SCIENCE CHINA
Information Sciences

• LETTER •

August 2020, Vol. 63 189204:1–189204:3
https://doi.org/10.1007/s11432-018-9645-3

Synthesis of model predictive control based on data-driven learning

Yuanqiang ZHOU[1,2,3], Dewei LI[2,3*], Yugeng XI[2,3] & Zhongxue GAN[1]

[1]*ENN Science & Technology Development Co., Ltd, Langfang* 065001*, China;*
[2]*Department of Automation, Shanghai Jiao Tong University, Shanghai* 200240*, China;*
[3]*Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai* 200240*, China*

Received 5 June 2018/Revised 5 September 2018/Accepted 19 October 2018/Published online 16 March 2020

Citation Zhou Y Q, Li D W, Xi Y G, et al. Synthesis of model predictive control based on data-driven learning. Sci China Inf Sci, 2020, 63(8): 189204, https://doi.org/10.1007/s11432-018-9645-3

Dear editor,
Model predictive control (MPC) is a practically effective and attractive approach in the field of industrial processes [1] owing to its excellent ability to handle constraints, nonlinearity, and performance/cost trade-offs. The core of all model-based predictive algorithms is to use "open-loop optimal control" instead of "closed-loop optimal control" within a moving horizon [2]. It is assumed in this study that the readers are familiar with MPC as a control design methodology.

Because the dynamic model of a system predicting its evolution is usually inaccurate, the actual behaviors may deviate significantly from the predicted ones. Thus, acquiring accurate knowledge of the physical model is essential to ensure satisfactory performance of MPC controllers. Owing to the well-developed information technology, copious amounts of measurable process data can be easily collected, and such data can then be employed to predict and assess system behaviors and make control decisions, especially for the establishment and development of learning MPC.

For the application of MPC design in online regulation or tracking control problems, several studies have attempted to develop an accurate model, and realize adequate uncertainty description of linear or nonlinear plants of the processes [3–5]. In this study, we employ the data-driven learning technique specified in [6] to iteratively approximate the dynamical parameters, without requiring a priori knowledge of system matrices. The proposed MPC approach can predict and optimize the future behaviors using multi-order derivatives of control input as decision variables. Because the proposed algorithm can obtain a linear system model at each sampling, it can adapt to the actual dynamics of time-varying or nonlinear plants. This methodology can serve as a data-driven identification tool to study adaptive optimal control problems for unknown complex systems.

Problem formulation. In this study, we consider a continuous-time industrial process given by

$$\dot{x}(t) = Ax(t) + Bu(t) \triangleq \mathcal{H}\left(x(t), u(t)\right)\Theta, \quad (1)$$

where $t \geqslant t_0$, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ are the system state and input, respectively. $\mathcal{H}(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times(n^2+mn)}$ is defined as $\mathcal{H}(x,u) \triangleq [(x\otimes I_n)^{\rm T}\ (u\otimes I_n)^{\rm T}]$, where $\otimes$ denotes the Kronecker product. Θ denotes the vector of the system parameters given by $\Theta \triangleq [\mathrm{vec}(A)^{\rm T}\ \mathrm{vec}(B)^{\rm T}]^{\rm T} \in \mathbb{R}^{n^2+nm}$, where $A \in \mathbb{R}^{n\times n}$ is the system matrix, $B \in \mathbb{R}^{n\times m}$ is the input matrix, and $\mathrm{vec}(\cdot)$ denotes the vectorization operator, that is, $\mathrm{vec}(P) = [p_1^{\rm T}, \ldots, p_m^{\rm T}]^{\rm T}$, where $p_i \in \mathbb{R}^n$ is the ith column of a matrix $P \in \mathbb{R}^{n\times m}$. We assume that (A, B) is controllable and (A, C) is observable.

In this study, we consider the following input constraint: $u \in \mathcal{U} \subset \mathbb{R}^m$, where $\mathcal{U}$ denotes a nonempty compact convex set and contains the

* Corresponding author (email: dwli@sjtu.edu.cn)

origin as its interior point. In this case, as Θ is unknown, the primary objective of this study is to design a data-driven MPC formulation to obtain an open-loop optimal control policy that tracks a given reference x_{d} and, at each sampling time $t_k, k = 1, 2, \ldots$, minimizes the following cost function

$$J(x(t_k), \hat{u}_k(\cdot)) = \int_{t_k}^{t_k+T} \left(\|e(\tau)\|_Q^2 + \|\hat{u}_k(\tau)\|_R^2\right)\mathrm{d}\tau + \Phi(e(t_k+T)), \tag{2}$$

where $e(\cdot) = x(\cdot) - x_{\mathrm{d}}(\cdot)$ denotes the error, $\Phi(\cdot)$ denotes the terminal cost, and $Q = Q^{\mathrm{T}} \succ 0$ and $R = R^{\mathrm{T}} \succeq 0$ are the symmetric weighting matrices.

Methodology. To facilitate MPC design, at time $t = t_k$, the state x and the parameter Θ of the predicted model over the moving horizon $[t, t+T]$ are both learned from the input-output measurements, using a data-driven learning technique. In this study, we consider two situations.

- All the states and input information are available to us. Then, by rearranging (1), we have the linear error system in the form

$$\mathcal{F}(t) = \mathcal{G}(t)\hat{\Theta}, \quad \forall t \in \mathbb{R}_{\geqslant 0}, \tag{3}$$

where $\hat{\Theta}$ is an estimate of the unknown parameter Θ; the matrices $\mathcal{F}(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^n$ and $\mathcal{G}(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^{n \times (n^2+mn)}$ are defined as

$$\mathcal{F}(t) = \begin{cases} x(t) - x(t-\delta), & t \in [\delta, \infty), \\ 0, & t < \delta, \end{cases}$$
$$\mathcal{G}(t) = \left[(\Xi_x(t) \otimes I_n)^{\mathrm{T}} \;\; (\Xi_u(t) \otimes I_n)^{\mathrm{T}}\right],$$

where δ denotes the sampling period and the vectors $\Xi_x(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^n$ and $\Xi_u(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^m$ are defined as

$$\Xi_x(t) = \begin{cases} \int_{t-\delta}^{t} x(\tau)\mathrm{d}\tau, & t \in [\delta, \infty), \\ 0, & t < \delta, \end{cases}$$
$$\Xi_u(t) = \begin{cases} \int_{t-\delta}^{t} u(\tau)\mathrm{d}\tau, & t \in [\delta, \infty), \\ 0, & t < \delta. \end{cases}$$

- Only partial states and input information are available. We assume the first $q = n/2 < n$ components of the states are available, denoted by $\xi \in \mathbb{R}^q$, and then the pair (A, B) has the form

$$A = \begin{bmatrix} 0_{q\times q} & I_q \\ A_1 & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0_{q\times m} \\ B_1 \end{bmatrix}. \tag{4}$$

Then, the linear error system is given by

$$\mathcal{F}_1(t) = \mathcal{G}_1(t)\hat{\Theta}_1, \quad \forall t \in \mathbb{R}_{\geqslant 0}, \tag{5}$$

where $\hat{\Theta}_1$ is an estimate of the unknown parameter $\Theta_1 = [\mathrm{vec}(A_1)^{\mathrm{T}} \; \mathrm{vec}(A_2)^{\mathrm{T}} \; \mathrm{vec}(B_1)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2q^2+mq}$. $\mathcal{F}_1(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^q$ and $\mathcal{G}_1(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^{q\times(2q^2+mq)}$ are defined as

$$\mathcal{F}_1(t) = \begin{cases} \xi(t-\delta_2-\delta_1) - \xi(t-\delta_1) + \xi(t) \\ \quad -\xi(t-\delta_2), & t \in [\delta_1+\delta_2, \infty), \\ 0, & t < \delta_1+\delta_2, \end{cases}$$
$$\mathcal{G}_1(t) = \left[(\Xi_p(t) \otimes I_n)^{\mathrm{T}} \;\; (\Xi_v(t) \otimes I_n)^{\mathrm{T}} \;\; (\Xi_u^1(t) \otimes I_n)^{\mathrm{T}}\right],$$

where δ_1 and δ_2 ($\delta_1 \neq \delta_2$) denote the different periods; $\Xi_p(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^q$, $\Xi_v(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^q$, and $\Xi_u^1(\cdot) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}^m$ are defined as

$$\Xi_p(t) = \begin{cases} \int_{t-\delta_2}^{t}\int_{\tau-\delta_1}^{\tau} \xi(\varsigma)\mathrm{d}\varsigma\mathrm{d}\tau, & t \in [\delta_1+\delta_2, \infty), \\ 0, & t < \delta_1+\delta_2, \end{cases}$$
$$\Xi_v(t) = \begin{cases} \int_{t-\delta_2}^{t} \xi(\tau)\mathrm{d}\tau - \int_{t-\delta_1-\delta_2}^{t-\delta_1} \xi(\tau)\mathrm{d}\tau, \\ \quad t \in [\delta_1+\delta_2, \infty), \\ 0, \quad t < \delta_1+\delta_2, \end{cases}$$
$$\Xi_u^1(t) = \begin{cases} \int_{t-\delta_2}^{t}\int_{\tau-\delta_1}^{\tau} u(\varsigma)\mathrm{d}\varsigma\mathrm{d}\tau, & t \in [\delta_1+\delta_2, \infty), \\ 0, & t < \delta_1+\delta_2. \end{cases}$$

Further, from (3) and using the measurements, for a positive integer $l \leqslant k$, we define the vector $\Gamma_k \in \mathbb{R}^{ln}$ and matrix $\Psi_k \in \mathbb{R}^{ln\times(n^2+mn)}$ such that

$$\Gamma_k \triangleq \left[\mathcal{F}^{\mathrm{T}}(t_0), \mathcal{F}^{\mathrm{T}}(t_1), \ldots, \mathcal{F}^{\mathrm{T}}(t_l)\right]^{\mathrm{T}},$$
$$\Psi_k \triangleq \left[\mathcal{G}^{\mathrm{T}}(t_0), \mathcal{G}^{\mathrm{T}}(t_1), \ldots, \mathcal{G}^{\mathrm{T}}(t_l)\right]^{\mathrm{T}},$$

where $0 \leqslant t_0 < t_1 < \cdots < t_l$ and $t_i = i\delta, i = 0, 1, \ldots, l$. Then, Eq. (3) implies the linear equation

$$\Gamma_k = \Psi_k\hat{\Theta}. \tag{6}$$

Notice that if Ψ_k has full column rank, Eq. (6) can be directly solved as

$$\hat{\Theta} = \left(\Psi_k^{\mathrm{T}}\Psi_k\right)^{-1}\Psi_k^{\mathrm{T}}\Gamma_k. \tag{7}$$

Similarly, for (5), we let $\delta_1 = \delta$ and $\delta_2 = 2\delta$; thus, we obtain the same results for $\hat{\Theta}_1$. To guarantee $\mathrm{rank}(\Psi_k^{\mathrm{T}}\Psi_k) = n^2 + nm$, we let the states and inputs collected over a sufficiently large number of data samples be $l \gg n^2 + nm$. In practice, we assume that there exists a nominal control input $u = -K_0 x$, where K_0 denotes a stabilizing feedback gain matrix, such that Γ_k and Ψ_k in (6) can be implemented using $2l$ integrators to collect information about the states and inputs. Using $\mathfrak{D}_k \triangleq \bigcup_{i=0}^{l}\{\mathcal{F}(t_i), \mathcal{G}(t_i)\}$ by (7), we have $\hat{\Theta} = [\mathrm{vec}(\hat{A})^{\mathrm{T}} \; \mathrm{vec}(\hat{B})^{\mathrm{T}}]^{\mathrm{T}}$, which can be used to predict and optimize the future behaviors over a finite horizon $[0, T]$.

First, let us consider the MPC formulation. As mentioned in [4], the majority (if not all) of existing formulations consider only $u(\cdot)$ as the decision variable. For (1), we can extend it to a higher order derivative of $u(\cdot)$, that is,

$$\boldsymbol{u}_k(t) \triangleq \left[\hat{u}_k^{\mathrm{T}}(t), (\hat{u}_k^{[1]})^{\mathrm{T}}(t), \ldots, (\hat{u}_k^{[r]})^{\mathrm{T}}(t)\right]^{\mathrm{T}}, \quad (8)$$

with some control order $r \in \mathbb{N}_+$ larger than $\rho \geqslant 1$, where ρ denotes the input relative degree of (1). This will improve the efficacy of our learning MPC, and the first term of $\boldsymbol{u}_k(\cdot)$ in (8) is the control input $\hat{u}_k(\cdot)$ that is to be optimized in (2). Then, we let $\tilde{B} = \mathrm{vec}^{-1}(\hat{B})$, $\tilde{A} = \mathrm{vec}^{-1}(\hat{A})$, where $\mathrm{vec}^{-1}(\cdot)$ denotes the inverse operation of $\mathrm{vec}(\cdot)$, and define the following matrices:

$$\mathcal{A}_1 \triangleq \left[I \ \tilde{A}^{\mathrm{T}} \ \cdots \ (\tilde{A}^{\rho-1})^{\mathrm{T}}\right]^{\mathrm{T}},$$

$$\mathcal{A}_2 \triangleq \left[(\tilde{A}^{\rho})^{\mathrm{T}} \ (\tilde{A}^{\rho+1})^{\mathrm{T}} \ \cdots \ (\tilde{A}^{r})^{\mathrm{T}}\right]^{\mathrm{T}}, \quad \text{and}$$

$$\mathcal{B} \triangleq \begin{bmatrix} \tilde{A}^{\rho-1}\tilde{B} & 0 & \cdots & 0 \\ \tilde{A}^{\rho}\tilde{B} & \tilde{A}^{\rho-1}\tilde{B} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{A}^{r}\tilde{B} & \tilde{A}^{r-1}\tilde{B} & \cdots & \tilde{A}^{\rho-1}\tilde{B} \end{bmatrix}.$$

At time instant $t = t_k$, the MPC formulation can be given by (9) (see Appendix A for details).

$$\boldsymbol{u}_k^{\star}(\cdot) = \arg\min_{\boldsymbol{u}_k(\cdot)} J(x(t_k), \boldsymbol{u}_k(\cdot)) \quad (9)$$

$$\text{s.t.} \quad x(t_k+\tau) = \left[T_1(\tau) \ T_2(\tau)\right] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \tau \in [0, T],$$

$$X_1 = \mathcal{A}_1 x, \quad X_2 = \mathcal{A}_2 x + \mathcal{B}\boldsymbol{u}_k,$$

$$\hat{x}(t_k) = x(t_k), \quad \hat{u}_k(t) \in \mathcal{U},$$

where $X_1 = [x^{\mathrm{T}}, (x^{[1]})^{\mathrm{T}}, \ldots, (x^{[\rho-1]})^{\mathrm{T}}]^{\mathrm{T}}$, $X_2 = [(x^{[\rho]})^{\mathrm{T}}, (x^{[\rho+1]})^{\mathrm{T}}, \ldots, (x^{[r]})^{\mathrm{T}}]^{\mathrm{T}}$, $T_1(\tau) = [1, \tau, \ldots, \frac{\tau^{\rho-1}}{(\rho-1)!}]$, $T_2(\tau) = [\frac{\tau^{\rho}}{\rho!}, \ldots, \frac{\tau^{r}}{r!}]$, $T_3(\tau) = [T_1(\tau), T_2(\tau)]$, $X_{1,\mathrm{d}} = [x_{\mathrm{d}}^{\mathrm{T}}, (x_{\mathrm{d}}^{[1]})^{\mathrm{T}}, \ldots, (x_{\mathrm{d}}^{[\rho-1]})^{\mathrm{T}}]$, $X_{2,\mathrm{d}} = [(x_{\mathrm{d}}^{[\rho]})^{\mathrm{T}}, (x_{\mathrm{d}}^{[\rho+1]})^{\mathrm{T}}, \ldots, (x_{\mathrm{d}}^{[r]})^{\mathrm{T}}]$, and

$$J(x(t_k), \boldsymbol{u}_k(\cdot)) = \tilde{X}_1^{\mathrm{T}}\mathcal{T}_{1,1}\tilde{X}_1 + 2\tilde{X}_1^{\mathrm{T}}\mathcal{T}_{1,2}\tilde{X}_2 + \tilde{X}_2^{\mathrm{T}}\mathcal{T}_{2,2}\tilde{X}_2 + \boldsymbol{u}_k{}^{\mathrm{T}}\mathcal{T}\boldsymbol{u}_k + \varPhi(\tilde{X}_i(t_k+T)),$$

with $\tilde{X}_i = X_i - X_{i,\mathrm{d}}$, $\Xi_i(\tau) = \sqrt{Q}T_i(\tau)$, $\mathcal{T}_{i,j} = \int_0^T \Xi_i^{\mathrm{T}}\Xi_j \mathrm{d}\tau$, $i, j \in \{1, 2\}$, and $\mathcal{T} = \int_0^T T_3^{\mathrm{T}} R T_3 \mathrm{d}\tau$. This is a standard quadratic programming (QP) problem that can be solved by many available tools. In particular, for the case with box constraints, with consideration of the possible model error with the data-driven method, we can handle the input constraints by the sub-optimal method in Appendix B. We thus have the optimal control policy $\hat{u}_k^{\star}(t) = I_u\boldsymbol{u}_k^{\star}$, where $I_u = [1, 0, \ldots, 0]_{1\times(r+1)}$. We summarize this proposed approach as Algorithm 1 in Appendix C.

Furthermore, we consider the linear error system with the control policy $\hat{u}_k^{\star}(t)$ applied to (1). For the actual state trajectory, we have the continuous error as

$$w(t) = \mathcal{H}\left(x(t), \hat{u}_k^{\star}(t)\right)(\hat{\Theta} - \Theta), \quad (10)$$

where $t \in [t_k, t_k+T]$. In Appendix D, we show that $w(t)$ is bounded and has the upper bounded rate of change with time t; if $\tilde{A} = A$ and $\tilde{B} = B$, then $w(t) = 0, t \geqslant 0$; with the updated control policy $\hat{u}_k^{\star}(t)$ at each time $t = t_k$, $\lim_{t\to\infty} w(t) = 0$, which implies the asymptotic stability of the closed-loop system.

Simulation results. An illustrative numerical example for two continuous stirred tank reactor (CSTR) systems is provided to validate the performance of the proposed approach. More details and discussion are presented in Appendix E.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61973214, 61590924, 61963030), Natural Science Foundation of Shanghai (Grant No. 19ZR1476200), and National Key Basic Research Special Foundation of China (Grant No. 2014CB249200). The authors would like to thank Prof. Zhong-Ping JIANG and his CAN Lab at Tandon School of Engineering, New York University, Brooklyn, NY, USA, for many inspirations and the help of this work.

Supporting information Appendixes A–E. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Xi Y G, Li D W, Lin S. Model predictive control — status and challenges. Acta Automatica Sin, 2013, 39: 222–236
2 Mayne D Q, Rawlings J B, Rao C V, et al. Constrained model predictive control: stability and optimality. Automatica, 2000, 36: 789–814
3 Li D, Xi Y, Gao F. Synthesis of dynamic output feedback RMPC with saturated inputs. Automatica, 2013, 49: 949–954
4 Chen W H, Ballance D J, Gawthrop P J. Optimal control of nonlinear systems: a predictive control approach. Automatica, 2003, 39: 633–641
5 Albin T. Benefits of model predictive control for gasoline airpath control. Sci China Inf Sci, 2018, 61: 070204
6 Jiang Y, Jiang Z P. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. Automatica, 2012, 48: 2699–2704