Chapter 1: part 1

Monday, November 9, 2015 9:45 PM

• Net

$(A,\preceq)$: directed set $\overset{def}{=} \left(\forall_{a\in A}\ a\preceq a,\ \forall_{a,b,c\in A}\ (a\preceq b \wedge b\preceq c) \Rightarrow a\preceq c,\ \forall_{a,b\in A}\ \exists_{c\in A}\ (a\preceq c \wedge b\preceq c)\right)$

$(x_a)_{a\in A}$: net in $\mathcal{X}$ $\overset{def}{\longleftrightarrow}$ operator from A to $\mathcal{X}$ (nonempty set)

$(y_b)_{b\in B}$: subnet of $(x_a)_{a\in A}$ $\overset{def}{\longleftrightarrow}$ $\exists_{k:B\to A}$ ($\cdot\ \forall_{b\in B}\ y_b = x_{k(b)}$ $\cdot\ \forall_{a\in A}\ \exists_{d\in B}\ \forall_{b\in B}\ b\succeq d \Rightarrow k(b)\succeq a$)

(k: some operator (not set-valued))

• Limit inferior of a net $(\xi_a)_{a\in A}$ in $[-\infty,\infty]$ is:

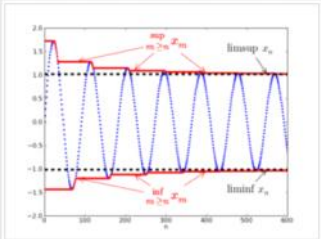
$$\underline{\lim}\ \xi_a = \sup_{c\in A}\ \inf_{b\in A:\, b\succeq c}\ \xi_b$$

Limit superior of a net $(\xi_a)_{a\in A}$ in $[-\infty,\infty]$ is:

$$\overline{\lim}\ \xi_a = \inf_{c\in A}\ \sup_{b\in A:\, b\succeq c}\ \xi_b$$

$\underline{\lim}\ \xi_a \leq \overline{\lim}\ \xi_a$

Limit inferior and limit superior always exists.



An illustration of limit superior and limit inferior. The sequence x_n is shown in blue. The two red curves approach the limit superior and limit inferior of x_n, shown as dashed black lines. In this case, the sequence *accumulates* around the two limits. The superior limit is the larger of the two, and the inferior limit is the smaller of the two. The inferior and superior limits agree if and only if the sequence is convergent (i.e., when there is a single limit).

6/9/2016 11:28 AM

* Lemma 1.6.

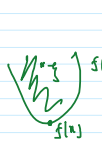
[[$(f_i)_{i\in I}$: family of functions from $\mathcal{X}$ to $[-\infty,+\infty]$]]

(i) $\mathrm{epi}\left(\sup_{i\in I} f_i\right) = \bigcap_{i\in I} \mathrm{epi}\, f_i$

(ii) I : finite $\Rightarrow$ $\mathrm{epi}\min_{i\in I} f_i = \bigcup_{i\in I} \mathrm{epi}\, f_i$

Proof:

(i) $\forall_{(x,\xi)}$



$(x,\xi)\in \mathrm{epi}\left(\sup_{i\in I} f_i\right)$

$\leftrightarrow \underbrace{\left(\sup_{i\in I} f_i\right)(x)}_{\sup_{i\in I} f_i(x)} \leq \xi$

$\leftrightarrow \forall_{i\in I}\ \underbrace{f_i(x)\leq\xi}_{(x,\xi)\in \mathrm{epi}\, f_i} \leftrightarrow (x,\xi)\in \bigcap_{i\in I} \mathrm{epi}\, f_i$

(ii) $\forall_{(x,\xi)}$ $(x,\xi)\in \mathrm{epi}\left(\min_{i\in I} f_i\right)$

$\leftrightarrow \min_{i\in I} f_i(x) \leq \xi$ # I : finite $\leftrightarrow \exists_{i\in I}\ \forall_{j\in I}\ i\leq j$

$\leftrightarrow \exists_{i\in I}\ \underbrace{f_i(x)\leq\xi}_{(x,\xi)\in \mathrm{epi}\, f_i} \leftrightarrow (x,\xi)\in \bigcup_{i\in I} \mathrm{epi}\, f_i$

∎

• Extended real line:

$[-\infty,\infty] = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$

[12/3/2015 7:57 AM]

1.7. Topological space: /* see Wiki article on [Topological space] */

$\mathcal{X}$: set

$\mathcal{T}$: family of subsets of $\mathcal{X}$ containing
- $\mathcal{X}$ /* $\mathcal{X}$ is generally open and close at the same time */
- ϕ /*open and closed at the same time */
- arbitrary union of elements of $\mathcal{X}$
- finite intersections " " "

} these elements are open sets

$(\mathcal{T},\mathcal{X})$: topological space

for more specific definition see the neighborhood axioms [Neighborhood definition] [12/3/2015 8:07 AM]

• V : neighborhood of x $\overset{def}{\longleftrightarrow} \exists_{U\in\mathcal{T}}\ (x\in U \wedge U\subseteq V)$ /* A neighborhood can be open or close */

$(T, \mathcal{X})$: topological space

for more specific definition see the neighborhood axioms [Neighborhood definition] [12/3/2015 8:07 AM]

- $\underbrace{V}_{\subseteq \mathcal{X}}$: neighborhood of $x \overset{def}{\leftrightarrow} \exists_{U \in T} (x \in U \wedge U \subseteq V)$ /* A neighborhood can be open or close */

  # for a neighborhood you can find a smaller neighborhood in the topology

- $\mathcal{V}(x)$: set of all neighborhoods of x

- $\underbrace{\mathcal{X}}_{\text{topological space}}$: Hausdorff space $\overset{def}{\leftrightarrow} \forall_{x_1, x_2 \in \mathcal{X} : x_1 \neq x_2} \exists_{V_1 \in \mathcal{V}(x_1)} \exists_{V_2 \in \mathcal{V}(x_2)} V_1 \cap V_2 = \phi$

- $\underbrace{C}_{\subseteq \mathcal{X}}$: compact $\overset{def}{\leftrightarrow}$ (C: contained in the union of a family of open sets) $\Rightarrow$ (C: contained in the union of a finite subfamily from that family)

- $\underbrace{(x_a)_A}_{\text{net in } \mathcal{X}}$: converges to $x \in \mathcal{X} \overset{def}{\leftrightarrow} \forall_{V \in \mathcal{V}(x)} \exists_{b \in A} \forall_{a \in A} (a \succeq b \Rightarrow x_a \in V)$

- $\underbrace{x}_{\in \mathcal{X}}$: cluster point of $(x_a)_A \overset{def}{\leftrightarrow} (x_a)_{a \in A}$ lies frequently in every neighborhood of x

  $\overset{def}{\leftrightarrow} \forall_{V \in \mathcal{V}(x)} \forall_{b \in A} \exists_{a \in A : a \succeq b} x_a \in V$

  /* Pick any neighborhood of x, and any index b∈A of the net, then there will be a subsequent index a ∈ A : a ≥ b, such that x_a lies in that neighborhood. If this algorithm evaluates true, then we have got ourselves a cluster point x. */

  $\overset{def}{\leftrightarrow} \exists_{(x_{k(b)})_{b \in A}} x_{k(b)} \to x$ /* $(y_b)_{b \in B}$: subnet of $(x_a)_{a \in A} \overset{def}{\leftrightarrow} \exists_{k : B \to A} \left(\begin{array}{l} \bullet\ y_b = x_{k(b)} \\ \bullet\ \forall_{a \in A} \exists_{d \in B} \forall_{b \in B : b \succeq d} k(b) \succeq a \end{array} \right)$

*

[Sequential Cluster Point]

11/17/2015 8:56 AM

$\underbrace{x}_{\in \mathcal{X}}$: sequential cluster point of $\underbrace{(x_n)_{n \in \mathbb{N}}}_{\text{sequence in } \mathcal{X}} \overset{def}{\leftrightarrow} (x_n)_{n \in \mathbb{N}}$ has a subsequence that converges to x

Lemma 1.10.

(C: subset of Hausdorff space, $x \in \mathcal{X}$)

$\underbrace{x \in \bar{C}}_{\text{closure of } C} \leftrightarrow \exists_{(x_a)_{a \in A}} (x_a \to x)$

Lemma 1.12.

[C: compact subset of $\overbrace{\mathcal{X}}^{\text{Hausdorff space}}$] $\Rightarrow$ C: closed, $\forall$subsets of C is compact

* Lemma 1.14.

[C: compact subset of a Hausdorff space $\mathcal{X}$

$(x_a)_{a \in A}$: net in C, x: unique cluster point in C of the net

$x_a \to x$

2/10/2016 2:24 PM

Proof: per absurdum,

$x_a \not\to x$

/* $x_a \to x \overset{def}{\leftrightarrow} \forall_{V \in \mathcal{V}(x)} \exists_{z \in A} \forall_{y \in A : y \succeq z} x_y \in V \leftrightarrow (x_a)_{a \in A}$: eventually lies in every neighborhood of x */

$\therefore x_a \not\to x \leftrightarrow \exists_{V \in \mathcal{V}(x)} \forall_{z \in A} \exists_{y \in A : y \succeq z} x_y \notin V$

As a result we can construct a subnet $(x_{k(a)})_{a \in A}$ of $(x_a)_{a \in A}$ such that $(x_{k(a)})_{a \in A}$ lies entirely outside V (e.g. take z, then find all those $y : y \succeq z : x_y \notin V$ then some $\tilde{z} \succeq z$ " " " $\tilde{y} : \tilde{y} \succeq \tilde{z} : x_{\tilde{y}} \notin V$ and so on)

but as $(x_a)_{a \in A} \subseteq C \Rightarrow (x_{k(a)})_{a \in A} \subseteq C$

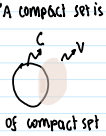
$\therefore (x_{k(a)})_{a \in A} \subseteq C \setminus V$

C: compact $\Rightarrow$ C: closed /* Lemma 1.12 */ A compact set is closed.

$C \setminus V \subseteq C$

if $C \cap V = \emptyset$, then $C \setminus V = C$: closed

if $C \cap V \neq \emptyset$, then $C \setminus V$: closed



So $C \setminus V$: compact /* closed subset of compact set is compact */ Lemma 1.12

/* recall that:

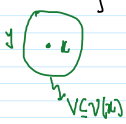
C: compact $\overset{\text{def}}{\Leftrightarrow}$ every net in C has a cluster point

(i.e., every net in C has a subnet that converges to a point in C */

as $(x_{k(a)})_{a\in A} \subseteq \underbrace{C\setminus V}_{\text{compact}}$, so $(x_{k(a)})_{a\in A}$ must have a cluster point in $C\setminus V$, say $y\in C\setminus V \Rightarrow y\notin V\subseteq V(x)$ $\Big\}\Rightarrow y\neq x\in V$

cluster point of $(x_{k(a)})_{a\in A}$ hence cluster point of $(x_a)_{a\in A}$

cluster point of $(x_a)_{a\in A}$



so we have found two different cluster points of $(x_a)_{a\in A}$: contradiction.

$\therefore x_a \to x$.

∎

2016-12-02 11:37 AM

Lemma 1.16

$[\,(\xi_a)_{a\in A}, (\eta_a)_{a\in A}$: nets in $[-\infty,+\infty]$, $\underline{\lim}\,\xi_a > -\infty$, $\underline{\lim}\,\eta_a > -\infty\,] \Rightarrow \underline{\lim}\,\xi_a + \underline{\lim}\,\eta_a \leq \underline{\lim}\,(\xi_a+\eta_a)$

Proof: $a\in A$

$\forall_{c\in A}\ c\succeq a \Rightarrow \underbrace{\inf_{\substack{b\in A\\ b\succeq a}} \xi_b}_{\leq \xi_{\tilde b}\ \forall_{\tilde b\in A:\tilde b\succeq a}} + \underbrace{\inf_{\substack{b\in A\\ b\succeq a}} \eta_b}_{\leq \eta_{\tilde b}\ \forall_{\tilde b\in A:\tilde b\succeq a}} \leq \xi_c+\eta_c$ $\quad\therefore \forall_{a\in A}\left(\inf_{b\in A:b\succeq a}\xi_b + \inf_{b\in A:b\succeq a}\eta_b \leq \xi_c+\eta_c \quad \forall_{c\in A:c\succeq a}\right) \Leftrightarrow \forall_{a\in A}\left(\inf_{b\in A:b\succeq a}\xi_b + \inf_{b\in A:b\succeq a}\eta_b \leq \inf_{\substack{c\in A:\\ c\succeq a}}(\xi_c+\eta_c)\right)$

/* set $\tilde b := c : c\succeq a$ */

$\Leftrightarrow \underbrace{\sup_{a\in A}\inf_{b\in A:b\succeq a}\xi_b}_{\underline{\lim}\,\xi_b} + \underbrace{\sup_{a\in A}\inf_{b\in A:b\succeq a}\eta_b}_{\underline{\lim}\,\eta_b} \leq \underbrace{\sup_{a\in A}\inf_{\substack{c\in A:\\ c\succeq a}}(\xi_c+\eta_c)}_{\underline{\lim}\,(\xi_c+\eta_c)}$

$\therefore \underline{\lim}\,\xi_b + \underline{\lim}\,\eta_b \leq \underline{\lim}\,(\xi_b+\eta_b)$

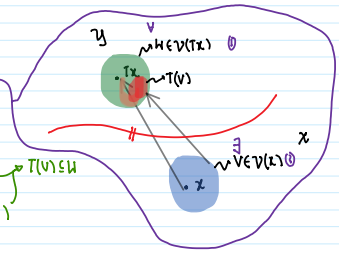
∎

*Continuity: [sec: Continuity] [11/14/2015 7:28 AM]

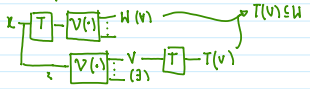
(X,T_X) : topological spaces

(Y,T_Y) " "

$\underbrace{T}_{:X\to Y}$: continuous at $x\in X \overset{\text{def}}{\Leftrightarrow} \forall_{W\in\mathcal{V}(Tx)} \exists_{V\in\mathcal{V}(x)}\ T(V)\subseteq W$



#



T: continuous $\overset{\text{def}}{\Leftrightarrow} (T$: continuous at $x)_{\forall_{x\in X}}$

*fact:

(X,T_X) : topological space

(Y,T_Y) : topological space

$T: X\to Y$

B_Y : base of T_Y /* $\underbrace{B}_{\text{subfamily of } T}$: base of $T \overset{\text{def}}{\Leftrightarrow} \forall_{x\in X}\ \forall_{V\in\mathcal{V}(x)}\ \exists_{\tilde B\in B}\ (x\in\tilde B \wedge \tilde B\subseteq V)$ */

then,

T: continuous $\Leftrightarrow \forall_{\tilde B\in B_Y}\ \underbrace{T^{-1}(\tilde B)}_{T^{-1}(B)=\{x\,|\,Tx\in\tilde B\}}\in T_X$

* fact: 1.19

X, Y : Hausdorff space

$T: X\to Y$

$x\in X$

[def: continuity using nets] [11/14/2015 7:28 AM]

T: continuous at $x \Leftrightarrow \left(\forall_{(x_a)_{a\in A}:\text{net in } X,\ x_a\to x}\ Tx_a\to Tx\right)$

/* When a continuous mapping takes a convergent net, the output net is also convergent with the limit being the operator acting on the *input sequence limit point* */

very intuitive

*Lemma 1.20:

X, Y: Hausdorff spaces

$T: X\to Y$, continuous

C : compact $\subseteq X$

$\Rightarrow \underbrace{T(C)}_{\{Tx\,:\,x\in C\}}$: compact

Proof:

$(y_a)_{a\in A}$: any net in $T(C)$

$\Rightarrow \forall_{a\in A}\ y_a\in T(C) = \{Tx : x\in C\}$

$\Leftrightarrow \forall_{a\in A}\ \exists_{x_a\in C}\ y_a = Tx_a$ $\quad\therefore$ Any net in $T(C)$ has form $(Tx_a)_{a\in A}$

given, C: compact # recall: $[\underbrace{C\subseteq X}_{\text{Hausdorff space}}]$ C: compact $\Leftrightarrow$ every net in C has a subnet that converges to a point in C

$\Updownarrow$

$\exists_{(x_{k(b)})_{b\in B}}$: subnet of $(x_a)_{a\in A}$ $\quad x_{k(b)}\to \underbrace{x}_{\in C}$

given, T: continuous

recall fact 1.19:

X, Y : Hausdorff space

$(x_{k(b)})_{b\in B}$: subnet of $(x_a)_{a\in A}$ $\in C$

given, T: continuous

$\Updownarrow$

$\forall_{x\in X}\ \forall_{(x_a)_{a\in A}: x_a\to x}\ Tx_a\to Tx$

$\ddagger$ recall fact 1.19:

X, Y : Hausdorff space
$T: X\to Y$
$x\in X$
T: continuous at $x \Leftrightarrow \left(\forall_{(x_a)_{a\in A}\text{: net in } X,\ x_a\to x}\ Tx_a\to Tx\right)$

as the subnet $(Tx_{k(b)})_{b\in B}$ comes from $(Tx_a)_{a\in A} \Rightarrow \left(Tx_{k(b)}\to \underbrace{Tx}_{\in T(C)}\right)$ /* $\because$ fact 1.9. says

$\left(\underbrace{(x_a)_{a\in A}}_{\text{net in Hausdorff space}}: x_a\to x,\ (x_{k(b)})_{b\in B}: \text{subnet of } (x_a)_{a\in A}\right) \Rightarrow \left(x_{k(b)}\to x\right)$ */

So, we have shown:

$\forall_{(Tx_a)_{a\in A}\text{: net in } T(C)}\ \exists_{(Tx_{k(b)})\text{: subnet of } (Tx_a)_{a\in A}}\ Tx_{k(b)}\to \underbrace{Tx}_{\in T(C)}$

$T(C)$: compact ☺

Chapter 1 : Part 2

Thursday, November 12, 2015 6:52 AM

1.10. Lower semicontinuity [sec: Lower Semicontinuity] [11/14/2015 7:28 AM]

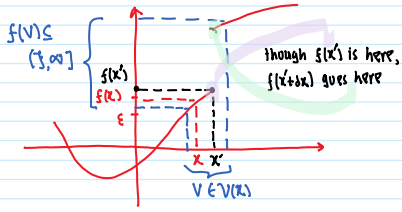
Definition.

$\boxed{\begin{array}{l} X: \text{Hausdorff space} \\ f: X \to [-\infty, \infty] \\ x \in X \end{array}}$

/* limit inferior of a net $(\xi_a)_{a\in A}$ in $[-\infty,\infty]$, $\underline{\lim}\, \xi_a = \sup_{c\in A} \inf_{\substack{b\in A \\ c\leq b}} \xi_b$ */

f: lower semicontinuous at $x \overset{\text{def}}{\leftrightarrow} \forall_{(x_a)_{a\in A}:\text{net in } X,\, x_a\to x}\ \underline{\lim}\, f(x_a) \geq f(x)$

$\overset{\text{def}}{\leftrightarrow} \forall_{\xi\in(-\infty, f(x))}\ \exists_{V\in\mathcal{V}(x)}\ f(V) \subseteq (\xi, \infty]$



f: upper semicontinuous at $x \leftrightarrow -f$: lower semicontinuous at x

f: continuous at $x \leftrightarrow f$: both upper and lower semicontinuous at x

* Domain of continuity: $\text{cont} f = \{x\in X : f(x)\in\mathbb{R} \text{ and } f: \text{continuous at } x\}$

$\subseteq \text{int dom } f$

* nice to know:
- A function cannot be continuous and real valued $\in(-\infty,\infty)$ on the boundary of its domain
- A continuous function can take values $-\infty$ and $+\infty$.

- $\boxed{\begin{array}{l} X: \text{Hausdorff space} \\ f: X \to [-\infty, \infty] \\ x \in X \end{array}}$ then $\underline{\lim}_{y\to x} f(y) = \sup_{V\in\mathcal{V}(x)} \inf f(V)$

$= \left(V\in\mathcal{V}(x) \,//\, f(V) \,//\, \inf_X (V) \right) \,//\, \sup_{V\in\mathcal{V}(x)} (V)$

Lemma 1.23.

$\boxed{\begin{array}{l} X: \text{Hausdorff space} \\ f: X \to [-\infty, \infty] \\ x \in X \\ N(x): \text{set of all nets in } X \text{ converging to } x \end{array}}$

/* recall limit inferior of a net $(\xi_a)_{a\in A}$ in $[-\infty,+\infty]$,

$\underline{\lim}\, \xi_a = \sup_{c\in A} \inf_{\substack{b\in A \\ b\geq c}} \xi_b = \left(\xi_{b:b\in A, b\geq c} \,//\, \inf f(V) \,//\, \sup_{c\in A} (V) \right)$ */

$\underline{\lim}_{y\to x} f(y) = \underbrace{\min_{\substack{(x_a)_{a\in A} \\ \in N(x)}} \underbrace{\underline{\lim}\, f(x_a)}_{\text{limit inferior for one net } (x_a)_{a\in A} \text{ in the neighborhood of } N(x)}}_{\text{minimum over all such nets in } N(x)}$

[Lemma 1.24] **Lemma 1.24.** /* Relates lower semicontinuity with epigraph and sublevel set */

$\boxed{\begin{array}{l} X: \text{Hausdorff space.} \\ f: X \to [-\infty, \infty] \end{array}} \Rightarrow$

(i) f: lower semicontinuous /* i.e. $(f: \text{lower semicontinuous at } x)_{\forall x\in X}$ */

$\Updownarrow$

(ii) $\text{epi} f$: closed in $X\times\mathbb{R}$

$\Updownarrow$

$\Updownarrow$

(ii) $\operatorname{epi} f$: closed in $\mathcal{X}\times\mathbb{R}$

$\Updownarrow$

(iii) $\forall_{\xi\in\mathbb{R}}$ $\operatorname{lev}_{\leq\xi} f$: closed in $\mathcal{X}$

<u>Proof</u>: (i)$\Rightarrow$(ii):

/* C: closed $\Leftrightarrow$ $\forall_{(x_a)_{a\in A}:\subseteq C,\, x_a\to x}$ $x\in C$ */

take $(x_a,\xi_a)_{a\in A}:\subseteq \operatorname{epi} f$, $(x_a,\xi_a)\to(x,\xi)\in\mathcal{X}\times\mathbb{R}$

as

$(x_a,\xi_a)\in\operatorname{epi} f \Leftrightarrow f(x_a)\leq\xi_a$

$x_a\to x$, f: lower semicontinuous $\Rightarrow f(x)\leq\underline{\lim}\, f(x_a)$

* Definition 1.21.
[$\mathcal{X}$: Hausdorff space, $f:\mathcal{X}\to[-\infty,+\infty]$, $x\in\mathcal{X}$]
f: lower semicontinuous at x $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x_a)_{a\in A}:\text{net in }\mathcal{X},\, x_a\to x}$ $\underline{\lim}\, f(x_a)\geq f(x)$ # definition in terms of a net

$\leq\underline{\lim}\,\xi_a$ /*: $x_a\geq y_a\Rightarrow \underline{\lim}\, x_a\geq\underline{\lim}\, y_a$ */

$=\lim\xi_a=\xi$

$\therefore f(x)\leq\xi\Leftrightarrow(x,\xi)\in\operatorname{epi} f$

So, $\operatorname{epi} f$: closed.

(ii)$\Rightarrow$(iii)

given: $\operatorname{epi} f$: closed $\Leftrightarrow$ $\forall_{(y_a)_{a\in A}:\subseteq\operatorname{epi} f,\, y_a\to y}$ $y\in\operatorname{epi} f$

goal: $\forall_{\xi\in\mathbb{R}}$ $\operatorname{lev}_{\leq\xi} f$: closed in $\mathcal{X}$

/* C: closed $\Leftrightarrow$ $\forall_{(x_a)_{a\in A}:\subseteq C,\, x_a\to x}$ $x\in C$ */

$\forall_{(x_a)_{a\in A}:\subseteq\operatorname{lev}_{\leq\xi} f,\, x_a\to x}$

$x_a\in\operatorname{lev}_{\leq\xi} f\Leftrightarrow f(x_a)\leq\xi\Leftrightarrow(x_a,\xi)\in\operatorname{epi} f$

now $(x_a,\xi)\to(x,\xi)$

$\Rightarrow(x,\xi)\in\operatorname{epi} f$

$\Leftrightarrow f(x)\leq\xi\Leftrightarrow x\in\operatorname{lev}_{\leq\xi} f$

$\therefore\forall_{\xi\in\mathbb{R}}$ $\operatorname{lev}_{\leq\xi} f$: closed in $\mathcal{X}$.

(iii)$\Rightarrow$(i)

given: $\forall_{\xi\in\mathbb{R}}$ $\operatorname{lev}_{\leq\xi} f$: closed in $\mathcal{X}$ $\Leftrightarrow$ $\forall_{\xi\in\mathbb{R}}$ $\forall_{(\tilde{x}_a)_{a\in A}:\subseteq\operatorname{lev}_{\leq\xi} f,\, \tilde{x}_a\to\tilde{x}}$ $\tilde{x}\in\operatorname{lev}_{\leq\xi} f$ [eq: 2017-02-08 2:23 PM]

/*

* Definition 1.21.
[$\mathcal{X}$: Hausdorff space, $f:\mathcal{X}\to[-\infty,+\infty]$, $x\in\mathcal{X}$]
f: lower semicontinuous at x $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x_a)_{a\in A}:\text{net in }\mathcal{X},\, x_a\to x}$ $\underline{\lim}\, f(x_a)\geq f(x)$ # definition in terms of a net

*/

take $(x_a)_{a\in A}:\subseteq\mathcal{X}$, $x_a\to x$

want to show $\underline{\lim}\, f(x_a)\geq f(x)$

set $\mu=\underline{\lim}\, f(x_a)\in[-\infty,+\infty]$

if $\mu=+\infty\Rightarrow$ goal is trivially proved.

now consider, $\mu\neq+\infty$, $\mu\in[-\infty,+\infty)$

Fact 1.15. ($\underline{\lim}$, $\overline{\lim}$ in different light) /* Conventionally this holds for bounded net */
[$(\xi_a)_{a\in A}$: net in $[-\infty,+\infty]$]
(i) $\underline{\lim}\,\xi_a=\lim\inf_{b\in A:b\geq a}\xi_b$; $\overline{\lim}\,\xi_a=\lim\sup_{b\in A:b\geq a}\xi_b$
(ii) $(\xi_a)_{a\in A}$: possesses subnets that converge to $\underline{\lim}\,\xi_a$ and $\overline{\lim}\,\xi_a$ respectively.
(iii) $(\xi_a)_{a\in A}$: converges $\Leftrightarrow$ $\underline{\lim}\,\xi_a=\overline{\lim}\,\xi_a$ /* in this case, $\underline{\lim}\,\xi_a=\lim\xi_a=\overline{\lim}\,\xi_a$ */

$\exists_{(x_{k(b)})_{b\in B}:\text{subnet of }(x_a)_{a\in A}}$ $f(x_{k(b)})\to\mu$

set $\xi\in]\mu,+\infty[\subseteq\mathbb{R}$

then $f(x_{k(b)})_{b\in B}$: eventually lies in $[-\infty,\xi]$

/* $(x_a)_{a\in A}$: eventually in $A\overset{\text{def}}{\Leftrightarrow}\exists_{c\in A}\ \forall_{a\in A:a\geq c}$ $x_a\in Y$ */

$\Leftrightarrow\exists_{c\in A}\ \forall_{b\in B:b\geq c}$ $f(x_{k(b)})\in[-\infty,\xi]$

/* $(x_a)_{a\in A}$: eventually in $A \overset{\text{def}}{\Leftrightarrow} \exists_{c\in A}\ \forall_{a\in A: a\succeq c}\ x_a \in Y$ */

$\Leftrightarrow \exists_{c\in B}\ \forall_{b\in B: b\succeq c}\ f(x_{k(b)}) \in [-\infty, \xi]$

$\Updownarrow$

$-\infty \leq f(x_{k(b)}) \leq \xi$

$\Updownarrow$

$x_{k(b)} \in \mathrm{lev}_{\leq\xi} f$

$\Rightarrow \{x_{k(b)} \mid b\succeq c,\ b\in B\} \subseteq \mathrm{lev}_{\leq\xi} f$

So. $(x_{k(b)})_{\substack{b\succeq c;\\ b\in B}} : \subseteq \mathrm{lev}_{\leq\xi} f,\ x_{k(b)} \to x$, as $\forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$

using [eq: 2017-02-08 2:23 PM] : $x \in \mathrm{lev}_{\leq\xi} f \Leftrightarrow f(x) \leq \xi \in]\mu, +\infty[$

letting $\xi \downarrow \mu$ we have $f(x) \leq \mu = \underline{\lim}\ f(x_a)$

So f: lower semicontinuous at every point in X.

$\boxtimes$

•Example: 1.25.

$$\iota_C : X \to [-\infty, \infty] : x \to \begin{cases} 0, & \text{if } x\in C \\ +\infty, & \text{else} \end{cases}.$$

ι_C: lower semi continuous $\Leftrightarrow$ C: closed

Proof: $\forall_{\xi\in\mathbb{R}}$ /* Recall $\xi\in\mathbb{R} \leftrightarrow -\infty<\xi<\infty$ */

$\mathrm{lev}_{\leq\xi} \iota_C = \{x\in X \mid f(x)\leq\xi\}$

$\xi<0 \to \mathrm{lev}_{\leq\xi} \iota_C = \{x\in X \mid f(x)<0\} = \emptyset$: closed /* empty set is both open and closed */

$\xi\in[0,\infty) \to \mathrm{lev}_{\leq\xi} \iota_C = \{x\in X \mid f(x)\leq 0\} = \{x\in X \mid f(x)=0\}$ /* by def $f(x)\geq 0$ */

$= \{x\in X \mid x\in C\} = C$: closed

$\therefore \forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: closed in $X \leftrightarrow \iota_C$: lower semicontinuous

*Lemma 1.26:

X: Hausdorff space
$(f_i)_{i\in I}$: family of lower semicontinuous functions from X to $[-\infty, +\infty]$

$\Rightarrow$

- $\sup_{i\in I} f_i$: lower semicontinuous
- I: finite $\Rightarrow \min_{i\in I} f_i$: lower semi continuous

* Lemma 1.27.

X: Hausdorff space
$(f_i)_{i\in I}$: finite family of lower semicontinus functions from $(-\infty, +\infty]$
$(\alpha_i)_{i\in I}$: $\alpha_i \in \mathbb{R}_{++}$

$\Rightarrow \sum_{i\in I} \alpha_i f_i$: lower semicontinuous

* Theorem 1.28 (Weierstrass) /* fundamental tool in proving existence of solutions of optimization problem */

X: Hausdorff space
$f: X \to [-\infty, +\infty]$, lower semicontinuous
C: compact subset of X
$C \cap \mathrm{dom} f \neq \emptyset$

$\Rightarrow$ f achieves its infimum over C

$\mathrm{dom} f = \{x \mid f(x) < +\infty\}$

1/18/2016 5:09 PM

Proof:

$\underbrace{\inf f(C)}_{\text{a set of real numbers}} = \inf_{x\in C} f(x) + \iota_C(x) \in]-\infty, +\infty[$ as. $C \cap \mathrm{dom} f \neq \emptyset$

so $\inf f(C)$ will exist, the question is whether it will exist in C, i.e., $\overset{?}{\exists} x\in C\ f(x) = \inf f(C)$

$\Rightarrow \exists_{\substack{\text{minimizing}\\ \text{sequence}}}$ of $\underbrace{f \text{ over } C}_{f+\iota_C}$; let the sequence be $\underbrace{(x_a)_{a\in A}}_{\text{also a net in } C}$ $\therefore f(x_a) \to \inf f(C)$ /* using *Fact 1.8.1. (Existence of a minimizing sequence for finite infimum)

$\inf f(C) \in \mathbb{R} \Rightarrow \exists_{(x_n)_{n\in\mathbb{N}}: \subseteq C}\ f(x_n) \to \inf f(C)$ */

C: compact $\leftrightarrow$ every net in C has a subnet that converges to a point in C /* fact 1.11 */

minimizing sequence of $\underbrace{f \text{ over } C}_{f+\iota_C}$; let the sequence be $\underbrace{(x_a)_{a\in A}}_{\text{also a net in } C}$ $\therefore f(x_a) \to \inf f(C)$ /* using

/* Fact 1.8.1. (Existence of a minimizing sequence for finite infimum)

$\inf f(C) \in \mathbb{R} \Rightarrow \exists_{(x_n)_{n\in\mathbb{N}} : \subseteq C} \; f(x_n) \to \inf f(C)$ */

C: compact $\leftrightarrow$ every net in C has a subnet that converges to a point in C /* fact 1.11 */

we can extract subnet $(x_{k(b)})_{b\in B}$ from $(x_a)_{a\in A}$ such that $x_{k(b)} \to \underbrace{x}_{\text{some point in } C}$

then. $f(x_{k(b)}) \to \inf f(C) \leq f(x) \; \forall_{x\in C}$ /* $x_{k(b)}$ is a subnet of the minimizing sequence. so $f(x_{k(b)}) \to \inf f(C)$ */

by definition of infimum $f(x) \geq \inf f(C) \; \forall x \in C$

now f: lower semicontinuous /* f: lower-semicontinuous at x $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x_a)_{a\in A}:\text{subnet in } X,\, x_a \to x}$ $f(x) \leq \underline{\lim} f(x_a)$ */

as subnet is also a net, we have $f(x) \leq \underline{\lim} f(x_{k(b)})$ /* $x_{k(b)} \to x$. so $\underline{\lim} f(x_{k(b)}) = \lim f(x_{k(b)}) = \overline{\lim} f(x_{k(b)})$ as for a converging sequence $\underline{\lim}, \lim, \overline{\lim}$ are sam */

$= \lim f(x_{k(b)})$ /* $\because f(x_{k(b)}) \to \inf f(C)$ $\therefore \lim f(x_{k(b)}) = \inf f(C)$

$= \inf f(C)$

$\therefore \inf f(C) \leq f(x) \leq \inf f(C) \leftrightarrow \underset{\in C}{f(x)} = \inf f(C)$ $\therefore$ f achieves its infimum over C. ■

* Lemma 1.29.

$\left(\begin{array}{l} X: \text{Hausdorff space} \\ C: \text{compact Hausdorff space} \\ \varphi: X \times C \to [-\infty, +\infty], \text{ lower semicontinuous} \end{array}\right) \Rightarrow$

- $(f: X \to [-\infty, +\infty] : x \mapsto \inf \varphi(x, C))$: lower semicontinuous /* f is called the marginal function */
- $\forall_{x\in X}$ $f(x) = \min \varphi(x, C)$

- Definition 1.30.

(X: Hausdorff space)

$\underbrace{\bar{f}}_{: X \to [-\infty, \infty]}$: lower semicontinuous envelope $\overset{\text{def}}{\Leftrightarrow}$ $\bar{f} = \sup\{g : X \to [-\infty, +\infty] \mid g \leq f,\ g: \text{lower semicontinuous}\}$

$= (g : (\leq f, \text{lower semicontinuous}, : X \to [-\infty, \infty]) \,//\, \sup(\cdot))\}$

- Lemma 1.31.

$\left(\begin{array}{l} X: \text{Hausdorff space} \\ f: X \to [-\infty, \infty] \end{array}\right) \Rightarrow$

(i) $\bar{f}$: largest lower semicontinuous function majorized by f.

(ii) epi $\bar{f}$: closed

(iii) dom $f \subseteq$ dom $\bar{f} \subseteq \overline{\text{dom}}\, f$, (iv) $\forall_{x\in X}$ $\bar{f}(x) = \underline{\lim}_{y\to x} f(y)$

(v) $x \in X \Rightarrow (f$: lower semicontinuous at $f \leftrightarrow \bar{f}(x) = f(x))$

(vi) epi $\bar{f} = \overline{\text{epi}}\, f$

Chapter 1: Part 3

Saturday, November 14, 2015 6:50 AM

[Sequentially closed set] 1.11. Sequential Topological Notions

$\left[\begin{array}{l} X: \text{Hausdorff space} \\ C: \text{subset of } X \end{array}\right]$ $(C: \text{sequentially closed}) \overset{\text{def}}{\leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}: \text{convergent, lies in } C} \left(\lim_{n\to\infty} x_n\right): \text{lies in } C$

/* example: $C := (0,1]$, consider $(1/n)_{n\in\mathbb{N}}$ which converges to 0, and $\forall_n\ (1/n) \in C$, but 0, the limit itself does not. So, $C = (0,1]$ is not sequentially closed. */

- A closed set is sequentially closed, but converse is not true

- Definition 1.32.

[Sequentially compact set] $[\ X: \text{Hausdorff space},\ C \subseteq X\ ]$

11/17/2015 8:56 AM

11/17/2015 9:20 AM

C: sequentially compact $\overset{\text{def}}{\leftrightarrow}$ $\forall_{\text{sequence in } C}\ \exists_{\text{subsequence}}$ subsequence converges to a point in C.

$\overset{\text{def}}{\leftrightarrow} \forall_{\text{sequence in } C}$ the sequence has a sequential cluster point

/* $(x_n)_{n\in\mathbb{N}}$ in X has a subsequence converging to $x \in X \overset{\text{def}}{\leftrightarrow} x$: sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ */

[11/14/2015 7:28 AM] In the notions of continuity and lower-semicontinuity replace nets by sequence, then we

[def: continuity using nets] [sec: Lower Semicontinuity]

get sequential continuity and sequential lower semicontinuity.

- [Sequentially Continuous operator] $[\ X, Y: \text{Hausdorff space},\ T: X \to Y,\ x \in X\ ]$

11/17/2015 9:20 AM

T: sequentially continuous at $x \overset{\text{def}}{\leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}: \text{sequence in } X,\ x_n \to x}\ Tx_n \to Tx$

[Sequentially lower semicontinuous at x]

- f: sequentially lower semicontinuous at $x \overset{\text{def}}{\leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}: \text{sequence in } X,\ x_n \to x}\ \underbrace{\underline{\lim}\ f(x_n) \geq f(x)}_{\sup_{c\in\mathbb{N}} \inf_{\substack{b\in\mathbb{N} \\ b \geq c}} x_b}$

11/14/2015 7:55 AM Sequential versions of Lemma 1.12, Lemma 1.14, Lemma 1.24:

[Lemma 1.24]

11/17/2015 8:56 AM Lemma 1.33. /*[Sequentially compact set]*/

$[C: \text{sequentially compact subset of } \underbrace{X}_{\text{Hausdorff space}}]$

[Sequentially closed set]

C: sequentially closed, $\forall_{\tilde{C}: \text{sequentially closed subset of } C}$ $\tilde{C}$: sequentially compact.

Lemma 1.34.

$[C: \text{sequentially compact subset of } \underbrace{X}]$ /* Sequentially compact subset implies that any sequence in the subset will have a subsequence that converges to some point in that subset */

[C: sequentially compact subset of $\underbrace{X}_{\text{Hausdorff space}}$]

/* Sequentially compact subset implies that any sequence in the subset will have a subsequence that converges to some point in that subset */

$\underbrace{(x_n)_{n\in\mathbb{N}}}_{\text{sequence in } C}$: has unique sequential cluster point $x \Rightarrow x_n \to x$

[Sequential Cluster Point]

/* Recall that a sequential cluster point is the converging point of a subsequence of the original sequence. Now this theorem is saying that if we have a sequentially compact set C, in which the sequence in consideration has a unique sequential cluster point, then the sequence itself converges to that cluster point! */

Lemma 1.35.

$\begin{pmatrix} X: \text{Hausdorff space} \\ f: X \to [-\infty, +\infty] \end{pmatrix} \Rightarrow$

(i) f: sequentially lower semicontinuous

(ii) epi f: sequentially closed in $X \times \mathbb{R}$

(iii) $\forall_{\varepsilon \in \mathbb{R}}$ $\text{lev}_{\leq \varepsilon} f$: sequentially closed in X

Chapter 1: Part 4

Sunday, November 15, 2015 10:48 PM

1.12 Metric Spaces.

X: metric space with metric (distance d)

$\underbrace{\dim C}_{\text{diameter of } C \subseteq X} = \sup_{(x,y)\in C\times C} d(x,y)$

$\underbrace{d_C}_{\text{distance to a set } C\subseteq X} : X \to [0,+\infty] : x \mapsto \inf_{y\in C} d(x,C)$

$\underbrace{B(x;r)}_{\text{closed ball with center } x\in X,\text{ radius } r\in\mathbb{R}_{++}} = \{y\in X : d(x,y)\leq r\}$

$\underbrace{B(x;r)}_{\text{open ball}} = \{y\in X : d(x,y) < r\}$

Metric topology of X: Topology that admits the family of all open balls as a base

/* A subfamily $\mathcal{B}$ of topology $\mathcal{T}$ is a base $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in X}\ \forall_{V\in\mathcal{V}(x)}\ \exists_{\tilde{B}\in\mathcal{B}}\ (x\in\tilde{B}\cap\tilde{B}\subseteq V)$ */

Metrizable topological space: topology coincides with metrix topology

$\underbrace{(x_n)_{n\in\mathbb{N}}}_{\text{sequence in } X}$: converges to $x\in X$ $\overset{\text{def}}{\Leftrightarrow}$ $d(x_n,x)\to 0$

Fact 1.37.

11/17/2015 9:20 AM

[X: metric space, Y: Hausdorff space, $T: X\to Y$]

T: continuous $\overset{\text{def}}{\Leftrightarrow}$ T: sequentially continuous

[Sequentially Continuous operator]

Fact 1.38.

[Sequentially compact set]

[C: subset of $\underbrace{X}_{\text{metric space}}$] C: compact $\Leftrightarrow$ C: sequentially compact

Lemma 1.39.

$C: \underbrace{\subseteq X}_{\text{metric space}}$, $\forall_{n\in\mathbb{N}}\ C\cap B(0;n) \Rightarrow C$: closed

Lemma 1.40.

C: compact subset of $\underbrace{X}_{\text{metric space}}$ $\overset{\text{def}}{\Leftrightarrow}$ C: closed, bounded

Lemma 1.41.

[X: metric space

$f: X\to[-\infty,+\infty]$

$x\in X$

$S(x)$: set of all sequences in X that converge to x

] $\Rightarrow$

$\underline{\lim}_{y\to x} f(y) = \min_{(x_n)_{n\in\mathbb{N}}\in S(x)} \underline{\lim}\ f(x_n)$

- Lemma 1.42. (Cantor)

A sequence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence if $d(x_m,x_n)\to 0$ as $m,n\to\infty$

(X: complete metric space /* all Cauchy sequence converges */

$(C_n)_{n\in\mathbb{N}}$: sequence of nonempty closed sets, $\forall_{n\in\mathbb{N}}\ C_{n+1}\subseteq C_n$, $\operatorname{diam} C_n\to 0$)

$\Rightarrow \bigcap_{n\in\mathbb{N}} C_n$: singleton

2/5/2016 7:32 PM

Proof:

$C=\bigcap_{n\in\mathbb{N}} C_n \subseteq C_n\ \forall_{n\in\mathbb{N}} \Rightarrow \operatorname{diam} C \leq \operatorname{diam} C_n\ \forall_{n\in\mathbb{N}}$; now given $\operatorname{diam} C_n\to 0 \Rightarrow 0\leq \operatorname{diam} C\leq \operatorname{diam} C_n\ \forall_{n\in\mathbb{N}}$

$\Rightarrow 0\leq \operatorname{diam} C\leq \lim \operatorname{diam} C_n = 0 \Leftrightarrow \boxed{\operatorname{diam} C = 0}$

Proof:

$C = \bigcap_{n\in\mathbb{N}} C_n \subseteq C_n \ \forall_{n\in\mathbb{N}} \Rightarrow \operatorname{diam} C \leq \operatorname{diam} C_n \ \forall_{n\in\mathbb{N}}$; now given $\operatorname{diam} C_n \to 0 \Rightarrow 0 \leq \operatorname{diam} C \leq \operatorname{diam} C_n \ \forall_{n\in\mathbb{N}}$

$\Rightarrow 0 \leq \operatorname{diam} C \leq \lim \operatorname{diam} C_n = 0 \leftrightarrow \boxed{\operatorname{diam} C = 0}$

Ursing Result 1.2.1. $(0 \leq a \leq b_n \to 0) \Rightarrow a = 0$ */

take $\forall_{n\in\mathbb{N}} \ x_n \in C_n$

$A_n = \{x_m\}_{m\in\mathbb{N}: m\geq n} = \{\underbrace{x_n}_{\in C_n}, \underbrace{x_{n+1}}_{\in C_{n+1}\subseteq C_n}, x_{n+2}, \dots\} \subseteq C_n$

as $\operatorname{diam} C_n \to 0 \Rightarrow \operatorname{diam} A_n \to 0$ // $\operatorname{diam} C = \sup_{(x,y)\in C\times C} d(x,y)$

$\Rightarrow \sup\{d(x_m, x_p) \mid m,p\in\mathbb{N},\ m,p \geq n\} \to 0$

$\Rightarrow d(x_m, x_p) \to 0$ as $m,p \to \infty$

$\Leftrightarrow (x_n)_{n\in\mathbb{N}}$: Cauchy sequence

As, X is a complete metric space $\exists_{x\in X} \ x_n \to x$

$\forall_{n\in\mathbb{N}} \ \forall_{p\in\mathbb{N}} \ x_{n+p} \subseteq C_{n+p} \subseteq C_n$, $x_{n+p} \to x$ as $p \to \infty$

as in a closed set every convergent net has its limit in the set. $x \in C_n$ $\therefore \forall_{n\in\mathbb{N}} \ C_n \ni x \Leftrightarrow \bigcap_{n\in\mathbb{N}} C_n = C \ni x$ $\boxed{\therefore x\in C}$

$\Rightarrow \boxed{x\in C,\ \operatorname{diam} C = 0}$

So, $C = \{x\}$

∎

Lemma 1.43 (Ursescu)

- [X: complete metric space]

(i) $(C_n)_{n\in\mathbb{N}}$: sequence of closed subsets of $X \Rightarrow \overline{\bigcup_{n\in\mathbb{N}} \operatorname{int} C_n} = \overline{\operatorname{int} \bigcup_{n\in\mathbb{N}} C_n}$

(ii) $(C_n)_{n\in\mathbb{N}}$: sequence of open subsets of $X \Rightarrow \operatorname{int} \bigcap_{n\in\mathbb{N}} \overline{C_n} = \operatorname{int} \overline{\bigcap_{n\in\mathbb{N}} C_n}$

Corollary 1.44.

[X: complete metric space,

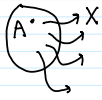
$(C_n)_{n\in\mathbb{N}}$: sequence of dense open subsets of X

]

/* In topology and related areas of mathematics, a subset A of a topological space X is called **dense** (in X) if every point x in X either belongs to A or is a limit point of A.[1] Informally, for every point in X, the point is either in A or arbitrarily "close" to a member of A — for instance, every real number is either a rational number or has one arbitrarily close to it
From <https://en.wikipedia.org/wiki/Dense_set> */

$\Rightarrow \bigcap_{n\in\mathbb{N}} C_n$: dense G_δ subset of X

/* G_δ : countable intersection of open sets in Hausdorff space */



Theorem 1.45. (Ekeland)

[

(X,d) : complete metric space

$f: X \to (-\infty, +\infty]$, proper, $\underbrace{\text{lower semi continuous}}_{\forall_{x\in X} \ \forall_{(x_a)_{a\in A}:\text{(net in } X,\ x_a\to x)} \ \liminf f(x_a) \geq f(x)}$, bounded below

$\alpha \in \mathbb{R}_{++}$, $\beta \in \mathbb{R}_{++}$

$y \in \operatorname{dom} f$: $f(y) \leq \alpha + \inf f(x)$ /* y is an arbitrarily close point to the infimum */

]

$\Rightarrow$

$\exists_{z\in X}$
- $f(z) + \frac{\alpha}{\beta} d(y,z) \leq f(y)$
- $d(y,z) \leq \beta$
- $\forall_{x\in X\setminus\{z\}} \ f(z) < f(x) + \frac{\alpha}{\beta} d(x,z)$

Definition 1.46. [Lipschitz Continuity] [11/19/2015 4:21 PM]

[$(X_1,d_1), (X_2,d_2)$: metric spaces

$T: X_1 \to X_2$

C : subset of X_1

]

- T: Lipschitz continuous with constant $\beta \in \mathbb{R}_+ \overset{\text{def}}{\longleftrightarrow} \forall_{x\in X_1} \ \forall_{y\in X_1} \ d_2(Tx,Ty) \leq \beta d_1(x,y)$
- T: (locally Lipschitz continuous near a point $x \in X_1$) $\overset{\text{def}}{\longleftrightarrow} \exists_{\rho\in\mathbb{R}_{++}} \ T|_{B(x;\rho)}$: Lipschitz continuous
- T: (locally Lipschitz continuous near [illegible]) $\overset{\text{def}}{\longleftrightarrow} \forall \ \forall \ d_2(Tx,Ty) \leq \beta d_1(x,y)$

(point $x \in X_1$] $\leftrightarrow \exists_{\rho \in \mathbb{R}_{++}}$ ' $B(x;\rho)$: Lipschitz continuous...

• T: (locally Lipschitz continuous near every point in C with constant $\beta \in \mathbb{R}_+$) $\overset{\text{def}}{\longleftrightarrow} \forall_{x \in C} \forall_{y \in C} \; d_2(Tx,Ty) \leq \beta d_1(x,y)$

11/19/2016 7:16 PM

Theorem 1.48. (Banach-Picard) [Convergence of contraction mapping iteration]

[(X,d) : complete metric space

$T: X \to X$: Lipschitz continuous with $\beta \in [0,1)$ /* it just means that T: contraction mapping */

$\updownarrow$

$\forall_{x,y \in X} \; d(Tx,Ty) \leq \beta d(x,y) \leftrightarrow \|Tx - Ty\| \leq \beta \|x-y\|$

$x_0 \in X$

$\forall_{n \in \mathbb{N}} \; x_{n+1} = Tx_n$

] $\Rightarrow$

$\exists x \in X$:

(i) x: unique fixed point of T

(ii) $\forall_{n \in \mathbb{N}} \; d(x_{n+1}, x) \leq \beta d(x_n, x)$ /* $\beta \in [0,1)$ so, the distance from the optimal point will strictly decrease */

(iii) $\forall_{n \in \mathbb{N}} \; d(x_n, x) \leq \beta^n d(x_1, x)$ /* converges linearly $\log \frac{d(x_n,x)}{d(x_0,x)} \leq n \log \beta$

(iv) A priori error estimate: $\forall_{n \in \mathbb{N}} \; d(x_n, x) \leq \frac{\beta^n}{(1-\beta)} d(x_0, x_1)$

(v) A posteriori error estimate:

$$\forall_{n \in \mathbb{N}} \; d(x_n, x) \leq \frac{d(x_n, x_{n+1})}{1-\beta}$$

(vi) $\frac{d(x_0,x_1)}{1+\beta} \leq d(x_n, x) \leq \frac{d(x_0,x_1)}{1-\beta}$ /* What a beautiful inequality! It shows how the distance from the original point is bounded */

/* Proof strategy:

We want to prove given $\Rightarrow$ ((i) $\wedge$ (ii) $\ldots \wedge$ (v)) $\leftrightarrow$ (given $\Rightarrow$ (i)) $\wedge \ldots \wedge$ (given $\Rightarrow$ (ii))

At first we prove $\underbrace{\text{given} \Rightarrow \text{(i)}, \text{ then (i)} \Rightarrow \text{(ii)}}_{\text{which implies given} \Rightarrow \text{(ii)}}$ ... and so on.

proof:

(ii)

The triangle inequality says: • $d(x,y) + d(y,z) \geq d(x,z)$

• $d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{n+m-1}, x_{n+m}) \geq d(x_n, x_{n+m})$

So, $\forall_{m \in \mathbb{N}} \; \forall_{n \in \mathbb{N}}$

$$d(x_n, x_{n+m}) \leq d(x_n, x_{n+1}) + \underbrace{d(x_{n+1}, x_{n+2})}_{\leq \beta^1 d(x_n, x_{n+1})} + \underbrace{d(x_{n+2}, x_{n+3})}_{\leq \beta^2 d(x_{n+1}, x_{n+2}) \leq \beta^2 d(x_n, x_{n+1})} + \ldots + \underbrace{d(x_{n+m-1}, x_{n+m})}_{\leq \beta d(x_{n+m-2}, x_{n+m-1}) \; \ldots \; \leq \beta^{m-1} d(x_n, x_{n+1})}$$

// $d(x_{n+1}, x_{n+2}) = d(Tx_n, Tx_{n+1}) \leq \beta (x_n, x_{n+1})$

$$\leq d(x_n, x_{n+1}) \underbrace{(1 + \beta + \beta^2 + \ldots + \beta^{m-1})}_{\frac{1(1-\beta^m)}{1-\beta}} = \frac{1-\beta^m}{1-\beta} d(x_n, x_{n+1}) \leq \frac{1}{1-\beta} d(x_n, x_{n+1}) \quad \ldots (1.68)$$

/* $\beta \in [0,1)$ */

/* $(1-\beta^m) \leq 1$, so removing it will only make the term bigger */

/* $d(x_n, x_{n+1}) \leq \beta d(x_{n-1}, x_n)$
$\leq \beta^2 d(x_{n-2}, x_{n-1})$
$\vdots$
$\leq \beta^n d(\underbrace{x_0}_{n=n}, x_1)$ */

$$\leq \frac{\beta^n}{1-\beta} d(x_0, x_1) \quad \ldots (1.69)$$

So, we have,

$$\forall_{m \in \mathbb{N}} \; \forall_{n \in \mathbb{N}} \; 0 \leq d(x_n, x_{n+m}) \leq \frac{\beta^n}{1-\beta} d(x_0, x_1)$$

if $n \to \infty$, then $d(x_n, x_{n+m}) \to 0$ $\therefore (x_n)_{n \in \mathbb{N}}$: Cauchy sequence

/* Recall, $(x_n)_{n \in \mathbb{N}}$: Cauchy sequence $\overset{\text{def}}{\leftrightarrow} d(x_m, x_n) \to 0$ as $m, n \to \infty$ */

Now,

(X,d) : complete metric space /* every Cauchy sequence converges */

$\therefore (x_n)_{n \in \mathbb{N}}$ converges to some $x \in X$.

now T: Lipschitz continuous /* recall fact 1.19. (X: Hausdorff space, Y: Hausdorff space, $T: X \to Y$, continuous at x) $\Rightarrow \forall_{(x_\alpha)_{\alpha \in A}}$: net in X, converges to x $Tx_\alpha \to Tx$ */

$\therefore Tx_n \to Tx$

$T: X \to Y$, continuous at x ... converges to x

$\therefore \underbrace{Tx_n}_{x_{n+1}} \to Tx$ /* $\because x_{n+1} = Tx_n$ */

$x_{n+1} \to Tx$ } /* In a Hausdorff space convergence of a sequence is always to a unique point */

but $x_{n+1} \to x$ } $\Rightarrow Tx = x \leftrightarrow x \in \mathrm{Fix}\,T$.

• (Uniqueness)

Consider $y \in \mathrm{Fix}\,T \setminus \{x\} \leftrightarrow x \neq y \wedge Ty = y$

$d(x,y) = d(Tx, Ty) \leq \beta\, d(x,y)$ /* [Lipschitz Continuity] [11/19/2015 4:21 PM] */

$\Leftrightarrow \beta \geq 1$ [$\because x \neq y \Rightarrow d(x,y) \neq 0$, so we can cancel]

$\downarrow$

contradiction as $\beta \in [0,1)$

$\therefore y = x$.

/* Proof strategy: for (ii)-(v):

$\underbrace{(A \to B)}_{\text{given (i)}} \wedge \underbrace{(B \to C)}_{\text{(i) (ii)}} \Rightarrow \underbrace{(A \to C)}_{\text{given (ii)}}$

*/

(ii)

By Lipschitz continuity, $\forall_{n \in \mathbb{N}}\ d(\underset{Tx_n}{\underset{\parallel}{x_{n+1}}}, \underset{Tx}{\underset{\parallel}{x}}) = d(Tx_n, Tx) \leq \beta\, d(x_n, x)$

(iii)

in (ii) we have shown, $d(x_{n+1}, x) \leq \beta\, d(x_n, x)$

so: $d(x_n, x) \leq \beta\, d(x_{n-1}, x) \leq \beta^2 d(x_{n-2}, x) \ldots \leq \beta^n d(\underbrace{x_{n-n}}_{0}, x)$

$d(x_{n-1}, x) \leq \beta\, d(x_{n-2}, x)$

$\vdots \qquad \vdots$

$d(x_1, x) \leq \beta\, d(\underbrace{x_0}_{n-n}, x)$

$\therefore \boxed{d(x_n, x) \leq \beta^n (x_0, x)}$

(iv)

in (1.69) we have:

$\forall_{m,n \in \mathbb{N}}\ d(x_n, x_{n+m}) \leq \frac{\beta^n}{1-\beta} d(x_0, x_1) \Rightarrow \forall_{n \in \mathbb{N}}\ \underline{\lim}\ d(x_n, x_{n+m}) \leq \frac{\beta^n}{1-\beta} d(x_0, x_1)$

take $m \to \infty$ then $x_{n+m} \to x \Rightarrow d(x_n, x_{n+m}) \to d(x_n, x)$ $\therefore \lim d(x_n, x_{n+m}) = \underline{\lim}\ d(x_n, x_{n+m})$

$\therefore d(x_n, x) \leq \frac{\beta^n}{1-\beta} d(x_0, x_1)$

(v) in (1.68) we have:

$\forall_{m,n \in \mathbb{N}}\ d(x_n, x_{n+m}) \leq \frac{1-\beta^m}{1-\beta} d(x_n, x_{n+1}) \leftrightarrow \forall_{m,n \in \mathbb{N}} \left(\frac{1}{1-\beta^m}\right) d(x_n, x_{n+m}) \leq \frac{1}{1-\beta} d(x_n, x_{n+1})$

$\Rightarrow \forall_{n \in \mathbb{N}}\ \underline{\lim} \left(\frac{1}{1-\beta^m}\right) d(x_n, x_{n+m}) \leq \frac{1}{1-\beta} d(x_n, x_{n+1})$

take $d(x_n, x_{n+m}) \to d(x_n, x)$ $\because \lim d(x_n, x_{n+m}) = d(x_n, x)$; $\lim \left(\frac{1}{1-\beta^m}\right) = 1$ $\therefore \underset{\underline{\lim} = \lim}{\underset{\parallel}{\lim}} \left(\frac{1}{1-\beta^m}\right) d(x_n, x_{n+m}) = d(x_n, x) \leq \frac{1}{1-\beta} d(x_n, x_{n+1})$

$\therefore d(x_n, x) \leq \frac{1}{1-\beta} d(x_n, x_{n+1})$

(vi) we want to show:

$\underbrace{\frac{d(x_0, x_1)}{1+\beta} \leq d(x_0, x)}\ \leq \frac{d(x_0, x_1)}{1-\beta}$

By triangle inequality, $d(x_0, x_1) \leq d(x_0, x) + \underbrace{d(x, x_1)}_{= d(x_1, x) \leq \beta\, d(x_0, x)} \leq d(x_0, x) + \beta\, d(x_0, x) = (1+\beta)\, d(x_0, x)$

/* in (ii): $\forall_{n \in \mathbb{N}}\ d(x_{n+1}, x) \leq \beta\, d(x_n, x)$ */

from (iv) we have $\forall_{n \in \mathbb{N}}\ d(x_n, x) \leq \beta^n d(x_0, x_1)/(1-\beta)$

$n = 0 \Rightarrow d(x_0, x) \leq d(x_0, x_1)/(1-\beta)$

■

• Theorem 1.49. (Banach-Picard variant)

[(X,d): complete metric space

$T: X \to X$,

$\exists_{(\beta_n)_{n\in\mathbb{N}}}$: summable sequence in $\mathbb{R}_+$ $\left(\forall_{x,y\in X} \forall_{n\in\mathbb{N}} \; d(T^n x, T^n y) \leq \beta_n d(x,y)\right)$ /* note the operator T is quite complicated here */

$T^n = \begin{cases} \tilde{I}, & \text{if } n=0 \\ n\text{-fold composition of } T, & \text{if } n>0 \end{cases}$ identity operator

$x_0 \in X$

$\forall_{n\in\mathbb{N}} \; x_{n+1} = Tx_n, \; \alpha_n = \sum_{k=n}^{+\infty} \beta_k$

]

$\Rightarrow$

$\exists_{x\in X} \begin{cases} \text{(i) } x: \text{unique fixed point of } T \\ \text{(ii) } x_n \to x \\ \text{(iii) } \forall_{n\in\mathbb{N}} \; d(x_n, x) \leq \alpha_n d(x_0, x_1) \end{cases}$

11/19/2015 9:16 PM /* Proof strategy: similar to Theorem 1.48 */

Proof:

given $\forall_{n\in\mathbb{N}} \; x_{n+1} = Tx_n, \; \alpha_n = \sum_{k=n}^{\infty} \beta_k$

Let's apply triangle inequality:

$\forall_{m\in\mathbb{N}} \forall_{n\in\mathbb{N}} \; d(x_n, x_{n+m}) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+m-1}, x_{n+m})$

$= \sum_{k=n}^{n+m-1} d(\underbrace{x_k}_{T^k x_0}, \underbrace{x_{k+1}}_{T^{k+1}x_0})$

$= T^k(Tx_0)$

$= T^k(x_1)$

$= \sum_{k=n}^{n+m-1} \underbrace{d(T^k x_0, T^k x_1)}_{\leq \beta_k d(x_0, x_1)}$

/* given, $x_{n+1} = Tx_n$

$x_1 = Tx_0$

$x_2 = Tx_1 = T^2 x_0$

$x_3 = Tx_2 = T^3 x_0$

$\vdots$

$x_k = T^k x_0$

*/

/* By given, $\forall_{x,y\in X} \forall_{n\in\mathbb{N}} \; d(T^n x, T^n y) \leq \beta_n d(x,y)$ */

$\leq \sum_{k=n}^{n+m-1} \beta_k d(x_0, x_1) = d(x_0, x_1) \underbrace{\sum_{k=n}^{n+m-1} \beta_k}_{\leq \sum_{k=n}^{\infty} \beta_k = \alpha_n} \leq d(x_0, x_1)\alpha_n = \alpha_n d(x_0, x_1) \dots (1.72)$

$\Rightarrow \boxed{\alpha_0 = \sum_{k=0}^{\infty} \beta_k}$

Proof of (i) & (ii):

By given $(\beta_n)_{n\in\mathbb{N}}$: summable $\Rightarrow \exists_{b\in\mathbb{R}} \; \tilde{a}_k = \sum_{j=0}^{k} \beta_j \to b \leftrightarrow$ as $k \to +\infty$, $\tilde{a}_k = \sum_{j=0}^{k} \beta_j \to b$ now $\alpha_0 = \sum_{k=0}^{\infty} \beta_k = \underbrace{\sum_{k=0}^{n} \beta_k}_{\tilde{a}_n} + \underbrace{\sum_{k=n+1}^{\infty} \beta_k}_{\alpha_{n+1}} \leftrightarrow \alpha_0 = \tilde{a}_n + \alpha_{n+1}$; now as $n \to \infty$, $\tilde{a}_n \to \alpha_0$ $\therefore \alpha_0 = \alpha_0 + \alpha_{n+1}$, as $n \to \infty$

$\lim \tilde{a}_k = \sum_{j=0}^{\infty} \beta_j = b = \alpha_0$

$\Rightarrow \boxed{\alpha_n \to 0 \text{ as } n \to +\infty}$

from (1.72) we have:

$\forall_{m,n\in\mathbb{N}} \; d(x_n, x_{n+m}) \leq \alpha_n d(x_0, x_1)$

As $n \to \infty$,

$d(x_n, x_{n+m}) \to 0$ /* as $\alpha_n \to 0$ as $n \to \infty$ */

/* Recall, $(x_n)_{n\in\mathbb{N}}$: Cauchy sequence $\overset{\text{def}}{\Leftrightarrow}$ $d(x_m, x_n) \to 0$ as $m,n \to \infty$ */

$\therefore (x_n)_{n\in\mathbb{N}}$: Cauchy sequence

Because (X,d): complete metric space, so every Cauchy sequence converges

$(x_n)_{n\in\mathbb{N}}$ converges to some point x in X. i.e., $x_n \to x$

from given, set $n=1$. then, $\forall_{x\in X} \forall_{y\in Y} \; d(Tx, Ty) \leq \beta_1 d(x,y)$

$\Rightarrow T$: Lipschitz continuous

now T: Lipschitz continuous /* recall fact 1.19. $\left(\begin{array}{l} X: \text{Hausdorff space} \\ Y: \text{Hausdorff space} \\ T: X \to Y, \text{ continuous at } x \end{array}\right) \Rightarrow \forall_{(x_a)_{a\in A}: \text{net in } X, \text{ converges to } x}$ $Tx_a \to Tx$ */

$\therefore \underbrace{Tx_n}_{x_{n+1}} \to Tx$ /* $\because x_{n+1} = Tx_n$ */

$x_{n+1} \to Tx$

but $x_{n+1} \to x$ $\Big\} \Rightarrow Tx = x \leftrightarrow x \in \text{Fix}T$.

/* In a Hausdorff space convergence of a sequence is always to a unique point */

• (Uniqueness)

Consider $y \in \text{Fix}T \setminus \{x\} \leftrightarrow x \neq y \wedge Ty = y$

now in given, put $\underbrace{x, y}_{\text{fixed points}}$ $\quad \forall_{n\in\mathbb{N}} \quad d(\underbrace{T^n x}_{x}, \underbrace{T^n y}_{y}) \leq \beta_n d(x,y)$

/* $T^n x = T..(\underbrace{Tx}_{x}) = \dots = x$

$T^n y = T\dots(\underbrace{Ty}_{y}) = y$ */

$\Leftrightarrow \forall_{n\in\mathbb{N}} \quad d(x,y) \leq \beta_n d(x,y)$

$\Leftrightarrow \forall_{n\in\mathbb{N}} \quad 1 \leq \beta_n \Leftrightarrow \lim_{n\to\infty} \beta_n \geq 1$, but if so, then $\sum_{n=0}^{\infty} \beta_n \to \infty$, but this is not possible as $(\beta_n)_{n\in\mathbb{N}}$: summable $\Rightarrow$ contradiction

$\therefore y = x$.

(iii) in 1.72 we have shown:

$\forall_{m\in\mathbb{N}} \forall_{n\in\mathbb{N}} \quad d(x_n, x_{n+m}) \leq \alpha_n d(x_0, x_1)$

But $\lim_{m\to\infty} d(x_n, \underbrace{x_{n+m}}_{\downarrow \; x \text{ as } m\to\infty}) = d(x_n, x) \leq \alpha_n d(x_0, x_1) \quad \forall_{n\in\mathbb{N}}$

$\therefore \forall_{n\in\mathbb{N}} \quad d(x_n, x) \leq \alpha_n d(x_0, x_1)$

$\blacksquare$

Chapter 1: Part 5

February 27, 2017 2:20 PM

Lemma 1.23: **

[X: Hausdorff space; $f: X \to [-\infty, +\infty]$; $x \in X$]

$$\underline{\lim}_{y \to x} f(y) = \min_{\substack{(x_a)_{a\in A}: \text{net in } X, \\ x_a \to x}} \underline{\lim} f(x_a)$$

Proof:

Define $N(x) = \{(x_a)_{a\in A} \subseteq X : x_a \to x\}$

take $(x_a)_{a\in A} \in N(x)$

$$\forall_{a\in A}\ \mu_a = \inf\{f(x_b) \mid b \succeq a\} = \inf_{b\in A: b\succeq a} f(x_b)$$

/* **
Fact 1.15: ($\underline{\lim}$, $\overline{\lim}$ in different light) /* conventionally this holds for bounded net */
[$(\xi_a)_{a\in A}$: net in $[-\infty, +\infty]$]
(i) $\underline{\lim}\, \xi_a = \lim \inf_{b\in A: b\succeq a} \xi_b$; $\overline{\lim}\, \xi_a = \lim \sup_{b\in A: b\succeq a} \xi_b$
(ii) $(\xi_a)_{a\in A}$: possesses subnets that converge to $\underline{\lim}\, \xi_a$ and $\overline{\lim}\, \xi_a$ respectively.
(iii) $(\xi_a)_{a\in A}$: converges $\Leftrightarrow \underline{\lim}\, \xi_a = \overline{\lim}\, \xi_a$ /* in this case, $\underline{\lim}\, \xi_a = \overline{\lim}\, \xi_a$ */
*/

now

$$\underline{\lim} f(x_a) = \lim \underbrace{\inf_{b\in A: b\succeq a} f(x_b)}_{\mu_a}$$

$$= \lim \mu_a$$

// so, $(\mu_a)_{a\in A}$: converging net

(1)

recall $\mathcal{V}(x)$ = set of all neighborhoods (open) of x

now $x_a \to x$ /* recall that:

• Convergence of a net
[$(x_a)_{a\in A}$: net in Hausdorff space X]
$x_a \to x \overset{\text{def}}{\Leftrightarrow} (x_a)_{a\in A}$ lies eventually in every neighborhood of x $\overset{\text{def}}{\Leftrightarrow}$
$\forall_{V_\varepsilon(x)\in\mathcal{V}(x)}\ \exists_{b\in A}\ \forall_{a\in A: a\succeq b}\ x_a \in V_\varepsilon(x)$
*/

using this:

$\forall_{V\in\mathcal{V}(x)}\ \exists a_V \in A\ \forall_{\tilde{a}\in A: \tilde{a}\succeq a_V}\ x_{\tilde{a}} \in V$ then

so, $\forall_{a\succeq a_V}\ \mu_a = \inf_{b\in A: b\succeq a} f(x_b)$ // $a \succeq a_V$, $b \succeq a \Rightarrow b \succeq a \succeq a_V \Rightarrow x_b \in V$

// so, $\{x_b \mid b\succeq a, b\in A\} \subseteq V$

$= \inf f(x_b)$

s.t. $x_b \in \{\tilde{x}_b \mid b\succeq a, b\in A\}$ /* as $\forall \{\tilde{x}_b \mid b\succeq a, b\in A\}$ we have $\inf f(V) \leq \inf f(x_b)$ s.t. $x_b \in \{\tilde{x}_b \mid b\succeq a, b\in A\}$

$\geq \inf f(x_b) = \inf f(V)$

s.t. $x_b \in V$

$\therefore \forall_{a\succeq a_V}\ \mu_a \geq \inf f(V)$ /* as $(\mu_a)_{a\in A}$: converging net $\Rightarrow (\mu_a)_{a\succeq a_V}$: subnet also must be converging to the same point

$\Rightarrow \lim_{a\succeq a_V} \mu_a \geq \inf f(V)$ /* using net version • $\forall_{n\in\mathbb{N}}\ a_n \leq a \Leftrightarrow \overline{\lim}\, a_n \leq a$ */

(2)

from (1), (2):

$$\forall_{\substack{(x_a)_{a\in A}: \\ x_a\to x}}\ \forall_{V\in\mathcal{V}(x)}\ \underbrace{\underline{\lim} f(x_a)}_{\text{function of } x_a} = \lim \mu_a = \lim_{a\succeq a_V} \mu_a \geq \inf f(V)$$

$$\Rightarrow \forall_{(x_a)_{a\in A}}\ \forall_{V\in\mathcal{V}(x)}\ \underbrace{\underline{\lim} f(x_a) + I_{x_a\to x}}_{\text{function of } x_a} \geq \underbrace{\inf f(V)}_{\text{function of } V\in\mathcal{V}(x)}$$

indicator function on the condition $x_a \to x$

$$\Rightarrow \inf\{\underline{\lim} f(x_a) \mid x_a \to x\} \geq \sup_{V\in\mathcal{V}(x)} \inf f(V)$$

// using: Fact 1.7.2: [$f, g: Y \to [-\infty, +\infty]$, $\forall_{x,y\in Y}\ f(x) \leq g(y)$] $\sup f(Y) \leq \inf g(Y)$

(1.37)

define: $B = \{(y, V) \mid y\in V, V\in\mathcal{V}(x)\}$:

↳ we impose order on the set elements, $\underbrace{(\tilde{y}, \tilde{V})}_{\in\mathcal{V}(x)}$, $\underbrace{(\tilde{z}, \tilde{W})}_{\in\mathcal{V}(x)}$ as follows: $(\tilde{y}, \tilde{V}) \leq (\tilde{z}, \tilde{W}) \Leftrightarrow \tilde{W} \supseteq \tilde{V}$

↳ this makes B directed

↳ which allows us to define a net on B. recall that for a net $(x_a)_{a\in A}$ directed set is required

now denote: $b = (y, V)$

$\forall_{b=(y,V)\in B}\ x_b = y \Rightarrow x_b \in V, V\in\mathcal{V}(x) \Rightarrow \boxed{x_b \to x}$ // as this is a subnet of $(x_a)_{a\in A}$ and $x_a \to x$ using Fact 1.9. * **

[$(x_a)_{a\in A}$: net in Hausdorff space X, $x_a \to \tilde{x} \in X$

$(x_{k(n)})_{b\in A}$: subnet of $(x_a)_{a\in A}$] $\Rightarrow x_{k(n)} \to x$

$y\in V$, $V\in\mathcal{V}(x)$; Similarly $\forall_{c=(z,W)}\ x_c = z$

also: $\underline{\lim} f(x_b) = \sup_{b\in B} \inf_{c\succeq b} f(x_c)$ // By definition: $\underline{\lim}\, \xi_a = \sup_{a\in A} \inf_{b\in A: b\succeq a} \xi_b$ also

$b \in B$ ↕ $(y, V): y\in V, V\in\mathcal{V}(x)$

/* now $b = (y, V): y\in V, V\in\mathcal{V}(x)$

$c = (z, W): z\in W, W\in\mathcal{V}(x)$

$c \succeq b \Leftrightarrow (z, W) \succeq (y, V) \Leftrightarrow W \supseteq V$ */

$$= \sup_{y\in V: V\in\mathcal{V}(x)} \inf_{z\in W\subseteq V} f(z)$$ /* using (•), (•), (•) */

$$= \sup_{y\in V: V\in\mathcal{V}(x)} \underbrace{\inf f(W)}_{\leq \inf f(V)}$$

// as $W \supseteq V \Rightarrow \inf f(W) \leq \inf f(V)$

$$\leq \sup_{y \in V : V \in \mathcal{V}(x)} \inf f(V) = \sup_{V \in \mathcal{V}(x)} \inf f(V)$$

$$\therefore \forall_{(x_b)_{b \in B} : x_b \to x} \quad \underline{\lim}\, f(x_b) \leq \sup_{y \in V : V \in \mathcal{V}(x)} \inf f(V)$$

$$\Rightarrow \quad \inf \{ \underline{\lim}\, f(x_b) \mid x_b \to x,\ b \in B \} \leq \sup_{V \in \mathcal{V}(x)} \inf f(V) \quad \ldots (1.38)$$

Finally recall that:

$$\underline{\lim}_{y \to x} f(y) = \sup_{V \in \mathcal{V}(x)} \inf f(V) \quad \ldots (1.39)$$

combining (1.39), (1.37), (1.38) we have:

$$\underline{\lim}_{y \to x} f(y) = \min_{(x_a)_{a \in A} : x_a \to x} \underline{\lim}\, f(x_a)$$

$\blacksquare$

Chapter 2: Part 1

Sunday, November 22, 2015 10:14 PM

$\mathcal{H}$: real Hilbert space with inner product $\langle\cdot|\cdot\rangle$,
norm $\|\cdot\|$, distance d, identity 1 /* in the textbook it is $\mathbb{1}$ */

$\forall_{x,y\in\mathcal{H}}\ \|x\|=\sqrt{\langle x|x\rangle}$ $\quad d(x,y)=\|x-y\|$

2.1. Notation and examples.

$\underbrace{C^{\perp}}_{\text{orthogonal complement of subset } C \text{ of } \mathcal{H}} = \{u\in\mathcal{H} \mid \forall_{x\in C}\ \langle x|u\rangle=0\}$

$\underbrace{C}_{\subseteq\mathcal{H}}$: orthonormal basis of $\mathcal{H}$ $\overset{\text{def}}{\Leftrightarrow}$ $\overline{\operatorname{span} C}=\mathcal{H}$ /* $\overline{\operatorname{span}}\, C$: smallest closed linear subspace of $\mathcal{X}$ containing C */

$\mathcal{H}$: seperable Hilbert space $\overset{\text{def}}{\Leftrightarrow}$ $\mathcal{H}$ has countable orthonormal basis

[

$(x_i)_{i\in I}$: family of vectors in $\mathcal{H}$ /* it is a net */

$\mathcal{I}$: class of nonempty finite subsets of I, $\subseteq I$

] $(x_i)_{i\in I}$: summable $\overset{\text{def}}{\Leftrightarrow}$ $\exists_{x\in\mathcal{H}}\ \underbrace{\left(\sum_{i\in J} x_i\right)_{J\in\mathcal{I}} \text{ converges to } x}_{\left(\forall_{\varepsilon\in\mathbb{R}_{++}}\ \exists_{K\in\mathcal{I}}\ \forall_{J\in\mathcal{I}:\,J\supseteq K}\ \|x-\sum_{i\in J} x_i\|\leq\varepsilon\right)}$

2.2. Basic Identities and inequalities:

Fact 2.10. (Cauchy-Schwarz)

[$x,y\in\mathcal{H}$]

- $|\langle x|y\rangle|\leq\|x\|\,\|y\|$
- $\langle x|y\rangle=\|x\|\,\|y\| \leftrightarrow \exists_{\alpha\in\mathbb{R}_+}\left(x=\alpha y \ \vee\ y=\alpha x\right)$

Lemma 2.11.

[$x,y,z\in\mathcal{H}$]

(i) $\|x+y\|^2=\|x\|^2+2\langle x|y\rangle+\|z\|^2$

(ii) Parallelogram identity:

$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2$

(iii) Polarization identity:

$4\langle x|y\rangle=\|x+y\|^2-\|x-y\|^2$

(iv)

$\|x-y\|^2=2\|z-x\|^2+2\|z-y\|^2-4\left\|z-\frac{x+y}{2}\right\|^2$

Lemma 2.12. /* Different ways of expressing orthogonal and obtuse vectors */

[$x,y\in\mathcal{H}$]

(i) $\langle x|y\rangle\leq 0 \leftrightarrow \forall_{\alpha\in\mathbb{R}_+}\ \|x\|\leq\|x-\alpha y\| \leftrightarrow \forall_{\alpha\in[0,1)}\ \|x\|\leq\|x-\alpha y\|$

(ii)

$x\perp y \leftrightarrow \forall_{\alpha\in\mathbb{R}}\ \|x\|\leq\|x-\alpha y\| \leftrightarrow \forall_{\alpha\in[-1,1]}\ \|x\|\leq\|x-\alpha y\|$

Lemma 2.13.

[

$(x_i)_{i\in I}, (u_i)_{i\in I}$: finite families in $\mathcal{H}$

$(\alpha_i)_{i\in I}$: family in $\mathbb{R}$, $\sum_{i\in I} \alpha_i = 1$

]

(i)

$$\left\langle \sum_{i\in I} \alpha_i x_i \,\middle|\, \sum_{j\in I} \alpha_j u_j \right\rangle + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i \alpha_j}{2} \langle x_i - x_j \mid u_i - u_j \rangle = \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle$$

(ii)

$$\left\| \sum_{i\in I} \alpha_i x_i \right\|^2 + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i \alpha_j}{2} \|x_i - x_j\|^2 = \sum_{i\in I} \alpha_i \|x_i\|^2$$

2015-12-03 10:06 AM

Proof:

(i)

$$2\left\langle \sum_{i\in I} \alpha_i x_i \,\middle|\, \sum_{j\in I} \alpha_j u_j \right\rangle$$

/* $\langle \sum_i a_i \mid b \rangle = \langle a_1 + a_2 + \dots + a_{final} \mid b \rangle = \langle a_1 | b \rangle + \dots + \langle a_{final} | b \rangle$

$= \sum_i \langle a_i | b \rangle$

say, $b = \sum_j c_j$ then $\langle a_i | \sum_j c_j \rangle = \langle \sum_j c_j \mid a_i \rangle = \sum_j \langle c_j | a_i \rangle$

$\therefore \langle \sum_i a_i \mid \sum_j c_j \rangle = \sum_i \sum_j \langle a_i | c_j \rangle$ */

$$= 2 \sum_{i\in I}\sum_{j\in I} \langle \alpha_i x_i \mid \alpha_j u_j \rangle$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \underbrace{2\langle x_i | u_j \rangle}_{\langle x_i | u_j \rangle + \underbrace{\langle x_i | u_j \rangle}_{\langle u_j | x_i \rangle}}$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i | u_j \rangle + \underbrace{\sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i | u_j \rangle}_{\text{/* interchanging index } i,j \text{ */}}$$

$$= \sum_{j\in I}\sum_{i\in I} \alpha_j \alpha_i \langle x_j | u_i \rangle = \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_j | u_i \rangle$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \left(\langle x_i | u_j \rangle + \langle x_j | u_i \rangle \right)$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \left(\langle x_i | u_i \rangle + \langle x_j | u_j \rangle - \langle x_i - x_j \mid u_i - u_j \rangle \right)$$

/* $\langle x_i - x_j \mid u_i - u_j \rangle = \langle x_i | u_i \rangle - \langle x_i | u_j \rangle - \langle x_j | u_i \rangle + \langle x_j | u_j \rangle$

$\therefore \langle x_i | u_j \rangle + \langle x_j | u_i \rangle = \langle x_i | u_i \rangle + \langle x_j | u_j \rangle - \langle x_i - x_j \mid u_i - u_j \rangle$ */

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i | u_i \rangle$$

$$+ \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_j | u_j \rangle$$

$$- \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i - x_j \mid u_i - u_j \rangle$$

/* $\sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i | u_i \rangle = \sum_{j\in I} \alpha_j \left(\sum_{i\in I} \alpha_i \langle x_i | u_i \rangle \right) = \left(\sum_{i\in I} \alpha_i \langle x_i | u_i \rangle \right) \underbrace{\left(\sum_{j\in I} \alpha_i \right)}_{=1} = \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle$

similarly, $\sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_j | u_j \rangle = \sum_{j\in I} \alpha_j \langle x_j | u_j \rangle$

$$= \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle + \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle - \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i - x_j \mid u_i - u_j \rangle$$

$$= 2\sum_{i\in I} \alpha_i \langle x_i | u_i \rangle - \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \langle x_i - x_j \mid u_i - u_j \rangle \quad \dots (2.15)$$

$$\therefore \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle = \left\langle \sum_{i\in I} \alpha_i x_i \,\middle|\, \sum_{j\in I} \alpha_j u_j \right\rangle + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i \alpha_j}{2} \langle x_i - x_j | u_i - u_j \rangle$$

(ii) In (2.15) we put : $(u_i)_{i\in I} := (x_i)_{i\in I}$ then we have

$$2\left\langle \sum_{i\in I} \alpha_i x_i \,\middle|\, \underbrace{\sum_{j\in I} \alpha_j u_j}_{\sum_{j\in I} \alpha_j x_j \,=\, \sum_{i\in I} \alpha_i x_i} \right\rangle = 2\sum_{i\in I} \alpha_i \underbrace{\langle x_i | u_i \rangle}_{\langle x_i | x_i \rangle = \|x_i\|^2} - \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \underbrace{\langle x_i - x_j | \underset{x_i}{u_i} - \underset{x_j}{u_j} \rangle}_{\|x_i - x_j\|^2}$$

$$\left\langle \left(\sum_{i\in I} \alpha_i x_i\right) \,\middle|\, \left(\sum_{i\in I} \alpha_i x_i\right) \right\rangle = \left\| \sum_{i\in I} \alpha_i x_i \right\|^2$$

$$\leftrightarrow \quad 2\left\| \sum_{i\in I} \alpha_i x_i \right\|^2 = 2\sum_{i\in I} \alpha_i \|x_i\|^2 - \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \|x_i - x_j\|^2$$

$$\leftrightarrow \quad \left\| \sum_{i\in I} \alpha_i x_i \right\|^2 + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i \alpha_j}{2} \|x_i - x_j\|^2 = \sum_{i\in I} \alpha_i \|x_i\|^2 \qquad \textcircled{:)}$$

Corollary : 2.14.

[$x, y \in H$, $\alpha \in \mathbb{R}$]

$$\|\alpha x + (1-\alpha) y\|^2 + \alpha(1-\alpha)\|x - y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2$$

2.3. Linear operators and functionals.

X, Y : normed vector space

$B(X,Y) = \{T : X \to Y \mid T : \text{linear, continuous}\}$ → a normed vector space

/* $B(X,Y)$: set of all linear and continuous operator */

$B(X) = B(X,X)$

norm of linear and continuous operator :

$\forall_{T \in B(X,Y)} \quad \|T\| = \sup \|T(B(0;1))\| = \sup_{\substack{x\in X \\ \|x\|\le 1}} \|Tx\|$ /* if T is a matrix then $\|T\|$ is the usual matrix norm */

$B(X,Y)$: Banach space if Y is a Banach space

• Fact 2.15.

[X, Y : real normed vector space

$\underset{:X\to Y}{T}$: linear]

T: continuous at a point in X ↔ T: Lipschitz continuous

/* for a linear operator continuity and Lipschitz continuity are same */

• Lemma 2.16. (uniform boundedness) /* called Banach-Steinhaus theorem */

[X: real Banach space, /* Banach space is a $\underbrace{\text{complete}}_{\text{every cauchy sequence converges}}$ normed vector space */

Y: real normed vector space

$\underbrace{(T_i)_{i\in I}}$: family of operators in $\mathcal{B}(X,Y)$.

T_i : pointwise bounded $\leftrightarrow \forall_{x\in X} \sup_{i\in I} \|T_i x\| < \infty$

]

$\Rightarrow$

$\sup_{i\in I} \|T_i\| < +\infty$ /* seems intuitive though */

[Fact 2.17: Riesz-Frechet representation theorem]

Fact 2.17.

• Riesz-Fréchet representation theorem. /* Says that any continuous linear functional on the real Hilbert space can be identified with a vector in that Hilbert space */

return to [Gateaux gradient]

$$f\in\mathcal{B}(\mathcal{H},\mathbb{R}) \Rightarrow \left(\underbrace{\exists_{!u\in\mathcal{H}}}_{\text{means unique}} \forall_{x\in\mathcal{H}} f(x)=\langle x|u\rangle\right) \wedge \left(\|f\|=\|u\|\right)$$

[Adjoint operator]

return [Frechet differentiable]

* Adjoint operator.

$\mathcal{K}$: real Hilbert space

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$

$\underbrace{T^*}_{\in\mathcal{B}(\mathcal{K},\mathcal{H})}$: adjoint of T $\overset{\text{def}}{\leftrightarrow} \forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{K}} \langle Tx|y\rangle = \langle x|T^*y\rangle$

/* adjoint operator is a unique operator */

* Properties of adjoint operator.

[$\mathcal{K}$: real Hilbert space

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$

$\ker T = \{x\in\mathcal{H} : Tx=0\}$

] $\Rightarrow$

(i) $T^{**}=T$ /* Double adjoint gives the same operator */

(ii) $\|T^*\| = \|T\| = \sqrt{\|T^*T\|}$ /* linear continuous operator and its adjoint has same norm */

(iii) $(\ker T)^\perp = \overline{\operatorname{ran}}\, T^*$

(iv) $(\operatorname{ran} T)^\perp = \ker T^*$

(v) $\ker T^*T = \ker T$

$\overline{\operatorname{ran}}\, TT^* = \overline{\operatorname{ran}}\, T$

} /* Finding range closure and kernel of a linear continuous operator using its adjoint */

Fact 2.19.

[$\mathcal{K}$: real Hilbert space,

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$]

$\operatorname{ran} T$: closed $\leftrightarrow \operatorname{ran} T^*$: closed $\leftrightarrow \operatorname{ran} TT^*$: closed

$\leftrightarrow \operatorname{ran} TT^*$: closed $\leftrightarrow \exists_{\alpha\in\mathbb{R}_{++}} \forall_{x\in(\ker T)^\perp} \|Tx\| \geq \alpha\|x\|$

• Hyperplane, halfspace etc.

$[u \in \mathcal{H} \setminus \{0\}, \eta \in \mathbb{R}]$

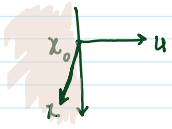
(closed hyperplane in $\mathcal{H}$) $= \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}$

(closed halfspace in $\mathcal{H}$) $= \{x \in \mathcal{H} : \langle x | u \rangle \leq \eta\}$ with outer normal u

If in $\mathbb{R}^n$: $u^T x \leq \eta = u^T x_0$

$\rightarrow u^T (x - x_0) \leq 0$

(open halfspace in $\mathcal{H}$ with outer normal u) $= \{x \in \mathcal{H} : \langle x | u \rangle < \eta\}$



d_C : (distance function to $C = \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}$, $\mathcal{H} \to \mathbb{R}_+$,

$$x \mapsto \underbrace{\frac{|\langle x | u \rangle - \eta|}{\|u\|}}$$

this is the minimum distance of point x from the hyperplane $C = \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}$

Chapter 2: Part 2

Wednesday, November 25, 2015 10:27 PM

2.4. Strong and weak topology:

$(\mathcal{H}, d)$: strong topology
(or norm topology)

$(x_a)_{a\in A}$ in $\mathcal{H}$ converges strongly to a point x $\overset{def}{\Leftrightarrow}$ $\underbrace{\|x_a - x\| \to 0}_{x_a \to x}$

Fact 2.21.

[U: closed linear subspace of $\mathcal{H}$
V: " " " " $\mathcal{H}$, has finite dimension or codimension
]

$U+V$: closed linear subspace.

Weak topology: The family of all finite intersections of open halfspaces of $\mathcal{H}$ forms the base of the weak topology of $\mathcal{H}$. the resulting topological space ($\mathcal{H}$, weak topology of $\mathcal{H}$) is called $\mathcal{H}_{weak}$ / we consider only sets in the weak topology in our consideration while dealing with $\mathcal{H}_{weak}$ */

topology: family of subsets of X that contains $X, \emptyset$, arbitrary union and finite intersections of elements of X

/* base. $\underbrace{B}_{\text{subfamily of } T}$: base of T $\overset{def}{\Leftrightarrow}$ $\forall_{x\in X}$ $\forall_{V\in\mathcal{V}(x)}$ $\exists_{S\in B}$ $(x\in S \wedge S\subseteq V)$

importance of base:

B: base of $T \Rightarrow \forall_{S\in T}$ S = union of elements of B */

2015-11-25 7:44 PM **Lemma 2.23.** $\mathcal{H}^{weak}$: Hausdorff space. [H^weak is Hausdorff] [12/3/2015 11:01 AM] /* $\overset{\text{topology of } X}{T}$: Hausdorff space $\Leftrightarrow \forall_{x,y\in X: x\neq y}$ $\exists_{V\in\mathcal{V}(x), W\in\mathcal{V}(y): V,W \text{ elements of base } T}$ $V\cap W=\emptyset$ */

/* What does this mean?
Recall the definition of Hausdorff space: for any two distinct points $x_1, x_2 \in X$ there exists $V_1\in\mathcal{V}(x_1), V_2\in\mathcal{V}(x_2)$ such that $V_1\cap V_2\neq\emptyset$
now if we confine the neighborhoods are associated with open halfspaces of $\mathcal{H}$ (more specifically union of finite intersections of $\mathcal{H}$, i.e., the weak topology of $\mathcal{H}$), and show that the Hausdorff condition holds, then we have $\mathcal{H}^{weak} \equiv (\mathcal{H}$, weak topology of $\mathcal{H})$ is a Hausdorff space */

Proof.

x, y: distinct points in $\mathcal{H}$, our goal is to show that the elements of $\mathcal{H}_{weak}$ = union of members of its base construct disjoint neighborhoods for different point.

$u := x-y$

$w = \frac{x+y}{2}$

Consider the set: $\{z\in\mathcal{H} \mid \langle z-w|u\rangle > 0\}$: this is a neighborhood of x, also a member of $\mathcal{H}_{weak}$'s base and hence $\mathcal{H}_{weak}$ because:

$$\langle \overset{:=x}{z} - w | u\rangle = \langle x - \tfrac{x+y}{2} \mid x-y\rangle = \langle \tfrac{x-y}{2} \mid x-y\rangle = \tfrac{1}{2}\langle x-y|x-y\rangle = \tfrac{1}{2}\|x-y\|^2 > 0 \quad (\because x\neq y)$$

$\therefore x \in \{z\in\mathcal{H} \mid \langle z-w|u\rangle > 0\}$

Similarly the set: $\{z\in\mathcal{H} \mid \langle z-w|u\rangle < 0\}$ is a neighborhood of y

$$\text{as } \langle \underbrace{z}_{:=y} - w | u\rangle = \langle y - \tfrac{x+y}{2} \mid x-y\rangle = \langle \tfrac{1}{2}(y-x) \mid x-y\rangle = \langle (-\tfrac{1}{2})(x-y) \mid (x-y)\rangle$$

$$= (-\tfrac{1}{2})\langle x-y|x-y\rangle = -\tfrac{1}{2}\|x-y\|^2 < 0$$

Clearly $\{z\in\mathcal{H} : \langle z-w|u\rangle > 0\}$ and $\{z\in\mathcal{H} : \langle z-w|u\rangle < 0\}$ are disjoint, so we have two disjoint neighborhoods of any two distinct points (definition of a Hausdorff space)

$\therefore \mathcal{H}^{weak}$: Hausdorff space. ∎

*Weakly open. (A subset of $\mathcal{H}$): weakly open $\overset{def}{\Leftrightarrow}$ it is union of finite intersections of open half-spaces

*Weakly convergent net.

$\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}}$: converges weakly to x $\overset{def}{\Leftrightarrow}$ $\forall_{u\in\mathcal{H}}$ $\langle x_a|u\rangle \to \langle x_a|x\rangle$ /* x is called the weak limit of $(x_a)_{a\in A}$ */

/* $x_a \to x$ */

*Weakly closed set. weakly compact set

$\underbrace{C}_{\text{subset of } \mathcal{H}}$: weakly closed $\overset{def}{\Leftrightarrow}$ weak limit of every weakly convergent net in C is also in C.

$\underbrace{C}_{\text{subset of } \mathcal{H}}$: weakly compact $\overset{def}{\Leftrightarrow}$ every net in C has a weak cluster point in C

/* Cluster point: $x: \in X$, cluster point of $(x_a)_{a\in A}$ $\overset{def}{\Leftrightarrow}$ $(x_a)_{a\in A}$ lies frequently in every neighborhood of x

$\Leftrightarrow \forall_{V\in\mathcal{V}(x)}$ $\forall_{b\in A}$ $\exists_{a\in A: a\geq b}$ $x_a \in V$ */

[D: nonempty set of $\mathcal{H}$,

$\mathcal{H}$: real Hilbert space.

$T: D \to \mathcal{H}$

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

]

T: weakly continuous $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } D}} : x_a \rightharpoonup x \wedge x \in D \quad Tx_a \rightharpoonup Tx$ /* $\underline{\lim}\, \xi_a = \left[\left(a \in A \,//\, (\xi_b : b \geq a) \,//\, \inf_{b\in A} (\cdot) \right) \,//\, \sup_{a\in A} (\cdot) \right]$ */

f: weakly lower semicontinuous at $x \in \mathcal{H}$ $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}}} : x_a \rightharpoonup x \quad \underline{\lim}\, f(x_a) \geq f(x)$

Lemma 2.35. [Norm in H: weakly lower semicontinuous] [12/3/2015 11:37 AM]

Norm of $\mathcal{H}$: weakly lower semicontinuous, i.e.,

2016-12-05 11:04 AM

$\forall_{x\in\mathcal{H}} \forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}}} : x_a \rightharpoonup x \quad \underline{\lim}\, \|x_a\| \geq \|x\|$

Proof:

take $x \in \mathcal{H}$, $(x_a)_{a\in A} : x_a \rightharpoonup x \overset{\text{def}}{\Leftrightarrow} \forall_{u\in\mathcal{H}} \quad \langle x_a | u \rangle \to \langle x | u \rangle \Leftrightarrow \lim \langle x_a | u \rangle = \langle x | u \rangle \Rightarrow \lim |\langle x_a | u \rangle| = |\langle x | u \rangle|$

$\overset{u:=x}{\Rightarrow} \lim |\langle x_a | x \rangle| = \|x\|^2$

now $\|x\|^2 = \lim |\langle x_a | x \rangle| = \underline{\lim}\, \underbrace{|\langle x_a | x \rangle|}_{\leq \|x_a\| \|x\| \text{ // using Cauchy-Schwarz}}$ /*recall that unlike $\lim \|x_a\|$ which has to be in finite norm, i.e., in $\mathbb{R}$, $\underline{\lim}\, \|x_a\| \in [-\infty, +\infty]$ and always exists in that sense */

$\leq \underline{\lim}\, \|x_a\| \|x\| = \|x\| \underline{\lim}\, \|x_a\|$

when $x = 0$, $\|x\| \leq \underline{\lim}\, \|x_a\|$ trivially holds

when $x \neq 0$, $\|x\|^2 \leq \|x\| \underline{\lim}\, \|x_a\|$

$\Leftrightarrow \|x\| \leq \underline{\lim}\, \|x_a\|$

$\therefore$ Norm of $\mathcal{H}$: weakly lower semicontinuous. ∎

Lemma 2.36. /* An important result */

[$(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \overset{\in\mathcal{H}}{x}$

$(u_a)_{a\in A}$: net in $\mathcal{H}$, $u_a \to \underbrace{u}_{\in\mathcal{H}}$]

$\Rightarrow$

$\langle x_a | u_a \rangle \to \langle x | u \rangle$

8/28/2016 3:30 PM

Proof:

$(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \overset{\in\mathcal{H}}{x}$

$\Rightarrow \sup_{a\in A} \|x_a\| < +\infty \ldots$ (i)

$u_a \to u \overset{\text{def}}{\Leftrightarrow} \|u_a - u\| \to 0 \ldots$ (ii)

$x_a \rightharpoonup x \overset{\text{def}}{\Leftrightarrow} \forall_{y\in\mathcal{H}} \quad \langle x_a | y \rangle \to \langle x | y \rangle \Rightarrow \langle x_a - x | y \rangle \to 0$

set $y := u \Rightarrow \langle x_a - x | u \rangle \to 0 \ldots$ (iii)

$\forall_{a\in A} \; 0 \leq |\langle x_a | u_a \rangle - \langle x | u \rangle|$

$= |\langle x_a | u_a \rangle - \langle x_a | u \rangle + \langle x_a | u \rangle - \langle x | u \rangle|$ /* this is the famous "one change at a time" trick in real analysis */

$= |\langle x_a | u_a - u \rangle + (-\langle x_a - x | u \rangle)|$

$\leq |\langle x_a | u_a - u \rangle| + |\langle x_a - x | u \rangle|$

$\leq \underbrace{\underbrace{\left(\sup_{a\in A} \|x_a\|\right)}_{<+\infty} \underbrace{\|u_a - u\|}_{\overset{\text{(ii)}}{\to} 0 \text{ /* by (ii) */}}}_{\downarrow 0} + \underbrace{|\langle x_a - x | u \rangle|}_{\to 0 \text{ /* now use}}$

$\downarrow 0$

Result 1.2.3. (Another important result, to establish convergence of a net indirectly)

[$(x_a)_{a\in A}, (y_a)_{a\in A}$] $(x \leq x_a \leq y_a, \; y_a \to x) \Rightarrow x_a \to x$ */

$\Rightarrow \langle x_a | u_a \rangle - \langle x | u \rangle \to 0$

∎

2.5. Weak convergence of sequences.

2015-11-25 11:45 AM

Lemma 2.37.

$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: posses a weakly convergent subsequence.

Proof:

$\mathcal{H}^{weak}$: Hausdorff space

set $\rho = \sup_{n\in\mathbb{N}} \|x_n\| \in (-\infty, \infty)$

$C = B(0;\rho)$: closed ball with 0 center and ρ radius. i.e., $B(0;\rho) = \{x \in \mathcal{H} \mid \|x\| < +\infty\}$

Fact 2.27. (Banach-Alaoglu): The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact

Fact 2.30. (Eberlein-Šmulian): [$C \subseteq \mathcal{H}$] (C: weakly compact $\Leftrightarrow$ C: weakly sequentially compact)

every net has a weak cluster point, i.e., every net has a subnet that weakly converges to a point in $\mathcal{H}$.

every sequence in C has a weak sequential cluster point, i.e., every sequence has a subsequence that weakly converges to a point in C.

$C = B(0;\rho)$: weakly sequentially compact

now $(x_n)_{n\in\mathbb{N}}$ lies in C by construction

/* Definition 1.32.

(C: subset of Hausdorff space X) : weakly sequentially compact

$\overset{def}{\Longleftrightarrow}$

every sequence in C has a subsequence which converges weakly to a point in C

*/

as C is in $\underbrace{\mathcal{H}}_{\text{Hausdorff space}}$ and is weakly sequentially compact $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence. ∎

2016-12-05 3:37 PM

Lemma 2.38.

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n\in\mathbb{N}}$: converges weakly $\Leftrightarrow$ $\begin{cases} \bullet\ (x_n)_{n\in\mathbb{N}} : \text{bounded} \\ \bullet\ (x_n)_{n\in\mathbb{N}} : \text{possesses almost one weak sequential cluster point.} \end{cases}$

Proof: ($\Rightarrow$)

take $\exists_{x\in\mathcal{H}}$ $x_n \rightharpoonup x$ /* now,

$\therefore$ x: unique sequential cluster point of $(x_n)_{n\in\mathbb{N}}$

now $x_n \rightharpoonup x \overset{def}{\Leftrightarrow}$ $\forall_{u\in\mathcal{H}}$ $\langle x_n | u\rangle \to \langle x | u\rangle$ ... (1)

*Lemma 2.23.
$\mathcal{H}^{weak}$: Hausdorff space.

• Fact 1.9. *
[$(a_\alpha)_{\alpha\in A}$: net in Hausdorff space X, $x_\alpha \to x$]
$(x_{\alpha(k)})_{k\in K}$: subnet of $(a_\alpha)_{\alpha\in A}$] $\Rightarrow$ $x_{\alpha(k)} \to x$
/* If a net in a Hausdorff space converges, then so does any of its subnet and to the same point. Mainly space is Hausdorff. */

$\Rightarrow$ $(\langle x_n | u\rangle)_{n\in\mathbb{N}}$: bounded $\forall_{u\in\mathcal{H}}$

$\therefore$ $\forall_{u\in\mathcal{H}} \sup_{n\in\mathbb{N}} |\langle x_n | u\rangle| < +\infty$

/*

*Lemma 2.16. (Banach-Steinhaus uniform boundedness theorem) [$\mathcal{X}$: real Banach space, $\mathcal{Y}$: real normed vector space ; $(T_i)_{i\in I}$: family of operators in $\mathcal{B}(\mathcal{X},\mathcal{Y})$, pointwise bounded $\overset{def}{\Leftrightarrow}$ $\forall_{x\in\mathcal{X}} \sup_{i\in I} \|T_i x\| < +\infty$] $\Rightarrow$ $\sup_{i\in I} \|T_i\| < +\infty$

set. $I = \mathbb{N}$, $T_i = \langle x_i | \cdot \rangle : \mathcal{H} \to \mathbb{R}$; $\mathcal{X} := \mathcal{H}$, $\mathcal{Y} := \mathbb{R}$; $\therefore$ $\forall_{x\in\mathcal{H}}$ $\sup_{i\in I} |\langle x_i | x\rangle| = \|T_i u\| < +\infty$

$\in \mathcal{B}(\mathcal{H},\mathbb{R})$

real normed space

$\sup_{n\in\mathbb{N}} \|T_n\| = \|\langle x_n | \cdot \rangle\| = \|x_n\| < +\infty$

$\leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded ... (2)

From (1), (2) we have proven $\Rightarrow$:)

($\Leftarrow$) given: $(x_n)_{n\in\mathbb{N}}$: bounded. possesses almost one sequential cluster point. say x

*Lemma 2.37. *
[$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$] $\Rightarrow$ $\exists_{(x_{n_k})_{k\in\mathbb{N}}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ $(x_{n_k})_{k\in\mathbb{N}}$: weakly convergent
$\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: possesses a weak sequential cluster point.

$\therefore$ $(x_n)_{n\in\mathbb{N}}$: posses exactly one sequential cluster point (3)

now $(x_n)_{n\in\mathbb{N}}$: bounded $\Rightarrow$ $\sup_{n\in\mathbb{N}} \|x_n\| = \rho < +\infty$, now construct closed ball $B(0;\rho)$ and $(x_n)_{n\in\mathbb{N}}$ lies in $B(0;\rho) = C$

$\therefore$ $(x_n)_{n\in\mathbb{N}} \subseteq B(0;\rho) = C$

Fact 2.27. (Banach-Alaoglu) * *
The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact.

now $B(0;\rho)$: weakly compact

$\Updownarrow$

$C = B(0;\rho)$: weakly sequentially compact (4)

Fact 2.30 (Eberlein-Šmulian)
[C: subset of $\mathcal{H}$]
C: weakly compact $\Leftrightarrow$ C: weakly sequentially compact

now

* *
Lemma 1.34. [C: sequentially compact, $\mathcal{H}$; $(x_n)_{n\in\mathbb{N}} : \subseteq C$, admits a unique sequential cluster point x] $x_n \rightharpoonup x$

// from (4) // from (4) Hausdorff space // weak // weakly // weak

$\mathcal{H}^{weak}$: Hausdorff space

$\therefore$ $x_n \rightharpoonup x$

∎

Lemma 2.39.

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$

C : nonempty, $\subseteq \mathcal{H}$

$\forall_{x\in C}\ (\|x_n - x\|)_{n\in\mathbb{N}}$: converges

$\forall_{y\in C}$ y : sequential cluster point of $(x_n)_{n\in\mathbb{N}}$

]

$\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C.

12/3/2015 10:54 PM

Proof: /* proof strategy: want to show $P\to(Q\wedge R)$: show $(P\to Q)$ and $(P\to R)$ seperately */

given that $\forall_{x\in C}\ \|x_n - x\|_{n\in\mathbb{N}}$: converges $\Rightarrow$ $\boxed{(x_n)_{n\in\mathbb{N}} : \text{bounded}}$ $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: will posses atleast one weak sequential cluster point // using Lemma 2.37

so we have to show that given $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: almost one weak sequential cluster point in C

/* By Lemma 2.38 */

$\Updownarrow$

all the subsequences converges weakly almost one point in C

$\Updownarrow$

$(x_n)_{n\in\mathbb{N}}$ cannot have two distinct weak sequential cluster points in C

x, y : weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ in C

$\Updownarrow$

$x_{k_n} \rightharpoonup \underset{\in C}{x}$, $x_{l_n} \rightharpoonup \underset{\in C}{x}$

$x, y \in C \Rightarrow$ /* by given */

$\left.\begin{array}{l}(\|x_n - x\|)_{n\in\mathbb{N}} : \text{converges} \\ (\|x_n - y\|)_{n\in\mathbb{N}} : \text{converges}\end{array}\right\}$ /* possibly to different numbers */

again,

$$\forall_{n\in\mathbb{N}}\quad \underbrace{\|x_n - y\|^2}_{\|x_n\|^2+\|y\|^2-2\langle x_n|y\rangle} - \underbrace{\|x_n - x\|^2}_{\|x_n\|^2+\|x\|^2-2\langle x_n|x\rangle} + \|x\|^2 - \|y\|^2$$

$$= \cancel{\|x_n\|^2} + \cancel{\|y\|^2} - 2\langle x_n|y\rangle - \cancel{\|x_n\|^2} - \cancel{\|x\|^2} + 2\langle x_n|x\rangle + \cancel{\|x\|^2} - \cancel{\|y\|^2}$$

$$= 2\langle x_n | x-y\rangle$$

$$\therefore\ \underbrace{\lim_{n\to\infty}\|x_n - y\|^2}_{\text{by given, converges to some number say } V_1} - \underbrace{\lim_{n\to\infty}\|x_n - x\|^2}_{\text{by given, converges to some number in } C\text{, say } V_2} + \underbrace{\lim_{n\to\infty}(\|x\|^2 - \|y\|^2)}_{\|x\|^2 - \|y\|^2 \text{ /* fixed term */}} = \lim_{n\to\infty} 2\langle x_n | x-y\rangle$$

$$\leftrightarrow\ \underbrace{\frac{1}{2}(V_1 - V_2 + \|x\|^2 - \|y\|^2)}_{\text{say } \ell} = \lim_{n\to\infty}\langle x_n | x-y\rangle$$

$$\leftrightarrow\ \lim_{n\to\infty}\langle x_n | x-y\rangle = \ell$$

$$\leftrightarrow\ \langle x_n | x-y\rangle \to \ell$$

as the sequence $(x_n)_{n\in\mathbb{N}}$ satisfies $\langle x_n | x-y\rangle \to \ell$, any two subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$ of it will also satisfy : $\langle x_{k_n} | x-y\rangle \to \ell$, $\langle x_{l_n} | x-y\rangle \to \ell$. /* Fact 1.9. $(x_a)_{a\in A}$: net in a Hausdorff space, converges to $x\in X$ $\Rightarrow$ $x_{\kappa(b)} \to x$; we apply this for sequence $(\langle x_n | x-y\rangle)_{n\in\mathbb{N}}$ */

$(x_{\kappa(b)})_{b\in B}$: subnet of $(x_a)_{a\in A}$

* recall : $(x_b)_{b\in B}$: subnet of net $(x_a)_{a\in A}$ def $\Leftrightarrow$ $\exists_{\kappa: B\to A}$ $\begin{cases} \cdot\ \forall_{a\in A}\ \exists_{b\in B}\ \forall_{b\in B:\ \kappa(b)\geq a}\ x_b = x_{\kappa(b)} \\ \cdot\ \forall_{a\in A}\ \exists_{b\in B}\ \forall_{b\in B:\ \kappa(b)\geq a}\ \kappa(b)\geq a \end{cases}$ KB1.7.9

but $\begin{cases} x_{k_n} \rightharpoonup x \leftrightarrow \forall_{\underset{\in H}{u}}\ \langle x_{k_n} | u\rangle \to \langle x | u\rangle \ : \text{set } \underset{u\to u}{u} := x-y \ \therefore\ \langle x_{k_n} | x-y\rangle \to \langle x | x-y\rangle \\ x_{l_n} \rightharpoonup y \leftrightarrow \forall_{u\in H}\ \langle x_{l_n} | u\rangle \to \langle x | u\rangle \ : \text{set } u := x-y \ \therefore\ \langle x_{l_n} | x-y\rangle \to \langle y | x-y\rangle \end{cases}$

using

/* If we have one weakly convergent bounded net, and one strongly convergent net with same index set, then the "inner product" net strongly converges */

** Lemma 2.36.

[$(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \underset{\in H}{x}$

$(u_a)_{a\in A}$: net in $\mathcal{H}$, $u_a \to \underset{\in H}{u}$] $\Rightarrow$ $\langle x_a | u_a\rangle \to \langle x | u\rangle$

$$\left.\begin{array}{l}\langle x | x-y\rangle = \ell \\ = \\ \langle y | x-y\rangle = \ell \\ =\end{array}\right\} \text{subtracting both} \quad \ell - \ell = \langle x | x-y\rangle - \langle y | x-y\rangle = \langle x-y | x-y\rangle = \|x-y\|^2 \leftrightarrow \|x-y\|^2 = 0 \leftrightarrow x = y. \ \blacksquare$$

$\boxed{(x_n)_{n\in\mathbb{N}} : \text{has almost 1 sequential cluster point}}$

so we have shown : $(x_n)_{n\in\mathbb{N}}$: bounded, and has almost 1 sequential cluster point $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: converges weakly

/* By Lemma 2.38 */ $\blacksquare$

Proposition 2.40

[$(e_i)_{i\in I}$: totally ordered family in $\mathcal{H}$, $\overline{\text{span}}\{e_i\}_{i\in I} = \mathcal{H}$ /* $\overline{\text{span}}\, C$: smallest linear subspace of X containing C */

/* By Lemma 2-38 */ ∎

Proposition 2-40.

$[(e_i)_{i\in I}$: totally ordered family in $\mathcal{H}$, $\overline{\mathrm{span}}\{e_i\}_{i\in I}=\mathcal{H}$ /* $\overline{\mathrm{span}}_{C\subseteq X} C$: smallest linear subspace of X containing C */

$(x_n)_{n\in\mathbb{N}}$: a sequence in $\mathcal{H}]\Rightarrow$

(i) $x_n \rightharpoonup x$

(ii) $(x_n)_{n\in\mathbb{N}}$: bounded $\wedge$

$\forall_{i\in I}$ $\langle x_n|e_i\rangle \to \langle x|e_i\rangle$ as $n\to\infty$

• Lemma 2-41.

$[(x_n)_{n\in\mathbb{N}}, (u_n)_{n\in\mathbb{N}}$: sequences in $\mathcal{H}$

$x, u\in\mathcal{H}\ ]\Rightarrow$

(i) $\left(x_n \rightharpoonup x \wedge \overline{\lim}\|x_n\| \leq \|x\|\right) \Leftrightarrow x_n \to x$

(ii) $\mathcal{H}$: finite dimensional $\Rightarrow$ $\left(x_n \rightharpoonup x \Leftrightarrow x_n \to x\right)$

(iii) $x_n \rightharpoonup x$, $u_n \to u$ $\Rightarrow$ $\langle x_n|u_n\rangle \to \langle x|u\rangle$

Corollary 2-42. /* Characterization of strong convergence */

$[(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$

$x\in\mathcal{H}]$

$\Rightarrow$

$(x_n\to x) \Leftrightarrow \left(x_n \rightharpoonup x \wedge \|x_n\|\to\|x\|\right)$

Chapter 2: Part 3

Sunday, November 29, 2015 5:45 PM

2.6. Differentiability.

$\mathcal{K}$: real Banach space.

A Banach space is a vector space X over the field $\mathbf{R}$ of real numbers, or over the field $\mathbf{C}$ of complex numbers, which is equipped with a norm and which is complete with respect to that norm, that is to say, for every Cauchy sequence $\{x_n\}$ in X, there exists an element x in X such that

$$\lim_{n\to\infty} x_n = x,$$

or equivalently:

$$\lim_{n\to\infty} \|x_n - x\|_X = 0.$$

$(x_n)_{n\in\mathbb{N}}$: Cauchy sequence $\overset{\text{def}}{\leftrightarrow}$ $d(x_m, x_n) \to 0$ as $m, n \to \infty$

The vector space structure allows one to relate the behavior of Cauchy sequences to that of converging series of vectors. A normed space X is a Banach space if and only if each absolutely convergent series in X converges,[2]

$$\sum_{n=1}^{\infty} \|v_n\|_X < \infty \quad \text{implies that} \quad \sum_{n=1}^{\infty} v_n \quad \text{converges in} \quad X.$$

Completeness of a normed space is preserved if the given norm is replaced by an equivalent one.

All norms on a finite-dimensional vector space are equivalent. Every finite-dimensional normed space over $\mathbf{R}$ or $\mathbf{C}$ is a Banach space.[3]

https://en.wikipedia.org/wiki/Banach_space#Definition

Definition 2.93. (Gâteaux differentiability)

[C : subset of $\mathcal{H}$

$T : C \to \mathcal{K}$

$x \in C$: $\forall_{y\in\mathcal{H}} \exists_{\alpha\in\mathbb{R}_{++}} \; [x, x+\alpha y] \subseteq C$ /* this condition just keeps the points involved in the definition of derivative in C */

]

T: Gâteaux differentiable at x $\overset{\text{def}}{\leftrightarrow}$ $\exists_{\underbrace{DT(x) \in \mathcal{B}(\mathcal{H},\mathcal{K})}_{\text{i.e. } DT(x) \text{ linear and continuous}}} \; \forall_{y\in\mathcal{H}} \quad DT(x)\, y = \lim_{\alpha\downarrow 0} \dfrac{T(x+\alpha y) - T(x)}{\alpha}$

unique when exists

- Defining Higher order Gâteaux derivative :

e.g. for second order:

$$D^2T(x) \in \mathcal{B}(\mathcal{H}, \mathcal{B}(\mathcal{H},\mathcal{K})) : \forall_{y\in\mathcal{H}} \quad D^2T(x)\, y = \lim_{\alpha\downarrow 0} \frac{DT(x+y) - DT(x)}{\alpha}$$

[Gateaux gradient]

- Notion of Gâteaux gradient:

[C: subset of $\mathcal{H}$,

$f : C \to \mathbb{R}$, Gâteaux differentiable at $x \in C$

]

/* recall Fact 2.17. [Fact 2.17: Riesz-Frechet */

(using Fact 2.17)

$\Rightarrow$

$\exists_{\underbrace{\nabla f(x)}_{T}}$: unique, $\in \mathcal{H}$ $\quad \forall_{y\in\mathcal{H}} \quad Df(x)\, y = \langle y \mid \nabla f(x) \rangle$

$\exists \underbrace{\nabla f(x)}_{\text{called Gâteaux gradient}}$: unique, $\in \mathcal{H}$ $\quad \forall_{y \in \mathcal{H}} \quad Df(x)y = \langle y | \nabla f(x) \rangle$

• Gâteaux Hessian.

f: twice Gâteaux differentiable at $x \in C$

$\Rightarrow \exists$ $\nabla^2 f(x) \in \mathcal{B}(\mathcal{H})$: unique $\quad \forall_{y \in \mathcal{H}} \forall_{z \in \mathcal{H}} \quad (D^2 f(x) y) z = \langle z | \underbrace{\nabla^2 f(x)}_{\text{Gâteaux Hessian of } f \text{ at } x} y \rangle$

[Frechet differentiable]

* Fréchet differentiable:

[$x \in \mathcal{H}$,
$C \in \mathcal{V}(x)$
$T: C \to \mathcal{K}$]

T: Fréchet differentiable at $x \overset{\text{def}}{\Leftrightarrow}$

$$\exists \underbrace{DT(x)}_{\text{linear and continuous}} \in \mathcal{B}(\mathcal{H}, \mathcal{K}) \quad \lim_{0 \neq \|y\| \to 0} \frac{\| T(x+y) - Tx - DT(x) y \|}{\|y\|} = 0 \qquad \{2\text{-}36\}$$

Example 2.46.

[$L \in \mathcal{B}(\mathcal{H})$
$u \in \mathcal{H}$
$x \in \mathcal{H}$
$f: \mathcal{H} \to \mathbb{R}$, $f(y) = \langle Ly | y \rangle - \langle y | u \rangle$]

$\Rightarrow$ f: twice Fréchet differentiable on $\mathcal{H}$, $\nabla f(x) = (L + L^*)x - u$, $\nabla^2 f(x) = L + L^*$

/* L^*: adjoint operator */ [Adjoint operator]

Proof:

$\forall_{y \in \mathcal{H}}$

$$f(x+y) - f(x) = \langle \overset{L \in \mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})}{L}(x+y) | (x+y) \rangle - \langle x+y | u \rangle - \langle Lx | x \rangle + \langle x | u \rangle$$

$$= \langle Lx + Ly | x+y \rangle - \langle x+y | u \rangle - \langle Lx | x \rangle + \langle x | u \rangle$$

$$= \langle Lx | x+y \rangle + \langle Ly | x+y \rangle - \langle x | u \rangle - \langle y | u \rangle - \langle Lx | x \rangle + \langle x | u \rangle$$

$$= \cancel{\langle Lx | x \rangle} + \langle Lx | y \rangle + \langle Ly | x \rangle + \langle Ly | y \rangle - \cancel{\langle x | u \rangle} - \langle y | u \rangle - \cancel{\langle Lx | x \rangle} + \cancel{\langle x | u \rangle}$$

$$= \underbrace{\langle Lx | y \rangle}_{\langle y | Lx \rangle} + \underbrace{\langle Ly | x \rangle}_{\langle y | L^* x \rangle} + \langle Ly | y \rangle - \langle y | u \rangle$$

$$= \langle y | \underbrace{Lx + L^* x}_{(L+L^*)(x)} \rangle + \langle Ly | y \rangle - \langle y | u \rangle$$

/* note that $(A+B)x = Ax + Bx$ is valid
but $A(x+y)$ is not necessarilly $Ax + Ay$! */

$$= \langle y \mid (L+L^*)(x)\rangle + \langle Ly \mid y\rangle - \langle y \mid u\rangle = \langle y \mid (L+L^*)x - u\rangle + \langle Ly \mid y\rangle$$

$$\therefore \quad f(x+y) - f(x) - \langle y \mid (L+L^*)x - u\rangle = \langle Ly \mid y\rangle$$

$$\rightarrow \quad \Big| f(x+y) - f(x) \underbrace{- \langle y \mid (L+L^*)x - u\rangle}_{DT(x)\,y} \Big| = \big|\langle Ly \mid y\rangle\big| \leq \underbrace{\|Ly\|}_{\leq \|L\|\,\|y\|} \|y\| \quad \text{/* Cauchy-Schwartz */}$$

/* By definition, $\|L\| = \sup_{y \neq 0} \frac{\|Ly\|}{\|y\|} \rightarrow \|Ly\| \leq \|L\|\,\|y\|$ */

$$\leq \|L\|\,\|y\|^2$$

$$\rightarrow \quad \frac{\| f(x+y) - f(x) - \langle (L+L^*)x - u \mid y\rangle \|}{\|y\|} \leq \|L\|\,\|y\|$$

$$\rightarrow \quad \lim_{0 \neq \|y\| \to 0} \frac{\| f(x+y) - f(x) - \langle (L+L^*)x - u \mid y\rangle \|}{\|y\|} \leq \lim_{0 \neq \|y\| \to 0} \|L\|\,\|y\| = 0$$

$\therefore$ By definition, {2-36} $DT(x)\,y = \langle (L+L^*)x - u \mid y\rangle$ ■

Lemma 2-49.

[$x \in \mathcal{H}$,
$C \in \mathcal{V}(x)$
$T: C \to \mathcal{K}$, Frechet differentiable at x]

(i) T: Gâteaux differentiable and the two derivatives coincide

(ii) T: continuous at x

Fact 2-50.

[$T: \mathcal{H} \to \mathcal{K}$,
$x \in \mathcal{H}$
$\exists$ Gateaux derivative T in neighborhood of x
DT: continuous at x]

$\Rightarrow$
T: Fréchet differentiable at x

Fact 2-51. /* Chain rule */

[
$x \in \mathcal{H}$,
U: neighborhood of x
$\mathcal{G}$: real Banach space
$T: U \to \mathcal{G}$, Fréchet differentiable
V: neighborhood of Tx
$R: V \to \mathcal{K}$, Gâteaux differentiable at Tx
]

- $R \circ T$: Gâteaux differentiable at x /* So the differentiability type of the composition is same as the outer operator */
- $D(R \circ T)(x) = (DR(Tx)) \circ (DT(x))$

$$= \left(DR(y)\big|_{y=Tx} \circ DT(x) \right)$$

- R: Fréchet differentiable at $x \Rightarrow R \circ T$: Fréchet differentiable at x

$$= \left(DR(y)\big|_{y=Tx} \circ DT(x) \right)$$

- R: Fréchet differentiable at $x \Rightarrow R \circ T$: Fréchet differentiable at x

Fact 2-53.

[$x \in \mathcal{H}$

U: neighborhood of x

K: real Banach space

$T: U \to K$, twice Fréchet differentiable at x]

$\Rightarrow$

$$\forall_{y,z \in \mathcal{H} \times \mathcal{H}} \quad \underbrace{\left(D^2T(x)y\right)z}_{\langle z | \nabla^2 T(x) y \rangle} = \underbrace{\left(D^2T(x)z\right)y}_{\langle y | \nabla^2 T(x) z \rangle}$$

/*

This is a generalization of the symmetric nature of Hessian

$\because \; z^T \nabla^2 T(x) y = z^T \nabla^2 T(x)^T y = (\nabla^2 T(x) z)^T y = y^T \nabla^2 T(x) z$ */

Chapter 3: Part 1

Tuesday, December 1, 2015 12:56 PM

Highlights:

- Fundamental notion of convexity
- Every nonempty closed convex subset C of $\mathcal{H}$: $\underbrace{\text{Chebyshev set}}$

  every point in $\mathcal{H}$ has a unique best approximation from C.

3.1. Definitions and Examples:

Definition 3.1. (Convex set)

$\underbrace{C}_{\subseteq \mathcal{H}}$: convex $\overset{\text{def}}{\leftrightarrow} \forall_{\alpha \in]0,1[} \quad \alpha C + (1-\alpha) C = C$ /* recall the notation: $\cdot\ C+D = \{x+y \mid x \in C, y \in C\}$

$\underbrace{\lambda}_{\in \mathbb{R}} C = \{\lambda x \mid x \in C\}$ */

$\overset{\text{def}}{\leftrightarrow} \forall_{x \in C} \forall_{y \in C} \quad]x,y[\subseteq C$

/* $]x,y[= \{(1-\alpha)x + \alpha y \mid 0 < \alpha < 1\}$ */

Example:

- $C = \mathcal{H}$
- $C = \emptyset$
- C : ball
- C : affine subspace /* C : affine subspace $\overset{\text{def}}{\leftrightarrow} \left(C \neq \emptyset \ \wedge \ \forall_{\lambda \in \mathbb{R}} \ C = \lambda C + (1-\lambda) C \right)$ */
- C : half-space
- $C = \bigcap_{i \in I} C_i$: $(C_i)_{i \in I}$ is a family of convex subsets of $\mathcal{H}$

Definition 3.3. (Convex hull of a set C)

[$C \subseteq \mathcal{H}$]

$\operatorname{conv} C$: convex hull of $C \overset{\text{def}}{\leftrightarrow}$ intersection of all the convex subsets of $\mathcal{H}$ containing C

$\uparrow$

smallest convex subset of $\mathcal{H}$ containing C

$\overline{\operatorname{conv}}\, C$: closed convex hull of $C \overset{\text{def}}{\leftrightarrow}$ smallest closed convex subset of $\mathcal{H}$ containing C

***Proposition 3.4.**

[$C \subseteq \mathcal{H}$,

D : set of all convex combinations of points in $C \overset{\text{def}}{\leftrightarrow} D = \left\{ \sum_{i \in I} \alpha_i x_i \;\middle|\; I : \text{finite}, \{x_i\}_{i \in I} \subseteq C, \{\alpha_i\}_{i \in I} \subseteq]0,1], \sum_{i \in I} \alpha_i = 1 \right\}$

]

$\Rightarrow$

$D = \operatorname{conv} C$

*** Proposition 3.5.**

[$\mathcal{H}$: real Hilbert space

$T : \mathcal{H} \to \mathcal{K}$, affine operator

C: convex subset of $\mathcal{H}$

D: " " of $\mathcal{K}$]

$\Rightarrow$

- $T(C)$: convex subset of $\mathcal{H}$
- $T^{-1}(D)$: convex subset of D.

*** Proposition 3.6.** /* totally ordered : (i) $\forall_{a \in A} a \in A$, (ii) $\forall_{a,b,c \in A} (a \in b, b \in c) \Rightarrow (a \in c)$ (iii) $\forall_{a,b \in A} (a \in b \wedge b \in a) \Rightarrow a = b$, (iv) $\forall_{a \in A} \forall_{b \in A} (a \in b \vee b \in a)$ */

[$(C_i)_{i \in I}$: totally ordered finite family of $\underset{\substack{\| \\ |I|}}{m}$ convex subsets of $\mathcal{H}$]

$\Rightarrow$

(i) $\left(\bigcap_{i \in I} C_i \right)$: convex

(ii) $\forall_{(\alpha_i)_{i \in I} \in \mathbb{R}^m} \quad \sum_{i \in I} \alpha_i C_i$: convex

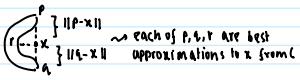
3.2 Best Approximation Property: /* Best approximation and projection are same things */

Definition 3.7.

[C: subset of $\mathcal{H}$

$x \in \mathcal{H}$

$p \in C$]



p: best approximation to x onto C $\overset{\text{def}}{\Leftrightarrow}$ $\|x-p\| = \underbrace{d_C(x)}_{\inf d(x,C)}$

(more commonly, projection of x onto C)

C: proximinimal $\overset{\text{def}}{\Leftrightarrow}$ every point in $\mathcal{H}$ has atleast one projection onto C

C: Chebyshev set $\overset{\text{def}}{\Leftrightarrow}$ every point in $\mathcal{H}$ has exactly one projection onto C

P_C: projection operator (projector) $\overset{\text{def}}{\Leftrightarrow}$ operator that maps every point in $\mathcal{H}$ to its unique projection onto C

if Chebyshev set

* Proposition 3.10.

[$\mathcal{H}$: finite dimensional

C: Chebyshev subset of $\mathcal{H}$]

$\Rightarrow P_C$: continuous

* Proposition 3.12.

(C: nonempty, weakly closed subset of $\mathcal{H}$) $\Rightarrow$ C: proximinimal.

/* A subset C of $\mathcal{H}$ is weakly closed $\overset{\text{def}}{\Leftrightarrow}$ weak limit of every weakly convergent net in C is also in C */

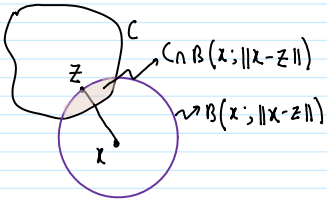
Prop 3.12 start

[Proof:

$x \in \mathcal{H}$,

$z \in C$

$D = C \cap B(x; \|x-z\|)$



$f: \mathcal{H} \to \mathbb{R}: y \mapsto \|x-y\|$

x: cluster point of sequence $(x_n)_{n\in\mathbb{N}}$ $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{V \in \mathcal{U}(x)}$ $\forall_{m \in \mathbb{N}}$ $\exists_{p \in \mathbb{N}: p \geq n}$ $x_p \in V$

C: weakly closed subset of $\mathcal{H}$ /* given */

$B(x; \|x-z\|)$: weakly compact /* weakly compact $\overset{\text{def}}{\Leftrightarrow}$ every sequence has a weak cluster point */

Recall:

Fact 2.27. The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact

$\Rightarrow B(x; \|x-z\|)$: weakly compact

/*

Lemma 1.12.

C: compact subset X (Hausdorff space) $\Rightarrow$ $\begin{cases} \bullet\ C: \text{closed} \\ \bullet\ \forall\ D: \text{closed subset of } C \quad D: \text{compact} \end{cases}$ */

So we can extend this result under weak topology

So, $B(x; \|x-z\|)$: weakly compact $\Rightarrow$ $B(x; \|x-z\|)$: weakly closed

C: weakly closed

$\therefore C \cap B(x; \|x-z\|)$: (weakly closed, /* arbitrary union and finite intersection of closed (open) sets are closed (open) */

a subset of $B(x; \|x-z\|)$ by definition)

$\Rightarrow C \cap B(x; \|x-z\|)$: weakly compact.

$\Rightarrow$ $C \cap B(x; \|x-z\|)$: weakly compact.

By construction, $z \in C \cap B(x; \|x-z\|)$: nonempty, $\because z \in B(x; \|x-z\|)$

now: $f(y) = \|x-y\|$ is weakly lower semicontinuous /* Lemma 2.35. Norm of $\mathcal{H}$: weakly lower semicontinuous [Norm in H: weakly lower semicontinuous] 12/3/2015 11:37 AM */

Now we apply Weirstrass theorem:

[[$\mathcal{H}^{weak}$: Hausdorff space /* Lemma 2.23. $\mathcal{H}^{weak}$: Hausdorff space see [H^weak is Hausdorff] [12/3/2015 11:01 AM] */

$f : \mathcal{H}^{weak} \to \mathbb{R}$, weakly lower semicontinuous

$C \cap B(x; \|x-z\|)$: weakly compact

$\text{dom} f = \mathcal{H}^{weak} \cap \underbrace{(C \cap B(x; \|x-z\|))}_{\text{because this lies in } \mathcal{H}^{weak}}$: nonempty]]

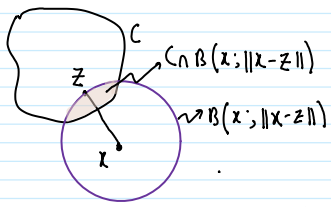
$\Rightarrow$

f achieves infimum over $C \cap B(x; \|x-z\|)$

So, if we define :

$$\left(\begin{array}{ll} \min\limits_{y \in \mathcal{H}^{weak}} & \|\overset{\text{any point in } \mathcal{H}}{x} - y\| \\ \text{s.t.} & y \in C \cap B(x; \|x - \overset{\text{any } z \in C}{z}\|) \end{array} \right)$$

then we have an finite infimum and a minimizer



lets denote the minimizer by $x^* \in \left(\begin{array}{ll} \text{Argmin}\limits_{y \in \mathcal{H}^{weak}} & \|x-y\| \\ \text{s.t.} & y \in C \cap B(x; \|x-z\|) \end{array} \right)$ we claim $x \in \left(\begin{array}{ll} \text{Argmin}\limits_{y \in \mathcal{H}^{weak}} & \|x-y\| \\ \text{s.t.} & y \in C \end{array} \right) = P_C(x)$

$\updownarrow$

$\therefore \forall_{z \in C}$ $\forall_{y \in C \cap B(x; \|x-z\|)}$ $\|x - x^*\| \leq \|x-y\|$ but per absurdum $\exists_{\tilde{x} \in C}$ $\|x - \tilde{x}\| < \|x - x^*\|$

$x^* \in C \cap B(x; \|x-z\|)$

set $z = \tilde{x} \in C$

$\therefore \forall_{y \in C \cap B(x; \|x-\tilde{x}\|)}$ $\|x - x^*\| \leq \|x-y\|$

as: $\tilde{x} \in C$, $\tilde{x} \in B(x; \|x-\tilde{x}\|) \Rightarrow \tilde{x} \in C \cap B(x; \|x-\tilde{x}\|)$

$\Rightarrow \|x-x^*\| \leq \|x-\tilde{x}\|$

$\to \|x-x^*\| \leq \|x-\tilde{x}\| < \|x-x^*\|$ $\forall_{x \in \mathcal{H}}$

by setting $x \in \mathcal{H} \setminus C$ we have $\|x - x^*\| \neq 0$

$\|x-x^*\| < \|x-x^*\|$

$\to 1 < 1$: contradiction

$\therefore x^* \in P_C(x)$

$\therefore$ So, C has atleast one projection onto C for every point in $\mathcal{H}$.

$\blacksquare$] prop 3.12 end

* Corollary 3.13.

[$\mathcal{H}$: finite dimensional

C: nonempty subset of $\mathcal{H}$]

C: proximinimal $\Leftrightarrow$ C: closed

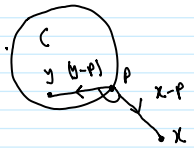
* Every Chebyshev set is proximinimal, but a proximinimal set may not be Chebyshev.

* Theorem 3.14. /* Defining property of projection on a nonempty closed convex set */

C: nonempty closed convex set of $\mathcal{H}$

$\Rightarrow$

- C: Chebyshev set
- $\forall_{x,p\in\mathcal{H}} \left(p = P_C(x) \Leftrightarrow \left(p\in C \wedge \forall_{y\in C} \langle y-p \mid x-p\rangle \leq 0 \right)\right)$



- From 3.14: every nonempty closed convex set is a Chebyshev set
- A Chebyshev set is nonempty and closed

Chebyshev problem: is every Chebyshev set convex?
- yes → if $\mathcal{H}$: finite dimensional
- do not know ☹ → if $\mathcal{H}$: infinite-dimensional : this is still an open problem.

* Proposition 3.17.

[C: nonempty closed convex subset of $\mathcal{H}$

$x, y \in \mathcal{H}$]

$\Rightarrow$

$P_{y+C}\, x = y + P_C(x-y)$

Proof:

By definition: $P_C(x-y) \in C$

$\Rightarrow$ $y + P_C(x-y) \in y+C$

/* Theorem 3.14.

C: nonempty closed convex set of $\mathcal{H}$

$\Rightarrow$

- C: Chebyshev set
- $\forall_{x,p\in\mathcal{H}} \left(p = P_C(x) \Leftrightarrow \left(p\in C \wedge \forall_{y\in C} \langle y-p \mid x-p\rangle \leq 0 \right)\right)$ */

We check this condition now: for $p = y + P_C(x-y)$, consider any $x\in\mathcal{H}$

$\forall_{\tilde{z}\in y+C} \langle \tilde{z} - p \mid x-p\rangle = \langle \tilde{z} - y - P_C(x-y) \mid x - y - P_C(x-y)\rangle \overset{?}{\leq} 0$

$\exists!\, z\in C \;\; \tilde{z} = y + z$

$\Leftrightarrow \forall_{z\in C} \langle \cancel{y} + z - \cancel{y} - P_C(x-y) \mid x - y - P_C(x-y)\rangle \overset{?}{\leq} 0$

$\Leftrightarrow \forall_{z\in C} \langle z - P_C(x-y) \mid (x-y) - P_C(x-y)\rangle \overset{\checkmark}{\leq} 0$

$z \in C$

$\Leftrightarrow \forall_{z \in C} \langle z - P_C(x-y) \mid (x-y) - P_C(x-y) \rangle \leq 0$

/*

now as $x, y \in \mathcal{H} \to x-y \in \mathcal{H}$

as $P_C(\tilde{x})$ is a projection onto C for any $\tilde{x} \in \mathcal{H}$

$\forall_{\tilde{x} \in \mathcal{H}} \ \forall_{\hat{y} \in C} \ \langle \hat{y} - P_C(\tilde{x}) \mid \tilde{z} - P_C(\tilde{x}) \rangle \leq 0$

$\tilde{x} := x-y \in \mathcal{H}, \ \hat{y} = z \in C \Rightarrow \langle z - P_C(x-y) \mid (x-y) - P_C(\tilde{x}) \rangle \leq 0$ */

$\therefore \forall_{\tilde{z} \in y+C} \ \langle \tilde{z} - (y + P_C(x-y)) \mid x - (y + P_C(x-y)) \rangle \leq 0$

So both of the defining properties for projection are satisfied:

$\therefore \ P_{y+C}\, x = y + P_C(x-y)$ ∎

Proposition 3.18.

$(C_n)_{n \in \mathbb{N}}$: (sequence of nonempty, bounded, closed convex subsets of $\mathcal{H}$, $\forall_{n \in \mathbb{N}} \ C_{n+1} \subseteq C_n$)

$\Rightarrow \bigcap_{n \in \mathbb{N}} C_n \neq \emptyset$

Proposition 3.19.

C : (nonempty closed convex subset of $\mathcal{H}$, $x \in \mathcal{H}$)

$\Rightarrow \forall_{\lambda \in \mathbb{R}_+} \ P_C(P_C x + \lambda(x - P_C x)) = P_C(x)$

* Additional properties for projections onto affine subspaces:

Corollary 3.20.

C : closed affine subspace of $\mathcal{H}$

$\Rightarrow$

(i) $[\![x, p \in \mathcal{H}]\!]$

$p = P_C(x) \Leftrightarrow \left(p \in C \wedge \forall_{y \in C} \forall_{z \in C} \ \langle y - z \mid x - p \rangle = 0 \right)$

(ii) P_C : affine operator

* Projections onto linear subspaces.

Corollary 3.22. /* A very important application of this corollary is least squares solution to linear equations */

[V : closed linear subspace of $\mathcal{H}$

$x \in \mathcal{H}$]

$\Rightarrow$

(i) $P_V(x) : \in V, \ x - P_V(x) \perp V$

(ii) $\|P_V(x)\|^2 = \langle P_V(x) \mid x \rangle$

(iii) • $V \neq \{0\} \Rightarrow (P_V \in \mathcal{B}(\mathcal{H}), \|P_V\| = 1)$

• $V = \{0\} \Rightarrow \|P_V\| = 0$

(iv) $V^{\perp\perp} = V$

(v) $P_{V^\perp} = I - P_V$

(vi) $P_V^* = P_V$

(vii)

$\|x\|^2 = \|P_V x\|^2 + \|P_{V^\perp} x\|^2 = d_V^2(x) + d_{V^\perp}^2(x)$

Proposition 3.23.

[C: nonempty subset of $\mathcal{H}$

$V = \overline{\operatorname{span} C}$

$\underbrace{\Pi_C}_{\text{set valued projector onto } C} : \mathcal{H} \to 2^C,\ x \mapsto \{p \in C \mid \|x-p\| = d_C(x)\}$]

- $\Pi_C = \Pi_C \circ P_V$
- C: proximinimal subset of $\mathcal{H} \Leftrightarrow C$: proximinimal subset of V.

*Notion of least-squares to linear equations

Definition. 3.24. (Least squares solution)

[K: real Hilbert space

$T: \in \mathcal{B}(\mathcal{H}, K)$

$y \in K$

$x \in \mathcal{H}$]

x: least-squares solution to $T\overset{\text{variable}}{z} = \overset{\text{fixed}}{y}$

$\overset{\text{def}}{\Leftrightarrow} \|Tx - y\| = \min_{z \in \mathcal{H}} \|Tz - y\|$

Proposition 3.25. (Characterization of least squares solution on closed range linear operator)

[K: real Hilbert space

$\underbrace{T}_{\in \mathcal{B}(\mathcal{H}, K)}$: ($\operatorname{ran} T$: closed)

$y \in K$] $\Rightarrow$

- $Tz = y$ has atleast one least squares solution
- $\forall_{x \in \mathcal{H}} \begin{cases} \text{(i) } x: \text{least squares solution} \Leftrightarrow \\ \text{(ii) } Tx = P_{\operatorname{ran} T} y \Leftrightarrow \\ \text{(iii) } T^*Tx = T^*y \text{ (normal equation)} \end{cases}$

Definition 3.26.

[K: real Hilbert space

$T \in \mathcal{B}(\mathcal{H}, K)$, $\operatorname{ran}(T)$: closed

$\forall_{y \in K}\ C_y = \{x \in \mathcal{H} \mid T^*Tx = T^*y\}$ /*set of least squares solution*/

]

$T^\dagger$: generalized inverse (Moore-Penrose inverse) $\overset{\text{def}}{\Leftrightarrow} \begin{cases} \cdot\ T^\dagger: K \to \mathcal{H} \\ \cdot\ T^\dagger y = P_{C_y}(0) \end{cases}$

*Proposition 3.28.

[K: real Hilbert space

$T: \in \mathcal{B}(\mathcal{H}, K)$, $\operatorname{ran}(T)$: closed] $\Rightarrow$

(i) $\forall_{y \in K}\ \{x \in \mathcal{H} \mid T^*Tx = T^*y\} \cap (\ker T)^\perp = \{T^\dagger y\}$

(ii) $P_{\operatorname{ran} T} = TT^\dagger$

(iii) $P_{\ker T} = I - T^*T^{*\dagger}$

(iv) $T^\dagger \in \mathcal{B}(K, \mathcal{H})$

(v) $\operatorname{ran} T^\dagger = \operatorname{ran} T^*$

(vi) $P_{\operatorname{ran} T^\dagger} = T^\dagger T$

*Proposition 3.29. (Moore-Desoer-Whalen)

[K: real Hilbert space,

$T: \in \mathcal{B}(\mathcal{H}, K)$, $\operatorname{ran}(T)$: closed,

$\tilde{T}: \in \mathcal{B}(K, \mathcal{H})$, $\operatorname{ran} \tilde{T}$: closed]

$\tilde{T} = T^\dagger \Leftrightarrow T\tilde{T} = P_{\operatorname{ran} T} \wedge \tilde{T}T = P_{\operatorname{ran} \tilde{T}} \Leftrightarrow \tilde{T}T|_{(\ker T)^\perp} = I \wedge \tilde{T}|_{(\operatorname{ran} T)^\perp} = 0$

* Corollary 3.30.

[K : real Hilbert space,

$T \in B(H, K)$, $\operatorname{ran} T$: closed] $\Rightarrow$

(i) $T^{\dagger\dagger} = T$

(ii) $T^{\dagger *} = T^{* \dagger}$

Chapter 3: Part 2

Saturday, December 5, 2015 10:06 PM

Hilbert space: metric space $\Rightarrow$ for strong topology closedness and sequential closedness are same

Theorem 3.32.

[C: convex subset of $\mathcal{H}$]

(i) C: weakly sequentially closed /* C: weakly sequentially closed $\overset{def}{\Leftrightarrow}$ weak limit of every weakly (in C) convergent sequence in C is also in C */

$\Leftrightarrow$ (ii) C: sequentially closed /* C: sequentially closed $\overset{def}{\Leftrightarrow}$ (strong) limit of every convergent sequence $(x_n)_{n\in\mathbb{N}}$ is in C */

$\Leftrightarrow$ (iii) C: closed /* C: closed $\overset{def}{\Leftrightarrow}$ (strong) limit of every convergent net that lies in C is also in C */

$\Leftrightarrow$ (iv) C: weakly closed /* C: weakly closed $\overset{def}{\Leftrightarrow}$ weak limit of every weakly convergent net that lies in C is also in C */

Proof: Suppose C: nonempty, else it is trivial.

(i) $\Rightarrow$ (ii):

/*
(One characterization of strong convergence) *
Corollary 2.42. [$(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$; $x\in\mathcal{H}$] $x_n \to x \Leftrightarrow (x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|)$ */

C: weakly sequentially closed

$\Leftrightarrow \forall_{(x_n)_{n\in\mathbb{N}}\subseteq C : x_n \rightharpoonup x}\ x\in C$ // so $\|x\|$: finite $\Rightarrow \|x_n\| \to \|x\|$ // as in norms $\in \mathbb{R}$: finite dimensional weak convergence = strong convergence

$\therefore (x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|)$

$\Leftrightarrow x_n \to x$

$\therefore \forall_{(x_n)_{n\in\mathbb{N}}\subseteq C : x_n \to x}\ x\in C \Leftrightarrow$ C: sequentially closed. ☺

(ii) $\Leftrightarrow$ (iii)

in Hilbert space closedness $\Leftrightarrow$ sequential closedness
continuity $\Leftrightarrow$ sequential continuity
lower semicontinuity $\Leftrightarrow$ sequential lower semicontinuity.
compactness $\Leftrightarrow$ sequential compactness

So, C: sequentially closed $\Leftrightarrow$ C: closed ☺

/*
/* Characterization of projection on closed convex nonempty set */ Theorem 3.14 ***
(C: nonempty closed convex subset of $\mathcal{H}$) $\Rightarrow$
- C: Chebyshev set, i.e. every point in $\mathcal{H}$ has exactly one projection on C
- $\forall_{x\in\mathcal{H}} \left(p = P_C(x) \Leftrightarrow (p\in C, \forall_{z\in C} \langle y-p | x-p\rangle \le 0)\right)$
*/

(iii) $\Rightarrow$ (v)

C: closed $\overset{def}{\Leftrightarrow} \forall_{(\tilde{x}_a)_{a\in A} : \subseteq C,\ \tilde{x}_a \to \tilde{x}}\ \tilde{x}\in C$: given

? C: weakly closed $\overset{def}{\Leftrightarrow} \underbrace{\forall_{(x_a)_{a\in A}\subseteq C : x_a \rightharpoonup x}}_{\text{given}}\ \underbrace{x\in C \leftrightarrow x = P_C x}_{\text{goal}}$

/* using
Fact 1.3.1
(Comes handy in dealing with sequences)
- $\forall_{n\in\mathbb{N}}\ a_n \le a \Leftrightarrow \overline{\lim}\, a_n \le a$
- $\forall_{n\in\mathbb{N}}\ b \le b_n \Rightarrow b \le \underline{\lim}\, b_n$
- $\forall_{n\in\mathbb{N}}\ a_1 \le a_n \le a_2 \Rightarrow a_1 \le \underline{\lim}\, a_n \le \overline{\lim}\, a_n \le a_2$
  $\lim a_n$ exists in finite norm
  $\overline{\lim}\, a_n$ exists in finite norm
*/

$P_C x$ will satisfy

$\forall_{z\in C}\ \langle z - P_C x | x - P_C x\rangle \le 0$

as $(x_a)_{a\in A}\subseteq C \Rightarrow \forall_{a\in A}\ \langle x_a - P_C x | x - P_C x\rangle \le 0 \Rightarrow \overline{\lim}\, \langle x_a - P_C x | x - P_C x\rangle \le 0$

now $x_a \rightharpoonup x$, $P_C x \to P_C x \Rightarrow (x_a - P_C x) \rightharpoonup (x - P_C x)$ and $x - P_C x \to x - P_C x$ $\Big\}$ $\langle x_a - P_C x | x - P_C x\rangle \to \langle x - P_C x | x - P_C x\rangle = \|x - P_C x\|^2$

using

$\therefore \lim \langle x_a - P_C x | x - P_C x\rangle = \overline{\lim}\, \langle x_a - P_C x | x - P_C x\rangle = \|x - P_C x\|^2$

/* if we have one weakly convergent bounded net, and one strongly convergent net with same index set, then the "inner product" net strongly converges */
** |
Lemma 2.36.
[$(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \to \underset{\in\mathcal{H}}{x}$
$(u_a)_{a\in A}$: net in $\mathcal{H}$, $u_a \to \underset{\in\mathcal{H}}{u}$] $\Rightarrow \langle x_a | u_a\rangle \to \langle x | u\rangle$

$0 \le \|x - P_C x\|^2 \le 0 \Leftrightarrow \|x - P_C x\|^2 = 0 \Leftrightarrow x = P_C x \Leftrightarrow x\in C$ [goal achieved]

$\therefore$ C: weakly closed. ☺

(iv) $\Rightarrow$ (i):

any net is a sequence

$\therefore$ C: weakly closed $\Rightarrow$ C: weakly sequentially closed.

■

Theorem 3.33.

[C: bounded, closed, convex subset of $\mathcal{H}$] $\Rightarrow$

C: weakly compact, weakly sequentially compact /* C: weakly compact $\overset{def}{\Leftrightarrow}$ Every net in C has a weak cluster point in C

C: weakly sequentially compact $\overset{def}{\Leftrightarrow}$ Every sequence in C has a weak sequential cluster point in C */

Proof:

C: weakly compact, weakly sequentially compact /* C: weakly compact $\overset{def}{\Leftrightarrow}$ Every net in C has a weak cluster point in C

C: weakly sequentially compact $\overset{def}{\Leftrightarrow}$ Every sequence in C has a weak sequential cluster point in C */

Proof:

C: closed, convex /* now

Theorem 3.32. /* this is a very important theorem which says that for a convex set all the different types of closedness coincides */

[C: convex subset of $\mathcal{H}$]

C: weakly sequentially closed $\Leftrightarrow$ C: sequentially closed $\Leftrightarrow$ C: closed $\Leftrightarrow$ C: weakly closed */

$\Rightarrow$ C: weakly closed

C: bounded : given

/* now, Fact 2.29. [$C \subseteq \mathcal{H}$] C: weakly compact $\Leftrightarrow$ C: weakly closed, bounded */

C: weakly compact /* now using (Eberlein-Šmulian) theorem: Fact 2.30 (Eberlein-Šmulian) **

[C: subset of $\mathcal{H}$]

C: weakly compact $\Leftrightarrow$ C: weakly sequentially compact */

$\Updownarrow$

C: weakly sequentially compact.

$\blacksquare$

Proposition 3.35.

C: convex subset of $\mathcal{H}$ $\Rightarrow$

$\forall_{x \in \operatorname{int} C} \ \forall_{y \in \bar{C}} \ [x, y[\subseteq C$ /* $[x,y[= \{(1-\alpha)x + \alpha y \mid 0 \leq \alpha < 1\}$ */

Proposition 3.36.

[C: convex subset of $\mathcal{H}$] $\Rightarrow$

(i) $\bar{C}$: convex

(ii) $\operatorname{int} C$: convex

(iii) $\operatorname{int} C \neq \emptyset \Rightarrow \begin{cases} \bullet \ \operatorname{int} C = \operatorname{int} \bar{C} \\ \bullet \ \bar{C} = \overline{\operatorname{int} C} \end{cases}$

3.4. Seperation

Definition 3.37. (Seperated sets)

[C, D : subsets of $\mathcal{H}$]

C, D: seperated $\overset{def}{\Leftrightarrow} \exists_{u \in \mathcal{H} \setminus \{0\}} \ \sup \langle C|u \rangle \leq \inf \langle D|u \rangle$ /* $\langle C|u \rangle = \{\langle x|u \rangle \mid x \in C\}$

$\langle D|u \rangle = \{\langle y|u \rangle \mid y \in C\}$ */

C, D: strongly seperated $\overset{def}{\Leftrightarrow} \exists_{u \in \mathcal{H} \setminus \{0\}} \ \sup \langle C|u \rangle < \inf \langle D|u \rangle$

$\underbrace{x}_{\text{a point in } \mathcal{H}}$: seperated from D $\overset{def}{\Leftrightarrow} \exists_{u \in \mathcal{H} \setminus \{0\}} \ \sup \langle x|u \rangle = \langle x|u \rangle \leq \inf \langle D|u \rangle$

x : strongly seperated from D $\overset{def}{\Leftrightarrow} \exists_{u \in \mathcal{H} \setminus \{0\}} \ \langle x|u \rangle < \inf \langle D|u \rangle$

Theorem 3.38

[C: nonempty closed convex subset of $\mathcal{H}$

$x \in \mathcal{H} \setminus C$]

x : strongly seperated from C

/* Two corollaries for strong seperateness of two sets */

Corollary 3.39.

[C, D : nonempty subsets of $\mathcal{H}$, $C \cap D = \emptyset$

C−D: closed, convex] /* $C - D = \{x - y \mid x \in C, y \in D\}$ */

C, D : strongly seperated

Corollary 3.40.

[C, D : nonempty closed convex subsets of $\mathcal{H}$,

$C \cap D = \emptyset$

D : bounded

[C, D: nonempty closed convex subsets of $\mathcal{H}$,
$C \cap D = \emptyset$
D: bounded
]

C, D: strongly seperated

Proof:

sketch: we show $C-D$: closed and convex $\overset{\text{Corollary 3.39}}{\Rightarrow}$ C, D strongly seperated

now C, D: nonempty closed convex

/* Proposition 3.6 (ii): [$(C_i)_{i\in I}$: totally ordered finite family of m convex sets, $\subseteq \mathcal{H}$]

$$\Rightarrow \forall_{(\alpha_i)_{i\in I}: \forall_{i\in I}\, \alpha_i \in \mathbb{R}} \quad \sum_{i\in I} \alpha_i C_i : \text{convex} \quad */$$

so for $\alpha = (1,-1)$, $I = \{1,2\}$, $C_1 = C$, $C_2 = D$ we have $\sum_{i\in I} \alpha_i C_i = C_1 - C_2 = C - D$: convex ✓

now we show $C-D$: closed $\Leftrightarrow$ $C-D$: weakly sequentially closed /* means weak limit of every convergent sequence in $C-D$ is also in $C-D$ */

[as we have shown $C-D$: convex, which we can use as a given now ($\because$ $(P \wedge P \Rightarrow Q) \Leftrightarrow (P \wedge Q)$), and for a convex set all types of closedness coincides */

take a convergent sequence $(x_n - y_n)_{n\in\mathbb{N}}$: $(x_n)_{n\in\mathbb{N}} \subseteq C$, $(y_n)_{n\in\mathbb{N}} \subseteq C$, $x_n - y_n \to \underbrace{z}_{\in \mathcal{H}}$

/* we can always construct such a sequence, the question is whether the limit is also in $C-D$ */

given, D: bounded, closed, convex, $\subseteq \mathcal{H}$ $\Rightarrow$ $\begin{cases} \bullet\ D\text{: weakly compact} \\ \bullet\ D\text{: weakly sequentially compact} \end{cases}$ using

Theorem 3.33.
[C: bounded, closed, convex subset of $\mathcal{H}$]
C: weakly compact, weakly sequentially compact

/* C: weakly compact $\overset{\text{def}}{\Leftrightarrow}$ Every net in has a weak cluster point in C

C: weakly sequentially compact $\overset{\text{def}}{\Leftrightarrow}$ Every sequence in C has a weak sequential cluster point in C

x: (strong) cluster point of $(x_a)_{a\in A}$ $\overset{\text{def}}{\Leftrightarrow}$ $(x_a)_{a\in A}$ has a subnet that (strongly) converges to $x \in X$

x: weak cluster point of $(x_a)_{a\in A}$ $\overset{\text{def}}{\Leftrightarrow}$ $(x_a)_{a\in A}$ has a subnet that weakly converges to $x \in X$ */

$\therefore$ D: weakly sequentially compact

$\Leftrightarrow \exists_{(y_{k_n})_{n\in\mathbb{N}}: \text{subsequence of } (y_n)_{n\in\mathbb{N}}} \quad (y_{k_n} \to y \wedge y \in D)$

as $x_n - y_n \to z$

$\Rightarrow$ any subsequence $x_{k_n} - y_{k_n} \to z$ /* as • Hilbert space is a Hausdorff space

$(z_{k_n})_{n\in\mathbb{N}} = (x_{k_n} - y_{k_n})_{n\in\mathbb{N}}$

• In a Hausdorff space every subnet of a convergent net converges to the same point */

$(x_{k_n})_{n\in\mathbb{N}} \subseteq C$ $(y_{k_n})_{n\in\mathbb{N}} \subseteq D$

$\Downarrow$

$x_{k_n} \to z + y \Rightarrow \underbrace{(x_{k_n})_{n\in\mathbb{N}}}_{\subseteq C}$: weakly convergent to $z+y$ using /* $(a_n \to a,\ b_n \to b) \Rightarrow a_n + b_n \to a + b$ */

But $\underbrace{C}_{\text{convex}}$: closed $\Rightarrow$ $\underbrace{C\text{: weakly sequentially closed}}_{\text{weak limit of every weakly convergent sequence in } C \text{ is also in } C.}$

/* Theorem 3.32. /* this is a very important theorem which says that for a convex set all the different types of closedness coincides */
[C: convex subset of $\mathcal{H}$]
C: weakly sequentially closed $\Leftrightarrow$ C: sequentially closed $\Leftrightarrow$ C: closed $\Leftrightarrow$ C: weakly closed */

$z + y \in C$
$\in D$

$\therefore$ $z \in C - D \Rightarrow$ closed

$\therefore$ (C, D: nonempty, $\subseteq \mathcal{H}$,
$C \cap D = \emptyset$
$C-D$: closed, convex)

$\overset{\text{Corollary 3.39}}{\Rightarrow}$
(C, D): strongly seperated ∎

Chapter 4: Part 1

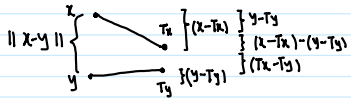
Wednesday, December 9, 2015 9:13 AM

4.1 Nonexpansive operators.

Definition 4.1. (Different types of nonexpansiveness)

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$]



- T: firmly nonexpansive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2 \leq \|x-y\|^2$
- T: nonexpansive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\| \leq \|x-y\|$
- T: quasinonexpansive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in D}\ \forall_{y\in \operatorname{Fix}T}\quad \|Tx-Ty\| \leq \|x-y\|$
- T: strictly quasinonexpansive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in D\setminus \operatorname{Fix}T}\ \forall_{y\in \operatorname{Fix}T}\quad \|Tx-y\| < \|x-y\|$

firmly nonexpansive $\Rightarrow$ nonexpansive $\Rightarrow$ quasiexpansive

$\Downarrow$

strictly quasinonexpansive $\Rightarrow$ quasiexpansive

- **Proposition 4.2.**

[D: nonempty subset of $\mathcal{H}$

$T: D\to\mathcal{H}$]

(i) T: firmly nonexpansive

$\Updownarrow$

(ii) $(I-T)$: firmly nonexpansive

$\Updownarrow$

(iii) $(2T-I)$: nonexpansive

$\Updownarrow$

(iv) $\forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2 \leq \langle x-y \mid Tx-Ty\rangle$

$\Updownarrow$

(v) $\forall_{x\in D}\ \forall_{y\in D}\quad 0 \leq \langle Tx-Ty \mid (I-T)x-(I-T)y\rangle$

$\Updownarrow$

(vi) $\forall_{x\in D}\ \forall_{y\in D}\ \forall_{\alpha\in[0,1]}\quad \|Tx-Ty\| \leq \|\alpha(x-y)+(1-\alpha)(Tx-Ty)\|$

Proof: (i) $\Leftrightarrow$ (ii):

T: firmly nonexpansive

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2 \leq \|x-y\|^2$

$(I-(I-T))x-(I-(I-T))y$

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\quad \|(I-T)x-(I-T)y\|^2+\|(I-(I-T))x-(I-(I-T))y\|^2 \leq \|x-y\|^2$

$\Leftrightarrow (I-T)$: firmly nonexpansive. $\therefore$ (i) $\Leftrightarrow$ (ii)

(i) $\Leftrightarrow$ (iii)

$\forall x,y\in D$

$R = 2T-I$

$\|Rx-Ry\|^2 = \|(2T-I)x-(2T-I)y\|^2 = \|2Tx-x-2Ty+y\|^2$

$R = 2T-1$

$\|Rx-Ry\|^2 = \|(2T-1)x-(2T-1)y\|^2 = \|2Tx-x-2Ty+y\|^2$

$= \|2(Tx-Ty)-(x-y)\|^2 = \|\underbrace{2}_{\alpha}(Tx-Ty)+\underbrace{(1-2)}_{\alpha}(x-y)\|^2$

/* Corollary 2.14.

$\forall_{x,y\in H}\ \forall_{\alpha\in\mathbb{R}}\quad \|\alpha x+(1-\alpha)y\|^2 = \alpha\|x\|^2+(1-\alpha)\|y\|^2-\alpha(1-\alpha)\|x-y\|^2$ */

$= 2\|Tx-Ty\|^2+(1-2)\|x-y\|^2-2(1-2)\underbrace{\|Tx-Ty-x+y\|^2}_{=\|(I-T)y-(I-T)x\|^2=\|(I-T)x-(I-T)y\|^2}$

$= 2\|Tx-Ty\|^2-\|x-y\|^2+2\|(I-T)x-(I-T)y\|^2$

$\leftrightarrow \|Rx-Ry\|^2-\|x-y\|^2 = 2\left(\|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2-\|x-y\|^2\right)$

// note that this is an identity

T: firmly nonexpansive

$\Leftrightarrow \|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2 \leq \|x-y\|^2$

$\Leftrightarrow \|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2-\|x-y\|^2 \leq 0$

$\Leftrightarrow 2\left(\|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2-\|x-y\|^2\right) \leq 0$

$\Leftrightarrow \|\underbrace{R}_{2T-1}x-Ry\|^2-\|x-y\|^2 \leq 0$

$\Leftrightarrow \|(2T-1)x-(2T-1)y\|^2 \leq \|x-y\|^2$

$\Leftrightarrow \|(2T-1)x-(2T-1)y\| \leq \|x-y\|$

$\Leftrightarrow 2T-1$: nonexpansive $\quad \therefore$ (i) $\Leftrightarrow$ (iii)

<u>(i) $\Leftrightarrow$ (iv):</u>

T: firmly nonexpansive

$\Leftrightarrow$

$\forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2+\|(I-T)x-(I-T)y\|^2 \leq \|x-y\|^2$

/* $\|(I-T)x-(I-T)y\|^2 = \|x-Tx-y+Ty\|^2 = \|(x-y)-(Tx-Ty)\|^2$

$= \|x-y\|^2+\|Tx-Ty\|^2-2\langle x-y|Tx-Ty\rangle$ */

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2+\cancel{\|x-y\|^2}+\|Tx-Ty\|^2-2\langle x-y|Tx-Ty\rangle \leq \cancel{\|x-y\|^2}$

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\quad 2\|Tx-Ty\|^2 \leq 2\langle x-y|Tx-Ty\rangle$

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2 \leq \langle x-y|Tx-Ty\rangle$

$\therefore$ (i) $\Leftrightarrow$ (iv)

• (iv) $\Leftrightarrow$ (v)

(iv):

$\forall_{x\in D}\ \forall_{y\in D}\quad \|Tx-Ty\|^2 = \langle Tx-Ty|Tx-Ty\rangle \leq \overbrace{\langle x-y|Tx-Ty\rangle}^{\langle Tx-Ty|x-y\rangle}$ /* $\because \langle a|b\rangle = \langle b|a\rangle$ */

$\Leftrightarrow \forall_{x\in D} \forall_{y\in D} \quad \langle Tx-Ty|x-y\rangle - \langle Tx-Ty|Tx-Ty\rangle = \langle Tx-Ty \,|\, \underbrace{x-y-Tx-Ty}_{(1-T)x-(1-T)y}\rangle \geq 0$ /* $\langle \alpha a+b\beta|c\rangle = \alpha\langle a|c\rangle+\beta\langle b|c\rangle$ */

$\Leftrightarrow \forall_{x\in D} \forall_{y\in D} \quad \langle Tx-Ty \,|\, (1-T)x-(1-T)y\rangle \geq 0$

$\therefore$ (iv) $\Leftrightarrow$ (v)

(v) $\Leftrightarrow$ (vi) :

/* Lemma 2.12.

(i) $\forall_{x,y\in\mathcal{H}} \quad \langle x|y\rangle \leq 0 \Leftrightarrow \forall_{\alpha\in\mathbb{R}_+} \|x\| \leq \|x-\alpha y\| \Leftrightarrow \forall_{\alpha\in[0,1]} \|x\| \leq \|x-\alpha y\|$

(ii) $\forall_{x,y\in\mathcal{H}} \quad x\perp y \Leftrightarrow \forall_{\alpha\in\mathbb{R}} \|x\| \leq \|x-\alpha y\| \Leftrightarrow \forall_{\alpha\in[-1,1]} \|x\| \leq \|x-\alpha y\|$

*/

(v):

$\forall_{x\in D} \forall_{y\in D} \quad \langle Tx-Ty \,|\, (1-T)x-(1-T)y\rangle \geq 0$

$\Leftrightarrow -\langle Tx-Ty \,|\, (1-T)x-(1-T)y\rangle \leq 0$

$\Leftrightarrow \langle -(Tx-Ty) \,|\, (1-T)x-(1-T)y\rangle \leq 0$

$\Leftrightarrow \forall_{\alpha\in[0,1]} \quad \|-(Tx-Ty)\| = \|Tx-Ty\| \leq \|(-Tx+Ty)-\alpha((1-T)x-(1-T)y)\|$

$= \|-Tx+Ty-\alpha(x-Tx-y+Ty)\|$

$= \|-Tx+Ty-\alpha x+\alpha Tx+\alpha y-\alpha Ty\|$

$= \|-\alpha(x-y)-(1-\alpha)Tx+(1-\alpha)Ty\|$

$= \|-(\alpha(x-y)+\underbrace{(1-\alpha)Tx-(1-\alpha)Ty}_{(1-\alpha)(Tx-Ty)})\|$

$= \|\alpha(x-y)+(1-\alpha)(Tx-Ty)\|$

$\Leftrightarrow \forall_{x\in D} \forall_{y\in D} \forall_{\alpha\in[0,1]} \quad \|Tx-Ty\| \leq \|\alpha(x-y)+(1-\alpha)(Tx-Ty)\|$: (vi)

$\therefore$ (v) $\Leftrightarrow$ (vi)

*Corollary 4.3.

[$T\in\mathcal{B}(\mathcal{H})$]

(i) T: firmly nonexpansive $\Leftrightarrow$

(ii) $\|2T-I\| \leq 1 \Leftrightarrow$

(iii) $\forall_{x\in\mathcal{H}} \quad \|Tx\|^2 \leq \langle x|Tx\rangle \Leftrightarrow$

(iv) T^*: firmly nonexpansive $\Leftrightarrow$

(v) $T+T^*-2T^*T$: positive

/* $\forall_{x\in\mathcal{H}} \quad \langle x|(T+T^*-2T^*T)x\rangle \geq 0$ */

* Definition: 4.4. (β-cocoercive / β-inverse strongly monotone)

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$, $\beta \in \mathbb{R}_{++}$]

T: β-cocoercive $\overset{\text{def}}{\Leftrightarrow}$ βT: firmly nonexpansive

(β-inverse strongly monotone)

$$\overset{\text{def}}{\Leftrightarrow} \forall_{x \in D} \forall_{y \in D} \quad \langle x-y \mid Tx-Ty \rangle \geq \beta \|Tx-Ty\|^2$$

/* Recall from Proposition 4.2. (ii), (iv): T: firmly nonexpansive $\Leftrightarrow \forall_{x \in D} \forall_{y \in D} \langle x-y \mid Tx-Ty \rangle \geq \|Tx-Ty\|^2$

$\therefore$ βT: firmly nonexpansive $\Leftrightarrow \forall_{x \in D} \forall_{y \in D} \underbrace{\langle x-y \mid \beta Tx - \beta Ty \rangle}_{\substack{\langle x-y \mid \beta(Tx-Ty) \rangle \\ = \underbrace{\beta}_{\in \mathbb{R}_{++}} \langle x-y \mid Tx-Ty \rangle \quad [\because \langle a \mid \alpha b \rangle = \alpha \langle a \mid b \rangle]}} \geq \underbrace{\|\beta Tx - \beta Ty\|^2}_{\|\beta(Tx-Ty)\|^2 = \beta^2 \|Tx-Ty\|^2}$

$\Leftrightarrow \beta \langle x-y \mid Tx-Ty \rangle \geq \beta^2 \|Tx-Ty\|^2$

$\Leftrightarrow \langle x-y \mid Tx-Ty \rangle \geq \beta \|Tx-Ty\|^2$: this is what is given in the definition. */

* Proposition 4.5.

[K: real Hilbert space

$\beta \in \mathbb{R}_{++}$

$T: K \to K$, β-cocoercive

$L \in \mathcal{B}(\mathcal{H}, K)$: $L \neq 0$, $\|L\|^2 = \frac{\beta}{\gamma}$]

$L^* T L$: γ-cocercive

[Proposition 4.5 Proof] Proof: $\forall_{x, y \in \mathcal{H}}$

We want to prove:

$L^* T L$: γ-cocercive

$$\Leftrightarrow \forall_{x, y \in \mathcal{H}} \quad \langle x-y \mid \underbrace{L^* T L x - L^* T L y} \rangle \geq \gamma \|L^* T L x - L^* T L y\|$$

now:

$$\forall_{x, y \in \mathcal{H}} \quad \underbrace{\langle x-y \mid \underbrace{L^* T L x - L^* T L y}_{L^*(TLx - TLy) \ [\because L^*: \text{linear, continuous}]} \rangle}_{\langle x-y \mid L^*(TLx-TLy) \rangle}$$

$= \langle L(x-y) \mid TLx - TLy \rangle$ [By definition of adjoint operator: $\forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}} \langle Tx \mid y \rangle = \langle x \mid T^* y \rangle$]

$= \langle Lx - Ly \mid TLx - TLy \rangle$ [$\because$ L: linear, continuous]

$= \langle (Lx) - (Ly) \mid T(Lx) - T(Ly) \rangle$ [$\because$ T: linear, continuous]

$\because$ $L \in \mathcal{B}(\mathcal{H}, K)$ $\therefore$ $Lx, Ly \in K$

$T: K \to K$, β-cocoercive $\overset{\text{def}}{\Leftrightarrow} \forall_{x \in K} \forall_{y \in K} \langle x-y \mid Tx-Ty \rangle \geq \beta \|Tx-Ty\|^2$

set: $x := Lx$, $y := Ly$

$\langle Lx - Ly \mid T(Lx) - T(Ly) \rangle \geq \beta \|T(Lx) - T(Ly)\|^2$

$= \gamma \|L\|^2 \|TLx - TLy\|^2$ /* $\because \beta = \gamma \|L\|^2$ */

$= \gamma (\|L\| \|TLx - TLy\|)^2$

$$= \gamma \|L\|^2 \|TLx - TLy\|^2 \quad /* \because \beta = \gamma \|L\|^2 */$$

$$= \gamma \Big(\underbrace{\|L\|}_{\substack{\text{operator,} \\ \in \mathcal{B}(\mathcal{H},\mathcal{K})}} \underbrace{\|TLx - TLy\|}_{\in \mathcal{K}} \Big)^2$$

[$\because x \in \mathcal{H}, L \in \mathcal{B}(\mathcal{H},\mathcal{K})$
$\therefore Lx \in \mathcal{K}$
$T \in \mathcal{B}(\mathcal{K},\mathcal{K}) \therefore TLx \in \mathcal{K} \Rightarrow TLx - TLy \in \mathcal{K}$]

/* with $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$ we cannot apply $\|T\| \|x\| \geq \|Tx\|$ as $TLx - TLy \in \mathcal{K}$, but if we take L^* as $\|L^*\| = \|L\|$ for linear continuous operator then it would work, as $L^* \in \mathcal{B}(\mathcal{K},\mathcal{H})$ */

$$= \gamma \Big(\overbrace{\|L^*\|}^{\in \mathcal{B}(\mathcal{K},\mathcal{H})} \underbrace{\|TLx - TLy\|}_{\in \mathcal{K}} \Big)^2$$

$$\geq \|\underbrace{L^*(TLx - TLy)}_{L^*TLx - L^*TLx)}\| \quad /*$$

$\forall_{T \in \mathcal{B}(\mathcal{H},\mathcal{K})} \|T\| = \sup_{\substack{x \in \mathcal{K} \\ \|x\| \leq 1}} \|Tx\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|}$

*An useful inequality: By definition $\forall_{x \neq 0} \|T\| \geq \frac{\|Tx\|}{\|x\|} \Leftrightarrow \forall_{x \neq 0} \|T\| \|x\| \geq \|Tx\|$

$\Leftrightarrow \boxed{\forall_x \|T\| \|x\| \geq \|Tx\|}$ /* as for $x=0$, equality will hold */ */

$$\geq \gamma \|L^*TLx - L^*TLy\|^2$$

$$\Leftrightarrow \forall_{x,y \in \mathcal{H}} \langle x - y | \underbrace{L^*TLx - L^*TLy} \rangle \geq \gamma \|L^*TLx - L^*TLy\|$$

$\therefore L^*TL$: γ-cocoercive. ∎

Corollary 4.6.

[$\mathcal{K}$: real Hilbert space
$T: \mathcal{K} \to \mathcal{K}$, firmly nonexpansive
$L \in \mathcal{B}(\mathcal{H},\mathcal{K}): \|L\| \leq 1$] $\Rightarrow L^*TL$: firmly nonexpansive

4.2. Projectors and Convex Sets.

Proposition 4.8.

[C: nonempty closed convex set of $\mathcal{H}$] $\Rightarrow P_C$: firmly nonexpansive

Corollary 4.10.

[C: nonempty closed convex set of $\mathcal{H}$] $\Rightarrow$ • $\mathcal{I} - P_C$: firmly nonexpansive
• $2P_C - \mathcal{I}$: nonexpansive

Proof: Comes from Proposition 4.8, Proposition 4.2.

Proposition 4.11.

[C: closed affine subspace of $\mathcal{H}$]

$\Rightarrow$

(i) P_C: weakly continuous

(ii) $\forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}} \quad \|P_Cx - P_Cy\|^2 = \langle x - y | P_Cx - P_Cy \rangle$

4.3. Fixed Points of Nonexpansive Operators:

Proposition 4.13.

4.3. Fixed Points of Nonexpansive Operators:

Proposition 4.13.

[D: nonempty convex subset of $\mathcal{H}$

$T: D \to \mathcal{H}$, quasinonexpansive] $\Rightarrow$

$\operatorname{Fix} T$: convex

Proof:

$x, y \in \operatorname{Fix} T$; $\alpha \in]0,1[$, $z = \alpha x + (1-\alpha) y \in D$ // $\because D$: nonconvex

now: $\|Tz - z\|^2$

$= \|\alpha Tz - \alpha Tz - \alpha x + \alpha x + Tz - z\|^2$

$= \|\alpha Tz - \alpha Tz - \alpha x + \cancel{\alpha x} + Tz - \cancel{\alpha x} - (1-\alpha) y\|^2$

$= \|\alpha Tz - \alpha x + (1-\alpha) Tz - (1-\alpha) y\|^2$

$= \|\alpha (Tz - x) + (1-\alpha)(Tz - y)\|^2$ /* Corollary 2.14. $\forall_{\alpha \in \mathbb{R}}$ $\|\alpha x + (1-\alpha) y\|^2 = \alpha \|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2$ */

$= \alpha \|Tz - \underbrace{x}_{Tx}\|^2 + (1-\alpha)\|Tz - \underbrace{y}_{Ty}\|^2 - \alpha(1-\alpha)\|\cancel{Tz} - x - \cancel{Tz} + y\|^2$ /+ $\because x, y \in \operatorname{Fix} T$ */

$= \alpha \underbrace{\|Tz - Tx\|^2}_{\leq \|z-x\|^2} + (1-\alpha) \underbrace{\|Tz - Ty\|^2}_{\leq \|z-y\|^2} - \alpha(1-\alpha)\|x-y\|^2$ /* T: quasinonexpansive $\overset{\text{def}}{\Leftrightarrow} \forall_{\tilde{x} \in D} \forall_{\tilde{y} \in \operatorname{Fix} T}$ $\|T\tilde{x} - T\tilde{y}\| \leq \|\tilde{x} - \tilde{y}\|^2$ */

$\leq \alpha \|z-x\|^2 + (1-\alpha)\|z-y\|^2 - \alpha(1-\alpha)\|x-y\|^2$ /+ now Corollary 2.14. in opposite direction */

$= \|\alpha(z-x) + (1-\alpha)(z-y)\|^2 = \|\cancel{\alpha z} - \alpha x + \underbrace{z}_{\alpha x + (1-\alpha) y} - y - \cancel{\alpha z} + \alpha y\|^2 = \|-\alpha x + \alpha x + (1-\alpha) y - (1-\alpha) y\|^2 = 0$

$\therefore Tz = z \Leftrightarrow z \in \operatorname{Fix} T$

$\therefore \left(\forall_{x, y \in \operatorname{Fix} T} \forall_{\alpha \in]0,1[} \; z \in \operatorname{Fix} T\right) \Leftrightarrow \operatorname{Fix} T$: convex

$\blacksquare$

Proposition 4.14.

[D: nonempty closed subset of $\mathcal{H}$,

$T: D \to \mathcal{H}$, continuous] $\Rightarrow$

$\operatorname{Fix} T$: closed.

Proof: /+ A: closed $\Rightarrow$ A: sequentially closed $\Leftrightarrow \forall_{(x_n)_{n \in \mathbb{N}} \subseteq A : x_n \to x} \; x \in A$ */

Suppose, $(x_n)_{n \in \mathbb{N}} : \subseteq \operatorname{Fix} T$, $x_n \to x$ // goal: $x \in D$

$\updownarrow$

$Tx_n = x_n \in D$

So, $(x_n)_{n \in \mathbb{N}} : \subseteq D$, $x_n \to x \overset{D:\text{closed}}{\Rightarrow} x \in D$

As, $T: D \to \mathcal{H}$, continuous $\overset{\text{def}}{\Leftrightarrow}$ T: continuous on every point in D

$\Rightarrow T$: continuous at $x \in D \Leftrightarrow \forall_{(x_n)_{n \in \mathbb{N}} : x_n \to x} \; Tx_n \to \underbrace{Tx}_{x} = x \in D$ // as $x \in \operatorname{Fix} T$

$\therefore x \in D$: goal achieved. $\blacksquare$

Corollary 4.15.

[D: nonempty closed convex subset of $\mathcal{H}$,

$T: D \to \mathcal{H}$, nonexpansive] $\Rightarrow$

$\operatorname{Fix} T$: closed, convex

Corollary 4.16.

Fix T: closed, convex

Corollary 4.16.

[D: nonempty closed convex subset of $\mathcal{H}$,

$T: D \to \mathcal{H}$, firmly nonexpansive] $\Rightarrow$

$\operatorname{Fix} T = \bigcap_{x \in D} \{y \in D : \langle y - Tx \mid x - Tx \rangle \leq 0\}$

Proof: /* Proposition 4.2. T: firmly nonexpansive $\leftrightarrow \forall_{\tilde{x} \in D} \forall_{\tilde{y} \in D} \langle T\tilde{x} - T\tilde{y} \mid \tilde{x} - T\tilde{x} - \tilde{y} + T\tilde{y} \rangle \geq 0$

$$\Rightarrow \forall_{x \in D} \forall_{y \in \operatorname{Fix} T} \langle \underbrace{Tx - Ty}_{y} \mid x - Tx - y + \underbrace{Ty}_{y} \rangle = \langle Tx - y \mid x - Tx \rangle = -\langle y - Tx \mid x - Tx \rangle \geq 0$$

$$\leftrightarrow \forall_{x \in D} \forall_{y \in \operatorname{Fix} T} \langle y - Tx \mid x - Tx \rangle \leq 0$$

$$\leftrightarrow \forall_{y \in \operatorname{Fix} T} \forall_{x \in D} \langle y - Tx \mid x - Tx \rangle \leq 0 \quad \ldots \text{(i)}$$

$C = \bigcap_{x \in D} \{y \in D : \langle y - Tx \mid x - Tx \rangle \leq 0\}$

$= \{y \in D : \forall_{x \in D} \langle y - Tx \mid x - Tx \rangle \leq 0\}$

$\therefore y \in C \leftrightarrow \forall_{x \in D} \langle y - Tx \mid x - Tx \rangle \leq 0 \ldots \text{(ii)}$

from (i), (ii): $(\forall_{y \in \operatorname{Fix} T} \; y \in C) \leftrightarrow \boxed{\operatorname{Fix} T \subseteq C}$

Now let's show $C \subseteq \operatorname{Fix} T$

$x := y \in D$ // $\because C \subseteq D$: by definition

$y \in C \leftrightarrow \forall_{x \in D} \langle y - Tx \mid x - Tx \rangle \leq 0 \Rightarrow \langle y - Ty \mid y - Ty \rangle = \|y - Ty\|^2 = 0 \leftrightarrow y = Ty \leftrightarrow y \in \operatorname{Fix} T$

$\therefore \forall_{y \in C} \; y \in \operatorname{Fix} T \leftrightarrow \boxed{C \subseteq \operatorname{Fix} T}$

$\therefore C = \operatorname{Fix} T$ ∎

Theorem 4.17. (demiclosedness principle)

[D: nonempty weakly sequentially closed subset of $\mathcal{H}$,

$T: D \to \mathcal{H}$, nonexpansive

$(x_n)_{n \in \mathbb{N}}$: sequence in D, $x_n \rightharpoonup \overset{\in \mathcal{H}}{x}$, $x_n - Tx_n \to u$] $\Rightarrow x - Tx = u$

Proof:

$(x_n)_{n \in \mathbb{N}} : \subseteq D$, $x_n \rightharpoonup \underbrace{x \in D}$

// $\because$ D: weakly sequentially closed

now: $T: D \to \mathcal{H}$, and $x \in D$ $\therefore$ dom $T = D$, thus Tx: well defined

T: nonexpansive

$\|x - Tx - u\|^2$ /* now: $\|x_n - Tx - u\|^2 = \|\overbrace{x_n - x} + \overbrace{x - Tx - u}\|^2 = \|x_n - x\|^2 + \|x - Tx - u\|^2 + 2\langle x_n - x \mid x - Tx - u \rangle$ */

$\therefore \|x - Tx - u\|^2 = \|x_n - Tx - u\|^2 - \|x_n - x\|^2 - \|x - Tx - u\|^2 - 2\langle x_n - x \mid x - Tx - u \rangle$ */

$= \|x_n - Tx - u\|^2 - \|x_n - x\|^2 - 2\langle x_n - x \mid x - Tx - u \rangle$

// one step at a time trick again

/* $\|x_n - Tx - u\|^2 = \|x_n - Tx_n + Tx_n - Tx - u\|^2$

$= \|(x_n - Tx_n - u) + (Tx_n - Tx)\|^2$

$= \|x_n - Tx_n - u\|^2 + \|Tx_n - Tx\|^2 + 2\langle x_n - Tx_n - u \mid Tx_n - Tx \rangle$ */

$= \|x_n - Tx_n - u\|^2 + \underbrace{\|Tx_n - Tx\|^2} + 2\langle x_n - Tx_n - u \mid Tx_n - Tx \rangle - \|x_n - x\|^2 - 2\langle x_n - x \mid x - Tx - u \rangle$

$$= \|x_n - Tx_n - u\|^2 + \underbrace{\|Tx_n - Tx\|^2}_{\leq \|x_n - x\|^2} + 2\langle x_n - Tx_n - u \mid Tx_n - Tx\rangle - \|x_n - x\|^2 - 2\langle x_n - x \mid x - Tx - u\rangle$$

/* T: nonexpansive */

$$\leq \|x_n - Tx_n - u\|^2 + \cancel{\|x_n - x\|^2} + 2\langle x_n - Tx_n - u \mid Tx_n - Tx\rangle - \cancel{\|x_n - x\|^2} - 2\langle x_n - x \mid x - Tx - u\rangle$$

$$= \|x_n - Tx_n - u\|^2 + 2\langle x_n - Tx_n - u \mid Tx_n - Tx\rangle - 2\langle x_n - x \mid x - Tx - u\rangle \quad \ldots (4.11)$$

Now, given that: $\left.\begin{array}{l} x_n - Tx_n \to u \\ x_n \to x \end{array}\right\} \ominus \Rightarrow (x_n - Tx_n) - x_n \to u - x \leftrightarrow -Tx_n \to u - x$

$\leftrightarrow Tx_n \to x - u$

$\leftrightarrow \boxed{Tx_n - Tx \to x - u - Tx} \quad \ldots (\text{eq } A)$

$\langle x_n - Tx_n - u \mid Tx_n - Tx\rangle \to 0$ /* $a_n \to 0,\ b_n \to 0 \Rightarrow \langle a_n \mid b_n\rangle \to 0$ */

and, $\|x_n - Tx_n - u\|^2 \to 0$

and, $\underbrace{\langle x_n - x}_{\to 0} \mid x - Tx - u\rangle \to 0$

$$\|x_n - Tx_n - u\|^2 + 2\langle x_n - Tx_n - u \mid Tx_n - Tx\rangle - 2\langle x_n - x \mid x - Tx - u\rangle \to 0 \quad (4.12)$$

From (4.11) and (4.12):

Result 1.2.1. (A very important result) /* verified, uses
[$(b_n)_{n\in\mathbb{N}} : \subseteq \mathbb{R}$] $(0 \leq a \leq b_n,\ b_n \to 0) \Rightarrow a = 0$

as, $n \to \infty$ $\quad \|x - Tx - u\|^2 \leq 0$

$\leftrightarrow x - Tx = u$ ∎

Corollary 4.18.

[D: nonempty closed convex subset of $\mathcal{H}$,
$T: D \to \mathcal{H}$, nonexpansive ; $x \in \mathcal{H}$
$(x_n)_{n\in\mathbb{N}}$: sequence in D, $x_n \rightharpoonup x$, $x_n - Tx_n \to 0$] $\Rightarrow$
$x \in \operatorname{Fix} T$

Proof:

D: convex, closed $\Rightarrow$ D: weakly sequentially closed /* for convex set all concepts of closedness are equivalent */

* Theorem 4.17. (demiclosedness principle) * set $u = 0$, rest are same
[[D: nonempty weakly sequentially closed subset of $\mathcal{H}$
$T: D \to \mathcal{H}$, nonexpansive
$(x_n)_{n\in\mathbb{N}}$: sequence in D, $x_n \rightharpoonup \overset{\in \mathcal{H}}{x}$, $x_n - Tx_n \to u$] $\Rightarrow \underbrace{x - Tx = u}$

$x - Tx = 0$

$\leftrightarrow Tx = x$

$\leftrightarrow x \in \operatorname{Fix} T$. ∎

[12/12/2015 4:45 PM]

Theorem 4.19. (Browder-Göhde-Kirk existence theorem)

[D: nonempty bounded closed convex subset of $\mathcal{H}$,
$T: D \to D$, nonexpansive] $\Rightarrow \operatorname{Fix} T \neq \emptyset$

Proof: /* The proof tries to construct a sequence in such a way that Corollary 4.18 can be applied */

D: nonempty bounded closed convex subset of $\mathcal{H}$

$\Rightarrow$ D: weakly sequentially closed /* Theorem 3.32 : for a convex set, all 4 types of closedness (weakly sequentially closed, sequentially closed, closed, weakly closed) collapses */

$\Rightarrow$ D: weakly sequentially compact /* Theorem 3.33. A bounded closed convex subset of $\mathcal{H}$ is weakly compact and weakly sequentially compact */

$\therefore$ D: weakly sequentially closed, and weakly sequentially compact /* $((P\Rightarrow Q)\wedge(P\Rightarrow R)) \Leftrightarrow (P\Rightarrow(Q\wedge R))$

$P\wedge(P\Rightarrow T)\Leftrightarrow(P\wedge T)$ */

define

$x_0\in D$

$(\alpha_n)_{n\in\mathbb{N}}$: sequence in $]0,1]$, $\alpha_0=1$, $\alpha_n\downarrow 0$

$\forall_{n\in\mathbb{N}}$ $T_n : D\to D : x\mapsto \alpha_n x_0+(1-\alpha_n)Tx$

now $T_n(\cdot)=\alpha_n x_0+(1-\alpha_n)T(\cdot)$: contraction as

$$\|T_nx-T_ny\|^2=\|\alpha_nx_0+(1-\alpha_n)Tx-\alpha_nx_0-(1-\alpha_n)Ty\|^2=\|(1-\alpha_n)(Tx-Ty)\|^2=(1-\alpha_n)^2\underbrace{\|Tx-Ty\|^2}_{\leq\|x-y\|^2\ [\because T:\text{nonexpansive}]}\leq(1-\alpha_n)^2\|x-y\|^2$$

$\therefore \|T_nx-T_ny\|\leq(1-\alpha_n)\|x-y\|$

$\because \forall_n\ (1-\alpha_n)<1$ $\therefore T_n$: strict contraction $\Rightarrow \exists\, x_n$: unique $T_nx_n=x_n$ /* Banach-Picard theorem says contraction mappings have unique fixed point */

x_n means it is associated with T_n

now, $\forall_{n\in\mathbb{N}}$ $\|\underset{\text{fixed point of } T_n}{x_n}-Tx_n\|=\|T_nx_n-Tx_n\|$

/* $T_nx=\alpha_nx_0+(1-\alpha_n)Tx$

$\therefore T_nx_n=\alpha_nx_0+(1-\alpha_n)Tx_n$

$\leftrightarrow T_nx_n-Tx_n=\alpha_nx_0-\alpha_nTx_n$ */

$=\|\alpha_nx_0-\alpha_nTx_n\|$

$=\alpha_n\|\underbrace{x_0}_{\in D}-\underbrace{Tx_n}_{\in D}\|$ /* $\because T:D\to D$ */

$\leq\alpha_n\,\mathrm{diam}(D)$ /* $\because$ both x_0, Tx_n are in D, so their total distance must be smaller than $\mathrm{diam}(D)$ = distance between furthest points in D */

$\therefore \forall_{n\in\mathbb{N}}\ 0\leq\|x_n-Tx_n\|\leq\alpha_n\,\mathrm{diam}(D)$

$$0\leq\underbrace{\underline{\lim}_{n\to\infty}\|x_n-Tx_n\|}_{a}\leq\overline{\lim_{n\to\infty}}\,\alpha_n\,\underbrace{\mathrm{diam}(D)}_{\text{finite as D: bounded}}=\overbrace{\lim_{n\to\infty}\alpha_n}^{0}=0$$

using

*Fact 1.5.1.
(comes handy in dealing with sequences)
- $\forall_{n\in\mathbb{N}}\ a_n\leq a \Rightarrow \underline{\lim}\, a_n\leq a$
- $\forall_{n\in\mathbb{N}}\ b\leq b_n \Rightarrow b\leq\overline{\lim}\, b_n$
- $\forall_{n\in\mathbb{N}}\ a_1\leq a_n\leq a_2 \Rightarrow$ $a_1\leq\underline{\lim}\, a_n\leq\overline{\lim}\, a_n\leq a_2$; $\underline{\lim}\, a_n$ exists in $[a_1,a_2]$; $\overline{\lim}\, a_n$ exists in $[a_1,a_2]$

Result 1.2.1. (A very important result) /* verified, uses
[$(b_n)_{n\in\mathbb{N}}:\subseteq\mathbb{R}$] $(0\leq a\leq b_n,\ b_n\to 0)\Rightarrow a=0$

$\therefore \lim_{n\to\infty}\|x_n-Tx_n\|=0$

$\leftrightarrow$ $\boxed{x_n\to Tx_n}$

Now D: weakly sequentially compact $\overset{\text{def}}{\leftrightarrow}$ every sequence in D has a weakly convergent subsequence with its weak limit in D

now $(x_n)_{n\in\mathbb{N}}$: sequence in D $\Rightarrow \exists\,(x_{k_n})_{n\in\mathbb{N}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ $\boxed{x_{k_n}\rightharpoonup x\in D}$

Again $(x_n-Tx_n)_{n\in\mathbb{N}}$ converges to zero

$\Rightarrow (x_{k_n}-Tx_{k_n})$ " " 0 /* if a net converges, so does its any subsequence to the same point */

$\Leftrightarrow$ $\boxed{x_{k_n}-Tx_{k_n}\to 0}$

/* recall Corollary 4.18. Corollary 4.18.
[D : nonempty closed convex subset of $\mathcal{H}$
T: $D\to\mathcal{H}$, nonexpansive
$(x_n)_{n\in\mathbb{N}}$: sequence in D, $x_n\rightharpoonup x$, $x_n-Tx_n\to 0$] $\Rightarrow x\in\mathrm{Fix}\,T$ */

$x\in\mathrm{Fix}\,T$ (proved) ☺

Chapter 4: Part 2

Sunday, December 13, 2015 9:07 AM

12/13/2015 10:40 AM

Proposition 4.21.

[$T_1 : \mathcal{H} \to \mathcal{H}$, firmly nonexpansive

$T_2 : \mathcal{H} \to \mathcal{H}$, firmly nonexpansive

$T = T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - T_2$] $\Rightarrow$

(i) $2T - 2\mathrm{d} = (2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d})$

(ii) T: firmly nonexpansive

(iii) $\mathrm{Fix}\,T = \mathrm{Fix}\,(2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d})$

(iv) T_1: projector onto a closed affine subspace $\Rightarrow \mathrm{Fix}\,T = \{x \in \mathcal{H} \mid T_1 x = T_2 x\}$

Proof: /* $(A+B)x = Ax + Bx$, $(A-B)x = Ax - Bx$ */

(i) $T = T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - T_2$

$\therefore 2T = 2T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - 2T_2$

$\Leftrightarrow 2T - 2\mathrm{d} = 2T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - 2T_2 - 2\mathrm{d} = 2T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - 2T_2$

$\therefore (2T - 2\mathrm{d})x = (2T_1(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - 2T_2)x = 2T_1(2T_2 - 2\mathrm{d})x + x - 2T_2 x$

$(2T_1 - 2\mathrm{d})\underbrace{\overbrace{(2T_2 - 2\mathrm{d})x}^{y}}_{(2T_2 x - x)} = (2T_1 y - y) = 2T_1(2T_2 x - x) - (2T_2 x - x) = 2T_1(2T_2 - 2\mathrm{d})x + x - 2T_2 x = (2T - 2\mathrm{d})x$

$\therefore (2T - 2\mathrm{d}) = (2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d})$

(ii) T_1: firmly nonexpansive $\Rightarrow (2T_1 - 2\mathrm{d})$: nonexpansive

T_2: firmly nonexpansive $\Rightarrow (2T_2 - 2\mathrm{d})$: nonexpansive

/* Proposition 4.2. (different representations of firmly nonexpansive operator)

[D: nonempty subset of $\mathcal{H}$.

$T : D \to \mathcal{H}$]

(i) T: firmly nonexpansive $\Leftrightarrow$ (ii) $(I - T)$: firmly nonexpansive $\Leftrightarrow$ (iii) $(2T - I)$: nonexpansive */

$\underbrace{(2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d})}_{(2T - 2\mathrm{d})\ [\text{from (i)}]}$: nonexpansive [$\because$ composition of nonexpansive operators is nonexpansive]

$\Leftrightarrow (2T - 2\mathrm{d})$: nonexpansive

$\Leftrightarrow$ T: firmly nonexpansive :)

(iii) first note that

$\forall_x \quad x \in \mathrm{Fix}\,T$

$\Leftrightarrow Tx = x$

$\Leftrightarrow 2Tx = 2x$

$\Leftrightarrow 2Tx - x = x$

$\Leftrightarrow (2T - 2\mathrm{d})x = x \Leftrightarrow x \in \mathrm{Fix}(2T - 2\mathrm{d})$ /* this is true for any T */

$\therefore \mathrm{Fix}(T) = \mathrm{Fix}(2T - 2\mathrm{d})$

$= \mathrm{Fix}\,((2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d}))$ /* from (i) */

$\therefore \mathrm{Fix}(T) = \mathrm{Fix}\,((2T_1 - 2\mathrm{d})(2T_2 - 2\mathrm{d}))$ ∎

(iv) recall:

Proposition 4.8.

[C: nonempty closed convex set of $\mathcal{H}$] $\Rightarrow P_C$: firmly nonexpansive

* Corollary 3.20.

[C: closed affine subspace of $\mathcal{H}$] $\forall_{\lambda \in \mathbb{R}} (C = \lambda C + (1-\lambda)C,\ C \neq \emptyset)$

(i) [$x, p \in \mathcal{H}$]

$p \in P_C(x) \Leftrightarrow \begin{cases} p \in C \\ \forall_{y \in C} \forall_{z \in C} \langle y - z \mid x - p \rangle = 0 \end{cases}$

(ii) P_C: affine operator /* $P_C x - P_C 0$: linear operator */

T: affine $\overset{\text{def}}{\Leftrightarrow} \forall_{x,y \in X} \forall_{\lambda \in \mathbb{R}} \quad T(\lambda x + (1-\lambda)y) = \lambda Tx + (1-\lambda)Ty$

let C: the closed $\underbrace{\text{affine subspace}}_{\Rightarrow \text{convex}}$ in given

$T_1 = P_C$

$x \in \mathcal{H}$

P_C: firmly nonexpansive, affine operator.

now suppose, $x \in \mathrm{Fix}\,T$

$\Leftrightarrow x = Tx$

$= (\underbrace{T_1}_{P_C}(2T_2 - 2\mathrm{d}) + 2\mathrm{d} - T_2)x$

$= (P_C(2T_2 - Id) + Id - T_2)x$

$\Leftrightarrow \cancel{x} = P_C(2T_2 - Id)x + \cancel{x} - T_2 x$

//affine

NOW

$\Leftrightarrow T_2 x = P_C(2T_2 - Id)x = P_C(2T_2 x - x) = P_C(2T_2 x + (1-2)x)$ /* T: affine $\overset{def}{\Leftrightarrow} \forall_{x,y\in X} \forall_{\lambda\in\mathbb{R}} \; T(\lambda x + (1-\lambda)y) = \lambda Tx + (1-\lambda)Ty \Leftrightarrow x \mapsto Tx - T_0$: linear */

$= 2\underbrace{P_C(\underbrace{T_2 x}_{\in H})}_{} + (1-2)\underbrace{P_C(\underbrace{x}_{\in H})}_{}$

/* $\because T_2: H \to H$ */

$\in C$ /* C: affine, $P_C(T_2)x \in C$, $P_C(x) \in C$

$\therefore$ affine combination $2(P_C T_2 x) + (1-2)P_C(x) \in C$ */

$\Leftrightarrow T_2 x = 2P_C(T_2 x) + (1-2)P_C(x) \in C$ /* so applying $P_C(\cdot)$ on $T_2 x$ will give $T_2 x$ */

$\Leftrightarrow P_C(T_2 x) = T_2 x = 2P_C(T_2 x) + (1-2)P_C(x)$

$\Leftrightarrow P_C(x) = P_C(T_2 x) = T_2 x$ /* projection on a closed convex set is unique */

$\therefore T_2 x = \underbrace{P_C(x)}_{T_1} \Leftrightarrow T_1 x = T_2 x$

so we have shown:

$\text{Fix}\, T = \{x \in H \mid T_1 x = T_2 x\}$ ■

* 4.22.

[T_1: projector onto a linear subspace of H

$T_2: H \to H$, firmly nonexpansive

$T = T_1 T_2 + (Id - T_1)(Id - T_2)$] $\Rightarrow$

- T: firmly nonexpansive
- $\text{Fix}\, T = \{x \in H : T_1 x = T_2 x\}$

4.4. Averaged Nonexpansive Operators:

- Definition. 4.23. (Averaged nonexpansive operator)

[D: nonempty subset of H,

$T: D \to H$,

$\alpha \in]0,1[$]

T: α-averaged $\overset{def}{\Leftrightarrow} T = (1-\alpha)Id + \alpha R$

- T: firmly nonexpansive $\Leftrightarrow$ T: $\frac{1}{2}$ averaged

Proof: T: $\frac{1}{2}$ averaged $\Leftrightarrow \exists_{R:\text{nonexpansive}} \; T = \frac{1}{2}Id + \frac{1}{2}R$

now R: nonexpansive $\Leftrightarrow Id + R - Id$: nonexpansive $\Leftrightarrow 2\overbrace{\left(\frac{1}{2}Id + \frac{1}{2}R\right)}^{T} - Id$: nonexpansive $\Leftrightarrow 2T - Id$: nonexpansive $\Leftrightarrow T$: firmly nonexpansive ■

[12/14/2015 10:05 AM]

* Proposition 4.25. ⇝ this is a very important theorem, especially in splitting algorithm.

[D: nonempty subset of H

$T: D \to H$, nonexpansive

$\alpha \in]0,1[$]

(i) T: α-averaged $\Leftrightarrow$

(ii) $(1 - \frac{1}{\alpha})Id + \frac{1}{\alpha}T$: nonexpansive $\Leftrightarrow$

(iii)

$\forall_{x\in D} \forall_{y\in D} \quad \|Tx - Ty\|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha}\|(Id - T)x - (Id - T)y\|^2 \Leftrightarrow$

(iv)

$\forall_{x\in D} \forall_{y\in D} \quad \|Tx - Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leq 2(1-\alpha)\langle x-y \mid Tx - Ty\rangle$

Proof: $\forall_{x,y\in D}$

(i) $\Leftrightarrow$ (ii) proof

T: α-averaged $\Leftrightarrow \exists_{R:\text{nonexpansive}} \; \underbrace{T = (1-\alpha)Id + \alpha R}$

$\alpha R = T - (1-\alpha)Id \Leftrightarrow R = \frac{1}{\alpha}T - (\frac{1}{\alpha} - 1)Id = (1 - \frac{1}{\alpha})Id + \frac{1}{\alpha}T$: nonexpansive

$\therefore T$: α-averaged $\Leftrightarrow (1 - \frac{1}{\alpha})Id + \frac{1}{\alpha}T$: nonexpansive

$\therefore$ (i) $\Leftrightarrow$ (ii) ■

Proof (ii) $\Leftrightarrow$ (iii)

$\therefore T : \alpha\text{-averaged} \Leftrightarrow (1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T : \text{nonexpansive}$

$\therefore$ (i) $\Leftrightarrow$ (ii) ∎

Proof (ii) $\Leftrightarrow$ (iii)

let $\lambda = \frac{1}{\alpha}$

then

$R = (1-\lambda)\mathrm{Id} + \lambda T.$

$\Leftrightarrow \lambda T = R - (1-\lambda)\mathrm{Id}$

$\Leftrightarrow T = \frac{1}{\lambda}R - (\frac{1}{\lambda}-1)\mathrm{Id}$

$\therefore T = \underbrace{\frac{1}{\lambda}}_{\alpha}R + \underbrace{(1-\frac{1}{\lambda})}_{\alpha}\mathrm{Id} = (1-\alpha)\mathrm{Id} + \alpha R$

Consider the identity: $\forall_{x,y\in D}$

$\|Rx-Ry\|^2 = \|((1-\lambda)\mathrm{Id}+\lambda T)x - ((1-\lambda)\mathrm{Id}+\lambda T)y\|^2$

$= \|(1-\lambda)x+\lambda Tx-(1-\lambda)y+\lambda Ty\|^2 = \|(1-\lambda)(x-y)+\lambda(Tx-Ty)\|^2$

$= (1-\lambda)\|x-y\|^2 + \lambda\|Tx-Ty\|^2 - \lambda(1-\lambda)\|\underbrace{x-y-Tx+Ty}_{(\mathrm{Id}-T)x-(\mathrm{Id}-T)y}\|^2$ /* $\because \|\alpha x+(1-\alpha)y\|^2 = \alpha\|x\|^2+(1-\alpha)\|y\|^2-\alpha(1-\alpha)\|x-y\|^2$ Corollary 2.14 */

$= (1-\lambda)\|x-y\|^2 + \lambda\|Tx-Ty\|^2 - \lambda(1-\lambda)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2$

$\Leftrightarrow \|Rx-Ry\|^2 = (1-\lambda)\|x-y\|^2 + \lambda\|Tx-Ty\|^2 - \lambda(1-\lambda)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2$

$\Leftrightarrow \underset{(\frac{1}{\alpha})}{\underset{\shortparallel}{\lambda}}\|x-y\|^2 - \underset{(\frac{1}{\alpha})}{\underset{\shortparallel}{\lambda}}\|Tx-Ty\|^2 + \underbrace{\lambda(1-\lambda)}_{\frac{1}{\alpha}(1-\frac{1}{\alpha}) = \frac{1}{\alpha}(\frac{\alpha-1}{\alpha}) = -\frac{1}{\alpha}(\frac{1-\alpha}{\alpha})}\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 = \|x-y\|^2 - \|Rx-Ry\|^2$

$\Leftrightarrow \frac{1}{\alpha}\|x-y\|^2 - \frac{1}{\alpha}\|Tx-Ty\|^2 - \frac{1}{\alpha}\left(\frac{1-\alpha}{\alpha}\right)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 = \|x-y\|^2 - \|Rx-Ry\|^2$

$\Leftrightarrow \|x-y\|^2 - \|Tx-Ty\|^2 - \left(\frac{1-\alpha}{\alpha}\right)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 = \alpha\left(\|x-y\|^2 - \|Rx-Ry\|^2\right)$

now $\underbrace{R}$: nonexpansive $\Leftrightarrow \|Rx-Ry\|^2 \leq \|x-y\|^2$

$(1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T$ $\Leftrightarrow \alpha\left(\|x-y\|^2 - \|Rx-Ry\|^2\right) \geq 0$ [$\because \alpha \in]0,1[$ by given]

$\Leftrightarrow \|x-y\|^2 - \|Tx-Ty\|^2 - \left(\frac{1-\alpha}{\alpha}\right)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \geq 0 \quad \forall_{x,y\in D}$

$\therefore (1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T$: nonexpansive $\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\ \|Tx-Ty\|^2 \leq \|x-y\|^2 - \left(\frac{1-\alpha}{\alpha}\right)\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2$

$\therefore$ (ii) $\Leftrightarrow$ (iii) ∎

(iii) $\Leftrightarrow$ (iv) proof:

(iii): $\forall_{x\in D}\ \forall_{y\in D}\ \|Tx-Ty\|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha}\underbrace{\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2}_{\shortparallel}$

/* $\|x-Tx-y+Ty\|^2 = \|(x-y)-(Tx-Ty)\|^2$

$= \|x-y\|^2 + \|Tx-Ty\|^2 - 2\langle x-y \mid Tx-Ty\rangle$ */

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\ \|Tx-Ty\|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha}\left(\|x-y\|^2 + \|Tx-Ty\|^2 - 2\langle x-y \mid Tx-Ty\rangle\right)$

$= \underbrace{\|x-y\|^2 - \frac{1-\alpha}{\alpha}\|x-y\|^2}_{(1-\frac{1-\alpha}{\alpha})\|x-y\|^2 = (\frac{\alpha-1+\alpha}{\alpha})\|x-y\|^2 = -\frac{1}{\alpha}(1-2\alpha)\|x-y\|^2 \text{: take to L.H.S}} - \frac{1-\alpha}{\alpha}\|Tx-Ty\|^2 + 2\frac{1-\alpha}{\alpha}\langle x-y \mid Tx-Ty\rangle$

$\Leftrightarrow \underbrace{\left(1+\frac{1-\alpha}{\alpha}\right)}_{\shortparallel\ (\frac{1}{\alpha})}\|Tx-Ty\|^2 + \frac{1}{\alpha}(1-2\alpha)\|x-y\|^2 \leq \frac{2(1-\alpha)}{\alpha}\langle x-y \mid Tx-Ty\rangle$

$\Leftrightarrow \frac{1}{\alpha}\|Tx-Ty\|^2 + \frac{1}{\alpha}(1-2\alpha)\|x-y\|^2 \leq \frac{2(1-\alpha)}{\alpha}\langle x-y \mid Tx-Ty\rangle$

$\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\ \|Tx-Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leq 2(1-\alpha)\langle x-y \mid Tx-Ty\rangle$

$\Leftrightarrow$ (iv)

$\therefore$ (iii) $\Leftrightarrow$ (iv) ∎

$\Leftrightarrow$ (iv)

$\therefore$ (iii) $\Leftrightarrow$ (iv) ∎

*Two implications of Proposition 4.25.

• Averaged operators are $\underbrace{\text{strictly quasi-nonexpansive}}$

$\forall_{x \in D \setminus \operatorname{Fix} T} \ \forall_{y \in \operatorname{Fix} T} \quad \|Tx - \underbrace{Ty}_{y}\| < \|x-y\|$

• $\underbrace{T}_{D \to \mathcal{H}}$: $\underbrace{\alpha}_{\alpha \in]0, \frac{1}{2}]}$-averaged $\Rightarrow$ T: firmly nonexpansive

* Proposition 4.28.

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$

$\alpha \in]0,1]$,

$\lambda =]0, \frac{1}{\alpha}[$]

T: α-averaged $\leftrightarrow$ $(1-\lambda)\mathrm{Id} + \lambda T$: $\lambda\alpha$ averaged

[12/15/2015 9:13 AM] * Corollary 4.29.

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$

$\lambda \in]0,2[$]

T: firmly nonexpansive $\leftrightarrow$ $(1-\lambda)\mathrm{Id} + \lambda T$: $\frac{\lambda}{2}$-averaged

Proof:

/* Proposition 4.28 : $\underbrace{T}_{: D \to \mathcal{H},}$: $\underbrace{\alpha}_{\in]0,1[}$ - averaged $\leftrightarrow$ $(1-\underset{\in]0, \frac{1}{\alpha}[}{\lambda})\mathrm{Id} + \lambda T$: $\lambda\alpha$ averaged */

set $\alpha = \frac{1}{2} \in]0,1[$

$\therefore$ T: $\frac{1}{2}$ averaged $\leftrightarrow$ $(1-\underset{\in]0, \frac{1}{\alpha}[=]0,2[}{\lambda})\mathrm{Id} + \lambda T$: $(\lambda\alpha = \frac{\lambda}{2})$-averaged [proved]

* Proposition 4.30.

[D: nonempty subset of $\mathcal{H}$;

$\underbrace{(T_i)_{i \in I}}_{: D \to \mathcal{H}}$: (• finite family of nonexpansive operators ;

• $\forall_{i \in I}$ T_i : $\underbrace{\alpha_i}_{\alpha_i \in]0,1]}$-averaged)

$\underbrace{(\omega_i)_{i \in I}}_{\in]0,1]}$: $\sum_{i \in I} \omega_i = 1$]

$\sum_{i \in I} \omega_i T_i$: $(\max_{i \in I} \alpha_i)$-averaged

Proof: set $T = \sum_{i \in I} \omega_i T_i$

$x, y \in D$

/*

Proposition 4.25. (different faces of an α-averaged operator) ***

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$, nonexpansive

$\alpha \in]0,1[$]

(i) T: α-averaged $\Leftrightarrow$

(ii) $(1-\frac{1}{\alpha})\mathrm{Id} + (\frac{1}{\alpha})T$: nonexpansive $\Leftrightarrow$

(iii)

$\forall_{x \in D} \ \forall_{y \in D} \quad \|Tx - Ty\|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha} \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2 \Leftrightarrow$

(iv)

$\forall_{x \in D} \ \forall_{y \in D} \quad \|Tx - Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leq 2(1-\alpha)\langle x-y \mid Tx - Ty\rangle$ */

as each T_i: α_i averaged

$\leftrightarrow \forall_{x,y \in D} \quad \|T_i x - T_i y\|^2 \leq \|x-y\|^2 - \frac{1-\alpha_i}{\alpha_i} \|(\mathrm{Id}-T_i)x - (\mathrm{Id}-T_i)y\|^2$

$\leftrightarrow \forall_{x,y \in D} \quad \|T_i x - T_i y\|^2 + \frac{1-\alpha_i}{\alpha_i} \|(\mathrm{Id}-T_i)x - (\mathrm{Id}-T_i)y\|^2 \leq \|x-y\|^2 \qquad (1)$

now: $\forall_{x,y \in D} \quad \|Tx - Ty\|^2 + \frac{1-\alpha}{\alpha} \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2$ /* want to show this $\leq \|x-y\|^2$

$\Leftrightarrow \forall_{x,y\in D} \quad \|T_i x - T_i y\|^2 + \frac{1-\alpha_i}{\alpha_i}\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2 \leq \|x-y\|^2 \qquad (1)$

now: $\forall_{x,y\in D} \quad \|Tx-Ty\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2$ /* want to show this $\leq \|x-y\|^2$

$= \|\sum_{i\in I}\omega_i T_i x - \sum_{i\in I}\omega_i T_i y\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-\sum_{i\in I}\omega_i T_i)x-(\mathrm{Id}-\sum_{i\in I}\omega_i T_i)y\|^2$ /* as $\sum_{i\in I}\omega_i = 1 \Rightarrow (\sum_{i\in I}\omega_i \mathrm{Id})x = \sum_{i\in I}\omega_i(\mathrm{Id}\,x) = \sum_{i\in I}\omega_i x = x\sum_{i\in I}\omega_i = x = \mathrm{Id}\,x \quad \therefore \mathrm{Id} = \sum_{i\in I}\omega_i \mathrm{Id}$ */

$= \underbrace{\|\sum_{i\in I}\omega_i T_i x - \sum_{i\in I}\omega_i T_i y\|^2}_{\|\sum_{i\in I}\omega_i(T_i x - T_i y)\|^2} + \frac{1-\alpha}{\alpha}\|\underbrace{(\sum_{i\in I}\omega_i \mathrm{Id} - \sum_{i\in I}\omega_i T_i)x}_{\sum_{i\in I}\omega_i(\mathrm{Id}-T_i)x} - \underbrace{(\sum_{i\in I}\omega_i \mathrm{Id} - \sum_{i\in I}\omega_i T_i)y}_{\sum_{i\in I}\omega_i(\mathrm{Id}-T_i)y}\|^2$

$\underbrace{\sum_{i\in I}\omega_i((\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y)}$

/* now,

*Proposition 2.5. [$f: \mathcal{H}\to[-\infty,+\infty]$]

f: convex $\overset{\text{def}}{\Leftrightarrow} \forall_{(x_i)_{i\in I}: \in \mathcal{H}, (\alpha_i)_{i\in I}: \in]0,1[, \sum_{i\in I}\alpha_i = 1 \text{ finite index set}} \forall_{(x_i)_{i\in I}: \in \mathrm{dom}\, f} \quad f(\sum_{i\in I}\alpha_i x_i) \leq \sum_{i\in I}\alpha_i f(x_i)$

as $\|\cdot\|^2$: convex $\Rightarrow$ $\text{norm_squared}(\sum_{i\in I}\omega_i(T_i x - T_i y)) \leq \sum_{i\in I}\omega_i\, \text{norm_squared}(T_i x - T_i y)$

$= \sum_{i\in I}\omega_i \|T_i x - T_i y\|^2$

similarly, $\|\sum_{i\in I}\omega_i((\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y)\|^2 \leq \sum_{i\in I}\omega_i\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2$ */

$\leq \sum_{i\in I}\omega_i\|T_i x - T_i y\|^2 + \frac{1-\alpha}{\alpha}\sum_{i\in I}\omega_i\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2$ /* now $\min_{i\in I}\left(\frac{1-\alpha_i}{\alpha_i} = \frac{1}{\alpha_i}-1\right) = \frac{1}{\max_{i\in I}\alpha_i}-1 = \frac{1}{\alpha}-1 = \frac{1-\alpha}{\alpha}$

$\therefore \forall_{i\in I} \quad \frac{1-\alpha}{\alpha} \leq \frac{1-\alpha_i}{\alpha_i}$ */

$\sum_{i\in I}\frac{1-\alpha}{\alpha}\omega_i\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2$

$\leq \sum_{i\in I}\frac{1-\alpha_i}{\alpha_i}\omega_i\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2$

$\leq \sum_{i\in I}\omega_i\|T_i x - T_i y\|^2 + \sum_{i\in I}\frac{1-\alpha_i}{\alpha_i}\omega_i\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2$

$= \sum_{i\in I}\omega_i\underbrace{\left(\|T_i x - T_i y\|^2 + \frac{1-\alpha_i}{\alpha_i}\|(\mathrm{Id}-T_i)x-(\mathrm{Id}-T_i)y\|^2\right)}_{\leq \|x-y\|^2}$ /* from (1) */

$\therefore \forall_{x,y\in D} \quad \|Tx-Ty\|^2 + \frac{(1-\alpha)}{\alpha}\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \leq \|x-y\|^2 \leftrightarrow T$: α-averaged

/* using

Proposition 4.25. (different faces of an α-averaged operator) ***

[D: nonempty subset of $\mathcal{H}$

$T: D\to\mathcal{H}$, nonexpansive

$\alpha\in]0,1[$]

(i) T: α-averaged $\Leftrightarrow$

(ii) $(1-\frac{1}{\alpha})\mathrm{Id} + (\frac{1}{\alpha})T$: nonexpansive $\Leftrightarrow$

(iii) $\forall_{x\in D}\forall_{y\in D} \quad \|Tx-Ty\|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \Leftrightarrow$

(iv) $\forall_{x\in D}\forall_{y\in D} \quad \|Tx-Ty\|^2 + (1-2\alpha)\|x-y\|^2 \leq 2(1-\alpha)\langle x-y \mid Tx-Ty\rangle$ */

* Proposition 4.32. (Composition of averaged operators)

[D: nonempty subset of $\mathcal{H}$

$I = \{1,\dots,m\}$ → strictly positive integer

$(\underbrace{T_i}_{:D\to D})_{i\in I}: (\forall_{i\in I} \quad T_i: \underbrace{\alpha_i}_{\in]0,1[}\text{-averaged})$]

$\left(T = T_1\cdots T_m, \quad \alpha = \frac{m}{m+\frac{1}{\max_{i\in I}\alpha_i}}\right) \Rightarrow T$: α-averaged

Proof: $\forall_{i\in I} \quad \alpha_i := \alpha_i/(1-\alpha_i)$

$\alpha = \max_{i\in I}\alpha_i$

$\forall x,y\in D$

now

$\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2/m$

$= \|(\mathrm{Id}-T_1T_2\cdots T_m)x-(\mathrm{Id}-T_1T_2\cdots T_m)y\|^2/m$ /* use one step at a time trick */

$= \|x-y-(T_1T_2\cdots T_m x - T_1\cdots T_m)y\|^2$

$= \|\underbrace{(x-y)-(T_m x - T_m y)}_{a_1} + \underbrace{(T_m x - T_m y)-(T_{m-1}T_m x - T_{m-1}T_m y)}_{a_2} + (T_{m-1}T_m x - T_{m-1}T_m y) - \dots - (T_2\cdots T_m x - T_2\cdots T_m y) + (T_2\cdots T_m x - T_2\cdots T_m y)$

$- (T_1\cdots T_m x - T_1\cdots T_m y)\|^2/m$

$$= \| \underbrace{(x-y)-(T_m x - T_m y)} + \underbrace{(T_m x - T_m y) - (T_{m-1}T_m x - T_{m-1}T_m y)} + (T_{m-1}T_m x - T_{m-1}T_m y) - \ldots - (T_2 \cdots T_m x - T_2 \cdots T_m y) + (T_2 \cdots T_m x - T_2 \cdots T_m y)$$

$$- (T_1 \cdots T_m x - T_1 \cdots T_m y) \|^2 / m$$

$$= \| \overbrace{(Id - T_m)x - (Id - T_m)y}^{a_1} + \overbrace{(Id - T_{m-1})T_m x - (Id - T_{m-1})T_m y}^{a_2} + \ldots + \overbrace{(Id - T_1)T_2 \cdots T_m x - (Id - T_1)T_2 \cdots T_m y}^{a_m} \|^2 / m$$

/* $\| a_1 + a_2 + \ldots + a_m \|^2 \frac{1}{m} = \underbrace{\| \frac{1}{m} a_1 + \frac{1}{m} a_2 + \ldots + \frac{1}{m} a_m \|^2 \cdot m}_{/* \leq \sum_{i=1}^{m} \frac{1}{m} \| a_i \|^2 \ */ \ \leq \sum_{i=1}^{m} \| a_i \|^2 \ */}$

Result 2.13.1: [$(x_i)_{i \in I}$: finite families in $\mathcal{H}$; $(\alpha_i)_{i \in I}$: $\subseteq \mathbb{R}$, $\sum_{i \in I} \alpha_i = 1$] $\| \sum_{i \in I} \alpha_i x_i \|^2 \leq \sum_{i \in I} \alpha_i \| x_i \|^2$

$$\leq \| (Id - T_m)x - (Id - T_m)y \|^2 + \| (Id - T_{m-1})T_m x - (Id - T_{m-1})T_m y \|^2 + \ldots + \| (Id - T_1)T_2 \cdots T_m x - (Id - T_1)T_2 \cdots T_m y \|^2 \ /*$$

$$\leq \underbrace{\frac{\kappa_m}{1-\kappa_m}}_{K_m} \left(\|x-y\|^2 - \| T_m x - T_m y \|^2 \right) + \underbrace{\frac{\alpha_{m-1}}{1-\alpha_{m-1}}}_{K_{m-1}} \left(\| T_m x - T_m y \|^2 - \| T_{m-1}T_m x - T_{m-1}T_m y \|^2 \right) + \ldots + \overbrace{\frac{\alpha_1}{1-\alpha_1}}^{K_1} \left(\| T_2 \cdots T_m x - T_2 \cdots T_m x \|^2 - \| T_1 \cdots T_m x - T_1 \cdots T_m y \|^2 \right)$$

$$= \overbrace{K_m \left(\|x-y\|^2 - \| T_m x - T_m y \|^2 \right) + K_{m-1} \left(\| T_m x - T_m y \|^2 - \| T_{m-1}T_m x - T_{m-1}T_m y \|^2 \right) + \ldots + K_1 \left(\| T_2 \cdots T_m x - T_2 \cdots T_m x \|^2 - \| T_1 \cdots T_m x - T_1 \cdots T_m y \|^2 \right)}^{\text{each of these are positive}}$$

/* now, $K = \max_{i \in I} K_i$ */

$$\leq K \left(\|x-y\|^2 - \| T_m x - T_m y \|^2 \right) + K \left(\| T_m x - T_m y \|^2 - \| T_{m-1}T_m x - T_{m-1}T_m y \|^2 \right) + \ldots + K \left(\| T_2 \cdots T_m x - T_2 \cdots T_m x \|^2 - \| \underbrace{T_1 \cdots T_m x}_{T} - \underbrace{T_1 \cdots T_m y}_{T} \|^2 \right)$$

$$= K \left(\|x-y\|^2 - \| Tx - Ty \|^2 \right)$$

$$\therefore \frac{1}{Km} \left(\| (Id - T)x - (Id - T)y \|^2 \right) \leq \cancel{K} \left(\|x-y\|^2 - \| Tx - Ty \|^2 \right)$$

$\leftrightarrow \| Tx - Ty \|^2 \leq \|x-y\|^2 - \frac{1}{Km} \| (Id - T)x - (Id - T)y \|^2$ /* now set $\frac{1-\beta}{\beta} = \frac{1}{Km} \leftrightarrow Km(1-\beta) = \beta \leftrightarrow Km = \beta(1+Km)$

$\leftrightarrow \beta = \frac{Km}{1+Km}$ */

$\therefore \| Tx - Ty \|^2 \leq \|x-y\|^2 - \frac{1-\beta}{\beta} \| (Id - T)x - (Id - T)y \|^2$ where $\beta = \frac{Km}{1+Km} = \alpha$ //given

$\leftrightarrow T$: α-averaged. ■

now

Proposition 4.25. (different faces of an α-averaged operator) ★★★

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$, nonexpansive

$\alpha \in]0,1[$]

(i) T: α-averaged $\Leftrightarrow$

(ii) $(1 - \frac{1}{\alpha}) Id + (\frac{1}{\alpha}) T$: nonexpansive $\Leftrightarrow$

(iii) $\forall_{x \in D} \forall_{y \in D} \ \alpha \| (Id - T)x - (Id - T)y \|^2 \leq \frac{\alpha}{1-\alpha} \left(\|x-y\|^2 - \| Tx - Ty \|^2 \right)$

$\forall_{x \in D} \forall_{y \in D} \ \| Tx - Ty \|^2 \leq \|x-y\|^2 - \frac{1-\alpha}{\alpha} \| (Id - T)x - (Id - T)y \|^2 \Leftrightarrow$

(iv) $\forall_{x \in D} \forall_{y \in D} \ \| Tx - Ty \|^2 + (1 - 2\alpha) \|x-y\|^2 \leq 2(1-\alpha) \langle x-y \mid Tx - Ty \rangle$] */

[12/15/2015 9:30 AM] ★proposition 4.33.

[D: nonempty subset of $\mathcal{H}$

$\beta \in \mathbb{R}_{++}$

$T: D \to \mathcal{H}$, β-cocoercive

$\gamma \in]0, 2\beta[$]

$(Id - \gamma T)$: $\frac{\gamma}{2\beta}$ averaged

Proof:

$\beta \in \mathbb{R}_{++}$

T: β-cocoercive $\leftrightarrow \beta T$: firmly nonexpansive

$\leftrightarrow \beta T$: $\frac{1}{2}$averaged

$\leftrightarrow \exists_{\underbrace{R: D \to \mathcal{H}}_{\because T: D \to \mathcal{H}}} \left(R: \text{nonexpansive} \wedge \underbrace{\beta T = \frac{1}{2} Id + \frac{1}{2} R}_{T = \frac{1}{2\beta}(Id + R)} \right)$

$Id - \gamma T \quad \{ \gamma \in]0, 2\beta[\}$

$= Id - \gamma \cdot \frac{1}{2\beta} (Id + R)$

$= (1 - \frac{\gamma}{2\beta}) Id + \underbrace{(\frac{\gamma}{2\beta})}_{\text{As } \gamma \in]0, 2\beta[\ \frac{\gamma}{2\beta} \in]0,1[\ */} \underbrace{(-R)}_{\text{nonexpansive}}$

/* R: nonexpansive $\Leftrightarrow \underbrace{\| Rx - Ry \|^2}_{= \| -(Rx - Ry) \|^2 = \| (-R)x - (-R)y \|^2} \leq \|x-y\|^2$

$\leftrightarrow \| (-R)x - (-R)y \|^2 \leq \|x-y\|^2$

$\therefore (-R)$: nonexpansive */

$\therefore Id - \gamma T$: $\frac{\gamma}{2\beta}$ averaged ■

Chapter 4: Part 3

Thursday, December 17, 2015 9:24 PM

* **Proposition 4.34.**

[D: nonempty subset of $\mathcal{H}$

$(T_i)_{i\in I}$: finite family of quasinonexpansive operators $: D \to \mathcal{H}$, $\bigcap_{i\in I} \operatorname{Fix} T_i \neq \emptyset$

$(\omega_i)_{i\in I}$: strictly positive real numbers, $\sum_{i\in I} \omega_i = 1$]

$\Rightarrow \operatorname{Fix} \sum_{i\in I} \omega_i T_i = \bigcap_{i\in I} \operatorname{Fix} T_i$

proof: set $T = \sum_{i\in I} \omega_i T_i$

first, $\bigcap_{i\in I} \operatorname{Fix} T_i \subseteq \operatorname{Fix} T$, as

$$\forall_x \left(x \in \bigcap_{i\in I} \operatorname{Fix} T_i \leftrightarrow \forall_{i\in I} \underbrace{x \in \operatorname{Fix} T_i}_{T_i x = x} \Rightarrow \forall_{i\in I}\ \omega_i T_i x = \omega_i x \Rightarrow \overbrace{\sum_{i\in I} \omega_i T_i}^{T} x = \sum_{i\in I} \omega_i x = x \overbrace{\sum_{i\in I} \omega_i}^{1} = x \leftrightarrow Tx = x \leftrightarrow x \in \operatorname{Fix} T \right)$$

$\leftrightarrow \bigcap_{i\in I} \operatorname{Fix} T_i \subseteq \operatorname{Fix} T$

Now show $\operatorname{Fix} T \subseteq \bigcap_{i\in I} \operatorname{Fix} T_i \leftrightarrow \forall_{\underbrace{x \in \operatorname{Fix} T}_{\text{given}}} \overbrace{\left(x \in \bigcap_{i\in I} \operatorname{Fix} T_i \leftrightarrow \forall_{i\in I}\ x \in \operatorname{Fix} T_i \leftrightarrow \forall_{i\in I}\ T_i x = x \right)}^{\text{goal}}$

now given, $\exists y \in \underbrace{\bigcap_{i\in I} \operatorname{Fix} T_i}_{A \neq \emptyset} \leftrightarrow \forall_{i\in I}\ y \in \operatorname{Fix} T_i$

now, T_i: quasinonexpansive $\forall_{i\in I}$

$$\leftrightarrow \forall_{i\in I}\ \forall_{\tilde{x}\in D}\ \forall_{\tilde{y}\in \operatorname{Fix} T_i}\quad \|T_i\tilde{x} - T_i\tilde{y}\| = \|T_i\tilde{x} - \tilde{y}\| \leq \|\tilde{x} - \tilde{y}\|$$

set $\tilde{x} := x \in \operatorname{Fix} T \subseteq D$, $\tilde{y} := y$

$$\Rightarrow \forall_{i\in I}\quad \|T_i x - y\|^2 \leq \|x - y\|^2$$

// a variant of "one step at a time" rule again

/* now $\|T_i x - y\|^2 = \|\underbrace{T_i x - x}_{} + \underbrace{x - y}_{}\|^2 = \|T_i x - x\|^2 + \|x - y\|^2 + 2\langle T_i x - x | x - y\rangle$ */

$$\forall_{i\in I}\quad \|T_i x - x\|^2 + \cancel{\|x - y\|^2} + 2\langle T_i x - x | x - y\rangle \leq \cancel{\|x - y\|^2}$$

$$\leftrightarrow \forall_{i\in I}\quad 2\langle T_i x - x | x - y\rangle \leq -\|T_i x - x\|^2$$

mult. both sides by ω_i

$$\leftrightarrow \forall_i\quad 2\omega_i\langle T_i x - x | x - y\rangle \leq -\omega_i\|T_i x - x\|^2$$

$$\Rightarrow \sum_{i\in I} 2\omega_i\langle T_i x - x | x - y\rangle = 2\left(\langle \omega_1 T_1 x - \omega_1 x | x - y\rangle + \dots + \langle \omega_m T_m x - \omega_m x | x - y\rangle \right) \leq -\sum_{i\in I} \omega_i \|T_i x - x\|^2$$

$$\leftrightarrow 2\langle (\omega_1 T_1 x + \dots + \omega_m T_m x) - \underbrace{(\omega_1 + \dots + \omega_m)}_{1} x | x - y\rangle \leq -\sum_{i\in I} \omega_i \|T_i x - x\|^2$$

$$\leftrightarrow 2\langle \underbrace{\underbrace{(\omega_1 T_1 + \dots + \omega_m T_m)}_{\sum_{i\in I} \omega_i T_i = T} x - x}_{Tx - x = 0 \ \text{/* } \because x \in \operatorname{Fix} T \text{ */}} | x - y\rangle \leq -\sum_{i\in I} \omega_i \|T_i x - x\|^2$$

$$\leftrightarrow 0 \leq -\sum_{i\in I} \omega_i \|T_i x - x\|^2 \leftrightarrow \sum_{i\in I} \omega_i \|T_i x - x\|^2 \leq 0 \leftrightarrow \forall_{i\in I}\ \|T_i x - x\|^2 = 0 \leftrightarrow \forall_{i\in I}\ x = T_i x$$

$\therefore$ goal achieved. ☺

■

* **Proposition 4.35.**

[D: nonempty subset of $\mathcal{H}$

T_1: quasinonexpansive operator $: D \to D$ } one of them is strictly quasinonexpansive,

T_2: " " $: D \to D$ } $\operatorname{Fix} T_1 \cap \operatorname{Fix} T_2 \neq \emptyset$

] $\Rightarrow$

- $\operatorname{Fix} T_1 T_2 = \operatorname{Fix} T_1 \cap \operatorname{Fix} T_2$
- $T_1 T_2$: quasinonexpansive
- (T_1, T_2): strictly quasinonexpansive $\Rightarrow$ $T_1 T_2$: strictly quasinonexpansive

* Corollary 4.36.

$\underbrace{\text{strictly quasinonexpansive}}_{}$, $: D \to D$ (nonempty), $\bigcap_{i \in I} \operatorname{Fix} T_i \neq \emptyset$ /* $I = \{1, \ldots, m\}$ */

- $T = T_1 T_2 \cdots T_m$: strictly quasinonexpansive $\wedge$ $\operatorname{Fix} T = \bigcap_{i \in I} \operatorname{Fix} T_i$

Proof: strong induction on m. /* Strong induction:

> **To prove a goal of the form $\forall n \in \mathbb{N} P(n)$:**
>
> Prove that $\forall n[(\forall k < n P(k)) \to P(n)]$, where both n and k range over the natural numbers in this statement. Of course, the most direct way to prove this is to let n be an arbitrary natural number, assume that $\forall k < n P(k)$, and then prove $P(n)$.

*/

for $m = 1 \Rightarrow T = T_1$: strictly quasinonexpansive (by given)

for $m = 1, 2 \Rightarrow T = T_1 T_2$: strictly quasinonexpansive /* from

> *Proposition 4.36. (Common fixed points of composition of two quasinonexpansive operators.*/
>
> [D : nonempty subset of $\mathcal{H}$
>
> T_1 : quasinonexpansive operator, $: D \to D$ } one of them strictly quasinonexpansive
>
> T_2 : " " , $: D \to D$ } $\operatorname{Fix} T_1 \cap \operatorname{Fix} T_2 \neq \emptyset$
>
>] $\Rightarrow$
>
> (i) $\operatorname{Fix} T_1 T_2 = \operatorname{Fix} T_1 \cap \operatorname{Fix} T_2$
>
> (ii) $T_1 T_2$: quasinonexpansive
>
> (iii) $\{T_1, T_2\}$: strictly quasinonexpansive $\Rightarrow$ $T_1 T_2$: strictly quasinonexpansive

*/

now assume:

$\forall i \in \{1, 2, \ldots, m\}$ $T = T_1 T_2 \cdots T_i$: strictly quasinonexpansive

and $\operatorname{Fix} T = \bigcap_{i \in \{1, \ldots, m\}} \operatorname{Fix} T_i$

now consider. $\underbrace{(T_i)_{i \in \{1, \ldots, m+1\}}}_{D \to D}$: strictly quasinonexpansive, $\bigcap_{i=1}^{m+1} \operatorname{Fix} T_i \neq \emptyset$

set, $\underbrace{R_1 = T_1 \cdots T_m}_{\text{strictly quasinonexpansive by base assumption (induction hypothesis)}}$, $R_2 = T_{m+1}$: strictly quasinonexpansive

and $\operatorname{Fix} R_1 = \bigcap_{i=1}^{m} \operatorname{Fix} T_i$, /* by induction hypothesis */

$\operatorname{Fix} R_2 = \operatorname{Fix} T_{m+1}$

/* using */

$R_1 R_2 = T_1 T_2 \cdots T_{m+1}$: strictly quasinonexpansive, and

$\operatorname{Fix} \underbrace{T_1 T_2 \cdots T_{m+1}}_{R_1 \; R_2} = \operatorname{Fix} R_1 \cap \operatorname{Fix} R_2 = \bigcap_{i=1}^{m+1} \operatorname{Fix} T_i$:)

■

Corollary 4.37

$\underbrace{\text{averaged nonexpansive}}_{}$, $: D \to D$ (nonempty), $\bigcap_{i \in I} \operatorname{Fix} T_i \neq \emptyset$

$T = T_1 T_2 \cdots T_m \Rightarrow \operatorname{Fix} T = \bigcap_{i \in I} \operatorname{Fix} T_i$

Proof: Because averaged operators are strictly quasinonexpansive

apply 4.36.

■

5.1 Fejer Monotone Sequences

Friday, December 18, 2015 9:29 PM

Definition 5.1. (Fejer monotonicity w.r.t a set)

[C: nonempty subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t C $\overset{\text{def}}{\Longleftrightarrow}$ $\forall_{x\in C}\ \forall_{n\in\mathbb{N}}\ \|x_{n+1}-x\| \leq \|x_n - x\|$ (5.1)

12/18/2015 10:29 PM

Proposition 5.4.

[C: nonempty subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$, Fejer monotone w.r.t C] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: bounded

(ii) $\forall_{x\in C}\ (\|x_n - x\|)_{n\in\mathbb{N}}$: converges

(iii) $(d_C(x_n))_{n\in\mathbb{N}}$: decreasing, converges.

Proof: From defn (5.1) $\forall_{x\in C}\ \forall_{n\in\mathbb{N}}\ \|x_{n+1}-x\| \leq \|x_n - x\|$

Consider any $x\in C$

$\|x_1 - x\| \leq \|x_0 - x\|$

$\|x_2 - x\| \leq \|x_1 - x\| \leq \|x_0 - x\|$

$\vdots$

$\|x_n - x\| \leq \|x_0 - x\|$

So, $\forall_{x\in C}\ \forall_{n\in\mathbb{N}}\ \|x_n - x\| \leq \|x_0 - x\|$ // recall that $B(c,r) = \{x \mid \|x-c\| \leq r\}$

$\leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$ lies in $B(x; \|x_0 - x\|)$ $\leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded ■

(ii)

$\forall_{x\in C}\ \forall_{n\in\mathbb{N}}\ \underbrace{\|x_{n+1}-x\|}_{y_{n+1}} \leq \underbrace{\|x_n - x\|}_{y_n}$: monotone

From (i) $(x_n)_{n\in\mathbb{N}}$ bounded, so $(\|x_n - x\|)_{n\in\mathbb{N}}$: bounded

} converges /* a bounded monotone sequence always converges */

(iii) given

$\forall_{n\in\mathbb{N}}\ \forall_{x\in C}\ \|x_{n+1}-x\| \leq \|x_n - x\|$

$\Leftrightarrow \forall_{n\in\mathbb{N}}\ \underbrace{\inf_{x\in C} \|x_{n+1}-x\|}_{d_C(x_{n+1})} \leq \underbrace{\inf_{x\in C} \|x_n - x\|}_{d_C(x_n)}$ /* order limit theorem */

$\Leftrightarrow \forall_{n\in\mathbb{N}}\ d_C(x_{n+1}) \leq d_C(x_n)$

From (ii) $\forall_{x\in C}\ \|x_n - x\|_{n\in\mathbb{N}}$: bounded $\Rightarrow d_C(x_n)$: bounded by def. /* a bounded monotone sequence always converges */

$\Rightarrow$ $d_C(x_n)$: converges ■

Theorem 5.5.

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$, Fejer monotone w.r.t C, Nonempty, $\subseteq \mathcal{H}$

every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C]

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C.

12/18/2015 10:59 PM

Proof: $\forall_{x\in C}\ \|x_n - x\|_{n\in\mathbb{N}}$: converges /* Proposition 5.4. (ii) */

recall:

/* Lemma 2.39.

C : nonempty, $\subseteq \mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$:
- sequence in $\mathcal{H}$
- $\forall_{x\in C}\ \|x_n - x\|$: converges
- Every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C.

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C. */

} Combining these two we see all the givens imply $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C.

- $\forall_{x\in C}\ \|x_n - x\|$: converges
- Every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C.

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C. */

■

***Proposition 5.7.**

[C : nonempty closed convex subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$, Fejer monotone w.r.t. C] $\Rightarrow$

$\underbrace{(P_C x_n)_{n\in\mathbb{N}}}_{\text{called the shadow sequence}}$: converges strongly to a point in C.

***Corollary 5.8.**

[C : nonempty closed convex subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$, Fejer monotone w.r.t. C, $x_n \to \overset{\in C}{x}$]

$\Rightarrow$

$P_C x_n \to \overset{\in C}{x}$.

Proof :

12/19/2015 10:46 PM

/* Proposition 5.7.

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $\overset{\text{nonempty, closed, convex}}{C}$, sequence in $\mathcal{H}$

$(P_C x_n)_{n\in\mathbb{N}}$: $P_C x_n \to y \in C$ */

$\exists_{y\in C}\ P_C x_n \to y$

$\Rightarrow$ $x - P_C x_n \to x - y$

given

$x_n \to x$

$\Rightarrow$ $x_n - P_C x_n \to x - y$

One of the defining property of projection is $\langle \overset{\in C}{\square} - \underset{\text{closed, convex, nonempty}}{P_C} \overset{\in \mathcal{H}}{\square} \mid \overset{\in C}{\square} - P_C \overset{\in C}{\square} \rangle \leq 0$

so in our case, $x \in C$, $x_n \in \mathcal{H}$, $P_C x_n \in C$

$\therefore \langle x - P_C x_n \mid x_n - P_C x_n \rangle \leq 0 \quad \forall_{n\in\mathbb{N}}$

/* Lemma 2.41. (iii) says :

$(u_n)_{n\in\mathbb{N}} : \subseteq \mathcal{H},\ u_n \to \overset{\in\mathcal{H}}{u}$

$\langle x_n \mid u_n \rangle \to \langle x \mid u \rangle$

$(x_n)_{n\in\mathbb{N}} : \subseteq \mathcal{H},\ x_n \to \underbrace{x}_{\in\mathcal{H}}$ */

now,

$$\left.\begin{array}{l} (x_n - P_C x_n) \to (x-y) \\ (x - P_C x_n) \to \underbrace{(x-y)}_{\in\mathcal{H}} \end{array}\right\} \langle x - P_C x_n \mid x_n - P_C x_n \rangle \to \langle x-y \mid x-y \rangle = \|x-y\|^2 \Rightarrow \|x-y\|^2 \leq 0 \Leftrightarrow \|x-y\|^2 = 0$$

[$\because$ C convex, $x, y \in C$

$x - y$ is not necessarily in C]

$\therefore x = y$

so, $P_C x_n \to x$ ■

12/20/2015 2:19 PM

Theorem 5.11.

[C : nonempty closed convex subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t C, $\subseteq \mathcal{H}$]

(i) $(x_n)_{n\in\mathbb{N}}$: converges strongly to a point in C $\Leftrightarrow$

(ii) $(x_n)_{n\in\mathbb{N}}$: posseses a strong sequential cluster point in C $\Leftrightarrow$

(iii) $\underline{\lim}\ d_C(x_n) = 0$

Proof : /* proof strategy : (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii), (iii) $\Rightarrow$ (iv) */

(i) $\to$ (ii) : $(x_n)_{n\in\mathbb{N}}$: converges strongly to a point in C (say x)

Proof: /* Proof strategy: (i) ⇒ (ii), (ii) ⇒ (iii), (iii) ⇒ (iv) */

(i) → (ii): $(x_n)_{n\in\mathbb{N}}$: converges strongly to a point in C (say x)

⇒ all subsequences of $(x_n)_{n\in\mathbb{N}}$ converges strongly to that point in C (say x) /*

⇒ $\exists$ subsequence of $(x_n)_{n\in\mathbb{N}}$ that converges strongly to x in C (say x)

$\overset{\text{def}}{\Leftrightarrow}$ $\underset{\in C}{x}$: strong sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ ☺

/* Fact 1.9.
[$(x_a)_{a\in A}$: net in Hausdorff space X, $x_a \xrightarrow{\text{ex}} x$
$(x_{a(n)})_{n\in A}$: subnet of $(x_a)_{a\in A}$] ⇒ $x_{a(n)} \to x$ /* is a net in a Hausdorff space converges, then so does its any subnet to the same point */ */

(ii) ⇒ (iii):

(ii): $\underset{\in C}{x}$: strong sequential cluster point of $(x_n)_{n\in\mathbb{N}}$, now $\underline{\lim}\, d_C(x_n)$ by definition will exist. say $\underline{\lim}\, d_C(x_n) = p \in \mathbb{R}\cup\{\pm\infty\} \Rightarrow \underline{\lim}\, d_C(x_{k_n}) = p$ (1) for all $(x_{k_n})_{n\in\mathbb{N}}$ that are subsequences of $(x_n)_{n\in\mathbb{N}}$ ... (0)

$\Leftrightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}}$ $x_{k_n} \to x \in C$

now, $\forall_{n\in\mathbb{N}}$ $d_C(x_{k_n}) = \inf d(x_{k_n}, C) = \inf_{y\in C} \|x_{k_n} - y\| \le \|x_{k_n} - x\|$ [by definition, and $x \in C$]

$\Leftrightarrow \forall_{n\in\mathbb{N}}$ $d_C(x_{k_n}) \le \|x_{k_n} - x\| \Rightarrow 0 \le \underline{\lim}\, d_C(x_{k_n}) \le \overline{\lim}\, \|x_{k_n} - x\|$

now, $\lim \|x_{k_n} - x\| = \overline{\lim}\, \|x_{k_n} - x\| = 0$ $\therefore \underline{\lim}\, d_C(x_{k_n}) = 0$ ... (2) using

Result 1.2.1. /* These are extension of 1.2 results on one of (comes handy in dealing with sequences)
- $\forall_{n\in\mathbb{N}}$ $a_n \le a \Rightarrow \overline{\lim}\, a_n \le a$
- $\forall_{n\in\mathbb{N}}$ $b \le b_n \Rightarrow b \le \underline{\lim}\, b_n$
- $\forall_{n\in\mathbb{N}}$ $a_n \le a_n \le a_n \Rightarrow a_n \le \underline{\lim}\, a_n \le \overline{\lim}\, a_n \le a_n$
  $\underline{\lim}\, a_n$ exists in finite norm
  $\overline{\lim}\, a_n$ exists in finite norm
- $\forall_{n\in\mathbb{N}}$ $a_n \le b_n \Rightarrow \underline{\lim}\, a_n \le \overline{\lim}\, b_n$

/** Result 1.2.1. (A very important result) /* verified, uses result 1.2.2, and definition of convergence of a sequence */
[$(b_n)_{n\in\mathbb{N}} : \subseteq \mathbb{R}$] $(0 \le a \le b_n,\ b_n \to 0) \Rightarrow a = 0$

Now,

from (0), (1) and (2) we have: $\underline{\lim}\, d_C(x_n) = 0$

$\therefore \underline{\lim}\, d_C(x_n) = 0$ ☺

(iii) ⇒ (i):

/* Proposition 5.4.(iii):

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t C; $\subseteq \mathcal{H}$,

$(d_C(x_n))_{n\in\mathbb{N}}$: decreasing and converges */

: nonempty; $\subseteq \mathcal{H}$

now (iii) says: $\underline{\lim}\, d_C(x_n) = 0$

as $d_C(x_n)$: converges $\Rightarrow \lim d_C(x_n) = \underline{\lim}\, d_C(x_n) = 0$

$\Leftrightarrow d_C(x_n) \to 0$

/* recall Fact: 1.15. (iii):

$\underbrace{(\xi_a)_{a\in A}}_{\text{net in } [-\infty,\infty]}$: converges $\Leftrightarrow (\underline{\lim}\, \xi_a = \lim \xi_a = \overline{\lim}\, \xi_a)$

as $d_C(x_n) \subseteq [-\infty, \infty]$, we can apply it */

now $d_C(x_n) = \inf_{y\in C} \|x_n - y\|$

$= \|x_n - P_C x_n\|$ /* By definition: $d_C(x) = \|x - P_C x\|$ */

$\to 0$

$\Leftrightarrow$ $x_n - P_C x_n \to 0$

/* now, the shadow sequence of a Fejer monotone sequence w.r.t to a nonempty closed convex set C, always converges strongly to a point in C (Proposition 5.7)

$\Leftrightarrow \exists_{x\in C}$ $P_C x_n \to x$ */

/* By adding */

$\therefore x_n \to x \in C$: (i) ☺

■

* A linear convergence result for Fejer monotone sequence.

12/20/2015 3:16 PM

Theorem 5.12.

[C : nonempty closed convex subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: $\subseteq \mathcal{H}$, Fejer monotone w.r.t C

$\exists_{\kappa\in[0,1[}$ $\forall_{n\in\mathbb{N}}$ $d_C(x_{n+1}) \le \kappa\, d_C(x_n)$] ⇒

More precisely:

$(x_n)_{n\in\mathbb{N}}$: $\subseteq \mathcal{H}$, Fejer monotone w.r.t C

$\exists_{k\in[0,1[}\ \forall_{n\in\mathbb{N}}\ d_C(x_{n+1}) \leq k\, d_C(x_n)\ ] \Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges linearly to a point $x\in C$

More precisely:

$\forall_{n\in\mathbb{N}}\ \|x_n - x\| \leq 2k^n d_C(x_0)$

Proof: As, $\forall_{n\in\mathbb{N}}\ d_C(x_{n+1}) \leq \underbrace{k}_{\in[0,1[}\, d_C(x_n) \Leftrightarrow \|x_{n+1} - P_C x_{n+1}\| \leq k\,\|x_n - P_C x_n\|$ (By defn $\|x - P_C x\| = d_C(x)$) #

Now, $0 \leq d_C(x_{n+1}) \leq k\, d_C(x_n) \leq k^2 d_C(x_{n-1}) \leq k^3 d_C(x_{n-2}) \leq \dots \leq k^{n+1} d_C(x_{n-n})$ (Eq:A)

$\therefore\ \forall_{n\in\mathbb{N}}\ 0 \leq d_C(x_{n+1}),\quad d_C(x_{n+1}) \leq k^{n+1} d_C(x_0)$ → /* now :

$\Rightarrow\ 0 \leq \underline{\lim}\, d_C(x_{n+1}),\quad \underline{\lim}\, d_C(x_{n+1}) \leq \overline{\lim}\, k^{n+1} d_C(x_0) = \lim k^{n+1} d_C(x_0) = 0$

$\therefore\ \underline{\lim}\, d_C(x_{n+1}) = 0$

/* now :

* Theorem 5.11:
[C: nonempty closed convex subset of $\mathcal{H}$
$(x_n)_{n\in\mathbb{N}}$: $\subseteq \mathcal{H}$, Fejer monotone w.r.t C]
(i) $(x_n)_{n\in\mathbb{N}}$: converges strongly to a point in C ⇔
(ii) $(x_n)_{n\in\mathbb{N}}$: possesses a strong sequential cluster point in C
(iii) $\underline{\lim}\, d_C(x_n) = 0$ /* recall: limit inferior : $\underline{\lim}\, x_n = \sup_{C\in\mathcal{A}} \left[\begin{array}{l} \inf_{n\in A} x_n \\ \text{s.t. } A\in C \end{array} \right]$ of net $(x_n)_{n\in\mathbb{N}}$ */

/* there is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ that converges strongly to a point in C */

#/

$\exists_{x\in C}\ x_n \to x$

Recall definition of Fejer monotone sequence :

$\forall_{x\in C}\ \forall_{n\in\mathbb{N}}\ \|x_{n+1} - x\| \leq \|x_n - x\|$ now $P_C(x_n) \in C$ by definition

$x := P_C(x_n) \Rightarrow \forall_{n\in\mathbb{N}}$

$$\left.\begin{array}{l} \|x_{n+1} - P_C x_n\| \leq \|x_n - P_C x_n\| \\ \|x_{n+2} - P_C x_n\| \leq \|x_{n+1} - P_C x_n\| \\ \vdots \\ \|x_{n+m} - P_C x_n\| \leq \|x_{n+m-1} - P_C x_n\| \end{array}\right\} \Rightarrow \|x_{n+m} - P_C x_n\| \leq \|x_n - P_C x_n\| = d_C(x_n)$$

now $\forall_{n,m\in\mathbb{N}}$

$$\begin{aligned} \|x_n - x_{n+m}\| &= \|x_n - P_C x_n - x_{n+m} + P_C x_n\| \\ &= \|(x_n - P_C x_n) + (-(x_{n+m} - P_C x_n))\| \\ &\leq \|(x_n - P_C x_n)\| + \|-(x_{n+m} - P_C x_n)\| \\ &= \underbrace{\|x_n - P_C x_n\|}_{= d_C(x_n)} + \underbrace{\|x_{n+m} - P_C x_n\|}_{\leq d_C(x_n)} \leq 2 d_C(x_n) \end{aligned}$$

let $m \to \infty$ then $x_{n+m} \to x$

$\therefore\ \forall_{n\in\mathbb{N}}\ \|x_n - x\| \leq 2 d_C(x_n)$

But given: $d_C(x_n) \leq k\, \underbrace{d_C(x_{n-1})}_{\leq k\, d_C(x_{n-2})} \leq k^2 d_C(x_{n-2}) \dots \leq k^n d_C(\underset{n-n}{x_0})$

From (Eq:A)

$\therefore\ \forall_{n\in\mathbb{N}}\ \|x_n - x\| \leq 2k^n d_C(x_0)$ ■

5.2 KM iteration

Monday, December 21, 2015 9:39 AM

$x_{n+1} = \underset{\text{nonexpansive}}{T} x_n$ may fail to produce a fixed point

e.g., $T = -Id$ want to $(-Id)x = x$ whic is 0

now $x_0 \neq 0$, then, $x_1 = -Id\, x_0 = -x_0$

$x_2 = -Id\, x_1 = -Id(-x_0) = x_0$

$x_3 = -Id\, x_2 = -x_0$

$\vdots$

so $x_n = \begin{cases} x_0, & \text{if } n: \text{even} \\ -x_0, & \text{if } n: \text{odd} \end{cases}$

so it is not going to 0, rather fluctuating between $\pm x_0$

* The reason behind not converging: as $x_n - Tx_n \not\to 0$

$\underbrace{x_n - Tx_n \to 0}$ /* in this case, $x_n - Tx_n = x_n - x_{n+1} = \pm 2x_0 \neq 0$

is called asympotic regularity property, this needs to hold for convergence.

if it holds then $x_{n+1} = Tx_n$ converges to a fixed point.

* Theorem 5.13. (Necessity of asympotic regularity for classic iteration scheme to work)

[D: nonempty closed convex subset of $\mathcal{H}$,

$T: D \to D$, nonexpansive, $\operatorname{Fix} T \neq \emptyset$, $x_n - Tx_n \to 0$, $\forall_{n \in \mathbb{N}}$ $x_{n+1} = Tx_n$

$x_0 \in D$] $\Rightarrow$

(i) $(x_n)_{n \in \mathbb{N}}$: converges weakly to a point in $\operatorname{Fix} T$.

(ii) $\left.\begin{array}{l} D = -D, \\ T: \text{odd} \overset{\text{def}}{\Leftrightarrow} \forall_{x \in D}\ T(-x) = -Tx \end{array}\right\} \Rightarrow (x_n)_{n \in \mathbb{N}}$: converges strongly to a point in $\operatorname{Fix} T$.

[12/22/2015 11:36 AM]

* Theorem 5.14. (KM algorithm)

[D: nonempty closed convex subset of $\mathcal{H}$

$T: D \to D$, nonexpansive, $\operatorname{Fix} T \neq \emptyset$,

$(\lambda_n)_{n \in \mathbb{N}}$: sequence in $[0,1]$, $\sum_{n \in \mathbb{N}} \lambda_n (1 - \lambda_n) = +\infty$

$x_0 \in D$

$\forall_{n \in \mathbb{N}}$ $x_{n+1} = (1 - \lambda_n) x_n + \lambda_n T x_n$ /* KM iteration */

] $\Rightarrow$

(i) $(x_n)_{n \in \mathbb{N}}$: Fejer monotone w.r.t $\operatorname{Fix} T$

(ii) $(Tx_n - x_n)_{n \in \mathbb{N}}$: converges strongly to 0 /*so, Asympotic regularity property holds*/

(iii) $(x_n)_{n \in \mathbb{N}}$: converges weakly to a point in $\operatorname{Fix} T$.

Proof: Before starting the proof we show that $(x_n)_{n \in \mathbb{N}}$: well defined sequence in D

$x_0 \in D$, $Tx_0 \in D$ ($\because T: D \to D$, so. $\operatorname{dom} T = D$, $\operatorname{ran} T \subseteq D$)

$x_1 = (1 - \lambda_0) \underbrace{x_0}_{\in D} + \lambda_0 \underbrace{Tx_0}_{\in D} \in D$ [$\because D$: convex]

$\vdots$

$x_{n+1} = (1 - \lambda_n) x_n + Tx_n \in D$

$\therefore (x_n)_{n \in \mathbb{N}}$ is a sequence in D.

(i)

$\forall_{y \in \operatorname{Fix} T}$ $\forall_{n \in \mathbb{N}}$

$\| x_{n+1} - y \|^2$

$= \| (1 - \lambda_n) x_n + \lambda_n T x_n \underbrace{- y}_{-(1-\lambda_n) y - \lambda_n y} \|^2$

$= \| (1 - \lambda_n) x_n - (1 - \lambda_n) y + \lambda_n T x_n - \lambda_n y \|^2$

$$\underbrace{\qquad\qquad}_{-(1-\lambda_n)y-\lambda_n y}$$

$$= \|(1-\lambda_n)x_n-(1-\lambda_n)y+\lambda_n Tx_n-\lambda_n y\|^2$$

$$= \|(1-\lambda_n)(x_n-y)+\lambda_n(Tx_n-y)\|^2$$

$$= (1-\lambda_n)\|x_n-y\|^2+\lambda_n\|Tx_n-y\|^2-\lambda_n(1-\lambda_n)\|x_n-\cancel{y}+\cancel{y}-Tx_n\|^2 \quad /* \text{ Corollary 2.14. } \forall_{x,y\in H}\ \forall_{\kappa\in R}\ \|\kappa x+(1-\kappa)y\|^2=\kappa\|x\|^2+(1-\kappa)\|y\|^2-\kappa(1-\kappa)\|x-y\|^2 \ */$$

$$= (1-\lambda_n)\|x_n-y\|^2+\lambda_n\|Tx_n-\underset{\substack{\| \\ Ty}}{y}\|^2-\lambda_n(1-\lambda_n)\|x_n-Tx_n\|^2 \quad [\because y\in \text{Fix}\,T]$$

$$\# \quad \|Tx_n-Ty\|^2\leq\|x_n-y\|^2 \quad [\because T: \text{nonexpansive}]$$

$$\leq (1-\lambda_n)\|x_n-y\|^2+\lambda_n\|x_n-y\|^2-\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2 \quad \ldots (5.13)$$

$$= \|x_n-y\|^2-\underbrace{\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2}_{\text{negative so removing it will increase the value}}$$

$$\leq \|x_n-y\|^2$$

$\therefore \forall_{y\in \text{Fix}T}\ \forall_{n\in N}\ \|x_{n+1}-y\|^2\leq\|x_n-y\|^2 \leftrightarrow (x_n)_{n\in N}$: Fejer monotone w.r.t Fix T. ☺

(ii) From 5.13 we have

$$\forall_{y\in \text{Fix}T}\ \forall_{n\in N}\ \|x_{n+1}-y\|^2\leq\|x_n-y\|^2-\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2$$

$$\leftrightarrow \lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2\leq\|x_n-y\|^2-\|x_{n+1}-y\|^2 \quad : \text{telescopic term}$$

$$\Rightarrow \sum_{n=0}^{m}\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2\leq\sum_{n=0}^{m}(\|x_n-y\|^2-\|x_{n+1}-y\|^2) = \|x_0-y\|^2-\cancel{\|x_1-y\|^2}+\cancel{\|x_1-y\|^2}-\cancel{\|x_2-y\|^2}+\cdots\cancel{\|x_m-y\|^2}-\|x_{m+1}-y\|^2$$

$$= \|x_0-y\|^2-\|x_{m+1}-y\|^2$$

$$\Leftrightarrow \sum_{n=0}^{m}\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2\leq\|x_0-y\|^2-\underbrace{\|x_{m+1}-y\|^2}_{\text{negative number so removing it only increase the rest}}$$

$$\Rightarrow \sum_{n=0}^{m}\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2\leq\|x_0-y\|^2$$

$$\Rightarrow \quad 0\leq\sum_{\substack{n\in N\\ \cap \\ n\in\{0,1,\ldots,\infty\}}}\lambda_n(1-\lambda_n)\|Tx_n-x_n\|^2\leq\|x_0-y\|^2$$

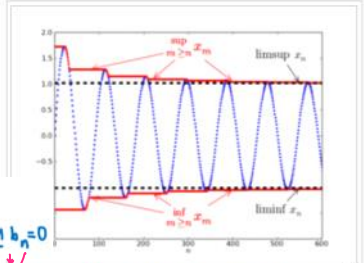
$\# \sum_{n\in N}\lambda_n(1-\lambda_n)=+\infty$: given

($\underline{\lim}$, $\overline{\lim}$ always exists for real number sequences)

now we donot know if $\lim\|Tx_n-x_n\|^2$ exists but $\underline{\lim}\|Tx_n-x_n\|^2$ must exist, and must be 0, otherwise the sequence wouldnot stay bounded. /*

$\therefore \underline{\lim}\|Tx_n-x_n\|^2=0 \leftrightarrow \underline{\lim}\|Tx_n-x_n\|=0$

/* Fact 1.3.2. (Establishing $\underline{\lim}(\cdot)=0$)
$[a_n,b_n\geq 0]\ (0\leq\sum_{n=0}^{\infty}a_nb_n\leq M,\ \sum_{n=0}^{\infty}a_n=+\infty)\Rightarrow \underline{\lim}\, b_n=0$ */



An illustration of limit superior and limit inferior. The sequence x_n is shown in blue. The two red curves approach the limit superior and limit inferior of x_n, shown as dashed black lines. In this case, the sequence *accumulates* around the two limits. The superior limit is the larger of the two, and the inferior limit is the smaller of the two. The inferior and superior limits agree if and only if the sequence is convergent (i.e., when there is a single limit).

now $\forall_{n\in N}$

$$\|Tx_{n+1}-\underbrace{x_{n+1}}_{\substack{\| \\ (1-\lambda_n)x_n+\lambda_nTx_n}}\| = \|Tx_{n+1}-(1-\lambda_n)x_n-\lambda_nTx_n\| = \|Tx_{n+1}-Tx_n+\underbrace{Tx_n-(1-\lambda_n)x_n-\lambda_nTx_n}_{\substack{(1-\lambda_n)Tx_n-(1-\lambda_n)x_n \\ =(1-\lambda_n)(Tx_n-x_n)}}\|$$

$$= \|Tx_{n+1}-Tx_n+(1-\lambda_n)(Tx_n-x_n)\|$$

$$\leq \underbrace{\|Tx_{n+1}-Tx_n\|}_{\leq\|x_{n+1}-x_n\| \ [T:\text{nonexpansive}]}+(1-\lambda_n)\|Tx_n-x_n\| \quad [\text{triangle inequality}]$$

$$\leq \underbrace{\|x_{n+1}-x_n\|}_{\lambda_n(Tx_n-x_n)}+(1-\lambda_n)\|Tx_n-x_n\| \quad /* \ x_{n+1}=(1-\lambda_n)x_n+\lambda_nTx_n=x_n+\lambda_n(Tx_n-x_n) \leftrightarrow x_{n+1}-x_n=\lambda_n(Tx_n-x_n) \ */$$

$\leq \|\underbrace{x_{n+1}-x_n}_{\lambda_n(Tx_n-x_n)}\| + (1-\lambda_n)\|Tx_n-x_n\| \qquad \Leftrightarrow x_{n+1}-x_n=\lambda_n(Tx_n-x_n)$ */

$= \lambda_n\|Tx_n-x_n\| + (1-\lambda_n)\|Tx_n-x_n\| = \|Tx_n-x_n\|$

$\therefore \forall_{n\in\mathbb{N}} \quad \|Tx_{n+1}-x_{n+1}\| \leq \|Tx_n-x_n\|$ So, $(\|Tx_n-x_n\|)_{n\in\mathbb{N}}$: monotonically decreasing sequence

$(\|Tx_n-x_n\|)_{n\in\mathbb{N}}$: bounded too, as $\|Tx_0-x_0\| \geq \|Tx_1-x_1\| \geq \cdots \geq \|Tx_n-x_n\| \geq \cdots \geq 0$

/* now

Result 1.24 (Convergence of monotonic sequence) /* $(x_n)_{n\in\mathbb{N}}$: monotonically increasing $\overset{\text{def}}{\Leftrightarrow} \forall_{n\in\mathbb{N}}\ x_{n+1} \geq x_n$
$(x_n)_{n\in\mathbb{N}}$: monotonically decreasing $\overset{\text{def}}{\Leftrightarrow} \forall_{n\in\mathbb{N}}\ x_{n+1} \leq x_n$ */

[$(x_n)_{n\in\mathbb{N}}$: monotonic]

$(x_n)_{n\in\mathbb{N}}$: converges $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded

$\therefore$ So, $(\|Tx_n-x_n\|)_{n\in\mathbb{N}}$: converges

$\therefore \lim \|Tx_n-x_n\| = \underbrace{\underline{\lim}\ \|Tx_n-x_n\|}_{=0} = \overline{\lim}\|Tx_n-x_n\| \Leftrightarrow Tx_n-x_n \to 0$

(ii)

in (i) we have shown that : $(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t FixT

$\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded /* Proposition 5.4.

• Proposition 5.4.
[C : nonempty subset of H
$(x_n)_{n\in\mathbb{N}}$: Fejer monotone sequence w.r.t C, $\subseteq H$] $\Rightarrow$
(i) $(x_n)_{n\in\mathbb{N}}$: bounded
(ii) $\forall_{z\in C}\ (\|x_n-z\|)_{n\in\mathbb{N}}$: converges
(iii) $(d_C(x_n))_{n\in\mathbb{N}}$: decreasing and converges

$\Rightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}}$ $(x_{k_n})_{n\in\mathbb{N}}$: weakly convergent /*

*Lemma 2.37.
[$(x_n)_{n\in\mathbb{N}}$: bounded sequence in H] $\Rightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ $(x_{k_n})_{n\in\mathbb{N}}$: weakly convergent */

now as $(x_n)_{n\in\mathbb{N}}$ stays in D, $(x_{k_n})_{n\in\mathbb{N}}$ will also stay in D

$\therefore$ x_{k_n} weakly converges to some point (say x) in D $\therefore x_{k_n} \rightharpoonup x \in D$

and by definition x : weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$

From (ii) we have $Tx_n - x_n \to 0$

$\Leftrightarrow x_n - Tx_n \to 0$

$\Rightarrow x_{k_n} - Tx_{k_n} \to 0$ /* If a net in a Hausdorff space converges, then so does any of its subnet and to the same point (Fact 1.9) */

Now we use Corollary 4.18.

A

Corollary 4.18. ★
[D : nonempty closed convex subset of H
T : (D → H, nonexpansive), x : some point in D
$(x_n)_{n\in\mathbb{N}}$: sequence in D, $x_n \rightharpoonup x$, $x_n - Tx_n \to 0$] $\Rightarrow x \in \text{Fix}\,T$ */

$\therefore x$: Fix T

So we have shown that any weakly convergent subsequence of $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in FixT

$\Updownarrow$

any weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in Fix T

Now we use Theorem 5.5.

Theorem 5.5.
[$(x_n)_{n\in\mathbb{N}}$: sequence in H, Fejer monotone w.r.t. $\underset{\neq\emptyset,\subseteq H}{C}$, every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in C] $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C

FixT [showed in (i)]

FixT

$x_n = (1-\lambda_{n-1})x_{n-1} + \lambda_{n-1}Tx_{n-1}$

$\therefore (x_n)_{n\in\mathbb{N}}$: converges weakly to a point in FixT. ☺

$\blacksquare$

* Proposition 5.15. (variant of KM algorithm for α-averaged operator)

[$\alpha \in]0,1[$

$T: \mathcal{H} \to \mathcal{H}$, α-averaged operator, $\operatorname{Fix} T \neq \emptyset$

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0, 1/\alpha]$, $\sum_{n\in\mathbb{N}} \lambda_n (1-\alpha\lambda_n) = +\infty$

$x_0 \in \mathcal{H}$

$\forall_{n\in\mathbb{N}} \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n)$] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $\operatorname{Fix} T$

(ii) $(Tx_n - x_n)_{n\in\mathbb{N}}$: converges strongly to 0

(iii) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Fix} T$.

Proof:

set, $R := (1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T$: nonexpansive /* T- α averaged $\overset{\text{(Proposition 4.25)}}{\Leftrightarrow}$ $(1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T$: nonexpansive */

$\mu_n := \alpha\lambda_n$

now $\operatorname{Fix} T = \operatorname{Fix} R$ /* $\forall x \quad x \in \operatorname{Fix} R \Leftrightarrow Rx = ((1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T)x = x$

$\Leftrightarrow (1-\frac{1}{\alpha})x + \frac{1}{\alpha}Tx = x - \frac{1}{\alpha}x + \frac{1}{\alpha}Tx = x$

$\Leftrightarrow Tx = x \Leftrightarrow x \in \operatorname{Fix} T \quad \therefore \operatorname{Fix} R = \operatorname{Fix} T$ */

now: the given iteration is:

$x_{n+1} = x_n + \lambda_n (Tx_n - x_n)$ /+ now, $R = (1-\frac{1}{\alpha})\mathrm{Id} + \frac{1}{\alpha}T$

$= x_n + \lambda_n \left(\underbrace{\alpha(R - (1-\frac{1}{\alpha})\mathrm{Id})x_n - x_n}_{\underbrace{\alpha(Rx_n - x_n + \frac{1}{\alpha}x_n)}_{(\alpha R x_n - \alpha x_n + x_n - x_n) = \alpha(Rx_n - x_n)}} \right)$ $\Leftrightarrow T = \alpha\left(R - (1-\frac{1}{\alpha})\mathrm{Id}\right)$ +/

$= x_n + \underbrace{\alpha\lambda_n}_{\mu_n \in [0,1]}(Rx_n - x_n) = x_n + \mu_n(\underbrace{R}_{\text{nonexpansive}}x_n - x_n)$

/+ : $\alpha \in [0,1]$, $\lambda_n \in [0, \frac{1}{\alpha}] \Rightarrow \underbrace{\alpha\lambda_n}_{\mu_n} \in \alpha[0, \frac{1}{\alpha}] = [0,1]$ +/

/+ set : $\tilde{D} := \mathcal{H}$, $\tilde{T} = R$, $\tilde{\lambda}_n = \mu_n$

#Theorem 5.14 (KM algorithm): A P

[$\tilde{D}$: nonempty closed convex subset of $\mathcal{H}$,
$\tilde{T}: \tilde{D} \to \tilde{D}$, nonexpansive, $\operatorname{Fix}\tilde{T} \neq \emptyset$
$(\tilde{\lambda}_n)_{n\in\mathbb{N}}$: sequence in $[0,1]$, $\sum_{n\in\mathbb{N}} \tilde{\lambda}_n(1-\tilde{\lambda}_n) = +\infty$
$x_0 \in \tilde{D}$
$\forall_{n\in\mathbb{N}} \quad x_{n+1} = (1-\tilde{\lambda}_n)x_n + \tilde{\lambda}_n \tilde{T}x_n$ /+ KM iteration +/] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $\operatorname{Fix}\tilde{T}$

(ii) $(\tilde{T}x_n - x_n)_{n\in\mathbb{N}}$: converges strongly to 0

(iii) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Fix}\tilde{T}$ +/

$\rightarrow$ the goals follow immediately. $\blacksquare$

*** Corollary 5.18.** (KM iteration for composition and convex combinations of nonexpansive operators)

[$(T_i)_{i\in I}$: finite family of $\underbrace{\alpha_i}_{\in]0,1[}$-averaged operators, $\forall_{i\in I}\ T_i : \mathcal{H}\to\mathcal{H}$, $\bigcap_{i\in I} \operatorname{Fix} T_i \neq \emptyset$,

$\tilde{i}$: $\begin{cases} \{(k,\ell) \mid k\in\{1,\dots,\overset{\in\mathbb{Z}_{++}}{p}\},\ \ell\in\{1,\dots,\overset{\in\mathbb{Z}_{++}}{m_k}\}\} \to I \\ \text{onto} \quad /* \ \operatorname{range}(\tilde{i}) = I \ */ \end{cases}$

$(\underbrace{\omega_k}_{>0})_{k=1}^{p}$: $\sum_{k=1}^{p} \omega_k = 1$

$\forall_{k\in\{1,\dots,p\}}$ $\begin{cases} I_k = \{\tilde{i}(k,1), \tilde{i}(k,2), \dots, \tilde{i}(k,m_k)\} \\ \rho_k = \dfrac{m_k}{(m_k-1) + \dfrac{1}{\max_{i\in I_k} \alpha_i}} \end{cases}$, $\alpha = \max_{1\le k\le p} \rho_k$

$(\lambda_n)_{n\in\mathbb{N}}$: $\subseteq [0, \frac{1}{\alpha}]$, $\sum_{n\in\mathbb{N}} \lambda_n (1-\alpha\lambda_n) = +\infty$

$\begin{cases} x_0 \in \mathcal{H} \\ \forall_{n\in\mathbb{N}} \quad x_{n+1} = (1-\lambda_n) x_n + \lambda_n \sum_{k=1}^{p} \omega_k \underbrace{\left(\prod_{j\in I_k} T_j\right)}_{T_{\tilde{i}(k,1)} \cdots T_{\tilde{i}(k,m_k)}} x_n \end{cases}$ /* KM iteration for composition of nonexpansive operators */

] $\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\bigcap_{i\in I} \operatorname{Fix} T_i$

[12/23/2015 1:31 PM]

Proof:

define $T = \sum_{k=1}^{p} \omega_k \left(\prod_{j\in I_k} T_j\right)$

then we can write the modified KM iteration as:

$$\forall_{n\in\mathbb{N}} \quad x_{n+1} = (1-\lambda_n) x_n + \lambda_n T x_n$$

12/23/2015 12:16 PM [Goal (i)]

*proof-strategy: we will show that, (i) T : α-averaged

once we have done that we can just plug in KM iteration convergence theorem for averaged operator:

* Proposition 5.15. (Revisit of KM algorithm for α-averaged operator)

[$\alpha \in]0,1[$

$T : \mathcal{H}\to\mathcal{H}$, α-averaged operator, $\operatorname{Fix} T \neq \emptyset$

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0, \frac{1}{\alpha}]$, $\sum_{n\in\mathbb{N}} \lambda_n (1-\alpha\lambda_n) = +\infty$

$x_0 \in \mathcal{H}$

$\forall_{n\in\mathbb{N}} \quad x_{n+1} = (1-\lambda_n) x_n + \lambda_n T x_n$]

(i) $(x_n)_{n\in\mathbb{N}}$: Fejer monotone wrt $\operatorname{Fix} T$

(ii) $(Tx_n - x_n)_n$: converges strongly to 0

(iii) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Fix} T$.

$\alpha \in]0,1[$, /* $\alpha \in]0,1[$

Proof: $\forall_k \quad \rho_k = \dfrac{m_k}{m_k - 1 + \dfrac{1}{\underbrace{\max_{i\in I_k} \alpha_i}_{\in]0,1[}}}$

so, $\left(\max_{i\in I_k} \alpha_i\right) \in]0,1[\Rightarrow \dfrac{1}{\max_{i\in I_k} \alpha_i} > 1$

$\Rightarrow \dfrac{1}{\max_{i\in I_k} \alpha_i} - 1 > 0 \Rightarrow \left(\underbrace{m_k}_{\in\mathbb{Z}_{++}} - 1 + \dfrac{1}{\max_{i\in I_k} \alpha_i}\right) > \underbrace{m_k}_{\in\mathbb{Z}_{++}}$

$\Rightarrow \rho_k = \dfrac{m_k}{m_k - 1 + \dfrac{1}{\max_{i\in I_k} \alpha_i}} \in]0,1[\Rightarrow \alpha = \max_{1\le k\le p} \rho_k \in]0,1[$ ■ */

12/23/2015 1:27 PM Goal (ii)

and then we show that (ii) $\operatorname{Fix} T = \bigcap_{i\in I} \operatorname{Fix} T_i$

Now let's start the proof: $\forall_{k\in\{1,\dots,p\}}$ $R_k = \prod_{j\in I_k} \overset{\alpha_j\text{-averaged}}{T_j}$

Recall:

* Proposition 4.32. (Composition of averaged operators) ★★★

[D: nonempty subset of $\mathcal{H}$

m: strictly positive integers

$I = \{1, \dots, m\}$

$(\alpha_i)_{i\in I}$: real numbers, $\alpha_i \in]0,1[$

$(T_i)_{i\in I}$: (family of operators, : $D \to D$, $\forall_{i\in I}$ T_i : α_i-averaged)]

$\left(T = T_1 \cdots T_m,\ \alpha = \dfrac{m}{m-1+\dfrac{1}{\max_{i\in I} \alpha_i}}\right) \Rightarrow T$: α-averaged

So: imposing our condition for this case we have:

* Proposition 4.32. (Composition of averaged operators) ★★★

[$\cancel{D}$: nonempty subset of $\mathcal{H}$

$\mathcal{H}$

m_k ~~m~~: strictly positive integers

I_k ~~I~~ $= \{1, \dots, \cancel{m}\}$ m_k

$(\alpha_i)_{i\in I}$: real numbers, $\alpha_i \in]0,1[$

$(T_i)_{i\in I}$: (family of operators, : $\cancel{D}\ \mathcal{H} \to \cancel{D}\ \mathcal{H}$, $\forall_{\cancel{i\in I}}$ T_i : α_i-averaged)]

$I_k = \{\dots\}$

$(\alpha_i)_{i\in I}$: real numbers, $\alpha_i \in]0,1[$

$(T_i)_{i\in I}$: (family of operators, $: \mathcal{H}\to\mathcal{H}$, $\forall_{i\in I_k}$ T_i : α_i-averaged)]

$$\left(\underset{R_k}{T} = \underbrace{T_1 \cdots T_{m_k}}_{T_{m_k}},\ \underset{\rho_k}{\alpha} = \frac{\alpha_1}{\alpha_1 - 1 + \frac{1}{m_k}\max_{i\in I_k}\alpha_i} \right) \Rightarrow \underset{R_k}{T} : \underset{\rho_k}{\alpha}\text{-averaged}$$

by given, and also $\rho_k \in]0,1[$

$\therefore R_k : \rho_k$-averaged

Now recall corollary 4.37:

Corollary 4.37. averaged operators, $: D\to D$, $\bigcap_{i\in I} \mathrm{Fix}\, T_i \neq \emptyset$

$$\mathrm{Fix}\left(\prod_{i=1}^{m} T_i\right) = \bigcap_{i=1}^{m} (\mathrm{Fix}\, T_i)$$

using this:

$$\mathrm{Fix}(R_k) = \mathrm{Fix}\left(\prod_{j\in I_k} T_j\right) = \bigcap_{j\in I_k} (\mathrm{Fix}\, T_j) \quad \left[\because \bigcap_{i\in I} \mathrm{Fix}\, T_i \neq \emptyset \text{ (given)} \quad \therefore \bigcap_{i\in I_k} \mathrm{Fix}\, T_i \neq \emptyset \text{ as } I_k \subseteq I\right]$$

So, we have $R_k : \rho_k$-averaged, $\mathrm{Fix}(R_k) = \bigcap_{i\in I_k} (\mathrm{Fix}\, T_i)$

Now, $T = \sum_{k=1}^{p} \omega_k \left(\prod_{j\in I_k} T_j\right) = \sum_{k=1}^{p} \omega_k R_k$

given : $\underbrace{(\omega_k)_{k=1}^{p}}_{>0} : \sum_{k=1}^{p} \omega_k = 1$

now we plug in proposition 4.30 regarding addition of averaged operator

Proposition 4.30. (Addition of averaged operators)

[D: nonempty subset of $\mathcal{H}$

$\underbrace{(T_i)_{i\in I}}_{:D\to\mathcal{H}}$: (• finite family of nonexpansive operators

• $\forall_{i\in I}$ T_i : $\underbrace{\alpha_i}_{\in]0,1]}$-averaged)

$\underbrace{(\omega_i)_{i\in I}}_{\in]0,1]}$: $\sum_{i\in I} \omega_i = 1$] $\Rightarrow$ $\sum_{i\in I} \omega_i T_i$: $\left(\max_{i\in I} \alpha_i\right)$: averaged.

Proposition 4.30. (Addition of averaged operators)

[D: nonempty subset of $\mathcal{H}$

$(T_i)_{i\in I}$: (• finite family of nonexpansive operators

$R_k : \rho_k$-averaged

• $\forall_{k\in\{1,\dots,p\}}$ T_i : α_i-averaged)

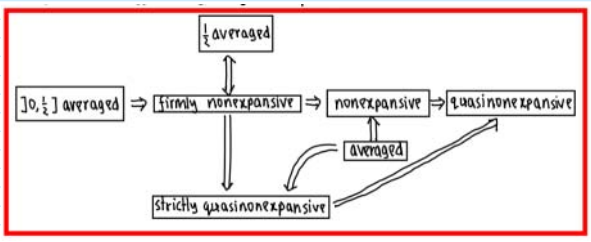
$\in]0,1]$

$(\omega_i)_{i\in I_k,\ k\in\{1,\dots,p\}}$: $\sum_{i\in I_k} \omega_i = 1$] $\Rightarrow$ $T = \sum_{k=1}^{p} \omega_k T_i$: $\left(\max_{1\le k\le p} \rho_k\right)$: averaged.

$\therefore T : \underbrace{\left(\max_{1\le k\le p} \rho_k\right)}_{\alpha}$: averaged

$\leftrightarrow$ $T : \alpha$-averaged 12/23/2015 12:16 PM [Goal (i) reached

Now let us work on goal: (ii): show $\mathrm{Fix}\, T = \bigcap_{i\in I} \mathrm{Fix}\, T_i \neq \emptyset$ first look at the diagram:



$R_k : \rho_k$-averaged
$\Downarrow$
nonexpansive
$\Downarrow$
quasinonexpansive

now recall proposition 4.34. and plug it in for our case

Proposition 4.34.

[D: nonempty subset of $\mathcal{H}$

$R_k : \rho_k$-averaged
$\Downarrow$

$(R_k)_{k\in\{1,\dots,p\}}$: finite family of quasinonexpansive operators, $: \mathcal{H}\to\mathcal{H}$, $\bigcap_{i\in I_k,\ k\in\{1,\dots,p\}} \mathrm{Fix}\, R_k \neq \emptyset$

$(\omega_k)_{k\in\{1,\dots,p\}}$: strictly positive real numbers, $\sum_{k\in\{1,\dots,p\}} \omega_k = 1$]

$\Rightarrow \mathrm{Fix} \sum_{k\in\{1,\dots,p\}} \omega_k R_k = \bigcap_{i\in I_k,\ k\in\{1,\dots,p\}} \mathrm{Fix}\, R_k \neq \emptyset$: given

proof: $\bigcap_{k=1}^{p} \underbrace{\mathrm{Fix}\, R_k}_{\bigcap_{i\in I_k} \mathrm{Fix}\, T_i} = \bigcap_{k=1}^{p} \bigcap_{i\in I_k} \mathrm{Fix}\, T_i = \bigcap_{i\in I} \mathrm{Fix}\, T_i \neq \emptyset$ [given]

as $\tilde{i}$ onto $\bigcap_{k=1}^{p} \bigcap_{i\in I_k} \equiv \bigcap_{i\in I}$

and as $\tilde{i}$: onto

$\Rightarrow \operatorname{Fix} \sum_{k\in\{1,\dots,p\}} \mu_k R_k = \bigcap_{k\in\{1,\dots,p\}} \operatorname{Fix} R_k \neq \emptyset$: given

$\therefore \operatorname{Fix} T = \operatorname{Fix}\left(\sum_{k=1}^{p} \mu_k R_k\right) = \bigcap_{k=1}^{p} \operatorname{Fix} R_k = \bigcap_{i\in I} \operatorname{Fix} T_i \neq \emptyset$ ∴ 12/23/2015 1:27 PM Goal (ii) reached

Now we apply proposition 5.15:

* Proposition 5.15. (Variant of KM algorithm for κ-averaged operator)

[$\kappa \in]0,1[$ ✓

$T: \mathcal{H} \to \mathcal{H}$, κ-averaged operator, $\operatorname{Fix} T \neq \emptyset$ ✓

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0, \frac{1}{\kappa}]$, $\sum_{n\in\mathbb{N}} \lambda_n(1-\kappa\lambda_n) = +\infty$ ✓

$x_0 \in \mathcal{H}$ ✓ ✓

$\forall_{n\in\mathbb{N}} \quad x_{n+1} = (1-\lambda_n)x_n + \lambda_n T x_n$] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $\operatorname{Fix} T$ ✓

(ii) $(Tx_n - x_n)_n$: converges strongly to 0 ✓

(iii) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Fix} T$. ✓

$\operatorname{Fix} T = \bigcap_{i\in I} \operatorname{Fix} T_i$

$\therefore (x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\bigcap_{i\in I} \operatorname{Fix} T_i$ ∎

∎

5.3 Iterating Compositions of Averaged Operators

Thursday, December 24, 2015 7:22 PM

Theorem 5.22.

[D: nonempty weakly sequentially closed subset of $\mathcal{H}$ /* weakly sequentially closed set: weak limit of every weakly convergent sequence in the set is also in the set.

$(T_i)_{i\in I}$: family of α_i-averaged operators, $: D \to D$, $\mathrm{Fix}(T_1 \dots T_m) \neq \emptyset$

$I = \{1, 2, \dots, m\}$, $\alpha_i \in]0,1[$

/* note that this is classic iteration, not the KM iteration */

$x_0 \in D$

$\forall_{n\in\mathbb{N}}\ x_{n+1} = (T_1 \cdots T_m) x_n$] $\Rightarrow$

- $x_n - (T_1 \cdots T_m) x_n \to 0$ (5.20)

- $\begin{pmatrix} \exists y_1 \in \mathrm{Fix}\, T_1 T_2 \dots T_{m-1} T_m \\ \exists y_2 \in \mathrm{Fix}\, T_2 T_3 \cdots T_m T_1 \\ \vdots \\ \exists y_m \in \mathrm{Fix}\, T_m T_1 \cdots T_{m-2} T_{m-1} \end{pmatrix}$:

$$x_n \to y_1 = T_1 y_2 \quad (5.21)$$
$$T_m x_n \to y_m = T_m y_1 \quad (5.22)$$
$$T_{m-1} T_m x_n \to y_{m-1} = T_{m-1} y_m \quad (5.23)$$
$$\vdots$$
$$T_3 \cdots T_m x_n \to y_3 = T_3 y_4 \quad (5.24)$$
$$T_2 T_3 \cdots T_m x_n \to y_2 = T_2 y_3 \quad (5.25)$$

[12/25/2015 12:37 PM] **Proof:**

Let $T = T_1 \cdots T_m$, then the iteration becomes $x_{n+1} = T x_n$

$\beta_i = (1-\alpha_i)/\alpha_i$

$y \in \mathrm{Fix}\, T$

then $\|x_{n+1} - y\|^2 = \|Tx_n - Ty\|^2 = \|T_1 \overbrace{(T_2 \cdots T_m) x_n}^{\alpha} - T_1 \overbrace{(T_2 \cdots T_m) y}^{\beta}\|^2$

$= \|T_1 \alpha - T_1 \beta\|^2 \leq \|\alpha - \beta\|^2 - \underbrace{\frac{1-\alpha_1}{\alpha_1}}_{\beta_1} \|(\mathrm{Id} - T_1)\alpha - (\mathrm{Id} - T_1)\beta\|^2$

$= \|(T_2 T_3 \cdots T_m) x_n - (T_2 T_3 \cdots T_m) y\|^2 -$

$\beta_1 \|(\mathrm{Id} - T_1)(T_2 \cdots T_m) x_n - (\mathrm{Id} - T_1)(T_2 \cdots T_m) y\|^2$

now $\|(T_2 T_3 \cdots T_m) x_n - (T_2 T_3 \cdots T_m) y\|^2$

$= \|T_2 (T_3 \cdots T_m) x_n - T_2 (T_3 \cdots T_m) y\|^2$

$\leq \|(T_3 \cdots T_m) x_n - (T_3 \cdots T_m) y\|^2 -$

$\beta_2 \|(\mathrm{Id} - T_2)(T_3 \cdots T_m) x_n - (\mathrm{Id} - T_2)(T_3 \cdots T_m) y\|^2$

(using Proposition 5.25 again)

Proposition 5.25. (Different faces of an α-averaged operator) ★★★

[D: nonempty subset of $\mathcal{H}$

$T: D \to \mathcal{H}$, nonexpansive

$\alpha \in]0,1[$]

(i) T: α-averaged $\Leftrightarrow$

(ii) $\left(1 - \frac{1}{\alpha}\right)\mathrm{Id} + \left(\frac{1}{\alpha}\right) T$: nonexpansive $\Leftrightarrow$

(iii) $\forall_{x\in D} \forall_{y\in D}\ \|Tx - Ty\|^2 \leq \|x - y\|^2 - \frac{1-\alpha}{\alpha} \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \Leftrightarrow$

(iv) $\forall_{x\in D} \forall_{y\in D}\ \|Tx - Ty\|^2 + (1 - 2\alpha)\|x - y\|^2 \leq 2(1-\alpha) \langle x - y \mid Tx - Ty \rangle$

then,

$\|x_{n+1} - y\|^2 \leq \|T_3 \cdots T_m x_n - T_3 \cdots T_m y\|^2$

$- \beta_2 \|(\mathrm{Id} - T_2)(T_3 \cdots T_m) x_n - (\mathrm{Id} - T_2)(T_3 \cdots T_m) y\|^2$

$- \beta_1 \|(\mathrm{Id} - T_1)(T_2 \cdots T_m) x_n - (\mathrm{Id} - T_1)(T_2 \cdots T_m) y\|^2$

$= \|T_3 \cdots T_m x_n - T_3 \cdots T_m y\|^2 - \sum_{i=2}^{1} \beta_i \|(\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) x_n - (\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) y\|^2$

$\vdots$

/* similar logic */

$\leq \|x_n - y\|^2 - \beta_m \|(\mathrm{Id} - T_m) x_n - (\mathrm{Id} - T_m) y\|^2$

$- \beta_{m-1} \|(\mathrm{Id} - T_{m-1}) T_m x_n - (\mathrm{Id} - T_{m-1}) T_m y\|^2$

$\vdots$

$- \beta_2 \|(\mathrm{Id} - T_2)(T_3 \cdots T_m) x_n - (\mathrm{Id} - T_2)(T_3 \cdots T_m) y\|^2$

$- \beta_1 \|(\mathrm{Id} - T_1)(T_2 \cdots T_m) x_n - (\mathrm{Id} - T_1)(T_2 \cdots T_m) y\|^2 \quad \ldots (\text{eq.}: 0)$

$= \|x_n - y\|^2 - \underbrace{\sum_{i=m}^{1} \beta_i \|(\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) x_n - (\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) y\|^2}_{\text{negative number, so removing it will make the rest only larger}} \quad (\text{eq.}: 1)$

$\leq \|x_n - y\|^2$

$\therefore \forall_{y\in \mathrm{Fix}\, T}\ \|x_{n+1} - y\|^2 \leq \|x_n - y\|^2 \Leftrightarrow (x_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $\mathrm{Fix}\, T$

$\Rightarrow \forall_{y\in \mathrm{Fix}\, T} (\|x_n - y\|)_{n\in\mathbb{N}}$: converges

Proposition 5.4. (Some key properties of Fejer monotone sequences)

[C: nonempty subset of $\mathcal{H}$

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone sequence w.r.t C, $\subseteq \mathcal{H}$] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: bounded

(ii) $\forall_{x\in C} (\|x_n - x\|)_{n\in\mathbb{N}}$: converges

now $\forall_{y\in \mathrm{Fix}\, T}\ \sum_{i=m}^{1} \beta_i \|(\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) x_n - (\mathrm{Id} - T_i)(\prod_{j=i+1}^{m} T_j) y\|^2 \leq$

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone sequence w.r.t C, 5.21] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: bounded

(ii) $\forall_{x\in C}\ (\|x_n-x\|)_{n\in\mathbb{N}}$: converges

(iii) $(d_C(x_n))_{n\in\mathbb{N}}$: decreasing and converges */

now $\forall_{y\in \mathrm{Fix}T}$ $\displaystyle\sum_{i=m}^{1}\beta_i\,\|(Id-T_i)(\prod_{j=i+1}^{m}T_j)x_n-(Id-T_i)(\prod_{j=i+1}^{m}T_j)y\|^2\leq$

$\|x_n-y\|^2-\|x_{n+1}-y\|^2$

(eq:0) $\Rightarrow$ // removing negative terms associated with

// $\beta_{m-1},\dots,\beta_1$

$\|x_{n+1}-y\|^2\leq\|x_n-y\|^2-\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2$

$\Leftrightarrow\ 0\leq\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2\leq\|x_n-y\|^2-\|x_{n+1}-y\|^2$

/* now Fact 13.1: $a_n\leq b_n\Rightarrow \underline{\lim}\,a_n\leq\underline{\lim}\,b_n,\ \overline{\lim}\,a_n\leq\overline{\lim}\,b_n$ */

$0\leq\underline{\lim}\,\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2\leq\underline{\lim}\,(\|x_n-y\|^2-\|x_{n+1}-y\|^2)=\lim(\|x_n-y\|^2-\|x_{n+1}-y\|^2)=0$

$0\leq\overline{\lim}\,\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2\leq\overline{\lim}\,(\|x_n-y\|^2-\|x_{n+1}-y\|^2)=\lim(\|x_n-y\|^2-\|x_{n+1}-y\|^2)=0$

$\underline{\lim}\,\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2=\overline{\lim}\,\beta_m\|(Id-T_m)x_n-(Id-T_m)y\|^2=\lim\underbrace{\beta_m}_{>0}\|(Id-T_m)x_n-(Id-T_m)y\|^2=0$

$\therefore\ \|(Id-T_m)x_n-(Id-T_m)y\|^2\to0\ \overset{eq.}{\Leftrightarrow}\ (Id-T_m)x_n-(Id-T_m)y=x_n-y-T_mx_n+T_my\to0$ (5.27)

Similarly. $(Id-T_{m-1})T_mx_n-(Id-T_{m-1})T_my=T_mx_n-T_my-T_{m-1}T_mx_n+T_{m-1}T_my\to0$ (5.28)

$\vdots$

$(Id-T_2)T_3\cdots T_mx_n-(Id-T_2)T_3\cdots T_my=T_3\cdots T_mx_n-T_3\cdots T_my-T_2\cdots T_mx_n+T_2\cdots T_my\to0$ (5.29)

$(Id-T_1)T_2\cdots T_mx_n-(Id-T_1)T_2\cdots T_my=T_2\cdots T_mx_n-T_2\cdots T_my-\underbrace{T_1\cdots T_m}_{T}x_n+\underbrace{T_1\cdots T_m}_{T}y\to0$ (5.30)

adding

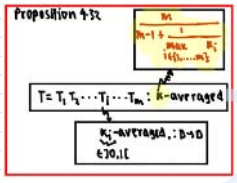
$x_n-y+Tx_n-\underbrace{Ty}_{y\ //\ y\in\mathrm{Fix}T}=x_n-Tx_n\to0$

$\therefore\ x_n-Tx_n\to0$: asymptotic regularity property holds.

recall that $T=T_1\cdots T_m$: $\dfrac{m}{m-1+\dfrac{1}{\max_{i=1,\dots,m}\alpha_i}}$ averaged

$\Rightarrow$ T: nonexpansive

Now we apply Theorem 5.13:



/* Theorem 5.13. (necessity of asymptotic regularity for classic nonexpansive iteration to work)

[D: nonempty closed convex subset of $\mathcal{H}$ ✓ /* Showing this property is the key */

$T: D\to D$, nonexpansive operator. $\mathrm{Fix}\,T\neq\emptyset$, $x_n-Tx_n\to0$ ✓

$x_0\in D$, $\forall_{n\in\mathbb{N}}\ x_{n+1}=Tx_n$] $\Rightarrow$ (given)

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\mathrm{Fix}\,T$ */

so, $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\mathrm{Fix}\,T$. Suppose the point is y_1.

so, $x_n\rightharpoonup y_1$, $y_1\in\mathrm{Fix}\,T=T_1\cdots T_m$ (goal 5.21 reached)

from 5.27.

$x_n-y+T_mx_n+T_my\to0\Rightarrow x_n-T_mx_n\to y-T_my\Leftrightarrow T_mx_n-x_n\to T_my-y\quad\forall y\in\mathrm{Fix}T$

set $y:=y_1\in\mathrm{Fix}T$

$\Rightarrow\ T_mx_n-x_n\to T_my_1-y_1$

$\Rightarrow\ T_mx_n-x_n\rightharpoonup T_my_1-y_1$

$\therefore\ T_mx_n\rightharpoonup T_my_1$ /* now $y_1\in\mathrm{Fix}T\Leftrightarrow Ty_1=T_1\cdots T_my_1=y_1$

$\Rightarrow T_m(T_1\cdots T_{m-1})T_my_1=T_my_1\Leftrightarrow \overbrace{T_my_1=y_m}^{//\text{say}}\in\mathrm{Fix}\,T_mT_1\cdots T_{m-1}$ */

$\therefore\ T_mx_n\rightharpoonup T_my_1=y_m$, $y_m\in\mathrm{Fix}\,T_mT_1\cdots T_{m-1}$: goal (5.22) proved

from (5.28)

$(Id-T_{m-1})T_mx_n-(Id-T_{m-1})T_my=T_mx_n-T_my-T_{m-1}T_mx_n+T_{m-1}T_my\to0$

from (5-28)

$$(Id-T_{m-1})T_m x_n-(Id-T_{m-1})T_m y = T_m x_n - T_m y - T_{m-1}T_m x_n + T_{m-1}T_m y \to 0$$

$$\Rightarrow \quad T_m x_n - T_{m-1}T_m x_n \to T_m y - T_{m-1}T_m y \qquad \forall y \in \mathrm{Fix}\,T$$

set $y := y_1 \in \mathrm{Fix}\,T$

$\Rightarrow$

$$T_m x_n - T_{m-1}T_m x_n \to T_m y_1 - T_{m-1}T_m y_1$$

$$\Leftrightarrow \quad T_{m-1}T_m x_n - T_m x_n \to T_{m-1}T_m y_1 - T_m y_1$$

$$\Rightarrow \quad T_{m-1}T_m x_n - T_m x_n \to T_{m-1}T_m y_1 - T_m y_1$$

$$T_{m-1}T_m x_n \to T_{m-1}\underbrace{T_m y_1}_{y_m:\ y_m \in \mathrm{Fix}\,T_m T_1 \cdots T_{m-1}} = T_{m-1} y_m \quad \text{At now} \quad T_m T_1 \cdots \underbrace{T_{m-1} y_m}_{y_{m-1}\ \text{is say}} = y_m \Rightarrow T_m T_1 \cdots T_{m-2} y_{m-1} = y_m \Rightarrow (T_{m-1}T_m T_1 \cdots T_{m-2}) y_{m-1} = T_{m-1} y_m = y_{m-1}$$

$$\therefore y_{m-1} = T_{m-1} y_m \in \mathrm{Fix}\,T_{m-1}T_m T_1 \cdots T_{m-2} \quad \#/$$

$$\therefore T_{m-1}T_m x_n \to y_{m-1} = T_{m-1} y_m,\ y_{m-1} \in \mathrm{Fix}\,T_{m-1}T_m T_1 \cdots T_{m-2}$$ (goal (5-23) proved)

Proceeding like this we can construct rest of 5-21-5-25. ∎

* Corollary 5-23 (POCS algorithm)

[$(C_i)_{i\in I}$: family of nonempty closed convex subsets of $\mathcal{H}$

$(P_i)_{i\in I}$: projectors on $(C_i)_{i\in I}$, $\mathrm{Fix}(P_1 P_2 \cdots P_m) \neq \emptyset$

$x_0 \in \mathcal{H}$ $\quad |I| = m$

$\forall_{n\in\mathbb{N}}\ x_{n+1} = P_1 P_2 \cdots P_m x_n$] $\Rightarrow$

$$\exists_{(y_1,\dots,y_m)\in C_1\times\cdots\times C_m} \begin{cases} x_n \to y_1 = P_1 y_2 \\ P_m x_n \to y_m = P_m y_1 \\ P_{m-1}P_m x_n \to y_{m-1} = P_{m-1} y_m \\ \vdots \qquad \vdots \\ P_3 \cdots P_m x_n \to y_3 = P_3 y_4 \\ P_2 \cdots P_m x_n \to y_2 = P_2 y_3 \end{cases}$$

6.1. Convex Cones

Monday, December 28, 2015 8:15 PM

convex cones : • lies between linear subspace and convex set.

• arises in the study of convex sets via tangent cone and normal cone operators.

6.1. Convex Cones.

$\underset{\in \mathcal{H}}{C}$: cone $\overset{\text{def}}{\Leftrightarrow}$ $C = \mathbb{R}_{++} C$

so $\mathcal{H}$ itself is a cone

cone C : conical hull of C $\overset{\text{def}}{\Leftrightarrow}$ smallest cone in $\mathcal{H}$ containing C

$\overline{\text{cone}}\, C$: closed conical hull of C $\overset{\text{def}}{\Leftrightarrow}$ smallest closed cone in $\mathcal{H}$ containing C

• Proposition 6.2.

[C: subset of $\mathcal{H}$] $\Rightarrow$

(i) cone $C = \mathbb{R}_{++} C$

(ii) $\overline{\text{cone}\, C} = \overline{\text{cone}}\, C$

(iii)
cone(conv C) = conv (cone C) : smallest convex cone containing C.

(iv) $\overline{\text{cone}}$(conv C) = $\overline{\text{conv}}$ (cone C) : smallest closed convex cone containing C.

Proof :

(i) $D = \mathbb{R}_{++} C = \{\lambda x \mid x \in C, \lambda > 0\}$ /* $y \in D \Leftrightarrow \exists_{\lambda > 0} \exists_{x \in C}\ \lambda x = y$ */

first show: cone $C \subseteq \mathbb{R}_{++} C$ . . . (1)

$\Rightarrow$ D: cone $\wedge$ $C = \{1x \mid x \in C, \lambda = 1\} \subseteq D$

$\Rightarrow$ cone$(C) \subseteq$ cone$(D) = D$

/* cone C : • $C \subseteq$ cone C
• $\mathbb{R}_{++}$ cone C = cone C
• $\forall_{E: \text{cone}}$ $C \subseteq E$, E: cone cone $C \subseteq E$ /*smallest cone containing C*/
now D: cone and $C \subseteq D \Rightarrow$ cone$(C) \subseteq D$ */

now show $D \subseteq$ cone(C)

$\Leftrightarrow \forall_{y \in D}\ y \in$ cone(C)

$\forall_y$

$y \in D \Leftrightarrow \exists_{\lambda > 0} \exists_{x \in C}\ y = \lambda x$

$\Rightarrow x \in C \subseteq$ cone $C \Rightarrow x \in$ cone $C \Rightarrow \underbrace{\lambda x}_{= y} \in$ cone C $\therefore y \in$ cone C

so, $\forall_{y \in D}\ y \in$ cone$(C) \Leftrightarrow D \subseteq$ cone C (2)

(1)+(2) $\Rightarrow$ cone $C = \mathbb{R}_{++} C$. ☺

(ii) $\overline{\text{cone}}\, C$: closed cone $\Rightarrow$ cone$(\overline{\text{cone}}\, C)$: the smallest cone containing $\overline{\text{cone}}\, C$ $= \overline{\text{cone}}\, C \Rightarrow \overline{\text{cone}(\overline{\text{cone}}\, C)} = \overline{\text{cone}}\, C$

$C \subseteq$ cone $C \subseteq \overline{\text{cone}}\, C \Rightarrow$ /* By definition of conic hull

now, cone C : smallest cone containing C,

and cone$(\overline{\text{cone}}\, C)$ is another cone containing C

cone $C \subseteq$ cone$(\overline{\text{cone}}\, C)$

$\Rightarrow \overline{\text{cone}}\, C \subseteq \overline{\text{cone}(\overline{\text{cone}}\, C)} = \overline{\text{cone}}\, C$ /* $\because A \subseteq B \Rightarrow \overline{A} = \overline{B}$ */
. . . (3)

Again, $\overline{\text{cone}}\, C$ = smallest closed cone containing C

cone C = " " " C

$\Rightarrow \overline{\text{cone}}\, C \subseteq \overline{\text{cone}}\, C$ . . . (4)

$\Rightarrow$ $\overline{\text{cone}\, C}$ = closure of C : a closed set, contains C

$\therefore$ (3)+(4) $\Rightarrow \overline{\text{cone}\, C} = \overline{\text{cone}}\, C$.

• Proposition 6.3.

[C: subset of $\mathcal{H}$]

(i) C: cone $\Rightarrow$ (C : convex $\Leftrightarrow C + C \subseteq C$)

(ii) C: convex, $0 \in C \Rightarrow$ (C : cone $\Leftrightarrow C + C \subseteq C$)

• Proposition 6.4.

[C: nonempty convex subset of $\mathcal{H}$]

(i) span C = cone C − cone C = cone C + cone $(-C)$

(ii) $C = -C \Rightarrow$ span C = cone C

• Definition 6.5

[K: convex cone in $\mathcal{H}$]

K: pointed $\overset{\text{def}}{\Leftrightarrow}$ $K \cap (-K) \subseteq \{0\}$

K: solid $\overset{\text{def}}{\Leftrightarrow}$ int $K \neq \emptyset$

Proposition 6.12.

[C: convex, $\subseteq \mathcal{H}$]

(int $C \neq \emptyset \vee C$: closed $\vee$ $\mathcal{H}$: finite-dimensional) $\Rightarrow$ int C = core C

Proof: $\operatorname{int} C \subseteq \operatorname{core} C$ always, so it suffices to show $\operatorname{core} C \subseteq \operatorname{int} C$

$\forall_{x \in \operatorname{core} C}$ $\quad x \in \operatorname{int} C \Leftrightarrow \exists_{\rho \in \mathbb{R}_{++}} B(0;\rho) \subseteq C - x$

$\{\tilde{x} \in C \mid \exists_{\rho \in \mathbb{R}_{++}} B(0;\rho) \subseteq C - x\}$

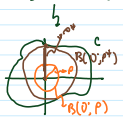
$\forall_{x \in C}: x \in \operatorname{core} C \Leftrightarrow \operatorname{cone}(C-x) = \mathcal{H}$

we can transform the coordinate system so that x becomes origin 0

in the new coordinate system C: convex, $0 \in C$, $\operatorname{cone}(C) = \mathbb{R}_{++} C = \mathcal{H}$

want to prove $\exists_{\rho \in \mathbb{R}_{++}} B(0;\rho) \subseteq C - 0 = C$

without loss of generality we assume $C = -C$ i.e. w.r.t. 0, C is symmetric as a closed ball, as if $B(0;\rho) \subseteq C \Leftrightarrow B(0;\rho) \subseteq \underbrace{B(0;\rho^*)}_{\text{(symmetric)}}$ where ρ^* is the closest point of C to the origin



So, our modified goal is: [$C = -C$, C: convex, $\subsetneq \mathcal{H}$, other antecedents] $0 \in \operatorname{core} C \Rightarrow 0 \in \operatorname{int} C$

(i) $\operatorname{int} C \neq \emptyset$, as $C = -C$, $\operatorname{int} C = \operatorname{int}(-C)$

if $y \in \operatorname{int} C \subseteq C \Rightarrow -y \in \operatorname{int}(-C) = \operatorname{int} C$

now C: convex $\Rightarrow \operatorname{int} C$: convex $\overset{\text{def}}{\Leftrightarrow} \forall_{\tilde{x}, \tilde{y} \in \operatorname{int} C} \forall_{\lambda \in]0,1[} \lambda \tilde{x} + (1-\lambda)\tilde{y} \in \operatorname{int} C$

set, $\tilde{x} = y$, $\tilde{y} := -y$, $\lambda = \frac{1}{2}$

$\Rightarrow \frac{1}{2} y + \frac{1}{2}(-y) = 0 \in \operatorname{int} C$ ☺

(ii)

First note that $\bigcup_{n \in \mathbb{N}} \underbrace{nC}_{C_n:\ \text{closed}} = \mathcal{H}$, /* e.g. take C to be a closed ball, than $2C$ is a closed ball with twice the radius and so on, thus $\bigcup_{n \in \mathbb{N}} nC = \mathcal{H}$ */

Lemma 1.43. (Orlescu) [X: complete metric space]

(i) ($(C_n)_{n \in \mathbb{N}}$: sequence of closed subsets of X) $\Rightarrow \overline{\bigcup_{n \in \mathbb{N}} \operatorname{int} C_n} = \overline{\operatorname{int} \bigcup_{n \in \mathbb{N}} C_n}$

(ii) ($(C_n)_{n \in \mathbb{N}}$: sequence of open subsets of X) $\Rightarrow \operatorname{int} \bigcap_{n \in \mathbb{N}} \overline{C_n} = \operatorname{int} \overline{\bigcap_{n \in \mathbb{N}} C_n}$

$\overline{\bigcup_{n \in \mathbb{N}} \operatorname{int} \underbrace{C_n}_{nC}} = \overline{\operatorname{int} \bigcup_{n \in \mathbb{N}} \underbrace{C_n}_{nC}} = \overline{\operatorname{int} \mathcal{H}} = \overline{\mathcal{H}} = \mathcal{H}$ /* $\overline{X} = \mathcal{H} \Leftrightarrow X = \mathcal{H}$ */

$\Rightarrow \bigcup_{n \in \mathbb{N}} \operatorname{int} nC = \mathcal{H}$

$\Rightarrow \operatorname{int} C \neq \emptyset$ /* else, $\operatorname{int} nC = \emptyset \Rightarrow \bigcup_{n \in \mathbb{N}} \operatorname{int} nC = \emptyset$ */

$\overset{\text{using (i)}}{\Rightarrow} 0 \in \operatorname{int} C$ ☺

(iii) $\mathcal{H}$: finite dimensional

let $(e_i)_{i \in I}$: orthonormal basis of $\mathcal{H}$, now as C: convex, $C = -C$, a scaled version of e_i, /* think about $B(0;\rho^+)$ analogy. take $\underbrace{\|\varepsilon e_i\|}_{\varepsilon \|e_i\| = \varepsilon} \leq \rho^+$ will belong to C */

say $\varepsilon e_i \in C \ \forall i \in I$,

$\because C = -C$

$\Rightarrow -\varepsilon e_i \in C \ \forall_{i \in I}$

$\therefore \{\varepsilon e_i, -\varepsilon e_i\}_{i \in I} \subseteq C \Rightarrow D = \operatorname{conv}\{\varepsilon e_i, -\varepsilon e_i\}_{i \in I} \subseteq C$ // as C is convex itself

But $\underbrace{B(0; \varepsilon/\sqrt{\dim \mathcal{H}})}_{\text{closed}} \subseteq D \subseteq C$ /* now recall that: $A \subseteq B \Rightarrow \operatorname{int} A \subseteq \operatorname{int} B$ fact 6.9.1.

$\Rightarrow 0 \in \underbrace{\operatorname{int}\left(B(0; \varepsilon/\sqrt{\dim \mathcal{H}})\right)}_{\text{open ball}} \subseteq \operatorname{int} C \Rightarrow 0 \in \operatorname{int} C.$ ∎

12/31/2015 10:07 PM

* Corollary 6.19.

[$\mathcal{H}$: finite-dimensional

K: finite dimensional real Hilbert space

$L \in \mathcal{B}(\mathcal{H}, K)$, C, D: nonempty convex subsets of $\mathcal{H}$] $\Rightarrow$

(i) $\operatorname{ri} LC = L(\operatorname{ri} C)$

(ii) $\operatorname{ri}(C - D) = (\operatorname{ri} C) - (\operatorname{ri} D)$

Proof: (i)

/* * Fact 6.14. [C: nonempty convex subset of $\mathcal{H}$]

(i) $\mathcal{H}$: finite dimensional $\Rightarrow$ $\begin{cases} \operatorname{ri} C: \text{interior of } C \text{ relative to } \operatorname{aff} C, \ \operatorname{ri} C \neq \emptyset \\ \overline{\operatorname{ri} C} = \overline{C}, \ \operatorname{ri} \overline{C} = \operatorname{ri} C, \ \operatorname{cl}(\operatorname{ri} C) = \operatorname{ri}(\operatorname{cl} C) = \operatorname{cl} C \end{cases}$ */

$\operatorname{sri} C = \operatorname{ri} C = \operatorname{qri} C \neq \emptyset$

C: nonempty convex subset of $\mathcal{H}$

/* (iv) [K: finite dimensional real Hilbert space, $L \in \mathcal{B}(\mathcal{H}, K)$, $\operatorname{qri} C \neq \emptyset$] $\Rightarrow \operatorname{ri} L(C) = L(\operatorname{qri} C)$ */

$\therefore \operatorname{ri} LC = L \operatorname{qri} C$ ☺

(ii) $L: \mathcal{H} \times \mathcal{H} \to \mathcal{H}: (x,y) \mapsto x - y$ i.e. $L(x,y) = x - y = [1 \ | -1] \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow L = [1 \ | -1]$

$\therefore \operatorname{ri}(C - D) = \operatorname{ri} L(C \times D) = L(\operatorname{ri}(C \times D))$ /* from (i) */ $= L(\operatorname{ri} C \times \operatorname{ri} D) = (\operatorname{ri} C) - (\operatorname{ri} D)$

$\{x - y \mid x \in C, y \in D\} = \{L(x,y) \mid (x,y) \in C \times D\} = L(C \times D)$

(ii) $L: \mathcal{H}\times\mathcal{H} \to \mathcal{H} : (x,y) \to x-y$ i.e., $L(x,y) = x-y = [1|-1]\begin{bmatrix} x \\ y \end{bmatrix} \leftrightarrow L = [1|-1]$

$\therefore$ $ri(C-D) = ri\, L((C\times D)) = L(ri(C\times D))$ /* from (i) */ $= L(ri\,C \times ri\,D) = (ri\,C) - (ri\,D)$

$\{x-y \mid x\in C, y\in D\} = \{L(x,y) \mid (x,y)\in C\times D\} = L(C\times D)$ ($L = [1|-1]$)

$\{(x,y) \mid (x,y)\in ri(C\times ri\,D)\}$ // $L = \{(x-y) \mid x\in ri\,C,\ y\in ri\,D\} = (ri\,C) - (ri\,D)$

/* $ri(C\times D) = ri\,C \times ri\,D$ as the sets C, D are on different spaces connected by their cross product */ ∎

* Proposition 6.17.

[C: convex subset of $\mathcal{H}$

$\operatorname{int} C \neq \emptyset$

$0 \in C$]

(i) $0 \in \operatorname{int} C \Leftrightarrow$

(ii) $\operatorname{cone}(\operatorname{int} C) = \mathcal{H} \Leftrightarrow$

(iii) $\operatorname{cone} C = \mathcal{H} \Leftrightarrow$

(iv) $\overline{\operatorname{cone}}\, C = \mathcal{H} \Leftrightarrow$

Proof: (i) $\Rightarrow$ (ii) $0 \in \operatorname{int} C = \{x\in C \mid \exists_{\rho\in\mathbb{R}_{++}}\ B(0,\rho) \subseteq C - x\}$

now, $\underbrace{\operatorname{int} C \neq \emptyset}_{\text{convex}} \Rightarrow \operatorname{int} C = \underbrace{\operatorname{core} C}_{\{x\in C \mid \operatorname{cone}(C-x) = \mathcal{H}\}} = \operatorname{sri} C = ri\, C = \operatorname{qri} C$ // as for a convex set with nonempty interior, the generalized interiors collapse.

$0 \in \operatorname{core} C \Leftrightarrow \operatorname{cone}(C-0) = \operatorname{cone}(C) = \mathcal{H}$

/* proposition 6.16. [C: convex subset of $\mathcal{H}$, $\operatorname{int} C \neq \emptyset$, $0\in C$] $\Rightarrow$ $\operatorname{int} \operatorname{cone}^{1} C = \operatorname{cone} \operatorname{int} C$ */

$\Rightarrow \operatorname{int} \operatorname{cone} C = \operatorname{cone} \operatorname{int} C = \operatorname{int} \mathcal{H} = \mathcal{H}$

$\therefore \operatorname{cone} \operatorname{int} C = \mathcal{H}$

(ii) $\Rightarrow$ (iii): $\operatorname{int} \operatorname{cone} C = \operatorname{cone} \operatorname{int} C = \mathcal{H} \Rightarrow \operatorname{closure}(\operatorname{int} \operatorname{cone} C) = \operatorname{cone} C = \operatorname{closure} \mathcal{H} = \mathcal{H}$.

(iii) $\Rightarrow$ (iv): $\operatorname{cone} C = \mathcal{H} \Rightarrow \overline{\operatorname{cone}}\, C = \mathcal{H}$ /* now, Proposition 6.2. [$C \subseteq \mathcal{H}$] $\overline{\operatorname{cone}}\, C = \overline{\operatorname{cone}}\, C$ */ $\therefore \overline{\operatorname{cone}}\, C = \mathcal{H}$

(iv) $\Rightarrow$ (ii):

$\overline{\operatorname{cone}}\, C = \overline{\operatorname{cone}}\, C = \mathcal{H}$

convex

$\therefore \operatorname{conv} C = C$

Proposition 6.2.

[C: subset of $\mathcal{H}$] $\Rightarrow$

(i) $\operatorname{cone} C = \mathbb{R}_{++} C$ // most intuitive way creating cone C out of C

(ii) $\overline{\operatorname{cone}}\, C = \overline{\operatorname{cone}}\, C$

(iii) $\operatorname{conv}(\operatorname{cone} C) = \operatorname{cone}(\operatorname{conv} C)$: smallest convex cone containing C

(iv) $\overline{\operatorname{cone}}(\operatorname{cone} C) = \overline{\operatorname{cone}}(\operatorname{cone} C)$: smallest closed convex cone containing C.

now: $\operatorname{cone}(\underbrace{\operatorname{conv} C}_{C}) = \underbrace{\operatorname{conv}(\operatorname{cone} C)}_{\text{convex by structure}}$

$\therefore \operatorname{cone} C$: convex

/* Proposition 6.16 [C: convex subset of $\mathcal{H}$, $\operatorname{int} C \neq \emptyset$, $0\in C$] $\Rightarrow$ $\operatorname{int}(\operatorname{cone} C) = \operatorname{cone}(\operatorname{int} C)$ */

Proposition 3.36.

C: convex subset of $\mathcal{H}$ $\Rightarrow$ (i) $\bar{C}$: convex; (ii) $\operatorname{int} C$: convex; (iii) $\operatorname{int} C \neq \emptyset \Rightarrow \begin{cases} \operatorname{int} C = \operatorname{int} \bar{C} \\ \bar{C} = \overline{\operatorname{int} C} \end{cases}$

now $\mathcal{H} = \operatorname{int} \mathcal{H} = \operatorname{int}(\overline{\operatorname{cone} C}) = \operatorname{int} \operatorname{cone} C$

$= \operatorname{cone} \operatorname{int} C$

$\therefore \operatorname{cone} \operatorname{int} C = \mathcal{H}$

(ii) $\Rightarrow$ (i)

$\operatorname{cone} \operatorname{int} C = \mathcal{H} \Rightarrow 0 \in \mathcal{H} = \operatorname{cone} \operatorname{int} C$ /* $\operatorname{cone} C = \mathbb{R}_{++} C$ */

$= \mathbb{R}_{++} \operatorname{int} C$

$= \bigcup_{\lambda\in\mathbb{R}_{++}} \lambda \operatorname{int} C$

$\Rightarrow \exists_{\lambda\in\mathbb{R}_{++}}\ 0 \in \lambda \operatorname{int} C$

$\Rightarrow 0 \in \operatorname{int} C$. ∎

* Proposition 6.20.

[m: integer, ≥ 2,

$I = \{1, \ldots, m\}$

$(C_i)_{i\in I}$: convex subsets of $\mathcal{H}$

one of the following holds:

(i) $\forall_{i\in\{2,\ldots,m\}}\ (C_i - \bigcap_{j=1}^{i-1} C_j)$: closed linear subspace

(ii) $(C_i)_{i\in I}$: linear subspaces, $\forall_{i\in\{2,\ldots,m\}}\ (C_i + \bigcap_{j=1}^{i-1} C_j)$: closed

(iii) $C_m \cap (\bigcap_{i=1}^{m} \operatorname{int} C_i) \neq \emptyset$

(iv) $\mathcal{H}$: finite-dimensional, $\bigcap_{i\in I} ri\, C_i \neq \emptyset$] $\Rightarrow$

$0 \in \bigcap_{i=2}^{m} \operatorname{sri}(C_i - \bigcap_{j=1}^{i-1} C_j)$

1/2/2016 9:27 PM

Proof: Apply proposition 6.19.

* Proposition 6.19. [$\mathcal{X}$: convex subset of $\mathcal{H}$, $\mathcal{H}$: real Hilbert space, $\mathcal{X} \in \mathcal{B}(\mathcal{H}, \mathcal{H})$, $\mathcal{X}$: convex subset of K]

Suppose one of the following holds:

(i) $\forall_{i\in\{2,\ldots,m\}}\ (C_i - \bigcap_{j=1}^{i-1} C_j)$: closed linear subspace ✓ $C_i + \mathcal{X}(\bigcap_{j=1}^{i-1} C_j) = C_i + \bigcap_{j=1}^{i-1} C_j$: closed

(ii) C_i, $\bigcap_{j=1}^{i-1} C_j$: linear subspaces and (• $D + L(C)$: closed ✓

• D: closed, $L(C)$: finite dimensional or finite co-dimensional ✓

• D: finite-dimensional or finite-codimensional and $L(C)$: closed)

(iii) D: cone, $D - L(C)$: closed linear subspace

(iv) $D = L(C)$, span D: closed

(v) $0 \in \operatorname{core}(D - L(C))$

(vi) $0 \in \operatorname{int}(D - L(C))$ $\therefore C_m \cap \operatorname{int} \mathcal{X}(\bigcap_{i=1}^{m-1} C_i)$

(vii) $C_m \cap \operatorname{int} L(C) \neq \emptyset$ ✓

(iv) $D = L(C)$, span D: closed

(v) $0 \in \operatorname{core}(D - L(C))$

(vi) $0 \in \operatorname{int}(D - L(C))$

(vii) $\cancel{\operatorname{ri} \operatorname{int} L(C) \neq \emptyset}$ ✓

$L(C) \cap \operatorname{int} D \neq \emptyset$

(viii) X: finite-dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

(ix) X: finite-dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

(x) (X, K): finite dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

Then $0 \in \operatorname{sri}(D - L(C))$.

$\bigcap_{i=1}^{m-1} C_i$

$\therefore C_m \cap \operatorname{int} \operatorname{Id}\left(\bigcap_{i=1}^{m-1} C_i\right)$

$= C_m \cap \operatorname{int}\left(\bigcap_{i=1}^{m-1} C_i\right) = C_m \cap \left(\bigcap_{i=1}^{m-1} \operatorname{int} C_i\right) \neq \emptyset$

$\operatorname{ri} C_1 \cap \left(\operatorname{ri} \operatorname{Id}\left(\bigcap_{j \in \{1\}^c} C_j\right)\right)$

$= \underbrace{\operatorname{ri} C_1}_{\text{convex}} \cap \left(\operatorname{ri} \underbrace{\bigcap_{j \in \{1\}^c} C_j}_{\text{convex}}\right)$

$= \bigcap_{j \in I} \operatorname{ri} C_j \neq \emptyset$

Fact 6.14

/+ now by (v) [X: finite dimensional, D: convex subset of X, $\operatorname{ri} \overset{\text{convex}}{C} \cap \operatorname{ri} D \neq \emptyset$] $\Rightarrow \operatorname{ri}(C \cap D) = \operatorname{ri} C \cap \operatorname{ri} D$

So, one of (i), (ii), (vii), (viii) are given

so, $0 \in \operatorname{sri}(\underbrace{D - L(C)}_{C_i}) = \operatorname{sri}\left(C_i - \bigcap_{j=1}^{i-1} C_j\right)$ ■

$\operatorname{Id}\left(\bigcap_{j=1}^{i-1} C_j\right) = \bigcap_{j=1}^{i-1} C_j$

6.3 Polar and Dual Cone, 6.4 Tangent and Normal Cone

Sunday, January 3, 2016 9:31 PM

Proposition 6.26.

[K_1, K_2: nonempty cones in $\mathcal{H}$] $\Rightarrow$

$(K_1+K_2)^\ominus = K_1^\ominus \cap K_2^\ominus$

in particular: [K_1, K_2: linear subspaces] $\Rightarrow (K_1+K_2)^\perp = K_1^\perp \cap K_2^\perp$

Proof: $x_1 \in K_1$, $x_2 \in K_2$

want to prove, $(K_1+K_2)^\ominus \subseteq K_1^\ominus \cap K_2^\ominus$

$\forall_u \; u \in (K_1+K_2)^\ominus = \{\tilde{x}+\tilde{y} : \tilde{x} \in K_1, \tilde{y} \in K_2\}^\ominus$

$= \{\tilde{u} \in \mathcal{H} \mid \sup\langle K_1+K_2 \mid \tilde{u}\rangle \leq 0\}$

$= \{\tilde{u} \in \mathcal{H} \mid \forall_{\tilde{x}+\tilde{y} \in K_1+K_2} \langle \tilde{x}+\tilde{y} \mid \tilde{u}\rangle \leq 0\}$ $\therefore u \in (K_1+K_2)^\ominus \Leftrightarrow \forall_{\tilde{x}+\tilde{y} \in K_1+K_2} \langle \tilde{x}+\tilde{y} \mid u\rangle \leq 0$ ... (P2.0)

now K_1, K_2: cones $\Leftrightarrow K_1 = \mathbb{R}_{++}K_1$, $K_2 = \mathbb{R}_{++}K_2 \Rightarrow K_1+K_2 = \mathbb{R}_{++}(K_1+K_2)$

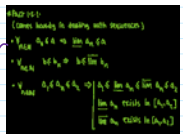
$\therefore \forall_{\lambda_1 \in \mathbb{R}_{++}, \lambda_2 \in \mathbb{R}_{++}} \; \lambda_1 x_1 \in K_1, \lambda_2 x_2 \in K_2 \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in \lambda_1 K_1 + \lambda_2 K_2 = K_1+K_2$

$\forall_{\lambda_1 \in \mathbb{R}_{++}, \lambda_2 \in \mathbb{R}_{++}} \langle \lambda_1 x_1 + \lambda_2 x_2 \mid u\rangle \leq 0$

set $\lambda_2 = 1$, $\lambda_1 := (\lambda_{1,n})_{n\in\mathbb{N}} : \subseteq \mathbb{R}_{++}, \lambda_{1,n} \to 0$

$\Rightarrow \langle \lambda_{1,n} x_1 + x_2 \mid u\rangle \leq 0$

$\therefore \forall_{x_2 \in K_2} \langle x_2 \mid u\rangle \leq 0$



now recall:

$u \in K_2^\ominus = \{\tilde{u} \in \mathcal{H} \mid \langle K_2 \mid u\rangle \leq 0\} \Leftrightarrow \forall_{x_2 \in K_2} \langle x_2 \mid u\rangle \leq 0$

$\therefore u \in K_2^\ominus$, similarly we can show: $u \in K_1^\ominus$

$\Leftrightarrow u \in K_1^\ominus + K_2^\ominus$

$\therefore (K_1+K_2)^\ominus \subseteq K_1^\ominus + K_2^\ominus$

Now, let us show

$K_1^\ominus \cap K_2^\ominus \subseteq (K_1+K_2)^\ominus$

$\Leftrightarrow \forall_{u \in K_1^\ominus \cap K_2^\ominus} \; u \in (K_1+K_2)^\ominus$

$u \in K_1^\ominus \cap K_2^\ominus \Leftrightarrow u \in K_1^\ominus, u \in K_2^\ominus$

$\Leftrightarrow \sup\langle K_1 \mid u\rangle \leq 0, \sup\langle K_2 \mid u\rangle \leq 0$

$\Leftrightarrow \forall_{x_1 \in K_1} \langle x_1 \mid u\rangle \leq 0, \forall_{x_2 \in K_2} \langle x_2 \mid u\rangle \leq 0$

adding

$\Rightarrow \forall_{x_1 \in K_1, x_2 \in K_2} \langle x_1 \mid u\rangle + \langle x_2 \mid u\rangle = \langle x_1+x_2 \mid u\rangle \leq 0 \Leftrightarrow u \in (K_1+K_2)^\ominus$ $\therefore K_1^\ominus \cap K_2^\ominus \subseteq (K_1+K_2)^\ominus$

$\therefore \boxed{(K_1+K_2)^\ominus = K_1^\ominus \cap K_2^\ominus}$

if K_1, K_2: linear subspaces $\Rightarrow K_1^\ominus = K_1^\perp$, $K_2^\ominus = K_2^\perp$ /* using **Proposition 6.22. [C: linear subspace of $\mathcal{H}$] $\Rightarrow C^\ominus = C^\perp$** */

$\Rightarrow K_1+K_2$: linear subspace $\Rightarrow (K_1+K_2)^\ominus = (K_1+K_2)^\perp$

$(K_1+K_2)^\perp = K_1^\perp \cap K_2^\perp$ ∎

1/7/2016 3:19 PM

* Proposition 6.27.

[K: nonempty closed convex cone in $\mathcal{H}$, $x, p \in \mathcal{H}$]

$p = P_K(x) \Leftrightarrow (p \in K, x-p \perp p, x-p \in K^\ominus)$

Proof: ($\Rightarrow$)

/* Recall Theorem 3.14

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

$C := K$, then $p = P_K(x) \Leftrightarrow \begin{cases} p \in K \\ \forall_{y \in K} \langle y-p \mid x-p\rangle \leq 0 \end{cases}$

$p \in K \Rightarrow \forall_{\alpha \in \mathbb{R}_{++}} \; \alpha p \in K$

set $y := \alpha p$

$\langle \alpha p - p \mid x-p\rangle = \langle p(\alpha-1) \mid x-p\rangle = (\alpha-1)\langle p \mid x-p\rangle \leq 0 \quad \forall \alpha \in \mathbb{R}_{++}$

as $(\alpha-1)$ can have either sign the only possibility is $\langle p \mid x-p\rangle = 0 \Leftrightarrow p \perp x-p$

(based on what $\alpha \in \mathbb{R}_{++}$ we pick)

$\forall_{y \in K} \langle y \mid x-p\rangle = \langle \underbrace{y-p}+\underbrace{p} \mid x-p\rangle = \underbrace{\langle y-p \mid x-p\rangle}_{\leq 0} + \underbrace{\langle p \mid x-p\rangle}_{=0}$ /* By definition, $\Box \in C^\ominus \Leftrightarrow \forall_{\Box \in C} \langle \Box \mid \Box\rangle \leq 0$

$= \langle y-p \mid x-p\rangle \leq 0$

$\Leftrightarrow \forall_{y \in K} \langle y \mid x-p\rangle \leq 0$

/*

now $K^\ominus = \{u \in \mathcal{H} : \forall_{y \in K} \langle y \mid u\rangle \leq 0\}$, so $u \in K^\ominus \Leftrightarrow \forall_{y \in K} \langle y \mid x-p\rangle \leq 0$ */

so, $(x-p) \in K^\ominus$

($\Leftarrow$:)

given, $p \in K$, $\underbrace{x-p \perp p}$ and $x-p \in K^{\ominus}$

$\Updownarrow$ $\qquad\qquad$ $\Updownarrow$

$\langle x-p|p\rangle = 0 \qquad \forall_{y\in K}\ \langle x-p|y\rangle \leq 0$

$\Downarrow$

$\forall_{y\in K}\ -\langle x-p|p\rangle + \langle x-p|y\rangle \leq 0$

$\leftrightarrow \forall_{y\in K}\ \langle x-p|y-p\rangle \leq 0$

$\therefore\ p\in K \wedge \forall_{y\in K}\ \langle x-p|y-p\rangle \leq 0.$

$\Leftrightarrow \quad p = P_K(x)$ ■

* Theorem 6.29. (Moreau)

1/7/2016 5:35 PM

[[K: nonempty convex cone in $\mathcal{H}$,
$x\in\mathcal{H}$]]

(i) $x = P_K x + P_{K^\ominus} x$

(ii) $P_K x \perp P_{K^\ominus} x$

(iii) $\|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x)$

/* Fact used : $K \subseteq K^{\ominus\ominus}$ */

Proof: (i)

$q = x - P_K x$ (A)

/*

* Proposition 6.27 [K: nonempty closed convex cone in $\mathcal{H}$; $x,p\in\mathcal{H}$] $\Leftrightarrow$ $p = P_K x \Leftrightarrow (p\in K,\ (x-p)\perp p,\ x-p\in K^\ominus)$

Characterization of projection on a closed convex cone.

/* projection on a closed convex cone is characterized by its polar cone */

*/

$\Downarrow$

[K: nonempty closed convex cone in $\mathcal{H}$; $x\in\mathcal{H}$]

$\underbrace{P_K x}_{x-q} \in K,\ \underbrace{(x-P_K x)}_{q} \perp \underbrace{P_K x}_{x-q},\ \underbrace{x-P_K x}_{q} \in K^\ominus$

$\leftrightarrow\ x-q\in K,\ q\perp x-q,\ q\in K^\ominus$

now, $K^\ominus$ is a cone too so $P_{K^\ominus}x$ will satisfy : $P_{K^\ominus}x\in K^\ominus$, $x-P_{K^\ominus}x \perp P_{K^\ominus}x$, $x-P_{K^\ominus}x \in \underbrace{K^{\ominus\ominus}}_{\{u\in\mathcal{H}|\ \forall_{\tilde{x}\in K^\ominus}\ \langle x|u\rangle\leq 0\}}$; so after comparison:

So, we have $q = P_{K^\ominus}x$, thus from A : $x = P_K x + P_{K^\ominus}x$

(ii) in $q = P_{K^\ominus}x$

$P_{K^\ominus}x \perp P_K x$

(iii) $\|x\|^2 = \|P_K x + P_{K^\ominus}x\|^2$ /* from (i) $x = P_K x + P_{K^\ominus}x$ */

$= \|P_{K^\ominus}x\|^2 + \|P_K x\|^2 + 2\underbrace{\langle P_K x|P_{K^\ominus}x\rangle}_{0}$ /* from (ii) $P_K x \perp P_{K^\ominus}x$ */

$= \underbrace{\|x-P_K x\|^2}_{d_K(x)} + \underbrace{\|x-P_{K^\ominus}x\|^2}_{d_{K^\ominus}x}$ /* using (ii) */

$= d_K^2(x) + d_{K^\ominus}^2(x)$ ■

Proposition 6.31.

[K: nonempty closed convex cone in $\mathcal{H}$;
$x\in\mathcal{H}$, $\langle x|x-P_K x\rangle \leq 0$] $\Rightarrow$ $x\in K$

$0 \leq$

$\|x-P_K x\|^2 = \langle x-P_K x|x-P_K x\rangle$

$= \underbrace{\langle x|x-P_K x\rangle}_{\leq 0 \text{ // given}} - \underbrace{\langle P_K x|x-P_K x\rangle}_{0}$ /* now.

* Proposition 6.27 [K: nonempty closed convex cone in $\mathcal{H}$; $x,p\in\mathcal{H}$] $\Leftrightarrow$ $p = P_K x \Leftrightarrow (p\in K,\ (x-p)\perp p,\ x-p\in K^\ominus)$

Characterization of projection on a closed convex cone.

/* projection on a closed convex cone is characterized by its polar cone */

*/

≤ 0

$\Rightarrow\ \|x-P_K x\|^2 = 0 \leftrightarrow x = P_K x \leftrightarrow x\in K.$ ■

1/8/2016 10:57 PM

* Proposition 6.32.

[[C: nonempty convex subset of $\mathcal{H}$] $C^{\ominus\ominus} = \overline{\text{cone}}\, C$

Proof : $K := \overline{\text{cone}}\, C$

Part 1: $K \subseteq C^{\ominus\ominus}$

/* as $\overline{\text{cone}}\, C$ smallest closed cone containing C, so set inclusion will not change */

/* fact : C: convex $\rightarrow$ $C \subseteq C^{\ominus\ominus}$ */ $\Rightarrow$ $\overline{\text{cone}}\, C \subseteq \overline{\text{cone}}\, C^{\ominus\ominus}$

* Proposition 6.23 [C: subset of $\mathcal{H}$] $\Rightarrow$ (i) $D\subseteq C \Rightarrow$ • $C^\ominus \subseteq D^\ominus$ • $C^{\ominus\ominus} \subseteq D^{\ominus\ominus}$ /* inclusion flips */

(ii) $C^\ominus, C^{\ominus\ominus}$: nonempty closed convex cones $\rightarrow$ $C^{\ominus\ominus}$: nonempty closed convex cone $\Rightarrow$ $\overline{\text{cone}}\, C^{\ominus\ominus} = C^{\ominus\ominus}$

(iii) $C^\ominus = (\text{cone}\, C)^\ominus = (\text{conv}\, C)^\ominus = \overline{C}^\ominus$ /* to the polar cone operator, a set, its cone, convex hull and closure all are same */

(iv) $\overline{\text{cone}}\, C = \overline{\text{cone}}\, C \Rightarrow C^{\ominus\ominus} = C^{\ominus\ominus\perp}$

$\overbrace{\boxed{\overline{\text{cone}}\, C \subseteq C^{\ominus\ominus}}}^{K}$

(i) $C^\ominus, C^\oplus$: nonempty closed convex cones $\Rightarrow$ $C^{\ominus\oplus}$: nonempty closed convex cone $\Rightarrow$ $\overline{\text{cone}}\, C^{\ominus\oplus} = C^{\ominus\oplus}$

(ii) $C^\ominus = (\text{cone}\, C)^\ominus = (\text{conv}\, C)^\ominus = \tilde{C}^\ominus$ /* to the polar cone operator, a set, its cone, convex hull and closure all are same */

(iii) $\overline{\text{cone}}\, C = \overline{\text{cone}}\, C \Rightarrow C^\ominus = C^\oplus = C^\perp$

Part 2: $C^{\ominus\ominus} \subseteq K$.

Take, $x \in C^{\ominus\ominus}$ /* now, $K = \overline{\text{cone}}\, C = \overline{\text{cone}}\, C$ /* Proposition 6.2: $\overline{\text{cone}}\, C = \overline{\text{cone}}\, C$ */

$K^\ominus = (\overline{\text{cone}}\, C)^\ominus = (\text{cone}\, C)^\ominus = C^\ominus$

$\Rightarrow K^{\ominus\ominus} = C^{\ominus\ominus}$ */

$\therefore x \in C^{\ominus\ominus} = K^{\ominus\ominus}$

* Proposition 6.27 [K: nonempty closed convex cone in $\mathcal{H}$; $x, p \in \mathcal{H}$] $\Rightarrow$ $p = P_K x \Leftrightarrow (p \in K,\ (x-p) \perp p,\ \underbrace{x-p}_{} \in K^\ominus)$

Characterization of projection on a closed convex cone.

/* projection on a closed convex cone is characterized by its polar cone */

$\langle x | x - P_K x \rangle \leq 0$ /* As by defn: $y \in C^\ominus \Leftrightarrow \forall_{x \in C} \langle x | y \rangle \leq 0$

$\Rightarrow \boxed{x \in K}$

$\because x \in (K^\ominus)^\ominus \Leftrightarrow \forall_{\tilde{y} \in K^\ominus} \langle \tilde{y} | x \rangle \leq 0$

$\tilde{y} := x - P_K x \in K^\ominus \Rightarrow \langle x | x - P_K x \rangle \leq 0$ */

/* As * Proposition 6.31. [K: nonempty closed convex cone in $\mathcal{H}$

$x \in \mathcal{H}$, $\langle x | x - P_K x \rangle \leq 0$] $\Rightarrow x \in K$ */

$\therefore \forall_{x \in C^{\ominus\ominus}}\ x \in K = \overline{\text{cone}}\, C \Leftrightarrow \boxed{C^{\ominus\ominus} \subseteq \overline{\text{cone}}\, C}$

$C^{\ominus\ominus} = \overline{\text{cone}}\, C$. ∎

Proposition 6.43. /* Relationship between tangent and normal cone */

[C: nonempty convex, $\subseteq \mathcal{H}$; $x \in C$]

(i) $T_C^\ominus(x) = N_C x$, $N_C^\ominus(x) = T_C(x)$

(ii) $x \in \text{core}\, C \Rightarrow T_C x = \mathcal{H} \Leftrightarrow N_C x = \{0\}$

Proof:

$$T_C \tilde{x} = \begin{cases} \overline{\text{cone}}\,(C - \tilde{x}), & \text{if } \tilde{x} \in C \\ \emptyset, & \text{else} \end{cases}$$

given, $x \in C$

$\Rightarrow T_C(x) = \underbrace{\overline{\text{cone}}\,(C-x)}_{\text{smallest closed cone containing } (-x)} \supseteq C - x$

$\Rightarrow T_C(x)^\ominus \subseteq (C-x)^\ominus$

/* now

* Proposition 6.25 [C: subset of $\mathcal{H}$] $\Rightarrow$ (i) $D \subseteq C \Rightarrow$ $C^\ominus \subseteq D^\ominus$ /* Inclusion lifts */, $C^\oplus \subseteq D^\oplus$

(ii) $C^\ominus, C^\oplus$: nonempty closed convex cones

(iii) $C^\ominus = (\text{cone}\, C)^\ominus = (\text{conv}\, C)^\ominus = \tilde{C}^\ominus$ /* to the polar cone operator, a set, its cone, convex hull and closure all are same */

(iv) $C^\ominus \cap C^\oplus = C^\perp$ /* intersection of polar and dual cone is orthogonal complement */

(v) $\overline{\text{cone}}\, C = -\overline{\text{cone}}\, C \Rightarrow C^\perp = C^\oplus$ /* sufficient condition for the coincide of polar, dual cone and orthogonal complement */

*/

now,

$$N_C \tilde{x} = \begin{cases} (C - \tilde{x})^\ominus, & \text{if } \tilde{x} \in C \\ \emptyset, & \text{else} \end{cases}$$

as $x \in C \Rightarrow (C-x)^\ominus = N_C x$ $\qquad T_C^\ominus x \subseteq N_C x$ (eq1)

now let's show: $N_C x \subseteq T_C^\ominus x \Leftrightarrow \forall_{u \in N_C x}\ u \in T_C^\ominus x = \{\tilde{u} \in \mathcal{H} \mid \sup \langle T_C x | \tilde{u} \rangle \leq 0\} = \{\tilde{u} \in \mathcal{H} \mid \langle \overline{\text{cone}}(C-x) | \tilde{u} \rangle \leq 0\}$

take $u \in N_C x$ // $x \in C$

$= (C-x)^\ominus = \{\tilde{u} \in \mathcal{H} \mid \sup \langle C - x | \tilde{u} \rangle \leq 0\}$

$\Leftrightarrow \sup \langle C - x | u \rangle \leq 0$

now $\overline{\text{cone}}\,(C-x) = \overline{\text{cone}}(C-x)$ /* Proposition 6.2 */

(iii) $\tilde{C}^\ominus = (\text{cone}\, \tilde{C})^\ominus = (\text{conv}\, \tilde{C})^\ominus = \tilde{\overline{C}}^\ominus$ /* to the polar cone operator, a set, its cone, convex hull and closure all are same */

$\tilde{C} := \text{cone}(C-x) \Rightarrow (\text{cone}(C-x))^\ominus = (\overline{\text{cone}\,(C-x)})^\ominus = (C-x)^\ominus$

now $x \in C$, $T_C x = \overline{\text{cone}}\,(C-x)$

$\Rightarrow T_C x^\ominus = (\text{cone}(C-x))^\ominus = (C-x)^\ominus = \{\tilde{u} \in \mathcal{H} \mid \sup \langle C - x | \tilde{u} \rangle \leq 0\}$

$u \in T_C x^\ominus$

$\therefore N_C x \subseteq T_C x^\ominus$ (eq2)

Combining both $\boxed{N_C x = T_C x^\ominus}$ // From (eq1), (eq2)

now $x \in C \Rightarrow T_C x = \overline{\text{cone}}\,(C-x)$: closed convex cone /* using:

* Corollary 6.33. [K: nonempty closed convex cone in $\mathcal{H}$] $\Rightarrow K^{\ominus\ominus} = K$

$\underbrace{T_C x^{\ominus\ominus}}_{\underbrace{((T_C x)^\ominus)^\ominus}_{N_C x \text{ // just proved}}} = T_C x$ $\qquad \therefore T_C x = N_C x^\ominus$

Proposition 6.2.

[C: subset of $\mathcal{H}$] $\Rightarrow$

(i) $\text{cone}\, C = K_{++} C$ // most intuitive way creating cone C out of C

(ii) $\overline{\text{cone}}\, C = \overline{\text{cone}}\, C$

(iii) * $\text{cone}(\text{conv}\, C) = \text{conv}(\text{cone}\, C)$: smallest convex cone containing C

as $C - x$: convex $\Rightarrow (-x) = \text{conv}(C-x)$

(iv) $\overline{\text{cone}}(\text{conv}\, C) = \overline{\text{cone}}(\text{cone}\, C)$: smallest closed convex cone containing C. */

(ii) $x \in \text{core}\, C = \{\tilde{x} \in \mathcal{H} \mid \text{cone}(C - \tilde{x}) = \mathcal{H}\}$

$\Leftrightarrow \text{cone}(C-x) = \mathcal{H} \Leftrightarrow \overline{\text{cone}}(C-x) = \overline{\mathcal{H}} = \mathcal{H} \Leftrightarrow \overline{\text{cone}}\,(C-x) = \mathcal{H}$

as $x \in C \Rightarrow T_C x = \overline{\text{cone}}\,(C-x) = \mathcal{H}$

from (i): $N_C x = (T_C x)^\ominus = \mathcal{H}^\ominus = \{\tilde{u} \in \mathcal{H} \mid \sup \langle \mathcal{H} | \tilde{u} \rangle \leq 0\} = \{0\}$

// this is the only value

$\therefore T_C x = \mathcal{H} \Rightarrow N_C(x) = \{0\}$

as $x \in C \Rightarrow T_C x = \text{cone}(C-x) = \mathcal{H}$

from (i): $N_C x = (T_C x)^\ominus = \mathcal{H}^\ominus = \{\tilde{u} \in \mathcal{H} \mid \sup \langle \mathcal{H} | \tilde{u} \rangle \leq 0\} = \underbrace{\{0\}}$

// this is the only value

$\therefore T_C x = \mathcal{H} \Rightarrow N_C(x) = \{0\}$

now take $N_C x = \{0\} \overset{\text{from (i): } N_C^\ominus x = T_C x}{\Rightarrow}$

$T_C x = N_C^\ominus x = \{0\}^\ominus = \{\tilde{u} \in \mathcal{H} \mid \sup \langle 0 | \tilde{u} \rangle = 0 \leq 0\} = \mathcal{H}$

$\therefore N_C x = \{0\} \Rightarrow T_C x = \{0\}^\ominus = \mathcal{H}$

$\therefore x \in \text{core}\, C \Rightarrow N_C x = \{0\} \Leftrightarrow T_C x = \mathcal{H}$.

■

1/9/2016 11:23 PM

*Proposition 6.46.

[C: nonempty closed convex subset of $\mathcal{H}$

$x, p \in \mathcal{H}$]

$p = P_C x \Leftrightarrow x - p \in N_C p$

Proof: from projection theorem: $\left(\underbrace{P_C}_{\text{closed convex set}} x \in C,\ \forall_{y \in C} \langle y - P_C x | x - P_C x \rangle \leq 0 \right)$ ... (eq:1)

from definition of normal cone to C at x: $N_C x = \begin{cases} \{u \in \mathcal{H} \mid \forall_{y \in C} \langle y - x | u \rangle \leq 0\}, & x \in C \\ \emptyset, & \text{otherwise} \end{cases}$ $\quad P_C x = p$

$(C-x)^\ominus$

say $p = P_C x$. then

$p \in C$, so $N_C p = \{u \in \mathcal{H} \mid \forall_{y \in C} \langle y - p | u \rangle \leq 0\}$

so. $\square \in N_C p \Leftrightarrow \forall_{y \in C} \langle y - p | \square \rangle \leq 0$

set, $\square = x - p \hookrightarrow \therefore x - p \in N_C p \Leftrightarrow \forall_{y \in C} \langle y - p | \underbrace{x - p}_{p \in C} \rangle \leq 0$, now comparing with (eq:1)

$\forall_{y \in C} \langle y - p | x - p \rangle \leq 0 \Leftrightarrow x - p \in N_C p$ ■

6.5 Recession and Barrier Cones

Sunday, January 10, 2016 11:15 AM

* Proposition 6.48.

[C: nonempty convex subset of $\mathcal{H}$]

(i) $\operatorname{rec} C$: convex cone, $0 \in \operatorname{rec} C$

(ii) $\operatorname{bar} C$: convex cone, $C^{\ominus} \subseteq \operatorname{bar} C$

(iii) C: bounded $\Rightarrow$ $\operatorname{bar} C = \mathcal{H}$

(iv) C: cone $\Rightarrow$ $\operatorname{bar} C = C^{\ominus}$

(v) C: closed $\Rightarrow$ $(\operatorname{bar} C)^{\ominus} = \operatorname{rec} C$

Proof: (not complete).

(i) $\operatorname{rec} C = \{ x \in \mathcal{H} \mid x + C \subseteq C \}$

so, $\boxed{\cdot} \in \operatorname{rec} C \Leftrightarrow \boxed{\cdot} \in \mathcal{H},\ \boxed{\cdot} + C \subseteq C$

as, $0 \in \mathcal{H},\ 0 + C = C \subseteq C \Leftrightarrow 0 \in \operatorname{rec} C$.

first, $\forall_{x_1 \in \operatorname{rec} C} \ \forall_{x_2 \in \operatorname{rec} C} \quad x_1 + x_2 = z_1 + z_2 - y_1 - y_2$

$\forall_{y_1 \in C} \ \exists_{z_1 \in C} \ x_1 + y_1 = z_1 \qquad \forall_{y_2 \in C} \ \forall_{z_2 \in C} \ x + y_2 = z_2$

By definition of a convex set, $\forall_{\alpha \in]0,1[} \ \alpha C + (1-\alpha) C = C$

$\forall_{x,y \in C} \ \forall_{\alpha \in]0,1[} \ \alpha x + (1-\alpha) y \in C$

suppose $x \in \operatorname{rec} C \Leftrightarrow \forall_{z \in C} \ x + z \in C \leftrightarrow \forall_{z \in C} \ \exists_{\eta_1 \in C} \ x + z = \eta_1 \rightarrow \alpha x + \alpha z = \alpha \eta_1$ ($\alpha \in]0,1[$)

$y \neq x$, $y \in \operatorname{rec} C \Leftrightarrow \forall_{w \in C} \ y + w \in C \leftrightarrow \forall_{w \in C} \ \exists_{\eta_2 \in C} \ y + w = \eta_2 \rightarrow (1-\alpha) y + (1-\alpha) w = (1-\alpha) \eta_2$

$$\alpha x + (1-\alpha) y + \underbrace{\alpha z + (1-\alpha) w}_{\bar{z} \in C} = \underbrace{\alpha \eta_1 + (1-\alpha) \eta_2}_{\bar{\eta} \in C}$$

[as $z, w \in \underbrace{C}_{\text{convex}}$] [as $\eta_1, \eta_2 \in C$]

$\leftrightarrow \alpha x + (1-\alpha) y + \bar{z} = \bar{\eta} \in C$

now $\bar{z}$: arbitrary as this is a convex combination of arbitrary points $z, w \in C$

So we have: $\forall_{x \in \operatorname{rec} C} \ \forall_{y \in \operatorname{rec} C} \quad \alpha x + (1-\alpha) y + C \subseteq C \leftrightarrow \alpha x + (1-\alpha) y \in \operatorname{rec} C$

$\therefore \operatorname{rec} C$: convex

/* Proposition 6.3. (ii)

[[C: convex, $0 \in C$]] $C + C \subseteq C \Leftrightarrow C$: cone */

We have shown, $\mathrm{rec}\, C$: convex, $\ni 0$, $\mathrm{rec}\, C + \mathrm{rec}\, C \subseteq \mathrm{rec}\, C$

7.1 Support Points

Sunday, January 10, 2016 10:31 PM

Proposition 7.2 [C, D: nonempty subsets of $\mathcal{H}$; $C \subseteq D$] $C \cap \text{spts}\, D \subseteq \text{spts}\, C = C \cap \text{spts}\, \bar{C}$

Proof:

$x \in C \cap \text{spts}\, D \Leftrightarrow x \in C \wedge x \in \text{spts}\, D = \{\tilde{x} \in D \mid \exists_{u_{\tilde{x}} \in \mathcal{H}\setminus\{0\}}\ \langle \tilde{x} | u_{\tilde{x}} \rangle \geq \sup \langle D | u_{\tilde{x}} \rangle\}$ (normal vector)

$\Updownarrow$

$\exists_{u \in \mathcal{H}\setminus\{0\}}\ \langle x|u \rangle \geq \sup \langle D|u \rangle$ /* $\sup \langle D|u \rangle = \sup_{x \in D} \langle x|u \rangle \geq \sup_{x \in C} \langle x|u \rangle = \sup \langle C|u \rangle$ */

$\geq \sup \langle C|u \rangle$ /* $D \supseteq C$ */

$\Rightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ \langle x|u \rangle \geq \sup \langle C|u \rangle$ /* now: $\text{spts}\, C = \{y \in C \mid \exists_{v_y \in \mathcal{H}\setminus\{0\}}\ \langle y | v_y \rangle \geq \sup \langle C|u \rangle\}$ */

$\Leftrightarrow x \in \text{spts}\, C$

$\therefore C \cap \text{spts}\, D \subseteq \text{spts}\, C$

/* required info: $\sup_{x \in C} f(x) = \sup_{x \in C} f(x)$ */

now let's show: $\text{spts}\, C = C \cap \text{spts}\, \bar{C}$

$\forall_x\ x \in \text{spts}\, C = \{\tilde{x} \in C \mid \exists_{u_{\tilde{x}} \in \mathcal{H}\setminus\{0\}}\ \langle \tilde{x} | u_{\tilde{x}} \rangle \geq \sup \langle C | u_{\tilde{x}} \rangle\}$

$\Leftrightarrow x \in C,\ \exists_{u \in \mathcal{H}\setminus\{0\}}\ \langle x|u \rangle \geq \sup \langle C|u \rangle = \sup \langle \bar{C}|u \rangle$

$\Leftrightarrow x \in C,\ \exists_{u \in \mathcal{H}\setminus\{0\}}\ \langle x|u \rangle \geq \sup \langle \bar{C}|u \rangle$

$\Leftrightarrow x \in C \cap \text{spts}\, \bar{C}$ /* as $\text{spts}\, \bar{C} = \{\tilde{x} \in \bar{C} \mid \exists_{u_{\tilde{x}} \in \mathcal{H}\setminus\{0\}}\ \langle \tilde{x} | u_{\tilde{x}} \rangle \geq \sup \langle \bar{C} | u_{\tilde{x}} \rangle\}$

$x \in C \Rightarrow x \in \bar{C}$

$\therefore \text{spts}\, C = C \cap \text{spts}\, \bar{C}$ ∎

Proposition 7.3.

[C: nonempty convex, $\subseteq \mathcal{H}$] $\text{spts}\, C = \{x \in \mathcal{H} \mid N_C x \setminus \{0\}\} = N_C^{-1}(\mathcal{H}\setminus\{0\})$

Proof: $\forall_x$ /* $\text{spts}\, C = \{y \in C \mid \exists_{u_y \in \mathcal{H}\setminus\{0\}}\ \langle y | u_y \rangle \geq \sup \langle C | u_y \rangle\}$

$x \in \text{spts}\, C \Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}} \left(\langle x|u \rangle \geq \sup \langle C|u \rangle \geq \langle z|u \rangle\ \forall_{z \in C} \right)$

$\Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ \underbrace{\forall_{z \in C}\ \langle z|u \rangle \leq \langle x|u \rangle}_{\langle z-x|u \rangle \leq 0,\ \sup \langle C-x|u \rangle \leq 0} \Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ \sup \langle C-x|u \rangle \leq 0$ /* recall, [C: nonempty convex, $\subseteq \mathcal{H}$; $x \in \mathcal{H}$] $N_C x = \begin{cases} (C-x)^{\ominus} = \{u \in \mathcal{H} \mid \sup \langle C-x|u \rangle \leq 0\}, & \text{if } x \in C \\ \emptyset, & \text{else} \end{cases}$ */

$\Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ u \in N_C x$

$\Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ N_C^{-1} u \ni x$

$\therefore \forall_x \left(x \in \text{spts}\, C \Leftrightarrow \exists_{u \in \mathcal{H}\setminus\{0\}}\ x \in N_C^{-1} u \right) \Leftrightarrow \text{spts}\, C = N_C^{-1}(\mathcal{H}\setminus\{0\})$ ∎

Theorem 7.4. (Bishop-Phelps)

[C: nonempty closed convex, $\subseteq \mathcal{H}$] $\Rightarrow$ $\text{spts}\, C = P_C(\mathcal{H}\setminus C)$, $\overline{\text{spts}\, C} = \text{bdry}\, C$

Proof: $C = \mathcal{H} \Rightarrow$ trivial

So, take $C \neq \mathcal{H}$ First we prove $\text{spts}\, C \subseteq P_C(\mathcal{H} \setminus C)$

$\forall \varepsilon \in \mathbb{R}_{++},\ x \in \text{spts}\, C$ /* $\text{spts}\, C = \{y \in C \mid \exists_{u_y \in \mathcal{H}\setminus\{0\}}\ \langle y | u_y \rangle \geq \sup \langle C | u_y \rangle\}$

$\Updownarrow$

$\exists_{u \in \mathcal{H}\setminus\{0\}}\ \underbrace{\langle x|u \rangle \geq \sup \langle C|u \rangle \geq \langle y|u \rangle\ \forall y \in C}_{\langle y-x|u \rangle \leq 0\ \forall y \in C}$

$\Updownarrow$

$\varepsilon \langle y-x|u \rangle \leq 0\ \forall y \in C \Leftrightarrow \varepsilon \langle y-x|u \rangle \leq 0\ \forall y \in C$

$\Updownarrow$

$\langle y-x|\varepsilon u \rangle \leq 0\ \forall y \in C$

$\Updownarrow$

$\langle y-x|(x+\varepsilon u)-x \rangle \leq 0\ \forall y \in C$ /* now

/* Characterization of projection on closed convex nonempty set */ Theorem 3.14 ***

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

$x \in \text{spts}\, C \subseteq C$

$\therefore P_C(x+\varepsilon u) = x$

now take $x + \varepsilon u \in C \Rightarrow P_C(x+\varepsilon u) = x + \varepsilon u \Rightarrow \varepsilon = 0 \vee u = 0 \Rightarrow$ contradiction

$\therefore \forall \varepsilon > 0\quad x + \varepsilon u \notin C \Rightarrow P_C(x+\varepsilon u) = x \in P_C(\mathcal{H}\setminus C)$

$\therefore \forall_{x \in \text{spts}\, C}\quad P_C(x+\varepsilon u) = x \in P_C(\mathcal{H}\setminus C)$

$\Leftrightarrow \boxed{\text{spts}\, C \subseteq P_C(\mathcal{H}\setminus C)}$

also $\forall_{\varepsilon > 0}\ x + \varepsilon u \notin C$

[Procedure 1]

Now let us show: $P_C(\mathcal{H}\setminus C) \subseteq \text{spts}\, C$

$\Updownarrow$

$\forall_{y \in \mathcal{H}\setminus C}\ P_C y \in \text{spts}\, C$

$y \in \mathcal{H}\setminus C,\ x := P_C y$

/* Proposition 6.46. [C: nonempty closed convex subset of $\mathcal{H}$, $x, p \in \mathcal{H}$] $p = P_C x \Leftrightarrow x - p \in N_C p$ */

$\forall_{y\in\mathcal{H}\setminus C}\ P_C y \in \operatorname{spts} C$

$y\in\mathcal{H}\setminus C,\ x:=P_C y$ /* now * Proposition 6.46. [C: nonempty closed convex subset of $\mathcal{H}$, $x,p\in\mathcal{H}$] $p=P_C x \Leftrightarrow x-p\in N_C p$ */

$\Updownarrow$

$y-x\in N_C x \leftrightarrow N_C^{-1}(\underbrace{y-x}_{\neq 0})\ni x$ /* now: Proposition 7.3. [C: nonempty convex set, $\in\mathcal{H}$] $\operatorname{spts} C = N_C^{-1}(\mathcal{H}\setminus\{0\})$

// as $y\notin C, x\in C$

$\Leftrightarrow x\in N_C^{-1}(y-x)\subseteq N_C^{-1}(\mathcal{H}\setminus\{0\})=\operatorname{spts} C$

So, $\forall_{y\in\mathcal{H}\setminus C}\ x=P_C y\in\operatorname{spts} C \Leftrightarrow \boxed{P_C(\mathcal{H}\setminus C)\subseteq \operatorname{spts} C}$

$P_C(\mathcal{H}\setminus C)=\operatorname{spts} C$

Now let us show $\operatorname{bdry} C\subseteq\overline{\operatorname{spts} C}$

$\forall z$

// From Procedure 2

$z\in\operatorname{bdry} C \Rightarrow \forall_{\varepsilon>0}\ \exists_{y\in\mathcal{H}\setminus C}\ \|z-y\|<\varepsilon,\ z\in C$

$z\in C$

now take $p:=P_C y\in P_C(\mathcal{H}\setminus C)=\operatorname{spts} C \Leftrightarrow p\in\operatorname{spts} C$

/* now

Proposition 4.2.

[C: nonempty closed convex set of $\mathcal{H}$] $\Rightarrow P_C$: firmly nonexpansive */

$\Rightarrow P_C$: nonexpansive

$\therefore \|\underbrace{P_C y}_{p}-\underbrace{P_C z}_{z}\|=\|p-z\|\leq\|y-z\|<\varepsilon$ $\therefore \|p-z\|<\varepsilon$

/* $z\in\operatorname{bdry} C\subseteq C \therefore P_C z=z$ */

$\forall_{\varepsilon>0}\ p\in\operatorname{spts} C,\ \|p-\overset{\in\operatorname{bdry} C}{z}\|<\varepsilon \Leftrightarrow p\in\operatorname{spts} C,\ p\in V_\varepsilon(z)\ \forall_{\varepsilon>0} \Rightarrow \forall_{\varepsilon>0}\ V_\varepsilon(z)\cap\operatorname{spts} C\ni p$ /* recall: $z\in\bar{S}\Leftrightarrow\forall_{\varepsilon>0}\ V_\varepsilon(z)=\{y\mid\|y-z\|<r\}\cap S\neq\emptyset$ */

$\Updownarrow$

$\forall_{\varepsilon>0}\ V_\varepsilon(z)\cap\operatorname{spts} C\neq\emptyset$

$\Updownarrow$

$z\in\overline{\operatorname{spts} C}$ So, $\forall z\in\operatorname{bdry} C\ z\in\overline{\operatorname{spts} C}$

$\therefore \boxed{\operatorname{bdry} C\subseteq\overline{\operatorname{spts} C}}$ [eq 2017-01-20 1:58 PM]

Now let us show: $\overline{\operatorname{spts} C}\subseteq\underbrace{\operatorname{bdry} C}_{C/\operatorname{int} C}$: lets simplify them

$\Leftrightarrow \forall_{z\in\overline{\operatorname{spts} C}}\ z\in\operatorname{bdry} C$

now: by def: $z\in\bar{S}\Leftrightarrow\forall_{\varepsilon>0}\ V_\varepsilon(z)\cap S\neq\emptyset$ $\therefore z\in\overline{\operatorname{spts} C}\Leftrightarrow\forall_{\varepsilon>0}\ V_\varepsilon(z)\cap\underbrace{\operatorname{spts} C}_{\neq\emptyset}\Leftrightarrow\forall_{\frac{\varepsilon}{2}>0}\ \exists_{q\in C,\ \|q-z\|<\frac{\varepsilon}{2}}\ \exists_{u\neq 0}\ \langle q|u\rangle\geq\sup\langle C|u\rangle$

$\{\tilde{q}\in C\mid\exists_{u_{\tilde{q}}\neq 0}\ \langle\tilde{q}|u_{\tilde{q}}\rangle\geq\sup\langle C|u_{\tilde{q}}\rangle\}$

$\Downarrow$ similar to [Procedure 1]

$\exists_{u\neq 0}\ \sup\langle C-q\mid(q+\frac{\varepsilon}{2}u)-q\rangle\leq 0$

$\Downarrow$

$\underbrace{q}_{\in C}=P_C(\underbrace{q+\frac{\varepsilon}{2}u}_{\notin C})$

Procedure 2

and $z\in\operatorname{bdry} C=C\setminus\operatorname{int} C\Leftrightarrow z\in C,\ z\notin\operatorname{int} C\Leftrightarrow z\in C,\ \forall_{\varepsilon>0}\ \exists_{y:\|y-z\|<\varepsilon}\ y\notin C$: this is what we want to show

$\Updownarrow$

/* $\neg(\exists_{\varepsilon>0}\ V_\varepsilon(z)\subseteq C)\Leftrightarrow\forall_{\varepsilon>0}\ \neg(\forall_{y:\|z-y\|<\varepsilon}\ y\in C)\Leftrightarrow\forall_{\varepsilon>0}\ \exists_{y:\|z-y\|<\varepsilon}\ y\notin C$ */

$\Updownarrow$

$(\forall_{\underset{\|z-y\|<\varepsilon}{y\in V_\varepsilon(z)}}\ y\in C)$

Proving this:

$\forall_{\varepsilon>0}$ set $y=q+\frac{\varepsilon}{2}u\notin C$, $\|y-z\|=\|z-u\|\leq\|q-z\|+\|q+\frac{\varepsilon}{2}u-q\|<\varepsilon$ ☑

$<\frac{\varepsilon}{2}$ $\quad\frac{\varepsilon}{2}\|u\|=\frac{\varepsilon}{2}$

now $\forall_{\varepsilon>0}\ \exists_{z^c\in C}\ \left(\frac{\varepsilon}{2}>\|z-z^c\|\right) \Rightarrow \forall_{\tilde{y}\in C}\ \|z-\tilde{y}\|\leq\|z-z^c\|$ $\tilde{y}=z\in C$

$\Rightarrow \forall_{\varepsilon>0}\ \varepsilon>\|z-z^c\|$ /*

Result 1.2.2 (Equality of two numbers, and two vectors)

[$a,b\in\mathbb{R}$] $(a=b)\Leftrightarrow\forall_{\varepsilon>0}\ |a-b|<\varepsilon$; [$x,y\in\mathcal{H}$] $x=y\Leftrightarrow\forall_{\varepsilon>0}\ \|x-y\|<\varepsilon$ */

$\leftrightarrow z=z^c$

$\leftrightarrow \boxed{z\in C}$ ☑

$\therefore z\in\operatorname{bdry} C$

So, $\forall_{z\in\overline{\operatorname{spts} C}}\ z\in\operatorname{bdry} C\Leftrightarrow\boxed{\overline{\operatorname{spts} C}\subseteq\operatorname{bdry} C}$ [eq 2017-01-20 1:59 PM]

[eq 2017-01-20 1:58 PM], [eq 2017-01-20 1:59 PM] $\Rightarrow \overline{\operatorname{spts} C}=\operatorname{bdry} C$ ■

1/11/2016 3:07 PM

Corollary 7.6.

[C: nonempty closed convex subset of $\mathcal{H}$

one of the following holds (i) $\operatorname{int} C\neq\emptyset$,

(ii) C: closed affine subspace

(iii) $\mathcal{H}$: finite dimensional]

[C: nonempty closed convex subset of $\mathcal{H}$
one of the following holds (i) $\operatorname{int} C \neq \emptyset$,
(ii) C: closed affine subspace
(iii) $\mathcal{H}$: finite dimensional]

$\Rightarrow$

$\operatorname{spts} C = \operatorname{bdry} C$

Proof:

(i)

/*

* Theorem 7.4. (Bishop-Phelps) [C: nonempty closed convex subset of $\mathcal{H}$] $\operatorname{spts} C = P_C(\mathcal{H}\setminus C)$

$\overline{\operatorname{spts} C} = \operatorname{bdry} C$

Proposition 7.5. [C: convex subset of $\mathcal{H}$, $\operatorname{int} C \neq \emptyset$] $\operatorname{bdry} C \subseteq \operatorname{spts} C$

$C \cap \operatorname{bdry} C \subseteq \operatorname{spts} C$

*/

$\overline{\operatorname{spts} C} = \operatorname{bdry} C \subseteq \overline{\operatorname{spts} C} = \operatorname{spts} C$ [$\because$ C: closed]

now we want to show that, $\operatorname{spts} C \subseteq \operatorname{bdry} C$

as $\operatorname{spts} C = P_C(\mathcal{H}\setminus C)$, so any point outside C will be projected on the boundary, i.e. bdry C

$\therefore \operatorname{spts} C = P_C(\mathcal{H}\setminus C) \subseteq \operatorname{bdry} C$ [1/11/2016 1:11 PM]

$\therefore \operatorname{spts} C = \operatorname{bdry} C$.

(ii) C: closed affine subspace

if $C = \mathcal{H}$ then obviously $\operatorname{spts} C = \operatorname{bdry} C$

if $C \neq \mathcal{H}$

$V = C - C$: closed linear subspace parallel to C [page 1, Bauschke]

$x \in \operatorname{bdry} C$, $u \in V^{\perp}\setminus\{0\}$

$\forall_{\tilde{x} \in C}$ $C = \tilde{x} + V$ [page 1, Bauschke] $\Rightarrow$ $\overset{\in \operatorname{bdry} C \subseteq C}{C = x + V}$

/* Proposition 3.17. $P_{\underbrace{y+C}_{\in \mathcal{H}}} \overset{\in \mathcal{H}}{x} = y + P_C(x-y)$ */ $\therefore P_C(x+u) = P_{x+V}(x+u) = x + P_V(x+u-x) = x + P_V u$

nonempty closed convex subset of $\mathcal{H}$

Now recall : /* (Corollary 3.18)

* Projection onto a closed linear subspace:
[V: closed linear subspace of $\mathcal{H}$, $x \in \mathcal{H}$]
(i) $P_V x \in V$, $x - P_V x \in V^{\perp}$
(ii) $\|P_V x\|^2 = \langle P_V x, x\rangle$
(iii) $[P_V x = x \Leftrightarrow x \in V]$, $[P_V x = 0 \Leftrightarrow x \in V^{\perp}]$
$[P_V x = 0 \Leftrightarrow V = \{0\}]$
(iv) $V^{\perp\perp} = V$
(v) $P_{V^{\perp}} = \mathrm{Id} - P_V$
(vi) $P_{V^{\perp}} = P_V$
(vii) $\|x\|^2 = \|P_V x\|^2 + \|x - P_V x\|^2 = \|P_V x\|^2 + \|P_{V^{\perp}} x\|^2$

*/

so, $\|u\|^2 = \|P_V u\|^2 + \|\underbrace{P_{V^{\perp}} u}_{u \ [\because u \in V^{\perp}]}\|^2 = \|P_V u\|^2 + \|u\|^2$

$\Rightarrow \|P_V u\|^2 = 0$

$\Leftrightarrow P_V u = 0$

$P_C(x+u) = x$, $\therefore \forall_{x \in \operatorname{bdry} C}$ $P_C(x+u) = x$

now, $x + u \in \mathcal{H}$, but $x + u \notin C$ /* if $\underbrace{x}_{\operatorname{bdry} C} + \underbrace{u}_{\in V^{\perp}\setminus\{0\}} \in C = x + V \Rightarrow u \in V \Rightarrow$ contradiction */

$\Rightarrow x + u \in \mathcal{H}\setminus C$

$\Rightarrow \underbrace{P_C(x+u)}_{= x} \in P_C(\mathcal{H}\setminus C) = \operatorname{spts} C$ /* using Recall Bishop-Phelps theorem [C: nonempty closed convex subset of $\mathcal{H}$] $\operatorname{spts} C = P_C(\mathcal{H}\setminus C)$

$\overline{\operatorname{spts} C} = \operatorname{bdry} C$ */

$\Rightarrow \forall_{x \in \operatorname{bdry} C}$ $x \in \operatorname{spts} C$ $\therefore \operatorname{bdry} C \subseteq \operatorname{spts} C$,

and, $\operatorname{spts} C \subseteq \operatorname{bdry} C$ From [1/11/2016 1:11 PM]

$\therefore \operatorname{bdry} C = \operatorname{spts} C$. ☺

(iii)

$\mathcal{H}$: finite dimensional

if $\operatorname{int} C \neq \emptyset \Rightarrow \operatorname{spts} C = \operatorname{bdry} C$ [proved in (i)]

now consider $\operatorname{int} C = \emptyset$

$\begin{cases} D = \operatorname{aff} C \\ x \in \operatorname{bdry} C \end{cases}$

/*

* Proposition 6.12
[C: convex subset of $\mathcal{H}$]
$(\operatorname{int} C \neq \emptyset \vee C$: closed $\vee$ $\mathcal{H}$: finite-dimensional$) \Rightarrow \operatorname{int} C = \operatorname{core} C$ */

$\because \operatorname{int} C = \operatorname{core} C = \{x \in C \mid \operatorname{cone}(C - x) = \mathcal{H}\} = \emptyset$

$\neg(\exists_{x \in C} \operatorname{cone}(C - x) = \mathcal{H})$

$\Leftrightarrow \forall_{x \in C} \operatorname{cone}(C - x) \neq \mathcal{H}$

$D = \operatorname{aff} C \neq H$ /* explanation needed */

/* Affine hull of a set in a finite-dimensional space is closed */

$\Rightarrow$ D: proper closed affine subspace of H

$\Rightarrow$ $x \in \operatorname{rbdry} D$

using (ii) : and Proposition 7.2.: $x \in C \cap \operatorname{spts} D \subseteq \operatorname{spts} C$. /* explanation needed */

$\blacksquare$

7.2 Support Functions, 7.3 Polar Sets

Monday, January 11, 2016 3:08 PM

1/11/2016 5:13 PM

Proposition 7.9.

[$C \subseteq \mathcal{H}$, $\forall_{u \in \mathcal{H}}$ $H_u = \{x \in \mathcal{H} \mid \langle x|u\rangle \leq \sigma_C(u)\}$] $\overline{\text{conv}}\, C = \bigcap_{u \in \mathcal{H}} H_u$

Proof:

- $C = \emptyset \Rightarrow$ obvious
- $C \neq \emptyset$.

$D := \bigcap_{u \in \mathcal{H}} H_u$ (intersection of closed sets)

$H_u = \{x \in \mathcal{H} \mid \langle x|u\rangle \leq \sigma_C(u)\}$

By definition H_u is a closed half-space containing C, i.e. $\forall_{u \in \mathcal{H}}$ $C \subseteq H_u \Rightarrow D$: closed, convex. (half-spaces are convex) $\underbrace{\overline{\text{conv}}\, C}_{\text{smallest closed convex set containing } C} \subseteq D$

/* $\sigma_C(u) = \sup_{\tilde{x} \in C} \langle \tilde{x}|u\rangle \geq \langle x|u\rangle$ $\forall_{x \in C}$ $\leftrightarrow \forall_{x \in C}$ $\underbrace{\langle x|u\rangle \leq \sigma_C(u)}_{\Updownarrow\ x \in H_u}$

$\Rightarrow \forall_{x \in C}$ $x \in H_u \leftrightarrow C \subseteq H_u$ */

/* as $C \subseteq H_u$ $\forall_{u \in \mathcal{H}}$ now each H_u: closed, convex

$\hookrightarrow H_u$ is a closed convex set containing C

$\Rightarrow \overline{\text{conv}}\, C \subseteq \overline{\text{conv}}\, H_u$ /* $A \subseteq B \Rightarrow \overline{\text{conv}}\, A \subseteq \overline{\text{conv}}\, B$

$= H_u$

$\Rightarrow \overline{\text{conv}}\, C \subseteq H_u$ $\forall_{u \in \mathcal{H}}$ $\Rightarrow \overline{\text{conv}}\, C \subseteq \bigcap_{u \in \mathcal{H}} H_u = D$ */

$\therefore \overline{\text{conv}}\, C \subseteq D = \bigcap_{u \in \mathcal{H}} H_u$

let's prove $D \subseteq \overline{\text{conv}}\, C$

let. $x \in D$, $p = P_{\overline{\text{conv}}\, C}\, x$. $\tilde{C} = \overline{\text{conv}}\, C$

/*

/* Characterization of projection on closed convex nonempty set */ Theorem 3.14 ***

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

$\forall_{y \in \tilde{C}}$ $\langle y-p|x-p\rangle = \langle y|x-p\rangle - \langle p|x-p\rangle \leq 0$

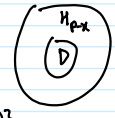
$\leftrightarrow \forall_{y \in \tilde{C}}$ $\langle y|x-p\rangle \leq \langle p|x-p\rangle$

$\leftrightarrow$ $\sup_{y \in \tilde{C}} \langle y|x-p\rangle = \sigma_{\tilde{C}}(x-p) \leq \langle p|x-p\rangle$

$\leftrightarrow$ $\sigma_{\overline{\text{conv}}\, C}(x-p) \leq \langle p|x-p\rangle$

now, $x \in D = \bigcap_{u \in \mathcal{H}} H_u \subseteq H_{x-p}$ [$\because p - x \in \mathcal{H}$]

$H_{x-p} = \{\underbrace{z}_{} \in \mathcal{H} \mid \langle z|x-p\rangle \leq \sigma_C(x-p)\}$



$\Rightarrow x \in H_{p-x} \leftrightarrow \langle x|x-p\rangle \leq \sigma_C(x-p)$

So. $\|x-p\|^2 = \langle x-p|x-p\rangle = \underbrace{\langle x|x-p\rangle}_{\leq \sigma_C(x-p)} - \underbrace{\langle p|x-p\rangle}_{\leq -\sigma_{\overline{\text{conv}}\, C}(x-p)} \leq \sigma_C(x-p) - \sigma_{\overline{\text{conv}}\, C}(x-p)$

$\leftrightarrow \|x-p\|^2 \leq \sigma_C(x-p) - \sigma_{\overline{\text{conv}}\, C}(x-p)$

now, $C \subseteq \overline{\text{conv}}\, C$

so. $\underbrace{\sup_{\tilde{x} \in \overline{\text{conv}}\, C} \langle \tilde{x}|x-p\rangle}_{= \sigma_{\overline{\text{conv}}\, C}(x-p)}$ is a relaxation of $\underbrace{\sup_{\tilde{x} \in C} \langle \tilde{x}|x-p\rangle}_{= \sigma_C(x-p)} \Rightarrow \sigma_{\overline{\text{conv}}\, C}(x-p) \geq \sigma_C(x-p)$

/* relaxation always have better objective value */

$\Rightarrow 0 \leq \|x-p\|^2 \leq 0 \leftrightarrow \|x-p\|^2 = 0 \leftrightarrow x = p = P_{\overline{\text{conv}}\, C}\, x \in \overline{\text{conv}}\, C$

$\therefore \forall_{x \in \underbrace{\bigcap_{u \in \mathcal{H}} H_u}_{= D}}$ $x \in \overline{\text{conv}}\, C \leftrightarrow \bigcap_{u \in \mathcal{H}} H_u \subseteq \overline{\text{conv}}\, C$

$\overline{\text{conv}}\, C = \bigcap_{u \in \mathcal{H}} H_u$ ∎

$\overline{\text{conv}}\, C = \bigcap_{u \in \mathcal{H}} H_u$ ∎

Theorem 7.16 [$C \subseteq \mathcal{H}$] $C^{\odot\odot} = \overline{\text{conv}}(C \cup \{0\})$

$\delta_C(u) = \sup\langle C|u\rangle$

Proof: // recall: [$C \subseteq \mathcal{H}$] $C^{\odot} = \text{lev}_{\leq 1}\, \delta_C = \{u \in \mathcal{H} \mid \delta_C(u) \leq 1\}$

* Proposition 7.14 [C, D: subsets of $\mathcal{H}$]
(i) $C^{\perp} \subseteq C^{\odot} \subseteq C^{\oplus}$
(ii) $0 \in C^{\odot}$, $C^{\odot}$: closed and convex
(iii) $C \cup \{0\} \subseteq C^{\odot\odot}$
(iv) $C \subseteq D \Rightarrow \underbrace{D^{\odot} \subseteq C^{\odot}}_{\text{inclusion flips by polar set operator}}$, $\underbrace{C^{\odot\odot} \subseteq D^{\odot\odot}}_{\text{double polar set operator maintains the inclusion}}$
(v) $C^{\odot\odot\odot} = C^{\odot}$
(vi) $(\overline{\text{conv}}\, C)^{\odot} = C^{\odot}$

$C \cup \{0\} \subseteq C^{\odot\odot}$ /* as $A \subseteq B \Rightarrow \overline{\text{conv}}\, A \subseteq \overline{\text{conv}}\, B$

$\Rightarrow \overline{\text{conv}}(C \cup \{0\}) \subseteq \overline{\text{conv}}(C^{\odot\odot})$ // now $C^{\odot}$: closed, convex

$= C^{\odot\odot}$ // $\Rightarrow C^{\odot\odot}$: closed, convex

$\therefore \overline{\text{conv}}(C \cup \{0\}) \subseteq C^{\odot\odot}$

lets prove $C^{\odot\odot} \subseteq \overline{\text{conv}}(C \cup \{0\})$

for absurdum assume $C^{\odot\odot} \not\subseteq \overline{\text{conv}}(C \cup \{0\})$

$\leftrightarrow \neg\left(\forall_{x \in C^{\odot\odot}}\ x \in \overline{\text{conv}}(C \cup \{0\})\right)$

$\leftrightarrow \exists_{x \in C^{\odot\odot}}\ x \notin \overline{\text{conv}}(C \cup \{0\})$

$\leftrightarrow \exists_{x}\ x \in C^{\odot\odot} \setminus \overline{\text{conv}}(C \cup \{0\})$

now $0 \in C^{\odot\odot}$ and $0 \in \overline{\text{conv}}(C \cup \{0\})$, $x \neq 0$

so more precisely:

$\exists_{x \in \mathcal{H} \setminus \{0\}}\ x \in C^{\odot\odot} \setminus \overline{\text{conv}}(C \cup \{0\})$

/*
Theorem 3.38.
[$\tilde{C}$: nonempty closed convex subset of $\mathcal{H}$
$x \in \mathcal{H} \setminus \tilde{C}$] $\Rightarrow$ x: strongly seperated from $\tilde{C}$ $\leftrightarrow \exists_{v \in \mathcal{H} \setminus \{0\}}\ \sup\langle \tilde{C}|v\rangle < \langle x|v\rangle$ */

$\tilde{C} := \overline{\text{conv}}(C \cup \{0\})$

$\exists_{v \in \mathcal{H} \setminus \{0\}}\ \langle x|v\rangle > \sup\langle \tilde{C}|v\rangle = \delta_{\tilde{C}}(v)$ // By definition: $\delta_C(u) = \sup\langle C|u\rangle$

$= \delta_{\overline{\text{conv}}(C \cup \{0\})}(v)$

now: $\delta_{\overline{\text{conv}}(C \cup \{0\})}(v) = \sup\langle \overline{\text{conv}}(C \cup \{0\})|v\rangle \geq \max\{\delta_C(v), 0\}$

sup over a larger set is larger set

/* using $\tilde{C} = \overline{\text{conv}}(C \cup \{0\}) \supseteq C \Rightarrow \forall_{\tilde{a}}\ \sup\langle \tilde{C}|\tilde{a}\rangle \geq \sup\langle C|\tilde{a}\rangle = \delta_C(\tilde{a})$
$\tilde{C} = \overline{\text{conv}}(C \cup \{0\}) \supseteq 0 \Rightarrow \forall_{\tilde{a}}\ \sup\langle \tilde{C}|\tilde{a}\rangle \geq \sup\langle 0|\tilde{a}\rangle = 0$
$\Rightarrow \forall_{\tilde{a}}\ \sup\langle \tilde{C}|\tilde{a}\rangle \geq \max\{\delta_C(\tilde{a}), 0\}$ */

$\therefore \langle x|v\rangle > \delta_{\overline{\text{conv}}(C \cup \{0\})}(v) \geq \max\{\delta_C(v), 0\} \geq 0$

$\beta > 0$ // say $\langle x|v\rangle$ is strictly positive

$\Rightarrow \langle x|v\rangle > \max\{\delta_C(v), 0\} \geq 0$ /* $C^{\odot} = \text{lev}_{\leq 1}\, \delta_C = \{\tilde{v} \in \mathcal{H} \mid \delta_C(\tilde{v}) \leq 1\}$ */

after scaling v if necessary: setting $v = \alpha u$: $\alpha > 0$

$\Rightarrow$

$\alpha \langle x|u\rangle = \beta > \max\{\delta_C(\alpha u), 0\} = \alpha \max\{\delta_C(u), 0\}$

$\Rightarrow \langle x|u\rangle = \frac{\beta}{\alpha} > \max\{\delta_C(u), 0\}$ /* for any β we can choose an α: such that $\langle x|u\rangle = \frac{\beta}{\alpha} > 1$ */

so, we can say:

$\langle x|u\rangle > 1 \geq \delta_C(u) = \sup\langle C|u\rangle$

$\langle x|u\rangle > 1 \geq \underbrace{\sigma_C(u) = \sup\langle C|u\rangle}$

$\Downarrow$

$u \in C^{\Theta}$

now $\sigma_{C^{\Theta}}(x) = \sup\langle C^{\Theta}|x\rangle \geq \langle u|x\rangle$

but $x \in \mathrm{lev}_{\leq 1}\sigma_{C^{\Theta}} \setminus \overline{\mathrm{conv}}(C\cup\{0\})$

$\Rightarrow \sigma_{C^{\Theta}}(x) \leq 1$

$\langle u|x\rangle \leq \sigma_{C^{\Theta}}(x) \leq 1$

$1 < \langle x|u\rangle \leq \sigma_{C^{\Theta}}(x) \leq 1 \Rightarrow 1<1 \Rightarrow$ contradiction

$\blacksquare$

Corollary 7.17.

[[$C \subseteq \mathcal{H}$]]

(i) (C: closed, convex, $0 \in C$) $\Leftrightarrow C^{\Theta\Theta} = C$

(ii) (C: nonempty, closed, convex) $\Leftrightarrow C^{\Theta\Theta} = C$

(iii) (C: closed linear subspace) $\Leftrightarrow C^{\perp\perp} = C$

1/12/2016 5:55 PM

Proof:

(i)

/* * Proposition 7.14. [[C, D: subsets of $\mathcal{H}$]] (i) $C^{\perp} \subseteq C^{\Theta} \subseteq C^{\Theta}$

(ii) $0 \in C^{\Theta}$, C^{Θ}: closed and convex */

($\Leftarrow$) $(C^{\Theta})^{\Theta} \Rightarrow$ closed, convex, $0 \in C^{\Theta\Theta}$

($\Rightarrow$)

C: closed, convex, $0 \in C$

$\Rightarrow C = C\cup\{0\}$: nonempty, convex, closed

$\Rightarrow \underbrace{\overline{\mathrm{conv}}(C\cup\{0\})}_{C^{\Theta\Theta}} = C$ /* Theorem 7.16. [[$C \subseteq \mathcal{H}$]] $C^{\Theta\Theta} = \overline{\mathrm{conv}}(C\cup\{0\})$ */ ☺

(ii)

* Proposition 6.23. [[C: subset of $\mathcal{H}$]] $\Rightarrow$ (i) $D \subseteq C \Rightarrow$ • $C^{\Theta} \subseteq D^{\Theta}$ /* inclusion flips */
• $C^{\Theta} \subseteq D^{\Theta}$

(ii) C^{Θ}, C^{Θ} : nonempty closed convex cones

(iii) $C^{\Theta} = (\mathrm{cone}\, C)^{\Theta} = (\mathrm{conv}\, C)^{\Theta} = \overline{C}^{\Theta}$ /* to the polar cone operator, a set, its cone, convex hull and closure all are same */

(iv) $\overline{\mathrm{cone}}\, C = \overline{\mathrm{conv}}\, C \Rightarrow C^{\Theta} = C^{\Theta} = C^{\perp}$

C: nonempty closed convex cone

$\Rightarrow C^{\Theta}$: nonempty closed convex cone $= C^{\Theta}$ /* $\because K$: cone in $\mathcal{H} \Rightarrow K^{\Theta} = K^{\Theta}$ */

$\Rightarrow C^{\Theta\Theta}$: nonempty closed convex cone, and contains 0 /* $\forall_{K \subseteq \mathcal{H}}$ $K^{\Theta} = \{u \in \mathcal{H} | \sup\langle C|u\rangle \leq 0\} \Rightarrow 0 \in K^{\Theta}$ as $\sup\langle C|0\rangle = 0$ */

$= (C^{\Theta})^{\Theta} = (C^{\Theta})^{\Theta} = C$ [from (i) : $C^{\Theta\Theta} = C$]

$\therefore C^{\Theta\Theta} = C$ ☺

(iii) C: closed linear subspace

$\Rightarrow C$: nonempty closed convex cone /* A closed linear subspace is also a nonempty closed convex cone */

$\Rightarrow C^{\Theta\Theta} = C$

$\Rightarrow C$: nonempty closed convex cone /* A closed linear subspace is also a nonempty closed convex cone */

$\Rightarrow C^{\ominus\ominus} = C$

But *Proposition 6.22. [C: linear subspace of $\mathcal{H}$] $\Rightarrow C^{\ominus} = C^{\perp}$

$\therefore C^{\perp\perp} = C$ ∎

1pChapter 8

Tuesday, January 12, 2016 9:29 PM

* Proposition 8.2.

$f: \mathcal{H} \to [-\infty, +\infty]$: convex $\Rightarrow$ $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$: convex

1/12/2016 9:50 PM

Proof: $L: \mathcal{H} \times \mathbb{R} \to \mathcal{H} : (x, \xi) \mapsto x$

$$\left.\begin{array}{l} L(x,\xi) = x \\ L(y,\xi) = y \end{array}\right\} L(\alpha x + \beta y, \xi) = \alpha x + \beta y = \alpha L(x,\xi) + \beta L(y,\xi)$$

$\therefore$ L : linear operator

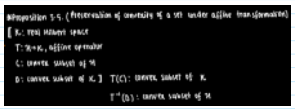
$$L(\operatorname{epi} f) = L(\{(x,\xi) \in \mathcal{H} \times \mathbb{R} \mid f(x) \leq \xi\})$$

$$= \{x \in \mathcal{H} \mid f(x) \leq \xi, \xi \in \mathbb{R}\} \quad \text{/* recall : } [-\infty, \infty] = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\} \text{ */}$$

$$= \{x \in \mathcal{H} \mid f(x) < +\infty\}$$

By definition $\operatorname{epi} f$: convex $\Rightarrow$ $L(\operatorname{epi} f)$: convex /* using :

linear hence affine operator


Proposition 5.5. (Preservation of convexity of a set under affine transformation)
[$\mathcal{H}$: real Hilbert space
$T: \mathcal{H} \to \mathcal{K}$, affine operator
C: convex subset of $\mathcal{H}$
D: convex subset of $\mathcal{K}$.] $T(C)$: convex subset of $\mathcal{K}$
$T^{-1}(D)$: convex subset of $\mathcal{H}$

*/

$\therefore$ $\operatorname{dom} f$: convex. ∎

1/13/2016 6:43 PM

* Proposition 8.12. (Convexity conditions for function on real line)

[[$\varphi: \mathbb{R} \to]-\infty, +\infty]$, proper function, differentiable on I]]

nonempty open interval in $\operatorname{dom} \varphi$

(i) φ' : increasing on $I \Rightarrow \varphi + L_I$: convex

(ii) φ' : strictly increasing on $I \Rightarrow \varphi$: strictly convex on I.

Proof:

$\forall_{x,y \in I}$ $\forall_{\alpha \in]0,1[}$

$\psi: \mathbb{R} \to]-\infty, +\infty] : z \mapsto \alpha\varphi(x) + (1-\alpha)\varphi(z) - \varphi(\alpha x + (1-\alpha) z)$

$\psi(\underbrace{z}_{\in I}) = \alpha\varphi(x) + (1-\alpha)\varphi(z) - \varphi(\alpha x + (1-\alpha) z)$

then $\underbrace{\psi'(z)}_{\text{will exist by given}} = 0 + (1-\alpha)\varphi'(z) - \underbrace{\frac{d}{d(\alpha x + (1-\alpha) z)} \varphi(\alpha x + (1-\alpha) z)}_{\varphi'(\alpha x + (1-\alpha) z)} \cdot \underbrace{\frac{d}{dz}(\alpha x + (1-\alpha) z)}_{(1-\alpha)}$

$$= (1-\alpha)\varphi'(z) - (1-\alpha)\varphi'(\alpha x + (1-\alpha) z)$$

$$= (1-\alpha)\left(\varphi'(z) - \varphi'(\underbrace{\alpha x + (1-\alpha) z}_{\alpha(x-z)+z})\right) \quad \ldots (8.10)$$

now $\psi'(x) = \psi'(z)\big|_{z:=x} = (1-\alpha)\left(\varphi'(x) - \varphi'(\cancel{\alpha x} + x - \cancel{\alpha x})\right)$

$$= (1-\alpha)\left(\varphi'(x) - \varphi'(x)\right) = 0$$

(i) go through the alternative proof:

if $z < x$, then $\psi'(z) = (1-\alpha)\left(\varphi'(z) - \varphi'(\underbrace{\alpha(x-z)}_{>0} + z)\right)$

$$= (1-\alpha)\underbrace{\left(\varphi'(z) - \varphi'(z_+)\right)}_{\leq 0} \leq 0$$

[$\because$ Given, φ' : increasing $\Leftrightarrow$ $\varphi'(z_+) - \varphi'(z) \geq 0$]

if $z > x$ then $\psi'(z) = (1-\alpha)\left(\varphi'(z) - \varphi'(\underbrace{\alpha(x-z)}_{<0} + z)\right)$

$$= (1-\alpha)\underbrace{\left(\varphi'(z) - \varphi'(z_-)\right)}_{\geq 0} \geq 0 \rightsquigarrow \forall_{z_1 < x < z_2} \quad \psi'(z_1) \leq \overset{\psi'(x)}{\overset{\shortparallel}{0}} \leq \psi'(z_2) \Rightarrow \psi'(z_1) \leq 0 \leq \psi'(z_2) \quad \therefore \forall_{z_1 < x < z_2} \psi'(z_1) \leq 0 \leq \psi'(z_2) \ldots (\text{eq: f})$$

When $z = x$, then $\psi'(z)_{z:=x} = 0$: local minimum achieved at $z = x$, the minimum value is $\psi(x) = \alpha\varphi(x) + (1-\alpha)\varphi(x) - \varphi(\alpha x + (1-\alpha) x) = \varphi(x) - \varphi(\alpha x + x - \alpha x) = 0$

So, from basic calculus, ψ achieves its global minimum of 0 on I at x

$\Leftrightarrow \forall_{z \in I, x \in I}$ $\psi(x) \leq \psi(z)$

/* $\psi(z) = \alpha\varphi(x) + (1-\alpha)\varphi(z) - \varphi(\alpha x + (1-\alpha) z)$

$\psi(x) = \alpha\varphi(x) + (1-\alpha)\varphi(x) - \varphi(\underbrace{\alpha x + (1-\alpha) x}_{\cancel{\alpha x} + x - \cancel{\alpha x} = x})$

$= \alpha\varphi(x) + \varphi(z) - \alpha\varphi(x) - \varphi(x) = 0$ */

$\Leftrightarrow \forall_{x,z \in I}$ $0 \leq \alpha\varphi(x) + (1-\alpha)\varphi(z) - \varphi(\alpha x + (1-\alpha) z)$

$\Leftrightarrow \forall_{x,z \in I}$ $\varphi(\underbrace{\alpha x + (1-\alpha) z}_{\alpha \in]0,1[}) \leq \alpha\varphi(x) + (1-\alpha)\varphi(z)$

$\Leftrightarrow$ φ : convex on I

$\Leftrightarrow$ $\varphi + L_I$: convex

(ii) just like (i)

$z < x \Rightarrow \psi'(z) < 0$

$z > x \Rightarrow \psi'(z) > 0$

$z := x \Rightarrow \psi'(x) = 0$

$\Rightarrow$ ψ : achieves strict minimum on I at x

$\forall_{z,x \in I}$ $\underset{0}{\underset{\shortparallel}{\psi(x)}} < \psi(z)$

$\Leftrightarrow \forall_{x,z \in I}$ $\varphi(\alpha x + (1-\alpha) z) < \alpha\varphi(x) + (1-\alpha)\varphi(z)$

$\Leftrightarrow$ φ : strictly convex on I

* alternative proof:

take $x_1, x_2, x_3 \in I$: $x_1 < x_2 < x_3$

By mean value theorem: $\exists_{\xi \in (x_1, x_2)}$ $\varphi'(\xi) = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$

$\exists_{\eta \in (x_2, x_3)}$ $\varphi'(\eta) = \dfrac{f(x_3) - f(x_2)}{x_3 - x_2}$ $\Rightarrow \xi < \eta$

now φ : increasing $\Rightarrow$ $\varphi'(\xi) \leq \varphi'(\eta)$

$\Rightarrow \dfrac{\varphi(x_2) - \varphi(x_1)}{x_2 - x_1} \leq \dfrac{\varphi(x_3) - \varphi(x_2)}{x_3 - x_2}$

f is convex on I if and only if:

$$\forall x_1, x_2, x_3 \in I : x_1 < x_2 < x_3 : \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}$$

this satisfies the definition of a convex function on $I \subseteq \mathbb{R}$

$\therefore$ φ : convex on $I \Rightarrow \varphi + L_I$: convex */

$\forall_{z,x\in I}$ $\overset{\Phi(x)}{\overset{\shortparallel}{0}}$

$\Leftrightarrow \forall_{x,z\in I} \quad \Phi(\kappa x+(1-\kappa)z) < \kappa\Phi(x)+(1-\kappa)\Phi(z)$

$\Leftrightarrow \Phi$: strictly convex on I

$\Leftrightarrow \Phi+L_1$: strictly convex ∎

\# Proposition 8.14.

$(f_i)_{i\in I}$: family of convex functions from $\mathcal{H}$ to $[-\infty,+\infty]$ $\Rightarrow \sup_{i\in I} f_i$: convex.

1/14/2016 5:32 AM

Proof:

/* recall

lemma 1.6. [$(f_i)_{i\in I}$: family of functions from $X\to[-\infty,+\infty]$] (i) $\operatorname{epi}(\sup_{i\in I} f_i)=\bigcap_{i\in I}\operatorname{epi} f_i$.

(ii) I : finite $\Rightarrow \operatorname{epi}(\min_{i\in I} f_i)=\bigcup_{i\in I}\operatorname{epi} f_i$. */

$\operatorname{epi}(\sup_{i\in I} f_i)=\underbrace{\bigcap_{i\in I}\operatorname{epi} f_i}_{\text{convex by given}}$: convex /* intersection of convex sets is convex */

$\Rightarrow \sup_{i\in I} f_i$: convex. ∎

* Proposition 8.19.

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

$\Phi:\mathbb{R}\to]-\infty,+\infty]$: convex.

$C=\operatorname{conv}(\mathbb{R}\cap\operatorname{ran} f)$ /* $\operatorname{ran} f=f(\mathcal{H})$, as $1\infty\notin\mathbb{R}$, so $C\not\ni 1\infty$ */

$\tilde{\Phi}$: extension of Φ, : $\underbrace{]-\infty,+\infty]}_{=\mathbb{R}\cup\{+\infty\}}\to]-\infty,+\infty]$, $\tilde{\Phi}(+\infty)=+\infty$

]

($C\subseteq\operatorname{dom}\Phi$, /* $\operatorname{dom}\Phi=\{x\in]-\infty,+\infty]\mid \Phi(x)<+\infty\}$ */

Φ : increasing on C)

$\Rightarrow$

$\tilde{\Phi}\circ f$: convex

1/14/2016 7:05 AM

Proof:

$\operatorname{dom}(\tilde{\Phi}\circ f)=\{x\in\mathcal{H}\mid(\tilde{\Phi}\circ f)(x)<+\infty\}=\{x\in\mathcal{H}\mid(\tilde{\Phi}\circ f)(x)<+\infty,\ \overset{x\in\operatorname{dom} f}{f(x)\neq\infty}\}=\{x\in\operatorname{dom} f\mid(\tilde{\Phi}\circ f)(x)<+\infty\}\subseteq\operatorname{dom} f$

$\tilde{\Phi}(f(x))<+\infty \Rightarrow f(x)\neq\infty$ [as $\tilde{\Phi}(+\infty)=+\infty$]

On $\operatorname{dom}(\tilde{\Phi}\circ f)$, $\tilde{\Phi}\circ f$ coincides with $\Phi\circ f$

as $x\in\operatorname{dom}(\tilde{\Phi}\circ f)\Leftrightarrow x\in\operatorname{dom} f,\ (\tilde{\Phi}\circ f)(x)<+\infty$

$\Leftrightarrow -\infty<f(x)<+\infty,\quad \tilde{\Phi}(\underbrace{f(x)}_{\in]-\infty,+\infty[=\mathbb{R}})=\Phi(f(x))$

/* $\because f:\mathcal{H}\to]-\infty,+\infty]$ */

$\therefore$ On $\operatorname{dom}(\tilde{\Phi}\circ f)$, $\tilde{\Phi}\circ f$ and $\Phi\circ f$ are same.

$\forall_{x,y\in\operatorname{dom}(\tilde{\Phi}\circ f)}\ \forall_{\kappa\in]0,1[}$

$f(\kappa x+(1-\kappa)y)\leq\kappa f(x)+(1-\kappa)f(y)$: clearly both L.H.S and R.H.S will belong to $\operatorname{conv}(\mathbb{R}\cap\operatorname{ran} f)$

now Φ: convex, increasing on $C\subseteq\operatorname{dom}\Phi$

$x\geq y\Leftrightarrow\Phi(x)\geq\Phi(y)$

so,

$\Phi(f(\kappa x+(1-\kappa)y)\leq\Phi(\kappa f(x)+(1-\kappa)f(y))\leq\kappa\Phi(f(x))+(1-\kappa)\Phi(f(y))$

[$\because \Phi$: convex on $\operatorname{dom}\Phi$ and $C\subseteq\operatorname{dom}\Phi$ by given

$\forall_{\tilde{x},\tilde{y}\in C}\ \forall_{\tilde{\kappa}\in]0,1[}\ \Phi(\tilde{\kappa}\tilde{x}+(1-\tilde{\kappa})\tilde{y})\leq\tilde{\kappa}\Phi(\tilde{x})+(1-\tilde{\kappa})\Phi(\tilde{y})$ */

$\Leftrightarrow(\tilde{\Phi}\circ f)(\kappa x+(1-\kappa)y)\leq\kappa(\tilde{\Phi}\circ f)(x)+(1-\kappa)(\tilde{\Phi}\circ f)(y)$ [$\because$ on $\operatorname{dom}(\tilde{\Phi}\circ f)$, $\tilde{\Phi}\circ f=\Phi\circ f$]

so, $\forall_{x,y\in\operatorname{dom}(\tilde{\Phi}\circ f)}\ \forall_{\kappa\in]0,1[}\ (\tilde{\Phi}\circ f)(\kappa x+(1-\kappa)y)\leq\kappa(\tilde{\Phi}\circ f)(x)+(1-\kappa)(\tilde{\Phi}\circ f)(y)$

$\Leftrightarrow(\tilde{\Phi}\circ f)$: convex. ∎

* Proposition 8.23.

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

$f:\mathbb{R}\times\mathcal{H}\to]-\infty,+\infty]$: $(\xi,x)\mapsto\begin{cases}\xi\Phi(x/\xi), & \text{if } \xi>0\\ +\infty, & \text{otherwise}\end{cases}$

]

1/15/2016 12:36 PM

Proof:

$C=\{1\}\times\operatorname{epi}\Phi$

$\Phi(\bar{x})\leq\bar{\xi}\ (<+\infty)\Leftrightarrow(\bar{x},\bar{\xi})\in\operatorname{epi}\Phi$

Φ: convex $\Rightarrow \operatorname{epi}\Phi=\{(\bar{x},\bar{\xi})\in\mathcal{H}\times\mathbb{R}\mid\Phi(\bar{x})\leq\bar{\xi}\}$: convex set /* By definition */

$\Rightarrow C:=\{1\}\times\operatorname{epi}\Phi=\{(1,\bar{x},\bar{\xi})\in\mathbb{R}\times\mathcal{H}\times\mathbb{R}\mid\Phi(\bar{x})\leq\bar{\xi}\}$: convex set on $\mathbb{R}\times\mathcal{H}\times\mathbb{R}$

$=\{(1,\bar{x},\bar{\xi})\in\mathbb{R}\times\mathcal{H}\times\mathbb{R}\mid(\bar{x},\bar{\xi})\in\operatorname{epi}\Phi\}$

/* By using Proposition 3.6. (i)
$(C_i)_{i\in I}$: totally ordered finite family of m convex subsets of $\mathcal{H}$

$\Rightarrow$

$\bigtimes_{i\in I} C_i$: convex set */

$f(\xi,x)=\begin{cases}\xi\Phi(x/\xi), & \text{if } \xi>0\\ +\infty, & \text{otherwise}\end{cases}$

now,

$\operatorname{epi} f=\{((\xi,x),\zeta)\in\mathbb{R}\times\mathcal{H}\times\mathbb{R}\mid\xi>0,\ f(\xi,x)\leq\zeta\}$

/* because else $f(\xi,x)=+\infty$, and by definition $\operatorname{epi} f$ only contains points <u>not</u> associated with $+\infty$ */

$=\{(\xi,x,\zeta)\in\mathbb{R}_{++}\times\mathcal{H}\times\mathbb{R}\mid\underbrace{\xi\Phi(x/\xi)}_{f}\leq\zeta\}=\{(\xi,x,\zeta)\in\mathbb{R}_{++}\times\mathcal{H}\times\mathbb{R}\mid\left(\frac{x}{\xi},\frac{\zeta}{\xi}\right)\in\operatorname{epi}\Phi\}$

/* $\Phi\left(\frac{x}{\xi}\right)\leq\frac{\zeta}{\xi}\Leftrightarrow\left(\frac{x}{\xi},\frac{\zeta}{\xi}\right)\in\operatorname{epi}\Phi$ *

$<+\infty$

$$\text{/* } \phi\left(\frac{x}{\xi}\right) \le \frac{\xi}{\xi} \underset{<+\infty}{} \leftrightarrow \left(\frac{x}{\xi}, \frac{\xi}{\xi}\right) \in \operatorname{epi}\phi \text{ */}$$

$= \{(\xi, x, \xi) \in \mathbb{R}_{++} \times \mathcal{H} \times \mathbb{R} \mid \left(\frac{x}{\xi}, \frac{\xi}{\xi}\right) \in \operatorname{epi}\phi\}$ /* let $\frac{x}{\xi} = y$, $\frac{\xi}{\xi} = \eta$, then, $\overset{\in \mathbb{R}_{++}}{(\xi, x, \xi)} = \xi(1, \frac{x}{\xi}, \frac{\xi}{\xi}) = \xi(1, y, \eta)$ */

$= \{\xi(1, y, \eta) \in \mathbb{R} \times \mathcal{H} \times \mathbb{R} \mid \xi \in \mathbb{R}_{++}, (y, \eta) \in \operatorname{epi}\phi\}$ & $\frac{x}{\xi} = y$, $\frac{\xi}{\xi} = \eta$ }

$= \underbrace{\xi}_{\in \mathbb{R}_{++}} \underbrace{\{(1, y, \eta) \in \mathbb{R} \times \mathcal{H} \times \mathbb{R} \mid (y, \eta) \in \operatorname{epi}\phi\}}_{C}$ & $\frac{x}{\xi} = y$, $\frac{\xi}{\xi} = \eta$ }

$= \mathbb{R}_{++} C$ { Recall : C : convex }

$= \operatorname{cone} C$ /* By Proposition 6.3. [C: subset of $\mathcal{H}$] $\Rightarrow$ $\operatorname{cone} C = \mathbb{R}_{++} C$ */

/* Recall Proposition 6.2.(iii)

[$C \subseteq \mathcal{H}$] $\operatorname{cone}(\operatorname{conv} C) = \operatorname{conv}(\operatorname{cone} C)$

had a consequence C: convex $\Rightarrow$ cone C: convex */

$\therefore$ $\operatorname{epi} f = \operatorname{cone} C$: convex

$\leftrightarrow$ f: convex function.

■

*** Proposition 8.26.**

[K: real Hilbert space

$F: \mathcal{H} \times K \to]-\infty, +\infty]$: convex]

$\underbrace{f: \mathcal{H} \to [-\infty, +\infty] : x \mapsto \inf F(x, K)}_{\text{marginal function}}$: convex

1/15/2016 8:58 PM

Proof :

$\forall x_1, x_2 \in \operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$

$\forall \kappa \in]0, 1[$

$\left.\begin{cases} f(x_1) = \inf F(x_1, K) \le F(x_1, \tilde{y}_1) & \forall \tilde{y}_1 \in K \\ f(x_2) = \inf F(x_2, K) \le F(x_2, \tilde{y}_2) & \forall \tilde{y}_2 \in K \end{cases}\right\}$ as $x_1 \in \operatorname{dom} f$, $f(x_1) = \inf F(x_1, K) < +\infty$, also, as by given $\forall \tilde{y}_1 \in K$ $F(x_1, \tilde{y}_1) > -\infty$

So, $f(x_1) = \text{finite}$ $\therefore$ $f(x_1) = \min F(x_1, K)$ $\exists y_1 \in K$ $y_1 = \underset{\tilde{y}_1 \in K}{\operatorname{argmin}} F(x_1, \tilde{y}_1) = \operatorname{argmin} F(x_1, K)$ $\therefore$ $f(x_1) = F(x_1, y_1)$

similarly, $\exists y_2 \in K$ $y_2 = \operatorname{argmin} F(x_2, K)$ $\therefore$ $f(x_2) = F(x_2, y_2)$

now,

$f(\kappa x_1 + (1-\kappa) x_2)$

$= \inf F(\underbrace{\kappa x_1 + (1-\kappa) x_2}_{\text{We can treat this as a column vector}}, K)$ K: real Hilbert space

$\le F(\kappa x_1 + (1-\kappa) x_2, \kappa y_1 + (1-\kappa) y_2)$ /* $\because$ $\kappa y_1 + (1-\kappa) y_2 \in K$ */

$\le \kappa F(x_1, y_1) + (1-\kappa) F(x_2, y_2)$ /* $\because$ $\kappa(x_1, y_1) + (1-\kappa)(x_2, y_2) = (\kappa x_1, \kappa y_1) + ((1-\kappa) x_2, (1-\kappa) y_2)$

$= (\kappa x_1 + (1-\kappa) x_2, \kappa y_1 + (1-\kappa) y_2)$ and F: convex */

$\le \kappa f(x_1) + (1-\kappa) f(x_2)$

now by letting $\xi_1 \downarrow f(x_1)$, $\xi_2 \downarrow f(x_2)$ we have.

$f(\kappa x_1 + (1-\kappa) x_2) \le \kappa f(x_1) + (1-\kappa) f(x_2) \overset{\text{def}}{\leftrightarrow} f$: convex ■

*** Theorem 8.29.**

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

$x_0 \in \operatorname{dom} f$]

(i) f: locally Lipschitz continuous near x_0 $\Leftrightarrow$

(ii) f: continuous at x_0 $\Leftrightarrow$

(iii) f: bounded on a neighborhood of x_0 $\Leftrightarrow$

(iv) f: bounded above on a neighborhood of x_0.

Moreover, if one of these conditions holds, then f: locally Lipschitz continuous on $\operatorname{int} \operatorname{dom} f$.

1/17/2016 10:58 AM

Proof: (i) means $\exists \rho > 0$ $\exists L > 0$ $\forall x, y \in B(x_0; \rho)$ $\|f(x) - f(y)\| \le L \|x - y\|$

(i) $\to$ (ii) $\to$ (iii) $\to$ (iv) : clear

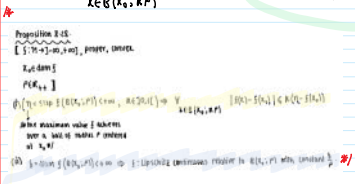
(iv) $\to$ (ii) : take $\rho \in \mathbb{R}_{++}$: $\eta = \sup f(B(x_0, \rho)) < +\infty$ /* this is allowed as f: bounded above on a neighborhood of x_0

$\Rightarrow \exists V_\xi(x_0)$ $\sup(f(V_\xi(x_0))) < +\infty$, so any $B(x_0; \rho) \subseteq V_\xi(x_0)$ will satisfy $\sup(f(B(x_0, \rho))) < +\infty$ */

$\kappa \in]0, 1[$

$x \in B(x_0; \kappa\rho)$

/*



$\forall x \in B(x_0; \kappa\rho)$ $|f(x) - f(x_0)| \le \kappa(\eta - f(x_0))$

/* recall, f: continuous at $x \overset{\text{def}}{\leftrightarrow} x_n \to x \Rightarrow f(x_n) \to f(x)$ */

$x \to x_0 \Leftrightarrow \kappa \to 0$, $x \in B(x_0; \kappa\rho)$

$\Rightarrow |f(x) - f(x_0)| \le \kappa(\underbrace{\eta}_{\sup f(B(x_0;\rho)) < +\infty} - f(x_0)) \to 0$ $< +\infty$, as $x_0 \in B(x_0, \rho)$!

$\Rightarrow |f(x) - f(x_0)| \to 0 \Rightarrow f(x) \to f(x_0)$ $\therefore$ f is continuous at x_0.

(iii) $\to$ (i) : follows from Proposition 8.28.(i)

Now we have shown (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv), (iv) $\Rightarrow$ (ii), (iii) $\Rightarrow$ (i)

(iii) → (i) : follows from Proposition 8.28.(i)

Now we have shown (i) ⇒ (ii) ⇒ (iii) ⇒ (iv), (iv) ⇒ (ii), (iii) ⇒ (i)

so, (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) : note that this infact establishes the following equivalence:

$$\underbrace{(i) \Leftrightarrow (ii) \Leftrightarrow (iii) \text{ and } (ii) \Leftrightarrow (iii) \Leftrightarrow (iv)}$$

$$\Rightarrow \quad (i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Leftrightarrow (iv)$$

Now let us prove that : f: locally Lipschitz continuous on int dom f. if one of (i)-(iv) holds (as they are equivalent)

Assume (iv) holds,

$$\exists_{\rho \in \mathbb{R}_{++}} \ \eta = \sup f(B(x_0;\rho)) < +\infty$$

then from (iv) ⇒ (i): f: locally Lipschitz continuous near x_0

take $\forall\, x \in \operatorname{int} \operatorname{dom} f \setminus \{x_0\}$, $\gamma \in \mathbb{R}_{++}$: $B(x;\gamma) \subseteq \operatorname{dom} f$.

set,

$$y = x_0 + \frac{1}{1-\alpha}(x - x_0), \quad \alpha = \frac{\gamma}{\gamma + \|x - x_0\|} \in]0,1[$$

then

$$y - x = x_0 + \frac{1}{1 - \frac{\gamma}{\gamma + \|x - x_0\|}}(x - x_0) - x$$

$$= \frac{\gamma + \|x - x_0\|}{\|x - x_0\|}(x - x_0) - (x - x_0)$$

$$= \left[\frac{\gamma + \|x - x_0\|}{\|x - x_0\|} - 1\right](x - x_0)$$

$$= \frac{\gamma}{\|x - x_0\|}(x - x_0) = \gamma \underbrace{\frac{x - x_0}{\|x - x_0\|}}_{\text{a vector with unit norm}}$$

$$\Rightarrow \|y - x\| = \gamma \underbrace{\left\|\frac{x - x_0}{\|x - x_0\|}\right\|}_{=1} = \gamma \Rightarrow y \in B(x;\gamma)$$

take, $z \in B(x;\alpha\rho)$

$$w = x_0 + \frac{z - x}{\alpha} \overset{?}{=} \frac{z - (1-\alpha)y}{\alpha}$$

$$\Leftrightarrow \quad x_0 + \cancel{\frac{z}{\alpha}} - \frac{x}{\alpha} \overset{?}{=} \cancel{\frac{z}{\alpha}} - \frac{(1-\alpha)y}{\alpha}$$

$$\Leftrightarrow \quad \frac{\alpha x_0 - x}{\cancel{\alpha}} \overset{?}{=} \frac{-(1-\alpha)y}{\cancel{\alpha}}$$

$$\Leftrightarrow \quad y \overset{?}{=} -\frac{\alpha x_0 - x}{1-\alpha} = \frac{x - \alpha x_0}{1-\alpha} = \frac{(1-\alpha)x_0 - x_0 + x}{1-\alpha} \overset{\checkmark}{=} x_0 + \frac{x - x_0}{1-\alpha}$$

$$w - x_0 = \frac{z - x}{\alpha}$$

$$\Rightarrow \|w - x_0\| = \frac{1}{\alpha}\|z - x\| \leq \rho \quad [\because z \in B(x;\alpha\rho) \Leftrightarrow \|z - x\| \leq \alpha\rho \Leftrightarrow \frac{\|z - x\|}{\alpha} \leq \rho]$$

$$\Rightarrow w \in B(x_0;\rho)$$

$$w = \frac{z - (1-\alpha)y}{\alpha}$$

$$\Leftrightarrow \quad z = \alpha w + (1-\alpha)y : \alpha \in]0,1[$$

now:

$$f(z) = f(\alpha w + (1-\alpha)y) \leq \alpha f(w) + (1-\alpha)f(y) \quad [\because f: \text{convex}]$$

now, $z \in B(x;\alpha\rho)$, $w \in B(x_0;\rho)$, $y \in B(x,\gamma)$

as $y \in B(x;\gamma) \subseteq \operatorname{dom} f \Rightarrow f(y) < +\infty$

$$\Rightarrow \quad f(z) \leq \alpha \underbrace{\sup_{w \in B(x_0;\rho)} f(w)}_{= \sup f(B(x_0;\rho)) = \eta} + (1-\alpha)\underbrace{f(y)}_{< +\infty \text{ [By construction]}} = \alpha\eta + (1-\alpha)f(y)$$

so $\forall_{z \in B(x;\alpha\rho)} \ f(z) < +\infty$

$$\Leftrightarrow \quad \sup(f(B(x;\alpha\rho))) < +\infty$$

$\Leftrightarrow$ f: bounded above on $B(x;\alpha\rho)$ so from (iv) ⇔ (iii) we have:

$x \in \operatorname{int} \operatorname{dom} f$

if one of (i)-(iv) holds, then f: locally Lipschitz continuous on int dom f. ∎

* Corollary 8.32.

[$\mathcal{H}$: finite dimensional,

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

C: nonempty, closed, bounded $\subseteq$ ri dom f /* ri $C = \{x \in C \mid \operatorname{cone}(C - x) = \operatorname{span}(C - x)\}$ */

] ⇒

f: Lipschitz continuous relative to C.

1/18/2016 3:59 PM Proof:

$z \in \operatorname{dom} f$

consider, $z - \operatorname{dom} f$ /* just z becomes the new origin in $z - \operatorname{dom} f$ */

construct

$\operatorname{span}(z - \operatorname{dom} f)$ /* span C: smallest linear subspace

$z \in \operatorname{dom} f$

consider, $z - \operatorname{dom} f$ /* just z becomes the new origin in $z - \operatorname{dom} f$ */

construct

$\operatorname{span}(z - \operatorname{dom} f)$ /* span C: smallest linear subspace containing C */

$C \subseteq \operatorname{ri} \operatorname{dom} f$

assume $C \subseteq \operatorname{int} \operatorname{dom} f$ without loss of generality /* need explanation */

// as a simple explanation, let us assume int dom f=nonempty, then ri dom f and int dom f will coincide, as dom f is convex set due to f being convex

/*

*Corollary 8.30.

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

one of the following holds

- f: bounded above on some neighborhood
- f: lower continuous,
- $\mathcal{H}$: finite-dimensional] $\Rightarrow$ $\operatorname{cont} f = \operatorname{int} \operatorname{dom} f$ /* $\operatorname{cont} f$: domain of continuity of a function f */

*/

now $\mathcal{H}$: finite-dimensional $\Rightarrow$ $\operatorname{cont} f = \operatorname{int} \operatorname{dom} f \leftrightarrow f$: continuous on $\operatorname{int} \operatorname{dom} f$

Take any $x_0 \in \operatorname{int} \operatorname{dom} f \subseteq \operatorname{dom} f$, then f: continuous at x_0

/*

* Theorem 8.29. f: proper $\nexists -\infty$ of $f(\mathcal{H})$, $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\} \neq \emptyset$

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

$x_0 \in \operatorname{dom} f$]

(i) f: locally Lipschitz continuous near x_0, $\Leftrightarrow$

(ii) f: continuous at x_0, $\Leftrightarrow$

(iii) f: bounded on a neighborhood of x_0, $\Leftrightarrow$

(iv) f: bounded above on a neighborhood of x_0.

Moreover, if one of these conditions holds, then f: locally Lipschitz continuous on $\operatorname{int} \operatorname{dom} f$ */

Now, as $C \subseteq \operatorname{int} \operatorname{dom} f$, f: locally Lipschitz continuous on $\operatorname{int} \operatorname{dom} f$

$\Rightarrow$ f: locally Lipschitz continuous on $C \subseteq \operatorname{int} \operatorname{dom} f$

C: closed, bounded $\Rightarrow$ C: compact /* on a finite dimensional space, closed and bounded sets are compact */

/* f: locally Lipschitz continuous on a compact set $\Rightarrow$
f: Lipschitz continuous on a compact set */ see proofs below:

$\therefore$ f: Lipschitz continuous on C ∎

Remark 10.10. If F is locally Lipschitz, then F is Lipschitz continuous on compact sets. Actually, for every compact set K there exists a neighborhood U of K such that F is Lipschitz on U.

Proof (by contradiction). Assume that F is locally Lipschitz, but that there is some compact set K such that F is not Lipschitz continuous on any neighborhood of K. In particular, F is continuous on X. Let $U_n = \{x \in X; d(x, K) < 1/n\}$. Let (Λ_n) be a sequence of positive numbers with $\Lambda_n \to \infty$ as $n \to \infty$. Since F is not Lipschitz on U_n, for each $n \in \mathbb{N}$ there exist $x_n, y_n \in U_n$ such that

$$\|F(x_n) - F(y_n)\| > \Lambda_n \|x_n - y_n\|, \qquad n \in \mathbb{N}.$$

For each $n \in \mathbb{N}$, there exists some $w_n, z_n \in K$ with $\|w_n - x_n\| < 1/n$ and $\|z_n - y_n\| < 1/n$. Since K is compact, after choosing subsequences, $w_n \to x$ and $z_n \to z$ for some $x, z \in K$. Then $x_n \to x$ and $y_n \to z$. Since F is continuous, $F(x_n) \to F(x)$ and $F(y_n) \to F(z)$. The inequality above implies that $x_n - y_n \to 0$ and so $x_n, y_n \to x$. Since F is locally Lipschitz, there exist some $\epsilon > 0$ and $\Lambda > 0$ such that

$$\|F(z) - F(y)\| \le \Lambda \|z - y\| \qquad \text{if } z, y \in K, \|z - x\| < \epsilon, \|y - z\| < \epsilon.$$

Choosing subsequences again, we can arrange that $\|x_n - x\| < \epsilon$ and $\|y_n - x\| < \epsilon$ for all $n \in \mathbb{N}$. This implies

$$\Lambda \ge \frac{\|F(x_n) - F(y_n)\|}{\|x_n - y_n\|} \ge \Lambda_n \to \infty,$$

a contradiction. □

Part 1

Tuesday, January 19, 2016 1:01 PM

Theorem 9.1.

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

(i) f: weakly sequentially lower semicontinuous $\Leftrightarrow$

(ii) f: sequentially lower semi continuous $\Leftrightarrow$

(iii) f: lower semicontinuous $\Leftrightarrow$

(iv) f: weakly lower semicontinuous.

1/19/2016 1:51 PM Proof:

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

$\Leftrightarrow$ epi f : convex [By definition]

Now recall:

Lemma 1.24

[X : Hausdorff space,
$f : X \to [-\infty, \infty]$]

f : lower semicontinuous $\Leftrightarrow$ epi f : closed in $X \times \mathbb{R}$ $\Leftrightarrow$ $\forall_{\xi \in \mathbb{R}}$ lev$_{\leq \xi} f$: closed in X

/* this Lemma tells us why in CCP function, f has to be closed, this essentially means that f : lower semicontinuous */

Lemma 1.35

[X : Hausdorff space
$f : X \to [-\infty, +\infty]$]

/* i.e., sequentially lower semicontinuous at every point at x */

(i) f: sequentially lower semicontinuous $\Leftrightarrow$

(ii) epi f: sequentially closed $\Leftrightarrow$

(iii) $\forall_{\xi \in \mathbb{R}}$ lev$_{\leq \xi} f$: sequentially closed in X

Theorem 3.32 /* this is a very important theorem which says that for a convex set all the different types of closedness coincides */

[C : convex subset of $\mathcal{H}$]

C : weakly sequentially closed $\Leftrightarrow$ C : sequentially closed $\Leftrightarrow$ C : closed $\Leftrightarrow$ C : weakly closed

/* C : weakly sequentially closed $\overset{def}{\Leftrightarrow}$ weak limit of every weakly convergent sequence in C is also in C

C : sequentially closed $\overset{def}{\Leftrightarrow}$ strong " " " strongly " " " C " " " C

C : closed $\overset{def}{\Leftrightarrow}$ strong " " " " " net " C " " " C

C : weakly closed $\overset{def}{\Leftrightarrow}$ weak " " " weakly " net " C " " " C */

f : lower semicontinuous $\Leftrightarrow$ epi f : closed, as epi f : convex we have

epi f : sequentially closed $\Leftrightarrow$ f : sequentially lower semicontinuous /* using Lemma 1.35 */

$\Updownarrow$

epi f : closed $\Leftrightarrow$ f : lower semicontinuous /* using Lemma 1.24 */

$\Updownarrow$

epi f : weakly closed $\Leftrightarrow$ f : weakly lower semicontinuous /* using Lemma 1.24 */

$\Updownarrow$

epi f : weakly sequentially closed $\Leftrightarrow$ f : weakly sequentially lower semicontinuous /* using Lemma 1.35 */

Proposition 9.3.

[$(f_i)_{i\in I}$: family in $\Gamma(\mathcal{H})$]

$\sup_{i\in I} f_i \in \Gamma(\mathcal{H})$

1/19/2016 2:03 PM Proof:

Recall:

/*

* Lemma 1.26.

[$\mathcal{X}$: Hausdorff space, $(f_i)_{i\in I}$: family of lower semicontinuous functions from $\mathcal{X}$ to $[-\infty,+\infty]$]

$\Rightarrow$

• $\sup_{i\in I} f_i$: lower semicontinuous

* Proposition 8.14. $(f_i)_{i\in I}$: family of convex functions from $\mathcal{H}$ to $[-\infty,+\infty]$ $\Rightarrow$ $\sup_{i\in I} f_i$: convex */

from them we have $\sup_{i\in I} f_i \in \Gamma(\mathcal{H})$ ∎

* Proposition 9.8.

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

(i) $\check{f}$: largest lower semicontinuous convex function majorized by f. // $\check{f} = \sup\{g \in \Gamma(\mathcal{H}) \mid g \le f\}$

(ii) $\forall_{x\in\mathcal{H}}$ $\check{f} = \underline{\lim}_{y\to x} \check{f}(y)$

$f : \mathcal{X} \to [-\infty,+\infty]$ a data structure containing all the values of f on V

$\underline{\lim}_{y\to x} f(y) = \sup_{V\in\mathcal{V}(x)} \inf f(V) = \left(V \in \mathcal{V}(x) \,//\, f(y) \,//\, \inf \boxed{Z} \right) \,//\, \sup_{V\in\mathcal{V}(x)} (y)$

(iii) epi $\check{f}$: closed, convex

(iv) conv dom $f \subseteq$ dom $\check{f} \subseteq \overline{\text{conv}}$ dom f

Proof:

(i) $\check{f} = \sup\{g \in \Gamma(\mathcal{H}) \mid g \le f\}$

$= \sup_{g\in\Gamma(\mathcal{H}) : g\le f} g \in \Gamma(\mathcal{H})$ /*

Proposition 9.3. /* $\Gamma(\mathcal{H})$: closed under sup */

[$(f_i)_{i\in I}$: family in $\Gamma(\mathcal{H})$] $\sup_{i\in I} f_i \in \Gamma(\mathcal{H})$ */

(ii) From (i) we have

$\check{f}$: largest lower semicontinuous convex function

* Lemma 1.31.

[$\mathcal{X}$: Hausdorff space,

$f : \mathcal{X} \to [-\infty,+\infty]$] $\Rightarrow$

(i) $\bar{f}$: largest lower semicontinuous function majorized by f

(ii) epi $\bar{f}$: closed

(iii) dom $f \subseteq$ dom $\bar{f} \subseteq \overline{\text{dom}}\, f$, (iv) $\forall_{x\in\mathcal{X}}$ $\bar{f}(x) = \underline{\lim}_{y\to x} f(y)$

(v) $x \in \mathcal{X} \Rightarrow$ (f : lower semicontinuous at f $\Leftrightarrow$ $\bar{f}(x) = f(x)$)

(vi) epi $\bar{f} = \overline{\text{epi}}\, f$.

$\forall_{x\in\mathcal{H}}$ $\check{f}(x) = \underline{\lim}_{y\to x} f(y)$

(iii) $\check{f}$: convex $\Rightarrow$ epi $\check{f}$: convex

$\check{f}$: lower semicontinuous $\Rightarrow$ epi $\check{f}$: closed

$\Rightarrow$ epi $\check{f}$: closed, convex.

Lemma 1.24.

[$\mathcal{X}$: Hausdorff space,

$f : \mathcal{X} \to [-\infty,\infty]$]

f : lower semicontinuous $\Leftrightarrow$ epi f : closed in $\mathcal{X}\times\mathbb{R}$ $\Leftrightarrow$ $\forall_{\xi\in\mathbb{R}}$ lev$_{\le\xi}$ f : closed in $\mathcal{X}$

Lemma 1.24.

[X : Hausdorff space, $f: X \to [-\infty, \infty]$]

f : lower semicontinuous $\Leftrightarrow$ epi f : closed in $X \times \mathbb{R}$ $\Leftrightarrow$ $\forall_{\xi \in \mathbb{R}}$ $\text{lev}_{\leq \xi} f$: closed in X

/* this lemma tells us why in CCP function, f has to be closed, this essentially means that f : lower semicontinuous */

(iv) By definition,

$$\check{f} = \sup\{g \in \Gamma(\mathcal{H}) \mid g \leq f\} = \sup_{\substack{g \in \Gamma(\mathcal{H}) \\ g \leq f}} g$$

so, $\check{f} \leq f$ [$\because$ $\check{f}$ is a feasible solution]

also $\check{f} \in \Gamma(\mathcal{H}) \Rightarrow \check{f}$: convex $\Rightarrow$ dom $\check{f}$: convex /* A convex function f has convex domain dom $f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$ */

$\therefore$ conv dom $\check{f}$ = dom $\check{f}$

suppose, $x \in \text{dom} f \Leftrightarrow f(x) < +\infty$

$\Rightarrow \check{f}(x) \leq f(x) < +\infty$

$\Leftrightarrow x \in \text{dom} \check{f}$

$\Rightarrow$ dom $f \subseteq$ dom $\check{f}$

$\Rightarrow$ conv dom $f \subseteq$ conv dom $\check{f}$

= dom $\check{f}$

/*

Fact 1.12 (set extension operator preserves inclusion)

[S : some set extension operator (e.g. span, $\overline{\text{span}}$, cone, $\overline{\text{cone}}$, conv, $\overline{\text{conv}}$ etc) that is the smallest set based on some property containing the operand set]

$A \subseteq B \Rightarrow S(A) \subseteq S(B)$

*/

now set $C = \overline{\text{conv}}$ dom f

$$g: \mathcal{H} \to [-\infty, +\infty] : x \mapsto \begin{cases} \check{f}(x), & \text{if } x \in C \\ +\infty, & x \notin C \end{cases}$$

$$g(x) = \begin{cases} \check{f}(x), & \text{if } x \in C \\ +\infty, & \text{if } x \notin C \end{cases} = \check{f}(x) + \iota_C(x)$$

epi $\check{f} = \{(x,t) \in \mathcal{H} \times \mathbb{R} \mid \check{f}(x) \leq t\}$: convex /* recall that in epi f $(x,t) \in \mathcal{H} \times \mathbb{R}$, so $t \neq \pm\infty$ */

epi $g = \{(x,t) \in \mathcal{H} \times \mathbb{R} \mid g(x) = \check{f}(x) + \iota_C(x) \leq t\}$

$= \underbrace{\{(x,t) \in \mathcal{H} \times \mathbb{R} \mid \check{f}(x) \leq t\}}_{\text{epi } \check{f}} \cap (C \times \mathbb{R})$ /* otherwise $g(x) = \infty$ */

$= \underbrace{\text{epi } \check{f}}_{\substack{\text{convex,} \\ \text{closed}}} \cap \underbrace{(C \times \mathbb{R})}_{\text{closed, convex}}$: closed, convex

/* lower semicontinuous function has closed epigraph */

epi g : closed $\Leftrightarrow$ g : lower semicontinuous /* " */

epi g : convex $\Leftrightarrow$ g : convex /* by definition */

$\therefore g \in \Gamma(\mathcal{H})$

$\forall_x$

$\begin{cases} x \in C \Rightarrow g(x) = \check{f}(x) \leq f(x), \text{ now if } x \notin C \Rightarrow g(x) = \check{f}(x) + \iota_C(x) = +\infty, \text{ also, dom } f \subseteq \overline{\text{conv}} \text{ dom } f = C, \text{ so } x \notin C \Rightarrow x \notin \text{dom} f \Leftrightarrow f(x) = +\infty \\ \therefore x \notin C \Rightarrow g(x) = f(x) = +\infty \end{cases}$

$\therefore \forall_{x \in \mathcal{H}}$ $g(x) \leq f(x)$

$\Leftrightarrow g \leq f$

and as $g \in \Gamma(\mathcal{H})$, $\check{f} = \sup\{\tilde{g} \in \Gamma(\mathcal{H}) \mid \tilde{g} \leq f\}$ $\quad \check{f} \geq g$

but $\check{f} \leq f$ $\quad \forall_{\tilde{g} \in \Gamma(\mathcal{H}) : \tilde{g} \leq f}$ $\quad \check{f} \geq \tilde{g}$ $\quad$ set $\tilde{g} := g$

$\therefore g \leq \check{f}$

$\Rightarrow$ dom $g \supseteq$ dom $\check{f}$ /* as, $\forall_{x \in \text{dom} \check{f}} \Leftrightarrow \check{f}(x) < +\infty$ $\quad g(x) \leq f(x) < +\infty \Rightarrow g(x) < +\infty$ $\quad \therefore$ dom $\check{f} \subseteq$ dom g */

again $\forall_{x \in \text{dom} g}$ $g(x) < +\infty \Leftrightarrow f(x) + \iota_C(x) < \infty \Rightarrow x \in C$ /* otherwise $\iota_C(x) = +\infty$ */

$\Rightarrow \operatorname{dom} g \supseteq \operatorname{dom} \check{f}$ /* as, $\forall_{x \in \operatorname{dom} \check{f}} \leftrightarrow \check{f}(x) < +\infty$ $\quad g(x) \leq f(x) < +\infty \Rightarrow g(x) < +\infty \quad \therefore \operatorname{dom} \check{f} \subseteq \operatorname{dom} g$ */

again $\forall_{x \in \operatorname{dom} g}$ $\quad \underset{\substack{\| \\ f(x)+L_C(x)}}{g(x)} < +\infty \leftrightarrow f(x) + L_C(x) < \infty \Rightarrow x \in C$ /* otherwise $L_C(x) = +\infty$ */

$\therefore \operatorname{dom} g \subseteq C$

$\operatorname{dom} \check{f} \subseteq \operatorname{dom} g \subseteq C = \overline{\operatorname{conv}} \operatorname{dom} f$

$\therefore \operatorname{conv} \operatorname{dom} f \subseteq \operatorname{dom} \check{f} \subseteq \overline{\operatorname{conv}} \operatorname{dom} f.$

$\blacksquare$

Part 2

Saturday, January 23, 2016 11:35 AM

Proposition 9.14.

[$f \in \Gamma_0(\mathcal{H})$, $x \in \mathcal{H}$, $y \in \operatorname{dom} f$;

$\forall_{\alpha \in]0,1[} \; x_\alpha := (1-\alpha)x + \alpha y$]

$\lim_{\alpha \downarrow 0} f(x_\alpha) = f(x)$

Proof :

* Definition 1.21.

[X: Hausdorff space, $f : X \to [-\infty, +\infty]$, $x \in X$]

f: lower semicontinuous at x $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x_\alpha)_{\alpha \in A} : \text{net in } X,\, x_\alpha \to x} \; \underline{\lim} f(x_\alpha) \geq f(x)$ ※ definition in terms of a net

$\overset{\text{def}}{\Leftrightarrow} \forall_{\xi \in]-\infty, f(x)[} \; \exists_{V \in \mathcal{V}(x)} \; f(V) \subseteq]\xi, +\infty]$

$\overset{\text{def}}{\Leftrightarrow} \forall_{\xi : \xi < f(x)} \; \exists_{V \in \mathcal{V}(x)} \; \forall_{f(y) \in f(V)} \; \xi < f(y)$ (see Thm 1.6)

---

first note that $(x_\alpha)_{\alpha \in]0,1[}$: net, now if we partially order them as: $x_{1-\ldots}, \ldots, x_{0_+}$, i.e., $\lim x_\alpha = \lim_{\alpha \downarrow 0} ((1-\alpha)x + \alpha y) = x$

then f: lower semicontinuous at $x \overset{\text{def}}{\Leftrightarrow}$ $f(x) \leq \underline{\lim} f(x_\alpha) \underset{\text{as } \alpha \downarrow 0}{=} \underline{\lim}_{\alpha \downarrow 0} f(x_\alpha)$

$= \underline{\lim}_{\alpha \downarrow 0} f((1-\alpha)x + \alpha y)$ /* now f: convex */

$\leq \underline{\lim}_{\alpha \downarrow 0} (1-\alpha) f(x) + \alpha f(y)$ /* now note that $\lim_{\alpha \downarrow 0} (1-\alpha) f(x) = f(x)$ $\lim_{\alpha \downarrow 0} (1-\alpha) = f(x) = \underline{\lim}_{\alpha \downarrow 0} (1-\alpha) f(x)$

$\lim_{\alpha \downarrow 0} \underbrace{\alpha f(y)}_{\text{finite}} = 0 = \lim_{\alpha \downarrow 0} \alpha f(y)$ /* as $y \in \operatorname{dom} f$, f: proper $\Leftrightarrow -\infty \notin f(\mathcal{H})$ $\Downarrow$ $f(y) < +\infty$ $\rightarrow$ $f(y) =$ finite */

$= f(x)$

So, the inbetween terms would collapse.

$\therefore \underline{\lim}_{\alpha \downarrow 0} f(x_\alpha) = f(x)$

$\therefore \underline{\lim}_{\alpha \downarrow 0} (1-\alpha) f(x) + \lim_{\alpha \downarrow 0} \alpha f(y)$

$= \lim_{\alpha \downarrow 0} (1-\alpha) f(x) + \alpha f(y) = f(x)$

$\therefore \underline{\lim}_{\alpha \downarrow 0} (1-\alpha) f(x) + \alpha f(y) = \lim_{\alpha \downarrow 0} (1-\alpha) f(x) + \alpha f(y) = f(x)$ */

∎

* Proposition 9.17.

[$f \in \Gamma_0(\mathcal{H})$, $(x, \xi) \in \mathcal{H} \times \mathbb{R}$, $(p, \pi) \in \mathcal{H} \times \mathbb{R}$]

$(p, \pi) = P_{\operatorname{epi} f}(x, \xi) \Leftrightarrow \begin{cases} \bullet \; \max\{\xi, f(p)\} \leq \pi \\ \bullet \; \forall_{y \in \operatorname{dom} f} \; \langle y - p \mid x - p \rangle + (f(y) - \pi)(\xi - \pi) \leq 0 \end{cases}$

1/23/2016 3:56 PM Proof:

$f \in \Gamma_0(\mathcal{H}) \Rightarrow \begin{cases} f: \text{convex} \Rightarrow \operatorname{epi} f: \text{convex} \quad \text{/* By def */} \\ f: \text{proper} \Rightarrow \operatorname{epi} f: \text{nonempty} \quad \text{/* } f: \text{proper} \overset{\text{def}}{\Leftrightarrow} -\infty \notin f(\mathcal{H}), \operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\} \neq \emptyset \text{ (} \exists_{(x,t) \in \mathcal{H} \times \mathbb{R}} \; f(x) \leq t \text{)}, \operatorname{epi} f = \{(x,t) \in \mathcal{H} \times \mathbb{R} \mid f(x) \leq t\} \neq \emptyset \text{ */} \\ f: \text{lower semicontinuous} \Rightarrow \operatorname{epi} f: \text{closed} \quad \text{/* Lemma 1.24 */} \end{cases}$

$\Rightarrow \operatorname{epi} f$: convex, nonempty, closed

$(p, \pi) := P_{\operatorname{epi} f}(x, \xi)$ well defined

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

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

now recall

/*

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

($\cancel{C}$ $\operatorname{epi} f$: nonempty closed convex subset of $\mathcal{H}$) $\Rightarrow \begin{cases} \bullet \; \cancel{C}\ \operatorname{epi} f: \text{Chebyshev set, i.e., every point in } \mathcal{H} \text{ has exactly one projection on } C \\ \bullet \; \forall_{\cancel{x \in \mathcal{H}}\ (x,\xi) \in \mathcal{H} \times \mathbb{R}} \; (\cancel{p = P_C(x)}\ (p,\pi) = P_{\operatorname{epi} f}(x,\xi) \Leftrightarrow (\cancel{p \in C}\ (p,\pi) \in \operatorname{epi} f, \cancel{\forall_{y \in C} \; \langle y - p \mid x - p \rangle \leq 0})) \end{cases}$ */

$(p, \pi) \in \operatorname{epi} f \Leftrightarrow f(p) \leq \pi$

$\forall_{(y,\eta) \in \operatorname{epi} f} \; \underbrace{\langle (y,\eta) - (p,\pi) \mid (x,\xi) - (p,\pi) \rangle \leq 0}_{\langle (y-p, \eta-\pi) \mid (x-p, \xi-\pi) \rangle = \langle y-p \mid x-p \rangle + (\eta-\pi)(\xi-\pi) \leq 0}$ [$\because x^T y = \sum_{i=1}^{n} x_i y_i = (x_1, \ldots, x_{n-1})^T (y_1, \ldots, y_{n-1}) + x_n y_n$ */

/* now, $\forall_y \forall_\eta \; (y,\eta) \in \operatorname{epi} f \Leftrightarrow \forall_{y \in \operatorname{dom} f} \forall_{\eta \in \mathbb{R}} \; f(y) \leq \eta$ /* $\forall_\eta \exists_{\lambda \in \mathbb{R}_+} \; \eta = f(y) + \lambda$ */ $\Leftrightarrow \forall_{y \in \operatorname{dom} f} \forall_{\lambda \in \mathbb{R}_+} \; (y, f(y) + \lambda) \in \operatorname{epi} f$ */

so, put $\eta = f(y) + \lambda$

$\Leftrightarrow \forall_{y \in \operatorname{dom} f} \forall_{\lambda \in \mathbb{R}_+} \; \langle y - p \mid x - p \rangle + (f(y) + \lambda - \pi)(\xi - \pi) \leq 0$

so if $\lambda \uparrow +\infty$ then we must have $\xi \leq \pi$ (otherwise nothing can stop the L.H.S > 0)

$\boxed{\max\{\xi, f(p)\} \leq \pi}$

$$\boxed{\max\{\xi, f(p)\} \leqslant \pi}$$

$\lambda := 0$ yields:

$$\boxed{\forall_{y \in \operatorname{dom} f} \quad \langle y-p | x-p \rangle + (f(y)-\pi)(\xi-\pi) \leqslant 0}$$

■

* Theorem 9.19.

[[$f \in \Gamma_0(\mathcal{H})$]]

f: possesses a continuous affine minorant.

1/23/2016 4:11 PM

Proof: Set,

$x \in \operatorname{dom} f \leftrightarrow f(x) < +\infty$

$\xi \in]-\infty, f(x)[\subseteq \mathbb{R} \Rightarrow \xi < f(x) \Rightarrow (x,\xi) \notin \operatorname{epi} f = \{(\tilde{x},\tilde{\xi}) \in \mathcal{H} \times \mathbb{R} \mid \tilde{\xi} \geqslant f(\tilde{x})\}$

$(p,\pi) = P_{\operatorname{epi} f}(x,\xi)$, as $(p,\pi) \in \operatorname{epi} f = \{(\tilde{x},\tilde{\xi}) \in \mathcal{H} \times \mathbb{R} \mid \tilde{\xi} \geqslant f(\tilde{x})\} \Rightarrow p \in \mathcal{H}$, $\pi \in \mathbb{R}$, $f(p) \leqslant \pi$

* recall:

* Proposition 9.18.

[$f \in \Gamma_0(\mathcal{H})$, $x \in \operatorname{dom} f$, $\xi \in]-\infty, f(x)[$, $(p,\pi) \in \mathcal{H} \times \mathbb{R}$]

$$(p,\pi) = P_{\operatorname{epi} f}(x,\xi) \Leftrightarrow \begin{cases} \bullet \ \xi < f(p) = \pi \\ \bullet \ \underbrace{\forall_{y \in \operatorname{dom} f} \ \langle y-p | x-p \rangle \leqslant (f(y)-f(p))\overbrace{(f(p)-\xi)}^{>0}}_{} \end{cases}$$ */

$f(p) = \pi \in \mathbb{R} \to p \in \operatorname{dom} f$

$$\langle y-p | \underbrace{x-p}_{= u} \rangle \leqslant f(y) - f(p) \leftrightarrow \langle y-p|u\rangle + f(p) \leqslant f(y) \leftrightarrow f(p) - \langle p|u\rangle \leqslant f(y) - \langle y|u\rangle$$

* u is a constant w.r.t variable y */

as, $f(p) = \pi$: finite
and $\langle p|u\rangle$: finite $\Big\}$ $\underbrace{f(p) - \langle p|u\rangle}_{\xi} = \text{finite} > -\infty$

$\therefore \forall_{y \in \operatorname{dom} f} \quad -\infty < \xi \leqslant f(y) - \langle y|u\rangle$

and also, $\forall_{y \in \mathcal{H} \setminus \operatorname{dom} f} \quad \underset{+\infty}{f(y)} - \underbrace{\langle y|u\rangle}_{\text{finite}} = +\infty > -\infty$

$\therefore \forall_{y \in \mathcal{H}} \quad f(y) - \langle y|u\rangle > -\infty$

$\leftrightarrow \quad f - \langle \cdot | u \rangle$: bounded below

$\overset{\text{def}}{\Leftrightarrow} \quad f$: has a continuous affine minorant

■

Corollary 9.20.

[$f \in \Gamma_0(\mathcal{H})$]

$\forall_{\text{nonempty bounded subset of } \mathcal{H}}$ f: bounded below on that set

1/23/2016 4:36 PM

Proof:

C: nonempty bounded subset of $\mathcal{H}$

$\beta = \sup_{x \in C} \|x\| \in \mathbb{R}_+ \leftrightarrow \forall_{x \in C} \quad \|x\| \leqslant \beta$

As by Theorem 9.19. [[$f \in \Gamma_0(\mathcal{H})$]] $\exists_{g:\ \text{continuous affine function}}$ $g \leqslant f$

say $g = \langle \cdot | u \rangle + \eta$

now, $\forall_{y \in \mathcal{H}} \quad g(y) \leqslant f(y)$

$$\begin{aligned} \Rightarrow \forall_{x \in C \subseteq \mathcal{H}} \quad f(x) &\geqslant g(x) \\ &= \langle x|u\rangle + \eta \\ &\geqslant -|\langle x|u\rangle| + \eta \\ &\geqslant -\|x\|\,\|u\| + \eta \qquad \text{/* Cauchy-Schwarz: } |\langle x|u\rangle| \leqslant \|x\|\,\|u\| \\ &\geqslant \underbrace{-\beta}_{\text{finite}}\,\underbrace{\|u\|}_{\text{finite}} + \underbrace{\eta}_{\text{finite}} \qquad \leftrightarrow -|\langle x|u\rangle| \geqslant -\|x\|\,\|u\| \ \text{*/} \\ &> -\infty \end{aligned}$$

$\therefore \forall_{x \in C} \quad f(x) > -\infty$

■

* Proposition 9.26.

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

* Proposition 9.26.

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

dom g: open. g: continuous on dom g

$$f: \mathcal{H} \to]-\infty, +\infty] : x \mapsto \begin{cases} g(x), & x \in \operatorname{dom} g \\ \underline{\lim}_{y \to x} g(y), & x \in \operatorname{bdry} \operatorname{dom} g \\ +\infty, & x \in \mathcal{H} \setminus \overline{\operatorname{dom}} g \end{cases}$$

]

$\Rightarrow$

$f = \breve{g}$. $f \in \Gamma_0(\mathcal{H})$.

1/25/2016 2:50 PM

Proof: Part 1: We will show that, $\forall_{x \in \operatorname{dom} g}$ $f = \breve{g}$

$C = \operatorname{dom} g$

Recall:

* Proposition 9.8. ★★

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

(i) $\breve{f}$: largest lower semicontinuous convex function majorized by f.

(ii) $\forall_{x \in \mathcal{H}}$ $\breve{f} = \underline{\lim}_{y \to x} f(y)$ /* $\underline{\lim}_{y \to x} f(y) = (\forall \in \mathcal{V}(x) \,//\, f(x) \,//\, \inf f \boxed{\mathcal{A}}) \,//\, \sup_{V \in \mathcal{V}(x)} \boxed{\mathcal{B}}$ */

(iii) epi $\breve{f}$: closed, convex

(iv) ~~conv dom $f \subseteq$ dom $\breve{f} \subseteq$ conv dom f~~

$\breve{g} = \sup\{\tilde{g} \in \Gamma(\mathcal{H}) \mid \tilde{g} \leq g\}$, $\breve{g} \in \Gamma(\mathcal{H})$

$\Rightarrow \breve{g} \leq g$ /* recall: $f \geq g \Rightarrow \operatorname{dom} f \subseteq \operatorname{dom} g$ */

$\Rightarrow \operatorname{dom} \breve{g} \supseteq \operatorname{dom} g = C$

as g: convex $\Rightarrow$ dom g: convex /* proposition 8.2 */

$\Rightarrow$ conv dom g = dom g, $\overline{\operatorname{conv}} \underbrace{\operatorname{dom} g}_{\text{convex}} = \overline{\operatorname{dom}} g$

$C = \operatorname{dom} g \subseteq \operatorname{dom} \breve{g} \subseteq \overline{\operatorname{dom}} g = \overline{C}$ [1/25/2016 2:15 PM]

Assume. $x \in C = \operatorname{dom} g = \{\tilde{x} \in \mathcal{H} \mid g(\tilde{x}) < +\infty\} \Rightarrow g(x) < +\infty$

then

$\underbrace{\breve{g}(x) \leq g(x)}_{(\because \breve{g} \leq g)} < +\infty$

/* recall Theorem 9.9. [$f: \mathcal{H} \to [-\infty, +\infty]$] $\Rightarrow$ epi $\breve{f} = \overline{\operatorname{conv}}$ epi f. */

now, g: convex $\Rightarrow$ epi g: convex by definition.

using Theorem 9.9. epi $\breve{g} = \overline{\operatorname{conv}}$ epi g = $\overline{\operatorname{epi}}\, g$ [2017-02-08 9:24 AM]

/* $(x, t) \in \operatorname{epi} \breve{g} \leftrightarrow \breve{g}(x) \leq t \in \mathbb{R}$

$(x, \breve{g}(x)) \in \operatorname{epi} \breve{g} \Leftrightarrow \breve{g}(x) \leq \breve{g}(x) < +\infty$ */

$\therefore (x, \breve{g}(x)) \in \operatorname{epi} \breve{g} = \overline{\operatorname{epi}}\, g$

recall that:

* Lemma 1.10. /* This is a basic useful. */
[C: subset of a Hausdorff space X
$x \in X$]
$x \in \overline{C} \Leftrightarrow \exists_{(x_\alpha)_{\alpha \in A}: \text{net in } C} \; x_\alpha \to x$.

Theorem 3.32. /* this is a very important theorem which says that for a convex set all the different types of closedness coincides w/
[C: convex subset of $\mathcal{H}$]
C: weakly sequentially closed $\Leftrightarrow$ C: sequentially closed $\Leftrightarrow$ C: closed $\Leftrightarrow$ C: weakly closed

[C: convex] $x \in \overline{C} \Leftrightarrow \exists_{(x_n)_{n \in \mathbb{N}}: \text{sequence in } C} \; x_n \to x$

as, $(x, \breve{g}(x)) \in \overline{\operatorname{epi}}\, g$

$\Rightarrow$

$\exists_{(x_n, \xi_n)_{n \in \mathbb{N}}: \text{sequence in epi } g}$ $(x_n, \xi_n) \to (x, \breve{g}(x))$ [1/25/2016 2:37 PM]

$\Rightarrow$ $x_n \to x$,

$\xi_n \to \breve{g}(x)$

$\Rightarrow$ $\breve{g}(x) = \lim \xi_n = \underline{\lim}\, \xi_n$ /* When the limit exists the lim sup, $\overline{\lim}$ and lim inf, $\underline{\lim}$ are same */ [1/25/2016 2:44 PM]

$\geq \underline{\lim}\, g(x_n)$ /* $\because (x_n, \xi_n) \in \operatorname{epi} g$ $\forall_{n \in \mathbb{N}}$

$\Leftrightarrow g(x_n) \leq \xi_n < +\infty$ $\forall_{n \in \mathbb{N}}$

$\Leftrightarrow \underline{\lim}\, g(x_n) \leq \underline{\lim}\, \xi_n$ */

$\geq \underline{\lim}\, \breve{g}(x_n)$ [$\because g \geq \breve{g}$ by definition]

$\geq \breve{g}(x)$ /* $\because f$: lower semicontinuous at x $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x_\alpha)_{\alpha \in A}: \text{net in } X: x_\alpha \to x}$ $\underline{\lim}\, f(x_\alpha) \geq f(x)$

as $\breve{g}$: lower semicontinuous, and $x_n \to x \Rightarrow \underline{\lim}\, \breve{g}(x_n) \geq \breve{g}(x)$ */

So, the inequality system will collapse. $\lim g(x_n) = \lim \breve{g}(x_n) = \breve{g}(x)$

as $\breve{g}$: lower semicontinuous, and $x_n \to x \Rightarrow \underline{\lim}\, \breve{g}(x_n) \geq \breve{g}(x)$ */

So, the inequality system will collapse $\therefore \underline{\lim}\, g(x_n) = \underline{\lim}\, \breve{g}(x_n) = \breve{g}(x)$

So, for $x \in C = \operatorname{dom} g$

$$\underbrace{f(x) = g(x)}_{\text{by given}} = \lim g(x_n) \quad /* \, x_n \to x, \ g: \text{continuous on } \operatorname{dom} g = C \text{ by given, by def}: g(x_n) = g(x) \, */$$

$$= \underline{\lim}\, g(x_n)$$

$$= \breve{g}(x) \qquad [\text{Part 1 proved}]$$

Part 2: We will show,

$\forall (x \in \mathcal{H} \setminus \bar{C} = \mathcal{H} \setminus \overline{\operatorname{dom} g}) \quad f(x) = \breve{g}(x) = +\infty$

$x \in \mathcal{H} \setminus \bar{C}$

[1/25/2016 2:15 PM] says:

$$\operatorname{dom} g \subseteq \operatorname{dom} \breve{g} \subseteq \overline{\operatorname{dom} g} = \bar{C}$$

$$\operatorname{dom} \breve{g} \subseteq \overline{\operatorname{dom} g}$$

$$\Rightarrow \mathcal{H} \setminus \operatorname{dom} \breve{g} \supseteq \mathcal{H} \setminus \overline{\operatorname{dom} g} = \mathcal{H} \setminus \bar{C}$$

$$\left.\begin{array}{l} x \in \mathcal{H} \setminus \bar{C} \Rightarrow x \in \mathcal{H} \setminus \operatorname{dom} \breve{g} \Rightarrow \breve{g}(x) = +\infty \\ \text{By construction, } x \in \mathcal{H} \setminus \bar{C} \Rightarrow f(x) = +\infty \end{array}\right\} \Rightarrow \forall_{x \in \mathcal{H} \setminus \bar{C}} \ f(x) = \breve{g}(x).$$

Part 3. in this part we will prove that: $\forall_{x \in \text{bdry dom } g} \ f(x) = \breve{g}(x)$; now $\text{bdry dom } g = (\text{bdry dom } g \setminus \overline{\operatorname{dom} g}) \cup (\text{bdry dom } g \cap \overline{\operatorname{dom} g})$:

so we will focus on them seperately

Consider $x \in (\text{bdry } C) \setminus (\overline{\operatorname{dom} g}) = (\text{bdry dom } g) \setminus (\overline{\operatorname{dom} g})$

$\Rightarrow g(x) = +\infty$, /* because $\operatorname{dom} g \subseteq \overline{\operatorname{dom} g}$, we are removing $\operatorname{dom} \bar{g}$, $g(x) = +\infty$ */

$$f(x) = \underbrace{\underline{\lim}_{y \to x} g(y)}_{} \geq \underline{\lim}_{y \to x} \breve{g}(y) \quad /* \because g \geq \breve{g} \, */$$

$$\underset{\text{by construction}}{\geq} \breve{g}(x) \quad /* \, \breve{g}: \text{lower semicontinuous} \overset{\text{def}}{\Leftrightarrow} y \to x \Rightarrow \underline{\lim}\, \breve{g}(y) \geq \breve{g}(x) \, */$$

$$= +\infty$$

$\forall_{x \in \text{bdry } C \setminus \operatorname{dom} \breve{g}} \ f(x) = \breve{g}(x) = +\infty$

Now consider,

$x \in \text{bdry } C \cap \operatorname{dom} \breve{g}$

$= \text{bdry dom } g \cap \operatorname{dom} \breve{g} \Rightarrow x \in \operatorname{dom} \breve{g} \Rightarrow (x, \breve{g}(x)) \in \operatorname{epi} \breve{g} = \overline{\operatorname{epi}}\, g$ /* Using [2017-02-08 9:24 AM] */

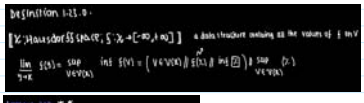
[1/25/2016 2:37 PM] says $\exists_{(x_n, \xi_n)_{n \in \mathbb{N}}}$: sequence in epi g $(x_n, \xi_n) \to (x, \breve{g}(x)) \leftrightarrow (x, \breve{g}(x)) \in \overline{\operatorname{epi}}\, g$

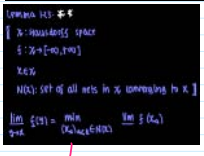
now, using similar logic as in [1/25/2016 2:44 PM] we have:

$$f(x) = \underbrace{\underline{\lim}_{y \to x} g(y) \geq \underline{\lim}_{y \to x} \breve{g}(y) = \breve{g}(x)}_{/* \ g \geq \breve{g} \ */} = \lim \xi_n = \underbrace{\underline{\lim}\, \xi_n}_{/* \ (x_n, \xi_n) \in \operatorname{epi} g \leftrightarrow g(x_n) \leq \xi_n \Rightarrow \underline{\lim}\, g(x_n) \leq \underline{\lim}\, \xi_n \, */} \geq \overbrace{\underline{\lim}\, g(x_n)}^{/* \text{ as } x_n \to x \in \operatorname{dom} \breve{g} \, */} \geq \underbrace{\underline{\lim}_{y \to x} g(y)}_{\text{lower semicontinuous}} = f(x) \Rightarrow \text{collapse of inequalities}$$

$$\Rightarrow f(x) = \breve{g}(x)$$

; $\underline{\lim}_{y \to x} g(y) = \min_{(x_a)_{a \in A} : x_a \to x} \underline{\lim}\, g(x_a) \leq \underline{\lim}\, g(x_n)$ */





$\therefore \forall_{x \in \text{bdry } C \cap \operatorname{dom} \breve{g}} \ f(x) = g(x)$

Combining Parts 1,2,3 we have $f = \breve{g} \in \Gamma_0(\mathcal{H})$

$\therefore f$: lower semicontinuous, convex, proper.

∎

Part 1

Monday, January 25, 2016 9:03 PM

Proposition 10.6.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, $\beta \in \mathbb{R}_{++}$]

f: strongly convex with constant β $\leftrightarrow$

$f - \frac{\beta}{2}\|\cdot\|^2$: convex

Proof:

$f - (\beta/2)\|\cdot\|^2$: convex

$$\leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \ \underbrace{\left(f - \frac{\beta}{2}\|\cdot\|^2\right)(\alpha x + (1-\alpha)y)}_{f(\alpha x+(1-\alpha)y) - \frac{\beta}{2}\underbrace{\|\alpha x + (1-\alpha)y\|^2}_{\alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2}} \leq \alpha \overbrace{\left(f - \frac{\beta}{2}\|\cdot\|^2\right)x}^{f(x) - \frac{\beta}{2}\|x\|^2} + (1-\alpha)\overbrace{\left(f - \frac{\beta}{2}\|\cdot\|^2\right)y}^{f(y) - \frac{\beta}{2}\|\cdot\|^2}$$

* Corollary 2.14. (Very useful identity)

$\forall_{x,y \in H} \ \forall_{\alpha \in \mathbb{R}} \quad \|\alpha x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2$

$$\leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad f(\alpha x + (1-\alpha)y) - \frac{\beta}{2}\alpha\|x\|^2 - \frac{\beta}{2}(1-\alpha)\|y\|^2 + \frac{\beta}{2}\alpha(1-\alpha)\|x-y\|^2 \leq \alpha f(x) - \frac{\beta}{2}\alpha\|x\|^2 + (1-\alpha)f(y) - \frac{\beta}{2}(1-\alpha)\|y\|^2$$

$$\leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\left(\frac{\beta}{2}\|x-y\|^2\right) \leq \alpha f(x) + (1-\alpha)f(y)$$

$\overset{\operatorname{dom} f}{\leftrightarrow}$ f: strongly convex with constant β.

∎

Corollary 10.11.

[f: proper, convex, exact modulus of convexity Φ]

f: uniformly convex $\Leftrightarrow$ Φ: vanishes only at 0 with modulus Φ

1/26/2016 3:01 PM

Proof:

($\Rightarrow$)

f: uniformly convex with modulus ψ

$$\Leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\psi(\|x-y\|) \leq \alpha f(x) + (1-\alpha)f(y)$$

$$\Leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad \psi(\|x-y\|) \leq \frac{\alpha f(x) + (1-\alpha)f(y) - f(\alpha x + (1-\alpha)y)}{\alpha(1-\alpha)}$$

$$\Leftrightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad \psi(\|x-y\|) \leq \inf_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t \\ \alpha \in]0,1[}} \frac{\alpha f(x) + (1-\alpha)f(y) - f(\alpha x + (1-\alpha)y)}{\alpha(1-\alpha)} = \Phi(t) \quad /* \because \forall_x \ f(x) \leq g(x) \Rightarrow \forall_x \ f(x) \leq \inf_x g(x)$$

$\forall_x \forall_z \quad f(x, g(z)) \leq h(x,z) \Rightarrow f(x, g(z)) \leq \inf_{x, g(z)=t} h(x,z)$ */

/*

Exact modulus of convexity.

$$\Phi(t) = \inf_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t \\ \alpha \in]0,1[}} \frac{\alpha f(x) + (1-\alpha)f(y) - f(\alpha x + (1-\alpha)y)}{\alpha(1-\alpha)} \quad */$$

$$\Rightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{t : t = \|x-y\|} \quad \psi(t) \leq \Phi(t)$$

$$\Rightarrow \psi \leq \Phi$$

now by definition,

* Definition 10.5. (Uniformly convex)

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

f: uniformly convex with modulus $\psi: \mathbb{R}_+ \to [0, +\infty]$ $\overset{\text{def}}{\leftrightarrow}$

- ψ: increasing
- ψ: vanishes only at 0
- $\forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\psi(\|x-y\|) \leq \alpha f(x) + (1-\alpha)f(y) \quad \ldots \quad (10.1)$

as ψ: vanishes only at zero, strictly positive elsewhere and increasing

Φ: vanishes at zero, increasing /* * Proposition 10.10. (Properties of exact modulus of convexity)

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex, Φ: exact modulus of convexity] $\Rightarrow$

- $\Phi(0) = 0$
- $\forall_{t \in \mathbb{R}_+} \forall_{r \in [1, +\infty]} \ \Phi(rt) \geq r^2 \Phi(t)$
- Φ: increasing. */

so. $\forall_t \quad \psi(0) = 0 = \Phi(0)$
$\forall_{t \neq 0} \quad 0 < \psi(t) \leq \Phi(t)$ $\Big\} \Rightarrow \Phi$: vanishes only at zero

($\Leftarrow$)

Φ: vanishes only at zero

now: $\underbrace{\Phi(t)}_{\|x-y\|} = \inf\limits_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t \\ \alpha \in]0,1[}} \dfrac{\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)}{\alpha(1-\alpha)} \leq \dfrac{\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)}{\alpha(1-\alpha)} \quad \forall_{\alpha \in]0,1[} \ \forall_{x,y \in \operatorname{dom} f : \|x-y\| = t}$

$\Rightarrow \forall_{x,y \in \operatorname{dom} f} \ \forall_{\alpha \in]0,1[} \quad f(\alpha x + (1-\alpha) y) + \alpha(1-\alpha)\, \Phi(\|x-y\|) \leq \alpha f(x) + (1-\alpha) f(y)$

also. Φ: increasing

$\therefore$ By definition $f(\cdot)$: uniformly convex with modulus Φ.

$\blacksquare$

* Proposition 10.12.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex, exact modulus of convexity φ

$\psi: \mathbb{R}_+ \to [0, +\infty]$, $t \mapsto \inf\limits_{\substack{x \in \operatorname{dom} f \\ y \in \operatorname{dom} f \\ \|x-y\| = t}} \ \frac{1}{2} f(x) + \frac{1}{2} f(y) - f\left(\frac{1}{2} x + \frac{1}{2} y\right)$

] $\Rightarrow$

(i) $2\psi \leq \varphi \leq 4\psi$

(ii) f: uniformly convex $\Leftrightarrow$ ψ: vanishes only at 0.

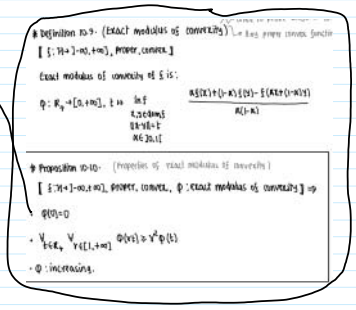
1/27/2016 1:33 PM

Proof:

(i) $t \in \mathbb{R}_+$

$\Phi(0) = 0$

if $t = 0$, then
$\|x-y\| = 0 \Leftrightarrow x = y$


* Definition 10.9. (Exact modulus of convexity)
[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex]
Exact modulus of convexity of f is:
$\varphi: \mathbb{R}_+ \to [0, +\infty]$, $t \mapsto \inf\limits_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t \\ \alpha \in]0,1[}} \dfrac{\alpha f(x) + (1-\alpha) f(y) - f(\alpha x + (1-\alpha) y)}{\alpha(1-\alpha)}$
* Proposition 10.10. (Properties of exact modulus of convexity)
[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex, Φ: exact modulus of convexity] $\Rightarrow$
$\Phi(0) = 0$
$\forall_{t \in \mathbb{R}_+} \forall_{r \in [1, +\infty]} \ \Phi(rt) \geq r^2 \Phi(t)$
Φ: increasing.

then

$\psi(0) = \inf\limits_{\substack{x,y \in \operatorname{dom} f \\ x = y}} \underbrace{\frac{1}{2} f(x) + \frac{1}{2} f(y) - f\left(\frac{1}{2} x + \frac{1}{2} y\right)}_{\frac{1}{2} f(x) + \frac{1}{2} f(x) - f(x) = 0} = 0$

$\therefore$ at $t = 0$, $\psi(0) = \Phi(0) = 0$.

assume $\operatorname{dom} f$: not a singleton /* recall f: convex $\Rightarrow$ $\operatorname{dom} f$: convex */

/* $\operatorname{dom} f$: singleton $\Rightarrow$ trivially the claim holds as $x = y \to t = \|x-y\| = 0$, for which case $\psi(0) = \Phi(0) = 0$ */

$t > 0$

$\alpha \in]0, \frac{1}{2}]$,

$x_0, y_0 \in \operatorname{dom} f : \|x_0 - y_0\| = t$ /* if no such point x, y exists, trivially we have the claim (i) as $\Phi(t) = \psi(t) = +\infty$ */

now,

$f(\alpha x_0 + (1-\alpha) y_0)$

$= f\left(2\alpha \cdot \frac{1}{2} x_0 + 2\alpha \cdot \frac{1}{2} y_0 + y_0 - 2\alpha y_0\right)$

$= f\left(2\alpha\left(\frac{1}{2} x_0 + \frac{1}{2} y_0\right) + (1-2\alpha) y_0\right)$ /* $2\alpha \in]0, 1]$, so, 2α is a convex parameter */

$\leq 2\alpha\, f\left(\frac{1}{2} x_0 + \frac{1}{2} y_0\right) + (1-2\alpha) f(y_0)$

/* now by definition,

$\psi(t) = \inf\limits_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t}} \frac{1}{2} f(x) + \frac{1}{2} f(y) - f\left(\frac{1}{2} x + \frac{1}{2} y\right) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y) - f\left(\frac{1}{2} x + \frac{1}{2} y\right) \quad \forall_{\substack{x,y \in \operatorname{dom} f \\ \|x-y\| = t}}$

$\to \quad \psi(\|x-y\|) \leq \frac{1}{2} f(x) + \frac{1}{2} f(y) - f\left(\frac{1}{2} x + \frac{1}{2} y\right) \quad \forall_{x,y \in \operatorname{dom} f}$

$$\psi(t) = \inf_{\substack{x,y\in \operatorname{dom} f\\ \|x-y\|=t}} \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y)-f\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right) \leq \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y)-f\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right) \quad \forall_{\substack{x,y\in\operatorname{dom} f\\ \|x-y\|=t}}$$

$$\rightarrow \quad \psi(\|x-y\|) \leq \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y)-f\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right) \quad \forall_{x,y\in\operatorname{dom} f}$$

$$\Rightarrow \forall_{x,y\in\operatorname{dom} f} \quad f\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right) \leq \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y)-\psi\left(\|x-y\|\right)$$

$$\therefore x_0,y_0\in\operatorname{dom} f \Rightarrow \quad f\left(\tfrac{1}{2}x_0+\tfrac{1}{2}y_0\right) \leq \tfrac{1}{2}f(x_0)+\tfrac{1}{2}f(y_0)-\psi\left(\|x_0-y_0\|\right)$$

$$\Rightarrow 2\alpha\, f\left(\tfrac{1}{2}x_0+\tfrac{1}{2}y_0\right) \leq \alpha f(x_0)+\alpha f(y_0)-2\alpha\,\psi\left(\|x_0-y_0\|\right) \quad */$$

$$f(\alpha x_0+(1-\alpha)y_0) \leq \alpha f(x_0)+\alpha f(y_0)-2\alpha\,\psi\left(\|x_0-y_0\|\right)+(1-2\alpha)f(y_0)$$

$$= \alpha f(x_0)+(1-\alpha)f(y_0)-2\alpha\,\psi\big(\underbrace{\|x_0-y_0\|}_{t}\big)$$

$$= \alpha f(x_0)+(1-\alpha)f(y_0)-2\alpha\,\psi(t)$$

$$\leq \alpha f(x_0)+(1-\alpha)f(y_0)-2\alpha\,\psi(t)+\underbrace{2\alpha^2}_{\geq 0}\underbrace{\psi(t)}$$

nonnegative function by construction, $\psi:\mathbb{R}_+\to[0,+\infty]$

$$= \alpha f(x_0)+(1-\alpha)f(y_0)-2\alpha(1-\alpha)\psi(t)$$

$$\Rightarrow f(\alpha x_0+(1-\alpha)y_0) \leq \alpha f(x_0)+(1-\alpha)f(y_0)-2\alpha(1-\alpha)\psi(t)$$

$$\Rightarrow \psi(t) \leq \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{2\alpha(1-\alpha)}$$

$$\Rightarrow 2\psi(t) \leq \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)}$$

$$\therefore \forall_{x_0,y_0\in\operatorname{dom} f},\ t=\|x_0-y_0\|,\ \alpha\in]0,1/2]$$

$$2\psi(t) \leq \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)}$$

/* Because the R.H.S is symmetric in x_0,y_0, so whatever is constructed in terms of $\alpha\in]0,1/2]$, can be created using $\alpha\in]0,1[$ (just by swapping x_0,y_0); say $\alpha=\frac{1}{4}$, then $2\psi(t) \leq \dfrac{\frac{1}{4}f(x_0)+\frac{3}{4}f(y_0)-f\left(\frac{1}{4}x_0+\frac{3}{4}y_0\right)}{\frac{1}{4}\cdot\frac{3}{4}}$, if we take $\tilde{\alpha}=\frac{3}{4}\in]0,1[$,

$$2\psi(t) \leq \frac{\tilde{\alpha} f(y_0)+(1-\tilde{\alpha})f(x_0)-f\left(\tilde{\alpha} f(y_0)+(1-\tilde{\alpha})f(x_0)\right)}{\tilde{\alpha}(1-\tilde{\alpha})}$$

now swap x_0,y_0 as it is upto us:

$$2\psi(t) \leq \left(\tilde{\alpha} f(x_0)+(1-\tilde{\alpha})f(y_0)-f\left(\tilde{\alpha} f(x_0)+(1-\tilde{\alpha})f(y_0)\right)\right) / \tilde{\alpha}(1-\tilde{\alpha}) \quad \forall\tilde{\alpha}\in]0,1[\quad */$$

$$\therefore \forall_{x_0,y_0\in\operatorname{dom} f},\ t=\|x_0-y_0\|,\ \alpha\in]0,1[$$

$$2\psi(t) \leq \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)}$$

$$\Rightarrow 2\psi(t) \leq \inf_{\substack{x_0,y_0\in\operatorname{dom} f\\ t=\|x_0-y_0\|\\ \alpha\in]0,1[}} \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0}{\alpha(1-\alpha)}$$

$$= \phi(t) \quad /* \text{ recall, } \quad \phi(t) = \inf_{\substack{t=\|x_0-y_0\|\\ \alpha\in]0,1[}} \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)} \quad */$$

so, $2\psi(t) \leq \phi(t)$

and again,

$$\phi(t) = \inf_{\substack{x_0,y_0\in\operatorname{dom} f\\ t=\|x_0-y_0\|\\ \alpha\in]0,1[}} \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)} \leq \left[\inf_{\substack{x_0,y_0\in\operatorname{dom} f\\ t=\|x_0-y_0\|}} \frac{\alpha f(x_0)+(1-\alpha)f(y_0)-f(\alpha x_0+(1-\alpha)y_0)}{\alpha(1-\alpha)}\right]_{\alpha=\frac{1}{2}}$$

$$= \inf_{\substack{x_0,y_0\in\operatorname{dom} f\\ t=\|x_0-y_0\|}} \frac{\frac{1}{2}f(x_0)+\frac{1}{2}f(y_0)-f\left(\frac{1}{2}x_0+\frac{1}{2}y_0\right)}{\frac{1}{4}}$$

$$= 4\underbrace{\inf_{\substack{x_0,y_0\in\operatorname{dom} f\\ t=\|x_0-y_0\|}} \left(\tfrac{1}{2}f(x_0)+\tfrac{1}{2}f(y_0)-f\left(\tfrac{1}{2}x_0+\tfrac{1}{2}y_0\right)\right)}_{\psi(t)\ \text{/* By definition */}}$$

$$= 4\psi(t)$$

$$\therefore \forall_{\substack{t:\, t=\|x_0-y_0\|,\\ x_0\in \operatorname{dom} f,\\ y_0\in \operatorname{dom} f}} \quad 2\psi(t)\leq \Phi(t)\leq 4\psi(t)$$

$$\Leftrightarrow \quad 2\psi\leq \Phi\leq 4\psi$$

(ii) from (i)

$$\forall_{\substack{t:\, t=\|x_0-y_0\|,\\ x_0\in \operatorname{dom} f,\\ y_0\in \operatorname{dom} f}} \quad 2\psi(t)\leq \Phi(t)\leq 4\psi(t)$$

when $t=0$, $\quad 2\psi(0)\leq \Phi(0)\leq 4\psi(0)$

Assume ψ: vanishes only at 0

$\Leftrightarrow \quad \psi(0)=0,\ 0\leq \Phi(0)\leq 0$

$\Leftrightarrow \quad \Phi(0)=0$ at 0 only

$\Leftrightarrow \quad f$: uniformly convex with modulus Φ

/* * Corollary 10.11.
[f: proper, convex, exact modulus of convexity Φ]
f: uniformly convex $\Leftrightarrow$ Φ: vanishes only at 0 */

$\blacksquare$

Proposition 10.15.

[$f: \mathcal{H}\to]-\infty,+\infty]$, proper, convex;
C: nonempty compact convex, $\subseteq \operatorname{dom} f$, f: strictly convex on C
$f|_C$: continuous] $\Rightarrow$

f: uniformly convex on C

Proof:

$g := f+\iota_C \Rightarrow \operatorname{dom} g = C$

define:

$$\psi: \mathbb{R}_+\to[0,+\infty]$$
$$t\mapsto \inf_{\substack{x,y\in \operatorname{dom} g\\ \|x-y\|=t}} \tfrac{1}{2}g(x)+\tfrac{1}{2}g(y)-g\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right)$$

So, $\psi(t)=\inf_{\substack{x,y\in \operatorname{dom} g\\ \|x-y\|=t}} \tfrac{1}{2}g(x)+\tfrac{1}{2}g(y)-g\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right)$

* Proposition 10.12. /* Finding bounds for exact modulus of convexity */
[$f: \mathcal{H}\to]-\infty,+\infty]$, proper, convex, exact modulus of convexity Φ

$\psi: \mathbb{R}_+\to[0,+\infty]$: $t\mapsto \inf_{\substack{x,y\in \operatorname{dom} f\\ \|x-y\|=t}} \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y)-f\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right)$] $\Rightarrow$

(i) $2\psi\leq\Phi\leq 4\psi$

(ii) f: uniformly convex $\Leftrightarrow$ ψ: vanishes only at 0.

* Proposition 10.10. (Properties of exact modulus of convexity)
[$f: \mathcal{H}\to]-\infty,+\infty]$, proper, convex, Φ: exact modulus of convexity] $\Rightarrow$

- $\Phi(0)=0$
- $\forall_{t\in\mathbb{R}_+}\ \forall_{r\in[1,+\infty]}\ \Phi(rt)\geq r^2\Phi(t)$
- Φ: increasing.

So, at $t=0 \Rightarrow 2\psi(0)\leq \underbrace{\Phi(0)}_{0}\leq 4\psi(0)$

$\leftrightarrow \boxed{\psi(0)\leq 0}\leq 2\psi(0)$ as $\operatorname{ran}\psi\subseteq[0,+\infty]$ $\Rightarrow \psi(0)=0 \Rightarrow$ so ψ has atleast one zero possibly many

We want to prove that ψ vanishes only at zero

take such a "t" that $\psi(t)=0$, and per absurdum take $t=\|x-y\|\neq 0 \Leftrightarrow x\neq y$ [Eq: 2017-02-14 8:53 PM]

$$\inf_{\substack{x,y\in \operatorname{dom} g=C\\ \|x-y\|=t}} \tfrac{1}{2}g(x)+\tfrac{1}{2}g(y)-g\left(\tfrac{1}{2}x+\tfrac{1}{2}y\right)=0 \quad // \text{finite}$$

now using

* Fact 1.8.1. (Existence of a minimizing sequence for finite infimum)
$\inf f(C)\in\mathbb{R} \Rightarrow \exists_{(x_n)_{n\in\mathbb{N}}:\subseteq C}\ f(x_n)\to \inf f(C)$

we have:

$$\exists_{\substack{(x_n)_{n\in\mathbb{N}},\ (y_n)_{n\in\mathbb{N}}:\ x_n,y_n\in \operatorname{dom} g=C,\\ \|x_n-y_n\|=t}} \quad \tfrac{1}{2}g(x_n)+\tfrac{1}{2}g(y_n)-g\left(\tfrac{1}{2}x_n+\tfrac{1}{2}y_n\right)\to 0$$

$(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} : x_n, y_n \in \operatorname{dom} g = C.$
$\|x_n - y_n\| = t$

now C: compact $\Rightarrow$ so every sequence in C has a convergent subsequence

/* C: compact $\Leftrightarrow$ $\forall_{(x_n)_{n\in A}:\text{ net in } C}\ \exists_{(x_{k(b)})_{b\in B}:\text{ subnet of } (x_n)_{n\in A};\ x\in C}\ x_{k(b)} \to x$ */

as $(x_n)_{n\in\mathbb{N}} \subseteq C$; $(y_n)_{n\in\mathbb{N}} \subseteq C$ $\quad \exists_{x\in C, y\in C}\ \exists_{(x_{k_n})_{n\in\mathbb{N}} \subseteq C,\ (y_{k_n})_{n\in\mathbb{N}} \subseteq C}\ x_{k_n} \to x \in C,\ y_{k_n} \to y \in C$

$\frac{1}{2} g(x_{k_n}) + \frac{1}{2} g(y_{k_n}) - g(\frac{1}{2} x_{k_n} + \frac{1}{2} y_{k_n}) \to 0$

/* If a sequence converges to something, any of its subsequence converges to the same thing */

also $\|x_{k_n} - y_{k_n}\| = t \Rightarrow t \leq \|x_{k_n} - y_{k_n}\| \leq t$

now, $(x_{k_n} \to x,\ y_{k_n} \to y) \Rightarrow x_{k_n} - y_{k_n} \to x - y$

$\Rightarrow \|x_{k_n} - y_{k_n}\| \to \|x - y\|$ /* using (One characterization of strong convergence) * Corollary 2.52. $[(x_n)_{n\in\mathbb{N}} : \subseteq H;\ x \in H]$ $x_n \to x \Leftrightarrow (x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|)$ */

/* using Fact 1.3.1. $\forall_{n\in\mathbb{N}}\ a_n \leq b_n \leq c_n \Rightarrow \left| \begin{array}{l} a_n \leq \underline{\lim}\, a_n \leq \overline{\lim}\, a_n \leq c_n \end{array} \right.$ */

$t \leq \|x - y\| \leq t$

$\therefore \|x - y\| = t$

now $g|_C = (f + \underbrace{\iota_C}_{0})|_C = f|_C$: continuous by given

/*

* Fact 1.19. * Definition of continuity using nets
[X, Y: Hausdorff spaces
$T: X \to Y$
$x \in X$]
T: continuous at $x \Leftrightarrow \forall_{\underbrace{(x_a)_{a\in A}}_{\subseteq X} : x_a \to x}\ Tx_a \to Tx.$ */

$\therefore \frac{1}{2} g(x_{k_n}) + \frac{1}{2} g(y_{k_n}) - g(\frac{1}{2} x_{k_n} + \frac{1}{2} y_{k_n}) \to \frac{1}{2} g(x) + \frac{1}{2} g(y) - g(\frac{1}{2} x + \frac{1}{2} y)$

Both of the limits must be the same thing

$\frac{1}{2} g(x) + \frac{1}{2} g(y) - g(\frac{1}{2} x + \frac{1}{2} y) = 0$

$\Leftrightarrow \frac{1}{2} g(x) + \frac{1}{2} g(y) = g(\frac{1}{2} x + \frac{1}{2} y)$

but $g|_C = f|_C$: strictly convex on C by given

$\therefore \forall_{x,y\in C : x \neq y}\ g(\frac{1}{2} x + \frac{1}{2} y) < \frac{1}{2} g(x) + \frac{1}{2} g(y)$

$x = y$

$\therefore t = \|x - y\| = 0$ contradiction with [Eq: 2017-02-14 8:53 PM]

$\therefore \psi$: vanishes only at zero

$\Updownarrow$

f: uniformly convex on C

/*

* Proposition 10.12. /* Finding bounds for exact modulus of convexity *
[$f: H \to]-\infty, +\infty]$, proper, convex, exact modulus of convexity φ

$\psi: \mathbb{R}_+ \to [0, +\infty]: t \mapsto \inf_{\substack{x,y\in\operatorname{dom} f \\ \|x-y\|=t}} \left[\frac{1}{2} f(x) + \frac{1}{2} f(y) - f(\frac{1}{2} x + \frac{1}{2} y) \right] \Rightarrow$

(i) $2\psi \leq \varphi \leq 4\psi$

(ii) f: uniformly convex $\Leftrightarrow \psi$: vanishes only at 0. */

$\boxtimes$

Part 2

Thursday, January 28, 2016 10:17 AM

Proposition. 10.22.

[$(f_i)_{i\in I}$: family of quasiconvex functions from $\mathcal{H}$ to $[-\infty,+\infty]$]

$\sup_{i\in I} f_i$: quasiconvex

1/28/2016 10:26 AM Proof:

$\forall_{\xi\in\mathbb{R}}$

$\mathrm{lev}_{\leq\xi} \sup_{i\in I} f_i$

$= \{x\in\mathcal{H} \mid \sup_{i\in I} f_i \leq \xi\}$

$\Updownarrow$

$\forall_{i\in I}\ f_i \leq \xi$

$= \bigcap_{i\in I} \{x\in\mathcal{H} \mid f_i \leq \xi\}$

$= \Big(\bigcap_{i\in I} \underbrace{\mathrm{lev}_{\leq\xi} f_i}_{\text{: convex /* By definition */}}\Big)$: convex /* As intersection of convex sets are always convex */

$\forall_{\xi\in\mathbb{R}} \left(\sup_{i\in I} f_i\right)$: convex $\overset{\text{def}}{\Leftrightarrow}$ $\sup_{i\in I} f_i$: quasiconvex

/* $\because$ By definition, f : quasiconvex $\overset{\text{def}}{\Leftrightarrow} \forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: convex

∎

* Proposition 10.23.

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

(i) f : weakly sequentially lower semicontinuous $\Leftrightarrow$

(ii) f : sequentially lower semicontinuous $\Leftrightarrow$

(iii) f : lower semicontinuous $\Leftrightarrow$

(iv) f : weakly lower semicontinuous

1/28/2016 10:13 PM Proof:

- f : quasiconvex $\overset{\text{def}}{\Leftrightarrow} \forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: convex

Assume

f : lower semicontinuous $\Rightarrow \forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: closed

/* using

Lemma 1.24.

[X : Hausdorff space,

$f: X\to[-\infty,\infty]$]

f : lower semicontinuous $\Leftrightarrow$ epi f : closed in $X\times\mathbb{R}$ $\Leftrightarrow \forall_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: closed in X

But /* Theorem 3.32. /* this is a very important theorem which says that for a convex set all the different types of closedness coincides */

[C : convex subset of $\mathcal{H}$]

C : weakly sequentially closed $\Leftrightarrow$ C : sequentially closed $\Leftrightarrow$ C : closed $\Leftrightarrow$ C : weakly closed */

Theorem 3.32. /* this is a very important theorem which says that for a convex set all the different types of closedness coincides */

[C : convex subset of $\mathcal{H}$]

C : weakly sequentially closed $\Leftrightarrow$ C : sequentially closed $\Leftrightarrow$ C : closed $\Leftrightarrow$ C : weakly closed */

$\forall_{\xi \in \mathbb{R}}$ $\mathrm{lev}_{\leq \xi} f$: weakly sequentially closed $\Leftrightarrow$ f : weakly sequentially lower semicontinuous

$\Updownarrow$

$\mathrm{lev}_{\leq \xi} f$: sequentially closed $\Leftrightarrow$ f : sequentially lower continuous

$\Updownarrow$

$\mathrm{lev}_{\leq \xi} f$: closed $\Leftrightarrow$ f : lower semicontinuous

$\Updownarrow$

$\mathrm{lev}_{\leq \xi} f$: weakly closed $\Leftrightarrow$ f : weakly lower semicontinuous

*Lemma 1.39.

[X : Hausdorff space

$f : X \to [-\infty, +\infty]$]

(i) f : sequentially lower semicontinuous $\Leftrightarrow$

(ii) epi f : sequentially closed $\Leftrightarrow$

(iii) $\forall_{\xi \in \mathbb{R}}$ $\mathrm{lev}_{\leq \xi} f$: sequentially closed in X

So, we have

f : weakly sequentially lower semicontinuous $\Leftrightarrow$

f : sequentially lower semicontinuous $\Leftrightarrow$

f : lower semicontinuous $\Leftrightarrow$

f : weakly lower semicontinuous

$\blacksquare$

Part 1

Saturday, January 30, 2016 12:46 PM

Proposition 11.1.

$[f:\mathcal{H}\to[-\infty,+\infty]\,;\ C:\text{ nonempty}\subseteq\mathcal{H}]$

(i) f: lower semicontinuous $\Rightarrow$ $\sup f(\overline{C})=\sup f(C)$

(ii) f: convex $\Rightarrow$ $\sup f(\operatorname{conv} C)=\sup f(C)$

(iii) $u\in\mathcal{H}\Rightarrow\begin{cases}\sup\langle\overline{\operatorname{conv}}\,C\,|\,u\rangle=\sup\langle C|u\rangle\\ \inf\langle\overline{\operatorname{conv}}\,C\,|\,u\rangle=\inf\langle C|u\rangle\end{cases}$

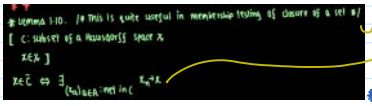
(iv) $(f\in\Gamma_0(\mathcal{H}),\ C:\text{ convex};\ \operatorname{dom}f\cap\operatorname{int}C\neq\emptyset)\Rightarrow\inf f(\overline{C})=\inf f(C)$

Proof:

(i) $C\subseteq\overline{C}$

$\Rightarrow\sup f(C)\leq\sup f(\overline{C})$ // Optimum objective value improves over a larger set

set $x\in\overline{C}$ /*



*/

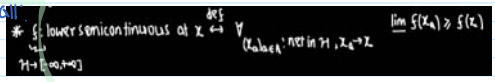
now a closed set is sequentially closed

$\Rightarrow\exists\,(x_n)_{n\in\mathbb{N}}:\subseteq C$ $\Rightarrow\exists\,(x_n)_{n\in\mathbb{N}}$: sequence in C $x_n\to x$

$x_n\to x$

now f: lower semicontinuous at x

/* recall:



*/

$\Rightarrow\underline{\lim}\,f(x_n)\geq f(x)$

/* now: $\underline{\lim}\,\xi_a=\sup\limits_{a\in A}\inf\limits_{b\in A:b\geq a}\xi_b$

$\therefore\underline{\lim}\,f(x_n)=\sup\limits_{n\in\mathbb{N}}\left(\inf\limits_{m\in\mathbb{N}:m\geq n}f(x_m)\right)\leq\sup\limits_{n\in\mathbb{N}}f(x_n)=\sup\{f(\tilde{x})\,|\,\tilde{x}\in\overbrace{(x_n)_{n\in\mathbb{N}}}^{\in C},\ \tilde{x}\in C\}\leq\sup\{f(x)\,|\,x\in C\}=\sup f(C)$

/* as this one is a supremum over a larger set */

$\forall_{x\in\overline{C}}\quad f(x)\leq\underline{\lim}\,f(x_n)\leq\sup f(C)$

$\Rightarrow\sup\limits_{x\in\overline{C}}f(x)=\sup f(\overline{C})\leq\sup f(C)$ ∎

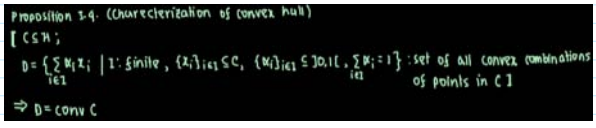
(ii)

$C\subseteq\operatorname{conv}C$

$\Rightarrow\sup f(C)\leq\sup f(\operatorname{conv}C)$

take $x\in\operatorname{conv}C$

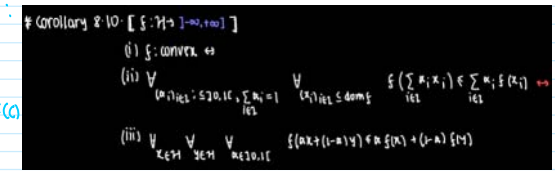
$\Leftrightarrow\exists\,(\alpha_i)_{i\in I}$: finite family in $]0,1[$, $\sum\limits_{i\in I}\alpha_i=1$ $\exists\,(x_i)_{i\in I}$: finite family in C $x=\sum\limits_{i\in I}\alpha_i x_i$ /* using



*/

now

$f(x)=f\left(\sum\limits_{i\in I}\alpha_i x_i\right)$ /* recall:



*/

$\leq\sum\limits_{i\in I}\alpha_i\underbrace{f(x_i)}_{\leq\sup\limits_{x\in C}f(x)=\sup f(C)}$

$\leq\sum\limits_{i\in I}\alpha_i\sup f(C)$

$=\sup f(C)\underbrace{\sum\limits_{i\in I}\alpha_i}_{1}=\sup f(C)$

$\therefore\forall_{x\in\operatorname{conv}C}\quad f(x)\leq\sup f(C)$

$\Rightarrow\sup\limits_{x\in\operatorname{conv}C}f(x)=\sup f(\operatorname{conv}C)\leq\sup f(C)$

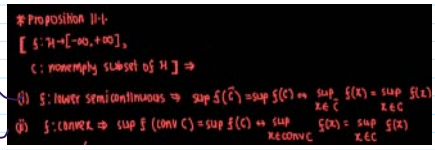
$\sup f(C)=\sup f(\operatorname{conv}C)$ ∎

(iii)

first note that $\left.\begin{matrix}\langle\cdot|u\rangle,\\ \langle\cdot|u\rangle\end{matrix}\right\}$ affine function $\Rightarrow$ convex, continuous in $\cdot$ $\Rightarrow$ lower semicontinuous $\Big\}\Rightarrow$ so we can apply (i) and (ii)

$\sup\limits_{x\in\overline{\operatorname{conv}}C}\langle x|u\rangle=\sup\limits_{x\in\operatorname{conv}C}\langle x|u\rangle$ using

$=\sup\limits_{x\in C}\langle x|u\rangle$ using



$\sup\langle\overline{\operatorname{conv}}\,C\,|\,u\rangle=\sup\langle C|u\rangle$

$= \sup_{x\in C} \langle z|u\rangle$ using

(i) f: lower semicontinuous $\Rightarrow \sup f(\bar{C}) = \sup f(C) \Leftrightarrow \sup_{x\in\bar{C}} f(x) = \sup_{x\in C} f(x)$

(ii) f: convex $\Rightarrow \sup f(\operatorname{conv} C) = \sup f(C) \Leftrightarrow \sup_{x\in \operatorname{conv} C} f(x) = \sup_{x\in C} f(x)$

$\therefore \sup \langle \overline{\operatorname{conv}}\, C|u\rangle = \sup \langle C|u\rangle$

as $\sup(\cdot) = -\inf(-\cdot)$, and $-\langle\cdot|u\rangle$ is convex, continuous we have: $\inf \langle \overline{\operatorname{conv}}\, C|u\rangle = \inf\langle C|u\rangle$

(iv) $\bar{C} \supseteq C$

$\Rightarrow \boxed{\inf f(\bar{C}) \leq \inf f(C)}$

take, $x_0 \in \bar{C}$, $x_1 \in \operatorname{dom} f \cap \operatorname{int} C$

set, $x_\alpha = (1-\alpha)x_0 + \alpha x_1$, $\forall_{\alpha\in]0,1[}$

/* Proposition 3.35.

C: convex subset of $\mathcal{H} \Rightarrow \forall_{x\in \operatorname{int} C} \ \forall_{y\in\bar{C}} \ [x,y[\subseteq C$ /* $[x,y[= \{(1-\alpha)x+\alpha y \mid 0<\alpha<1\}$ */ */

using this we have: $x_\alpha = (1-\alpha)x_0 + \alpha x_1 \in C$

$\therefore \forall_{\alpha\in]0,1[} \ x_\alpha = (1-\alpha)x_0 + \alpha x_1 \in C$

$\Rightarrow \forall_{\alpha\in]0,1[} \ f(\underbrace{x_\alpha}_{\in C}) \geq \inf_{x\in C} f(x)$

$\Rightarrow \lim_{\alpha\downarrow 0} f(x_\alpha) \geq \inf_{x\in C} f(x)$

/*

*Proposition 3.14.

[$f\in\Gamma_0(\mathcal{H})$, $\tilde{z}\in\mathcal{H}$, $\tilde{y}\in\operatorname{dom} f$

$\forall_{\alpha\in]0,1[} \ \tilde{z}_\alpha := (1-\alpha)\tilde{z} + \alpha\tilde{y}$]

$\lim_{\alpha\downarrow 0} f(\tilde{z}_\alpha) = f(\tilde{z})$

*/

/*

Fact 1.1.1. / these are extendable to nets as well */

(comes handy in dealing with sequences)

- $\forall_{n\in\mathbb{N}} \ a_n \leq a \Rightarrow \overline{\lim}\, a_n \leq a$
- $\forall_{n\in\mathbb{N}} \ b \leq b_n \Rightarrow b \leq \underline{\lim}\, b_n$
- $\forall_{n\in\mathbb{N}} \ a_1 \leq a_n \leq a_2 \Rightarrow a_1 \leq \underline{\lim}\, a_n \leq \overline{\lim}\, a_n \leq a_2$
  - $\lim a_n$ exists in finite norm
  - $\overline{\lim}\, a_n$ exists in finite norm
- $\forall_{n\in\mathbb{N}} \ a_n \leq b_n \Rightarrow (\underline{\lim}\, a_n \leq \underline{\lim}\, b_n,\ \overline{\lim}\, a_n \leq \overline{\lim}\, b_n)$
- $\forall_{n\in\mathbb{N}} \ a_n \geq 0,\ b_n \geq 0 \Rightarrow \overline{\lim}\, a_n b_n \leq \overline{\lim}\, a_n \overline{\lim}\, b_n$

*/

using this: $\lim_{\alpha\downarrow 0} f(x_\alpha) = f(x_0)$

$\therefore \underline{\lim}_{\alpha\downarrow 0} f(x_\alpha) = \lim_{\alpha\downarrow 0} f(x_\alpha) = f(x_0)$

$f(x_0) \geq \inf_{x\in C} f(x)$

So, $\forall_{x_0\in\bar{C}} \ f(x_0) \geq \inf_{x\in C} f(x) \Leftrightarrow \inf_{x_0\in\bar{C}} f(x_0) \geq \inf_{x\in C} f(x)$

$\therefore \boxed{\inf f(\bar{C}) \geq \inf f(C)}$

So, $\inf f(\bar{C}) = \inf f(C)$ ∎

* Proposition 11.4.

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

x: local minimizer of $f \Leftrightarrow x$: minimizer of f.

1/30/2016 2:09 PM

Proof:

($\Leftarrow$): trivial

($\Rightarrow$): x: local minimizer of f

$\Leftrightarrow \exists_{\rho\in\mathbb{R}_{++}} \ f(x) = \min f(\underbrace{B(x;\rho)}_{\text{closed ball of center } x \text{ and radius } \rho})$

pick $y \in \operatorname{dom} f \setminus B(x;\rho) \Rightarrow f(y) < +\infty$, $\|x-y\| > \rho \to 1 > \frac{\rho}{\|x-y\|} > 0 \Rightarrow -1 < -\frac{\rho}{\|x-y\|} < 0$

$\Rightarrow 0 < \underbrace{1 - \frac{\rho}{\|x-y\|}}_{\alpha} < 1$

$\therefore \alpha \in]0,1[$

construct $z = \alpha x + (1-\alpha)y$

note that $\|z-x\| = \|\alpha x + (1-\alpha)y - x\|$

$= \|(\alpha-1)x + (1-\alpha)y\|$

$= \|-(1-\alpha)x + (1-\alpha)y\|$

$= \|(1-\alpha)x - (1-\alpha)y\|$

$= (1-\alpha)\|x-y\|$

$= \left(1 - \left(1 - \frac{\rho}{\|x-y\|}\right)\right)\|x-y\|$

$= \frac{\rho}{\|x-y\|}\|x-y\| = \rho$

$$= \left(1-\left(1-\frac{\rho}{\|x-y\|}\right)\right)\|x-y\|$$

$$= \frac{\rho}{\|x-y\|}\|x-y\| = \rho$$

$\therefore z \in B(x;\rho)$

now, $f(x) = \min f(B(x;\rho))$

$\Rightarrow \forall_{\tilde{x}\in B(x;\rho)}\ f(x) \leq f(\tilde{x})$

recall $y \in \operatorname{dom} f \setminus B(x;\rho)$

as $z \in B(x;\rho)$ $\quad f(x) \leq f(z) = f(\alpha x + (1-\alpha)y)$

$\leq \alpha f(x) + (1-\alpha) f(y)$

$\Rightarrow (1-\alpha) f(x) \leq (1-\alpha) f(y)$

$\Rightarrow f(x) \leq f(y)$

$\therefore \forall_{y\in \operatorname{dom} f\setminus B(x;\rho)}\ f(x) \leq f(y)$

by given, $\forall_{\tilde{x}\in B(x;\rho)}\ f(x) \leq f(\tilde{x})$

$\Big\}\ \forall_{y\in \operatorname{dom} f}\ f(x) \leq f(y) \Leftrightarrow x$: minimizer of f.

/* Note, we donot have to worry about $y \in \mathcal{H}\setminus \operatorname{dom} f$ as $f(y) = +\infty$ then */

∎

Proposition. 11.5.

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

C: subset of $\mathcal{H}$

x: minimizer of f over C, $x \in \operatorname{int} C$]

x: minimizer of f.

1/30/2016 2:33 PM

Proof:

$x \in \operatorname{int} C \Rightarrow C \neq \emptyset$

$\Rightarrow \exists_{\rho\in \mathbb{R}_{++}}\ B(x;\rho) \subseteq C$

as x: minimizer of f over C

$\Rightarrow x$: minimizer of f over $B(x;\rho)$

$\therefore f(x) = \inf f(B(x;\rho))$

so x: local minimizer of f

$\Rightarrow x$: global minimizer of f.

∎

/* f: quasiconvex $\Leftrightarrow (\operatorname{lev}_{\leq\alpha} f)_{\alpha\in\mathbb{R}}$: convex set */

* Proposition 11.6.

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

Argmin f: convex

1/31/2016 1:22 PM

Proof:

$\operatorname{Argmin} f = \{x\in\mathcal{H} \mid \forall_{y\in\mathcal{H}}\ f(x) \leq f(y)\}$

$= \{x\in\mathcal{H} \mid \forall_{y\in\operatorname{dom} f}\ f(x) \leq f(y),\ \overbrace{\forall_{y\in\mathcal{H}\setminus\operatorname{dom} f}\ f(x) \leq f(y) = +\infty}^{\text{redundant}}\}$

$= \{x\in\mathcal{H} \mid \forall_{y\in\operatorname{dom} f}\ f(x) \leq f(y) \in \mathbb{R}\}$

$\Updownarrow$

$f(x) \leq \inf_{y\in\operatorname{dom} f} f(y) = \alpha \in \mathbb{R}$ $\quad$ /* $\forall_{y\in\operatorname{dom} f}\ f(y) < +\infty \Rightarrow f(\operatorname{dom} f)$: bounded $\Rightarrow \inf f(\operatorname{dom} f) =$ finite */

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

$= \{x\in\mathcal{H} \mid f(x) \leq \alpha, \alpha\in\mathbb{R}\} = \operatorname{lev}_{\underbrace{\alpha}_{\in\mathbb{R}}} f$: convex by definition.

∎

* Proposition 11.7. (Existence of atleast one minimizer in an optimization problem)

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

C: convex subset of $\mathcal{H}$, $C \cap \operatorname{dom} f \neq \emptyset$

One of the following holds:

(i) $f+\iota_C$: strictly quasiconvex

(ii) f: convex, $C\cap \operatorname{Argmin} f=\emptyset$, C: strictly convex $\overset{\text{def}}{\Leftrightarrow}$

$$\forall_{x,y\in C}\ x\neq y \Rightarrow \frac{x+y}{2}\in \operatorname{int} C$$

]

$\Rightarrow$ f: atmost one minimizer over C.

1/31/2016 2:04 PM

Proof:

Assume C: not a singleton, otherwise proof is trivial

$\mu=\inf f(C)$ /* as $C\cap \operatorname{dom} f\neq\emptyset \Rightarrow \exists_{z\in C\cap \operatorname{dom} f}\ f(z)<+\infty$, so $\inf f(C)$: finite */

Assume $\exists_{x,y\in C\cap \operatorname{dom} f}: x\neq y$ $\quad f(x)=f(y)=\mu$ /* Per absurdum */

note that, $x\in C\cap \operatorname{lev}_{\leq\mu} f$

$y\in C\cap \operatorname{lev}_{\leq\mu} f$

/* $x\in C\cap \operatorname{dom} f \Leftrightarrow x\in C,\ f(x)<+\infty$

$f(x)=\mu \Rightarrow x: f(x)\leq\mu<+\infty \Leftrightarrow x\in \operatorname{lev}_{\leq\mu} f$

$\therefore x\in C\cap \operatorname{lev}_{\leq\mu} f$

similarly, $y\in C\cap \operatorname{lev}_{\leq\mu} f$ */

as f: convex $\Rightarrow \operatorname{lev}_{\leq\mu} f$: convex

C: convex

$\therefore C\cap \operatorname{lev}_{\leq\mu} f$: convex /* intersection of convex sets are always convex */

$\underbrace{\tfrac{1}{2}x+\tfrac{1}{2}y}_{=z\ (\text{say})}\in C\cap \operatorname{lev}_{\leq\mu} f \Rightarrow f(z)\leq\mu$ but $\mu=\inf f(C)\leq f(z)$

So, $f(z)=\mu$ (i.e., z: another different minimizer of $f|_C$)

$=\inf f(C)$

(i)

recall, $\tilde{f}$: strictly quasiconvex $\overset{\text{def}}{\Leftrightarrow} \forall_{\tilde{x}\in \operatorname{dom}} \forall_{\tilde{y}\in \operatorname{dom} f} \forall_{\kappa\in]0,1[}\ \tilde{x}\neq\tilde{y} \Rightarrow \tilde{f}(\kappa\tilde{x}+(1-\kappa)\tilde{y})<\max\{\tilde{f}(\tilde{x}),\tilde{f}(\tilde{y})\}$

$f(x)=f(y)=\underset{\tfrac{1}{2}x+\tfrac{1}{2}y}{\underset{\parallel}{f(z)}}=\inf f(C)$: finite $\Rightarrow x,y,z\in \operatorname{dom}(f+\iota_C)$

also, $x\neq y$

so, $\underbrace{(f+\iota_C)(\overset{z}{\overset{\parallel}{\tfrac{1}{2}x+\tfrac{1}{2}y}})}_{\underbrace{f(z)}_{\mu}+\underbrace{\iota_C(z)}_{0}=f(z)}<\max\{\underbrace{(f+\iota_C)(x)}_{\underset{\mu}{\underset{\parallel}{f(x)}}+\underset{0}{\underset{\parallel}{\iota_C(x)}}},\underbrace{(f+\iota_C)(y)}_{\underset{\mu}{\underset{\parallel}{f(y)}}+\underset{0}{\underset{\parallel}{\iota_C(y)}}}\}$

$\Rightarrow \mu<\max\{\mu,\mu\}=\mu$: contradiction

$\therefore$ There cannot be more than one minimizer.

(ii)

$z=\tfrac{1}{2}x+\tfrac{1}{2}y\in C\cap \operatorname{lev}_{\leq\mu} f$, $f(z)=\inf f(C)$

as $x,y\in C$, $x\neq y \Rightarrow \frac{x+y}{2}=z\in \operatorname{int} C$ /* C: strictly convex set so, applying the definition */

/* Recall

* Proposition 11.5. /* The statement is not trivial at all */

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

C: subset of $\mathcal{H}$ ✓

z ~~x~~: minimizer of f over C, ~~x~~ $z\in \operatorname{int} C$ ✓ ✓

$\Rightarrow$ z ~~x~~ minimizer of f.

*/

z: minimizer of $f \Leftrightarrow z\in \operatorname{Argmin} f$

$z\in C\cap \operatorname{Argmin} f$

$\therefore C\cap \operatorname{Argmin} f\neq\emptyset$: contradiction (According to given, $C\cap \operatorname{Argmin} f=\emptyset$)

$\therefore f$ has atmost one minimizer over C.

∎

2/2/2016 9:36 PM

*Theorem 11.9. (Existence of minimizers)

[$f: \mathcal{H} \to]-\infty, +\infty]$, lower semicontinuous, quasiconvex

C: closed convex subset of $\mathcal{H}$

$\exists_{\xi \in \mathbb{R}}$ $C \cap \mathrm{lev}_{\leq \xi} f$: nonempty, bounded] $\Rightarrow$

f: has a minimizer over C.

Proof: Recall

/*
* Proposition 10.23.
[$f: \mathcal{H} \to]-\infty, +\infty]$, quasiconvex]
(i) f: weakly sequentially lower semicontinuous $\Leftrightarrow$
(ii) f: sequentially lower semicontinuous $\Leftrightarrow$
(iii) f: lower semicontinuous $\Leftrightarrow$
(iv) f: weakly lower semicontinuous */

f: lower semicontinuous, quasiconvex $\Rightarrow$ f: weakly lower semicontinuous.

C: closed convex

$\mathrm{lev}_{\leq \xi} f$: closed convex

/* A quasiconvex function has convex lower levelset by definition

A lower semicontinuous f has closed lower level set as

Lemma 1.24. ★
[X: Hausdorff space,
$f: X \to [-\infty, \infty]$]

f: lower semicontinuous $\Leftrightarrow$ epi f: closed in $X \times \mathbb{R}$ $\Leftrightarrow$ $\forall_{\xi \in \mathbb{R}}$ $\mathrm{lev}_{\leq \xi} f$: closed in X */

$\therefore C \cap \mathrm{lev}_{\leq \xi} f$: closed, convex

/* closedness and convexness are preserved under intersection */

So, $C \cap \mathrm{lev}_{\leq \xi} f$: closed, convex, $\underbrace{\text{bounded}}_{\text{given}}$ $\Rightarrow$ $C \cap \mathrm{lev}_{\leq \xi} f$: weakly compact, weakly sequentially compact.

$\Leftrightarrow$ $C \cap \mathrm{lev}_{\leq \xi} f$: bounded, weakly closed

/* Lemma 2.29: C: weakly compact $\Leftrightarrow$
C: bounded, weakly closed */

/* using: Theorem 3.33.
[C: bounded, closed, convex subset of $\mathcal{H}$]
C: weakly compact, weakly sequentially compact

/* C: weakly compact $\overset{\text{def}}{\Leftrightarrow}$ Every net in has a weak cluster point in C

C: weakly sequentially compact $\overset{\text{def}}{\Leftrightarrow}$ Every sequence in C has a weak sequential cluster point in C

x: (strong) cluster point of $(x_\alpha)_{\alpha \in A}$ $\overset{\text{def}}{\Leftrightarrow}$
$(x_\alpha)_{\alpha \in A}$ has a subnet that (strong)ly converges to $x \in X$

x: weak cluster point of $(x_\alpha)_{\alpha \in A}$ $\overset{\text{def}}{\Leftrightarrow}$
$(x_\alpha)_{\alpha \in A}$ has a subnet that weakly converges to $x \in X$ */

*/

Note that

$$\begin{pmatrix} \min\limits_{x \in \mathcal{H}} & f(x) \\ \text{s.t.} & x \in C \end{pmatrix} = \begin{pmatrix} \min\limits_{x \in \mathcal{H}} & f(x) + L_C(x) \end{pmatrix}$$

$$= \begin{pmatrix} \min\limits_{x \in \mathcal{H}} & f(x) + L_C(x) + L_{\mathrm{lev}_{\leq \xi} f}(x) \end{pmatrix}$$

$$= \begin{pmatrix} \min\limits_{x \in \mathcal{H}} & f(x) \\ \text{s.t.} & x \in C \cap \mathrm{lev}_{\leq \xi} f \end{pmatrix}$$

/* Given, $C \cap \mathrm{lev}_{\leq \xi} f$: nonempty, bounded

$\Leftrightarrow$ $\exists_{x \in C}$ $f(x) \overset{\in \mathbb{R}}{\leq \xi}$

So we can search for the minimizer in $\mathrm{lev}_{\leq \xi} f$ where $L_C(x)$ will evaluate to zero */

So, minimizing f over C is equivalent to minimizing f over $C \cap \mathrm{lev}_{\leq \xi} f$.

/* Recall.

- Lemma 2.23. $\mathcal{H}^{\text{weak}}$: Hausdorff space

- * Theorem 1.28. (Weierstrass) /* fundamental tool in proving the existence of solutions of optimization problem */

[X: Hausdorff space,
$f: X \to [-\infty, +\infty]$, lower semicontinuous
C: compact subset of X
$C \cap \mathrm{dom}\, f \neq \emptyset$] $\Rightarrow$

f: achieves its infimum over C.

*/

in our case we have:

$\mathcal{H}^{\text{weak}}$: Hausdorff space

$f: \mathcal{H} \to [-\infty, +\infty]$, lower semicontinuous

$C \cap \mathrm{lev}_{\leq \xi} f$: weakly compact

$(C \cap \mathrm{lev}_{\leq \xi} f) \cap \mathrm{dom}\, f \neq \emptyset$ /* as $\mathrm{lev}_{\leq \xi} f \subseteq \mathrm{dom}\, f$, $\Rightarrow$ $\mathrm{dom}\, f \cap \mathrm{lev}_{\leq \xi} f = \mathrm{lev}_{\leq \xi} f \neq \emptyset$
given $C \cap \mathrm{lev}_{\leq \xi} f \neq \emptyset$ */

$\therefore$ f achieves a minimum over $C \cap \mathrm{lev}_{\leq \xi} f$

$\Leftrightarrow$ f " " " " C.

∎

Proposition 11.11. (Coercivity of a function in terms of lower levelset)

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

f: coercive $\Leftrightarrow$ $(\mathrm{lev}_{\leq \xi} f)_{\xi \in \mathbb{R}}$: bounded

Proof:

($\Rightarrow$) Proof by contrapositive

$$\neg\left((\mathrm{lev}_{\leq\xi} f)_{\xi\in\mathbb{R}}\right) : \text{bounded} \Rightarrow f : \text{not coercive} \leftrightarrow \lim_{\|x\|\to+\infty} f(x) = +\infty$$

Per absurdum,

$$\exists_{\xi\in\mathbb{R}} \quad \mathrm{lev}_{\leq\xi} f : \text{unbounded}$$

$$\Rightarrow \exists_{(x_n)_{n\in\mathbb{N}} \subseteq \mathrm{lev}_{\leq\xi} f} \quad \|x_n\| \to +\infty$$

$$\updownarrow$$

$$f(x_n) \leq \xi \in \mathbb{R}$$

$$\leftrightarrow \quad \|x_n\| \to +\infty, \left(f(x_n) \leq \xi \Rightarrow f(x_n) \neq +\infty\right)$$

$$\Rightarrow \underline{\lim}_{\|x_n\|\to+\infty} f(x_n) \leq \overline{\lim}_{\|x_n\|\to+\infty} f(x_n) \leq \xi < +\infty$$

/* using

- $\forall_{n\in\mathbb{N}} \quad a_1 \leq a_n \leq a_2 \Rightarrow$
  - $a_1 \leq \underline{\lim}\, a_n \leq \overline{\lim}\, a_n \leq a_2$
  - $\underline{\lim}\, a_n$ exists in finite norm
  - $\overline{\lim}\, a_n$ exists in finite norm

*/

as, $\underline{\lim} \leq \lim \leq \lim$

$$\lim_{\|x\|\to+\infty} f(x) \neq +\infty$$

$$\updownarrow$$

f: not coercive.

($\Leftarrow$) /* $(\mathrm{lev}_{\leq\xi} f)_{\xi\in\mathbb{R}}$: bounded $\Rightarrow$ f: coercive */

$$\forall_{\xi\in\mathbb{R}} \quad \mathrm{lev}_{\leq\xi} f = \{x\in\mathcal{H} \mid f(x)\leq\xi\} : \text{bounded} \Rightarrow \exists_{t\in\mathbb{R}_+} \forall_{x\in\mathrm{lev}_{\leq\xi} f \leftrightarrow f(x)\leq\xi} \quad \|x\|\leq t$$

$$\therefore \forall_{\xi\in\mathbb{R}} \exists_{t\in\mathbb{R}_+} \forall_x \left(f(x)\leq\xi \Rightarrow \|x\|\leq t\right)$$

$$\updownarrow \text{contrapositive}$$

$$\left(\|x\| > t \Rightarrow f(x) > \xi\right)$$

take a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H} : \|x_n\|\to+\infty \overset{\text{def}}{\leftrightarrow} \forall_{a\in\mathbb{R}} \exists_{N\in\mathbb{N}} \forall_{n\in\mathbb{N}: n\geq N} \quad \|x_n\| > a$

$$\overset{a:=t}{\Rightarrow} \exists_{N\in\mathbb{N}} \forall_{n\in\mathbb{N}: n\geq N} \quad \|x_n\| > t \Rightarrow f(x_n) > \xi$$

$$\therefore \exists_{N\in\mathbb{N}} \forall_{n\in\mathbb{N}: n\geq N} \quad f(x_n) > \xi$$

$$\Rightarrow \inf_{n\geq N} f(x_n) \geq \xi$$

but ξ was arbitrary positive number

$$\therefore \forall_{\xi\in\mathbb{R}_+} \forall_{(x_n)_{n\in\mathbb{N}}: \|x_n\|\to+\infty} \exists_N \inf_{n\geq N} f(x_n) \geq \xi$$

$$\Downarrow$$

$$f(x_n) \to +\infty$$

$$\therefore \lim_{\|x\|\to+\infty} f(x) = +\infty \overset{\text{def}}{\leftrightarrow} f : \text{coercive}.$$

$\blacksquare$

Proposition 11.12.

[$\mathcal{H}$: finite-dimensional

$f \in \Gamma_0(\mathcal{H})$]

$$f : \text{coercive} \leftrightarrow \exists_{\xi\in\mathbb{R}} \quad \mathrm{lev}_{\leq\xi} f : \text{nonempty, bounded}$$

Proof:

($\Rightarrow$)

*** Proposition 11.11. (Coercivity of a function in terms of lowerlevel set)**

[$f : \mathcal{H} \to [-\infty, +\infty]$] f: coercive $\leftrightarrow$ $(\mathrm{lev}_{\leq\xi} f)_{\xi\in\mathbb{R}}$: bounded

using this:

$$\exists_{\xi\in\mathbb{R}} \quad \mathrm{lev}_{\leq\xi} f : \text{nonempty, bounded}$$

as long as $\mathrm{dom} f$: nonempty we can find a nonempty sublevel set.

($\Leftarrow$) $\exists_{\xi\in\mathbb{R}} \mathrm{lev}_{\leq\xi} f$: nonempty bounded $\leftrightarrow$ $\mathrm{lev}_{\leq\xi} f = \{x\in\mathcal{H} \mid f(x)\leq\xi\}$: nonempty bounded $\leftrightarrow$ $\exists_{x'\in\mathcal{H}} \; f(x')\leq\xi$

$$\exists_{\eta>0} \forall_{x\in\mathcal{H}: f(x)\leq\xi} \quad \|x\|\leq\eta$$

$\Rightarrow \forall_{\xi'\leq\xi} \quad \mathrm{lev}_{\leq\xi'} f$: bounded /* note that we are focusing on boundedness only as we want to use Proposition 11.11 */

$\exists_{\xi\in\mathbb{R}}\ \mathrm{lev}_{\leq\xi} f$: nonempty bounded $\Leftrightarrow$ $\mathrm{lev}_{\leq\xi} f = \{x\in\mathcal{H} \mid f(x)\leq\xi\}$: nonempty bounded $\Leftrightarrow$ $\exists_{x'\in\mathcal{H}}\ f(x')\leq\xi$

$\Rightarrow \forall_{\xi'\leq\xi}\ \mathrm{lev}_{\leq\xi'} f$: bounded /+ note that we are focusing on boundedness only as we want to use Proposition 11.11 */

$\exists_{\eta>0}\ \forall_{x\in\mathcal{H}:\ f(x)\leq\xi}\ \|x\|\leq\eta$

$\hookrightarrow$ A $\forall_{y\in\mathrm{lev}_{\leq\xi'} f}$

$f(y)\leq\xi' \Rightarrow f(y)\leq\xi'\leq\xi \Rightarrow y\in\mathrm{lev}_{\leq\xi} f \Rightarrow \|y\|\leq\eta$ $\quad\therefore\ \mathrm{lev}_{\leq\xi'} f$: bounded */

now take $\eta\in\,]\xi,+\infty[$ we want to show that $\mathrm{lev}_{\leq\eta} f$: bounded

Per absurdum let us assume : $\mathrm{lev}_{\leq\eta} f$: unbounded /+ now to be unbounded it has to be nonempty ;

also $f\in\Gamma_0(\mathcal{H}) \Rightarrow \mathrm{lev}_{\leq\eta} f$: convex */

/+ *Corollary 6.51. [$\mathcal{H}$: finite dimensional, C : nonempty convex subset of $\mathcal{H}$] $\Rightarrow$ C : bounded $\Leftrightarrow$ $\mathrm{rec}\, C=\{0\}$ +/

A $\mathrm{rec}\, C=\{z\in\mathcal{H} \mid z+C\subseteq C\}$ +/

$\mathrm{lev}_{\leq\eta} f$: convex, nonempty, bounded

$\Rightarrow \mathrm{rec}\ \mathrm{lev}_{\leq\eta} f \neq \{0\} \Rightarrow \mathrm{rec}\ \mathrm{lev}_{\leq\eta} f$: unbounded as this is a cone

// $\mathbb{R}_{++}$ cone = cone

A $\{\tilde{z}\in\mathcal{H} \mid \tilde{z}+\mathrm{lev}_{\leq\eta} f\subseteq\mathrm{lev}_{\leq\eta} f\}\neq\{0\}$

$\therefore\ \forall_{\tilde{y}:\ \tilde{y}\in\mathrm{rec}\,\mathrm{lev}_{\leq\eta} f}$ $\quad\therefore$ so, $\tilde{y}+x'\in\mathrm{lev}_{\leq\eta} f$

$\Updownarrow$

$\tilde{y}+\underbrace{\mathrm{lev}_{\leq\eta} f}_{\ni x'}\subseteq\mathrm{lev}_{\leq\eta} f$

/+ $x'\in\mathrm{lev}_{\leq\xi} f \Leftrightarrow f(x')\leq\xi\leq\eta \Rightarrow x'\in\mathrm{lev}_{\leq\eta} f$ */

now $\mathrm{rec}\ \mathrm{lev}_{\leq\eta} f$ is a cone, so $\forall_{\lambda\in\mathbb{R}_{++}}\ \lambda\tilde{y}\in\mathrm{rec}\ \mathrm{lev}_{\leq\eta} f$

$\lambda\tilde{y}+x'\in\mathrm{lev}_{\leq\eta} f\ \ \forall_{\lambda\in\mathbb{R}_{++}}$

now take $\lambda\in\,]1,+\infty[$

$\forall_{\lambda\in\,]1,+\infty[}$ $\Downarrow$ $\frac{1}{\lambda}\in\,]0,1[$

$x'+\tilde{y}=x'-\frac{1}{\lambda}x'+\frac{1}{\lambda}x'+\frac{1}{\lambda}\lambda\tilde{y}=(1-\frac{1}{\lambda})\underbrace{x'}_{\in\mathrm{lev}_{\leq\eta} f}+\frac{1}{\lambda}\underbrace{(x'+\lambda\tilde{y})}_{\in\mathrm{lev}_{\leq\eta} f}$: convex combination of x', $x'+\lambda\tilde{y}$

also. $\mathrm{lev}_{\leq\eta} f\subseteq\mathrm{dom}\, f$ and $f\in\Gamma_0(\mathcal{H})$

So,

$$f(x'+\tilde{y})=f\Big((1-\tfrac{1}{\lambda})\underbrace{x'}_{\in\mathrm{dom} f}+\tfrac{1}{\lambda}\underbrace{(x'+\lambda\tilde{y})}_{\in\mathrm{dom} f}\Big)\leq(1-\tfrac{1}{\lambda})f(x')+\tfrac{1}{\lambda}\underbrace{f(x'+\lambda\tilde{y})}_{\leq\eta}$$

$$\leq(1-\tfrac{1}{\lambda})f(x')+\tfrac{1}{\lambda}\eta$$

$\Rightarrow f(x'+\tilde{y})-f(x')\leq\frac{1}{\lambda}(\eta-f(x'))$

$\Rightarrow \forall_{\lambda\in\,]1,+\infty[}\ \lambda(f(x'+\tilde{y})-f(x'))\leq\underbrace{\eta-f(x')}_{\geq 0}$ /* $x'\in\mathrm{lev}_{\leq\eta} f \Rightarrow f(x)\leq\eta \Rightarrow \eta-f(x)\geq 0$ and $\eta-f(x)$: finite */

and finite

$f\in\Gamma_0(\mathcal{H})$ are from $\mathcal{H}$ to $]-\infty,+\infty]$

$\Rightarrow \forall_{\lambda\in\,]1,+\infty[}\ f(x'+\tilde{y})-f(x')\leq\frac{1}{\lambda}(\eta-f(x')) \Leftrightarrow f(x'+\lambda\tilde{y})-f(x')\leq\inf_{\lambda\in\,]1,+\infty[}\frac{1}{\lambda}(\eta-f(x'))$

$=(\eta-f(x'))\underbrace{\inf_{\lambda\in\,]1,+\infty[}\frac{1}{\lambda}}_{=0}=0$

$\therefore\ f(x'+\tilde{y})-f(x')\leq 0$

$\Rightarrow f(x'+\tilde{y})\leq\underbrace{f(x')\leq\xi}_{\text{A : } x'\in\mathrm{lev}_{\leq\xi} f\ */}$

$\therefore\ \forall_{\tilde{y}\in\mathrm{rec}\,\mathrm{lev}_{\leq\eta} f}$ $\quad f(x'+\tilde{y})\leq\xi$

$\Updownarrow$

$x'+\tilde{y}\in\mathrm{lev}_{\leq\xi} f$

$\Rightarrow \underbrace{\underbrace{x'+\mathrm{rec}\,\mathrm{lev}_{\leq\eta} f}_{\text{unbounded}}}_{\text{unbounded}}\subseteq\underbrace{\mathrm{lev}_{\leq\xi} f}_{\text{bounded}}$

So, we have contained an unbounded set in a bounded set.

$\Downarrow$

contradiction $\Rightarrow \mathrm{lev}_{\leq\eta} f$: can not be unbounded

$\therefore\ \mathrm{lev}_{\leq\eta} f$: bounded

So, we have proven that :

$\forall\ \mathrm{lev}_{\leq\eta} f$: bounded A

*Proposition 11.11. (Coercivity of a function in terms of lower level set)

So, we have proven that:

$\forall_{\eta\in\mathbb{R}}$ $\mathrm{lev}_{\leq\eta} f$: bounded

$\Leftrightarrow$ f : coercive. ■

*Proposition 11.11. (Coercivity of a function in terms of lower level set)
[$f:\mathcal{H}\to(-\infty,+\infty]$] f : coercive $\Leftrightarrow$ $(\mathrm{lev}_{\leq\xi} f)_{\xi\in\mathbb{R}}$: bounded

*Proposition 11.17. (Asymptotic center)

[$(z_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$

C : nonempty closed convex subset of $\mathcal{H}$

$T: C\to C$, nonexpansive

$f: \mathcal{H}\to\mathbb{R}: x\mapsto \overline{\lim}\, \|x-z_n\|^2$]

(i) f : strongly convex with constant 2.

(ii) f : supercoercive

(iii) $f+\iota_C$: strongly convex, supercoercive,
its unique minimizer z_C is called the asymptotic center of $(z_n)_{n\in\mathbb{N}}$ relative to C.

(iv) $(z\in\mathcal{H}, z_n\to z) \Rightarrow \forall_{x\in\mathcal{H}}$ $f(x)=\|x-z\|^2+f(z)$, $z_C=P_C z$

(v) $(z_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t $C \Rightarrow P_C z_n = z_C$

(vi) $\forall_{n\in\mathbb{N}}$ $z_{n+1}=Tz_n \Rightarrow z_C\in \mathrm{Fix}\,T$

(vii) $z_n - Tz_n \to 0 \Rightarrow z_C \in \mathrm{Fix}\,T$.

*Corollary 11.18.

[C : nonempty closed convex subset of $\mathcal{H}$

$T: C\to C$, nonexpansive

$z_0\in C$, $\forall_{n\in\mathbb{N}}$ $z_{n+1}=Tz_n$ /*Classical Banach-Picard iteration*/]

(i) $\mathrm{Fix}\,T\neq\emptyset \Leftrightarrow$

(ii) $\forall_{z_0\in C}$ $(z_n)_{n\in\mathbb{N}}$: bounded $\Leftrightarrow$

(iii) $\exists_{z_0\in C}$ $(z_n)_{n\in\mathbb{N}}$: bounded.

2/3/2016 11:23 AM

Proof:

(i) $\Rightarrow$ (ii)

T : nonexpansive $\Rightarrow$ T : quasinonexpansive.

Recall Example 5.3, Proposition 5.9.(i)

/*

*Example 5.3. * (Used later)

[C : nonempty subset of $\mathcal{H}$ ✓

$T: C\to C$, quasinonexpansive ✓ /* T : quasinonexpansive $\Leftrightarrow \forall_{x\in D}\ \forall_{y\in \mathrm{Fix}\,T}\ \|Tx-\overset{y}{Ty}\|\leq\|x-y\|$ */

$\mathrm{Fix}\,T\neq\emptyset$ ✓ [Given] ✓

$z_0\in C$, $\forall_{n\in\mathbb{N}}$ $x_{n+1}=Tx_n$ /* Banach-Picard iteration */

] $(z_n)_{n\in\mathbb{N}}$

$(z_n)_{n\in\mathbb{N}}$: Fejer monotone w.r.t. $\mathrm{Fix}\,T$ ✓

Proposition 5.9. (Some key properties of Fejer monotone sequences)

[C nonempty subset of $\mathcal{H}$ ✓ $\mathrm{Fix}\,T$ ✓

(Given) $(z_n)_{n\in\mathbb{N}}$ Fejer monotone sequence w.r.t. C, $S\mathcal{H}$] $\Rightarrow$

(i) $(z_n)_{n\in\mathbb{N}}$: bounded

(ii) $\forall_{x\in C}$ $(\|x_n-x\|)_{n\in\mathbb{N}}$: converges /*note that it does not necessarily mean convergence in C*/

(iii) $(d_C(x_n))_{n\in\mathbb{N}}$: decreasing and converges

*/

So, $\forall_{z_0\in C}$ $(z_n)_{n\in\mathbb{N}}$: bounded ☺

(ii) $\Rightarrow$ (iii) : obvious. ☺

(iii) $\Rightarrow$ (i) :

(ii) ⇒ (iii) : obvious. ①

(iii) ⇒ (i) :

given $\exists_{z_0 \in C}$ $(z_n)_{n\in\mathbb{N}}$: bounded

Recall Proposition 11.17.(vi) says:

[$(z_n)_{n\in\mathbb{N}}$: bounded sequence, ✓

C : nonempty closed convex subset of $\mathcal{H}$ ✓

$T : C \to C$, nonexpansive ✓

$\forall_{n\in\mathbb{N}}$ $z_n = Tz_n$ ✓]

⇒

$z_c : \underbrace{(P_C z_n \to z_c)}_{\text{/*convergent point of the shadow sequence*/}}$, $z_c \in \operatorname{Fix} T$

⇒ $\operatorname{Fix} T \neq \emptyset$

∎

* Proposition 11.19.

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

⇒

$\forall_{(x_n)_{n\in\mathbb{N}}}$: minimizing sequence of f $\quad (x_n)_{n\in\mathbb{N}}$: bounded

2/3/2016 12:04 PM

Proof: first note that $f : \mathcal{H} \to]-\infty, +\infty]$, $\underbrace{\text{proper}}_{\operatorname{dom} f \neq \emptyset,\ -\infty \notin f(\mathcal{H})}$ ⇒ $\inf f(x)$: finite

/* Minimizing sequence :

[$f : X \to [-\infty, +\infty]$; $(x_n)_{n\in\mathbb{N}}$: sequence in $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$]

$(x_n)_{n\in\mathbb{N}}$: minimizing sequence of $f \overset{\text{def}}{\Leftrightarrow} f(x_n) \to \inf f(x)$ */

* Proposition 11.11. (Coercivity of a function in terms of lower level set)

[$f : \mathcal{H} \to [-\infty, +\infty]$] f : coercive ↔ $(\operatorname{lev}_{\leq \xi} f)_{\xi \in \mathbb{R}}$: bounded

$\forall_{\xi\in\mathbb{R}}$ $(\operatorname{lev}_{\leq \xi} f)$: bounded

↔ $\operatorname{lev}_{\leq \xi} f = \{x \in \mathcal{H} \mid f(x) \leq \xi\}$: bounded

per absurdum assume $(x_n)_{n\in\mathbb{N}}$: not bounded /* $(x_n)_{n\in\mathbb{N}}$: sequence in C, bounded $\overset{\text{def}}{\Leftrightarrow} \exists_{m\in\mathbb{N}} \forall_{n\in\mathbb{N}}$ $x_n \in C \cap B(0;m)$ */

So, $\|x_n\| \to +\infty$

now, f : coercive $\overset{\text{def}}{\Leftrightarrow} \lim_{\|x\|\to\infty} f(x) = +\infty$

⇒ $\lim_{\|x_n\|\to\infty} f(x_n) = +\infty$ ⇒ $f(x_n) \to +\infty$

but $f(x_n) \to \inf f(\mathcal{H}) < +\infty$ } : contradiction

∴ Any minimizing sequence of f : bounded.

∎

Part 2

Thursday, February 4, 2016 6:50 PM

*Proposition 11.25.

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

$(x_n)_{n\in\mathbb{N}}$: minimizing sequence of f

$\exists_{\xi \in]\inf f(\mathcal{H}), +\infty[}$ $C = \mathrm{lev}_{\leq \xi} f$: bounded]

(i) $(x_n)_{n\in\mathbb{N}}$: has a weak sequential cluster point,

and such point is a minimizer

(ii) $(f+L_C)$: strictly quasiconvex $\Rightarrow$ f : has unique minimizer, $x_n \rightharpoonup x$

(iii) $(f+L_C)$: uniformly quasiconvex $\Rightarrow$ f : has a unique minimizer, $x_n \to x$

2/5/2016 12:33 PM

Proof: Assume $(x_n)_{n\in\mathbb{N}} \subseteq C = \mathrm{lev}_{\leq \xi} f$ without loss of generality /* $\mathrm{lev}_{\leq \xi} f$: closed, convex, bounded */

as f: lower semicontinuous, f: quasiconvex, bounded: given

/* Explanation: Rudin 3.2.(a)

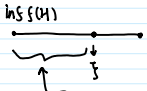
2/5/2016 2:41 PM

$\{p_n\}$ converges to $p \in X \Leftrightarrow$ any neighborhood of p contains p_n for all but finitely many n

now, $(x_n)_{n\in\mathbb{N}} \subseteq C = \mathrm{lev}_{\leq \xi} f = \{x \in \mathcal{H} \mid f(x) \leq \xi\}$

$\Leftrightarrow \forall_{n\in\mathbb{N}}$ $f(x_n) \leq \xi$

Assume $\exists_{n\in\mathbb{N}}$ $f(x_n) > \xi$



$(f(x_n))_{n\in\mathbb{N}}$: convergent sequence in $\mathbb{R}$, take the open interval, which will contain all but a finite numbers of points of $(f(x_n))_{n\in\mathbb{N}}$. Discard those outside and rename the new (sub) sequence $f(x_n)_{n\in\mathbb{N}}$: $f(x_n) \leq \xi$ $\forall_{n\in\mathbb{N}} \Leftrightarrow \forall_{n\in\mathbb{N}}$ $x_n \in \mathrm{lev}_{\leq \xi} f$

$\Leftrightarrow (x_n)_{n\in\mathbb{N}} \subseteq C = \mathrm{lev}_{\leq \xi} f$

*/

(i)

Recall Lemma 2.37. [$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$] $\Rightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ (x_{k_n}) : weakly convergent

$\Leftrightarrow (x_n)_{n\in\mathbb{N}}$ has a weak sequential cluster point

as, $(x_n)_{n\in\mathbb{N}} \subseteq C = \mathrm{lev}_{\leq \xi} f$: bounded $\Rightarrow (x_n)_{n\in\mathbb{N}}$: has a weak sequential cluster point

$\Leftrightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ $\exists_{x\in\mathcal{H}}$ $x_{k_n} \rightharpoonup x$

as, given that $(x_n)_{n\in\mathbb{N}}$: minimizing sequence of f $\overset{\text{def}}{\Leftrightarrow}$ $f(x_n) \to \inf f(\mathcal{H})$

$\Rightarrow f(x_{k_n}) \to \inf f(\mathcal{H})$ /* As, $(f(x_{k_n}))_{n\in\mathbb{N}}$: subsequence of $\underbrace{(f(x_n))_{n\in\mathbb{N}}}_{\text{minimizing sequence}}$

for a convergent sequence, any subsequence converges to the same point */

$\therefore (x_{k_n})_{n\in\mathbb{N}}$: minimizing sequence.

/* As a result, in face any convergent subsequence of the minimizing sequence will be another minimizing subsequence */

$x_{k_n} \rightharpoonup x \in \mathcal{H}$

Recall:

*Proposition 11.20.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper lower semicontinuous quasiconvex function

$(x_{k_n})_{n\in\mathbb{N}}$: minimizing sequence of f, converges weakly to $x \in \mathcal{H}$]

$f(x) = \inf f(\mathcal{H})$

$\therefore f(x) = \inf f(\mathcal{H})$ (eq:2017-02-22 8:44 PM)

$\Leftrightarrow x$: minimizer of f by definition.

(ii)

Recall:

* Proposition 11.7. (Existence of almost one minimizer in an optimization problem)

[$f: \mathcal{H} \to]-\infty, +\infty]$, quasiconvex ✓

C: convex subset of $\mathcal{H}$, $C \cap \operatorname{dom} f \neq \emptyset$ ✓ /* there are atleast one point in C where f is finite */

/* $C = \operatorname{lev}_{\leq \xi} f$: convex, closed, $\operatorname{lev}_{\leq \xi} f \subseteq \operatorname{dom} f$ */

One of the following holds:

(i) $f + \iota_C$: strictly quasiconvex ✓

(ii) f: convex, $\underbrace{C \cap \operatorname{Argmin} f = \emptyset}_{\text{this means}}$, C: strictly convex set /* C: strictly convex set $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x,y \in C}\ x \neq y \Rightarrow \frac{x+y}{2} \in \operatorname{int} C$ */

]

$\Rightarrow$ f: has atmost one minimizer over C. ✓

so, f: has atmost one minimizer over C

in (i) we have shown the existence of minimizer

$\therefore$ f: posseses a unique minimizer

in (i) we have shown that $\forall_{(x_{k_n})_{n \in \mathbb{N}}}$: convergent subsequence of $(x_n)_{n \in \mathbb{N}}$ $x_{k_n} \rightharpoonup x$: unique minimizer of f.

recall:

because any weak sequential cluster point is the unique minimizer of f.

* lemma 2.38. ☆☆

[$(x_n)_{n \in \mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n \in \mathbb{N}}$: converges weakly $\Leftrightarrow$ $(x_n)_{n \in \mathbb{N}}$: bounded ✓, posseses at most one weak sequential cluster point ✓

$\Uparrow$

as $(x_n)_{n \in \mathbb{N}} \subseteq \operatorname{lev}_{\leq \xi} f$: bounded

$\therefore$ so, $(x_n)_{n \in \mathbb{N}}$: converges weakly

$(x_n)_{n \in \mathbb{N}}$: will also converge weakly to the minimizer of f, as otherwise the subsequence cannot go to the minimizer !

/* As if a net converges to some point, so does any subsequences Fact 1.9 */

$\therefore$ f: posseses a unique minimizer of x, $x_n \rightharpoonup x$.

(iii)

$f + \iota_C$: uniformly quasiconvex $\Rightarrow$ $f + \iota_C$: strictly quasiconvex $\Rightarrow$ f: has a unique minimizer over C.

/* uniform quasiconvexity $\Rightarrow$ strict quasiconvexity */

recall:

/*

* Definition 10.25.

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

(i) f: strictly quasiconvex $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x \in \operatorname{dom} f}\ \forall_{y \in \operatorname{dom} f}\ \forall_{\alpha \in]0,1[}\ x \neq y \Rightarrow f(\alpha x + (1-\alpha) y) < \max\{f(x), f(y)\}$

(ii) f: uniformly quasiconvex with modulus $\underbrace{\varphi}_{: \mathbb{R}_+ \to [0, +\infty]}$ $\overset{\text{def}}{\Leftrightarrow}$
- φ: increasing
- φ: vanishes only at zero
- $\forall_{x \in \operatorname{dom} f}\ \forall_{y \in \operatorname{dom} f}\ \forall_{\alpha \in]0,1[}$ $f(\alpha x + (1-\alpha) y) + \alpha(1-\alpha)\varphi(\|x-y\|) \leq \max\{f(x), f(y)\}$

*/

fix $\alpha \in]0,1[$.

as, $(x_n)_{n \in \mathbb{N}} \subseteq \operatorname{lev}_{\leq \xi} f$, $x \in \operatorname{lev}_{\leq \xi} f$ $\in \operatorname{Argmin} f$ and $\operatorname{lev}_{\leq \xi} f \subseteq \operatorname{dom} f$

So, $\forall_{n \in \mathbb{N}}\ \forall_{\alpha \in]0,1[}$ $f(\alpha x_n + (1-\alpha) x) + \alpha(1-\alpha)\varphi(\|x_n - x\|) \leq \max\{f(x_n), f(x)\} = f(x_n)$ /* as $f(x) = \inf f(\mathcal{H}) \leq f(y)\ \forall_y$ */

now, $f(x) + \alpha(1-\alpha)\varphi(\|x_n - x\|) = \inf f(\mathcal{H}) + \alpha(1-\alpha)\varphi(\|x_n - x\|)$

now, $\underbrace{f(x)}_{\inf f(H)} + \alpha(1-\alpha)\,\Phi(\|x_n - x\|) = \inf f(H) + \alpha(1-\alpha)\,\Phi(\|x_n - x\|)$

/* $\leq f(y) \quad \forall_{y \in H}$

set $y := \alpha x_n + (1-\alpha)x$ */

$\leq f(\alpha x_n + (1-\alpha)x) + \alpha(1-\alpha)\,\Phi(\|x_n - x\|)$

$\leq f(x_n)$

So. $\forall_{n\in\mathbb{N}} \; \forall_{x_n} \; \forall_{\alpha \in]0,1[} \quad f(x) + \alpha(1-\alpha)\,\Phi(\|x_n - x\|) \leq f(x_n)$

Fact 1.3.1

/* $\forall_{n\in\mathbb{N}} \; a_n \leq b_n \Rightarrow (\underline{\lim}\, a_n \leq \underline{\lim}\, b_n,\ \overline{\lim}\, a_n \leq \overline{\lim}\, b_n)$ */

$\Rightarrow \forall_{\alpha \in]0,1[}$

$$\begin{cases} \underline{\lim}\ f(x) + \alpha(1-\alpha)\,\Phi(\|x_n - x\|) \leq \underline{\lim}\ f(x_n) = \lim f(x_n) = \inf f(H) = f(x) & // \text{from} \\ \overline{\lim}\ f(x) + \alpha(1-\alpha)\,\Phi(\|x_n - x\|) \leq \overline{\lim}\ f(x_n) = f(x) \end{cases}$$

(eq:2017-02-22 8:44 PM)

$\Rightarrow \forall_{\alpha \in]0,1[} \quad \cancel{f(x)} + \underline{\lim}\ \alpha(1-\alpha)\,\Phi(\|x_n - x\|) \leq \cancel{f(x)} \;:\; \cancel{f(x)} + \overline{\lim}\ \alpha(1-\alpha)\,\Phi(\|x_n - x\|) \leq \cancel{f(x)}$

$\Rightarrow \alpha(1-\alpha)\ \underline{\lim}\ \Phi(\|x_n - x\|) \leq 0, \quad \alpha(1-\alpha)\ \overline{\lim}\ \Phi(\|x_n - x\|) \leq 0$

$\Rightarrow \quad \underline{\lim}\ \Phi(\|x_n - x\|) \leq 0, \quad \overline{\lim}\ \Phi(\|x_n - x\|) \leq 0$

but $\Phi : \mathbb{R}_+ \to [0, +\infty]$

$\therefore \quad \underline{\lim}\ \Phi(\|x_n - x\|) = \overline{\lim}\ \Phi(\|x_n - x\|) = \lim \Phi(\|x_n - x\|) = 0$

/* but Φ vanishes only at zero */

$\Rightarrow \quad \|x_n - x\| \to 0$ /* in (ii) we have shown that: $x_n \rightharpoonup x$ i.e., $(x_n - x) \to 0$ */

/* (one characterization of strong convergence using weak convergence) **

Corollary 2.42. $[(x_n)_{n\in\mathbb{N}} \subseteq H;\ z \in H]$ $x_n \to x \Leftrightarrow (x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|)$ */

$\Rightarrow \quad x_n - x \to 0 \quad$ so. $x_n \to x$

$\therefore$ f: posseses a unique minimizer x. and $x_n \to x$. ∎

*Proposition 11.27.

[

$f \in \Gamma_0(H)$

C: [bounded closed convex subset of H, $C \cap \operatorname{dom} f \neq \emptyset$

$C \cap \operatorname{Argmin} f = \emptyset$

uniformly convex set $\overset{\text{def}}{\Leftrightarrow} \exists\ \Phi : [0, \operatorname{diam} C] \to \mathbb{R}_+$ $\begin{cases} \cdot \text{ increasing function} \\ \cdot \text{ vanishes only at zero} \\ \cdot\ \forall_{x\in C}\ \forall_{y \in C}\ B\left(\frac{x+y}{2}, \Phi(\|x-y\|)\right) \subseteq C \end{cases}$ (11.10)

]

$(x_n)_{n\in\mathbb{N}}$: minimizing sequence of $f + \iota_C$

]⇒

f: has a unique minimizer of x over C, $x_n \to x$.

2/5/2016 4:00 PM **Proof:** without loss of generality, assume that $(x_n)_{n\in\mathbb{N}} \subseteq C$ /*same logic as 2/5/2016 2:41 PM */

if C = singleton ⇒ trivial

consider $C \neq$ singleton

recall

* Proposition 11.14.

[$f \in \Gamma_0(\mathcal{H})$, ✓

C: closed convex subset of $\mathcal{H}$, $C \cap \operatorname{dom} f \neq \emptyset$ ✓

One of the following holds:

(i) f: coercive

(ii) C: bounded] ⇒

f: has a minimizer over C. ✓ /*So, a coercive objective function is guranteed to provide a minimizer even when the constraint set is not bounded */

So, f has a minimizer over C.

recall:

* Proposition 11.7. (Existence of atmost one minimizer in an optimization problem)

[$f: \mathcal{H} \to]-\infty, +\infty]$, quasiconvex ✓ ($f \in \Gamma_0(\mathcal{H})$)

C: convex subset of $\mathcal{H}$, $C \cap \operatorname{dom} f \neq \emptyset$ /* there are atleast one point in C where f is finite*/

One of the following holds:

(i) ~~$f+\iota_C$: strictly quasiconvex~~

(ii) f: convex, $C \cap \operatorname{Argmin} f = \emptyset$, C: strictly convex set /* C: strictly convex set $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x,y \in C}$ $x \neq y \Rightarrow \frac{x+y}{2} \in \operatorname{int} C$ */

(C: uniformly convex /* uniform convexity implies strict convexity, both in functions and in sets */)

this means

]

⇒ f: has atmost one minimizer over C. ✓

∴ So, f: has atmost one minimizer over C.

f: has a unique minimizer over C.

Now let us show the convergence

recall:

* Proposition 11.29. ★ ($f \in \Gamma_0(\mathcal{H})$)

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, lower semicontinuous, quasiconvex ✓ ($f+\iota_C$)

$(x_n)_{n\in\mathbb{N}}$: minimizing sequence of f ✓

$\exists_{\xi \in]\inf f(\mathcal{H}), +\infty[}$ $\mathcal{L} = \operatorname{lev}_{\leq \xi} f$: bounded] ✓ /* Explanation: C: bounded $\Rightarrow \exists_{\xi \in \mathbb{R}} \forall_{x \in C}$ $f(x) \leq \xi$

$D = \operatorname{lev}_{\leq \xi} f+\iota_C = \{x \in C \mid f(x)+\iota_C(x) \leq \xi\} = \{x \in \mathcal{H} \mid f(x) \leq \xi, x \in C\}$

$= \{x \in C \mid f(x) \leq \xi\} \subseteq C$ */ → as C: bounded, D must be as well

$(f+\iota_C)(\mathcal{H})$

(i) $(x_n)_{n\in\mathbb{N}}$: has a weak sequential cluster point ⇔ $\exists_{x} \exists_{(x_{k_n})_{n\in\mathbb{N}}}$ $x_{k_n} \rightharpoonup x$

every such point is a minimizer of f. ✓

(ii) ($f+\iota_C$: strictly quasiconvex) ⇒ f: has a unique minimizer, $x_n \to x$

(iii) ($f+\iota_C$: uniformly quasiconvex) ⇒ f: has a unique minimizer, $x_n \to x$

∴ $(x_n)_{n\in\mathbb{N}}$: has a weak sequential cluster point, where such point is a unique minimizer of $f+\iota_C$

⇒ x_n : x_n has unique weak sequential cluster point which happens to be the unique minimizer x of $f+\iota_C$.

using

* Lemma 2.38. ★★

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$] ($(x_n)_{n\in\mathbb{N}} \subseteq C$: bounded ✓)

$(x_n)_{n\in\mathbb{N}}$: converges weakly ⇔ $(x_n)_{n\in\mathbb{N}}$: bounded, possesses at most one weak sequential cluster point. ✓

we have:

$x_n \to x$.

$C \cap \operatorname{Argmin} f = \emptyset$

but, $x \in C$, $x \in \operatorname{Argmin}(f+\iota_C)$, $x \notin \operatorname{Argmin} f$

recall:

* Proposition 11.5. /* The statement is

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

C: subset of $\mathcal{H}$

x: minimizer of f over C, $x \in \operatorname{int} C$] → using contrapositive

⇒ x: minimizer of f.

$x \notin \operatorname{int} C$, but $x \in C$

∴ $x \in C \setminus \operatorname{int} C = \operatorname{bdry} C$

definition of uniform convexity tells us that $\operatorname{int} C \neq \emptyset$

/* recall:

* Corollary 7.6. [C: nonempty closed convex subset of $\mathcal{H}$

one of the following holds:

(i) $\operatorname{int} C \neq \emptyset$ ✓

(ii) C: closed affine subspace

* Corollary 7.6. [C: nonempty closed convex subset of $\mathcal{H}$
one of the following holds:
(i) $\operatorname{int} C \neq \emptyset$
(ii) C: closed affine subspace
(iii) $\mathcal{H}$: finite dimensional] $\Rightarrow$ $\operatorname{spts} C = \operatorname{bdry} C$

*/

$\therefore x \in \operatorname{spts} C$

denote u: $\|u\|=1$ which is the associated normal vector u A [C: nonempty subset of $\mathcal{H}$; $x \in C$; $u \in \mathcal{H} \setminus \{0\}$]

x: support point of C with normal vector u $\overset{\text{def}}{\Leftrightarrow}$ $\langle x|u\rangle \geq \sup\langle C|u\rangle$ */

$\forall_{n\in\mathbb{N}}$ $z_n := \frac{x_n+x}{2} + \Phi(\|x_n-x\|)u$

now, $x_n \in C$, $x \in C$ then

$$z_n \overset{?}{\in} B\left(\frac{x_n+x}{2}; \Phi(\|x_n-x\|)\right) \subseteq C$$

$$\Leftrightarrow \left\| \underbrace{z_n}_{} - \frac{x_n+x}{2} \right\| \overset{?}{\leq} \Phi(\|x_n-x\|)$$

$$= \frac{x_n+x}{2} + \Phi(\|x_n-x\|)u$$

$$\Leftrightarrow \quad \|\Phi(\|x_n-x\|)u\| = \underbrace{\|u\|}_{=1} \|\Phi(\|x_n-x\|)\| \overset{\checkmark}{\leq} \Phi(\|x_n-x\|)$$

$\therefore z_n \in C$

RECALL

* Definition 7.1. (Support point)
[C: nonempty subset of $\mathcal{H}$; $x \in C$; $u \in \mathcal{H} \setminus \{0\}$]
x: support point of C with normal vector u $\overset{\text{def}}{\Leftrightarrow}$ $\langle x|u\rangle \geq \sup\langle C|u\rangle$
supporting hyperplane of C at support point $x = \{y \in \mathcal{H} \mid \langle y|u\rangle = \langle x|u\rangle\}$
$\operatorname{spts} C$ = set of support points of C,
$\overline{\operatorname{spts} C}$ = set of closure of $\operatorname{spts} C$

as $x \in \operatorname{spts} C$, u: $\|u\|=1 \neq \{0\}$ we have:

$$\langle x|u\rangle \geq \sup\langle C|u\rangle \quad \begin{matrix} = \langle x|u\rangle \\ = \sup\langle C|u\rangle \end{matrix}$$

$$\Rightarrow \forall_{\tilde{x}\in C} \quad \langle x|u\rangle \geq \langle \tilde{x}|u\rangle$$

$$\because z_n \in C \quad \therefore \quad \langle x|u\rangle \geq \langle \underbrace{z_n}_{\frac{x_n+x}{2}+\Phi(\|x_n-x\|)u}|u\rangle = \left\langle \frac{x_n+x}{2} + \Phi(\|x_n-x\|)u \,\middle|\, u\right\rangle$$

$$\Rightarrow \left\langle \frac{x_n+x}{2} + \Phi(\|x_n-x\|)u \,\middle|\, u\right\rangle - \langle x|u\rangle \leq 0$$

$$\Leftrightarrow \left\langle \frac{x_n+x}{2} + \Phi(\|x_n-x\|)u - x \,\middle|\, u\right\rangle = \left\langle \frac{x_n-x}{2} + \Phi(\|x_n-x\|)u \,\middle|\, u\right\rangle \quad /* \ \langle x+y|z\rangle = \langle x|z\rangle + \langle y|z\rangle \ */$$

$$= \left\langle \frac{x_n-x}{2} \,\middle|\, u\right\rangle + \underbrace{\Phi(\|x_n-x\|)}_{\text{number}} \underbrace{\langle u|u\rangle}_{=\|u\|^2=1} \quad /* \ \langle cx|y\rangle = c\langle x|y\rangle \ */$$

$$\leq 0$$

$$\Leftrightarrow \left\langle \frac{x_n-x}{2} \,\middle|\, u\right\rangle + \Phi(\|x_n-x\|) \leq 0$$

$$\Leftrightarrow \quad \Phi(\|x_n-x\|) \leq \left\langle \frac{x-x_n}{2} \,\middle|\, u\right\rangle$$

So, $\forall_n \quad \Phi(\|x_n-x\|) \leq \left\langle \frac{x-x_n}{2} \,\middle|\, u\right\rangle = \frac{1}{2}\langle x-x_n|u\rangle$

But $x_n \to x \overset{\text{def}}{\Leftrightarrow} \forall_u \langle x_n - x|u\rangle \to 0$

$\Phi(\|x_n-x\|) \to 0$ /* Φ: vanishes only at 0 */

$\Leftrightarrow \|x_n-x\| \to 0$

$\Leftrightarrow x_n \to x$.

☺

Part 1

Saturday, February 6, 2016 3:36 PM

*Proposition 12.6.

[[$f, g, h : \mathcal{H} \to]-\infty, +\infty]$]] $\Rightarrow$

(i) f, g : have continuous affine minorant with slope $u \in \mathcal{H}$

$\Rightarrow f \square g$: has " " " " " $u \in \mathcal{H}$, $-\infty \notin (f \square g)(\mathcal{H})$

(ii) $\operatorname{dom} f \square g = \operatorname{dom} f + \operatorname{dom} g$

(iii) $f \square g = g \square f$

2/7/2016 7:39 AM

Proof:

(i) f : has continuous affine minorant with slope $u \in \mathcal{H} \overset{\text{def}}{\Leftrightarrow} \exists_{\eta \in \mathbb{R}} \quad f - \langle \cdot | u \rangle \geq \eta > -\infty \Leftrightarrow \exists_{\eta \in \mathbb{R}} \forall_{\tilde{x} \in \mathcal{H}} \quad f(\tilde{x}) - \langle \tilde{x} | u \rangle \geq \eta$

g : " " " " " " $u \in \mathcal{H} \overset{\text{def}}{\Leftrightarrow} \exists_{\mu \in \mathbb{R}} \quad g - \langle \cdot | u \rangle \geq \mu > -\infty \Leftrightarrow \exists_{\mu \in \mathbb{R}} \forall_{\tilde{y} \in \mathcal{H}} \quad g(\tilde{y}) - \langle \tilde{y} | u \rangle \geq \mu$

$\tilde{x} := y$

$$\left.\begin{array}{l} f(y) - \langle y | u \rangle \geq \eta \\ \tilde{y} := x - y \\ g(x-y) - \langle x - y | u \rangle \geq \mu \end{array}\right\} \oplus \Rightarrow f(y) + g(x-y) \underbrace{- \langle y | u \rangle - \langle x - y | u \rangle}_{\substack{-\langle y|u\rangle - \langle x|u\rangle + \langle y|u\rangle \\ = -\langle x|u\rangle}} \geq \eta + \mu > -\infty$$

$$f(y) + g(x-y) - \langle x | u \rangle \geq \eta + \mu > -\infty$$

now y : arbitrary. x : arbitrary

$$\therefore \forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}} \quad f(y) + g(x-y) - \langle x | u \rangle \geq \eta + \mu > -\infty$$

$$\Leftrightarrow \forall_{x \in \mathcal{H}} \quad \underbrace{\inf_{y \in \mathcal{H}} \left(f(y) + g(x-y) \right)}_{(f \square g)(x)} \underbrace{- \langle x | u \rangle}_{\text{constant w.r.t } y} \geq \eta + \mu > -\infty$$

$$\therefore \forall_{x \in \mathcal{H}} \quad (f \square g)(x) - \langle x | u \rangle \geq \eta + \mu > -\infty$$

$$\Leftrightarrow \quad (f \square g) - \langle \cdot | u \rangle \geq \eta + \mu > -\infty$$

So, $f \square g$: has a continuous affine minorant.

$$\Rightarrow \forall_{x \in \mathcal{H}} \quad (f \square g)(x) \geq (\eta + \mu) + \langle x | u \rangle > -\infty$$

$$\Rightarrow \quad -\infty \notin (f \square g)(\mathcal{H}) \quad \text{(i)}$$

(ii), (iii)

$$f \square g : \mathcal{H} \to [-\infty, +\infty] : x \mapsto \inf_{y \in \mathcal{H}} \left(f(y) + g(x-y) \right)$$

$$(f \square g)(x) = \inf_{u \in \mathcal{H}} \quad f(u) + g(x-u)$$

$$= \inf_{u} \quad f(u) + g(\underbrace{x-u}_{v})$$

$$\text{s.t.} \quad u \in \mathcal{H}$$

v : new decision vector $= x - u$

$$= \left(\begin{array}{ll} \inf_{u,v} & f(u) + g(v) \\ \text{s.t.} & u + v = x \\ & u, v \in \mathcal{H} \end{array} \right) \rightsquigarrow (\text{eq}, 1)$$

note that this inf problem stays same if we switch f and g, so,

$$(f \square g)(x) = (g \square f)(x) \quad \forall_{x \in \mathcal{H}}$$

$\Leftrightarrow f \square g = g \square f$ (iii) proved

$(f\square g)(x) = (g\square f)(x) \quad \forall_{x\in\mathcal{H}}$

$\Leftrightarrow \quad f\square g = g\square f \qquad$ (iii) proved

(ii): want to prove: $\underbrace{\mathrm{dom}(f\square g) = \mathrm{dom} f + \mathrm{dom} g}_{// \; x\in \mathrm{dom} f + \mathrm{dom} g \Leftrightarrow \exists_{u,v:\, x=u+v} \; u\in\mathrm{dom} f,\; v\in \mathrm{dom} g}$ $\quad // \; (\mathrm{dom} f + \mathrm{dom} g)(x) = (\mathrm{dom} f)(x) + (\mathrm{dom} g)(x)$

$\forall_x$

$x\in \mathrm{dom}(f\square g) \Leftrightarrow (f\square g)(x) = \begin{pmatrix} \inf\limits_{u,v\in\mathcal{H}} & f(u)+g(v) \\ \text{s.t.} & u+v=x \end{pmatrix} < +\infty \quad$ // from (eq 1)

$\Rightarrow \forall_{\substack{u,v\in\mathcal{H}: \\ u+v=x}} \quad f(u)+g(v) < +\infty \quad$ // as $f,g,h : \mathcal{H} \to]-\infty,+\infty]$, this implies : $f(u)<+\infty$, $f(v)<+\infty$

$\Rightarrow \forall_{u,v\in\mathcal{H}:\, u+v=x} \quad f(u)<+\infty,\; g(v)<+\infty$

$\Leftrightarrow \forall_{u,v\in\mathcal{H}:\, u+v=x} \quad u\in\mathrm{dom} f,\; v\in\mathrm{dom} g$

$\Rightarrow \exists_{u,v\in\mathcal{H}:\, u+v=x} \quad u\in\mathrm{dom} f,\; v\in\mathrm{dom} g \Leftrightarrow x\in\mathrm{dom} f+\mathrm{dom} g \quad \therefore \; \mathrm{dom}(f\square g) \subseteq \mathrm{dom} f + \mathrm{dom} g \quad (1)$

now: $\forall_x$

$x\in\mathrm{dom} f+\mathrm{dom} g \Leftrightarrow \exists_{\tilde{u},\tilde{v}:\, x=\tilde{u}+\tilde{v}} \quad \tilde{u}\in\mathrm{dom} f,\; \tilde{v}\in\mathrm{dom} g$

$\qquad f(\tilde{u})<+\infty \qquad g(\tilde{v})<+\infty \qquad$ // as $f,g,h : \mathcal{H}\to]-\infty,+\infty]$, this implies : $f(u)+g(v)<+\infty$

$\Rightarrow \exists_{\tilde{u},\tilde{v}:\, x=\tilde{u}+\tilde{v}} \quad f(\tilde{u})+g(\tilde{v}) < +\infty$

now by definition: $\left.\begin{pmatrix} \inf\limits_{u,v\in\mathcal{H}:} & f(u)+g(v) \\ u+v=x & \end{pmatrix} = (f\square g)(x) < f(\tilde{u})+g(\tilde{v})\right\} \Rightarrow (f\square g)(x) < +\infty \Leftrightarrow x\in\mathrm{dom}(f\square g)$

// from eq 1

$\therefore \; \mathrm{dom} f + \mathrm{dom} g \subseteq \mathrm{dom}(f\square g) \quad (2)$

Thus from (1), (2) : $\mathrm{dom}(f\square g) = \mathrm{dom} f + \mathrm{dom} g$. ☺

(v) f,g,h: continuous affine minorants with slope 4

$\Rightarrow f\square g,\; g\square h : -\infty \notin \mathrm{dom}\, f\square g,\; -\infty\notin \mathrm{dom}\, g\square h$ /* from (i) */

so. $f\square(g\square h), (f\square g)\square h$: well defined

now $(g\square h)(y) = \begin{array}{ll} \inf\limits_{v,w} & g(v)+h(w) \\ \text{s.t.} & v+w=y \\ & v,w\in\mathcal{H} \end{array}$

now. $f\square(g\square h)x = \begin{array}{ll} \inf\limits_{u,y\in\mathcal{H}} & f(u)+(g\square h)(y) \\ \text{s.t.} & u+y=x \end{array}$

constant w.r.t v,w so we can take it inside

$= \inf\limits_{u,y\in\mathcal{H}} \left(\left(f(u)+\iota_{u+y=x}(u,y)\right) + \inf\limits_{v,w} \left(g(v)+h(w)+\iota_{v+w=y}(v,w)\right)\right)$

$= \inf\limits_{u,y\in\mathcal{H}} \; \inf\limits_{v,w} \left[f(u)+g(v)+h(w) + \underbrace{\iota_{u+y=x}(u,y)+\iota_{v+w=y}(v,w)}_{\iota_{u+v+w=x}(u,v+w)} \right]$

$= \inf\limits_{u,v+w\in\mathcal{H}} \; \inf\limits_{v,w} \left[f(u)+g(v)+h(w)+\iota_{u+v+w=x}(u,v+w)\right]$

$$= \inf_{u,v,w} \; f(u)+g(v)+h(w) \quad \text{s.t.} \quad u+v+w=x, \; u,v,w \in \mathcal{H}$$

$= ((f \square g) \square h)(x)$ // as: if we started with $((f \square g) \square h)(x)$ we would end up with this $(\cdot)$ term.

■

2/8/2016 6:55 PM

*Corollary 12.19.

[C: nonempty subset of $\mathcal{H}$

$h: C \to \mathbb{R}$, β-Lipschitz continuous for some $\beta \in \mathbb{R}_+$

$f: \mathcal{H} \to]-\infty, +\infty] : x \mapsto \begin{cases} h(x), & \text{if } x \in C \\ +\infty, & \text{otherwise} \end{cases}$

]

$f \square \beta\|\cdot\|$: β-Lipschitz continuous extension of h

Proof:

$g = f \square \beta\|\cdot\|$

$\forall_{x \in C} \quad g(x) = \inf_{y \in \mathcal{H}} \left(f(y) + \beta\|x-y\| \right)$

$= \inf_{y \in \mathcal{H}} \left(h(y)[y \in C] + \infty[y \notin C] + \beta\|x-y\| \right)$

$= \inf_{y \in C} \left(h(y) + \beta\|x-y\| \right)$ /* h: β-Lipschitz continuous on C
$\Rightarrow \|h(x)-h(y)\| \leq \beta\|x-y\| \Rightarrow h(x)-h(y) \leq \|h(x)-h(y)\| \leq \beta\|x-y\|$ */

$\geq \inf_{y \in C} \left(h(y) + \|h(x)-h(y)\| \right)$

$\geq \inf_{y \in C} \left(h(y) + h(x) - h(y) \right)$ /* $\forall_{a \in \mathbb{R}} \; \|a\| \geq a \Leftrightarrow b + \|a\| \geq b + a$ */

$= \inf_{y \in C} \underbrace{h(x)}_{\text{const w.r.t. } y} = h(x) > -\infty$ /* As, $\forall_{x \in C} \; h(x) \in \mathbb{R}$ ($\because h: C \to \mathbb{R}$)
$\Rightarrow \forall_{x \in C} \; h(x) > -\infty$ */

/* So we have, $\forall_{x \in C} \; g(x) \geq h(x)$ */

$\therefore \forall_{x \in C} \; g(x) > -\infty$

Recall that.

* Proposition 12.17.
[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, $\beta \in \mathbb{R}_+$
g: β-Pasch-Hausdorff envelope of f]
Exactly one of the following holds:
(i) f: possesses a β-Lipschitz continuous minorant.
g: largest β-Lipschitz continuous minorant of f.
(ii) f: posses no β-Lipschitz continuous minorant, $g \equiv -\infty$ → so (ii) doesnot hold

as (ii) doesnot hold, (i) must hold

so, f: posses a β-Lipschitz continuous minorant,

g: largest β-Lipschitz continuous minorant of f. $\therefore \forall_{x \in \mathcal{H}} \; f(x) \geq g(x)$

$\Rightarrow \forall_{x \in C} \; g(x) \leq f(x)$

$\forall_{x \in C} \; g(x) = f(x) = h(x)$ /* as on C, $f = h$ */

as g: β-Lipschitz continuous in general /* $\forall_{x,y \in \mathcal{H}} \; \|g(x)-g(y)\| \leq \beta\|x-y\|$ */

$g = f \square (\beta\|\cdot\|)$: β-Lipschitz continuous extension of h.

■

*Proposition 12.17.

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

$\beta \in \mathbb{R}_+$

g: β-Pasch-Hausdorff envelope of f /* $g = f \square \underbrace{\beta}_{\in \mathbb{R}_+}\|\cdot\|$ */

1

Exactly one of the following holds:

(i) f: posseses a β-Lipschitz continuous minorant, in this case
g: largest β-Lipschitz continuous minorant of f

(ii) f: posses no β-Lipschitz continuous minorant, in this case $g \equiv -\infty$

2/9/2016 9:43 AM

Proof: it is true that no matter what
Exactly one of (i) or (ii) is apparent, as f either has a β-Lipschitz continuous minorant or it does not.

g: β-Pasch-Hausdorff envelope of f

$\Leftrightarrow$

$$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad g(x) = \inf_{z\in\mathcal{H}} \underbrace{\left(f(z) + \beta\|x-z\|\right)}_{\substack{f(z)+\beta\|x-y+y-z\| \\ \leq f(z)+\beta\|x-y\|+\beta\|y-z\| \ /*\text{using triangle inequality}*/}} \overset{\text{constant w.r.t } z}{\leq} \inf_{z\in\mathcal{H}} \left(f(z)+\beta\|x-y\|+\beta\|y-z\|\right) = \beta\|x-y\| + \underbrace{\inf_{z\in\mathcal{H}}\left(f(z)+\beta\|y-z\|\right)}_{g(y)} = \overbrace{\beta\|x-y\|}^{\text{finite}} + g(y)$$

as power norm infimal convolution for a general function has full domain, i.e.,

(Power norm infimal convolution of a general function and its properties)

* Proposition 12.9.

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

$p \in [1,+\infty[$

$\forall_{\gamma\in\mathbb{R}_{++}} \quad g_\gamma = f \,\square\, \left(\frac{1}{\gamma p}\|\cdot\|^p\right)$: power norm infimal convolution]

$\forall_{\gamma\in\mathbb{R}_{++}} \forall_{x\in\mathcal{H}}$
- $\operatorname{dom} g_\gamma = \mathcal{H}$
- $\mu \in]\gamma,+\infty[\Rightarrow \inf f(\mathcal{H}) \leq g_\mu(x) \leq g_\gamma(x) \leq f(x)$
- $\inf g_\gamma(\mathcal{H}) = \inf f(\mathcal{H})$
- $g_\mu(x) \downarrow \inf f(\mathcal{H})$ as $\mu \uparrow +\infty$
- g_γ: bounded on every ball in $\mathcal{H}$

As β-Pasch-Hausdorff envelope satisfies all the properties of power norm infimal convolution,

$\operatorname{dom} g = \mathcal{H} \Leftrightarrow \forall_{x\in\mathcal{H}} \; g(x) < +\infty$

$$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad g(x) = \overbrace{\beta\|x-y\|}^{\text{finite}} + \overbrace{g(y)}^{<+\infty} < +\infty \quad \ldots \quad (12.15)$$

it is true that no matter what
Exactly one of (i) or (ii) is apparent, as f either has a β-Lipschitz continuous minorant or it does not.

(i) Now consider the first possibility:

f: posseses a β-Lipschitz continuous minorant

f: Posses a β-Lipschitz continuous minorant $h: \mathcal{H} \to \mathbb{R} \Leftrightarrow$

- h satisfies
  $\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \; |h(x)-h(y)| \leq \beta\|x-y\| \Rightarrow h(x) \leq h(y) + \beta\|x-y\|$
- h: minorant of f
  $\Leftrightarrow \forall_{\tilde{x}\in\mathcal{H}} \; h(\tilde{x}) \leq f(\tilde{x})$

$$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad h(x) \leq h(y) + \beta\|x-y\| \leq f(y) + \beta\|x-y\|$$

$$\Leftrightarrow \forall_{x\in\mathcal{H}} \quad h(x) \leq \inf_{y\in\mathcal{H}} \left(f(y) + \beta\|x-y\|\right)$$

$$= \left(f \,\square\, (\beta\|\cdot\|)\right)(x)$$

$\forall_{x\in H}$ $\quad \forall_{y\in H}$

$$= \underbrace{\left(f \,\square\, (\beta\|\cdot\|)\right)}_{g}(x)$$

$$= g(x)$$

$\therefore \forall_{x\in H}\ h(x) \leq g(x) < +\infty$

but, $h : H \to \mathbb{R}$, $\not\ni \pm\infty$

so, $\forall_{x\in H}\ -\infty < h(x) \leq g(x) < +\infty$

so, $g : H \to \mathbb{R}$,

also, $\forall_{x\in H}\ \forall_{y\in H}\ g(x) - g(y) = \beta\|x-y\|$ (from (12.15))

similarly, swapping x and y:

$\forall_{x\in H}\ \forall_{y\in H}\ g(y) - g(x) = \beta\|x-y\|$

$\Rightarrow \forall_{x\in H}\ \forall_{y\in H}\ |g(x) - g(y)| = \beta\|x-y\| \Rightarrow g$: β-Lipschitz-continuous

so, we have: $h \leq g \leq f$, g: β-Lipschitz continuous minorant

/* any β-Lipschitz continuous minorant */

/* also, Proposition 12.8(ii) : says $g \leq f$ */ so, g: largest β-Lipschitz continuous minorant of f

(ii)

f: possesses no β-Lipschitz continuous minorant but per absurdum g be real valued at some point in H, say the point is $\bar{y}$

now (12.15) says

$\forall_{x\in H}\ \forall_{y\in H}\ g(x) - g(y) \leq \beta\|x-y\|$

$\forall_{y\in H}\ \forall_{x\in H}\ g(y) - g(x) \leq \beta\|x-y\|$

$\Rightarrow \forall_{x\in H}\ \forall_{y\in H}\ |g(x) - g(y)| \leq \beta\|x-y\|$

so, g: β-Lipschitz continuous

set $y := \bar{y}$

$$\Rightarrow \forall_{x\in H}\ \underbrace{|g(x) - \underbrace{g(\bar{y})}_{\text{finite}}|}_{\geq |g(x)| - |g(\bar{y})|} \leq \underbrace{\beta\|x-\bar{y}\|}_{\text{finite}}$$

/* using reverse triangle inequality

$\|a-b\| + \|b\| \geq \|a-b+b\| = \|a\|$

$\Leftrightarrow \|a-b\| \geq \|a\| - \|b\|$ */

$$\Rightarrow \forall_{x\in H}\ |g(x)| - \underbrace{|g(\bar{y})|}_{\text{finite}} \leq |g(x) - g(\bar{y})| \leq \beta\|x-\bar{y}\|$$

$$\therefore \forall_{x\in H}\ |g(x)| \leq \underbrace{\beta\|x-\bar{y}\| + |g(\bar{y})|}_{\text{finite}}$$

so, g: everywhere real valued

$\therefore g$: real valued everywhere and β-Lipschitz continuous.

now suppose $\beta = 0$. then by definition

$$g = f \,\square\, \underbrace{\beta}_{0}\|\cdot\|$$

$$\therefore \forall_{x\in H}\ g(x) = \inf_{y\in H}\left(f(y) + \overbrace{\beta}^{0}\|x-y\|\right) = \inf_{y\in H} f(y) = \inf f(H)$$

$$\Leftrightarrow \forall_{x\in H}\ \forall_{y\in H}\ g(x) \leq f(y)$$

set $y := x \Rightarrow$

$$\forall_{x\in H}\ g(x) \leq f(x)$$

$\Leftrightarrow g \leq f \Leftrightarrow g$: β-Lipschitz continuous minorant of $f \Rightarrow$ contradiction

$\forall_{x \in H} \quad g(x) \leq f(x)$

$\leftrightarrow \quad \underbrace{g}_{\substack{\text{real valued everywhere} \\ \beta\text{- Lipschitz continuous} \\ 0}} \leq f \leftrightarrow g$: β-Lipschitz continuous minorant of $f \Rightarrow$ contradiction

now consider $\beta > 0$, then

$$g = f \,\square\, \overset{>0}{\beta} \|\cdot\|$$

$$= f \,\square\, \frac{1}{(1/\beta)\,1} \|\cdot\|^1$$: this is power norm infimal convolution with 1 norm, so we can apply Proposition 12.9 (see above) property 2 which says

$$= g_{1/\beta}$$

$$\forall_{x \in H} \quad \underbrace{g_{1/\beta}}_{g}(x) \leq f(x)$$

$$\leftrightarrow \quad g \leq f$$

$\therefore$ g: β-Lipschitz continuous minorant of $f \Rightarrow$ contradiction

$\therefore$ g cannot be real valued anywhere, but $\operatorname{dom} g = H$ (Proposition 12.9 (i))

$\forall_{x \in H} \quad g(x) < +\infty$

$\therefore \forall_{x \in H} \quad g(x) = -\infty \;\&\; g \equiv -\infty.$

$\blacksquare$

Part 2

Wednesday, February 10, 2016 7:30 AM

2/10/2016 8:28 AM

* Proposition 12.22:

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

$\gamma \in \mathbb{R}_{++}$

$\mu \in \mathbb{R}_{++}$]

(i) ${}^{\mu}(\gamma f) = \gamma\,({}^{\gamma\mu} f)$

(ii) $\gamma\,({}^{\mu} f) = {}^{(\gamma+\mu)} f$

Proof:

Take $x \in \mathcal{H}$

(i) $\gamma f = f \,\square\, \frac{1}{2\gamma}\|\cdot\|^2$

$$
\begin{aligned}
{}^{\mu}(\gamma f)(x) &= \left(\gamma f \;\square\, \tfrac{1}{2\mu}\|\cdot\|^2\right)(x) \\
&= \inf_{y \in \mathcal{H}} \left(\gamma f(y) + \frac{1}{2\mu}\|x-y\|^2\right) \\
&= \gamma \underbrace{\inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2\gamma\mu}\|x-y\|^2\right)}_{\left(f \,\square\, \frac{1}{2(\gamma\mu)}\|\cdot\|^2\right)(x) = {}^{\gamma\mu}(f)} \\
&= \gamma\,({}^{\gamma\mu} f)(x)
\end{aligned}
$$

$\therefore\ {}^{\mu}(\gamma f) = \gamma\,({}^{\gamma\mu} f)$

(ii)

$$
\begin{aligned}
&\gamma\,({}^{\mu} f)(x) \\
&= \left({}^{\mu} f \,\square\, \tfrac{1}{2\gamma}\|\cdot\|^2\right)(x) \\
&= \inf_{z \in \mathcal{H}} \Big(\underbrace{{}^{\mu} f(z)}_{\left(f \,\square\, \frac{1}{2\mu}\|\cdot\|^2\right)(z) = \inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2\mu}\|z-y\|^2\right)} + \frac{1}{2\gamma}\|z-x\|^2\Big) \\
&= \inf_{z \in \mathcal{H}} \left(\inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2\mu}\|z-y\|^2\right) + \underbrace{\frac{1}{2\gamma}\|z-x\|^2}_{\text{constant w.r.t } y}\right) \\
&= \inf_{z \in \mathcal{H}} \inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2\mu}\|z-y\|^2 + \frac{1}{2\gamma}\|z-x\|^2\right) \quad /\!\!* \ \inf_{(a,b) \in \mathcal{H} \times \mathcal{H}} = \inf_{(b,a) \in \mathcal{H} \times \mathcal{H}} \ *\!\!/ \\
&= \inf_{y \in \mathcal{H}} \inf_{z \in \mathcal{H}} \left(f(y) + \frac{1}{2\mu}\|z-y\|^2 + \frac{1}{2\gamma}\|z-x\|^2\right) \\
&= \inf_{y \in \mathcal{H}} f(y) + \inf_{z \in \mathcal{H}} \left(\frac{1}{2\mu}\|z-y\|^2 + \frac{1}{2\gamma}\|z-x\|^2\right)
\end{aligned}
$$

A trick worth remembering

/* define $\alpha = \frac{\mu}{\mu+\gamma} \in (0,1)$, $1-\alpha = \frac{\gamma}{\mu+\gamma} \in (0,1)$

then $\frac{\alpha}{2\alpha\gamma} = \frac{1}{2\gamma}$, $\frac{1-\alpha}{2\alpha\gamma} = \frac{\cancel{\gamma}}{\cancel{\mu+\gamma}} \cdot \frac{1}{2 \cdot \frac{\mu}{\cancel{\mu+\gamma}} \cdot \cancel{\gamma}} = \frac{1}{2\mu}$ */

$$
\begin{aligned}
&= \inf_{z \in \mathcal{H}} \frac{(1-\alpha)}{2\alpha\gamma}\|z-y\|^2 + \frac{\alpha}{2\alpha\gamma}\|z-x\|^2 \\
&= \frac{1}{2\alpha\gamma} \inf_{z \in \mathcal{H}} \underbrace{(1-\alpha)\|z-y\|^2 + \alpha\|z-x\|^2}_{\substack{\|(1-\alpha)(z-y) + \alpha(z-x)\|^2 + \\ \alpha(1-\alpha)\|z-x-z+y\|^2 \\ = \|(1-\alpha)z - \alpha(z-y) + \alpha z - \alpha x\|^2 + \alpha(1-\alpha)\|x-y\|^2 \\ = \|z - (\alpha x + (1-\alpha)y)\|^2 + \alpha(1-\alpha)\|x-y\|^2}}
\end{aligned}
$$

/* recall, corollary 2.14:

$\|\underbrace{\alpha}_{\in \mathbb{R}} x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2$ */

$$
\begin{aligned}
&= \frac{1}{2\alpha\gamma} \inf_{z \in \mathcal{H}} \left(\|z - (\alpha x + (1-\alpha)y)\|^2 + \alpha(1-\alpha)\|x-y\|^2\right) \\
&= \frac{1}{2\alpha\gamma}\alpha(1-\alpha)\|x-y\|^2 + \frac{1}{2\alpha\gamma} \underbrace{\inf_{z \in \mathcal{H}} \|z - (\alpha x + (1-\alpha)y)\|^2}_{0 \ \text{/* by setting } z := \alpha x + (1-\alpha)y \ *\!\!/} \\
&= \frac{1-\alpha}{2\gamma}\|x-y\|^2 \quad /\!\!* \ (1-\alpha) = \frac{\gamma}{\gamma+\mu} \ *\!\!/ \\
&= \frac{\cancel{\gamma}}{(\gamma+\mu)} \cdot \frac{1}{2\cancel{\gamma}}\|x-y\|^2 = \frac{1}{2(\gamma+\mu)}\|x-y\|^2
\end{aligned}
$$

$$
\begin{aligned}
&= \inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2(\gamma+\mu)}\|x-y\|^2\right) \\
&= \left(f \,\square\, \frac{1}{2(\gamma+\mu)}\|\cdot\|^2\right)(x) \\
&= {}^{(\gamma+\mu)} f\,(x)
\end{aligned}
$$

$\therefore\ \gamma\,({}^{\mu} f) = {}^{(\gamma+\mu)} f$

∎

$= \cdots f(x)$

$\therefore \gamma(\mu f) = (\gamma+\mu) f$

■

Proposition 12.27.

[$f \in \Gamma_0(\mathcal{H})$]

prox_f : firmly nonexpansive

$\mathrm{Id} - \mathrm{prox}_f$: firmly nonexpansive

2/11/2016 10:17 PM

Proof:

$\forall_{x,y \in \mathcal{H}}$

$\begin{cases} p = \mathrm{prox}_f x \\ q = \mathrm{prox}_f y \end{cases}$

Proposition 12.26. (Defining properties of prox operator)

[$f \in \Gamma_0(\mathcal{H})$; $x, p \in \mathcal{H}$]

$p = \mathrm{prox}_f x \Leftrightarrow \forall_{y \in \mathcal{H}} \langle y-p \mid x-p \rangle + f(p) \leq f(y)$

$z = \mathrm{prox}_f v \Leftrightarrow \forall_{\pi \in \mathcal{H}} \langle \pi - z \mid v - z \rangle + f(z) \leq f(\pi)$

$\forall_{\pi \in \mathcal{H}} \langle \pi - p \mid x-p \rangle + f(p) \leq f(\pi) \overset{\pi := q}{\Rightarrow} \langle q-p \mid x-p \rangle + f(p) \leq f(q)$

$\forall_{\xi \in \mathcal{H}} \langle \xi - q \mid y-q \rangle + f(q) \leq f(\xi) \overset{\xi := p}{\Rightarrow} \langle p-q \mid y-q \rangle + f(q) \leq f(p)$

$\oplus$

$\langle q-p \mid x-p \rangle + \langle p-q \mid y-q \rangle + \cancel{f(p)} + \cancel{f(q)} \leq \cancel{f(p)} + \cancel{f(q)}$

/* this cancellation is valid as $p, q \in \mathrm{dom}\, f$

$p = \mathrm{prox}_f(x) \Leftrightarrow \forall_{z \in \mathcal{H}} \langle z-p \mid x-p \rangle + f(p) \leq f(z)$

take $z \in \mathrm{dom}\, f \setminus \{p\}$ then.

$\underbrace{\langle z-p \mid x-p \rangle}_{\text{finite}} + f(p) \leq \underbrace{f(z)}_{\text{finite}}$

$\Rightarrow f(p) \leq \text{finite} < +\infty \Rightarrow p \in \mathrm{dom}\, f$

similarly $q \in \mathrm{dom}\, f$

*/

$\Leftrightarrow -\langle p-q \mid x-p \rangle + \langle p-q \mid y-q \rangle \leq 0$

$\Leftrightarrow -\langle p-q \mid (x-p)-(y-q) \rangle \leq 0$

$\Leftrightarrow \langle p-q \mid (x-p)-(y-q) \rangle \geq 0$

/* denote $\mathrm{prox}_f \equiv T$ */

$\Leftrightarrow \langle Tx - Ty \mid \underbrace{(x-Tx)}_{(2\mathrm{Id}-T)x} - \underbrace{(y-Ty)}_{(2\mathrm{Id}-T)y} \rangle \geq 0$

$\therefore \forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}} \langle Tx - Ty \mid (2\mathrm{Id}-T)x - (2\mathrm{Id}-T)y \rangle \geq 0$

$\Leftrightarrow$ T: firmly nonexpansive

$\Leftrightarrow$ $(2\mathrm{Id}-T)$: firmly nonexpansive

/* Equivalence follows from the theorem.

Proposition 4.2. (Defining properties of firmly nonexpansive operator)

[E: nonempty subset of $\mathcal{H}$.

$T: E \to \mathcal{H}$]

(i) T: firmly nonexpansive $\Leftrightarrow$ (ii) $(\mathrm{Id}-T)$: firmly nonexpansive $\Leftrightarrow$ (iii) $(2T-\mathrm{Id})$: nonexpansive

$\Leftrightarrow$ (iv) $\forall_{x \in E} \forall_{y \in E} \|Tx-Ty\|^2 \leq \langle x-y \mid Tx-Ty \rangle$

$\Leftrightarrow$ (v) $\forall_{x \in E} \forall_{y \in E} \langle Tx-Ty \mid (\mathrm{Id}-T)x-(\mathrm{Id}-T)y \rangle \geq 0$

$\Leftrightarrow$ (vi) $\forall_{x \in E} \forall_{y \in E} \forall_{\alpha \in [0,1]} \|Tx-Ty\| \leq \|\alpha(x-y)+(1-\alpha)(Tx-Ty)\|$

*/

■

Proposition 12.26.

[$f \in \Gamma_0(\mathcal{H})$; $x, p \in \mathcal{H}$]

$p = \mathrm{prox}_f x \Leftrightarrow \forall_{y \in \mathcal{H}} \langle y-p \mid x-p \rangle + f(p) \leq f(y)$

2/11/2016 10:58 PM

Proof:

($\Rightarrow$)

take any $y \in \mathcal{H}$

$p = \mathrm{prox}_f x$

$\forall_{\alpha \in]0,1[}$ $p_\alpha = \alpha y + (1-\alpha) p$, now as $f \in \Gamma_0(\mathcal{H}) \Rightarrow f(p_\alpha) = f(\alpha y + (1-\alpha)p) \leq \alpha f(y) + (1-\alpha) f(p)$

now. $\mathrm{prox}_f x$ by definition satisfies the following

$^1 f(x) = \min_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2}\|x-y\|^2 \right) = f(\mathrm{prox}_f x) + \frac{1}{2}\|x - \mathrm{prox}_f x\|^2$

$= f(p) + \frac{1}{2}\|x-p\|^2$

$\Rightarrow \forall_{\tilde{y} \in \mathcal{H}} \quad f(y) + \frac{1}{2}\|x-y\|^2 \geq f(p) + \frac{1}{2}\|x-p\|^2$

set $\tilde{y} := p_\alpha = \alpha y + (1-\alpha)p$

$f(p) + \frac{1}{2}\|x-p\|^2 \leq f(p_\alpha) + \frac{1}{2}\|x-p_\alpha\|^2$

$\Leftrightarrow f(p) \leq f(p_\alpha) + \frac{1}{2}\|x-p_\alpha\|^2 - \frac{1}{2}\|x-p\|^2$

$\leq \alpha f(y) + (1-\alpha) f(p) + \frac{1}{2}\underbrace{\|x-p_\alpha\|^2}_{} - \frac{1}{2}\|x-p\|^2$

/* $\|x - \alpha y - (1-\alpha)p\|^2 = \|x - \alpha y - p + \alpha p\|^2$

$= \|(x-p) - \alpha(y-p)\|^2 = \|x-p\|^2 + \alpha^2\|y-p\|^2 - 2\alpha\langle x-p \mid y-p \rangle$ */

$= \alpha f(y) + (1-\alpha) f(p) + \frac{1}{2}\cancel{\|x-p\|^2} + \frac{\alpha^2}{2}\|y-p\|^2 - \alpha\langle x-p \mid y-p \rangle - \frac{1}{2}\cancel{\|x-p\|^2}$

$\Rightarrow f(p) \leq \alpha f(y) + (1-\alpha) f(p) + \frac{\alpha^2}{2}\|y-p\|^2 - \alpha\langle x-p \mid y-p \rangle$

$\Leftrightarrow \cancel{\alpha} f(p) \leq \cancel{\alpha} f(y) + \frac{\alpha^{\cancel{2}}}{2}\|y-p\|^2 - \cancel{\alpha}\langle x-p \mid y-p \rangle \quad [\because \alpha \in]0,1[\,]$

$\therefore \forall_{\alpha \in]0,1[} \quad f(p) + \langle x-p \mid y-p \rangle \leq f(y) + \frac{\alpha}{2}\|y-p\|^2$

$\alpha \downarrow 0 \Rightarrow \langle y-p \mid x-p \rangle + f(p) \leq f(y) \quad \forall_{y \in \mathcal{H}}$

$\therefore \forall_{\alpha \in]0,1[} \quad f(p) + \langle x-p | y-p \rangle \leq f(y) + \frac{\alpha}{2} \|y-p\|^2$

$\alpha \downarrow 0 \Rightarrow \quad \langle y-p | x-p \rangle + f(p) \leq f(y) \quad \forall_{y \in \mathcal{H}}$

$(\Leftarrow)$

$\forall_{y \in \mathcal{H}} \quad \langle y-p | x-p \rangle + f(p) \leq f(y)$

$\Leftrightarrow \quad f(p) \leq f(y) - \langle y-p | x-p \rangle$

$\Leftrightarrow \quad f(p) + \frac{1}{2}\|x-p\|^2 \leq f(y) + \frac{1}{2}\|x-p\|^2 + \langle x-p | p-y \rangle$

$\leq \underbrace{f(y) + \frac{1}{2}\|x-p\|^2 + \langle x-p | p-y \rangle + \underbrace{\frac{1}{2}\|p-y\|^2}_{\geq 0}}_{\frac{1}{2}\left(\|x-p\|^2 + 2\langle x-p | p-y \rangle + \|p-y\|^2\right) = \frac{1}{2}\left(\|x-p+p-y\|^2\right) = \frac{1}{2}\|x-y\|^2}$ so adding it will make r.h.s bigger

$= f(y) + \frac{1}{2}\|x-y\|^2$

$\therefore \forall_{y \in \mathcal{H}} \quad f(p) + \frac{1}{2}\|x-p\|^2 \leq f(y) + \frac{1}{2}\|x-y\|^2$

$\Leftrightarrow \quad f(p) + \frac{1}{2}\|x-p\|^2 = \min_{y \in \mathcal{H}} f(y) + \frac{1}{2}\|x-y\|^2$

$\Leftrightarrow \quad p = \mathrm{Prox}_f x$

■

* Proposition 12.29.

[[$f \in \Gamma_0(\mathcal{H})$

$\gamma \in \mathbb{R}_{++}$]]

${}^{\gamma}f : \mathcal{H} \to \mathbb{R}$. Fréchet differentiable, /*[[$x \in \mathcal{H}$. $C \in \mathcal{V}(x)$, $T : C \to \mathcal{K}$]] T: Fréchet differentiable at $x \overset{\text{def}}{\Leftrightarrow} \exists \; DT(x) \in \underbrace{\mathcal{B}(\mathcal{H}, \mathcal{K})}_{\text{space of bounded linear operator from } \mathcal{H} \text{ to } \mathcal{K} \text{ with domain } \mathcal{H}} \quad \lim_{0 \neq \|y\| \to 0} \frac{\|T(x+y) - Tx - DT(x)y\|}{\|y\|} = 0$

$\nabla({}^{\gamma}f) = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f}) : \gamma^{-1}$ Lipschitz continuous.

2/12/2016 4:37 PM

<u>Proof</u>:

take, $x, y \in \mathcal{H} : x \neq y$

$p = \mathrm{Prox}_{\gamma f} x, \quad q = \mathrm{Prox}_{\gamma f} y$

/* ${}^{\gamma}f(x) = f(\mathrm{Prox}_{\gamma f} x) + \frac{1}{2\gamma}\|x - \mathrm{Prox}_{\gamma f} x\|^2$ */

so,

${}^{\gamma}f(x) = f(p) + \frac{1}{2\gamma}\|x-p\|^2$

${}^{\gamma}f(y) = f(q) + \frac{1}{2\gamma}\|y-q\|^2$

${}^{\gamma}f(y) - {}^{\gamma}f(x) = f(q) - f(p) + \frac{1}{2\gamma}\left(\|y-q\|^2 - \|x-p\|^2\right)$

/* Proposition 12.26 :

$\mathfrak{p} = \mathrm{Prox}_f(\mathfrak{x}) \Leftrightarrow \forall_{\mathfrak{y} \in \mathcal{H}} \quad \langle \mathfrak{y} - \mathfrak{p} | \mathfrak{x} - \mathfrak{p} \rangle + f(\mathfrak{p}) \leq f(\mathfrak{y})$ /* 특히 $\mathfrak{x} = \mathfrak{y}$ 일 경우 */

$p = \mathrm{Prox}_{\gamma f} x \Leftrightarrow \forall_{\mathfrak{z} \in \mathcal{H}} \quad \langle \mathfrak{z} - p | x - p \rangle + \gamma f(p) \leq \gamma f(\mathfrak{z})$

$\mathfrak{z} := q$

$\langle q - p | x - p \rangle + \gamma f(p) \leq \gamma f(q)$

$\Leftrightarrow \quad \gamma f(q) - \gamma f(p) \geq \langle q - p | x - p \rangle$

$\Leftrightarrow \quad f(q) - f(p) \geq \frac{1}{\gamma}\langle q - p | x - p \rangle$

*/

${}^{\gamma}f(y) - {}^{\gamma}f(x) \geq \frac{1}{\gamma}\langle q-p | x-p \rangle + \frac{1}{2\gamma}\left(\|y-q\|^2 - \|x-p\|^2\right)$

$= \frac{1}{2\gamma}\left(2\langle q-p | x-p \rangle + \|y-q\|^2 - \|x-p\|^2\right)$ /* now

$= \frac{1}{2\gamma}\left(\underbrace{\|(y-q)-(x-p)\|^2}_{\geq 0 \text{, so removing it will give smaller term}} + 2\langle y-x | x-p \rangle\right)$

$\geq \frac{1}{\gamma}\langle y-x | x-p \rangle$

$\|(y-q)-(x-p)\|^2 + 2\langle y-x | x-p \rangle$

$= \|y-q\|^2 + \|x-p\|^2 - \underbrace{2\langle y-q | x-p \rangle + 2\langle y-x | x-p \rangle}_{+\, 2\langle y-x-y+q | x-p \rangle}$

$= \|y-q\|^2 + \|x-p\|^2 + 2\langle q-x | x-p \rangle$

$= \|y-q\|^2 - \|x-p\|^2 + \underbrace{2\|x-p\|^2 + 2\langle q-x | x-p \rangle}_{2\langle x-p | x-p \rangle + 2\langle q-x | x-p \rangle}$

$= 2\langle \cancel{x-p+q-x} | x-p \rangle$

$= \|y-q\|^2 - \|x-p\|^2 + 2\langle q-p | x-p \rangle$

*/

$\therefore \forall \quad {}^{\gamma}f(y) - {}^{\gamma}f(x) \geq \frac{1}{\gamma}\langle y-x | x-p \rangle$

$$\frac{1}{\gamma}\langle y-x|x-p\rangle$$

$$= \|y-q\|^2 - \|x-p\|^2 + 2\langle q-p|x-p\rangle \quad */$$

$$\therefore \forall_{x,y\in\mathcal{H}:\, x\neq y} \quad {}^{\gamma}f(y) - {}^{\gamma}f(x) \geq \frac{1}{\gamma}\langle y-x|x-p\rangle$$

$$p = \mathrm{Prox}_{\gamma f}\, x$$

/* in free notation

$$\forall_{z,y\in\mathcal{H}:\, z\neq y} \quad {}^{\gamma}f(y) - {}^{\gamma}f(z) \geq \frac{1}{\gamma}\langle y-z|z-\mathrm{Prox}_{\gamma f}\, z\rangle$$

so, $z := y$, $y := x$

$$\forall_{y,x\in\mathcal{H}:\, y\neq x} \quad {}^{\gamma}f(x) - {}^{\gamma}f(y) \geq \frac{1}{\gamma}\langle x-y|\, \overbrace{y-\mathrm{Prox}_{\gamma f}\, y}^{q}\rangle$$

$$\leftrightarrow \quad {}^{\gamma}f(y) - {}^{\gamma}f(x) \leq \frac{1}{\gamma}\langle y-x|y-q\rangle \quad */$$

$$\forall_{x,y\in\mathcal{H}:\, x\neq y} \quad 0 \leq \underbrace{{}^{\gamma}f(y) - {}^{\gamma}f(x)}_{\leq \frac{1}{\gamma}\langle y-x|y-q\rangle} - \frac{1}{\gamma}\langle y-x|x-p\rangle$$

$$\leq \frac{1}{\gamma}\langle y-x|y-q\rangle - \frac{1}{\gamma}\langle y-x|x-p\rangle$$

$$= \frac{1}{\gamma}\langle y-x|(y-q)-(x-p)\rangle$$

$$= \frac{1}{\gamma}\langle y-x|(y-x)-(q-p)\rangle$$

$$= \frac{1}{\gamma}\Big(\underbrace{\langle y-x|y-x\rangle}_{\|y-x\|^2} - \langle y-x|q-p\rangle\Big)$$

$$\leq \frac{1}{\gamma}\Big(\|y-x\|^2 - \underbrace{\|q-p\|^2}_{\leq 0}\Big)$$

≤ 0, so removing it would give a larger term

$$\leq \frac{1}{\gamma}\|y-x\|^2$$

/* set $\mathrm{Prox}_{\gamma f} = T$: firmly nonexpansive then by Proposition 4.2. (iv)

$$\forall_{x\in\mathcal{H}}\ \forall_{y\in\mathcal{H}} \quad \|\underbrace{Ty}_{q}-\underbrace{Tx}_{p}\|^2 \leq \langle y-x|\underbrace{Ty}_{q}-\underbrace{Tx}_{p}\rangle$$

$$\leftrightarrow \quad \|q-p\|^2 \leq \langle y-x|q-p\rangle$$

$$\leftrightarrow \quad -\|q-p\|^2 \geq -\langle y-x|q-p\rangle \quad */$$

so we have got

$$\forall_{x,y\in\mathcal{H}:\, x\neq y} \quad 0 \leq {}^{\gamma}f(y) - {}^{\gamma}f(x) - \langle y-x|\tfrac{1}{\gamma}(x-p)\rangle \leq \frac{1}{\gamma}\|y-x\|^2$$

$$\leftrightarrow \quad 0 \leq \frac{{}^{\gamma}f(y) - {}^{\gamma}f(x) - \langle y-x|\frac{1}{\gamma}(x-p)\rangle}{\|y-x\|} \leq \frac{1}{\gamma}\|y-x\|$$

set $y-x = h \neq 0$ then $y = x+h$, which makes the above as follows:

$$0 \leq \frac{{}^{\gamma}f(x+h) - {}^{\gamma}f(x) - \langle h|\frac{1}{\gamma}(x-\mathrm{Prox}_{\gamma f}\, x)\rangle}{\|h\|} \leq \frac{1}{\gamma}\|h\|$$

thus,

$$\lim_{0\neq\|h\|\to 0} \frac{{}^{\gamma}f(x+h) - {}^{\gamma}f(x) - \langle \overbrace{\tfrac{1}{\gamma}(x-\mathrm{Prox}_{\gamma f}\, x)}^{(D^{\gamma}f)(x)}|h\rangle}{\|h\|} = 0 \quad \text{/* Checking with the definition */}$$

$$\therefore \nabla({}^{\gamma}f)(\cdot) = (D^{\gamma}f)^{\overset{\text{transpose}}{T}}(\cdot) = \frac{1}{\gamma}(\cdot - \mathrm{Prox}_{\gamma f}\,\cdot)^{TT} = \frac{1}{\gamma}\underbrace{(\cdot - \mathrm{Prox}_{\gamma f}\,\cdot)}_{\text{column vector}}$$

$$\text{so, } \nabla({}^{\gamma}f) = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f}) \quad \blacksquare$$

Proposition 12.31.

[[K: nonempty closed convex cone in $\mathcal{H}$

$q = \frac{1}{2}\|\cdot\|^2$]]

$$\nabla(q\cdot P_K) = \nabla\Big(\frac{1}{2}d_{K^\ominus}^2\Big) = P_K$$

2/13/2016 8:53 AM

Proof:

/* $C^\ominus$: polar cone of $C \overset{\text{def}}{\leftrightarrow} C^\ominus = \{u\in\mathcal{H} \mid \sup\langle C|u\rangle \leq 0\}$ */

Recall Moreaus Theorem (Theorem 6.29)

/*

* Theorem 6.29. (Moreau) /* relates projection, cone and polar cone */

[[K: nonempty convex cone in $\mathcal{H}$

$x\in\mathcal{H}$]]

(i) $x = P_K x + P_{K^\ominus} x$ /* this is amazing, any vector in $\mathcal{H}$ can be decomposed into two vectors associated with the projection on a cone and its polar */

(ii) $P_K x \perp P_{K^\ominus} x$

(iii) $\|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x)$

and Corollary 12.30.

* Corollary 12.30.

[[C: nonempty closed convex subset of $\mathcal{H}$]]

d_C^2: Fréchet differentiable /* d_C: distance function to a set : $d_C(x) = \inf d(x,C)$ */

$\nabla(d_C^2) = 2(\mathrm{Id} - P_C)$

[[C: nonempty closed convex subset of $\mathcal{H}$]]

d_C^2: Fréchet differentiable /* d_C: distance function to a set: $d_C(x) = \inf d(x,C)$ */

~~$\nabla d_C^2 = 2(Id - P_C)$~~

*/

$$\forall_{x\in\mathcal{H}} \quad \|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x)$$

$$\Leftrightarrow \quad d_K^2(x) = \|x\|^2 - d_{K^\ominus}^2(x)$$

$$\Rightarrow \quad \nabla_x d_K^2(x) = \nabla_x\left(\|x\|^2 - d_{K^\ominus}^2(x)\right)$$

$$= 2x - \nabla_x d_{K^\ominus}^2(x)$$

$$\Leftrightarrow \quad \nabla_x d_K^2(x) = 2(Id - P_K)(x) = 2x - \nabla_x d_{K^\ominus}^2(x)$$

$$\Rightarrow \quad 2(x - P_K x) = 2x - \nabla_x d_{K^\ominus}^2(x)$$

$$\Leftrightarrow \quad 2P_K x = \nabla_x d_{K^\ominus}^2(x)$$

$$\Leftrightarrow \quad \tfrac{1}{2}\nabla_x d_{K^\ominus}^2(x) = P_K(x)$$

$$\therefore \quad P_K = \nabla\left(\tfrac{1}{2} d_{K^\ominus}^2\right)$$

$\blacksquare$

Proposition 12.32.

[[$f \in \Gamma_0(\mathcal{H})$

$x \in \mathcal{H}$]]

$$\left({}^{\gamma}f(x)\right)_{\gamma\in\mathbb{R}_{++}} : \begin{cases} \bullet \text{ decreasing} \\ \bullet\ {}^{\gamma}f(x) \uparrow f(x) \text{ as } \gamma \downarrow 0 \\ \bullet\ {}^{\gamma}f(x) \downarrow \inf f(\mathcal{H}) \text{ as } \gamma \uparrow +\infty \end{cases}$$

2/13/2016 2:32 PM

Proof:

Recall that ${}^{\gamma}f$ is the power 2-norm infimal convolution

So, ${}^{\gamma}f$ will satisfy all the properties of power norm infimal convolution:

(Power norm infimal convolution of a general function and its properties)

Proposition 12.9. ★★★

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

$p \in [1, +\infty[$

$\forall_{\gamma\in\mathbb{R}_{++}} \quad g_\gamma = f \,\square\, \left(\frac{1}{\gamma p}\|\cdot\|^p\right)$: power norm infimal convolution]]

$\operatorname{dom} g_\gamma = \mathcal{H}$

~~$\forall_{\gamma\in\mathbb{R}_{++}} \forall_{x\in\mathcal{H}} \quad \inf g_\gamma(\mathcal{H}) = \inf f(\mathcal{H})$~~

g_γ: bounded on every ball in $\mathcal{H}$

/* e-Pasch-Hausdorff envelope is the power 1-norm infimal convolution for $\beta > 0$.

• Moreau envelope is power 2-norm infimal convolution */

in this theorem we can replace g_γ with ${}^{\gamma}f$.

as $\forall_{\mu>\gamma} \quad g_\mu(x) \leq g_\gamma(x)$

$$\Rightarrow \quad \forall_{\mu>\gamma} \quad {}^{\mu}f(x) \leq {}^{\gamma}f(x)$$

$$\therefore \quad \left({}^{\gamma}f(x)\right)_{\gamma\in\mathbb{R}_{++}} : \text{decreasing} \quad ☺$$

$g_\mu(x) \downarrow \inf f(\mathcal{H})$ as $\mu \to +\infty$

$\therefore \ {}^{\mu}f(x) \downarrow \inf f(\mathcal{H})$ as $\mu \to +\infty$ ☺

So, we are left with proving ${}^{\gamma}f(x) \uparrow f(x)$ as $\gamma \downarrow 0$

set $\mu = \sup_{\gamma\in\mathbb{R}_{++}} {}^{\gamma}f(x)$.

From Proposition 12.9 (ii):

$$\forall_{\gamma\in\mathbb{R}_{++}} \forall_{\tilde{\mu}\in]\gamma,+\infty[} \quad \inf f(\mathcal{H}) \leq {}^{\tilde{\mu}}f(x) \leq {}^{\gamma}f(x) \leq f(x) \Rightarrow \forall_{\gamma\in\mathbb{R}_{++}} \quad {}^{\gamma}f(x) \leq f(x) \Rightarrow \underbrace{\sup_{\gamma\in\mathbb{R}_{++}} {}^{\gamma}f(x)}_{\mu} \leq \underbrace{\sup_{\gamma\in\mathbb{R}_{++}} f(x)}_{f(x)} \Rightarrow \mu \leq f(x)$$

$$\therefore \ \forall_{\gamma\in\mathbb{R}_{++}} \forall_{\tilde{\mu}\in]\gamma,+\infty[} \quad \inf f(\mathcal{H}) \leq {}^{\tilde{\mu}}f(x) \leq {}^{\gamma}f(x) \leq \mu \leq f(x) \rightarrow \underline{\lim}_{\gamma\to 0} {}^{\gamma}f(x) \leq \overline{\lim}_{\gamma\to 0} {}^{\gamma}f(x) \leq \mu \leq f(x) \quad (1)$$

as ${}^{\gamma}f(x)$: decreasing in γ, the sup will be achieved as $\gamma \downarrow 0$ /* A bounded set has a finite supremum */

assume $\mu < +\infty$ /*otherwise, $\mu = +\infty \leq f(x) \Rightarrow f(x) = +\infty \Rightarrow x \notin \operatorname{dom} f$ the claim holds trivially */

Now if we show.

$\underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \geq f(x)$ it is enough as we have already shown (•) then from (1) and $\underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \geq f(x) \Rightarrow \underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \leq \overline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \leq \mu \leq f(x) \leq \underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x)$ which means: $\lim_{\gamma\downarrow 0} {}^{\gamma}f(x) = f(x)$ */

/* trick move: as $\gamma \downarrow 0$, we can confine γ to $]0,1[$ instead of $\mathbb{R}_{++}$ in our analysis */

[Eq:2017-03-02 10:36 AM]

assume $\mu<+\infty$ /* otherwise, $\mu=+\infty \leq f(x) \Rightarrow f(x)=+\infty \Rightarrow x\notin \operatorname{dom} f$ the claim holds trivially */

Now if we show.

$\underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \geq f(x)$ /* it is enough as we have already shown (•) then from (1) and $\underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \geq f(x) \Rightarrow \underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \leq \overline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x) \leq \mu \leq f(x) \leq \underline{\lim}_{\gamma\downarrow 0} {}^{\gamma}f(x)$ which means: $\lim_{\gamma\downarrow 0} {}^{\gamma}f(x) = f(x)$ */

/* trick move: as $\gamma\downarrow 0$, we can confine γ to $]0,1[$ instead of $\mathbb{R}_{++}$ in our analysis */

[Eq:2017-03-02 10:36 AM]

First, recall from definition of ${}^{\gamma}f$:

$$ {}^{\gamma}f(x) = f(\operatorname{Prox}_{\gamma f} x) + \frac{1}{2\gamma}\|x-\operatorname{Prox}_{\gamma f} x\|^2 $$

so,

$$ \sup_{\gamma\in\mathbb{R}_{++}} {}^{\gamma}f(x) = \mu \geq {}^{\gamma}f(x) = f(\operatorname{Prox}_{\gamma f} x) + \frac{1}{2\gamma}\|x-\operatorname{Prox}_{\gamma f} x\|^2 $$

denote $g := f + \frac{1}{2}\|x-\cdot\|^2$

$\therefore \operatorname{lev}_{\leq\mu} g = \{y\in\mathcal{H} \mid g(y) = f(y) + \frac{1}{2}\|x-y\|^2 \leq \mu\}$

suppose $\gamma\in]0,1]$ then,

$$ \frac{1}{2} \leq \frac{1}{2}\cdot\frac{1}{\gamma} $$

$$ \Rightarrow \frac{1}{2}\|x-y\|^2 \leq \frac{1}{2\gamma}\|x-y\|^2 $$

$$ \Leftrightarrow \underbrace{f(y) + \frac{1}{2}\|x-y\|^2}_{g(y)} \leq f(y) + \frac{1}{2}\frac{1}{\gamma}\|x-y\|^2 $$

$$ \therefore \forall_{\gamma\in]0,1[} \ \forall_{y\in\mathcal{H}} \quad g(y) \leq f(y) + \frac{1}{2}\frac{1}{\gamma}\|x-y\|^2 $$

setting $y := \operatorname{Prox}_{\gamma f} x$

$$ \forall_{\gamma\in]0,1[} \quad g(\operatorname{Prox}_{\gamma f} x) \leq f(\operatorname{Prox}_{\gamma f} x) + \frac{1}{2\gamma}\|x-\operatorname{Prox}_{\gamma f} x\|^2 = {}^{\gamma}f(x) \leq \mu $$

$$ \Rightarrow \forall_{\gamma\in]0,1[} \quad \operatorname{Prox}_{\gamma f}(x) \in \operatorname{lev}_{\leq\mu} g $$

now note that

g: coercive /* from Corollary 11.15 */

* Corollary 11.15. *

[$\tilde{f}, \tilde{g} \in \Gamma_0(\mathcal{H})$ /* set $\tilde{g} := f$, $\tilde{f} := \frac{1}{2}\|x-\cdot\|^2$

$\operatorname{dom}\tilde{f} \cap \operatorname{dom}\tilde{g} \neq \emptyset$ /* holds */

One of the following holds:

(i) $\tilde{f}$: supercoercive

(ii) $\tilde{f}$: coercive, $\tilde{g}$: bounded below]

• $\tilde{f}+\tilde{g}$: coercive, has a minimizer over $\mathcal{H}$

• $\tilde{f}$ or $\tilde{g}$: strictly convex $\Rightarrow$ $\tilde{f}+\tilde{g}$: has exactly one minimizer over $\mathcal{H}$.

/* $\tilde{f} = \|x-\cdot\|^2$: supercoercive as $\lim_{\|x\|\to+\infty} \frac{\tilde{f}(x)}{\|x\|} = \lim_{\|x\|\to+\infty} \frac{\frac{1}{2}\|x-\cdot\|^2}{\|x\|} = \lim_{\|x\|\to+\infty} \frac{\frac{1}{2}\|x-\cdot\|^2}{\|x\|}$ (Proposition 1.1.1.

$$^{\gamma}f(x) = f(\mathrm{Prox}_{\gamma f}x) + \frac{1}{2\gamma}\|x - \mathrm{prox}_{\gamma f}x\|^2$$

$$\Rightarrow \lim_{\gamma \downarrow 0} {}^{\gamma}f(x) = \lim_{\gamma \downarrow 0} f(\mathrm{Prox}_{\gamma f}x) + \underbrace{\frac{1}{2\gamma}\|x - \mathrm{prox}_{\gamma f}x\|^2}_{\geq 0\text{, so removing will give a smaller number}} \quad /* \; \overset{]a,b[}{f(\gamma,x) = g(\gamma,x)} \Rightarrow \lim_{\gamma \to a} f(\gamma,x) = \lim_{\gamma \to a} g(\gamma,x) \; */$$

$$\geq \lim_{\gamma \downarrow 0} f(\mathrm{Prox}_{\gamma f}x)$$

$$= \lim_{\gamma \downarrow 0} f(\mathrm{Prox}_{\gamma f}x) \quad /* \; \underline{\lim}\, x_a = \lim x_a = \overline{\lim}\, x_a \text{ when the limit exists } */$$

$$\geq f(x) \quad /* \; \mathrm{Prox}_{\gamma f}x \to x \text{ as } \gamma \to 0 \Rightarrow \lim_{\gamma \downarrow 0} f(\mathrm{Prox}_{\gamma f}x) \geq fx \; */$$

[Eq:2017-03-02 10:36 AM] implies:

as f : lower semicontinuous

$$\therefore \lim_{\gamma \downarrow 0} {}^{\gamma}f(x) = f(x)$$

[Goal reached :)]

∎

Part 3

February 24, 2017 8:37 AM

(Power norm infimal convolution of a general function and its properties)

*Proposition 12.9. ***

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

$p\in[1,+\infty]$

$\forall_{\gamma\in\mathbb{R}_{++}}\quad g_\gamma = f\,\square\left(\frac{1}{\gamma p}\|\cdot\|^p\right)$:power norm infimal convolution]

$\Rightarrow \forall_{\gamma\in\mathbb{R}_{++}}\ \forall_{x\in\mathcal{H}}$

- $\operatorname{dom} g_\gamma = \mathcal{H}$

- $\mu\in]\gamma,+\infty[\Rightarrow \inf f(\mathcal{H}) \leq g_\mu(x) \leq g_\gamma(x) \leq f(x)$

- $\inf g_\gamma(\mathcal{H}) = \inf f(\mathcal{H})$

- $g_\mu(x) \downarrow \inf f(\mathcal{H})$ as $\mu\uparrow+\infty$

- g_γ: bounded on every ball in $\mathcal{H}$.

Proof:

take $x\in\mathcal{H}$, $\gamma\in\mathbb{R}_{++}$

(i) $g_\gamma = f\,\square\,\frac{1}{\gamma p}\|\cdot\|^p$

$\operatorname{dom} g_\gamma = \operatorname{dom} f + \underbrace{\operatorname{dom}\frac{1}{\gamma p}\|\cdot\|^p}_{\mathcal{H}}$

$= \underbrace{\operatorname{dom} f}_{\subseteq\mathcal{H}} + \mathcal{H} = \mathcal{H}$

/*

*Proposition 12.6.

[$f,g,h:\mathcal{H}\to]-\infty,+\infty]$] $\Rightarrow$

(i) f,g: have continuous affine minorant with slope $u\in\mathcal{H}$

$\Rightarrow f\,\square\,g$: has continuous affine minorant with slope u, $-\infty\notin(f\,\square\,g)(\mathcal{H})$

(ii) $\operatorname{dom}(f\,\square\,g) = \operatorname{dom} f + \operatorname{dom} g$

(iii) $f\,\square\,g = g\,\square\,f$ /*symmetry*/

(iv) (f,g,h: have continuous affine minorant with same slope)

$\Rightarrow f\,\square\,(g\,\square\,h) = (f\,\square\,g)\,\square\,h$ /*Associativity*/

*/

(ii) /*required info

Fact 1.7.1. [$f,g:\mathcal{H}\to[-\infty,+\infty]$, $\forall_{x\in\mathcal{H}}\ f(x)\leq g(x) \Leftrightarrow f\leq g$] $\sup f(\mathcal{H})\leq\sup g(\mathcal{H})$, $\inf f(\mathcal{H})\leq\inf g(\mathcal{H})$

Fact 1.7.2. [$f,g:\mathcal{H}\to[-\infty,+\infty]$, $\forall_{x,y\in\mathcal{H}}\ f(x)\leq g(y)$] $\sup f(\mathcal{H})\leq\inf g(\mathcal{H})$

*/

take $\mu\in]\gamma,+\infty[$

$\Leftrightarrow \gamma<\mu<+\infty$

$\Rightarrow \frac{1}{\gamma} > \frac{1}{\mu}$

$\Rightarrow \frac{1}{\gamma}\cdot\underbrace{\frac{1}{p}\|\cdot\|^p}_{\text{positive}} \geq \frac{1}{\mu}\,\frac{1}{p}\|\cdot\|^p$ /*equality when $\cdot=0$*/

$$\Rightarrow \forall_{x,y\in\mathcal{H}} \quad f(y)+\frac{1}{\gamma}\frac{1}{p}\|x-y\|^p \geqslant f(y)+\frac{1}{\mu}\cdot\frac{1}{p}\|x-y\|^p$$

$$\Rightarrow \forall_{x\in\mathcal{H}} \quad \underbrace{\inf_{y\in\mathcal{H}} f(y)+\frac{1}{\gamma}\frac{1}{p}\|x-y\|^p}_{\left(f\,\square\,\frac{1}{\gamma}\cdot\frac{1}{p}\|\cdot\|^p\right)(x)=g_\gamma(x)} \geqslant \underbrace{\inf_{y\in\mathcal{H}} f(y)+\frac{1}{\mu p}\|x-y\|^p}_{\left(f\,\square\,\frac{1}{\mu p}\|\cdot\|^p\right)(x)=g_\mu(x)}$$

$$\Rightarrow \forall_{x\in\mathcal{H}} \quad g_\gamma(x)\geqslant g_\mu(x) \qquad (1)$$

now

$$\forall_{x\in\mathcal{H}}\ \forall_{y\in\mathcal{H}} \quad f(y)\leqslant f(y)+\underbrace{\frac{1}{\gamma p}\|x-y\|^p}_{\substack{\text{positive number}\\ \text{so adding it would}\\ \text{make the side larger}}}$$

$$\Rightarrow \forall_{x\in\mathcal{H}} \quad \underbrace{\inf_{y\in\mathcal{H}} f(y)}_{\inf f(\mathcal{H})} \leqslant \underbrace{\inf_{y\in\mathcal{H}}\left(f(y)+\frac{1}{\gamma p}\|x-y\|^p\right)}_{g_\gamma(x)} \overset{\text{set } y:=x}{\leqslant} f(x)+\frac{1}{\gamma p}\|x-x\|^p = f(x) \qquad (2)$$

now γ in (2) was arbitrary so, $\forall_{\gamma\in\mathbb{R}_{++}} \quad \inf f(\mathcal{H})\leqslant g_\gamma(x)\leqslant f(x)$

and $\forall_{\mu\in]\gamma,+\infty[} \quad \inf f(\mathcal{H})\leqslant g_\mu(x)\leqslant f(x)$

But (1) $\Rightarrow \quad g_\mu(x)\leqslant g_\gamma(x)$ $\Rightarrow$

$$\forall_{x\in\mathcal{H}}\ \forall_{\gamma\in\mathbb{R}_{++}}\ \forall_{\mu\in]\gamma,+\infty[} \quad \inf f(\mathcal{H})\leqslant g_\mu(x)\leqslant g_\gamma(x)\leqslant f(x) \quad ☺$$

(iii) from (ii)

$$\forall_{\gamma\in\mathbb{R}_{++}}\ \forall_{x\in\mathcal{H}} \quad \inf f(\mathcal{H})\leqslant g_\gamma(x)\leqslant f(x)$$

$$\Rightarrow \forall_{\gamma\in\mathbb{R}_{++}}\ \forall_{x\in\mathcal{H}} \quad g_\gamma(x)\leqslant f(x), \quad \text{and} \quad \inf f(\mathcal{H})\leqslant \inf_{x\in\mathcal{H}} g_\gamma(x) = \inf g_\gamma(\mathcal{H})$$

$$\Rightarrow \forall_{\gamma\in\mathbb{R}_{++}} \quad \underbrace{\inf_{x\in\mathcal{H}} g_\gamma(x)}_{\inf g_\gamma(\mathcal{H})}\leqslant \underbrace{\inf_{x\in\mathcal{H}} f(x)}_{\inf f(\mathcal{H})} \quad \text{and} \quad \inf f(\mathcal{H})\leqslant \inf g_\gamma(\mathcal{H})$$

$$\Rightarrow \forall_{\gamma\in\mathbb{R}_{++}} \quad \inf f(\mathcal{H})\leqslant \inf g_\gamma(\mathcal{H})\leqslant \inf f(\mathcal{H})$$

$$\because \forall_{\gamma \in \mathbb{R}_{++}} \quad \inf g_\gamma(\mathcal{H}) = \inf f(\mathcal{H}) \quad \because$$

(iv)

$\forall_{y \in \mathcal{H}}$

$$\forall_{\mu \in \mathbb{R}_{++}} \quad g_\mu(x) = \inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{\mu p} \|x-y\|^p \right) \leq f(y) + \frac{1}{\mu p} \|x-y\|^p$$

recall the definition of functional limit superior:

Definition 1.23.2. (Functional limit superior)

[$\mathcal{X}$: Hausdorff space; $f: \mathcal{X} \to [-\infty, +\infty]$; $x \in \mathcal{X}$] $\overline{\lim}_{y \to x} f(y) = \inf_{V \in \mathcal{V}(x)} \sup f(V)$

$$= \max_{(x_a)_{a \in A}:\ \text{net in } \mathcal{X},\ x_a \to x} \overline{\lim}\, f(x_a)$$

so, $$\overline{\lim}_{\mu \uparrow +\infty} g_\mu(x) = \max_{(\tilde{x}_a)_{a \in A}:\ \text{net in } \mathbb{R},\ \tilde{x}_a \to +\infty} \overline{\lim}\, \underbrace{g_{\tilde{x}_a}(x)}_{\inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{\tilde{x}_a p} \|x-y\|^p \right)}$$

$$= \max_{(\tilde{x}_a)_{a \in A}:\ \text{net in } \mathbb{R},\ \tilde{x}_a \to +\infty} \underbrace{\overline{\lim} \inf_{y \in \mathcal{H}} \left(f(y) + \overbrace{\frac{1}{\tilde{x}_a p} \|x-y\|^p}^{\to 0} \right)}_{\inf_{y \in \mathcal{H}} f(y) = \inf f(\mathcal{H})}$$

$$= \max_{(\tilde{x}_a)_{a \in A}:\ \text{net in } \mathbb{R},\ \tilde{x}_a \to +\infty} \inf f(\mathcal{H}) = \inf f(\mathcal{H})$$

now recall the definition of limit inferior for functions:

Definition 1.23.1. (Functional limit inferior) **

[$\mathcal{X}$: Hausdorff space; $f: \mathcal{X} \to [-\infty, +\infty]$; $x \in \mathcal{X}$] $\underline{\lim}_{y \to x} f(y) = \sup_{V \in \mathcal{V}(x)} \inf f(V)$

$$= \min_{(x_a)_{a \in A}:\ \text{net in } \mathcal{X},\ x_a \to x} \underline{\lim}\, f(x_a)$$

lemma 1.23.

so, $$\underline{\lim}_{\mu \uparrow +\infty} g_\mu(x) = \min_{(\tilde{x}_a)_{a \in A}:\ \text{net in } \mathbb{R},\ \tilde{x}_a \to +\infty} \lim\, \underbrace{g_{\tilde{x}_a}(x)}_{\inf_{y \in \mathcal{H}} \left(f(y) + \frac{1}{\tilde{x}_a p} \|x-y\|^p \right)}$$

$$= \min_{(\tilde{x}_a)_{a \in A}:\ \text{net in } \mathbb{R},\ \tilde{x}_a \to +\infty} \underbrace{\underline{\lim} \inf_{y \in \mathcal{H}} \left(f(y) + \overbrace{\frac{1}{\tilde{x}_a p} \|x-y\|^p}^{\to 0} \right)}_{\|}$$

$$= \underset{\substack{(\tilde{x}_a)_{a\in A}:\text{net in } \mathbb{R},\\ \tilde{x}_a \to +\infty}}{\text{—}} \underbrace{\underset{y\in H}{\text{—}} \quad x_a, p}_{\displaystyle \inf_{y\in H} f(y) = \inf f(H)}$$

$$= \min_{\substack{(\tilde{x}_a)_{a\in A}:\text{net in } \mathbb{R},\\ \tilde{x}_a \to +\infty}} \inf f(H) = \inf f(H)$$

$$\therefore \underline{\lim}_{\mu\uparrow+\infty} g_\mu(x) = \overline{\lim}_{\mu\uparrow+\infty} g_\mu(x) = \lim_{\mu\uparrow+\infty} g_\mu(x) = \inf f(H)$$

(4) $z \in \operatorname{dom} f$. $\rho \in \mathbb{R}_{++}$

$$\forall_{y\in B(x;\rho)} \quad g_\gamma(y) = \inf_{z\in H} f(z) + \frac{1}{\gamma p}\|y-z\|^p$$

$$\leq f(z) + \frac{1}{\gamma p}\|y-z\|^p$$

/+ We use one step at a time trick:

- $p=1$

$$\|y-z\|^1 = \|y-x+x-z\|^1 \leq \|y-x\|+\|x-z\|$$

- $p>1$ Recall Example 8.21. ** [$p\in]1,+\infty[$] $\Rightarrow$
  - $\|\cdot\|^p$: strictly convex
  - $\forall_{x\in H}\ \forall_{y\in H}\quad \|x+y\|^p \leq 2^{p-1}(\|x\|^p+\|y\|^p)$ +/

$$\|y-z\|^p = \|(y-x)+(x-z)\|^p \leq 2^{p-1}(\|y-x\|^p+\|x-z\|^p) \quad +/$$

$$\leq f(z) + \frac{1}{\gamma p}2^{p-1}(\|y-x\|^p + \|x-z\|^p)$$

/+ now $y\in B(x;\rho)$
$\Leftrightarrow \|y-x\|\leq\rho$ +/

$$\leq \underbrace{f(z)}_{\substack{\text{finite}\\ \text{as } z\in\operatorname{dom} f}} + \underbrace{\frac{1}{\gamma p}2^{p-1}}_{\text{finite}}\Big(\underbrace{\rho^p}_{\text{finite}} + \underbrace{\|x-z\|^p}_{\substack{\text{fixed, so finite}\\ \text{on the ball } B(x;\rho)}}\Big) < +\infty$$

$$\therefore \forall_{y\in B(x;\rho)} \quad g_\gamma(y) < +\infty$$

$\Leftrightarrow$ g_γ: bounded above on every ball in H. ∎

Part 1

Sunday, February 14, 2016 10:09 PM

Example 13.9.

[[$\phi : \mathcal{H} \to]-\infty, +\infty]$, proper;

$\gamma \in \mathbb{R}_{++}$;

$f = \phi + \frac{1}{2\gamma}\|\cdot\|^2$]]

$f^* = \frac{\gamma}{2}\|\cdot\|^2 - {}^{\gamma}\phi \circ \gamma\mathrm{Id} = \frac{\gamma}{2}\|\cdot\|^2 - (\phi \,\square\, \frac{1}{2\gamma}\|\cdot\|^2) \circ \gamma\mathrm{Id}$

Proof:

$u \in \mathcal{H}$

$$f^*(u) = \sup_{x \in \mathcal{H}} \left(\langle x|u\rangle - f(x) \right)$$

$$= -\inf_{x \in \mathcal{H}} \left(-\langle x|u\rangle + f(x) \right) \quad /\!\!* \text{using } \sup(\cdot) = -\inf(-\cdot) \; */$$

$$= -\inf_{x \in \mathcal{H}} \left(-\langle x|u\rangle + \phi(x) + \frac{1}{2\gamma}\|x\|^2 \right) \quad /\!\!* \; \frac{1}{2\gamma}\|x\|^2 - \langle x|u\rangle = \frac{1}{2\gamma}\left(\|x\|^2 - 2\gamma\langle x|u\rangle + \gamma^2\|u\|^2 - \gamma^2\|u\|^2 \right) = \frac{1}{2\gamma}\left(\|x - \gamma u\|^2 - \gamma^2\|u\|^2 \right) \; */$$

$$= -\inf_{x \in \mathcal{H}} \left(\phi(x) + \frac{1}{2\gamma}\|x - \gamma u\|^2 - \underbrace{\frac{\gamma}{2}\|u\|^2}_{\text{constant w.r.t } x} \right)$$

so বাইরে রাখা যাবে

$$= \frac{\gamma}{2}\|u\|^2 - \underbrace{\inf_{x \in \mathcal{H}} \left(\phi(x) + \frac{1}{2\gamma}\|x - \gamma u\|^2 \right)}_{{}^{\gamma}\phi(\gamma u)} \quad /\!\!* \text{ now: Moreau envelope, } {}^{\gamma}f(x) = \inf_{y \in \mathcal{H}} \left(f(x) + \frac{1}{2\gamma}\|x - y\|^2 \right) \; */$$

$$= \frac{\gamma}{2}\|u\|^2 - {}^{\gamma}\phi(\gamma u)$$

∎

Example 13.8. [$\phi : \mathcal{H} \to]-\infty, +\infty]$; f: perspective function of ϕ : $\mathbb{R} \times \mathcal{H} \to]-\infty, +\infty] : (\xi, x) \mapsto \begin{cases} \xi\,\phi(x/\xi), & \text{if } \xi > 0 \\ +\infty, & \text{else} \end{cases}$

$C = \{(v,u) \in \mathbb{R} \times \mathcal{H} \mid v + \phi^*(u) \leq 0\}$] $\Rightarrow f^* = \iota_C$

Proof:

$$f^*(v,u) = \sup_{(\xi,x) \in \mathbb{R} \times \mathcal{H}} \left(-f(\xi, x) + \langle (\xi, x) | (v,u) \rangle \right) \quad /\!\!* \; \tilde{f}^*(\tilde{u}) = \sup_{\tilde{x} \in \mathcal{H}} \left(-\tilde{f}(\tilde{x}) + \langle \tilde{x}|\tilde{u}\rangle \right) \; */$$

$$= \sup_{(\xi,x) \in \mathbb{R} \times \mathcal{H}} \left(-\left[\begin{cases} \xi\,\phi(x/\xi), & \text{if } \xi > 0 \\ +\infty, & \text{else} \end{cases} \right] + \langle \xi|v\rangle + \langle x|u\rangle \right)$$

$$= \sup_{\substack{(\xi,x) \in \mathbb{R} \times \mathcal{H} \\ \xi > 0}} \left(-\xi\,\phi(x/\xi) + \langle \xi|v\rangle + \langle x|u\rangle \right) \quad /\!\!* \text{ now } (x, \xi) \text{ independent, so we can sup over } x \text{ first and then over } \xi \; */$$

/* otherwise the value is $-\infty$ */

$$= \sup_{\xi \in \mathbb{R}_{++}} \sup_{x \in \mathcal{H}} \left(-\xi\,\phi(x/\xi) + \overbrace{\langle \xi|v\rangle}^{\text{constant w.r.t. } x} + \langle x|u\rangle \right)$$

$$= \sup_{\xi \in \mathbb{R}_{++}} \left(\langle \xi|v\rangle + \underbrace{\sup_{x \in \mathcal{H}} \left(\langle x|u\rangle - \xi\,\phi(x/\xi) \right)}_{\sup_{x \in \mathcal{H}} \xi \left(\langle \frac{x}{\xi}|u\rangle - \phi(x/\xi) \right) = \xi \sup_{x \in \mathcal{H}} \left(-\phi(\frac{x}{\xi}) + \langle \frac{x}{\xi}|u\rangle \right) = \xi \overbrace{\sup_{y = \frac{x}{\xi} \in \mathcal{H}} \left(-\phi(y) + \langle y|u\rangle \right)}^{\phi^*(u)} = \xi\,\phi^*(u)} \right)$$

$$= \sup_{\xi \in \mathbb{R}_{++}} \left(\underbrace{\langle \xi|v\rangle}_{\xi v \;/\!\!/ \text{ as both } \xi, v \text{ are real numbers}} + \xi\,\phi^*(u) \right) = \sup_{\xi \in \mathbb{R}_{++}} \left(\xi v + \xi\,\phi^*(u) \right) = \sup_{\xi \in \mathbb{R}_{++}} \xi \left(v + \phi^*(u) \right) = \left(v + \phi^*(u) \right) \sup_{\xi \in \mathbb{R}_{++}} \xi = \begin{cases} 0, & \text{if } v + \phi^*(u) \leq 0 \\ +\infty, & \text{else} \end{cases} = \iota_C \;\&\; C = \{(v,u) \in \mathbb{R} \times \mathcal{H} \mid v + \phi^*(u) \leq 0\} \;\&$$

$\therefore f^* = \iota_C \;\&\; C = \{(v,u) \in \mathbb{R} \times \mathcal{H} \mid v + \phi^*(u) \leq 0\}$

∎

2/15/2016 2:44 PM

* Proposition 13.10.

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

(i) [$(u,v) \in \mathcal{H} \times \mathbb{R}$]

$(u,v) \in \operatorname{epi} f^* \leftrightarrow \langle \cdot|u\rangle - v \leq f$

(ii) $f^* = +\infty \leftrightarrow f$: possesses no continuous affine minorant

(iii) [$\operatorname{dom} f^* \neq \emptyset$]

f : bounded below on every bounded subset of $\mathcal{H}$.

Proof:

(iii) [$\operatorname{dom} f^* \neq \emptyset$]

f: bounded below on every bounded subset of $\mathcal{H}$.

Proof:

(i) $(u,v) \in \operatorname{epi} f^*$

$\Leftrightarrow f^*(u) \leq v$ /* $f^*(u) = \sup_{x \in \mathcal{H}} \left(\langle u|x \rangle - f(x) \right)$ */

$\Leftrightarrow \sup_{x \in \mathcal{H}} \left(\langle u|x \rangle - f(x) \right) \leq v$

$\Leftrightarrow \forall_{x \in \mathcal{H}} \quad \langle u|x \rangle - f(x) \leq v$

$\Leftrightarrow \forall_{x \in \mathcal{H}} \quad \langle u|x \rangle - v \leq f(x)$

$\Leftrightarrow \langle u|\cdot \rangle - v \leq f$

☺

(ii)

$f^* = +\infty$

$\Leftrightarrow \operatorname{epi} f^* = \{ (x,t) \in \mathcal{H} \times \mathbb{R} \mid f(x) \leq t \} = \emptyset$

$\Leftrightarrow \neg \left(\exists_{(u,v) \in \operatorname{epi} f^*} \right)$

$\Leftrightarrow \neg \left(\exists_{\substack{(u,v) \\ \in \mathcal{H} \times \mathbb{R}}} \langle u|\cdot \rangle - v \leq f \right)$ /* from (i) */

$\Leftrightarrow \forall_{(u,v) \in \mathcal{H} \times \mathbb{R}} \quad \langle u|\cdot \rangle - v > f$

$\Leftrightarrow f$: possesses no continuous affine minorant.

(iii)

$\operatorname{dom} f^* = \{ x \in \mathcal{H} \mid f(x) < +\infty \} \neq \emptyset$

$\Leftrightarrow \exists_{u \in \mathcal{H}} \quad f^*(u) < +\infty$

$\Leftrightarrow \exists_{(u,v) \in \mathcal{H} \times \mathbb{R}} \quad f^*(u) = v < +\infty$

$\Leftrightarrow (u,v) \in \operatorname{epi} f^*$

$\Leftrightarrow \langle \cdot | u \rangle - v \leq f$ /* from (i) */

Consider a bounded set C in $\mathcal{H}$

take $\beta = \sup_{x \in C} \|x\|$ then $\forall_{x \in C} \quad \beta = \sup_{x \in C} \|x\| \geq \|x\| \Rightarrow \forall_{x \in C} \quad -\beta \leq -\|x\|$ . . . (1)

$\forall_{x \in C} \quad f(x) \geq \langle x|u \rangle - v$

$\geq -\|u\| \, \|x\| - v$

/* Cauchy Schwartz:

$|\langle x|u \rangle| \leq \|x\| \, \|u\|$

$\max \{ \langle x|u \rangle, -\langle x|u \rangle \}$

So, $\langle x|u \rangle \leq \|x\| \, \|u\|$

$-\langle x|u \rangle \leq \|x\| \, \|u\|$

$\Leftrightarrow \langle x|u \rangle \geq -\|x\| \, \|u\|$ */

$\Leftrightarrow \forall_{x \in C} \quad f(x) \geq -\|x\| \, \|u\| - v$

$\geq -\beta \|u\| - v$

/* from (1) */

$= \underbrace{-\beta \|u\|}_{\text{finite}} \underbrace{-v}_{\text{finite}}$

$> -\infty$

$\therefore f$: bounded below on every bounded subset of $\mathcal{H}$.

∎

Proposition 13.13. (Fenchel-Young inequality)

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

$\forall_{x \in \mathcal{H}} \; \forall_{u \in \mathcal{H}} \quad f(x) + f^*(u) \geq \langle x|u \rangle$

2/16/2016 2:48 PM

Proof:

$\forall_{x \in \mathcal{H}} \; \forall_{u \in \mathcal{H}}$

f: proper $\Rightarrow -\infty \notin f(\mathcal{H})$, $\operatorname{dom} f = \{ \tilde{x} \in \mathcal{H} \mid f(\tilde{x}) < +\infty \} \neq \emptyset$

if $f(x) = +\infty$, the inequality trivially holds

$f(x) < +\infty \Rightarrow f^*(u) = \sup_{\tilde{x} \in \mathcal{H}} \left(\langle \tilde{x}|u \rangle - f(\tilde{x}) \right) \geq \langle \tilde{x}|u \rangle - f(\tilde{x}) \quad \forall_{\tilde{x} \in \mathcal{H}}$

$\tilde{x} := x$

$\Rightarrow f^*(u) \geq \langle x|u \rangle - f(x)$

$\therefore f^*(u) + f(x) \geq \langle x|u \rangle$

∎

* Proposition 13.12.

2/16/2016 3:06 PM

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

f^*: even

Proof: $\forall u \in \mathcal{H}$

$f^*(-u) = \sup_{x \in \mathcal{H}} \left(\langle x|-u \rangle - \underbrace{f(x)}_{\substack{\| \\ f(-x)}} \right)$

$f(-x)$ /* f: even */

$= \sup_{x\in H} \left(\langle -x|u\rangle - f(-x) \right)$

$= \sup_{(-x)\in H} \left(\langle -x|u\rangle - f(-x) \right)$ /* e.g. $\max_{x\in H} f(x) = \max_{(-x)\in H} f(-x)$ */

$= \sup_{y\in H} \left(\langle y|u\rangle - f(y) \right)$

$= f^*(u)$

$\therefore \forall_{u\in H}\ f^*(-u) = f^*(u)$

$\leftrightarrow f^*$: even.

■

* Proposition 13.19.

[[$f: H \to]-\infty, +\infty]$, $f \geqslant f(0) = 0$]]

$f^* \geqslant f^*(0) = 0$

2/16/2016 3:23 PM

Proof:

$f^*(u) = \sup_{x\in H} \left(\langle x|u\rangle - f(x) \right)$

$\therefore f^*(0) = \sup_{x\in H} \left(\underbrace{\langle x|0\rangle}_{=0} - f(x) \right)$

$= \sup_{x\in H} \left(-f(x) \right)$

$= -\inf_{x\in H} \left(f(x) \right)$ /* $\inf(\cdot) = -\sup(-\cdot)$ */

$= -\inf f(H)$

now given, $f \geqslant f(0) = 0$

$\leftrightarrow \forall_{x\in H}\ f(x) \geqslant f(0) = 0$

$\leftrightarrow \inf_{x\in H} f(x) \geqslant f(0) = 0$, as $f(0) = 0$, and $\inf f(H) \geqslant 0$, the minimum value will be achieved at $x = 0$

$\leftrightarrow \inf_{x\in H} f(x) = \min_{x\in H} f(x) = f(0)$ /* $0 \in H$, so equality will hold */
$= 0$

$\therefore \inf f(H) = 0$

$\therefore f^*(0) = 0 = f(0)$

Again, $\forall_{u\in H}\ f^*(u) = \sup_{x\in H} \left(\langle x|u\rangle - f(x) \right) \geqslant \langle x|u\rangle - f(x)\ \ \forall_{x\in H}$

$\overset{x:=0}{\Rightarrow} \forall_{u\in H}\ f^*(u) \geqslant \underbrace{\langle 0|u\rangle}_{=0} - f(0) = -f^*(0) = 0 = f^*(0)$

$\therefore \forall_{u\in H}\ f^*(u) \geqslant f^*(0) = 0$

$\leftrightarrow f^* \geqslant f^*(0) = 0$

■

Proposition 13.21.

[[$f, g: H \to]-\infty, +\infty]$]]

(i) $(f \square g)^* = f^* + g^*$

(ii) f, g: proper $\Rightarrow (f+g)^* \leqslant f^* \square g^*$

(iii) $\forall_{\gamma\in\mathbb{R}_{++}}\ (\gamma f)^* = f^* + \frac{\gamma}{2}\|\cdot\|^2$

(iv) $L \in B(H, K) \Rightarrow (L \triangleright f)^* = f^* \circ L^*$

(v) $L \in B(K, H) \Rightarrow (f \circ L)^* \leqslant L^* \triangleright f^*$

2/17/2016 4:48 PM

Proof:

(i) denote $f \square g = h$ /* $(f \square g)(z) = \inf_{y\in H} \left(f(y) + g(z-y) \right)$ */

$\therefore h(x) = \inf_{y\in H} \left(f(y) + g(x-y) \right)$

$\forall_{u\in H}$

$h^*(u) = \sup_{x\in H} \left(\langle x|u\rangle - h(x) \right)$

$= \sup_{x\in H} \left(\langle x|u\rangle - \inf_{y\in H} \left(f(y) + g(x-y) \right) \right)$

$-\sup_{y\in H} \left(-f(y) - g(x-y) \right)$ /* $\because \inf(\cdot) = -\sup(-\cdot)$ */

$= \sup_{x\in H} \left(\langle x|u\rangle + \sup_{y\in H} \left(-f(y) - g(x-y) \right) \right)$

$= \sup_{x\in H} \sup_{y\in H} \left(\langle x|u\rangle - f(y) - g(x-y) \right)$

$$= \sup_{y\in H} \sup_{x\in H} \left(\langle x|u\rangle - f(y) - g(x-y) \right)$$

/* $\langle y|u\rangle - \langle y|u\rangle + \langle x|u\rangle - f(y) - g(x-y)$

$= \langle y|u\rangle + \langle x-y|u\rangle - f(y) - g(x-y)$ /* $\langle a+b|c\rangle = \langle a|c\rangle + \langle b|c\rangle$ */

$= \left(\langle y|u\rangle - f(y)\right) + \left(\langle x-y|u\rangle - g(x-y)\right)$ */

$$= \sup_{y\in H} \sup_{x\in H} \left(\left(\langle y|u\rangle - f(y)\right) + \left(\langle x-y|u\rangle - g(x-y)\right) \right)$$

$$= \sup_{y\in H} \left[\left(\langle y|u\rangle - f(y)\right) + \sup_{x\in H} \left(\langle x-y|u\rangle - g(x-y)\right) \right]$$

$= \sup_{x-y\in H} \left(\langle x-y|u\rangle - g(x-y)\right)$ /* as y is a constant w.r.t x we can optimize w.r.t $x-y\in H$ and the $\sup(\cdot)$ value would be the same */

$= g^*(u)$

$$= \sup_{y\in H} \left(\langle y|u\rangle - f(y) + g^*(u) \right)$$

$$= g^*(u) + \underbrace{\sup_{y\in H} \left(\langle y|u\rangle - f(y) \right)}_{f^*(u)}$$

$= f^*(u) + g^*(u)$

$\therefore \forall_{u\in H} \quad (f \square g)^*(u) = f^*(u) + g^*(u)$

$\Leftrightarrow (f\square g)^* = f^* + g^*$ ☺

(ii)

/* Recall

* Proposition 13.14.

$[\![f, g : H \to [-\infty, +\infty]]\!]$

(i) $f^{**} \leq f$

(ii) $f \leq g \Rightarrow (f^* \geq g^*,\ f^{**} \leq g^{**})$

(iii) $f^{***} = f^*$

(iv) $(\check{f})^* = f^*$ /* $\check{f} = \sup\{g \in \Gamma(H) : g \leq f\}$: lower semicontinuous convex envelope of f */

*/

$f \geq f^{**}$

$g \geq g^{**}$

$f + g \geq f^{**} + g^{**}$

$\Rightarrow (f+g)^* \leq (f^{**} + g^{**})^*$

$=$ /* in (i) we have shown $f^* + g^* = (f \square g)^*$

so, $f^{**} + g^{**} = (f^*)^* + (g^*)^* = (f^* \square g^*)^*$ */

$= \left((f^* \square g^*)^* \right)^*$

$= (f^* \square g^*)^{**}$

$\leq f^* \square g^*$

$\therefore (f+g)^* \leq f^* \square g^*$

☺

(iii)

recall that, ${}^{\gamma}f = f \square \frac{1}{2\gamma}\|\cdot\|^2$

so, $({}^{\gamma}f)^* = \left(f \square \frac{1}{2\gamma}\|\cdot\|^2\right)^*$

$= f^* + \frac{1}{\gamma}\left(\frac{1}{2}\|\cdot\|^2\right)^*$ /* from (i) $(f \square g)^* = f^* + g^*$ */

$= f^* + \frac{1}{2\gamma}\|\cdot\|^2$ /* $\frac{1}{2}\|\cdot\|^2$ is the only self-conjugate function */

(iv)

$L \in \mathcal{B}(H,K)$ /* recall that: $(L \triangleright f)(y) = \inf_{x\in H : Lx = y} f(x)$: infimal postcomposition of f by L

$L \in \mathcal{B}(H,K) \Leftrightarrow L$: continuous linear operator */

$\forall_{v\in K}$

$(L \triangleright f)^*(v)$ /* say $L \triangleright f = h$ */

$= h^*(v)$

$(L \triangleright f)^*(v)$ /* say $L \triangleright f = h$ */

$= h^*(v)$

$= \sup_{y \in K} \left(\langle y|v\rangle - \underbrace{h(y)}_{\shortparallel} \right)$

/* $(L \triangleright f)(y) = \inf_{x \in H : Lx = y} f(x)$ */

$= \sup_{y \in K} \left(\langle y|v\rangle - \inf_{x \in H : Lx = y} f(x) \right)$

$\underbrace{}_{-\sup_{x \in H : Lx = y} -f(x)}$ /* $\because \inf(\cdot) = -\sup(-\cdot)$ */

$= \sup_{y \in K} \left(\langle y|v\rangle + \sup_{x \in H : Lx = y} (-f(x)) \right)$

$= \sup_{y \in K} \sup_{x \in H : Lx = y} \left[\langle y|v\rangle - f(x) \right]$

$= \begin{bmatrix} \sup_{(x,y) \in H \times K} & \langle y|v\rangle - f(x) \\ \text{s.t.} & Lx = y \end{bmatrix}$

$= \sup_{(x,y) \in H \times K} \langle y|v\rangle - f(x) + \iota_{Lx=y}(x,y)$

/*

* Supremum and infimum of functions : Results :

- $[f, g : H \to \mathbb{R} ; \forall_{x,y \in H} |f(x) - f(y)| \leq |g(x) - g(y)|]$ $\sup f(H) - \inf f(H) \leq \sup g(H) - \inf g(H)$
- $[\text{term} \in \{\sup, \inf, \overline{\lim}, \underline{\lim}, \underbrace{\lim}_{\text{if exists}}\} ; f : H \times K \to \mathbb{R}]$ $\text{term}_{x \in X} \sup_{y \in Y} f(x,y) \geq \sup_{y \in Y} \text{term}_{x \in X} f(x,y)$

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

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

*/

$= \sup_{x \in H} \sup_{y \in K} \left(\langle y|v\rangle - f(x) + \iota_{Lx=y}(x,y) \right)$

$= \sup_{x \in H} -f(x) + \underbrace{\sup_{y \in K} \left(\langle y|v\rangle + \iota_{Lx=y}(x,y) \right)}_{\langle Lx|v\rangle}$ /* as : $y = Lx$ for a fixed x is a single vector, so we are supremizing over a single vector

$\therefore \sup_{y \in K} \left(\langle y|v\rangle + \iota_{Lx=y}(x,y) \right) = \langle Lx|v\rangle$ */

$= \sup_{x \in H} \langle Lx|v\rangle - f(x)$ /* recall that, a linear bounded (continuous) operator $L \in \mathcal{B}(H,K)$ has its adjoint defined as:

$\forall_{x \in H} \forall_{y \in H} \langle Lx|y\rangle = \langle x|L^*y\rangle$ */

$= \sup_{x \in H} \langle x|L^*v\rangle - f(x)$

$= f^*(L^*v)$ /* $f^*(u) = \sup_{x \in H} \langle x|u\rangle - f(x)$ */

$\forall_{v \in K} \; (L \triangleright f)^*(v) = f^*(L^*v) = f^* L^* v = (f^* \circ L^*) v$

$\leftrightarrow \; (L \triangleright f)^* = (f^* \circ L^*)$ /* Caution : f^*: conjugate of f, L^*: adjoint of L */

☺

these stars stand for conjugate

these stars stand for adjoint

(v)

$(L^* \triangleright f^*)^* = (f^*)^* \circ (L^*)^*$ /* from (iv) $(L \triangleright f)^* = f^* \circ L^*$ */

$= f^{**} \circ \underbrace{L^{**}}_{L}$ /* for a continuous linear operator L, $L^{**} = L$: Fact 2.18 - (i) */

$= f^{**} \circ L$

$\leq f \circ L$

/* for any $g : H \to [-\infty, +\infty]$, $g^{**} \leq g \leftrightarrow \forall_{x \in H} \; g^{**}(x) \leq g(x)$

now, $\forall_{x \in H} \; f^{**}(x) \leq f(x)$

set $x := Ly$

then $\forall_{y \in K}$ $f^{**}(Ly) \leq f(Ly) \leftrightarrow f^{**} \circ L \leq f \circ L$ */

So,

$(L^* \triangleright f^*)^* \leq f \circ L$ /* from Proposition 13.14 (ii) $f \leq g \Rightarrow (f^* \geq g^*, f^{**} \leq g^{**})$

$\Rightarrow (L^* \triangleright f^*)^{**} \geq (f \circ L)^*$ /* from Proposition 13.14 (i) : $h^{**} \leq h$

$\therefore (L^* \triangleright f^*)^{**} \leq L^* \triangleright f^*$ */

$\therefore L^* \triangleright f^* \geq (f \circ L)^*$

∎

* Proposition 13.25.

[[$(f_i)_{i \in I}$: family of proper functions, : $H \to]-\infty, +\infty]$]]

(i) $\left(\inf_{i \in I} f_i \right)^* = \left(\sup_{i \in I} f_i^* \right)$

(ii) $\left(\sup_{i\in I} f_i\right)^* \leq \inf_{i\in I} f_i^*$

2/17/2016 5:03 PM

Proof:

(i) $\forall_{u\in H}$ /* take $\left(\inf_{i\in I} f_i\right) = h$ */

$$h^*(u) = \sup_{x\in H}\left(\langle x|u\rangle - h(x)\right)$$

/* interchangable as x, i independant */

$$= \sup_{x\in H}\left(\langle x|u\rangle - \underbrace{\left(\inf_{i\in I} f_i(x)\right)}_{-\sup_{i\in I} -f_i(x)}\right) = \sup_{x\in H}\left(\langle x|u\rangle + \sup_{i\in I} -f_i(x)\right) = \sup_{x\in H}\sup_{i\in I}\left(\langle x|u\rangle - f_i(x)\right)$$

$$= \sup_{i\in I}\underbrace{\sup_{x\in H}\left(\langle x|u\rangle - f_i(x)\right)}_{f_i^*(u)} = \sup f_i^*(u)$$

$$\therefore \quad \left(\inf_{i\in I} f_i\right)^*(u) = \left(\sup_{i\in I} f_i^*\right)(u) \Leftrightarrow \left(\inf_{i\in I} f_i\right)^* = \left(\sup_{i\in I} f_i^*\right)$$

(ii) take $g = \sup_{i\in I} f_i$

by definition, $g \geq f_i \ \forall_{i\in I}$

$\Rightarrow g^* \leq f_i^* \ \forall_{i\in I}$ /* By proposition 13.14 (ii): $g \geq h \Rightarrow g^* \leq h^*$

$\Leftrightarrow \underbrace{g^*}_{\left(\sup_{i\in I} f_i\right)^*} \leq \inf_{i\in I} f_i^*$

$$\therefore \left(\sup_{i\in I} f_i\right)^* \leq \inf_{i\in I} f_i^*$$

Part 2

Thursday, February 18, 2016 1:45 PM

Proposition 13.28.

[K: real Hilbert space;

$F: \mathcal{H}\times K \to]-\infty,+\infty]$, proper

$f: \mathcal{H}\to[-\infty,+\infty]: x \mapsto \inf F(x,K) = \inf_{y\in K} F(x,y)$]

$f^* = F^*(\cdot, 0)$

Proof:

Fix $u\in\mathcal{H}$:

$$f^*(u) = \sup_{x\in\mathcal{H}} \big(\langle x|u\rangle - \overbrace{f(x)}^{\inf_{y\in K} F(x,y)}\big)$$

$$= \sup_{x\in\mathcal{H}} \big(\langle x|u\rangle - \underbrace{\inf_{y\in K} F(x,y)}_{-\sup_{y\in K} -F(x,y)}\big)$$

$$= \sup_{x\in\mathcal{H}} \big(\langle x|u\rangle + \overbrace{\sup_{y\in K} -F(x,y)}^{\text{constant w.r.t } y}\big)$$

$$= \sup_{x\in\mathcal{H}} \sup_{y\in K} \big(\langle x|u\rangle - F(x,y)\big)$$ /* recall $\sup_{x\in X}\sup_{y\in Y} f(x,y) = \sup_{y\in Y}\sup_{x\in X} f(x,y) = \sup_{(x,y)\in X\times Y} f(x,y)$ */

$$= \sup_{(x,y)\in\mathcal{H}\times K} \big(\underbrace{\langle x|u\rangle}_{\langle (x,y)|(u,0)\rangle = \langle x|u\rangle + \langle y|0\rangle = \langle x|u\rangle} - F(x,y)\big) = \sup_{(x,y)\in\mathcal{H}\times K} \big(\langle (x,y)|(u,0)\rangle - F(x,y)\big) = F^*(u,0)$$ ∴ ∎

(0: 0 vector)

Proposition 13.31.

[[$F\in\Gamma(\mathcal{H}\times\mathcal{H})$: autoconjugate]] /* [$F: \mathcal{H}\times\mathcal{H}\to[-\infty,+\infty]$] F: autoconjugate $\overset{\text{def}}{\leftrightarrow} F^* = F^T$ */

$F^T: (u,x) \mapsto F(x,u)$ (transposition operator)

$F \geq \langle\cdot|\cdot\rangle$

$F \geq \langle\cdot|\cdot\rangle$

$\therefore \boxed{F^*(u,x) = F^T(u,x) = F(x,u)}$

Proof:

* 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! */

$$\forall_{x\in\mathcal{H}} \forall_{u\in\mathcal{H}} \quad f(x) + f^*(u) \geq \langle x|u\rangle$$

$\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \forall_{(y,v)\in\mathcal{H}\times\mathcal{H}}$

$$F(x,u) + F^*(y,v) \geq \langle (x,u)|(y,v)\rangle$$

set $(y,v) := (u,x)$

then $F(x,u) + \underbrace{F^*(u,x)}_{F^T(u,x) = F(x,u)} \geq \langle (x,u)|(u,x)\rangle = \langle x|u\rangle + \langle u|x\rangle = 2\langle x|u\rangle$

$\leftrightarrow \quad 2F(x,u) \geq 2\langle x|u\rangle \quad \therefore F \geq \langle\cdot|\cdot\rangle$

as $F^*(u,x) = F(x,u)$ we also have $F^*(u,x) \geq \langle x|u\rangle = \langle u|x\rangle$

$\therefore F^* \geq \langle\cdot|\cdot\rangle$ ∎

2/19/2016 5:43 PM

* Theorem 13.32. (Fenchel-Moreau Theorem)

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

- f: lower semicontinuous, convex $\leftrightarrow f = f^{**}$
- f: lower semicontinuous, convex $\Rightarrow f^*$: proper

Proof: ($\Leftarrow$)

given: $f = f^{**}$ /* recall that, $\forall_{f:\mathcal{H}\to[-\infty,+\infty]}$ $f^* \in \Gamma(\mathcal{H})$ */ (proposition 13.32; set of all lower semicontinuous and convex functions, $:\mathcal{H}\to[-\infty,+\infty]$)

$\Rightarrow \underbrace{f^{**}}_{\text{irrespective of what } f^* \text{ is}} = f$: lower semicontinuous and convex

($\Rightarrow$)

/* Proof sketch:

Part 1: We prove that:

$$[\![f\in\Gamma_0(\mathcal{H}),\ \overbrace{(\bar{x},\xi)}^{\text{edom} f}\in\mathcal{H}\times\mathbb{R},\ \xi\in]-\infty, f(\bar{x})[,\ (p,\pi) = P_{epi}(\bar{x},\xi)]\!] \quad f^{**}(\bar{x}) > \xi$$

Proof sketch for part 1: (i) use

$\in \operatorname{dom} f$

$[\![f \in \Gamma_0(\mathcal{H}),\ (x,\xi) \in \mathcal{H} \times \mathbb{R},\ \xi \in]-\infty, f(x)[,\ (p,\pi) = P_{epi}(x,\xi)]\!]\quad f^{**}(x) > \xi$

Proof sketch for part 1: (i) use

* Proposition 9.18. * (characterization of projection of the epigraph of a Γ_0 function)

$[\![f \in \Gamma_0(\mathcal{H}),\ x \in \operatorname{dom} f,\ \xi \in]-\infty, f(x)[,\ (p,\pi) \in \mathcal{H} \times \mathbb{R}]\!]$

$$(p,\pi) = P_{epi f}(x,\xi) \Leftrightarrow \begin{cases} \bullet\ \boxed{\xi < f(p) = \pi} \ \checkmark \\ \bullet\ \forall_{y \in \operatorname{dom} f}\ \langle y-p \mid x-p \rangle \leq (f(y)-f(p))(\pi-\xi) \end{cases}$$

then (ii) prove:

$$\pi > \xi \Rightarrow f^{**}(x) > \xi \quad */$$

. . . (goal 1.(i))

Part 2: using part 1: we show that ($\operatorname{dom} f \neq \emptyset$ as $f \in \Gamma_0(\mathcal{H})$ for this case)

$$\forall_{x \in \operatorname{dom} f}\ f(x) = f^{**}(x) \quad \ldots \text{(goal 2)}$$

Part 3: we prove that:

$$\forall_{x \notin \operatorname{dom} f}\ f(x) = f^{**}(x) = +\infty \quad */$$

thus part 2 and part 3 proves that:

$$\forall_{x \in \mathcal{H}}\ f(x) = f^{**}(x)$$

Part 4: $f = f^{*} \Rightarrow f^{*}$: proper

*/

/+ proving goal 1.(i) */

given: f: lower semicontinuous, convex, $\underbrace{\text{proper}}_{\text{universal given}} \Leftrightarrow f \in \Gamma_0(\mathcal{H})$

/* recall:

Proposition 9.17.

$[\![f \in \Gamma_0(\mathcal{H}),\ (x,\xi) \in \mathcal{H} \times \mathbb{R},\ (p,\pi) \in \mathcal{H} \times \mathbb{R}]\!]$

$$(p,\pi) = P_{epi f}(x,\xi) \Leftrightarrow \begin{cases} \max\{\xi, f(p)\} \leq \pi \\ \forall_{y \in \operatorname{dom} f}\ \langle y-p \mid x-p \rangle + (f(y)-\pi)(\xi-\pi) \leq 0 \end{cases}$$

*/

take any $x \in \mathcal{H}$, take $\xi \in]-\infty, f(x)[$ *we are allowed to pick any $\xi \in \mathbb{R}$ */ otherwise $f(x) \leq \xi \Leftrightarrow (x,\xi) \in epi\, f$ and $P_{epi f}(x,\xi) = (x,\xi)$ a noninteresting case

$$(p,\pi) = P_{epi f}(x,\xi) \Leftrightarrow \begin{cases} \max\{\xi, f(p)\} \leq \pi \Rightarrow \xi \leq \pi \quad \text{*} \ \pi = f(p) \ \xi \ \text{*/} \\ \forall_{y \in \operatorname{dom} f}\ \langle y-p \mid x-p \rangle + (f(y)-\pi)(\xi-\pi) \leq 0 \end{cases} \quad \text{[Eq:13.20]}$$

/* $P \wedge (P \Rightarrow Q) \Leftrightarrow (P \wedge Q)$ */ /+ now we can have two cases $\pi > \xi$ and $\pi = \xi$

case $\pi > \xi$, then

$$\forall_{y \in \operatorname{dom} f}\ \langle y-p \mid x-p \rangle \leq \underbrace{(\pi-\xi)}_{>0}(f(y)-\pi)$$

$$\Leftrightarrow \left\langle y-p \,\middle|\, \frac{x-p}{\pi-\xi} \right\rangle \leq (f(y)-\pi)$$

$$\text{*set } v := \frac{x-p}{\pi-\xi} \text{ *}$$

$$\Leftrightarrow \langle y-p \mid v \rangle = \langle y \mid v \rangle - \langle p \mid v \rangle \leq f(y) - \pi$$

$$\Leftrightarrow \langle y \mid v \rangle - f(y) \leq \langle p \mid v \rangle - \pi \qquad \text{*} \ v = \frac{x-p}{\pi-\xi} \Leftrightarrow x-p = v(\pi-\xi) \Leftrightarrow p = x - v(\pi-\xi) \text{ *}$$

$$= \langle x - v(\pi-\xi) \mid v \rangle - \pi$$

$$= \langle x \mid v \rangle - (\pi-\xi)\underbrace{\langle v \mid v \rangle}_{\|v\|^2} - \pi$$

$$= \langle x \mid v \rangle \underbrace{- \underbrace{(\pi-\xi)\|v\|^2}_{\geq 0} - \pi}_{\leq 0 \text{ : so removing it will yield larger stuff}}$$

$$\leq \langle x \mid v \rangle - \pi$$

So, $\forall_{y \in \operatorname{dom} f}\ \langle y \mid v \rangle - f(y) \leq \langle x \mid v \rangle - \pi$

$$\Leftrightarrow \underbrace{\sup_{y \in \operatorname{dom} f} \langle y \mid v \rangle - f(y)}_{} \leq \langle x \mid v \rangle - \pi$$

$\leftrightarrow$ $\underbrace{\sup_{y\in \mathrm{dom} f} \langle y|v\rangle - f(y)}$ $\leq \langle x|v\rangle - \pi$

$f^*(v) = \sup_{y\in\mathcal{H}} \langle y|v\rangle - f(y)$ /* Because if $y \notin \mathrm{dom} f$ $f(y)=+\infty$

$\langle y|v\rangle - f(y) = -\infty \leq \langle x|v\rangle - \pi$ */

$\therefore$ $f^*(v) \leq \langle x|v\rangle - \pi$

$\Rightarrow$ $\pi \leq \langle x|v\rangle - f^*(v) \overset{\text{obviously}}{\leq} \sup_{w\in\mathcal{H}} \langle x|w\rangle - f^*(w) = f^{**}(x)$ /* now $\pi > \xi \Rightarrow f^{**} \geq \pi > \xi$ */

$v = \dfrac{x-p}{\pi-\xi}$

so we have shown that: $\boxed{\pi > \xi \Rightarrow f^{**}(x) > \xi}$ $\forall_{x\in \mathrm{dom} f}$ /* goal 1.(i) proved */

consider the case $x \in \mathrm{dom} f$

/* recall that.

* Proposition 9.18. * (characterization of projection of the epigraph of a $\subset$ function)

$[\![f\in\Gamma_0(\mathcal{H}),\ x\in\mathrm{dom} f,\ \xi\in]-\infty, f(x)[,\ (p,\pi)\in\mathcal{H}\times\mathbb{R}]\!]$

$(p,\pi) = P_{\mathrm{epi} f}(x,\xi) \Leftrightarrow \begin{cases} \cdot\ \xi \leq f(x) = \pi \\ \cdot\ \forall_{y\in\mathrm{dom} f}\ \langle y-p|x-p\rangle \leq (f(y)-\pi)(\xi-\pi) \end{cases}$

goal 1(i)

Combining both we have proved Part 1:

$\boxed{[\![f\in\Gamma_0(\mathcal{H}),\ \overbrace{(x,\xi)}^{\in \mathrm{dom} f}\in\mathcal{H}\times\mathbb{R},\ \xi\in]-\infty, f(x)[,\ (p,\pi) = P_{\mathrm{epi}}(x,\xi)]\!]\ \ f^{**}(x) > \xi}$

*/

/* Proof of part 2 starts here */ want to prove: $\forall_{x\in \mathrm{dom} f}$ $f^{**}(x) = f(x)$, assume $\mathrm{dom} f \neq \emptyset$

/* recall

* Proposition 13.14. * *

$[\![f,g : \mathcal{H}\to[-\infty,+\infty]]\!]$ /* these properties are actually quite useful */

(i) $f^{**} \leq f$

(ii) $f\leq g \Rightarrow (f^* \geq g^*,\ f^{**}\leq g^{**})$

(iii) $f^{***} = f^*$

(iv) $(\check{f})^* = f^*$ /* $\check{f} = \sup\{g\in\Gamma(\mathcal{H}) : g\leq f\}$: lower semicontinuous convex envelope of f */ */

$\boxed{f^{**}(x) \leq f(x)}$

$\xi < f^{**}(x) \leq f(x)$ /* $\xi \in]-\infty, f(x)[$ */

$\Rightarrow \underbrace{\sup_{\xi\in]-\infty,f(x)[} \xi}_{f(x)} \leq f^{**}(x) \leq f(x)$ /* e.g., $\forall_{x:x^2<2}$ $x < \sqrt{2}$

$\leftrightarrow \sup_{x:x^2<2} x \leq \sqrt{2}$ */

$\leftrightarrow$ $f(x) \leq f^{**}(x) \leq f(x)$

$\boxed{\forall_{x\in \mathrm{dom} f}\ f(x) = f^{**}(x)}$ (Part 2 proved) [Part_2_proved]

/* Part 3 proof started */

our goal is to prove that: $\forall_{x\notin \mathrm{dom} f}$ $f(x) = f^{**}(x)$

Proposition 9.17. so, $x\notin \mathrm{dom} f$ এর জন্য formula টি বসাই

$[\![f\in\Gamma_0(\mathcal{H}),\ (x,\xi)\in\mathcal{H}\times\mathbb{R},\ (p,\pi)\in\mathcal{H}\times\mathbb{R}]\!]$

$(p,\pi) = P_{\mathrm{epi} f}(x,\xi) \Leftrightarrow \begin{cases} \max\{\xi, f(p)\} \leq \pi \Rightarrow \xi \leq \pi \\ \forall_{y\in\mathrm{dom} f}\ \langle y-p|x-p\rangle + (f(y)-\pi)(\xi-\pi) \leq 0 \end{cases}$ [Eq13.20]

so, there are two cases: $\boxed{\xi<\pi}$ case1, $\boxed{\xi=\pi}$ case2

Consider case 1: $\pi > \xi$

$\pi > \xi \Rightarrow f^*(x) > \xi$ $\forall_{x\in\mathcal{H}}$ $\forall_{\xi\in]-\infty, f(x)[}$

if $x\notin \mathrm{dom} f \Rightarrow f(x) = +\infty$ $\quad f(x) = +\infty$

$\forall_{x\notin \mathrm{dom} f}$ $f^*(x) > \sup_{\xi\in]-\infty,+\infty[} \xi = +\infty \Rightarrow f^*(x) = +\infty$

$\forall_{x\notin \mathrm{dom} f}$ $f(x) = f^*(x) = +\infty$ for $\xi<\pi$ (case 1)

Now consider case 2: $\pi = \xi$

Now let us consider $\pi = \xi$

Now consider case 2. $\pi = \xi$

Now let us consider $\pi = \xi$

now, $\underbrace{(p,\pi)}_{\in \operatorname{epi} f} = \underbrace{p}_{\operatorname{epi} f}$ $\underbrace{(x,\xi)}_{\xi \in]-\infty, f(x)[}$

/* By definition */ so. $\xi < f(x) \Leftrightarrow (x,\xi) \notin \operatorname{epi} f$

so, $\underbrace{(p,\pi)}_{\in \operatorname{epi} f} \neq \underbrace{(x,\xi)}_{\notin \operatorname{epi} f}$ and $\pi = \xi$

$\Rightarrow p \neq x \Leftrightarrow \|x-p\| > 0$

take $w \in \operatorname{dom} f^* \Leftrightarrow f^*(w) < +\infty$,

$u = \|x-p\|$

[Eq13.20] says:

$\xi \leq \pi$, $\forall_{y \in \operatorname{dom} f}$ $\langle y-p | x-p \rangle + (f(y)-\pi)(\xi-\pi) \leq 0$

now $\xi = \pi$ so $\forall_{y \in \operatorname{dom} f}$ $\langle y-p | \underbrace{x-p}_{u} \rangle \leq 0$

$\langle y|u\rangle - \langle p|u\rangle \leq 0$

$\Leftrightarrow \langle y|u\rangle \leq \langle p|u\rangle$

$\therefore \forall_{y \in \operatorname{dom} f}$ $\langle y|u\rangle \leq \langle p|u\rangle$

$w \in \operatorname{dom} f^* \Rightarrow f^*(w) = \sup_{y \in H} \langle y|w\rangle - f(y) < +\infty$

$\Rightarrow \forall_{y \in H}$ $\langle y|w\rangle - f(y) \leq f^*(w) < +\infty$

$\Leftrightarrow \forall_{y \in \operatorname{dom} f}$ $\langle y|w\rangle - f(y) \leq f^*(w)$ /* As outside dom f, $\langle y|w\rangle - \overset{+\infty}{f(y)} = -\infty < +\infty$. so we can confine y to dom f */

So, we have $\boxed{\forall_{y \in \operatorname{dom} f} \quad \langle y|w\rangle - f(y) \leq f^*(w), \ \langle y|u\rangle \leq \langle p|u\rangle}$

take $\lambda \in \mathbb{R}_{++}$ then $\lambda\langle y|u\rangle \leq \lambda\langle p|u\rangle$

$\Leftrightarrow \langle y|\lambda u\rangle \leq \langle p|\lambda u\rangle$

$\langle y|w\rangle - f(y) \leq f^*(w)$ ①

$\forall_{y \in \operatorname{dom} f}$ $\langle y|w+\lambda u\rangle - f(y) \leq f^*(w) + \langle p|\lambda u\rangle$

$\Leftrightarrow \underbrace{\sup_{y \in \operatorname{dom} f} \langle y|w+\lambda u\rangle - f(y)}_{f^*(w+\lambda u)} \leq f^*(w) + \langle p|\lambda u\rangle$

$\therefore f^*(w+\lambda u) \leq f^*(w) + \underbrace{\langle p|\lambda u\rangle}$ /* $x-p = u \Leftrightarrow p = x-u$ */

$\lambda\langle p|u\rangle = \lambda\langle x-u|u\rangle$

$= \lambda\langle x|u\rangle - \lambda\langle u|u\rangle$

$= \langle x|\lambda u\rangle - \lambda\|u\|^2$

$= \langle x|w\rangle + \langle x|\lambda u\rangle - \langle x|w\rangle - \lambda\|u\|^2$

$= \langle x|w+\lambda u\rangle - \langle x|w\rangle - \lambda\|u\|^2$

$\Leftrightarrow f^*(w+\lambda u) \leq f^*(w) + \langle x|w+\lambda u\rangle - \langle x|w\rangle - \lambda\|u\|^2$

$\Leftrightarrow \langle x|w+\lambda u\rangle - f^*(w+\lambda u) \geq \langle x|w\rangle + \lambda\|u\|^2 - f^*(w)$

now. $f^{**}(x) = \sup_{\tilde{y} \in H} \langle x|\tilde{y}\rangle - f^*(\tilde{y}) \geq \langle x|\tilde{y}\rangle - f^*(\tilde{y})$ $\forall_{\tilde{y} \in H}$

set $\tilde{y} := w + \lambda u \in H$, then:

$\forall_{\lambda \in \mathbb{R}_{++}}$ $f^{**}(x) \geq \langle x|w+\lambda u\rangle - f^*(w+\lambda u) \geq \langle x|w\rangle + \lambda\|u\|^2 - f^*(w)$

$f^{**}(x) \geq \langle x|w\rangle + \lambda\|u\|^2 - f^*(w)$

$\Leftrightarrow f^{**}(x) \geq \sup_{\lambda \in \mathbb{R}_{++}} \underbrace{\langle x|w\rangle}_{\text{finite}} + \lambda\|u\|^2 - \underbrace{\underbrace{f^*(w)}_{<+\infty}}_{>-\infty} = +\infty$

$\therefore f^{**}(x) = +\infty$

$\forall_{x \notin \operatorname{dom} f}$ $f(x) = f^{**}(x) = +\infty$ for $\xi = \pi$ (case 2)

$f(x) = +\infty$

So, for both cases we have: $\forall_{x \notin \operatorname{dom} f}$ $f(x) = f^{**}(x) = +\infty$ ... (Part 3 proved) [Part_3_proved]

So, from (Part 2) and (Part 3) we have: $\forall_{x \in H}$ $f(x) = f^{**}(x)$:)

$f(x) = f^{**}(x)$

[Part_2_proved] [Part_3_proved]

So, from (Part 2) and (Part 3) we have: $\forall_{x\in\mathcal{H}}\ f(x)=f^{**}(x)$:)

[Part_2_proved] [Part_3_proved] $\Updownarrow$ $f(\mathcal{H})=f^{**}(\mathcal{H})$

(Part 4 proof start)

Now let us prove that: when $\left(\forall_{x\in\mathcal{H}}\ f(x)=f^{**}(x)\right)\Rightarrow f^{*}$: proper /* to prove this we use Proposition 13.9. g^{*}: proper $\Rightarrow g$: proper */

also, $f\in\Gamma_0(\mathcal{H})\Rightarrow -\infty\notin f(\mathcal{H})=f^{**}(\mathcal{H})$ ✓

also, $f(\operatorname{dom} f)=f^{**}(\operatorname{dom} f)$, $f\in\Gamma_0(\mathcal{H})\Rightarrow \operatorname{dom} f^{**}\neq\emptyset$ ✓ ($\Downarrow$ $\operatorname{dom} f\neq\emptyset$)

$\left.\right\}\ (-\infty\notin f^{**}(\mathcal{H}),\ \operatorname{dom} f^{**}\neq\emptyset)\Leftrightarrow f^{**}$ proper $\Rightarrow f^{*}$: proper. ∎

2/19/2016 6:38 PM

*Proposition 13.39.

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

- f: has a continuous affine minorant $\Rightarrow \check{f}=f^{**}$
  $\Updownarrow$
  $\operatorname{dom} f^{*}\neq\emptyset$
- f: doesnot have a continuous affine minorant $\Rightarrow f^{**}=-\infty$

Proof:

Case 1: $f=+\infty$ In this case f itself is affine, so we will show that $f^{**}=+\infty$

$f=+\infty$: convex $\Rightarrow \check{f}=f$ /* trivially */

$\hookrightarrow \operatorname{dom} f=\emptyset$ /* Sup over an empty set is $-\infty$ */

then $f^{*}(u)=\underbrace{\sup_{x\in\operatorname{dom} f}\left(\langle x|u\rangle-f(x)\right)}_{=\emptyset}=-\infty \Leftrightarrow f^{*}\equiv-\infty$

/* Proposition 13.9. (iv) */

$\therefore f^{**}(v)=\sup_{u\in\mathcal{H}}\ \langle u|v\rangle-\underbrace{f^{*}(u)}=+\infty \Leftrightarrow f^{**}=+\infty$

$\Rightarrow f^{**}=\check{f}$:)

Case 2: $f\not\equiv+\infty$ /* Now in this case we show that, if f has a continuous affine minorant then, $f^{**}=f^{\smile}$ and if does not then $f^{**}=-\infty$

/* recall Proposition 13.10. (i)

$f^{*}\equiv+\infty \Leftrightarrow f$: possesses no continuous affine minorant */

- f: possesses no continuous affine minorant $\Rightarrow f^{*}=+\infty$

$\Leftrightarrow f^{**}=-\infty$ /* using Proposition 13.9 (ii)

$-\infty\in f^{*}(\mathcal{H})\Leftrightarrow f\equiv+\infty\Leftrightarrow f^{*}\equiv-\infty$ */

- f: possesses a continuous affine minorant

$\Leftrightarrow \exists\ a:\mathcal{H}\to\mathbb{R}$, affine function, $a\leq f$

/* continuous affine function

e.g. $c^{T}x+d$ cannot shoot to $\pm\infty$ */

naturally $a=\check{a}$ /* as a: affine continuous */

$\Downarrow$ convex $\Downarrow$ lower semicontinuous

so, $a=\check{a}$: lower semicontinuous, convex function

$\Leftrightarrow a=\check{a}\in\Gamma(\mathcal{H})$

by definition, $\check{f}=\sup\{g\in\Gamma(\mathcal{H})\mid g\leq f\}\geq a=\check{a}$

$\therefore a=\check{a}\leq\check{f}\leq f$

now, $\check{a}$: proper as $\pm\infty\notin\check{a}(\mathcal{H})$

and $f:\mathcal{H}\to]-\infty,+\infty]$, $f\not\equiv+\infty$

$\Leftrightarrow f:\mathcal{H}\to]-\infty,+\infty[=\mathbb{R}\to f$: proper

so we have $\check{f}$: proper as it is sandwiched between two proper functions

So, $\check{f}$: proper, and by definition $\check{f}\in\Gamma(\mathcal{H})$

So we have $\check{f}$: proper as it is sandwiched between two proper functions

So, $\check{f}$: proper, and by definition $\check{f} \in \Gamma(\mathcal{H})$

$\rightarrow \check{f} \in \Gamma_0(\mathcal{H})$

/* recall Proposition 13.14 (iv) $(\check{f})^* = f^*$

Corollary 13.32. $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H}),\ f^{**} = f$ */

So, $\check{f} \in \Gamma_0(\mathcal{H}) \Rightarrow (\check{f})^* = f^* \in \Gamma_0(\mathcal{H}),\ \check{f}^{**} = f^{**} = f = \check{f}$

$\therefore f^{**} = (\check{f})^{**} = \check{f}$

$\blacksquare$

Proposition 13.41.

[Let $(f_i)_{i\in I} : \subseteq \Gamma_0(\mathcal{H}),\ \sup_{i\in I} f_i \neq +\infty$] $(\sup_{i\in I} f_i)^* = (\inf_{i\in I} f_i^*)^\vee$

Proof:

/* use Fenchel-Moreau theorem

* Theorem 13.32 (Fenchel-Moreau Theorem) *

[$f : \mathcal{H} \to]-\infty, +\infty]$, proper] /* f: proper $\Leftrightarrow -\infty \notin f(\mathcal{H})$, $\operatorname{dom} f = [x \in \mathcal{H} \mid f(x) < +\infty] \neq \emptyset$ */

- f: lower semicontinuous, convex $\Leftrightarrow f = f^{**}$
- f: lower semicontinuous, convex $\Rightarrow f^*$: proper

* Proposition 13.25.

[$(f_i)_{i\in I}$: family of proper functions, $: \mathcal{H} \to]-\infty, +\infty]$]

(i) $(\inf_{i\in I} f_i)^* = (\sup_{i\in I} f_i^*)$

(ii) $(\sup_{i\in I} f_i)^* \leq \inf_{i\in I} f_i^*$

*/

$\sup_{i\in I} f_i = \sup_{i\in I} f_i^{**}$ /* $f_i \in \Gamma_0(\mathcal{H}) \Rightarrow f_i^{**} = f_i$ */

$= \sup_{i\in I} (f_i^*)^*$

$= (\inf_{i\in I} f_i^*)^* \Rightarrow \boxed{(\sup_{i\in I} f_i)^* = (\inf_{i\in I} f_i^*)^{**}}$

first note that:

given: $\sup_{i\in I} f_i \neq +\infty$

$\Rightarrow \forall_{i\in I}\ f_i(\mathcal{H}) \in \mathbb{R}$ // as f_i: proper, $-\infty \notin f_i(\mathcal{H})$

$\Rightarrow \forall_{i\in I}\ f_i^*(u) = \sup_{x\in\mathcal{H}} \underbrace{(\langle x|u\rangle - f(x))}_{\text{finite}} = \text{finite}\ \ \forall_{u\in\mathcal{H}}$

$\Rightarrow \inf_{i\in I} f_i^* = \text{finite}$

using a similar logic

$\Rightarrow (\inf_{i\in I} f_i^*)^* = \text{finite}$

$\Rightarrow (\inf_{i\in I} f_i^*)^* \neq +\infty$ /* now

/* using

* Theorem 9.19.

[$f \in \Gamma_0(\mathcal{H})$] f: possesses a continuous affine minorant. /*the proof is constructive */

*/

* Proposition 13.10 *

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

(i) [nontrivial] (nontrivial $\Leftrightarrow (-\mathcal{H}) \cap \operatorname{dom} f \neq \emptyset$) /* Any continuous affine minorant of a function has the form $\langle \cdot | u \rangle - \beta$ in the notation of the original function. */

(ii) $f^* = +\infty \Leftrightarrow f$: possesses no continuous affine minorant

*/

$\Leftrightarrow (\inf_{i\in I} f_i^*)$: has a continuous affine minorant.

/* use

Proposition 13.29. *

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

- f: has a continuous affine minorant $\Rightarrow f^{**} = \check{f}$
  $\operatorname{dom} f^* \neq \emptyset$ (from nontrivial condition)
- f: doesnot have a continuous affine minorant $\Rightarrow f^{**} = -\infty$

*/

$\underbrace{(\inf_{i\in I} f_i^*)^{**}}_{(\sup_{i\in I} f_i)^*} = (\inf_{i\in I} f_i^*)^\vee$

$$\underbrace{\left(\inf_{i\in I} f_i^*\right)^{**}}_{\left(\sup_{i\in I} f_i\right)^*} = \left(\inf_{i\in I} f_i^*\right)^{\cup}$$

$$\therefore \left(\sup_{i\in I} f_i\right)^* = \left(\inf_{i\in I} f_i^*\right)^{\cup} \quad \blacksquare$$

Part 0

March 17, 2017 10:10 AM

Proposition 14.7.

[$L: \mathcal{H}\times\mathcal{H} \to \mathcal{H} : (y,z) \mapsto \frac{1}{2}(y+z)$;

$f, g \in \Gamma_0(\mathcal{H})$;

$F: \mathcal{H}\times\mathcal{H} \to]-\infty, +\infty] : (y,z) \mapsto \frac{1}{2}f(y) + \frac{1}{2}g(z) + \frac{1}{8}\|y-z\|^2$]

(i) $\operatorname{pav}(f,g) = \operatorname{pav}(g,f)$

(ii) $\operatorname{pav}(f,g) = L \triangleright f$

(iii) $\operatorname{dom} \operatorname{pav}(f,g) = \frac{1}{2}\operatorname{dom} f + \frac{1}{2}\operatorname{dom} g$

(iv) $\operatorname{pav}(f,g)$: proper, convex function

<u>Proof</u>:

(i)
$$\operatorname{pav}(f,g) \overset{\text{def}}{=} \left[\begin{array}{ll} \inf\limits_{(y,z)} & \frac{1}{2}\left(f(y) + g(z) + \frac{1}{4}\|y-z\|^2\right) \\ \text{s.t.} & y+z = 2x \\ & (y,z) \in \mathcal{H}\times\mathcal{H} \end{array}\right] \quad \ldots (1)$$

from definition,

it is clear that: $\operatorname{pav}(f,g) = \operatorname{pav}(g,f)$

(ii) recall that

$$(L \triangleright f)(\tilde{y}) = \inf_{\tilde{x}} f(\tilde{x}) \quad \text{s.t. } L\tilde{x} = \tilde{y}$$

So,
$$(L \triangleright F)(x) = \left(\begin{array}{ll} \inf\limits_{(y,z)} & F(y,z) \\ \text{s.t.} & L(y,z) = x \\ & (y,z) \in \mathcal{H}\times\mathcal{H} \end{array}\right)$$
/* given, $F(y,z) = \frac{1}{2}f(y) + \frac{1}{2}g(z) + \frac{1}{8}\|y-z\|^2$

$L(y,z) = \frac{1}{2}(y+z)$ */

$$= \left(\begin{array}{ll} \inf\limits_{(y,z)} & \frac{1}{2}f(y) + \frac{1}{2}g(z) + \frac{1}{8}\|y-z\|^2 \\ \text{s.t.} & \frac{1}{2}(y+z) = x \leftrightarrow y+z = 2x \\ & (y,z) \in \mathcal{H}\times\mathcal{H} \end{array}\right) = \operatorname{pav}(f,g)(x) \quad \text{// from (1)}$$

first note that:

$$\operatorname{pav}(f,g)(x) = \inf_{(y,z)} \frac{1}{2}f(y) + \frac{1}{2}g(z) + \frac{1}{4}\|y-z\|^2 + \iota_{y+z=2x}(y,z)$$

$$= \frac{1}{2}\left(\inf_{y\in\mathcal{H}} \inf_{z\in\mathcal{H}} f(y) + g(z) + \frac{1}{4}\|y-z\|^2 + \iota_{y+z=2x}(y,z)\right)$$ // recall $\inf\limits_{(x,y)\in X\times Y} f(x,y) = \inf\limits_{x\in X}\inf\limits_{y\in Y} f(x,y) = \inf\limits_{y\in Y}\inf\limits_{x\in X} f(x,y)$ Fact 1.8.3.

$$= \frac{1}{2}\left(\inf_{y\in\mathcal{H}} f(y) + \underbrace{\inf_{z\in\mathcal{H}} g(z) + \frac{1}{4}\|y-z\|^2 + \iota_{z=2x-y}(y,z)}_{\|}\right)$$

/* $\left(\inf\limits_{z\in\mathcal{H}: z=\underbrace{2x-y}_{\text{this is a single vector, so inf will be achieved at } z=2x-y}} g(z) + \frac{1}{4}\|y-z\|^2\right) = g(2x-y) + \frac{1}{4}\|y-2x+y\|^2 = g(2x-y) + \frac{1}{4}\|2(y-x)\|^2 = g(2x-y) + \frac{1}{\cancel{4}}\cdot\cancel{4}\|y-x\|^2$ */

$$= \frac{1}{2}\left(\inf_{y\in\mathcal{H}} \underbrace{f(y) + g(2x-y) + \|y-x\|^2}_{h(y,x)}\right) = \frac{1}{2}\inf_{y\in\mathcal{H}} h(y,x) \quad \ldots (2)$$

define: $h : h(y,x) = f(y) + g(2x-y) + \|x-y\|^2$

$$= \underbrace{\tfrac{1}{2}\cdot 2\|y\|^2}_{\frac{\gamma}{2}\|y\|^2 : \gamma = 2} + \underbrace{\left(f(y) + g(2x-y) - 2\langle y|x\rangle + \|x\|^2\right)}_{\in \Gamma_0(\mathcal{H}) \text{ in } y}$$

so, $h \in \Gamma_0(\mathcal{H})$, strongly convex in y // as $\underbrace{f}_{\in\Gamma_0(\mathcal{H})} + \frac{\gamma}{2}\|\cdot\|^2$: strongly convex

. . . (3)

/* from (2), (3) and using

Corollary 11.16.
[$f \in \Gamma_0(\mathcal{H})$, strongly convex] $\Rightarrow$ f: supercoercive, has exactly one minimizer over $\mathcal{H}$.

we have:

$$\mathrm{pav}(f,g)(x) = \tfrac{1}{2}\inf_{y\in\mathcal{H}} h(y,x) \text{ has a unique minimizer}$$

so,

$$\mathrm{pav}(f,g) = L \triangleright F$$

(iii)

from (ii):

$$\mathrm{pav}(f,g) = L \triangleright F$$

recall:

Proposition 12.34.
[$f : \mathcal{H} \to]-\infty, +\infty]$,
$\mathcal{K}$: real Hilbert space
$L : \mathcal{H} \to \mathcal{K}$]

(i) $\mathrm{dom}(L \triangleright f) = L(\mathrm{dom}\, f)$

(ii) (f: convex, L: affine) $\Rightarrow$ $L \triangleright f$: convex /*wow!*/

$\therefore \mathrm{dom}\, \mathrm{pav}(f,g) = \mathrm{dom}\,(L \triangleright F) = L\, \mathrm{dom}\, F$

$$= L\left\{(y,z) \in \mathcal{H}\times\mathcal{H} \,\middle|\, F(y,z) = \tfrac{1}{2}f(y) + \tfrac{1}{2}g(z) + \tfrac{1}{8}\|y-z\|^2 < +\infty\right\}$$

$$= L\left\{(y,z) \in \mathcal{H}\times\mathcal{H} \,\middle|\, f(y) < +\infty,\ g(z) < +\infty\right\} \quad // L(y,z) = \tfrac{1}{2}y + \tfrac{1}{2}z$$

$$= \left\{\tfrac{1}{2}y + \tfrac{1}{2}z \in \mathcal{H}\times\mathcal{H} \,\middle|\, y \in \mathrm{dom}\, f,\ z \in \mathrm{dom}\, g\right\}$$

$$= \tfrac{1}{2}\left\{y + z \in \mathcal{H}\times\mathcal{H} \,\middle|\, y \in \mathrm{dom}\, f,\ z \in \mathrm{dom}\, g\right\}$$

$$= \tfrac{1}{2}\left(\{y \in \mathcal{H} \mid y \in \mathrm{dom}\, f\} + \{z \in \mathcal{H} \mid z \in \mathrm{dom}\, g\}\right)$$

$$= \tfrac{1}{2}\,\mathrm{dom}\, f + \tfrac{1}{2}\,\mathrm{dom}\, g$$

(iv) F: convex, L: linear $\Rightarrow$ L: affine $\therefore (L \triangleright f) = \mathrm{pav}(f,g)$: convex

now $f, g \in \Gamma_0(\mathcal{H}) \Rightarrow f, g$: proper $\Rightarrow \mathrm{dom}\, f \neq \emptyset$, $\mathrm{dom}\, g \neq \emptyset$

in (iii)

$$\mathrm{dom}\, \mathrm{pav}(f,g) = \tfrac{1}{2}\,\mathrm{dom}\, f + \tfrac{1}{2}\,\mathrm{dom}\, g \neq \emptyset$$

also, as for $\forall_{\tilde{x}\in\mathcal{H}}$ $\mathrm{pav}(f,g)(\tilde{x}) = (L \triangleright f)(\tilde{x}) \Rightarrow \mathrm{pav}(f,g)(\tilde{x}) \neq -\infty$

$\therefore \mathrm{dom}\, \mathrm{pav}(f,g) \neq \emptyset$, $\mathrm{pav}(f,g)(\mathcal{H}) \not\ni -\infty$

$\updownarrow$ def

$\mathrm{pav}(f,g)$: proper

So, $\mathrm{pav}(f,g)$: proper, convex function. $\blacksquare$

* Proposition 11.11. (Coercivity of a function in terms of lower level set)
[$f : H \to [-\infty, +\infty]$] f : coercive $\leftrightarrow$ $(\mathrm{lev}_{\leq \xi} f)_{\xi \in \mathbb{R}}$: bounded

Part 1

Sunday, February 21, 2016 1:51 PM

* Theorem 14.3.

[$f \in \Gamma_0(\mathcal{H})$, $\gamma \in \mathbb{R}_{++}$]

(i) $\frac{1}{2\gamma}\|\cdot\|^2 = \left(f \,\square\, \frac{1}{2\gamma}\|\cdot\|^2\right) + \left(f^* \,\square\, \frac{\gamma}{2}\|\cdot\|^2\right) \circ \frac{1}{\gamma}\mathrm{Id} = {}^{\gamma}f + {}^{(\frac{1}{\gamma})}(f^*) \circ \frac{1}{\gamma}\mathrm{Id}$

(ii) $\mathrm{Id} = \mathrm{Prox}_{\gamma f} + \gamma\, \mathrm{Prox}_{\frac{1}{\gamma}f^*} \circ \frac{1}{\gamma}\mathrm{Id}$

(iii) $\forall_{x \in \mathcal{H}}$

$$f(\mathrm{Prox}_{\gamma f} x) + f^*\left(\mathrm{Prox}_{\frac{1}{\gamma}f^*}\left(\frac{x}{\gamma}\right)\right) = \left\langle \mathrm{Prox}_{\gamma f} x \,\middle|\, \mathrm{Prox}_{\frac{1}{\gamma}f^*} \frac{x}{\gamma} \right\rangle$$

2/21/2016 3:35 PM

Proof: take $\tilde{\gamma} = \frac{1}{\gamma} \in \mathbb{R}_{++}$, take $\hat{f} = f^* \in \Gamma_0(\mathcal{H})$ /* $\because f : \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H})$ using Fenchel-Moreau theorem */

/* Recall:

* Proposition 12.15. /* Properties of power norm infimal convolution of a convex function */

[$\hat{f} \in \Gamma_0(\mathcal{H})$
$\tilde{\gamma} \in \mathbb{R}_{++}$
$p \in]1, +\infty[$

$\hat{f} \,\square\, \frac{1}{\tilde{\gamma}p}\|\cdot\|^p : \mathcal{H} \to]-\infty, +\infty] : x \mapsto \inf_{y \in \mathcal{H}} \left(\hat{f}(y) + \frac{1}{\tilde{\gamma}p}\|x - y\|^p\right)$]

- $\hat{f} \,\square\, \frac{1}{\tilde{\gamma}p}\|\cdot\|^p$: convex, real-valued, continuous, exact, infimum uniquely attained.

*

Example 13.4

[$\hat{f} : \mathcal{H} \to]-\infty, +\infty]$, proper;
$\tilde{\gamma} \in \mathbb{R}_{++}$;
$\tilde{f} = \hat{f} + \frac{1}{2\tilde{\gamma}}\|\cdot\|^2$]

$\tilde{f}^* = \frac{\tilde{\gamma}}{2}\|\cdot\|^2 - {}^{\tilde{\gamma}}\hat{f} \circ \tilde{\gamma}\mathrm{Id} = \frac{\tilde{\gamma}}{2}\|\cdot\|^2 - \underbrace{\left(\hat{f} \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right)}_{{}^{\tilde{\gamma}}\hat{f}} \circ \tilde{\gamma}\mathrm{Id}$

* Proposition 14.1.

[$\hat{f} \in \Gamma_0(\mathcal{H})$; $\gamma \in \mathbb{R}_{++}$]

$\left(\hat{f} + \frac{\gamma}{2}\|\cdot\|^2\right)^* = \hat{f}^* \,\square\, \frac{1}{2\gamma}\|\cdot\|^2 = {}^{\gamma}(f^*)$

$$\tilde{f}^* = \left(\hat{f} + \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right)^* = \frac{\tilde{\gamma}}{2}\|\cdot\|^2 - \underbrace{{}^{\tilde{\gamma}}\hat{f}}_{\left(\underbrace{\hat{f}}_{\in \Gamma_0(\mathcal{H})} \,\square\, \underbrace{\frac{1}{2\tilde{\gamma}}}_{\in \mathbb{R}_{++}}\|\cdot\|^2\right)} \circ \tilde{\gamma}\mathrm{Id} = \frac{\tilde{\gamma}}{2}\|\cdot\|^2 - \left(\hat{f} \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right) \circ \tilde{\gamma}\mathrm{Id}$$

$\Downarrow$

${}^{\tilde{\gamma}}\hat{f} = \hat{f} \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2$: convex, continuous, real valued, exact

$\Rightarrow {}^{\tilde{\gamma}}\hat{f} = f \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2$

$$\Rightarrow \underbrace{\left(\hat{f} + \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right)^*}_{\hat{f}^* \,\square\, \frac{1}{2\tilde{\gamma}}} = \frac{\tilde{\gamma}}{2}\|\cdot\|^2 - \left(\hat{f} \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right) \circ \tilde{\gamma}\mathrm{Id}$$

$\leftrightarrow \frac{\tilde{\gamma}}{2}\|\cdot\|^2 = \left(\hat{f}^* \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right)^* + \left(\hat{f} \,\square\, \frac{1}{2\tilde{\gamma}}\|\cdot\|^2\right) \circ \tilde{\gamma}\mathrm{Id}$ /* set $\gamma = \frac{1}{\tilde{\gamma}}$. $\tilde{f} = f^*$, and recall from Fenchel-Moreau that, $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^{**} = f$ */

$\leftrightarrow \frac{1}{2\gamma}\|\cdot\|^2 = \left(\underbrace{f^{**}}_{f} \,\square\, \frac{\gamma}{2}\|\cdot\|^2\right)^* + \left(f^* \,\square\, \frac{\gamma}{2}\|\cdot\|^2\right) \circ \frac{1}{\gamma}\mathrm{Id}$

$= \underbrace{\left(f \,\square\, \frac{\gamma}{2}\|\cdot\|^2\right)^*}_{{}^{\gamma}f} + \underbrace{\left(f^* \,\square\, \frac{\gamma}{2}\|\cdot\|^2\right)}_{{}^{\gamma}f^*} \circ \frac{1}{\gamma}\mathrm{Id}$ ☺

(ii) this part of the proof is incomplete. I need to comeback to it later.

* Proposition 12.29.

[$f \in \Gamma_0(\mathcal{H})$

$\gamma \in \mathbb{R}_{++}$]

${}^{\gamma}f : \mathcal{H} \to \mathbb{R}$, Fréchet differentiable

$\nabla({}^{\gamma}f) = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f})$; γ^{-1} Lipschitz continuous.

so, $\nabla^{\gamma} f = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f}) \Leftrightarrow \nabla^{\gamma} f(\cdot) = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f})(\cdot) = \frac{1}{\gamma}(\cdot - \mathrm{Prox}_{\gamma f}\cdot)$

$\nabla^{\gamma} f^* = \frac{1}{\gamma}(\mathrm{Id} - \mathrm{Prox}_{\gamma f^*})$

now $\left({}^{\gamma}f^* \circ \frac{1}{\gamma}\mathrm{Id}\right)(x) = {}^{\gamma}f^*\left(\frac{1}{\gamma}x\right)$

recall, $(\nabla \cdot) = (D \cdot)^T$

chain rule says: $D(R \circ T)(x) = DR(Tx) \circ DT(x)$

so, $D(\underbrace{{}^{\gamma}f^* \circ \frac{1}{\gamma}\mathrm{Id}}_{\nabla})(x) = D\left[{}^{\gamma}f^*\left(\frac{1}{\gamma}\mathrm{Id}\,x\right)\right] \circ D\frac{1}{\gamma}\mathrm{Id}(x)$

$= D\left[{}^{\gamma}f^*\left(\frac{x}{\gamma}\right)\right] \circ \frac{1}{\gamma} \underbrace{Dx}_{\mathrm{Id}}$

$= D\left[{}^{\gamma}f^*\left(\frac{x}{\gamma}\right)\right] \circ \frac{1}{\gamma}\mathrm{Id}$

$\therefore D\nabla(x)\,y = \left(D\,{}^{\gamma}f^*\left(\frac{x}{\gamma}\right) \circ \frac{1}{\gamma}\mathrm{Id}\right)(y)$

$= \underbrace{D\,{}^{\gamma}f^*\left(\frac{x}{\gamma}\right)}_{\text{linear operator}}\left(\frac{y}{\gamma}\right) = \frac{1}{\gamma} D\,{}^{\gamma}f^*\left(\frac{x}{\gamma}\right)(y)$

* $DT(x) \overset{\text{def}}{\Leftrightarrow} \lim_{0 \neq y \to 0} \frac{\|T(x+y) - Tx - DT(x)y\|}{\|y\|} = 0$

here $T = \mathrm{Id}$ $\quad \lim_{0 \neq y \to 0} \frac{\|x + y - x - \overbrace{(Dx)}^{\mathrm{Id}\ \text{satisfies the equation}} y\|}{\|y\|} = 0$

$\therefore Dx = \mathrm{Id}$ */

now the relation between ∇ and D says

$\forall_{y \in \mathcal{H}} \quad D\nabla(x)y = \langle y | \nabla\nabla(x) \rangle$

$\Leftrightarrow \frac{1}{\gamma} \underbrace{D\,{}^{\gamma}f^*\left(\frac{x}{\gamma}\right)}(y) = \langle y | \nabla\nabla(x) \rangle$

from (i)

$\frac{1}{2\gamma}\|\cdot\|^2 = {}^{\gamma}f + {}^{\gamma}f^* \circ \frac{1}{\gamma}\mathrm{Id}$

taking derivative on both sides.

$\mathrm{Id} = \mathrm{Prox}_{\gamma f} + \gamma\, \mathrm{Prox}_{\frac{1}{\gamma}f^*} \circ \frac{1}{\gamma}\mathrm{Id}$.

(iii) incomplete.

$\blacksquare$ ☺

Theorem 14.17.
(Moreau-Rockafellar theorem)

2/24/2016 7:43 AM [[$f \in \Gamma_0(\mathcal{H})$, $u \in \mathcal{H}$]]

$f - \langle \cdot | u \rangle$: coercive $\Leftrightarrow u \in \operatorname{int} \operatorname{dom} f^*$

Proof:

Required info:

* Proposition 14.16. (Alternative characterization of coercive functions */

[$f \in \Gamma_0(\mathcal{H})$] /* info: f : coercive $\Leftrightarrow \lim_{\|x\| \to +\infty} f(x) = +\infty$ */

(i) f : coercive $\Leftrightarrow$

(ii) $(\mathrm{lev}_{\leq \xi} f)_{\xi \in \mathbb{R}}$: bounded $\Leftrightarrow$

(iii) $\varliminf_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} > 0$ $\Leftrightarrow$

(iv) $\exists_{(\alpha, \beta) \in \mathbb{R}_{++} \times \mathbb{R}}$ $f \geq \alpha\|\cdot\| + \beta$ $\Leftrightarrow$

(v) f^* : bounded above on a neighborhood of 0 $\Leftrightarrow$

(vi) $0 \in \operatorname{int} \operatorname{dom} f^*$

* Proposition 13.20. (Easy conjugate formulas)

(vi) $0 \in \operatorname{int} \operatorname{dom} f^*$

* Proposition 13.20. (Easy conjugate formulas)

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

(i) $\forall_{\alpha \in \mathbb{R}_{++}} \quad (\alpha f)^* = \alpha f^*(\cdot/\alpha)$

(ii) $\forall_{\alpha \in \mathbb{R}_{++}} \quad (\alpha f(\cdot/\alpha))^* = \alpha f^*$

(iii) $\forall_{y \in \mathcal{H}} \forall_{\alpha \in \mathbb{R}} \quad (\tau_y f + \langle \cdot | v \rangle + \alpha)^* = \tau_v f^* + \langle y | \cdot \rangle - \langle y | v \rangle - \alpha$ /* where $\tau_y f$: translation of f by y, $x \mapsto f(x-y)$ */

(iv) $L : \in \mathcal{B}(\mathcal{H})$, bijective $\Rightarrow (f \circ L)^* = f^* \circ L^{-1}$

(v) $f^{\vee *} = f^{*\vee}$ /* $f^\vee$: reversal of a function f i.e. $f^\vee : x \mapsto f(-x)$ */

(vi) (V: closed linear subspace of $\mathcal{H}$, $\operatorname{dom} f \subseteq V$) $\Rightarrow (f|_V)^* + \iota_V = f^* + \iota_V$

$f - \langle \cdot | u \rangle$: coercive

$\Leftrightarrow 0 \in \operatorname{int} \operatorname{dom} (f - \langle \cdot | u \rangle)^*$

/* now lets us show, $(f - \langle \cdot | u \rangle)^* = \tau_{-u} f^*$

$(\tau_y f + \langle \cdot | v \rangle + \alpha)^* = \tau_v f^* + \langle y | \cdot \rangle - \langle y | v \rangle - \alpha \quad \forall_{y, v, \alpha}$

set $y := 0$ then $\tau_y f(\cdot) = f(\cdot - \ ) = f(\cdot)$

$v := -u$, $\alpha := 0$

then

$(f - \langle \cdot | u \rangle)^* = \tau_{-u}$ */

$\Leftrightarrow 0 \in \operatorname{int} \operatorname{dom} (\underbrace{\tau_{-u} f^*}) = \operatorname{int} \operatorname{dom} f^*(\cdot + u)$

$\tau_{-u} f^* = f^*(\cdot + u)$

$\Leftrightarrow u \in \operatorname{int} \operatorname{dom} f^*$

/* As, $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H}) \Rightarrow$

$-\infty \notin f^*(\mathcal{H})$.

$\operatorname{dom} f^* = \{x \in \mathcal{H} \mid f^*(x) < +\infty\}$

$\Rightarrow \operatorname{dom} f^* = \{x \in \mathcal{H} \mid -\infty < f^*(x) < +\infty\}$

$\exists_{M \in \mathbb{R}} \forall_{x \in \operatorname{dom} f} \quad f^*(x) \leq M$

$\therefore \operatorname{dom} f^* = \{x \in \mathcal{H} \mid f^*(x) \leq M\}$

$\Rightarrow \operatorname{int} \operatorname{dom} f^* = \{x \in \mathcal{H} \mid f^*(x) < M\}$

So. $0 \in \operatorname{int} \operatorname{dom} f^*(\cdot + u) = \{x \in \mathcal{H} \mid f^*(x+u) < M\}$

$\Leftrightarrow f^*(0+u) < M$

$\Leftrightarrow f^*(u) < M \Leftrightarrow u \in \operatorname{int} \operatorname{dom} f^*$ */

$\blacksquare$

Proposition. (14.19) (Conjugate of the difference)

[[$g : \mathcal{H} \to]-\infty, +\infty]$; $h \in \Gamma_0(\mathcal{H})$;

$f : \mathcal{H} \to [-\infty, +\infty] : x \mapsto \begin{cases} g(x) - h(x), & x \in \operatorname{dom} g = \{\tilde{x} \in \mathcal{H} \mid g(\tilde{x}) < +\infty\} \\ +\infty, & x \notin \operatorname{dom} g \end{cases}$

]] $\Rightarrow$

$\forall_{u \in \mathcal{H}} \quad f^*(u) = \sup_{v \in \operatorname{dom} h^*} \left(g^*(u+v) - h^*(v) \right)$

<u>Proof</u>: $\forall u \in \mathcal{H}$

/* recall

***

Corollary 13.33. (An immediate consequence of Fenchel-Moreau theorem)

[$f \in \Gamma_0(\mathcal{H})$]

- $f^* \in \Gamma_0(\mathcal{H})$
- $f^{**} = f$

*/

$f^*(u) = \sup_{x \in \mathcal{H}} \left(\langle x | u \rangle - f(x) \right)$

$= \sup_{x \in \mathcal{H}} \left(\langle x | u \rangle - (g(x) - h(x)) \ [x \in \operatorname{dom} g] \right)$

$x \in H$

$$= \sup_{x \in H} \left(\langle x|u \rangle - (g(x) - h(x)) \, [x \in \operatorname{dom} g] \right)$$

$$= \sup_{x \in \operatorname{dom} g} \left(\langle x|u \rangle - g(x) + h(x) \right) \quad /* \text{But } h \in \Gamma_0(H), \text{ so } h^{**} = h \; */$$

$$= \sup_{x \in \operatorname{dom} g} \left(\langle x|u \rangle - g(x) + \underbrace{h^{**}(x)}_{\sup_{v \in \operatorname{dom} h^*} \left(-h^*(v) + \langle v|x \rangle \right)} \right) \quad /* \text{By def: } f^*(u) = \sup_{x \in H} \left(-f(x) + \langle x|u \rangle \right) = \sup_{x \in \operatorname{dom} f} \left(-f(x) + \langle x|u \rangle \right) \; */$$

$$= \sup_{x \in \operatorname{dom} g} \left(\underbrace{\langle x|u \rangle - g(x)}_{\text{constant w.r.t } v\text{, so can be taken inside}} + \sup_{v \in \operatorname{dom} h^*} \left(\langle v|x \rangle - h^*(v) \right) \right)$$

interchangeable

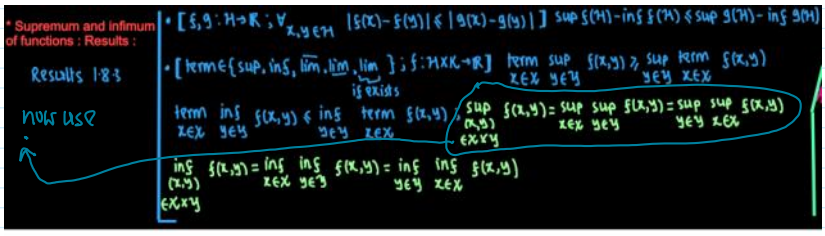
$$= \sup_{x \in \operatorname{dom} g} \sup_{v \in \operatorname{dom} h^*} \left(\langle x|u \rangle - g(x) + \underbrace{\langle v|x \rangle}_{\langle x|v \rangle} - h^*(v) \right) \quad /*$$

$$= \sup_{v \in \operatorname{dom} h^*} \sup_{x \in \operatorname{dom} g} \left(\langle x|u+v \rangle - g(x) - h^*(v) \right)$$

$$= \sup_{v \in \operatorname{dom} h^*} \left(-h^*(v) + \underbrace{\sup_{x \in \operatorname{dom} g} \left(-g(x) + \langle x|u+v \rangle \right)}_{g^*(u+v)} \right) \quad */$$

$$= \sup_{v \in \operatorname{dom} h^*} \left(g^*(u+v) - h^*(v) \right)$$

$\blacksquare$



*/

Part 1

Thursday, February 25, 2016 6:42 AM

2/25/2016 7:53 AM

* Proposition 15.1.

[$f, g \in \Gamma_0(\mathcal{H})$: $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$]

$f^* \square g^*$: proper, convex

$f^* \square g^*$: has continuous affine minorant.

$(f+g)^* = (f^* \square g^*)^{**} = (f^* \square g^*)^{\vee}$

Proof:

/*

* Proposition 13.21. (Very important)

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

(i) $(f \square g)^* = f^* + g^*$

(ii) f, g : proper $\Rightarrow (f+g)^* \leq f^* \square g^*$

(iii) $\forall_{\gamma \in \mathbb{R}_{++}} \; (\gamma f)^* = f^* + \frac{\gamma}{2} \|\cdot\|^2$

(iv) $L \in \mathcal{B}(\mathcal{H}, \mathcal{K}) \Rightarrow (L \triangleright f)^* = f^* \circ L^*$

(v) $L \in \mathcal{B}(\mathcal{K}, \mathcal{H}) \Rightarrow (f \circ L)^* \leq L^* \triangleright f^*$

Corollary 13.33. (Fenchel-Moreau theorem)

[$f \in \Gamma_0(\mathcal{H})$]

- $f^* \in \Gamma_0(\mathcal{H})$
- $f^{**} = f$

*/

$f, g \in \Gamma_0(\mathcal{H}) \Rightarrow f^*, g^* \in \Gamma_0(\mathcal{H})$

$\Rightarrow (f^* \square g^*)^* = f^{**} + g^{**}$ (1)

$= f + g \in \Gamma_0(\mathcal{H})$ /* Γness is preserved under addition when number of functions are finite and none contains $-\infty$ (corollary 9.4), what is left is properness, now $\operatorname{dom} f+g = \operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$ given ✓ so, $\operatorname{dom}(f+g) \neq \emptyset$, $-\infty \notin (f+g)(\mathcal{H})$, and $f+g$: proper $\therefore f+g \in \Gamma_0(\mathcal{H})$ */

$\therefore (f^* \square g^*)^*$: proper $\Rightarrow (f^* \square g^*)$: proper

as, $f^*, g^* \in \Gamma_0(\mathcal{H}) \Rightarrow (f^* \square g^*)$: convex

$\therefore (f^* \square g^*)$: proper, convex ✓

now $(f^* \square g^*)^* \in \Gamma_0(\mathcal{H}) \Rightarrow (f^* \square g^*)^* \neq +\infty$

(1) $\Leftrightarrow f^* \square g^*$: has a continuous affine minorant ✓

as $f^* \square g^*$ has a continuous affine minorant

$\Rightarrow (f^* \square g^*)^{**} = (f^* \square g^*)^{\vee}$ /* $\check{\mathfrak{f}}$: lower semicontinuous convex envelope of $\mathfrak{f}$ */

/* $(f^* \square g^*)^{**} = [(f^* \square g^*)^*]^* = (f+g)^*$

$f^{**} + g^{**}$ From (1)

$= f + g$ /* $\because f, g \in \Gamma_0(\mathcal{H}) \Rightarrow f^*, g^* \in \Gamma_0(\mathcal{H})$ using Fenchel-Moreau theorem */

$\therefore (f^* \square g^*)^{**} = (f+g)^* = (f^* \square g^*)^{\vee}$

$\blacksquare$

* Proposition 12.9. ★

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

(i) $f^*(0) = -\inf f(\mathcal{H})$ /* 즉, $\inf f(\mathcal{H})$ 의 음수 취한 conjugate function 의 0 에 evaluate 하면 -1 되는 것과 동치! */

(ii) $-\infty \in f^*(\mathcal{H}) \Leftrightarrow f = +\infty \Leftrightarrow f^* = -\infty$

(iii) f^* : proper $\Rightarrow f$: proper

(iv) [$u \in \mathcal{H}$]

$f^*(u) = \sup_{x \in \operatorname{dom} f} \left(\langle x | u \rangle - f(u) \right) = \sup_{(x, \xi) \in \operatorname{epi} f} \left(\langle x | u \rangle - \xi \right)$

(v) $f^* = L^*_{\operatorname{gra} f}(\cdot, -1) = L^*_{\operatorname{epi} f}(\cdot, -1)$

Proposition 12.11. (infimal convolution between convex functions preserves convexity)

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

$f \square g$: convex

* Proposition 13.10. ★

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

(i) [$(u, v) \in \mathcal{H} \times \mathbb{R}$] $(u, v) \in \operatorname{epi} f^* \leftrightarrow \langle \cdot | u \rangle - v \leq f$ /* Any continuous affine minorant of a function has the (slope, shift) in the epigraph of the conjugate function! */

(ii) $f^* = +\infty \leftrightarrow f$: possesses no continuous affine minorant

(iii) [$\operatorname{dom} f^* \neq \emptyset$] f : bounded below on every bounded subset of $\mathcal{H}$.

Proof: See 2/15/2016 2:44 PM.

Proposition 13.39.

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

- f : has a continuous affine minorant $\Rightarrow f^{**} = \check{f}$

  $\Updownarrow$ $\operatorname{dom} f^* \neq \emptyset$ (from proposition 13.10(iii))

- f : does not have a continuous affine minorant $\Rightarrow f^{**} = -\infty$

* Theorem 15.3. (Attouch-Brézis theorem)

[$f, g \in \Gamma_0(\mathcal{H})$

$0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom} g) \leftrightarrow \operatorname{cone}(\operatorname{dom} f - \operatorname{dom} g) = \overline{\operatorname{span}}(\operatorname{dom} f - \operatorname{dom} g)$]

$(f+g)^* = f^* \boxdot g^* \in \Gamma_0(\mathcal{H})$

Proof:

Proof:

$f, g \in \Gamma_0(\mathcal{H})$

$\Rightarrow f+g \in \Gamma_0(\mathcal{H})$ /* Γ_0 : closed under addition */

$\Rightarrow (f+g)^* \in \Gamma_0(\mathcal{H})$ /* Fenchel-Moreau corollary: $f \in \Gamma_0(\mathcal{H}) \Rightarrow \begin{cases} f^* \in \Gamma_0(\mathcal{H}) \\ f^{**} = f \end{cases}$ */

$\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$ because

$0 \in \operatorname{sri}(\underbrace{\operatorname{dom} f - \operatorname{dom} g}_{\text{ガ}}) = \{ x \in \text{ガ} \mid \operatorname{cone}(\text{ガ}) = \overline{\operatorname{span}}(\text{ガ}) \}$

$\Leftrightarrow 0 \in \underbrace{\text{ガ}}_{\operatorname{dom} f - \operatorname{dom} g} \Leftrightarrow 0 \in \operatorname{dom} f - \operatorname{dom} g \Leftrightarrow \exists_{\substack{\text{ガ} \in \operatorname{dom} f \\ f(\text{ガ}) < +\infty}} \exists_{\substack{\text{ヌ} \in \operatorname{dom} g \\ g(\text{ヌ}) < +\infty}} \quad 0 = \text{ガ} - \text{ヌ} \Leftrightarrow \underbrace{\text{ガ}}_{\in \operatorname{dom} f} = \overbrace{\text{ヌ}}^{\in \operatorname{dom} g} \Leftrightarrow \exists_{\text{ガ} = \text{ヌ}} \ \text{ガ} \in \operatorname{dom} f \cap \operatorname{dom} g.$

$\therefore \operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$

set, $\phi : x \mapsto f(x+z)$

$\psi : y \mapsto g(y+z)$

Note that, $0 \in \operatorname{dom} \phi - \operatorname{dom} \psi$, $\operatorname{dom} \phi - \operatorname{dom} \psi = \operatorname{dom} f - \operatorname{dom} g$

$$
\begin{aligned}
0 \in \operatorname{dom} f - \operatorname{dom} g &\Leftrightarrow \exists_z \quad z \in \operatorname{dom} f \cap \operatorname{dom} g \\
&\Leftrightarrow \quad f(z) < +\infty, \ g(z) < +\infty \\
&\Leftrightarrow \quad f(0+z) = \phi(0) < +\infty, \ g(0+z) = \psi(0) < +\infty \\
&\Leftrightarrow \quad 0 \in \operatorname{dom} \phi, \ 0 \in \operatorname{dom} \psi \\
&\Leftrightarrow \quad 0 \in \operatorname{dom} \phi \cap \operatorname{dom} \psi
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{dom} f - \operatorname{dom} g &= \{ x - y \mid x \in \operatorname{dom} f, \ y \in \operatorname{dom} g \} \\
&= \{ x - y \mid f(x) < +\infty, \ g(y) < +\infty \} \\
&= \{ x - y \mid f(x - z + z) = \phi(x - z) < +\infty
\end{aligned}
$$

(incomplete...)

Part 2

Sunday, August 7, 2016 8:44 PM

* Proposition 15.12. 8/7/2016 9:15 PM

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

$\mu = \inf\{f+g\}(\mathcal{H})$,

$\mu^* = \inf\{f^* + g^{*}\}(\mathcal{H})$]

(i) $\mu \geq -\mu^*$ (essentially the weak duality condition)

(ii) $\Delta(f,g) \in [0, +\infty]$

(iii) $\mu = -\mu^* \Leftrightarrow \Delta(f,g) = 0$

Proof:

(i) $\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}}$ $f(x) + g(x) + f^*(u) + g^*(-u)$

$$= \underbrace{f(x) + f^*(u)}_{\geq \langle x, u \rangle} + \underbrace{g(x) + g^*(-u)}_{\geq \langle x, -u \rangle = -\langle x, u \rangle}$$

// Fenchel-Young inequality: [$f: \mathcal{H} \to]-\infty, +\infty]$, proper] $\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}}$ $f(x) + f^*(u) \geq \langle x, u \rangle$

$$\geq \langle x, u \rangle - \langle x, u \rangle = 0$$

$\Leftrightarrow \forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}}$ $f(x) + g(x) \geq -f^*(u) - g^*(-u)$

$\Leftrightarrow \forall_{u \in \mathcal{H}}$ $\inf_{x \in \mathcal{H}} f(x) + g(x) \geq -f^*(u) - g^*(-u)$

$\Leftrightarrow \underbrace{\inf_{x \in \mathcal{H}} f(x) + g(x)}_{\mu} \geq \underbrace{\sup_{u \in \mathcal{H}} -f^*(u) - g^*(-u)}_{= -\inf_{u \in \mathcal{H}} f^*(u) + g^*(-u)} = -\mu^*$

$\therefore \mu \geq -\mu^*$

(ii), (iii) follows from definition. ■

Proposition 15.13. (Strong duality) 8/7/2016 10:00 PM

[$f, g \in \Gamma_0(\mathcal{H})$: $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom} g)$]

$\inf\{f+g\}(\mathcal{H}) = -\min(f^* \square g^*)(\mathcal{H})$

Proof:

$f: \mathcal{H} \to]-\infty, +\infty] \Rightarrow f^*(0) = -\inf f(\mathcal{H})$ // Proposition 13.3.(i)

$f, g \in \Gamma_0(\mathcal{H})$: $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom} g) \Rightarrow (f+g)^* = f^* \square g^* \in \Gamma_0(\mathcal{H})$ // Attouch-Brézis theorem

$\inf(f+g)(\mathcal{H}) = -(f+g)^*(0)$

$= -(f^* \square g^*)(0)$

$= -\min_{x \in \mathcal{H}} (f^*(x) + g^*(0-x))$ // $(f \square g)(y) = \min_{x \in \mathcal{H}} (f(x) + g(x-y))$

$= -\min_{x \in \mathcal{H}} (f^*(x) + \underbrace{g^*(-x)}_{= g^{*}(x)})$

$= -\min_{x \in \mathcal{H}} (f^*(x) + g^{*}(x))$

$\therefore \inf(f+g)(\mathcal{H}) = -\min(f^* + g^{*})(\mathcal{H})$ ■

Corollary 15.14. 8/22/2016 9:11 AM

[$f \in \Gamma_0(\mathcal{H})$; K: closed convex cone in $\mathcal{H}$, $0 \in \operatorname{sri}(K - \operatorname{dom} f)$] $\inf f(K) = -\min f^*(K^{\oplus})$

Proof:

// Proposition 15.13:

[$f, g \in \Gamma_0(\mathcal{H})$: $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom} g)$]

$\inf(f+g)(\mathcal{H}) = -\min(f^* + g^{*})(\mathcal{H})$ */ ... (i)

K: closed convex cone

$\Rightarrow g = \iota_K$

$\Rightarrow g^* = \iota_{K^{\oplus}}$ /* K: nonempty cone in $\mathcal{H} \Rightarrow g^* = \iota_{K^{\oplus}}$

(remember in 5.5(ii)) */ (ii)

$K^{\oplus}$: polar cone of K

$= \{u \in \mathcal{H} \mid \sup(C)(u) \leq 0\}$ */

$= \{u \in \mathcal{H} \mid \forall_{x \in C} \langle x, u \rangle \leq 0\}$ */ (v)

$\therefore \inf(f+g)(\mathcal{H})$

$= \inf(f + \iota_K)(\mathcal{H})$

$= \inf_{x \in K} f(x)$ /* comes from definition */

$= \inf f(K)$ ... (iii)

now, $\min(f^* + g^{*})(\mathcal{H}) = \min(f^* + \iota_{K^{\oplus}})(\mathcal{H})$ /* from (ii) */

$= \min_{x \in \mathcal{H}} f^*(x) + \iota_{K^{\oplus}}(x)$ /* $f^*(x) = f(-x)$ */

$= \min_{x \in \mathcal{H}} f^*(x) + \iota_{K^{\oplus}}(-x)$

$= \min_{x \in \mathcal{H} : -x \in K^{\oplus}} f^*(x)$

$-x \in K^{\oplus}$
$\Leftrightarrow \sup_{x \in K} \langle -x, x \rangle = -\|x\|^2 = 0 \Leftrightarrow x = 0$
$\Leftrightarrow \sup_{x \in K} \langle x, x \rangle = 0$
$\Leftrightarrow \sup_{x \in K} \|x\|^2 = \|x\|^2 = 0$
$\Leftrightarrow x = 0$ (since $\sup(C) = 0 \Leftrightarrow C \subseteq \{0\}$)

$= \min_{x \in K^{\oplus}} f^*(x)$

$= \min f^*(K^{\oplus})$ ... (iv)

from (i), (iii), (iv) $\Rightarrow$

$\inf f(K) = -\min f^*(K^{\oplus})$

■

* Conditions for zero duality gap for Fenchel-Rockafellar duality: 8/23/2016 11:56 AM

Proposition 15.22

[$f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{K})$, $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$]

$0 \in \operatorname{core}(\operatorname{dom} g - L(\operatorname{dom} f)) \Rightarrow$

$\inf(f + g \circ L)(\mathcal{H}) = -\min(f^* + \iota_{\mathcal{H}}^{\oplus} + g^{*})(\mathcal{K})$

this seperable sum is valid as we are talking about functions, so $f: \mathcal{H} \to]-\infty, +\infty]$, $g: \mathcal{K} \to]-\infty, +\infty]$

Proof:

Denote $\varphi = f \oplus g$: $\forall_{x \in \mathcal{H}, y \in \mathcal{K}}$ $\varphi(x, y) = f(x) + g(y)$ (-i)

$V = \operatorname{gra} L = \{(x, Lx) \mid x \in \mathcal{H}\}$ (0)

take any $(x, y) \in \mathcal{H} \times \mathcal{K}$

$\Rightarrow y - Lx \in K$ ... (i)

given: $0 \in \operatorname{core}(\operatorname{dom} g - L(\operatorname{dom} f))$ // $\operatorname{core}(C = \{x \in C \mid \operatorname{cone}(C - x) = \mathcal{H}\}$

// $0 \in \operatorname{core} C \Leftrightarrow \operatorname{cone}(C) = \mathcal{H}$

$\Leftrightarrow \mathbb{R}_{++} \underbrace{(\operatorname{dom} g - L(\operatorname{dom} f))}_{\{\tilde{b} - L\tilde{a} \mid \tilde{b} \in \operatorname{dom} g, \tilde{a} \in \operatorname{dom} f\}} = K$ ... (ii) $\Leftrightarrow \mathbb{R}_{++} C = \mathcal{H}$ // from Proposition 6.2. ($\forall C$: cone, $C \subseteq \mathcal{H}$) cone $C = \mathbb{R}_{++} C$

$\subseteq K$ $\subseteq \mathcal{H}$

From (i), (ii):

$\exists_{\gamma \in \mathbb{R}_{++}, b \in \operatorname{dom} g, a \in \operatorname{dom} f}$ $\gamma(b - La) = y - Lx$ // set $z = a - \frac{x}{\gamma} \Rightarrow x = \gamma(a - z)$

$\Leftrightarrow y = Lx + \gamma b - \gamma La$

$= L\gamma(a - z) + \gamma b - \gamma La$

$= \gamma L(a - z) + \gamma b - \gamma La$

$= \gamma(\cancel{La} - Lz + b - \cancel{La})$

$= \gamma(b - Lz)$

$\therefore \forall_{(x,y) \in \mathcal{H} \times \mathcal{K}} \exists_{\gamma \in \mathbb{R}_{++}, b \in \operatorname{dom} g, a \in \operatorname{dom} f}$ $\left| \begin{array}{l} x = \gamma(a - z) \\ y = \gamma(b - Lz) \end{array} \right.$ $\Leftrightarrow (x, y) = \gamma(a - z, b - Lz) = \gamma \begin{bmatrix} a - z \\ b - Lz \end{bmatrix} = \gamma\left(\begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} z \\ Lz \end{bmatrix}\right)$ // so, $\varphi\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = f(a) + g(b) < +\infty \Leftrightarrow \begin{bmatrix} a \\ b \end{bmatrix} \in \operatorname{dom} \varphi$

$\in \operatorname{dom} f \Rightarrow f(a) < +\infty$

$= \gamma(b - Lz)$

$\therefore \forall_{(x,y)\in \mathcal{H}\times\mathcal{K}} \; \exists_{\gamma\in\mathbb{R}_{++},\, b\in\operatorname{dom} g,\, a\in\operatorname{dom} f} \left| \begin{array}{l} x = \gamma(a-z) \\ y = \gamma(b-Lz) \end{array} \right. \Leftrightarrow (x,y) = \gamma(a-z, b-Lz) = \gamma \begin{bmatrix} a-z \\ b-Lz \end{bmatrix} = \gamma\left(\begin{bmatrix} a \\ b \end{bmatrix} - \underbrace{\begin{bmatrix} z \\ Lz \end{bmatrix}}_{\in V}\right)$ // so, $\Phi\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = f(a) + g(b) < +\infty \Leftrightarrow \begin{bmatrix} a \\ b \end{bmatrix} \in \operatorname{dom}\Phi$

$\in \operatorname{dom} f \Rightarrow f(a) < +\infty$

$\in \operatorname{dom} g \Rightarrow g(b) < +\infty$

$\in \underbrace{\gamma}_{\in\mathbb{R}_{++}}(\operatorname{dom}\Phi - V) \in \operatorname{cone}(\operatorname{dom}\Phi - V)$

$\therefore$ we have shown that, $\forall_{(x,y)\in\mathcal{H}\times\mathcal{K}} \; (x,y) \in \operatorname{cone}(\operatorname{dom}\Phi - V)$

$\Leftrightarrow \mathcal{H}\times\mathcal{K} \subseteq \operatorname{cone}(\operatorname{dom}\Phi - V)$

Naturally, $\operatorname{cone}(\operatorname{dom}\Phi - V) \subseteq \mathcal{H}\times\mathcal{K}$

$\Rightarrow \operatorname{cone}((\operatorname{dom}\Phi) - V) = \mathcal{H}\times\mathcal{K} \quad \ldots (\text{iii})$

so, $\vec{0} = (0,0) \in \operatorname{cone}((\operatorname{dom}\Phi) - V)$

$\Leftrightarrow 0 \in \operatorname{cone}(V - \operatorname{dom}\Phi) \subseteq \operatorname{sri}(V - \operatorname{dom}\Phi)$ // V: convex set

$\operatorname{int} C \subseteq \operatorname{core} C \subseteq \operatorname{sric} \subseteq \operatorname{ric} \subseteq \operatorname{bric} \subseteq C$

recall corollary 15.19:

[$f \in \Gamma_0(\mathcal{H})$, K: closed convex cone $\subseteq \mathcal{H}$

$0 \in \operatorname{sri}(K - \operatorname{dom} f)$] $\inf f(K) = -\min f^*(K^\oplus)$

in our case set $K := V$: closed convex cone

$\Phi = f \oplus g \in \Gamma_0(\mathcal{H}\times\mathcal{K})$, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{K})$

$\Rightarrow \inf \Phi(V) = -\min \Phi^*(V^\oplus)$ // As V: linear $\Rightarrow V^\oplus = V^\perp$ Proposition

subspace

orthogonal complement of V

$= -\min \Phi^*(V^\perp) \quad \ldots (\text{iv})$

Now, $V^\perp = \{(x,y)\in\mathcal{H}\times\mathcal{K} \mid \forall_{(\tilde{x},\tilde{y})\in V} \langle (x,y) | (\tilde{x},\tilde{y})\rangle = 0\}$

$\tilde{y} = L\tilde{x}$, $\mathcal{H} \xrightarrow{L} \mathcal{K}$, $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$

$\Updownarrow \; \forall_{(\tilde{x},\tilde{y})\in\mathcal{H}\times\mathcal{K}:\, \tilde{y}=L\tilde{x}} \; \left\langle \begin{bmatrix} x \\ y \end{bmatrix} \middle| \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} \right\rangle$

$\Updownarrow \; \forall_{\tilde{x}\in\mathcal{H}} \; \left\langle \begin{bmatrix} x \\ y \end{bmatrix} \middle| \begin{bmatrix} \tilde{x} \\ L\tilde{x} \end{bmatrix} \right\rangle = \langle x | \tilde{x}\rangle + \langle$

$\Updownarrow \; \forall_{\tilde{x}\in\mathcal{H}} \; \langle \tilde{x} | x + L^*y \rangle = 0$

$\Updownarrow \; x + L^*y = 0$ // As this is the

$\Updownarrow \; x = -L^*y$

also, $\Phi = f \oplus g \Rightarrow \Phi^* = f^* \oplus g^*$ // Proposition

$\ldots (\text{vi})$

now: $\Phi(V) = (f+g)(V) \rightsquigarrow \{(x, Lx) \mid x \in \mathcal{H}\}$

$= \left\{ (f+g)\begin{bmatrix} x \\ Lx \end{bmatrix} : x \in \mathcal{H} \right\}$

$= \{ f(x) + g(Lx) \mid x \in \mathcal{H} \}$

$\Rightarrow \inf \Phi(V) = \inf_{x\in\mathcal{H}} f(x) + g(Lx)$

$= \inf (f + g\circ L)(\mathcal{H})$

$\therefore \inf (f+g\circ L)(\mathcal{H}) = \inf \Phi(V)$ // use (iv) now

$= -\min \Phi^*(V^\perp)$ // use (v) now

$= -\min \Phi^*(\{(x,y)\in\mathcal{H}\times\mathcal{K} \mid x = -L^*y\})$

$= -\min \Phi^*(\{(-L^*y, y) : y \in K\})$

$= -\min_{y\in K} \Phi^*(-L^*y, y)$

$\Phi^* = (f^* \oplus g^*)$

$(f^* \oplus g^*)(-L^*y, y) =$

$= -\min_{y\in K} \left((f^* \circ L^*)(-y) + g^*(y) \right)$

// now set $y := -z \in K$

$= -\min_{-z\in K} \left((f^* \circ L^*)(z) + \underbrace{g^*(-z)}_{= g^{*\vee}(z)} \right)$

$= -\min_{-z\in K} \left((f^* \circ L^*) + g^{*\vee} \right)(z)$

$= -\min \left((f^* \circ L^*) + g^{*\vee} \right)(K)$

$\therefore \inf (f+g\circ L)(\mathcal{H}) = -\min (f^* \circ L^* + g^{*\vee})(K)$

■

Fact 15.25:

[$(f \in \Gamma_0(\mathcal{H}), g \in \Gamma_0(K))$: polyhedral 8/24/2016 4:21 PM

$L \in \mathcal{B}(\mathcal{H},K)$

K: finite dimensional and one of the following holds:

(i) $\operatorname{dom} g \cap \operatorname{ri} L(\operatorname{dom} f) \neq \emptyset$

(ii) $\mathcal{H}$ finite-dimensional, f polyhedral, $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$]

$\Rightarrow \inf (f+g\circ L)(\mathcal{H}) = -\min ((f^* \circ L^*) + g^{*\vee})(K)$

Proof:

$\mu = \inf (f+g\circ L)(\mathcal{H})$

assume $\mu > -\infty$ then, $\ldots (0)$

$\mu = \inf (f+g\circ L)(\mathcal{H})$

$= \inf_{x\in\mathcal{H}} f(x) + g(Lx)$ // recall graph $L = \{(\tilde{x}, L\tilde{x}) : \tilde{x} \in \mathcal{H}\}$

6:22

$\{(x,y) \in \mathcal{H} \times \mathcal{K} \mid x = -L^* y\}$ . . . (V)

$= 0$

$\langle y | L\tilde{x} \rangle = \langle x | \tilde{x} \rangle + \langle L^* y | \tilde{x} \rangle = \langle x + L^* y | \tilde{x} \rangle = 0$

$\langle L^* y | \tilde{x} \rangle$

only possibility */

13:27

$y\})$

$)$

this is linear too

$g^*(-L^* y) + g^*(y) = (f^* \circ L^*)(-y) + g^*(y) = \left((f^* \circ L^*)^\vee + g^* \right)(y)$

$= L^*(-y)$

$\parallel$

$(f^* \circ L^*)^\vee(y)$ // By definition of the reversal function

$= \{(Lx, y) : y = Lx\}$

$$= \inf_{\substack{x \in H \\ \{y \in K : y = Lx\} \\ (x,y) \in \operatorname{gra} L}} f(x) + g(y)$$

$$= \inf_{(x,y) \in \operatorname{gra} L} f(x) + g(y) = \inf_{\substack{y \in K \\ x \in H}} g(y) + f(x) : x \in L^{-1}y$$

$\Leftrightarrow (y,x) \in \operatorname{gra} L^{-1}$

$\Leftrightarrow L^{-1}y \ni x$

$$= \underbrace{\inf_{y \in K}}_{\text{independent variable}} \underbrace{\inf_{x \in H}}_{\text{dependent variable}} \left(g(y) + f(x) : x \in L^{-1}y \right)$$

$$= \inf_{y \in K} \Big(g(y) + \underbrace{\inf_{x \in L^{-1}y} f(x)}_{(L \triangleright f)(y)} \Big)$$

// recall the definition of infimal post composition:

$$(L \triangleright f) : y \mapsto \inf_{\substack{x \in H \\ Lx = y \Leftrightarrow (x,y) \in L \\ \Leftrightarrow (y,x) \in L^{-1} \\ \Leftrightarrow L^{-1}y \ni x}} f(x)$$

$$= \inf_{y \in K} \left(g(y) + (L \triangleright f)(y) \right) \quad \ldots (i)$$

/* recall Proposition 12.19.

$$\left[\begin{array}{l} f : X \to]-\infty, +\infty] \\ L : X \to K \end{array} \right] \Rightarrow \operatorname{dom}(L \triangleright f) = L(\operatorname{dom} f) \quad */$$

real Hilbert space

as $f \in \Gamma_0(H) \Leftrightarrow f : H \to]-\infty, +\infty] \Leftrightarrow \operatorname{dom}(L \triangleright f) = L(\operatorname{dom} f) \quad \ldots (ii)$

/* recall Proposition 13.21.

$[\, (f,g) : H \to]-\infty, +\infty] \,] \Rightarrow (L \triangleright f)^* = f^* \circ L^*$ // what a compact and beautiful statement

*/

Similarly, $(L \triangleright f)^* = f^* \circ L^* \quad \ldots (iii)$

Now

- assume antecedent (i): $\operatorname{dom} g \cap \operatorname{ri} L(\operatorname{dom} f) \neq \emptyset$ holds

  let $z \in \operatorname{dom} f$: $Lz \in \operatorname{dom} g \cap \operatorname{ri} L \operatorname{dom} f$

  $= \operatorname{dom} g \cap \operatorname{ri} \operatorname{dom}(L \triangleright f)$ // as this is nonempty, there will be one such z

  $\ldots (iv)$

- Assume antecedent (ii): H: finite-dimensional, f: polyhedral, $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$

  let $z \in \operatorname{dom} f$: $Lz \in \operatorname{dom} g \cap L(\operatorname{dom} f) = \operatorname{dom} g \cap \operatorname{dom}(L \triangleright f)$

  $\ldots (v)$

From (iv) and (v): $\mu = \inf_{y \in K} \left(g(y) + (L \triangleright f)(y) \right)$

$\leq g(Lz) + (L \triangleright f)(Lz)$ // $\because Lz \in K$

/* $(L \triangleright f)(y) = \inf_{\substack{x \in H \\ Lx = y}} f(x) \leq f(z) \ \forall x : Lx = y$

$\therefore (L \triangleright f)(Lz) \leq f(\tilde{z}) : \forall \tilde{x} : L\tilde{x} = Lz$

$\Downarrow$

choose $\tilde{x} = z$ satisfies $L\tilde{x} = Lz$

$\leq f(z)$ */

$\leq \underbrace{g(Lz)}_{<+\infty} + \underbrace{f(z)}_{<+\infty}$ // as $z \in \operatorname{dom} f$, $Lz \in \operatorname{dom} g$ from (iv) and (v)

$< +\infty \quad \ldots (vi)$

From (0) and (vi) we have

$-\infty < \mu < +\infty$

$\Leftrightarrow \mu \in \mathbb{R} \quad \ldots (vii)$

Assume $(L \triangleright f)(Lz) = \inf_{\substack{\tilde{x} \in H \\ L\tilde{x} = Lz}} f(\tilde{x}) = -\infty$

$\Rightarrow \exists (x_n)_{n \in \mathbb{N}} \subseteq H : Lx_n = Lz$ $\quad f(x_n) \to -\infty$

now, $\mu = \inf (f + g \circ L)(H)$

$\leq f(x) + g(Lx) \quad \forall x \in H$

$\leq f(x_n) + g(Lx_n)$

$\to -\infty \quad \ldots (viii)$

$\Downarrow$

contradicts with (vii)

$\therefore$ the assumption $(L \triangleright f)(Lz) \in \mathbb{R}$

Antecedent (i) holds $\Rightarrow (L \triangleright f)$: convex, proper // f: convex, L: affine $\Rightarrow (L \triangleright f)$: convex

if Antecedent (ii) $\Rightarrow (L \triangleright f)$: polyhedral, proper // By Rockafellar's convex analysis book

using Rockafellar's convex analysis book: Theorem 31.1 yields:

$$\inf_{y \in K} \left(g(y) + (L \triangleright f)(y) \right) = -\min_{v \in K} \left(g^{*v}(v) + (L \triangleright f)^*(v) \right)$$

$$= -\min_{v \in K} \left(g^{*v}(v) + f^*(L^* v) \right)$$

using this with (i) we arrive at the conclusion.

∎

Part 1

Friday, August 26, 2016 5:16 PM

2017-03-31 10:23 AM

Proposition 16.9:

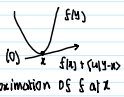
[[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \mathcal{H}$]]

$$x \in \operatorname{dom} f \Rightarrow \begin{cases} f^{**}(x) = f(x) \\ \partial f^{**}(x) = \partial f(x) \end{cases}$$

Proof:

take $u \in \partial f(x)$

$$\overset{\text{def}}{\Leftrightarrow} \forall_{y \in \mathcal{H}} \quad \underbrace{\langle y-x | u \rangle + f(x)}_{\text{so this first order approximation of } f \text{ at } x \text{ acts as a continuous affine minorant of } f} \leq f(y) \qquad (0)$$



/* recall:

Proposition 11.39. *

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

- f: has a continuous affine minorant $\Rightarrow f^{**} = \bar{f}$
  $\operatorname{dom} \bar{f} \neq \emptyset$ (from Proposition 11.27(i))
- f: doesn't have a continuous affine minorant $\Rightarrow f^{**} = -\infty$

*/

(* This theorem can have a practical significance, when the first inequality holds, even if the function f is highly nonconvex, its convex envelope can be found by bi-conjugating the original problem, as conjugating may be a relatively easy optimization problem, that is very good practically *)

$$f^{**} = \check{f}$$

/* recall:

---

*Corollary 9.10.

[$f: \mathcal{H} \to [-\infty, +\infty]$, convex] $\Rightarrow \bar{f} = \check{f}$

/* $\check{f} = (g \in \Gamma(\mathcal{H}) \mid g \leq f)$ // $\sup$ [2]

lower semicontinuous convex envelope

$\bar{f} = (g$: lower semicontinuous $\mid g \leq f)$ // $\sup$ [1]

lower semicontinuous envelope

*/

* Lemma 1.31:

[X: hausdorff space.

$f: X \to [-\infty, +\infty]$] $\Rightarrow$

(i) $\bar{f}$: largest lower semicontinuous function majorized by f

(ii) $\operatorname{epi} \bar{f}$: closed

(iii) $\operatorname{dom} \bar{f} \subseteq \operatorname{dom} f \subseteq \overline{\operatorname{dom} f}$, (iv) $\forall_{x \in X}$ $\bar{f}(x) = \lim_{y \to x} f(y)$

(v) $x \in X \Rightarrow$ (f: lower semicontinuous at $x \Leftrightarrow \bar{f}(x) = f(x)$)

(vi) $\operatorname{epi} \bar{f} = \overline{\operatorname{epi} f}$.

and finally recall that:

Proposition 16.3. [$f \in \mathcal{H} \to]-\infty, +\infty]$, proper, $x \in \operatorname{dom} f$] $\Rightarrow$

(i) $\operatorname{dom} \partial f \subseteq \operatorname{dom} f$

(ii) $\partial f(x) = \bigcap_{y \in \operatorname{dom} f} \{u \in \mathcal{H} \mid \langle y-x | u \rangle \leq f(y) - f(x)\}$

(iii) $\partial f(x)$: closed, convex

(iv) $x \in \operatorname{dom} \partial f \Rightarrow f$: lower semicontinuous at x */

$f = \bar{f}$ /* i.e. f will be its own lower-semicontinuous envelope at x */

at x: $f(x) = \bar{f}(x)$

$$f^{**}(x) = \check{f}(x) = \bar{f}(x) = \lim_{y \to x} f(y) = f(x) \qquad (0)$$

thus from (0) and (1) we have:

$$\forall_{y \in \mathcal{H}} \quad \langle y-x | u \rangle + f^{**}(x) \leq f^{**}(y)$$

$$\Leftrightarrow \quad u \in \partial f^{**}(x) \qquad (2)$$

$$\therefore \quad \partial f(x) \subseteq \partial f^{**}(x)$$

Now consider the other side: take $u \in \partial f^{**}(x) \Leftrightarrow \forall_{y \in \mathcal{H}} \langle y-x | u \rangle + f^{**}(x) \leq f^{**}(y)$ /* recall for any function, f its bi-conjugate is a lower bound for the function:

So, $f^{**} \leq f$ /* Proposition 16.9 */ */

$$\Rightarrow \forall_{y \in \mathcal{H}} \quad \langle y-x | u \rangle + f^{**}(x) \leq f(y)$$

now as f: proper, convex, $x \in \operatorname{dom} \partial f$ // By given

$\Updownarrow$

f: lower semicontinuous at x

$\Leftrightarrow$ f: proper, convex, lower-semicontinuous at x

$\Leftrightarrow$ $f^{**}(x) = f(x)$ // from Fenchel-Moreau theorem $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^{**} = f$

[3]

$$\forall_{y \in \mathcal{H}} \quad \langle y-x | u \rangle + f(x) \leq f(y)$$

$$\Leftrightarrow \quad u \in \partial f(x)$$

So, we have proven, $\partial f^{**}(x) \subseteq \partial f(x)$ (4)

from (0), (2), (3), (4):

$$f^{**}(x) = f(x)$$

$$\partial f^{**}(x) = \partial f(x) \qquad \blacksquare$$

2017-04-04 9:05 AM

Proposition 16.21: **

[$f \in \Gamma_0(\mathcal{H})$] $\operatorname{int} \operatorname{dom} f = \operatorname{cont} f \subseteq \operatorname{dom} \partial f \subseteq \operatorname{dom} f$

Proof:

/* Recall: as $f \in \Gamma_0(\mathcal{H}) \Rightarrow$

*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] $\Rightarrow$ $\boxed{\operatorname{cont} f = \operatorname{int} \operatorname{dom} f}$ /* cont f: domain of continuity of a function f */

☑

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

/* now

Proposition 16.3 [$f \in \mathcal{H} \to]-\infty, +\infty]$, proper, $x \in \operatorname{dom} f$] $\Rightarrow$ ☑ ☑ ☑

(i) $\operatorname{dom} \partial f \subseteq \operatorname{dom} f$ ☑ $\qquad \operatorname{dom} \partial f \subseteq \operatorname{dom} f$ (2)

(ii) $\partial f(x) = \bigcap_{y \in \operatorname{dom} f} \{u \in \mathcal{H} \mid \langle y-x | u \rangle \leq f(y) - f(x)\}$

(iii) $\partial f(x)$: closed, convex

(iv) $x \in \operatorname{dom} \partial f \Rightarrow f$: lower semicontinuous at x

Proposition 16.14. [$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \operatorname{dom} f$] $\Rightarrow$

(i) [$\operatorname{int} \operatorname{dom} f \neq \emptyset$, $x \in \operatorname{bdry} \operatorname{dom} f$] $\Rightarrow \partial f(x)$: empty or unbounded

(ii) $x \in \operatorname{cont} f \Rightarrow \partial f(x)$: nonempty, weakly compact

(iii) $x \in \operatorname{cont} f \Rightarrow \exists_{\delta \in \mathbb{R}_{++}}$ $\partial f(B(x, \delta))$: bounded

(iv) $\operatorname{cont} f \neq \emptyset \Rightarrow \operatorname{int} \operatorname{dom} f \subseteq \operatorname{dom} \partial f$ ☑

*/

(ii) $x \in \operatorname{cont} f \Rightarrow \partial f(x)$: nonempty, weakly compact

(iii) $x \in \operatorname{cont} f \Rightarrow \exists_{\delta \in \mathbb{R}_{++}}$ $\partial f(B(x;\delta))$: bounded

(iv) $\underbrace{\operatorname{cont} f \neq \emptyset}$ $\Rightarrow$ $\operatorname{int} \operatorname{dom} f \subseteq \operatorname{dom} \partial f$ ☑

*/ assume $\operatorname{int} \operatorname{dom} f = \operatorname{cont} f \neq \emptyset$ /*otherwise, $\operatorname{int} \operatorname{dom} f \subseteq \operatorname{dom} \partial f$ trivially */
(from (1))

then $\operatorname{int} \operatorname{dom} f = \operatorname{cont} f \subseteq \operatorname{dom} \partial f$ (3)

from (2), (3):

$$\operatorname{int} \operatorname{dom} f = \operatorname{cont} f \subseteq \operatorname{dom} \partial f \subseteq \operatorname{dom} f$$

■

8/26/2016 5:55 PM

Proposition 16.28.

$f \in \Gamma_0(\mathcal{H}) \Rightarrow \operatorname{gra}(f + \iota_{\operatorname{dom} \partial f})$: dense subset of $\operatorname{gra} f$

$$\leftrightarrow \forall_{x \in \operatorname{dom} f} \exists_{(x_n)_{n \in \mathbb{N}} \subseteq \operatorname{dom} \partial f} (x_n \to x \wedge f(x_n) \to f(x))$$

Proof:

$x \in \operatorname{dom} f$, $\varepsilon \in \mathbb{R}_{++}$

$(p, \pi) = P_{\operatorname{epi} f}(x, f(x) - \varepsilon)$ . . . (i)

/* proposition 3.18.
[$f \in \Gamma_0(\mathcal{H})$, $x \in \operatorname{dom} f$, $\eta \in]-\infty, f(x)[$]

$$(p, \pi) = P_{\operatorname{epi} f}(x, \eta) \leftrightarrow \begin{cases} \eta < f(p) = \pi \\ \forall_{y \in \operatorname{dom} f} \langle y - p \mid x - p \rangle \leq (f(y) - f(p))(f(p) - \eta) \end{cases}$$

*/

using this: (i) is equivalent to:

$f(x) - \varepsilon < f(p) = \pi \leftrightarrow f(x) - f(p) < \varepsilon$ . . . (o)

$\forall_{y \in \operatorname{dom} f} \langle y - p \mid x - p \rangle \leq (f(y) - f(p))(f(p) - f(x) + \eta)$

/* recall: $u \in \partial f(x) \overset{\text{def}}{\leftrightarrow} \forall_{y \in \operatorname{dom} f} f(y) \geq f(x) + \langle u \mid y - x \rangle$ */ . . . (ii)

$$\leftrightarrow \forall_{y \in \operatorname{dom} f} \quad (f(y) - \cancel{f(p)})\cancel{(f(p) - f(x) + \eta)} \geq \langle y - p \mid \cancel{x - p} \rangle \frac{x - p}{f(p) - f(x) + \eta} \rangle + f(p)$$

$$\leftrightarrow \forall_{y \in \operatorname{dom} f} \quad f(y) \geq f(p) + \left\langle \frac{x - p}{f(p) - f(x) + \eta} \,\middle|\, y - p \right\rangle \quad \ldots (\text{ii})$$

from (ii) and (iii):

$$\underbrace{\frac{x - p}{f(p) - f(x) + \eta}}_{\text{finite}} \in \partial f(p) \quad // \operatorname{dom} A = \{x : Ax \neq \emptyset\}$$

$\Rightarrow$ $p \in \operatorname{dom} \partial f$ as $\partial f(p) \neq \emptyset$ . . . (iv)

in (iii) set $y := x$ we have:

$$f(x) \geq f(p) + \frac{1}{f(p) - f(x) + \eta} \|x - p\|^2$$

$$\leftrightarrow \quad f(x) - f(p) \geq \frac{1}{-(f(x) - f(p) - \eta)} \|x - p\|^2$$

$$\leftrightarrow \quad -\|f(x) - f(p)\|^2 + \eta(f(x) - f(p)) \geq \|x - p\|^2$$

$$\leftrightarrow \quad \|(x - p)\|^2 + \|f(x) - f(p)\|^2 \leq \eta(f(x) - f(p))$$

$$\leftrightarrow \quad \|(x, f(x)) - (p, f(p))\|^2 \leq \varepsilon \underbrace{(f(x) - f(p))}_{< \varepsilon \ /* \text{from (o)} */} < \varepsilon^2 \quad \ldots (\text{v})$$

So, essentially we have proven that

$$\forall_{x \in \operatorname{dom} f} \forall_{\varepsilon \in \mathbb{R}_{++}} \quad (p, \pi) = P_{\operatorname{epi} f}(x, f(x) - \varepsilon) \Rightarrow p \in \operatorname{dom} \partial f \wedge \|(x, f(x)) - (p, f(p))\|^2 < \varepsilon$$

for $\varepsilon = \frac{1}{n}$ we can construct sequence $\underbrace{(x_n)_{n \in \mathbb{N}}}_{\in \operatorname{dom} f}$: $(x_n, \pi_n) = P_{\operatorname{epi} f}(x, f(x) - \frac{1}{n})$ which will

satisfy: $x_n \in \operatorname{dom} \partial f$ and $\forall_n \|(x, f(x)) - (x_n, f(x_n))\|^2 < \frac{1}{n^2}$ /* by taking $n \to \infty$, $x_n \to x$, $f(x_n) \to f(x)$ */

$\Rightarrow$ $x_n \to x$, $f(x_n) \to f(x)$

■

Proposition 16.13.

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

$$\left.\begin{aligned} u \in \partial f(x) &\Leftrightarrow (u, -1) \in N_{\operatorname{epi} f}(x, f(x)) \\ &\Leftrightarrow f(x) + f^*(u) = \langle x \mid u \rangle \end{aligned}\right] \Rightarrow x \in \partial f^*(u)$$

2017-03-27 8:55 AM

Proof:

f : convex, proper

$\Rightarrow$ $\operatorname{epi} f$: convex, nonempty

now: $(u, -1) \in N_{\operatorname{epi} f}(x, f(x))$ /* recall that: $N_C(x) = \begin{cases} (C - x)^\ominus = \{u \in \mathcal{H} \mid \sup \langle C - x \mid u \rangle \leq 0\}, & \text{if } x \in C \\ \emptyset, & \text{else} \end{cases}$ */

$$\therefore N_{\operatorname{epi} f}(x, f(x)) = \begin{cases} (\operatorname{epi} f - (x, f(x)))^\ominus = \{(\tilde{u}, \tilde{v}) \in \mathcal{H} \times \mathbb{R} \mid \sup \langle \operatorname{epi} f - (x, f(x)) \mid (\tilde{u}, \tilde{v}) \rangle \leq 0\}, & \text{if } (x, f(x)) \in \operatorname{epi} f \\ \emptyset, & \text{else} \end{cases}$$

this cannot be the case as $(u, -1) \in N_{\operatorname{epi} f}(x, f(x))$

So, $(u, -1) \in N_{\operatorname{epi} f}(x, f(x))$

$$\Leftrightarrow \sup \langle \operatorname{epi} f - (x, f(x)) \mid (u, -1) \rangle \leq 0, \quad (\overbrace{x}^{\in \mathcal{H}}, \overbrace{f(x)}^{\in \mathbb{R}}) \in \operatorname{epi} f$$

$$\Updownarrow$$

$$f(x) \leq f(x) \in \mathbb{R}$$

$$\Updownarrow$$

$$x \in \operatorname{dom} f, \text{ as } f : \mathcal{H} \to]-\infty, \infty]$$

$$\Leftrightarrow \forall_{(y, \eta) \in \operatorname{epi} f} \underbrace{\langle (y, \eta) - (x, f(x)) \mid (u, -1) \rangle \leq 0}, \quad x \in \operatorname{dom} f$$

$x \in \operatorname{dom} f$, as $f : \mathcal{H} \to]-\infty, \infty]$

$$\Leftrightarrow \forall_{(y,\eta)\in \operatorname{epi} f} \underbrace{\langle (y,\eta)-(x,f(x)) \,|\, (u,-1)\rangle \leq 0}, \quad x \in \operatorname{dom} f$$

$$\left\langle \begin{bmatrix} y-x \\ \eta - f(x) \end{bmatrix} \middle| \begin{bmatrix} u \\ -1 \end{bmatrix} \right\rangle = \langle y-x \,|\, u\rangle + \underbrace{\langle \eta - f(x) \,|\, -1\rangle}_{\text{number number}} = \langle y-x \,|\, u\rangle - \eta + f(x)$$

$$(-1)(\eta - f(x)) = -\eta + f(x)$$

$$\Leftrightarrow \forall_{(y,\eta)\in \operatorname{epi} f} \langle y-x \,|\, u\rangle + f(x) \leq \eta, \quad x \in \operatorname{dom} f$$

$f(y) \leq \eta \in \mathbb{R}$, $y \in \operatorname{dom} f$

$$\Leftrightarrow \forall_{y \in \operatorname{dom} f} \langle y-x \,|\, u\rangle + f(x) \leq f(y)$$

/* miniproof: $\forall_{(y,\eta)\in \operatorname{epi} f} \langle y-x \,|\, u\rangle + f(x) \leq \eta \Leftrightarrow \forall_{y \in \operatorname{dom} f} \langle y-x \,|\, u\rangle + f(x) \leq f(y)$

($\Rightarrow$): set $\eta = f(y)$ then $(y, f(y)) \in \operatorname{epi} f$ which gives $\forall_{y \in \operatorname{dom} f} \langle y-x \,|\, u\rangle + f(x) \leq f(y)$

($\Leftarrow$): have: $\forall_{y \in \operatorname{dom} f} \langle y-x \,|\, u\rangle + f(x) \leq f(y)$, $\forall_{(y,\eta)} (y,\eta) \in \operatorname{epi} f \Leftrightarrow f(y) \leq \eta$, $\forall_{(y,\eta)\in \operatorname{epi} f} \langle y-x \,|\, u\rangle + f(x) \leq f(y) \leq \eta$

$\therefore \forall_{(y,\eta)\in \operatorname{epi} f} \langle y-x \,|\, u\rangle + f(x) \leq \eta$ */

$$\Leftrightarrow u \in \partial f(x)$$

/* Proposition 16.9.

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

$x \in \mathcal{H}$; $u \in \mathcal{H}$]

$u \in \partial f(x) \Leftrightarrow f(x) + f^*(u) = \langle x \,|\, u\rangle \Rightarrow x \in \partial f^*(u)$ */

the rest of the proof follows from this ∎

2017-03-27 4:03 PM

Proposition 16.9.

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

$u \in \partial f(x) \Leftrightarrow f(x) + f^*(u) = \langle x \,|\, u\rangle \Rightarrow x \in \partial f^*(u)$

Proof:

$$u \in \partial f(x) \overset{\text{def}}{\leftrightarrow} \forall_{y \in \operatorname{dom} f} \overbrace{\langle y-x \,|\, u\rangle}^{\langle y|u\rangle - \langle x|u\rangle} + f(x) \leq f(y)$$

$$\leftrightarrow \forall_{y \in \operatorname{dom} f} \langle y|u\rangle - f(y) \leq \langle x|u\rangle - f(x)$$

$$\leftrightarrow \underbrace{\sup_{y \in \operatorname{dom} f} \langle y|u\rangle - f(y)}_{f^*(u)} \leq \langle x|u\rangle - f(x)$$

$$\leftrightarrow f^*(u) \leq \langle x|u\rangle - f(x)$$

$$\leftrightarrow f(x) + f^*(u) \leq \langle x|u\rangle \quad \ldots (1)$$

But from Fenchel-Moreau 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! */

$$\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}} \quad f(x) + f^*(u) \geq \langle x|u\rangle$$

$$\leftrightarrow f(x) + f^*(u) = \langle x|u\rangle$$ /* $\because$ Fenchel Moreau is an independent piece of result */

(2)

* Proposition 13.14. ★ ★ ★

[[$f, g : \mathcal{H} \to [-\infty, +\infty]$]] /* these properties are actually quite useful */

(i) $f^{**} \leq f$ /* Biconjugate is always a lower bound */

(ii) $f \leq g \Rightarrow (f^* \geq g^*,\ f^{**} \leq g^{**})$ /* Conjugation flips the relation between two functions */

(iii) $f^{***} = f^*$ /* Conjugating thrice and once is the same */

(iv) $(\breve{f})^* = f^*$ /* $\breve{f} = \sup\{g \in \Gamma(\mathcal{H}) : g \leq f\}$: lower semicontinuous convex envelope of f */

setting $f := f^*$

$$\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}} \quad \langle x|u\rangle \leq f^*(x) + f^{**}(u)$$

interchanging x, u

$$\Leftrightarrow \forall_{u \in \mathcal{H}} \forall_{x \in \mathcal{H}} \quad \underbrace{\langle u|x\rangle \leq f^*(u) + f^{**}(x) \leq f^*(u) + f(x) = \langle u|x\rangle}_{\text{collapse}}$$ /* using (2) which is equivalent to $u \in \partial f(x)$ */

$$\therefore f^*(u) + f^{**}(x) = \langle u|x\rangle$$ /* use (2) again */

$$\Leftrightarrow x \in \partial f^*(u)$$

∎

Part 2

April 6, 2017 10:23 PM

Proposition 16.34.

[[$f \in \Gamma_0(\mathcal{H})$; $x, p \in \mathcal{H}$]]

$p = \mathrm{Prox}_f x \Leftrightarrow x - p \in \partial f(p)$

i.e., $\mathrm{Prox}_f = (\mathrm{Id} + \partial f)^{-1}$

Proof: /* 2017-04-06 10:51 PM */

Proposition 8.26. (Defining properties of Prox. operator) ★ /* This shows the similarities between projection and proximal operator */

[$f \in \Gamma_0(\mathcal{H})$; $x, p \in \mathcal{H}$]

$p = \mathrm{Prox}_f x \Leftrightarrow \forall_{y \in \mathcal{H}} \ \langle y - p | x - p \rangle + f(p) \leq f(y)$

Definition of subdifferential $\quad \forall_{y \in \mathcal{H}} \ \langle y - p | \overbrace{x - p}^{u} \rangle + f(p) \leq f(y) \Leftrightarrow (x - p) \in \partial f(p)$

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

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

now $x - p \in \partial f(p) \Leftrightarrow x \in \partial f(p) + p = (\mathrm{Id} + \partial f)(p)$

$\therefore \ (\mathrm{Id} + \partial f)^{-1}(x) \ni p \Leftrightarrow p = \mathrm{Prox}_f x \quad (1)$

So, $(\mathrm{Id} + \cdot)^{-1} = \mathrm{Prox}_f x \quad (2)$ $\quad$ But $\mathrm{Prox}_f(\cdot)$: is a singleton operator

so, $(\mathrm{Id} + \partial f)^{-1}$ " " " "

thus (1) becomes: $p = (\mathrm{Id} + \partial f)^{-1}(x)$ ∎

Proposition 16.32.

[$\mathcal{K}$: real Hilbert space ; $f \in \Gamma_0(\mathcal{H})$; $g \in \Gamma_0(\mathcal{K})$; $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$: $L(\mathrm{dom} f) \cap (\mathrm{dom} g) \neq \emptyset$; $(f + g \circ L)^* = f^* \boxminus (L^* \triangleright g^*)$]]

$\partial(f + g \circ L) = \partial f + L^* \circ \partial g \circ L$

2017-04-07 2:33 PM

Proof:

*

Proposition 16.5. [$\mathcal{K}$: real Hilbert space ; $f : \mathcal{H} \to]-\infty, +\infty]$; $g : \mathcal{K} \to]-\infty, +\infty]$, proper ; $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$; $\lambda \in \mathbb{R}_{++}$]

(i) $\partial(\lambda f) = \lambda \partial f$

(ii) $\mathrm{dom}\, g \cap L(\mathrm{dom} f) \neq \emptyset \Rightarrow \partial f + L^* \circ (\partial g) \circ L \subseteq \partial(f + g \circ L)$

so, we have: $\partial f + L^* \circ (\partial g) \circ L \subseteq \partial(f + g \circ L) \quad (-1)$

and we need to prove. $\partial(f + g \circ L) \subseteq \partial f + L^* \circ (\partial g) \circ L$

take $(x, u) \in \mathrm{gra}\, \partial(f + g \circ L)$ (0)

Proposition 16.9.
[$f : \mathcal{H} \to]-\infty, +\infty]$, proper ; $x \in \mathcal{H}$; $u \in \mathcal{H}$]
$u \in \partial f(x) \Leftrightarrow f(x) + f^*(u) = \langle x | u \rangle \Rightarrow x \in \partial f^*(u)$

$\Leftrightarrow \quad (f + g \circ L)(x) + (f + g \circ L)^*(u) = \langle x | u \rangle \quad (1)$

also, given: $(f + g \circ L)^* = f^* \boxminus (L^* \triangleright g^*)$

$\therefore \ (f + g \circ L)^*(u) = (f^* \boxminus (L^* \triangleright g^*))(u)$

$= \min_{\tilde{y} \in \mathcal{H}} \left((L^* \triangleright g^*)(\tilde{y}) + f^*(u - \tilde{y}) \right)$ /* say the optimizer is y^* */

$= \underbrace{(L^* \triangleright g^*)(y^*)}_{\parallel} + f^*(u - y^*)$

/* $\left(\begin{array}{l} \min \ g^*(\tilde{x}) \\ \text{s.t. } L^* \tilde{x} = y^* \end{array} \right)$ say the minimizer is v, then $y^* = L^* v$ */

$= g^*(v) + f^*(u - y^*) = g^*(v) + f^*(u - L^* v)$ $\quad (L^* v = y^*)$

So, we have:

$(f + g \circ L)^*(u) = g^*(v) + f^*(u - L^* v) \quad (2)$

$\underbrace{(f + g \circ L)(x)}_{f(x) + g(Lx)} + g^*(v) + f^*(u - L^* v) = \underbrace{\langle x | u \rangle}_{\langle x | u - L^* v + L^* v \rangle = \langle x | u - L^* v \rangle + \langle x | L^* v \rangle}$

$\Leftrightarrow \ \underbrace{\left(f(x) + f^*(u - L^* v) - \langle x | u - L^* v \rangle \right)}_{\geq 0} + \underbrace{\left(g(Lx) + g^*(v) - \langle x | L^* v \rangle \right)}_{\geq 0} = 0$ /* now using 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! */ */

$$\Leftrightarrow \underbrace{\left(f(x)+f^*(u-L^*v)-\langle x|u-L^*v\rangle\right)}_{\geq 0}+\underbrace{\left(g(Lx)+g^*(v)-\langle x|L^*v\rangle\right)}_{\geq 0}=0$$

/* now using 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! */ */

$$\forall_{x\in\mathcal{H}}\ \forall_{u\in\mathcal{H}}\quad f(x)+f^*(u)\geq\langle x|u\rangle$$

So both of these terms must be seperately zero :

/* if sum of two nonegative numbers is zero, then they seperately must be zero */

$$f(x)+f^*(u-L^*v)=\langle x|u-L^*v\rangle \quad \text{and}$$

$$g(Lx)+g^*(v)=\langle x|L^*v\rangle \overset{\text{def}}{=} \langle v|Lx\rangle$$

using

Proposition 16.9
[$f:\mathcal{H}\to]-\infty,+\infty]$, proper ; $x\in\mathcal{H}$; $u\in\mathcal{H}$]
$u\in\partial f(x)\Leftrightarrow f(x)+f^*(u)=\langle x|u\rangle \Rightarrow x\in\partial f^*(u)$

$$\left.\begin{aligned} u-L^*v &\in \partial f(x)\\ v &\in \partial g(Lx)\\ \Rightarrow L^*v &\in L^*\partial g(Lx)\end{aligned}\right\}\Rightarrow u\in L^*v+\partial f(x)\subseteq L^*g(Lx)+\partial f(x)=\left(L^*\circ g\circ L+\partial f\right)(x)$$

$$\Updownarrow$$

$$(x,u)\in \operatorname{gra}\left(\partial f+L^*\circ(\partial g)\circ L\right) \qquad (3)$$

So, from (0) and (3) we have:

$$\partial(f+g\circ L)\subseteq \operatorname{gra}\left(\partial f+L^*\circ(\partial g)\circ L\right) \qquad (4)$$

So, from (-1), (4) we have:

$$\partial(f+g\circ L)=\partial f+L^*\circ(\partial g)\circ L$$

∎

Proposition 16.46.

[K: real Hilbert space; $F\in\Gamma_0(\mathcal{H}\times K)$; $f:\mathcal{H}\to[-\infty,+\infty]: x\mapsto \inf F(x,K)$;
f: proper; $f(x)=F(x,y)$; $u\in\mathcal{H}$] $u\in\partial f(x)\Leftrightarrow (u,0)\in\partial F(x,y)$

$$\Updownarrow$$

$$y=\underset{\tilde{y}\in K}{\operatorname{argmin}}\ F(x,\tilde{y})$$

2017-04-07 4:23 PM

Proof:

*Proposition 13.28: **
[K: real Hilbert space
$F:\mathcal{H}\times K\to]-\infty,+\infty]$, proper
$f:\mathcal{H}\to[-\infty,+\infty]: x\mapsto \inf F(x,K)$]
$f^*=F^*(\cdot,0)$

we have: $f^*(u)=F^*(u,0)$

Proposition 16.9. (Very important theorem, subdifferential membership can be tested using it */
[$f:\mathcal{H}\to]-\infty,+\infty]$, proper ; $x\in\mathcal{H}$; $u\in\mathcal{H}$]
$u\in\partial f(x)\Leftrightarrow f(x)+f^*(u)=\langle x|u\rangle \Rightarrow x\in\partial f^*(u)$

$$u\in\partial f(x)\Leftrightarrow \underbrace{f(x)}_{F(x,y)}+\underbrace{f^*(u)}_{F^*(u,0)}=\langle x|u\rangle$$

$$\Leftrightarrow F^*(u,0)=\langle x|u\rangle-F(x,y)=\langle (x,y)|(u,0)\rangle-F(x,y)$$

$$\Leftrightarrow F(x,y)+F^*(u,0)=\langle (x,y)|(u,0)\rangle$$

$$\Leftrightarrow (u,0)\in\partial F(x,y)$$ ∎

Proposition 16.47.

[$f \in \Gamma_0(\mathcal{H})$; $\mathcal{K}$: real Hilbert space ; $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$;
$y \in \mathrm{dom}(L \triangleright f)$; $x \in \mathcal{H}$; $Lx = y$] $\Rightarrow$

(i) $(L \triangleright f)(y) = f(x) \Rightarrow \partial(L \triangleright f)(y) = (L^*)^{-1}(\partial f(x))$

(ii) $(L^*)^{-1}(\partial f(x)) \neq \emptyset \Rightarrow (L \triangleright f)(y) = f(x)$

<u>Proof</u>: $v \in \mathcal{K}$

suppose,

$f(x) + (L \triangleright f)^*(v) = \langle y | v \rangle$ (-1) then this is equivalent to saying:

$\leftrightarrow f(x) + (f^* \circ L^*)(v) = \langle Lx | v \rangle$ /* $\because Lx = y$: given, and

$\leftrightarrow f(x) + f^*(L^* v) = \langle Lx | v \rangle = \langle x | L^* v \rangle$ . . . (0)

/* using the definition of adjoint operator */

Proposition 15.11. (very important)
[$f, g : \mathcal{H} \to]-\infty, +\infty]$]
(i) $(f \square g)^* = f^* + g^*$
(ii) f, g: proper $\Rightarrow (f+g)^* \leq f^* \square g^*$
(iii) $\forall_{\gamma \in \mathbb{R}_{++}} (\gamma f)^* = f^* \circ \frac{1}{\gamma} \|\cdot\|^2$
(iv) $L \in \mathcal{B}(\mathcal{H}, \mathcal{K}) \Rightarrow (L \triangleright f)^* = f^* \circ L^*$
(v) $L \in \mathcal{B}(\mathcal{K}, \mathcal{H}) \Rightarrow (f \circ L)^* \leq L^* \triangleright f^*$ */

/* using

Proposition 16.9 (Very important theorem, unconditional monotonicity (can be proved using 3.4))
[$f : \mathcal{H} \to]-\infty, +\infty]$, proper ; $x \in \mathcal{H}$; $u \in \mathcal{H}$]
$u \in \partial f(x) \Leftrightarrow \langle f(x) + f^*(u) - \langle x | u \rangle \rangle \leq 0 \in \mathbb{R}_{\mathbb{R}}^{\mathbb{R}}(u)$ */

$\leftrightarrow L^* v \in \partial f(x)$ (1)

we are going to use this later

Proposition 13.15. (Fenchel-Young inequality)
[$f : \mathcal{H} \to]-\infty, +\infty]$, proper] /* wow! see how general the function is, infact any suitable function would satisfy this! */
$\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}} \quad f(x) + f^*(u) \geq \langle x | u \rangle$

(i)

$v \in \partial(L \triangleright f)(y) \leftrightarrow (L \triangleright f)(y) + \underbrace{(L \triangleright f)^*}_{(f^* \circ L^*)}(v) = \langle y | v \rangle$ /* given $Lx = y$, and $(L \triangleright f)(y) = f(x)$ */

$\leftrightarrow f(x) + f^*(L^* v) = \langle Lx | v \rangle$ /* this is (0), so use (0) */

$\leftrightarrow L^* v \in \partial f(x) \leftrightarrow \exists_{\eta \in \partial f(x)} \; L^* v = \eta$

$\leftrightarrow \exists_{\eta \in \partial f(x)} \; v = (L^*)^{-1} \eta$ /* Provided that L^{*-1} exists */

$\leftrightarrow \boxed{v = (L^*)^{-1} \eta \in (L^*)^{-1} \partial f(x)}$ $\therefore \partial(L \triangleright f)(y) = (L^*)^{-1}(\partial f(x))$

/* in a manner similar to this */

(ii) take $v \in (L^*)^{-1}(\partial f(x)) \leftrightarrow L^* v \in \partial f(x) \leftrightarrow f(x) + (L \triangleright f)^*(v) = \langle y | v \rangle$ (from (1)) . . . (4)

/* using

Proposition 13.15. (Fenchel-Young inequality)
[$f : \mathcal{H} \to]-\infty, +\infty]$, proper] /* wow! see how general the function is, infact any suitable function would satisfy this! */
$\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}} \quad f(x) + f^*(u) \geq \langle x | u \rangle$
*

$\langle y | v \rangle \leq \underbrace{(L \triangleright f)(y)}_{\parallel} + (L \triangleright f)^*(v)$

/* $\left(\begin{array}{l} \min_{\tilde{x}} \; f(\tilde{x}) \\ \text{s.t.} \; L\tilde{x} = y \end{array} \right) \leq f(x)$, as $Lx = y$, x is a feasible point for the optimization problem */

$\leq f(x) + (L \triangleright f)^*(v) = \langle y | v \rangle$ /* from (4) */

So, the inequalities collapse:

$\therefore \langle y | v \rangle = \underbrace{(L \triangleright f)(y) + (L \triangleright f)^*(v) = f(x)} + (L \triangleright f)^*(v) = \langle y | v \rangle$

$\therefore (L \triangleright f)(y) = f(x)$ ∎

/* info: $[\![F: \mathcal{H}\times\mathcal{H} \to [-\infty,+\infty]]\!]$ F: autoconjugate $\overset{\text{def}}{\Leftrightarrow} F^* = F^T$
$F^T(u,x) = F(x,u)$ $(\cdot)^T$: transposition operator */

2017-04-12 9:39 AM

Proposition 16.52. (Subdifferential of autoconjugate function)

$[\![F: \in \Gamma_0(\mathcal{H}\times\mathcal{H})$, autoconjugate ; $(x,u)\in\mathcal{H}\times\mathcal{H}]\!]$

$F(x,u) = \langle x|u\rangle \leftrightarrow F^*(u,x) = \langle x|u\rangle \leftrightarrow (u,x)\in\partial F(x,u)$

$\leftrightarrow (x,u) = \mathrm{Prox}_F(x+u, x+u)$

Proof: (i) $\Leftrightarrow$ (ii):

(i) $\Leftrightarrow$ $F(x,u) = \langle x|u\rangle$ /* By definition: $F^T(u,x) = F(x,u)$ */

$\leftrightarrow$ $F^T(u,x) = F(x,u) = \langle x|u\rangle$ /* F: autoconjugate $\overset{\text{def}}{\Leftrightarrow}$ $F^T(u,x) = F^*(u,x)$ */

$\leftrightarrow$ $F^*(u,x) = \langle x|u\rangle = F^T(u,x)$

$\Leftrightarrow$ (ii)

(i) $\Leftrightarrow$ (iii)

$F(x,u) = \langle x|u\rangle$

$\leftrightarrow 2F(x,u) = 2\langle x|u\rangle \leftrightarrow F(x,u) + \underbrace{F(x,u)}_{\overset{\shortparallel}{F^T(u,x) = F^*(u,x)}} = \langle x|u\rangle + \overbrace{\langle x|u\rangle}^{\langle u|x\rangle}$ /* F: autoconjugate */

$\leftrightarrow F(x,u) + F^*(u,x) = \langle (x,u) | (u,x)\rangle$ /*

$\leftrightarrow (u,x)\in\partial F(x,u) \Leftrightarrow$ (iii) */

(iii) $\Leftrightarrow$ (iv)

(iii): $(u,x)\in\partial F(x,u)$ /* recall: $f\in\Gamma_0(\mathcal{H}) \Rightarrow \mathrm{Prox}_f = (\mathrm{Id}+\partial f)^{-1}$ */

$\leftrightarrow (u,x) + (x,u) \in \partial F(x,u) + (x,u) = (\partial F + \mathrm{Id})(x,u)$

$\leftrightarrow \underbrace{(\partial F + \mathrm{Id})^{-1}}_{\mathrm{Prox}_F}\Big(\underbrace{(u,x)+(x,u)}_{\begin{bmatrix} u \\ x\end{bmatrix} + \begin{bmatrix} x \\ u\end{bmatrix} = \begin{bmatrix} u+x \\ x+u\end{bmatrix} = \begin{bmatrix} x+u \\ x+u\end{bmatrix} = (x+u, x+u)}\Big) \ni (x,u)$

$\leftrightarrow \mathrm{Prox}_F(x+u, x+u) \ni (x,u)$ /* But: $\mathrm{Prox}_{(\cdot)}(\cdot)$: is a single valued operator */

$\leftrightarrow \mathrm{Prox}_F(x+u, x+u) = (x,u) \Leftrightarrow$ (iv)

$\blacksquare$

2017-04-16 12:36 PM

Proposition 16.53.

$[\![L: \mathcal{H}\times\mathcal{H} \to \mathcal{H}\times\mathcal{H} : (x,u)\mapsto(u,x)$; $F\in\Gamma_0(\mathcal{H}\times\mathcal{H})]\!]$

$\mathrm{Prox}_{F^{*T}} = (\mathrm{Id} - L\,\mathrm{Prox}_F\, L)$

Proof:

$L\begin{bmatrix} x \\ u\end{bmatrix} = \begin{bmatrix} u \\ x\end{bmatrix}$, now $\langle L(x,u) | (\tilde{x},\tilde{u})\rangle = \langle \begin{bmatrix} u \\ x\end{bmatrix} | \begin{bmatrix} \tilde{x} \\ \tilde{u}\end{bmatrix}\rangle$

$= \langle u|\tilde{x}\rangle + \langle x|\tilde{u}\rangle$

$\langle (x,u) | L(\tilde{x},\tilde{u})\rangle = \langle \begin{bmatrix} x \\ u\end{bmatrix} | \begin{bmatrix} \tilde{u} \\ \tilde{x}\end{bmatrix}\rangle = \langle x|\tilde{u}\rangle + \langle u|\tilde{x}\rangle$

$= \langle u|\tilde{x}\rangle + \langle x|\tilde{u}\rangle$

so, $\langle L(x,u) | (\tilde{x},\tilde{u})\rangle = \langle (x,u) | L(\tilde{x},\tilde{u})\rangle$ /* recall: $\langle Lx|y\rangle = \langle x|L^*y\rangle$ */

$\therefore L = L^*$

also, $L = L^{-1}$ as $L\left(L\begin{bmatrix} x \\ u\end{bmatrix}\right) = L\begin{bmatrix} u \\ x\end{bmatrix} = \begin{bmatrix} x \\ u\end{bmatrix}$

so, $L^* = L = L^{-1}$ $\therefore \mathrm{Id} = LL^{-1} = LL$ (0.1)

also, $L = L^{-1}$ as $L\left(L\begin{bmatrix} x \\ u \end{bmatrix}\right) = L\begin{bmatrix} u \\ x \end{bmatrix} = \begin{bmatrix} x \\ u \end{bmatrix}$

so, $L^{*} = L = L^{-1}$ (0) $\therefore Id = LL^{-1} = LL$ (0.1)

now:

$F^{*} \in \Gamma_0(\mathcal{H}\times\mathcal{H})$

Theorem 16.39. [K: real Hilbert space; $f \in \Gamma_0(\mathcal{H})$; $g \in \Gamma_0(K)$; $L \in \mathcal{B}(\mathcal{H}, K)$; one of the following holds: $0 = \mathcal{H}$

(i) $0 \in \operatorname{sri}(\operatorname{dom} g - L(\operatorname{dom} f)) = \operatorname{sri}(\mathcal{H}) = \mathcal{H}$ ✓

(ii) K: finite-dimensional; g: polyhedral; $\operatorname{dom} g \cap \operatorname{ri} L(\operatorname{dom} f) \neq \emptyset$

(iii) $\mathcal{H}, K$: finite-dimensional; f, g: polyhedral; $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$] $\Rightarrow \partial(f + g \circ L) = \partial f + L^{*} \circ \partial g \circ L$

so, $\partial(F^{*} \circ L) = \underbrace{L^{*}}_{=L} \circ (\partial F^{*}) \circ L = L \circ (\partial F^{*}) \circ L$ $\cdots (1)$

now, $F^{*T}(x,u) = (F^{*})^{T}(x,u) = F^{*}(u,x) = F^{*}\left(L\begin{bmatrix} x \\ u \end{bmatrix}\right)$

$= (F^{*} \circ L)(x,u)$

so, $F^{*T} = F^{*} \circ L$

$\Rightarrow \partial F^{*T} = \partial(F^{*} \circ L) = L \circ (\partial F^{*}) \circ L$

so, $(Id + \partial F^{*T}) = \underbrace{L \circ L}_{\text{// from (0.1)}} + \underbrace{L \circ (\partial F^{*}) \circ L}_{\text{// from (0)}} = L \circ Id \circ L + L \circ (\partial F^{*}) \circ L = L(Id + \partial F^{*})L$

now: recall $\operatorname{Prox}_{f}(\cdot) = (Id + \partial f)^{-1}$

so, $\operatorname{Prox}_{F^{*T}} = (Id + \partial F^{*T})^{-1} = (L \circ (Id + \partial F^{*}) \circ L)^{-1}$ /* all of L, $(Id + \partial F^{*})$ are matrices, where $(AB)^{-1} = B^{-1}A^{-1}$ */

$= \underbrace{L^{-1}}_{L} \circ \underbrace{(Id + \partial F^{*})^{-1}}_{\operatorname{Prox}_{F^{*}}} \circ \underbrace{L^{-1}}_{L \text{ // from (0)}}$

$= L \circ \operatorname{Prox}_{F^{*}} \circ L$

/* now using Moreau's decomposition:

$[f \in \Gamma_0(\mathcal{H})]$ $\operatorname{Prox}_{f} + \operatorname{Prox}_{f^{*}} = Id$ */

$= L \circ (Id - \operatorname{Prox}_{F}) \circ L = L \circ (\underbrace{Id \circ L}_{L} - \operatorname{Prox}_{F} \circ L) = \underbrace{(L \circ L)}_{Id \text{ // from (0.1)}} - L \circ \operatorname{Prox}_{F} \circ L$

$= Id - L \operatorname{Prox}_{F} L$

$\therefore \operatorname{Prox}_{F^{*T}} = Id - L \operatorname{Prox}_{F} L$ ∎

Part 1

March 23, 2017 2:13 PM

Proposition 17.3.

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

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

2017-04-20 9:21 PM

Proof:

($\Rightarrow$)

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

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

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

/* Proposition 17.2. (Properties of directional derivative)
[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \operatorname{dom} f$; $y \in \mathcal{H}$] $\Rightarrow$
(i) $\varphi: \mathbb{R}_{++} \to]-\infty, +\infty]: \alpha \mapsto (f(x+\alpha y) - f(x))/\alpha$: 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) $f'(x; y-x) + f(x) \leq f(y)$ // kind 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 */

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

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

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

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

$\tilde{y} := y - x$

$\Rightarrow \quad f'(x; y-x) \geq 0$

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

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

2017-04-21 7:44 PM

Corollary 17.6.

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

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

Proof:

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

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

$p = \operatorname{Prox}_f x = (2\lambda + \partial f)^{-1} x \quad // \because \operatorname{Prox}_f = (2\lambda + \partial f)^{-1}$

$\Leftrightarrow (2\lambda + \partial f) p \ni x$

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

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

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

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

2017-04-26 10:21 PM

Proof:

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

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

/* Proposition 17.2. (Properties of directional derivative)
[$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \operatorname{dom} f$; $y \in \mathcal{H}$] $\Rightarrow$
(i) $\varphi: \mathbb{R}_{++} \to]-\infty, +\infty]: \alpha \mapsto (f(x+\alpha y) - f(x))/\alpha$: increasing
(ii) $f'(x;y)$ exists in $[-\infty, +\infty]$ and $f'(x;y) = \inf_{\alpha \in \mathbb{R}_{++}} \dfrac{f(x+\alpha y) - f(x)}{\alpha}$ #/

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

$= \sup \left(\sup \left(f(x) - f(x+\alpha y) \right) + \langle y | u \rangle \right)$

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

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

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

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

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

Results 18.5

- $[\operatorname{lemm}\{\sup,\inf,\overline{\lim},\underline{\lim},\lim\}; f:H\times K\to\mathbb{R}]$ $\operatorname{lemm}\sup_{x\in X}\sup_{y\in Y} f(x,y) \geq \sup_{y\in Y}\operatorname{lemm}\sup_{x\in X} f(x,y)$

  if exists

  $\operatorname{lemm}\inf_{x\in X}\inf_{y\in Y} f(x,y) \leq \inf_{y\in Y}\operatorname{lemm}\inf_{x\in X} f(x,y)$; $\sup_{(x,y)\in X\times Y} f(x,y) = \sup_{x\in X}\sup_{y\in Y} f(x,y) = \sup_{y\in Y}\sup_{x\in X} f(x,y)$

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

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

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

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

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

new decision variable

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

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

Fenchel-Young inequality:

* Proposition 13.13. (Fenchel-Young inequality)

[$f: H\to]-\infty,+\infty]$, proper] /* wow! see how general the function is, infact any sensible function would satisfy this! */

$\forall_{x\in H}\ \forall_{u\in H}\quad f(x)+f^*(u)\geq\langle x|u\rangle \quad \ldots (2)$ */

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)

[$f: H\to]-\infty,+\infty]$, proper, convex ; $x\in\operatorname{cont} f\setminus\operatorname{Argmin} f$; $u\in P_{\partial f(x)}(0)$; $z=-\frac{u}{\|u\|}$]

z: unique minimizer of $f'(x;\cdot)$ over $B(0;1)$

2017-04-28 6:13 PM

Proof:

A recall:

Proposition 16.14. [$f: H\to]-\infty,+\infty]$, proper, convex ; $x\in\operatorname{dom} f$] $\Rightarrow$

(i) [$\operatorname{int}\operatorname{dom} f\neq\emptyset$, $x\in\operatorname{bary}\operatorname{dom} f$] $\Rightarrow \partial f(x)$: empty or unbounded

(ii) $x\in\operatorname{cont} f \Rightarrow \partial f(x)$: nonempty, weakly compact ✓ (1)

(iii) $x\in\operatorname{cont} f \Rightarrow \exists_{\delta\in\mathbb{R}_{++}}\ \partial f(B(x;\delta))$: bounded

(iv) $\operatorname{cont} f\neq\emptyset \Rightarrow \operatorname{int}\operatorname{dom} f\subseteq\operatorname{dom}\partial f$

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

Proposition 16.3. [$f\in H\to]-\infty,+\infty]$, proper, $x\in\operatorname{dom} f$] $\Rightarrow$

(i) $\operatorname{dom}\partial f\subseteq\operatorname{dom} f$

(ii) $\partial f(x)=\bigcap_{z\in\operatorname{dom} f}\{u\in H \mid \langle z-x|u\rangle\leq f(z)-f(x)\}$

(iii) $\partial f(x)$: closed, convex (2)

(iv) $x\in\operatorname{dom}\partial f \Rightarrow f$: lower semicontinuous at x

Theorem 16.2. (Fermat's rule) $x\notin\operatorname{Argmin} f \Rightarrow \partial f(x)\not\ni 0$ (3)

[$f: H\to]-\infty,+\infty]$, proper] $\operatorname{Argmin} f=\operatorname{zer}\partial f=\{x\in H \mid 0\in\partial f(x)\}$ */

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.

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

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

*/

$u \in \partial f(x)$

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

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

$\geq -\|y\|\,\|u\|$ /*

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

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

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

* Cauchy Schwarz inequality: $|\langle x\mid y\rangle| \leq \|x\|\,\|y\|$

$\max\{\langle x\mid y\rangle, -\langle x\mid y\rangle\}$

$\begin{cases} \text{Flavor 1: } \langle x\mid y\rangle \leq \|x\|\,\|y\| \\ \text{Flavor 2: } -\langle x\mid y\rangle \leq \|x\|\,\|y\| \Leftrightarrow \langle x\mid y\rangle \geq -\|x\|\,\|y\| \end{cases}$ */

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

2017-03-24 11:11 AM

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

$\nabla f(x) \in \partial f(x)$, 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 $y \overset{\text{def}}{\Leftrightarrow}$

$f'(x;y) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha y) - f(x)}{\alpha}$ { 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) $u \in \partial f(x) \Leftrightarrow \langle \cdot | u \rangle \leq f'(x; \cdot)$

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

$u \in \partial f(x) \Leftrightarrow \langle \cdot | u \rangle \leq f'(x; \cdot)$

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

$\Rightarrow$

$\langle u - \nabla f(x) | u \rangle \leq f'(x; u - \nabla f(x)) = \langle u - \nabla f(x) | \nabla f(x) \rangle$

$\therefore \langle u - \nabla f(x) | u \rangle \leq \langle u - \nabla f(x) | \nabla f(x) \rangle$

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

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

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

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

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 */

$\exists\, !\, \nabla f(x) \in \mathcal{H} \quad \forall\, y \in \mathcal{H} \quad f'(x;y) = \langle y | \nabla f(x) \rangle = \lim_{\alpha \to 0} \frac{f(x+\alpha y) - f(x)}{\alpha}$

unique

Gateaux gradient of f at x

(ii)

/*

Theorem 17.19. (Max formula)

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

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

*/

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

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

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

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

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

$\Rightarrow u = \nabla f(x)$ ∎

Part 2

May 1, 2017 1:30 PM

2017-05-01 6:54 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): $x \in \operatorname{cont} f = \operatorname{int} \operatorname{dom} f$ ☑

Part 1

March 28, 2017 8:42 AM

Proposition 18.1:

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

f: Frechet differentiable at $x \Leftrightarrow$

$\forall_{\varepsilon\in\mathbb{R}_{++}}\ \exists_{\eta\in\mathbb{R}_{++}}\ \forall_{y\in\mathcal{H}:\|y\|=1}\quad f(x+\eta y)+f(x-\eta y)-2f(x)\leq\eta\varepsilon$

Proof: Assume $\mathcal{H}\neq\{0\}$ else it is trivial

($\Rightarrow$)

f: Frechet differentiable at x

/*

* (Frechet differentiability of a function)

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

f: Frechet differentiable at $x \overset{\text{def}}{\Leftrightarrow} \lim_{0\neq y\to 0}\dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}=0$

Definition 1.17.0 · (Functional limit)

[X: Hausdorff space; $f: X\to[-\infty,+\infty]$; $x\in X$]

$\lim_{y\to x} f(y)=\ell \overset{\text{def}}{\Leftrightarrow} \forall_{\varepsilon>0}\ \exists_{\delta>0}\ \forall_{y\in X:\ d(x,y)\in]0,\delta[}\ |f(y)-\ell|<\varepsilon$

$\overset{\text{def}}{\Leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}:\text{net in } X,\ x_n\neq x,\ x_n\to x}\ \lim f(x_n)=\ell$

$\Leftrightarrow \forall_{\varepsilon>0}\ \exists_{V_\delta(x)}\ \forall_{y\in V_\delta(x)}\quad f(y)\in V_\varepsilon(\ell)$

*/

define $g_x(y)=\dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}=0$

$\therefore \lim_{0\neq y\to 0} g_x(y)=0 \Leftrightarrow \forall_{\tilde\varepsilon>0}\ \exists_{\tilde\delta>0}\ \forall_{y:\|y-x\|<\tilde\delta}\ |g_x(y)-0|<\tilde\varepsilon$

$\Leftrightarrow \forall_{\tilde\varepsilon>0}\ \exists_{\tilde\delta>0}\ \forall_{y:\|y-x\|<\tilde\delta}\ \left|\dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}-0\right| = \underbrace{\left|\dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}\right|}_{=\ \dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}}<\tilde\varepsilon$ /*because $f(\cdot)$: convex, $\forall_y\ f(x+y)-f(x)-\langle\nabla f(x)|y\rangle\geq 0$ */

$\Leftrightarrow \forall_{\tilde\varepsilon>0}\ \exists_{\tilde\delta>0}\ \forall_{y:\|y-x\|<\tilde\delta}\ \dfrac{f(x+y)-f(x)-\langle\nabla f(x)|y\rangle}{\|y\|}<\tilde\varepsilon$

(incomplete)

2017-03-28 3:27 PM

Proposition 18.6:

[[$f,g\in\Gamma_0(\mathcal{H})$; $x,y\in\mathcal{H}$; $x\in\operatorname{dom}\partial(f\square g)$; $(f\square g)(x)=f(y)+g(x-y)$; f: Gateaux differentiable at y]]

$\partial(f\square g)(x)=\{\nabla f(y)\}$

Proof: recall

/* Proposition 16.48:

[[$f,g\in\Gamma_0(\mathcal{H})$; $x\in\operatorname{dom}(f\square g)$; $y\in\mathcal{H}$]] $\Rightarrow$

(i) $(f\square g)(x)=f(y)+g(x-y) \Rightarrow \partial(f\square g)(x)=\partial f(y)\cap\partial g(x-y)$

↳ given in our case

(ii) $\partial f(y)\cap\partial g(x-y)\neq\emptyset \Rightarrow (f\square g)(x)=f(y)+g(x-y)$ */

** Proposition 17.26. [[$f:\mathcal{H}\to]-\infty,+\infty]$, proper, convex; $x\in\operatorname{dom} f$]]

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

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

$\partial(f\square g)(x)=\partial f(y)\cap\partial g(x-y)=\{\nabla f(y)\}\cap\partial g(x-y)\subseteq\{\nabla f(y)\}$ // $\because A\cap B\subseteq A$

/* given: f: Gateaux differentiable at y

$\therefore \partial f(y)=\{\nabla f(y)\}$ */

given, $\partial(f\square g)(x)\neq\emptyset$, $\{\nabla f(y)\}\subseteq\partial(f\square g)(x)$ // 왜?

$\therefore \partial(f\square g)(x)=\{\nabla f(y)\}$

2017-07-12 5:13 PM

Corollary 18.8.

[[$f,g\in\Gamma_0(\mathcal{H})$; f: real valued, supercoercive, frechet differentiable on $\mathcal{H}$]] $\Rightarrow f\square g$: frechet differentiable on $\mathcal{H}$.

Proof:

recall:

* Proposition 12.14. (Sufficient condition for exact infimal convolution)

[$f, g \in \Gamma_0(\mathcal{H})$

One of the following holds:

(i) f: supercoercive

(ii) f: coercive, bounded below]

$f \square g = f \boxed{\square} g \in \Gamma_0(\mathcal{H})$

/* f: supercoercive $\overset{\text{def}}{\Leftrightarrow} \lim_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} = +\infty$

f: coercive $\overset{\text{def}}{\Leftrightarrow} \lim_{\|x\| \to +\infty} f(x) = +\infty$ */

So, $f \square g = f \boxed{\square} g \in \Gamma_0(\mathcal{H})$, as f: supercoercive

take $x \in \mathcal{H}$

/*Proposition 12.6. (ii): [$f, g \in \mathcal{H} \to]-\infty, +\infty]$] $\operatorname{dom}(f \square g) = \operatorname{dom} f + \operatorname{dom} g$

as f: real valued $\Rightarrow \forall y \in \mathcal{H}$ $f(y) \in \mathbb{R} \Rightarrow \operatorname{dom} f = \{\tilde{x} \in \mathcal{H} \mid f(\tilde{x}) < +\infty\} = \mathcal{H}$

$\operatorname{dom}(f \square g) = \overbrace{\operatorname{dom} f}^{\mathcal{H}} + \operatorname{dom} g = \mathcal{H}$

$\therefore \exists_{y \in \mathcal{H}}$ $y = \underset{\tilde{x} \in \mathcal{H}}{\operatorname{argmin}} \; f(\tilde{x}) + g(x - \tilde{x})$

$(f \square g)(x) = (f \boxed{\square} g)(x) = f(y) + g(x-y)$

Proposition 18.7. [$f, g \in \Gamma_0(\mathcal{H})$; $x, y \in \mathcal{H}$; $f \square g \in \Gamma_0(\mathcal{H})$, $(f \square g)(x) = f(y) + g(x-y)$;

f: Gateaux differentiable at y] $\Rightarrow$

- $x \in \operatorname{cont}(f \square g)$
- $(f \square g)$: Gateaux differentiable at x
- $\nabla(f \square g)(x) = \nabla f(y)$
- f: Frechet differentiable at $y \Rightarrow$ $\boxed{f \square g: \text{Frechet differentiable at } x}$ ☺

$\therefore f \square g$: Frechet differentiable on $\mathcal{H}$.

∎

2017-07-13 9:16 AM

Corollary 18.11.

[$f: \in \Gamma_0(\mathcal{H})$, $\operatorname{dom} \partial f = \operatorname{int} \operatorname{dom} f$]

f: Gateaux differentiable on $\operatorname{int} \operatorname{dom} f \Leftrightarrow$

f^*: strictly convex on every nonempty convex subset of $\operatorname{dom} \partial f^*$

$\Rightarrow$

$\operatorname{dom} \partial f^* = \nabla f(\operatorname{int} \operatorname{dom} f)$

Proof: ($\Leftarrow$)

/*Proposition 18.9. given

Proposition 18.9 [$f \in \Gamma_0(\mathcal{H})$; f^*: strictly convex on every nonempty convex subset of $\operatorname{dom} \partial f^*$] $\Rightarrow f$: Gateaux differentiable on $\operatorname{int} \operatorname{dom} f$.

*/

$\therefore f$: Gateaux differentiable on $\operatorname{int} \operatorname{dom} f$ (O)

($\Rightarrow$)

given: f: Gateaux differentiable on $\operatorname{int} \operatorname{dom} f$

take $\forall x \in \operatorname{int} \operatorname{dom} f$ $\therefore f$: Gateaux differentiable at x

Proposition 17.26 [$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex; $x \in \operatorname{dom} f$]

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

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

$\therefore \partial f(x) = \{\nabla f(x)\}$ $\forall x \in \operatorname{int} \operatorname{dom} f$

So $\nabla f(\operatorname{int} \operatorname{dom} f) = \partial f(\operatorname{int} \operatorname{dom} f)$

$\partial f(\operatorname{dom} \partial f)$ /* given: $\operatorname{dom} \partial f = \operatorname{int} \operatorname{dom} f$

Proposition 17.26 [$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex ; $x \in \operatorname{dom} f$]

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

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

$\therefore \partial f(x) = \{\nabla f(x)\}$
$\forall x \in \operatorname{int} \operatorname{dom} f$

So $\nabla f(\operatorname{int} \operatorname{dom} f) = \partial f(\operatorname{int} \operatorname{dom} f)$

$= \partial f(\operatorname{dom} \partial f)$ /* given: $\operatorname{dom} \partial f = \operatorname{int} \operatorname{dom} f$ */

$= \operatorname{ran} \partial f$ /* For any T: set valued operator

$T(\operatorname{dom} T) = \operatorname{ran} T$

$T\{x : T(x) \neq \emptyset\} = \bigcup_{x \in \operatorname{dom} T} Tx$

$= \bigcup_{x \in \operatorname{dom} T} Tx \cup \bigcup_{x \notin \operatorname{dom} T} \overbrace{Tx}^{\emptyset}$

$= \bigcup_{x \in X} Tx = \operatorname{ran} T$ */

$= \operatorname{dom} (\partial f)^{-1}$ /* For any set-valued operator: $\operatorname{ran} T = \operatorname{dom}(T^{-1})$ */

$= \operatorname{dom} \partial f^*$ /*

Corollary 6.24 [$f \in \Gamma_0(\mathcal{H})$] $(\partial f)^{-1} = \partial f^*$ */

given

$\therefore f$: Gateaux differentiable on
$\operatorname{int} \operatorname{dom} f \Rightarrow \nabla f(\operatorname{int} \operatorname{dom} f) = \operatorname{dom} \partial f^*$ (1)

also from

given of ($\Rightarrow$)

Proposition 18.10 [$f: \mathcal{H} \to \Gamma_0(\mathcal{H})$, Gateaux differentiable on $\operatorname{int} \operatorname{dom} f$] $\Rightarrow$ [f^*: strictly convex on every nonempty convex subset of $\nabla f(\operatorname{int} \operatorname{dom} f)$]

$\therefore f$: Gateaux differentiable on $\operatorname{int} \operatorname{dom} f$
$\Rightarrow f^*$: strictly convex on every nonempty convex subset of

$\nabla f(\operatorname{int} \operatorname{dom} f) = \operatorname{dom} \partial f^*$ (from (1)) ... (2)

So the equivalence is established from (0), (2)
the implication follows from (1)

$\blacksquare$

Corollary 18.17.

2017-07-14 1:52 PM

[$L \in \mathcal{B}(\mathcal{H})$, self-adjoint, positive ; $x \in \mathcal{H}$] $\Rightarrow$
$\|L\| \langle Lx | x \rangle \geq \|Lx\|^2$

Proof:

Denote: $f: \mathcal{H} \to \mathbb{R} : y \mapsto \frac{1}{2} \langle Ly | y \rangle$

recall:

set $u := 0$, $\tilde{L} := \frac{L}{2}$

Example 2.46 [$L \in \mathcal{B}(\mathcal{H})$, $u \in \mathcal{H}$, $x \in \mathcal{H}$, $f: \mathcal{H} \to \mathbb{R}$, $y \mapsto \langle \tilde{L}y | y \rangle - \langle y | u \rangle$] $\Rightarrow$
f: twice Frechet differentiable on $\mathcal{H}$; $\nabla f(x) = (\tilde{L} + \tilde{L}^*)x - u$; $\nabla^2 f(x) = \tilde{L} + \tilde{L}^*$

f: twice Frechet differentiable on $\mathcal{H}$.

$\nabla f(x) = (\tilde{L} + \tilde{L}^*)x - u$

$= 2\tilde{L}x$ // L: self adjoint.

$= Lx$

So, $\|\nabla f(x) - \nabla f(y)\| = \|Lx - Ly\| = \|L(x-y)\|$ // L: linear

$\leq \|L\| \, \|x - y\|$ // any linear T: $\forall_x \ \|T\| \, \|x\| \geq \|Tx\|$

$\therefore$ ∇f: Lipschitz continuous with constant $\|L\| = \beta$
f: Frechet differentiable on $\mathcal{H}$

Theorem 18.15. [$f \in \Gamma_0(\mathcal{H})$; $\beta \in \mathbb{R}_{++}$; $h = f - \frac{1}{\beta} q$: $q = \frac{1}{2} \|\cdot\|^2$]
The following are equivalent:

(i) $\boxed{f: \text{Frechet differentiable on } \mathcal{H} \ \wedge \ \nabla f: \beta\text{-Lipschitz continuous}}$ $\Leftrightarrow$

(ii) f: Frechet differentiable on $\mathcal{H}$ $\wedge$
$\forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}} \ \langle x - y | \nabla f(x) - \nabla f(y) \rangle \leq \beta \|x - y\|^2$ $\Leftrightarrow$

(iii) (descent lemma)
f: Frechet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad \langle x-y \mid \nabla f(x)-\nabla f(y)\rangle \geqslant \beta \|x-y\|^2 \Leftrightarrow$

(iii) (descent lemma)

f: Frechet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad f(y) \leqslant f(x) + \langle y-x \mid \nabla f(x)\rangle + \frac{\beta}{2}\|x-y\|^2 \Leftrightarrow$

(iv) f: Frechet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad f^*(\nabla f(y)) \geqslant f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\beta}\|\nabla f(x)-\nabla f(y)\|^2 \Leftrightarrow$

(v) $\boxed{f: \text{Frechet differentiable on } \mathcal{H} \wedge \nabla f : \tfrac{1}{\beta} \text{ cocoercive}} \Leftrightarrow$

(vi) $\beta q - f = \frac{\beta}{2}\|\cdot\|^2 - f$: convex /* this is quite amazing, because in generally this is subtraction of two convex functions */ $\Leftrightarrow$

(vii) $f^* - \frac{1}{\beta} q$: convex /* same as saying f^*: $\frac{1}{\beta}$ strongly convex */ $\Leftrightarrow$

(viii) $h \in \Gamma_0(\mathcal{H}) \wedge f = \left(\frac{1}{\beta}\right)(h^*) = \beta q - {}^{\beta}h = \beta \square d \Leftrightarrow$

(ix) $h \in \Gamma_0(\mathcal{H}) \wedge \nabla f = \operatorname{Prox}_{\beta h} = \beta \square d = \beta\left(\mathrm{Id} - \operatorname{Prox}_{h^*/\beta}\right)$

$\Rightarrow \nabla f : \left[\frac{1}{\beta}\right]$ cocoercive $=$
$\approx \left[1/\|L\|\right]$

Definition 4.4 (β-cocoercive / β-inverse strongly monotone)

[D: nonempty subset of $\mathcal{H}$, $T: D \to \mathcal{H}$, $\beta \in \mathbb{R}_{++}$]

T: β-cocoercive (or β-inverse strongly monotone) $\overset{\text{def}}{\Leftrightarrow}$ βT: firmly nonexpansive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x\in D} \forall_{y\in D} \quad \langle x-y \mid Tx-Ty\rangle \geqslant \beta\|Tx-Ty\|^2$

$$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \quad \underbrace{\langle x-y \mid \underbrace{\nabla f(x)-\nabla f(y)}_{Lx-Ly}\rangle}_{\langle x-y \mid Lx-Ly\rangle} \geqslant \frac{1}{\|L\|}\underbrace{\|\nabla f(x)-\nabla f(y)\|^2}_{\|Lx-Ly\|^2}$$

$= \langle x-y \mid L(x-y)\rangle$

now set $y=0$ then

$\forall_{x\in\mathcal{H}} \quad \langle x \mid Lx\rangle \geqslant \frac{1}{\|L\|}\|Lx\|^2$

$\Leftrightarrow \forall_{x\in\mathcal{H}} \quad \|L\|\,\langle x \mid Lx\rangle \geqslant \|Lx\|^2$

$\blacksquare$

Corollary 18.19.

[$(f_i)_{i\in I}$: finite family of functions in $\Gamma_0(\mathcal{H})$;
$(\alpha_i)_{i\in I}$: finite family of real numbers, $\subseteq [0,1]$, $\sum_{i\in I}\alpha_i = 1$:
$q = \frac{1}{2}\|\cdot\|^2$; $h = \left(\sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)\right)^* - q$]

$\operatorname{Prox}_h = \sum_{i\in I}\alpha_i \operatorname{Prox}_{f_i}$

2017-07-17 9:17 AM

Proof:

$f := \sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)$

now: $\forall_{i\in I} \quad f_i \in \Gamma_0(\mathcal{H})$

$\Rightarrow \forall_{i\in I} \quad f_i^* \in \Gamma_0(\mathcal{H})$ /*: **Corollary 11.15** (Or immediate consequence of Fenchel-Moreau theorem)
[$f(x) \in \mathbb{R}$
$\cdot\ f^*\in\Gamma_0(\mathcal{H})$
$\cdot\ f^{**} = f$]
*/

$\Rightarrow \forall_{i\in I} \quad (f_i^* \,\square\, q) = \left(f_i^* \,\square\, \frac{1}{2}\|\cdot\|^2\right) = {}^{1}f_i^*$: /* convex, real-valued $\Rightarrow$ proper, continuous, $\in \Gamma_0(\mathcal{H})$ and frechet differentiable on $\mathcal{H}$ */

/* **Proposition 11.15.** /* Properties of power norm infimal convolution of a convex function */
[$f \in \Gamma_0(\mathcal{H})$
$\gamma \in \mathbb{R}_{++}$
$\theta \in]1,+\infty[$
$\cdot\ q_\theta = f \,\square\, \frac{1}{\theta\gamma}\|\cdot\|^\theta : \mathcal{H} \to]-\infty,+\infty] : x \mapsto \inf_{u\in\mathcal{H}}\left(f(u) + \frac{1}{\theta\gamma}\|x-u\|^\theta\right)$]
$\cdot\ q_\theta = f \,\square\, \frac{1}{\theta\gamma}\|\cdot\|^\theta$: convex, real-valued, continuous, exact, infimum uniquely attained.
/* Moreau envelope satisfies all these properties */

Proposition 11.29
[$f \in \Gamma_0(\mathcal{H})$
$\gamma \in \mathbb{R}_{++}$]
$\gamma f : \mathcal{H} \to \mathbb{R}$, Frechet differentiable : $\gamma f = f \,\square\, \frac{1}{2\gamma}\|\cdot\|^2$
$\nabla(\gamma f) = \frac{1}{\gamma}(\mathrm{Id} - \operatorname{Prox}_{\gamma f}) = \frac{1}{\gamma}\operatorname{Prox}_{\gamma f^*}$: $\frac{1}{\gamma}$ Lipschitz continuous */

$\Rightarrow \sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)$: $\in \Gamma_0(\mathcal{H})$ and frechet differentiable on $\mathcal{H}$
$\underbrace{\qquad}_{f}$

/* **Proposition 2.16** For all α strong functions from $\mathcal{H}$ to $\mathbb{R}$ and α is closed under addition and multiplication by α, α strong and convex $\Rightarrow$ α strong convex
$\begin{cases} \sum_{i=1}^{n} \alpha_i \\ \prod_{i=1}^{n} \alpha_i \end{cases}$ strong convex
*/

with $\nabla f = \nabla\left(\sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)\right)$

$= \sum_{i\in I}\alpha_i\,\nabla\underbrace{(f_i^* \,\square\, q)}_{{}^{1}f_i^*}$

$= \sum_{i\in I}\alpha_i\,\nabla({}^{1}f_i^*)$

$= \sum_{i\in I}\alpha_i \operatorname{Prox}_{f_i^{**}}$

$= \sum_{i\in I}\alpha_i \operatorname{Prox}_{f_i}$ /* $f_i \in \Gamma_0(\mathcal{H}) \Rightarrow f_i^{**} = f_i$ */

$\therefore f := \sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)$: $\in \Gamma_0(\mathcal{H})$ and frechet differentiable on $\mathcal{H}$ with $\nabla f = \sum_{i\in I}\alpha_i \operatorname{Prox}_{f_i}$ and 1-Lipschitz continuous . . . (1)

also, $\operatorname{Prox}_{f_i}$: firmly nonexpansive $\Leftrightarrow$ $\frac{1}{2}$ averaged /*

/* **Proposition 11.37** & ***
[$f \in \Gamma_0(\mathcal{H})$]
Prox_f : firmly nonexpansive
$\operatorname{Id} - \operatorname{Prox}_f$: firmly nonexpansive $\Leftrightarrow$ $\operatorname{Prox}_{f^*}$: firmly nonexpansive
*/

$\therefore \nabla f = \sum_{i\in I}\alpha_i \operatorname{Prox}_{f_i}(\cdot)$: $\frac{1}{2}$ averaged (2)
($\sum_{i\in I}\alpha_i = 1$)
firmly nonexpansive

/* **Proposition 4.30** (Addition of averaged operators) /* A convex combination of α_i-averaged operators is $\max \alpha_i$-averaged */
[D: nonempty subset of $\mathcal{H}$
$(T_i)_{i\in I}$: ($\cdot$ finite family of nonexpansive operators
$\cdot\ \forall_{i\in I}\ T_i$: α_i-averaged, $\alpha \in]0,1[$)
$(w_i)_{i\in I}$: $\sum_{i\in I} w_i = 1$] $\Rightarrow$ $\sum_{i\in I} w_i T_i$: ($\max_{i\in I}\alpha_i$) averaged.
*/

now use.

/* set: $f = \sum_{i\in I}\alpha_i\,(f_i^* \,\square\, q)$, $\beta := 1$,

now use.

$\cdot \forall_{i\in I}\ T_i : \underset{\in[0,1]}{\alpha_i}$-averaged)

$(\alpha_i)_{i\in I} : \sum_{i\in I} \alpha_i = 1$] $\Rightarrow$ $\sum_{i\in I} \alpha_i T_i$: ($\max_{i\in I} \alpha_i$) averaged.

*/

A set: $f = \sum_{i\in I} \alpha_i (f_i^* \square q)$, $\beta := 1$,

Theorem 18.15 [$f \in \Gamma_0(\mathcal{H})$, $\beta \in \mathbb{R}_{++}$; $h = f^* - \frac{1}{\beta} q$: $h + \frac{1}{2}\|\cdot\|^2$]

The following are equivalent:

holds by (1)

(i) f: Fréchet differentiable on $\mathcal{H}$ ✓ $\wedge$ ∇f : β-Lipschitz continuous $\Leftrightarrow$

(ii) f: Fréchet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}}$ $\langle x-y | \nabla f(x) - \nabla f(y)\rangle \geq \beta \|\nabla f(x) - \nabla f(y)\|^2$ $\Leftrightarrow$

(iii) (descent lemma)

f: Fréchet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}}$ $f(y) \leq f(x) + \langle y-x | \nabla f(x)\rangle + \frac{\beta}{2}\|x-y\|^2$ $\Leftrightarrow$

(iv) f: Fréchet differentiable on $\mathcal{H}$ $\wedge$

$\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}}$ $f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle y | \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\beta}\|\nabla f(x) - \nabla f(y)\|^2$ $\Leftrightarrow$

(v) f: Fréchet differentiable on $\mathcal{H}$ $\wedge$ ∇f : $\frac{1}{\beta}$-cocoercive $\Leftrightarrow$

(vi) $\beta q - f = \frac{\beta}{2}\|\cdot\|^2 - f$: convex. // this is quite amazing, because in generally this is subtraction of two convex functions */ $\Leftrightarrow$

(vii) $f^* - \frac{1}{\beta} q$: convex // same as saying f^*: $\frac{1}{\beta}$-strongly convex */ $\Leftrightarrow$

(viii) $h \in \Gamma_0(\mathcal{H})$ $\wedge$ $f = \square^{(\frac{1}{\beta})}(h^*) = \beta q - {}^{*}h + \beta \frac{1}{2}$ $\Leftrightarrow$

(ix) $h \in \Gamma_0(\mathcal{H})$ $\wedge$ $\nabla f = \mathrm{Prox}_{\beta h} \circ \beta \mathrm{Id} = \beta(\mathrm{Id} - \mathrm{Prox}_{h^*/\beta})$

$h = \sum_{i\in I} \alpha_i (f_i^* \square q)^* - q \in \Gamma_0(\mathcal{H})$

$\underbrace{\nabla f = (\mathrm{Prox}_h \circ \mathrm{Id})}_{} = (\mathrm{Id} - \mathrm{Prox}_{h^*})$

$\Rightarrow$ $\boxed{\mathrm{Prox}_h = \sum \alpha_i \mathrm{Prox}_{f_i}}$ ∴

Part 1

July 18, 2017 8:53 AM

Theorem 19.1. [$f \in \Gamma_0(\mathcal{H})$; $g \in \Gamma_0(\mathcal{K})$; $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$; $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$;
$\mu = \inf(f + g \circ L)(\mathcal{H})$; $\mu^* = \inf(f^* \circ L^* + g^{*\vee})(\mathcal{K})$; $x \in \mathcal{H}$; $v \in \mathcal{K}$]

The following are equivalent:

(i) x: primal solution; v: dual solution; $\mu = -\mu^*$

(ii) $L^* v \in \partial f(x)$, $-v \in \partial g(Lx)$;

(iii) $x \in \partial f^*(L^* v) \cap L^{-1}(\partial g^*(-v))$

2017-07-18 9:48 AM

Proof:

recall:

- **Theorem 16.23.** [$f \in \Gamma_0(\mathcal{H})$; $x \in \mathcal{H}$; $u \in \mathcal{H}$] The following are equivalent:

(i) $(x,u) \in \operatorname{gra} \partial f \Leftrightarrow$

(ii) $(u,-1) \in N_{\operatorname{epi} f}(x, f(x)) \Leftrightarrow$

(iii) $f(x) + f^*(u) = \langle x | u \rangle \Leftrightarrow$

(iv) $(u,x) \in \operatorname{gra} \partial f^*$

(i) $\Leftrightarrow$ (ii):

x: primal solution, v: dual solution, $\mu = -\mu^*$

$\Leftrightarrow$ $f(x) + g(Lx) = \mu = -\mu^* = -(f^*(L^* v) + g^*(-v))$

$\Leftrightarrow$ $f(x) + f^*(L^* v) + g(Lx) + g^*(-v) = 0$

$\Leftrightarrow$ $\big(f(x) + f^*(L^* v) - \langle x | L^* v \rangle\big) + \big(g(Lx) + g^*(-v) + \underbrace{\langle x | L^* v \rangle}_{\langle Lx | v \rangle = -\langle Lx | -v \rangle}\big) = 0$

$\Leftrightarrow$ $\underbrace{\big(f(x) + f^*(L^* v) - \langle x | L^* v \rangle\big)}_{(\mathfrak{F})} + \underbrace{\big(g(Lx) + g^*(-v) - \langle Lx | -v \rangle\big)}_{(\mathfrak{G})} = 0$ /* using Fenchel-Moreau 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! */

$\forall_{x \in \mathcal{H}} \forall_{u \in \mathcal{H}} \quad f(x) + f^*(u) \geq \langle x | u \rangle$

so both term $(\mathfrak{F})$, $(\mathfrak{G})$ are non-negative, whose addition are zero, so both of them will be 0 seperately */

$\Leftrightarrow$ $f(x) + f^*(L^* v) - \langle x | L^* v \rangle = 0 \;\wedge\; g(Lx) + g^*(-v) - \langle Lx | -v \rangle = 0$

$\Leftrightarrow$ $L^* v \in \partial f(x) \;\wedge\; -v \in \partial g(Lx)$ /*using (·) */

(ii) $\Leftrightarrow$ (iii)

(ii): $L^* v \in \partial f(x) \;\wedge\; -v \in \partial g(Lx)$

Corollary 16.24. [$f \in \Gamma_0(\mathcal{H})$] $(\partial f)^{-1} = \partial f^*$

$\Updownarrow \qquad\qquad \Updownarrow$

$\underbrace{(\partial f)^{-1}}_{(\partial f)^*}(L^* v) \ni x \;\Big|\; \underbrace{(\partial g)^{-1}}_{(\partial g)^*}(-v) \ni Lx \leftrightarrow (\partial g)^*(-v) \ni Lx \leftrightarrow L^{-1}(\partial g)^*(-v) \ni x$

$\Updownarrow$

$x \in (\partial f)^*(L^* v)$

$\Leftrightarrow$ $x \in \partial f^*(L^* v) \cap L^{-1}(\partial g^*(-v))$

$\blacksquare$

Part 2

July 25, 2017 8:01 AM

2017-07-25 9:01 AM

Corollary 19.6.

[$\Phi \in \Gamma_0(\mathcal{H})$; $z \in \mathcal{H}$; $r \in \mathcal{K}$; $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$, $\|L\| \leq 1$; $-r \in \operatorname{sri}(L(\operatorname{dom}\Phi))$;

$T : \mathcal{K} \to \mathcal{K} : v \mapsto r + v - L(\operatorname{Prox}_\Phi(L^* v + z))$] $\Rightarrow$

- T: firmly nonexpansive
- $\operatorname{Fix} T \neq \emptyset$
- $\forall_{v \in \operatorname{Fix} T}$ $x = \operatorname{Prox}_\Phi(L^* v + z)$: unique solution to $\min_{\substack{x \in \mathcal{H} : \\ Lx = r}} \Phi(x) + \frac{1}{2}\|x - z\|^2$

Proof:

$S = \{ v \in \mathcal{K} \mid L(\operatorname{Prox}_\Phi(L^* v + z)) = r \}$

recall: set $\psi = \iota_{\{0\}} : x \mapsto \begin{cases} 0, & \text{if } x = 0 \\ +\infty, & \text{else} \end{cases}$, this has conjugate function: $\psi^*(u) = \sup_{\substack{x \in \operatorname{dom} f \\ = \{0\}}} (\langle x | u \rangle - \psi(x)) = 0$

**

Proposition 19.4. [$\Phi \in \Gamma_0(\mathcal{H})$; $\psi \in \Gamma_0(\mathcal{K})$; $z \in \mathcal{H}$; $r \in \mathcal{K}$; $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$; $r \in \operatorname{sri}(\operatorname{dom}\psi - L(\operatorname{dom}\Phi))$;

Consider the problem:

$$\min_{x \in \mathcal{H}} \Phi(x) + \underbrace{\psi(Lx - r)}_{(Lx - r)} + \frac{1}{2}\|x - z\|^2 \qquad (19.6)$$

$= \min_{\substack{x \in \mathcal{H} \\ Lx = r}} \Phi(x) + \frac{1}{2}\|x - z\|^2$: which is same as the problem (*)

together with the problem:

$$\left(\min_{v \in \mathcal{K}} {}^{\square}(\Phi^*)(L^* v + z) + \psi^*(-v) - \langle v | r \rangle \right) \qquad (19.7)$$

$$= \left(\min_{v \in \mathcal{K}} \frac{1}{2}\|L^* v + z\|^2 - {}^{\square}\Phi(L^* v + z) + \psi^*(-v) - \langle v | r \rangle \right) \qquad (19.8)$$

v: solution to (19.7)] (19.6) has unique solution: $x = \operatorname{Prox}_\Phi(L^* v + z)$

becomes

$$\min_{v \in \mathcal{K}} {}^{\square}(\Phi^*)(L^* v + z) + \underbrace{\psi^*(-v)}_{0} - \langle v | r \rangle$$

$$= \min_{v \in \mathcal{K}} \underbrace{{}^{\square}(\Phi^*)(L^* v + z) - \langle v | r \rangle}_{\text{gradient} = 0 \text{ at optimal } v} \qquad (0)$$

/*

* Some important implications of Moreau's theorem:

- $(f \boxdot \frac{1}{2}\|\cdot - x\|^2) + (f^* \boxdot \frac{1}{2}\|\cdot - x^*\|^2) = \frac{1}{2}\|\cdot - x\|^2$
- $\operatorname{Prox}_f + \operatorname{Prox}_{f^*} = \mathrm{Id}$
- $\operatorname{Prox}_f = \mathrm{Id} - \nabla(f \boxdot \frac{1}{2}\|\cdot - x\|^2) = \nabla(f^* \boxdot \frac{1}{2}\|\cdot - x^*\|^2)$ */

i.e., $0 = \nabla_v \left[{}^{\square}(\Phi^*)(L^* v + z) - \langle v | r \rangle \right]$ /* ${}^{\square}(f^*) = (f + \frac{1}{2}\|\cdot\|^2)^* = f^* \boxdot \frac{1}{2}\|\cdot\|^2$

$$= \underbrace{L(\operatorname{Prox}_\Phi(L^* v + z))}_{\text{firmly nonexpansive}} - r$$

$\nabla({}^{\square}(f^*)) = \nabla(f^* \boxdot \frac{1}{2}\|\cdot\|^2) = \operatorname{Prox}_f$ */

so. $\nabla_v {}^{\square}(\Phi^*)(v) = \operatorname{Prox}_\Phi(\cdot)$

$\Rightarrow \nabla_v {}^{\square}(\Phi^*)(L^* v + z) = L \operatorname{Prox}_\Phi(L^* v + z)$

also $\nabla_v \langle v | r \rangle = r$ */

$\therefore$ $T : v \mapsto r + v - L(\operatorname{Prox}_\Phi(L^* v + z))$: firmly nonexpansive

/* from

Exercise 4.7 Let $\mathcal{K}$ be a real Hilbert space, let $T : \mathcal{K} \to \mathcal{K}$ be firmly nonexpansive, let $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ be such that $\|L\| \leq 1$, let $z \in \mathcal{H}$, and let $r \in \mathcal{K}$. Show that $x \mapsto z + L^* T(z + Lx)$ is firmly nonexpansive. */

$v \in \operatorname{Fix} T \leftrightarrow r + v - L(\operatorname{Prox}_\Phi(L^* v + z)) = v$

$\leftrightarrow L(\operatorname{Prox}_\Phi(L^* v + z)) = v$

$\leftrightarrow v \in S$

$\therefore \operatorname{Fix} T = S =$ solution set of (0)

$\therefore$ $\forall_{v \in \operatorname{Fix} T}$ $x = \operatorname{Prox}_\Phi(L^* v + z)$: unique solution to $\min_{\substack{x \in \mathcal{H} : \\ Lx = r}} \Phi(x) + \frac{1}{2}\|x - z\|^2$

$\blacksquare$

Parametric Duality:

- 2017-07-28 10:40 AM

Proposition 19.11.

[[$F : \mathcal{H} \times K \to]-\infty, +\infty]$;

$\vartheta : K \to [-\infty, +\infty] : y \mapsto \inf F(\mathcal{H}, y)$, the associated value function]] $\Rightarrow$

(i) $\vartheta^* = F^*(0, \cdot)$

(ii) $-\inf F^*(0, K) = \vartheta^{**}(0) \leq \vartheta(0) = \inf F(\mathcal{H}, 0)$

<u>Proof</u> : Assume F : proper

/* recall

(Definitions for parametric duality)

Definition 19.10 - [$F : \mathcal{H} \times K \to]-\infty, +\infty]$]

Primal problem : $\min_{x \in \mathcal{H}} F(x, 0)$; solution to this is called the primal solution

Dual problem : $\min_{v \in K} F^*(0, v)$; " " " " " " dual solution

Value function : $\vartheta : K \to [-\infty, +\infty] : y \mapsto \inf F(\mathcal{H}, y)$ // ϑ : vartheta

*/

(i) /* * Proposition 13.28 : * *

[K : real Hilbert space

$F : \mathcal{H} \times K \to]-\infty, +\infty]$, proper

$f : \mathcal{H} \to [-\infty, +\infty] : x \mapsto \inf F(x, K)$]]

$f^* = F^*(\cdot, 0)$ */

in our case, we have $\vartheta : K \to [-\infty, +\infty] : y \mapsto \inf F(\mathcal{H}, y)$

so, $\vartheta^* = F^*(0, \cdot)$

(ii)

* Proposition 12.1 - * *

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

(i) $f^*(0) = -\inf f(\mathcal{H})$ /* this is the most easy conjugate function to be evaluated since it is just a flat line. */

(ii) $-\infty \in f^*(\mathcal{H}) \Leftrightarrow f = +\infty \Leftrightarrow f^* = -\infty$

(iii) f^* : proper $\Rightarrow f$: proper

(iv) [$u \in \mathcal{H}$]

$$f^*(u) = \sup_{x \in \operatorname{dom} f} \left(\langle x | u \rangle - f(x) \right) = \sup_{(x, \xi) \in \operatorname{epi} f} \left(\langle x | u \rangle - \xi \right)$$

(v) $f^* = \iota^*_{\operatorname{gra} f}(\cdot, -1) = \iota^*_{\operatorname{epi} f}(\cdot, -1)$

from (i) we have : $\vartheta^*(x) = F^*(0, x)$

so, $\vartheta^*(0) = -\inf \vartheta^*(K)$

$$= -\inf_{x \in K} F^*(0, x) = -\inf F^*(0, K) \quad \ldots (1)$$

again : by definition :

$$\vartheta(y) = \inf F(\mathcal{H}, y)$$

$$\Rightarrow \vartheta(0) = \inf F(\mathcal{H}, 0) \quad \ldots (2)$$

/*

* Proposition 13.14 . * * *

[$f, g : \mathcal{H} \to [-\infty, +\infty]$] /* these properties are actually quite useful */

(i) $f^{**} \leq f$ /* Biconjugate is always a lower bound */

(ii) $f \leq g \Rightarrow (f^* \geq g^*,\ f^{**} \leq g^{**})$ /* Conjugation flips the relation between two functions */

(iii) $f^{***} = f^*$ /* Conjugating thrice and once is the same */

(iv) $(\check{f})^* = f^*$ /* $\check{f} = \sup\{g \in \Gamma(\mathcal{H}) : g \leq f\}$: lower semicontinuous convex envelope of f */

/* Conjugating a function and its lower continuous convex envelope gives the same function */

*/

apply to ϑ at 0

$$\vartheta^{**}(0) \leq \vartheta(0)$$

$$\Rightarrow -\inf F^*(0, K) = \vartheta^{**}(0) \leq \vartheta(0) = \inf F(\mathcal{H}, 0)$$

$\blacksquare$

Proposition 19.14 :

Proposition 19.14. 2017-07-31 10:51 AM

[$F \in \Gamma_0(\mathcal{H}, \mathcal{K})$; $(x,v) \in \mathcal{H} \times \mathcal{K}$] The following are equivalent:

(i) (x: primal solution,
v: dual solution,
$\inf F(\mathcal{H},0) = -\inf F^*(0,\mathcal{K}) \in \mathbb{R}$) $\Leftrightarrow$

(ii) $F(x,0) + F^*(0,v) = 0$ $\Leftrightarrow$

(iii) $(0,v) \in \partial F(x,0)$ $\Leftrightarrow$

(iv) $(x,0) \in \partial F^*(0,v)$

<u>Proof</u>:

(i) $\Rightarrow$ (ii):

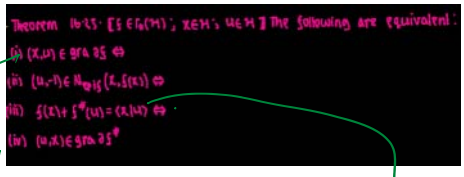
given $\begin{cases} (x\text{: primal solution} \leftrightarrow x = \operatorname{argmin}_{\tilde{x}} F(\tilde{x},0) \\ v\text{: dual solution} \leftrightarrow v = \operatorname{argmin}_{\tilde{v}} F^*(0,\tilde{v}) \\ \inf F(\mathcal{H},0) = -\inf F^*(0,\mathcal{K}) \in \mathbb{R}) \end{cases}$

So, $F(x,0) = \inf F(\mathcal{H},0) = -\inf F^*(0,\mathcal{K})$
$= -F^*(0,v) \in \mathbb{R}$

$\Rightarrow F(x,0) + F^*(0,v) = 0$

(ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv):

/* recall Theorem 16.23.


Theorem 16.23. [$f \in \Gamma_0(\mathcal{H})$; $x \in \mathcal{H}$; $u \in \mathcal{H}$] The following are equivalent:
(i) $(x,u) \in \operatorname{gra} \partial f$ $\Leftrightarrow$
(ii) $(u,-1) \in N_{\operatorname{epi} f}(x, f(x))$ $\Leftrightarrow$
(iii) $f(x) + f^*(u) = \langle x | u \rangle$ $\Leftrightarrow$
(iv) $(u,x) \in \operatorname{gra} \partial f^*$

*/

$F(x,0) + F^*(0,v) = 0$

$= \langle (x,0) | (0,v) \rangle$

$\Leftrightarrow ((x,0),(0,v)) \in \operatorname{gra} \partial F$

$\Leftrightarrow \partial F(x,0) \ni (0,v)$ /* this is (iii) */

$\Leftrightarrow ((0,v),(x,0)) \in \operatorname{gra} \partial F^*$

$\Leftrightarrow \partial F^*(0,v) \ni (x,0)$ /* this is (iv) */

Finally: (iv) $\Rightarrow$ (i)

but (iv) $\Leftrightarrow$ (ii) } so we will prove (ii) $\Rightarrow$ (i)

(ii) says

$$F(x,0) + F^*(0,v) = 0$$

/*

Proposition 19.11. [[$F: \mathcal{H} \times K \to]-\infty,+\infty]$; $\vartheta : K \to [-\infty,+\infty] : y \mapsto \inf F(\mathcal{H},y)$, the associated value function]] $\Rightarrow$

(i) $\vartheta^* = F^*(0,\cdot)$

(ii) $\underbrace{-\inf F^*(0,K)}_{\text{optimal value of the dual problem}} = \vartheta^{**}(0) \leq \vartheta(0) = \underbrace{\inf F(\mathcal{H},0)}_{\text{optimal value of the primal problem}}$

*/

$$-F^*(0,v) \leq -\inf F^*(0,K) \leq \inf F(\mathcal{H},0) \leq F(x,0) \quad \ldots (0)$$

so v: dual solution,
x: primal solution, and

$F(x,0) + F^*(0,v) \geq 0$ but

we already have
$F(x,0) + F^*(0,v) = 0$

so, inequality (0) collapses: and we have:

$$\inf F(\mathcal{H},0) = -\inf F^*(0,K) \in \mathbb{R}$$

$\blacksquare$

Part 1

Thursday, August 25, 2016 5:02 PM

2017-03-29 2:36 PM

Proposition 20.2.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$]

(i) A: monotone $\Leftrightarrow$

(ii) A: accretive $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \forall_{\alpha\in[0,1]}\ \|x-y+\alpha(u-v)\| \geq \|x-y\| \Leftrightarrow$

(iii) $\forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \|x-u\|^2+\|y-v\|^2 \leq \|x-v\|^2+\|y-u\|^2$

Proof:

(ii) $\Leftrightarrow$ (i)

A: monotone $\Leftrightarrow$ $\forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \langle x-y|u-v\rangle \geq 0$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \langle x-y|v-u\rangle \leq 0$

/* Lemma 2.12. (Characterizes obtuse and perpendicular vectors)

(i) $\forall_{x,y\in\mathcal{H}}\ \langle x|y\rangle \leq 0 \Leftrightarrow \forall_{\alpha\in\mathbb{R}_+}\ \|x\| \leq \|x-\alpha y\| \Leftrightarrow \forall_{\alpha\in[0,1]}\ \|x\| \leq \|x-\alpha y\|$

(ii) $\forall_{x,y\in\mathcal{H}}\ x \perp y \Leftrightarrow \forall_{\alpha\in\mathbb{R}}\ \|x\| \leq \|x-\alpha y\| \Leftrightarrow \forall_{\alpha\in[-1,1]}\ \|x\| \leq \|x-\alpha y\|$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \forall_{\alpha\in[0,1]}$

$$\|x-y\| \leq \|(x-y)-\alpha(v-u)\|$$

$$= \|x-y+\alpha(u-v)\| \qquad \text{(ii)} \Leftrightarrow \text{(i)}$$

(iii) $\Leftrightarrow$ (ii):

$\forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \|x-u\|^2+\|y-v\|^2 \leq \|x-v\|^2+\|y-u\|^2$

$\Leftrightarrow \cancel{\|x\|^2}+\cancel{\|u\|^2}-2\langle x|u\rangle+\cancel{\|y\|^2}+\cancel{\|v\|^2}-2\langle y|v\rangle \leq \cancel{\|x\|^2}+\cancel{\|v\|^2}-2\langle x|v\rangle+\cancel{\|y\|^2}+\cancel{\|u\|^2}-2\langle y|u\rangle$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ -2\langle x|u\rangle-2\langle y|v\rangle+2\langle x|v\rangle+2\langle y|u\rangle \leq 0$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \langle x|v-u\rangle+\langle y|u-v\rangle \leq 0$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ -\langle x|u-v\rangle+\langle y|u-v\rangle = \langle -(x-y)|(u-v)\rangle \leq 0$

$\Leftrightarrow \forall_{(x,u)\in \operatorname{gra} A}\ \forall_{(y,v)\in \operatorname{gra} A}\ \langle x-y|u-v\rangle \leq 0$

$\Leftrightarrow$ A: monotone

■

Example 20.8.

[$D: \neg\emptyset, \subseteq \mathcal{H}$

$T: D \to \mathcal{H}$

$A = \mathrm{Id}-T$]

A: monotone $\Leftrightarrow$ T: pseudononexpansive

/* T: pseudononexpansive $\Leftrightarrow \forall_{x\in D}\ \forall_{y\in D}\ \|Tx-Ty\|^2 \leq \|x-y\|^2+\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2$ */

8/25/2016 5:15 PM

Proof: take T: pseudononexpansive $\Leftrightarrow$

$\forall_{x,y\in D}$

$\|x-y\|^2+\|(\mathrm{Id}-T)x-(\mathrm{Id}-T)y\|^2 \geq \|Tx-Ty\|^2$

$\forall x, y \in D$

$$\|x-y\|^2 + \underbrace{\|(2d-T)x-(2d-T)y\|^2}_{\|(x-y)-(Tx-Ty)\|^2 = \|x-y\|^2 - 2\langle x-y|Tx-Ty\rangle + \|Tx-Ty\|^2} \geq \|Tx-Ty\|^2$$

$$\leftrightarrow \quad 2\|x-y\|^2 - 2\langle x-y|Tx-Ty\rangle + \cancel{\|Tx-Ty\|^2} \geq \cancel{\|Tx-Ty\|^2}$$

$$\leftrightarrow \quad \underbrace{\|x-y\|^2}_{\langle x-y|x-y\rangle} - \langle x-y|Tx-Ty\rangle \geq 0$$

$$\leftrightarrow \quad \langle x-y|\underbrace{x-y-Tx+Ty}_{\underbrace{(2d-T)}_{A}x-\underbrace{(2d-T)}_{A}y}\rangle \geq 0 \leftrightarrow \langle x-y|Ax-Ay\rangle \geq 0$$

$\therefore$ T: nonexpansive $\leftrightarrow \forall_{x,y\in D} \langle x-y|Ax-Ay\rangle \geq 0 \overset{\text{def}}{\leftrightarrow} A$: monotone.

$\blacksquare$

Example 20.12. (Projection operator of any set is monotone)

8/25/2016 5:46 PM

[C: nonempty, $\subseteq H$] Π_C: monotone

Proof:

$\forall_{(x,u),(y,v)\in \operatorname{gra}\Pi_C}$ // recall: $\Pi_C(x) = \{p \in C \mid \|x-p\| = d_C(x) = \inf_{y\in C}\|x-y\|\}$

$\Updownarrow$ // $\therefore p \in \Pi_C(x) \leftrightarrow \|x-p\| = d_C(x)$

$\Pi_C x \ni u \leftrightarrow \|x-u\| = d_C(x) = \inf_{\tilde{u}\in C}\|x-\tilde{u}\| \leq \|x-\tilde{u}\| \ \forall_{\tilde{u}\in C} \leq \|x-v\|$ // setting $\tilde{u} := v$

$\Pi_C y \ni v \leftrightarrow \|y-v\| = d_C(y) = \inf_{\tilde{v}\in C}\|y-\tilde{v}\| \leq \|y-\tilde{v}\| \ \forall_{\tilde{v}\in C} \leq \|y-u\|$ // setting $\tilde{v} := u$

$$\Rightarrow \quad \underbrace{\|x-u\|^2}_{\leq \|x-v\|} + \underbrace{\|y-v\|^2}_{\leq \|y-u\|} \leq \|x-v\|^2 + \|y-u\|^2$$

$$\leftrightarrow \quad \cancel{\|x\|^2} + \cancel{\|u\|^2} - 2\langle x|u\rangle + \cancel{\|y\|^2} + \cancel{\|v\|^2} - 2\langle y|v\rangle \leq \cancel{\|x\|^2} + \cancel{\|y\|^2} - 2\langle x|v\rangle + \cancel{\|y\|^2} + \cancel{\|u\|^2} - 2\langle y|u\rangle$$

$$\leftrightarrow \quad -\langle x|u\rangle - \langle y|v\rangle \leq -\langle x|v\rangle - \langle y|u\rangle$$

$$\leftrightarrow \quad \langle x|u\rangle + \langle y|v\rangle \geq \langle x|v\rangle + \langle y|u\rangle$$

$$\leftrightarrow \quad \langle x|u-v\rangle - \langle y|u-v\rangle \geq 0$$

$$\leftrightarrow \quad \langle x-y|u-v\rangle \geq 0$$

So, $\forall_{(x,u),(y,v)\in \operatorname{gra}\Pi_C} \langle x-y|u-v\rangle \geq 0 \leftrightarrow \Pi_C$: monotone.

$\blacksquare$

Part 2

Monday, March 21, 2016 9:07 AM

3/21/2016 9:38 AM **Theorem 20-21.** // Any monotone operator has a maximally monotone extension.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone]

$\exists \tilde{A}: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone $\quad \operatorname{gra} A \subseteq \operatorname{gra} \tilde{A}$

Proof:

(case 1)

$\operatorname{gra} A \neq \emptyset$

$\mathcal{M} = \{ B: \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \subseteq \operatorname{gra} B \}$: the set of all monotone operators containing A

$\mathcal{M}$: nonempty (contains A)

$\mathcal{M}$: partially ordered /+ notation. $\forall_{B_1, B_2 \subseteq \mathcal{M}}$ $\operatorname{gra} B_1 \subseteq \operatorname{gra} B_2 \Leftrightarrow B_1 \leq B_2$ #/

verified the definition in 1,2,3

$\mathcal{C}$: (chain in $\mathcal{M}$ // $\mathcal{C} \subseteq \mathcal{M}$

(totally ordered subset of a partially ordered set)

$\Rightarrow \forall_{B_1, B_2 \in \mathcal{C}}$ $\operatorname{gra} B_1 \subseteq \operatorname{gra} B_2$ or $\operatorname{gra} B_2 \subseteq \operatorname{gra} B_1$

So, $\underbrace{\bigcup_{C \in \mathcal{C}} \operatorname{gra} C}_{\operatorname{gra} C_{max, \mathcal{C}}} \supseteq \operatorname{gra} B \quad \forall B \in \mathcal{C}$ // $\because \mathcal{C} \subseteq \mathcal{M}$, and each member of $\mathcal{M}$ contains $\operatorname{gra} B$, so $\bigcup_{C \in \mathcal{C}} C$ must contain B.

/+ caution. C_{max} is not necessarily in

$\Rightarrow \forall B \in \mathcal{C} \quad B \leq C_{max, \mathcal{C}}$ // i.e., $\forall_{B \in \mathcal{C}}$ $\operatorname{gra} B \subseteq \operatorname{gra} \bigcup_{C \in \mathcal{C}} C = C_{max, \mathcal{C}}$

So, for $\forall \mathcal{C}$: chain in $\mathcal{M}$ $\forall B \in \mathcal{C}$ $B \leq C_{max, \mathcal{C}}$

/+ Zorns lemma: [A: partially ordered set] ($\forall \mathcal{C}$: chain in A $\exists$ upper bound for $\mathcal{C}$) // if

$\Rightarrow$ (A : contains a maximal element a) // then

$\forall c \in A \quad a \leq c \Rightarrow c = a$ #/

So, using Zorn's lemma we have : $\mathcal{M}$ will contain a maximal element

$\exists \tilde{A} \in \mathcal{M} \quad \forall B \in \mathcal{M} \quad B \geq \tilde{A}$

$\operatorname{gra} B \supseteq \operatorname{gra} \tilde{A}$

$\Rightarrow \operatorname{gra} \tilde{A} = \operatorname{gra} A \quad [\because A \in \mathcal{M}]$

now by definition of $\mathcal{M}$, there cannot be any other $B \in \mathcal{M}$ containing $\tilde{A}$

So, $\tilde{A}$: maximally monotone.

(case 2)

$\operatorname{gra} A = \emptyset$

$\tilde{A}$: $\operatorname{gra} \tilde{A} = \{(0,0)\}$ is a maximally monotone extension of A.

■

8/28/2016 10:32 AM **Proposition 20-31.**

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, $x \in \mathcal{H}$]

Ax: closed, convex

Proof:

case 1: $x \notin \operatorname{dom} A \Rightarrow$ $Ax = \emptyset$: closed, convex and noninteresting

case 2: $x \in \operatorname{dom} A \Rightarrow$ $Ax \neq \emptyset = \{\ldots, u, \ldots\} \Rightarrow \exists_u$ $u \in Ax \Leftrightarrow (x,u) \in \operatorname{gra} A \Leftrightarrow \forall_{(y,v) \in \operatorname{gra} A} \langle x-y | u-v \rangle \geq 0$: true for $u \in Ax$

/# [$A: \mathcal{H} \to 2^{\mathcal{H}}$] A: maximally monotone $\Leftrightarrow \forall_{(x,u)} \left((x,u) \in \operatorname{gra} A \Leftrightarrow \forall_{(y,v) \in \operatorname{gra} A} \langle x-y | u-v \rangle \geq 0 \right)$ #/

So, $\forall_u \left(u \in Ax \Leftrightarrow \forall_{(y,v) \in \operatorname{gra} A} \langle x-y | u-v \rangle \geq 0 \right)$ ($(x,u) \in \operatorname{gra} A$)

$\Rightarrow \quad Ax = \{ u \in \mathcal{H} | \forall_{(y,v) \in \operatorname{gra} A} \langle x-y | u-v \rangle \geq 0 \}$

$= \underbrace{\bigcap_{(y,v) \in \operatorname{gra} A} \underbrace{\{ u \in \mathcal{H} | \langle x-y | u-v \rangle \geq 0 \}}_{\text{closed, convex set in } u}}_{\text{intersection of a bunch of closed, convex sets} \Rightarrow \text{convex, closed}}$

$\therefore Ax$: closed, convex ■

8/28/2016 1:15 PM # **Proposition 20-32.**

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone operator

$(x_k, u_k)_{k \in \mathbb{N}}$: bounded net in $\operatorname{gra} A$

$(x,u) \in \mathcal{H} \times \mathcal{H}$]

(i) $(x_k \to x, u_k \to u) \Rightarrow (x,u) \in \operatorname{gra} A$

(ii) $(x_k \rightharpoonup x, u_k \rightharpoonup u) \Rightarrow (x,u) \in \operatorname{gra} A$

Proof:

$\forall_{(y,v) \in \operatorname{gra} A}$

$\Leftrightarrow Ay \ni v$

/# A: maximally monotone $\Leftrightarrow \forall_{(x,u)} \left((x,u) \in \operatorname{gra} A \Leftrightarrow \forall_{(y,v) \in \operatorname{gra} A} \langle x-y | u-v \rangle \geq 0 \right)$ . . . (0)

$\because (x_k, u_k)_{k \in \mathbb{N}} \subseteq \operatorname{gra} A$

↳ index set of the net $\mathbb{N}$

$\forall_{k \in \mathbb{N}} \langle x_k - y | u_k - v \rangle \geq 0$ . . . (i)

Lemma 2-36: [$(x_\alpha)_{\alpha \in A}$: net in $\mathcal{H}$, bounded, $x_\alpha \xrightarrow{\mathcal{H}} x$,

$(u_\alpha)_{\alpha \in A}$: net in $\mathcal{H}$, $u_\alpha \xrightarrow[\mathcal{H}]{} u$] $\Rightarrow \langle x_\alpha | u_\alpha \rangle \to \langle x | u \rangle$ #/

using this in our case : $(x_k)_{k \in \mathbb{N}}, (u_k)_{k \in \mathbb{N}}$: bounded, $x_k \to x$, $u_k \to u \Rightarrow z_k = y \to x - y$, $u_k = v \to u - v$

$\overset{\text{Theorem 2-36}}{\Rightarrow} \langle z_k - y | u_k - v \rangle \to \langle x - y | u - v \rangle$

$\overset{(0)(i)}{\Rightarrow} \langle x - y | u - v \rangle \geq 0$

$\Rightarrow (x,u) \in \operatorname{gra} A$ ■

(ii) Apply (i) to A^{-1}

■

Proposition 20-33.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

(i) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{strong} \times \mathcal{H}^{weak}$

$\Leftrightarrow \forall_{(x_n, u_n)_{n \in \mathbb{N}} \subseteq \operatorname{gra} A}, \forall_{(x,u) \in \mathcal{H} \times \mathcal{H}} : x_n \to x, u_n \rightharpoonup u \quad (x,u) \in \operatorname{gra} A$

(ii) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{weak} \times \mathcal{H}^{strong}$

$\Leftrightarrow \forall_{(x_n, u_n)_{n \in \mathbb{N}} \subseteq \operatorname{gra} A}, \forall_{(x,u) \in \mathcal{H} \times \mathcal{H}} : x_n \rightharpoonup x, u_n \to u \quad (x,u) \in \operatorname{gra} A$

(iii) $\operatorname{gra} A$: closed in $\mathcal{H}^{strong} \times \mathcal{H}^{strong}$

3/21/2016 2:28 PM Proof: (i)

Lemma 2-32 : [$(x_n)_{n \in \mathbb{N}} \subseteq \mathcal{H}$] $x_n \rightharpoonup \square \Leftrightarrow (x_n)_{n \in \mathbb{N}}$: bounded, has almost one weak sequential cluster point

(20-32 (i))

Proposition (Behavior of a bounded net in a maximally monotone operator : strong

(ii) Let x_n: closed in H $\Rightarrow$ x_n

3/21/2016 2:28 PM

Proof: (i)

Lemma 2.38: $[(x_n)_{n\in\mathbb{N}} \subseteq H]$ $x_n \rightharpoonup \square \Leftrightarrow (x_n)_{n\in\mathbb{N}}$: bounded, has almost one weak sequential cluster point]

(20.32 (i))

Proposition (Behavior of a bounded net in a maximally monotone operator : strong convergence of only one point in a pair suffices)

$[A: \mathcal{H} \to 2^{\mathcal{H}}; (x_n, u_n)_{n\in\mathbb{N}}$: bounded net in $\operatorname{gra} A; (x,u) \in \mathcal{H} \times \mathcal{H}]$

✓ $(x_n \to x; u_n \to u) \Rightarrow (x,u) \in \operatorname{gra} A$

• $(x_n \to x; u_n \to u) \Leftrightarrow (x,u) \in \operatorname{gra} A$ */

given: $(x_n, u_n) \in \operatorname{gra} A$, $(x,u) \in \mathcal{H}\times\mathcal{H}$: $(x_n, u_n) \to (x,u)$

/* $\left.\begin{array}{l} x_n \rightharpoonup x \Rightarrow x_n \rightharpoonup x \\ u_n \rightharpoonup u \end{array}\right\} \Rightarrow (x_n, u_n) \rightharpoonup (x,u)$ */

So Lemma 2.38 says $(x_n, u_n)_{n\in\mathbb{N}}$: bounded net, has almost one weak sequential cluster point]

now Proposition 20.32 (i) can be applied:

$(x,u) \in \operatorname{gra} A$

(ii) A^{-1}: maximally monotone // A: maximally monotone $\Rightarrow A^{-1}$: maximally monotone

now apply (i)'s logic to A^{-1}

(iii) follows trivially.

$\square$

Proposition 20.35.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, at most single-valued

$\operatorname{dom} A$: linear subspace /* $\forall_{\alpha,\beta\in\mathbb{R}} \forall_{x,y\in\operatorname{dom}A} \Leftrightarrow Ax, Ay \neq \emptyset$ $\alpha x + \beta y \in \operatorname{dom} A$ */

$A|_{\operatorname{dom} A}$: linear /* $\forall_{\alpha,\beta\in\mathbb{R}} \forall_{x,y\in\operatorname{dom}A}$ $A(\alpha x+\beta y) = \alpha Ax + \beta Ay$ */

$\forall_{x\in\operatorname{dom}A} \forall_{y\in\operatorname{dom}A}$ $\langle x|Ay\rangle = \langle Ax|y\rangle$ /* on Euclidean space it means A: symmetric matrix */

$h: \mathcal{H} \to]-\infty, +\infty]: x \mapsto \begin{cases} \frac{1}{2}\langle x|Ax\rangle, & \text{if } x\in\operatorname{dom}A \\ +\infty \end{cases}$ // Another way of saying this is: $h(x) = \frac{1}{2}\langle x \mid Ax\rangle + \iota_{\operatorname{dom}A}$

$f: \mathcal{H} \to]-\infty, +\infty]$, $x \mapsto \sup_{y\in\operatorname{dom}A} \left(\langle x|Ay\rangle - h(y)\right)$]

$\Rightarrow$
- (i) $f + \iota_{\operatorname{dom}A} = h$
- (ii) $f \in \Gamma_0(\mathcal{H})$
- (iii) $\partial f = A$
- (iv) $f = h^{**}$

8/30/2016 8:38 AM

Proof: $\operatorname{dom} A = \operatorname{dom} h$ /* caution: the first one is the domain of an operator, second one " " " " " function. $\operatorname{dom} A = \{x\in\mathcal{H}: Ax \neq \emptyset\}$, $\operatorname{dom} h = \{x\in\mathcal{H}: h(x) < +\infty\}$ however definitionwise they are the same */

Take $x \in \operatorname{dom} A = \operatorname{dom} h$

(i) $\forall_{y\in\operatorname{dom}A}$ $0 \leq \langle x-y|Ax-Ay\rangle$ // A: maximally monotone we have just used the definition here

$= \langle x|Ax\rangle + \langle y|Ay\rangle - \underbrace{\langle y|Ax\rangle}_{=} - \langle x|Ay\rangle$

// $\langle x|Ay\rangle$ as $x, y \in \operatorname{dom} A$ by given

$= 2h(x) + 2h(y) - 2\langle x|Ay\rangle$ /* using definition of A */

$\Rightarrow \forall_{\substack{x\in\operatorname{dom}A \\ =\operatorname{dom}h}} \forall_{\substack{y\in\operatorname{dom}A \\ =\operatorname{dom}h}}$ $0 \leq h(x) + h(y) - \langle x|Ay\rangle$

$\Leftrightarrow$ " " $\langle x|Ay\rangle - h(y) \leq h(x)$

$\Leftrightarrow \forall_{x\in\operatorname{dom}A} \underbrace{\sup_{y\in\operatorname{dom}A} \langle x|Ay\rangle - h(y)}_{f(x)\ //\text{ by definition}} \leq h(x)$

$\Leftrightarrow \forall_{x\in\operatorname{dom}A = \operatorname{dom}h}$ $f(x) \leq h(x)$

$\Leftrightarrow \forall_{x\in\operatorname{dom}A}$ $f(x) + \iota_{\operatorname{dom}A}(x) \leq h(x)$ /* as $\forall x\in\operatorname{dom}A$ $\iota_{\operatorname{dom}A}(x) = 0$ */

note that even $\forall_{x\in\operatorname{dom}A}$ $f(x) + \underbrace{\iota_{\operatorname{dom}A}(x)}_{=+\infty} = +\infty = h(x)$ // $\because \operatorname{dom}A = \operatorname{dom}h$

So, $\forall_{x\in\mathcal{H}}$ $f(x) + \iota_{\operatorname{dom}A}(x) \leq h(x)$ . . . (i)

again by definition,

$\forall_{\substack{x\in\operatorname{dom}A \\ \subseteq\mathcal{H}}} f(x) = \sup_{y\in\operatorname{dom}A} \left(\langle x|Ay\rangle - h(y)\right) \geq \langle x|Ay\rangle - h(y)$ $\forall_{y\in\operatorname{dom}A}$

set $y := x \in \operatorname{dom} A$

$\Rightarrow$ $f(x) \geq \underbrace{\langle x|Ax\rangle}_{2h(x)} - h(x) = h(x)$ // $\because x\in\operatorname{dom}A$, using the definition

$\Rightarrow \forall_{x\in\operatorname{dom}A = \operatorname{dom}h}$ $f(x) - h(x) \geq 0$

$\Rightarrow \inf_{x\in\operatorname{dom}A} \left(f(x) - h(x)\right) \geq 0$

$\Leftrightarrow \inf_{x\in\mathcal{H}} \left(f(x) - h(x) + \iota_{\operatorname{dom}A}\right) \geq 0$ //just using the indicator function trick

$\Leftrightarrow \forall_{x\in\mathcal{H}}$ $f(x) - h(x) + \iota_{\operatorname{dom}A} \geq 0$

$\Leftrightarrow \forall_{x\in\mathcal{H}}$ $f(x) + \iota_{\operatorname{dom}A}(x) \geq h(x)$ . . . (ii)

From (i) and (ii):

$\forall_{x\in\mathcal{H}}$ $f(x) + \iota_{\operatorname{dom}A}(x) = h(x)$

$\Leftrightarrow \boxed{f + \iota_{\operatorname{dom}A} = h}$ $\blacksquare$

(ii)

$\forall_{x\in\mathcal{H}}$ $f(x) = \sup_{y\in\operatorname{dom}A} \underbrace{\left(\langle x|Ay\rangle - h(y)\right)}_{\text{continuous affine function over } x \text{ (the variable) with parameter } y}$

$\Rightarrow$ lower semicontinuous convex $\in \Gamma(\mathcal{H})$

sup over the parameter $y \in \operatorname{dom} A \Rightarrow$ so $f(\cdot)$: convex

// $\sup(-\cdot) = -\inf(-\cdot)$

by definition, $f: \mathcal{H} \to]-\infty, +\infty] \Rightarrow -\infty \notin f(\mathcal{H})$, also $\operatorname{dom} f \neq \emptyset$ as $f(0) = \sup_{y\in\operatorname{dom}A} \left(-h(y)\right) = -\underbrace{\inf_{y\in\operatorname{dom}A} \left(-h(y)\right)}_{=\operatorname{dom}h} = -\underbrace{\inf_{y\in\operatorname{dom}h} h(y)}_{=0} = -\frac{1}{2}\underbrace{\inf_{y\in\operatorname{dom}A} \langle y|Ay\rangle}_{=0}$ /*because A: maximally monotone, $A|_{\operatorname{dom}A}$: linear $\Rightarrow A$: positive semidefinite on $\operatorname{dom}A$ comes from: Example 20.15. A: monotone, linear on $S \Leftrightarrow \forall_{x\in S} \langle Ax|x\rangle \geq 0$ so, $y = 0$ will yield the minimum value of 0 */

$\Downarrow$

f: proper

$\therefore f$: proper, $\in \Gamma(\mathcal{H}) \Leftrightarrow f \in \Gamma_0(\mathcal{H})$ $\blacksquare$

(iii) $\forall_{y\in\mathcal{H}}$ $f(x) + \langle y-x|Ax\rangle$ // recall that we have taken $x \in \operatorname{dom} A$

// using (i)

$= h(x) - \underbrace{\iota_{\operatorname{dom}A}(x)}_{0} + \langle y-x|Ax\rangle$ // $x\in\operatorname{dom}A \Rightarrow \iota_{\operatorname{dom}A}(x) = 0$

$= \underbrace{h(x)}_{\frac{1}{2}\langle x|Ax\rangle} + \underbrace{\langle y-x|Ax\rangle}_{\langle y|Ax\rangle - \langle x|Ax\rangle} = \langle y|Ax\rangle - \overbrace{\frac{1}{2}\langle x|Ax\rangle}^{h(x)} = \langle y|Ax\rangle - h(x) \leq \underbrace{\sup_{\tilde{x}\in\operatorname{dom}A} \left(\langle y|A\tilde{x}\rangle - h(\tilde{x})\right)}_{f(y)\ /\text{* by definition *}} = f(y)$

/* $\forall x\in\operatorname{dom}A$ $h(x) = \frac{1}{2}\langle x|Ax\rangle$ */

So, we have:

$\forall_{y\in\mathcal{H}}$ $f(y) \geq f(x) + \langle Ax|y-x\rangle \Leftrightarrow Ax \in \partial f(x) \Leftrightarrow (x, Ax) \in \operatorname{gra} \partial f$

/* recall that $Ax \in \partial f(x) \Leftrightarrow \forall y\in\mathcal{H}$ $f(y) \geq f(x) + \langle Ax|y-x\rangle$ */

$\{\langle x|Ax\rangle \quad \langle y|Ay\rangle \quad \langle z|Az\rangle\}$ $\qquad \| \qquad f(y)$ /* by definition */

/* $\forall x \in \operatorname{dom} A \quad h(x) = \frac{1}{2}\langle x|Ax\rangle$ */

So, we have:

$\forall_{y\in\mathcal{H}} \quad f(y) \geq f(x) + \langle Ax|y-x\rangle \leftrightarrow Ax \in \partial f(x) \quad \leftrightarrow (x, Ax) \in \operatorname{gra}\partial f$

/* recall that $\quad y \in \partial f(x) \leftrightarrow (\forall y \in \mathcal{H}) \quad f(y) \geq f(x) + \langle y|y-x\rangle$ */

Now, $\quad \operatorname{gra} A = \{(x, Ax) \mid x \in \operatorname{dom} A\}$ /* for $x \notin \operatorname{dom} A \quad (x, Ax) = (x, \emptyset)$ which is a trivial member of all set, so we do not need to consider that */

So, $\forall_{(x,Ax)\in\operatorname{gra}A} \quad (x, Ax) \in \operatorname{gra}\partial f$ /* ∂f: monotone for any proper function */

$\leftrightarrow \quad \operatorname{gra} A \subseteq \operatorname{gra}\partial f \quad \ldots (i)$

now A: maximally monotone

$\leftrightarrow \neg\left(\exists_{B:\mathcal{H}\to 2^{\mathcal{H}}, \text{monotone}} \ \operatorname{gra} A \subsetneq \operatorname{gra} B\right) \ldots (ii)$

From (i) and (ii) the only possibility is:

$\operatorname{gra} A = \operatorname{gra}\partial f$

$\leftrightarrow \quad A = \partial f$

(*) from (ii) we have proven: $\quad f \in \Gamma_0(\mathcal{H}) \Rightarrow \underbrace{f = (f + \iota_{\operatorname{dom}\partial f})^{**}}_{\text{proof}} = \overbrace{(f + \iota_{\operatorname{dom}\partial f})^{**}}^{h \text{ (from (i))}} = h^{**}$

/* from corollary 16.31: $\quad f \in \Gamma_0(\mathcal{H}) \Rightarrow f = (f + \iota_{\operatorname{dom}\partial f})^{**}$ */ ∎

3/21/2016 9:34 PM

*Corollary 20.39.

[$F: \mathcal{H}\times\mathcal{H} \to]-\infty, +\infty]$, autoconjugate

$A: \operatorname{gra} A = \{(x,u) \in \mathcal{H}\times\mathcal{H} : F(x,u) = \langle x|u\rangle\}$ /* $(x,u) \in A \leftrightarrow \forall x \exists u \leftrightarrow F(x,u) = \langle x|u\rangle$ */

A: maximally monotone

Proof: // Proof blueprint: Direct application of Theorem 16.52

/* Proposition 13.31: [$F: \mathcal{H}\times\mathcal{H} \to]-\infty, +\infty]$, autoconjugate] $\forall (x,u)\in\mathcal{H}\times\mathcal{H} \quad F(x,u) \geq \langle x|u\rangle$, $F^*(u,x) \geq \langle x|u\rangle$

Proposition 16.52: (Sum characterization of autoconjugate function)

[$F: \mathcal{H}\times\mathcal{H} \to]-\infty, +\infty]$, autoconjugate

$(x,u) \in \mathcal{H}\times\mathcal{H}$]

$F(x,u) = \langle x|u\rangle \Leftrightarrow F^*(u,x) = \langle x|u\rangle \Leftrightarrow (x,u) \in \partial F(x,u) \Leftrightarrow (x,u) \in \operatorname{Fix}_{\mathcal{H}} (\mathrm{Id}_{\mathcal{H}}, \mathrm{Id}_{\mathcal{H}})$

* Theorem 16.52:

[$F: \mathcal{H}\times\mathcal{H} \to]-\infty, +\infty]$, convex function, F^*: proper, $\overbrace{F \geq \langle\cdot|\cdot\rangle}^{\mathcal{E}}$

$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\operatorname{gra} A = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid F^*(u,x) = \langle x|u\rangle\}$]

- A: monotone
- $F \geq \langle\cdot|\cdot\rangle \Rightarrow A$: maximally monotone

Proof: $\quad f \in \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H})$ /* using Fenchel-Moreau corollary */

$\underbrace{(f \oplus f^*)}_{\in \Gamma_0(\mathcal{H}\times\mathcal{H})} : (x,y) \mapsto f(x) + f^*(y)$ /* $\oplus$: Hilbert direct sum */

/* Proposition 6.15. convexity is preserved under Hilbert direct sum */

$\ldots (0)$

$(f \oplus f^*)$: autoconjugate. (1) /* F: autoconjugate $\leftrightarrow F^* = F^T$ where $F^T(u,x) = F(x,u)$ */

/* proof $(f \oplus f^*)(x,y) = f(x) + f^*(y)$ /* recall: Hilbert direct sum: $(f_1 \oplus f_2)(x,y) = f_1(x) + f_2(y)$ */

$(f \oplus f^*)^T(x,y) = (f \oplus f^*)(y,x) = f(y) + f^*(x)$

$(f \oplus f^*)^{**}(x,y) = (f^{**} \oplus f^{**})(x,y)$ /* proposition 13.23. $\left(\bigoplus_{i\in I} f_i\right)^* = \bigoplus_{i\in I} f_i^*$ */ (I: finite, $f_i: \mathcal{H}\to]-\infty, +\infty]$)

$= f^*(x) + \underbrace{f^{**}(y)}_{f(y)}$ /* Fenchel-Moreau corollary: $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^{**} = f$ */

$= f^*(x) + f(y)$

$\therefore (f \oplus f^*)^T(x,y) = (f \oplus f^*)^*(x,y) \leftrightarrow (f \oplus f^*)$: autoconjugate */

/* Proposition 16.9.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper] $\quad u \in \partial f(x) \leftrightarrow x \in \partial f^*(u) \leftrightarrow \underbrace{f(x) + f^*(u) = \langle x|u\rangle}_{(f\oplus f^*)(x,u) = \langle x|u\rangle}$ */ ($(x,u) \in \operatorname{gra}\partial f$) this is a beautifully symmetric result.

So, $\quad \operatorname{gra}\partial f = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid (f \oplus f^*)(x,u) = \langle x|u\rangle\} \ldots (2)$

$(0), (1), \quad \Rightarrow$

$\left.\begin{array}{l} (f \oplus f^*)^T: \text{autoconjugate}, \in \Gamma_0(\mathcal{H}\times\mathcal{H}) \\ (2) \end{array}\right\} \Rightarrow \partial f$: maximally monotone /

∎

Example 20.41

[$f \in \Gamma_0(\mathcal{H}\times\mathcal{H})$, autoconjugate $\quad$ /* F: autoconjugate $\leftrightarrow F^* = F^T$ */ /* $F^T(u,x) = F(x,u)$ */

$A: \operatorname{gra} A = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid F(x,u) = \langle x|u\rangle\}$] /* $(x,u) \in A \leftrightarrow F(x,u) = \langle x|u\rangle$ */

A: maximally monotone

3/22/2016 2:55 PM

*Theorem 20.40.

[$f \in \Gamma_0(\mathcal{H})$]

∂f: maximally monotone

Proof Blueprint: Essentially applies Corollary 20.39 which says:

[$F \in \Gamma_0(\mathcal{H}\times\mathcal{H})$, autoconjugate, $A: \operatorname{gra} A = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid F(x,u) = \langle x|u\rangle\}$] A: maximally monotone

We set the bivariate function $F := f \oplus f^*$, and $A := \partial f$, which satisfies the antecedent, thus making ∂f a maximally monotone operator.

Proof: $\quad f \in \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H})$ /* using Fenchel-Moreau corollary */

$\underbrace{(f \oplus f^*)}_{\in \Gamma_0(\mathcal{H}\times\mathcal{H})} : (x,y) \mapsto f(x) + f^*(y)$ /* $\oplus$: Hilbert direct sum */

/* Proposition 6.15. convexity is preserved under Hilbert direct sum */

$\ldots (0)$

$(f \oplus f^*)$: autoconjugate. (1) /* F: autoconjugate $\leftrightarrow F^* = F^T$ where $F^T(u,x) = F(x,u)$ */

/* proof $(f \oplus f^*)(x,y) = f(x) + f^*(y)$ /* recall: Hilbert direct sum: $(f_1 \oplus f_2)(x,y) = f_1(x) + f_2(y)$ */

$(f \oplus f^*)^T(x,y) = (f \oplus f^*)(y,x) = f(y) + f^*(x)$

$(f \oplus f^*)^{**}(x,y) = (f^{**} \oplus f^{**})(x,y)$ /* proposition 13.23. $\left(\bigoplus_{i\in I} f_i\right)^* = \bigoplus_{i\in I} f_i^*$ */ (I: finite, $f_i: \mathcal{H}\to]-\infty, +\infty]$)

$= f^*(x) + \underbrace{f^{**}(y)}_{f(y)}$ /* Fenchel-Moreau corollary: $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^{**} = f$ */

$= f^*(x) + f(y)$

$\therefore (f \oplus f^*)^T(x,y) = (f \oplus f^*)^*(x,y) \leftrightarrow (f \oplus f^*)$: autoconjugate */

/* Proposition 16.9.

[$f: \mathcal{H} \to]-\infty, +\infty]$, proper] $\quad u \in \partial f(x) \leftrightarrow x \in \partial f^*(u) \leftrightarrow \underbrace{f(x) + f^*(u) = \langle x|u\rangle}_{(f\oplus f^*)(x,u) = \langle x|u\rangle}$ */ ($(x,u) \in \operatorname{gra}\partial f$) this is a beautifully symmetric result.

So, $\quad \operatorname{gra}\partial f = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid (f \oplus f^*)(x,u) = \langle x|u\rangle\} \ldots (2)$

$(0), (1), \quad \Rightarrow$

$\left.\begin{array}{l} (f \oplus f^*)^T: \text{autoconjugate}, \in \Gamma_0(\mathcal{H}\times\mathcal{H}) \\ (2) \end{array}\right\} \Rightarrow \partial f$: maximally monotone /

∎

Example 20.41

[C: nonempty closed convex subset of $\mathcal{H}$]

N_C: maximally monotone

Proof:

$f \in \Gamma_0(\mathcal{H}) \Rightarrow \partial f$: maximally monotone $\quad$ (Theorem 20.40)

set $\left.\begin{array}{l} f = \iota_C \in \Gamma_0(\mathcal{H}) \\ \therefore \partial f = N_C \quad \text{/* } C\text{: nonempty} \end{array}\right\}$

N_C: maximally monotone.

∎

3/22/2016 3:13 PM

* Example 20.49.

[A: monotone, $\in \mathcal{B}(\mathcal{H})$; $q_A: \mathcal{H} \to \mathbb{R}: x \mapsto \frac{1}{2}\langle x|Ax\rangle$]

$\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \quad f_A(x,u) = 2q_A^*\left(\frac{1}{2}u + \frac{1}{2}A^*x\right)$

Proof Blueprint: Uses definition of conjugate function, adjoint operator. The rest is the direct manipulation of the definition of the Fitzpatrick function.

Proof: $\forall (x,u) \in \mathcal{H}\times\mathcal{H}$

$f_A(x,u) = \sup_{\substack{(y,v)\in\operatorname{gra}A \\ Ay = v}} \left(\langle y|u\rangle + \langle x|v\rangle - \langle y|v\rangle\right)$ /* $A \in \mathcal{B}(\mathcal{H}) \leftrightarrow Ax$: single valued from $\mathcal{H}$ to $\mathcal{H}$. */

$= \sup_{(y,v):Ay=v} \left(\langle y|u\rangle + \langle x|Ay\rangle - \langle y|Ay\rangle\right)$

$= \sup_{y\in\mathcal{H}} \left(\langle y|u\rangle + \langle x|Ay\rangle - \langle y|Ay\rangle\right)$

/* As y is the only independent variable */

$= 2\sup_{y\in\mathcal{H}} \left(\underbrace{\frac{1}{2}\langle y|u\rangle + \frac{1}{2}\langle x|Ay\rangle}_{\frac{1}{2}\langle y|(u + A^*x)\rangle} - \frac{1}{2}\langle y|Ay\rangle\right)$

$\langle x|Ay\rangle = \langle A^*x|y\rangle = \langle y|A^*x\rangle$ /* Adjoint of a linear bounded operator $\langle x|Ay\rangle = \langle A^*x|y\rangle$ */

$= 2\sup_{y\in\mathcal{H}} \left(\frac{1}{2}\langle y|u + A^*x\rangle - \frac{1}{2}\langle y|Ay\rangle\right)$

/* recall the definition of conjugate function: $f^*(u) = \sup_x \left(-c(x) + \langle x|u\rangle\right)$

$\langle A^*x|y\rangle$ /* adjoint of a linear bounded operator

$= \langle y|Ax\rangle$

$\langle x|Ay\rangle = \langle A^*x|y\rangle$ */

$\frac{1}{2}\langle y|(u+A^*x)\rangle$

$= \sup_{y\in\mathcal{H}} \left(\frac{1}{2}\langle y|u+A^*x\rangle - \frac{1}{2}\langle y|Ay\rangle \right)$

$\langle y|\frac{1}{2}(u+A^*x)\rangle \qquad f_A(y)$

/* recall the definition of conjugate function:

$f^*(u) = \sup_{x\in\mathcal{H}} \left(-f(x)+\langle x|u\rangle\right)$

now $f_A(x) = \frac{1}{2}\langle x|Ax\rangle$

$f_A^*(u) = \sup_{x\in\mathcal{H}} \left(-\frac{1}{2}\langle x|Ax\rangle + \langle x|u\rangle\right)$

$= \sup_{y\in\mathcal{H}} \left(-f_A(y) + \langle y|\frac{1}{2}(u+A^*x)\rangle\right)$

$= f_A^*\left(\frac{1}{2}u+\frac{1}{2}A^*x\right)$

∎

9/2/2016 10:14 AM

Proposition 20.98. (Fitzpatrick function of maximally monotone operators)

$[A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone operator$]$

$\cdot F_A \geq \langle\cdot|\cdot\rangle$

$\cdot \operatorname{gra} A = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u) = \langle x|u\rangle\}$

Proof:

$(x,u)\in\mathcal{H}\times\mathcal{H}$

$(x,u)\in\operatorname{gra} A \Leftrightarrow F_A(x,u) = \langle x|u\rangle$ (1)

Proposition 20.98 (Fitzpatrick function of monotone operator)

$[A:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\operatorname{gra}A\neq\emptyset$

$(\forall (x,u)\in\mathcal{H}\times\mathcal{H})$

(i) $(x,u)\in\operatorname{gra}A \Rightarrow F_A(x,u)=\langle x|u\rangle$

now if $(x,u)\notin\operatorname{gra}A \Leftrightarrow \{(x,u)\}\cup\operatorname{gra}A$: not monotone (2)

/* A : maximal monotone $\overset{\text{def}}{\Leftrightarrow} \forall(\tilde{x},\tilde{u})\ \left((\tilde{x},\tilde{u})\in\operatorname{gra}A \Leftrightarrow \forall(y,v)\in\operatorname{gra}A\ \langle\tilde{x}-y|\tilde{u}-v\rangle\geq 0\right)$

so, $(x,u)\notin\operatorname{gra}A \Leftrightarrow \exists(y,v)\in\operatorname{gra}A\ \langle x-y|u-v\rangle<0$

so if $\operatorname{gra}B = \{(x,u)\}\cup\operatorname{gra}A$ then

$(x,u)\in\operatorname{gra}B,\ (y,v)\in\operatorname{gra}A\subseteq\operatorname{gra}B$ but $\langle x-y|u-v\rangle<0$

so, $\operatorname{gra}B = \{(x,u)\}\cup\operatorname{gra}A$: not monotone */

From (2) and proposition 20.97 (iii) /*

Proposition 20.98 (Fitzpatrick function of monotone operator)

$[A:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\operatorname{gra}A\neq\emptyset$

(iii) $(x,u)\in\mathcal{H}\times\mathcal{H}$ $\quad F_A(x,u)\leq\langle x|u\rangle \Leftrightarrow \{(x,u)\}\cup\operatorname{gra}A$: monotone */

$F_A(x,u) > \langle x|u\rangle$

So, we have:

$(x,u)\in\operatorname{gra}A \Leftrightarrow F_A(x,u)=\langle x|u\rangle$ . . . (5)

$(x,u)\notin\operatorname{gra}A \Rightarrow F_A(x,u) > \langle x|u\rangle$ // this is the interesting part

. . (13)

$S = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u)=\langle x|u\rangle\} \subseteq \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u)\leq\langle x|u\rangle\} = S'$ . . . (4)

(5) says $\operatorname{gra}A\subseteq S$ . . . (5)

contrapositive (3): $F_A(x,u)\leq\langle x|u\rangle \Rightarrow (x,u)\in\operatorname{gra}A$

$\Leftrightarrow S'\subseteq\operatorname{gra}A$

$\Rightarrow S\subseteq S'\subseteq\operatorname{gra}A$ (from (4))

. . (6)

from (5) and (6):

$\operatorname{gra}A = S = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u)=\langle x|u\rangle\}$

∎

Corollary 20.99. /* $\underline{\lim}\, x_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} x_k$ */

$[A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone, $x,u\in\mathcal{H}$, $(x_n,u_n)_{n\in\mathbb{N}}\subseteq\operatorname{gra}A$: $(x_n,u_n)\to(x,u)]$

(i) $\langle x|u\rangle \leq \underline{\lim}\,\langle x_n|u_n\rangle$

(ii) $\underline{\lim}\,\langle x_n|u_n\rangle = \langle x|u\rangle \Rightarrow (x,u)\in\operatorname{gra}A$

(iii) $\overline{\lim}\,\langle x_n|u_n\rangle \leq \langle x|u\rangle \Rightarrow \begin{cases}\langle x_n|u_n\rangle\to\langle x|u\rangle\\ (x,u)\in\operatorname{gra}A\end{cases}$

3/23/2016 8:44 AM

Proof:

(i)

Proposition 20.97 (ii): $[A:\mathcal{H}\to 2^{\mathcal{H}}$, $\operatorname{gra}A\neq\emptyset]$ $F_A\in\Gamma_0(\mathcal{H}\times\mathcal{H}) \Rightarrow F_A$: lower semicontinuous

maximally monotone — set of proper lower semicontinuous convex functions

Theorem 7.1: For a convex function all types of lower semicontinuity are equivalent. $] \Rightarrow F_A$: weakly sequentially lower semicontinuous

/* f: weakly sequentially lower semicontinuous $\Leftrightarrow \forall x\in\mathcal{H}\ \forall(x_n)_{n\in\mathbb{N}}\subseteq\mathcal{H}: x_n\rightharpoonup x\ \ f(x)\leq\underline{\lim}\, f(x_n)$ */ . . . (Rk: 1)

Proposition 20.98. $[A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone$]$ $\left[\begin{array}{l}\forall(x,u)\ \ F_A(x,u)\geq\langle x|u\rangle\\ \operatorname{gra}A = \{(x,u) \mid F_A(x,u)=\langle x|u\rangle\}\end{array}\right.$ . . . (4)

now, $\langle x|u\rangle \leq F_A(x,u) \leq \underline{\lim}\, F(x_n,u_n)$ // from (Rk: 1) /* $(x_n,u_n)\to(x,u)$ */

$= \underline{\lim}\,\langle x_n|u_n\rangle$ /* $(x_n,u_n)\in\operatorname{gra}A \Leftrightarrow F_A(x_n,u_n) = \langle x_n|u_n\rangle$ */

. . . (5) (i)

(ii) given: $\langle x|u\rangle = \underline{\lim}\,\langle x_n|u_n\rangle$, so the inequality in (5) collapses, and we have:

$F_A(x,u)=\langle x|u\rangle \Leftrightarrow (x,u)\in\operatorname{gra}A$ /* (4) */

(iii) given: $\overline{\lim}\,\langle x_n|u_n\rangle \leq \langle x|u\rangle$ . . . (6)

(5),(6) $\Rightarrow$

$\overline{\lim}\,\langle x_n|u_n\rangle \leq \langle x|u\rangle \leq \underline{\lim}\,\langle x_n,u_n\rangle \leq \overline{\lim}\,\langle x_n,u_n\rangle \Rightarrow$ collapse

by assn

$\Rightarrow \overline{\lim}\,\langle x_n|u_n\rangle = \underline{\lim}\,\langle x_n|u_n\rangle = \lim\langle x_n|u_n\rangle = \langle x|u\rangle \Rightarrow (x,u)\in\operatorname{gra}A$ ✓ [from (i)]

$\Leftrightarrow \langle x_n|u_n\rangle \to \langle x|u\rangle$ ✓ ∎

3/23/2016 1:12 PM

Theorem 20.93. (Extending a monotone operator to maximally monotone operator */

$[A:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\operatorname{gra}A\neq\emptyset$

$G = \operatorname{pav}(F_A, F_A^{*T})$

$B:\mathcal{H}\to 2^{\mathcal{H}}$, $\operatorname{gra}B = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid G(x,u)=\langle x|u\rangle\}]$

B: maximally monotone extension of A

Proof Blueprint: This proof has two parts. First, we show that A : maximally monotone, and then we show that $\operatorname{gra}A\subseteq\operatorname{gra}B$. To prove the first part we use Corollary 20.99:

$$[F\in\Gamma_0(\mathcal{H}\times\mathcal{H}), \text{autoconjugate}; \operatorname{gra}A = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F(x,u)=\langle x|u\rangle\}]$$
$$A: \text{maximally monotone}$$

To this goal, we show that $G := \operatorname{pav}(F_A, F_A^{*T})\in\Gamma_0(\mathcal{H}\times\mathcal{H})$, autoconjugate, then apply Corollary 20.99. Then to prove $\operatorname{gra}A\subseteq\operatorname{gra}B$, we use proximal properties of both Fitzpatrick and proximal average function, and show that

$$(\forall(x,u)\in\operatorname{gra}A)\quad \operatorname{Prox}_G(x+u,x+u) = (x,u) \Leftrightarrow (x,u)\in\operatorname{gra}B$$
* using Proposition 16.52 *

Proof:

Proposition 20.97 (ii): Fitzpatrick function of a monotone operator is Γ_0 (ccp)

i.e. $F_A\in\Gamma_0(\mathcal{H}\times\mathcal{H})$ . . . (1)

$\Rightarrow F_A^*\in\Gamma_0(\mathcal{H}\times\mathcal{H})$ /* Fenchel-Moreau corollary 13.33 */

$\Rightarrow F_A^{*T} \in \Gamma_0(\mathcal{H}\times\mathcal{H})$ /* $\square^T(x,y) = \square(y,x)$ so convexity stays the same

. . . (2) as only the variables are being reindexed ! */

From (1), (2):

$G = \mathrm{pav}(F_A, F_A^{*T}) \in \Gamma_0(\mathcal{H}\times\mathcal{H})$

$G^* = \mathrm{pav}(F_A, F_A^{*T})^* = \mathrm{pav}(F_A^*, F_A^{*T*})$

/* Using Corollary 14.8 (i), (ii) : $[f, g \in \Gamma_0(\mathcal{H})]$ $\mathrm{pav}(f,g) \in \Gamma_0(\mathcal{H})$ $(\mathrm{pav}(f,g))^* = \mathrm{pav}(f^*, g^*)$ */

$= \mathrm{pav}(F_A^*, F_A^{T**})$ /* $F_A \in \Gamma_0(\mathcal{H}) \Rightarrow F_A^{**} = F_A$ (Fenchel Moreau corollary)

Proposition 13.30. $F \in \Gamma(\mathcal{H}\times\mathcal{H}) \Rightarrow F^{*T} = F^{T*}$

$\because F_A \in \Gamma_0(\mathcal{H}\times\mathcal{H}) \subseteq \Gamma(\mathcal{H}\times\mathcal{H}) \Rightarrow F_A^{*T} = F_A^{T*}$ */

$= \mathrm{pav}(F_A^*, F_A^T)$ /* $F_A \in \Gamma_0(\mathcal{H}\times\mathcal{H})$

$\Rightarrow F_A^T \in \Gamma_0(\mathcal{H}\times\mathcal{H}) \Rightarrow F_A^{T**} = F_A^T$ /* Fenchel Moreau corollary */

$= \mathrm{pav}(F_A^T, F_A^*)$ /* $\mathrm{pav}(f,g) = \mathrm{pav}(g,f)$ */

$= \mathrm{pav}(F_A^T, F_A^{*TT})$ /* transpositioning a bivariate function twice brings back the original function */

$= \mathrm{pav}(\boxed{F_A}^T, \boxed{F_A^{*T}}^T)$

$= \underbrace{\mathrm{pav}(F_A, F_A^{*T})^T}_{G}$ /* Proposition 14.10. $[F, G \in \Gamma_0(\mathcal{H}\times\mathcal{H})]$ $(\mathrm{pav}(F,G))^T = \mathrm{pav}(F^T, G^T)$ */

$= G^T$

$\therefore G^* = G^T$ /* A bivariate function F is autoconjugate if $F^* = F^T$ (transposition) */

$\leftrightarrow G$: autoconjugate . . . (3)

/* Corollary 20.39.

$[F \in \Gamma_0(\mathcal{H}\times\mathcal{H})$, autoconjugate

A: gra $A = \{(x,u) \in \mathcal{H}\times\mathcal{H} \mid F(x,u) = \langle x, u\rangle\}]$

A: maximally monotone */

given : gra $B = \{(x,u) \in \mathcal{H}\times\mathcal{H} : G(x,u) = \langle x|u\rangle\}$ . . . (4)

(3), (4), corollary 20.39 $\Rightarrow$

B: maximally monotone . . . (5)

define $L : \mathcal{H}\times\mathcal{H} \to \mathcal{H}\times\mathcal{H} : (x,u) \mapsto (u,x)$ // Another way of putting it is $L = \mathrm{id}^T : (x,y) \mapsto \mathrm{id}^T(x,y) = (y,x)$

$(x,u) \in \mathrm{gra}\, A$

$\Rightarrow (x,u) = \mathrm{Prox}_{F_A}(x+u, x+u)$

. . . (6)

/* Proposition 20.40 (Fitzpatrick function of maximal operator) */

$[A : \mathcal{H} \to 2^{\mathcal{H}}$, monotone, gra $A \neq \emptyset$

(AUGCHECK)

(iii) $(x,u) \in \mathrm{gra}\, A \Leftrightarrow (x,u) = \mathrm{Prox}_{\varphi_A}(x+u, x+u)$ /* Here (and-valued) pair (x,u) is a maximal monotone operator is related to the resolvent map of the Fitzpatrick function at (input-valued, input-valued) */

now, $\forall (x,u) \in \mathrm{gra}\, A$

$\mathrm{Prox}_G(x+u, x+u)$ /* $G = \mathrm{pav}(F_A, F_A^{*T})$ */

$= \mathrm{Prox}_{\mathrm{pav}(F_A, F_A^{*T})}(x+u, x+u)$

$= \frac{1}{2}\mathrm{Prox}_{F_A}(x+u, x+u) + \frac{1}{2}\mathrm{Prox}_{F_A^{*T}}(x+u, x+u)$

/* Corollary 14.8. (ii) $[f, g \in \Gamma_0(\mathcal{H})$, $\varphi = \frac{1}{2}\|\cdot\|^2]$ $\mathrm{Prox}_{\mathrm{pav}(f,g)} = \frac{1}{2}\mathrm{Prox}_f + \frac{1}{2}\mathrm{Prox}_g$ */

$= \frac{1}{2}\mathrm{Prox}_{F_A}(x+u, x+u) + \frac{1}{2}(\mathrm{Id} - L\,\mathrm{Prox}_{F_A})(x+u, x+u)$

/* Proposition 16.53. $[L : \mathcal{H}\times\mathcal{H} \to \mathcal{H}\times\mathcal{H} : (x,u) \mapsto (u,x)$, $F \in \Gamma_0(\mathcal{H})]$ $\mathrm{Prox}_{F^{*T}} = \mathrm{Id} - L\,\mathrm{Prox}_F$ */

$= \frac{1}{2}\overbrace{\mathrm{Prox}_{F_A}(x+u, x+u)}^{(x,u)}$ /* From (6) */

$+ \frac{1}{2}(x+u, x+u) - \frac{1}{2}L\,\underbrace{\mathrm{Prox}_{F_A}(x+u, x+u)}_{=(x,u)}$ /* From (6) */

$= \frac{1}{2}(x,u) + \frac{1}{2}(x+u, x+u) - \frac{1}{2}\underbrace{L(x,u)}_{(u,x)}$

$= \frac{1}{2}(x+x+u-u,\ u+x+u-x)$

$= (x,u)$

$\Leftrightarrow \mathrm{Prox}_G(x+u, x+u) = (x,u)$

$\leftrightarrow G(x,u) = \langle x|u\rangle$

/* Proposition 16.52. $[G$: autoconjugate $\in \Gamma_0(\mathcal{H}\times\mathcal{H})]$ $(x,u) = \mathrm{Prox}_G(x+u, x+u) \Leftrightarrow G(x,u) = \langle x|u\rangle$ */

$\leftrightarrow (x,u) \in \mathrm{gra}\, B$

/* By definition of B */

. . . (6)

$\therefore \forall_{(x,u)\in \mathrm{gra}\, A}\ (x,u) \in \mathrm{gra}\, B$

$\leftrightarrow \mathrm{gra}\, A \subseteq \mathrm{gra}\, B$ . . . (7)

maximally monotone (from (5))

$\therefore B$: maximally monotone extension of A ∎

Part 3

Tuesday, August 30, 2016 9:18 PM

8/31/2016 11:41 AM

Theorem 20.38-

[$F: \mathcal{H}\times\mathcal{H} \to]-\infty,+\infty]$, convex function

F^*: proper, $F^* \geq \langle\cdot|\cdot\rangle \Leftrightarrow F^*(a,b) \geq \langle a|b\rangle$

$A: \mathcal{H} \rightrightarrows \mathcal{H}$, gra $A = \{(x,u)\in\mathcal{H}\times\mathcal{H} : F^*(x,u) = \langle x|u\rangle\}$]

i) A: monotone

ii) $F \geq \langle\cdot|\cdot\rangle \Rightarrow A$: maximally monotone.

Proof:

$\forall_{(x,u),(y,v)\in \text{gra}\,A}$

$$\langle x|u\rangle + \langle y|v\rangle = F^*(u,x) + F^*(v,y)$$

$$\geq 2F^*\left(\tfrac{1}{2}(u,x)+\tfrac{1}{2}(v,y)\right)$$

$$= 2F^*\left(\tfrac{1}{2}\begin{bmatrix}u\\x\end{bmatrix}+\tfrac{1}{2}\begin{bmatrix}v\\y\end{bmatrix}\right)$$

$$= 2F^*\left(\tfrac{1}{2}\begin{bmatrix}u+v\\x+y\end{bmatrix}\right)$$

$$= 2F^*\left(\tfrac{1}{2}(u+v,x+y)\right)$$

$$\geq 2\cdot\left(\tfrac{1}{2}\langle u+v|x+y\rangle\right) \quad /\!/ \because F^* \geq \langle\cdot|\cdot\rangle \; //$$

/* conjugate function is always convex

$\Rightarrow F^*(\cdot,\cdot)$: convex

$\Rightarrow F^*\left(\theta\begin{bmatrix}x\\x\end{bmatrix}+(1-\theta)\begin{bmatrix}y\\y\end{bmatrix}\right) \leq \theta F^*\left(\begin{bmatrix}u\\x\end{bmatrix}\right)+(1-\theta)F^*\left(\begin{bmatrix}v\\y\end{bmatrix}\right)$

$\theta \in [0,1]$

set $\theta := \frac{1}{2} \Rightarrow F^*\left(\frac{1}{2}(u,x)+\frac{1}{2}(v,y)\right) \leq \frac{1}{2}F^*(u,x)+\frac{1}{2}F^*(v,y)$

$F^*(u,x)+F^*(v,y) \geq 2F^*\left(\frac{1}{2}(u,x)+\frac{1}{2}(v,y)\right)$ */

$$\Rightarrow \langle x|u\rangle + \langle y|v\rangle \geq \tfrac{1}{2}\left(\langle u|x\rangle + \langle v|y\rangle + \langle v|x\rangle + \langle u|y\rangle\right)$$

$$\Leftrightarrow \tfrac{1}{2}\langle u|x\rangle + \tfrac{1}{2}\langle v|y\rangle \geq \tfrac{1}{2}\langle v|x\rangle + \tfrac{1}{2}\langle u|y\rangle$$

$$\Leftrightarrow \langle u-v|x\rangle + \langle u-v|y\rangle \geq 0$$

$$\Leftrightarrow \langle u-v|x-y\rangle = \langle x-y|u-v\rangle \geq 0$$

$\therefore$ So, we have shown $\forall_{(x,u)}\left((x,u)\in\text{gra}\,A \Rightarrow \forall_{(y,v)\in\text{gra}\,A} \langle x-y|u-v\rangle \geq 0\right)$

$\Leftrightarrow A$: monotone ∎

ii)

Proof blueprint: In (i), we have already proven that A is monotone, in this part we want to prove that it is maximal too.

Given: $\forall(x,u)\in \text{gra}A \;.\; \langle x-x \mid w-u\rangle \geq 0$

Goal: $(z,w)\in \text{gra}A$

So the given is true for any $(x,u)\in\text{gra}A$, so the antecedent would still be true for a specially designed $(x,u)\in\text{gra}A$.

To do this, we set a very specific $(x,u) := (y,v)\in\text{gra}A : F^*(x,y) = \langle x|y\rangle$, and using this we show that $(z,w)\in\text{gra}A$. To do this we first prove that, $F^* \geq \langle\cdot|\cdot\rangle$. Then we proceed according to the blueprint.

given F^*: proper $\overset{\text{def}}{\Rightarrow}$ $-\infty \notin S^*(\mathcal{H})$ $\Rightarrow$ S: posses a continuous affine minorant $\overset{\text{def}}{\Leftrightarrow}$ $\exists_{u\in\mathcal{H},\,\eta\in\mathbb{R}}$ $S \geq \overbrace{\langle u|\cdot\rangle + \eta}^{\text{continuous affine term}}$

F^*: convex /tell the time?/

/*

$\therefore$ $0 \in \operatorname{ri}(\operatorname{dom} F^* - \operatorname{dom} G)$ ☑

so we can apply the corollary !

which is:

$\exists_{(x,y)\in\mathcal{H}\times\mathcal{H}} \; F^*(x,y) + G(\underbrace{-L(x,y)}_{(y,-v)}) \leq 0 \Leftrightarrow F^*(x,y) + G(y,v) \leq 0$

$(x,y)\in\mathcal{H}\times\mathcal{H}$ $\quad F^*(x,y) + G(-L(x,y)) \leq 0$

now given : $\langle y|v\rangle \leq F^*$

$\underbrace{\langle v|y\rangle}_{\langle y|v\rangle} + G(y,v) \leq F^*(v,y) + G(y,v) \leq 0$

$\exists_{(v,y)\in\mathcal{H}\times\mathcal{H}} \; G(y,v) + \langle y|v\rangle \geq 0 \Rightarrow \exists_{(v,y)\in\mathcal{H}\times\mathcal{H}} \; 0 \leq G(y,v) + \langle y|v\rangle \leq F^*(v,y) + G(y,v) \leq 0$

$\Leftrightarrow \exists_{(v,y)\in\mathcal{H}\times\mathcal{H}} \; \boxed{G(y,v) + \langle y|v\rangle = 0 \;\wedge\; F^*(v,y) + G(y,v) = 0}$ $\quad$ (20.29)

$G(y,v) = -\langle y|v\rangle \qquad F^*(v,y) = -G(y,v) = \langle y|v\rangle$

$\therefore \exists_{(v,y)\in\mathcal{H}\times\mathcal{H}} \; F^*(v,y) = \langle y|v\rangle$ $\quad$ /* recall : $\operatorname{gra} A = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F^*(u,x) = \langle x|u\rangle\}$

$\Downarrow$ $\quad$ so. $(y,v) \in \operatorname{gra} A$ */

$(y,v)\in\operatorname{gra} A$ $\quad$ /* note that this is coming from the definition of $\operatorname{gra} A$ */

From (eq:antecedent prop 20.38) $\forall_{(z,w)\in\mathcal{H}\times\mathcal{H}} \; \forall_{(x,u)} \; \big((x,u)\in\operatorname{gra} A \Rightarrow \langle z-x|w-u\rangle \geq 0\big)$

$\langle z-y|w-v\rangle \geq 0$ . . . (20.29)

from (20.29), (20.29) and (eq:Lemma 20.37 (iii)) :

$(z,w) = (y,v)$

but we have shown that : $(z,w)\in\operatorname{gra} A$ $\quad\therefore$ (goal (iii)) reached

$\therefore$ A : maximally monotone.

■

9/1/2016 9:12 AM

Proposition 20.43.

[$A : \mathcal{H} \to 2^{\mathcal{H}}$, monotone : $\operatorname{gra} A \neq \emptyset$

$(x,u)\in\mathcal{H}\times\mathcal{H}$]

(i) $(x,u)\in\operatorname{gra} A \Rightarrow F_A(x,u) = \langle x|u\rangle$

(ii) $F_A = (\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle) \in \Gamma_0(\mathcal{H}\times\mathcal{H})$

(iii) $F_A(x,u) \leq \langle x|u\rangle \Leftrightarrow \{(x,u)\}\cup\operatorname{gra} A$: monotone

(iv) $F_A(x,u) \leq F_A^*(u,x)$

(v) $(x,u)\in\operatorname{gra} A \Rightarrow F_A^*(u,x) = \langle x|u\rangle$

(vi) $F_A(x,u) = F_{A^{-1}}(u,x)$

(vii) $\gamma\in\mathbb{R}_{++} \Rightarrow F_{\gamma A}(x,u) = \gamma F_A(x, u/\gamma)$

(viii) $(x,u)\in\operatorname{gra} A \Rightarrow (x,u) = \operatorname{prox}_{F_A}(x+u, x+u)$

Proof:

(i)

A : monotone $\overset{\text{def}}{\Leftrightarrow} \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \; \big((x,u)\in\operatorname{gra} A \Rightarrow \forall_{\substack{(y,v)\\ \in\operatorname{gra} A}} \langle x-y|u-v\rangle \geq 0\big)$

$\Leftrightarrow \inf_{(y,v)\in\operatorname{gra} A} \langle x-y|u-v\rangle \geq 0$

clearly $(y,v) := (x-u)$ will yield the infimum 0.

$\Rightarrow \inf_{(y,v)\in\operatorname{gra} A} \langle x-y|u-v\rangle = 0$ . . . (eq:1)

now. $F_A(x,u) \overset{\text{def}}{=} \langle x|u\rangle - \inf_{(y,v)\in\operatorname{gra} A} \langle x-y|u-v\rangle$

$= \langle x|u\rangle$ ☺

(ii)

By definition,

$F_A(x,u) \overset{\text{def}}{=} \sup_{(y,v)\in\operatorname{gra} A} \big(\langle y|v\rangle + \langle x|v\rangle - \langle y|v\rangle\big)$

now consider

$\big(\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle\big)^*(u,x)$ $\quad$ /* $f^*(u) = \sup_{x\in\mathcal{H}} \big(-f(x) + \langle x|u\rangle\big)$ */

$= \sup_{(v,y)\in\mathcal{H}\times\mathcal{H}} \Big(-(\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle)(v,y) + \underbrace{\langle (v,y)|(u,x)\rangle}_{\left\langle \begin{bmatrix} y \\ v \end{bmatrix} \middle| \begin{bmatrix} u \\ x \end{bmatrix} \right\rangle = \langle y|u\rangle + \langle v|x\rangle}\Big)$ $\quad$ /* e.g., in finite dimension: $\langle (y,v)|(x,u)\rangle = \begin{bmatrix} y_1 \\ \vdots \\ y_m \\ v_1 \\ \vdots \\ v_n \end{bmatrix}^T \begin{bmatrix} x_1 \\ \vdots \\ x_m \\ u_1 \\ \vdots \\ u_n \end{bmatrix} = (y_1x_1 + \dots + y_mx_m) + (v_1u_1 + \dots + v_nu_n) = y^Tx + u^Tv = \langle y|x\rangle + \langle u|v\rangle$ */

. . . (note: 1)

$= \sup_{(v,y)\in\mathcal{H}\times\mathcal{H}} \Big(-\iota_{\operatorname{gra} A^{-1}}(v,y) - \langle v|y\rangle + \langle y|u\rangle + \langle v|x\rangle\Big)$

$= -\inf_{(v,y)\in\mathcal{H}\times\mathcal{H}} \Big(\iota_{\operatorname{gra} A^{-1}}(y,v) - \big(\langle y|u\rangle + \langle v|x\rangle - \langle v|y\rangle\big)\Big)$ $\quad$ /* $\sup(\cdot) = -\inf(-\cdot)$ */

$= -\inf_{(v,y)\in\operatorname{gra} A^{-1}} \Big(-\big(\langle y|u\rangle + \langle v|x\rangle - \langle v|y\rangle\big)\Big)$

$= \sup_{\substack{(v,y)\in\operatorname{gra} A^{-1}\\ \Leftrightarrow (y,v)\in\operatorname{gra} A}} \Big(\langle y|u\rangle + \langle v|x\rangle - \langle v|y\rangle\Big)$

$= F_A(x,u)$

$\therefore F_A = (\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle)^* \in \Gamma(\mathcal{H}\times\mathcal{H})$ /* any conjugate function is lower-semicontinuous and convex */ (yes, any)

. . . (eq: 2)

now, we have to prove properness, i.e., $\operatorname{dom} F_A \neq \emptyset$, $-\infty \notin F_A(\mathcal{H})$

now (in (i)) : for A : monotone $F_A(x,u) = \langle x|u\rangle$ when $(x,u)\in\operatorname{gra} A$.

now $\operatorname{gra} A \neq \emptyset$ by given, so $\exists (\tilde{x},\tilde{u})\in\operatorname{gra} A \Rightarrow F_A(\tilde{x},\tilde{u}) = \langle\tilde{x}|\tilde{u}\rangle = \text{finite} \Rightarrow \operatorname{dom} F_A \neq \infty$ . . . (eq: 3)

and similarly, $-\infty \notin F_A(\operatorname{dom} F_A)$, and when $(x,u)\in\mathcal{H}\times\mathcal{H}\setminus\operatorname{dom} F_A$ we have $F_A(x,u) = +\infty \neq -\infty$, so combining both we have $-\infty \notin F_A(\mathcal{H}\times\mathcal{H})$. So from this and eq: 3 we have F_A is proper.

F_A : proper, lower-semicontinuous and convex $\Leftrightarrow F_A \in \Gamma_0(\mathcal{H}\times\mathcal{H})$

☺

(iii)

$F_A(x,u) \overset{\text{def}}{=} \langle x|u\rangle - \inf_{(y,v)\in\operatorname{gra} A} \langle x-y|u-v\rangle$

$\Leftrightarrow F_A(x,u) - \langle x|u\rangle = -\underbrace{\inf_{(y,v)\in\operatorname{gra} A} \langle x-y|u-v\rangle}_{\geq 0 \;\Leftrightarrow\; \{(x,u)\}\cup\operatorname{gra} A \text{ : monotone}}$

$\Leftrightarrow \forall_{(\tilde{x},\tilde{u})\in\mathcal{H}} \Big((\tilde{x},\tilde{u})\in\operatorname{gra}(\{(x,u)\}\cup\operatorname{gra} A)\Big) \Rightarrow \forall_{(y,v)\in\operatorname{gra}(\{(x,u)\}\cup\operatorname{gra} A)} \langle\tilde{x}-y|\tilde{u}-v\rangle \geq 0$ */

$\Leftrightarrow$

$F_A(x,u) - \langle x|u\rangle \leq 0 \Leftrightarrow \{(x,u)\}\cup\operatorname{gra} A$: monotone

(iv)

$F_A(x,u) \overset{\text{def}}{=} \sup \big(\langle y|u\rangle + \langle x|v\rangle - \langle y|v\rangle\big)$

$\leftrightarrow$

$F_A(x,u) - \langle x|u\rangle \leq 0 \Leftrightarrow \{(x,u)\} \cup \operatorname{gra} A$: monotone

(iv)

$$F_A(x,u) \overset{\text{def}}{=} \sup_{(y,v)\in \operatorname{gra} A} \left(\langle y|u\rangle + \langle x|v\rangle - \langle y|v\rangle \right)$$

/* as, $(y,v)\in \operatorname{gra} A \Rightarrow F_A(y,v) = \langle y|v\rangle$ from (6). */

$$= \sup_{(y,v)\in \operatorname{gra} A} \left(\langle y|u\rangle + \langle x|v\rangle - F_A(y,v) \right)$$

$$\leq \sup_{(y,v)\in \mathcal{H}\times\mathcal{H}} \left(\langle y|u\rangle + \langle x|v\rangle - F_A(y,v) \right)$$ /* using the fact that optimizing over a larger set gives better objective value */

$$= \sup_{(y,v)\in \mathcal{H}\times\mathcal{H}} \left(\langle (y,v) \mid (u,x)\rangle - F_A(y,v) \right)$$ /* using (note: 1) */

$$= F_A^*(u,x)$$

$$\therefore \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \quad F_A(x,u) = F_A^*(u,x)$$

(v)

By (ii):

$$\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \quad F_A(x,u) = \left(\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle\right)^*(x,u)$$

$$\Rightarrow F_A^* = \left(\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle\right)^{**}$$

$$\leq \iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle$$ /* Biconjugate is always a lower bound (minorant) of the original function */

In (iv) we have shown:

$$\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \quad F_A(x,u) \leq F_A^*(u,x) \leq \left(\iota_{\operatorname{gra} A^{-1}} + \langle\cdot|\cdot\rangle\right)(u,x) = \iota_{\operatorname{gra} A^{-1}}(u,x) + \langle u|x\rangle$$

$$\iota_{\operatorname{gra} A^{-1}}(u,x) = \iota_{\operatorname{gra} A}(x,u)$$

now $(x,u)\in\operatorname{gra} A \Rightarrow$
- $\iota_{\operatorname{gra} A}(x,u) = 0$
- $F_A(x,u) = \langle x|u\rangle$ (from 6)

$$\forall_{(x,u)\in\operatorname{gra} A} \quad F_A(x,u) \leq F_A^*(u,x) \leq \langle u|x\rangle = F_A(x,u)$$

$$\leftrightarrow \forall_{(x,u)\in\operatorname{gra} A} \quad F_A(x,u) = F_A^*(u,x)$$ ☺

(vi), (vii): direct application of the definition: $F_A(x,u) = \sup_{(y,v)\in\operatorname{gra} A} \left(\langle y|u\rangle + \langle x|v\rangle - \langle y|v\rangle \right)$

(viii)

In (i): $\forall_{(x,u)\in\operatorname{gra} A} \quad F_A(x,u) = \langle x|u\rangle$

In (v): $\forall_{(x,u)\in\operatorname{gra} A} \quad F_A^*(u,x) = \langle x|u\rangle$

$$\forall_{(x,u)\in\operatorname{gra} A} \quad \underbrace{F_A(x,u)}_{a} + \underbrace{F_A^*(u,x)}_{b} = 2\langle x|u\rangle = \langle \underbrace{(x,u)}_{a} | \underbrace{(u,x)}_{b}\rangle$$ /* clever rewriting using (note: 1) */

/* recall Theorem 14.2.3.:

$[f\in\Gamma_0(\mathcal{H}),\ x\in\mathcal{H},\ u\in\mathcal{H}]\quad (x,u)\in\operatorname{gra}\partial f \Leftrightarrow f(x) + f^*(u) = \langle x|u\rangle$ */

now in (ii) we have shown $F_A\in\Gamma_0(\mathcal{H}\times\mathcal{H})$

$$F_A(a) + F_A(b) = \langle a|b\rangle$$

$$\leftrightarrow (a,b)\in\operatorname{gra}\partial F_A$$

$$\leftrightarrow \partial F_A(a) \ni b, \quad a = (x,u),\ b = (u,x)$$

$$\leftrightarrow \partial F_A(x,u) \ni (u,x)$$

$$\leftrightarrow \underbrace{(x,u) + \partial F_A(x,u)}_{(\mathrm{Id}+\partial F_A)(x,u)} \ni (u,x) + (x,u) = \begin{bmatrix} u \\ x \end{bmatrix} + \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} u+x \\ x+u \end{bmatrix} = (u+x,\ x+u)$$

$$\leftrightarrow (\mathrm{Id}+\partial F_A)(x,u) \ni (x+u,\ x+u)$$

$$\leftrightarrow (x,u) \in (\mathrm{Id}+\partial F_A)^{-1}(x+u,\ x+u)$$

$$(\mathrm{Id}+\partial F_A)^{-1} = \operatorname{Prox}_{F_A}$$ /* $\operatorname{Prox}_f = (\mathrm{Id}+\partial f)^{-1}$ */

$$\leftrightarrow (x,u) \in \operatorname{Prox}_{F_A}(x+u,\ x+u)$$

$$\leftrightarrow (x,u) = \operatorname{Prox}_{F_A}(x+u,\ x+u)$$ /* because Prox is a function */

∎

Proposition 20.50.

[C,D: closed affine subspaces of $\mathcal{H}$: $D-D = (C-C)^\perp$

A: maximally monotone

$(x_n,u_n)\subseteq\operatorname{gra} A$

$(x,u)\in\mathcal{H}\times\mathcal{H}$]

$$\left.\begin{array}{l} x_n\to x \\ u_n\to u \\ x_n - P_C x_n\to 0 \\ u_n - P_D u_n\to 0 \end{array}\right] \Rightarrow \left[\begin{array}{l} (x,u)\in (C\times D)\cap\operatorname{gra} A \\ \langle x_n|u_n\rangle \to \langle x|u\rangle \end{array}\right.$$

9/3/2016 10:38 AM

Proof: // This proof is a beast, though the proof blueprint is relatively simple, uses several facts about affine space and some ofthe previous results.

$V := C-C$: linear subspace parallel to C

given, $x_n - P_C x_n \to 0 \leftrightarrow P_C x_n - x_n \to 0$

$$\leftrightarrow P_C x_n \to x_n \to x \quad \text{// given } x_n\to x$$

$$\therefore P_C x_n \to x \quad (\text{eq.:1})$$

given, C: closed affine subspaces of $\mathcal{H}$

$\Rightarrow C$: closed convex set of $\mathcal{H}$ /* affine sets are convex */

$\leftarrow C$: weakly sequentially closed set of $\mathcal{H}$ /* for a convex set, all notions of closedness collapses */

recall the definition of a weakly sequentially closed set via sequences: C: weakly sequentially closed $\overset{\text{def}}{\leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}:\ S\subset,\ x_n\to x} \quad x\in C \quad \ldots (\text{eq.:2})$

B.S. $(P_C x_n)_{n\in\mathbb{N}}\in C$ by definition of $P_C(\cdot)$, using (eq.:1) and (eq.:2) we have: $x\in C \Rightarrow C = x + \overset{\text{convex}}{\widetilde{C-C}}$ // C: affine subspace $\Rightarrow \forall_{x\in C} \quad C = x + \underbrace{V}_{C-C}$: page 1: Bauschke */

using similar logic, $\underbrace{u\in D}_{(\text{eq.:5})} \Rightarrow D = u + D - D$

$$= u + (C-C)^\perp \quad \text{// given: } D-D = (C-C)^\perp \text{ */}$$

$$= u + V^\perp$$

$$\therefore \text{so, } \left.\begin{array}{l} C = x + V \\ D = u + V^\perp \end{array}\right\} (\text{eq.:7})$$

$$\because P_C w = P_{x+V}\, w = x + P_V(w-x) \quad \text{// Proposition 3.17: [} C \text{: nonempty closed convex set of } \mathcal{H};\ x\in\mathcal{H} \text{] } P_{x+C}\, z = x + P_C(z-x)$$

affine

$$= x + P_V(w-x) \quad \text{// } V \text{: linear subspace} \Rightarrow P_V(0) = 0$$

V: affine subspace too $\Rightarrow P_V$: affine operator $\Leftrightarrow \widetilde{x} \mapsto P_V\widetilde{x} - P_V(0)$: linear

$\Leftrightarrow x \mapsto P_V\widetilde{x}$: linear $\Leftrightarrow P_V$: linear */

$$= x + P_V w - P_V(x)$$

$$= P_V w + \underbrace{(\mathrm{Id} - P_V)(x)}_{P_{V^\perp}x} \quad \text{// Corollary 3.22: (v): [} V \text{: closed linear subspace of } \mathcal{H};\ x\in\mathcal{H} \text{] } P_{V^\perp} = \mathrm{Id} - P_V \text{ */}$$

$$\forall_{x\in\mathcal{H}} \quad P_{V^\perp}x = x - P_V x \Leftrightarrow x = P_V x + P_{V^\perp}x \quad \ldots (\text{eq.:4})$$

$$\Rightarrow \left.\begin{array}{l} P_C w = P_V w + P_{V^\perp}x \\ \text{similarly,} \\ P_D w = P_{V^\perp}u + P_V u \end{array}\right\} \forall_{w\in\mathcal{H}} \quad \ldots (\text{eq.:5})$$

/* Proposition 4.11.

[C: closed affine subspace of $\mathcal{H}$] P_C: weakly continuous */

using this:

P_C, P_D: weakly continuous // $T: \mathcal{H}\to\mathcal{H}$: weakly continuous $\overset{\text{def}}{\Leftrightarrow} \forall_{(x_n)_{n\in\mathbb{N}}:\ S\subset,\ x_n\to x} \quad Tx_n\to Tx$

$$\therefore \langle x_n|u_n\rangle = \langle P_V x_n + P_{V^\perp}x_n \mid P_V u_n + P_{V^\perp}u_n\rangle$$

$$= \langle P_V x \mid P_V u\rangle + \langle P_V x \mid P_{V^\perp}u\rangle + \langle P_{V^\perp}x \mid P_V u\rangle + \langle P_{V^\perp}x \mid P_{V^\perp}u\rangle$$ // same logic */

[C : closed affine subspace of $\mathcal{H}$] P_C : weakly continuous */

using this :

P_C, P_B : weakly continuous || $T: B \to \mathcal{H}$: weakly continuous $\overset{\text{def5}}{\Leftrightarrow} \forall_{\{x_n\}_{\mathbb{N}} \subset B,\, x_n \to x} \; Tx_n \rightharpoonup Tx$

$\therefore \langle x_n | u_n \rangle = \langle P_V x_n + P_{V^\perp} x_n | P_V u_n + P_{V^\perp} u_n \rangle$

$= \langle P_V x_n | P_V u_n \rangle + \underbrace{\langle P_{V^\perp} x_n | P_V u_n \rangle}_{0} + \overbrace{\langle P_V x_n | P_{V^\perp} u_n \rangle}^{0 \text{ /same logic */}} + \langle P_{V^\perp} x_n | P_{V^\perp} u_n \rangle$

$\underbrace{0}_{\text{/* as } P_{V^\perp} x_n \in V^\perp \text{ and } P_V u_n \in V,\ \langle P_{V^\perp} x_n | P_V u_n \rangle = 0 \text{ */}}$

$= \langle P_V x_n | \underbrace{P_V u_n}_{(Id - P_{V^\perp})} \rangle + \langle \underbrace{P_{V^\perp} x_n}_{(Id - P_V)} | P_{V^\perp} u_n \rangle$

$= \langle P_V x_n | u_n - P_{V^\perp} u_n \rangle + \langle x_n - P_V x_n | P_{V^\perp} u_n \rangle$

$\langle P_V u_n - P_V u_n | \rangle \quad \langle P_{V^\perp} x_n - P_{V^\perp} x | \rangle$ /using eq:5 */

$= \langle P_V x_n | u_n - P_V u_n \rangle + \langle P_V x_n | P_V u_n \rangle + \langle x_n - P_V x_n | P_{V^\perp} u_n \rangle + \langle P_{V^\perp} x | P_{V^\perp} u_n \rangle$

$0 \qquad \langle P_V x | P_V u \rangle \qquad 0 \qquad \langle P_{V^\perp} x | P_{V^\perp} u \rangle$

$\to \langle P_V x | P_V u \rangle + \langle P_{V^\perp} x | P_{V^\perp} u \rangle$

$= \langle x | u \rangle$

/* subproofs.

$\langle x | u \rangle = \langle P_V x + P_{V^\perp} x \,|\, P_V u + P_{V^\perp} u \rangle$ /using eq:4 */

$= \langle P_V x | P_V u \rangle + \langle P_{V^\perp} x | P_{V^\perp} u \rangle$

$+ \underbrace{\langle P_V x | P_{V^\perp} u \rangle}_{0} + \underbrace{\langle P_{V^\perp} x | P_V u \rangle}_{0}$

/* as $P_V x \in V$, $P_{V^\perp} u \in V^\perp$ */

$= \langle P_V x | P_V u \rangle + \langle P_{V^\perp} x | P_{V^\perp} u \rangle$ */

$\langle P_V x_n | u_n - P_V u_n \rangle \to 0$

subproof: $\underbrace{x_n \rightharpoonup x}_{\text{given}} \Rightarrow P_V x_n \rightharpoonup P_V x$, $\underbrace{u_n - P_V u_n \to 0}_{\text{given}}$, so $\langle P_V x_n | u_n - P_V u_n \rangle \to 0$ By using

similarly, $\langle x_n - P_V x_n | P_{V^\perp} u_n \rangle \to 0$

subproof: given $\left| \begin{array}{l} u_n \to u \\ x_n - P_V x_n \to 0 \end{array} \right| \Rightarrow P_{V^\perp} u_n \to P_V u$

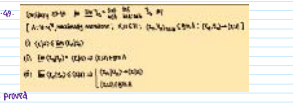
$\langle x_n - P_V x_n | P_{V^\perp} u_n \rangle \to 0$

similarly, $P_V x_n \rightharpoonup P_V x$, $P_V u \to P_V u \Rightarrow \langle P_V x_n | P_V x \rangle \to \langle P_V x | P_V u \rangle$

and $\underbrace{P_{V^\perp} x \to P_{V^\perp} x}_{\text{they are the same !}}$, $P_{V^\perp} u_n \to P_{V^\perp} u \Rightarrow \langle P_{V^\perp} x | P_{V^\perp} u_n \rangle \to \langle P_{V^\perp} x | P_{V^\perp} u \rangle$

$\therefore \underbrace{\langle x_n | u_n \rangle \to \langle x | u \rangle}_{\text{one goal proved}} \Leftrightarrow \lim \langle x_n | u_n \rangle = \langle x | u \rangle \Rightarrow \underline{\lim} \langle x_n | u_n \rangle = \langle x | u \rangle \Rightarrow (x, u) \in \operatorname{gra} A$ ... (eq: 6)

/using proposition 20-40.



So, using (eq:2), (eq:3-5) and (eq:6) :

$(x, u) \in (C \times D) \cap \operatorname{gra} A$ $\therefore$ other goal proved

■

Part 4

Sunday, September 4, 2016 1:41 PM

9/5/2016 11:36 AM

Proposition 20.51.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$]

(i) $F_A^* = (\iota_{\operatorname{gra} A^{-1}} + \langle \cdot | \cdot \rangle)^{**}$

(ii) $\operatorname{conv} \operatorname{gra} A^{-1} \subseteq \operatorname{dom} F_A^* \subseteq \overline{\operatorname{conv}} \operatorname{gra} A^{-1} \subseteq \overline{\operatorname{conv}} \operatorname{ran} A \times \overline{\operatorname{conv}} \operatorname{dom} A$

(iii) $F_A^* \geqslant \langle \cdot | \cdot \rangle$

(iv) A: maximally monotone $\Rightarrow$ $\operatorname{gra} A = \{(x,u) \in \mathcal{H} \times \mathcal{H} \mid F_A^*(u,x) = \langle x | u \rangle\}$

Proof Blueprint: (i), (ii) are just coming from definitions and previous results. In (iii) we construct a maximally monotone extenstion of A, denoted by B, then we show a vital inequality:

$$(\forall (x,u) \in \mathcal{H} \times \mathcal{H}) \quad F_A^*(u,x) \geq F_B^*(u,x) \geq F_B(x,u) \geq \langle x \mid u \rangle, \qquad (1)$$

which contains the goal (iii). In (iv) we prove, $F_B^*(u,x) \leq \langle x \mid u \rangle \Rightarrow (x,u) \in \mathbf{gra} B$ (in fact its contrapositive) and then $(x,u) \in \mathbf{gra} B \Rightarrow F_B^*(u,x) \leq \langle x \mid u \rangle$ using previously proven results. Thus we have $(x,u) \in \mathbf{gra} B \Leftrightarrow F_B^*(u,x) \leq \langle x \mid u \rangle$. Now in (1), we have shown $(\forall (x,u) \in \mathcal{H} \times \mathcal{H}) \quad F_B^*(u,x) \geq \langle x \mid u \rangle$, so combining both we have $(x,u) \in \mathbf{gra} B \Leftrightarrow F_B^*(u,x) = \langle x \mid u \rangle$. Now if A : maximally monotone, then its maximally monotone extension will be itself, *i.e.*, $B = A$. Thus we arrive at the claim.

Proof:

(i) /* Proposition 20.47.(ii): [$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$] $F_A = (\iota_{\operatorname{gra} A^{-1}} + \langle \cdot | \cdot \rangle)^*$ */

$\Rightarrow F_A^* = (\iota_{\operatorname{gra} A^{-1}} + \langle \cdot | \cdot \rangle)^{**}$ ☺

(ii) /* recall: Proposition 9.8.(iv). [$f: \mathcal{H} \to [-\infty, +\infty]$] $\operatorname{conv} \operatorname{dom} f \subseteq \operatorname{dom} \breve{f} \subseteq \overline{\operatorname{conv}} \operatorname{dom} f$ */

$\breve{f}$: largest lower semicontinuous convex function majorized by f.

Proposition 18.39. *

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

- f: has a continuous affine minorant $\Rightarrow f^{**} = \breve{f}$
  $\operatorname{dom} f^* \neq \emptyset$ (from monotone 18.18(i))
- f: doesn't have a continuous affine minorant $\Rightarrow f^{**} = -\infty$

define $G: (u,x) \mapsto (\iota_{\operatorname{gra} A^{-1}} + \langle \cdot | \cdot \rangle)(u,x)$

$$= \begin{cases} \langle \cdot | \cdot \rangle, & \text{if } (u,x) \in \operatorname{gra} A^{-1} \\ +\infty, & \text{if } (u,x) \notin \operatorname{gra} A^{-1} \end{cases}$$

also $\operatorname{dom} G = \{(u,x) \in \mathcal{H} \times \mathcal{H} \mid G(x,u) < +\infty\} = \operatorname{gra} A^{-1}$

So, $\langle \cdot | \cdot \rangle \leqslant G \Leftrightarrow G$: has a continuous affine minorant $\Rightarrow G^{**} = \breve{G} = (\iota_{\operatorname{gra} A^{-1}} + \langle \cdot | \cdot \rangle)^{**} = F_A^*$

in proposition 9.8(iv) set $f := G \Rightarrow$

$\operatorname{conv} \operatorname{dom} G \subseteq \operatorname{dom} \breve{G} \subseteq \overline{\operatorname{conv}} \operatorname{dom} G$

$\leftrightarrow \operatorname{conv} \operatorname{gra} A^{-1} \subseteq \operatorname{dom} F_A^* \subseteq \overline{\operatorname{conv}} \operatorname{gra} A^{-1} \ldots$ (eq:1)

$\left.\begin{array}{l} \operatorname{dom} A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\} \\ \operatorname{ran} A = \{u \in \mathcal{H} \mid \exists_{x \in \operatorname{dom} A} Ax \ni u\} \end{array}\right\} \Rightarrow \operatorname{gra} A \subseteq \operatorname{dom} A \times \operatorname{ran} A$

$\leftrightarrow \operatorname{gra} A^{-1} \subseteq \operatorname{ran} A \times \operatorname{dom} A$

$\Rightarrow \overline{\operatorname{conv}} \operatorname{gra} A^{-1} \subseteq \overline{\operatorname{conv}} \operatorname{ran} A \times \overline{\operatorname{conv}} \operatorname{dom} A \ldots$ (eq:2)

From (eq:1) and (eq:2) we have:

$\operatorname{conv} \operatorname{gra} A^{-1} \subseteq \operatorname{dom} F_A^* \subseteq \overline{\operatorname{conv}} \operatorname{gra} A^{-1} \subseteq \overline{\operatorname{conv}} \operatorname{ran} A \times \overline{\operatorname{conv}} \operatorname{dom} A$ ☺

(iii)

Assume B: maximally monotone extension of A //By the virtue of Theorem 20.21. which says that // a monotone operator has a maximally monotone //extension

$\operatorname{gra} A \subseteq \operatorname{gra} B$

now, $F_A(x,u) = \sup_{(y,v) \in \operatorname{gra} A} (\langle y | u \rangle + \langle x | v \rangle - \langle y | v \rangle)$

$\forall_{(x,u) \in \mathcal{H} \times \mathcal{H}} \quad F_B(x,u) = \sup_{(y,v) \in \operatorname{gra} B} (\langle y | u \rangle + \langle x | v \rangle - \langle y | v \rangle) \geqslant \sup_{(y,v) \in \operatorname{gra} A} (\langle y | u \rangle + \langle x | v \rangle - \langle y | v \rangle)$

$= F_A(x,u)$

$\leftrightarrow F_B \geqslant F_A$

$\Rightarrow F_B^* \leqslant F_A^*$ // taking conjugate flips the relationship between functions ; Proposition 13.14.(ii)

$\ldots$ (eq:3)

/*Proposition 20.47.(iv): [$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$] $\forall_{(x,u) \in \mathcal{H} \times \mathcal{H}} \; F_A(x,u) \leqslant F_A^*(u,x)$

$\because B$: monotone, $\operatorname{gra} B \neq \emptyset \Rightarrow \forall_{(x,u) \in \mathcal{H} \times \mathcal{H}} \; F_B(x,u) \leqslant F_B^*(u,x)$

Proposition 20.47.(iv): [$A:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$] $\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ F_A(x,u) \leq F_A^*(u,x)$

$\because B$: monotone, $\operatorname{gra} B \neq \emptyset \Rightarrow \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ F_B(x,u) \leq F_B^*(u,x)$

$\forall_{(x,u)\in\mathcal{H}}\ F_A^*(u,x) \geq F_B^*(u,x) \geq \underbrace{F_B(x,u) \geq \langle x|u\rangle}_{\text{// Fitzpatrick function of a maximally monotone operator has the inner product as its minorant}} = \langle u|x\rangle \quad \ldots (\text{eq}:4)$ Proposition 20.48

$\Rightarrow \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ F_A^*(u,x) \geq \langle u|x\rangle$

(iv)

consider $(x,u)\notin \operatorname{gra} B$

B: maximally monotone $\Rightarrow (x,u)\in\operatorname{gra} B \leftrightarrow F_B(x,u) = \langle x|u\rangle$ // Proposition 20.48

$\leftrightarrow (x,u)\notin\operatorname{gra} B \leftrightarrow F_B(x,u) \neq \langle x|u\rangle$

But (eq:4) $\Rightarrow$ $F_B^*(u,x) \geq F_B(x,u) \geq \langle x|u\rangle$ $\Big\} \Rightarrow F_B^*(u,x) \geq F_B(x,u) > \langle x|u\rangle \quad \ldots (\text{eq}:5)$

$\therefore (x,u)\notin\operatorname{gra} B \Rightarrow F_B^*(u,x) > \langle x|u\rangle$

$\leftrightarrow F_B^*(u,x) \leq \langle x|u\rangle \Rightarrow (x,u)\in\operatorname{gra} B \quad \ldots (\text{eq}:6)$

$\Rightarrow \left(F_B^*(u,x) = \langle x|u\rangle \Rightarrow (x,u)\in\operatorname{gra} B \right) \quad \ldots (\text{eq}:6.5)$

now consider $(x,u)\in\operatorname{gra} B$, take $(y,v)\in\mathcal{H}\times\mathcal{H} \Rightarrow F_B(y,v) = \sup_{(\tilde{x},\tilde{u})\in\operatorname{gra} B} \left(\langle y|\tilde{u}\rangle + \langle \tilde{x}|v\rangle - \langle \tilde{x}|\tilde{u}\rangle \right)$

$\geq \langle y|u\rangle + \langle x|v\rangle - \langle x|u\rangle$ // now: $\langle y|u\rangle + \langle x|v\rangle = y^T u + x^T v = y^T u + v^T x = \begin{bmatrix} y^T & v^T \end{bmatrix} \begin{bmatrix} u \\ x \end{bmatrix} = \begin{bmatrix} y \\ v \end{bmatrix}^T \begin{bmatrix} u \\ x \end{bmatrix} = \langle (y,v)|(u,x)\rangle$ *

$= \langle (y,v)|(u,x)\rangle - \langle x|u\rangle$

$\Rightarrow F_B(y,v) \geq \langle (y,v)|(u,x)\rangle - \langle x|u\rangle \quad \forall_{(y,v)\in\mathcal{H}\times\mathcal{H}}$

$\leftrightarrow \langle x|u\rangle \geq \langle (y,v)|(u,x)\rangle - F_B(y,v) \quad \forall_{(y,v)\in\mathcal{H}\times\mathcal{H}}$

$\Leftrightarrow \langle x|u\rangle \geq \sup_{(y,v)\in\mathcal{H}\times\mathcal{H}} \left(\langle (y,v)|(u,x)\rangle - F_B(y,v) \right)$ /* now: $f^*(u) = \sup_{x\in\mathcal{H}} \left(\langle x|u\rangle - f^*(x) \right)$, so:

$\therefore \sup_{(y,v)\in\mathcal{H}\times\mathcal{H}} \left(\langle (y,v)|(u,x)\rangle - F_B(y,v) \right) = F_B^*(u,x)$ */

$= F_B^*(u,x)$

$\therefore (x,u)\in\operatorname{gra} B \Rightarrow F_B^*(u,x) \leq \langle x|u\rangle \quad \ldots (\text{eq}:7)$

But $F_B^* \geq \langle \cdot|\cdot \rangle$ for any monotone operator by (ii) $\Big\} \Rightarrow \left((x,u)\in\operatorname{gra} B \Rightarrow F_B^*(u,x) = \langle x|u\rangle \right) \quad \ldots (\text{eq}:7.5)$

$\therefore B$ = maximally monotone $\overset{(\text{eq}:6.5,\,7.5)}{\Rightarrow} \operatorname{gra} B = \{ (x,u)\in\mathcal{H}\times\mathcal{H} \mid F_B^*(u,x) = \langle x|u\rangle \}$

now recall that, B := maximal monotone extension of A

given in (iv): A = maximally monotone $\Rightarrow A = B$

$\therefore A$: maximally monotone $\Rightarrow \operatorname{gra} A = \{ (x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A^*(u,x) = \langle x|u\rangle \}$

■

Part 1

Friday, March 25, 2016 11:05 AM

Theorem 21.1. (Minty's theorem)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone]

A: maximally monotone $\leftrightarrow$ $\operatorname{ran}(Id+A)=\mathcal{H}$

3/25/2016 1:48 PM

Proof:

($\Leftarrow$) given $\operatorname{ran}(Id+A)=\mathcal{H}$

goal A: maximally monotone $\leftrightarrow \forall (x,u)\ \left((x,u)\in \operatorname{gra}A \leftrightarrow \forall (\tilde{y},\tilde{v})\in \operatorname{gra}A\ \ \langle x-\tilde{y} \mid u-\tilde{v}\rangle \geq 0 \right)$

A: monotone $\leftrightarrow \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} \left((x,u) \in \operatorname{gra}A \Rightarrow \forall_{(\tilde{y},\tilde{v})\in \operatorname{gra}A}\ \ \langle x-\tilde{y} \mid u-\tilde{v}\rangle \geq 0 \right)$

So, more specifically we want to show, $\underbrace{\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ \forall_{(\tilde{y},\tilde{v})\operatorname{gra}A}\ \langle x-\tilde{y} \mid u-\tilde{v}\rangle \geq 0}_{\text{given (1)}} \Rightarrow (x,u)\in \operatorname{gra}A$ (goal (0))

first note that $\forall (x,u)\in\mathcal{H}\times\mathcal{H}\ \ \tilde{x}=x+u\in\mathcal{H}$

now given,

$\operatorname{ran}(Id+A)=\mathcal{H}$ /* $\operatorname{ran}\tilde{A}=\tilde{A}(\mathcal{H})$ */

$\leftrightarrow \forall \tilde{x}\in\mathcal{H}\ \exists y\in\mathcal{H}\ \ (Id+A)y \ni \tilde{x}$

now $\tilde{x}=x+u$ $\ \exists y\in\mathcal{H}\ \ (Id+A)y \ni \tilde{x}=x+u$

$\leftrightarrow y+Ay \ni x+u$

$\leftrightarrow \exists_{v\in Ay}\ \ y+v=x+u\in(Id+A)y$

$\leftrightarrow \exists_{\substack{v\in Ay,\\ (y,v)\in\operatorname{gra}A}}\ \ y+v=x+u\in(Id+A)y$

So, for $x+u\in\mathcal{H}$ $\ \exists_{(y,v)\in\mathcal{H}\times\mathcal{H}} \left((y,v)\in\operatorname{gra}A \wedge y+v=x+u \right)$

in given (1): setting $(\tilde{y},\tilde{v}):=(y,v)\in\operatorname{gra}A \Rightarrow$

$0 \leq \langle x-y \mid u-v\rangle = \langle x-y \mid y-x\rangle$ [from (2)]

$= -\|x-y\|^2 \leq 0$

$\leftrightarrow \|x-y\|^2=0 \leftrightarrow x=y$ (3)

(2),(3) $\Rightarrow$ $u=v$ ...(4)

$\therefore (x,u)=(y,v)\in\operatorname{gra}A$ /*from (2),(3),(4)*/ // It does say something interesting $\forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ \forall_{(\tilde{y},\tilde{v})\in\operatorname{gra}A}\ \langle x-\tilde{y} \mid u-\tilde{v}\rangle \geq 0 \Rightarrow (\tilde{y},\tilde{v})=(x,y)\in\operatorname{gra}A$ (1')

$\leftrightarrow$ goal (0) achieved.

($\Leftarrow$ direction showed) ☺

($\Rightarrow$)

given A: maximally monotone $\leftrightarrow \forall (x,u)\ \left((x,u)\in\operatorname{gra}A \leftrightarrow \forall (\tilde{y},\tilde{v})\in\operatorname{gra}A\ \ \langle x-\tilde{y} \mid u-\tilde{v}\rangle \geq 0 \right)$ . . . (1)

goal: $\operatorname{ran}(Id+A)=\mathcal{H}$

/*recall maximally monotone operator's Fitzpatrick function representation:

[$A: \mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone] $\left| \begin{array}{l} \operatorname{gra}A = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u)=\langle x \mid u\rangle\} \ \therefore (x,u)\in\operatorname{gra}A \leftrightarrow F_A(x,u)=\langle x \mid u\rangle \\ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ \ F_A(x,u)\geq\langle x \mid u\rangle \end{array} \right\}$ (5)

$F_A(x,y)=\langle x \mid y\rangle - \inf_{(y,v)\in\operatorname{gra}A} \langle x-y \mid u-v\rangle$; [A: monotone] $F_A\in\Gamma_0(\mathcal{H}\times\mathcal{H})$ . . . (6)

*/

$\forall (x,u)\in\mathcal{H}\times\mathcal{H}$

$2F_A(x,u)+\|(x,u)\|^2$

$=2F_A(x,u)+\|x\|^2+\|u\|^2$ /*$\|(x,u)\|^2=\sum x_i^2+\sum u_i^2=\|u\|^2+\|x\|^2$*/

$\geq 2\langle x \mid u\rangle+\|x\|^2+\|u\|^2$

$=\|x+u\|^2\geq 0$

$\leftrightarrow \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\ \ \underbrace{F_A(x,u)}_{\in\Gamma_0(\mathcal{H}\times\mathcal{H})\ [\text{from (6)}]}+\frac{1}{2}\|(x,u)\|^2\geq 0$ // dividing both sides by 2

$\leftrightarrow F_A+\frac{1}{2}\|\cdot\|^2\geq 0,\ F_A\in\Gamma_0$ ...(7)

/* Corollary 15.17. [$f \in \Gamma_0(\mathcal{H})$; $q = \frac{1}{2}\|\cdot\|^2$; $f+q \geq 0$] $\exists\, w \in \mathcal{H}\ \forall x \in \mathcal{H}\quad f(x)+q(x) \geq q(x-w)$ */

From (7) and Corollary 15.17 we have: /* $w \in \mathcal{H} \Leftrightarrow \exists\, p \in \mathcal{H}\ \ w=-p$ */

$$\exists_{(v,-y)\in\mathcal{H}\times\mathcal{H}}\ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\quad F_A(x,u)+\frac{1}{2}\|(x,u)\|^2 \geq \frac{1}{2}\|(x,u)+(v,y)\|^2$$

$$= \frac{1}{2}\|(x,u)\|^2+\frac{1}{2}\|(v,y)\|^2+\underbrace{\langle (x,u)|(v,y)\rangle}_{\langle x|v\rangle+\langle u|y\rangle} \quad /* \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} v \\ y \end{bmatrix} = [x^T\ u^T]\begin{bmatrix} v \\ y \end{bmatrix} = x^Tv+u^Ty\ */$$

$$\Leftrightarrow F_A(x,u) \geq \underbrace{\frac{1}{2}\|(v,y)\|^2}_{\frac{1}{2}\|v\|^2+\frac{1}{2}\|y\|^2 \geq -\langle v|y\rangle}+\langle x|v\rangle+\langle u|y\rangle$$

/* $\|v+y\|^2 = \|v\|^2+\|y\|^2+2\langle v|y\rangle \geq 0$

$\Rightarrow \frac{1}{2}\|v\|^2+\frac{1}{2}\|y\|^2 \geq -\langle v|y\rangle$ */

$$\geq -\langle v|y\rangle+\langle x|v\rangle+\langle u|y\rangle \quad \ldots (8)$$

$$\therefore \exists_{(v,y)\in\mathcal{H}\times\mathcal{H}}\ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}}\quad F_A(x,u) \geq -\langle v|y\rangle+\langle x|v\rangle+\langle u|y\rangle \quad \ldots (9)$$

if $(x,u)\in \operatorname{gra} A \Leftrightarrow F_A(x,u) = \langle x|u\rangle$ [from (5)] ... (10)

(9), (10) $\Rightarrow$

$$\exists_{(v,y)\in\mathcal{H}\times\mathcal{H}}\ \forall_{(x,u)\in \operatorname{gra} A}\quad \langle x|u\rangle \geq -\langle v|y\rangle+\langle x|v\rangle+\langle u|y\rangle$$

$$\Leftrightarrow \langle x|u\rangle-\langle x|v\rangle+\langle v|y\rangle-\langle u|y\rangle \geq 0$$

$$\Leftrightarrow \langle x|u-v\rangle-\langle u-v|y\rangle = \langle x|u-v\rangle-\langle y|u-v\rangle = \langle x-y|u-v\rangle \geq 0$$

A: maximally monotone $\Rightarrow (y,v)\in \operatorname{gra} A \Leftrightarrow Ay \ni v$ [From (11)]

$\vdots$ (12)

$$\exists_{(v,y)\in \operatorname{gra} A}\ \forall_{(x,u)\in \operatorname{gra} A}\quad \langle x|u\rangle \geq \frac{1}{2}\|v\|^2+\langle x|v\rangle+\frac{1}{2}\|y\|^2+\langle y|u\rangle$$

/* set $(x,u) := (y,v) \in \operatorname{gra} A$ [from (11)] */

$$\langle y|v\rangle \geq \frac{1}{2}\|v\|^2+\langle y|v\rangle+\frac{1}{2}\|y\|^2+\langle y|v\rangle$$

$$\Rightarrow 0 \geq \frac{1}{2}\|v\|^2+\frac{1}{2}\|y\|^2+\langle y|v\rangle$$

$$\Leftrightarrow 0 \geq \|v\|^2+\|y\|^2+2\langle y|v\rangle = \|v+y\|^2$$

$$\Leftrightarrow v+y=0$$

$$\Leftrightarrow v=-y \quad \ldots (13)$$

(12), (13) $\Rightarrow$

$(y,-y)\in \operatorname{gra} A$

$\Leftrightarrow Ay \ni -y \Leftrightarrow (Id+A)y \ni 0$

$\therefore 0 \in (Id+A)y$

$\Leftrightarrow \exists y\in\mathcal{H}\ \ (Id+A)(y) \ni 0$

$\Leftrightarrow 0 \in \operatorname{ran}(Id+A)$ /* note that this holds for any maximally monotone operator */ ... (14)

take any $w\in\mathcal{H}$,

define $B:\mathcal{H}\to 2^{\mathcal{H}} : x \mapsto -w+Ax$: maximally monotone /* [$A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone]

$\underbrace{u+\gamma A(\cdot+z)}_{\in\mathbb{R}_{++}} : \mathcal{H}\to\mathcal{H}$: maximally monotone */

Here comes the magic step:

now, (14) was derived for any maximally monotone operator, so B, too will satisfy (14)

$0 \in \operatorname{ran}(Id+B)$

$\Leftrightarrow \exists z\in\mathcal{H}\ \ (Id+B)z = z+Bz = 0$

$\Leftrightarrow z-w+Az = 0$

$\Leftrightarrow (Id+A)z = w$

$\Leftrightarrow w \in \operatorname{ran}(Id+A)$

$\therefore \forall_{w\in\mathcal{H}}\ \ w\in\operatorname{ran}(Id+A)$

$\Leftrightarrow \mathcal{H} \subseteq \operatorname{ran}(Id+A)$

also trivially $\operatorname{ran}(Id+A) \subseteq \mathcal{H}$ $\Rightarrow$ $\boxed{\operatorname{ran}(Id+A)=\mathcal{H}}$

■

3/26/2016 7:01 PM

Theorem 21.7. (Debrunner-Flor)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$]

$$\forall_{w \in \mathcal{H}} \ \exists_{x \in \overline{\operatorname{conv}} \operatorname{dom} A} \ \inf_{(y,v) \in \operatorname{gra} A} \langle y-x \mid v-(w-x) \rangle \geq 0$$

<u>Proof:</u>

$C = \overline{\operatorname{conv}} \operatorname{dom} A$

goal:

$$\forall_{w \in \mathcal{H}} \ \exists_{x \in \overline{\operatorname{conv}} \operatorname{dom} A} \ \inf_{(y,v) \in \operatorname{gra} A} \langle y-x \mid v-(w-x) \rangle \geq 0 \quad \ldots (1)$$

now $F_A(x,u) = \langle x \mid u \rangle - \inf_{(y,v) \in \operatorname{gra} A} \langle x-y \mid u-v \rangle$ /* By definition */

$$\therefore F_A(x, w-x) = \langle x \mid w-x \rangle - \inf_{(y,v) \in \operatorname{gra} A} \underbrace{\langle x-y \mid w-x-v \rangle}_{\frac{1}{(-1)^2} \langle -(x-y) \mid -(w-x-v) \rangle = \langle y-x \mid v-w+x \rangle}$$

$$\leftrightarrow \quad \langle x \mid w-x \rangle - F_A(x, w-x) = \inf_{(y,v) \in \operatorname{gra} A} \langle x-y \mid w-x+v \rangle \quad \ldots (2)$$

from (1),(2) the goal is:

$$\forall_{w \in \mathcal{H}} \ \exists_{x \in C} \ \underbrace{\langle x \mid w-x \rangle - F_A(x, w-x)}_{= -\underbrace{\langle x \mid x \rangle}_{\|x\|^2} + \langle x \mid w \rangle - F_A(x, w-x)} \geq 0$$

$$\leftrightarrow \forall_{w \in \mathcal{H}} \ \exists_{x \in C} \ F_A(x, w-x) + \|x\|^2 - \langle x \mid w \rangle \leq 0$$

$$\leftrightarrow \forall_{w \in \mathcal{H}} \ \min_{x \in \mathcal{H}} \left(F_A(x, w-x) + \|x\|^2 - \langle x \mid w \rangle + \iota_C(x) \right) \leq 0 \quad /* \ \exists_{x \in C} \ P(x) \leq 0 \leftrightarrow \min_{x \in \mathcal{H}} P(x) + \iota_C(x) \leq 0 \ */$$

$$\ldots \text{ goal (3)}$$

We divide the proof into two cases: $w = 0$, $w \neq 0$

<u>case 1:</u> $w = 0$

then goal (3) becomes:

$$\min_{x \in \mathcal{H}} \left(F_A(x, -x) + \|x\|^2 + \iota_C(x) \right) \leq 0 \quad \ldots (\text{goal } (4))$$

set. $q = \frac{1}{2} \|\cdot\|^2$

$f: \mathcal{H} \times \mathcal{H} \to]-\infty, +\infty]$: $f(y,x) = \frac{1}{2} \underbrace{F_A^*}_{\text{conjugate of the Fitzpatrick function}} (2y, 2x) \in \Gamma_0(\mathcal{H} \times \mathcal{H})$

$g = (q + \iota_C)^* = q - \frac{1}{2} d_C^2 \in \Gamma_0(\mathcal{H})$ /* Example 13.5. $\left(\frac{1}{2}\|\cdot\|^2 + \iota_C\right)^* = \frac{1}{2}\left(\|\cdot\|^2 - d_C^2\right)$ */

$L: \mathcal{H} \times \mathcal{H} \to \mathcal{H}$: $L(y,x) = x - y \in \mathcal{B}(\mathcal{H} \times \mathcal{H}, \mathcal{H})$

now we show that, $\inf_{(y,x) \in \mathcal{H} \times \mathcal{H}} \left(f(y,x) + g(L(y,x)) \right) \geq 0 \quad \ldots (4)$

$$\leftrightarrow \forall_{(y,x) \in \mathcal{H} \times \mathcal{H}} \ f(y,x) + g(L(y,x)) \geq 0$$

lets expand the objective:

$$(f + g \circ L)(y,x) = f(y,x) + g(\underbrace{L(y,x)}_{x-y}) = \frac{1}{2} F_A^*(2y, 2x) + \underbrace{q(x-y)}_{\frac{1}{2}\|x-y\|^2} - \frac{1}{2} \underbrace{d_C^2(x-y)}_{/* d_C(\square) = \|\square - P_C \square\| \leq \|\square - z\| \ \forall z \in C \ */} \quad \ldots (5)$$

in (4)

We can confine $(y,x) \in \operatorname{dom} f \cap \operatorname{dom} g \circ L$ /*outside it we will have $+\infty$ as the value */
$\subseteq \operatorname{dom} f$

now, $(y,x) \in \operatorname{dom} f \leftrightarrow (2y, 2x) \in \operatorname{dom} F_A^*$ [from (5)] $\ldots (6)$

Now we show (4) as follows:

$\forall_{(y,x)\in\mathcal{H}\times\mathcal{H}}$

/* Proposition 20.51. (ii) : conjugate of Fitzpatrick function

[$A:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A\neq\emptyset$] $\left|\begin{array}{l}\operatorname{conv}\operatorname{gra}A^{-1}\subseteq\operatorname{dom}F_A^*\subseteq\overline{\operatorname{conv}}\operatorname{gra}A^{-1}\subseteq\overline{\operatorname{conv}}\operatorname{ran}A\times\overline{\operatorname{conv}}\operatorname{dom}A\\ F_A^*\geqslant\langle\cdot|\cdot\rangle\end{array}\right.$ */

$\operatorname{dom}F_A^*\subseteq\overline{\operatorname{conv}}\operatorname{ran}A\times\underbrace{\overline{\operatorname{conv}}\operatorname{dom}A}_{C}$. (7)

from (6),(7) $\Rightarrow$ $(2y,2x)\in\overline{\operatorname{conv}}\operatorname{ran}A\times C\Rightarrow\boxed{2x\in C}$ ...(8)

now: $0=\underbrace{4\langle y|x\rangle}_{\langle 2y|2x\rangle}+\|x-y\|^2-\|x+y\|^2$ /* $=4\langle y|x\rangle+\|x\|^2-2\langle x|y\rangle+\|y\|^2-\|x\|^2-2\langle x|y\rangle-\|y\|^2$ */

$=\langle 2y|2x\rangle+\|x-y\|^2-\|(x-y)-2x\|^2$

$\leqslant F_A^*(2y,2x)+\|x-y\|^2-\|(x-y)-2x\|^2$

$\leqslant F_A^*(2y,2x)+\|x-y\|^2-d_C^2(x-y)$

$=2\left(f(y,x)+g(L(y,x))\right)$ [from (5)]

/* (8) $\Rightarrow$ $2x\in C$

now $d_C(x-y)=\|(x-y)-\underbrace{P_C(x-y)}_{\in C}\|\leqslant\|x-y-z\|\quad\forall z\in C$

$\Rightarrow\ -d_C^2(x-y)\geqslant-\|x-y-z\|^2\quad\forall z\in C$ */

so,

$\forall_{(x,y)\in\mathcal{H}\times\mathcal{H}}\quad f(y,x)+g(L(y,x))\geqslant 0\ \Leftrightarrow\ \inf(f+g\circ L)(\mathcal{H}\times\mathcal{H})\geqslant 0$ ...(9)

now $g(x)=\frac{1}{2}\|x\|^2-\frac{1}{2}d_C^2(x)<+\infty\quad\forall_{x\in\mathcal{H}}$

$\Rightarrow\ \operatorname{dom}g=\mathcal{H}$

as a result $\operatorname{dom}g-L\operatorname{dom}f=\mathcal{H}-L\operatorname{dom}f=\mathcal{H}$ /* recall $C-D=\{x-y\mid x\in C,y\in D\}$, so $C=\mathcal{H}$ 라면 some other set C 라면 substract 하면서 range 를 change 하는 거 */

so, $\operatorname{int}(\operatorname{dom}g-L\operatorname{dom}f)=\operatorname{int}\mathcal{H}=\mathcal{H}$

now $0\in\mathcal{H}=\operatorname{int}(\operatorname{dom}g-L\operatorname{dom}f)\subseteq\operatorname{sri}(\operatorname{dom}g-L\operatorname{dom}f)$ /* $\operatorname{int}C\subseteq\operatorname{sri}C$ for C: convex $\Leftarrow$ (6.11) */

Theorem 15.23: $f\in\Gamma_0(\mathcal{H}\times\mathcal{H})$, $g\in\Gamma_0(\mathcal{H})$, $L\in\mathcal{B}(\mathcal{H},\mathcal{H})$

[$f\in\Gamma_0(\mathcal{H}\times\mathcal{H})$, $g\in\Gamma_0(\mathcal{H})$, $L\in\mathcal{B}(\mathcal{H},\mathcal{H})$

$0\in\operatorname{sri}(\operatorname{dom}g-L(\operatorname{dom}f))$]

$\inf(f+g\circ L)(\mathcal{H}\times\mathcal{H})=-\min(f^*\circ L^*+g^{*\vee})(\mathcal{H})$ ✓ /* $\varphi^\vee$: reversal of $\varphi\overset{\text{def}}{\leftrightarrow}\varphi^\vee:x\mapsto\varphi(-x)$ */

$\inf(f+g\circ L)(\mathcal{H})=-\min(f^*\circ L^*+g^{*\vee})(\mathcal{H})$

$f:\mathcal{H}\times\mathcal{H}\to]-\infty,+\infty]$: $f(y,x)=\frac{1}{2}\underbrace{F_A^*}_{\text{conjugate of the Fitzpatrick function}}(2y,2x)\in\Gamma_0(\mathcal{H}\times\mathcal{H})\Rightarrow f^*=\frac{1}{2}F_A$

$g=(q+\iota_C)^*=q-\frac{1}{2}d_C^2\in\Gamma_0(\mathcal{H})\Rightarrow g^*=q+\iota_C$

$L:\mathcal{H}\times\mathcal{H}\to\mathcal{H}$: $L(y,x)=x-y\in\mathcal{B}(\mathcal{H}\times\mathcal{H},\mathcal{H})\Rightarrow L^*:\mathcal{H}\to\mathcal{H}\times\mathcal{H}:x\mapsto(x,-x)$

using 15.23 we have

$\min_{x\in\mathcal{H}}\left(f^*(-L^*x)+g^*(x)\right)\leqslant 0$ [could not figure this step]

$\Rightarrow\ \min_{x\in\mathcal{H}}\left(F_A(x,-x)+(\|x\|^2+\iota_C(x))\right)\leqslant 0$

$\therefore$ goal (4) achieved. ☺

(b)

$w\neq 0$

$B:\mathcal{H}\to 2^{\mathcal{H}}$: $\operatorname{gra}B=-(0,w)+\operatorname{gra}A$

as in (a) $\exists_{(x,-x)\in C\times\mathcal{H}}$ $\{(x,-x)\}\cup\operatorname{gra}B$: monotone

$\leftrightarrow \{(x, w-x)\} \cup \operatorname{gra} A$: monotone

$\blacksquare$

3/29/2016 9:42 AM

Theorem 21.15. (Rockafellar–Veselý)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone

$x \in \mathcal{H}$]

A: locally bounded at $x \leftrightarrow x \notin \operatorname{bdry} \operatorname{dom} A$

Proof:

First we prove $x \notin \operatorname{bdry} \operatorname{dom} A \leftrightarrow x \in \operatorname{int} \overline{\operatorname{dom}} A \Rightarrow A$: locally bounded at x ... (goal 1)

S: set of all points at which A is locally bounded $\Leftrightarrow x \in S \leftrightarrow \exists_{M_x > 0} \exists_{\delta > 0} \forall_{y : \|y - x\|} \|Ay\| < M_x$

$\mathcal{H} \setminus \overline{\operatorname{dom}} A \subseteq S$ ... (1)

/* $\forall x \quad x \in \mathcal{H} \setminus \overline{\operatorname{dom}} A \leftrightarrow x \in \mathcal{H},\ x \notin \overline{\operatorname{dom}} A = \operatorname{dom} A \cup \text{limpoints}(\operatorname{dom} A)$

$\Updownarrow$

$x \notin \operatorname{dom} A,\ x \notin \text{limpoints}(\operatorname{dom} A)$ /*recall: $x \in \text{limpoints}(E) \leftrightarrow \exists_{(x_a)_{a \in A} \subseteq E : x_a \neq x} \ x_a \to x$ */

$\Updownarrow \qquad \Updownarrow$

$Ax = \{\} \qquad \forall (x_a)_{a \in A} \subseteq \operatorname{dom} A : x_a \neq x \quad x_a \neq x$

$\Updownarrow$

$(x_a)_{a \in A} : Ax_a \neq \{\} : x_a \neq x \qquad \exists \varepsilon > 0\ \forall N\ \exists b \geq N\ \|x_b - x\| \geq \varepsilon$

So, $x \in \mathcal{H} \setminus \overline{\operatorname{dom}} A \leftrightarrow x \in \mathcal{H},\ Ax = \{\},\ \forall_{(x_a)_{a \in A} : Ax_a \neq \{\}, x_a \neq x} \exists_{\varepsilon > 0} \forall_N \exists_{b \geq N} \|x_b - x\| \geq \varepsilon$

clearly at such x, Ax will be bounded (as $Ax = \{\}$ does not even have a point, so it cannot be unbounded */

$x \in S$ */

Proposition 21.10. (Sufficient condition for local boundedness of an operator)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone;

$\theta_1 : \mathcal{H} \times \mathcal{H} \to \mathcal{H} : (x, u) \mapsto x$

$z \in \operatorname{int} \theta_1(\operatorname{dom} F_A)$]

A: locally bounded at $z \leftrightarrow z \in S \quad \therefore \operatorname{int} \theta_1(\operatorname{dom} F_A) \subseteq S$

Proposition 21.11. (Representing domain of a maximally monotone operator via Fitzpatrick function)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone

$\theta_1 : \mathcal{H} \times \mathcal{H} \to \mathcal{H} : (x, u) \mapsto x$]

$\operatorname{int} \operatorname{dom} A \subseteq \operatorname{int} \theta_1(\operatorname{dom} F_A) \subseteq \operatorname{dom} A \subseteq \theta_1(\operatorname{dom} F_A) \subseteq \overline{\operatorname{dom}} A$

$\operatorname{int} \operatorname{dom} A = \operatorname{int} \theta_1(\operatorname{dom} F_A)$

$\overline{\operatorname{dom} A} = \overline{\theta_1(\operatorname{dom} F_A)}$

$\operatorname{int} \operatorname{dom} A \subseteq S$

Now we claim

$S \cap \overline{\operatorname{dom}} A = S \cap \operatorname{dom} A$ ... (goal 0)

* Showing $S \cap \overline{\operatorname{dom}} A \subseteq S \cap \operatorname{dom} A$

let, $x \in S \cap \overline{\operatorname{dom}} A \leftrightarrow x \in \overline{\operatorname{dom}} A,\ x \in S$

$\uparrow$ A locally bounded at x

now $\overline{\operatorname{dom}} A$: convex, closed

$\Rightarrow \overline{\operatorname{dom}} A$: sequentially closed $\leftrightarrow \forall_{(x_n)_{n \in \mathbb{N}} \subseteq \overline{\operatorname{dom}} A : x_n \to \tilde{x}} \ \tilde{x} \in \overline{\operatorname{dom}} A$

$\Rightarrow \exists_{\substack{(x_n, u_n)_{n \in \mathbb{N}} : x_n \to x \\ \subseteq \operatorname{gra} A}} \ x \in \overline{\operatorname{dom}} A \wedge (u_n)_{n \in \mathbb{N}}$: bounded

A bounded sequence has a weakly convergent sequence (Lemma 2.37)

So, we can construct a subsequence (still expressed by (x_n, u_n) for convenience)

such that:

$\exists_{\substack{(x_n, u_n)_{n \in \mathbb{N}} : x_n \to x,\ u_n \to u \\ \subseteq \operatorname{gra} A}} \quad (x, u) \in \operatorname{gra} A$, also u: bounded $\Rightarrow x \in \operatorname{dom} A$, locally bounded

$\Rightarrow x \in S \cap \operatorname{dom} A$

$\therefore S \cap \overline{\operatorname{dom}} A \subseteq S \cap \operatorname{dom} A$ ... (2)

Proposition 20.33. (Used heavily by Davis)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

(i) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{weak}}$

$\overset{\text{def}}{\leftrightarrow} \forall_{(x_n, u_n)_{n \in \mathbb{N}} \subseteq \operatorname{gra} A} \ \forall_{(x, u) \in \mathcal{H} \times \mathcal{H} : x_n \to x,\ u_n \to u} \ (x, u) \in \operatorname{gra} A$

now as $\operatorname{dom} A \subseteq \overline{\operatorname{dom}} A$

$\Rightarrow S \cap \operatorname{dom} A \subseteq S \cap \overline{\operatorname{dom}} A$ ... (3)

(i) gra A : sequentially closed in $\mathcal{H}^{strong} \times \mathcal{H}^{weak}$

$\overset{def}{\leftrightarrow} \forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \text{gra}\, A} \ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} : x_n \to x, u_n \rightharpoonup u \quad (x,u) \in \text{gra}\, A$

(ii) gra A : sequentially closed in $\mathcal{H}^{weak} \times \mathcal{H}^{strong}$

$\overset{def}{\leftrightarrow} \forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \text{gra}\, A} \ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} : x_n \rightharpoonup x, u_n \to u \quad (x,u) \in \text{gra}\, A$

(iii) gra A : closed in $\mathcal{H}^{strong} \times \mathcal{H}^{strong}$

now as $\text{dom}\, A \subseteq \text{dom}\, A$

$\Rightarrow \quad S \cap \text{dom}\, A \subseteq S \cap \overline{\text{dom}}\, A \quad \ldots (3)$

So, from (2),(3):

$S \cap \overline{\text{dom}}\, A = S \cap \text{dom}\, A \quad \ldots$ (goal (0) achieved)

$\Downarrow$

(goal (1) achieved)

($\Rightarrow$)

Now, we show

$S \cap \text{bdry dom}\, A = \emptyset$

$\leftrightarrow \neg (\exists_x \ x \in S \wedge x \in \text{bdry dom}\, A)$

$\leftrightarrow \forall_x \ \underbrace{x \notin S}_{\neg(x\in S)} \vee x \notin \text{bdry dom}\, A$

$\leftrightarrow \forall_x \ x \in S \Rightarrow x \notin \text{bdry dom}\, A$

$\leftrightarrow \forall_{x : x\in S} \ x \notin \text{bdry dom}\, A$

Per absurdum $\exists_{x : x\in S} \ x \in \text{bdry dom}\, A$

$\updownarrow$

$\exists\, \delta \in \mathbb{R}_{++} \ A(B(x;2\delta))$: bounded

$C = \overline{\text{dom}}\, A$: convex, closed, nonempty [Corollary 21.12].

$\Rightarrow \text{spts}\, C = P_C(\underbrace{\mathcal{H} \setminus \overline{\text{dom}}\, A}_{\text{bdry}\, C}) = P_C(\text{bdry}\, C)$

$\overline{\text{spts}}\, C = \text{bdry}\, C$

존재하는 점?

$\Rightarrow \quad \exists_{z \in \text{spts}\, C} \ z \in \text{bdry}\, C \cap B(x;\delta)$

$w \in N_C z \setminus \{0\}$

$B(z;\delta) \subseteq B(x;2\delta) \Rightarrow z \in \text{bdry}\, C$

$\vdots$

(incomplete)

$\blacksquare$

3/28/2016 9:33 PM

⊗ Corollary 21.17.

[$A : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, at most single-valued]

A : strong-to-weak continuous everywhere on int dom A

<u>Proof</u>: We want to show:

$\forall_{x \in \text{int dom}\, A} \ \forall_{(x_n)_{n\in\mathbb{N}} : x_n \to x} \ Ax_n \rightharpoonup Ax$

$\leftrightarrow$ A : strong-to-weak continuous everywhere on int dom A.

Fix

$x \in \text{int dom}\, A \Rightarrow x \notin \text{bdry dom}\, A \leftrightarrow A$: locally bounded at $x \leftrightarrow \exists_{\delta \in \mathbb{R}_{++}} \ A(B(x;\delta))$: bounded

Theorem 21.15. (Rockafellar-Veselý)
[$A : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone
$x \in \mathcal{H}$]
A : locally bounded at $x \Leftrightarrow x \notin \text{bdry dom}\, A$

construct, $(x_n)_{n\in\mathbb{N}} \subseteq B(x;\delta) : x_n \to x$

$\therefore (Ax_n)_{n\in\mathbb{N}}$: bounded

$\Rightarrow \exists$ subsequence that weakly converges [Lemma 2.37]

$\leftrightarrow \exists_{(Ax_{k_n})_{n\in\mathbb{N}}} \ \exists_y \ Ax_{k_n} \rightharpoonup y$

$x_{k_n} \to x$ /* any subsequence of a convergent sequence goes to the same limit point */

$(x_{k_n}, Ax_{k_n}) \in \text{gra}\, A : x_{k_n} \to x, Ax_{k_n} \rightharpoonup y$

$\Rightarrow \boxed{(x,y) \in \text{gra}\, A}$ /* using

Proposition 20.33. (Used heavily by Davis)

$(x_{k_n}, Ax_{k_n}) \in \operatorname{gra} A : x_{k_n} \to x,\ Ax_{k_n} \to y$

$\Rightarrow \boxed{(x, y) \in \operatorname{gra} A}$ /* using

Proposition 20.33. (Used heavily by Davis)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

(i) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{weak}}$

$\overset{\text{def}}{\Leftrightarrow} \forall_{(x_n, u_n)_{n\in\mathbb{N}} \subseteq \operatorname{gra} A} \ \forall_{(x,u) \in \mathcal{H}\times\mathcal{H} : x_n \to x,\ u_n \to u} \ (x,u) \in \operatorname{gra} A$

(ii) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{\text{weak}} \times \mathcal{H}^{\text{strong}}$

$\overset{\text{def}}{\Leftrightarrow} \forall_{(x_n, u_n)_{n\in\mathbb{N}} \subseteq \operatorname{gra} A} \ \forall_{(x,u) \in \mathcal{H}\times\mathcal{H} : x_n \to x,\ u_n \to u} \ (x,u) \in \operatorname{gra} A$

(iii) $\operatorname{gra} A$: closed in $\mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{strong}}$ */

$\Rightarrow y = Ax$

[$\because A$: almost single-valued]

$Ax_{k_n} \to Ax$

now note that $(Ax_{k_n})_{n\in\mathbb{N}}$ is an arbitrary subsequence of $(Ax_n)_{n\in\mathbb{N}}$

so: any subsequence of $(x_n)_{n\in\mathbb{N}}$ will weakly converge to $Ax = y$.

$\therefore Ax$: unique cluster point of $(Ax_n)_{n\in\mathbb{N}}$

$Ax_n \rightharpoonup Ax$

using lemma 2.38.

* lemma 2.38. ★★★

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n\in\mathbb{N}}$: converges weakly $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded, possesses at most one weak sequential cluster point.

So, we have proven that,

$\forall_{x \in \operatorname{int}\operatorname{dom} A} \ \forall_{(x_n)_{n\in\mathbb{N}} : x_n \to x} \ Ax_n \rightharpoonup Ax$

$\Leftrightarrow$ A: strong-to-weak continuous everywhere on $\operatorname{int}\operatorname{dom} A$.

$\blacksquare$

Corollary 21.20.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$; maximally monotone; $\lim_{\|x\| \to +\infty} \inf \|Ax\| = +\infty$]

A: surjective $\Leftrightarrow$ A: onto

3/28/2016 10:33 PM

Proof:

/* Corollary 21.19.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

A: surjective $\Leftrightarrow$ A^{-1}: locally bounded everywhere on $\mathcal{H}$ */

goal: A^{-1}: locally bounded on $\mathcal{H}$

Per absurdum A^{-1}: not locally bounded at $u \in \mathcal{H}$

$\Rightarrow \exists_{(x_n, u_n)_{n\in\mathbb{N}} \in \operatorname{gra} A} : \underbrace{u_n \to u}_{\text{finite}} \quad \|x_n\| \to +\infty$

Hence,

$+\infty = \lim_{\|x\| \to +\infty} \inf \|Ax\|$

$= \lim \inf \|Ax_n\| \leq \lim \|Ax_n\| = \lim \|u_n\| = \|u\|$

contradiction.

$\blacksquare$

Part 2

Friday, September 9, 2016 4:23 PM

9/10/2016 7:43 PM

Theorem 21.8.

[[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone]]

$\exists_{\tilde{A}:\ \text{maximally monotone extension of } A}$ $\quad \operatorname{dom} \tilde{A} \subseteq \overline{\operatorname{conv}} \operatorname{dom} A$

proof:

set $C = \overline{\operatorname{conv}} \operatorname{dom} A$

$\mathcal{M} = \{ B \mid B: \text{monotone extension of } A,\ \operatorname{dom} B \subseteq C = \overline{\operatorname{conv}} \operatorname{dom} A \}$ //this set is not empty as //atleast A is there

order $\mathcal{M}$ partially: $\forall_{B_1 \in \mathcal{M}} \forall_{B_2 \in \mathcal{M}} \quad \preceq B_2 \leftrightarrow (B_1 \ \operatorname{gra} B_1 \subseteq \operatorname{gra} B_2)$

Every chain $\mathcal{C}$ in $\mathcal{M}$ has its union as an upper bound of $\mathcal{C}$ //see a detailed explanation of this logic at 3/21/2016 9:38 AM

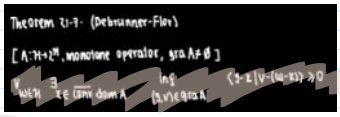
/+ Recall Zorn's lemma: [[A: partially ordered set; $\forall_{C:\ \text{chain in } A} \exists_{\text{upper bound for } C}$] A: contains a maximal element m

$\leftrightarrow \forall_{x \in A.\ x \succeq m} \ x = m$ +/

Using Zorn's lemma we will have a maximal element $\tilde{A} \in \mathcal{M}$. now we show that $\tilde{A}$: maximally monotone. $\tilde{A}$: monotone by construction

take $w \in \mathcal{H}$, assume $w \in \mathcal{H} \setminus \operatorname{ran}(\mathrm{Id} + \tilde{A})$

Now use, Debrunner-Flor theorem:



$\Rightarrow$

$$\exists_{x \in \overline{\operatorname{conv}} \operatorname{dom} \tilde{A}} \ \inf_{(y,v) \in \operatorname{gra} \tilde{A}} \langle y - x \mid v - (w - x) \rangle \geq 0 \quad \ldots (1)$$

Now by construction, $\operatorname{dom} \tilde{A} \subseteq \overline{\operatorname{conv}} \operatorname{dom} A$

$\Rightarrow \overline{\operatorname{conv}} \operatorname{dom} \tilde{A} \subseteq \overline{\operatorname{conv}} (\overline{\operatorname{conv}} \operatorname{dom} A) = \overline{\operatorname{conv}} \operatorname{dom} A$ $\quad // \because A \subseteq B \Rightarrow \overline{\operatorname{conv}} A \subseteq \overline{\operatorname{conv}} B$

$= C$, so (1) becomes:

$$\exists_{x \in C} \ \inf_{(y,v) \in \operatorname{gra} \tilde{A}} \underbrace{\langle y - x \mid v - (w - x) \rangle \geq 0}_{\langle x - y \mid (w - x) - v \rangle \geq 0}$$

now, $w \notin \operatorname{ran}(\mathrm{Id} + \tilde{A})$

$\leftrightarrow \forall_{\tilde{x} \in \mathcal{H}} \ \underbrace{(\mathrm{Id} + \tilde{A})\tilde{x} \not\ni w}$

$\overset{\tilde{x} := x}{\Rightarrow} \quad x + \tilde{A}x \not\ni w \leftrightarrow \tilde{A}x \not\ni w - x \leftrightarrow (x, w - x) \notin \operatorname{gra} \tilde{A}$

so, $\{(x, w-x)\} \cup \operatorname{gra} \tilde{A}$: graph of an operator in $\mathcal{M}$ that properly extends $\tilde{A}$

this contradicts the maximality of $\tilde{A}$

$\therefore \operatorname{ran}(\mathrm{Id} + \tilde{A}) = \mathcal{H}$

$\uparrow$ Minty's theorem

$\tilde{A}$: maximally monotone.

$\blacksquare$

Part 1

Tuesday, September 6, 2016 9:53 AM

Definition. 23.1.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$]

J_A : resolvent of $A \overset{\text{def}}{\leftrightarrow} J_A = (\mathrm{Id}+A)^{-1}$

${}^{\gamma}A$: Yosida approximation of A with index $\gamma \leftrightarrow {}^{\gamma}A = \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma A})$

Proposition 23.2.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$, $x \in \mathcal{H}$, $p \in \mathcal{H}$]

(i) $\operatorname{dom} J_{\gamma A} = \operatorname{dom} {}^{\gamma}A = \operatorname{ran}(\mathrm{Id} + \gamma A)$;

$\operatorname{ran} J_{\gamma A} = \operatorname{dom} A$

(ii) $p \in J_{\gamma A}(x) \leftrightarrow x \in p + \gamma A p$

$\leftrightarrow (x-p) \in \gamma A p$

$\leftrightarrow (p, \frac{1}{\gamma}(x-p)) \in \operatorname{gra} A$

(iii) $p \in {}^{\gamma}A x \leftrightarrow p \in A(x - \gamma p) \leftrightarrow (x - \gamma p, p) \in \operatorname{gra} A$

Example 23.3.

$f \in \Gamma_0(\mathcal{H})$, $\gamma \in \mathbb{R}_{++} \Rightarrow \begin{cases} \operatorname{Prox}_{\gamma f} = J_{\gamma \partial f} \\ \nabla({}^{\gamma}f) = {}^{\gamma}(\partial f) \end{cases}$

9/6/2016 10:14 AM

Proof:

Proposition 16.34:

[$f \in \Gamma_0(\mathcal{H})$; $x, p \in \mathcal{H}$]

$p = \operatorname{Prox}_f x \Leftrightarrow x - p \in \partial f(p)$, i.e.,

$\operatorname{Prox}_f = (\mathrm{Id} + \partial f)^{-1}$

$\Rightarrow \operatorname{Prox}_{\gamma f} = (\mathrm{Id} + \partial \gamma f)^{-1} = (\mathrm{Id} + \gamma \partial f)^{-1} = J_{\gamma \partial f}$ /* By def */

Proposition 12.29.

[$f \in \Gamma_0(\mathcal{H})$; $\gamma \in \mathbb{R}_{++}$]

${}^{\gamma}f : \mathcal{H} \to \mathbb{R}$, Frechet differentiable // ${}^{\gamma}f$: Moreau envelope of f : ${}^{\gamma}f = f \,\square\, (\frac{1}{2\gamma}\|\cdot\|^2)$

$\nabla({}^{\gamma}f) = \frac{1}{\gamma}(\mathrm{Id} - \operatorname{Prox}_{\gamma f})$: γ^{-1} Lipschitz continuous

$= \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma \partial f}) = \underbrace{{}^{\gamma}(\partial f)}_{\text{Yosida approximation of } \partial f}$

∎

Example 23.4.

[C: nonempty closed convex subset $\subseteq \mathcal{H}$

$\gamma \in \mathbb{R}_{++}$]

$J_{N_C} = P_C$ // N_C : normal cone operator

${}^{\gamma}N_C = \frac{1}{\gamma}(\mathrm{Id} - P_C)$

Proof:

/* Example 12.25 : [C: nonempty closed convex subset of $\mathcal{H}$] $\operatorname{Prox}_{\iota_C} = P_C$ */

in Example 23.4. set $f := \iota_C$ we have:

$J_{\gamma \partial f} = J_{\gamma N_C}$ // $\because N_C = \partial \iota_C$

$= \operatorname{Prox}_{\gamma f} = P_C$

$\therefore J_{N_C} = P_C$

and ${}^{\gamma}N_C = \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma \partial f}) = \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma N_C}) = \frac{1}{\gamma}(\mathrm{Id} - P_C)$

∎

9/6/2016 3:37 PM

Proposition 23.7.

[D: nonempty, $\subseteq \mathcal{H}$

$T: D \to \mathcal{H}$

$A = T^{-1} - \mathrm{Id}$]

(i) $T = J$

(i) $T = J_A$

(ii) T: firmly nonexpansive $\Leftrightarrow$ A: monotone

(iii) $\left(\begin{array}{l} T\text{: firmly nonexpansive} \\ D = \mathcal{H} \end{array}\right) \Rightarrow A$: maximally monotone

Proof:

(i) From Def $J_A = (Id + A)^{-1}$

$\forall_{(x,u)}$ $(x,u) \in \operatorname{gra} J_A$

$\Leftrightarrow J_A x \ni u$

$\Leftrightarrow (Id+A)^{-1} x \ni u$

$\Leftrightarrow x \in (Id+A)u$

$\Leftrightarrow x \in u + Au$

$\Leftrightarrow x \in u + (T^{-1} - Id)u$

$\Leftrightarrow x - \cancel{u} \in T^{-1}u - \cancel{u}$

$\Leftrightarrow x \in T^{-1}u$

$\Leftrightarrow Tx \ni u$

$\Leftrightarrow (x,u) \in \operatorname{gra} T$

$\therefore \operatorname{gra} J_A = \operatorname{gra} T \Leftrightarrow T = J_A$.

(ii)

First prove: T: firmly nonexpansive $\Rightarrow$ A: monotone

given T: firmly nonexpansive

$\Leftrightarrow \forall_{x \in D} \forall_{y \in D}$ $\|Tx - Ty\|^2 \leq \langle x-y \mid Tx - Ty \rangle$

take $\left| \begin{array}{l} (x,u) \in \operatorname{gra} A \Leftrightarrow Ax \ni u \Leftrightarrow (T^{-1} - Id)x = T^{-1}x - x \ni u \Leftrightarrow T^{-1}x \ni x+u \Leftrightarrow x \in T(x+u) \Leftrightarrow x = T(x+u) \\ (y,v) \in \operatorname{gra} A \text{, similarly } y = T(y+v),\ y+v \in D \end{array} \right.$

// as $T: D \to \mathcal{H}$ so its a single-valued // operator

as $T(x+u) = x \in \mathcal{H}$, we can say that $(x+u) \in D$

// recall that By def

// $T: D \to \mathcal{H}$ means that T maps every point $\tilde{x}$ in D

// to a point Tx in $\mathcal{H}$, as $T(x+u) = x \in \mathcal{H}$, $x+u \in D$

now, Proposition 4.2.(v) says: T: firmly nonexpansive

$\Leftrightarrow \forall_{x \in D} \forall_{y \in D}$ $\langle Tx - Ty \mid (Id - T)x - (Id - T)y \rangle \geq 0$

$x+u \in D$, $y+v \in D$

$\Rightarrow \underbrace{\langle T(x+u) - T(y+v) \mid (Id - T)(x+u) - (Id - T)(y+v) \rangle}$ ≥ 0

$\langle x-y \mid x+u - T(x+u) - (y+v) + T(y+v) \rangle$ // ingenerally $T(x+y) \neq Tx + Ty$, if so then T: linear

// but $(T+R)x = Tx + Rx$

$= \langle x-y \mid x+u-x-y-v+y \rangle = \langle x-y \mid u-v \rangle$

$\Leftrightarrow \langle x-y \mid u-v \rangle \geq 0$

$\therefore$ So we have proven that, $\forall_{(x,u) \in \operatorname{gra} A} \forall_{(y,v) \in \operatorname{gra} A}$ $\langle x-y \mid u-v \rangle \geq 0 \overset{\text{def}}{\Leftrightarrow} A$: monotone ($\Rightarrow$ proved)

Now let us show that,

A: monotone $\Rightarrow T$: firmly nonexpansive

take $x, y \in D$

now, $T: D \to \mathcal{H}$

$\Rightarrow \exists_{u \in \mathcal{H}}$ $u = Tx = J_A x$ // By (i) $T = J_A = (Id + A)^{-1}$

$= (Id + A)^{-1} x$

$\Leftrightarrow (Id + A)^{-1} x = u$ // this says: $\forall_x \left(x \in D \Leftrightarrow x \in \operatorname{ran}(Id + A) \right) \Leftrightarrow D = \operatorname{ran}(Id + A)$

$\Leftrightarrow x \in (Id + A)y = y + Ay$

$\Leftrightarrow x - u \in Au$ // now $u = Tx$

$\Leftrightarrow x - Tx \in A(Tx) \Leftrightarrow (Tx, x - Tx) \in \operatorname{gra} A$

similarly, $y - Ty \in A(Ty) \Leftrightarrow (Ty, y - Ty) \in \operatorname{gra} A$

A: monotone $\overset{\text{def}}{\Leftrightarrow} \forall_{(x,u) \in \operatorname{gra} A} \forall_{(y,v) \in \operatorname{gra} A}$ $\langle x-y \mid u-v \rangle \geq 0$

$\langle Tx - Ty \mid (Id - T)x - (Id - T)y \rangle \geq 0$

So, we have shown that: Proposition 4.2(v)

$\forall_{x\in D} \forall_{y\in D} \langle Tx-Ty \mid (Id-T)x-(Id-T)y\rangle \geq 0 \longleftrightarrow$ T: firmly nonexpansive

($\Rightarrow$ proved)

(iii) $T: D\to\mathcal{H} \Rightarrow \operatorname{dom} T = D$ //see note page 2 of Bauschke

$\leftrightarrow \operatorname{ran} T^{-1} = \operatorname{dom} T = D$

$\leftrightarrow \operatorname{ran}(Id+A) = D$ // $\because T^{-1} = Id+A$

given $D=\mathcal{H}$

Minty's theorem

$\operatorname{ran}(Id+A)=\mathcal{H} \leftrightarrow$ A: maximally monotone

Proposition 23.9.

[$A:\mathcal{H}\to 2^{\mathcal{H}}$, $\operatorname{dom} A \neq \emptyset$,

$D = \operatorname{ran}(Id+A)$

$T = J_{A|D}$]

(i) $A = T^{-1} - Id$

(ii) A: monotone $\leftrightarrow$ T: firmly nonexpansive

(iii) A: maximally monotone $\leftrightarrow$ (T: firmly nonexpansive $\wedge$ $D=\mathcal{H}$)

9/6/2016 6:14 PM

Proof:

(i) $(x,u)\in \operatorname{gra} A \leftrightarrow Ax \ni u \leftrightarrow Ax+x \ni u+x \leftrightarrow (A+Id)(x) \ni u+x \Rightarrow u+x \in \operatorname{ran}(Id+A) = D$ . . . (1)

$(x,u)\in \operatorname{gra}(T^{-1}-Id) \leftrightarrow (T^{-1}-Id)(x) \ni u$

$\leftrightarrow T^{-1}x - x \ni u$

$\leftrightarrow T^{-1}x \ni x+u$

$\leftrightarrow x \in T(x+u) = J_A|_D(x+u)$ // As $(x+u)\in D$, we have $J_{(A|D)}(x+u) = J_A(x+u)$

$= (Id+A)^{-1}(x+u)$ {$(x+u)\in D$} // From (1)

$\leftrightarrow x = (Id+A)^{-1}(x+u)$

$\leftrightarrow (Id+A)(x) \ni x+u$ {$x+u\in D$}

$\leftrightarrow x+Ax \ni x+u$ { $x+u\in D$ }

$\leftrightarrow Ax \ni u \leftrightarrow (x,u)\in \operatorname{gra} A$ { $x+u\in D$ }

$\therefore \forall_{(x,u)} \left((x,u)\in \operatorname{gra} A \leftrightarrow (x,u)\in (T^{-1}-Id)\right)$

$\leftrightarrow A = T^{-1} - Id$ ☺

(ii)

First we prove, A: monotone $\Rightarrow$ T: firmly nonexpansive

take, // At first we will prove that J_A is single valued, i.e., for $(x,u)\in \operatorname{gra} J_A$, $(y,v)\in \operatorname{gra} J_A$, if $x=y$, then $u=v$. To do this we find out the realtionship between x,u,y,v with A, and use its monotonicity.

take $(x,u)\in \operatorname{gra}(J_A)$

$\leftrightarrow J_A(x)\ni u$

$\leftrightarrow (Id+A)^{-1}x \ni u$

$\leftrightarrow x\in (Id+A)u = u+Au$ //note that as $\exists_u (Id+A)u \ni x \Rightarrow x\in \operatorname{ran}(Id+A)=D$

$\leftrightarrow (x-u)\in Au \leftrightarrow (u, x-u)\in \operatorname{gra} A$

similarly, $(y,v)\in \operatorname{gra} J_A \leftrightarrow (y-v)\in Av \leftrightarrow (v, y-v)\in \operatorname{gra} A$ // similarly, $y\in \operatorname{ran}(Id+A)=D$

A: monotone $\overset{\text{def}}{\leftrightarrow} \forall_{(a,b)\in \operatorname{gra} A} \forall_{(c,d)\in \operatorname{gra} A} \langle a-c \mid b-d\rangle \geq 0$

$\langle u-v \mid (x-u)-(y-v)\rangle \geq 0$

$\langle u-v \mid x-y-(u-v)\rangle = -\|u-v\|^2 + \langle u-v \mid x-y\rangle$

$\leftrightarrow \langle u-v \mid x-y\rangle \geq \|u-v\|^2$ . . . (eq:3)

set $x:=y$ $\Rightarrow \|u-v\|^2 \leq 0 \leftrightarrow \|u-v\|^2 = 0 \leftrightarrow \underbrace{u}_{Tx}=\underbrace{v}_{Ty}$

So, $x=y \Rightarrow u=v \Rightarrow J_A$: single-valued $\Rightarrow T = J_{A|D}$: single valued

thus (eq:3) becomes: // by setting $u=Tx$, $v=Ty$:

$\langle Tx-Ty \mid x-y\rangle \geq \|Tx-Ty\|^2 \quad \forall_{x\in D} \forall_{y\in D}$

//Recall Proposition 4.2.(iv): T: firmly nonexpansive $\leftrightarrow \forall_{x\in D} \forall_{y\in D} \langle x-y \mid Tx-Ty\rangle \geq \|Tx-Ty\|^2$

$x \in D \quad y \in D$

//Recall Proposition 4.2.(iv) : T: firmly nonexpansive $\leftrightarrow \forall_{x \in D} \forall_{y \in D} \langle x-y | Tx-Ty \rangle \geq \|Tx-Ty\|^2$

T: firmly nonexpansive. ($\Rightarrow$ direction proved)

($\Leftarrow$)

T: firmly nonexpansive $\Rightarrow$ A: monotone

From 23.7 (ii) : A: monotone

Proposition 23.7.

[D: nonempty subset of $\mathcal{H}$ ✓

$T: D \to \mathcal{H}$ ✓

$A = T^{-1} - Id$] ✓

(i) $T = J_A$ ✓

(ii) T: firmly nonexpansive $\leftrightarrow$ A: monotone

(iii) T: firmly nonexpansive, $D = \mathcal{H}$ $\leftrightarrow$ A: maximally monotone

(iii)

follows from (i),(ii) and

∎

Corollary 23.10.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, $\gamma \in \mathbb{R}_{++}$]

(i) $J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone

$(Id - J_{\gamma A}): \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone

(ii) Reflected resolvent (Cayley operator)

$R_{\gamma A}: \mathcal{H} \to \mathcal{H}: x \mapsto 2J_{\gamma A}x - x$: nonexpansive

(iii) ${}^{\gamma}A: \mathcal{H} \to \mathcal{H}$: γ-cocoercive

(iv) ${}^{\gamma}A$: maximally monotone

(v) ${}^{\gamma}A: \mathcal{H} \to \mathcal{H}$: $\frac{1}{\gamma}$ Lipschitz continuous

9/7/2016 1:33 PM

Proof : /* Corollary 23.8.

T: firmly nonexpansive $\leftrightarrow \exists_{A: \mathcal{H} \to 2^{\mathcal{H}}, \text{ maximally monotone}}$ $T = J_A = (Id + A)^{-1}$ */

using this

A: maximally monotone, $\gamma \in \mathbb{R}_{++}$

$\Downarrow$

$J_{\gamma A}$: firmly nonexpansive // : $J_{\gamma A} = (Id + \gamma A)^{-1}$: (γA: maximally monotone)

$\leftrightarrow (Id - J_{\gamma A})$: firmly nonexpansive /* Proposition 4.2.

[[D: nonempty, $\subseteq \mathcal{H}$; $T: D \to \mathcal{H}$]] T: firmly expansive $\leftrightarrow$

$Id - T$: firmly nonexpansive */

/* Example 20.27.

($T: \mathcal{H} \to \mathcal{H}$, $]0, \frac{1}{2}]$ averaged) $\Rightarrow$ T: maximally monotone

recall that $]0, \frac{1}{2}]$ averaged operators are firmly nonexpansive

*/

$J_{\gamma A}$, $Id - J_{\gamma A}$: firmly nonexpansive, $\mathcal{H} \to \mathcal{H}$

$\Rightarrow J_{\gamma A}$, $Id - J_{\gamma A}$: maximally monotone

(ii)

/* recall proposition 4.2. (iii)

[[$T: D \to \mathcal{H}$]] T : firmly nonexpansive $\leftrightarrow$ $2T - Id$: nonexpansive

*/

$J_{\gamma A}$: firmly nonexpansive // in (i)

$\leftrightarrow 2J_{\gamma A} - Id = R_{\gamma A}$: nonexpansive

(iii) in (i) we have proven :

$(Id - J_{\gamma A})$: firmly nonexpansive

now By the definition of Yosida approximation : ${}^{\gamma}A = \frac{1}{\gamma}(Id - J_{\gamma A}) \leftrightarrow (Id - J_{\gamma A}) = \gamma\, {}^{\gamma}A$

$\Leftrightarrow \gamma\, {}^{\gamma}A$: firmly nonexpansive

$\Leftrightarrow$ ${}^{\gamma}A$: γ-cocoercive // T : β-cocoercive $\Leftrightarrow$ βT : firmly nonexpansive

$\because$

(iv) /* Example 2028.

[[$T : \mathcal{H} \to \mathcal{H}$, β-cocoercive, $\beta \in \mathbb{R}_{++}$]] T : maximally monotone

in (iii') : ${}^{\gamma}A$: $\underbrace{\gamma}_{\in \mathbb{R}_{++}}$-cocoercive $\Rightarrow$ ${}^{\gamma}A$: maximally monotone

(v) in (iii'), we have proven that : ${}^{\gamma}A$: γ-cocoercive

Cauchy-Schwarz

$$\overset{\text{def}}{\Leftrightarrow} \quad \forall_{x \in \mathcal{H}} \ \forall_{y \in \mathcal{H}} \quad \gamma \, \| {}^{\gamma}Ax - {}^{\gamma}Ay \|^2 \leq \langle x-y \mid {}^{\gamma}Ax - {}^{\gamma}Ay \rangle \leq \| x-y \| \, \| {}^{\gamma}Ax - {}^{\gamma}Ay \|$$

$$\Leftrightarrow \quad \forall_{x \in \mathcal{H}} \ \forall_{y \in \mathcal{H}} \quad \| {}^{\gamma}Ax - {}^{\gamma}Ay \| \leq \frac{1}{\gamma} \| x-y \|$$

$\therefore$ ${}^{\gamma}A$: $\frac{1}{\gamma}$ Lipschitz continuous.

$\blacksquare$

Part 2

Thursday, September 8, 2016 7:37 AM

Proposition 23.11:

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone

$\beta \in \mathbb{R}_{++}$]

A: strongly monotone with constant β $\Leftrightarrow$

J_A: $(\beta+1)$ cocoercive $\Rightarrow$ J_A: Lipschitz continuous with constant $\frac{1}{\beta+1} \in]0,1[$

9/8/2016 9:04 AM

Proof:

$x, y, u, v \in \mathcal{H}$

($\Rightarrow$ direction)

A: β-strongly monotone $\Rightarrow$ A: monotone $\Rightarrow$ $T = J_A$: firmly nonexpansive (single valued operator) on its domain

/* (Proposition 23.7. (ii): J_A: firmly nonexpansive $\leftrightarrow$ A: monotone */

* Proposition 23.2. //basic relationships between A, $\mathcal{H}$, $J_{\gamma A}$

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$, $x \in \mathcal{H}$, $p \in \mathcal{H}$]

(i): $\operatorname{dom} J_{\gamma A} = \operatorname{dom} {}^{\gamma}A = \operatorname{ran}(\mathrm{Id} + \gamma A)$; $\operatorname{ran} J_{\gamma A} = \operatorname{dom} A$

(ii): $p \in J_{\gamma A}(x) \Leftrightarrow x \in p + \gamma A p \Leftrightarrow x - p \in \gamma A p \Leftrightarrow (p, \frac{1}{\gamma}(x-p)) \in \operatorname{gra} A$

(iii): $p \in {}^{\gamma}A x \Leftrightarrow p \in A(x - \gamma p) \Leftrightarrow (x - \gamma p, p) \in \operatorname{gra} A$ */

take (x,u): $J_A x = u \Leftrightarrow x - u \in Au \Leftrightarrow (u, x-u) \in \operatorname{gra} A$

(y,v): $J_A y = v \Leftrightarrow y - v \in Av \Leftrightarrow (v, y-v) \in \operatorname{gra} A$

$\because$ A: β strongly monotone $\overset{\text{def}}{\Leftrightarrow}$ $\forall_{(x,u) \in \operatorname{gra} A} \forall_{(y,v) \in \operatorname{gra} A}$ $\langle x-y | u-v \rangle \geq \beta \|x-y\|^2$

$$\underbrace{\langle u-v | \underbrace{(x-u) - (y-v)}_{(x-y)-(u-v)} \rangle}_{= \langle u-v | x-y \rangle - \|u-v\|^2} \geq \beta \|u-v\|^2$$

$\Leftrightarrow$ $\langle u-v | x-y \rangle \geq (\beta+1) \|u-v\|^2$ $\quad \therefore \forall_{x \in \operatorname{dom} J_A} \forall_{y \in \operatorname{dom} J_A}$ $\langle x-y | J_A x - J_A y \rangle \geq (\beta+1) \|J_A x - J_A y\|^2 \overset{\text{def}}{\Leftrightarrow} J_A$: $(\beta+1)$ cocoercive on $\operatorname{dom} J_A = \operatorname{ran}(\mathrm{Id} + A) = (\mathrm{Id} + A)(\mathcal{H})$ //from 23.2 (i): $\operatorname{dom} J_{\gamma A}$

$\Rightarrow$ $\|u-v\| \|x-y\| \geq \langle u-v | x-y \rangle \geq (\beta+1) \|u-v\|^2$

$\Leftrightarrow$ $\|x-y\| \geq (\beta+1) \|u-v\|$ $\Leftrightarrow$ $\|J_A x - J_A y\| \leq \frac{1}{\beta+1} \|x-y\|$

$\therefore$ J_A: $(\beta+1)$ cocoercive $\Rightarrow$ J_A: $\underbrace{\frac{1}{\beta+1}}_{\in]0,1[}$ Lipschitz continuous. :)

($\Leftarrow$)

J_A: $(\beta+1)$ cocoercive

$\overset{\text{def}}{\Leftrightarrow}$ $\forall_{x \in \operatorname{dom} J_A} \forall_{y \in \operatorname{dom} J_A}$ $\langle x-y | J_A x - J_A y \rangle \geq (\beta+1) \|J_A x - J_A y\|^2$

now take $(x,u) \in \operatorname{gra} A$ $\Leftrightarrow$ $Ax \ni u = (u+x) - x$ // now: Proposition 23.2. (ii) $\boxed{x}$ $-p \in \gamma A p \Leftrightarrow p \in J_{\gamma A} \boxed{x}$ */

$\Leftrightarrow$ $J_A(u+x) \ni x$ // now A: monotone $\Rightarrow$ J_A: single valued, firmly nonexpansive operator

$\Leftrightarrow$ $J_A(u+x) = x$

similarly $(y,v) \in \operatorname{gra} A \Leftrightarrow J_A(y+v) = y$

$(x,y) = (J_A(u+x), J_A(y+v))$ // now, clearly $(u+x) \in \operatorname{dom} J_A$, $(u+v) \in \operatorname{dom} J_A$

$\Rightarrow$

$\langle (u+x) - (y+v) | \underbrace{J_A(u+x)}_{x} - \underbrace{J_A(y+v)}_{y} \rangle \geq (\beta+1) \|\underbrace{J_A(u+x)}_{x} - \underbrace{J_A(y+v)}_{y}\|^2$

$\Leftrightarrow$ $\langle (u-v) + (x-y) | x-y \rangle \geq (\beta+1) \|x-y\|^2$

$\Leftrightarrow$ $\langle u-v | x-y \rangle + \|x-y\|^2 \geq (\beta+1) \|x-y\|^2$

$\Leftrightarrow$ $\langle u-v | x-y \rangle \geq \beta \|x-y\|^2$

$\therefore \forall_{(x,u) \in \operatorname{gra} A} \forall_{(y,v) \in \operatorname{gra} A}$ $\langle x-y | u-v \rangle \geq \beta \|x-y\|^2 \overset{\text{def}}{\Leftrightarrow} A$: β-strongly monotone

$A: \mathcal{H} \to 2^{\mathcal{H}}$: strongly monotone with constant $\beta \in \mathbb{R}_{++}$ $\overset{\text{def}}{\Leftrightarrow}$ $(A - \beta \mathrm{Id})$: monotone $\Leftrightarrow$ $(\forall (x,u) \in \operatorname{gra} A)(\forall (y,v) \in \operatorname{gra} A)$ $\langle x-y | u-v \rangle \geq \beta \|x-y\|^2$

$\boxtimes$

/* Extending a firmly nonexpansive operator on $D \subseteq \mathcal{H}$ (corresponding to a monotone operator) to a

" " " " " $\mathcal{H}$ (" " " maximally monotone operator) */

Theorem 23.13.

[D: nonempty $\subseteq \mathcal{H}$; $T: D \to \mathcal{H}$, firmly nonexpansive] $\Rightarrow$

$\exists_{\tilde{T}: \mathcal{H} \to \mathcal{H}}$ ($\tilde{T}$: firmly nonexpansive, $\tilde{T}|_D = T$, $\operatorname{ran} T \subseteq \overline{\operatorname{conv}} \operatorname{ran} T$)

9/8/2016 12:17 PM

Proof:

*An important assertion of Proposition 23.7. (ii)

[$T: \underbrace{D}_{\mathcal{H}, \subseteq \mathcal{H}} \to \mathcal{H}$: firmly nonexpansive] $\exists_{\substack{A: \mathcal{H} \to 2^{\mathcal{H}} \\ \Downarrow \\ A = T^{-1} - \mathrm{Id}}}$ (A: monotone, $\operatorname{ran}(\mathrm{Id} + A) = D$, $T = J_A$)

using this we have: $\exists_{A: \mathcal{H} \to 2^{\mathcal{H}}: A = T^{-1} - \mathrm{Id}}$ (A: monotone, $\operatorname{ran}(\mathrm{Id} + A) = D$, $T = J_A$) ✓

Theorem 21.8. *

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone]

$\exists_{\tilde{A}}$: maximally monotone extension of A $\quad \operatorname{dom} \tilde{A} \subseteq \overline{\operatorname{conv}} \operatorname{dom} A$

$\rightsquigarrow \exists_{\tilde{A}}$: maximally monotone extension of A $\quad \operatorname{dom} \tilde{A} \subseteq \overline{\operatorname{conv}} \operatorname{dom} A$

Theorem 21.8.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone]
$\exists\, \tilde{A}$: maximally monotone extension of A $\quad \operatorname{dom} \tilde{A} \subset \overline{\operatorname{conv}} \operatorname{dom} A$

$\rightsquigarrow \exists\, \tilde{A}$: maximally monotone extension of A

$\operatorname{dom} A \subseteq \operatorname{conv} \operatorname{dom} A$

define $\tilde{T} = J_{\tilde{A}}$

$\Rightarrow J_{\tilde{A}}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive
$\Rightarrow$ single valued operator so. $\operatorname{dom} \tilde{T} = \mathcal{H}$

/* for a single valued operator $T: X \to Y$. $\operatorname{dom} T = X$ by definition. */

/*
Corollary 23.10. [$A: \mathcal{H} \to 2^{\mathcal{H}}$: maximally monotone; $\gamma \in \mathbb{R}_{++}$]
(i) $J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone
$Id - J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone
(ii) $R_{\gamma A}: \mathcal{H} \to \mathcal{H}: x \mapsto 2J_{\gamma A}x - x$: nonexpansive
reflected resolvent
(iii) ${}^{\gamma}A: \mathcal{H} \to \mathcal{H}$: γ-cocoercive
(iv) ${}^{\gamma}A$: maximally monotone
(v) ${}^{\gamma}A: \mathcal{H} \to \mathcal{H}$: $\frac{1}{\gamma}$ Lipschitz continuous.
*/

$\operatorname{ran} \tilde{T} = \operatorname{ran} J_{\tilde{A}} = \operatorname{dom} \tilde{A}$ /*

/* Proposition 23.2. Basic relationships between A, ${}^{\gamma}A$, $J_{\gamma A}$
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$, $x \in \mathcal{H}$, $p \in \mathcal{H}$]
(i) $\operatorname{dom} J_{\gamma A} = \operatorname{dom} {}^{\gamma}A = \operatorname{ran}(Id + \gamma A)$; $\operatorname{ran} J_{\gamma A} = \operatorname{dom} A$
(ii) $p \in J_{\gamma A}(x) \Leftrightarrow x \in p + \gamma Ap \Leftrightarrow x - p \in \gamma Ap \Leftrightarrow (p, \frac{1}{\gamma}(x-p)) \in \operatorname{gra} A$
(iii) $p \in {}^{\gamma}Ax \Leftrightarrow p \in A(x - \gamma p) \Leftrightarrow (x - \gamma p, p) \in \operatorname{gra} A$
*/

$\subseteq \overline{\operatorname{conv}} \operatorname{dom} A$

$= \overline{\operatorname{conv}} \operatorname{ran} J_A$

$= \overline{\operatorname{conv}} \operatorname{ran} T$

$\therefore \operatorname{ran} \tilde{T} \subseteq \overline{\operatorname{conv}} \operatorname{ran} T$

Finally. let $x \in D \Rightarrow Tx = J_A x = u$ //say

// Using Proposition 23.2 (ii)

$\Leftrightarrow \underbrace{x - u}_{Tx} \in Au \Leftrightarrow x - Tx \in A(Tx) \subseteq \tilde{A}(Tx)$ // $\tilde{A}$: maximally monotone extension of A
// $\therefore A(Tx) \subseteq \tilde{A}(Tx)$

$\Rightarrow x - Tx \in \tilde{A}(Tx)$

$\Leftrightarrow Tx \in J_{\tilde{A}}(x)$ // now, $J_{\tilde{A}}$: firmly nonexpansive

$\Leftrightarrow Tx = J_{\tilde{A}}(x) = \tilde{T}x$

$\therefore \forall_{x \in D}\ \tilde{T}x = Tx$

$\Leftrightarrow \tilde{T}|_D = T$ ∎

9/8/2016 2:08 PM

Corollary 23.14. (Kirszbraun-Valentine theorem)
[$D: \neq \emptyset, \subseteq \mathcal{H}$; $T: D \to \mathcal{H}$, nonexpansive operator] $\Rightarrow \exists_{\tilde{T}: \mathcal{H} \to \mathcal{H}}$ ($\tilde{T}$: nonexpansive, $\tilde{T}|_D = T$, $\operatorname{ran} \tilde{T} \subseteq \overline{\operatorname{conv}} \operatorname{ran} T$)

Proof:

set $R := \frac{1}{2}(Id + T)$ // Proposition 4.2. (iii) T: firmly nonexpansive $\Leftrightarrow \tilde{T} = 2T - Id$ expansive, i.e. $\tilde{T}$: expansive $\Leftrightarrow T = \frac{1}{2}(\tilde{T} + Id)$: firmly nonexpansive

$\Rightarrow R: D \to \mathcal{H}$: firmly nonexpansive /* now use Theorem 23.13.

/* Theorem 23.13. [D: nonempty, $\subseteq \mathcal{H}$; $T: D \to \mathcal{H}$, firmly nonexpansive] $\Rightarrow \exists_{\tilde{T}: \mathcal{H} \to \mathcal{H}}$ ($\tilde{T}$: firmly nonexpansive, $\tilde{T}|_D = T$, $\operatorname{ran} T \subseteq \overline{\operatorname{conv}} \operatorname{ran} T$)

$\Rightarrow \exists_{\tilde{R}: \mathcal{H} \to \mathcal{H}}$ ($\tilde{R}$: firmly nonexpansive, $\tilde{R}|_D = R$, $\operatorname{ran} \tilde{R} \subseteq \underbrace{\overline{\operatorname{conv}} \operatorname{ran} R}_{= C}$)

$\tilde{R}$: firmly nonexpansive, $\tilde{R}|_D = R$, ran

$\Leftrightarrow \tilde{T} = (2\tilde{R} - Id)$: expansive
$\tilde{R} = \frac{1}{2}(Id + \tilde{T})$

$\tilde{R}|_D = R \Leftrightarrow \frac{1}{2}(\tilde{T} + Id)|_D = R = \frac{1}{2}(T + Id) \Leftrightarrow \frac{1}{2}\tilde{T}|_D + \cancel{\frac{1}{2}Id|_D} = \frac{1}{2}T + \frac{1}{2}\cancel{Id} \Rightarrow \tilde{T}|_D = T$ on D we have $Id|_D = Id$

$\operatorname{ran} \tilde{R} \subseteq \overline{\operatorname{conv}} \operatorname{ran} R$

$\operatorname{ran} \tilde{R} = \operatorname{ran} \frac{1}{2}(Id + \tilde{T}) \subseteq \overline{\operatorname{conv}} \operatorname{ran} \frac{1}{2}(T + Id)$ //this part kind of gets complicated, that is why they have defined $\tilde{T} := P_C \circ (2\tilde{R} - Id)$

now define:
$\tilde{T} := P_C \circ (2\tilde{R} - Id)$ it would have the required properties $\tilde{T}|_D = T$
$\operatorname{ran} \tilde{T} \subseteq \overline{\operatorname{conv}} \operatorname{ran} T = C$

∎

Proposition 23.18.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone; $\gamma \in \mathbb{R}_{++}$] $\Rightarrow$
$Id = J_{\gamma A} + \gamma J_{\gamma^{-1}A^{-1}} \circ \gamma^{-1} Id$
$J_{A^{-1}} = Id - J_A$

9/9/2016 12:38 PM

Proof: A: maximally monotone $\Rightarrow J_{\gamma A}$: firmly nonexpansive, $\operatorname{dom} J_{\gamma A} = \mathcal{H}$ //using Proposition 23.3.
$\Rightarrow J_{\gamma A}$:

A^{-1}: maximally monotone $\Rightarrow J_{\gamma^{-1}A^{-1}}$: firmly nonexpansive, $\operatorname{dom} J_{\gamma^{-1}A^{-1}} = \mathcal{H}$
/* A: maximally monotone $\Rightarrow A^{-1}$: maximally monotone: proposition 20.22 */

/*
[D: nonempty subset of $\mathcal{H}$
$T: D \to \mathcal{H}$ // for a single-valued operator $D = \operatorname{dom} T$
$A = T^{-1} - Id$]
(i) $T = J_A \rightsquigarrow A = J_A^{-1} - Id$
(ii) T: firmly nonexpansive $\Leftrightarrow A$: monotone // J_A: firmly nonexpansive $\Leftrightarrow A$: monotone
(iii) (T: firmly nonexpansive, $D = \mathcal{H}$) $\Leftrightarrow A$: maximally monotone
*/

now ${}^{\gamma}A := \frac{1}{\gamma}(Id - J_{\gamma A})$ def

/* Proposition 23.6 (ii):
[$A: \mathcal{H} \to 2^{\mathcal{H}}$; $\gamma \in \mathbb{R}_{++}$; $\mu \in \mathbb{R}_{++}$] ${}^{\gamma}A = (\gamma Id + A^{-1})^{-1} = (J_{\gamma^{-1}A^{-1}}) \circ \gamma^{-1} Id$ */

$\frac{1}{\gamma}(Id - J_{\gamma A}) = (J_{\gamma^{-1}A^{-1}}) \circ \gamma^{-1} Id$

$\Leftrightarrow Id = J_{\gamma A} + \gamma(J_{\gamma^{-1}A^{-1}}) \circ \gamma^{-1} Id$

$\because J_{\gamma A}$: single-valued $\Rightarrow {}^{\gamma}A = \frac{1}{\gamma}(Id - J_{\gamma A})$: single valued $\therefore \forall_{x \in \operatorname{dom} {}^{\gamma}A}\ \exists_{u \in \mathcal{H}}\ {}^{\gamma}Ax = u$

$\leftrightarrow \quad Id = J_{\gamma A} + \gamma(J_{\gamma^{-1}A^{-1}}) \circ \gamma^{-1} Id$

$\because J_{\gamma A}$: single-valued $\Rightarrow {}^{\gamma}A = \frac{1}{\gamma}(Id - J_{\gamma A})$: single valued $\therefore \forall_{x \in dom\, {}^{\gamma}A} \exists_{u \in H}\ {}^{\gamma}Ax = u$

$\frac{1}{\gamma}(Id - J_{\gamma A}) = (\gamma Id + A^{-1})^{-1}$

$\gamma := 1$

$\Rightarrow \quad (Id - J_A) = (Id + A^{-1})^{-1}$ //By definition, $J_{A^{-1}} = (Id + A^{-1})^{-1}$

$= J_{A^{-1}}$

$\leftrightarrow \quad \boxed{Id = J_A + J_{A^{-1}}}$

∎

Proposition 23.20.

[[$T: H \to H$; $\gamma \in R_{++}$]]

T : γ-cocoercive $\leftrightarrow \exists_{A: H \to 2^H,\ \text{maximally monotone}}$ $\quad T = {}^{\gamma}A = \frac{1}{\gamma}(Id - J_{\gamma A})$

9/9/2016 1:27 PM

Proof: ($\Leftarrow$ direction)

/* Corollary 23.10 (iii)

[[$A: H \to 2^H$, maximally monotone ; $\gamma \in R_{++}$]] ${}^{\gamma}A: H \to H$: γ-cocoercive */

using this $T = {}^{\gamma}A$: γ-cocoercive

($\Rightarrow$ direction)

given: T : γ-cocoercive

goal : $A: H \to 2^H$, maximally monotone , $T = {}^{\gamma}A = \frac{1}{\gamma}(Id - J_{\gamma A})$

/* Corollary 23.8 [[$\tilde{T}: H \to H$]] $\tilde{T}$: firmly nonexpansive $\leftrightarrow \exists_{A: H \to 2^H:\ \text{maximally monotone}}$ $\quad \tilde{T} = J_A$ */

γT : firmly nonexpansive //By definition

cor 23.8

$\Leftrightarrow \exists_{B: H \to 2^H:\ \text{maximally monotone}}$ $\quad \gamma T = J_B = Id - J_{B^{-1}}$ // Coming from

Proposition 23.12.
[$A: H \to 2^H$, maximally monotone ; $\gamma \in R_{++}$] $\Rightarrow$
$Id = J_{\gamma A} + \gamma J_{\gamma^{-1}A^{-1}} \circ \gamma^{-1} Id$
$J_{A^{-1}} = Id - J_A$

$\leftrightarrow \quad T = \frac{1}{\gamma}(Id - J_{B^{-1}}) = {}^{\gamma}(B^{-1})$

as B : maximally monotone $\leftrightarrow B^{-1}$: maximally monotone

$\therefore A := B^{-1} : H \to 2^H$, maximally monotone , $T = {}^{\gamma}A$

∎

Proposition 23.27.

[$A: H \to 2^H$, maximally monotone ; $\gamma \in R_{++}$; $B = \gamma^{-1} Id - {}^{\gamma}A = \gamma^{-1} J_{\gamma A}$]

$B: H \to H$: maximally monotone

$J_B = Id - \frac{1}{\gamma} J_{\frac{\gamma^2}{\gamma+1}A} \circ \left(\frac{\gamma}{\gamma+1} Id\right)$

9/10/2016 10:23 PM

Proof: Corollary 23.10 [$A: H \to 2^H$, maximally monotone ; $\gamma \in R_{++}$]
(i) $J_{\gamma A}: H \to H$: firmly nonexpansive, maximally monotone
(ii) $Id - J_{\gamma A}: H \to H$: firmly nonexpansive, maximally monotone

So, $J_{\gamma A}$: firmly nonexpansive $\Rightarrow$ single valued
maximally monotone , $dom\, J_{\gamma A} = H$

$\Rightarrow B = \gamma^{-1} J_{\gamma A}$: single valued,
maximally monotone
$dom\, \gamma^{-1} J_{\gamma A} = H$

/* $\because$ [$A: H \to 2^H$, maximally monotone ; $\lambda, \mu \in H$; $\gamma \in R_{++}$] $\left[\begin{array}{l} \cdot\ A^{\gamma}: \text{maximally monotone} \\ \cdot\ (\lambda + \gamma A)(\mu + \gamma \lambda): \text{maximally monotone} \end{array}\right.$ */

$\forall_{x \in H} \forall_{p \in H} \quad p = J_B x$

$\leftrightarrow \quad x = p + \underbrace{Bp}_{\text{single valued}}$

$\leftrightarrow \quad x - p = Bp = \gamma^{-1} J_{\gamma A} p$

$\leftrightarrow \quad \gamma(x - p) = J_{\gamma A} p \quad \leftrightarrow \quad p \in \gamma(x - p) + \gamma A(\gamma(x - p)) \quad \leftrightarrow \quad p - \gamma(x - p) \in \gamma A(\gamma(x - p))$ // now multiply both sides γ our end goal is putting it in $\tilde{x} - \tilde{p} \in \tilde{\gamma} \tilde{A} \tilde{p}$ form to use the equivalence $\tilde{p} \in J_{\tilde{\gamma}\tilde{A}} \tilde{x}$

/* Proposition 23.2: Basic relationships between A, ${}^{\gamma}A$, $J_{\gamma A}$
[$A: H \to 2^H$, $\gamma \in R_{++}$, $x \in H$, $p \in H$]
(i) $dom\, J_{\gamma A} = dom\, {}^{\gamma}A = ran\, (Id + \gamma A)$; $ran\, J_{\gamma A} = dom\, A$
(ii) $p \in J_{\gamma A}(x) \Leftrightarrow x \in p + \gamma A p \Leftrightarrow x - p \in \gamma A p \Leftrightarrow (p, \frac{1}{\gamma}(x - p)) \in gra\, A$
(iii) $p \in {}^{\gamma}A x \Leftrightarrow p \in A(x - \gamma p) \Leftrightarrow (x - \gamma p, p) \in gra\, A$ */

$\leftrightarrow \quad \gamma^2 A(\gamma(x - p)) \ni \gamma p - \gamma^2 x + \gamma^2 p + \gamma x - \gamma x = \gamma x - \gamma^2(x - p) - \gamma x + \gamma p = \gamma x \underbrace{- \gamma^2(x - p) - \gamma(x - p)}_{-(x-p)\gamma(\gamma+1)} = \gamma x - \gamma(\gamma + 1)(x - p)$ // now divide both sides by $(\gamma+1)$

$\leftrightarrow \quad \underbrace{\frac{\gamma^2}{\gamma+1}}_{\tilde{\gamma}} \underbrace{A}_{\tilde{A}} \underbrace{(\gamma(x - p))}_{\tilde{p}} \ni \underbrace{\frac{\gamma}{\gamma+1} x}_{\tilde{x}} - \underbrace{\gamma(x - p)}_{\tilde{p}}$ // now $\tilde{x} - \tilde{p} \in \tilde{\gamma} \tilde{A} \tilde{p} \leftrightarrow \tilde{p} \in J_{\tilde{\gamma}\tilde{A}} \tilde{x}$

$\leftrightarrow \quad \gamma(x - p) \in J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1} x\right)$ // now $J_{\frac{\gamma^2}{\gamma+1}A}(\cdot)$: single valued, firmly nonexpansive, monotone, so "$\in$" can be replaced with "$=$"

$\leftrightarrow \quad \gamma(x - p) = J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1} x\right)$

$\leftrightarrow \quad x - p = \frac{1}{\gamma} \cdot J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1} x\right)$

$$\Leftrightarrow \quad x - p = \frac{1}{\gamma} \cdot J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1}x\right)$$

$$\Leftrightarrow \quad p = x - \frac{1}{\gamma} J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1}x\right)$$

So, $\forall_{x\in\mathcal{H}} \forall_{p\in\mathcal{H}} \quad p = J_B x \quad \Leftrightarrow \quad p = x - \frac{1}{\gamma} J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1}x\right) = J_B x$

$$\therefore \quad J_B = \mathrm{Id} - \frac{1}{\gamma} J_{\frac{\gamma^2}{\gamma+1}A}\left(\frac{\gamma}{\gamma+1}\mathrm{Id}\right) \quad \blacksquare$$

Proposition 23.32:

[$\mathcal{K}$: real Hilbert space

$L \in \mathcal{B}(\mathcal{H},\mathcal{K})$: $LL^* = \mu\,\mathrm{Id}$; $f \in \Gamma_0(\mathcal{K})$; $x \in \mathcal{H}$]

$f \circ L \in \Gamma_0(\mathcal{H})$

$\mathrm{Prox}_{f\circ L}\, x = x + \mu^{-1} L^*(\mathrm{Prox}_{\mu f}(Lx) - Lx)$

9/12/2016 3:02 PM

Proof:

Fact 2.18: [Important info on linear continuous operator]

[$T \in \mathcal{B}(\mathcal{H},\mathcal{K})$] $\Rightarrow$ (real Hilbert space)

(i) $T^{**} = T$

(ii) $\|T^*\| = \|T\| = \sqrt{\|TT^*\|}$

(iii) $\ker T^* = \overline{\operatorname{ran} T}$

(iv) $\overline{\operatorname{ran} T} = \ker T^*$

(v) $\ker T^*T = \ker T$, $\overline{\operatorname{ran} TT^*} = \overline{\operatorname{ran} T}$.

Fact 2.19: [$\mathcal{K}$: real Hilbert space ; $T \in \mathcal{B}(\mathcal{H},\mathcal{K})$] $\operatorname{ran} T$: closed $\Leftrightarrow$ $\operatorname{ran} T^*$: closed $\Leftrightarrow$ $\operatorname{ran} TT^*$: closed

$\Leftrightarrow$ $\operatorname{ran} T^*T$: closed $\Leftrightarrow$ $\exists_{\mu\in\mathbb{R}_{++}} \forall_{x\in(\ker T)^\perp} \|Tx\| \geq \mu \|x\|$

$\operatorname{ran} L$: closed $\Leftrightarrow$ $\operatorname{ran} L = \overline{\operatorname{ran} L}$

$\operatorname{ran} LL^*$: closed $\Leftrightarrow$ $\operatorname{ran} LL^* = \overline{\operatorname{ran} LL^*}$

$\overline{\operatorname{ran} LL^*} = \overline{\operatorname{ran} L}$

$\Rightarrow \quad \underbrace{\operatorname{ran} \underbrace{LL^*}_{\mu\mathrm{Id}}}_{=\mathcal{K}} = \operatorname{ran} L$

$\therefore \operatorname{ran} L = \mathcal{K}$

Proposition 16.42: [$f \in \Gamma_0(\mathcal{K})$; $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$; one of the following holds: (i) $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{ran} L)$

(ii) $\mathcal{K}$: finite-dimensional, f: polyhedral, $\operatorname{dom} f \cap \operatorname{ran} L \neq \emptyset$] $\partial(f\circ L) = L^* \circ (\partial f) \circ L$

Now $f \in \Gamma_0(\mathcal{K})$; $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$; lets check: $0 \overset{?}{\in} \operatorname{sri}(\operatorname{dom} f - \operatorname{ran} L) = \operatorname{sri}(\operatorname{dom} f - \mathcal{K})$ // $\because \operatorname{ran} L = \mathcal{K}$

Proposition 16.24: [$f \in \Gamma_0(\mathcal{K})$, $g \in \Gamma_0(\mathcal{K})$, $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$] $\operatorname{dom} f \cap L(\operatorname{dom} g) \neq \emptyset$ and one of the following holds:

Clearly. $(\operatorname{dom} f \cap \mathrm{Id}(\mathcal{K})) = (\operatorname{dom} f \cap \mathcal{K}) \neq \emptyset$

(i) $\operatorname{cone}(\operatorname{dom} g - L(\operatorname{dom} f)) = \overline{\operatorname{span}}(\operatorname{dom} g - L(\operatorname{dom} f))$

(ii) $\operatorname{dom} g - L(\operatorname{dom} f)$: closed linear subspace

(iii) $(\operatorname{dom} f, \operatorname{dom} g)$: linear subspaces,

$\operatorname{dom} g + L(\operatorname{dom} f)$: closed

(iv) $\operatorname{dom} g$: cone, $\operatorname{dom} g - \operatorname{cone} L(\operatorname{dom} f)$: closed linear subspaces

(v) $0 \in \operatorname{core}(\operatorname{dom} g - L(\operatorname{dom} f))$

(vi) $0 \in \operatorname{int}(\operatorname{dom} g - L(\operatorname{dom} f))$

(vii) $\operatorname{cont} f \cap L(\operatorname{dom} f) \neq \emptyset$ $\quad \operatorname{cont} f \cap \mathrm{Id}(\mathcal{K}) = \operatorname{cont} f \cap \mathcal{K} \neq \emptyset$

(viii) $\mathcal{K}$: finite dimensional $\wedge$ $(\operatorname{ri} \operatorname{dom} g) \cap \operatorname{ri} L(\operatorname{dom} f) \neq \emptyset$

(ix) $\mathcal{K}$: finite dimensional $\wedge$ $(\operatorname{ri} \operatorname{dom} g) \cap L(\operatorname{ri} \operatorname{dom} f) \neq \emptyset$)

$\Rightarrow \quad 0 \in \operatorname{sri}(\operatorname{dom} f - \mathrm{Id}(\operatorname{dom} f))$ // which will imply $\inf(f + g\circ L)(\mathcal{H}) = -\min(f^* \circ L^* + g^*)(\mathcal{K})$

$\Leftrightarrow 0 \in \operatorname{sri}(\operatorname{dom} f - \mathcal{K})$

So, all the antecedents of Proposition 16.42 holds, so: $\partial(f\circ L) = L^* \circ \partial f \circ L = L^* \partial f L$ . . . (eq: 1)

Now:

Proposition 23.25:

[$\mathcal{K}$: real Hilbert space,

$L: \in \mathcal{B}(\mathcal{H},\mathcal{K})$, LL^*: invertible as $LL^* = \mu\,\mathrm{Id}$

∂f $A: \mathcal{K} \to 2^{\mathcal{K}}$: maximally monotone as $f \in \Gamma_0(\mathcal{K})$

$B = L^* A L$] $\Rightarrow$

(i) $B: \mathcal{H} \to 2^{\mathcal{H}}$: maximally monotone

(ii) $J_B = \mathrm{Id} - L^*(LL^* + A^{-1})^{-1} L$

(iii) $\exists_{\mu\in\mathbb{R}_{++}} \; LL^* = \mu\,\mathrm{Id} \Rightarrow J_B = \mathrm{Id} - L^* A_\mu L$

$\therefore \underbrace{J_{(L^*\partial f L)}}_{J_{\partial(f\circ L)}} = \mathrm{Id} - L^* \circ \underbrace{{}^{\mu}\partial f}_{=\frac{1}{\mu}(\mathrm{Id} - J_{\mu A})} \circ L = \mathrm{Id} - L^* \circ \frac{1}{\mu}(\mathrm{Id} - J_{\mu A}) \circ L$

// from (eq: 1)

$= \mathrm{Id} - \frac{1}{\mu} L^* \circ (\mathrm{Id} - J_{\mu\partial f}) \circ L$

$\leftrightarrow \; J_{\partial(f\circ L)} = Id - \frac{1}{\mu} L^* \circ (Id - \underbrace{J_{\mu\partial f}}_{\partial(\mu f)}) \circ L$ // now $J_{\partial f} = \mathrm{Prox}_f$ for Γ_0 type f

$\leftrightarrow \; \mathrm{Prox}_{f\circ L} = Id - \frac{1}{\mu} L^* \circ (Id - \mathrm{Prox}_{\mu f}) \circ L$

$\therefore \; \forall_{x\in\mathcal{H}} \quad \mathrm{Prox}_{f\circ L}\, x = x - \frac{1}{\mu} L^* (Lx - \mathrm{Prox}_{\mu f}(Lx))$

$= x + \frac{1}{\mu} L^* (Lx + \mathrm{Prox}_{\mu f}(Lx))$ ✓

and by composition law : $f\circ L \in \Gamma_0(\mathcal{H})$. ✓

$\boxtimes$

Part 3

Wednesday, September 14, 2016 1:00 PM

Zeros of monotone operator:

Proposition 23.35.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, strictly monotone] zer A: atmost a singleton

9/14/2016 1:08 PM

Proof:

Per absurdum lets assume zer A has atleast two elements $x, y : x \neq y$

$$\begin{cases} x \in \operatorname{zer} A \leftrightarrow Ax \ni 0 \leftrightarrow (x,0) \in \operatorname{gra} A \\ y \in \operatorname{zer} A \leftrightarrow Ay \ni 0 \leftrightarrow (y,0) \in \operatorname{gra} A \end{cases}, \; x \neq y$$

A: strictly monotone $\overset{\text{def}}{\leftrightarrow} \forall_{(x,u) \in \operatorname{gra} A} \; \forall_{(y,v) \in \operatorname{gra} B} \; (x \neq y \Rightarrow \langle x-y \mid u-v \rangle > 0)$

so, $\langle x-y \mid 0-0 \rangle = 0 > 0 \to$ contradiction

$\therefore$ zer A: atmost a singleton.

∎

Proposition 23.36.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone,

one of the following holds:

(i) A^{-1}: locally bounded everywhere

(ii) $\lim\inf_{\|x\| \to +\infty} \|Ax\| = +\infty$

(iii) dom A: bounded] $\Rightarrow$

zer $A \neq \emptyset$

9/14/2016 1:22 PM

Proof: zer $A \neq \emptyset \leftrightarrow \exists_x \; Ax \ni 0$, if we show that A: surjective (onto) $\leftrightarrow \forall_{y \in \mathcal{H}} \; \exists_{x \in \mathcal{H}} \; Ax \ni y$, this would imply zer $A \neq \emptyset$

Now we have the following results regarding surjectivity of A: maximally monotone operators

Corollary 21.19. /* surjective : onto */
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]
A: surjective $\leftrightarrow A^{-1}$: locally bounded everywhere on $\mathcal{H}$

Corollary 21.20.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, $\lim\inf_{\|x\| \to +\infty} \|Ax\| = +\infty$] A: surjective

Corollary 21.21.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, bounded domain]
A: surjective

Thus $\big((i) \vee (ii) \vee (iii) \big) \Rightarrow A$: surjective $\Rightarrow$ zer $A \neq \emptyset$

∎

Corollary 23.37.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone

one of the following holds:

(i) A: uniformly monotone with a supercoercive modulus

(ii) A: strongly monotone]

A: singleton

9/14/2016 3:25 PM

Proof:

Proposition 22.8.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, one of the following holds | A: uniformly monotone with a supercoercive modulus; A: strongly monotone]

$\lim\inf_{\|x\| \to +\infty} \|Ax\| = +\infty$, A: surjective (onto) /* A: onto $\leftrightarrow \forall_{x \in \mathcal{H}} \; \exists_{y \in \mathcal{H}} \; Ay \ni x$ */

using this:

A: onto

$\leftrightarrow \forall_{x \in \mathcal{H}} \; \exists_{y \in \mathcal{H}} \; Ay \ni x$

$\Rightarrow$ zer $A \neq \emptyset$ . . . (eq:1)

strong monotonicity $\Rightarrow$ uniform monotonicity $\Rightarrow$ strict monotonicity $\Rightarrow$ paramonotonicity $\Rightarrow$ monotonicity

A: strictly monotone

$\Downarrow$

zero A: atmost a singleton (eq:2)

so, from (eq:1) and (eq:2) we have:

zer A: singleton.

∎

Proposition 23.38.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone;

$\gamma \in \mathbb{R}_{++}$]

Fix $J_{\gamma A}$ = zer A = zer γA

9/14/2016 3:54 PM

Proof: $x \in \mathcal{H}$

$x \in \operatorname{zer} A \leftrightarrow Ax \ni 0 \leftrightarrow \gamma Ax \ni 0 = x - x \leftrightarrow x \in x + \gamma A \leftrightarrow x \in J_{\gamma A}(x)$ /* now Proposition 23.37 J_A: firmly nonexpansive $\leftrightarrow A$: monotone */

$\leftrightarrow x = J_{\gamma A}(x) \leftrightarrow x \in \operatorname{Fix} J_{\gamma A}$ . . . (eq: 1)

$x \in \operatorname{zer} A \Leftrightarrow Ax \ni 0 \Leftrightarrow \gamma Ax \ni 0 = x - x \Leftrightarrow x \in x + \gamma A \Leftrightarrow x \in J_{\gamma A}(x)$ now /* Proposition 23.37 J_A: firmly nonexpansive $\Leftrightarrow$ A: monotone */

$\Leftrightarrow x = J_{\gamma A}(x) \Leftrightarrow x \in \operatorname{Fix} J_{\gamma A}$ ... (eq.: 1)

***

* Proposition 23.2. Basic relationships between $A, {}^{\gamma}A, J_{\gamma A}$

$[A: \mathcal{H} \to 2^{\mathcal{H}}, \gamma \in \mathbb{R}_{++}, x \in \mathcal{H}, p \in \mathcal{H}]$

(i) $\operatorname{dom} J_{\gamma A} = \operatorname{dom} {}^{\gamma}A = \operatorname{ran}(\mathrm{Id} + \gamma A)$; $\operatorname{ran} J_{\gamma A} = \operatorname{dom} A$

(ii) $p \in J_{\gamma A}(x) \Leftrightarrow x \in p + \gamma A p \Leftrightarrow x - p \in \gamma A p \Leftrightarrow (p, \frac{1}{\gamma}(x-p)) \in \operatorname{gra} A$

(iii) $p \in {}^{\gamma}Ax \Leftrightarrow p \in A(x - \gamma p) \Leftrightarrow (x - \gamma p, p) \in \operatorname{gra} A$

again $0 \in Ax \Leftrightarrow 0 \in A(x - \gamma 0) \Leftrightarrow 0 \in {}^{\gamma}Ax \Leftrightarrow x \in \operatorname{zer} {}^{\gamma}A$ ... (eq.: 2)

From (i), (ii): $\forall_x \; (x \in \operatorname{zer} A \Leftrightarrow x \in \operatorname{Fix} J_{\gamma A} \Leftrightarrow x \in \operatorname{zer} {}^{\gamma}A)$

$\Leftrightarrow \operatorname{zer} A = \operatorname{Fix} J_{\gamma A} = \operatorname{zer} {}^{\gamma}A$

∎

Proposition 23.39.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone] $\operatorname{zer} A$: closed, convex

9/14/2016 4:05 PM

Proof: A: maximally monotone

$\Rightarrow A^{-1}$: maximally monotone ... (eq.: 1)

$x \in \operatorname{zer} A \Leftrightarrow Ax \ni 0 \Leftrightarrow x \in A^{-1}0$ $\therefore \operatorname{zer} A = A^{-1}0$ ... (eq.: 2)

* Proposition 20.31. (Just beautiful! Output set of a maximally monotone operator on a point is closed, convex */

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $x \in \mathcal{H}$]

Ax: closed, convex

$\tilde{A} := A^{-1}$, $\tilde{x} = 0$ using this, (eq.: 1), (eq.: 2)

$A^{-1}0$: closed convex

∎

Example 23.40. (Classical Proximal Point Algorithm)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $\gamma \in \mathbb{R}_{++}$; $x_0 \in \mathcal{H}$; $\forall_{n \in \mathbb{N}} \; x_{n+1} = J_{\gamma A} x_n$] $\Rightarrow$

(i) $\operatorname{zer} A \neq \emptyset \Rightarrow x_n \to x$; $x \in \operatorname{zer} A$

(ii) A: strongly monotone with $\beta \in \mathbb{R}_{++} \Rightarrow x_n \xrightarrow{\text{linearly}} \underbrace{x}_{\text{unique point in zer}A}$

9/14/2016 7:18 PM

Proof:

(i) Proposition 23.38. [$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone ; $\gamma \in \mathbb{R}_{++}$] $\operatorname{Fix} J_{\gamma A} = \operatorname{zer} A = \operatorname{zer} {}^{\gamma}A$

so, $\operatorname{Fix} J_{\gamma A} = \operatorname{zer} A = \operatorname{zer} {}^{\gamma}A$

Also

**

Corollary 23.10. [$A: \mathcal{H} \to 2^{\mathcal{H}}$: maximally monotone ; $\gamma \in \mathbb{R}_{++}$]

(i) $J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone

$\mathrm{Id} - J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive, maximally monotone

$\therefore J_{\gamma A}: \mathcal{H} \to \mathcal{H}$: firmly nonexpansive

recall Example 5.17. [$T: \mathcal{H} \to \mathcal{H}$, firmly nonexpansive, $\operatorname{Fix} T \neq \emptyset$; $x_0 \in \mathcal{H}$; $\forall_{n \in \mathbb{N}} \; x_{n+1} = Tx_n$] $x_n \to \underbrace{x}_{\text{point in Fix}T}$

$T := J_{\gamma A}$ ($T = J_{\gamma A}$, $\operatorname{Fix} T \neq \emptyset$ by given)

we arrive at the claim that $x_n \to \underbrace{x}_{\text{points}}$ ∴

(ii)

Proposition 23.11.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone ; $\beta \in \mathbb{R}_{++}$]

A: strongly monotone with constant $\beta \Leftrightarrow J_A$: $(\beta+1)$-cocoercive

$\Rightarrow J_A$: Lipschitz continuous with constant $\frac{1}{\beta+1} \in]0,1[$

A: strongly monotone with constant $\beta \in \mathbb{R}_{++}$

$\Leftrightarrow$

$\forall_{(x,u) \in \operatorname{gra} A} \; \forall_{(y,v) \in \operatorname{gra} A} \; \langle x-y \mid u-v \rangle \geq \beta \|x-y\|^2$

$(x,u) \in \operatorname{gra} \gamma A \Leftrightarrow \gamma Ax \ni u \Leftrightarrow A(\gamma x) \ni \gamma u \Leftrightarrow (\gamma x, u) \in \operatorname{gra} A$ ($\gamma \in \mathbb{R}_{++}$)

$(y,v) \in \operatorname{gra} \gamma A \Leftrightarrow (\gamma u, v) \in \operatorname{gra} A$

$\langle \gamma x - \gamma y \mid u - v \rangle \geq \beta \|\gamma x - \gamma y\|^2 = \beta \gamma^2 \|x-y\|^2$

$\Leftrightarrow \gamma \langle x-y \mid u-v \rangle \geq \beta \gamma^2 \|x-y\|^2$

$\Leftrightarrow \langle x-y \mid u-v \rangle \geq \beta \gamma \|x-y\|^2$

$\therefore \forall_{(x,y) \in \operatorname{gra} \gamma A} \; \forall_{(u,v) \in \operatorname{gra} \gamma A} \; \langle x-y \mid u-v \rangle \geq \beta \gamma \|x-y\|^2$

$\therefore \forall_{(x,y)\in \operatorname{gra}\gamma A} \forall_{(u,v)\in \operatorname{gra}\gamma A} \quad \langle x-y \mid u-v\rangle \geq \beta\gamma \|x-y\|^2$

$\leftrightarrow$ γA: strongly monotone with constant $\beta\gamma$

$\Rightarrow J_{\gamma A}$: lipschitz continuous with constant $\frac{1}{\beta\gamma+1} \in]0,1[$

γA: strongly monotone $\Rightarrow J_{\gamma A}$: firmly nonexpansive

$\therefore J_{\gamma A}$: contractive with contraction parameter $\frac{1}{\beta\gamma+1} \in]0,1[$

$\therefore \forall_{n\in\mathbb{N}} \quad x_{n+1} = J_{\gamma A} x_n$: will be equivalent to classical Banach-Picard iteration.

(Banach-Picard iteration for contractive operator)

Theorem 1.48. [(X,d): complete metric space; $T: X \to X$: Lipschitz continuous with constant $\beta \in [0,1[$

$x_0 \in X$; $\forall_{n\in\mathbb{N}} \quad x_{n+1} = Tx_n$]

(i) $\exists_{x}$ x: unique fixed point of T

(ii) $\forall_{n\in\mathbb{N}} \quad d(x_{n+1}, x) \leq \beta\, d(x_n, x)$

(iii) $\forall_{n\in\mathbb{N}} \quad d(x_n, x) \leq \beta^n (x_0, x)$ // $x_n \to x$ linearly

(iv) Priori error estimate: $\forall_{n\in\mathbb{N}} \quad d(x_n, x) \leq \beta^n d(x_0, x_1)/(1-\beta)$

(v) Posteriori error estimate: $\forall_{n\in\mathbb{N}} \quad d(x_n, x) \leq d(x_n, x_{n+1})/(1-\beta)$

(vi) $d(x_0, x_1)/(1+\beta) \leq d(x_0, x) \leq d(x_0, x_1)/(1-\beta)$

Proof: see orienate 11/19/2016 7:16PM

we arrive at the claim.

∎

Theorem 23.41. (Proximal-point algorithm)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, $\operatorname{zer} A \neq \emptyset$;

$(\gamma_n)_{n\in\mathbb{N}}$: $\subseteq \mathbb{R}_{++}$, $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$;

$x_0 \in \mathcal{H}$;

$\forall_{n\in\mathbb{N}} \quad x_{n+1} = J_{\gamma_n A} x_n$]

$\therefore$ (eq:1)

(i) $x_n \to \underbrace{x}_{\in \operatorname{zer} A}$

(ii) A: uniformly monotone on every bounded subset of $\mathcal{H} \Rightarrow x_n \to \underbrace{x}_{\text{unique point in } \operatorname{zer} A}$

9/16/2016 7:32 PM

Proof:

Proof Sketch for Theorem 23.41

First we work on the gap sequence $u_n = \frac{x_n - x_{n+1}}{\gamma_n}$, and show $(x_{n+1}, u_n) \in \mathbf{gra}A$. Then using the monotonicity of A applied to (x_{n+1}, u_n) and (x_{n+2}, u_{n+1}) in $\mathbf{gra}A$, we show that $\|u_{n+1}\| \leq \|u_n\|$, which implies $\|u_n\|$ converges.

Next, we show that $(x_n)_{n\in\mathbb{N}}$ is Fejer monotone w.r.t $\mathbf{zer}A = \mathbf{Fix}J_{\gamma_n A}$ using the firmly nonexpansive nature of $J_{\gamma_n A}$. As $(x_n)_{n\in\mathbb{N}}$ is bounded due to Fejer monotonicity, it will have a weak sequential cluster point.

Third, going back to the initial inequality to prove the Fejer monotone nature of the $(x_n)_{n\in\mathbb{N}}$, we sum over all the indices and show $\sum_{n\in\mathbb{N}} \gamma_n^2 \|u_n\|^2 < +\infty$, which along with given $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$ implies $\underline{\lim}\|u_n\| = 0$, which along with $\lim \|u_n\|$ exists implies $u_n \to 0$.

Now we are in business. We use Theorem 5.5 as our key theorem with $C := \mathbf{zer}A$. Take any weak sequential cluster point x of $(x_n)_{n\in\mathbb{N}}$. Using Proposition 20.3, regarding closedness of the $\mathbf{gra}A$, we show that $(x, 0) \in \mathbf{zer}A$, which suffices for claim (i). Claim (ii) follows from claim (i), Lemma 2.41 and definition of uniform monotonicity of A.

Theorem 5.5.
[$(x_n)_{n\in\mathbb{N}} :\subseteq \mathcal{H}$, Fejer monotone w.r.t. C, every weak sequential cluster point of the sequence is in C]
$\Rightarrow (x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C.

Proposition 20.33.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]
$\mathbf{gra}A$: sequentially closed in $\mathcal{H}^{weak} \times \mathcal{H}^{strong}$, i.e.,
$\forall_{((x_n,u_n))_{n\in\mathbb{N}} \subseteq \mathbf{gra}A} \forall_{(x,u)\in\mathcal{H}\times\mathcal{H}} (x_n, u_n) \to (x,u) \in \mathbf{gra}A$.

Lemma 2.41. [$(x_n)_{n\in\mathbb{N}}, (u_n)_{n\in\mathbb{N}} :\subseteq \mathcal{H}$; $x, u \in \mathcal{H}$]
$(x_n \rightharpoonup x,\ u_n \to x) \Rightarrow (x_n \mid u_n) \to (x \mid u)$.

(stage 0) before we prove (i),(ii)

Set: $\forall_{n\in\mathbb{N}} \quad u_n = \frac{x_n - x_{n+1}}{\gamma_n}$ ... (eq:0) : our first goal is showing that $u_n \to 0$, before we start proving (i) and (ii).

now from (eq:1) $\quad x_{n+1} = J_{\gamma_n A} x_n$

$\leftrightarrow \quad x_n \in x_{n+1} + \gamma_n A x_{n+1}$ // Proposition 23.2.

$\leftrightarrow \quad \underbrace{\frac{x_n - x_{n+1}}{\gamma_n}}_{u_n} \in A x_{n+1}$

▶ ▶▶

\# Proposition 23.2. //Basic relationships between A, ${}^{\gamma}A$, $J_{\gamma A}$

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$, $x \in \mathcal{H}$, $p \in \mathcal{H}$]

(i) $\operatorname{dom} J_{\gamma A} = \operatorname{dom} {}^{\gamma}A = \operatorname{ran}(Id + \gamma A)$; $\operatorname{ran} J_{\gamma A} = \operatorname{dom} A$

(ii) $p \in J_{\gamma A}(x) \Leftrightarrow x \in p + \gamma A p \Leftrightarrow x - p \in \gamma A p \Leftrightarrow (p, \frac{1}{\gamma}(x-p)) \in \operatorname{gra} A$

(iii) $p \in {}^{\gamma}A x \Leftrightarrow p \in A(x - \gamma p) \Leftrightarrow (x - \gamma p, p) \in \operatorname{gra} A$

$\therefore u_n \in A x_{n+1}$ ... (eq:2)

from (eq:0): $\quad (x_{n+1}, u_n) \in \operatorname{gra} A \quad \forall_{n\in\mathbb{N}}$ [eq: 2.5]

$$u_{n+1} = \frac{x_{n+1} - x_{n+2}}{\gamma_{n+1}}$$

$\leftrightarrow \quad x_{n+1} - x_{n+2} = \gamma_{n+1} u_{n+1} \quad \ldots$ (eq:3)

A: maximally monotone $\leftrightarrow \forall_{(x,u)} \left((x,u) \in \operatorname{gra} A \leftrightarrow \forall_{(y,v) \in \operatorname{gra} A} \langle x-y \mid u-v \rangle \geq 0 \right)$

$\because (x_{n+1}, u_n), (x_{n+2}, u_{n+1}) \in \operatorname{gra} A \quad \forall_{n \in \mathbb{N}}$

$$\Rightarrow \forall_{n \in \mathbb{N}} \quad \underbrace{\langle \underbrace{x_{n+1} - x_{n+2}}_{\gamma_{n+1} u_{n+1} \ \parallel \text{ from eq:3}} \mid u_n - u_{n+1} \rangle}_{\shortparallel} \geq 0$$

$$\gamma_{n+1} \langle u_{n+1} \mid u_n - u_{n+1} \rangle = \gamma_{n+1} \left(\underbrace{\langle u_{n+1} \mid u_n \rangle}_{\text{cauchy-schwarz}} - \|u_{n+1}\|^2 \right) \leq \gamma_{n+1} \left(\|u_{n+1}\| \|u_n\| - \|u_{n+1}\|^2 \right) = \gamma_{n+1} \|u_{n+1}\| \left(\|u_n\| - \|u_{n+1}\| \right)$$

$\leq \|u_{n+1}\| \|u_n\|$

$$\therefore \forall_{n \in \mathbb{N}} \quad \underbrace{\gamma_{n+1}}_{\in \mathbb{R}_{++}} \|u_{n+1}\| \left(\|u_n\| - \|u_{n+1}\| \right) \geq \langle x_{n+1} - x_{n+2} \mid u_n - u_{n+1} \rangle \geq 0$$

$$\leftrightarrow \forall_{n \in \mathbb{N}} \quad \underbrace{\|u_{n+1}\|}_{\geq 0} \left(\|u_n\| - \|u_{n+1}\| \right) \geq 0$$

$\Rightarrow \forall_{n \in \mathbb{N}} \quad \|u_n\| - \|u_{n+1}\| \geq 0 \leftrightarrow \|u_{n+1}\| \leq \|u_n\| \quad \ldots$ (eq: 4)

(Rudin)

*** Theorem 3.14**
[$(s_n)_{n \in \mathbb{N}}$: monotonic] $(s_n)_{n \in \mathbb{N}}$: converges $\leftrightarrow$ $(s_n)_{n \in \mathbb{N}}$: bounded

so, the sequence $(\|u_n\|)_{n \in \mathbb{N}}$: monotonic decreasing, and clearly $\|u_n\| \geq 0$. so, $\|u_n\|$ converges as
and $\|u_n\| \leq \|u_0\|$
$\therefore (\|u_n\|)_{n \in \mathbb{N}}$: bounded in $[0, \|u_0\|]$

$\therefore \exists_{C \geq 0} \quad \|u_n\| = \frac{1}{\gamma_n} \|x_n - x_{n+1}\| \to C \quad \ldots$ (eq: 4.5)

$\leftrightarrow \quad \lim \|u_n\| = C \quad \ldots$ (eq: 4.6)

given $\operatorname{zer} A \neq \emptyset$

$\leftrightarrow \exists_z \quad z \in \operatorname{zer} A \quad \ldots$ (eq:5)

now: **Proposition 23.38.** [$A: \mathcal{H} \to 2^{\mathcal{H}}$, monotone ; $\gamma \in \mathbb{R}_{++}$] fix $J_{\gamma A} = \operatorname{zer} A = \operatorname{zer} \overset{\gamma}{A}$

$\therefore \forall_{n \in \mathbb{N}} \quad \operatorname{fix} J_{\gamma_n A} = \operatorname{zer} A \quad \ldots$ (eq: 6)

from (eq:5) and (eq:6):

$z \in \operatorname{Fix} J_{\gamma_n A} \quad \forall_{n \in \mathbb{N}}$

$\leftrightarrow \quad z \in \bigcap_{n \in \mathbb{N}} \operatorname{Fix} J_{\gamma_n A} \quad \ldots$ (eq: 7)

Now, A: maximally monotone $\leftrightarrow \underbrace{J_{\gamma_n A}}_{\mathcal{H} \to \mathcal{H}}$: firmly nonexpansive, maximally monotone $\forall_{n \in \mathbb{N}}$ /* Corollary 23.10 */

$\updownarrow$

$$\forall_{n \in \mathbb{N}} \quad \forall_{x \in \mathcal{H}} \quad \forall_{y \in \mathcal{H}} \quad \|J_{\gamma_n A} x - J_{\gamma_n A} y\|^2 + \|(\mathrm{Id} - J_{\gamma_n A}) x - (\mathrm{Id} - J_{\gamma_n A}) y\|^2 \leq \|x - y\|^2$$

set $x := x_n$, $y := z \in \operatorname{Fix} J_{\gamma_n A}$

$$\|J_{\gamma_n A} x_n - J_{\gamma_n A} z\|^2 \leq \|x_n - z\|^2 - \|(\mathrm{Id} - J_{\gamma_n A}) x_n - \underbrace{(\mathrm{Id} - J_{\gamma_n A}) z}_{(z - J_{\gamma_n A} z) = 0 \ \text{/* } \because z = J_{\gamma_n A} z \text{ */}}\|^2 = \|x_n - z\|^2 - \|x_n - J_{\gamma_n A} x_n\|^2$$

now $\underbrace{J_{\gamma_n A} x_n}_{\shortparallel} \in \operatorname{Fix} J_{\gamma_n A} \leftrightarrow J_{\gamma_n A} z = z$ $\quad$ $\parallel$ by iteration scheme

$\forall_{n \in \mathbb{N}} \quad \|x_{n+1} - z\|^2$

$$= \|J_{\gamma_n A} x_n - J_{\gamma_n A} z\|^2 \leq \|x_n - z\|^2 - \underbrace{\|x_n - \underbrace{J_{\gamma_n A} x_n}_{x_{n+1} \ \parallel \text{ by the iteration scheme}}\|^2}_{\|x_n - x_{n+1}\|^2 = \gamma_n^2 \|u_n\|^2 \ \text{/* } \because u_n = \frac{x_n - x_{n+1}}{\gamma_n} \text{ */}} = \|x_n - z\|^2 - \|x_n - x_{n+1}\|^2 = \|x_n - z\|^2 - \underbrace{\gamma_n^2 \|u_n\|^2}_{\text{negative number. so removing it will give a larger number}} \leq \|x_n - z\|^2$$

$\therefore \forall_{n \in \mathbb{N}} \quad \forall_{z \in \operatorname{Fix} J_{\gamma_n A}} \quad \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2$

Definition. (Fejer monotone sequence)
Let

$$\|x_n - x_{n+1}\|^2 = \gamma_n^2 \|u_n\|^2 \quad /\!* \because u_n = \frac{x_n - x_{n+1}}{\gamma_n} *\!/$$

$\therefore \forall_{n\in\mathbb{N}} \ \forall_{z \in \operatorname{Fix} J_{\gamma_n A}} \ \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2$

$\leftrightarrow$

$(x_n)_{n\in\mathbb{N}}$: Fejer monotone sequence in $\mathcal{H}$ w.r.t. $\operatorname{Fix} J_{\gamma_n A} \ \forall_{n\in\mathbb{N}}$

$\operatorname{Fix} J_{\gamma_n A} = \operatorname{zer} A$ /* from eq:6 */

$\operatorname{zer} A \neq \emptyset$

Definition. (Fejer monotone sequence)
Let
- C : nonempty subset of $\mathcal{H}$,
- $(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$.

Then $(x_n)_{n\in\mathbb{N}}$ is Fejer monotone with respect to C if

$$(\forall x \in C) \ (\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|.$$

• Proposition 5.9. (Some key properties of Fejer monotone sequences)
[C : nonempty subset of $\mathcal{H}$
$(x_n)_{n\in\mathbb{N}}$: Fejer monotone sequence w.r.t C, $S\mathcal{H}$] $\Rightarrow$
(i) $(x_n)_{n\in\mathbb{N}}$: bounded
(ii) $\forall_{z\in C} (\|x_n - z\|)_{n\in\mathbb{N}}$: converges /*note that it does not necessarily mean convergence in C */
(iii) $(d_C(x_n))_{n\in\mathbb{N}}$: decreasing and converges

so, $(x_n)_{n\in\mathbb{N}}$: bounded ... (eq:8), so,

*Lemma 2.57. ★★
[$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$] $\Rightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}}$: subsequence of $(x_n)_{n\in\mathbb{N}}$ $(x_{k_n})_{n\in\mathbb{N}}$: weakly convergent $\leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: possesses a weak sequential cluster point

... (eq:85)

also, $\forall_{n\in\mathbb{N}} \ \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 - \gamma_n^2 \|u_n\|^2$

$\leftrightarrow \ \forall_{n\in\mathbb{N}} \ \gamma_n^2 \|u_n\|^2 \leq \|x_n - z\|^2 - \|x_{n+1} - z\|^2$

$$\leftrightarrow \sum_{n=0}^{m} \gamma_n^2 \|u_n\|^2 \leq \sum_{n=0}^{m} \|x_n - z\|^2 - \sum_{0\leq n\leq m} \|x_{n+1} - z\|^2 = \sum_{n=0}^{m} \|x_n - z\|^2 - \sum_{\tilde{n}=1}^{m+1} \|x_{\tilde{n}} - z\|^2$$

$\leftrightarrow 1 \leq \underbrace{n+1}_{\tilde{n}} \leq m+1$

$\leftrightarrow 1 \leq \tilde{n} \leq m+1$

$$= \|x_0 - z\|^2 + \cancel{\sum_{n=1}^{m} \|x_n - z\|^2} - \cancel{\sum_{\tilde{n}=1}^{m} \|x_{\tilde{n}} - z\|^2} - \|x_{m+1} - z\|^2$$

$$\leftrightarrow \sum_{n=0}^{m} \gamma_n^2 \|u_n\|^2 \leq \underbrace{\|x_0 - z\|^2}_{\text{finite}} - \underbrace{\|x_{m+1} - z\|^2}_{\substack{\text{finite as } (x_n)_{n\in\mathbb{N}} \text{ : bounded per (eq:8)} \\ \forall m}}$$

$$\Rightarrow \lim_{m\to\infty} \sum_{n=0}^{m} \gamma_n^2 \|u_n\|^2 \leq \|x_0 - z\|^2 - \operatorname*{Lt}_{m\to\infty} \|x_{m+1} - z\|^2 = \text{finite}$$

$\therefore \sum_{n\in\mathbb{N}} \gamma_n^2 \|u_n\|^2 < +\infty$

But given $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$

$\Rightarrow \underline{\lim} \|u_n\| = 0$, but $\lim \|u_n\| = c$, now when $\lim$ exists $\underline{\lim} \|u_n\| = \lim \|u_n\| = \overline{\lim} \|u_n\| = 0$ (via eq:9.6)

$\therefore \underline{\lim} \|u_n\| = 0$

$\Leftrightarrow \|u_n\| \to 0$

$\Leftrightarrow u_n \to 0$

*An important result used in the proof of KM iteration
[$(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$; $(b_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$; $\sum_{n\in\mathbb{N}} a_n b_n < +\infty$; $\sum_{n\in\mathbb{N}} a_n = +\infty$] $\underline{\lim}\, b_n = 0$.

(i)
assume x : any weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ /* from (eq:5)

$\leftrightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}} \ x_{k_n} \rightharpoonup x \Leftrightarrow x_{k_n+1} \rightharpoonup x$

now in (eq:2) $\forall_{n\in\mathbb{N}} \ Ax_{n+1} \ni u_n$, and $u_n \to 0$

$\Rightarrow Ax_{k_n+1} \ni u_{k_n}$, and $u_{k_n} \to 0$

$\therefore (x_{k_n+1}, u_{k_n}) \in \operatorname{gra} A$, $u_{k_n} \to 0$, $x_{k_n+1} \rightharpoonup x$

denote $x_{k_n+1} \equiv y_{k_n}$

$(y_{k_n}, u_{k_n}) \in \operatorname{gra} A$, $u_{k_n} \to 0$, $y_{k_n} \rightharpoonup x$ /* now recall :

Proposition 20.33. (Used heavily by Davis) ★★
[$A : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]
(i) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{weak}}$
$\overset{\text{def}}{\leftrightarrow} \forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \operatorname{gra} A} \ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H} : x_n \to x, u_n \rightharpoonup u} \ (x,u) \in \operatorname{gra} A$
(ii) $\operatorname{gra} A$: sequentially closed in $\mathcal{H}^{\text{weak}} \times \mathcal{H}^{\text{strong}}$
$\overset{\text{def}}{\leftrightarrow} \forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \operatorname{gra} A} \ \forall_{(x,u)\in\mathcal{H}\times\mathcal{H} : x_n \rightharpoonup x, u_n \to u} \ (x,u) \in \operatorname{gra} A$
(iii) $\operatorname{gra} A$: closed in $\mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{strong}}$ */

$\therefore (x, 0) \in \operatorname{gra} A \leftrightarrow Ax \ni 0 \leftrightarrow x \in \operatorname{zer} A$

now use Theorem 5.5.

Theorem 5.5. ★★★ ✓ $C := \operatorname{zer} A$ $\quad C \neq \emptyset, \subseteq \mathcal{H}$

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$, Fejer monotone w.r.t. C, every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ is in $\underset{\operatorname{zer} A}{\underset{\parallel}{C}}$] $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} A$

$\therefore x_n$ converges weakly to a point in $\operatorname{zer} A$ ☺

(ii) A: uniformly monotone on every bounded subset of $\mathcal{H}$

strong monotonicity $\Rightarrow$ uniform monotonicity $\Rightarrow$ strict monotonicity $\Rightarrow$ paramonotonicity $\Rightarrow$ monotonicity

Proposition 23.39. [$A: \mathcal{H} \to 2^{\mathcal{H}}$, strictly monotone] $\operatorname{zer} A$: almost a singleton

$\operatorname{zer} A$: nonempty by given

$\operatorname{zer} A$: singleton

let, x: weak limit of $(x_n)_{n\in\mathbb{N}}$ in (i)

as $x \in \operatorname{zer} A \leftrightarrow Ax \ni 0$

Also, [eq: 2.5] $\Rightarrow$ $(x_{n+1}, u_n) \in \operatorname{gra} A \leftrightarrow Ax_{n+1} \ni u_n$

also $(x_n)_{n\in\mathbb{N}}$: bounded from (eq: 8) $\Rightarrow \exists_{\tilde{C}}$ $(x_n)_{n\in\mathbb{N}}$: bounded on $\tilde{C}$

$\therefore (x_n)_{n\in\mathbb{N}} \cup \{x\}$: bounded on $C := \tilde{C} \cup \{x\}$

A: uniformly monotone on C $\leftrightarrow$ $\exists \phi: \mathbb{R}_+ \to [0, +\infty]$, increasing, vanishing only at zero
$\forall_{x\in C} \forall_{y\in C} \forall_{u\in Ax} \forall_{v\in Ay} \langle x-y | u-v \rangle \geq \phi(\|x-y\|)$

$x := x$, $y := x_{n+1}$;
$u := 0$, $v := u_n$

$\langle x - x_{n+1} | 0 - u_n \rangle \geq \phi(\|x - x_{n+1}\|)$

$\therefore \forall_{n\in\mathbb{N}} \langle x_{n+1} - x | u_n \rangle \geq \phi(\|x_{n+1} - x\|)$ // now $\phi \geq 0$. ϕ : vanishes only at zero

now $u_n \to 0$, $x_{n+1} \rightharpoonup x$, so $\quad x_{n+1} - x \rightharpoonup 0$

Lemma 2.41. ★★
[$(x_n)_{n\in\mathbb{N}}, (u_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$; $x, u \in \mathcal{H}$]
(i) $(x_n \rightharpoonup x, \overline{\lim}\|x_n\| \leq \|x\|) \leftrightarrow x_n \to x$
(ii) [$\mathcal{H}$: finite dimensional] $x_n \rightharpoonup x \Leftrightarrow x_n \to x$
} when weak convergence is equivalent to strong convergence
(iii) $(x_n \rightharpoonup x, u_n \to u) \Rightarrow \langle x_n | u_n \rangle \to \langle x | u \rangle$
} one weak convergence and one strong convergence implies strong convergence of the inner product.

says, $\lim \langle x_{n+1} - x | u_n \rangle = 0$

$$\Rightarrow \underbrace{\lim_{n\to\infty} \langle x_{n+1} - x | u_n \rangle}_{= 0} \geq \lim_{n\to\infty} \phi(\|x_{n+1} - x\|) \geq 0$$

$$\Rightarrow \lim_{n\to\infty} \phi(\|x_{n+1} - x\|) = 0$$

$$\Leftrightarrow \|x_{n+1} - x\| \to 0$$

$$\Leftrightarrow x_{n+1} \to x$$

$\therefore x_n \to \underbrace{x}_{\text{unique point in } \operatorname{zer} A}$ ☺

■

Corollary 23.46.

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $x \in \mathcal{H}$]

Exactly one of the following holds:

(i) $x \in \operatorname{dom} A$, ${}^{\gamma}Ax \to {}^{0}Ax$ and $\|{}^{\gamma}Ax\| \uparrow \|{}^{0}Ax\|$ as $\gamma \downarrow 0$

(ii) $x \notin \operatorname{dom} A$, $\|{}^{\gamma}Ax\| \uparrow +\infty$ as $\gamma \downarrow 0$

9/17/2016 10:40 AM

Proof:

define $B: y \mapsto By = A^{-1}y - x$

Proposition 23.21. (Maximally monotonicity preserving operation)
[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $z, x \in \mathcal{H}$; $\gamma \in \mathbb{R}_{++}$]
- A^{-1}: maximally monotone
- $\gamma Ax + (z - x)$: maximally monotone

So. B: maximally monotone

$\forall_z \left(z \in \operatorname{zer} B \leftrightarrow Bz \ni 0 \leftrightarrow A^{-1}z - x \ni 0 \leftrightarrow A^{-1}z \ni x \right.$

$\left. \leftrightarrow z \in Ax \right)$

$\forall_z (z \in \operatorname{zer} B \leftrightarrow z \in Ax) \Leftrightarrow \operatorname{zer} B = Ax$ . . . (eq: 1)

now using Theorem 23.44

// perturbation result: set $\tilde{A} := B$, $\tilde{x} = 0$
Theorem 23.44. [$\tilde{A}: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $\tilde{x} \in \mathcal{H}$] $\Rightarrow$
- $\forall_{\gamma \in]0,1[}$ $\overset{B}{\gamma x} + \overset{B}{\gamma x_\gamma} + (\overset{0}{x_\gamma} - \overset{0}{\tilde{x}}) \ni 0$ define a unique curve $(x_\gamma)_{\gamma \in]0,1[}$
- Exactly one of the following holds:
(i) $\operatorname{zer} \overset{\gamma B}{\tilde{A}} \neq \emptyset$, $x_\gamma \to \underset{\operatorname{zer} \tilde{A}}{P} \overset{0}{\tilde{x}}$ as $\gamma \downarrow 0$
(ii) $\operatorname{zer} \overset{\gamma B}{\tilde{A}} = \emptyset$, $\|x_\gamma\| \uparrow +\infty$ as $\gamma \downarrow 0$

$\forall_{\gamma \in]0,1[}$ $B x_\gamma + \gamma x_\gamma \ni 0$ will define a unique curve $(x_\gamma)_{\gamma \in]0,1[}$

[eq:0]

Proposition 23.6.
[$A: \mathcal{H} \to 2^{\mathcal{H}}$; $\gamma \in \mathbb{R}_{++}$; $x \in \mathbb{R}_{++}$]
(i) $\operatorname{gra} {}^{\gamma}A \subseteq \operatorname{gra}(Ax)_{\mathcal{H}}$
(ii) ${}^{\gamma}A = (\gamma Id + A^{-1})^{-1} = (J_{\gamma^{-1}A^{-1}})^{\circ} \frac{1}{\gamma} Id$. now ${}^{\gamma}A = (J_{\underbrace{\gamma^{-1}A^{-1}}_{\text{maximally monotone}}})^{\circ} \cdot \frac{1}{\gamma} Id$: $w \mapsto {}^{\gamma}Aw = J_{\gamma^{-1}A^{-1}}(\frac{1}{\gamma}w)$: single valued
firmly nonexpansive
(iii) ${}^{\gamma}A = {}^{\gamma}({}^{0}A)$
(iv) $J_{\gamma(A)} = Id + \frac{\gamma}{\gamma+1}(J_{(\gamma A)^{-1}} - Id)$
firmly nonexpansive $\Rightarrow$ single valued

define a unique curve $(x_\gamma)_{\gamma\in]0,1[}$

(i) $\operatorname{zer}\tilde{B}=\emptyset$, $\|x_\gamma\|\uparrow+\infty$ as $\gamma\downarrow 0$ [eq:0]

(ii) $\operatorname{zer}\tilde{B}\neq\emptyset$, $\|x_\gamma\|\uparrow$ as $\gamma\downarrow 0$

(iii) $\lim_{\gamma\downarrow 0} x_\gamma = \gamma^{(0)}A$

(iv) $\exists\, \gamma_{\gamma^{(0)}A} = \lambda + \frac{\gamma}{\gamma+\lambda}(\gamma_{\operatorname{dom}A}-\lambda)$

$\underbrace{\gamma^{-1}A^{-1}}_{\text{maximally monotone}}$

strictly nonexpansive $\Rightarrow$ single valued

$$0\in Bx_\gamma+\gamma x_\gamma = A^{-1}x_\gamma - x + \gamma x_\gamma = (A^{-1}+\gamma\mathrm{Id})x_\gamma - x \Leftrightarrow x\in(A^{-1}+\gamma\mathrm{Id})x_\gamma \Leftrightarrow \underbrace{(A^{-1}+\gamma\mathrm{Id})^{-1}}_{{}^\gamma A}x \ni x_\gamma \Leftrightarrow {}^\gamma Ax\ni x_\gamma \Leftrightarrow {}^\gamma Ax = x_\gamma \quad \dots(eq:2)$$

(i) suppose $x\in\operatorname{dom}A$

$\Leftrightarrow Ax=\operatorname{zer}B\neq\emptyset$ (from eq:1) $\Rightarrow \underbrace{x_\gamma \to P_{\operatorname{zer}B}0}$ as $\gamma\downarrow 0$

${}^\gamma Ax$ (from eq:2) $= x_\gamma$

as this is projection on $\operatorname{zer}B$. $P_{\operatorname{zer}B}0\in\operatorname{zer}B$

now, ${}^0Ax = \min_{y\in Ax} \|y\| = \min_{y\in\operatorname{zer}B}\|y\|$, $Ax = \operatorname{zer}B$

and by definition: $P_{\operatorname{zer}B}0 = \operatorname{argmin}_{y\in\operatorname{zer}B}\frac{1}{2}\|y-0\|^2 = \operatorname{argmin}_{y\in\operatorname{zer}B}\|y\| = {}^0Ax$

$\therefore {}^\gamma Ax \to {}^0Ax$ as $\gamma\downarrow 0$ ✓

Proposition 13.45: [$A:X\rightrightarrows X^*$, maximally monotone; $\gamma\in\mathbb{R}_{++}$; $x\in\mathbb{R}_{++}$; $x\in\operatorname{dom}A$]

(i) $\|{}^\gamma Ax\|\le\inf\|Ax\|$

(ii) $\|{}^{\gamma+\mu}Ax\|\le\|{}^\gamma Ax\|\le\|{}^0Ax\|$

$\Rightarrow \|{}^\gamma Ax\|\uparrow\|{}^0Ax\|$ as $\gamma\downarrow 0$ ✓

(ii) suppose $x\notin\operatorname{dom}A$

$\Leftrightarrow Ax=\operatorname{zer}B=\emptyset$ (from eq:1)

From [eq:0]: $\|x_\gamma\|\uparrow+\infty$ as $\gamma\downarrow 0$

$\|x_\gamma\| = \|{}^\gamma Ax\|$ By (eq:2)

$\therefore \|{}^\gamma Ax\|\uparrow+\infty$ as $\gamma\downarrow 0$ ☺ ∎

Part 1

Saturday, September 17, 2016 7:31 PM

Theorem 24.3: [[A, B: maximally monotone from $\mathcal{H}$ to $2^{\mathcal{H}}$,

$\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)$]] $\Rightarrow$ $A+B$: maximally monotone

Proof blueprint: Our goal is showing that $0 \in \operatorname{sri}(\operatorname{dom} F_A - \operatorname{dom} F_B)$ which will is the antecedent for the addition of two maximally monotone operators to be maximally monotone

9/17/2016 9:38 PM

Proof: fact. $A \subseteq B \Rightarrow \operatorname{cone}(A) \subseteq \operatorname{cone}(B)$, also, $\forall_{C:\text{set}}$ $\operatorname{cone}(C) \subseteq \overline{\operatorname{span}}\, C$

/* recall:

Proposition 23.11. (Representing domain of a maximally monotone operator via Fitzpatrick function)

[[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone

$\varphi_A: \mathcal{H} \times \mathcal{H} \to \mathbb{R}: (x,u) \mapsto \ast$]]

$\operatorname{int} \operatorname{dom} A \subseteq \operatorname{int} \varphi_A(\operatorname{dom} F_A) \subseteq \operatorname{dom} A \subseteq \varphi_A(\operatorname{dom} F_A) \subseteq \overline{\operatorname{dom}} A$

$\operatorname{int} \operatorname{dom} A = \operatorname{int} \varphi_A(\operatorname{dom} F_A)$

$\overline{\operatorname{dom}} A = \varphi_A(\operatorname{dom} F_A)$ */

So, $\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B)$ /* now $\tilde{A} \subseteq \tilde{B}, \tilde{C} \subseteq \tilde{D} \Rightarrow \tilde{A} - \tilde{C} \subseteq \tilde{B} - \tilde{D}$ */

$\subseteq \operatorname{cone}(\varphi_A(\operatorname{dom} F_A) - \varphi_B(\operatorname{dom} F_B))$

$\subseteq \overline{\operatorname{span}}(\underbrace{\varphi_A(\operatorname{dom} F_A)}_{\subseteq \overline{\operatorname{dom}} A} - \underbrace{\varphi_B(\operatorname{dom} F_B)}_{\subseteq \overline{\operatorname{dom}} B})$

$\subseteq \overline{\operatorname{span}}(\overline{\operatorname{dom}} A - \overline{\operatorname{dom}} B)$

$= \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)$ // as we remove closure from $\overline{\operatorname{dom}} A, \overline{\operatorname{dom}} B$, the final closure on $\overline{\operatorname{span}}$ will // make the end result same

$= \operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B)$ // given

So, the sandwiched sets collapses!

so, $\operatorname{cone}(\varphi_A(\operatorname{dom} F_A - \operatorname{dom} F_B)) = \overline{\operatorname{span}}(\varphi_A(\operatorname{dom} F_A - \operatorname{dom} F_B))$

now recall, $\operatorname{sri} C = \{x \in C \mid \overline{\operatorname{span}}(C - x) = \operatorname{cone}(C - x)\}$

i.e., $x \in C \Leftrightarrow \overline{\operatorname{span}}(C - x) = \operatorname{cone}(C - x)$

$\Rightarrow 0 \in \operatorname{sri} \varphi_A(\operatorname{dom} F_A - \operatorname{dom} F_B)$

now recall that

Theorem 24.2: [[$\varphi_A: \mathcal{H} \times \mathcal{H} \to \mathbb{R}: (x,u) \mapsto \ast$; A, B: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, $0 \in \operatorname{sri} \varphi_A(\operatorname{dom} F_A - \operatorname{dom} F_B)$]] $\Rightarrow$ $A+B$: maximally monotone

So, $A+B$: maximally monotone ∎

Corollary 24.4: [[A, B: maximally monotone, $: \mathcal{H} \to 2^{\mathcal{H}}$; one of the following holds:

(i) $\operatorname{dom} B = \mathcal{H}$

(ii) $\operatorname{dom} A \cap \operatorname{int} \operatorname{dom} B \neq \emptyset$

(iii) $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$

(iv) $\forall_{x \in \mathcal{H}} \exists_{\varepsilon \in \mathbb{R}_{++}} [0, \varepsilon x] \subseteq \operatorname{dom} A - \operatorname{dom} B$

(v) $\operatorname{dom} A, \operatorname{dom} B$: convex, $0 \in \operatorname{sri}(\operatorname{dom} A - \operatorname{dom} B)$]] $\Rightarrow$ $A+B$: maximally monotone

Proof:

9/18/2016 11:45 AM

$\operatorname{dom} B = \mathcal{H}$ // by (i)

$\Leftrightarrow \forall_{x \in \mathcal{H}} Bx \neq \emptyset$

$\Rightarrow \operatorname{int} \operatorname{dom} B = \mathcal{H}$ /* as $\mathcal{H}$ is both open and closed

then $\operatorname{dom} A \cap \operatorname{int} \operatorname{dom} B = \operatorname{dom} A \cap \mathcal{H} \neq \emptyset$ /* $\operatorname{dom} A \neq \emptyset$ */ /* this is (ii) */

$\operatorname{int} \operatorname{dom} B = \mathcal{H} \Rightarrow \operatorname{dom} A \subseteq \operatorname{dom} B$

$C := \operatorname{dom} A - \operatorname{dom} B = \{x - y \mid Ax \neq \emptyset, y \in \mathcal{H}\}$, note that no matter what $0 \in C$ as $y \in \mathcal{H}$

note that $C := \mathcal{H}$ as $\forall_{z \in \mathcal{H}}$ setting $y = -z + x \in \mathcal{H}$, will give $x - y = x - (-z + x) = z$

so, $\operatorname{int} C = \mathcal{H} \to 0 \in \operatorname{int} C = \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$ /* this is (iii) */

as $\operatorname{dom} A - \operatorname{dom} B = \mathcal{H} \Rightarrow \forall_{x \in \mathcal{H}} \forall_{\varepsilon \in \mathbb{R}_{++}} [0, \varepsilon x] \subseteq \operatorname{dom} A - \operatorname{dom} B$ /* this is (iv) */

clearly (iv) $\Rightarrow \operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B) \Rightarrow A+B$: maximally monotone

/* using Theorem 24.3: [[A, B: maximally monotone, $: \mathcal{H} \to 2^{\mathcal{H}}$; $\operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)$]] $A+B$: maximally monotone */

now consider (v): $\operatorname{dom} A, \operatorname{dom} B$: convex, $0 \in \operatorname{sri}(\operatorname{dom} A - \operatorname{dom} B)$ /* recall: $\operatorname{sri} C = \{x \in C \mid \operatorname{cone}(C - x) = \overline{\operatorname{span}}(C - x)\}$ */

$\Leftrightarrow \operatorname{cone}(\operatorname{dom} A - \operatorname{dom} B) = \overline{\operatorname{span}}(\operatorname{dom} A - \operatorname{dom} B)$

$\Rightarrow A+B$: maximally monotone ∎

/* **A: uniformly monotone with modulus φ** $\overset{\text{def}}{\Leftrightarrow}$ **φ: increasing, vanishes only at 0, $\forall_{(x,u),(y,v) \in \operatorname{gra} A}$ $\langle x - y \mid u - v \rangle \geq \varphi(\|x - y\|)$** */

Example 24.11: [[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, uniformly monotone with supercoercive modulus φ]] A: 3^* monotone $\overset{\text{def}}{\Leftrightarrow}$ $\operatorname{dom} A \times \operatorname{ran} A \subseteq \operatorname{dom} F_A$

9/18/2016 1:27 PM

Proof: set. $(x,u) \in \operatorname{gra} A$, $w \in \mathcal{H}$, /* $\lim_{t \to +\infty} \frac{\varphi(t)}{t} = +\infty$ */

$\gamma := \|w - u\|$, $\psi: \mathbb{R}_+ \to [-\infty, +\infty[: t \mapsto \gamma t - \varphi(t)$ /* $\varphi(0) = 0$ as φ: vanishes only at zero */

fixed for fixed w, u $\Leftrightarrow \psi(t) = \gamma t - \varphi(t) = t\left(\gamma - \frac{\varphi(t)}{t}\right) \Rightarrow \psi(0) = 0 \cdot 0 - \varphi(0) = 0$

$\Rightarrow \lim_{t \to +\infty} \psi(t) = \lim_{t \to +\infty} t\left(\gamma - \frac{\varphi(t)}{t}\right)$ /* recall that, if $\lim_{x \to a} f(x)$, $\lim_{x \to a} g(x)$ exists $\Rightarrow \lim_{x \to a} f(x) g(x) = \lim_{x \to a} f(x) \lim_{x \to a} g(x)$ */

now, as $\lim_{t \to \infty} \frac{\varphi(t)}{t} = +\infty$, $\lim_{t \to +\infty} t = +\infty$, $\Rightarrow \lim_{t \to +\infty} \psi(t) = +\infty(\gamma - \infty) = -\infty$

$\Rightarrow$ we can find a z: $\forall_t$ $t > z \Rightarrow \psi(t) \leq 0 = \psi(0)$, so, the graph will look somewhat like this:

thus from the figure, $\sup_{t \in \mathbb{R}_{++}} \psi(t) = \sup_{t \in [0, z]} \psi(t) \leq \gamma z - \varphi(z)$

$\Rightarrow \sup_{t \in \mathbb{R}_{++}} \psi(t) = \sup_{t \in [0, z]} \psi(t) \leq \gamma z - \varphi(z)$ /* now $\varphi(0) = 0$, and increasing $\therefore \forall_{t \geq 0}$ $\varphi(t) \geq 0$, so removing $-\varphi(z)$ will give a larger number */

$\leq \gamma z < +\infty$ ... (eq. 1)

Cauchy Schwarz Inequality: $|\langle x, y \rangle| \leq \|x\| \|y\|$

$\max\{\langle x, y \rangle, -\langle x, y \rangle\}$

Fisher 1: $\langle x, y \rangle \leq \|x\| \|y\|$

Fisher 2: $-\langle x, y \rangle \leq \|x\| \|y\| \Leftrightarrow \langle x, y \rangle \geq -\|x\| \|y\|$

So, $F_A(x,u) - \langle x, u \rangle = \langle x, u \rangle - \inf_{(y,v) \in \operatorname{gra} A} \langle x - y \mid u - v \rangle = -\inf_{(y,v) \in \operatorname{gra} A} \langle x - y \mid u - v \rangle = \sup_{(y,v) \in \operatorname{gra} A} -\langle x - y \mid u - v \rangle = \sup_{(y,v) \in \operatorname{gra} A} -\langle x - y \mid (w - u) + (u - v) \rangle = \sup_{(y,v) \in \operatorname{gra} A} \left(-\langle x - y \mid w - u \rangle - \langle x - y \mid u - v \rangle\right)$ /* now using Cauchy-Schwarz Fisher 2: $-\langle x - y \mid w - u \rangle \leq \underbrace{\|x - y\| \|w - u\|}_{r} = r\|x - y\|$

and as **A: uniformly monotone with modulus φ** $\overset{\text{def}}{\Leftrightarrow}$ **φ: increasing, vanishes only at 0, $\forall_{(x,u),(y,v) \in \operatorname{gra} A}$ $\langle x - y \mid u - v \rangle \geq \varphi(\|x - y\|)$**

$\Rightarrow (x,u), (y,v) \in \operatorname{gra} A \Rightarrow \langle x - y \mid u - v \rangle \geq \varphi(\|x - y\|) \Leftrightarrow -\langle x - y \mid u - v \rangle \leq -\varphi(\|x - y\|)$ */

$\overset{\text{def}}{\Leftrightarrow} \langle x, u \rangle - \inf_{(y,v) \in \operatorname{gra} A} \langle x - y \mid u - v \rangle$

$= \sup \left(\|x - y\| \left(\|w - u\| + \|u - v\|\right)\right)$

// watch out: trick move!

$\leq \gamma z < +\infty$ ... (Eq:1)

So. $\underbrace{F_A(x,u)}_{\text{fixed}} - \langle x|u\rangle = \langle x|u\rangle - \inf_{(y,v)\in \mathrm{gra}A} \langle x-y|u-v\rangle - \langle x|u\rangle = -\inf_{(y,v)\in\mathrm{gra}A} \langle x-y|u-v\rangle = \sup_{(y,v)\in\mathrm{gra}A} -\langle x-y|u-v\rangle = \sup_{(y,v)\in\mathrm{gra}A} -\langle x-y|\overbrace{(u-u)+(u-v)}^{u-u+u-v}\rangle = \sup_{(y,v)\in\mathrm{gra}A} \Big(-\langle x-y|u-u\rangle - \langle x-y|u-v\rangle\Big)$ /* now using Cauchy-Schwarz Inequ: $-\langle x-y|u-v\rangle \leq \|x-y\|\,\overbrace{\|u-v\|}^{r} = r\|x-y\|$

$\langle x|u\rangle - \inf_{(y,v)\in\mathrm{gra}A}\langle x-y|u-v\rangle$

$= \sup_{(y,v)\in\mathrm{gra}A} \Big(\|x-y\|\big(\|v-u\|+\|u-u\|\big)\Big)$

// watch out: trick move!

and as A: uniformly monotone with modulus $\phi \overset{\text{def}}{\Leftrightarrow} \phi$: increasing, vanishes only at 0, $\forall_{(x,u),(y,v)\in\mathrm{gra}A}\ \langle x-y|u-v\rangle \geq \phi(\|x-y\|)$

$\langle x,u\rangle, \langle y,v\rangle \in \mathrm{gra}A \Rightarrow \langle x-y|u-v\rangle \geq \phi(\|x-y\|) \Leftrightarrow -\langle x-y|u-v\rangle \leq -\phi(\|x-y\|)$ */

$\leq \sup_{\substack{(y,v)\in\mathrm{gra}A \\ \Downarrow \\ u\in\mathrm{dom}A}} \Big(\underbrace{r\|x-y\| - \phi(\|x-y\|)}_{\psi(\|x-y\|)}\Big) = \sup_{y\in\mathrm{dom}A} \psi(\|x-y\|) = \sup_{\substack{\|x-y\|\in\mathbb{R}_{+} \\ x:\text{fixed}}} \psi(t) < +\infty$ [from (Eq:1)]

/* as v absent in objective */

$\therefore F_A(x,u) - \langle x|u\rangle = \text{finite}$

$\Rightarrow F_A(x,u) = \text{finite} + \underbrace{\langle x|u\rangle}_{\text{finite}} = \text{finite} < +\infty \Leftrightarrow (x,u)\in\mathrm{dom}F_A$ /* for a function f. $\mathrm{dom}f = \{x\in\mathcal{H} \mid f(x) < +\infty\}$ */

So. $\forall_{x\in\mathrm{dom}A}\ \forall_{u\in\mathcal{H}}\ (x,u)\in\mathrm{dom}F_A$

$\Leftrightarrow \forall_{(x,u)\in\mathrm{dom}A\times\mathcal{H}}\ (x,u)\in\mathrm{dom}F_A \Leftrightarrow \mathrm{dom}A\times\mathcal{H}\subseteq\mathrm{dom}F_A$

$\because \mathrm{dom}A\times\mathrm{ran}A \subseteq \mathrm{dom}A\times\mathcal{H}\subseteq\mathrm{dom}F_A \Rightarrow \mathrm{dom}A\times\mathrm{ran}A\subseteq\mathrm{dom}F_A$

Definition 24.3. (3^* monotone operator)
[A:$\mathcal{H}\to 2^{\mathcal{H}}$, monotone] A: 3^* monotone $\overset{\text{def}}{\Leftrightarrow}$ $\mathrm{dom}A\times\mathrm{ran}A\subseteq\mathrm{dom}F_A$

A: 3^* monotone ■

Proposition 24.15.

[[A:$\mathcal{H}\to 2^{\mathcal{H}}$, monotone, $\gamma\in\mathbb{R}_{++}$]]

(i) A: 3^*-monotone $\leftrightarrow$ A^{-1}: 3^*-monotone

(ii) A: 3^*-monotone $\leftrightarrow$ γA: 3^*-monotone

9/19/2016 6:48 PM

Proof.

(i) /* Proposition 20.47. (vii). $F_A(x,u) = F_{A^{-1}}(u,x)$, now $F_A^T(u,x) = F_A(x,u)$ $\therefore F_A^T(u,x) = F_{A^{-1}}(u,x)$ $\forall_{(u,x)}$ $\Leftrightarrow F_A^T = F_{A^{-1}} \Leftrightarrow \mathrm{dom}F_A^T = \mathrm{dom}F_{A^{-1}}$

A: 3^*-monotone $\Leftrightarrow$

$\mathrm{dom}A\times\mathrm{ran}A\subseteq\mathrm{dom}F_A = \{(x,u)\in\mathcal{H}\times\mathcal{H} \mid F_A(x,u) < +\infty\}$ /* now. $\underbrace{\mathrm{dom}F_A^T}_{:(x,u)\mapsto(u,x)} = \{(u,x) \mid F_A^T(x,u) = F_A(x,u) < +\infty\}$, so if we swap the coordinates of the vectors in $\mathrm{dom}F_A$ we get $\mathrm{dom}F_A^T$ */

$\Leftrightarrow \underbrace{\mathrm{ran}A}_{\mathrm{dom}A^{-1}} \times \underbrace{\mathrm{dom}A}_{\mathrm{ran}A^{-1}} \subseteq \mathrm{dom}F_A^T$ /* $\mathrm{ran}A^{-1}=\mathrm{dom}A$; $\mathrm{dom}A^{-1}=\mathrm{ran}A$ page 2 */

$\Leftrightarrow \mathrm{dom}A^{-1}\times\mathrm{ran}A^{-1}\subseteq\mathrm{dom}F_A^T = \mathrm{dom}F_{A^{-1}}$

$\therefore \mathrm{dom}A^{-1}\times\mathrm{ran}A^{-1}\subseteq\mathrm{dom}F_{A^{-1}}$

$\overset{\text{def}}{\Leftrightarrow} A^{-1}$: 3^* monotone :)

(ii)

A: 3^* monotone

$\Leftrightarrow \mathrm{dom}A\times\mathrm{ran}A\subseteq\mathrm{dom}F_A$ /* $\gamma\,\mathrm{ran}A = \gamma\{u\in\mathcal{H} \mid \exists_{x\in\mathcal{H}}\ Ax\ni u\}$;

$\Leftrightarrow \mathrm{dom}A\times\gamma\,\mathrm{ran}A\subseteq(\mathrm{Id}\times\gamma\mathrm{Id})(\mathrm{dom}F_A)$ /* now, according to Proposition 20.47. (vii)

$\Leftrightarrow \mathrm{dom}\gamma A\times\mathrm{ran}\gamma A\subseteq\mathrm{dom}F_{\gamma A}$

$\Leftrightarrow \gamma A$: 3^* monotone

[A: monotone, $\mathrm{gra}A\neq\emptyset$] $F_{\gamma A}(x,\tilde{u}) = \gamma F_A(x,\overset{u}{\tfrac{\tilde{u}}{\gamma}}) \Leftrightarrow F_A(x,u) = \frac{1}{\gamma}F_{\gamma A}(x,u\gamma)$ $\because \tilde{u} = u\gamma$

$\mathrm{dom}F_A = \{(x,u) \mid F_A(x,u) < +\infty\}$

$\Rightarrow (\mathrm{Id}\times\gamma\mathrm{Id})(\mathrm{dom}F_A) = \{(x,\gamma u) \mid \underbrace{F_A(x,u) < +\infty}_{\Leftrightarrow} \} = \{(x,\overbrace{\gamma u}^{\tilde{u}}) \mid F_{\gamma A}(x,u\gamma) < +\infty\} = \{(x,\tilde{u}) \mid F_{\gamma A}(x,\tilde{u}) < +\infty\} = \mathrm{dom}F_{\gamma A}$

$\frac{1}{\gamma}F_{\gamma A}(x,u\gamma) < +\infty \Leftrightarrow F_{\gamma A}(x,u\gamma) < +\infty$

$\therefore (\mathrm{Id}\times\gamma\mathrm{Id})(\mathrm{dom}F_A) = \mathrm{dom}F_{\gamma A}$

also, $\mathrm{dom}A = \{x\in\mathcal{H} \mid Ax\neq\emptyset\}$

now $x\in\mathrm{dom}A \Leftrightarrow Ax\neq\emptyset \Leftrightarrow \gamma Ax\neq\emptyset \Leftrightarrow x\in\mathrm{dom}\gamma A$ $\therefore \mathrm{dom}A = \mathrm{dom}\gamma A$ */

$\gamma\,\mathrm{ran}A = \gamma\{y \mid \exists_{x\in\mathcal{H}}\ Ax\ni y\} = \{\gamma y \mid \exists_{x\in\mathcal{H}}\ Ax\ni y\} = \{v \mid v = \gamma y \wedge \exists_{x\in\mathcal{H}}\ Ax\ni y\} = \{v \mid \exists_{x\in\mathcal{H}}\ Ax\ni \overset{y}{\tfrac{v}{\gamma}}\}$

$= \{v \mid \exists_{x\in\mathcal{H}}\ \gamma Ax\ni v\}$

$= \mathrm{ran}\gamma A$

$\therefore \gamma\,\mathrm{ran}A = \mathrm{ran}\gamma A$ */

■

Proposition 24.17.

[[A,B: monotone operators, $\mathcal{H}\to 2^{\mathcal{H}}$]]

$\forall_{x\in\mathcal{H}}\ \forall_{y\in\mathcal{H}}\ F_{A+B}(x,u) \leq \big(F_A(x,\cdot)\,\square\,F_B(x,\cdot)\big)(u)$

9/20/2016 10:36 PM

Proof:

Fix $(x,u)\in\mathcal{H}\times\mathcal{H}$

$\exists\ (u_1,u_2)\in\mathcal{H}\times\mathcal{H}: u = u_1+u_2$

take, $\left.\begin{array}{l}(y,v_1)\in\mathrm{gra}A \leftrightarrow v_1\in Ay \\ (y,v_2)\in\mathrm{gra}B \leftrightarrow v_2\in By\end{array}\right\} \Rightarrow v_1+v_2\in Ay+By = (A+B)y \Leftrightarrow (y,v_1+v_2)\in\mathrm{gra}(A+B)$

$F_{A+B}(x,u) = F_{A+B}(x,u_1+u_2) = \sup_{(y,v)\in\mathrm{gra}(A+B)} \Big(\langle x|v\rangle + \langle y|u\rangle - \langle y|v\rangle\Big)$ //now, $(y,v_1+v_2)\in\mathrm{gra}(A+B)$

$\geq \langle x|v_1+v_2\rangle + \langle y|u\rangle - \langle y|v_1+v_2\rangle$

$= \langle x|v_1+v_2\rangle + \langle y|u\rangle - \langle y|v_1+v_2\rangle$ // $u = u_1+u_2$

$= \langle x|v_1\rangle + \langle x|v_2\rangle + \langle y|u_1\rangle + \langle y|u_2\rangle - \langle y|v_1\rangle - \langle y|v_2\rangle$

$= \big(\langle x|v_1\rangle + \langle y|u_1\rangle - \langle y|v_1\rangle\big) + \big(\langle x|v_2\rangle + \langle y|u_2\rangle - \langle y|v_2\rangle\big)$ /* recall, $F_A(z,t) = \sup_{(z,t)\in\mathrm{gra}A} \Big(\langle z|t\rangle + \langle z|t\rangle - \langle z|t\rangle\Big)$ */

$\leq \underbrace{\sup_{(y,v_1)\in\mathrm{gra}A} \Big(\langle x|v_1\rangle + \langle y|u_1\rangle - \langle y|v_1\rangle\Big)}_{F_A(x,u_1)} + \underbrace{\sup_{(y,v_2)\in\mathrm{gra}A} \Big(\langle x|v_2\rangle + \langle y|u_2\rangle - \langle y|v_2\rangle\Big)}_{F_B(x,u_2)}$

$= F_A(x,u_1) + F_B(x,u_2)$

I donot follow how $F_{A+B}(x,u) \leq F_A(x,u_1) + F_B(x,u_2)$? কিভাবে হবে?

■

Theorem 24.20. (Brézis-Haraux)

[A,B: monotone, $\mathcal{H}\to 2^{\mathcal{H}}$, A+B: maximally monotone

one of the following holds

(i) $A, B : 3^*$ monotone

(ii) $\operatorname{dom} A \subseteq \operatorname{dom} B$, $B : 3^*$ monotone $] \Rightarrow$

- $\overline{\operatorname{ran}}\,(A+B) = \overline{\operatorname{ran} A + \operatorname{ran} B}$
- $\operatorname{int} \operatorname{ran}(A+B) = \operatorname{int}(\operatorname{ran} A + \operatorname{ran} B)$

9/21/2016 10:21 AM

Proof: From Simon's theorem

> Theorem 24.17 (Simon)
> $[\![A, B$: monotone, $\mathcal{H} \to 2^{\mathcal{H}}$, $A+B$: maximally monotone ;
> $\forall_{u \in \operatorname{ran} A} \forall_{v \in \operatorname{ran} B} \exists_{x \in \mathcal{H}} \; (x,u) \in \operatorname{dom} F_A \wedge (x,v) \in \operatorname{dom} F_B]\!]$
> $\Rightarrow \overline{\operatorname{ran}}(A+B) = \overline{\operatorname{ran} A + \operatorname{ran} B}$, $\operatorname{int} \operatorname{ran}(A+B) = \operatorname{int}(\operatorname{ran} A + \operatorname{ran} B)$

so if we can prove $\forall_{u \in \operatorname{ran} A} \forall_{v \in \operatorname{ran} B} \exists_{x \in \mathcal{H}} \; (x,u) \in \operatorname{dom} F_A \wedge (x,v) \in \operatorname{dom} F_B$

then Simon's theorem will take us to the goal.

suppose antecedent (i) holds: i.e., $A, B : 3^*$ monotone

$\overset{\text{def}}{\leftrightarrow} \operatorname{dom} A \times \operatorname{ran} A \subseteq \operatorname{dom} F_A$, $\operatorname{dom} B \times \operatorname{ran} B \subseteq \operatorname{dom} F_B$

take any (u,v): $u \in \operatorname{ran} A$, $v \in \operatorname{ran} B$

$A+B$: maximally monotone $\Rightarrow$ $\operatorname{gra}(A+B)$: nonempty $\Rightarrow \operatorname{dom}(A+B)$: nonempty

$\leftrightarrow \exists_{\tilde{x} \in \mathcal{H}} \; (A+B)\tilde{x} = \tilde{u} + \tilde{v} \neq \emptyset$

$\Rightarrow \tilde{x} \in \operatorname{dom} A \cap \operatorname{dom} B$

$\therefore \forall_{u \in \operatorname{ran} A} \forall_{v \in \operatorname{ran} A} \exists_{x := \tilde{x} \in \operatorname{dom} A \cap \operatorname{dom} B}$ $\quad (\tilde{x}, u) \in \operatorname{dom} A \times \operatorname{ran} A \subseteq \operatorname{dom} F_A$, $\quad (\tilde{x}, v) \in \operatorname{dom} B \times \operatorname{ran} B \subseteq \operatorname{dom} F_B$

$\cdots$ (Simon's antecedent)

$\Downarrow$

goal :)

now suppose antecedent (ii) holds:

$\forall u \in \operatorname{ran} A$

$\exists_{\tilde{x}} \; A\tilde{x} \ni u \Rightarrow \tilde{x} \in \operatorname{dom} A$

$\therefore \forall_{u \in \operatorname{ran} A} \exists_{\tilde{x} \in \operatorname{dom} A} \; (\tilde{x}, u) \in \operatorname{gra} A \subseteq \operatorname{dom} F_A$

$\therefore \forall_{u \in \operatorname{ran} A} \exists_{\tilde{x} \in \operatorname{dom} A} \; (\tilde{x}, u) \in \operatorname{dom} F_A$

> * Proposition 20.49 (Fitzpatrick function of monotone operator) ★★★★★
> $[\![A : \mathcal{H} \to 2^{\mathcal{H}}$, monotone, $\operatorname{gra} A \neq \emptyset$ // This proposition characterizes points in the graph of
> $(x,u) \in \mathcal{H} \times \mathcal{H}]\!]$ // a monotone operator via Fitzpatrick function Nicely
> (i) $(x,u) \in \operatorname{gra} A \Rightarrow F_A(x,u) = \langle x, u \rangle$ /* i.e., $(x,u) \in \operatorname{gra} A \Leftrightarrow (x,u) \in \operatorname{dom} F_A$ */
> $\therefore \operatorname{gra} A \subseteq \operatorname{dom} F_A$

now. $\tilde{x} \in \operatorname{dom} A \subseteq \operatorname{dom} B$ //given

$\therefore \forall_{v \in \operatorname{ran} B} \; (\tilde{x}, v) \in \operatorname{dom} B \times \operatorname{ran} B \subseteq \operatorname{dom} F_B$ // as $B : 3^*$ monotone

$\forall_{u \in \operatorname{ran} A} \forall_{v \in \operatorname{ran} B} \exists_{\tilde{x} \in \mathcal{H}} \; (\tilde{x}, u) \in \operatorname{dom} F_A \wedge (\tilde{x}, v) \in \operatorname{dom} F_B \Rightarrow$ Simon's antecedent

$\Downarrow$

goal :) ∎

Corollary 24.23.

$[\![A, B$: maximally monotone, $\mathcal{H} \to 2^{\mathcal{H}}$, $A+B$: maximally monotone

B: uniformly monotone with a supercoercive modulus φ,

One of the following holds:

(i) $A : 3^*$ monotone

(ii) $\operatorname{dom} A \subseteq \operatorname{dom} B]\!] \Rightarrow$

(i) $\operatorname{ran}(A+B) = \mathcal{H}$

(ii) $\operatorname{zer}(A+B)$: singleton.

9/23/2016 9:21 PM

Proof:

/*

> Proposition 22.8.
> $[\![\tilde{A} : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, one of the following holds | $\tilde{A}$: uniformly monotone with a supercoercive modulus
> $\tilde{A}$: strongly monotone $]\!]$
>
> $\lim_{\|x\| \to \infty} \inf \{\langle x, z \rangle \mid z \in \tilde{A}x\} = +\infty$, $\tilde{A}$: surjective (onto) /* $\tilde{A}$: onto $\Leftrightarrow \forall_{x \in \mathcal{H}} \exists_{z \in \mathcal{H}} \; \tilde{A}y \ni x$ */

*/

> * $S : \mathcal{H} \to [-\infty, +\infty]$: supercoercive $\overset{\text{def}}{\Leftrightarrow} \lim_{\|x\| \to \infty} (S(x) / \|x\|) = +\infty$
> $\tilde{A}$: uniformly monotone with modulus $\varphi \overset{\text{def}}{\Leftrightarrow} \varphi$: increasing, vanishes only at 0, $\forall_{(x,u),(y,v) \in \operatorname{gra} A} \; \langle x-y, u-v \rangle \geq \varphi(\|x-y\|)$
> Example 24.11. $[\![\tilde{A} : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, uniformly monotone with supercoercive modulus
> $\varphi]\!] \Rightarrow \tilde{A} : 3^*$ monotone.

*/

setting $\tilde{A} := B$

$B : 3^*$ monotone, surjective (onto)

> *
> Corollary 24.22. $[\![A, B$: monotone, $\mathcal{H} \to 2^{\mathcal{H}}$, $A+B$: maximally monotone, A or B: surjective (onto)
> One of the following holds: (i) $A, B : 3^*$ monotone, (ii) $\operatorname{dom} A \subseteq \operatorname{dom} B$, $B : 3^*$ monotone
> $]\!] \Rightarrow A+B$: surjective

$\neg$ given (i) $\vee$ (ii)

$\Rightarrow A+B$: surjective $\leftrightarrow \operatorname{ran}(A+B) = \mathcal{H}$.

(ii) B: uniformly monotone $\Rightarrow B$: strictly monotone

A: monotone

$A+B$: strictly monotone $\quad$ set $\tilde{A} := A+B$

> Proposition 23.15. $[\![\tilde{A} : \mathcal{H} \to 2^{\mathcal{H}}$, strictly monotone $]\!]$ $\operatorname{zer} \tilde{A}$: almost a singleton

$\operatorname{zer}(A+B)$: singleton (as $\operatorname{zer}(A+B) \neq \emptyset$ by (i))

∎

// Parallel sum of resolvents:

Proposition 24.28.

$[\![A, B$: monotone operators, $\mathcal{H} \to 2^{\mathcal{H}}$, $A+B$: maximally monotone $]\!]$

$J_A \square J_B = J_{\frac{1}{2}(A+B)} \circ \frac{1}{2} \mathrm{Id}$

9/23/2016 11:15 AM

Proof:

Proposition 24.26:

[A, B: monotone operators, $\mathcal{H} \to 2^{\mathcal{H}}$, $A+B$: maximally monotone]

$J_A \square J_B = J_{\frac{1}{2}(A+B)} \circ \frac{1}{2} \mathrm{Id}$

9/23/2016 11:15 AM

PROOF:

$(x,u) \in \mathrm{gra}(J_A \square J_B) \leftrightarrow (J_A \square J_B)x \ni u$ /* now,

/* Helpful properties of parallel sum

$\tilde{A} := J_A \circ R_A$, $\tilde{B} := J_B \circ R_B$

Proposition 23.18: [A, B: monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$] $\Rightarrow$

(i) $(A \square B)x = \bigcup_{y \in \mathcal{H}} (A_y \square B(x-y))$

(ii) $u \in (\tilde{A} \square \tilde{B})x \Leftrightarrow \exists_{u \in \mathcal{H}} \left(y \in \tilde{A}^{-1}u \wedge x-y \in \tilde{B}^{-1}u \right)$ */

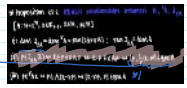
$\leftrightarrow \exists_{y \in \mathcal{H}} \; J_A y \ni u \wedge J_B(x-y) \ni u \; \ldots$ (eq:0)

/* now, A, B: monotone $\Rightarrow J_A, J_B$: firmly nonexpansive $\Rightarrow$ single-valued operator */

$\leftrightarrow \exists_{y \in \mathcal{H}} \; J_A y = u \wedge J_B(x-y) = u$

$\leftrightarrow \exists_{y \in \mathcal{H}} \; u = J_A y = J_B(x-y) \; \ldots$ (eq:1)

/* now



*/

$\leftrightarrow (u, y-u) \in \mathrm{gra}\, A \wedge (u, x-y-u) \in \mathrm{gra}\, B \; \ldots$ (eq:2)

now: $A+B$: maximally monotone $\Rightarrow \mathrm{gra}(A+B) \neq \emptyset$

let. $(z,w) \in \mathrm{gra}(A+B)$

$\leftrightarrow (A+B)z = Az + Bz \ni w$

$\Leftrightarrow \exists_{w_A \in Az} \; \exists_{w_B \in Bz} \; w = w_A + w_B \in Az + Bz$

$\leftrightarrow \exists_{(z,w_A) \in \mathrm{gra}\, A} \; \exists_{(z,w_B) \in \mathrm{gra}\, B} \; w = w_A + w_B \; \ldots$ (eq:3)

A, B: monotone /* A: monotone $\overset{\text{def}}{\Leftrightarrow} \forall_{(\tilde{x},\tilde{u}),(\tilde{y},\tilde{v}) \in \mathrm{gra}\, A} \langle \tilde{x}-\tilde{y} | \tilde{u}-\tilde{v} \rangle \geq 0$ */

$$\left.\begin{array}{l} \langle u-z | y-u-w_A \rangle \geq 0 \\ \langle u-z | x-y-u-w_B \rangle \geq 0 \end{array}\right\} \overset{\oplus}{\Rightarrow} \langle u-z | \underbrace{\cancel{y}-u-w_A + x-\cancel{y}-u-w_B}_{\substack{u \\ x-2u-(w_A+w_B) \\ \underbrace{}_{w}}} \rangle \geq 0 \leftrightarrow \langle u-z | (x-2u)-w \rangle \geq 0 \; \forall_{(z,w) \in \mathrm{gra}(A+B)}$$

$= x-2u-w$

/* now $(z,w) \in \mathrm{gra}(A+B)$, and

recall that $\tilde{A}$: maximally monotone $\overset{\text{def}}{\Leftrightarrow} \forall_{(\tilde{x},\tilde{u})} \left((\tilde{x},\tilde{u}) \in \mathrm{gra}\, \tilde{A} \leftrightarrow \forall_{(\tilde{y},\tilde{v}) \in \mathrm{gra}\, \tilde{A}} \langle \tilde{x}-\tilde{y} | \tilde{u}-\tilde{v} \rangle \geq 0 \right)$ */

set $\tilde{A} := A+B$

$\Leftrightarrow \underbrace{(u, x-2u)}_{(u,\, 2(\frac{x}{2}-u))} \in \mathrm{gra}(A+B)$ /* recall that according to Proposition 23.2 (ii): $(p, \frac{1}{\gamma}(x-p)) \in \mathrm{gra}\, A \leftrightarrow p \in J_{\gamma A}(x)$ */

$\leftrightarrow u = J_{\frac{1}{2}(A+B)}\left(\frac{1}{2}x\right) = \left(J_{\frac{1}{2}(A+B)} \circ \frac{1}{2}\mathrm{Id}\right)(x) \leftrightarrow (x,u) \in \mathrm{gra}\left(J_{\frac{1}{2}(A+B)} \circ \frac{1}{2}\mathrm{Id}\right)$

So, we have proven that:

$\forall_{(x,u)} \left((x,u) \in \mathrm{gra}(J_A \square J_B) \leftrightarrow (x,u) \in \mathrm{gra}\left(J_{\frac{1}{2}(A+B)} \circ \frac{1}{2}\mathrm{Id}\right) \right)$ // as along the way we have maintained $\Leftrightarrow$

$\leftrightarrow \; J_A \square J_B = J_{\frac{1}{2}(A+B)} \circ \frac{1}{2}\mathrm{Id}$

$\blacksquare$

Part 1: Charecterization

Saturday, September 24, 2016 3:34 PM

Proposition 25.1. $[\![A:\mathcal{H}\to 2^{\mathcal{H}},\ B:\mathcal{H}\to 2^{\mathcal{H}},\ \gamma\in\mathbb{R}_{++}]\!] \Rightarrow$

(i) $\operatorname{zer}(A+B)=\operatorname{dom}(A\cap(-B))$

(ii) A,B: monotone $\Rightarrow \operatorname{zer}(A+B)=J_{\gamma B}(\operatorname{Fix} R_{\gamma A} R_{\gamma B})$

(iii) $(C$: closed affine subspace of $\mathcal{H}$; $V:=C-C$; $A=N_C)\Rightarrow$
$\operatorname{zer}(A+B)=\{x\in C \mid V^{\perp}\cap Bx\neq\emptyset\}$

(iv) $(A$: monotone, B: almost single-valued$)\Rightarrow \operatorname{zer}(A+B)=\operatorname{Fix} J_{\gamma A}\circ(\mathrm{Id}-\gamma B)$

9/24/2016 5:27 PM

Proof:

(i) $\forall x\in\mathcal{H}$

$x\in\operatorname{zer}(A+B) \leftrightarrow 0\in(A+B)x=\underbrace{Ax+Bx}_{\exists_u\, u\in Ax} \leftrightarrow \exists_{u\in Ax}\ u+Bx\ni 0 \leftrightarrow \exists_{u\in Ax}\ \underbrace{Bx}_{(-B)x\ni u}\ni -u \leftrightarrow \exists_u\ u\in Ax\cap u\in -Bx$

$\leftrightarrow \exists_u\ u\in(Ax)\cap((-B)x)=(A\cap(-B))(x)$

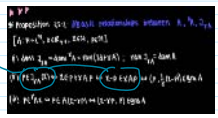
$\leftrightarrow x\in\operatorname{dom}(A\cap(-B))$

$\therefore \operatorname{zer}(A+B)=\operatorname{dom}(A\cap(-B))$

(ii) $\forall x\in\mathcal{H}$

$x\in\operatorname{zer}(A+B) \leftrightarrow 0\in(A+B)x=Ax+Bx \leftrightarrow \gamma Ax+\gamma Bx\ni 0 \leftrightarrow \exists_{u\in\mathcal{H}}\ u\in\gamma Ax,\ -u\in\gamma Bx$ /* let, $x-u=\overset{\exists y}{y} \leftrightarrow u=x-y$ */

$\leftrightarrow \exists_{y\in\mathcal{H}}\ x-y\in\gamma Ax,\ \underbrace{y-x\in\gamma Bx}_{x\in J_{\gamma B}y,\ x=J_{\gamma B}y}$



/* $\because B$: monotone $\Rightarrow J_{\gamma B}$: firmly nonexpansive $\to$ single-valued */

$\leftrightarrow \exists_{y\in\mathcal{H}}\ \boxed{\underbrace{x+x-y}_{2x-y}\in\underbrace{x+\gamma Ax}_{(\mathrm{Id}+\gamma A)x}}$, $x=J_{\gamma B}y$

/* $2x=2J_{\gamma B}y \leftrightarrow 2x-y=2J_{\gamma B}y-y=\underbrace{(2J_{\gamma B}-\mathrm{Id})}_{R_{\gamma B}}y=R_{\gamma B}y$ */ // By definition

$2x-y\in(\mathrm{Id}+\gamma A)x$

$\leftrightarrow \underbrace{(\mathrm{Id}+\gamma A)^{-1}}_{J_{\gamma A}}(2x-y)\ni x \leftrightarrow J_{\gamma A}\underbrace{(2x-y)}_{R_{\gamma B}y}=x \leftrightarrow J_{\gamma A}(R_{\gamma B}y)=x$ // By definition

$\leftrightarrow \exists_{y\in\mathcal{H}}\ x=J_{\gamma A}(R_{\gamma B}y)\ \wedge\ x=J_{\gamma B}y\ \wedge\ 2x-y=R_{\gamma B}y$

$2J_{\gamma A}(R_{\gamma B}y)-y=R_{\gamma B}y$

$\leftrightarrow\ y=2J_{\gamma A}(R_{\gamma B}y)-(R_{\gamma B}y)=\underbrace{(2J_{\gamma A}-\mathrm{Id})}_{R_{\gamma A}}(R_{\gamma B}y)=R_{\gamma A}R_{\gamma B}y$ /* By definition */

$\leftrightarrow \exists_{y\in\mathcal{H}}\ y=R_{\gamma A}R_{\gamma B}y\ \wedge\ x=J_{\gamma B}y$

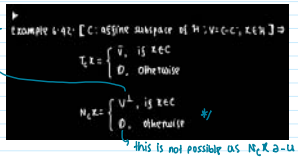
$\leftrightarrow \exists_{y\in\mathcal{H}}\ y\in\operatorname{Fix} R_{\gamma A}R_{\gamma B},\ x=J_{\gamma B}y$

$\leftrightarrow \exists_{y\in\operatorname{Fix} R_{\gamma A}R_{\gamma B}}\ x=J_{\gamma B}y \leftrightarrow x\in J_{\gamma B}(\operatorname{Fix} R_{\gamma A}R_{\gamma B})$

$\therefore \operatorname{zer}(A+B)=J_{\gamma B}(\operatorname{Fix} R_{\gamma A}R_{\gamma B})$

(iii) $\forall x\in\mathcal{H}$

$x\in\operatorname{zer}(\underbrace{A+B}_{N_C})=\operatorname{zer}(N_C+B) \leftrightarrow (N_C+B)x=\underbrace{N_Cx}_{\exists_u\ \exists -u}+\underbrace{Bx}_{\exists u}\ni 0 \leftrightarrow \exists_{u\in Bx}\ (-u\in N_Cx=V^{\perp}\ \wedge\ x\in C)$ /*:



this is not possible as $N_Cx\ni -u$

$\leftrightarrow (x\in C\ \wedge\ \exists_{u\in Bx}\ u\in V^{\perp})$ // as $V^{\perp}$ linear $u,-u\in V^{\perp}$

$\therefore \operatorname{zer}(A+B)=\{x\in C \mid V^{\perp}\cap Bx\neq\emptyset\}$

(iv) $\forall x\in\mathcal{H}$

$x\in\operatorname{zer}(A+B)$

$\leftrightarrow (A+B)x=Ax+Bx\ni 0$

$\leftrightarrow Ax\ni -Bx$ /* as B: single valued $Bx\ni Ax$ not possible for general A */

$\leftrightarrow \gamma Ax\ni\gamma Bx \leftrightarrow \underbrace{x-\gamma Bx}_{(\mathrm{Id}-\gamma B)x}\in\underbrace{x+\gamma Ax}_{(\mathrm{Id}+\gamma A)x} \leftrightarrow \underbrace{\underbrace{(\mathrm{Id}+\gamma A)^{-1}}_{J_{\gamma A}}\underbrace{(\mathrm{Id}-\gamma B)}_{\text{single-valued as } B \text{ single-valued}}x}_{\text{single-valued}}\ni x \leftrightarrow J_{\gamma A}(\mathrm{Id}-\gamma B)x=x \leftrightarrow x\in\operatorname{Fix}(J_{\gamma A}(\mathrm{Id}-\gamma B))$

$\therefore \operatorname{zer}(A+B)=(J_{\gamma A}\circ(\mathrm{Id}-\gamma B))$

■

2016-09-26 12:45 PM

Proposition 25.2. $[\![A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone; $B:\mathcal{H}\to\mathcal{H}$; $\gamma\in\mathbb{R}_{++}]\!]$ $\operatorname{Fix} J_A\circ(\mathrm{Id}+\gamma(B-\mathrm{Id}))=\operatorname{Fix} J_{\gamma^{-1}A}\circ B$

Proof:

$\forall x\in\mathcal{H}$

$x\in\operatorname{Fix} J_{\gamma^{-1}A}\circ B$

$\leftrightarrow (J_{\gamma^{-1}A}\circ B)x=\underbrace{J_{\gamma^{-1}A}}_{(\mathrm{Id}+\gamma A)^{-1}\ \text{// By definition}}(Bx)\ni x \leftrightarrow \overbrace{J_{\gamma^{-1}A}(Bx)}^{\text{single-valued as } A \text{: maximally monotone}}=x \leftrightarrow \underbrace{Bx}_{\text{single-valued}}-x\in\gamma^{-1}Ax$ // using $p\in J_{\gamma A}x \leftrightarrow x-p\in\gamma Ap$

$\leftrightarrow \gamma\underbrace{(B-\mathrm{Id})x}_{}\in Ax$

$\leftrightarrow \underbrace{x+\gamma(B-\mathrm{Id})x}_{(\mathrm{Id}+\gamma(B-\mathrm{Id}))x}\in x+\gamma Ax=(\mathrm{Id}+\gamma A)x$

$\leftrightarrow \underbrace{\underbrace{(\mathrm{Id}+\gamma A)^{-1}}_{=J_A\ \text{// By definition}}\underbrace{((1-\gamma)\mathrm{Id}+\gamma B)}_{\text{single-valued}}}_{\text{single-valued}}x=x \leftrightarrow (J_A\circ((1-\gamma)\mathrm{Id}+\gamma B))x=x$

$\leftrightarrow x\in\operatorname{Fix}(J_A\circ((1-\gamma)\mathrm{Id}+\gamma B))$

$\therefore \operatorname{Fix} J_{\gamma^{-1}A\circ B}=\operatorname{Fix} J_A\circ(\mathrm{Id}+\gamma(B-\mathrm{Id}))$

■

Proposition 25.3.

2016-09-26 9:05 PM

[[V: closed linear subspace of $\mathcal{H}$

$B:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone

$(x_n,u_n)_{n\in\mathbb{N}}\subseteq \operatorname{gra} B$

$(x,u)\in\mathcal{H}\times\mathcal{H}$: $x_n\to x, u_n\to u$, $P_{V^\perp}x_n\to 0$, $P_V u_n\to 0$]] $\Rightarrow$

(i) $x\in \operatorname{zer}(N_V+B)$

(ii) $(x,u)\in (V\times V^\perp)\cap \operatorname{gra} B$

(iii) $\langle x_n|u_n\rangle\to\langle x|u\rangle=0$

Proof: we will (ii) and (iii) first

* Proposition 20.50. linear & hence affine $V-V=V$: for a linear subspace $V=V-V$

[C,D: closed affine subspaces of $\mathcal{H}$: $\underbrace{D-D}_{V}=\underbrace{(C-C)^\perp}_{V^\perp}$ $\therefore V^\perp=V^\perp$: vacuously true ✓

A,B: maximally monotone

$(x_n,u_n)_{n\in\mathbb{N}}\subseteq \operatorname{gra} A/B$

$(x,u)\in\mathcal{H}\times\mathcal{H}$]

✓ ✓ ✓ ✓

$(x_n\to x;\ u_n\to u;\ x_n-P_{V^\perp}x_n\to 0;\ u_n-P_V u_n\to 0)\Rightarrow$ | $(x,u)\in(C\times D)\cap \operatorname{gra} A \leftrightarrow (x,u)\in(V^\perp\times V)\cap \operatorname{gra} B$ $\therefore$ (ii) proved ✓

| $\langle x_n|u_n\rangle\to\langle x|u\rangle \leftrightarrow \langle x|u\rangle$ $\therefore$ (iii) proved ✓

(i)

1* Proposition 25.1.(iii). [[C: closed affine subspace of $\mathcal{H}$; $V=C-C$; $A=N_C$;]] $\operatorname{zer}(A+B)=\{\tilde{x}\in C \mid V^\perp\cap B\tilde{x}\neq\emptyset\}$ *1

$V=V-V$ has V: linear

$\updownarrow$

$\operatorname{zer}(N_V+B)=\{\tilde{x}\in V \mid V^\perp\cap B\tilde{x}\neq\emptyset\}$

from (ii)

$(x,u)\in V^\perp\times V$ $\wedge$ $(x,u)\in \operatorname{gra} B$

$\Downarrow$ $\Downarrow$

$x\in V^\perp$ $Bx\ni u\Rightarrow Ax\neq\emptyset$

$Bx\cap V^\perp\neq\emptyset\Rightarrow x\in\operatorname{zer}(N_V+B)$

2016-09-27 5:40 PM

Proposition 25.4. // Notation here $\tilde{x}=(\tilde{x}_i)_{i\in I}\in\tilde{\mathcal{H}}=\bigoplus_{i\in I}\mathcal{H}$, $\tilde{x}_i\in\mathcal{H}$

[$m\in\mathbb{N}$: $m\geq 2$; $I=\{1,\dots,m\}$; $(A_i)_{i\in I}$: maximally monotone, $\mathcal{H}\to 2^{\mathcal{H}}$;

$\tilde{\mathcal{H}}=\bigoplus_{i\in I}\mathcal{H}_i$; $\tilde{D}=\{(x,\dots,x)\in\tilde{\mathcal{H}}: x\in\mathcal{H}\}$; $\tilde{J}: x\to\tilde{D}: x\mapsto(x,x,\dots,x)$; $\tilde{B}=\bigtimes_{i\in I}A_i$]

$\forall\tilde{x}=(\tilde{x}_i)_{i\in I}\in\tilde{\mathcal{H}}$

(i) $\tilde{D}^\perp=\{\tilde{u}\in\tilde{\mathcal{H}} \mid \sum_{i\in I}\tilde{u}_i=0\}$

(ii) $N_{\tilde{D}}\tilde{x}=\begin{cases}\{\tilde{u}\in\tilde{\mathcal{H}}: \sum_{i\in I}u_i=0\}, & \text{if } \tilde{x}\in\tilde{D}\\ \emptyset, & \text{otherwise}\end{cases}$

(iii) $P_{\tilde{D}}\tilde{x}=\tilde{J}\left(\frac{1}{m}\sum_{i\in I}\tilde{x}_i\right)$

(iv) $P_{\tilde{D}^\perp}\tilde{x}=\left(\tilde{x}_i-\frac{1}{m}\sum_{j\in I}\tilde{x}_j\right)_{i\in I}$

(v) $J_{\tilde{B}}\tilde{x}=(J_{A_i}\tilde{x}_i)_{i\in I}$

(vi) $\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)=\operatorname{zer}(N_{\tilde{D}}+\tilde{B})$

Proof.

(i) $\tilde{D}=\{(x,x,\dots,x)\in\tilde{\mathcal{H}}: x\in\mathcal{H}\}$ now by definition $\tilde{D}^\perp=\{\tilde{u}\in\tilde{\mathcal{H}} \mid \forall_{\tilde{x}\in\tilde{D}}\ \langle\tilde{x}|\tilde{u}\rangle=\sum_i\langle x_i|\tilde{u}_i\rangle=\sum_i\langle x|\tilde{u}_i\rangle=\langle x|\sum_i\tilde{u}_i\rangle=0\Leftrightarrow\sum\tilde{u}_i=0\}$ $\therefore\tilde{D}^\perp=\{\tilde{u}\in\tilde{\mathcal{H}} \mid \sum_{i\in I}\tilde{u}_i=0\}$ ☺

$(\tilde{x}_i)_{i\in I}=(x,\dots,x)=(x)_{i\in I}$, $x_i=x$ for $\tilde{x}\in\tilde{D}$

Example 2.1: (Hilbert direct sum notation)

$\bigoplus_{i\in I}\mathcal{H}_i=\{\tilde{x}=(x_i)_{i\in I}\in\bigtimes_{i\in I}\mathcal{H}_i \mid \sum_{i\in I}\|x_i\|^2<+\infty\}$: Hilbert direct sum // when I finite $\Rightarrow\bigoplus_{i\in I}\mathcal{H}_i=\bigtimes_{i\in I}\mathcal{H}_i$

- addition on Hilbert direct sum: $(\tilde{x},\tilde{y})\mapsto(x_i+y_i)_{i\in I}$ // $\tilde{x}+\tilde{y}=(x_i+y_i)_{i\in I}$
- scalar multiplication on Hilbert direct sum: $(\alpha,\tilde{x})\mapsto(\alpha x_i)_{i\in I}$ // $\alpha\tilde{x}=(\alpha x_i)_{i\in I}$
- scalar product on Hilbert direct sum: $(\tilde{x},\tilde{y})\mapsto\sum_{i\in I}\langle x_i|y_i\rangle$, // $\langle\tilde{x}|\tilde{y}\rangle=\sum_{i\in I}\langle x_i|y_i\rangle$
- $\forall_{i\in I}$ $f_i:\mathcal{H}_i\to]-\infty,+\infty]$; I: finite // $\inf_{i\in I}f_i\geq 0$] $\Rightarrow\bigoplus_{i\in I}f_i:\bigoplus_{i\in I}\mathcal{H}_i\to]-\infty,+\infty]:(x_i)_{i\in I}\mapsto\sum_{i\in I}f_i(x_i)$

(ii) first note that $\tilde{D}$: linear subspace of $\tilde{\mathcal{H}}$ as $\forall_{\tilde{x},\tilde{y}\in\tilde{D}}$ $\forall_{\alpha,\beta\in\mathbb{R}}$ $z=\alpha\tilde{x}+\beta\tilde{y}=(\alpha x+\beta y,\dots,\alpha x+\beta y)\in\tilde{D}$, by structure $\tilde{D}$: closed too

$\tilde{x}=(x,\dots,x)$, $\tilde{y}=(y,\dots,y)$

now let us use Example 6.42.

$\tilde{C}$: linear, $\tilde{C}$ has $\tilde{D}$ linear

Example 6.42: [[$\tilde{C}$: affine subspace of $\tilde{\mathcal{H}}$ [$V=\tilde{C}-\tilde{C}$; $\tilde{x}\in\tilde{\mathcal{H}}$] $\Rightarrow$

$\tilde{D}^\perp$, if $\tilde{x}\in\tilde{D}$

$\emptyset$, otherwise

$N_{\tilde{D}}\tilde{x}=\begin{cases}\tilde{D}^\perp, & \text{if } \tilde{x}\in\tilde{D}\\ \emptyset, & \text{otherwise}\end{cases}$

$\therefore\forall_{\tilde{x}\in\tilde{\mathcal{H}}}$ $N_{\tilde{D}}\tilde{x}=\begin{cases}\tilde{D}^\perp, & \text{if } \tilde{x}\in\tilde{D}\\ \emptyset, & \text{else}\end{cases}$ ☺

(iii) take any $\tilde{x}=(\tilde{x}_i)_{i\in I}\in\tilde{\mathcal{H}}$, we want to show that the operator $\tilde{x}=(\tilde{x}_i)_{i\in I}\mapsto\tilde{J}\left(\frac{1}{m}\sum_{i\in I}\tilde{x}_i\right)=P_{\tilde{D}}\tilde{x}$

$p=\frac{1}{m}\sum_{i\in I}\tilde{x}_i$

$\therefore\tilde{p}=\tilde{J}(p)=(p,p,\dots,p)\in\tilde{D}$

denote, $\tilde{y}=\tilde{J}(y): y\in\mathcal{H}$

$=(y,\dots,y)\in\tilde{D}$ $\therefore$ so. $\tilde{y}$: arbitrary point in $\tilde{D}$

*(Corollary 3.22) ★★☆ let $\tilde{V}:=\tilde{D}$

* Projection onto a closed linear subspaces.

[V: closed linear subspace of $\tilde{\mathcal{H}}$, $\tilde{x}\in\tilde{\mathcal{H}}$]

(i) $P_V\tilde{x}$: $P_V\tilde{x}\in\tilde{V}$, $\tilde{x}-P_V\tilde{x}\perp\tilde{V}$

lets test this one.

want to see if $\langle\tilde{x}-\tilde{p}|\tilde{y}\rangle=0$

now: $\langle\tilde{x}-\tilde{p}|\tilde{y}\rangle=\sum_{i\in I}\langle\tilde{x}_i-\underset{\underset{\frac{1}{m}\sum_{i\in I}x_i}{\parallel}}{p}|\underset{\underset{y}{\parallel}}{\tilde{y}_i}\rangle=\sum_{i\in I}\langle\tilde{x}_i-p|y\rangle=\sum_{i\in I}\langle\tilde{x}_i|y\rangle-\underbrace{\sum_{i\in I}\langle p|y\rangle}_{\langle p|y\rangle\sum_{i\in I}1=m\langle p|y\rangle=\langle mp|y\rangle}=\langle\sum_{i\in I}\tilde{x}_i|y\rangle-\langle mp|y\rangle=\langle\underbrace{(\sum_{i\in I}\tilde{x}_i)-mp}_{0}|y\rangle=0$

$\tilde{x}_1^Ty+\tilde{x}_2^Ty+\dots+\tilde{x}_m^Ty=(\tilde{x}_1+\tilde{x}_2+\dots+\tilde{x}_m)^Ty=(\sum_{i\in I}\tilde{x}_i)^Ty=\langle\sum_{i\in I}\tilde{x}_i|y\rangle$

$\therefore P_{\tilde{D}}\tilde{x}=\tilde{J}\left(\frac{1}{m}\sum_{i\in I}\tilde{x}_i\right)=\left(\frac{1}{m}\sum_{i\in I}\tilde{x}_i,\dots,\frac{1}{m}\sum_{i\in I}\tilde{x}_i\right)$

(iv)

*(Corollary 3.22) ★★☆

using the closed linear subspace decomposition rule: $\forall_{\tilde{x}\in\tilde{\mathcal{H}}}$ $\tilde{x}=P_{\tilde{D}}\tilde{x}+P_{\tilde{D}^\perp}\tilde{x}$

(v) $\frac{1}{m}\sum_{i\in I} x_i$ $\qquad$ $\underbrace{\qquad}_{m}$

***(Corollary 3.22)** ★★☆

*** Projection onto a closed linear subspaces.**

[V: closed linear subspace of $\mathcal{H}$, $x\in\mathcal{H}$]

(i) $P_V x$: $P_V x\in V$, $x-P_V x\perp V$

(ii) $\|P_V x\|^2=\langle P_V x|x\rangle$

(iii) $(P_V\in\mathcal{B}(\mathcal{H}),\ \|P_V\|=1)\Leftarrow V\neq\{0\}$

$\|P_V\|=0\Leftarrow V=\{0\}$

(iv) $V^{\perp\perp}=V$

(v) $P_{V^\perp}=Id-P_V$ // i.e., $\forall_{x\in\mathcal{H}}$ $x=P_V x+P_{V^\perp}x$ (Linear subspace decomposition rule)

(vi) $P_{V^\perp}=P_V$

(vii) $\|x\|^2=\|P_V x\|^2+\|P_{V^\perp}x\|^2=d_V^2(x)+d_{V^\perp}^2(x)$

using the closed linear subspace decomposition rule: $\forall_{\tilde{x}\in\tilde{\mathcal{H}}}$ $\tilde{x}=P_{\tilde{D}}\tilde{x}+P_{\tilde{D}^\perp}\tilde{x}$

$\leftrightarrow P_{\tilde{D}^\perp}\tilde{x}=\tilde{x}-P_{\tilde{D}}\tilde{x}=(\tilde{x}_i-\frac{1}{m}\sum_{i\in I}x_i)_{i\in I}$ ☺

(v) Recall

Proposition 23.16 (Resolvent on Hilbert Sum space)

[$(\mathcal{H}_i)_{i\in I}$: totally ordered finite family of real Hilbert spaces

$\mathcal{H}=\bigoplus_{i\in I}\mathcal{H}_i$

$\forall_{i\in I}$ $A_i:\mathcal{H}_i\to 2^{\mathcal{H}_i}$: maximally monotone

$A=\bigtimes_{i\in I}A_i$] $\Rightarrow$

A: maximally monotone

$J_A=\bigtimes_{i\in I}J_{A_i}$

using this $J_{\gamma\tilde{B}}\tilde{x}=(\bigtimes_{i\in I}J_{\gamma A_i})(\tilde{x}_1,\dots,\tilde{x}_m)=(J_{\gamma A_i}\underbrace{\tilde{x}_i}_{\in\mathcal{H}})_{i\in I}$ ☺

$\tilde{B}=(\bigtimes_{i\in I}A_i)$

(vi) $\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)=\tilde{J}(\{x\in\mathcal{H}\mid x\in\operatorname{zer}\sum_{i\in I}A_i\})=\{\tilde{J}(x)=(x,\dots,x)\in\tilde{\mathcal{H}}\mid x\in\operatorname{zer}\sum_{i\in I}A_i\}$

$\therefore$ $\tilde{J}(x)\in\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)$

$\leftrightarrow$ $x\in\operatorname{zer}(\sum_{i\in I}A_i)\leftrightarrow(\sum_{i\in I}A_i)x=(\sum_{i\in I}A_i x)\ni 0\leftrightarrow\exists_{u_1\in A_1x,\dots,u_m\in A_mx}\ u_1+\dots+u_m=0$

$\leftrightarrow\exists_{\underbrace{(u_i)_{i\in I}}_{u}\in(\bigtimes_{i\in I}(A_ix))}\ \sum_{i\in I}u_i=0$ /+ now from (ii) $\tilde{D}^\perp=\{\tilde{u}\in\tilde{\mathcal{H}}\mid\sum_{i\in I}\tilde{u}_i=0\}$ $\therefore u\in\tilde{D}^\perp\leftrightarrow -u\in\tilde{D}^\perp$ */

(as $\tilde{D}^\perp$: linear subspace)

$\underbrace{(\bigtimes_{i\in I}A_i)}_{\tilde{B}}\underbrace{(x,x,\dots,x)}_{m\text{th}}=\tilde{B}\tilde{J}(x)$

$\tilde{J}(x)=(x,\dots,x)\in\tilde{\mathcal{H}}$ by def

$\leftrightarrow\exists_{u\in\tilde{\mathcal{H}}:u\in\tilde{B}\tilde{J}(x)}\ -u\in\tilde{D}^\perp$

/+ now by definition, $\tilde{J}(x)\in\tilde{D}\Rightarrow N_{\tilde{D}}\tilde{J}(x)=\{u\in\tilde{\mathcal{H}}\mid\sum_{i\in I}\tilde{u}_i=0\}=\tilde{D}^\perp$ by (ii) */

$\leftrightarrow\exists_{u\in\tilde{\mathcal{H}}}$ $u\in\tilde{B}\tilde{J}(x)$, $-u\in\tilde{D}^\perp=N_{\tilde{D}}\tilde{J}(x)$

$\leftrightarrow$ $N_{\tilde{D}}\tilde{J}(x)+\tilde{B}\tilde{J}(x)\ni 0$

$\leftrightarrow$ $(N_{\tilde{D}}+\tilde{B})(\tilde{J}(x))\ni 0\leftrightarrow\tilde{J}(x)\in\operatorname{zer}(N_{\tilde{D}}+\tilde{B})$

$\therefore$ $\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)=\operatorname{zer}(N_{\tilde{D}}+\tilde{B})$ ☺

■

Corollary 25.5. [$m\in\mathbb{N}$: $m\geq 2$; $I=\{1,\dots,m\}$; $(A_i)_{i\in I}$: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$;

$\forall_{i\in I}$ $(x_{i,n},u_{i,n})_{n\in\mathbb{N}}\subseteq\operatorname{gra}A_i$; $(x_i,u_i)\in\mathcal{H}\times\mathcal{H}$;

$\sum_{i\in I}u_{i,n}\to 0$; $\forall_{i\in I}\begin{cases}x_{i,n}\to x_i\\ u_{i,n}\to u_i\\ mx_{i,n}-\sum_{j\in I}x_{j,n}\to 0\end{cases}$] $\Rightarrow\exists_{\tilde{x}\in\operatorname{zer}\sum_{i\in I}A_i}\begin{cases}\text{(i) }\tilde{x}=x_1=\dots=x_m\\ \text{(ii) }\sum_{i\in I}u_i=0\\ \text{(iii) }\forall_{i\in I}\ (\tilde{x},u_i)\in\operatorname{gra}A_i\\ \text{(iv) }\sum_{i\in I}\langle x_{i,n}|u_{i,n}\rangle\to\langle\tilde{x}|\sum_{i\in I}u_i\rangle=0\end{cases}$

2016-09-28 9:46 PM

Proof:

define: $\tilde{\mathcal{H}}=\bigoplus_{i\in I}\mathcal{H}$

$\tilde{D}=\{(x,\dots,x)\in\tilde{\mathcal{H}}:x\in\mathcal{H}\}$: consensus linear space

$\tilde{J}:\mathcal{H}\to\tilde{D}:x\mapsto(x,x,\dots,x)$; $\tilde{B}=\bigtimes_{i\in I}A_i$ // $\tilde{B}\underbrace{\tilde{x}}_{(\tilde{x}_1,\dots,\tilde{x}_m)}=(A_1\tilde{x}_1,A_2\tilde{x}_2,\dots,A_m\tilde{x}_m)$, $\tilde{x}_i\in\mathcal{H}$

now $\tilde{B}=\underbrace{\bigtimes_{i\in I}A_i}_{\text{maximally monotone}}$: maximally monotone using:

Proposition 23.16 (Resolvent on Hilbert Sum space) ★★

[$(\mathcal{H}_i)_{i\in I}$: totally ordered finite family of real Hilbert spaces

$\mathcal{H}=\bigoplus_{i\in I}\mathcal{H}_i$

$\forall_{i\in I}$ $A_i:\mathcal{H}_i\to 2^{\mathcal{H}_i}$: maximally monotone

$A=\bigtimes_{i\in I}A_i$] $\Rightarrow$

A: maximally monotone $\quad J_A=\bigtimes_{i\in I}J_{A_i}$, $\underbrace{(\tilde{x}_i)_{i\in I}}_{\in\mathcal{H}}\in\mathcal{H}$

Define. $(x_i)_{i\in I}=x$, $(u_i)_{i\in I}=u$

$\forall_{n\in\mathbb{N}}\begin{cases}x_n=(x_{i,n})_{i\in I}\\ u_n=(u_{i,n})_{i\in I}\end{cases}\Rightarrow(x_n,u_n)\subseteq\operatorname{gra}B$

monotone

Define. $(x_i)_{i\in I}=x$, $(u_i)_{i\in I}=u$

$$\forall_{n\in\mathbb{N}} \left.\begin{cases} x_n=(x_{i,n})_{i\in I} \\ u_n=(u_{i,n})_{i\in I}\end{cases}\right\} \Rightarrow (x_n,u_n)\subseteq \operatorname{gra} B$$

now given

$$\left(\sum_{i\in I} u_{i,n}\to 0;\ \forall_{i\in I}\begin{cases} x_{i,n}\to x_i \\ u_{i,n}\to u_i \\ m x_{i,n}-\sum_{j\in I} x_{j,n}\to 0\end{cases}\right)$$

applying notation $\boxed{x_n\to x}$ $\boxed{u_n\to u}$

$\Rightarrow \sum_{i\in I} u_{i,n}\to \sum_{i\in I} u_i$, but $\sum_{i\in I} u_{i,n}\to 0$ by given $\Rightarrow \boxed{\sum_{i\in I} u_i=0}$ // as the other possibility is not possible

(ii) proved. :)

$$\left[\forall_{i\in I}\left(x_{i,n}-\frac{1}{m}\sum_{j\in I} x_{j,n}\right)\to 0\right] \leftrightarrow \underbrace{\left(x_{i,n}-\frac{1}{m}\sum_{j\in I} x_{j,n}\right)_{i\in I}}_{P_{\tilde{D}^\perp} x_n}\to 0 \quad \therefore \boxed{P_{\tilde{D}^\perp} x_n\to 0}$$

$$\leftrightarrow \underbrace{\left(\frac{1}{m}\sum_{i\in I} u_{i,n},\dots,\frac{1}{m}\sum_{i\in I} u_{i,n}\right)}_{\tilde{J}_{\tilde{B}}\left(\frac{1}{m}\sum_{i\in I} u_{i,n}\right)}\to(0,\dots,0)$$

$$\leftrightarrow \underbrace{\tilde{J}_{\tilde{B}}\left(\frac{1}{m}\sum_{i\in I} u_{i,n}\right)}_{P_{\tilde{B}} u_n}\to 0\in\tilde{\mathcal{H}} \quad \therefore \boxed{P_{\tilde{B}} u_n\to 0}$$

Proposition 25.4. ♦♦

[$m\in\mathbb{N}$: $m\geq 2$; $I=\{1,\dots,m\}$; $(A_i)_{i\in I}$: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$.

We may name $\sum$ as the distribution operator over cartesian space

$\tilde{\mathcal{H}}=\bigoplus_{i\in I}\mathcal{H}$; $\tilde{D}=\{(x,\dots,x)\in\tilde{\mathcal{H}} : x\in\mathcal{H}\}$; $\tilde{J}:\mathcal{H}\to\tilde{D}: x\mapsto(x,x,\dots,x)$; $B=\bigtimes_{i\in I} A_i$]

We can name $\tilde{D}$ as the cartesian space, as every component exists are the same

$\forall \tilde{x}=(\tilde{x}_i)_{i\in I}\in\tilde{\mathcal{H}}$

(i) $\tilde{D}^\perp=\{\tilde{u}\in\tilde{\mathcal{H}} \mid \sum_{i\in I}\tilde{u}_i=0\}$; (ii) $N_{\tilde{D}}\tilde{x}=\begin{cases}\{\tilde{u}\in\tilde{\mathcal{H}} \mid \sum_{i\in I}\tilde{u}_i=0\}, & \text{if } \tilde{x}\in\tilde{D} \\ \emptyset, & \text{otherwise}\end{cases}$; (iii) $P_{\tilde{D}}\tilde{x}=\tilde{J}\left(\frac{1}{m}\sum_{i\in I}\tilde{x}_i\right)$;

(iv) $P_{\tilde{D}^\perp}\tilde{x}=\left(\tilde{x}_i-\frac{1}{m}\sum_{j\in I}\tilde{x}_j\right)_{i\in I}$; (v) $J_{\gamma\tilde{B}}\tilde{x}=(J_{\gamma A_i}\tilde{x}_i)_{i\in I}$; (vi) $\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)=\operatorname{zer}(N_{\tilde{D}}+\tilde{B})$

Now we use Proposition 25.3.

Proposition 25.3. [V: closed linear subspace of $\mathcal{H}$; $B:\mathcal{H}\to 2^{\mathcal{H}}$ maximally monotone ;
$(x_n,u_n)\subseteq\operatorname{gra} B$; $(x,u)\in\mathcal{H}\times\mathcal{H}$: $x_n\to x$, $u_n\to u$, $P_{V^\perp}x_n\to 0$, $P_V u_n\to 0$]

(i) $x\in\operatorname{zer}(N_V+B)$

(ii) $(x,u)\in(V\times V)\cap\operatorname{gra} B$

(iii) $\langle x_n|u_n\rangle\to\langle x|u\rangle=0$

then in our case $\underset{(\tilde{x},\dots,\tilde{x})}{\tilde{x}}\in\operatorname{zer}(N_{\tilde{D}}+\tilde{B})=\tilde{J}(\operatorname{zer}\sum_{i\in I}A_i)=\tilde{J}\{\tilde{x}\in\mathcal{H} \mid \operatorname{zer}(\sum_{i\in I}A_i)(\tilde{x})\ni 0\}$

$=\{(\tilde{x},\dots,\tilde{x}) \mid \operatorname{zer}(\sum_{i\in I}A_i\tilde{x})\ni 0\}$

$$\leftrightarrow \underbrace{\exists_{\tilde{x}}\ x_1=\dots=x_m=\tilde{x}}_{\text{(i) proved :)}} \wedge \underbrace{\tilde{x}\in\operatorname{zer}\left(\sum_{i\in I}A_i\right)}_{\text{initial goal proven}}$$

now given

$\forall_{i\in I}\ (x_{i,n},u_{i,n})_{n\in\mathbb{N}}\subseteq\operatorname{gra} A_i$; $(x_i,u_i)\in\mathcal{H}\times\mathcal{H}$; $x_{i,n}\to x_i=\tilde{x}$, $u_{i,n}\to u_i$

$(x,u)\in(\tilde{D}^\perp\times\tilde{D})\cap\operatorname{gra}\tilde{B}$

$$\Rightarrow \underbrace{(x,u)}_{(\tilde{x},\dots,\tilde{x})}\in\operatorname{gra}\underbrace{\tilde{B}}_{\bigtimes_{i\in I}A_i}$$

$\leftrightarrow (\bigtimes_{i\in I}A_i)(\tilde{x},\dots,\tilde{x})\ni(u_1,\dots,u_m)$

$\leftrightarrow (A_1\tilde{x},A_2\tilde{x},\dots,A_m\tilde{x})\ni(u_1,\dots,u_m)$

$\leftrightarrow \forall_{i\in I}\ A_i\tilde{x}\ni u_i \leftrightarrow \forall_{i\in I}\ (\tilde{x},u_i)\in\operatorname{gra} A_i$ ...(iii) proved :)

$\langle x_n|u_n\rangle\to\langle x|u\rangle=0$

$$\langle x_n|u_n\rangle=\langle (x_{i,n})_{i\in I}|(u_{i,n})_{i\in I}\rangle=\langle (x_{1,n},\dots,x_{m,n})|(u_{1,n},\dots,u_{m,n})\rangle=(x_{1,n},\dots,x_{m,n})^T(u_{1,n},\dots,u_{m,n})$$

$$=x_{1,n}^T u_{1,n}+\dots+x_{m,n}^T u_{m,n}$$

$$=\sum_{i\in I}\langle x_{i,n}|u_{i,n}\rangle$$

$$\langle x|u\rangle=\langle \underbrace{(x_i)_{i\in I}}_{\tilde{x}}|(u_i)_{i\in I}\rangle \overset{\text{similarly}}{=} \tilde{x}^T u_1+\dots+\tilde{x}^T u_n=\tilde{x}^T(u_1+\dots+u_m)$$

$$=\langle \tilde{x}|\sum_{i\in I}u_i\rangle$$

$\sum_{i\in I}\langle x_{i,n}|u_{i,n}\rangle\to\langle\tilde{x}|\sum_{i\in I}u_i\rangle=0$:) (iv) proved

∎

Part 2: Douglas Rachford Algorithm

September 29, 2016 4:32 PM

2016-10-05 9:02 AM

Theorem 25.4. (Douglas-Rachford algorithm)

[A,B: maximally monotone operators from $\mathcal{H}$ to $\mathcal{E}^{\mathcal{H}}$, $\mathrm{zer}(A+B)\neq\emptyset$;

$(\lambda_n)_{n\in\mathbb{N}}$: sequence in $[0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n)=+\infty$; $x\in\mathcal{H}_{n+1}$: $x_n\in\mathcal{H}$

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{array} \right\} \text{(eq:25.9)}$$

] $\Rightarrow \exists_z$ $z\in \mathrm{Fix}\, R_{\gamma A} R_{\gamma B}$

(i) $J_{\gamma B} z \in \mathrm{zer}(A+B)$

(ii) $(x_n - y_n)_{n\in\mathbb{N}}$: $(y_n - z_n)$ $\to 0$

(iii) $(x_n)_{n\in\mathbb{N}}$: $x_n \to x$

(iv) $(y_n)_{n\in\mathbb{N}}$: $y_n \to J_{\gamma B} z$

(v) $(z_n)_{n\in\mathbb{N}}$: $z_n \to J_{\gamma B} z$

(vi) (C: closed affine subspace of $\mathcal{H}$; $\mathrm{int}\, H_C \neq \emptyset$) $\Rightarrow$ $(P_C z_n)_{n\in\mathbb{N}}$: $P_C z_n \to J_{\gamma B} z$

(vii) (one of the following holds:

(a) A: uniformly monotone on every nonempty bounded subset of $\mathrm{dom}\, A$

(b) B: uniformly monotone on every nonempty bounded subset of $\mathrm{dom}\, B$) $\Rightarrow$

$(y_n)_{n\in\mathbb{N}}$, $(z_n)_{n\in\mathbb{N}}$: converge strongly to the unique point in $\mathrm{zer}(A+B)$

Proof:

A,B : maximally monotone if now

$\Rightarrow R_{\gamma A}, R_{\gamma B}$: nonexpansive

$\therefore T = R_{\gamma A} R_{\gamma B}$: nonexpansive

// composition of two nonexpansive operators are nonexpansive

Proposition 25.1. [A: $\mathcal{H}\to\mathcal{E}^{\mathcal{H}}$, B: $\mathcal{H}\to\mathcal{E}^{\mathcal{H}}$, $\gamma\in\mathbb{R}_{++}$] $\Rightarrow$

(i) $\mathrm{zer}(A+B) = \mathrm{zer}(A+\gamma B)$

(ii) A,B monotone $\Rightarrow$ $\mathrm{zer}(A+B) = J_{\gamma B}(\mathrm{Fix}\, R_{\gamma A} R_{\gamma B})$ // Note: very special, this is not valid in non-convex setup!

$\therefore \mathrm{zer}(A+B) = J_{\gamma B}(\mathrm{Fix}\, R_{\gamma A} R_{\gamma B}) \neq \emptyset$ ... (eq:C)

$\Uparrow$ // B is given w/

$\therefore \mathrm{Fix}\, T = \mathrm{Fix}\, R_{\gamma A} R_{\gamma B} \neq \emptyset$ // as otherwise (*) cannot hold

$\Leftrightarrow \exists_z$ $z\in \mathrm{Fix}\, T = \mathrm{Fix}\, R_{\gamma A} R_{\gamma B}$ (eq:D) : this proves the initial consequent

now using

Proposition 4.21 // A special composition of two firmly nonexpansive operators being firmly nonexpansive *]

[T_1: $\mathcal{H}\to\mathcal{H}$, firmly nonexpansive

T_2: $\mathcal{H}\to\mathcal{H}$, firmly nonexpansive

$T = \frac{1}{2}(T_1 - I) + \frac{1}{2}(T_2 - I)$] $\Rightarrow$

(i) T: firmly nonexpansive

(ii) $\mathrm{Fix}\, T = \mathrm{Fix}\,(T_1 - I)(T_2 - I)$

(iii) T_1: maximally monotone $\Rightarrow \mathrm{Fix}\, T = \mathrm{Fix}\,(T_1 T_2 - T_1)$

set $J_{\gamma A}, J_{\gamma B}$ both firmly nonexpansive as A,B: maximal monotone

define. $T_1 = J_{\gamma A}$, $T_2 = J_{\gamma B}$ in Proposition 4.21.

define $\tilde{T} = T_1(2T_2 - I) + I - T_2 = J_{\gamma A}(\overbrace{2J_{\gamma B} - I}^{R_{\gamma B}}) + (I - J_{\gamma B}) = J_{\gamma A} R_{\gamma B} + I - J_{\gamma B}$

then: $2\tilde{T} - I = 2(J_{\gamma A} R_{\gamma B} + I - J_{\gamma B}) - I = 2J_{\gamma A} R_{\gamma B} + 2I - 2J_{\gamma B} - I = 2J_{\gamma A} R_{\gamma B} - 2J_{\gamma B} + I = 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B}) + I$

but $2\tilde{T} - I = (2T_1 - I)(2T_2 - I) = (\underbrace{2J_{\gamma A} - I}_{R_{\gamma A}})(\underbrace{2J_{\gamma B} - I}_{R_{\gamma B}}) = 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B}) + I$

$\Leftrightarrow R_{\gamma A} R_{\gamma B} = 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B}) + I \Leftrightarrow J_{\gamma A} R_{\gamma B} - J_{\gamma B} = \frac{1}{2}(R_{\gamma A} R_{\gamma B} - I)$ [eq: 2016-09-30 7:32 AM]

this sort of operators can be manipulated via + like normal numbers

// Check : $R_{\gamma A} R_{\gamma B} x = \left(2(J_{\gamma A} R_{\gamma B} - J_{\gamma B}) + I\right)(x) = 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B})x + x$

$\Leftrightarrow (J_{\gamma A} R_{\gamma B} - J_{\gamma B})x = \frac{1}{2}(R_{\gamma A} R_{\gamma B} x - x) = \frac{1}{2}(R_{\gamma A} R_{\gamma B} - I)x$ ✓!

from (eq:25.9):

$y_n = J_{\gamma B} x_n$

$z_n = J_{\gamma A}(2y_n - x_n) = J_{\gamma A}(2J_{\gamma B} x_n - x_n) = \underbrace{J_{\gamma A}(2J_{\gamma B} - I)}_{R_{\gamma B}} x_n = J_{\gamma A} R_{\gamma B} x_n$ // Abs definition w/ [Eq: 2016-09-30 7:36 AM]

$x_{n+1} = x_n + \lambda_n(z_n - y_n)$

$= x_n + \lambda_n(J_{\gamma A} R_{\gamma B} x_n - J_{\gamma B} x_n)$

$= x_n + \lambda_n(J_{\gamma A} R_{\gamma B} - J_{\gamma B})(x_n)$

$= x_n + \underbrace{\lambda_n}_{// \text{ (1*)}} \frac{1}{2}(R_{\gamma A} R_{\gamma B} - I)(x_n)$

$= x_n + \underbrace{\lambda_n \frac{1}{2}(R_{\gamma A} R_{\gamma B} - I)}_{= T \text{ // Our notation}} x_n = x_n + \lambda_n(Tx_n - x_n)$ $\therefore \forall_{n\in\mathbb{N}}$ $x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$ // $\lambda_n = \frac{\lambda_n}{2}$ ... (25.10)

$= (1-\lambda_n)x_n + \lambda_n Tx_n$ [eq: 2016-09-30 7:36 AM]

For convenience set:

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = R_{\gamma A} x_n \\ u_n = 2y_n - x_n - z_n \end{array} \right. \quad \text{(Eq:25.11)}$$

$$\overline{y_n + u_n = x_n - z_n} \quad \text{(eq: A)}$$

... (25.11)

Theorem 25.5. (Basic inclusion between $A, J^{\gamma}, J_{\gamma B}$ [A: $\mathcal{H}\to\mathcal{E}^{\mathcal{H}}$, $\mathrm{zer}(A+B)\neq\emptyset$, $\gamma\in\mathbb{R}_{++}$] $\Rightarrow$

(i) $\mathrm{dom}\, J_{\gamma A} = \mathrm{dom}\, J^{\gamma A} = \mathrm{ran}(I + \gamma A) = \mathcal{H}$, $\mathrm{ran}\, J_{\gamma A}^{\gamma A}$ closed

(ii) $\mathrm{Fix}\, J_{\gamma A} = \mathrm{Fix}\, J^{\gamma A} = \mathrm{zer}(A)$ // $\mathrm{Fix}\, J_{\gamma A} = \mathrm{Fix}\, J^{\gamma A} = \mathrm{zer}(A+\gamma A) = \mathrm{zer}(A)$

(iii) $\mathrm{zer}(A+B) = \mathrm{Fix}\, R_{\gamma A} R_{\gamma B}$ // $\mathrm{zer}(A+B) = \mathrm{zer}(A+\gamma B) = \mathrm{Fix}\, J_{\gamma A} R_{\gamma B}$ ✓

Main from (eq:25.9):

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = J_{\gamma B} x_n \Leftrightarrow x_n - y_n \in \gamma B y_n \Leftrightarrow (y_n, \overbrace{\tfrac{x_n - y_n}{\gamma}}^{v_n}) = (y_n, v_n) \in \mathrm{gra}\, \gamma B \\ z_n = J_{\gamma A}(2y_n - x_n) \Leftrightarrow \underbrace{2y_n - x_n - z_n}_{u_n} \in \gamma A z_n \Leftrightarrow (z_n, 2y_n - x_n - z_n) = (z_n, u_n) \in \mathrm{gra}\, \gamma A \end{array} \right\} \cdots \text{(eq:B)} \quad \text{(Eq: B)}$$

(i) from (eq:C) and (eq:D) : $\left| \begin{array}{l} \mathrm{zer}(A+B) = J_{\gamma B}(\underbrace{\mathrm{Fix}\, R_{\gamma A} R_{\gamma B}}_{= T}) \neq \emptyset \\ \exists_{z\in\mathcal{H}} \quad z\in \mathrm{Fix}\, T \end{array} \right.$

$\mathrm{zer}(A+B) = J_{\gamma B}(\{\tilde{z} : \tilde{z}\in \mathrm{Fix}\, T\}) = \{J_{\gamma B}\tilde{z} \mid \tilde{z}\in \mathrm{Fix}\, T\} \ni J_{\gamma B} z$

$\therefore J_{\gamma B} z \in \mathrm{zer}(A+B)$ ... (i) proved ↓

(ii) From (eq:25.10):

$\forall_{n\in\mathbb{N}}$ $x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$ // $\lambda_n \in \{\frac{\lambda_n}{2}\}$

λ_n : relaxation $\in [0,1]$ as $\lambda_n \in [0,2]$

So, via KM iteration convergence

$\therefore (Tx_n - x_n)_{n\in\mathbb{N}}$: strongly converges to 0.

$R_{\gamma A} R_{\gamma B}$

Theorem 5.14. (KM algorithm):

[D: nonempty closed convex subset of $\mathcal{H}$,

T: $D \to D$, nonexpansive, $\mathrm{Fix}\, T \neq \emptyset$

$(\lambda_n)_{n\in\mathbb{N}}$: sequence in $[0,1]$, $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$

$x_1 \in D$

$\forall_{n\in\mathbb{N}}$ $x_{n+1} = (1-\lambda_n)x_n + \lambda_n Tx_n$ // KM iteration w/] $\Rightarrow$

(i) $(x_n - Tx_n)_{n\in\mathbb{N}}$: $x_n + \lambda_n(Tx_n - x_n)$

(ii) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\mathrm{Fix}\, T$

(iii) $(x_n - Tx_n)_{n\in\mathbb{N}}$: converges strongly to 0

(iv) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\mathrm{Fix}\, T$

now: $(R_{\gamma A} R_{\gamma B} x_n - x_n)$

$= (R_{\gamma A} R_{\gamma B} - I)x_n$ // [eq: 2016-09-30 7:32 AM] says: $(R_{\gamma A} R_{\gamma B} - I) = 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B})$

$= 2(J_{\gamma A} R_{\gamma B} - J_{\gamma B})x_n$

$= 2(\underbrace{J_{\gamma A} R_{\gamma B} x_n}_{z_n} - \underbrace{J_{\gamma B} x_n}_{y_n})$ // using. [Eq: 2016-09-30 7:36 AM]

$= 2(z_n - y_n)$

$\therefore 2(z_n - y_n)$: strongly converges to zero

$\Leftrightarrow (z_n - y_n)$: strongly converges to zero (B) proved ↓

(iii)

[eq: 25.10] and (*) says : $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\mathrm{Fix}\, T$. ↓

$x_{n+1} = x_n + \underbrace{\lambda_n}_{\in[0,1]}(Tx_n - x_n)$

(iv) want to show: $y_n \to J_{\gamma B} x$

now B: maximally monotone, also recall $y_n = J_{\gamma B} x_n$

$\Rightarrow J_{\gamma B}$: firmly nonexpansive, maximally monotone, $\mathcal{H}\to\mathcal{H}$

$\epsilon_{n+1} = \epsilon_n \alpha^{-1} \mu_n (1-\epsilon_n - \epsilon_n)$
$\in [0,1]$

(iv) now to show: $y_n \rightharpoonup J_{rB} x$.

Now B: maximally monotone, also recall $y_n = J_{rB} z_n$

$\Rightarrow J_{rB}$: firmly nonexpansive, maximally monotone, $\mathcal{H} \to \mathcal{H}$

$\Downarrow$

J_{rB}: nonexpansive on $\mathcal{H}$ $\therefore \forall_{n \in \mathbb{N}}$ $\|y_n - y_n\| = \|J_{rB} x_n - J_{rB} x_n\| \leq \|x_n - x_n\|$ /*

$\Rightarrow \|y_n\| - \|y_n\| \leq \|y_n - y_n\| \leq \|x_n - x_n\|$ */

$\therefore \forall_{n \in \mathbb{N}}$ $\|y_n\| \leq \|y_n\| + \underbrace{\|x_n - x_n\|}_{\text{bounded}}$ // now, $x_n \rightharpoonup x \in \mathrm{Fix}\, T \Rightarrow (x_n)_{n \in \mathbb{N}}$: bounded

$\Rightarrow (y_n)_{n \in \mathbb{N}}$: bounded [Eq: 2016-09-30 9:06 AM]

$\Rightarrow (y_n)_{n \in \mathbb{N}}$: will have a sequential cluster point. //using

$\Leftrightarrow \exists_{(y_{k_n})_{n \in \mathbb{N}}}$ $\boxed{y_{k_n} \rightharpoonup y}$

Now: [Eq: 25.11]: $v_n = x_n - z_n$

$v_n = 2y_n - x_n - z_n$

from (i),(ii): $y_n - z_n \to 0 \Rightarrow y_{k_n} - z_{k_n} \to 0$

$x_n \rightharpoonup x \Rightarrow x_{k_n} \rightharpoonup x$

$\boxed{v_{k_n} = x_{k_n} - z_{k_n} \rightharpoonup x - y}$

$\boxed{z_{k_n} \rightharpoonup y}$ // as $(a_n - b_n \to 0,\ b_n \rightharpoonup b) \Rightarrow a_n \rightharpoonup b$

$w_{k_n} = 2y_{k_n} - x_{k_n} - z_{k_n} \rightharpoonup 2y - x - y = y - x$ $\therefore \boxed{w_{k_n} \rightharpoonup y - x}$

So collecting the results above:

$y_{k_n} \rightharpoonup y,\ z_{k_n} \rightharpoonup y,\ v_{k_n} \rightharpoonup x - y,\ w_{k_n} \rightharpoonup y - x$ ... (eq: 25.12)

Now we are going to use corollary 25.5

$v_n + w_n = y_n - z_n$ // $\because$ (eq: A)

$(z_n, w_n) \in \mathrm{gra}\, A \ \forall n$

$(y_n, v_n) \in \mathrm{gra}\, B$ // $\because$ (eq: B)

(Eq: 25.12)

$(y_n - z_n \to 0$ // from (i)

$v_n + w_n \to 0 \Rightarrow v_{k_n} + w_{k_n} \to 0$, $x_{k_n} \rightharpoonup y$, $z_{k_n} \rightharpoonup y$, $\|z_{k_n} - (y_{k_n} + z_{k_n})\| = \|y_{k_n} - z_{k_n}\| \to 0$ // from (i)

and $\|z_{k_n} - (y_{k_n} + z_{k_n})\| = \|z_{k_n} - y_{k_n}\| \to 0$

result from

So, all of the antecedents of corollary 25.5 hold:

$\therefore \exists_{\bar{y} \in \mathrm{zer}(\mathrm{rA} + \mathrm{rB}) = \mathrm{zer}(\mathrm{A+B})}$

$\bar{y} = y = y$

$(\bar{y}, y - x) = (y, y - x) \in \mathrm{gra}\, \mathrm{rA}$

$(\bar{y}, x - y) = (y, x - y) \in \mathrm{gra}\, \mathrm{rB}$

$\Leftrightarrow x - y \in \mathrm{rB}\, y \Leftrightarrow y \in J_{rB} x$ // but J_{rB}: firmly nonexpansive i.e.; single-valued

$\Leftrightarrow y = J_{rB} x \wedge y \in \mathrm{dom}\, \mathrm{rA} \Leftrightarrow y \in \mathrm{dom}\, A$ ... (eq: 25.14) (eq: 25.14)

$\therefore y_{k_n} \rightharpoonup y = J_{rB} x$

But $(y_n)_{n \in \mathbb{N}}$: arbitrary convergent subsequence converging to one point $J_{rB} x$

$\Leftrightarrow (y_n)_{n \in \mathbb{N}}$: possesses one weak sequential cluster point [Eq: 2016-09-30 9:04 AM]

So, [Eq: 2016-09-30 9:06 AM] and [Eq: 2016-09-30 9:04 AM] says: $(y_n)_{n \in \mathbb{N}}$: bounded, $(y_n)_{n \in \mathbb{N}}$: possesses one weak sequential cluster point /*

$\Leftrightarrow y_n \rightharpoonup y = J_{rB} x$ ☺

(v) from (i): $\left.\begin{array}{l} y_n - z_n \to 0 \\ y_n \rightharpoonup J_{rB} x \end{array}\right\} \Rightarrow z_n \rightharpoonup J_{rB} x$ // $(a_n - b_n \to 0,\ b_n \rightharpoonup b) \Rightarrow a_n \rightharpoonup b$

☺

(vi) from (ii): $x_n \rightharpoonup x$

Approach proposition 4.11(i): [illegible]

as C: closed affine subspace of $\mathcal{H}$

$\Rightarrow P_C$: weakly continuous // $x_n \rightharpoonup x \Rightarrow P_C x_n \rightharpoonup P_C x$ // $\therefore P_C x_n \rightharpoonup P_C x$

C: closed affine subspace

Let now

[illegible]

$x_{n+1} = \frac{1}{2}(x_n + T x_n)$

$P_C x_n \rightharpoonup J_{rB} x$ // now given: $N_C = A$, also: $N_C = \mathrm{ri}\, N_C$

as $N_C x = \begin{cases} \{u \in \mathcal{H} \mid \sup \langle c - x, u \rangle \leq 0\}, & x \in C \\ \emptyset, & \text{else} \end{cases}$

$\mathrm{ri}\, N_C x = \begin{cases} \{u \in \mathcal{H} \mid \sup \langle c - x, u \rangle < 0\} \Leftrightarrow \sup \langle c - x, u \rangle = 0 \} = \{\vec{u} \in \mathcal{H} \mid \sup \langle c - x, \vec{u} \rangle > 0\} = N_C x, & x \in C \\ \emptyset & \text{else} \end{cases}$ $\therefore N_C = \mathrm{ri}\, N_C$

$x^{\sharp}_{rN_C} x \rightharpoonup J_{rN_C} x$ $\therefore \boxed{P_C x \rightharpoonup J_{rB}}$

$\boxed{P_C x_n \rightharpoonup P_C x = J_{rB} x}$

now let us use:

$T = \mathrm{rA}, T^{\sharp} = \mathrm{rB}$

// [illegible]

$\therefore x \in \mathrm{Fix}\, T \Leftrightarrow x \in \{\bar{x} \in \mathcal{H} \mid \underbrace{J_{rB}}_{P_C} \bar{x} = x, P_C \bar{x} = x\} \Leftrightarrow \boxed{P_C x = J_{rB} x}$

$\boxed{P_C x_n \rightharpoonup J_{rB} x}$ ☺

(vii) let $y = J_{rB} x$

from (i): $J_{rB} x \in \mathrm{zer}(\mathrm{A+B})$

from (ii): $x_n \rightharpoonup x$

now as uniform monotonicity $\Rightarrow$ strong monotonicity

as $A \vee B$: uniformly monotone (by (a) or (b)) we have $A \vee B$: strictly monotone

$\therefore A + B$: strictly monotone // A: monotone, B: strictly monotone $\Rightarrow A + B$: strictly monotone

$\Rightarrow \mathrm{zer}(\mathrm{A+B})$ = singleton // proposition 23.35: $[A: \mathcal{H} \to 2^{\mathcal{H}}]$ zer A: singleton */

$\mathrm{zer}(\mathrm{A+B}) = \{y\} = \{J_{rB} x\}$

now from (ii): $x_n \rightharpoonup x$

- (iv): $y_n \rightharpoonup J_{rB} x = y$
- (v): $z_n \rightharpoonup J_{rB} x = y$

(eq: 25.15)

[Eq: 25.11] says: $v_n = y_n - z_n \rightharpoonup x - y$

$w_n = 2y_n - x_n - z_n \rightharpoonup 2y - x - y = y - x$ } (25.15)

from [Eq: B]:

$(z_n, w_n) \in \mathrm{gra}\, \mathrm{rA}$: $(z_n, w_n) \rightharpoonup (y, y - x)$

$(y_n, v_n) \in \mathrm{gra}\, \mathrm{rB}$: $(y_n, v_n) \rightharpoonup (y, x - y)$

} using corollary 25.5

$(y, y - x) \in \mathrm{gra}\, \mathrm{rA}$

$(y, x - y) \in \mathrm{gra}\, \mathrm{rB}$ } (eq: 25.16)

Corollary 25.5 [illegible]

now let us consider possibility (vii)(a): A uniformly monotone on every nonempty bounded subset of $\mathrm{dom}\, A$

set. $C = \{x = J_{rB} x\} \cup \{z_n\}_{n \in \mathbb{N}}$

from (vi): $z_n \rightharpoonup J_{rB} x \Rightarrow z_n$: bounded sequence

from [eq: A]: $(z_n, w_n) \in \mathrm{gra}\, \mathrm{rA} \Leftrightarrow w_n \in A z_n \Leftrightarrow A z_n \ni w_n \Leftrightarrow z_n \in \mathrm{dom}\, A$

$\Leftrightarrow (z_n)_{n \in \mathbb{N}}$: bounded sequence on $\mathrm{dom}\, A$

From [eq: 25.14]: $x \in \mathrm{dom}\, A \leftrightarrow C = \{x\} \cup \{z_n\}_{n \in \mathbb{N}}$: bounded subset on $\mathrm{dom}\, A$

[redacted]

Set. $C := \{x \in_{FB} X\} \cup \{z_n\}_{n\in\mathbb{N}}$

from (iv): $z_n \to \mathcal{Z}_{FB} x \Rightarrow z_n$: bounded sequence

from (eq:A): $(u_n, w_n) \in \operatorname{gra} A \Leftrightarrow x \in z_n \operatorname{bin} u_n \Leftrightarrow A z_n \ni \frac{1}{2} w_n \Leftrightarrow z_n \in \operatorname{dom} A$

$\Rightarrow \{z_n\}_{n\in\mathbb{N}}$: bounded sequence on $\operatorname{dom} A$

From (eq:25.16): $x \in \operatorname{dom} A$ — $C := \{x\} \cup \{z_n\}_{n\in\mathbb{N}}$: bounded subset on $\operatorname{dom} A$

Now: recall

$(y,w) \in \operatorname{gra} A \Leftrightarrow y \in Ax$ & $w \in Ay$

$(y,w) \in \operatorname{gra} A \Leftrightarrow y \in Ax$ & $w \in Ay$

$(x,y) \in \operatorname{gra} A \Leftrightarrow y \in Ax$

$(x,y) \in \operatorname{gra} A \Leftrightarrow y \in Ax$

$\forall$ monotone operator A $\forall x \in \operatorname{dom} A$, $\forall y \in Ax$, $\forall w \in Ay$, $\forall \lambda \in \mathbb{R}$ $\langle y-x \mid w-\lambda y \rangle \geq \lambda \|y\|^2$

$\Rightarrow \forall x \in \operatorname{dom} A$, $\forall u_n \in Ax$, $\forall w_n \in Ay$, $\langle y-x \mid w_n - \lambda y \rangle \geq \lambda \|y\|^2$

$\Rightarrow \forall x \in \operatorname{dom} A$, $\forall z_n \in Ax$, $\forall \frac{1}{2} w_n \in Ay$, $\langle y-x \mid \frac{1}{2} w_n - \lambda y \rangle \geq \lambda \|y\|^2$

$\Rightarrow \forall x \in \operatorname{dom} A$, $\forall z_n \in Ax$, $\forall \frac{1}{2} w_n \in Ay$, $\langle y-x \mid \frac{1}{2} w_n - \lambda y \rangle \geq \lambda \|y\|^2$

$\Leftrightarrow \forall x \in \operatorname{dom} A$, $\forall z_n \in Ax$, $\forall \frac{1}{2} w_n \in Ay$, $\langle y-x \mid \frac{1}{2} w_n - \lambda y \rangle \geq \lambda \|y\|^2$

set $C := \operatorname{dom} A$,

$\Rightarrow \langle z_n - y \mid \frac{1}{2}(w_n - 2\lambda x) \rangle \geq \varphi_A(\|z_n - y\|)$ $\forall n \in \mathbb{N}$ — $\varphi_A : \mathbb{R}_+ \to [0,+\infty]$, vanishes only at 0

$\Leftrightarrow \langle z_n - y \mid w_n - 2\lambda x \rangle \geq 2\varphi_A(\|z_n - y\|)$ $\forall n \in \mathbb{N}$ ..... (eq:25.19)

As γA: monotone

$\Rightarrow \forall n \in \mathbb{N}$ $\langle y_n - y \mid w_n - 2\lambda y \rangle \geq 0$ ..... (eq:26.18) — (1)

$\forall n \in \mathbb{N}$ $\forall \varphi_A(\|z_n - y\|) \leq \langle z_n - y \mid w_n - 2\lambda x \rangle + \langle y_n - y \mid w_n - 2\lambda y \rangle$

$(z_n - y_n) + (y_n - y)$ Again, the gra chasser ad a time rule from real analysis

$= \langle z_n - y_n \mid w_n - 2\lambda y + 2\lambda \rangle + \langle y_n - y \mid w_n - 2\lambda y \rangle + \langle y_n - y \mid w_n - 2\lambda x \rangle = \langle z_n - y_n \mid w_n - 2\lambda y + 2\lambda \rangle - \langle z_n - y_n \mid w_n - 2\lambda y \rangle + \langle z_n - y_n \mid w_n - 2\lambda x \rangle + \langle z_n - y_n \mid w_n - 2\lambda y \rangle$

$\langle y_n - y \mid w_n - 2\lambda x - 2\lambda y \rangle + \langle y_n - y \mid w_n - 2\lambda y \rangle = \langle y_n - y \mid w_n - 2\lambda x \rangle$

$\therefore \forall n \in \mathbb{N}$ $0 \leq \gamma \varphi_A(\|z_n - y\|) \leq \langle z_n - y_n \mid w_n - 2\lambda x \rangle$

& now: from (i): $z_n - y_n \to 0$

from (26.19): $y_n \to y$

$w_n \to 2x$, $y_n \to y$

$\Rightarrow w_n - y_n \to -2\lambda y - x + x$

$\Leftrightarrow w_n - y_n \to 0$

Lemma 1.41: $\forall (x_n, y_n) \in \operatorname{gra} A$, $\lim_{n\to\infty} \|x_n - y_n\| = 0$ $\Rightarrow$ the same sequence and cluster point of (x_n) and (y_n) must be fixed point

$\langle z_n - y_n \mid w_n - 2\lambda x \rangle \to 0$ */

$\Rightarrow$ as $n \to \infty$,

$0 \leq \gamma \varphi_A(\|z_n - y\|) \leq 0$

$\therefore$ as $n \to \infty$, $\varphi_A(\|z_n - y\|) = 0$

$\Leftrightarrow$ as $n \to +\infty$ $\|z_n - y\| = 0$

$\Leftrightarrow$ $z_n \to y = \mathcal{Z}_{FB} x$

$\therefore z_n \to \mathcal{Z}_{FB}^{-1}$; $x_n \to \mathcal{Z}_{FB}^{-1}$ As $z_n \to 0$ */

When possibility (viii) (b): the basic is same except roles of A & B are interchanged ∴

∎

Part 3: Parallel splitting alg, Forward-backward alg

October 8, 2016 6:37 PM

Proposition 25.7 (Parallel splitting algorithm)

[$m \in \mathbb{N}: m \geq 2$; $I = \{1, \dots, m\}$; $(A_i)_{i \in I}$: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$; $\operatorname{zer} \sum_{i \in I} A_i \neq \emptyset$;

$(\lambda_n)_{n \in \mathbb{N}}$: $\subset [0,1]$; $\sum_{n \in \mathbb{N}} \lambda_n (1-\lambda_n) = +\infty$; $\forall \in \mathbb{R}_{++}$; $(\gamma_i)_{i \in I} \in \mathbb{R}^m$,

$\forall_{n \in \mathbb{N}}$

$$p_n = \frac{1}{m} \sum_{i \in I} x_{i,n}$$

$$\forall_{i \in I} \quad y_{i,n} = J_{\gamma_i A_i} x_{i,n} \qquad (9.25.21)$$

$$q_n = \frac{1}{m} \sum_{i \in I} y_{i,n}$$

$$\forall_{i \in I} \quad x_{i,n+1} = x_{i,n} + \lambda_n (2q_n - p_n - y_{i,n}) \quad] \Rightarrow$$

$(p_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\operatorname{zer} \sum_{i \in I} A_i$

2016-10-08 9:18 PM

Proof:

define: $\tilde{\mathcal{H}} = \bigoplus_{i \in I} \mathcal{H}$

$\tilde{D} = \{(x, \dots, x) \in \tilde{\mathcal{H}} : x \in \mathcal{H}\}$: consensus closed linear space $\Rightarrow$ closed convex set

$\tilde{J} : \mathcal{H} \to \tilde{D} : x \mapsto (x, x, \dots, x)$: $\tilde{B} = \times_{i \in I} A_i$ // $\tilde{B}\tilde{x} = (\underbrace{b_1 \tilde{x}_1}_{\in \mathcal{H}}, \underbrace{b_2 \tilde{x}_2}_{\in \mathcal{H}}, \dots, \underbrace{b_m \tilde{x}_m}_{\in \mathcal{H}})$

now $\tilde{B} = \times_{i \in I} A_i$: maximally monotone using:

maximally monotone

Proposition 23.16 Parallel sum of finitely many maximally monotone operators is maximally monotone.

[$(A_i)_{i \in I}$: finitely many maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$;

$\forall_{i \in I}$ A_i : maximally monotone

$A = \sum_{i \in I} A_i$] $\Rightarrow$

1. A : maximally monotone

$\forall_{i \in I}$ $\sum_{i \in I} A_i$ $\quad (\sum_{i \in I} A_i)^{-1} = \sum_{i \in I} A_i^{-1}$

define $\tilde{A} = N_{\tilde{D}}$: maximally monotone

// : the normal cone operator of a

//nonempty closed convex set is maximally

//monotone : Example 20.41.

Using Proposition 25.4:

Monotone 25.4 ★★★

[$m \in \{1, 2, \dots\}$; $\mathcal{H}_1, \dots, \mathcal{H}_m$: Hilbert spaces ; $(A_i)_{i \in I}$: maximally monotone operator from $\mathcal{H}_i$ to $2^{\mathcal{H}_i}$;

$\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i$; $\tilde{D} = \{(x_1, \dots, x_m) \in \mathcal{H} : x_1 = \dots = x_m\}$; $\tilde{J} : \mathcal{H} \to \tilde{D} : x \mapsto (x, x, \dots, x)$; $A = \sum_{i \in I} A_i$]

$\tilde{A} = \sum_{i \in I} A_i$; $\tilde{D} = \{(x, \dots, x) \in \mathcal{H} : x \in \mathcal{H}\}$; $\tilde{J} : \mathcal{H} \to \tilde{D} : x \mapsto (x, x, \dots, x)$; $\tilde{A} = N_{\tilde{D}}$; $\tilde{J}^* : \mathcal{H} \to \mathcal{H} : (x_1, \dots, x_m) \mapsto \frac{1}{m} \sum_{i \in I} x_i$

$\forall \tilde{x} = (\tilde{x}_i)_{i \in I} \in \mathcal{H}$

$$J_{\tilde{\gamma} \tilde{A}} \tilde{x} = J_{\tilde{\gamma} \sum_{i \in I} A_i} \tilde{x} = J_{\frac{1}{m} \sum_{i \in I} A_i} \tilde{x} \qquad (9.25.18)(i)$$

$$J_{\tilde{\gamma} \tilde{B}} \tilde{x} = (J_{\gamma_i A_i} \tilde{x}_i)_{i \in I} \qquad (9.25.22.a)$$

// using Proposition 23.4

$\tilde{J}(\operatorname{zer} \sum_{i \in I} A_i) = \operatorname{zer}(\tilde{A} + \tilde{B})$

set: $\forall_{n \in \mathbb{N}}$

$$\tilde{x}_n = (x_{i,n})_{i \in I} \in \tilde{\mathcal{H}}$$

$$\tilde{p}_n = \tilde{J}(p_n) \in \tilde{\mathcal{H}}$$

$$\tilde{y}_n = (y_{i,n})_{i \in I} \in \tilde{\mathcal{H}}$$

$$\tilde{q}_n = \tilde{J}(q_n) \in \tilde{\mathcal{H}}$$

// So we have extended the original sequence to the Hilbert Sum space.

now using (9.25.21) and (9.25.21)

setting $\tilde{x} := \tilde{x}_n$ in (9.25.22.b)

$$J_{\tilde{\gamma} \tilde{B}} \tilde{x}_n = \tilde{J}\Big(\underbrace{\frac{1}{m} \sum x_{i,n}}_{p_n \text{ Eq. (25.21)}}\Big) = \tilde{J}(p_n) = \tilde{p}_n$$

setting $\tilde{x} := \tilde{x}_n \in \tilde{\mathcal{H}}$ in (9.25.22.(ii)) we have : $J_{\tilde{\gamma} \tilde{B}} \tilde{x}_n = (J_{\gamma_i A_i} x_{i,n})_{i \in I} = (y_{i,n})_{i \in I} = \tilde{y}_n$

// by (9.25.21)

n $\quad \tilde{x} := \tilde{x}_n \in \tilde{\mathcal{H}}$ in (9.25.22.(i)) we have :

$$J_{\tilde{\gamma} \tilde{A}} \tilde{x}_n = \tilde{J}\Big(\underbrace{\frac{1}{m} \sum_{i \in I} y_{i,n}}_{q_n \text{ by (9.25.21)}}\Big) = \tilde{J}(q_n) = \tilde{q}_n$$

Finally concatenating $(x_{i,n+1} = x_{i,n} + \lambda_n (2q_n - p_n - y_{i,n}))_{i \in I}$ // the individual components are the last given iteration

we have:

$$\tilde{x}_{n+1} = \tilde{x}_n + \lambda_n (2\tilde{J}(q_n) - \tilde{J}(p_n) - \tilde{y}_n) = \tilde{x}_n + \lambda_n (2\tilde{q}_n - \tilde{p}_n - \tilde{y}_n)$$

So, we have the concatenated sequence:

$\forall_{n \in \mathbb{N}}$

$$\tilde{p}_n = J_{\tilde{\gamma} \tilde{A}} \tilde{x}_n$$

$$\tilde{y}_n = J_{\tilde{\gamma} \tilde{B}} \tilde{x}_n$$

$$\tilde{q}_n = J_{\tilde{\gamma} \tilde{A}} \tilde{y}_n$$

$$\tilde{x}_{n+1} = \tilde{x}_n + \lambda_n (2\tilde{q}_n - \tilde{p}_n - \tilde{y}_n) = \tilde{x}_n + \lambda_n (J_{\tilde{\gamma} \tilde{A}}(2\tilde{y}_n - \tilde{x}_n) - \tilde{y}_n)$$

these two are used in $\tilde{x}_{n+1}$, so not needed again

$\tilde{y}_n$

$$2 J_{\tilde{\gamma} \tilde{A}} \tilde{y}_n - J_{\tilde{\gamma} \tilde{A}} \tilde{x}_n - \tilde{y}_n = J_{\tilde{\gamma} \tilde{A}}(2\tilde{y}_n - \tilde{x}_n) - \tilde{y}_n$$

// now $J_{\gamma \tilde{A}} = P_{\tilde{A}}$ where $\tilde{A}$: closed linear subspace

$\Rightarrow J_{\gamma \tilde{A}} = P_{\tilde{A}}$: linear // Corollary 3.22.(iii) : projector on a closed linear subspace is a linear operator */

$\therefore 2 J_{\tilde{\gamma} \tilde{A}} \tilde{y}_n - J_{\tilde{\gamma} \tilde{A}} \tilde{x}_n = J_{\tilde{\gamma} \tilde{A}}(2\tilde{y}_n - \tilde{x}_n)$

$\leftrightarrow \forall_{n \in \mathbb{N}}$

$$\tilde{y}_n = J_{\tilde{\gamma} \tilde{B}} \tilde{x}_n$$

$$\tilde{q}_n = J_{\tilde{\gamma} \tilde{A}}(2\tilde{y}_n - \tilde{x}_n)$$

$$\tilde{x}_{n+1} = \tilde{x}_n + \lambda_n (\tilde{q}_n - \tilde{y}_n)$$

Proposition 25.6 (Douglas-Rachford algorithm)

[A, B : maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, $\operatorname{zer}(A+B) \neq \emptyset$;

$(\lambda_n)_{n \in \mathbb{N}}$: sequence in $[0,1]$; $\sum_{n \in \mathbb{N}} \lambda_n (1-\lambda_n) = +\infty$; $\gamma \in \mathbb{R}_{++}$; $x_0 \in \mathcal{H}$

$\forall_{n \in \mathbb{N}}$

$$y_n = J_{\gamma B} x_n$$

$$z_n = J_{\gamma A}(2y_n - x_n)$$

$$x_{n+1} = x_n + \lambda_n (z_n - y_n) \quad] \Rightarrow \exists_{x \in \operatorname{Fix} R_{\gamma A} R_{\gamma B}} \; R_{\gamma A} R_{\gamma B} x = x$$

$\exists_{x \in \operatorname{Fix} R_{\gamma A} R_{\gamma B}}$ $(J_{\gamma B} x_n) = y_n$

(i) $(y_n)_{n \in \mathbb{N}}$: $y_n = x_n$

(ii) $(z_n)_{n \in \mathbb{N}}$: $z_n = x_n$

(iii) $(x_n)_{n \in \mathbb{N}}$: $x_n = x_n$

(iv) (C : closed affine subspace of $\mathcal{H}$; $A + N_C$) $\Rightarrow$ $(J_{\gamma A} z_n)_{n \in \mathbb{N}}$: $P_C x_n = J_{\gamma A} x$

(v) (one of the following holds:

(a) A : uniformly monotone on every nonempty bounded subset of dom A

(b) B : uniformly monotone on every nonempty bounded subset of dom B) $\Rightarrow$

$(x_n)_{n \in \mathbb{N}}$: $(x_n)_{n \in \mathbb{N}}$: converges strongly to the unique point in $\operatorname{zer}(A+B)$

So, all the results of D-R splitting will hold:

now in (viii) we have:

$\tilde{A}$: closed linear subspace, $\tilde{A} = N_{\tilde{D}}$, $\tilde{A}$ is the

$\Downarrow$ define

other operator

$\therefore \underbrace{P_{\tilde{A}} \tilde{x}_n}_{\tilde{p}_n} \to J_{\tilde{\gamma} \tilde{A}} \tilde{x}^* \in \operatorname{zer}(\tilde{A} + \tilde{B}) = \tilde{J}(\operatorname{zer} \sum_{i \in I} A_i)$

$\Downarrow$

$\tilde{p}_n = \tilde{J}(p_n)$

//where $\tilde{x}^* \in \operatorname{Fix} R_{\tilde{\gamma} \tilde{A}} R_{\tilde{\gamma} \tilde{B}} \tilde{x}$

now $J_{\tilde{\gamma} \tilde{B}} \tilde{x}^* \in \tilde{J}(\operatorname{zer} \sum_{i \in I} A_i) = (\operatorname{zer} \sum_{i \in I} A_i, \dots, \operatorname{zer} \sum_{i \in I} A_i)$

$\leftrightarrow \exists_{x \in \operatorname{zer} \sum_{i \in I} A_i} \; J_{\tilde{\gamma} \tilde{B}} \tilde{x}^* = \tilde{J}(x)$

$\Rightarrow (\tilde{p}_n) = \tilde{J}(p_n) \to \tilde{J}(x) : x \in \operatorname{zer} \sum_{i \in I} A_i$

// by def

$\Leftrightarrow \quad p_n \to x : x \in \operatorname{zer} \sum_{i \in I} A_i$ ■

2016-10-10 6:51 PM

Theorem 25.8 (Forward-backward algorithm)

[$A : \mathcal{H} \to 2^{\mathcal{H}}$: maximally monotone ; $\beta \in \mathbb{R}_{++}$; $B : \mathcal{H} \to \mathcal{H}$: β-cocercive : $\forall \in]0, 2\beta[$; $\delta = \min\{1, \frac{\beta}{\gamma} + \frac{1}{2}\}$;

$(\lambda_n)_{n \in \mathbb{N}}$: sequence in $[0,\delta]$; $\sum_{n \in \mathbb{N}} \lambda_n (\delta - \lambda_n) = +\infty$; $x_0 \in \mathcal{H}$; $\operatorname{zer}(A+B) \neq \emptyset$;

$$\forall_{n \in \mathbb{N}} \left[\begin{array}{l} y_n = x_n - \gamma B x_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n) \end{array} \right\} \text{(9.25.26)} \quad] \Rightarrow$$

(i) $(x_n)_{n \in \mathbb{N}}$: converges weakly to a point in $\operatorname{zer}(A+B)$

(ii) ($\inf_{n \in \mathbb{N}} \lambda_n > 0$, $x \in \operatorname{zer}(A+B)$) $\Rightarrow$ $(Bx_n)_{n \in \mathbb{N}}$: $Bx_n \to Bx$

(iii) ($\inf_{n \in \mathbb{N}} \lambda_n > 0$, one of the following holds:

(a) A : uniformly monotone on every nonempty bounded subset of dom A

(b) B : uniformly monotone on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$(x_n)_{n \in \mathbb{N}}$: converges strongly to the unique point in $\operatorname{zer}(A+B)$

Proof:

Set $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$

$\alpha = \frac{1}{\delta}$

Recall:

/* Remark 4.24.(ii) : T : firmly nonexpansive $\Leftrightarrow$ $T = \frac{1}{2}$ averaged

$\Leftrightarrow J_{\gamma A}$: nonexpansive $\quad T = \frac{1}{2} \mathrm{Id} + \frac{1}{2} R$ */

/* Corollary 4.8 : ($T : \mathcal{H} \to \mathcal{H}$, firmly nonexpansive) $\Leftrightarrow \exists_{A : \mathcal{H} \to 2^{\mathcal{H}}}$ maximally monotone, $\gamma \in \mathbb{R}_{++}$ $\quad T = J_{\gamma A}$ */

Now A : maximally monotone $\Rightarrow J_{\gamma A}$: firmly nonexpansive /* Proposition 23.3 (ii).(b) */

$\Leftrightarrow J_{\gamma A} = \frac{1}{2}$ averaged

using this

$(\mathrm{Id} - \gamma B)$: $\frac{1}{\gamma \beta}$ averaged $\quad \therefore T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$: $\frac{1}{2}$ averaged

using Proposition 4.24

$\frac{1}{2} \quad = \quad \frac{1}{2} \quad = \quad \frac{1}{2} \quad \times 1 \quad = \quad \frac{1}{2}$: averaged

Now A: maximally monotone $\Leftrightarrow J_{\gamma A}$: firmly nonexpansive /* Proposition 23.7 $T(i)$ (i) */

$\Leftrightarrow J_{\gamma A}$: $\frac{1}{2}$-averaged

using this

$(1-\gamma B)$: $\frac{\gamma}{\xi\beta}$-averaged

$\therefore T = J_{\gamma A} \circ (1-\gamma B)$: $\frac{1}{\beta}$-averaged

using Proposition 4.25

$$2-1+\frac{1}{\max\{1,\frac{\gamma}{\xi\beta}\}} = \frac{2}{1+2\min\{1,\frac{\beta}{\gamma\xi}\}} = \frac{2}{2\left(\frac{1}{2}+\min\{1,\frac{\beta}{\gamma\xi}\}\right)} = \frac{1}{\delta} = \kappa : \text{averaged}$$

$\frac{\gamma}{\xi\beta}$-averaged

$\frac{1}{2}\max\{1,\frac{\gamma}{\xi\beta}\} = \frac{1}{2}\frac{1}{\min\{1,\frac{\beta}{\gamma\xi}\}}$

/* as $\max\{a,b\} = \frac{1}{\min\{\frac{1}{a},\frac{1}{b}\}}$

$\Leftrightarrow$ $\boxed{T: \kappa\text{-averaged}}$

/* Proposition 5.15 (iv):

(A: monotone, B: always single-valued) $\Rightarrow$ $\operatorname{zer}(A+B) = \operatorname{Fix} J_{\gamma A}(1-\gamma B)$ */ $\therefore$ $\boxed{\operatorname{Fix} \underbrace{J_{\gamma A}(1-\gamma B)}_{T} = \operatorname{Fix} T = \operatorname{zer}(A+B)}$

from (24.25-26):

$$\forall_{n\in\mathbb{N}} \begin{cases} y_n = x_n - \gamma B x_n = (1-\gamma B)x_n \\ x_{n+1} = x_n + \lambda_n(J_{\gamma A} y_n - x_n) \end{cases}$$

$$= x_n + \lambda_n\big(\underbrace{(J_{\gamma A}\circ(1-\gamma B))}_{T:\ \kappa\text{-averaged}} x_n - x_n\big)$$

$$= x_n + \lambda_n(Tx_n - x_n) \quad \ldots \text{(24.27)}$$

now we use Proposition 5.15 (all the notations are the same)

thus (i), (ii), (iii) will hold:

(i): from Proposition 5.15 (iii):

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer}(A+B) = \operatorname{Fix} T$ and $(x_n)_{n\in\mathbb{N}}$ is fejer monotone w.r.t $\operatorname{Fix} T$. $\cap\Omega$

$Tx_n - x_n \to 0 \quad \ldots \text{(24.28)}$ (eq:B)

$\therefore x_n \to x$: $x \in \operatorname{zer}(A+B)$

(ii): given, $\varepsilon = \inf_{n\in\mathbb{N}} \lambda_n > 0$, $x \in \operatorname{zer}(A+B)$

$$\forall_{n\in\mathbb{N}} \quad \|\underbrace{Tx_n}_{J_{\gamma A}(1-\gamma B)} - x\|^2 = \|J_{\gamma A}(x_n - \gamma B x_n - x) - x\|^2$$

/* $x \in \operatorname{Fix} T = \operatorname{Fix} J_{\gamma A}(1-\gamma B)$

$\Leftrightarrow J_{\gamma A}\underbrace{(1-\gamma B)}_{\text{single-valued}} x = x$ */ single-valued

$$= \|J_{\gamma A}(x_n - \gamma B x_n) - J_{\gamma A}(x - \gamma B x)\|^2 \quad \text{/* } J_{\gamma A}\text{: nonexpansive. */}$$

$$\leq \|(x_n - \gamma B x_n) - (x - \gamma B x)\|^2 \quad \text{//applying the nonexpansiveness to } (x_n - \gamma B x_n), (x - \gamma B x)$$

$$= \|(1-\gamma B)x_n - (1-\gamma B)x\|^2 \quad \text{//now } (1-\gamma B)\text{: } \tfrac{\gamma}{\xi\beta}\text{-averaged}$$

/* Proposition 4.25 (iii): [B: nonempty subset of $\mathbb{H}$; $\tilde{T}: D \to \mathbb{H}$, nonexpansive; $\kappa \in]0,1[$]

set: $\tilde{T} := (1-\gamma B)$, $\tilde{x} := x_n$, $\tilde{y} := x$

$$\forall_{x\in D}\ \forall_{\kappa\in B} \quad \|\tilde{T}\tilde{x} - \tilde{\kappa}\|^2 \leq \|\tilde{x} - \tilde{y}\|^2 - \frac{1-\kappa}{\kappa}\|(1-\tilde{T})\tilde{x} - (1-\tilde{T})\tilde{y}\|^2 \quad */$$

$$\leq \|x_n - x\|^2 - \frac{1-\frac{\gamma}{\xi\beta}}{\frac{\gamma}{\xi\beta}}\|(1-(1-\gamma B))x_n - (1-(1-\gamma B))x\|^2 \quad \text{// } \kappa = \gamma B$$

$$\left(1-\frac{\gamma}{\xi\beta}\right)/\frac{\gamma}{\xi\beta} = \frac{\xi\beta-\gamma}{\xi\beta}\cdot\frac{\xi\beta}{\gamma} = \frac{\xi\beta-\gamma}{\gamma}$$

$$= \|x_n - x\|^2 - \frac{\xi\beta-\gamma}{\gamma}\|\gamma B x_n - \gamma B x\|^2 \quad \text{// } \gamma\text{: linear operator LOL}$$

$$= \|x_n - x\|^2 - \frac{\xi\beta-\gamma}{\gamma}\gamma^2\|Bx_n - Bx\|^2$$

$$= \|x_n - x\|^2 - (\xi\beta-\gamma)\gamma\|Bx_n - Bx\|^2$$

$$\therefore \forall_{n\in\mathbb{N}} \quad \|Tx_n - x\|^2 \leq \|x_n - x\|^2 - \gamma(\xi\beta-\gamma)\|Bx_n - Bx\|^2 \quad \ldots \text{(24.25-27)}$$

$\delta = \min\{1, \frac{\beta}{\gamma}\xi + \frac{1}{2}\}$ now note that $\delta > 1$ as $\delta < \underbrace{1+\frac{1}{2}}_{>1}$ or $\underbrace{\frac{\beta}{\gamma}\xi + \frac{1}{2}}_{\gamma\in]0,2\beta[} = \frac{\beta}{\xi\gamma} + \frac{1}{2} = \overbrace{\left(\frac{\beta}{\gamma}\right)}^{>1}\frac{1}{\xi} + \frac{1}{2} > 1$

$\gamma \in]0,2\beta[\Rightarrow \exists_{\xi\in]0,1[}$ $\gamma = \kappa\cdot 2\beta$

$\gamma \in]0,2\beta[$

//another way of seeing this is $\delta = \frac{1}{\alpha}$ now $\alpha \in]0,1[$ as T: κ-averaged

$\Rightarrow \delta > 1$ */

So, we have: $\forall_{n\in\mathbb{N}}$

$$\|x_{n+1} - x\|^2$$

$$= \|x_n + \lambda_n(Tx_n - x_n) - x\|^2$$

$$= \|\underbrace{x_n + \lambda_n Tx_n - \lambda_n x_n - (1-\lambda_n+\lambda_n)x}_{(1-\lambda_n)x_n - (1-\lambda_n)x + \lambda_n Tx_n - \lambda_n x}\|^2$$

$$= \|(1-\lambda_n)(x_n - x) + \lambda_n(Tx_n - x)\|^2 \quad \text{/* Using Corollary 2.14: } \forall_{x\in\mathbb{H}}\ \forall_{y\in\mathbb{H}}\ \forall_{\alpha\in\mathbb{R}}\ \|\alpha x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2 \quad */$$

$$= (1-\lambda_n)\|x_n - x\|^2 + \lambda_n\|Tx_n - x\|^2 - \lambda_n(1-\lambda_n)\|Tx_n - x - x_n + x\|^2$$

$$\leq \underbrace{(1-\lambda_n)\|x_n - x\|^2}_{\|x_n - x\|^2} + \lambda_n\|x_n - x\|^2 - \lambda_n\gamma(\xi\beta-\gamma)\|Bx_n - Bx\|^2 - \lambda_n\underbrace{(1-\lambda_n)}_{\in[0,1]}\|Tx_n - x_n\|^2$$

$\delta > 1$

$$= \|x_n - x\|^2 - \gamma(\xi\beta-\gamma)\lambda_n\|Bx_n - Bx\|^2 + \lambda_n(\lambda_n - 1)\|Tx_n - x_n\|^2 \quad \|$$

/* as $(\lambda_n)_{n\in\mathbb{N}}$: is bounded below $\Rightarrow$ $\inf$ will exist

$\gamma\beta > 0$ as $\gamma \in]0,2\beta[$

$\varepsilon = \inf_{n\in\mathbb{N}} \lambda_n > 0 \Leftrightarrow \forall_{n\in\mathbb{N}}\ \varepsilon \leq \lambda_n \Leftrightarrow \forall_{n\in\mathbb{N}}\ -\varepsilon \geq -\lambda_n \Leftrightarrow \boxed{\forall_{n\in\mathbb{N}}\ -\gamma(\xi\beta-\gamma)\varepsilon \geq -\gamma(\xi\beta-\gamma)\lambda_n}$

$\|\ \lambda_n \in [0,1] \Leftrightarrow \lambda_n \leq \delta \Leftrightarrow \lambda_n - 1 \leq \delta - 1$ // as $\delta > 1$, $\lambda_n \leq \delta$, $\lambda_n - 1$, $\delta - 1$ are positive

$$\leq \|x_n - x\|^2 - \gamma(\xi\beta-\gamma)\varepsilon\|Bx_n - Bx\|^2 + \delta(\delta-1)\|Tx_n - x_n\|^2$$

$\boxed{\lambda_n(\lambda_n - 1) \leq \delta(\delta-1)}$ recall, $a,b,c,d \in \mathbb{R}$: $(a \geq b, c \geq d) \Rightarrow (ac \geq bd)$

$$\therefore \forall_{n\in\mathbb{N}} \quad \|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \gamma(\xi\beta-\gamma)\varepsilon\|Bx_n - Bx\|^2 + \delta(\delta-1)\|Tx_n - x_n\|^2 \quad \text{(24.25-28)}$$

$$\Leftrightarrow \gamma(\xi\beta-\gamma)\varepsilon\|Bx_n - Bx\|^2 \leq \underbrace{\|x_n - x\|^2 - \|x_{n+1} - x\|^2}_{\downarrow\ 0} + \underbrace{\delta(\delta-1)\|Tx_n - x_n\|^2}_{\downarrow\ 0} \quad \|$$

$\therefore$ as $n \to \infty$

$$\gamma(\xi\beta-\gamma)\varepsilon\|Bx_n - Bx\|^2 \to 0$$

$\Leftrightarrow$ $\boxed{Bx_n \to Bx}$ //

$\therefore \forall_{x\in\operatorname{Fix} T}\ \|x_n - x\|$: converges

$\Leftrightarrow \boxed{\|x_n - x\|^2 - \|x_{n+1} - x\|^2 \to 0}$

$\boxed{\|x_n - x_{n+1}\| \to 0}$

(iii) in (i) we have proven that $\exists_{x\in\operatorname{zer}(A+B)}$ $x_n \to x$

$\Downarrow$

$Ax + Bx \ni 0 \Leftrightarrow \boxed{\forall x \in \exists -\gamma Bx}$ $\boxed{\|\text{as } Bx\text{: singleton}}$

$\boxed{x \in \operatorname{dom} A,\ -Bx \in Ax}$

suppose (iii)(a) holds:

set, $\forall_{n\in\mathbb{N}}$ $z_n \in J_{\gamma A} y_n \Leftrightarrow \forall A\, z_n \ni y_n - z_n$ /* as $p \in J_{\gamma A} x \Leftrightarrow x - p \in \gamma Ap$ using Proposition 23.2-(ii) */

$\Rightarrow$ $\boxed{\forall_{n\in\mathbb{N}}\ \underbrace{x_n - \gamma Bx_n - z_n}_{y_n \text{ via iteration definition}} = y_n - z_n \in \gamma A z_n}$ $\boxed{z_n \in \operatorname{dom} A}$

$\frac{1}{\gamma}(x_n - \gamma Bx_n - z_n) \in Az_n$

now:

$$z_n - x_n = J_{\gamma A}\underbrace{y_n}_{(1-\gamma B)x_n} - x_n = \left(J_{\gamma A}\circ(1-\gamma B)\right)x_n - x_n = Tx_n - x_n \to 0 \quad \|\text{from (eq:B)}$$

$(1-\gamma B)x_n$ // by iteration scheme

$\Leftrightarrow$ $z_n \to x_n$, but $x_n \to x$

$\Rightarrow$ $\boxed{z_n \to x}$ $\Rightarrow$ clearly $\boxed{(z_n)_{n\in\mathbb{N}}\text{: bounded}}$ $(z_n)_{n\in\mathbb{N}}$: bounded in $\operatorname{dom} A$

now: $C = \{z_n\} \cup \{z_n\}_{n\in\mathbb{N}}$: bounded subset of $\operatorname{dom} A$

given, A: uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$

set $C = \operatorname{dom} A$,

$\exists \phi_A : \mathbb{R}_+ \to [0,+\infty]$, increasing, vanishing only at zero

$\{z_n \in C,\ x \in C,\ \frac{1}{\gamma}(x_n - \gamma Bx_n - z_n) \in Az_n,\ -Bx \in Ax\} \Rightarrow \langle z_n - x \mid \frac{1}{\gamma}(x_n - \gamma Bx_n - z_n) + Bx\rangle \geq \phi_A(\|z_n - x\|)$

$$\Leftrightarrow \langle z_n - x \mid \frac{1}{\gamma}(x_n - \gamma Bx_n - z_n) + Bx\rangle \geq \phi_A(\|z_n - x\|)$$

$$\therefore \forall_{n\in\mathbb{N}} \quad \langle z_n - x \mid x_n - z_n + \gamma(Bx - Bx_n)\rangle \geq \gamma\phi_A(\|z_n - x\|) \geq 0$$

$\{\bar{z}_n \in C,\ x \in C,\ \frac{1}{\gamma}(z_n - \gamma B z_n - \bar{z}_n) \in Az_n,\ -Bx \in Ax\} \Rightarrow \langle \bar{z}_n - x \mid \frac{1}{\gamma}(z_n - \gamma B z_n - \bar{z}_n) + Bx\rangle \geq \phi_A(\|\bar{z}_n - x\|)$

$\Leftrightarrow \langle \bar{z}_n - x \mid \frac{1}{\gamma}(z_n - \gamma B z_n - \bar{z}_n + \gamma Bx)\rangle \geq \phi_A(\|\bar{z}_n - x\|)$

$\therefore \forall_{n\in\mathbb{N}} \quad \langle \bar{z}_n - x \mid z_n - \bar{z}_n + \gamma(Bx - Bz_n)\rangle \geq \gamma\phi_A(\|\bar{z}_n - x\|) \geq 0$

now: $\boxed{\begin{array}{l} \bar{z}_n \to \bar{z},\ z_n \to z,\ \underbrace{Bz_n \to Bz}_{\text{given (1)}} \\ \underbrace{\qquad}_{\text{B}} \\ z_n - \bar{z}_n + \gamma(Bz - Bz_n) \to 0 \end{array}}$ $\Rightarrow \langle \bar{z}_n - x \mid z_n - \bar{z}_n + \gamma(Bz - Bz_n)\rangle \to 0$ // using

[redacted]

As $n \to \infty$, $\phi_A(\|\bar{z}_n - x\|) \to 0$

$\Leftrightarrow \|\bar{z}_n - x\| \to 0$ // as ϕ_A: vanishes only at zero */

$\Leftrightarrow \bar{z}_n \to x : x \in \mathrm{zer}(A+B)$ // A: uniformly monotone ⇒ A: strictly monotone, B: monotone } ⇒ A+B: strictly monotone ⇒ zer(A+B): almost singleton. given: zer(A+B) ≠ ∅ ∴ zer(A+B) = {x} */

// $\bar{z}_n \to z$, $Bz_n \to Bz$ // in (1)

$\therefore \boxed{z_n \to x : \mathrm{zer}(A+B) = \{x\}}$ ☺

(ii)(a) holds:
now B: uniformly monotone on every nonempty bounded subset of $\mathcal{H}$
in (i) we have proven that $\boxed{z_n \to x : x \in \mathrm{zer}(A+B)}$

Clearly $(z_n)_{n\in\mathbb{N}}$: bounded w.r.t $\mathrm{zer}(A+B)$
$\Rightarrow (z_n)_{n\in\mathbb{N}}$ " " $\mathcal{H}$
$\Rightarrow C = \{x\} \cup \{z_n\}_{n\in\mathbb{N}}$: bounded in $\mathcal{H}$

$\forall$ uniformly monotone map $\Rightarrow \exists \phi_B : [0,+\infty[$ increasing, vanishes only at 0.

$\{x \in C,\ z_n \in C,\ \underbrace{Bx \in Bx,\ Bz_n \in Bz_n}_{\text{trivial right!}}\} \Rightarrow \langle \bar{z}_n - x \mid Bz_n - Bx\rangle \geq \phi_B(\|\bar{z}_n - x\|) \geq 0$

now $z_n \to x$ // in (i)

$\{Bz_n \to Bx$ // in (ii)$\}$ using

[redacted]

$\Rightarrow \langle \bar{z}_n - x \mid Bz_n - Bx\rangle \to 0$

$\therefore$ As n→∞, $\phi_B(\|\bar{z}_n - x\|) \to 0$

$\Leftrightarrow \|\bar{z}_n - x\| \to 0$ // as ϕ_B: vanishes only at zero

$\Leftrightarrow z_n \to x : \mathrm{zer}(A+B) = \{x\}$ // B: uniformly monotone ⇒ B: strictly monotone, A: monotone } ⇒ A+B: strictly monotone ⇒ zer(A+B): almost singleton. given: zer(A+B) ≠ ∅ ∴ zer(A+B) = {x} */

// A strictly monotone operator has a zero set that is at most singleton

∎

Proposition 25.9: (Linear convergence of forward-backward algorithm)
[D: nonempty closed subset of $\mathcal{H}$;
A: maximally monotone, $\mathrm{dom}A \subseteq D$;
$Q: D \to \mathcal{H}$; $\kappa \in \mathbb{R}_{++}$: $B \in \mathbb{R}_{++}$;
One of the following holds:
(i) A: κ-strongly monotone; Q: β-cocoercive; $\gamma \in]0, 2\beta[$
(ii) B∘Q: B: κ-strongly monotone, β-Lipschitz continuous; $\gamma \in]0, \frac{2\kappa}{\beta^2}[$;
$z_0 \in D$
$\forall_{n\in\mathbb{N}} \left[\begin{array}{l} y_n = z_n - \gamma B z_n \\ z_{n+1} = J_{\gamma A} y_n \end{array}\right\}$ (EQ: 25.34)] ⇒

$(z_n)_{n\in\mathbb{N}}$: converges linearly to the unique point in zer(A+B).

2016-10-11 7:08 PM

Proof: set, $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$ // now [redacted] */

⇒ Fix T = zer(A+B)

[redacted] ⇒ so $\mathrm{dom}\, J_{\gamma A} = \mathcal{H}$, $\mathrm{ran}\, J_{\gamma A} \subseteq \mathcal{H}$ } as $J_{\gamma A}$: single valued

[redacted]

Also by def (1.5): $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$:

$\mathrm{dom}\, T = \underbrace{\mathrm{dom}\, J_{\gamma A}}_{\mathcal{H}} \cap \underbrace{\mathrm{dom}(\mathrm{Id} - \gamma B)}_{\mathrm{dom}B = D}$

$= \mathcal{H} \cap D = D$

[redacted]

$\therefore \mathrm{dom}\, T = D$

$\mathrm{ran}\, T = \bigcup_{x\in D} (J_{\gamma A} \circ (\mathrm{Id} - \gamma B))x = \bigcup_{x\in D} J_{\gamma A}((\mathrm{Id} - \gamma B)x) \subseteq \bigcup_{x\in D} J_{\gamma A} x$ // because in $(\mathrm{Id} - \gamma B)x$ only those $x \in \mathrm{dom}(\mathrm{Id} - \gamma B) \subseteq D$ will be summed */

$\subseteq \mathrm{ran}\, J_{\gamma A}$ // as by definition $\mathrm{ran}\, J_{\gamma A} = J_{\gamma A}(\mathcal{H}) = \bigcup_{x\in\mathcal{H}} J_{\gamma A}(x) \supseteq \bigcup_{x\in D} J_{\gamma A}(x)$ */

$= \mathrm{dom}\, A$ // as $\mathrm{ran}\, J_{\gamma A} = \mathrm{dom}\, A$ by Proposition 23.2 */

$\subseteq D$ // given $\mathrm{dom}\, A \subseteq D$ */

$\therefore \mathrm{ran}\, T \subseteq D$

$\therefore$ T: single valued operator, $D \to D$ // i.e., $\mathrm{dom}\, T = D$, $\mathrm{ran}\, T \subseteq D$ */

Now D: closed subset of $\mathcal{H} \Rightarrow$ D: complete metric space equipped with Euclidean norm
i.e., $(D, \|\cdot\|)$: complete metric space

Now from (EQ: 25.34):

$\forall_{n\in\mathbb{N}} \left[\begin{array}{l} y_n = x_n - \gamma B x_n = (\mathrm{Id} - \gamma B) x_n \\ x_{n+1} = J_{\gamma A} y_n = \underbrace{J_{\gamma A} \circ (\mathrm{Id} - \gamma B)}_{T} x_n = T x_n \end{array}\right\} \Leftrightarrow \forall_{n\in\mathbb{N}}\ z_{n+1} = T z_n$: note that this iteration looks like Banach-Picard iteration, so we are going to use the associated theorem

[redacted]

Notice that T: Lipschitz continuous with constant $[0,1[$, then the results of Banach-Picard iteration will hold and we will be done

Let us assume that (i) holds i.e., A: κ-strongly monotone; Q: β-cocoercive; $\gamma \in]0, 2\beta[$

set $\tau := \frac{1}{\kappa\gamma + 1}$ ($\tau \in \mathbb{R}_{++}$, $\tau \in]0, 2\beta[$)

[redacted]

using this and results modifying the theorem

$J_{\gamma A}$: Lipschitz continuous with constant $\tau = \frac{1}{\kappa\gamma+1}$ and $J_{\gamma A}: D \to D \overset{\text{def}}{\Leftrightarrow} \forall_{x,y\in D} \|J_{\gamma A} x - J_{\gamma A} y\| \leq \tau \|x - y\|$

now B: β-cocoercive, $\gamma \in]0, 2\beta[$ — clearly $\tau \in [0,1[$ as $\kappa\gamma > 0$

[redacted]

$(\mathrm{Id} - \gamma B)$: $\frac{\gamma}{2\beta}$-averaged ⇒ $(\mathrm{Id} - \gamma B)$: nonexpansive, // An averaged operator is nonexpansive.

$(\mathrm{Id} - \gamma B): D \to \mathcal{H}$

$\overset{\text{def}}{\Leftrightarrow} \forall_{x,y\in D} \|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\| \leq \|x - y\|$

We have shown that $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B): D \to D$

$\therefore \forall_{x\in D} \forall_{y\in D} \quad \|Tx - Ty\| = \|J_{\gamma A}\underbrace{(\mathrm{Id} - \gamma B)x}_{\in D} - J_{\gamma A}\underbrace{(\mathrm{Id} - \gamma B)y}_{\in D}\| \leq \tau \underbrace{\|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\|}_{\leq \|x-y\|} \leq \tau \|x - y\| \Leftrightarrow$ T: Lipschitz continuous with constant $\tau \in [0,1[$ ☺

// dom B = dom(Id − γB)

Now let us assume that (ii) holds, i.e., B∘Q: κ-strongly monotone, β-Lipschitz continuous, $\gamma \in]0, 2\kappa/\beta^2[$

A: maximally monotone ⇒ $J_{\gamma A}: \mathcal{H} \to \mathcal{H}$, $J_{\gamma A}$: nonexpansive ⇔ $\forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{H}} \|J_{\gamma A} x - J_{\gamma A} y\|^2 \leq \|x - y\|^2$

now $\forall_{x\in D} \forall_{y\in D} \|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\|^2 = \|(x - y) - \gamma(Bx - By)\|^2$

$= \|x - y\|^2 + \gamma^2 \underbrace{\|Bx - By\|^2}_{\leq \beta^2\|x-y\|^2} - 2\gamma\underbrace{\langle x - y \mid Bx - By\rangle}_{\geq \kappa\|x-y\|^2}$ // B: κ-strongly monotone ⇒ $\langle x - y \mid Bx - By\rangle \geq \kappa\|x - y\|^2$ */

$\leq \|x - y\|^2 + \beta^2\gamma^2\|x - y\|^2 - 2\gamma\kappa\|x - y\|^2$

$= \|x - y\|^2 (1 + \beta^2\gamma^2 - 2\gamma\kappa)$ // now: given, $\gamma \in]0, \frac{2\kappa}{\beta^2}[\Leftrightarrow \gamma < \frac{2\kappa}{\beta^2} \Leftrightarrow \beta^2\gamma^2 < 2\kappa\gamma \Leftrightarrow 1 + \beta^2\gamma^2 - 2\kappa\gamma < 1$

also, $0 \leq \beta \Leftrightarrow -\kappa\gamma\beta \leq -2\beta\gamma\delta - 2\kappa\gamma$

$\therefore 1 + \beta^2\gamma^2 - 2\kappa\gamma \geq 1 + \beta^2\gamma^2 - 2\beta\gamma = (1 - \beta\gamma)^2 \geq 0$

$0 \leq 1 + \beta^2\gamma^2 - 2\kappa\gamma < 1$

$\overset{\text{def}}{\Leftrightarrow} \tau^2 = 1 + \beta^2\gamma^2 - 2\kappa\gamma \in [0,1[$ */

$= \tau^2 \|x - y\|^2 : \tau \in [0,1)$

$\therefore \forall_{x\in D} \forall_{y\in D} \|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\|^2 \leq \tau^2 \|x - y\|^2$

We have shown that $T = J_{\gamma A} \circ (\mathrm{Id} - \gamma B): D \to D$

$\therefore \forall_{x\in D} \forall_{y\in D} \quad \|Tx - Ty\| = \|J_{\gamma A}(\mathrm{Id} - \gamma B)x - J_{\gamma A}(\mathrm{Id} - \gamma B)y\| \leq \tau \|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\| \leq \tau \|x - y\| \Leftrightarrow$ T: Lipschitz continuous with constant $\tau \in [0,1[$ ☺

We have shown that $T = J_{\gamma A^\circ}(\mathrm{Id} - \gamma B) : D \to D$

$\because \forall_{x \in D} \forall_{y \in D} \quad \|Tx - Ty\| = \| J_{\gamma A}\underbrace{(\mathrm{Id} - \gamma B)x}_{\in D} - J_{\gamma A}\underbrace{(\mathrm{Id} - \gamma B)y}_{\in D}\| \leq \underbrace{\|(\mathrm{Id} - \gamma B)x - (\mathrm{Id} - \gamma B)y\|}_{\underset{\in]0,1[}{\leq \tau} \|x - y\|} \leq \tau \|x - y\| \Leftrightarrow T$: Lipschitz continuous with constant $\tau \in]0,1[$ $\because$

$\|\operatorname{dom} B = \operatorname{dom}(\mathrm{Id} - \gamma B)$

■

Part 4: Tseng's algorithm

October 12, 2016 8:14 AM

Theorem 25.10 : (Tseng's algorithm)

[D: nonempty subset of H; A: $D \to 2^H$, maximally monotone, $\operatorname{dom} A \subseteq D$ & $B: H \to 2^H$, monotone, single-valued on D;
A+B: maximally monotone ; C: closed convex, $\subseteq D$, $\operatorname{Cr} \operatorname{zer}(A+B) \neq \emptyset$;
B: $\frac{1}{\beta}$ Lipschitz continuous relative to $\operatorname{Cr} \operatorname{dom} A$, $\beta \in \mathbb{R}_{++}$;
$\gamma \in]0, \frac{1}{\beta}[$;

$$\forall_{n \in \mathbb{N}} \left| \begin{array}{l} y_n = x_n - \gamma B x_n \\ z_n = J_{\gamma A} y_n \\ r_n = z_n - \gamma B z_n \\ x_{n+1} = P_C(x_n - y_n + r_n) \end{array} \right. \quad \text{(eq: 25.36)} \quad]\Rightarrow$$

(i) $(x_n - z_n)_{n \in \mathbb{N}}$: converges strongly to 0

(ii) $(x_n)_{n \in \mathbb{N}}$ and $(z_n)_{n \in \mathbb{N}}$ converges weakly to a point in $\operatorname{Cr} \operatorname{zer}(A+B)$

(iii) (A or B : uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$) $\Rightarrow$

$(x_n)_{n \in \mathbb{N}}$, $(z_n)_{n \in \mathbb{N}}$: converge strongly to the unique point in $\operatorname{Cr} \operatorname{zer}(A+B)$

2016-10-14 10:31 PM

Proof:

(Setup)

First, we will show that in (eq: 25.36) the iterations are well defined and "=" makes sense [as in general they mislike "∈"

// $x_C(x)$ is projection on C

$\forall_{n \in \mathbb{N}} \quad x_n \in C \subseteq D$ [$\because C \subseteq D$

$\Rightarrow \boxed{x_n \in D}$ (eq: 2016-10-13 11:14 AM)

now. $y_n = \underbrace{x_n}_{\in D} - \gamma \underbrace{B x_n}_{\text{one vector as } B: \text{single valued on } D}$

$\in D$: one vector as B: single valued on D

$B x_n$: one vector

so the reality is one defined [as otherwise $\in$ will be more general case.

Proposition 23.6 : [basic relationships between $A, J_{\gamma A}$]

[$A: H \to 2^H$, $x \in H$, $y \in H$]

i) $\operatorname{dom} J_{\gamma A} = \operatorname{dom}^H A = \operatorname{ran}(Id + \gamma A)$; $\operatorname{ran} J_{\gamma A} \subseteq \operatorname{dom} A$

(ii) $y \in J_{\gamma A} x \Leftrightarrow x - y \in \gamma A y \Leftrightarrow x - y \in \gamma A y \Leftrightarrow (y, \frac{1}{\gamma}(x-y)) \in \operatorname{gra} A$

(iii) $y \in J_{\gamma A} x \Leftrightarrow y \in (Id + \gamma A)^{-1} x \Leftrightarrow (x, y) \in \operatorname{gra}(Id + \gamma A)$

Now $z_n = J_{\gamma A} y_n \Rightarrow \boxed{z_n \in \operatorname{ran} J_{\gamma A} = \operatorname{dom} A \subseteq D} \Rightarrow B z_n$: single vector [as B: single valued on D

(eq: 2016-10-13 11:14 AM)

$\Rightarrow \underbrace{z_n - \gamma B z_n}_{\text{single vector}} = r_n$: well-defined

β = makes sense

finally, $x_{n+1} = \underbrace{P_C}_{\text{projection on a closed convex set is unique}}(x_n - y_n + r_n) \in C$

$\therefore$ By induction, $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}}, (z_n)_{n \in \mathbb{N}}, (r_n)_{n \in \mathbb{N}}$: well defined

Define: $v_{n \in \mathbb{N}} \quad v_n = \gamma^{-1}(x_n - z_n) + B z_n - B x_n$ [from (eq: 25.36) $z_n = J_{\gamma A} y_n \Leftrightarrow \gamma A z_n \ni y_n - z_n \Leftrightarrow \boxed{(z_n, y_n - z_n) \in \operatorname{gra}(\gamma A)}$

$\Leftrightarrow A z_n \ni \frac{1}{\gamma}(y_n - z_n)$ #/

[eq: 25.37]

$\therefore v_{n \in \mathbb{N}} \quad v_n = \underbrace{\gamma^{-1}(x_n - z_n) + B z_n}_{\in A z_n} - B x_n = \underbrace{\frac{1}{\gamma}(y_n + \gamma B x_n - z_n)}_{\in A z_n \text{ from eq: 25.36}} + B z_n - B x_n \in A z_n + B z_n = (A+B) z_n \Leftrightarrow (z_n, v_n) \in \operatorname{gra}(A+B)$ (eq: 25.38)

given $\operatorname{Cr} \operatorname{zer}(A+B) \neq \emptyset \Rightarrow \exists_z \quad z \in C \cap \operatorname{zer}(A+B)$

$\Rightarrow z \in C \Leftrightarrow z = J_C z$; $\boxed{z \in D}$; $z \in \operatorname{zer}(A+B) \Leftrightarrow A z + B z \ni 0$

$\Leftrightarrow \gamma A z + \gamma B z \ni 0 \Leftrightarrow \gamma A z \ni -\gamma B z \Leftrightarrow \boxed{(z, -\gamma B z) \in \operatorname{gra} \gamma A}$

// single vector [as B: single-valued on D ; $C \subseteq D$

A: maximally monotone

$\Rightarrow \gamma A$: $\Rightarrow \forall_{(x,u),(y,v) \in \operatorname{dom} A} \quad \langle x - y \mid u - v \rangle \geq 0$

$\langle z_n - z \mid y_n - z_n + \gamma B z \rangle \geq 0$

$\Leftrightarrow \boxed{\langle z_n - z \mid z_n - y_n - \gamma B z \rangle \leq 0}$

Again. B: monotone, single-valued on $D \Rightarrow \gamma B$: monotone, single valued on D

Also $\boxed{z_n \in D}$, $\boxed{z \in D}$

$\Leftrightarrow (z, \gamma B z) \in \operatorname{gra} \gamma B$, $(z_n, \gamma B z_n) \in \operatorname{gra} \gamma B$

$\xRightarrow{\text{B: monotone}} \langle z_n - z \mid \gamma B z_n - \gamma B z \rangle \geq 0$

$\Leftrightarrow \boxed{\langle z_n - z \mid \gamma B z - \gamma B z_n \rangle \leq 0}$ (+)

$\boxed{\langle z_n - z \mid z_n - y_n - \gamma B z \rangle \leq 0}$

Now:

$2\gamma \langle z_n - z \mid B x_n - B z_n \rangle$

$= 2 \langle z_n - z \mid \gamma B x_n - \gamma B z_n \rangle$

$= 2 \langle z_n - z \mid \underbrace{\gamma B x_n + y_n - z_n + z_n - y_n - \gamma B z_n}_{} \rangle$

$= 2 \langle z_n - z \mid \gamma B x_n + y_n - z_n \rangle + 2 \langle z_n - z \mid z_n - y_n - \gamma B z_n \rangle$

//negative number, so removing it will give a larger number

$\leq 2 \langle z_n - z \mid \underbrace{\gamma B x_n + y_n - z_n}_{\gamma B x_n + x_n - \gamma B x_n - x_n = x_n - x_n} \rangle$ /+ from (eq: 25.36) : $y_n = x_n - \gamma B x_n$ #/

$= 2 \langle z_n - z \mid x_n - z_n \rangle$ // now $\|(z_n - z) + (x_n - z_n)\|^2 = \|z_n - z\|^2 + \|x_n - z_n\|^2 + 2 \langle z_n - z \mid x_n - z_n \rangle$ */

$= \|x_n - z\|^2 - \|z_n - z\|^2 - \|x_n - z_n\|^2$

$\therefore 2\gamma \langle z_n - z \mid B x_n - B z_n \rangle \leq \|x_n - z\|^2 - \|z_n - z\|^2 - \|x_n - z_n\|^2 \quad \ldots$ (eq: 25.41)

Now Proposition 4.8 says: [C: nonempty closed convex, $\subseteq H$] P_C: firmly nonexpansive

$\Downarrow$

nonexpansive

So. $\underbrace{\|x_{n+1} - z\|^2}_{P_C(z_n + r_n - y_n) \text{ for } z = z} = \|P_C(z_n + r_n - y_n) - P_C z\|^2 \leq \|z_n + \underbrace{r_n}_{z_n - \gamma B z_n} - \underbrace{y_n}_{x_n - \gamma B x_n} - z\|^2 = \|z_n - \gamma B z_n + \gamma B x_n - z\|^2 = \|(z_n - z) + \gamma(B x_n - B z_n)\|^2$

[By iteration scheme (eq: 25.36)]

$= \|(z_n - z)\|^2 + \gamma^2 \|B x_n - B z_n\|^2 + 2\gamma \langle z_n - z \mid B x_n - B z_n \rangle$

$\leq \|z_n - z\|^2 + \gamma^2 \|B x_n - B z_n\|^2 + 2\gamma \|x_n - z\|^2 - \|z_n - z\|^2 - \|x_n - z_n\|^2$

$= \|x_n - z\|^2 - \|x_n - z_n\|^2 + \gamma^2 \|B x_n - B z_n\|^2$ /+ : B: $\frac{1}{\beta}$ Lipschitz continuous w.r.t $C \subseteq \operatorname{dom} A$. now $x_n \in C$, $z \in C$ (given)

$\leq \frac{1}{\beta^2} \|x_n - z_n\|^2$ $\therefore \|B x_n - B z_n\| \leq \frac{1}{\beta} \|x_n - z_n\|$

$\leq \|x_n - z\|^2 - \underbrace{\left(1 - \frac{\gamma^2}{\beta^2}\right)}_{\text{/+ negative number, so removing it will give only larger number #/}} \|x_n - z_n\|^2$ /+ now $\gamma \in]0, \beta[$, $(1 - \frac{\gamma^2}{\beta^2}) \geq 0$ +/ $\ldots$ (eq: 25.42)

$\leq \|x_n - z\|^2$

$\therefore \forall_{z \in \operatorname{Cr} \operatorname{zer}(A+B)} \quad \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 \xrightarrow{\text{def}} (x_n)_{n \in \mathbb{N}}$: Fejer monotone wrt $\operatorname{Cr} \operatorname{zer}(A+B)$

$\boxed{\text{(i)}}$ from (eq: 25.42) $\forall_{z \in \operatorname{Cr} \operatorname{zer}(A+B)} \forall_{n \in \mathbb{N}} \quad \|x_{n+1} - z\|^2 \leq \|x_n - z\|^2 - (1 - \frac{\gamma^2}{\beta^2}) \|x_n - z_n\|^2$

$\Rightarrow \forall_{n \in \mathbb{N}} \quad 0 \leq (1 - \frac{\gamma^2}{\beta^2}) \|x_n - z_n\|^2 \leq \|x_n - z\|^2 - \|x_{n+1} - z\|^2$

as $n \to \infty$. $\|x_n - z\|^2 - \|x_{n+1} - z\|^2 \to 0$ /+ as.

$\therefore$ as $n \to \infty$. $(1 - \frac{\gamma^2}{\beta^2}) \|x_n - z_n\|^2 \to 0$

$\Leftrightarrow$ as $n \to \infty$, $\|x_n - z_n\|^2 \to 0$

$\therefore x_n - z_n \to 0$ #/

Proposition 4.4 : [weak seq convergence of fejer monotone sequences]

[C: nonempty subset of H

Fix(T) : fejer monotone sequence wrt C, T]

i) $\lim_{n \to \infty} \|x_n - z\|$ exists

ii) $\forall_{z \in \operatorname{Fix}(T)} \|x_n - z\|_{n \in \mathbb{N}}$: decreasing and converges

$\boxed{\text{(ii)}}$ $\forall_{n \in \mathbb{N}}$ $\text{Df } \|z_n - B z_n\| \leq \frac{1}{\beta} \underbrace{\|x_n - z_n\|}_{\to 0 \text{ from (i)}}$

$\therefore$ as $n \to \infty$. $(B x_n - B z_n) \to 0$

now. from (eq: 25.37) $u_n = \frac{1}{\gamma}(x_n - z_n) + (B x_n - B z_n)$

$\therefore u_n \to 0$ (eq: 25.44)

now: $(x_n)_{n \in \mathbb{N}}$: bounded [as $(x_n)_{n \in \mathbb{N}}$: Fejer monotone w.r.t $\operatorname{Cr} \operatorname{zer}(A+B)$ and Fejer monotone sequences are bounded via

// $x_0 \in C$ in (eq: 25.36)

x: a weak sequential cluster point of $(x_n)_{n \in \mathbb{N}}$ // Any bounded sequence has a weakly convergent subsequence: Lemma 2.43

$\Leftrightarrow \exists_{(x_n)_{n \in \mathbb{N}}} \quad x_n \rightharpoonup x$,

6 downA

*1

$\therefore y_n \to 0$ (eq: 25.44)

now: $(x_n)_{n\in\mathbb{N}}$: bounded (as $(x_n)_{n\in\mathbb{N}}$: fejer monotone w.r.t. $C \cap \mathrm{zer}(A+B)$) and fejer monotone sequences are bounded via

// $x_n \in C$ in (eq: 25.36)

x: a weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ // Any bounded sequence has a weakly convergent subsequence: Lemma 2.5

$\hookrightarrow \exists (x_{k_n})_{n\in\mathbb{N}}$ $x_{k_n} \rightharpoonup x$

now we will show $x \in C \cap \mathrm{zer}(A+B)$

first: C: closed, convex, $\subseteq D$ /

$\Rightarrow$ C: sequentially closed

$\Rightarrow$ $\boxed{x \in C}$

now we will show that:

$x \in \mathrm{zer}(A+B) \Leftrightarrow (A+B)x \ni 0$

$\hookrightarrow (x,0) \in \mathrm{gra}(A+B)$

Proposition 2.5: (Weak sequential closedness) Let C be a non-empty convex subset of X. Then C is closed if and only if it is weakly closed.

[C: closed $\Leftrightarrow$ C: weakly closed]

C: weakly sequentially closed $\Leftrightarrow$ C: sequentially closed $\Leftrightarrow$ C: closed $\Leftrightarrow$ C: weakly closed

A: weakly sequentially closed $\Leftrightarrow$ $\forall$ seq. (x_n) in X and every weakly convergent sequence in C to some $x \in C$

C: sequentially closed $\Leftrightarrow$ every $x_n \in C$ and $x_n \to x$ $\Rightarrow$ $x \in C$

C: closed $\Leftrightarrow$ $C = \mathrm{cl}(C)$

C: weakly closed $\Leftrightarrow$ $C = \mathrm{wcl}(C)$ */

in (i) we have shown: $z_{k_n} - z_{k_n} \to 0$

$\Rightarrow$ $\boxed{z_{k_n} - z_{k_n} \to 0}$ // now, $(a_n - b_n \to 0,\ a_n \to a) \Rightarrow b_n \to a$

$\therefore$ $\boxed{z_{k_n} \to x}$

$\boxed{u_{k_n} \to 0}$

(eq: 25.38): $(z_n, u_n) \in \mathrm{gra}(A+B) \Rightarrow (z_{k_n}, u_{k_n}) \in \mathrm{gra}(A+B)$ /

Proposition 23.11: (Graph analysis of sums)

[A: $\mathrm{ran}\,A$, maximally monotone]

(i) $\mathrm{gra}\,A$: sequentially closed in $\mathbb{X}^{\text{strong}} \times \mathbb{X}^{\text{weak}}$

aka $\forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \mathrm{gra}\,A}$ $\forall_{(x,u)\in\mathbb{X}\times\mathbb{X}:\ x_n \to x,\ u_n \rightharpoonup u}$ $(x,u) \in \mathrm{gra}\,A$

(ii) $\mathrm{gra}\,A$: sequentially closed in $\mathbb{X}^{\text{weak}} \times \mathbb{X}^{\text{strong}}$

aka $\forall_{(x_n,u_n)_{n\in\mathbb{N}} \subseteq \mathrm{gra}\,A}$ $\forall_{(x,u)\in\mathbb{X}\times\mathbb{X}:\ x_n \rightharpoonup x,\ u_n \to u}$ $(x,u) \in \mathrm{gra}\,A$

(iii) $\mathrm{gra}\,A$: closed in $\mathbb{X}^{\text{strong}} \times \mathbb{X}^{\text{strong}}$

*/

$(x,0) \in \mathrm{gra}(A+B) \Leftrightarrow$ $\boxed{x \in \mathrm{zer}(A+B)}$

$x \in C \cap \mathrm{zer}(A+B)$

$\therefore \forall_{\underbrace{(x_{k_n})_{n\in\mathbb{N}}}_{\text{any subsequence of } (x_n)_{n\in\mathbb{N}}:\ \text{Fejér monotone w.r.t. } C\cap\mathrm{zer}(A+B)}}$ $x_{k_n} \to x \in C \cap \mathrm{zer}(A+B)$

Subsequence of Bounded Sequence: (2.4)

[Subsequence of Bounded Seq. (x_n) is bounded. Any one subsequence (x_{k_n}) of (x_n) is bounded. Given seq. (x_n) is bounded, $\exists M > 0$ $\|x_n\| \le M$]

$\therefore x_n \to x \in C \cap \mathrm{zer}(A+B)$

as $z_n - z_n \to 0$ //from (i)

$\boxed{x_n \to x \in C \cap \mathrm{zer}(A+B)}$ // $\because (a_n - b_n \to 0,\ a_n \to a) \Rightarrow b_n \to a$ ✓

(eq: 2016-10-14 10:27 PM)

[iii] $A+B$: strictly monotone // $\because$ A: uniformly monotone $\Rightarrow$ strictly monotone

$\Rightarrow$ $\mathrm{zer}(A+B)$: almost a singleton // $\overline{B}$: monotone

// $\overline{A+B}$: strictly monotone

// $\because$ A strictly monotone operator's zero set is almost singleton

given $\mathrm{zer}(A+B) : \neq \emptyset$

$\mathrm{zer}(A+B)$: singleton $\Leftrightarrow$ $C \cap \mathrm{zer}(A+B)$: singleton (eq: 2016-10-14 10:27 PM) (eq: 25.44)

// $\because$ given, $C \cap \mathrm{zer}(A+B) \neq \emptyset$ $\Downarrow$ $z_{k_n} \to x$; and $\Downarrow$ $u_{k_n} \to 0$

in (ii) we have shown that: $\exists_{x \in C \cap \mathrm{zer}(A+B) \subseteq \mathrm{dom}\,A}$

given. $A \cup B$: uniformly monotone on every nonempty bounded subset of $\mathrm{dom}\,A$

$\Rightarrow$ $A+B$: uniformly $\quad\Rightarrow\quad$ $\Rightarrow$ $\{x\} \cup \{x_n\}_{n\in\mathbb{N}} \subseteq \mathrm{dom}\,A$ // as $(x_n)_{n\in\mathbb{N}} \subseteq \mathrm{dom}\,A$, $x \in \mathrm{dom}\,A$

now, $x \in \mathrm{zer}(A+B)$

$\hookrightarrow (x,0) \in \mathrm{gra}(A+B)$

(eq: 25.38) $(z_n, u_n) \in \mathrm{gra}(A+B)$

(eq 22.7): let $K = (A+B)$, $C = \mathrm{dom}\,A$

K: monotone, maximal on C $\Leftrightarrow$ $\forall_{(x,u)\in C\times X}$ $\forall_{(y,v)\in\mathrm{gra}\,K}$ $\langle x-y, u-v\rangle \ge 0 \Rightarrow (x,u) \in \mathrm{gra}\,K$

$\exists_{\phi: \mathbb{R}_+ \to [0,+\infty]}$, vanishes only at zero, increasing $\quad (z_n \in \mathrm{dom}\,A,\ x \in \mathrm{dom}\,A,\ u_n \in (A+B)z_n,\ 0 \in (A+B)x)$ $\underbrace{\langle z_n - x, u_n - 0\rangle}_{\to 0} \ge \chi(\phi(\|z_n - x\|)) \ge 0 \Rightarrow \|z_n - x\| \to 0 \Leftrightarrow z_n \to x$ $\therefore x \in C \cap \mathrm{zer}(A+B) = \{x\}$

$\phi: \mathbb{R}_+ \to [0,+\infty]$, $\phi(t) = t\chi(t)$, $\chi(t) = \begin{cases} 1 & \text{if } t \ge \delta \\ 0 & \text{otherwise} \end{cases}$ $\quad$ ϕ is lsc, convex and vanishes only at zero. $\quad$ $\phi(t) = 0 \Leftrightarrow t = 0$ $\quad$ $\phi(t) \ge \delta\chi(t) \ge \delta\chi(\delta) = \delta$ $\quad$ $\phi(t) \ge \delta$ for $t \ge \delta$ */

Part 5: Variational inequality problem

October 17, 2016 9:05 AM

Definition 25.12. (Variational inequality problem)
[$f \in \Gamma_0(\mathcal{H})$; $B: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]
Associated variational inequality problem:

find $x \in \mathcal{H}$
such that $\exists_{u \in Bx} \ \forall_{y \in \mathcal{H}} \ \langle x-y \mid u \rangle + f(x) \leq f(y)$ (eq.:25.49)

Example 25.13. [$z \in \mathcal{H}$; $B: x \mapsto x - z$] Associated variational inequality problem:

find $x \in \mathcal{H}$
such that $\forall_{y \in \mathcal{H}} \ \langle x-y \mid x-z \rangle + f(x) \leq f(y)$

• solution: $x = \operatorname{prox}_f z$

Example 25.14. [set: $f = \iota_C$ in Definition 25.12., C: nonempty, closed, convex, $\subseteq \mathcal{H}$; $B: \mathcal{H} \to \mathcal{H}$, maximally monotone] Classical variational inequality problem:

find $x \in C$
such that $\forall_{y \in C} \ \langle x-y \mid Bx \rangle \leq 0$ . . . (eq.: 25.51)

// Dual cone:
[$C \subseteq \mathcal{H}$] $C^{\oplus}$: dual cone of $C = \{u \in \mathcal{H} \mid \inf \langle C \mid u \rangle \geq 0\}$
$\forall_{x \in C} \ \langle x \mid u \rangle \geq 0$ #/

Example 25.15. (Complementarity problem) [In example 25.14., set $C := K$: K : nonempty closed convex cone in $\mathcal{H}$] Complementarity problem: $B: \mathcal{H} \to \mathcal{H}$, maximally monotone

find $x \in K$
2016-10-17 11:57 AM such that $x \perp Bx$, $Bx \in K^{\oplus}$

Proof: We want to show that, under our antecedent, Example 25.14 reduces to our goal.

K: nonempty closed convex cone $\overset{\text{def}}{\Leftrightarrow}$ $K = \mathbb{R}_{++} K$

$\therefore x \in K \Rightarrow \begin{cases} x/2 \in K \\ 2x \in K \end{cases}$

from (eq.:25.51) $\forall_{y \in C} \ \langle x-y \mid Bx \rangle \leq 0$ // now $C := K$

$\Leftrightarrow \forall_{y \in K} \ \langle x-y \mid Bx \rangle \leq 0$ (eq.: 1)

$y := x/2 \Rightarrow \langle x - \frac{x}{2} \mid Bx \rangle = \langle \frac{x}{2} \mid Bx \rangle = \frac{1}{2} \langle x \mid Bx \rangle \leq 0 \Leftrightarrow \langle x \mid Bx \rangle \leq 0$

$y := 2x \Rightarrow \langle x - 2x \mid Bx \rangle = \langle -x \mid Bx \rangle = (-1) \langle x \mid Bx \rangle \leq 0 \Leftrightarrow \langle x \mid Bx \rangle \geq 0$

$\Rightarrow \langle x \mid Bx \rangle = 0 \Leftrightarrow \boxed{x \perp Bx}$ // now we can use this as a given $\because (A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow \begin{cases} A \Rightarrow B \\ A \Rightarrow C \end{cases} \Leftrightarrow A \Rightarrow (B \wedge C)$

$\forall_{\underset{K}{y \in C}} \ \langle x-y \mid Bx \rangle = \underbrace{\langle x \mid Bx \rangle}_{0} - \langle y \mid Bx \rangle = -\langle y \mid Bx \rangle \leq 0$

$\Leftrightarrow \forall_{y \in C = K} \ \langle y \mid Bx \rangle \geq 0$ // $K^{\oplus} = \{u \in \mathcal{H} \mid \forall_{\underset{K}{x \in C}} \ \langle x \mid u \rangle \geq 0\}$ $\therefore u \in K^{\oplus} \Leftrightarrow \forall_{\underset{K}{\tilde{x} \in C}} \ \langle \tilde{x} \mid u \rangle \geq 0$

$\Leftrightarrow \boxed{Bx \in K^{\oplus}}$

So, (eq.:25.51) becomes:

find $x \in K$
such that $x \perp Bx$, $Bx \in K^{\oplus}$

∎

Remark 25.16. Restructuring of the variational inequality problem:
[$f \in \Gamma_0(\mathcal{H})$; $B: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]
Associated variational inequality problem: // Restructured version of (eq.: 25.49)
find $x \in \mathcal{H}$
such that $x \in \operatorname{zer}(\partial f + B)$

2016-10-17 2:36 PM Proof:

$x \in \operatorname{zer}(\partial f + B)$

$\Leftrightarrow (\partial f + B) x \ni 0 \Leftrightarrow \partial f(x) + Bx \ni 0 \Leftrightarrow \exists_{u \in \mathcal{H}} \ u \in Bx, \ -u \in \partial f(x)$

$\Leftrightarrow \exists_{u \in Bx} \ -u \in \partial f(x)$

// By definition $\tilde{u} \in \partial f(x) \overset{\text{def}}{\Leftrightarrow} \forall_{y \in \mathcal{H}} \ \langle y - x \mid \tilde{u} \rangle + f(x) \leq f(y)$

$\therefore -u \in \partial f(x) \Leftrightarrow \forall_{y \in \mathcal{H}} \ \underbrace{\langle y - x \mid -u \rangle}_{\substack{\text{// } (-1)\langle y-x \mid u \rangle \\ = \langle x-y \mid u \rangle \text{ //}}} + f(x) \leq f(y)$

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

$\Leftrightarrow \exists_{u \in Bx} \ \langle x-y \mid u \rangle + f(x) \leq f(y)$ . . . (eq.: 1)

$\therefore$ Variational inequality problem:=
// from (eq.:25.49)

$\left(\begin{array}{l} \text{find } x \\ \text{such that } \exists_{u \in Bx} \ \forall_{y \in \mathcal{H}} \ \langle x-y \mid u \rangle + f(x) \leq f(y) \end{array} \right)$ // now using (eq.: 1)

$= \left(\begin{array}{l} \text{find } x \\ \text{such that } x \in \operatorname{zer}(\partial f + B) \end{array} \right)$ ∎

Proposition 25.17. (Douglas-Rachford algorithm for variational inequality)

[$f \in \Gamma_0(\mathcal{H})$; $B: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone; the variational inequality problem: find $x \in \mathcal{H}$ such that $\exists_{u \in Bx} \forall_{y \in \mathcal{H}} \langle x-y \mid u \rangle + f(x) \leq f(y)$ has atleast one solution; $(\lambda_n)_{n\in\mathbb{N}}: \subseteq [0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$; $\gamma \in \mathbb{R}_{++}$; $x_0 \in \mathcal{H}$; set

$$\forall_{n\in\mathbb{N}} \quad \left| \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = \mathrm{Prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{array} \right. \qquad (\text{eq:}25.59)$$

] $\Rightarrow \exists_{x \in \mathcal{H}}$

(i) $J_{\gamma B} x$: solution to the variational inequality problem

(ii) $(x_n)_{n\in\mathbb{N}}$: $x_n \rightharpoonup x$

(iii) $(y_n)_{n\in\mathbb{N}}$: $y_n \to J_{\gamma B} x$

$(z_n)_{n\in\mathbb{N}}$: $z_n \to J_{\gamma B} x$

(iv) (One of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of dom ∂f

(b) B: uniformly monotone on every nonempty bounded subset of dom B) $\Rightarrow$

$(y_n)_{n\in\mathbb{N}}$: $y_n \to J_{\gamma B} x$: unique solution of the variational inequality

$(z_n)_{n\in\mathbb{N}}$: $z_n \to J_{\gamma B} x$: unique solution of the variational inequality

2016-10-18 7:57 AM

Proof: Proof is straightforward and follows from the following results:

Definition 25.12. (Variational inequality problem)

[$f \in \Gamma_0(\mathcal{H})$; $B: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

Associated variational inequality problem:

find $x \in \mathcal{H}$

such that $\exists_{u \in Bx} \forall_{y \in \mathcal{H}} \langle x-y \mid u \rangle + f(x) \leq f(y)$ (eq:25.49)

Remark 25.16. (Restructuring of Variational Inequality Problem)

[$f \in \Gamma_0(\mathcal{H})$, $B: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone]

Associated variational inequality problem: // restructured version of (eq:25.49)

find $x \in \mathcal{H}$

such that $x \in \mathrm{zer}(\partial f + B)$ // $f \in \Gamma_0(\mathcal{H}) \Rightarrow \partial f$: maximally monotone

Example 23.3.

[$f \in \Gamma_0(\mathcal{H})$, $\gamma \in \mathbb{R}_{++}$] $\Rightarrow$ $\left| \begin{array}{l} \mathrm{Prox}_{\gamma f} = J_{\gamma \partial f} \\ \nabla({}^{\gamma}f) = {}^{\gamma}(\partial f) \end{array} \right.$ // here ${}^{\gamma}f$: Moreau envelope of f :

Example 22.4. set $C := \mathrm{dom}\, \partial f$

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

C : nonempty, $\subseteq \mathrm{dom}\, \partial f$]

∂f : uniformly monotone on C

Theorem 25.6. (Douglas-Rachford algorithm) set $A := \partial f$: maximally monotone

[[A, B : maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$; $\mathrm{zer}(A+B) \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}}$: sequence in $[0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$; $\gamma \in \mathbb{R}_{++}$; $x_0 \in \mathcal{H}$

$$\forall_{n\in\mathbb{N}} \quad \left| \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) = \mathrm{Prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{array} \right. \text{]] } \Rightarrow \exists_x \; x \in \mathrm{Fix}\, R_{\gamma A} R_{\gamma B}$$

(i) $J_{\gamma B} x \in \mathrm{zer}(A+B)$

(ii) $(y_n - z_n)_{n\in\mathbb{N}}$: $(y_n - z_n) \to 0$

(iii) $(z_n)_{n\in\mathbb{N}}$: $z_n \rightharpoonup x$

(iv) $(y_n)_{n\in\mathbb{N}}$: $y_n \to J_{\gamma B} x$

(v) $(z_n)_{n\in\mathbb{N}}$: $z_n \to J_{\gamma B} x$

(vi) C: closed affine subspace of $\mathcal{H}$; $A = N_C$) $\Rightarrow$ $(P_C x_n)_{n\in\mathbb{N}}$: $P_C x_n \to J_{\gamma B} x$

(vii) One of the following holds:

(a) A: uniformly monotone on every nonempty bounded subset of dom A

(vii) One of the following holds:

(a) A: uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$

(b) B: uniformly monotone on every nonempty bounded subset of $\operatorname{dom} B$) $\Rightarrow$

$(y_n)_{n\in\mathbb{N}}, (z_n)_{n\in\mathbb{N}}$: converge strongly to the unique point in $\operatorname{zer}(A+B)$

Proposition 25.18. (Forward-backward algorithm for variational inequality)

[$f\in\Gamma_0(\mathcal{H})$; $\beta\in\mathbb{R}_{++}$; $B:\mathcal{H}\to\mathcal{H}$, β-cocoercive ; $\gamma\in\,]0,2\beta[$; $\delta=\min\{1,\beta/\gamma\}+\frac{1}{2}$; the variational inequality problem: find $x\in\mathcal{H}$ such that $\forall_{y\in\mathcal{H}}\ \langle x-y|Bx\rangle+f(x)\leq f(y)$ admits atleast one solution ;

$(\lambda_n)_{n\in\mathbb{N}}:\subseteq[0,\delta]$, $\sum_{n\in\mathbb{N}}\lambda_n(\delta-\lambda_n)=+\infty$; $x_0\in\mathcal{H}$; set

$$\forall_{n\in\mathbb{N}}\ \left|\begin{array}{ll} y_n=x_n-\gamma Bx_n & \text{(eq.: 25.56)}\\ x_{n+1}=x_n+\lambda_n(\operatorname{Prox}_{\gamma f}y_n-x_n) & \end{array}\right.$$

] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a solution to the variational inequality

(ii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$; x: solution to the variational inequality) $\Rightarrow$ $(Bx_n)_{n\in\mathbb{N}}$: $Bx_n\to Bx$

(iii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$; One of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of $\operatorname{dom}\partial f$

(b) B: uniformly convex on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: x_n converges strongly to the unique solution to the variational inequality.

2016-10-18 8:47 AM

Proof: Proof is strightforward again using the following results:

Example 23.5.

[$f\in\Gamma_0(\mathcal{H})$, $\gamma\in\mathbb{R}_{++}$] $\Rightarrow$ $\left|\begin{array}{l}\operatorname{Prox}_{\gamma f}=J_{\gamma\partial f}\\ \nabla({}^{\gamma}f)={}^{\gamma}(\partial f)\end{array}\right.$ // here ${}^{\gamma}f$: Moreau envelope of f ;

Example 22.4. set $C:=\operatorname{dom}\partial f$

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

C: nonempty, $\subseteq\operatorname{dom}\partial f$]

∂f: uniformly monotone on C

Remark 25.16. (Restructuring of Variational Inequality Problem)

[$f\in\Gamma_0(\mathcal{H})$, $B:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone]

Associated variational inequality problem: // restructured version of (eq.25.49)

find $x\in\mathcal{H}$

such that $x\in\operatorname{zer}(\partial f+B)$ // $f\in\Gamma_0(\mathcal{H})\Rightarrow\partial f$: maximally monotone

Theorem 25.8. (Forward-backward algorithm) set $A:=\partial f$

// BT: firmly nonexpansive

[$A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone ; $\beta\in\mathbb{R}_{++}$; $B:\mathcal{H}\to\mathcal{H}$, β-cocoercive ; $\gamma\in\,]0,2\beta[$; $\delta=\min\{1,\frac{\beta}{\gamma}\}+\frac{1}{2}$;

$(\lambda_n)_{n\in\mathbb{N}}$: sequence in $[0,\delta]$, $\sum_{n\in\mathbb{N}}\lambda_n(\delta-\lambda_n)=+\infty$; $x_0\in\mathcal{H}$; $\operatorname{zer}(A+B)\neq\emptyset$;

$$\forall_{n\in\mathbb{N}}\ \left|\begin{array}{l} y_n=x_n-\gamma Bx_n\\ x_{n+1}=x_n+\lambda_n(\underbrace{J_{\gamma A}}_{\operatorname{prox}_{\gamma f}}y_n-x_n)\end{array}\right.$$] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer}(A+B)$

(ii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$, $x\in\operatorname{zer}(A+B)$) $\Rightarrow$ $(Bx_n)_{n\in\mathbb{N}}$: $Bx_n\to Bx$

(iii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$; one of the following holds:

(a) A: uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$

(b) B: uniformly monotone on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges strongly to the unique point in $\operatorname{zer}(A+B)$

Example 25.19. [C: nonempty closed convex subset of $\mathcal{H}$; $\beta\in\mathbb{R}_{++}$; $B:\mathcal{H}\to\mathcal{H}$, β-cocoercive ; the variational inequality problem: find x such that $\forall_{y\in C}\ \langle x-y|Bx\rangle\leq 0$ admits atleast one solution x to the variational inequality ; $x_0\in\mathcal{H}$; set:

$$\forall_{n\in\mathbb{N}}\ x_{n+1}=P_C(x_n-\beta Bx_n)\ ]\Rightarrow\left|\begin{array}{l}(x_n)_{n\in\mathbb{N}}:x_n\to x\\ (Bx_n)_{n\in\mathbb{N}}:Bx_n\to Bx\end{array}\right.$$

2016-10-19 8:37 AM

Proof: Proof follows from the following results:

$\forall_{n\in\mathbb{N}}\ x_{n+1}=P_C(x_n-\beta Bx_n)\ ] \Rightarrow \left| \begin{array}{l} (x_n)_{n\in\mathbb{N}} : x_n\to x \\ (Bx_n)_{n\in\mathbb{N}} : Bx_n\to Bx \end{array}\right.$

2016-10-19 8:37 AM

Proof: Proof follows from the following results:

here set $f=\iota_C\in\Gamma_0(\mathcal{H})$; $\gamma:=\beta$; $\lambda_n=1$; // recall example 12.25. C: nonempty closed convex, $\subseteq\mathcal{H}\Rightarrow(\underbrace{\mathrm{Prox}_{\gamma\iota_C}}_{\in\mathbb{R}_{++}}=P_C)$

Proposition 25.18. (Forward-backward algorithm for variational inequality)
[$f\in\Gamma_0(\mathcal{H})$; $\beta\in\mathbb{R}_{++}$; $B:\mathcal{H}\to\mathcal{H}$, β-cocoercive; $\gamma\in\,]0,2\beta[$; $\delta=\min\{1,\underbrace{\beta/\gamma}_{=1/2}\}+\frac{1}{2}$; the variational inequality problem: find $x\in\mathcal{H}$ such that $\forall_{y\in\mathcal{H}}\ \langle x-y|Bx\rangle+f(x)\le f(y)$ admits atleast one solution; $(\lambda_n)_{n\in\mathbb{N}}:\subseteq[0,\delta]$, $\sum_{n\in\mathbb{N}}\lambda_n(\delta-\lambda_n)=+\infty$; $x_0\in\mathcal{H}$; set

$\forall_{y\in C}\ \langle x-y|Bx\rangle\le 0$ // this is the modified // VI problem

$\forall_{n\in\mathbb{N}}\ \left|\begin{array}{l} y_n=x_n-\gamma Bx_n=x_n-\beta Bx_n \\ x_{n+1}=x_n+\underbrace{\lambda_n}_{=1}(\underbrace{\mathrm{Prox}_{\gamma f}}_{=P_C}y_n-x_n)=x_n+P_Cy_n-x_n=P_Cy_n=P_C(x_n-\beta Bx_n) \end{array}\right.$ (eq: 25.56) ✓ this is our iteration scheme

] $\Rightarrow$

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a solution to the variational inequality ✓ // this is (i)

(ii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$; x: solution to the variational inequality) $\Rightarrow$ $(Bx_n)_{n\in\mathbb{N}}$: $Bx_n\to Bx$ ✓ // this is (ii)

(iii) ($\inf_{n\in\mathbb{N}}\lambda_n>0$; One of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of dom ∂f

(b) B: uniformly convex on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: x_n converges strongly to the unique solution to the variational inequality

Example 25.20.

[C: nonempty closed convex subset of $\mathcal{H}$; $\beta\in\mathbb{R}_{++}$; $B:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone operator, single-valued and β-Lipschitz continuous relative to C; $\mathrm{cone}(C-\mathrm{dom}B)=\overline{\mathrm{span}}(C-\mathrm{dom}B)$; variational inequality problem: find $x\in C$ such that $\forall_{y\in C}\ \langle x-y|Bx\rangle\le 0$ admits atleast solution; $x_0\in C$; $\gamma\in\,]0,1/\beta[$; set:

$$\forall_{n\in\mathbb{N}}\ \left|\begin{array}{l} y_n=x_n-\gamma Bx_n \\ z_n=P_Cy_n \\ r_n=z_n-\gamma Bz_n \\ x_{n+1}=P_C(x_n+y_n-r_n)\ ] \Rightarrow \end{array}\right.$$

(i) $(x_n-z_n)\to 0$

(ii) $(x_n)_{n\in\mathbb{N}}, (z_n)_{n\in\mathbb{N}}$: converge weakly to a solution of the variational inequality

(iii) B: uniformly monotone on every nonempty bounded subset of $C\Rightarrow$

$(x_n)_{n\in\mathbb{N}}, (z_n)_{n\in\mathbb{N}}$: converge strongly to the unique solution to variational inequality.

2016-10-19 10:01 PM

Proof:

First recall the following results:

/* Example 20.41. [C: nonempty closed convex, $\subseteq\mathcal{H}$] $\left|\begin{array}{l} J_{N_C}=P_C \text{ (Example 23.4)} \\ N_C: \text{maximally monotone; By definition: } \mathrm{dom}\,N_C=C \end{array}\right.$

Theorem 24.5. [A, B: maximally monotone from $\mathcal{H}$ to $2^{\mathcal{H}}$, set $A:=N_C$

$\underbrace{\mathrm{cone}(\underbrace{\mathrm{dom}A}_{C}-\mathrm{dom}B)=\overline{\mathrm{span}}(\underbrace{\mathrm{dom}A}_{C}-\mathrm{dom}B)}_{\text{this is given}}$] $\Rightarrow$ $A+B$: maximally monotone */ $\Rightarrow$ $\therefore N_C+B$: maximally monotone

Now recall Tseng's algorithm:

Theorem 25.10. (Tseng's algorithm) set $A=N_C$, $D=C\Rightarrow \mathrm{dom}\,N_C=C$

[D: nonempty subset of $\mathcal{H}$; $A:\mathcal{H}\to 2^{\mathcal{H}}$, maximally monotone, $\mathrm{dom}A\subseteq D$; $B:\mathcal{H}\to 2^{\mathcal{H}}$, monotone, single-valued on D; $A+B$: maximally monotone; C: closed convex, $\subseteq D$, $C\cap\mathrm{zer}(A+B)\neq\emptyset$; B: $\frac{1}{\beta}$ Lipschitz continuous relative to $\underbrace{C\cup\underbrace{\mathrm{dom}A}_{C}}_{C}$, $\beta\in\mathbb{R}_{++}$; $\gamma\in\,]0,\beta[$;

$\underbrace{C\cap\mathrm{dom}A}_{\neq\emptyset}$ follows from the existence of solution of VI

$\forall_{n\in\mathbb{N}}\ \left|\begin{array}{l} y_n=x_n-\gamma Bx_n \\ z_n=J_{\gamma N_C}y_n=P_Cy_n \\ r_n=z_n-\gamma Bz_n \\ x_{n+1}=P_C(x_n-y_n+r_n)\ ] \Rightarrow \end{array}\right.$ (eq: 25.36)

(i) $(x_n-z_n)_{n\in\mathbb{N}}$: converges strongly to 0

(ii) $(x_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\underbrace{C\cap\mathrm{zer}(A+B)}_{\mathrm{zer}(N_C+B)}$

(iii) ($A\cup B$: uniformly monotone on every nonempty bounded subset of dom A) $\Rightarrow$

$(x_n)_{n\in\mathbb{N}}, (z_n)_{n\in\mathbb{N}}$: converge strongly to the unique point in $\underbrace{C\cap\mathrm{zer}(A+B)}_{\mathrm{zer}(N_C+B)}$

Part 1

October 21, 2016 7:22 AM

Notation 26.1: T: lower differentiable $\Rightarrow$ $\underset{\text{STANDARD}}{\nabla_{\mathcal{H}}} T(x) := \lim_{h \to 0} \frac{T(x+h) - T(x)}{h}$
Gateaux derivative

(Formula rule to the minimization of two convex functions satisfying constraint qualification)

Theorem 26.2: [$f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(K)$, $L \in \mathcal{B}(\mathcal{H}, K)$, $\bar{x} \in \mathcal{H}$, one of the following holds:

(a) $\operatorname{ri}(\operatorname{dom} g - L(\operatorname{dom} f))$

(b) K: finite-dimensional, g: polyhedral, and $\operatorname{dom} g \cap \operatorname{ri} L(\operatorname{dom} f) \neq \emptyset$

(c) $\mathcal{H}, K$: finite-dimensional, f, g: polyhedral, $\operatorname{dom} f \cap L^{-1}(\operatorname{dom} g) \neq \emptyset$] The following are equivalent:

(i) $\bar{x}$: solution to $\min_{x \in \mathcal{H}} f(x) + g(Lx)$

(ii) $\bar{x} \in \operatorname{zer}(\partial f + L^* \circ \partial g \circ L)$

(iii) $\exists_{v \in \partial g(L\bar{x})}$ $-L^* v \in \partial f(\bar{x})$

(iv) $\exists_{v \in \partial g(L\bar{x})} \forall_{x \in \mathcal{H}}$ $\langle \bar{x} - x, L^* v \rangle + f(\bar{x}) \leq f(x)$

Moreover, g: Gateaux differentiable at $L\bar{x}$. then (i)~(iv) is equivalent to each of the following:

(v) $-L^*(\nabla g(L\bar{x})) \in \partial f(\bar{x})$, (vi) $\forall_{x \in \mathcal{H}}$ $\langle \bar{x} - x, L^*(\nabla g(L\bar{x})) \rangle + f(\bar{x}) \leq f(x)$, (vii) $\forall_{x \in \mathcal{H}_{++}}$ $\bar{x} = \operatorname{prox}_{\frac{1}{\gamma} f}(\bar{x} - \gamma L^*(\nabla g(L\bar{x})))$

2016-10-21 9:22 AM

Proof:

(i) $\Leftrightarrow$ (ii):

$\bar{x}$: solution to $\min_{x \in \mathcal{H}} f(x) + g(Lx)$

$\Leftrightarrow \bar{x} \in \operatorname{Argmin}(f + g \circ L)$ // Proposition 26.1: $\operatorname{Argmin} \overset{\in \Gamma_0(\mathcal{H})}{f} = \operatorname{zer} \partial f = \operatorname{fix} \operatorname{prox}_f = \operatorname{zer} \operatorname{prox}_{\frac{1}{\gamma} f}$ #/

// $\operatorname{zer}(\underbrace{\partial f + L^* \circ \partial g \circ L}_{\partial f + L^* \circ \partial g \circ L})$ // using

Theorem 16.47: [K: real Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(K)$, $L \in \mathcal{B}(\mathcal{H}, K)$, one of the following holds:

(a) $\operatorname{ri}(\operatorname{dom} g - L(\operatorname{dom} f))$

(b) K: finite-dimensional, g: polyhedral, $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$

(c) $\mathcal{H}, K$: finite-dimensional, f, g: polyhedral, $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset$] $\Rightarrow \partial(f + g \circ L) = \partial f + L^* \circ (\partial g) \circ L$ #/

$\Leftrightarrow \bar{x} \in \operatorname{zer}(\partial f + L^* \circ \partial g \circ L)$

(ii) $\Leftrightarrow$ (iii):

$\bar{x} \in \operatorname{zer}(\partial f + L^* \circ \partial g \circ L)$

$\Leftrightarrow (\partial f + L^* \circ \partial g \circ L)(\bar{x}) = \partial f(\bar{x}) + L^*(\underbrace{\partial g(L\bar{x})}_{\exists_{v \in \partial g(L\bar{x})}}) \ni 0 \Leftrightarrow \exists_{v \in \partial g(L\bar{x})} \underbrace{\partial f(\bar{x})}_{L^*: \text{linear} \Rightarrow \text{single-valued} \Rightarrow L^* v: \text{single vector}} + L^* v \ni 0 \Leftrightarrow \exists_{v \in \partial g(L\bar{x})} \partial f(\bar{x}) \ni -L^* v$;

(iii) $\Leftrightarrow$ (iv)

$\exists_{v \in \partial g(L\bar{x})} -L^* v \in \partial f(\bar{x})$ // recall: $u \in \partial f(x) \overset{\text{def}}{\Leftrightarrow} \forall_{y \in \mathcal{H}} \langle y - x, u \rangle + f(x) \leq f(y)$ #/

$\Leftrightarrow \exists_{v \in \partial g(L\bar{x})} \forall_{y \in \mathcal{H}} \langle y - \bar{x}, L^* v \rangle + f(\bar{x}) \leq f(y)$;

Now assume, g: Gateaux differentiable at $L\bar{x}$; recall

Proposition 17.24: [$f: \mathcal{H} \to]-\infty, +\infty]$, proper, convex, $L \in \operatorname{dom} \partial f$]

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

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

(iii) $\Leftrightarrow$ (iv):

$\exists_{v \in \partial g(L\bar{x})} -L^* v \in \partial f(\bar{x}) \Leftrightarrow -L^*(\nabla g(L\bar{x})) \in \partial f(\bar{x})$

$\Leftrightarrow \partial g(L\bar{x}) = \{\nabla g(L\bar{x})\}$

(iv) $\Leftrightarrow$ (vi):

$\exists_{v \in \partial g(L\bar{x})} \forall_{y \in \mathcal{H}} \langle \bar{x} - y, L^* v \rangle + f(\bar{x}) \leq f(y) \Leftrightarrow \forall_{y \in \mathcal{H}} \langle \bar{x} - y, L^*(\nabla g(L\bar{x})) \rangle + f(\bar{x}) \leq f(y)$

$\Leftrightarrow \partial g(L\bar{x}) = \{\nabla g(L\bar{x})\}$

(v) $\Leftrightarrow$ (vii):

(v): $-L^*(\nabla g(L\bar{x})) \in \partial f(\bar{x})$ // take $\operatorname{prox}_{\gamma f}$

$\Leftrightarrow -\gamma L^*(\nabla g(L\bar{x})) \in \gamma \partial f(\bar{x}) = \partial(\gamma f)(\bar{x})$

$\Leftrightarrow \underbrace{\bar{x} - \gamma L^*(\nabla g(L\bar{x}))}_{\tilde{x}} - \underbrace{\bar{x}}_{\tilde{y}} \in \partial(\gamma f)(\bar{x})$ // now: Proposition A.34: $\forall$ $\begin{cases} f \in \Gamma_0(\mathcal{H}), \gamma \in \mathbb{R}_{++} \\ x \in \operatorname{Fix} \operatorname{prox}_{\gamma f} \Leftrightarrow x - \operatorname{prox}_{\gamma f}(x) \in \partial(\gamma f)(x) \end{cases}$

$\Leftrightarrow \bar{x} = \operatorname{prox}_{\gamma f}(\bar{x} - \gamma L^*(\nabla g(L\bar{x})))$;

Proposition 26.5: [C: closed convex subset of $\mathcal{H}$, $\bar{x} \in \operatorname{ri} C$, $\bar{x} \in \mathcal{H}$, $v \in \mathcal{H}_{++}$, one of the following holds:

(a) $\operatorname{Ri}(C - \operatorname{dom} f)$

(b) $\mathcal{H}$: finite dimensional, C: polyhedral, $C \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset$

(c) $\mathcal{H}$: finite dimensional, C: polyhedral set, f: polyhedral function, $C \cap \operatorname{dom} f \neq \emptyset$]

The following are equivalent:

(i) $\bar{x}$: a solution to the problem $\min_{x \in C} f(x)$; (ii) $\bar{x} \in \operatorname{zer}(N_C + \partial f)$

(iii) $\bar{x} \in \operatorname{Prox}_{\gamma f}(\bar{x} - \gamma(\nabla \operatorname{prox}_{\gamma f}(\bar{x}))$;

(iv) $\exists_{u \in \mathcal{H}} \overset{\in \mathcal{H}}{u} \in \partial f(\bar{x})$, for $\exists_{\bar{x} \in \mathcal{H}} \bar{x} \in C$, for $\bar{x} \in C \wedge \exists_{u \in \partial f(\bar{x})} \forall_{x \in C} \langle \bar{x} - x, u \rangle \leq 0$

Moreover, f: Gateaux differentiable at $\bar{x}$, each of the items (i)~(iv) is also equivalent to each of the following:

(v) $-\nabla f(\bar{x}) \in N_C(\bar{x})$

(vi) $\bar{x} \in C$ and $\forall_{x \in C} \langle \bar{x} - x, \nabla f(\bar{x}) \rangle \leq 0$

(vii) $\bar{x} = P_C(\bar{x} - \gamma \nabla f(\bar{x}))$ // Projected gradient method

2016-10-23 8:50 AM

Proof. Proof follows from the following results: 1) C: nonempty closed convex subset of $\mathcal{H} \Rightarrow$ $\operatorname{prox}_{\iota_C} = P_C$, $\partial \iota_C = N_C$

(Formula of minimization for a sum of two Γ_0 functions) // $\bar{x}$: $\bar{x} = \bar{x}$

Condition 16.5: [$\bar{x}, \bar{y} \in \mathcal{H}$, $\bar{x} \in \operatorname{ri} C$, $\bar{y} \in \operatorname{ri} D$, one of the following holds: $\forall \operatorname{dom} \bar{x} - \operatorname{dom} \iota_C \ni 0$] // $\because \iota_C = \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{else} \end{cases}$, $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$

(a) $0 \in \operatorname{ri}(\operatorname{dom} \bar{x} - \operatorname{dom} \bar{y})$ ✓

(b) $\mathcal{H}$: finite-dimensional, $\bar{x}$: polyhedral, $\operatorname{dom} \bar{y} \cap \operatorname{ri}(\operatorname{dom} \bar{x}) \neq \emptyset$ ✓ ($\bar{y}$: polyhedral $\Rightarrow$ f: polyhedral)

(c) $\mathcal{H}$: finite-dimensional, $\bar{x}, \bar{y}$ polyhedral, $\operatorname{dom} \bar{y} \cap \operatorname{dom} \bar{x} \neq \emptyset$ ✓ ($\bar{x}$: polyhedral $\Rightarrow$ C: polyhedral)

(1) $\mathcal{H}$: finite-dimensional, $\bar{x}, \bar{y}$ polyhedral, $\operatorname{dom} \bar{y} \cap \operatorname{dom} \bar{x} \neq \emptyset$ ✓ ($\bar{x}$: polyhedral $\Rightarrow$ C: polyhedral) ✓ The following are equivalent:

(i) $\bar{x}$: solution to the problem $\min_{x \in \mathcal{H}} \{\bar{x}(x) + \bar{y}(x)\}$; (ii) $\bar{x} \in \operatorname{zer}(\partial \bar{x} + \partial \bar{y})$ ✓ ($\partial \bar{x} = N_C$, $\partial \bar{y} = \partial f$)

(iii) $\bar{x} \in \operatorname{prox}_{\gamma \bar{x}}(\bar{x} - \gamma \nabla \bar{y}(\bar{x}))$ ✓ ($\operatorname{prox}_{\gamma \bar{x}} = \operatorname{prox}_{\gamma \iota_C} = P_C$, $\nabla \bar{y} = \nabla f$)

(iv) $\exists_{u \in \partial \bar{x}(\bar{x})} -u \in \partial \bar{y}(\bar{x})$ ✓ ($\partial \bar{x}(\bar{x}) = N_C(\bar{x})$, $\partial \bar{y}(\bar{x}) = \partial f(\bar{x})$)

(v) $\exists_{u \in \partial \bar{x}(\bar{x})} \forall_{x \in \mathcal{H}}$ $\langle \bar{x} - x, u \rangle + \bar{y}(\bar{x}) \leq \bar{y}(x) \Leftrightarrow \forall_{x \in \mathcal{H}} \langle \bar{x} - x, \nabla \bar{y}(\bar{x}) \rangle + \bar{y}(\bar{x}) \leq \bar{y}(x)$ ✓ ($\bar{x} = \iota_C$, $\nabla \bar{y} = \nabla f$)

Moreover, $\bar{y}$: Gateaux differentiable at $\bar{x}$, each of the items (i)~(v) is also equivalent to each of the following:

(vi) $-\nabla \bar{y}(\bar{x}) \in N_C(\bar{x})$ ✓ ($\nabla \bar{y} = \nabla f$)

(vii) $\bar{x} \in C \wedge \forall_{x \in C} \langle \bar{x} - x, \nabla \bar{y}(\bar{x}) \rangle \leq 0$ ✓ ($\bar{x} = \iota_C$, $\nabla \bar{y} = \nabla f$)

(viii) $\bar{x} = \operatorname{Prox}_{\gamma \bar{x}}(\bar{x} - \gamma \nabla \bar{y}(\bar{x}))$ ✓ ($\bar{x} = \iota_C$, $\nabla \bar{y} = \nabla f$)

Example 26.8 (Stampacchia's theorem which is applied a lot in the study of partial differential equation)

F: monotone, bilinear form, $\exists_{M, \beta \in \mathbb{R}_{++}} \forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}}$ $\begin{cases} |F(x,y)| \leq M \|x\| \|y\| \\ F(x,x) \geq \beta \|x\|^2 \end{cases}$ (eq. 26.5)

C: nonempty closed convex subset of $\mathcal{H}$;

$\beta \in \operatorname{Ri}(\mathcal{H}, K)$] $\Rightarrow$

(i) $\exists_{\bar{x} \in \mathcal{H}}$ $\bar{x} \in C \wedge \forall_{y \in C}$ $F(\bar{x}, y - \bar{x}) \geq \beta(y - \bar{x})$ (eq. 26.6)

(ii) [F: symmetric ; $\bar{x} \in \mathcal{H}$; $f: \mathcal{H} \to K : x \mapsto \frac{1}{2} F(x,x) - \beta(x)$] The following are equivalent: (a) $\bar{x}$ solves (eq. 26.6)

(b) $\operatorname{Argmin}_C f = \{\bar{x}\}$

2016-10-25 7:55 PM

Proof:

F: bilinear $\overset{\text{def}}{\Leftrightarrow}$ $\begin{cases} \forall_{x \in \mathcal{H}} \{F(x, \cdot) \in \mathcal{B}(\mathcal{H}, K)\} \\ \forall_{y \in \mathcal{H}} \{F(\cdot, y) \in \mathcal{B}(\mathcal{H}, K)\} \end{cases}$

Fact 1.11. (Riesz-Frechet representation) [$\mathcal{H}$: $\mathcal{B}(\mathcal{H}, K)$] $\exists_{!\, z \in \mathcal{H}} \forall_{x \in \mathcal{H}}$ $\begin{cases} T(x) = \langle z, x \rangle \\ \|z\| = \|T\| \end{cases}$ set $\bar{z} := z$

$\Rightarrow \exists_{!\, z \in \mathcal{H}} \forall_{x \in \mathcal{H}} \{T(x) = \langle z, x \rangle\}$

// let us define an onto linear operator $L: \mathcal{H} \to \mathcal{H}$ $\exists_{!\, L \in \mathcal{B}(\mathcal{H}, \mathcal{H})}$ such that $Lx = v$

$\Rightarrow \forall_{x \in \mathcal{H}} F(x, \cdot) = \langle \cdot, Lx \rangle$

// now, Example 20.15. [$A: \mathcal{H} \to \mathcal{H}$, linear] A: monotone $\Leftrightarrow \forall_{x \in \mathcal{H}} \langle x, Ax \rangle \geq 0$ #/

given: $\forall_{x \in \mathcal{H}} \forall_{y \in \mathcal{H}}$ $\begin{cases} |F(x,y)| \leq \beta \|x\| \|y\| \\ |F(x,y)| = |\langle y, Lx \rangle| \leq \beta \|x\| \|y\| \end{cases}$ $\wedge$ $F(x,x) \geq \alpha \|x\|^2$ $\Rightarrow F(x,x) = \langle x, Lx \rangle \geq \alpha \|x\|^2 \geq 0 \Rightarrow$ L: monotone // from Example 20.29. [$A: \mathcal{H} \to \mathcal{H}$, monotone] A: maximally monotone #/

$\Downarrow$ L: maximally monotone

// now, from Example 22.7. [$A: \mathcal{H} \to \mathcal{H}$, linear] A: strongly monotone with constant $\beta \in \mathbb{R}_{++} \Leftrightarrow \forall_{x \in \mathcal{H}} \langle x, Ax \rangle \geq \beta \|x\|^2$ #/

space of linear continuous operators from $\mathcal{H}$ to $\mathcal{H}$

L: strongly monotone with constant α

// now recall that:

Proposition 17.11: [Maximally monotonicity preserving operation]

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $\operatorname{dom} A = \mathcal{H}$; $\begin{cases} \cdot \alpha A^{-1}: \text{maximally monotone} \\ \cdot \operatorname{dom}(A + \alpha \mathrm{Id}): \text{maximally monotone} \end{cases}$] #/

$\underbrace{B}_{\text{B: maximally monotone}}$; $\operatorname{dom} B = \mathcal{H}$; $\underbrace{S}_{\text{S: strongly valued operator}}: \mathcal{H} \to \mathcal{H}$, $\operatorname{dom} S = \mathcal{H}$ // $\boxed{B: \text{strongly monotone with constant } \alpha}$ because $\forall_{(x,u) \in \operatorname{gra} B} \forall_{(y,v) \in \operatorname{gra} B}$ $\langle x - y, u - v \rangle \geq \alpha \|x - y\|^2$

$\langle x - y, Sx - Sy \rangle = \langle x - y, Lx - Ly \rangle \geq \alpha \|x - y\|^2$

B: strongly monotone with constant α

the given assumptions we want to solve from Ex.26.13: $\begin{cases} \bar{x} \in C \\ \forall_{y \in C} \{F(\bar{x}, y - \bar{x}) \geq \beta(y - \bar{x})\} \end{cases}$ now we will show that $\bar{x} \in \operatorname{zer}(N_C + B)$ will solve this.

$(N_C + B)\bar{x} \ni 0$

$\Leftrightarrow N_C \bar{x} + B\bar{x} \ni 0$ ($B\bar{x}$: single-vector as B: affine)

$\Rightarrow N_C \bar{x} + (L\bar{x} - u) \ni 0$

$\Leftrightarrow N_C \bar{x} \ni u - L\bar{x}$ // $N_C \bar{x} \neq \emptyset$ as $u - L\bar{x}$ is unit. now recall

$\Leftrightarrow \{v \in \mathcal{H} \mid \forall_{y \in C} \langle y - \bar{x}, v \rangle \leq 0\} \ni u - L\bar{x}$; $\bar{x} \in C$

$\Leftrightarrow \exists_{v \in \mathcal{H}}$ $v = u - L\bar{x}$ $\wedge$ $\forall_{y \in C}$ $\langle y - \bar{x}, u - Lx \rangle \leq 0$ $\wedge$ $\bar{x} \in C$

$\Leftrightarrow N_C\bar{x} + (L\bar{x}-u) \ni 0$

$\Leftrightarrow N_C\bar{x} \ni u - L\bar{x}$ // $N_C\bar{x} \neq \emptyset$ as $u-L\bar{x}$ is int. now recall

$\Leftrightarrow \{v \in N \mid \forall_{y\in C}\ \langle y-\bar{x}, v\rangle \leq 0\} \ni u - L\bar{x}$, $\bar{x} \in C$

$\Leftrightarrow \exists_{v\in N}\ v = u - L\bar{x} \ \wedge\ \forall_{y\in C}\ \langle y-\bar{x}, u-L\bar{x}\rangle \leq 0 \ \wedge\ \bar{x}\in C$

$\Leftrightarrow \forall_{y\in C}\ \underbrace{\langle y-\bar{x}, u\rangle}_{f(y-\bar{x})} - \underbrace{\langle y-\bar{x}, L\bar{x}\rangle}_{F(\bar{x},y-\bar{x})} \leq 0 \ \wedge\ \bar{x}\in C$

$\Leftrightarrow \forall_{y\in C}\ F(\bar{x}, y-\bar{x}) \geq f(\bar{x}, y-\bar{x}) \ \wedge\ \bar{x}\in C$

$\therefore\ \bar{x} \in \mathrm{zer}(N_C f) \Leftrightarrow \left(\forall_{y\in C}\ F(\bar{x}, y-\bar{x}) \geq f(\bar{x}, y-\bar{x}) \ \wedge\ \bar{x}\in C\right)$

Now: recall:

Lemma 16.5 [A: maximally monotone, $\mathrm{int}\,\mathrm{dom}\,A \neq \emptyset$, then A is strongly mono.] [Bo1] $A := N_C$

$\therefore$ $N_C f$: maximally monotone

$N_C f$: strongly monotone /* : f : strongly monotone, and adding a monotone and strongly monotone operator creates a maximally monotone operator in linear intersected domain */

// now $\tilde{A} := N_C f$

$\therefore$ $\mathrm{zer}(N_C f)$: singleton

$\exists_{!\bar{x}\in H}\left(\forall_{y\in C}\ F(\bar{x}, y-\bar{x}) \geq f(\bar{x}, y-\bar{x}) \ \wedge\ \bar{x}\in C\right)$ ↯

(b) given F: symmetric,

$f : H \to \mathbb{R} : x \mapsto \frac{1}{2}F(x,x) - \langle u, x\rangle$

now $\begin{cases} f \in \Gamma_0(H) \text{ // need to figure out how} \\ \mathrm{dom}\, f = H \end{cases}$

Example 2.47: $Df(x) = F(x,\cdot) - u$

$\Downarrow$

result follows from (a) and proposition 16.5.

■

(KKT sufficient conditions)

Proposition 16.18: [$f \in \Gamma_0(H)$; $m \in \mathbb{N}_{++}$; $I = \{1,\dots,m\}$; $(g_i)_{i\in I} : \Gamma_0(H)$, Slater condition: $\begin{cases} \forall_{i\in I}\ \mathrm{lev}_{<0}\, g_i \subseteq \mathrm{int}\,\mathrm{dom}\, g_i \\ \mathrm{dom}\, f \cap \left(\bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i\right) \neq \emptyset \end{cases}$] [eq: 26.44]

is satisfied ; $\bar{x} \in H$; Problem in consideration: $\min_{x\in H} f(x)$ s.t. $g_i(x) \leq 0$ $\forall i \in I$ (26.46)] $\Rightarrow$

- $\bar{x}$: solution to (26.46) $\Leftrightarrow$

$\left(\exists_{(\bar{v}_i)_{i\in I}\in\mathbb{R}^m_+}\ \exists_{(u_i)_{i\in I}\in\mathbb{R}^m}\ \bigcap_{i\in I} \partial g_i(\bar{x}) \ni -\sum_{i\in I} \bar{v}_i u_i \in \partial f(\bar{x}) \ \wedge\ \forall_{i\in I}\ g_i(\bar{x}) \leq 0 \ \wedge\ \forall_{i\in I}\ \bar{v}_i g_i(\bar{x}) = 0\right)$

in which case $(\bar{v}_i)_{i\in I}$: Lagrange multipliers associated with $\bar{x}$, and $\bar{x}$: solves $\min_{x\in H} f(x) + \sum_{i\in I} \bar{v}_i g_i(x)$ becomes

- Moreover

$\left(f,\ (g_i)_{i\in I} : \text{Gateaux differentiable at } \bar{x}\right) \Rightarrow \exists_{(\bar{v}_i)_{i\in I}\in\mathbb{R}^m_+}$ $\begin{cases} \nabla f(\bar{x}) = -\sum_{i\in I} \bar{v}_i \nabla g_i(\bar{x}) \\ \forall_{i\in I}\ g_i(\bar{x}) \leq 0 \\ \forall_{i\in I}\ \bar{v}_i g_i(\bar{x}) = 0 \end{cases}$

2016-11-01 1:31 PM

Proof:

Define : $\forall_{i\in I}\ C_i = \mathrm{lev}_{\leq 0}\, g_i$ and $C = \bigcap_{i\in I} C_i$ // (1) this is the constraint set $= \{x \in H \mid g_i(x) \leq 0\}$

Now: recall

Proposition 8.38: If $\tilde{f} := g_i$

[$\tilde{f} : H \to]-\infty, +\infty]$ proper, convex, ✓ // as $g_i \in \Gamma_0(H)$

TFAE of the following holds:

- $\tilde{f}$: bounded under some neighborhood.
- $\tilde{f}$: lower semicontinuous ✓ // as $g_i \in \Gamma_0(H)$
- H : finite-dimensional] $\Rightarrow \mathrm{cont}\,\tilde{f} = \mathrm{int}\,\mathrm{dom}\,\tilde{f}$ // $\mathrm{cont}\,\tilde{f}$: domain of continuity of a function $\tilde{f}$ */

$\therefore$ $\mathrm{cont}\, g_i = \mathrm{int}\,\mathrm{dom}\, g_i = \mathrm{int}\left(\{x \in H \mid g_i(x) < +\infty\}\right)$

$\therefore$ g_i : continuous on $\mathrm{int}\,\mathrm{dom}\, g_i$ and all subsets of it

$\Rightarrow g_i$: upper-semicontinuous on all subsets of $\mathrm{int}\,\mathrm{dom}\, g_i$

now checking $\mathrm{lev}_{<0}\, g_i = \{x \in H \mid g_i(x) < 0\} \subseteq \mathrm{int}\,\mathrm{dom}\, g_i$

$\therefore$ g_i : upper-semicontinuous on $\mathrm{lev}_{<0}\, g_i$

given Slater's condition holds $\forall_{i\in I}\ \mathrm{lev}_{<0}\, g_i \subseteq \mathrm{int}\,\mathrm{dom}\, g_i$

$\mathrm{dom}\, f \cap \left(\bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i\right) \neq \emptyset \Rightarrow \bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i \neq \emptyset \Rightarrow \forall_{i\in I}\ \mathrm{lev}_{<0}\, g_i \neq \emptyset$

Corollary 8.38: If set $\tilde{f} := g_i$

[$\tilde{f} : H \to]-\infty, +\infty]$ proper, convex, $\mathrm{lev}_{<0}\, \tilde{f} \neq \emptyset$ ✓

TFAE of the following holds:

(i) $\tilde{f}$: upper semicontinuous on $\mathrm{lev}_{<0}\, \tilde{f}$ ✓

(ii) $\tilde{f}$: lower semicontinuous, $\mathrm{dom}\, \tilde{f}$: open

(iii) H : finite-dimensional, $\mathrm{dom}\, \tilde{f}$: open] $\Rightarrow$ $\boxed{\mathrm{int}\,\mathrm{lev}_{<0}\, \tilde{f} = \mathrm{lev}_{<0}\, \tilde{f}}$ $\Rightarrow \mathrm{int}\,\mathrm{lev}_{<0}\, g_i = \mathrm{lev}_{<0}\, g_i$ // i.e. $\mathrm{int}\{x \mid g_i(x) < 0\} = \{x \mid g_i(x) < 0\}$

now, $\mathrm{int}\, C = \bigcap_{i\in I} \mathrm{int}\, C_i = \bigcap_{i\in I} \mathrm{int}\,\mathrm{lev}_{<0}\, g_i = \bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i$

// $\mathrm{int} \cap_i C_i = \cap_i \mathrm{int}\, C_i$

$\mathrm{int}\, C \cap \mathrm{dom}\, f \neq \emptyset \Rightarrow C \cap \mathrm{dom}\, f \neq \emptyset$

now C : convex

$\mathrm{dom}\, f$: convex // $\because f \in \Gamma_0(H)$

$\Rightarrow 0 \in \mathrm{int}(C - \mathrm{dom}\, f)$ // using Result 3.1.m. [A, B: convex, $A \cap \mathrm{int}\, B \neq \emptyset \Rightarrow \mathrm{int}(A-B)$]

$\subseteq \mathrm{sri}(C - \mathrm{dom}\, f)$ // relative interior for general sets of dimensions

$\mathrm{sri}\, C \subseteq \mathrm{int}\, C \subseteq \mathrm{ri}\, C \subseteq \mathrm{sri}\, C$ */

Now we use Proposition 26.5.

Proposition 26.5 [C: closed convex subset of H; $f \in \Gamma_0(H)$; $\bar{x}, u \in H$, one of the following holds:

(i) $\mathrm{zer}(\partial f - u) \neq \emptyset$

(ii) H: finite dimensional, C: polyhedral, $C \cap \mathrm{ri}(\mathrm{dom}\, f) \neq \emptyset$

(iii) H: finite dimensional, C: polyhedral set, f: polyhedral function, $C \cap \mathrm{dom}\, f \neq \emptyset$]

The following are equivalent:

(i) $\bar{x}$ is solution to the problem $\min_{x\in H} f(x) - \langle x, u\rangle$: (26.50) ; $\bar{x} \in \mathrm{zer}(A + N_C)$

(ii) $\bar{x} \in \mathrm{Fix}\left(\mathrm{Prox}_{\gamma f} \circ \mathrm{Proj}_C \circ (\mathrm{Id} - \gamma u)\right)$

for $\gamma \in]0, +\infty[$; or $\exists_{\text{weak}\,\bar{x}\in H}\ \bar{x} \in \mathrm{zer}\, \tilde{A}$; or $\exists_{\bar{x}\in H}\ \bar{x} = \mathrm{J}_{\text{weak}\,\tilde{A}}\, \mathrm{V}_{\tilde{A}}\, (\mathrm{C}\circ\mathrm{J}_{\tilde{A}})\circ\mathrm{C}$

Moreover, f : Gateaux differentiable at $\bar{x}$, each of the items (i)-(ii) is also equivalent to each of the following

(iv) $-u \in \partial f(\bar{x})$

(v) $\bar{x}$ is a $\mathrm{J}_{\text{weak}\,\tilde{A}}\, \mathrm{V}_{\tilde{A}}\, (\mathrm{C}\circ\mathrm{J}_{\tilde{A}})\circ\mathrm{C}$

(vi) $\bar{x} = \mathrm{Prox}_{\gamma f}(\bar{x} - \gamma\nabla f(\bar{x}) - \gamma u)$ // Proximal gradient method

$\boxed{\bar{x} \in \mathrm{Argmin}_C f \Leftrightarrow \exists_{u\in N_C\bar{x}}\ -u \in \partial f(\bar{x})}$ (eq: 26.50) [eq:26.50]

Suppose, $\bar{x} \in C$ // as $\bar{x} \in \mathrm{Argmin}_C f$

We want to calculate $N_C\bar{x}$

define,

$\begin{cases} I_0 = \{i \in I \mid g_i(\bar{x}) = 0\} \\ I_1 = \{i \in I \mid g_i(\bar{x}) < 0\} \end{cases} \Rightarrow \boxed{I = I_0 \cup I_1}$

now $\bigcap_{i\in I} \mathrm{int}\,\mathrm{dom}\, \iota_{C_i}$ // $\mathrm{dom}\, \iota_{C_i} = C_i$ */

$= \bigcap_{i\in I} \mathrm{int}\, C_i = \bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i \neq \emptyset$ . . . (eq: 26.51) [eq: 26.52]

// $\forall_C\ \mathrm{int}\, C \subseteq \mathrm{core}\, C$ */

now $\bar{x} \in C = \bigcap_{i\in I} \mathrm{lev}_{<0}\, g_i \subseteq \bigcap_{i\in I} \mathrm{int}\,\mathrm{dom}\, g_i \Rightarrow \boxed{\forall_{i\in I}\ \bar{x} \in \mathrm{int}\,\mathrm{dom}\, g_i} \Rightarrow \bar{x} \in \mathrm{core}\,\mathrm{dom}\, g_i$

$\subseteq \mathrm{int}\,\mathrm{dom}\, g_i$ // from [eq: 26.44]

$\tilde{C} := \mathrm{dom}\, g_i$: convex

Proposition 4.45: [Relationship between "weak" and "strong" core */

[$\tilde{C}$: strongly convex subset of H, $\bar{x} \in \tilde{C}$] (1) $\bar{x} \in \mathrm{core}\,\tilde{C}$, $N_{\tilde{C}}\bar{x} = \{0\}$

(2) assume $\tilde{C} = \mathrm{cone}\,\tilde{C}$

$\Rightarrow \tilde{C} = \{0\}$

$\therefore$ $N_{\mathrm{dom}\, g_i}\, \bar{x} = \{0\}$ . . . (eq: 26.53)

now: $\mathrm{lev}_{<0}\, g_i$

$N_C\bar{x}$ // $C = \bigcap_{i\in I} C_i$ ∧ Example 4.11 [C_1, C_2 closed convex, $C_1 \cap C_2 \neq \emptyset$] */

$= \partial\iota_C\,\bar{x} = \partial\iota_{\bigcap_{i\in I} C_i}(\bar{x})$ // by definition, $\iota_{\bigcap_{i\in I} C_i} = \sum_{i\in I} \iota_{C_i}$ */

$= \partial\left(\sum_{i\in I} \iota_{C_i}\right)(\bar{x})$ // Theorem 16.39 [m: mean, $m \geq 1$, $I = \{1,\dots,m\}$; now set: $\tilde{f}_i := \partial_{C_i} \in \Gamma_0(H)$

$\{f_1,\dots,f_m\} \subset \Gamma_0(H)$, some of the following holds:

Now:

$N_C\bar{x} \;/\!/\; C = \bigcap_{i\in I} C_i \;\wedge\;$ [illegible]

$= \partial_C \bar{x} = \partial_{\bigcap_{i\in I} C_i}(\bar{x})$ // By definition, $C = \bigcap_{i\in I} C_i$ //

$= \partial\Big(\sum_{i\in I} \partial_{C_i}\Big)(\bar{x})$

$= \sum_{i\in I} \underbrace{\partial_{C_i}(\bar{x})}_{N_{C_i}}$

$= \sum_{i\in I} N_{C_i}(\bar{x})$

$= \sum_{i\in I_+} N_{C_i}(\bar{x}) + \sum_{i\in I_-} N_{C_i}(\bar{x})$

$= \sum_{i\in I_+} \big(\bigcup_{v_i\in\mathbb{R}_+} v_i\,\partial g_i(\bar{x})\big)$

$\subseteq \sum_{i\in I_+} \bigcup_{v_i\in\mathbb{R}_+} v_i\,\partial g_i(\bar{x})$

Theorem 18.19 [m: main, aux ; I = {1, ..., m} ; now set: $\tilde{f}_i := \partial_{C_i} \in \Gamma_0(M)$, $\tilde{g}_i : \to 0$, closed, convex]
$(\tilde{f}_i)_{i\in I} \subseteq \Gamma_0(M)$, one of the following holds:
(i) $\operatorname{int}\bigcap_{i\in I}(\operatorname{dom}\tilde{f}_i - \bigcap_{i\in I}\operatorname{dom}\tilde{f}_i)$
(ii) $\forall_{i=1,\dots,m}$ $\operatorname{dom}\tilde{f}_i - \bigcap_{i=1}^{m}\operatorname{dom}\tilde{f}_i$: closed linear subspace
(iii) $(\operatorname{dom}\tilde{f}_i)_{i\in I}$: linear subspaces, $\forall_{i=1,\dots,m}$ $\operatorname{dom}\tilde{f}_i + \bigcap_{i=1}^{m}\operatorname{dom}\tilde{f}_i$: closed
(iv) $\operatorname{dom}\tilde{f}_m \cap \big(\bigcap_{i=1}^{m-1}\operatorname{int}\operatorname{dom}\tilde{f}_i\big) \neq \emptyset$ ✓ /!/ : Jay: 26.52 $\bigcap_{i\in I}\operatorname{int}\operatorname{dom}\tilde{f}_i = \bigcap_{i\in I}\operatorname{int}\operatorname{dom} f_i$, $\operatorname{int}\operatorname{dom} f_m \subseteq \operatorname{dom} f_m$ $\Rightarrow \operatorname{dom} f_m \cap \big(\bigcap_{i=1,\dots,m-1}\operatorname{int}\operatorname{dom} f_i\big) \neq \emptyset$ //
(v) M : finite-dimensional, $\bigcap_{i=1}^{m}\operatorname{ri}\operatorname{dom}\tilde{f}_i \neq \emptyset$] $\partial\big(\sum_{i\in I}\tilde{f}_i\big) = \sum_{i\in I}\partial\tilde{f}_i$

$\therefore \partial\sum_{i\in I} f_i = \sum_{i\in I}\partial f_i$ //

/! set: $\tilde{g} := g$, $\tilde{f} := \mathrm{Ry}_{\mathbb{R}_+}\tilde{g} = \mathrm{Ry}_{\mathbb{R}_+} g = c$; lives our def

$\tilde{x} := \bar{x}$

Lemma 18.17 [$\tilde{f}: M \to]-\infty,+\infty]$: convex, $\operatorname{Im}_{\mathbb{R}_+}\tilde{f} \neq \emptyset$; $\tilde{f} = \mathrm{Ry}_{\mathbb{R}_+}\tilde{g}$; $x \in C$]

$N_C\tilde{x} = \begin{cases} N_{\operatorname{dom}\tilde{f}}\tilde{x} \cup \operatorname{cone}\partial\tilde{g}(\tilde{x}), & \text{if } \tilde{g}(\tilde{x}) = 0 \\ N_{\operatorname{dom}\tilde{f}}\tilde{x}, & \text{if } \tilde{g}(\tilde{x}) < 0 \end{cases}$

$N_{\mathbb{R}_+}\bar{x} = \begin{cases} N_{\operatorname{dom}\mathrm{Ry}_{\mathbb{R}_+}}\bar{x} \cup \operatorname{cone}\partial g(\bar{x}), & \text{if } g(\bar{x}) = 0 \Leftrightarrow i \in I_+ \text{ ; Ray def} \\ N_{\operatorname{dom}\mathrm{Ry}_{\mathbb{R}_+}}\bar{x} = \{0\}, & \text{if } g(\bar{x}) < 0 \Leftrightarrow i \in I_- \text{ ; } \dots \end{cases}$

$\therefore \forall_{i\in I_+}$ $N_{\mathbb{R}_+}\bar{x} = \{0\} \cup \operatorname{cone}\partial g(\bar{x}) = \operatorname{conv}(\underbrace{\partial g(\bar{x})}_{\text{closed, convex as } g \in \Gamma_0(\text{Jay: proposition 11.5})}) = R_{+\infty}\,\partial g(\bar{x}) \cup \{0\}$ // $\forall_{v_i\in\mathbb{R}_+}$ $\operatorname{cone} C = \mathbb{R}_{++}C$; $C = \bigcup_{v_i\in\mathbb{R}_+} v_i\,\partial g(\bar{x}) \cup \{0\} = \bigcup_{v_i\in\mathbb{R}_+} v_i\,\partial g(\bar{x})$ // this '+' is to avoid lack of associated g_i

$\forall_{i\in I_-}$ $N_{\mathbb{R}_+}\bar{x} = \{0\}$

/! By def: $[\bigcup_{\lambda\in\Lambda} R] \cap C = \bigcup_{\lambda\in\Lambda} \lambda C$ //

Now, Jay: 26.50 : $\bar{x} \in \operatorname{Argmin}_C f \Leftrightarrow \exists_{u\in N_C\bar{x}}$ $-u \in \partial f(\bar{x}) \Leftrightarrow \exists_{u = \sum_{i\in I_+} v_i u_i : v_i \in \mathbb{R}_+,\ u_i \in \partial g_i(\bar{x})}$ $-u \in \partial f(\bar{x})$

$\therefore N_C\bar{x} = \sum_{i\in I_+} \bigcup_{v_i\in\mathbb{R}_+} v_i\,\partial g_i(\bar{x}) = \sum_{i\in I_+} \underbrace{\{v_i u_i \mid v_i \in \mathbb{R}_+,\ u_i \in \partial g_i(\bar{x})\}}_{\{u_i \mid u_i \in \partial g_i(\bar{x})\}} = \Big\{\sum_{i\in I_+} v_i u_i \;\Big|\; \forall_{i\in I_+}\ v_i \in \mathbb{R}_+,\ u_i \in \partial g_i(\bar{x})\Big\}$

/! makes sense, visualize small examples //

/! now we want to extend the index set to $I = I_+ \cup I_-$, so that still holds

now we design $(v_i)_{i\in I_-}, (u_i)_{i\in I_-}$ as follows:

$\left.\begin{array}{l} (u_i)_{i\in I_-} \in \bigtimes_{i\in I_-} \partial g_i(\bar{x}) \\ \text{now we want } \forall_{i\in I_-}\ v_i g_i(\bar{x}) = 0 \Rightarrow (v_i)_{i\in I_-} = (0) \end{array}\right\}$ ISO, $\sum_{i\in I_-} v_i u_i = 0$ is maintained

$g_i(\bar{x}) < 0$

again $\forall_{i\in I_-,\ v_i g_i(\bar{x}) = 0}$ $v_i g_i(\bar{x}) = 0$ $\Big\}$ $\boxed{\forall_{i\in I_-}\ v_i g_i(\bar{x}) = 0}$

also, as $\bar{x} \in \operatorname{Argmin}_C f \Rightarrow \bar{x} \in C \Leftrightarrow \boxed{\forall_{i\in I_-}\ g_i(\bar{x}) \leq 0}$ //

$\Leftrightarrow \exists_{u = \sum_{i\in I} v_i u_i : \forall_{i\in I}\ v_i \in \mathbb{R}_+,\ u_i \in \partial g_i(\bar{x})}$ $-u \in \partial f(\bar{x})$ $\wedge$ $\forall_{i\in I}$ $\left\{\begin{array}{l} g_i(\bar{x}) \leq 0 \\ v_i g_i(\bar{x}) = 0 \end{array}\right.$

$\therefore$ We have arrived at the key claim:

$\bar{x}$: solves $\begin{pmatrix} \min_{x\in M} f(x) \\ \text{s.t. } g_1(x) \leq 0 \\ \vdots \\ g_m(x) \leq 0 \end{pmatrix} \Leftrightarrow \exists_{(v_i)_{i\in I} \in \mathbb{R}^m} \exists_{(u_i)_{i\in I} \in \bigtimes_{i\in I} \partial g_i(\bar{x})}$ $-\sum_{i\in I} v_i u_i \in \partial f(\bar{x})$ $\wedge$ $\forall_{i\in I} \left\{\begin{array}{l} g_i(\bar{x}) \leq 0 \\ \bar{v}_i g_i(\bar{x}) = 0 \end{array}\right.$

naturally this means: $\sum_{i\in I} \underbrace{\bar{v}_i u_i}_{\in \partial g_i(\bar{x})} + \partial f(\bar{x}) \ni 0$

$\Leftrightarrow \sum_{i\in I} \bar{v}_i \partial g_i(\bar{x}) + \partial f(\bar{x}) \Leftrightarrow \partial\Big(\sum_{i\in I} \bar{v}_i g_i(\bar{x}) + f(\bar{x})\Big) \ni 0$

$\Leftrightarrow$ $\bar{x}$: solves $\min_{x\in M} f(x) + \sum_{i\in I} \bar{v}_i g_i(x)$

clearly $\bar{v}_i$ are the associated Lagrange multipliers.

The claims for Gateaux differentiability follows from

Proposition 17.24 [$f: M \to]-\infty,+\infty]$, proper, convex ; $x \in \operatorname{dom} f$]
(i) f : Gateaux differentiable at x $\Rightarrow$ $\partial f(x) = \{\nabla f(x)\}$
$\operatorname{ri}\big(\{x \in \operatorname{dom} f \mid \partial f(x) = \{u\}\}\big) \Rightarrow \big\{f : \text{Gateaux differentiable at } x \text{ ; } u = \nabla f(x)\big\}$

set $k = \mathbb{R}^m$, $K = \mathbb{R}^m_+$; $R: x \mapsto (g_i(x))_{i=1}^m$

Proposition 19.15 [$F \in \Gamma_0(M)$, K: nonempty, closed, convex ; $F: M \times K \to]-\infty,+\infty]$: $(x,k) \mapsto \begin{cases} f(x), & \text{if } Rx \in -K \\ +\infty, & \text{if } Rx \notin -K \end{cases}$
(i) FEQ(MOC)
(ii) Primal problem: $\min_{x\in M} f(x)$ s.t. $Rx \in -K$
(iii) Dual problem: $\min_{v\in K^\oplus} \Phi(v)$ solves
(iv) The Lagrangian: $L: M \times K \to [-\infty,+\infty]$: $(x,v) \mapsto \begin{cases} -\infty & \\ f(x) + \langle Rx, v\rangle & \\ +\infty & \end{cases}$
(v) $(\bar{x},\bar{v})$: saddle point of L $\Leftrightarrow$
$\bar{x} \in \operatorname{dom} f$
$R\bar{x} \in K$
$\bar{v} \in K^\oplus$
$f(\bar{x}) = \inf_{x\in M}\big(f(x) + \langle Rx, \bar{v}\rangle\big)$
in which case: $\langle R\bar{x}, \bar{v}\rangle = 0$ ∧ L: primal
Remark 19.16 $(\bar{x},\bar{v})$: saddle point of L $\Leftrightarrow$ $\bar{x} \in \operatorname{FEQ}(\text{MOC})$

using

$\Phi(\bar{x})$:

[illegible] convex cone in E : E: continuous linear w.r.t. K, $E(\Omega(\mathbb{R}^n))\neq\emptyset$;

$\Rightarrow$

$\Phi : \forall x \in \bigcup_{x\in K} \{ (\Phi(x)-\Phi(\bar{x})$

- if $x\in\mathrm{int}\,K$, $v\in K^{\oplus}$
- if $x\in\mathrm{bd}\,K$, $v\in K^{\oplus}$
- if $x\notin K$

[illegible] solution

Φ | $\bar{\gamma}$: Lagrange multiplier to $\bar{x}$

$\bar{x}$: solves $\min_{x\in\Omega} \{\Phi(x)+\langle x,\bar{\gamma}\rangle\}$

Part 1

November 2, 2016 5:18 PM

Corollary 27.4 (Douglas-Rachford algorithm)

$[\, f,g \in \Gamma_0(\mathcal{H}),\ \operatorname{ri}(\operatorname{dom} f - \operatorname{dom} g) \neq \emptyset;\ (\lambda_n)_{n\in\mathbb{N}} : \subseteq [0,2],\ \sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty;\ \forall \varepsilon \in \mathbb{R}_{++} : x_\varepsilon \in \mathcal{H};$ set:

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = \operatorname{Prox}_{\gamma g} x_n \\ z_n = \operatorname{Prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{array} \right] \Rightarrow \exists_{x\in\mathcal{H}} \left[\begin{array}{l} \text{(i) } \operatorname{Prox}_{\gamma g} x \in \operatorname{Argmin}(f+g) \\ \text{(ii) } x_n - z_n \to 0 \\ \text{(iii) } x_n \to 0 \\ \text{(iv) } y_n \to \operatorname{Prox}_{\gamma g} x \\ \quad z_n \to \operatorname{Prox}_{\gamma f} x \\ \text{(v) (one of the following holds:} \\ \quad \text{(a) } f \text{: uniformly convex on every nonempty bounded subset of } \operatorname{dom} f) \\ \quad \text{(b) } g \text{: uniformly convex on every nonempty bounded subset of } \operatorname{dom} g) \\ \Rightarrow \\ \quad y_n \to \operatorname{Prox}_{\gamma g} x : \operatorname{Argmin}(f+g) = \{\operatorname{Prox}_{\gamma g} x\} \\ \quad z_n \to \operatorname{Prox}_{\gamma f} x : \operatorname{Argmin}(f+g) = \{\operatorname{Prox}_{\gamma f} x\} \end{array} \right]$$

2016-11-03 10:14 AM

Proof: Set $A := \partial f$, $B := \partial g$

$f, g \in \Gamma_0(\mathcal{H}) \Rightarrow \partial f, \partial g$: maximally monotone

$\Downarrow$

$f+g \in \Gamma_0(\mathcal{H}) \Rightarrow \partial(f+g) = \partial f + \partial g$: maximally monotone

$L = Id$

/* Set $K := \mathcal{H}$; $L = Id$. from Proposition 23.2: $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$

Proposition 16.5 [K: real Hilbert space; $f : \mathcal{H} \to]-\infty, +\infty]$; $g : K \to]-\infty, +\infty]$, proper; $L \in \mathcal{B}(\mathcal{H}, K)$; $\lambda \in \mathbb{R}_{++}$]

(i) $\partial(\lambda f) = \lambda \partial f$

(ii) $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset \Rightarrow \partial f + L^*(\partial g) \circ L \subseteq \partial(f + g \circ L)$

$\partial f + \partial g \subseteq \partial(f+g)$

$\Rightarrow \operatorname{zer}(\partial f + \partial g) \subseteq \operatorname{zer} \partial(f+g) = \operatorname{Argmin}(f+g)$

/* from Fermat's rule */

now the claims follow from the result below:

Example 15.5

$[\, f \in \Gamma_0(\mathcal{H}),\ x \in \operatorname{dom} f \Rightarrow \left| \begin{array}{l} \operatorname{Prox}_f x = J_{\partial f} x \\ \partial(f^*)(x) = J_{\partial f}(x) \end{array} \right.$ Since J: Normal cone is 1-1.

Proposition 23.7 [Douglas-Rachford operator] Set $A := \partial f$, $B := \partial g$

[A, B: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$; $\operatorname{zer}(A+B) \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$; $\forall \varepsilon \in \mathbb{R}_{++} : x_\varepsilon \in \mathcal{H}$;

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \end{array} \right] \Rightarrow x_n \to x$$

$\Rightarrow$ [x is a solution

$\Rightarrow (z_n - y_n)_{n\in\mathbb{N}} \to 0$

$\Rightarrow y_n \to x$

$\Rightarrow z_n \to x$

$\Rightarrow (x_n)_{n\in\mathbb{N}}$ converges to a point in $\operatorname{zer}(A+B)$]

Proposition 16.4

[K: real Hilbert space; $f : K \to]-\infty, +\infty]$, proper, strongly convex, $\gamma \in \mathbb{R}_{++}$

$\Rightarrow \partial f$: strongly monotone w.c]

■

Proposition 23.8. (Parallel Splitting Algorithm) [m: $\mathbb{N}\setminus\{1\}$; $1 : \{1, \dots, m\}$; $(\varepsilon_i)_{i\in I} : \subseteq]0, 1[$;

Problem: $\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x)$: has atleast one solution; one of the following holds:

(i) $0 \in \bigcap_{i=1}^{m} \operatorname{sri}(\operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j)$, (ii) $\operatorname{dom} f_1 \cap (\bigcap_{i=1}^{m} \operatorname{int} \operatorname{dom} f_i) \neq \emptyset$, (iii) $\mathcal{H}$: finite-dimension

and $\bigcap_{i\in I} \operatorname{ri} \operatorname{dom} f_i \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$; $\forall \varepsilon \in \mathbb{R}_{++}$; $(x_{i\varepsilon})_{i\in I} \in \mathcal{H}^m$; set:

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} p_n = \frac{1}{m} \sum_{i\in I} x_{i,n} \\ \forall_{i\in I}\ y_{i,n} = \operatorname{Prox}_{\varepsilon f_i} x_{i,n} \\ q_n = \frac{1}{m} \sum_{i\in I} y_{i,n} \\ \forall_{i\in I}\ x_{i,n+1} = x_{i,n} + \lambda_n(2q_n - p_n - y_{i,n}) \end{array} \right] \Rightarrow p_n \to p : p \in \operatorname{Argmin}_{x\in\mathcal{H}} \sum_{i\in I} f_i(x)$$

2016-11-05 10:08 AM

Proof:

Recall Proposition 23.9. Set $A_i := \partial f_i$ and recall $J_{\gamma A_i} = J_{\gamma \partial f_i} = \operatorname{Prox}_{\gamma f_i}$ for A_i: maximally monotone

Proposition 23.7. (Parallel Splitting Algorithm)

[$m \in \mathbb{N}\setminus\{1\}$; $I = \{1, \dots, m\}$; $(A_i)_{i\in I}$: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, $\operatorname{zer} \sum_{i\in I} A_i \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}} : \subseteq [0,2]$, $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$; $\forall \varepsilon \in \mathbb{R}_{++}$; $(x_{i\varepsilon})_{i\in I} \in \mathcal{H}^m$; $(x_{i,n})_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$;

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} p_n = \frac{1}{m} \sum_{i\in I} x_{i,n} \\ \forall_{i\in I}\ y_{i,n} = J_{\gamma A_i} x_{i,n} = \operatorname{Prox}_{\gamma f_i} x_{i,n} \\ q_n = \frac{1}{m} \sum_{i\in I} y_{i,n} \\ \forall_{i\in I}\ x_{i,n+1} = x_{i,n} + \lambda_n(2q_n - p_n - y_{i,n}) \end{array} \right] \Rightarrow (p_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} \sum_{i\in I} A_i$

$\operatorname{zer} \sum_{i\in I} \partial f_i$

$\therefore (p_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} \sum_{i\in I} \partial f_i$

now

(sufficient condition for additivity of subdifferentials) ★

Theorem 16.39 [$m : \mathbb{N}\setminus\{1\}$; $I = \{1, \dots, m\}$;

$(f_i)_{i\in I} : \subseteq \Gamma_0(\mathcal{H})$, one of the following holds:

(i) $0 \in \bigcap_{i=1}^{m} \operatorname{sri}(\operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j)$

(ii) $\forall_{i\in\{1,\dots,m\}}\ \operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j$: closed linear subspace

(iii) $(\operatorname{dom} f_i)_{i\in I}$: linear subspaces, $\forall_{i\in\{1,\dots,m\}}\ \operatorname{dom} f_i = \bigcap_{j=1}^{i-1} \operatorname{dom} f_j$: closed

(iv) $\operatorname{dom} f_m \cap (\bigcap_{i=1}^{m-1} \operatorname{int} \operatorname{dom} f_i) \neq \emptyset$

(v) $\mathcal{H}$: finite-dimensional, $\bigcap_{i=1}^{m} \operatorname{ri} \operatorname{dom} f_i \neq \emptyset$] $\partial(\sum_{i=1}^{m} f_i) = \sum_{i=1}^{m} \partial f_i$

$\sum_{i\in I} \partial f_i = \partial \sum_{i\in I} f_i$

$\therefore (p_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} \partial \sum_{i\in I} f_i = \operatorname{Argmin} \sum_{i\in I} f_i$

■

Corollary 23.9. (Forward-backward algorithm)

[$f \in \Gamma_0(\mathcal{H})$; $g : \mathcal{H} \to \mathbb{R}$, convex, differentiable, $\frac{1}{\beta}$-Lipschitz-continuous gradient with $\beta \in \mathbb{R}_{++}$;

$\forall \varepsilon \in]0, \iota\beta[$; $\delta = \min\{1, \beta/\iota\} + \frac{\varepsilon}{2}$; $(\lambda_n)_{n\in\mathbb{N}} : \subseteq [\varepsilon, \delta]$, $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$; $x_\varepsilon \in \mathcal{H}$; $\operatorname{Argmin}(f+g) \neq \emptyset$;

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = x_n - \gamma \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n(\operatorname{Prox}_{\gamma f} y_n - x_n) \end{array} \right] \Rightarrow$$

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Argmin}(f+g)$

(ii) $(\inf_{n\in\mathbb{N}} \lambda_n > 0$; $x \in \operatorname{Argmin}(f+g)) \Rightarrow \nabla g(x_n) \to \nabla g(x)$

(iii) $(\inf_{n\in\mathbb{N}} \lambda_n > 0$; one of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of $\operatorname{dom} \partial f$

(b) g: uniformly convex on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$x_n \to$ unique minimizer of $f+g$

2016-11-07 10:08 AM

Proof:

Set: $\left. \begin{array}{l} A = \partial f \\ B = \nabla g \end{array} \right\} \Rightarrow A, B$: maximally // subdifferential operator of a Γ_0 function is maximally monotone

now: $\operatorname{dom} g = \mathcal{H}$ /* per absurdum: $\exists_{\tilde{x}\in\mathcal{H}}\ g(\tilde{x}) = +\infty$

$\Rightarrow \operatorname{int} \operatorname{dom} g = \mathcal{H}$

g: cannot be differentiable at $\tilde{x}$ #!

$\therefore \underbrace{\operatorname{dom} f}_{\text{convex}} \cap \underbrace{\operatorname{int}(\operatorname{dom} g)}_{\mathcal{H}\text{: convex}} \neq \emptyset$

as $f \in \Gamma_0(\mathcal{H})$

else

(sufficient condition for 0 belonging to the strong relative interior of a convex set) ★

$\hookrightarrow$ g: cannot be differentiable at $\bar{x}$ #!

$\therefore \underbrace{\operatorname{dom} f}_{\text{convex}} \cap \underbrace{\operatorname{int}(\operatorname{dom} g)}_{\text{convex}} \neq \emptyset$

as $f \in \Gamma_0(\mathcal{H})$

$\mathcal{H}$: convex

$\Downarrow$ SR

(sufficient condition for a subdifferential to be strong relative interior of a convex set) #!

Set: $K := \mathcal{H}$, $L := \mathrm{Id}: \mathcal{H} \to \mathcal{H}$; $C := \operatorname{dom} f$, $D := \operatorname{dom} g$

Proposition 6.19 [C: convex subset of $\mathcal{H}$, K: real Hilbert space, $L \in \mathcal{B}(\mathcal{H}, K)$, D: convex subset of K]

Suppose that C, D satisfy the following holds:

(i) $D - L(C)$: closed linear subspace

(ii) C, D: linear subspaces and [• $D + L(C)$: closed ✓

• D: closed, $L(C)$: finite-dimensional or finite co-dimensional ✓

• D: finite-dimensional or finite-codimensional and $L(C)$: closed]

(iii) D: cone, $D - L(C)$: closed linear subspace

(iv) $D = L(C)$, given D: closed

(v) $0 \in \operatorname{core}(D - L(C))$

(vi) $0 \in \operatorname{int}(D - L(C))$

(vii) $\operatorname{ri} D \cap \operatorname{ri} L(C) \neq \emptyset$ ✓

$\boxed{L(C) \cap \operatorname{int} D \neq \emptyset}$

(viii) K: finite-dimensional and $(\operatorname{ri} D) \cap \operatorname{ri} L(C) \neq \emptyset$

(ix) K: finite-dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

(x) $(\mathcal{H}, K)$: finite dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

then $0 \in \operatorname{ri}(D - L(C))$.

$0 \in \operatorname{ri}(\operatorname{dom} g - \operatorname{dom} f)$

now recall:

(Properties of minimizer for a sum of two Γ_0 functions) #!

Corollary 26.5 [$f, g \in \Gamma_0(\mathcal{H})$; $\bar{x} \in \mathcal{H}$; $\gamma \in \mathbb{R}_{++}$; one of the following holds:

(a) $0 \in \operatorname{sri}(\operatorname{dom} g - \operatorname{dom} f)$ #! see Proposition 6.19 for sufficient conditions for this #!

(b) $\mathcal{H}$: finite-dimensional, g: polyhedral, $\operatorname{dom} g \cap \operatorname{ri} \operatorname{dom} f \neq \emptyset$

(c) $\mathcal{H}$: finite-dimensional; f, g: polyhedral; $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$]. The following are equivalent

(i) $\bar{x}$: solution to the problem: $\min_{x \in \mathcal{H}} f(x) + g(x)$; (ii) $\bar{x} \in \operatorname{zer}(\partial f + \partial g)$ $\Rightarrow$ (iii) $\operatorname{Argmin}(f+g) = \operatorname{zer}(\partial f + \partial g)$ #!

(iii) $\bar{x} \in \operatorname{prox}_{\gamma g}(\operatorname{prox}_{\gamma f}(\bar{x} - \gamma \bar{x}) + (\operatorname{prox}_{\gamma f}(\bar{x} - \gamma \bar{x}))$

(iv) $\exists_{u \in \partial f(\bar{x})} \quad -u \in \partial g(\bar{x})$

(v) $\exists_{u \in \partial f(\bar{x})} \ \forall_{x \in \mathcal{H}} \quad \langle \bar{x} - x, u \rangle + f(\bar{x}) \leq g(x)$

Moreover, if g: Gâteaux differentiable at $\bar{x}$, each of the items (i)-(v) is also equivalent to each of the following:

(vi) $-\nabla g(\bar{x}) \in \partial f(\bar{x})$

(vii) $\forall_{x \in \mathcal{H}} \quad \langle \bar{x} - x, \nabla g(\bar{x}) \rangle + f(\bar{x}) \leq f(x)$

(viii) $\bar{x} = \operatorname{prox}_{\gamma f}(\bar{x} - \gamma \nabla g(\bar{x}))$

$\therefore$ $\operatorname{Argmin}(f+g) = \operatorname{zer}(A+B)$

now:

(proximal splitting algorithm) #!

[$f: \mathcal{H} \to \mathbb{R}$, finite differentiable, convex; $f \in \Gamma_0(\mathcal{H})$ #!

• ∇f: $\frac{1}{\beta}$-Lipschitz continuous $\Leftrightarrow$ ∇f: cocoercive

• ∇f: nonexpansive $\Leftrightarrow$ ∇f: firmly nonexpansive

$\hookrightarrow$ $B = \nabla f$: β-cocoercive

finally we use the Theorem 25.8: which is the forward-backward algorithm in its general form:

Theorem 25.8 (Forward-backward algorithm) #! B: firmly nonexpansive

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone; $\beta \in \mathbb{R}_{++}$; $B: \mathcal{H} \to \mathcal{H}$, β-cocoercive; $\gamma \in]0, 2\beta[$; $\delta := \min\{1, \frac{\beta}{\gamma}\} + \frac{1}{2}$;

$(\lambda_n)_{n \in \mathbb{N}}$: sequence in $[0, \delta]$, $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$; $x_0 \in \mathcal{H}$; $\operatorname{zer}(A+B) \neq \emptyset$;

$$\forall_{n \in \mathbb{N}} \quad \begin{cases} y_n = x_n - \gamma B x_n = x_n - \gamma \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n(J_{\gamma A} y_n - x_n) \end{cases} \Rightarrow$$

$= x_n + \lambda_n(\operatorname{prox}_{\gamma g} y_n - x_n)$

(i) $(x_n)_{n \in \mathbb{N}}$: converges weakly to a point in $\operatorname{zer}(A+B)$

(ii) $(\inf_{n \in \mathbb{N}} \lambda_n > 0$, $x \in \operatorname{zer}(A+B)) \Rightarrow (Bx_n)_{n \in \mathbb{N}}$: $Bx_n \to Bx$

(iii) $(\inf_{n \in \mathbb{N}} \lambda_n > 0$; one of the following holds:

$\Rightarrow$ (a) A: uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$

(b) B: uniformly monotone on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$(x_n)_{n \in \mathbb{N}}$: converges strongly to the unique point in $\operatorname{zer}(A+B)$

Example 25.5

[$f \in \Gamma_0(\mathcal{H})$, $\forall \in \mathbb{R}_{++}$] $\Rightarrow$ $\begin{cases} \operatorname{prox}_{\gamma f}: \mathcal{H} \to \mathcal{H} \\ \operatorname{id} - \operatorname{prox}_{\gamma f}: \mathcal{H} \to \mathcal{H} \end{cases}$

Example 23.4

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

$\Rightarrow$ $\operatorname{dom} S = \mathcal{H}$

S: uniformly monotone in C

Proposition 27.15 (Primal-dual algorithm) [$\Phi \in \Gamma_0(\mathcal{H})$; $\Psi \in \Gamma_0(K)$; $z \in \mathcal{H}$; $r \in K$; $L \in \mathcal{B}(\mathcal{H}, K)$, $L \neq 0$, $\operatorname{resri}(L(\operatorname{dom} \Phi) - \operatorname{dom} \Psi)$;

Primal problem:

$$\min_{x \in \mathcal{H}} \ \Phi(x) + \Psi(Lx - r) + \frac{1}{2}\|x - z\|^2 \qquad \text{[eq: 27.32]}$$

Dual problem:

$$\min_{v \in K} \ (\Phi^*)(L^*v + z) + \Psi^*(-v) - \langle v | r \rangle \qquad \text{[eq: 27.33]}$$

$\forall \in]0, \frac{2}{\|L\|^2}[$; $\delta := \min\{1, \frac{1}{\forall \|L\|^2} + \frac{1}{2}\}$; $(\lambda_n)_{n \in \mathbb{N}}: \subseteq]0, \delta[$, $\inf_{n \in \mathbb{N}} \lambda_n > 0$, $\sup_{n \in \mathbb{N}} \lambda_n < \delta$; $v_0 \in K$;

Set:

$$\forall_{n \in \mathbb{N}} \quad \begin{cases} x_n = \operatorname{prox}_{\Phi}(L^* v_n + z) \\ v_{n+1} = v_n - \lambda_n(\operatorname{prox}_{\forall \Psi^*}(\forall(Lx_n - r) - v_n) + v_n) \end{cases} \qquad \text{[eq: 27.34]}$$

$\bar{x}$: unique solution to the primal problem] $\Rightarrow$

(i) $v_n \to \bar{v}$: solution to the dual problem

$\bar{x} = \operatorname{prox}_{\Phi}(L^* \bar{v} + z)$

(ii) $x_n \to \bar{x}$

2016-11-17 10:55 AM

Proof:

Set $h: \mathcal{H} \to]-\infty, +\infty]: x \mapsto \Phi(x) + \frac{1}{2}\|x - z\|^2 \in \Gamma_0(\mathcal{H}) \Rightarrow \operatorname{dom} \Phi = \operatorname{dom} h$ (z: fixed vector)

$j: K \to]-\infty, +\infty]: y \mapsto \Psi(y - r) \in \Gamma_0(K) \Rightarrow \operatorname{dom} j = \operatorname{dom} \Psi$

$\therefore j(y) = \Psi(\underbrace{y - r}_{\tilde{y}}) \Leftrightarrow \Psi(\tilde{y}) = j(\tilde{y} + r)$

then primal problem: $\left(\min_{x \in \mathcal{H}} \underbrace{\Phi(x) + \frac{1}{2}\|x - z\|^2}_{h(x)} + \underbrace{\Psi(Lx - r)}_{j(Lx)}\right) = \left(\min_{x \in \mathcal{H}} h(x) + j(Lx)\right) = \min(h + j \circ L)(\mathcal{H})$ [eq: 27.35]

given $\operatorname{resri}(L(\operatorname{dom} \Phi) - \operatorname{dom} \Psi)$ #! $\bar{x} \in \operatorname{sri} C \Leftrightarrow \operatorname{cone}(C - \bar{x}) = \overline{\operatorname{span}}(C - x)$ #!

$\Leftrightarrow \operatorname{cone}(L(\operatorname{dom} \Phi) - \operatorname{dom} \Psi - r) = \overline{\operatorname{span}}(L(\operatorname{dom} \Phi) - \operatorname{dom} \Psi - r)$

now $L(\operatorname{dom} \Phi) - \operatorname{dom} \Psi - r$

$= L\{x \in \mathcal{H} \mid \Phi(x) < +\infty\} - \{y \in K \mid \Psi(y) < +\infty\} - r$

$= \{Lx \in K \mid \Phi(x) < +\infty\} - \{y \in K \mid \Psi(y) < +\infty\} - r$

$= \{Lx - y - r \mid \Phi(x) < +\infty,\ \Psi(y) < +\infty\}$

$\Phi(x) + \frac{1}{2}\|x - z\|^2 < +\infty$ ($= h(x)$); $j(y + r) < +\infty$

$= \{Lx - (y + r) \mid h(x) < +\infty,\ j(y + r) < +\infty\}$

$= L\{x \mid h(x) < +\infty\} - \{\underbrace{y + r}_{\tilde{y} \text{ say}} \mid j(y + r) < +\infty\}$

$= \underbrace{L\{x \mid h(x) < +\infty\}}_{\operatorname{dom} h} - \underbrace{\{\tilde{y} \mid j(\tilde{y}) < +\infty\}}_{\operatorname{dom} j} = L(\operatorname{dom} h) - \operatorname{dom} j$

$\operatorname{cone}(L(\operatorname{dom} h) - \operatorname{dom} j) = \overline{\operatorname{span}}(L(\operatorname{dom} h) - \operatorname{dom} j)$

$\Leftrightarrow \boxed{0 \in \operatorname{sri}(L(\operatorname{dom} h) - \operatorname{dom} j)}$ #! $\bar{x} \in \operatorname{sri} C \Leftrightarrow \operatorname{cone}(C - \bar{x}) = \overline{\operatorname{span}}(C - x)$ #!

$\Leftrightarrow 0 \in \operatorname{sri}(\operatorname{dom} j - L(\operatorname{dom} h))$

now recall:

$\operatorname{cone}(L(\operatorname{dom} h) - \operatorname{dom} j) = \overline{\operatorname{span}}(L(\operatorname{dom} h) - \operatorname{dom} j)$

$\Leftrightarrow \boxed{0 \in \operatorname{sri}(L(\operatorname{dom} h) - \operatorname{dom} j)}$ // $\bar{x} \in \operatorname{sri} C \Leftrightarrow \operatorname{cone}(C - \bar{x}) = \overline{\operatorname{span}}(C - x)$ //

$\Leftrightarrow 0 \in \operatorname{sri}(\operatorname{dom} j - L(\operatorname{dom} h))$

now recall

$\bar{\xi}^{\mathcal{H}} \xi = h \quad \bar{\eta} = \xi$

Theorem 15.23 // [Fenchel Rockafellar]
$\{\xi \in \xi(\mathcal{H}), \eta \in \xi(K), L \in \mathcal{B}(\mathcal{H}, K)\}$
$0 \in \operatorname{sri}(\operatorname{dom} \xi - L(\operatorname{dom} \xi))$ } Measuring the question a some $b \in \operatorname{sri}(\operatorname{dom} h - L\operatorname{dom} \xi)$ } makes sense because of the next proposition
$\inf(\xi + \eta \circ L)(\mathcal{H}) = -\min(\xi^* \circ L^* + \eta^*)(K)$ // ξ^*: Fenchel conj of ξ, $\xi \mapsto \xi(x)$ //

$$\inf(h + j \circ L)(\mathcal{H}) = -\min_{v \in K}(\underbrace{h^*}_{\text{conjugate function}} \circ \underbrace{L^*}_{\text{adjoint of the linear operator}} + \underbrace{j^*}_{j^*(-v)})(v) = -\min_{v \in K}(h^*(L^* v) + j^*(-v)) = -\min_{v \in K}(h^*(L^* v) + j^*(-v))$$

minimizer exists in K, so, inf on R.H.S has a minimizer in $\mathcal{H}$, $\Rightarrow$ can be replaced with min.

[eq: 27.35.1]

[eq: 27.35] : primal problem: $\min(h + j \circ L)(\mathcal{H}) = \inf(h + j \circ L)(\mathcal{H}) = -\min_{v \in K}(h^*(L^* v) + j^*(-v))$: has atleast one solution in K

This is the dual problem in min. form, now we want to show that it coincides with [eq: 27.33]

now: $h(x) = \Phi(x) + \frac{1}{2}\|x - z\|^2 = \Phi(x) + \frac{1}{2}\|x\|^2 + \frac{1}{2}\|z\|^2 - \langle x | z \rangle$

$= \Phi(x) + \frac{1}{2}\|x\|^2 + \langle x | -z \rangle + \frac{1}{2}\|z\|^2$

$\Leftrightarrow h = \Phi + \frac{1}{2}\|\cdot\|^2 + \langle \cdot | -z \rangle + \frac{1}{2}\|z\|^2$

$\therefore \boxed{h^* = (\Phi^* \,\square\, \tfrac{1}{2}\|\cdot\|^2)(\cdot + z) - \frac{1}{2}\|z\|^2}$

Proposition 15.18 // [Fenchel conjugate infimal.] //
$\{\xi \in \xi(\mathcal{H})\}$
$\forall x \in \mathcal{H}$, $\xi^*(x) = \sup_{y \in \mathcal{H}}(\langle x | y \rangle - \xi(y))$
$\forall \xi \in \xi(\mathcal{H})$, $\forall \lambda > 0$, $\forall x \in \mathcal{H}$, $(\lambda \xi)^*(x) = \lambda \xi^*(\lambda^{-1} x)$
$\forall \xi, \zeta \in \xi(\mathcal{H})$, $(\xi \,\square\, \zeta)^* = \xi^* + \zeta^*$
$\forall \xi \in \xi(\mathcal{H})$, $\forall \zeta \in \mathcal{B}(\mathcal{H}, \mathbb{R})$, $(\xi + \zeta)^* = \xi^* - \zeta$
$\forall \xi \in \xi(\mathcal{H})$, $\forall \zeta \in \mathcal{B}(\mathcal{H}, \mathbb{R})$, $(\xi \circ L + \zeta)^* = \xi^* \circ L^* + \zeta^*$
$\forall \xi \in \xi(\mathcal{H})$, $\forall \lambda > 0$, $\forall x \in \mathcal{H}$, $\xi^*(x) = \lambda \xi^*(\lambda^{-1} x)

$$\boxed{\therefore\ \nabla g(v) = L\nabla(\Phi^*)(L^*v+z)}$$ [eq: 27.37]

L: linear operator

$$\therefore\ \forall_{v\in K}\ \forall_{w\in K}\quad \|\nabla g(v)-\nabla g(w)\| = \|L\nabla(\Phi^*)(L^*v+z) - L\nabla(\Phi^*)(L^*w+z)\|$$

$$= \|L\left(\nabla(\Phi^*)(L^*v+z) - \nabla(\Phi^*)(L^*w+z)\right)\| \quad /\!*\text{ use the def of linear-operator norm }*/$$

$$\leq \|L\|\ \underbrace{\|\nabla(\Phi^*)(L^*v+z) - \nabla(\Phi^*)(L^*w+z)\|}_{/\!*\ \leq \|L^*v+z - L^*w - z\| = \|L^*v - L^*w\|\ */} \quad /\!*\text{ now from [eq: 2016-11-15 1:42 PM] } \nabla(\Phi^*):\text{ 1-Lipschitz continuous }*/$$

$$\leq \|L\|\ \underbrace{\|L^*v - L^*w\|}_{\|\cdot\|} \quad /\!*\ \|L^*(v-w)\| \leq \underbrace{\|L^*\|}_{\|L\|}\|v-w\|\ */$$

$\|L\|$ // using Fact 2.18 $\|L\| = \|L^*\|$

$$\leq \|L\|^2\|v-w\|$$

$$\therefore\ \forall_{v\in K}\ \forall_{w\in K}\quad \|\nabla g(v)-\nabla g(w)\| \leq \|L\|^2\|v-w\| \Rightarrow \nabla g:\ \|L\|^2\text{-Lipschitz continuous}$$

$$\boxed{\therefore\ \nabla g(v) = L\underbrace{\nabla(\Phi^*)}_{\mathrm{Prox}_\Phi}(L^*v+z) = L\,\mathrm{prox}_\Phi(L^*v+z)}$$ [eq: 24.37.2]

[eq: 27.38]

[eq: 27.34], [eq: 27.37] and [eq: 27.38]

$$x_n = \mathrm{Prox}_\Phi(L^*v_n+z) \Rightarrow Lx_n = L\,\mathrm{Prox}_\Phi(L^*v_n+z) = \nabla g(v_n)$$

$$v_{n+1} = v_n - \lambda_n\left(\mathrm{Prox}_{\gamma g^*}(v_n - \gamma(Lx_n - r)) - v_n\right) \quad /\!*\ \mathrm{Prox}_{\gamma g^*}(\gamma(\underbrace{Lx_n - r}_{\nabla g(v_n)})) - v_n = \mathrm{Prox}_{\gamma g^*}(\gamma(\nabla g(v_n) - r) - v_n) = \mathrm{Prox}_{\gamma g^*}(-(v_n - \gamma\nabla g(v_n) + \gamma r))$$

$$= v_n - \lambda_n\left(-\mathrm{Prox}_{\gamma f}(v_n - \gamma\nabla g(v_n)) - v_n\right) \quad /\!*\text{ using }*/$$

$$= v_n + \lambda_n\left(\mathrm{Prox}_{\gamma f}(v_n - \gamma\nabla g(v_n)) + v_n\right)$$

$$= -\mathrm{Prox}_{(\gamma g^*)^*}(v_n - \gamma\nabla g(v_n) + \gamma r) \quad /\!*\text{ using: }*/$$

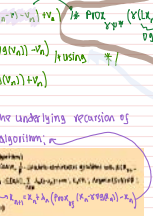
$$= -\mathrm{Prox}_{(\gamma g^*)^*}\left(v_n - \gamma\nabla g(v_n) - (-\gamma r)\right)$$

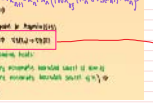
$$= -\mathrm{Prox}_{(\gamma g^*)^*}(v_n - \gamma\nabla g(v_n)) \quad /\!*\text{ using }*/$$

$$/\!*\ \gamma(\gamma^{-1}r)^* - \gamma(-(-r)) = \gamma(\underbrace{\gamma^{-1*}r - (-r)}_{\text{// by definition}}) = \gamma f\ */$$

$$= -\mathrm{Prox}_{\gamma f}(v_n - \gamma\nabla g(v_n)) \quad */$$

notice that this is the underlying recursion of forward-backward algorithm;





(i) follows from

v_n: converges weakly to a point in $\mathrm{Argmin}(f+g)$ // $f: v\mapsto\gamma^*(-v)-\langle v|r\rangle$

// $g: v\mapsto\gamma'(\Phi^*)(L^*v+z)$

denote it by $\bar{v}$

[eq: dual problem]: says solution of our dual problem is $\mathrm{Argmin}(f+g)$

$$\therefore\ \bar{x} = \mathrm{Prox}_\Phi(L^*\bar{v}+z)$$

(ii) In (i) we have shown that $v_n \to \bar{v}$: $\bar{v}\in\mathrm{Argmin}(f+g)$ defined in [eq: dual problem] (solution to dual problem)

and $\bar{x} = \mathrm{Prox}_\Phi(L^*\bar{v}+z)$: unique solution to dual problem

set $\rho = \sup_{n\in\mathbb{N}}\|v_n - \bar{v}\|$ /\!* as $(v_n)_{n\in\mathbb{N}}$: bounded $\Leftarrow$ $(v_n)_{n\in\mathbb{N}}$: converges weakly $\Leftrightarrow$ $(v_n)_{n\in\mathbb{N}}$: bounded, possesses at most one weak sequential cluster point */

$\rho < +\infty$

$$\therefore\ \forall_{n\in\mathbb{N}}\quad \|x_n - \bar{x}\|^2 = \|\mathrm{Prox}_\Phi(L^*v_n+z) - \mathrm{Prox}_\Phi(L^*\bar{v}+z)\|^2 \quad /\!*\ \mathrm{Prox}_\Phi(L^*v_n+z)\text{ by definition}$$

$$\leq \langle L^*v_n + z^* - L^*\bar{v} - z^* \mid \mathrm{Prox}_\Phi(L^*v_n+z) - \mathrm{Prox}_\Phi(L^*\bar{v}+z)\rangle$$

$$= \langle L^*(v_n - \bar{v}) \mid \mathrm{Prox}_\Phi(L^*v_n+z) - \mathrm{Prox}_\Phi(L^*\bar{v}+z)\rangle \quad /\!*\text{ linear operator }*/$$

$$= \langle v_n - \bar{v} \mid L(\mathrm{Prox}_\Phi(L^*v_n+z) - \mathrm{Prox}_\Phi(L^*\bar{v}+z))\rangle \quad /\!*\text{ now use: } \langle L^*x|y\rangle = \langle x|Ly\rangle\ */$$

$$= \langle v_n - \bar{v} \mid \underbrace{L\,\mathrm{Prox}_\Phi(L^*v_n+z)}_{\nabla g(v_n)} - \underbrace{L\,\mathrm{Prox}_\Phi(L^*\bar{v}+z)}_{\nabla g(\bar{v})}\rangle \quad /\!*\text{ from [eq: 24.37.2]: } \nabla g(v) = L\,\mathrm{Prox}_\Phi(L^*v+z)\ */$$

$$= \langle v_n - \bar{v} \mid \nabla g(v_n) - \nabla g(\bar{v})\rangle \leq \|v_n - \bar{v}\|\,\|\nabla g(v_n) - \nabla g(\bar{v})\| \quad \forall_{n\in\mathbb{N}}\ /\!*\text{ Cauchy-Schwarz: } \langle x|y\rangle \leq \|x\|\,\|y\|\ */$$

$$\leq \underbrace{\left(\sup_{n\in\mathbb{N}}\|v_n - \bar{v}\|\right)}_{\rho:\text{ finite, positive}}\|\nabla g(v_n) - \nabla g(\bar{v})\|$$

$$\therefore\ \forall_{n\in\mathbb{N}}\quad \|x_n - \bar{x}\|^2 \leq \rho\underbrace{\|\nabla g(v_n) - \nabla g(\bar{v})\|}_{\to 0\text{ as }n\to\infty}$$

$$\therefore\ \|x_n - \bar{x}\| \to 0 \text{ as } n\to\infty$$

$$\Leftrightarrow \boxed{x_n \to x}$$

■

Part 1

November 18, 2016 7:34 AM

Proposition 28.7.

[$(C_n)_{n\in\mathbb{N}}$: sequence of nonempty closed convex subsets of $\mathcal{H}$,

$\forall_{n\in\mathbb{N}}\ C_n \subseteq C_{n+1}$; $C = \overline{\bigcup_{n\in\mathbb{N}} C_n}$; $x\in\mathcal{H}$] $P_{C_n}x \to P_C x$

2016-11-18 9:28 AM

Proof: $\forall_{n\in\mathbb{N}}\ \underbrace{\bigcup_{n\in\mathbb{N}} C_n}_{\text{convex, nonempty}}$ /* using $\Rightarrow$ $\bigcup_{n\in\mathbb{N}} C_n$: convex, nonempty

/* Proposition 3.26 (i) : [C : convex, $\subseteq\mathcal{H}$] $\overline{C}$: convex */

$\therefore C = \overline{\bigcup_{n\in\mathbb{N}} C_n}$: convex, nonempty /* * A consequence of Theorem 3.32

[C : convex subset of $\mathcal{H}$] $x\in\overline{C} \Leftrightarrow \exists_{(x_n)_{n\in\mathbb{N}}\subseteq C}\ x_n \to x$ /* Relation : between n & C $x_n \to x$ */

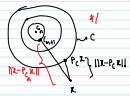
$P_C x \in C \Leftrightarrow \exists_{(\tilde{y}_n)_{n\in\mathbb{N}}\subseteq C}\ \tilde{y}_n \to P_C x$

now as $C = \overline{\bigcup_{n\in\mathbb{N}} C_n}$ we can cherrypick components of $(\tilde{y}_n)_{n\in\mathbb{N}}$ and create a subsequence

$(y_n)_{n\in\mathbb{N}}$: $y_n \in C_n$

$\therefore \exists_{(y_n)_{n\in\mathbb{N}}\subseteq C}\ (y_n \to P_C x,\ y_n \in C_n)$

$\forall_{n\in\mathbb{N}}\ \|x-P_C x\| \leq \|x-P_{C_n}x\|$ /* consider the picture:



$\leq \|x-y_n\|$ /* as by definition: $\|x-P_{C_n}x\| \leq \|x-y\|\ \ \forall_{y\in C_n}$ $\Rightarrow \|x-P_{C_n}x\| \leq \|x-y_n\|$ */

$= \|x-P_C x + P_C x - y_n\|$

$\leq \|x-P_C x\| + \|y_n - P_C x\|$

sub this from all sides

$\Rightarrow \forall_{n\in\mathbb{N}}\ \|x-P_C x\| \leq \|x-P_{C_n}x\| \leq \|x-P_C x\| + \|y_n - P_C x\|$

$\Rightarrow \forall_{n\in\mathbb{N}}\ 0 \leq \|x-P_{C_n}x\| - \|x-P_C x\| \leq \|y_n - P_C x\|$

as $n\to\infty$ $0 \leq \|x-P_{C_n}x\| - \|x-P_C x\| \leq \underbrace{\|y_n - P_C x\|}_{\to 0}$

$\Rightarrow \|x-P_{C_n}x\| - \|x-P_C x\| \to 0$

$\therefore \boxed{\|x-P_{C_n}x\| \to \|x-P_C x\|} \Rightarrow \forall_{(P_{C_{n_k}}x)_k} \|x-P_{C_{n_k}}x\| \to \|x-P_C x\|$

$\Rightarrow \boxed{(P_{C_n}x)_{n\in\mathbb{N}} : \text{bounded}}$

$\subseteq C$ /* $P_{C_n}x \in C_n \subseteq C_{n+1} \subseteq \ldots C$ */

/* now

[$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$] $\Rightarrow \exists_{(x_{n_k})_{k\in\mathbb{N}}}$ subsequence of $(x_n)_{n\in\mathbb{N}}$ $(x_{n_k})_{k\in\mathbb{N}}$: weakly convergent

$\Leftrightarrow$ $(x_{n_k})_{k\in\mathbb{N}}$ possesses a weak sequential cluster point */

/* Lemma 2.55 /* Norm of $\mathcal{H}$: weakly lower semicontinuous */ *

[$(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ $\forall_{x\in\mathcal{H}}$ $x_n \rightharpoonup x \Rightarrow \|x\| \leq \underline{\lim}\, \|x_n\|$]

$\underbrace{(P_{C_n}x)_{n\in\mathbb{N}}}_{\subseteq C}$: possesses a weak sequential cluster point z (say) $\Leftrightarrow \exists_{(P_{C_{n_k}}x)_{k\in\mathbb{N}}\subseteq C}\ P_{C_{n_k}}x \rightharpoonup z \Leftrightarrow \boxed{P_{C_{n_k}}x - x \rightharpoonup z - x} \Rightarrow \boxed{\|z-x\| \leq \underline{\lim}\, \|P_{C_{n_k}}x - x\|}$

now C : convex, closed

$\Rightarrow C$: sequentially closed

/* now

[C : convex subset of $\mathcal{H}$] $\Rightarrow$ C : sequentially closed $\Leftrightarrow$ C : weakly closed

$\boxed{z\in C}$

now : by def

$\forall_{\tilde{x}\in C}\ \|x-P_C x\| \leq \|x-\tilde{x}\|$

$\overset{\tilde{x}=z\in C}{\Rightarrow} \|x-P_C x\| \leq \|x-z\| \leq \underline{\lim}\, \|x-P_{C_{n_k}}x\| = \lim \|x-P_{C_{n_k}}x\| = \|x-P_C x\|$

$\Rightarrow \boxed{\|x-z\| = \|x-P_C x\|}$

now By def: $\|x-P_C x\| = \min_{\tilde{x}\in C} \|x-\tilde{x}\|$ and $P_C x$: unique as C : closed, convex

$\Rightarrow \boxed{P_C x = z}$ But: z : sequential cluster point of an arbitrary subsequence of $(P_{C_n}x)_{n\in\mathbb{N}}$

$\Rightarrow \boxed{z = P_C x : \text{only weak sequential cluster point of } (P_{C_n}x)_{n\in\mathbb{N}}}$

/* using

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n\in\mathbb{N}}$ converges weakly $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$ bounded, possesses at most one weak sequential cluster point. */

$\therefore P_{C_n}x \rightharpoonup z = P_C x$

$\Rightarrow \boxed{x - P_{C_n}x \rightharpoonup x - P_C x}$

/* using

(One characterization of strong convergence)

Corollary 2.58: [$(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$; $x\in\mathcal{H}$] $x_n \rightharpoonup x \Leftrightarrow (x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|)$ */

$\therefore x - P_{C_n}x \to x - P_C x$

$\Rightarrow \boxed{P_{C_n}x \to P_C x}$ ∎

Proposition 28.8.

[$(C_n)_{n\in\mathbb{N}}$: sequence of nonempty, closed, convex, subsets of $\mathcal{H}$;

$C = \bigcap_{n\in\mathbb{N}} C_n \neq \emptyset$; $\forall_{n\in\mathbb{N}}\ C_{n+1} \subseteq C_n$; $x\in\mathcal{H}$] $P_{C_n}x \to P_C x$

C_n seems very intuitive

2016-11-21 4:46 PM

Proof: 1 $\forall_{n\in\mathbb{N}}\ p_n = P_{C_n}x$

$\forall_{n\in\mathbb{N}}\ \|x-p_n\| \leq \|x-p_{m+1}\| \leq \|x-P_C x\|$

$\Rightarrow (p_n)_{n\in\mathbb{N}}$: bounded

/* Lemma 2.54

[$(x_n)_{n\in\mathbb{N}}$: monotonic] $(x_n)_{n\in\mathbb{N}}$: convergent $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: bounded */

now $(\|x-p_n\|)_{n\in\mathbb{N}}$: monotonic, bounded $\Rightarrow (\|x-p_n\|)_{n\in\mathbb{N}}$: converges

$\forall_{m,n\in\mathbb{N}} : m\leq n$ $\{p_m, p_n\} \subseteq C_m$

/* Assume $C_1 \subseteq C_m$ (Since) by def: $p_m \in C_m$, $\overline{p_n} \in C_n \subseteq C_m$ now C_m : convex $\Rightarrow \{p_m, \overline{p_n}\} \subseteq C_m$ */

now:

$\|p_n - p_m\|^2$

$= 2(\|p_n - x\|^2 + \|p_m - x\|^2) - 4\|x - \frac{p_n + p_m}{2}\|^2$

/* now: $p_m = P_{C_m}x$:

$\forall_{y\in C_m}\ \|x-p_m\| \leq \|x-y\|$ /* by def */

$y = \frac{p_n+p_m}{2} \in C_m$ $\Rightarrow \|x-p_m\| \leq \|x - \frac{p_n+p_m}{2}\|$

$\leftrightarrow \|x-p_m\|^2 \leq \|x - \frac{p_n+p_m}{2}\|^2$

$\leftrightarrow -4\|x-p_m\|^2 \geq -4\|x - \frac{p_n+p_m}{2}\|^2$ */

$\leq 2\|p_n - x\|^2 + 2\|p_m - x\|^2 - 4\|x - p_m\|^2$

$= 2(\|p_n - x\|^2 - \|p_m - x\|^2)$ /* now as $(\|p_n - x\|)_{n\in\mathbb{N}}$ converges $\Rightarrow$ as $n,m\to\infty$, $\|p_n - x\|^2 - \|p_m - x\|^2 \to 0$ */

$\therefore$ as $m,n\to+\infty$, $\|p_n - p_m\|^2 \to 0$

$\leftrightarrow (p_n)_{n\in\mathbb{N}}$: Cauchy sequence in $\mathcal{H}$ /* $\mathcal{H}$: complete metric space */

$\therefore \exists_{p\in\mathcal{H}}\ p_n \to p$

$\forall_{n\in\mathbb{N}}\ (p_k)_{k\geq n}$ lies in $C_n \Rightarrow p\in C_n$ /* C_n : closed, so the limit point of any convergent sequence will lie in C_n */

p /* as $(p_n)_{k\geq n}$: subsequence of $(p_n)_{n\in\mathbb{N}}$ */

now, as $C = \bigcap_{n\in\mathbb{N}} C_n \Rightarrow p\in C$

$\therefore$ By definition $\|x-P_C x\| \leq \|x-\tilde{y}\|\ \ \forall_{\tilde{y}\in C}$ $\overset{\tilde{y}=p}{\Rightarrow} \|x-P_C x\| \leq \|x-p\| = \lim \|x-p_n\| \leq \|x-P_C x\|$

$\Rightarrow \|x-P_C x\| = \|x-p\|$, But $P_C x$: unique minimizer of $\min_{y\in C} \|x-y\|$

$\therefore p = P_C x$

$\therefore P_{C_n}x \to p = P_C x$ ∎

Proposition 28.28. [$f : \mathcal{H}\to\mathbb{R}$; convex, continuous

$C = \operatorname{argmin} f$

$(z,\xi) \in (\mathcal{H}\times\mathbb{R})\setminus C$] $\Rightarrow$

- $z \in x + (\partial f(x) - \xi)\partial f(x)$: has unique solution (say $\bar{x}$)
- $P_C(z,\xi) = (\bar{x}, f(\bar{x}))$

2016-11-23 5:47 PM

Proof: $\underbrace{(z-x, \xi-f)}$ /* as, $\|(x,y)\|^2 = \|x\|^2 + \|y\|^2$ */

$(z,\zeta) \in (\mathcal{H}\times\mathbb{R})\setminus C$] $\Rightarrow$

- $z \in x + (f(x)-\zeta)\,\partial f(x)$: has unique solution (say $\bar{x}$)
- $P_C(z,\zeta) = (\bar{x}, f(\bar{x}))$

2016-11-23 5:47 PM

Proof:

$$(\bar{x},\bar{\zeta}) = P_C(z,\zeta) \overset{\text{def}}{\Leftrightarrow} (\bar{x},\bar{\zeta}) = \begin{cases} \operatorname{argmin} \frac{1}{2}\|\overbrace{(z,\zeta)-(x,\zeta)}^{(z-x,\,\zeta-\zeta)}\|^2 = \frac{1}{2}\|x-z\|^2 + \frac{1}{2}\|\zeta-\zeta\|^2 = \frac{1}{2}\|x-z\|^2 + \frac{1}{2}|\zeta-\zeta|^2 \\ \text{s.t.} \quad (x,\zeta)\in C = \operatorname{epi} f \Leftrightarrow f(x) \le \zeta \end{cases}$$

// as, $\|(x,y)\|^2 = \|x\|^2 + \|y\|^2$ //

$\rightsquigarrow$ Lagrangian: $L((x,\zeta),v) = \frac{1}{2}\|x-z\|^2 + \frac{1}{2}|\zeta-\zeta|^2 + v(f(x)-\zeta)$

if we apply KKT condition: then

primal feasibility: $f(\bar{x}) \le \bar{\zeta}$

dual feasibility: $\bar{v} \ge 0$

complementary slackness: $\bar{v}(f(\bar{x})-\bar{\zeta}) = 0$

minimizer of the Lagrangian: $\partial_{(\bar{x},\bar{\zeta})} L((\bar{x},\bar{\zeta}),\bar{v}) \ni 0$

$$\Leftrightarrow \begin{bmatrix} \frac{\partial}{\partial \bar{x}} \\ \frac{\partial}{\partial \bar{\zeta}} \end{bmatrix} \left(\frac{1}{2}\|\bar{x}-z\|^2 + \frac{1}{2}|\bar{\zeta}-\zeta|^2 + \bar{v}(f(\bar{x})-\bar{\zeta}) \right) = \begin{bmatrix} \bar{x}-z + \bar{v}\,\partial f(\bar{x}) \\ \bar{\zeta}-\zeta-\bar{v} \end{bmatrix} \ni \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Updownarrow$$

$$\boxed{\bar{x}-z+\bar{v}\,\partial f(\bar{x}) \ni 0, \quad \bar{\zeta}-\zeta-\bar{v}=0}$$

assume: $\bar{v}=0 \Rightarrow \bar{x}=z,\ \bar{\zeta}=\zeta \Leftrightarrow \underbrace{(\bar{x},\bar{\zeta})}_{\in C} = \underbrace{(z,\zeta)}_{\notin C} \Rightarrow$ contradiction $\Rightarrow \bar{v} \ne 0 \Rightarrow \bar{v} > 0 \Rightarrow$ then using complementary slackness: $f(\bar{x})-\bar{\zeta}=0 \Leftrightarrow \boxed{\bar{\zeta}=f(\bar{x})}$ $\therefore \bar{v} = f(\bar{x})-\zeta$

$$z-\bar{x} \in \bar{v}\,\partial f(\bar{x}) = (f(\bar{x})-\zeta)\,\partial f(\bar{x})$$

$$\Leftrightarrow \boxed{z \in \bar{x} + (f(\bar{x})-\zeta)\,\partial f(\bar{x})}$$ $\rightsquigarrow$ now $\bar{x}$: unique as the Lagrangian is strongly convex in x

$\therefore P_C(z,\zeta) = (\bar{x}, f(\bar{x}))$: ($\bar{x}$: solves $z \in x + (f(x)-\zeta)\,\partial f(x)$ uniquely)

■

2016-11-24 1:17 PM

Proposition 28.30. [$f : \mathcal{H} \to \mathbb{R}$, convex, continuous ; $z \in \mathcal{H}$; $\zeta \in \inf f(\mathcal{H})$, $f(z) < \zeta = \operatorname{lev}_{\le f} f$] $\Rightarrow$

- $f(\operatorname{prox}_{vf} z) = \zeta$: has atleast one solution in $\mathbb{R}_{++}$ say $\bar{v}$
  ↳ variable
- $P_C z = \operatorname{prox}_{\bar{v}f} z$

Proof: Take $z \notin C$, as otherwise it is just $P_C z = z$, trivial. Say $\bar{x} = P_C z$

By definition, $P_C z = \begin{cases} \operatorname{argmin} \frac{1}{2}\|x-z\|^2 \\ x \in \mathcal{H} \\ \text{s.t.} \quad x \in C = \operatorname{lev}_{\le f} f \Leftrightarrow f(x) \le \zeta \end{cases}$ $\rightsquigarrow$ Lagrangian:

$L(x,v) = \frac{1}{2}\|x-z\|^2 + v(f(x)-\zeta)$ // convex and proper

KKT conditions: for primal-dual pair $(\bar{x},\bar{v}) = (P_C z, \bar{v})$

Primal feasibility:	$f(\bar{x}) \le \zeta$
Dual feasibility:	$\bar{v} \ge 0$
Complementary slackness:	$\bar{v}(f(\bar{x})-\zeta) = 0$
Vanishing gradient of the Lagrangian:	$\nabla_x L(\bar{x},\bar{v}) = 0$

$$\Updownarrow$$

$$(\bar{x}-z) + \bar{v}\,\partial f(\bar{x}) \ni 0$$ // there is atleast one solution

suppose $\bar{v}=0 \Rightarrow \underbrace{\bar{x}}_{\in C} = \underbrace{z}_{\notin C} \Rightarrow$ contradiction $\therefore \bar{v} > 0 \Rightarrow f(\bar{x}) = f(P_C z) = \zeta$

$$\bar{x} - z + \bar{v}\,\partial f(\bar{x}) \ni 0$$

$$\Leftrightarrow \bar{v}\,\partial f(\bar{x}) \ni z - \bar{x} \Leftrightarrow \partial(\bar{v} f(\bar{x})) \ni z - \bar{x} \quad // \ x - p \in \partial f(p) \Leftrightarrow p = \operatorname{prox}_f x$$

$$\Leftrightarrow \boxed{\bar{x} = \operatorname{prox}_{\bar{v}f} z = P_C z}$$ //the primal variable has been written in terms of the dual
// variable

$f(\bar{x}) = f(P_C z) = f(\operatorname{prox}_{\bar{v}f} z) = \zeta$ $\therefore f(\operatorname{prox}_{vf} z) = \zeta$. must have atleast one solution in $\mathbb{R}_{++}$; and if $\bar{v}$ is such a solution then: $\boxed{P_C z = \operatorname{prox}_{\bar{v}f} z}$ ■

↳ variable

Part 1

November 25, 2016 1:06 PM

• Halpern's Q operator:

$Q: \mathcal{H}^4 \to \mathcal{H}$ /* $X = (x-y|z-x)$; $\mu = \|x-y\|^2$; $V = \|z-x\|^2$; $\rho = \|x-v\|^2$ */

$$(x,y,z) \mapsto \begin{cases} x, & \text{if } \rho = 0,\ \mu \geq 0 \\ x + \left(1 + \frac{\mu}{V}\right)(x-y), & \text{if } \rho > 0,\ \mu \geq \rho \\ y + \frac{V}{\rho}\left(X(x-y) + \mu(z-y)\right), & \text{if } \rho > 0,\ \mu < \rho \end{cases}$$

(Strongly convergent proximal point algorithm that finds the zero of a maximally monotone operator at minimal distance from the starting point)

[$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone, $0 \in \operatorname{ran} A$ /i.e. $\operatorname{zer} A \neq \emptyset$ */;

$x_0 \in \mathcal{H}$; $\forall_{n \in \mathbb{N}}\ x_{n+1} = Q(x_0, x_n, J_A x_n)$] $x_n \to \underbrace{P_{\operatorname{zer} A} x_0}_{\text{this is the zero of } A \text{ that is closest to } x_0}$

2016-11-25 1:25 PM

Proof:

set $T_1 = J_A$ // A: maximally monotone $\Rightarrow J_A$: firmly nonexpansive (corollary 23.8)

T_1: firmly nonexpansive

$\Rightarrow \operatorname{zer} A = \operatorname{fix} T_1 = \operatorname{fix} J_A \neq \emptyset$

now: from Halpern's algorithm:

(Halpern's algorithm to find the best approximation to a point from the set of common fixed points of firmly nonexpansive operator) set $T = \{T_1\}$, $i(n) = 1$,

Theorem 23.3. [$(T_i)_{i \in I}$: finite family of firmly nonexpansive operators from $\mathcal{H}$ to $\mathcal{H}$,

$C = \bigcap_{i \in I} \operatorname{Fix} T_i \neq \emptyset$; $x_0 \in \mathcal{H}$; $i: \mathbb{N} \to I$: $\exists_{m \in \mathbb{N}_{++}}\ \forall_{i \in I}\ \forall_{n \in \mathbb{N}}\ i \in \{i(n), \ldots, i(n+m)\}$; satisfied

$\forall_{n \in \mathbb{N}}\ x_{n+1} = Q(x_0, x_n, \underbrace{T_{i(n)} x_n}_{T_1 x_n = J_A x_n})$] $x_n \to P_C x_0 = P_{\operatorname{zer} A} x_0$

$\therefore x_n \to P_{\operatorname{zer} A} x_0$

■

(Strongly convergent forward-backward algorithm) /*of course can be applied to solving VI problem */

Corollary 29.5. [$A: \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone; $\beta \in \mathbb{R}_{++}$; $B: \mathcal{H} \to \mathcal{H}$, β-cocoercive; $\gamma \in]0, 2\beta[$;

$\operatorname{zer}(A+B) \neq \emptyset$; $x_0 \in \mathcal{H}$; Set:

$$\forall_{n \in \mathbb{N}} \begin{cases} y_n = x_n - \gamma B x_n \\ z_n = \frac{1}{2}(x_n + J_{\gamma A} y_n) \\ x_{n+1} = Q(x_0, x_n, z_n) \end{cases} \quad (29.41)$$

] $\Rightarrow x_n \to \underbrace{P_{\operatorname{zer}(A+B)} x_0}_{\text{zero of } (A+B) \text{ that is closest to } x_0}$

2016-11-25 2:25 PM

Proof: From (Eq. 29.41): $y_n = x_n - \gamma B x_n$

$z_n = \frac{1}{2}(x_n + J_{\gamma A} y_n) = \frac{1}{2}(x_n + \underbrace{J_{\gamma A}(x_n - \gamma B x_n)}_{(J_{\gamma A} \circ (\mathrm{Id} - \gamma B))(x_n)}) = \frac{1}{2}(x_n + (J_{\gamma A} \circ (\mathrm{Id} - \gamma B)) x_n) = \underbrace{\frac{1}{2}(\mathrm{Id} + J_{\gamma A} \circ (\mathrm{Id} - \gamma B))}_{T_1 \text{ /\!/say}} x_n = T_1 x_n$

$x_{n+1} = Q(x_0, x_n, z_n) = Q(x_0, x_n, T_1 x_n)$

now T_1: firmly nonexpansive

/* A: maximally monotone $\Rightarrow J_{\gamma A}$: firmly nonexpansive $\Rightarrow J_{\gamma A}$: nonexpansive

B: β-cocoercive $\Rightarrow (\mathrm{Id} - \gamma B)$: nonexpansive /* a.s

Proposition 4.33.
[D: nonempty subset of $\mathcal{H}$
$\beta \in \mathbb{R}_{++}$
$T: D \to \mathcal{H}$, β-cocoercive
$\gamma \in]0, 2\beta[$]

$(\mathrm{Id} - \gamma T)$: $\frac{\gamma}{2\beta}$ averaged and averaged operators are nonexpansive */

$\therefore J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$: nonexpansive

now set

$2\tilde{T} - \mathrm{Id} = J_{\gamma A} \circ (\mathrm{Id} - \gamma B)$: nonexpansive

$\Leftrightarrow T_1 = \tilde{T} = \frac{1}{2}(\mathrm{Id} + J_{\gamma A} \circ (\mathrm{Id} - \gamma B))$

: firmly nonexpansive

*/

Proposition 4.4. (Averaged representations of firmly nonexpansive operator)
[D: nonempty subset of $\mathcal{H}$,
$T: D \to \mathcal{H}$]
(i) T: firmly nonexpansive $\Leftrightarrow$ (ii) $(2T - \mathrm{Id})$: firmly nonexpansive $\Leftrightarrow$ (iii) $(2T - \mathrm{Id})$: nonexpansive
$\Leftrightarrow$ (iv) $\forall_{x \in D} \forall_{y \in D}\ \|Tx - Ty\|^2 \leq (x-y|Tx-Ty)$
$\Leftrightarrow$ (v) $\forall_{x \in D} \forall_{y \in D}\ (Tx - Ty|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y) \geq 0$
$\Leftrightarrow$ (vi) $\forall_{x \in D} \forall_{y \in D} \forall_{\alpha \in [0,1]}\ \|Tx - Ty\|^2 \leq \|\alpha(x-y) + (1-\alpha)(Tx - Ty)\|$

Proposition 25.1. [$A: \mathcal{H} \to 2^{\mathcal{H}}$, $B: \mathcal{H} \to 2^{\mathcal{H}}$, $\gamma \in \mathbb{R}_{++}$] $\Rightarrow$

(i) $\operatorname{zer}(A+B) = \operatorname{dom}(A \cap (-B))$

(ii) A, B: monotone $\Rightarrow \operatorname{zer}(A+B) = J_{\gamma B}(\operatorname{Fix} R_{\gamma A} R_{\gamma B})$ // note: very general, can it be used in nonconvex setup?

(iii) (C: closed affine subspace of $\mathcal{H}$; $V := C - C$; $A = N_C$) $\Rightarrow$
$\operatorname{zer}(A+B) = \{x \in C \mid V^\perp \cap Bx \neq \emptyset\}$

(iv) (A: monotone, B: almost single-valued) $\Rightarrow \operatorname{zer}(A+B) = \operatorname{Fix} \underbrace{J_{\gamma A} \circ (\mathrm{Id} - \gamma B)}_{2T_1 - \mathrm{Id} \text{ // in our case}}$

B: β-cocoercive

$\therefore \operatorname{zer}(A+B) = \operatorname{fix}(2T_1 - \mathrm{Id}) = \operatorname{fix} T_1$

/* $x \in \operatorname{Fix} T \Leftrightarrow Tx \ni x$
$\Leftrightarrow 2Tx \ni 2x$
$\Leftrightarrow 2Tx - \mathrm{Id}x \ni 2x - x = x$
$\Leftrightarrow (2T - \mathrm{Id})x \ni x \Leftrightarrow x \in \operatorname{fix}(2T - \mathrm{Id})$ */

(Halpern's algorithm to find the best approximation to a point from the set of common fixed points of firmly nonexpansive operator) set: $T = \{T_1\}$, $i: n \mapsto 1$,

Theorem 23.3. [$(T_i)_{i \in I}$: finite family of firmly nonexpansive operators from $\mathcal{H}$ to $\mathcal{H}$,

$C = \bigcap_{i \in I} \operatorname{Fix} T_i \neq \emptyset$; $x_0 \in \mathcal{H}$; $i: \mathbb{N} \to I$: $\exists_{m \in \mathbb{N}_{++}}\ \forall_{i \in I}\ \forall_{n \in \mathbb{N}}\ i \in \{i(n), \ldots, i(n+m)\}$; ✓

$= \operatorname{Fix} T = \operatorname{zer}(A+B)$

$\forall_{n \in \mathbb{N}}\ x_{n+1} = Q(x_0, x_n, \underbrace{T_{i(n)}}_{T_1} x_n)$] $x_n \to P_C x_0$

$x_n \to P_C x_0 = P_{\operatorname{zer}(A+B)} x_0$

■