1 (1.3.3) Subdifferention of Distance functions

- X is a Banach space, $\Omega \subset X$. The distance function $d_\Omega : X \to \mathbb{R}$ is defined by

$$d_\Omega(x) = d(x, \Omega).$$

- $d_\Omega(\cdot)$ is nonsmooth and globally Lipschitz continuous (with modulus $\ell = 1$) on X.
- In what follows, $\widehat{\partial}_\varepsilon := \widehat{\partial}_{a\varepsilon}$.
- We will study subgradients of the function $d_\Omega(\cdot)$ at points in Ω and outside Ω.

Proposition 1 *(1.95 ε-subgradients of distance functions at in-set points)*
Suppose $\Omega \subset X$, $\overline{x} \in \Omega$, $\varepsilon \geq 0$. Then

$$\begin{aligned} \widehat{\partial}_\varepsilon d_\Omega(\overline{x}) &\subset \left\{ x^* \in \widehat{N}_\varepsilon(\overline{x}; \Omega) : \|x^*\| \leq 1 + \varepsilon \right\}, \\ \widehat{\partial}_\varepsilon d_\Omega(\overline{x}) &\supset \left\{ x^* \in \widehat{N}_{\varepsilon/4}(\overline{x}; \Omega) : \|x^*\| \leq 1 + \varepsilon/4 \right\} \end{aligned}$$

Proof. The first part is easy to check. For the second part, let $x^* \in \widehat{N}_{\varepsilon/4}(\overline{x}; \Omega)$ and let $\{x_k\}$ be an arbitrary sequence converging to $\overline{x}$. For each k choose a $u_k \in \Omega$ such that

$$\|x_k - u_k\| \leq d_\Omega(x_k) + \frac{1}{2}\|x_k - \overline{x}\|^2.$$

Then $u_k \to \overline{x}$ and $\frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|} \leq \frac{3}{2}$. Now

$$\begin{aligned} \frac{\langle x^*, x_k - \overline{x}\rangle - (d_\Omega(x_k) - d_\Omega(\overline{x}))}{\|x_k - \overline{x}\|} &= \frac{\langle x^*, x_k - \overline{x}\rangle}{\|x_k - \overline{x}\|} - \frac{d_\Omega(x_k)}{\|x_k - \overline{x}\|} \\ &\leq \frac{\langle x^*, x_k - \overline{x}\rangle}{\|x_k - \overline{x}\|} - \frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|} + \frac{1}{2}\|x_k - \overline{x}\| \\ &= \frac{\langle x^*, u_k - \overline{x}\rangle}{\|u_k - \overline{x}\|}\frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|} + \frac{\langle x^*, x_k - u_k\rangle}{\|u_k - \overline{x}\|} \\ &\quad - \frac{\|x_k - \overline{x}\|}{\|x_k - \overline{x}\|} + \frac{1}{2}\|x_k - \overline{x}\| \\ &\leq \frac{\langle x^*, u_k - \overline{x}\rangle}{\|u_k - \overline{x}\|}\left(1 + \frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|}\right) + \|x^*\|\frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|} \\ &\quad - \frac{\|x_k - \overline{x}\|}{\|x_k - \overline{x}\|} + \frac{1}{2}\|x_k - \overline{x}\| \\ &= \frac{\langle x^*, u_k - \overline{x}\rangle}{\|u_k - \overline{x}\|}\left(1 + \frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|}\right) + (\|x^*\| - 1)\frac{\|x_k - u_k\|}{\|x_k - \overline{x}\|} \\ &\quad + \frac{1}{2}\|x_k - \overline{x}\| \\ &\leq \frac{5}{2}\frac{\langle x^*, u_k - \overline{x}\rangle}{\|u_k - \overline{x}\|} + \frac{3\varepsilon}{8} + \frac{1}{2}\|x_k - \overline{x}\|. \end{aligned}$$

Therefore,

$$\begin{aligned}\overline{\lim_{k\to\infty}}\frac{\langle x^*, x_k-\overline{x}\rangle-(d_\Omega(x_k)-d_\Omega(\overline{x}))}{\|x_k-\overline{x}\|} &\leq \frac{5}{2}\overline{\lim_{k\to\infty}}\frac{\langle x^*, u_k-\overline{x}\rangle}{\|u_k-\overline{x}\|}+\frac{3\varepsilon}{8}\\ &\leq \frac{5}{2}\overline{\lim_{u\xrightarrow{\Omega}\overline{x}}}\frac{\langle x^*, u_k-\overline{x}\rangle}{\|u_k-\overline{x}\|}+\frac{3\varepsilon}{8}\\ &\leq \frac{5\varepsilon}{8}+\frac{3\varepsilon}{8}=\varepsilon.\end{aligned}$$

Since $\{x_k\}$ is an arbitrary sequence converging to $\overline{x}$, it follows that

$$\overline{\lim_{x\to\overline{x}}}\frac{\langle x^*, x_k-\overline{x}\rangle-(d_\Omega(x_k)-d_\Omega(\overline{x}))}{\|x_k-\overline{x}\|}\leq\varepsilon$$

and thus, $x^*\in\widehat{\partial}_\varepsilon d_\Omega(\overline{x})$. ■

Corollary 2 *(1.96 Fréchet subgradients of distance functions at in-set points)*
Suppose $\Omega\subset X$, $\overline{x}\in\Omega$. Then

$$\begin{aligned}\widehat{\partial}d_\Omega(x) &= \widehat{N}(\overline{x};\Omega)\cap\mathbf{B}^*,\\ \widehat{N}(\overline{x};\Omega) &= \bigcup_{\lambda>0}\lambda\widehat{\partial}d_\Omega(x).\end{aligned}$$

Theorem 3 *(1.97 basic normals via subgradients of distance functions at in-set points)*
Suppose $\Omega\subset X$ is nonempty and closed. Then, for any $\overline{x}\in\Omega$,

$$N(\overline{x};\Omega)=\bigcup_{\lambda>0}\lambda\partial d_\Omega(\overline{x}).$$

Corollary 4 *(1.98 regularity of sets and distance functions at in-set points)*
Suppose $\Omega\subset X$ is nonempty and closed and $\overline{x}\in\Omega$. Then Ω is normally regular at $\overline{x}$ if and only if $d_\Omega(\cdot)$ is lower regular at $\overline{x}$.

- To study subgradients at out-of-set points we define the ρ-enlargement $\Omega(\rho)$ of a set Ω relative to $\overline{x}$:
$$\Omega(\rho):=\{x\in X: d_\Omega(x)\leq\rho\},$$
where $\rho:=d_\Omega(\overline{x})$.
- $\Omega(\rho)$ is always closed even if Ω is not.
- $\Omega(\rho)=\Omega+\rho\overline{\mathbf{B}}$ if either Ω is compact in a Banach space or is closed in a finite dimensional space.

Theorem 5 *(1.99 ε-subgradients of distance functions at out-of-set points)*
Suppose $\Omega \subset X$ is nonempty, $\overline{x} \notin \Omega$ and $\varepsilon \geq 0$. Then

$$\begin{aligned}\widehat{\partial}_\varepsilon d_\Omega(\overline{x}) &\subset \left\{x^* \in \widehat{N}_\varepsilon(\overline{x};\Omega(\rho)) : 1-\varepsilon \leq \|x^*\| \leq 1+\varepsilon\right\}, \\ \widehat{\partial}_\varepsilon d_\Omega(\overline{x}) &\supset \left\{x^* \in \widehat{N}_{\varepsilon/4}(\overline{x};\Omega(\rho)) : 1-\varepsilon/4 \leq \|x^*\| \leq 1+\varepsilon/4\right\}.\end{aligned}$$

In particular, for $\varepsilon = 0$,

$$\widehat{\partial} d_\Omega(\overline{x}) = \widehat{N}(\overline{x};\Omega(\rho)) \cap \mathbf{S}^*.$$

- Fof $\Omega = \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \geq 1\}$, $\overline{x} = (0,0)$, $d_\Omega(\overline{x}) = 1 = \rho$ and $\Omega(\rho) = \mathbb{R}^2$, hence, $\widehat{N}(\overline{x};\Omega(\rho)) = \{0\}$. On the other hand, $d_\Omega(x) = 1 - \sqrt{x_1^2 + x_2^2}$. Therefore, $\partial d_\Omega(\overline{x}) = \mathbf{S}$. Therefore, in this example $\partial d_\Omega(\overline{x}) \nsubseteq \widehat{N}(\overline{x};\Omega(\rho)) \cap \mathbf{B}^*$.
- To derive the correct inclusions we need to introduce the definition of "right-sided" subdifferentials.

Definition 6 *(1.100 right-sided subdifferential)*
Suppose $\varphi : X \to \overline{\mathbb{R}}$, $\overline{x} \in X$, $|\varphi(\overline{x})| < \infty$. The right-sided subdifferential of φ at $\overline{x}$ is defined by

$$\partial_{\geq}\varphi(\overline{x}) := \mathop{\overline{\lim}}_{\substack{x \xrightarrow{\varphi^+} \overline{x}, \\ \varepsilon \to 0^+}} \widehat{\partial}_\varepsilon \varphi(\overline{x}),$$

where $x \xrightarrow{\varphi^+} \overline{x}$ means that $x \to \overline{x}$, $\varphi(x) \to \varphi(\overline{x})$ and $\varphi(x) \geq \varphi(\overline{x})$.

- For the above example, $\partial_{\geq}\varphi(\overline{x}) = \phi$.
- If φ attains a local minimum at $\overline{x} \in \Omega$ then

  $$\partial_{\geq}\varphi(\overline{x}) = \partial\varphi(\overline{x})$$

  and thus $0 \in \partial_{\geq}\varphi(\overline{x})$.

Theorem 7 *(1.101 right sided sugradients of distance functions at out-of-set points)*
Suppose $\Omega \subset X$ is nonempty and closed, $\overline{x} \notin \Omega$. Then

(i) *We have*

$$\partial_{\geq} d_\Omega(\overline{x}) \subset \left[\widehat{N}(\overline{x};\Omega(\rho)) \cap \mathbf{B}^*\right].$$

If Ω is SNC at $\overline{x}$ then

$$\partial_{\geq} d_\Omega(\overline{x}) \subset \left[\widehat{N}(\overline{x};\Omega(\rho)) \cap \mathbf{B}^*\right] \setminus \{0\}.$$

(ii) *We always have the equality*

$$\widehat{N}\left(\overline{x};\Omega\left(\rho\right)\right)=\bigcup_{\lambda\geq0}\lambda\partial d_{\Omega}\left(\overline{x}\right).$$

Theorem 8 *(1.103 ε-subgradients of distance functions and ε-normals to perturbed projections)*

Suppose $\Omega\subset X$ is nonempty and closed, $\overline{x}\notin\Omega$. Then for $\varepsilon\in[0,1]$,

$$\widehat{\partial}_{\varepsilon}d_{\Omega}\left(\overline{x}\right)\subset\bigcap_{\eta>0}\bigcup_{w\in\Pi_{\eta}\left(\overline{x};\Omega\right)}\left\{\widehat{N}_{\varepsilon}\left(w;\Omega\right)\cap\left[1-\varepsilon,1+\varepsilon\right]\mathbf{S}^{*}\right\},$$

where

$$\Pi_{\eta}\left(\overline{x};\Omega\right):=\left\{w\in\Omega:\left\|\overline{x}-w\right\|\leq d_{\Omega}\left(\overline{x}\right)+\eta\right\}.$$

Definition 9 *(1.104 well posedness of best approximations)*

Suppose $\Omega\subset X$ is nonempty and $\overline{x}\notin\Omega$. We say that the best approximation problem to Ω from $\overline{x}$ is well posed if either one of the following properties holds:

(a) *Every sequence $x_k\in\Omega$ satisfying*

$$\left\|\overline{x}-x_k\right\|\to d_{\Omega}\left(\overline{x}\right)$$

contains a convergent subsequence.

(b) *For every sequence $x_k\to\overline{x}$ with $\widehat{\partial}_{\varepsilon_k}d_{\Omega}\left(x_k\right)\neq\phi,\varepsilon_k\to0^{+}$ there is a sequence $w_k\in\Pi\left(x_k;\Omega\right)$ that contains a convergent subsequence.*

Theorem 10 *(1.105 projection formulas for basic subgradients of distance functions at out-of-set points)*

Suppose $\Omega\subset X$ is nonempty and closed, $\overline{x}\notin\Omega$. Assume also that the best approximation problem to Ω from $\overline{x}$ is well posed. Then

$$\partial d_{\Omega}\left(\overline{x}\right)\subset\bigcup_{w\in\Pi\left(\overline{x};\Omega\right)}\left\{N\left(w;\Omega\right)\cap\mathbf{B}^{*}\right\}.$$

If Ω is SNC at every point $w\in\Pi\left(\overline{x};\Omega\right)$ then

$$\partial d_{\Omega}\left(\overline{x}\right)\subset\bigcup_{w\in\Pi\left(\overline{x};\Omega\right)}\left\{N\left(w;\Omega\right)\cap\mathbf{B}^{*}\right\}\setminus\left\{0\right\}.$$

Furthermore, if $\dim X<\infty$ then

$$\partial d_{\Omega}\left(\overline{x}\right)\subset\bigcup_{w\in\Pi\left(\overline{x};\Omega\right)}\left\{N\left(w;\Omega\right)\cap\mathbf{S}^{*}\right\}.$$

2 (1.3.4) Subdifferential Calculus in Banach Spaces

- For $\varphi : X \to \overline{\mathbb{R}}$, $|\varphi(\overline{x})| < \infty$,

$$\partial(\lambda\varphi)(\overline{x}) = \begin{cases} \lambda\partial\varphi(\overline{x}), & \lambda \geq 0 \\ \lambda\partial^{+}\varphi(\overline{x}), & \lambda < 0. \end{cases}$$

  Similarly for ∂^{∞} and $\widehat{\partial}$.

Proposition 11 *(1.107 subdifferential sum rules with equalities)*
Suppose $\psi : X \to \overline{\mathbb{R}}$, $|\psi(\overline{x})| < \infty$. Then

(i) *If $\varphi : X \to \overline{\mathbb{R}}$ is Fréchet differentiable at $\overline{x}$ then*

$$\widehat{\partial}(\varphi + \psi)(\overline{x}) = \varphi'(\overline{x}) + \widehat{\partial}\psi(\overline{x}).$$

(ii) *If $\varphi : X \to \overline{\mathbb{R}}$ is strictly differentiable at $\overline{x}$ then*

$$\partial(\varphi + \psi)(\overline{x}) = \varphi'(\overline{x}) + \partial\psi(\overline{x}).$$

Moreover, $\varphi + \psi$ is lower (respectively epigraphical) regular at $\overline{x}$ if and only if ψ has the corresponding property at $\overline{x}$.

(iii) *If $\varphi : X \to \overline{\mathbb{R}}$ is Lipschitz continuous around $\overline{x}$ then*

$$\partial^{\infty}(\varphi + \psi)(\overline{x}) = \partial^{\infty}\psi(\overline{x}).$$

- Let $\varphi : X \times Y \to \overline{\mathbb{R}}$ be a cost function and $G : X \rightrightarrows Y$ be a constraint multifunction.
  - The marginal function $\mu : X \to \overline{\mathbb{R}}$ is defined by

$$\mu(x) := \inf\{\varphi(x, y) : y \in G(x)\}.$$

  - The restriction to the graph of G is the function $v : X \times Y \to \mathbb{R}$ defined by

$$v(x, y) := \varphi(x, y) + \delta((x, y); \mathrm{gr}\, G).$$

  - The argmin function multifunction $M : X \rightrightarrows Y$ is defined by

$$M(x) := \{y \in G(x) : \varphi(x, y) = \mu(x)\}.$$

- Recall that G is closed graph at $\overline{x}$ if for all $x_k \to \overline{x}$ and $y_k \in G(x_k)$ such that $y_k \to y$ we have $y \in G(\overline{x})$.

Theorem 12 *(1.108 subdifferentiation of marginal functions)*
Suppose $\varphi : X \times Y \to \overline{\mathbb{R}}$ is a cost function, $G : X \rightrightarrows Y$ is a constraint multifunction and $\overline{x} \in X$ such that $|\mu(\overline{x})| < \infty$ and $M(\overline{x}) \neq \phi$. Then

(i) *If $\overline{y} \in M(\overline{x})$ and M is inner semicontinuous at $(\overline{x}, \overline{y})$ then*

$$\begin{aligned} \partial \mu(\overline{x}) &\subset \{x^* \in X^* : (x^*, 0) \in \partial v(\overline{x}, \overline{y})\}, \\ \partial^\infty \mu(\overline{x}) &\subset \{x^* \in X^* : (x^*, 0) \in \partial^\infty v(\overline{x}, \overline{y})\}. \end{aligned}$$

(ii) *If M is inner semicompact at $\overline{x}$, G is closed-graph at $\overline{x}$ and φ is lower semicontinuous on $\operatorname{gr} G$ when $x = \overline{x}$ then*

$$\begin{aligned} \partial \mu(\overline{x}) &\subset \left\{ x^* \in X^* : (x^*, 0) \in \bigcup_{y \in M(\overline{x})} \partial v(\overline{x}, y) \right\}, \\ \partial^\infty \mu(\overline{x}) &\subset \left\{ x^* \in X^* : (x^*, 0) \in \bigcup_{y \in M(\overline{x})} \partial^\infty v(\overline{x}, y) \right\}. \end{aligned}$$

Corollary 13 *(1.109 marginal functions with smooth costs)*
Assume that M is inner semicontinuous at $(\overline{x}, \overline{y})$ with $\overline{y} \in M(\overline{x})$ and φ is strictly differentiable at $(\overline{x}, \overline{y})$. Then

$$\begin{aligned} \partial \mu(\overline{x}) &\subset \varphi_x(\overline{x}, \overline{y}) + D_N^* G(\overline{x}, \overline{y}) \left(\varphi_y(\overline{x}, \overline{y})\right), \\ \partial^\infty \mu(\overline{x}) &\subset D_N^* G(\overline{x}, \overline{y})(0). \end{aligned}$$

Theorem 14 *(1.110 subdifferentiation of compositions: equalities)*
Suppose $\varphi : X \times Y \to \overline{\mathbb{R}}$ is a cost function, $g : X \to Y$ is a constraint function, $|\varphi(\overline{x}, \overline{y})| < \infty$ with $\overline{y} = g(\overline{x})$. Assume further that g is Lipschitz continuous around $\overline{x}$. Then

(i) *If either g is strictly differentiable at $\overline{x}$ or $\dim Y < \infty$ then*

$$\begin{aligned} \partial(\varphi \circ g)(\overline{x}) &= \{x^* \in X^* : (x^*, 0) \in \partial v(\overline{x}, g(\overline{x}))\}, \\ \partial^\infty(\varphi \circ g)(\overline{x}) &= \{x^* \in X^* : (x^*, 0) \in \partial^\infty v(\overline{x}, g(\overline{x}))\} \end{aligned}$$

If $\dim Y < \infty$ then $\varphi \circ g$ is lower (resp. epigraphically) regular at $\overline{x}$ if v has the corresponding property at $(\overline{x}, \overline{y})$.

(ii) *Assume that φ is strictly differentiable at $(\overline{x}, \overline{y})$. Then*

$$\begin{aligned} \partial(\varphi \circ g)(\overline{x}) &= \varphi_x(\overline{x}, \overline{y}) + D_M^* g(\overline{x}) \left(\varphi_y(\overline{x}, \overline{y})\right) \\ &= \varphi_x(\overline{x}, \overline{y}) + \partial \left\langle \varphi_y(\overline{x}, \overline{y}), g \right\rangle (\overline{x}). \end{aligned}$$

Moreover, $\varphi \circ g$ is lower regular at $\overline{x}$ if g is M-regular at $\overline{x}$.

Corollary 15 *(1.111 subdifferentiation of products and quotients)*
Suppose $\varphi_i : X \times Y \to \overline{\mathbb{R}}$, $i = 1, 2$ are lipschitz continuous around $\overline{x}$. Then

(i) *We always have*

$$\partial\left(\varphi_1\cdot\varphi_2\right)(\overline{x})=\partial\left(\varphi_1\cdot\varphi_2\left(\overline{x}\right)+\varphi_1\left(\overline{x}\right)\cdot\varphi_2\right)(\overline{x}).$$

Furthermore, if φ_1 *is strictly differentiable at* $\overline{x}$ *then*

$$\partial\left(\varphi_1\cdot\varphi_2\right)(\overline{x})=\varphi_1'\left(\overline{x}\right)\cdot\varphi_2\left(\overline{x}\right)+\partial\left(\varphi_1\left(\overline{x}\right)\cdot\varphi_2\right)(\overline{x}).$$

In the latter case $\varphi_1\cdot\varphi_2$ *is lower regular at* $\overline{x}$ *if and only if the mapping* $x\longmapsto\varphi_1(x)\cdot\varphi_2(x)$ *is lower regular at* $\overline{x}$ *.*

(ii) *Assume* $\varphi_2(\overline{x})\neq 0$*. Then*

$$\partial\left(\varphi_1/\varphi_2\right)(\overline{x})=\frac{\partial\left(\varphi_2\left(\overline{x}\right)\cdot\varphi_1-\varphi_1\left(\overline{x}\right)\cdot\varphi_2\right)(\overline{x})}{\left[\varphi_2\left(\overline{x}\right)\right]^2}.$$

Furthermore, if φ_1 *is strictly differentiable at* $\overline{x}$ *then*

$$\partial\left(\varphi_1/\varphi_2\right)(\overline{x})=\frac{\varphi_2\left(\overline{x}\right)\cdot\varphi_1'\left(\overline{x}\right)+\partial\left(-\varphi_1\left(\overline{x}\right)\cdot\varphi_2\right)(\overline{x})}{\left[\varphi_2\left(\overline{x}\right)\right]^2}.$$

In the latter case φ_1/φ_2 *is lower regular at* $\overline{x}$ *if and only if the mapping* $x\longmapsto\varphi_1(x)/\varphi_2(x)$ *is lower regular at* $\overline{x}$*.*

(iii) *Suppose* $\varphi : X\times Y\to\overline{\mathbb{R}}$ *is lipschitz continuous around* $\overline{x}$ *and* $\varphi(\overline{x})\neq 0$*. Then*

$$\partial\left(1/\varphi\right)(\overline{x})=-\frac{\partial^+\varphi\left(\overline{x}\right)}{\left[\varphi\left(\overline{x}\right)\right]^2}.$$

Moreover, $1/\varphi$ *is lower regular at* $\overline{x}$ *if and only if* φ *is upper regular at* $\overline{x}$*.*