Homework #2: LQR

1. Theoretical Exercises

(a) Consider the following optimal control problem, which considers a linear system with additive noise and quadratic cost:

$$\begin{aligned}
&\min_{x,u} \mathbb{E}[\sum_{t=0}^{T-1} x_t^\top Q x_t + u_t^\top R u_t] + \mathbb{E}[x_T^\top Q x_T] \\
&\text{s.t. } x_{t+1} = Ax_t + Bu_t + w_t, \forall t = 0, 1, 2, ..., T-1
\end{aligned}$$

with w_t independent random vectors with $\mathbb{E}[w_t] = 0$, and $\mathbb{E}[w_t w_t^\top] = \Sigma_w$.

Find an LQR-like sequence of matrix updates that computes the optimal cost-to-go at all times and the optimal feedback controller at all times. Describe the expected cost incurred in excess of the expected cost in the case when there is no noise.

Solution:
Define $J_i(x)$ as done in lecture to be the cost for i steps to go and perform Value Iteration:

$$\begin{aligned}
\textit{Initialize } J_0(x) &= x^\top P_0 x \\
\Rightarrow J_1(x) &= \min_u x^\top Q x + u^\top R u + \mathbb{E}[J_0(Ax + Bu + w_{T-1})] \\
&= \min_u x^\top Q x + u^\top R u + \mathbb{E}[(Ax + Bu + w_{T-1})^\top P_0 (Ax + Bu + w_{T-1})] \\
&= \min_u x^\top Q x + u^\top R u + (Ax + Bu)^\top P_0 (Ax + Bu) + \mathbb{E}[w_{T-1}^\top P_0 w_{T-1}]
\end{aligned}$$

Where we use that the expectation of the noise variable is 0. To find the minimum over u, we set the gradient of $J_1(x)$ with respect to u and see that $\mathbb{E}[w^\top P_0 w]$ does not contribute:

$$\begin{aligned}
\nabla_u[...] &= 2Ru + 2B^\top P_0(Ax + Bu) = 0, \\
\textit{hence } u^* &= -(R + B^\top P_0 B)^{-1} B^\top P_0 A x = K_1 x
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
J_1(x) &= x^\top Q x + (K_1 x)^\top R (K_1 x) + (Ax + B(K_1 x))^\top P_0 (Ax + B(K_1 x)) + \mathbb{E}[w_{T-1}^\top P_0 w_{T-1}] \\
&= x^\top Q x + (K_1 x)^\top R (K_1 x) + ((A + BK_1)x)^\top P_0 ((A + BK_1)x) + \mathbb{E}[w_{T-1}^\top P_0 w_{T-1}] \\
&= x^\top [Q + K_1^\top R K_1 + (A + BK_1)^\top P_0 (A + BK_1)] x + \mathbb{E}[w_{T-1}^\top P_0 w_{T-1}] \\
&= x^\top P_1 x + \mathbb{E}[w_{T-1}^\top P_0 w_{T-1}] \\
&= x^\top P_1 x + \mathbb{E}[\text{tr}(w_{T-1}^\top P_0 w_{T-1})] \\
&= x^\top P_1 x + \text{tr}(P_0 \Sigma_w)
\end{aligned}$$

Where we use the linearity and cyclic properties of the trace operator to simplify the expectation term. We can see that we'll carry this trace term when we use J_1 in our expression *for* J_2 *and so on, so let's assume that*

$$\begin{aligned}
J_i(x) &= x^\top P_i x + \sum_{j=0}^{i-1} \mathrm{tr}(P_j \Sigma_w) \\
K_i &= -(R + B^\top P_{i-1} B)^{-1} B^\top P_{i-1} A \\
P_i &= Q + K_i^\top R K_i + (A + BK_i)^\top P_{i-1} (A + BK_i)
\end{aligned}$$

We can confirm that J_{i+1} *follows. Here we drop the subscript on* w *for clarity since the first and second moments of* w_t *are the same for all* t*:*

$$\begin{aligned}
J_{i+1}(x) &= \min_u x^\top Q x + u^\top R u + \mathbb{E}[J_i(Ax + Bu + w)] \\
&= \min_u x^\top Q x + u^\top R u + \mathbb{E}[(Ax + Bu + w)^\top P_i (Ax + Bu + w) + \sum_{j=0}^{i-1} \mathrm{tr}(P_j \Sigma_w)] \\
&= \min_u x^\top Q x + u^\top R u + (Ax + Bu)^\top P_i (Ax + Bu) + \mathbb{E}[w^\top P_i w] + \sum_{j=0}^{i-1} \mathrm{tr}(P_j \Sigma_w) \\
&= \min_u x^\top Q x + u^\top R u + (Ax + Bu)^\top P_i (Ax + Bu) + \sum_{j=0}^{i} \mathrm{tr}(P_j \Sigma_w)
\end{aligned}$$

Clearly, we have the same form for the gradient with respect to u *as we did for* J_1 *and we can confirm by plugging in* u^**, exactly as we did before, that the claim is true. Thus, the expected excess cost incurred compared to the no-noise case for an* i*-step horizon is*

$$\sum_{j=0}^{i-1} \mathrm{tr}(P_j \Sigma_w)$$

(b) Now consider a linear system with multiplicative noise and quadratic cost:

$$\begin{aligned}
&\min_{x,u} \mathbb{E}[x_T^\top Q x_T] \\
&\text{s.t. } x_{t+1} = Ax_t + (B + W_t)u_t, \ \forall t = 0, 1, 2, ..., T-1
\end{aligned}$$

Here $Q \in R^{n_X \times n_X}, A \in R^{n_X \times n_X}, B \in R^{n_X \times n_U}$ are given and fixed. $W_t \in R^{n_X \times n_U}, t = 0, 1, ..., T-1$ are independent random matrices with $\mathbb{E}[W_t] = 0$. Higher-order expectations involving W_t will show up. These higher-order expectations are not assumed to be zero, and you should just keep these expectations around, i.e., no need to try to simplify these (and not possible anyway unless additional assumptions are made).

Find an LQR-like sequence of matrix updates that computes the optimal cost-to-go at all times and the optimal feedback controller at all times.
Solution:

Again, we initialize $J_0(x) = x^\top P_0 x$

$$\begin{aligned}
\Rightarrow J_1(x) &= \min_u \mathbb{E}[J_0(Ax + (B + W_{T-1})u)] \\
&= \min_u \mathbb{E}[(Ax + (B + W_{T-1})u)^\top P_0(Ax + (B + W_{T-1})u)] \\
&= \min_u \mathbb{E}[(Ax + Bu + W_{T-1}u)^\top P_0(Ax(Ax + (B + W_{T-1})u) + Bu + W_{T-1}u)] \\
&= \min_u (Ax + Bu)^\top P_0(Ax + Bu) + \mathbb{E}[(W_{T-1}u)^\top P_0(W_{T-1}u)] \\
&= \min_u (Ax + Bu)^\top P_0(Ax + Bu) + u^\top \mathbb{E}[W_{T-1}^\top P_0 W_{T-1}]u
\end{aligned}$$

Just as we did previously, take the gradient and set it to 0 to minimize over u

$$\begin{aligned}
\nabla_u[...] &= 2B^\top P_0 Bu + 2B^\top P_0 Ax + 2\mathbb{E}[W_{T-1}^\top P_0 W_{T-1}]u = 0, \\
\textit{hence } u^* &= -(B^\top P_0 B + \mathbb{E}[W_{T-1}^\top P_0 W_{T-1}])^{-1} B^\top P_0 Ax = K_1 x \\
\Rightarrow \; K_i &= -(B^\top P_{i-1} B + \mathbb{E}[W_{T-i}^\top P_{i-1} W_{T-i}])^{-1} B^\top P_{i-1} A
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
J_1(x) &= (Ax + B(K_1 x))^\top P_0(Ax + B(K_1 x)) + (K_1 x)^\top \mathbb{E}[W_{T-1}^\top P_0 W_{T-1}]K_1 x \\
&= ((A + BK_1)x)^\top P_0(A + BK_1)x) + (K_1 x)^\top \mathbb{E}[W_{T-1}^\top P_0 W_{T-1}]K_1 x \\
&= x^\top[(A + BK_1)^\top P_0(A + BK_1) + K_1^\top \mathbb{E}[W_{T-1}^\top P_0 W_{T-1}]K_1]x \\
&= x^\top P_1 x \\
\Rightarrow P_i &= (A + BK_i)^\top P_{i-1}(A + BK_i) + K_i^\top \mathbb{E}[W_{T-i}^\top P_{i-1} W_{T-i}]K_i
\end{aligned}$$

And the cost-to-go for i-step given by:

$$J_i(x) = x^\top P_i x$$