Aalto University

LQR from HJB

CHEM-E7225 (was E7195), 2023

Francesco Corona (¬_¬)

Chemical and Metallurgical Engineering
School of Chemical Engineering

The LQR from the HJB

Consider a linear time-varying dynamical system, with initial condition $x(t_0) = \overline{x}_0$,

$$\dot{x}(t) = \underbrace{A(t)x(t) + B(t)u(t)}_{f(x(t),u(t),t)}$$

The cost to transfer the state $x(t_0)$ to $x(T)$ using control $u(t)$ with $t \in [t_0, t_f]$,

$$\underbrace{\frac{1}{2}x(t_f)^T Q_f x(t_f)}_{E(x(t_f))} + \frac{1}{2}\int_{t_0}^{t_f} \underbrace{x(t)^T Q(t)x(t) + u^T(t)R(t)u(t)}_{2L(x(t),u(t),t)}\, dt$$

- $\rightsquigarrow$ $Q(t) = Q^T(t) \succeq 0$
- $\rightsquigarrow$ $R(t) = R^T(t) \succ 0$
- $\rightsquigarrow$ $Q_f = Q_f^T \succeq 0$

The quadratic cost is very reasonable, since both Q and R are positive (semi)definite matrices, both the size of the state vector and the size of the control vector are penalised

- Matrices Q and R retain their relative relevance

We are interested in the optimal control $u^*(t)$, for all $t \in [t_0, t_f]$

- $\rightsquigarrow$ From the Hamilton-Jacobi-Bellman equation, we have

$$u^*(t) = u^*\left(x(t), \frac{\partial J^*(x(t))}{\partial x}\right)$$

The LQR from the HJB (cont.)

We defined the Hamilton-Jacobi-Bellman equation, as the partial differential equation

$$\frac{\partial J^*\left(x^*(t)\right)}{\partial t} + \underbrace{H\left(x^*(t), u^*(t), \frac{\partial J^*\left(x^*(t)\right)}{\partial x^*}\right)}_{\text{Optimal value of the Hamiltonian}} = 0$$

It contains the partial derivatives of the value function with respect to state and time

The HJB PDE is integrated backwards, from the boundary condition

- The terminal stage-cost

$$J^*\left(x^*(t_f), t_f\right) = E\left(x^*(t_f)\right)$$

The terminal cost does not appear in the HJB PDE itself

Solving the HJB equation analytically is a challenging task, even for simple problems

The solution of the HJB equation is the value function,

$$J^*\left(x(t), t\right) \quad x(t) \in \mathcal{X} \;\; t \in [0, T]$$

The LQR from the HJB (cont.)

$$\frac{\partial J^* (x^*(t))}{\partial t} + \underbrace{H\left(x^*(t), u^*(t), \frac{\partial J^* (x^*(t))}{\partial x^*}\right)}_{\text{Optimal value of the Hamiltonian}} = 0$$

In the Hamilton-Jacobi-Bellman equation, we defined the Hamiltonian,

$$H\left(x(t), u(t), \frac{\partial J^* (x(t))}{\partial x}\right) = L\left(x(t), u(t)\right) + \left[\frac{\partial J^* (x(t))}{\partial x(t)}\right]^T f\left(x(t), u(t), t\right)$$

For linear time-varying systems in continuous-time and quadratic costs,

$$H\left(x(t), u(t), \frac{\partial J^* (x(t))}{\partial x(t)}\right) = \underbrace{\frac{1}{2}\left(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\right)}_{L(x(t),u(t),t)}$$

$$+ \left[\frac{\partial J^* (x(t))}{\partial x(t)}\right]^T \underbrace{\left(A(t)x(t) + B(t)u(t)\right)}_{f(x(t),u(t),t)}$$

The LQR from the HJB (cont.)

$$H\left(x(t), u(t), \frac{\partial J^*(x(t))}{\partial x(t)}\right) = \frac{1}{2}\left(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\right) + \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T (A(t)x(t) + B(t)u(t))$$

The optimal value of the Hamiltonian is obtained from first-order optimality conditions,

$$\frac{\partial H\left(x(t), u(t), \frac{\partial J^*(x(t))}{\partial u}\right)}{\partial u} = 0$$

Differentiating the Hamiltonian with respect to $u(t)$, we get

$$R(t)u(t) + B^T(t)\left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T = 0$$

The LQR from the HJB (cont.)

$$R(t)u(t) + B^T(t)\frac{\partial J^*(x(t))}{\partial x(t)} = 0$$

The gradient of the Hamiltonian witch respect to $u(\cdot)$ must vanish along the trajectory

From the first-order optimality conditions, we solve for the optimal control and get

$$u^*(t) = -R^{-1}(t)B^T(t)\frac{\partial J^*(x(t))}{\partial x(t)}$$

- We used the assumption that $R(t)$ is invertible

The LQR from the HJB (cont.)

$$H\left(x(t), u(t), \frac{\partial J^*(x(t))}{\partial x(t)}\right) = \frac{1}{2}\Big(x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\Big)$$
$$+ \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T (A(t)x(t) + B(t)u(t))$$

We get the optimal value of the Hamiltonian, by substituting the optimal control $u^*(t)$

$$u^*(t) = -R^{-1}(t)B^T(t)\frac{\partial J^*(x(t))}{\partial x(t)}$$

We get,

$$H(x(t), u(t), \frac{\partial J^*(x(t))}{\partial x(t)}) = \frac{1}{2}x^T(t)Q(t)x(t)$$
$$+ \frac{1}{2}\left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T B(t)R^{-1}(t)R(t)R^{-1}(t)B^T(t)\frac{\partial J^*(x(t))}{\partial x(t)}$$
$$+ \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right] A(t)x(t)$$
$$- \left[\frac{\partial J^*(x(t))}{\partial x(t)}\right]^T B(t)R^{-1}B^T(t)\frac{\partial J^*(x(t))}{\partial x(t)}$$

The LQR from the HJB

After grouping terms and rearranging, we get the optimal value of the Hamiltonian

$$
\begin{aligned}
H\left(x(t), u(t), \frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right) = \frac{1}{2}x^T(t)Q(t)x(t) \\
- \frac{1}{2}\left[\frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right]^T B(t)R^{-1}B^T(t)\frac{\partial J^*\left(x(t)\right)}{+}\partial x(t) \\
+ \left[\frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right]^T A(t)x(t)
\end{aligned}
$$

The LQR from the HJB (cont.)

$$\frac{\partial J^*\left(x^*(t)\right)}{\partial t} + \underbrace{H\left(x^*(t), u^*(t), \frac{\partial J^*\left(x^*(t)\right)}{\partial x^*}\right)}_{\text{Optimal value of the Hamiltonian}} = 0$$

Given the optimal value of the Hamiltonian, we can re-write the HJB equation

$$\frac{\partial J^*\left(x^*(t)\right)}{\partial t} = -\frac{1}{2}x^T(t)Q(t)x(t) + \frac{1}{2}\left[\frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right]^T B(t)R^{-1}B^T(t)\frac{\partial J^*\left(x(t)\right)}{\partial x(t)} - \left[\frac{\partial J^*\left(x(t)\right)}{\partial x(t)}\right]^T A(t)x(t)$$

The boundary condition is given by the terminal stage-cost,

$$J^*\left(x^*(t_f), t_f\right) = \frac{1}{2}x^T(t_f)Q_f x^*(t_f)$$

The LQR from the HJB (cont.)

Assume that the Hamilton-Jacobi-Bellman has a quadratic solution in the state,

$$J^*\left(x(t), t\right) = \frac{1}{2} x^T(t) P(t) x(t)$$

The candidate solution mimics the quadratic form of the boundary condition,

$$J^*\left(x^*(t_f), t_f\right) = \frac{1}{2} x^T(t_f) Q_f x^*(t_f)$$

- $P(t) = P^T(t) \succeq 0$, as the cost must be non-negative

---

By taking the partial derivative of the candidate solution with respect to time, we get

$$\frac{\partial J^*\left(x(t)\right)}{\partial t} = \frac{1}{2} x(t) \dot{P} x(t)$$

Similarly, by taking the partial derivative with respect to the state we get

$$\frac{\partial J^*\left(x(t)\right)}{\partial x} = P(t) x(t)$$

The LQR from the HJB (cont.)

After substituting the partial derivatives in the Hamilton-Jacobi-Bellman equation,

$$\frac{1}{2}\underbrace{x^T(t)\dot{P}x(t)}_{\text{quadratic in } x(t)} + \frac{1}{2}\underbrace{x^T(t)Q(t)x(t)}_{\text{quadratic in } x(t)} - \frac{1}{2}\underbrace{x^T(t)P(t)B(t)R^{-1}(t)B^T(t)P(t)x(t)}_{\text{quadratic in } x(t)} + \underbrace{x^T(t)P(t)A(t)x(t)}_{\text{quadratic in } x(t)} = 0$$

- $\dot{P}(t)$ is symmetric
- $Q(t)$ is symmetric
- $P(t)B(t)R^{-1}(t)B^T(t)P(t)$ is symmetric
- Matrix $P(t)A(t)$ is not necessarily symmetric

The LQR from the HJB (cont.)

For any (not necessarily symmetric) state matrix A, we have

$$A = \underbrace{A_1}_{\text{symmetric}} + \underbrace{A_2}_{\text{skew-symmetric}}$$

Then, we can write

$$A_1 = \frac{A + A^T}{2}$$
$$A_2 = \frac{A - A^T}{2}$$

We re-write $P(t)A(t)$ in $x^T(t)P(t)A(t)x(t)$

$$P(t)A(t) = \frac{1}{2}\underbrace{\Big(P(t)A(t) + (P(t)A(t))^T\Big)}_{\text{symmetric}} + \frac{1}{2}\underbrace{\Big(P(t)A(t) - (P(t)A(t))^T\Big)}_{\text{skew-symmetric}}$$

In the quadratic form, the skew-symmetric part will vanish

The LQR from the HJB (cont.)

After substituting $P(t)A(t)$ with $\frac{1}{2}\Big(P(t)A(t) + (P(t)A(t))^T\Big)$ in the HJB, we get

$$\frac{1}{2}x^T(t)\dot{P}(t)x(t) + \frac{1}{2}x^T(t)Q(t)x(t) - \frac{1}{2}x^T(t)P(t)B(t)R^{-1}(t)B^T(t)P(t)x(t) + \frac{1}{2}x^T(t)P(t)A(t)x(t) + \frac{1}{2}x^T(t)A^T(t)P(t)x(t) = 0$$

That is, we have

$$\frac{1}{2}x^T(t)\Big(\dot{P}(t) + Q(t) - P(t)B(t)R^{-1}(t)B^T(t)P(t) + P(t)A(t) + A^T(t)P(t)\Big)x(t) = 0$$

As a result, for any $x(t)$ we get the matrix ordinary differential equation for $P(t)$

$$\dot{P}(t) + Q(t) - P(t)B(t)R^{-1}(t)B^T(t)P(t)A(t) + P(t)A(t) + A^T(t)P(t) = 0$$

The LQR from the HJB (cont.)

The matrix differential equation is the **matrix differential Riccati equation**

$$\dot{P}(t) = -P(t)A(t) - A^T(t)P(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t) - Q(t)$$

The boundary condition $P(t_f)$ is the terminal state weight-matrix Q_f

---

Once matrix $P(t)$ is determined along the trajectory, we get the optimal control

$$\begin{aligned} u(t) &= -R^{-1}(t)B^T(t)J^*\left(x(t), t\right) \\ &= -R^{-1}(t)B^T(t)P(t)x(t) \\ &= -K(t)x(t) \end{aligned}$$

The optimal control is given in linear state feedback form

- The time-varying feedback gain,

$$K(t) = R^{-1}(t)B^T(t)P(t)$$

- (Also for LTI systems)

A remarkable conclusion, though we did not prove the global optimality of the control