

一种适用于流程工业预测控制的模型自主学习方法

技术领域

本发明涉及流程工业预测技术领域，具体涉及一种适用于流程工业预测控制的模型自主学习方法。

背景技术

自 20 世纪末期以来，流程工业得到了快速发展，其中也包括模型预测控制在流程工业中的成功应用。模型预测控制与传统的 PID 控制相比，不仅适用于多耦合、大时滞等难控的被控对象，而且可以对工艺条件和经济指标等约束建模，配合优化技术完成控制任务。故，模型预测控制在流程工业中具有广泛的应用前景。然而，在实际生产中，随着生产目标改变、生产原料配比改变、设备老化、催化剂活性不足等多种因素，过程的动态特性会发生变化。这种变化会引发预测控制的预测模型与真实过程动态特性的失配问题。因此，如果能增加预测模型与真实过程动态特性的匹配程度，则控制器可以实现更加准确的过程输出预测，进而有利于控制效果的提升。目前，行业已有一些思想或者方法来应对变化的动态特性问题，比如非线性建模、多工况建模、在线辨识等方法。但现有的方法在实施时存在以下不足。

第一点，对于非线性建模和多工况建模。流程工业大量的三传一反过程增加了机理建模的难度，且现场运行数据难以覆盖所有工况，因此数据驱动的建模方法也只能建立局部模型或者部分工况模型，如果投运时运行在未知工况，仍然需要一种模型在线学习的方法。

第二点，对于在线辨识的方法。一方面，现有方法多采用线性回归对动态特性建模，参数缺乏实际的物理意义，辨识时也缺少对参数的合理约束，导致辨识参数的自由度过高；另一方面，过程的动态特性往往是缓慢变化的，当前正在运行的预测模型应该对新的动态特性参数具有指导意义。然而，暂未发现能有效兼顾这两方面的技术方法。

发明内容

本发明为了克服以上技术的不足，提供了一种适用于流程工业预测控制的模型自主学习方法。

该方法的特点在于：第一，可在线自主学习过程的动态特性，以多输入单输出子系统划分的方式可适用于多输入多输出系统，当然也适用于比多输入多输出系统更简单的其他系统；第二，可根据预设的模型参数自动选择待学习参数的类型，使待学习参数具有真实的物理意义，且在学习过程中，预设模型对模型学习的结果具有约束作用，以提高模型的可靠性；第三，对参数的学习结果进行检验，对模型的预测结果进行评价，进一步提高模型的可靠性；第四，采用参数死区的方式增加了模型参数的平稳性，减少因模型参数频繁变化而引起的控制效果波动。本发明的方法是一种融合模型知识和过程数据的在线学习方法，适用于流程工业预测控制的模型自主学习问题。

术语解释：

- 1、增益：放大倍数。
- 2、自然角频率：描述物体振动快慢的物理量，与振动系统的固有属性有关。
- 3、阻尼系数：表征二阶以及二阶以上的系统在运动过程中系统能量减少这一特性的参数。
- 4、对数正态分布：一个随机变量的对数服从正态分布，则该随机变量服从对数正态分布。
- 5、预测模型：指预测控制中用于预测的，用数学语言或公式所描述的事物间的数量关系。
- 6、参数死区：指参数变化时，如果变化幅度未超过该值，则认为参数没有发生变化。

本发明克服其技术问题所采用的技术方案是：

一种适用于流程工业预测控制的模型自主学习方法，通过多输入单输出系统作为基本子系统予以实现，包括如下步骤：

步骤一、模型设置及学习参数自选择：设置所有传递函数模型，对每个传

递函数模型进行计算并自动选择待学习参数。

其中，设置传递函数模型具体包括：

在固定的二阶系统加一阶积分加纯滞后传递函数模型模板下，设置每个输入变量到每个输出变量之间的传递函数模型。

假设传递函数模型有 M 个输入变量和 P 个输出变量，则需要按照式 (1) 所示的模板设置 $M \times P$ 个传递函数模型，且额外设置标志位 $\text{Int}_{p,m}$ 以说明该传递函数模型 $G_{p,m}$ 是否额外串联一个一阶积分环节；

$$G_{p,m} = \frac{K_{p,m} e^{-\tau_{p,m}s}}{a_{p,m}s^2 + b_{p,m}s + 1}, a_{p,m} \geq 0, p=1,2,\dots,P, m=1,2,\dots,M \quad (1)$$

上式中， $K_{p,m}$ 为模型增益， $\tau_{p,m}$ 为延迟时间， s 为拉式算子， $a_{p,m}$ 和 $b_{p,m}$ 分别为传递函数特征多项式的二次项系数和一次项系数；

然后，对每个传递函数模型进行计算并自动选择待学习参数，*表示模型自学习前设定的模型参数值，具体包括：

先对每个传递函数模型计算，然后进行如下分析：

如果某个传递函数模型的增益 $K_{p,m}^* = 0$ ，则该传递函数模型 $G_{p,m}$ 不参与模型在线学习；

否则，判断二阶系统的极点位置：（1）如果存在实部为正的根或者纯虚数根，则该传递函数模型 $G_{p,m}$ 也不参与模型在线学习；（2）如果存在两个实部为负数的共轭复数根，则计算传递函数模型 $G_{p,m}$ 的自然角频率 $\omega_{n,p,m}^*$ 和阻尼系数 $\zeta_{p,m}^*$ ，并将模型增益 $K_{p,m}$ 、自然角频率 $\omega_{n,p,m}$ 和阻尼系数 $\zeta_{p,m}$ 作为待学习参数。

除了上述（1）和（2）两种情况外，如果传递函数模型的分母多项式 $a_{p,m}^*s^2 + b_{p,m}^*s + 1$ 可以实现式（2）的因式分解，则将非零的一次项系数 $T_{1,p,m}$ 、非零的一次项系数 $T_{2,p,m}$ 和模型增益 $K_{p,m}$ 作为传递函数模型 $G_{p,m}$ 的待学习参数。

$$a_{p,m}^*s^2 + b_{p,m}^*s + 1 = (T_{1,p,m}^*s + 1)(T_{2,p,m}^*s + 1), T_{1,p,m}^* \geq 0, T_{2,p,m}^* \geq 0 \quad (2)$$

步骤二、参数分布建模：对于参与学习的每个传递函数模型，采用对数正

态分布对待学习参数建立对数正态分布模型。具体包括如下：

在传递函数模型 $G_{p,m}$ 参与学习的情况下，采用对数正态分布 $p(\theta|\mu_\theta, \sigma_\theta^2)$ 对模型参数值的分布建立对数正态分布模型，即使得传递函数模型参数服从对数正态分布，其中， μ_θ 和 σ_θ^2 是关于参数 θ 对数正态分布的分布参数；根据步骤一所述的待学习参数的种类，由用户设定与分布形状相关的系数 $\{\beta_\theta|\beta_\theta>1\}$ ， $\theta \in \{K, T_1, T_2, \omega_n, \zeta\}$ ， $\theta_{p,m}^* \in \{K_{p,m}^*, T_{1,p,m}^*, T_{2,p,m}^*, \omega_{n,p,m}^*, \zeta_{p,m}^*\}$ 为设定的传递函数模型的参数值，根据式（3）解算对数正态分布的参数。

$$\begin{cases} E(\theta) = \beta_\theta \times \theta_{p,m}^* \\ \arg \max(p(\theta)) = \theta_{p,m}^* \end{cases} \quad (3)$$

步骤三、数据采集和处理：在线采集输入变量和输出变量数据及输入变量和输出变量对应的质量码，并对质量码为坏值的情况进行处理。具体包括如下：

在线采集传感器的过程变量和质量码，具体是，采集输入变量和输出变量数据及输入变量和输出变量对应的质量码，其中，质量码用于表示采集到的输入变量和输出变量是否有效，记输入变量为 $u_m(k)$ ，输出变量为 $y_p(k)$ ， k 为采样时刻；

若 $u_m(k)$ 的质量码为坏，则清空采集的所有变量历史数据并重新采集；

若 $y_p(k)$ 的质量码为坏，则继续采集数据，直到 y_p 出现好值，若期间连续坏值的个数大于等于预设的 N_{bad} ，则清空采集的所有变量历史数据并重新采集；否则，利用坏值数据段两端的好值数据点对坏值数据段进行线性插值以在线补全历史数据，继续采集数据。

步骤四、模型参数学习：基于贝叶斯优化理论建立传递函数模型自主学习的数学优化模型，利用差分进化算法优化传递函数模型参数。具体包括如下：

当输入变量或输出变量的数据量（输出变量和输入变量的测量次数始终是一样）累积到预设个数 N_s 时，针对每个多输入单输出子系统，基于贝叶斯优化理论建立传递函数模型自主学习的数学优化模型。

根据传递函数模型计算系统脉冲响应序列，利用脉冲响应序列和系统输入变量的卷积计算系统输出变量的预测序列。

重复执行 L 次基于差分进化算法的传递函数模型参数学习， $10 \leq L \leq 30$ 。记当前情况为第 p 个多输入单输出子系统的第 r 次模型参数学习， N_{popu} 为用户设置的种群数量；随机选取长度为 N_{DS} 的连续数据片段，视待学习参数类型随机初始化种群参数，根据学习参数计算系数 $a_{p,m}$ 和 $b_{p,m}$ 并代入式 (1)，采用双线性变换离散化，获取传递函数模型 $G_{p,m}$ 的脉冲响应 $\{g_{p,m,k}, k=1,2,\dots,N_{mdl}+d\}$ ， N_{mdl} 为脉冲响应序列阶段长度， d 为延迟点个数，并计算传递函数模型 $G_{p,m}$ 输出的预测值 $\Delta\hat{y}_{p,m}(k)$ ；如果 $\text{Int}_{p,m}=1$ ，则额外利用一阶积分环节的标准离散化计算方法更新 $\Delta\hat{y}_{p,m}(k)$ 序列；按照式 (4) 计算多输入单输出子系统的输出预测值 $\Delta\hat{y}_p(k)$ ，视待学习参数类型按式 (5) 计算目标函数值并记录在 $\{J_{i_{popu}} | i_{popu}=1,\dots,N_{popu}\}$ ，其中， $\sigma^2 > 0$ 为用户设置的噪声水平， $(\theta_1, \theta_2) \in \{(T_1, T_2), (\omega_n, \xi)\}$ 。

$$\Delta\hat{y}_p(k) = \sum_{m=1}^M \Delta\hat{y}_{p,m}(k) \quad (4)$$

$$J = \sum_{k=N_{mdl}+d}^{N_{DS}} \frac{(\Delta y_p(k) - \Delta\hat{y}_p(k))^2}{\sigma^2} - \sum_{m=1}^M \ln(p(K_{p,m})p(\theta_{1,p,m})p(\theta_{2,p,m})) \quad (5)$$

接下来，反复执行多轮标准差分进化算法中的三个步骤：种群进化、个体变异和个体选择，以极小化目标函数的原则记录并更新 $\{J_{i_{popu}} | i_{popu}=1,\dots,N_{popu}\}$ 和对应的模型参数；最终将最优个体对应的参数学习结果记录到 $\{K_{p,m,r}^{best}, T_{1,p,m,r}^{best}, T_{2,p,m,r}^{best} | m=1,\dots,M\}$ 或 $\{K_{p,m,r}^{best}, \omega_{n,p,m,r}^{best}, \xi_{p,m,r}^{best} | m=1,\dots,M\}$ 中，其中， $K_{p,m,r}^{best}, T_{1,p,m,r}^{best}, T_{2,p,m,r}^{best}, \omega_{n,p,m,r}^{best}, \xi_{p,m,r}^{best}$ 分别为第 r 次模型参数学习的最优个体的模型增益、两个一次项系数、自然角频率和阻尼比。

步骤五、模型参数检验：对每个传递函数模型的每个参数执行参数检验，包括正态性检验、参数死区判断和均值检验。具体包括如下：

对 $\{G_{p,m} | p=1,2,\dots,P, m=1,2,\dots,M\}$ 的每个参数的 L 次学习结果进行检验。

环节一、正态性检验：正态性检验方法采用 GB/T 4882-2001 中的夏皮洛-威尔克 (Shapiro-Wilk) 检验方法，根据被学习的参数类型按照式 (6) 计算特征 x_r 的样本数据，若正态性检验不通过，则该参数保持 $\theta_{p,m}^*$ 不变；否则，进入环节二。

$$x_r = \ln(|\theta_{p,m,r}^{best}|) - \ln(|\theta_{p,m}^*|) \quad (6)$$

环节二、参数死区判断：计算 $\{x_r | r=1, \dots, L\}$ 的平均值 $\bar{x}$ ，如果 $|\bar{x}| \leq \ln(1+\eta_1)$ ，则该参数保持 $\theta_{p,m}^*$ 不变，其中， $\eta_1 \geq 0$ 是用户设定的参数死区范围；否则，进入环节三。

环节三、均值检验：由用户设定均值检验系数 $\eta_2 \in [0,1]$ ，如果 $\bar{x} > 0$ ，则原假设为总体均值 $\mu \leq \eta_2 \bar{x}$ ；反之，如果 $\bar{x} < 0$ ，则原假设为总体均值 $\mu \geq \eta_2 \bar{x}$ ；如果均值检验不通过，则该参数保持 $\theta_{p,m}^*$ 不变；否则，将 $x_r = \eta_2 \bar{x}$ 代入式 (6) 反算出参数的学习结果。

步骤六、模型预测误差对比：用原传递函数模型计算预测均方根误差，用学习后的传递函数模型计算预测均方根误差，根据两者的大小关系决定是否更新多输入单输出系统的传递函数模型，然后返回步骤一继续执行。具体包括如下：

对每个多输入单输出子系统，用原传递函数模型参数为准，计算第 p 个多输入单输出系统的预测均方根误差并记为 $RMSE_p^1$ ；用学习后的传递函数模型参数为准，计算第 p 个多输入单输出系统的预测均方根误差并记为 $RMSE_p^2$ ；

如果 $RMSE_p^2 < RMSE_p^1$ ，则更新该多输入单输出系统的传递函数模型；反之，则该多输入单输出系统的传递函数模型保持不变；

然后返回步骤一继续运行。

本发明的有益效果是：

1、本发明是一种支持在线运行的模型参数的自主学习技术，可利用在线数据主动地学习模型参数。

2、本发明中，传递函数模板采用了二阶传递函数的一般形式，不仅可以学习稳定的二阶实特征根系统模型，而且可以学习稳定的二阶共轭复特征根系统模型，使学习参数具有明确的物理意义。

3、本发明中，通过两种方法增加了传递函数模型在线学习结果的可靠性：第一种，将过程的先验知识融合进了传递函数模型的学习过程，先验知识包括正在投运的控制器传递函数模型的结构、参数值、增益的符号和由系数自动计算出的、具有物理意义的传递函数模型待学习参数类型；第二种，包含参数检验和传递函数模型预测误差结果对比两个步骤，只有参数通过了假设检验且传递函数模型通过了预测误差结果对比，传递函数模型参数才会被自动更新为学习后的结果。

4、本发明中，采用参数死区的方法增加了传递函数模型参数的平稳性，当传递函数模型参数的变化位于用户设定的死区内，则传递函数模型参数不发生改变，以此增加传递函数模型参数的平稳性，有利于控制效果的稳定性。

5、本发明中，采用了对数正态分布的传递函数模型参数分布假设，一方面，对数正态分布的随机变量取值范围是正数域，有效的 $|K|$ 、 T_1 、 T_2 、 ω_n 、 ζ 的值域也是正数域，即二者吻合；另一方面，对于传递函数模型的增益而言，增益偏大会使控制偏慢，但增益偏小可能会使系统不稳定，故减小增益应更谨慎；而对数正态分布的概率密度函数曲线在最大值左侧形状较为陡峭，概率密度随着增益减小衰减较快；故，这种先验模型适用于对增益偏小带来的风险进行建模。

附图说明

图 1 示出了本发明所述的适用于流程工业预测控制的模型自主学习方法的流程框图。

图 2 示出了模型参数检验步骤中的正态性检验、参数死区判断和均值检验三环节流程图。

图 3 示出了不同 η_2 时的传递函数模型学习结果与原传递函数模型的阶跃响应曲线。

图 4 示出了无参数死区方法时的传递函数模型部分学习结果与原传递函数模型的阶跃响应曲线。

具体实施方式

为了便于本领域人员更好的理解本发明，下面结合附图和具体实施例对本发明做进一步详细说明，下述仅是示例性的不限定本发明的保护范围。

本发明所述的一种适用于流程工业预测控制的模型自主学习方法，可适用于多输入多输出系统，具体是基于多输入单输出子系统划分的方式进行的，故具体实施方式中，采用多输入单输出子系统进行详细说明。

下面以两输入单输出子系统为例进行说明，但并不局限于本实施所述的两输入单输出子系统，其他多输入单输出子系统均适用。

假设两输入单输出系统的当前传递函数模型通过下式 (7) 表示。

$$\begin{aligned} Y(s) &= G_{1,1}(s)U_1(s) + G_{1,2}(s)U_2(s) \\ &= \frac{2}{(5s+1)s}U_1(s) + \frac{3.5}{4s^2+0.8s+1}U_2(s) \end{aligned} \quad (7)$$

步骤一、模型设置及学习参数自选择：设置所有传递函数模型，对每个传递函数模型进行计算并自动选择待学习参数。

根据式 (1) 和式 (7)，本实施例中两个传递函数模型设置参数如下：

$$K_{1,1}^* = 2, a_{1,1}^* = 0, b_{1,1}^* = 5, \text{Int}_{1,1} = 1; K_{1,2}^* = 3.5, a_{1,2}^* = 4, b_{1,2}^* = 0.8, \text{Int}_{1,2} = 0$$

对于传递函数模型 $G_{1,1}$ ，其特征多项式（不含纯积分）可以在实数域内被分解为一阶多项式乘积的形式，故按照式 (2) 计算， $G_{1,1}$ 的待学习参数为 $K_{1,1}$ 和 $T_{1,1}$ ，且 $K_{1,1}^* = 2$ 和 $T_{1,1}^* = 5$ 。

对于传递函数模型 $G_{1,2}$ ，其特征多项式（不含纯积分）包含两个共轭复数根，故 $G_{1,2}$ 的待学习参数为 $K_{1,2}$ 、 $\omega_{n,2}$ 和 $\zeta_{1,2}$ ，且将特征多项式转化为零极点形式后，可知 $K_{1,2}^* = 3.5$ 、 $\omega_{n,2}^* = 0.5$ 和 $\zeta_{1,2}^* = 0.2$ 。

步骤二、参数分布建模：对于参与学习的每个传递函数模型，采用对数正态分布对待学习参数建立对数正态分布模型。

由于两个传递函数模型的增益均不为零，且特征根实部非正且无纯虚数根，故两个传递函数模型均应被学习；

假设设定 $\beta_\theta = 1.1$ ， $\theta \in \{K, T_1, T_2, \omega_n, \zeta\}$ ，则根据式(3)可以解算出 $u_{K_{1,1}} \approx 0.756687$ ，
 $u_{T_{1,1}} \approx 1.672978$ ， $u_{K_{1,2}} \approx 1.316303$ ， $u_{\omega_{n,2}} \approx -0.629607$ ， $u_{\zeta_{1,2}} \approx -1.545898$ ，
 $\sigma_K^2 = \sigma_{T_1}^2 = \sigma_{\omega_n}^2 = \sigma_\zeta^2 \approx 0.06354$ ，其中， $(u_{K_{1,1}}, \sigma_K^2)$ 、 $(u_{T_{1,1}}, \sigma_{T_1}^2)$ 、 $(u_{K_{1,2}}, \sigma_K^2)$ 、 $(u_{\omega_{n,2}}, \sigma_{\omega_n}^2)$ 和
 $(u_{\zeta_{1,2}}, \sigma_\zeta^2)$ 分别为参数 $K_{1,1}$ 、 $T_{1,1}$ 、 $K_{1,2}$ 、 $\omega_{n,2}$ 和 $\zeta_{1,2}$ 的对数正态分布的分布参数。

步骤三、数据采集和处理：在线采集输入变量和输出变量数据及输入变量和输出变量对应的质量码，并对质量码为坏值的情况进行处理。

在线输入变量和输出变量数据及输入变量和输出变量对应的质量码，并判断数据质量码好坏，其中，记输入变量为 $u_1(k)$ 和 $u_2(k)$ ，输出变量为 $y_1(k)$ ， k 为采样时刻；

若 $u_1(k)$ 或 $u_2(k)$ 的质量码为坏，则清空采集的所有变量历史数据并重新采集；

若 $y_1(k)$ 的质量码为坏，则继续采集数据，直到 y_1 出现好值，若期间连续坏值的个数大于等于预设的 N_{bad} ，则清空采集的所有变量历史数据并重新采集，若期间连续坏值的个数小于 N_{bad} ，则利用坏值数据段两端的好值数据点对坏值数据段进行线性插值后在线补全历史数据，继续采集数据。比如 $N_{bad} = 3$ ， $y_1(5) = 1, y_1(6) = 1, y_1(7) = 1, y_1(8) = 4$ ，第6个和第7个样本点的质量码为坏，第5个和第8个样本点的质量码为好，则在 $k = 8$ 时，利用 $k = 5$ 和 $k = 8$ 的数据点对 $k = 6$ 和 $k = 7$ 进行线性插值，即 $y_1(6) = 2, y_1(7) = 3$ 。

步骤四、模型参数学习：基于贝叶斯优化理论建立传递函数模型自主学习的数学优化模型，利用差分进化算法优化传递函数模型参数。

当输入变量或输出变量的数据量累积到预设个数 $N_s = 3000$ 个时，对该两输入单输出系统重复执行 $L = 20$ 次基于差分进化算法的模型参数学习。

记当前情况为该两输入单输出系统的第 r 次模型参数学习：

随机选取长度为 $N_{DS} = 1000$ 的连续数据片段；设置种群数目 $N_{popu} = 10$ ，随机

初始化种群参数 $\{K_{1,1,i_{popu}}, T_{1,1,i_{popu}} | i_{popu} = 1, \dots, 10\}$ 和 $\{K_{1,2,i_{popu}}, \omega_{n1,2,i_{popu}}, \zeta_{1,2,i_{popu}} | i_{popu} = 1, \dots, 10\}$, 根据学习参数反算系数 $a_{1,1}, b_{1,1}, a_{1,2}$ 和 $b_{1,2}$ 并代入传递函数模型的模板, 采用双线性变换离散化, 获取传 $G_{1,1}$ 和 $G_{1,2}$ 的脉冲响应, 分别记为 $\{g_{1,1,k}, k=1, 2, \dots, N_{mdl}+1\}$ 和 $\{g_{1,2,k}, k=1, 2, \dots, N_{mdl}+1\}$, 递函数模型按照式 (8) 计算传递函数模型 $G_{1,1}$ 和 $G_{1,2}$ 输出的预测值 $\Delta \hat{y}_{1,1}(k)$ 和 $\Delta \hat{y}_{1,2}(k)$, 其中, Δ 为差分算子。

$$\Delta \hat{y}_{p,m}(k) = \sum_{i=1}^{N_{mdl}+d} g_{p,m,i} \Delta u_m(k-i), \Delta u_m(0) = 0, k = N_{mdl}+1, \dots, 1000 \quad (8)$$

因为 $\text{Int}_{1,1}=1$, 故利用一阶纯积分环节的标准离散化计算方法更新 $\Delta \hat{y}_{1,1}(k)$, 按照式 (9) 计算多输入单输出系统的输出预测值 $\Delta \hat{y}_1(k)$, 按照式 (10) 计算目标函数值并记录在 $\{J_{i_{popu}} | i_{popu} = 1, \dots, 10\}$, 其中, $\sigma^2 > 0$ 为用户设置的噪声水平, $(\theta_1, \theta_2) \in \{(T_1, T_2), (\omega_n, \xi)\}$ 。

$$\Delta \hat{y}_p(k) = \sum_{m=1}^2 \Delta \hat{y}_{p,m}(k) \quad (9)$$

$$J = \sum_{k=N_{mdl}+d}^{N_{ps}} \frac{(\Delta y_p(k) - \Delta \hat{y}_p(k))^2}{\sigma^2} - \sum_{m=1}^2 \ln(p(|K_{p,m}|)p(\theta_{1,p,m})p(\theta_{2,p,m})) \quad (10)$$

接下来, 反复执行多轮标准差分进化算法中的三个步骤: 种群进化、个体变异和个体选择, 以极小化目标函数的原则记录并更新 $\{J_{i_{popu}} | i_{popu} = 1, \dots, 10\}$ 和对应的模型参数, 最终将最优个体对应的参数学习结果记录到 $\{K_{1,1,r}^{best}, T_{1,1,r}^{best}\}$ 和 $\{K_{1,2,r}^{best}, \omega_{n1,2,r}^{best}, \zeta_{1,2,r}^{best}\}$ 中。

步骤五、模型参数检验: 对每个传递函数模型的每个参数执行参数检验, 包括正态性检验、参数死区判断和均值检验。

如图 2 所示, 对递函数模型 $G_{1,1}$ 和 $G_{1,2}$ 的每个参数的 20 次学习结果进行检验。

环节一、正态性检验: 根据被学习的参数类型按照式 (6) 计算特征 x_r 的样本数据。比如, 参数 $T_{1,p,m}$ 检验结果不通过, 则参数 $T_{1,p,m}$ 保持 5 不变; 其它参数通过, 进入环节二。

环节二、参数死区判断：计算 $\{x_r | r=1, \dots, 20\}$ 的平均值 $\bar{x}$ ，如果 $|\bar{x}| \leq \ln(1+\eta_1)$ ，则该参数保持 $\theta_{p,m}^*$ 不变；比如 $\eta_1 < 0.2$ ，如果在执行 $K_{1,2}$ 的参数死区判断时， $|\bar{x}| = 0.15 \leq \ln(1+0.2)$ ，则 $K_{1,2}$ 保持 3.5 不变；其它参数通过，进入环节三。

环节三、均值检验：如果 $\bar{x} > 0$ ，则原假设为总体均值 $\mu \leq \eta_2 \bar{x}$ ；反之，如果 $\bar{x} < 0$ ，则原假设为总体均值 $\mu \geq \eta_2 \bar{x}$ ；如果检验结果不通过，则该参数保持 $\theta_{p,m}^*$ 不变；比如在执行参数 $\omega_{m,2}$ 的均值检验时， $\eta_2 = 0.8$ ， $\bar{x} = 0.2$ ，如果检验结果为无法拒绝 $\mu \leq 0.16$ 的假设，则 $\omega_{m,2}$ 维持 0.5 不变；否则，拒绝原假设，将 $x_r = \eta_2 \bar{x}$ 代入式 (6) 反算出参数的学习结果。

步骤六、模型预测误差对比：用原传递函数模型计算预测均方根误差，用学习后的传递函数模型计算预测均方根误差，根据两者的大小关系决定是否更新多输入单输出系统的传递函数模型，然后返回步骤一继续执行。

具体是，用原传递函数模型参数为准，计算该两输入单输出系统的输出预测值并按照式 (11) 计算预测均方根误差，记为 $RMSE_1^1$ ；以学习后的传递函数模型为准，再次计算该两输入单输出系统的输出预测值并按照式 (11) 计算预测均方根误差，记为 $RMSE_1^2$ 。

$$RMSE_p = \sqrt{\frac{1}{N_S - (N_{mdl} + d) + 1} \sum_{k=N_{mdl}+d}^{N_S} (\Delta y_p(k) - \Delta \hat{y}_p(k))^2} \quad (11)$$

如果 $RMSE_1^2 < RMSE_1^1$ ，则更新该两输入单输出系统的模型参数 $\{K_{1,m}^*, a_{1,m}^*, b_{1,m}^* | m=1, 2\}$ ；反之，该两输入单输出系统的所有模型保持不变。

然后返回步骤一继续重复运行。

现结合另一具体实施例继续说明，该具体实施例为式 (12) 所示的多输入单输出系统，输入变量为两个支路的水温设定值，输出变量为干路水温，控制方案是用两路支路控制干路的水温。作为多输入单输出系统，其具体实施方式与上文所述相同，故不赘述实施过程，而是重点说明本发明的效果。

$$Y(s) = G_{1,1}(s)U_1(s) + G_{1,2}(s)U_2(s) \quad (12)$$

假设正在投运的控制器传递函数模型为式 (13)：

$$\begin{aligned} Y(s) &= G_{1,1}^*(s)U_1(s) + G_{1,2}^*(s)U_2(s) \\ &= \frac{0.395e^{-75s}}{300s+1}U_1(s) + \frac{0.312e^{-45s}}{480s+1}U_2(s) \end{aligned} \quad (13)$$

分别设置 $\eta_2 = 0.3$ 、 $\eta_2 = 0.6$ 和 $\eta_2 = 0.8$ ，进行模型参数自学习，将 $G_{1,1}(s)$ 的学习结果的阶跃响应曲线和 $G_{1,1}^*(s)$ 的阶跃响应曲线绘制在图 3 中。当 η_2 取值 0.3 时， $G_{1,1}(s)$ 的学习结果的阶跃响应曲线更接近 $G_{1,1}^*(s)$ 的阶跃响应曲线，模型参数的变化程度小；当 η_2 取值 0.8 时， $G_{1,1}(s)$ 的学习结果的阶跃响应曲线更远离 $G_{1,1}^*(s)$ 的阶跃响应曲线，模型参数的变化程度大。归纳上述现象，即 η_2 会影响 $G_{1,1}(s)$ 传递函数模型的学习结果，在模型自主学习的迭代过程中，用户设定不同的 η_2 ，可以控制模型自主学习的参数变化速度，在不确定工况中增加模型的可靠性，防止模型的大幅度变化对过程平稳性的影响。

图 4 画出了不含参数死区方法时， $G_{1,2}(s)$ 的学习结果的阶跃响应曲线与 $G_{1,2}^*(s)$ 的阶跃响应曲线。这些学习结果的传递函数模型参数与 $G_{1,2}^*(s)$ 的传递函数模型参数相对差别小于 10%。在这种情况下，由于预测控制器具有反馈校正环节，预测控制器便具有一定范围内的鲁棒性，故为了减小预测控制器的传递函数模型的小幅波动，此时可以设置参数死区范围 $\eta_1 = 0.1$ ，则模型学习结果将保持 $G_{1,2}^*(s)$ 不发生变化。归纳上述现象，用户设定不同的 η_1 ，可以借助控制器的鲁棒性，增加传递函数模型参数的平稳性，有利于控制效果的稳定性。

以上仅描述了本发明的基本原理和优选实施方式，本领域人员可以根据上述描述做出许多变化和改进，这些变化和改进应该属于本发明的保护范围。

权 利 要 求 书

1.一种适用于流程工业预测控制的模型自主学习方法，其特征在于，通过多输入单输出系统作为基本子系统予以实现，包括步骤：

步骤一、设置所有传递函数模型，对每个传递函数模型进行计算并自动选择待学习参数；

步骤二、对于参与学习的每个传递函数模型，采用对数正态分布对待学习参数建立对数正态分布模型；

步骤三、在线采集输入变量和输出变量数据及输入变量和输出变量对应的质量码，并对质量码为坏值的情况进行处理；

步骤四、基于贝叶斯优化理论建立传递函数模型自主学习的数学优化模型，利用差分进化算法优化传递函数模型参数；

步骤五、对每个传递函数模型的每个参数执行参数检验，包括正态性检验、参数死区判断和均值检验；

步骤六、对每个多输入单输出子系统，用原传递函数模型计算预测均方根误差，用学习后的传递函数模型计算预测均方根误差，根据两者的大小关系决定是否更新多输入单输出系统的传递函数模型，然后返回步骤一继续执行。

2.根据权利要求1所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤一中，设置传递函数模型具体包括：

在固定的二阶系统加一阶积分加纯滞后传递函数模型模板下，设置每个输入变量到每个输出变量之间的传递函数模型。

3.根据权利要求2所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤一中，传递函数模型的数量根据输入变量和输出变量的数量而定，设输入变量为 M 个、输出变量为 P 个，则需设置 $M \times P$ 个传递函数模型。

4.根据权利要求1所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤一中，对每个传递函数模型进行计算并自动选择待学习参数具体包括：

先对每个传递函数模型计算，然后进行如下分析：

如果某个传递函数模型的增益为0，则该传递函数模型不参与模型在线学

习；

否则，判断二阶系统的极点位置：（1）如果存在实部为正的根或者纯虚数根，则该传递函数模型不参与模型在线学习；（2）如果存在两个实部为负数的共轭复数根，则计算传递函数模型的自然角频率和阻尼系数，并将模型增益、自然角频率和阻尼系数作为待学习参数；除上述（1）和（2）两种情况外，则将传递函数模型的分母多项式在实数域中进行因式分解，取分解后的所有的非零的一次项系数和模型增益作为待学习参数。

5.根据权利要求1所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤二具体包括：

对于参与学习的每个传递函数模型，根据用户设定的与分布形状相关的系数，计算对数正态分布参数，利用对数正态分布对待学习参数建模。

6.根据权利要求1所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤三具体包括：

在线采集输入变量和输出变量数据及输入变量和输出变量对应的质量码，并进行如下判断和处理：

对于输入变量，只要出现坏质量码数据，则清空采集的所有变量历史数据并重新采集；

对于输出变量，当连续出现预设个数 N_{bad} 或 N_{bad} 以上个坏质量码数据时，则清空采集的所有变量历史数据并重新采集，当连续出现的坏质量码数据少于 N_{bad} 时，则利用坏值数据段两端的好值数据点采用线性插值方法在线补全历史数据。

7.根据权利要求1所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤四具体包括：

当输入变量或输出变量的数据量累积到预设个数时，针对每个多输入单输出子系统，基于贝叶斯优化理论建立传递函数模型自主学习的数学优化模型；

根据传递函数模型计算系统脉冲响应序列，利用脉冲响应序列和系统输入变量的卷积计算系统输出变量的预测序列；

利用差分进化算法优化传递函数模型参数。

8.根据权利要求 7 所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，差分进化算法包括三个步骤：种群进化、个体变异和个体选择。

9.根据权利要求 1 所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤五具体包括：

对每个传递函数模型的每个参数按顺序依次执行正态性检验、参数死区判断和均值检验三个环节，当三个环节全部通过时，参数学习结果有效，否则，参数学习结果无效，保持原参数不变。

10.根据权利要求 1 所述的适用于流程工业预测控制的模型自主学习方法，其特征在于，步骤六具体包括：

对每个多输入单输出子系统，以原传递函数模型参数为准，计算多输入单输出系统的预测均方根误差并记为 $RMSE^1$ ，以学习后的传递函数模型参数为准，计算多输入单输出系统的预测均方根误差并记为 $RMSE^2$ ；

如果 $RMSE^2 < RMSE^1$ ，则更新该多输入单输出子系统的传递函数模型，反之，则该多输入单输出子系统的传递函数模型保持不变；

然后，返回步骤一继续执行。