ECE276B: Planning & Learning in Robotics
Lecture 14: Linear Quadratic Control

Lecturer:
Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:
Tianyu Wang: tiw161@eng.ucsd.edu
Yongxi Lu: yol070@eng.ucsd.edu

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Globally Optimal Closed-Loop Control

- **Deterministic finite-horizon continuous-time optimal control**:

$$\begin{aligned}
\min_{\pi \in PC^0([t_0,T],\mathcal{U})} \quad & J^\pi(t_0, x_0) := \int_{t_0}^{T} g(x(t), \pi(t, x(t)))dt + g_T(x(T)) \\
\text{s.t.} \quad & \dot{x}(t) = f(x(t), u(t)), \;\; x(t_0) = x_0 \\
& x(t) \in \mathcal{X}, \; \pi(t, x(t)) \in \mathcal{U}
\end{aligned}$$

- **Hamiltonian**: $H(x, u, p) := g(x, u) + p^T f(x, u)$

HJB PDE: Sufficient Conditions for Optimality

If $V(t, x)$ satisfies the HJB PDE:

$$-\frac{\partial V(t,x)}{\partial t} = \min_{u \in \mathcal{U}} H(x(t), u, \nabla_x V(t, \cdot)), \;\; V(T, x) = g_T(x), \;\; \forall x \in \mathcal{X}, t \in [t_0, T]$$

then it is the optimal cost-to-go and the policy $\pi(t, x)$ that attains the minimum is an optimal policy.

Locally Optimal Open-Loop Control

- **Deterministic finite-horizon continuous-time optimal control**:

$$
\begin{aligned}
\min_{\pi \in PC^0([t_0,T],\mathcal{U})} \quad & J^\pi(t_0, x_0) := \int_{t_0}^{T} g(x(t), \pi(t, x(t)))dt + g_T(x(T)) \\
\text{s.t.} \quad & \dot{x}(t) = f(x(t), u(t)), \;\; x(t_0) = x_0 \\
& x(t) \in \mathcal{X}, \; \pi(t, x(t)) \in \mathcal{U}
\end{aligned}
$$

- **Hamiltonian**: $H(x, u, p) := g(x, u) + p^T f(x, u)$

PMP ODE: Necessary Conditions for Optimality

If $(x^*(t), u^*(t))$ for $t \in [t_0, T]$ is a trajectory from an optimal policy $\pi^*(t, x)$:

$$
\begin{aligned}
&\dot{x}^*(t) = f(x^*(t), u^*(t)), && x^*(t_0) = x_0 \\
&\dot{p}^*(t) = -\nabla_x g(x^*(t), u^*(t)) - [\nabla_x f(x^*(t), u^*(t))]^T p^*(t), && p^*(T) = \nabla_x g_T(x^*(T)) \\
&u^*(t) = \arg\min_{u \in \mathcal{U}} H(x^*(t), u, p^*(t)), && \forall t \in [t_0, T] \\
&H(x^*(t), u^*(t), p^*(t)) = constant, && \forall t \in [t_0, T]
\end{aligned}
$$

Tractable Problems

- Consider a deterministic finite-horizon problem with dynamics and cost:

$$\dot{x} = a(x) + Bu \qquad g(x,u) = q(x) + \frac{1}{2}u^T R u$$

- **Hamiltonian**:

$$H(x,u,p) = q(x) + \frac{1}{2}u^T R u + p^T a(x) + p^T B u$$

$$\nabla_u H(x,u,p) = Ru + B^T p \qquad \nabla_u^2 H(x,u,p) = R$$

- **HJB PDE**: obtains globally optimal cost-to-go and policy:

$$\begin{aligned}
\pi^*(t,x) &= \operatorname*{arg\,min}_{u \in \mathcal{U}} H(x,u,V_x(t,x)) = -R^{-1}B^T V_x(t,x), && t \in [t_0,T], x \in \mathcal{X} \\
V(T,x) &= g_T(x), && x \in \mathcal{X} \\
-V_t(t,x) &= q(x) + a^T V_x(t,x) - \frac{1}{2}V_x(t,x)^T B R^{-1} B^T V_x(t,x), && t \in [t_0,T], x \in \mathcal{X}
\end{aligned}$$

- **PMP**: both necessary and sufficient for local min as long as $R \succ 0$:

$$\begin{aligned}
u(t) &= \operatorname*{arg\,min}_{u \in \mathcal{U}} H(x,u,p) = -R^{-1}B^T p(t), && t \in [t_0,T] \\
\dot{x} &= a(x) - BR^{-1}B^T p, && x(0) = x_0 \\
\dot{p} &= -q_x(x)^T - a_x(x)^T p, && p(T) = \nabla_x g_T(x(T))
\end{aligned}$$

Example: Pendulum

$$\dot{x} = \begin{bmatrix} x_2 \\ k\sin(x_1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x(0) = x_0$$

$$a_x(x) = \begin{bmatrix} 0 & 1 \\ k\cos(x_1) & 0 \end{bmatrix}$$

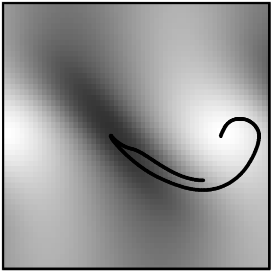
- Cost:

$$g(x,u) = 1 - e^{-2x_1^2} + \frac{r}{2}u^2 \text{ and } g_T(x) = 0$$

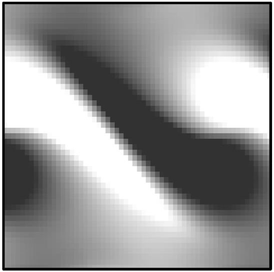
- PMP: locally optimal controller:

$$\begin{aligned}
u(t) &= -r^{-1}p_2(t), & t &\in [t_0, T] \\
\dot{x}_1 &= x_2, & x_1(0) &= 0 \\
\dot{x}_2 &= k\sin(x_1) - r^{-1}p_2, & x_2(0) &= 0 \\
\dot{p}_1 &= -4e^{-2x_1^2}x_1 - p_2, & p_1(T) &= 0 \\
\dot{p}_2 &= -k\cos(x_1)p_1, & p_2(T) &= 0
\end{aligned}$$

- Cost-to-go and trajectories:



- Optimal policy (from HJB):



Linear Quadratic Control

- The key assumptions that allowed us to minimize the Hamiltonian analytically were:
  - The system dynamics are linear in the control u
  - The stage-cost is quadratic in the control u

- Let us study the simplest such setting in which a deterministic time-invariant linear system needs to minimize a quadratic cost over a finite horizon:

$$\min_{\pi \in PC^0([t_0,T],\mathcal{U})} J^\pi(t_0,x_0) := \int_{t_0}^{T} \underbrace{\frac{1}{2}x(t)^T Q x(t) + \frac{1}{2}u(t)^T R u(t)}_{g(x(t),u(t))} dt + \underbrace{\frac{1}{2}x(T)^T Q_T x(T)}_{g_T(x(T))}$$

$$\begin{aligned} \text{s.t.} \quad & \dot{x} = Ax + Bu, \;\; x(t_0) = x_0 \\ & x(t) \in \mathbb{R}^n, \; u(t) = \pi(t, x(t)) \in \mathbb{R}^m \end{aligned}$$

where $Q = Q^T \succeq 0$, $Q_T = Q_T^T \succeq 0$, and $R = R^T \succ 0$

- This problem is called the **Linear Quadratic Regulator** (LQR)

LQR via the PMP

- Hamiltonian: $H(x,u,p) = \frac{1}{2}x^TQx + \frac{1}{2}u^TRu + p^TAx + p^TBu$
- Canonical equations with boundary conditions:

$$\dot{x} = \nabla_p H(x,u,p) = Ax + Bu, \qquad x(t_0) = x_0$$
$$\dot{p} = -\nabla_x H(x,u,p) = -Qx - A^Tp, \qquad p(T) = \nabla_x g_T(x(T)) = Q_T x(T)$$

- Minimum principle:

$$\nabla_u H(x,u,p)^T = Ru + B^Tp = 0 \quad \Rightarrow \quad u^*(t) = -R^{-1}B^Tp(t)$$
$$\nabla_u^2 H(x,u,p) = R \succ 0 \quad \Rightarrow \quad u^*(t) \text{ is a minimum}$$

- **Hamiltonian matrix**: the canonical equations can now be simplified to a linear time-invariant (LTI) system with two-point boundary conditions:

$$\begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}, \qquad \begin{matrix} x(t_0) = x_0 \\ p(T) = Q_T x(T) \end{matrix}$$

LQR via the PMP

- **Claim**: There exists a matrix $M(t) = M(t)^T \succeq 0$ such that $p(t) = M(t)x(t)$ for all $t \in [t_0, T]$

- We can solve the LTI system described by the Hamiltonian matrix backwards in time:

$$\begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = \underbrace{e^{\begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix}(t-T)}}_{\Phi(t,T)} \begin{bmatrix} x(T) \\ Q_T x(T) \end{bmatrix}$$

$$x(t) = (\Phi_{11}(t, T) + \Phi_{12}(t, T)Q_T)x(T)$$

$$p(t) = (\Phi_{21}(t, T) + \Phi_{22}(t, T)Q_T)x(T)$$

- It turns out that $D(t, T) := \Phi_{11}(t, T) + \Phi_{12}(t, T)Q_T$ is invertible for $t \in [t_0, T]$ and thus:

$$p(t) = \underbrace{(\Phi_{21}(t, T) + \Phi_{22}(t, T)Q_T)D^{-1}(t, T)}_{=:M(t)} x(t), \quad \forall t \in [t_0, T]$$

LQR via the PMP

- From $x(t_0) = D(t_0, T)x(T)$, we obtain an **open-loop control policy**:

$$u(t) = -R^{-1}B^T(\Phi_{21}(t, T) + \Phi_{22}(t, T)Q_T)D(t_0, T)^{-1}x_0$$

- From the claim that $p(t) = M(t)x(t)$, however, we can also obtain a **linear state feedback** control policy:

$$u(t) = -R^{-1}B^T M(t)x(t)$$

- We can obtain a better description of $M(t)$ by differentiating $p(t) = M(t)x(t)$ and using the canonical equations:

$$\begin{aligned} \dot{p}(t) &= \dot{M}(t)x(t) + M(t)\dot{x}(t) \\ -Qx(t) - A^T p(t) &= \dot{M}(t)x(t) + M(t)Ax(t) - M(t)BR^{-1}B^T p(t) \\ -\dot{M}(t)x(t) &= Qx(t) + A^T M(t)x(t) + M(t)Ax(t) - M(t)BR^{-1}B^T M(t)x(t) \end{aligned}$$

which needs to hold for all $x(t)$ and $t \in [t_0, T]$ and satisfy the boundary condition $p(T) = M(T)x(T) = Q_T x(T)$

LQR via the PMP (Summary)

- A unique candidate $u(t) = -R^{-1}B^T M(t)x(t)$ satsifies the necessary conditions of the PMP for optimality

- The candidate policy is linear in the state and the matrix $M(t)$ satisfies a quadratic **Riccati differential equation** (RDE):

$$-\dot{M}(t) = Q + A^T M(t) + M(t)A - M(t)BR^{-1}B^T M(t), \quad M(T) = Q_T$$

- Other tools (e.g., the HJB PDE) are needed to decide whether $u(t)$ is a globally optimal policy

LQR via the HJB PDE

- Hamiltonian: $H(x, u, p) = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru + p^T Ax + p^T Bu$
- HJB PDE:

$$\begin{aligned}
\pi^*(t,x) &= \arg\min_{u \in \mathcal{U}} H(x, u, V_x(t,x)) = -R^{-1}B^T V_x(t,x), && t \in [t_0, T], x \in \mathcal{X} \\
-V_t(t,x) &= \frac{1}{2}x^T Qx + x^T A^T V_x(t,x) - \frac{1}{2}V_x(t,x)^T BR^{-1}B^T V_x(t,x), && t \in [t_0, T], x \in \mathcal{X} \\
V(T,x) &= \frac{1}{2}x^T Q_T x
\end{aligned}$$

- Guess a solution to the HJB PDE based on the intuition from the PMP:

$$\begin{aligned}
\pi(t,x) &= -R^{-1}B^T M(t)x \\
V(t,x) &= \frac{1}{2}x^T M(t)x \\
V_t(t,x) &= \frac{1}{2}x^T \dot{M}(t)x \\
V_x(t,x) &= M(t)x
\end{aligned}$$

LQR via the HJB PDE

- Substituting the candidate $V(t, x)$ into the HJB PDE leads to the same **RDE** as before and we know that $M(t)$ satisfies it!

$$\frac{1}{2}x^T M(T)x = \frac{1}{2}x^T Q_T x$$
$$-\frac{1}{2}x^T \dot{M}(t)x = \frac{1}{2}x^T Qx + x^T A^T M(t)x - \frac{1}{2}x^T M(t)BR^{-1}B^T M(t)x, \ t \in [t_0, T], x \in \mathcal{X}$$

- **Conclusion**: Since $M(t)$ satisfies the RDE, $V(t, x) = x^T M(t)x$ is the unique solution to the HJB PDE and is the optimal cost-to-go for the linear quadratic problem with an associated optimal policy $\pi(t, x) = -R^{-1}B^T M(t)x$.
- **General Strategy for Continuous-time Optimal Control Problems**:
  1. Identify a candidate policy using the PMP
  2. Use the intuition from 1. to guess a candidate cost-to-go
  3. Verify that the candidate policy and cost-to-go satisfy the HJB PDE

Continuous-time Finite-horizon LQG

- **Linear Quadratic Gaussian** (LQG) regulation problem:

$$\begin{aligned} \min_{\pi \in PC^0([t_0,T],\mathcal{U})} J^\pi(t_0,x_0) &:= \frac{1}{2}\mathbb{E}\left\{\int_{t_0}^{T} e^{-\frac{t}{\gamma}} \begin{bmatrix} x^T(t) & u^T(t)\end{bmatrix} \begin{bmatrix} Q & P^T \\ P & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt + e^{-\frac{T}{\gamma}} x(T)^T Q_T x(T)\right\} \\ \text{s.t.} \quad & dx = (Ax + Bu)dt + C d\omega, \;\; x(t_0) = x_0 \\ & x(t) \in \mathcal{X}, \; u(t) = \pi(t, x(t)) \in \mathcal{U} \end{aligned}$$

- **Discount factor**: $\gamma \in [0, \infty]$
- **Optimal cost-to-go**: $J^*(t,x) = \frac{1}{2}x^T M(t) x + m(t)$
- **Optimal policy**: $\pi^*(t,x) = -R^{-1}(P + B^T M(t))x$
- **Riccati Equation**:

$$-\dot{M}(t) = Q + A^T M(t) + M(t)A - (P + B^T M(t))^T R^{-1}(P + B^T M(t)) - \frac{1}{\gamma}M(t), \qquad M(T) = Q_T$$

$$-\dot{m} = \frac{1}{2}\operatorname{tr}(CC^T M(t)) - \frac{1}{\gamma}m(t), \qquad m(T) = 0$$

- $M(t)$ is independent of the noise amplitude C, which implies that the optimal policy $\pi^*(t,x)$ is **the same for the stochastic (LQG) and deterministic (LQR) problems**!

Continuous-time Infinite-horizon LQG

- **Linear Quadratic Gaussian** (LQG) regulation problem:

$$\min_{\pi \in PC^0(\mathcal{X},\mathcal{U})} J^\pi(x_0) := \frac{1}{2}\mathbb{E}\left\{\int_{t_0}^{\infty} e^{-\frac{t}{\gamma}} \begin{bmatrix} x^T(t) & u^T(t) \end{bmatrix} \begin{bmatrix} Q & P^T \\ P & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \right\}$$

$$\text{s.t.} \quad dx = (Ax + Bu)dt + Cd\omega, \;\; x(t_0) = x_0$$

$$x(t) \in \mathcal{X}, \; u(t) = \pi(x(t)) \in \mathcal{U}$$

- **Discount factor**: $\gamma \in [0, \infty)$
- **Optimal cost-to-go**: $J^*(x) = \frac{1}{2}x^T M x + m$
- **Optimal policy**: $\pi^*(x) = -R^{-1}(P + B^T M)x$
- **Riccati Equation** ('care' in Matlab):

$$\frac{1}{\gamma}M = Q + A^T M + MA - (P + B^T M)^T R^{-1}(P + B^T M)$$

$$m = \frac{\gamma}{2}\operatorname{tr}(CC^T M)$$

- M is independent of the noise amplitude C, which implies that the optimal policy $\pi^*(x)$ is **the same for LQG and LQR**!

Discrete-time Finite-horizon LQG

- **Linear Quadratic Gaussian** (LQG) regulation problem:

$$\min_{\pi_{0:T-1}} \quad J_0^{\pi}(x) := \frac{1}{2}\mathbb{E}\left\{\sum_{t=0}^{T-1} \gamma^t \left(x_t^T Q x_t + 2u_t^T P x_t + u_t^T R u_t\right) + \gamma^T x_T Q_T x_T\right\}$$

$$\text{s.t.} \quad x_{t+1} = Ax_t + Bu_t + Cw_t, \;\; x_0 = x, \; w_t \sim \mathcal{N}(0, I)$$

$$x(t) \in \mathcal{X}, \; u_t = \pi_t(x_t) \in \mathcal{U}$$

- **Discount factor**: $\gamma \in [0, 1]$
- **Optimal cost-to-go**: $J_t^*(x) = \frac{1}{2}x^T M_t x + m_t$
- **Optimal policy**: $\pi_t^*(x) = -(R + \gamma B^T M_{t+1} B)^{-1}(P + \gamma B^T M_{t+1} A)x$
- **Riccati Equation**:

$$M_t = Q + \gamma A^T M_{t+1} A - (P + \gamma B^T M_{t+1} A)^T (R + \gamma B^T M_{t+1} B)^{-1} (P + \gamma B^T M_{t+1} A), \quad M_T = Q_T$$

$$m_t = \gamma m_{t+1} + \gamma \frac{1}{2} \operatorname{tr}(CC^T M_{t+1}), \qquad m_T = 0$$

- M_t is independent of the noise amplitude C, which implies that the optimal policy $\pi_t^*(x)$ is **the same for LQG and LQR**!

Discrete-time Infinite-horizon LQG

- **Linear Quadratic Gaussian** (LQG) regulation problem:

$$\min_{\pi} \; J^{\pi}(x) := \frac{1}{2}\mathbb{E}\left\{\sum_{t=0}^{\infty}\gamma^{t}\left(x_t^T Q x_t + 2u_t^T P x_t + u_t^T R u_t\right)\right\}$$

$$\text{s.t.} \quad x_{t+1} = Ax_t + Bu_t + Cw_t, \;\; x_{t_0} = x_0, \;\; w_t \sim \mathcal{N}(0, I)$$

$$x(t) \in \mathcal{X}, \;\; u_t = \pi(x_t) \in \mathcal{U}$$

- **Discount factor**: $\gamma \in [0, 1)$
- **Optimal cost-to-go**: $J^*(x) = \frac{1}{2}x^T M x + m$
- **Optimal policy**: $\pi^*(x) = -(R + \gamma B^T M B)^{-1}(P + \gamma B^T M A)x$
- **Riccati Equation** ('dare' in Matlab):

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

$$m = \frac{\gamma}{2(1-\gamma)}\operatorname{tr}(CC^T M)$$

- M is independent of the noise amplitude C, which implies that the optimal policy $\pi^*(x)$ is **the same for LQG and LQR**!

Relation between Continuous- and Discrete-time LQR

- The continuous-time system:

$$\dot{x} = Ax + Bu$$
$$g(x,u) = \frac{1}{2}x^T Qx + \frac{1}{2}u^T Ru$$

  can be discretized with time step τ:

$$x_{t+1} = (I + \tau A)x_t + \tau B u_t$$
$$\tau g(x,u) = \frac{\tau}{2}x^T Qx + \frac{\tau}{2}u^T Ru$$

- In the limit as $\tau \to 0$, the discrete-time Riccati equation reduces to the continuous one:

$$\begin{aligned}
M &= \tau Q + (I + \tau A)^T M (I + \tau A) \\
&\qquad - (I + \tau A)^T M \tau B (\tau R + \tau B^T M \tau B)^{-1} \tau B^T M (I + \tau A) \\
M &= \tau Q + M + \tau A^T M + \tau M A - \tau M B (R + \tau B^T M B)^{-1} B^T M + o(\tau^2) \\
0 &= Q + A^T M + MA - MB(R + \tau B^T M B)^{-1} B^T M + \frac{1}{\tau} o(\tau^2)
\end{aligned}$$

Encoding Goals as Quadratic Costs

- In the finite-horizon case, the matrices A, B, Q, R can be time-varying which is useful for specifying reference trajectories x_t^* and for approximating non-LQG problems

- The cost $\|x_t - x_t^*\|^2$ can be captured in the LQG formulation by modifying the state and cost as follows:

$$\tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix} \qquad \tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}, \text{etc.}$$

$$\frac{1}{2}\tilde{x}^T \tilde{Q}_t \tilde{x} = \frac{1}{2}\tilde{x}^T (D_t^T D_t)\tilde{x} \qquad D_t \tilde{x}_t := \begin{bmatrix} I & -x_t^* \end{bmatrix} \tilde{x}_t = x_t - x_t^*$$

- If the target/goal is stationary, we can instead include it in the state $\tilde{x}$ and use $D := \begin{bmatrix} I & -I \end{bmatrix}$. This has the advantage that the resulting policy is independent of x^* and can be used for any target x^*.