

A Note on Open Problems and Challenges in Optimization Theory and Algorithms



A. Migdalas and P. M. Pardalos

Abstract In this note, we review some open problems and challenges concerning optimization theory and the algorithms.

Keywords Combinatorial optimization · Computational complexity · Global optimization · Meta-heuristics

1 Introduction

Optimization is an important subject with a wide field of applications in decision-making problems that arise in economics, engineering, logistics and transportation, traffic planning, location and layout of facilities, telecommunications, social and biological networks, machine learning, and other fields (see, e.g., [12]). According to Werner [25], Cantor [9] observed that the first formal optimization problem in history was formally stated, ca. 300 BC, in Euclid's Elements of Geometry (see, e.g., [11]), Book VI, and is concerned with the identification of a point on a triangle's side such that the resulting parallelogram, obtained by drawing the parallel lines to the other two triangle sides from this point, has maximal area. Werner cites also another ancient Greek mathematician, Heron (ca. 100 BC), who solved another optimization problem in geometry, that of finding a point on a given line such that

A. Migdalas (✉)

Industrial Logistics, ETS Institute, Lulea University of Technology, Norrbotten, Sweden

Aristotle University of Thessaloniki, Department of Civil Engineering, Thessaloniki, Central Macedonia, Greece

e-mail: athmig@ltu.se; samig@civil.auth.gr

P. M. Pardalos

Industrial and Systems Engineering Department, University of Florida,

Center For Applied Optimization, Gainesville, FL, USA

e-mail: pardalos@ise.ufl.edu

© Springer Nature Switzerland AG 2018

P. M. Pardalos, A. Migdalas (eds.), *Open Problems in Optimization and Data Analysis*, Springer Optimization and Its Applications 141,

https://doi.org/10.1007/978-3-319-99142-9_1

the sum of the distances from two other given points is minimal. Thus, it seems that historically, optimization has its roots in geometry, and actually, optimization still plays an important role in geometry (see, e.g., [5]) and particularly in computational geometry (see, e.g., [4, 8]).

While some solution algorithms have been developed for most interesting optimization problems, and although several different solution approaches have been proposed for the same optimization problems, their performance may not always be satisfactory in practice and/or in theory. Thus, while every NP-hard problem can be solved in exponential time by exhaustive search, the question whether we can do better than a trivial enumeration is naturally raised. Could it ever be possible to solve NP-hard problems in quasi-polynomial time?

Modern heuristic and meta-heuristic approaches (from the Greek “Heuriskein”—*εὕρισκειν*—to find, to discover) have become so popular that a new optimization branch has been born that devises algorithms inspired from physics or nature, with such names as simulated annealing, ant colony, or particle swarm optimization. However, their mathematical properties and convergence remain largely unaddressed, and many open problems need attention (see, e.g., [29]). Moreover, the “No Free Lunch Theorem” shows that all non-resampling algorithms perform equally, averaged over all problems, that is, no such algorithm can outperform any other under any metric over all problems (see, e.g., [18]). Similar practical and theoretical considerations have led to the development of algorithm portfolios for certain optimization problems (see, e.g., [13]) whose theoretical and computational properties remain, however, largely unexplored.

Under suitable convexity assumptions in nonlinear optimization, “exact” optimization algorithms find the global optimal solution if one exists. However, such “exact” algorithms are of limited use for global optimization problems, although they may be adopted as heuristics in “local” search techniques (see, e.g., [10]). Moreover, there are several issues here: One is the fact that there are optimization problems for which only global optimization matters [16]. Another is the complexity of determining the convexity of the problem [2, 23]. Black-box optimization, where the objective function is known only by observing different sets of input–output pairs from a computational simulation or experiment is the reign of heuristics and meta-heuristics. So, questions can be raised both with respect to problem convexity and to the “No Free Lunch Theorem” in continuous domain [3, 24].

Computing power is essential for all optimization algorithms. But, what are the limits of what the humans can compute and what are the limits of the machines? How far can emerging technologies, including quantum computers, stretch these limits, and what will be their impact on optimization algorithms? Such issues are discussed in, e.g., [7, 20] and [21].

2 Some Open and Challenging Problems

Nonlinear optimization is a great source of challenges and open problems. It is well known in this context that a global optimum can be provably attained by optimization algorithms based on local properties only under suitable convexity assumptions. But, how easy is it to prove convexity? One of the seven open problems in complexity theory for numerical optimization listed by Pardalos and Vavasis in [23] is the following:

Given a degree 4 polynomial in n variables, what is the complexity of determining whether this polynomial describes a convex function?

It was shown by Ahmadi et al. [2] that unless $P=NP$, there exists no polynomial time, not even pseudo-polynomial time, algorithm that can decide whether a multivariate polynomial of degree four or higher even degree is globally convex. They even show that deciding strict, strong, quasi- and pseudo-convexity of polynomials of degree four or higher even degree is strongly NP-hard, while quasi- and pseudo-convexity of odd degree polynomials can be decided in polynomial time. So, the question whether determining convexity of a general function is a “decidable problem” is open. Another important open problem related to convexity is whether in the case of d.c. optimization it is possible to characterize the “best” d.c. decomposition of a function into the difference of two convex functions.

The problem of minimizing a convex function f over a convex set $X \subseteq \mathbb{R}^n$, where the only access to f is via a stochastic gradient oracle, which given a point $\mathbf{x} \in X$ returns a random vector $\mathbf{g}(\mathbf{x})$ such that $\mathcal{E}[\mathbf{g}(\mathbf{x})] = \nabla f(\mathbf{x})$, is known as the “stochastic exp-concave optimization problem” and is of great importance as it captures several fundamental problems in machine learning [1, 19]. Optimization algorithms, such as the stochastic gradient descent algorithm, are used in order to obtain a point $\hat{\mathbf{x}}$ for which $f(\hat{\mathbf{x}}) - \min_{\mathbf{x} \in X} f(\mathbf{x}) \leq \epsilon$, for given target accuracy ϵ , either in expectation or with high probability. Despite the importance of this problem, current algorithms scale poorly with the dimension n of the problem. Therefore, algorithms with fast rate and which scale better with the dimension are sought. Attempts are discussed in, e.g., [1, 19].

There are several challenging problems in continuous global optimization, both theoretical and algorithmic ones, such as whether it is possible to derive general optimization conditions, whether it is possible to decide upon the feasibility in the case of large constrained problems, and how to utilize sparsity and other inherent structures in order to attack large-scale problems. However, even certain fundamental questions regarding optimality of global optimization problems may be indeed very hard to answer. Consider, for instance, the quadratic problem:

$$\begin{aligned} \min f(\mathbf{x}) &= \mathbf{c}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} \\ \text{s.t. } \mathbf{x} &\geq \mathbf{0}, \end{aligned}$$

where \mathbf{Q} is an arbitrary $n \times n$ symmetric matrix, and $\mathbf{x} \in \mathbb{R}^n$.

The Karush–Kuhn–Tucker optimality conditions for this problem become a so-called linear complementarity problem, $LC P(\mathbf{Q}, \mathbf{c})$, which is formulated as follows: Find $\mathbf{x} \in \mathbb{R}^n$, or prove that none such exists, that satisfies the system:

$$\begin{aligned} \mathbf{Q}\mathbf{x} + \mathbf{c} &\geq \mathbf{0}, \quad \mathbf{x} \geq \mathbf{0}, \\ \mathbf{x}^T (\mathbf{Q}\mathbf{x} + \mathbf{c}) &= 0 \end{aligned}$$

In 1994, it was shown by Horst et al. that the $LC P(\mathbf{Q}, \mathbf{c})$ is an NP-hard problem (see, e.g., [17]). In fact, the problem of checking local optimality for a feasible point is not that easy either. Indeed, consider the linearly constrained problem:

$$\begin{aligned} \min f(\mathbf{x}) \\ \text{s.t.} \quad \mathbf{A}\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned}$$

where $f(\mathbf{x})$ is indefinite quadratic function. The same researchers have shown that the problem of checking the strict local optimality of a feasible point \mathbf{x} for the above problem is also NP-hard. However, even if local optimality can be proven, it may be pointless, since as Hiriart-Urruty [16] has shown, there exist problems for which every feasible point is also a local optimizer. Two such problems are:

The problem of minimizing the rank of a matrix,

$$\begin{aligned} \min f(\mathbf{A}) &= \text{rank}(\mathbf{A}) \\ \text{s.t.} \quad \mathbf{A} &\in \mathcal{C}, \end{aligned}$$

where $\mathcal{C} \subset \mathcal{M}_{m,m}(\mathbb{R})$, the vector space of $m \times n$ real matrices, and the related problem of minimizing the so-called counting function of nonzero components,

$$\begin{aligned} \min c(\mathbf{x}) \\ \text{s.t.} \quad \mathbf{x} &\in S, \end{aligned}$$

where $S \subset \mathbb{R}^n$ and $c(\mathbf{x})$ = number of $x_i \neq 0$ in \mathbf{x} , have been shown in [16] to possess the property that “every feasible point is a local minimizer.” Clearly, “only global optimization matters” for such problems [16].

Meta-heuristics provide the means to decide which part of the search space should be explored next. The local exploration is typically performed by some local minimization algorithms or some other heuristic approaches. Meta-heuristics have shown to be successful in practice for a large number of important combinatorial and global optimization problems. Of particular, fruitful importance for meta-heuristic application have been the so-called black-box optimization problems. Such a problem can be stated as follows:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in X, \end{aligned}$$

where $X \subset \mathbb{R}^n$, and $f : X \rightarrow \mathbb{R}$ is function which is known only through a set of pairs of input and output values obtained through an experiment or simulation, that is, it is only known as a set $D = \{(\mathbf{x}^1, f(\mathbf{x}^1)), (\mathbf{x}^2, f(\mathbf{x}^2)), \dots, (\mathbf{x}^n, f(\mathbf{x}^n))\}$. There are quite a few challenges and open questions in relation to these very important problems in engineering design. One important issue is with respect to the “no free lunch theorem,” which in the case of combinatorial optimization roughly states that all non-resampling optimization algorithms perform equally, averaged over all problems, and therefore no optimization algorithm can outperform any other under any metric over all problems. In [18], Joyce and Herrmann summarize the following results from the literature which emphasize different aspects of the theorem:

1. The average performance of any pair of algorithms across all possible problems is identical.
2. For all possible metrics, no search algorithm is better than another when its performance is averaged over all possible discrete functions.
3. On average, no algorithm is better than random enumeration in locating the global optimum.
4. The histogram of values seen, and thus any measures or performance based on it, is independent of the algorithm if all functions are considered equally likely.
5. No algorithm performs better than any other when their performance is averaged over all possible problems of a particular type.
6. With no prior knowledge about the function f , in a situation where any functional form is uniformly admissible, the information provided by the value of the function in some points in the domain will not say anything about the value of the function in other regions of its domain.

The last statement (6) was proved by Serafino [24] for exactly the case of black-box optimization. Hence, the prior knowledge about the objective function landscape is the key for success. Since all meta-heuristics assume such a priori model, their success largely depends upon the fitness of the model geometry to the geometry of the problem under consideration [24]. Thus, although new meta-heuristics are often introduced as a panacea, the “no free lunch theorem” says otherwise and emphasizes the need of obtaining prior knowledge of the problems geometry. Understanding this type of Bayesian prior and its relation to the “no free lunch theorem” seems as important for the development of successful optimization algorithm as is prior knowledge important for successful learning in machine learning [18]. Indeed, the importance of prior information on the objective function that would permit to choose algorithms that perform better than pure blind search is emphasized in the case of continuous search domain where the necessary conditions for the “no free lunch theorem” are shown to be even stronger and far more restrictive [3].

Evaluation of heuristic and meta-heuristic algorithms raises some interesting challenging computational problems with respect to experimental testing, supply of good lower- and upper-bound techniques, supply of benchmark instances with known optimal solution, and also derivation of techniques for automatic identification of parameter values. Concerning theory, the mathematical properties of almost all meta-heuristic algorithms remain largely unaddressed or unsatisfactorily investigated and therefore constitute challenging issues [29]. Population-based meta-heuristics can be addressed by studying the interaction of multiple Markov chains corresponding to the search formations. Theoretical development along these lines has already been initiated but it is in early stages. The mathematical analysis concerning the rate of convergence of population-based continues to constitute a challenging issue. Obtaining strategies and techniques that lead to a balanced trade-off between local intensification and global diversification is another important issue. Deriving combination of algorithms into algorithm portfolios that adaptively fit the assumed model geometry to the real geometry of the optimized problem is of interest. Studying the approach in relation to the “no free lunch theorem” is of theoretical importance. Conditions under which algorithm portfolios can provide computational advantages over other approaches are also of interest. In [13], it is shown that a “risk-seeking” strategy can be advantageous in a portfolio setting. What other strategies could be proven advantageous and for which kind of problems?

3 Concluding Remarks

We have indicated above a few directions along which interesting challenges and open problems can be identified for further research. There are a lot more such challenges and open questions in several other sub-subject areas. In [15], several conjectures and open problems, including nonlinear optimization, are presented. Concerning open problems about exact algorithms and their worst time bounds and worst case bounds for NP-hard problems, the reader should consult [28]. Open problems concerning the theory of approximation algorithms for NP-hard discrete optimization problems are discussed in [27]. West [26] has gathered 38 open problems concerning both theory and optimization in the context of graph theory and combinatorics. A lengthy and well-structured list of open combinatorial optimization problems on graphs, including important problems of broadcasting and gossiping, as well as open problems concerning complexity is provided by Hedetniemi [14]. Computational challenges of cliques and related problems are the subjection in [22]. Finally, a source of open problems with respect to current and future algorithmic development for the solution of complex optimization problems is the book by Battiti et al. [6]

References

1. Agarwal, N., Gonen, A.: Effective dimension of exp-concave optimization (2018). <https://arxiv.org/abs/1805.08268>
2. Ahmadi, A.A., Olshevsky, A., Parrilo, P.A., Tsitsiklis, J.N.: NP-hardness of deciding convexity of quartic polynomials and related problems. *Math. Program.* **137**, 453–476 (2013)
3. Alabert, A., Berti, A., Caballero, R., Ferrante, M.: No-free-lunch theorem in continuum. *Theor. Comput. Sci.* **600**, 98–106 (2015)
4. Allen-Zhu I, Z., Liao, Z., Yuan, Y.: Optimization algorithms for faster computational geometry. In: Chatzigiannakis, I., Mitzenmacher, M., Rabani, Y., Sangiorgi, D. (eds.) 43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016), Article No. 53, pp. 53:1–53:6. (2016)
5. Andreatt, M., Bezdek, A., Boronski, J.P.: The problem of Malfatti: two centuries of debate. *Math. Intell.* **33**, 72–76 (2011)
6. Battiti, R., Brunato, R., Mascia, F.: *Reactive Search and Intelligent Optimization*. Operations Research/Computer Science Interfaces Series. Springer, Berlin (2009)
7. Bennet, C.H., Landauer, R.: The fundamental physical limits of computation. *Sci. Am.* **253**, 48–56 (2014)
8. Bezdek, K., Deza, A., Ye, Y.: Selected open problems in discrete geometry and optimization. In: Bezdek, K., et al. (eds.) *Discrete Geometry and Optimization*, pp. 321–336. Springer, Berlin (2013)
9. Cantor, M.: *Vorlesungen über Geschichte der Mathematik, Band 1*. B.G. Teubner, Leipzig (1880)
10. D’Apuzzo, M., Marino, M., Migdalas, A., Pardalos, P.M., Toraldo, G.: Parallel computing in global optimization. In: Kontogiorghes, E.J. (ed.) *Handbook of Parallel Computing and Statistics*, pp. 225–258. Chapman and Hall, Boca Raton (2006)
11. *Euclid’s Elements of Geometry* (Edited, and provided with a modern English translation, by Richard Fitzpatrick). <http://farside.ph.utexas.edu/Books/Euclid/Elements.pdf>
12. Floudas, C.A., Pardalos, P.M. (eds.): *Encyclopedia of Optimization*, 2nd edn. Springer, Berlin (2009)
13. Gomes, C.P., Selman, B.: Algorithm portfolios. *Artif. Intell.* **126**, 43–62 (2001)
14. Hedetniemi, S.: Open problems in combinatorial optimization (1998). <https://people.cs.clemson.edu/~hedet/preface.html>
15. Hiriart-Urruty, J.-B.: Potpourri of conjectures and open questions in nonlinear analysis and optimization. *SIAM Rev.* **49**, 255–273 (2007)
16. Hiriart-Urruty, J.-B.: When only global optimization matters. *J. Glob. Optim.* **56**, 761–763 (2013)
17. Horst, R., Pardalos, P.M., Thoai, N.V.: *Introduction to Global Optimization*. Nonconvex Optimization and Its Applications, vol. 3. Springer, Berlin (2000)
18. Joyce, T., Herrmann, J.M.: A review of no free lunch theorems, and their implications for metaheuristic optimisation. In: Yang, X.-S. (ed.) *Nature-Inspired Algorithms and Applied Optimization*. Studies in Computational Intelligence, vol. 744, pp. 27–51. Springer, Berlin (2018)
19. Kore, T.: Open problem: fast stochastic exp-concave optimization. In: *JMLR: Workshop and Conference Proceedings*, vol. 30, pp. 1–3 (2013)
20. Markov, I.L.: Limits on fundamental limits to computation. *Nature* **512**, 147–154 (2014)
21. Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge Series on Information and Natural Sciences. Cambridge University Press, Cambridge (2000)
22. Pardalos, P.M., Rebennack, S.: Computational challenges with cliques, quasi-cliques and clique partitions in graphs. In: Festa, P. (ed.) *International Symposium on Experimental Algorithms, SEA 2010: Experimental Algorithms*. Lecture Notes in Computer Science, vol. 6049, pp. 13–22. Springer, Berlin (2010)

23. Pardalos, P.M., Vavasis, S.A.: Open questions in complexity theory for numerical optimization. *Math. Programm.* **57**, 337–339 (1992)
24. Serafino, L.: Optimizing without derivatives: what does the no free lunch theorem actually say? *Not. AMS* **61**, 750–755 (2014)
25. Werner, J.: *Optimization Theory and Applications*. Vieweg Advanced Lectures in Mathematics. Friedr. Vieweg & Son, Braunschweig (1984)
26. West, D.B.: Open problems - graph theory and combinatorics (2018). <https://faculty.math.illinois.edu/~west/openp/>
27. Williamson, D.P., Shmoys, D.B.: *The Design of Approximation Algorithms*. Cambridge University Press, Cambridge (2010)
28. Woeginger, G.J.: Open problems around exact algorithms. *Discrete Appl. Math.* **156**, 397–405 (2008)
29. Yang, X.-S.: Metaheuristic Optimization: Algorithm Analysis and Open Problems. In: Pardalos, P.M., Rebennack, S. (eds.) *Experimental Algorithms*. SEA 2011. Lecture Notes in Computer Science, vol. 6630. Springer, Berlin (2011)