model-predictive control (MPC)

Annual Review of Control, Robotics, and Autonomous Systems

Learning-Based Model Predictive Control: Toward Safe Learning in Control

Lukas Hewing, Kim P. Wabersich, Marcel Menner, and Melanie N. Zeilinger

Institute for Dynamic Systems and Control, ETH Zurich, Zurich 8092, Switzerland; email: lhewing@ethz.ch, wkim@ethz.ch, mmenner@ethz.ch, mzeilinger@ethz.ch

---

- consider stochastic optimal control problem

$$\text{(SOCP)} \quad \min_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \mathcal{L}_t(x_t, u_t)\right]$$

$$\text{s.t.} \quad x^+ = f(x, u, w; \theta) \quad \text{- dynamics}$$

$$u_t = \pi(x_0, \cdots, x_t) \quad \text{- policy}$$

$u_t = \pi(x_0, \cdots, x_t)$ – policy

$P(x \in X) \geq p$ – safe states w/ high probability

$P(u \in \mathcal{U}) \geq q$ – safe inputs w.h.p.

where $w \in W$ is random variable and $\theta \in \Theta$ is parameter

idea: solve & re-solve simpler shorter time horizon problem ("receding-horizon" control)

- let $z_{\tau|t}$ denote "τ-step-ahead" prediction of z_t

$$(\text{MPC}) \quad \min_u \; \ell(x_{N|t}, u_{N|t}) + \sum_{\tau=0}^{N-1} \mathcal{L}_{t+\tau}(x_{\tau|t}, u_{\tau|t})$$

$$\text{s.t.} \quad x_{\tau+1|t} = \tilde{f}(x_{\tau|t}, u_{\tau|t})$$

$$x \in X, \; u \in \mathcal{U}, \; x_{0|t} = x_t, \; \overbrace{x_{N|t} \in X_f}^{\text{terminal state constraint}}$$

- two key properties sought in MPC theory:
  - (i) feasibility: if x_t is feasible for (MPC) then x_{t+1} generated by $u_{0|t}$ is feasible
  - (ii) stability: optimal value v^* for (MPC) is a Lyapunov function for closed-loop dynamics
- can consider learning in key ways:

1°. learning f (Aswani et al 2013)

- assume $x^+ = f(x,u) = Ax + Bu + g(x,u)$
  where $x \in X$, $u \in \mathcal{U}$, $g(x,u) \in W$, X & $\mathcal{U}$ & W are <u>polytopes</u>
  (ie intersections of half-spaces)
- use robustness inherent in MPC to guarantee feasibility
  even with incorrect model ($x^+ = Ax + Bu$)
- learn model $\tilde{g}$ using basis functions, neural networks,
  or nonparametric estimator that guarantees boundedness

<u>ex</u>: Nadaraya-Watson: parameters $\lambda, \eta > 0$

- let $z_i = \begin{bmatrix} x_i \\ u_i \end{bmatrix}$, $y_i = x_{i+1} - (Ax_i - Bu_i)$, $\xi_i(z) = \| z - z_i \|^2 \frac{1}{\eta}$

  and $\kappa : \mathbb{R} \to \mathbb{R}_{\geq 0}$ smooth, even, finite support

- then $\tilde{g}(x,u) = \dfrac{1}{\lambda + \sum_i \kappa(\xi_i(x,u))} \sum_i y_i \kappa(\xi_i(x,u))$

$$= \arg\min_d \sum_i \kappa(\xi_i(x,u)) \| y_i - d \|^2 + \lambda \| d \|^2$$

2°. guaranteeing safety while learning

<u>idea</u>: given learned policy π_0, pass through "safety filter" MPC

$$\min_{\pi} \| \pi - \pi_0 \|$$
$$\text{s.t. } x^+ = f(x, u, w; \theta)$$

''

$$
\begin{aligned}
\text{s.t. } \quad & x^+ = f(x, u, w; \theta) \\
& u_t = \pi(x_0, \cdots, x_t) \\
& P(x \in X) \geq p \\
& P(u \in U) \geq q
\end{aligned}
$$