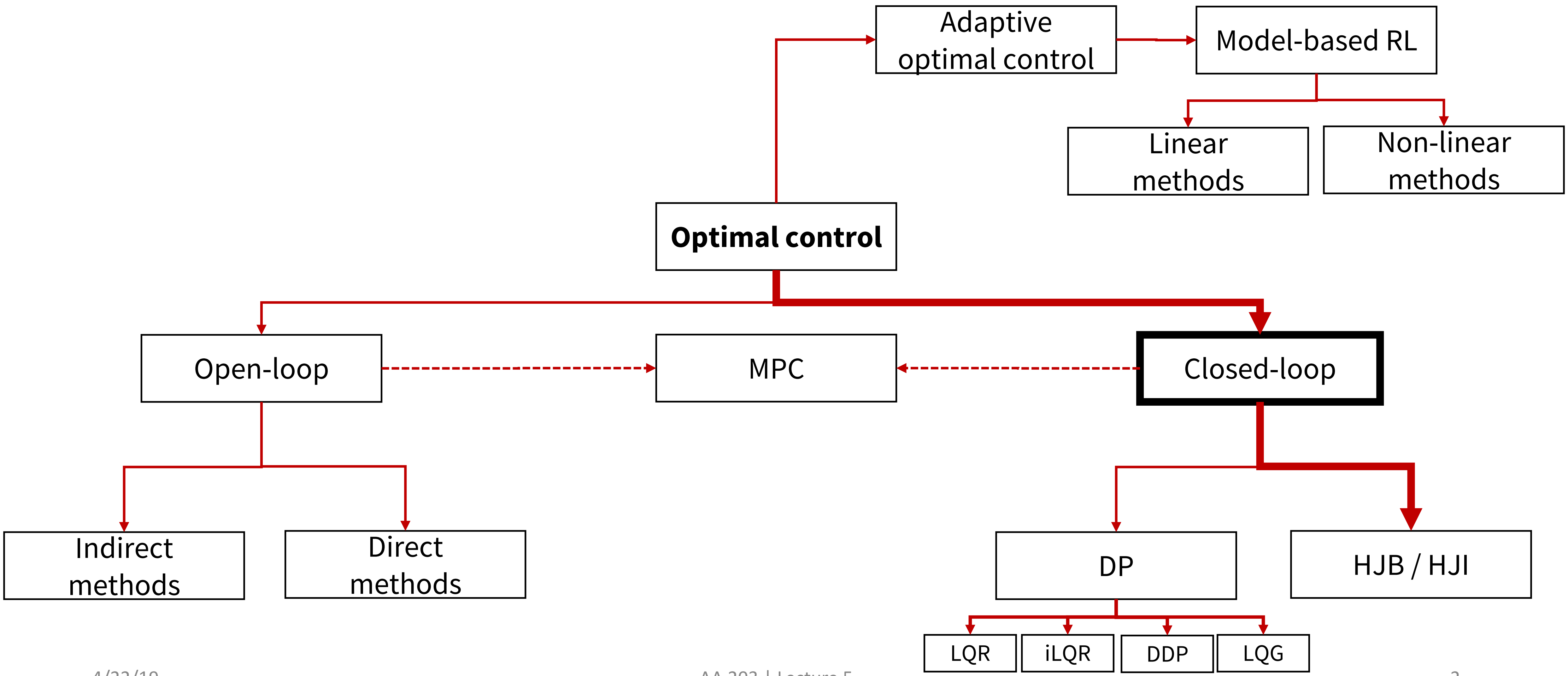
AA203
Optimal and Learning-based Control

HJB Equation and Continuous LQR

Stanford University

Autonomous Systems Lab

Roadmap



Adaptive optimal control
Model-based RL
Linear methods
Non-linear methods
Optimal control
Open-loop
MPC
Closed-loop
Indirect methods
Direct methods
DP
HJB / HJI
LQR
iLQR
DDP
LQG

LQR extensions

Last week: focus on discrete-time setting

This week: focus on continuous-time setting

- dynamic programming approach leads to HJB equation: non-linear partial differential equation
- solution to continuous LQR problem
- differential games
- reachability analysis

Continuous-time model

Last time:

- Model: $\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k, k)$,
- Cost: $J(\mathbf{x}_0) = h_N(\mathbf{x}_N) + \sum_{k=0}^{N-1} g(\mathbf{x}_k, \mathbf{u}_k, k)$

This time:

- Model: $\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t), t)$,
- Cost: $J\big(\mathbf{x}(0)\big) = h\big(\mathbf{x}\big(t_f\big), t_f\big) + \int_{t_0}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau$

where t_0 and t_f are fixed

HJB equation

Key idea: apply principle of optimality

The "truncated" problem is

$$J(\mathbf{x}(t), \mathbf{u}(\tau)_{t \le \tau \le t_f}, t) = h(\mathbf{x}(t_f), t_f) + \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau$$

where $t \le t_f$ and $\mathbf{x}(t)$ is an admissible state value

J^* is the optimal cost-to-go at time t and state $\mathbf{x}(t)$

The optimal solution is

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(\tau)_{t \le \tau \le t_f}} \left\{ h(\mathbf{x}(t_f), t_f) + \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau \right\}$$

HJB equation

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(\tau)_{t \le \tau \le t_f}} \left\{ \int_t^{t+\Delta t} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau + h(\mathbf{x}(t_f), t_f) + \int_{t+\Delta t}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau \right\}$$

$$= \min_{\mathbf{u}(\tau)_{t \le \tau \le t+\Delta t}} \left\{ \int_t^{t+\Delta t} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau + J^*(\mathbf{x}(t+\Delta t), t+\Delta t) \right\}$$

where

- the second equality follows from the principle of optimality
- $J^*(\mathbf{x}(t+\Delta t), t+\Delta t)$ is the minimum cost of the process for the time interval $t+\Delta t \le \tau \le t_f$ with initial state $\mathbf{x}(t+\Delta t)$

HJB equation

For small Δt

$$J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(\tau)_{t \leq \tau \leq t+\Delta t}} \Bigg\{ \int_t^{t+\Delta t} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau + J^*(\mathbf{x}(t), t)$$
$$+ \underbrace{\frac{\partial J^*}{\partial t}(\mathbf{x}(t), t)}_{:= J_t^*(\mathbf{x}(t), t)} \Delta t + \underbrace{\left[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x}(t), t)\right]'}_{:= J_{\mathbf{x}}^*(\mathbf{x}(t), t)'} [\mathbf{x}(t + \Delta t) - \mathbf{x}(t)] + o(\Delta t) \Bigg\}$$

where

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

HJB equation

Carrying this one step further

$$J^*(\mathbf{x}(t),t) = \min_{\mathbf{u}(t)}\Big\{ g(\mathbf{x}(t),\mathbf{u}(t),t)\Delta t + J^*(\mathbf{x}(t),t)$$

$$+ J_t^*(\mathbf{x}(t),t)\Delta t + J_{\mathbf{x}}^*(\mathbf{x}(t),t)' \left[\mathbf{x}(t+\Delta t) - \mathbf{x}(t)\right] + o(\Delta t)\Big\}$$

Hence, we obtain the equation

$$0 = J_t^*(\mathbf{x}(t),t)\Delta t + \min_{\mathbf{u}(t)}\Big\{ g(\mathbf{x}(t),\mathbf{u}(t),t)\Delta t$$

$$+ J_{\mathbf{x}}^*(\mathbf{x}(t),t)' \left[\mathbf{x}(t+\Delta t) - \mathbf{x}(t)\right] + o(\Delta t)\Big\}$$

HJB equation

Dividing by Δt, and taking $\Delta t \to 0$, we obtain the Hamilton-Jacobi-Bellman (HJB) equation

$$\boxed{0 = J_t^*(\mathbf{x}(t), t) + \min_{\mathbf{u}(t)} \left\{ g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^*(\mathbf{x}(t), t)' \, \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \right\}}$$

with boundary condition

$$\boxed{J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f)}$$

- Partial differential equation that should be satisfied for all state-time pairs $(\mathbf{x}, t)$ by the cost-to-go functions $J^*(\mathbf{x}(t), t)$
- Note: we assumed differentiability of $J^*(\mathbf{x}(t), t)$

Hamiltonian formalism

Define the Hamiltonian

$$\mathcal{H}(\mathbf{x}(t), \mathbf{u}(t), J^*_{\mathbf{x}}, t) := g(\mathbf{x}(t), \mathbf{u}(t), t) + J^*_{\mathbf{x}}(\mathbf{x}(t), t)' \, \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)$$

and

$$\mathcal{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J^*_{\mathbf{x}}, t), J^*_{\mathbf{x}}, t) := \min_{\mathbf{u}(t)} \mathcal{H}(\mathbf{x}(t), \mathbf{u}(t), J^*_{\mathbf{x}}, t)$$

Then, HJB equation can be written as

$$0 = J^*_t(\mathbf{x}(t), t) + \mathcal{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J^*_{\mathbf{x}}, t), J^*_{\mathbf{x}}, t)$$

Sufficiency theorem

Suppose $V(\mathbf{x}, t)$ is a solution to the HJB equation, that is V is C^1 in t and $\mathbf{x}$, and is such that

$$0 = V_t(\mathbf{x}, t) + \min_{\mathbf{u} \in U} \left\{ g(\mathbf{x}, \mathbf{u}, t) + V_{\mathbf{x}}(\mathbf{x}, t)' \, \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \right\}$$

$$V(\mathbf{x}, t_f) = h(\mathbf{x}, t_f) \qquad \text{for all } \mathbf{x}$$

Suppose also that $\pi^*(\mathbf{x}, t)$ attains the minimum in this equation for all t and $\mathbf{x}$. Let $\{\mathbf{x}^*(t) | \ t \in \ [t_0, \ t_f]\}$ be the state trajectory obtained from the given initial condition $\mathbf{x}(0)$ when the control trajectory $\mathbf{u}^*(t) = \pi^*(\mathbf{x}^*(t), t), t \in \ [t_0, t_f]$ is used. Then V is equal to the optimal cost-to-go function, i.e.,

$$V(\mathbf{x}, t) = J^*(\mathbf{x}, t) \qquad \text{for all } \mathbf{x} \text{ and } t$$

Furthermore, the control trajectory $\{\mathbf{u}^*(t) | \ t \in \ [t_0, t_f]\}$ is optimal.

Example: continuous LQR

Continuous LQR: select control inputs to minimize

$$J(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}(t_f)' H \mathbf{x}(t_f) + \frac{1}{2}\int_{t_0}^{t_f} \left[\mathbf{x}(t)'Q(t)\mathbf{x}(t) + \mathbf{u}(t)'R(t)\mathbf{u}(t)\right] dt$$

subject to the dynamics

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

Assumptions:

- $H = H' \geq 0,\ Q(t) = Q(t)' \geq 0, R(t) = R(t)' > 0$
- t_0 and t_f specified
- $\mathbf{x}(t)$ and $\mathbf{u}(t)$ unconstrained

Example: continuous LQR

Consider the Hamiltonian

$$\mathcal{H} = \frac{1}{2}\mathbf{x}(t)'Q(t)\mathbf{x}(t) + \frac{1}{2}\mathbf{u}(t)'R(t)\mathbf{u}(t) + J_{\mathbf{x}}^*(\mathbf{x}(t), t)'\left[A(t)\mathbf{x}(t) + B(t)\mathbf{u}(t)\right]$$

Necessary condition to minimize $\mathbf{u}(t)$ is that $\frac{\partial H}{\partial \mathbf{u}} = 0$, that is:

$$\frac{\partial \mathcal{H}}{\partial \mathbf{u}} = R(t)\mathbf{u}(t) + B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t), t) = 0$$

Since $\frac{\partial^2 H}{\partial \mathbf{u}^2} = R(t) > 0$, the control that satisfies the NOC is a global minimum; solving:

$$\mathbf{u}^*(t) = -R^{-1}(t)B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t), t)$$

Example: continuous LQR

Plugging in:

$$
\begin{aligned}
\mathcal{H} &= \frac{1}{2}\mathbf{x}(t)'Q(t)\mathbf{x}(t) + \frac{1}{2}J_{\mathbf{x}}^*(\mathbf{x}(t),t)'B(t)R(t)^{-1}B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t),t) \\
&\quad + J_{\mathbf{x}}^*(\mathbf{x}(t),t)'A(t)\mathbf{x}(t) - J_{\mathbf{x}}^*(\mathbf{x}(t),t)'B(t)R(t)^{-1}B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t),t) \\
&= \frac{1}{2}\mathbf{x}(t)'Q(t)\mathbf{x}(t) - \frac{1}{2}J_{\mathbf{x}}^*(\mathbf{x}(t),t)'B(t)R(t)^{-1}B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t),t) \\
&\quad + J_{\mathbf{x}}^*(\mathbf{x}(t),t)'A(t)\mathbf{x}(t)
\end{aligned}
$$

Example: continuous LQR

The HJB equation is then

$$
\begin{aligned}
0 = J_t^*(\mathbf{x}(t), t) + \frac{1}{2}\mathbf{x}(t)'Q(t)\mathbf{x}(t) \\
- \frac{1}{2}J_{\mathbf{x}}^*(\mathbf{x}(t), t)'B(t)R(t)^{-1}B(t)'J_{\mathbf{x}}^*(\mathbf{x}(t), t) + J_{\mathbf{x}}^*(\mathbf{x}(t), t)'A(t)\mathbf{x}(t)
\end{aligned}
$$

with boundary conditions

$$J^*(\mathbf{x}(t_f), t_f) = \frac{1}{2}\mathbf{x}(t_f)'H\mathbf{x}(t_f)$$

Example: continuous LQR

Ansatz (lucky guess):

$$J^*(\mathbf{x}(t), t) = \frac{1}{2}\mathbf{x}(t)' K(t) \mathbf{x}(t) \qquad \text{for } \textit{some } K(t) > 0$$

Substituting

$$0 = \frac{1}{2}\mathbf{x}(t)' \dot{K}(t)\mathbf{x}(t) + \frac{1}{2}\mathbf{x}(t)' Q(t)\mathbf{x}(t) \\ - \frac{1}{2}\mathbf{x}(t)' K(t) B(t) R(t)^{-1} B(t)' K(t)\mathbf{x}(t) + \mathbf{x}(t)' K(t) A(t)\mathbf{x}(t)$$

Note:

- $J^*_{\mathbf{x}} = K(t)\mathbf{x}(t)$ and $J^*_t = \frac{1}{2}\mathbf{x}(t)'\dot{K}(t)\mathbf{x}(t)$
- $\mathbf{x}(t)' K(t) A(t)\mathbf{x}(t) = \frac{1}{2}\mathbf{x}(t)' K(t) A(t)\mathbf{x}(t) + \frac{1}{2}\mathbf{x}(t)' A(t)' K(t)\mathbf{x}(t)$

Example: continuous LQR

So we obtain

$$0 = \frac{1}{2}\mathbf{x}(t)'\Big(\dot{K}(t) + Q(t)$$

$$- K(t)B(t)R(t)^{-1}B(t)'K(t) + K(t)A(t) + A(t)'K(t)\Big)\mathbf{x}(t)$$

Since this equation must hold for all $\mathbf{x}(t)$, we obtain

$$-\dot{K}(t) = Q(t) - K(t)B(t)R(t)^{-1}B(t)'K(t) + K(t)A(t) + A(t)'K(t)$$

with boundary condition $K(t_f) = H$

Example: continuous LQR

- The HJB equation reduces to a set of differential equation (the Riccati equation)
- The Riccati equation is integrated backwards
- Once we find $K(t)$, the control policy is

$$\mathbf{u}^*(t) = -R(t)^{-1}B(t)'K(t)\mathbf{x}(t)$$

- Analogously to the discrete case, under some additional assumptions, $K(t) \rightarrow$ constant in the infinite horizon setting

Next time

- HJI and reachability analysis