Beyond Convex? Global Optimization Is Feasible Only for Convex Objective Functions: A Theorem

R. Baker Kearfott[1] and Vladik Kreinovich[2]

[1]Department of Mathematical Sciences
University of Lafayette at Louisiana
Lafayette, LA 70504-1010, rbk@louisiana.edu

[2]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, vladik@cs.utep.edu

Abstract

It is known that there are feasible algorithms for minimizing convex functions, and that for general functions, global minimization is a difficult (NP-hard) problem. It is reasonable to ask whether there exists a class of functions that is larger than the class of all convex functions for which we can still solve the corresponding minimization problems feasibly. In this paper, we prove, in essence, that no such more general class exists. In other words, we prove that global optimization is always feasible only for convex objective functions.

1 Introduction

It is well known that in general, global optimization is a difficult-to-solve problem. In particular, it is known that even the problem of minimizing an objective function $f(x_1, \ldots, x_n)$ on a box ("hyper-rectangle") $[\underline{x}_1, \overline{x}_1] \times \ldots \times [\underline{x}_n, \overline{x}_n]$ – a problem of interest in *interval computations* [1, 2, 3] – is NP-hard; see, e.g., [4, 6]. Crudely speaking, this means as the number of variables n increases, in the worst case, the computation time required to solve the corresponding optimization problem grows exponentially with n – and so, for large n, it is not possible to have an algorithm that correctly solves all possible global optimization problems.

It is also well known that there exist feasible algorithms for minimizing *convex* objective functions $f(x_1, \ldots, x_n)$; see, e.g., [6]. A natural question is: can we extend these algorithms to a larger class of objective functions? In other words, can we extend the class of all convex functions to a larger class for which minimization is still feasible?

Of course, if we take this question literally, the answer is clearly "yes": we can extend the class of all convex functions by adding one or more objective functions for which we already know the solutions to the corresponding minimization problems.

This answer is not very interesting from a practical viewpoint. Indeed, the class of all convex functions is not simply a collection of unrelated functions, it is *closed* under several useful operations such as addition, multiplication by a positive constant, substitution of linear combinations of variables instead of the original variables, etc. It is therefore reasonable to ask: is there a class of functions that is similarly closed and for which global minimization is feasible?

In this paper, we show that convex functions are the only ones for which this is possible – once we have a single non-convex function in our closed class, the corresponding global minimization problem becomes NP-hard.

2 Definitions and the Main Result

In this paper, we consider continuous functions $f(x_1, \ldots, x_n)$ from R^n to R for different n.

Definition. *We say that a class of functions F is* closed *if it satisfies the following four conditions:*

- *F contains all linear functions;*
- *F is* closed under addition, *i.e., if $f \in F$ and $g \in F$, then $f + g \in F$;*
- *F is* closed under multiplication by a positive constant, *i.e., if $f \in F$ and $c > 0$, then $c \cdot f \in F$;*
- *F is* closed under linear substitution *if whenever $f(x_1, \ldots, x_k) \in F$ and c_{ik} are real numbers, we have*

$$f(c_{10} + c_{11} \cdot x_1 + \ldots + c_{1n} \cdot x_n, \ldots, c_{k0} + c_{k1} \cdot x_1 + \ldots + c_{kn} \cdot x_n) \in F.$$

It is easy to see that the class of all linear functions is closed, and that the class of all convex functions is closed.

By a *minimization problem*, we mean the following problem: given a function $f \in F$ and a box B, find the minimal value of the function f on the box B.

Theorem 1. *If a closed class F contains at least one non-convex function and at least one non-linear convex function, then for this class, the problem of finding the minimum of a given function $f \in F$ on a given box is NP-hard.*

Since for convex functions, minimization is feasible, this theorem can be reformulated as follows: for a closed class F that contains at least one non-linear convex function, the following two conditions are equivalent to each other:

- all functions from the class F are convex;
- for the lass F, global minimization is feasible.

The same result holds if we consider a ε-*minimization problem,* i.e., if we fix some real number $\varepsilon > 0$ (called *accuracy*), and, instead of looking for the exact minimum m of a function f, we look for the value $\widetilde{m}$ that is ε-close to m, i.e., for which $|\widetilde{m} - m| \le \varepsilon$.

Theorem 2. *Let $\varepsilon > 0$ and let F be a closed class that contains at least one non-convex function and at least one non-linear convex function. Then, the problem of finding an ε-approximation to the minimum of a given function $f \in F$ on a given box is NP-hard.*

3 Proofs

Let us first prove Theorem 1. Let F be a closed class that contains a non-convex function $f_0(x_1, \ldots, x_k)$ and a non-linear convex function $f_1(x_1, \ldots, x_m)$.

1°. Let us first prove that F contains a non-convex function of one variable.

Indeed, by definition, a function f of k variables is convex if

$$f(\alpha \cdot a + (1-\alpha) \cdot b) \le \alpha \cdot f(a) + (1-\alpha) \cdot f(b) \tag{1}$$

for all $a, b \in R^k$ and for all $\alpha \in (0,1)$. Thus, non-convexity of f_0 means that there exist points $a = (a_1, \ldots, a_k)$ and $b = (b_1, \ldots, b_k)$, and a number $\alpha \in (0,1)$ for which

$$f_0(\alpha \cdot a + (1-\alpha) \cdot b) > \alpha \cdot f_0(a) + (1-\alpha) \cdot f_0(b). \tag{2}$$

Since the class F is closed under linear substitution, the function

$$f_2(x_1) \stackrel{\text{def}}{=} f_0(a_1 + x_1 \cdot (b_1 - a_1), \ldots, a_k + x_1 \cdot (b_k - a_k)). \tag{3}$$

also belongs to the class F. In terms of $f_2(x_1)$, the inequality (2) takes the form $f_2(\alpha) > \alpha \cdot f_2(0) + (1-\alpha) \cdot f_2(1)$. Thus, the function $f_2(x_1)$ is non-convex. The statement is proven.

2°. Let us now prove that F contains a function $f_3(x_1)$ of one variable for which $f_3(0) = f_3(1) = 0$ and $f_3(\alpha) > 0$ for some $\alpha \in (0,1)$.

We will construct this function f_3 from the above function f_2, as $f_3(x) = f_2(x) - f_2(0) - x \cdot (f_2(1) - f_2(0))$. Since F is closed, it contains all linear functions, and it is closed under addition; thus, $f_3 \in F$. It is easy to check that $f_3(0) = f_3(1) = 0$, and that (3) implies $f_3(\alpha) > 0$.

3°. Let us now prove that F contains a function $f_4(x_1)$ of one variable for which $f_4(0) = f_4(1) = 0$ and $f_4(x) > 0$ for *all* $x \in (0, 1)$.

We will construct this function f_4 from the above function f_3. We know that $f_3(\alpha) > 0$ and that $f(0) = 0$. Let α^- denote the supremum of all the values $x < \alpha$ for which $f_3(x) \le 0$. By definition of α^-, we have $f_3(x) > 0$ for all $x \in (\alpha^-, \alpha]$. The supremum α^- is a limit point of non-positive values $f_3(x)$, $x \le \alpha^-$, and it is also a limit point of positive values $f_3(x)$, $x > \alpha^-$. Thus, $f_3(\alpha^-) = 0$.

Similarly, if we take, as α^+, the infimum of all the values $x > \alpha$ for which $f_3(x) \le 0$. Then, $f_3(\alpha^+) = 0$ and $f_3(x) > 0$ for all $x \in [\alpha, \alpha^+)$. So, $f_3(\alpha^-) = f_3(\alpha^+) = 0$ and $f_3(x) > 0$ for all $x \in (\alpha^-, \alpha^+)$. Thus, the function $f_4(x) \stackrel{\text{def}}{=} f_3(\alpha^- + x \cdot (\alpha^+ - \alpha^-))$ belongs to the class F and has the desired property.

4°. Let us now prove that F contains a non-linear convex function $f_5(x)$ of one variable.

This can be done similarly to Part 1° of this proof. Indeed, one can easily see that a function f of m variables is linear if

$$f(\alpha \cdot a + (1 - \alpha) \cdot b) = \alpha \cdot f(a) + (1 - \alpha) \cdot f(b) \tag{4}$$

for all $a, b \in R^m$ and for all $\alpha \in (0, 1)$. Thus, non-linearity of f_1 means that there exist points $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$, and a number $\alpha \in (0, 1)$ for which

$$f_1(\alpha \cdot a + (1 - \alpha) \cdot b) \ne \alpha \cdot f_1(a) + (1 - \alpha) \cdot f_1(b). \tag{5}$$

Since the function f_1 is convex, we conclude that

$$f_1(\alpha \cdot a + (1 - \alpha) \cdot b) < \alpha \cdot f_1(a) + (1 - \alpha) \cdot f_1(b). \tag{6}$$

Since the class F is closed under linear substitution, the function

$$f_5(x_1) \stackrel{\text{def}}{=} f(a_1 + x_1 \cdot (b_1 - a_1), \ldots, a_k + x_1 \cdot (b_k - a_k)) \tag{7}$$

also belongs to the class F. In terms of $f_5(x_1)$, the inequality (6) takes the form $f_5(\alpha) < \alpha \cdot f_5(0) + (1 - \alpha) \cdot f_5(1)$. Thus, the function $f_5(x_1)$ is non-linear. The class of all convex functions is closed under linear substitution, so the function $f_5(x)$ is also convex. The statement is proven.

5°. Let us now prove that F contains a function $f_6(x_1)$ of one variable for which $f_6(0) = f_6(1) = 0$ and $f_6(\alpha) < 0$ for some $\alpha \in (0, 1)$.

Similarly to Part 2° of this proof, we take

$$f_6(x) = f_5(x) - f_5(0) - x \cdot (f_5(1) - f_5(0)). \tag{8}$$

6°. Let us now prove that F contains a function $f_7(x_1)$ of one variable for which $f_7(0) = 0$, $f_7(x) \geq 0$ for all $x \in (-1, 1)$ and $f_7(x) > 0$ for all $x \in [0, 1)$.

We will construct this function f_7 from the above function f_6. We know that $f_6(0) = f_6(1) = 0$ and that $f_6(\alpha) < 0$. Let m denote the minimum value of the function $f_6(x)$ on the interval $[0, 1]$, and let α_0 denote the supremum of all the values $x \in (0, 1)$ at which the function $f_6(x)$ attains this minimum value m. Then, $f_6(x) \geq m$ for all $x \in [0, 1]$, and $f_6(x) > m$ for all $x > \alpha_0$.

If we take $\Delta \stackrel{\text{def}}{=} \min(\alpha_0, 1 - \alpha_0)$, then $[\alpha_0 - \Delta, \alpha_0 + \Delta] \subseteq [0, 1]$. Thus, for $f_7(x) \stackrel{\text{def}}{=} f_7(\alpha_0 + x \cdot \Delta) - m$, we have $f_7 \in F$, $f_7(0) = 0$, $f_7(x) \geq 0$ for all $x \in [-1, 1]$ and $f_7(x) > 0$ for all $x \in (0, 1]$.

7°. Finally, let us now prove that F contains a function $f_8(x_1)$ of one variable for which $f_8(0) = 0$ and $f_8(x) > 0$ for all $x \in [-1, 1]$ for which $x \neq 0$.

Indeed, we can take $f_8(x) \stackrel{\text{def}}{=} f_7(x) + f_7(-x)$.

8°. We are now ready to prove Theorem 1.

To prove NP-hardness of the global minimization problem for the class F, we will reduce, to this problem, a known NP-problem, namely, the following *subset sum* problem [4, 5]: Given n positive integers $s_1, \ldots, s_n$ and an integer $s > 0$, check whether it is possible to find a subset of this set of integers whose sum is equal to exactly s.

For each i, we can take $x_i = 0$ if we do not include the i-th integer in the subset, and $x_i = 1$ if we do. Then, the subset problem takes the following form: check whether there exist values $x_i \in \{0, 1\}$ for which $\sum s_i \cdot x_i = s$.

We will reduce each instance of this problem to the problem of minimizing a function $f_9(x_1, \ldots, x_n)$ on the box $[0, 1]^n$, where f is defined as follows:

$$f_9(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_4(x_i) + f_8\left(\sum_{i=1}^{n} s'_i \cdot x_i - s'\right), \tag{9}$$

where $s'_i \stackrel{\text{def}}{=} s_i / S$, $s' \stackrel{\text{def}}{=} s/S$, and $S \stackrel{\text{def}}{=} s + \sum s_i$.

Since the class F is closed, the function (9) belongs to the class F. Let us prove that the minimum of the function (9) is equal to 0 if and only if the original subset problem has a solution.

Indeed, due to the choice of S, we have $|\sum s'_i \cdot x_i - s'| \leq 1$. Thus, due to Part 6 of this proof, we have $f_8(\sum s'_i \cdot x_i - s') \geq 0$. Due to Part 3 of this proof, we have $f_4(x_i) \geq 0$. Thus, the function f_9, as a sum of non-negative terms, is always non-negative. The only way for this function to be equal to 0 is when all the non-negative terms are equal to 0. Due to Parts 3 and 6, this is possible only if for every i, $x_i = 0$ or $x_i = 1$, and if $\sum s'_i \cdot x_i = s'$ – hence $\sum s_i \cdot x_i = s$. Thus, if the minimum is 0, the subset sum problem has a solution.

Vice versa, if the subset sum problem has a solution $x_1, \ldots, x_n$, then for these values x_i, we will have $f_9(x_1, \ldots, x_n) = 0$. Hence, in this case, the minimum of the function (9) is equal to 0. The reduction is proven, thus the minimization proven is NP-hard. Theorem 1 is proven.

9°. We will start proving Theorem 2 by showing that for every $\delta > 0$, there exists a $\beta > 0$ for which $f_8(x) \le \delta$ and $x \in [-1,1]$ implies $|x| \le \delta$.

We can prove this by reduction to a contradiction. If the statement that we try to prove is false, this means that there exists a $\delta > 0$ such that for every β, there exists an $x(\beta)$ for which $f_8(x(\beta)) \le \beta$ and $|x(\beta)| \ge \delta$. All the values $x(\beta)$ belong to the same compact set $[-1,1]$, thus, from the sequence $x(\beta)$, we can extract a converging subsequence $x(\beta_k) \to x_0$. For its limit x_0, we have $f_8(x_0) = 0$ and $|x_0| \ge \delta$, which contradicts to what we proved in Part 6. This contradiction shows that our statement is indeed true.

10°. Similarly, we can prove that for every $\delta > 0$, there exists a $\gamma > 0$ such that if $f_4(x) \le \gamma$ and $x \in [0,1]$, then either $x \le \delta$ or $x \ge 1-\delta$.

11°. Let us now fix $\varepsilon > 0$, and let us reduce the subset sum problem to the problem of finding the minimum of functions $f \in F$ with accuracy ε.

For every instance of the subset sum problem, we will take $\delta \stackrel{\rm def}{=} 0.2/S$. For this δ, let β and γ denote the values described in Parts 9 and 10 of this proof. Let us denote $\varepsilon_0 \stackrel{\rm def}{=} \min(\beta,\gamma)$. Then, as the desired function $f \in F$, we take a function $f_{10} = (3\varepsilon/\varepsilon_0) \cdot f_9$, where f_9 is described by the formula (9).

If the subset sum problem has a solution, then the minimum of the function f_{10} is equal to 0. Let us show that if the minimum of the function f_{10} is smaller or equal than 3ε, then the subset problem is equal to 0. Thus, the minimum is either equal to 0, or larger than 3ε.

- In the first case, if we compute the ε-approximation $\widetilde{m}$ to the minimum m, we get $\widetilde{m} \le \varepsilon$.

- In the second case, if we compute the ε-approximation $\widetilde{m}$ to the minimum m, we get $\widetilde{m} > 2\varepsilon$.

Thus, by comparing $\widetilde{m}$ with ε, we will be able to tell whether the original instance of the subset sum problem has a solution.

So, to complete the proof of Theorem 2, we must show that if the minimum m of the function f_{10} is not larger than 3ε, then the original instance of the subset problem has a solution. Indeed, this minimum is attained for some inputs $x_1, \ldots, x_n \in [0,1]$. Since $f_{10} = (3\varepsilon/\varepsilon_0) \cdot f_9$, for these inputs, the function f_9 takes the value $\le \varepsilon_0$.

The expression (9) that defines the function f_9 is the sum of non-negative terms. Thus, each of these terms is not larger than $\varepsilon_0 = \min(\beta,\gamma)$, and thus, not larger than β and not larger than γ. From Part 10 and $f_4(x_i) \le \gamma$, we conclude that either $x_i \le \delta$ or $x_i \ge 1-\delta$. In other words, if by $\widetilde{x}_i$, we denote the integer that is closest to x_i, we conclude that

$$|x_i - \widetilde{x}_i| \le \delta. \tag{10}$$

Similarly, from $f_8(\sum s_i' \cdot x_i - s') \le \beta$, we conclude that $|\sum s_i' \cdot x_i' - s'| \le \delta$. Multiplying both sides of this inequality by S, we get

$$\left|\sum s_i \cdot x_i - s\right| \le S \cdot \delta. \tag{11}$$

¿From (10), we conclude that

$$\left|\left(\sum s_i \cdot \widetilde{x}_i - s\right) - \left(\sum s_i \cdot \widetilde{x}_i - s\right)\right| = \left|\sum (\widetilde{x}_i - x_i) \cdot s_i\right| \le$$

$$\delta \cdot \sum s_i \le \delta \cdot S. \tag{12}$$

¿From (11) and (12), we conclude that

$$\left|\sum s_i \cdot \widetilde{x}_i - s\right| \le 2 \cdot \delta \cdot S. \tag{13}$$

By definition of δ, the product $2 \cdot \delta \cdot S$ is equal to 0.4. Thus, the absolute value of the integer $\sum s_i \cdot \widetilde{x}_i - s$ does not exceed 0.4. The only such integer is 0. Hence, $\sum s_i \cdot \widetilde{x}_i - s = 0$, i.e., the original instance of the subset sum problem indeed has a solution. The theorem is proven.

Acknowledgments

This work was partially supported by the NASA grant NCC5-209, by NSF grants EAR-0112968 and EAR-0225670, by the Air Force Office of Sponsored Research grant F49620-00-0365, by the Army Research Laboratories grant DATM-05-02-C-0046, and by the University of Louisiana at Lafayette.

References

[1] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*, Springer-Verlag, London, 2001.

[2] R. B. Kearfott, *Rigorous Global Search: Continuous Problems*, Kluwer, Boston, MA, 1996.

[3] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Boston, Massachusetts, 1996.

[4] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1997.

[5] C. H. Papadimitriou, *Computational Complexity*, Addison-Wesley, Reading, Massachusetts, 1994.

[6] S. A. Vavasis, *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.