

51-55

非线性系统的神经网络辨识和控制方法

薛家祥 黄石生

(华南理工大学自动化系 广州510641)

TP271

摘要 Basing on the properties, architectures and performance of artificial neural networks, This paper analyzes its applications in modelling, identification and control of nonlinear systems. A variety of neural network architectures in control are supplied. We also discuss the theory problems to be resolved and the areas for future research.

关键词 Neural networks, Nonlinear systems, Modeling, Identification, Control.

1 引言

在 Wiener 开创控制论的伊始,就将控制、信息和神经科学作为一个共同的课题。此后,控制学科、计算科学和神经生理学趋于分开发展。自从80年代初期以来,神经网络有了长足的进步,在人工智能和复杂的非线性系统的控制中其应用潜力很大。神经网络能用于控制领域是由于有如下特点:^[1,2]

- (1)非线性系统 由于非线性的多样性,其普遍适应的非线性控制器设计的理论还未产生。神经网络的任意近似非线性映射能力,比其它近似方法能得到更简洁的模型。
- (2)并行分布处理 神经网络有着高度并行结构,可望得到高度的容错能力,加之并行处理,整个处理速度很快。
- (3)硬件实现 神经网络不仅进行并行运算,而且能由 VLSI 硬件实现,这既增加了速度,又增大了能够实现的网络规模。
- (4)学习和适应性 神经网络能用所研究系统过去的记录数据来加以训练,能推广到训练数据中未出现的输入的情况,也能在线修正。
- (5)数据融合 神经网络能同时运算定量和定性数据,这相当于神经网络介于有定量数据的工程系统和有符号数据的人工智能系统之间。
- (6)多变量系统 神经网络自然地能处理多输入和多输出,可以很容易地应用到多变量系统中。

2 神经网络结构

神经网络是由叫做神经元(亦称感知器)的基本

处理单元互连构成的,单个感知器能形成简单的决定边界和逻辑函数,其结构见图1。

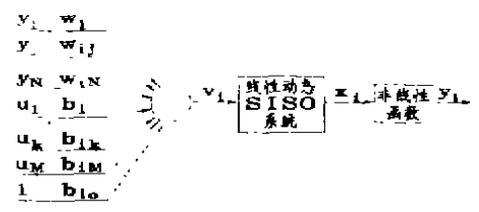


图1 神经元的结构

一般地神经元可以分为三个部分^[3]:

- (1)权累加器,它将神经元的所有赋予一定权值的输入累加在一起;
- (2)线性动态 SISO 系统,可用如下传递函数形式描述:

$$X_1(s) = H(s)V_1(s) \quad (1)$$

其中 H(s)有五种形式:

- $H(s) = 1,$
- $H(s) = 1/s,$
- $H(s) = 1/(1+sT),$
- $H(s) = 1/(a_0s - a_1),$
- $H(s) = e^{-sT}.$

(3)非线性激活函数。

神经网络可分为两种基本类型:静态神经网络和动态神经网络。这取决于方程(2)中 H(s)的值,若 H(s)=1,则为静态神经网络;否则,取其它形式为动态神经网络。

2.1 静态神经网络

薛家祥 博士生,研究兴趣为神经网络、模糊控制和自适应控制。黄石生 教授,博士生导师,研究兴趣为焊接过程智能控制。

在静态神经网络中,多层感知器(MLP)应用最广泛,它由几层感知器级联在一起构成,其特征是节点方程无记忆,即它们的输出只是当前输入的函数,与过去或未来的输入或输出无关。另一种静态神经网络模型是近年来发展的径向函数(RBF)神经网络。其特点是学习速度快,并收敛于全局最小,但因其只有一个隐含层,所以只适于较简单的对象,这两种网络统称为多前向网络。

多层感知器(MLP)的一般学习方法为基于梯度搜寻原理的BP算法,该算法的学习速度较慢,且易于陷入误差函数的局部最小点。为此,出现了多种改进方法,包括线性搜寻和二阶法等。

多层前馈网络有如下三个功能:(1)实现 Boolean 逻辑函数;(2)分类问题中区分模式空间;(3)函数近似中实现非线性传递^[1]。

2.2 动态神经网络

动态神经网络的节点方程是有记忆的,它由微分或差分方程描述。动态神经网络有三种不同的类型:(1)前向动力学神经网络,(2)有输出反馈神经网络,(3)有状态反馈神经网络。由于现实世界中许多非线性动态系统,因此研究动态神经网络尤为重要。

(1)时延神经网络。多层感知器也能用来处理时间序列数据,只要简单地将瞬时序列转变为静态模式即可,这里,时间就被看成问题的另一维,此则时延神经网络,见图2^[2]。系统的输出与有限个瞬时输入有关,即:

$$y(k) = f[x(k), x(k-1), \dots, x(k-n)] \quad (3)$$

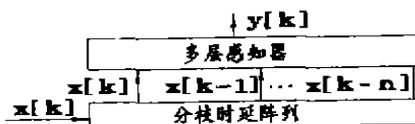


图2 时延神经网络

(2)输出反馈神经网络。将输出反馈加到神经网络结构中的一种简单办法是通过分枝延迟阵列,见图3。这种神经网络结构主要用在非线性辨识和控制中,理论上它能对式(4)表示的任何系统建立模型。其学习算法亦为梯度下降法。

$$y(k) = f[x(k), x(k-1), \dots, x(k-n), y(k-1), \dots, y(k-m)] \quad (4)$$

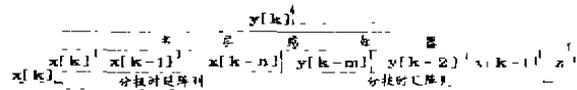


图3 输出反馈神经网络

(3)状态反馈神经网络。这种神经网络典型的结构是节点间有反馈连接的单层网络,最一般的情况是节点完全互连,这类动态神经网络以 Hopfield 神经网络最为著名。它是联想记忆型,主要用于模式识别和优化领域。

上述两种带反馈的动态神经网络也称为循环(Recurrent)神经网络,反馈动态系统比纯前馈系统有很多优点,一个小的反馈系统相当于一个大的或可能无限大的前馈系统,反馈系统特别适于辨识(建模)、控制和滤波。训练循环神经网络的关键在于寻求一种有意义的学习算法,大部分的学习算法是类似于BP算法所用的梯度搜寻技术,由于节点的输出是其前一步时间输出的递归函数,梯度算法也必须进行递归计算,因此这些学习算法非常复杂。

3 神经网络的学习行为

人工神经网络的应用中,若用一组合适的数据来训练网络,那么它就有定常系数,因此,表现出固定的行为,这就是静态神经网络模型,在神经网络训练期间,若提供一种算法能自动调整网络的系数,此则为动态神经网络模型。

学习算法可分为两组:(1)监督学习,它配合外部参考信号(教师)和/或局部系统信息;(2)无监督学习,它无外部参考信号,而依靠局部信息和内部信号。

(1)静态多层感知器的学习。一般地多层感知器采用BP算法训练,该算法的学习过程由正向传播和反向传播组成,在正向传播过程中,输入信息从输入层经隐层逐层处理并传向输出层,每一层的神经元的状态只影响下一层神经元的状态。如果输出层不能达到期望的输出,则转入反向传播,误差信号沿原来的通路返回,通过修改各层神经元的权值使得误差信号最小。

(2)动态神经网络的学习。由于动态神经网络的神经元将信号以变增益(权值)反馈到其它神经元或自身的输入中,这是固有动态的,它有强大的表示

能力。在动态神经网络训练中存在两种学习方式：(1)固定点学习，目的在于使神经网络达到规定的平衡点或完成稳定的匹配；(2)轨迹学习，训练神经网络及时地跟踪期望的轨迹。

在非线性系统控制中，神经网络可看作为过程模型化形式，对象动力学和映射特点的知识被隐含地存储在神经网络中。

4 神经网络的非线性映射能力

神经网络的非线性映射能力是极为重要的，它直接决定了其能否用于非线性系统的控制中。训练神经网络的学习结构表示非线性系统的前向和逆向动力学，它决定了神经网络的训练效率及用于实时控制的能力。

有一个隐层的多层感知器能任意近似任意的多变量连续函数，但这一结论对于隐含层节点是不加限制的，对于复杂的问题，隐节点数可能大到令人难以接受，实践表明两隐层比单隐层的多层感知器能以较少的处理单元获得更高的精度和更好的一般化能力。

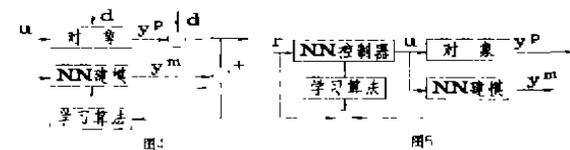
在多层前馈神经网络中，RBF神经网络(如高斯网络)有最好的函数近似性能，由于RBF只有一个隐含层，能用线性化技术来训练以确保获得全局最优解，但这也是其缺点所在，对于多输入的高维问题，其所需的节点太多，因此，它只适于简单的非线性对象^[5]。

对于具有联想记忆功能的神经网络，如Hopfield和ART等，多用于模式分类和优化等领域。

5 神经网络用于过程辨识的方法

系统辨识的重要问题是系统的可辨识性，即给定一个特殊的模型结构，所研究的系统能否用其适当地表示出来。建模的方法有两种：前向建模和逆向建模。

(1) 前向建模。训练神经网络表示系统的前向动力学就称为前向建模，由图4可见，神经网络模型与系统并置，系统与神经网络输出的误差用作为训练信息。



在辨识过程中，与控制相关的问题是系统的动力学性质，在神经网络中引入动力学有两种方法，即应用循环神经网络和直接将动力学引入到神经元中。

假设非线性离散时间系统的关系方程为：

$$y^p(k+1) = f[y^p(t), \dots, y^p(t-n+1), u(t), \dots, u(t-m+1)] \quad (5)$$

显然对于这一系统的建模方法就是神经网络的输入输出结构与系统相同，则其输出可表示为：

$$y^m(k+1) = f[y^p(k), \dots, y^p(k-n+1), u(k), \dots, u(k-n+1)] \quad (6)$$

其中，f表示神经网络的非线性输入输出映射，即是f的近似，在神经网络的输入中包含真实系统的过去输出值，如果神经网络经过适当的训练后，能给出对象的很好表示，即 $y^m \approx y^p$ ，那么随后网络的训练就能用自身输出反馈成为其输入的一部分，这样神经网络可以独立于对象被使用。该模型可描述如下：

$$y^m(k+1) = f[y^m(k), \dots, y^m(k-n+1), u(k), \dots, u(k-m+1)] \quad (7)$$

方程(6)的结构由于其稳定性结果在辨识中得到青睐；在处理有噪声系统时，方程(7)的结构可以避免噪声作用于实际所引起的偏差问题。

(2) 逆向建模。动态系统的逆模型在控制结构中起着至关重要的作用。图5示出了一种逆向建模的结构，神经网络逆模型位于系统之前，输入的训练信息包含了所控制系统的期望操作输出空间，即系统的参考信息。这种学习结构还含有与对象并置的前向模型。训练算法的误差信号是训练信号和系统输出的差，在有噪声系统，误差信号也可以是训练信号和前向模型输出的差。

这种建模方法的特点是：(1)建模过程的目标直接基于期望的系统输出和实际系统输出之间的误差。(2)在系统前向模型非一一对应的情况下，可找到某一特定的逆。

由方程(8)的反函数 f^{-1} 产生 $u(k+1)$ 需要知道系统未来的输出值 $y^p(k+1)$ ，为了克服这一问题，可用 $r(k+1)$ 代替 $y^p(k+1)$ ， $r(\cdot)$ 是一步超前可知的。因此，用神经网络建立对象逆模型的非线性输入输出关系为：

$$u(k) = f^{-1}[y^p(k), \dots, y^p(k-n+1), r(k+1), u(k-1), \dots, u(k-m+1)] \quad (8)$$

神经网络逆模型接收的输入为系统当前和过去的输出，在无真实系统或系统有噪声的情况下，可用前向模型的输出 y^m 代替系统的真实值^[5]，则式(8)

可改写为:

$$u(k) = f^{-1}[y^m(k), \dots, y^m(k-n-1), r(k+1), u(k-1), \dots, u(k-m-1)] \quad (9)$$

6 非线性系统的神经网络控制方法

无论是静态系统还是动态系统都能用神经网络来控制,因为神经网络的参数不是由分析计算而是通过输入输出数据训练得到的。

在有了动态系统的前向模型及逆模型之后,就可用于控制该系统,非线性控制结构有很多种,我们将针对神经网络能够实现的系统前向模型和逆模型来建立其控制结构。

6.1 直接逆模型控制

将神经网络逆模型直接作为对象的控制器,使得组成的系统在期望的响应和系统的输出间成为恒等映射。学习过程的目的在于调整神经网络的权值,以便在对象期望的运行范围内,产生正确的映射。这种控制方法可有不同的学习结构^[6,7]。

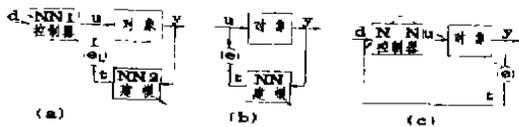


图6 (a)间接学习结构 (b)一般学习结构 (c)特殊学习结构

间接学习结构见图6(a),期望响应经过神经网络控制器 NN1 产生对象的输入 u , 对象对这一输入的响应被用作为 NN1 的复制网络 NN2 的输入。将神经网络 NN1 和 NN2 输出的差 $e_1 = u - t$ 作为训练数据,通过使误差 e_1 最小化来得到期望的输出,这种学习结构的优点是直接基于神经网络输出来调节权值使误差最小。但是 $e_1 \rightarrow 0$, 并不能保证 $e \rightarrow 0$, 当神经网络 NN1 将所有期望响应映射成唯一的对象输入时,即 $u = u_0$ 。同样,神经网络 NN2 也将所有的 y 映射成 $t = u_0$, 显然 $e_1 = 0$, 但此时的 e 并非是 0。

一般学习结构,见图6(b),选择输入 u 作用于对象获得相应的输出 y ,神经网络被训练使之从 y 再生 u 。然后,这个神经网络被用作为控制器,就能由期望的响应 d 产生近似的 u ,使实际对象的输出 y 接近 d 。这种训练是离线进行的,不用考虑稳定性问题,但不适于实时控制。

特殊学习结构见图6(c),用期望响应 d 作为神经网络的输入,训练网络找到对象输入 u ,使系统输

出 y 达到期望的 d ,在每次迭代中,对象的期望响应和实际响应的差被用来调整网络的权值,以最大限度地减少误差,由于误差要经过对象反传,因此,这一过程需要对象的某些知识,它能在线对感兴趣的特殊区域学习。这里,必须考虑稳定性的问题,学习率不能取得过大。

直接逆模型结构通常用于机器人的控制中。显然,这种方法的控制效果主要依赖于逆模型的精度。由于缺乏反馈,所以其鲁棒性较差。

6.2 内模控制

图7示出了内模控制结构,其中系统的前向和逆向模型被直接用作为反馈环中的单元,系统模型和控制器(逆模型)由神经网络实现,在系统输出和模型输出之间的差被用于反馈,该信号随后被前向通路控制器子系统处理。在控制器前通常加一线性滤波器,能将其设计成引入期望的鲁棒性和跟踪闭环系统的响应。

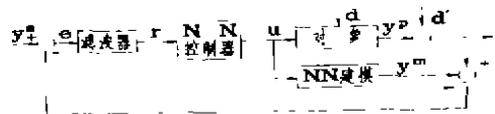


图7 内模控制结构

6.3 模型参考控制

图8示出了模型参考控制结构^[8],其中参考模型是稳定的,它规定闭环系统期望的性能,控制系统的目的就在于使对象的输出 $y^p(k)$ 匹配参考模型的输出,即对于给定的常数 $\epsilon \geq 0$,有:

$$\lim \|y^p(k) - y^r(k)\| \leq \epsilon \quad (10)$$

上式被用来训练神经网络控制器,一般地,训练过程将使这一控制器成为按参考模型确定的“解调”逆。

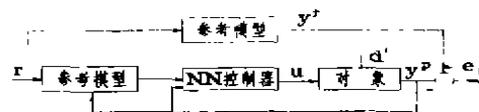


图8 模型参考控制结构

6.4 自适应控制

自适应控制理论在过去近20年里取得了丰硕的理论成果,尽管它是建立在线性时不变系统假设的基础上的,但其概念和理论并行存在于神经网络领域中。

对于一离散非线性多变量系统,其状态方程可写为:

$$S: \begin{cases} x(k+1) = f[x(k), u(k)], \\ y(k) = h[x(k)] \end{cases} \quad (11)$$

这里, $u(k), y(k) \in R^m, x(k) \in R^n$, 分别为系统在 k 时刻的输入、输出和状态, $f, h \in C^\infty$, 参考输入为 $r(k) \in R^m$. 对系统的控制问题就是确定输入 $u(k)$ 使 $y(k)$ 跟踪参考输入 $r(k)$. 在 f 和 h 未知的情况下, 就成为自适应控制问题. 若神经网络作为控制器和辨识器, 则可用于非线性系统, 见图9所示^[6].

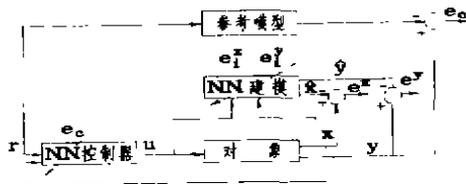


图9 自适应控制结构

神经网络自适应控制器的设计过程如下:

(1) 模型辨识, 由于 f 和 h 未知, 可用输入输出数据训练神经网络来辨识系统的模型, 并使辨识误差最小化, 这种用开环方式离线训练得到的神经网络模型, 其精度不能达到在线控制的要求, 但它提供了合理的待辨识参数的初始值。

(2) 控制器的离线设计, 为了确保控制参数在线调整的稳定性, 控制参数的初始值是至关重要的, 这些值可由第(1)步所获得的近似模型离线确定。

(3) 在线调整辨识器和控制器, 在期望的均匀分布的参数输入范围内, 在每一瞬时根据辨识器和控制器的误差分别对它们加以更新, 辨识器采用静态神经网络, 而控制器则用动态神经网络。

6.5 其它控制方法

非线性控制设计的最新方法包括非线性算子理论和优化技术, 神经网络可望为其提供非线性系统模型. 神经网络既有定量近似的能力也可用于表示符号. 因此, 它与专家系统相结合是很有前途的人工智能研究课题. 模糊控制与神经网络相结合用于控制领域已取得显著的成果, 有关这方面的研究报道很多, 这里不再尽述。

7 讨论

由上述分析可知, 神经网络控制器的应用潜力

很大, 但由于在控制领域的应用时间不长, 因此还有许多理论和技术问题尚待解决。

(1) 神经网络的稳定性和收敛性, 仍未给出完整的理论证明。

(2) 对于一个具体的控制对象, 如何选择合适的神经网络拓扑结构还未有统一的指导性理论。

(3) 神经网络的学习速度较慢, 如何优化其算法, 使之适于实时控制的更高要求尚需进一步的研究。

(4) 神经网络控制器的硬件实现及商品化还需作出艰苦的努力。

可以相信, 随着神经网络理论研究的深入和完善, 它将在控制领域中发挥愈来愈大的作用。

参考文献

- [1] D. E. Rumelhart et al., Learning Representation by Back-propagating Errors, Nature, Vol. 323 9, Oct. 1986
- [2] R. P. Lipmann, An Introduction to Computing with Neural Nets, IEEE ASSP Magazine, Apr. 1987
- [3] R. Hecht-Nielsen, Theory of the Backpropagation Neural Networks, IJCNN' 1989, Vol. 1
- [4] D. R. Hush et al., Progress in Supervised Neural Networks, IEEE Signal Processing Magazine, Jan. 1993
- [5] K. S. Narendra et al., Identification and Control for Dynamical Systems Using Neural Networks, IEEE Trans. on Neural Networks, Vol. 1, 1990
- [6] Yichuang Jin and Tong Pipe, Neural Net Versus Control Theory, ICJNN' 1991
- [7] D. Psaltis et al., A Multilayered Neural Network Controller, IEEE Control Systems Magazine, Apr. 1988
- [8] K. S. Narendra et al., Adaptive Control of Nonlinear Multivariable Systems Using Neural Networks, Neural Networks Vol. 7, No. 5, 1994