6.7920 Pset 2

Olivia Dias

September 24, 2025

1 Traveling Repairman

1.1

We define our DP algorithm with a state space, action space, transition function, cost function, and recursive step.
We define the **state space** as

$s \in \mathbb{S} := (i, j, k)$ for $1 \leq i \leq k \leq j \leq n$

where i is the last site serviced to the left of s, j is the last site serviced to the right of s, and k is the last serviced site.
We define the **action space** as

$a \in \mathbb{A} := ((i-1), (j+1))$ where the repairman takes action $i-1$ if $1 < i$ and $k = j$, otherwise the repairman takes action $j+1$ if $j < n$ and $k = i$
We define the **transition function** as

$$f(s'|s,a) = \begin{cases} (i-1, j, i-1) & \text{if } 2 \leq i \text{ and } k = j \\ (i, j+1, j+1) & \text{if } j \leq n-1 \text{ and } k = i \end{cases}$$

You transition to the next states based on if $1 < i$, $j < n$, $k = i$, and $k = j$. The base case is starting at position at s and you can either move $s-1$ or $s+1$.
We define the **cost function** as

$C(s) = t_{ij} + \sum_{b=1, b \notin [i,j]}^{n} c_b$

You incur a travel cost for travel to repair site i to j or vice versa and a waiting cost for all sites i that have not been serviced.
We define the **recursive solution** to the DP as

$V_t(s_t)^* = min_{a \in \mathbb{A}}[C(s,a) + V_{t+1}^*(s,a)]$

Thus, we have the DP algorithm for finding the minimum-cost service schedule.

2 Deterministic LQR

2.1

In order to show that V is convex in the variables of $x_0, x_1, ..., x_T, u_0, u_1, ..., u_{T-1}$ we will find the second derviative of V and show that it's greater than zero. We

are told that $V := \frac{1}{2}\sum_{t=0}^{T-1}(x_t^T Q x_t + u_t^T R u_t) + \frac{1}{2}x_T^T Q_T x_T$
First we will take the second order partial derivative with respect to x_t:

$\frac{\delta^2 V}{\delta^2 x_t} = Q$

Now we will take the second order partial derivative with respect to u_t:

$\frac{\delta^2 V}{\delta^2 u_t} = R$

Finally we will take the second order partial derivative with respect to x_T:

$\frac{\delta^2 V}{\delta^2 x_T} = Q_T$

We have the second order partial derivatives for V and now need to show that the second order partial derivatives are > 0.

We are told to assume that $Q \succ 0$, $R \succ 0$, $Q_T \succ 0$. Since all the second order partial derivatives of V are $\succ 0$, then based on lecture then $\Delta^2 V \succ 0$. therefore we can say that V is strictly convex.

2.2

We are told to take the gradient of L and then set it to 0. L is defined as

$L := V + \sum_{t=0}^{T-1}\lambda_t^T(Ax_t + Bu_t - x_{t-1})$

$= \frac{1}{2}\sum_{t=0}^{T-1}(x_t^T Q x_t + u_t^T R u_t) + \frac{1}{2}x_T^T Q_T x_t + \sum_{t=0}^{T-1}\lambda_t^T(Ax_t + Bu_t - x_{t+1})$

And we are told that $x_{t+1} = Ax_t + Bu_t, \forall t = 0, ..., T-1$

Now we will take the partial derivative of L with respect to x_t and set it equal to zero.

$\frac{\delta L}{\delta x_t} = Qx_t + A^T\lambda_t - \lambda_{t-1}$

$0 = Qx_t + A^T\lambda_t - \lambda_{t-1}$

Now we will take the partial derivative of L with respect to u_t and set it equal to zero.

$\frac{\delta L}{\delta u_t} = Ru_t + B^T\lambda_t$

$0 = Ru_t + B^T\lambda_t$

Now we will take the partial derivative of L with respect to x_T and set it equal to zero.

$\frac{\delta L}{\delta x_T} = Q_T x_T - \lambda_{T-1}$

$0 = Q_T x_T - \lambda_{T-1}$

And we define $\overline{A}, z$, and $\bar{b}$ to be

$$\overline{A} = \begin{bmatrix}
Q & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & A & 0 & 0 & 0 & \cdots & 0 \\
0 & Q & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & -1 & A & 0 & 0 & \cdots & 0 \\
0 & 0 & Q & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & A & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \cdots & \cdots & \vdots & \cdots & \cdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & -1 & A \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & R & 0 & \cdots & 0 & B & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \vdots & 0 & \ddots & \ddots & \vdots & 0 & \ddots & \cdots & \cdots & \cdots & 0 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & R & 0 & \cdots & \cdots & \cdots & 0 & B \\
0 & \cdots & \cdots & \cdots & \cdots & \mathrm{Q}_T & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & -1 & 0 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \vdots & 0 & \ddots & \ddots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & 0 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \mathrm{Q}_T & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 & -1 & 0
\end{bmatrix}$$

$$z = \begin{bmatrix} x_0 \\ \vdots \\ x_T \\ u_0 \\ \vdots \\ u_{T-1} \\ \lambda_0 \\ \vdots \\ \lambda_T \end{bmatrix}$$

for $t = 0, ..., T-1$

$$\bar{b} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

2.3

From 2.2 we know that we have the following equations

$0 = Qx_t + A^T\lambda_t - \lambda_{t-1}$

$0 = Ru_t + B^T\lambda_t$

$0 = Qx_T - \lambda_{T-1}$

To determine the DP recursive formula we will solve for the different variables as we get:

$\lambda_{t-1} = Qx_t + A^T\lambda_t$

We now have an equation for λ_{t-1}
$-B^T\lambda_t = Ru_t$
$u_t = -R^{-1}B^T\lambda_t$
We now have an equation for u_t
$\lambda_{T-1} = Qx_T$
We now have an equation for λ_{T-1}

For the DP recursive problem we know that it takes the form
$min_{u_0,\ldots,u_{T-1}} V := \frac{1}{2}\sum_{t=0}^{T-1}(x_t^TQx_t + u_t^TRu_t) + \frac{1}{2}x_T^TQ_Tx_T$
and so the cost to go function at the terminal time step T is
$V_T(x_T) = \frac{1}{2}x_T^TQ_Tx_T$
We know that x_{t+1} is constrained to be equal to $Ax_t + Bu_t$ and see a resemblance of this x_{t+1} constraint in the langrangian multiplier. So the langrangian multiplier sits on the next state constraint. And from N3 lecture notes we know that the solution to the LQ problem takes the form: $V_t^*(x_t) = x^TP_tx$ for every $t \leq T$ and for some n x n positive semidefinite matrices P_t. And at terminal stage T, we have that $P_T = Q_T$, so we will use all this information to get deduce the following
$V_{t+1}(x_{t+1}) = \frac{1}{2}x_{t+1}^TP_{t+1}x_{t+1}$
$\nabla_{x_{t+1}}V_{t+1}(x_{t+1}) = P_{t+1}x_{t+1}$
Since langrangian measures the marginal value of changing x_{t+1}, i.e. its applied to the $(Ax_t + Bu_t - x_{t+1})$ and $x_{t+1} = Ax_t + Bu_t$, then we can say that $\lambda_t = \nabla_{x_{t+1}}V_{t+1}$, thus
$\lambda_t = \nabla_{x_{t+1}}V_{t+1}(x_{t+1}) = P_{t+1}x_{t+1}$
$\lambda_t = P_{t+1}x_{t+1}$
The λ_{T-1} equation we found, $\lambda_{T-1} = Q_Tx_T$, matches the terminal condition for $V_T(x_T) = Q_Tx_T = P_Tx_T$.

So we can plug our new value for λ_t, $P_{t+1}x_{t+1}$ into our equation for u_t and we get
$u_t = -R^{-1}B^T\lambda_t$
$u_t = -R^{-1}B^T(P_{t+1}x_{t+1})$
We know that $x_{t+1} = Ax_t + Bu_t$ so we now have
$u_t = -R^{-1}B^TP_{t+1}(Ax_t + Bu_t)$
And we will move all the u_t to one side of the equation and we get
$u_t = -R^{-1}B^TP_{t+1}Ax_t - R^{-1}B^TP_{t+1}Bu_t$
$u_t + R^{-1}B^TP_{t+1}Bu_t = -R^{-1}B^TP_{t+1}Ax_t$
We can multiply both sides by R to get rid of the R^{-1} and then solve for u_t and get
$Ru_t + B^TP_{t+1}Bu_t = -B^TP_{t+1}Ax_t$
$u_t(R + B^TP_{t+1}B) = -B^TP_{t+1}Ax_t$
$u_t = -(R + B^TP_{t+1}B)^{-1}B^TP_{t+1}Ax_t$
We now have a part of the equation that resembles K_t, which is the optimal gain at a specific time step t
$K_t = (R + B^TP_{t+1}B)^{-1}B^TP_{t+1}A$
So we now have

$u_t = -K_t x_t$

And this is the optimal policy from N3 page 6 notes.

Now we will derive the Ricatti equation using λ_{t-1} equation and our definition of λ_t and x_{t+1}.

$\lambda_{t-1} = Qx_t + A^T \lambda_t$

$\lambda_{t-1} = Qx_t + A^T (P_{t+1} x_{t+1})$

$\lambda_{t-1} = Qx_t + A^T P_{t+1}(Ax_t + Bu_t)$

We plug in the value we found for u_t and simplify the equation, so we get

$\lambda_{t-1} = Qx_t + A^T P_{t+1}(Ax_t + B(-K_t x_t))$

$\lambda_{t-1} = Qx_t + A^T P_{t+1}(Ax_t - BK_t x_t)$

$\lambda_{t-1} = Qx_t + A^T P_{t+1}((A - BK_t)x_t)$

$\lambda_{t-1} = Qx_t + (A^T P_{t+1} A - A^T P_{t+1} BK_t)x_t$

$\lambda_{t-1} = (Q + A^T P_{t+1} A - A^T P_{t+1} BK_t)x_t$

And if we expand out our K_t term we get

$\lambda_{t-1} = (Q + A^T P_{t+1} A - A^T P_{t+1} B(R + BP_{t+1} B)^{-1} BP_{t+1} A)x_t$

Finally we replace the λ_{t-1} term with $P_t x_t$ and cancel out the x_t term, which equals

$P_t x_t = (Q + A^T P_{t+1} A - A^T P_{t+1} B(R + BP_{t+1} B)^{-1} BP_{t+1} A)x_t$

$P_t = (Q + A^T P_{t+1} A - A^T P_{t+1} B(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A)$

Now, we have the recursive Ricatti equation and the terminal P_T equation.

Thus, we have found the DP recursion formula in closed form is

$u_t = -K_t x_t$

$K_t = (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$

$P_t = (Q + A^T P_{t+1} A - A^T P_{t+1} B(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A)$

$P_T = Q_T$

2.4

From 2.3 we know that the DP recursion formula in closed form is

$u_t = -K_t x_t$

$K_t = (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$

$P_t = (Q + A^T P_{t+1} A - A^T P_{t+1} B(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A)$

$P_T = Q_T$

And this solution as $T \to \infty$ becomes

$P = (Q + A^T PA - A^T PB(R + B^T PB)^{-1} B^T PA)$

$u_t = -Kx_t$

$K = (R + B^T PB)^{-1} B^T PA$

This solution is the same equations stated in 2.4. This is because our closed form from 2.3 converges to a unique solution $P \succ 0$ from a fixed Q, Q_T, R as $T \to \infty$.

3 Stability of LQR

3.1

We are told the following:
$D = A + BL$
$L = -(R + B^T PB)^{-1} B^T PA$
$P = A^T (P - PB(R + B^T PB)^{-1} B^T P)A + Q.$
And we want to show that $P = D^T PD + Q + L^T RL$

So we plug in the value of D into the $P = D^T PD + Q + L^T RL$ and get
$P = (A + BL)^T P(A + BL) + Q + L^T RL$
$P = (A^T + L^T B^T)P(A + BL) + Q + L^T RL$
$P = (A^T P + L^T B^T P)(A + BL) + Q + L^T RL$
$P = A^T PA + A^T PBL + L^T B^T PA + L^T B^T PBL + Q + L^T RL$

Now we plug in $L = -(R+B^T PB)^{-1} B^T PA$ for the $P^T DP$ part of the equation
$P = A^T PA+A^T PB(-(R+B^T PB)^{-1} B^T PA)+(-(R+B^T PB)^{-1} B^T PA)^T B^T PA+$
$(-(R + B^T PB)^{-1} B^T PA)^T B^T PB(-(R + B^T PB)^{-1} B^T PA) + Q + L^T RL$
To make it easier to simplify the equation we will define a new matrix $W :=$
$R + B^T PB$ and the equation becomes

$P = A^T PA+A^T PB(-W^{-1} B^T PA)+(-W^{-1} B^T PA)^T B^T PA+(-W^{-1} B^T PA)^T B^T PB(-W^{-1} B^T PA)+$
$Q + L^T RL$

$P = A^T PA-A^T PBW^{-1} B^T PA-A^T PBW^{-1} B^T PA-A^T PBW^{-1} B^T PB(-W^{-1} B^T PA)+$
$Q + L^T RL$

$P = A^T PA-2A^T PBW^{-1} B^T PA+A^T PBW^{-1} B^T PBW^{-1} B^T PA+Q+L^T RL$

Now we will plug in the $L = W^{-1} B^T PA$ into the $L^T RL$ part of the equation.
$L^T RL = (W^{-1} B^T PA)^T R(W^{-1} B^T PA) = A^T PBW^{-1} RW^{-1} B^T PA$

We now will combine $D^T PD$ and $L^T RL$ and we get
$P = Q + A^T PA - 2A^T PBW^{-1} B^T PA + A^T PBW^{-1} B^T PBW^{-1} B^T PA +$
$A^T PBW^{-1} RW^{-1} B^T PA$

And we can factor out a $A^T PBW^{-1}$ and $W^{-1} B^T PA$ and the equation now becomes
$P = Q+A^T PA-2A^T PBW^{-1} B^T PA+A^T PBW^{-1}(R+B^T PB)W^{-1} B^T PA$
And since we defined $W = R + B^T PB$ we can plug it in and we can cancel out of the W^{-1} terms and the equation becomes:
$P = Q + A^T PA - 2A^T PBW^{-1} B^T PA + A^T PBW^{-1}(W)W^{-1} B^T PA$
$P = Q + A^T PA - 2A^T PBW^{-1} B^T PA + A^T PBW^{-1} B^T PA$
$P = Q + A^T PA - A^T PBW^{-1} B^T PA$
We can now factor out A^T and A and the equation now becomes:

$P = Q + A^T(P - PBW^{-1}B^TP)A$
This is the Ricatti equation. Thus, we have shown that $P = D^TPD + L^TRL$.

3.2

From the problem statement we know that
$u_t = Lx_t$
$x_{t+1} = Ax_t + Bu_t$
We can rearrange the equation for moving to the next state and we get
$= Ax_t + B(Lx_t)$
$= (A + BL)x_t$
$= Dx_t$

Now we can use the $x_t^TPx_t$ and write it in terms of the $P = D^TPD + L^TRL + Q$ equation, and we get
$x_t^TPx_t$
$x_t^TPx_t = x_t^T(D^TPT + L^TRL + Q)x_t$
$x_t^TPx_t = x_t^TD^TPDx_t + x_t^TL^TRLx_t + x_t^TQx_t$
$x_t^TPx_t = (Dx_t)^TP(Dx_t)x_t^TL^TRLx_t + x_t^TQx_t$
$x_t^TPx_t = x_{t+1}^TPx_{t+1} + x_t^TL^TRLx_t + x_t^TQx_t$

We are told to assume that Q is positive semidefinite, R is positive definite, and P is positive definite. If $x_t \neq 0$ for any t, then because we are adding a positive semidefinite and a positive definite term to $x_{t+1}^TPx_{t+1}$, that means $x_t^TPx_t$ must be greater than $x_{t+1}^TPx_{t+1}$. Therefore,
$x_{t+1}^TPx_{t+1} < x_t^TPx_t$

4 Computational Problem: Deterministic LQR

4.1

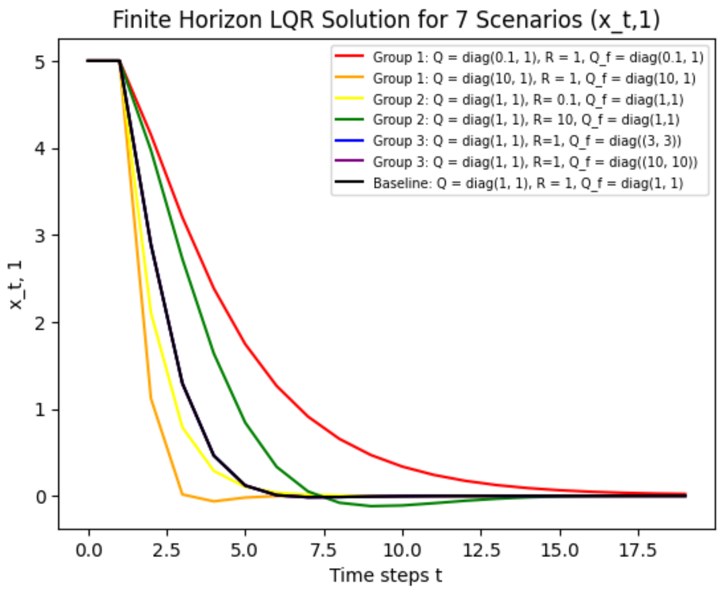
In order to explain this system in physical terms, we will think of the system as a car. We will of a car driving on a highway and at each minute or time step, t, the car has a specific position and speed/velocity. In order to change the controls of the current state you have to press the gas to accelerate the car and this effects the position and speed of the car. Therefore the position corresponds to x_1, the velocity corresponds to x_2, and the acceleration corresponds to u_t in this system.

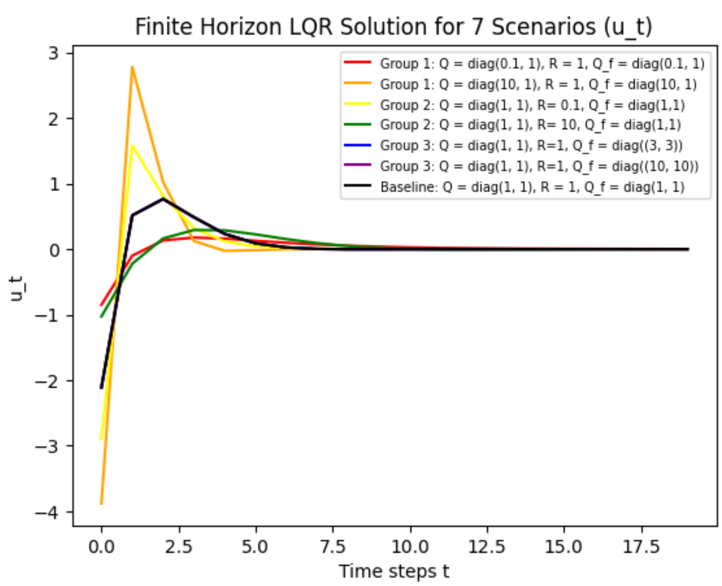
4.2

(a) From problem 2.3, we know that the discrete-time Ricatti recursion for P_t is
$u_t = -K_tx_t$
$K_t = (R + B^TP_{t+1}B)^{-1}B^TP_{t+1}A$
$P_t = (Q + A^TP_{t+1}A - A^TP_{t+1}B(R + B^TP_{t+1}B)^{-1}B^TP_{t+1}A)$

$P_T = Q_T$

(b)


Finite Horizon LQR Solution for 7 Scenarios (x_t,1)
Group 1: Q = diag(0.1, 1), R = 1, Q_f = diag(0.1, 1)
Group 1: Q = diag(10, 1), R = 1, Q_f = diag(10, 1)
Group 2: Q = diag(1, 1), R= 0.1, Q_f = diag(1,1)
Group 2: Q = diag(1, 1), R= 10, Q_f = diag(1,1)
Group 3: Q = diag(1, 1), R=1, Q_f = diag((3, 3))
Group 3: Q = diag(1, 1), R=1, Q_f = diag((10, 10))
Baseline: Q = diag(1, 1), R = 1, Q_f = diag(1, 1)
x_t, 1
Time steps t

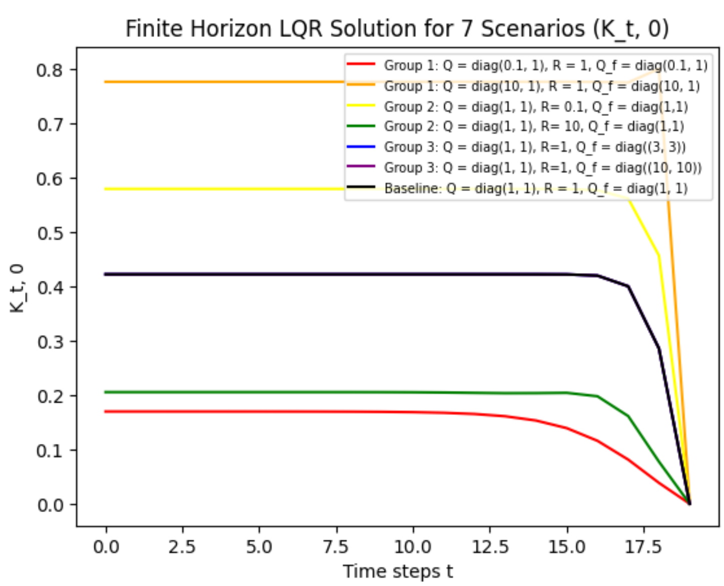

Finite Horizon LQR Solution for 7 Scenarios (u_t)
Group 1: Q = diag(0.1, 1), R = 1, Q_f = diag(0.1, 1)
Group 1: Q = diag(10, 1), R = 1, Q_f = diag(10, 1)
Group 2: Q = diag(1, 1), R= 0.1, Q_f = diag(1,1)
Group 2: Q = diag(1, 1), R= 10, Q_f = diag(1,1)
Group 3: Q = diag(1, 1), R=1, Q_f = diag((3, 3))
Group 3: Q = diag(1, 1), R=1, Q_f = diag((10, 10))
Baseline: Q = diag(1, 1), R = 1, Q_f = diag(1, 1)
u_t
Time steps t

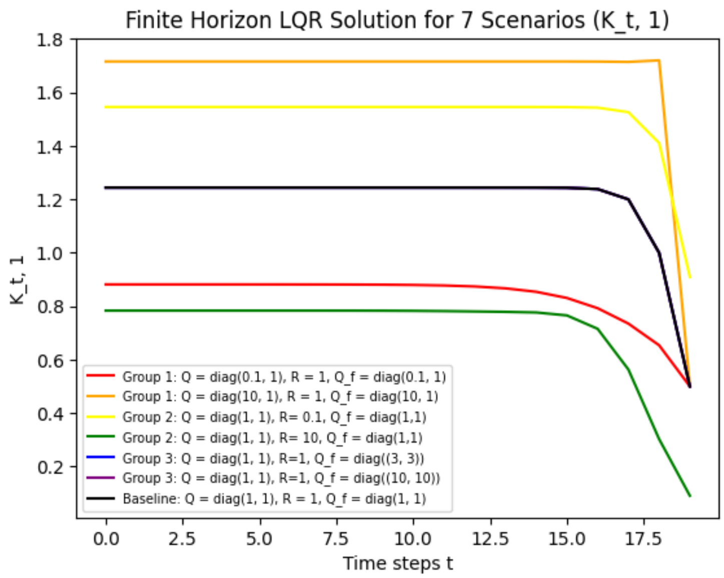
4.3

The decay speed of $x_{t,1}$ for group 1 with q = 0.1 decays slowest. The decay for group 1 with q = 10 decays the fastest. The decay for group 2 with r = 0.1 decays the second fastest. The decay for group 2 with r = 10 decays the second slowest. The decay for group 3 with $q_f = 3$ and 10 have the same decay as the baseline, which is in the middle of the different scenarios. This shows that the change in the Q value has the greatest effect on the position of the system. Overall, the Q, R, Q_f values cause the position, $x_{t,1}$ to decrease over time but large values for Q and small values for R will cause the $x_{t,1}$ to decay the quickest and then plateau.

The magnitude of the control effect u_t for group 1 with q = 0.1 shows the smallest amount of change and was the least aggressive. The magnitude for group 1 with q = 10 had the biggest change and was most aggressive. The magnitude for group 2 with r = 0.1 had the second biggest change and was the second most aggressive. The magnitude for group 2 with r = 10 has the second smallest change and was the second least aggressive control effect. The magnitude for group 3 with $q_f = 3$ and 10 have the same magnitude as the baseline, which is in the middle of the different scenarios. This shows that the change in the Q value has the greatest effect on the position of the system. The Q, R, Q_f values cause the control effect, u_t to increase in magnitude and then plateau. However, really big Q values and really small values of R will cause the u_t to increase the quickest in magnitude at the beginning of the finite horizon.

4.4

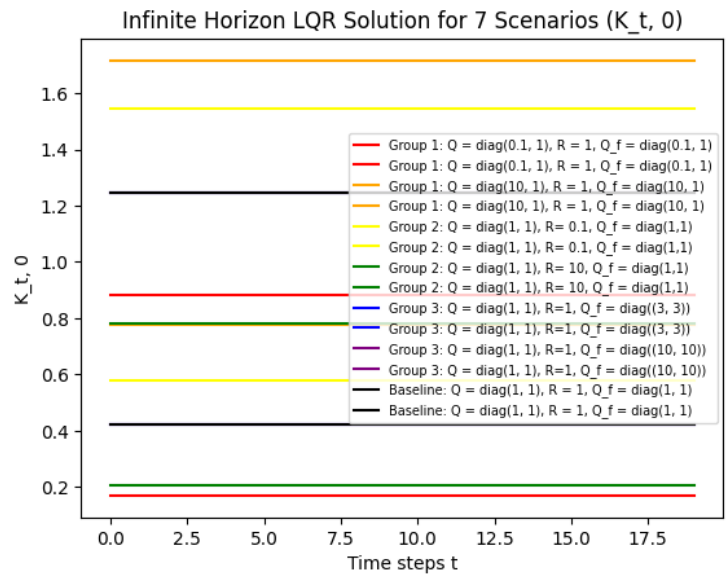


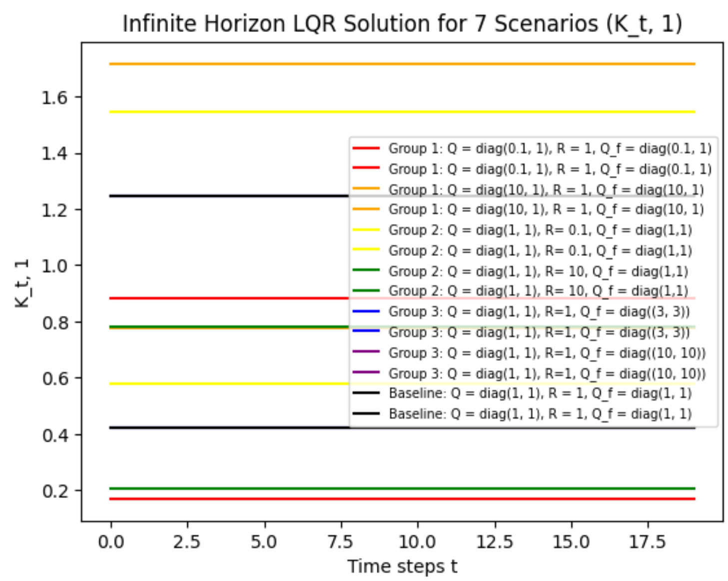


We see that the convergence for $K_{t,0}$ and $K_{t,1}$ are similar. The position and velocity feedback convergence was the quickest when the scenario was group 2 with $R = 10$. The slowest convergence occurred with group 1 with $q = 10$. This shows that convergence occurs quicker with bigger R values and it also had the lowest initial $k_{t,1}$ with group 2 and $R = 10$. However, with $K_{t,0}$ the lowest initial value occurred with group 1 and $q = 0.1$. In general both had more aggressive convergence when the q value was really big or when the r value was really small. This observation is validated by the behavior we saw for $x_{t,1}$ and u_t in the previous section.

4.5

We found the infinite horizon solutions by iterating through the Ricatti equation until convergence. That means we have a constant P and K value for the LQR control system.





Since we run the ricatti recursion until convergence we have constant value for K and P at every time step as $T \to \infty$. As we ran the infinite horizon LQR system we say that for each scenario it took the following number of time steps to converge with tolerance $1e-12$:

- infinite horizon: group 1: q: 0.1

Number of iterations to convergence: 43

- infinite horizon: group 1: q: 10

  Number of iterations to convergence: 12

- infinite horizon: group 2: r: 0.1

  Number of iterations to convergence: 15

- infinite horizon: group 2: r: 10

  Number of iterations to convergence: 37

- infinite horizon: group 3: q_f: 3

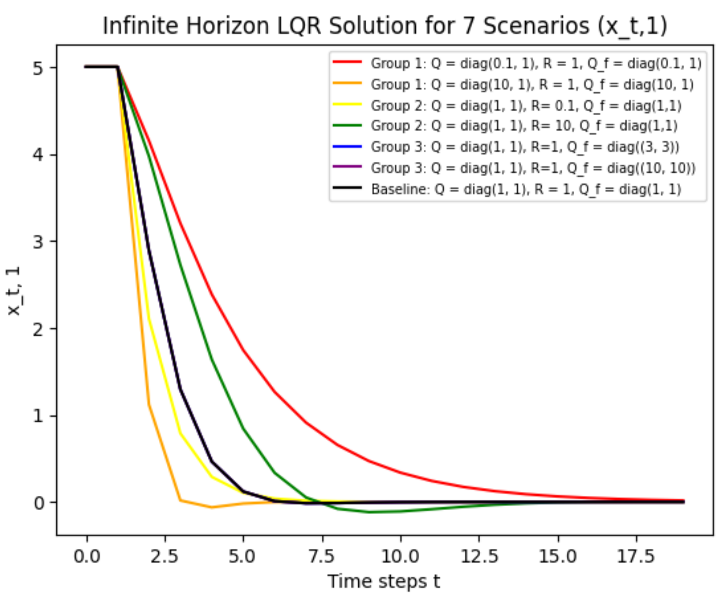
  Number of iterations to convergence: 18

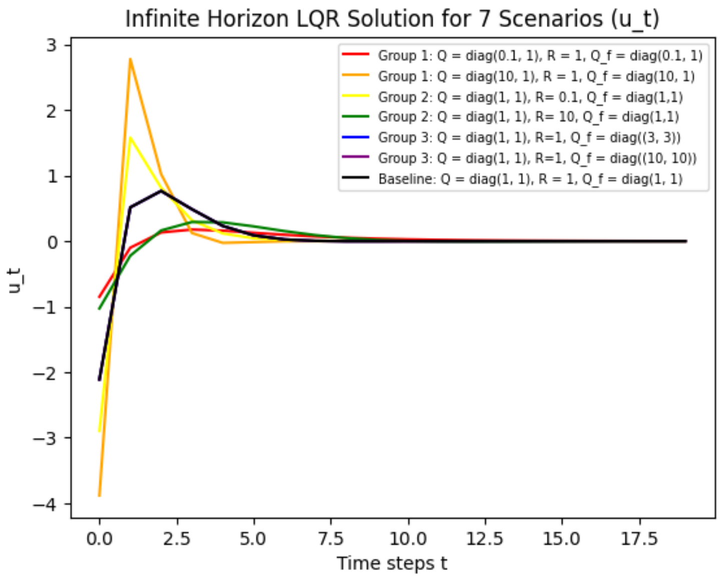
- infinite horizon: group 3: q_f: 10

  Number of iterations to convergence: 18

- infinite horizon: baseline

  Number of iterations to convergence: 18





For the infinite horizon versions of the $x_{t,1}$ and the u_t graphs have similar behavior to the finite horizon versions fo the $x_{t,1}$ and the u_t. Therefore, we can describe the decay speed of the position of the infinite horizon $x_{t,1}$ the same way we described the decay speed of the position of the finite horizon $x_{t,1}$ in part 4.3. Similarly, we can describe the magnitude of the control effect of the infinite horizon u_t the same way we described the magnitude of the control effect of the finite horizon u_t in part 4.3.