COCO II DeePC Exercises

November 5, 2023

Exercise 1

Goals: develop intuition for

- why the rank of the block-Hankel matrices for MIMO systems is at most $mL + n$ for noiseless data

- what the data requirements are to get a rank-$(mL + n)$ Hankel matrix (persistency of excitation)

Let us consider a MIMO system given by the state-space equations

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k) + Du(k), \end{aligned} \tag{1}$$

with $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^m$, $y(k) \in \mathbb{R}^p$ for all $k \geq 0$, and matrices A, B, C, D of appropriate dimensions. The input-output behavior of the system for a trajectory of length L can be written as

$$\begin{bmatrix} u(0) \\ \vdots \\ u(L-1) \\ y(0) \\ \vdots \\ y(L-1) \end{bmatrix} = M \begin{bmatrix} u(0) \\ \vdots \\ u(L-1) \\ x(0) \end{bmatrix},$$

with some matrix $M \in \mathbb{R}^{L(m+p) \times Lm+n}$.

a) Express the matrix M with the extended observability and convolution matrices defined as

$$\mathcal{O}_L := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{bmatrix} \in \mathbb{R}^{Lp \times n}; \quad \mathcal{G}_L := \begin{bmatrix} D & 0 & \cdots & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{L-2}B & CA^{L-3}B & \cdots & CB & D \end{bmatrix} \in \mathbb{R}^{Lp \times Lm}.$$

For which values of L does M have full column rank, i.e. $\text{rank}(M) = Lm + n$?

b) We measure a T_d long trajectory (inputs and outputs) of the system given by the state-space equations in Eq. (1), and build the Hankel matrix $\mathcal{H}_L$ of length L. How should we choose T_d, so that the Hankel matrix is square?[1]

c) In the following scenarios what does the Hankel matrix look like in terms of the state-space matrices A and B, assuming $C = I_n$ and $D = 0_{p \times m}$?
What is the maximum possible rank of H_L in each case?

---

[1] Note, the Hankel matrix does not need to be square to have rank $mL + n$. In practice, additional columns do not hurt and may add some robustness to noise in the data.

1

i) zero initial condition $x(0) = 0$ and zero inputs $u(k) = 0$ for all $k \geq 0$

ii) nonzero initial condition $x(0) = x_0 \neq 0$ and zero inputs $u(k) = 0$ for $k \geq 0$

iii) zero initial condition $x(0) = 0$ and constant input $u(k) = c_u \in \mathbb{R}^m$ for all $k \geq 0$

iv) zero initial condition $x(0) = 0$ and time-varying input $u(k)$.

What is more important for getting an $(mL + n)$–rank Hankel matrix, the initial condition or the input?

Exercise 2

Goals: develop intuition for

- the effect of noise on Hankel matrices
- regularization in DeePC

A template for the code necessary for this exercise is available here:
`https://colab.research.google.com/drive/1d0d4Vl2zQhN4qQQin7vKQC6CuafMuMqM?usp=sharing`

Consider the double integrator system of the form

$$x(k+1) = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0.005 \\ 0.1 \end{bmatrix} u(k), \tag{2}$$

$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(k) + d(k). \tag{3}$$

a) Simulate the system for $T_d = 115$ time steps with

- random initial conditions $x(0) \sim \mathbb{N}(0, 4)$,
- random input $u(k) \sim \mathcal{N}(0, 4)$,

and either

- $d(k) \sim \mathcal{N}(0, 0.01)$, or
- $d(k) = 0$

for the measurement noise. Build a Hankel matrix $\mathcal{H}_L(w)$, $w = [u \; y]^\top$, with $L = 35$ for both the noisy and noise-free trajectories.

    i) What is the rank of the Hankel matrix with and without noise? Plot the singular values of the matrix in both cases.

    ii) What trajectories can DeePC possibly predict if the Hankel matrix has full row rank? Do all of these trajectories correspond to the true system dynamics?

b) Implement the following DeePC controller

$$\min_{u,y,g} \sum_{i=1}^{T_f} Q(y_i - y_{ref})^2 + Ru_i^2 + \lambda_\sigma \|\sigma\|_2^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} g = \begin{bmatrix} u_{ini} \\ y_{ini} + \sigma \\ u \\ y \end{bmatrix},$$

where

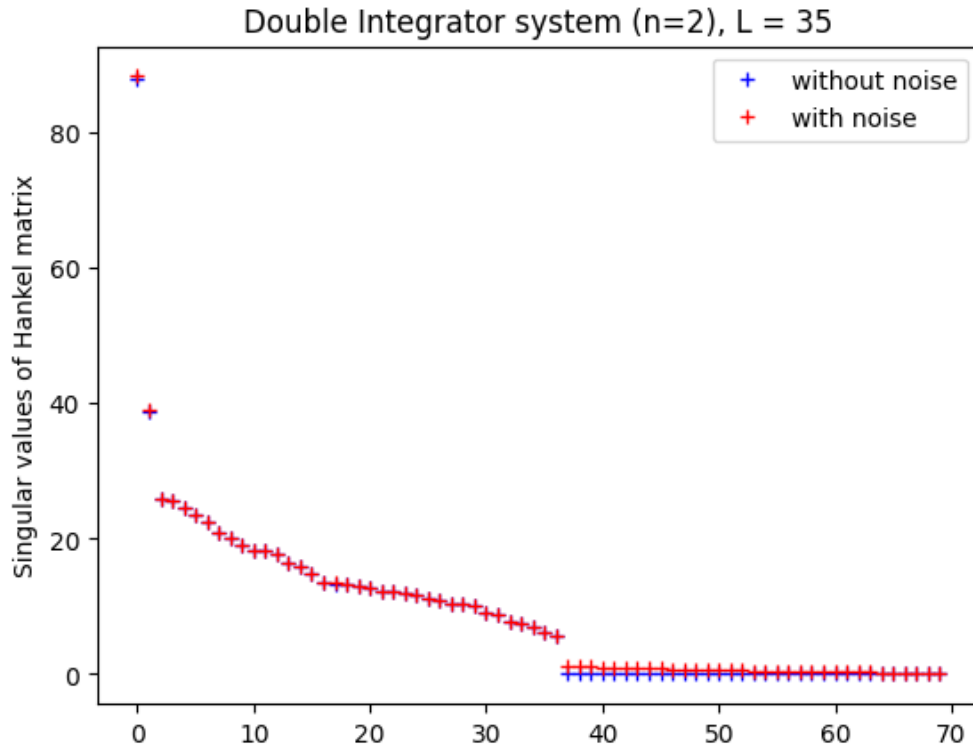- $y_{ref} = 1$ is the reference output to be tracked

Figure 1: Singular values of the Hankel matrix $\mathcal{H}_L(w)$ with and without noise

- $Q = 200$ and $R = 1$ are the weights for the output and input
- $T_f = 25$ is the prediction horizon
- $T_{ini} = 10$ steps is the initial trajectory
- The matrices U_p, U_f, Y_p, and Y_f are the Hankel matrices for the past and future input, and the **noisy** past and future output, respectively.

Since the initial output trajectory y_{ini} is noisy, we introduce the slack variable σ, with a 2-norm regularizer $\lambda_\sigma \|\sigma\|_2^2$ in the cost. *Hint: try large values for λ_σ ($10^3 - 10^5$).*

Simulate the closed-loop system with the DeePC controller for 100 steps starting using the noise-free system model (2), (3) as your groundtruth simulator and plot the closed-loop trajectory. Use

- zero initial input and output trajectory for the DeePC controller ($u_{\text{ini}} = \mathbf{0}$, $y_{\text{ini}} = \mathbf{0}$)
- an initial condition of zero for the groundtruth simulator ($x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$)

We do not expect this controller to work. Why?

c) Implement the following DeePC controller

$$\min_{u,y,g} \sum_{i=1}^{T_f} Q(y_i - y_{ref})^2 + Ru_i^2 + \lambda_g\|g\|_1^2 + \lambda_\sigma\|\sigma\|_2^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} g = \begin{bmatrix} u_{ini} \\ y_{ini} + \sigma \\ u \\ y \end{bmatrix},$$

To mitigate the effect of noise in the data matrices Y_p and Y_f, we include a 1-norm regularizer term $\lambda_g\|g\|_1^2$ in the cost.

Once again, simulate the closed-loop system with the DeePC controller for 100 steps starting with the same settings as part (b). Try different values for λ_g until the performance of the controller is satisfactory.

What do you observe when λ_g is small? Why?

What do you observe when λ_g is large? Why?

4

d) **Bonus:** Change the regularization term $\lambda_g \|g\|_1^2$ to the projection based regularizer[2] $\lambda_g \|(I - \Pi)g\|_2^2$, with $\Pi = \begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}^\dagger \begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}$. Increase the weight of the regularizer λ_g. What do you observe?

---

[2]Further details in Sec. IV/B in Dörfler et al. *Bridging direct and indirect data-driven control formulations via regularizations and relaxations*, IEEE Transactions on Automatic Control, 2022

Exercise 3

Goal: develop intuition for applying DeePC to nonlinear systems

Standard DeePC theory is derived for noiseless linear, time-invariant (LTI) systems. In the previous exercise we demonstrated that DeePC can be implemented on systems with (some) noise if regularizations on the g decision variable are used (penalties on the magnitude of g are added to the cost function). In this exercise we will show that DeePC can also be used on (some) nonlinear systems if the g decision variable is regularized.

a) Conceptual questions

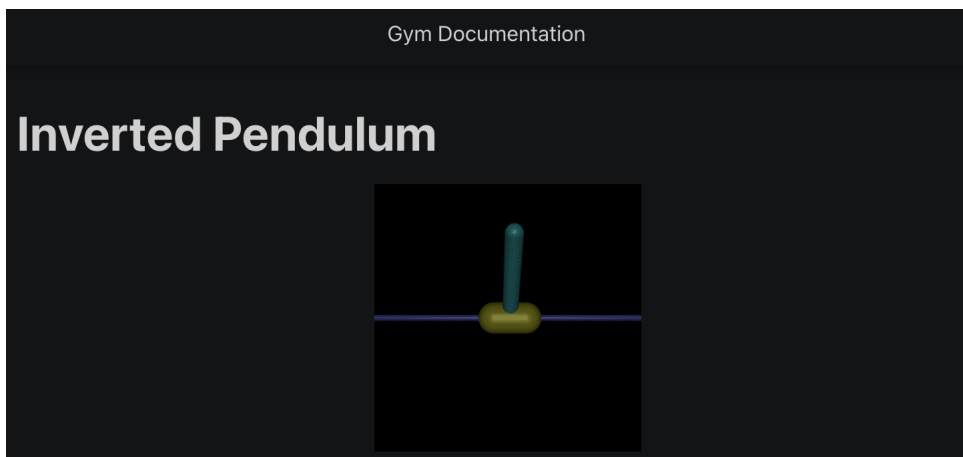   i) Why might we expect regularization to allow DeePC to be applied nonlinear systems?

    ii) When might we expect DeePC to be applicable to nonlinear systems?

b) Simulation

This link provides code for simulating and controlling an inverted pendulum using DeePC.

`https://colab.research.google.com/drive/1OoThlkNjFqXcwg8YXKzRCl6VIC6pj4vU?usp=sharing`

The simulation uses the Mujoco gym environment to simulate the inverted pendulum.

    i) Complete the code and find lambda_g (and/or lam_g_PI) such that the DeePC controller works.

c) Bonus questions

    i) What should you do to keep the system in the linear regime when you are gathering data to build the Hankel matrix used by DeePC?

ii) Can you use data from multiple simulation runs to build a DeePC Hankel matrix?

iii) Do you need σ_y for nonlinear systems?