



A line search technique for a class of multi-objective optimization problems using subgradient

Dinesh Kumar¹ · Geetanjali Panda¹

Received: 1 February 2024 / Accepted: 13 April 2024 / Published online: 8 May 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

This paper proposes a line search technique to solve a special class of multi-objective optimization problems in which the objective functions are supposed to be convex but need not be differentiable. This is an iterative process to determine Pareto critical points. A suitable sub-problem is proposed at every iteration of the iterative process to determine the direction vector using the sub-differential of every objective function at that point. The proposed method is verified in numerical examples. This methodology does not bear any burden of selecting suitable parameters like the scalarization methods.

Keywords Line search technique · Convex optimization · Pareto optimal point · Multi-objective optimization · Sub-gradient method

Mathematics Subject Classification 49J52 · 90C29 · 65K10 · 49M37 · 90C30 · 90C25

1 Introduction

The formulation of a general unconstrained multi-objective optimization problem is

$$\text{MP : } \min_{x \in \mathbb{R}^n} \mathcal{F}(x),$$

where $\mathcal{F}(x) = (f_1(x), f_2(x), f_3(x), \dots, f_m(x))^T$, $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, 3, \dots, m$, $m \geq 2$. In an ideal situation, a single point minimizes all the objective functions. But, in real-life situations, at least one objective is compromised. Therefore, the solution of MP is termed as a compromising/Pareto optimal/efficient solution.

✉ Geetanjali Panda
geetanjali@maths.iitkgp.ac.in

Dinesh Kumar
dineshkumar022000@gmail.com

¹ Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

The scalarization approaches [1–3] are the well-known techniques to solve MP, which transform MP into a single objective optimization problem using some parameters, provided by the decision maker. These methods often have difficulties in finding a well-diversified Pareto front. Heuristic methods [4–9] are widely used to solve MP, for which convergence property is not guaranteed. However, these methods usually provide approximate Pareto front.

In recent years, line search techniques for single objective cases have been extended to MP to avoid these difficulties. Some of the line search techniques for MP are the steepest descent methods [10, 11], projected gradient methods [12, 13], Newton method [14], quasi-Newton methods [15], q-Newton's method [16], q-steepest descent method [17], Newton's method for uncertain MP [18] and some known references for constrained MP are sequential quadratic programming techniques [19], sequential quadratically constrained quadratic programming techniques [20, 21] etc.. The general line search techniques for MP, which are developed in these reference papers, generate a sequence of points $\{x^k\}$ as $x^{k+1} = x^k + \alpha_k d^k$, which converges to a Pareto critical point under some assumptions. A suitable sub-problem is formulated at every iterate x^k , whose solution is the direction vector d^k , and the step length α_k is determined using Armijo-like conditions. The BFGS update formula is modified to generate a sequence of positive definite matrices to ensure the descent property. All the above-stated methods are gradient-based algorithms and applicable only if all functions of MP are continuously differentiable.

In recent years, some line search techniques have been developed for nonsmooth MP, in which all the functions may not be differentiable. Recently the concept of subdifferential has been used by Tanabe et. al [22] to develop the proximal gradient method for a specific class of optimization problems, in which each objective function is composed of two parts. One part is a proper convex function that is lower semicontinuous but not necessarily differentiable. The other part is a continuously differentiable function. The subproblem at every iterating point takes care of the first-order information of the differentiable function. The proximal sub-gradient method is also extended by Bento et. al [23] to constrained MP, in which the subproblems are constructed using one element of the subdifferential set. Da Cruz Neto et al. [24] develop a subgradient method for a class of MP, in which all the objective functions are quasi-convex, locally Lipschitz functions, and not necessarily differentiable. These papers have studied many theoretical results related to the existence of solutions to MP using subdifferential concepts.

The present work aims to develop an iterative process for MP in which the gradient information of all the functions is unavailable. The basic concept of this method is a variation of the existing gradient-based methods cited above. First, the direction vector at every iterate is obtained as the solution of a subproblem using sub-differentials of the objective functions, which are not differentiable. The step-length selection process is provided to ensure the descent property of every objective function along this direction. Moreover, under some standard assumptions, the global convergence of this iterative process is justified.

The following is an outline for the paper. In Sect. 2, certain prerequisites pertaining to MP are described. Section 3 develops some theoretical results for converging the iterative process. In Sect. 4, the methodology is verified in some numerical examples.

2 Prerequisites

Throughout this manuscript, we use the following notations.

MP: Multi-objective programming

SP: Single objective programming

PF: Pareto front

PO: Pareto optimal

PC: Pareto critical

$\Lambda_m = \{1, 2, 3, \dots, m\}$, $\mathbb{R}_+^m := \{a \in \mathbb{R}^m \mid 0 \leq a_i \text{ for all } i \in \Lambda_m\}$

$\mathbb{R}_{++}^m := \{a \in \mathbb{R}^m \mid 0 < a_i \text{ for all } i \in \Lambda_m\}$.

For $a^1, a^2 \in \mathbb{R}^m$, the vector inequalities are defined as:

$a^2 \geq a^1 \Leftrightarrow a^2 - a^1 \in \mathbb{R}_+^m \setminus \{0\}$ and $a^2 > a^1 \Leftrightarrow a^2 - a^1 \in \mathbb{R}_{++}^m$.

$\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called a convex vector function if for each $y, z \in \mathbb{R}^n, b \in [0, 1]$

$$\mathcal{F}(by + (1 - b)z) \leq b\mathcal{F}(y) + (1 - b)\mathcal{F}(z).$$

$\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called a strict convex vector function if for each $y, z \in \mathbb{R}^n, b \in (0, 1)$

$$\mathcal{F}(by + (1 - b)z) < b\mathcal{F}(y) + (1 - b)\mathcal{F}(z).$$

Subdifferential of a convex function f_i on \mathbb{R}^n at x is defined as

$$\partial f_i(x) := \{g_i \in \mathbb{R}^n \mid g_i^T(y - x) \leq f_i(y) - f_i(x) \ \forall y \in \mathbb{R}^n\}.$$

The column vector $g_i := g_i(x)$ is the subgradient of f_i at x . If f_i is differentiable then $g_i = \nabla f_i(x)$.

Subdifferential of a convex vector function $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ at x is

$$\partial \mathcal{F}(x) := \{A \in L(\mathbb{R}^n, \mathbb{R}^m) \mid A(y - x) \leq \mathcal{F}(y) - \mathcal{F}(x) \text{ for all } y \in \mathbb{R}^n\}.$$

Here $L(\mathbb{R}^n, \mathbb{R}^m)$ represents the space of linear continuous mappings from \mathbb{R}^n into \mathbb{R}^m , it is also known as the $(m \times n)$ real matrix space. If all f_i are differentiable at x then $J\mathcal{F}(x) = \partial \mathcal{F}(x)$, $J\mathcal{F}(x)$ represents the Jacobian of \mathcal{F} at point x . In fact the matrix $A \in \partial \mathcal{F}(x)$ is $(g_1, g_2, \dots, g_m)^T$.

Let every f_i be locally Lipschitz continuous at y . The Clarke generalized directional derivative of the function f_i at y along the direction of $d \in \mathbb{R}^n$ is stated as follows.

$$f_i^o(y; d) = \limsup_{z \rightarrow y, t \downarrow 0} \frac{f_i(z + td) - f_i(z)}{t}$$

The subsequent symbols are employed throughout this manuscript. d_x is the direction vector at x , d_{x^k} is the direction vector at x^k , $g_i := g_i(x) \in \partial f_i(x)$, $i \in \Lambda_m$ and $g_i^k := g_i(x^k) \in \partial f_i(x^k)$, $i \in \Lambda_m$.

3 Methodology

A point x^* is called the PO solution of MP if no other point exists x such that $\mathcal{F}(x) \leq \mathcal{F}(x^*)$ and $\mathcal{F}(x) \neq \mathcal{F}(x^*)$, and x^* is said to be the weak PO solution of MP if no other point exists x for which $\mathcal{F}(x) < \mathcal{F}(x^*)$. Suppose x^* is not a PO point of MP. Then $\exists x \in \mathbb{R}^n$ such that $\mathcal{F}(x) \leq \mathcal{F}(x^*)$ and $\mathcal{F}(x) \neq \mathcal{F}(x^*)$. Using the subdifferential's definition of \mathcal{F} at x^* ,

$$0 \geq \mathcal{F}(x) - \mathcal{F}(x^*) \geq A(x - x^*) \text{ for all } A \in \partial\mathcal{F}(x^*).$$

For $x = x^* + \alpha d$, where $\alpha > 0$ and the direction vector d at x^* , the above inequality becomes $Ad \leq 0$. Hence, if x^* is not a PO point then there exists $d \in \mathbb{R}^n$ such that $Ad \in -\mathbb{R}_+^m - \{0\}$ for all $A \in \partial\mathcal{F}(x^*)$. A point x^* is a PO point if $\exists d \in \mathbb{R}^n$, $Ad \notin -\mathbb{R}_+^m - \{0\}$ for some $A \in \partial\mathcal{F}(x^*)$. Therefore if for all $d \in \mathbb{R}^n$, $Ad \notin -\mathbb{R}_+^m$, for some $A \in \partial\mathcal{F}(x^*)$ then x^* is a local weak PO point of MP. We say a point that satisfies this necessary condition is a PC point. Hence, the definition of the PC point defined in [25] is similar to the one that follows.

Definition 1 A point $x^* \in \mathbb{R}^n$ is said to be a PC point of MP if $\exists A \in \partial\mathcal{F}(x^*)$ so that $(Ad)_i \geq 0$ for at least one $i \in \Lambda_m$ and for all $d \in \mathbb{R}^n$, where $(Ad)_i$ denotes the i^{th} component of vector Ad .

Theorem 1 If \mathcal{F} is a convex vector function then the PC point of MP is a weak PO point. If \mathcal{F} is a strictly convex vector function then the PC point of MP is a PO point.

Proof Suppose \mathcal{F} is a convex vector function and x^* is a PC point of MP. If possible, x^* is not a weak PO point. It is consequent that there exists $y \in \mathbb{R}^n$, $y \neq x^*$ such that $f_i(y) < f_i(x^*)$ for all $i \in \Lambda_m$. Therefore

$$0 > f_i(y) - f_i(x^*) \geq g_i^T(y - x^*) \text{ for all } i \in \Lambda_m, \text{ for all } g_i \in \partial f_i(x).$$

Hence $\exists d (= y - x^*)$ such that $Ad < 0$ for all $A \in \partial\mathcal{F}(x^*)$. This implies x^* is not a PC point, which follows from Definition 1. So x^* is a weak PO point.

Suppose \mathcal{F} is a strictly convex vector function and x^* is a PC point of MP. If possible, suppose x^* is not a PO point. Therefore, $\exists a y \in \mathbb{R}^n$ such that $f_i(y) < f_i(x) \forall i \in \Lambda_m$. Since all f_i are strictly convex functions, for every $i \in \Lambda_m$,

$$f_i(x^*) + \alpha(f_i(y) - f_i(x^*)) > f_i(x^* + \alpha(y - x^*)) \forall \alpha \in (0, 1).$$

Consequently, for all $i \in \Lambda_m$ and for all $g_i \in \partial f_i(x^*)$,

$$0 \geq \alpha(f_i(y) - f_i(x^*)) > f_i(x^* + \alpha(y - x^*)) - f_i(x^*) \geq g_i^T((x^* + \alpha(y - x^*)) - x^*).$$

This implies $g_i^T(y - x^*) < 0 \forall i \in \Lambda_m$. Hence $\exists d = y - x^*$ such that $Ad < 0 \forall A \in \partial\mathcal{F}(x^*)$. This contradicts the idea that x is a PC point. So x^* must be a PO point. \square

For $x \in \mathbb{R}^n$ define $\mathcal{F}_x : \mathbb{R}^n \rightarrow \mathbb{R}$ as

$$\mathcal{F}_x(d) = \max\{(Ad)_1, (Ad)_2, (Ad)_3, \dots, (Ad)_m \ \forall A \in \partial\mathcal{F}(x)\},$$

and consider a subproblem at x as

$$\min_{d \in \mathbb{R}^n} \mathcal{F}_x(d) + \frac{1}{2} \|d\|^2 \tag{1}$$

A simple reformulation of the above subproblem is

$$(P_x) : \begin{cases} \min_{t \in \mathbb{R}, d \in \mathbb{R}^n} t + \frac{1}{2} \|d\|^2 \\ \text{s.t. } (Ad)_i \leq t, \ \forall i \in \Lambda_m, \ \forall A \in \partial\mathcal{F}(x) \end{cases}.$$

Solution of (P_x) : (P_x) is a convex quadratic semi-infinite programming problem characterized by linear inequality constraints. It is difficult to solve a general semi-infinite programming problem. However, with a known subdifferential set, some heuristic approaches can solve (P_x) using the discretization process over the subdifferential set. The MATLAB command 'fseminf' also takes care of the discretization process, as discussed in Subsection 3.1. This heuristic approach can be implemented if the subdifferential set is known in advance. To compute the subdifferential set, a separate methodology has to be developed for general nonsmooth functions, which is beside the context of the current work. The proposed methodology focuses only on the theoretical developments behind an iterative process based on a line search technique for MP using subdifferential set information.

Theorem 2 *Let (d_x, t_x) be the unique solution of the Subproblem (P_x) and v_x be the optimal value.*

1. For any $x \in \mathbb{R}^n$, optimal value of (P_x) is non positive and vanishes if x is a PC point of MP.
2. If x is a PC point of MP then $(t_x, d_x) = (0, 0)$ is the solution of (P_x) .
3. If x is not a PC then $v_x < 0$ and $\mathcal{F}_x(d_x) \leq -(1/2)\|d_x\|^2 < 0$.
4. The mappings $x \rightarrow d_x$ and $x \rightarrow v_x$ are continuous.

Proof 1. For any $x \in \mathbb{R}^n$, suppose v_x is the optimal value of (P_x) , which is acquired at the optimal solution (t_x, d_x) . We have $\mathcal{F}_x(d_x) + \frac{1}{2}\|d_x\|^2 \leq \mathcal{F}_x(d) + \frac{1}{2}\|d\|^2 \ \forall d \in \mathbb{R}^n$. At $d = 0$, $\mathcal{F}_x(d) + \frac{1}{2}\|d\|^2 = 0$.

Hence $\mathcal{F}_x(d_x) + \frac{1}{2}\|d_x\|^2 \leq 0$. That is, $v_x \leq 0$.

Suppose x is a PC point for MP and $v_x < 0$. That is, $\mathcal{F}_x(d_x) + \frac{1}{2}\|d_x\|^2 < 0$ or $\mathcal{F}_x(d_x) < 0$. So $Ad_x < 0 \ \forall A \in \partial\mathcal{F}(x^*)$, which follows from the definition of $\mathcal{F}_x(d_x)$. This contradicts that x is the PC point for which $Ad \notin -\mathbb{R}_{++}^m$ for some $A \in \partial\mathcal{F}(x)$ and $\forall d \in \mathbb{R}^n$. Hence $v_x = 0$.

2. From (1), $v_x = 0$ at the PC point x and (t_x, d_x) is the solution of (P_x) , so $t_x + \frac{1}{2}\|d_x\|^2 = 0$. That is,

$$\mathcal{F}_x(d_x) \leq t_x \leq t_x + \frac{1}{2}\|d_x\|^2 = 0$$

Since x is a PC point so $\mathcal{F}_x(d_x) \geq 0$. Hence $t_x = 0, d_x = 0$.

3. Consider x to be a non-PC point. It follows that there is $d \in \mathbb{R}^n$ such that $Ad < 0 \forall A \in \partial \mathcal{F}(x)$. That is, $\mathcal{F}_x(d) < 0$ for some $d \in \mathbb{R}^n$.

Since $\mathcal{F}_x(d)$ is positive homogeneous of degree 1 we can consider $\frac{\mathcal{F}_x(d)}{\|d\|^2}$, denote $\overline{\mathcal{F}_x(d)} = -\frac{\mathcal{F}_x(d)}{\|d\|^2}$, and $\bar{d} = \overline{\mathcal{F}_x(d)}. d$. Hence

$$\begin{aligned} \mathcal{F}_x(\bar{d}) + \frac{1}{2}\|\bar{d}\|^2 &= \overline{\mathcal{F}_x(d)} \cdot \mathcal{F}_x(d) + \frac{1}{2}(\overline{\mathcal{F}_x(d)})^2\|d\|^2 \\ &= -\frac{(\mathcal{F}_x(d))^2}{\|d\|^2} + \frac{1}{2}\left(-\frac{\mathcal{F}_x(d)}{\|d\|^2}\right)^2\|d\|^2 \\ &= -\frac{1}{2}\frac{(\mathcal{F}_x(d))^2}{\|d\|^2} \\ &< 0. \end{aligned}$$

Hence $v_x < 0$. This implies $\min_{d \in \mathbb{R}^n} \mathcal{F}_x(d) + \frac{1}{2}\|d\|^2 < 0$. That is, $\exists d \in \mathbb{R}^n$ such that $\mathcal{F}_x(d) + \frac{1}{2}\|d\|^2 < 0$. The result follows.

4. For $\epsilon > 0$ and $\bar{x} \in \mathbb{R}^n$, consider the set $S(\bar{x}) := \{d \in \mathbb{R}^n \mid \|d_{\bar{x}} - d\| = \epsilon\}$. Let $d_{\bar{x}}$ be the minimizer for the Subproblem (1) at \bar{x} . The objective function of Subproblem (1) is strongly convex with a modulus of $1/2$ since $\mathcal{F}_x(\cdot)$ is a convex function. Therefore,

$$\mathcal{F}_{\bar{x}}(d) + (1/2)\|d\|^2 \geq \mathcal{F}_{\bar{x}}(d_{\bar{x}}) + (1/2)\|d_{\bar{x}}\|^2 + (1/2)\epsilon^2 \quad \forall d \in S(\bar{x}).$$

Since $S(\bar{x})$ is a compact set and $\mathcal{F}_x(d)$ is a continuous function, $\exists \delta > 0$ such that

$$\mathcal{F}_x(d) + (1/2)\|d\|^2 > \mathcal{F}_x(d_{\bar{x}}) + (1/2)\|d_{\bar{x}}\|^2 \quad \forall d \in S, \text{ whenever } \|x - \bar{x}\| \leq \delta.$$

Since $\mathcal{F}_x(d) + (1/2)\|d\|^2$ is a convex function and d_x is the minimizer of $\mathcal{F}_x(d) + (1/2)\|d\|^2$, therefore, from the above inequality we can conclude that d_x does not lie in the region $\|d - d_{\bar{x}}\| > \epsilon$ for $\|x - \bar{x}\| \leq \delta$. Hence $\|d_x - d_{\bar{x}}\| \leq \epsilon$. This justifies that d_x is continuous.

For the second part, it is sufficient to prove that v_x is continuous on every compact subset of \mathbb{R}^n . Consider a function $\gamma_x^i : U \rightarrow \mathbb{R}$ at every $x \in \mathbb{R}^n$, where U is an arbitrary compact subset of $\mathbb{R}^n, i = 1, 2, 3, \dots, m$, as

$$\gamma_x^i(z) = \max_{A \in \partial \mathcal{F}(z)} \{(Ad_x)_i\} + \frac{1}{2}\|d_x\|^2$$

and

$$\Gamma_x(z) = \max\{\gamma_x^1(z), \gamma_x^2(z), \dots, \gamma_x^m(z)\},$$

where $z \in U$. For every $x \in \mathbb{R}^n, \gamma_x^i$ is a continuous function for each i and, hence Γ_x is a continuous function. Therefore for $\epsilon > 0, \exists \delta > 0$ such that for all

$y, z \in U, \|y - z\| < \delta \implies |\Gamma_x(y) - \Gamma_x(z)| < \epsilon$. From the Subproblem (1),

$$\begin{aligned} v_z &= \mathcal{F}_z(d_z) + \frac{1}{2}\|d_z\|^2 \leq \mathcal{F}_z(d_y) + \frac{1}{2}\|d_y\|^2, \text{ where } d_y \neq d_z \\ &= \max_{A \in \partial \mathcal{F}(z)} \{(Ad_y)_i : i = 1, 2, \dots, m\} + \frac{1}{2}\|d_y\|^2 \\ &= \Gamma_y(z) \leq \Gamma_y(y) + |\Gamma_y(z) - \Gamma_y(y)| \\ &\leq v_y + \epsilon, \text{ when } \|y - z\| < \delta \end{aligned}$$

That is, $v_z - v_y < \epsilon$ whenever $\|y - z\| < \delta$.

Similarly interchanging y and z in the above inequality, we get $v_y - v_z < \epsilon$ whenever $\|y - z\| \leq \delta$. Hence $|v_y - v_z| < \epsilon$ whenever $\|y - z\| < \delta$. So v_x is continuous. □

The subsequent outcome is proved to compute the step length α like Armijo type condition. This result justifies the descent property of every objective function.

Lemma 1 *Suppose \mathcal{F} be a convex vector function and $\beta \in (0, 1)$. If x is not a PC point, then there exists a positive value $\hat{\alpha}$ satisfies,*

$$\mathcal{F}(x + \alpha d_x) \leq \mathcal{F}(x) + \alpha \beta t_x \hat{1} \quad \forall \alpha \in (0, \hat{\alpha}),$$

where $\hat{1}$ is the vector of dimension $n \times 1$ and all of its entries are equal to 1.

Proof Given that x does not represent a PC point, the directional derivative for every objective function f_i at x along the direction d_x is

$$f'_i(x; d_x) = \lim_{\alpha \downarrow 0} \frac{f_i(x + \alpha d_x) - f_i(x)}{\alpha} = \max_{g_i \in \partial f_i(x)} g_i^T d_x < 0.$$

That is,

$$\lim_{\alpha \downarrow 0} \frac{f_i(x + \alpha d_x) - f_i(x)}{\alpha} = \max_{g_i \in \partial f_i(x)} g_i^T d_x \leq \beta \max_{g_i \in \partial f_i(x)} g_i^T d_x.$$

Hence $\exists \hat{\alpha} > 0$ such that

$$f_i(x + \alpha d_x) - f_i(x) \leq \alpha \beta \max_{g_i \in \partial f_i(x)} g_i^T d_x \quad \forall \alpha \in (0, \hat{\alpha}),$$

holds for every $i \in \Lambda_m$. Since (t_x, d_x) is the solution of the subproblem (P_x) at x , so $f_i(x + \alpha d_x) - f_i(x) \leq \alpha \beta t_x \quad \forall \alpha \in (0, \hat{\alpha})$ holds for each $i \in \Lambda_m$. This implies $\mathcal{F}(x + \alpha d_x) \leq \mathcal{F}(x) + \alpha \beta t_x \hat{1} \quad \forall \alpha \in (0, \hat{\alpha})$. □

In the remaining part of the manuscript, we utilize the notations: $d^k := d_{x^k}, t_k := t_{x^k}, A_k := A \in \partial \mathcal{F}(x^k)$ and $f_i^k := f_i(x^k)$.

Theorem 3 (Convergence property) *Suppose $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a convex vector function and $\{x^k\}$ be a sequence generated as $x^{k+1} = x^k + \alpha_k d^k$ starting with the initial point x^0 , d^k is an approximate solution of the Subproblem (1) and α_k corresponds to the first element of $\{\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots\}$ satisfying*

$$f_i^{k+1} \leq f_i^k + \alpha_k \beta t_k \quad \text{for all } i \in \Lambda_m, \tag{2}$$

for some $\beta \in (0, 1)$. Further, suppose $M^0 = \{x \in \mathbb{R}^n : \mathcal{F}(x) \leq \mathcal{F}(x^0)\}$ is convex and bounded, $f_i(x)$ is bounded below for every $i \in \Lambda_m$. Then every accumulation point of $\{x^k\}$ is a PC point of (MP).

Proof If Inequality (2) holds then \exists at least one $i \in \Lambda_m$ such that

$$\frac{f_i(x^k + \frac{1}{2\omega}d^k) - f_i(x^k)}{\frac{1}{2\omega}} > \beta t_k \text{ for sufficiently large } \omega.$$

Taking $\omega \rightarrow \infty$, $f'_i(x^k; d^k) \geq \beta t_k$. That is,

$$\max_{g_i^k \in \partial f_i(x^k)} g_i^{kT} d^k \geq \beta t_k \text{ for at least one } i \in \Lambda_m. \tag{3}$$

From the constraints of the Subproblem (P_x) we have $(A_k d^k)_i \leq t_k, \forall A_k \in \partial \mathcal{F}(x^k), \forall i \in \Lambda_m$. This implies $g_i^{kT} d^k \leq t_k \forall g_i^k \in \partial f_i(x^k), \forall i \in \Lambda_m$. Hence

$$\max_{g_i^k \in \partial f_i(x^k)} (g_i^{kT} d^k) \leq t_k \quad \forall i \in \Lambda_m.$$

Using the above relation and (3),

$$\begin{aligned} \beta \max_{g_i^k \in \partial f_i(x^k)} g_i^{kT} d^k &\leq \beta t_k \leq \max_{g_i^k \in \partial f_i(x^k)} g_i^{kT} d^k \text{ for at least one } i \in \Lambda_m \\ \Rightarrow (1 - \beta) \max_{g_i^k \in \partial f_i(x^k)} g_i^{kT} d^k &\geq 0 \text{ for at least one } i \in \Lambda_m. \end{aligned}$$

Hence $\max_{g_i^k \in \partial f_i(x^k)} g_i^{kT} d^k \geq 0$ for at least one $i \in \Lambda_m$. This contradicts $A_k(d^k) < 0 \forall A_k \in \partial \mathcal{F}(x^k)$.

Hence \exists some $\hat{\alpha} > 0$ such that $\hat{\alpha} \leq \alpha_k$ holds for each k . Since x^k is not PC point so $\max_{i \in \Lambda_m} (A_k(d^k))_i < 0$ and from Theorem 2-(3), $t_k \leq -\frac{1}{2} \|d^k\|^2$. Substituting this Inequality in (2),

$$f_i(x^{k+1}) \leq f_i(x^k) - \frac{1}{2} \beta \alpha_k \|d^k\|^2$$

$$\begin{aligned}
 &\leq f_i(x^{k-1}) - \frac{1}{2}\beta\alpha_{k-1} \|d^{k-1}\|^2 - \frac{1}{2}\beta\alpha_k \|d^k\|^2 \\
 &\dots \\
 &\leq f_i(x^0) - \frac{1}{2}\beta \sum_{j=0}^k \alpha_j \|d^j\|^2 \\
 &\Rightarrow f_i(x^0) - f_i(x^{k+1}) \geq \frac{1}{2}\beta \sum_{j=0}^k \alpha_j \|d^j\|^2. \tag{4}
 \end{aligned}$$

$\{f_i(x^k)\}$ is a monotonically decreasing sequence and bounded below. Let $\{f_i(x^k)\}$ converges to some f_i^* as $k \rightarrow \infty$. Taking $k \rightarrow \infty$ in Inequality(4). We have,

$$\infty > f_i(x^0) - f_i^* \geq \frac{1}{2}\beta \sum_{j=0}^{\infty} \alpha_j \|d^j\|^2 \geq \frac{1}{2}\beta \hat{\alpha} \sum_{j=0}^{\infty} \|d^j\|^2.$$

Suppose $\|d^j\| \neq 0$ for all j and $\inf_{j \in \mathbb{N}} \|d^j\| = m$. Then

$$\infty > f_i(x^0) - f_i^* \geq \frac{1}{2}\beta \hat{\alpha} m^2 \sum_{j=0}^{\infty} 1.$$

This is not possible since $\beta \neq 0$ and $m \neq 0$. So $d^k \rightarrow 0$ as $k \rightarrow \infty$. Hence $t_k \rightarrow 0$ as $k \rightarrow \infty$.

Since $t_k = \max\{(A_k d^k)_1, (A_k d^k)_2, (A_k d^k)_3, \dots, (A_k d^k)_m \ \forall A_k \in \partial \mathcal{F}(x^k)\}$, $\exists g_i^k \in \partial f_i(x^k)$ such that $g_i^{kT} d^k \rightarrow 0$ as $k \rightarrow \infty$. Since M^0 is a bounded set, $\{x^k\}$ must have at least one accumulation point. Hence \exists a sub sequence $\{x^{k_l}\}$ such that $x^{k_l} \rightarrow x^*$ and $\lim_{k_l \rightarrow \infty} d^{k_l} = 0$ and $\lim_{k_l \rightarrow \infty} t^{k_l} = 0$. That is, x^* is a PC point of MP. □

From Theorem 2-(2), one can conclude that if at x^k , $t_k \neq 0$, $d^k \neq 0$ then x^k is not a PC point. In this situation, the next iterating point should be calculated. Note that since v_{x^k} is non-positive, from Theorem 2-(3), one can conclude that either $v_{x^k} = 0$ (or $|v_{x^k}| = 0$) or $\mathcal{F}_x(d_x) = 0$. Then x^k is a PC point, and iterating process can stop here.

Observe that if the point x^k is not a PC point, then using the Theorem (2)-(3),

$$\max_i (A(d^k))_i \leq \max_i (A(d^k))_i + \frac{1}{2} \|d^k\|^2 < 0 \ \forall A \in \partial F(x^k).$$

Hence $T_k := \{\alpha = 1/2^j \mid j \in \mathbb{N}, \mathcal{F}(x^k + \alpha d^k) \leq \mathcal{F}(x^k) + \beta \alpha t_k \hat{1}\} \neq \emptyset$. When \mathcal{F} is a scalar-valued differentiable function, the determination of the step length is reduced to an Armijo-like rule. The steps of this method are demonstrated through the following examples.

4 Numerical examples

The theoretical results of the previous section can be summarized in the following steps.

- Starting at any point x_0 , a sequence $\{x^k\}$ is generated in the proposed methodology, which converges to a PC point. Convergence of the sequence is guaranteed in Theorem 3.
- (t_k, d^k) is the solution of the Subproblem (P_x) and α_k satisfies (2) at every iterating point x^k , corresponding to the first element of $\{\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots\}$ for some $0 < \beta < 1$. Thus, the iterative process can be generated as $x^{k+1} = x^k + \alpha_k d^k$.
- From Theorem 2, it is clear that if x^k is a PC point then the solution (t_{x^k}, d_{x^k}) of the Subproblem (P_x) at x^k satisfies $(t_{x^k}, d_{x^k}) = (0, 0)$ or $v_{x^k} = 0$. The converse part of this result can be justified as follows. From Theorem 2(3), one can conclude that if $v_{x^k} \geq 0$ or $F_{x^k}(d_{x^k}) \geq 0$ then x^k is a PC point. From Theorem 2(1), $v_{x^k} \leq 0$ for all x^k . Hence $v_{x^k} = 0$. This implies $(t_{x^k}, d_{x^k}) = (0, 0)$. Therefore, x^k is a PC point if and only if $v_{x^k} = 0$ or $(t_{x^k}, d_{x^k}) = (0, 0)$.

In a practical situation, an error term ϵ can be considered so that $|v_{x^k}| < \epsilon$.

Example 1 Consider an optimization problem: $\min_{x \in \mathbb{R}} \mathcal{F}(x)$, where $\mathcal{F}(x) = (f_1(x), f_2(x))^T$, $f_1(x) = |x|$ and $f_2(x) = |x| + |x + 1| + |x + 2|$. The objective functions are not differentiable everywhere. Hence this problem can not be solved by the existing line search techniques proposed in the papers, which are stated in the introduction section. From Fig. 1, it is clear that the PO set is $[-1, 0]$. Using the proposed methodology, this can be verified as follows.

Consider $x^0 = -2$.

\mathcal{F} is not differentiable at x^0 and $\partial \mathcal{F}(x^0) = \left\{ \begin{pmatrix} -1 \\ a \end{pmatrix} : a \in [-3, -1] \right\}$. To compute the descent direction d^0 , consider

$$(P_{x^0}) : \begin{cases} \min_{t, d \in \mathbb{R}} t + \frac{1}{2} \|d\|^2 \\ \text{s.t. } (Ad)_i < t \quad i = 1, 2 \quad \forall A \in \partial \mathcal{F}(-2) \end{cases},$$

which is similar to

$$\begin{cases} \min_{t, d \in \mathbb{R}} t + \frac{1}{2} \|d\|^2 \\ \text{s.t. } -d \leq t \text{ and } ad \leq t, \quad \forall a \in [-3, -1] \end{cases}$$

The subproblem is solved using the MATLAB command “optimproblem” after discretizing the value of a . The solution of this subproblem is found as $d_{x^0} = 1$ and $t_{x^0} = -1$ and the optimal value is $v(x^0) = -0.5$. Since $v(x^0) < 0$ so x^0 is not a PC point, which can also be verified from Figure 1 as in the neighbourhood of x^0 , both f_1 and f_2 are decreasing. For $\beta = 1/2$,

$$\alpha_0 = \max T_0 = \max \left\{ \frac{1}{2^j} \mid j \in \mathbb{N}, \mathcal{F}(x^0 + \frac{1}{2^j} d^0) \leq \mathcal{F}(x^0) + \beta \frac{1}{2^j} t^0 \hat{1} \right\}$$

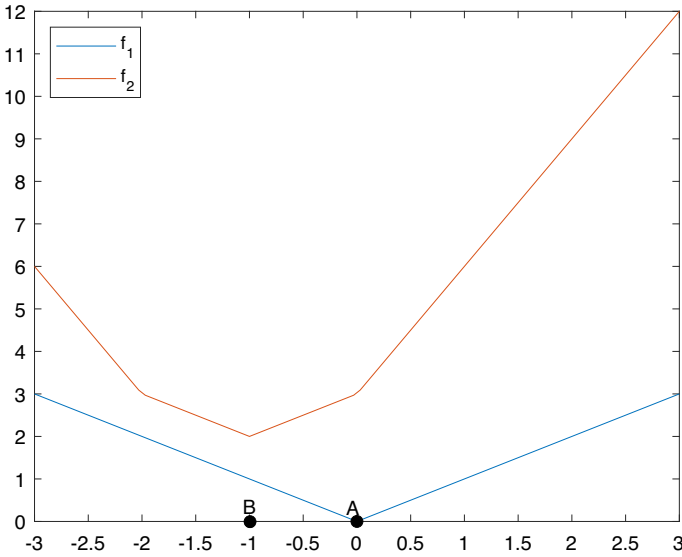


Fig. 1 Set of solutions of Example 1

$$= \max \left\{ \frac{1}{2^j} \mid j \in \mathbb{N}, \begin{pmatrix} 2 - \frac{1}{2^j} \\ 3 - \frac{1}{2^j} \end{pmatrix} \leq \begin{pmatrix} 2 - \frac{1}{2^{j+1}} \\ 3 - \frac{1}{2^{j+1}} \end{pmatrix} \right\} = \frac{1}{2}$$

$x^1 = x^0 + \alpha_1 d^1 = -1.5$ and $\mathcal{F}(x^1) = \begin{pmatrix} 1.5 \\ 2.5 \end{pmatrix}$. Here $\mathcal{F}(x^1) < \mathcal{F}(x^0)$. Other points are computed in a similar manner. At x^1 , both the functions are differentiable. $\partial \mathcal{F}(x^1) = \left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\}$. (P_{x^1}) is a convex quadratic problem which has unique solution $(t_1, d^1) = (-1, 1)$, optimal value is $v(x^1) = -0.5$ and $\alpha_1 = \frac{1}{2}$. Hence $x^2 = x^1 + \alpha_1 d^1 = -1$.

$$\text{At } x^2, \partial \mathcal{F}(x^1) = \left\{ \begin{pmatrix} -1 \\ a \end{pmatrix} \mid a \in [-1, 1] \right\} \text{ and } (P_{x^2}) = \begin{cases} \min_{t, d \in \mathbb{R}} t + \frac{1}{2} \|d\|^2 \\ \text{s.t. } -d \leq t \\ \text{and } ad \leq t, \forall a \in [-1, 1] \end{cases}.$$

Solution of this subproblem (P_{x^2}) is $(t_1, d^2) = (3.2195 \times 10^{-4}, 3.2195 \times 10^{-4})$ and optimal value is $v(x^2) \approx 3.2201 \times 10^{-4}$. Hence x^2 is an approximate PC point. One can also verify this from Fig. 1.

The set of all PC points and efficient frontier can be determined by repeating the same process with different initial points and a suitable spreading technique. In Fig. 2a, the PF generated by this methodology is compared with the PF generated by NSGA-II. One may observe that both are approximately the same.

Example 2 Consider the problem.

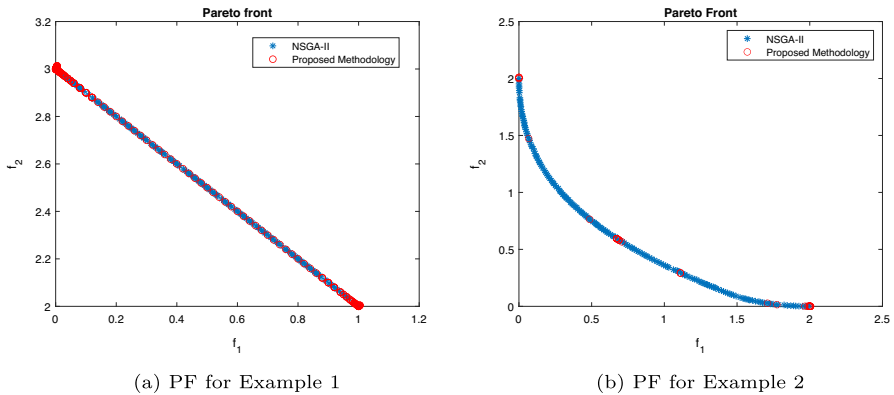


Fig. 2 Red circles and blue stars denote the PF using the proposed methodology and NSGA-II, respectively

$$\min_{x \in \mathbb{R}^2} \mathcal{F}(x) = \min_{x \in \mathbb{R}^2} ((x_1 - 1)^2 + (x_2 - 1)^2, x_1^2 + |x_2|) \text{ where } f_1(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \text{ and } f_2(x) = x_1^2 + |x_2|.$$

The steps of the proposed methodology for this example are summarized in the following Table 1.

Using the proposed methodology, the solution to this problem is found as $x^4 = \begin{pmatrix} 0.5911 \\ 0.64598 \end{pmatrix}$, starting with an initial point $x^0 = \begin{pmatrix} 1.5 \\ 0 \end{pmatrix}$. In Fig. 2b, the PF generated by this methodology is compared with the PF generated by NSGA-II. One may observe that both are approximately the same.

Some more problems are solved in a similar way using the proposed methodology and summarized in Table 2. The objective functions of these problems are borrowed from [26–28].

All these problems are nonsmooth and convex. The proposed method is implemented in MATLAB. To find the solution to the subproblem at every iteration, first, we created the subproblem using the MATLAB command 'optimproblem'. This is a semi-infinite programming problem, which is solved using the discretization process and then solved using the MATLAB command 'solve'. We have observed that the solution of the subproblem using the MATLAB command 'fsemif' provided the same solution set by solving the subproblem using the commands 'optimproblem' and 'solve' with discretization. The pre-specified constant β and the stopping parameter are taken $\frac{1}{2}$ and 10^{-3} respectively. In each case, we have verified that the PF is generated using the proposed methodology with the NSGA-II method.

5 Concluding remarks

A methodology for unconstrained MP, which uses the subdifferential information of the nonsmooth objective functions, is proposed in this paper. The iterative process develops a sequence, converging to the PC point under reasonable assumptions, preserving the descent property. The proposed methodology is applied in a set of

Table 1 Iterations of proposed methodology for Example 2

x^k	$\partial \mathcal{F}(x^k)$	(P_{x^k})	v_{x^k}	d^k	α^k
$x^0 = \begin{pmatrix} 1.5 \\ 0 \end{pmatrix}$	$\left\{ \begin{pmatrix} 1 & -2 \\ 3 & a \end{pmatrix} \mid a \in [-1, 1] \right\}$	$\begin{cases} \min_{r \in \mathbb{R}, d \in \mathbb{R}^2} r + \frac{1}{2} \ d\ ^2 \\ s.t. \ d_1 - 2d_2 \leq t \\ \text{and } 3d_1 + ad_2 \leq t, \ a \in [-1, 1] \end{cases}$	-1.89	$\begin{pmatrix} -1.615 \\ 1.077 \end{pmatrix}$	$\frac{1}{2}$
$x^1 = \begin{pmatrix} 0.6925 \\ 0.5385 \end{pmatrix}$	$\left\{ \begin{pmatrix} -0.615 & -0.923 \\ 1.385 & 1 \end{pmatrix} \right\}$	$\begin{cases} \min_{r \in \mathbb{R}, d \in \mathbb{R}^2} r + \frac{1}{2} \ d\ ^2 \\ s.t. \ -0.615d_1 - 0.923d_2 \leq t \\ 1.385d_1 + d_2 \leq t \end{cases}$	-0.0286	$\begin{pmatrix} -0.1657 \\ 0.1723 \end{pmatrix}$	$\frac{1}{2}$
$x^2 = \begin{pmatrix} 0.60965 \\ 0.62465 \end{pmatrix}$	$\left\{ \begin{pmatrix} -0.7807 & -0.7507 \\ 1.2193 & 1 \end{pmatrix} \right\}$	$\begin{cases} \min_{r \in \mathbb{R}, d \in \mathbb{R}^2} r + \frac{1}{2} \ d\ ^2 \\ s.t. \ -0.7807d_1 - 0.7507d_2 \leq t \\ 1.2193d_1 + d_2 \leq t \end{cases}$	-0.0013	$\begin{pmatrix} -0.0334 \\ 0.0381 \end{pmatrix}$	$\frac{1}{4}$
$x^3 = \begin{pmatrix} 0.6013 \\ 0.634225 \end{pmatrix}$	$\left\{ \begin{pmatrix} -0.7974 & -0.73155 \\ 1.2026 & 1 \end{pmatrix} \right\}$	$\begin{cases} \min_{r \in \mathbb{R}, d \in \mathbb{R}^2} r + \frac{1}{2} \ d\ ^2 \\ s.t. \ -0.7974d_1 - 0.73155d_2 \leq t \\ 1.12026d_1 + d_2 \leq t \end{cases}$	-0.000485	$\begin{pmatrix} -0.0204 \\ 0.0235 \end{pmatrix}$	$\frac{1}{2}$
$x^4 = \begin{pmatrix} 0.5911 \\ 0.64598 \end{pmatrix}$	$\left\{ \begin{pmatrix} -0.8178 & -0.70804 \\ 1.1822 & 1 \end{pmatrix} \right\}$	$\begin{cases} \min_{r \in \mathbb{R}, d \in \mathbb{R}^2} r + \frac{1}{2} \ d\ ^2 \\ s.t. \ -0.8178d_1 - 0.70804d_2 \leq t \\ 1.1822d_1 + d_2 \leq t \end{cases}$	$-\frac{2.6765}{10^5}$	$\begin{pmatrix} -0.0048 \\ 0.0056 \end{pmatrix}$	

Table 2 Numerical Examples

Problem	Objective functions	Initial point (x ₀)	$\mathcal{F}(x_0)$	No. of iterations (k)	x	$\mathcal{F}(x)$
1	$f_1(x) = x $ $f_2(x) = x + x + 1 + x + 2 $	-2	$\begin{pmatrix} 2 \\ 3 \end{pmatrix}$	3	-1	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$
2	$f_1(x) = (x_1 - 1)^2 + (x_2 - 1)^2$ $f_2(x) = x_1^2 + x_2 $	$\begin{pmatrix} 1.5 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1.25 \\ 2.25 \end{pmatrix}$	5	$\begin{pmatrix} 0.5911 \\ 0.64598 \end{pmatrix}$	$\begin{pmatrix} 0.2925 \\ 0.9954 \end{pmatrix}$
3	$f_1(x) = x_1 + x_2 + 2x_1$ $f_2(x) = x_1 + x_2 + 2x_2$	$\begin{pmatrix} 0 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$	2	$\begin{pmatrix} -0.5 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}$
4	$f_1(x) = \max\{(x_1 - 2)^2 + (x_2 + 2)^2, x_1^2 + 8x_2\}$ $f_2(x) = \max\{5x_1 + x_2, x_1^2 + x_2^2\}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 17 \\ 7 \end{pmatrix}$	3	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 0 \end{pmatrix}$
5	$f_1(x) = \max\{x_1^2 + (x_2 - 1)^2, (x_1 + 1)^2\}$ $f_2(x) = \max\{2x_1 + 2x_2, x_1^4 + x_2^2\}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	2	$\begin{pmatrix} -0.5 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.25 \\ 1.0625 \end{pmatrix}$
6	$f_1(x) = \max\{x_1^2 + x_2^4, (2 - x_1)^2 + (2 - x_2)^2, 2e^{(x_2 - x_1)}\}$ $f_2(x) = \max\{x_1^4 + x_2^2, (2 - x_1)^2 + (2 - x_2)^2, 2e^{(x_2 - x_1)}\}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 20 \\ 20 \end{pmatrix}$	9	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$
7	$f_1(x) = \max\{5x_1 + x_2, -5x_1 + x_2, x_1^2 + x_2^2 + 4x_2\}$ $f_2(x) = \max\{x_1^4 + x_2^2, (2 - x_1)^2 + (2 - x_2)^2, 2e^{(x_2 - x_1)}\}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 16 \\ 20 \end{pmatrix}$	9	$\begin{pmatrix} 0.562977 \\ 0.678975 \end{pmatrix}$	$\begin{pmatrix} 3.493862 \\ 3.810140 \end{pmatrix}$

Table 2 continued

Problem	Objective functions	Initial point (x_0)	$\mathcal{F}(x_0)$	No. of iterations (k)	x	$\mathcal{F}(x)$
8	$f_1(x) = \{x_1^4 + x_2^2, (2 - x_1)^2 + (2 - x_2)^2, 2e^{(x_2 - x_1)}\}$ $f_2(x) = \max\{-x_1 - x_2, -x_1 - x_2 + x_1^2 + x_2^2 - 1\}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 20 \\ 3 \end{pmatrix}$	3	$\begin{pmatrix} 0.875 \\ 0.875 \end{pmatrix}$	$\begin{pmatrix} 2.53125 \\ -1.21875 \end{pmatrix}$
9	$f_1(x) = \max\{-x_1 - x_2, -x_1 - x_2 + x_1^2 + x_2^2 - 1\}$ $f_2(x) = -x_1 + 20 \max\{x_1^2 + x_2^2 - 1, 0\}$	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 78 \end{pmatrix}$	15	$\begin{pmatrix} 0.8 \\ 0.4 \end{pmatrix}$	$\begin{pmatrix} -1.4 \\ -0.8 \end{pmatrix}$
10	$f_1(x) = \max\{2x_1 + 2x_2, x_1^4 + x_2^2\}$ $f_2(x) = \max\{5x_1 + x_2, x_1^2 + x_2^2\}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 20 \\ 12 \end{pmatrix}$	7	$\begin{pmatrix} -0.160347 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 6.6 \times 10^{-4} \\ 0.02571 \end{pmatrix}$
11	$f_1(x) = \max\{x_1^2 + (x_2 - 1)^2, (x_1 + 1)^2\}$ $f_2(x) = \max\{5x_1 + x_2, x_1^2 + x_2^2\}$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	3	$\begin{pmatrix} -0.375 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.390625 \\ 0.390625 \end{pmatrix}$
12	$f_1(x) = \max\{x, x^2\}$ $f_2(x) = x + x + 1 + x + 2 $	1	$\begin{pmatrix} 1 \\ 6 \end{pmatrix}$	2	0	$\begin{pmatrix} 0 \\ 3 \end{pmatrix}$
13	$f_1(x) = \max\{x_1^2 + x_2^2, x_1^2 + x_2^2 + 10(-4x_1 - x_2 + 4), x_1^2 + x_2^2 + 10(-x_1 - 2x_2 + 6)\}$ $f_2(x) = \max\{-x_1 - x_2, -x_1 - x_2 + (x_1^2 + x_2^2 - 1)\}$	$\begin{pmatrix} 2 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 20 \\ 13 \end{pmatrix}$	3	$\begin{pmatrix} 1.31448 \\ 2.21946 \end{pmatrix}$	$\begin{pmatrix} 9.11991 \\ 2.11991 \end{pmatrix}$
14	$f_1(x) = x_1 + x_1 + 1 + x_1 + 2 $ $f_2(x) = x_1 + x_2 $	$\begin{pmatrix} -2 \\ -2 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 4 \end{pmatrix}$	3	$\begin{pmatrix} -1 \\ -2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 3 \end{pmatrix}$

numerical examples. However, getting information on the subdifferential set of all types of functions isn't easy. Some computational processes should be developed to generate the subdifferential set to determine the descent direction explicitly. Generating a well-distributed Pareto front is also a challenge to this aspect. These are the possible future scope of the current work

Acknowledgements The authors wish to extend their heartfelt appreciation to the anonymous referees for their insightful comments and valuable feedback, which improved the quality and clarity of this paper.

Author Contributions Both the authors have equal contribution to this work.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest No potential conflict of interest have been declared by the authors.

References

1. Tammer, C., Weidner, P.: *Scalarization and Separation by Translation Invariant Functions*. Springer, Cham (2020)
2. Miettinen, K., Mäkelä, M.: Interactive bundle-based method for nondifferentiable multiobjective optimization: Nimbus. *Optimization* **34**(3), 231–246 (1995)
3. Miettinen, K., Mäkelä, M.M.: Comparative evaluation of some interactive reference point-based methods for multi-objective optimisation. *J. Oper. Res. Soc.* **50**(9), 949–959 (1999)
4. Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining convergence and diversity in evolutionary multiobjective optimization. *Evol. Comput.* **10**(3), 263–282 (2002)
5. Mostaghim, S., Branke, J., Schmeck, H.: Multi-objective particle swarm optimization on computer grids. In: *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pp. 869–875 (2007)
6. Coello, C.A.C.: *Evolutionary Algorithms for Solving Multi-objective Problems*. Springer, New York (2007)
7. Deb, K.: Recent advances in evolutionary multi-criterion optimization (emo). In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 702–735 (2017)
8. Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. *IEEE Trans. Evol. Comput.* **18**(4), 577–601 (2013)
9. Seada, H., Deb, K.: Non-dominated sorting based multi/many-objective optimization: Two decades of research and application. *Multi-Objective Optimization: Evolutionary to Hybrid Framework*, 1–24 (2018)
10. Fliege, J., Svaiter, B.F.: Steepest descent methods for multicriteria optimization. *Math. Methods Oper. Res.* **51**, 479–494 (2000)
11. Drummond, L.G., Svaiter, B.F.: A steepest descent method for vector optimization. *J. Comput. Appl. Math.* **175**(2), 395–414 (2005)
12. Fukuda, E.H., Drummond, L.G.: On the convergence of the projected gradient method for vector optimization. *Optimization* **60**(8–9), 1009–1021 (2011)
13. Drummond, L.G., Iusem, A.N.: A projected gradient method for vector optimization problems. *Comput. Optim. Appl.* **28**, 5–29 (2004)
14. Fliege, J., Drummond, L.G., Svaiter, B.F.: Newton's method for multiobjective optimization. *SIAM J. Optim.* **20**(2), 602–626 (2009)
15. Qu, S., Goh, M., Chan, F.T.: Quasi-newton methods for solving multiobjective optimization. *Oper. Res. Lett.* **39**(5), 397–399 (2011)
16. Mishra, S.K., Panda, G., Ansary, M.A.T., Ram, B.: On q-newton's method for unconstrained multiobjective optimization problems. *J. Appl. Math. Comput.* **63**, 391–410 (2020)

17. Lai, K.K., Mishra, S.K., Panda, G., Ansary, M.A.T., Ram, B.: On q-steepest descent method for unconstrained multiobjective optimization problems. *AIMS Math.* **5**(6), 5521–5540 (2020)
18. Kumar, S., Ansary, M.A.T., Mahato, N.K., Ghosh, D., Shehu, Y.: Newton's method for uncertain multiobjective optimization problems under finite uncertainty sets. *J. Nonlinear Var. Anal.* **7**(5) (2023)
19. Fliege, J., Vaz, A.I.F.: A method for constrained multiobjective optimization based on sqp techniques. *SIAM J. Optim.* **26**(4), 2091–2119 (2016)
20. Ansary, M.A.T., Panda, G.: A sequential quadratically constrained quadratic programming technique for a multi-objective optimization problem. *Eng. Optim.* **51**(1), 22–41 (2019)
21. Ansary, M.A.T., Panda, G.: A globally convergent sqcqp method for multiobjective optimization problems. *SIAM J. Optim.* **31**(1), 91–113 (2021)
22. Tanabe, H., Fukuda, E.H., Yamashita, N.: Proximal gradient methods for multiobjective optimization and their applications. *Comput. Optim. Appl.* **72**, 339–361 (2019)
23. Bento, G.D.C., Cruz Neto, J.X., López, G., Soubeyran, A., Souza, J.: The proximal point method for locally lipschitz functions in multiobjective optimization with application to the compromise problem. *SIAM J. Optim.* **28**(2), 1104–1120 (2018)
24. Da Cruz Neto, J.X., Da Silva, G., Ferreira, O.P., Lopes, J.O.: A subgradient method for multiobjective optimization. *Comput. Optim. Appl.* **54**(3), 461–472 (2013)
25. Ansary, M.A.T.: A newton-type proximal gradient method for nonlinear multi-objective optimization problems. *Optim. Methods Softw.* **38**, 570–590 (2023)
26. Lukšan, L., Váček, J.: Test problems for nonsmooth unconstrained and linearly constrained optimization. Technical report (2000)
27. Montonen, O., Karmitsa, N., Mäkelä, M.: Multiple subgradient descent bundle method for convex nonsmooth multiobjective optimization. *Optimization* **67**(1), 139–158 (2018)
28. Gebken, B., Peitz, S.: An efficient descent method for locally lipschitz multiobjective optimization problems. *J. Optim. Theory Appl.* **188**, 696–723 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.