

基于神经网络的S参数估计法¹

周力 毛钧杰 姚德森

(国防科技大学四院 长沙 410073)

摘 要 该文讨论了神经网络用于微波电路的一种新的训练算法,是基于误差反传算法与模拟退火算法相结合的。本算法可以解决BP网的局部最小问题。并且不用对神经网络模型的结构作任何改动。通过一个具体算例的计算验证了这种算法。

关键词 神经网络, 误差反传算法, BP网, 局部最优, 模拟退火算法

中图分类号 TN-052, TN011

1 前言

近几年来,随着神经网络在目标识别、模式识别、数据融合等方面获得的巨大成功,在国外,神经网络也被应用于微波领域。就目前看,主要用于微波电路的设计及优化上。

同传统的微波电路CAD相比,神经网络是通过对数据的训练学习来模拟输入输出之间的关系,无需求解复杂的方程,实际上是一种泛函逼近的方法。但是,同一般的逼近方法相比,神经网络显示出极大的优越性。首先,神经网络良好的非线性性能,处理多维问题的方法很简单,就算法来看,几乎同一维问题一样。其次,神经网络有很强的“反传”能力,对给定模型,在训练结束后,保持模型结构不变,仅改变输入与输出之间的方向,再对它进行训练,很快便可出结果。这样在微波电路的设计中不仅可以由电路的物理参数得到其性能参数,亦可由其性能参数得到其物理参数,使得微波电路的设计更加方便灵活。目前,国外已经在微波电路设计中成功地对神经网络的方法进行了一些成功的应用。例如对微波阻抗的匹配^[1], 共面波导电路的设计^[2], 及其它^[3-5]等等。

由于BP(Backward Propagation)网的原理及算法比较简单,易于掌握。所以,目前国外微波电路设计中大多用BP网。但是,BP网仍有一些问题,尤其是BP网的算法是基于梯度下降法的,存在着局部最小的问题,且该问题还是不可预测及控制的,往往训练结束后仍不知是否达到全局最优。而模拟退火算法是可以解决局部最小问题的,但其训练速度又太慢。因此,本文给出了一种BP算法与模拟退火算法相结合的方法。一方面,可以保持其训练速度,另一方面,又能找到全局最小。而且,国外已经成功地用这种方法训练了语音识别系统,其可行性是不用怀疑的。本文也用这种方法对一个具体的实例进行了计算,并将结果同一般的BP算法所得结果作了比较,比较结果表明这种方法的效果是不错的。

2 具体算法

2.1 BP网模型及其算法 一个BP网是由输入层、中间层、输出层三部分组成的,其结构如图1所示。其中 X, Z, Y 分别为输入矢量、中间层输出矢量和输出层矢量, W_{ij}, W_{jk} 分别为连接输入与中间层及中间层与输出层之间的连接权。 $X = (x_1, x_2, \dots, x_n)^T$ 为 n 维列矢量, $Z = (z_1, z_2, \dots, z_m)^T$ 为 m 维列矢量, $Y = (y_1, y_2, \dots, y_p)^T$ 为 p 维列矢量。中间层与输入层之间的关系为 $z_j = f(\sum_{i=1}^n W_{ij} \cdot x_i - \theta_j)$,其中 θ_j 为中间层第 j 个神经元的偏置门限, $f(\cdot)$ 为非线性可微函数,通常取 $f(u) = 1/[1 + \exp(-u)]$ 。输出层输出为 $y_k = f(\sum_{j=1}^m W_{jk} \cdot z_j - \theta_k)$,误差 $E = (1/2)(Y - D)^T \cdot (Y - D)$, $D = (y_{01}, y_{02}, \dots, y_{0p})$ 为期望输出。这样,根据误差反传算法可以通过调整权值来逐步减小误差,使实际输出逼近期望输出。权值调整方程为

¹ 1999-04-14收到, 1999-10-25定稿

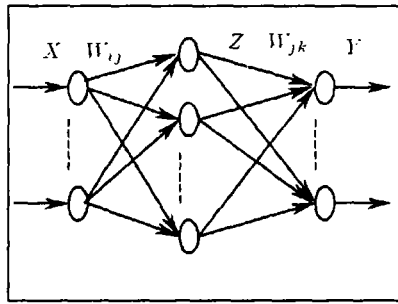


图 1

$$W_{jk}^{t+1} = W_{jk}^t + \eta \cdot (-\partial E / \partial W_{jk})$$

$$W_{ij}^{t+1} = W_{ij}^t + \eta \cdot (-\partial E / \partial W_{ij})$$

其中 $0 < \eta < 1$ 为调整的步长。为加快收敛速度可加一 momentum 项 $\alpha \cdot (W^t - W^{t-1})$, 其中 $0 < \alpha < 1$ 。训练中网络的所有权值及偏置值可初始化为 $[-1, 1]$ 之间均匀分布的随机数。

在选取模型时需注意中间层的选取, 因为中间层的层数与其神经元的个数决定模型所能反映的复杂程度。一般说来, 含三层非线性输出的网络可模拟任意复杂的非线性关系, 通常含一个中间层便可以了。关键是中间层神经元个数的选择, 要求在满足复杂性要求的前提下, 中间层神经元个数越少越好。可采取逐步增加神经元的方法, 但训练速度太慢。于是有用稀疏矩阵的方法^[6]。由于模型是基于数据训练的, 要求大量的、典型的数据, 对数据要求较高。有人提出了用神经网络结合一定微波电路的知识来建立模型^[7-8]。但是还没人讨论过对设计中的局部最小问题的解决方法。本文将给出一种 BP 结合模拟退火的方法。

2.2 模拟退火算法及其与 BP 算法的结合 模拟退火算法的原理可以这样理解, 一平板上有很多凹点, 其上有一小球滚动, 为保证其停在最深点而不是其它点, 给小球一足够大的动能, 使其能从所有凹点中跳出。显然小球跳出最深点所需能量最大, 逐步降低小球的动能, 只要保证下降速率足够慢, 就能找到一动能, 使小球能够跳出其它的凹点而不能从最深的凹点跳出。便找到了全局最小。该算法步骤为: (1) 从一足够高的初始温度 $T = T_0$ 出发, 随机地选择初态 s_0 (即 W_{ij}^0 、 W_{jk}^0), 计算其相应的能量 $E(s_0)$ (即误差); (2) 在 s_0 基础上随机移到一新态 s_1 , 计算能量变化 $\Delta E = E(s_1) - E(s_0)$; (3) 若 $\Delta E < 0$, 接受 s_1 , 否则, 计算 $p = 1/[1 + \exp(\Delta E/T)]$, 并与给定概率 ξ 比较; (4) 若 $\xi < p$, 接受 s_1 , 否则仍停留在 s_0 。返回到第 (2) 步; (5) 若已达到最优状态则停止, 否则, 降低温度 T , 重复以上过程。

从算法的步骤中可看出该算法收敛速度很慢。为了既保证一定的速度而又能找到全局最小, 把模拟退火算法与 BP 算法结合起来考虑。即在 BP 算法中权值的调整项中加一模拟退火调整项, 对是否接受模拟退火调整项仍由概率决定。公式为

$$W^{t+1} = W^t + \eta \cdot (-\partial E / \partial W) + (1 - \eta) \cdot rc$$

其中 $rc = T \cdot \text{tg}(rc_0)$ 为模拟退火项的取值, rc_0 为 $(-\pi/2, \pi/2)$ 之间一随机数, $T = T_0/(1+t)$ 为现时温度, T_0 为初始温度, t 为模拟退火调整项的有效调整次数。而是否接受 rc 的调整要由概率决定, 当 $\Delta E < 0$ 时, 接受调整, 否则, 计算概率 $p = 1/(1 + \exp(\Delta E/T))$ 并与给定概率 ξ 比较, 当 $p > \xi$ 时, 接受, 否则不接受。可看出, 当 $\eta = 1$ 时其为一般的 BP 算法, 当 $\eta = 0$ 时是纯粹的模拟退火算法。在训练中需注意, rc 的调整范围很大, 可能会使 W 很大, 而导致该层的输出值 y_{out} 饱和 (即为 0 或 1), 由于 BP 算法中的调整项含 $(1 - y_{out})$ 的

乘积项,使 W 的调整值为零,引起训练中止。当某一神经元的输出饱和时,需对其进行调整使 $W = -5 + 10/(1 + \exp(-W))$, 这样的调整对绝对值大的 W 影响大,而对绝对值小的 W 影响小,故能使误差的调整不受太大影响。

3 具体实例

本文用 BP 算法与模拟退火算法结合的方法对一含并联介质柱的微带传输线的 S 参数进行了训练。该结构如图 2 所示,其各项参数如图。其中 D 为微带宽度, d 为介质柱直径, h 为介质基带厚度。固定 $h=0.635$, $\epsilon_s = 100$, 设输入为 $X = (f, w, d, \epsilon_p)^T$, 输出为 $Y = (|s_{11}|, |s_{21}|)^T$, 中间层神经元个数为 15 个,用 100 组数据(数据取自文献[9])来建立计算该微带线 S 参数的神经网络模型。用 50 组数据作训练样本, 50 组数据作检验样本。用纯 BP 算法与本文算法两种方法分别地进行训练学习,并用检验样本进行了检验。两种方法的训练误差如图 3 所示。其中虚线为纯 BP 算法的训练误差曲线,实线为本文算法的训练误差曲线。训练与检验结果的比较如表 1 所示。表中数据为平方误差。可看到,纯 BP 算法在训练中陷入了局部最小,导致其估计能力不强。而本文算法可越过局部最小,并最终达到预期的精度。且本模型处理的是多维问题,比只能处理一维问题的一般逼近方法要更具优势。

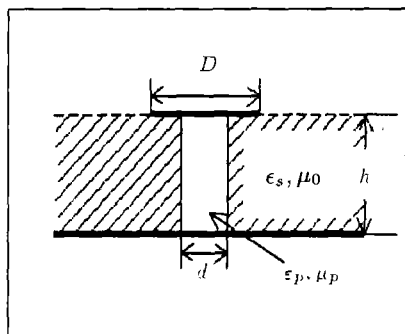


图 2

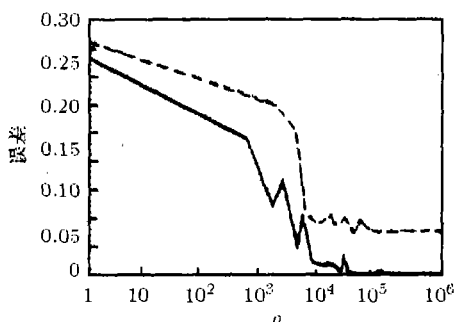


图 3

表 1

	平均训练误差	最大训练误差	平均检验误差	最大检验误差
BP 算法	0.05	0.13	0.075	0.14
本文算法	0.0001	0.0017	0.0003	0.0028

4 结 论

本文介绍了一种无论是从原理或是从实现方法来看都是比较简单的算法。而且,从具体使用的结果来看,效果不错,确实可以解决局部最小问题。

参 考 文 献

- [1] M. Vai, S. Prasad, Automatic impedance matching with a neural network. IEEE Microwave Guided Wave Lett., 1993, 3(3), 353-354.
- [2] P. M. Watson, K. C. Gupta, Design and optimizasion of CPW circuits, IEEE Trans. on MTT, 1997, 45(12), 2515-2523.
- [3] B. G. Hoskins, M. R. Haskard, Artificial neural network techniques for the estimation of thick-film resistance, Microelectronics Journal, 1995, 25(1), 9-16.
- [4] A. H. Zaabab, Zhang Q. J, M. Nakhla, A neural network modeling approach to circuits optimization and statistical design, IEEE Trans. on MTT, 1995, 43(6), 1349-1358.

- [5] A. Veluswami, M. S. Nakhla, Zhang Q. J, The application of neural networks to EM-base simulation and optimization of interconnects in high-speed VLSI circuits, IEEE Trans. on MTT, 1997, 45(5), 712-722.
- [6] A. H. Zaabab, Zhang Q. J, M. S. Nakhla, Device and circuit-level modeling using neural networks with faster training base on network sparsity, IEEE Trans. on MTT, 1997, 45(10), 1696-1704.
- [7] P. M. Watson, K. G. Gupta, EMM-ANN models for microstrip via and interconnects in multilayer circuits, IEEE Trans. on MTT, 1997, 45(12), 2495-2503.
- [8] F. F. Wang, Zhang Q J, Knowledge-based neural models for microwave design, IEEE Trans. on MTT, 1997, 45(12), 2333-2343.
- [9] K. L. Finch, N. Alexopoulos, Shunt post in microstrip transmission lines, IEEE Trans. on MTT, 1990, 38(11), 1585-1594.

A METHOD BASED ON NEURAL NETWORK TO ESTIMATE SCATTERING PARAMETER

Zhou Li Mao Junjie Yao Demiao

(National University of Defence Technology, Changsha 410073 China)

Abstract In this paper, a new neural network's training method combining back error propagation with simulated annealing used to microwave circuit design is discussed. This method can solve the local minimum question of the BP neural network. The new training technique does not require any structure change in neural network model. An example is calculated and the results verified the proposed technique.

Key words Neural network, Back error propagation, BP network, Local minimum, Simulated annealing

周 力: 男, 1975 年生, 博士生, 主要对宽带天线、微波电路 CAD、神经网络用于微波电路设计及计算电磁学感兴趣。

毛钧杰: 男, 1943 年生, 教授, 博士生导师, 主要从事雷达、微波毫米波技术与天线方面的研究与教学工作。

姚德森: 男, 1938 年生, 教授, 主要从事无线电制导、雷达、微波毫米波技术与天线方面的研究与教学工作。