Part 1

2:13 PM

Proposition 17.5.

$[\![\ f:\mathcal{H}\to\,]-\infty,+\infty],\ \text{proper, convex};\ x\in\operatorname{dom} f\ ]\!]$

$x\in\operatorname{Argmin} f \Leftrightarrow f'(x;\cdot)\geq 0$

Proof:

$(\Rightarrow)$

$x\in\operatorname{Argmin} f \Rightarrow \forall_{y\in\mathcal{H}}\ \forall_{\alpha\in\mathbb{R}_{++}}\quad f(x)\leq f(x+\alpha y)$

$\Leftrightarrow \forall_{y\in\mathcal{H}}\ \forall_{\alpha\in\mathbb{R}_{++}}\quad \dfrac{f(x+\alpha y)-f(x)}{\alpha}\geq 0$

$\Leftrightarrow \forall_{y\in\mathcal{H}}\quad \underbrace{\inf_{\alpha\in\mathbb{R}_{++}} \dfrac{f(x+\alpha y)-f(x)}{\alpha}}_{f'(x;y)}\geq 0$

/* Proposition 17.2. (Properties of directional derivative)

$[\ f:\mathcal{H}\to\,]-\infty,+\infty],\ \text{proper, convex};\ x\in\operatorname{dom} f;\ y\in\mathcal{H}\ ]\Rightarrow$

(i) $\phi:\mathbb{R}_{++}\to\,]-\infty,+\infty]: \alpha\mapsto (f(x+\alpha y)-f(x))/\alpha$: increasing

(ii) $f'(x;y)$ exists in $[-\infty,+\infty]$ and $f'(x;y)=\inf_{\alpha\in\mathbb{R}_{++}}\dfrac{f(x+\alpha y)-f(x)}{\alpha}$

(iii) $f'(x;y-x)+f(x)\leq f(y)$ // wow a subgradient like property!

(iv) $f'(x;\cdot)$: sublinear; $f'(x;0)=0$.

(v) $f'(x;\cdot)$: proper, convex; $\operatorname{dom} f'(x;\cdot)=\operatorname{cone}(\operatorname{dom} f - x)$

(vi) $x\in\operatorname{core}\operatorname{dom} f\Rightarrow f'(x;\cdot)$: real valued, sublinear */

$\Leftrightarrow f'(x;\cdot)\geq 0$

$\therefore\ x\in\operatorname{Argmin} f \Rightarrow f'(x;\cdot)\geq 0$

$(\Leftarrow)\quad f'(x;y-x)+f(x)\leq f(y)\quad \forall_{y\in\mathcal{H}}\ \forall_{x\in\operatorname{dom} f}\qquad (1)$

$f'(x;\cdot)\geq 0 \Leftrightarrow \forall_{\tilde{y}\in\mathcal{H}}\quad f'(x;\tilde{y})\geq 0$

$\overset{\tilde{y}:=y-x}{\Rightarrow}\quad f'(x;y-x)\geq 0$

$\Leftrightarrow\quad f(x)+f'(x;y-x)\geq f(x)$

$\overset{\text{from}(1)}{\Rightarrow}\quad f(y)\geq f(x)+f'(x;y-x)\geq f(x)$

$\therefore\ \forall_{y\in\mathcal{H}}\quad f(y)\geq f(x) \Leftrightarrow x\in\operatorname{Argmin} f$

So, $f'(x;\cdot)\geq 0 \Rightarrow x\in\operatorname{Argmin} f$

■

Corollary 17.6.

$[\ f\in\Gamma_0(\mathcal{H});\ x,p\in\mathcal{H};\ \operatorname{dom} f: \text{open};\ f:$ Gateaux differentiable on $\operatorname{dom} f\ ]\!]$

$p=\operatorname{Prox}_f x \Leftrightarrow \nabla f(p)+p-x=0$

Proof:

set $\ g: y\mapsto f(y)+\frac{1}{2}\|x-y\|^2 \Rightarrow \forall_{y\in\operatorname{dom} f}\quad \nabla g(y)=\nabla f(y)-(y-x)\qquad \cdots(2)$

So, g: convex, Gateaux differentiable on $\operatorname{dom} f$;

$p=\operatorname{Prox}_f x=(\mathrm{Id}+\partial f)^{-1}x\quad // \because\ \operatorname{Prox}_f=(\mathrm{Id}+\partial f)^{-1}$

$\Leftrightarrow (\mathrm{Id}+\partial f)p\ni x$

$\Leftrightarrow p+\partial f(p)\ni x\quad // \partial f(p)=\nabla f(p)$ and " $\ni$ " becomes " $=$ "

$\Leftrightarrow p+\nabla f(p)=x$

$\therefore\ \nabla f(p)+p-x=0$

■

Proposition 17.18. $[\![\ f:\mathcal{H}\to\,]-\infty,+\infty],\ \text{proper, convex};\ x\in\operatorname{dom} f\ ]\!]\Rightarrow (f'(x;\cdot))^*=\iota_{\partial f(x)}$

Proof:

Define: $\ \phi:\mathcal{H}\to\,]-\infty,+\infty]: y\mapsto f'(x;y)=\lim_{\alpha\downarrow 0}\dfrac{f(x+\alpha y)-f(x)}{\alpha}$

So. $\underbrace{\phi^*(u)}_{(f'(x;y))^*}\overset{\text{def}}{=}\sup_{y\in\mathcal{H}}\big(-\underbrace{\phi(y)}_{f'(x;y)}+\langle y|u\rangle\big)$

/* Proposition 17.2. (Properties of directional derivative)

$[\ f:\mathcal{H}\to\,]-\infty,+\infty],\ \text{proper, convex};\ x\in\operatorname{dom} f;\ y\in\mathcal{H}\ ]\Rightarrow$

(i) $\phi:\mathbb{R}_{++}\to\,]-\infty,+\infty]: \alpha\mapsto (f(x+\alpha y)-f(x))/\alpha$: increasing

(ii) $f'(x;y)$ exists in $[-\infty,+\infty]$ and $f'(x;y)=\inf_{\alpha\in\mathbb{R}_{++}}\dfrac{f(x+\alpha y)-f(x)}{\alpha}$ */

$=\sup_{y\in\mathcal{H}}\Big(\underbrace{-\inf_{\alpha\in\mathbb{R}_{++}}\Big(\dfrac{f(x+\alpha y)-f(x)}{\alpha}\Big)}_{\left(-\sup_{\alpha\in\mathbb{R}_{++}}\frac{-f(x+\alpha y)+f(x)}{\alpha}\right)}+\langle y|u\rangle\Big)$

$$\left(-\sup_{\alpha\in\mathbb{R}_{++}} \frac{-f(x+\alpha y)+f(x)}{\alpha}\right)$$

$$= \sup_{y\in\mathcal{H}}\left(\sup_{\alpha\in\mathbb{R}_{++}}\left(\frac{f(x)-f(x+\alpha y)}{\alpha}\right)+\langle y|u\rangle\right)$$

$\langle \alpha y|u\rangle = \langle x+\alpha y - x|u\rangle = \langle x+\alpha y|u\rangle - \langle x|u\rangle$

$$= \sup_{y\in\mathcal{H}}\ \sup_{\alpha\in\mathbb{R}_{++}}\left(\frac{f(x)-f(x+\alpha y)+\alpha\langle y|u\rangle}{\alpha}\right)$$

$$= \sup_{y\in\mathcal{H}}\ \sup_{\alpha\in\mathbb{R}_{++}}\left(\frac{f(x)-f(x+\alpha y)+\langle x+\alpha y|u\rangle-\langle x|u\rangle}{\alpha}\right)$$ // now

Results 18.4

• [term ∈ {sup, inf, $\overline{\lim}$, $\underline{\lim}$, $\underbrace{\lim}_{\text{if exists}}$}; $f: \mathcal{X}\times\mathcal{Y}\to\mathbb{R}$] $\text{term}_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} f(x,y) \geq \sup_{y\in\mathcal{Y}} \text{term}_{x\in\mathcal{X}} f(x,y)$

$\text{term}_{x\in\mathcal{X}} \inf_{y\in\mathcal{Y}} f(x,y) \leq \inf_{y\in\mathcal{Y}} \text{term}_{x\in\mathcal{X}} f(x,y)$; $\sup_{(x,y)\in\mathcal{X}\times\mathcal{Y}} f(x,y) = \sup_{x\in\mathcal{X}}\sup_{y\in\mathcal{Y}} f(x,y) = \sup_{y\in\mathcal{Y}}\sup_{x\in\mathcal{X}} f(x,y)$

$\inf_{(x,y)\in\mathcal{X}\times\mathcal{Y}} f(x,y) = \inf_{x\in\mathcal{X}}\inf_{y\in\mathcal{Y}} f(x,y) = \inf_{y\in\mathcal{Y}}\inf_{x\in\mathcal{X}} f(x,y)$

$$= \sup_{\alpha\in\mathbb{R}_{++}}\ \sup_{y\in\mathcal{H}} \frac{1}{\alpha}\left(f(x)-f(x+\alpha y)+\langle x+\alpha y|u\rangle-\langle x|u\rangle\right)$$

$$= \sup_{\alpha\in\mathbb{R}_{++}}\left[\frac{1}{\alpha}\sup_{y\in\mathcal{H}}\left(\langle x+\alpha y|u\rangle - f(x+\alpha y) + f(x) - \langle x|u\rangle\right)\right]$$

/* $\sup_{y\in\mathcal{H}}\left(\langle \underbrace{x+\alpha y}_{z}|u\rangle - f(x+\alpha y)\right) + f(x)-\langle x|u\rangle = \sup_{\substack{y=\frac{z-x}{\alpha}\in\mathcal{H}\\ \Leftrightarrow\\ z\in\mathcal{H}}}\left(\langle z|u\rangle - f(z)\right) + f(x) - \langle x|u\rangle = f^*(u)+f(x)-\langle x|u\rangle$ */

//then $z = x+\alpha y \Leftrightarrow y=(z-x)/\alpha$
new decision variable

$$= \sup_{\alpha\in\mathbb{R}_{++}} \frac{f(x)+f^*(u)-\langle x|u\rangle}{\alpha} \quad \ldots \quad (17.20)$$

/* Proposition 16.9. $u\in\partial f(x) \Leftrightarrow f(x)+\overset{*}{f}(u)=\langle x|u\rangle$

Fenchel-Young inequality:

* Proposition 13.13. (Fenchel-Young inequality)
[$f:\mathcal{H}\to]-\infty,+\infty]$, proper] /* wow! see how general the function is, infact any sensible function would satisfy this! */
$\forall_{x\in\mathcal{H}}\ \forall_{u\in\mathcal{H}}\quad f(x)+f^*(u)\geq\langle x|u\rangle$ .. (2)

*/

there are two possibilities: $u\in\partial f(x)$ or $u\notin\partial f(x)$

if $u\in\partial f(x) \Leftrightarrow f(x)+f^*(u)-\langle x|u\rangle=0$

$\Leftrightarrow \phi^*(u)=0$

if $u\notin\partial f(x) \Leftrightarrow f(x)+f^*(u)-\langle x|u\rangle>0$

$$\Leftrightarrow \phi^*(u)=\sup_{\alpha\in\mathbb{R}_{++}}\frac{f(x)+f^*(u)-\langle x|u\rangle}{\alpha}=+\infty$$

So, $$\phi^*(u)=\left(f'(x;u)\right)^*=\begin{cases}0, & \text{if } u\in\partial f(x)\\ +\infty, & \text{if } u\notin\partial f(x)\end{cases} = \iota_{\partial f(x)}$$

So, $\left(f'(x;u)\right)^*=\iota_{\partial f(x)}$ ■

Proposition 17.22. (Steepest descent direction)
[$f:\mathcal{H}\to]-\infty,+\infty]$, proper, convex ; $x\in\text{cont} f\setminus\text{Argmin} f$; $u\in P_{\partial f(x)}(0)$; $z=-\frac{u}{\|u\|}$]
z: unique minimizer of $f'(x;\cdot)$ over $B(0;1)$

Proof:

/* recall:

Proposition 16.14. [$f:\mathcal{H}\to]-\infty,+\infty]$, proper, convex ; $x\in\text{dom} f$] $\Rightarrow$
(i) [int dom $f\neq\emptyset$, $x\in$ bdry dom f] $\Rightarrow$ $\partial f(x)$: empty or unbounded
(ii) $x\in\text{cont} f \Rightarrow \partial f(x)$: nonempty, weakly compact ✓ (1)
(iii) $x\in\text{cont} f\Rightarrow \exists_{\delta\in\mathbb{R}_{++}}\ \partial f(B(x;\delta))$: bounded
(iv) cont $f\neq\emptyset \Rightarrow$ int dom $f\subseteq$ dom ∂f

also given, $x\in\text{cont} f\subseteq\text{dom} f$; f: proper

Proposition 16.3. [$f\in\mathcal{H}\to]-\infty,+\infty]$, proper, $x\in\text{dom} f$] $\Rightarrow$
(i) dom $\partial f\subseteq$ dom f
(ii) $\partial f(x)=\bigcap_{y\in\text{dom} f}\{u\in\mathcal{H}\mid\langle y-x|u\rangle\leq f(y)-f(x)\}$
✓ (iii) $\partial f(x)$: closed, convex (2)
(iv) $x\in\text{dom}\,\partial f\Rightarrow f$: lower semicontinuous at x

Theorem 16.2. (Fermat's rule)
[$f:\mathcal{H}\to]-\infty,+\infty]$, proper] $\text{Argmin} f=\text{zer}\,\partial f=\{x\in\mathcal{H}\mid 0\in\partial f(x)\}$ */

$x\notin\text{Argmin} f\Rightarrow\partial f(x)\not\ni 0$ (3)

So, from (1), (2), (3) we have:

$\partial f(x)$: nonempty, closed, convex, weakly compact, $\not\ni 0$

So, $P_{\partial f(x)} 0 = u \neq 0$ (4)

/*

Theorem 3.14.

/* Charecterization of projection on closed convex nonempty set */ Theorem 3.14. ***

(C: nonempty closed convex subset of $\mathcal{H}$) $\Rightarrow$ $\begin{cases} \bullet\ C\text{: Chebyshev set, i.e., every point in } \mathcal{H} \text{ has exactly one projection on } C \\ \bullet\ \forall_{x \in \mathcal{H}} \left(p = P_C(x) \Leftrightarrow (p \in C, \forall_{y \in C} \langle y-p \mid x-p \rangle \leq 0) \right) \end{cases}$

*/

$u \in \partial f(x)$

$$\forall_{y \in \partial f(x)} \quad \langle y-u \mid 0-u \rangle = \langle y-u \mid -u \rangle = -\langle y \mid u \rangle + \|u\|^2 \leq 0$$

$\Rightarrow \max_{y \in \partial f(x)} -\langle y \mid u \rangle + \|u\|^2 \leq 0$ // But for $y = u \in \partial f(x)$ the max value $= 0$

$\therefore \max_{y \in \partial f(x)} -\langle y \mid u \rangle + \|u\|^2 = 0 \Leftrightarrow \max \langle -u \mid \partial f(x) \rangle = -\|u\|^2$

/* Theorem 17.19 (Max formula)

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex; $x \in \operatorname{cont} f$]

$f'(x; \cdot) = \max \langle \cdot \mid \partial f(x) \rangle$ */

So, we have:

$f'(x; z) = \max \langle z \mid \partial f(x) \rangle$ // given $z = -\frac{u}{\|u\|}$

$= \max \langle -\frac{u}{\|u\|} \mid \partial f(x) \rangle = -\frac{1}{\|u\|} \max \langle u \mid \partial f(x) \rangle = -\frac{1}{\|u\|} \|u\|^2 = -\|u\|$ (4)

and $\forall_{y \in B(0;1)}$ $f'(x; y) = \max \langle y \mid \partial f(x) \rangle \geq \langle y \mid u \rangle$ // as $u \in \partial f(x)$

$\geq -\|y\| \, \|u\|$ /* * Cauchy Schwarz inequality: $\underbrace{|\langle x \mid y \rangle|}_{\max\{\langle x|y \rangle, -\langle x|y \rangle\}} \leq \|x\| \, \|y\|$

Flavor 1: $\langle x \mid y \rangle \leq \|x\| \, \|y\|$

Flavor 2: $-\langle x \mid y \rangle \leq \|x\| \, \|y\| \Leftrightarrow \langle x \mid y \rangle \geq -\|x\| \, \|y\|$ */

/* as $y \in B(0;1)$

$\forall_{y \in B(0;1)}$ $0 \leq \|y\| \leq 1$

$\Rightarrow -1 \leq -\|y\| \leq 0$ */

$\geq (-1) \|u\| = -\|u\| = f'(x; z)$ // from (4)

So, $\forall_{y \in B(0;1)}$ $f'(x; y) \geq f'(x; z)$ // But $z = -\frac{u}{\|u\|} \in B(0;1)$

So, $\min_{y \in B(0;1)} f'(x; y) = f'(x; z) = -\|u\|$ i.e., $z = \operatorname{argmin}_{y \in B(0;1)} f'(x; y)$

also, by construction $z = -\frac{u}{\|u\|}$ is the unique point which can produce minimum objective value $-\|u\|$

So, it is also the unique minimizer

■

Proposition 17.26.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex

$x \in \operatorname{dom} f$] $\Rightarrow$

(i) f: Gateaux differentiable at $x \Rightarrow \partial f(x) = \{\nabla f(x)\}$ — the Gateaux derivative

(ii) [$x \in \operatorname{cont} f$, $\partial f(x) = \{u\}$] $\Rightarrow$ (f: Gateaux differentiable at x ; $u = \nabla f(x)$)

Proof:

(i)

Proposition 17.9. [$x \in \mathcal{H}$; $f: \mathcal{H} \to]-\infty, +\infty]$, convex, Gateaux differentiable at x]

$\forall_{y \in \mathcal{H}}$ $\langle y - x \mid \nabla f(x) \rangle + f(x) \leq f(y)$

Definition 16.1. [$f: \mathcal{H} \to]-\infty, +\infty]$, proper] (Definition of subdifferential)

∂f : subdifferential of $f \overset{\text{def}}{\Leftrightarrow} \partial f: \mathcal{H} \to 2^{\mathcal{H}}: x \mapsto \{u \in \mathcal{H} \mid \forall_{y \in \mathcal{H}} \langle y - x \mid u \rangle + f(x) \leq f(y)\}$

u: subgradient

∂f : subdifferential of $f \overset{\text{def}}{\leftrightarrow} \partial f : \mathcal{H} \to 2^{\mathcal{H}} : x \mapsto \{\underbrace{u}_{u\text{: subgradient}} \in \mathcal{H} \mid \forall_{y \in \mathcal{H}} \ \langle y-x \mid u \rangle + f(x) \le f(y)\}$

$\nabla f(x) \in \partial f(x)$, now we show $\nabla f(x)$ is the only element

now suppose, $u \in \partial f(x)$ such that $u \neq \nabla f(x)$

Definition 17.1. (Directional derivative)

$[\![f : \mathcal{H} \to]-\infty, +\infty]$, proper; $x \in \operatorname{dom} f$; $y \in \mathcal{H}]\!]$

$f'(x;y)$: directional derivative of f at x in the direction of $y \overset{\text{def}}{\leftrightarrow}$

$f'(x;y) = \lim_{\alpha \downarrow 0} \dfrac{f(x+\alpha y) - f(x)}{\alpha}$ {provided that the limit exists in $[-\infty, +\infty]$}

Proposition 17.17. $[\![f : \mathcal{H} \to]-\infty, +\infty]$, proper, convex; $x \in \operatorname{dom} \partial f$; $u \in \mathcal{H}]\!] \Rightarrow$

(i) $u \in \partial f(x) \Leftrightarrow \langle \cdot \mid u \rangle \le f'(x;\cdot)$

(ii) $f'(x;\cdot)$: proper, sublinear

$u \in \partial f(x) \leftrightarrow \langle \cdot \mid u \rangle \le f'(x;\cdot)$

$\cdot := u - \nabla f(x)$

$\Rightarrow$

$\langle u - \nabla f(x) \mid u \rangle \le f'(x; u - \nabla f(x)) = \langle u - \nabla f(x) \mid \nabla f(x) \rangle$

Definition 17.2.1. (Gateaux gradient of a function at a point)

$[\![x \in \operatorname{dom} f$; f: real valued, $\mathcal{H} \to]-\infty, +\infty]$; $f'(x;\cdot)$: linear and continuous on $\mathcal{H}]\!]$

- f : Gateaux differentiable at x
- /* using Riesz-Frechet representation */

$\exists! \underbrace{\nabla f(x)}_{\text{Gateaux gradient of } f \text{ at } x} \in \mathcal{H} \ \forall_{y \in \mathcal{H}} \ f'(x;y) = \langle y \mid \nabla f(x) \rangle = \lim_{\alpha \to 0} \dfrac{f(x+\alpha y) - f(x)}{\alpha}$

unique

$\therefore \langle u - \nabla f(x) \mid u \rangle \le \langle u - \nabla f(x) \mid \nabla f(x) \rangle$

$\Rightarrow \langle u - \nabla f(x) \mid u - \nabla f(x) \rangle = \| u - \nabla f(x) \|^2 \le 0$

$\Rightarrow \| u - \nabla f(x) \| = 0 \leftrightarrow u = \nabla f(x) \Rightarrow$ contradiction

$\therefore \nabla f(x)$ is the only element in $\partial f(x)$

$\leftrightarrow \partial f(x) = \{ \nabla f(x) \}$ ☺

(ii)

/*

Theorem 17.19. (Max formula)

$[\![f : \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \operatorname{cont} f]\!]$

$f'(x;\cdot) = \max \langle \cdot \mid \partial f(x) \rangle$

*/

given: $x \in \operatorname{cont} f$

$\partial f(x) = \{u\}$

$f'(x;y) = \max \langle y \mid \underbrace{\partial f(x)}_{= \{u\}} \rangle$

$= \max \langle y \mid u \rangle = \langle y \mid u \rangle$

But, $f'(x;y) = \langle y \mid \underbrace{\nabla f(x)}_{\text{unique Gateaux gradient}} \rangle$

$\Rightarrow u = \nabla f(x)$ ∎

Part 2

1:30 PM

Proposition 17.41 [[$f : \in \Gamma_0(\mathcal{H})$, Gateaux differentiable at $x \in \operatorname{dom} f$]] $\Rightarrow$
$x \in \operatorname{int} \operatorname{dom} f \wedge f$: continuous on $\operatorname{int} \operatorname{dom} f$

Proof: $f \in \Gamma_0(\mathcal{H}) \Rightarrow f$: lower semicontinuous

*Corollary 8.30. *

[$f : \mathcal{H} \to]-\infty, +\infty]$, proper, convex,
one of the following holds
- f: bounded above on some neighborhood
- ☑ f: lower semicontinuous
- $\mathcal{H}$: finite-dimensional] $\Rightarrow$ $\operatorname{cont} f = \operatorname{int} \operatorname{dom} f$ ☑ /* $\operatorname{cont} f$: domain of continuity of a function f */

So, $\operatorname{cont} f = \operatorname{int} \operatorname{dom} f$ (1)

define $g : \mathcal{H} \to [0, \infty] : x \mapsto \begin{cases} 0, & \text{if } x = 0 \\ +\infty, & \text{else} \end{cases}$ $\Rightarrow g \in \Gamma_0(\mathcal{H})$ by construction $\Rightarrow \operatorname{dom} g = \{0\}$

* Fact 9.16.

[[$f, g \in \Gamma_0(\mathcal{H})$]] $\operatorname{int}(\operatorname{dom} f - \underbrace{\operatorname{dom} g}_{\{0\}}) = \operatorname{core}(\operatorname{dom} f - \underbrace{\operatorname{dom} g}_{=\{0\}})$ $\Rightarrow \operatorname{int} \operatorname{dom} f = \operatorname{core} \operatorname{dom} f$ (2)

From (1), (2) $\Rightarrow$ $\operatorname{cont} f = \operatorname{int} \operatorname{dom} f = \operatorname{core} \operatorname{dom} f$ (3)

$\therefore f$: continuous on $\operatorname{int} \operatorname{dom} f$ ✓

also. f: Gateaux differentiable and $x \in \operatorname{dom} f$

$\overset{?}{\Rightarrow}$ $x \in \operatorname{core} \operatorname{dom} f$

from (3): $x \in \operatorname{cont} f = \operatorname{int} \operatorname{dom} f$ ☒