FIRST- AND SECOND-ORDER DIRECTIONAL DERIVATIVES OF A MAX-TYPE FUNCTION INDUCED FROM AN INEQUALITY STATE CONSTRAINT

Kawasaki, Hidefumi
Graduate School of Mathematics, Kyushu University

FIRST- AND SECOND-ORDER DIRECTIONAL DERIVATIVES OF A MAX-TYPE FUNCTION INDUCED FROM AN INEQUALITY STATE CONSTRAINT *

By

Hidefumi KAWASAKI[†]

Abstract

In this paper, we deal with a max-type function $S(x) := \max_{t \in T} f(x(t), t)$, where x is a n-dimensional vector-valued continuous functions. This max-type function is induced from an inequality state constraint $f(x(t), t) \leq 0$, which appears in variational problems and optimal control problems. We give formulae for first- and second-order directional derivatives of $S(x)$. We show that the one-side state constraint $x(t) \geq a(t)$ always forms an envelope except two trivial cases.

1. Introduction

In this paper, we deal with a max-type function:

$$S(x) := \max_{t \in T} f(x(t), t) \quad x \in X, \tag{1.1}$$

where T is a compact metric space, X is a subspace of the set of all n-dimensional vector-valued continuous functions $C(T)^n$ equipped with the uniform norm, and $f(x, t)$ is a continuous function defined on $R^n \times T$. This max-type function is induced from an inequality state constraint:

$$f(x(t), t) \leq 0 \quad \text{for all } t \in T, \tag{1.2}$$

which appears in variational problems and optimal control problems. For instance, a variational problem to find the shortest path in R^2 joining two given points P and Q that does not transverse the unit ball is formulated as follows:

$$\text{Minimize} \quad \int_0^1 \sqrt{\dot{x}_1^2 + \dot{x}_2^2} dt$$

$$\text{subject to} \quad (x_1(0), x_2(0)) = P, \quad (x_1(1), x_2(1)) = Q,$$

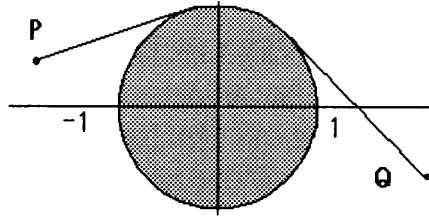$$1 - x_1(t)^2 - x_2(t)^2 \leq 0 \quad {}^\forall t \in [0, 1].$$

Fig. 1.1

There are two aims in this paper. First one is to give formulae for first- and second-order directional derivatives of $S(x)$. The other is to show that the one-sided state constraint

$$x(t) \geq a(t) \quad \text{for every } t, \tag{1.3}$$

where $a(t)$ is a given continuous function, always forms an envelope except two trivial cases:

$$x(t) \equiv a(t), \tag{1.4}$$

$$x(t) > a(t) \quad \text{for every } t. \tag{1.5}$$

By the way, there are many papers that dealt with another max-type function:

$$S_0(x) := \max_{t \in T} f(x, t), \quad x \in R^n \tag{1.6}$$

or an optimal value function:

$$S_1(x) := \max_{t \in T_x} f(x, t), \quad x \in R^n, \tag{1.7}$$

where T_x is a subset of T that depends on x, see Clarke (1975), Correa (1985), Danskin (1967), Dem'yanov (1974), Demyanov and Zabrodin (1986), Hettich and Jongen (1978), Ioffe (1989), Kawasaki (1988a) (1988b) (1991) (1992) (1995), Shiraishi (1995), Seeger (1988), and Wetterling (1970). The functions $S_0(x)$ and $S_1(x)$ appear in Tchebycheff approximation and sensitivity analysis, respectively. If we take as X the set of all constant functions $\{x \in C(T)^n \mid x(t) \equiv \text{constant } n\text{-vector}\}$, then $S(x)$ reduces to $S_0(x)$. So $S(x)$ inherits some properties from $S_0(x)$, which will be discussed in Sections 2 and 3. By the way, the max-type function $S_0(x)$ often forms an envelope, and so does $S(x)$. Concerning the one-sided state constraint, we will prove, in Section 4, a stronger result, that is, it always forms an envelope except two trivial cases.

2.   Continuity and Differentiability

We begin this section with our notation. We denote by f_x and f_{xx} the gradient (row) vector and the Hesse matrix of f w.r.t. x, respectively. We denote by $T(x)$ the set of all extreme points of $f(x(t), t)$, that is,

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

THEOREM 2.1. *The max-type function $S(x)$ is continuous on X.*

PROOF. Assume that x_n converges to x. Then, since $f(x(t), t)$ is continuous on the compact set T, there exists $t_n \in T$ such that

$$S(x_n) = f(x(t_n), t_n). \tag{2.1}$$

Now, let $x_{n'}$ be a subsequence satisfying

$$\limsup_{n \to +\infty} S(x_n) = \lim_{n' \to +\infty} S(x_{n'}). \tag{2.2}$$

By choosing a subsequence of $x_{n'}$ again, we may assume that $t_{n'}$ converges to some $\bar{t} \in T$. Then it is easily seen that

$$\lim_{n' \to +\infty} f(x_{n'}(t_{n'}), t_{n'}) = f(x(\bar{t}), \bar{t}) \leq S(x). \tag{2.3}$$

Combining (2.1), (2.2), and (2.3), we have $\limsup_{n \to +\infty} S(x_n) \leq S(x)$. On the other hand, since $S(x_n) \geq f(x_n(t), t)$ for every $t \in T(x)$, we have

$$\liminf_{n \to +\infty} S(x_n) \geq \liminf_{n \to +\infty} f(x_n(t), t) = f(x(t), t) = S(x).$$

Hence $S(x)$ is continuous.

THEOREM 2.2. *If $f_x(x, t)$ is continuous on $R^n \times T$, then $S(x)$ is directionally differentiable in any direction $y \in X$, and its directional derivative is given by*

$$S'(x; y) = \max\{f_x(x(t), t)y(t); \ t \in T(x)\}. \tag{2.4}$$

PROOF. For any $t \in T(x)$ and $\varepsilon > 0$, it holds that

$$\frac{S(x + \varepsilon y) - S(x)}{\varepsilon} \geq \frac{f(x(t) + \varepsilon y(t), t) - f(x(t), t)}{\varepsilon} = f_x(x(t) + \rho y(t), t)y(t)$$

for some $0 < \rho < \varepsilon$. Tending ε to $+0$, we have

$$\liminf_{\varepsilon \to +0} \frac{S(x + \varepsilon y) - S(x)}{\varepsilon} \geq f_x(x(t), t)y(t). \tag{2.5}$$

On the other hand, there exists a sequence $\varepsilon_n \to +0$ such that

$$\limsup_{\varepsilon \to +0} \frac{S(x + \varepsilon y) - S(x)}{\varepsilon} = \lim_{n \to \infty} \frac{S(x + \varepsilon_n y) - S(x)}{\varepsilon_n}. \tag{2.6}$$

For each ε_n, there exists $t_n \in T$ such that

$$S(x + \varepsilon_n y) = f(x(t_n) + \varepsilon_n y(t_n), t_n). \tag{2.7}$$

By choosing a subsequence, we may assume that t_n converges to some $\bar{t} \in T$. Tending n to $+\infty$ in (2.7), we get from the continuity of $S(x)$ that

$$S(x) = f(x(\bar{t}), \bar{t}).$$

Thus $\bar{t} \in T(x)$. Furthermore, it holds for some $0 < \rho_n < \varepsilon_n$ that

$$
\begin{aligned}
\frac{S(x + \varepsilon_n y) - S(x)}{\varepsilon_n} &\leq \frac{f(x(t_n) + \varepsilon_n y(t_n), t_n) - f(x(t_n), t_n)}{\varepsilon_n} \\
&= f_x(x(t_n) + \rho_n y(t_n), t_n) y(t_n).
\end{aligned}
\tag{2.8}
$$

Since f_x is continuous, the right hand-side of (2.8) converges to $f_x(x(\bar{t}), \bar{t}) y(\bar{t})$. Hence we have

$$\limsup_{\varepsilon \to +0} \frac{S(x + \varepsilon y) - S(x)}{\varepsilon} \leq f_x(x(\bar{t}), \bar{t}) y(\bar{t}). \tag{2.9}$$

Combining (2.5) and (2.9), we get the conclusion.

As a special case, let us consider the one-sided state constraint:

$$x(t) \geq a(t) \quad \text{for every } t \in T \tag{2.10}$$

where $a \in C(T)$ is given. In order to apply Theorem 2.2, we take $f(x, t) := a(t) - x$. Then, since $f_x = -1$, we readily get the following result:

COROLLARY 2.3. *The directional derivative of $S(x)$ induced from the one-sided state constraint (1.3) is expressed as:*

$$S'(x; y) = -\min\{y(t) \ ; \ t \in T(x)\}$$

for any $y \in C(T)$.

Taking constant functions as $x(t)$ and $y(t)$ in Theorem 2.2, we get Danskin's formula (1967):

COROLLARY 2.4. *If $f_x(x, t)$ is continuous on $R^n \times T$, then $S(x)$ is directionally differentiable in any direction $y \in R^n$ and its directional derivative is given by*

$$S'(x; y) = \max\{f_x(x, t) y; \ t \in T(x)\}. \tag{2.11}$$

3. Second-order directional derivative

In this section, we consider a second-order directional derivative of $S(x)$. The upper second-order directional derivative of $S(x)$ at x in the direction y is defined by

$$\overline{S}''(x; y) = \limsup_{\varepsilon \to +0} \frac{S(x + \varepsilon y) - S(x) - \varepsilon S'(x; y)}{\varepsilon^2} \tag{3.1}$$

In order to describe the second-order directional derivative, we need the following sets $K(u,v)$, $K(y)$, and a function $E(t)$ that were introduced in Kawasaki (1988b). For any u, $v \in C(T)$, we define

$$K(u,v) := \{w \in C(T) \; ; \; {}^{\exists}o(1) \; s.t. \; \theta^2 u + \theta v + w + o(1) \in C_+(T) \;\; {}^{\forall}\theta > 0\}, \qquad (3.2)$$

where $C_+(T) := \{z \in C(T) \; ; \; z(t) \geq 0 \;\; {}^{\forall}t \in T\}$ and $o(1)$ is an element of $C(T)$ such that $o(1) \to 0$ as $\theta \to +\infty$. It is evident from the definition that $w \in K(u,v)$ if and only if

$$w(t) \geq -\theta^2 u(t) - \theta v(t) - o(1) \quad \text{for all } t \in T. \qquad (3.3)$$

For any $u \in C_+(T)$ and $v \in C(T)$ that satisfies

$$v(t) \geq 0 \quad \text{if} \quad u(t) = 0, \qquad (3.4)$$

the set $K(u,v)$ is characterized by an upper semicontinuous function $E(t)$ that is defined by (3.5) below, see Kawasaki (1988b):

$$E(t) := \begin{cases} \max\left\{\text{limsup} \frac{v(t_n)^2}{4u(t_n)}; \; \{t_n\} \text{ satisfies } (3.6)\right\}, & \text{if } t \in T_0, \\ 0 & \text{if } t \in T_1 \backslash T_0, \\ -\infty & \text{otherwise,} \end{cases} \qquad (3.5)$$

where

$$T_0 := \left\{t \in T \mid {}^{\exists}t_n \to t \; s.t. \; u(t_n) > 0, \; -\frac{v(t_n)}{u(t_n)} \to +\infty\right\}, \qquad (3.6)$$

$$T_1 := \{t \in T \mid u(t) = v(t) = 0\}. \qquad (3.7)$$

That is

$$w \in K(u,v) \quad \text{if and only if} \quad w(t) \geq E(t) \quad \text{for every } t. \qquad (3.8)$$

THEOREM 3.1. *Let $x(t)$ and $y(t)$ be arbitrary functions in $C(T)^n$. Assume that f_x and f_{xx} are continuous on $R^n \times T$. Then it holds that*

$$\overline{S}''(x;y) = \max\left\{\frac{1}{2}y(t)^T f_{xx}(x(t),t)y(t) + E(t) \; ; \; t \in T(x;y)\right\}, \qquad (3.9)$$

where $T(x;y) := \{t \in T \; ; \; S'(x;y) = f_x(x(t),t)y(t)\}$ and $E(t)$ is defined via (3.5) by taking

$$u(t) = S(x) - f(x(t),t), \quad v(t) = S'(x;y) - f_x(x(t),t)y(t). \qquad (3.10)$$

PROOF. For each t and for any $\varepsilon > 0$, there exists $0 < \rho_t < 1$ such that

$$f(x(t) + \varepsilon y(t), t)$$
$$= \; f(x(t),t) + \varepsilon f_x(x(t),t)y(t) + \frac{\varepsilon^2}{2}y(t)^T f_{xx}(x(t) + \rho_t \varepsilon y(t), t)y(t)$$
$$= \; f(x(t),t) + \varepsilon f_x(x(t),t)y(t) + \frac{\varepsilon^2}{2}y(t)^T f_{xx}(x(t),t)y(t)$$
$$+ \frac{\varepsilon^2}{2}y(t)^T \{f_{xx}(x(t) + \rho_t \varepsilon y(t), t) - f_{xx}(x(t),t)\}y(t)$$
$$= \; f(x(t),t) + \varepsilon f_x(x(t),t)y(t) + \frac{\varepsilon^2}{2}y(t)^T f_{xx}(x(t),t)y(t) + o(\varepsilon^2). \qquad (3.11)$$

Here we note that $o(\varepsilon^2)$ is uniform w.r.t. t. Thus we have

$$-\frac{u(t)}{\varepsilon^2} - \frac{v(t)}{\varepsilon}$$

$$= \frac{1}{\varepsilon^2}\{f(x(t),t) + \varepsilon f_x(x(t),t)y(t) - S(x) - \varepsilon S'(x;y)\}$$

$$= -\frac{1}{2}y(t)^T f_{xx}(x(t),t)y(t) + \frac{1}{\varepsilon^2}\{f(x(t)+\varepsilon y(t),t) - S(x) - \varepsilon S'(x;y)\}, \quad (3.12)$$

where the last equality follows from (3.11). Now, put $w_r(t) := r - \frac{1}{2}y(t)^T f_{xx}(x(t),t)y(t)$ for any $r > 0$. Then, regarding $\frac{1}{\varepsilon}$ as θ in (3.3), we see from (3.12) that w_r belongs to $K(u,v)$ if and only if

$$r \geq \frac{1}{\varepsilon^2}\{f(x(t)+\varepsilon y(t),t) - S(x) - \varepsilon S'(x;y)\} + o(1) \quad \text{for every } t, \quad (3.13)$$

which is equivalent to

$$r \geq \frac{1}{\varepsilon^2}\{S(x+\varepsilon y) - S(x) - \varepsilon S'(x;y)\} + o(1), \quad (3.14)$$

which is, furthermore, equivalent to

$$r \geq \overline{S}''(x;y). \quad (3.15)$$

On the other hand, we see from (3.8) that w_r belongs to $K(u,v)$ if and only if

$$r \geq \frac{1}{2}y(t)^T f_{xx}(x(t),t)y(t) + E(t) \quad \text{for every } t \in T. \quad (3.16)$$

Combining (3.15) and (3.16), we get the desired result.

Let us consider the one-sided state constraint (1.3) again. Then, since $f(x,t) = a(t) - x$, we have $f_{xx} = 0$. Hence we readily get the following result:

COROLLARY 3.2. *The upper second-order directional derivative of $S(x)$ induced from the one-sided state constraint (1.3) is expressed as:*

$$\overline{S}''(x;y) = \max\{E(t) \; ; \; t \in T(x;y)\}, \quad (3.17)$$

where $T(x;y) = \{t \in T \; ; \; S'(x;y) = -y(t)\}$ and $E(t)$ is defined via (3.5) by taking

$$u(t) = S(x) + x(t) - a(t), \quad v(t) = S'(x;y) + y(t). \quad (3.18)$$

Taking constant functions as $x(t)$ and $y(t)$ in Theorem 3.1, we get the following formula due to Kawasaki (1988b):

COROLLARY 3.3. *Let x and y be arbitrary points in R^n. Then it holds that*

$$\overline{S}_0''(x;y) = \max\left\{\frac{1}{2}y^T f_{xx}(x,t)y + E(t) \; ; \; t \in T(x;y)\right\}, \quad (3.19)$$

where $E(t)$ is defined via (3.5) by taking

$$u(t) = S_0(x) - f(x,t), \quad v(t) = S_0'(x;y) - f_x(x,t)y. \quad (3.20)$$

4.   Envelope

In formula (3.19), the function $E(t)$ fills the gap between the second derivatives of $f(x,t)$ and that of the envelope $S_0(x)$. A typical example was given in (1988b), see Example 4.1 below. In this section, we consider an envelope formed by inequality constraints.

EXAMPLE 4.1. Let us consider a family of straight lines $f(x,t) = 2tx - t^2$, where $t \in [0,1]$ and $x \in R$. It is evident that it forms an envelope $S_0(x) = x^2$, $(0 \le x \le 1)$, see Fig. 4.1. Hence $S_0''(0;y) = y^2$ for any $y \ge 0$ and $f_{xx}(0,t) \equiv 0$. This fact indicates that there is a gap between the second-order directional derivatives of the max-type function $S_0(x)$ and those of the constituent functions $f(x,t)$. On the other hand, it is directly computed from the definition that $E(0) = y^2$ for every $y > 0$, which fills the gap.

$$\overline{S}_0''(0;y) = \max\left\{\frac{1}{2}y^T f_{xx}(0,t)y + E(t) \; ; \; t \in T(0;y)\right\}$$
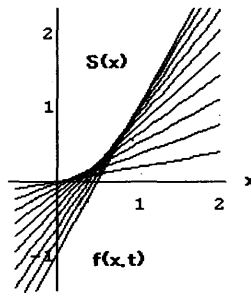$$= E(0) = y^2$$

Fig. 4.1

As was noted in Section 1, the max-type function $S(x)$ is an extension of $S_0(x)$. Hence inequality state constraints sometimes form envelopes, also. Concerning the one-sided state constraint $x(t) \ge a(t)$, we can prove a stronger result. That is, it forms an envelope not sometimes but always except two trivial cases (4.1) and (4.2) below.

THEOREM 4.2. *Let T be a connected compact metric space. Assume that $\bar{x}(t)$ does not satisfy neither*

$$x(t) \equiv a(t), \tag{4.1}$$

nor

$$x(t) > a(t) \quad \text{for every } t. \tag{4.2}$$

Then there exists a function $y \in C(T)$ such that the one-sided state constraint forms an envelope at $\bar{x}$ in the direction y.

PROOF. Let $y(t) := -2\sqrt{x(t) - a(t)}$ and put for $\xi \in R$

$$
\begin{aligned}
s(\xi) \;\; &:= \;\; S(\bar{x} + \xi y) = \max_t \{a(t) - \bar{x}(t) - \xi y(t)\} \\
&= \;\; \max_t \{a(t) - \bar{x}(t) + 2\xi\sqrt{\bar{x}(t) - a(t)}\}
\end{aligned}
$$

Then $s(\xi)$ is a standard max-function:

$$
s(\xi) = \max_{\tau \in T'} \{2\xi\tau - \tau^2\}
$$

Furthermore, by the assumption, the image of T by the continuous function $\sqrt{x(t) - a(t)}$ is a compact interval $T' := [0, t_1]$ with $t_1 > 0$. Hence $s(\xi)$ is same with the function in Example 4.1, so that an envelope is surely formed.

We close this paper with considering a special case where T is a finite set $\{1, \ldots, m\}$. In this case, $C(T) = R^m$ and the one-sided state constraint reduces to

$$
x(j) \geq a(j) \;\; \text{for every } j = 1, \ldots, m,
$$

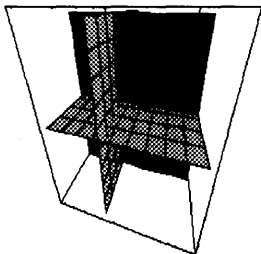which means m cuttings like Fig. 4.2 below $(m = 3)$.

Fig. 4.2

Hence no envelope is formed by them. Similarly, when T is infinite set, for instance, $T = [0, 1]$, the one-sided state constraint can be regarded an infinite number of cuttings in $C[0, 1]$. So the author thought that no envelope was formed by them, first. But that was not true. As we have proved in Theorem 4.1, they always form an envelope.

References

Clarke, F.H. (1975), "Generalized gradients and applications", *Trans. Amer. Math. Society.*, vol. 205, 247–262.

Correa, R and Seeger, A. (1985), "Directional derivative of a minimax function" *Nonlinear Analysis, Theory and Appl.*, vol. 9, 13–22.

Danskin, J.M. (1967), *The Theory of Max-Min and its Applications to Weapons Allocations Problems.* Springer, New York.

Dem'yanov, V.F. and Malozemov, V.N. (1974), *Introduction to Minimax.* John Wiley and Sons, New York.

Demyanov, V.F. and Zabrodin, I.S. (1986), "Directional differentiability of a continual maximum function of quasidifferentiable functions", Math. Program. Study, vol. 29, 108–117.

Hettich, R.P. and Jongen, H.Th. (1978), *Semi-infinite programming: conditions of optimality and applications* in *J. Stoer (ed.) Optimization Techniques 2*, Springer.

Ioffe, A. (1989), *On some recent developments in the theory of second order optimality conditions* in *S. Dolezki (ed.) Optimization*, Lecture Notes in Math., Vol. 1405, 55–68, Springer, New York.

Ioffe, A. (1991), "Variational analysis of a composite function: A formula for the lower second order epi-derivative," *J. Math. Anal. Appl.,* vol. 160, 379–405.

Kawasaki, H. (1988a), "An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems" *Mathematical Programming,* vol. 41, 73–96.

Kawasaki, H. (1988b), "The upper and lower second order directional derivatives of a sup-type function" *Mathematical Programming,* vol. 41, 327–339.

Kawasaki, H. (1991), "Second order necessary optimality conditions for minimizing a sup-type function" *Mathematical Programming,* vol. 49, 213–229.

Kawasaki, H. (1992), "Second-order necessary and sufficient optimality conditions for minimizing a sup-type function" *Applied Mathematics and Optimization,* vol. 26, 195–220.

Kawasaki, H. (1994), "A second-order property of spline functions with one free knot", *J. Approx. Theory,* 78, 293–297.

Kawasaki, H. (1995), "A first-order envelope-like effect of nonsmooth functions with an application to best approximation by polygonal curves with free knots", in Proceedings of APORS'94 (eds. M. Fushimi and K. Tone), World Scientific, New Jersey, 490-496.

Kawasaki, H. and Koga, S. (1995), "Legendre condition for a variational problem with one-sided phase constraints", J. Oper. Res. Soc. Japan, vol. 38, 483–492.

Koga, S. and Kawasaki, H. (1995), "Legendre conditions for variational problems with inequality phase constraints", in Proceedings of APORS'94 (eds. M. Fushimi and K. Tone), World Scientific, New Jersey, 484–489.

Seeger, A. (1988), "Second order directional derivatives in parametric optimization problems", Math. Oper. Res. vol. 13, 124–139.

Shiraishi, S. (1995), "Directional differentiability of max-functions and its applications to convex functions", in Proceedings of APORS'94, (eds. M. Fushimi and K Tone), World Scientific, New Jersey, 477–483.

Wetterling, W. (1970), Definitheits bedingungen für relative Extrema bei Optimierungs- und Approximationsaufgaben *Numer. Math.,* 15, 122–136.