Varational analysis approach for composite optimality problem

Wen Song
School of Mathematical and Sciences, Harbin Normal University,
Harbin, 150025, P.R. China

Contents

1 Classical optimality conditions

Let $f\colon S \subset \mathbb{R}^n \to \mathbb{R}$, where $S \subset \mathbb{R}^n$ is a subset. Consider the following minimization problem:

$$(P) \qquad \min_{x\in S} f(x).$$

A point $\bar{x} \in S$ is called a local minimum of this problem if there exists $\delta > 0$ such that

$$f(x) \geq f(\bar{x}) \quad \text{for all } x \in S \ \text{ such that } \|x - \bar{x}\| \leq \delta.$$

The point $\bar{x}$ is called the global minimum if $f(x) \geq f(\bar{x})$ for all $x \in S$.

We say that f satisfying the γ-order growth condition at $\bar{x}$, if there exists some $\kappa, \delta > 0$ such that

$$f(x) - f(\bar{x}) \geq \kappa\|x - \bar{x}|^{\gamma} \quad \text{for all } \|x - \bar{x}\| \leq \delta.$$

In particular, if $\gamma = 1, 2$ we say f satisfying the first (second) order growth condition at $\bar{x}$.

When the space is $\mathbb{R}^1$, the problem becomes the classical extreme problem, and the first necessarily condition is given by Fermat 1638 and Newton 1670.

Lemma 1.1 (Fermat 1638; Newton 1670)**.** *Let $\varphi\colon O \subset \mathbb{R} \to \mathbb{R}$ be a differentiable function defined on an open set O. If φ attains it local minimum(or maximum) at $\bar{x}$, then $\varphi^{'}(\bar{x}) = 0$.*

If f is differentiable at a point $x \in S$, then the gradient and Hessian matrix are denoted by

$$\nabla f(x) = (\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}),$$

$$\nabla^2 f(x) = [\frac{\partial^2 f(x)}{\partial x_i \partial x_j}], \quad i, j = 1, \ldots, n.$$

Theorem 1.1. *(Euler 1755)(The first order necessarily condition) Let f be differentiable at $\bar{x} \in \text{int}S$. If f attains it local minimum(or maximum) at $\bar{x}$, then*

$$\nabla f(\bar{x}) = 0. \tag{1.1}$$

If we assume that f is convex, but not necessarily differentiable, then the necessarily condition of a local minimizer becomes $0 \in \partial f(\bar{x})$.

When $\bar{x}$ is not an interior of S, in order to characterize the necessary property of local minimizer, we need the concepts of the tangent cone of S to $\bar{x}$.

Definition 1.1. *Let $S \subset \mathbb{R}^n$ be a nonempty set and $\bar{x} \in S$. The contingent (Bouligand) cone of S to $\bar{x}$ is defined as*

$$\begin{aligned} T_S(\bar{x}) : &= \{d \in X : \ \exists\ t_k \downarrow 0, \exists S \ni x_k \to \bar{x}, \frac{x_k - \bar{x}}{t_k} \to d\} \\ &= \{d \in X : \ \exists\ t_k \downarrow 0, d_k \to d \ \text{ s.t. } \ \bar{x} + t_k d_k \in S\}. \end{aligned}$$

Theorem 1.2. *Assume that $\bar{x}$ is a local minimum of problem (P) and that f is differentiable at $\bar{x}$. Then*

$$\nabla f(\bar{x})^T d \geq 0, \quad \text{for all} \quad d \in T_S(\bar{x}) \Leftrightarrow -\nabla f(\bar{x}) \in [T_S(\bar{x})]^0$$

Usually, the negative polar of $T_S(\bar{x})$ is called Fréchet normal cone of S to $\bar{x}$, denoted by $\hat{N}_S(\bar{x})$. Eqivalently,

$$\hat{N}_S(\bar{x}) = \{x^* \in X^* | \limsup_{x' \xrightarrow{S} \bar{x}} \frac{\langle x^*, x' - \bar{x}\rangle}{\|x' - \bar{x}\|} \leq 0\}.$$

When S is convex,

$$T_S(\bar{x}) = \overline{\text{cone}}(S - \bar{x})$$

and

$$\hat{N}_S(\bar{x}) = N_S(\bar{x}) = \{\bar{x} \in \mathbb{R}^n \colon \langle \bar{x}, x - \bar{x}\rangle \leq 0, \ \ \forall x \in S\}.$$

When S is defined by some constraint functions, for instance,

$$S := \{x \in \mathbb{R}^n \colon g_i(x) \leq 0, i = 1, 2, \ldots, p; h_j(x) = 0, j = 1, 2, \ldots, q\}, \tag{1.2}$$

where g_i, h_j are differentiable functions defined on $\mathbb{R}^n$. or more general form

$$S := \{x \in \mathbb{R}^n \colon F(x) \in K\}, \tag{1.3}$$

where $F \colon \mathbb{R}^n \to \mathbb{R}^m$ is a differentiable mapping, $K \subset \mathbb{R}^m$.

Question: how to characterize the tangent and the normal cone of S to $\bar{x} \in S$?

It is not hard to know

$$T_S(\bar{x}) \subset \{d \in \mathbb{R}^n \colon \nabla g_i(\bar{x})^T d \leq 0, i \in I(\bar{x}); \nabla h_j(\bar{x})^T d = 0, j = 1, 2, \ldots, q\} =: L_S(\bar{x})$$

or

$$T_S(\bar{x}) \subset \{d \in \mathbb{R}^n \colon DG(\bar{x})^T d \in T_K(\bar{x})\} =: L_S(\bar{x}).$$

Under what condition, the above inclusions become as equality?

$$T_S(\bar{x}) = L_S(\bar{x})$$

Such condition is called constraint qualification, for example, Mangasarian-Fromovitz, Robinson condition, metric regularity condition and so on.

Mangasarian-Fromovitz constraint qualification:

$$\begin{array}{c} \nabla h_j(\bar{x}), \ j = 1, \ldots, q, \ \text{are linearly independent}, \\ \exists d \in X : \nabla g_i(\bar{x}) d = 0, \ i = 1, \ldots q, \nabla g_i(\bar{x}) d < 0, \forall i \in I(\bar{x}), \end{array} \tag{1.4}$$

where $I(\bar{x})$ denotes the index set of active at $\bar{x}$ inequality constraints.

Robinson constraint qualification:

$$0 \in \text{int}\{F(\bar{x}) + DF(\bar{x})(\mathbb{R}^n) - K\}. \tag{1.5}$$

Furthermore, In the case of the equality holds, how to compute the normal cone $[T_S(\bar{x})]^o$? This is related to Farkas lemma.

From the normal analysis point of view, we easily show that

$$\hat{N}_S(\bar{x}) \supset \{DF(\bar{x})^* y^* | y^* \in \hat{N}_K(F(\bar{x})\},$$

and

$$N_S(\bar{x}) \subset \{DF(\bar{x})^* y^* | y^* \in N_K(F(\bar{x})\},$$

whenever the Robinson's condition holds, where $N_S(\bar{x})$ denote the (Morduchovich) limit normal cone to S at $\bar{x}$.

On the other hand, the problem (P) can be equivalent to write as

$$(P') \qquad \min_{x \in X} f(x) + I_S(x).$$

The necessarily condition of $\bar{x}$ being a local minimizer of (P') is $0 \in \hat{\partial}(f + I_S)(\bar{x})$. When f is a continuous convex function and S is a convex set, we have

$$\partial(f + I_S)(\bar{x}) = \partial f(\bar{x}) + \partial I_S(\bar{x}) = \partial f(\bar{x}) + N_S(\bar{x}).$$

How to calculate the subdifferential of the sum of two functions?

In optimization, second derivatives help significantly in the understanding of optimality, especially the formulation of sufficient conditions for local optimality in the absence of convexity. Such conditions form the basis for numerical methodology and assist in studies of what happens to optimal solutions when the parameters on which a problem depends are perturbed.

Theorem 1.3. *(The second optimality condition) Suppose that f is twice differentiable at $\bar{x} \in \mathrm{int} S$. If f attains it local minimum at $\bar{x}$, then*

$$\nabla f(\bar{x}) = 0, \quad \langle d, \nabla f(\bar{x}) d \rangle \geq 0, \quad \forall d \in \mathbb{R}^n.$$

Consider the second order tangent vector to a set S at $\bar{x} \in S$.

$$\begin{aligned} T_S^2(x, d) &= \{w \in X : \exists t_k \downarrow 0 \quad \text{such that} \quad \mathrm{d}(x + t_k d + \frac{1}{2} t_k^2 w, S) = o(t_k^2)\} \\ &= \{d \in X : \ \exists\ t_k \downarrow 0, w_k \to w \ \text{ s.t. } \ \bar{x} + t_k d + \frac{1}{2} t_k w_k \in S\}. \end{aligned}$$

Theorem 1.4. *Assume that $\bar{x}$ is a local minimum of problem (P) and that f is twice continuously differentiable at $\bar{x}$. Then for every $d \in T_S(\bar{x})$ with $\nabla f(\bar{x})^T d = 0$, we have*

$$\nabla f(\bar{x}) w + \langle d, \nabla^2 f(\bar{x}) d \rangle \geq 0, \quad \textit{for all } w \in T_S^2(\bar{x}, d), \tag{1.6}$$

Question: How to calculate the second order tangent set and how to reformulate the inequality (1.6)?

When S is defined by some constraint functions, for instance, has the form of (1.2) or (1.3), we have the following result.

Lemma 1.2. *Assume that Mangasarian-Fromivicz condition is satisfied at $\bar{x} \in S$, i.e. Then for every $d \in T_S(\bar{x})$,*

$$\begin{aligned} T_S^2(\bar{x}, d) &= \{w \in \mathbb{R}^n \colon \langle \nabla g_i(\bar{x}), w\rangle \leq -\langle d, \nabla^2 g_i(\bar{x}) d\rangle, i \in I^{00}(\bar{x}, d), \\ &\quad \langle \nabla h_j(\bar{x}), w\rangle = -\langle d, \nabla^2 h_j(\bar{x}) d\rangle, j = 1, ..., q.\} \end{aligned} \tag{1.7}$$

with $I^{00}(\bar{x}, d) = \{i \in I_0(\bar{x}) \colon \langle \nabla g_i(\bar{x}), d\rangle = 0\}$.

Similarly, if Robinsons condition is satisfied at $\bar{x} \in S$ then

$$T_S^2(\bar{x}, d) = DF(\bar{x})^{-1}[T_K^2(F(\bar{x}), DF(\bar{x})d) - D^2F(\bar{x})(d, d)]. \tag{1.8}$$

It follows by Lemma 1.2 that the inequality (1.6) becomes

$$\begin{aligned} &\inf_w \nabla f(\bar{x}) w + \langle d, \nabla^2 f(\bar{x}) d\rangle \geq 0, \\ &\text{s.t.} \quad \langle \nabla g_i(\bar{x}), w\rangle \leq -\langle d, \nabla^2 g_i(\bar{x}) d\rangle, i \in I^{00}(\bar{x}, d), \\ &\qquad \langle \nabla h_j(\bar{x}), w\rangle = -\langle d, \nabla^2 h_j(\bar{x}) d\rangle, j = 1, ..., q. \end{aligned}$$

or

$$\begin{aligned} &\inf_w \nabla f(\bar{x}) w + \langle d, \nabla^2 f(\bar{x}) d\rangle \geq 0, \\ &\text{s.t.} \quad DF(\bar{x})w + DF(\bar{x})(d, d) \in T_K^2(F(\bar{x}), DF(\bar{x})d). \end{aligned}$$

By using duality of linear optimization problem, we can reformulate the second order necessary condition.

2 Preliminaries from variational analysis

A great amount of functions involved in optimization problems are not differentiable. Maximization and minimization are often useful in constructing new functions and mappings from given ones, but, in contrast to addition and composition, they commonly fail to preserve smoothness. For example, the maximization of a finite many affine functions is convex, but not differentiable. However, there are directional differentiable.

Let $f \colon \mathbb{R}^n \to \bar{\mathbb{R}}$ be a function,

$$\text{dom} f = \{x \in \mathbb{R}^n | f(x) < \infty\}, \;\; \text{epi} f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} | f(x) \leq r\}. \tag{2.9}$$

Assume that $f(\bar{x})$ is finite. The direction derivative of f at $\bar{x}$ in the direction w is defined as

$$f^{'}(\bar{x}, w) = \lim_{t \to 0} \frac{f(\bar{x} + tw) - f(\bar{x})}{t} \tag{2.10}$$

If f is convex, then $t \to \frac{f(\bar{x}+tw)-f(\bar{x})}{t}$ is in non-increasing as $t \to 0$, so the limit in (2.10) and

$$f^{'}(\bar{x}, w) = \inf_{t > 0} \frac{f(\bar{x} + tw) - f(\bar{x})}{t}.$$

The directional derivative $f^{'}(\bar{x}, w)$ in (2.10) dependents only in the direction of w but not others. Instead, we can consider

$$f^{'}(\bar{x}, w) = \lim_{\substack{t \downarrow 0 \\ w' \to w}} \frac{f(\bar{x} + tw) - f(\bar{x})}{t}. \tag{2.11}$$

If the limit exists in (2.11), we say that f is semidifferentaible(Hadamard directional differentiable) at $\bar{x}$ in the direction w. In this case $f^{'}(\bar{x}, w)$ is continuous and positively homogeneous in w.

Clearly, the existence of the limit in (2.11) if and only if

$$\liminf_{\substack{t \downarrow 0 \\ w' \to w}} \frac{f(\bar{x} + tw) - f(\bar{x})}{t} = \limsup_{\substack{t \downarrow 0 \\ w' \to w}} \frac{f(\bar{x} + tw) - f(\bar{x})}{t}.$$

When f is semidifferentiable at $\bar{x}$, its continuity there, forces it to be finite on a neighborhood of of $\bar{x}$. So, $\bar{x} \in \text{int}(\text{dom} f)$. Therefore, semidifferentiablity can't assist in the study of situation where, for instance $\bar{x}$ is a boundary point of dom$f$.

Different choices of a mode of convergence for the difference quotient functions will lead to different kind of derivatives. The situation can be avoid when the differentiablity introduced by an approach through epi-convergence instead of continuous convergence.

2.1 Epi-derivatives and epi-differentiability

Let us first recall the following notions of *upper* and *lower limits*, in the sense of Painlevé-Kuratowski, of a parameterized family A_t of subsets of $\mathbb{R}^n$, where t can be real valued or, more generally, an element of a metric space.

Definition 2.1. *The following sets are called the upper (outer) and lower (inner) limits of a parameterized family A_t, of subsets of $\mathbb{R}^n$,*

$$\operatorname*{Lim\,sup}_{t \to t_0} A_t := \{x \in X : \liminf_{t \to t_0} \mathrm{d}(x, A_t) = 0\}$$

$$\operatorname*{Lim\,inf}_{t \to t_0} A_t := \{x \in X : \limsup_{t \to t_0} \mathrm{d}(x, A_t) = 0\}$$

respectively.

It easily follows from the definition that the upper and lower limit sets are both closed. These sets can be also described in terms of sequences as follows.

$$\operatorname*{Lim\,sup}_{t \to t_0} A_t = \{x | \exists t_k \to t_0, \exists x_k \in A_{t_k}, x_k \to x\},$$
$$\operatorname*{Lim\,inf}_{t \to t_0} A_t = \{x | \forall t_k \to t_0, \exists x_k \in A_{t_k}, x_k \to x\}.$$

If the equality in the above holds, we say A_t has a limit at t_0.

Now let $\varphi_t \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a family of extended real valued functions. The lower and upper *epi-limits* of φ_t, as $t \to t_0$, are defined as

$$\mathrm{epi}(\text{e-}\liminf_{t\to t_0} \varphi_t(\cdot)) = \operatorname*{Lim\,sup}_{t\to t_0} \mathrm{epi}\varphi_t, \tag{2.12}$$

$$\mathrm{epi}(\text{e-}\limsup_{t\to t_0} \varphi_t(\cdot)) = \operatorname*{Lim\,inf}_{t\to t_0} \mathrm{epi}\varphi_t. \tag{2.13}$$

Note that since the lower and upper set-limits are closed sets, the lower and upper epi-limit functions have closed epigraphs and hence are lower semicontinuous.

We define the *lower* and *upper directional epiderivatives* of an extended real valued function $f : \mathbb{R}^n \to \mathbb{R}$, at a point $x \in \mathbb{R}^n$ such that $f(x)$ is finite, as follows

$$f_-^\downarrow(x, \cdot) := \text{e-}\liminf_{t\downarrow 0} \frac{f(x + t\cdot) - f(x)}{t}, \tag{2.14}$$

$$f_+^\downarrow(x, \cdot) := \text{e-}\limsup_{t\downarrow 0} \frac{f(x + t\cdot) - f(x)}{t}. \tag{2.15}$$

Equivalently, we can write these derivatives in the following equivalent forms

$$f_-^\downarrow(x, w) = \liminf_{\substack{t\downarrow 0\\ w'\to w}} \frac{f(x + tw') - f(x)}{t}, \tag{2.16}$$

$$f_+^\downarrow(x, w) = \sup_{\{t_k\}\in\Sigma_0} (\liminf_{\substack{k\to\infty\\ w'\to w}} \frac{f(x + t_k w') - f(x)}{t_k}), \tag{2.17}$$

where Σ_0 denotes the set of positive real sequences $\{t_k\}$ converging to zero.

Since epi-limit functions are lower semicontinuous, we have that $f_-^\downarrow(x, \cdot)$ and $f_+^\downarrow(x, \cdot)$ are l. s. c. positively homogeneous functions.

We also have that

$$f_-^\downarrow(x, w) \le f_+^\downarrow(x, w), \qquad f_-^\downarrow(x, w) \le f'_-(x, w), \qquad f_+^\downarrow(x, w) \le f'_+(x, w). \tag{2.18}$$

We say that f is directionally epidifferentiable at x, in a direction w, if $f_-^\downarrow(x, w) = f_+^\downarrow(x, w)$, and in that case we denote $f^\downarrow(x, w)$ the common value. Note that $f^\downarrow(x, w)$ can be different from $f'(x, w)$ even if f is convex.

If f is directionally differentiable, its second order directional derivative is defined as

$$f''(x; h, w) := \lim_{t\downarrow 0} \frac{f(x + th + \frac{1}{2}t^2 w) - f(x) - tf'(x, h)}{\frac{1}{2}t^2}, \tag{2.19}$$

provided the above limit exists. We can also define

$$f^{''}(x; h, w) := \lim_{\substack{t\downarrow 0\\ w'\to w}} \frac{f(x + th + \frac{1}{2}t^2 w') - f(x) - tf'(x, h)}{\frac{1}{2}t^2}. \tag{2.20}$$

Note that if f has the second-order Taylor expansion at x,

$$f(x + h) = f(x) + \nabla f(x)h + \frac{1}{2}\nabla^2 f(x)(h, h) + o(\|h\|^2), \tag{2.21}$$

then

$$f''(x; h, w) = \nabla f(x)w + \nabla^2 f(x)(h, h). \tag{2.22}$$

Definition 2.2. *Assuming that $f(x)$ and the respective directional epiderivatives $f_-^{\downarrow}(x,h)$ and $f_+^{\downarrow}(x,h)$ are finite, we call*

$$f_-^{\downarrow\downarrow}(x;h,\cdot) := e\text{-}\liminf_{t\downarrow 0} \frac{f(x+th+\frac{1}{2}t^2\cdot) - f(x) - tf_-^{\downarrow}(x,h)}{\frac{1}{2}t^2}, \tag{2.23}$$

$$f_+^{\downarrow\downarrow}(x;h,\cdot) := e\text{-}\limsup_{t\downarrow 0} \frac{f(x+th+\frac{1}{2}t^2\cdot) - f(x) - tf_+^{\downarrow}(x,h)}{\frac{1}{2}t^2} \tag{2.24}$$

the lower and upper second order epidervatives. The lower second order epidervatives can be characterized pointwisely (see Ben-Tal and Zowe, 1982)

$$f_-^{\downarrow\downarrow}(x;h,w) := \liminf_{\substack{t\downarrow 0 \\ w'\to w}} \frac{f(x+th+\frac{1}{2}t^2w') - f(x) - tf_-^{\downarrow}(x,h)}{\frac{1}{2}t^2}. \tag{2.25}$$

We say that f is second order parabolicly directionally epidifferentiable at x in a direction h, if $f_-^{\downarrow\downarrow}(x;h,\cdot) = f_+^{\downarrow\downarrow}(x;h,\cdot)$.

f is second order parabolicly directionally epidifferentiable at x in a direction h if and only if for every $w \in \mathbb{R}^n$ and for every $t_k \downarrow 0$, there exists sequence $w_k \to w$ such that

$$f_-^{\downarrow\downarrow}(x;h,w) = \lim_{k\to\infty} \frac{f(x+t_kh+\frac{1}{2}t_k^2w_k) - f(x) - t_kf_-^{\downarrow}(x,h)}{\frac{1}{2}t_k^2}. \tag{2.26}$$

Note again that if $f(\cdot)$ is Lipschitz continuous and directionally differentiable at x, then for all $h, w \in X$ we have $f_-^{\downarrow\downarrow}(x;h,w) = f_-^{''}(x;h,w)$ and $f_+^{\downarrow\downarrow}(x;h,w) = f_+^{''}(x;h,w)$.

We can also consider another kind of epi-derivatives. Denote by

$$\Delta_t^2 f(x)(h) = \frac{f(x+th) - f(x) - tf_-^{\downarrow}(x,h)}{\frac{1}{2}t^2},$$
$$\Delta_t^2 f(x,v)(h) = \frac{f(x+th) - f(x) - t\langle v, h\rangle}{\frac{1}{2}t^2}, \quad \text{for } v \in \mathbb{R}^n.$$

Definition 2.3. *The second subderivative of f at x is defined by*

$$\mathrm{d}^2 f(x)(\cdot) := e - \liminf_{t\downarrow 0} \Delta_t^2 f(x)(\cdot), \tag{2.27}$$

the second subderivative of f at x for v is defined by

$$\mathrm{d}^2 f(x,v)(\cdot) = e - \liminf_{t\downarrow 0} \Delta_t^2 f(x,v)(\cdot). \tag{2.28}$$

If the second order difference quotient function $h \to \Delta_t^2 f(x)(h)$ (resp. $h \to \Delta_t^2 f(x,v)(h)$), epi-converges to some function as $t \downarrow 0$, we say that f is twice epi-differentiable at x (resp. for v); it is properly twice epi-differentiable at x (for v) if $\mathrm{d}^2 f(x)(\cdot)$ (resp. $\mathrm{d}^2 f(x,v)(\cdot)$) is proper.

The second subderivative can be equivalently written as

$$\mathrm{d}^2 f(x)(h) = \liminf_{\substack{t\downarrow 0\\ h'\to h}} \frac{f(x+th') - f(x) - tf_-^{\downarrow}(x,h')}{\frac{1}{2}t^2}, \tag{2.29}$$

and

$$\mathrm{d}^2 f(x,v)(h) = \liminf_{\substack{t\downarrow 0\\ h'\to h}} \frac{f(x+th') - f(x) - t\langle v,h'\rangle}{\frac{1}{2}t^2}. \tag{2.30}$$

Clearly, f is twice epi-differentiable at x (for v) if and only if for every sequence $t_k \downarrow 0$ and every $h \in \mathbb{R}^n$, there exists a sequence $h_k \to h$ such that

$$\mathrm{d}^2 f(x)(h) = \lim_{k\to\infty} \frac{f(x+t_k h_k) - f(x) - t_k f_-^{\downarrow}(x,h_k)}{\frac{1}{2}t_k^2}, \tag{2.31}$$

and

$$\mathrm{d}^2 f(x,v)(h) = \lim_{k\to\infty} \frac{f(x+t_k h_k) - f(x) - t_k\langle v,h_k\rangle}{\frac{1}{2}t_k^2}. \tag{2.32}$$

If f is twice epi-differentiable at x relative to v, then the second-order epi-derivative function $\mathrm{d}^2 f(x,v)(\cdot)$ is lower semicontinuous and positively homogeneous of degree 2.

$f_-^{\downarrow}(x,\cdot)$ is also called the subderivative of f at x and write as $\mathrm{d}f(x)$.

2.2 Tangent and Normal cones

Definition 2.4. *For $S \subset \mathbb{R}^n$ and a point $\bar{x} \in S$, define the following set:*
the contingent(Bouligand) cone

$$T_S(\bar{x}) := \limsup_{t\downarrow 0} \frac{S-\bar{x}}{t} = \{h \in \mathbb{R}^n \colon \exists\, t_k \downarrow 0,\ \mathrm{d}(\bar{x}+t_k h, S) = o(t_k)\}, \tag{2.33}$$

the inner tangent cone

$$T_S^i(\bar{x}) := \liminf_{t\downarrow 0} \frac{S-\bar{x}}{t} = \{h \in \mathbb{R}^n \colon \mathrm{d}(\bar{x}+th, S) = o(t),\ t \geq 0\}, \tag{2.34}$$

Clarke tangent cone

$$T_S^c(\bar{x}) := \liminf_{\substack{x' \xrightarrow{S} \bar{x}\\ t\downarrow 0}} \frac{S-x'}{t} \tag{2.35}$$

It is clear that if $\bar{x} \in S$, then $0 \in T_S^c(\bar{x}) \subset T_S^i(\bar{x}) \subset T_S(\bar{x})$. In general, these cones can be different, and the Clarke tangent cone are convex, but the inner tangent cones and contingent cone can be nonconvex.

If $T_S(\bar{x}) = T_S^i(\bar{x})$, we say that S is geometrically derivable at $\bar{x}$. It can be equivalently characterized by there exist a constant $\epsilon > 0$ and an arc $\xi\colon [0,\epsilon] \to S$ such that $\xi(0) = \bar{x}$ and $\xi_+'(0) = w$. For convex sets, however, the contingent, inner and Clarke tangent cones are equal to each other, so must be geometrically derivable.

Definition 2.5.

$$T_S^{i,2}(\bar{x}, h) := \liminf_{t\downarrow 0} \frac{S - \bar{x} - th}{\frac{1}{2}t^2}, \tag{2.36}$$

$$T_S^{2}(\bar{x}, h) := \limsup_{t\downarrow 0} \frac{S - \bar{x} - th}{\frac{1}{2}t^2}. \tag{2.37}$$

are called the inner and outer second order tangent sets, respectively, to the set S at the point $\bar{x}$ and in the direction h.

Alternatively these tangent sets can be written in the form

$$T_S^{i,2}(\bar{x}, h) = \{w \in \mathbb{R}^n : \mathrm{d}(\bar{x} + th + \frac{1}{2}t^2 w, S) = o(t^2), t \geq 0\}, \tag{2.38}$$

$$T_S^{2}(\bar{x}, h) = \{w \in \mathbb{R}^n : \exists t_k \downarrow 0 \quad \text{such that} \quad \mathrm{d}(\bar{x} + t_k h + \frac{1}{2}t_k^2 w, S) = o(t_k^2)\}. \tag{2.39}$$

Cleraly, $T_S^{i,2}(\bar{x}, h) \subset T_S^{2}(\bar{x}, h)$.

If $T_S^{i,2}(\bar{x}, h) = T_S^{2}(\bar{x}, h)$ for all h, we say that S be parabolically derivable at $\bar{x}$ for h. Equivalently, if for each $w \in T_S^{2}(\bar{x}, h)$ there are $\epsilon > 0$ and an arc $\xi\colon [0, \epsilon] \to S$ such that $\xi(0) = \bar{x}$ and $\xi'_+(0) = h$, $\xi''_+(0) = w$.

It is well-known that if C is convex, then $T_S^{i,2}(\bar{x}, h)$ is convex. If in addition, C is parabolically derivable at $\bar{x}$ for h, then $T_C^2(x, h)$ is convex. For a convex set C the following inclusions hold

$$T_C^{i,2}(x, h) + T_{T_C(x)}(h) \subset T_C^{i,2}(x, h) \subset T_{T_C(x)}(h), \tag{2.40}$$

$$T_C^{2}(x, h) + T_{T_C(x)}(h) \subset T_C^{2}(x, h) \subset T_{T_C(x)}(h). \tag{2.41}$$

It follows that if $0 \in T_C^2(x, h)$, then $T_C^2(x, h) = T_{T_C(x)}(h)$. Moreover, if $0 \in T_C^{i,2}(x, h)$, i.e. $\mathrm{d}(x + th, C) = o(t^2)$, all three sets coincide, that is

$$T_C^{i,2}(x, h) = T_C^{2}(x, h) = T_{T_C(x)}(h).$$

If $T_S^{i,2}(x, h) \neq \emptyset$ (resp, $T_S^2(x, h) \neq \emptyset$) only if $h \in T_S^i(x, h)$ (resp. $h \in T_S(x)$).

Given the function $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$, $f(\bar{x})$ is finite. The regular subdifferential of f at $\bar{x} \in \mathrm{dom} f$ is defined by

$$\hat{\partial} f(\bar{x}) = \{v \in \mathbb{R}^n | \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - t\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq 0. \tag{2.42}$$

It is easy to observe that the regular subgradients admits the dual representation

$$\hat{\partial} f(\bar{x}) = \{v \in \mathbb{R}^n | \langle v, w\rangle \leq f_-^{\downarrow}(\bar{x}, w), \ \forall w \in \mathbb{R}^n\}. \tag{2.43}$$

The subdifferential of f at $\bar{x}$ is given by

$$\partial f(\bar{x}) = \{v \in \mathbb{R}^n | \exists x_k \xrightarrow{f} \bar{x}, v_k \to v \text{ with } v_k \in \hat{\partial} f(x_k), \tag{2.44}$$

where $x_k \xrightarrow{f} \bar{x}$ stands for $x_k \to \bar{x}$ and $f(x_k) \to f(\bar{x})$.

The proximal subdiffferntial of f at $\bar{x}$ is given by

$$\partial^p f(\bar{x}) = \{v \in \mathbb{R}^n | \exists \sigma > 0, \delta > 0 \text{ s. t. } f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle - \frac{\sigma}{2}\|x - \bar{x}\|^2, \forall \|x - \bar{x}| < \delta\} \tag{2.45}$$

It is well-known that the inclusions $\partial^p f(\bar{x}) \subset \hat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$ always hold and

$$\partial f(\bar{x}) = \limsup_{x \xrightarrow{f} \bar{x}} \partial^p f(x).$$

Given a nonempty set $S \subset \mathbb{R}^n$, the proximal and regular normal cones to S at $\bar{x} \in S$ are defined, respectively, by

$$N_S^p(\bar{x}) = \partial^p \delta S(\bar{x}), \tag{2.46}$$

$$\hat{N}_S(\bar{x}) = \hat{\partial} \delta S(\bar{x}). \tag{2.47}$$

Similarly, we define the (limiting/Mordukhovich) normal cone of S at $\bar{x}$ by $N_S(\bar{x}) := \partial \delta_S(\bar{x})$.

2.3 Preliminary properties

Proposition 2.1. *Let $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ be an extended real valued function and let $x \in X$ be a point such that $f(x)$is finite. Then*

$$T_{\mathrm{epi}f}(x, f(x)) = \mathrm{epi} f_-^{\downarrow}(x, \cdot) \tag{2.48}$$

$$T^i_{\mathrm{epi}f}(x, f(x)) = \mathrm{epi} f_+^{\downarrow}(x, \cdot). \tag{2.49}$$

Proof. By the definition, we have that the epigraph of $f_-^{\downarrow}(x, \cdot)$ coincides with the upper set limit of epi$\{\frac{f(x+t\cdot)-f(x)}{t}\} = \frac{\mathrm{epi}f - (\bar{x}, f(\bar{x}))}{t}$, as $t \to 0$. Together with the definition of the contingent cones, this implies the first equation. The second equation can be proved similarly. □

Similarly, we can prove the following results.

Proposition 2.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be an extended real valued function taking a finite value at a point $x \in X$. Then*

$$T^{i,2}_{\mathrm{epi}f}[(x, f(x)), (h, f_+^{\downarrow}(x, h))] = \mathrm{epi} f_+^{\downarrow\downarrow}(x; h, \cdot), \tag{2.50}$$

$$T^2_{\mathrm{epi}f}[(x, f(x)), (h, f_-^{\downarrow}(x, h))] = \mathrm{epi} f_-^{\downarrow\downarrow}(x; h, \cdot), \tag{2.51}$$

provided the respective values $f_-^{\downarrow}(x, h)$ and $f_+^{\downarrow}(x, h)$ are finite ($h \in T_{\mathrm{dom}f}(\bar{x})$).

It is clear that epi$f$ is derivable at $(x, f(x)$ if and only if f is parabolically epi-differentiable at $\bar{x}$ in h and epi$f$ is parabolically derivable at $(x, f(x)$ for $(h, f_-^{\downarrow}(x, h))$ if and only if f is second order parabolically epi-differentiable at $\bar{x}$ in h.

We say a function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is called Lipschitz continuous around $\bar{x}$ relative to $C \subset \mathrm{dom}f$ with constant $\ell > 0$ if $\bar{x} \in C$ and there exists a neighborhood U of $\bar{x}$ such that

$$\|f(x_1) - f(x_2)| \leq \ell \|x_1 - x_2\| \;\; \forall x_1, x_2 \in U \cap C. \tag{2.52}$$

Such a function is called locally Lipschitz continuous relative to C if for every $\bar{x} \in C$, this function is Lipschitz continuous around $\bar{x}$ relative to C. Piecewise linear-quadratic functions (not necessarily convex) and an indicator function of a nonempty set are important examples of functions that are locally Lipschitz continuous relative to their domains.

Proposition 2.3. *Suppose that $f\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is Lipschitz continuous around $\bar{x}$ relative to its domain. Then*

(i) $\mathrm{dom} f_-^{\downarrow}(\bar{x}, \cdot) = T_{\mathrm{dom} f}(\bar{x})$ *and for every $h \in T_{\mathrm{dom} f}(\bar{x})$, $f_-^{\downarrow}(\bar{x}, h)$ is finite.*

(ii) If, in addition, f is parabolically epi-differentiable at $\bar{x}$ for h. then $\mathrm{dom} f$ *is parabolically derivable at $\bar{x}$ for h and*

$$T^2_{\mathrm{dom} f}(\bar{x}, h) = \mathrm{dom} f_-^{\downarrow\downarrow}(x, h, \cdot). \tag{2.53}$$

Example 2.1. *Let S be a non empty closed convex subset of $\mathbb{R}^n$, $\delta_S(\cdot)$ is a proper l.s.c. convex function. Consider the set consider the set $K := \mathrm{epi}\delta_S = S \times \mathbb{R}_+$ and $x \in S$. Then $\delta^{\downarrow}(x, \cdot) = \delta_{T_S(x)}(\cdot)$. Given a vector $h \in T_S(x)$. It is not difficult to see that*

$$\delta_+^{\downarrow\downarrow}(x; h, w) = \begin{cases} 0, & \text{if } w \in T_S^{i,2}(x, h), \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.54}$$

$$\delta_-^{\downarrow\downarrow}(x; h, w) = \begin{cases} 0, & \text{if } w \in T_S^{2}(x, h), \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.55}$$

Moreover, we have

$$T_K^{i,2}((x,0),(h,\gamma)) = \begin{cases} T_S^{i,2}(x,h) \times \mathbb{R}, & \text{if } \gamma > 0, \\ T_S^{i,2}(x,h) \times \mathbb{R}_+, & \text{if } \gamma = 0, \\ \emptyset, & \text{if } \gamma < 0, \end{cases} \tag{2.56}$$

and

$$T_K^{2}((x,0),(h,\gamma)) = \begin{cases} T_S^{2}(x,h) \times \mathbb{R}, & \text{if } \gamma > 0, \\ T_S^{2}(x,h) \times \mathbb{R}_+, & \text{if } \gamma = 0, \\ \emptyset, & \text{if } \gamma < 0, \end{cases} \tag{2.57}$$

Therefore the following conditions are equivalent:(i) the set $K := \mathrm{epi}\delta_S$ is parabolically derivable at $(x, 0)$, (ii) the set S is parabolically derivable at x, and (iii) the function δ_S is parabolically epidifferentiable at $\bar{x}$.

Proposition 2.4. *Let $\bar{x}$ be such that $f(\bar{x})$ is finite, for a given $h \in \mathbb{R}^n$, let $v \in \mathbb{R}^n$ be such that $\langle v, h\rangle = f_-^{\downarrow}(x, h)$. Then*

$$\inf_{w \in X} \{ f_-^{\downarrow\downarrow}(\bar{x}; h, w) - \langle v, w\rangle \} \geq \mathrm{d}^2 f(\bar{x}, v)(h). \tag{2.58}$$

Proof. Let $w_k \to w$, $t_k \downarrow 0$ be such that

$$f_-^{\downarrow\downarrow}(x; h, w) = \lim_{k\to\infty} \frac{f(x + t_k h + \frac{1}{2}t_k^2 w_k) - f(x) - t_k \mathrm{d}f(x)(h)}{\frac{1}{2}t_k^2} \tag{2.59}$$

and take $h_k^{'} := h + \frac{1}{2}t_k w_k$ in the definition of $\mathrm{d}^2 f(x, v)(h)$. We obtain then that

$$\mathrm{d}^2 f(x, v)(h) \leq \liminf_{k\to\infty} \frac{f(x + t_k h + \frac{1}{2}t_k^2 w_k) - f(x) - t_k \mathrm{d}f(x)(h) - \frac{1}{2}t_k^2\langle v, w_k\rangle}{\frac{1}{2}t_k^2}.$$

It follows that for any $w \in \mathbb{R}^n$,

$$\mathrm{d}^2 f(x,v)(h) \le f_-^{\downarrow\downarrow}(x;h,w) - \langle v, w\rangle.$$

By taking the infimum of the right hand side of the above inequality over all $w \in \mathbb{R}^n$, we obtain the desired result. □

Proposition 2.5. *(properties of second subderivative). Let $f : \mathbb{R}^n \to \mathbb{R}$ taking finite value at $\bar{x}$, $\bar{v} \in \mathbb{R}^n$. Then the following conditions hold:*

(i) *if $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function, then we always have*

$$\mathrm{dom}\mathrm{d}^2 f(\bar{x}, \bar{v}) \subset \{h \in \mathbb{R}^n | f_-^{\downarrow}(\bar{x}, h) = \langle \bar{v}, h\rangle\}, \tag{2.60}$$

Moreover, the equality holds if, in addition, $\mathrm{dom} f_-^{\downarrow\downarrow}(x;h,\cdot) \neq \emptyset$;

(ii) if $\bar{v} \in \hat{\partial} f(\bar{x})$, then for any $w \in \mathbb{R}^n$ we have $\mathrm{d}^2 f(\bar{x}, \bar{v})(h) \ge -\sigma\|h\|$. In particular, $\mathrm{d}^2 f(\bar{x}, \bar{v})$ is a proper function.

Proof. Note that

$$\begin{aligned} \mathrm{d}^2 f(\bar{x}, \bar{v})(w) &= \liminf_{\substack{t\downarrow 0 \\ w' \to w}} \frac{f(\bar{x} + tw') - f(\bar{x}) - t\langle \bar{v}, w'\rangle}{\frac{1}{2}t^2} \\ &= \liminf_{\substack{t\downarrow 0 \\ w' \to w}} \frac{\frac{f(\bar{x}+tw')-f(\bar{x})}{t} - \langle \bar{v}, w'\rangle}{\frac{1}{2}t} \end{aligned}$$

It is easily to see that (i) holds by the definition and (2.58). By the definition of proximal subdifferential (2.45), as $t \downarrow 0$, $w' \to w$, we have $\|\bar{x} + tw' - \bar{x}\| < \delta$, the assertion (ii) follows.

Proposition 2.2 implies that $f_-^{\downarrow\downarrow}(x;h,\cdot)$ and $f_+^{\downarrow\downarrow}(x;h,\cdot)$ are lower semicontinuous functions. If, in addition, f is parabolically derivable at $\bar{x}$ for h, then $f_-^{\downarrow\downarrow}(x;h,\cdot) = f_+^{\downarrow\downarrow}(x;h,\cdot)$ is lower semicontinuous and convex. It follows from Proposition 2.4 and Proposition 2.5 that $f_-^{\downarrow\downarrow}(x;h,\cdot)$ is also proper whenever $\mathrm{dom} f_-^{\downarrow\downarrow}(x;h,\cdot) \neq \emptyset$ and there exists $\bar{v} \in \hat{\partial} f(\bar{x})$ such $h \in K_f(\bar{x}, \bar{v})$.

3 Twice epi-Differetiability for composite function

3.1 Parabolic regularity

Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, $f(\bar{x})$ is finite and pick $\bar{v}\mathbb{R}^n$. The critical cone of f at $(\bar{x}, \bar{v})$ is defined by

$$K_f(\bar{x}, \bar{v}) := \{h \in \mathbb{R}^n | \mathrm{d}f(\bar{x})(h) = \langle v, h\rangle\}. \tag{3.61}$$

Definition 3.1. *It is said that the function f is parabolically regular at $\bar{x}$ for $\bar{v}$ in $h \in R^n$ if*

$$\inf_{w\in\mathbb{R}^n} \{f_-^{\downarrow\downarrow}(\bar{x};h,w) - \langle \bar{v}, w\rangle\} = \mathrm{d}^2 f(\bar{x}, \bar{v})(h). \tag{3.62}$$

A nonempty set $S \subset \mathbb{R}^n$ is said to be parabolically regular at $\bar{x}$ for $\bar{v}$ if the indicator function δ_S is parabolically regular at $\bar{x}$ for $\bar{v}$.

The parabolical regularity of f at $\bar{x}$ can be equivalent characterized as following.

Proposition 3.1. *f is parabolically regular at $\bar{x}$ for $\bar{v} \in \hat{\partial} f(\bar{x})$ if for every $h \in \operatorname{dom} d^2 f(\bar{x}, \bar{v})$, there exist, among the sequences $t_k \downarrow 0$ and $h_k \to h$ with $\Delta^2_{t_k} f(x, v)(h_k) \to d^2 f(\bar{x}, \bar{v})(h)$, ones with additional property that*

$$\limsup_{k\to\infty} \frac{\|h_k - h\|}{t_k} < \infty. \tag{3.63}$$

Moreover, for every $h \in \operatorname{dom} d^2 f(\bar{x}, \bar{v})$, there exists $w \in \mathbb{R}^n$ such that

$$d^2 f(\bar{x}, \bar{v})(h) = f_-^{\downarrow\downarrow}(\bar{x}; h, w) - \langle \bar{v}, w\rangle.$$

Proof Since $\bar{v} \in \hat{\partial} f(\bar{x})$, we know that $d^2 f(\bar{x}, \bar{v})(h) \geq -\sigma\|h\| > -\infty$. If $h \in K_f(\bar{x}, \bar{v}) \setminus \operatorname{dom} d^2 f(\bar{x}, \bar{v})$ then, by (2.58), we have that (3.62) holds. Therefore, we can assume that $d^2 f(\bar{x}, v)(h)$ is finite. Let $h_k \to h$ and $t_k \to 0$ be sequences at which the limit in the definition of $d^2 f(\bar{x}, v)(h)$ is attained. Consider $w_k := (\frac{1}{2}t_k)^{-1}(h_k - h)$, i.e., $h_k = h + \frac{1}{2}t_k w_k$, and $x_k = x + t_k h + \frac{1}{2}t_k^2 w_k$. Then

$$\begin{aligned} d^2 f(\bar{x}, \bar{v})(h) &= \lim_{k\to\infty} \frac{f(\bar{x} + t_k h_k) - f(\bar{x}) - t_k\langle \bar{v}, h_k\rangle}{\frac{1}{2}t_k^2} \\ &= \lim_{k\to\infty} \frac{f(\bar{x} + t_k h + \frac{1}{2}t_k^2 w_k) - f(\bar{x}) - t_k\langle \bar{v}, h\rangle}{\frac{1}{2}t_k^2} - \langle \bar{v}, w_k\rangle. \end{aligned}$$

Since $\{w_k\}$ is bounded, without loss of generality, we may assume that $w_k \to w$. Hence

$$d^2 f(\bar{x}, \bar{v})(h) \geq f_-^{\downarrow\downarrow}(\bar{x}; h, w) - \langle \bar{v}, w\rangle.$$

Combining (2.4), we get

$$d^2 f(\bar{x}, \bar{v})(h) = \inf_w \{f_-^{\downarrow\downarrow}(\bar{x}; h, w) - \langle \bar{v}, w\rangle\}.$$

The left side of (3.62) is identical to the lowest limit attainable for

$$\lim_{k\to\infty} \frac{f(\bar{x} + t_k h + \frac{1}{2}t_k^2 w_k) - f(\bar{x}) - t_k\langle \bar{v}, h\rangle}{\frac{1}{2}t_k^2} - \langle \bar{v}, w_k\rangle.$$

relative to $t_k \downarrow 0$ and a bounded sequence of vectors w_k (as seen from the cluster points w of such a sequence). In terms of $h_k = h + \frac{1}{2}t_k w_k$, which corresponds to $\{\frac{h_k - h}{t_k}\}$ is bounded. □

Definition 3.2. *We say that the set K is outer second order regular at a point $y \in K$ in direction $d \in T_K(y)$, if for any sequence $y_k \in K$ of the form $y_k = y + td + \frac{1}{2}t_k^2 w_k$, satisfying $t_k \downarrow 0$, $t_k w_k \to 0$, the following condition holds:*

$$\lim_{k\to\infty} d(w_k, T_K^2(y, d)) = 0. \tag{3.64}$$

We say that K is second order regular at y if K is parabolically derivable and outer second order regular at y in all direction $d \in T_K(y)$.

For example, the cone of positive semidefinite matrices, the second order cone are second order regular.

Definition 3.3. *Let* $f\colon X \to \bar{\mathbb{R}}$. *We say that* $f(\cdot)$ *is (outer) second order regular at* $\bar{x}$ *in the direction* h *if* $\mathrm{d}f(\bar{x})(h)$ *is finite and the set* $K = \mathrm{epi}f$ *is (outer) second order at the point* $(\bar{x}, f(\bar{x}))$ *in the direction* $(h, \mathrm{d}f(\bar{x})(h))$.

For example, the leading eigenvalue functions of symmetric matrix is the second order regular function.

Proposition 3.2. *Let* $f\colon X \to \bar{\mathbb{R}}$, $(\bar{x}, \bar{v}) \in \mathrm{gph}\partial f$, *and* $h \in K_f(\bar{x}, \bar{v})$ *with* $\mathrm{d}^2 f(x, v)(h) > -\infty$. *Then* f *is parabolically regular at* x *for* $\bar{v}$ *in the direction* h, *if the function* f *is outer second order regular at* $\bar{x}$ *in the direction* h.

Proof. If $\mathrm{d}^2 f(\bar{x}, v)(h) = +\infty$ then, by (2.58), we have that (3.62) holds. Therefore, we can assume that $\mathrm{d}^2 f(\bar{x}, v)(h)$ is finite. Note that because of $\langle v, h\rangle = \mathrm{d}f(\bar{x})(h)$, we have that $\mathrm{d}f(\bar{x})(h)$ is finite.

Suppose that f is outer second order regular at $\bar{x}$ in the direction h. Let $h_k \to h$ and $t_k \to 0$ be sequences at which the limit in the definition of $\mathrm{d}^2 f(\bar{x}, v)(h)$ is attained. Consider $w_k := (\frac{1}{2}t_k)^{-1}(h_k - h)$, i.e., $h_k = h + \frac{1}{2}t_k w_k$, and $x_k = x + t_k h + \frac{1}{2}t_k^2 w_k$. Then $t_k w_k \to 0$ and

$$\mathrm{d}^2 f(\bar{x}, v)(h) = \lim_{k\to\infty} \frac{f(x_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(h) - \frac{1}{2}t_k^2\langle v, w_k\rangle}{\frac{1}{2}t_k^2}. \tag{3.65}$$

Consider

$$c_k := \frac{f(x_k) - f(\bar{x}) - t_k \mathrm{d}f(\bar{x})(h)}{\frac{1}{2}t_k^2}.$$

Since $t_k w_k \to 0$ and $\mathrm{d}^2 f(x, v)(h)$ is finite, because of (3.65) we have that $t_k c_k \to 0$. Then by formula

$$T^2_{\mathrm{epi}f}[(\bar{x}, f(x)), (h, \mathrm{d}f(\bar{x})(h)] = \mathrm{epi}f_-^{\downarrow\downarrow}(\bar{x}; h, \cdot),$$

it follows from the outer second order regularity of f that

$$d((w_k, c_k), \mathrm{epi}f_-^{\downarrow\downarrow}(\bar{x}, h, \cdot)) \to 0.$$

It follows that there exists $(w_k', c_k') \in \mathrm{epi}f_-^{\downarrow\downarrow}(\bar{x}, h, \cdot)$ such that

$$(w_k, c_k) - (w_k', c_k') \to (0, 0).$$

This implies that

$$f(x_k) \geq f(\bar{x}) + t_k \mathrm{d}f(\bar{x})(h) + \frac{1}{2}t_k^2 f_-^{\downarrow\downarrow}(\bar{x}, h, w_k') + o(t_k^2).$$

It follows from (3.65) that

$$\mathrm{d}^2 f(\bar{x}, v)(h) \geq \liminf_{k\to\infty}\{f_-^{\downarrow\downarrow}(\bar{x}; h, w_k') - \langle v, w_k'\rangle + \langle v, w_k - w_k'\rangle\}.$$

Since $\langle v, w_k - w_k'\rangle \to 0$, it follows that $\mathrm{d}^2 f(\bar{x}, v)(h)$ is greater than or equal to the left hand side of (3.62), and hence the equality follows. $\square$

Note that if $v \in \hat{\partial} f(\bar{x})$, then the numerator of the radio inside the limit in (3.65) is nonnegative, and hence $\mathrm{d}^2 f(\bar{x}, v)(h) \geq 0$. Therefore, in this case, the assumption $\mathrm{d}^2 f(\bar{x}, v)(h) > -\infty$ in the above proposition is superfluous.

Theorem 3.1. *(twice epi-differenitability of parabolically regular functions) Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar v \in \hat\partial f(\bar x)$ and let f be parabolically epi-differentiable at $\bar x$ for every $h \in K_f(\bar x, \bar v)$. If f is parabolically regular at $\bar x$ for $\bar v$, then it is properly twice epi-differentiable at $\bar x$ for $\bar v$ with*

$$\mathrm{d}^2 f(\bar x, \bar v)(h) = \begin{cases} \min_{w\in\mathbb{R}^n}\{f_-^{\downarrow\downarrow}(\bar x; h, w) - \langle \bar v, w\rangle\} & \text{if } h \in K_f(\bar x, \bar v)\ , \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.66}$$

Proof. It follows from the parabolic epi-differentiability of f at $\bar x$ for every $h \in K_f(\bar x, \bar v)$ and Proposition 2.5 that $\mathrm{dom}\mathrm{d}^2 f(\bar x, \bar v) = K_f(\bar x, \bar v)$. This together with proposition 3.1 justifies the second subderivative formula (3.66). To establish the twice epi-differentiability of f at $\bar x$ for $\bar v$, we are going to show that (2.32) holds for all $h \in \mathbb{R}^n$. Pick $h \in K_f(\bar x, \bar v)$ and an arbitrary sequence $t_k \downarrow 0$. Since f is parabolically regular at $\bar x$ for $\bar v$, by Proposition 3.1, we find $w \in \mathbb{R}^n$ such that

$$\mathrm{d}^2 f(\bar x, \bar v)(h) = f_-^{\downarrow\downarrow}(\bar x; h, w) - \langle \bar v, w\rangle. \tag{3.67}$$

By the parabolic epi-differentiability of f at $\bar x$ for h, we find a sequence $w_k \to w$ for which we have

$$f_-^{\downarrow\downarrow}(\bar x; h, w) = \lim_{k\to\infty} \frac{f(x + t_k h + \frac{1}{2}t^2 w_k) - f(\bar x) - t_k \mathrm{d}f(\bar x)(h)}{\frac{1}{2}t_k^2}. \tag{3.68}$$

Define $h_k := h + \frac{1}{2}w_k$ for all k. Using this and $h \in K_f(\bar x, \bar v)$, we obtain

$$\begin{aligned} \Delta_{t_k}^2 f(\bar x, \bar v)(h_k) &= \lim_{k\to\infty} \frac{f(\bar x + t_k h_k) - f(\bar x) - t_k\langle \bar v, h_k\rangle}{\frac{1}{2}t_k^2} \\ &= \lim_{k\to\infty} \frac{f(\bar x + t_k h + \frac{1}{2}t_k^2 w_k) - f(\bar x) - t_k\langle \bar v, h\rangle}{\frac{1}{2}t_k^2} - \langle \bar v, w_k\rangle. \end{aligned}$$

This together with (3.67) and (3.68) results in

$$\lim_{k\to\infty} \Delta_{t_k}^2 f(\bar x, \bar v)(h_k) = f_-^{\downarrow\downarrow}(\bar x; h, w) - \langle \bar v, w\rangle = \mathrm{d}^2 f(\bar x, \bar v)(h)$$

which justifies (2.32) for every $h \in K_f(\bar x, \bar v)$. Finally, we are going to show the validity of (2.32) for every $h \notin K_f(\bar x, \bar v)$. For any such a h, we see that $\mathrm{d}^2 f(\bar x, \bar v)(h) = \infty$. Hence $\mathrm{d}^2 f(\bar x, \bar v)(h) = e - \limsup_{t\downarrow 0} \Delta_t^2 f(\bar x, \bar v)(h) = \infty$. This completes the proof of the Theorem. □

Proposition 3.3. *(conjugate of twice parabolic epiderivatives) Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper lower semicontinuous and convex function, and $\bar v \in \hat\partial f(\bar x)$, and let f be parabolically epi-differentiable at $\bar x$ for every $h \in K_f(\bar x, \bar v)$. If f is parabolically regular at $\bar x$ for $\bar v$, then $\phi(w) := f_-^{\downarrow\downarrow}(\bar x; h, w)$ is proper l.s.c. and convex functions and its conjugate function is given by*

$$\phi^*(v) = \begin{cases} -\mathrm{d}^2 f(\bar x, \bar v)(h) & \text{if } v \in \mathcal{A}(\bar x, h)\ , \\ +\infty, & \text{otherwise,} \end{cases} \tag{3.69}$$

where $\mathcal{A}(\bar x, h) = \{v \in \partial f(\bar x) | \mathrm{d}f(\bar x)(h) = \langle v, h\rangle\}$.

Proof. We know that ϕ is proper lower semicontinuous and convex function. Pick $v \in \mathcal{A}(\bar{x}, h)$, the formula (3.69) is clearly true duo to Theorem 3.1. Assume now that $v \notin \mathcal{A}(\bar{x}, h)$. This means either $\bar{v} \notin \partial f(\bar{x})$ or $\mathrm{d}f(\bar{x})(h) \neq \langle \bar{v}, h\rangle$. Define the parabolic difference quotients for f at $\bar{x}$ for h by

$$\Delta_{t,\bar{x},h} f(w) = \frac{f(\bar{x} + th + \frac{1}{2}t^2 w) - f(\bar{x}) - \mathrm{d}f(\bar{x})(h)}{\frac{1}{2}t^2}, \quad w \in \mathbb{R}^n, t > 0.$$

It is not hard to see that $\Delta_{t,\bar{x},h} f(w)$ are proper, convex, and

$$(\Delta_{t,\bar{x},h} f)^*(v) = \frac{f(\bar{x}) + f^*(v) - \langle v, \bar{x}\rangle}{\frac{1}{2}t^2} + \frac{\mathrm{d}f(\bar{x})(h) - \langle v, h\rangle}{\frac{1}{2}t}, \quad v \in \mathbb{R}^n$$

Remember that by the parabolic epi-differentiability of f at $\bar{x}$ for h amounts to the sets $\mathrm{epi}\Delta_{t,\bar{x},h} f(\cdot)$ converging to $\mathrm{epi}\phi$ as $t \downarrow 0$ and that the functions $\Delta_{t,\bar{x},h} f(\cdot)$ and ϕ are proper, l.s.c. and convex. Appealing to [14, Theorem 11.34] tells us that the former is equivalent to the sets $\mathrm{epi}(\Delta_{t,\bar{x},h} f)^*$ converging to $\mathrm{epi}\phi^*$ as $t \downarrow 0$ This, in particular, means that for any sequence $t_k \downarrow 0$, there exists a sequence $v_k \to \bar{v}$ such that

$$\phi^*(\bar{v}) = \lim_{k\to\infty} (\Delta_{t_k,\bar{x},h} f)^*(v_k).$$

If $\bar{v} \notin \partial f(\bar{x})$, then we have

$$f(\bar{x}) + f^*(v) - \langle v, \bar{x}\rangle > 0.$$

Since f^* is l.s.c., we get

$$\liminf_{k\to\infty} \frac{f(\bar{x}) + f^*(v_k) - \langle v_k, \bar{x}\rangle}{\frac{1}{2}t_k} + \frac{\mathrm{d}f(\bar{x})(h) - \langle v_k, h\rangle}{\frac{1}{2}} \geq \infty$$

which in turn confirms that

$$\phi^*(\bar{v}) = \lim_{k\to\infty} (\Delta_{t_k,\bar{x},h} f)^*(v_k) = \infty.$$

If $\bar{v} \in \partial f(\bar{x})$ but $\langle v, h\rangle < \mathrm{d}f(\bar{x})(h)$. Since we always have

$$f(\bar{x}) + f^*(v_k) - \langle v_k, \bar{x}\rangle \geq 0,$$

we arrive at

$$\phi^*(\bar{v}) \geq \lim \frac{\mathrm{d}f(\bar{x})(h) - \langle v_k, h\rangle}{\frac{1}{2}t_k} = \infty.$$

□

Example 3.1. *(piecewise linear-quadratic functions). Assume that the function $f\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ with is convex piecewise linear-quadratic. That is if $\mathrm{dom} f = \cup_{i=1}^p C_i$ with C_i being polyhedral convex sets for $i = 1, ..., p$, and if f has a representation of the form*

$$f(x) = \frac{1}{2}\langle A_i x, x_i\rangle + \langle a_i, x_i\rangle + \alpha_i \quad \textit{for all } x \in C_i, \tag{3.70}$$

Where A_i is an $n \times n$ symmetric matrix, $a_i \in \mathbb{R}^n$, and $\alpha_i \in \mathbb{R}$ for $i = 1, , p$. It was proven in [14, Propsoition 13.9] that the second subderivative of f at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ can be calculated by

$$\mathrm{d}^2 f(\bar{x}, \bar{v})(h) = \begin{cases} \langle A_i, h \rangle & \text{if } h \in T_{C_i}(\bar{x}) \cap \{\bar{v}_i\}^{\perp} \\ \infty & \text{otherwise,} \end{cases} \tag{3.71}$$

where $\bar{v}_i = \bar{v} - A_i\bar{x} - a_i$. To prove the parabolic regularity of f at $\bar{x}$ for $\bar{v}$, pick a vector $h \in \mathbb{R}^n$ with $\mathrm{d}^2 f(\bar{x}, \bar{v})(h) < \infty$. This implies that there is an i with $1 \leq i \leq p$ such that $h \in T_{C_i}(\bar{x}) \cap \{\bar{v}_i\}^{\perp}$ Since C_i is a polyhedral convex set, there exists an $\tau > 0$ such that $\bar{x} + th \in C_i$ for all $t \in [0, \tau]$. Pick a sequence $t_k \downarrow 0$ such that $t_k \in [0, \tau]$ and let $h_k := h$ for all $k = 1, 2, \ldots,$. Thus a simple calculation tells us that

$$\Delta^2_{t_k} f(\bar{x}, \bar{v})(h_k) = \langle A_i h, h \rangle + \frac{\langle h_k, \bar{v} - A_i\bar{x} - a_i \rangle}{\frac{1}{2}t_k^2} = \langle A_i h, h \rangle, \tag{3.72}$$

which clearly implies that $\Delta^2_{t_k} f(\bar{x}, \bar{v})(h_k) \to \mathrm{d}^2 f(\bar{x}, \bar{v})(h)$ as $k \to \infty$ with $\limsup_{k\to\infty} \frac{\|h_k - h\|}{t_k} = 0 < \infty$. Hence, f is parabolic regular at $\bar{x}$ for $\bar{v}$. This function is also parabolical epi-differentiable at $\bar{x}$ for $\bar{v}$.

Another example for parabolic regularity and parabolical epi-differentiable function is the sum of the k largest eigenvalues functions for symmetric matrix is parabolically regular.

3.2 First/second second order chain rules of subderivatives

Definition 3.4. *We say that the set-valued mapping $\Psi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is metric regular at a point $(\bar{x}, \bar{y}) \in \mathrm{gr}\Psi$, at a rate c, if there exist neighbourhood U of $\bar{x}$ and V of $\bar{y}$,*

$$d(x, \Psi^{-1}(y)) \leq cd(y, \Psi(x)) \;\; \forall x \in U, y \in V. \tag{3.73}$$

If $y = \bar{y}$ in (3.73), the set-valued mapping Ψ is said to be metric subregular $(\bar{x}, \bar{y})$. The infimum of the set of values c for which this holds is the modulus of metric regularity, denoted by $\mathrm{reg}\Psi(\bar{x}, \bar{y})$.

Consider the constraint system $\Psi(x) = F(x) - K$, where $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differentiable mapping and $K \subset \mathbb{R}^m$ a closed set.

Definition 3.5. *We say that Robinson's constraint qualification holds at a point $\bar{x} \in \mathbb{R}^n$ such that $F(\bar{x}) \in K$, with respect to the mapping $F(\cdot)$ and the set K, if the following condition holds:*

$$0 \in \mathrm{int}\{F(\bar{x}) + DF(\bar{x})\mathbb{R}^n - K\} \tag{3.74}$$

The set-valued mapping $\Psi(x) = F(x) - K$ is metric regular at $(\bar{x}, 0)$, i.e., there exist neighborhoods U of $\bar{x}$ and V of 0 and a constant number $c > 0$ such that

$$d(x, F^{-1}(K - y)) \leq cd(F(x) + y, K) \quad \forall x \in U, y \in V. \tag{3.75}$$

It is well known that Ψ is metric regular at $(\bar{x}, 0)$ if and only if

$$y^* \in N_K(F(\bar{x})), DF(\bar{x})^* y^* = 0 \Rightarrow y^* = 0 \tag{3.76}$$

and both are equivalent to (3.74) whenever K is convex.

Set $C := \{x \in \mathbb{R}^n | F(x) \in K\}$ and $\bar{x} \in C$. Then we always have

$$T_C(\bar{x}) \subset \{h \in \mathbb{R}^n | DF(\bar{x})h \in T_K(F(\bar{x}))\}; \quad (3.77)$$
$$T_C^2(\bar{x}, h) \subset \{w \in \mathbb{R}^n | DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle \in T_K^2(F(\bar{x}), DF(\bar{x})h)\}, \quad (3.78)$$

and the equality hold whenever (3.75) is satisfied. C is parabolically derivable at x for h whenever K is Clarke regular at $F(\bar{x})$ and parabolically derivable at $F(\bar{x})$ for $DF(\bar{x})h$ (in particular K is convex) under the assumption of (3.76).

Proposition 3.4. *Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be a continuously differentiable mapping and $g\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ be l.s.c. convex function takeing a finite value at point $\bar{y} = F(\bar{x})$. Suppose that Robinson's constraint qualification condition*

$$0 \in \operatorname{int}\{F(\bar{x}) + DF(\bar{x})\mathbb{R}^n - \operatorname{dom} g\} \quad (3.79)$$

holds. Then the lower and upper directional epiderivatives of the composite function $\psi := g \circ F$ coincide at $\bar{x}$, and

$$(\psi)^{\downarrow}(\bar{x}, h) = g^{\downarrow}(F(\bar{x}), DF(\bar{x})h). \quad (3.80)$$

Proof. Let $\hat{K} := \operatorname{epi} g$ and $\hat{F}(x, \alpha) := (F(x), \alpha), \alpha \in \mathbb{R}$. It is not difficult to verify that $\hat{F}^{-1}(\hat{K}) = \operatorname{epi}(g \circ F)$ and that (3.79) implies that Robinson's constraint qualification for the set $\hat{K}$ and the mapping $\hat{F}$, at $(\bar{x}, g(F(\bar{x})))$, i.e.

$$0 \in \operatorname{int}\Big\{ \begin{bmatrix} F(\bar{x}) \\ g(F(\bar{x})) \end{bmatrix} + \begin{bmatrix} DF(\bar{x})\mathbb{R}^n \\ \mathbb{R} \end{bmatrix} - \operatorname{epi} g \Big\}.$$

It follows that

$$T_{\operatorname{epi}(g \circ F)}(\bar{x}, (g \circ F)(\bar{x})) = D\hat{F}(\bar{x}, \alpha)^{-1} T_{\operatorname{epi}(g)}(\hat{F}(\bar{x}), g(F(\bar{x}))). \quad (3.81)$$

□

The following second order chain rules can be proved in a similar way.

Theorem 3.2. *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a twice continuously differentiable mapping and $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ be a l.s.c. convex function taking a finite value at a point $\bar{y} := F(\bar{x})$. Suppose that Robinson's constraint qualification $0 \in \operatorname{int}\{F(\bar{x}) + DF(\bar{x})\mathbb{R}^n - \operatorname{dom} g\}$ holds. Then, we have the following formula for twice parabolically lower epiderivative of composite function $\psi = g \circ F$ provided $g^{\downarrow}(F(\bar{x}), DF(\bar{x})h)$ is finite.*

$$\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h, h)), \quad (3.82)$$

$$\psi_+^{\downarrow\downarrow}(\bar{x}; h, w) = g_+^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h, h)), \quad (3.83)$$

We have the following result concerning outer second order regularity of composite function

Proposition 3.5. *Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be a twice continuously differentiable mapping, $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ a lower semicontinuous convex function taking a finite value at a point $\bar{y} := F(\bar{x})$, and $h \in \mathbb{R}^n$ and $\lambda \in \partial g(\bar{y})$ satisfying $DF(\bar{x})h \in K_g(F(\bar{x}), \lambda)$. Suppose that g is outer second order regular at $\bar{y}$ in the direction $DF(\bar{x})h$, and that Robinson's constraint qualification $0 \in \operatorname{int}\{F(\bar{x}) + DF(\bar{x})X - \operatorname{dom} g\}$ holds. Then the composite function $\psi = g \circ F$ is outer second order regular at $\bar{x}$ in the direction h.*

Theorem 3.3. *Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be a twice continuously differentiable mapping, $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ a lower semicontinuous convex function taking a finite value at a point $\bar{y} := F(\bar{x})$, $h \in \mathbb{R}^n$ and $\lambda \in \partial g(\bar{y})$ satisfying $DF(\bar{x})h \in K_g(F(\bar{x}), \lambda)$. Suppose that g is outer second order regular at $\bar{y}$ in the direction $DF(\bar{x})h$, and that Robinson's constraint qualification $0 \in \operatorname{int}\{F(\bar{x}) + DF(\bar{x})\mathbb{R}^n - \operatorname{dom} g\}$ holds. Then the composite function $\psi = g \circ F$ is parabolically regular at $\bar{x}$ in the direction h for $v := [DF(\bar{x})]^* \lambda$, and*

$$\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h,h)); \tag{3.84}$$

and

$$\mathrm{d}^2\psi(\bar{x}, v)(h) = \inf_{w \in X}\{g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h,h)) - \langle v, w\rangle\}. \tag{3.85}$$

Proof. By Proposition 3.5, it follows from outer second order regularity of g that the composite function $g \circ F$ is outer second order regular at $\bar{x}$ in direction h. Also, since $\lambda \in \partial g(\bar{y})$ and $F(\bar{x} + th^{'}) = F(\bar{x}) + tDF(\bar{x})h^{'} + o(t)$ for $h^{'} \to h$ and $t \downarrow 0$, we have

$$g(F(\bar{x} + th^{'})) - g(F(\bar{x})) - t\langle \lambda, DF(\bar{x})h^{'}\rangle \geq o(t),$$

and hence $\mathrm{d}^2(g \circ F)(\bar{x}, v)(h) \geq 0$. By Proposition 3.2, the parabolic regularity of $g \circ F$ then follows. Formula (3.85) follows from (3.84) and the corresponding formula from Theorem 3.2 for the lower second order epiderivative of the composite function. □

Consider the composite function $\psi(x) = g(F(x))$, where $F\colon \mathbb{R}^n \to \mathbb{R}^m$ a twice differentiable at $\bar{x}$ and $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is proper, l.s.c., convex, and Lipschitz continuous around $F(\bar{x})$ relative to its domain. It is easy to see that

$$\operatorname{dom}\psi = \{x \in \mathbb{R}^n | F(x) \in \operatorname{dom} g\}.$$

It is clear that the Robinson condition (3.79) implies that $\Psi(x) = F(x) - \operatorname{dom} g$ is metric subregular at $(\bar{x}, 0)$, i.e., if there exist a constant $c > 0$ and a neighbourhood U of $\bar{x}$,

$$d(x, \operatorname{dom}\psi)) \leq \kappa d(F(x), \operatorname{dom} g) \;\; \forall x \in U. \tag{3.86}$$

Theorem 3.4. *Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\bar{x}$ and $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ be Lipschitz continuous around $F(\bar{x})$ relative to its domain. If the metric subregularity constraint qualification (3.86) hold, then the following hold:*

(i) for any $h \in \mathbb{R}^n$, the following subderivative chain rule for ψ at $\bar{x}$ holds:

$$\mathrm{d}\psi(\bar{x})(h) = \mathrm{d}g(F(\bar{x}))(DF(\bar{x})h);$$

(ii) we have the chain rules

$$T_{\operatorname{dom}\psi}(\bar{x}) = \{h \in \mathbb{R}^n | DF(\bar{x})h \in T_{\operatorname{dom} g}(F(\bar{x}))\}.$$

If, in addition, F is continuously differentiable at $\bar{x}$, g is convex, then

$$\partial\psi(\bar{x}) = DF(\bar{x})^*\partial g(F(\bar{x})).$$

Proof. Since $F(x)$ is continuously differentiable at $\bar{x}$,

$$F(\bar{x}+th) = F(\bar{x}) + tDF(\bar{x})h + o(t\|h\|).$$

It is easy to show that

$$T_{\mathrm{dom}\psi}(\bar{x}) \subset \{h \in \mathbb{R}^n | DF(\bar{x})h \in T_{\mathrm{dom}g}(F(\bar{x}))\}.$$

Take h being such that $DF(\bar{x})h \in \mathrm{dom}\ \mathrm{d}g(F(\bar{x}))$, there exist sequences $t_k \downarrow 0$ and $v_k \to DF(\bar{x})h$ such that $F(\bar{x}) + t_k v_k \in \mathrm{dom}g$ for all $k \geq 1$. Then the (3.86) yields

$$d(\bar{x}+t_k h, \mathrm{dom}\psi) \leq \kappa d(F(\bar{x}+t_k h), \mathrm{dom}g), \quad k \geq 1, \tag{3.87}$$

which in turn implies

$$\begin{aligned} \frac{d(\bar{x}+t_k h, \mathrm{dom}\psi)}{t_k} &\leq \frac{\kappa}{t_k} d(F(\bar{x}) + t_k DF(\bar{x})h + o(t_k), \mathrm{dom}g) \\ &\leq \frac{\kappa}{t_k}\|F(\bar{x}) + t_k DF(\bar{x})h + o(t_k) - F(\bar{x}) - t_k v_v\|) \\ &= \kappa\|DF(\bar{x})h - v_k + \frac{o(t_k)}{t_k}\| \ \ \forall k \geq 1. \end{aligned} \tag{3.88}$$

This implies that $h \in T_{\mathrm{dom}\psi}(\bar{x})$.

For any $h \in \mathbb{R}^n$, we have

$$\begin{aligned} \mathrm{d}\psi(\bar{x})(h) &= \liminf_{\substack{t\downarrow 0 \\ h' \to h}} \frac{g(F(\bar{x}+th')) - g(F(\bar{x}))}{t} \\ &= \liminf_{\substack{t\downarrow 0 \\ h' \to h}} \frac{g(F(\bar{x}) + tDF(\bar{x})h' + o(t\|h\|)) - g(F(\bar{x}))}{t} \\ &= \liminf_{\substack{t\downarrow 0 \\ h' \to h}} \frac{g(F(\bar{x}) + tDF(\bar{x})h' + o(t\|h\|)) - g(F(\bar{x}))}{t} \\ &\geq \mathrm{d}g(F(\bar{x}))(DF(\bar{x})h). \end{aligned}$$

Take any $h \in R^n$ and observe from the Lipschitz continuity of g around $F(\bar{x})$ relative to its domain that $\mathrm{d}g(F(\bar{x}))(DF(\bar{x})h) > -\infty$. Since the converse inequality is obvious if $\mathrm{d}g(F(\bar{x}))(DF(\bar{x})h) = \infty$, we may assume that the value $\mathrm{d}g(F(\bar{x}))(DF(\bar{x})h)$ is finite, i,e., By definition, there exist sequences $t_k \downarrow 0$ and $v_k \to DF(\bar{x})h$ such that

$$\mathrm{d}g(F(\bar{x}))(DF(\bar{x})h) = \lim_{k\to\infty} \frac{g(F(\bar{x}) + t_k v_k)) - g(F(\bar{x}))}{t_k} < \infty.$$

Suppose without lost of generality that $F(\bar{x}) + t_k v_k \in \mathrm{dom}g$ for all $k \geq 1$. From the above proof, we see that there exist $h_k \in \frac{\mathrm{dom}\psi - \bar{x}}{t_k}$ satisfying

$$\|h_k - h\| \leq \kappa\|DF(\bar{x})h - v_k + \frac{o(t_k)}{t_k}\| + \frac{1}{k}.$$

It follows that $h_k \to h \in$ as $k \to \infty$. Therefore,

$$\begin{array}{rl} & \mathrm{d}g(F(\bar{x}))(DF(\bar{x})h) \\ = & \lim_{k\to\infty} [\frac{g(F(\bar{x}+t_kh_k)) - g(F(\bar{x}))}{t_k} + \frac{g(F(\bar{x})+t_kv_k) - g(F(\bar{x}+t_kh_k))}{t_k}\|] \\ \geq & \liminf_{k\to\infty} \frac{g(F(\bar{x}+t_kh_k)) - g(F(\bar{x}))}{t_k} - \ell \lim_{k\to\infty} \|\frac{F(\bar{x}+t_kh_k) - F(\bar{x})}{t_k} - v_k\| \\ \geq & \mathrm{d}\psi(\bar{x})(h) \end{array}$$

We remain to prove that

$$\partial\psi(\bar{x}) = DF(\bar{x})^*\partial g(F(\bar{x}))$$

whenever F is continuously differentiable at $\bar{x}$, g is convex.

The inclusion $\partial\psi(\bar{x}) \supset DF(\bar{x})^*\partial g(F(\bar{x}))$ always hold. The inclusion

$$\partial\psi(\bar{x}) \subset DF(\bar{x})^*\partial g(F(\bar{x}))$$

is a very fundamental calculus formula in variational analysis. It is well know that this is true under Robinson's condition (3.79). It was proved under the weaker condition (3.86) in [6] very recently. This inclusion has also been proved under (3.86) in special case g is a indicator function of a closed convex set and F is Lipschitz continuous in [4, 5].

This Theorem remains true if we replace the metric subregularity condition (3.86) by the assumption that Ψ is metric-regula at $(\bar{x}, 0)$ and epi$g$ is Clarke regular at $(\bar{x}, g(\bar{x}))$.

Theorem 3.5. *Suppose that $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is twice differentiable at $\bar{x}$ and $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is proper, l.s.c., convex and Lipschitz continuous around $F(\bar{x})$ relative to its domain and suppose that the metric subregularity constraint qualification (3.86) hold. If g is parabolically epi-differentiable at $\bar{y} := F(\bar{x})$ in the direction $DF(\bar{x})h$ with $h \in T_{\mathrm{dom}f}(\bar{x})$, then f is parabolically epi-differentiable at $\bar{x}$ for h and the following conditions hold:*

(i) for every $w \in \mathbb{R}^n$,

$$\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h,h)); \tag{3.89}$$

(ii)

$$\begin{array}{rcl} \mathrm{dom}\psi_-^{\downarrow\downarrow}(\bar{x}; h, \cdot) & = & T^2_{\mathrm{dom}\psi}(\bar{x}, h) \\ & = & \{w \in \mathbb{R}^n | DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle \in T^2_{\mathrm{dom}g}(F(\bar{x}), DF(\bar{x})h)\}. \end{array}$$

Proof. It is easy to show that

$$T^2_{\mathrm{dom}\psi}(\bar{x}, h) \subset \{w \in \mathbb{R}^n | DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle \in T^2_{\mathrm{dom}g}(F(\bar{x}), DF(\bar{x})h)\}.$$

Let w be satisfying

$$z := DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle \in T^2_{\mathrm{dom}g}(F(\bar{x}), DF(\bar{x})h). \tag{3.90}$$

Then there exists $t_k \downarrow 0$, $z_k \to z$ such that

$$F(\bar{x}) + t_kDF(\bar{x})h + \frac{1}{2}t_k^2 z_k \in \mathrm{dom}g \tag{3.91}$$

Let $x_k := \bar{x} + t_k h + \frac{1}{2}t_k^2 w$. Then by (3.86) we have

$$\begin{aligned} d(x_k, \text{dom}\psi) &\leq cd(F(x_k), \text{dom}g) \\ &\leq c\|F(x_k) - F(\bar{x}) - t_k DF(\bar{x})h - \frac{1}{2}t_k^2 z_k\| \\ &= c\|\frac{1}{2}t_k^2(DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle) + o(t_k^2) - \frac{1}{2}t_k^2 z_k\| \end{aligned}$$

It follows that $\frac{d(x_k,\text{dom}\psi)}{t_k^2} \to 0$ and so $w \in T^2_{\text{dom}\psi}(\bar{x}, h)$.

Since F is twice continuously differentiable at $\bar{x}$, it is not hard to prove that for every $w \in \mathbb{R}^n$, $z = DF(\bar{x})w + \langle h, DF(\bar{x})h\rangle$,

$$g^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, z) \leq \psi_-^{\downarrow\downarrow}(\bar{x}, h, w). \tag{3.92}$$

For every $w \in T^2_{\text{dom}\psi}(\bar{x}, h)$, there exist exists $t_k \downarrow 0$, $w_k \to w$ such that

$$x_k := \bar{x} + t_k h + \frac{1}{2}t_k^2 w_k \in \text{dom}\psi.$$

Since g is parabolically epi-differentiable at $\bar{y} := F(\bar{x})$ in the direction $DF(\bar{x})h$, we have $\text{dom}g_-^{\downarrow\downarrow}(F(\bar{x}), DF(\bar{x})h), \cdot) \neq \emptyset$ and for z and the above $t_k \downarrow 0$, there exist $z_k \to z$ such that

$$g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h), z) = \lim_{k\to\infty} \frac{g(F(\bar{x}) + t_k DF(\bar{x})h + \frac{1}{2}t_k^2 z_k) - g(F(\bar{x})) - t_k \text{d}g(F(\bar{x}), DF(\bar{x})h))}{\frac{1}{2}t_k^2}.$$

Since $z \in T^2_{\text{dom}g}(F(\bar{x}), DF(\bar{x})h) = \text{dom}g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h), \cdot)$ due to Proposition 2.3, we have $g_-^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h), z) < \infty$. It follows that $y_k := F(\bar{x}) + t_k DF(\bar{x})h + \frac{1}{2}t_k^2 z_k \in \text{dom}g$ for k sufficiently large. By the Lipschitz property of g we obtain

$$\begin{aligned} &\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) \\ &\quad= \liminf_{k\to\infty} \frac{\psi(\bar{x} + t_k h + \frac{1}{2}w_k)) - \psi(\bar{x}) - t_k \text{d}\psi(\bar{x}, h)}{\frac{1}{2}t_k^2} \\ &\quad\leq \limsup_{k\to\infty} \frac{g(F(x_k)) - g(F(\bar{x})) - t_k \text{d}g(F(\bar{x}), DF(\bar{x})h))}{\frac{1}{2}t_k^2} \\ &\quad\leq \limsup_{k\to\infty} \frac{g(y_k) - g(F(\bar{x}) - t_k \text{d}g(F(\bar{x})), DF(\bar{x})h))}{\frac{1}{2}t_k^2} + \limsup_{k\to\infty} \frac{g(F(x_k)) - g(y_k)}{\frac{1}{2}t_k^2} \\ &\quad\leq g^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, z). \end{aligned}$$

Therefore, we have prove that

$$\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = g^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, z)$$

for the case $w \in T^2_{\text{dom}\psi}(\bar{x}, h)$.

For $w \notin T^2_{\text{dom}\psi}(\bar{x}, h)$, since $\text{dom}\psi_-^{\downarrow\downarrow}(\bar{x}; h\cdot) \subset T^2_{\text{dom}\psi}(\bar{x}, h)$ always hold, we have $\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = +\infty$ and $z \notin T^2_{\text{dom}g}(F(\bar{x}), DF(\bar{x})h) = \text{dom}g_-^{\downarrow\downarrow}(F(\bar{x}), DF(\bar{x})h), \cdot)$. This implies that $g^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, z) = +\infty$ and $w \notin \text{dom}\psi_-^{\downarrow\downarrow}(\bar{x}; h, \cdot)$ duo to (3.92). This show that

$$\psi_-^{\downarrow\downarrow}(\bar{x}; h, w) = g^{\downarrow\downarrow}(F(\bar{x}); DF(\bar{x})h, z) = +\infty$$

for the case $w \notin T^2_{\mathrm{dom}\psi}(\bar{x}, h)$ and $\mathrm{dom}\psi^{\downarrow\downarrow}_{-}(\bar{x}; h\cdot) = T^2_{\mathrm{dom}\psi}(\bar{x}, h)$. Hence, we complete the proofs of (i) and (ii). □

It is worth mentioning that a chain rule for parabolic subderivatives for the composite form (3.89) was achieved in [14, Exercise 13.63] and [2, Proposition 3.42] when g is merely a proper l.s.c. function and the assumption that $\Psi(x) = F(x) - \mathrm{dom} g$ is metric regular at $(\bar{x}, 0)$.

Given $\bar{v} \in \partial\psi(\bar{x})$, we define the set of Lagrangian multipliers associated with $(\bar{x}, \bar{v})$ by

$$\Lambda(\bar{x}, \bar{v}) := \{\lambda \in \mathbb{R}^m | DF(\bar{x})^*\lambda = \bar{v}, \lambda \in \partial g(F(\bar{x})). \tag{3.93}$$

It is easy to see that $h \in K_\psi(\bar{x}, \bar{v}) \Leftrightarrow DF(\bar{x})h \in K_g(F(\bar{x}), \lambda) \quad \forall \lambda \in \Lambda(\bar{x}, \bar{v})$ whenever either the Robinson's condition (3.79) holds or the metric subreguarity (3.86) and g is Lipschitz continuous around $F(\bar{x})$ relative to its domain.

Proposition 3.6. *Suppose that $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is twice differentiable at $\bar{x}$ and $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is proper, l.s.c., convex and Lipschitz continuous around $\bar{y} = F(\bar{x})$(with $g(\bar{y})$ finite) relative to its domain and suppose that the metric subregularity constraint qualification (3.86) hold. If for every $\bar{v} \in \partial\psi(\bar{x})$ and $\lambda \in \Lambda(\bar{x}, \bar{v})$, g is parabolically epi-differentiable at $\bar{y}$ in every direction $d \in K_g(\bar{y}, \lambda)$, then for every $h \in \mathbb{R}^n$ we have the lower estimate*

$$\mathrm{d}^2\psi(\bar{x}, \bar{v})(h) \geq \sup_{\lambda \in \Lambda(\bar{x}, \bar{v})} \{\langle \lambda, D^2F(\bar{x})(h, h)\rangle + \mathrm{d}^2 g(F(\bar{x}), \lambda)(DF(\bar{x})h)\}; \tag{3.94}$$

while for every $h \in K_f(\bar{x}, \bar{v})$ we obtain the upper estimate

$$\mathrm{d}^2\psi(\bar{x}, \bar{v})(h) \leq \inf_{w \in \mathbb{R}^n} \{g^{\downarrow\downarrow}_{-}(F(\bar{x}); DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h, h)) - \langle \bar{v}, w\rangle\} < \infty. \tag{3.95}$$

Proof. For any $\lambda \in \Lambda(\bar{x}, \bar{v})$, we can write

$$\begin{aligned}
\Delta_t^2\psi(\bar{x}, \bar{v})(h) &= \frac{g(F(\bar{x} + th)) - g(F(\bar{x})) - t\langle \bar{v}, h\rangle}{\frac{1}{2}t^2} \\
&= \frac{g(F(\bar{x}) + t\frac{F(\bar{x}+th) - F(\bar{x})}{t}) - g(F(\bar{x})) - t\langle \lambda, DF(\bar{x})h\rangle}{\frac{1}{2}t^2} \\
&= \frac{g(F(\bar{x}) + t\frac{F(\bar{x}+th) - F(\bar{x})}{t}) - g(F(\bar{x})) - t\langle \lambda, \frac{F(\bar{x}+th) - F(\bar{x})}{t}\rangle}{\frac{1}{2}t^2} \\
&+ \frac{\langle \lambda, F(\bar{x} + th) - F(\bar{x}) - tDF(\bar{x})h\rangle}{\frac{1}{2}t^2}
\end{aligned}$$

Because $\frac{F(\bar{x}+th') - F(\bar{x})}{t} \to DF(\bar{x})h$ as $t \to 0$ and $h' \to h$, while at the same time

$$\frac{\langle \lambda, F(\bar{x} + th') - F(\bar{x}) - tDF(\bar{x})h'\rangle}{\frac{1}{2}t^2} \to \langle \lambda, D^2F(\bar{x})(h, h)\rangle,$$

we have

$$\mathrm{d}^2\psi(\bar{x}, \bar{v})(h) \geq \mathrm{d}^2 g(F(\bar{x}), \lambda)(DF(\bar{x})h) + \langle \lambda, D^2F(\bar{x})(h, h)\rangle.$$

This complete the proof of (3.94). The upper estimation (3.95) follows from Proposition 3.5 and Proposition 2.4 □

The above result carries important information by which we can achieve a chain rule for the second subderivative. To do so, we should look for conditions under which the lower and upper estimates (3.94) and (3.95), respectively, coincide. This motivates us to consider the unconstrained optimization problem

$$\min_{w\in\mathbb{R}^n}\{g_-^{\downarrow\downarrow}(F(\bar{x}), DF(\bar{x})h, DF(\bar{x})w + D^2F(\bar{x})(h,h)) - \langle\bar{v}, w\rangle\} \tag{3.96}$$

Proposition 3.7. *Suppose that $F\colon\mathbb{R}^n\to\mathbb{R}^m$ is twice differentiable at $\bar{x}$ and $g\colon\mathbb{R}^m\to\bar{\mathbb{R}}$ is proper, l.s.c., convex and Lipschitz continuous around $\bar{y}=F(\bar{x})$(with $g(\bar{y})$ finite) relative to its domain and suppose that the metric subregularity constraint qualification (3.86) hold. If for every $\bar{v}\in\partial\psi(\bar{x})$ and $\lambda\in\Lambda(\bar{x},\bar{v})$, g is parabolically epi-differentiable at $\bar{y}$ in every direction $d\in K_g(\bar{y},\lambda)$, and parabolically regular at $F(\bar{x})$ for λ , then for every $h\in K_\psi(\bar{x},\bar{v})$, the dual problem of (3.96) is given by*

$$\mathrm{d}^2\psi(\bar{x},\bar{v})(h) = \max_{\lambda\in\Lambda(\bar{x},\bar{v})}\langle\lambda, D^2F(\bar{x})(h,h)\rangle + \mathrm{d}^2g(F(\bar{x}),\lambda)(DF(\bar{x})h), \tag{3.97}$$

Proof. (i) By the classical Fenchel-Rockafellar duality theorem, we know that the dual problem of (3.96) is given by

$$\max_{\lambda\in\Lambda(\bar{x},\bar{v})}\langle\lambda, D^2F(\bar{x})(h,h)\rangle + \mathrm{d}^2g(F(\bar{x}),\lambda)(DF(\bar{x})h). \tag{3.98}$$

□

Theorem 3.6. *Suppose that $F\colon\mathbb{R}^n\to\mathbb{R}^m$ is twice differentiable at $\bar{x}$ and $g\colon\mathbb{R}^m\to\bar{\mathbb{R}}$ is proper, l.s.c., convex and Lipschitz continuous around $\bar{y}=F(\bar{x})$(with $g(\bar{y})$ finite) relative to its domain and suppose that the metric subregularity constraint qualification (3.86) hold. Suppose that for every $\bar{v}\in\partial\psi(\bar{x})$ and $\lambda\in\Lambda(\bar{x},\bar{v})$, g is parabolically epi-differentiable at $\bar{y}$ in the direction $d\in K_g(F(\bar{x}),\lambda)$, and parabolically regulat at $F(\bar{x})$ for λ. Then ψ is parabolically regular at $\bar{x}$ for $\bar{v}$. Furthermore, for every $h\in\mathbb{R}^n$, the second subderivative of ψ at $\bar{x}$ for $\bar{v}$ is calculated by*

$$\mathrm{d}^2\psi(\bar{x},\bar{v})(h) = \max_{\lambda\in\Lambda(\bar{x},\bar{v})}\{\langle\lambda, D^2F(\bar{x})(h,h)\rangle + \mathrm{d}^2g(F(\bar{x}),\lambda)(DF(\bar{x})h)\} \tag{3.99}$$

Proof. If $h\in K_\psi(\bar{x},\bar{v})$, It follows from Proposition 3.6 and Proposition 3.7 and Theorem 3.5 that (3.100) is satisfied and

$$\mathrm{d}^2\psi(\bar{x},\bar{v})(h) = \inf_{w\in\mathbb{R}^n}\psi_-^{\downarrow\downarrow}(\bar{x};h,w) - \langle\bar{v},w\rangle\}. \tag{3.100}$$

By Theorem 3.5, ψ is parabolic epi-differntiable at $\bar{x}$ for every $h\in K_\psi(\bar{x},\bar{v})$ and $\mathrm{dom}\psi_-^{\downarrow\downarrow}(\bar{x},h,\cdot)\neq\emptyset$ for every $h\in K_\psi(\bar{x},\bar{v})$. So, since $\mathrm{d}^2\psi(\bar{x},\bar{v})$ is proper, by Proposition 2.5, $\mathrm{domd}^2\psi(\bar{x},\bar{v}) = K_\psi(\bar{x},\bar{v})$. Thus, if $h\notin K_\psi(\bar{x},\bar{v})$, then $\mathrm{d}^2\psi(\bar{x},\bar{v})(h)=\infty$.

On the other hand, since $h\notin K_\psi(\bar{x},\bar{v}) \Leftrightarrow DF(\bar{x})h\notin K_g(F(\bar{x}),\lambda)\ \ \forall\lambda\in\Lambda(\bar{x},\bar{v})$, we have $\mathrm{d}^2g(F(\bar{x}),\lambda)(DF(\bar{x})h)=\infty$. Therefore, both side in (3.100) are ∞. □

When g is a convex piecewise linear-quadratic (or more case where g is fully amenable), the parabolic regularity of the composite $\psi = g\circ F$ and chain rule 3.100) were established in [14, Theorem 13.67] under the stronger condition the metric regularity.

Corollary 3.1. *(chain rule for twice epi-differentiability). Suppose all the assumptions of Theorem 3.6. Then ψ is twice epi-differentiable at $\bar{x}$ for $\bar{v}$.*

4 The second order optimality conditions for composite optimization problem

We consider the following composite optimization problem:

$$(P) \qquad \min_{x\in\mathbb{R}^n}\{f(x)+g(F(x))\},$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is twice differentiable on $\mathbb{R}^n$, $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a twice differentiable mapping, and $g\colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is a proper lower semicontinuous extended real-valued function. Clearly, if $g(\cdot) := \delta_K(\cdot)$ is the indicator function of a nonempty set $K \subset Y$, then the problem (P) reduces to (CP).

$$(CP) \qquad \min_{x\in\mathbb{R}^n} f(x) \quad \text{subject to } F(x) \in K.$$

The Lagrangian function for (P) is defined by

$$L(x,\lambda) = f(x) + \langle \lambda, F(x)\rangle,$$

and the Lagrangian multiplier set is defined by

$$\Lambda(\bar{x}) = \{\lambda \in \mathbb{R}^m | \nabla f(\bar{x}) + DF(\bar{x})^*(\lambda) = 0, \lambda \in \partial g(F(\bar{x}))\}.$$

If $\bar{x}$ is a locally optimal solution of (P), then $0 \in \partial(f+g\circ F)(\bar{x}) = \nabla f(\bar{x}) + \partial(g\circ F)(\bar{x})$. Hence $-\nabla f(\bar{x}) \in \partial(g\circ F)(\bar{x})$. If the assumption Theorem 3.4 is satisfied, then $-\nabla f(\bar{x}) \in DF(\bar{x})^*\partial g(F(\bar{x}))$. It is clear in this case that

$$\Lambda(\bar{x}) = \Lambda(\bar{x}, -\nabla f(\bar{x})). \tag{4.101}$$

4.1 The second order optimality conditions for optimization problem

Proposition 4.1. *Let $\phi\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ be an extended lower semicontinuous real-valued function.* (i) *If $\bar{x}$ is a locally minimizer of ϕ on $\mathbb{R}^n$, then $\mathrm{d}^2\phi(\bar{x},0)(h) \geq 0$ for all $h \in \mathbb{R}^n$.*

(ii) *If the second order growth condition holds at $\bar{x}$ iff*

$$\mathrm{d}^2\phi(\bar{x},0)(h) > 0, \ \ \forall h \in \mathbb{R}^n \setminus \{0\}. \tag{4.102}$$

Proposition 4.2. *Assume that the first order necessary condition $\mathrm{d}\phi(\bar{x}) \geq 0$. Then*

(i) *If $\bar{x}$ is a local minimizer of ϕ, then*

$$\inf_{w\in\mathbb{R}^n} \phi_-^{\downarrow\downarrow}(\bar{x}; h, w) \geq 0, \ \ \forall h \in K_\phi(\bar{x},0). \tag{4.103}$$

(ii) *If, in addition, f is parabolically regular at $\bar{x}$, for $\bar{v} = 0$, then the second order growth condition holds at $\bar{x}$ iff*

$$\inf_{w\in\mathbb{R}^n} \phi_-^{\downarrow\downarrow}(\bar{x}; h, w) > 0, \ \ \forall h \in\in K_\phi(\bar{x},0). \tag{4.104}$$

Example 4.1. *Let S be a convex closed subset of X and consider the corresponding indicator function $g(\cdot) := \delta_S(\cdot)$, a point $x \in S$ and a direction $h \in T_S(x)$. Recall that $\mathrm{d}g(x,h) = \delta_{T_S(x)}(h)$, that $g_-^{\downarrow\downarrow}(x;h,\cdot) = \delta_{T_S^2(x,h)}$, and that the indicator function g is (outer) second order regular at x iff the set S is (outer) second order regular at x. Let $v \in N_S(x)$ be such that $\langle v,h\rangle = 0$. By proposition 3.2 we have that if S is outer second order regular at x in the direction h, then g is parabolically regular at x in the direction h, for v, and*

$$\mathrm{d}^2 g(x,v)(h) = \inf_{w\in T_S^2(x,h)} (-\langle v,w\rangle) = -\sigma(v, T_S^2(x,h)).$$

It is interesting to note that if $v = 0$, then $\mathrm{d}^2 g(x,0)(h) = 0$ whether S is second order regular or not. On the other hand, $\inf_w g_-^{\downarrow\downarrow}(x,h,w)$ is equal to 0 iff $T_S^2(x,h)$ is nonempty, and is $+\infty$ otherwise. Therefore, g is parabolically regular at x in the direction h , for $v = 0$, iff $T_S^2(x,h)$ is nonempty.

Results presented in this section can be used to derive second order optimality conditions for constrained problems. Let S be a closed set of $\mathbb{R}^n$, let $f\colon X \to \mathbb{R}$ be twice continuously differentiable function, and consider the problem

$$\min_{x\in S} f(x). \tag{4.105}$$

Clearly, the above problem is equivalent to minimization of the extended real valued function $\phi(\cdot) = f(\cdot) + \delta_S(\cdot)$ over $\mathbb{R}^n$.

Let $\bar{x} \in S$. We have then that for any $h \in X$,

$$\mathrm{d}\phi(\bar{x})(h) = \nabla f(\bar{x})h + \delta_{T_S(\bar{x})}(h).$$

It follows that if $\bar{x}$ is a local minimizer of f over S, then $\nabla f(\bar{x})h \geq 0$ for all $h \in T_S(\bar{x})$. Moreover, we have that

$$\phi_-^{\downarrow\downarrow}(\bar{x};h,w) = \nabla f(\bar{x})w + \nabla^2 f(\bar{x})(h,h) + \delta_{T_S^2(\bar{x},h)}(w),$$

for all $h \in T_S(\bar{x})$ and $w \in \mathbb{R}^n$, and hence

$$\inf_{w\in X} \phi_-^{\downarrow\downarrow}(\bar{x};h,w) = \nabla^2 f(\bar{x})(h,h) - \sigma(-\nabla f(\bar{x}), T_S^2(\bar{x},h)).$$

It follows then by Proposition 4.2 that if $\bar{x}$ is a local minimizer of f over S, then

$$\nabla^2 f(\bar{x})(h,h) - \sigma(-\nabla f(\bar{x}), T_S^2(\bar{x},h)) \geq 0, \;\; \forall h \text{ s. t. } \;\nabla f(\bar{x})h \in T_S(\bar{x}). \tag{4.106}$$

Condition (4.106) hold irrespective of S being convex or not. Also, we have that the function $\bar{f}$ is out second order regular at $\bar{x}$ iff the set S is outer second order regular at $\bar{x}$. Therefore it follows from Proposition 4.2 that if the space X is finite dimensional, the set S is outer second order regular at $\bar{x}$ and $\bar{x}$ satisfies the first order necessary optimality conditions, then the second order growth condition holds at $\bar{x}$ iff

$$\nabla^2 f(\bar{x})(h,h) - \sigma(-\nabla f(\bar{x}), T_S^2(\bar{x},h)) > 0, \;\; \forall\, h \neq 0 \text{ s. t. } \;\nabla f(\bar{x})h \in T_S(\bar{x}). \tag{4.107}$$

Suppose now that the set S is given in the form $S := F^{-1}(K)$, where $K \subset \mathbb{R}^m$ is a closed convex and $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a twice continuously differentiable mapping. Suppose further that Robinson's constraint qualification holds at $\bar{x}$. Then by the the chain rule (1.8) we have

$$T_S^2(\bar{x}, h) = \{w \colon DF(\bar{x})(w) + D^2F(\bar{x})(h, h) \in T_K^2(F(\bar{x}), DF(\bar{x})h)\}. \tag{4.108}$$

Combing (4.108), the second necessarily condition (4.106) becomes for every $h \in C(\bar{x}) := \{h \in \mathbb{R}^n | DF(\bar{x})h \in T_K(F(\bar{x}), \nabla f(\bar{x})h = 0\}$,

$$\nabla^2 f(\bar{x})(h, h) + \nabla f(\bar{x})w \geq 0, \tag{4.109}$$

$$\text{subject to } DF(\bar{x})(w) + D^2F(\bar{x})(h, h) \in T_K^2(F(\bar{x}), DF(\bar{x})h). \tag{4.110}$$

We now consider the second order optimality condition for (P). Let $\phi(x) = f(x) + g(F(x))$.

Theorem 4.1. *Suppose that $g \colon \mathbb{R}^m \to \bar{\mathbb{R}}$ is proper, l.s.c., convex and Lipschitz continuous around $\bar{y} = F(\bar{x})$(with $g(\bar{y})$ finite) relative to its domain and suppose that the metric subregularity constraint qualification (3.86) hold. Let $\bar{v} = -\nabla f(\bar{x}) \in \partial(g \circ F)(\bar{x})$ Suppose that for every $\lambda \in \Lambda(\bar{x}, \bar{v}) = \Lambda(\bar{x})$, g is parabolically epi-differentiable at $\bar{y}$ in every direction $d \in K_g(\bar{y}, \lambda)$, and parabolically regular at $F(\bar{x})$ for λ. Then the following second-order optimality conditions for the composite problem (P) are satisfied:*

(i) if $\bar{x}$ is a local minimum of (P), then the second-order necessary condition

$$\max_{\lambda \in \Lambda(\bar{x})} \{\langle \nabla_{xx}^2 L(\bar{x}, \lambda)h, h\rangle + \mathrm{d}^2 g(F(\bar{x}), \lambda)(DF(\bar{x})h)\} \geq 0 \tag{4.111}$$

holds for all $h \in C(\bar{x}) = \{h \in \mathbb{R}^n | \langle -\nabla f(\bar{x}), h\rangle = \mathrm{d}g(F(\bar{x}))(DF(\bar{x})h)\}$;

(ii) the second-order condition

$$\max_{\lambda \in \Lambda(\bar{x})} \{\langle \nabla_{xx}^2 L(\bar{x}, \lambda)h, h\rangle + \mathrm{d}^2 g(F(\bar{x}), \lambda)(DF(\bar{x})h)\} > 0 \quad \text{for all } h \in C(\bar{x}) \setminus \{0\} \tag{4.112}$$

is equivalent to the second order growth condition for $\phi(x)$.

4.2 The second order optimality condition for augmented Lagrangian function

We consider the augmented Lagrangian function for the following composite optimization problem:

$$(P) \qquad \min_{x \in \mathbb{R}^n} \{f(x) + g(F(x))\}.$$

The Moreau envelope $e_\nu \phi(\cdot)$ of a convex function ϕ at $u \in \mathbb{R}^m$ for parameter $\nu > 0$, defined as

$$e_\nu \phi(u) = \inf_{w \in \mathbb{R}^m} \{\phi(w) + \frac{1}{2\nu}\|u - w\|^2\}, \tag{4.113}$$

is real-valued, convex and continuous, and the infimum in (4.113) is uniquely attained for every $u \in \mathbb{R}^m$. We denote its unique minimizer by $P_\nu \phi(u)$, i.e.,

$$P_\nu \phi(u) := \mathrm{argmin}_{w \in \mathbb{R}^m} \{\phi(w) + \frac{1}{2\nu}\|u - w\|^2\} = (I + \nu \partial \phi)^{-1}(u), \tag{4.114}$$

which is called the *proximal mapping* of ϕ. Moreover, it was well known that that $e_\nu\phi(u)$ is differentiable, and its gradient

$$\nabla e_\nu\phi(u) = \nu^{-1}(u - P_\nu\phi(u)) \tag{4.115}$$

is ν^{-1}-Lipschitz continuous. Following [14], the augmented Lagrangian for the problem (P) can be expressed by

$$\begin{aligned}\mathcal{L}(x,\lambda,\tau) &= f(x) + \inf_{y\in\mathbb{R}^m}\{g(F(x)+y) + \frac{\tau}{2}\|y\|^2 - \langle\lambda, y\rangle\}\\ &= f(x) + \inf_{y\in\mathbb{R}^m}\{\theta(F(x)+y) + \frac{\tau}{2}\|\tau^{-1}\lambda - y\|^2\} - \frac{1}{2\tau}\|\lambda\|^2\\ &= f(x) + e_{\tau^{-1}}\theta(\tau^{-1}\lambda + F(x)) - \frac{1}{2\tau}\|\lambda\|^2. \end{aligned}\tag{4.116}$$

Proposition 4.3. *The set $\Lambda(\bar{x})$ of the Lagrange multipliers of the problem (P) at $\bar{x}$ can be expressed as*

$$\Lambda(\bar{x}) := \{\lambda \in \mathbb{R}^m : \nabla_x\mathcal{L}(\bar{x},\lambda,\tau) = 0 \text{ and } \nabla_\lambda\mathcal{L}(\bar{x},\lambda,\tau) = 0\},$$

for any $\tau > 0$.

In the following we assume that the function $g : \mathbb{R}^m \to \bar{\mathbb{R}}$ is proper lower semicontinuous convex and twice epi-differentiable at $F(\bar{x})$.

Lemma 4.1. *Suppose that the set $\Lambda(\bar{x})$ of the Lagrange multipliers is nonempty and $\lambda \in \Lambda(\bar{x})$. Then, for any $\xi \in \mathbb{R}^n$, one has*

$$\mathrm{d}^2 e_{\tau^{-1}}g(\tau^{-1}\lambda + F(\cdot))(\bar{x}, DF(\bar{x})^*\lambda)(h) = 2d_{\tau^{-1}}(DF(\bar{x})h) + \langle\lambda, D^2F(\bar{x})(h,h)\rangle, \tag{4.117}$$

where

$$d_{\tau^{-1}}(DF(\bar{x})h) := \inf_{y\in\mathbb{R}^m}\{\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(y) + \frac{\tau}{2}\|DF(\bar{x})h - y\|^2\} \tag{4.118}$$

denotes the Moreau envelope of $\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$ at $DF(\bar{x})h$.

Denote by $\ell_\tau(x) = \mathcal{L}(x,\lambda,\tau)$. From the proof of Proposition 4.3, we have $\nabla\ell_\tau(\bar{x}) = \nabla_x L(\bar{x},\lambda) = 0$. By Proposition 2.10 in [11], the second-order epi-derivative of $\ell_\tau(x)$ at $\bar{x}$ relative to $w = D\ell_\tau(\bar{x}) = 0$ can be expressed as

$$\mathrm{d}^2(\ell_\tau)(\bar{x},0)(h) = \nabla^2_{xx}L(\bar{x},\lambda)(h,h) + 2d_{\tau^{-1}}(DF(\bar{x})h). \tag{4.119}$$

The necessary and sufficient conditions are stated as follows:

Theorem 4.2. *Let $\lambda \in \Lambda(\bar{x})$. Then, $0 \in \hat{\partial}\ell_\tau(\bar{x})$ for any $\tau > 0$, and the following assertions hold:*

(a) *(Necessary condition). If $\ell_\tau(\cdot)$ has a local minimum at $\bar{x}$, then*

$$\nabla^2_{xx}L(\bar{x},\lambda)(h,h) + 2d_{\tau^{-1}}(DF(\bar{x})h) \geq 0, \text{ for all } h \in \mathbb{R}^n. \tag{4.120}$$

(b) *(Sufficient condition). If the condition*

$$\nabla^2_{xx}L(\bar{x},\lambda)(h,h) + 2d_{\tau^{-1}}(DF(\bar{x})h) > 0, \text{ for all } h \in \mathbb{R}^n \setminus \{0\}, \tag{4.121}$$

holds, then $\ell_\tau(\cdot)$ has a local minimum at $\bar{x}$ in the sense of quadratic growth condition.

Proposition 4.4. *Let $\lambda \in \Lambda(\bar{x})$ be a Lagrange multiplier of the problem (P) at $\bar{x}$. Then (4.121) is equivalent to the following condition*

$$\nabla^2_{xx}L(\bar{x},\lambda)(h,h) + \mathrm{d}^2 g(F(\bar{x}),\lambda)(DF(\bar{x})h) > 0, \ \forall\ h \in C(\bar{x}) \setminus \{0\} \tag{4.122}$$

for τ sufficiently large.

Proof. For any $h \in C(\bar{x})$, by taking $y = DF(\bar{x})h$, we have that

$$d_{\tau^{-1}}(DF(\bar{x})h) \le \frac{1}{2} d^2 g(F(\bar{x}),\lambda)(DF(\bar{x})h).$$

Hence, the condition (4.121) implies that

$$\nabla^2_{xx}L(\bar{x},\lambda)(h,h) + \mathrm{d}^2 g(F(\bar{x}),\lambda)(DF(\bar{x})h) > 0, \ \forall\ h \in C(\bar{x}) \setminus \{0\}. \tag{4.123}$$

Suppose that (4.121) doesn't hold. Then, there exist some sequence $\tau_k \to +\infty$ as $k \to \infty$ and $h_k \in \mathbb{R}^n$ with $\|h_k\| = 1$ such that

$$\nabla^2_{xx}L(\bar{x},\lambda)(h_k,h_k) + 2d_{\tau_k^{-1}}(DF(\bar{x})h_k) \le 0.$$

Observe that $d_{\tau_k^{-1}}(\cdot)$ defined in (4.118) is the Moreau envelope of $\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$. Since $\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$ is proper lower semicontinuous and convex, the infimum of $d_{\tau_k^{-1}}(DF(\bar{x})h_k)$ can be attained uniquely at $\bar{z}_k := \mathrm{Prox}_{\tau_k^{-1}}(\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(DF(\bar{x})h_k) \in \mathrm{dom}\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$. It follows that

$$\nabla^2_{xx}L(\bar{x},\lambda)(h_k,h_k) + \tau_k\|DF(\bar{x})h_k - \bar{z}_k\|^2 + \frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\bar{z}_k) \le 0. \tag{4.124}$$

Since $\{\nabla^2_{xx}L(\bar{x},\lambda)(h_k,h_k)\}$ is bounded, the term $\mathrm{d}^2 g(F(\bar{x}),\lambda)(\bar{z}_k)$ is nonnegative and $\tau_k \to +\infty$ as $k \to +\infty$, it follows from (4.124) that

$$\lim_{k\to+\infty} \|DF(\bar{x})h_k - \bar{z}_k\| = 0. \tag{4.125}$$

Because $\|h_k\| = 1$, by passing to a subsequence if necessary, we assume that $h_k \to h$. Hence, we have $\|h\| = 1$ and by (4.125), $\bar{z}_k \to DF(\bar{x})h$ as $k \to +\infty$. The lower semicontinuity of $\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$ implies that $DF(\bar{x})h \in \mathrm{dom}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$, and so that

$$\mathrm{d}g(F(\bar{x})(DF(\bar{x})h) = \langle \lambda, DF(\bar{x})h\rangle = \langle -\nabla f(\bar{x}), h\rangle.$$

This implies that $h \in C(\bar{x})$. Again, by the lower semicontinuity of $\frac{1}{2}\mathrm{d}^2 g(F(\bar{x}),\lambda)(\cdot)$, we have from (4.124) that

$$\nabla^2_{xx}L(\bar{x},\lambda)(h,h) + \mathrm{d}^2 g(F(\bar{x}),\lambda)(DF(\bar{x})h) \le 0,$$

which contradicts (4.123). □

References

[1] H. Attouch and R. J.-B. Wets, Epigraphical analysis, Ann. Inst. Henri Poincaré Anal.Non-Linéaire, 6(1989), pp. 73-100.

[2] J.F. Bonnans and A. Shapiro, Perturbation Analysis of Optimization Problems. Springer, New York (2000)

[3] A.L. Dontchev, R.T. Rockafellar, Implicit Functions and Solution Mappings, Springer, Berlin, 2009.

[4] R. Henrion, A. Jourani and J. V. Outrata, On the calmness of a class of multifunctions, SIAM J. Optim., 13 (2002), 603-618.

[5] H, Gfrerer, B. S. Mordukhovich, Robinson stability of parametric variational systems via variational analysis. SIAM J. Optim. 27(2017), 438-465.

[6] A. Mohammadi, B. S. Mordukhovich. Variational Analysis in Normed Spaces with Applications to Constrained Optimization. SIAM Journal on Optimization 31(2021):1, 569-603.

[7] A. Mohammadi and M. E. Sarabi, Twice epi-Differentiability of extended-real-valued functions with applications in composite optimization, SIAM J. Optim., 30(2020):3, 2379-2409.

[8] Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I: Basic Theory, II: Applications, Springer, Berlin (2006)

[9] R. A. Poliquin and R. T. Rockafellar, A calculus of epi-derivatives applicable to optimization, Canadian J. Math. 45 (1993), 879-896.

[10] R. Poliquin and R. T. Rockafellar, Generalized Hessian properties of regularized nonsmooth functions, SIAM J. Optim. 6 (1996), 1121-1137.

[11] R.T. Rockafellar, First- and second-order epi-differentiability in nonlinear programming, Trans. Amer. Math. Soc., 307(1988), pp. 75-108.

[12] R.T. Rockafellar, Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives, Math. Oper. Res., 14(1989), pp. 462-484.

[13] R.T. Rockafellar, Generalized second derivatives of convex functions and saddle function, Trans.Amer. Math. Soc., 322(1990), pp. 51-77.

[14] R.T. Rockafellar and R.J.-B. Wets, Variational Analysis, Springer-Verlag, New York, 1998.