16-899 Adaptive Control and Reinforcement Learning

Homework 1

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, PA 37235

October 10, 2020

Carnegie Mellon University
The Robotics Institute

1 Problem 1: System

1.1

Yes, the system is control-affine.

1.2

The system is:

Nonlinear

Continuous time

Time invariant

Deterministic

Continuous state

1.3

To linearize the system we need to write it in the form:

$$\dot{x} = Ax + Bu$$

Using Taylor expansion on the nonlinear terms of the dynamics equation:

$$\begin{aligned}\dot{p_x} = v\cos\theta =& v^r\cos\theta^r + \cos\theta^r(v - v^r) - v^r\sin\theta^r(\theta - \theta^r)\\ =& \theta^r v^r \sin\theta^r + (\cos\theta^r)v - (v^r\sin\theta^r)\theta\end{aligned}$$

$$\begin{aligned}\dot{p_y} = v\sin\theta =& v^r\sin\theta^r + \sin\theta^r(v - v^r) + v^r\cos\theta^r(\theta - \theta^r)\\ =& -\theta^r v^r \cos\theta^r + (\sin\theta^r)v + (v^r\cos\theta^r)\theta\end{aligned}$$

Thus, the linearized dynamics can be written as:

$$\begin{bmatrix}\dot{p_x}\\ \dot{p_y}\\ \dot{v}\\ \dot{\theta}\end{bmatrix} = \begin{bmatrix}\theta^r v^r\sin\theta^r\\ -\theta^r v^r\cos\theta^r\\ 0\\ 0\end{bmatrix} + \begin{bmatrix}0 & 0 & \cos\theta^r & -v^r\sin\theta^r\\ 0 & 0 & \sin\theta^r & v^r\cos\theta^r\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}p_x\\ p_y\\ v\\ \theta\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\dot{v}\\ \dot{\theta}\end{bmatrix}$$

1.4

At $x = [p_x, p_y, v, \theta] = [1, 1, 0, 0]$, the above equation reduces to:

$$\begin{bmatrix} \dot{p_x} \\ \dot{p_y} \\ \dot{v} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ v \\ \theta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{\theta} \end{bmatrix}$$

Necessary and sufficient condition for controllability:
Rank of controllability matrix $C = rank[B\ AB\ A^2B\ ...A^{n-1}B] = n$ where n is the size of the state space. In our case, n = 4,

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Rank(C) = rank$[B\ AB\ A^2B\ A^3B] = 3$ hence it is not full rank and the system is uncontrollable.

1.5

$$\begin{aligned} C &= [B\ AB\ A^2B\ A^3B] \\ &= \begin{bmatrix} 0 & 0 & \cos\theta & -v\sin\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & \sin\theta & v\cos\theta & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

To have rank(C) = 4, all the rows of C should be linearly independent. Obviously, row 3 and 4 are independent of each other and of rows 1 and 2. Thus, we only have to see if rows 1 and 2 are independent or not.
Row 1 and row 2 of C will be independent if there does not exist a constant a, $a \neq 0$, such that:

$$\cos\theta = a\sin(\theta) \tag{1}$$

And

$$-v\sin\theta = av\cos(\theta) \tag{2}$$

Case 1: $v \neq 0$

Solving (1) and (2) simultaneously for a gives:

$$\frac{\cos\theta}{\sin\theta} = -\frac{\sin\theta}{\cos\theta}$$
$$\implies \sin^2\theta + \cos^2\theta = 0$$

The above equation holds true only if $\sin\theta = 0$ and $\cos\theta = 0$ which is impossible. Thus, if $v \neq 0$ the system is controllable.

Case 2: $v = 0$

In this case, (2) holds true for all values of a. Thus, a feasible value of a would exist that satisfies (1), that is $a = \frac{\sin\theta}{\cos\theta}$. Thus, in this case, the system is uncontrollable.

2 Problem 2: Continuous Time Optimal Control

2.1

Cost-to-go function:

$$\begin{aligned}
J(x_0, u_{0:T}) &= \int_0^T l(x_t, u_t)dt + l_T(x_T, u_T) \\
&= \int_0^T \left(\frac{1}{2}(x_t - x_G)^\top Q(x_t - x_G) + \frac{1}{2}u_t^\top R u_t\right)dt + \frac{1}{2}(x_T - x_G)^\top S(x_T - x_G)
\end{aligned}$$

Value function:

$$\begin{aligned}
V(t, x_t) &= \min_{u_{t:T}} J(x_t, u_{t:T}) \\
&= \min_{u_{t:T}} \int_{\tau=t}^T \left(\frac{1}{2}(x_\tau - x_G)^\top Q(x_\tau - x_G) + \frac{1}{2}u_\tau^\top R u_\tau\right)d\tau + \frac{1}{2}(x_T - x_G)^\top S(x_T - x_G)
\end{aligned}$$

2.2

Hamilton-Jacobi-Bellman equation:

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}V(t, x_t) + \min_{u_t}\left\{l(x_t, u_t) + \frac{\partial}{\partial x}V(t, x_t)f(x_t, u_t)\right\} \\
\implies 0 &= \frac{\partial}{\partial t}V(t, x_t) + \min_{u_t}\left\{\left(\frac{1}{2}(x_t - x_G)^\top Q(x_t - x_G) + \frac{1}{2}u_t^\top R u_t\right) + \frac{\partial}{\partial x}V(t, x_t)f(x_t, u_t)\right\}
\end{aligned}$$

where $\dot{x} = f(x_t, u_t)$ is the system dynamics.

2.3

Hamiltonian and co-state:

$$\begin{aligned}
H(t, x_t, u_t, \lambda_t) &= l(x_t, u_t) + \lambda_t^\top f(x_t, u_t) \\
&= \frac{1}{2}(x_t - x_G)^\top Q(x_t - x_G) + \frac{1}{2}u_t^\top R u_t + \lambda_t^\top f(x_t, u_t) \\
\lambda_t &= \frac{\partial}{\partial x}V(t, x_t)
\end{aligned}$$

2.4

Dynamics of co-state:

$$
\begin{aligned}
-\dot{\lambda}_t^* &= \frac{\partial}{\partial x} l(x_t^*, u_t^*) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_t^* \\
&= Q(x_t^* - x_G) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_t^* \\
\text{where, } u_t^* &= arg \min_u H(t, x_t^*, u, \lambda_t^*)
\end{aligned}
$$

Boundary condition:

$$
\lambda_T^* = \frac{\partial}{\partial x} l(x_T^*) = S(x_T - x_G)
$$

2.5

For quadratic cost function as given in the problem, the optimal control law:

$$
u_t^* = arg \min_u H(t, x_t^*, u, \lambda_t^*)
$$

Taking derivative of RHS and setting to zero:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial u}\left(\frac{1}{2}(x_t - x_G)^\top Q(x_t - x_G) + \frac{1}{2} u_t^\top R u_t + \lambda_t^\top f(x_t, u_t)\right) \\
&= R_t u + \left[\frac{\partial}{\partial u} f(x_t, u_t)\right]^\top \lambda_t \\
&= R_t u + B_t^\top \lambda_t \\
\implies u_t^* &= -R_t^{-1} B_t^\top \lambda_t
\end{aligned}
$$

2.6

For infinite horizon, all dependence of V (value function) on t drops out (value function converges), and the HJB equation reduces to:

$$
\begin{aligned}
0 &= \min_{u_t} \left\{ l(x_t, u_t) + \frac{\partial}{\partial x} V(x_t) f(x_t, u_t) \right\} \\
&= \min_{u_t} \left\{ \left(\frac{1}{2}(x_t - x_G)^\top Q(x_t - x_G) + \frac{1}{2} u_t^\top R u_t\right) + \frac{\partial}{\partial x} V(x_t) f(x_t, u_t) \right\}
\end{aligned}
$$

2.7

Optimal value function cannot depend on time in the infinite horizon case, therefore, $V(t, x_t) = V(x_t)$. Thus, we can write the equation for the co-state as:

$$\lambda_t = \frac{\partial}{\partial x} V(t, x_t) = \frac{\partial}{\partial x} V(x_t)$$

$$-\dot{\lambda}_t^* = \frac{\partial}{\partial x} l(x_t^*, u_t^*) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_t^*$$

2.8

Optimal control law:

$$u_t = -R_t^{-1} B_t^\top \lambda_t$$

2.9

The co-state also does not explicitly depend on time.

$$\lambda_t = P(x_t - x_G)$$

where P is constant over time. P can be derived using the algebraic Ricatti equation:

$$PA_t + A_t^\top P - PB_t R_t^{-1} B_t^\top P + Q_t = 0$$

We can make our substitution by assuming a form of λ

$$u_t = -R_t^{-1} B_t^\top P(x_t - x_G)$$

2.10

For the first point, there will be no resulting optimal control law because LQR will not be able to find a stabilizing controller for the given linearization point.

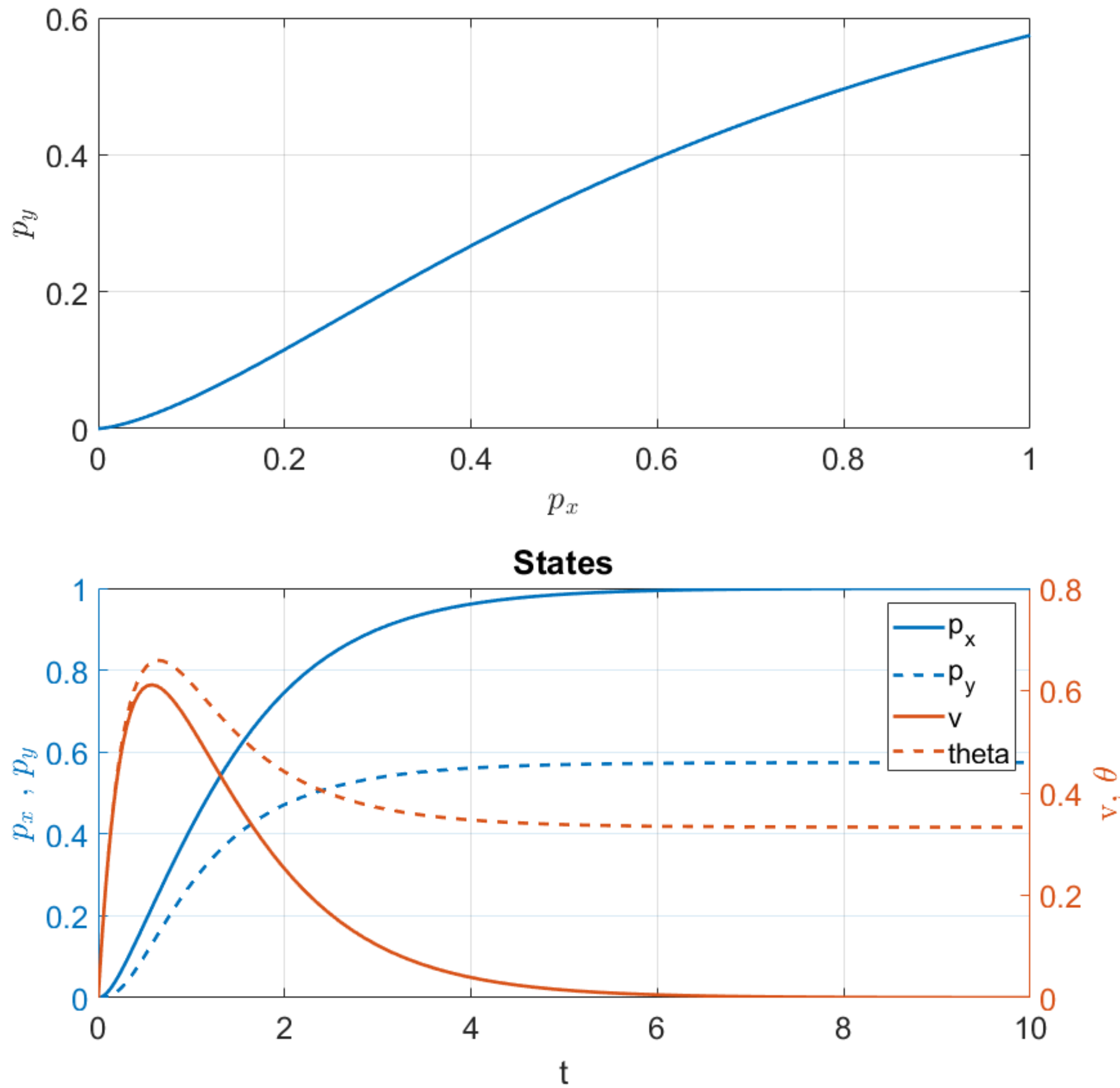


Figure 1: Infinite horizon LQR

3 Problem 3: Discrete Time Optimal Control

3.1 Discrete-time dynamics

$$
\begin{aligned}
\dot{x} = \frac{x_{t+\Delta t} - x_t}{\Delta t} &= \begin{bmatrix} v\cos\theta \\ v\sin\theta \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{\theta} \end{bmatrix} \\
\implies x_{t+1} &= \begin{bmatrix} p_x \\ p_y \\ v \\ \theta \end{bmatrix} + \begin{bmatrix} v\Delta t\cos\theta \\ v\Delta t\sin\theta \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{\theta} \end{bmatrix} \\
&\approx \begin{bmatrix} \theta^r v^r \sin\theta^r \Delta t \\ -\theta^r v^r \cos\theta^r \Delta t \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \Delta t\cos\theta^r & -v^r\Delta t\sin\theta^r \\ 0 & 1 & \Delta t\sin\theta^r & v^r\Delta t\cos\theta^r \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{x_t} \\ p_{y_t} \\ v_t \\ \theta_t \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} \dot{v}_t \\ \dot{\theta}_t \end{bmatrix}
\end{aligned}
$$

3.2 Cost-to-go function and value function

$$
\begin{aligned}
J(x_0, u_{0:N-1}) &= \sum_{t=0}^{N-1} l(x_t, u_t) + l_N(x_N, u_N) \\
&= \sum_{t=0}^{N-1} \left(\frac{1}{2}(x_t - x_G)^\top Q (x_t - x_G) + \frac{1}{2} u_t^\top R u_t \right) + \frac{1}{2}(x_N - x_G)^\top S (x_N - x_G)
\end{aligned}
$$

Value function:

$$
\begin{aligned}
V(x_t) &= \min_{u_{t:N-1}} J(x_t, u_{t:N-1}) \\
&= \min_{u_{t:N-1}} \sum_{\tau=t}^{N-1} \left(\frac{1}{2}(x_\tau - x_G)^\top Q (x_\tau - x_G) + \frac{1}{2} u_\tau^\top R u_\tau \right) + \frac{1}{2}(x_N - x_G)^\top S (x_N - x_G)
\end{aligned}
$$

Bellman equation:

$$
V(x_t) = \min_{u_t} \left(l(x_t, u_t) + V(x_{t+1}) \right)
$$

3.3

Hamiltonian:

$$
\begin{aligned}
H_t(x_t, u_t, \lambda_t) &= l(x_t, u_t) + \lambda_{t+1}^\top f(x_t, u_t) \\
&= \frac{1}{2}(x_t - x_G)^\top Q (x_t - x_G) + \frac{1}{2} u_t^\top R u_t + \lambda_{t+1}^\top f(x_t, u_t)
\end{aligned}
$$

where, $x_{t+1} = f(x_t, u_t)$ is the discrete-time dynamics of the system.

$$\lambda_t = \frac{\partial}{\partial x} H(x_t, u_t, \lambda_{t+1})$$

Dynamics of the co-state:

$$\begin{aligned}
\lambda_t^* = \frac{\partial}{\partial x} H(x_t^*, u_t^*, \lambda_{t+1}^*) &= \frac{\partial}{\partial x} l(x_t^*, u_t^*) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_{t+1}^* \\
&= Q_t(x_t^* - x_G) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_{t+1}^*
\end{aligned}$$

Boundary conditions:

We know that:

$$J_N(x_N^*) = l_N(x_N^*) = \frac{1}{2}(x_N^* - x_G)^\top S(x_N^* - x_G)$$

Thus, the boundary condition for the co-state:

$$\lambda_N^*(x_N) = \frac{\partial}{\partial x} J_N(x_N^*) = S(x_N^* - x_G)$$

Now, we can guess:

$$\begin{aligned}
&\lambda_N^{*\top}(x_N^* - x_G) + \gamma_N^* = J_N^*(x_N^*) \\
\implies &(x_N^* - x_G)^\top S(x_N^* - x_G) + \gamma_N^* = \frac{1}{2}(x_N^* - x_G)^\top S(x_N^* - x_G) \\
\implies &\gamma_N^* = -\frac{1}{2}(x_N^* - x_G)^\top S(x_N^* - x_G)
\end{aligned}$$

3.4 Optimal control law

$$u_t^* = arg \min_u H(x_t^*, u_t, \lambda_{t+1}^*)$$

To find u_t^* we can take the derivative of $H(x_t^*, u_t^*, \lambda_{t+1}^*)$ with respect to u and set it to zero:

$$R_t u_t^* + \left[\frac{\partial}{\partial u} f(x_t^*, u_t^*)\right]^\top \lambda_{t+1}^* = 0$$

Thus, solution of the above equation will give us the optimal control.

Also, in our case:

$$\frac{\partial}{\partial u} f(x_t^*, u_t^*) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix} = B$$

We can substitute it back in the equation for control law to get:

$$
\begin{aligned}
R_t u_t + \left[\frac{\partial}{\partial u} f(x_t^*, u_t^*)\right]^\top \lambda_{t+1}^* &= 0 \\
\implies R_t u_t^* + B^\top \lambda_{t+1}^* &= 0 \\
\implies u_t^* &= -R_t^{-1} B^\top \lambda_{t+1}^*
\end{aligned}
$$

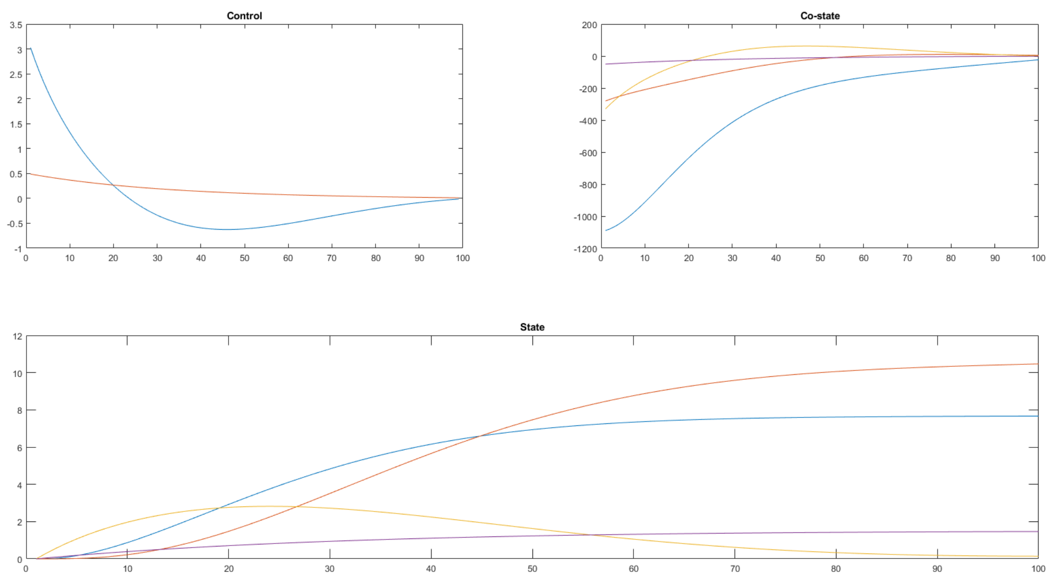
3.5

Dynamics of state and co-state:

$$
\begin{aligned}
\lambda_t^* &= Q_t(x_t^* - x_G) + \left[\frac{\partial}{\partial x} f(x_t^*, u_t^*)\right]^\top \lambda_{t+1}^* \\
x_{t+1} &= f(x_t, u_t) \\
u_t^* &= -R_t^{-1} B^\top \lambda_{t+1}^*
\end{aligned}
$$

Boundary conditions:

$$
\begin{aligned}
x_0 &= [0, 0, 0, 0]^\top \\
\lambda_N^* &= S(x_N^* - x_G)
\end{aligned}
$$



3.6

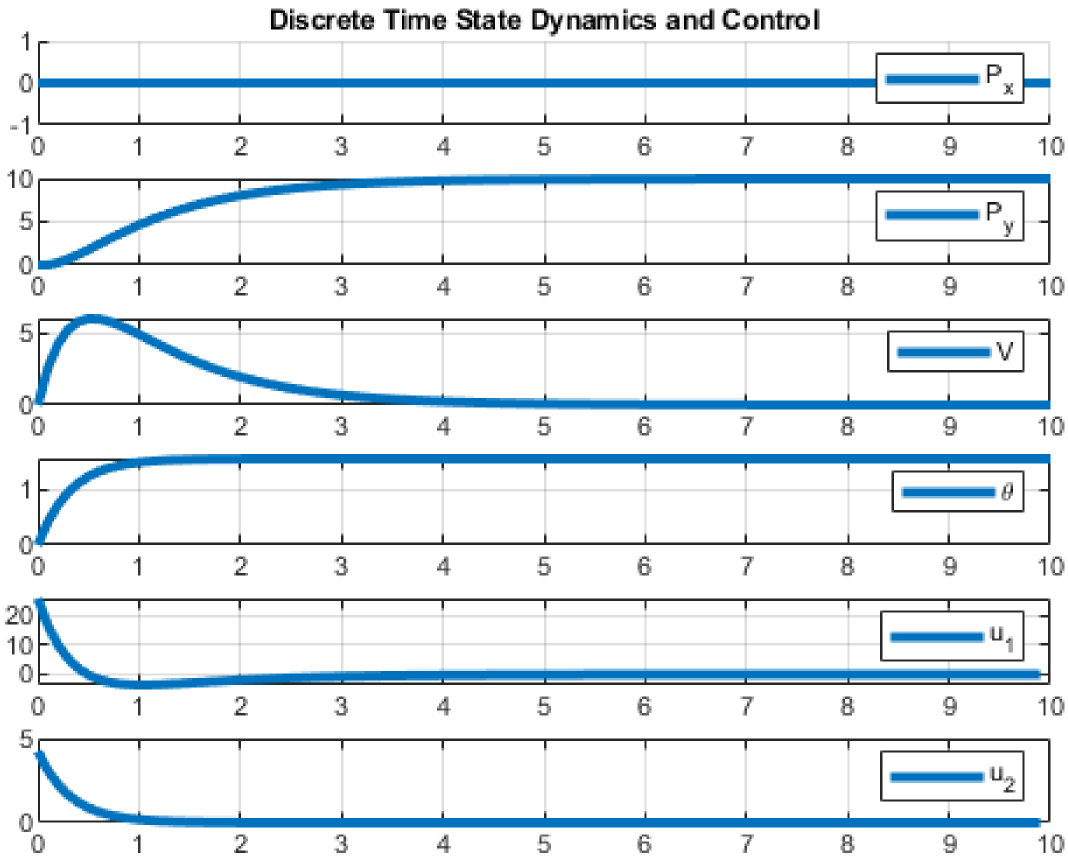


Figure 2: Discrete time LQR with linearized dynamics

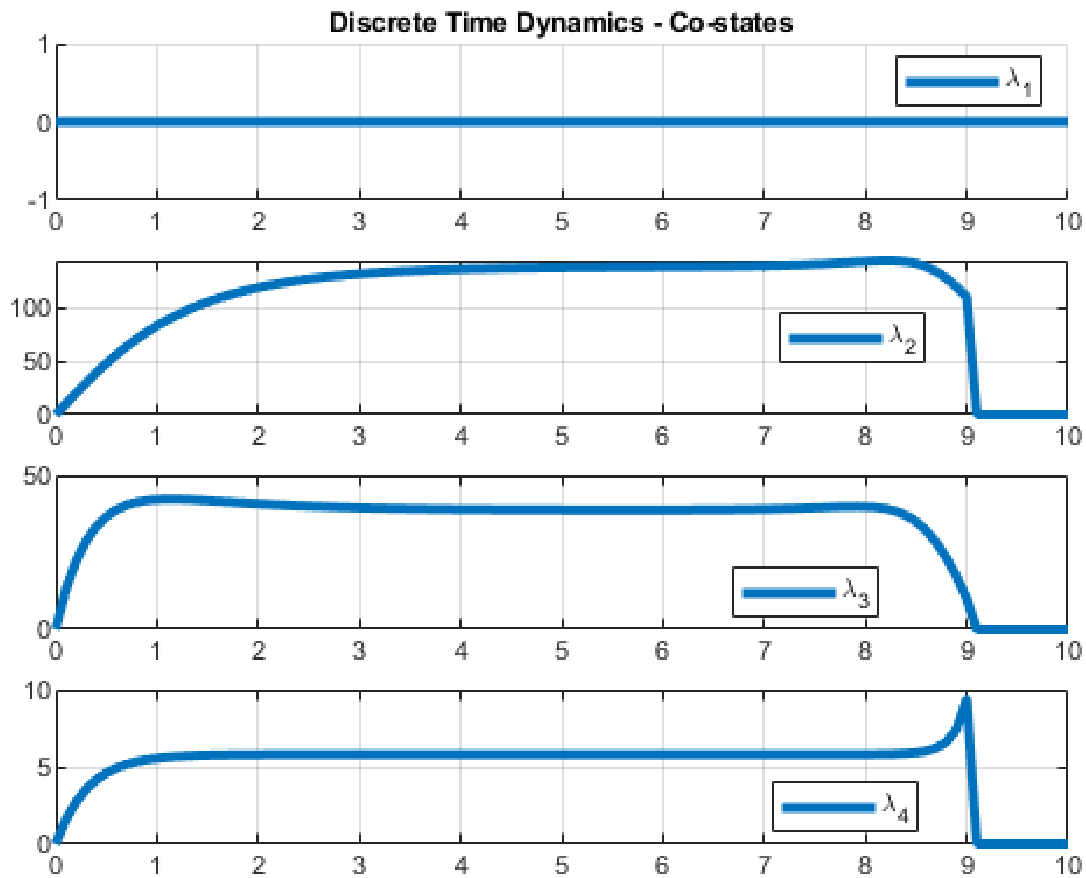


Figure 3: Discrete time LQR with linearized dynamics

Please note, you could have graphed you co-states a little differently here and it would still have been fine. The results here are shown as $\lambda = Px$, where you could have done $\lambda = P(x - x_g)$. This last approach would have yielded something along the lines of:

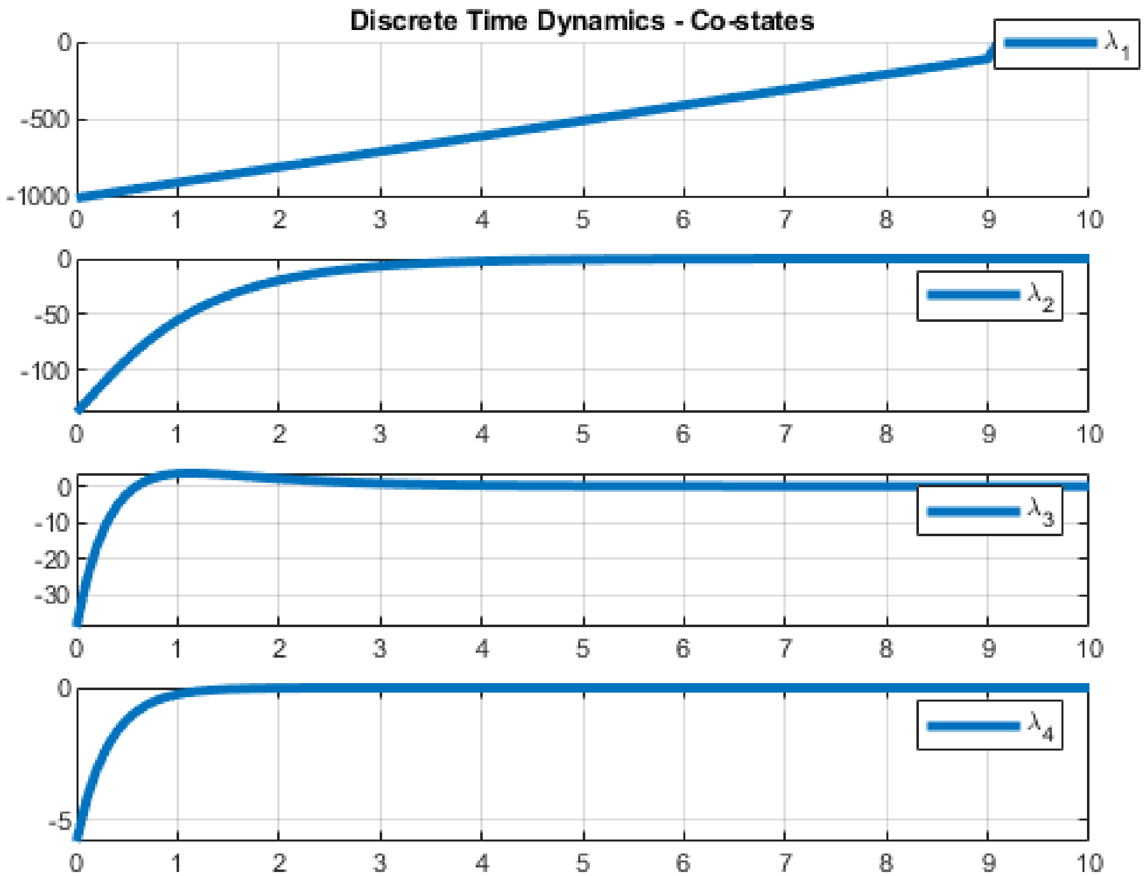


Figure 4: Discrete time LQR with linearized dynamics

3.7

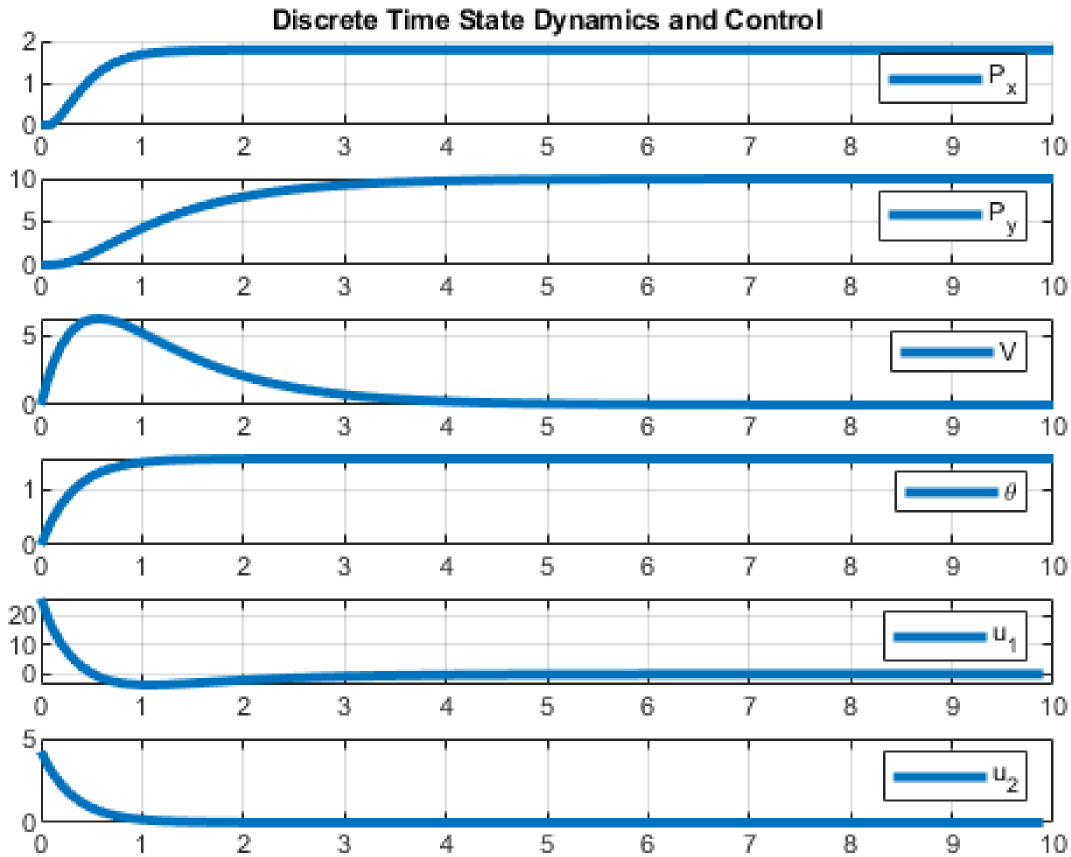


Figure 5: Discrete time LQR with non-linear dynamics

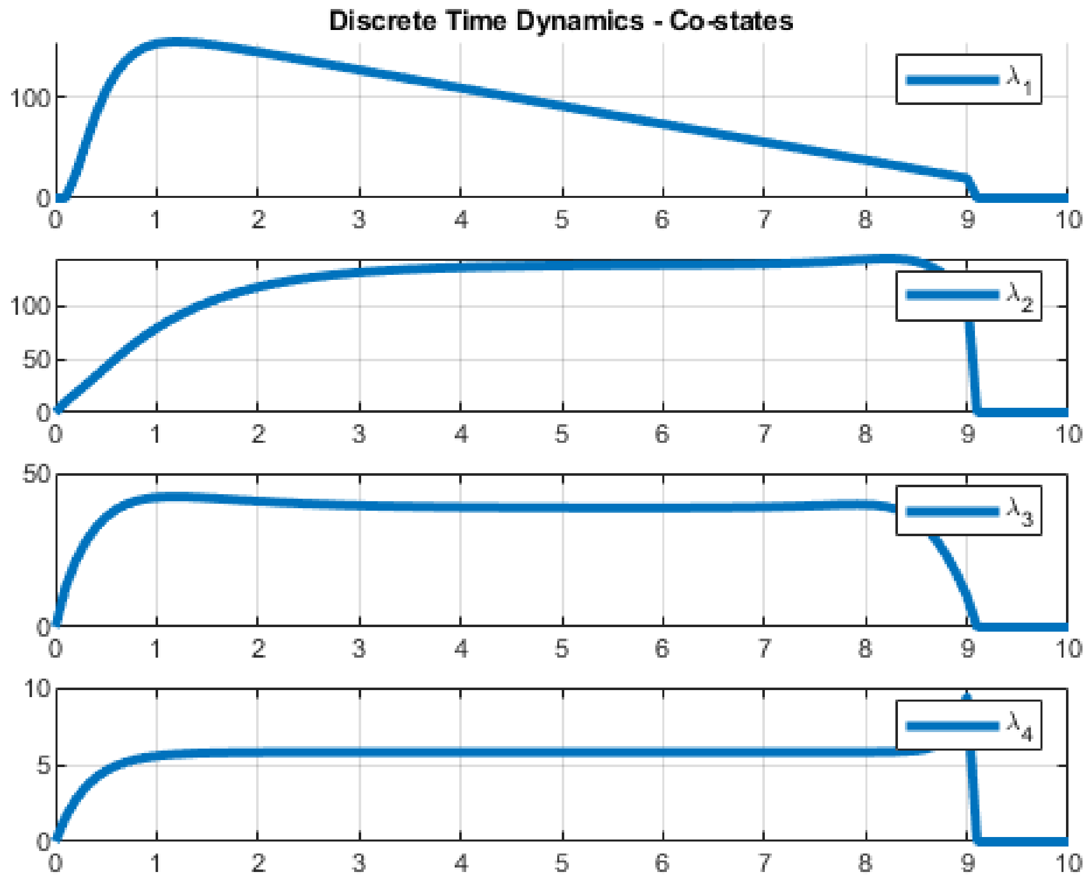


Figure 6: Discrete time LQR with non-linear dynamics

Please note, you could have graphed you co-states a little differently here and it would still have been fine. The results here are shown as $\lambda = Px$, where you could have done $\lambda = P(x - x_g)$. This last approach would have yielded something along the lines of:

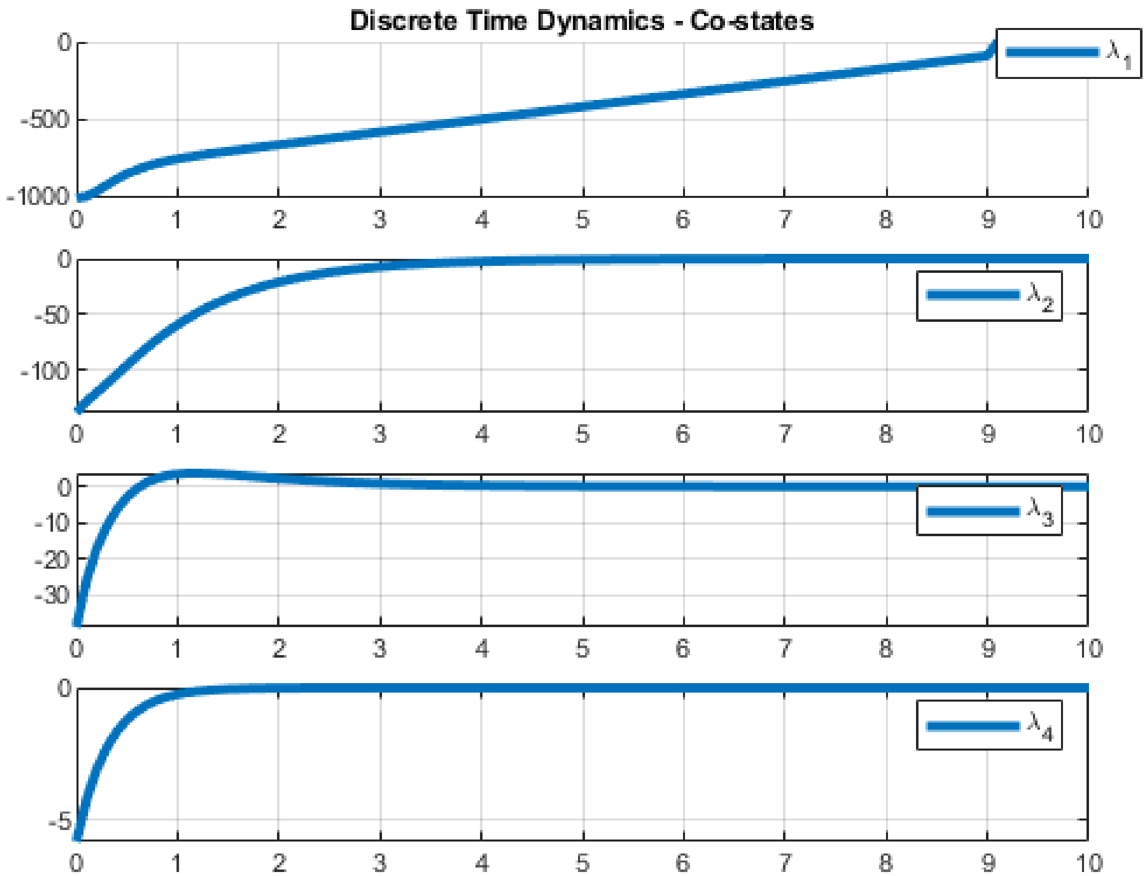


Figure 7: Discrete time LQR with linearized dynamics

4 Question 4

4.1

Hamiltonian is $H(x_k, u_k, \lambda_{k+1}) = x_k^\top Q x_k + 2x_k^\top N u_k + u_k^\top R u_k + \lambda_{k+1}^\top (Ax_k + Bu_k)$.

4.2

Assume $\lambda_{k+1} = 2Px_{k+1}$. Deriving optimal control

$$0 = \frac{\partial H}{\partial u} = 2N^\top x_k + 2Ru_k + B^\top \lambda_{k+1} = 2N^\top x_k + 2Ru_k + 2B^\top P(Ax_k + Bu_k) \tag{3}$$

By rearrangement, we have the control law as:

$$u_k = -[R + B^\top PB]^{-1}[B^\top PA + N^\top]x_k \tag{4}$$

Define $K = [R + B^\top PB]^{-1}[B^\top PA + N^\top]$. The control law can be simply written as $u_k = Kx_k$. **Note: You could also have done the substitution with $\lambda = Px$ as long as you carried it all the way through.**

4.3

Then we can compute $\lambda_k = 2Px_k$ as

$$\lambda_k = \frac{\partial H}{\partial x} = 2Qx_k + 2Nu_k + A^\top \lambda_{k+1} = 2Qx_k - 2NKx_k + 2A^\top P[A - BK]x_k \tag{5}$$

Hence,

$$\begin{aligned} P &= Q - NK + A^\top P[A - BK] && (6) \\ &= Q + A^\top PA - [N + A^\top PB]K && (7) \end{aligned}$$

So the Riccati equation is

$$P = Q + A^\top PA - [N + A^\top PB][R + B^\top PB]^{-1}[B^\top PA + N^\top] \tag{8}$$

4.4

We need to make sure $x_k^\top Q x_k + 2x_k^\top N u_k + u_k^\top R u_k$ is a convex function on x_k and u_k. By rearrangement, the run time cost becomes

$$x_k^\top Q x_k + (u_k + R^{-1}N^\top x_k)^\top R(u_k + R^{-1}N^\top x_k) - x_k^\top N R^{-1} N^\top x_k \tag{9}$$

$$k = \begin{bmatrix} 0.4408 & 1.0183 \end{bmatrix} \tag{10}$$

$$P = \begin{bmatrix} 2.3103 & 1.0136 \\ 1.0136 & 2.5558 \end{bmatrix} \tag{11}$$

Then N should satisfy that $Q - NR^{-1}N^\top \succ 0$. At minimum, discuss that N is a mixing term that couples the cost/interaction between state and control.