

ECE276B: Planning & Learning in Robotics

Lecture 17: Linear Quadratic Control

Instructor:

Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:

Thai Duong: tduong@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:**

$$\min_{\pi} V^{\pi}(0, \mathbf{x}_0) := \int_0^T \ell(\mathbf{x}(t), \pi(t, \mathbf{x}(t))) dt + q(\mathbf{x}(T))$$

$$\text{s.t. } \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0$$

$$\mathbf{x}(t) \in \mathcal{X}, \quad \pi(t, \mathbf{x}(t)) \in PC^0([0, T], \mathcal{U})$$

- ▶ **Hamiltonian:** $H(\mathbf{x}, \mathbf{u}, \mathbf{p}) := \ell(\mathbf{x}, \mathbf{u}) + \mathbf{p}^{\top} \mathbf{f}(\mathbf{x}, \mathbf{u})$

HJB PDE: Sufficient Conditions for Optimality

If $V(t, \mathbf{x})$ satisfies the HJB PDE:

$$V(T, \mathbf{x}) = q(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}$$

$$-\frac{\partial}{\partial t} V(t, \mathbf{x}) = \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}(t), \mathbf{u}, \nabla_{\mathbf{x}} V(t, \cdot)), \quad \forall \mathbf{x} \in \mathcal{X}, t \in [0, T]$$

then it is the optimal value function and the policy $\pi(t, \mathbf{x})$ that attains the minimum is an optimal policy.

Locally Optimal Open-Loop Control

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

$$\min_{\pi} V^{\pi}(0, \mathbf{x}_0) := \int_0^T \ell(\mathbf{x}(t), \pi(t, \mathbf{x}(t))) dt + q(\mathbf{x}(T))$$

$$\text{s.t. } \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0$$

$$\mathbf{x}(t) \in \mathcal{X}, \quad \pi(t, \mathbf{x}(t)) \in PC^0([0, T], \mathcal{U})$$

- **Hamiltonian:** $H(\mathbf{x}, \mathbf{u}, \mathbf{p}) := \ell(\mathbf{x}, \mathbf{u}) + \mathbf{p}^{\top} \mathbf{f}(\mathbf{x}, \mathbf{u})$

PMP ODE: Necessary Conditions for Optimality

If $(\mathbf{x}^*(t), \mathbf{u}^*(t))$ for $t \in [0, T]$ is a trajectory from an optimal policy $\pi^*(t, \mathbf{x})$, then it satisfies:

$$\dot{\mathbf{x}}^*(t) = \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t)),$$

$$\mathbf{x}^*(0) = \mathbf{x}_0$$

$$\dot{\mathbf{p}}^*(t) = -\nabla_{\mathbf{x}} \ell(\mathbf{x}^*(t), \mathbf{u}^*(t)) - [\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t))]^{\top} \mathbf{p}^*(t), \quad \mathbf{p}^*(T) = \nabla_{\mathbf{x}} q(\mathbf{x}^*(T))$$

$$\mathbf{u}^*(t) = \arg \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}^*(t), \mathbf{u}, \mathbf{p}^*(t)),$$

$$\forall t \in [0, T]$$

$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)) = \text{constant},$$

$$\forall t \in [0, T]$$

Tractable Problems

- ▶ Control-affine dynamics and quadratic-in-control cost:

$$\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}) + B(\mathbf{x})\mathbf{u} \quad \ell(\mathbf{x}, \mathbf{u}) = q(\mathbf{x}) + \frac{1}{2}\mathbf{u}^\top R(\mathbf{x})\mathbf{u} \quad R(\mathbf{x}) \succ 0$$

- ▶ **Hamiltonian:**

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

- ▶ **HJB PDE:** obtains the globally optimal value function and policy:

$$\pi^*(t, \mathbf{x}) = \arg \min_{\mathbf{u}} H(\mathbf{x}, \mathbf{u}, V_x(t, \mathbf{x})) = -R(\mathbf{x})^{-1} B(\mathbf{x})^\top V_x(t, \mathbf{x}),$$
$$V(T, \mathbf{x}) = q(\mathbf{x}),$$
$$-V_t(t, \mathbf{x}) = q(\mathbf{x}) + \mathbf{a}(\mathbf{x})^\top V_x(t, \mathbf{x}) - \frac{1}{2} V_x(t, \mathbf{x})^\top B(\mathbf{x}) R(\mathbf{x})^{-1} B(\mathbf{x})^\top V_x(t, \mathbf{x}).$$

Tractable Problems

- ▶ Control-affine dynamics and quadratic-in-control cost:

$$\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}) + B(\mathbf{x})\mathbf{u} \quad \ell(\mathbf{x}, \mathbf{u}) = q(\mathbf{x}) + \frac{1}{2}\mathbf{u}^\top R(\mathbf{x})\mathbf{u} \quad R(\mathbf{x}) \succ 0$$

- ▶ **Hamiltonian:**

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

- ▶ **PMP:** both necessary and sufficient for a local minimum:

$$\mathbf{u} = \arg \min_{\mathbf{u}} H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = -R(\mathbf{x})^{-1} B(\mathbf{x})^\top \mathbf{p},$$

$$\dot{\mathbf{x}} = \mathbf{a}(\mathbf{x}) - B(\mathbf{x})R^{-1}(\mathbf{x})B^\top(\mathbf{x})\mathbf{p},$$

$$\mathbf{x}(0) = \mathbf{x}_0$$

$$\dot{\mathbf{p}} = -q_x(\mathbf{x}) - a_x(\mathbf{x})^\top \mathbf{p},$$

$$\mathbf{p}(T) = q_x(\mathbf{x}(T))$$

Example: Pendulum

$$\dot{\mathbf{x}} = \begin{bmatrix} x_2 \\ k \sin(x_1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad \mathbf{x}(0) = \mathbf{x}_0$$

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

► Cost:

$$\ell(\mathbf{x}, u) = 1 - e^{-2x_1^2} + \frac{r}{2} u^2 \quad \text{and} \quad q(\mathbf{x}) = 0$$

► PMP locally optimal trajectories:

$$u(t) = -r^{-1} p_2(t), \quad t \in [0, T]$$

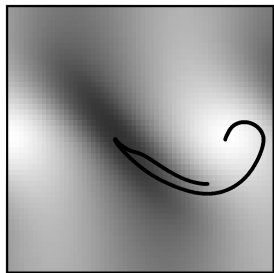
$$\dot{x}_1 = x_2, \quad x_1(0) = 0$$

$$\dot{x}_2 = k \sin(x_1) - r^{-1} p_2, \quad x_2(0) = 0$$

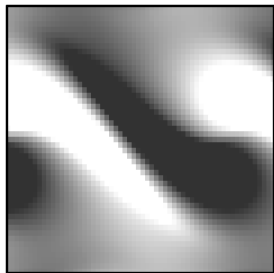
$$\dot{p}_1 = -4e^{-2x_1^2} x_1 - p_2, \quad p_1(T) = 0$$

$$\dot{p}_2 = -k \cos(x_1) p_1, \quad p_2(T) = 0$$

► Optimal value from HJB:



► Optimal policy from HJB:



Linear Quadratic Regulator

- ▶ Key assumptions that allowed minimizing the Hamiltonian analytically:
 - ▶ The system dynamics are linear in the control u
 - ▶ The stage-cost is quadratic in the control u
- ▶ **Linear Quadratic Regulator** (LQR): a deterministic time-invariant linear system needs to minimize a quadratic cost over a finite horizon:

$$\min_{\pi} V^{\pi}(0, \mathbf{x}_0) := \int_0^T \underbrace{\frac{1}{2} \mathbf{x}(t)^{\top} \mathbf{Q} \mathbf{x}(t) + \frac{1}{2} \mathbf{u}(t)^{\top} \mathbf{R} \mathbf{u}(t)}_{\ell(\mathbf{x}(t), \mathbf{u}(t))} dt + \underbrace{\frac{1}{2} \mathbf{x}(T)^{\top} \mathbf{Q} \mathbf{x}(T)}_{q(\mathbf{x}(T))}$$

$$\text{s.t. } \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad \mathbf{x}(0) = \mathbf{x}_0$$

$$\mathbf{x}(t) \in \mathbb{R}^n, \quad \mathbf{u}(t) = \pi(t, \mathbf{x}(t)) \in \mathbb{R}^m$$

where $\mathbf{Q} = \mathbf{Q}^{\top} \succeq 0$, $\mathbf{Q} = \mathbf{Q}^{\top} \succeq 0$, and $\mathbf{R} = \mathbf{R}^{\top} \succ 0$

LQR via the PMP

▶ Hamiltonian: $H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\top R\mathbf{u} + \mathbf{p}^\top A\mathbf{x} + \mathbf{p}^\top B\mathbf{u}$

▶ Canonical equations with boundary conditions:

$$\begin{aligned}\dot{\mathbf{x}} &= \nabla_{\mathbf{p}} H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = A\mathbf{x} + B\mathbf{u}, & \mathbf{x}(0) &= \mathbf{x}_0 \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{x}} H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = -Q\mathbf{x} - A^\top \mathbf{p}, & \mathbf{p}(T) &= Q\mathbf{x}(T)\end{aligned}$$

▶ PMP:

$$\begin{aligned}\nabla_{\mathbf{u}} H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = R\mathbf{u} + B^\top \mathbf{p} = 0 & \Rightarrow \mathbf{u}^*(t) = -R^{-1}B^\top \mathbf{p}(t) \\ \nabla_{\mathbf{u}}^2 H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = R \succ 0 & \Rightarrow \mathbf{u}^*(t) \text{ is a minimum}\end{aligned}$$

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

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^\top \\ -Q & -A^\top \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}, \quad \begin{aligned} \mathbf{x}(0) &= \mathbf{x}_0 \\ \mathbf{p}(T) &= Q\mathbf{x}(T) \end{aligned}$$

LQR via the PMP

- ▶ **Claim:** There exists a matrix $M(t) = M(t)^T \succeq 0$ such that $\mathbf{p}(t) = M(t)\mathbf{x}(t)$ for all $t \in [0, T]$
- ▶ We can solve the LTI system described by the Hamiltonian matrix backwards in time:

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

$$\mathbf{x}(t) = (\Phi_{11}(t, T) + \Phi_{12}(t, T)\mathbb{Q})\mathbf{x}(T)$$

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

- ▶ It turns out that $D(t, T) := \Phi_{11}(t, T) + \Phi_{12}(t, T)\mathbb{Q}$ is invertible for $t \in [0, T]$ and thus:

$$\mathbf{p}(t) = \underbrace{(\Phi_{21}(t, T) + \Phi_{22}(t, T)\mathbb{Q})D^{-1}(t, T)}_{=:M(t)}\mathbf{x}(t), \quad \forall t \in [0, T]$$

LQR via the PMP

- ▶ From $\mathbf{x}(0) = D(0, T)\mathbf{x}(T)$, we obtain an **open-loop control policy**:

$$\mathbf{u}(t) = -R^{-1}B^{\top}(\Phi_{21}(t, T) + \Phi_{22}(t, T)Q)D(0, T)^{-1}\mathbf{x}_0$$

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

$$\mathbf{u}(t) = -R^{-1}B^{\top}M(t)\mathbf{x}(t)$$

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

$$\begin{aligned}\dot{\mathbf{p}}(t) &= \dot{M}(t)\mathbf{x}(t) + M(t)\dot{\mathbf{x}}(t) \\ -Q\mathbf{x}(t) - A^{\top}\mathbf{p}(t) &= \dot{M}(t)\mathbf{x}(t) + M(t)A\mathbf{x}(t) - M(t)BR^{-1}B^{\top}\mathbf{p}(t) \\ -\dot{M}(t)\mathbf{x}(t) &= Q\mathbf{x}(t) + A^{\top}M(t)\mathbf{x}(t) + M(t)A\mathbf{x}(t) - M(t)BR^{-1}B^{\top}M(t)\mathbf{x}(t)\end{aligned}$$

which needs to hold for all $\mathbf{x}(t)$ and $t \in [0, T]$ and satisfy the boundary condition $\mathbf{p}(T) = M(T)\mathbf{x}(T) = Q\mathbf{x}(T)$

LQR via the PMP (Summary)

- ▶ A unique candidate $\mathbf{u}(t) = -R^{-1}B^T M(t)\mathbf{x}(t)$ satisfies 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):

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

- ▶ The HJB PDE is needed to decide whether $\mathbf{u}(t)$ is globally optimal

LQR via the HJB PDE

▶ Hamiltonian: $H(\mathbf{x}, \mathbf{u}, \mathbf{p}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\top R\mathbf{u} + \mathbf{p}^\top A\mathbf{x} + \mathbf{p}^\top B\mathbf{u}$

▶ HJB PDE:

$$\pi^*(t, \mathbf{x}) = \arg \min_{\mathbf{u} \in \mathcal{U}} H(\mathbf{x}, \mathbf{u}, V_x(t, \mathbf{x})) = -R^{-1}B^\top V_x(t, \mathbf{x}), \quad t \in [0, T], \mathbf{x} \in \mathcal{X}$$

$$-V_t(t, \mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \mathbf{x}^\top A^\top V_x(t, \mathbf{x}) - \frac{1}{2}V_x(t, \mathbf{x})^\top B R^{-1} B^\top V_x(t, \mathbf{x}), \quad t \in [0, T], \mathbf{x} \in \mathcal{X}$$

$$V(T, \mathbf{x}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x}$$

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

$$\pi(t, \mathbf{x}) = -R^{-1}B^\top M(t)\mathbf{x}$$

$$V(t, \mathbf{x}) = \frac{1}{2}\mathbf{x}^\top M(t)\mathbf{x}$$

$$V_t(t, \mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \dot{M}(t)\mathbf{x}$$

$$V_x(t, \mathbf{x}) = M(t)\mathbf{x}$$

LQR via the HJB PDE

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

$$\begin{aligned}\frac{1}{2}\mathbf{x}^\top M(T)\mathbf{x} &= \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} \\ -\frac{1}{2}\mathbf{x}^\top \dot{M}(t)\mathbf{x} &= \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \mathbf{x}^\top A^\top M(t)\mathbf{x} - \frac{1}{2}\mathbf{x}^\top M(t)BR^{-1}B^\top M(t)\mathbf{x}, \quad t \in [0, T], \mathbf{x} \in \mathcal{X}\end{aligned}$$

- ▶ **Conclusion:** Since $M(t)$ satisfies the RDE, $V(t, \mathbf{x}) = \mathbf{x}^\top M(t)\mathbf{x}$ is the unique solution to the HJB PDE and is the optimal value function for the linear quadratic problem with an associated optimal policy $\pi(t, \mathbf{x}) = -R^{-1}B^\top M(t)\mathbf{x}$.

- ▶ General strategy for continuous-time optimal control problems:
 1. Identify a candidate policy using the PMP
 2. Use intuition from 1. to guess a candidate value function
 3. Verify that the candidate policy and value function satisfy the HJB PDE

Continuous-time Finite-horizon LQG

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

$$\min_{\pi} V^{\pi}(0, \mathbf{x}_0) := \frac{1}{2} \mathbb{E} \left\{ \int_0^T e^{-\frac{t}{\gamma}} [\mathbf{x}^{\top}(t) \quad \mathbf{u}^{\top}(t)] \begin{bmatrix} Q & P^{\top} \\ P & R \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} dt + e^{-\frac{T}{\gamma}} \mathbf{x}(T)^{\top} Q \mathbf{x}(T) \right\}$$

$$\text{s.t. } \dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} + C\boldsymbol{\omega}, \quad \mathbf{x}(0) = \mathbf{x}_0 \\ \mathbf{x}(t) \in \mathbb{R}^n, \quad \mathbf{u}(t) = \pi(t, \mathbf{x}(t)) \in \mathbb{R}^m$$

- ▶ **Discount factor:** $\gamma \in [0, \infty]$

- ▶ **Optimal value:** $V^*(t, \mathbf{x}) = \frac{1}{2} \mathbf{x}^{\top} M(t) \mathbf{x} + m(t)$

- ▶ **Optimal policy:** $\pi^*(t, \mathbf{x}) = -R^{-1}(P + B^{\top} M(t)) \mathbf{x}$

- ▶ **Riccati Equation:**

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

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

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

Continuous-time Infinite-horizon LQG

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

$$\min_{\pi} V^{\pi}(\mathbf{x}_0) := \frac{1}{2} \mathbb{E} \left\{ \int_0^{\infty} e^{-\frac{t}{\gamma}} [\mathbf{x}^{\top}(t) \quad \mathbf{u}^{\top}(t)] \begin{bmatrix} Q & P^{\top} \\ P & R \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} dt \right\}$$

$$\text{s.t. } \dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} + C\boldsymbol{\omega}, \quad \mathbf{x}(0) = \mathbf{x}_0$$
$$\mathbf{x}(t) \in \mathbb{R}^n, \quad \mathbf{u}(t) = \pi(\mathbf{x}(t)) \in \mathbb{R}^m$$

- ▶ **Discount factor:** $\gamma \in [0, \infty)$
- ▶ **Optimal value:** $V^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^{\top} M \mathbf{x} + m$
- ▶ **Optimal policy:** $\pi^*(\mathbf{x}) = -R^{-1}(P + B^{\top} M)\mathbf{x}$
- ▶ **Riccati Equation** ('care' in Matlab):

$$\frac{1}{\gamma} M = Q + A^{\top} M + M A - (P + B^{\top} M)^{\top} R^{-1} (P + B^{\top} M)$$
$$m = \frac{\gamma}{2} \text{tr}(C C^{\top} M)$$

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

Discrete-time Linear Quadratic Control

Discrete-time Finite-horizon Linear Quadratic Regulator

- ▶ **Linear Quadratic Regulator** (LQR) problem:

$$\min_{\pi_{0:T-1}} V_0^\pi(\mathbf{x}) := \frac{1}{2} \left\{ \sum_{t=0}^{T-1} \left(\mathbf{x}_t^\top Q \mathbf{x}_t + \mathbf{u}_t^\top R \mathbf{u}_t \right) + \mathbf{x}_T^\top Q \mathbf{x}_T \right\}$$

$$\text{s.t. } \mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t, \quad \mathbf{x}_0 = \mathbf{x}$$
$$\mathbf{x}(t) \in \mathbb{R}^n, \quad \mathbf{u}_t = \pi_t(\mathbf{x}_t) \in \mathbb{R}^m$$

- ▶ Since this is a discrete-time finite-horizon problem, we can use Dynamic Programming
- ▶ At $t = T$, there are no control choices and the value function is quadratic in \mathbf{x} :

$$V_T^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top M_T \mathbf{x} := \frac{1}{2} \mathbf{x}^\top Q \mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

- ▶ Iterate backwards in time $t = T - 1, \dots, 0$:

$$V_t^*(\mathbf{x}) = \min_{\mathbf{u}} \left\{ \frac{1}{2} \left(\mathbf{x}^\top Q \mathbf{x} + \mathbf{u}^\top R \mathbf{u} \right) + V_{t+1}^*(A\mathbf{x} + B\mathbf{u}) \right\}$$

Discrete-time Finite-horizon Linear Quadratic Regulator

- ▶ At $t = T - 1$:

$$V_{T-1}^*(\mathbf{x}) = \min_{\mathbf{u}} \frac{1}{2} \left\{ \mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{u}^\top \mathbf{R} \mathbf{u} + (\mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u})^\top M_T (\mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u}) \right\}$$

- ▶ $V_{T-1}^*(\mathbf{x})$ is a positive-definite quadratic function of \mathbf{u} since $R \succ 0$
- ▶ Taking the gradient and setting it equal to 0:

$$\pi_{T-1}^*(\mathbf{x}) = - \left(\mathbf{B}^\top \mathbf{Q} \mathbf{B} + \mathbf{R} \right)^{-1} \mathbf{B}^\top \mathbf{Q} \mathbf{A} \mathbf{x}$$

$$V_{T-1}^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top M_{T-1} \mathbf{x}$$

$$M_{T-1} = \mathbf{A}^\top M_T \mathbf{A} + \mathbf{Q} - \mathbf{A}^\top M_T \mathbf{B} \left(\mathbf{B}^\top M_T \mathbf{B} + \mathbf{R} \right)^{-1} \mathbf{B}^\top M_T \mathbf{A}$$

Discrete-time Finite-horizon Linear Quadratic Regulator

- ▶ At $t = T - 2$:

$$V_{T-2}^*(\mathbf{x}) = \min_{\mathbf{u}} \frac{1}{2} \left\{ \mathbf{x}^\top Q \mathbf{x} + \mathbf{u}^\top R \mathbf{u} + (A\mathbf{x} + B\mathbf{u})^\top M_{T-1} (A\mathbf{x} + B\mathbf{u}) \right\}$$

- ▶ $V_{T-2}^*(\mathbf{x})$ is a positive-definite quadratic function of \mathbf{u} since $R \succ 0$
- ▶ Taking the gradient and setting it equal to 0:

$$\pi_{T-2}^*(\mathbf{x}) = - \left(B^\top M_{T-1} B + R \right)^{-1} B^\top M_{T-1} A \mathbf{x}$$

$$V_{T-2}^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top M_{T-2} \mathbf{x}$$

$$M_{T-2} = A^\top M_{T-1} A + Q - A^\top M_{T-1} B \left(B^\top M_{T-1} B + R \right)^{-1} B^\top M_{T-1} A$$

Discrete-time Finite-horizon Linear Quadratic Regulator

- ▶ **Batch Approach:** instead of using the DP algorithm, express the system evolution as a large matrix system

$$\underbrace{\begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_T \end{bmatrix}}_{\mathbf{s}} = \underbrace{\begin{bmatrix} I \\ A \\ \vdots \\ A^T \end{bmatrix}}_A \mathbf{x}_0 + \underbrace{\begin{bmatrix} 0 & \cdots & \cdots & 0 \\ B & 0 & \cdots & 0 \\ AB & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ A^{T-1}B & \cdots & \cdots & B \end{bmatrix}}_B \underbrace{\begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{T-1} \end{bmatrix}}_{\mathbf{v}}$$

- ▶ Write the objective function in terms of \mathbf{s} and \mathbf{v} :

$$V_0^\pi(\mathbf{x}_0) = \frac{1}{2} \left(\mathbf{s}^T \mathcal{Q} \mathbf{s} + \mathbf{v}^T \mathcal{R} \mathbf{v} \right) \quad \mathcal{Q} := \mathbf{diag}(Q, \dots, Q, Q) \succeq 0$$
$$\mathcal{R} := \mathbf{diag}(R, \dots, R) \succ 0$$

Discrete-time Finite-horizon Linear Quadratic Regulator

- ▶ Express $V_0^\pi(\mathbf{x}_0)$ only in terms of the initial condition \mathbf{x}_0 and the control sequence \mathbf{v} by using the batch dynamics $\mathbf{s} = \mathcal{A}\mathbf{x}_0 + \mathcal{B}\mathbf{v}$:

$$V_0^\pi(\mathbf{x}_0) = \frac{1}{2} \left(\mathbf{v}^\top \left(\mathcal{B}^\top \mathcal{Q} \mathcal{B} + \mathcal{R} \right) \mathbf{v} + 2\mathbf{x}_0^\top \left(\mathcal{A}^\top \mathcal{Q} \mathcal{A} \right) \mathbf{v} + \mathbf{x}_0^\top \mathcal{A}^\top \mathcal{Q} \mathcal{A} \mathbf{x}_0 \right)$$

- ▶ $V_0^\pi(\mathbf{x}_0)$ is a positive-definite quadratic function of \mathbf{v} since $\mathcal{R} \succ 0$
- ▶ Taking the gradient wrt \mathbf{v} and setting it equal to 0:

$$\mathbf{v}^* = - \left(\mathcal{B}^\top \mathcal{Q} \mathcal{B} + \mathcal{R} \right)^{-1} \mathcal{B}^\top \mathcal{Q} \mathcal{A} \mathbf{x}_0$$

$$V_0^*(\mathbf{x}_0) = \frac{1}{2} \mathbf{x}_0^\top \left(\mathcal{A}^\top \mathcal{Q} \mathcal{A} - \mathcal{A}^\top \mathcal{Q} \mathcal{B} \left(\mathcal{B}^\top \mathcal{Q} \mathcal{B} + \mathcal{R} \right)^{-1} \mathcal{B}^\top \mathcal{Q} \mathcal{A} \right) \mathbf{x}_0$$

- ▶ The optimal sequence of control inputs $\mathbf{u}_{0:T-1}^*$ is a linear function of \mathbf{x}_0
- ▶ The optimal value function $V_0^*(\mathbf{x}_0)$ is a quadratic function of \mathbf{x}_0

Discrete-time Finite-horizon LQG

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

$$\min_{\pi_{0:T-1}} V_0^\pi(\mathbf{x}) := \frac{1}{2} \mathbb{E} \left\{ \sum_{t=0}^{T-1} \gamma^t \left(\mathbf{x}_t^\top Q \mathbf{x}_t + 2\mathbf{u}_t^\top P \mathbf{x}_t + \mathbf{u}_t^\top R \mathbf{u}_t \right) + \gamma^T \mathbf{x}_T^\top Q \mathbf{x}_T \right\}$$

$$\text{s.t. } \mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + C\mathbf{w}_t, \quad \mathbf{x}_0 = \mathbf{x}, \quad \mathbf{w}_t \sim \mathcal{N}(0, I)$$

$$\mathbf{x}_t \in \mathbb{R}^n, \quad \mathbf{u}_t = \pi_t(\mathbf{x}_t) \in \mathbb{R}^m$$

- ▶ **Discount factor:** $\gamma \in [0, 1]$

- ▶ **Optimal value:** $V_t^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top M_t \mathbf{x} + m_t$

- ▶ **Optimal policy:** $\pi_t^*(\mathbf{x}) = -(R + \gamma B^\top M_{t+1} B)^{-1} (P + \gamma B^\top M_{t+1} A) \mathbf{x}$

- ▶ **Riccati Equation:**

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

$$m_t = \gamma m_{t+1} + \gamma \frac{1}{2} \text{tr}(C C^\top M_{t+1}), \quad m_T = 0$$

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

Discrete-time Infinite-horizon LQG

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

$$\min_{\pi} V^{\pi}(\mathbf{x}) := \frac{1}{2} \mathbb{E} \left\{ \sum_{t=0}^{\infty} \gamma^t \left(\mathbf{x}_t^{\top} \mathbf{Q} \mathbf{x}_t + 2 \mathbf{u}_t^{\top} \mathbf{P} \mathbf{x}_t + \mathbf{u}_t^{\top} \mathbf{R} \mathbf{u}_t \right) \right\}$$

$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A} \mathbf{x}_t + \mathbf{B} \mathbf{u}_t + \mathbf{C} \mathbf{w}_t, \quad \mathbf{x}_0 = \mathbf{x}, \quad \mathbf{w}_t \sim \mathcal{N}(0, I)$$
$$\mathbf{x}_t \in \mathbb{R}^n, \quad \mathbf{u}_t = \pi(\mathbf{x}_t) \in \mathbb{R}^m$$

- ▶ **Discount factor:** $\gamma \in [0, 1)$

- ▶ **Optimal value:** $V^*(\mathbf{x}) = \frac{1}{2} \mathbf{x}^{\top} \mathbf{M} \mathbf{x} + m$

- ▶ **Optimal policy:** $\pi^*(\mathbf{x}) = -(\mathbf{R} + \gamma \mathbf{B}^{\top} \mathbf{M} \mathbf{B})^{-1} (\mathbf{P} + \gamma \mathbf{B}^{\top} \mathbf{M} \mathbf{A}) \mathbf{x}$

- ▶ **Riccati Equation** ('dare' in Matlab):

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

$$m = \frac{\gamma}{2(1-\gamma)} \text{tr}(\mathbf{C} \mathbf{C}^{\top} \mathbf{M})$$

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

Relation between Continuous- and Discrete-time LQR

- ▶ The continuous-time system:

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$$

$$\ell(\mathbf{x}, \mathbf{u}) = \frac{1}{2}\mathbf{x}^\top Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\top R\mathbf{u}$$

can be discretized with time step τ :

$$\mathbf{x}_{t+1} = (I + \tau A)\mathbf{x}_t + \tau B\mathbf{u}_t$$

$$\tau\ell(\mathbf{x}, \mathbf{u}) = \frac{\tau}{2}\mathbf{x}^\top Q\mathbf{x} + \frac{\tau}{2}\mathbf{u}^\top R\mathbf{u}$$

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

$$M = \tau Q + (I + \tau A)^\top M(I + \tau A) - (I + \tau A)^\top M \tau B (\tau R + \tau B^\top M \tau B)^{-1} \tau B^\top M (I + \tau A)$$

$$M = \tau Q + M + \tau A^\top M + \tau M A - \tau M B (R + \tau B^\top M B)^{-1} B^\top M + o(\tau^2)$$

$$0 = Q + A^\top M + M A - M B (R + \tau B^\top M B)^{-1} B^\top M + \frac{1}{\tau} o(\tau^2)$$

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 \mathbf{x}_t^* and for approximating non-LQG problems
- ▶ The cost $\|\mathbf{x}_t - \mathbf{x}_t^*\|^2$ can be captured in the LQG formulation by modifying the state and cost as follows:

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

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

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