Spring 2020

Stanford
AA 203: Optimal and Learning-based Control
Problem set 3, due on May 8

Problem 1: Consider the scalar system

$$\dot{x}(t) = u(t),$$

with the constraint $|u(t)| \leq 1$ for all $t \in [0, T]$. The cost function is

$$J = \Big(x(T)\Big)^2 + \int_0^T \Big(u(t)\Big)^2 dt.$$

Find an optimal control policy for this problem by using the HJB equation.

Hint: Consider the following cost-to-go function candidate:

$$J^*(t, x) = \begin{cases} (x - T + t)^2 + T - t & \textit{if } x > 1 + T - t, \\ (x + T - t)^2 + T - t & \textit{if } x < -(1 + T - t), \\ x^2/(1 + T - t) & \textit{if } |x| \leq 1 + T - t, \end{cases}$$

and then verify that it is indeed a solution to the HJB equation.

Problem 2: Consider the continuous LQR problem:

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t),$$

with cost function:

$$J = \frac{1}{2}\Bigg\{ x^T(t_f) \begin{pmatrix} 0 & 0 \\ 0 & h \end{pmatrix} x(t_f) + \int_0^{t_f} \left[x^T(t) \begin{pmatrix} q & 0 \\ 0 & 0 \end{pmatrix} x(t) + r\, u^2(t) \right] dt \Bigg\},$$

where $q = 1$, $r = 3$, $h = 4$, $t_f = 10$. Write a script to solve the corresponding Riccati differential equation. Plot the time-varying gains and the state solution with initial condition $x(0) = [1, 1]^T$. Repeat the problem with $t_f = 100$ and compare the results with those for the case $t_f = 10$.

Problem 3: Given the plant dynamics,

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= x_1(t) + u(t) + w(t) \\ y(t) &= x_2(t) + v(t) \end{aligned}$$

and cost function

$$J = \mathbb{E}\left\{ \frac{1}{2} \int_0^{t_f} \left(3x_1^2(t) + 3x_2^2(t) + u^2(t)\right) dt \right\}$$

where $w(t) \sim \mathcal{N}(0, 4)$ and $v(t) \sim \mathcal{N}(0, 0.5)$ are Gaussian, white noises and $t_f = 15$:

(a) Numerically integrate the Riccati equations for the LQR and the LQE to find the time-varying regulator and estimator gains. *Hint: use the results for continuous time LQR and LQE in Chapter 3 of the course notes.*

(b) The full stochastic linear optimal output feedback problem involves using $u(t) = -F\hat{x}(t)$. For this control policy, the compensator (i.e., the system mapping sensor measurements into actuator commands) is of the form

$$\begin{aligned} \dot{x}_c(t) &= A_c\, x_c(t) + B_c\, y(t) \\ u(t) &= -C_c\, x_c(t). \end{aligned}$$

Find A_c, B_c, and C_c.

(c) Write down the dynamic equations for the combined plant and compensator system $\begin{bmatrix} x \\ x_c \end{bmatrix}$. Simulate the full closed-loop system in MATLAB using the gains found in Part (a) and the initial conditions $x_0 = \begin{bmatrix} 10 \\ -10 \end{bmatrix}$ and $\hat{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(d) Assume steady-state values. Show that the eigenvalues of the closed-loop dynamics are equal to those of the steady-state regulator and estimator. Compare the performance of the compensator in Part (c) to the performance you obtain when you use these steady state values of the regulator and estimator gains.

Problem 4: Consider a kinematic model of a fixed-wing aircraft flying in the presence of wind, given as follows:

$$\dot{x} = \nu \cos(\theta) + d_x \tag{1}$$
$$\dot{y} = \nu \sin(\theta) + d_y \tag{2}$$
$$\dot{\theta} = u + d_\theta \tag{3}$$

The state (x, y, θ) consists of x-position, y-position and heading θ. The aircraft flies at a constant speed of $\nu = 0.4$ km/s, and controls its turn rate u with the control effort constraint of $|u| \leq 0.5$ rad/s. The effect of wind is modeled by the disturbance $d = (d_x, d_y, d_\theta)$ with $|d_x|, |d_y| \leq 0.05$km/s and $|d_\theta| \leq 0.005$ rad/s.
The initial position and heading of the aircraft is $(x, y, \theta) = (0, 0, 0)$, and the aircraft is flying towards its next waypoint (x_w, y_w, θ_w) where $x_w = 2$ km, $y_w = 2$ km and

$\theta_w = \frac{3\pi}{8}$ rad, with an acceptable tolerance of 0.2 km in distance and $\frac{\pi}{10}$. The aircraft needs to reach the waypoint within tolerance by time $t = 0$.

(a) What is the target set T representing the acceptable set of states for reaching the waypoint within the tolerances? Find a suitable function $I(x, y, \theta)$ so that $I(x, y, \theta) \leq 0$ if and only if $(x, y, \theta) \in T$.

Hint: This may be written as the pointwise maximum of multiple functions

(b) The value function $V(t, x, y, \theta)$ represents the set of states from which the aircraft can reach the target set T within t sections is the solution to the Hamilton-Jacobi-Isaacs (HJI) Partial Differential Equation:

$$\frac{\partial V}{\partial t}(t, x, y, \theta) + \min_{u \in U} \max_{d \in D} \nabla V(t, x, y, \theta)^\top f(x, y, \theta, u, d) = 0 \tag{4}$$

Where $f(x, y, \theta, u, d)$ represents the system dynamics. The sets

$$U := \{u : |u| \leq 0.5\} \tag{5}$$

$$D := \{(d_x, d_y, d_\theta) : |d_x| \leq 0.05, |d_y| \leq 0.05, |d_\theta| \leq 0.005\} \tag{6}$$

represent the bounds on the control and disturbances. First, given a state (x, y, θ) and a control u, find the analytical form of the worst case disturbance for this state and control via

$$d^*(t, x, y, \theta, u) := \arg\max_{d \in D} \nabla V(t, x, y, \theta)^\top f(x, y, \theta, u, d). \tag{7}$$

Then, find the analytical form of the optimal worst-case control u^* via

$$u^* := \arg\min_{u \in U} \left(\max_{d \in D} \nabla V(t, x, y, \theta)^\top f(x, y, \theta, u, d) \right). \tag{8}$$

Hint: u^*, d^* can be written as a function of the derivatives of the value function with respect to the state

(c) Compute $V(t, x, y, \theta)$ from $t = -10$ to $t = 0$ using the helperOC toolbox, which can be found at `https://github.com/HJReachability/helperOC`.

Visualize and provide an explanation of the following functions: $V(t = -10, x, y, \theta), V(t = -5, x, y, \theta), V(t = -10, x, y, \theta = 0), V(t = -10, x, y = 0, \theta)$.

Hint: A good starting point is `tutorial.m` in the helperOC repository. We recommend the following approach:

(i) Read through `tutorial.m` and replace the grid, time vector and target set to match the problem setting. Specifically, determine which shape function can be used to represent the target set found in part (a). A list of candidates can be found in the Level Set Methods toolbox dependency, more precisely in the `BasicShapes` folder.

For your plots, we recommend ranges $x \in [-2, 5]$ with 45 grid points, $y \in [-2, 4]$ with 45 grid points, and $\theta \in [-\pi, \pi]$ with 35 grid points.

(ii) Fill in the problem parameters to match the problem setting and define the corresponding dynamical system. The Dubins car model seen in lecture should be used. Read through the matlab functions corresponding to the dynamics (in the `dynSys` folder). Make sure to provide the disturbance bounds defining D for the Dubins Car model.

(iii) Make sure to include the objective for both the control and the disturbance (the objective for the control has already been specified as `'min'` in the tutorial).

(iv) Use the `visSetIm.m` and `proj.m` functions to visualize the functions specified above

Problem 5: Consider the optimal control problem

$$(\mathbf{OCP})_1 \begin{cases} \min \; h(y(1)) = -y(1) \\ \dot{x}(t) = -x(t)\,u(t) + y(t)\,u^2(t), \;\; \dot{y}(t) = x(t)\,u(t) - 3y(t)\,u^2(t) \\ x(0) = 1, \;\; y(0) = 0 \\ 0 \leq u(t) \leq 1, \;\; t \in [0,1] \end{cases}$$

where $0 \leq x(t) \leq 1$ and $0 \leq y(t) \leq 1$ represent concentrations of chemical substances that react according to the above differential equations, under a temperature control action represented by $0 \leq u(t) \leq 1$. The final time is fixed, namely: $t_f = 1$. The objective consists of maximizing the concentration of the second substance y starting from a maximal concentration of the first substance x.

The effectiveness of direct methods for optimal control problems heavily depends on the rule used to numerically integrate the differential equations (and the cost). Herein, we describe a simple rule for integration known as *trapezoidal rule*, which resembles the classical forward Euler scheme, but is much more efficient. Specifically, consider a dynamical system $\dot{x} = f(x(t), u(t))$ and select two points $a < b$ in $[0, t_f]$. By the fundamental theorem of calculus, one has

$$x(b) = x(a) + \int_a^b f(x(t), u(t)) \; dt.$$

When a and b are "close enough," the previous integral can be approximated by the area of the trapezoid with vertices a, $f(a)$, $f(b)$, and b (see figure 1). One then obtains the approximation

$$x(b) \simeq x(a) + \frac{f(a) + f(b)}{2}(b - a).$$

Using such an approximation (referred to as trapezoidal approximation), and given a time discretization $0 = t_0 < t_1 < \cdots < t_N = t_f$, the differential constraints can then be transcribed into the following set of constraints:

$$x(t_{i+1}) - x(t_i) - (t_{i+1} - t_i)\,\frac{f(t_i) + f(t_{i+1})}{2} = 0, \quad i = 0, \ldots, N-1 \; . \tag{9}$$

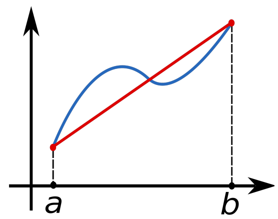

Figure 1: Trapezoidal approximation.

(a) Numerically solve problem $(\mathbf{OCP})_1$ using a direct method leveraging the trapezoidal rule (9). Specifically, implement the dynamics discretization, specify a cost and impose the constraints by filling in the `Matlab` scripts `fDyn.m`, `cost.m`, and `constraint.m` available in the folder `ChemicalReaction`, and then run the script `collocation.m` to obtain a solution. Provide plots for the time evolutions of x, y and u.
Note: We strongly encourage you to read through the provided matlab script (`collocation.m` to understand the inner workings of these methods. In the next Problem Set you will have the opportunity to fully implement a direct method.

(b) What is the optimal quantity of the second substance y at the final time $t_f = 1$?

Learning goals for this problem set:

Problem 1: To learn how to solve optimal control problems by solving the HJB equation.

Problem 2: To gain insights into the "fundamental" LQR problem.

Problem 3: To gain experience with LQG control.

Problem 4: To gain experience with HJI Reachability.

Problem 5: To get an introduction to direct methods leveraging the Trapezoidal integration scheme.