文章编号: 1001-0920(2012)01-0143-04

基于量子门线路的量子神经网络模型及算法

李盼池[a,b], 宋考平[a], 杨二龙[a]

(东北石油大学 a. 石油与天然气工程博士后科研流动站，b. 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算法. 该模型为一组量子门线路. 输入信息用量子位表示, 经量子旋转门进行相位旋转后作为控制位, 控制隐层量子位的翻转; 隐层量子位经量子旋转门进行相位旋转后作为控制位, 控制输出层量子位的翻转. 以输出层量子位中激发态的概率幅作为网络输出, 基于梯度下降法构造了该模型的学习算法. 仿真结果表明, 该模型及算法在收敛能力和鲁棒性方面均优于普通 BP 网络.

关键词: 量子计算；量子旋转门；受控非门；量子线路；量子神经网络

中图分类号: TP18 **文献标识码:** A

Quantum neural networks model and algorithm based on quantum gates circuit

LI Pan-chi[a,b], *SONG Kao-ping*[a], *YANG Er-long*[a]

(a. Post-doctoral Research Center of Oil and Gas Engineering, b. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China. Correspondent: LI Pan-chi, E-mail: lipanchi@vip.sina.com)

Abstract: A quantum neural networks model and its algorithm are proposed. The model is a group of the quantum gate circuits. The input information is expressed as the qubits, which, as the control qubits after rotated by the rotation gate, control the qubits in hidden layer to reversal. The qubits in hidden layer, as the control qubits after rotated by the rotation gate, control the qubits in output layer to reversal. The final output is described by the probability amplitudes of excited states in output layer. The learning algorithm is presented based on the gradient descent algorithm. The simulation results show that the proposed algorithm is superior to the common BP neural networks in both convergence capability and robustness.

Key words: quantum computing; quantum rotation gate; controlled-not gate; quantum circuit; quantum neural networks

1 引 言

量子神经网络的概念出现于 20 世纪 90 年代, 人们在这个崭新的领域进行了不同方向的探索, 提出了很多想法和初步的模型, 充分体现了量子神经网络研究的巨大潜力. 1995 年, Kak[1]首次提出了量子神经计算的概念. 随后, Gopathy 等人[2]借用量子理论中量子态叠加的思想, 提出了基于多级激励函数的量子神经网络模型, 在 3 层网络结构中, 隐层量子神经元的激励函数采用多个传统激励函数的叠加, 使网络具有一种固有的模糊性. Dan 等人[3]提出了量子联想存储算法, 相对于传统存储, 量子联想存储具有指数级的存储容量. Ajit 等人[4]从多宇宙的量子理论观点, 提出了构建叠加的多宇宙量子神经网络模型的思想. 解光军和庄镇泉[5]首次在国内撰文阐述了量子神经计算的概念, 对国内量子神经网络研究起到了一定的先导作用. 解光军等人[6-7]深入研究了以通用量子逻辑门组作为计算基函数, 构造量子神经网络模型的新方法, 研究结果表明该模型的性能优于传统神经网络. Matsui 等人[8]根据量子线路结构提出一种求解异或问题的量子神经元学习算法. 文献 [9] 提出一种基于量子输入和量子权值的量子自组织特征映射网络模型及聚类算法, 可明显提高普通自组织网络的聚类性能. [10] 基于量子旋转门和量子受控非门的物理意义, 提出一种基于通用量子门组演化的量子 BP 网络模型. 随后, [11] 提出一种基于量子权值和量子活性值

收稿日期: 2010-09-07；**修回日期:** 2010-12-08.

基金项目: 国家自然科学基金项目(61170132)；中国博士后科学基金特别项目(201003405)；中国博士后科学基金项目(20090460864)；黑龙江省博士后科学基金项目(LBH-Z09289)；黑龙江省教育厅科学技术研究项目(11551015).

作者简介: 李盼池(1969–), 男, 副教授, 博士, 从事量子搜索、量子智能优化等研究；宋考平(1962–), 男, 教授, 博士生导师, 从事油气田开发工程等研究.

的量子神经网络模型, 同时基于梯度下降法构造了该模型的超线性收敛学习算法. 目前量子计算与神经计算的融合正逐步成为一个崭新的研究方向.

本文在分析量子受控运算的基础上, 提出一种基于受控量子旋转门的量子神经元模型和量子神经网络模型, 并基于量子计算原理构造了该模型的学习算法. 仿真结果表明, 该模型及算法在收敛能力和鲁棒性两个方面明显优于普通BP(CBP)网络.

2 量子比特及量子门的概念

2.1 量子比特

在量子计算中, 量子比特有两个可能的状态$|0\rangle$和$|1\rangle$. 与经典比特的区别在于, 量子比特的状态可以落在$|0\rangle$和$|1\rangle$之外, 即可以是状态的线性叠加态, 其形式为

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1}$$

其中α和β是一对复数, 称为量子态的概率幅. 量子态也可用概率幅表示为$|\phi\rangle = [\alpha, \beta]^{\mathrm{T}}$.

2.2 量子门及线路表示

量子门是物理实现量子计算的基础, 它包含了量子计算的特点.

2.2.1 量子旋转门

量子旋转门定义为

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{2}$$

令$|\phi\rangle = \begin{bmatrix} \cos\theta_0 \\ \sin\theta_0 \end{bmatrix}$, 由$R(\theta)|\phi\rangle = \begin{bmatrix} \cos(\theta_0+\theta) \\ \sin(\theta_0+\theta) \end{bmatrix}$知, $R(\theta)$实现了$|\phi\rangle$的相位旋转.

2.2.2 多位受控非门

设有$n+1$位量子比特, $|x_1\rangle, |x_2\rangle, \cdots, |x_n\rangle$是控制比特, $|\phi\rangle$是目标比特, X是单比特量子非门, 多位受控非门如图1所示, 其运算式$C^n(X)$定义为

$$C^n(|x_1x_2\cdots x_n\rangle|\phi\rangle) = |x_1x_2\cdots x_n\rangle X^{x_1x_2\cdots x_n}|\phi\rangle. \tag{3}$$

即若前n个量子比特全为1, 则非门X作用到$|\phi\rangle$使其翻转; 否则, $|\phi\rangle$状态不变.

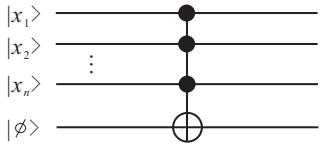

图 1 多位受控非门

3 量子神经网络模型算法

3.1 网络模型

本文提出的基于量子门线路的量子神经网络(QNN)如图2所示. 其中: $|x_1\rangle, |x_2\rangle, \cdots, |x_n\rangle$为输入; $|h_1\rangle, |h_2\rangle, \cdots, |h_p\rangle$为隐层输出; $|y_1\rangle, |y_2\rangle, \cdots, |y_m\rangle$为网络输出.

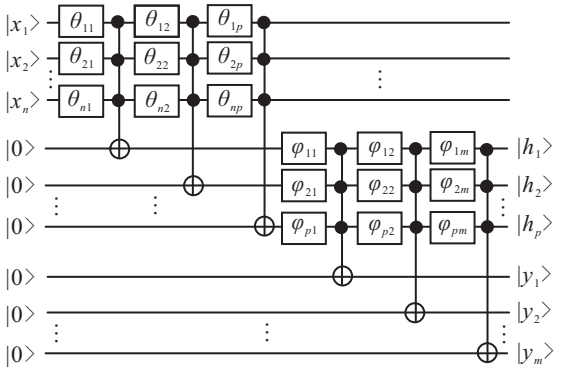

图 2 量子神经网络模型

本文约定, 以各层量子神经元输出中状态$|1\rangle$的概率幅作为该层的实际输出. 令$|x_i\rangle = \cos\theta_i|0\rangle + \sin\theta_i|1\rangle$, 根据量子计算原理, 网络各层的实际输出为

$$h_j = \sin(\varphi_j) = \prod_{i=1}^{n}\sin(\theta_i + \theta_{ij}), \tag{4}$$

$$y_k = \prod_{j=1}^{p}\sin(\theta_j + \theta_{jk}) = \prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\theta_i+\theta_{ij})\Big) + \varphi_{jk}\Big). \tag{5}$$

3.2 学习算法

设归一化后的期望输出为$\overline{y}_1, \overline{y}_2, \cdots, \overline{y}_m$, 定义误差函数为

$$E = \frac{1}{2}\sum_{k=1}^{m}(\overline{y}_k - y_k)^2. \tag{6}$$

令

$$\widehat{\theta}_{ij} = h_j\cot(\theta_i+\theta_{ij})\Big/\sqrt{1-h_j^2},$$

$$\delta = \sum_{k=1}^{m}\delta_k,\ \delta_k = (\overline{y}_k - y_k)y_k.$$

根据梯度下降法, 各层旋转相位的更新式为

$$\theta_{ij}(t+1) = \theta_{ij}(t) + \eta\delta\cot(\varphi_j+\varphi_{jk})\widehat{\theta}_{ij}, \tag{7}$$

$$\varphi_{jk}(t+1) = \varphi_{jk}(t) + \eta\delta_k\cot(\varphi_j+\varphi_{jk}). \tag{8}$$

其中: t为迭代步数, η为学习速率.

3.3 多重吸引子定理

由式(5)可以看出, 两个迭代序列$\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的全局最优解并不唯一. 事实上, 该模型存在大量全局吸引子, 从而可显著扩充最优解的数量. 具体表述为如下定理.

定理 1 设$\overline{\theta}_{ij}, \overline{\varphi}_{jk}$是迭代序列$\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的全局最优解, 则当$n_1$, n_2为任意整数时, 由如下两式确定的$\widehat{\theta}_{ij}$和$\widehat{\varphi}_{jk}$也是$\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的全局最优解:

$$\begin{cases}\widehat{\theta}_{ij}=2n_1\pi+\overline{\theta}_{ij},\\ \widehat{\varphi}_{jk}=2n_2\pi+\overline{\varphi}_{jk};\end{cases}\tag{9}$$

$$\begin{cases}\widehat{\theta}_{ij}=(2n_1+1)\pi-(2\theta_i+\overline{\theta}_{ij}),\\ \widehat{\varphi}_{jk}=2n_2\pi+\overline{\varphi}_{jk}.\end{cases}\tag{10}$$

证明　设 $\widehat{y}_k$ 为与实际输出 y_k 对应的期望输出. 根据式(5), 有

$$y_k=\prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\theta_i+2n_1\pi+\overline{\theta}_{ij})\Big)+2n_2\pi+\overline{\varphi}_{jk}\Big)=$$
$$\prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\theta_i+\overline{\theta}_{ij})\Big)+\overline{\varphi}_{jk}\Big)=\overline{y}_k.$$

故式(9)确定的 $\widehat{\theta}_{ij}$ 和 $\widehat{\varphi}_{jk}$ 是全局最优解. 因

$$y_k=\prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\theta_i+(2n_1+1)\pi-(2\theta_i+\overline{\theta}_{ij}))\Big)+2n_2\pi+\overline{\varphi}_{jk}\Big)=$$
$$\prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\pi-(\theta_i+\overline{\theta}_{ij}))\Big)+\overline{\varphi}_{jk}\Big)=$$
$$\prod_{j=1}^{p}\sin\Big(\arcsin\Big(\prod_{i=1}^{n}\sin(\theta_i+\overline{\theta}_{ij})\Big)+\overline{\varphi}_{jk}\Big)=\overline{y}_k.$$

故式(10)确定的 $\widehat{\theta}_{ij}$ 和 $\widehat{\varphi}_{jk}$ 是全局最优解. □

定理 2　对于任意 $\phi\in R$, 在 $[\phi,\phi+2\pi]$ 内存在迭代序列 $\{\theta_{ij}(t)\}$ 和 $\{\varphi_{jk}(t)\}$ 的所有全局最优解.

证明　设 $\overline{\theta}_{ij}$ 和 $\overline{\varphi}_{jk}$ 分别是迭代序列 $\{\theta_{ij}(t)\}$ 和 $\{\varphi_{jk}(t)\}$ 的全局最优解, 且 $\overline{\theta}_{ij}\notin[\phi,\phi+2\pi]$, $\overline{\varphi}_{jk}\notin[\phi,\phi+2\pi]$. 则存在整数 n_1 和 n_2 满足

$$\begin{cases}\widehat{\theta}_{ij}=2n_1\pi+\overline{\theta}_{ij},\\ \widehat{\varphi}_{jk}=2n_2\pi+\overline{\varphi}_{jk};\end{cases}\quad\begin{cases}\widehat{\theta}_{ij}\in[\phi,\phi+2\pi],\\ \widehat{\varphi}_{jk}\in[\phi,\phi+2\pi].\end{cases}$$

根据定理 1, $\widehat{\theta}_{ij}$ 和 $\widehat{\varphi}_{jk}$ 也是 $\{\theta_{ij}(t)\}$ 和 $\{\varphi_{jk}(t)\}$ 的全局最优解. 故对于任意 $\phi\in R$, 在 $[\phi,\phi+2\pi]$ 内存在 $\{\theta_{ij}(t)\}$ 和 $\{\varphi_{jk}(t)\}$ 的所有全局最优解. □

根据定理 2, $\{\theta_{ij}(t)\}$ 和 $\{\varphi_{jk}(t)\}$ 的全局最优解能够扩展到长度为 2π 的任意闭区间上. 这相当于 QNN 在全体实数域内均匀分布着众多全局吸引子, 从而可显著提高收敛率.

3.4　样本量子态描述

对实值样本 $\overline{X}=[\overline{x}_1,\overline{x}_2,\cdots,\overline{x}_n]^{\mathrm{T}}$, 定义如下量子态转换公式:

$$|X\rangle=[|x_1\rangle,|x_2\rangle,\cdots,|x_n\rangle]^{\mathrm{T}},\tag{11}$$

其中

$$|x_i\rangle=\Big[\cos\Big(\frac{2\pi}{1+\mathrm{e}^{-x_i}}\Big),\ \sin\Big(\frac{2\pi}{1+\mathrm{e}^{-x_i}}\Big)\Big]^{\mathrm{T}}.$$

4　算例仿真

为验证 QNN 算法的性能, 本文选取函数逼近进行仿真, 同时与相同结构的 CBP 作性能对比. 具体使用以下常用的三变量非线性函数:

$$y=\Big(1.0+\sqrt{x_1}+\frac{1}{x_2}+x_3^{-1.5}\Big)^2,\tag{12}$$

其中 x_1, x_2, x_3 取 $\{1,2,3,4,5\}$ 中的整数值. 本文选用其中的 20 组数据(如表 1 所示).

表 1　函数逼近仿真的样本数据

x_1	x_2	x_3	y	x_1	x_2	x_3	y
1	3	1	11.111	3	2	1	17.910
1	3	2	7.219 0	3	2	2	12.857
1	3	3	6.380 0	3	2	3	11.727
1	3	4	6.043 0	3	2	4	11.270
1	3	5	5.870 0	3	2	5	11.032
2	5	1	13.063	5	3	1	20.879
2	5	2	8.808 0	5	3	2	15.390
2	5	3	7.877 0	5	3	3	14.152
2	5	4	7.503 0	5	3	4	13.649
2	5	5	7.310 0	5	3	5	13.387

首先, 通过改变学习速率考察模型的逼近能力. 网络取 3-8-1 型, 迭代步数取 2 000, 学习速率 η 在集合 {0.1, 0.2, ⋯, 1.0} 中取值. 分别用 QNN 和 CBP 对每种学习速率仿真 100 次, 记录模型的最大误差和平均误差. 当学习速率改变时, QNN 平均误差的最大值为 0.026 7 ($\eta=0.1$), 最小值为 0.008 3 ($\eta=1.0$), 波动范围仅为 0.018 4. CBP 平均误差的最大值为 0.195 8 ($\eta=0.1$), 最小值为 0.066 0 ($\eta=1.0$), 波动范围高达 0.129 8. 因此, QNN 不仅具有较强的逼近能力, 同时对参数变化不敏感, 具有较好的鲁棒性能. 对比结果如图 3 所示.

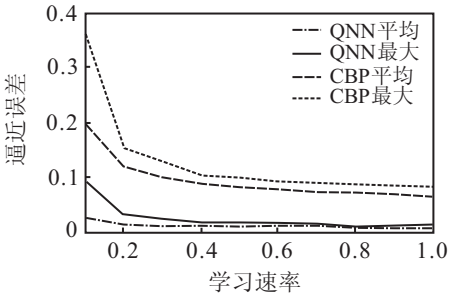

图 3　逼近误差与学习速率的关系

然后, 考察学习速率的变化对迭代步数的影响. 网络仍取 3-8-1 型, 限定误差取 0.05, 限定步数取 50 000, 学习速率 η 在集合 {0.1, 0.2, ⋯, 1.0} 中取值. 对每种学习速率, 分别用 QNN 和 CBP 仿真 100 次, 并计算平均迭代步数. 当学习速率改变时, QNN 平均迭代步数最多为 2 684.91 ($\eta=0.1$), 最少为 111.64 ($\eta=1.0$), 变化范围仅为 2 573.27. 而 CBP 最多为 29 846.60 ($\eta=0.1$), 最少为 3 767.81 ($\eta=0.9$), 变化范围高达

26 078.79. 这表明QNN不仅迭代步数明显少于CBP, 而且当学习速率改变时其波动范围很小. 对比结果如图4所示.

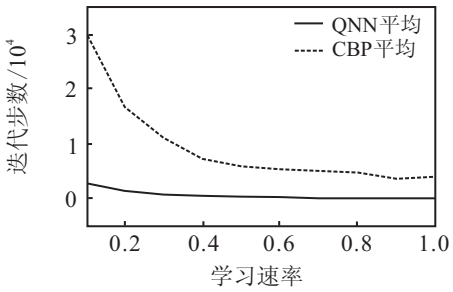

图 4 迭代步数与学习速率的关系

最后, 通过改变学习速率考察QNN的收敛能力. 隐层仍取8个节点, 限定误差取0.05, 限定步数取2 000, 学习速率η的取值范围同上. 对每种学习速率, 分别用QNN和CBP仿真100次, 并记录收敛次数. 当学习速率改变时, QNN收敛率均为100%. 而CBP当$\eta = 0.1$和$\eta = 0.2$时不收敛; 当$\eta > 0.2$时, 收敛率最高仅为64% ($\eta = 1.0$). 这表明QNN不仅收敛能力强, 而且当学习参数改变时具有较好的稳定性. 对比结果如图5所示.

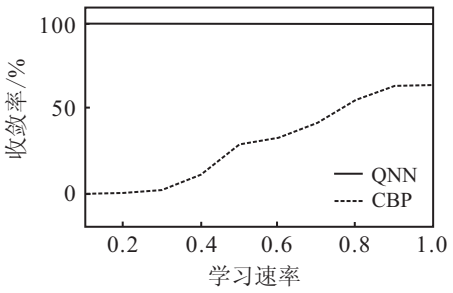

图 5 迭代收敛率与学习速率的关系

在以上仿真中, $\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的初值均在$[-\pi, \pi]$中随机选取. 根据定理2, 在$[-\pi, \pi]$中存在$\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的所有全局最优解, 即QNN在$[-\pi, \pi]$中存在众多吸引子, 因此有效提高了QNN的收敛率. 对于CBP, 所有初始权值均从$[-1, 1]$中随机选取, 由于全局最优解不会重复出现, 其数量远小于QNN, 同时, 单一的吸引子使迭代过程不易收敛.

在以下仿真中, $\{\theta_{ij}(t)\}$和$\{\varphi_{jk}(t)\}$的初值均在$[\pi, 3\pi]$中随机选取. CBP的所有初始权值均在$[1, 3]$中随机选取. 当$\eta = 0.9$时, 两种模型分别仿真100次, QNN的收敛率仍为100%, 而CBP的收敛率仅为37%. 仿真结果表明, QNN比CBP拥有更多的全局吸引子数目, 多重吸引子是本文QNN模型的主要特色, 正是多重全局吸引子的存在, 提高了QNN的收敛率, 这与定理2的结论是一致的. 综合以上仿真结果, QNN的收敛能力及鲁棒性能均比CBP有较大程度的提高.

5 结 论

本文从一个全新角度提出了一种基于量子门线路的QNN模型及算法. 不同于传统神经网络采用的加权、聚合、激励等方法, 在本文QNN模型中, 调整参数为量子旋转门的旋转角度, 各层输出由量子旋转门和多位受控非门共同决定. 仿真结果表明, 本文提出的QNN模型在收敛率、逼近能力、收敛速度和鲁棒性等方面均明显优于普通BP网络.

参考文献(References)

[1] Kak S. On quantum neural computing[J]. Information Sciences, 1995, 83(3): 143-160.

[2] Gopathy P, Nicolaos B. Quantum neural networks(QNNs) inherently fuzzy feedforward neural networks[J]. IEEE Trans on Neural Networks, 1997, 8(3): 679-693.

[3] Ventura D, Tony M. Quantum associative memory with exponential capacity[C]. Proc of the Int Joint Conf on Neural Networks. Alaska: Anchorage Press, 1998: 509-513.

[4] Ajit N, Tammy M. Quantum artificial neural network architectures and components[J]. Information Sciences, 2000, 128: 231-255.

[5] 解光军, 庄镇泉. 量子神经网络[J]. 计算机科学, 2001, 28(7): 1-6.
(Xie G J, Zhuang Z Q. Quantum neural networks[J]. Computer Science, 2001, 28(7): 1-6.)

[6] 解光军, 范海秋, 操礼程. 一种量子神经计算网络模型[J]. 复旦学报: 自然科学版, 2004, 43(5): 700-703.
(Xie G J, Fan H Q, Cao L C. A quantum neural computational networks model[J]. J of Fudan University: Natural Science, 2004, 43(5): 700-703.)

[7] 解光军, 周典, 范海秋. 基于量子门组单元的神经网络及其应用[J]. 系统工程理论与实践, 2005, 25(5): 113-117.
(Xie G J, Zhou D, Fan H Q. A quantum neural networks model based on quantum gates cell and its applications[J]. System Engineering-Theory & Practice, 2005, 25(5): 113-117.)

[8] Maeda M, Suenaga M, Miyajima H. Qubit neuron according to quantum circuit for XOR problem[J]. Applied Mathematics and Computation, 2007, 185(2): 1015-1025.

[9] 李盼池, 李士勇. 一种量子自组织特征映射网络模型及聚类算法[J]. 量子电子学报, 2007, 24(4): 463-468.
(Li P C, Li S Y. A quantum self-organization feature mapping networks model and clustering algorithm[J]. Chinese J of Quantum Electronics, 2007, 24(4): 463-468.)

(下转第151页)