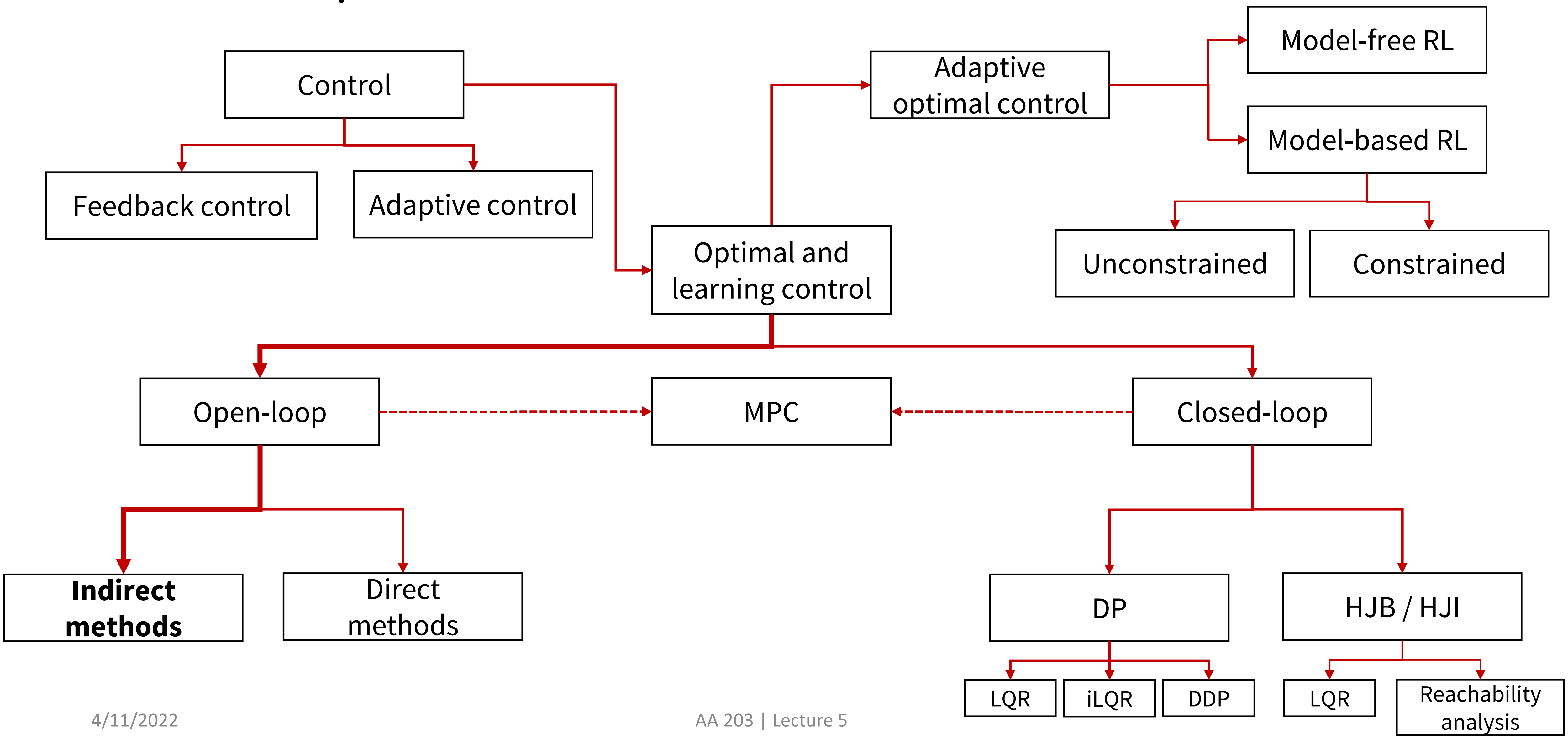
AA203
Optimal and Learning-based Control

CoV extensions, NOC for optimal control

Stanford University

Autonomous Systems Lab

Roadmap


Control
Feedback control
Adaptive control
Optimal and learning control
Adaptive optimal control
Model-free RL
Model-based RL
Unconstrained
Constrained
Open-loop
MPC
Closed-loop
Indirect methods
Direct methods
DP
HJB / HJI
LQR
iLQR
DDP
LQR
Reachability analysis

CoV extension I: generalized boundary conditions

- Let $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ be a vector-valued function, where each component x_i is in the class of functions with continuous first derivatives. It is desired to find the function $\mathbf{x}^*$ for which the functional

$$J(\mathbf{x}) = \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)dt$$

  has a relative extremum
- Assumptions:
  - $g \in C^2$
  - t_0 and $\mathbf{x}(0)$ are fixed
  - t_f might be fixed or free, and each component of $\mathbf{x}(t_f)$ might be fixed or free
- Reading:
  - D. E. Kirk. *Optimal Control Theory: An Introduction*, 2004.

CoV extension I: generalized boundary conditions

- Regardless of the boundary conditions, the Euler equations

$$g_{\mathbf{x}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), t) - \frac{d}{dt} g_{\dot{\mathbf{x}}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), t) = \mathbf{0}$$

  must be satisfied

CoV extension I: generalized boundary conditions

- Regardless of the boundary conditions, the Euler equations

$$g_{\mathbf{x}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), t) - \frac{d}{dt} g_{\dot{\mathbf{x}}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), t) = \mathbf{0}$$

  must be satisfied

- The required boundary conditions are found from the equation

$$g_{\dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)^T \delta\mathbf{x}_f + \left[g(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f) - g_{\dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)^T \dot{\mathbf{x}}^*(t_f)\right] \delta t_f = 0$$

  by making the "appropriate" substitutions for $\delta\mathbf{x}_f$ and δt_f

CoV extension I: generalized boundary conditions

- $\delta \mathbf{x}_f$ and δt_f capture the notion of "allowable" variations at the end point, thus $\delta t_f = 0$ if the final time is fixed, and $\delta x_i(t_f) = 0$ if the end value of state variable $x_i(t_f)$ is fixed
- For example, suppose that t_f is fixed, $x_i(t_f), i = 1, \dots, r$ are fixed, and $x_j(t_f), j = r+1, \dots, n$ are free. Then the substitutions are:

$$\delta t_f = 0$$

$$\delta x_i(t_f) = 0, \qquad i = 1, \dots, r$$

$$\delta x_j(t_f) \text{ arbitrary}, \qquad j = r+1, \dots, n$$

CoV extension I: generalized boundary conditions

Problem description	*Substitution*	*Boundary conditions*	*Remarks*
1. $\mathbf{x}(t_f)$, t_f both specified (*Problem 1*)	$\delta\mathbf{x}_f = \delta\mathbf{x}(t_f) = \mathbf{0}$ $\delta t_f = 0$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\mathbf{x}^*(t_f) = \mathbf{x}_f$	$2n$ equations to determine $2n$ constants of integration
2. $\mathbf{x}(t_f)$ free; t_f specified (*Problem 2*)	$\delta\mathbf{x}_f = \delta\mathbf{x}(t_f)$ $\delta t_f = 0$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\frac{\partial g}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f) = \mathbf{0}$	$2n$ equations to determine $2n$ constants of integration
3. t_f free; $\mathbf{x}(t_f)$ specified (*Problem 3*)	$\delta\mathbf{x}_f = \mathbf{0}$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\mathbf{x}^*(t_f) = \mathbf{x}_f$ $g(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)$ $- \left[\frac{\partial g}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)\right]^T \dot{\mathbf{x}}^*(t_f) = 0$	$(2n + 1)$ equations to determine $2n$ constants of integration and t_f
4. t_f, $\mathbf{x}(t_f)$ free and independent (*Problem 4*)	—	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\frac{\partial g}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f) = \mathbf{0}$ $g(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f) = 0$	$(2n + 1)$ equations to determine $2n$ constants of integration and t_f
5. t_f, $\mathbf{x}(t_f)$ free but related by $\mathbf{x}(t_f) = \boldsymbol{\theta}(t_f)$ (*Problem 4*)	$\delta\mathbf{x}_f = \frac{d\boldsymbol{\theta}}{dt}(t_f)\,\delta t_f$†	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\mathbf{x}^*(t_f) = \boldsymbol{\theta}(t_f)$ $g(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)$ $+ \left[\frac{\partial g}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t_f), \dot{\mathbf{x}}^*(t_f), t_f)\right]^T \left[\frac{d\boldsymbol{\theta}}{dt}(t_f) - \dot{\mathbf{x}}^*(t_f)\right] = 0$†	$(2n + 1)$ equations to determine $2n$ constants of integration and t_f

From Kirk, Section 4.3

Example

- Determine the smooth curve of smallest length connecting the point $x(0) = 1$ to the line $t = 5$
  - Solution: $x(t) = 1$

CoV extension II: constrained extrema

- Let $\mathbf{w} : \mathbb{R} \rightarrow \mathbb{R}^{n+m}$ be a vector-valued function, where each component w_i is in the class of functions with continuous first derivatives. It is desired to find the function $\boldsymbol{w}^*$ for which the functional

$$J(\mathbf{w}) = \int_{t_0}^{t_f} g(\mathbf{w}(t), \dot{\mathbf{w}}(t), t)dt$$

has a relative extremum, subject to the constraints

$$f_i(\mathbf{w}(t), \dot{\mathbf{w}}(t), t) = 0, \qquad i = 1, \dots, n$$

- Assumptions:
  - $g \in C^2$
  - t_0 and $\mathbf{w}(0)$ are fixed

CoV extension II: constrained extrema

- Because of the n differential constraints, only m of the $n + m$ components of $\mathbf{w}$ are independent
- Constraints of this type may represent the state equation constraints in optimal control problems where $\mathbf{w}$ corresponds to the $n + m$ vector $\mathbf{w} = [\mathbf{x}, \mathbf{u}]^{\mathbf{T}}$
- Similar to the case of constrained optimization, define the augmented integrand function

$$g_a(\mathbf{w}(t), \dot{\mathbf{w}}(t), \mathbf{p}(t), t) \coloneqq$$
$$g(\mathbf{w}(t), \dot{\mathbf{w}}(t), t) + \boxed{\mathbf{p}(t)^T} \mathbf{f}(\mathbf{w}(t), \dot{\mathbf{w}}(t), t)$$

Lagrange multipliers (now functions of time!), the "costate"

CoV extension II: constrained extrema

- A necessary condition for optimality is then

$$\frac{\partial g_a}{\partial \mathbf{w}}(\mathbf{w}^*(t), \dot{\mathbf{w}}^*(t), \mathbf{p}^*(t), t) - \frac{d}{dt}\frac{\partial g_a}{\partial \dot{\mathbf{w}}}(\mathbf{w}^*(t), \dot{\mathbf{w}}^*(t), \mathbf{p}^*(t), t) = \mathbf{0}$$

along with

$$\mathbf{f}(\mathbf{w}^*(t), \dot{\mathbf{w}}^*(t), t) = \mathbf{0}$$

- That is, to determine the necessary conditions for an extremal we simply form the augmented integrand g_a and write the Euler equations *as if* there were no constraints among the functions $\mathbf{w}(t)$
- Note the similarity with the case of constrained optimization!

The variational approach to optimal control

Roadmap:

1. We will first derive necessary conditions for optimal control assuming that the admissible controls are not bounded
2. Next, we will heuristically introduce Pontryagin's Minimum Principle as a generalization of the fundamental theorem of CoV
3. Finally, we will consider special cases of problems with bounded controls and state variables

Necessary conditions for optimal control
(with unbounded controls)

- The problem is to find an *admissible control* $\mathbf{u}^*$ which causes the system

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

to follow an *admissible trajectory* $\mathbf{x}^*$ that minimizes the *functional*

$$J(\mathbf{u}) = h\big(\mathbf{x}(t_f), t_f\big) + \int_{t_0}^{t_f} g\,(\mathbf{x}(t), \mathbf{u}(t), t)\, dt$$

- Assumptions: $h \in C^2$, state and control regions are unbounded, t_0 and $\mathbf{x}(0)$ are fixed

Necessary conditions for optimal control
(with unbounded controls)

- Define the Hamiltonian

$$H(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}(t), t) \coloneqq g(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{p}(t)^T \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)$$

- The necessary conditions for optimality (proof to follow) are

$$\left.\begin{aligned} \dot{\mathbf{x}}^*(t) &= \frac{\partial H}{\partial \mathbf{p}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \\ \dot{\mathbf{p}}^*(t) &= -\frac{\partial H}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \\ \mathbf{0} &= \frac{\partial H}{\partial \mathbf{u}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \end{aligned}\right\} \text{ for all } t \in [t_0, t_f]$$

with boundary conditions

$$\left[\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f), t_f) - \mathbf{p}^*(t_f)\right]^T \delta\mathbf{x}_f + \left[H(\mathbf{x}^*(t_f), \mathbf{u}^*(t_f), \mathbf{p}^*(t_f), t_f) + \frac{\partial h}{\partial t}(\mathbf{x}^*(t_f), t_f)\right] \delta t_f = 0$$

Necessary conditions for optimal control
(with unbounded controls)

Problem	Description	Substitution in Eq. (5.1-18)	Boundary-condition equations	Remarks
t_f fixed	1. $\mathbf{x}(t_f) = \mathbf{x}_f$ specified final state	$\delta \mathbf{x}_f = \delta \mathbf{x}(t_f) = \mathbf{0}$ $\delta t_f = 0$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\mathbf{x}^*(t_f) = \mathbf{x}_f$	$2n$ equations to determine $2n$ constants of integration
	2. $\mathbf{x}(t_f)$ free	$\delta \mathbf{x}_f = \delta \mathbf{x}(t_f)$ $\delta t_f = 0$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f)) - \mathbf{p}^*(t_f) = \mathbf{0}$	$2n$ equations to determine $2n$ constants of integration
	3. $\mathbf{x}(t_f)$ on the surface $\mathbf{m}(\mathbf{x}(t)) = \mathbf{0}$	$\delta \mathbf{x}_f = \delta \mathbf{x}(t_f)$ $\delta t_f = 0$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f)) - \mathbf{p}^*(t_f) = \sum_{i=1}^{k} d_i \left[\frac{\partial m_i}{\partial \mathbf{x}}(\mathbf{x}^*(t_f))\right]$ $\mathbf{m}(\mathbf{x}^*(t_f)) = \mathbf{0}$	$(2n + k)$ equations to determine the $2n$ constants of integration and the variables $d_1, \ldots, d_k$
t_f free	4. $\mathbf{x}(t_f) = \mathbf{x}_f$ specified final state	$\delta \mathbf{x}_f = \mathbf{0}$	$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\mathbf{x}^*(t_f) = \mathbf{x}_f$ $\mathscr{H}(\mathbf{x}^*(t_f), \mathbf{u}^*(t_f), \mathbf{p}^*(t_f), t_f) + \frac{\partial h}{\partial t}(\mathbf{x}^*(t_f), t_f) = 0$	$(2n + 1)$ equations to determine the $2n$ constants of integration and t_f
	5. $\mathbf{x}(t_f)$ free		$\mathbf{x}^*(t_0) = \mathbf{x}_0$ $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f), t_f) - \mathbf{p}^*(t_f) = \mathbf{0}$ $\mathscr{H}(\mathbf{x}^*(t_f), \mathbf{u}^*(t_f), \mathbf{p}^*(t_f), t_f) + \frac{\partial h}{\partial t}(\mathbf{x}^*(t_f), t_f) = 0$	$(2n + 1)$ equations to determine the $2n$ constants of integration and t_f

From Kirk, Section 5.1

Interlude: Pontryagin's minimum principle
(with bounded controls)

- Assuming bounded controls $\mathbf{u} \in U$, the necessary optimality conditions are (H is the Hamiltonian)

$$\dot{\mathbf{x}}^*(t) = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t)$$

$$\dot{\mathbf{p}}^*(t) = -\frac{\partial H}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t)$$

$$H(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \leq H(\mathbf{x}^*(t), \mathbf{u}(t), \mathbf{p}^*(t), t), \text{ for all } \mathbf{u}(t) \in U$$

for all $t \in [t_0, t_f]$

along with the boundary conditions:

$$\left[\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f), t_f) - \mathbf{p}^*(t_f)\right]^T \delta \mathbf{x}_f + \left[H(\mathbf{x}^*(t_f), \mathbf{u}^*(t_f), \mathbf{p}^*(t_f), t_f) + \frac{\partial h}{\partial t}(\mathbf{x}^*(t_f), t_f)\right] \delta t_f = 0$$

Necessary conditions for optimal control
(with unbounded controls)

- Necessary conditions consist of a set of $2n$, *first-order*, differential equations (state and costate equations), and a set of m algebraic equations (control equations)
- The solution to the state and costate equations will contain $2n$ constants of integration
- To obtain values for the constants, we use the n equations $\mathbf{x}(t_0) = \mathbf{x}_0$, and an additional set of n (or $n + 1$) equations from the boundary conditions
- Once again: *2-point boundary value problem*

Example

Find optimal control $u(t)$ to steer the system

$$\ddot{x}(t) = u(t)$$

from $x(0) = 10, \dot{x}(0) = 0$ to the origin $x(t_f) = 0, \dot{x}(t_f) = 0$, and to minimize

$$J = \frac{1}{2}\alpha t_f^2 + \frac{1}{2}\int_{t_0}^{t_f} b\, u^2(t) dt\,, \quad \alpha, b > 0$$

(note: the final time t_f is free)

Example

Find optimal control $u(t)$ to steer the system

$$\ddot{x}(t) = u(t)$$

from $x(0) = 10, \dot{x}(0) = 0$ to the origin $x(t_f) = 0, \dot{x}(t_f) = 0$, and to minimize

$$J = \frac{1}{2}\alpha t_f^2 + \frac{1}{2}\int_{t_0}^{t_f} b\, u^2(t)dt\,, \quad \alpha, b > 0$$

- Solution: optimal time is

$$t_f = \left(\frac{1800b}{\alpha}\right)^{1/5}$$

Necessary conditions for optimal control
(with unbounded controls)

We want to prove that, with unbounded controls, the necessary optimality conditions are ($H = g + \mathbf{p}^{\mathrm{T}}\mathbf{f}$ is the Hamiltonian)

$$\left.\begin{aligned} \dot{\mathbf{x}}^*(t) &= \frac{\partial H}{\partial \mathbf{p}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \\ \dot{\mathbf{p}}^*(t) &= -\frac{\partial H}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \\ \mathbf{0} &= \frac{\partial H}{\partial \mathbf{u}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \end{aligned}\right\} \text{ for all } t \in [t_0, t_f]$$

along with the boundary conditions:

$$\left[\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}^*(t_f), t_f) - \mathbf{p}^*(t_f)\right]^T \delta\mathbf{x}_f + \left[H(\mathbf{x}^*(t_f), \mathbf{u}^*(t_f), \mathbf{p}^*(t_f), t_f) + \frac{\partial h}{\partial t}(\mathbf{x}^*(t_f), t_f)\right]\delta t_f = 0$$

Proof sketch of NOC

- For simplicity, assume that the terminal penalty is equal to zero, and that t_f and $\mathbf{x}(t_f)$ are fixed and given
- Consider the augmented cost function

$$g_a(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{p}(t), t) \coloneqq g(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{p}(t)^T [\mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) - \dot{\mathbf{x}}(t)]$$

  where the $\{p_i(t)\}$'s are Lagrange multipliers

- Note that we have simply added zero to the cost function!
- The augmented cost function is then

$$J_a(\mathbf{u}) = \int_{t_0}^{t_f} g_a(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{p}(t), t) \; dt$$

Proof sketch of NOC

On an extremal, by applying the fundamental theorem of the CoV

By the CoV theorem

$$0 = \delta J_a(\mathbf{u}) = \int_{t_0}^{t_f} \Bigg(\left[\frac{\partial g_a}{\partial \mathbf{x}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) - \frac{d}{dt}\frac{\partial g_a}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{x}(t)$$
$$+ \left[\frac{\partial g_a}{\partial \mathbf{u}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{u}(t) + \left[\frac{\partial g_a}{\partial \mathbf{p}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{p}(t) \Bigg) dt$$

Proof sketch of NOC

On an extremal, by applying the fundamental theorem of the CoV

By the CoV theorem

$$= \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t) + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t)^T \mathbf{p}^*(t) \qquad = -\frac{d}{dt}(-\mathbf{p}^*(t))$$

$$0 = \delta J_a(\mathbf{u}) = \int_{t_0}^{t_f} \left(\left[\frac{\partial g_a}{\partial \mathbf{x}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) - \frac{d}{dt}\frac{\partial g_a}{\partial \dot{\mathbf{x}}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{x}(t) \right.$$

$$\left. + \left[\frac{\partial g_a}{\partial \mathbf{u}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{u}(t) + \left[\frac{\partial g_a}{\partial \mathbf{p}}(\mathbf{x}^*(t), \dot{\mathbf{x}}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t), t) \right]^T \delta \mathbf{p}(t) \right) dt$$

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

Proof sketch of NOC

Considering each term in sequence,

- $\mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t), t) - \dot{\mathbf{x}}^*(t) = \mathbf{0}$, on an extremal
- The Lagrange multipliers are arbitrary, so we can select them to make the coefficient of $\delta\mathbf{x}(t)$ equal to zero, that is
$$\dot{\mathbf{p}}^*(t) = -\frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t) - \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t)^T \mathbf{p}^*(t)$$
- The remaining variation $\delta\mathbf{u}(t)$, is independent, so its coefficient must be zero; thus
$$\frac{\partial g}{\partial \mathbf{u}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t) + \frac{\partial \mathbf{f}}{\partial \mathbf{u}}(\mathbf{x}^*(t), \mathbf{u}^*(t), t)^T \mathbf{p}^*(t) = \mathbf{0}$$

By using the Hamiltonian formalism, one obtains the claim

Next time

- Pontryagin's Minimum Principle
- Intro to dynamic programming