Journal of Statistical and Mathematical Sciences
ISSN 3065-100X
Volume 13 Issue 3
July- September 2025
Impact Factor: 8.99
Enquiry: contact@continentalpub.online
Published by Continental Publication | http://continentalpub.online/index.php/jsms

HIGH-DIMENSIONAL OPTIMIZATION OF NON-SMOOTH CONVEX FUNCTIONS: ALGORITHMS AND THEORY

Dr. Chukwuemeka Obinna Nwosu and Dr. Ifeoma Grace Okafor
Abia State Polytechnic, Aba, Nigeria.

Abstract: *In this study, we address the optimization of convex functions in N-dimensional spaces, a problem with widespread applications in various fields. Convex functions exhibit unique characteristics, such as having a single minimum value X* when they possess a finite minimum and the gradient vanishing at X* when the function is both differentiable and strictly convex.*
To tackle this optimization problem, we explore the use of the descent (steepest) method and Newton's method, two well-established techniques in the field. The core challenge lies in minimizing the non-linear convex function subject to constraints of the form ( ), where i ranges from 1 to n.
We also consider the problem from the perspective of minimizing f over a closed convex subset. To achieve this, we introduce the projection map T, which maps elements in the N-dimensional space to a subset such that the Euclidean norm difference between the two sets is minimized, as expressed by the equation ‖ − ‖ = ‖ − ‖.

Keywords: *convex optimization, descent method, Newton's method, constraint optimization, projection map.*

1. Introduction

We consider the optimization of a convex function in N-space which is a special case of the non-linear optimization problem of minimizing a non-linear function ( ) over the n-dimensional Euclidean space R. The range of applications in which determination of at which ( ) attains its minimum are important is extremely wide.

The convex function is specially shaped so that if it possesses a finite minimum, the minimizing value X*, say, is unique and the gradient of the function vanishes at X* when is differentiable and strictly convex. A number of finite terminating algorithms for obtaining approximate values of X are in literature. We implore the use of descent
(steepest) method and the Newton's method. The basic problem is that of minimizing the non-linear convex function
( ) subject to constraints

( ), = 1, 2, … .

Then one can view the problem as that of minimizing f over a closed convex subset. In other words, Let T be the projection map of onto , that is for , is that elements in such that
‖ − ‖ = ‖ − ‖.
So that the sequence of elements is then defined as follows

= − __ (1)

Finally, the optimizer can be assumed to exist and the problem is to find it with minimum functional evaluation. We try to locate this X*of the non-differentiable convex function by exploiting the connection between a convex function and the accretive operator and central in this formulation is that of optimal experimental design.

2. Constrained Optimization.

An analysis of the multivariable unconstrained non-linear multivariable unconstrained non-linear maximization problems set the stage for the analysis of constrained models. The algorithmic difficulties to be overcome here are present also in the constrained case and the techniques below can be suitably modified when constraints are imposed. However, a constrained problem can often be solved by first converting to an unconstrained problem.

Many of the techniques for solving the general variable non-linear optimization actually employ simple variable optimization in one of the steps for example, a linear function

( ) = +

Has its optimal solutions at the extreme points, end points, If in a closed interval i.e.

( ) = ( ), [,]

To guarantee that solution techniques are valid, we impose certain assumptions.

2.1. Assumptions of Constrained Optimization.

1. For all values of , ( ) , is uniquely defined and finite.
2. For all , ( ) is uniquely defined, finite and continuous.
3. ( ) possess a finite optimum
4. for any possible value of, ( ) , say C, there exist an associated finite number . Such that every

≤ , ( ) ≥ C

2.1.1 The Search for Optimal Solution.

In solving non-linear programming problems, it might appear a bit difficult but there are several fundamental theorems that can be utilized to guide our search even in the face of such difficulties. However, if such conditions as convexities or concavity are met, the characterization of the optimal solution becomes relatively well defined. But we are dealing with bounded continuous functions, by Weierstrass theorem guarantees us that a maximum or minimum will always exist either at a point interior to the boundaries of the feasible solution variable or at the boundaries itself. This is intuitively clear, since a bounded function must always possess maximum or minimum values somewhere within the region of interest. If the function is continuous over the domain of interest, stationary points can be located through the use of differential calculus provided all derivations can be found.

2.2. Steepest Descent Method.

The impossibility in finding the minimum of a function analytically paves way for an iterative method for obtaining an approximate solution to it also the Newton's method though being effective but it is unreliable. Hence we consider the steepest descent approach. Given a function that is differentiable at , the direction of the steepest descent is the vector $-\nabla$ ( ).

( ) = (+) (1.1)

Where u is a unit vector.

$^{|}$( ) = —— + ⋯ …. . +

— + ⋯. . + —

∇ (+) . ——

: . $^{|}$(0) = ∇ ( ). = |∇ ( )|

DOI:https://doi.org/10.5281/zenodo.15913964

Where is the angle between∇ ( ) and U. it follows that ′(0) is minimized when = which yields

$= \frac{-\nabla\ (\ ).}{|\nabla\ (\ )|}\quad (0) = -|\nabla\ (\ )|$

We can therefore reduce the problem of minimizing a function of several variable to a single variable minimization problem, by finding the minimum of ( ) for this choice. ie, we can find the value of , for > 0, that minimizes

( ) = − ∇ ( ) (1.2)

After finding the minimizer , we can set

= − ∇ ( )

and continue the process by searching from in the direction of−∇ ( ) to obtain by minimizing

( ) = (− ∇ ( ))
and so on

This is the method of steepest descent qiven an initial guess . The method computes a sequence of iterates, where

= − ∇ ( ) , = 0, 1,2 … … (1.3)

Where 0 minimizes the function

( ) = − ∇ ( ) (1.4)

Example;
Consider the non-linear minimization problem

Minimize (,) = − + 2 + 2 + (1.5)

Using steepest descent method with the initial point at = (0,0)

—

Solution:

$(\ ) = \begin{bmatrix} \frac{(\ )}{(\ )} & \frac{(\ )}{(\ )\ (\ )} \\ \frac{(\ )}{} & \frac{(\ )}{} \end{bmatrix} = \begin{matrix} 4 & 2 \\ 2 & 2 \end{matrix} = 4 > 0$

Hence convex.

$\nabla\ \_ = \frac{(\ )}{}, \frac{(\ )}{} = (1 + 4\ \ + \ \ . -1 + 2\ \ + 2\ \ )$

$\nabla\ \_^{*} = (1, -1)$
$= (0, 0) - (1, -\ ) = -,$

Journal of Statistical and Mathematical Sciences
Vol. 13 No. 3 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15913964

Substituting in (1.5) we obtain

$$= 1, \quad = -1, \quad = 1$$

$$= (-1, 1) - \ (-1, -1) = -1 + \ , 1 +$$

$$(-1 + \ ) - (1 + \ ) + 2(-1 + \ ) \ + [(-1 + \ )(1 + \ ) + (1 + \ )$$

$$5 \quad - 2 \quad - 1$$

$$\frac{(\ )}{} = 10 \quad - 2 \rightarrow \quad = \frac{1}{5}$$

$$\rightarrow (-1, 1) = \frac{1}{5}(-1, -1)$$

$$= -1 + \frac{1}{5} = -0.8 =$$

So from this we can proceed to get the result in the table below.

Table 1: Results of the minimization problem using the steepest descent method

Iteration			Step size
0	0	0	1
1	-0.8	1.2	0.5
2	-1	1.4	1
3	-0.6	1.8	0.2
4	-0.86	1.34	0.12
5	0.993	1.352	0.3
6	-0.922	1.409	0.367
7	-0.9632	1.4172	0.3170
8	-0.9567	1.4497	0.3527692
9	-0.9722	1.4526	0.2136219
10	-0.9701	1.4541	0.38931
11	-0.0017	1.4905	1.1367029
12	-0.9967	1.4949	0.1952688
13	-0.0002	1.4992	1.1827957
14	-0.9997	1.4995	0.532258
15	-0.9998	1.4998	0.666666
16	-0.9999	1.4997	0.09375

Optimal value -1.0, 1.5

2.3 Nature of the Objective Function

Suppose that the function is now restricted further by adding an assumption about its shape. A general variable function ( ) is defined as convex if the inequality is replaced by $0 < > 1$ so that we have the sufficient conditions for a minimum.

Given assumption (1) to (4) in 2.1 and that ( ) is convex, if each $\frac{}{} = 0$ at a point then (*) is the minimum value for ( ).

Further, if ( ) is strictly convexthen * is unique.

Therefore, local minimum of a convex function is also a global minimum.

Thus if one applies the method at steepest descent using an optimal step size, then the sequence f descent using an optimal step size, then the sequence ( ) decrease the limit to the minimum value of ( )

Journal of Statistical and Mathematical Sciences
Vol. 13 No. 3 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15913964

If the function is strictly convex, the entire sequence converges to the unique optimal solution .

3. Locating the Optimizer of a Non-Differentiable Convex Function in N-Space

A convex function in n-space is defined as; for any two points and0 ≤ ≤ 1 , [(+ (1 −))] ≤ = [(+ (1 −) ( )] .

Where is non-differentiable convex function, a unique minimizing value can be assumed to exist and the problem is to find it with minimum functional evaluation. We try to locate of the differentiable convex function f by exploiting the connection between a convex function and the accretive operator. Central in this formulation is the method of optimal experimental design.

3.1 Accretive Operator's

A mapping T with domain ( ) and range ( ) is accretive if the inequality ⟨ − *, − *⟩ ≥ 0 holds for every , * ( ) , where ⟨⟩ denotes the inner product in

If ≥ this is replaced by > we say that T is strongly accretive. For a convex function, ( ) satisfies:

(+ *) ≤ ( ) + ( )

≥ 0 , ≥ 0 + = 1 *

Minty [10] for every , we can associate the vector * ,such that

( ) − (*) ≥ ⟨ − *, *⟩, (1.6)

So that

(*) − ( ) ≥ ⟨ * − , ⟩ (1.7)

Adding equations 1.6 and 1.7, we have

0 ≥ ⟨ − *, *⟩ + ⟨ * − , ⟩

= ⟨ − *, − *⟩

Thus, we can see that the A associated with the convex function(3.1) is`accretive and that = 0. Again from equation 1.6 we have ( ) − (*) = 0 so that ( ) ≥ (*) Hence is the optimizer of F when * = 0

.

If F differentiable, is identifiable with the gradient of ( ) at .

Let's denote the kernel of A by

= { : = 0}

Then the kernel of the accretive operator A associated with the convexn function turns out to be the optimizer of . Hence the problem of locating the optimizer of is equivalent to that of obtaining the kernel of the accretive operator A.

Chidume(2) showed that given a sequence { }∞ satisfies A if

: = 1 , 0 < < 1 ≥ 1

: ∑∞ = ∞

: ∑∞( ) < ∞

The sequence { }∞ generated by ( ), = − , ≥ 0 converges strongly to the solution of the equation = 0 where A is strongly accretive with error estimates ‖− *‖ = 0 ()

However, the main constraint is that in a given situation, we may not be able to compute the vector AX but only observe it at a point. Thus we employ the method of response surface exploration to estimate it. This method is optimal because it minimizes the Euclidean distance between the true and estimated accretive operator.

3.1.1. Estimating the Accretive Operator.

Let ( ) = (*) = ( )such that equation 3.2becomes

( ) ≥ ⟨ − *, *⟩ (1.8)

DOI:https://doi.org/10.5281/zenodo.15913964

Suppose that the design is chosen in the neighbourhood of, the relation between Y(( ) the vector = * is well represented by the hyperplane.

= − *, * + (1.9)

Where isan observableand it is error used to account for one's inequality to describe − *, * which the so called response surface is.

Let us suppose also that as a result of the experimental design , ,it is possible to construct an estimate of "indirectly"based on the measured values ( ), … … … . . ( ) such that the Euclidean distance between the true accretive operator and the estimated accretive operator is minimized. This is very possible for each observable Y, we associate a positive linear operator P such that

i. If ≥ 0 *h* (( ) ≥ 0

ii. (, ( ) + ( ) = [( ] + [( ] iii. − [( ][( ] − [(= ⟨ (− (, ( ) − (( )⟩

= − (( ) ≤ : =

0 : ≠ :

iv. (= 0

So that ( )= ⟨ − *, *⟩

Let − * =

So

= | * (2)

and

=

is a symmetric matrix.

Thus, when m is non-singular, the unique solution of the equation

(*) = (*), *

Where

(*) = ∑ (( ) − (*) (2.1)

Is

* = (2.2)

Which turns out to be the least square estimate of

Then

(*) = = | * = *

So that

‖ * − *‖ = (*) =

(where is the identity matrix).

is not known and has no influence on the estimation of and on the design used so that without loss of generality we assume = and hence magnitude of the Euclidean distance between the estimated and true accretive operator depends only on the design used.

3.2 Numerical example

Consider the convex function

(,) = − + 2 + + 2

The optimizer is * = (*, *) = (−1, 1.5) let the searching point be and let the design be = (0,0)

Journal of Statistical and Mathematical Sciences
Vol. 13 No. 3 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15913964

and let the design be

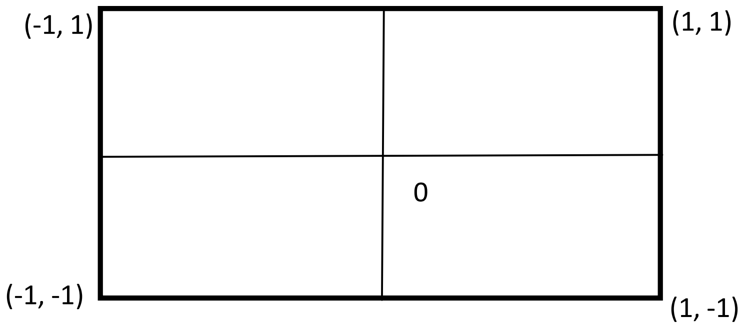
$= 0. \frac{1}{} = 1 \qquad h.$

Where $= (\quad , \quad , \ldots \ldots \ldots . \quad)$.

Design:

Choose the four vertices of a square of a unit radius centered origin.

Figure 1: The design



We can state the problem as

Minimize $\quad , \quad = \quad - \quad + 2 \quad + 2 \quad + \quad , \quad = (\quad , \quad)$

So that if $= (\quad ,\ldots \ldots \ldots . . , \quad)$ is known and $(\quad) = \quad (\quad) - ({}^{*}) = \langle \quad - \quad {}^{*} , \quad {}^{*} \rangle$

Our design point constitutes the following:

$= (1,1), \qquad = (1, -1), \qquad = (-1, 1), \qquad = (-1, -1)$

Then estimate for the accretive operator T denoted by A* is given by

$$(2.3)$$

We denote the sequence $\{ \quad \}^{\infty}$ Iterated by $= \quad - \quad {}_{1}$ With error estimate given as $\| \quad {}^{*}\| = 0 (\quad)$

We see such that so that $\| \quad - \quad \| <$ the sequence $\{ \quad \}^{\infty}$ will converge to the solution of $= 0$ for a finite n.

Let the response vector be

$$= \begin{pmatrix} (\quad) \\ (\quad) \\ (\quad) \\ (\quad) \end{pmatrix}$$

So that

$$\begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix} = \left| -1 \right| \qquad (2.4)$$

In order to estimate A*, we compute = − = 1, 2 … … . . . , = 1, 2 , … … …

From the design so that

=

= ( ) = (2.5)

So that = * = ( ) h (2.6)

From the design

1 1 1 −1 (2.7)

−1 1
−1 − 1
1 1
= 1 1 −1 −1 1 −1 4 0
= −1 1 = 0 4
So that 1 −1 1 −1
−1 −1

* =

Hence, * = 1 −1

0.25 = 0 0 1 0.25 1 1 −1 −1 1 −1 −1 = 5 3 −1 5 = 0.25 0.25 −0.25 −0.25 0.25 −0.25 0.25 − 0,25 5 3 −1 5 = 1 −1

Having obtained A*, we approximate along

Thus = − *, ≥ 0

Where = ___ ≥ 0

For the first iteration we have

= − *, = 1/− = 1 * = 1 −1

The starting point is

0 =

DOI:https://doi.org/10.5281/zenodo.15913964

0

So that

$$\mid = \begin{matrix} -1 \\ 1 \end{matrix}$$

We continue in this manner for the second, third and so on. So the response

$$3\ 1 = \quad 0.75 + 0.25 - 0.25 - 1.75 = \quad -1$$
$$1 \quad 0.75 - 0.25 + 0.25 - 1.75 \quad -1$$
$$7$$

$$\begin{matrix} 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & -0.25 & 0.25 & -0.25 \end{matrix}$$

$$\therefore\ ^* = \begin{matrix} -1 \\ -1 \end{matrix}$$

$$(\ ) = (\ ,\ ) - (\ ^{\mid},\ ^{\mid})$$

Which we summarized in a column vector as

$$3 = 1\quad 1\quad 7$$

The blue A* for the accretive operator / is then

$$\mid = \frac{1}{2}$$

$$\therefore = \mid - \mid ^*$$

$$= \begin{matrix} -1 \\ -1 \end{matrix} - \frac{1}{2} \begin{matrix} -1 \\ -1 \end{matrix} = \begin{matrix} -0.5 \\ 1.5 \end{matrix}$$

$$\therefore \quad = \begin{matrix} -0.5 \\ 1.5 \end{matrix}$$

Hence we have the following result in the table

$$= \frac{1}{+1}$$

Table 2: Result of the design

0	0

-1	1
-0.5	1.5
-1.16	1.17
-0.9686	1.335
-0.9277	1.3906
-0.9393	1.40306
-0.9461	1.4136
-0.9514	1.4217
-0.9557	1.4283
-0.9591	1.4338
-0.9619	1.4384
-0.9643	1.4423
-0.9664	1.4457
-0.9682	1.4487
-0.9698	1.4513
-0.9713	1.4536
-0.9726	1.4557
-0.9738	1.4576
-0.9749	1.4593
-0.9759	1.4605

The performance of the steepest descent for the estimated accretive operator relative to the steepest descent
method is summarized in the table below.

Table 3: Steepest descent method for estimated accretive operator

iterations	Steepest descent method	Steepest descent method	Steepest descent for method estimated accretive operator	Steepest for descent method estimated accretive operator
0	0	0	0	0
1	-0.8	1.2	-1	1
2	-1.0	1.4	-0.5	1.5
3	-0.6	1.8	-1.16	1.17
4	-0.86	1.34	-0.9686	1.335
5	-0.933	1.352	-0.9277	1.3906
6	-0.922	1.409	-0.9393	1.40306
7	-0.9632	1.4172	-0.9461	1.4136
8	-0.9567	1.4496	-0.9515	1.4217
9	-0.9722	1.4526	-0.9557	1.4283
10	-0.9701	1.4521	-0.9591	1.4338
11	-1.0017	1.4905	-0.9619	1.4348
12	-0.9967	1.4949	-0.9643	1.4423
13	-1.0002	1.4992	-0.9664	1.4457
14	-0.9997	1.4995	-0.9682	1.4487
15	-09998	1.4998	-0.9698	1.4513
16	-0.9999	1.4997	-0.9713	1.4536

17			-0.9726	1.4557
18			-0.9738	1.4576
19			-0.9749	1.4593
20			-0.9759	1.4609

4. Conclusion

The steepest descent method for the estimated accretive operator solves the minimization problem with no reference to the derivative of the function. However, if the design are optimized the formulation of the steepest descent for the estimated accretive operator is the generalization of the ordinary steepest descent method.

References

Achi Gods will Uche, "Locating The Optimizer Of A Non Differentiable Convex Function In N Space". Unpublished project, Department of Mathematics, school of Physical Science, Abia State University, Uturu, July 1995.

Chidume C.E (1987) "Iterative approximation of fixed points of Lipschitzian strictly pseudo-contrative mappings" Proc. Amer. Math.sol 99, No 2, 283-288.

Chidume C.E E "Steepest decent approximations for accretive operator equations" ICTP, Trieste preprint IC/93/94.

Claudeio Marales. Nonlinear Equations Involving m-Accretive Operators. Journal OF Mathematica Analysis and Applications 97, 329-336 (1983).L

Dimtri P. Bertsekas, Sanjoy K. Mitter (1973). A Descent Numerical Method For Optimization Problems With Nondifferentiable Cost Functionals. SIAM J. CONTROL VOL 11, NO4.

Everette, H. "Generalised Lagrangian Multipliers for Solving Problems of Optimal Allocations of Resources, Operation Research II, 399 (1969).

Harvey M. Wagner "Principles of operations research with application to managerial decisions" 2nd Ed. New Delli – 110001 525-587.

Jim Lambers, MAT419/519 Lecture 10 Note. Summer Session 2011-2012.

Lee, Harvey, "An experimental study of solving convex quadratic programming problems". Unpublished M.S. project, School of Industrial Engineering, PurdueUniversity. W. Lafayette, August, 1975.

Minty, G.J. (1964) "On the monotonicity of gradient of convex function" Pacific J. of maths vol Xiv 234-247.

Pazman A. (1987) "Fundamentals of optimal experimental designs" O. Reidel publishing company.

Robert M. Freund, (2014) "The Steepest Descent Algorithm for Unconstrained Optimization and a Bisection Linesearch Method" February, 2014.