ELEC-E8101: Digital and optimal control

Model predictive control and convex optimization

Abdullah Tokmak

Cyber-physical systems group, Aalto University
Office number 2554, Maarintie 8

November 25, 2024

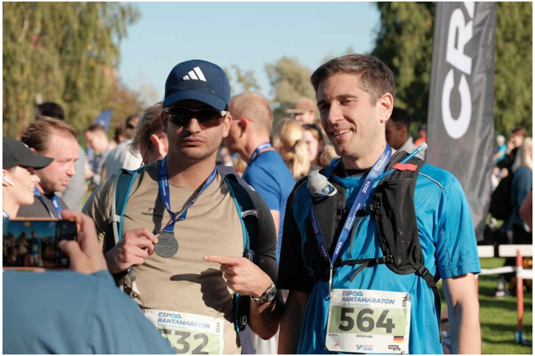
Aalto University
School of Electrical
Engineering

About me: Abdullah Tokmak

- Ph.D. student under Dominik Baumann

Research interests

- Machine learning algorithms for safe and automatic control
- Nonparametric function approximation, e.g., Gaussian processes



Marathon with my boss

In the previous lecture…

We

- Explained the basic working principle of model predictive control
- Explained its advantages and drawbacks compared to optimal control and pole placement

Learning outcomes

By the end of this lecture, you should be able to. . .

- Recognize whether an MPC problem is linear and understand what this implies for the underlying optimization problem
- Know the properties of convex optimization problems
- Derive KKT conditions to solve convex constrained optimization problems
- Explain the basic idea of interior point methods
- Know that the world of nonlinear MPC is huge

Questions during the lecture?

Ask in Presemo: `https://presemo.aalto.fi/digoptctrl`

Linear MPC

- If the **system dynamics are linear**, we can formulate a linear MPC
- Many real-world systems are nonlinear but we can approximate linearly

Linear MPC example: Temperature control in a room

$$\min_{[u_0,\ldots,u_{N-1}]^\top \in \mathbb{R}^N} J(x,u) = \sum_{k=0}^{N-1} \left[(x_k - x_{\mathrm{ref}})^\top Q (x_k - x_{\mathrm{ref}}) + u_k^\top R u_k \right]$$

E.ON Energy Research Center | RWTH AACHEN UNIVERSITY

$$\text{s.t.} \quad x_{k+1} = A x_k + B u_k \qquad \texttt{\% Linear dynamics}$$

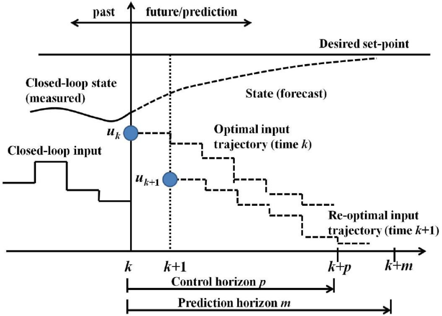
$$x_k \in [x_{\min}, x_{\max}] \qquad \texttt{\% Box-constraints}$$

$$u_k \in [u_{\min}, u_{\max}] \qquad \texttt{\% Box-constraints}$$

- Cost function J **penalizes** deviation from reference temperature x_{ref} and heat flow u_k
- For **appropriately chosen weight matrices**, the optimization problem is **simpler**

Receding horizon

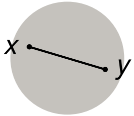
- In MPC, we solve this optimization problem at every time step $k \in [0, N-1]$ for the whole horizon N and only apply the first action $\rightarrow$ **Receding horizon**
- Commonly chosen horizon: $N \approx 120$
  $\rightarrow$ 120 optimization variables per time-step
- To meet **real-time requirements**, solving optimization problems should be simple



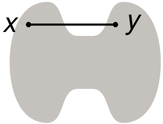
Linear MPC

Linear MPC yields **convex optimization problems**, which are significantly simpler to solve than non-convex optimization problems.

Convexity



Convex set Ω



Non-convex set Ω

Convex sets

The set Ω is convex if, and only if (iff), $\forall x, y \in \Omega, \forall t \in [0, 1] : (1 - t)x + ty \in \Omega$.

Convex functions

Let Ω be a convex set. Then, a twice-differentiable function f is convex iff the three **equivalent** statements hold:

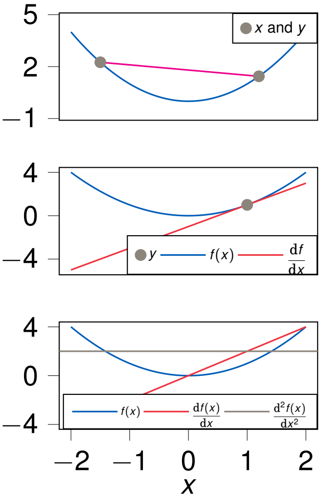
1. $f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y), \quad \forall x, y \in \Omega, \forall t \in [0, 1]$
2. $f(y) \geq f(x) + \dfrac{\mathrm{d}f(x)}{\mathrm{d}x}(x - y), \quad \forall x, y \in \Omega$
3. $\dfrac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} \geq 0, \quad \forall x \in \Omega$

Interpretation of statements for convex functions

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \ \forall x, y \in \Omega, \forall t \in [0,1]$$

$$f(y) \geq f(x) + \frac{\mathrm{d}f(x)}{\mathrm{d}x}(x-y), \quad \forall x, y \in \Omega$$

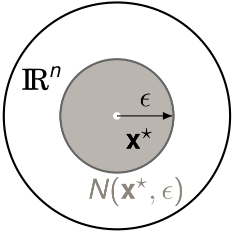
$$\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} \geq 0, \quad \forall x \in \Omega$$



Convex unconstrained optimization: Optimality

- We consider, without loss of generality, only **minimization problems**
- If we want to maximize a function f, we can equivalently minimize its negation $-f$

- The point $\mathbf{x}^\star \in \mathbb{R}^n$ is called a local minimizer if $f(\mathbf{x}^\star) \leq f(\mathbf{x})\ \forall \mathbf{x} \in N(\mathbf{x}^\star, \epsilon)$, where $N(\mathbf{x}^\star, \epsilon)$ is called an ϵ-neighborhood around the minimizer $\mathbf{x}^\star$
- The point $\mathbf{x}^\star \in \mathbb{R}^n$ is called the global minimizer if $f(\mathbf{x}^\star) \leq f(\mathbf{x})\ \forall \mathbf{x} \in \mathbb{R}^n$



MPC and optimality

The input trajectory u_{k} is the solution of the optimization problem at any time step $k \in [0, N-1]$. A **sub-optimal solution**, i.e., not the global minimizer, will lead to **sub-optimal operation** of the plant and, in the worst case, may cause **instability**.

Convex unconstrained optimization

- For a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient evaluated at $\mathbf{x} \in \mathbb{R}^n$ is
$$\nabla f(\mathbf{x}) = \left[\frac{\mathrm{d}f}{\mathrm{d}x_1}(\mathbf{x}), \ldots, \frac{\mathrm{d}f}{\mathrm{d}x_\mathrm{n}}(\mathbf{x})\right]^\top$$
- A point $\mathbf{x} \in \mathbb{R}^n$ is called **stationary** if $\nabla f(\mathbf{x}) = [0, \ldots, 0]^\top =: \mathbf{0}$
- For any direction $\mathbf{p} \in \mathbb{R}^n$, the directional derivative is $\nabla_\mathbf{p} f(\mathbf{x}) := \lim_{\eta \to 0} \frac{f(\mathbf{x}+\eta\mathbf{p})-f(\mathbf{x})}{\eta}$
- The directional derivative is a **projection** of $\nabla f(x)$ onto $\mathbf{p}$, i.e., $\nabla_\mathbf{p} f(\mathbf{x}) = \mathbf{p}^\top \nabla f(\mathbf{x})$
- Although MPC requires constrained optimization, **unconstrained optimization** serves as the **foundation** to solve the underlying problem

Theorem 1: Global minimizer of unconstrained convex functions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable and convex. Then, $\mathbf{x}^\star \in \mathbb{R}^n$ is the **global minimizer**, i.e., $f(\mathbf{x}^\star) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, iff $\nabla f(\mathbf{x}^\star) = \mathbf{0}$, i.e., iff $\mathbf{x}^\star$ is a stationary point.

Proof of Theorem 1: Logic

Handwritten

Theorem 1: Global minimizer of unconstrained convex functions

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable and convex. Then, $\mathbf{x}^\star \in \mathbb{R}^n$ is the **global minimizer**, i.e., $f(\mathbf{x}^\star) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, iff $\nabla f(\mathbf{x}^\star) = \mathbf{0}$, i.e., iff $x^\star$ is a stationary point.

- Statement A: $\nabla f(\mathbf{x}^\star) = \mathbf{0}$ (stationarity)
- Statement B: $f(\mathbf{x}^\star) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in N(\mathbf{x}^\star, \epsilon)$ (local minimum)
- Statement C: $f(\mathbf{x}^\star) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^n$ (global minimum)
- We want to prove $A \iff C$; Through transitive logical relation of the statements:

$$\begin{aligned} A \iff C &= (A \implies C) \wedge (C \implies A) \\ &= \underbrace{(A \implies C)}_{\text{Lemma 1}} \wedge (\underbrace{(C \implies B)}_{\text{Trivial}} \wedge \underbrace{(B \implies A)}_{\text{Lemma 2}}) \end{aligned}$$

Proof of Theorem 1: Lemma 1

Handwritten

Lemma 1 ($A \implies C$)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, convex, and $\nabla f(\mathbf{x}^\star) = 0$. Then, $f(\mathbf{x}^\star) \leq f(\mathbf{x}), \ \forall \mathbf{x} \in \mathbb{R}^n$.

Proof of Lemma 1

- Since f is convex, $f(\mathbf{x}) \geq f(\mathbf{y}) + \nabla f(\mathbf{y})^\top (\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
- Choose $\mathbf{y} = \mathbf{x}^\star$: $f(\mathbf{x}) \geq f(\mathbf{x}^\star) + \nabla f(\mathbf{x}^\star)^\top (\mathbf{x}^\star - \mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^n$
- Since $\nabla f(\mathbf{x}^\star) = 0$, $f(\mathbf{x}) \geq f(\mathbf{x}^\star) \quad \forall \mathbf{x} \in \mathbb{R}^n$ $\square$

Proof of Theorem 1: Lemma 2

Handwritten

Lemma 2 ($B \implies A$)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, convex, and $f(\mathbf{x}^\star) \leq f(\mathbf{x}) \, \forall \mathbf{x} \in N(\mathbf{x}^\star, \epsilon)$. Then, $\nabla f(\mathbf{x}^\star) = 0$.

Proof of Lemma 2

- Since $\mathbf{x}^\star$ is a local minimizer, $f(\mathbf{x}^\star + \eta \cdot \mathbf{p}) \geq f(\mathbf{x}^\star), \quad \forall \eta \in [0, \epsilon], \forall \mathbf{p} \in \mathbb{R}^n$
- Therefore, $0 \geq \lim_{\eta \to 0} \frac{f(\mathbf{x}^\star + \eta \cdot \mathbf{p}) - f(\mathbf{x}^\star)}{\eta} = \nabla f(\mathbf{x}^\star)^\top \mathbf{p} \quad \forall \mathbf{p} \in \mathbb{R}^n$
- Choose $\mathbf{p} = -\nabla f(\mathbf{x}^\star) : \nabla f(\mathbf{x}^\star)^\top \mathbf{p} = -\nabla f(\mathbf{x}^\star)^\top \nabla f(\mathbf{x}^\star) =: -\|\nabla f(\mathbf{x}^\star)\|_2^2 \leq 0$
- Since $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} \geq 0 \, \forall \mathbf{p}$ and $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} \leq 0$, $\mathbf{p} = -\nabla f(\mathbf{x}^\star) \implies \nabla f(\mathbf{x}^\star) = 0$ $\square$

Proof of Theorem 1: Summary

Theorem 1: Global minimizer of unconstrained convex functions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable and convex. Then, $\mathbf{x}^\star \in \mathbb{R}^n$ is the **global minimizer**, i.e., $f(\mathbf{x}^\star) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, iff $\nabla f(\mathbf{x}^\star) = \mathbf{0}$, i.e., iff $x^\star$ is a stationary point.

Proof of Theorem 1 (sketch)

- Lemma 1: If $\mathbf{x}^\star \in \mathbb{R}^n$ is a stationary point, then $\mathbf{x}^\star$ is a global minimizer
- Trivial: If $\mathbf{x}^\star \in \mathbb{R}^n$ is a global minimizer, then $\mathbf{x}^\star \in \mathbb{R}^n$ is a local minimizer
- Lemma 2: If $\mathbf{x}^\star \in \mathbb{R}^n$ is a local minimizer, then $\mathbf{x}^\star$ is a stationary point

- First order necessary conditions are necessary and sufficient
- Every local minimum is a global minimum
- These are properties that distinguish convex and non-convex optimization

Convex constrained optimization problem

- The optimization problem of MPC is a **constrained** optimization problem
- **Equality constraints** through initial condition and system dynamics
- **Inequality constraints** through constraints on input and state (box constraints)

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \quad \text{subject to} \quad c_i(\mathbf{x}) \leq 0 \quad \forall i \in \mathcal{I}, \quad c_i(\mathbf{x}) = 0 \quad \forall i \in \mathcal{E}$$

- Inequality constraints $c_i, i \in \mathcal{I} \subseteq \mathbb{N}$, equality constraints $c_i, i \in \mathcal{E} \subseteq \mathbb{N}$
- Equivalent formulation by defining the **feasible domain** Ω:

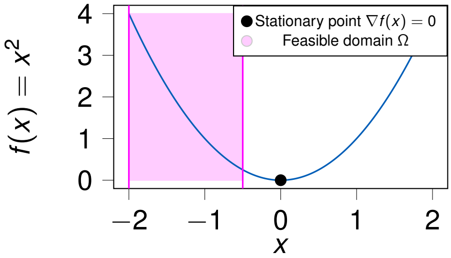
$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}), \quad \Omega := \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) \leq 0 \; \forall i \in \mathcal{I} \cap c_i(\mathbf{x}) = 0 \; \forall i \in \mathcal{E}\}$$

Definition: Convex constrained optimization problem

A constrained optimization problem is convex iff the **objective function f is convex** and the resulting **feasible domain Ω is convex**, i.e., the **equality constraints are linear** and the **inequality constraints are convex**.

Stationarity for convex constrained optimization problems

- Stationarity condition $\nabla f(\mathbf{x}^\star)$ is necessary and sufficient condition for $\mathbf{x}^\star$ being a global minimizer for **convex unconstrained optimization**
- Unfortunately, this does not (directly) translate to the **constrained case**
- Example: $\min_{x \in \Omega} x^2, \quad \Omega := [-2, -0.5]$, where $\nabla f(0) = 0$ but $0 \notin \Omega$



Stationarity conditions

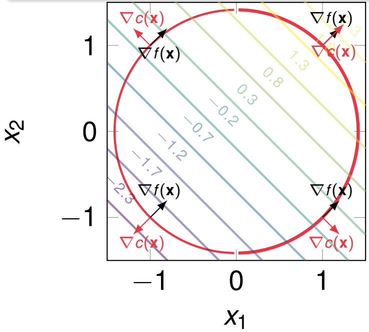
Which stationarity conditions are fulfilled at the solution of constrained optimization problems?

Stationarity with single equality constraint (graphical)

Stationarity conditions

Which stationarity conditions are fulfilled at the solution of constrained optimization?

$$\min_{\mathbf{x}\in\mathbb{R}^2} x_1 + x_2, \quad \text{subject to } x_1^2 + x_2^2 - 2 = 0$$



- $f(\mathbf{x}) = x_1 + x_2 \implies \nabla f(\mathbf{x}) = [1, 1]^\top$
- $c(\mathbf{x}) = x_1^2 + x_2^2 - 2 \implies \nabla c(\mathbf{x}) = [2x_1, 2x_2]^\top$
- Equivalent: $c(\mathbf{x}) = -x_1^2 - x_2^2 + 2$ (sign-switch)
- At global minimum $\mathbf{x}^\star = [-1, -1]^\top$, gradients $\nabla f(\mathbf{x})$ and $\nabla c(\mathbf{x})$ are **parallel**; the sign of λ does not matter because of the equivalent sign-switch

$$\nabla f(\mathbf{x}) + \lambda \nabla c(\mathbf{x}) \stackrel{!}{=} 0, \quad \lambda \in \mathbb{R}$$

Stationarity with single equality constraint (mathematical)

- A feasible point $\mathbf{x}^\star$ is not optimal if there exists a step in direction $\mathbf{p}$ such that **feasibility is retained** and **the value of the objective is decreased**

Feasibility

- Convex feasible set $\iff$ linear constraints: $c(\mathbf{x}^\star + \mathbf{p}) = c(\mathbf{x}^\star) + \nabla c(\mathbf{x}^\star)^\top \mathbf{p}$
- To remain feasible: $c(\mathbf{x}^\star + \mathbf{p}) = 0 \Rightarrow \nabla c(\mathbf{x}^\star)^\top \mathbf{p} = 0$ since $c(\mathbf{x}^\star) = 0$

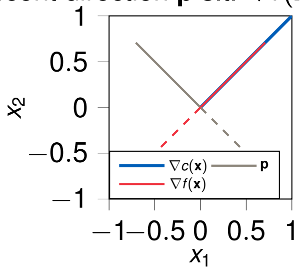
Decreasing value of objective function

- Decrease in the objective function obtained in descent direction $\mathbf{p}$ s.t. $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} < 0$

Optimality Handwritten

- At the optimum $\mathbf{x}^\star$, for any $\mathbf{p}$ with $\nabla c(\mathbf{x})^\top \mathbf{p} = 0$ (orthogonality), we need $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} \geq 0$ ($\iff \nabla f(\mathbf{x}^\star)^\top \mathbf{p} = 0$)

$$\nabla f(\mathbf{x}) + \lambda \nabla c(\mathbf{x}) \stackrel{!}{=} 0, \quad \lambda \in \mathbb{R}$$



Stationarity with equality constraints (summary)

For one equality constraint, we require...

$$\nabla f(\mathbf{x}) + \lambda \nabla c(\mathbf{x}) \stackrel{!}{=} 0, \quad \lambda \in \mathbb{R}$$

For multiple equality constraints, we require...

$$\nabla f(\mathbf{x}) + \textstyle\sum_{i \in \mathcal{E}} \lambda_i \nabla c_i(\mathbf{x}) \stackrel{!}{=} 0, \quad \lambda_i \in \mathbb{R}$$

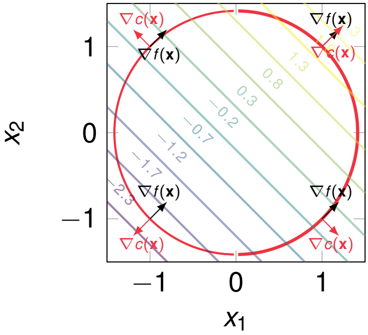
First-order necessary conditions (equality constraints)

$$\nabla f(\mathbf{x}) + \sum_{i \in \mathcal{E}} \lambda_i \nabla c_i(\mathbf{x}) = 0 \qquad \texttt{\% Stationarity}$$

$$c_i(\mathbf{x}) = 0, \ \forall i \in \mathcal{E} \qquad \texttt{\% Primal feasibility}$$

Stationarity with single inequality constraint (graphical)

$$\min_{\mathbf{x}\in\mathbb{R}^2} x_1 + x_2, \quad \text{subject to } x_1^2 + x_2^2 - 2 \leq 0$$



- $f(\mathbf{x}) = x_1 + x_2 \implies \nabla f(\mathbf{x}) = [1, 1]^\top$
- $c(\mathbf{x}) = x_1^2 + x_2^2 - 2 \implies \nabla c(\mathbf{x}) = [2x_1, 2x_2]^\top$
- In constrast to equality constraints, a sign-switch of inequality constraints is not equivalent
- At global minimum $\mathbf{x}^\star = [-1, -1]^\top$, gradient of objective function $\nabla f(\mathbf{x})$ and of constraint $\nabla c(\mathbf{x})$ are **parallel** and show in **opposite directions**

$$\nabla f(\mathbf{x}) + \mu \nabla c(\mathbf{x}) \stackrel{!}{=} 0, \quad \mu \geq 0$$

Stationarity with single inequality constraint (mathematical)

- A feasible point $\mathbf{x}^\star$ is not optimal if there exists a step in direction $\mathbf{p}$ such that **feasibility is retained** and **the value of the objective is decreased**

Feasibility

- $c(\mathbf{x}^\star + \mathbf{p}) = c(\mathbf{x}^\star) + \nabla c(\mathbf{x}^\star)^\top \mathbf{p} + \mathcal{O}(\mathbf{p}^\top \mathbf{p}) = c(\mathbf{x}^\star) + \nabla c(\mathbf{x}^\star)^\top \mathbf{p}$ for sufficiently small $\mathbf{p}$
- To remain feasible: $0 \geq c(\mathbf{x}^\star + \mathbf{p}) = c(\mathbf{x}^\star) + \nabla c(\mathbf{x}^\star)^\top \mathbf{p}$

Decreasing value of objective function

- Decrease in the objective function obtained in descent direction $\mathbf{p}$ s.t. $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} < 0$

Optimality Handwritten

If $c(\mathbf{x}^\star) < 0$ (inequality constraint is inactive)... $\implies \nabla f(\mathbf{x}^\star) \overset{!}{=} 0$

- $c(\mathbf{x}^\star + \mathbf{p}) \leq 0$ is always satisfied for sufficiently small $\mathbf{p}$
- $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} \not< 0$ if $\nabla f(\mathbf{x}^\star) = 0$

If $c(\mathbf{x}^\star) = 0$ (inequality constraint is active)... $\implies \nabla f(\mathbf{x}^\star) + \mu \nabla c(\mathbf{x}^\star) \overset{!}{=} 0, \;\; \mu \geq 0$

- $0 \geq c(\mathbf{x}^\star + \mathbf{p}) = c(\mathbf{x}^\star) + \nabla c(\mathbf{x}^\star)^\top \mathbf{p} = \nabla c(\mathbf{x}^\star)^\top \mathbf{p}$
- $\nabla f(\mathbf{x}^\star)^\top \mathbf{p} \not< 0$ if $\nabla f(\mathbf{x}^\star) + \mu c(\mathbf{x}^\star) = 0, \quad \mu \geq 0$ (parallelity in opposite directions)

Complementary slackness of inequalities

- If $c(\mathbf{x}^\star) < 0$ (inequality constraint is inactive) $\implies \nabla f(\mathbf{x}^\star) \stackrel{!}{=} 0$
- If $c(\mathbf{x}^\star) = 0$ (inequality constraint is active)... $\implies \nabla f(\mathbf{x}) + \mu \nabla c(\mathbf{x}) \stackrel{!}{=} 0,\ \mu \geq 0$
- This if-else relationship is achieved with...

$$
\begin{aligned}
\nabla f(\mathbf{x}) + \mu c(\mathbf{x}) &= 0 \quad \texttt{\% Stationarity} \\
\mu &\geq 0 \quad \texttt{\% Dual feasibility} \\
\mu c(\mathbf{x}) &= 0 \quad \texttt{\% Complementary slackness}
\end{aligned}
$$

- Vista: Complementary slackness is **non-smooth** constraint, making it difficult for numerical solvers

Stationarity with inequality constraints (summary)

First-order necessary conditions (single inequality constraint)

$$\nabla f(\mathbf{x}) + \mu \nabla c(\mathbf{x}) = 0 \quad \texttt{\% Stationarity}$$
$$c(\mathbf{x}) \leq 0 \quad \texttt{\% Primal feasibility}$$
$$\mu \geq 0 \quad \texttt{\% Dual feasibility}$$
$$\mu c(\mathbf{x}) = 0 \quad \texttt{\% Complementary slackness}$$

First-order necessary conditions (multiple inequality constraints)

$$\nabla f(\mathbf{x}) + \textstyle\sum_{i \in \mathcal{I}} \mu_i \nabla c_i(\mathbf{x}) = 0 \quad \texttt{\% Stationarity}$$
$$c_i(\mathbf{x}) \leq 0, \ \forall i \in \mathcal{I} \quad \texttt{\% Primal feasibility}$$
$$\mu_i \geq 0, \ \forall i \in \mathcal{I} \quad \texttt{\% Dual feasibility}$$
$$\mu_i c_i(\mathbf{x}) = 0 \quad \texttt{\% Complementary slackness}$$

Karush-Kahn-Tucker (KKT) conditions

Theorem 2: KKT conditions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a **differentiable and convex objective function**. Let the constraints $c_i(\mathbf{x}) = 0 \forall i \in \mathcal{E}$ and $c_i(\mathbf{x}) \leq 0 \forall i \in \mathcal{I}$ cause **a convex feasible domain** and additionally have certain regularity properties. Then, $\mathbf{x} \in \mathbb{R}^n$ is the **global optimum of the constrained minimization problem** with corresponding multipliers $\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star$ iff:

$$
\begin{aligned}
\nabla f(\mathbf{x}^\star) + \textstyle\sum_{i\in\mathcal{E}} \lambda_i^\star \nabla c_i(\mathbf{x}) + \sum_{i\in\mathcal{I}} \mu_i^\star \nabla c_i(\mathbf{x}) &= 0 && \texttt{\% Stationarity} \\
c_i(\mathbf{x}) &= 0, \ \forall i \in \mathcal{E} && \texttt{\% Primal feasibility} \\
c_i(\mathbf{x}) &\leq 0, \ \forall i \in \mathcal{I} && \texttt{\% Primal feasibility} \\
\mu_i^\star &\geq 0, \ \forall i \in \mathcal{I} && \texttt{\% Dual feasibility} \\
\mu_i^\star c_i(\mathbf{x}) &= 0 && \texttt{\% Complementary slackness}
\end{aligned}
$$

Optimization methods

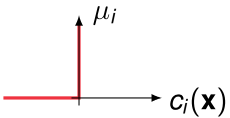
- We have derived the **KKT conditions**, which are **necessary and sufficient** conditions for the global minimizer of convex constrained problems like the one of **linear MPC**
- In linear MPC, we solve a convex optimization problem at every time-step and require the global minimum for, e.g., performance guarantees
- How do **solvers find the global minimizer** at each time-step **in practice**?
- **qpOASES**:[1] Solver based on active-set methods, which have exponential runtime (in the worst case)
- **IPOPT**:[2] Solver based on interior point-methods, which have polynomial runtime

[1] Ferreau et al. "qpOASES: A parametric active-set algorithm for quadratic programming," 2014.

[2] Wächter et al., "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," 2006.

Interior point methods

- Interior point methods move along the **interior of the feasible domain** to find global the minimum $\mathbf{x}^\star$
- They can be seen as a **numerical way to solve KKT conditions**
- Main numerical challenge for solvers is the complementary slackness $\mu_i c_i(\mathbf{x}) = 0$
- Complementary slackness results in a corner with no interior $\rightarrow$ **non-smooth constraint**
- This non-smooth constraint is extremely **challenging** to handle for numerical solvers

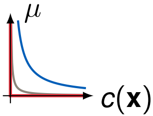


Algorithms based on interior point methods

How can an algoritm converge to the solution of the KKT conditions and circumvent the issue caused by the complementary slackness?

Algorithms based on interior point methods

Handwritten



- Instead of $\mu c(\mathbf{x}) = 0$, we require $\mu c(\mathbf{x}) = \tau, \ \tau > 0$
- Example: Start with $\tau = 0.01$, then $\tau = 0.001$, continue decreasing $\tau \to 0$ iteratively
- Another technique of interior point methods is to replace inequality constraints by a **logarithmic barrier** in the objective function: Instead of $\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$ s.t. $c(\mathbf{x}) \leq 0$, we solve $\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) - \tau \log(-c(\mathbf{x}))$
- Since $\log(\cdot)$ is only defined for positive arguments, we enforce $c(\mathbf{x}) < 0$, i.e., we stay in the **interior of the feasible domain**
- Larger τ causes smaller values of $c(\mathbf{x})$, whereas smaller τ enables $c(\mathbf{x}) \to 0$
- The solution path for decreasing τ is called the **central path** and is guaranteed to **converge to the KKT conditions** (under suitable assumptions)

Nonlinear MPC

Recap:

- If the system dynamics are linear, we can formulate a linear MPC
- Many real-world systems are nonlinear but we could approximate linearly
- Linear MPC yields convex optimization problems, which are simpler to solve, thus satisfying real-time requirements for systems with high sampling rates

However, **most MPC formulations remain nonlinear** because:

- System dynamics are too nonlinear to approximate linearly
- Simple approximations of the dynamics would lead to **losing closed-loop guarantees of MPC** like stability, recursive feasibility, and constraint satisfaction

Nonlinear MPC

Nonlinear MPC yields a nonlinear program (NLP), which requires the solution of a non-convex optimization problem.

Approximate nonlinear MPC

Non-convex optimization

"The great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity," Rockafellar, *Convex Analysis*, 1970.

- In non-convex optimization problems, we can have local minima that are not the global minimum $\rightarrow$ KKT conditions are necessary but **not sufficient** anymore
- For nonlinear MPC, instead of approximating the system dynamics, we can directly approximate the solution of the NLP, i.e., the **implicitly defined control law** $u = g(x)$
- Problem definition: Compute **explicit function** h that approximates g with $|h(x) - g(x)| \leq \epsilon, \forall x \in \Omega, \forall \epsilon > 0$
- ALKIA-X:[3] yielding fast-to-evaluate nonlinear MPC with performance guarantees

[3] A. Tokmak, C. Fiedler, M. N. Zeilinger, S. Trimpe, J. Köhler, "Automatic nonlinear MPC approximation with closed-loop guarantees," IEEE Transactions on Automatic Control (submitted), 2024.

Learning outcomes

By the end of this lecture, you should be able to. . .

- Recognize whether an MPC problem is linear and understand what this implies for the underlying optimization problem
- Know the properties of convex optimization problems
- Derive KKT conditions to solve convex constrained optimization problems
- Explain the basic idea of interior point methods
- Know that the world of nonlinear MPC is huge

Exam

- Content: the second part of the course (state-space representations, stability analysis, controllability, observability, pole placement, optimal control, stochastic optimal control, model predictive control, convex optimization)
- As in the previous exam, we will have calculation exercises and some where you need to explain something
- A calculator is not allowed (and not needed)
- You can use either the databook provided in MyCourses without annotations or added formulas or a handwritten, one-sided A4 page on which you can write whatever you feel might help you during the exam
- We will collect the sheets/databooks after the exams—please stick to the rules!
- The exam will be Tuesday, 3.12., AS2, 13:00 – 16:00

Feedback

Feedback

Please leave some feedback for today's lecture: `https://presemo.aalto.fi/digoptctrl`

References

1. A. Mitsos, "Applied numerical optimization," Lecture 2, 2021.
2. A. Mitsos, "Applied numerical optimization," Lecture 5, 2021.
3. A. Mitsos, "Applied numerical optimization," Lecture 7, 2021.
4. D. Abel, "Model predictive control of energy systems," Lecture 2, 2021.
5. H.J. Ferreau et al. “qpOASES: A parametric active-set algorithm for quadratic programming,” 2014.
6. A. Wächter et al., “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” 2006.
7. A. Tokmak et al., "Automatic nonlinear MPC approximation with closed-loop guarantees," 2024.
8. R.T. Rockafellar, "Convex Analysis," 1970.