A Note on Directional Differentiability of Max-Functions

白石俊輔

富山大学経済学部富大経済論集　第38巻第3号抜刷（1993年3月）

A Note on Directional Differentiability of Max-Functions

Shunsuke Shiraishi

1. Introduction.

In this paper we are concerned with the *max* (*sup*)*-function* defined as:

$$S(x) := \sup\{f(x,t) \mid t \in T\},$$

where T is a closed set in a metric space and $f: \mathbb{R}^n \times T \to \mathbb{R}$ is a *continuous* function. Max-functions often appear in the area of mathematical programming: e.g., in game theory, in duality theory, in numerical optimization as a sort of penalty functions, in Chebychev approximation problems, etc. and form a very important class in *Nonsmooth Analysis*. In particular, we emphasize that *convex functions* are also in this class. So it is worth to study their properties. The main interests are in their continuity and differentiability. For max-functions, non-differentiability arises naturally. Fortunately, however, it occurs frequently that the directional derivative of the max-functions exists and, if it does, we are content enough in the nonsmooth context (see Zowe [8] and Lemarechal [6]). A most well-known directional differentiability result, which seems to be due to Danskin, is:

Theorem 1. (Danskin [2]) *Assume that:*

(A) *T is a compact set,*

(B) *for each $t \in T$, $\nabla_x f(x,t)$ exists and is continuous on $\mathbb{R}^n \times T$.*

Then the (one-sided) directional derivative of S

$$S'(x;d) := \lim_{\lambda \to 0^+} [S(x+\lambda d) - S(x)] / \lambda$$

exists at all x in every direction d, and it can be expressed as

$$S'(x;d) = \max \{ < \nabla_x f(x,t), d > \mid t \in M(x) \}, \tag{1}$$

where $M(x) := \{ t \in T \mid f(x,t) = S(x) \}$.

In the sequel, we fix the base point x and the direction d. Generalization of this result can be done in two ways. First one is to weaken the assumption (A). Instead of assuming compactness of T, an assumption, which is just like compactness, will be made on the multifunction $M: \mathbb{R}^n \to T$.

Assumption (A'): The multifunction M is nonempty-valued and *uniformly compact* near x,i.e., there exists a neighborhood U of x such that $M(y)$ is nonempty for all $y \in U$ and cl $\cup_{y \in U} M(y)$ is compact, where cl C denotes the closure of a set C.

The following theorem is due to Auslender:

Theorem 2. (Auslender [1]) *Under the assumptions* (A') *and* (B), *S is directionally differentiable at the considered point x in the direction d and the expression formula* (1) *holds.*

Remark. One may think that the assumption (A') seems to be rather artificial. But it is automatically satisfied if T is compact, of course, and functions satisfying (A') for noncompact T exist. For example, convex functions, typically distance functions to a nonempty closed convex set, satisfy (A') (see Auslendr [1] and Shiraishi [7]). Auslender uses the uniform boundedness condition of M instead of the uniform compactness condition of it,i.e. $\cup_{y\in U}M(y)$ is bounded. However, since he works in a finite dimensional Euc lidean space setting, both conditions are the same.

Second generalization is to weaken the differentiability assumption (B) of the constituents f. Given a metric space (X, ρ), for a nonempty subset Ω of X and a positive number δ, we define the closed δ-neighborhood of Ω by

$$B(\Omega;\delta):=\{s\in X \mid \exists t\in\Omega;\ \rho(s,t)\leqq\delta\}.$$

Set $T(\delta):=B(M(x);\delta)\cap T$. Set also $f_t(\cdot)=f(\cdot,t)$ for $t\in T$.

Assumption (B'): For each $t\in T$, the directional derivative $f_t'(x;d)$ exists and for some positive number δ it holds that

$$[f_t(x+\lambda d)-f_t(x)]/\lambda \rightarrow f_t'(x;d) \text{ uniformly in } t\in T(\delta) \text{ as } \lambda\rightarrow 0^+.$$

Theorem 3. (Furukawa [3]) *Under the assumptions* (A) and (B'), *S is directionally differentiable at the considered point x in the direction d and the following expression formula holds instead of* (1):

$$S'(x;d)=\max\{f_t'(x;d) \mid t\in M(x)\} \tag{2}$$

The purpose of this note is to show the directional differentiability of S under the assumptions (A') and (B') in place of (A) and/or (B). In the next section, we investigate the continuity and the directional differentiabiliy of S. Examples which illustrate the necessity of (A') and (B') are also displayed.

2. Main results.

We begin with a definition of (semi-) continuity of multifunctions ([4]). Let $\Gamma : \mathbb{R}^n \to T$ be a multifunction. Γ is *closed* or *upper semi-continuous* at $x \in \mathbb{R}^n$ provided that $\mathbb{R}^n \supset \{x_k\}$, $x_k \to x$, $t_k \in \Gamma(x_k)$, and $t_k \to t$ imply $t \in \Gamma(x)$.

Proposition. *Under the assumptions* (A´), *the function S is continuous and the multifunction M is closed at x.*

Proof. Since S is a pointwise supremum of continuous functions, it is lower semi-continuous. Let $\mathbb{R}^n \supset \{x_k\}$ be such that $x_k \to x$. It follows from (A´) that there exist $T \supset \{t_k\}$ with $t_k \in M(X_k)$. Also, by taking a subsequence, we may assume $t_k \to t \in T$. Then by the continuity of f and the definition of S itself,

$$S(x) \geq f(x,t) = \lim_{k \to \infty} f(x_k, t_k) = \limsup_{k \to \infty} S(x_k),$$

which means S is upper semi-continuous at x. Closedness of M is a immediate consequence of the continuitity of S and f.

Q.E.D.

Before proving the directional differentiability, we need a small lemma.

Lemma. *Assume* (A´). *Given a positive number* δ_0, *there exists a neighborhood V of x such that* $T(\delta_0) \supset \cup_{y \in V} M(y)$.

Proof. If the conclusion were false, then there exist $\mathbb{R}^n \supset \{x_k\}$ with $x_k \to x$ and $t_k \in M(x_k)$ such that $\rho(t_k, s) > \delta_0$ for all $s \in M(x)$. From the uniform compactness and the closedness of M at x, we may assume t_k

$\to t \in M(x)$. By taking a limit, we have $\rho(t,s) \geq \delta_0$ for all $s \in M(x)$, *which leads to a contradiction.*

Q.E.D.

Theorem 4 . *Under the assumptions* (A´) *and* (B´) , *S is directionally differentiable at the considered point x in the direction d and the expression formula* (2) *holds.*

Proof. Since $f_t'(x;d)$ is a uniform convergent limit of continuous functions $t \to [f_t(x+\lambda d) - f_t(x)] / \lambda$, the function $t \to f_t'(x;d)$ is also continuous on $T(\delta)$. Hence for any $\varepsilon > 0$, there exists $\delta' > 0$ such that

$$\max\{f_t'(x;d) \mid t \in T(\delta')\} \leq \max\{f_t'(x;d) \mid t \in M(x)\} + \varepsilon / 2. \tag{3}$$

If we set $\delta_0 := \min(\delta, \delta')$, $W := U \cap V$ and $\hat{T} := \mathrm{cl} \cup_{y \in W} M(y)$, then the above lemma says that $T(\delta_0) \supset \hat{T}$ and $S(y) := \max\{f(y,t) \mid t \in \hat{T}\}$ holds for all $y \in W$. Also, by (B´), there exists $\lambda' > 0$ such that

$$| f_t(x+\lambda d) - f_t(x) - \lambda f_t'(x;d) | \leq \lambda \varepsilon / 2, \tag{4}$$

for all $\lambda \in]0, \lambda'[$ and $t \in T(\delta_0)$. For sufficiently small $\lambda > 0$, we have

$$\begin{aligned} S(x+\lambda d) &= \max_{t \in \hat{T}} f_t(x+\lambda d) \\ &\leq \max_{t \in \hat{T}} f_t(x) + \lambda \max_{t \in \hat{T}} f_t'(x;d) + \lambda \varepsilon / 2 \quad \text{(by (4))} \\ &= S(x) + \lambda \max_{t \in \hat{T}} f_t'(x;d) + \lambda \varepsilon / 2 \\ &\leq S(x) + \lambda \max_{t \in M(x)} f_t'(x;d) + \lambda \varepsilon. \quad \text{(by (3))} \end{aligned}$$

Thus for sufficiently small $\lambda > 0$:

$$S(x+\lambda d) \leq S(x) + \lambda \max_{t \in M(x)} f_t'(x;d) + \lambda \varepsilon. \tag{5}$$

On the other hand, using the relation $\hat{T} \supset M(x)$, for sufficiently small $\lambda > 0$, we have

$$S(x+\lambda d) = \max_{t\in\hat{T}} f_t'(x;d)$$
$$\geq \max_{t\in M(x)} \{f_t(x) + \lambda f_t'(x;d) - \lambda \varepsilon / 2\}$$
$$= S(x) + \lambda \max_{t\in M(x)} f_t'(x;d) - \lambda \varepsilon / 2$$
$$> S(x) + \lambda \max_{t\in M(x)} f_t'(x;d) - \lambda \varepsilon . \tag{6}$$

Combining (5) and (6) yields that for sufficiently small $\lambda > 0$:

$$| [S(x+\lambda d) - S(x)] / \lambda - \max_{t\in M(x)} f_t'(x;d) | \leq \varepsilon .$$

Since ε is arbitrary, the last inequality shows

$$\lim_{\lambda \to 0^+} [S(x+\lambda d) - S(x)] / \lambda = \max_{t\in M(x)} f_t'(x;d).$$

Q.E.D.

3. Examples.

The first example shows the necessity of uniform boundedness in the assumption (A´)

Example 1.

Let $T := \mathbb{R}_+$ and $f(x,t) := -1/2(t\sqrt{t} + t^2\sqrt{t})x^2 + (3/2\sqrt{t} + t\sqrt{t})x - 1/2\sqrt{t}$.
Then an easy calculation shows that:

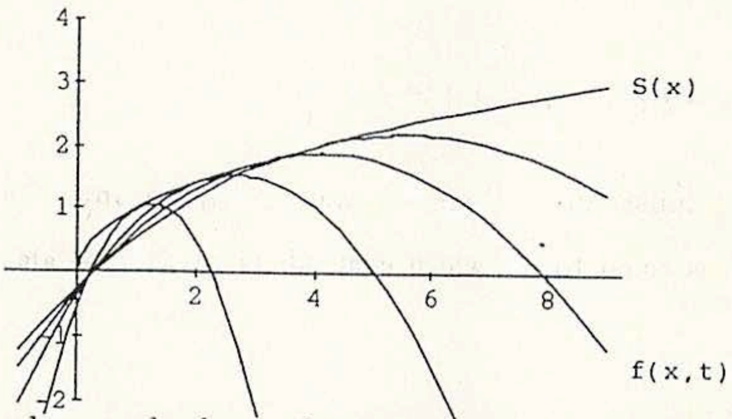
$$S(x) = \begin{cases} \sqrt{x}, & \text{if } x \geq 0, \\ 0, & \text{if } x \leq 0, \end{cases}$$

and

$$M(x) = \begin{cases} 1/x, & \text{if } x \geq 0, \\ 0, & \text{if } x \leq 0. \end{cases}$$

Hence $M(x)$ is not uniformly bounded near $x=0$ and S is not directionally differentiable at x in the direction $d=1$. See Figure 1.

Figure 1



The second example shows the necessity of the uniform convergence in the assumption (B').

Example 2. (Kawasaki [5])

Set $t_n = 3^{-n}$, $P_n := (t_n, 0)$, and $Q_n := (2t_n/3, -2t_n/3)$ for n=0, 1, 2,⋯. Let G denote the polygonal curve determined by P_0, Q_0, P_1, Q_1, P_2, Q_2,⋯. Let $g(t)$ denote the continuous function defined on $T := [0,1]$ whose graph is G. Define $f: \mathbb{R}^n \times T \to \mathbb{R}$ by

$$f(x,t) := g(t) - 2|x-t|.$$

Then it is easily seen that $S(x) = g(x)$ for all $x \in [0,1]$. Hence S is not directionally differentiable at $x=0$ in the direction $d=1$. The reason of the failure of the directional differentiability is that the assumption (B′) is not satisfied. Indeed, for any $s>0$ and $t \in [0,1]$,

$$f(0+s,t) - f(0,t) = 2t - 2|s-t|$$

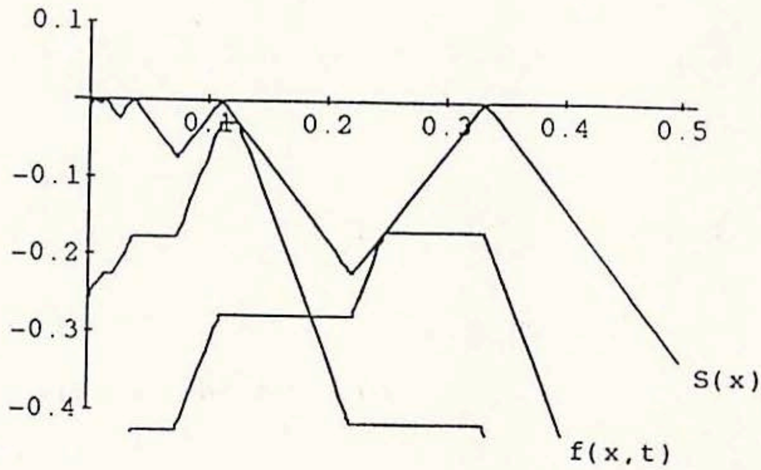
$$= \begin{cases} 4t - 2s \text{ if } s \geq t, \\ 2s, \qquad \text{if } s \leq t. \end{cases}$$

Suppose that (B̸) is satisfied. Then, there exists $t_0 > 0$ satisfying the following property: for any $\varepsilon > 0$, there exists $t_0 > s_0 > 0$ such that for evey $0 < s < s_0$ it hold that

$$| [2t - 2 | s - t |] / s - 2 | < \varepsilon \tag{7}$$

for all $t \in] 0, t_0]$. Substitute $s = s_0 / 2$ and $t = s_0 / 4$ into (7). Then the left-hand side of (7) is equal to 2 , which contradicts to (7). See also Figure 2 .

Figure 2



References.

[1] A.Auslender (1976), *Optimisation. Méthodes Numérique*. (Masson)

[2] J.M.Danskin (1967), *The theory of max-min*. (Springer)

[3] N.Furukawa (1983), "Optimality conditions in nondifferentiable programming and their applications to best approximations," Applied Math. Optim.9,337-371.

[4] W.W. Hogan (1973), "Point-to-set maps in mathematical programming," SIAM Review 15, 591-603.

[5] H.Kawasaki (1988), *Personal communication*.

[6] C. Lemaréchal (1989), "Nondifferentiable optimization," in: G.L. Nemhauser et al eds., *Handbooks in OR&MS, Vol.1* (North-Holland), 529-572.

[7] S. Shiraishi (1993), "Directional differentiability of convex functions," (in Japanese) to appear in RIMS kokyuroku.

[8] J. Zowe (1985), "Nondifferentiable optimization," in :K. Schittkowski ed., *Computational Mathematical Programming*. (Springer), 323-356.