

MINIMIZERS OF OPTIMIZATION PROBLEMS

MARKUS GRASMAIR

In the following we will always consider a minimization problem of the form

$$\min_{x \in \Omega} f(x),$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is an extended real valued function (the *cost function* or *objective function*) and $\Omega \subset \mathbb{R}^n$ is some set (the *feasible set*). In the case where $\Omega = \mathbb{R}^n$, we speak of *unconstrained optimization*, else of *constrained optimization*. In general, the solution of constrained optimization problems is *much* harder than the solution of unconstrained ones. However, the question whether a solution exists at all can be answered for both types of problems with (more or less) the same methods.

NOTIONS OF MINIMIZERS

First we have to clarify what we mean by a solution of an optimization problem.

Definition 1 (Global minimizer). A point $x^* \in \mathbb{R}^n$ is called a *global minimizer* of the optimization problem $\min_{x \in \Omega} f(x)$, if $x^* \in \Omega$ and

$$f(x^*) \leq f(x)$$

for all $x \in \Omega$.

In other words, a global minimizer is *feasible*, and its value is not larger than the value of any other feasible point.

Note:

- Global minimizers need not exist, as one can see (for instance) in the following examples:
 - Minimize the function $f(x) = 1/x^2$ for $x \in \mathbb{R} \setminus \{0\}$.
 - Minimize the function $f(x) = e^{-x^2}$ for $x \in \mathbb{R}$.
 - Minimize the function $f(x) = x$ for $x > 0$.
- Global minimizers need not be unique. One example is the function $f(x) = (x^2 - 1)^2$ with two global minimizers $x^* = \pm 1$. A more extreme example is the function $f(x) = 0$, where every point $x \in \mathbb{R}$ is a global minimizer.

One problem of global minimizers is that they are incredibly hard to recognize in general. In order to verify that a point x^* is a global minimizer, one would have to compare $f(x^*)$ with every other value $f(x)$, no matter how large the distance between x and x^* is. In actual applications, however, one usually may only obtain the value of f (and, possibly, some of its derivatives) at a small number of selected points. With only this information available, only in very special cases is it possible to prove that a given point x^* is really a global minimizer.

As an alternative, we therefore consider local minimizers:

Definition 2 (Local minimizer). A point $x^* \in \mathbb{R}^n$ is called a *local minimizer* of the optimization problem $\min_{x \in \Omega} f(x)$, if there exists a neighbourhood N of x^* such that x^* is a global minimizer of the problem $\min_{x \in \Omega \cap N} f(x)$.

In other words, x^* is a local minimizer of $\min_{x \in \Omega} f(x)$, if $x^* \in \Omega$, and there exists $\varepsilon > 0$ such that

$$f(x^*) \leq f(x)$$

whenever $x \in \Omega$ satisfies $\|x - x^*\| \leq \varepsilon$.

Slightly strengthening this notation, we obtain:

Definition 3 (Strict local minimizer). A point x^* is called a *strict local minimizer* of $\min_{x \in \Omega} f(x)$, if $x^* \in \Omega$, and there exists $\varepsilon > 0$ such that

$$f(x^*) < f(x)$$

whenever $x \in \Omega \setminus \{x^*\}$ satisfies $\|x - x^*\| \leq \varepsilon$.

That is, we replace the inequality \leq by the strict inequality $<$ in the definition of the local minimizer.

In addition, it makes sometimes sense to strengthen this notion further:

Definition 4 (Isolated local minimizer). A point $x^* \in \Omega$ is called an *isolated local minimizer* of $\min_{x \in \Omega} f(x)$, if $x^* \in \Omega$ and there exists a neighbourhood N of x^* such that x^* is the only local minimizer of the optimization problem $\min_{x \in \Omega \cap N} f(x)$.

Note that every isolated local minimizer is a strict local minimizer, but the converse does not necessarily hold. As an example consider the (rather pathological) function

$$f(x) = \begin{cases} 2x^2 + x^2 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

This function has a strict local minimizer at $x = 0$ (which is at the same time the unique global minimizer of f), but there exists a sequence of (isolated!) local minimizers converging to 0. Thus the minimizer at 0 is not isolated.

EXISTENCE OF MINIMIZERS

For the following definition, recall that the lower limit of a sequence of real numbers z_k is defined as

$$\liminf_{k \rightarrow \infty} z_k := \lim_{k \rightarrow \infty} \inf_{\ell \geq k} z_\ell.$$

This is equivalent to defining $\liminf_{k \rightarrow \infty} z_k$ as the smallest possible limit of convergent subsequences of z_k . Note that, in contrast to the limit, every sequence has a lower limit (the sequence $(\inf_{\ell \geq k} z_\ell)_{k \in \mathbb{N}}$ is increasing, and therefore its limit exists).

Definition 5 (Lower semi-continuity). A function $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is called *lower semi-continuous*, if for every $x \in \Omega$ and every sequence $x_k \in \Omega$ converging to x we have

$$f(x) \leq \liminf_{k \rightarrow \infty} f(x_k).$$

This means that, whenever we have a sequence x_k converging to x , the sequence of values $f(x_k)$ cannot have a limit that is smaller than $f(x)$. For instance:

- Every continuous function is lower semi-continuous.
- The function $f: \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 & \text{if } x > 0, \\ x^2 - 1 & \text{if } x \leq 0, \end{cases}$$

is lower semi-continuous.

- The function $f: \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0, \\ -1 & \text{if } x = 0, \end{cases}$$

is lower semi-continuous.

- The function $f: \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} +1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0, \end{cases}$$

is *not* lower semi-continuous.

- If $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I$, is *any* family of continuous functions, then the function

$$f(x) := \sup_{i \in I} f_i(x)$$

is lower semi-continuous. (Note that we do not require that the family is finite!)

This last property of lower semi-continuous functions turns out to be very important in certain branches of optimization, where so called min-max (or inf-sup) problems appear naturally, that is, problems of the form

$$\inf_{x \in \Omega} \sup_{y \in W} g(x, y).$$

Remark 6. An alternative (equivalent) definition of lower semi-continuity is the following:

A function $f: \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous, if the *lower level set*

$$\Omega_\alpha := \{x \in \Omega : f(x) \leq \alpha\}$$

is relatively closed in Ω for every $\alpha \in \mathbb{R}$. In other words: Whenever $\alpha \in \mathbb{R}$ and $x_k \in \Omega_\alpha$ is a sequence that converges to some $x \in \Omega$, we have that $x \in \Omega_\alpha$. Because this definition does not rely directly on sequences but rather on the notion of closedness, it can, in some situations, be less cumbersome to handle.

Definition 7 (Minimizing sequence). Assume that $\Omega \neq \emptyset$ and let

$$f^* := \inf_{x \in \Omega} f(x).$$

A minimizing sequence for the optimization problem $\min_{x \in \Omega} f(x)$ is a sequence x_k in Ω satisfying

$$\lim_{k \rightarrow \infty} f(x_k) = f^*.$$

That is, a minimizing sequence is a sequence the function values of which converge to the minimal (infimal) value of f . Note that this does not say anything about the convergence of the sequence x_k itself. For example, the sequence $x_k := (2k+1)\pi + 1/k$ is a (obviously diverging) minimizing sequence for the function $f(x) = \cos(x)$.

Theorem 8 (Existence of a solution, part I). *Assume that $\Omega \subset \mathbb{R}^n$ is non-empty, closed, and bounded, and that $f: \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous. Then the optimization problem $\min_{x \in \Omega} f(x)$ admits at least one global minimizer x^* .*

Proof. Let $x_k \in \Omega$ be a minimizing sequence for the optimization problem. Because Ω is closed and bounded, it follows that the sequence x_k admits a sub-sequence, say $x'_k \in \Omega$, converging to some point $x^* \in \Omega$ (Heine–Borel Theorem). Thus the definitions of x_k , x'_k , and x^* , and the lower semi-continuity of f imply that

$$\inf_{x \in \Omega} f(x) = f^* = \lim_{k \rightarrow \infty} f(x_k) = \lim_{k \rightarrow \infty} f(x'_k) = \liminf_{k \rightarrow \infty} f(x'_k) \geq f(x^*),$$

which shows that $f(x^*) \leq f(x)$ for every $x \in \Omega$. In other words, x^* is a global minimizer of f in Ω . \square

Note the similarity between this theorem and the extreme value theorem, which states that every continuous function on a closed and bounded set admits both its maximum and minimum. Because we are only interested in finding minima, we can relax the assumption of continuity and replace it by the weaker assumption of lower semi-continuity. Doing so, however, we lose (possibly) the existence of a maximizer.

In order to obtain an existence result that is also applicable to unconstrained optimization problems (that is, the situation where $\Omega = \mathbb{R}^n$, which is obviously an unbounded set), we have to introduce another definition:

Definition 9 (Coercivity). A function $f: \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ is called coercive, if every sequence $x_k \in \Omega$ with $\|x_k\| \rightarrow \infty$ satisfies $f(x_k) \rightarrow \infty$.

Theorem 10 (Existence of a solution, part II). *Assume that $\Omega \subset \mathbb{R}^n$ is non-empty and closed, and that $f: \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and coercive. Then the optimization problem $\min_{x \in \Omega} f(x)$ admits at least one global minimizer x^* .*

Proof. If $f(x) = +\infty$ for every $x \in \Omega$, it follows that every point $x \in \Omega$ is a global minimizer. Thus we may assume without loss of generality that there exists some $\hat{x} \in \Omega$ with $f(\hat{x}) < +\infty$. In this case, it follows that $f^* := \inf_{x \in \Omega} f(x) \leq f(\hat{x}) < +\infty$.

Now let $x_k \in \Omega$ be a minimizing sequence for the optimization problem. Then the sequence $f(x_k)$ does not converge to ∞ , and therefore the coercivity of f implies that the sequence x_k is bounded. Thus it admits a sub-sequence $x'_k \in \Omega$ converging to some point $x^* \in \mathbb{R}^n$. Because Ω is closed, it follows that, actually, $x^* \in \Omega$. Thus we have (as above) that

$$\inf_{x \in \Omega} f(x) = f^* = \lim_{k \rightarrow \infty} f(x_k) = \lim_{k \rightarrow \infty} f(x'_k) = \liminf_{k \rightarrow \infty} f(x'_k) \geq f(x^*),$$

which shows that x^* is a global minimizer of f in Ω . \square

DEPARTMENT OF MATHEMATICS, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY, 7491 TRONDHEIM, NORWAY

E-mail address: markus.grasmair@math.ntnu.no