EEE 446/546 **Summer 2019**

MATLAB Assignment

Problem 1: The Kalman Filter and Matlab

Let a linear system driven by Gaussian noise be given by the following:

$$x_{t+1} = Ax_t + w_t$$

$$y_t = Cx_t + v_t$$

Here $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1 \end{bmatrix}$, and the i.i.d. noise processes satisfy $w_t \sim \mathcal{N}(0, I)$, $v_t \sim \mathcal{N}(0, 1)$. Suppose that $\Sigma_{0|-1} = I$ (here, we use the notation in Chapter 6 of the Lecture notes). Compute the Riccati recursions and by Matlab verify that the recursions converge to a unique fixed point.

Problem 2: Policy Iteration/Value Iteration/Q-Learning and Linear Programming

Consider the following problem: Let $\mathbb{X} = \{1, 2\}, \mathbb{U} = \{1, 2\}$, where $\mathbb{X}$ denotes whether a fading channel is in a good state ($x = 2$) or a bad state ($x = 1$). There exists an encoder who can either try to use the channel ($u = 2$) or not use the channel ($u = 1$). The goal of the encoder is send information across the channel.

Suppose that the encoder's cost (to be minimized) is given by:

$$c(x, u) = -1_{\{x=2,u=2\}} + \alpha(u - 1),$$

for $\alpha = 1/2$ (if you view this as a maximization problem, you can see that the goal is to maximize information transmission efficiency subject to a cost involving an attempt to use the channel; the model can be made more complicated but the idea is that when the channel state is good, $u = 2$ can represent a channel input which contains data to be transmitted and $u = 1$ denotes that the channel is not used).

Suppose that the transition kernel is given by:

$$P(x_{t+1} = 2|x_t = 2, u_t = 2) = 0.8, \qquad P(x_{t+1} = 1|x_t = 2, u_t = 2) = 0.2$$

$$P(x_{t+1} = 2|x_t = 2, u_t = 1) = 0.2, \qquad P(x_{t+1} = 1|x_t = 2, u_t = 1) = 0.8$$

$$P(x_{t+1} = 2|x_t = 1, u_t = 2) = 0.5, \qquad P(x_{t+1} = 1|x_t = 1, u_t = 2) = 0.5$$

$$P(x_{t+1} = 2|x_t = 1, u_t = 1) = 0.9, \qquad P(x_{t+1} = 1|x_t = 1, u_t = 1) = 0.1$$

We will consider either a discounted cost criterion for some $\beta \in (0, 1)$ (you can fix an arbitrary value)

$$\inf_{\Pi} E_x^{\Pi}[\sum_{t=0}^{\infty} \beta^t c(x_t, u_t)] \tag{1}$$

or the average cost criterion

$$\inf_{\Pi} \limsup_{T\to\infty} \frac{1}{T} E_x^{\Pi}[\sum_{t=0}^{T-1} c(x_t,u_t)]. \tag{2}$$

a) Using Matlab or some other program, obtain a solution to the problem given above in (1) through the following:

(i) [10 Points] Policy Iteration

(ii) [10 Points] Value Iteration.

(iii) [15 Points] Q-Learning. Note that a common way to pick α coefficients in the Q-learning algorithm is to take for every x, u pair:

$$\alpha_t(x,u) = \frac{1}{1+\sum_{k=0}^{t} 1_{\{x_k=x,u_k=u\}}}$$

For Policy Iteration, see the discussion in *Chapter 8*, Section 8.1 of the Lecture Notes on the course web site. For Value Iteration, see Theorem 5.4.2. For Q-Learning, see Section 8.2.1 of the Lecture Notes. You may also consult resources on the Internet.

b) [15 Points] Consider the criterion given in (2). Apply the convex analytic method, by solving the corresponding linear program, to find the optimal policy? In Matlab, the command *linprog* can be used to solve linear programming problems.