DeePC and DRO Notes

Alessandro Quattrociocchi

December 4, 2024

Contents

1 Problem Formulation

Consider a deterministic LTI system whose model is unknown. Let $T_f \in \mathbb{Z}_{>0}$ be the prediction horizon. Let $f_1 : \mathbb{R}^{mT_f} \to \mathbb{R}_{\geq 0}$, and $f_2 : \mathbb{R}^{pT_f} \to \mathbb{R}_{\geq 0}$ be a cost function on the future inputs and, respectively, outputs. We aim at finding a input sequece $\text{col}(u_t, \ldots, u_{t+T_f-1}) \in \mathbb{R}^{mT_f}$, such that $\text{col}(y_t, \ldots, y_{t+T_f-1}) \in \mathbb{R}^{pT_f}$ minimizes the cost $f_1 + f_2$, and the contraints satisfied, i.e., $u \in \mathcal{U}$ and $y \in \mathcal{Y}$, with $\mathcal{U} \subseteq \mathcal{R}^{mT_f}$ and $\mathcal{Y} \subseteq \mathcal{R}^{pT_f}$.

1.1 Conventional Formulation

Using conventional input output state space representation.

$$\begin{aligned}
\min_{u,y,x} \quad & f_1(u) + f_2(y) \\
\text{s.t.} \quad & \forall k \in \{0, \ldots, T_f - 1\} \\
& x_{k+1} = Ax_k + Bu_k \\
& y_k = Cx_k + Du_k \\
& x_0 = \hat{x}_t \\
& u \in \mathcal{U},\ y \in \mathcal{Y}
\end{aligned} \tag{1}$$

1.2 Deterministic DeePC Formulation

From data we contruct $\hat{U}_p, \hat{U}_f, \hat{Y}_p, \hat{Y}_f$.

$$\begin{aligned} \min_g \quad & f_1(u) + f_2(y) \\ \text{s.t.} \quad & \begin{bmatrix} \hat{U}_p \\ \hat{Y}_p \\ \hat{U}_f \\ \hat{Y}_f \end{bmatrix} g = \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{bmatrix} \\ & u \in \mathcal{U},\ y \in \mathcal{Y} \end{aligned} \tag{2}$$

- U_p, Y_p represent the past input-output trajectories.
- U_f, Y_f represent the future input-output trajectories.

1.2.1 Remark I: *DeepC representation of predictor.*

$$\begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{bmatrix} = H(w)g, \quad \text{with} \quad H(w) = \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} \tag{3}$$

where $H(w)$ is the suitably partitioned Hankel matrix, or

$$y = Y_f g \tag{4}$$

where g is computed from:

$$\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} g = \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix}.$$

The explicit solution for y is the superposition of a particolar solition and the homogenous term:

$$y = Y_f \cdot \left(\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}^+ \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix} + \ker\left(\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} \right) \right),$$

where $\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}^+$ is the pseudoinverse of the Hankel matrix, and ker represents the kernel of the matrix.

1.2.2 Remark: *Regularization in DeePC(rank-deficient Hankel matrices)*

When the Hankel matrix H is rank-deficient (i.e., H has more columns than rows), there are infinitely many solutions g that satisfy the system equation:

$$Hg = \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix}. \tag{5}$$

To find a noise-robust prediction, the Projection Regularizer is introduced. The regularizer penalizes the portion of g that lies in the nullspace of H. This is achieved by minimizing:

$$\|(I-\Pi)g\|_2^2, \tag{6}$$

where:

- Π is the orthogonal projection matrix onto the row space of H,
- $(I-\Pi)g$ isolates the component of g in $\ker(H)$, the nullspace of H.

This ensures that the computed g is the smallest norm solution to:

$$y = Y_f g. \tag{7}$$

In the noiseless case ($\lambda_s = 0$) with regularization constant $\lambda_p \geq 0$, the optimization problem ensures that both u and y match the reference trajectory u_r, y_r, making it beneficial to use the regularizer with a sufficiently large λ_p.

2 Stochastic DeePC Formulation

Following the formulation in [2], we now extend the DeePC definition of the unknow LTI system by considering a disturbance vector $w_t \in \mathbb{R}^q$, such as:

$$\begin{cases} x_{t+1} = Ax_t + Bu_t + Ew_t \\ y_t = Cx_t + Du_t + Fw_t \end{cases} \tag{8}$$

where $E \in \mathbb{R}^{n\times q}$ and $F \in \mathbb{R}^{p\times q}$. Therefore, the system is subject to an **unknow** and **uncontrollable** disturbance, whose past trajectory can be measured but whose future evolutions are unknown. Similar to u^d and y^d, we can build the Hankel matrix of the disturbance $\mathcal{H}_{T_{ini+N}}$, splitted as follows:

$$\begin{bmatrix} W_{\mathrm{P}} \\ W_{\mathrm{F}} \end{bmatrix} := \mathcal{H}_{T_{\mathrm{ini}}+N}(w^{\mathrm{d}}) \tag{14}$$

with $W_{\mathrm{P}} \in \mathbb{R}^{qT_{\mathrm{ini}}\times H_c}$ and $W_{\mathrm{F}} \in \mathbb{R}^{qN\times H_c}$. Therefore, we assume exist $g \in \mathbb{R}^{H_c}$, so that:

$$\begin{bmatrix} U_{\mathrm{P}} \\ W_{\mathrm{P}} \\ Y_{\mathrm{P}} \\ U_{\mathrm{F}} \\ W_{\mathrm{F}} \\ Y_{\mathrm{F}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ w_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ w \\ y \end{bmatrix} \tag{9}$$

Unlike [2], we DO NOT assume $w_t \in [\underline{w}, \overline{w}]^q$. Why this? We assume that in optimal scheduling problems, a time series forecasting model is always subject to an error which in very few toy examples converges to white noise. Therefore, we assume the general case where the error term distribution w_t is not known a priori. Therefore, we assume $\hat{W}_{\mathrm{P}}$

and $\hat{W}_{\mathrm{F}}$ realizations of random variables W_{P} and W_{F}. Reformulating from [1], we define $\xi = (\xi_1^T, \ldots, \xi_{p,(T_{\mathrm{ini}}+T_{\mathrm{f}})})$, with ξ_i denoting the ith rows $(W_{\mathrm{p}}^T, W_{\mathrm{f}}^T)$.

Objective Function

The objective function formulated in Eq.1, is therefore reformulated into its robust conterpart, namely:

$$\min_{g} \sup_{\mathbb{Q}\in\hat{\mathcal{P}}} \mathbb{E}_{\mathbb{Q}} \left[f_1(\hat{U}_{\mathrm{f}} g) + f_2(Y_{\mathrm{f}} g)) \right]$$

$$\text{s.t.} \begin{bmatrix} U_{\mathrm{P}} \\ W_{\mathrm{P}} \\ Y_{\mathrm{P}} \\ U_{\mathrm{F}} \\ W_{\mathrm{F}} \\ Y_{\mathrm{F}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ w_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ w \\ y \end{bmatrix}$$

$$\sup_{\mathbb{Q}\in\hat{\mathcal{P}}} \mathrm{CVaR}_{1-\alpha}^{\mathbb{Q}}(h(y)) \leq 0. \tag{10}$$

References

[1] Jeremy Coulson, John Lygeros, and Florian Dörfler. Distributionally robust chance constrained data-enabled predictive control. *IEEE Transactions on Automatic Control*, 67(7):3289–3304, 2021.

[2] Linbin Huang, Jeremy Coulson, John Lygeros, and Florian Dörfler. Decentralized data-enabled predictive control for power system oscillation damping. *IEEE Transactions on Control Systems Technology*, 30(3):1065–1077, 2021.