3.4 Computing Subgradients.

{ weak res → compute some subgrad.
strong res. → full characterize subdifferential set.

3.4.1 Positive Scalar Mult.

Thm 3.35. $f: E \to (-\infty, \infty]$. be proper function, $\alpha > 0$, then for $\forall x \in dom(f)$,

$$\partial(\alpha f)(x) = \alpha \partial f(x).$$

proof: $g \in \partial f(x) \Leftrightarrow \forall y \in E,\ f(y) \geq f(x) + \langle g, y-x \rangle$.

That is, $\alpha f(y) \geq \alpha f(x) + \langle \alpha g, y-x \rangle$. So $\alpha g \in \partial(\alpha f)(x)$.

$\uparrow$
$\alpha > 0$.

Since $dom(\alpha f) = dom(f)$, $\partial(\alpha f)(x) = \alpha \partial f(x)$.

3.4.2 Summation.

Thm 3.36.

Let $f_1, f_2 : E \to (-\infty, \infty]$ be proper convex functions, and let $x \in dom(f_1) \cap dom(f_2)$.

(a) The following inclusion holds:

$$\partial f_1(x) + \partial f_2(x) \subseteq \partial(f_1 + f_2)(x).$$

Proof of (a):
Only need def of subgrad.

(b). If $x \in int(dom(f_1)) \cap int(dom(f_2))$, then

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x).$$

proof of (b):
utilizes the max formula.

proof:

(a). Let $g \in \partial f_1(x) + \partial f_2(x)$, then $\exists\, g_1 \in \partial f_1(x),\ g_2 \in \partial f_2(x)$ s.t. $g = g_1 + g_2$

$\Rightarrow \forall y \in E,\quad f_1(y) \geq f_1(x) + \langle g_1, y-x \rangle$

$f_2(y) \geq f_2(x) + \langle g_2, y-x \rangle$.

$\Rightarrow \quad (f_1+f_2)(y) \geq (f_1+f_2)(x) + \langle g, y-x \rangle \quad \Rightarrow g \in \partial(f_1+f_2)(x)$.

(b). Define $f = f_1 + f_2$. $d \in E$.

$x \in \operatorname{int}(\operatorname{dom}(f)) = \operatorname{int}(\operatorname{dom}(f_1)) \cap \operatorname{int}(\operatorname{dom}(f_2))$

By Max formula, $f'(x;d) = \max\{\langle g, d\rangle : g \in \partial f(x)\} = \sigma_{\partial f(x)}(d)$,

$$\sigma_{\partial f(x)}(d) = f'(x;d) = f_1'(x;d) + f_2'(x;d)$$

$$= \max\{\langle g_1, d\rangle : g_1 \in \partial f_1(x)\} + \max\{\langle g_2, d\rangle : g_2 \in \partial f_2(x)\}.$$

$$= \max\{\langle g_1+g_2, d\rangle : \underline{g_1 \in \partial f_1(x),\ g_2 \in \partial f_2(x)}\}$$

$$= \sigma_{\partial f_1(x) + \partial f_2(x)}(d)$$

$\Downarrow$

$g_1 + g_2 \in \partial f_1(x) + \partial f_2(x)$.

$\langle g_1+g_2, d\rangle$ 在 g_1+g_2 fixed 时也fixed,

故不需这些条件.

By Lemma 2.34 (If A, B closed & convex, then $\sigma_A = \sigma_B \Leftrightarrow A = B$).

$\partial f(x)$, $\underline{\partial f_1(x),\ \partial f_2(x)}$ are compact, convex

$\Downarrow$ $\partial f_1(x) + \partial f_2(x)$ is compact, convex.

$\Rightarrow \partial f(x) = \partial f_1(x) + \partial f_2(x)$

Remark: part (a) does not require convexity assumption on f_1, f_2.

Collary 3.38.

$f_1, f_2, \cdots f_m : E \to (-\infty, \infty]$ be proper convex functions. let $x \in \bigcap_{i=1}^{m} \mathrm{dom}(f_i)$. Then.

(a) $\sum_{i=1}^{m} \partial f_i(x) \subseteq \partial\left(\sum_{i=1}^{m} f_i\right)(x)$.

(b) If $x \in \bigcap_{i=1}^{m} \mathrm{int}(\mathrm{dom}(f_i))$, then $\partial\left(\sum_{i=1}^{m} f_i\right)(x) = \sum_{i=1}^{m} \partial f_i(x)$.

Collary 3.39.

[$f_1 \sim f_m$ are real-valued.]

$f_1, f_2, \cdots f_m : E \to \mathbb{R}$, real-valued convex functions. Then for $\forall x \in E$, $\partial\left(\sum_{i=1}^{m} f_i\right)(x) = \sum_{i=1}^{m} \partial f_i(x)$.

Thm 3.40. 更強假設的 Thm 3.38. $\bigcap_{i=1}^{m} \mathrm{ri}(\mathrm{dom}(f_i)) \neq \phi$.

Let $f_1, f_2, \cdots f_m : E \to (-\infty, \infty]$ be proper convex functions, and assume that $\bigcap_{i=1}^{m} \mathrm{ri}(\mathrm{dom}(f_i)) \neq \phi$. Then for any $x \in E$,

$$\partial\left(\sum_{i=1}^{m} f_i\right)(x) = \sum_{i=1}^{m} \partial f_i(x).$$

[108, Thm 23.8].

Eg: l_1-norm.

$f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i|$.

Then $f = \sum_{i=1}^{n} f_i$, where $f_i(x) = |x_i|$. We have

$(f_i : \mathbb{R}^n \to \mathbb{R})$ 拆开.

$$\partial f_i(x) = \begin{cases} \{\operatorname{sgn}(x_i) e_i\}, & x_i \neq 0 \\ [-e_i, e_i], & x_i = 0 \end{cases}$$

$$\Rightarrow \partial f(x) = \sum_{i=1}^{n} \partial f_i(x) = \sum_{i \in I_{\neq}(x)} \operatorname{sgn}(x_i) e_i + \sum_{i \in I_0(x)} [-e_i, e_i].$$

$I_{\neq}(x) = \{i : x_i \neq 0\}$.

$I_0(x) = \{i : x_i = 0\}$.

Hence,

$$\partial f(x) = \{z \in \mathbb{R}^n : \quad z_i = \operatorname{sgn}(x_i), \ i \in I_{\neq}(x), \ |z_j| \leq 1, \ j \in I_0(x)\}.$$

3.4.3 Affine Transformation

$$h'(x; d) = \lim_{\alpha \to 0^+} \frac{h(x + \alpha d) - h(x)}{\alpha}$$