

# 基于广义模拟退火的人工神经网络学习方法

严又生\*

(石油地球物理勘探局物探地质研究院)

## 摘 要

严又生. 基于广义模拟退火的人工神经网络学习方法. 石油地球物理勘探, 1996, 31(4): 476~482

本文利用广义模拟退火法(简称 GSA, 下同)能将非线性多极值目标函数较快收敛于全局极值的特点, 替换人工神经网络学习过程中基于梯度下降原理的误差反向传播算法(简称 BP, 下同)。该方法将由神经网络的学习输出与期待输出之差的平方和构成的目标函数视为一整体能量系统, 模拟热物理学中金属退火处理过程, 调整网络中的连接权值, 使系统能量尽可能收敛于全局极小。与 BP 法相比, 本方法无需计算梯度, 输出响应可采用不可微分的激励函数; 另外, 无需作误差反向传播计算, 因而在神经网络学习中可使用局部反馈连接的网络结构。该方法的应用为神经网络学习提供了一种新途径。

**主题词** 广义模拟退火 神经元 网络 反向 传输 目标函数

## ABSTRACT

Yan Yousheng. Neural-network learning method based on generalized simulated annealing. *OGP*, 1996, 31(4): 476~482

Generalized simulated annealing algorithm can make nonlinear multi-extremum object function converge to global extremum property, thus replacing error back-propagation algorithm based on gradient decline principle in learning of neural network. In the method described here, the object function which was constructed by using square sum of the differences between learning outputs and expected outputs of neural network is considered as a whole energy system to simulate metal annealing processing and to regulate joint weight values in network, so that energy in the system converges to global minimum. It is better than BP algorithm. This method needs no gradient computation, offers the output response in the form of nondifferentiable excitation function, and performs no error backpropagation computation; therefore, local feedback-jointed network structure can be introduced in the learning of neural network. This method is a new way for neural network learning.

\* Yan Yousheng, Geophysical & Geological Research Institute, Bureau of Oil Geophysical Prospecting, Zhuozhou City, Hebei Province, Postcode: 072751  
本文于 1995 年 12 月 11 日收到。

Subject heading: generalized simulated annealing, neuron, network, back direction, propagation, object function

## 引言

人工神经网络学习过程中常用的BP算法是基于梯度下降原理<sup>[1],[4]</sup>。对于具有极小值的能量函数,梯度下降法是沿能量函数的负梯度方向寻找能量极小值。对于单极值函数,该算法是精确的。但对于多极值非线性函数,其缺点就暴露无遗,即能量函数在迭代过程中容易产生振荡,或迭代收敛在函数的局部极值上。究其原因,主要有二:其一,梯度下降算法的导出是对目标函数采取了线性化处理,这种处理对于较复杂的目标函数是不合适的,由此导致了该种算法在非线性多极值函数全局寻优方面自身的缺陷;其二,对于目标函数梯度为零处(即极值处)及其较小的邻域内,该算法具有迭代方向的不确定性,使得迭代难以跳出局部极值,并可能形成迭代振荡。在现用的BP算法中,尽管对梯度下降算法做了许多修改,如选用连续可导的Sigmoid激励函数,对学习率、冲量及阈值进行调节,但仍未能在进行非线性反演的实质方面有所突破。本文所用的广义模拟退火法,从方法实质上摆脱了线性反演的框架,它是通过自变量随机搜索调整网络节点间权值和接收概率估算约束相结合的方法进行非线性反演,从而防止搜索过程收敛于局部最优解,有较大把握求得全局最优解。此外,从理论上讲,由于本方法的迭代过程是采用独立调节各节点之间的权值,不作误差回传,使得优化过程既可在全局空间上进行又可在局部子空间上进行。由此得到的神经网络系统的群体行为更为准确,诊断推理的适应性和灵活性更强。

## 基本原理

### 系统能量函数的构成

假设一神经网络由一个输入层、一个输出层及NH个隐蔽层组成。同时假定输出层节点数为NO,输入层节点数为NI,连接网络各节点之间的权值为 $W_{ij}$ ( $i, j$ 表示上层第 $i$ 个节点与下层第 $j$ 个节点的连接序号),学习样本数为NS,并规定输入层为底层,输出层为顶层,其结构如图1所示。这样,NS个学习样本通过该神经网络学习,在输出节点产生的 $NS * NO$ 个输出响应 $Y_{mn}$ 与其期待响应 $D_{mn}$ 构成了一个独立的能量系统,其能量函数 $E$ 可表示为

$$E = \frac{1}{2NS} \sum_{m=1}^{NS} \sum_{n=1}^{NO} (Y_{mn} - D_{mn})^2 \quad (1)$$

显然,该函数主要与连接网络各节点的权值 $W_{ij}$ 等参数有关。这样,网络学习的过程归结为求取该能量函数最小的过程,即是网络输出与期待输出的最佳拟合。实际上,该过程最终体现在如何调整权值上。

### 广义模拟退火法优化原理

模拟退火法是用数学手段模拟热物理学中金属退火过程<sup>[3]</sup>,即将待优化的目标函数设计成能量函数,通过调节“温度”及其它变量参数,使能量函数趋于最小能量状态(注:目标函数也可构造成求相关极大的形式),由此得到所求问题的解。标准的模拟退火法对变量修改是采取

随机游动采样,即从某一初始状态点(该点可任意给定也可特别给定)开始,计算出系统能量的初始状态  $E_0$ 。然后以一定步长  $\Delta r$ ,在独立的  $N$  维超平面上随机选择一个点( $N$  为所求问题的维数),计算该系统能量新的状态  $E_1$ 。该随机搜索点可按下式确定是否作为一个接收点,即

$$P = \begin{cases} 1 & \text{当 } \Delta E = E_1 - E_0 \leq 0 \\ \exp(-\beta \Delta E) & \text{当 } \Delta E > 0 \end{cases} \quad (2)$$

式中,  $\beta = K/T$ , 这里  $K$  为 Boltzmann 常数,  $T$  为绝对温度。当  $\Delta E \leq 0$ , 则该点为无条件接收, 说明系统能量向减小的方向变化; 当  $\Delta E > 0$ , 其接收情况必须根据辅助手段或经验来确定。通常采用的作法是在均匀分布的  $(0, 1)$  区间里产生一随机数  $\gamma$ , 如果  $\gamma < \exp(-\beta \Delta E)$ , 则该点为有条件接收; 否则, 该搜索点就放弃, 必须重新选择其它点。这意味着, 有条件接收的情况是由于能量减少的概率(或接受概率)较大, 尽管当前  $\Delta E > 0$ , 但根据其概率分布规律, 认为是可接收的。正是由于这种接收表面视为不利点( $\Delta E > 0$ )的概率总有大于零的性质, 使得迭代路径能跳出局部极小甚至全局极小。实验表明, 标准模拟退火法在优化过程中需要较多次迭代方能收敛到全局极小。为了克服这种缺点, 提出了广义模拟退火法<sup>[2]</sup>。其作法是对式(2)进行如下修改

$$P = \begin{cases} 1 \\ \exp(-\beta E^g \Delta E) \end{cases} \quad (3)$$

式中,  $g$  为一任意负数。显然当  $g=0$  时, 为标准模拟退火法表达式。通常取  $g=-1$ 。这样就可使得当能量函数趋于全局极小时, 接收不利搜索点的概率可能等于零, 从而保证了搜索点终止于全局极小点位置或在其附近。对于复杂的目标函数, 在迭代过程中  $\beta$  值应是可调的, 即温度参数是可控的, 这里主要包括起始温度, 临界温度和过程温度的控制。关于这方面的内容, 已有不少文献论述<sup>[6]、[7]</sup>, 这里不再赘述。此外, 迭代步长的选择取决于目标函数的性质以及优化过程期待的精度和分辨率, 必须根据

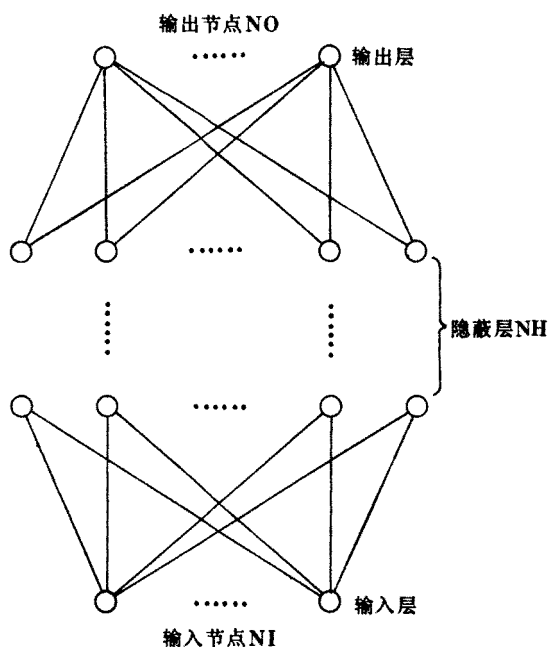


图1 神经元网络结构图

实际情况通过实验选定。

## 实 现 步 骤

综上所述, 对于图1所示的神经元网络, 其学习过程是使式(1)所示的系统能量逐渐减小的迭代过程。学习质量的好坏取决于系统能量减少的程度。在迭代过程中, 假定需要修改的参数主要是连接网络各节点的权值(这里忽略了阈值)。因此, 权值系数的修改方法成为关键。在具体的实施过程中发现, 若神经元网络的层数和节点数太少, 则严重影响学习和诊断识别的效果, 若太多则带来计算上的问题。因此必须根据能量函数的性质确定层数和节点数。通常情况

下,需调整的权系数个数较多,例如对于由输入层有4个节点、隐蔽层有10个节点和输出层有1个节点组成的三层网络,共有50个连接权值。这使得直接使用模拟退火法有困难,主要体现在计算量过大,难以通过整体调节权系数使能量降到最小。通过进一步分析发现,神经网络在结构上存在一定的局部独立性,即可把每个节点和与其相连的权值视为一个子体系。利用这种特性,本文采取的方法是从第一个隐蔽层的第一个节点开始,直到最终输出层(顶层)的最后一个节点,分别用广义模拟退火法进行局部调整权系数,在调整某个节点的权系数时,其它节点的权系数不变。这样从下至上逐步进行调整,使系统能量减小到给定范围。具体实现步骤为:

(1)根据实际情况,构造所需的神经网络系统(层数及各层的节点数),对该网络中所有权值初始化,确定学习样本集中的特征参数及个数。

(2)向网络提供学习样本,输入特征参数及期待输出。

(3)从网络最底隐蔽层的第一个节点开始,将由每一节点及其相连的权值构成的局部网络视为一个只有一个输出的独立子体系,用广义模拟退火法在 $N$ 维空间( $N$ 为该节点所连接的权系数个数)随机流动搜索调整权值,使得系统总能量下降。同时保证其它节点上的权值不变。将该过程遍及除输入层外的各节点,直至输出层的最后一个节点。

(4)若样本集中的样本学习尚未完成,转向第2步;否则比较现有的系统能量是否在给定误差范围之内。如果系统能量小于给定误差,则迭代终止,该网络可用于判断识别;否则转向第2步,重复上述过程。

用广义模拟退火法调整权值的主要过程归纳如下。假定权值 $W_{ij} \in \Omega$ ,  $\Omega$ 为 $N$ 维变量空间一子集。若 $E_m$ 为全局极小值,则:

(1)在 $\Omega$ 中,任意给定起始点 $W_0$ (也可特别指定)和初始步长 $\Delta r$ ;

(2)根据式(1)计算能量函数 $E_0(W_0)$ ,如果 $|E_0 - E_m| < \epsilon$ ( $\epsilon$ 为给定误差),则程序终止;

(3)在 $[-1, 1]$ 内随机选择方向 $U$ ,由此确定新的搜索点(权值),即

$$W^* = W_0 + (\Delta r)U$$

(4)如果 $W^* \notin \Omega$ ,转向第(3)步;

(5)计算 $E_1(W^*)$ 和 $\Delta E = E_1 - E_0$ ;

(6)如果 $E_1 \leq E_0$ ,则 $W_0 = W^*$ 和 $E_0 = E_1$ ;

(7)如果 $|E_0 - E_m| < \epsilon$ ,程序终止;否则转向第3步;

(8)如果 $E_1 > E_0$ ,则计算 $P = \exp(-\beta E^* \Delta E)$ ;

(9)在 $(0, 1)$ 之间产生随机数 $\gamma$ ,如果 $\gamma \geq P$ ,转向第3步;否则为条件接收,使 $W_0 = W^*$ ,  $E_0 = E_1$ ,转向第3步。

## 模型验证

神经网络技术已在地球物理领域得到了广泛应用。出于对本文提出的方法进行验证的考虑,这里选用了在井间地震观测系统下五层地质模型的波场计算结果中的一共炮点道集。如图2所示,其中共80道,每道600个样点数据。最初将本方法用于初至拾取,由于模型结果的初至特性明显,拾取效果很好。分别采用BP和GSA方法训练的神经元识别网络的拾取正确率均为100%。为了进一步验证本方法,选用该道集中来自地表反射的同相轴(波谷)作为拾取

目标,该同相轴无论在振幅、相位以及频率特性等方面都较初至复杂。这里使用的三层网络,输入层节点为 4 个,输出层节点为 1 个,隐蔽层节点为 10 个。学习过程中选用了 4 个特征参数,即主峰值、主瓣峰值与旁瓣峰值之差、相邻时窗的均方根振幅之比及视周期。学习样本为 9 个,从炮集中的 5 道数据中(第 1、10、30、50 和 70 道)提取,表 1 给出了用 BP 和 GSA 算法的学习结果。显然 GSA 法学习得到的精度较高。图 3 为用 GSA 算法学习拾取的结果,准确率为 100%。图 4 是用 BP 算法学习拾取的结果,有两道拾取错。为了检验两种算法训练的识别网络的容错性和灵敏性,本文将表 1 中的第 6 个样本去掉,留下 8 个样本学习。图 5 为用 GSA 算法学习拾取的结果,有 45 道拾取正确,其准确率为 56%。图 6 为用 BP 算法学习的结果,只有 1 个道拾取正确(第 26 道)。根据这些计算结果,说明了两个问题:其一,特征参数刻化的准确程度及个数在神经网络的识别过程中是如此的重要;其二,在相同条件下,GSA 比 BP 算法更能适用于较为复杂的神经网络学习和识别过程。

表 1 GSA 和 BP 算法计算结果比较

样本序号	主峰值	峰值差	振幅比	视周期	BP 法	GSA 法	目的输出
1	-9.07E-2	-1.17E-1	2.25	3.08E-1	0.9999	0.9999	1
2	-3.26E-2	-6.43E-2	4.19	2.19E-1	2.40E-4	7.77E-7	0
3	-5.54E-2	-1.08E-1	1.85	2.58E-1	0.9950	0.9999	1
4	-8.42E-2	-1.10E-1	2.56	2.36E-1	0.9995	0.9999	1
5	2.73E-3	-8.12E-3	1.57	2.92E-1	3.12E-3	1.54E-6	0
6	1.48E-1	2.25E-1	1.31	1.0	1.89E-6	3.84E-24	0
7	-7.03E-2	-1.08E-1	2.31	2.42E-1	0.9990	0.9999	1
8	-4.45E-2	-1.02E-1	1.61	4.11E-1	0.9993	0.9999	1
9	-9.09E-2	-1.54E-1	1.58	1.90E-1	3.87E-3	8.23E-7	0

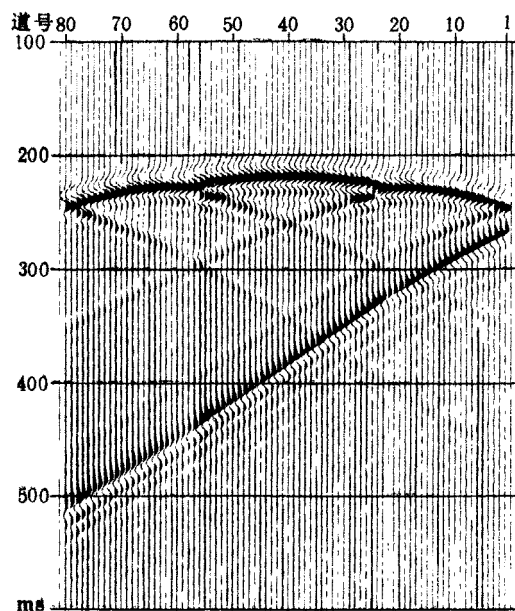


图 2 井间地震测量共炮点道集模型数据



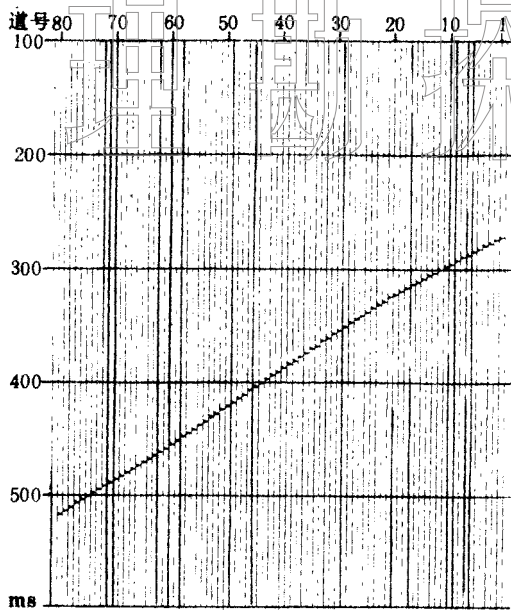


图 3 GSA 算法(9 个样点)拾取结果

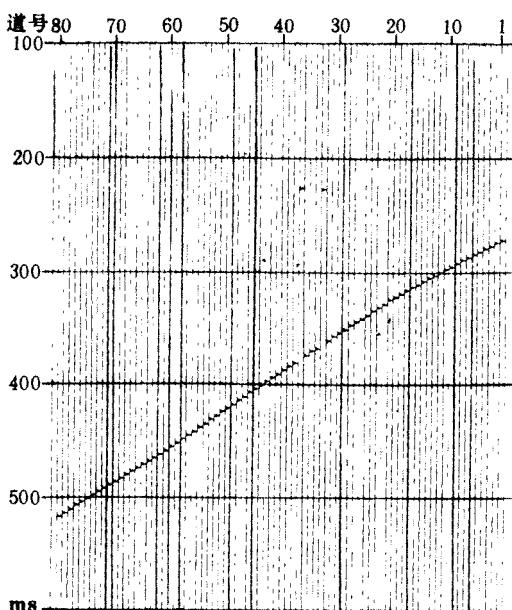


图 1 BP 算法(9 个样点)拾取结果

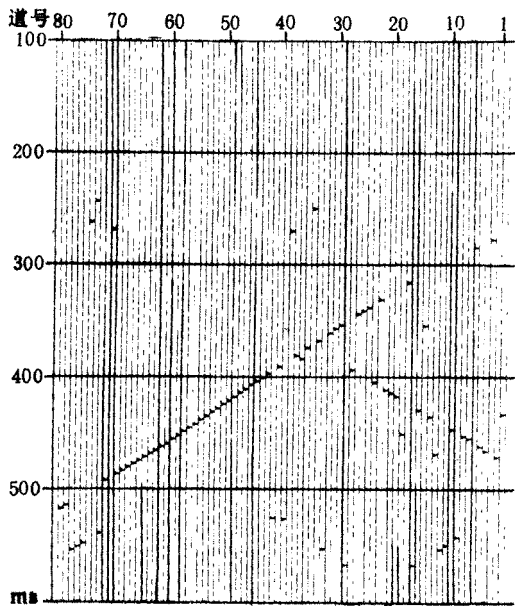


图 5 GSA 算法(8 个样点)拾取结果  
共 45 道正确

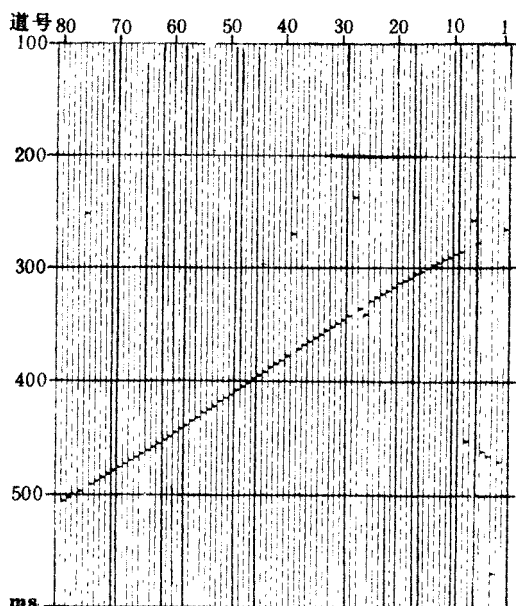


图 6 BP 算法(8 个样点)拾取结果  
仅有 1 道(第 26 道)正确,其余各道同相轴均存在时差

## 结 束 语

通过对广义模拟退火法原理及神经网络学习过程的分析,本文实现了将 GSA 算法替换 BP 算法用于神经网络学习,并根据大量模型的计算结果,得到以下认识:

(1)GSA 算法能克服 BP 算法易收敛于目标函数局部极值的缺陷。由 GSA 算法学习得到

的神经网络识别系统准确性更高、容错性更好。

(2)GSA 算法无需计算梯度,因而可使用不可微分的激励函数,拓宽了神经网络的学习和识别的领域。

(3)GSA 算法不做误差反向传播运算,因而允许使用局部反馈连接的网络系统,使其能适用于更多层和节点的神经网络学习。

(4)GSA 算法的迭代收敛过程是跳跃的,不如 BP 算法平稳,使得对不同问题的迭代时间、收敛程度较难预测。而且在 GSA 方法中对减少计算次数、最佳调节迭代步长和温度等仍是需要进一步探讨的问题。

应该指出,要将神经网络技术很好地用于生产,除了对学习算法进行改进外,更重要的是对不同问题的地球物理特征参数的刻画要准确。要注意应用范围和条件,不能滥用。

石油物探局物探地质研究院许增魁同志提供了有关资料,在此表示感谢。

#### 参 考 文 献

- 1 Michael E Murat and Albert J Rudman. Automated first arrival picking: A neural network approach. *Geophysical Prospecting*, 1992, 40(6): 587~604
- 2 Ihor O Bohachevsky et al. Generalized simulated annealing for function optimization. *Technometrics*, 1986, (28): 209~217
- 3 Kirkpatrick S et al. Optimization by simulated annealing. *Science*, 1983, 220: 617~679
- 4 杨叔子, 丁洪等. 基于知识的诊断推理. 清华大学出版社, 广西科学技术出版社
- 5 姚姚. 用人工神经网络实现同相轴自动拾取. 石油地球物理勘探, 1994, 29(1): 111~116
- 6 唐建侯, 张金山. 高效模拟退火剩余静校正. 石油地球物理勘探, 1994, 29(3): 382~387
- 7 陈均, 钱泉, 唐权钧. 用模拟退火法做叠后层速度计算. 石油物探, 1994, 32(2): 11~20

#### · 消息 ·

### SEG 北京联络部组成第二届执行委员会

SEG 北京联络部于 1996 年 6 月 12 日在涿州市召开了第二届执行委员会会议。孟尔盛、潘媛、陈祖传、裘慰庭、李衍达、许云、李光文、谢剑鸣、张淑敏和陈文菊、张平等 11 人出席了会议。王小牧、赵化昆和陆邦干因事请假未能到会。

会议开始由张淑敏女士通报 SEG 北京联络部第一届执委会的工作情况。随后由上届执委会主席孟尔盛向与会代表介绍已在杭州召开的“电磁方法研讨会”的情况及预定今年 9 月 12 日~15 日在新疆库尔勒市召开西部地区地震方法研讨会的准备情况。会议最后讨论和决定 SEG 北京联络部第二届执行委员会委员的分工。

SEG 北京联络部第二届执委会由下列人员组成:

现任主席: 陈祖传      当选(下任)主席: 王小牧  
第一副主席: 裘慰庭      第二副主席: 赵化昆  
执行委员: 李衍达 李光文 许云 谢剑鸣 张淑敏  
顾问: 潘媛 孟尔盛

立 早