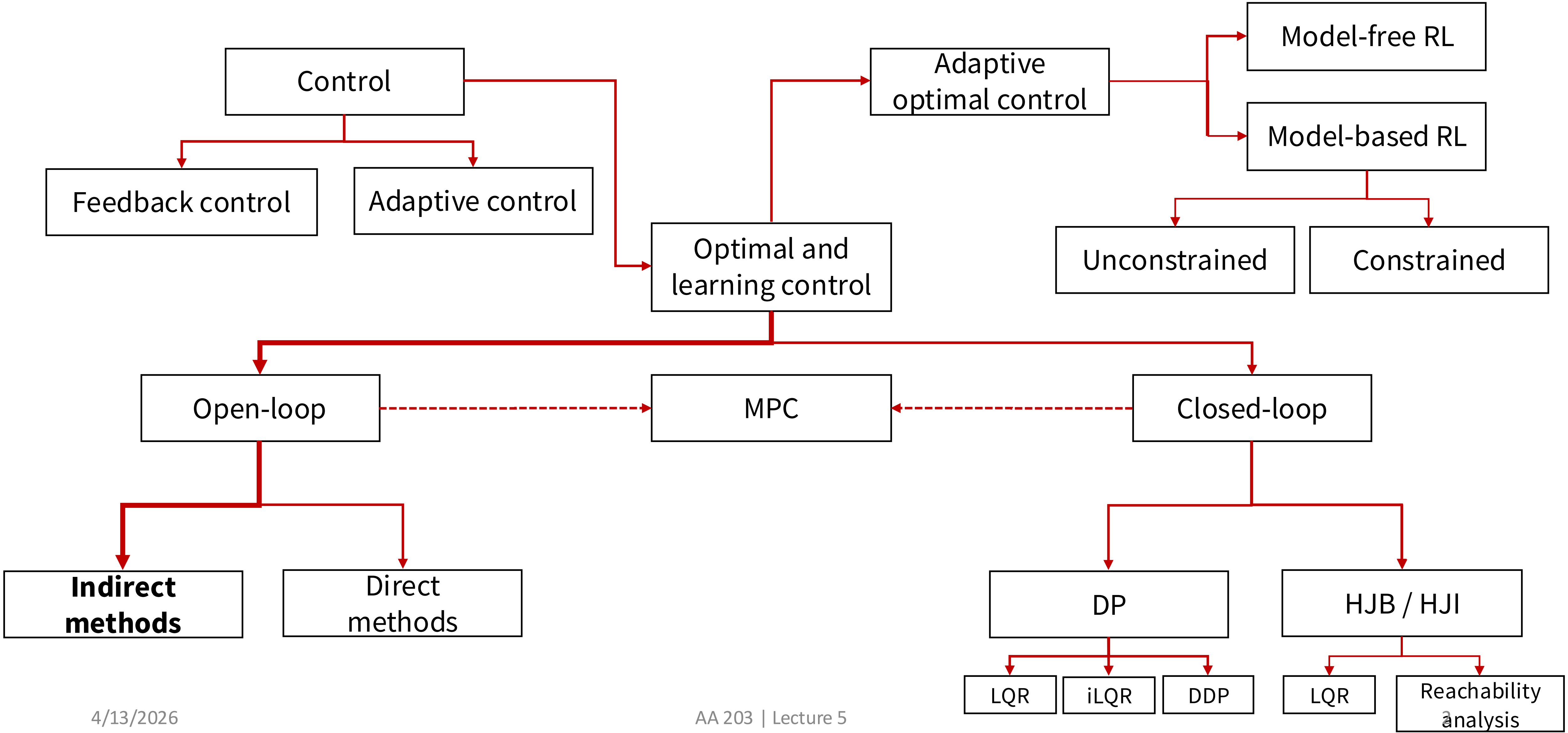
AA203
Optimal and Learning-based Control

Pontryagin's Minimum Principle (PMP);
computational methods

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

Outline

- Necessary conditions for optimal control with bounded controls:
  - Pontryagin's Minimum Principle (PMP)
- Examples: Applications of PMP (and insights we can derive from the analysis)
- Computational methods

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 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 bounded controls)

- So far, we have assumed that the admissible controls and states are not constrained by any boundaries
- However, in realistic systems, such constraints do commonly occur
  - control constraints often occur due to actuation limits
  - state constraints often occur due to safety considerations
- We will now consider the case with control constraints, which will lead to the statement of the Pontryagin's minimum principle

Why do control constraints complicate the analysis?

- By definition, the control $\mathbf{u}^*$ causes the functional J to have a relative minimum if

$$J(\mathbf{u}) - J(\mathbf{u}^*) = \Delta J \geq 0$$

for all admissible controls "close" to $\mathbf{u}^*$

- If we let $\mathbf{u} = \mathbf{u}^* + \delta\mathbf{u}$, the increment in J can be expressed as

$$\Delta J(\mathbf{u}^*, \delta\mathbf{u}) = \delta J(\mathbf{u}^*, \delta\mathbf{u}) + \text{higher order terms}$$

- The variation $\delta\mathbf{u}$ is arbitrary *only if* the extremal control is strictly within the boundary for all time in the interval $[t_0, t_f]$
- In general, however, an extremal control lies on a boundary during at least one subinterval of the interval $[t_0, t_f]$

Why do control constraints complicate the analysis?

- As a consequence, admissible control variations $\delta \mathbf{u}$ exist whose negatives $(-\delta \mathbf{u})$ are not admissible
- This implies that a necessary condition *for* $\mathbf{u}^*$ to minimize J is

$$\delta J(\mathbf{u}^*, \delta \mathbf{u}) \geq 0$$

for all admissible variations with $||\delta \mathbf{u}||$ small enough

Pontryagin's minimum principle

- 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:

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

Pontryagin's minimum principle

- $\mathbf{u}^*(t)$ is a control that causes $H(\mathbf{x}^*(t), \mathbf{u}(t), \mathbf{p}^*(t), t)$ to assume its *global* minimum
- Harder condition in general to analyze
- Example: consider the system having dynamics:

$$\dot{x}_1(t) = x_2(t), \qquad \dot{x}_2(t) = -x_2(t) + u(t);$$

it is desired to minimize the functional

$$J = \int_{t_0}^{t_f} \frac{1}{2}[x_1^2(t) + u^2(t)]dt$$

subject to the control constraint $|u(t)| \leq 1$ with t_f fixed and the final state free.

Pontryagin's minimum principle

Solution:

- If the control is unconstrained,

$$u^*(t) = -p_2^*(t)$$

- If the control is constrained as $|u(t)| \leq 1$, then

$$u^*(t) = \begin{cases} -1 & \text{for } 1 < p_2^*(t) \\ -p_2^*(t), & -1 \leq p_2^*(t) \leq 1 \\ +1 & \text{for } p_2^*(t) < -1 \end{cases}$$

- To determine $u^*(t)$ explicitly, the state and co-state equations must still be solved

Additional necessary conditions

1. If the final time is fixed and the Hamiltonian does not depend explicitly on time, then

$$H\big(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)\big) = c \quad \text{for all } t \in \left[t_0, t_f\right]$$

2. If the final time is free and the Hamiltonian does not depend explicitly on time, then

$$H\big(\mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}^*(t)\big) = 0 \quad \text{for all } t \in [t_0, t_f]$$

Minimum time problems

- Find the control input sequence

$$M_i^- \leq u_i(t) \leq M_i^+ \text{ for } i = 1, \dots, m$$

that drives the control affine system

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

from an arbitrary state $\mathbf{x}_0$ to the origin, and minimizes time

$$J = \int_{t_0}^{t_f} 1 \, dt$$

Minimum time problems

- Form the Hamiltonian

$$H = 1 + \mathbf{p}(t)^T\{\mathbf{a}(\mathbf{x},t) + B(\mathbf{x},t)\mathbf{u}(t)\}$$

$$= 1 + \mathbf{p}(t)^T\{\mathbf{a}(\mathbf{x},t) + [\mathbf{b}_1(\mathbf{x},t)\ \ \mathbf{b}_2(\mathbf{x},t) \cdots \mathbf{b}_m(\mathbf{x},t)]\mathbf{u}(t)\}$$

$$= 1 + \mathbf{p}(t)^T\mathbf{a}(\mathbf{x},t) + \sum_{i=1}^{m} \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t)u_i(t)$$

- By the PMP, select $u_i(t)$ to minimize H, which gives

$$u_i^*(t) = \begin{cases} M_i^+ \text{ if } \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t) < 0 \\ M_i^- \text{ if } \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t) > 0 \end{cases}$$

"Bang-bang" control

Minimum time problems

- Note: we showed what to do when $\mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) \neq 0$
- Not obvious what to do if $\mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) = 0$
- If $\mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) = 0$ for some finite time interval, then the coefficient of $u_i(t)$ in the Hamiltonian is zero, so the PMP provides no information on how to select $u_i(t)$
- The treatment of such a *singular condition* requires a more sophisticated analysis
- The analysis in the linear case is significantly easier, see Kirk Sec. 5.4

Minimum fuel problems

- Find the control input sequence

$$M_i^- \le u_i(t) \le M_i^+ \text{ for } i = 1, \dots, m$$

that drives the control affine system

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

from an arbitrary state $\mathbf{x}_0$ to the origin in a fixed time, and minimizes

$$J = \int_{t_0}^{t_f} \sum_{i=1}^{m} c_i \, |u_i(t)| \, dt$$

Minimum fuel problems

- Form the Hamiltonian

$$H = \sum_{i=1}^{m} c_i \, |u_i(t)| + \mathbf{p}(t)^T \{\mathbf{a}(\mathbf{x}, t) + B(\mathbf{x}, t)\mathbf{u}(t)\}$$

$$= \sum_{i=1}^{m} c_i \, |u_i(t)| + \mathbf{p}(t)^T \mathbf{a}(\mathbf{x}, t) + \sum_{i=1}^{m} \mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) u_i(t)$$

$$= \sum_{i=1}^{m} [c_i \, |u_i(t)| + \mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) u_i(t)] + \mathbf{p}(t)^T \mathbf{a}(\mathbf{x}, t)$$

- By the PMP, select $u_i(t)$ to minimize H, that is

$$\sum_{i=1}^{m} [c_i \, |u_i^*(t)| + \mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) u_i^*(t)] \leq \sum_{i=1}^{m} [c_i \, |u_i(t)| + \mathbf{p}(t)^T \mathbf{b}_i(\mathbf{x}, t) u_i(t)]$$

Minimum fuel problems

- Since the components of $\mathbf{u}(t)$ are independent, then one can just look at

$$c_i\,|u_i^*(t)| + \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t)u_i^*(t) \le c_i\,|u_i(t)| + \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t)u_i(t)$$

- The resulting control law is

$$u_i^*(t) = \begin{cases} M_i^- & \text{if } c_i < \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t) \\ 0 & \text{if } -c_i < \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t) < c_i \\ M_i^+ & \text{if } \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t) < -c_i \end{cases}$$

"Bang-off-bang" control

Minimum energy problems

- Find the control input sequence

$$M_i^- \leq u_i(t) \leq M_i^+ \text{ for } i = 1, \dots, m$$

that drives the control affine system

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

from an arbitrary state $\mathbf{x}_0$ to the origin in a fixed time, and minimizes

$$J = \frac{1}{2}\int_{t_0}^{t_f} \mathbf{u}(t)^T R\mathbf{u}(t)dt \,,$$

where $R > 0$ and diagonal

Minimum energy problems

- Form the Hamiltonian

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

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

- By the PMP, we need to solve

$$\mathbf{u}^*(t) = \arg\min_{\mathbf{u}(t)\in U}\left[\sum_{i=1}^{m}\frac{1}{2}R_{ii}u_i(t)^2 + \mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x},t)u_i(t)\right]$$

Minimum energy problems

- As in the first example today, in the unconstrained case, the optimal solution for each component of $\mathbf{u}(t)$ would be

$$\hat{u}_i(t) = -R_{ii}^{-1}\,\mathbf{p}(t)^T\mathbf{b}_i(\mathbf{x}, t)$$

- Considering the input constraints, the resulting control law is

$$u^*(t) = \begin{cases} M_i^- & \text{if} \quad \hat{u}_i(t) < M_i^- \\ \hat{u}_i(t) & \text{if} \quad M_i^- < \hat{u}_i(t) < M_i^+ \\ M_i^+ & \text{if} \quad M_i^+ < \hat{u}_i(t) \end{cases}$$

"Saturating" control

Uniqueness and existence

- Note: uniqueness and existence are not in general guaranteed!

- Example 1 (non uniqueness): find a control sequence $u(t)$ to transfer the system $\dot{x}(t) = u(t)$ from an arbitrary initial state x_0 to the origin, and such that the functional $J = \int_0^{t_f} |u(t)| dt$ is minimized. The final time is free, and the admissible controls are $|u(t)| \leq 1$

- Example 2 (non existence): find a control sequence $u(t)$ to transfer the system $\dot{x}(t) = -x(t) + u(t)$ from an arbitrary initial state x_0 to the origin, and such that the functional $J = \int_{t_0}^{t_f} |u(t)| dt$ is minimized. The final time is free, and the admissible controls are $|u(t)| \leq 1$

Computational methods

- Until now, we derived necessary conditions for optimality and *analytically* studied a few special cases
- We now focus on *numerical* techniques to solve two-point boundary value problems; popular methods:
  - Indirect shooting method
  - Indirect collocation method

Let's revisit our 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

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

Indirect methods: practical aspects

To obtain solution to the necessary conditions for optimality, one needs to solve two-point boundary value problems

- In python, we'll be using `scipy.integrate.solve_bvp` to solve problems in "standard" form

$$\dot{z} = g(z, t, \boldsymbol{p}), \qquad BC\left(z(t_0), z(t_f)\right) = 0$$

where $\boldsymbol{p}$ are extra variables that can also be optimized

- Syntax: `sol = solve_bvp(fun, bc, t, z, p=None)`

Example: $\dot{z}_1 = z_2, \quad \dot{z}_2 = -|z_1|, \quad z_1(0) = 0, \quad z_1(4) = -2$

*`solve_bvp` uses a collocation formula (three-stage Lobatto)

Extensions

- What about problems whose necessary conditions to not fit directly the "standard" form (e.g., free end time problems)?
- Handy tricks exist to convert problems into standard form :
  - Ascher, U., & Russell, R. D. (1981). Reformulation of boundary value problems into "standard" form. SIAM review, 23(2), 238-254.

Important case: free final time

1. Rescale time so that $\tau = t/t_f$, then $\tau \in [0,1]$
2. Change derivatives $\frac{d}{d\tau} := t_f \frac{d}{dt}$
3. Introduce dummy state r that corresponds to t_f with dynamics $\dot{r} = 0$
4. Replace all instances of t_f with r

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

1. Define state as $\boldsymbol{z} = [\boldsymbol{x}, \boldsymbol{p}, r]$
2. BC are: $x_1(0) = 10, x_2(0) = 0, x_1(t_f) = 0, x_2(t_f) = 0, -\frac{p_2(t_f)^2}{2b} + \alpha t_f = 0$
3. BVP becomes: $\frac{dz}{d\tau} = t_f \frac{dz}{dt} = z_5 \begin{bmatrix} A & -B[0\ 1]/b & 0 \\ 0 & -A' & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{z}$
4. BC become $z_1(0) = 10, z_2(0) = 0, z_1(1) = 0, z_2(1) = 0, -\frac{z_4(1)^2}{2b} + \alpha z_5(1) = 0$

Next time

- Direct methods