033 299 5	0.032 503 3	0.033 627 1	0.033 4557	
	$[p_0/3]$	0.033 553 1	0.033 553 1	0.034 272 4	0.034 272 4	
	$[p_0/4]$	0.034 272 4	0.034 272 4	0.034 272 4	0.033 553 1	
	$[p_0/5]$	0.034 377 8	0.035 377 8	0.035 740 1	0.035 740 1	
	$[p_0/6]$	0.035 377 8	0.035 377 8	0.035 740 1	0.035 740 1	
	$[p_0/10]$	0.037 664 6	0.037 664 6	0.037 664 6	0.037 664 6	
$p_0 = 2m + 9$	$[p_0/2]$	0.033 299 5	0.033 183 0	0.033 455 7	0.033 455 7	0.034 064 5
	$[p_0/3]$	0.034 064 5	0.034 197 9	0.034 633 5	0.034 633 5	0.034 633 5
	$[p_0/4]$	0.034 633 5	0.033 553 1	0.033 553 1	0.033 553 1	0.035 740 1
	$[p_0/5]$	0.034 272 4	0.035 740 1	0.035 377 8	0.035 377 8	0.035 377 8
	$[p_0/6]$	0.034 176 2	0.034 176 2	0.037 664 6	0.037 664 6	0.037 664 6
	$[p_0/8]$	0.037 664 6	0.034 176 2	0.034 176 2	0.034 176 2	0.034 176 2
	$[p_0/11]$	0.036 970 2	0.036 970 2	0.036 970 2	0.036 970 2	—

6 雷 鸣,朱心飏,尹申明,等.自构形神经网络及其应用.计算机科学.1994,21(1): 52~ 54

7 余新新,黎志成,田俊彦.前馈式神经网络模型变量选择问题的研究.见:神经网络理论及应用—'94最新进展.武汉:华中理工大学出版社,35~ 38

8 高 峰,李人厚.一种有效的隐节点渐增式人工神经网络的學習算法.见:神经网络理论及应用—'94最新进展.武汉:华中理工大学出版社,88~ 91

9 阎平凡.现在人工神经网络模型的局限性与结构自适应问题.见:神经网络理论及应用—'94最新进展.武汉:华中理工大学出版社,703~ 707

10 洪 沁,何振亚.手写体数字的神经网络识别方法.模式识别与人工智能,1994,7(1): 68

11 洛伦茨 G G著.函数逼近论.谢庭藩,施威亮译.上海:上海科学技术出版社,1981

The Study of the Method to Determining the Number of Hidden Units of Three-layer BP Neural Networks

Wu Chunxiang Liu Lin Wang Baoyuan

(School of Information Engineering, W TU SM, 129 Luoyu Road, Wuhan, China, 430079)

Abstract This paper studies the theoretical basis and the current methods of determining the number of the hidden units of three-layer back-propagation neural networks, and puts forward a two-way method to solve this problem.

Key words BP neural network; the hidden units; two-way method