

变结构神经网络*

毛宗源 姚尹武

胡尚信

(华南理工大学自动控制工程系·广州, 510640) (华东船舶工业学院自动化系·镇江, 212000)

摘要: 针对神经网络结构、性能不佳的普通状况, 本文提出变结构的神经网络, 先后研究了神经网络单隐层和多隐层的变化情况, 仿真试验结果证明是有效的。

关键词: 变结构神经网络; 单隐层; 多隐层; 层数; 维数

The Variable Structure Neural Network

Mao Zongyuan and Yao Yinwu

(Department of Automatic Control Engineering, South China University of Technology, Guangzhou, 510640, P. R. China)

Hu Shangxin

(Department of Automation, East China Ship industrial College, Zhenjiang, 212000, P. R. China)

Abstract: In this paper, the variable structure neural network is proposed for the reason that the network structure and performance there are universality not optimum. We have studied that the variable structure neural network for single hidden layer and multi-hidden layer. The simulation experiment results demonstrate that the variable structure neural networks are effective.

Key words: the variable structure neural network; single hidden layer; multi-hidden layer; layer number; nodal point number

1 引言(Introduction)

文献[1]成功设计了基于神经网络的自组织模糊控制方法, 然而作为其中关键部分的神经网络目前仍然存在不尽人意的地方, 常常发现它的性能并不是那么理想, 竟然训练不下去! 而且, 总会发现选定一个合适结构的神经网络是件相当头疼的事, 用变结构的神经网络可能是解决问题的有效途径。

所谓变结构的神经网络指的是在学习过程中, 隐层的层数和维数不断变化, 结构呈现为动态变化的神经网络, 实现变结构神经网络的目的在于使网络性能在训练中始终保持最优, 最终达到最佳结构。本文研究了神经网络单隐层和多隐层层数及每层维数变化的实现算法。

2 变结构神经网络的实现步骤(Achieving steps for variable structure neural network)

变结构神经网络的实现步骤如下:

1) 网络初始化: 给定任意初始结构, 将各连接权值 $W_{l,j,k}$ 随机置到小的随机数, 给定精度和学习次数, 并设定学习速度 η , 一般取 0.01 到 1 之间。

2) 输入一组训练样本 $X_{o,k,p}$ 。

3) 计算网络各层的输出

$$\begin{cases} Y_{o,k,p} = x_{o,k,p}, \\ Y_{l,k,p} = f\left(\sum_{j=1}^{N(l-1)} W_{l,j,k} \cdot Y_{l-1,j,p}\right). \end{cases} \quad (1)$$

4) 计算各层输出结点的误差信号

$$\begin{cases} \delta_{l-1,k,p} = (Y_{l-1,k,p}^d - Y_{l-1,k,p}) Y_{l-1,k,p} (1 - Y_{l-1,k,p}), \\ \delta_{l,k,p} = \sum_{i=1}^{N(l+1)} \delta_{l+1,i,p} W_{l+1,i,k} \cdot Y_{l,k,p} (1 - Y_{l,k,p}). \end{cases} \quad (2)$$

5) 再输入一组训练样本, 返回 3), 直到训练样本用完为止。

6) 计算误差 $E(W)$, 是否达到精度要求? 否, 继续; 是, 训练结束。

$$E(W) = \frac{1}{2} \sum_{p=1}^P \sum_{k=1}^{N(l-1)} (Y_{l-1,k,p} - Y_{l-1,k,p}^d)^2. \quad (3)$$

式中 $Y_{l-1,k,p}$, $Y_{l-1,k,p}^d$ 分别为输出层第 k 结点的实际输出和期望输出, P 为网络训练样本数。

7) 学习次数是否小于设定次数? 否, 继续; 是, 训练结束。

8) 计算各连接权值的调节量

* 广东省自然科学基金(960304)资助项目。

本文于 1996 年 10 月 25 日收到, 1997 年 9 月 17 日收到修改稿。

$$\Delta W_{l,j,k} = \eta \sum_{p=1}^P \delta_{l,k,p} \cdot Y_{l-1,j,p} \tag{4}$$

即第 $l-1$ 层第 j 结点与第 l 层第 k 结点,连接权值调
节量.

9) 计算连接权值: $W_{l,j,k} = W_{l,j,k} + \Delta W_{l,j,k}$.

10) 根据神经网络的动态算法重新确定神经网络
结构与连接权值.

11) 返回 2).

3 三层神经网络的变结构训练实例 (Take
three layers variable structure neural network
learning for example)

对于 BP 网有一个重要定理,即 Kolmogorol 定
理^[2]:

给定任意学习率 $\zeta > 0$ 和任意函数 $f: [0,1]^n \subset$
 $\mathbb{R}_n \rightarrow \mathbb{R}_m$, 存在一个三层 BP 网络,它可以在任意精度
的平方误差内逼近 f .

对于只有一个隐层的 BP 网络来说,实现变结
构只需要变隐层的结点数,于是,问题归结在选择隐
层的维数大小上,其算法在文献[1]中作了详细论
述,本文仅仅举一个三层变结构神经网络的训练
实例.

进行模糊推理与实现自组织算法的变结构 BP
网络起始结构定义为 1-2-1,学习率取 0.2, D_s, D_f
分别取 $5.0e^{-0.4}$ 和 $1.0e^{-0.3}$, T 和 T_k 分别取 200 和 50,
收敛阈值为 0.001. 当输出均方差低于 0.0005 时,表
示训练达到精度要求. 训练的最终结构分别为 26-
8-13 和 13-10-13. 表 1, 表 2 给出三层变结构神
经网络训练的关键过程^[5].

表 1 三层变结构神经网络的训练过程 $(E, C) \rightarrow U$ ^[5]
Table 1 Three layers variable structure neural network learning
process($E, C) \rightarrow U$

序 号	次数	均方差	衰减率	隐层节点数
1	200	0.756249	$5.069e^{-0.4}$	2
2	400	0.710168	$2.067e^{-0.4}$	2
3	650	0.382570	$9.495e^{-0.4}$	3
4	850	0.352882	$3.669e^{-0.4}$	3
⋮	⋮	⋮	⋮	⋮
18	11000	0.000967	$1.965e^{-0.4}$	10
19	11250	0.000923	$1.863e^{-0.4}$	9
20	11500	0.000887	$1.75e^{-0.4}$	8
21	26500	0.000499	$5.440e^{-0.4}$	8

表 2 三层变结构神经网络的训练过程 $(J \rightarrow V)$
Table 2 Three layers variable structure neural network learning
process($J \rightarrow V$)

序 号	次数	均方差	衰减率	隐层节点数
1	200	0.657442	$1.376e^{-0.3}$	2
2	400	0.571347	$4.514e^{-0.4}$	2
3	650	0.358595	$3.886e^{-0.4}$	3
4	2050	0.035480	$9.830e^{-0.4}$	3
5	2300	0.027231	$1.020e^{-0.3}$	4
⋮	⋮	⋮	⋮	⋮
28	12850	0.000935	$1.248e^{-0.4}$	12
29	13100	0.000910	$1.207e^{-0.4}$	11
30	13350	0.000889	$1.180e^{-0.4}$	11
31	13600	0.000868	$1.155e^{-0.4}$	10
32	21600	0.000496	$6.779e^{-0.4}$	10

4 多层神经网络的动态结构 (Dynamic struc-
ture of multi-layer neural network)

下面讨论多层神经网络的变结构,因为多层结
构的神经网络可能更优. 多层变结构神经网络的算
法归纳如下:

- 1) 对于任意给定的初始结构,当网络学习次数
达到 T 次后,计算 E, D_1, D_2 .
- 2) 如果 $E < \epsilon$,说明网络已经收敛,进入网络的
删减过程.
- 3) 如果 $D_2 > D_f$,说明网络性能很好,网络结构
保持不变.
- 4) 如果 $D_2 < D_s$,说明网络性能很差,进入网络
的增加过程.
- 5) 如果 $D_s < D_2 < D_f$ 且 $D_2 < D_1$,说明网络的
性能很差,进入网络的增加过程.
- 6) 如果 $D_s < D_2 < D_f$ 且 $D_2 > D_1$,说明网络性
能很好,网络保持不变.

根据文献[5]可取网络输出的均方差为网络的
性能评判标准,其表达式为:

$$E = \frac{\sum_{p=1}^P \sum_{k=1}^{N(l-1)} (Y_{l-1,k,p} - Y_{l-1,k,p}^d)^2}{P} \tag{5}$$

式中 $Y_{l-1,k,p}, Y_{l-1,k,p}^d$ 分别为输出层第 k 结点的
实际输出和期望输出. 在 BP 网络训练过程中,每学
习一定次数 T 后 E 及 E 的衰减率 D :

$$D = \frac{E(T-1) - E(T)}{E(T-1)} \tag{6}$$

式中 $E(T-1)$, $E(T)$ 分别 $(T-1)$ 及 T 时刻的均方误差. 设 D_1, D_2 分别为在 T 次学习前和学习后的衰减率, D_s, D_f ($D_f > D_s > 0$) 为两个给定的衰减

率. $\epsilon > 0$ 为给定的 E 的收敛阈值. 用 $E(j)$ 表示输出层 j 结点的输出误差均方值, 其表达

$$E(j) = \frac{\sum_{p=1}^P \sum_{k=1}^{N(l-1)} (f(\sum_{i=1}^{N(l)} W_{l-1,i,k} \cdot Y_{l,i,k} - W_{l-1,i,k} \cdot Y_{l,i,p}) - Y_{l-1,k,p}^d)^2}{P} \quad (7)$$

为保证网络结构变化时能够“平滑”过渡, 在增结点时, 与之相关的连接权值取很小或零; 在删除结点时, 删去对输出影响最小的结点. 即删去 i 点, 满足

$$i: E(i) = \min_{j=1}^{N(1)} E(j). \quad (8)$$

1) 增加过程.

多层变结构神经网络的增加过程包括增加隐层维数与层数两个方面. 首先增加隐层维数, 在能够促使学习效率最高的隐层中增加结点. 即在 i 层中增加结点, 满足

$$i: D(i) = \max_{j=1}^{l-2} D(j). \quad (9)$$

其中 $D(j)$ 表示在第 j 层增加结点后的衰减率. 如果增加结点并经过 T_k 次学习有 $E_k > E$ 或 $D_k < D_2$ 或 $D_k < D_s$ ($D_k > D_s$), 说明增加结点对网络的学习性能没有明显的提高. 这时通过增加网络的隐层层数提高网络的学习效率, 在最后隐层与输出层之间增加一层, 其维数定为

$$N(\text{new}) = \text{int}(\log_2(N(l-1) + N(l-2))). \quad (10)$$

如果经过 T_k 次学习仍然有 $E_k > E$ 或 $D_k < D_2$ 或 $D_k < D_s$, 则保持神经网络的原状态.

2) 删减过程.

在使神经网络学习效率最低的隐层中删去对输出影响最小的结点, 及在 i 层删除结点, i 满足

$$i: D(i) = \min_{j=1}^{l-2} D(j). \quad (11)$$

如果经过 T_k 次学习仍然有 $E_k > E$, 则保持神经网络的原状态.

同三层神经网络训练实例一样, 用上述算法进行模糊推理与实现自组织. 图 1 是它们的训练过程. 从图中可知, 实现 $(E, C) \rightarrow U$ 与 $J \rightarrow V$ 的最终结构分别为 $26-7-5-4-13$, $13-5-3-13$, 连接权值分别为 $26 \times 7 + 7 \times 5 + 5 \times 4 + 4 \times 13 = 289$ 个与 $13 \times 5 + 5 \times 3 + 3 \times 13 = 119$ 个, 比三层网络最终 $26-8-13$ 和 $13-10-13$ 所使用的连接权值 ($26 \times 8 + 8 \times 13 = 312$ & $13 \times 10 + 10 \times 13 = 260$ 个) 要少. 因而它们的结构更简单、更优. 值得强调的是, 这并不排除三层网络的实用性, 而且, 在多层神经网络的动

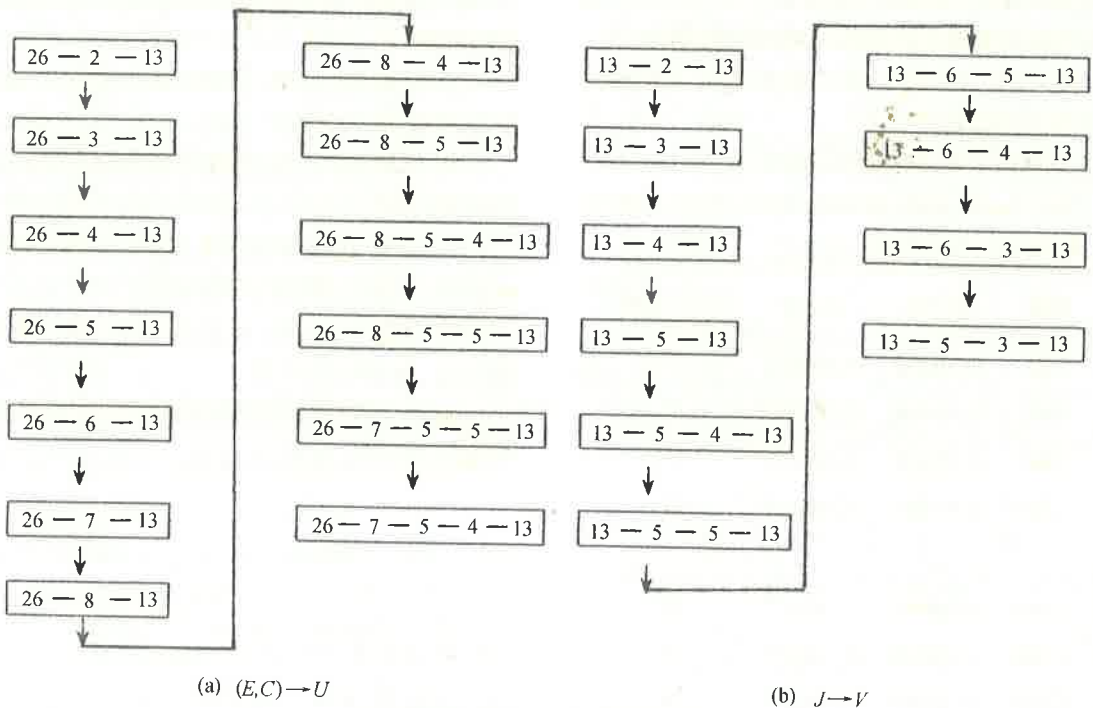


图 1 多层变结构神经网络训练过程

Fig. 1 Multi-layers variable structure neural network learning process

态算法中把 D_s 取小一些, 比如说取 $1.0e^{-4}$, 它的训练结果将与三层动态神经网络的训练结果一样。

5 仿真结果^[5] (Simulation result)

5.1 一阶延迟系统 (First order delay system)

对象特性为:

$$G(s) = \frac{K_0 \cdot e^{-\tau s}}{T_0 s + 1} \quad (12)$$

其离散形式为

$$y(k+1) = a \cdot y(k) + b \cdot u(k-N), \quad (13)$$

其中 y 为输出, u 为控制量, $a = e^{-\tau/T_0}$, $b = K_0 \cdot (1 - e^{-\tau/T_0})$, $N = \text{int}(\frac{\tau}{T})$ 。

图 2 是当采样周期取 $T = 0.1s$, 延迟时间 τ 取

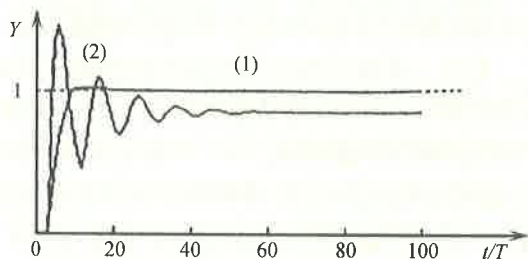
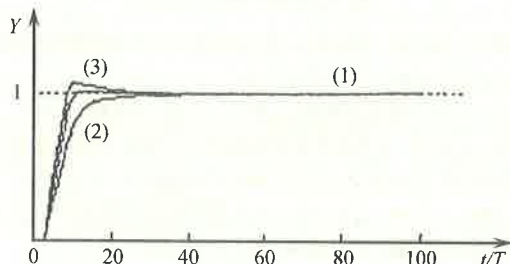


图 2 一阶延迟系统的控制曲线

Fig. 2 Control curve of first order delay system



(1) $K_0=5.0, T_0=1.0$
(a) (2) $K_0=7.0, T_0=1.0$
(3) $K_0=3.0, T_0=1.0$

0.2, $K_0 = 5.0$, $T_0 = 1.0$ 的控制曲线. 曲线(2)为系统响应曲线, 曲线(1)为基于变结构神经网络的自组织模糊控制结果。

5.2 抗干扰实验 (The test for resistance interference)

以一阶延迟系统为例. 在系统中加随机干扰, 干扰峰值取 0.1, 得到控制曲线如图 3 所示。

5.3 鲁棒性实验 (Robustness test)

鲁棒性是指当对象特性发生变化时, 系统保持性能不变的能力. 对一个控制系统而言, 希望其鲁棒性越强越好. 为了验证基于变结构神经网络的自组织模糊控制具有良好的鲁棒性, 以一阶延迟系统为例作鲁棒性实验. 在其它条件不变的情况下, 将对象的参数改变, 得到控制的结果如图 4 所示。

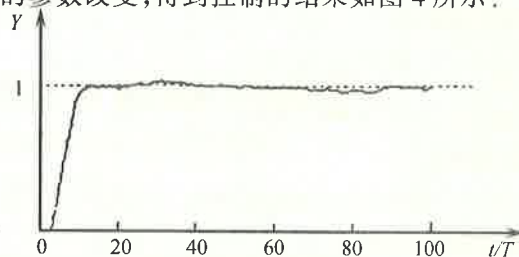
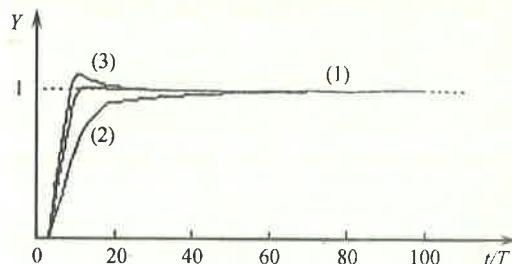


图 3 加干扰的一阶延迟系统的控制曲线

Fig. 3 The control curve for first order delay system under interference



(1) $K_0=5.0, T_0=1.0$
(b) (2) $K_0=5.0, T_0=2.0$
(3) $K_0=5.0, T_0=0.7$

图 4 一阶延迟系统的鲁棒性实验

Fig. 4 Robustness test for first order delay system

5.4 机器人控制 (Robot control)

机器人是高度非线性和强耦合的动力学系统, 无法用精确的数学模型描述, 因而用基于精确的控制方法进行控制便显得困难. 在文中, 采用基于变结构神经网络的自组织模糊控制方法进行控制尝试, 其控制结构见图 5. 图中, $q_d, \dot{q}_d$ 分别为给定的机器人关节的位移和速度, $q_k, \dot{q}_k$ 分别是机器人实际输出的位移与速度, D 是干扰。

假定用于仿真实验的 PUMA560 机器人的参数如下:

$\frac{K_a}{R_a} = 1.0$, R_a 为电机枢电阻 (Ω), K_a 为电机 -

力矩比 ($\text{kg} \cdot \text{m/A}$),

$K_b = -5.0$ 为比例常数 ($\text{V} \cdot \text{s/rad}$),

$J_{\text{eff}} = 1.0$, 折合到电机轴的电机和负载综合有效惯性矩 ($\text{kg} \cdot \text{m}^2$),

$f_{\text{eff}} = 3.0$, 折合到电机和负载综合有效粘摩擦系数 ($\text{kg} \cdot \text{m}^2/\text{s}$),

$\omega_n = 10$, 固有角频率 ($1/\text{s}$),

$n = 0.5$, 齿轮比。

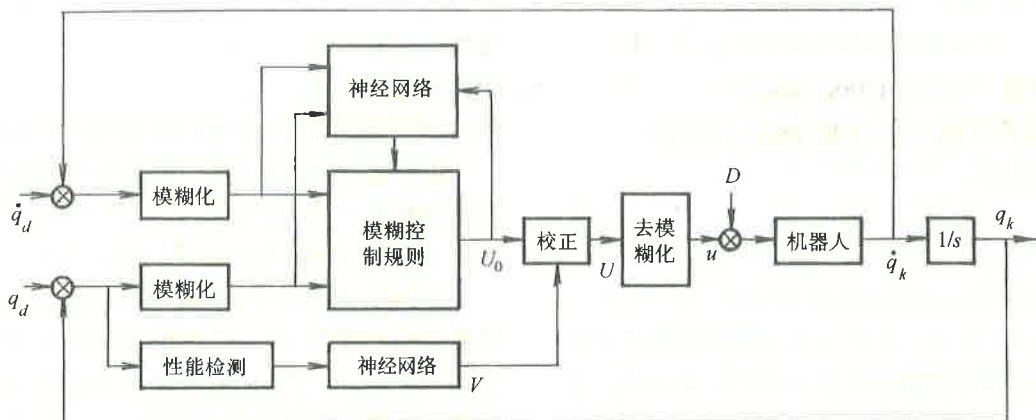


图 5 基于变结构神经网络的机器人自组织模糊控制系统

Fig. 5 Self-organizing fuzzy control system based on dynamic neural network for robot

期望作业轨迹

$$\begin{cases} q_d(t) = 0.3\cos(2\pi t), \\ \dot{q}_d(t) = 0.6\pi\sin(2\pi t), \end{cases} \quad 0 \leq t \leq 1. \quad (14)$$

干扰力矩为

$$D(t) = -\sin(0.5t). \quad (15)$$

系统初始条件为

$$q_d(0) = 0.3, \quad \dot{q}_d(0) = 0. \quad (16)$$

控制的结果见图 6. 图中, 虚线表示期望的运行轨迹, 实线为实际的位移轨迹, 很明显, 控制的效果相当不错.

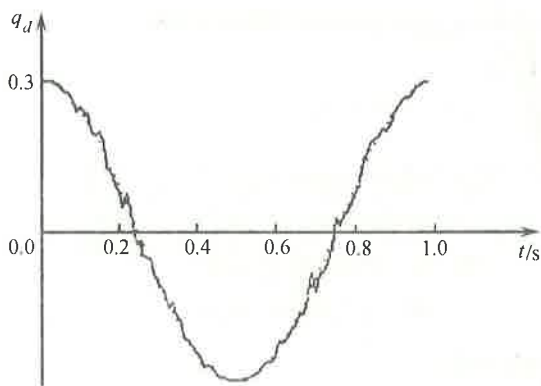


图 6 机器人仿真控制的位移曲线

Fig. 6 The simulation result for robot position control

6 结束语(Conclusion)

变结构神经网络算法可根据问题的复杂程度, 产生相应的最优网络结构. 用这样的神经网络进行模糊推理和规则校正, 可达到最佳的效果.

在仿真实验中, 作了不同对象的控制、抗干扰控制、鲁棒性控制和机器人控制等实验, 说明用变结构神经网络

络组成控制是可行的, 适应性、实用性等都比较好的.

需进一步探讨的是用变结构神经网络的训练过程速度并不理想, 有待从算法和硬件上去探讨. 网络学习后的结构虽然最优, 但计算量过大, 花费时间较长, 实时控制亟待解决. 神经网络应用于控制有很多结合方式, 到底何种方式最佳, 有待深入研究. 稳定性有待研究.

参考文献(References)

- 1 姚尹武, 熊金涛, 毛宗源. 一种神经网络自组织模糊控制. 控制理论与应用, 1996, 13(6): 738 - 744
- 2 焦李成. 神经网络系统理论. 西安: 西安电子出版社, 1993
- 3 葛红, 毛宗源. 采用改进的遗传算法学习基于神经网络结构的模糊辨识器. 控制理论与应用, 1996, 13(Suppl, 1): 155 - 158
- 4 王隆杰, 毛宗源. 用神经网络进行模糊推理的模糊控制器. 控制理论与应用, 1994, 11(4): 507 - 512
- 5 姚尹武. 基于动态神经网络的自组织模糊控制的研究. [硕士学位论文]. 广州: 华南理工大学, 1996

本文作者简介

毛宗源 1936 年生. 1962 年毕业于大连工学院. 现为华南理工大学自动控制工程系教授, 博士生导师. 研究兴趣: 工业自动化, 智能控制.

姚尹武 1993 年获西安交通大学自动控制专业学士学位, 1996 年获华南理工大学工业自动化专业硕士学位. 现在广东省珠海市国家税务局从事计算机网络工作, 研究兴趣为计算机管理及控制, 智能控制.

胡尚信 1938 年生. 1962 年毕业于大连工学院, 现为华东船舶工业学院自动化系副教授, 系主任. 主要研究方向为船舶自动化.