Model Predictive Control

3. Fundamentals of Optimization

Jun.-Prof. Dr.-Ing. Daniel Görges
Juniorprofessur für Elektromobilität
Technische Universität Kaiserslautern

Concepts from Calculus

Gradient, Hessian, and Jacobian

Tutorial

Definition 3.1 The gradient of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as $\nabla f(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix}^T$.

Definition 3.2 The Hessian of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as $\boldsymbol{H}_f(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$.

Definition 3.3 The Jacobian of a function $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$ is defined as $\boldsymbol{J}_f(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$.

Nonlinear Optimization Problem

Problem 3.1 A nonlinear optimization problem is defined in standard form as

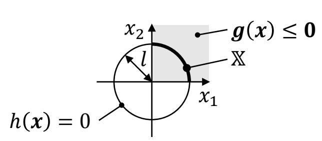
$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \quad \text{with } f: \mathbb{R}^n \to \mathbb{R} \quad \text{cost function or objective function} \tag{3.1}$$

$$\text{subject to } \begin{cases} \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} & \text{with } \boldsymbol{h}: \mathbb{R}^n \to \mathbb{R}^m \quad \text{equality constraints} \quad (3.2) \\ \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} & \text{with } \boldsymbol{g}: \mathbb{R}^n \to \mathbb{R}^p \quad \text{inequality constraints} \quad (3.3) \end{cases}$$

- **Symbols**
  - The vector $\boldsymbol{x} = (x_1 \quad x_2 \quad \cdots \quad x_n)^T \in \mathbb{R}^n$ is denoted as decision variable or optimization variable
  - The solution $\boldsymbol{x}^* \in \mathbb{R}^n$ of Problem 3.1 is denoted as minimizer
- **Remark**
  - For $m < n$ the equality constraints (3.2) are underdetermined ✓
  - For $m = n$ the equality constraints (3.2) are determined for $h_i, i \in \{1, \dots, m\}$ independent ✗
  - For $m > n$ the equality constraints (3.2) are overdetermined ✗

Nonlinear Optimization Problem

- **Assumption**
  - Cost function $f \in \mathcal{C}^2$, functions $h_i \in \mathcal{C}^1, i \in \{1, \dots, m\}$ and $g_j \in \mathcal{C}^1, j \in \{1, \dots, p\}$ where $\mathcal{C}^j$ is the set of j times continuously differentiable functions
- **Remarks**
  - Nonsmooth optimization if assumption not fulfilled (not considered in this lecture)
  - Integer optimization if $\boldsymbol{x} \in \mathbb{Z}^n$ (not considered in this lecture)
- **Example**
  - Maximization of the area of a right triangle with legs x_1 and x_2 and a given hypotenuse l
  - Cost function $f(\boldsymbol{x}) = -\frac{1}{2} x_1 x_2$
  - Equality constraint $h(\boldsymbol{x}) = x_1^2 + x_2^2 - l^2 = 0$
  - Inequality constraints $g_1(\boldsymbol{x}) = -x_1 \leq 0, g_2(\boldsymbol{x}) = -x_2 \leq 0$



Nonlinear Optimization

Nonlinear Optimization Problem

Problem 3.2 A nonlinear optimization problem is defined as

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \qquad \text{with } f: \mathbb{R}^n \to \mathbb{R} \qquad \text{cost function} \qquad (3.4)$$

$$\text{subject to } \boldsymbol{x} \in \mathbb{X} \qquad \text{with } \mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}\} \qquad \text{feasible set} \qquad (3.5)$$

- **Symbols**
  - A vector $\boldsymbol{x} \in \mathbb{X}$ is denoted as feasible point
- **Remarks**
  - Problem 3.2 is an alternative formulation of Problem 3.1
  - Problem 3.2 can be written even more briefly as $\min_{\boldsymbol{x} \in \mathbb{X}} f(\boldsymbol{x})$
  - Note that considering a minimization problem is not restrictive since a maximization problem can be transformed into a minimization problem using $\max_{\boldsymbol{x} \in \mathbb{X}} f(\boldsymbol{x}) = \min_{\boldsymbol{x} \in \mathbb{X}} -f(\boldsymbol{x})$

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Nonlinear Optimization

Local Minimum and Global Minimum

Definition 3.4 The cost function $f(\boldsymbol{x})$ has a local minimum at the point $\boldsymbol{x}^* \in \mathbb{X}$ if there exists an $\varepsilon > 0$ such that $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{X} \backslash \{\boldsymbol{x}^*\}$ and $\|\boldsymbol{x} - \boldsymbol{x}^*\| < \varepsilon$. If $\leq$ is replaced by $<$, then the local minimum is a strict local minimum.

Definition 3.5 The cost function $f(\boldsymbol{x})$ has a global minimum at the point $\boldsymbol{x}^* \in \mathbb{X}$ if $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{X} \backslash \{\boldsymbol{x}^*\}$. If $\leq$ is replaced by $<$, then the global minimum is a unique or strict global minimum.

Theorem 3.1 A global minimum exists if

(1) the feasible set $\mathbb{X}$ is bounded, i.e. $\exists \alpha \in \mathbb{R}: \|\boldsymbol{x}\| \leq \alpha \ \forall \boldsymbol{x} \in \mathbb{X}$,

(2) the feasible set is not empty, i.e. $\mathbb{X} \neq \emptyset$.

- **Remark**
  - Note the Theorem 3.1 is only sufficient

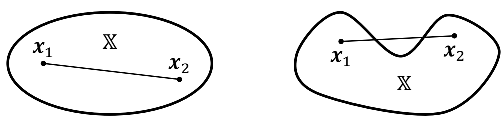
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Convex Optimization

Convex Sets

Definition 3.6 A set $\mathbb{X}$ is convex if $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbb{X}$ for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$ and $\alpha \in [0,1]$.

- **Interpretation**
  - Note that $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ with $\alpha \in [0,1]$ represents the line segment between the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$
  - A set is thus convex if the line segment connecting two arbitrary points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is also in the set



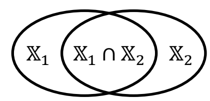
convex | not convex

- **Properties**

(1) $\mathbb{X}$ convex, $\beta \in \mathbb{R}$ $\Rightarrow$ $\beta\mathbb{X} = \{\boldsymbol{x} | \boldsymbol{x} = \beta\boldsymbol{v}, \boldsymbol{v} \in \mathbb{X}\}$ convex

(2) $\mathbb{X}_1, \mathbb{X}_2$ convex $\Rightarrow$ $\mathbb{X}_1 + \mathbb{X}_2 = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1 \in \mathbb{X}_1, \boldsymbol{v}_2 \in \mathbb{X}_2\}$ convex

(3) $\mathbb{X}_1, \mathbb{X}_2$ convex $\Rightarrow$ $\mathbb{X}_1 \cap \mathbb{X}_2$ convex



TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Convex Optimization

Convex Functions

Definition 3.7 A function $f : \mathbb{X} \to \mathbb{R}$ is convex on a convex set $\mathbb{X}$ if

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X} \quad \forall \alpha \in [0,1].$$

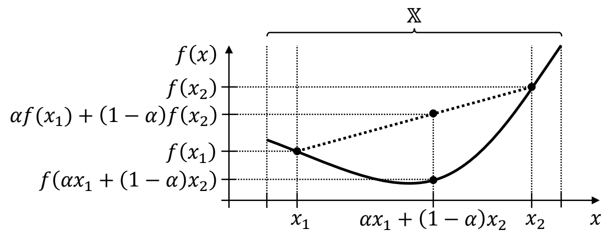
Definition 3.8 A function $f : \mathbb{X} \to \mathbb{R}$ is strictly convex on a convex set $\mathbb{X}$ if

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) < \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}, \boldsymbol{x}_1 \neq \boldsymbol{x}_2 \quad \forall \alpha \in (0,1).$$

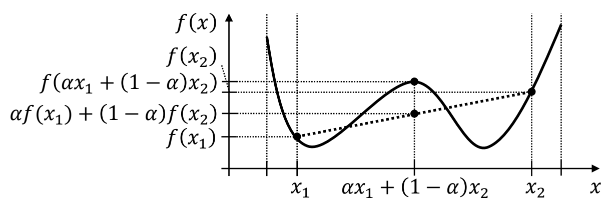
Definition 3.9 A function $f : \mathbb{X} \to \mathbb{R}$ is (strictly) concave on a convex set $\mathbb{X}$ if $-f$ is (strictly) convex.

- **Interpretation**
  - A function f is convex if the secant connecting two arbitrary points $(\boldsymbol{x}_1, f(\boldsymbol{x}_1))$ and $(\boldsymbol{x}_2, f(\boldsymbol{x}_2))$ lies on or above the graph of f

Convex Functions



convex



not convex

Convex Functions

- **Example**
  - Is the function $f(\boldsymbol{x}) = x_1 x_2$ convex on $\mathbb{X} = \{\boldsymbol{x} | x_1 \geq 0, x_2 \geq 0\}$?
  - Consider the points $\boldsymbol{x}_1 = (1 \quad 2)^T \in \mathbb{X}$ and $\boldsymbol{x}_2 = (2 \quad 1)^T \in \mathbb{X}$, then

$$\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 = \alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha + 2 - 2\alpha \\ 2\alpha + 1 - \alpha \end{pmatrix} = \begin{pmatrix} 2-\alpha \\ 1+\alpha \end{pmatrix}$$

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) = (2-\alpha)(1+\alpha) = 2 + \alpha - \alpha^2$$

$$\alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) = 2\alpha + 2(1-\alpha) = 2$$

  - Consider e.g. $\alpha = \frac{1}{2}$, then

$$f\left(\tfrac{1}{2}\boldsymbol{x}_1 + \tfrac{1}{2}\boldsymbol{x}_2\right) = 2 + \tfrac{1}{2} - \tfrac{1}{4} = \tfrac{9}{4} > \tfrac{1}{2} f(\boldsymbol{x}_1) + \tfrac{1}{2} f(\boldsymbol{x}_2) = 2$$

  - The function $f(\boldsymbol{x})$ is not convex on $\mathbb{X}$

Convex Optimization

Convex Functions

- **Properties**

  (1) $f_i(\boldsymbol{x})$ convex on $\mathbb{X}$, $\alpha_i \geq 0$, $i \in \{1, \dots, N\}$ $\Rightarrow$ $f(\boldsymbol{x}) = \sum_{i=1}^{N} \alpha_i f_i(\boldsymbol{x})$ convex on $\mathbb{X}$

  (2) $f(\boldsymbol{x})$ convex on $\mathbb{X}$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$ $\Rightarrow$ $f(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2)$ convex on $\mathbb{X}$ for $\alpha \in [0,1]$

  (3) $f(\boldsymbol{x})$ convex on $\mathbb{X}$ $\Rightarrow$ $\{\boldsymbol{x} \in \mathbb{X} | f(\boldsymbol{x}) \leq 0\}$ convex

  (4) $\{\boldsymbol{x} \in \mathbb{X} | f(\boldsymbol{x}) \leq 0\}$ convex $\nRightarrow$ $f(\boldsymbol{x})$ convex on $\mathbb{X}$

  (5) $f(\boldsymbol{x}) \in \mathcal{C}^1$ convex on $\mathbb{X}$ $\Leftrightarrow$ $f(\boldsymbol{x}_2) \geq f(\boldsymbol{x}_1) + (\boldsymbol{x}_2 - \boldsymbol{x}_1)^T \nabla f(\boldsymbol{x}_1)$ $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$

  (6) $f(\boldsymbol{x}) \in \mathcal{C}^1$ strictly convex on $\mathbb{X}$ $\Leftrightarrow$ $f(\boldsymbol{x}_2) > f(\boldsymbol{x}_1) + (\boldsymbol{x}_2 - \boldsymbol{x}_1)^T \nabla f(\boldsymbol{x}_1)$ $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}, \boldsymbol{x}_1 \neq \boldsymbol{x}_2$

  (7) $f(\boldsymbol{x}) \in \mathcal{C}^2$ convex on $\mathbb{X}$ $\Leftrightarrow$ $\boldsymbol{H}_f(\boldsymbol{x}) \succcurlyeq \boldsymbol{0}$ $\forall \boldsymbol{x} \in \mathbb{X}$

  (8) $f(\boldsymbol{x}) \in \mathcal{C}^2$ strictly convex on $\mathbb{X}$ $\Leftarrow$ $\boldsymbol{H}_f(\boldsymbol{x}) \succ \boldsymbol{0}$ $\forall \boldsymbol{x} \in \mathbb{X}$

- **Example**
  - When is the quadratic form $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$ with $\boldsymbol{P} = \boldsymbol{P}^T$ convex and strictly convex on $\mathbb{R}^n$?
  - It is $\boldsymbol{H}_f(\boldsymbol{x}) = \boldsymbol{P}$. Thus, $f(\boldsymbol{x})$ is convex on $\mathbb{R}^n$ iff $\boldsymbol{P} \succcurlyeq \boldsymbol{0}$ and strictly convex on $\mathbb{R}^n$ iff $\boldsymbol{P} \succ \boldsymbol{0}$ (!)

Technische Universität Kaiserslautern

Convex Optimization

Convex Optimization Problem

Problem 3.3 Consider the nonlinear optimization problem

$\min\limits_{\boldsymbol{x}} f(\boldsymbol{x})$	with $f: \mathbb{R}^n \to \mathbb{R}$	cost function	(3.6)
subject to $\boldsymbol{x} \in \mathbb{X}$	with $\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}\}$	feasible set	(3.7)

The problem is convex if the feasible set $\mathbb{X}$ is convex and the cost function f is convex on the feasible set $\mathbb{X}$.
It is furthermore strictly convex if the cost function f is also strictly convex on the feasible set $\mathbb{X}$.

- **Remark**
  - Proving convexity of the feasible set $\mathbb{X}$ is very involved except in special cases
  - For example, if the functions $h_i(\boldsymbol{x}), i \in \{1, \dots, m\}$ are linear and the functions $g_j(\boldsymbol{x}), j \in \{1, \dots, p\}$ are convex on $\mathbb{X}$, then the feasible set $\mathbb{X}$ is an intersection of convex sets and therefore convex

Theorem 3.2 Let $f: \mathbb{X} \to \mathbb{R}$ be a convex function defined on the convex set $\mathbb{X}$. Then each local minimum of f on $\mathbb{X}$ is also a global minimum of f on $\mathbb{X}$ and the set of global minima of f on $\mathbb{X}$ is convex.

Optimality Conditions

Definitions

Definition 3.10 An inequality constraint $g_j(\boldsymbol{x}) \leq 0$ is denoted as active at a feasible point $\boldsymbol{x} \in \mathbb{X}$ if $g_j(\boldsymbol{x}) = 0$ and as inactive at a feasible point $\boldsymbol{x} \in \mathbb{X}$ if $g_j(\boldsymbol{x}) < 0$.

- **Remark**
  - Active inequality constraints will be denoted in the following by $\boldsymbol{g}^{\mathrm{a}}: \mathbb{R}^n \to \mathbb{R}^{p^{\mathrm{a}}}, \boldsymbol{g}^{\mathrm{a}}(\boldsymbol{x}) = \boldsymbol{0}$
  - Inactive inequality constraints will be denoted in the following by $\boldsymbol{g}^{\mathrm{i}}: \mathbb{R}^n \to \mathbb{R}^{p^{\mathrm{i}}}, \boldsymbol{g}^{\mathrm{i}}(\boldsymbol{x}) < \boldsymbol{0}$
  - Note that $p^{\mathrm{a}} + p^{\mathrm{i}} = p$

Definition 3.11 The feasible point $\boldsymbol{x} \in \mathbb{X}$ is denoted as regular point if the vectors

$$\nabla h_i(\boldsymbol{x}), i \in \{1, \dots m\} \text{ and } \nabla g_j^{\mathrm{a}}(\boldsymbol{x}), j \in \{1, \dots, p^{\mathrm{a}}\}$$

are linearly independent.

Optimality Conditions

Karush-Kuhn-Tucker (KKT) Conditions

Theorem 3.3 Let $\boldsymbol{x}^* \in \mathbb{R}^n$ be a regular point and a local minimizer to Problem 3.1 and introduce the function $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x})$. Then there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that

(1) $\nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \nabla f(\boldsymbol{x}^*) + \boldsymbol{J}_{\boldsymbol{h}}^T(\boldsymbol{x}^*)\boldsymbol{\lambda}^* + \boldsymbol{J}_{\boldsymbol{g}}^T(\boldsymbol{x}^*)\boldsymbol{\mu}^* = \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^m \nabla \boldsymbol{h}_i(\boldsymbol{x}^*)\lambda_i^* + \sum_{j=1}^p \nabla \boldsymbol{g}_j(\boldsymbol{x}^*)\mu_i^* = \boldsymbol{0}$

(2) $\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$

(3) $\boldsymbol{g}(\boldsymbol{x}^*) \leq \boldsymbol{0}$

(4) $\boldsymbol{g}^T(\boldsymbol{x}^*)\boldsymbol{\mu}^* = 0$

(5) $\boldsymbol{\mu}^* \geq \boldsymbol{0}$.

- **Remarks**
  - No constraints? Only the green term is relevant.
  - Only equality constraints? Only the green term and blue terms are relevant.
  - Condition (4) can also be written as $g_j(\boldsymbol{x}^*)\mu_j^* = 0, j \in \{1, \dots, p\}$

Optimality Conditions

Karush-Kuhn-Tucker (KKT) Conditions

- **Symbols**
  - The function $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x})$ is called Lagrangian
  - The vector $\boldsymbol{\lambda}$ is called Lagrange multiplier
  - The vector $\boldsymbol{\mu}$ is called Karush-Kuhn-Tucker multiplier
- **Properties**
  - $\mu_j^* = 0$ if $g_j(\boldsymbol{x}^*) < 0$ (i.e. if the inequality constraint is inactive) due to conditions (3) to (5)
  - $\mu_j^* \geq 0$ if $g_j(\boldsymbol{x}^*) = 0$ (i.e. if the inequality constraint is active) due to conditions (3) to (5)
  - $\mu_j < 0$ and $g_j(\boldsymbol{x}) = 0$ (i.e. the inequality constraint is active) while (1) to (4) fulfilled indicates that the cost $f(\boldsymbol{x})$ can be reduced by setting $g_j(\boldsymbol{x}) < 0$ (i.e. by setting the inequality constraint inactive)
- **Remarks**
  - The KKT conditions presume constraint qualification. Constraint qualification is ensured in most optimization problems, e.g. if $\boldsymbol{h}$ and $\boldsymbol{g}^{\mathrm{a}}$ are linear, see [PLB12, p. 78] for details.

Optimality Conditions

Karush-Kuhn-Tucker (KKT) Conditions

- **Remarks**
  - The KKT conditions are only necessary for general nonlinear optimization problems (Problem 3.1)
  - The KKT conditions are necessary and sufficient for convex optimization problems (Problem 3.3)
  - The KKT conditions can usually be evaluated analytically for simple optimization problems
  - The KKT conditions must generally be evaluated numerically for complex optimization problems
- **Example**
  - Maximization of the area of a right triangle with legs x_1 and x_2 and a given hypotenuse l (Slide 3-4)
  - Cost function $f(\boldsymbol{x}) = -\frac{1}{2} x_1 x_2$
  - Constraints $h(\boldsymbol{x}) = x_1^2 + x_2^2 - l^2 = 0,\ g_1(\boldsymbol{x}) = -x_1 \leq 0,\ g_2(\boldsymbol{x}) = -x_2 \leq 0$
  - Lagrangian $L(\boldsymbol{x}, \lambda, \boldsymbol{\mu}) = -\frac{1}{2} x_1 x_2 + \lambda(x_1^2 + x_2^2 - l^2) - \mu_1 x_1 - \mu_2 x_2$
  - An analytical solution can be obtained by analyzing all combinations of active and inactive inequality constraints to determine candidate solutions and then comparing the candidate solutions w.r.t. cost

Optimality Conditions

Karush-Kuhn-Tucker (KKT) Conditions

- **Example**
  - Case 1 $\quad g_1(\boldsymbol{x}^*) < 0$ (inactive), $g_2(\boldsymbol{x}^*) < 0$ (inactive), then $\mu_1^* = \mu_2^* = 0$

$$\left.\begin{array}{l}\left.\begin{array}{l}\frac{\partial}{\partial x_1}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_2^* + 2\lambda^* x_1^* = 0\\ \frac{\partial}{\partial x_2}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_1^* + 2\lambda^* x_2^* = 0\end{array}\right\} x_1^* = x_2^*\\ \frac{\partial}{\partial \lambda}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = {x_1^*}^2 + {x_2^*}^2 - l^2 = 0\end{array}\right\} x_1^* = x_2^* = \frac{l}{\sqrt{2}}, \lambda^* = \frac{1}{4} \checkmark$$

  - Case 2 $\quad g_1(\boldsymbol{x}^*) = 0$ (active), $g_2(\boldsymbol{x}^*) < 0$ (inactive), then $\mu_1^* \geq 0, \mu_2^* = 0$

$$\left.\begin{array}{l}\frac{\partial}{\partial x_1}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_2^* + 2\lambda^* x_1^* - \mu_1^* = 0\\ \frac{\partial}{\partial x_2}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_1^* + 2\lambda^* x_2^* = 0\\ \frac{\partial}{\partial \lambda}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = {x_1^*}^2 + {x_2^*}^2 - l^2 = 0\\ \left.-x_1^*\mu_1^* = 0 \right\} x_1^* = 0 \text{ or } \mu_1^* = 0\end{array}\right\} \begin{array}{l} x_1^* = 0, x_2^* = \pm l, \mu_1^* = \mp\frac{1}{2}l \times\\ \text{or}\\ \mu_1^* = 0, x_1^* = x_2^* = \frac{l}{\sqrt{2}}, \lambda^* = \frac{1}{4} \checkmark\end{array}$$

Optimality Conditions

Karush-Kuhn-Tucker (KKT) Conditions

- **Example**
  - Case 3 $\quad g_1(\boldsymbol{x}^*) < 0$ (inactive), $g_2(\boldsymbol{x}^*) = 0$ (active), then $\mu_1^* = 0, \mu_2^* \geq 0$

    Analogous to Case 2
  - Case 4 $\quad g_1(\boldsymbol{x}^*) = 0$ (active), $g_2(\boldsymbol{x}^*) = 0$ (active), then $\mu_1^* \geq 0, \mu_2^* \geq 0$

$$\left.\begin{array}{l}\frac{\partial}{\partial x_1}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_2^* + 2\lambda^* x_1^* - \mu_1^* = 0\\ \frac{\partial}{\partial x_2}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = -\frac{1}{2}x_1^* + 2\lambda^* x_2^* - \mu_2^* = 0\\ \frac{\partial}{\partial \lambda}L(\boldsymbol{x}^*,\lambda^*,\boldsymbol{\mu}^*) = {x_1^*}^2 + {x_2^*}^2 - l^2 = 0\\ \left.\begin{array}{l}-x_1^*\mu_1^* = 0\\ -x_2^*\mu_2^* = 0\end{array}\right\} \begin{array}{l}x_1^* = 0 \text{ or } x_2^* = 0 \times\\ \mu_1^* = 0 \text{ and } \mu_2^* = 0\end{array}\end{array}\right\} \begin{array}{l}\mu_1^* = 0, \mu_2^* = 0,\\ x_1^* = x_2^* = \frac{l}{\sqrt{2}}, \lambda^* = \frac{1}{4} \checkmark\end{array}$$

  - The maximum area is obtained for the legs $x_1^* = x_2^* = \frac{l}{\sqrt{2}}$ and has the value $\frac{1}{2}x_1^* x_2^* = \frac{l^2}{4}$

Concepts from Geometry

Hyperplanes and Half-Spaces

Tutorial

Definition 3.12 The set $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^T\boldsymbol{x} = b\}$ with $\boldsymbol{a} = (a_1 \ a_2 \ \cdots \ a_n)^T \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$, $b \in \mathbb{R}$ is called hyperplane.

- **Remarks**
  - The vector $\boldsymbol{a}$ is orthogonal to the hyperplane and therefore called normal
  - For $b = 0$ the hyperplane contains the origin and thus is a subspace of $\mathbb{R}^n$
  - For $n = 2$ the hyperplane becomes $a_1x_1 + a_2x_2 = b$ and thus describes a line in $\mathbb{R}^2$
  - For $n = 3$ the hyperplane becomes $a_1x_1 + a_2x_2 + a_3x_3 = b$ and thus describes a plane in $\mathbb{R}^3$
  - A hyperplane is a convex set

Definition 3.13 The set $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^T\boldsymbol{x} \leq b\}$ with $\boldsymbol{a} = (a_1 \ a_2 \ \cdots \ a_n)^T \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$, $b \in \mathbb{R}$ is called half-space.

- **Remarks**
  - Partly $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^T\boldsymbol{x} \geq b\}$ is called positive half-space and $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{a}^T\boldsymbol{x} \leq b\}$ negative half-space
  - A half-space is a convex set

Concepts from Geometry

Linear Varieties

Tutorial

Definition 3.14 The set $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ with $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ is called linear variety or flat.

- **Remarks**
  - A linear variety can also be written as $\boldsymbol{a}_i^T\boldsymbol{x} = b_i, i \in \{1, \dots, m\}$ ($\boldsymbol{a}_i^T$ rows of $\boldsymbol{A}$, b_i components of $\boldsymbol{b}$)
  - A linear variety is therefore the intersection of m hyperplanes
  - A linear variety is therefore a convex set (intersection of convex sets, cf. Slide 3-7, Property (3))
- **Examples**

$$(a_1 \quad a_2 \quad a_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = b \quad \Leftrightarrow \quad a_1x_1 + a_2x_2 + a_3x_3 = b$$

describes a plane in $\mathbb{R}^3$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \Leftrightarrow \quad \begin{matrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{matrix}$$

describes a line in $\mathbb{R}^{3*}$

* if $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are linearly independent

Concepts from Geometry

Polyhedra and Polytopes

Tutorial

Definition 3.15 The set $\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ with $\mathbf{A} \in \mathbb{R}^{m \times n}, \boldsymbol{b} \in \mathbb{R}^m$ is called polyhedron.

- **Remarks**
  - A polyhedron can also be written as $\boldsymbol{a}_i^T \boldsymbol{x} \leq b_i, i \in \{1, \dots, m\}$ ($\boldsymbol{a}_i^T$ rows of $\boldsymbol{A}$, b_i components of $\boldsymbol{b}$)
  - A polyhedron is therefore the intersection of m half-spaces
  - A polyhedron is therefore a convex set (intersection of convex sets, cf. Slide 3-7, Property (3))
  - The $0, 1, \dots, (k-1)$-dim. polyhedra forming the boundary of a k-dim. polyhedron are called faces
  - The faces of dimension 0, 1, $(k-2)$, and $(k-1)$ are called vertices, edges, ridges, and facets

Definition 3.16 A polytope is a bounded polyhedron (i.e. $\exists \alpha \in \mathbb{R}: ||\boldsymbol{y}|| \leq \alpha \; \forall \boldsymbol{y} \in \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$).

- **Remark**
  - Note that the definition of a polytope is not unique in the literature
  - Definition 3.16 is based on [BBM15, Section 3.1]

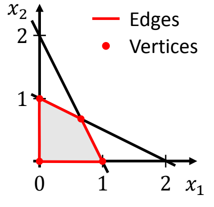
Concepts from Geometry

Polyhedra and Polytopes

Tutorial

Definition 3.17 The set $\left\{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x} = \sum_{i=1}^{V} \alpha_i \boldsymbol{V}_i, 0 \leq \alpha_i \leq 1, \sum_{i=1}^{V} \alpha_i = 1\right\}$ is called polytope where $\boldsymbol{V}_i \in \mathbb{R}^n$ are the vertices and V is the number of vertices.

- **Remark**
  - The representation according to Definition 3.15 is called half-space representation (H-representation)
  - The representation according to Definition 3.17 is called vertex representation (V-representation)
- **Example**



$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 2 & 1 \\ 0.5 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} \quad \Leftrightarrow \quad \begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ x_2 &\leq -2x_1 + 2 \\ x_2 &\leq -0.5x_1 + 1 \end{aligned}$$

The polyhedron is bounded and therefore a polytope

The polyhedron is unbounded if the first or second row are removed

Linear Programming

Linear Programming Problem

Problem 3.4 The linear programming problem is defined as

$$\min_{\boldsymbol{x}} \boldsymbol{c}^T\boldsymbol{x} \qquad \text{with } \boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n \qquad \text{linear cost function} \qquad (3.8)$$

$$\text{subject to} \begin{cases} \boldsymbol{A}_{\text{eq}}\boldsymbol{x} = \boldsymbol{b}_{\text{eq}} & \text{with } \boldsymbol{A}_{\text{eq}} \in \mathbb{R}^{m\times n}, \boldsymbol{b}_{\text{eq}} \in \mathbb{R}^m \quad \text{linear equality constraints} \quad (3.9) \\ \boldsymbol{A}_{\text{ieq}}\boldsymbol{x} \leq \boldsymbol{b}_{\text{ieq}} & \text{with } \boldsymbol{A}_{\text{ieq}} \in \mathbb{R}^{p\times n}, \boldsymbol{b}_{\text{ieq}} \in \mathbb{R}^p \quad \text{linear inequality constraints} \quad (3.10) \end{cases}$$

- **Remarks**
  - The linear cost function is a convex function. The linear equality constraints (linear variety) and the linear inequality constraints (polyhedron) are convex sets and thus the feasible set is a convex set.
  - The linear programming problem is therefore convex.
  - Several methods exist for solving the linear programming problem. The most important are the simplex method (exponential complexity) and Karmarkar‘s method (polynomial complexity)
  - The linear programming problem can be solved in MATLAB/Optimization Toolbox with `linprog`

Linear Programming

Characterization of the Solution

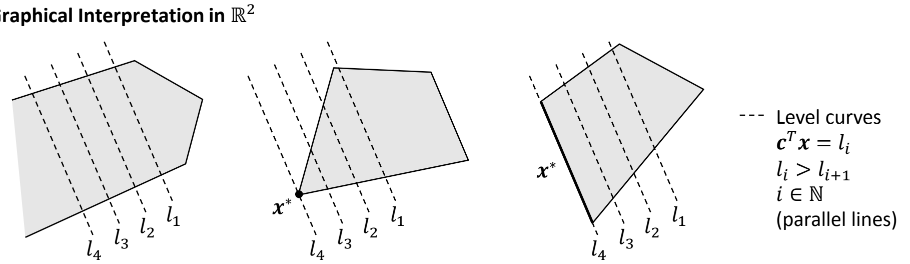
- **Cases**

  (1) The cost is unbounded, i.e. $\boldsymbol{c}^T\boldsymbol{x}^* = -\infty$

  (2) The cost is bounded, i.e. $\boldsymbol{c}^T\boldsymbol{x}^* > -\infty$, the minimizer $\boldsymbol{x}^*$ unique (vertex of the feasible set for $\mathbb{R}^2$)

  (3) The cost is bounded, i.e. $\boldsymbol{c}^T\boldsymbol{x}^* > -\infty$, the minimizer $\boldsymbol{x}^*$ not unique (edge of the feasible set for $\mathbb{R}^2$)

- **Graphical Interpretation in $\mathbb{R}^2$**



Case (1) Case (2) Case (3)

Quadratic Programming

Quadratic Programming Problem

Problem 3.5 The quadratic programming problem is defined as

$$\min_{x} \frac{1}{2} \boldsymbol{x}^T \boldsymbol{H} \boldsymbol{x} + \boldsymbol{f}^T \boldsymbol{x} \qquad \text{with } \boldsymbol{H} \in \mathbb{R}^{n \times n}, \boldsymbol{H} = \boldsymbol{H}^T \succcurlyeq \boldsymbol{0}, \boldsymbol{f} \in \mathbb{R}^n \quad \text{quadratic cost function} \qquad (3.11)$$

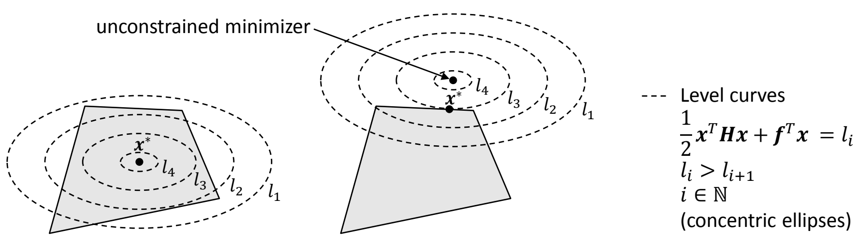
$$\text{subject to } \begin{cases} \boldsymbol{A}_{\text{eq}} \boldsymbol{x} = \boldsymbol{b}_{\text{eq}} & \text{with } \boldsymbol{A}_{\text{eq}} \in \mathbb{R}^{m \times n}, \boldsymbol{b}_{\text{eq}} \in \mathbb{R}^m \quad \text{linear equality constr.} \qquad (3.12) \\ \boldsymbol{A}_{\text{ieq}} \boldsymbol{x} \le \boldsymbol{b}_{\text{ieq}} & \text{with } \boldsymbol{A}_{\text{ieq}} \in \mathbb{R}^{p \times n}, \boldsymbol{b}_{\text{ieq}} \in \mathbb{R}^p \quad \text{linear inequality constr.} \qquad (3.13) \end{cases}$$

- **Remarks**
  - The quadratic cost function is a convex function for $\boldsymbol{H} \succcurlyeq \boldsymbol{0}$ and a strictly convex function for $\boldsymbol{H} \succ \boldsymbol{0}$. The linear equality constraints (linear variety) and the linear inequality constraints (polyhedron) are convex sets and thus the feasible set is a convex set.
  - The quadratic programming problem is therefore convex for $\boldsymbol{H} \succcurlyeq \boldsymbol{0}$ and strictly convex for $\boldsymbol{H} \succ \boldsymbol{0}$
  - The quadratic programming problem can be solved in MATLAB/Optimization Toolbox with `quadprog`

Quadratic Programming

Characterization of the Solution

- **Cases**
  (1) The cost is bounded and the minimizer $\boldsymbol{x}^*$ lies strictly inside the feasible set
  (2) The cost is bounded and the minimizer $\boldsymbol{x}^*$ lies on the boundary of the feasible set
- **Graphical Interpretation in $\mathbb{R}^2$**



Case (1) Case (2)

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Quadratic Programming

Solution based on the Active Set Method

- **Approach**
  - Consider that a feasible point $\boldsymbol{x}^{(i)}$ and related active inequality constraints $\boldsymbol{A}_{\text{ieq}}^{\text{a}}\boldsymbol{x}^{(i)} = \boldsymbol{b}_{\text{ieq}}^{\text{a}}$ are known
  - Find an improved point $\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ considering only $\boldsymbol{A}_{\text{eq}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0}$ and $\boldsymbol{A}_{\text{ieq}}^{\text{a}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0}$

    For the improved point $\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ the cost function becomes

$$
\begin{aligned}
f\left(\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}\right) &= \tfrac{1}{2}\left(\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}\right)^T \boldsymbol{H}\left(\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}\right) + \boldsymbol{f}^T\left(\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}\right) \\
&= f\left(\boldsymbol{x}^{(i)}\right) + \tfrac{1}{2}\Delta\boldsymbol{x}^{(i)T}\boldsymbol{H}\Delta\boldsymbol{x}^{(i)} + \underbrace{\left(\boldsymbol{f}^T + \boldsymbol{x}^{(i)T}\boldsymbol{H}\right)}_{\boldsymbol{f}^{(i)T}}\Delta\boldsymbol{x}^{(i)} \\
&= f\left(\boldsymbol{x}^{(i)}\right) + \tfrac{1}{2}\Delta\boldsymbol{x}^{(i)T}\boldsymbol{H}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{f}^{(i)T}\Delta\boldsymbol{x}^{(i)}
\end{aligned}
$$

    The improved point thus results from the optimization problem

$$
\begin{aligned}
&\min_{\Delta\boldsymbol{x}^{(i)}} \tfrac{1}{2}\Delta\boldsymbol{x}^{(i)T}\boldsymbol{H}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{f}^{(i)T}\Delta\boldsymbol{x}^{(i)} \\
&\text{subject to } \boldsymbol{A}_{\text{eq}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0},\ \boldsymbol{A}_{\text{ieq}}^{\text{a}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0}
\end{aligned}
\tag{3.14}
$$

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Quadratic Programming

Solution based on the Active Set Method

- **Approach**

  The Lagrangian to the optimization problem (3.14) obeys

$$
L\left(\Delta\boldsymbol{x}^{(i)}, \boldsymbol{\lambda}^{(i+1)}, \boldsymbol{\mu}^{(i+1)}\right) = \tfrac{1}{2}\Delta\boldsymbol{x}^{(i)T}\boldsymbol{H}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{f}^{(i)T}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{\lambda}^{(i+1)T}\boldsymbol{A}_{\text{eq}}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{\mu}^{(i+1)T}\boldsymbol{A}_{\text{ieq}}^{\text{a}}\Delta\boldsymbol{x}^{(i)}
$$

  The KKT conditions (only (1) and (2) relevant) to optimization problem (3.14) are then given by

$$
\begin{aligned}
\nabla_{\Delta\boldsymbol{x}^{(i)}} L\left(\Delta\boldsymbol{x}^{(i)}, \boldsymbol{\lambda}^{(i+1)}, \boldsymbol{\mu}^{(i+1)}\right) &= \boldsymbol{H}\Delta\boldsymbol{x}^{(i)} + \boldsymbol{f}^{(i)} + \boldsymbol{A}_{\text{eq}}^T\boldsymbol{\lambda}^{(i+1)} + \boldsymbol{A}_{\text{ieq}}^{\text{a}\,T}\boldsymbol{\mu}^{(i+1)} = \boldsymbol{0} \\
\nabla_{\boldsymbol{\lambda}^{(i+1)}} L\left(\Delta\boldsymbol{x}^{(i)}, \boldsymbol{\lambda}^{(i+1)}, \boldsymbol{\mu}^{(i+1)}\right) &= \boldsymbol{A}_{\text{eq}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0} \\
\nabla_{\boldsymbol{\mu}^{(i+1)}} L\left(\Delta\boldsymbol{x}^{(i)}, \boldsymbol{\lambda}^{(i+1)}, \boldsymbol{\mu}^{(i+1)}\right) &= \boldsymbol{A}_{\text{ieq}}^{\text{a}}\Delta\boldsymbol{x}^{(i)} = \boldsymbol{0}
\end{aligned}
$$

  which can be written as a system of linear equations (SLE)

$$
\begin{pmatrix} \boldsymbol{H} & \boldsymbol{A}_{\text{eq}}^T & \boldsymbol{A}_{\text{ieq}}^{\text{a}\,T} \\ \boldsymbol{A}_{\text{eq}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{A}_{\text{ieq}}^{\text{a}} & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}
\begin{pmatrix} \Delta\boldsymbol{x}^{(i)} \\ \boldsymbol{\lambda}^{(i+1)} \\ \boldsymbol{\mu}^{(i+1)} \end{pmatrix}
=
\begin{pmatrix} -\boldsymbol{f}^{(i)} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}
\tag{3.15}
$$

Quadratic Programming

Solution based on the Active Set Method

- **Approach**

  The solution of the optimization problem (3.14) finally follows by solving the SLE (3.15), e.g. based on the inverse (slow) or QR/LU decomposition (fast), cf. [Mac02, Section 3.3], [PLB12, Section 5.4.3]

  – Check if the improved point $\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ is a minimizer of the original quadratic programming problem (Problem 3.5) by evaluating the KKT conditions (1) to (5)

  – If not, then consider another improved point

- **Remarks**

  – Solving a quadratic programming problem with only equality constraints is obviously quite easy

  – The active set method is based on solving quadratic programming problems with equality constraints iteratively for different combinations of active inequality constraints (active sets)

  – This can be formalized as an algorithm

Quadratic Programming

Solution based on the Active Set Method

- **Algorithm**

  1. Determine initial feasible point $\boldsymbol{x}^{(0)}$ and active inequality constraints $\boldsymbol{A}_{\mathrm{ieq}}^{\mathrm{a}}\boldsymbol{x}^{(0)} = \boldsymbol{b}_{\mathrm{ieq}}^{\mathrm{a}}$ (active set)
  2. Set $i \coloneqq 0$
  3. Determine $\Delta\boldsymbol{x}^{(i)}$, $\boldsymbol{\lambda}^{(i+1)}$, and $\boldsymbol{\mu}^{(i+1)}$ by solving the SLE (3.15)
  4. Evaluate the KKT conditions (1) to (5) for Problem 3.5
     a. If $\Delta\boldsymbol{x}^{(i)} = \mathbf{0}$ and $\boldsymbol{\mu}^{(i+1)} \geq \mathbf{0}$, then stop since $\boldsymbol{x}^{(i)}$ is a feasible global minimizer for Problem 3.5
     b. If $\Delta\boldsymbol{x}^{(i)} = \mathbf{0}$ and at least one $\boldsymbol{\mu}^{(i+1)} < \mathbf{0}$, then set $\boldsymbol{x}^{(i+1)} \coloneqq \boldsymbol{x}^{(i)}$ and remove the active inequality constraint with the smallest $\boldsymbol{\mu}^{(i+1)}$ from the active set
     c. If $\Delta\boldsymbol{x}^{(i)} \neq \mathbf{0}$ and $\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ feasible, then set $\boldsymbol{x}^{(i+1)} \coloneqq \boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ and retain the active set
     d. If $\Delta\boldsymbol{x}^{(i)} \neq \mathbf{0}$ and $\boldsymbol{x}^{(i)} + \Delta\boldsymbol{x}^{(i)}$ infeasible, then find the largest $\alpha^{(i)} > 0$ for which $\boldsymbol{x}^{(i+1)} \coloneqq \boldsymbol{x}^{(i)} + \alpha^{(i)}\Delta\boldsymbol{x}^{(i)}$ is feasible and add resulting active inequality constraint to active set
  5. Set $i \coloneqq i + 1$ and go to 3.

Quadratic Programming

Solution based on the Active Set Method

- **Remarks**
  - An initial feasible point can be determined from a linear or quadratic programming problem, see [Mac02, Section 3.3] for details
  - The variables $\boldsymbol{x}^{(i+1)}$ resulting after each iteration i are feasible points of the original quadratic programming problem (Problem 3.5), allowing an early termination (relevant for MPC)
  - A warm start, i.e. an initialization of the iteration with a point which is known to be close to the minimizer (initial guess) for reducing the number of iterations, is straightforward (relevant for MPC)
  - The active set method has exponential complexity

Technische Universität Kaiserslautern

Quadratic Programming

Solution based on the Interior Point Method

- **Approach**
  - Transform the constrained optimization problem to an unconstrained optimization problem, i.e.

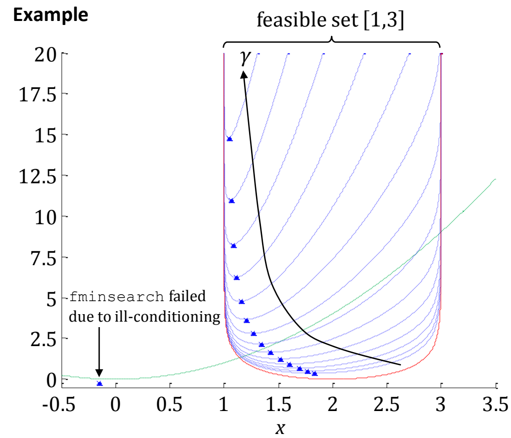
$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{H}\boldsymbol{x} + \boldsymbol{f}^T\boldsymbol{x} \quad \text{subject to } \boldsymbol{A}_{\text{ieq}}\boldsymbol{x} \leq \boldsymbol{b}_{\text{ieq}} \qquad \rightarrow \qquad \min_{\boldsymbol{x}} \gamma\left(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{H}\boldsymbol{x} + \boldsymbol{f}^T\boldsymbol{x}\right) \underbrace{- \textstyle\sum_{j=1}^{p} \ln\left(b_{\text{ieq},j} - \boldsymbol{a}_{\text{ieq},j}^T\boldsymbol{x}\right)}_{\text{barrier function}}$$

  - The barrier function is finite in the interior but infinite on the boundary of the feasible set
  - Let $\boldsymbol{x}_\gamma^*$ be the minimizer of the unconstrained optimization problem for some $\gamma > 0$ and $\boldsymbol{x}^*$ be the minimizer of the constrained optimization problem. It can be shown that $\boldsymbol{x}_\gamma^* \rightarrow \boldsymbol{x}^*$ as $\gamma \rightarrow \infty$. However, the unconstrained optimization problem becomes ill-conditioned as $\gamma \rightarrow \infty$.
  - The interior point method is based on solving the unconstrained optimization problem iteratively for an increasing γ until $\boldsymbol{x}_\gamma^*$ does not change significantly anymore
  - The path followed by $\boldsymbol{x}_\gamma^*$ is denoted as central path

Quadratic Programming

Solution based on the Interior Point Method

- **Example**



$\min_{x\in[1,3]} x^2$

$\min_{x} \gamma x^2 - \ln(-1+x) - \ln(3-x)$

x^2

$-\ln(-1+x) - \ln(3-x)$

$\gamma x^2 - \ln(-1+x) - \ln(3-x)$
(for $\gamma =$ `logspace(-1,2,20)`)

x_γ^* (obtained with `fminsearch`)

Quadratic Programming

Solution based on the Interior Point Method

- **Remarks**
  - The equality constraint $\boldsymbol{A}_{\text{eq}}\boldsymbol{x} = \boldsymbol{b}_{\text{eq}}$ can be regarded in the interior point method by reformulation into two inequality constraints $\boldsymbol{A}_{\text{eq}}\boldsymbol{x} \leq \boldsymbol{b}_{\text{eq}}$ and $-\boldsymbol{A}_{\text{eq}}\boldsymbol{x} \leq -\boldsymbol{b}_{\text{eq}}$
  - The minimizers $\boldsymbol{x}_\gamma^*$ of the unconstrained optimization problem are feasible points of the constrained optimization problem, allowing an early termination of the iterations (relevant for MPC)
  - The interior point method requires modifications to address ill-conditioning. The minimizers $\boldsymbol{x}_\gamma^*$ are usually no feasible points under the these modifications, not allowing an early termination (MPC)
  - A warm start, i.e. an initialization of the iteration with a point which is known to be close to the minimizer (initial guess) for reducing the number of iterations, is usually difficult (relevant for MPC)
  - The interior point method has polynomial complexity

Optimization Software

Remarks on Optimization Software

- **Overviews**
  - plato.asu.edu/guide.html
  - yalmip.github.io/allsolvers/
  - neos-guide.org/optimization-tree
  - https://www.coin-or.org/
- **Modeling Languages and Solvers**
  - YALMIP (yalmip.github.io/)
  - CVX (cvxr.com/cvx/)
  - CVXGEN (cvxgen.com)
  - FORCES (forces.ethz.ch)
  - µAO-MPC (ifatwww.et.uni-magdeburg.de/syst/muAO-MPC/)