International Journal of Automation, Control and Intelligent Systems
Vol. 1, No. 4, 2015, pp. 97-101
http://www.aiscience.org/journal/ijacis
ISSN: 2381-7526 (Print); ISSN: 2381-7534 (Online)

Data-Driven Constrained PID Controller Design for CSTR

Xu Dezhi*, Ji Nan

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi Jiangsu, China

Abstract

Since most chemical processes exhibit severe nonlinear and time-varying behavior, the control of such processes is challenging. In this paper, we propose data-driven constrained controller design method based on lazy learning for chemical processes. Using a lazy learning algorithm, a local valid linear model denoting the current state of system is automatically exacted for adjusting the PID controller parameters based on input/output data. This scheme can adjust the constrained PID parameters in an online manner even if the system has nonlinear properties. At same time, in order to solve the control input saturation problem, a dynamic anti-windup compensator is proposed for accommodating the reference. The simulation results on the dynamic model of Continuous Stirred Tank Reactor (CSTR) are provided to demonstrate the effectiveness of the proposed new control techniques.

Keywords

Data-Driven, Lazy Learning, PID, Dynamic Anti-Windup, Continuous Stirred Tank Reactor (CSTR)

Received: October 11, 2015 / Accepted: October 28, 2015 / Published online: December 6, 2015

1. Introduction

With the development of control theory and technology, many scholars put forward the adaptive online adjustment strategy, such as: based on the optimization strategy parameters self-tuning [1-3], based on the generalized minimum variance parameter adjustment [4-6], etc. When using these advanced strategy, it requires to identify the current system conditions of the model, which has two kinds of online and offline. In actual application of these strategies are achieved good results. And in the face of a wide range, strong nonlinear, or jumping change system, if we adopt the rolling strategy, it tend to throw away the past data, relying only on the current some of the data obtained, can't let the system to establish a more accurate model, it is difficult to achieve good control effect. If the offline global model is set up, because the new point, often need to be trained, the offline model is set up again, big workload [7-9].

In order to overcome the above problem, we use the real-time algorithm adaptive, the nature of the K-VNN search strategy, accumulated data to find out from the system and the current point matching the data set, a local polynomial fitting method is adopted to establish the local model of the system. And the system is based on tracking error of the least performance index which is deduced with the PID control law in the form of structure, it is easy to realize the adaptive PID control for nonlinear systems. At same time, in order to satisfy the requirements for control constraints, a dynamic constraint unit and an anti-windup scheme are adopted.

2. Data-Driven Modeling Method (Lazy Learning)

Lazy learning algorithm based on "input to produce output of similitude" principle. It is generally the sample data memory in memory, and then according to the input point, find similar

* Corresponding author
E-mail address: xudezhi@jiangnan.edu.cn (Xu Dezhi)

data in the sample data, according to the sample data to get the input to the corresponding output [10-12]. Therefore, it is also called "Based on the study of Memory" (the MBL: Memory-based Learning). Describe the input data and sample data correlation criterion generally USES the distance function, which recently with the input point data has the higher correlation [12-13].

Considering an unknown nonlinear mapping $f: R^n \to R$, we assumed that can get system observable input and output data $\{(X_i, y_i)\}_{i=1}^{N}$. And this set data has function relation:

$$y_i = f(x_i) + \varepsilon$$

Where: the independent variable $X \in R^n$, the dependent variable $y_i \in R$, zero mean $\varepsilon_i \in R$ and variance σ^2 for the independent random distribution variables. The question is whether any vector X_q of the input space according to the existing data sets can create a mapping. It can pass the mapping and get the system corresponding estimate output $\hat{y}_q$. This problem can be summed up in solving the following optimization problem.

$$\min \sum_{(X_i, Y_i) \in \Omega_k} \left(y_i - f(X_i, \theta)\right)^2 w_i$$

Where: Ω_k is the local space X_k of the nearest k samples. $f(\cdot)$ describes the nonlinear mapping function of the input and output vector. w_i is weight, it is the local space within the sample data to the influence degree of the output vector. Effects on the system output of different sample data in local space are different. Intuitively, the closest to the input vector corresponding to the sample vector output can most reflect the current output. It is also the basic principle of lazy learning algorithm: analog input produces analog output.

By using the principle of the above algorithm, we assume that the controlled system available can use NARX model to represent as follows:

$$y(t) = f(\phi(t-1)) + \varepsilon(t) \tag{1}$$

Where:

$\phi(t-1) = \left[y(t-1), \cdots, y(t-n_y), u(t-p), \cdots, u(t-p-n_u)\right]^T$ It is system state vector on t moment. $m(m = n_y + n_u); n_y, n_u, p > 0$ is the system output, the system input order and the system delay. $y(t)$, $u(t)$, $\varepsilon(t)$ are the system output, the system input, the zero mean white noise. $f(\cdot)$ is an unknown nonlinear function.

The control input magnitude and rate constraints is redefined as

$$u_{\min} \le u_f \le u_{\max} \quad \dot{u}_{\min} \le \dot{u} \le \dot{u}_{\max} \tag{2}$$

About the system (1) the physical description of the system is unknown, assuming that there are N sets of input and output data $\{y(i), \phi(i)\}_{i=1}^{N}$, at the current time t, system input information $\phi(t)$, the K-VNN search strategy is adopted, in the system of existing N sets of data to find the most similar data (K<< N), specific as follows:

1, When $\cos\beta(\phi(i), \phi(t)) < 0$, we think $\phi(i)$ diverge from the current input $\phi(t)$, which is unfavorable for system modeling, so we should discard the data;

2, Otherwise, the function of the index nuclear of $\phi(i)$ and $\phi(t)$ and the angle cosine weighted sum is defined as

$$D(\phi(i), \phi(t)) = \alpha \cdot e^{-d(\phi(i), \phi(t))} + (1-\alpha) \cdot \cos\beta(\phi(i), \phi(t)) \tag{3}$$

Where:

$$\begin{cases} \cos\beta(\phi(i), \phi(t)) = \dfrac{\phi^T(i)\phi(t)}{\|\phi(i)\|_2 \bullet \|\phi(t)\|_2} \\ d(\phi(i), \phi(t)) = \|\phi(i) - \phi(t)\|_2, \alpha \in [0,1] \end{cases}$$

By (3), the weighted selection criteria $D(\phi(i), \phi(t))$ directly reflects $\phi(i)$ and $\phi(t)$ similarity. If the two vectors are the closer, d is the more small, and $\cos\beta$ is the greater, thereby $D(\phi(i), \phi(t))$ is bigger. So in the existing data information, (3) selects the largest k set data values $D(\cdot)$, in descending order, constructed learning sets:

$$\begin{aligned} &\{(\phi(1), y(1)), \cdots, (\phi(k), y(k))\} \\ &D(\phi(1), \phi(t)) > \cdots > D(\phi(k), \phi(t)). \end{aligned} \tag{4}$$

The information vector is linear regression, the system can get the local linear model of the current conditions. But at different working conditions, in accordance with the current point data $\phi(t)$ may not be the same density. The data number used in modeling is also uncertain, namely: modeling neighborhood values are variable, in order to get the best linear model $\hat{\theta}$, reduce the amount of calculation, set up neighborhood range $k \in [k_m, k_M] (k_m < k_M)$ in advance. When calculating the local model $\hat{\theta}_{k+1}$ of the neighboring $k+1$, we directly use model values θ_k of the neighbor k and the recursive least squares method. So can get the system model.

$$\begin{cases} \psi_{k+1} = \Phi_{k+1}^T P_k \Phi_{k+1} + 1/D_{k+1}, \\ \gamma_{k+1} = P_k \Phi_{k+1} \psi_{k+1}^{-1}, \\ e_{k+1} = y(k+1) - \Phi_{k+1}^T \hat{\theta}_k, \\ \hat{\theta}_{k+1} = \hat{\theta}_k + \gamma_{k+1} e_{k+1}, \\ P_{k+1} = P_k - \gamma_{k+1} \gamma_{k+1}^T \psi_{k+1}. \end{cases} \quad (5)$$

We get the model θ_{k+1} of the near $\mathrm{k}+1$. At the same time,we can get a cross error value of the near $\mathrm{k}+1$:

$$e_{k+1,j}^{loo} = y(j) - \Phi_j \hat{\theta}_{k+1}^{-j} = \frac{y(j) - \Phi_j^T \hat{\theta}_{k+1}}{1 - \Phi_j^T P_{k+1} \Phi_j}, \; j = 1, 2 \cdots, k+1. \quad (6)$$

Where: in $\mathrm{k}+1$ sets data, it represents local model $\hat{\theta}_{k+1}^{-j}$ by removing the $j-th$ data. $e_{k+1,j}^{loo}$ is the error value between the actual value $y(j)$ and the model $\hat{\theta}_{k+1}^{-j}$. In this way, we can get cross error $\{e_{k+1,j}^{loo}\}_{j=1}^{k+1}, k+1 \le k_M$, mean square, and $e_{k+1,j}^{loo}$ of the neighbor $\mathrm{k}+1$

$$E^{loo}(k+1) = \frac{\sum_{j=1}^{k+1} w_j \left(e_{k+1,j}^{loo}\right)^2}{\sum_{j=1}^{k+1} w_j}. \quad (7)$$

Where: the weighted factor $w = \sqrt{D(\phi(j), \phi(t))}$ directly reflects the size of the cross error $E^{loo}(k+1)$ with each $\phi(j)$. The $\phi(j)$ is closer $\phi(t)$, its contribution is the greater, whereas the smaller.

$$E^{loo}(k+1) > E^{loo}(k), k+1 \in [k_m, k_M]. \quad (8)$$

From it, we think model variation and stop the regression calculation, then we can use the model θ_k as the system optimal model of the current time. Otherwise, by using the recursive least squares algorithm model, we select new information vector from learning sets, and continue to iteration, until $k = k_M$. So, we can judge the merits of the local model in time, get in line with the current moment system input and output relationship of the best local linear model.

3. PID Parameter Tuning Principles

In this paper, we propose adaptive constrained PID control based on lazy learning identification. Its structure is shown in Fig. 1.

Define the new tracking error $e(k) = y_r(k) - y_m(k) - \varsigma$. Since the dynamic constraints in the close-loop SOFC control, an anti-windup compensator is designed to accommodate the reference trajectory $y_r(k)$. The compensation signal ς is designed as following

$$\varsigma(k) = \mu\varsigma(k-1) + \frac{\partial y_m(k)}{\partial u(k)}(u_c(k) - u(k)) \quad (9)$$

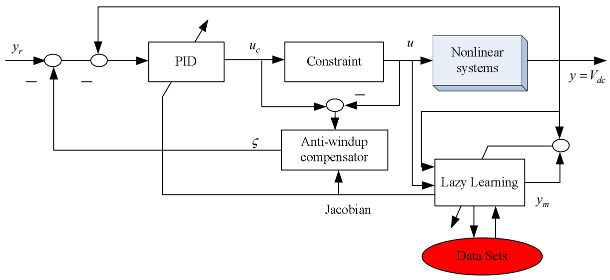

Fig. 1. Structure of adaptive constrained PID control based on lazy learning.

where μ is chosen in the unit circle. The PID parameters are modified on-line using of results of lazy learning identification.

PID input is:

$$\begin{cases} x_{c1} = e(k) - e(k-1) \\ x_{c2} = e(k) \\ x_{c3} = e(k) - 2e(k-1) + e(k-2) \end{cases}$$

Control algorithm is:

$$u_c(k) = u(k-1) + K_p x_{c1} + K_i x_{c2} + K_d x_{c3} \quad (10)$$

The input constraints (2), then adaptive constrained controller is described as

$$u(t) = Sat\left\{\left(u(k-1) + Sat\left\{\left(u_c(k) - u(k-1)\right), T\dot{u}_{\min}, T\dot{u}_{\max}\right\}\right) u_{\min}, u_{\max}\right\}$$

where $Sat(\cdot)$ function is defined as

$$Sat(a,b,c) = \begin{cases} b & a \le b \\ a & b < a < c \\ c & a \ge c \end{cases}$$

The indicators of lazy learning algorithm are

$$J(k) = \frac{1}{2} e^2(k) \quad (11)$$

Adjusted K_p, K_i, K_d, we use the gradient descent method

$$\begin{aligned}\Delta K_p &= -\eta_p \frac{\partial E_1}{\partial K_p} = -\eta_p \frac{\partial E_1}{\partial u_c}\frac{\partial u_c}{\partial K_p} = \eta_p e(k)\frac{\partial y_m}{\partial u} xc_1(k)\\ \Delta K_i &= -\eta_i \frac{\partial E_1}{\partial K_i} = -\eta_i \frac{\partial E_1}{\partial u_c}\frac{\partial u_c}{\partial K_i} = \eta_i e(k)\frac{\partial y_m}{\partial u} xc_2(k) \qquad (12)\\ \Delta K_d &= -\eta_d \frac{\partial E_1}{\partial K_d} = -\eta_d \frac{\partial E_1}{\partial u_c}\frac{\partial u_c}{\partial K_d} = \eta_d e(k)\frac{\partial y_m}{\partial u} xc_3(k)\end{aligned}$$

Where: $\frac{\partial y_m}{\partial u}$ is Jacobian information of the controlled object. It is gotten by lazy learning method in Section 2.

4. Simulation Results

CSTR is an important unit of chemical process. It has strong nonlinear characteristics, and it is a typical nonlinear object in chemical system. Dynamic equations for the input and output are

$$\begin{cases}\dfrac{\mathrm{d}C_a}{\mathrm{d}t} = \dfrac{q}{V}\left(C_{af} - C_a\right) - k_0 C_a \exp\left(-E/RT\right),\\ \dfrac{\mathrm{d}T}{\mathrm{d}t} = \dfrac{q}{V}\left(T_f - T\right) + \dfrac{(-\Delta H)k_0 C_a}{\rho C_p}\exp\left(-E/RT\right)\\ \qquad + \dfrac{\rho_c C_{pc}}{Q_c C_p} Q_c\left[1 - \exp\left(\dfrac{-hA}{Q_c \rho_c C_{pc}}\right)\right]\left(T_{cf} - T\right).\end{cases} \qquad (13)$$

Here: the density C_a of the product A in the reactor, the reactants temperature T, the in and out of the material flow q, the coolant flow Q_c, the coolant inlet temperature T_{cf}, the coolant outlet temperature T_c, the feed concentration C_{af}, the feed temperature T_f. By adjusting the size of the coolant flow, we can control the temperature T in the reactor and the reactants concentration C_a. According to process requirements, we determine the process output C_a and T. Where: the control variables Q_c, the external disturbance variables T_c, q, C_{af}. The physical parameters are shown in table 1.

In the table, the static working point of CSTR corresponding steady-state value is C_{c0}, T_0, C_{a0}. In the static working point, when it changes $\pm 20\%$ of the coolant flow $Q_{c0} = 103.41\ L/\min$, it will produce 2000 set samples.

Input vector of the model is:

$$\phi(t-1) = \left[T(t-1), T(t-2), Q_c(t-1), Q_c(t-2)\right]$$

The local model order is $n_y = 2$, $n_u = 1$. Lazy learning parameter is $\alpha = 0.85$, $k \in [12,100]$. Punishment of PID controller is $Q = 9.3$. $T(446.5 \to 444.5 \to 438.5)$ output variable change value is shown in Fig. 2.

Fig. 2 shows that when the system working point move, lazy learning method selects the modeling data due to the nature of time and space adjacent to the change of the fast tracking system, so gain parameters of PID controller has better adaptive. In the whole tracking trajectory, PID control parameters are shown in Fig. 3.

Table 1. Model parameters.

Parameter	Value
q	100 L/min
T_f	350 K
V	100 L
k_0	7.2×10^{10} L/min
$-\Delta H$	2×10^5 cal/mol
C_p, C_{pc}	1 cal/(g·k)
T_0	440.2 K
T_s	0.1 min
C_{af}	1.0 mol/L
T_{cf}	350 K
hA	7×10^5 cal/(min·k^{-1})
E/R	9.95×10^3 K
ρ, ρ_c	1000 g/L
Q_{c0}	103.41 L/min
C_{a0}	0.0836 mol/L

5. Conclusions

We have carried out a systematic study on the data driven PID control of a CSTR in this paper. The lazy learning algorithm, a local valid linear model denoting the current state of system is automatically exacted for adjusting the PID controller parameters based on input/output data. A dynamic constraint unit with anti-windup scheme is adopted to keep saturation range as long as possible. This scheme can adjust the PID parameters in an online manner even if the system has nonlinear properties. Finally, simulation results are provided on CSTR to show the effective and advantages of the control strategy.

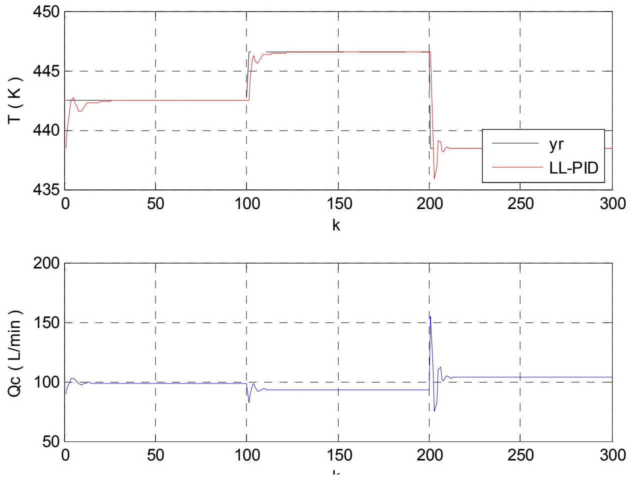

Fig. 2. The CSTR response curve.

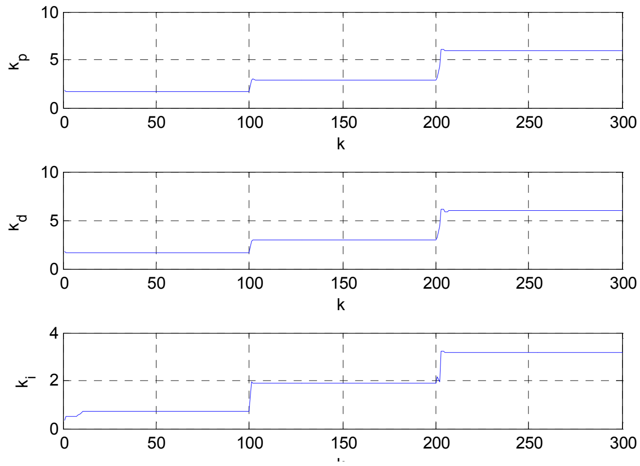

Fig. 3. The change curve of PID controller parameters.

Acknowledgements

This work is supported by National Natural Science Foundation of China (61503156), the Fundamental Research Funds for the Central Universities (JUSRP11562).

References

[1] XU M, LI S Y, CAI W J. Receding horizon optimization approach to PID controller parameters auto tuning. *Acta Automatica Sinica*, 2005, 31(3): 459-463.

[2] CHEN J H, HUANG T C. Applying neural networks to online updated PID controller for nonlinear process control. *Journal of Process Control*, 2004, 14(2): 211-230.

[3] ZHANG Y, CHEN Z Q, YUAN Z Z. Nonlinear system PID-type multi-step predictive control. *Journal of Control Theory & Applications*, 2004, 21(2): 201-204.

[4] XU D, JIANG B and SHI P. A novel model free adaptive control design for multivariable industrial processes. *IEEE Trans. on Industrial Electronics,* 2014, 61(11), 6391-6398.

[5] XU D, JIANG B and SHI P. Adaptive observer based data-driven control for nonlinear discrete-time processes. *IEEE Trans. on Automation Science and Engineering,* 2014, 11(4), 1037-1045.

[6] WANG W. Research on Control Methods of CSTR. *Beijing Jiao tong university*,2013.

[7] WANG T. Research on soft sensing methods for estimating parameters of the concentration in continuous stirred tank reactor. *Northeaster University*, 2008.

[8] BONTEMPI G, BIRATTARI M. From linearization to lazy learning: a survey of divide-and-conquer techniques for nonlinear control. *International Journal of Computational Cognition,* 2005, 3(1): 56-73.

[9] SUN W, WANG W, ZHU R. Iterative learning control for nonlinear system using lazy learning method. *Control and Decision,* 2003, 3(18): 263-266.

[10] BOADA M J L, Calvo J A, BOADA B L, et al. Modeling of a magnetorheological damper by recursive lazy learning. *International Journal of Non-Linear Mechanics*, 2011, 46(3): 479-485.

[11] WEI C. Comparing lazy and eager learning models for water level forecasting in river-reservoir basins of inundation regions. *Environmental Modelling & Software*, 2015, 63: 137-155

[12] ZhANG M L, Zhou Z H. ML-KNN: A lazy learning approach to multi-label learning. *Pattern recognition*, 2007, 40(7): 2038-2048.

[13] VALL J M, GALVAN I M, ISASI P. Learning radial basis neural networks in a lazy way: A comparative study. *Neurocomputing*, 2008, 71(13): 2529-2537.