Module 05 — Introduction to Optimal Control

Ahmad F. Taha

EE 5243: Introduction to Cyber-Physical Systems

Email: ahmad.taha@utsa.edu

Webpage: http://engineering.utsa.edu/~taha/index.html

October 8, 2015

Outline

In this Module, we discuss the following:

- What is optimal control? How is it different than regular optimization?
- A general optimal control problem
- Dynamic programming & principle of optimality + example
- HJB equation, PMP + example
- LQR for LTV systems, important remarks + example

Motivation & Intro

- **Functionals:** mappings from a set of functions to real numbers
- Often expressed as definite integrals involving functions
- **Calculus of variations:** maximizing or minimizing functionals
- Example: find a curve of shortest length connecting two points under constraints
- **Optimal control:** extension of **calculus of variations** — *a mathematical optimization method for deriving control policies*
- **Pioneers**: Pontryagin and Bellman

Your Daily Optimal Control Problem

- Optimal control: finding a control law s.t. an optimality criterion is achieved
- OCP: **cost functional** + **differential equations** + **bounds on control & state** (constraints)
- OC law: derived using Pontryagin's maximum principle (a necessary condition), or by solving the HJB equation (a sufficient condition)
- Example: driving on a hilly road — *how should the driver drive such that traveling time is minimized?*
- Control: driving way (pedaling, steering, gearing)
- Constraints: car & road dynamics, speed limits, fuel, ICs
- Objective: $\text{minimize } (t_{final} - t_{initial})$

Your Daily Drive — In Equations

- Can we translate the optimal driving route to equations? **Yes!**
- Your optimal drive problem can be, hypothetically, written as:

$$\text{minimize } J = \underbrace{\Phi\left[x(t_0), t_0, x(t_f), t_f\right] + \int_{t_0}^{t_f} \mathcal{L}\left[x(t), u(t), t\right] \mathrm{d}\,t}_{\text{minimal cost-functional}}$$

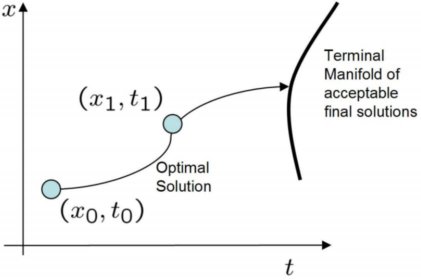
$$\text{subject to } \underbrace{\dot{x}(t) = f\left[x(t), u(t), t\right]}_{\text{state-space dynamics: your car dynamics}}$$

$$\underbrace{g\left[x(t), u(t), t\right] \leq 0}_{\text{algebraic constraints: the road-constraints, pedalling, steering, gearing}}$$

$$\underbrace{\phi\left[x(t_0), t_0, x(t_f), t_f\right]}_{\text{final \& initial speeds, location}} = 0$$

Principle of Optimality (PoO)

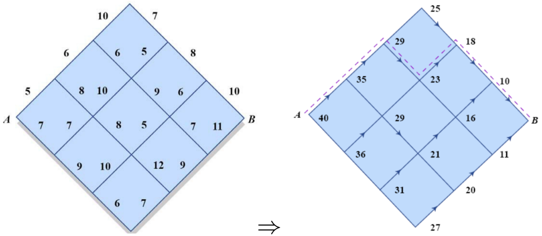
- **Principle of Optimality:** optimal solution for a problem passes through $(x_1, t_1) \Rightarrow$ optimal solution starting at (x_1, t_1) must be continuation of the same path



- This paved the way to numerical solutions, such as **dynamic programming**

Dynamic Programming

- **DP:** solving a large-scale, complex **problem** by solving small-scale, less complex **subproblems**
- DP combines optimization + computer science methods, uses PoO
- **Example**: travel from A to B with least cost (robot navigation or aircraft path)



- 20 possible options, trying all would be so tedious
- Strategy: start from B, and go backwards, invoking PoO

Discrete LQR + DP

- Many DP problems are solved numerically
- Discrete LQR can be solved **analytically**
- **Objective:** select optimal control inputs to minimize J:

$$\min \; J \;=\; \frac{1}{2}x_N^\top H x_N + \sum_{k=0}^{N-1} \underbrace{\frac{1}{2}\left[x_k^\top Q_k x_k + u_k^\top R_k u_k\right]}_{=g(x_k,u_k)}$$

$$\text{subject to} \quad x_{k+1} = A_k x_k + B_k u_k$$

$$H = H^\top, Q = Q^\top \succeq 0, R = R^\top \succ 0$$

- Use DP to solve the LQR for LTV systems. **How?**

$$\boxed{J_{k-1}^*[x_{k-1}] = \min_{u_{k-1}} \left\{g(x_{k-1}, u_{k-1}) + J_k^*[x_k]\right\}}$$

- Start from $k = N$. What is $J_N^*[x_N]$? Clearly, it is: $\boxed{J_N^*[x_N] = \frac{1}{2}x_N^\top H x_N}$

Discrete LQR

$$\boxed{J^*_{k-1}[x_{k-1}] = \min_{u_{N-1}} \{g(x_{k-1}, u_{k-1}) + J^*_k[x_k]\}}$$

- We now know that $J^*_N[x_N] = \frac{1}{2} x_N^\top H x_N \Rightarrow$

$$J^*_{N-1}[x_{N-1}] = \min_{u_{N-1}} \{g(x_{N-1}, u_{N-1}) + J^*_N[x_N]\}$$

$$= \frac{1}{2} \min_{u_{N-1}} \left\{x_{N-1}^\top Q_{N-1} x_{N-1} + u_{N-1}^\top R_{N-1} u_{N-1} + x_N^\top H x_N\right\}$$

- From state-dynamics: $x_N = A_{N-1} x_{N-1} + B_{N-1} u_{N-1}$, thus:

$$J^*_{N-1}[x_{N-1}] = \frac{1}{2} \min_{u_{N-1}} \{x_{N-1}^\top Q_{N-1} x_{N-1} + u_{N-1}^\top R_{N-1} u_{N-1}$$

$$+(A_{N-1} x_{N-1} + B_{N-1} u_{N-1})^\top H (A_{N-1} x_{N-1} + B_{N-1} u_{N-1})\}$$

- Find optimal control by taking derivative of J_{N-1} with respect to u_{N-1}:

$$\frac{\partial J^*_{N-1}}{\partial u_{N-1}} = u_{N-1}^\top R_{N-1} + (A_{N-1} x_{N-1} + B_{N-1} u_{N-1})^\top H B_{N-1} = 0$$

Optimality Conditions

- Optimality condition at step $N-1$ yields:

$$(R_{N-1} + B_{N-1}^\top H B_{N-1}) u_{N-1}^* + B_{N-1}^\top H A_{N-1} x_{N-1} = 0$$

- Therefore, candidate optimal u_{N-1}^* can be written as:

$$\boxed{u_{N-1}^* = -\underbrace{\left(R_{N-1} + B_{N-1}^\top H B_{N-1}\right)^{-1} B_{N-1}^\top H A_{N-1}}_{F_{N-1}} x_{N-1}}$$

- *What is that?* It's simply an **optimal, time-varying linear state feedback!**

- Second order necessary condition are satisfied:

$$\frac{\partial^2 J_{N-1}^*}{\partial u_{N-1}^2} = R_{N-1} + B_{N-1}^\top H B_{N-1} \succ 0$$

Discrete LTV LQR Solutions

$$u^*_{N-1} = -\underbrace{\left(R_{N-1} + B^\top_{N-1} H B_{N-1}\right)^{-1} B^\top_{N-1} H A_{N-1}}_{F_{N-1}} x_{N-1}$$

- Given this optimal control action at $N-1$, what is the optimal cost? By substitution,

$$J^*_{N-1}[x_{N-1}] = \frac{1}{2}\left\{x^\top_{N-1} Q_{N-1} x_{N-1} + (u^*_{N-1})^\top R_{N-1} u^*_{N-1} + x^\top_N H x_N\right\}$$

- Therefore,

$$J^*_{N-1}[x_{N-1}] = \frac{1}{2} x^\top_{N-1} P_{N-1} x_{N-1} \quad \text{, where}$$

$$P_{N-1} = Q_{N-1} + F^\top_{N-1} R_{N-1} F_{N-1} + (A_{N-1} - B_{N-1} F_{N-1})^\top H (A_{N-1} - B_{N-1} F_{N-1})$$

- Since $P_N = H$, then:

$$F_{N-1} = \left(R_{N-1} + B^\top_{N-1} P_N B_{N-1}\right)^{-1} B^\top_{N-1} P_N A_{N-1}$$

Discrete LTV LQR Algorithm

For $k = N-1 \to 0$:

1. $P_N = H$
2. $F_k = \left(R_k + B_k^\top P_{k+1} B_k\right)^{-1} B_k^\top P_{k+1} A_k$
3. $P_k = Q_k + F_k^\top R_k F_k + (A_k - B_k F_k)^\top P_{k+1} (A_k - B_k F_k)$

Remarks:

- The optimal solution is a time-varying control law, for time-varying A, B, Q, R
- Result can be easily applied to LTI systems
- Assumption that $R_k \succ 0$ can be relaxed
- P_k and F_k can be computed offline — both independent on x and u
- Can eliminate F_k

DP Example + LQR

For this dynamical system,

$$x_{k+1} = bu_k, \;\; b \neq 0,$$

find u_0^*, u_1^* such that $\boxed{J = (x_2 - 1)^2 + 2\sum_{k=0}^{1} u_k^2}$ is minimized.

- In DP, we start from the terminal conditions
- By definition, $J^*(x_k) \equiv$ optimal cost of transfer from x_k to x_2
- We know that: $J^*(x_2) = (x_2 - 1)^2 = (bu_1 - 1)^2$

$$J^*(x_1) = \min_{u_1}(2u_1^2 + J^*(x_2)) = \min_{u_1}(2u_1^2 + (bu_1 - 1)^2)$$

- Setting $\dfrac{\partial J^*(x_1)}{\partial u_1} = 4u_1 + 2b(bu_1 - 1) = 0 \rightarrow \boxed{u_1^* = \dfrac{b}{b^2 + 2}}$
- Similarly: $J^*(x_0) = \min_{u_0}(2u_0^2 + J^*(x_1)) = \min_{u_0}(2u_0^2 + \dfrac{b}{b^2+2})$
- Therefore, $\boxed{u_0^* = 0}$

HJB Equation

- Previous approach is relatively easy for DT systems
- But what if we want to consider closed-form, exact solutions for CT NL ODEs?

$$\begin{aligned} \text{minimize } J &= h\,(x(t_f), t_f\,) + \int_{t_0}^{t_f} g\,(\,x(t), u(t), t\,)\,dt \\ \text{subject to } & \quad \dot{x}(t) = f\,(\,x, u, t\,),\;\; x(t_0) = x_{t_0} \end{aligned}$$

- Objective: find $u^*(t)$, $t_0 \leq t \leq t_f$, such that the cost is minimized
- **Hamiltonian:** $\boxed{\mathcal{H}(x, u, \lambda^*(x,t), t) = g(x,u,t) + \lambda^*(x,t) f(x,u,t)}$

HJB Equation and PMP

- **Value function**, optimal cost-to-go: $\boxed{V(x,t) = \min_{u} J(x,u,t)}$
- Value function properties:
  1. $V_x(x,t) = \frac{\partial V}{\partial x} = \lambda^*(x,t)$
  2. $-V_t(x,t) = -\frac{\partial V}{\partial t} = \min_{u \in \mathcal{U}} \mathcal{H}(x,u,\lambda^*(x,t),t) = \left(\frac{\partial \mathcal{H}}{\partial x}\right)^\top$
- The **HJB Equation:**

$$\boxed{-V_t^*(x,t) = -\frac{\partial V}{\partial t} = \min_{u \in \mathcal{U}} \mathcal{H}(x,u,\lambda^*(x,t),t) = \left(\frac{\partial \mathcal{H}}{\partial x}\right)^\top}$$

- *What is this?* It's a PDE.

Pontryagin's Maximum Principle (PMP)

Optimal control u^* must satisfy:

$$\boxed{H(x^*(t),u^*(t),\lambda^*(x,t),t) \leq H(x^*(t),u(t),\lambda^*(x,t),t), \quad \forall u \in \mathcal{U}, \quad t \in [t_0,t_f]}$$

HJB-Equation Example

Compute the optimal $u^*(t)$ and $x^*(t)$ for the following optimal control problem:

$$\begin{aligned} \text{minimize} \quad & \int_1^2 \sqrt{1+u^2(t)}\, dt \\ \text{subject to} \quad & \dot{x}(t) = u(t), \;\; x(1) = 3, \;\; x(2) = 5 \end{aligned}$$

- First, construct the Hamiltonian: $\mathcal{H}(x, u, J_x, t) = \sqrt{1+u^2(t)} + \lambda(x,t)u(t)$
- Since there are no constraints on $u(t)$, the optimal controller candidate is:

$$0 = \frac{\partial \mathcal{H}}{\partial u} = \lambda(x,t) + \frac{u}{\sqrt{1+u^2}} \Rightarrow u^*(t) = \frac{\lambda(x,t)}{\sqrt{1-\lambda^2(x,t)}}$$

- HJB equation: $-V_t(x,t) = \left(\frac{\partial \mathcal{H}}{\partial x}\right)^\top = 0 \Rightarrow V(t,x) = v$ is constant
- Therefore, $u^*(t) = \frac{\lambda(x,t)}{\sqrt{1-\lambda^2(x,t)}} = \frac{\lambda}{\sqrt{1-\lambda^2}} = c$ is also constant
- Since $x(1)$ and $x(2)$ are given, we can determine $u^*(t) = c$, as follows:

$$x(2) = x(1) + \int_1^2 c\, d\tau \Rightarrow u^*(t) = c = 2, \Rightarrow x(t) = x(1) + \int_1^t 2\, d\tau = 2t+1$$

Continuous LTV, LQR

- How about the continuous LQR?

$$\begin{aligned} \text{minimize } J &= \frac{1}{2}x_{t_f}^\top H x_{t_f} + \frac{1}{2}\int_{t_0}^{t_f} \left[x(t)^\top Q(t)x(t) + u(t)^\top R(t)u(t)\right] dt \\ \text{subject to } & \quad \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ & \quad H = H^\top, Q = Q^\top \succeq 0, R = R^\top \succ 0 \end{aligned}$$

- Construct the **Hamiltonian:**
$\mathcal{H}(x, u, \lambda^*(x,t), t) = g(x,u,t) + \lambda^*(x,t) f(x,u,t)$

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

- Minimum of $\mathcal{H}$ w.r.t. u:

$$\frac{\partial \mathcal{H}}{\partial u} = u(t)^\top R(t) + \lambda^*(x,t)B(t) = 0 \Rightarrow \boxed{u^*(t) = -R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top}$$

- Note that $\dfrac{\partial^2 \mathcal{H}}{\partial u^2} = R(t) \succ 0$

Optimal Control for LTV Systems

- What do we have now? Optimal control law as a function of $\lambda^*(x,t)$:

$$\boxed{u^*(t) = -R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top}$$

- Write the **Hamiltonian** in terms of $u^*(t) : \mathcal{H}(x,u,\lambda^*(x,t),t) =$

$$\frac{1}{2}\left[x(t)^\top Q(t)x(t) + \left(R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top\right)^\top R(t)\left(R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top\right)\right]$$
$$+\lambda^*(x,t)\left[A(t)x(t) + B(t)R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top\right]$$
$$= \frac{1}{2}x(t)^\top Q(t)x(t) + \lambda^*(x,t)A(t)x(t) - \frac{1}{2}\lambda^*(x,t)B(t)R^{-1}(t)B^\top(t)\lambda^*(x,t)^\top \quad (*)$$

- Consider a candidate VF: $\boxed{V^*(x,t) = \frac{1}{2}x^\top(t)P(t)x(t),}$ $P(t) = P^\top(t)$

- Properties of VF (see previous slides):
  1. $V_x^*(x,t) = \lambda^*(x,t) = x^\top(t)P(t)$[1]
  2. $V_t^* = \frac{1}{2}x^\top(t)\dot{P}(t)x(t) = -\min_{u\in\mathcal{U}} \mathcal{H}(x,u,\lambda^*(x,t),t) = -(*)$

[1] The partial derivatives taken w.r.t. one variable assuming the other is fixed. Note that there are two independent variables in this problem x and t: x is time-varying, but not a function of t.

Solution for LTV, LQR

$$\lambda^*(x,t) = x^\top(t)P(t)$$

- Substitute $\lambda^*(x,t)$ into $(*)$:

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

$$= \frac{1}{2}x(t)^\top \left(Q(t) + P(t)A(t) + A^\top(t)P(t) - P(t)B(t)R^{-1}(t)B(t)^\top P(t)\right) x(t) \ (**)$$

- But $-V_t^*(x,t) = (*) = (**) = -\frac{1}{2}x^\top(t)\dot{P}(t)x(t)$
- Hence, for $V^*(x,t) = \frac{1}{2}x^\top(t)P(t)x(t)$ to be an optimal VF, we require:
  1. $\boxed{-\dot{P}(t) = Q(t) + P(t)A(t) + A^\top(t)P(t) - P(t)B(t)R^{-1}(t)B(t)^\top P(t)}$
  2. $P(t_f) = H$
  3. 1. and 2. generate a solution $P(t)$ for a **Differential Riccati Equation**

Remarks on LTV, LQR Solution

- Recall that $u^*(t) = -R^{-1}(t)B(t)^\top \lambda^*(x,t)^\top = -\underbrace{R^{-1}(t)B(t)^\top P(t)}_{=F(t)} x(t)$
- Hence, solution is (again) a time-varying, LSF control law
- Real-time gains $(K(t))$ can be generated offline
- What happens when $t_f \to \infty$? Well...DRE saturates $\Rightarrow \dot{P}(t) = 0$
- Hence, we can solve the **continuous algebraic Riccati equation** (**CARE**):

$$\boxed{Q + P_{ss}A + A^\top P_{ss} - P_{ss}BR^{-1}B^\top P_{ss} = 0}$$

- CARE solves for $P = P^\top \succeq 0$ — can we write this as an LMI? (it looks like a **bilateral** matrix inequality, not an LMI, though)
- **Fact:** If (A, B, C) are stabilizable and detectable $\Rightarrow$ steady state solution P_{ss} approaches unique PSD CARE solution

LTI, CT LQR Example

Find the optimal LSF controller, $u = -Kx$, that minimizes:

$$J = \int_0^\infty u^2(t)\, dt, \;\; \text{subject to} \;\; \dot{x}(t) = x(t) + 2u(t), \;\; x(0) = 1.$$

- From the previous slide, if $t_f = \infty$, we can solve `CARE`
- For the given J and dynamics, we have: $Q = 0, R = I, A = 1, B = 2$
- CARE (variable is $P \in \mathbb{R}^{1\times 1}$):

$$Q + PA + A^\top P - PBR^{-1}B^\top P = 0 + 1 \cdot p + p \cdot 1 - p^2(2)(1)(2) = 0$$

- Or: $2p - 4p^2 = 0 \Rightarrow p = \dfrac{1}{2}$, ($p = 0$ is not positive definite)
- Thus, $u^*(t) = -R^{-1}B^\top Px(t) = -x(t)$
- Optimal cost: $J_{\min} = \int_0^\infty (u^*(t))^2\, dt = \int_0^\infty x^\top(t)x(t)\, dt = x_0^\top P x_0 = \dfrac{1}{2}$

Questions And Suggestions?

Thank You!

Please visit

engineering.utsa.edu/~taha

IFF you want to know more ☺

References I

1. Bertsekas, Dimitri P., and Dimitri P. Bertsekas. *Dynamic programming and optimal control*. Vol. 1, no. 2. Belmont, MA: Athena Scientific, 1995.
2. Żak, Stanislaw H. *Systems and control*. New York: Oxford University Press, 2003.
3. Jonathan How. *16.323 Principles of Optimal Control, Spring 2008.* (Massachusetts Institute of Technology: MIT OpenCourseWare), http://ocw.mit.edu