Chapter 1

Introduction to optimal control

1.1 Optimal control problem formulation

Consider the continuous-time system ($t \in \mathbb{R}$)

$$\dot{x}(t) = f(x(t), u(t), t) \tag{1.1}$$
$$y(t) = h(x(t), u(t), t) \tag{1.2}$$

- $x(t) \in \mathbb{R}^n$ state of the system at time t
- $u(t) \in \mathbb{R}^m$ input of the system at time t
- $y(t) \in \mathbb{R}^p$ output of the system at time t

We will mainly work with time invariant systems, $\dot{x}(t) = f(x(t), u(t))$.

We consider nonlinear, discrete-time systems described by

$$x(t+1) = f_t(x(t), u(t)) \quad t \in \mathbb{N}_0$$

but from now on we will use the compact notation

$$x_{t+1} = f_t(x_t, u_t) \quad t \in \mathbb{N}_0$$

where $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ are the state and the input of the system at time t.

Consider a nonlinear, discrete-time system on a finite time horizon

$$x_{t+1} = f_t(x_t, u_t) \quad t = 0, \ldots, T-1$$

We use $\mathbf{x} \in \mathbb{R}^{nT}$ and $\mathbf{u} \in \mathbb{R}^{mT}$ to denote, respectively, the stack of the states x_t for all $t \in \{1, \ldots, T\}$ and the unputs u_t for all $t \in \{0, \ldots, T-1\}$, that is:

$$\mathbf{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_T \end{bmatrix} \qquad \mathbf{u} := \begin{bmatrix} u_0 \\ \vdots \\ u_{T-1} \end{bmatrix}$$

Trajectory of a system

Definition: A pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{R}^{nT} \times \mathbb{R}^{mT}$ is called a trajectory of system (1) if $\bar{x}_{t+1} = f_t(\bar{x}_t, \bar{u}_t)$ for all $t \in \{0, \ldots, T-1\}$., That is, if $\bar{\mathbf{x}}, \bar{\mathbf{u}})$ satisfies the system dynamics (the same holds for continuous time systems with proper adjustments). In particular, $\bar{\mathbf{x}}$ is the state trajectory, while $\bar{\mathbf{u}}$ is the input trajectory.

Equilibrium

Definition: A state-input pair $(x_e, u_e) \in \mathbb{R}^n \times \mathbb{R}^m$ is called an equilibrium pair of (1) if $(x_t, u_t) = (x_e, u_e), t \in \mathbb{N}_0$ is a trajectory of the system.
Equilibria of time-invariant systems satisfy $x_e = f(x_e, u_e)$

Linearization of a system about a trajectory

Given the dynamics (1)and a trajectory $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$, the linearization of (1) about $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is given by the linear (possibly) time-varying system

$$\Delta x_{t+1} = A_t \Delta x_t + B_t \Delta u_t \quad t \in \mathbb{N}_0$$

with A_t and B_t the Jacobians of f_t, with respect to state and input respectively, evaluated at $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$A_t = \frac{\partial}{\partial x} f(\bar{x}_t, \bar{u}_t)\Big|_{(\bar{\mathbf{x}}, \bar{\mathbf{u}})} \qquad B_t = \frac{\partial}{\partial u} f(\bar{x}_t, \bar{u}_t)\Big|_{(\bar{\mathbf{x}}, \bar{\mathbf{u}})}$$

1.1.1 Optimization

Main ingredients

- Decision variable: $x \in \mathbb{R}^n$
- Cost function: $\ell(x) : \mathbb{R}^n \to \mathbb{R}$ cost associated to decision x
- Constraints (constraint sets): for some given functions $h_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m,$ and $g_j : \mathbb{R}^n \to \mathbb{R}$, the decision vector $x \in \mathbb{R}^n$ needs to satisfy

  $$h_i(x) = 0 \quad i = 1, \ldots, m g_j(x) = 0 \quad j = 1, \ldots, r$$

  equivalently we can say that we require $x \in X$ with

  $$X = \{x \in \mathbb{R}^n | h(x) = 0, g(x) \leq 0\},$$

  where we compactly denoted $h(x) = \text{col}(h_1(x), \ldots, h_m(x))$ and $g(x) = \text{col}(g_1(x), \ldots, g_r(x))$

Minimization

We can write our optimization problem as

$$\min_{x \in \mathbb{R}^n} \ell(x) \tag{1.3}$$

$$\text{subj. to } h_i(x) = 0 \quad i = 1, \ldots, m \tag{1.4}$$

$$g_j(x) \leq 0 \quad j = 1, \ldots, r \tag{1.5}$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ and $g_j : \mathbb{R}^n \to \mathbb{R}$
We can write it more compactly as

$$\min_{x \in \mathbb{R}^n} \ell(x)$$
$$\text{subj. to } h(x) = 0 g(x) \leq 0$$

where $h : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^r$

1.1.2 Discrete-time optimal control

main ingredients

- Dynamics: a discrete-time system in state space form

  $$x_{t+1} = f_t(x_t, u_t) \quad t = 0, 1, \ldots, T-1$$

- the dynamics introduce T equality constraints

  $$\begin{array}{lll} x_1 = f(x_0, u_0) & \text{i.e.} & x_1 - f_t(x_0, u_0) = 0 \\ x_2 = f(x_1, u_1) & \text{i.e.} & x_1 - f_t(x_1, u_1) = 0 \\ \vdots & & \\ x_T = f(x_{T-1}, u_{T-1}) & \text{i.e.} & x_T - f_t(x_{T-1}, u_{T-1}) = 0 \end{array}$$

  This is equivalent to nT scalar constraints

- Cost function: a cost "to be payed" for a chosen trajectory. We consider an additive structure in time

$$\ell(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T)$$

where $\ell_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is called stage-cost, while $\ell_T : \mathbb{R}^n \to \mathbb{R}$ is the terminal cost.

- End-point constraints: function of the state variable prescribed at initial and/or final point

$$r(x_0, x_T) = 0$$

- Path constraints: point-wise (in time) constraints representing possible limits on states and inputs at each time t

$$g_t(x_t, u_t) \leq 0, \quad t \in \{0, \ldots, T-1\}$$

A discrete-time optimal control problem can be written as

$$\begin{aligned}
\min_{\substack{x_0, x_1, \ldots, x_T \\ u_0, \ldots, u_{T-1}}} \; & \sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T) \\
\text{subj. to} \quad & x_{t+1} = f_t(x_t, u_t), \quad t \in \{0, \ldots, T-1\} \\
& r(x_0, x_T) = 0 \\
& g_t(x_t, u_t) \leq 0, \quad t \in \{0, \ldots, T-1\}
\end{aligned}$$

Optimal control for trajectory generation

We can pose a trajectory generation problem as

$$\min_{\mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^m} \sum_{t=0}^{T-1} \frac{1}{2} \|x_t - x_t^{\text{des}}\|_Q^2 + \frac{1}{2} \|u_t - u_t^{\text{des}}\|_R^2 + \frac{1}{2} \|x_T - x_T^{\text{des}}\|_{P_f}^2$$

Continuous-time Optimal Control problem

A continuous-time optimal control problem, i.e., $t \in \mathbb{R}$ can be written as

$$\begin{aligned}
\min_{(x(\cdot), u(\cdot)) \in \mathcal{F}} \; & \int_0^T \ell_\tau(x(\tau), u(\tau)) d\tau + \ell_T(x(T)) \\
\text{subj. to} \quad & \dot{x}(t) = f_t(x(t), u(t)) \quad t \in [0, T] \\
& r(x(0), x(T)) = 0 \\
& g_t(x(t), u(t)) \leq 0 \quad t \in [0, T)
\end{aligned}$$

Note that $\mathcal{F}$ is a space of functions (function space). This is an infinite dimensional optimization problem

- Cost functional $\ell : \mathcal{F} \to \mathbb{R}$

$$\ell(x(\cdot), u(\cdot)) = \int_0^T \ell_\tau(x(\tau), u(\tau)) d\tau + \ell_T(x(T))$$

- Space of trajectories (or trajectory manifold)

$$\mathcal{T} = \{(x(\cdot), u(\cdot)) \in \mathcal{F} | \dot{x}(t) = f_t(x(t), u(t)), t \geq 0\}$$

Chapter 2

Nonlinear Optimization

2.1 Unconstrained Optimization

Consider the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \ell(x)$$

with $\ell : \mathbb{R}^n \to \mathbb{R}$ a cost function to be minimized and x a decision vector

We say that x^* is a

- global minimum if $\ell(x^*) \leq \ell(x)$ for all $x \in \mathbb{R}^n$
- strict global minimum if $\ell(x^*) < \ell(x)$ for all $x \neq x^*$
- local minimum if there exists $\epsilon > 0$ such that $\ell(x^*) \leq \ell(x)$ for all $x \in B(x^*, \epsilon) = \{x \in \mathbb{R}^n | \|x - x^*\| < \epsilon\}$
- strict local minimum if there exists $\epsilon > 0$ such that $\ell(x^*) < \ell(x)$ for all $x \in B(x^*, \epsilon)$

Notation

We denote $\ell(x^*)$ the optimal (minimum) value of a generic optimization problem, i.e.

$$\ell(x^*) = \min_{x \in \mathbb{R}^n} \ell(x)$$

where x^* is the minimum point (optimal value for the optimization variable) i.e.

$$x^* = \arg\min_{x \in \mathbb{R}^n} \ell(x)$$

Gradient and Hessian

Gradient of a function: for a function $r : \mathbb{R}^n \to \mathbb{R}$ the gradient is denoted as

$$\nabla r(x) = \begin{bmatrix} \frac{\partial r(x)}{\partial x_1} \\ \vdots \\ \frac{\partial r(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

Hessian matrix of a function: for a fcuntion $r : \mathbb{R}^n \to \mathbb{R}$ the Hessian matrix is denoted as

$$\nabla^2(r(x)) = \begin{bmatrix} \frac{\partial^2 r(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 r(x)}{\partial x_1 x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 r(x)}{\partial x_n x_1} & \cdots & \frac{\partial^2 r(x)}{\partial x_1^2} \end{bmatrix}$$

Gradient of a vector-valued function: for a vector field $r : \mathbb{R}^n \to \mathbb{R}^m$, the gradient is denoted as

$$\nabla r(x) = \begin{bmatrix} \nabla r_1(x) & \cdots & \nabla r_m(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial r_1(x)}{\partial x_1} & \cdots & \frac{\partial r_m(x)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_1(x)}{\partial x_n} & \cdots & \frac{\partial r_m(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times m}$$

which is the transpose of the Jacobian matrix of r

2.1.1 Conditions of optimality

First order necessary condition (FNC) of optimality (unconstrained)

Let x^* be an unconstrained local minimum of $\ell : \mathbb{R}^n \to \mathbb{R}$ and assume that ℓ is continuously differentiable ($\mathcal{C}^1$) in $B(x^*, \epsilon)$ for some $\epsilon > 0$. Then $\nabla \ell(x^*) = 0$

Second order necessary condition (FNC) of optimality (unconstrained)

If additionally ℓ is twice continuously differentiable ($\mathcal{C}^2$) in $B(x^*, \epsilon)$, then $\nabla^2 \ell(x^*) \geq 0$ (The Hessian of ℓ is positive semidifinite)

Second order sufficient conditions of optimality (unconstrained)

Let $\ell : \mathbb{R}^n \to \mathbb{R} \in \mathcal{C}^2$ in $b(x^*, \epsilon)$ for some $\epsilon > 0$. Suppose that $x^* \in \mathbb{R}^n$ satisfies

$$\nabla \ell(x^*) = 0 and \nabla^2 \ell(x^*) > 0$$

Then x^* is a strict (unconstrained) local minimum of ℓ

Convex set

A set $X \subset \mathbb{R}^n$ is convex if for any two points x_A and x_B in X and for all $\lambda \in [0, 1]$, then

$$\lambda x_a + (1 - \lambda) x_B \in X$$

Convex functions

Let $X \subset \mathbb{R}^n$ be a convex set. A function $\ell : X \to \mathbb{R}$ is convex if for any two points x_A and x_B in X and for all $\lambda \in [0, 1]$, then

$$\ell(\lambda x_A + (1 - \lambda) x_B) \leq \lambda \ell(x_A) + (1 - \lambda) \ell(x_B)$$

2.1.2 Minimization of convex functions

Proposition

Let $X \subset \mathbb{R}^n$ be a convex set and $\ell : X \to \mathbb{R}$ a convex function. Then a local minimum of ℓ is also a global minimum
Proof: not done in class but present in slides for funsies

Necessary and sufficient condition of optimality (unconstrained)

For the unconstrained minimization of a convex function it can be shown that the first order necessary condition of optimality is also sufficient (for a global minimum).

Proposition

Let $\ell^n_{\mathbb{R}} \to \mathbb{R}$ be a convex function. Then x^* is a global minimum if and only if $\nabla \ell(x^*) = 0$

Proof: not done in class but present in slides for funsies

Chapter 3

Optimality conditions for optimal control

Chapter 4

Linear Quadratic (LQ) optimal control

Chapter 5

Dynamic Programming

Chapter 6

Numerical methods for nonlinear optimal control

Chapter 7

Optimization-based predictive control