DTU

Industrial IoT for Digitization of Electronis Assets

Model Predictive Control and Imitation Learning

Agenda

- Introduction
- Overview of MPC
- System Model
- Objective Function
- Constraints and Optimization
- Learning Agent
- Conclusion

Overview of MPC

- Model Predictive Control (MPC) is an advanced method of process control that predicts the future behavior of a system.
- MPC uses a mathematical model to make predictions and optimize control actions.
- It handles multi-variable control problems with constraints effectively.

System Model

- The system is typically represented by a state-space model:

$$x_{k+1} = Ax_k + Bu_k + w_k$$

$$y_k = Cx_k + v_k$$

- x_k: state vector, u_k: control input, y_k: output.
- A, B, C: system matrices, w_k, v_k: process and measurement noise.

MPC: The Objective Function

Objective function to be minimized over a prediction horizon T:

$$\begin{aligned} \min_{u,x,y} \quad & \sum_{k=0}^{T} \|y_k - r_k\|_Q^2 + \|u_k\|_R^2 \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k, \quad \forall\, \mathbf{k} \in \{\mathbf{1}, \dots, \mathbf{T}\} \\ & y_k = Cx_k + Du_k, \quad \forall\, \mathbf{k} \in \{\mathbf{1}, \dots, \mathbf{T}\} \end{aligned}$$

- y_k: predicted output, r_k: reference output, u_k: predicted control input.
- Q, R: weighting matrices for tracking error and control effort.

https://www.youtube.com/watch?v=sKQqAeJMzBk&t=1838s&ab_channel=ControlMeetsLearning

Constraints and Optimization

- MPC can handle various constraints like input, state, and output constraints.
- Optimization problem solved at each step to find the best control sequence.
- Receding horizon principle: Only the first control action is implemented and then the horizon is updated.

https://www.youtube.com/watch?v=YwodGM2eoy4&ab_channel=SteveBrunton

Example

Let's consider a simple example, where the goal is to control the temperature of a room.

- The temperature of the room is the output variable.
- The control input is the power of the heater.
- The reference output is the desired temperature.

And that the system is described by the following equation:

$$T_{new} = T_{current} + P_{[\%]\ heater} \times 0.1 \times \Delta T$$

Example

Let's consider a simple example, where the goal is to control the temperature of a room.

- The temperature of the room is the output variable.
- The control input is the power of the heater.
- The reference output is the desired temperature.

And that the system is described by the following equation:

$$T_{new} = T_{current} + P_{[\%]\ heater} \times 0.1 \times \Delta T$$

Question: How can we control the temperature of the room?

Example

$$T_{new} = T_{current} + P_{[\%]\ heater} \times 0.1 \times \Delta T$$

Given:

- $T_{current} = 20\,^{\circ}\mathrm{C}$
- $T_{desired} = 22\,^{\circ}\mathrm{C}$
- $T_{output} = 18\,^{\circ}\mathrm{C}$
- $\Delta T = T_{output} - T_{current} = -2\,^{\circ}\mathrm{C}$

$$\textbf{HOUR 1:} \quad T_{new} = 20^{\circ}\mathrm{C} + \mathbf{U}\% \times 0.1\frac{^{\circ}\mathrm{C}}{\%} - 2 \longrightarrow \mathbf{U} = 40\%$$

So we need to set the heater to 40% for the first hour. Once the first hour is over, we can update the system model and repeat the process.

A MPC controller to increase the Energy Efficiency of a Wastewater Station

$$\min_{\omega_1,\omega_2} \sum_{k=1}^{N} (E_{PV1,k} + E_{PV2,k}) + w_h f_h(h_k)$$

subject to:

$$\begin{bmatrix} Q_{l,k} \\ \vdots \\ Q_{l,k+N} \end{bmatrix} = f_{Ql} \begin{bmatrix} \omega_{l,k} \\ \vdots \\ \omega_{l,k+N} \end{bmatrix}, \quad l \in \{1,2\}$$

$$\begin{bmatrix} h_{k+1} \\ \vdots \\ h_{k+N+1} \end{bmatrix} = \begin{bmatrix} h_k \\ \vdots \\ h_{k+N} \end{bmatrix} + \frac{T_s}{A} \begin{bmatrix} \hat{Q}_{in,k} - \sum_l Q_{l,k} \\ \vdots \\ \hat{Q}_{in,k+N} - \sum_l Q_{l,k+N} \end{bmatrix}$$

$$0 \leq \begin{bmatrix} \omega_{l,k} & \dots & \omega_{l,k+N} \end{bmatrix}^T \leq \omega_{l\,\max}$$

$$h_{\min} \leq \begin{bmatrix} h_k & \dots & h_{k+N} \end{bmatrix}^T \leq h_{\max}$$

Summary:

- MPC is a powerful control strategy for systems with predictive models.
- Its ability to anticipate and optimize future behavior makes it applicable in various fields.
- The optimization formulation is key to its effectiveness.

Challenges in MPC Deployment

- Solving optimization problems online is computationally demanding.
- High-dimensional systems pose a challenge due to the complexity and required computational resources.
- Strict latency requirements and limited computational or energy resources can impede the deployment of MPC.

Ahn, Kwangjun, et al. "Model Predictive Control via On-Policy Imitation Learning." Learning for Dynamics and Control Conference. PMLR, 2023.

Interactive Data Collection Scheme

- A scheme is proposed to interactively collect data from a system in feedback with an MPC controller.
- The goal is to learn an explicit controller that directly maps states to inputs.
- This methodology aligns with imitation learning approaches in the reinforcement learning domain.

Ahn, Kwangjun, et al. "Model Predictive Control via On-Policy Imitation Learning." Learning for Dynamics and Control Conference. PMLR, 2023.

Imitation Learning and MPC

- Imitation learning involves learning an explicit controller that maps states to inputs.
- It is suitable for MPC as it can query the MPC for the next input at any state by solving the optimization problem.
- This process aligns with explicit MPC, which pre-computes solutions to optimization problems for runtime efficiency.

Ahn, Kwangjun, et al. "Model Predictive Control via On-Policy Imitation Learning." Learning for Dynamics and Control Conference. PMLR, 2023.

Learning Controllers with High Fidelity to MPC

- The goal is to learn a map from states to inputs that encapsulates the strategy of an MPC controller.
- Unlike methods that collect data pre-learning, our approach interacts with the system dynamics to avoid distribution shift.
- This interaction prevents sub-optimal performance and error compounding, which are common in non-interactive imitation learning.
- Our approach aims for a learned controller that matches MPC performance with high probability.

Ahn, Kwangjun, et al. "Model Predictive Control via On-Policy Imitation Learning." Learning for Dynamics and Control Conference. PMLR, 2023.

Imitation Learning from an Expert

Imitation learning aims to learn the optimal controller $\hat{\pi}$ by minimizing the loss function $L(\pi)$ with respect to the MPC controller.

$$\min_{W,b} J(W,b) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\omega}_i^2 - \omega_i^{opt}) + \lambda \sum_k \sum_j w_{k,j}^2 \quad l \in \{1,2\}$$

Quattrociocchi, Alessandro, et al. "Energy Efficiency Optimization of a Wastewater Pumping Station Through IoT and AI: A Real-World Application of Digital Twins." IECON 2023-49th Annual Conference of the IEEE Industrial Electronics Society. IEEE, 2023.

All the Loss function seen so far...

Paramaters Estimation in ARX Model

$$\mathcal{L}(\theta, Z^N) = \sum_{k=0}^{N-1} (y(t) - \hat{y}(t|\theta))^2 = \sum_{k=0}^{N-1} (y(t) - \varphi'(t)\theta)^2$$

All the Loss function seen so far...

MPC Objective Function

$$\min_{u,x,y} \quad \sum_{k=0}^{T} \|y_k - r_k\|_Q^2 + \|u_k\|_R^2$$

All the Loss function seen so far...

Neural Network Loss Function

$$\min_{W,b} J(W,b) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\omega}_i^2 - \omega_i^{opt}) + \lambda \sum_k \sum_j w_{k,j}^2 \quad l \in \{1,2\}$$

DTU

Some Resources:

- [Model Predictive Control in a Nutshell]
- [Visualize and Draw the Structure of a Neural Network]
- [Visualize the Training of a Neural Network Online]
- [Reinforcement Learning Agent Simulation]