

Methods Of Optimal Control

Arash Fahim

FSU Fall 2025

September 2, 2025

Structure of the course

- Short lecture (10-15 minutes)
- Team work (50 minutes)
 - * Problem solving
 - * Programming (Python)
 - * Managerial skills
 - * Reporting skills
- Delivery (5 minutes)

Structure of the course

Python

Each student is expected to bring a computer to the classroom with a *Python 3.12*, *IPython*, and *Jupyter Notebook* installed.

- <https://www.python.org/downloads>
- <https://www.anaconda.com/docs/main>
- Virtual environment:

<https://docs.python.org/3/library/venv.html>

[https://www.anaconda.com/docs/tools/
working-with-conda/environments](https://www.anaconda.com/docs/tools/working-with-conda/environments)

GitHub

Each student is required to have a *GitHub* account.

Rules of teamwork

- Be respectful.
- Allow all team members to engage and express their opinion. No interruption.
- Arrange chair to make sure everyone is involved.
- No one knows everything. Make sure that everyone learns what you know.
- Break down the tasks between group members based on ones abilities.
- Ask for feedback from me frequently.

Composition of teams

Each team requires one or two members to take the following roles:

- Problem solving
- Programming (Python)
- Managerial skills
- Reporting skills

Optimization versus control

Example1

$\alpha : [0, T] \rightarrow \mathbb{R}$ is given.

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

where the infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$.

Optimization versus control

Example1

$\alpha : [0, T] \rightarrow \mathbb{R}$ is given.

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

where the infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$.

$$\inf_{x \in \mathbb{R}^d} \left\{ x^2 - \alpha_t x \right\}$$

Optimization versus control

Example1

$\alpha : [0, T] \rightarrow \mathbb{R}$ is given.

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

where the infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$.

$$\inf_{x \in \mathbb{R}^d} \left\{ x^2 - \alpha_t x \right\}$$

$$x_t^* = \frac{\alpha_t}{2} = \operatorname{argmin}_{x \in \mathbb{R}^d} \left\{ x^2 - \alpha_t x \right\}$$

Optimization versus control

Dynamic x_t

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

Infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$ such that for some function $u : [0, T] \rightarrow \mathbb{R}$

$$dx_t = (-\beta x_t + u_t)dt, \quad x_0 = x$$

Optimization versus control

Dynamic x_t

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

Infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$ such that for some function $u : [0, T] \rightarrow \mathbb{R}$

$$dx_t = (-\beta x_t + u_t) dt, \quad x_0 = x$$

Can we find u_t such that $x_t = \frac{\alpha_t}{2}$? (For simplicity, take $\beta = 0$.)

Optimization versus control

Dynamic x_t

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

Infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$ such that for some function $u : [0, T] \rightarrow \mathbb{R}$

$$dx_t = (-\beta x_t + u_t) dt, \quad x_0 = x$$

Can we find u_t such that $x_t = \frac{\alpha_t}{2}$? (For simplicity, take $\beta = 0$.)

Check it for $\alpha_t = \mathbb{1}_{\{0 \leq t \leq \frac{T}{2}\}}$.

Optimization versus control

Dynamic x_t

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\}$$

Infimum is over all functions $x : [0, T] \rightarrow \mathbb{R}$ such that for some function $u : [0, T] \rightarrow \mathbb{R}$

$$dx_t = (-\beta x_t + u_t)dt, \quad x_0 = x$$

Can we find u_t such that $x_t = \frac{\alpha_t}{2}$? (For simplicity, take $\beta = 0$.)

Check it for $\alpha_t = \mathbb{1}_{\{0 \leq t \leq \frac{T}{2}\}}$.

For $\alpha_t = \mathbb{1}_{\{0 \leq t \leq \frac{T}{2}\}}$, what is the value of the infimum? Is it

$$\inf \left\{ \int_0^T (x_t^2 - \alpha_t x_t) dt \right\} = -\frac{1}{4} \int_0^T \alpha_t^2 dt = -\frac{T}{8}?$$

Optimization versus control

A control problem without a myopic solution

$$\inf \int_0^T \left(x_t^2 - \alpha_t x_t + u_t^2 \right) dt \quad (1)$$

Infimum is over all functions $u : [0, T] \rightarrow \mathbb{R}$

$$dx_t = (-\beta x_t + u_t) dt, \quad x_0 = x$$

Trade-off:

- Trying to send $x_t \rightarrow \frac{\alpha_t}{2}$ may cause $\int_0^T u_t^2 dt$ to grow.
- Trying to keep cost $\int_0^T u_t^2 dt$ near zero, does not bring x_t close to $\frac{\alpha_t}{2}$.

What is the sweet spot for u_t ?

A generic control problem

Definition

$$\inf_{u \in \mathcal{U}} \int_0^T C(t, x_t, u_t) dt + g(x_T), \quad dx_t = f(x_t, u_t) dt \quad (2)$$

- $C : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}$: *running cost*
- $g : \mathbb{R}^d \rightarrow \mathbb{R}$: *terminal cost*
- \mathcal{U} : *an admissible set of functions $u : [0, T] \rightarrow \mathbb{R}^n$, control variable.*

A generic control problem

Admissible controls

\mathcal{U} is chosen to fit the proper application and/or to make the control problem wellposed.

Admissibility for wellposedness

$$\inf_{u \in \mathcal{U}} \int_0^T (x_t - u_t^2) dt = -\infty, \quad dx_t = (x_t - u_t) dt \quad (3)$$

\mathcal{U} to be the set of all functions $u : [0, T] \rightarrow \mathbb{R}$

If we restrict \mathcal{U} to the set of functions $u : [0, T] \rightarrow [-1, \infty)$ (some lower bound on the value), then

$$\inf_{u \in \mathcal{U}} \int_0^T (x_t - u_t^2) dt > -\infty, \quad dx_t = (x_t - u_t) dt. \quad (4)$$

Infinite horizon

Infinite horizon

An infinite horizon control problem is accommodated by setting $T = \infty$. For example,

$$\inf_{u \in \mathcal{U}} \int_0^{\infty} e^{-t} (x_t^2 + u_t^2) dt, \quad C(t, x, u) = e^{-t} (x^2 + u^2) \quad (5)$$

Exercise

Write the following problem as a generic control problems by associating the horizon T , the running cost $C(t, x, u)$ and terminal cost $g(x)$

(Shortest time to exit a bounded domain) Given a bounded domain $D \subset \mathbb{R}^d$, find

$$\inf_u \{t \geq 0 : x_t \notin D\} \quad (6)$$

where $dx_t = u_t dt$ with control $|u_t| \leq 1$ and $u_t \in \mathbb{R}^d$ and initial position $x_0 = x \in D$.

Infinite horizon

Exercise

Write the following problem as a generic control problems by associating the horizon T , the running cost $C(t, x, u)$ and terminal cost $g(x)$

(Shortest time to exit a bounded domain) Given a bounded domain $D \subset \mathbb{R}^d$, find

$$\inf_u \{t \geq 0 : x_t \notin D\} \quad (7)$$

where $dx_t = u_t dt$ with control $|u_t| \leq 1$ and $u_t \in \mathbb{R}^d$ and initial position $x_0 = x \in D$.

Solution

$$\inf_u \int_0^\infty \mathbb{1}_{\{x_t \in D\}} dt, \quad dx_t = u_t dt \quad \text{with} \quad |u_t| \leq 1$$

An optimal control is described by existing D as fast as possible, $|u| = 1$, and stop as soon as we exit, $|u| = 0$.

Dynamic programming principle (DPP)

Value function

Fix $x_t = x$.

$$V(t, x) := \inf_{u \in \mathcal{U}_t} \int_t^T C(s, x_s, u_s) ds + g(x_T), \quad dx_s = f(x_s, u_s) ds$$

\mathcal{U}_t : the set of admissible controls restricted to $[t, T]$.

Dynamic programming principle (DPP)

$$V(t, x) = \inf_{u \in \mathcal{U}_{t,s}} \int_t^s C(r, x_r, u_r) dr + V(s, x_s), \quad dx_r = f(x_r, u_r) dr$$

$\mathcal{U}_{t,s}$: the set of admissible controls restricted to $[t, s]$.

DPP

Balance of cost in DPP

$$V(t, x) = \inf_{u \in \mathcal{U}_{t,s}} \int_t^s C(r, x_r, u_r) dr + V(s, x_s), \quad x_s = x + \int_t^s f(x_r, u_r) dr$$

Proof of DPP

$$\begin{aligned} V(t, x) &= \inf_{u \in \mathcal{U}_t} \int_t^T C(r, x_r, u_r) dr + g(x_T) \\ &= \inf_{u_1 \in \mathcal{U}_{t,s}} \inf_{u_2 \in \mathcal{U}_s} \int_t^s C(r, x_r, u_1(r)) dr + \int_s^T C(r, x_r, u_2(r)) dr + g(x_T) \\ &= \inf_{u_1 \in \mathcal{U}_{t,s}} \int_t^s C(r, x_r, u_1(r)) dr + \inf_{u_2 \in \mathcal{U}_s} \int_s^T C(r, x_r, u_2(r)) dr + g(x_T) \end{aligned}$$

Note that $\inf_{u_2 \in \mathcal{U}_s} \int_s^T C(r, x_r, u_2(r)) dr + g(x_T) = V(s, x_s)$. Therefore,

$$V(t, x) = \inf_{u_1 \in \mathcal{U}_{t,s}} \int_t^s C(r, x_r, u_1(r)) dr + V(s, x_s)$$

Numerical DPP

Discretization


$$t_0 \quad t_1 \quad t_2 \quad \cdots \quad t_{N-1} \quad t_N = T \quad \Delta t = \frac{T}{N}$$

$$\begin{cases} V(t_N, x) = g(x) \\ V(t_i, x) = \inf_{u \in \mathcal{U}_{t_i, t_{i+1}}} \int_{t_i}^{t_{i+1}} C(r, x_r, u_r) dr + V(t_{i+1}, x_{t_{i+1}}), \\ x_{t_{i+1}} = x + \int_{t_i}^{t_{i+1}} f(s, x_s, u_s) ds \end{cases}$$

Approximation

$$\begin{cases} \hat{V}(t_N, x) = g(x) \\ \hat{V}(t_i, x) := \inf_u C(t_i, x_{t_i}, u) \Delta t + \hat{V}(t_{i+1}, x_{t_{i+1}}), \quad x_{t_{i+1}} = x + f(t_i, x, u) \Delta t \end{cases}$$

Numerical DPP

Approximation

$$\begin{cases} \hat{V}(t_N, x) = g(x) \\ \hat{V}(t_i, x) := \inf_u C(t_i, x_{t_i}, u)\Delta t + \hat{V}(t_{i+1}, x_{t_{i+1}}), \quad x_{t_{i+1}} = x + f(t_i, x, u)\Delta t \end{cases}$$

Simplification of one-step approximate DPP

The approximation is not over the control $u : [t_i, t_{i+1}] \rightarrow \mathbb{R}^m$, but over values $u \in \mathbb{R}^m$. The optimal value \hat{u}^* is a constant approximately optimal control over $[t_i, t_{i+1}]$.

Algorithm

Algorithm 1: Numerical DPP

Parameter $T, N, f(t, x, u), C(t, x, u),$ and $g(x);$

:

$$\Delta t = \frac{T}{N}$$

Data: $\hat{V}(t_N, x) = g(x);$

x_i^j for $j = 1, \dots, J$ and $i = 0, \dots, N - 1$

1 **for** $i \leftarrow N - 1$ **to** 0 **do**

2 $\hat{x}_{i+1}^j = x_i^j + f(t_i, x_i^j, u)\Delta t;$

3 $\tilde{V}(t_i, x_i^j) \leftarrow \inf_u C(t_i, x_i^j, u)\Delta t + \hat{V}(t_{i+1}, \hat{x}_{i+1}^j);$

4 $\hat{V}(t_i, x)$ obtained from interpolation on $\tilde{V}(t_i, x_i^j)$ for $j = 1, \dots, J;$

5 $\hat{u}^*(x_i^j) \in \underset{u}{\operatorname{argmin}} C(t_i, x_i^j, u)\Delta t + \hat{V}(t_{i+1}, \hat{x}_{i+1}^j);$

6 **return** \hat{V} and $\hat{u}^*.$

DPP algorithm

Exercise

Why interpolation is required in Algorithm 1?

Can we perform the algorithm by only knowing $\hat{V}(t_{i+1}, x_{i+1}^j)$ for all $j = 1, \dots, J$?

Note the difference between $\hat{V}(t_{i+1}, x_{i+1}^j)$ and $\hat{V}(t_{i+1}, \hat{x}_{i+1}^j)$ and the difference between x_{i+1}^j and $\hat{x}_{i+1}^j = x_i^j + f(t_i, x_i^j, u)\Delta t$.

$$\inf_u C(t_i, x_i^j, u) + \hat{V}(t_i, x_i^j + f(t_i, x_i^j, u)\Delta t)$$

Quadratic example

Example

Value function:

$$V(t, x) := \inf_{u \in \mathcal{U}_t} \int_t^T (x_s^2 + u_s^2) ds + \frac{1}{2}x_T^2 - x_T, \quad dx_s = (x_s - u_s)ds. \quad (8)$$

We cannot find value functions using a myopic argument.

Exercise

- 1) In example above, write the approximate DPP from time t_i to t_{i+1} .
- 2) Assume that $\hat{V}(t_{i+1}, x) = a_{i+1}x^2 + b_{i+1}x + c_{i+1}$ for some known values a_{i+1} , b_{i+1} , and c_{i+1} .
- 3) Use optimization of a quadratic function to find $\hat{V}(t_i, x)$. Note that you need to use $\hat{x}_{t_{i+1}} = x + (x - u)\Delta t$.
- 4) Does $\hat{V}(t_i, x)$ is of the form $a_ix^2 + b_ix + c_i$? What is the relation between (a_i, b_i, c_i) and $(a_{i+1}, b_{i+1}, c_{i+1})$?

Individual project

Due end October

In your area of study, find an optimal control problem. Then, write the cost functions and the control variable and determine a set of admissible controls.