MINIMISERS OF OPTIMISATION PROBLEMS

MARKUS GRASMAIR

In this note we will discuss the existence of solutions of minimisation problems of the form

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

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is a real valued function (the *cost function* or *objective function*) and $\Omega \subset \mathbb{R}^n$ is some set (the *feasible set*).

1. Notions of Minimisers

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

Definition 1 (Global minimiser)**.** A point $x^* \in \Omega$ is called a *global minimiser* (or *global minimum*) of the optimisation problem $\min_{x\in\Omega} f(x)$, if

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

for all $x \in \Omega$.

The point x^* is *strict global minimiser*, if $f(x^*) < f(x)$ for all $x \in \Omega$, $x \neq x^*$.

Note:

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

One problem of global minimisers is that they are incredibly hard to recognise in general. In order to verify that a point x^* is a global minimiser, 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 minimiser.

As an alternative, we therefore consider local minimisers:

Definition 2 (Local minimiser)**.** A point $x^* \in \Omega$ is called a *local minimiser* of the optimisation problem $\min_{x\in\Omega} f(x)$, if there exists $\varepsilon > 0$ such that $f(x^*) \le f(x)$ whenever $x \in \Omega$ satisfies $\|x - x^*\| \le \varepsilon$.

Slightly strengthening this notation, we obtain:

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

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

whenever $x \in \Omega$, $x \neq x^*$ satisfies $\|x - x^*\| \le \varepsilon$.

Date: January 2017.

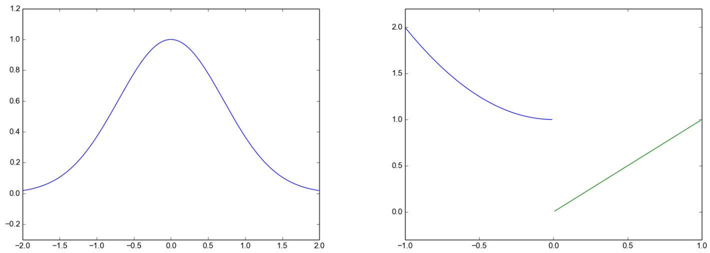

FIGURE 1. *Left:* The function $f(x) = e^{-x^2}$ obviously does not attain its minimum, because of the drop-off of the function values near infinity. *Right:* The existence of a minimiser of the function f defined by $f(x) = x^2 + 1$ for $x < 0$ and $f(x) = x$ for $x > 0$ depends on its value at 0. If $f(0) \leq 0$, the point $x = 0$ is the unique global and local minimum. If, however, $f(0) > 0$, the function does not attain its minimum.

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

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

Definition 4 (Isolated local minimiser)**.** A point $x^* \in \Omega$ is called an *isolated local minimiser* of the problem $\min_{x \in \Omega} f(x)$, if there exists $\varepsilon > 0$ such that x^* is the only local minimiser of f in Ω an ε-ball around x^*. That is, if $y^* \neq x^*$ is another local minimiser of f in Ω, then $\|x^* - y^*\| > \varepsilon$.

Note that every isolated local minimiser is a strict local minimiser, 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 minimiser at $x = 0$ (which is at the same time the unique global minimiser of f), but there exists a sequence of (isolated!) local minimisers converging to 0. Thus the minimiser at 0 is not isolated. See also Figure 2.

2. Existence of Minimisers

We have seen above that an optimisation problem need not necessarily have a solution: As seen above, the function $f(x) = e^{-x^2}$ does not attain a minimum, and nor does the function f defined by $f(x) = x$ if $x > 0$ and $f(x) = x^2 + 1$ if $x \leq 0$. It turns out, however, that these two examples are in a sense the typical counter-examples to the existence of minimisers: In the case of the function e^{-x^2}, the problem is that the function to be minimised becomes smaller as the argument increases. In the case of the other counter-example, the problem is a discontinuity at the point where we would "naturally" expect the minimum. By excluding these two possibilities, that is, by requiring the function f to be continuous and to grow as its argument tends to infinity, we can indeed guarantee the existence of a minimiser. Because discontinuous functions can be important in some applications, it makes sense to try to obtain results for this type of functions as well, though.

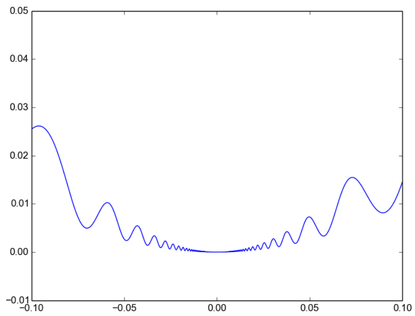

FIGURE 2. A close-up view of the function $f(x) = 2x^2 + x^2 \sin(1/x)$ near 0. The point $x = 0$ is the unique global minimum, but is also an accumulation point of isolated local minima.

For the following definition, recall that the lower limit of a sequence of real numbers $(z_k)_{k\in\mathbb{N}}$ is defined as

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

This is equivalent to defining $\liminf_{k\to\infty} z_k$ as the smallest possible limit of convergent subsequences of z_k. Equivalently, $\liminf_{k\to\infty}$ is the infimum of all accumulation points of the sequence $(z_k)_{k\in\mathbb{N}}$ in the extended real line $\mathbb{R}\cup\{\pm\infty\}$.

Moreover, we recall some properties of the lower limit:

- The lower limit of a sequence $(z_k)_{k\in\mathbb{N}}$ always exists (in $\mathbb{R}\cup\{\pm\infty\}$).
- If the sequence $(z_k)_{k\in\mathbb{N}}$ converges, then $\liminf_{k\to\infty} z_k = \lim_{k\to\infty} z_k$.
- If $(y_k)_{k\in\mathbb{N}}$, $(z_k)_{k\in\mathbb{N}}$ are two sequences, then
$$\liminf_{k\to\infty}(y_k + z_k) \geq \liminf_{k\to\infty} y_k + \liminf_{k\to\infty} z_k.$$
Here we define $+\infty + (-\infty) := -\infty$.
- If $(y_k)_{k\in\mathbb{N}}$ is a sequence and $\lambda \geq 0$, then
$$\liminf_{k\to\infty} \lambda y_k = \lambda \liminf_{k\to\infty} y_k.$$
Here $\lambda(\pm\infty) = \pm\infty$ for $\lambda > 0$, and $0\cdot(\pm\infty) := 0$.
- If $(y_k^{(i)})_{k\in\mathbb{N}}$, $i \in I$, is a family of sequences (with an arbitrary index set I), then
$$\liminf_{k\to\infty} \sup_{i\in I} y_k^{(i)} \geq \sup_{i\in I} \liminf_{k\to\infty} y_k^{(i)}.$$

Definition 5 (Lower semi-continuity)**.** A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called *lower semi-continuous*, if for every $x \in \mathbb{R}^n$ and every sequence $(x_k)_{k\in\mathbb{N}}$ converging to x we have

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

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 $f\colon \mathbb{R}\to\mathbb{R}$ defined by
$$f(x) = \begin{cases} x & \text{if } x > 0,\\ x^2+1 & \text{if } x \leq 0,\end{cases}$$
is *not* lower semi-continuous.

- The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by
$$f(x) = \begin{cases} x & \text{if } x \geq 0, \\ x^2 + 1 & \text{if } x < 0, \end{cases}$$
is lower semi-continuous.
- The function $f\colon \mathbb{R} \to \mathbb{R}$ 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 $f\colon \mathbb{R} \to \mathbb{R}$ 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 $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i \in I$, 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 optimisation, 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 $f\colon \mathbb{R}^n \to \mathbb{R}$ is lower semi-continuous, if the *lower level set*
$$\Omega_\alpha := \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$
is closed for every $\alpha \in \mathbb{R}$.

In other words: Whenever $\alpha \in \mathbb{R}$ and $(x_k)_{k\in\mathbb{N}}$ is a sequence that converges to some $x \in \mathbb{R}^n$ and $x_k \in \Omega_\alpha$ for all k (that is, $f(x_k) \leq \alpha$), we have that $x \in \Omega_\alpha$ (that is, $f(x) \leq \alpha$). 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.

A further characterisation of lower semi-continuity can be obtained by considering convergence of points on the graph of f. One can show that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is lower semi-continuous, if and only if for every $x \in \mathbb{R}^n$ and every sequence $(x_k)_{k\in\mathbb{N}}$ such that the values $f(x_k)$ converge in the extended real line we have that
$$f(x) \leq \lim_{k\to\infty} f(x_k).$$
In that sense, the difference between continuity and lower semi-continuity is that the equal-sign in the definition of continuity is replaced by a lower-than-or-equal sign.

Proposition 7. *Assume that* $f\colon \mathbb{R}^n \to \mathbb{R}$ *is lower semi-continuous and that* $\Omega \subset \mathbb{R}^n$ *is compact. Then the optimisation problem*
$$\min_{x \in \Omega} f(x)$$

admits a solution.

Proof. Denote
$$f^* := \inf_{x\in\Omega} f(x).$$
Then there exists a sequence $(x_k)_{k\in\mathbb{N}} \subset \Omega$ such that
$$\lim_{k\to\infty} f(x_k) = f^*.$$
Because of the compactness of Ω, the sequence $(x_k)_{k\in\mathbb{N}}$ has a subsequence $(x_{k'})_{k'}$ converging to some $x^* \in \Omega$. From the lower semi-continuity of f we now obtain that
$$f(x^*) \le \liminf_{k'\to\infty} f(x_{k'}) = f^* = \inf_{x\in\Omega} f(x).$$
This shows that x^* actually is a global minimum of f in Ω. □

This result covers the case of constrained optimisation problems where the feasible set Ω is compact. However, it also can be used in order to prove existence of solutions in more general settings as long as f is lower semi-continuous. The basic idea there is to show that all possible minimisers can only ever be found in some compact set, which makes it possible to reduce the problem to one with compact constraints. One possibility for this approach is shown in the following.

Definition 8. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called coercive, if we have for every sequence $(x_k)_{k\in\mathbb{N}}$ with $\|x_k\| \to \infty$ that $f(x_k) \to \infty$.

Theorem 9. *Assume that the function $f\colon \mathbb{R}^n \to \mathbb{R}$ is lower semi-continuous and coercive. Then the optimisation problem $\min_{x\in\mathbb{R}^n} f(x)$ admits at least one global minimiser x^*.*

Proof. By assumption, the function f is coercive, which implies that there exists some $R > 0$ such that
$$f(x) > f(0) \qquad \text{whenever } \|x\| > R.$$
As a consequence, if the problem $\min_{x\in\mathbb{R}^n} f(x)$ admits a solution, this solution necessarily has to be contained in the ball $B_R := \{x \in \mathbb{R}^n : \|x\| \le R\}$, as every point $x \notin B_R$ is minorised by 0. This, however, implies that the problem $\min_{x\in\mathbb{R}^n} f(x)$ is equivalent to the constrained optimisation problem $\min_{x\in B_R} f(x)$. From Proposition 7 we now obtain the existence of a solution of the latter problem, which, because of the equivalence of the two problems, immediately implies the existence of the original, unconstrained one. □

Department of Mathematics, Norwegian University of Science and Technology, 7491 Trondheim, Norway

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