CME LIVE OPEN ACCESS

Journal of Biomedical and Engineering Research

Research Article

ISSN: 3065-8780

Plane Based Optimization Method: A Method for Better Global Minimum Cost Function Localization

Dalton H Bermudez*

Department of Medical Physics, University of Wisconsin-Madison

Corresponding Author: Dalton H Bermudez, Department of Medical Physics, University of Wisconsin-Madison

Received: 2025 Jan 28 **Accepted:** 2025 Feb 17 **Published:** 2025 Mar 30

Abstract

Background: *As a popular optimization algorithm, gradient descent is more prone to converge at a local minimum rather than a global one since it is highly dependent on how initial values are instantiated and an iterative process of updates, reducing its efficiency for non-convex cost functions. This limitation is of utmost importance in domains such as medical physics, where there exist many treatment parameters that need to be optimized.*

Purpose: *We introduce the flat plane with maximum perpendicular distance method, a new optimization algorithm which identifies global minima robustly and efficiently in convex cost functions. This facilitates an initialization-free, non-iterative update method that serves as a complementary approach to classic methods.*

Methods: *1D and 2D convex cost functions, with both global and relative minima, were tested using the flat plane method. Further, the approach was employed to determine optimal locations for radiation beams in a medical physics problem. Results were compared to existing gradient descent methods for effectiveness assessment.*

Results: *The flat plane algorithm found the global minimum in all the scenarios considered while the use of gradient descent would have failed to do so in many cases due to initialization issues. However, due to a less direct but highly parallelizable strategy (i.e. flat plane method), these techniques are certainly scalable for high-dimensional problems with potentially higher expense (e.g. dimensionality reduction, sampling, parallel computing).*

Conclusion: *In low-dimensional problems, the flat plane method provides a strong alternative to gradient descent for the global optimization problem. We are yet to adapt the method for high-dimensional applications but will work in that direction to increase its applicability over more complex optimization problems.*

Keywords: Global Minimum, Optimization, Engineering, Parabolic Function

1. Introduction

Optimization techniques are essential for solving many complex scientific and engineering problems, especially when identifying minima of cost functions for decision-making. Gradient descent is a simple and computationally efficient method (in terms of number of operations) used in many applications. Nevertheless, such approaches are usually limited in that they depend on local knowledge and they are susceptible to getting stuck in local minima. This limitation becomes even more significant in problems where the optimization landscape follows a complex, high-dimensional, and irregular geometry. We address this gap through a simple optimization method: the flat plane with maximum perpendicular distance. A flat plane is created on the maximum point of the function of cost; then the perpendicular distances of all points below is measured. This way, it does not depend on any initialization or iterative updates, therefore being resilient in finding global minima given that there may exist several local minima. To show the effectiveness of this approach, we evaluated it on both simulated convex functions in one and two dimensions, as well as a higher dimensional medical physics optimization problem to determine radiation beam placement. Such robustness and potential point to the use of such an approach for optimization in scientific and engineering domains.

2. Methods/Theory

2.1 Mathematical Proof: Flat Plane with Maximum Perpendicular Distance for Convex Problems

Definition and Assumptions

Let f(x) be a convex function defined over a domain $x\epsilon\mathbb{R}^n$.
A function is convex if for any $x1,x2\epsilon\mathbb{R}^n$ and λϵ[0,1], the following holds:

$$f(\lambda x1+(1-\lambda)x2)\leq\lambda f(x1)+(1-\lambda)f(x2)$$

The global minimum of f(x), denoted as x*, satisfies f(x*)≤f(x) for all $x\epsilon\mathbb{R}^n$.

The flat plane method defines a horizontal plane P at the maximum value of the functions, fman, and calculates the perpendicular distance d(x)= P-f(x) for all x. The global minimum is identified as the point where d(x) is maximized. For convex functions, the point with the maximum perpendicular distance below the plane P corresponds to the global minimum of the function. This result holds because of the properties of convexity and the monotonic relationship between f(x) and d(x).

2.1 Proof

Flat plane is defined as

$$P = \text{fmax} = \max_{x\epsilon\mathbb{R}^n} f(x)$$

The perpendicular distance from the flat plane to the function at point x is:

D(x) = P-f(x)

Since f(x)≤fmax for all x, d(x)≥0.

Maximization of the perpendicular distance by maximizing d(x) is equivalent to minimizing f(x), as:

$$arg \max_{x\epsilon\mathbb{R}^n} d(x) = \arg \min_{x\epsilon\mathbb{R}^n} f(x)$$

At the global minimum x*, f(x*) is minimized, and d(x*) = P-f(x*) is maximized.

Of convexity, f(x) has a unique global minimum x* if it is strictly convex. For non-strictly convex functions, the global minimum might not be unique, but all global minima have the same value f(x*), resulting in the same maximum d(x*). The flat plane P is constant across the domain. The maximum perpendicular distance occurs where f(x) is minimized, as the distance d(x) decreases monotonically with increase f(x).

2.3 Simulations of Cost Functions

The new optimization method explores two distinct approaches for locating the global minimum of convex cost functions: a gradient descent and the flat plane with maximum perpendicular distance method. These techniques are applied to both one-dimensional and two-dimensional convex cost functions, including scenarios with relative and global minima, offering a robust framework for optimization analysis. First a simple convex function, $f(x) = x^2$, and a two-hump convex function, which models a global minimum at one center and a relative minimum at another by combining two parabolic functions with varying scales and centers. For the simulations the two-hump convex functions had a shift of 5, a global minimum parabolic function had a scale of 0.5 and the relative minimum parabolic function had a scale of 1. For the 1-dimensional convex problem a parameter space from -10 to 10 composed of 500 points was utilized. For the 2-dimensional convex problem a parameter space from 10 to 10 with 100 points in x and y directions was utilized.

A non-convex multidimensional cost function was modeled with the following equations:

$$Z = \cos(\text{sqrt}(X^2 + Y^2)) + 0.5 \sin(2X)\sin(2Y);$$

The cost functions surface parameter space was defined in a grid -5 to 5 in both x and y domain with 100 sampled points. The proposed flat plane with maximum perpendicular distance method was use to see how robust is this method is at finding the global minimum in nonconvex scenarios.

2.4 Radiation Beam Optimization

In radiation therapy, the goal is to minimize the dose to surrounding healthy tissue and maximize dose to tumor. The defined cost function that includes Dh which is the dose to heathy tissue, Dt dose to tumor, and a penalty term for exceeding a certain dose threshold in healthy tissue.

2.4.1 The Corresponding Cost Function is

$$C(\mathbf{x}) = whDh(\mathbf{x})\text{-}wtDt(\mathbf{x})+\lambda\max(0,Dh(\mathbf{x})\text{-}Dh^{threshold})$$

Where x is the beam parameters (e.g positions, angles, intensity), wh, wt are the weights for healthy tissue and tumor doses, respectively. λ is the penalty weight for exceeding the healthy tissue dose threshold. The $Dh^{threshold}$ was set to 0.5, initial weights wh and wt were set to 1. The initial beam positions was set to -5.

$Dh(x) = \exp(-((x\text{-healthy center})^2))$

$Dt(x) = \exp(-((x\text{-tumor center})^2))$

The tumor center and healthy center were set to 2 and -2 respectively. This optimization problem was solved using standard gradient descent with learning rate of 0.1 and convergence tolerance of $1*10^{-5}$. To Update the beam parameters iteratively the following functions was used.

$$X^{k+1} = x^k - \eta\Delta C(x^k)$$

Where η is the learning rate.

For this optimization the gradient descent method was compared with the proposed new flat plane with maximum perpendicular distance method. For this method a flat plane at the maximum value of the cost function was created and the corresponding perpendicular distances from the plane to all the points in the cost function computed. Then the global minimum was identified as the point with the maximum perpendicular distance below the plane.

3. Results

Citation: Bermudez, D. H. (2025). Plane Based Optimization Method: A Method for Better Global Minimum Cost Function Localization. *J Biomed Eng Res, 3*(1), 1-5.

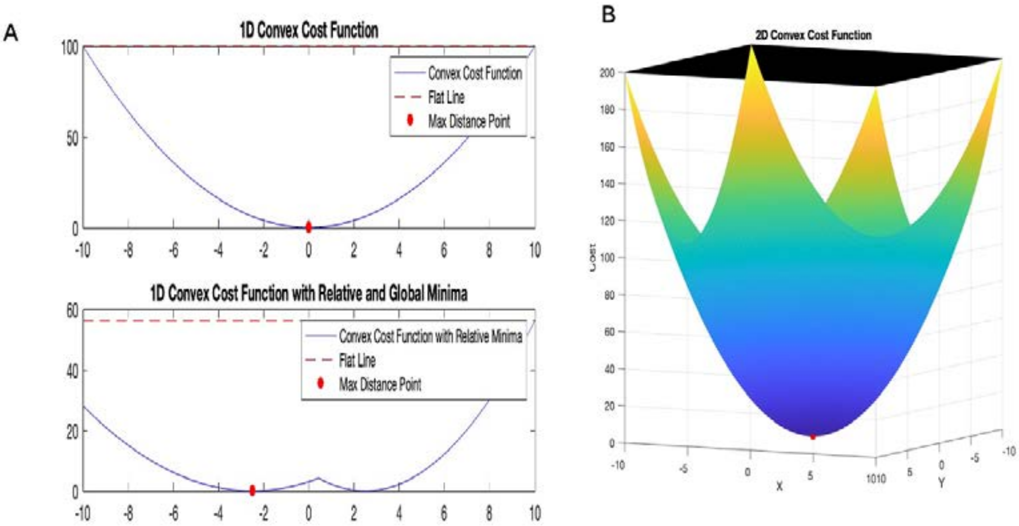

Figure 1: Simulated Convex Cost Functions in A) 1 Dimensional and B) Two Dimensional Where the Proposed Flat Plane with Maximum Perpendicular Distance Method was Used to Find the Global Minimum Specified by the Red Dot Being the Point Found by the Proposed Method

In the simulated 1-dimensional and 2-dimensional convex cost functions the proposed flat plane with maximum perpendicular distance method found efficiently the global minimum in both cases where the const function had a unique minimum and the cost functions where the cost function had a relative minimum and a global minimum. Visualization is employed to compare the performance of the two methods. In one-dimensional scenarios, the cost function, flat plane, and identified minima are plotted for direct comparison. For two-dimensional functions, the cost surface and flat plane are displayed as 3D surfaces, highlighting the global minimum with a marker.

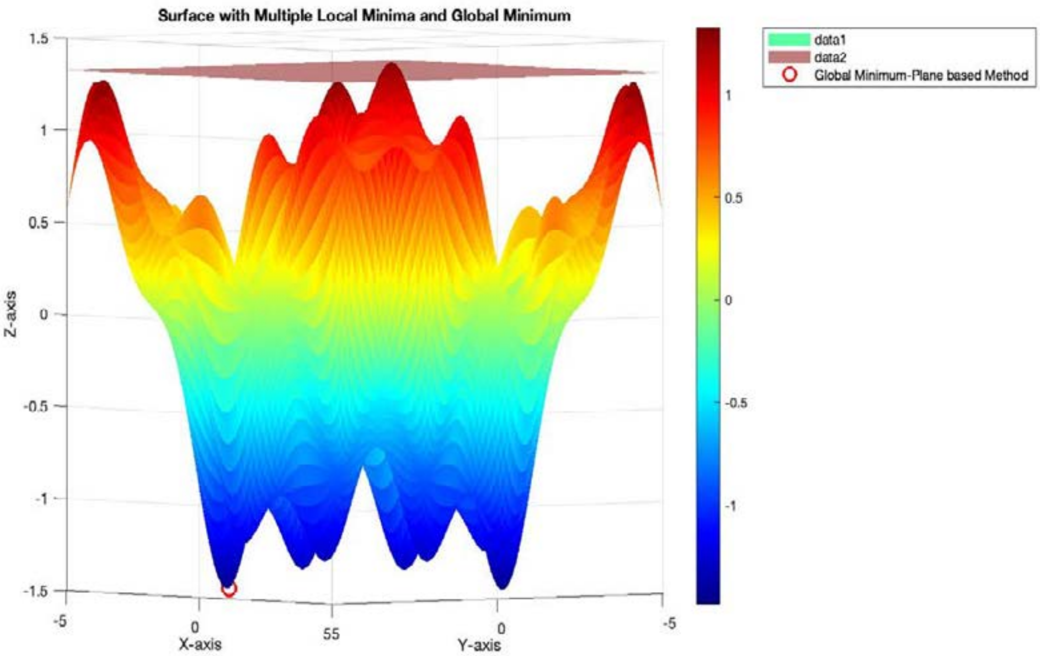

Figure 2: In A Multidimensional Cost Function with Multiple Local Minima and A Global Maximum the Plane-Based Method with Maximum Perpendicular Distance Below the Plane Successfully Found the Global Minimum of the Cost Functions

The proposed methods flat plane-based method was able to robustly locate the global minimum in a nonconvex example of a cost function with multiple local minima and a global minimum. The cost surface and flat plane are displayed as 3D surfaces, highlighting the found global minimum with a marker.

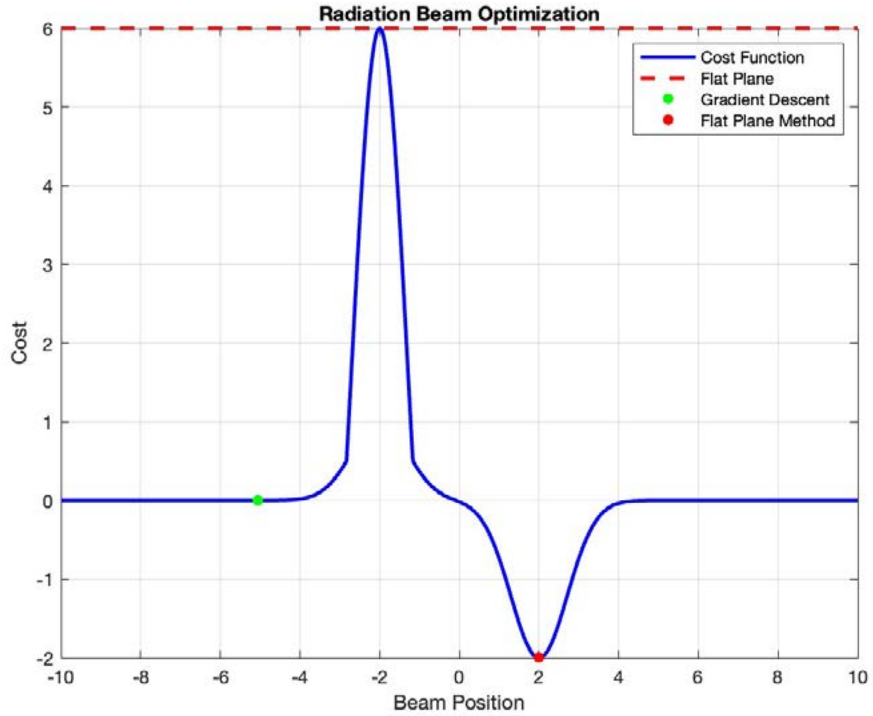

Figure 3: Radiation Beam Optimizations Cost Functions to Minimize the Dose to Surrounding Healthy Tissue and Maximize Dose to Tumor. Where Optimized Minimal Solution From Gradient Descent is Indicated by the Green Dot and the Optimized Minimal Solution From the Proposed Flat Plane with Maximum Perpendicular Distance Method is Indicated by the Red Dot

The medical physics tumor optimization script applies these optimization methods to a clinically relevant problem: optimizing radiation beam placement to minimize exposure to healthy tissues while maximizing the dose to the tumor. The cost function is defined as a combination of weighted healthy tissue and tumor doses, with an additional penalty term for exceeding a predefined healthy tissue dose threshold. The dose distributions for healthy tissues and tumors are modeled as exponential decay functions, simulating Gaussian-like dose profiles centered on their respective regions. In the Radiation Beam Optimization where the proposed flat plane method was compared to a standard gradient descent method. The proposed flat plane method was able to query all points in the cost functions at the same time and find the global minimum of the cost function in an efficient manner. While the gradient descent method query adjacent points close to the chosen initial value and the gradient descent method was unsuccessful finding the global minimum of the cost function.

4. Discussion

In the gradient descent method, the position is iteratively updated by moving in the direction of the negative gradient of the cost function. Gradients are computed numerically by evaluating the function at small perturbations from the current position [1]. The optimization process continues until the position updates fall below a specified tolerance, indicating convergence. This approach is efficient for navigating smooth cost landscapes and identifying minima within the local vicinity.

The flat plane method involves constructing a flat plane at the maximum value of the cost function and calculating the perpendicular distances from this plane to all points below it. The global minimum is identified as the point with the maximum perpendicular distance beneath the plane. This method is particularly effective in scenarios with multiple local minima, as it examines the entire cost function surface without being constrained to local regions. Gradient descent is used to iteratively adjust the beam position, with numerical gradients calculated via small perturbations. Convergence is determined by the magnitude of position updates, ensuring efficient identification of optimal beam placement. Similarly, the flat plane method calculates the distances from a flat plane to the cost function, identifying the global minimum as the point with the largest perpendicular distance. This highlights the importance of exploring different optimization algorithms and understanding how they fit a given problem. One iterative optimization approach is the gradient descent technique, which is just relying on the update in the direction of cost function as local minimum. But its performance highly relies on the initialization point, as well as choice of the learning rate and step size. If the initialization is poor or if we pick ineffective values for the learning rate than we can have slow convergence or inconsistencies, or we may get stuck local minimum and not be able to reach the global minimum [2,3]. It is computationally efficient for high-dimensional problems, but convergence may take more iterations and careful parameter tuning.

In contrast, the flat plane provides robustness in terms of finding the global minimum as it looks at all points in the cost function. The solution strategy involves creating a plane at the highest value of the cost function and measuring the perpendicular distances from this plane down to all the points beneath it. Additionally, the approach is different from gradient descent, which involves convergence to local minima, as there is very little dependence on iterations of points or initialization points. This has a high computational cost though, as the distance has to be calculated for each point in the search space, and its cost grows rapidly with the dimensionality of the optimization problem. While this works well for 1D or 2D, its use for high-dimensional problems is not much as it is computationally difficult. It should be noted however that certain general adaptations can be made to efficiently apply the flat plane method on high-dimensional optimization problems. First, dimensionality reduction methods like principal component analysis (PCA) may be used to project the optimization problem into a lower-dimensional subspace while maintaining most features of the cost function. This would cut down the distance calculations needed. Second, the cost function can be approximated by sampling strategies (e.g., Monte Carlo or Latin hypercube sampling) to find distances at discrete points instead of uniformly over the entire space [4]. Lastly, parallel computing and specifically GPU acceleration, can speed up distance computations, distributing the required calculations on several processors. These approaches enable the flat plane method to be scaled up to high-dimensional optimization problems without excessive computational cost.

5. Conclusion

Notably, the flat plane with maximum perpendicular distance method is highly demanded alternative approach method compared to optimization techniques to accurately search global minima in convex cost function. It showed clear benefits to gradient descent using 1D and 2D cost function simulations and its application to a medical physics problem, such as initialization independence and ability to escape from local minima. Due to the computational cost associated with the algorithm, it cannot be used for high dimensional problems, but there are suggested modifications to the method involving dimensionality reduction, sampling strategies or parallel methods that demonstrate good potential scalability. Not only does this work provide the theory behind the flat plane method, but it also emphasizes its usefulness in solving more challenging optimization problems, enabling future developments in computational optimization and multidisciplinary applications. Potential future work could involve improving the computational efficiency of the flat plane method in high dimensional problems.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work the author(s) used Chat GPT to provide concrete scientific literature that supported the results and for grammar corrections. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgments

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research in this publication was supported by the National GEM Consortium Fellowship and the SciMed GRS provided by the Graduate School, part of the Office of Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison, with funding from the Wisconsin Alumni Research Foundation and the UW-Madison.

Data Availability

All the data used in this manuscript were simulated data. All the scripts used to generate the data could be provided upon request.

References

1. Wang, P., Lei, Y., Ying, Y., & Zhou, D. X. (2024). Differentially private stochastic gradient descent with low-noise. *Neurocomputing, 585,* 127557.
2. Swenson, B., Murray, R., Poor, H. V., & Kar, S. (2022). Distributed stochastic gradient descent: Nonconvexity, nonsmoothness, and convergence to local minima. *Journal of Machine Learning Research, 23*(328), 1-62.
3. Lee, J. D., Simchowitz, M., Jordan, M. I., & Recht, B. (2016, June). Gradient descent only converges to minimizers. *In Conference on learning theory* (pp. 1246-1257). PMLR.
4. Rosenthal, J. S. (2000). Parallel computing and Monte Carlo algorithms. *Far east journal of theoretical statistics, 4*(2), 207-236.