

一种量子神经网络模型学习算法及应用

李盼池

(大庆石油学院 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要: 提出一种量子神经网络模型及学习算法. 首先基于生物神经元信息处理机制和量子计算原理构造出一种量子神经元, 该神经元由加权、聚合、活化、激励四部分组成. 然后由量子神经元构造出三层量子神经网络模型, 其输入和输出为实值向量, 权值和活性值为量子比特. 基于梯度下降法构造了该模型的超线性收敛学习算法. 通过模式识别和函数逼近两种仿真结果表明该模型及算法是有效的.

关键词: 量子计算; 量子神经元; 量子神经网络; 超线性收敛

中图分类号: TP183 **文献标识码:** A

A learning algorithm and its applications to the quantum neural network model

LI Pan-chi

(School of Computer and Information Technology, Daqing Petroleum Institute, Daqing Heilongjiang 163318, China)

Abstract: A quantum neural network model and its learning algorithm are presented. According to the information processing mode of the biology neuron and the quantum computing theory, we first propose a quantum neuron model which includes weighting, aggregating, activating, and prompting. Secondly, the quantum neural network model based on quantum neuron is constructed in which both the input and the output are real vectors and both the linked weight and the activation value are qubits. Using gradient descent algorithm, we also propose a super-linearly convergent learning algorithm of the quantum neural network. Finally, the availability of the approach is illustrated by two application examples of pattern recognition and function approximation.

Key words: quantum computing; quantum neuron; quantum neural network; super-linear convergence

1 引言(Introduction)

自从1989年多层神经网络被证明是连续函数的一致逼近器以来, 人工神经网络研究得到了长足发展, 现已被广泛应用于各个研究领域. 但也暴露了不少问题, 主要是因为它建立在及其简化的神经元模型之上, 因此难于更好地满足信息量和信息复杂度增加的需求^[1]. 量子计算的概念自20世纪80年代初被首次提出之后, 以其独特的计算性能引起了广泛瞩目, 并迅速成为研究的热点^[2,3]. 目前越来越多的研究者注意到, 量子理论在大脑神经系统中可能扮演一种至关重要的角色, 将ANN和量子理论相结合会更好模拟人脑的信息处理过程^[4]. 因此量子神经网络(quantum neural networks, QNN)现已成为一个极富前景的崭新研究领域. 本文提出一种具有量子特性权值和活性值的QNN模型及算法. 仿真结

果表明, 该模型及算法优于普通3层BP网络.

2 量子神经网络模型(Quantum neural networks model)

2.1 量子神经元模型(Quantum neuron model)

本文提出的量子神经元模型包括加权、聚合、活化、激励4部分. 其中, 加权是对神经细胞突触之间结合强度的模拟; 聚合是对多个树突接收到的刺激进行时空整合的模拟; 活化是对神经细胞由刺激引发的膜电位变化与当前活性值相互作用效果的模拟; 激励是对神经细胞的兴奋、抑制、疲劳、不应期以及阈值等非线性特性的模拟^[5]; 模型如图1所示. 图中 $|\phi_i\rangle$ 为对 x_i 的加权, $|\varphi\rangle$ 为活性值, Σ 为聚合算子, F 为活化函数, f 为激励函数.

记 $X = (x_1, x_2, \dots, x_n)^T$ 为输入向量; y 为输

出实数; $|\phi\rangle = (|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_n\rangle)^T$ 为权值, 其中 $|\phi_i\rangle = (\cos\theta_i, \sin\theta_i)^T$, $i = 1, 2, \dots, n$; $|\varphi\rangle = (\cos\xi, \sin\xi)^T$ 为活性值; τ 为阈值. 量子神经元的输入输出关系可描述为式(1):

$$y = f(F(X^T|\phi\rangle, |\varphi\rangle) - \tau) = f\left(\sum_{i=1}^n x_i \cos(\theta_i - \xi) - \tau\right). \quad (1)$$

其中: F 是内积算子, f 是sigmoid函数.

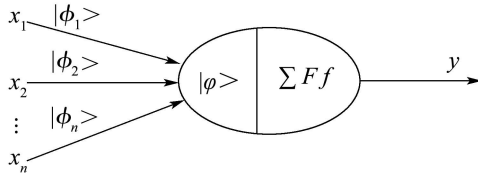


图1 量子神经元模型

Fig. 1 Quantum neuron model

2.2 量子神经网络模型(Quantum neural networks model)

由若干个量子神经元按一定的拓扑结构构成的网络称量子神经网络. 三层前馈量子神经网络模型如图2所示, 其中输入层、隐层、输出层各有 n , p , m 个量子神经元.

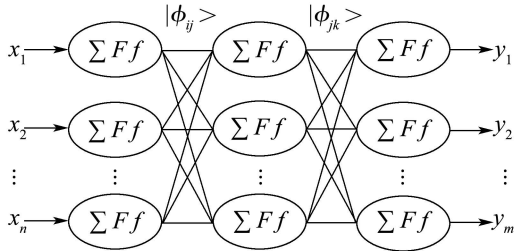


图2 量子神经网络模型

Fig. 2 Quantum neural networks model

记 x_i 为网络输入, h_j 为隐层输出, y_k 为网络输出, $|\phi_{ij}\rangle$ 为隐层权值, $|\varphi_j\rangle$ 为隐层活性值, τ_j 为隐层阈值, $|\phi_{jk}\rangle$ 为输出层权值, $|\varphi_k\rangle$ 为输出层活性值, τ_k 为输出层阈值. 量子神经网络的输入输出关系可描述为式(2):

$$\begin{aligned} y_k &= f\left(\sum_{j=1}^p h_j \langle \phi_{jk} | \varphi_k \rangle - \tau_k\right) = \\ &= f\left(\sum_{j=1}^p f\left(\sum_{i=1}^n x_i \langle \phi_{ij} | \varphi_j \rangle - \tau_j\right) \langle \phi_{jk} | \varphi_k \rangle - \tau_k\right) = \\ &= f\left(\sum_{j=1}^p f\left(\sum_{i=1}^n x_i \cos(\theta_{ij} - \xi_j) - \tau_j\right) \cos(\theta_{jk} - \xi_k) - \tau_k\right). \end{aligned} \quad (2)$$

其中: $i = 1, \dots, n$; $j = 1, \dots, p$; $k = 1, \dots, m$.

3 量子神经网络学习算法(Quantum neural networks algorithm)

在QNN模型中, 权值量子位和活性值量子位的修正可由单比特量子旋转门实现, 修正公式为

$$|\phi(t+1)\rangle = \begin{bmatrix} \cos \Delta\theta_t & -\sin \Delta\theta_t \\ \sin \Delta\theta_t & \cos \Delta\theta_t \end{bmatrix} |\phi(t)\rangle, \quad (3)$$

$$|\varphi(t+1)\rangle = \begin{bmatrix} \cos \Delta\xi_t & -\sin \Delta\xi_t \\ \sin \Delta\xi_t & \cos \Delta\xi_t \end{bmatrix} |\varphi(t)\rangle. \quad (4)$$

因此, 设计算法的关键是如何确定量子旋转门的转角 $\Delta\theta(t)$, $\Delta\xi(t)$ 和阈值修正量 $\Delta\tau(t)$, 以使如下3个迭代序列 $\{\theta(t)\}$, $\{\xi(t)\}$, $\{\tau(t)\}$ 快速收敛:

$$\theta(t+1) = \theta(t) + \Delta\theta(t), \quad (5)$$

$$\xi(t+1) = \xi(t) + \Delta\xi(t), \quad (6)$$

$$\tau(t+1) = \tau(t) + \Delta\tau(t). \quad (7)$$

定义网络误差函数为

$$E = (\hat{Y} - Y)^2 / 2. \quad (8)$$

其中 $\hat{Y}$ 和 Y 分别表示归一化后的期望输出和实际输出. 显然 E 是 $\theta(t)$, $\xi(t)$, $\tau(t)$ 的一致连续函数.

引理 1^[6] 设 $\theta(t) \rightarrow \theta^*$, 记

$$\rho(t) = \theta(t) - \theta^*,$$

$$\delta(t) = \theta(t+1) - \theta(t) = \rho(t+1) - \rho(t),$$

则 $\{\theta(t)\}$ 超线性收敛等价于 $\|\rho(t+1)\| = O(\|\delta(t)\|)$.

关于 $\Delta\theta(t)$, $\Delta\xi(t)$, $\Delta\tau(t)$ 的计算, 有如下定理.

定理 1 当 $\Delta\theta(t)$, $\Delta\xi(t)$, $\Delta\tau(t)$ 分别按式(9)(10)(11)计算且 $r > 2$ 时, 迭代序列 $\{\theta(t)\}$, $\{\xi(t)\}$, $\{\tau(t)\}$ 是超线性收敛的:

$$\Delta\theta(t) = -\eta \left(\frac{\partial E(t)}{\partial \theta(t)} \right) (E(t))^{\frac{1}{r}}, \quad (9)$$

$$\Delta\xi(t) = -\eta \left(\frac{\partial E(t)}{\partial \xi(t)} \right) (E(t))^{\frac{1}{r}}, \quad (10)$$

$$\Delta\tau(t) = -\eta \left(\frac{\partial E(t)}{\partial \tau(t)} \right) (E(t))^{\frac{1}{r}}. \quad (11)$$

其中 η 是学习速率.

下面借鉴文献[6]的思路给出证明.

证 首先证明其收敛性. 由泰勒公式可得

$$\begin{aligned} E(t+1) - E(t) &= \\ &= \begin{bmatrix} \Delta\theta(t) \\ \Delta\xi(t) \\ \Delta\tau(t) \end{bmatrix}^T \begin{bmatrix} \frac{\partial E(t)}{\partial \theta(t)} \\ \frac{\partial E(t)}{\partial \xi(t)} \\ \frac{\partial E(t)}{\partial \tau(t)} \end{bmatrix} + O\left(\left\| \begin{bmatrix} \Delta\theta(t) \\ \Delta\xi(t) \\ \Delta\tau(t) \end{bmatrix} \right\|\right) \approx \end{aligned}$$

$$-\eta \left(\left(\frac{\partial E(t)}{\partial \theta(t)} \right)^2 + \left(\frac{\partial E(t)}{\partial \xi(t)} \right)^2 + \left(\frac{\partial E(t)}{\partial \tau(t)} \right)^2 \right) (E(t))^{\frac{1}{r}}.$$

因此, $E(t+1) - E(t) \leq 0$, 即 $\{\theta(t)\}, \{\xi(t)\}, \{\tau(t)\}$ 单调下降; 因 $E(t)$ 有下界, 故极限存在; 又 $E(t)$ 关于 $\theta(t), \xi(t), \tau(t)$ 一致连续, 必有

$$\lim_{t \rightarrow \infty} (\theta(t), \xi(t), \tau(t)) = (\theta^*, \xi^*, \tau^*),$$

收敛性得证.

下面证明其收敛速度是超线性的. 记

$$A = \left(\frac{\partial E(t)}{\partial \theta(t)} \right)_{\max},$$

则有

$$\frac{\|\rho(t+1)\|}{\|\delta(t)\|} = \frac{\|\theta(t+1) - \theta^*\|}{\left\| -\eta \left(\frac{\partial E(t)}{\partial \theta(t)} \right) (E(t))^{\frac{1}{r}} \right\|} \geq \frac{\|\theta(t+1) - \theta^*\|}{\eta A (E(t))^{\frac{1}{r}}},$$

因 Y 是 $\theta(t), \xi(t), \tau(t)$ 的一次函数, 根据式(8)可知, $E(t)$ 为 $\theta(t), \xi(t), \tau(t)$ 的二次函数. 因此, 当 $r > 2$ 时, 有 $\|\theta(t+1) - \theta^*\| = O(E(t)^{\frac{1}{r}})$. 所以, $\|\rho(t+1)\| = O(\|\delta(t)\|)$. 由引理1可知, $\{\theta(t)\}$ 超线性收敛. 同理, $\{\xi(t)\}, \{\tau(t)\}$ 超线性收敛.

定理 2 若 θ^*, ξ^* 是序列 $\{\theta(t)\}, \{\xi(t)\}$ 的收敛解, 则任何满足 $\hat{\theta}^* - \hat{\xi}^* = \pm(\theta^* - \xi^*)$ 的 $\hat{\theta}^*, \hat{\xi}^*$ 也是序列 $\{\theta(t)\}, \{\xi(t)\}$ 的收敛解.

证 由 θ^*, ξ^* 是序列 $\{\theta(t)\}, \{\xi(t)\}$ 的收敛解, 根据式(2), 网络输出为

$$\begin{aligned} y_k &= f\left(\sum_{j=1}^p f\left(\sum_{i=1}^n x_i \cos(\theta_{ij}^* - \xi_j^*)\right) \cos(\theta_{jk}^* - \xi_k^*) - \tau_k\right) = \\ &f\left(\sum_{j=1}^p f\left(\sum_{i=1}^n x_i \cos(\pm(\theta_{ij}^* - \xi_j^*))\right) \cos(\pm(\theta_{jk}^* - \xi_k^*)) - \tau_k\right) = \\ &f\left(\sum_{j=1}^p f\left(\sum_{i=1}^n x_i \cos(\hat{\theta}_{ij}^* - \hat{\xi}_j^*)\right) \cos(\hat{\theta}_{jk}^* - \hat{\xi}_k^*) - \tau_k\right) = \hat{y}_k. \end{aligned}$$

因此, $\hat{\theta}^*, \hat{\xi}^*$ 也是序列 $\{\theta(t)\}, \{\xi(t)\}$ 的收敛解. 其中 $\hat{y}_k$ 为与 y_k 对应的期望输出.

4 仿真结果(Simulation results)

本节仿真将与普通BP网络(common BP, CBP)对比. 为便于比较, 两种网络取相同的结构及参数.

4.1 模式识别问题(Pattern recognition)

例 1 九点模式分类: 样本数据见表1. 网络结构取2-10-1型, 学习速率取0.8, 选取3种不同的误差精度, 分别用QNN和CBP训练50次, 然后对迭代步数取平均值. 训练结果见表2.

表 1 九点模式分类问题样本数据

Table 1 The datum of nine-sample patterns

输入	-1, -1	-1, 0	-1, 1	0, -1	0, 0	0, 1	1, -1	1, 0	1, 1
输出	1	0	1	0	1	0	1	0	1

表 2 QNN和CBP训练结果

Table 2 Training results of QNN and CBP

限定误差	限定步数	QNN		CBP		
		平均步数	收敛次数	平均步数	收敛次数	收敛需要步数
0.10	100	65.78	49	100	不收敛	605.36
0.05	500	221.56	47	500	不收敛	1693.21
0.01	5000	3735.64	43	5000	不收敛	14238.53

例 2 双螺旋线分类: 在双螺旋线

$$\begin{cases} x_1 = \theta \cos \theta, \\ y_1 = \theta \sin \theta, \\ x_2 = \theta \cos(\pi + \theta), \\ y_2 = \theta \sin(\pi + \theta), \end{cases} \theta \in (0, 3\pi)$$

上分别取48个点, 构成两类样本集合, 用本文QNN给出正确的分类. 网络结构取2-10-1型, 学习速率取0.6, 选取两种误差精度分别用QNN和CBP训练50次, 并对迭代步数取平均值. 训练结果见表3.

表 3 QNN和CBP训练结果

Table 3 Training results of QNN and CBP

限定误差	限定步数	QNN		CBP		
		平均步数	收敛次数	平均步数	收敛次数	收敛需要步数
0.10	2000	1580.38	45	2000	不收敛	4471.36
0.05	10000	6380.86	42	9311.36	16	18930.83

4.2 函数逼近问题(Function approximation)

以三角函数 $f(x) = \cos(1/x)$ 为例, 该函数在区间 $[1/(2\pi + 1), 1]$ 上11个离散点的函数值见表4, 用本文QNN逼近表中数据. 网络结构取1-10-1型, 学习速率0.5, 限定误差0.05, 限定迭代步数15000. 实际迭代12242次后收敛, 逼近误差为0.0499. 用相同结构的CBP逼近, 迭代15000次后, 没有收敛, 逼近误差为0.0627.

表4 函数 $f(x) = \cos(1/x)$ 的11组数据
Table 4 11 groups datum of $f(x) = \cos(1/x)$

x	$f(x)$	x	$f(x)$
0.1373	0.5403	0.2846	-0.9317
0.1503	0.9317	0.3466	-0.9672
0.1659	0.9672	0.4431	-0.6333
0.1852	0.6333	0.6141	-0.0575
0.2096	0.0575	1.0000	0.5403
0.2415	-0.5403	—	—

5 结论(Conclusion)

量子神经网络是量子计算与神经计算相融合的产物. 本文基于生物神经元信息处理方式及其呈现出的量子特性, 提出了一种量子神经网络模型, 证明了该模型在一定条件下是超线性收敛的, 且能增加收敛的概率. 量子神经元的权值和活性值采用量子位表示, 并用量子门更新其相位, 使其两个概率幅都参与了优化过程, 因此, 量子神经网络和优化能力优于普通神经网络. 仿真结果表明本文QNN模型及算法是有效的, 从而为量子计算与神经计算的融合开辟了一条新途径.

参考文献(References):

[1] 焦李成, 杜海峰, 刘芳, 等. 免疫优化计算、学习与识别[M]. 北京:

科学出版社, 2006: 180 – 181.

(JIAO Licheng, DU Haifeng, LIU Fang, et al. *Immune Optimization: Computation, Learning and Recognition*[M]. Beijing: Science Press, 2006: 180 – 181.)

[2] SHOR P W. Algorithms for quantum computation: Discrete logarithms and factoring[C] // *Proceedings of the 35th Annual Symp on Foundations of Computer Science*. New York, USA: IEEE Computer Society Press, 1994, 11: 124 – 134.

[3] GROVER L K. A fast quantum mechanical algorithm for database search[C] // *Proceedings of the 28th annual ACM Symp on Theory of Computing*. New York, USA: ACM Press, 1996, 6: 212 – 219.

[4] YUKINARI K. Indispensable role of quantum theory in the brain dynamics[J]. *BioSystems*, 2005: 263 – 272.

[5] 李士勇. 模糊控制·神经控制和智能控制论[M]. 哈尔滨: 哈尔滨工业大学出版社, 1998, 9: 77 – 81.

(LI Shiyong. *Fuzzy control, Neurocontrol and Intelligent Cybernetics*[M]. Harbin: Harbin Institute of Technology Press, 1998, 9: 77 – 81.)

[6] 梁久祯, 何新贵, 黄德双. 前馈网络的一种超线性收敛BP学习算法[J]. 软件学报, 2000, 11(8): 1094 – 1096.

(LIANG Jiuzhen, HE Xin'gui, HUANG Deshuang. Super-linearly convergent BP learning algorithm for feed forward neural networks[J]. *Journal of Software*, 2000, 11(8): 1094 – 1096.)

作者简介:

李盼池 (1969—), 男, 副教授, 博士研究生, 研究方向为量子优化算法, E-mail: lipanchi@vip.sina.com.