基于遗传算法的 RBF 神经网络学习算法

上海交通大学自动化系　　凌云 谢剑英 陈应麟

摘要：本文提出了用遗传算法选择 RBF 神经网络隐含层神经元的中心值的方法和 RBF 神经网络离线学习算法。

关键词：RBF 神经网络；遗传算法；Givens 最小二乘法

Learning Algorithm of RBF Neural Networks Based on Genetic Algorithm

Ling Yun，Xie Jianying and Chen Yinglin

Abstract: In this paper, a method of selecting the centers of hidden layer neurons of RBF neural networks is proposed. A new off-line learning algorithm of RBF neural networks is proposed.

Key word：RBF neural networks; Genetic algorithm; Givens least squares algorithm

1 引言

径向基函数理论为前向神经网络的学习提供了一种新颖而有效的手段。由于径向基函数神经网络训练的快速性、泛化能力好和结构简单，RBF 神经网络正吸引着越来越多的关注。在使用 RBF 神经网络之前，需要确定 RBF 神经网络的结构和参数。首先需要确定的是 RBF 神经网络的结构：

1)RBF 神经网络隐层神经元的激励函数的类型

根据理论的证明，RBF 神经网络隐层神经元的激励函数的类型对神经网络性能的影响并不很大，所以几种常用的径向基函数都可以被选作隐层激励函数。

2)RBF 神经网络的隐层节点数

RBF 神经网络的隐层节点数一般是根据经验来确定。对于一些特殊的方法，如正交最小二乘法，隐层节点数的初始值可设定为数据点的数目，然后在学习过程中根据设定的参数确定适当的隐层节点数。

在确定了 RBF 神经网络的结构后，就需要根据一定的神经网络学习算法来确定 RBF 神经网络的参数，需要确定的参数有：

1)高斯径向基函数的宽度 σ

根据经验，高斯径向基函数的宽度 σ 一般应是高斯径向基函数的中心点平均间距的两倍。当高斯径向基函数的中心点在数据集合中均匀分布时，宽度 σ 可按下式求得

$$\sigma = \frac{2D}{M} \qquad (1)$$

其中，D 是数据点之间的最大距离，M 是隐层神经元数目。

2)高斯径向基函数的中心值 c^m

高斯径向基函数的中心值应能描述神经网络输入空间的分布性质，即高斯径向基函数的中心值应在输入向量空间中均匀分布。中心值的选取对 RBF 神经网络的性能有极大的影响，只有在中心值选取适当的情况下，RBF 神经网络才能发挥良好的性能。

3)隐层与输出层之间的连接权值 w^m

RBF 神经网络的输出层的输出是隐层输出的加权值。当隐层神经元的中心值固定后，隐层与输出层之间的连接权值 w^m 的求解问题就成为一个线性问题，可以采用适当的学习算法求得。

理想的 RBF 神经网络的学习算法应选定一个误差函数，根据使误差函数最小的原则，同时确定合适的高斯径向基函数的宽度 σ，中心值 c^m 和隐层与输出层之间的连接权值 w^m。因为 RBF 神经网络的非线性性质，这样的学

习算法一般求解困难，所以在实践中常采用分步的方法，首先根据一定的算法确定合适的高斯径向基函数的宽度 σ 和中心值 c^m，然后在固定宽度 σ 和中心值 c^m 的前提下，求解隐层与输出层之间的连接权值 w^m。因为高斯径向基函数的宽度 σ 可根据式(1)来选定，而在固定宽度 σ 和中心值 c^m 的前提下，隐层与输出层之间的连接权值 w^m 的求解问题就成为一个线性问题，所以 RBF 神经网络学习算法的主要困难在于中心值 c^m 的选择方法。

2 常用中心值选择方法

常用的中心值 c^m 的选择方法有：

1)随机选择方法

最简单的中心值 c^m 的选择方法为随机选择方法，即从训练集中随机地选择 M 个输入向量作为中心值 c^m。如果训练集中的输入向量在神经网络的输入空间中均匀分布且隐层神经元的数目 M 较大时，这样选择出的中心值 c^m 可能会满足一定的要求，但是在大多数的情况下，这种方法都过于简单。

2)自组织选择方法

所谓自组织选择方法，就是根据训练集中的 N 个输入向量的分布特征，把训练集中的输入向量分成 M 个簇，然后在每个簇中选择一个簇中心，使得该簇中的所有向量到该簇中心的距离和最小。这种方法简单快速，但是在大多数情况下只能收敛到次优解，而且只利用了输入向量的信息，却没有利用输出向量的信息。

3)梯度方法

可以根据使一定的误差函数最小的准则，来求解最优的中心值。如果用梯度方法来求解中心值，很可能收敛到局部极小点，而无法得到最优的中心值。

4)正交最小二乘法

正交最小二乘法是最常用的也是最有效的 RBF 神经网络学习方法，它能够根据一定的误差要求自动地确定中心点的数目和相应的中心值。但是根据正交最小二乘法得到的中心值必定是训练集中的输入向量集的一个子集，所以在大多数情况下得到的中心值不是最优的中心值。

从以上常用的中心值 c^m 的选择方法来看，它们都存在一定的缺陷，至多只能得到中心值 c^m 的次优解。为了尽可能地发挥 RBF 神经网络的优越性，我们下面将提出一种基于遗传算法的 RBF 神经网络的中心值选择方法。

3 基于遗传算法的 RBF 神经网络的中心值选择方法

我们提出一种基于遗传算法的 RBF 神经网络的中心值的选择方法，因为遗传算法的收敛比较慢，我们只把它作为离线辨识算法，具体的步骤如下：

1)编码

我们用二进制无符号整数形式来表示一个 RBF 神经网络的中心值选择 c。如果 $c=\left[c^1 \quad \ldots \quad c^M\right]$，其中 $c^m=\left[c_1^m \quad \ldots \quad c_d^m\right]\in R^d$，则表示一个中心值选择 c 的数串由 M 个子串组成，每个子串应选择一个适当的码长。

2)初始中心值的生成

在生成初始中心值之前，首先要得到训练集中的输入向量的分布空间。如果训练集中的输入向量 $x^n=\left[x_1^n \quad \ldots \quad x_d^n\right]\in R^d$，则输入向量的第 i 个分量在训练集中有一个取值空间 $D^i=\left[x_{\min}^i, x_{\max}^i\right]$。在 $D^i=\left[x_{\min}^i, x_{\max}^i\right]$ 中选取一个随机数作为第 m 个中心值的第 i 个分量 c_i，然后把它编成一个二进制串，d 个子串就组成了初始中心值 c^m，M 个初始中心值构成一个中心值选择 c。初始群体由若干个这样的数串组成。

3)适应度评估检测

遗传算法需要一个适应度函数来评估个体或解的优劣，这里我们选取误差函数作为适应度函数

$$E[h(x)]=\frac{1}{2}\sum_{n=1}^{N}\left[\sum_{m=1}^{M}w^{m}g(\|x-c^{m}\|)-t^{n}\right]^{2}=\frac{1}{2}\sum_{n=1}^{N}\left[\sum_{m=1}^{M}w^{m}e^{-\frac{\|x-c^{m}\|^{2}}{2\sigma^{2}}}-t^{n}\right]^{2} \qquad (2)$$

由上式可见，对一个确定的个体 c，只要宽度 σ 和权值 w^{m} 给定，就可以得到相应的评估函数值 $E[h(x)]$，评估函数值 $E[h(x)]$ 越小，相应的个体的适应度越大。在求取适应度函数的过程中，不仅用到了训练集中的输入向量集合，也用到了输出向量集合。

式(2)中的宽度 σ 可根据式(1)求得，权值 w^{m} 根据 Givens 最小二乘法求得。

4)遗传操作

根据群体中个体的适应度的值，对群体进行选择、交叉和变异操作，得到新的群体。判断新的群体是否满足一定的收敛条件，如果未满足收敛条件，则重复以上各步；如果满足收敛条件，迭代中止，满足收敛条件的中心值 c^{m} 和权值 w^{m} 则为所求解。

基于遗传算法的 RBF 神经网络的中心值选择方法的流程图如下：

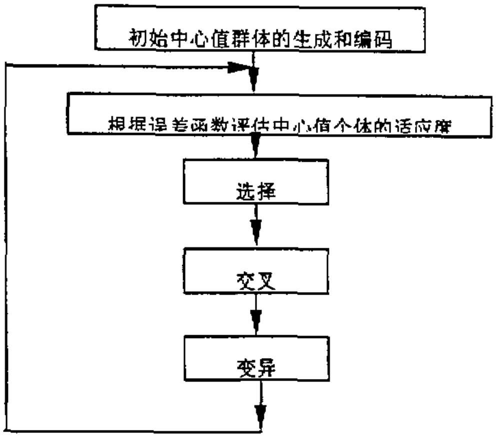

图 1 基于遗传算法的 RBF 神经网络的中心值选择方法

4 仿真实例

若非线性函数为

$$f(x)=\cos(2\pi x)+\sin(4\pi x)$$

其中：$x\in[0.0,1.0]$，在 $[0.0,1.0]$ 上均匀采样 41 个点作为训练集。

用高斯径向基函数神经网络对非线性函数 $f(x)$ 建模，隐层神经元有 6 个神经元，高斯径向基函数的宽度 σ 取为 0.34，用上述的基于遗传算法的 RBF 神经网络的学习算法进行离线训练，训练结果如图 2 所示，训练误差为 0.1725。图 2 中 "*" 代表 41 个数据点 x^{n}，虚线代表曲线 $f(x)$，实线代表训练得出的 RBF 神经网络

$h(x)=\sum_{m=1}^{6}w^{m}e^{-\frac{|x-c^{m}|^{2}}{2\sigma^{2}}}$，其中中心值为 $c=[0.180\quad 0.251\quad 0.404\quad 0.569\quad 0.737\quad 0.800]^{T}$，权值 $W=[-1163.935\quad 2369.456\quad -2458.621\quad 2258.159\quad -2193.457\quad 1188.267]^{T}$。

上述问题如果用最常用的 RBF 神经网络的学习算法一正交最小二乘法进行离线训练，训练结果如图 3 所示，训练误差为 0.3445。图 3 中 "*" 代表 41 个数据点 x^{n}，虚线代表曲线 $f(x)$，实线代表训练得出的 RBF 神经网络

$h(x)=\sum_{m=1}^{6}w^{m}e^{-\frac{|x-c^{m}|^{2}}{2\sigma^{2}}}$，其中中心值为 $c=[0\quad 0.025\quad 0.825\quad 0.05\quad 0.075\quad 1]^{T}$，权值 $W=[-2028.2\quad 5265\quad 0.472\quad -4497.4\quad 1250.6\quad -5.596]^{T}$。

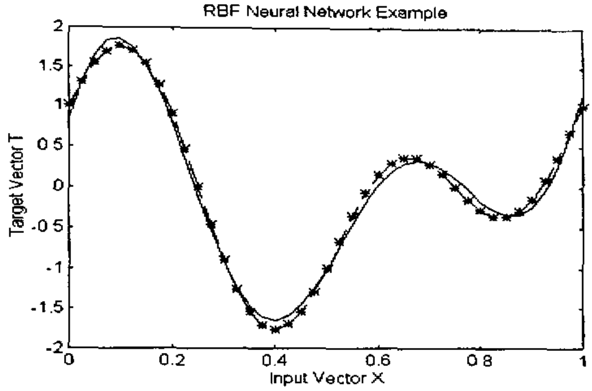

图 2 用基于遗传算法的 RBF 神经网络学习算法训练的 RBF 神经网络

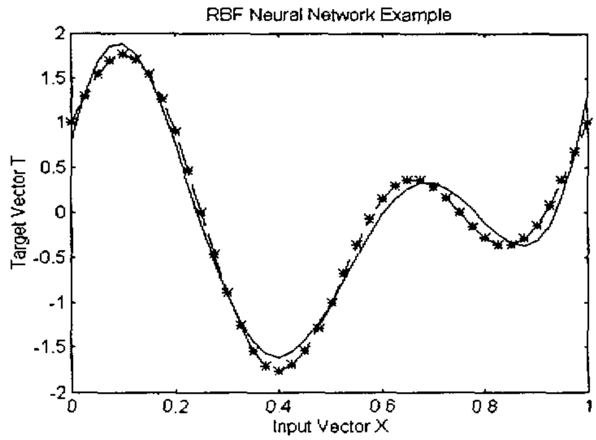

图 3 用正交最小二乘法训练的 RBF 神经网络

由上面的例子可见，基于遗传算法的 RBF 神经网络学习算法的训练误差为 0.1725,要比正交最小二乘法的训练误差 0.3445 的效果好。

5 结论

本文首先介绍了 RBF 神经网络学习算法，分析了常用 RBF 神经网络隐含层神经元的中心值选择方法的不足之处，提出了用遗传算法选择 RBF 神经网络隐含层神经元的中心值的方法及其对应的 RBF 离线学习方法，并与最常用的 RBF 神经网络学习算法一正交最小二乘法进行了比较，仿真结果证明了基于遗传算法的 RBF 神经网络学习能得到更好的结果。

参考文献

[1]W. Morven Gentleman, "Least Squares Computations by Givens Transformations without Square Roots", Journal of the Institute of Mathmatics and Its Applications, Vol.12, No.3, pp.329-336, 1973.

[2]S. Chen, P. M. Grant, C. F. N. Cowan, "Orthogonal Least-Squares Algorithm for Training Multioutput Radial Basis Function Networks", IEEE Proceedings-F: Radar and Signal Processing, Vol.139, No.6, pp.378-384, 1992.

[3]Luc Boullart, Stefan Sette, "Genetic Algorithms: Theory & Applications", Journal A, Vol.38, No.2, pp.13-23, 1997.