

AE 4803 Robotics and Autonomy
Professor Evangelos Theodorou
Homework 1

Luis Pimentel Jackson Crandell
lpimentel3@gatech.edu jackcrandell@gatech.edu

October 2, 2020

Part 1.

1.1) First we linearize the dynamics $\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, \mathbf{u}, t)$:

$$\begin{aligned}\frac{d\mathbf{x}}{dt} &= f(\mathbf{x}, \mathbf{u}, t) \\ &= f(\bar{\mathbf{x}} + (\mathbf{x} - \bar{\mathbf{x}}), \bar{\mathbf{u}} + (\mathbf{u} - \bar{\mathbf{u}}), t) \\ &= f(\bar{\mathbf{x}} + (\mathbf{x} - \bar{\mathbf{x}}), \bar{\mathbf{u}} + (\mathbf{u} - \bar{\mathbf{u}}), t) \\ &= f(\bar{\mathbf{x}} + \delta\mathbf{x}, \bar{\mathbf{u}} + \delta\mathbf{u}, t) \\ &= f(\bar{\mathbf{x}}, \bar{\mathbf{u}}, t) + \nabla_{\mathbf{x}} f \delta\mathbf{x} + \nabla_{\mathbf{u}} f \delta\mathbf{u} \\ \frac{d\mathbf{x}}{dt} - f(\bar{\mathbf{x}}, \bar{\mathbf{u}}, t) &= \nabla_{\mathbf{x}} f \delta\mathbf{x} + \nabla_{\mathbf{u}} f \delta\mathbf{u} \\ \frac{d\delta\mathbf{x}}{dt} &= \nabla_{\mathbf{x}} f \delta\mathbf{x} + \nabla_{\mathbf{u}} f \delta\mathbf{u}\end{aligned}$$

Then we time discretize the linearized dynamics:

$$\begin{aligned}\frac{\delta\mathbf{x}(t_{k+1}) - \delta\mathbf{x}(t_k)}{t_{k+1} - t_k} &= \nabla_{\mathbf{x}} f \delta\mathbf{x}(t_k) + \nabla_{\mathbf{u}} f \delta\mathbf{u}(t_k) \\ \frac{\delta\mathbf{x}(t_{k+1}) - \delta\mathbf{x}(t_k)}{dt} &= \nabla_{\mathbf{x}} f \delta\mathbf{x}(t_k) + \nabla_{\mathbf{u}} f \delta\mathbf{u}(t_k) \\ \delta\mathbf{x}(t_{k+1}) - \delta\mathbf{x}(t_k) &= \nabla_{\mathbf{x}} f dt \delta\mathbf{x}(t_k) + \nabla_{\mathbf{u}} f dt \delta\mathbf{u}(t_k) \\ \delta\mathbf{x}(t_{k+1}) &= \nabla_{\mathbf{x}} f dt \delta\mathbf{x}(t_k) + \delta\mathbf{x}(t_k) + \nabla_{\mathbf{u}} f dt \delta\mathbf{u}(t_k) \\ \delta\mathbf{x}(t_{k+1}) &= (\mathbf{I}_{n \times n} + \nabla_{\mathbf{x}} f dt) \delta\mathbf{x}(t_k) + (\nabla_{\mathbf{u}} f dt) \delta\mathbf{u}(t_k) \\ \delta\mathbf{x}(t_{k+1}) &= \mathbf{\Phi}(t_k) \delta\mathbf{x}(t_k) + \mathbf{B}(t_k) \delta\mathbf{u}(t_k)\end{aligned}$$

Where:

$$\Phi(t_k) = \mathbf{I}_{n \times n} + \nabla_{\mathbf{x}} f(t_{k+1} + t_k)$$

$$\mathbf{B}(t_k) = \nabla_{\mathbf{u}} f(t_{k+1} + t_k)$$

1.2) Next we want to express the second-order expansion of $Q(\mathbf{x}(t_k), \mathbf{u}(t_k))$. First we consider the general 2nd Order Taylor Series Expansion of a single-variable function and a multi-variable function:

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2$$

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

$$+ \frac{1}{2}f_{xx}(a, b)(x - a)^2 + f_{xy}(a, b)(x - a)(y - b) + \frac{1}{2}f_{yy}(a, b)(y - b)^2$$

We can generalize these forms in matrix-algebra for our state-action value function.

$$Q(\mathbf{x}(t_k), \mathbf{u}(t_k)) = \mathcal{L}(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) + V(\mathbf{x}(t_{k+1}), t_{k+1})$$

We will start by writing the expansion of the running cost and the value function individually.

First we do the second order expansion of the running cost along the nominal trajectory $\bar{\mathbf{x}}$ and $\bar{\mathbf{u}}$:

$$\begin{aligned} & \mathcal{L}(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) \\ &= \mathcal{L}(\mathbf{x}(t_k) + (\bar{\mathbf{x}}(t_k) - \bar{\mathbf{x}}(t_k)), \mathbf{u}(t_k) + (\bar{\mathbf{u}}(t_k) - \bar{\mathbf{u}}(t_k)), t_k) \\ &= \mathcal{L}(\bar{\mathbf{x}}(t_k) + (\mathbf{x}(t_k) - \bar{\mathbf{x}}(t_k)), \bar{\mathbf{u}}(t_k) + (\mathbf{u}(t_k) - \bar{\mathbf{u}}(t_k)), t_k) \\ &= \mathcal{L}(\bar{\mathbf{x}}(t_k) + \delta\mathbf{x}(t_k), \bar{\mathbf{u}}(t_k) + \delta\mathbf{u}(t_k), t_k) \\ &= \underbrace{\ell(\bar{\mathbf{x}}(t_k), \bar{\mathbf{u}}(t_k), t_k)dt}_{\mathcal{L}} + \underbrace{\left(\nabla_{\mathbf{x}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)^T}_{\mathcal{L}_{\mathbf{x}}} \delta\mathbf{x}(t_k) + \underbrace{\left(\nabla_{\mathbf{u}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)^T}_{\mathcal{L}_{\mathbf{u}}} \delta\mathbf{u}(t_k) \\ &+ \frac{1}{2}\delta\mathbf{x}(t_k)^T \underbrace{\left(\nabla_{\mathbf{x}\mathbf{x}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)}_{\mathcal{L}_{\mathbf{x}\mathbf{x}}} \delta\mathbf{x}(t_k) + \frac{1}{2}\delta\mathbf{u}(t_k)^T \underbrace{\left(\nabla_{\mathbf{u}\mathbf{u}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)}_{\mathcal{L}_{\mathbf{u}\mathbf{u}}} \delta\mathbf{u}(t_k) \\ &+ \frac{1}{2}\delta\mathbf{u}(t_k)^T \underbrace{\left(\nabla_{\mathbf{x}\mathbf{u}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)}_{\mathcal{L}_{\mathbf{x}\mathbf{u}}} \delta\mathbf{x}(t_k) + \frac{1}{2}\delta\mathbf{x}(t_k)^T \underbrace{\left(\nabla_{\mathbf{u}\mathbf{x}}\ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k)dt\right)}_{\mathcal{L}_{\mathbf{u}\mathbf{x}}} \delta\mathbf{u}(t_k) \end{aligned}$$

We denote:

$$\begin{aligned}
\mathcal{L} &= \ell(\bar{\mathbf{x}}(t_k), \bar{\mathbf{u}}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{x}} &= \nabla_{\mathbf{x}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{u}} &= \nabla_{\mathbf{u}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{xx}} &= \nabla_{\mathbf{xx}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{uu}} &= \nabla_{\mathbf{uu}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{xu}} &= \nabla_{\mathbf{xu}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt \\
\mathcal{L}_{\mathbf{ux}} &= \nabla_{\mathbf{ux}} \ell(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) dt
\end{aligned}$$

With this we write second order expansion of the running cost as:

$$\begin{aligned}
&\mathcal{L}(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) \\
&= \mathcal{L} + \mathcal{L}_{\mathbf{x}}^T \delta \mathbf{x}(t_k) + \mathcal{L}_{\mathbf{u}}^T \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathcal{L}_{\mathbf{xx}} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathcal{L}_{\mathbf{uu}} \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathcal{L}_{\mathbf{xu}} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathcal{L}_{\mathbf{ux}} \delta \mathbf{u}(t_k)
\end{aligned}$$

Next we do the second order expansion of the value function along the nominal trajectory $\bar{\mathbf{x}}$:

$$\begin{aligned}
&V(\mathbf{x}(t_{k+1}), t_{k+1}) \\
&= V(\mathbf{x}(t_{k+1}) + (\bar{\mathbf{x}}(t_{k+1}) - \bar{\mathbf{x}}(t_{k+1})), t_{k+1}) \\
&= V(\bar{\mathbf{x}}(t_{k+1}) + (\mathbf{x}(t_{k+1}) - \bar{\mathbf{x}}(t_{k+1})), t_{k+1}) \\
&= V(\bar{\mathbf{x}}(t_{k+1}) + \delta \mathbf{x}(t_{k+1}), t_{k+1}) \\
&= V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1}) + \nabla_{\mathbf{x}} V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1})^T \delta \mathbf{x}(t_{k+1}) + \frac{1}{2} \delta \mathbf{x}(t_{k+1})^T \nabla_{\mathbf{xx}} V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1}) \delta \mathbf{x}(t_{k+1})
\end{aligned}$$

We denote:

$$\begin{aligned}
\delta \mathbf{x}(t_{k+1}) &= \mathbf{x}(t_{k+1}) - \bar{\mathbf{x}}(t_{k+1}) \\
V(t_{k+1}) &= V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1}) \\
V_{\mathbf{x}}(t_{k+1}) &= \nabla_{\mathbf{x}} V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1}) \\
V_{\mathbf{xx}}(t_{k+1}) &= \nabla_{\mathbf{xx}} V(\bar{\mathbf{x}}(t_{k+1}), t_{k+1})
\end{aligned}$$

Therefore:

$$= V(t_{k+1}) + V_{\mathbf{x}}(t_{k+1})^T \delta \mathbf{x}(t_{k+1}) + \frac{1}{2} \delta \mathbf{x}(t_{k+1})^T V_{\mathbf{xx}}(t_{k+1}) \delta \mathbf{x}(t_{k+1})$$

We recall from the linearized dynamics in Part 1.1 that $\delta \mathbf{x}(t_{k+1}) = \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \mathbf{B}(t_k)\delta \mathbf{u}(t_k)$. Therefore:

$$\begin{aligned}
&= V(t_{k+1}) + V_{\mathbf{x}}(t_{k+1})^T \left(\mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \right) \\
&+ \frac{1}{2} \left(\mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \right)^T V_{\mathbf{xx}}(t_{k+1}) \left(\mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \right) \\
&= V(t_{k+1}) + V_{\mathbf{x}}(t_{k+1})^T \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + V_{\mathbf{x}}(t_{k+1})^T \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathbf{\Phi}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathbf{\Phi}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k)\delta \mathbf{u}(t_k)
\end{aligned}$$

The last two terms can be combined as they evaluate to equal scalars. Therefore, we can simplify the second order expansion of the value function as:

$$\begin{aligned}
&V(\mathbf{x}(t_{k+1}), t_{k+1}) \\
&= V(t_{k+1}) + V_{\mathbf{x}}(t_{k+1})^T \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + V_{\mathbf{x}}(t_{k+1})^T \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathbf{\Phi}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k)\delta \mathbf{u}(t_k) \\
&+ \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{\Phi}(t_k)\delta \mathbf{x}(t_k)
\end{aligned}$$

We can now consider the second order expansion of the action-value function itself:

$$\begin{aligned}
&Q(\mathbf{x}(t_k), \mathbf{u}(t_k), t_k) \\
&= Q(\mathbf{x}(t_k) + (\bar{\mathbf{x}}(t_k) - \bar{\mathbf{x}}(t_k)), \mathbf{u}(t_k) + (\bar{\mathbf{u}}(t_k) - \bar{\mathbf{u}}(t_k)), t_k) \\
&= Q(\bar{\mathbf{x}}(t_k) + (\mathbf{x}(t_k) - \bar{\mathbf{x}}(t_k)), \bar{\mathbf{u}}(t_k) + (\mathbf{u}(t_k) - \bar{\mathbf{u}}(t_k)), t_k) \\
&= Q(\bar{\mathbf{x}}(t_k) + \delta \mathbf{x}(t_k), \bar{\mathbf{u}}(t_k) + \delta \mathbf{u}(t_k), t_k) \\
&= Q_0 + Q_{\mathbf{x}}^T \delta \mathbf{x}(t_k) + Q_{\mathbf{u}}^T \delta \mathbf{u}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T Q_{\mathbf{xx}} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T Q_{\mathbf{uu}} \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{u}(t_k)^T Q_{\mathbf{xu}} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T Q_{\mathbf{ux}} \delta \mathbf{u}(t_k)
\end{aligned}$$

Similarly to the value function expansion, the last two terms can be combined as they evaluate to equal scalars, we can simplify the second order expansion of the state-action value function as:

$$\begin{aligned}
&Q(\bar{\mathbf{x}}(t_k) + \delta \mathbf{x}(t_k), \bar{\mathbf{u}}(t_k) + \delta \mathbf{u}(t_k), t_k) \\
&= Q_0 + Q_{\mathbf{x}}^T \delta \mathbf{x}(t_k) + Q_{\mathbf{u}}^T \delta \mathbf{u}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T Q_{\mathbf{xx}} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T Q_{\mathbf{uu}} \delta \mathbf{u}(t_k) \\
&+ \delta \mathbf{x}(t_k)^T Q_{\mathbf{ux}} \delta \mathbf{u}(t_k)
\end{aligned}$$

Finally we can combine the second order expansions of the running cost and value function to express the second order expansion of the state-action value function:

$$\begin{aligned}
& Q(\bar{\mathbf{x}}(t_k) + \delta \mathbf{x}(t_k), \bar{\mathbf{u}}(t_k) + \delta \mathbf{u}(t_k), t_k) \\
&= \mathcal{L} + \mathcal{L}_x^T \delta \mathbf{x}(t_k) + \mathcal{L}_u^T \delta \mathbf{u}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathcal{L}_{xx} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathcal{L}_{uu} \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathcal{L}_{xu} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T \mathcal{L}_{ux} \delta \mathbf{u}(t_k) \\
&+ V(t_{k+1}) + V_{\mathbf{x}}(t_{k+1})^T \Phi(t_k) \delta \mathbf{x}(t_k) + V_{\mathbf{x}}(t_{k+1})^T \mathbf{B}(t_k) \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \Phi(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \Phi(t_k) \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k) \delta \mathbf{u}(t_k) \\
&+ \delta \mathbf{u}(t_k)^T \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \Phi(t_k) \delta \mathbf{x}(t_k)
\end{aligned}$$

We can group the zero, first, and second order terms as:

$$\begin{aligned}
&= \underbrace{\left(\mathcal{L} + V(t_{k+1}) \right)}_{Q_0} + \underbrace{\left(\mathcal{L}_x + \Phi(t_k)^T V_{\mathbf{x}}(t_{k+1}) \right)^T}_{Q_x} \delta \mathbf{x}(t_k) + \underbrace{\left(\mathcal{L}_u + \mathbf{B}(t_k)^T V_{\mathbf{x}}(t_{k+1}) \right)^T}_{Q_u} \delta \mathbf{u}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \underbrace{\left(\mathcal{L}_{xx} + \Phi(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \Phi(t_k) \right)}_{Q_{xx}} \delta \mathbf{x}(t_k) \\
&+ \frac{1}{2} \delta \mathbf{u}(t_k)^T \underbrace{\left(\mathcal{L}_{uu} + \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k) \right)}_{Q_{uu}} \delta \mathbf{u}(t_k) \\
&+ \delta \mathbf{u}(t_k)^T \underbrace{\left(\frac{1}{2} \mathcal{L}_{ux} + \frac{1}{2} \mathcal{L}_{xu}^T + \frac{1}{2} \mathbf{B}(t_k)^T \left(V_{\mathbf{xx}}(t_{k+1}) + V_{\mathbf{xx}}(t_{k+1})^T \right) \Phi(t_k) \right)}_{Q_{ux}} \delta \mathbf{x}(t_k)
\end{aligned}$$

Where we denote:

$$\begin{aligned}
Q_0 &= \mathcal{L} + V(t_{k+1}) \\
Q_x &= \mathcal{L}_x + \Phi(t_k)^T V_{\mathbf{x}}(t_{k+1}) \\
Q_u &= \mathcal{L}_u + \mathbf{B}(t_k)^T V_{\mathbf{x}}(t_{k+1}) \\
Q_{xx} &= \mathcal{L}_{xx} + \Phi(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \Phi(t_k) \\
Q_{uu} &= \mathcal{L}_{uu} + \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \mathbf{B}(t_k) \\
Q_{ux} &= \mathbf{B}(t_k)^T V_{\mathbf{xx}}(t_{k+1}) \Phi(t_k) + \mathcal{L}_{ux}
\end{aligned}$$

- 1.3) Next we compute the optimal control corrections $\delta \mathbf{u}^*(t_k)$. We do this by minimizing $\delta \mathbf{u}(t_k)$ with respect to the state-action value function $Q(\mathbf{x}(t_k), \mathbf{u}(t_k))$. This is the same as minimizing $\mathbf{u}(t_k)$ to compute the locally optimal control $\mathbf{u}^*(t_k)$ in the Bellman Principle. This so since $\mathbf{u}^*(t_k) = \bar{\mathbf{u}}(t_k) + \delta \mathbf{u}^*(t_k)$.

$$\begin{aligned} & \min_{\delta \mathbf{u}(t_k)} \left[Q(\mathbf{x}(t_k), \mathbf{u}(t_k)) \right] \\ &= \min_{\delta \mathbf{u}(t_k)} \left[\left(\mathcal{L} + V(t_{k+1}) \right) + \left(\mathcal{L}_{\mathbf{x}} + \Phi(t_k)^T V_{\mathbf{x}}(t_{k+1}) \right)^T \delta \mathbf{x}(t_k) + \left(\mathcal{L}_{\mathbf{u}} + \mathbf{B}(t_k)^T V_{\mathbf{x}}(t_{k+1}) \right)^T \delta \mathbf{u}(t_k) \right. \\ &+ \frac{1}{2} \delta \mathbf{x}(t_k)^T \left(\mathcal{L}_{\mathbf{x}\mathbf{x}} + \Phi(t_k)^T V_{\mathbf{x}\mathbf{x}}(t_{k+1}) \Phi(t_k) \right) \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{u}(t_k)^T \left(\mathcal{L}_{\mathbf{u}\mathbf{u}} + \mathbf{B}(t_k)^T V_{\mathbf{x}\mathbf{x}}(t_{k+1}) \mathbf{B}(t_k) \right) \delta \mathbf{u}(t_k) \\ &+ \left. \delta \mathbf{u}(t_k)^T \left(\frac{1}{2} \mathcal{L}_{\mathbf{u}\mathbf{x}} + \frac{1}{2} \mathcal{L}_{\mathbf{x}\mathbf{u}}^T + \frac{1}{2} \mathbf{B}(t_k)^T \left(V_{\mathbf{x}\mathbf{x}}(t_{k+1}) + V_{\mathbf{x}\mathbf{x}}(t_{k+1})^T \right) \Phi(t_k) \right) \delta \mathbf{x}(t_k) \right] \end{aligned}$$

To solve this take the gradient of $Q(\mathbf{x}(t_k), \mathbf{u}(t_k))$ with respect to $\delta \mathbf{u}(t_k)$ and set it equal to zero:

$$\begin{aligned} 0 &= \underbrace{\left(\mathcal{L}_{\mathbf{u}} + \mathbf{B}(t_k)^T V_{\mathbf{x}}(t_{k+1}) \right)}_{Q_{\mathbf{u}}} + \underbrace{\left(\frac{1}{2} \mathcal{L}_{\mathbf{u}\mathbf{x}} + \frac{1}{2} \mathcal{L}_{\mathbf{x}\mathbf{u}}^T + \frac{1}{2} \mathbf{B}(t_k)^T \left(V_{\mathbf{x}\mathbf{x}}(t_{k+1}) + V_{\mathbf{x}\mathbf{x}}(t_{k+1})^T \right) \Phi(t_k) \right)}_{Q_{\mathbf{u}\mathbf{x}}} \delta \mathbf{x}(t_k) \\ &+ \underbrace{\left(\mathcal{L}_{\mathbf{u}\mathbf{u}} + \mathbf{B}(t_k)^T V_{\mathbf{x}\mathbf{x}}(t_{k+1}) \mathbf{B}(t_k) \right)}_{Q_{\mathbf{u}\mathbf{u}}} \delta \mathbf{u}(t_k) \end{aligned}$$

Therefore:

$$\begin{aligned} 0 &= Q_{\mathbf{u}} + Q_{\mathbf{u}\mathbf{x}} \delta \mathbf{x}(t_k) + Q_{\mathbf{u}\mathbf{u}} \delta \mathbf{u}(t_k) \\ - Q_{\mathbf{u}\mathbf{u}} \delta \mathbf{u}(t_k) &= Q_{\mathbf{u}} + Q_{\mathbf{u}\mathbf{x}} \delta \mathbf{x}(t_k) \\ \delta \mathbf{u}(t_k) &= \underbrace{-Q_{\mathbf{u}\mathbf{u}}^{-1} Q_{\mathbf{u}}}_{\mathbf{l}_{\mathbf{u}}(t_k)} + \underbrace{\left(-Q_{\mathbf{u}\mathbf{u}}^{-1} Q_{\mathbf{u}\mathbf{x}} \right)}_{\mathbf{L}_{\mathbf{u}}(t_k)} \delta \mathbf{x}(t_k) \end{aligned}$$

Further, we denote $\mathbf{l}_{\mathbf{u}}(t_k)$ as the feedforward term and $\mathbf{L}_{\mathbf{u}}(t_k)$ as the feedback term:

$$\begin{aligned} \mathbf{l}_{\mathbf{u}}(t_k) &= -Q_{\mathbf{u}\mathbf{u}}^{-1} Q_{\mathbf{u}} \\ \mathbf{L}_{\mathbf{u}}(t_k) &= -Q_{\mathbf{u}\mathbf{u}}^{-1} Q_{\mathbf{u}\mathbf{x}} \end{aligned}$$

And determine the optimal control corrections to be:

$$\delta \mathbf{u}^*(t_k) = \mathbf{l}_{\mathbf{u}}(t_k) + \mathbf{L}_{\mathbf{u}}(t_k) \delta \mathbf{x}(t_k)$$

- 1.4) We can compute the equations $V(t_k)$, $V_x(t_k)$, $V_{xx}(t_k)$ in backward time by substituting in our optimal control correction $\delta \mathbf{u}^*(t_k)$ into the expanded state-action value function $Q(\mathbf{x}(t_k), \mathbf{u}(t_k))$.

$$\begin{aligned}
& Q(\mathbf{x}(t_k), \mathbf{u}(t_k)) \Big|_{\delta \mathbf{u}^*(t_k) = \mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k)} \\
&= Q_0 + Q_x^T \delta \mathbf{x}(t_k) + Q_u^T \left(\mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k) \right) \\
&+ \frac{1}{2} \delta \mathbf{x}(t_k)^T Q_{xx} \delta \mathbf{x}(t_k) + \frac{1}{2} \left(\mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k) \right)^T Q_{uu} \left(\mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k) \right) \\
&+ \frac{1}{2} \left(\mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k) \right)^T Q_{ux} \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T Q_{xu} \left(\mathbf{l}_u(t_k) + \mathbf{L}_u(t_k) \delta \mathbf{x}(t_k) \right)
\end{aligned}$$

To compute the backward terms we can denote:

$$V(\bar{\mathbf{x}}(t_k) + \delta \mathbf{x}(t_k), t) = V(t_k) + V_x(t_k)^T \delta \mathbf{x}(t_k) + \frac{1}{2} \delta \mathbf{x}(t_k)^T V_{xx}(t_k) \delta \mathbf{x}(t_k)$$

Therefore the zero order terms of the state-action value function in $\delta \mathbf{x}$ are equal to $V(t_k)$, the first order terms are equal to $V_x(t_k)$, and the second order terms are equal to $V_{xx}(t_k)$.

First can group the zero order terms that equate to $V(t_k)$:

$$V(t_k) = Q_0 + Q_u^T \mathbf{l}_u(t_k) + \frac{1}{2} \mathbf{l}_u(t_k)^T Q_{uu} \mathbf{l}_u(t_k)$$

Next we can group the first order terms that equate to $V_x(t_k)$:

$$V_x(t_k) = Q_x + Q_u^T \mathbf{L}_u(t_k) + Q_{xu} \mathbf{l}_u(t_k) + \mathbf{L}_u(t_k)^T Q_{uu} \mathbf{l}_u(t_k)$$

Finally we can group the first order terms that equate to $V_{xx}(t_k)$:

$$V_{xx}(t_k) = Q_{xx} + \mathbf{L}_u(t_k)^T Q_{uu} \mathbf{L}_u(t_k) + \mathbf{L}_u(t_k)^T Q_{ux} + Q_{xu} \mathbf{L}_u(t_k)$$

In summary, the backward equations can be expressed as:

$$\begin{aligned}
V(t_k) &= Q_0 + Q_u^T \mathbf{l}_u(t_k) + \frac{1}{2} \mathbf{l}_u(t_k)^T Q_{uu} \mathbf{l}_u(t_k) \\
V_x(t_k) &= Q_x + Q_u^T \mathbf{L}_u(t_k) + Q_{xu} \mathbf{l}_u(t_k) + \mathbf{L}_u(t_k)^T Q_{uu} \mathbf{l}_u(t_k) \\
V_{xx}(t_k) &= Q_{xx} + \mathbf{L}_u(t_k)^T Q_{uu} \mathbf{L}_u(t_k) + \mathbf{L}_u(t_k)^T Q_{ux} + Q_{xu} \mathbf{L}_u(t_k)
\end{aligned}$$

Further we can express these simply as functions of state-action value function by substituting in $\mathbf{l}_u(t_k)$ and $\mathbf{L}_u(t_k)$:

We do this for the value function:

$$\begin{aligned}
 V(t_k) &= Q_0 + Q_u^T \left(-Q_{uu}^{-1} Q_u \right) + \frac{1}{2} \left(-Q_{uu}^{-1} Q_u \right)^T Q_{uu} \left(-Q_{uu}^{-1} Q_u \right) \\
 &= Q_0 - Q_u^T Q_{uu}^{-1} Q_u + \frac{1}{2} Q_u^T Q_{uu}^{-1} \underbrace{Q_{uu} Q_{uu}^{-1}}_I Q_u \\
 &= Q_0 - Q_u^T Q_{uu}^{-1} Q_u + \frac{1}{2} Q_u^T Q_{uu}^{-1} Q_u \\
 &= Q_0 - \frac{1}{2} Q_u^T Q_{uu}^{-1} Q_u
 \end{aligned}$$

Then, we do this for the gradient of the value function:

$$\begin{aligned}
 V_x(t_k) &= Q_x + Q_u^T \left(-Q_{uu}^{-1} Q_{ux} \right) + Q_{xu} \left(-Q_{uu}^{-1} Q_u \right) + \left(-Q_{uu}^{-1} Q_{ux} \right)^T Q_{uu} \left(-Q_{uu}^{-1} Q_u \right) \\
 &= Q_x - Q_u^T Q_{uu}^{-1} Q_{ux} - Q_{xu} Q_{uu}^{-1} Q_u + Q_{ux}^T Q_{uu}^{-1} \underbrace{Q_{uu} Q_{uu}^{-1}}_I Q_u \\
 &= Q_x - Q_u^T Q_{uu}^{-1} Q_{ux} - Q_{xu} Q_{uu}^{-1} Q_u + Q_{ux}^T Q_{uu}^{-1} Q_{ux} \\
 &= Q_x - Q_{xu} Q_{uu}^{-1} Q_u
 \end{aligned}$$

Finally, we do this for the Hessian of the value function:

$$\begin{aligned}
 V_{xx}(t_k) &= Q_{xx} + \mathbf{L}_u(t_k)^T Q_{uu} \mathbf{L}_u(t_k) + \mathbf{L}_u(t_k)^T Q_{ux} + Q_{xu} \mathbf{L}_u(t_k) \\
 &= Q_{xx} + \left(-Q_{uu}^{-1} Q_{ux} \right)^T Q_{uu} \left(-Q_{uu}^{-1} Q_{ux} \right) + \left(-Q_{uu}^{-1} Q_{ux} \right)^T Q_{ux} + Q_{xu} \left(-Q_{uu}^{-1} Q_{ux} \right) \\
 &= Q_{xx} + Q_{ux}^T Q_{uu}^{-1} \underbrace{Q_{uu} Q_{uu}^{-1}}_I Q_{ux} - Q_{ux}^T Q_{uu}^{-1} Q_{ux} - Q_{xu} Q_{uu}^{-1} Q_{ux} \\
 &= Q_{xx} + Q_{ux}^T Q_{uu}^{-1} Q_{ux} - Q_{ux}^T Q_{uu}^{-1} Q_{ux} - Q_{xu} Q_{uu}^{-1} Q_{ux} \\
 &= Q_{xx} - Q_{xu} Q_{uu}^{-1} Q_{ux}
 \end{aligned}$$

In summary, the backward equations can be expressed as functions of the state-action value function as:

$$\begin{aligned}
 V(t_k) &= Q_0 - \frac{1}{2} Q_u^T Q_{uu}^{-1} Q_u \\
 V_x(t_k) &= Q_x - Q_{xu} Q_{uu}^{-1} Q_u \\
 V_{xx}(t_k) &= Q_{xx} - Q_{xu} Q_{uu}^{-1} Q_{ux}
 \end{aligned}$$

Part 2.

- 2.1) See MATLAB code implementation.
- 2.2) See MATLAB code implementation.
- 2.3) In this problem we are asked to test the robustness of our DDP policy against stochastic forces that act as disturbances in our dynamics.

$$I\ddot{\theta} + b\dot{\theta} + mgl\sin(\theta) = u + f_{stochastic}$$

We do this by running the DDP optimization until it converges. This results in an optimal trajectory $\mathbf{x}^*(t_k)$, optimal control input $\mathbf{u}^*(t_k)$, and optimal feedback gains $\mathbf{L}^*(t_k)$. We then initialize a new trajectory and propagate the dynamics forward using the following controller: $\mathbf{u}(t_k) = \mathbf{u}^*(t_k) + \mathbf{L}^*(t_k)(\mathbf{x}(t_k) - \mathbf{x}^*(t_k))$

This process is repeated for many trajectories through the following algorithm:

Algorithm 1: DDP Robustness Test Against Stochastic Forces

```

initialize:  $\mathbf{x}^*(t_k), \mathbf{u}^*(t_k), \mathbf{L}^*(t_k)$ ;
initialize:  $\mathbf{H}$ : time horizon,  $dt$ : time step,  $num\_traj$ : number of trajectories;
initialize:  $\mathbf{F}_x(\mathbf{x}(t_k)), \mathbf{F}_u(\mathbf{u}(t_k))$ : dynamics;
initialize:  $\mathbf{x}(t_k) = \mathbf{0}$ ;
initialize:  $\sigma$ : noise level;
for  $i = 1$  to  $num\_traj$  do
    for  $k = 1$  to  $\mathbf{H} - 1$  do
         $R = \text{randn}$ ;
         $\mathbf{u}(t_k) = \mathbf{u}^*(t_k) + \mathbf{L}^*(t_k)(\mathbf{x}(t_k) - \mathbf{x}^*(t_k))$ ;
         $\mathbf{x}(t_{k+1}) = \mathbf{F}_x(\mathbf{x}(t_k))dt + \mathbf{F}_u(\mathbf{u}(t_k))dt + \mathbf{F}_u(\mathbf{u}(t_k))R\sigma\sqrt{dt}$ 
    end
end
end

```

The following plot shows the results of DDP after optimization:

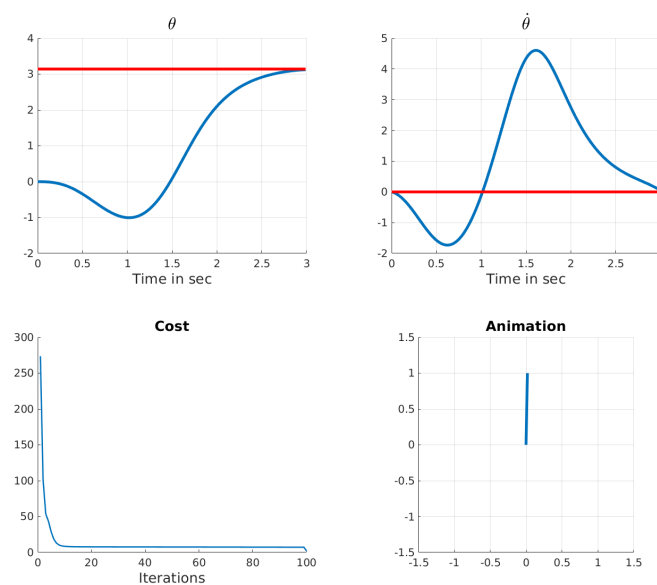


Figure 1: DDP optimal results.

The following plot shows the results of running the Robustness Test algorithm for 20 trajectories using the optimal DDP trajectory, controller, and feedback gains:

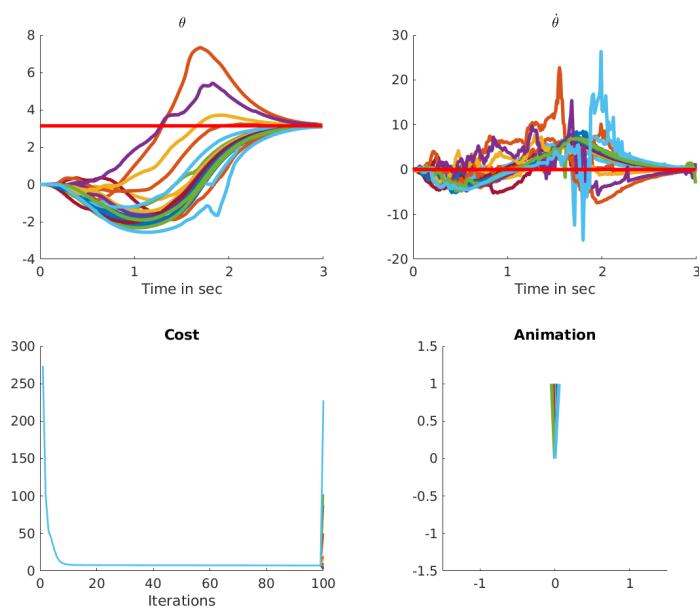


Figure 2: Robustness Test results.