Problem Set 2

Statistical Contraction Theory for Learning-based Control of Nonlinear Systems

- You are welcome to collaborate, but you must write up your own solutions.
- You can use any programming language, but you must add comments to each section of the codes to clearly explain your intent.
- If you use any resources other than your brain and lecture notes to solve the problems, you must cite them.
- In the following, $\|\cdot\|$ denotes the Euclidean norm for vectors and the induced 2-norm for matrices, i.e., $\|\cdot\| = \|\cdot\|_2$, unless otherwise noted.

1 Conformal Inference

In Problem Set 1, we designed a neural network for approximating the residual dynamics of a spacecraft orbiting around an unknown astronomical body, the complete dynamics of which is given as follows:

$$\dot{æ}_c = \frac{d}{dt}\begin{bmatrix} r \\ v \\ h \\ \Omega \\ i \\ \theta \end{bmatrix} = \begin{bmatrix} v \\ -\frac{\mu}{r^2} + \frac{h^2}{r^3} \\ 0 \\ 0 \\ 0 \\ \frac{h}{r^2} \end{bmatrix} + \begin{bmatrix} 0 \\ \zeta_v(æ_c) \\ \zeta_h(æ_c) \\ \zeta_\Omega(æ_c) \\ \zeta_i(æ_c) \\ \zeta_\theta(æ_c) \end{bmatrix} \tag{1}$$

where $\mu = 1$, $æ_c = [r, v, h, \Omega, i, \theta]^\top$ is the spacecraft state, $\zeta(æ_c) = [\zeta_v(æ_c), \zeta_h(æ_c), \zeta_\Omega(æ_c), \zeta_i(æ_c), \zeta_\theta(æ_c)]^\top$ represents the uncertain part of the dynamics, and the definitions of all the symbols are given in Table 1.

1. The files in `data` contain a calibration dataset consisting of `input_cb.npy` and `output_cb.npy`:

   `input_cb.npy` = `(1000,11240,6) numpy array` representing 1000 input trajectory samples (each with 11240 data points of $æ_c \in \mathbb{R}^{1\times 6}$)

   `output_cb.npy` = `(1000,11240,5) numpy array` representing 1000 output trajectory samples (each with 11240 data points of $\zeta(æ_c)^\top \in \mathbb{R}^{1\times 5}$).

   The trajectories are sampled with the time interval $\Delta t = 0.001$. For your trained neural network $\zeta_{nn}(æ_c(t))$ in Problem Set 1, compute the prediction nonconformity score defined as

   $$S_i = \sup_{t\in[0,T]} \|\zeta_i - \zeta_{nn}(æ_{c,i}(t))\|,\ i = 1, \cdots, 1000 \tag{2}$$

   and plot the results as a histogram (i.e., score vs. frequency). Here, T is the time horizon of each trajectory, i.e., $T = 11240 \times \Delta t = 11.240$, $æ_{c,i}(t)$ denotes the input trajectory data samples in `input_cb.npy`, and $\zeta_i = \zeta(æ_{c,i}(t))$ denotes the output trajectory data samples in `output_cb.npy`. For simplicity, you can assume that $æ_c(t)$ remains constant within each sampling interval of duration Δt. (Hint: If your neural network is well-trained, the scores should be around 0.5; otherwise, there may be an issue with your implementation.)

2. Suppose that the trajectory data samples $\{æ_{c,i}(t)\}_{i=1}^{1000}$ in `input_cb.npy` and unseen real-world trajectories $æ_c(t)$ are all exchangeable with each other. For conformal inference, compute the empirical quantile $Q_{1-\delta}$ given by

   $$Q_{1-\delta} = \lceil (1-\delta)(N+1) \rceil\text{-}th\ smallest\ value\ in\ (S_1, \cdots, S_N) \tag{3}$$

   where $N = 1000$ and $\delta = 0.1$, and overlay it on the score histogram as a vertical line.

3. The files in `data` posted on Canvas contain a validation dataset consisting of `input_emp.npy` and `output_emp.npy`:

$$\begin{aligned}\texttt{input\_emp.npy} &= \texttt{(100,11240,6) numpy array} \text{ representing 100 input trajectory samples}\\ &\quad \text{(each with 11240 data points of } æ_c \in \mathbb{R}^{1\times 6}\text{)}\\ \texttt{output\_emp.npy} &= \texttt{(100,11240,5) numpy array} \text{ representing 100 output trajectory samples}\\ &\quad \text{(each with 11240 data points of } \zeta(æ_c)^\top \in \mathbb{R}^{1\times 5}\text{).}\end{aligned}$$

For the nonconformity scores $S_i,\ i = 1, \cdots, 100$, defined as in (2) **for the samples in the validation datasets**, compute

$$empirical\ failure\ probability = \frac{number\ of\ samples\ satisfying\ S_i > Q_{1-\delta}}{100},\ i = 1, \cdots, 100.$$

Is *empirical failure probability* exactly equal to δ? Briefly justify your answer. (Hint: Our calibration dataset is fixed https://www.stat.berkeley.edu/ ryantibs/statlearn-s23/lectures/conformal.pdf.)

4. The goal of conformal inference is to find $\bar{Q}_{1-\delta}$ that satisfies

$$\mathbb{P}\left[\sup_{t\in[0,T]} \|\zeta(æ_c(t)) - \zeta_{nn}(æ_c(t))\| \le \bar{Q}_{1-\delta}\right] \ge 1 - \delta.$$

Based on the observation in 3, is $\bar{Q}_{1-\delta}$ exactly equal to $Q_{1-\delta}$ of (3)? Briefly justify your answer.

Table 1: Summary of the symbols in (1)

Symbol	Description
$æ_c$	State of spacecraft, $æ_c = [r, v, h, \Omega, i, \theta]^\top$
r	Radial distance from center of target body
v	Radial velocity
h	Specific angular momentum
Ω	Right ascension of ascending node
i	Inclination of orbit
θ	True anomaly
μ	Gravitational constant of target body