

MECH 6326 - Optimal Control and Dynamic Programming

Homework 4

Jonas Wagner
jonas.wagner@utdallas.edu

2023, April 13th

Contents

Problem 1: Navigation in an uncertain environment	2
1a)	2
1b)	3
Problem 2: Zero Mean Disturbances	4

Problem 1 Robot navigation in an uncertain environment

PROBLEM:

You are designing a controller to navigate a robot through a cluttered and uncertain environment. We use a very simplified 2D grid model, where the robot starts in the bottom-left corner, and must move to the top-right corner, and can move either up or to the right at each step. Due to unpredictable robot-environment interaction, if the robot moves in the same direction as in the previous time-step, there is a higher probability of moving in the desired direction than if it tries to change direction. If the direction selected is the same as in the previous step, the robot lands in the desired state with probability 0.8 and lands in the undesired state with probability 0.2. If the direction selected is different from in the previous step, the robot lands in the intended state with probability 0.6 and lands in the unintended state with probability 0.4. The first move is treated as though there was a move in the same direction previously.

The environment is a 41×41 grid. Assume that at each time step, the robot knows its location in the grid and can use the information for feedback. The environment has a number of obstacles whose locations are defined in the file *robot_nav.mat* available on eLearning. Hitting an obstacle or the environment boundary results in a crash, and the mission is failed (i.e, you can treat the obstacle locations as absorbing states).

SOLUTION:

A summary of the problem is as follows:

Let the state be the robot position:

$$x_k \in \mathcal{X} = \{1, \dots, 41\}^2 \subset \mathbb{Z}^2$$

Obstacles exist within that result in a crash (along with the boundary), $\mathcal{X}_{obstacles}$, thus a safe region can be denoted as $\mathcal{X}_{safe} = \mathcal{X} \setminus \mathcal{X}_{obstacles}$. The initial state is in the SW corner: $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Additionally, the goal state is in the NE corner: $x_{goal} = \begin{bmatrix} 41 \\ 41 \end{bmatrix}$.

Let the input be the desired heading:

$$u_k \in \mathcal{U} = \{N, E\} = \left\{ \begin{bmatrix} 0 \\ +1 \end{bmatrix}, \begin{bmatrix} +1 \\ 0 \end{bmatrix} \right\}$$

The update equation for the position is as follows:

$$x_{k+1} = f(x_k, u_k) = x_k + w_k(u_{k-1})u_k, \quad w_k(u_k, u_{k-1}) \sim \begin{cases} \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} & u_k = u_{k-1} \\ \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix} & u_k \neq u_{k-1} \end{cases}$$

1a)

PROBLEM:

Determine and plot the optimal policy and value function for the stage cost $\forall_{k=1, \dots, N}$

$$g_k(x) = \begin{cases} 1 & x = x_{goal} \\ -1 & x \in \mathcal{X}_{obstacles} \\ 0 & \text{otherwise} \end{cases}$$

Comment on your result.

SOLUTION:

1b)

Simulate the system to estimate (via either distribution propagation or Monte Carlo) the success rate and plot a sample trajectory of a successful mission.

Code: See attached matlab code.

Linear Quadratic Problems

Problem 2 Non-zero Mean Disturbances

PROBLEM:

Use dynamic programming to derive the optimal cost functions and policies for a linear quadratic problem with non-zero mean disturbances. The dynamics are

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \quad \forall k=0, \dots, N-1$$

where $\mathbf{E}[w_k] = \bar{w}_k$ and $\mathbf{E}[w_k w_k^T] = W_k$, and the cost function is the linear quadratic cost:

$$\sum_{k=0}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k) + x_N^T Q_N x_N$$

Compute the optimal policy and optimal cost function for the problem instance with constant problem data $\forall k=0, \dots, N-1$

$$A_k = \begin{bmatrix} 0.4 & -0.3 & 0 & 0.6 \\ 0.1 & -0.7 & 0.2 & 0 \\ 0.5 & 0.2 & -0.8 & 0.1 \\ 0 & 0.3 & -0.4 & 0.9 \end{bmatrix}, \quad B_k = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.3 \\ 0 & 0.1 \\ 0.2 & 0 \end{bmatrix},$$

$Q_k = \mathbf{I}$, $R_k = \mathbf{I}$, $\mathbf{E}[w_k] = \bar{w}_k = [0.1 \ -0.1 \ 0.3 \ -0.3]^T$, $\mathbf{E}[w_k w_k^T] = W_k = 0.1\mathbf{I}$, and $N = 30$.

Report the optimal policy coefficients at time 0, $k = 0$, and the optimal cost if the initial state is zero, $x_0 = \mathbf{0}$.

SOLUTION:

For the general case, the dynamic program can be defined by the algorithm with stage cost, $g_k(x_k, u_k) = x_k^T Q_k x_k + u_k^T R_k u_k$, and terminal cost, $g_N(x_N) = x_N^T Q_N x_N$.