Exercises for Lecture Course on Numerical Optimal Control (NOC)
Albert-Ludwigs-Universität Freiburg – Summer Term 2017

Exercise 7: Dynamic Programming

Prof. Dr. Moritz Diehl, Andrea Zanelli, Dimitris Kouzoupis

In this exercise, we will use dynamic programming (DP) to implement a controller for the inverted pendulum from Exercise 6:

$$\begin{aligned} \dot{\phi} &= \omega \\ \dot{\omega} &= \sin(\phi) + \tau, \end{aligned} \tag{1}$$

where ϕ is the angle describing the orientation of the pendulum, ω is its angular velocity and τ is the input torque. The goal is to design a feedback policy capable of swinging up the pendulum starting from $\phi = \pi$. Moreover, we will prove the Schur Complement Lemma, which can be used to derive the formulation for the LQR controller.

1. **Dynamic programming:** Consider the following optimal control problem:

$$\begin{aligned} \min_{x,\,u} \quad & \sum_{i=0}^{N-1} \left(x_i^T Q x_i + u_i^T R u_i \right) + x_N^T Q_N x_N \\ \text{s.t.} \quad & x_0 = \bar{x}_0 \\ & x_{i+1} = f(x_i, u_i),\ i = 0, ..., N-1, \\ & -10 \leq u_i \leq 10,\ i = 0, ..., N-1, \end{aligned} \tag{2}$$

where f describes the discretized dynamics obtained by applying one step of the explicit RK4 integrator with step-size $h = 0.1$ to (1).

(a) Using the template provided, implement the DP algorithm and use it to compute the cost-to-go associated with (2). Choose $N = 20$, $Q = \text{diag}(100,\, 0.01)$, $R = 0.001$ and Q_N equal to the cost matrix associated with the LQR controller for the dynamics linearized at $x_{\text{lin}} = [0,\, 0]^T$. Discretize the angle ϕ into 200 values between $-\frac{\pi}{2}$ and 2π using the MATLAB command `linspace`. Analogously, discretize the angular velocity into 40 values between -10 and 10 and the torque τ into 20 values between 10 and -10. Plot the obtained cost-to-go and the feedback policy associated with it. *Remark: in order to compute the cost-to-go you will have to project the state obtained by simulating the dynamics forward onto the defined discretization grid. To this end, use in your code the MATLAB function* `project` *provided with this exercise.*

(3 points)

(b) In the figures obtained at the previous point, plot the cost-to-go associated with the LQR controller designed for the linearized system and the corresponding feedback policy. Where is the LQR policy similar to the one obtained with DP? Where is it different? Why?

(2 points)

(c) Following the templates provided, use the LQR controller designed to control the system in closed loop starting from the initial state $\bar{x}_0 = [\pi,\, 0]^T$ and plot the trajectories obtained and the associated closed-loop cost.

(1 point)

(d) Analogously, use the DP policy to control the system starting from the same initial state. Compare the trajectories and the closed-loop cost using the templates provided. Which of the two controllers achieves the best performance? Why?

(2 points)

2. **Schur Complement Lemma:** Consider the following lemma:

Lemma 1 *(Schur Complement Lemma) Let R be a positive-definite matrix. Then, the following holds:*

$$\min_u \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q & S^T \\ S & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = x^T(Q - S^T R^{-1} S)x \tag{3}$$

and the minimizer $u^*(x)$ is given by $u^*(x) = -R^{-1}Sx$.

(a) Prove Lemma 1.

(2 points)

This sheet gives in total 10 points