

## ◎ 工程与应用 ◎

## PID 神经元网络之权值直接确定法研究

张雨浓<sup>1</sup>, 杨逸文<sup>1,2</sup>ZHANG Yu-nong<sup>1</sup>, YANG Yi-wen<sup>1,2</sup>

1.中山大学 信息科学与技术学院, 广州 510275

2.中山大学 软件学院, 广州 510275

1.School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China

2.School of Software, Sun Yat-Sen University, Guangzhou 510275, China

E-mail: zhynong@mail.sysu.edu.cn

ZHANG Yu-nong, YANG Yi-wen. On weights-direct-determination method of PID neurons' network. Computer Engineering and Applications, 2009, 45(19): 189-191.

**Abstract:** Weights-determination is an important issue in the research of artificial neural networks. The traditional PID neurons' network adopts mostly the standard error Back-Propagation (BP) algorithm which obtains the connecting weights via iterative training. In contrast, a pseudo-inverse based method is proposed, which could directly determine the neural-network weights just in one step without iterative BP training. Computer-simulation results show that this weights-direct-determination method could be more efficient and effective than standard BP-training algorithms.

**Key words:** PID neurons' network; weights-direct-determination; system identification; pseudo-inverse

**摘要:** 网络权值如何确定, 是人工神经网络研究中的一个重要课题。传统 PID 神经元网络在该问题的研究上, 大多数采用误差回传(BP)的思想, 通过迭代训练而估算出该网络的连接权值。针对 PID 神经元网络, 对其进行简单巧妙的转化, 可提出一种基于矩阵伪逆表述的直接计算权值的方法, 从而避免了冗长的迭代训练过程。计算机仿真结果表明, 该权值直接确定方法不仅有更快的学习/计算速度, 而且能达到更高的计算精度。

**关键词:** PID 神经元网络; 权值直接确定; 系统辨识; 矩阵伪逆

**DOI:** 10.3778/j.issn.1002-8331.2009.19.058 **文章编号:** 1002-8331(2009)19-0189-03 **文献标识码:** A **中图分类号:** TP183

## 1 引言

人工神经网络(Artificial Neural Networks)因其具有高度的并行性、分布式存储、非线性自适应信息处理能力等显著特点, 在模式识别、智能控制、组合优化、系统辨识与预测等领域都得到成功应用<sup>[1-7]</sup>。在实际的应用中, 人工神经网络大多数采用的是一种基于误差回传(Back Propagation, BP)思想的神经网络及其改进模型。该模型可有效应用于多输入多输出系统的辨识, 其辨识模型如图 1 所示<sup>[4,8-13]</sup>。

PID 控制是最早发展起来的控制策略之一, 由于其算法简单、可靠性高以及抗干扰能力强, 被广泛应用于工业过程控制中<sup>[8]</sup>。加之, PID 控制器具有比例实时、积分滞后、微分超前的特性, 且结构简单规范、容易实现<sup>[8]</sup>。而前人提出的 PID 神经元网络<sup>[9-13]</sup>是将神经网络和 PID 控制规律融为一体而形成的一种内含动、静特性的多层前向神经网络系统, 可应用于(未知)目标系统辨识并已取得较好的效果。

人工神经网络研究中的一个关键方面就是神经网络权值

的快速确定: 无论是传统的 BP 算法还是其在迭代基础上的改进算法, 都存在迭代学习过程过长, 易陷入极小点等一系列固有缺陷<sup>[14-18]</sup>。考虑到文献[9-13]提出的 PID 神经元网络仍采用基于梯度下降的 BP 迭代学习算法以调整权值, 因此难免仍存在上述内在弱点。为了解决这些问题, 本文对传统 PID 神经网络辨识模型进行简单而巧妙的转化, 可提出一种基于矩阵伪逆的权值直接确定方法, 从而避免了传统 BP 学习算法的冗长迭代过程<sup>[14-18]</sup>。计算机仿真结果显示, 该方法在学习速度和计算精度上都体现出更好的性能。

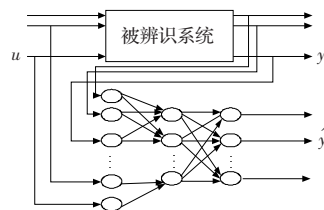


图 1 用于多输入多输出系统辨识的 BP 神经网络模型

**基金项目:** 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60775050);“新世纪优秀人才支持计划”基金(No.NCET-07-0887)。

**作者简介:** 张雨浓(1973-), 男, 博士, 教授, 博导, 研究方向: 神经网络、机器人、高斯过程、科学计算与优化。

**收稿日期:** 2009-01-09 **修回日期:** 2009-02-19

## 2 PID 神经网络与系统辨识

在研究权值直接确定法之前,我们对用于系统辨识的传统 PID 神经网络<sup>[9-13]</sup>进行必要的探讨与转化。

### 2.1 传统 PID 神经网络辨识模型

文献[9-13]提出了一种 PID 神经网络辨识模型,如图 2 所示,其中,  $u$  和  $y$  分别表示被辨识系统的输入和输出,  $\hat{y}$  表示 PID 神经网络的输出。

在图 2 中输入层有 2 个输入神经元,隐层有 3 个隐神经元(分别为 1 个比例神经元、1 个积分神经元和 1 个微分神经元)。对任意采样时刻  $k$ :

比例神经元输出

$$x_1'(k) = u_1'(k) \quad (1)$$

积分神经元输出

$$x_2'(k) = u_2'(1) + u_2'(2) + \dots + u_2'(k) \quad (2)$$

微分神经元输出

$$x_3'(k) = u_3'(k) - u_3'(k-1) \quad (3)$$

其中  $u_j'(k)$  和  $x_j'(k)$  ( $j=1, 2, 3$  为隐层神经元序号) 分别代表隐层神经元的输入和输出。

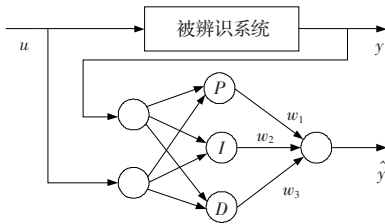


图 2 用于单输入单输出系统辨识的 PID 神经网络

### 2.2 PID 神经网络辨识模型的转化

在提出权值直接确定法之前,对图 2 的 PID 神经网络进行必要的思考与转化:对被辨识系统的输入与输出进行比较(即,求差),并以之作为 PID 神经网络的输入,则图 2 所示的 PID 神经网络辨识模型可转化为图 3,且转化后输入层和输出层神经元均使用线性激励函数(其权值固定为 1、阈值为 0)。

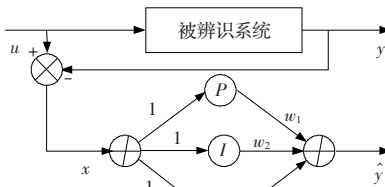


图 3 转化后的 PID 神经网络辨识模型

输入层神经元:

$$x(k) = u(k) - y(k) \quad (4)$$

隐层神经元:

$$\phi_1(k) = x(k) \quad (5)$$

$$\phi_2(k) = x(1) + x(2) + \dots + x(k) \quad (6)$$

$$\phi_3(k) = x(k) - x(k-1) \quad (7)$$

输出层神经元:

$$\hat{y}(k) = \sum_{j=1}^3 w_j \phi_j(k) \quad (8)$$

其中  $j=1, 2, 3$  为隐层神经元序号。隐神经元与输出层神经元之

间的连接权值记为  $w_j$ 。

### 3 权值迭代与直接确定

对于图 3 所给出的 PID 神经网络辨识模型,其隐层神经元个数为 3, 隐层神经元与输出层神经元之间的连接权值为  $w_j, j=1, 2, 3$ 。该神经网络输入与输出的关系可以描述为公式(8)。如果采用样本集  $\{(u_k, y_k), k=1, \dots, n\}$  对该神经网络进行批处理训练,并可定义其总体误差函数  $E$  如下:

$$E = \frac{1}{2} \sum_{k=1}^n (y_k - \hat{y}_k)^2 \quad (9)$$

其中  $n$  为采样点总数。把公式(8)代入公式(9),可以得到网络的误差函数  $E$  如下:

$$E = \frac{1}{2} \sum_{k=1}^n \left( y_k - \sum_{j=1}^3 w_j \phi_j(x_k) \right)^2 \quad (10)$$

则图 3 所示的 PID 神经网络基于负梯度 BP 学习算法的权值  $w_i$  修正公式可如定理 1 所述<sup>[14-18]</sup>。

定理 1 图 3 所示的 PID 神经网络的权值迭代公式可如下设计为:

$$w_i(m+1) = w_i(m) - \eta \frac{\partial E}{\partial w_i} \Big|_{w_i=w_i(m)} = w_i(m) - \eta \sum_{k=1}^n \left( \phi_i(x_k) \left( \sum_{j=1}^3 w_j(m) \phi_j(x_k) - y_k \right) \right) \quad (11)$$

其中  $m=1, 2, 3, \dots$  为迭代次数,学习步长  $\eta > 0$ 。

证明  $\frac{\partial E}{\partial w_i} = \frac{\partial \left( \frac{1}{2} \sum_{k=1}^n \left( y_k - \sum_{j=1}^3 w_j \phi_j(x_k) \right)^2 \right)}{\partial w_i} = \sum_{k=1}^n \left( y_k - \sum_{j=1}^3 w_j \phi_j(x_k) \right) \frac{\partial \left( y_k - \sum_{j=1}^3 w_j \phi_j(x_k) \right)}{\partial w_i} = \sum_{k=1}^n \left( \phi_i(x_k) \left( \sum_{j=1}^3 w_j \phi_j(x_k) - y_k \right) \right)$

再把  $\partial E / \partial w_i$  代入  $w_i(m+1) = w_i(m) - \eta \partial E / \partial w_i |_{w_i=w_i(m)}$  即可推导得到上述标量迭代公式(11)。证明完毕。

定理 2 标量迭代公式(11)可转化为如下矩阵形式:

$$\mathbf{w}(m+1) = \mathbf{w}(m) - \eta \Phi^T [\Phi \mathbf{w}(m) - \boldsymbol{\gamma}] \quad (12)$$

其中权值列向量  $\mathbf{w}$ 、输入受激励矩阵  $\Phi$  和目标输出向量  $\boldsymbol{\gamma}$  可分别记为:

$$\mathbf{w} := [w_1 \ w_2 \ w_3]^T \in R^{3 \times 1}$$

$$\Phi := \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \phi_3(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \phi_3(x_2) \\ \vdots & \vdots & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \phi_3(x_n) \end{bmatrix} \in R^{n \times 3}$$

$$\boldsymbol{\gamma} := [y_1 \ y_2 \ \dots \ y_n]^T \in R^{n \times 1}$$

证明 在公式(11)中,  $E$  对  $w_i$  的偏导为:

$$\frac{\partial E}{\partial w_i} = \sum_{k=1}^n \left( \phi_i(x_k) \left( \sum_{j=1}^3 w_j \phi_j(x_k) - y_k \right) \right)$$

其中  $i=1, 2, 3$ 。对应的矩阵向量表达式可因此推出:

$$\frac{\partial \mathbf{E}}{\partial \mathbf{w}} = \Phi^T (\Phi \mathbf{w} - \boldsymbol{\gamma})$$

则权值迭代公式(11)的矩阵向量形式可以写为  $\mathbf{w}(m+1) = \mathbf{w}(m) - \eta \Phi^T [\Phi \mathbf{w}(m) - \boldsymbol{\gamma}]$ , 即公式(12)。

证明完毕。

**定理 3** 定义  $\Phi, \gamma$  如上述定理 2。图 3 所示的 PID 神经网络之最优权值可直接给定为:

$$\omega = (\Phi^T \Phi)^{-1} \Phi^T \gamma \quad (13)$$

或简记为  $\omega = \text{pinv}(\Phi) \gamma$ , 其中  $\text{pinv}(\Phi)$  表示输入受激励矩阵  $\Phi$  的伪逆。

**证明** 根据定理 2 和权值向量迭代公式(12)

$$w(m+1) = w(m) - \eta \Phi^T [\Phi w(m) - \gamma]$$

当网络训练达到稳态时(此时  $k$  足够大, 使得  $w(m+1) = w(m) = \omega$ ), 由公式(12)可得:

$$\Phi^T [\Phi w - \gamma] = 0 \quad (14)$$

因此公式(14)的  $\omega$  即为该 PID 神经网络的稳态最优权值  $\lim_{m \rightarrow \infty} w(m)$ 。利用矩阵伪逆, 公式(14)的直接解为  $\omega = (\Phi^T \Phi)^{-1} \Phi^T \gamma$ , 从而可直接确定上述 PID 神经网络的最优权值  $\omega$  如定理 3 所示。证明完毕。

### 4 计算机仿真验证

对图 3 所示的 PID 神经网络进行计算机仿真验证, 并可采用如下非线性动态目标系统<sup>[10-13]</sup>:

$$y_k = 0.4y_{k-1} + 0.54y_{k-2} + f(u_{k-1})$$

其中  $k=2, 3, 4, 5, \dots, 200, y_k$  为  $k$  时刻系统输出(且令初值  $y_1 = y_2 = 0$ ),  $u_k$  为  $k$  时刻目标系统输入且有  $f(u) = u^3 + u^2 - 2.5u$ , 并设目标系统(也即, 待辨识系统)输入函数为:

$$u(k) = 0.2 \sin \frac{0.02\pi k}{25} + 0.3 \sin \frac{0.01\pi k}{75} + e^{-0.02k}$$

其中  $k=0, 1, 2, \dots, 200$ 。如图 4 所示, 样本数据点  $u_k$  作为系统输入, 而对应的数据值  $y_k$  为目标输出。

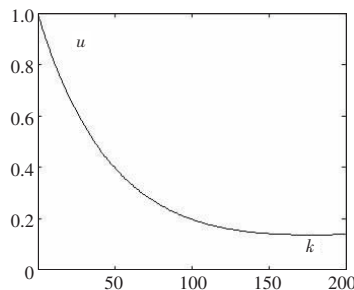
#### 4.1 三种权值方法的分别验证

为了考察权值直接确定法与两种迭代学习法的性能区别, 分别就三种方法[迭代法公式(11)和公式(12)、权值直接确定法公式(13)]对上述目标系统进行相应的计算机仿真验证。其中, 对于迭代法(公式(11)和公式(12)), 学习步长  $\eta$  需满足如下条件<sup>[19]</sup>:

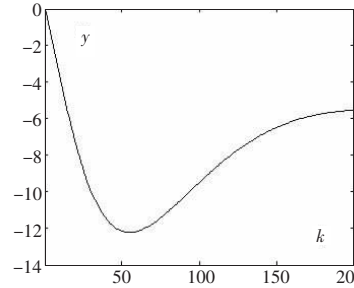
$$0 < \eta < \frac{2}{\|\Phi\|_2^2} = \frac{2}{\lambda_{\max}(\Phi^T \Phi)} \approx \frac{2}{\text{trace}(\Phi^T \Phi)}$$

图 5(a)和图 5(b)分别对应式(11)与式(12)所得到的网络逼近效果。由图 5 可知, 与权值直接确定法相比(如图 5(c)所示), 基于迭代法训练 120 000 次而得到的网络输出与目标输出的误差仍比较大, 而基于矩阵伪逆的权值直接确定方法可以一步就直接得到与目标输出完全重合的网络输出值。因此可见, 该

PID 神经网络模型, 在使用权值直接确定法的情况下<sup>[14-18]</sup>, 能够表现出更好的学习能力和系统辨识能力。



(a) 系统输入函数曲线



(b) 被辨识系统目标输出曲线

图 4 被辨识目标系统的输入与输出情况

表 1 分别采用标量迭代法(式(11))、矩阵向量迭代法(式(12))与权值直接确定法(式(13))的仿真结果比较

权值方法	迭代次数	训练时间/s	总误差 E	均差 E/m	校验均差
标量迭代法	120 000 次	$5.46 \times 10^2$	$5.95 \times 10^0$	$2.96 \times 10^{-2}$	$2.89 \times 10^{-2}$
矩阵向量迭代法	120 000 次	$6.49 \times 10^0$	$5.92 \times 10^0$	$2.94 \times 10^{-2}$	$2.88 \times 10^{-2}$
权值直接确定法	一步直接计算	$1.12 \times 10^{-3}$	$9.95 \times 10^{-1}$	$4.95 \times 10^{-3}$	$4.66 \times 10^{-3}$

#### 4.2 仿真结果分析

从图 5 和表 1 可知, 采用权值直接确定方法, 无论在训练速度还是计算精度上, 都要明显优于迭代学习方法。(1)训练时间方面: 两种迭代学习法的网络训练时间分别为  $5.46 \times 10^2$  s 和  $6.49 \times 10^0$  s, 而权值直接确定法仅需  $1.12 \times 10^{-3}$  s。(2)计算精度方面: 两种迭代学习法训练 120 000 次后的均方误差分别为  $2.96 \times 10^{-2}$  和  $2.94 \times 10^{-2}$ , 而权值直接确定法直接计算得到的均方误差仅为  $4.95 \times 10^{-3}$ 。网络训练完毕后, 也采用了一组未经训练的校验样本(改变了输入值  $u_k$ )对该 PID 神经网络进行测试。如表 1 最后一列所示, 迭代学习法对应的校验均方误差为  $2.89 \times 10^{-2}$  和  $2.88 \times 10^{-2}$ , 而权值直接确定法对应的校验均方误差为  $4.66 \times 10^{-3}$ , 由此可见该 PID 神经网络在使用权值直接确定法的情况下具有更好的学习和逼近能力。

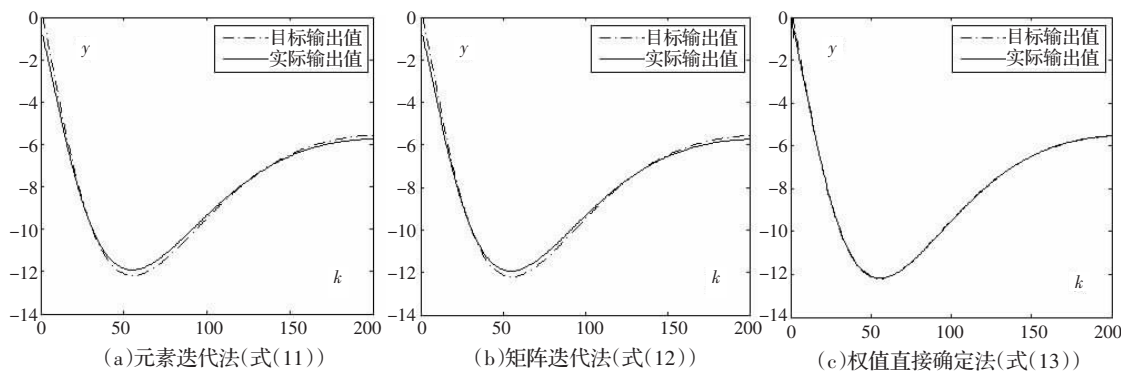


图 5 基于三种权值方法(分别对应权值公式(11)、(12)和(13))的 PID 神经网络系统逼近效果

(下转 197 页)