

人工神经网络在岩石变形温度估算中的应用

李三忠 朱晓军

(长春地质学院, 长春 130026)

主题词 人工神经网络 BP 网络 肖钠长石 菱切面角 变形温度

提 要 Christie (1962)、Smith (1972) 和张翊钧 (1985) 都提出了钠质斜长石的 $\Delta_{131-131}$ (或 σ)、 An 与温度 T 的关系图, 后来又提出了钠质斜长石变形温度的计算公式。但由于图解的不便之处及目前应用的公式中推断分界点的不确定性和 σ 、 An 与 T 之间存在一种非线性关系的特点, 本文基于具有高度非线性映射能力的人工神经网络, 提出了求解岩石变形温度的新方法。文中先介绍了变形温度计的研究概况, 然后阐述了 ANN (Artificial Neural Network) 模型, 最后应用该模型估算岩石变形时的温度, 经对比得出, 利用人工神经网络在岩石变形温度估算中具有良好的应用效果。

1 引言

很早以前, 人们就开始寻求度量地质过程中温度、压力变化的方法, 于是地质温度计、地质压力计应运而生了, 然而它们只能解释变质作用发生时的温度和压力情况, 因为变形过程与变质过程并不总是同时发生的, 当变形作用与变质作用同时发生时, 我们可以利用变质作用形成的变质矿物, 采用适当的方法来求得温度和压力; 当变形和变质作用不同时发生时, 或者虽然变质与变形同时发生, 但同构造变质矿物没有相应的地质温压计时, 人们便希望从构造形迹上来推知变形时的温度。为此, 构造相观点提出了不同温压范围内各种矿物的变形特征和各种构造形迹的表现特征, 但都只能大致估计温度, 而不能获得准确数据。为了获得精确数据, 人们从显微乃至超显微构造上得到了突破。

19 世纪人们发现了位错现象。位错是晶体晶格中扩展的缺陷, 在大多数天然晶体中都存在。它们可能由晶体自身生长过程产生或由于晶体变形所造成。位错可产生攀移蠕变或滑动蠕变, 主要的蠕变机制取决于晶粒大小、温度、压力和应力大小等参数, 因此人们有望在显微构造与变形温度之间建立某种关系。此外, 矿物的晶胞具有热胀冷缩性质, 随着温度变化, 晶胞 a_0 、 b_0 、 c_0 轴的参数也变化, 同时经研究表明, 在构造应力双晶中, 肖钠双晶最普遍, 且肖钠双晶的菱切面角 σ [/(010) 面的解理在 (010) 面上与肖钠双晶的交角] 与晶胞参数 α^* 及 γ 相关, 即 $\text{Cot}\sigma = \text{Cos}\alpha^*/\text{Cot}\gamma$, 因此, 菱切面角与变形时的温度存在一种相关性。

由于作为测定变形温度的矿物必须满足以下两个条件: 一是矿物成分不能太复杂, 否则成分变化引起的结构变化与温度变化引起的晶体结构变化将无法区分, 因而普通角闪石、帘石等矿物不适合; 二是矿物的晶体对称程度不能高, 对称程度高的晶体容易调整它的结构状态,

它保留原结构状态的可能性就小,因而石榴石、石英、方解石、橄榄石不宜应用。目前发现斜长石比较适宜,斜长石是地球上最普遍的矿物之一,尤其是在花岗岩中。

斜长石的 $\Delta^{131-1\bar{3}1}$ 、成分 An 与温度 T 的关系图解最早由 Christie (1962) 据高温高压实验结果得出,但因据此测得的温度介于 $400-550^{\circ}\text{C}$ 之间,作为形成温度又偏低,又不知是否为变形温度,得不到合理的地质解释,影响了它的推广。因 $\Delta^{131-1\bar{3}1}$ 间的距离反映了晶胞参数,晶胞参数又与菱切面角有关,故人们找到了 σ 与 $\Delta^{131-1\bar{3}1}$ 之间存在的如下线性关系: $\Delta = 1.86563 - 0.0233358\sigma$ 。由此,张翊钧提出了 σ 、 An 和 T 的关系图解,因 σ 较 $\Delta^{131-1\bar{3}1}$ 易于测定,且测定 σ 比较经济,故得以推广应用。但图解有其不便之处,精度较低,从而提出了斜长石菱切面角变形温度计的计算公式^[4]。

变形温度计由 550°C 为分界点,利用回归分析拟合为两个计算公式。实际上分界点的确定具有推断性,而且公式的应用上,应事先对变形时温度的大致范围有所估计再选用。此外,与变形温度相关的参数 σ 、 An 与变形温度之间的相关关系是一种复杂的非线性映射关系。

基于上述情况,本文提出求解岩石变形温度估算的新方法——人工神经网络。人工神经网络在近 30 年的发展过程中已被广泛应用到自然科学与社会科学上,得到良好应用效果,BP (Back-propagation) 是 ANN 中的一种层状非线性前馈型网络,它由大量的处理单元——神经元广泛互连而成,构成一非线性系统,具有良好的非线性映射能力。我们将 BP 网络应用到变形温度的估算中,经实验验证、对比,该方法具有较高的精度,它既不需要确定分界点,也不需要事先了解变形温度的范围,比图解法也更为方便。

2 BP 网络的拓扑空间及学习方法

2.1 网络的基本单元——神经元

神经元是一个多输入、单输出的非线性单元,其结构如图 1。

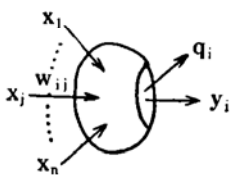


图 1 神经元结构

Fig. 1 Neuron structure

x_1, x_2, \dots, x_n 为神经元的输入;

w_{ij} 为神经元之间的连接权;

q_i 为神经元的阈值;

y_i 为神经元的输出。

上述神经元模型可描述为:

$$\sigma_i = \sum_{j=1}^n w_{ij}x_j + q_i$$

$$y_i = f(\sigma_i)$$

其中 $f(x)$ 是神经元的特性函数,在此我们用 S 形函数:

$$f(\sigma_i) = 1/(1 + e^{-\sigma_i})$$

2.2 网络拓扑结构

BP 网络是一种由神经元为基本单元构成的层次前馈型网络,其基本网络由三层构成,即输入层 (I)、中间隐含层 (H)、输出层 (O)。网络的连接方式如图 2 所示。信息从 I 层输入网络后单向流动,经 H 层变换,最后由 O 层输出,网络 I、O 层单元数是由具体问题的 I/O 参数来定, H 层的单元数由具体问题的复杂程度及可供学习的样本数的多少来定。Necht-Niclson

已证明对任意 L_2 空间上从 $[0, 1]^n$ 到 $[0, 1]^m$ 的映射 G , 都存在一个三层神经网络可以逼近 G , 一般情况下, 一个简单的三层网络可以满足要求, 网络的拓扑结构如图 2 所示。

x_1, x_2, \dots, x_n 为网络输入;

\tilde{W}_1, \tilde{W}_2 为单元间连接权;

\tilde{W}_1 : 输入层与中间隐含层间神经元连接权 ($w_{11}, w_{12}, \dots, w_{21}, \dots, w_{ij}, \dots, w_{nh}$);

\tilde{W}_2 : 中间隐含层与输出层间神经元的连接权 ($w'_{11}, w'_{12}, \dots, w'_{21}, \dots, w'_{ij}, \dots, w'_{hm}$)。

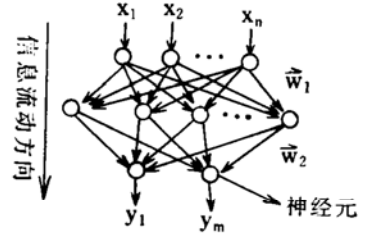


图 2 网络结构

Fig. 2 Network structure

2.3 BP 网络的学习算法

BP 网络间连接权 \tilde{W}_1, \tilde{W}_2 在网络的学习过程中不断地得到修正, 使得这两组权所构成的网络能实现学习样本中输入参数与输出参数间的特定的映射关系, 权的分布体现了各输入分量在输入矢量中所占特征强度的分布。网络学习的方法有多种, 我们采用了梯度下降算法。

设有 L 对学习样本 $(\tilde{x}_k, \tilde{O}_k) (k = 1, 2, \dots, L)$, 设一个输入 \tilde{x}_k 经网络传播后得到的实际输出为 \tilde{y}_k , 则它与要求的期望输出 \tilde{O}_k 之间的平方误差为:

$$E_k = \frac{1}{2} \sum_j^M (Y_{k,j} - O_{k,j})^2$$

M 为输出层单元数, $Y_{k,j}$ 为第 k 对样本第 j 分量的实际输出, $O_{k,j}$ 为第 k 对样本的第 j 分量的期望输出。

这样整个样本集的总误差为:

$$E = \sum_{k=1}^L E_k$$

由梯度下降法修改网络的权系数, 使得 E 取得极小值, 这样

$$\Delta w_{ij}^{(k)} = -\eta \left(\frac{\partial E_k}{\partial w_{ij}} \right)$$

η 为学习率, 可取介于 0 到 1 间的数值, 所有学习样本对对权 w_{ij} 的修正为:

$$\Delta w_{ij} = \sum_k^L \Delta w_{ij}^{(k)}$$

通常为增加学习过程的稳定性, 用下式对 w_{ij} 进行修正。

$$\Delta w_{ij}(t) = \sum_k^L \Delta w_{ij}^{(k)} + \beta(w_{ij}(t) - w_{ij}(t-1))$$

β 为冲量常数, $w_{ij}(t-1)$ 为网络前一次迭代循环训练后的连接权值。

在网络学习的过程中, 先调整输出层与中间隐含层间的连接权 \tilde{W}_2 , 然后再调整中间隐含层和输出层间的连接权 \tilde{W}_1 。

3 网络的建立及计算

我们采用三层 BP 网络, 网络的输入单元有 2 个, 分别对应着 (An, σ) , 一个输出单元

表1 网络连接权和神经元阈值参数表

Table 1 Parameters of network weight and neural threshold

连接权 w_1	连接权 w_2	H层神经元阈值	O层神经元阈值
w_{11} : 28.29412	w_{11} : 11.09991	θ_1 : -17.02844	θ_1 : -18.24490
w_{12} : -30.35948	w_{21} : 119.0905	θ_2 : 46.37627	
w_{13} : -120.5831	w_{31} : 4.552830	θ_3 : -12.18922	
w_{14} : -91.36080	w_{41} : -27.54686	θ_4 : -21.52176	
w_{15} : 11.32743	w_{51} : 19.96205	θ_5 : 50.19433	
w_{21} : 5.762324			
w_{22} : 19.51446			
w_{23} : 33.23441			
w_{24} : -180.7298			
w_{25} : -19.64627			

表2 实验值 T_0 、BP 网络计算值 T_1 及公式*计算值 T_2 对比表

Table 2 Comparison of experimental values (T_0), BP network calculated values (T_1) and formulae T_2

σ	An	T_0	T_1	T_2
33.4306	0	400	400.1	403.2632
30.0024	0	450	448.2	448.2717
23.6174	0	500	502.1	507.6635
28.7597	5	400	401.1	403.2049
25.9100	5	450	448.1	443.7025
20.3820	5	500	496.8	504.2000
24.8815	10	400	402.4	399.5957
21.8818	10	450	451.4	446.4739
17.5966	10	500	499.3	501.2677
21.1319	15	400	404.8	401.3022
18.4751	15	450	451.9	447.3645
14.8112	15	500	497.7	501.8525
17.5537	20	400	407.5	406.8599
15.6682	20	450	442.7	443.3391
12.0257	20	500	499.7	505.9546
14.8112	25	400	403.6	400.6522
12.5614	25	450	449.2	449.1690
10.2045	25	500	501.0	495.7617
11.5972	30	400	405.2	409.3660
9.9902	30	450	445.0	447.6266
7.9547	30	500	502.8	493.1983
8.3618	35	400	406.0	424.0142
6.8620	35	450	445.3	462.7511
5.1693	35	500	506.9	504.3609
11.2758	0	550	546.9	532.3260
10.2044	5	550	550.0	553.2426
9.4546	10	550	551.1	565.9147
8.9189	15	550	551.5	567.5200
8.3189	20	550	551.9	559.0516

续表 2

σ	An	T_0	T_1	T_2
7.7404	25	550	544.1	539.8395
5.8121	30	550	549.5	537.6774
4.4194	35	550	539.6	522.0812
7.3334	0	600	605.8	602.7283
5.0622	0	650	649.5	639.6768
7.0977	5	600	598.5	607.3838
4.6337	5	650	649.5	652.4052
6.8834	10	600	593.6	607.0375
4.4622	10	650	648.6	656.0719
6.4549	15	600	598.6	606.7607
4.5265	15	650	646.1	649.6312
6.0263	20	600	602.1	602.6847
4.2694	20	650	644.6	645.2669
5.3836	25	600	601.4	600.4082
3.7980	25	650	643.5	642.0692
4.4622	30	600	597.4	603.0664
3.0267	30	650	644.0	643.7802
3.2409	35	600	608.1	613.0627
1.7411	35	650	648.7	658.8251

* 据张翊钧公式计算。

表 3 五台地区变形温度计得出的值 (T_0) 与 BP 网络得出的值 (T_1) 对比表
 Table 3 Comparison between values calculated by deformation thermometer (T_0) and values estimated by BP network (T_1) for Wutai area

σ	An	T_0	T_1
13	20	490	489.87
16	8	520	512.69
11.5	18	520	518.47
11	20	520	517.99
10	22	520	523.99
23	0	505	510.40
15	14	505	497.50
11	26	470	474.43

(T), 中间层单元数由实际计算时误差下降情况定为 5 个。这样就形成了 2 个输入单元、5 个中间单元、1 个输出单元的三层 BP 网络。取 α 值为 0.15, 对样本集学习达 6000 次以后, 网络的平均误差小于 1.0%。经学习结束后, 形成的参数见表 1, 计算结果如表 2 所示, 最后采用五台地区已有的资料用该文方法计算结果见表 3。

由计算结果表明, BP 网络计算结果的精度比公式计算^[2]的结果更接近实验值。

4 结论

通过对实际资料的计算和对比, 我们可以得到以下几点结论:

(1) 人工神经网络所具有的高度非线性函数映射功能能较好地应用到岩石变形温度估算中,比以往的方法有更高的精度及实用性。一组网络经训练后即能记下训练集样本所包含的非线性函数关系。随着样本集的扩大,可以形成一种神经网络计算系统,从而得到一种通用性强的岩石变形温度估算的工具。

(2) 样本集与应用对象间的相似性是该方法的基础,同时在对输入参数及输出参数的归一化方式上要保持样本集和应用对象间的一致性。前者要求样本的选取要有代表性、分布均匀、覆盖范围越大越好,同时样本集要满足一定的数量。后者是因神经网络在 $[0,1]$ 空间工作,故归一化是预处理中的最重要的一步。

(3) 在网络学习的过程中,中间层单元数的选取有一定的要求,这一点主要是由经验给定。所选用的单元数越多可得到的函数关系可以越复杂,但单元数越多增加了每个学习周期的计算量。同时要求的样本数也越多,以保证所有的连接权都得到充分学习。一般情况又以能够使学习误差下降较快的所要的最少单元数来定义中间层单元数为宜。

(4) 由第2条结论,我们知道这种方法所得结果的应用范围是随着样本集的扩大、完善而不断进化的,随着我们占有资料的丰富,网络的性能会不断提高。提高的方法有二种:一种是针对不同特性的岩石形成特定的BP网络,这样形成的网络具有针对性,因而网络的中间隐含层单元数不必太多,且计算的精度及正确性都能得到保证,但这种网络的应用范围受到限制;另一种方法是,通过增加网络的复杂性,即增加中间隐含层单元数目,或增加网络的层数来达到目的,这种网络应用范围广,其缺点是学习时间长,且更易于陷入局部极小从而影响计算的精度和正确性。

致谢: 作者感谢天津地质矿产研究所的张翊钧先生提供了所有原始数据资料和岩石变形温度计理论方面的指导。

参 考 文 献

- 1 张翊钧. 菱切面角温度计作为变形温度计在五台山区的应用. 岩石矿物学杂志,1987,6(4).
- 2 张翊钧. 斜长石变形温度计. 科学通报,1992,(1).
- 3 焦李成. 神经网络系统理论. 西安电子科技大学出版社,1990.

The Application of Artificial Neural Network to the Estimation of Rock Deformation Temperature

Li Sanzhong, Zhu Xiaojun

(Changchun College of Geology, Changchun 130026)

Key words: artificial neural network, BP network, pericline twin, angle of rhombic section, deformation thermometer

Abstract

Diagrams showing relationship of the temperature (T), An and $\Delta 131-131$ (OR σ) of sodic plagioclase were advanced in succession by Christie (1962), Smith (1972) and Zhang Yijun (1985). Later, the formulae for deformation temperature of sodic plagioclase were put forward by Zhang Yijun. Nevertheless, in view of the inconvenience of the diagram, the uncertainty of the dividing point in the formulae and the nonlinear relationship among σ , An and T , we present a new method for calculating the temperature of the deformation—artificial neural network, which has accurate nonlinear projection. In this paper, an account is first given about the study of the deformation thermometer, followed by a description of the ANN (Artificial Neural Network) model. With this model, the temperature of the deformation is estimated. A comparison shows that the artificial neural network is very effective in the estimation of the deformation temperature.