

基于非线性人工神经网络模型的优化策略 及其在润滑油生产中的应用

冯恩波 俞金寿 蒋慰孙

(华东化工学院自动化所, 上海)

摘要 本文基于人工神经网络的计算特性, 提出了一种具有学习能力的在线优化策略. 这种方法不需要对优化对象进行机理建模或变量间的相关性分析, 而直接从工艺管线的拓扑结构出发建立神经网络模型. 整个优化过程采用增量跟踪运算方式, 因而能自动适应工况的变化. 文中以润滑油生产过程为例进行了仿真研究, 结果表明本方法是有效的.

关键词: 神经网络, 优化, 建模, 相关性分析, 增量跟踪运算

1 引言

溶剂脱腊是润滑油生产的主要过程, 它的优化操作具有明显的经济效益. 对脱腊工序完全用理论建模几乎是不可能的, 过去通常用回归模型来表征系统特性并进行优化计算, 然而带来了许多棘手的问题.

本文试图运用人工神经网络(NN)模型的运算特性, 来构成具有学习能力的在线优化方案. NN模型可以使我们在计算过程中同时考虑到约束的影响, 并能将人的经验知识直接结合进去. 另一特点是, 将过程NN模型学习与操作优化计算统一进行.

2 NN模型及其学习算法

神经网络有多种形式⁽¹⁻⁴⁾, 本文基于分层感知机结构(如图1), 建立过程的静态网络模型. 它由输入、输出和中间层组成, 虽然此结构会丧失某些并行特征, 但具有更简便的学习能力. 在中间层, 我们还可以完成某些非线性运算, 以此提高网络的适应性. 神经元的输入可表达成加权和式

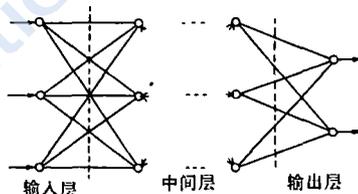


图1 分层感知机一般结构

$$n_{i,k} = \sum_j w_{i,j,k} o_{j,k-1} + t_{i,k} \quad (1)$$

其中 $t_{i,k}$ —第 k 层第 i 个神经元的阈值; $w_{i,j,k}$ —第 k 层第 i 个神经元与第 $k-1$ 层第 j 个神经元的突触联系强度; $n_{i,k}$ —第 k 层第 i 个神经元的输入; $o_{i,k}$ —第 k 层第 i 个神经元的输出. 它的输出可表示为

$$o_{i,k} = f(n_{i,k}) \quad (2)$$

一般地, 当 $f(\cdot)$ 为线性函数时, 称网络为线性的, 否则称为非线性的. 不难证明, 任何线性多层网络均可等效成两层线性神经网络, 从而不具备多层性的优点. 最普遍地, 常将

函数 $f(\cdot)$ 取为 S 形, 即

$$f(n_{i,k}) = \theta_{i,k} / (1 + e^{-\beta_{i,k} \cdot n_{i,k}}) \quad (3)$$

此处参数 $\theta_{i,k}$, $\beta_{i,k}$ 均为正数. 显然, 当 $\beta_{i,k}$ 很小时, (3) 式呈现出近似线性范围. 整个网络学习目标, 是使下式

$$E = \frac{1}{2} \sum_c \sum_i (k_i - o_{i,c})^2 \quad (4)$$

取得极小值. 此处指标 c , i 分别取遍训练数据组数和输出神经元数. 学习算法可采用最速梯度下降法

$$\Delta w_i = -\eta(\partial E / \partial w_i) \quad (5)$$

文献[6]曾指出, 对于非线性网络, 如果从较小的中间层神经元数目开始学习过程, 再逐步调整增加神经元数目, 则基本上可保证目标(4)式达到全局最小, 虽然这个问题并未从理论上彻底解决.

3 基于神经网络模型的优化策略

对于已经具有合理直接控制层的实际系统, 其静态特性可表示成控制量 u 、系统参数 p 、设定值 s_p 及随机干扰 z 的函数

$$y = \varphi(u, p, s_p, z) \quad (6)$$

过程优化, 即寻求直接控制层的设定值集合 s_p , 使得

$$\min J = \varphi(u, y, z, s_p) \quad (7)$$

基于神经网络模型的优化策略如图 2 所示.

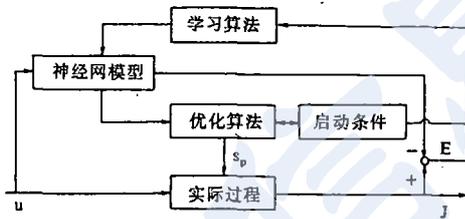


图 2 基于 NN 的优化控制策略

需要指出的是, 我们以优化目标函数作为输出量, 使得神经网络模型直接跟踪目标 $\varphi(\cdot)$. 事实上, 只要在网络中增加一个(单目标优化情况)或几个(多目标优化情况)输出神经元即可, 处理起来非常方便. 此外, 在优化算法启动以前, 需有足够的训练时间, 以使误差 E 小到一定程度.

4 神经网络学习过程的进一步讨论

神经网络的学习过程, 关键是计算出各待训练参数对目标函数(4)的梯度. 由式(4), 对于某步计算, 我们有

$$\begin{aligned} \partial E / \partial w_{i,k} &= (\partial E / \partial n_{i,k})(\partial n_{i,k} / \partial w_{i,k}); \quad \partial E / \partial t_{i,k} = (\partial E / \partial n_{i,k})(\partial n_{i,k} / \partial t_{i,k}); \\ \partial E / \partial \theta_{i,k} &= (\partial E / \partial o_{i,k})(\partial o_{i,k} / \partial \theta_{i,k}); \quad \partial E / \partial \beta_{i,k} = (\partial E / \partial o_{i,k})(\partial o_{i,k} / \partial \beta_{i,k}) \end{aligned} \quad (8)$$

由式(1)和式(2)不难得出

$$\begin{aligned} \partial n_{i,k} / \partial w_{i,k} &= o_{i,k-1}; \quad \partial n_{i,k} / \partial t_{i,k} = 1; \quad \partial o_{i,k} / \partial \theta_{i,k} = o_{i,k} / \theta_{i,k}; \\ \partial o_{i,k} / \partial \beta_{i,k} &= o_{i,k}(1 - o_{i,k} / \theta_{i,k})n_{i,k}; \quad \partial o_{i,k} / \partial n_{i,k} = o_{i,k}(1 - o_{i,k} / \theta_{i,k})\beta_{i,k} \end{aligned} \quad (9)$$

如果学习过程从输出层向输入层逐层进行, 则可直接利用后一层的计算结果, 为此先考虑输出层(设为第 n 层), 由学习目标及式(9)可得

$$\begin{cases} \frac{\partial E}{\partial n_{i,n}} = \left(\frac{\partial E}{\partial o_{i,n}}\right)\left(\frac{\partial o_{i,n}}{\partial n_{i,n}}\right) = -(k_i - o_{i,n})o_{i,n}\left(1 - \frac{o_{i,n}}{\theta_{i,n}}\right)\beta_{i,n} \\ \frac{\partial E}{\partial o_{i,n}} = -(k_i - o_{i,n}) \end{cases} \quad (10)$$

对于第 k 层的修正量, 则有

$$\frac{\partial E}{\partial o_{i,k}} = \sum_j (\frac{\partial E}{\partial n_{j,k+1}})(\frac{\partial n_{j,k+1}}{\partial o_{i,k}}) \quad (11)$$

此处 j 取遍第 $k-1$ 层的神经元个数, 再由式(1), 上式易写作

$$\frac{\partial E}{\partial o_{i,k}} = \sum_j (\frac{\partial E}{\partial n_{j,k+1}})w_{ji,k+1} \quad (12)$$

类似地, 可推得

$$\frac{\partial E}{\partial n_{i,k}} = (\frac{\partial E}{\partial o_{i,k}})o_{i,k}\left(1 - \frac{o_{i,k}}{\theta_{i,k}}\right)\beta_{i,k} \quad (13)$$

利用式(8)~(13)可立即算出各参数的修正量, 值得注意的是, 在第 k 层的修正量计算中, 计算过程仅与第 $k+1$ 层有关, 这意味着, 可采用并行运算方式, 以加快计算速度, 此点于在线应用时具有重要意义。

5 润滑油溶剂脱蜡过程优化方案

5.1 工艺简况

某大型炼油厂的脱蜡过程如图 3 所示, 图中共有五路并联的原料输入和结晶装置, 最后混合并加入第三次溶剂, 经冷洗过滤, 得到脱蜡油。

工艺操作目标要求脱蜡油的收率最大, 并且要反映凝固点的温度 T_3 在一定范围内, 操作经验告诉我们, 影响收率的主要因素是溶剂比、原料量、冷洗比和温度 T_1, T_2 , 其中处理量决定于前面的工段, 故只能作为优化时的参变量, 而非优化变量。

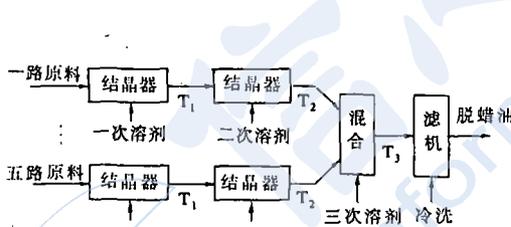


图 3 脱蜡过程工艺简图

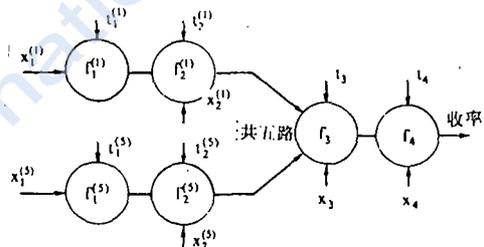


图 4 脱蜡过程 NN 模型结构

5.2 脱蜡过程 NN 模型和优化步骤

由工艺要求, 图 3 中的温度 T_1, T_2 可不予调整, 在优化时一般仅调整一次溶剂、二次溶剂、三次溶剂、冷洗溶剂与原料量的比值, 以下分别称之为一次比、二次比、三次比和冷洗比, 并记为 $x_i (i=1, 2, 3, 4)$ 。如前所述, 可直接从工艺管线的结构出发来构造 NN 模型, 如图 4。

实际工况表明, 五路操作条件差别不大, 为进一步简化网络结构, 减少待学习参数, 将五路的一、二次比均值作为每路操作值, 且假定各路工况相同, 则图 4 的结构可再简化为四层单神经元结构, 如图 5 所示。

从图 5 可以看出, 这个 NN 模型的特点是每一个神经元都有外界输入, 即它不仅作

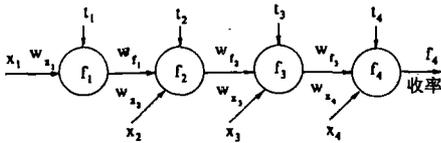


图5 脱蜡过程 NN 模型简化结构

式(2)中的参数 β 并入到加权量 w 中, 则可把整个模型表成下列递归关系式

$$f_{i+1} = \theta_{i+1} / (1 + e^{-n_{i+1}}) \quad (14a)$$

$$n_{i+1} = w_{x_{i+1}} x_{i+1} + w_{f_i} f_i + t_{i+1} \quad (14b)$$

$$f_0 = 0, w_{f_0} = 0, i = 0, 1, 2, 3$$

模型的学习目标依然取为二次型

$$\min E = (1/2)(f_4 - k)^2 \quad (15)$$

式中 k 是实际测量到的收率值. 显然, 对某参数的梯度学习算法一般地可写为

$$Q^{(k+1)} = Q^{(k)} - \eta^{(k)} \left(\frac{\partial E}{\partial Q} \right)^{(k)} \left(\frac{\partial f_4}{\partial Q} \right)^{(k)} = Q^{(k)} - \eta^{(k)} \left(\frac{\partial f_4}{\partial Q} \right)^{(k)} (f_4^{(k)} - k) \quad (16)$$

此处 $Q^{(k)}$ 表示诸参数 θ, w, t 中任一个的第 k 步计算值.

由式(16)看出, 我们的主要困难在于如何得到 $(\partial f_4 / \partial Q)$ 的可计算表达式. 为此给出几个由网络(14)决定的关系式, 它们的证明是容易的.

关系式1 $(\partial f_j / \partial n_j) = f_j - f_j^2 / \theta_j, j = 1, 2, 3, 4$

关系式2 $(\partial f_j / \partial \theta_j) = f_j / \theta_j, j = 1, 2, 3, 4$

关系式3 $(\partial f_{j+1} / \partial f_j) = (f_{j+1} - f_{j+1}^2 / \theta_{j+1}) w_{f_j}, j = 1, 2, 3$

证明 由关系式1, 再注意到 $(\partial n_{j+1} / \partial f_j) = w_{f_j} (j = 1, 2, 3)$, 即得.

关系式4 $(\partial f_4 / \partial \theta_j) = \prod_{i=j}^3 (f_{i+1} - f_{i+1}^2 / \theta_{i+1}) w_{f_i} (f_j / \theta_j), j = 1, 2, 3$

证明 因 $(\partial f_4 / \partial \theta_j) = \prod_{i=j}^3 (\partial f_{i+1} / \partial f_i) (\partial f_j / \partial \theta_j)$, 由关系式2和3即得.

关系式5 $(\partial f_4 / \partial w_{x_j}) = \prod_{i=j}^3 (f_{i+1} - f_{i+1}^2 / \theta_{i+1}) w_{f_i} x_j, j = 1, 2, 3, 4$

证明 因 $(\partial f_4 / \partial w_{x_j}) = \prod_{i=j}^3 (\partial f_{i+1} / \partial f_i) (\partial f_i / \partial n_i) (\partial n_i / \partial w_{x_j})$, 由关系式3和1,

并注意到 $(\partial n_j / \partial w_{x_j}) = x_j$, 立即得证.

关系式6 $(\partial f_4 / \partial w_{f_j}) = \prod_{i=j}^3 (f_{i+1} - f_{i+1}^2 / \theta_{i+1}) w_{f_i} f_j, j = 1, 2, 3$

证明 因 $(\partial f_4 / \partial w_{f_j}) = \prod_{i=j}^3 (\partial f_{i+1} / \partial f_i) (\partial f_j / \partial n_j) (\partial n_j / \partial w_{f_j})$, $j = 1, 2, 3$, 利用关系式3和1, 并注意到 $(\partial n_j / \partial w_{f_j}) = f_j$, 得证.

关系式7 $(\partial f_4 / \partial t_j) = \prod_{i=j}^3 (f_{i+1} - f_{i+1}^2 / \theta_{i+1}) w_{f_i}, j = 1, 2, 3, 4$

证明 与关系式5同理, 由式(14), 立即得证.

事实上, 将关系式4, 5, 6, 7代入式(16)中, 就得到了任一个参数的梯度学习算法.

为计算层(中间层)或输出层, 而且同时作输入层, 这是与图1所示网络结构的不同之处.

现在讨论它的学习算法. 若图5中各神经元的输出函数 $f(\cdot)$ 均取为S型, 并将

值得指出的是,由上边诸关系式形成学习算法,使得某层的学习过程只用到且仅用到后层的信息,这使得我们可以利用少数样本输入,对某一层的某一个参数进行反复学习,直到目标函数不再减小。事实上,这种学习算法在结构上是并行的,这不仅能加快运算速度,而且更重要的是,利用较少的样本输入就可同时使全部参数达到稳定状态,此点是传统数据回归方法所不具备的。

有了上面的 NN 模型以后,可依下列程序进行优化控制。

(1) 给网络模型赋初值; (2) 采集过程数据,并作为 NN 模型训练样本; (3) 运行网络学习程序; (4) 判断 NN 模型精度,若不满足要求,转(3),否则继续; (5) 启动优化子程序,以模型的输入变量作为优化变量; (6) 结合约束条件和经验知识对实际过程施加控制; (7) 转向(2)。

6 仿真研究

为说明本文方法的有效性和便于结果比较,我们以某炼厂脱蜡过程数学模型⁽⁷⁾作为实际过程静态特性模拟器,整个系统采用增量运行方式,即边对网络参数进行学习,边进行输入参数的优化,其结果又反馈到过程模拟器,以产生新的过程数据。此过程如图 6 所示。图 7 是仿真结果,其中曲线 1, 2 是以学习次数为横坐标,从此可以看出实际收率的变化和 NN 模型的学习和跟踪过程,曲线 3 是以时间为横坐标(单位为控制次数)的实际收率变化,此曲线显示出,经过约 40 个控制周期作用,收率事实上已达到极限稳定状态。

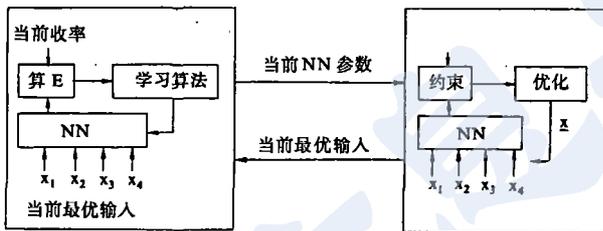


图 6 增量运行示意图

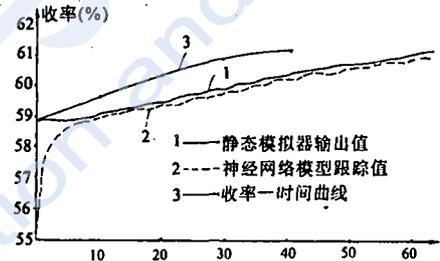


图 7 仿真结果

参 考 文 献

- 1 Hopfield J J, Tank D W. Neural Computation of Decisions Optimization Problems. Biol Cyber, 1985; 52:141-152
- 2 Rosenblatt R. Principle of Neurodynamics. New York: Spartan Books, 1961
- 3 Rumelhart D E, McClelland J L. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. I. II, MIT Press, 1986
- 4 Psaltis D et al. A Multilayered Neural Network Controller. IEEE Contr Syst Mag, 1988; (4):17-21
- 5 Kohonen T. Associative Memory: A System Theoretical. Approach. New York: Springer, 1977
- 6 Rangwala S S, Dornfeld D A. Learning and Optimization of Machining Operations Using Computing Abilities of Neural Networks. IEEE Trans on Syst Man and Cyber, 1989; 19(2): 299-314
- 7 刘喜梅, 冯恩波, 俞金寿. 大型炼油厂酮苯脱蜡过程模型开发. 第六届全国系统与控制科学青年讨论会, 杭州, 1990