一种基于 Gaussian 函数的双向选择径向基函数神经网络算法

黄宏涛

（广东财经职业学院信息管理系　广州 510420）

摘　要　径向基函数神经网络是一类重要的神经网络算法。本文对现有的径向基函数神经网络算法进行了总结分析，将现有算法分为前向选择和后向选择两类。在分析各自优缺点的基础上从提高神经网络泛化能力的角度提出了一种新的基于 Gaussian 函数的双向选择径向基函数神经网络算法——BSRBF，从数理角度研究了神经元选择的基本技术方法，并对算法的基本思想和具体步骤进行了阐述。最后，用一个实验对比验证了双向选择算法的有效性。

关键词　神经网络，径向基函数，Gaussian 函数，双向选择，算法

A New Bi-directional Selection RBF NN Algorithm Based on Gaussian Function

HUANG Hong-Tao

(Department of Information Management, Guangdong Vocational College of Finace and Economy, Guangzhou 510420)

Abstract　RBF NN is an important Neutral Network algorithm. This paper summarizes the existing RBF NN algorithms and classifies it into forward-direction selection algorithm and back-direction selection algorithm. A new bi-directional selection RBF NN Algorithm based on Gaussian Function (BSRBF) is proposed from the view of enhancing the capacity of generalization of Neutral Network. On the base of analyzing their advantages and disadvantages, this paper researches the elementary technologies and methods of the selection of nerve cell and set forth basis thought and steps of the algorithm. In the end, the validity of BSRBF algorithm is validated through an experiment.

Keywords　NN, RBF, Gaussian function, Bi-directional selection, Algorithm

神经网络是20世纪中期提出来的一种有效的科学计算方法，它借鉴人类大脑的组织结构和运行机制，进行神经元与神经元间的复杂运算，产生神经元与神经元之间的信号加权值，使神经网络具有人工智能性。它从模拟人脑智能的角度出发，探寻新的信息表示、存储和处理方式，设计全新的计算机处理结构模型，构造了一种更接近人类智能的信息处理系统来解决实际工程和科学研究领域中传统的冯·诺依曼计算机难以解决的问题。经过几十年的发展，神经网络技术目前已经得到了较为广泛的应用，各种以神经网络为核心的软件理论和应用系统层出不穷，为人类的科技发展提供了强有力的支撑。

1　基于 Gaussian 函数的径向基函数简述

在各种神经网络方法中，径向基函数是一种简单而应用广泛的方法。径向基函数神经网络由于其结构简单、算法简便，被广泛地用于函数逼近、系统识别、时间序列预测、语音识别、自动控制、数据挖掘等许多领域。它不仅解决了神经网络非线性的问题，而且能将训练样本的信息系统储地存在隐藏层神经元中，且可使用简单的矩阵运算来计算网络输出加权值，不需要在训练之前设定大批参数的值，只要适当地决定训练停止条件即可。径向基函数方法最初是被用于严格多变量插值问题中的一种方法。对一个严格多变量插值问题，要求插值空间（函数）穿过所有训练样本数据点，用数学语言可以表示如下：

给定一个 n 个不同点的集合 $\{x_k\in R^q\,|\,k=1,2,\cdots,n\}$ 和一个对应的 n 个实数组成的集合 $\{y_k\in R^1\,|\,k=1,2,\cdots,n\}$，寻找一个函数 $F:R^q\to R^1$ 满足以下条件：$F(x_k)=y_k, k=1,2,\cdots,n$。径向基函数就是选取以下形式的函数来构成 F：

$$F(x)=\sum_{k=1}^{n}a_k\phi(\|x-x_k\|)$$

在这里，$\{\phi(\|x-x_k\|)\,|\,k=1,2,\cdots,n\}$ 是一个由 n 个随机函数（通常是非线性的）构成的集合，被称为径向基函数（径向基函数）；$\|\ \|$ 表示范数，通常取为欧氏范数；已知点 $x_k\in R^q, k=1,2,\cdots,n$ 作为径向基函数的中心；$a_k, k=1,2,\cdots,n$ 是一个线性权值集合。径向基函数是不规则分布数据点插值逼近的有效工具。显然，基函数具有以下特点：

$$\phi(-r)=\phi(r)$$

$$\lim_{r\to\pm\infty}\phi(r)=0$$

选择不同的 $F(x)$ 形式就够成了不同的径向基函数神经网络。Gaussian 函数是一种最简单易用函数，其一般形式为：$\phi(r)=e^{-(\frac{r^2}{\beta^2})}$。此处，$r$ 表示数据点 x 到基函数中心 μ 的欧氏距离，即：$r=\|x-\mu\|$，β 是一个由用户定值的基函数参数。其中，$\phi_{ij}=\phi(\|x_i-x_j\|), i,j=1,2,\cdots,n$ 称为插值矩阵，由插值矩阵的逆矩阵和对应的输出值 y_k，$(k=1,2,\cdots,n)$ 可以计算得出线性权值。

2　径向基函数神经网络的泛化问题

与普通神经网络类似，径向基函数神经网络包含输入层、隐层、输出层。输入层神经元个数（m）与输入向量（$X\in R^m$）维数相同；输出层神经元数（n）与输出向量（$Y\in R^n$）维数相同。隐层的作用是对输入向量进行非线性变换。输出层神经元对隐层输出进行线性组合，得到网络输出为：

黄宏涛　讲师，硕士，主要研究方向为：计算机网络，数据库，数据挖掘。

$$\hat{y}_k(x^p)=\sum_{i=0}^{N_h} w_{ki}\phi_i(x^p) \quad k=1,\cdots,n$$

其中,$\hat{y}_k$ 为网络的实际输出;$w=\{w_{ki} \mid k=1,2,\cdots,n;i=0,1,\cdots,N_h\}$为隐层到输出层的权值;$N_h$ 为隐层神经元数;$\phi_i(x^p)$为所选径向基函数函数,在此本文取为 Guassian 函数:$\phi_i(x^p)=\exp\left[-\frac{(\| x_p-c_i \|)^2}{\sigma_i{}^2}\right]$,$c_i$ 为基函数中心,σ_i 为宽度;w_{k0} 实质为输出层神经元的偏置,一般将其作为一特殊的权值与其它权值同时调整。$(x^p,y^p)\in R^m\times R^n$,$p=1,2,\cdots,N$ 构成了一个训练样本集。

从径向基函数网络的结构可以看出,构造和训练一个径向基函数网络就是要使它通过学习,确定出隐层神经元的个数 N_h、每个隐层神经元基函数的中心 c_i、宽度 σ_i 以及隐层到输出层的权值 w 这些参数的过程,从而可以完成所需的输入到输出的映射。在 N_h 确定的情况下,训练一个径向基函数神经网络就是要确定三组参数:基函数的中心 c_i、宽度 σ_i 以及隐层到输出层的权值 w,达到最小化某一个函数的要求。在插值问题中,基函数的数目是与样本数据点的数目完全一致的,因为只有这样才能保证插值函数严格穿越每一个样本点。而在神经网络的训练过程中,训练样本数目常常是远大于所学习问题的自由维度的,并且训练样本中基本都包含噪音。如果还对每一个样本点都引入一个基函数,就将导致网络由于过度庞大而瘫痪。事实上,网络训练的根本目的不是寻求对训练样本的完全拟和,而是为了对产生样本的潜在统计规律进行建模。因此,学习算法的重点在于网络的泛化(推广)能力,即它对训练集以外数据的正确反映能力。泛化能力有时也称为推广能力,是指经训练后的神经网络对未在训练样本集中出现(但具有同一规律性)的样本作出正确反映的能力。

基于以上的分析,有学者为去除严格插值问题中的基函数数目限制,建立起了一种含有一个隐层的三层结构的径向基函数神经网络,并对严格插值中的径向基函数方法做了一些改动,如控制使径向基函数神经网络使用的基函数数目 N_h 远小于训练样本数 N;将选择基函数中心作为径向基函数神经网络训练过程的一个重要步骤,而不是限定在输入样本点上;训练时调整每个基函数的宽度,而不是对所有基函数预先设定一个固定的宽度。经过这些改进,径向基函数神经网络的输入→输出映射模型就相应地变为:

$$F(x)=\sum_{k=1}^{n} w_k\phi(\| x-x_k \|)+w_0$$

一旦基函数的中心、宽度参数确定下来,输入层到隐层的变换也就确定了。这时径向基函数神经网络可以看作是一个输入和输出是线性关系的简单的两层结构神经网络。对于停止标准问题,可以使用经典的最小二乘法(LS)求解隐层一输出层连接权 w:$w^T=\Phi^* T$。其中,$(T)_{pk}=t_k^p$,$(\Phi)_{pj}=\phi_j(x^p)$,$\Phi^*=(\Phi^T\Phi)^{-1}\Phi^T$,表示 Φ 的伪拟矩阵。

3 基于 Gaussian 函数的双向选择径向基函数神经网络算法——BSRBF

基于径向基函数的局部响应特性和径向基函数神经网络分阶段的训练方法,目前出现了两类以 Gaussian 函数为核心的径向基函数神经网络算法,分别是前向选择算法和后向选择算法。

前向选择算法的基本思想是:假设给定一个具有初始结构的径向基函数神经网络和一个候选基函数(隐单元)集,一般来说就是以训练数据点为中心的 Gaussian 函数集。每一步,找到一个对误差函数(如误差平方和)下降贡献最大的隐单元(基函数),将其加入到网络隐层中,并从候选集中删除它。这一加入隐单元的过程一直进行到某个标准,不再下降为止。虽然前向选择不是一个非线性最优化技术,但它有不必预先设定隐层单元数的优点,并且有一个可操作的选择标准。此外,这种算法的计算效率也是比较高的。与前向选择算法不同,后向选择算法的初始网络隐层包含全部候选基函数。其基本思想是:每一步,删除一个带来最少误差下降的隐节点。重复执行这一过程直到预定的模型选择指标函数值不再下降。此时,认为网络的隐层节点数目已足够表达目标问题的复杂程度。这两类算法在提高神经网络泛化能力方面的一个共同点是,随着训练过程对网络动态进行增加或删减隐层节点。对于径向基函数神经网络来说,隐层节点的数目是影响网络泛化能力的最关键因素。选择数目不同的隐层节点数,就是选择不同的网络模型。隐层节点数目过少或过多,网络都不能对目标问题准确建模。在单纯的前向选择或后向选择算法中,为了追求停止标准函数——通常是误差平方和——的下降,不断加入新的神经元。往往会出现当神经元继续增加下去,虽然训练误差逐渐下降,但验证误差却逐渐上升的情况。出现这种情况的原因是径向基函数神经元过多,造成迁就噪音现象,形成所谓过拟和,此时径向基函数虽有较低的训练误差,但泛化能力却不佳。解决这个问题的一个最直接的办法是分别训练含有不同隐层结点数的网络,然后通过反馈检验选择最优的隐层结点数,确定最优的网络结构。

基于这种思想,本文提出了一个基于 Gaussian 函数的双向选择径向基函数神经网络算法(Bi-directional Selection RBF NN Algorithm Based on Gaussian Function)——BSRBF,目的在于吸收这两种算法的各自优点,在增加或删除隐层节点的问题上做到"适度",以此来提高径向基函数神经网络的泛化能力。此算法的基本思想是,径向基函数神经网络隐层神经元数目及其中心点与宽度值被选定后,可将径向基函数神经网络表示为如下的线性回归模式:

$$y_i(t)=\sum_{j=1}^{N_h} w_{ji}\phi_j(t)+e_i(t),i=1,2,\cdots,n$$

其中,t 表示训练样本 $x(t)$ 的计数指标,y_i 和 e_i 分别为期望输出和训练误差。如果将所有训练样本代入上式,则可以得到一个线性联立方程:

$$Y=\Phi W+E$$

输出权值可以通过求解误差平方和函数的最小化问题求出:

$$M=\frac{1}{2}\sum_{k=1}^{N}\sum_{i=1}^{n}(y_{ki}-\hat{y}_{ki})^2$$

其中,N 是训练样本总数,n 是输出层神经元数,y_{ki} 和 $\hat{y}_{ki}$ 分别表示第 k 个训练样本的第 i 个期望输出值和网络实际输出值。

在运行过程中,这种训练算法逐步增加网络隐藏层神经元数目,每增加一个隐藏层神经元,尝试将每个训练样本的输入部份当作中心点,逐一比较每个训练样本在成为中心点后,得到误差下降比,然后选择具有最大误差下降比的训练样本成为新的径向基函数中心点,利用线性最小平方法求解网络输出层与隐藏层间输出加权值,进而逐步扩展网络的结构。算法的具体步骤如下:

①设隐层中心点计数指标 $j=1$。

②令 $k=1$,则 Gaussian 函数的宽度值 $\delta=$

$\sqrt{d_{max}^2/(j+1)}$，其中 d_{max} 表示训练样本中输入向量间的最大距离。

③设 $c_j = x_k$，也就是将第 k 个训练样本的输入向量当作第 j 个径向基函数隐层神经元的中心。

④将 M 函数方程式中的 E 值的问题转换成线性最小平方问题，然后计算此时所造成的误差下降比。

⑤令 $k = k+1$，返回步骤 3，直至所有尚未成为中心点的训练样本都经过步骤 3 的测试。

⑥选择能获得最大误差下降比的 x_k 当作径向基函数新隐藏层神经元的中心点，并计算网络输出加权值及训练误差。

⑦令 $j = j+1$，并回到步骤 2，直到获得满意的训练误差或到达额定的隐藏层神经元数目为止。

该算法的优点是：(1)以每个资料点为中心点的候选人，故能系统地选取训练样本中具有代表性的资料点来当作隐藏层神经元的中心点；(2)能依据所设定的容忍误差，来决定径向基函数隐藏层神经元的个数。缺点是算法太过复杂，当样本点较多时，速度可能会比较慢。

4 实验对比

为验证 BSRBF 算法的有效性，本文采用 UCI 机器学习数据库中的 Breast Cancer(乳腺癌)数据集作为实验对比平台。这个数据集是一个典型的分类问题数据集，包括 699 个样本数据，其中有 16 个样本各含有一个缺失的属性值。数据集包含 11 个属性列，其中一个为样本序号列，一个为诊断结果列(决策属性)，九个病例属性列(条件属性)，Breast Cancer 分类问题就是根据其余 8 个条件属性判断一个病例是良性的还是恶性的。为了产生可信的对比效果，验证 BSRBF 算法的有效性，本文以一种典型的前向选择算法为对比基准，即文[2] Karayiannis 所提出的实验结果作为对照。

实验之前，本文对实验数据集进行了数据预处理。对于 Breast Cancer 数据集中存在 16 处属性值缺失的样本数据点，本文采用以该属性的所有现存属性值的均值填充的办法补足。此外，为了使网络结构更紧凑和计算简便，本文将该数据集中的决策属性用一个二进制数据表示输出，0 表示良性；1 表示恶性。

实验过程中，将数据集分为训练集和测试集两部分，分别采用两种算法在训练集上对神经网络进行训练，然后在测试集中进行测试。神经网络的泛化能力我们可以用分类误差率来表示，因为分类误差率就是不同算法在同一个测试集中的运行效果，可以间接反映出该神经网络算法的泛化能力。表 1 列出了 BSRBF 算法与 Karayiannis 算法解决 Breast Cancer 分类问题的结果。可以看出，BSRBF 算法取得了与 Karayiannis 算法几乎相同的均方误差，但是其分类误差率要小得多，说明双向选择神经网络算法具有较高的泛化能力。

表 1　BSRBF 算法与 Karayiannis 算法求解乳腺癌决 Breast Cancer 问题的实验结果对比

实验数据集	神经网络算法	均方误差	分类误差率
训练集	Karayiannis	1.32	1.38
	双向选择	1.57	1.68
测试集	Karayiannis	3.47	4.77
	双向选择	3.12	4.22

以上实验表明，双向选择神经网络算法与前向选择神经网络相比，在解决数据挖掘问题中的分类问题具有较大的优越性。解决同样的问题双向选择算法具有较高的网络泛化能力，对新数据集有更强的适应能力。当然，此次实验中也存在不足之处。由于训练集较小，双向选择径向基函数神经网络算法在大训练集中的应用效果还不能完全确定；本实验中未出现相似隐结点合并的情况，使双向选择神经网络算法的一个重要的优越性未能体现出来。此外，BSRBF 算法与后向选择神经网络算法的对比效果也有待进一步确定。

参考文献

1　Lu Y W，Sundererarajan N，Saratchandran P．A Sequential Learning Scheme for Function Approximation Using Minimal Radial Basis Function Networks. Neural Computation, 1997, 9: 461～478

2　Karayiannis N B. Reformulated Radial Basis Neural Networks Trained by Gradient Descent. IEEE Transactions on Neural Networks, 1999, 10(3): 657～671

3　杨根兴，高大启．改进的 RBF 神经网络模式分类方法研究．华东理工大学学报，2001，27(6)：24～28

4　邓继雄，李志舜，梁红．确定 RBF 神经网络参数的新方法，微处理机，2006(4)：18～22

5　阎平凡，张长水．人工神经网络与模拟进化计算．北京：清华大学出版社，2000

6　苏小红，侯秋香，马培军，等．RBF 神经网络的混合学习算法．哈尔滨工业大学学报，2006(9)：34～37

(上接第 183 页)

do Step2 到 Step5

else end-if

结论与展望　通过对战术计划识别的领域特性分析，本文给出了多 Agent 战术意图识别问题的一般描述，并通过对战术意图识别问题的详细展开，实现了计划库的知识组织与逻辑描述，从而确定了战术计划识别的逻辑推理框架。但值得注意的是，该框架反映的仅仅是识别推理过程中各要素之间的逻辑联系，并不意味着战术计划识别就是一个纯逻辑的保真推理过程。因为，从战场感知获得的信息源中包含的客观不确定性，到战术目标与各任务之间主观分解所包含的主观不确定性，都决定了不确定性是多 Agent 意图识别的根本特性，随着这种特性与逻辑推理框架的连接，以及与时间要素的结合，未来的多 Agent 战术意图识别问题必将是一个基于 PL 的动态不确定性传播与集成过程，而这也将成为下一步的研究方向与工作重点。

参考文献

1　Sridharan S C F. The plan recognition problem: An intersection of psychology and artifical intelligence [J]. AI-1978, 11: 45～83

2　Kautz H. A formal theory of plan recognition and its implementation[A]. In: Reasoning about Plans [C], Morgan Kaufman, San Mateo, CA, US, 1991. 69～125

3　Azarewicz J. Plan recognition for airborne tactical decision making [A]. In: Proceedings of the National Conferenceon Artificial Intelligence [C]. Menlo Park, Calif.: AAAI press, 1986. 805～811

4　Saria S. Probabilistic Plan Recognition in Multi-agent Systems [A]. ICAPS [C], 2004. 89～87

5　Wobcke W. Two logical theories of plan recognition [J]. Journal of Logic Computation, 2002. 371～412

6　Bui H, Venkatesh S, West G. Policy Recognition in the Abstract Hidden Markov Model [J]. Journal of Artificial Intelligence 2002, 17: 451～499