The Pure Sub-gradient Method for Constraint Optimization problems

Iliyasu,Abubakar Hamma'adama

[{abubakariliyasu229@gmail.com}](mailto:abubakariliyasu229@gmail.com)

Modibbo Adama University (MAU) Yola, Adamawa¹, No. 2, Opp. Central Mosque Bazza, Michika Adamawa State, Nigeria²

Abstract

In this paper, pure sub-gradient method has been investigated and a comprehensive study of high-level research knowledge about the concept of sub-gradient method have been developed. This method arises as a result of in-efficiency of direct method and conjugate gradient method to handle large sparse optimization problems, and therefore, sub-gradient optimization is one of the most important topics in the field of optimization. The method has been applied effectively in solving concave and convex optimization problems using programing system code language, as in the case of large-scale practical problems in the maximization of linear and nonlinear integer programming problem. This was achieved through the use of its properties and algorithms to taste the efficiency for convergence in finding an optimal solution to optimization problems. Hence, sub-gradient methods can be employed to effectively find solutions to large sparse optimization problems which are too large for direct methods or conjugate gradient methods to handle.

Keywords: Constraint Optimization. Concave function. Convex function. Pure sub-gradient method. Linear and nonlinear

methods. Iterative algorithms.

1 Introduction

This research paper is targeted on the concept of pure sub-gradient optimization method for the solution of linear and nonlinear constraint optimization problems, through the use of programing system code language. This study captured two numerical problems on which we used the methods of pure sub-gradient in finding solutions to the problems.

2 Optimization

This is a process of iterative procedures for finding optimal solution to an optimization problem.

2.1 Convex set:

Let the set $S \subset \mathbb{R}^n$. Then we said that S is a convex set if, for any $x_1, x_2 \in S$, we have $\alpha x_1 +$ $(1 - \alpha)x_2 \in S$, $\forall \alpha \in [0, 1]$, In geometry, this definition shows that for any two points $x_1, x_2 \in S$, the line segment joining x_1 and x_2 is completely contained in S. It also states that S is a path connected by segmented points, i.e., two arbitrary points in S can be linked by a continuous path. This can also be shown by induction, that the set $S \subset \mathbb{R}^n$ is convex if and only if, for any $x_1, x_2, ..., x_m \in S$, there exist

$$
y = \sum_{i=1}^{m} \alpha_i x_i \in S,
$$
\n(1.1)

which is a linear combination of *vectors.* Where $\sum_{i=1}^{m} \alpha_i = 1$, $\alpha i \geq 0$, $i = 1, \dots, m$. Therefore, $x = \alpha x_1$, $+ (1 - \alpha)x_2$) where $\alpha \in [0, 1]$, is called a convex combination of x_1 and x_2 . While $\sum_{i=1}^{m} \alpha_i x_i$ is called a convex combination of x_1, x_2, \dots, x_m , where $\sum_{i=1}^m \alpha_i = 1, \alpha_i \geq 0, i = 1, \dots, m$. Diagrammatically, these sets can be represented as:

2.2 Concave set

Let $m, v: [0, S] \rightarrow \Re$ be two Lebasque integrable, monotone functions, say m decreasing and v increasing and set:

$$
M(s) := \int_0^s m(\sigma) d\sigma \text{ and } V(s) := \int_0^s v(\sigma) d\sigma. \tag{1.2}
$$

Obviously M is concave and V is convex. Furthermore, a set is said to be concave if and only if the complement of the set is convex.

Example 1.1

The hyper plane $H = \{x \in \mathbb{R}^n | P^T x = \alpha\}$ is a convex set, where a nonzero vector $P \in \mathbb{R}^n$ is referred to as the normal vector to the hyper plane and α is a scalar.

In fact, for any $x_1, x_2 \in H$ and each $\theta \in [0, 1]$,

$$
P^T \left[\theta x_1 + (1 - \theta) x_2 \right] = \alpha, \tag{1.3}
$$

then,

 $x_1 + (1 - \theta)x_2 \in H$.

Note that, in the hyper plane $H = \{x \in \mathbb{R}^n \mid P^T x = \alpha\}$ if $\alpha = 0$, then it can be reduced to a subspace of vectors that are orthogonal to P .

Similarly, the closed half spaces

$$
H^- = \{x \in \mathbb{R}^n \mid P^T x \le \beta\} \text{ and } H^+ = \{x \in \mathbb{R}^n \mid P^T x \ge \beta\}
$$
 (1.4)

are closed convex sets. The open half spaces

 $(\dot{H})^{-} = \{x \in \mathbb{R}^{n} \mid P^{T}x < \beta\}$ and $(\dot{H})^{+} = \{x \in \mathbb{R}^{n} \mid P^{T}x > \beta\}$ (1.5) are open convex sets.

Example 1.2

The ray $S = \{x \in \mathbb{R}^n | x = x_0 + \lambda d, \lambda \ge 0\}$ is a convex set, where $d \in \mathbb{R}^n$ nonzero vector and $x_0 \in \mathbb{R}^n$ is a fixed point. In fact, for any $x_1, x_2 \in S$ and each $\lambda \in [0, 1]$, we have

$$
x_1 = x_0 + \lambda_1 d, \quad x_2 = x_0 + \lambda_2 d,
$$

where $\lambda_1, \lambda_2 \in [0, 1].$ (1.6)

Hence,

$$
\begin{aligned} \lambda x_1 + (1 - \lambda)x_2 &= \lambda (x_0 + \lambda_1 d) + (1 - \lambda)(x_0 + \lambda_2 d) \\ &= x_0 + [\lambda \lambda_1 + (1 - \lambda)\lambda_2]d. \end{aligned} \tag{1.7}
$$

Since

 $\lambda \lambda_1 + (1 - \lambda)\lambda_2 \geq 0$, then $\lambda x_1 + (1 - \lambda)x_2 \in S$.

The finite intersection of closed half spaces is

 $S = \{x \in \mathbb{R}^n | P_i^T x \le \beta_i, i = 1, \dots, m\}$ (1.8)

which is called a polyhedral set, where P_i are nonzero vector and β_i is a scalar. The polyhedral is a convex set. Since equality can be represented by two inequalities, the following sets are examples of polyhedral sets:

$$
S = \{x \in \mathbb{R}^n \mid Ax = b, x \ge 0\},\tag{1.9}
$$

$$
S = \{x \in \mathbb{R}^n \mid Ax \ge 0, x \ge 0\}.
$$
 (1.10)

The intersection of two convex sets is convex, the algebraic sum of two convex sets is convex, the interior of a convex set is convex, and the closure of a convex set is convex.

Theorem 1.1

Let S_1 and S_2 be two convex sets in \mathbb{R}^n . Then,

i. $S_1 \cap S_2$ is convex;

ii. $S_1 \pm S_2 = \{x_1 \pm x_2 \mid x_1 \in S_1, x_2 \in S_2\}$ is convex.

Proof: Following the definition of convex set above, it is obvious that condition (i) is true for any intersection of two convex sets to be convex also. Similarly, it holds for condition (ii), that is for the sum or difference of two convex sets their union or difference must also be convex.

Theorem 1.2

Let $S \subset \mathbb{R}^n$ be a convex set. Then

- i. The interior *int*, S of S is a convex set:
- ii. The closure \hat{S} of S is a convex set.

Proof: (i)

Let x and x' be in int. S and $x'' = \alpha x + (1 - \alpha) x'$, $\alpha \in (0, 1)$. Choose $\delta > 0$ such that $B(x', \delta) \subset S$, where $B(x', \delta)$ is the δ -neighborhood of x'. It is easy to see that $\frac{x''-x}{x'-x} = 1 - \alpha$. We know that $B(x'', (1 - \alpha)\delta)$ is just the set $\alpha x + (1 - \alpha)B(x', \delta)$ which is in S. Therefore $B(x'', (1 - \alpha)\delta) \subset S$ which shows that $x'' \in intS$.

Proof: (ii)

Take $x, x' \in \hat{S}$. Select two sequences in S as $\{x_k\}$ and $\{x'_{k}\}$ converging to x and x' respectively. Then, for $\alpha \in [0, 1]$, we have

$$
[\alpha x_k + (1 - \alpha)x'_k] - [\alpha x + (1 - \alpha)x']
$$

= $\alpha(x_k - x) + (1 - \alpha)(x'_k - x')$
 $\le \alpha x_k - x + (1 - \alpha)x'_k - x'.$ (1.11)

Taking the limit of both sides yields

lim $|\left[\alpha x k + (1 - \alpha) x' k\right] - \left[\alpha x + (1 - \alpha) x'\right]\right| = 0,$ (1.12) k→∞

which shows that $\alpha x + (1 - \alpha) x' \in \hat{S}$.

2.3 Concave function

Let f be a function on a convex set S and $S \subset \mathbb{R}^n \to \mathbb{R}$. Then f is said to be concave and closed if for any $x_1, x_2 \in S$ and $\alpha \in [0,1]$, we have that

$$
f((1 - \alpha)x_1 + \alpha x_2) \ge (1 - \alpha)f(x_1) + \alpha f(x_2), \text{ for } x_1 \ne x_2
$$
 (1.13)
We say that, f is strictly concave if

$$
f((1 - \alpha)x_1 + \alpha x_2) > (1 - \alpha)f(x_1) + \alpha f(x_2).
$$
 (1.14)

2.4 Convex function

Let f be a function on a convex set S and $S \subset \mathbb{R}^n \to \mathbb{R}$. Then f is said to be concave if for any $x_1, x_2 \in$ S and $\alpha \in (0.1)$, we have that

$$
f(\alpha x_1 + (1 - \alpha)x_2) \le \alpha f(x_1) + (1 - \alpha)f(x_2), \text{ for } x_1 \ne x_2
$$
 (1.15)
We say that, f is strictly convex if

$$
f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2). \tag{1.16}
$$

3 Differentiability of a function

A function $f(u)$ is said to be differentiable at the point *u*, if the derivative $f'(u)$ exists at every point u in its domain. And also, f is *continuously differentiable* if its derivatives exist continuously over its domain.

3.1 A Close function

A function $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to be closed if it is lower semi-continuous everywhere, or if its epigraph is closed, or if its level sets are also closed. Consequently, the indicator function I_s is closed if and only if S is closed. Also, epi $I_s = S \times \mathbb{R}^+$. The support function σ_s is closed as well.

3.2 Sub-differential

The Sub-differential of a function f at x' , is the set of all sub-gradients of f at x' which is given by $\partial f(x') = \{s : f(x') + s(x - x') \ge f(x) \forall x \in \mathbb{R}^n\}.$ (1.17) Furthermore, if $\delta f(x)$ is non-empty, then f is said to be sub-differentiable at a point x', which implies that a concave function is sub-differentiable at every point in its domain, and the sub-differential is a non-empty convex, closed and bounded set. Note that, a concave function is not always differentiable at all points in its domain but closed.

3.3 Sub-gradient

Let $f: \mathbb{R}^n \to \mathbb{R}$ be concave. The vector $g \in \mathbb{R}^n$ is called a Sub-gradient of f at $x' \in \mathbb{R}^n$ if $f(x') + g(x - x') \ge f(x), \forall x \in \mathbb{R}^n$ (1.18)

Similarly,

 $f(x') - f(x) \ge \langle g, x - x' \rangle$, $\forall x \in \mathbb{R}^n$ (1.19) if $f(x') = g$ and $f(x') = 0$, then it implies that $g = 0$ and $(0, g) \in \partial f(x') \forall x \in \mathbb{R}^n$.

3.4 Important Results

Let $u = (u_1, u_2, l, u_n)$ and $v = (v_1, v_2, ..., v_n)$ be two vector spaces in \mathbb{R}^n . Then we say that,

- i. $u_i v_i := P_n \sum_{i=1}^m u_i v_i$, is a linear combination of independent vectors, whether they are raw or column vectors.
- ii. $u \geq 0$ means $u_i \geq 0$ for each $i = 1, 2, \ldots n$.
- iii. $||u||: = \sqrt{uu}$ the Euclidean norm.
- iv. $P_{\Omega}(u) = argmin z \in \Omega \{||z u||\}$ is the Euclidean projection of u onto a closed convex set Ω ; i.e. the point in closed set Ω to u .
- v. int(Ω) = { $x \in \Omega$: $||y x|| < \varepsilon \implies y \in \Omega$, for some $\varepsilon > 0$ } is the set of interior points of Ω , where $\Omega \subseteq \mathbb{R}^n$.
- vi. Ω^* is the set of optimal solutions of an optimization problem whose feasible set is Ω .
- vii. ϕ^* Optimal objective value of an optimization problem whose objective function is $\phi(.)$.

3.5 Constraints

These are logical criteria or conditions that a solution of an optimization problem must satisfy.

4 Iterative Algorithms for Sub-gradient Methods

In this chapter, we considered two iteration procedures as methodologies used for understanding and finding of solutions to both linear and nonlinear optimization problems. These methods are:

4.1 The Pure Sub-gradient Method Algorithm

Consider the integer programing (IP) problem of the form

(IP) $\max{\{\phi(u): u \in \Omega\}},$

using the following generic procedure:

- i. Choose an initial point $u^0 \in \Omega$.
- ii. Construct a sequence of points $\{u^k\} \in \Omega$, which eventually converges to an optimal solution following the rule:

$$
u^{k+1} = P_{\Omega} (u^k + \lambda_k s^k) \text{ for } k = 0, 1, 2, ...
$$
 (2.1)

Step 1: Set u^0 for $k=0$ $u^{k+1} = P_{\Omega} (u^k + \lambda_k s^k), k = 0, 1, 2, ...$ $u^1 = P_{\Omega} (u^0 + \lambda_0 s^0)$ $k = 0, 1, 2, ...$ **Step 2:** $u^{k+1} = P_{\Omega} (u^k + \lambda_k s^k), k = 1, 2, ...$ $u^2 = P_{\Omega}(u^1 + \lambda_1 s^1), k = 1, 2, ...$

Step 3: repeat the process continuously until the solution converges to an optimal solution. Where s^k is a sub-gradient of the concave function ϕ at a point u^k , which is been determined at each iterate point until some certain stopping condition, $\lambda_k > 0$ is an appropriately chosen step length and P_Ω(.) is the Euclidean projection on the feasible set $Ω$.

4.2 Numerical Solutions

In this chapter, we will find and discuss of two numerical examples, using the iterative algorithms of the stated methods and compare the results of the sub-gradient methods to see their efficiency for finding optimum solutions to optimization problems.

Example 2.1: Let

$$
f(x) = \max_{x \in \mathfrak{R}} \left\{ 2x, \ x + 2, \frac{5}{3}x + 5 \right\} \tag{3.1}
$$

Then f is a piecewise linear concave function given by

$$
f(x) = \begin{cases} 2x & x \le 1 \\ x + 2 & 1 \le x \le 3 \\ \frac{5}{3}x + 5 & x \ge 3 \end{cases}
$$
 (3.2)

f is differentiable at every point $\bar{x} \in \mathcal{R} \setminus [1,3]$. Hence, for any $\bar{x} \notin [1,3]$ the sub-gradient $s(\bar{x})$ of f at \bar{x} is given by $s(\bar{x}) = f'(x)$ That is,

$$
s(\bar{x}) = \begin{cases} 2, & \bar{x} < 1 \\ 1, & 1 < \bar{x} < 3 \\ \frac{5}{3}, & \bar{x} > 3 \end{cases}
$$
(3.3)

However, at $\bar{x} = 1$ both $S_1 = 3$ and $S_2 = 1$ are sub-gradients of f. Moreover, any convex combination of S₁ and S₂ is also a sub-gradient of f at $\bar{x} = 1$. Similarly, both S₂ = 1 and S₃ = $\frac{5}{3}$ $\frac{3}{3}$ as well as any of their convex combinations are the sub-gradients of f at $\bar{x} = 3$.

Hence, the pure sub-gradient procedure is best applied to solve these type of non-smooth concave problems, because it compute the sub-gradient vector say $\bar{s} \in \partial \phi(x)$ at each iterate point say $\bar{x} \in \Omega$ by fulfilling that $\Omega = \mathbb{R}^m_+$ for any linear concave but nondifferentiable function and maintaining the search direction.

Example 2.2

Let consider the non-linear constraint optimization problem $\min_{x,y\in\Re} f(x,y)$ defined by $f(x, y) = (1-x)^2 + 100(y - x^2)^2$, $\forall x, y \in (1, 2)$,

s. t. $x^2 + y^2 \le 2$.

The computation of the results for optimum solutions of the problem, we take the range of the variables as: $x \in (1, 2)$, $y \in (1, 2)$.

References

- [1] Bertsekas, D. P. (2015). Convex optimization algorithms (pp. 54–104). Athena Scientific. Retrieved from [www.athenasc.com,](http://www.athenasc.com/) Nashua, USA.
- [2] Caprara, A., Fischetti, M., & Toth, P. (1999). A heuristic method for the set covering problem. Operations Research, 47(5), 730–743.<https://doi.org/10.1287/opre.47.5.730>
- [3] Geoffrion, A. M. (1974). Lagrangian relaxation for integer programming. Mathematical Programming Study, 2, 82–114.
- [4] Guta, B. (2003). Subgradient optimization methods in integer programming with an application to a radiation therapy problem. VDM Verlag Dr. Müller.
- [5] Sun, W., & Yuan, Y. (2006). Optimization Theory and Methods, Nonlinear Programming. Springer Science + Business Media, New York, NY, USA.