Lecture: Nonlinear optimization without constraints

1. General Nonlinear optimization problems
2. Convex sets
3. Convex functions
4. Convex optimization problems

General nonlinear optimization problems

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F} \end{aligned} \tag{1}$$

- $\mathcal{F} \subset \mathbf{R}^n$ is called the *feasible region*.
- $f : \mathcal{F} \to \mathbf{R}$ is called the objective function.

Definition 1.

(i) *The point* $\mathbf{x} \in \mathbf{R}^n$ *is called* feasible *if* $\mathbf{x} \in \mathcal{F}$.

(ii) *The point* $\hat{\mathbf{x}} \in \mathcal{F}$ *is a* local optimal solution *to (2) if there exists a* $\delta > 0$ *such that* $f(\hat{\mathbf{x}}) \le f(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{F}$ *such that* $|\mathbf{x} - \hat{\mathbf{x}}| \le \delta$.

(iii) *The point* $\hat{\mathbf{x}} \in \mathcal{F}$ *is a* global optimal solution *to (2) if* $f(\hat{\mathbf{x}}) \le f(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{F}$.

Some examples

Linear optimization: $f(\mathbf{x}) = \mathbf{c}^{\mathsf{T}}\mathbf{x}$ and $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$, i.e.

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

Quadratic optimization: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ and $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$, i.e.

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

Nonlinear optimization: $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0,\ i = 1, \ldots, m\}$, i.e.

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_i(\mathbf{x}) \leq \mathbf{0},\ i = 1, \ldots, m. \end{aligned}$$

- Linear optimization problems are well posed since they can be solved using the simplex method.
- Quadratic optimization problems are well posed if $\mathbf{H}$ is positive semi-definite.
- When are general nonlinear optimization problems well posed ?
  - One class of nonlinear optimization problems that is well posed is convex optimization problems.

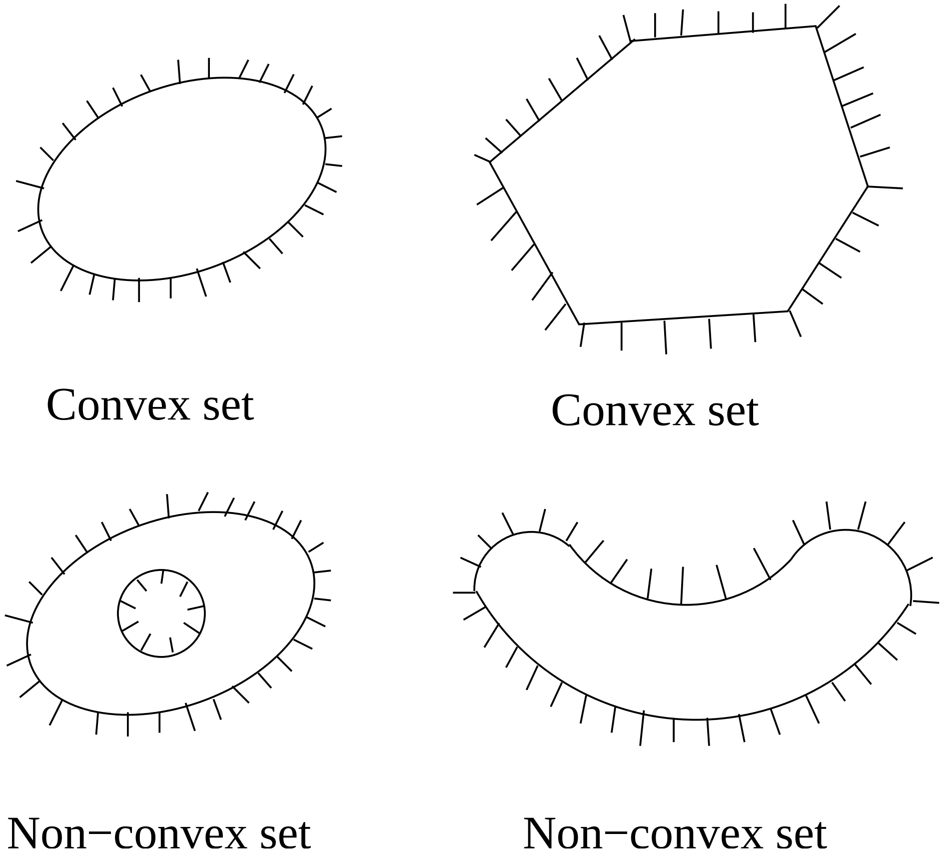
Comment 1. *We will see that linear optimization problems and quadratic optimization problems with positive semi-definite* $\mathbf{H}$ *are special cases of convex optimization problems.*

Convex sets

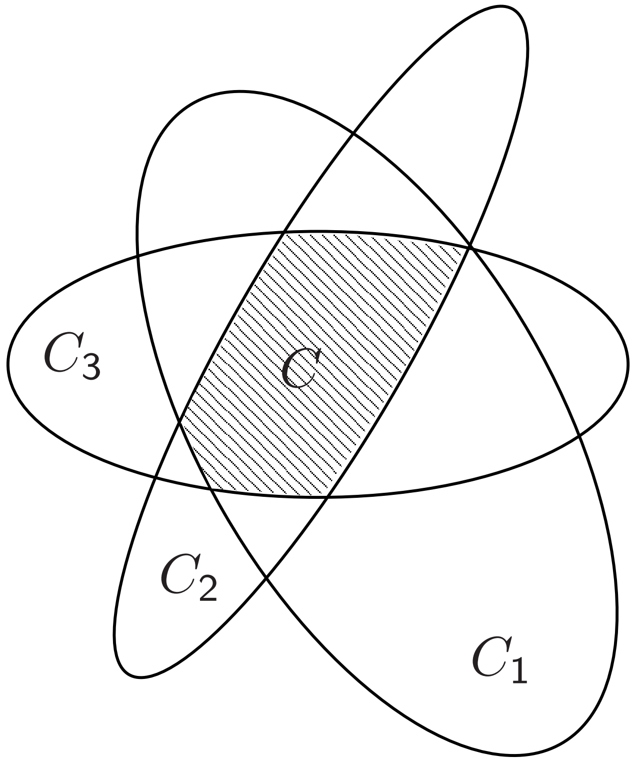
Definition 2. *A set $C \subset \mathbf{R}^n$ is convex if*

$$(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in C$$

for every choice of $\mathbf{x}_1 \in C$, $\mathbf{x}_2 \in C$ and $t \in (0,1)$.



Theorem 1. *Let $C_1, \dots C_n \subset \mathbf{R}^n$ be convex sets. Then the intersection $C = \cap_{k=1}^n C_k$ of the sets is also a convex set.*



Exempel 1. *Let*

$$\mathbf{A} = \begin{bmatrix} \bar{\mathbf{a}}_1^{\mathsf{T}} \\ \vdots \\ \bar{\mathbf{a}}_m^{\mathsf{T}} \end{bmatrix} \in \mathbf{R}^{m\times n} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbf{R}^m$$

The set $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ *is convex. This follows since*

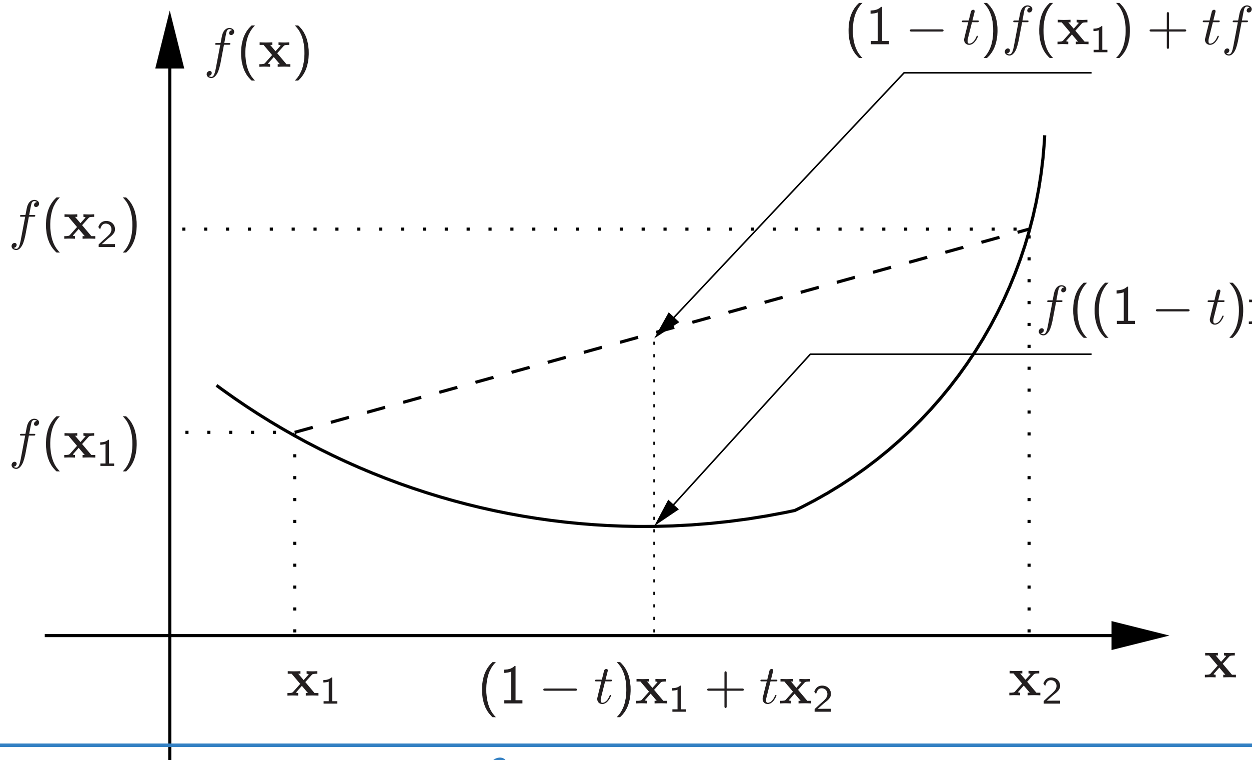
$$\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \bar{\mathbf{a}}_j^{\mathsf{T}}\mathbf{x} \geq b_j,\ j = 1,\ldots,m\} = \cap_{j=1}^{n}\{\mathbf{x} \in \mathbf{R}^n : \bar{\mathbf{a}}_j^{\mathsf{T}}\mathbf{x} \geq b_j\}$$

The sets $\{\mathbf{x} \in \mathbf{R}^n : \bar{\mathbf{a}}_j^{\mathsf{T}}\mathbf{x} \geq b_j\}$ *are half-planes and thereby obviously convex.*

Convex functions

Definition 3. *Let $C \subset \mathbf{R}^n$ be a convex set. A real valued function $f : C \to \mathbf{R}$ is convex if for every choice of $\mathbf{x}_1 \in C$, $\mathbf{x}_2 \in C$ and $t \in (0,1)$*

$$f((1-t)\mathbf{x}_1 + t\mathbf{x}_2) \leq (1-t)f(\mathbf{x}_1) + tf(\mathbf{x}_2)$$



Theorem 2. *If $f_1, \ldots, f_m$ are convex functions on C and $\gamma_1, \ldots, \gamma_m$ are non-negative real constants, then the function*

$$f(\mathbf{x}) = \gamma_1 f_1(\mathbf{x}) + \ldots + \gamma_m f_m(\mathbf{x})$$

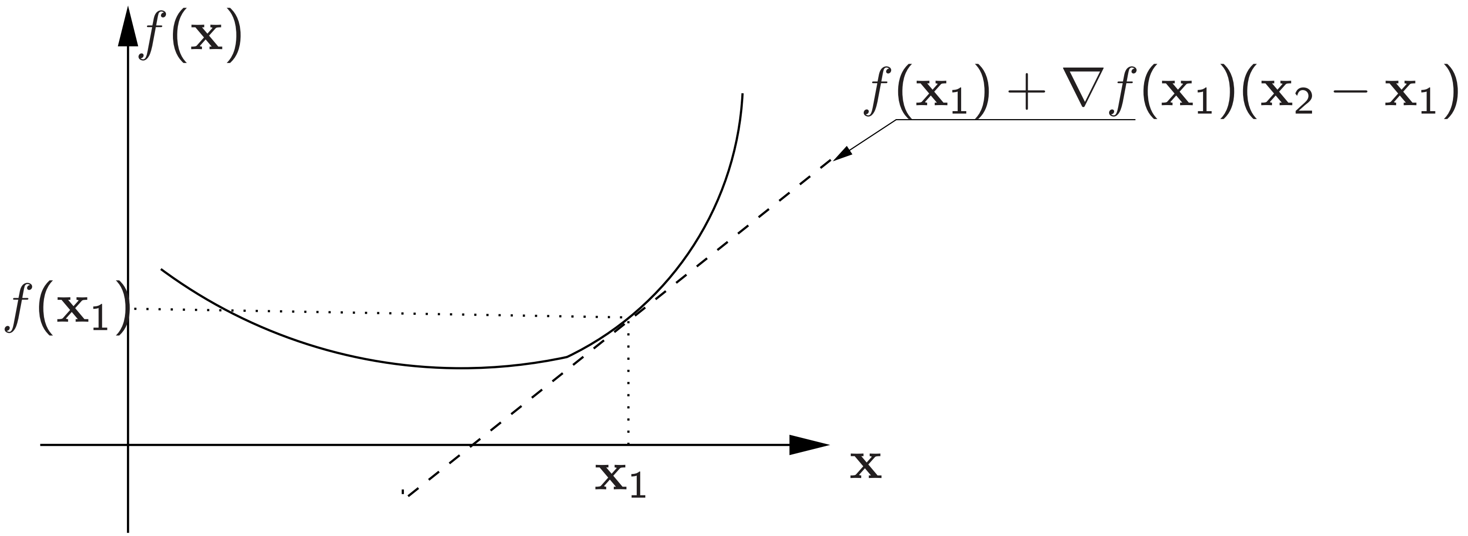
is convex on C.

Proof: *Let $\mathbf{x}_1, \mathbf{x}_2 \in C$ and $t \in (0, 1)$. According to Definition 3 we have that*

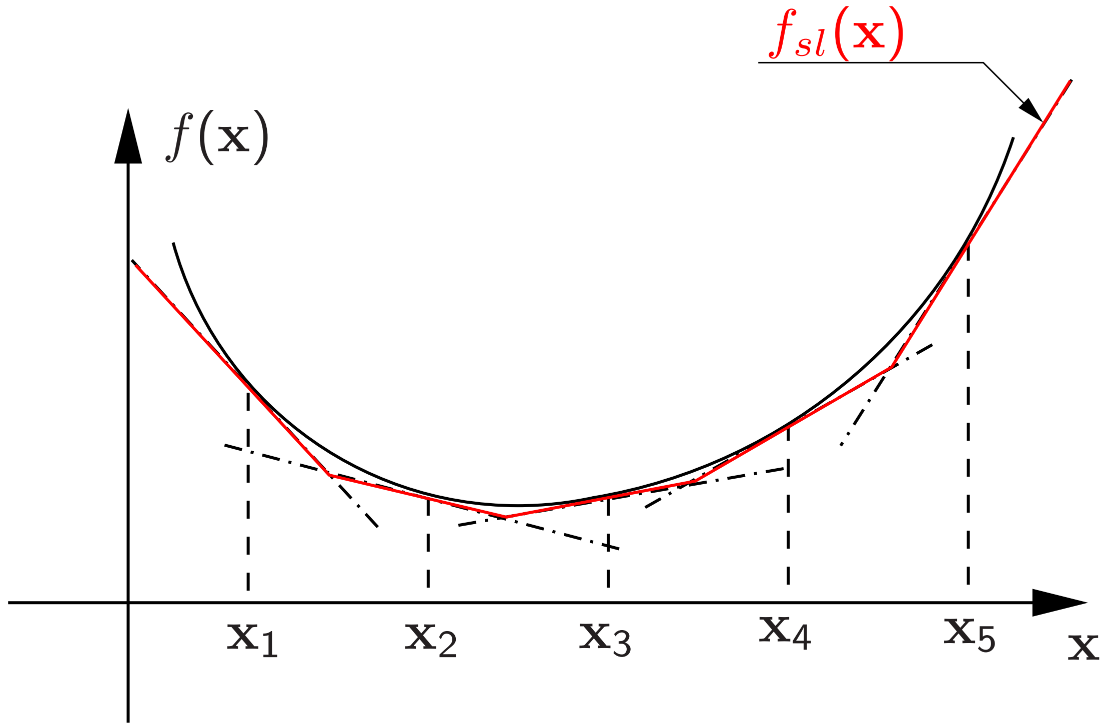
$$\begin{aligned} f((1-t)\mathbf{x}_1 + t\mathbf{x}_2) &= \sum_{k=1}^{m} \gamma_k f_k((1-t)\mathbf{x}_1 + t\mathbf{x}_2) \\ &\leq \sum_{k=1}^{m} \gamma_k [(1-t) f_k(\mathbf{x}_1) + t f_k(\mathbf{x}_2)] \\ &= (1-t) \sum_{k=1}^{m} \gamma_k f_k(\mathbf{x}_1) + t \sum_{k=1}^{m} \gamma_k f_k(\mathbf{x}_2) \\ &= (1-t) f(\mathbf{x}_1) + t f(\mathbf{x}_2), \quad \text{which shows that } f \text{ is convex.} \end{aligned}$$

Theorem 3. *Let $C \subset \mathbf{R}^n$ be a convex set. A continuously differentiable function $f : C \to \mathbf{R}$ is convex if and only if, for every $\mathbf{x}_1 \in C$ and $\mathbf{x}_2 \in C$*

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1)$$



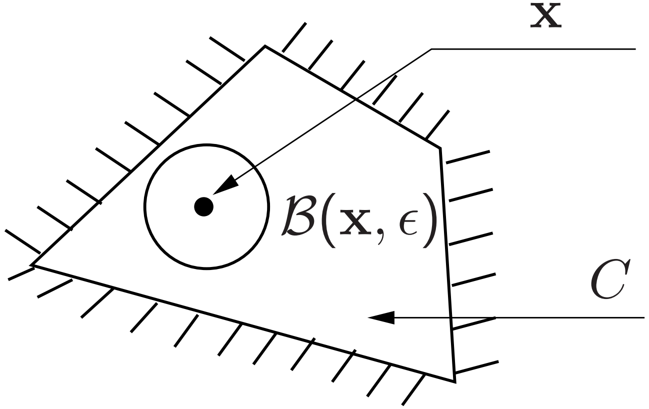
A direct consequence of the Theorem is that a convex function can be estimated from below with a piecewise linear convex function.



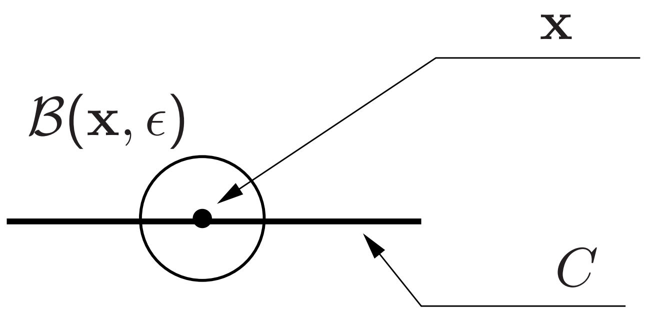
$$f_{sl}(\mathbf{x}) = \max_{\mathbf{x}}\{f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k)\}$$

Theorem 4. *Assume that $C \subset \mathbf{R}^n$ is a given convex set with at least one interior point. A two times continuously differentiable function $f : C \to \mathbf{R}$ is then convex, if and only if, the Hessian $\nabla^2 f(\mathbf{x})$ is positive semi-definite for every $\mathbf{x} \in C$.*

A point $\mathbf{x} \in C$ is an interior point if there exists a ball with center in $\mathbf{x}$, $\mathcal{B}(\mathbf{x}, \epsilon) = \{\mathbf{y} \in \mathbf{R}^n : |\mathbf{x} - \mathbf{y}| < \epsilon\}$, such that $\mathcal{B}(\mathbf{x}, \epsilon) \subset C$.



Convex set with interior points



Convex set without an interior point

Note that it in general is important that the function f is defined on a (convex) subset $C \subset \mathbf{R}^n$.

Exempel 2. *Let $C = (0, \infty)$, which is a convex set. The Function $f(x) = -\ln(x)$ is convex on C since $\nabla^2 f(\mathbf{x}) = \frac{1}{x^2}$ is positive for all $x \in C = (0, \infty)$. The function is not definied for $x \leq 0$ and is therefore not convex on $C = (-\infty, \infty)$.*

Exempel 3. *The function $f(x) = x^3$ is convex if it is defined on $C = [0, \infty)$, but it is not convex if it is defined on $C = (-\infty, \infty)$.*

Sometimes, there are no interior point, and then the following theorem is useful:

Theorem 5. *Assume that $C \subset \mathbf{R}^n$ is a given convex set and that $f : C \to \mathbf{R}$ is a two times continuously differentiable function on C. Then, f is convex on C if, and only if, the following inequality is satisfied for every choice of $\hat{\mathbf{x}} \in C$ and $\mathbf{x} \in C$:*

$$(\mathbf{x} - \hat{\mathbf{x}})^{\mathsf{T}} \nabla^2 f(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) \geq 0$$

Exempel 4. *Assume that*

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$$
$$C = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}.$$

If $\mathbf{x}, \hat{\mathbf{x}} \in C$, *it follows that* $\mathbf{x} - \hat{\mathbf{x}} \in \mathcal{N}(\mathbf{A})$ *since* $\mathbf{A}(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$. *If* $\mathbf{Z}$ *is a matrix spanning the nullspace, i.e.* $\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 & \dots & \mathbf{z}_k \end{bmatrix}$ *where the columns form a basis for* $\mathcal{N}(\mathbf{A})$ *it follows that* $\mathbf{x} - \hat{\mathbf{x}} = \mathbf{Z}\mathbf{v}$ *for some* $\mathbf{v} \in \mathbf{R}^k$. *Therefore, it holds that*

$$(\mathbf{x} - \hat{\mathbf{x}})^\mathsf{T}\nabla^2 f(\hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{v}^\mathsf{T}\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}\mathbf{v}$$

which is positive for an arbitrary $\mathbf{v} \in \mathbf{R}^k$ *if, and only if, if the reduced Hessian* $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}$ *is positive semi-definite. According to the previous theorem* f *is thus convex on* C *if, and only if,* $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}$ *is positive semi-definite.*

Convex optimization problems

The optimization problem

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F} \end{aligned} \tag{2}$$

is called convex if

- the feasible region $\mathcal{F} \subset \mathbf{R}^n$ is convex.
- the objective function $f : \mathcal{F} \to \mathbf{R}$ is convex.

Examples of convex optimization problems

Linear optimization:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

We have already shown that $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ is convex and the objective function $f(\mathbf{x}) = \mathbf{c}^{\mathsf{T}}\mathbf{x}$ is convex since it satisfies the inequality in Definition 3 with equality.

Quadratic optimization:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

This problem is convex if $\mathbf{H}$ is positive semi-definite since the objective function $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{H}\mathbf{x} + \mathbf{c}^{\mathsf{T}}\mathbf{x}$ according to Theorem 4 then is convex.

Quadratic optimization under linear equality constraints:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

It is easy to show that the feasible region $\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$ is convex. This problem is therefore, according to Example 4 convex if the reduced Hessian $\mathbf{Z}^\mathsf{T}\mathbf{H}\mathbf{Z}$ is positive semi-definite.

Nonlinear optimization:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_i(\mathbf{x}) \leq \mathbf{0}, \;\; i = 1, \ldots, m. \end{aligned}$$

This problem is convex if $g_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, n$ are convex functions and f is a real valued convex function on

$$\mathcal{F} = \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0, \;\; i = 1, \ldots, m\}.$$

For proving this, we need to show that $\mathcal{F}$ is convex. Since,

$$\mathcal{F} = \cap_{i=1}^{m} \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0\}$$

it is enough to show that $\mathcal{F}_i = \{\mathbf{x} \in \mathbf{R}^n : g_i(\mathbf{x}) \leq 0\}$ is convex.

If $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{F}_i$, it holds that $g_i(\mathbf{x}_1) \leq 0$ and $g_i(\mathbf{x}_2) \leq 0$. Since g_i is convex, it follows that

$$g_i((1-t)\mathbf{x}_1 + t\mathbf{x}_2) \leq (1-t)g_i(\mathbf{x}_1) + tg_i(\mathbf{x}_2) \leq 0$$

for $t \in (0,1)$. This shows that $(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in \mathcal{F}_i$ and $\mathcal{F}_i$ is therefore a convex set.

The following theorem demonstrated a very nice property of convex optimization problems

Theorem 6. *If $\hat{\mathbf{x}} \in \mathcal{F}$ is a local optimal solution to the convex optimization problem (2), then it is also a global optimal solution.*

Proof: *Assume that $\hat{\mathbf{x}} \in \mathcal{F}$ is not a global optimal solution. Then there exists a $\mathbf{x} \in \mathcal{F}$ such that $f(\mathbf{x}) < f(\hat{\mathbf{x}})$. Since $\mathcal{F}$ is convex it holds that $\mathbf{x}(t) = (1-t)\hat{\mathbf{x}} + t\mathbf{x} \in \mathcal{F}$ for $t \in (0,1)$. Since f is convex we have*

$$f(\mathbf{x}(t)) \leq (1-t)f(\hat{\mathbf{x}}) + tf(\mathbf{x}) = f(\hat{\mathbf{x}}) + t(f(\mathbf{x}) - f(\hat{\mathbf{x}})) < f(\hat{\mathbf{x}})$$

But since $\mathbf{x}(t) = \hat{\mathbf{x}} + t(\mathbf{x} - \hat{\mathbf{x}})$, there are feasible points arbitrarily close to $\hat{\mathbf{x}}$ (i.e. for small t) where the objective function value is smaller than $f(\mathbf{x})$. This contradicts that $\hat{\mathbf{x}}$ is a local optimal solution. Therefore, it must hold that $f(\hat{\mathbf{x}}) < f(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{F}$.

The following corollary is very useful:

Theorem 7. *Assume that the function f is continuously differentiable and convex on* $\mathbf{R}^n$. *Then* $\hat{\mathbf{x}} \in \mathbf{R}^n$ *is a global minimum if, and only if,* $\nabla f(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$.

Proof: *Since* $\hat{\mathbf{x}} \in \mathbf{R}^n$ *is a global minimum it is also a local minimum, hence it follows from the first order necessary optimality conditions that* $\nabla f(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$.

Conversely, assume that $\nabla f(\hat{\mathbf{x}}) = \mathbf{0}^\mathsf{T}$. *Then according to Theorem 3 it holds that*

$$f(\mathbf{x}) \geq f(\hat{\mathbf{x}}) + \underbrace{\nabla f(\hat{\mathbf{x}})}_{=0}(\mathbf{x} - \hat{\mathbf{x}}) = f(\hat{\mathbf{x}})$$

for all $\mathbf{x} \in \mathbf{R}^n$. *This shows that* $\hat{\mathbf{x}}$ *is a global minimum to* f.

Exempel 5. *Consider the optimization problem*

$$\text{minimize } f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{H}\mathbf{x} + \mathbf{c}^\mathsf{T}\mathbf{x}$$

where $\mathbf{H}$ *is positive semi-definite. Then* f *is a convex function and the minimizing point satisfies*

$$\nabla f(\hat{\mathbf{x}})^\mathsf{T} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{c} = \mathbf{0}$$

This is the same optimality condition that was derived earlier.

Comments

Theorem 6 shows that every local optimal solution to a convex optimization problem also is a global optimal solution. Note that the optimal solution is not necessarily unique, and that there is not always a (finite) optimal solution to a convex optimization problem.

Exempel 6.

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1 \geq 1 \end{aligned}$$

is unbounded from below and hence lacks an optimal solution.

The convex optimization problem below lacks a unique optimal solution

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 1 \end{aligned}$$