Enhanced Gaussian Process Regression for Active Learning Model-based Predictive Control

Rui Ren[1], Shaoyuan Li [2]

1. Shanghai Jiao Tong University, Shanghai, 200240, China
2. Shanghai Jiao Tong University, Shanghai, 200240, China

E-mail: syli@sjtu.edu.cn

Abstract

Learning-based MPC methods consider automatic adjustment of the system model during operation using machine learning methods, which passively leverage the available system data and result in slow learning with lacking of informative data. In this paper, an active learning-based MPC scheme is proposed with introducing an information content cost in the optimization problem, aiming to improve the learning ability as well as the closed-loop control performance.

Methods

In this paper, firstly, a simple state-space model is used to represent the prior model and a GPR model is applied to learn the residual uncertainty, and the conventional LB-MPC scheme is proposed. Then, an information objective, which can be used to measure the information gain of the explorative data, is introduced. Based on this, the optimal control problem can be reformulated with the additive information objective so that the inputs can not only regulate the system dynamics but also have a probing effect that generates informative closed-loop data. Finally, the active learning-based MPC(AL-MPC) scheme is proposed and illustrated in Fig.1.

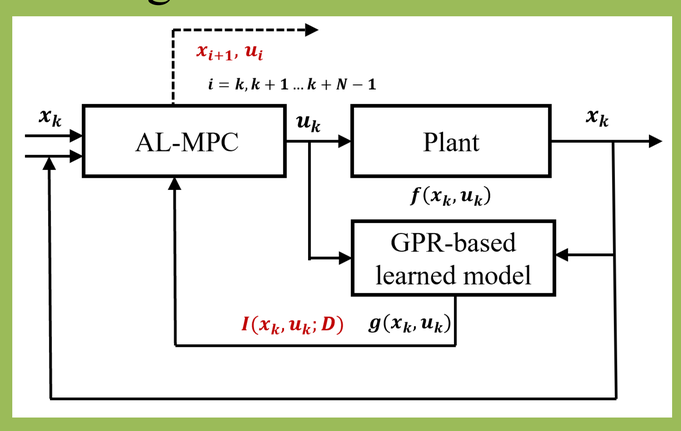

Fig. 1. block diagram illustrating AL-MPC

(1)**Passive learning part**: Automatic adjustment of the system model during operation using Gaussian Process Regression.

(2)**Active learning part**: Information content is introduced as dual objective for explicitly rewarding the system probing. The dual objective is derived from information theory:

$$I(z_i, D) = H(D) - H(D \cup z_i)$$

And the cost function of the OCP is:

$$\min_U J = \sum_{i=0}^{N-1} [x_{i|k}^T Q x_{i|k} + u_{i|k}^T R u_{i|k}] - \omega \sum_{i=0}^{N_e-1} c(I(z_i, D))$$

Experimental Results

Experiment is respectively carried out in a Van der Pol Oscillator. The AL-MPC algorithm is compared to Learning-based MPC which only passively leverage the available system data. The equation of the system dynamics is as follows:

$$x_1 = (1 - x_2^2)x_1 - x_2 + u$$
$$x_2 = x_1$$

In fig.2, the comparison of the control input obtained using different algorithms suggests that AL-MPC scheme does introduce extra active excitation to the system. In Fig.3, the difference in state trajectory shows that the better learning ability and the effectively improvement of tracking effect of AL-MPC.

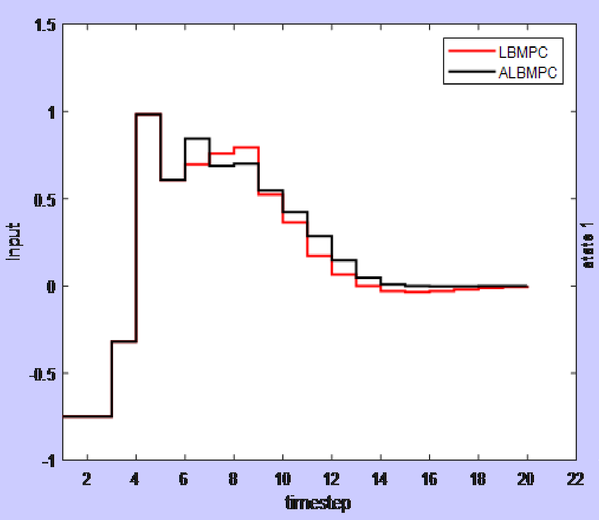

Fig.2. Comparison of control input of two algorithms

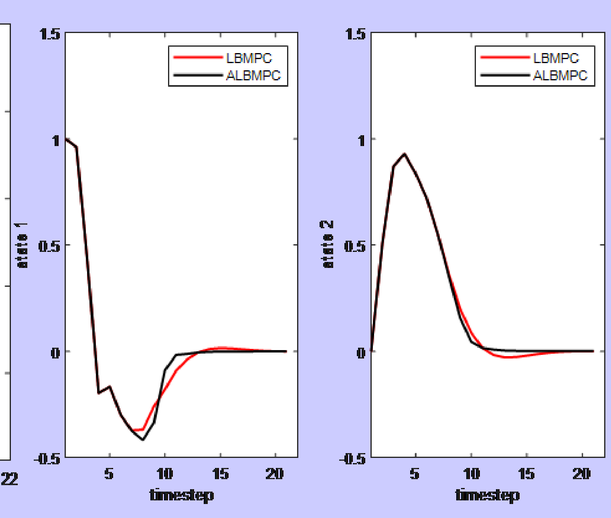

Fig.3. Comparison of state trajectory of two algorithms

Conclusion

The proposed algorithm could effectively improve the learning ability as well as the closed-loop control performance. Experimental results show that the AL-MPC scheme does introduce active excitation to the system and the tracking effect of the controller is improved effectively.

Further research will focus on the satisfaction of input and state constraints which is lack in most learning-based methods. Then the safety guarantees on the active learning-based MPC will be studied.