Aalto University

Discrete-time optimal control

CHEM-E7225 (was E7195), 2023

Francesco Corona (¬_¬)

Chemical and Metallurgical Engineering
School of Chemical Engineering

Overview

We combine the notions on dynamic systems and simulation with the notions on nonlinear programming, to formulate a general **discrete-time optimal control** problem

- We understand and treat them as special forms of nonlinear programs

Overview (cont.)

Consider a system f which maps an initial state vector x_k onto a final state vector x_{k+1}

- We also consider the presence of a control u_k that modifies the transition

$$x_{k+1} = f\left(x_k, u_k | \theta_x\right), \quad (k = 0, 1, \dots, K-1)$$

We consider transitions over a time-horizon, from time $k = 0$ to time $k = K$

$$0 \cdots 1 \cdots\cdots (k-1) \cdots k \cdots (k+1) \cdots\cdots (K-1) \cdots K$$

Over the time-horizon of interest, we thus have the sequences

⇝ States $\{x_k\}_{k=0}^{K}$, with $x_k \in \mathcal{R}^{N_x}$

⇝ Controls $\{u_k\}_{k=0}^{K-1}$, with $u_k \in \mathcal{R}^{N_u}$

For notational simplicity, we used time-invariant dynamics f

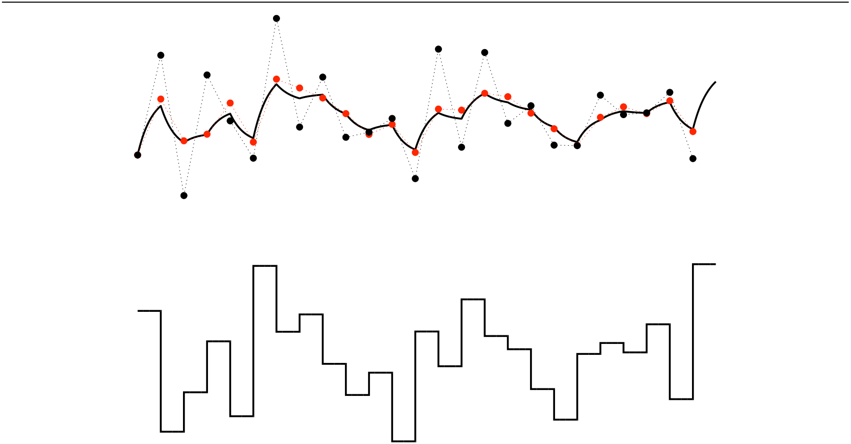
- In general, we have $x_{k+1} = f_k\left(x_k, u_k | \theta_x\right)$

Overview (cont.)

$$x_{k+1} = f\left(x_k, u_k | \theta_x\right), \quad (k = 0, 1, \dots, K-1)$$

The dynamics f are often derived from the discretisation of a continuous-time system

- As result of a numerical integration schemes, under piecewise constant controls



Overview (cont.)

$$x_{k+1} = f\left(x_k, u_k | \theta_x\right), \quad (k = 0, 1, \dots, K-1)$$

Given an initial state x_0 and any sequence of controls $\{u_k\}_{k=0}^{K-1}$, we know all the states

The forward simulation function determines the sequence of states $\{x_k\}_{k=0}^{K}$

$$\begin{aligned} f_{\text{sim}} &: \mathcal{R}^{N_x + (K \times N_u)} \rightarrow \mathcal{R}^{(K+1)N_x} \\ &: \left(x_0, u_0, u_1, \dots, u_{K-1}\right) \mapsto \left(x_0, x_1, \dots, x_K\right) \end{aligned}$$

For arbitrary systems, the forward simulation map is built recursively

$$\begin{aligned} x_0 &= x_0 \\ x_1 &= f\left(x_0, u_0\right) \\ x_2 &= f\left(x_1, u_1\right) \\ &= f\left(f\left(x_0, u_0\right), u_1\right) \\ x_3 &= f\left(x_2, u_2\right) \\ &= f\left(f\left(f\left(x_0, u_0\right), u_1\right), u_2\right) \\ \cdots &= \cdots \end{aligned}$$

Overview (cont.)

$$x_{k+1} = f\left(x_k, u_k | \theta_x\right), \quad (k = 0, 1, \dots, K-1)$$

In optimal control, the dynamics can be used as equality constraints in optimisation

In this case, the initial state vector x_0 is not necessarily known, or fixed

- It can be one of the decision variables to be determined
- Moreover, certain constraints would apply to it

Similarly, also the final state x_K can be treated as decision variable in an optimisation

Overview (cont.)

Initial and terminal state constraints

We express the constraints on initial and terminal states in terms of function $r(x_0, x_K)$

$$r : \mathcal{R}^{N_x+N_x} \rightarrow \mathcal{R}^{N_r}$$

We express the desire to reach certain initial and terminal states as equality constraints

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

For fixed initial state $x_0 = \overline{x}_0$, we have

$$r(x_0, x_K) = x_0 - \bar{x}_0$$

For fixed terminal state $x_K = \overline{x}_K$, we have

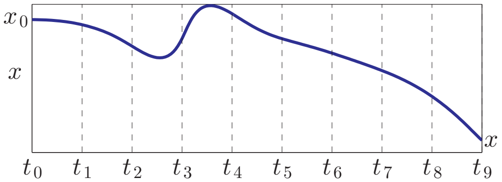
$$r(x_0, x_K) = x_K - \bar{x}_K$$

For fixed both initial and terminal states, $x_0 = \overline{x}_0$ and $x_K = \overline{x}_K$, we have

$$r(x_0, x_K) = \begin{bmatrix} x_0 - \bar{x}_0 \\ x_K - \bar{x}_K \end{bmatrix}$$

Overview (cont.)

For fixed both initial and terminal states, $x_0 = \overline{x}_0$ and $x_K = \overline{x}_K$, we have



$$r(x_0, x_K) = \underbrace{\begin{bmatrix} x_0^{(1)} - \overline{x}_0^{(1)} \\ x_0^{(2)} - \overline{x}_0^{(2)} \\ \vdots \\ x_0^{(N_x)} - \overline{x}_0^{(N_x)} \\ \hline x_K^{(1)} - \overline{x}_K^{(1)} \\ x_K^{(2)} - \overline{x}_K^{(2)} \\ \vdots \\ x_K^{(N_x)} - \overline{x}_K^{(N_x)} \end{bmatrix}}_{N_r \times 1}$$

Overview (cont.)

Path constraints

We can express certain constraints on arbitrary state and control values, x_k and u_k

- These constraints often represent certain technological restrictions
- They are expressed in terms of inequality constraints
- The main idea is to use them to avoid violations

$$h\left(x_k, u_k\right) \leq 0, \quad k = 0, 1, \ldots, K-1$$

For notational simplicity, we used time-invariant inequality constraint functions h

---

For upper and lower bounds on the controls, $u_{\min} \leq u_k \leq u_{\max}$, we have

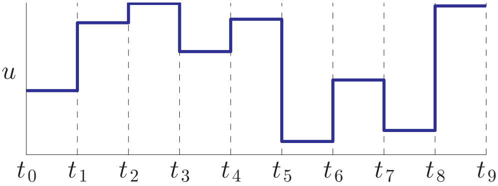
$$h\left(x_k, u_k\right) = \begin{bmatrix} u_k - u_{\max} \\ u_{\min} - u_k \end{bmatrix}$$

For upper and lower bounds on the states, $x_{\min} \leq x_k \leq x_{\max}$, we have

$$h\left(x_k, u_k\right) = \begin{bmatrix} x_k - x_{\max} \\ x_{\min} - x_k \end{bmatrix}$$

Overview (cont.)

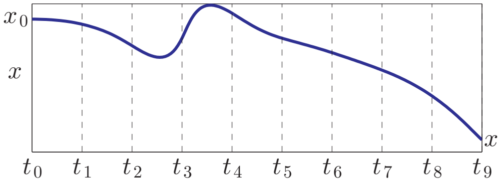
For upper and lower bounds on the controls, $u_{\min} \geq u_k \geq u_{\max}$, we have



$$h\left(x_k, u_k\right) = \left[\begin{array}{c} u_k^{(1)} - u_{\max}^{(1)} \\ u_k^{(2)} - u_{\max}^{(2)} \\ \vdots \\ u_k^{(N_u)} - u_{\max}^{(N_u)} \\ \hline u_{\min}^{(1)} - u_k^{(1)} \\ u_{\min}^{(2)} - u_k^{(2)} \\ \vdots \\ u_{\min}^{(N_u)} - u_k^{(N_u)} \end{array}\right]$$

Overview (cont.)

For upper and lower bounds on the states, $x_{\min} \geq x_k \geq x_{\max}$, we have



$$h\left(x_k, u_k\right) = \begin{bmatrix} x_k^{(1)} - x_{\max}^{(1)} \\ x_k^{(2)} - x_{\max}^{(2)} \\ \vdots \\ x_k^{(N_)} - x_{\max}^{(N_x)} \\ \hline x_{\min}^{(1)} - x_k^{(1)} \\ x_{\min}^{(2)} - x_k^{(2)} \\ \vdots \\ x_{\min}^{(N_x)} - x_k^{(N_x)} \end{bmatrix}$$

Problem formulations

Discrete-time optimal control

Problem formulations

We have the system dynamics and the specifications on the state and control constraints

We use them to formulate the control problem, as constrained nonlinear optimisation

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \qquad k = 0, 1, \dots, K-1 \\
& h(x_k, u_k) \leq 0, \qquad k = 0, 1, \dots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}
$$

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \quad && k = 0, 1, \dots, K-1 \\
& h(x_k, u_k) \leq 0, && k = 0, 1, \dots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}
$$

The **objective function**, two terms

$$\sum_{k=0}^{K-1} L(x_k, u_k) + E(x_K)$$

The **decision variables**, two sets

$$
\begin{aligned}
& x_0, x_1, \dots, x_{K-1}, x_K \\
& u_0, u_1, \dots, u_{K-1}
\end{aligned}
$$

The **equality constraints**, two sets

$$
\begin{aligned}
x_{k+1} - f(x_k, u_k | \theta_x) = 0 \quad & (k = 0, \dots, K-1) \\
r(x_0, x_K) = 0 \quad &
\end{aligned}
$$

The **inequality constraints**

$$h(x_k, u_k) \leq 0 \quad (k = 0, 1, \dots, K-1)$$

Problem formulations (cont.)

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \quad && k = 0, 1, \dots, K-1 \\
& h(x_k, u_k) \leq 0, && k = 0, 1, \dots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}
$$

The **objective function** is the sum of all **stage costs** $L(x_k, u_k)$ and a **terminal cost** $E(x_K)$

$$
\underbrace{\sum_{k=0}^{K-1} L(x_k, u_k) + E(x_K)}_{f(w) \in \mathcal{R}}
$$

That is,

$$
L(x_0, u_0) + L(x_1, u_1) + \cdots + L(x_{K-1}, u_{K-1}) + E(x_K)
$$

Stage cost is a (potentially nonlinear and time-varying) function of state and controls

The **decision variables**, $K \times N_u$ **control** and $(K+1) \times N_x$ **state variables**

$$
\underbrace{(x_0, x_1, \dots, x_{K-1}, x_K) \cup (u_0, u_1, \dots, u_{K-1})}_{w \in \mathcal{R}^{K \times N_u + (K+1) \times N_x}}
$$

Problem formulations (cont.)

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \qquad k = 0, 1, \dots, K-1 \\
& h(x_k, u_k) \leq 0, \qquad k = 0, 1, \dots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}
$$

The **equality constraints**, the K **dynamics** and the N_r **boundary conditions**

$$
\underbrace{\begin{matrix} x_{k+1} - f(x_k, u_k | \theta_x) = 0 \quad (k = 0, \dots, K-1) \\ r(x_0, x_K) = 0 \end{matrix}}_{g(w) \in \mathcal{R}^{N_g}}
$$

The **inequality constraints**

$$
\underbrace{h(x_k, u_k) \leq 0 \quad (k = 0, 1, \dots, K-1)}_{h(w) \in \mathcal{R}^{N_h}}
$$

Problem formulations (cont.)

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \ldots, x_K \\ u_0, u_1, \ldots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \qquad k = 0, 1, \ldots, K-1 \\
& h(x_k, u_k) \leq 0, \qquad k = 0, 1, \ldots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}
$$

The discrete-time optimal control problem is a potentially very large nonlinear program

- In principle, its solution can be approached using any generic NLP solver

We discuss the two approaches used to solve discrete-time optimal control problems

- The **simultaneous approach**
- The **sequential approach**

The simultaneous approach

Problem formulations

Problem formulations | Simultaneous approach

$$\begin{aligned}
\min_{\substack{x_0, x_1, \ldots, x_K \\ u_0, u_1, \ldots, u_{K-1}}} \quad & E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right) \\
\text{subject to} \quad & x_{k+1} - f\left(x_k, u_k | \theta_x\right) = 0, \quad && k = 0, 1, \ldots, K-1 \\
& h\left(x_k, u_k\right) \leq 0, && k = 0, 1, \ldots, K-1 \\
& r\left(x_0, x_K\right) = 0
\end{aligned}$$

The **simultaneous approach** solves the problem in the space of all the decision vars

$$w = \left(x_0, u_0, x_1, u_1, \ldots, x_{K-1}, u_{K-1}, x_K\right)$$

Thus, there are $(K \times N_u) + ((K+1) \times N_x)$ decision variables

Problem formulations | Simultaneous approach

The Lagrangian function of the problem,

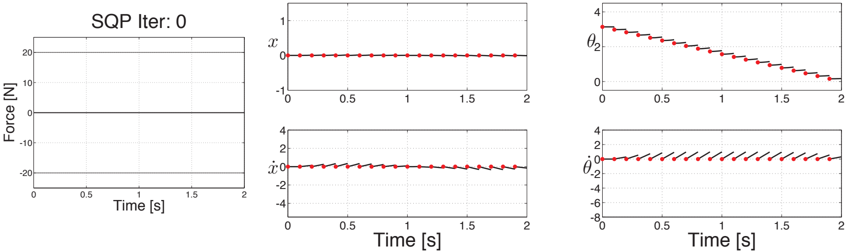
$$\mathcal{L}(w, \lambda, \mu) = f(w) + \lambda^T g(w) + \mu^T h(w)$$

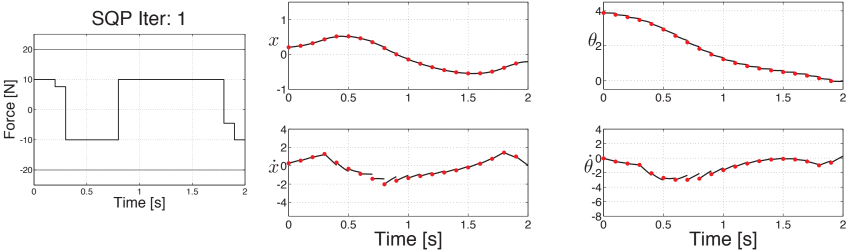
The Karush-Kuhn-Tucker conditions,

$$\begin{aligned}
\nabla f(w^*) + \nabla g(w^*)\lambda^* + \nabla h(w^*)\mu^* &= 0 \\
g(w^*) &= 0 \\
h(w^*) &\leq 0 \\
\mu^* &\geq 0 \\
\mu^*_{n_h} h_{n_h}(w^*) &= 0, \quad n_h = 1, \dots, N_h
\end{aligned}$$

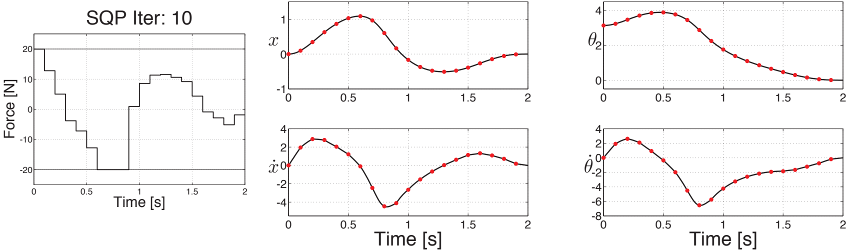
If point $w^* = (x_0^*, u_0^*, \dots, x_{K-1}^*, u_{K-1}^*, x_K^*)$ is a local minimiser of the nonlinear program and if LICQ holds at w^*, there there exist two vectors, the Lagrange multipliers $\lambda \in \mathcal{R}^{N_g}$ and $\mu \in \mathcal{R}^{N_h}$, such that the Karhush-Kuhn-Tucker conditions are verified

Problem formulations | Simultaneous approach (cont.)





Problem formulations | Simultaneous approach (cont.)



Problem formulations | Simultaneous approach (cont.)

To understand more closely the structure and sparsity properties, consider an example

$$\begin{aligned}\min_{\substack{x_0,x_1,\dots,x_K\\u_0,u_1,\dots,u_{K-1}}} \quad & E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right)\\ \text{subject to} \quad & x_{k+1} - f\left(x_k, u_k|\theta_x\right) = 0, \quad k = 0, 1, \dots, K-1\\ & r\left(x_0, x_K\right) = 0\end{aligned}$$

This optimal control problem in discrete-time has no inequality constraints

- Inequality constraints are omitted for notational simplicity

The objective $f\left(w\right) = E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right)$ of the decision variables,

$$w = \left(x_0, u_0, x_1, u_1, \dots, x_{K-1}, u_{K-1}, x_K\right)$$

Problem formulations | Simultaneous approach (cont.)

$$\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \quad k = 0, 1, \dots, K-1 \\
& r(x_0, x_K) = 0
\end{aligned}$$

We define the equality constraint function by concatenation

$$g(w) = \begin{bmatrix} g_1(w) \\ g_2(w) \\ \vdots \\ g_{N_g}(w) \end{bmatrix} = \underbrace{\left[\begin{array}{c} x_1 - f(x_0, u_0) \\ x_2 - f(x_1, u_1) \\ \vdots \\ x_K - f(x_{K-1}, u_{K-1}) \\ \hline r(x_0, x_K) \end{array}\right]}_{((K \times N_x) + N_r) \times 1}$$

Problem formulations | Simultaneous approach (cont.)

$$\begin{aligned} \min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\ \text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \quad k = 0, 1, \dots, K-1 \\ & r(x_0, x_K) = 0 \end{aligned}$$

The Lagrangian function for equality constrained problems,

$$\mathcal{L}(w) = f(w) + \lambda^T g(w)$$

The equality multipliers,

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_K, \lambda_{N_r})$$

The KKT conditions,

$$\begin{aligned} \nabla_w \mathcal{L}(w, \lambda) &= 0 \\ g(w) &= 0 \end{aligned}$$

Problem formulations | Simultaneous approach (cont.)

$$\underbrace{\begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_K & \lambda_{N_r} \end{bmatrix}}_{\lambda^T} \underbrace{\left[\begin{array}{c} x_1 - f\left(x_0, u_0\right) \\ x_2 - f\left(x_1, u_1\right) \\ \vdots \\ x_K - f\left(x_{K-1}, u_{K-1}\right) \\ \hline \\ r\left(x_0, x_K\right) \end{array}\right]}_{g(w)}$$

After expanding the terms in the inner product, we re-write the Lagrangian function

$$\mathcal{L}\left(w, \lambda\right) = \underbrace{E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right)}_{f(w)} + \underbrace{\left(\sum_{k=0}^{K-1} \lambda_{k+1}^T \left(f\left(x_k, u_k\right) - x_{k+1}\right) + \lambda_{N_r}^T r\left(x_0, x_K\right)\right)}_{\lambda^T g(w)}$$

Problem formulations | Simultaneous approach (cont.)

Consider one of the dynamic constraints,

$$x_{k+1} - f\left(x_k, u_k\right) = 0$$

More explicitly, we have

$$\underbrace{\begin{bmatrix} x_{k+1}^{(1)} - f_1\left(x_k, u_k\right) \\ x_{k+1}^{(2)} - f_2\left(x_k, u_k\right) \\ \vdots \\ x_{k+1}^{(n_x)} - f_{n_x}\left(x_k, u_k\right) \\ \vdots \\ x_{k+1}^{(N_x)} - f_{N_x}\left(x_k, u_k\right) \end{bmatrix}}_{N_x \times 1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Problem formulations | Simultaneous approach (cont.)

Consider the corresponding product with the equality multiplier,

$$\underbrace{\lambda_{k+1}^T \underbrace{\left(f\left(x_k, u_k\right) - x_{k+1}\right)}_{N_x \times 1}}_{1 \times 1}$$

More explicitly, we have

$$\underbrace{\begin{bmatrix} \lambda_{k+1}^{(1)} & \lambda_{k+1}^{(2)} & \cdots & \lambda_{k+1}^{(n_x)} & \cdots & \lambda_{k+1}^{(N_x)} \end{bmatrix}}_{1 \times N_x} \underbrace{\begin{bmatrix} x_{k+1}^{(1)} - f_1\left(x_k, u_k\right) \\ x_{k+1}^{(2)} - f_2\left(x_k, u_k\right) \\ \vdots \\ x_{k+1}^{(n_x)} - f_{n_x}\left(x_k, u_k\right) \\ \vdots \\ x_{k+1}^{(N_x)} - f_{N_x}\left(x_k, u_k\right) \end{bmatrix}}_{N_x \times 1}$$

Problem formulations | Simultaneous approach (cont.)

Similarly, consider the boundary constraint,

$$r\left(x_0, x_K\right) = 0$$

In more detail, we have,

$$r\left(x_0, x_N\right) = \underbrace{\left[\begin{array}{c} x_0^{(1)} - \overline{x}_0^{(1)} \\ x_0^{(2)} - \overline{x}_0^{(2)} \\ \vdots \\ x_0^{(N_x)} - \overline{x}_0^{(N_x)} \\ \hline x_K^{(1)} - \overline{x}_K^{(1)} \\ x_K^{(2)} - \overline{x}_K^{(2)} \\ \vdots \\ x_K^{(N_x)} - \overline{x}_K^{(N_x)} \end{array}\right]}_{N_r \times 1}$$

Problem formulations | Simultaneous approach (cont.)

For the product $\lambda_{N_r}^T r(x_0, x_K)$ with the equality multiplier, we have

$$\underbrace{\underbrace{\lambda_{N_r}^T}_{1 \times N_r} \underbrace{r(x_0, x_K)}_{N_r \times 1}}_{1 \times 1}$$

More explicitly, we have

$$\underbrace{\begin{bmatrix} \lambda_{N_r}^{(1)} & \cdots & \lambda_{N_r}^{(N_x)} & \lambda_{N_r}^{(N_x+1)} & \cdots & \lambda_{N_r}^{(2N_x)} \end{bmatrix}}_{1 \times N_r} \underbrace{\left[\begin{array}{c} x_0^{(1)} - \overline{x}_0^{(1)} \\ x_0^{(2)} - \overline{x}_0^{(2)} \\ \vdots \\ x_0^{(N_x)} - \overline{x}_0^{(N_x)} \\ \hline x_K^{(1)} - \overline{x}_K^{(1)} \\ x_K^{(2)} - \overline{x}_K^{(2)} \\ \vdots \\ x_K^{(N_x)} - \overline{x}_K^{(N_x)} \end{array}\right]}_{N_r \times 1}$$

Problem formulations | Simultaneous approach (cont.)

For the Lagrangian function for equality constrained problems, we thus have

$$\mathcal{L}(w, \lambda) = \underbrace{f(w)}_{1 \times 1} + \underbrace{\begin{bmatrix} \underbrace{\lambda_1}_{1 \times N_x} & \underbrace{\lambda_2}_{1 \times N_x} & \cdots & \underbrace{\lambda_K}_{1 \times N_x} & \underbrace{\lambda_{N_r}}_{1 \times N_r} \end{bmatrix}}_{\underbrace{\lambda^T}_{1 \times ((K \times N_x) + N_r)}} \underbrace{\left[\begin{array}{c} \underbrace{x_1 - f(x_0, u_0)}_{N_x \times 1} \\ \underbrace{x_2 - f(x_1, u_1)}_{N_x \times 1} \\ \vdots \\ \underbrace{x_K - f(x_{K-1}, u_{K-1})}_{N_x \times 1} \\ \hline \underbrace{r(x_0, x_K)}_{N_r \times 1} \end{array}\right]}_{\underbrace{g(w)}_{((K \times N_x) + N_r) \times 1}}$$

Problem formulations | Simultaneous approach (cont.)

$$\begin{aligned} \nabla_w \mathcal{L}(w, \lambda) &= 0 \\ g(w) &= 0 \end{aligned}$$

The second KKT condition,

$$\begin{aligned} x_{k+1} - f(x_k, u_k) &= 0 \quad (k = 0, \dots, K-1) \\ r(x_0, x_K) &= 0 \end{aligned}$$

The first KKT condition regards the derivative of $\mathcal{L}$ with respect to the primal vars w

$$w = (x_0, u_0, x_1, u_1, \dots, x_{K-1}, u_{K-1}, x_K)$$

The Lagrangian function in structural form,

$$\underbrace{\underbrace{E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k)}_{f(w)} + \underbrace{\left(\sum_{k=0}^{K-1} \lambda_{k+1}^T \left(f(x_k, u_k) - x_{k+1} \right) + \lambda_{N_r}^T r(x_0, x_K) \right)}_{\lambda^T g(w)}}_{\mathcal{L}(w, \lambda)}$$

Problem formulations | Simultaneous approach (cont.)

$$g\left(w\right) = 0$$

For the second KKT condition, we have

$$\begin{aligned} x_{k+1} - f\left(x_k, u_k\right) &= 0 \quad (k = 0, \dots, K-1) \\ r\left(x_0, x_K\right) &= 0 \end{aligned}$$

That is,

$$\left[\begin{array}{c} \underbrace{x_1 - f\left(x_0, u_0\right)}_{N_x \times 1} \\ \underbrace{x_2 - f\left(x_1, u_1\right)}_{N_x \times 1} \\ \vdots \\ \underbrace{x_K - f\left(x_{K-1}, u_{K-1}\right)}_{N_x \times 1} \\ \hline \underbrace{r\left(x_0, x_K\right)}_{N_r \times 1} \end{array}\right] = \left[\begin{array}{c} \underbrace{0}_{N_x \times 1} \\ \underbrace{0}_{N_x \times 1} \\ \vdots \\ \underbrace{0}_{N_x \times 1} \\ \hline \underbrace{0}_{N_r \times 1} \end{array}\right]$$

Problem formulations | Simultaneous approach (cont.)

$$\nabla_w \mathcal{L}(w, \lambda) = 0$$

Consider the gradient of the Lagrangian function, it is a concatenation of gradients

$$\nabla_w \mathcal{L}(w, \lambda) = \left[\begin{array}{c} \nabla_{x_0} \mathcal{L}(w, \lambda) \\ \nabla_{x_1} \mathcal{L}(w, \lambda) \\ \vdots \\ \nabla_{x_K} \mathcal{L}(w, \lambda) \\ \hline \nabla_{u_0} \mathcal{L}(w, \lambda) \\ \nabla_{u_1} \mathcal{L}(w, \lambda) \\ \vdots \\ \nabla_{u_{K-1}} \mathcal{L}(w, \lambda) \end{array}\right]$$

For the second KKT conditions, it is necessary to determine/evaluate the derivatives

Problem formulations | Simultaneous approach (cont.)

$$\underbrace{E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right) + \left(\sum_{k=0}^{K-1} \lambda_{k+1}^T \left(f\left(x_k, u_k\right) - x_{k+1}\right) + \lambda_{N_r}^T r\left(x_0, x_K\right)\right)}_{\mathcal{L}(w,\lambda)}$$

The derivatives of the Lagrangian function with respect to the state variables x_k

- For $k = 0$, we have

$$\nabla_{x_0} \mathcal{L}\left(w, \lambda\right) = \nabla_{x_0} L\left(x_0, u_0\right) + \frac{\partial f\left(x_0, u_0\right)}{\partial x_0}^T \lambda_1 + \frac{\partial r\left(x_0, x_K\right)}{\partial x_0}^T \lambda_{N_r}$$

- For $k = 1, \dots, K-1$, we have

$$\nabla_{x_k} \mathcal{L}\left(w, \lambda\right) = \nabla_{x_k} L\left(x_k, u_k\right) + \frac{\partial f\left(x_k, u_k\right)}{\partial x_k}^T \lambda_{k+1} - \lambda_k$$

- For $k = K$, we have

$$\nabla_{x_K} \mathcal{L}\left(w, \lambda\right) = \nabla_{x_K} E\left(x_N\right) - \lambda_K + \frac{\partial r\left(x_0, x_K\right)}{\partial x_K}^T \lambda_{N_r}$$

Problem formulations | Simultaneous approach (cont.)

Consider the generic term $\nabla_{x_k}\mathcal{L}(w,\lambda)$,

$$\nabla_{x_k}\mathcal{L}(w,\lambda) = \underbrace{\begin{bmatrix} \dfrac{\partial \mathcal{L}(w,\lambda)}{\partial x_k^{(1)}} \\ \dfrac{\partial \mathcal{L}(w,\lambda)}{\partial x_k^{(2)}} \\ \vdots \\ \dfrac{\partial \mathcal{L}(w,\lambda)}{\partial x_k^{(N_x)}} \end{bmatrix}}_{N_x \times 1}$$

Problem formulations | Simultaneous approach (cont.)

Consider the derivative of the dynamics,

$$\frac{\partial f\left(x_k, u_k\right)}{\partial x_k}$$

Remember the dynamics,

$$f\left(x_k, u_k\right) = \begin{bmatrix} f_1\left(x_k^{(1)}, \dots, x_K^{(N_x)}, u_k\right) \\ \vdots \\ f_{n_x}\left(x_k^{(1)}, \dots, x_K^{(N_x)}, u_k\right) \\ \vdots \\ f_{N_x}\left(x_k^{(1)}, \dots, x_K^{(N_x)}, u_k\right) \end{bmatrix}$$

Problem formulations | Simultaneous approach (cont.)

For the derivative of the dynamics, we have

$$\frac{\partial f\left(x_k^{(1)}, \dots, x_k^{(N_x)}, u_k\right)}{\partial x_k} = \begin{bmatrix} \dfrac{\partial f_1\left(x_k^{(1)}, \dots, x_k^{(N_x)}, u_k\right)}{\partial x_k} \\ \vdots \\ \dfrac{\partial f_{n_x}\left(x_k^{(1)}, \dots, x_k^{(N_x)}, u_k\right)}{\partial x_k} \\ \vdots \\ \dfrac{\partial f_{N_x}\left(x_k^{(1)}, \dots, x_k^{(N_x)}, u_k\right)}{\partial x_k} \end{bmatrix}$$

Problem formulations | Simultaneous approach (cont.)

In more detail, we have

$$\frac{\partial f\left(x_k, u_k\right)}{\partial x_k} = \underbrace{\begin{bmatrix} \dfrac{\partial f_1\left(x_k, u_k\right)}{\partial x_k^{(1)}} & \dfrac{\partial f_1\left(x_k, u_k\right)}{\partial x_k^{(2)}} & \cdots & \dfrac{\partial f_1\left(x_k, u_k\right)}{\partial x_k^{(N_x)}} \\ \dfrac{\partial f_2\left(x_k, u_k\right)}{\partial x_k^{(1)}} & \dfrac{\partial f_2\left(x_k, u_k\right)}{\partial x_k^{(2)}} & \cdots & \dfrac{\partial f_2\left(x_k, u_k\right)}{\partial x_k^{(N_x)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_{N_x}\left(x_k, u_k\right)}{\partial x_k^{(1)}} & \dfrac{\partial f_{N_x}\left(x_k, u_k\right)}{\partial x_k^{(2)}} & \cdots & \dfrac{\partial f_{N_x}\left(x_k, u_k\right)}{\partial x_k^{(N_x)}} \end{bmatrix}}_{N_x \times N_x}$$

For the product with the equality multiplier, we get

$$\underbrace{\underbrace{\frac{\partial f\left(x_k, u_k\right)}{\partial x_k}^T}_{N_x \times N_x} \underbrace{\lambda_{k+1}}_{N_x \times 1}}_{N_x \times 1}$$

Problem formulations | Simultaneous approach (cont.)

Consider the derivatives of the boundary conditions, we have the terms

- $$\frac{\partial r\left(x_0, x_K\right)}{\partial x_0}$$
- $$\frac{\partial r\left(x_0, x_K\right)}{\partial x_K}$$

Remember the boundary constraints

$$r\left(x_0, x_K\right) = \underbrace{\begin{bmatrix} x_0^{(1)} - \overline{x}_0^{(1)} \\ x_0^{(2)} - \overline{x}_0^{(2)} \\ \vdots \\ x_0^{(N_x)} - \overline{x}_0^{(N_x)} \\ \hline x_K^{(1)} - \overline{x}_K^{(1)} \\ x_K^{(2)} - \overline{x}_K^{(2)} \\ \vdots \\ x_K^{(N_x)} - \overline{x}_K^{(N_x)} \end{bmatrix}}_{N_r \times 1}$$

Problem formulations | Simultaneous approach (cont.)

For the derivative of the boundary constraints with respect to x_0, we have

$$\frac{\partial r\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} = \left[\begin{array}{c} \dfrac{\partial r_1\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \\ \dfrac{\partial r_2\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \\ \vdots \\ \dfrac{\partial r_{N_x}\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \\ \hline \dfrac{\partial r_{N_x+1}\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \\ \dfrac{\partial r_{N_x+2}\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \\ \vdots \\ \dfrac{\partial r_{2N_x}\left(x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(N_x)}, x_K\right)}{\partial x_0} \end{array}\right]$$

Problem formulations | Simultaneous approach (cont.)

In more detail, we have

$$\frac{\partial r(x_0, x_K)}{\partial x_0} = \underbrace{\begin{bmatrix} \dfrac{\partial r_1(x_0, x_K)}{\partial x_0^{(1)}} & \dfrac{\partial r_1(x_0, x_K)}{\partial x_0^{(2)}} & \cdots & \dfrac{\partial r_1(x_0, x_K)}{\partial x_0^{(N_x)}} \\ \dfrac{\partial r_2(x_0, x_K)}{\partial x_0^{(1)}} & \dfrac{\partial r_2(x_0, x_K)}{\partial x_0^{(2)}} & \cdots & \dfrac{\partial r_2(x_0, x_K)}{\partial x_0^{(N_x)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial r_{2N_r}(x_0, x_k)}{\partial x_0^{(1)}} & \dfrac{\partial r_{2N_r}(x_0, x_K)}{\partial x_0^{(2)}} & \cdots & \dfrac{\partial r_{2N_r}(x_0, x_K)}{\partial x_0^{(N_x)}} \end{bmatrix}}_{2N_r \times N_x}$$

For the product with the equality multiplier, we get

$$\underbrace{\underbrace{\frac{\partial r(x_0, x_K)}{\partial x_0}^T}_{N_x \times 2N_r} \underbrace{\lambda_{k+1}}_{2N_r \times 1}}_{N_x \times 1}$$

Problem formulations | Simultaneous approach (cont.)

$$\underbrace{E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right) + \left(\sum_{k=0}^{K-1} \lambda_{k+1}^T \left(f\left(x_k, u_k\right) - x_{k+1}\right) + \lambda_{N_r}^T r\left(x_0, x_K\right)\right)}_{\mathcal{L}(w,\lambda)}$$

The derivatives of the Lagrangian function with respect to the control variables u_k

- For $k = 0, \ldots, K-1$, we have

$$\nabla_{u_k} \mathcal{L}\left(w, \lambda\right) = \nabla_{u_k} L\left(x_k, u_k\right) + \frac{\partial f\left(x_k, u_k\right)}{\partial u_k}^T \lambda_{k+1}$$

Problem formulations | Simultaneous approach (cont.)

$$\begin{aligned} \nabla_w \mathcal{L}(w, \lambda) &= 0 \\ g(w) &= 0 \end{aligned}$$

We can collect all the KKT conditions and solve them using a Newton-type method

- The approach solves the problem in the full space of the decision variables

Problem formulations | Simultaneous approach (cont.)

The approach can be extended to more general discrete-time optimal control problems

$$\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L_k(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f_k(x_k, u_k | \theta_x) = 0, && k = 0, 1, \dots, K-1 \\
& h_k(x_k, u_k) \leq 0, && k = 0, 1, \dots, K-1 \\
& R_K(x_K) + \sum_{k=0}^{K-1} r_k(x_k, u_k) = 0 \\
& h_K(x_K) \leq 0
\end{aligned}$$

All problem functions are explicitly time-varying and we have also a terminal inequality

- Moreover, the boundary conditions are expressed in general form

---

By collecting all variables in the vector w, we have the complete Lagrangian function

$$\mathcal{L}(w, \lambda, \mu) = f(w) + \lambda^T g(w) + \mu^T h(w)$$

The sequential approach

Problem formulations

Problem formulations | Sequential approach

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \ldots, x_K \\ u_0, u_1, \ldots, u_{K-1}}} \quad & E(x_K) + \sum_{k=0}^{K-1} L(x_k, u_k) \\
\text{subject to} \quad & x_{k+1} - f(x_k, u_k | \theta_x) = 0, \qquad k = 0, 1, \ldots, K-1 \\
& h(x_k, u_k) \leq 0, \qquad k = 0, 1, \ldots, K-1 \\
& r(x_0, x_N) = 0
\end{aligned}
$$

The **sequential approach** solves the same problem in a reduced space of variables

The idea is to eliminate all the state variables $x_1, x_2, \ldots, x_K$ by a forward simulation

$$
\begin{aligned}
x_0 &= x_0 \\
x_1 &= f(x_0, u_0) \\
x_2 &= f(x_1, u_1) \\
&= f(f(x_0, u_0), u_1) \\
x_3 &= f(x_2, u_2) \\
&= f(f(f(x_0, u_0), u_1), u_2) \\
\cdots &= \cdots
\end{aligned}
$$

Problem formulations | Sequential approach (cont.)

We can express the states as function of the initial condition and previous controls

$$
\begin{aligned}
x_0 &= \underbrace{x_0}_{\overline{x}_0(x_0)} \\
x_1 &= \underbrace{f(x_0, u_0)}_{\overline{x}_1(x_0, u_0)} \\
x_2 &= f(x_1, u_1) \\
&= \underbrace{f(f(x_0, u_0), u_1)}_{\overline{x}_2(x_0, u_0, u_1)} \\
x_3 &= f(x_2, u_2) \\
&= \underbrace{f(f(f(x_0, u_0), u_1), u_2)}_{\overline{x}_3(x_0, u_0, u)_1, u_2)} \\
\cdots &= \cdots
\end{aligned}
$$

More generally, the dependence is on all the control variables and the initial condition

$$
\begin{aligned}
\overline{x}_0(x_o, u_0, u_1, \dots, u_{K-1}) &= x_0 \\
\overline{x}_{k+1}(x_0, u_0, u_1, \dots, u_{K-1}) &= f(\overline{x}_k(x_0, u_0, u_1, \dots, u_{K-1}), u_k), \quad k = 0, 1, \dots, K-1
\end{aligned}
$$

Problem formulations | Sequential approach

$$
\begin{aligned}
\min_{\substack{x_0, x_1, \dots, x_K \\ u_0, u_1, \dots, u_{K-1}}} \quad & E\left(x_K\right) + \sum_{k=0}^{K-1} L\left(x_k, u_k\right) \\
\text{subject to} \quad & x_{k+1} - f\left(x_k, u_k | \theta_x\right) = 0, \qquad k = 0, 1, \dots, K-1 \\
& h\left(x_k, u_k\right) \leq 0, \qquad k = 0, 1, \dots, K-1 \\
& r\left(x_0, x_N\right) = 0
\end{aligned}
$$

We can re-write the general discrete-time optimal control problem in reduced form

$$
\begin{aligned}
\min_{\substack{x_0 \\ u_0, u_1, \dots, u_{K-1}}} \quad & E\left(\overline{x}_K\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) + \sum_{k=0}^{K-1} L\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \\
\text{subject to} \quad & h\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \leq 0, k = 0, 1, \dots, K-1 \\
& r\left(x_0, \overline{x}_N\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) = 0
\end{aligned}
$$

Problem formulations | Sequential approach (cont.)

$$\begin{aligned} \min_{\substack{x_0 \\ u_0, u_1, \dots, u_{K-1}}} \quad & E\left(\overline{x}_K\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) + \sum_{k=0}^{K-1} L\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \\ \text{subject to} \quad & h\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \leq 0, k = 0, 1, \dots, K-1 \\ & r\left(x_0, \overline{x}_N\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) = 0 \end{aligned}$$

The **objective function**, sum of **stage costs** $L\left(\overline{x}_k, u_k\right)$ and a **terminal cost** $E\left(\overline{x}_K\right)$

$$\underbrace{\sum_{k=0}^{K-1} L\left(\overline{x}_k, u_k\right) + E\left(\overline{x}_K\right)}_{f(w) \in \mathcal{R}}$$

That is,

$$L\left(x_0, u_0\right) + L\left(\overline{x}_1, u_1\right) + \cdots + L\left(\overline{x}_{K-1}, u_{K-1}\right) + E\left(\overline{x}_K\right)$$

Stage cost is a (potentially nonlinear and time-varying) function of state and controls

The **decision variables**, $K \times N_u$ **control** and N_x **state variables**

$$\underbrace{\left(x_0\right) \cup \left(u_0, u_1, \dots, u_{K-1}\right)}_{w \in \mathcal{R}^{K \times N_u + N_x}}$$

Problem formulations | Sequential approach (cont.)

$$\begin{aligned}
\min_{\substack{x_0 \\ u_0, u_1, \ldots, u_{K-1}}} \quad & E\left(\overline{x}_K\left(x_0, u_0, u_1, \ldots, u_{K-1}\right)\right) + \sum_{k=0}^{K-1} L\left(\overline{x}_k\left(x_0, u_0, u_1, \ldots, u_{K-1}\right), u_k\right) \\
\text{subject to} \quad & h\left(\overline{x}_k\left(x_0, u_0, u_1, \ldots, u_{K-1}\right), u_k\right) \leq 0, k = 0, 1, \ldots, K-1 \\
& r\left(x_0, \overline{x}_N\left(x_0, u_0, u_1, \ldots, u_{K-1}\right)\right) = 0
\end{aligned}$$

The **equality constraints**, the N_r **boundary conditions**

$$\underbrace{r\left(x_0, \overline{x}_K\right) = 0}_{g(w) \in \mathcal{R}^{N_g}}$$

The **inequality constraints**

$$\underbrace{h\left(\overline{x}_k, u_k\right) \leq 0 \quad (k = 0, 1, \ldots, K-1)}_{h(w) \in \mathcal{R}^{N_h}}$$

Problem formulations | Sequential approach (cont.)

$$\begin{aligned}
\min_{\substack{x_0 \\ u_0, u_1, \dots, u_{K-1}}} \quad & E\left(\overline{x}_K\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) + \sum_{k=0}^{K-1} L\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \\
\text{subject to} \quad & h\left(\overline{x}_k\left(x_0, u_0, u_1, \dots, u_{K-1}\right), u_k\right) \leq 0, k = 0, 1, \dots, K-1 \\
& r\left(x_0, \overline{x}_N\left(x_0, u_0, u_1, \dots, u_{K-1}\right)\right) = 0
\end{aligned}$$

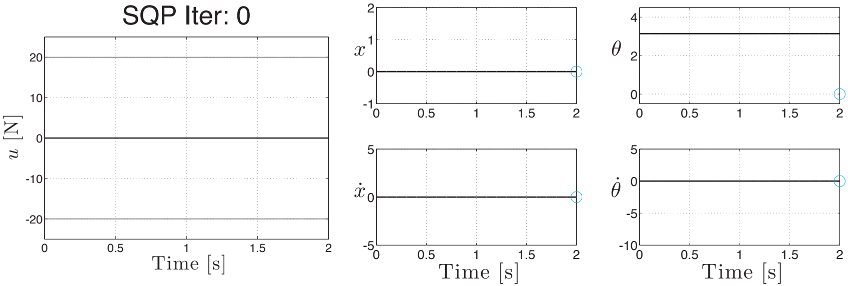
The Lagrangian function of the problem,

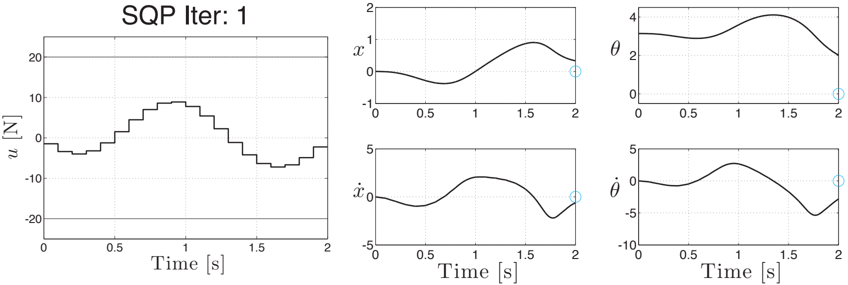
$$\mathcal{L}\left(w, \lambda, \mu\right) = f\left(w\right) + \lambda^T g\left(w\right) + \mu^T h\left(w\right)$$

The Karush-Kuhn-Tucker conditions,

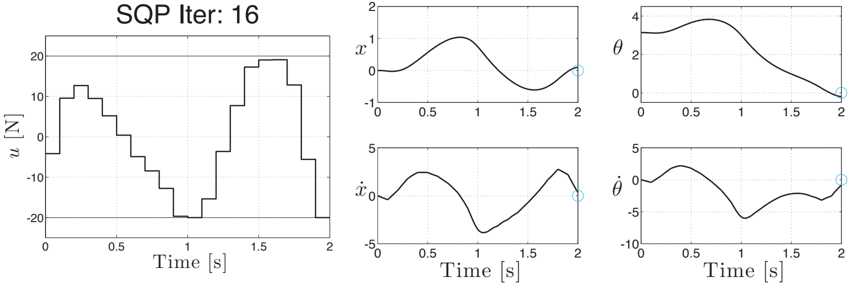
$$\begin{aligned}
\nabla f\left(w^*\right) - \nabla g\left(w^*\right)\lambda^* - \nabla h\left(w^*\right)\mu^* &= 0 \\
g\left(w^*\right) &= 0 \\
h\left(w^*\right) &\geq 0 \\
\mu^* &\geq 0 \\
\mu_{n_h}^* h_{n_h}\left(w^*\right) &= 0, \quad n_h = 1, \dots, N_h
\end{aligned}$$

Problem formulations | Sequential approach (cont.)





Problem formulations | Sequential approach (cont.)



For computational efficiency, it is preferable to use specific structure-exploiting solvers

- Such solvers recognise the sparsity properties of this class of problems