A SUBGRADIENT METHODS FOR NON-SMOOTH VECTOR OPTIMIZATION PROBLEMS

Arunima Kumari
Department of Mathematics
Bhagwan Parshuram Institute of Technology
Guru Govind Singh Indraprastha University
Rohini, Delhi 110089, INDIA

Abstract: Vector optimization problems are a significant extension of scalar optimization and have wide range of application in various fields of economics, decision theory, game theory, information theory and optimal control theory. In this paper, unlike general subgradient methods, bundle methods and gradient sampling methods used to solve nonsmooth vector optimization problems which include scalarization approach, a subgradient method without usual scalarization approached is proposed for minimizing a non-differentiable convex function which works directly with vector-valued function. A general sub-gradient method for non-smooth convex optimization that includes regularization and interior point variants of Newton’s Method are proposed. This algorithm builds a sequence of efficient points at the interior of epigraph of objective function which satisfies KKT conditions. In this paper, under the suitable conditions it is proved that the sequence generated by algorithm converges to \( \epsilon \)-efficient point.

AMS Subject Classification: 90C29, 90C30
Key Words: subgradient method, nonsmooth optimization, vector optimization, \( \epsilon \)-efficient points
Vector optimization problems are a significant extension of scalar optimization and have many real life applications. Which also referred multiobjective optimization problems, deals with the problem where several objective functions have to be simultaneously minimized or maximized. For nontrivial problems, no single point minimizes or maximizes all given objective at a time, and so the concept of maximum point or minimum point is to be replaced by the concept of Pareto Optima or efficient solution. Multiobjective optimization problem are traditionally solved by two different approaches: scalarization and ordering techniques [22, 29]. The scalarization approach enables the computation of efficient (Pareto) or weakly efficient solutions by formulating single objective optimization problems (with or without constraints) and choosing good parameters in advance. Other scalarization techniques that are free of establishing parameters can compute a discrete approximation of the optimal set. The solution strategies that use ordering techniques require a prior ranking specification for the objectives. For solving smooth type multiobjective optimization problem, descent, steepest descent and Newton methods, which neither require parameter nor ordering information, are proposed by several authors ([11, 12, 25]). Unlike to smooth vector optimization problem, non-smooth vector optimization deals with the problems where the functions involved are not continuously differentiable. The gradient does not exist implying that the function may have kinks or corner points and therefore the functions cannot be maximized or minimized by general proposed analytical process. The Feasible Direction Interior Point Algorithm and the projected subgradient methods for non-smooth vector optimization are proposed by several authors [16, 17, 20, 21]. But most of the existing numerical methods for non-smooth vector optimization problems are either based on subgradient using scalarization approach and space dilation type algorithms [20, 22, 29]. In this paper we deal with subgradient method.

Subgradient method was developed in 1962 by N.Z. Shor [36] and used by him for solving large-scale transportation problem of Linear Programming Problems. His work extensively expanded both in theory and in practice by several authors [8, 12, 13, 14, 15, 16, 18, 23, 26]. Held and Karp [19], unaware of the work of Shor, develop a method for the travelling salesman problem that uses subgradient optimization to compute a bound in a Lagrangean Relaxation scheme. This seminal contribution also led to others, see Fisher [10]. Research work in this direction led by Allen et al. [2], Ermoliev [8], Ermoliev and Shor [9], Goffine [13, 14], Goffine and Kiwiel [15] and Polyak [31] gave a precise statement of the method and its convergence theorems which may be taken in to account
as the culmination of the first stage in developing subgradient techniques. All of their work in this field are well reported in their book, Refer Shor [37, 38, 39] and Polyak [34] as well as in the second book by Shor [40]. Also refer the book by Nesterov [27].

The outline of this article is as follows. In section 2 first we briefly describe our notations and basic definitions. In section 3 we presented subgradient method for vector optimization and state some basic preliminaries materials. The convergence of standard derivative free methods like Powel’s method and genetic algorithms has been proved only for smooth functions in section 4 we describe the convergence analysis of the algorithm for non-smooth function.

2. Preliminaries and Basic Definitions

Let $F$ be a non-empty, closed and convex subset of $\mathbb{R}^n$. The inner product in $\mathbb{R}^n$ is denoted by $\langle \cdot , \cdot \rangle$ and the norm generated by the inner product by $\| \cdot \|$. Given a proper, closed, convex and pointed cone $K$ of $\mathbb{R}^m$, we consider the Partial Order ‘$\leq$’ defined as $x \leq y$ if and only if $y - x \in K$ ($y - x \in \text{int}(K)$). In similar way, we define the Partial Order ‘$\geq$’. Throughout in this article, the zero vector $(0, 0, 0, 0, \ldots)$ in $\mathbb{R}^n$ is denoted by $0$. The non-negative orthant in $\mathbb{R}^n$ is denoted by $\mathbb{R}^+_n = \{x \in \mathbb{R}^n; x \geq 0\}$.

Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. In this paper we are interested in the following problem

\[
\begin{align*}
\text{(P)} & \quad \text{Min } f(x) = (f_1(x), f_2(x), f_3(x), \ldots, f_m(x)) \\
& \text{s.t. } x \in F
\end{align*}
\]  

**Definition 2.1** (Pareto Optimum or Efficient). A point $x_0 \in F$ is called efficient solution of (P) (or Pareto Optimal) if there is no $x \in F$ such that $f(x) \leq f(x_0)$ and $f(x) \neq f(x_0)$ in $\mathbb{R}^m$.

**Definition 2.2** ($\epsilon$-Efficient Solution). Let $\epsilon$ be an element of $\mathbb{R}^n_+$. A point $x_0 \in F$ is said to be an $\epsilon$-efficient solution for (P) if there is no $x \in F$ such that

\[f(x) \leq f(x_0) - \epsilon \text{ and } f(x) \neq f(x_0) - \epsilon.\]

**Solution set.** We denote the solution set of our problem by $F^*$ and define as

\[F^* = \{x : \exists \hat{x} \in F \text{ such that } f(\hat{x}) \leq f(x^*) - \epsilon\} \]
When \( m = 1 \) and \( K = R_+ \), the problem will be called a scalar valued optimization problem.

**Definition 2.3.** The positive Polar cone of \( K \), denoted by \( K^* \) is given by
\[
K^* = \{ y \in R^m : y^T x \geq 0, \text{ for all } x \in K \}
\]

**Definition 2.4.** A function \( f : R^n \to R^m \) is \( K \)-convex, i.e.
\[
f(\lambda x + (1 - \lambda)y) \leq_K \lambda f(x) + (1 - \lambda)f(y) \text{ for all } x, y \in R^n \text{ and } \lambda \in [0,1].
\]

**Definition 2.5.** A vector \( g \in R^n \) is said to be subgradient of a scalar-valued function \( f : R^n \to R \) at \( x \in \text{dom}(f) \) if for all \( z \in \text{dom}(f) \),
\[
f(z) \geq f(x) + \langle g, z - x \rangle.
\]
Geometrically, \( g \) is subgradient of \( f \) at \( x \) if \((g, -1)\) support epigraph at \((x, f(x))\). A function \( f \) is called subdifferentiable at \( x \) if there exists at least one subgradient at \( x \). The set subgradients of \( f \) at \( x \) is called the subdifferential of \( f \) at \( x \) and is denoted by \( \partial f(x) \). i.e.
\[
\partial f(x) = \{ g \in R^n : f(x) + g^T(z - x) \leq f(z), \text{ for all } z \in \text{dom}(f) \}
\]

At differentiable points we have unique subgradient and that is gradient of given function at that point but at non-differentiable point there are infinite number of subgradients i.e. there exists infinite number of supporting hyper plane at that point.

**Example 2.6.** Consider \( f(x) = |x| \) for \( x < 0 \), the subgradient is unique and \( \partial f = \{-1\} \). Similarly for \( x > 0 \), \( \partial f = \{1\} \). That means there exist unique gradient. At \( x = 0 \), \(|z| \geq g(z - 0) = gz \), or, \( gz \leq |z| \) for all \( z \). which satisfied if and only if \( g \in [-1,1] \).

**Example 2.7.** Let \( f(x) = \text{Min}[f_1(x), f_2(x)] \), where \( f_1 \) and \( f_2 \) are defined below:
\[
\begin{align*}
f_1(x) &= 4 - |x|, \quad x \in R \\
f_2(x) &= 4 - (x - 2)^2, \quad x \in R.
\end{align*}
\]
Since \( f_2(x) \geq f_1(x) \) for \( 1 \leq x \leq 4 \), \( f \) can be represented as follows:
\[
f(x) = \begin{cases} 
4 - x, & 1 \leq x \leq 4 \\
4 - (x - 2)^2 & \text{otherwise}.
\end{cases}
\]
In the open interval \((1,4)\), the subgradient \( g = -1 \) and it is unique. Similarly if \( x < 1 \) or \( x > 4 \), then \( g = -2(x - 2) \) is also a unique subgradient of \( f \).
But at the point \( x = 1 \) and \( x = 4 \), the subgradients are not unique because many supporting hyperplanes exist. At \( x = 1 \) any \( g \) in the interval \([-1, 2]\) is a subgradient of \( f \). Similarly any \( g \) in the interval \([-4, -1]\) is a subgradient of \( f \) at \( x = 4 \).

**Definition 2.8** (Subgradient of Vector Valued Function). Note that for \( z \in K^* \setminus \{0\} \), we have \( \varphi : R^n \to R \) such that \( \varphi(x) = z^T f(x) \), where \( f : R^n \to R^m \), is a convex scalar function, then \( g \in \partial \varphi \) if and only if

\[
\varphi(\hat{x}) \geq \varphi(x) + g^T(\hat{x} - x) \quad \text{for all } \hat{x} \in R^n.
\]

It follows that there exists \( U \in R^{m \times n} \) such that \( g = U^T z \), and therefore

\[
\varphi(\hat{x}) \geq \varphi(x) + g^T(\hat{x} - x).
\]

implies

\[
z^T f(\hat{x}) \geq z^T f(x) + z^T U(\hat{x} - x),
\]

since \( z \) is an arbitrary vector in \( K^* \setminus \{0\} \), we get

\[
f(\hat{x}) \geq f(x) + U(\hat{x} - x) \quad (2.2)
\]

For every \( x, \hat{x} \in R^n \). Therefore we consider \( \partial f : R^n \to R^{m \times m} \) defined as

\[
\partial f = \{U \in R^{m \times m} : f(\hat{x}) \geq f(x) + U(\hat{x} - x) \text{ for all } \hat{x} \in R^{m \times n}\}.
\]

In analog with the scalar case we will call this set the subdifferential of vector valued function \( f \) and its elements subgradients of \( f \) at \( x \). If we consider \( K = R^n_+ \) and all the components of \( f \) i.e. \( f_i \) are convex in the usual sense then any matrix \( U \in R^{m \times n} \) with the property that its \( i \)th line is a subgradient of \( f_i \) at \( x \) for \( i = 1, 2, 3, \ldots, m \), belongs to \( \partial f(x) \).

### 3. A Sub-gradient Method for Vector Optimization

Our goal is to minimize \( f : R^n \to R^m \) which is component wise \( K \)-convex. To do this, the subgradient method uses the simple iteration to update \( x_{i+1} \) as

\[
x_{i+1} = x_i - t_i g_i \quad (3.1)
\]

Where \( x_i \) is the \( i \)th iterate, \( g_i \) is any subgradient of \( f_i \) and \( t_i \) is the step size which generate a sequence that satisfy

\[
t_i \geq 0; \quad \sum_{i=1}^{\infty} t_i^2 < \infty, \quad \sum_{i=1}^{\infty} t_i = \infty. \quad (3.2)
\]
This rule has been considered by several authors for similar methods [1, 6, 34, 35].

**Algorithm 1.**

**Initialization:** Take $x_0 \in F$.

**Iterative Step:** Given $x_i \in F$, take $U_i \in \partial f_i$ and compute

$$g_i = \arg\min_{w \in F_i} \left\{ \frac{1}{2} \|w\|^2 + \frac{t_i}{\mu_i} \max_{y \in K^*} \{ y^T U_i w \} \right\}, \quad (3.3)$$

with

$$F_i = F - x_i \quad (3.4)$$

and $t_i$ satisfy (3.2) and

$$\mu_i = \max_{y \in K^*} \{ \|U_i\|^T y \}. \quad (3.5)$$

**Stopping Criterion.** If $g_i = 0$ or $x_i \in F^*$ then stop, otherwise compute

$$x_{i+1} = x_i + g_i. \quad (3.6)$$

The proposed algorithm generates a Quasi-Fejer convergent bounded sequence.

**Definition 3.1.** Let $S$ be nonempty subset of $R^n$. A sequence $\{x_i\}$ is said to Quasi-Fejerand convergence to $S$ if and only if for all $x \in S$, there exist $\bar{i}$ and a summable sequence $\{\delta_i\} \subset R^+$ such that

$$\|x_{i+1} - x\|^2 = \|x_i - x\|^2 + d_i \quad \text{for all } i \geq \bar{i}. \quad (3.7)$$

**Lemma 3.2.** Let $\{x_i\}$ be a bounded sequence and $f$ a $K$-convex function. Then, any sequence $\{U_i\} \in \partial f(x_i)$ is bounded.

**Remark.** A very popular strategy for solving vector optimization problem is the scalarization approach where vector valued function is converted into scalar valued function by different techniques. Most popular techniques are “parametric” and “weight” method. Disadvantage of the parametric approach is that the choice of the parameters is not known in advance, leaving the modeler and the decision maker with the burden of choosing them [6]. Another method is “weight” method. Basically in this technique one minimize a linear combination of the objectives where the vector of “weights” is not known a priori. This procedure may lead to unbounded numerical problems, which therefore may lack minimizers, see [7, 17, 22]. In this proposed method we do not used scalarization. Therefore we neither required parameter to scalarize nor weight parameter. We have explored the structure of the vectors problem.
4. Convergence Result

There are many results given by numerous researcher [2, 5, 6, 7] on convergence of the subgradient method for nonsmooth convex optimization problem, here functions are scalar functions. Even the convergence of subgradient methods for vector optimization by using scalarization technique has been also studied by several authors [14, 15]. In this case of content step size (or step length) the subgradient algorithm is guaranteed to converge to within some range of the optimal value, i.e.

\[
\lim_{k \to \infty} f_k - f^* < \epsilon. \tag{4.1}
\]

Here \( f^* \) is optimal value.

In this section we attempt to establish convergence of the sequence generated by the algorithm by vector valued function without using scalarization, to a point in the solution set. From now on we denote by \( \{x_k\} \), the sequence generated by our algorithm starting from \( x_0 \in F \). To establish the convergence we also propose and prove followings:

**Proposition 4.1** (Belongingness). \( \{x_i\} \in F \) where \( \{x_i\} \) is the sequence generated by the Algorithm A.

**Proposition 4.2** (Validity of Stopping Criterion). If \( g_i = 0 \) then \( x_i \in F^* \).

**Proof.** Let \( x_i \notin F^* \). Then there exist \( z \in F \) such that \( fz < f(x_i) \). Let us take \( g = z - x_i \). It is true that \( g \neq 0 \). By \( K \)-convexity of \( f \), \( f(z) \succ f(x_i) + U_i g \). Implies that \( U_i g \in -\text{int}(K) \), and therefore \( \max_{y \in K^*} \{y^T U_i g\} < 0 \).

Considering \( z(\lambda) = \lambda z + (1 - \lambda) x_i \), \( 0 < \lambda < 1 \).

We get

\[
\frac{\|g_i\|^2}{2} + \frac{t_i}{\mu_i} \max_{y \in K^*} \{y^T U_i g_i\} \\
\leq \frac{\|z_\lambda - x_i\|^2}{2} + \frac{t_i}{\mu_i} \max_{y \in K^*} \{y^T (z_\lambda - x_i)\} \\
\leq \lambda \left[ \frac{\|g\|^2}{2} + \frac{t_i}{\mu_i} \max_{y \in K^*} \{y^T U_i g_i\} \right] \text{ for all } \lambda \in [0, 1]. \tag{4.2}
\]
Then
\[
\frac{\|g_i\|^2}{2} + t_i \max_{y \in K^*} \{y^T U_i g_i\} < 0, \text{ and hence } g_i \neq 0. \tag{4.3}
\]

Here we define and make the following assumption

**Definition 4.3.** For the sequence \(\{x_i\}\) generated by the algorithm we define the set
\[
T = \{x \in F : \min_{y \in K^*} (f(x_i) - f(x) - \epsilon)^T y \geq 0, \ i = 0, 1, 2, 3, \ldots\} \tag{4.4}
\]

**Assumption 4.4.** For all sequence \(\{x_i\} \subset R^n \setminus F^*\), there exists \(x \in R^n\) such that \(f(x) \leq f(x_i) - \epsilon\) for \(i = 0, 1, 2, 3, \ldots\).

Under the above assumption set \(T\) is non-empty.

Next we will prove that the sequence generated by the algorithm is a Quasi-Fejer convergent sequence and hence bounded.

**Proposition 4.5.** Let \(\{x_i\}\) be the sequence generated by the Algorithm A and \(g_i\) and \(\mu_i\) is as defined in (3.3) and (3.5) respectively. Then for all \(x \in F\)
\[
\|x_{i+1} - x\|^2 \leq \|x_i - x\|^2 + 3t_i^2 - \frac{t_i}{\mu_i} \min_{y \in K} \{[f(x_i) - f(x)]^T y\}
\]
As a result \(\|x_{i+1} - x\|^2 = \|x_i - x\|^2 + 3t_i^2\) and \(\{x_i\}\) is bounded.

**Proof.** First we prove the following inequality
\[
\|x_{i+1} - x_i\| = t_i, \quad i = 0, 1, 2, 3, \ldots
\]
In order to prove this inequality note that
\[
x_{i+1} = x_i + g_i \quad \text{(as in (3.6)) and } x_{i+1} - x_i = g_i
\]
\[
\Rightarrow x_{i+1} - x_i = g_i
\]
\[
\Rightarrow \|x_{i+1} - x_i\| = g_i
\]
To prove
\[
\|x_{i+1} - x_i\| = t_i \tag{4.5a}
\]
We have to prove that
\[
\|g_i\| = t_i \tag{4.5b}
\]
According to definition of $g_i$ i.e.

$$g_i = \text{argmin}_{w \in F_i} \left\{ \frac{1}{2} \| w \|^2 + \frac{t_i}{\mu_i} \max_{y \in K^*} \{ y^T U_i w \} \right\}$$  \hspace{1cm} (4.6)$$

and by the triangular inequality of norm

$$\| g_i \|^2 = \frac{t_i}{\mu_i} \max_{y \in K^*} y^T U_i g_i \quad \text{where} \quad \mu_i = \max_{y \in K^*} \| U_i \|^T y$$

$$= \frac{t_i}{\mu_i} \|(U_i)^T \bar{y}_i \| \| g_i \| \quad \text{where} \quad \bar{y}_i \in K^*$$

$$\leq t_i \| g_i \|.$$  

Since $\| g_i \| \neq 0$. $\| g_i \| \leq t_i$.

Therefore, we get that

$$\| x_{i+1} - x_i \| \leq t_i.$$  \hspace{1cm} (4.7)$$

Now by simple algebraic manipulation, for all $x \in F$

$$\| x_{i+1} - x \|^2 = \| x_{i+1} - x_i + x_i - x \|^2$$

$$= \| x_{i+1} - x_i \|^2 + \| x_i - x \|^2 - 2 \langle x_{i+1} - x_i, x - x_i \rangle$$

$$\| x_{i+1} - x \|^2 - \| x_{i+1} - x_i \|^2 + 2 \langle x_{i+1} - x_i, x - x_i \rangle$$

$$= \| x_{i+1} - x_i \|^2 \leq t_i^2$$

or

$$2 \langle x_{i+1} - x_i, x - x_i \rangle \leq t_i^2 + \| x_i - x \|^2 - \| x_{i+1} - x \|^2$$

or

$$2 \langle g_i, x - x_i \rangle \leq t_i^2 + \| x_i - x \|^2 - \| x_{i+1} - x \|^2.$$  \hspace{1cm} (4.8)$$

Now by formula for the subdifferential of a maximum of convex function and definition of $g_i$ (as in (4.6)) there exists a positive integer $I(i)$ and $y \in K^*$ and $\lambda_k^i > 0$ with $1 \leq k \leq I(i)$,

$$\sum_{k=1}^{I(i)} \lambda_k^i = 1.$$  \hspace{1cm} (4.9)$$
and
\[(y^k_i)^T U_i g_i = \max_{y \in K^*} \{ y^T U_i g \} \quad (1 \leq k \leq I(i)) \quad (4.10)\]

By definition of \(g_i\)
\[\|g_i\|^2 \leq \frac{t_i}{\mu_i} (\bar{y}_i)^T U_i g_i, \text{ where } \bar{y}_i \in K^* \quad (4.11)\]

\[\langle g_i, x - x_i \rangle \geq \frac{t_i}{\mu_i} (\bar{y}_i)^T U_i g_i + \frac{t_i}{\mu_i} \sum_{k=1}^{I(i)} \lambda^k_i (y^k_i)^T U_i (x_i - x). \quad (4.12)\]

Where \(\bar{y}_i\) is max of \(y\).

Therefore
\[t^2_i + \|x_i - x\|^2 - \|x_{i+1} - x\|^2 \geq 2 \left( \frac{t_i}{\mu_i} (\bar{y}_i)^T U_i g_i + \frac{t_i}{\mu_i} \sum_{k=1}^{I(i)} \lambda^k_i (y^k_i)^T U_i (x_i - x) \right),\]

\[t^2_i + \|x_i - x\|^2 - \|x_{i+1} - x\|^2 \geq -2 \frac{t_i}{\mu_i} \|(U_i)^T \bar{y}_i\| \|g_i\| + 2 \frac{t_i}{\mu_i} \sum_{k=1}^{I(i)} \lambda^k_i (y^k_i)^T U_i (x_i - x) \quad (4.13)\]

From (4.5b), we have
\[t^2_i + \|x_i - x\|^2 - \|x_{i+1} - x\|^2 \geq -2t^2_i + 2 \frac{t_i}{\mu_i} \sum_{k=1}^{I(i)} \lambda^k_i (y^k_i)^T U_i (x_i - x) \quad (4.13)\]

By using (2.2) for \(K\)-convex function we can say
\[U_i(x_i - x) \geq f(x_i) - f(x). \quad (4.14)\]

Therefore from (4.13) and (4.14), we get
\[t^2_i + \|x_i - x\|^2 - \|x_{i+1} - x\|^2 \geq -2t^2_i + 2 \frac{t_i}{\mu_i} \sum_{k=1}^{I(i)} \lambda^k_i (y^k_i)(f(x_i) - f(x)) \]
\[ \geq -2t_i^2 + \frac{2t_i}{\mu_i} \min_{y \in K^*} \{(f(x_i) - f(x))^Ty\} \]

Implies that

\[ 3t_i^2 + \|x_i - x\|^2 - \|x_{i+1} - x\|^2 \geq 0. \]

Hence

\[ \|x_{i+1} - x\|^2 \leq \|x_i - x\|^2 + 3t_i^2 \text{ for all } i \geq \bar{i} \quad (4.15) \]

**Boundedness**

Since \( t_i \) satisfy (3.2) and \( T \neq \Phi \), therefore the sequence \( \{x_i\} \) generated by the algorithm is quasi-Fejer convergent in \( T \) and hence by property of quasi-Fejer, sequence \( \{x_i\} \) bounded for all \( i = \bar{i} \) [6].

**Theorem 4.6.** The sequence \( \{x_i\} \) generated by algorithm converge to some cluster \( x^* \in T \).

**Proof.** Since \( \mu_i \leq \rho \), where \( \rho \) is radius of closed ball centered at 0. Therefore for \( x \in F \) we obtain

\[ 2t_i \frac{t_i}{\mu_i} \min_{y \in K^*} \{(f(x_i) - f(x) - \epsilon)^Ty\} \leq \|x_i - x\|^2 - \|x_{i+1} - x\|^2 + 3t_i^2. \quad (4.16) \]

Since

\[ \mu_i \leq \rho \]

\[ \Rightarrow \quad 1 \leq \frac{1}{\mu_i} \geq \frac{1}{\rho}. \quad (4.17) \]

For sake of a simpler notation let us define the sequence

\[ \gamma_i = \min_{y \in K^*} \{(f(x_i) - f(x) - \epsilon)^Ty\} \]

and rewrite above inequality (4.16) as

\[ 2t_i \frac{\gamma_i}{\rho} \leq \|x_i - x\|^2 - \|x_{i+1} - x\|^2 + 3t_i^2. \]

Taking summation on both sides from \( i = 0 \) to \( i = m \), we get

\[ 2 \sum_{i=0}^m t_i \gamma_i \leq \sum_{i=0}^m (\|x_i - x\|^2 - \|x_{i+1} - x\|^2) + 3 \sum_{i=0}^m t_i^2. \]
Taking \( m \to \infty \), we get

\[
\text{R.H.S.} \leq \|x_0 - x\|^2 + 3 \sum_{i=0}^{\infty} t_i^2.
\]

implies that

\[
\sum_{i=0}^{\infty} t_i \gamma_i < \infty. \tag{4.18}
\]

Since \( \sum_{i=0}^{\infty} t_i = \infty \). Therefore from (4.18) we conclude that \( \liminf \gamma_i = 0 \). Since \( \gamma_i \geq 0 \), therefore there exists a subsequence \( \{\gamma_{i_j}\} \) such that

\[
\lim_{j \to \infty} \gamma_{i_j} = 0.
\]

By definition of \( \gamma_i \), we obtain that

\[
\lim_{j \to \infty} f_k(x_{i_j}) = f_k(x) - \epsilon, \quad k = 1, 2, 3, 4, \ldots, m.
\]

Let \( x^* \) be cluster point of \( \{x_{i_j}\} \) such that \( x^* = \lim_{i \to \infty} x_{i_j} \) then

\[
f_k(x^*) = \lim_{i \to \infty} f_k(x_{i_j}) = \lim_{j \to \infty} f_k(x_{i_j}) = f_k(x) \tag{4.19}
\]

From (4.19) it follows that \( x^* \in T \) hence \( \{x_{i_j}\} \to x^* \). \( \square \)

The following theorem is our main result.

**Theorem 4.7.** Sequence \( \{x_i\} \) converge to a \( \epsilon \)-efficient solution of the problem (2.1).

*Proof.* From Theorem 4.1, we have that \( \{x_i\} \) converge to cluster point \( x^* \in T \). Also \( x^* \) is a cluster point of \( \{x_{i_j}\} \) associated to the subsequence \( \{x_{i_j}\} \). Follows from Preposition 4.1 that \( x^* \in F \). Now we have to prove that \( x^* \in F^* \). Let us suppose that \( x^* \notin F^* \). Therefore there exists \( \bar{x} \) and an index \( j \) such that

\[
f_k(\bar{x}) \leq f_k(x^*) - \epsilon, \quad k = 1, 2, 3, 4, \ldots, m \text{ and } f_j(\bar{x}) < f_j(x^*) - \epsilon.
\]

Implies that

\[
[f_j(x^*) - f_j(\bar{x}) - \epsilon]y \geq 0, \quad \text{for all } y \in K. \tag{4.20}
\]
Therefore by definition of $T \bar{x} \in T$.

Let $\bar{y} \in k^*$ be minimize of all $y$. Then

$$0 \geq \lim_{i \to \infty} \{\min_{y \in k^*}\{(f(x_i) - f(\bar{x}) - \epsilon)^T y\}\}$$

$$\geq \lim_{i \to \infty} \{(f(x_i) - f(\bar{x}) - \epsilon)^T y\}$$

Since $x^* \in T$ is a cluster point of $\{x_i\}$ associated with subsequence $\{x_{i_j}\}$.

$$\Rightarrow 0 \geq \{(f(x^*) - f(\bar{x}) - \epsilon)^T \bar{y}\},$$

Which contradict (4.20). Therefore $x^* \in F^*$. Hence proved. 

5. Conclusion

In this paper our algorithms do not use scalarization method, a very popular strategy for solving multi-objective problem. Our methods strongly explored the vector structure of the problem. But as we observe during journey of discussion, the vector subgradient method out performs the simpler method in many instances. But if method work good (under suitable initialization and step size like Newton’s method) it’s work very good and give suitable optimal solution but if it is bad (as in case of ill-conditioning or in case of gully-shaped functions) it is very bad and always far away from efficient solution. Therefore with suitable step-size and good initialization the idea discussed in this paper may be explored to deal with more efficient method to solve vector optimization problem like bundle method, cutting plane method etc.

References

[1] Ya.I. Alber and A.N. Iusem, Extension of subgradient techniques for non-smooth optimization in Banach space, Set-Valued Analysis, 9 (2001), 315-335.

[2] E. Allen, R. Helgason and J. Kennigton, A generalization of Polyak’s convergence result for subgradient optimization, Mathematical Programming, 37 (1987), 309-317.

[3] S. Bolintineanu, Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization, Journal of Optimization and Applications, 106 (2000), 265-296.
[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, (2004).

[5] F.E. Browder, Convergence theorem for sequence of non linear operators in Banach spaces, *Mathematische Zeitschrift*, 100 (1967), 201-329.

[6] R.S. Burachik, L.M. Grana Drummond, A.N. Lusem and B.F. Svaiter, Full Convergence of the steepest descent method with inexact line search, *Optimization*, 32 (1995), 137-146.

[7] J.Y. Cruz Bello, L.R. Lucambio Perez and J.G. Melo, Convergence of the projection gradient method for quasiconvex multiobjective optimization, *Nonlinear Analysis*, 74 (2011), 5268-5273.

[8] Yu.M. Ermoliev, Methods of solutions of nonlinear extremal problems, *Cybernetics*, 2 (4), 1-16.

[9] Yu.M. Ermoliev, On the minimization of nondifferentiable functions, *Cybernetics*, 3 (4), 72.

[10] M.L. Fisher, The Lagrangian relaxation method for solving integer programming problem, *Management Science*, 27 (1981), 1-18.

[11] J. Fliege and B.F. Svaiter, Steepest descent methods for multicriteria optimization, *Math. Meth. Oper. Res.*, 51 (3) (2000), 479-494.

[12] J. Fliege, L.M.G. Drummond and B.F. Svaiter, Newton’s method for multiobjective optimization, *SIAM J. Opt.*, 20 (2) (2009), 602-626.

[13] J.L. Goffin, Subgradient optimization in non-smooth optimization, *Documenta Mathematica*, Extra Volume, ISMP (2012), 277-290.

[14] J.L. Goffin, On convergent rate of subgradient optimization methods, *Mathematical Programming*, 13 (1971), 329-347.

[15] J.L. Goffin and K.C. Kiwiel, Convergence of a simple subgradient method, *Mathematical Programming*, 85 (1999), 207-211.

[16] J.L. Goffin, Non-differentiable optimization and the relaxation method, *Nonsmooth Proceeding of the IIASA workshop*, March 28-April 8, Pergamon Press, 31-50 (1977).

[17] L.M. Grana Drummond and A.N. Iusem, A projected gradient method for vector optimization problems, *Computational Optimization and Applications*, 28 (2004), 5-30.
[18] L.M. Grana Drummond, N. Maculan and B.F. Svaiter, On the choice of parameters for the weighting method in vector optimization, *Mathematical Programming*, 111 (2008), 201

[19] M. Held and R.M. Karp, The travelling salesman problem and minimum spanning tree, 1 (1971), 6-25, 20.

[20] M. Held, P. Wolf and H.P. Crowder, Validation of subgradient optimization, *Mathematical Programming*, 6 (1971), 62-88.

[21] J. Jahn, *Vector Optimization-Theory, Application and Extensions*, Springer, Erlangen (2003).

[22] D.T. Luc, Scalarization of vector optimization problems, *Journal of Optimization Theory and Applications*, 55 (1987), 85-102.

[23] D.T. Luc, *Theory of Vector Optimization*, Lecture Notes in Economics and Mathematical Systems, Springer Berlin (1989).

[24] D.T. Luc, N.X. Tan and P.N. Tinh, Convex vector functions and their subdifferential, *ACTA Mathematica Vietnamica*, 23 (1998), 107-127.

[25] K.M. Miettinen and M.M. Makela, An iterative methods for non-smooth multiobjective optimization with an application to optimal control, *Optimization Methods Softw.*, 2 (1993), 31-44.

[26] K.M. Miettinen, *Nonlinear Multiobjective Optimization*, Kluwer Academic publishers, Norwell (1999).

[27] Yu Nesterov, *Introductory Lecture on Convex Optimization. A Basic Course*, Kluwer Academic Publication, Boston, London (2004).

[28] Yu Nesterov, Primal-dual subgradient methods for convex problems, *Mathematical Programming Series B*, 120 (2009), 221-259.

[29] Yu Nesterov, Subgradient methods for Huge-Scale Optimization problems, *CORE Discussion Paper*, 2 (2012).

[30] G.D. Pillo, L. Grippo and S. Lucidi, A smooth method for the finite minimax problem, *Mathematical Programming*, 60 (1993), 187-214.

[31] E. Polak, D. Mayne and J. Higgins, A superlinear convergent algorithm formin th max problems, *Proceedings of the 28 Conference on Decision and Control*, Tampa, Florida, 894-898 (1989).
[32] B.T. Polyak, A general method of solving external problems, *Soviet Maths, Doklady*, 8 (1967), 593-597.

[33] B.T. Polyak, Minimization of unsmooth functions, *USSR Computational Mathematics and Mathematical Physics*, 9 (1969), 14-29.

[34] B.T. Polyak, *Subgradient Methods: A survey of Soviet Research*, Pergamon Press (1987).

[35] H. Sawaragi and T. Tanino, Theory of Multiobjective Optimization, Academic Press, Orlando (1985).

[36] N.Z. Shor, An application of the method of gradient descent to the solution of the network transportation problem, in: *Materialy Naucnovo seminara po Teoret i Prikland Voprosam Kibernet. i Issled Operacii, Nucnyi Sov. Po Kibernet*, Akad. Nauk Ukrain SSSR, vyp. 1 (1962), 9-17

[37] N.Z. Shor, On the structure of algorithm for numerical solution of problems of optimal planning and design, *Diss Doctor. Philos. Kiev* (1964).

[38] N.Z. Shor, Generalization of gradient methods for nonsmooth functions and their applications to mathematical programming, *Economic and Mathematical Methods*, 12 (2) (1976), 337-356.

[39] N.Z. Shor, *Minimization Methods for Non-differentiable Functions*, Springer Berlin (1985).

[40] N.Z. Shor, *Non-Differentiable Optimization and polynomial Problems*, Kluwer Academic Publication, Boston, Doordrecht, London (1998).