Chapter VI
Nonsmooth Optimization

March 11, 2023

1 Generalized Derivatives

6.1.2 Continuity of Dini derivative

For a point in $\mathbb{E}$, prove the function $f^{-}(x;.)$ is Lipschitz if f is locally Lipschitz around x.

Proof: Note that

$$|f^{-}(x;h_1) - f^{-}(x;h_2)| = |\liminf_{t\downarrow 0} \frac{f(x+th_1)-f(x)}{t} - \liminf_{t\downarrow 0} \frac{f(x+th_2)-f(x)}{t}|.$$

But, note that $\limsup$ is sublinear and thus

$$\liminf x^r \geq \liminf x^r + y^r + \liminf -y^r \Rightarrow \liminf x^r + y^r - \liminf x^r \leq \limsup y^r.$$

Thus,

$$|\liminf_{t\downarrow 0} \frac{f(x+th_1)-f(x)}{t} - \liminf_{t\downarrow 0} \frac{f(x+th_2)-f(x)}{t}| \leq ||h_1 - h_2||.$$

Similarly,

$$|\liminf_{t\downarrow 0} \frac{f(x+th_2)-f(x)}{t} - \liminf_{t\downarrow 0} \frac{f(x+th_1)-f(x)}{t}| \leq ||h_1 - h_2||.$$

Hence, $|f^{-}(x;h_1) - f^{-}(x;h_2)| \leq ||h_1 - h_2||$.

6.1.4 Surjective Dini subdifferentials

Suppose the continuous function $f : \mathbb{E} \to \mathbb{R}$ satisfies the growth condition

$$\lim_{||x||\to+\infty} \frac{f(x)}{||x||} = +\infty. \tag{1}$$

For any element $\phi \in \mathbb{E}$, prove there is a point x in $\mathbb{E}$ with $\phi \in \partial_{-} f(x)$.

Proof: Note that $\langle \phi, . \rangle - f(.)$ also satisfies 1. Thus without loss of generality suppose $\phi = 0$. Thus, we need to show f has a local minimum if f satisfies 1. Suppose not! Then let $B_i := \{x \in \mathbb{E} : ||x|| \leq i\}$. Then B_i is compact and thus $f|_{B_i}$ obtains its minimum on B_i and assume it happens at $x_i \in B_i$. If $||x_i|| < i$, then x_i is a local minimum for f and we are done. Thus, suppose $||x_i|| = i$. Hence, we have $f(x_{i+1}) < f(x_i)$ by definition of x_{i+1} and the fact that $x_i \in B_{i+1}$ and also x_i is not a local minimum for $f|_{B_{i+1}}$ (and thus the strict inequality). Now note that $||x_i|| \to +\infty$ and thus $\lim_{i\to+\infty} \frac{f(x_i)}{||x_i||} = +\infty$ which is a contradiction as $f(x_i)$ is decreasing.

6.1.6. Failure of Dini calculus

Show that the inclusion

$$\partial_{-}(f+g)(x) \subseteq \partial_{-} f(x) + \partial_{-} g(x)$$

can fail for locally Lipschitz functions f and g.

Proof: Let $f(x) = ||x|| - ||x||^2$ and also $g(x) = ||x||^2$. Then we claim

$$\partial_{-} g(0) = \{0\}.$$

If $\phi \in \partial_- g(0)$ then for any $||h|| = 1$, and small enough $t > 0$ we have

$$\langle s, th \rangle \leq t^2 ||h||^2 \Rightarrow \langle s, h \rangle \leq t||h||^2 \Rightarrow \langle s, h \rangle = 0.$$

Thus, $s = 0$. It is clear that $0 \in \partial_- g(0)$. However, $(f+g)(x) = ||x||$ and hence $\partial_-(f+g)(0) = B$. Now if $\partial_-(f+g)(0) \subseteq \partial_- f(0) + \partial_- g(0)$ then $B \subseteq \partial_- f(0)$. Now, let $s \in B$ with $||s|| = 1$ and hence $\langle s, ts \rangle \leq t - t^2$ for small enough t. Thus, $t \leq t - t^2$ for small enough t, which is a contradiction.

Side: Note that the function

$$f : B_{\frac{1}{2}} \subseteq \mathbb{E} \to \mathbb{R}, f(x) = \sqrt{1 - ||x||^2}$$

is Lipschitz and also has no Dini subgradient at 0. In fact, suppose $\phi \in \partial_- f(0)$ then for each h with $||h|| = 1$,

$$\langle \phi, th \rangle \leq f(th) - 1 \text{ for small enough } t > 0.$$

But, $f(th) - 1 \leq 0$ and thus $\langle \phi, h \rangle \leq 0$ for all h and hence $\phi = 0$. Thus, $f(th) = 1$ if and only if $th = 0$ or $h = 0$.

6.1.9. Mean value theorem

- Suppose the function $f : \mathbb{E} \to \mathbb{R}$ is locally Lipschitz. For any points x and y, prove there is a real t in $(0, 1)$ satisfying

$$f(x) - f(y) \in \langle x - y, \partial_\diamond f(tx + (1-t)y) \rangle$$

- **Monocity and convexity** If the set C in $\mathbb{E}$ is open and convex and the function $f : C \to \mathbb{R}$ is locally Lipschitz, prove f is convex if and only if it satisfies

$$\langle x - y, \phi - \psi \rangle \geq 0 \text{ for all } x, y \in C, \phi \in \partial_\diamond f(x) \ \&\ \psi \in \partial_\diamond f(y).$$

- If $\partial_\diamond f(y) \subseteq kB$ for all points y near x, prove f has local Lipschitz constant k about x.

Proof:

- To be done!
- Suppose that f is convex on C, then $\partial_\diamond f(x) = \partial f(x)$ and thus

$$f(y) \geq f(x) + \langle \phi, y - x \rangle, f(x) \geq f(y) + \langle \psi, y - x \rangle \Rightarrow \langle \phi - \psi, x - y \rangle \geq 0.$$

Conversely, suppose that the above statement holds. Then, let $\phi \in \partial_\diamond f(x)$, note that this set is nonempty since f is locally Lipschitz on $\mathbb{E}$. Then,

$$f(y) - f(x) = \langle y - x, \psi \rangle \text{ for some } \psi \in f(ty + (1-t)x) \text{ for some } t \in (0, 1).$$

Now it suffices to prove that $\langle y - x, \psi \rangle \geq \langle y - x, \phi \rangle$. But we have

$$\langle tx + (1-t)y - x, \psi - \phi \rangle \geq 0 \Rightarrow \langle y - x, \psi - \phi \rangle \geq 0.$$

- Let y, z be in a small neighborhood about x, then

  $f(y) - f(z) = \langle y - z, \phi\rangle$ for some $\phi \in \partial_\diamond f(w)$ wherein w lies on the line segment $[y, z]$.

  Thus, $|f(y) - f(x)| \leq k||y - z||$.

6.1.11 Order statistics

Calculate the Dini, the Michel-Penot, and the Clarke directional derivatives and differentials of the function

$$x \in \mathbb{R}^n \to [x]_k.$$

Proof:

Dini directional derivative

Suppose that

$$[x]_1 = \cdots = [x]_{l_1} > [x]_{l_1+1} = \cdots = [x]_{l_1+l_2} > \cdots > [x]_{l_1+l_2+\cdots+l_{t-1}+1} = \cdots = [x]_{l_1+\cdot+l_t},$$

and assume

$$h_{i_{1,1}} \geq h_{i_{1,2}} \geq \cdots \geq h_{i_{1,l_1}}, h_{i_{2,1}} \geq \cdots \geq h_{i_{2,l_2}}, \cdots, h_{i_{t,1}} \geq \cdots \geq h_{i_{t,l_t}},$$

wherein, for all $1 \leq j \leq t$

$S_j := \{i_{j,l_l+\cdots+l_{j-1}+1}, \cdots, i_{j,l_l+\cdots+l_{j-1}+l_j}\} = \{l_1+\cdots+l_{j-1}+1, l_1+\cdots+l_{j-1}+2, \cdots, l_1+\cdots+l_{j-1}+l_j\}$.

Then, for $t > 0$ small enough,

$$[x + th]_k = [x]_k + th_{i_{j,k}} \text{ where } k \in S_j.$$

Thus, $[.]_k^-(x; h) = h_{i_{j,k}}$.

Michel-Penot directional derivative

From the above discussion we have $[.]_k^\circ(x; h) = h_{i_{j,1}}$ where $k \in S_j$.

6.1.12 Closed subdifferentials

- Suppose the function $f : \mathbb{E} \to (\infty, +\infty]$ is convex, and the point x lies in $\text{int}(\text{dom}\, f)$. Prove the convex subdifferential $\partial f(.)$ is closed at x; in other words, $x^r \to x$ and $\phi^r \to \phi$ in $\mathbb{E}$ with ϕ^r in $\partial f(x^r)$ implies $\phi \in \partial f(x)$.

  Suppose the real function f is locally Lipschitz around the point x in $\mathbb{E}$.

- For any direction h in $\mathbb{E}$, prove the Clarke directional derivative has the property that $-f^\circ(.; h)$ is lower semicontinuous at x.

- Deduce the Clarke subdifferential is closed at x.

- Deduce further the inclusion $\subseteq$ in the Intrinsic Clarke subdifferential theorem:

  $$\partial_\circ f(x) = \text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\},$$

  wherein outside of the measure zero set S, f is Gateaux differentiable.

- Show that Dini and Michel-Penot subdifferentials are not necessary closed.

Proof:

- Note that f on int(dom f) is continuous and thus closed. Hence, the proof is complete due to 4.2.8.

- We need to show for any $x^r \to x$ we have
$$\liminf_r -f^\circ(x^r;h) \geq -f^\circ(x;h) \iff \limsup_r f^\circ(x^r;h) \leq f^\circ(x;h).$$
This holds if and only if
$$\limsup_r \limsup_{t\downarrow 0, y\to x^r} \frac{f(y+th)-f(y)}{t}.$$
Let $\epsilon > 0$, then let y^r, t_r be such that $||y^r - x^r|| \leq \frac{1}{r}$ and $t_r < \frac{1}{r}$.
$$|\frac{f(y^r+t_rh)-f(y^r)}{t_r} - \limsup_{t\downarrow 0, y\to x^r} \frac{f(y+th)-f(y)}{t}| \leq \frac{\epsilon}{2^r}.$$
Then,
$$\limsup_r \limsup_{t\downarrow 0, y\to x^r} \frac{f(y+th)-f(y)}{t} = \limsup_r \frac{f(y^r+t_rh)-f(y^r)}{t_r} \leq \limsup_{t\downarrow 0, y\to x} \frac{f(y+th)-f(y)}{t} = f^\circ(x;h)$$

- Let $x^r \to x$ and also $\phi^r \to \phi$ where $\phi^r \in \partial_\circ f(x^r)$. Then we wish to show that $\phi \in \partial_\circ f(x)$. This holds true if and only if
$$\langle \phi, h\rangle \leq f^\circ(x;h).$$
But,
$$\langle \phi, h\rangle = \lim_r \langle \phi^r, h\rangle \leq \limsup_r f^\circ(x^r,h) \leq f^\circ(x;h).$$
This completes the proof.

- Note that $\text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\} \subseteq \partial_\circ f(x)$ as Clarke subdifferntials are convex and closed. Then we claim that $\text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\}$ is compact. In fact, $\text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\} \subseteq \partial_\circ f(x)$ and thus $\text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\}$ is bounded as $\partial_\circ f(x)$ is compact. Let $s_i = \lim_r \nabla f(x_i^r) \in \partial_\circ f(x)$ with $x_i^r \to x$. Let $||x_i^{r_i} - x|| < \frac{1}{i}$ and $||s_i - \nabla f(x_i^{r_i})|| < \frac{1}{i}$. Now
$$\lim_j \nabla f(x_j^{r_j}) = \lim_j s_j.$$
Thus, $\lim_j s_j \in \text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\}$.

Now let $s \in \partial_\circ f(x) \setminus \text{conv}\{\lim_r \nabla f(x^r) : x^r \to x, x^r \notin S\}$, then there exists $\phi \in \mathbb{E}$ such that
$$\langle s, \phi\rangle < a < b \leq \langle \lim_r \nabla f(x^r), \phi\rangle \text{ wherein } x^r \to x, x^r \notin S.$$
Let $\phi = y - x$. Then choose $x^r \to x$ with $x^r \notin S$ and $x^r \in [x, y]$. Thus,
$$0 < b - a \leq \langle \nabla f(x^r) - s, y - x\rangle.$$
We obtain a contradiction as we tend r to infinity.

2 Regularity and Strict Differentiability

6.2.6.

Prove that a unique Clarke subgradient implies regularity. Note that the function is Lipschitz about the point x.

Proof: Recall that Clarke subgradient is a unique vector ϕ if and only if

$$\limsup_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)}{t} = \langle \phi, h\rangle.$$

Note that

$$f^{-}(x;h) = -\limsup_{t\downarrow 0} \frac{f((x+th)-th)-f(x+th)}{t} = -\langle \phi, -h\rangle = \langle \phi, h\rangle.$$

This completes the proof.

6.2.7 Strict differentiability

A real function f has strict derivative ϕ at a point x in $\mathbb{E}$ if and only if it is locally Lipschitz around x with

$$\lim_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)}{t} = \langle \phi, h\rangle$$

for all direction h in $\mathbb{E}$. In particular, this holds if f is continuously differentiable around x with $\nabla f(x) = \phi$.

Proof: First, suppose that f has strict derivative ϕ at x. Then if f is not locally Lipschitz around x, then for any fixed $C \in \mathbb{R}_{++}$ and for every i, there exists $y_i, z_i \in B_{\frac{1}{i}}(x)$ such that $|f(y_i)-f(z_i)| > C||y_i - z_i||$. However,

$$0 = \lim_{i\to+\infty, t\downarrow 0} |\frac{f(y_i)-f(z_i)-\langle \phi, y_i - z_i\rangle}{||y_i - z_i||}| \geq C - \limsup_{i\to+\infty} \frac{\langle \phi, y_i - z_i\rangle}{||y_i - z_i||},$$

so,

$$||\phi|| \geq \limsup_{i\to+\infty} \frac{\langle \phi, y_i - z_i\rangle}{||y_i - z_i||} \geq C,$$

which is a contradiction. Thus, f is locally Lipschitz around x. Now, fix h and let $y \leftarrow y+th$ and $z \leftarrow y$. Thus,

$$0 = \lim_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)-t\langle \phi, h\rangle}{t} \Rightarrow \lim_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)}{t} = \langle \phi, h\rangle.$$

Conversely, suppose that

$$\lim_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)}{t} = \langle \phi, h\rangle$$

for all $\phi \in \mathbb{E}$ and also f is locally Lipschitz around x, then we wish to prove

$$\lim_{y,z\to x, y\neq z} \frac{f(y)-f(z)-\langle \phi, y-z\rangle}{||y-z||} = 0.$$

However, the above equals to,

$$\lim_{z\to x, w\in S^1, t\downarrow 0} \frac{f(z+tw)-f(z)-\langle\phi, tw\rangle}{t} = \lim_{z\to x, w\in S^1, t\downarrow 0} \frac{f(z+tw)-f(z)}{t} - \langle\phi, w\rangle = \lim_{z\to x, w\in S^1, t\downarrow 0} g(z,w,t).$$

Now suppose the above does not hold, then there exists (z_i, w_i, t_i) with $t_i \downarrow 0$ and $w_i \in S^1$ and also $||w_i|| \to ||w||$ such that $|g(z_i, w_i, t_i)| \geq \epsilon$ for some $\epsilon > 0$. However,

$$|g(z_i, w_i, t_i) - g(z_i, w, t_i)| \leq ||w - w_i|| + |\langle\phi, w - w_i\rangle|.$$

However, $|g(z_i, w, t_i)| \to 0$ and thus $g(z_i, w_i, t_i) \to 0$ as desired.

Now if f is continuously differentiable then $||\nabla f||$ is bounded above in a neighborhood of x and thus

$$||f(x+h) - f(x)|| \leq ||\nabla f(x+th)|| ||h|| \leq C||h||,$$

for some constant C; note that $t \in (0,1)$ comes from the Taylor expansion. Now for each $t \in (0, \epsilon)$ for some small enough ϵ, there exists $t^* \in (0,t)$ such that

$$\frac{f(y+th)-f(y)}{t} = \nabla f(y+t^*h)^T h.$$

Now if $y \to x$ and $t \downarrow 0$, then the above tends to $\nabla f(x)^T h = \langle\phi, h\rangle$.

6.1.8

Prove the following results:

- $f^\circ(x; -h) = (-f)^\circ(x; h)$
- $(\lambda f)^\circ(x; h) = \lambda f^\circ(x; h)$ for $0 \leq \lambda \in \mathbb{R}$.
- $\partial_\circ(\lambda f)(x) = \lambda \partial_\circ f(x)$ for all λ in $\mathbb{R}$.

Proof:

- Note that

  $$f^\circ(x; -h) = \limsup_{y\to x, t\downarrow 0} \frac{f(y-th)-f(y)}{t},$$

  and

  $$(-f)^\circ(x; h) = \limsup_{y\to x, t\downarrow 0} \frac{-f(y+th)+f(y)}{t} = \limsup_{y-th\to x, t\downarrow 0} \frac{-f((y-th)+th)+f(y-th)}{t}.$$

  Now note that $y \to x$ is the same as $y - th \to x$.

- $$(\lambda f)^\circ(x; h) = \limsup_{y\to x, t\downarrow 0} \frac{\lambda f(y+th)-\lambda f(y)}{t} = \lambda \limsup_{y\to x, t\downarrow 0} \frac{f(y+th)-f(y)}{t} = \lambda f^\circ(x; h)$$

- $$\partial_\circ(\lambda f)(x) = \{\phi : \langle\phi, h\rangle \leq \lambda f^\circ(x; h)\ \forall h \in \mathbb{E}\}$$

6.2.9. Mixed sum rules

Suppose that the real function f is locally Lipschitz around the point x in $\mathbb{E}$ and that the function $g : \mathbb{E} \to (infty, +\infty]$ is convex with $xin\text{int}(\text{dom}\, g)$. Prove:

- $\partial_{\diamondsuit}(f+g)(x) = \nabla f(x) + \partial g(x)$ if f is Gateaux differentiable at x.
- $\partial_{\circ}(f+g)(x) = \nabla f(x) + \partial g(x)$ if f is strictly differentiable at x.

Proof:

- We have

$$(f+g)^{\diamondsuit}(x;h) = \sup_{u\in\mathbb{E}} \limsup_{t\downarrow 0} \frac{(f+g)(x+th+tu) - (f+g)(x+th)}{t}$$

$$= \langle \nabla f(x), h\rangle + \sup_{u\in\mathbb{E}} \limsup_{t\downarrow 0} \frac{g(x+th+tu) - g(x+th)}{t} = \langle \nabla f(x), h\rangle + g'(x;h).$$

Thus $\nabla f(x) + \phi \in \partial_{\diamondsuit}(f+g)(x)$ if and only if $\langle \phi, h\rangle \leq g'(x;h)$.

- We have

$$(f+g)^{\circ}(x;h) = \limsup_{y\to x, x\downarrow 0} \frac{(f+g)(y+th) - (f+g)(y)}{t}$$

$$= \limsup_{y\to x, t\downarrow 0} \frac{f(y+th) - f(y)}{t} + \limsup_{y\to x, t\downarrow 0} \frac{g(y+th) - g(y)}{t}.$$

$$= \langle \nabla f(x), h\rangle + g'(x;h).$$

Thus $\nabla f(x) + \phi \in \partial_{\circ}(f+g)(x)$ if and only if $\langle \phi, h\rangle \leq g'(x;h)$.

6.2.13 Dense Dini subgradients

Suppose the real function f is locally Liptschitz around the point x in $\mathbb{E}$. By considering the closet point in $\text{epi}\, f$ to the point $(x, f(x) - \delta)$ (for a small real $\delta > 0$), prove there are Dini Subgradients at points arbitrary close to x.

Proof:

Lemma: Let $B_{\frac{\sqrt{2}}{2}}$ be the ball of radius $\frac{\sqrt{2}}{2}$ around the origin. Then the function $f : \mathbb{E} \to \mathbb{R}$ with $f(x) = \sqrt{1 - ||x||^2}$ is Lipschitz on $B_{\frac{\sqrt{2}}{2}}$.

Proof of Lemma: Note that

$$|\sqrt{1-||x||^2} - \sqrt{1-||y||^2}| \leq |\sqrt{1-||x||^2} - \sqrt{1-||y||^2}||\sqrt{1-||x||^2} + \sqrt{1-||y||^2}| = |||x||^2 - ||y||^2| \leq 2|||x|| -$$

Note that f has no Dini subgradients at 0.

Now let (y, r) to be the closest point on the epigraph from $(x, f(x) - \delta)$. We claim $r = f(y)$. In fact, suppose that $r > f(y)$ and therefore

$$d^2 = (r - f(x) + \delta)^2 + ||x-y||^2 \leq (f(y) - f(x) + \delta)^2 + ||x-y||^2 \Rightarrow (r - f(y))(r + f(y) + 2\delta - 2f(x)) \leq 0.$$

Thus, $2f(y) + 2\delta - 2f(x) \le r + f(y) + 2\delta - 2f(x) \le 0$. Thus, $f(x) - \delta \ge f(y)$. Note that if $f(y) < f(x) - \delta$, then there exists y' closed enough to y such that $f(y') \le f(x) - \delta$ and also $||x - y'|| < ||x - y||$. So, $(y', f(x) - \delta)$ is closer to $(x, f(x) - \delta)$ than $(y, f(x) - \delta)$. Thus, $f(y) = f(x) - \delta$. After all, $r = f(x) - \delta$ which is a contradiction. Thus, $r = f(y)$ and also $f(y) \ge f(x) - \delta$. Now if $y \ne x$, then choose $||x - y|| > \epsilon > 0$ small enough such that $|f(z) - f(x)| < \delta$ for all $||z - x|| < \epsilon$. Then we have

$$(f(z) - f(x) + \delta)^2 + ||x - z||^2 \ge (f(y) - f(x) + \delta)^2 + ||x - y||^2 \ge (f(y) - f(x) + \delta)^2 + ||x - z||^2.$$

Thus, $f(z) - f(x) + \delta \ge f(y) - f(x) + \delta$, or $f(z) \ge f(x)$. Thus, f is a local minimum of hence $0 \in \partial_- f(x)$. Thus suppose that $x = y$ and $d = \delta$. Hence,

$$(f(y) - f(x) + \delta)^2 + ||x - y||^2 \ge \delta^2 \Rightarrow f(y) - f(x) \ge \sqrt{\delta^2 - ||x - y||^2} - \delta \text{ for } y \text{ close enough to } x.$$

So for y closed enough to x we have

$$f(y) - f(x) \ge \delta[\sqrt{1 - (\frac{||y - x||}{\delta})^2} - 1].$$

But, the RHS has subgradients for points arbitrary close to x.

---

3 Tangent Cones

6.3.1 Exact penalization

For a set $U \subseteq \mathbb{E}$, suppose that the function $f : U \to \mathbb{R}$ has Lipschitz constant L', and that the set $S \subseteq U$ is closed. For any $L > L'$, if the point x minimizes $f + Ld_S$ on U, prove $x \in S$.

Proof: Suppose $x \in U$ is not in S and also $y \in S$ such that $||y - x|| = d_S(x)$. Then we have

$$(f + Ld_S)(x) \le (f + Ld_S)(y) = f(y) \Rightarrow Ld_S(x) \le f(y) - f(x) \le L'||y - x|| < L||y - x||.$$

Thus, $d_S(x) < ||y - x||$. This contradiction completes the proof.

6.3.3 Examples of tangent cones

For the following sets $S \subseteq \mathbb{R}^2$, calculate $T_S(0)$ and $K_S(0)$:

- $\{(x, y) : y \ge x^3\}$.
- $\{(x, y) : x \ge 0, y \ge 0\}$.
- $\{(x, y) : x = 0 \text{ or } y = 0\}$.
- $\{r(\cos\theta, \sin\theta) : 0 \le r \le 1, \frac{\pi}{4} \le \theta \le \frac{7\pi}{4}\}$.

6.3.4 Topology of contingent cone

Prove that the contingent cone is closed, and derive the following topological description: Suppose $x \in S$. The contingent cone $K_S(x)$ consists of those vectors h in $\mathbb{E}$ such that there are sequences $t_r \downarrow 0$ in $\mathbb{R}$ and $h^r \to h$ in $\mathbb{E}$ such that $x + t_r h^r$ lies in S for all r.

Proof: Recall that

$$K_S(x) = \{h : d_S^-(x; h) = 0\}.$$

Since, $x \in S$, x is a local minimum for d_S and thus $0 \leq d_S^-(s; h)$ for all $h \in \mathbb{E}$, so $T_S(x) \subseteq K_S(x)$. However, suppose $d_S^-(x; h) = 0$ and hence $\liminf_{t \downarrow 0} \frac{d_S(x+th)}{t} = 0$. Thus, there exists $t_r \downarrow 0$ such that $d_S(x + t_r h)/t_r \to 0$. Thus, if $x + t_r h^r \in S$ such that $||x + t_r h - x - t_r h^r|| \leq d_S(x + t_r h) + t_r^2$. Thus, $||h - h^r|| \to 0$ and hence $h^r \to h$.

Conversely, suppose that $x + t_r h^r \in S$ and $h^r \to h$ and also $t_r \downarrow 0$. Then wish to show that $h \in K_S(x)$. Note that

$$d_S(x+t_r h) \leq ||x+t_r h-x-t_r h^r|| = t_r ||h-h^r|| \Rightarrow 0 \leq \liminf_{t \downarrow 0} \frac{d_S(x+th)}{t} \leq \lim_{t_r \to +\infty} \frac{d_S(x+t_r h^r)}{t_r} = 0.$$

Thus, $d_S^-(x; h) = 0$ and $h \in K_S(x)$.

6.3.5 Topology of Clarke cone

Suppose that x lies in the set $S \subseteq \mathbb{E}$.

- Prove $d_S^\circ(x; .) \geq 0$.
- Prove
$$d_S^\circ(x; h) = \limsup_{y \to x \text{ in } S, t \downarrow 0} \frac{d_S(y + th)}{t}.$$
- Prove that the Clarke Tangent cone consists of those vectors h in $\mathbb{E}$ such that for any sequence $t_r \downarrow 0$ in $\mathbb{R}$ and $x^r \to x$ in S, there is a sequence $h^r \to h$ such that $x^r + t_r h^r$ lies in S for all r.

Proof:

- Let $h \in \mathbb{E}$, then
$$d_S^\circ(x; h) = \limsup_{y \to x, t \downarrow 0} \frac{d_S(y + th) - d_S(y)}{t} \geq \limsup_{t \downarrow 0} \frac{d_S(x + th)}{t} \geq 0.$$
- Note that
$$d_S^\circ(x; h) = \limsup_{y \to x, t \downarrow 0} \frac{d_S(y + th) - d_S(y)}{t} \geq \limsup_{y \to x \text{ in } S, t \downarrow 0} \frac{d_S(y + th)}{t}.$$
  Now fix some $\epsilon > 0$. Note that $||y'+th-y''|| \leq d_S(y'+th)+\frac{1}{2}\epsilon$ and $||y-y'|| \leq d_S(y)+\frac{1}{2}\epsilon$ for some $y', y'' \in S$. Thus,
$$d_S(y+th) \leq ||y+th-y''|| \leq d_S(y'+th)+d_S(y)+\epsilon \Rightarrow d_S(y'+th)+\epsilon \geq d_S(y+th)-d_S(y).$$

Thus,

$$\limsup_{y\to x, t\downarrow 0} \frac{d_S(y+th)-d_S(y)}{t} \le \limsup_{y'\to x \text{ in } S, t\downarrow 0} \frac{d_S(y'+th)}{t} + \epsilon.$$

Thus,

$$\limsup_{y\to x, t\downarrow 0} \frac{d_S(y+th)-d_S(y)}{t} \le \limsup_{y\to x \text{ in } S, t\downarrow 0} \frac{d_S(y+th)}{t}.$$

This completes the proof.

- Suppose h has the aforementioned properties then

$$d_S^\circ(x;h) = \limsup_{y\to x \text{ in } S, t\downarrow 0} \frac{d_S(y+th)}{t} \le \limsup_{x^r\to x \text{ in } S, t_r\downarrow 0, h^r\to h} \frac{||x^r + t_r h - x^r - t_r h^r||}{t_r} = 0.$$

Thus, $h \in T_S(x)$. Conversely, suppose that $d_S^\circ(x;h) = 0$. Then for every $x^r \to x$ and every $t_r \to 0$, we must have $\lim_{r\to+\infty} d_S(x^r + t_r h)/t_r = 0$. Let $y^r \in S$ such that $||x^r + t_r h - y^r|| \le d_S(x^r + t_r h) + t_r^2$. Suppose $h^r \in \mathbb{E}$ such that $y^r = x^r + t_r h^r$. Then $||h - h^r|| \le d_S(x^r + t_r h)/t_r + t_r$. Thus, $h^r \to h$. Since, $x^r + t_r h^r \in S$, we are done.

6.3.8 Isotonicity

Suppose $x \in U \subseteq V \subseteq \mathbb{E}$. Prove $K_U(x) \subseteq K_V(x)$, but give an example where $T_U(x) \not\subseteq T_V(x)$.

Proof: Recall that

$$K_S(x) = \{h : d_S^-(x;h) = 0\}.$$

Now we want to show that $d_U^-(x;h) = 0$ implies $d_V^-(x;h) = 0$. Note that

$$0 \le d_V^-(x;h) = \liminf_{t\downarrow 0} \frac{d_V(x+th)}{t} \le \liminf_{t\downarrow 0} \frac{d_U(x+th)}{t} = 0.$$

This completes the proof.

Now recall that

$$d_S^\circ(x;h) = \limsup_{y\to x \text{ in } S, t\downarrow 0} \frac{d_S(y+th)}{t}.$$

So we wish to find $U \subseteq V$ and some $h \in \mathbb{E}$ such that

$$\limsup_{y\to x \text{ in } U, t\downarrow 0} \frac{d_U(y+th)}{t} = 0 \text{ but } \limsup_{w\to x \text{ in } V, t\downarrow 0} \frac{d_V(w+th)}{t} \ne 0.$$

Let $U = S^1$ and $V = S^1 \cup \{(x,y) : x \ge 1, y = 0\}$ and also $h = (0,1)$ and $x = (1,0)$. Then we first show that

$$d_U(x;h) = 0.$$

In fact, let $p_i = (x_i, y_i) \to x$ inside U, then $x_i^2 + y_i^2 = 1$ and $x_i \to 1$ and $y_i \to 0$. Note that $d_U(p_i + th) = \sqrt{x_i^2 + (y_i+t)^2} - 1$. Hence,

$$d_U^\circ(x;h) = \limsup_{i\to+\infty, t\to 0} \frac{x_i^2 + (y_i+t)^2 - 1}{t(\sqrt{x_i^2 + (y_i+t)^2} + 1)} = \limsup_{i\to+\infty, t\to 0} \frac{t^2 + 2ty_i}{t(\sqrt{1 + t^2 + 2ty_i} + 1)},$$

which equals to

$$\limsup_{i\to+\infty, t\to 0} \frac{t+2y_i}{\sqrt{1+t^2+2ty_i}+1} = 0.$$

Now let $p_t = (\sqrt{2t+1}, 0)$. Note that $p_t + th = (\sqrt{2t+1}, t)$ with distance $\sqrt{2t+1+t^2}-1 = t$ to U. Hence, $d_V(p_t+th) = t$. Hence,

$$\lim_{t\downarrow 0} \frac{d_V(p_t+th)}{t} = 1 \leq \limsup_{w\to x \text{ in } V, t\downarrow 0} \frac{d_V(w+th)}{t} \neq 0.$$

6.4.3 Local minimizers

Consider a function $f : \mathbb{E} \to [-\infty, +\infty]$ which is finite at the point $x \in \mathbb{E}$.

- If x is local minimizer, prove $0 \in \partial_- f(x)$.
- If $0 \in \partial_- f(x)$, prove for any $\delta > 0$ that x is a strict local minimizer of the function $f(.) - \delta||. - x||$.

Proof:

- We know that $0 \in \partial_- f(x)$ if and only if $f^-(x; h) \geq 0$ for all $h \in \mathbb{E}$. However,

$$f^-(x; h) = \liminf_{t\downarrow 0, h'\to h} \frac{f(x+th') - f(x)}{t} \geq 0 \text{ as } x \text{ is a local minimizer.}$$

- Now suppose that $0 \in \partial_- f(x)$. Then if x is not a strict local minimizer for $f() + \delta||.||$, then there exists $x_i \to x$ such that

$$f(x_i) + \delta||x_i|| \leq f(x).$$

Let $x_i = x + t_i u_i$ where $u_i = \frac{x_i - x}{||x_i - x||}$ and also $t_i = ||x_i - x|| \to 0$. Also, assume $u_i \to u$. Then

$$0 \leq f^-(x; u) = \liminf_{t\downarrow 0, v\to u} \frac{f(x+tv) - f(x)}{t} \leq \liminf_i \frac{f(x+t_i u_i) - f(x)}{t_i} \leq \liminf_i \frac{-\delta||x_i||}{t_i}.$$

If $x = 0$ then $\frac{-\delta||x_i||}{t_i} = -\delta < 0$ which is a contradiction. If $x \neq 0$, then $\liminf_i \frac{-\delta||x_i||}{t_i} = -\infty$, again a contradiction.

6.4.6. Prove a limiting sub differential sum rule for a finite number of lower semi continuous functions, with all but one being locally Lipschitz.

Proof:

Let $f_1, \cdots, f_k$ be lower semicontinuous at x and also g locally Lipschitz around x. Recall the Fuzzy sum rule:

Fuzzy sum rule: Fix $\delta > 0$. Then

$$\partial_-(\sum_{i=1}^k f_i + g)(x) \subseteq \delta B + \sum_{i=1}^k \partial_-(f_i)(U(f_i, x, \delta)) + \partial_-(g)(U(f_i, x, \delta)).$$

Let $\phi^r \in \partial_-(\sum_{i=1}^k f_i + g)$ and also $\phi_i^r \in \partial_-(f_i)(x^r)$ and $\psi^r \in \partial_- g(y^r)$ such that

$$||\phi^r - \sum_{i=1}^k \phi_i^r - \psi^r|| < \frac{1}{r} \ , \ ||x^r - x|| < \frac{1}{r}, ||f_i(x^r) - f_i(x)|| < \frac{1}{r}, ||g(y^r) - g(x)|| < \frac{1}{r}, ||y^r - y|| < \frac{1}{r}$$

Note that for all $\psi \in \partial_- g(x')$ we have $\langle \psi, v \rangle \leq C||v||$. Hence, $||\psi|| \leq C$. So, suppose that $\psi^r \to \psi$. 0

6.4.7 Limiting and Clarke sub differentials
Suppose the real function f is locally Lipschitz around the point x in $\mathbb{E}$.

- Use the fact that the Clarke sub differential is a closed multi-function to show $\partial_a f(x) \subseteq \partial_\circ f(x)$.
- Deduce from the Intrinsic Clarke sub differential theorem the property $\partial_\circ f(x) = \operatorname{conv} \partial_a f(x)$.
- Prove $\partial_a f(x) = \{\phi\}$ if and only if ϕ is the strict derivative of f at x.

Proof:

- Let $\phi \in \partial_a f(x)$, then there exists $\phi_i \in \partial_- f(x^i)$ for some $x^i \to x$ such that $\phi_i \to \phi$. Then $\phi_i \in \partial_\circ f(x^i)$ and thus $\phi \in \partial_a f(x)$ as Clarke sub differentials are closed under limit.
- Now since $\partial_\circ f(x)$ is convex we have $\operatorname{conv} \partial_a f(x) \subseteq \partial_\circ f(x)$. On the other hand,

$$\begin{aligned} \partial_\circ f(x) = \operatorname{conv}\{\lim \nabla f(x^i) : f \text{ is differentiable at } x^i \text{ and also } x^i \to x\} \subseteq \\ \operatorname{conv}\{\lim \phi_i : \phi_i \in \partial_- f(x^i) \text{ and also } x^i \to x\} = \operatorname{conv} \partial_a f(x). \end{aligned}$$

- Note that $\partial_\circ f(x)$ is a singleton if and only if $\partial_a f(x)$ is a singleton.