

NUMERICAL ANALYSIS

Numerical methods for solving unrestricted optimisation problems

Prof. Dr. Andreas Fischer

University of Technology Dresden

Department of Mathematics

Date: 23.05.2019

Author: Ilina Tuneska

Index number: 4737573

Summer semester 2019

Lecture 5

THE ROLE OF CONVEX FUNCTIONS IN THE OPTIMISATION

In this lecture we take a close look at the role of convex functions in the optimization problems.

First we review some key definitions from the previous Proseminars. Afterwards, we take a look at some examples and lastly, we define theorems which we are going to use as a base for explaining the correlation between the convexity of functions and the optimisation problems.

We begin this topic with revision of some of the key terms defined in the previous lectures of this Proseminar.

Definition 5.1 (Convex set)

A set $X \subseteq \mathbb{R}^n$ is a **convex set**, if for all $x_1, x_2 \in X$ and all $\lambda \in (0, 1)$ it holds:

$$\lambda x_1 + (1 - \lambda)x_2 \in X \quad (5.1)$$

Definition 5.2 (Convex, strictly and uniformly convex function:)

Let $X \subseteq \mathbb{R}^n$ be a convex set. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is:

1. **convex** on X , if for all $x_1, x_2 \in X$ and all $\lambda \in (0, 1)$ holds:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (5.2)$$

2. **strictly convex** on X , if for all $x_1, x_2 \in X$ and all $\lambda \in (0, 1)$ holds

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (5.3)$$

3. **uniformly convex** on X , if there exists $\mu > 0$ with:

$$f(\lambda x_1 + (1 - \lambda)x_2) + \mu \lambda(1 - \lambda) \|x_1 - x_2\|^2 \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (5.4)$$

Now we begin with defining our first theorem:

Theorem 5.3

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Then, the level sets

$$\begin{aligned} \mathcal{L}(x_0) &:= \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\} \\ \mathcal{L}(x_0) &:= \{x \in \mathbb{R}^n | f(x) < f(x_0)\} \end{aligned}$$

are for all $f(x_0) \in \mathbb{R}$, convex.

Proof. • The objective of this proof is to show that the set $\mathcal{L}(x_0)$ is convex, that is, for all $x_1, x_2 \in \mathcal{L}(x_0)$, $\lambda \in (0, 1) : \lambda x_1 + (1 - \lambda)x_2 \in \mathcal{L}(x_0)$.

Let $x_1, x_2 \in X$, $\lambda \in (0, 1)$. Looking at the preconditions of this theorem, we see that f is convex, which according to inequality (5.2) means:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

Then for $x := \lambda x_1 + (1 - \lambda)x_2$, it holds:

$$\begin{aligned}
 f(x) &= f(\lambda x_1 + (1 - \lambda)x_2) \\
 &\stackrel{(5.2)}{\leq} \lambda f(x_1) + (1 - \lambda)f(x_2) \\
 &\stackrel{(*)}{\leq} \lambda f(x_0) + (1 - \lambda)f(x_0) \\
 &= f(x_0)
 \end{aligned}$$

(2) it follows from the convexity of the function f

(*) since $x_1, x_2 \in \mathcal{L}(x_0)$, it follows that $f(x_1), f(x_2) \leq f(x_0)$

- The proof that the level set $\mathcal{L}(x_0) := \{x \in \mathbb{R}^n | f(x) < f(x_0)\}$ is for all $f(x_0) \in \mathbb{R}$, convex, follows analogously with the use of the proof of the first part with (" $<$ ") instead of (" \leq ").

Next, we go back to the optimisation problems and explain why the class of (strict, uniform) convex functions play an important role in the optimisation.

Theorem 5.4

Let $f : X \rightarrow \mathbb{R}$ be continuously differentiable and $X \subseteq \mathbb{R}^n$ be convex with $x \in X$. We consider the restricted optimization problem

$$\min(f(x)) \tag{5.5}$$

Then the following statements hold:

1. If the function f is **convex** on X , then the solution set of (5.5) is **convex (or empty)**
2. If the function f is **strictly convex** on X , then the solution set of (5.5) has **at most one solution**
3. If the function f is **uniformly convex** on X , $X \neq \emptyset$ and X is a closed set, then (5.5) has **exactly one solution**
 - (a) If the function f is **uniformly convex** on X but X is not empty, then (5.5) has no solution
 - (b) If the function f is **uniformly convex** on X but X is not a closed set, then (5.5) has no solution.

Proof. (1) Objective: to show that for $x_1, x_2 \in \min f(x)$ and $\lambda \in (0, 1)$, it holds: $\lambda x_1 + (1 - \lambda)x_2 \in \min(f(x))$.

Let $x_1, x_2 \in \min(f(x))$ (which means that x_1 and x_2 are two solutions of (5.5), that is, $f(x_1) = f(x_2) = \min(f(x))$). From the convexity of X , it holds: $\lambda x_1 + (1 - \lambda)x_2 \in X$, for $\lambda \in (0, 1)$. Therefore, from the convexity of f it follows:

$$\begin{aligned}
 f(\lambda x_1 + (1 - \lambda)x_2) &\stackrel{(5.2)}{\leq} \lambda f(x_1) + (1 - \lambda)f(x_2) \\
 &= \lambda f(x_1) + f(x_2) - \lambda f(x_2) \quad \text{because } f(x_1) = f(x_2) \\
 &= \lambda f(x_1) + f(x_1) - \lambda f(x_1) \\
 &= f(x_1) \\
 &= \min(f(x))
 \end{aligned}$$

from where it follows that $f(\lambda x_1 + (1 - \lambda)x_2) \leq \min f(x)$ (minimum of the function f and therefore $\lambda x_1 + (1 - \lambda)x_2$ is a minimum of (1): $\lambda x_1 + (1 - \lambda)x_2 \in \min(f(x))$).

- (2) We prove this next part of the Theorem through contradiction.

Let the optimization problem (5.5) have two different solutions: x_1 and x_2 with $x_1 \neq x_2$ and $f(x_1) = f(x_2) = \min(f(x))$. Since X is a convex set, for $\lambda \in (0, 1)$ is $\lambda x_1 + (1 - \lambda)x_2 \in X$. From the strict convexity of the function f it holds that

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &\stackrel{(5.3)}{<} \lambda f(x_1) + (1 - \lambda)f(x_2) \\ &\stackrel{(*)}{=} \lambda f(x_1) + f(x_1) - \lambda f(x_1) \\ &= f(x_1) = \min(f(x)). \end{aligned}$$

where $(*)$ follows from the fact that $f(x_1) = f(x_2)$.

Therefore, we found an element of the set X , which is strictly smaller than all elements in the set $\min(f(x))$, which is in contradiction to the assumed minimality of $f(x_1)$.

- (3) Let $x_0 \in X$ be arbitrarily selected. From Lemma 3.9 it follows that the level set $\mathcal{L}(x_0)$ is compact. So the set $X \cap \mathcal{L}(x_0)$ is also compact and, by definition, non empty. With the extreme value theorem, the continuous function f has a global minimum on the compact set $X \cap \mathcal{L}(x_0)$, which from obvious reasons must also be a minimum of (1).

We now take a look just at the third part of the Theorem and ask ourselves, what would happen if the function f was convex (and therefore strictly convex) and:

- (a) the set X was empty:

If X was empty, then we would not have any points in $X \cap \mathcal{L}(x_0)$, which would trivially mean that the function f would not have a minimum (as seen in example 5.5 (1))

- (b) the set X was not closed.

If X was not closed, then the function f would again trivially, not have a minimum (as seen in example 5.5 (2))

Now we take a look at some counterexamples regarding theorem 5.4 (3):

Example 5.5

- (1) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ with $f(x) = e^x$.

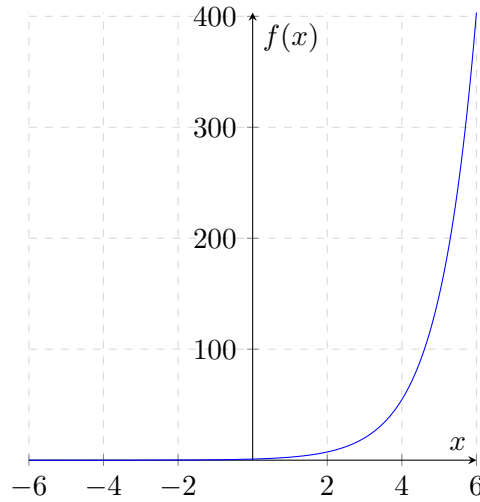
Proof. (The function $f(x) = e^x$ is a convex function:)

Since \mathbb{R} is per definition a convex set, $f(x) = e^x$ is a continuously differentiable function as composition of continuously differentiable functions, and $\nabla f(x) = \nabla e^x = e^x$ is a monotone function, it follows that $f(x) = e^x$ is a convex function.

Then, since the convex set \mathbb{R} is not bounded, the function $f(x) = e^x$ is decreasing on the domain \mathbb{R} for $x \rightarrow 0$ and never reaches a point of minimum, which means that the function does not have a minimum on the domain \mathbb{R} .

- (2) Let $f : X \rightarrow \mathbb{R}$ with $X = \emptyset$.

Then, since the convex set X is empty, it does not contain any values $f(x)$, which trivially means that the function does not have a minimum.



We continue with defining an inequality for uniformly convex functions, which will be useful later on.

Lemma 5.6

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuously differentiable, $x_0 \in \mathbb{R}^n$, the level set $\mathcal{L}(x_0)$ convex, f uniformly convex on $\mathcal{L}(x_0)$ and $x \in \mathbb{R}$, according to Theorem 2, 3), unique global minimum of f . Then there exists $\mu > 0$ with

$$\mu \|x - x^*\|^2 \leq f(x) - f(x^*) \quad \forall x \in \mathcal{L}(x_0)$$

.

Proof. Using:

- The inequality (5.3) from the begining of this lecture and
- the fact that x^* is, as a result of being a global Minimum of f , a stationary point of f and therefore $\nabla f(x^*) = 0$,

we conclude the needed inequality.

Finally, we analyze in which way the existence of stationary points of a function are connected to the existence of global minimum of that function.

Theorem 5.7

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuously differentiable and convex function, $x \in \mathbb{R}^n$ a stationary point of f . Then x^* is a global minimum of f on \mathbb{R}^n .

Proof. First, we remind ourselves that a point $x^* \in X$ is called a stationary point of a continously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, if $\nabla f(x^*) = 0$. From Proposition 3.5. a) follows:

$$f(x) - f(x^*) \leq \nabla f(x^*)^t (x - x^*) = 0$$

and thus

$$f(x) \geq f(x^*), \forall x \in R$$

It follows that x^* is a global minimum of f .