

量子过程神经网络模型算法及应用*

李盼池[†] 王海英 戴庆 肖红

(东北石油大学计算机与信息技术学院, 大庆 163318)

(2011年9月30日收到; 2012年2月8日收到修改稿)

为提高过程神经网络的逼近和泛化能力, 从研究过程神经元信息处理的量子计算实现机理入手, 提出基于量子旋转门及多位受控非门的物理意义构造量子过程神经元的新思想. 将离散化后的过程式输入信息作为受控非门的控制位, 经过量子旋转门作用后控制目标量子位的状态, 以目标量子位处于状态 $|1\rangle$ 的概率幅作为量子过程神经元的输出. 以量子过程神经元为隐层, 普通神经元为输出层, 可构成量子过程神经网络. 基于量子计算机理推导了该模型的学习算法. 将该模型用于太阳黑子数年均值预测, 应用结果表明, 所提方法与普通过程神经网络相比, 预测精度有所提高, 对于复杂预测问题具有一定理论意义和实用价值.

关键词: 量子计算, 量子过程神经元, 量子过程神经网络, 算法设计

PACS: 03.65.-w

1 引言

长期的神经生理学实验和生物学研究表明, 生物神经系统的信息处理特征主要包括空间聚合、多因素聚合、时间累积效应、激励阈值特性、自适应性、兴奋与抑制特性、延时特性和传导与输出特性^[1]等8个方面. 由M-P神经元模型的定义可以看出, 传统人工神经元较好地模拟了生物神经元的空间加权聚合、自适应、传导和输出等大部分特性, 但缺乏对于时间延迟和累积效应以及多因素聚合作用的描述. 由于生物神经系统在实际信息处理过程中, 生物神经元的记忆和输出既依赖于各个输入信息的空间聚合作用, 同时也与时间的延迟和累积效应有关, 乃至与其他多因素聚合作用有关. 因此, 诞生于21世纪初, 由He等^[2]提出的过程神经网络(process neural networks, PNN)是迄今为止最能够模拟生物神经元这些重要信息处理特性的神经网络模型. 一般神经网络只能用来描述输入输出值之间的瞬时映射关系, 而过程神经网络则可描述输出对于输入在时间轴上的累积效应或

聚合效果, 它是传统神经元网络在时域上的扩展. 过程神经网络的研究内容丰富, 而且其理论和模型提出的时间都不长, 因此, 在模型构造和算法设计方面尚有很多问题有待于进一步深入研究和完善.

对于过程神经网络的训练问题, 文献[3, 4]给出了基于正交基展开的学习算法, 利用基函数的正交性, 有效简化了时域聚合运算. 在网络性能方面, 文献[5]研究了PNN的连续性、逼近能力、计算能力等理论性质. 在模型构造方面, 文献[6—12]先后提出了自组织、径向基、级联、双并联、分段、反馈、概率等多种PNN理论模型. 在上述所有模型中, PNN的输入均为时变的连续函数, 然而在许多实际问题中, 系统输入均为离散化的数据表. 目前的处理方法是先将离散化数据拟合成解析函数, 然后再对解析函数实施正交基函数展开^[13]. 这种处理方法的缺点, 一是解析函数的形式(即经验公式)不易确定; 二是采用高次多项式函数拟合时, 因存在龙格震荡现象会导致拟合失真; 三是正交基函数的个数选取尚无理论依据, 通常结合具体问题由多次试验确定. 上述缺点直接导致了PNN逼近

* 国家自然科学基金(批准号: 61170132)、中国博士后科学基金(批准号: 20090460864, 201003405)、黑龙江省博士后科学基金(批准号: LBH-Z09289)和黑龙江省教育厅科学基金(批准号: 11551015)资助的课题.

[†] E-mail: lipanchi@vip.sina.com

和泛化能力的下降.

目前,将量子计算与神经网络融合的量子神经网络是当前人工神经网络理论研究的前沿课题,已引起国内外学者的高度重视,并已提出一些有意义的理论模型^[14,15].为了避免PNN处理离散输入时的拟合及正交基展开运算,并有效提高其泛化能力,本文提出将量子计算与过程神经网络相融合的新思想.采用多位受控非门中目标量子位的循环演化模拟PNN的时域聚合运算,基于量子旋转门和受控非门的物理意义设计了量子过程神经网络(quantum process neural networks, QPNN)模型和算法.实验结果表明,与PNN相比,QPNN的逼近和泛化能力均有所提高.

2 量子比特及量子门概述

2.1 量子比特

在量子计算中,量子比特可用量子态波函数 $|\phi(t)\rangle$ 描述.在固定时刻 $t \in [0, T]$,其中 T 为时间(或过程)的长度,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其状态可以落在 $|0\rangle$ 和 $|1\rangle$ 之外,即可以处于下式所示的线性叠加态:

$$|\phi(t)\rangle = \alpha(t)|0\rangle + \beta(t)|1\rangle, \quad (1)$$

其中 $\alpha(t), \beta(t)$ 称为量子态概率幅.量子比特也可用概率幅表示为 $|\phi(t)\rangle = [\alpha(t), \beta(t)]^T$.

2.2 单比特量子旋转门

量子门是实现量子计算的基础,它包含了量子计算的特点.单比特量子旋转门定义为

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

令 $|\phi(t)\rangle = \begin{bmatrix} \cos(\theta_0(t)) \\ \sin(\theta_0(t)) \end{bmatrix}$,由 $\mathbf{R}(\theta(t))$ 对 $|\phi(t)\rangle$ 的作用 $\mathbf{R}(\theta(t))|\phi(t)\rangle = \begin{bmatrix} \cos(\theta_0(t) + \theta(t)) \\ \sin(\theta_0(t) + \theta(t)) \end{bmatrix}$ 知,其作用为实现对 $|\phi(t)\rangle$ 的相位旋转.

2.3 多比特受控非门

在真实的量子系统中,单个量子比特的

状态往往同时受多位量子比特 $|x_i(t)\rangle$ 的联合控制,多比特受控非门 $C^n(\mathbf{X}(t))$ 是这种控制方式的模型.多比特量子系统可用波函数 $|x_1(t)x_2(t)\cdots x_n(t)\rangle$ 来描述.在 $n+1$ 比特量子系统 $|x_1(t)x_2(t)\cdots x_n(t)\rangle|\phi(t)\rangle$ 中,当目标比特 $|\phi(t)\rangle$ 同时受 n 位输入比特 $|x_1(t)x_2(t)\cdots x_n(t)\rangle$ 的联合控制时,系统的动态特性可用多比特受控非门 $C^n(\mathbf{X}(t))$ 模型描述,如图1所示.

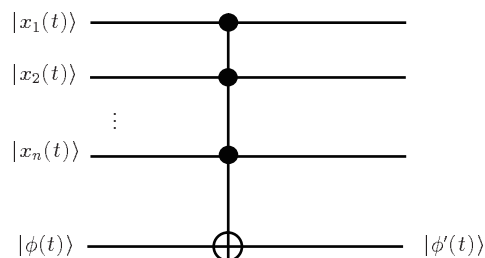


图1 多比特受控非门

设在 $t \in [0, T]$ 时刻, $n+1$ 位量子系统的状态为 $|x_1(t)x_2(t)\cdots x_n(t)\rangle|\phi(t)\rangle$, $\mathbf{X}(t)$ 是单比特非门,则多比特受控非门 $C^n(\mathbf{X}(t))$ 的运算关系可定义为

$$\begin{aligned} & C^n(\mathbf{X}(t))|x_1(t)x_2(t)\cdots x_n(t)\rangle|\phi(t)\rangle \\ &= |x_1(t)x_2(t)\cdots x_n(t)\rangle \\ & \quad \times \mathbf{X}(t)^{x_1(t)x_2(t)\cdots x_n(t)}|\phi(t)\rangle, \end{aligned} \quad (3)$$

即若前 n 个控制比特全为1,则非门 $\mathbf{X}(t)$ 作用到目标比特使其翻转,否则非门不起作用.

令 n 个控制比特为 $|x_i(t)\rangle = a_i|0\rangle + b_i|1\rangle$,目标比特为 $|\phi(t)\rangle = c|0\rangle + d|1\rangle$,根据(3)式,多比特受控非门的输出为

$$\begin{aligned} & C^n(\mathbf{X}(t))|x_1(t)x_2(t)\cdots x_n(t)\rangle|\phi(t)\rangle \\ &= |x_1(t)\rangle \otimes |x_2(t)\rangle \otimes \cdots \otimes |x_n(t)\rangle \otimes |\phi(t)\rangle \\ & \quad - b_1 b_2 \cdots b_n c \underbrace{|11 \cdots 10\rangle}_{n \uparrow} \\ & \quad + b_1 b_2 \cdots b_n d \underbrace{|11 \cdots 10\rangle}_{n \uparrow} \\ & \quad + b_1 b_2 \cdots b_n c \underbrace{|11 \cdots 11\rangle}_{n \uparrow} \\ & \quad - b_1 b_2 \cdots b_n d \underbrace{|11 \cdots 11\rangle}_{n \uparrow}. \end{aligned} \quad (4)$$

由此可知,多比特受控非门 $C^n(\mathbf{X}(t))$ 的输出处于 $n+1$ 个量子比特的纠缠态中,且目标比特 $|\phi'(t)\rangle$

处于状态 $|1\rangle$ 的概率为

$$P = (b_1 b_2 \cdots b_n)^2 (c^2 - d^2) + d^2. \quad (5)$$

此时, 经 n 个输入比特联合控制后, 目标比特可定义为

$$|\phi'(t)\rangle = \sqrt{1-P}|0\rangle + \sqrt{P}|1\rangle. \quad (6)$$

3 量子过程神经网络模型

3.1 量子过程神经元模型

本文提出的量子过程神经元由量子旋转门和多比特受控非门组成, 如图 2 所示. 其输入为定义在时域区间 $[0, T]$ 上的量子态波函数 $|x_i(t)\rangle$, 输出为量子过程神经元在 $[0, T]$ 上的聚合结果 $|y\rangle$, 控制参数为量子旋转门 $R(\bar{\theta}_i(t))$ 的旋转角度 $\bar{\theta}_i(t)$, $i = 1, 2, \dots, n$.

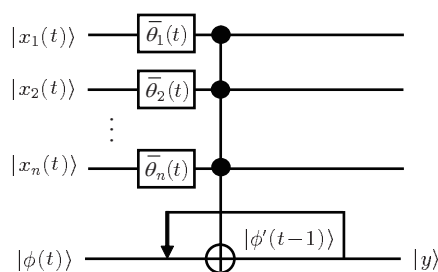


图 2 量子过程神经元

设时域区间 $[0, T]$ 上的采样时刻为 $0 = t_1 < t_2 < \cdots < t_q = T$, 则 $|x_i(t)\rangle$ 在 $t \in [0, T]$ 上可离散化为

$$|x_i(t_r)\rangle = \cos \theta_i(t_r)|0\rangle + \sin \theta_i(t_r)|1\rangle, \quad (7)$$

其中 $1 \leq r \leq q$.

设 $|\phi(t_1)\rangle = |0\rangle$, 根据量子旋转门和多比特受

控非门的运算关系得

$$|\phi'(t_1)\rangle = \cos \varphi(t_1)|0\rangle + \sin \varphi(t_1)|1\rangle, \quad (8)$$

其中

$$\varphi(t_1) = \arcsin \left(\prod_{i=1}^n \sin(\theta_i(t_1) + \bar{\theta}_i(t_1)) \right).$$

令 $t = t_r, 1 < r \leq q$, 由 $|\phi(t)\rangle = |\phi'(t-1)\rangle$, 得 $|\phi(t_r)\rangle = |\phi'(t_{r-1})\rangle$, 根据量子旋转门和多比特受控非门运算关系得

$$|\phi'(t_r)\rangle = \cos \varphi(t_r)|0\rangle + \sin \varphi(t_r)|1\rangle, \quad (9)$$

其中

$$\varphi(t_r) = \arcsin \left(\sqrt{S_r \cos 2\varphi(t_{r-1}) + \sin^2 \varphi(t_{r-1})} \right),$$

$$S_r = \prod_{i=1}^n \sin^2(\theta_i(t_r) + \bar{\theta}_i(t_r)).$$

量子过程神经元在 $[0, T]$ 上的聚合结果为

$$|y\rangle = |\phi'(t_q)\rangle = \cos \varphi(t_q)|0\rangle + \sin \varphi(t_q)|1\rangle, \quad (10)$$

其中

$$\varphi(t_q) = \arcsin \left(\sqrt{S_q \cos 2\varphi(t_{q-1}) + \sin^2 \varphi(t_{q-1})} \right),$$

$$S_q = \prod_{i=1}^n \sin^2(\theta_i(t_q) + \bar{\theta}_i(t_q)).$$

本文规定以 $|y\rangle$ 处于状态 $|1\rangle$ 的概率幅作为量子过程神经元的输出, 因此, 输出为

$$y = \sqrt{S_q \cos 2\varphi(t_{q-1}) + \sin^2 \varphi(t_{q-1})}. \quad (11)$$

3.2 量子过程神经网络模型

本文提出的量子过程神经网络模型如图 3 所示. 其中 $|x_1(t)\rangle, |x_2(t)\rangle, \dots, |x_n(t)\rangle$ 为输入, h_1, h_2, \dots, h_p 为隐层输出, y_1, y_2, \dots, y_m 为网络输出. 隐层由量子过程神经元组成, 输出层由普通神经元组成. 输出层激励采用 sigmoid 函数.

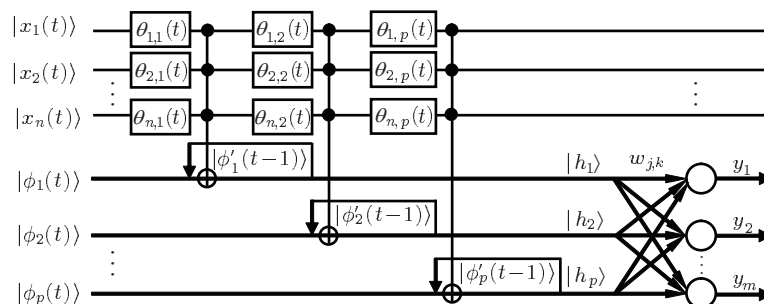


图 3 量子过程神经网络

设图 3 所示 QPNN 模型中, 输入量子位 $|x_i(t) = \cos \theta_i(t)|0\rangle + \sin \theta_i(t)|1\rangle$, 时域采样区间为 $[0, T]$, 采样时刻为 $0 = t_1 < t_2 < \dots < t_q = T$, $|\phi_j(t_1) = |0\rangle$, 输出层权值为 w_{jk} .

记 $\bar{h}_{j,r} = \prod_{i=1}^n \sin(\theta_i(t_r) + \theta_{i,j}(t_r))$, 根据量子过程神经元的输入输出关系, 隐层第 j 个量子过程神经元在 $[0, t_r]$ 上的聚合结果为

$$h_j(t_1) = \bar{h}_{j,1}, \quad (12)$$

$$h_j(t_r) = \sqrt{\bar{h}_{j,r}^2(1 - 2h_j^2(t_{r-1})) + h_j^2(t_{r-1})}, \quad (13)$$

该神经元的输出 (即 $[0, T]$ 上的聚合结果) 为

$$h_j = h_j(t_q), \quad (14)$$

输出层第 k 个神经元输出为

$$y_k = \frac{1}{1 + e^{(-\sum_{j=1}^p w_{j,k} h_j)}}. \quad (15)$$

以上各式中, $1 \leq i \leq n, 1 \leq j \leq p, 1 \leq k \leq m$.

4 量子过程神经网络算法

4.1 样本离散化及量子态描述

设时域采样区间为 $[0, T]$, 采样时刻为 $0 = t_1 < t_2 < \dots < t_q = T$. 对 n 维欧氏空间中的过程式样本 $\bar{\mathbf{X}}(t) = (\bar{x}_1(t), \bar{x}_2(t), \dots, \bar{x}_n(t))^T$, 首先将其离散化为

$$\bar{\mathbf{X}}(t_r) = (\bar{x}_1(t_r), \bar{x}_2(t_r), \dots, \bar{x}_n(t_r))^T, \quad (16)$$

其中 $1 \leq r \leq q$.

为进行样本量子态描述, 记

$$M_{\max i} = \max(\bar{x}_i(t_1), \bar{x}_i(t_2), \dots, \bar{x}_i(t_q)), \quad (17)$$

$$M_{\min i} = \min(\bar{x}_i(t_1), \bar{x}_i(t_2), \dots, \bar{x}_i(t_q)), \quad (18)$$

$$\theta_{ir} = \frac{\bar{x}_i(t_r) - M_{\min i}}{M_{\max i} - M_{\min i}} \frac{\pi}{2}, \quad (19)$$

其中 $1 \leq i \leq n$. 该样本可用如下定义的量子态向量来描述:

$$|X(t_r)\rangle = [|x_1(t_r)\rangle, |x_2(t_r)\rangle, \dots, |x_n(t_r)\rangle]^T, \quad (20)$$

其中 $|x_i(x_r)\rangle = \cos(\theta_{ir})|0\rangle + \sin(\theta_{ir})|1\rangle$.

4.2 网络各层参数的调整

在图 3 模型中, 可调参数为隐层量子旋转门的旋转角度和输出层权值. 设归一化后的期望输出为 $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$, 定义误差函数为

$$E = \frac{1}{2} \sum_{k=1}^m (\bar{y}_k - y_k)^2, \quad (21)$$

记 $\bar{h}_{j,r} = \prod_{i=1}^n \sin(\theta_i(t_r) + \theta_{i,j}(t_r))$, $h_j(t_1) = \bar{h}_{j,1}$, $S_{jr} = \bar{h}_{j,r}^2(1 - 2h_j^2(t_{r-1}))$. 根据梯度下降法, 隐层旋转角度 $\theta_{i,j}(t_r)$ 的梯度可按下列各式计算

$$-\frac{\partial E}{\partial h_j(t_q)} = \sum_{k=1}^m (\bar{y}_k - y_k) y_k (1 - y_k) w_{j,k}, \quad (22)$$

$$\frac{\partial h_j(t_r)}{\partial h_j(t_{r-1})} = (1 - 2\bar{h}_{j,r}^2) h_j(t_{r-1}) / h_j(t_r), \quad (23)$$

$$\frac{\partial h_j(t_r)}{\partial \theta_{i,j}(t_r)} = S_{jr} \cot(\theta_i(t_r) + \theta_{i,j}(t_r)) / h_j(t_r), \quad (24)$$

综合以上三式可得

$$\begin{aligned} & -\frac{\partial E}{\partial \theta_{i,j}(t_r)} \\ &= -\frac{\partial E}{\partial h_j(t_q)} \dots \frac{\partial h_j(t_{r+1})}{\partial h_j(t_r)} \frac{\partial h_j(t_r)}{\partial \theta_{i,j}(t_r)} \\ &= \sum_{k=1}^m (\bar{y}_k - y_k) y_k (1 - y_k) w_{j,k} \\ & \quad \times \prod_{s=r+1}^q (1 - 2\bar{h}_{j,s}^2) h_j(t_{s-1}) \\ & \quad \times \bar{h}_{j,r}^2 (1 - 2h_j^2(t_{r-1})) \\ & \quad \times \cot(\theta_i(t_r) + \theta_{i,j}(t_r)) / \prod_{s=r}^q h_j(t_s), \end{aligned} \quad (25)$$

其中 $1 \leq r \leq q$. 输出层权值 $w_{j,k}$ 的调整量计算式为

$$-\frac{\partial E}{\partial w_{j,k}} = (\bar{y}_k - y_k) y_k (1 - y_k) h_j(t_q), \quad (26)$$

综上, 网络参数可按如下两式调整

$$\theta_{i,j}^{(s+1)}(t_r) = \theta_{i,j}^{(s)}(t_r) - \alpha \frac{\partial E}{\partial \theta_{i,j}(t_r)}, \quad (27)$$

$$w_{j,k}^{(s+1)} = w_{j,k}^{(s)} - \beta \frac{\partial E}{\partial w_{j,k}}, \quad (28)$$

其中 s 为迭代步数, α, β 为学习速率, $1 \leq i \leq n, 1 \leq j \leq p, 1 \leq k \leq m, 1 \leq r \leq q$.

5 实际应用

5.1 丙烯酰胺均相聚合反应过程建模

丙烯酰胺均相聚合反应是化学工业中一类重要的化学反应, 在这类反应的诸多影响因素中, 温度是影响生成物浓度的一个重要因素. 为了得到温度变化与生成物浓度之间的映射关系, 传统方法是借助于微分动力学方程和热力学方程. 但由于

物理化学模型的复杂性, 往往难以找到两者之间的真实关系, 而 QPNN 恰好适合该问题求解. 表 1 列出了丙烯酰胺均相聚合反应实验的部分结果, 其中 t 表示从 0 时刻开始以分钟计的累计时间, $T_i (i = 1, 2, \dots, 9)$ 表示第 i 组实验的温度 (单位为 $^{\circ}\text{C}$) 测试结果, N_m 为每组实验对应生成物的分子数 (单位为 $10^6/\text{cm}^3$), 即丙烯酰胺均相聚合反应生成物的浓度.

表 1 9 组聚合化学反应实验结果

t/min	$T_1/^{\circ}\text{C}$	$T_2/^{\circ}\text{C}$	$T_3/^{\circ}\text{C}$	$T_4/^{\circ}\text{C}$	$T_5/^{\circ}\text{C}$	$T_6/^{\circ}\text{C}$	$T_7/^{\circ}\text{C}$	$T_8/^{\circ}\text{C}$	$T_9/^{\circ}\text{C}$
0	17	18	18	18	14	15	15	14	13
10	19	16	21	18	15	16	16	15	15
20	23	16	23	19.5	17	16	18	16	16
30	27	17	26	21.5	20	16.5	21	17.5	18
40	30	18	31	24.5	22	18	25	19.5	19
50	34	19	36	29.5	25	20	30	22	21
60	39	20	43.5	34	28	22	36	24	23
70	45	21	52.5	38	33	24	40	26.5	25
80	52	22.5	59	46	38	29	54	30.5	27
90	64	24	68	56	46	32	64	35.5	30
100	73	26	73	63	52	36	73	43	34
110	81	28	77	74	60	40	76	54.5	41
120	82	32	79	78	66	44.5	77	67	44
130	84	36	83	93	72	50.5	78	79.5	49
140	88	45	87	95	76	57.5	79	83	54
150	92	54	96	97	81	66.5	84	87	65
160	101	64	98	100	83	67	86	93	77
170	113	68	102	110	94	69	89	98	80
$N_m/10^6 \text{ cm}^{-1}$	1811	1838	1894	1853	1496	1701	1754	1915	1644

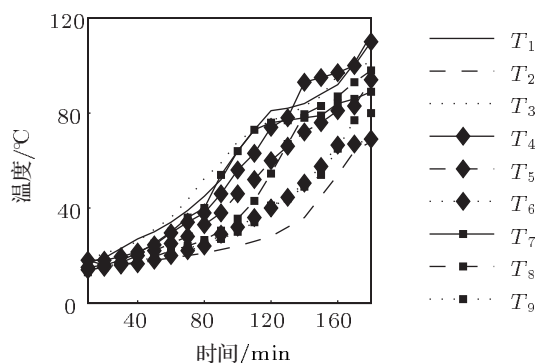


图 4 丙烯酰胺聚合反应温度时间变化图

图 4 给出了描述丙烯酰胺均相聚合反应实验温度随时间变化的 9 组曲线, 对于 QPNN 而言, 每条曲线对应一个与时间相关的输入序列, 而生成物的分子数为 QPNN 的输出.

本实验分别用 QPNN 和 PNN 建模. 由问题性质可知, 网络输入和输出均为 1 个节点, 离散采样点数为 18. 为检验两种网络的逼近能力, 隐层均分别取 5, 6, \dots , 20 个节点, 归一化后的最大绝对误差均限定为 10^{-3} , 迭代步数均限定为 1000. QPNN 隐层旋转角度初始化为 $(-\pi/2, \pi/2)$ 之间的随机数, 输出层权值初始化为 $(-1, 1)$ 之间的随

机数. 在 PNN 参数设置中, 拟合采用拉格朗日多项式函数, 基展开采用 Walsh 正交基函数, 基函数个数为 32, 两层权值均初始化为 $(-1, 1)$ 之间的随机数. 为评价两种模型的逼近能力, 本文定义评价指标 F 为最大绝对误差与迭代步数的乘积. 显然绝对误差越小, 迭代步数越少, 逼近能力就越好. 本实验中, 隐层节点数 p 为各种取值下的迭代步数 N 和评价指标 F 对比如图 5 和图 6 所示.

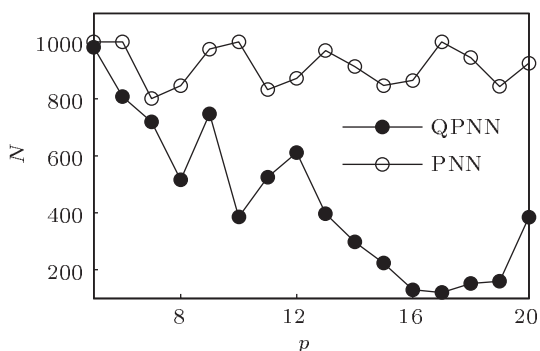


图 5 QPNN 和 PNN 的迭代步数对比

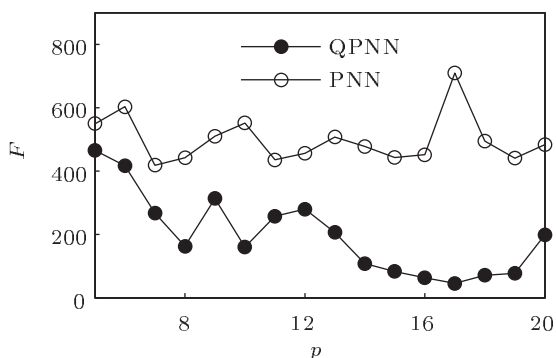


图 6 QPNN 和 PNN 的逼近能力对比

实验结果表明, 对于隐层节点的各种取值,

QPNN 的逼近能力比 PNN 均有明显提高.

5.2 太阳黑子数年均值预测

本实验采用太阳黑子数作为实验数据, 采用 QPNN 和 PNN 实现太阳黑子数时间序列的非线性预测.

预测方案 本实验采用连续 259 年 (1749—2007 年) 太阳黑子月数据和年数据. 月数据指一个月的太阳黑子平均数, 年数据指全年的太阳黑子平均数. 预测方案为用前 n 年的月数据预测第 $n+1$ 年的年数据. 为合理确定 n 的取值, 本文采取计算太阳黑子第 $k-k+n$ 年和第 $k+1-k+n+1$ 年数据之间自相关系数的方法. 计算结果如图 7 所示.

由图 7 可知, $n \geq 10$ 时相关系数接近最大值. 故预测方案为采用前 10 年月数据预测第 11 年的年数据.

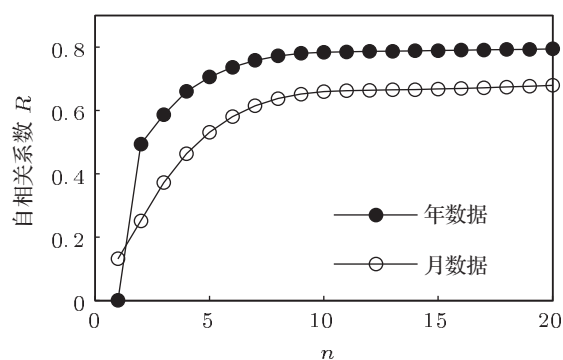


图 7 太阳黑子数序列的自相关图

样本构造 太阳黑子数序列的原始数据如表 2 所示. 根据预测方案, 可由表 2 构造出用于网络训练和预测的样本数据, 数据模式如表 3 所示.

表 2 太阳黑子数序列的原始数据

年份	月数据												年数据
	1	2	3	4	5	6	7	8	9	10	11	12	
1749	58.0	62.6	70.0	55.7	85.0	83.5	94.8	66.3	75.9	75.5	158.6	85.2	80.9
1750	73.3	75.9	89.2	88.3	90.0	100.0	85.4	103.0	91.2	65.7	63.3	75.4	83.4
1751	70.0	43.5	45.3	56.4	60.7	50.7	66.3	59.8	23.5	23.2	28.5	44.0	47.7
1752	35.0	50.0	71.0	59.3	59.7	39.6	78.4	29.3	27.1	46.6	37.6	40.0	47.8
1753	44.0	32.0	45.7	38.0	36.0	31.7	22.0	39.0	28.0	25.0	20.0	6.7	30.7
...
2007	16.8	10.7	4.5	3.4	11.7	12.1	9.7	6.0	2.4	0.9	1.7	10.1	7.5

表3 太阳黑子数序列的样本数据模式

序号	输入年份(月数据)										预测年份 (年数据)
	$x_1(t)$	$x_2(t)$	$x_3(t)$	$x_4(t)$	$x_5(t)$	$x_6(t)$	$x_7(t)$	$x_8(t)$	$x_9(t)$	$x_{10}(t)$	
1	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759
2	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760
3	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761
4	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762
5	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763
...
249	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007

网络结构 根据预测方案, QPNN 和 PNN 的输入节点为 10 个, 离散采样点数为 12, 输出节点为 1 个. 为充分对比 QPNN 和 PNN 的逼近和泛化能力, 隐层节点取 5—20 个. 对于 PNN, 离散样本采用拉格朗日多项式函数拟合, 并用 16 个 Walsh 基函数做正交基展开处理. 对于两种网络的训练过程, 样本归一化后的最大绝对误差均限定为 10^{-5} , 迭代步数均限定为 1000.

训练过程 在所有样本中, 选取前 180 年 (1749—1928 年) 的数据用于训练, 使其逼近太阳黑子数序列在不同时段之间的映射关系, 后 79 年 (1929—2007 年) 的数据用于测试网络的泛化能力. 定义评价指标 F 为最大绝对误差与迭代步数的乘积, QPNN 和 PNN 在隐层节点各种取值下的评价指标 F 对比如图 8 所示.

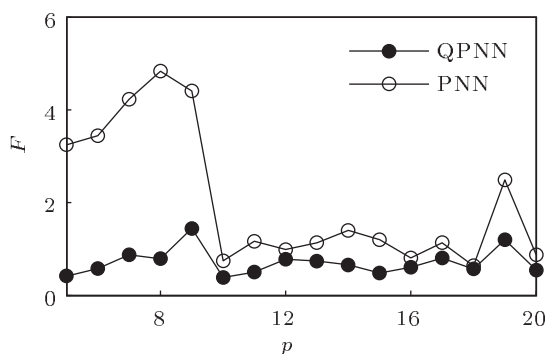


图8 QPNN 和 PNN 的逼近能力对比

预测过程 用 1929—2007 年的样本数据, 对训练好的网络进行测试, 以检验其泛化能力. QPNN 和 PNN 在隐层节点数的各种取值下, 预测结果的最大绝对误差 E 和方差 σ 对比如图 9 和图 10 所示. 以 15 个隐层节点为例, QPNN 和 PNN 的预测结果对比如图 11 所示. 预测结果表明 QPNN 的泛

化能力优于 PNN.

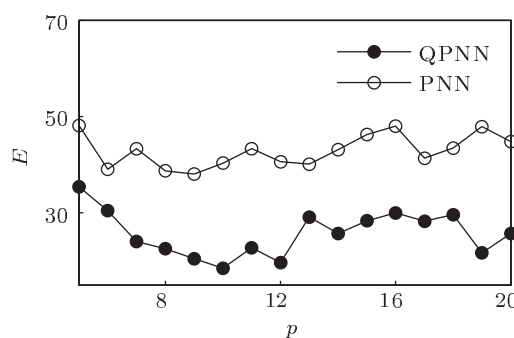


图9 QPNN 和 PNN 的绝对误差对比

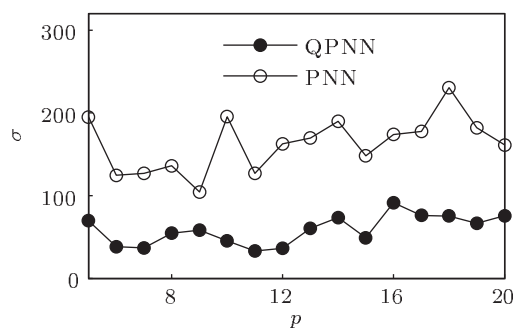


图10 QPNN 和 PNN 的方差对比

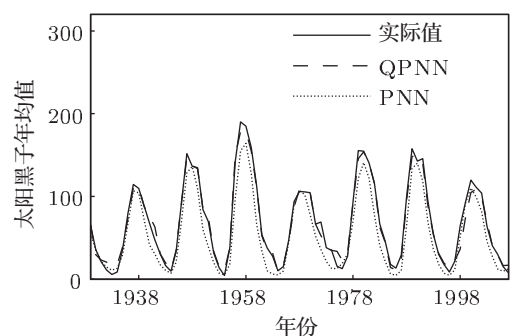


图11 QPNN 和 PNN 的预测结果对比

以上两个实验结果表明, QPNN 的逼近能力和预测能力比 PNN 均有所提高. 对此结果可做如下分析. QPNN 和 PNN 对于时域聚合运算采取了不同的处理方式. QPNN 直接接收离散输入序列, 采用量子信息处理机理将其循环映射为隐层量子受控非门的输出, 由于受控非门的输出呈现为多个量子位的纠缠态, 因此这种映射具有高度非线性, 从而使量子过程神经元的时域聚合结果具有更强的模式记忆能力. 对于 PNN, 由于不能直接处理离散输入, 而需要对离散样本进行拟合, 从而存在拟合误差. 在随后的正交基展开中, 理论上基函数的个数应为无限个, 有限项的基函数展开又必然会带来截断误差, 因此在 PNN 信息处理过程中不可避免地存在模式特征的丢失. 从而在网络结构和参数设置相似的情况下, QPNN 的逼近能力和泛化能力均

优于 PNN.

6 结论

本文提出了一种基于量子计算原理的量子过程神经网络模型和算法. 实验结果揭示出, 借助多位量子受控非门的信息处理机理可有效处理普通过程神经网络中的时域聚合运算, 并可避免普通过程神经网络由于离散样本拟合及样本函数的正交基展开运算带来的信息丢失, 从而可提高网络的逼近及泛化能力. 实际应用效果验证了提出模型和算法的有效性. 关于量子过程神经网络的连续性、逼近能力、泛化能力、计算复杂度等若干理论性质是值得进一步深入研究的问题.

-
- [1] Tsoi A C 1994 *IEEE Trans. Neural Networ.* **7** 229
- [2] He X G, Liang J Z 2000 *Proceedings of the 16th World Computer Conferences on Intelligent Information Processing* Beijing, China, August 12–15, 2000 p143
- [3] He X G, Liang J Z 2000 *Eng. Sci.* **2** 40 (in Chinese) [何新贵, 梁久祯 2000 中国工程科学 **2** 40]
- [4] He X G, Liang J Z, Xu S H 2001 *Eng. Sci.* **3** 31 (in Chinese) [何新贵, 梁久祯, 许少华 2001 中国工程科学 **3** 31]
- [5] Xu S H, He X G, Liu K 2006 *Acta Electron. Sin.* **34** 1838 (in Chinese) [许少华, 何新贵, 刘坤 2006 电子学报 **34** 1838]
- [6] Xu S H, He X G, Li P C 2003 *J. Comput. Res. Dev.* **40** 1612 (in Chinese) [许少华, 何新贵, 李盼池 2003 计算机研究与发展 **40** 1612]
- [7] Xu S H, He X G 2004 *J. BUAA* **30** 14 (in Chinese) [许少华, 何新贵 2004 北京航空航天大学学报 **30** 14]
- [8] Xu S H, He X G 2004 *Pattern Recogn. Artif. Intell.* **17** 201 (in Chinese) [许少华, 何新贵 2004 模式识别与人工智能 **17** 201]
- [9] Zhong S S, Ding G 2005 *Control Decis.* **20** 764 (in Chinese) [钟诗胜, 丁刚 2005 控制与决策 **20** 764]
- [10] Liang J Z 2006 *Pattern Recogn. Artif. Intell.* **19** 295 (in Chinese) [梁久祯 2006 模式识别与人工智能 **19** 295]
- [11] He X G, Xu S H 2004 *Acta Automatica Sin.* **30** 801 (in Chinese) [何新贵, 许少华 2004 自动化学报 **30** 801]
- [12] Xu S H, Li P C, He X G 2009 *CAAI Trans. Intell. Syst.* **4** 283 (in Chinese) [许少华, 李盼池, 何新贵 2009 智能系统学报 **4** 283]
- [13] Ding G, Zhong S S 2007 *Acta Phys. Sin.* **56** 1224 (in Chinese) [丁刚, 钟诗胜 2007 物理学报 **56** 1224]
- [14] Michiharu M, Masaya S, Hiromi M 2007 *Appl. Math. Comput.* **185** 1015
- [15] Li P C, Song K P, Yang E L 2010 *Neural Netw. World* **20** 189

Quantum process neural networks model algorithm and applications*

Li Pan-Chi[†] Wang Hai-Ying Dai Qing Xiao Hong

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

(Received 30 September 2011; revised manuscript received 8 February 2012)

Abstract

To enhance the approximation and generalization ability of process neural networks (PNNs), by studying the quantum implementation mechanism of information processing of process neuron, a new idea of designing quantum process neuron is proposed in this paper, based on the quantum rotation gates and the quantum controlled-non gates. In the proposed approach, the discrete process inputs are expressed by the qubits, which, as the control qubits of controlled-non gates after being rotated by the quantum rotation gates, control the target qubits to reverse. The model outputs are described by the probability amplitude of state $|1\rangle$ in the target qubits. Then the quantum process neural networks (QPNNs) are designed by the quantum process neurons for the hidden layer and the normal neurons for the output layer. The algorithm of QPNN is derived through the quantum computing. The proposed approach is utilized to predict the smoothed yearly mean sunspot numbers, and the results indicate that the QPNN has higher prediction accuracy than the normal PNN, thus it has a certain theoretical meaning and practical value for the complex prediction.

Keywords: quantum computation, quantum process neuron, quantum process neural networks, algorithm design

PACS: 03.65.—w

* Project supported by the National Natural Science Foundation of China (Grant No. 61170132), the China Postdoctoral Science Foundation (Grant Nos. 20090460864, 201003405), the Postdoctoral Science Foundation of Heilongjiang Province, China (Grant No. LBH-Z09289), and the Scientific Research Foundation of the Education Department of Heilongjiang Province, China (Grant No. 11551015).

[†] E-mail: lipanchi@vip.sina.com