

人工神经网络预报模型的过拟合研究*

金 龙

(广西壮族自治区气象减灾研究所, 南宁, 530022)

况雪源

(广西壮族自治区气候中心, 南宁, 530022)

黄海洪

(广西壮族自治区气象减灾研究所, 南宁, 530022)

覃志年

(广西壮族自治区气候中心, 南宁, 530022)

王业宏

(南京气象学院, 南京, 210044)

摘 要

针对神经网络方法在预报建模中存在的“过拟合”(overfitting)现象和提高泛化性能(generalization capability)问题,提出了采用主成分分析构造神经网络低维学习矩阵的预报建模方法。研究表明,这种新的神经网络预报建模方法,通过浓缩预报信息,降维去噪,使得神经网络的预报建模不需要进行适宜隐节点数的最优网络结构试验,没有“过拟合”现象,并且与传统的神经网络预报建模方法及逐步回归预报模型相比泛化能力有显著提高。

关键词: 神经网络, 泛化性能, 过拟合现象, 预报建模。

1 引 言

20 世纪 90 年代以来,国内外在气象学科中开展了很多有关神经网络预报建模和气候分析等应用研究^[1~5]。然而随着神经网络方法在大气科学领域应用研究的不断深入,发现人工神经网络方法在实际业务天气预报应用中存在一个很重要的“瓶颈”问题,即在利用神经网络方法进行气象预报建模时,对已知训练样本集的学习训练达到什么样的拟合精度,才能使预报模型对未知样本具有最好的预报能力。该问题的研究不仅关系到在大气科学中能否进一步深入开展有关神经网络方法的业务预报应用,并且也是目前人工神经网络应用理论研究中尚未得到很好解决的有关神经网络模型的过拟合和泛化性能的很关键问题^[6,7]。本文针对实际气象预报问题,尝试从如何构造一种能够全面反映预报因子与预报量关系的低维学习矩阵的“主动”性方法,来探讨神经网络的泛化性能和过拟合现象。

2 影响泛化性能的因子分析

2.1 网络隐层节点

人工神经网络在众多学科的应用研究中已在理论上证明,一个 3 层的前馈网络能够实现任意精度的连续函数映射。并且人们在神经网络的学习算法方面进行了大量的研究,提出了很多关于前馈网络的改进学习算法,使得大多数有关神经网络学习精度的问题都能达到要求。但是在神经网络的预报建模研究中很关键的问题是关于神经网络预报模型的泛化能力,具体在天气预报问题中,就是神经网络天气预报模型根据训练样本建立预报模型后,对未参加学习训练样本(即独立样本)的预报能力,这显然是神经网络预报模型好坏及该方法能否应用于实际业务预报的核心问题。因此,如何提高神经网络预报模型的泛化能力成为神经网络理论研究和实际应用研究最为重要的研究课题。该问题研究的主要内容包括神经网络的网络结构(主要是适宜隐节点数的确定)和网络结构的学习算法对泛化性能和过拟合现象的影响研究^[8~11]。而一般关于网络结构的影响研究认为网络结构增大,容易产生过拟合现象,使网络泛化性能下降。一般只要在网络收敛的条件下,小的网络结构泛化能力强,且训练速度快。并且认为,隐节点数应小于样本数,但是在隐节点数小于

* 初稿时间:2003 年 2 月 13 日;修改稿时间:2003 年 4 月 25 日。

资助课题:国家自然科学基金项目(40075021)。

