

软测量模型的建立

徐文骞, 曾文华

(杭州电子工业学院计算机分院, 浙江 杭州 310037)

摘要:软测量是采用过程中比较容易测量的辅助变量,构造推断估计器来推算出难以测量或根本无法测量的关键工艺参数;也就是说软测量是根据某种最优准则,选择一组既与主导变量有密切联系又容易测量的变量,即辅助变量,通过构造某种数学关系,用计算机软件实现对主导变量的在线估计。本文介绍了软测量的核心技术,并重点阐述了软测量模型建立的方法,同时还给出了建立一个完整的软测量模型的步骤。

关键词:软测量;反向传播网络;径向基函数网络;部分最小二乘

中图分类号:TP391.4

文献标识码:A

文章编号:1001-9146(2002)01-0016-05

0 引言

在过程控制中,存在一个重要问题就是对过程输出变量的估计,这些变量难以用传感器进行检测,如石油产品的某些质量指标,化学反应的反应深度,反应物的浓度分布及催化剂的活性等。而另一方面,这些变量的实时检测对于保证产品质量,生产装置的连续平稳操作以及充分发挥装置的生产能力都具有举足轻重的作用。为了解决这个问题,一般可以采用两种方法:一是间接测量法,利用可测辅助变量去估计难以测量的变量。二是利用在线分析仪直接测量所需参数。但是它需要很大的投资,并且往往存在测量滞后。在第一种方法中利用可测辅助变量来推断不可测变量正是软测量的核心。软测量就是采用过程中比较容易测量的辅助变量(Secondary Variable),构造推断估计器来推算出难以测量或根本无法测量的关键工艺参数。换句话说,软测量就是根据某种最优准则,选择一组与主导变量(Primary Variable)有密切联系又容易测量的变量,即辅助变量,通过构造某种数学关系,用计算机软件实现对主导变量的在线估计^[2]。

1 软测量技术简介

软测量技术主要包括 4 个方面:

1.1 输入变量的选择

目前软测量对象都是灰箱系统,得不到系统的准确模型。一般来说,原始辅助变量数目很多,往往有数十个,并且相关程度差异较大,为了实时运行方便,有必要对输入变量进行适当的降维处理。一种方法是通过机理分析的方法,找到那些对被测变量影响大的相关变量;另一种方法是采用主元分析,部分最小二乘法等统计方法进行数据相关性分析,剔除冗余的变量,降低系统的维数。也可以将两者结合

收稿日期:2001-10-22

作者简介:徐文骞(1976-),男,江苏镇江人,在读研究生,智能控制。

起来,利用机理分析方法将一些变量作为预选变量,运用相关性理论选取相关性大的作为输入变量。

1.2 输入数据处理

软测量在线运行时,其输入为DCS(集散型计算机控制系统)中采集到的现场测量值和被测变量的人工化验值,而它们常因自身特点或外部干扰不能直接作为软测量器的输入。因此输入数据的预处理是必要的。数据预处理包括数据变换和误差处理,数据变换包括标度转换和权函数3个方面。数据变换不仅直接影响着过程模型的精度和非线性映射能力,而且影响着数值优化算法的运行效果。数据误差包括随机误差和过失误差两类。对于随机误差除剔除跳变信号外,常采用滤波的方法来解决。过失误差将大大影响软测量的在线运行精度,因此及时侦破剔除和校正这类数据是误差处理的首要任务,其常用的方法有统计假设校验法、广义似然法和贝叶斯法等。

1.3 软测量模型的建立

从软测量技术的发展过程可以看出,软测量技术的核心是建立对象的数学模型,对象数学模型的好坏,将直接关系到软测量器的计算结果。本文将详细讨论软测量模型的建立。

1.4 软测量模型的在线校正

由于工业对象并非一成不变的,在长期运行中对象特征会因操作条件变化,原料改变、装置改造等原因而发生改变,从而导致软测量器的精度下降。因此软测量器必须具备在线校正功能,以使软测量器能够跟踪系统的变化,提高模型的适应性。软测量模型的在线校正可表示为模型结构和模型参数的优化过程。为解决模型结构修正耗时长和在线校正的矛盾,提出了短期学习和长期学习的校正方法。短期学习是指利用最近一次的离线测量值与估计值的误差对当前估计值进行修正。长期学习是指由于对象特性发生很大改变,软测量器已无法通过短期学习来保证精度,必须依赖运行中所积累的历史数据重新构建模型。

2 软测量的建模

数学建模是软测量器的核心,一个好的软测量器必须具备一个完善的实用的数学模型。建立对象的数学模型主要有以下几种:

2.1 基于机理分析的方法

机理模型是基于对被测对象的深刻认识,通过对被测对象的机理分析,找出不可测主导变量和可测辅助变量之间的关系,以数学表达式的形式进行计算。从理论上来说,机理模型是最精确的模型,然而它要求对被测对象的内部特性完全了解。由于实际工业过程的复杂性,往往难以完全通过机理分析得到软测量模型。因此,基于机理分析的方法建模非常困难,需要与其他方法配合使用。事实上,机理分析对于其他任何一种方法也都是必需的。

2.2 基于控制对象动态模型的方法

由于软测量技术是控制工程的一个分支,因此在软测量技术发展之初,有很多是使用控制中的建模或辨识等方法来建立系统的动态模型,以此来设计软测量器。

如果假设系统的状态空间模型为:

$$\begin{cases} \dot{x} = Ax + Bu + Ev \\ y = Cx \\ \theta = C_0x + W \end{cases} \quad (1)$$

式中: x 为过程的状态向量, y 和分别为过程的主导变量和辅助变量, u 为控制变量, v 和 w 为白噪声。

并且系统的状态对于辅助变量完全可观,那么软测量问题就转化为经典的状态观测和状态估计问题,用Kalman滤波器可以从辅助变量得到主导变量的估计值^[3-4]。

2.3 基于回归方法

这类方法就是通过对生产过程历史数据的回归分析,建立质量指标的软测量模型,根据所采用的数

学方法的不同,又可以将回归方法分为线性回归和非线性回归。

2.3.1 线性回归方法

在线性回归中,多变量线性回归(MLR)应用是最广泛的,主元回归法(PCR)和部分最小二乘法(PLS)都是从MLR方法中派生出来的。

MLR方法的目标是建立一个从 m 个不相关可测变量 $x_i(i=1,2,\dots,m)$ 到估计变量 y 之间的线性映射。如果考虑一个有 n 组样本的MLR回归问题,可用下式表达:

$$Y = Xb + b_0 \quad (2)$$

式中: Y 是 n 维的样本向量, X 是 $n \times m$ 维的输入变量测量值, b 是 m 维线性系数向量; b_0 是 n 维常数向量。

用最小二乘估计法,可以推导出回归系数的估计公式,从而使

$$Q = (Y - Xb)^T(Y - Xb) \quad (3)$$

为最小值。要估计 b 并使 Q 值最小,可以对 b 求导,

$$\frac{\partial Q}{\partial b} = -2X^T Y + 2X^T Xb \quad (4)$$

令导数为零,就可以得到:

$$b = (X^T X)^{-1} X^T Y \quad (5)$$

显然,该式所表达的MLR问题的求解取决于矩阵 $X^T X$ 的逆阵是否存在。当 X 中包含有多余的信息时,也就是输入变量存在线性相关变量, X 就会变为病态矩阵,MLR方法就无法应用了。因工业数据中变量相关还是比较普遍的,此时需要采用部分最小二乘法(PLS)。最小二乘法是一种将高维空间信息投影到由很少量特征向量组成的低维空间的多变量回归方法,这些特征向量包含了原始数据序列中所有的重要信息。最小二乘法对于克服由于数据之间的严重关联和共线性问题,有较明显的优势。1984年,Wold等人提出了部分最小二乘法(PLS)^[7]。顾名思义,部分最小二乘法和标准最小二乘法有着一定的联系。线性回归只是简单地求取各输入向量的线性组合关系,可以用以下的形式描述,假设:

$$\hat{Y} = Xr \quad (6)$$

式中: \hat{Y} 是 n 维的样本向量预测值, X 是 $n \times m$ 维的输入变量测量值; r 是 m 维线性系数向量。

通常最小二乘法使用的目标函数是使实测值 Y 和预测值 \hat{Y} 的方差最小,即:

$$\min(Y - \hat{Y})^T(Y - \hat{Y}) = \min(Y - Xr)^T(Y - Xr) \quad (7)$$

在最小二乘法中,对回归变量 r 是毫无限制的,即 $r \in \mathbf{R}^m$, m 为输入变量的维数。而在PLS中,回归系数限制在由输入变量矩阵的列数 \mathbf{R}^m 决定的一个 d 维子空间($d < m$)。因此,相对于式6,它的形式为:

$$\hat{Y} = XWb \quad (8)$$

式中: $W \in \mathbf{R}^{m \times d}$, $b \in \mathbf{R}^d$ 。

如果所有的方向都考虑到, $d = m$,那么 W 的列向量将充满输入空间,PLS方法和MLR方法就统一了。在这里,充成一个空间的意义与其在线性代数中的意义略有不同,从数学观点严格来说,输入数据总能张成输入空间。但是,实际上这种分布很大程度上是噪声,如果可以得到无噪声的数据,就会发现实际上这些数据并没有张成输入空间。于是,可以将这些数据压缩到低维的空间上,如果方法正确,多余的信息可以用来减轻噪声的干扰。PLS方法的关键在于目标函数被用来决定特征变量的方向数及矩阵 W 的列数。由于回归系数被分成两个部分(W 和 b),这两个参数集可以用来满足不同的目标。将式7展开:

$$\min(\hat{Y}^T \hat{Y} - 2Y^T \hat{Y} + Y^T Y) = \min(r^T X^T Xr - 2Y^T Xr + Y^T Y) \quad (9)$$

通过对目标函数的进一步考察,式中最后一项即输出变量的方差不会受到回归变量的影响。因此标准最小二乘法的目标函数由两个子优化目标组成,即减小模型预测值方差的同时增大输出量与预测值的协方差。标准最小二乘法确实可以减小模型预测值的方差,但由于对 r 的大小毫无限制,它不能有效地增大输出量与预测值的协方差。如果用式8选择参数进行分解的方法:选择方向矩阵 W 使输出量与预测值的协方差最大,同时选择 b 直接减小模型预测值的方差。实际也证明通PLS方法的确可以做

到在减小模型预测值的方差的同时增大输出量与预测值的协方差。

线性回归的实质是实际对象函数关系在操作点附近忽略了其高阶项的一阶泰勒展开式。但是如果对象的非线性特性比较严重,或者对象的操作点变动范围比较大,泰勒表达式中的高阶项就不能被忽略了。此时线性回归方法建立的模型精度就会显著下降,这样我们就应该选用非线性回归的方法。

2.3.2 非线性回归方法

非线性回归方法主要使用基于人工神经网络的方法。人工神经网络在近 10 年来发展非常迅速,在过程控制领域的应用也越来越广泛。这类方法就是将辅助变量作为人工神经网络的输入,主导变量作为其输出,通过网络学习来解决不可测变量的软测量问题。由于实际工业过程中,许多对象具有复杂的不确定性、实时性和高度的非线性,这必然造成根本无法或者很难精确建模,然而人工神经网络具有逼近任意非线性关系的能力,可以很好地解决这些问题。目前许多化工过程软测量建模都是采用基于神经网络的方法来实现,常用的有反向传播(Back Propagation)神经网络和径向基函数(Radial Basis Function)神经网络。BP 网络不仅有输入层节点、输出层节点,而且有一层或多层隐含节点。对于输入信息,要先向前传播到隐含层的节点上,经过各单元的激活函数运算后,把隐含节点的输出信息传播到输出节点,最后给出输出结果。网络的学习过程由正向和反向传播两部分组成。在正向传播过程中,每一层神经元的状态只能影响下一层的神经网络。如果输出层不能得到期望的输出,那么转入反向传播过程,将误差信号沿原来的连接通路返回,通过修改各层神经元的权值,逐次地向输入层传播去进行计算。经过正向反向这两个过程的反复运用,使得误差信号最小,当误差达到要求时,学习过程就结束了。

径向基函数神经网络是单隐层的前向网络,网络输出权值、基函数的中心、隐含单元数目需学习确定。基函数的中心可以采用 K-means 法、竞争学习等方法^[8],网络输出权值可以采用正交最小二乘法 OLS(Orthogonal Least Squares)^[9]来实现。而网络隐含层函数类型要针对具体问题选用不同的径向基函数。RBF 网络的主要学习算法有 Moody 与 Darken 算法,局部训练算法,正交优选算法,聚类算法与 Givens 最小二乘联合迭代法等。由于 RBF 网络的学习算法不存在局部最优问题,具有全局逼近性质,而且它的参数调整是线性的,所以训练方法快速易行,非常适合于动态系统的实时辨识。

此外 RBF 网络采用局部逼近的方法,因而比 BP 网络具有更好的适应性,RBF 网络在一定程度上解决了 BP 网络训练时间长和存在局部最小的问题,因而目前在建模研究中应用很多。但是由于 BP 网络本身具有很好的逼近非线性映射的能力,而且结构与算法容易实现,在大量的实际应用中不断被改进和完善,因此基于 BP 网络的软测量模型仍旧是实际工程中常用的方法。除上述线性和非线性回归方法外,还可以考虑将线性回归方法与非线性回归方法结合起来,在前面提到的 PLS 线性回归方法能抽取原始数据中的特征数据,因此可以去掉相关的数据信息,且对关联数据序列具有较好的补偿作用。但是 PLS 法仅能从原始数据中抽取线性信息,而许多实际对象数据却具有非线性特性。基于 ANN 的非线性回归方法能近似任何非线性特征,但当输入数据存在关联时,神经网络方面则会导致模型的过拟合现象,从而出现训练精度较高,但测试精度却很低的现象。但若将 PLS 的抽取特征向量和神经网络的非线性近似能力有效地组合,则既能近似非线性特征,又能处理相关联的输入数据。

3 软测量模型的设计步骤

针对一个具体的工业对象设计一个成功的软测量模型大致可分为以下几个步骤:

(1) 对主导变量进行机理分析,寻找辅助变量。目前大多数的软测量对象都属于灰箱系统,可以通过机理分析找到被测变量的相关变量,将其中的可测变量暂时定为原始辅助变量,以及这些变量对主导变量影响的大小。同时还要决定系统是否需要降维操作,以减少辅助变量的个数,这里往往采用一些数学和机理分析相结合的方法来确定如何进行系统降维;

(2) 收集工业数据。收集被测量的主导变量的实测值和既定的原始辅助变量在相应时刻的测量数据。数据不仅可以用来研究对象的特性,还可以用来训练模型和检验模型。如果系统需要降维,还需要计算出降维后的辅助变量值;

(3) 对数据进行预处理,因为在一个生产过程中可能包括不同的工况,如果对不同的工况用同一个软测量模型加以训练,拟合精度会大大降低,这时就需要对数据进行分类处理。同时还要对在线输入数据进行一些其他的处理,如剔除人为测量误差造成的“野点”等等;

(4) 分析对象特性,选择相应的软测量模型。可以用 MLR PLS 和 ANN 等方法分别拟合,以观察各自的效果,选择出既简单又有效的方法作为软测量模型;

(5) 离线训练模型。将已知的工业数据分为两个部分,一部分用来训练模型,另一部分用来检验模型。如果模型的检验结果达到设计要求,则转入下一步,否则修改模型的结构;

(6) 为模型设计完善的在线学习功能,使模型的精度在投入运行后能得到长期的保证;

(7) 软测量器的实现。软测量器是以软件形式构建在不同的 DCS 平台上,必须根据平台的要求编制相应的软件。

4 结束语

软测量技术是通过数学模型来估计工程上难以检测的变量值,也即通过建立过程数学模型来实现对主导变量的在线估计。因此,软测量技术核心是建立软测量模型以实现二次变量对主导变量的最佳估计。本文介绍了软测量的核心技术,以及建立软测量模型的方法。软测量是一门新兴的工业控制技术,随着软测量技术在模型构建上的不断完善,其在工业控制中将得到更广泛的运用。

参考文献:

- [1] 袁曾任. 人工神经网络及其应用[M]. 北京:清华大学出版社,1999.66-130.
- [2] 于静江,周春晖. 过程控制中的软测量技术[J]. 控制理论与应用,1996,21(4):139-143.
- [3] Guilandoust Mtet al. Adaptive Inferential Control[J]. IEEE Trans. on Neural Networks, IEE Proceeding - D, 1987,134(3):171-179.
- [4] Tham MTet al. Soft - Sensors for Process Estimation and Inferential Control[J]. IEEE Trans. on Neural Networks, 1991,1(1):3-14.
- [5] 徐敏,俞金寿. 软测量技术[J]. 石油化工自动化,1998,13(2):1-3.
- [6] 王旭东,邵惠鹤. 神经网络建模与软测量技术[J]. 化工自动化及仪表,1996,23(2):703-706.
- [7] Wold S, et al. The Collinearity problem in linear regression: the partial least squares approach to generalized inverses. SIAM J. Sci. Stat. Comput. 5, 1984: 743-753.
- [8] 王旭东,邵惠鹤. 基于神经网络的通用软测量技术[J]. 自动化学报,1998,24(5):702-706.
- [9] Chen S, Cowan C F N, Grant P M. Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks[J]. IEEE Trans. on Neural Networks, 1991,2(2):302-309.
- [10] 孙欣,王金春,何声亮. 基于神经网络的过程软测量[J]. 自动化仪表,1996,17(9):7-10.
- [11] 朱军. 线性模型分析原理[M]. 北京:科学出版社,1999.20-66.
- [12] 汪永生,邵惠鹤. 通用软测量软件包的开发与应用[J]. 化工自动化及仪表,2000,27(2):34-36.

The Design of Soft Sensor Model

XU Wen-qian ZENG Wen-hua

(School of Computer, Hangzhou Institute

Of Electronics and Engineering, Hangzhou Zhejiang 310037, China)

Abstract: The technique of Soft Sensor can assess the variable that is difficultly measured by designing the model of math. This paper give a general idea about the kernal technique about Soft Sensor, concentrating on the method of how to design the Soft Sensor model, and giving a process of designing a integrated Soft Sensor system.

Key words: Soft Sensor; BP; RBF; PLS