

动态驱动神经网络辨识永磁直线同步电动机模型

吕 刚, 范 瑜, 李国国

(北京交通大学 电气工程学院, 北京 100044)

摘要: 永磁直线同步电动机(PMLSM)模型的建立对研究其稳态特性、动态特性和控制策略都是非常重要的. 本文利用动态驱动神经网络对其进行建模, 并在代价函数一致的基础上加入残差分析法来辨识模型的阶次, 使得神经网络具有自动识别阶次的能力. 为了克服神经网络结构依靠人工试凑的不足, 使用基于Hessian矩阵的修剪法来优化其结构. 考虑到改进BP算法(学习速率自适应、动量项的方法)的一些固有缺点, 使用NDEKF(基于节点的解耦扩展Kalman滤波器算法)来训练网络. 实验证明, 混合网络能够准确辨识出试验样机的阶次并且输出结果与实际结果十分接近; 同时将NDEKF与改进BP算法进行对比, NDEKF算法具有收敛较快、泛化能力强等特点.

关键词: 神经网络; 永磁直线同步电动机; 辨识; 混合神经网络; NDEKF

中图分类号: TP273 **文献标识码:** A

Hybrid nonlinear autoregressive neural networks for permanent-magnet linear synchronous motor identification

LÜ Gang, FAN Yu, LI Guo-guo

(Electric Engineering school, Beijing Jiaotong University, Beijing 100044, China)

Abstract: The modeling of permanent-magnet linear synchronous motor is crucial to the control, static and dynamic characters analysis for the system. The model of permanent-magnet linear synchronous motor is presented in this paper by using neural networks of the nonlinear autoregressive with exogenous inputs. For the same cost function, residual signal analysis is employed to identify motor's order automatically. Some shortages of back-propagation algorithm are considered, so NDEKF (node-decoupled extended Kalman filter) is applied to train networks. Finally, experiment results show that the hybrid neural networks of the nonlinear autoregressive with exogenous inputs can identify object's order precisely, and the output of networks is very close to the experimental result. In the experiment, the performance of NDEKF is often superior to that of BP, such as it requires significantly fewer presentations of training data and shorter training time than BP does, and has the better generalization ability.

Key words: neural networks; permanent-magnet linear synchronous motor; identification; hybrid nonlinear autoregressive neural network; NDEKF

1 引言(Introduction)

PMLSM(永磁直线同步电动机)具有结构简单、加工方便、高效、节能等显著优点, 在地面高速运输线、往复伺服系统、垂直提升系统等场合正越来越受到重视, 具有广阔的应用前景. 与永磁旋转电机相比, 永磁直线同步电动机稳态分析的理论尚不完善, 导致不能准确分析直线电动机的稳态运行特性以及实现高性能的稳态控制. PMLSM动态方面的研究几乎还是空白, 这就更使得无法了解PMLSM运行的动态特性, 也不能实现良好的动态控制.

监督神经网络具有很强的随机逼近能力, 通过学习它可以拟合任意一个高维空间的曲线. 针对PMLSM的动态特性和稳态特性, 本文使用动态驱动神经网络逼近其实际的动态和稳态特性. 针对目前大部分神经网络的辨识都是在已知模型的阶次

前提下进行, 在与神经网络具有相同的代价函数基础上, 使用残差分析法和神经网络相结合的策略, 使得神经网络具有自动辨识对象阶次的能力. 使用DEKF算法, 其不仅可以按照解耦的不同程度来调节算法的复杂度, 而且具有收敛较快、在达到相同的逼近效果时所需数据比改进BP算法少等特点.

2 离散化的ARMA数学模型 (ARMA discrete model)

由于PMLSM是非线性、强耦合系统, 因此选用带控的非线性自回归滑动平均(ARMA)模型作为其辨识模型, 其表达为

$$\mathbf{y}(t) = f(\mathbf{y}(t-1), \dots, \mathbf{y}(t-n_y), \mathbf{u}(t-1), \dots, \mathbf{u}(t-n_u), \mathbf{e}(t-1), \dots, \mathbf{e}(t-n_\epsilon)) + \mathbf{e}(t). \quad (1)$$

其中: $\mathbf{y}(t) = [y(t), \dots, y_m(t)]^T$, $\mathbf{u}(t) = [u(t), \dots,$

$u_r(t)^T, e(t) = [e(t), \dots, e_m(t)]^T$ 分别是系统的输入、输出、噪声, n_y, n_u, n_ξ 分别是其阶次, $\{e(t)\}$ 是零均值的独立过程, $f(\cdot)$ 是向量非线性函数. 式(1)可以分解成 m 个标量方程, 并将 $f(\cdot)$ 展成 L_i 阶的多项式, 可得

$$y_i = \sum_{j=1}^{L_i} \theta_{ij} x_{ij}(t) + e_i(t) + \theta_{i0}, \quad i = 1, 2, \dots, m. \quad (2)$$

其中

$$\begin{cases} n_i = \sum_{j=1}^{n_i} n_{ij}, \quad n_{i0} = 1, \quad n_{i1} = m(n_y + n_\xi) + rn_u, \\ n_{ij} = n_{i1} + P_{j-1}^1 C_{n_{i1}}^2 + \dots + P_{j-[j-2]}^{[j/2]} C_{n_{i1}}^{[j/2]+1} + \\ P_{j-[j-1]}^{j-[j/2]-2} C_{n_{i1}}^{[j/2]+2} + \dots + P_2^1 C_{n_{i1}}^{j-1} + C_{n_{i1}}^j, \end{cases}$$

$j = 1, 2, \dots, L_i; \theta_{i0}$ 是常数项; $x_{ij}(t) (i = 1, 2, \dots, m, j = 1, 2, \dots, n_i)$ 是单项式, 每一项均由带延迟的输出、输入和噪声构成; n_{ij} 是 j 阶多项式的项的个数. 式(2)写成矩阵形式:

$$Y_i = X_i \theta_i + e_i. \quad (3)$$

3 残差分析法确定结构参数 (Structural parameter identification using residual signal analysis)

下面在模型(3)的基础上利用不同结构模型的残差变化情况来确定结构参数, 其本质是通过检验输出的残差变化情况来确定数据的相关性.

考虑取如下作为参数估计值:

$$\hat{\theta} = X^+ y, \quad (4)$$

其中 X^+ 表示 X 的伪逆, 那么向量 y 的估计值为

$$\hat{y} = X X^+ Y. \quad (5)$$

残差为

$$\varepsilon(\hat{\theta}) = Y^T (I - X X^+) y. \quad (6)$$

引理 当且仅当 $\varepsilon(\hat{\theta}) = 0$ 时, 向量 y 是 (x_1, x_2, \dots, x_n) 的线性组合, 并且向量集合 (x_1, x_2, \dots, x_n) 不要

求是相互独立的.

下面来考虑加入噪声之后的情况, 设向量 y 含有均值为零的噪声 ω , 则有

$$z = y + \omega. \quad (7)$$

这时, z 是可测的, 而 y 是未知的. 由式(6)定义的输出残差, 有

$$\begin{aligned} \varepsilon(\hat{\theta}) &= z^T (I - X X^+) z = \\ &= y^T (I - X X^+) y - 2y^T (I - X X^+) \omega + \\ &= \omega^T (I - X X^T) \omega. \end{aligned} \quad (8)$$

设 ω 是均值为零且与向量 y 不相关的的噪声向量, 若 y 是向量集合 (x_1, x_2, \dots, x_n) 的线性组合, 则有

$$E\{\varepsilon(\hat{\theta})\} = E\{\omega^T (I - X X^+) \omega\}. \quad (9)$$

否则

$$E\{\varepsilon(\hat{\theta})\} = E\{\omega^T (I - X X^+) \omega\} + y^T (I - X X^+) y. \quad (10)$$

这样, 当 \hat{n}_y, \hat{n}_u 分别是其阶次, y 是向量集合 (x_1, x_2, \dots, x_n) 的线性组合时, 在数据的采样长度 N 充分大的条件下, 有

$$E\{\varepsilon(\hat{\theta})\} = E\{\omega^T (I - X X^+) \omega\} \rightarrow \sigma^2. \quad (11)$$

否则

$$E\{\varepsilon(\hat{\theta})\} = E\{\omega^T (I - X X^+) \omega\} > \sigma^2. \quad (12)$$

4 NARX神经网络模型及其网络结构确定 (NARX neural networks and its structure identification)

对应于数学模型(1), 使用NARX神经网络模型. 如图1所示, 实际对象的输出通过 \hat{n}_y 个抽头延迟线记忆反馈到输入端、每一个输入通过对应的 \hat{n}_u 个抽头延迟线记忆到输入端. 网络系统的输入和输出 $u(t) = [u_1(t), \dots, u_r(t)]^T, y(t) = [y_1(t), \dots, y_m(t)]^T$ 根据具体的情况可以确定. 由2, 3节设定的 \hat{n}_y, \hat{n}_u 可以得到多层感知器神经网络的实际输入序列^[1,2].

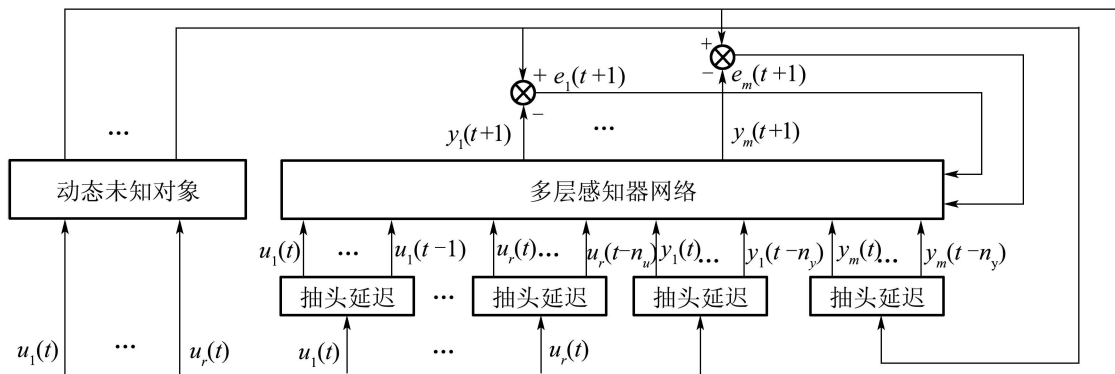


图1 NARX神经网络模型

Fig. 1 Model of NARX neural networks

目前, 隐含层和隐含节点的选择基本还是依靠设计人员的经验, 这就使得网络结构的冗余很严重^[3,4]. 采用基于Hessian矩阵的网络修剪方法, 从一个很大的具有足够解决当前问题的多层感知器(初始网络)开始, 然后通过有选择的、有序的方式削或消除某些突触权值来修剪网络, 从而使得在保持网络的良好性能的同时使网络规模最小化.

5 基于节点的解耦扩展Kalman滤波器算法 (Node-decoupled extended Kalman filter algorithm)

使用DEKF(解耦扩展的Kalman滤波器算法), 考虑到计算的简单性和收敛的快速性, 以每一个神经元节点来解耦, 从而形成NDEKF(基于节点的解耦扩展Kalman滤波器算法).

首先, 根据自适应滤波的思想建立网络状态模型, 而后线性化次模型, 可得^[5]

$$\mathbf{d}(n) = \mathbf{C}(n)\mathbf{w}(n) + \mathbf{e}(n), \quad (13)$$

其中: 噪声度量向量 $\mathbf{e}(n)$ 是一个零均值且对角协方差矩阵 $\mathbf{R}(n)$ 的多元白噪声; $\mathbf{C}(n)$ 是线性模型 $m \times W$ 的度量矩阵, 表示整个网络的输出对 W 个权值的偏微分, 即

$$\mathbf{C}(n) = \begin{bmatrix} \frac{\partial c_1}{\partial w_1} & \cdots & \frac{\partial c_1}{\partial w_W} \\ \vdots & & \vdots \\ \frac{\partial c_m}{\partial w_1} & \cdots & \frac{\partial c_m}{\partial w_W} \end{bmatrix}. \quad (14)$$

上式的矩阵中的元素仍然是应用偏导的链式法则仿照BP(误差反向传播算法)计算得出. 其中 $c_i (i = 1, 2, 3, \dots, m)$ 表示未线性化之前非线性函数 $\mathbf{c}(\mathbf{w}(n), \mathbf{u}(n), \mathbf{v}(n))$ 的第 i 个元素^[7]. 若以 s 个神经元节点来分组, 那么每一个节点在 $\mathbf{C}(n)$ 的内部作为一个独立的分矩阵 $\mathbf{C}_i(n)$. 这样矩阵 $\mathbf{C}(n)$ 就是 W 个 $\mathbf{C}_i(n)$ 的并置, 表示如下:

$$\mathbf{C}(n) = [\mathbf{C}_1(n), \mathbf{C}_2(n), \dots, \mathbf{C}_s(n)]. \quad (15)$$

那么, NDEKF可以表述为

$$\begin{cases} \boldsymbol{\alpha}_i(n) = \mathbf{K}_i(n, n-1)\mathbf{u}_i(n), \\ \boldsymbol{\Gamma}(n) = [\mathbf{R}(n) + \sum_{i=1}^s \{\mathbf{u}_i^T(n)\boldsymbol{\alpha}_i(n)\} \{\dot{\varphi}_i(n)\dot{\varphi}_i^T(n)\}]^{-1}, \\ \boldsymbol{\varsigma}(n) = \mathbf{d}(n) - \hat{\mathbf{d}}(n | n-1), \\ \hat{\mathbf{w}}_i(n+1|n) = \hat{\mathbf{w}}_i(n|n-1) + \{\dot{\varphi}_i^T(n)\boldsymbol{\Gamma}(n)\boldsymbol{\varsigma}(n)\}\boldsymbol{\alpha}_i(n), \\ \mathbf{K}_i(n+1, n) = \\ \mathbf{K}_i(n, n-1) - \{\dot{\varphi}_i(n)\boldsymbol{\Gamma}(n)\dot{\varphi}_i^T(n)\}\{\boldsymbol{\alpha}_i(n)\boldsymbol{\alpha}_i^T(n)\}. \end{cases} \quad (16)$$

其中: $\mathbf{K}(1, 0) = \delta^{-1}\mathbf{I}$, δ 是小的正常数; $i = 1, 2, \dots, s$; $\boldsymbol{\Gamma}(n) : m \times m$ 矩阵, 表示整个网络的转换因子; $\boldsymbol{\alpha}_i(n) : w_i \times m$ 矩阵, 表示第 i 个神经元的Kalman增益; $\dot{\varphi}_i(n) : m \times 1$ 向量, 表示 m 个输出对节点净输入的偏导; $\boldsymbol{\varsigma}(n) : m \times 1$ 向量, 表示线性化系统的期望响应和其估计值的差; $\hat{\mathbf{w}}_i(n+1 | n) : W \times 1$ 向量, 表示在给定直到时间 $n-1$ 的观测数据情况下, 对第 i 组权值矩阵的估计; $\mathbf{K}_i(n, n-1) : k_i \times k_i$ 的矩阵, 表示第 i 组神经元的误差协方差矩阵, k_i 是权值数.

6 辨识PMLSM模型 (Hybrid nonlinear autoregressive neural network for PMLSM identification)

6.1 试验样机和样本获取 (Experiment equipments and train data)

试验样机主要参数为: 初级电枢绕组(定子)5组 (10台); 单台定子的槽数39; 极数12; 磁体长15 mm; 极距 $\tau=24$ mm; 磁体宽80mm; 转子材料: 钕铁硼(NdFeB). 电机的额定电压为100 V, 额定电流为25 A, 额定频率为10 Hz.

训练样本包括三相电压 u_a, u_b, u_c 和三相电流 i_a, i_b, i_c 以及推力 F_x 和运行频率 f . 电源频率直接取变频器的设定频率(4 Hz). 采样周期选为1.25 ms, 每个参数均采样500点, 采样的时间区间包括电机启动直到稳态运行一段时间的全过程, 以保证建立的神经网络模型为稳态和动态的综合模型. 取三批实验数据, 分别对应于负载为15 kg, 20 kg, 35 kg. 随机将采集样本分为两部分, 其中1/3为测试集、2/3为训练集^[10].

6.2 辨识的步骤 (Identification steps)

可知电机的推力与三相电压、电流以及运行频率存在一种映射关系, 这种关系可以建立一个多输入单输出的神经网络来逼近这种映射. 输出节点数可以确定为1, 但输入节点数要根据阶次才可定出^[6,7]. 具体步骤和结果如下:

步骤1 确定模型的阶次.

辨识结果为: $\hat{n}_u = 0, \hat{n}_i = 1, \hat{n}_F = 2$; 残差的变化与阶次的关系如图2示:

步骤2 优化网络结构, 这里采用基于Hessian矩阵的网络修剪方法.

步骤3 训练所得的神经网络及算法比较.

首先, 将优化的网络初始化, 其中权值和偏置均取[0,1]的随机数. 数据均采用随机抽取的送给方式^[8,9]. 分别采取NDEKF和BP算法来对网络进

行训练,其中BP算法采用学习速率自适应和加入动量项的方法,动量常数取0.6;NDEKF算法中初始 $K(n, n-1) = 0.011$, $Q(n) = 1 \times 10^{-4}$.限定训练次数为100,分别用两种方法进行训练.误差收敛曲线如图3.由图可见,NDEKF算法在第20步就非常接近目标;而改进BP算法的误差才接近50.计算可得NDEKF的最终误差为 5.391×10^{-4} ,改进BP算法的最终误差为 4.524×10^{-3} .

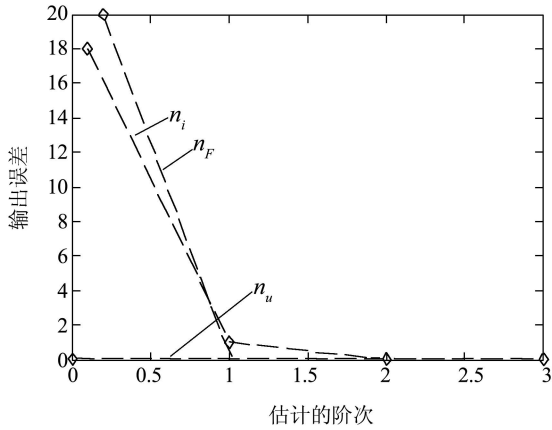


图2 结构参数辨识结果

Fig. 2 Results of structure parameter

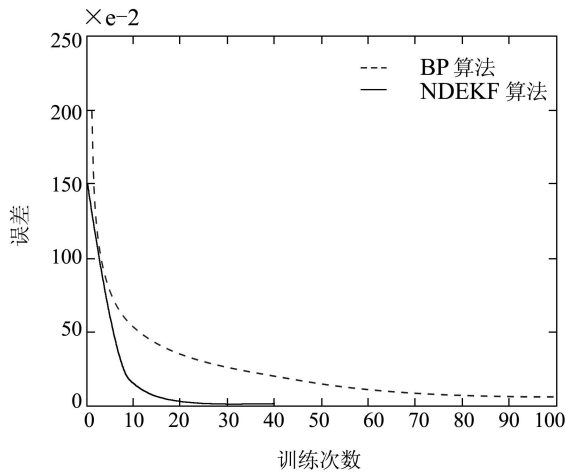


图3 BP算法与NDEKF算法学习曲线的对比

Fig. 3 Learning curves of BP and NDEKF

步骤4 模型的和输出结果分析.

电机从启动到稳定运行的推力曲线如图4所示,分别为电机实际输出推力波形和滤波之后的波形.将解析模型计算出的推力、训练后神经网络的输出推力和实测的推力加以对比,如图5所示.可见,解析模型由于是在一定的假设条件和理想状况下得出的,因此对电机的推力波动较难反映出来;而神经网络所建的模型由于是用实测的推力数据训练所得,所以能够较好的反映电机的动态过程.另外,在电机进入稳态之后,解析模型和神经网络的

输出推力均可以较好的反映电机的实际推力.

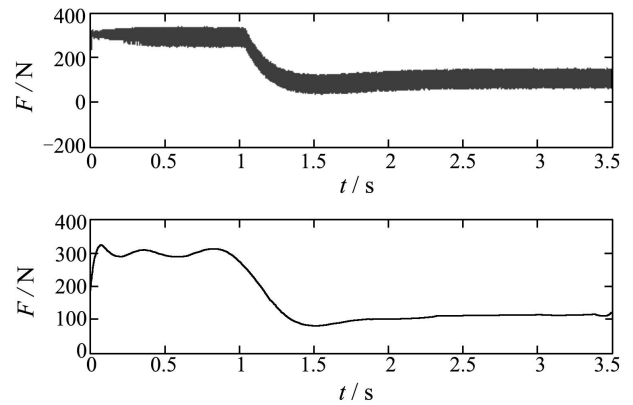


图4 永磁直线同步电动机的实际输出推力

Fig. 4 Real output force of PMLSM

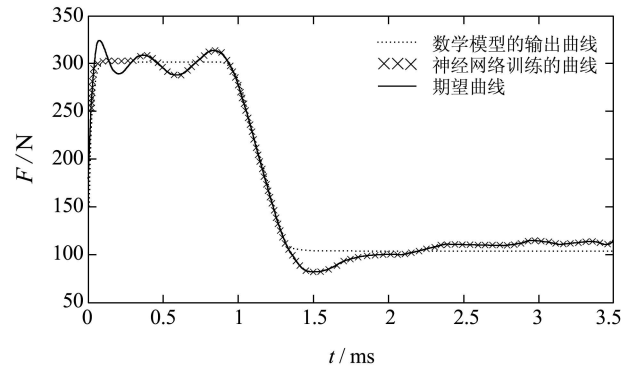


图5 两种模型输出的推力和实测推力的对比

Fig. 5 Real output forces from two models and test

7 结束 (Conclusion)

本文在相同的代价函数基础上,将动态驱动神经网络与残差分析法相结合,形成具有强逼近能力和自动识别阶次功能的混合神经网络.使用的方法和得到的结论如下:

- 1) 利用动态驱动神经网络建立的PMLSM模型,从而可对特定的PMLSM准确的建立其稳态和动态模型.
- 2) 使用混合动态驱动神经网络,克服了利用神经网络辨识非线性系统时必须知道其阶次的不足.从而使得神经网络具有识别系统阶次的能力.
- 3) 使用NDEKF算法,与改进BP算法对比具有收敛较快、泛化能力强等特点.

与在理想情况下用解析法建立的数学模型相比,这种模型更接近实际,因此更有实际应用价值.混合动态驱动神经网络辨识的推导过程不失一般性,所以该方法可以推广至更多非线性系统的情况.