



XI'AN JIAOTONG UNIVERSITY

# 神经网络导论

---

## 径向基神经网络





XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络简介


1985年，Powell提出了多变量插值的径向基函数（Radial-Basis Function）方法。1988年，Broomhead和Lowe首先将RBF用于神经网络设计，提出了径向基函数神经网络，即RBF神经网络。

RBF神经网络的基本思想是：用径向基函数作为隐单元的“基”，构成隐含层空间，隐含层对输入矢量进行变换，将低维的输入模式变换到高维空间，使得在低维空间内的线性不可分问题在高维空间中变为线性可分问题。

RBF神经网络由于拓扑结构简单、训练简洁，能逼近任意非线性函数，已在信号处理、模式识别、非线性控制、时间序列预测以及图像处理等方面得到广泛应用。

RBF神经网络是前向三层神经网络（包括输入层），其中隐含层神经元采用径向基函数作为神经元功能函数。

2013-1-5
《数据通信与计算机网络》——数据链路层
2



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络简介

### 径向基函数


径向基函数的值只与自变量到某个固定点之间的欧氏距离有关。比如在N维空间中，设 $X_0$ 为空间中的某个固定点，则 $f(\|X - X_0\|)$ 为径向基函数。常见的径向基函数有：

高斯函数：
$$\phi(X, X_0) = \exp\left(-\frac{\|X - X_0\|^2}{2\sigma^2}\right)$$

多二次函数：
$$\phi(X, X_0) = (\|X - X_0\|^2 + c^2)^{1/2}$$

逆多二次函数：
$$\phi(X, X_0) = (\|X - X_0\|^2 + c^2)^{-1/2}$$

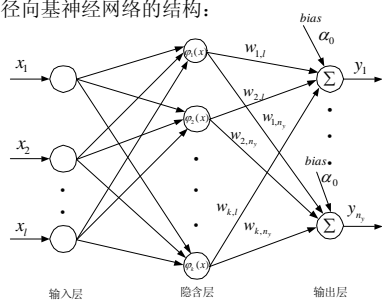
2013-1-5
《数据通信与计算机网络》——数据链路层
3



XI'AN JIAOTONG UNIVERSITY


## 径向基神经网络结构

如图为径向基神经网络的结构：



根据隐含层神经元的个数不同，RBF神经网络可以分为正规化网络（Regularization Network）和广义网络（Generalized Network）。

2013-1-5
《数据通信与计算机网络》——数据链路层
4



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法

RBF网络需要通过学习确定的参数包括：隐含层神经元的个数、中心和方差（宽度）以及输出层神经元的权值。


RBF网络的学习方法很多：有监督选取中心法、自组织选取中心法、基于资源分配的学习算法等。

基于资源分配的方法主要有：资源分配网络RAN（Resource Allocating Network）及其改进型算法RANEKF（RAN via Extended Kalman Filter）和最小资源分配网络MRAN（Minimal RAN）。GAP-RBF（Growing and Pruning RBF）和GGAP-RBF（Generalized GAP-RBF）。

极速学习机ELM（Extreme Learning Machine）。

下面简单介绍其中的几种算法。

2013-1-5 《数据通信与计算机网络》——数据链路层 5




XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法

### 自组织选取中心法

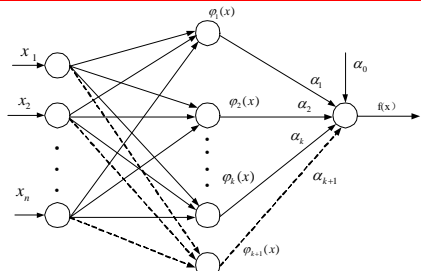
1. 自组织学习阶段。使用聚类算法确定输入样本的类别以及每个类别的样本中心。用学习得到的样本中心作为RBF网络隐含层基函数的中心。
2. 确定每个样本中心的方差。根据类别之间的距离确定每个中心的方差（宽度）。
3. 学习输出层的权值。可以使用LMS算法。

2013-1-5 《数据通信与计算机网络》——数据链路层 6



XI'AN JIAOTONG UNIVERSITY


## 径向基神经网络学习算法



$$\{(\mathbf{x}_n, y_n), n = 1, 2, \dots, N\}$$

$$f(\mathbf{x}) = \alpha_0 + \sum_{k=1}^K \alpha_k \phi_k(\mathbf{x}) \quad \phi_k(\mathbf{x}) = \exp\left(-\frac{1}{\sigma_k^2} \|\mathbf{x} - \boldsymbol{\mu}_k\|^2\right)$$

2013-1-5 《数据通信与计算机网络》——数据链路层 7



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法

### 资源分配网络RAN

Platt, 1991


传统方法，隐含层神经元的个数根据输入数据的性质先验确定，输出层权值用LMS算法估计。而RAN算法基于数据的新颖性添加隐含层神经元。该算法适合序贯学习。

1. 网络初始时无隐含层神经元。
2. 随着观察矢量的输入，网络使用其中的一些作为隐含层神经元。增加的准则为：

$$\min_{1 \leq k \leq K} \|\mathbf{x}_n - \boldsymbol{\mu}_k\| > \varepsilon_n \quad e_n = [y_n - f(\mathbf{x}_n)] > e_{\min}$$

3. 如果输入的观察矢量不满足增加条件，用LMS算法调整（更新）已有的权系数。

2013-1-5 《数据通信与计算机网络》——数据链路层 8



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法


### 使用扩展卡尔曼滤波的资源分配网络RANEKF

增加策略与RAN算法相同。区别只在于当输入的观察矢量不满足增加为隐含层神经元中心时，不是使用LMS算法，而是使用扩展的卡尔曼滤波方法更新已有的权系数。

RAN与RANEKF的缺点是：一旦添加了一个隐含层神经元，这个神经元就不会被删除。但会出现这样的情况，某些加入的神经元在随后的学习中对网络输出的贡献很小。如果我们能检测并删除这些神经元，就可以得到更简约的网络拓扑。

最小资源分配网络MRAN就是在RANEKF算法的基础上加入了删除策略。

2013-1-5
《数据通信与计算机网络》——数据链路层
9



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法

### 最小资源分配网络MRAN


MRAN在RANEKF的基础上增加了如下的删除策略。

如果对连续M个输入，某个隐含层神经元的归一化的输出值都小于某个门限，那么就删除这个神经元。具体步骤为：

1. 对每个输入矢量，计算所有隐含层神经元的输出。
2. 找出输出的最大绝对值，并用这个值对所有的输出归一化。
3. 如果某个隐含层神经元的连续M个输入的归一化输出小于某个门限值，删除此神经元。
4. 调整网络参数。

为了使隐含层神经元数目的变化更平滑，除上面的新颖性准则外，MRAN还增加了基于滑动数据窗上的输出误差的均方根准则。

2013-1-5
《数据通信与计算机网络》——数据链路层
10



XI'AN JIAOTONG UNIVERSITY

## 径向基神经网络学习算法

前面的算法，必须通过试凑的办法确定各种门限。

### GAP-RBF (Growing and Pruning RBF)

GAP-RBF引入神经元“重要度”（Significance）的概念，并把它与学习精度联系在一起。


神经元的“重要度”定义为神经元对目前为止所有输入数据的网络输出的平均贡献。

$$f_1(\mathbf{x}_i) = \sum_{k=1}^K \alpha_k \phi_k(\mathbf{x}_i) \quad f_2(\mathbf{x}_i) = \sum_{k=1}^{m-1} \alpha_k \phi_k(\mathbf{x}_i) + \sum_{k=m+1}^K \alpha_k \phi_k(\mathbf{x}_i)$$

$$E(m, i) = |f_1(\mathbf{x}_i) - f_2(\mathbf{x}_i)| = |\alpha_m| \phi_m(\mathbf{x}_i), i = 1, 2, \dots, N$$

$$E_{ave}(m) = \frac{\sum_{i=1}^N E(m, i)}{n} = \frac{|\alpha_m|}{n} \sum_{i=1}^N \phi_m(\mathbf{x}_i)$$

2013-1-5
《数据通信与计算机网络》——数据链路层
11



XI'AN JIAOTONG UNIVERSITY

$$y(x) = \exp\left[-\frac{(x_1-0.3)^2 + (x_2-0.2)^2}{0.01}\right] + \exp\left[-\frac{(x_1-0.7)^2 + (x_2-0.2)^2}{0.01}\right]$$

$$+ \exp\left[-\frac{(x_1-0.1)^2 + (x_2-0.5)^2}{0.02}\right] + \exp\left[-\frac{(x_1-0.9)^2 + (x_2-0.5)^2}{0.02}\right]$$

$$+ \exp\left[-\frac{(x_1-0.3)^2 + (x_2-0.8)^2}{0.01}\right] + \exp\left[-\frac{(x_1-0.7)^2 + (x_2-0.8)^2}{0.01}\right]$$

2013-1-5
《数据通信与计算机网络》——数据链路层
12