ECE 602: LUMPED LINEAR SYSTEMS

Professor Jianghai Hu

Linear Quadratic Regulation: Solution Algorithm

Back to LQR Problem

A discrete-time LTI system

$$x[k+1] = Ax[k] + Bu[k], \quad x[0] = x_0$$

Problem: Given a time horizon $k \in \{0, 1, \ldots, N\}$, find the optimal input sequence $U = \{u[0], \ldots, u[N-1]\}$ that minimizes the cost function

$$J(U) = \sum_{k=0}^{N-1} \underbrace{\left(x[k]^T Q x[k] + u[k]^T R u[k]\right)}_{\text{running cost}} + \underbrace{x[N]^T Q_f x[N]}_{\text{terminal cost}}.$$

Can we apply dynamic programming method to LQR problem?

Value Function of LQR Problem

Value function (cost-to-go) at time $t \in \{0, 1, \ldots, N\}$ and state $x \in \mathbb{R}^n$ is

$$V_t(x) = \min_{u[t],\ldots,u[N-1]} \sum_{k=t}^{N-1} \left(x[k]^T Q x[k] + u[k]^T R u[k] \right) + x[N]^T Q_f x[N]$$

with the initial condition $x[t] = x$

- Optimal cost over the time horizon $\{t, t+1, \ldots, N\}$, starting from $x[t] = x$
- $V_0(x)$ is the optimal cost of the original LQR problem

Preview of Dynamic Programming Solution

- Value function at any time t is quadratic: $V_t(x) = x^T P_t x$
- P_t can be obtained recursively from P_{t+1}

Solution algorithm:

1. Start from $P_N = Q_f$ at time $t = N$;
2. For $t = N-1, \cdots, 1, 0$, compute P_t from P_{t+1}
3. For $t = 0, 1, \ldots, N-1$, compute $u^*[t]$ using P_{t+1}

Bellman Equation

Optimality principle: to achieve optimal cost-to-go $V_t(x)$ from $x[t] = x$, cost-to-go from next state $x[t+1]$ should be optimal

- $V_t(x)$ is the optimal cost-to-go from current position $x[t] = x$
- Assume the control adopted at time t is $u[t] = v$. Then cost over time horizon $\{t, t+1, \ldots, N\}$ can be decomposed into
  - Current running cost: $x^T Qx + v^T Rv$
  - Cost-to-go from next state $x[t+1] = Ax + Bv$ over $\{t+1, \ldots, N\}$

(Hamilton-Jacobi-)Bellman equation:

$$V_t(x) = \min_{u[t]=v} \left[x^T Qx + v^T Rv + V_{t+1}(\underbrace{Ax + Bv}_{x[t+1]}) \right]$$

$t = N$ Case

Value function at time N is quadratic:

$$V_N(x) = x^T P_N x, \quad \forall x \in \mathbb{R}^n, \quad \text{where } P_N = Q_f$$

$t = N - 1$ Case

Value function at time $N-1$ is also quadratic:

$$
\begin{aligned}
V_{N-1}(x) &= \min_v \left[x^T Q x + v^T R v + V_N(Ax + Bv) \right] \\
&= \min_v \begin{bmatrix} x \\ v \end{bmatrix}^T \begin{bmatrix} Q + A^T P_N A & A^T P_N B \\ B^T P_N A & R + B^T P_N B \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} \\
&= x^T \underbrace{\left(Q + A^T P_N A - A^T P_N B (R + B^T P_N B)^{-1} B^T P_N A \right)}_{P_{N-1}} x
\end{aligned}
$$

with the optimal control $v^* = -\underbrace{(R + B^T P_N B)^{-1} B^T P_N A}_{K_{N-1}} x$

- Optimal control $u^*[N-1] = -K_{N-1} \cdot x[N-1]$ at time $N-1$ is a linear state feedback controller with gain determined from P_N (not P_{N-1}!)

Schur Complement

Given $X = X^T \in \mathbb{R}^{m \times m}$, $Y = Y^T \in \mathbb{R}^{n \times n}$ ($Y \succ 0$), $Z \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^m$

$$\min_{y \in \mathbb{R}^n} \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} X & Z \\ Z^T & Y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^T (X - Z Y^{-1} Z^T) x$$

- $X - ZY^{-1}Z^T$ is called the Schur complement of Y in $\begin{bmatrix} X & Z \\ Z^T & Y \end{bmatrix}$
- Minimum is achieved at $y^* = -Y^{-1} Z^T x$

Fact (Schur Complement Lemma)

$\begin{bmatrix} X & Z \\ Z^T & Y \end{bmatrix} \succ 0$ *if and only if* $Y \succ 0$ *and* $X - ZY^{-1}Z^T \succ 0$

General Case

Suppose value function at time $t+1$ is quadratic: $V_{t+1}(x) = x^T P_{t+1} x$

- Value function at time t is also quadratic: $V_t(x) = x^T P_t x$
- P_t obtained from P_{t+1} according to the **Riccati recursion**:

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

- Optimal control at time t is a linear state feedback controller:

$$u^*[t] = -\underbrace{(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A}_{\textbf{Kalman gain } K_t} x$$

whose gain is determined from P_{t+1} (not P_t!)

LQR Solution Algorithm

Set $P_N = Q_f$
for $t = N-1, N-2, \ldots, 0$ **do**
Compute the value functions backward in time:

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

end for
Return $V_0(x_0)$ as the optimal cost

Set $x^*[0] = x_0$
for $t = 0, 1, \ldots, N-1$ **do**
Recover the optimal control and state trajectory forward in time:

$$u^*[t] = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A x^*[t]$$
$$x^*[t+1] = A x^*[t] + B u^*[t]$$

end for
Return u^* and x^* as the optimal control and state sequences

Remarks

- Value function at any time is quadratic (easy numeric representation)
- Optimal controls are linear state feedback with time-varying gains
- Yields the optimal solutions for all initial conditions x_0 and all intial times $t_0 \in \{0, 1, \ldots, N\}$ simultaneously
- Easily extended to time-varying dynamics and costs cases