A New Practical Generalized Derivative for Nonsmooth Functions

M. H. Noori Skandari
e-mail: hadinoori344@yahoo.com
Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

A. V. Kamyad
e-mail: avkamyad@yahoo.com
Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

H. R.Erfanian
e-mail: erfanianhamidreza@yahoo.com
Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

In this paper, we introduce a new generalized derivative (GD) for nonsmooth functions which is based on optimization. This type of GD is an extension of the concept of the usual derivative of smooth functions and is defined as an optimal solution of a special optimization problem. Here, the optimization problem is approximated with a linear programming problem by solving of which, we can obtain this GD as simple as other approaches. This GD is utilized for some nonsmooth functions and is shown in some examples the efficiency of the proposed approach.

Keywords: Nonsmooth functions, Generalized derivative, Optimization, Linear programming.

1. Introduction

Nonsmooth analysis is a branch of mathematics proposed by Clarke in 1973 (see [3,4]). In nonsmooth analysis, we deal with nonsmooth functions, the sets with nonsmooth boundaries, and the set valued mappings. The nonsmooth functions are not differentiable in at least one point of their domain. The researchers endeavored to generalize the derivative of these functions in nonsmoothness points (see Jeyakumar [12]). The discrete and continuous-time optimization problems are the most important ones which need generalized derivatives (GDs) of nonsmooth functions, since the GDs in these problems give us the optimal solutions of these problems. Thus, not only we present a GD, but we must also be able to use it to obtain the optimal solution of the optimization problems. Up to now, many generalized derivatives for nonsmooth functions have been presented and in this section, we are going to make the reader familiar with one of the most well known GDs, that is Clarke's GDs. Note that, by using Clarke's GDs and other GDs, necessary and sufficient conditions for optimality in optimization problems are presented, but these conditions are only used to test the optimal solution and by these, we usually cannot obtain the optimal solution (see [12,17]).

In this section, suppose that $\varphi:\mathbb{R}^n \to \mathbb{R}$ is a real-valued function. We know that the partial derivative of the smooth function $\varphi(.)$ with respect to the j^{th} component of $x \in \mathbb{R}^n$ is denoted by $\frac{\partial \varphi(x)}{\partial x_j}$ and defined as $\frac{\partial \varphi}{\partial x_j} = \lim_{t\to 0} \frac{\varphi(x+te_j)-\varphi(x)}{t}$ where e_j is the unit j^{th} coordinate direction in $\mathbb{R}^n$. Moreover, the gradient of the function $\varphi(.)$ at the point x is denoted by $\nabla\varphi(x)$ and defined as

$\nabla\varphi(x) = (\frac{\partial\varphi(x)}{\partial x_1},...,\frac{\partial\varphi(x)}{\partial x_n})$. But, for nonsmooth functions, the Clarke generalized derivatives are defined as follows. Suppose that $\varphi(.)$ is locally Lipschitz function at point $x \in \mathbb{R}^n$, that is, there exists a neighbourhood B of x and a positive M such that

$$\left|\varphi(x) - \varphi(y)\right| \le M\left|x - y\right|, \quad y \in B.$$

The *Clarke directional derivative* of the function $\varphi(.)$ at x in the direction unit vector $u \in \mathbb{R}^n$ is defined as follows:

$$\varphi^0(x;u) = \lim_{\substack{t\to 0 \\ y\to x}} \frac{\varphi(y+tu) - \varphi(y)}{t}.$$

Moreover, the *Clarke subdifferential* of $\varphi(.)$ at x is defined by

$$\partial^C\varphi(x) = \quad \xi = (\xi_1,...,\xi_n) \in \mathbb{R}^n : \xi.u \le \varphi^0(x;u), \text{ for } \quad u = (u_1,...,u_n) \in \mathbb{R}^n \quad .$$

Indeed, the *Clarke generalized gradient* of function $\varphi(.)$ at x is defined by

$$\partial^C\varphi(x) = co\left\{\lim_{x_i\to x} \nabla\varphi(x_i) : x_i \in \Omega\right\}$$

where Ω is the set of points in neighbourhood B at which $\varphi(.)$ is differentiable, and $co(A)$ is the convex hall of set $A \subset \mathbb{R}^n$.

Many other generalized derivatives have been defined as needed in various other situations. Several of more well known GDs are collected and listed by Jeyakumar [12], including the Mordokhovich subgradient and coderivative [16,17], the Warga derivative containers [20,21], the Ioff prederivative and approximate subdifferentials [10,11], the Gowda and Ravindran H-differentials [8], the Clarke-Rockafellar subdifferential [18], the Demyanov-Robinov quasi-differentials [7]. Note that the majority of the GDs are examples of Jeyacumar's pseudo-jacobians (see chapter 1 of Jeyakumar [12]). Moreover, the semi-smooth Newton derivative is defined for nonsmooth functions (see [6,9,13,14,22]) which is near to the Frechet and strict (Hadamard) derivatives.

Each of the above the above-mentioned GDs of nonsmooth functions, has some conditions and restrictions. Some of these conditions and restrictions are as follows:

i) For obtaining GDs of nonsmooth function $\varphi(.) : \mathbb{R}^n \to \mathbb{R}$ at point $x \in \mathbb{R}^n$, function $\varphi(.)$ must usually be continuous and locally Lipschitz at point x .

ii) The GDs of nonsmooth function $\varphi(.) : \mathbb{R}^n \to \mathbb{R}$ are presented in a given and known point $x \in \mathbb{R}^n$.

iii) Usually, the concept of lim, limsup or liminf is used to define GDs of nonsmooth function $\varphi(.) : \mathbb{R}^n \to \mathbb{R}$ at point $x \in \mathbb{R}^n$ (see chapter 1 of Jeyakumar [12]).

iv) The set of smoothness and nonsmoothness points for nonsmooth function $\varphi(.) : \mathbb{R}^n \to \mathbb{R}$ must be known.

Thus, it is important that we introduce a practical and applicable GD for nonsmooth integrable functions that avoids these restricting conditions. For this porpuse, Kamyad et al. proposed a useful and practical GD for nonsmooth functions (see [15]). We now present a different approach and definition for the GD of nonsmooth functions, which is based on a special optimization problem. Here, we approximate the optimization problem by a linear programming problem.

The structure of this paper is as follows. In Section 2, we state some theorems and introduce an optimization problem for GD of one-variable nonsmooth integrable functions. In particular, we obtain a linear programming (LP) problem corresponding to the above-mentioned optimization problem the solution of which is an approximation to the GD. In Section 3, using the GD of the one-variable nonsmooth integrable functions, we generalize this concept of GD to multi-variable nonsmooth integrable functions. In Section 4, we illustrate the GD of some nonsmooth functions by using this approach. Finally, the conclusions in Section 5 include several applications where the GD could be utilized.

2. Generalized derivative of nonsmooth one-variable functions

We introduce the GD of one-variable nonsmooth integrable functions $f:\Omega\subset\mathbb{R}\to\mathbb{R}$. Assume that $C^1(\Omega)$ is the space of continuous differentiable functions on set Ω, and $N_\delta(s)$ is the neighbourhood of s with the radius δ. For facility and without loss of generality, we assume that $\Omega=(0,1)$ and initiate with the following lemma:

Lemma 2.1: Let $\phi:(0,1)\to\mathbb{R}$ be a function such that $\lim_{x\to c}\phi(x)=L$ where $L\in\mathbb{R}$ and $c\in(0,1)$. Then for all $K\in\mathbb{R}, K\neq L$ there exists $\rho_c>0$ such that $\left|\phi(x)-L\right|<\left|\phi(x)-K\right|$ for all $x\in(c-\rho_c,c+\rho_c)\setminus\{c\}$.

Proof: Let $\varepsilon_0=\dfrac{\left|L-K\right|}{2}>0$. Since $\lim_{x\to c}\phi(x)=L$, there is $\rho_c>0$ such that for all $x\in N_{\rho_c}(c)\setminus\{c\}$ we have

$$\left|\phi(x)-L\right|<\varepsilon_0. \tag{1}$$

In addition,

$$\begin{aligned}\left|\phi(x)-K\right|&=\left|\phi(x)-K+L-L\right|\\&>\left|L-K\right|-\left|\phi(x)-L\right|\end{aligned}$$

Thus by (1)

$$\begin{aligned}\left|\phi(x)-K\right|&>\left|L-K\right|-\frac{\left|L-K\right|}{2}\\&=\frac{\left|L-K\right|}{2}\\&>\left|\phi(x)-L\right|.\ \square\end{aligned}$$

Proposition 2.2: Let $f(.)\in C^1[0,1]$ and $m\in\mathbb{N}$. Then there exists $\delta>0$ such that for all arbitrary $s_i\in(\frac{i-1}{m},\frac{i}{m})$ and $g_i\in\mathbb{R}$, $i=1,2,\dots,m$ we have

$$\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)f'(s_i)\right|dx<\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx \tag{2}$$

Proof: By lemma 2.1, since $f'(s_i)=\lim_{x\to s_i}\dfrac{f(x)-f(s_i)}{x-s_i}$, $i=1,2,\dots,m$, there is $\rho_{s_i}>0$ such that for all $x\in N_{\rho_{s_i}}(s_i)\setminus\{s_i\}$ and $g_i\neq f'(s_i)$ we have

$$\left|\frac{f(x)-f(s_i)}{x-s_i}-f'(s_i)\right|<\left|\frac{f(x)-f(s_i)}{x-s_i}-g_i\right|,$$

$$\left|f(x)-f(s_i)-(x-s_i)f'(s_i)\right|<\left|f(x)-f(s_i)-(x-s_i)g_i\right|, \tag{3}$$

Suppose that $\delta=\min\ \rho_{s_i}:1,2,\ldots,m$. Thus $(s_i,s_i+\delta)\subseteq N_{\rho_{s_i}}(s_i)\setminus\{s_i\}$ for $i=1,2,\ldots,m$ and by (3) we have:

$$\int_{s_i}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)f'(s_i)\right|dx<\int_{s_i}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx, \tag{4}$$

In addition, $(s_i-\delta,s_i)\subseteq N_{\rho_{s_i}}(s_i)\setminus\{s_i\}$ and by (3)

$$\int_{s_i-\delta}^{s_i}\left|f(x)-f(s_i)-(x-s_i)f'(s_i)\right|dx<\int_{s_i-\delta}^{s_i}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx, \tag{5}$$

Hence, using (4) and (5) for all $i=1,2,\ldots,m$

$$\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)f'(s_i)\right|dx<\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx,\quad g_i\neq f'(s_i). \quad \square$$

Let $f:[0,1]\to\mathbb{R}$ and let $m\in\mathbb{N}$ be a given large number. Also, assume that $s_i\in(\frac{i-1}{m},\frac{i}{m})$, for $i=1,2,\ldots,m$ are arbitrary numbers. Define the following optimization problem:

$$\underset{(g_1,\ldots,g_m)\in\Delta}{Minimize}\quad T(g_1,\ldots,g_m)=\sum_{i=1}^{m}\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx \tag{6}$$

where $\delta>0$ is a sufficiently small given number and Δ is a subspace of $\mathbb{R}^m$.

Theorem 2.3 : Let $f:[0,1]\to\mathbb{R}$ be an integrable function, $m\in\mathbb{N}$ is a given large number, $s_i\in(\frac{i-1}{m},\frac{i}{m})$ for $i=1,2,\ldots,m$ are arbitrary numbers, and $\delta>0$ is a sufficiently small number . The function $T:\Delta\subseteq\mathbb{R}^n\to\mathbb{R}$ defined in problem (6) is continuous.

Proof: Let $\bar{g}\in\mathbb{R}$ and $\varepsilon>0$ be given. Define

$$\psi_i(g)=\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g\right|dx,\qquad g\in\mathbb{R},\ i=1,2,\ldots,m,$$

and assume that $\delta_0=\frac{\varepsilon}{\delta^2}$. Now, for all $g\in\mathbb{R}$, if $\left|g-\bar{g}\right|<\delta_0$ then

$$\left|\psi_i(g)-\psi_i(\bar{g})\right|=\left|\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g_i\right|dx-\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)\bar{g}_i\right|dx\right|$$

$$=\left|\int_{s_i-\delta}^{s_i+\delta}\left|f(x)-f(s_i)-(x-s_i)g\right|-\left|f(x)-f(s_i)-(x-s_i)\bar{g}\right|dx\right|$$

$$
\begin{aligned}
&\le \int_{s_i-\delta}^{s_i+\delta} \Big| \left|f(x)-f(s_i)-(x-s_i)g\right| - \left|f(x)-f(s_i)-(x-s_i)\overline{g}\right| \Big| dx \\
&\le \int_{s_i-\delta}^{s_i+\delta} \Big| \ f(x)-f(s_i)-(x-s_i)g \ - \ f(x)-f(s_i)-(x-s_i)\overline{g} \ \Big| dx \\
&= \int_{s_i-\delta}^{s_i+\delta} \left|(x-s_i)\right|\left|g-\overline{g}\right| dx \\
&= \delta^2 \left|g-\overline{g}\right| \\
&< \varepsilon.
\end{aligned}
$$

Thus functions $\psi_i(.), i=1,2,...,m$ are continuous and hence function $T(g_1,...,g_m)=\sum_{i=1}^{m}\psi_i(g_i)$, $(g_1,...,g_m)\in\Delta$ is continuous. □

Theorem 2.4*:* Let $f\in C^1[0,1]$ and $s_i\in(\frac{i-1}{m},\frac{i}{m})$ for $i=1,2,...,m$ be arbitrary numbers. Then there is sufficiently small number $\delta>0$ such that the unique optimal solution of the optimization problem (6) is $f'(s_1),...,f'(s_m)$ where $\Delta=\mathbb{R}^m$.

Proof: Let $g_1,...,g_m \in\mathbb{R}^m$ be an arbitrary point and $g_i\ne f'(s_i)$ for $i=1,2,...,m$. By Proposition 2.2 and relation (2), there is $\delta>0$ such that

$$T\ \ f'(s_1),...,f'(s_m)\ < T\ \ g_1,...,g_m\ \ . \tag{7}$$

Thus

$$T\ \ f'(s_1),...,f'(s_m)\ < \underset{g_1,...,g_m\ \in\mathbb{R}^m,\ g_i\ne f'(s_i)}{Minimize}\ T\ \ g_1,...,g_m\ \ .$$

On the other hand, $f'(s_1),...,f'(s_m)\ \in\mathbb{R}^m$. Thus the point $f'(s_1),...,f'(s_m)$ is the unique optimal solution of the optimization problem (6) and

$$T\ \ f'(s_1),...,f'(s_m)\ = \underset{g_1,...,g_m\ \in\mathbb{R}^m}{Minimize}\ T\ \ g_1,...,g_m\ \ .$$

Theorem 2.5*:* Let $f:[0,1]\to\mathbb{R}$ be a nonsmooth integrable function, $m\in\mathbb{N}$ is a given large number, $s_i\in(\frac{i-1}{m},\frac{i}{m})$ for $i=1,2,...,m$ are arbitrary numbers, and $\delta>0$ is a sufficiently small number. The optimization problem (6) has an optimal solution on $\Delta=[-l,l]^m$ where $l>0$ is sufficiently big number.

Proof: By Theorem 2.3, function $T:[-l,l]^m\to\mathbb{R}$, defined in problem (6), is continuous and so has a minimum point $g_1^*,...,g_m^*$ on compact set $[-l,l]^m$. □

Now, define the GD of a nonsmooth integrable function $f:(0,1)\to\mathbb{R}$.

Definition 2.6: Let $f:(0,1)\to\mathbb{R}$ be an arbitrary integrable function, $m\in\mathbb{N}$ is a given large number and $s_i\in(\frac{i-1}{m},\frac{i}{m})$ for $i=1,2,...,m$ are arbitrary numbers. Moreover, suppose that $g_1^*,...,g_m^*$ is an optimal solution for the optimization problem (6). The GD of $f(.)$ on $(0,1)$ is denoted by $\partial f(.)$ and defined as $\partial f(s_i)=g_i^*$, $i=1,2,...,m$.

Remark 2.7: Considering the Theorem 2.4, if $f:(0,1)\to\mathbb{R}$ is a one-variable smooth function then $\partial f(s_i)=f'(s_i)$ where $s_i\in(0,1), i=1,2,...,m$. So in this case the GD is unique. Further, if $f(.)$ is a one-variable nonsmooth integrable function, then by theorem 2.5, the GD of $f(.)$ is an approximation for the derivative of $f(.)$. Moreover, we can show that there is $\lambda\in[0,1]$ such that

$$\partial f(s_i)=\lambda f'(s_i^-)+(1-\lambda)f'(s_i^+).$$

where $f'(s_i^-)$ and $f'(s_i^+)$ are the left and right derivative of function $f(.)$ in point s_i, respectively. Here, we are going to present an LP problem, corresponding to the optimization problem (6), for approximating the optimal solution of problem (6), i.e., the GD. For this, let $\delta>0$ be a sufficiently small number. Choose arbitrary points $s_i\in(\frac{i-1}{m},\frac{i}{m})$ for $i=1,2,...,m$ and suppose that $\varphi_i(x)=\left|f(x)-f(s_i)-(x-s_i)\,g_i\right|$ for all $x\in N_\delta(s_i)$. By the trapezoidal approximation

$$\int_{s_i-\delta}^{s_i+\delta}\varphi_i(x)\,dx\ \simeq\delta\ \ \varphi_i(x_{i1})+\varphi_i(x_{i2}) \tag{8}$$

where $x_{i1}=s_i-\delta$ and $x_{i2}=s_i+\delta$ for all $i=1,2,...,m$. We assume that $\varphi_{ij}=\varphi_i(x_{ij})$, $f_i=f(s_i)$, $f_{ij}=f(x_{ij})$ for $i=1,...,m$ and $j=1,2$. By approximation (8), optimization problem (6) is approximated with the following nonlinear optimization problem (see [1, 2]):

$$\begin{array}{ll} Minimize & \delta\sum_{i=1}^{m}\varphi_{i1}+\varphi_{i2} \\ subject\ to & \varphi_{ij}=\left|\,f_{ij}-f_i-(x_{ij}-s_i)\,g_i\right|, \\ & (g_1,g_2,...,g_m)\in\Delta,\ \ i=1,2,...,m\ \ \ ,j=1,2. \end{array} \tag{9}$$

where g_i and φ_{ij} for $i=1,2,...,m$ and $j=1,2$ are decision variables of the problem.

Lemma 2.8: Let the pair (u^*,v^*) be the optimal solution of the following LP problem:

$$\begin{array}{ll} Minimize & v \\ subject\ to & v\ge u,\ v\ge -u,\ u\in I \end{array}$$

where I is a compact set. Then u^* is the optimal solution of the following nonlinear problem:

$$\underset{u\in I}{Minimize}\ \left|u\right|.$$

Proof: It is obvious that $v^*\ge u^*, v^*\ge -u^*$ and thus $\left|u^*\right|\le v^*$. Now, Assume there exists $\bar{u}\in I$ such that $\left|\bar{u}\right|<\left|u^*\right|$. Define, $v=\left|\bar{u}\right|$ then we have $\bar{v}\ge\bar{u}, \bar{v}\ge-\bar{u}$. Thus $\bar{v}=\left|\bar{u}\right|<\left|u^*\right|\le v^*$ and so $\bar{v}<v^*$ which is a contradiction. □

According to Lemma 2.8, the nonlinear problem (9) is equivalent to the following LP problem:

$$\begin{array}{ll} Minimize & \delta\sum_{i=1}^{m} \varphi_{i1}+\varphi_{i2} \\ subject\ to & -\varphi_{ij}+(x_{ij}-s_i)\,g_i \leq f_{ij}-f_i, \\ & -\varphi_{ij}-(x_{ij}-s_i)\,g_i \leq -f_{ij}+f_i, \\ & (g_1,g_2,\dots,g_m)\in\Delta,\ i=1,\dots,m\ ,\ j=1,2. \end{array} \tag{10}$$

By solving LP problem (10), we obtain optimal solutions g_i^* and φ_{ij}^* for $i=1,2,\dots,m$ and $j=1,2$.

Thus we have $\partial f(s_i)=g_i^*$ for all $i=1,2,\dots,m$.

In the next section, we generalize the optimization problem (6) to the multi-variable nonsmooth functions.

3. Generalized derivative of multi-variable nonsmooth functions

In this section, we generalize the GD of nonsmooth one-variable functions to multi-variable nonsmooth functions. For this goal, assume that $q\in\ 1,2,\dots,n$ is fixed. Here, we define the GD of $f(.):\Omega\subset\mathbb{R}^n\to\mathbb{R}$ with respect to x_q, i.e. q^{th} component of $x=\ x_1,\dots,x_n\ \in\Omega$. Without loss of generality, assume $\Omega=[0,1]^n$ and define $\bar{\Omega}\subset\Omega$ as follows:

$$\bar{\Omega}=\ (x_1,\dots,x_{q-1},x_{q+1},\dots,x_n):x_j\in[0,1],\ j=1,2,\dots,n\ ,\ j\neq q\ .$$

Now, select N as a sufficiently large number and divide $\bar{\Omega}$ in to similar grids $\bar{\Omega}_j$, $j=1,2,\dots,N^{n-1}$ such that these grids cover $\bar{\Omega}$. In the next step, consider arbitrary points $s_j\in\bar{\Omega}_j,\ j=1,2,\dots,N^{n-1}$ as $s_j=(s_{j_1},\dots,s_{j_{q-1}},s_{j_{q+1}},\dots,s_{j_n})$. Moreover, define the following vectors for all $s_j\in\bar{\Omega}_j$:

$$r_j(t)=(s_{j_1},\dots,s_{j_{q-1}},t,s_{j_{q+1}},\dots,s_{j_n}),\qquad t\in(0,1),\ \ j=1,2,\dots,N^{n-1}. \tag{11}$$

Let $m\in\mathbb{N}$ be a given large number and $h_j^f(t)=f(r_j(t)),t\in[0,1]$ for $j=1,2,\dots,N^{n-1}$. Note that functions $h_j^f(.)$, $j=1,2,\dots,N^{n-1}$ are one-variable. Assume that $w_i\in(\frac{i-1}{m},\frac{i}{m}),i=1,2,\dots\ \dots,m$ are arbitrary numbers and define the optimization problem for $j=1,2,\dots,N^{n-1}$:

$$\underset{g_1,\dots,g_m\ \in\Delta}{Minimize}\quad T\ \ g_1,\dots,g_m\ =\sum_{i=1}^{m}\int_{w_i-\delta}^{w_i+\delta}\left|h_j^f(t)-h_j^f(w_i)-(t-w_i)g_i\right|dt \tag{12}$$

where $\delta>0$ is a sufficiently small number and Δ is a subspace of $\mathbb{R}^m$.

Theorem 3.1: Let $m\in\mathbb{N}$ be a given large number, $f\in C^1([0,1]^n)$ and $h_j^f(t)=f(r_j(t)),t\in[0,1]$ for $j=1,2,\dots,N^{n-1}$ where $r_j(.)$ is satisfied in (11). Then there is a sufficiently small number $\delta>0$ such that the unique optimal solution of the optimization problem

(12) for $j=1,2,\dots,N^{n-1}$, is $\left(\frac{d}{dt}h_j^f(w_1),\dots,\frac{d}{dt}h_j^f(w_m)\right)$ and we have

$$\frac{d}{dt}h_j^f(w_i) = \frac{\partial f}{\partial x_q}(r_j(w_i)),\ i = 1,2,\dots,m.$$

Proof: The proof is similar to the proof of Theorem 2.4.

Definition 3.2: Let function $f:[0,1]^n \to \mathbb{R}$ be a nonsmooth integrable function and $g_i^{j,*}(.)$, $i = 1,2,\dots,m$ are the optimal solutions of the functional optimization problem (12), for , $j = 1,2,\dots,N^{n-1}$. We denote the GD of $f(.)$ with respect to variable x_q by $\partial_q f(.)$ and define as $\partial_q f(r_j(w_i)) = g_i^{j,*}$ for $i = 1,2,\dots,m$ and $j = 1,2,\dots,N^{n-1}$.

Remark 3.3: Considering Theorem 3.1, if $f:[0,1]^n \to \mathbb{R}$ is a multi-variable smooth function then for $s_j = (s_{j_1},\dots,s_{j_{q-1}},s_{j_{q+1}},\dots,s_{j_n}) \in \bar{\Omega}_j$, $j = 1,2,\dots,N^{n-1}$

$$\partial_q f(s_{j_1},\dots,s_{j_{q-1}},w_i,s_{j_{q+1}},\dots,s_{j_n}) = \frac{\partial f}{\partial x_q}(s_{j_1},\dots,s_{j_{q-1}},w_i,s_{j_{q+1}},\dots,s_{j_n}), \qquad w_i \in (0,1).$$

Further, if $f(.)$ is a multi-variable nonsmooth integrable function, then the GD of $f(.)$ with respect to x_q is an approximation for partial differentiation of $f(.)$ with respect to x_q .

Here, as in Section 2, we approximate the optimization problem (12) for $j = 1,2,\dots,N^{n-1}$ with a LP problem. For this, let $\delta > 0$ be a given small number. Choose arbitrary points $w_i \in (\frac{i-1}{m},\frac{i}{m})$ for $i = 1,2,\dots,m$ and suppose that $\varphi_{ij}(t) = \left| h_j^f(t) - h_j^f(w_i) - (t - w_i)\,g_i \right|$ for all $t \in [w_i - \delta, w_i + \delta]$. Moreover, for $j = 1,2,\dots,N^{n-1}$ and $i = 1,2,\dots,m$, we define

$$h_{ij}^f = h_j^f(w_i),\ h_{ij1}^f = h_j^f(w_i - \delta),\ h_{ij2}^f = h_j^f(w_i + \delta),$$
$$t_{i1} = w_i - \delta,\ \ t_{i2} = w_i + \delta,\ \ \varphi_{ij1} = \varphi_{ij}(w_i - \delta), \varphi_{ij2} = \varphi_{ij2}(w_i + \delta).$$

Now, following Section 2, we can approximate the optimization problem (12) by an LP problem with the decision variables φ_{ijk} and g_i for $i = 1,\dots,m$, $k = 1,2$, as follows:

$$\begin{aligned} &Minimize && \delta\sum_{i=1}^{m} \varphi_{ij1} + \varphi_{ij2} \\ &subject\ to && -\varphi_{ijk} + (t_{ik} - w_i)\,g_i \le h_{ijk}^f - h_{ij}^f \\ & && -\varphi_{ijk} - (t_{ik} - w_i)\,g_i \le -h_{ijk}^f + h_{ij}^f \\ & && (g_1, g_2,\dots,g_m) \in \Delta,\ i = 1,\dots,m\ ,\ k = 1,2. \end{aligned} \tag{13}$$

Note that if $g_i^{j,*}$, $i = 1,\dots,m$ are the optimal solutions of LP problems (13) for $j = 1,2,\dots,N^{n-1}$, then $\partial_q f(s_{j_1},\dots,s_{j_{q-1}},w_i,s_{j_{q+1}},\dots,s_{j_n}) = g_i^{j,*}$ for $i = 1,\dots,m$. Moreover, in LP problems (10) and (13), we can assume that $\Delta = \mathbb{R}^m$, in other words, we can suppose that the variables g_i for $i = 1,2,\dots,m$ are free variables.

4. Simulation results

Here, we obtain the GD of some nonsmooth functions. The LP problems (10) and (13) are solved by revised simplex method and LINPROG function in MATLAB software.

4.1. Obtaining the GD of some nonsmooth one-variable functions

For solving problem (10), in the examples of this subsection, we assume that

$$\delta = 0.01, \ m = 99, \ s_i = 0.01i, \ x_{i1} = s_i - 0.01, \ x_{i2} = s_i + 0.01, \ i = 1, 2, \ldots, 99.$$

Note that for all $m \geq 99$, in the following examples, we don't see tangible change in the GD of nonsmooth functions.

Example 4.1.1: Consider the nonsmooth continuous function $f(x) = \left|4x^2 - 4x + \frac{3}{4}\right|$ on the interval $\Omega = (0,1)$. This function is illustrated in Figure 1. It is obvious that the nonsmoothness points of this function are $x = 0.25$ and $x = 0.75$. By solving the corresponding problem (10) for this function, we obtain the optimal solutions $\partial f(s_i) = g_i^*$ and φ_{ij}^* for $i = 1, 2, \ldots, 99$ and $j = 1, 2$. Figure 2 shows the resulting GD, i.e. $\partial f(.)$.

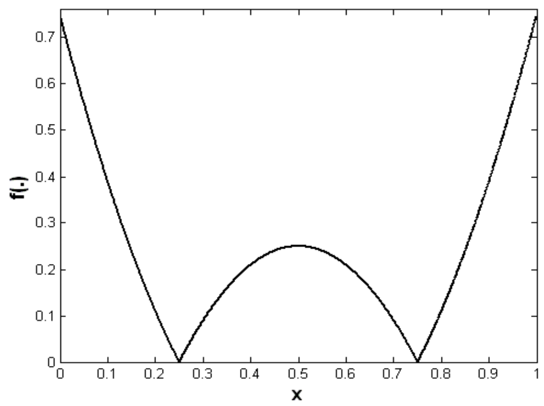

Fig. 1: The graph of function $f(.)$ for Ex. 4.1.1.

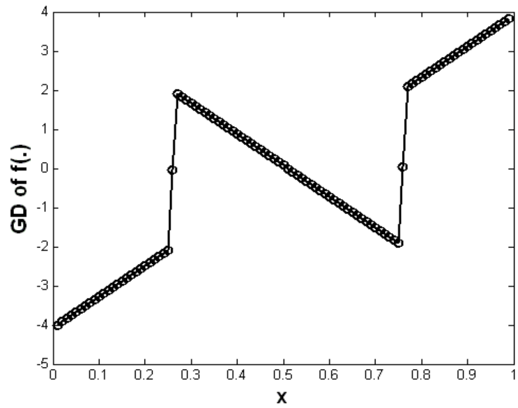

Fig. 2: The graph of $\partial f(.)$ for Ex. 4.1.1.

Example 4.1.2: Consider the nonsmooth continuous function $f(x) = |\sin(3\pi x)\cos(4\pi x)|$ on the interval $(0,1)$. It is differentiable on (0,1) except at the points in the set $x : \sin(3\pi x) = 0 \text{ or } \cos(4\pi x) = 0$. In Figure 3, we show this function on interval (0,1) which it is continuous but not differentiable. We solve the problems (10) corresponding to this function and gain the optimal solutions $\partial f(s_i) = g_i^*$ and φ_{ij}^* for $i = 1, 2, \ldots, 99$ and $j = 1, 2$. Figure 4 shows the graph of the computed GD of function $f(.)$.

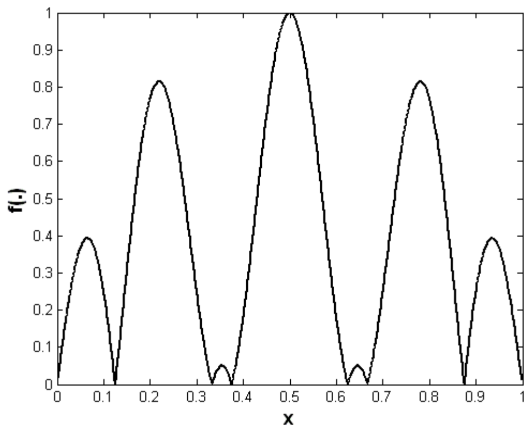

Fig. 3: The graph of function $f(.)$ for Ex. 4.1.2.

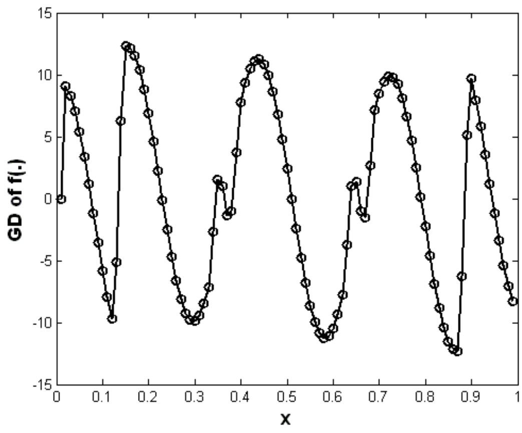

Fig. 4: The graph of $\partial f(.)$ for Ex. 4.1.2.

Example 4.1.3: Consider the following functions

$$f_1(x) = |x-0.5|,\ f_2(x) = -|x-0.5|,\ f_3(x) = |x-0.5|^{0.5},\ f_4(x) = \text{sign}(x-0.5)|x-0.5|^{0.5}$$

on $x \in [0,1]$. We obtain the GD of these functions using the LP problem (10) on the interval $[0,1]$ and compare it with proximal subdifferential, strict subdifferential, and limiting subdifferential at point $x = 0.5$ (see page 135 of Vinter [19]). Table 1 shows the obtained results. Note that by attention to Table 1, our GD is better than the above-mentioned subdifferentials, since it has a specific numeric value (corresponding to the definition of GD) and it is not an empty or infinite set.

Table 1: Comparison of the GD with the subdifferentials of functions $f_i(.), i = 1,2,3,4$ at point $x = 0.5$ for Ex. 4.1.3.

	$f_1(.)$	$f_2(.)$	$f_3(.)$	$f_4(.)$
Proximal subdifferential	$[-1,1]$	ϕ	$(-\infty,\infty)$	ϕ
Strict subdifferential	$[-1,1]$	ϕ	$(-\infty,\infty)$	ϕ
Limiting subdifferential	$[-1,1]$	$\{-1,1\}$	$(-\infty,\infty)$	ϕ
GD for $\delta = 0.1$ and $m = 9$	-9×10^{-11}	-3×10^{-12}	3×10^{-9}	3.162
GD for $\delta = 0.02$ and $m = 49$	3×10^{-11}	7×10^{-11}	1×10^{-9}	7.071
GD for $\delta = 0.01$ and $m = 99$	2×10^{-9}	-2×10^{-9}	3×10^{-8}	10.000
GD for $\delta = 0.005$ and $m = 199$	5×10^{-10}	-4×10^{-9}	8×10^{-7}	14.142

4.2. Obtaining the GD of some multi-variable nonsmooth functions

In this subsection, we find the GD of some nonsmooth functions $f = f(x_1,x_2),\ (x_1,x_2) \in (0,1)^2$ with respect to x_1 (namely $\partial_1 f(.,.)$) in some examples using LP problem (13). We assume $N = 100$ and devide $\bar{\Omega} = (0,1)$ to the similar grids $\bar{\Omega}_j,\ j = 1,2,\ldots,N$ and select points $s_j \in \bar{\Omega}_j$ as $s_j = 0.005 + 0.01(j-1),\ j = 1,2,\ldots,N$. Moreover, we suppose

$$\delta = 0.01,\ \text{m=99},\ \text{w}_i = 0.01i,\ t_{i1} = \text{w}_i - 0.01,\ t_{i1} = \text{w}_i + 0.01,\ i = 1,2,\ldots,99.$$

Indeed, we set $r_j(t) = (t,s_j)$ for $j = 1,2,\ldots,N$ and $t \in (0,1)$.

Example 4.2.1: Consider the function $f(x_1,x_2) = |x_2 - 0.5| - x_2|x_1 - 0.5|$ on $(x_1,x_2) \in (0,1)^2$ which is a nonsmooth function in the points of the set $\{(x_1,x_2) \in (0,1)^2 : x_1 = 0.5 \text{ or } x_2 = 0.5\}$. This function is graphed in Figure 5. The GD of this function with respect to x_1 i.e. $\partial_1 f(.,.)$, is found Using the LP problem (13). The graphed of $\partial_1 f(.,.)$ is shown in Figure 6.

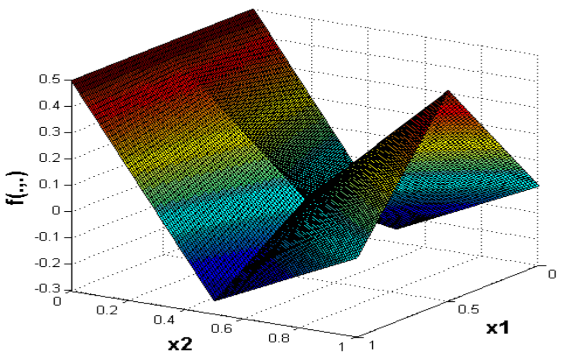

Fig. 5: The graph of function $f(.,.)$ for Ex. 4.2.1.

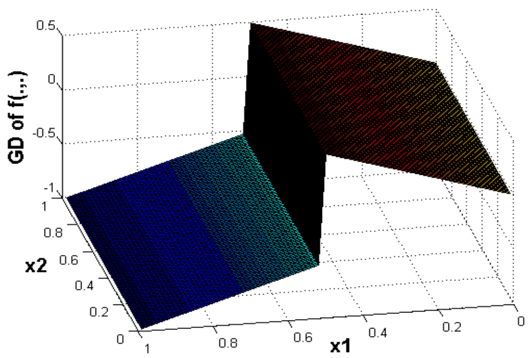

Fig. 6: The graph of $\partial_1 f(.,.)$ for Ex. 4.2.1.

Example 4.2.2: Consider the function $f(x_1,x_2)=x_2\left|\sin(4\pi x_1)\right|-\left|\cos(2\pi x_2)\right|$ on $(x_1,x_2)\in(0,1)^2$. Its graph is shown this function in Figure 7. Observe that this is nonsmooth in points of the set $\{(x_1,x_2)\in(0,1)^2 : \sin(4\pi x_1)=0 \text{ or } \cos(2\pi x_2)=0\}$. The LP problem (13) is used to obtain the GD of this function with respect to x_1, i.e. $\partial_1 f(.,.)$. It is graphed in Figure 8.

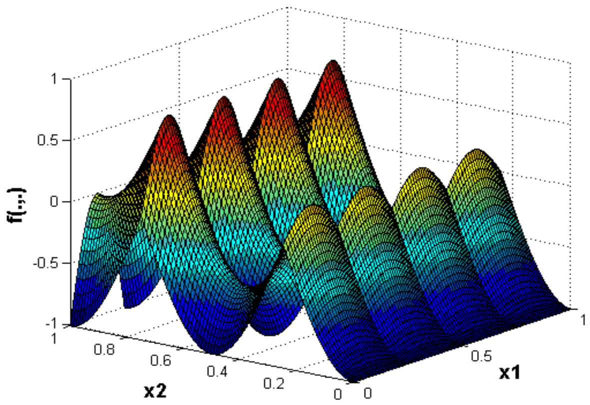

Fig. 7: The graph of function $f(.,.)$ for Ex. 4.2.2.

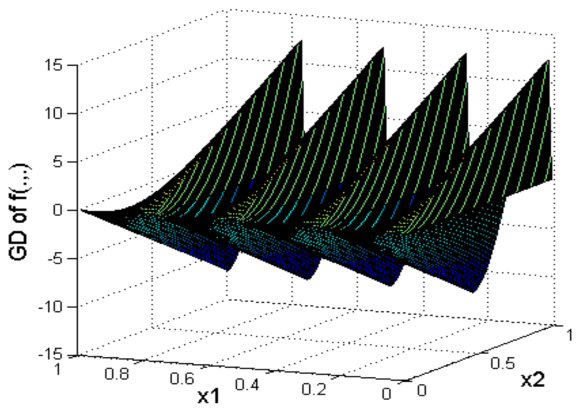

Fig. 8: The graph of $\partial_1 f(.,.)$ for Ex. 4.2.2.

5. Conclusions and suggestions

In this paper, we defined a new practical GD for multi-variable nonsmooth functions as the optimal solution of an optimization problem on an interval. We solved the optimization problem, which corresponds to the original nonsmooth function, by using discretization method where finally is converted to an LP problem. This GD for smooth functions coincides with the exact differentiations of these functions and for nonsmooth functions it is an approximation for differentiations. Four advantages of our GD with respect to the other approaches are as follows:

I. The GD of a nonsmooth function by our approach does not depend on the nonsmoothness points of function. Thus we can use this GD for many cases that we do not know the points of non-differentiability of the function.

II. The GD of nonsmooth functions by our approach gives a good global approximate derivative as on the domain of functions, whereas in the other approach the GD is calculated at one given point.

III. The GD by our approach is defined for nonsmooth integrable functions, whereas the other approaches is defined usually for locally Lipschiptz or convex functions.

IV. The GD in our approach is obtained by solving an LP problem, whereas the other approaches for GD, are based on lim, liminf and limsup (see chapter 1 of Jeyakumar [12]) and calculating the GD is very hard and usually is not possible.

We suggest that the GD of this paper can be used to approximate the optimal solutions of nonsmooth discrete and continuous-time optimization problems, practically and applicably. Moreover, we can apply this GD to approximate the solutions of nonsmooth algebraic systems.

References:

[1] Bazaraa, M.S., Javis, J.J. and Sheralli, H.D., *Linear programming*, New York; published by Wiley & Sons, 1990.

[2] Bazaraa, M. S., Sheralli, H.D., Shetty, C.M., Nonlinear programming: Theory and Application, New York; published by Wiley & Sons, 2006.

[3] Clarke, F. H., *Necessary conditions for nonsmooth problem in optimal control and the calculus of variations*, Ph.D. dissertation, Department of mathematics, university of Washington, Seattle, 1973.

[4] Clarke, F. H., *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[5] Clarke, F. H., *Nonsmooth analysis and Control theory,* Springer, New York, 1998.

[6] De Los Reyes, J. and Kunisch, K., *A semi-smooth newton method for control constrained boundary optimal control of the navier-stokes equations*, Nonlinear Anal., Theory Methods Appl., Vol.62, pp. 1289-1316, 2005.

[7] Demyanov, V. F. and Rubinov, A. M., *Constructive Nonsmooth Analysis*, Verlag Peter Lang, 1995.

[8] Gowda, M. S. and Ravindran, G., *Algebraic univalence theorems for nonsmooth functions,* J. Math. Anal. Appl. Vol. 252, pp. 917–935, 2000.

[9] Hintermuller, M., Ito, K. and Kunisch, K., *The primal-dual active set strategy as a semi-smooth Newton method*, SIAM Journal on Optimization,Vol. 13, pp. 865-888, 2002.

[10] Ioff, A. D., *Nonsmooth analysis: Differential calculus of nondifferentiable mappin*gs, Trans. Amer. Math. Soc. Vol. 266, pp. 1-56, 1981.

[11] Ioff, A. D., *Necessary conditions in nonsmooth optimization*, Math. Oper. Res. Vol. 9. pp. 159-188, 1984.

[12] Jeyakumar, V. and Luc, D.T., *Nonsmooth vector functions and continuous optimization*, Springer, 2008.

[13] Kroner, A., Kunisch, K. and Vexler, B., *Semismooth Newton Methods for optimal control of wave equation with control constraints*, SIAM j. control optim., Vol. 49, No. 2, pp. 830–858, 2011.

[14] Kunisch, K., *Semi-smooth Newton Methods for nondifferentiable optimization problems*, Lipschitz Lectures, February 19, 2008.

[15] Kamyad, A. V., Noori Skandari, M. H. and Erfanian, H. R., *A new definition for generalized first derivative of nonsmooth functions*, Applied Mathematics, Vol. 2, pp. 1252-1257, 2011.

[16] Mordokhovich, B. S., *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl. Vol. 183, pp. 250-288, 1994.

[17] Mordokhovich, B., *Variatinal analysis and generalized differentiation*, Vol 1 and 2, Springer, New York, 2006.

[18] Rockafellar, R. T. and Wets, R.J.B., *Variational Analysis,* Springer, New York, 1997.

[19] Vinter, R., Optimal Control, Birkhauser, 2000.

[20] Warga, J., *Derivative containers, inverse functions, and controllability, in Calculus of Variations and Control Theor*y, D.L. Russel, Ed., Academic Press, New York, 1976.

[21] Warga, J., *A implicit function theorem without differentiability,* Proc. Amer. Math. Soc. Vol. 69, pp. 65-69, 1978.

[22] Sun*, Z.*, Zeng, J. *and* Xu, H., *Generalized Newton-iterative method for semismooth equations*, Numer Algor , Vol. 58, pp. 333–349, 2011.