J. Nonlinear Var. Anal. 11 (2027), No. 1, pp. 5-28
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.11.2027.1.02

ℓ_0-NORM MULTIOBJECTIVE OPTIMIZATION MODELS MOTIVATED BY APPLICATIONS TO PROTON THERAPY

XIAODA CONG[1], XUANFENG DING[1], BORIS MORDUKHOVICH[2,*], ANH VU NGUYEN[2], LEWEI ZHAO[3]

[1]Proton Therapy Center, Corewell William Beaumont Hospital, Royal Oak, MI 48073, USA
[2]Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
[3]Department of Radiation Oncology, MedStar Georgetown University Hospital, Washington DC, 20007, USA

Dedicated to Professor Christiane Tammer in honor of her 70th birthday

Abstract. The paper is devoted to investigating single-objective and multiobjective optimization problems involving the ℓ_0-norm function, which is nonconvex and nondifferentiable. Our motivation comes from proton beam therapy models in cancer research. The developed approach uses subdifferential tools of variational analysis and the Gerstewitz (Tammer) scalarization function in multiobjective optimization. Based on this machinery, we propose several algorithms of the subgradient type and conduct their convergence analysis. The obtained results are illustrated by numerical examples, which reveal some characteristic features of the proposed algorithms and their interactions with the gradient descent.

Keywords. Multiobjective optimization; Gerstewitz scalarization function; Nonsmooth and nonconvex optimization; Proton beam therapy; Subgradient algorithms; Variational analysis and generalized differentiation.

1. INTRODUCTION

This paper revolves around variational analysis and optimization of single-objective and multiobjective models involving the so-called ℓ_0-*norm function* $\|\cdot\|_0$, which signifies the number of nonzero components in a given vector. The function $\|\cdot\|_0$ (it is not actually a norm) is nondifferentiable at the origin being also nonconvex. Therefore, the standard machinery of classical and convex analysis is not applicable to the study of the ℓ_0-norm function, which stimulate us to use for these purposes advanced tools of variational analysis and generalized differentiation dealing with nonsmoothness and nonconvexity.

A profound interest to optimization problems involving the ℓ_0-norm function has arisen in more recent years from applications to machine learning, signal processing, compressed sensing, etc.; see, e.g., [1] and the references therein. It has been realized that such problems are *NP-hard* and reflect the *sparsity* in optimization, which is a challenging issue. To this end, we refer the reader to the primal-dual active set with continuation (PDASC) algorithm developed in [2] for solving the ℓ_0-regularized least-squares problem that frequently arises in compressed

*Corresponding author.
E-mail address: aa1086@wayne.edu (B. Mordukhovich)
Received 2 September 2025; Accepted 22 March 2026; Published online 10 November 2026.
©2027 Journal of Nonlinear and Variational Analysis

sensing. A practical method to solve ℓ_0-norm problems for neural networks was developed in [3]. In [4], ℓ_0-sparse optimization was used in minimizing the number of features in deep learning.

In recent years, the ℓ_0-norm function has gained significant attention in the field of medical physics. Using the ℓ_0-norm encourages sparsity and enables the extraction of meaningful information from the noisy one and/or reducing the unnecessary delivery, which plays a pivotal role to ensure accurate diagnostics and patient treatment efficiency. The study in [5] compared ℓ_p-regularization for $p = 0, 1/2, 2/3, 1$ in image-domain multimaterial decomposition for dual-energy computed tomography. It is shown therein that the smaller is p, the more nonconvex will be the problem, and thus it is more difficult to find optimal sparsity solutions. In particular, ℓ_0-sparsity optimization techniques have been leveraged in radiation therapy.

Proton therapy is an advanced type of radiation treatment for cancer that uses high-energy protons to precisely kill tumor, while minimizing damage to surrounding healthy tissues. In proton therapy, especially proton arc therapy, treatment delivery efficiency plays a crucial role in the routine clinical implementation. Previous preliminary investigations indicated that the number of energy layers and spot impact not only affects plan quality but also the treatment delivery time. Therefore, finding an optimal sparsity level of the energy layer or spot to ensure fast treatment delivery while maintaining clinically acceptable plan quality remains a significant challenge. In terms of the energy layer sparsity solution, Gu et al. [6] formulated an optimization problem by integrating an $\ell_{1/2}$-regularization term for energy layer selection. Meanwhile, in the direction of spot number sparsity, [7] introduce the ℓ_0-norm concept to reduce the unnecessary spots in the complicated SPArc treatment plan. Recently, the alternating direction method of multipliers has been developed and compared with the PDASC method, which shows promising results reported in [8].

Proton arc therapy delivers proton beams while the treatment machine rotates around the patient. A challenge to proton arc therapy is how to reduce beam delivery time. One approach is to decrease spot number. The alternating direction method of multipliers and the aforementioned PDASC were applied in [9], [8], and [10] from the viewpoint of ℓ_0-sparsity optimization in proton arc therapy to reduce the corresponding spot number. These strategies prioritize the principle of sparsity while assuming that optimal solutions are either nearly sparse, or can be made sparse through an appropriate transformation.

Reducing spot number will reduce degree of freedom, which may degrade the plan quality. Thus minimizing beam delivery time and optimizing plan quality becomes two conflict goals, making the model a multicriterial optimization problem. Motivated by such practical multiobjective models arising in proton beam therapy that are unavoidably contain the ℓ_0-norm function and the like, we aim here to provide some simplified descriptions of such model and then design novel *subgradient-based algorithms* to find their *efficient/Pareto optimal* solutions by using the precise calculation of the *basic/limiting subdifferential* of the ℓ_0-norm function in the sense of Mordukhovich; see [11, 12] and the references therein. Our strategy is to start with a single-objective optimization problem of minimizing the sum of a convex differentiable function and the ℓ_0-norm. Then we formulate a multiobjective optimization problem whose vector objective consists of scalar components of the above type. To deal with the latter problem, we employ the two scalarization techniques: the standard *weight-sum* approach and the (more general) *Gerstewitz scalarization* [13, 14]. In this way, the designed algorithm in the scalar ℓ_0-norm

case induces the corresponding algorithms for the multiobjective problem in question by using the aforementioned two scalarization techniques. A detailed convergence analysis is conducted fore all the proposed algorithms. Numerical calculations by using the code developed in the Proton Beam Therapy Group of the William Beaumont Hospital demonstrate the reliability and efficacy of our algorithms and thus signify a promising direction of our study to handle multi-faceted optimization problems with potential wide-ranging impacts, particularly in the areas of cancer research and medical physics.

The rest of the paper is structured as follows. Section 2 presents basic definitions of the normal cone, subdifferential, and ℓ_0-norm functions with brief discussions of their underlying properties. In Section 3, we provide a complete calculation of the basic subdifferential of the ℓ_0-norm function important for the subsequent algorithmic design. Section 4 provides the formulation of both single-objective and multiobjective optimization problems by using weight-sum scalarization. In Section 5, we explore some properties of the Gerstewitz scalarization function needed for our algorithmic developments, This allows us to formulate and investigate in Section 6 the scalarized version of ℓ_0 multiobjective optimization problem involving the Gerstewitz function with an ℓ_0 addition. Based on the above, the algorithms for ℓ_0 scalar and multiobjective optimization are designed in Section 7. Their detailed convergence analysis is conducted in Section 8. In Section 9, we implement the designed algorithms in numerical calculations for typical examples and discuss their characteristic features. Finally, Section 10 summarizes the main achievements of the paper and lists some topics of our future research.

2. Basic Definitions and Discussions

Using a geometric approach to variational analysis and generalized differentiation [11, 15], we define first generalized normals to sets. Given a nonempty set $\Omega \subset \mathbb{R}^n$, the (Fréchet) *regular normal cone* to Ω at $\bar{x} \in \Omega$ is given by

$$\widehat{N}(x;\Omega) := \left\{ v \in \mathbb{R}^n \;\middle|\; \limsup_{u \xrightarrow{\Omega} x} \frac{\langle v, u-x\rangle}{\|u-x\|} \le 0 \right\}, \tag{2.1}$$

where $u \xrightarrow{\Omega} x$ means that $u \to x$ with $u \in \Omega$. If $x \notin \Omega$, we put $\hat{N}(x,\Omega) := \emptyset$. The (Mordukhovich) *basic/limiting normal cone* $N(\bar{x};\Omega)$ to Ω at $\bar{x} \in \Omega$ is defined by

$$N(\bar{x};\Omega) := \Big\{ v \in \mathbb{R}^n \;\Big|\; \begin{array}{l} \exists x_k \xrightarrow{\Omega} \bar{x},\ \exists v_k \to v \ \text{ as } \ k \to \infty \\ \text{such that } \ v_k \in \widehat{N}(x_k;\Omega) \ \text{ for all } \ k \in \mathbb{N} := \{1,2,\} \end{array} \Big\} \tag{2.2}$$

with $N(\bar{x};\Omega) := \emptyset$ whenever $\bar{x} \notin \Omega$. Both normals cones in (2.1) and (2.2) reduce to the classical normal cone of convex analysis

$$N(\bar{x};\Omega) = \big\{ v \in \mathbb{R}^n \;\big|\; \langle v, x-\bar{x}\rangle \le 0 \ \text{ for all } \ x \in \Omega \big\}$$

if the set Ω is convex. If Ω is a nonconvex set, then properties of the limiting normal cone (2.2) are much better than those for its regular counterpart (2.1), which may be even trivial (i.e., $\widehat{N}(\bar{x};\Omega) = \{0\}$) while $\bar{x}$ is a boundary point as for the set $\Omega := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \ge -|x_1|\}$ at the origin $\bar{x} = (\bar{x}_1,\bar{x}_2) = (0,0)$.

Consider next an extended real-valued function $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$ and the associated *domain* and *epigraph* sets of φ given by

$$\operatorname{dom}\varphi := \big\{x \in \mathbb{R}^n \;\big|\; \varphi(x) < \infty\big\} \;\text{ and }\; \operatorname{epi}\varphi := \big\{(x,\mu) \in \mathbb{R}^{n+1} \;\big|\; \mu \geq \varphi(x)\big\},$$

respectively. We define the (Mordukhovich) *basic/limiting subdifferential* of φ geometrically via the basic normal cone (2.2) to the epigraph by

$$\partial\varphi(\bar{x}) := \big\{v \in \mathbb{R}^n \;\big|\; (v,-1) \in N\big((\bar{x},\varphi(\bar{x})); \operatorname{epi}\varphi\big)\big\} \tag{2.3}$$

if $\bar{x} \in \operatorname{dom}\varphi$ and $\partial\varphi(\bar{x}) := \emptyset$ otherwise. There are various analytic representations of (2.3), which can be found in the books [11, 12, 15] and the references therein. These books, as well as the quite recent one [16], contain comprehensive calculus rules and other results for the subdifferential (2.3) and related constructions with a broad variety of applications that are mainly based on the variational/extremal principles of variational analysis. Note that the subgradient set (2.3) reduces to the classical gradient $\{\nabla\varphi(\bar{x})\}$ for smooth functions and to the subdifferential of convex analysis if φ is convex.

The primary object of our study and applications in this paper is the following real-valued function defined on finite-dimensional spaces.

Definition 2.1. The ℓ_0-NORM FUNCTION $\|\cdot\|_0 : \mathbb{R}^n \to \mathbb{R}$ is given by

$$\|x\|_0 := \text{number of nonzero components of } x. \tag{2.4}$$

This function is clearly nondifferentiable and nonconvex, and we intend to utilize its basic subdifferential (2.3) for its study and applications to optimization model. To begin with, let us illustrate the calculation of (2.4) in the three-dimensional space.

Example 2.1. For the ℓ_0-norm function (2.4) on $\mathbb{R}^3$, we have

$$\|(0,-1,4)\|_0 = 2,\; \|(1,0,0)\|_0 = 1,\; \|(0,0,0)\|_0 = 0,\; \|(1,2,3)\|_0 = 3.$$

To further illustrate (2.4), we interpret it as a piecewise function (say, in $\mathbb{R}^2$) defined by

$$f(x,y) := \begin{cases} 0 & \text{if} \quad x = y = 0, \\ 1 & \text{if either} \quad x \neq 0,\; y = 0, \quad \text{or} \quad \text{if} \quad x = 0,\; y \neq 0, \\ 2 & \text{otherwise.} \end{cases} \tag{2.5}$$

Based on (2.5), the graph of the ℓ_0-norm function on $\mathbb{R}^2$ is

$$\operatorname{gph}\|\cdot\|_0 = \big\{(0,0,0)\big\} \bigcup \big\{(x,0,1) \mid x \neq 0\big\} \bigcup \big\{(0,y,1) \mid y \neq 0\big\} \bigcup \big\{(x,y,2) \mid x,y \neq 0\big\},$$

which is depicted in Figure 1.

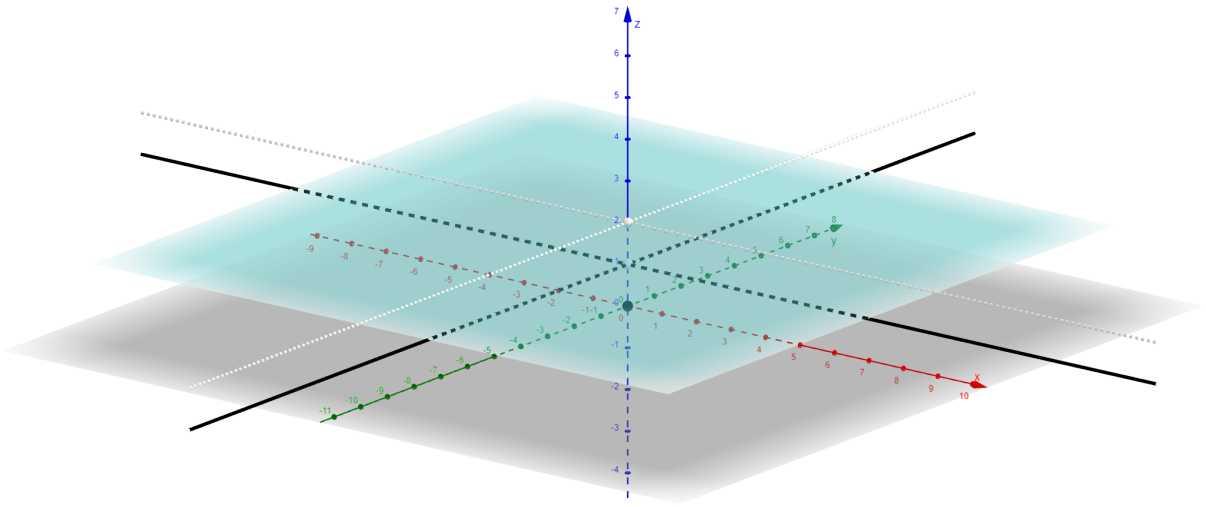


FIGURE 1. Graph of the ℓ_0-norm function

3. SUBGRADIENT CALCULATION FOR THE ℓ_0-NORM FUNCTION

In this section, we completely calculate the basic subdifferential (2.3) of the ℓ_0-norm function (2.4) at any point of an arbitrary finite-dimensional space.

Theorem 3.1. *The subdifferential of* $\|\cdot\|_0$ *at any* $\bar{x}=(\bar{x}_1,\ldots,\bar{x}_n)\in\mathbb{R}^n$ *is calculated by*

$$\partial\|\cdot\|_0(\bar{x})=\big\{v=(v_1,\ldots,v_n)\in\mathbb{R}^n\big\},$$

$$\text{where } v_i \text{ are defined as } \begin{cases} v_i=0 & \text{if } \bar{x}_i\neq 0,\\ v_i\in\mathbb{R} & \text{if } \bar{x}_i=0. \end{cases}$$

Proof. For simplicity, we verify the claimed formula for the function $\|\cdot\|_0$ defined on $\mathbb{R}^2$. Based on the definitions in (2.3) and (2.4), split the proof into the following cases:

Case 1: $\bar{x}=(\bar{x}_1,\bar{x}_2)$ *with* $\bar{x}_1,\bar{x}_2\neq 0$. Taking a sequence $x_{1,k}x_{2,k}\to\bar{x}_1,\bar{x}_2$, we get that $x_{1,k},x_{2,k}\neq 0$ when $x_{1,k},x_{2,k}$ are sufficiently close to $\bar{x}_1,\bar{x}_2$. This tells us by (2.4) that $\|(x_{1,k},x_{2,k})\|_0=2$. Picking now a sequence $((x_{1,k},x_{2,k},2)\to(\bar{x}_1,\bar{x}_2,2)$ in epi$\|\cdot\|_0$, we deduce from the definitions of the subdifferential (2.3) and the basic normal cone (2.2) that

$$\limsup_{(x_{1,k},x_{2,k})\to(\bar{x}_1,\bar{x}_2)}\frac{\langle(v_1,v_2,-1),(x_{1,k},x_{2,k},2)-(\bar{x}_1,\bar{x}_2,2)\rangle}{\|(x_{1,k},x_{2,k})-(\bar{x}_1,\bar{x}_2)\|}\leq 0,$$

which is equivalently written as

$$\limsup_{(x_{1,k},x_{2,k})\to(\bar{x}_1,\bar{x}_2)}\frac{\langle(v_1,v_2),(x_{1,k},x_{2,k})-(\bar{x}_1,\bar{x}_2)\rangle}{\|(x_{1,k},x_{2,k})-(\bar{x}_1,\bar{x}_2)\|}\leq 0.$$

Choosing $x_{i,k}=\bar{x}_i$ for $i=1,2$ gives us

$$\limsup_{x_{j,k}\to\bar{x}_j}\frac{v_j(x_{j,k}-\bar{x}_j)}{\|x_{j,k}-\bar{x}_j\|}\le 0 \ \text{ for } \ j\neq i,$$

which readily implies that $v_j=0$, and thus $v=(0,0)$ as claimed.

Case 2: $\bar{x}=(0,\bar{x}_2)$ *with* $\bar{x}_2\neq 0$. Choosing $x_{1,k}=0$ allows us to conclude, similarly to the proof in Case 1, that $v_2=0$. Now let $x_{2,k}=\bar{x}_2$, and let $\{x_{1,k}\}$ be a a sequence on nonzero numbers converging to 0. Then we have $\|(x_{1,k},x_{2,k})\|_0=2$ and

$$\limsup_{(x_{1,k},x_{2,k})\to(\bar{x}_1,\bar{x}_2)}\frac{\langle(v_1,v_2,-1),(x_{1,k},x_{2,k},2)-(\bar{x}_1,\bar{x}_2,2)\rangle}{\|(x_{1,k},x_{2,k})-(\bar{x}_1,\bar{x}_2)\|}\le 0,$$

which can be equivalently rewritten as

$$\limsup_{(x_{1,k},x_{2,k})\to(\bar{x}_1,\bar{x}_2)}\frac{\langle(v_1,v_2,-1),(x_{1,k},x_{2,k},2)-(\bar{x}_1,\bar{x}_2,1)\rangle}{\|(x_{1,k},x_{2,k})-(\bar{x}_1,\bar{x}_2)\|}\le 0$$

and brings us therefore to the inequality

$$\limsup_{x_{1,k}\to\bar{x}_1}\frac{v_1(x_{1,k}-\bar{x}_1)-1}{\|x_{1,k}-\bar{x}_1\|}\le 0.$$

For any v_1, we choose k to be so large that $v_1(x_{1,k}-\bar{x}_1)-1\le 0$, which ensures that $\{v=(v_1,v_2)\}=\{(v_1,0)\mid v_1\in\mathbb{R}\}$ as claimed in this case.

Case 3: $\bar{x}=(\bar{x}_1,0)$. The calculation of $\partial\|\cdot\|_0(\bar{x})$ is similar to the proof in Case 2.

Case 4: $\bar{x}=(0,0)$. The proof is similar to Case 2 when we fix $x_{i,k}=0$ for either $i=1$ or $i=2$. This completes the proof of the theorem. □

4. Scalar and Vector Problems of ℓ_0 Optimization

First we formulate here the following single-objective problem of optimization involving the ℓ_0-norm function. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a smooth and convex function whose gradient is Lipschitz continuous with a constant L. Consider the ℓ_0 *optimization problem*:

$$(\ell_0\text{NOP}) \qquad \min_{x\in\mathbb{R}^n} f(x)+\|x\|_0.$$

The objective function in (ℓ_0NOP) is nonconvex and nondifferentiable. Let us illustrate its behavior by the following example.

Example 4.1. Consider the quadratic function $f:\mathbb{R}^2\to\mathbb{R}$ defined by $f(x,y):=(x-\frac{1}{2})^2+\frac{1}{4}(y+1)^2$. The graph of $f(x,y)+\|(x,y)\|_0$ is depicted in Figure 2.

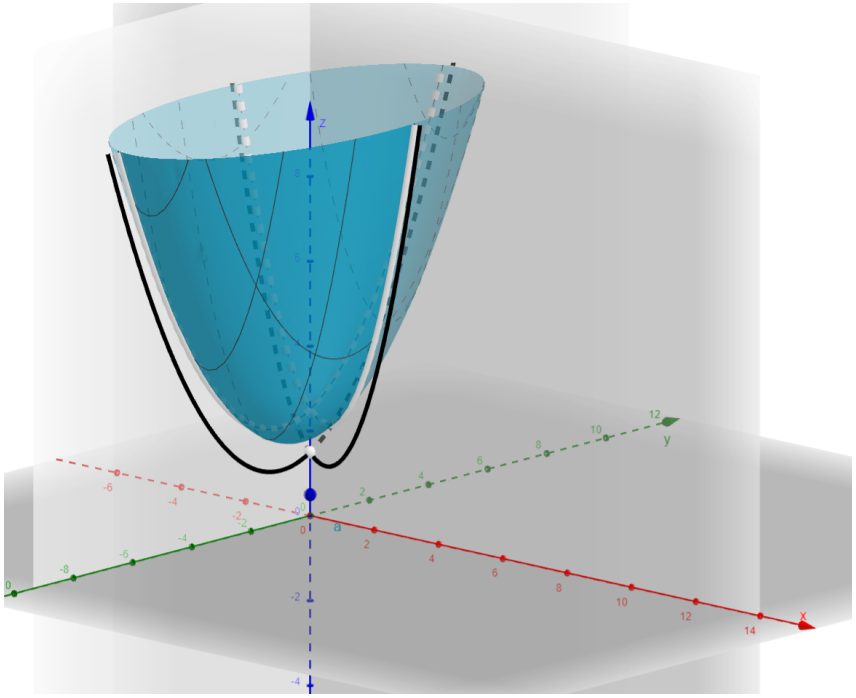


FIGURE 2. Graph of $f(\cdot)+\|\cdot\|_0$

As seen from Figure 2, the objective function in Example 4.1 has multiply local minimizers. Our goal concerning problem (ℓ_0NOP) is design an efficient subgradient algorithm, which allows us to find at least one *local optimal solution* to this problem. To proceed in this direction, we first construct, based on the subdifferential calculation in Theorem 3.1, a certain *projection matrix* associated with $x \in \mathbb{R}^n$ over the hyperplane where some of the components of x are zero. Define the *projection mapping* $\tilde{I}_i : \mathbb{R}^n \to \mathbb{R}^n$ as follow: for $x_i = (x_{1,i}, \ldots, x_{n,i}) \in \mathbb{R}^n$, let $\tilde{I}_i$ be an $n \times n$ matrix such that

$$\begin{aligned} a_{j,k} &= 0 \quad \text{if } j \neq k, \\ a_{j,j} &= 0 \quad \text{if } x_{j,i} = 0, \\ a_{j,j} &= 1 \quad \text{if } x_{j,i} \neq 0. \end{aligned}$$

The next example illustrates the construction of the projection matrix in $\mathbb{R}^3$.

Example 4.2. For $x_i = (0, 2, -3)$, we have

$$\tilde{I}_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Observe that the projection mapping $\tilde{I}_i$ is defined by the value of x_i with $\tilde{I}_i x_i = x_i$. For each projection mapping, we have the corresponding *projection hyperplane* $I_i := \tilde{I}_i(x_1, \ldots, x_n)$. This does not make any change in the value of x_i while helping us to determine the direction for each iteration of our algorithms designed in Section 7.

The following important result allows us to find a local minimizer of problem (ℓ_0NOP) by minimizing the smooth and convex function f therein over the corresponding projection hyperplane generated by the subgradient calculation for the ℓ_0-norm function.

Theorem 4.1. *If $\bar{x}$ solves the* PROJECTION MINIMIZATION PROBLEM

$$\text{(PP)} \qquad \min_{x\in I_i} \quad f(x),$$

then $\bar{x}$ is a local minimizer of the original ℓ_0 optimization problem (ℓ_0OP).

Proof. Picking $x \in I_i$ and taking into account that $\bar{x}$ minimizes f over the hyperplane I_i ensures that $\|\bar{x}\|_0 = \|x\|_0$ implies that $f(\bar{x}) + \|\bar{x}\|_0 \le f(x) + \|x\|_0$. If $x \notin I_i$ being sufficiently close to $\bar{x}$, then it follows for each component $\bar{x}_j \neq 0$ of $\bar{x}$ that $x_i \neq 0$. In the case where $\bar{x}_i = 0$, there clearly exist some nonzero components of x_i. This yields

$$\|\bar{x}\|_0 \le \|x\|_0 - 1.$$

By the continuity of f, for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(\bar{x})| < \varepsilon \ \text{ whenever } \ \|x - \bar{x}\| < \delta,$$

and therefore $f(\bar{x}) < f(x) + \varepsilon$. In this way, we arrive at the estimates

$$\begin{aligned} f(\bar{x}) + \|\bar{x}\|_0 &\le f(x) + \|x\|_0 + \varepsilon - 1 \\ &\le f(x) + \|x\|_0 \end{aligned}$$

and thus complete the proof of the theorem. □

Now we formulate ℓ_0 multiobjective optimization problem, which is a multiobjective/vector version of the scalar ℓ_0 optimization problem (ℓ_0NOP). Consider the vector mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ with $F(x) := \{F_1(x), \ldots, F_m(x)\}$, where each $F_i(x) := f_i(x) + \|x\|_0$ as $i = 1, \ldots, m$ is such that f_i are convex differentiable functions with Lipschitz continuous gradients. Denote by L_i the Lipschitz constant of the corresponding gradient and set $L := \max\{L_i \mid i = 1, \ldots, m\}$. The ℓ_0 *multiobjective optimization problem* is defined by

$$\text{(MO-}\ell_0\text{NP)} \qquad \min_{x\in\mathbb{R}^n} F(x) = (F_1(x), \ldots, F_m(x)),$$

where 'min' is understood in the sense of *Pareto optimality/efficiency*.

Definition 4.1. A vector $\bar{x} \in \mathbb{R}^n$ is called a LOCAL PARETO OPTIMAL SOLUTION to the multiobjective optimization problem (MO-ℓ_0NP) if there exists no other point x in a neighborhood U of $\bar{x}$ for which we have

$$F_i(x) \le F_i(\bar{x}), \quad i = 1, \ldots, m,$$

and

$$F_j(x) < F_j(\bar{x}) \quad \text{for some} \quad j \in \{1, \ldots, m\}.$$

A natural way to study multiobjective optimization problems is to use a certain *scalarization*, which converts the multiobjective problem in question to some problem of scalar (single-objective) optimization. the most simple scalarization approach is employing the *weight-sum scalarization*. For problem (MO-ℓ_0NP), such a scalarization looks as follows:

$$\min_{x\in\mathbb{R}^n} \qquad \text{weight sum}(x) := \sum_{i=1}^{m} \omega_i F_i(x) \tag{4.1}$$

with the weight coefficients satisfying $\omega_i \ge 0$ and $\sum_{i=1}^m \omega_i = 1$. It is easy to observe the following relation between optimal solutions of the multiobjective and scalarized problems.

Proposition 4.1. *A local minimizer of the weight sum function in* (4.1) *is a local Pareto optimal solution to the multiobjective optimization problem* (MO-ℓ_0NP).

Proof. Pick an arbitrary local minimizer $\bar{x}$ in (4.1) and show that it is a local Pareto solution to the multiobjective problem (MO-ℓ_0NP). Assuming the contrary, we find a vector x in some neighborhood U of $\bar{x}$ satisfying the conditions

$$\begin{aligned} F_i(x) &\le F_i(\bar{x}) \quad \text{for all} \quad i = 1,\ldots,m,\\ F_i(x) &< F_i(\bar{x}) \quad \text{for some} \quad i. \end{aligned}$$

This readily implies, by using the properties of the weight coefficients in (4.1) that

$$\text{weight sum}(x) < \text{weight sum}(\bar{x}).$$

The latter contradicts the local optimality of $\bar{x}$ in (4.1) and thus completes the proof. □

Using the weight condition $\sum_{i=1}^m \omega_i = 1$ and the form of the components F_i in (MO-ℓ_0NP), the weight sum function in (4.1) can be written as

$$\text{weight sum}(x) = \sum_{i=1}^m \omega_i f_i(x) + \|x\|_0,$$

which shows that (4.1) is a problem of scalar ℓ_0 optimization of type (ℓ_0NOP).

5. Gerstewitz Scalarization Function

In this section, we start exploring another approach to scalarization the ℓ_0 multiobjective optimization problem (MO-ℓ_0NP) that is based on using the *Gerstewitz scalarization function* introduced in [13] and then developed in many publications; see, e.g., [14] and the references therein. The Gerstewitz approach is more general than the weight-sum one and allows us, in particular, to handle the multiobjective problem (MO-ℓ_0NP) with *Lipschitz continuous* (not just smooth) functions f_i, which is important for the design and justification of subgradient-type methods vs. the gradient descent. To begin with, we review several properties of the Gerstewitz function needed for developing our algorithms. Some of these properties are known, but nevertheless their simplified proofs are provided for completeness and the reader's convenience. Our main novelty here is the choice and verification of the *directions* in the Gerstewitz function in order to adjust her scalarization technique to the nonsmooth and nonsmooth ℓ_0-norm setting in multiobjective ℓ_0 optimization, which will implemented in the subsequent sections.

Definition 5.1. Given a closed convex set $C \subset \mathbb{R}^m$, a nonzero direction $k^0 \in C\backslash(-C)$, and an nonempty set $A \subset \mathbb{R}^m$ with $A - \mathbb{R}_+ k^0 \subset A$, the Gerstewitz function $\phi_A := \phi_{A,k^0} : \mathbb{R}^m \to \overline{\mathbb{R}}$, associated with A and k^0 is defined by

$$\phi_{A,k^0} := \inf\big\{t \in \mathbb{R} \mid y \in tk^0 + A\big\}. \tag{5.1}$$

Note that the choice of the set A in Definition 5.1 provides some flexibility in scalarization and contains the weight-sum scalarization (4.1) with different weight coefficients as special cases. In particular, the case where $A = \{(y_1,y_2) \in \mathbb{R}^2 \mid y_1 + y_2 \le 0\}$ corresponds to (4.1) with $(\omega_1,\omega_2) = (1/2,1/2)$, the choice $\{(y_1,y_2) \in \mathbb{R}^2 \mid 2y_1 + y_2 \le 0\}$ gives us (4.1) with $(\omega_1,\omega_2) = (2/3,1/3)$. Other examples for the choice of the A, which provide nonsmooth scalarizations are given in Section 9. Figure 3 illustrates the construction of the Gerstewitz function, where the set A has a nonsmooth boundary.

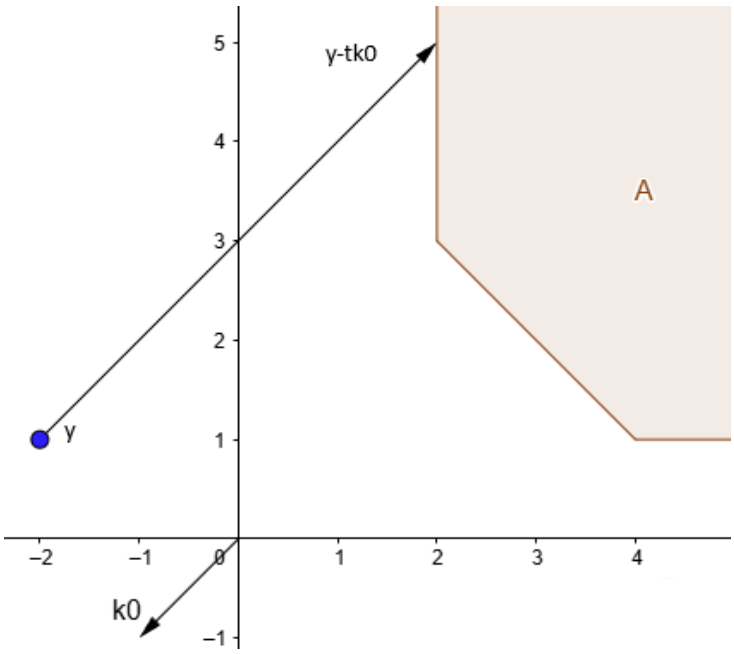


FIGURE 3. Gerstewitz function with $A = \{(y_1, y_2) \mid y_1 \geq 2, y_2 \geq 1, y_1 + y_2 \geq 5\}$

The first statement of this section provides conditions ensuring the properness and finiteness of the Gerstewitz scalarization function (5.1).

Proposition 5.1. *In addition to the conditions on* (5.1) *in Definition* 5.1*, assume that C is a cone. The following assertions hold:*

(i) *ϕ_A is proper if and only if the set A does not contain any line parallel to k^0, i.e.,*

$$\text{for all } \ y \in \mathbb{R}^m \ \text{ there exists } \ t \in \mathbb{R} \ \text{ with } \ y + tk^0 \notin A. \tag{5.2}$$

(ii) *ϕ_A is finite if and only if A does not contain any line parallel to k^0 and*

$$\mathbb{R}k^0 + A = \mathbb{R}^m.$$

Proof. To verify (i), suppose that there exists y such that

$$\phi_A(y) = -\infty \Longleftrightarrow y \in tk^0 + A \ \text{ for all } \ t \in \mathbb{R} \Longleftrightarrow \{y + tk^0 \mid t \in \mathbb{R}\} \subset A.$$

This directly implies that ϕ_A is proper if and only if (5.2) is satisfied.

To prove (ii), check first that if A does not contain any line parallel to k^0, then ϕ_A is proper. Indeed, we get that $\operatorname{dom}\phi_A = \mathbb{R}k^0 + A = \mathbb{R}^m$, which ensures that $\phi_A(y) < \infty$, and thus ϕ_A is finite. Conversely, if ϕ_A is finite, then $\mathbb{R}^m \subset \operatorname{dom}\phi_A$ and the properness of ϕ_A implies by (i) that A does not contain any line parallel to k^0. We clearly have $\mathbb{R}k^0 + A = \mathbb{R}^m$, which completes the proof of the proposition. □

The next theorem provides sufficient conditions ensuring *lower semicontinuity* (l.s.c.) and *continuity* of Gerstewitz function (5.1) and contains also a property crucial for applying (5.1) to ℓ_0 multiobjective optimization; see Remark 5.1.

Theorem 5.1. *In addition to the assumptions on* (5.1) *in Definition* 5.1*, suppose that C is a cone and that A is closed with $A \neq \mathbb{R}^m$. The following assertions hold:*

(i) *If $A-C=A$, then ϕ_A is l.s.c. and we have*

$$\big\{y\in\mathbb{R}^m \mid \phi_A(y)\le\lambda\big\}=\lambda k^0+A \ \text{ for all } \ \lambda\in\mathbb{R}. \tag{5.3}$$

(ii) *If $A-(C\backslash\{0\})=\operatorname{int}A$, then ϕ_A continuous and we have*

$$\big\{y\in\mathbb{R}^m \mid \phi_A(y)<\lambda\big\}=\lambda k^0+\operatorname{int}A \ \text{ for all } \ \lambda\in\mathbb{R}, \tag{5.4}$$

$$\big\{y\in\mathbb{R}^n \mid \phi_A(y)=\lambda\big\}=\lambda k^0+\operatorname{bd}A \ \text{ for all } \ \lambda\in\mathbb{R}. \tag{5.5}$$

Proof. To verify (i), denote $A':=\{(y,t)\in\mathbb{R}^m\times\mathbb{R}\mid y\in tk^0+A\}$ and pick $(y,t)\in A'$ with $t'\ge t$. Our goal is to show that $(y,t')\in A'$. Indeed, we get that

$$tk^0+A=t'k^0+A-(t'-t)k^0\subset t'k^0+A,$$

and hence $(y,t')\in A'$. Define $T:\mathbb{R}^m\times\mathbb{R}\to Y$ by $T(y,t):=y-tk^0$, which is a linear continuous mapping. It is clear that $A'=T^{-1}(A)$ and that the assumed closedness of A yields the closedness of A'. Since $A'\subset\operatorname{epi}\phi_A\subset\operatorname{cl}A'$ and A' is closed, we get $A'=\operatorname{epi}\phi_A$, which justifies that ϕ_A is l.s.c. Considering further $y\in\lambda k^0+A$ and $\lambda\in\{t\in\mathbb{R}\mid y\in tk^0+A\}$ tells us that $\lambda\ge\phi_A(y)$, i.e., $y\in\{z\in Y\mid\phi_A(z)\le\lambda\}$, which justifies the inclusion "$\supset$" in (5.3). To verify the opposite inequality therein, we take $t:=\phi_A(y)\le\lambda$ and get $y\in tk^0+A=\lambda k^0+A-(\lambda-t)k^0\subset A+\lambda k^0$. This gives us that

$$\begin{aligned}\phi_A(y+\lambda k^0)&=\inf\big\{t\in\mathbb{R}\mid y+\lambda k^0\in tk^0+A\big\}\\&=\inf\big\{t\in\mathbb{R}\mid y\in(t-\lambda)k^0+A\big\}.\end{aligned}$$

Denoting further $t':=t-\lambda$, we arrive at

$$\begin{aligned}\phi_A(y+\lambda k^0)&=\inf\big\{t'+\lambda\in\mathbb{R}\mid y\in t'k^0+A\big\}\\&=\inf\big\{t'\in\mathbb{R}\mid y\in t'k^0+A\big\}+\lambda\\&=\phi_A(y)+\lambda,\end{aligned}$$

which thus completes the verification of assertion (i).

To prove (ii), we pick any $\lambda\in\mathbb{R}$ and choose $y\in\lambda k^0+\operatorname{int}A$. Since $y-tk^0\in\operatorname{int}A$, there exists $\varepsilon>0$ with $y-tk^0+\varepsilon k^0\in A$. which yields $\phi_A(y)\le\lambda-\varepsilon<\lambda$. This shows that the inclusion "$\supset$" holds in 5.4. To verify the reverse inclusion, let $\lambda\in\mathbb{R}$ and $y\in Y$ be such that $\phi_A(y)<\lambda$. There exists $t\in\mathbb{R}$ with $t<\lambda$ such that $y\in tk^0+A$. Then $y\in\lambda k^0+A-(\lambda-t)k^0\subset\lambda k^0+\operatorname{int}A$, and hence (5.4) holds. Moreover, this shows that ϕ_A is upper semicontinuous. Combining the latter with (i) tells us that ϕ_A is continuous. Since $\operatorname{bd}A=A\backslash(\operatorname{int}A)$, it follows from (5.3) and (5.4) that (5.5) is satisfied, which therefore completes the proof of the theorem. □

Remark 5.1. It follows from the proof of Theorem 5.1 that

$$\phi_A(y+\lambda k^0)=\phi_A(y)+\lambda. \tag{5.6}$$

This is a *key property* of the Gerstewitz scalarization to handle ℓ_0 multiobjective optimization and develop a subgradient algorithm to find local Pareto solutions; see below.

Now we present useful characterizations of the convexity and positive homogeneity properties of the Gerstewitz scalarization function.

Proposition 5.2. *Suppose that all the assumptions of Theorem 5.1*(i) *are satisfied. Then:*
(i) ϕ_A *is a convex function if and only if* A *is a convex set.*
(ii) ϕ_A *is a positively homogeneous function if and only if* A *is a cone.*

Proof. We clearly have the representation $\operatorname{epi}\phi_A = T^{-1}(A)$, where the linear operator T is defined in the proof of Theorem 5.1(i). Then ϕ_A is convex (resp. positively homogeneous) if and only if $A = T(\operatorname{epi}\phi_A)$ is a convex (resp. conic) set. □

Given a closed, convex, and pointed cone $B \subset \mathbb{R}^m$, recall that the *partial order* relation $\le_B$ on $\mathbb{R}^m$ induced by B is defined by

$$x \le_B y \quad \text{if and only if} \quad y - x \in B. \tag{5.7}$$

It is easy to check that that the partial order $\le_B$ in (5.7) satisfies the properties:
- *Reflexivity*: $x \le_B x$.
- *Antisymmetry*: If $x \le_B y$ and $y \le_B x$, then $x = y$.
- *Transitivity*: If $x \le_B y$ and $y \le_B z$, then $x \le_B z$.
- *Convergence order*: If there are sequences $\{x_k\}$ and $\{y_k\}$ with $x_k \le_B y_k$ for all $k \in \mathbb{N}$ such that $x_k \to x$ and $y_k \to y$ as $k \to \infty$, then $x \le_B y$.

The next proposition presents characterizations of the standard *monotonicity* property of the Gerstewitz function with respect to the partial order (5.7); see the proof.

Proposition 5.3. *Let A be a closed subset of $\mathbb{R}^m$ with $A \ne \mathbb{R}^m$, let $B \subset \mathbb{R}^m$ be a cone that induces the partial order (5.7), and let $A - C = A$. Then the Gerstewitz function ϕ_A is monotone with respect to B if and only if $A - B \subset A$.*

Proof. First we check that the inclusion $A - B \subset A$ yields the monotonicity of ϕ_A. Pick any $y_1, y_2 \in \mathbb{R}^m$ with $y_2 - y_1 \in B$ and choose $t \in \mathbb{R}$ such that $y_2 \in tk^0 + A$. Then we get $y_1 \in y_2 - B \subset tk^0 + (A - B) \subset tk^0 + A$ and $\phi_A(y_1) \le t$. Therefore, $\phi_A(y_1) \le \phi_A(y_2)$, which justifies the monotonicity of the Gerstewitz function ϕ_A with respect to B.

To verify the reverse implication, suppose that ϕ_A is monotone with respect to B and then pick $y \in A$ and $b \in B$. Since $\phi_A(y) \le 0$, $y - (y - b) \in B$, and ϕ_A is monotone with respect to B, we have $\phi_A(y - b) \le \phi_A(y) \le 0$. The latter shows that $A - B \subset A$ and thus completes the proof. □

Now we derive a simple condition, which ensures the Lipschitz continuity of the Gerstewitz function on the entire space $\mathbb{R}^m$. The property plays an important role in verifying the convergence of our subgradient algorithm proposed below.

In what follows, we *identify the cone C* in definition (5.1) of the Gerstewitz function with the *ordering cone B* in (5.7). In these terms, recall the useful relationship from [17]:

$$\phi_A(y) \le \phi_A(y') + \phi_{-C}(y - y') \ \text{ for all } \ y, y' \in \mathbb{R}^m. \tag{5.8}$$

Theorem 5.2. *If $k^0 \in \operatorname{int} C$, then $\phi_A = \phi_{A,k^0}$ is Lipschitz continuous on $\mathbb{R}^m$.*

Proof. Given $k^0 \in \operatorname{int} C$, let $V \subset \mathbb{R}^m$ be a closed ball centered at 0 such that $k^0 + V \subset C$, and let $p_V : Y \to \mathbb{R}$ be the Minkowski functional associated with V. It is well known that p_V is a continuous seminorm and that $V = \{y \in \mathbb{R}^m \mid p_V(y) \le 1\}$. Pick $y \in \mathbb{R}^m$ and $t > 0$ such that $y \in tV$. Then $t^{-1}y \in V \subset k^0 - C$, and so $y \in tk^0 - C$ ensuring that $\phi_{-C}(y) \le t$. Therefore,

$\phi_{-C}(y) \le p_V(y)$, which confirms that $\mathbb{R}k^0 = \operatorname{dom}\phi_{-C} = \mathbb{R}^m$. It follows from the convexity and positive homogeneity of ϕ_C that $\phi_{-C}(y) \le \phi_{-C}(y') + p_V(y-y')$, which implies the inequality

$$|\phi_{-C}(y) - \phi_{-C}(y')| \le p_V(y-y') \ \text{ for all } \ y, y' \in \mathbb{R}^m. \tag{5.9}$$

This justifies the Lipschitz continuity of ϕ_{-C}. Let us show that ϕ_A does not take the value $-\infty$. If on the contrary $\phi_A(y_0) = -\infty$ for some $y_0 \in \mathbb{R}^m$, then $y_0 + \mathbb{R}k^0 \subset A$ yielding

$$A = A - C \supset y_0 + \mathbb{R}k^0 - C = y_0 + \mathbb{R}^m = \mathbb{R}^m,$$

a contradiction. Since $\operatorname{dom}\phi_A = \mathbb{R}k^0 + A = \mathbb{R}k^0 - C + A = \mathbb{R}^m + A = \mathbb{R}^m$ and the assumptions of Proposition 5.1 are clearly satisfied, we deduce from assertion (ii) therein that ϕ_A is finite. Using finally (5.8) together with (5.9) tells us that

$$|\phi_A(y) - \phi_A(y')| \le p_V(y-y') \ \text{ for all } \ y, y' \in \mathbb{R}^m,$$

which justifies the Lipschitz continuity of ϕ_A on $\mathbb{R}^m$ and thus completes the proof. □

We need the (convex) subdifferential calculation taken from [17, Corollary 4.2], where A_∞ stands for the horizon/recession cone associated with the convex set A; see [12].

Proposition 5.4. *Let A be a convex set, and let $k^0 \notin A_\infty$. Then, for all $\bar{y} \in \mathbb{R}^m$,*

$$\partial\phi_A(\bar{y}) = \big\{y^* \in \operatorname{cl} A \mid \langle k^0, y^*\rangle = 1,\ \langle \bar{y},\ y^*\rangle - \phi_A(\bar{y}) \ge \langle y, y^*\rangle \ \text{ whenever } \ y \in A\big\},$$

where cl$A$ *signifies the closure of the set A.*

6. ℓ_0 Optimization via Gerstewitz Scalarization

Now we are in a position to define the *scalarized ℓ_0 Gerstewitz function* associated with problem (MO-ℓ_0NP) as follows:

$$\phi_A(f_1(x) + \|x\|_0, \ldots, f_m(x) + \|x\|_0), \quad x \in \mathbb{R}^n. \tag{6.1}$$

Choosing $k^0 = (1, \ldots, 1) \in \mathbb{R}^m$, assuming that $\|x\|_0 = \lambda$, and using the key property (5.6) allow us to rewrite (6.1) in the equivalent form

$$\begin{aligned}\phi_A(f_1(x) + \|x\|_0, \ldots, f_m(x) + \|x\|_0) &= \phi_A((f_1(x), \ldots, f_m(x)) + \lambda k^0)\\ &= \phi_A(f_1(x), \ldots, f_m(x)) + \lambda\\ &= \phi_A(f_1(x), \ldots, f_m(x)) + \|x\|_0.\end{aligned}$$

This leads us to the *Gerstewitz-scalarized* (MO-ℓ_0NP) *problem* defined by

$$\min \phi_A(f_1(x), \ldots, f_m(x)) + \|x\|_0, \quad x \in \mathbb{R}^n. \tag{6.2}$$

Observe that the above choice of the direction k^0 is instrumental to take the ℓ_0-norm out of the composition in the Gerstewitz function ϕ_A as in (6.1) and place the ℓ_0-norm as an *additive term* in (6.2). This plays a crucial role in the development and justification of our algorithm to solve the multiobjective problem (MO-ℓ_0NP) in what follows.

To deal with the Gerstewitz composition $\phi_A(f_1(x), \ldots, f_m(x))$ in (6.2), we need the following two major properties. The first one concerns the convexity of $\phi_A(f_1(x), \ldots, f_m(x))$ with respect to the *nonpositive cone partial order* on $\mathbb{R}^m$; cf. Proposition 5.3.

Proposition 6.1. *In the setting of (6.2), assume that the functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1,\dots,m$ are convex and the set $A \neq \mathbb{R}^m$ is convex being such that $A - C = A$ and $A - \mathbb{R}^m_+ \subset A$. Then $\phi_A(f_1,\dots,f_m)$ is convex with respect to the partial order $\leq_{\mathbb{R}^m_+}$ in (5.7).*

Proof. By the convexity of f_i, we have $f_i(\lambda x+(1-\lambda)y) \leq \lambda f_i(x)+(1-\lambda)f_i$ for all $x,y \in \mathbb{R}^m$ and $i = 1,\dots,m$, $\lambda \in [0,1]$. Therefore,

$$\begin{aligned}&(f_1(\lambda x+(1-\lambda)y),\dots,f_m(\lambda x+(1-\lambda)y))\\ &\leq_{\mathbb{R}^m_+} \lambda(f_1(x),\dots,f_m(x))+(1-\lambda)(f_1(y),\dots,f_m(y)).\end{aligned}$$

It follows from Proposition 5.3 that

$$\begin{aligned}&(\phi_A((f_1(\lambda x+(1-\lambda)y),\dots,f_m(\lambda x+(1-\lambda)y))\\ &\leq \phi_A(\lambda(f_1(x),\dots,f_m(x))+(1-\lambda)(f_1(y),\dots,f_m(y))),\end{aligned}$$

which completes the proof of the proposition. □

The second proposition concerns the Lipschitz continuity of the Gerstewitz composition $\phi_A(f_1(x),\dots,f_m(x))$ in the ℓ_0 optimization problem (6.2).

Proposition 6.2. *Assume that the function f_i, $i = 1,\dots,m$, are Lipschitz continuous on $\mathbb{R}^n$. Then the Gerstewitz composition $\phi_A(f_1,\dots,f_m)$ is also Lipschitz continuous on $\mathbb{R}^n$.*

Proof. Let L_i be the Lipschitz constants of f_i for $i = 1,\dots,m$. Then we have

$$\|(f_1(x),\dots,f_m(x))-(f_1(y),\dots,f_m(y))\| \leq \sum_{i=1}^m L_i\|x-y\|.$$

It follows from Theorem 5.2 that, under the choice of the direction k^0 above, the Gerstewitz function ϕ_{A,k^0} is Lipschitz continuous on $\mathbb{R}^m$ with some constant M. Having all of this together leads us to the inequalities

$$\begin{aligned}&\|\phi_A((f_1(x),\dots,f_m(x)))-\phi_A((f_1(y),\dots,f_m(y))\|\\ &\leq M\|(f_1(x),\dots,f_m(x))-(f_1(y),\dots,f_m(y))\|\\ &\leq M\sum_{i=1}^m L_i\|x-y\|,\end{aligned}$$

which justifies the claimed Lipschitz continuity of the composition in (6.2). □

7. Algorithms for ℓ_0 Optimization

In this section, we design our novel algorithms to solve single-objection and multiobjective optimization problems containing the ℓ_0-norm function. Our algorithms are of the subgradient type that are based, due to the heavy nonconvexity of the ℓ_0-norm function, on the complete calculation of its limiting subdifferential given in Theorem 3.1.

Our further strategy is as follows. First we design two algorithms to solve the scalar problem of ℓ_0 optimization (ℓ_0NOP). Considering ℓ_0 multiobjective optimization, the results above allow us to find their Pareto solutions by employing the two scalarization approaches via the weight-sum and Gerstewitz functions. Accordingly, we get the two subgradient-type algorithms for the ℓ_0 multiobjective optimization problem while dealing with its scalarized versions in (4.1) and (6.2), respectively,

The first algorithm to find local minimizers of the scalar ℓ_0 optimization problem (ℓ_0NOP) by using the subdifferential calculation in Theorem 3.1 with the projection notation described in Section 4. Observe that the term $\tilde{I}_k\nabla f(x_k)$ in the algorithm is the gradient of the function f over the projection $(f\circ\tilde{I}_k)$ at x_k, and we have

$$\tilde{I}_k\nabla f(x_k)=\nabla f(x_k)+(v_{1,k},\ldots,v_{n,k})\in\nabla f(x_k)+\partial\|\cdot\|_0(x_k),$$

where $v_{i,k}$ are defined by

$$v_{i,k}:=\begin{cases}-\dfrac{\partial f}{\partial x_i}(x_{i,k}) & \text{if} \quad x_{i,k}=0,\\ 0 & \text{if} \quad x_{i,k}\neq 0.\end{cases}$$

Clearly, this choice of v is a subgradient of the ℓ_0-norm function.

Algorithm 1 to find a local minimizer in ℓ_0 scalar optimization

Require: Given x_0 as starting point, $\varepsilon>0$, stepsize $t<\dfrac{1}{L}$
if $\|\tilde{I}_k\nabla f(x_k)\|\geq\varepsilon$ **then**
$x_{k+1}=x_k-t\cdot\tilde{I}_k\nabla f(x_k)$
else Stop
end if

Algorithm 1 ensures that if x_k is in one hyperplane at some k-th iteration, then all x_j are in the same hyperplane for every $j>k$. Observe also that we get the inclusion $\tilde{I}_j\subset\tilde{I}_k$ by taking into account that for $x_{i,k}=0$, the vector $\tilde{I}_k\nabla f(x_k)$ has the same zero value i-th component. By Theorem 4.1, it is important to consider each hyperplane I_k, since some of our local minimizers lie on these hyperplanes.

We see that Algorithm 1 stops at a local minimizer. However, problem (ℓ_0NOP) may have many local minimizers, and it may be of interest to find other local minimizers of the problem as well. The second algorithm allows us to proceed further in this direction where I stands for the identity matrix.

Algorithm 2 to find local multiply minimizers of (ℓ_0NOP)

Require: Given x_0 as starting point, $\varepsilon>0$, stepsize $t<\dfrac{1}{L}$
if $\|\nabla f(x_k)\|\geq\varepsilon$ **then**
if $\|\tilde{I}_k\nabla f(x_k)\|\geq\varepsilon$ **then**
$x_{k+1}=x_k-t\cdot\tilde{I}_k\nabla f(x_k)$
else
$x_{k+1}=x_k-t\cdot(I-\tilde{I}_k)\nabla f(x_k)$
end if
else Stop
end if

Algorithm 2 can be interpreted as follows. Given any starting point, we employ Algorithm 1, which stops at some local minimizer. Then Algorithm 2 forces the iteration out of the projection that the obtained local minimizer is in by using the matrix $I-\tilde{I}_k$. This creates a new

starting point and gives us a new direction to reach a new local minimizer. In this way, we may eventually arrive at a global solution to the problem.

The next algorithm addresses solving the multiobjective ℓ_0 optimization problem via the weight-sum scalarization.

Algorithm 3 ℓ_0 multiobjective optimization via weight-sum scalarization

Require: Given x_0 as starting point, $\varepsilon > 0$, stepsize $t < \dfrac{1}{L}$, where $L := \max_{i=1,\ldots,m} L_i$, choose ω_i in weight sum (4.1)
if $\|\tilde{I}_k \nabla(\text{weight sum})(x_k)\| \geq \varepsilon$ **then**
$x_{k+1} = x_k - t \cdot \tilde{I}_k \nabla(\text{weight sum})(x_k)$
else Stop
end if

The last algorithm addresses finding local minimizers of the Gerstewitz-scalarized ℓ_0 minimization problem (6.2) with arriving at a local Pareto solution to the ℓ_0 multiobjective optimization problem (MO-ℓ_0NP) under a special choice of the set A.

Algorithm 4 ℓ_0 multiobjective optimization via Gerstewitz scalarization

Require: Given x_0 as starting point, $\varepsilon > 0$, step size t, $A \subset \mathbb{R}^m$ as a convex set with $A - \mathbb{R}^m_+ \subset A$, $k^0 = (1, \ldots, 1)$
if $\|\phi_A(f_1, \ldots, f_m)(x^{(k)} + 1) - \phi_A(f_1, \ldots, f_m)(x^{(k)})\| \geq \varepsilon$ **then**
$x_{k+1} = x_k - t \cdot \tilde{I}_k J_f(x^{(k)})^\top g^{(k)}$, where $J_f(x^{(k)})$ is a Jacobian matrix at $x^{(k)}$ and $g^{(k)} \in \partial \phi_A(f_1(x^{(k)}), \ldots f_m(x^{(k)}))$
$\phi_A \circ f_{best} = \min\{\phi_A \circ f_{best}, \phi_A \circ f(x^{k+1})\}$
else Stop
end if

8. Convergence Analysis

This section conducts convergence analysis of the proposed algorithms to find local solutions to the ℓ_0 of single-objective and multiobjective optimization problems. It is natural to start with scalar ℓ_0 optimization while restricting our attention to the convergence proof for Algorithm 1. The convergence proof for Algorithm 2 can be done similarly, and we leave the details to the reader.

To begin with, let us first verify the *monotonicity property* of Algorithm 1, which is crucial for the proof of convergence while being of its own interest.

Theorem 8.1. *Under the standing assumptions on the function f formulated in Section 4, We have the following monotonicity property of iterates in Algorithm 1 with respect to the objective function of problem* (ℓ_0NOP)*:*

$$f(x_{k+1}) + \|x_{k+1}\|_0 \leq f(x_k) + \|x_k\|_0.$$

Proof. It follows from the convexity and smoothness of f with the Lipschitz gradient that

$$\begin{aligned} f(y) &\le f(x)+\nabla f(x)^\top(y-x)+\frac{1}{2}\nabla^2 f(x)\|y-x\|_2^2 \\ &\le f(x)+\nabla f(x)^\top(y-x)+\frac{1}{2}L\|y-x\|_2^2, \end{aligned}$$

where L is a Lipschitz constant of f. Putting there $y=x_{k+1}=x_k-t\cdot\tilde{I}_k\nabla f(x_k)$ and $x=x_k$ brings us to the estimate

$$\begin{aligned} f(x_{k+1}) &\le f(x_k)+\nabla f(x_k)^\top(x_{k+1}-x_k)+\frac{1}{2}L\|x_{k+1}-x_k\|_2^2 \\ &= f(x_k)+\nabla f(x_k)(x_k-t\cdot\tilde{I}_k\nabla f(x_k)-x_k)+\frac{1}{2}L\|x_k-t\cdot\tilde{I}_k\nabla f(x_k)-x_k\|_2^2 \\ &= f(x_k)-\nabla f(x_k)t\cdot\tilde{I}_k\nabla f(x_k)+\frac{1}{2}L\|t\cdot\tilde{I}_k\nabla f(x_k)\|_2^2 \\ &= f(x_k)-t\|\tilde{I}_k\cdot\nabla f(x_k)\|_2^2+\frac{1}{2}Lt^2\|\cdot\tilde{I}_k\nabla f(x_k)\|_2^2 \\ &= f(x_k)-t\left(1-\frac{1}{2}Lt\right)\|\tilde{I}_k\nabla f(x_k)\|_2^2. \end{aligned}$$

Using $t\le\dfrac{1}{L}$, we know that $-\left(1-\dfrac{1}{2}Lt\right)=\dfrac{1}{2}Lt-1\le\dfrac{1}{2}L\cdot\dfrac{1}{L}-1=\dfrac{1}{2}-1=-\dfrac{1}{2}$. Plugging the latter into the last inequality tells us that

$$f(x_{k+1})\le f(x_k)-\frac{1}{2}t\|\tilde{I}_k\nabla f(x_k)\|_2^2. \tag{8.1}$$

Since $\dfrac{1}{2}t\|\tilde{I}_k\nabla f(x_k)\|_2^2>0$ unless $\tilde{I}_k\nabla f(x_k)=0$, which is the stopping condition for optimal solutions to problem (PP) in Theorem 4.1, we get

$$f(x_{k+1})\le f(x_k). \tag{8.2}$$

As discussed after the formulation of Algorithm 1, it follows from $\tilde{I}_j\subset\tilde{I}_k$ for $j>k$ that $\|x_j\|_0\le\|x_k\|_0$, and hence $\|x_{k+1}\|_0\le\|x_k\|_0$. Combining with (8.2) yields

$$f(x_{k+1})+\|x_{k+1}\|_0\le f(x_k)+\|x_k\|_0,$$

which therefore completes the proof of the theorem. □

Now we are ready to establish the convergence of Algorithm 1 to a local minimizer of problem (ℓ_0NOP) with a rate estimate.

Theorem 8.2. *Let x_k be the k-th iteration of Algorithm 1 of Theorem 8.1, and let I_k be the associated smallest subspace. Then the sequence $\{x_k\}$ converges to a local minimizer $\bar{x}$ of problem (ℓ_0NOP) with the rate estimate*

$$f(x_{(k+s)})+\|x_{(k+s)}\|-f(\bar{x})-\|\bar{x}\|\le\frac{\|x_k-\bar{x}\|_2^2}{2st}.$$

Proof. The iterative process stops after k-th iteration if we have $x_k\in I_k$ for the corresponding hyperplane. Our ℓ_0 optimization problem (ℓ_0OP) turns into a projection problem (PP), which is the convex problem by the convexity of the function $f\circ\tilde{I}_k$. Starting from this iteration,

Algorithm 1 becomes the classical gradient descent method with x_k being the starting point of the algorithm, and thus its the convergence follows.

To proceed in this direction, let x_k be the starting point of subproblem (PP). The assumption that x_k is the k-th iteration of (ℓ_0OP) with I_k being the associated smallest subspace ensures that $\|x_{k+j}\|_0 = \|x_k\|_0$ for all j. Since (PP) is a convex problem, we have

$$f(\bar{x}) - f(x) \geq \nabla f(x)^\top (x - \bar{x}),$$
$$f(x) \leq f(\bar{x}) + \nabla f(x)^\top (x - \bar{x}).$$

Denote $x^+ := x - t \cdot \tilde{I}(x)\nabla f(x)$, where $\tilde{I}(x)$ signifies the projection matrix associated with x, and deduce from (8.1) the relationships

$$\begin{aligned}
f(x_+) &\leq f(x) - \frac{1}{2}t\|\tilde{I}(x)\nabla f(x)\|_2^2 \\
&\leq f(\bar{x}) + \nabla f(x)^\top (x - \bar{x}) - \frac{1}{2}t\|\tilde{I}(x)\nabla f(x)\|_2^2 \\
&= f(\bar{x}) + \frac{1}{2t}\left(2t\nabla f(x)^\top (x - \bar{x}) - t^2\|\tilde{I}(x)\nabla f(x)\|_2^2 - \|x - \bar{x}\|_2^2 + \|x - \bar{x}\|_2^2\right) \\
&= f(\bar{x}) + \frac{1}{2t}\left(\|x - \bar{x}\|_2^2 - \|x - t\tilde{I}(x)\nabla f(x) - \bar{x}\|_2^2\right) \\
&= f(\bar{x}) + \frac{1}{2t}\left(\|x - \bar{x}\|_2^2 - \|x^+ - \bar{x}\|_2^2\right).
\end{aligned}$$

Summing over iterations leads us to the estimates

$$\begin{aligned}
\sum_{i=1}^{s}(f(x_{k+i}) - f(\bar{x})) &\leq \sum_{i=1}^{s}\frac{1}{2t}\left(\|x_{(k+i-1)} - \bar{x}\|_2^2 - \|x_{k+i} - \bar{x}\|_2^2\right) \\
&\leq \frac{1}{2t}\|x_k - \bar{x}\|_2^2.
\end{aligned}$$

It follows from Theorem 8.1 that

$$f(x_{k+s}) - f(\bar{x}) \leq \frac{1}{s}\sum_{i=1}^{s}(f(x_{k+i}) - f(\bar{x})) \leq \frac{\|x_k - \bar{x}\|_2^2}{2st},$$

which therefore completes the proof of the theorem. □

The next theorem verifies the convergence and the rate estimates for Algorithm 2 to find a local Pareto solution of the ℓ_0 multiobjective optimization problem (MO-ℓ_0NP).

Theorem 8.3. *Consider problem* (MO-ℓ_0NP) *under the standing assumptions. Then Algorithm 3 converges to a local Pareto solution of* (MO-ℓ_0NP) *with the rate estimate*

$$(\text{weight sum})(x_{(k+s)}) + \|x_{(k+s)}\| - (\text{weight sum})(\bar{x}) - \|\bar{x}\| \leq \frac{\|x_k - \bar{x}\|_2^2}{2st},$$

Proof. Proposition 4.1 tells us that we can find a local Pareto solution to problem (MO-ℓ_0NP) as a local minimizer of the weight-sum ℓ_0 minimization problem (4.1). Then the claimed results follow from Theorem 8.2 applied to (4.1). □

The last theorem of this section verifies the convergence and the rate estimates for Algorithm 4 to find a local minimizer of the Gerstewitz-scalarized l^0 multiobjective optimization problem in (6.2) under the general choice of the set A therein. A particular choice of A allows

us to find a local Pareto solution to the ℓ_0 multiobjective problem (MO-ℓ_0NP). The proof of this theorem is based on the properties of the Gerstewitz scalarization function (5.1) established in Sections 5 and 6.

Theorem 8.4. *Consider the Gerstewitz-scalarized ℓ_0 optimization problem (6.2) in the setting of Algorithm 4. Then this algorithm converges to a local minimizer $\bar{x}$ of (6.2) with the following rate estimate:*

$$\phi_A \circ f_{best}^{(k+s)} - \phi_A \circ \bar{f} \le \frac{\operatorname{dist}(x^{(1+s)},\bar{x})^2 + M^2 t^2 k}{2tk}, \tag{8.3}$$

where M is a Lipschitz modulus of $\phi_A \circ f$, and where $\bar{f} := f(\bar{x})$. Moreover, the choice of the set A satisfying $A - \mathbb{R}^m_+ \subset A$ ensures that $\bar{x}$ is a local Pareto solution to the ℓ_0 multiobjective optimization problem (MO-ℓ_0NP).

Proof. Similarly to proof of Theorem 8.2, we suppose that the process of reducing the dimension stops at the $s-$th iteration. Let $\bar{x}$ be any optimal solution in the smallest hyperplane over the iterations. To simplify the proof, we denote $\tilde{g}^{(k)} := J_f(x^{(k)})^\top g^{(k)}$ and get

$$\begin{aligned} \|x^{(k+s1)} - \bar{x}\|_2^2 &= \|x^{(k+s)} - t \cdot \tilde{I}_{k+s}\tilde{g}^{(k+s)} - \bar{x}\|_2^2 \\ &= \|x^{(k+s)} - \bar{x}\|_2^2 - 2t \cdot \tilde{I}_{k+s}\tilde{g}^{(k+s)}(x^{(k+s)} - \bar{x}) + t^2\|\tilde{g}^{(k+s)}\|_2^2 \\ &\le \|x^{k+s} - \bar{x}\|_2^2 - 2t(\phi_A \circ f(x^{(k+s)}) - \phi_A \circ \bar{f}) + t^2\|g^{(k+s)}\|_2^2 \end{aligned} \tag{8.4}$$

with the subgradient inequality coming from the definition

$$\cdot\tilde{I}_{k+s}\tilde{g}^{(k+s)}(\bar{x} - x^{(k+s)}) = t \cdot \tilde{g}^{(k+s)}(\bar{x} - x^{(k+s)}) \le \phi_A \circ \bar{f} - \phi_A \circ f(x^{(k+s)}).$$

Applying (8.4) recursively leads us to the estimate

$$\|x^{(k+s+1)} - \bar{x}\|_2^2 \le \|x^{(s+1)} - \bar{x}\|_2^2 - 2t\sum_{i=1}^{k}(\phi_A \circ f(x^{(i+s)} - \phi_A \circ \bar{f}) + t^2\sum_{i=1}^{k}\|\tilde{g}^{(i+s)}\|_2^2.$$

The usage of $\|x^{(k+s+1)} - \bar{x}\|_2^2 \ge 0$ yields

$$2t\sum_{i=1}^{k}(\phi_A \circ f(x^{(i+s)} - \phi_A \circ \bar{f}) \le \|x^{(s+1)} - \bar{x}\|_2^2 + t^2\sum_{i=1}^{k}\|\tilde{g}^{(i+s)}\|_2^2,$$

which being combined with $t\sum_{i=1}^{k}(\phi_A \circ f(x^{(i+s)} - \phi_A \circ \bar{f} \ge (\sum_{i=1}^{k} t)(\phi_A \circ f_{best}^{(k+s)} - f(\bar{x}))$ results in the composition estimate

$$\phi_A \circ f_{best}^{(k+s)} - \phi_A \circ \bar{f} \le \frac{\|x^{(s+1)} - \bar{x}\|_2^2 + t^2\sum_{i=1}^{k}\|\tilde{g}^{(k+s)}\|_2^2}{2\sum_{i=1}^{k} t}.$$

Furthermore, the imposed Lipschitz continuity $\|\tilde{g}^{(k+s)}\|_2 \le M$ ensures that

$$\phi_A \circ f_{best}^{(k+s)} - \phi_A\bar{f} \le \frac{\|x^{(s+1)} - \bar{x}\|_2^2 + t^2 k M^2}{2tk},$$

and therefore we arrive at the condition $\phi_A \circ f_{\text{best}}^{(k+s)} - \phi_A \circ \bar{f} \le \dfrac{\operatorname{dist}(x^{(1+s)},\bar{x})^2 + M^2 t^2 k}{2tk}$, which verifies the convergence of Algorithm 4 to the local minimizer $\bar{x}$ of (6.2) with the claimed rate estimate in (8.3).

Finally, the choice of A such that $A-\mathbb{R}^m_+ \subset A$ and the monotonicity of the Gerstewitz function with respect to partial order ensure that $\bar{x}$ is a local Pareto solution to the ℓ_0 multiobjective optimization problem (MO-ℓ_0NP). This completes the proof. □

9. Numerical Illustrations

In this section, we demonstrate the performance of the designed algorithms by considering typical examples. All of the calculations were conducted by using Jupyter Notebook.

Example 9.1. Let $f(x,y) := x^2+2y^2-2x-2xy+3+\|(x,y)\|_0$ for $(x,y)\in\mathbb{R}^2$.

It is easy to check that the function in this example has 3 local minimizers: $(2,1),(1,0),$ $(0,0)$. We will use 3 different starting points to compare the convergences of Algorithm 1 to optimal points and to see us how the algorithm performs in these settings, which is graphically illustrated in the figures below.

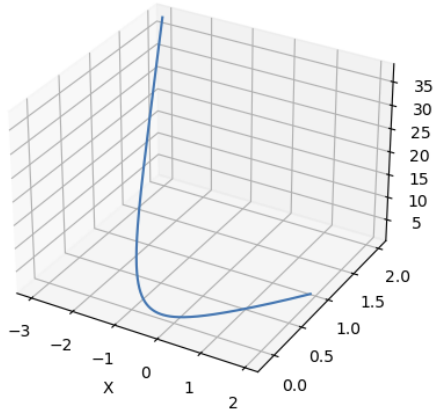


FIGURE 4. Initial point $(-3,2)$

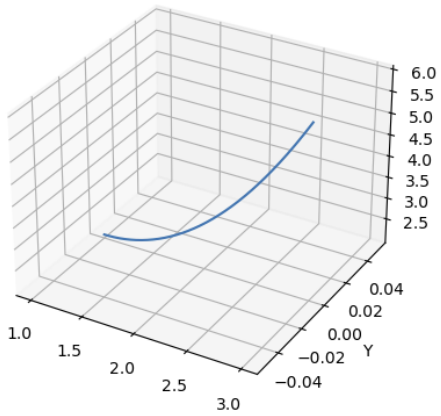


FIGURE 5. Initial point $(3,0)$

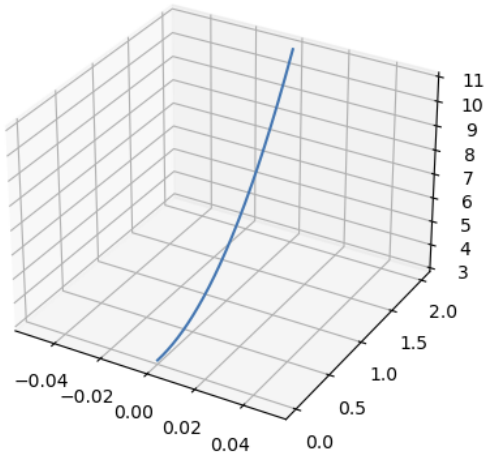


FIGURE 6. Initial point $(0,2)$

As we can see in the the Figure 5 and Figure 6, the iterations remain in the hyperplanes and converge to the corresponding local minimizers.

Let us further use an alternative definition of the ℓ_0-norm function to demonstrate the possibility of the components going to zero. Instead of defining the ℓ_0-norm by the condition that x_i exactly equal to zero, we modify it by $|x_i| \leq \varepsilon$, with $\varepsilon = 10^{-6}$ in the next figure, to see the differences.

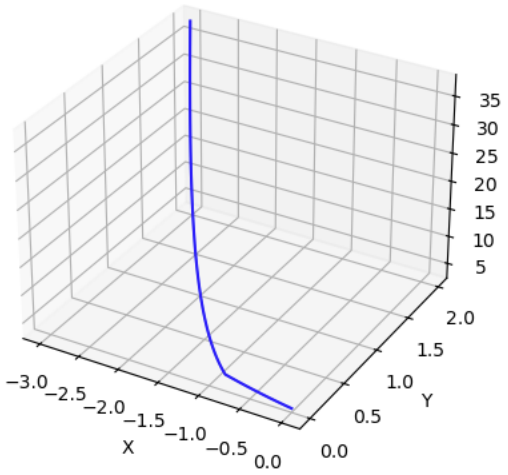


FIGURE 7. Initial point $(-3,2)$ with alternative definition

Compared to Figure 4, which is similar to the classical gradient descent method, the modified iterations go into the hyperplane $y = 0$ with the subsequent iterations remaining in it. To see the drop of the value and guarantee that Algorithm 1 is a descent algorithm, we can check the graph of the function values.

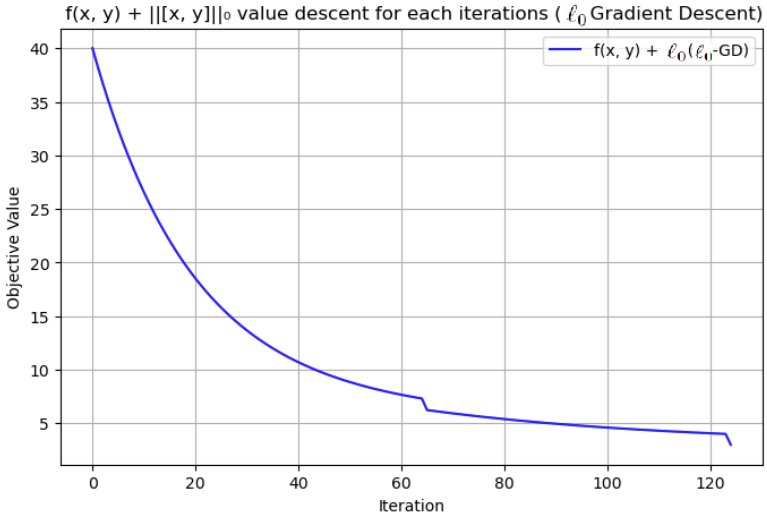


FIGURE 8. Adaptive definition for the ℓ_0 gradient descent

The next example addresses ℓ_0 multiobjective optimization problems.

Example 9.2. Let $f_1(x,y) := (x-1)^2 + y^2$ and $f_2(x) := x^2 + (y-2)^2$.

Apply the Gerstewitz scalarization, we choose the set $A = \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 + y_2 \leq 0\}$ with $k^0 = (1,1)$ in Algorithm 4. Definition (5.1) of Gerstewitz function gives us $\phi_A(f_1, f_2) = \inf\{t \mid (f_1, f_2) - tk^0 \in A\}$, and so $f_1 - t + f_2 - t \leq 0$, which yields $t \geq \dfrac{f_1 + f_2}{2}$. Taking the infimum of t, we get $t = \dfrac{f_1 + f_2}{2}$. Therefore, this scalarization agrees with to the case of the weighted-sum method.

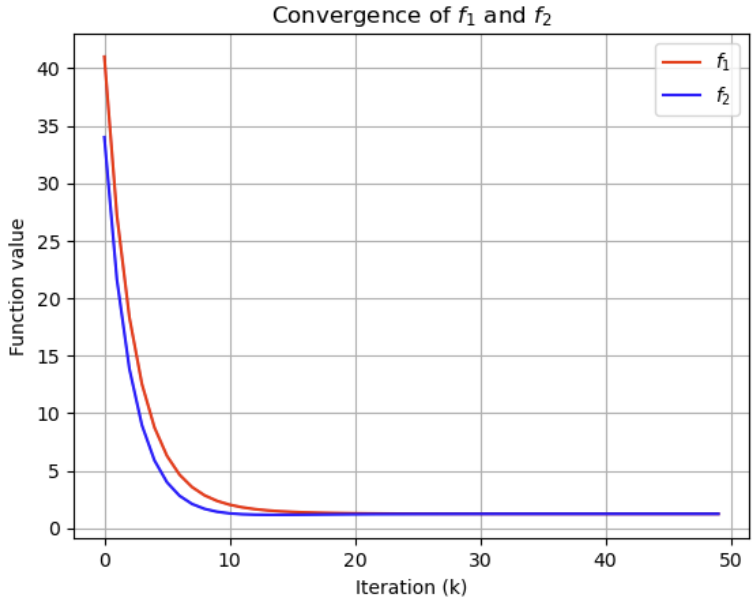


FIGURE 9. Weighted-Sum Method

The last example illustrates the Gerstewitz scalarization method bringing us a nonsmooth resulting function in contrast to the weight-sum approach.

Example 9.3. Let $f_1(x) := (x-2)^2$ and $f_2(x) := (x+1)^2+1$.

In this example, we choose the set $A := \{(y_1,y_2) \in \mathbb{R}^2 \,|\, y_1 \text{ and } y_2 \le 0\}$. We have $\phi_A(f_1,f_2) = \inf\{t \mid (f_1,f_2) - tk^0 \in A\}$. Then $t \ge f_1$ and $t \ge f_2$. Thus $t \ge \max\{f_1,f_2\}$ with the infimum of t calculated by $t = \max\{f_1,f_2\}$.

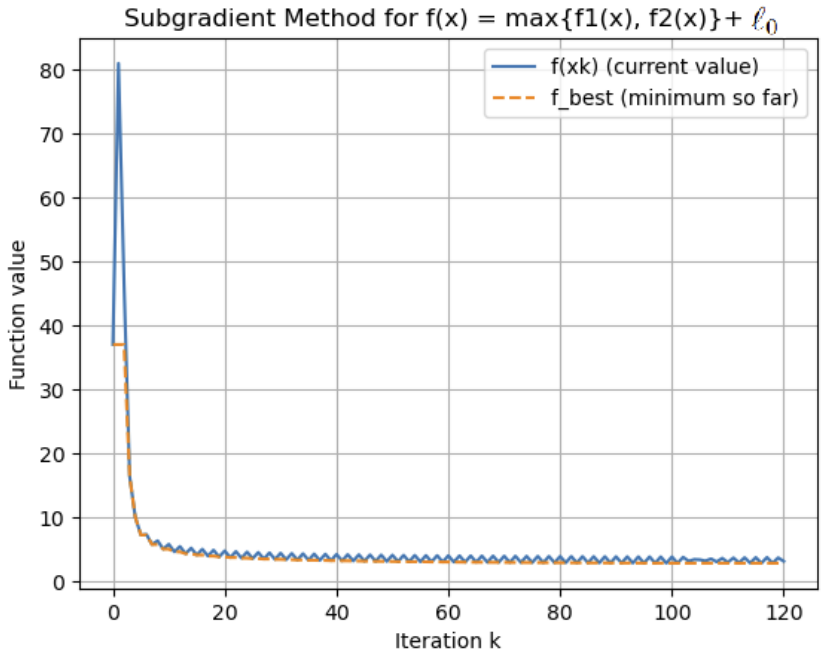


FIGURE 10. Maximum Function

10. CONCLUDING REMARKS AND FUTURE RESEARCH

This paper concerns single-objective and multiobjective optimization problems involving the ℓ_0-norm function in their objectives. Problems of this type arise in models of proton therapy, which have been the main practical motivations of our research. Being intrinsically nonsmooth and nonconvex, such problems require the usage advanced tools of variational analysis for their study and applications. The main results obtained in the paper revolve around novel subgradient algorithms to solve both scalar and multiobjective versions based on the limiting subdifferential by Mordukhovich. In this way, we implement two scalarization techniques to deal with multi-objective problems: the weigh-sum approach and mainly Gerstewitz scalarization. The desired convergence properties of the designed algorithms are rigorously verified, and the performance of these algorithms are illustrated by numerical calculations for typical examples.

While this study introduces a novel framework for ℓ_0 optimization in proton radiation therapy, we acknowledge that the numerical examples provided are simplified relative the complexity of real clinical scenarios. These examples primarily serve to illustrate the feasibility and potential of proposed methodology. The main direction of our future research to implement the designed algorithms to solving realistic practical models of proton beam therapy as well as related models

of cancer research. These tasks will definitely require some adjustments and modifications, which will bring us to new mathematical results.

Acknowledgements
Research of Xuanfeng Ding was partly supported by NIH under grant R01CA301448. Research of Boris Mordukhovich was partly supported by the US National Science Foundation under grant DMS-2204519, by the Australian Research Council under Discovery Project DP-190100555, and by Project 111 of China under grant D21024. Research of Anh Vu Nguyen was partly supported by the US National Science Foundation under grant DMS-2204519.

REFERENCES

[1] P. D. Khanh, B. S. Mordukhovich, V. T. Phat, Coderivative-based newton methods in structured nonconvex and nonsmooth optimization, arXiv:2403.04262v2, 2025.
[2] Y. Jiao, B. Jin, X. Lu, A primal dual active set with continuation algorithm for the ℓ_0-regularized optimization problem, Appl. Comput. Harmon. Anal. 39 (2015), 400–426.
[3] C. Louizos, M. Welling, D. P. Kingma, Learning sparse neural networks through ℓ_0 regularization, In: Proc. Inter. Conf. Learn. Reprersen. arXiv:1712.01312, 2018.
[4] C. Lee, F. Imrie, M. van der Schaar, Self-supervision enhanced feature selection with correlated gates, In: Proc. Inter. Conf. Learn. Represen. 2022.
[5] Q. Lyu, et al., Image-domain multimaterial decomposition for dual-energy computed tomography with non-convex sparsity regularization, J. Med. Imag. 6 (2-19) 044004.
[6] W. Gu, et al., A novel energy layer optimization framework for spot-scanning proton arc therapy, Med. Phys. 47 (2020), 2072–2084.
[7] L. Zhao, et al., The first direct method of spot sparsity optimization for proton arc therapy, Acta Oncol. 62 (2023), 48–52.
[8] Q. Fan, et al., Optimizing spot-scanning proton arc therapy with a novel spot sparsity approach, Med. Phys. 52 (2025), 1789–1797.
[9] L. Zhao, et al., A novel simultaneous plan quality and beam delivery time SPArc optimization platform using primal dual active set with continuation (PDASC), Med Phys. 49(6):E116, 2022.
[10] L. Zhao, et al., The first direct method of spot sparsity optimization for proton arc therapy, Acta Oncol. 62 (2023), 48–52.
[11] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications, Springer, Berlin, 2006.
[12] R. T. Rockafellar, R. J-B. Wets, Variational Analysis, Springer, Berlin, 1998.
[13] C. Gerstewicz (Tammer), Nichtknvexe dualität in der vektoroptimierung, Wissenschaftitiche Zeitschrift der TH Leuna-Merseburg. 25 (1983), 357–364.
[14] A. A. Khan, C. Tammer, C. Zălinescu, Set-valued Optimization. An Introduction and Applications, Springer, Berlin, 2015.
[15] B. S. Mordukhovich, Variational Analysis and Applications, Springer, Cham, 2018.
[16] B. S. Mordukhovich, Second-order Variational Analysis in Optimization, Variational Stability, and Control: Theory, Algorithms, Applications, Springer, Cham, 2018.
[17] C. Tammer, C. Zălinescu, Lipschitz properties of the scalarization function and applications, Optimization, 59 (2010), 305–319.