Finding global minima of non-convex functions via swarm based gradient descent

Moritz Danzebrink
mdanzeb1@uni-koeln.de

Janina Tikko
jtikko2@uni-koeln.de

Department of Mathematics and Computer Science, Division of Mathematics
University of Cologne

Abstract

We present a Swarm-Based Gradient Descent method (SBGD) introduced in [JTZ]. The underlying problem is to find global minima and can be formulated as

$$\operatorname*{argmin}_{\mathbf{x}\in\Omega\subset\mathbb{R}^d} f(\mathbf{x})$$

for a function $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$. The challenge is to overcome basins of local minima in order to find the true global minima. Other methods like classical Gradient Descent depend heavily on starting positions, hence, they struggle to escape local minima. Our presented approach combines Gradient Descent with swarm-behavior in order to overcome the difficulties of local minima.

A swarm contains a number of agents which are each defined by a position $\mathbf{x}$ and a mass m. The agents explore the ambient space by moving along the direction of the gradient. As stepsize-protocol the backtracking-line-search method is used, as proposed by Tadmor et al [JTZ]. Key aspect in this method is communication between agents through transferring parts of their mass. Each agent receives a different stepsize $h = h(\mathbf{x}, m)$, depending on its individual mass. This leads to lighter agents exploring regions which are further away. On the contrary heavier agents receive smaller stepsizes approaching potential minima. Naturally the swarm is divided in lighter explorers which improve the swarms overall position and a global leader. The global leader might change through the process due to an improved minimizer.

During our research we demonstrated the operating principle [T]. We give a more visual understanding of how the new method works and how different parameters impact the algorithm. In particular we analyzed the influence of the relative mass on the backtracking protocol [D]. Moreover, we studied the quality of the error with respect to the number of agents.

References

[D] M. Danzebrink, *Optimization approach for Swarm-Based Gradient Descent in multiple arguments*, Bachelor Thesis, Department of Mathematics and Computer Science, Division of Mathematics, University of Cologne, November 2024.

[JTZ] L. Jingcheng, E. Tadmor, A. Zenginoğlu, *Swarm-based gradient descent method for non-convex optimization*, Communications of the American Mathematical Society, Volume 4, Pages 787-822, 2024, doi:10.1090/cams/42.

[T] J. Tikko, *Introduction: Swarm-based gradient descent for non convex optimization*, 2024, https://arxiv.org/abs/2404.00005.