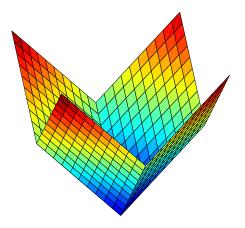
Data processing and networks optimization Part III: Subdifferential and conjugate

Pierre Borgnat¹, Jean-Christophe Pesquet², Nelly Pustelnik¹

¹ ENS Lyon – Laboratoire de Physique – CNRS UMR 5672 pierre.borgnat@ens-lyon.fr, nelly.pustelnik@ens-lyon.fr

> ² LIGM – Univ. Paris-Est – CNRS UMR 8049 jean-christophe.pesquet@univ-paris-est.fr

Non-smooth convex optimization



2/19

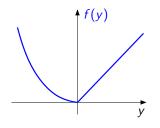
A pioneer

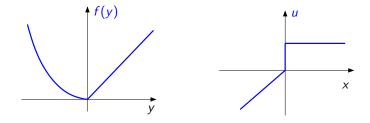


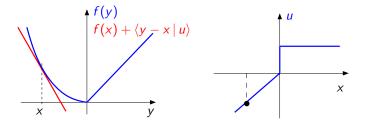
Jean-Jacques Moreau (1923–2014)

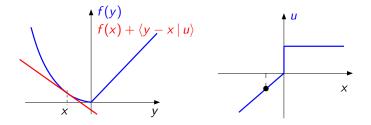
The (Moreau) subdifferential of f, denoted by ∂f ,

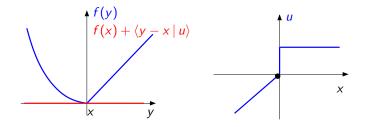
Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function. The (Moreau) subdifferential of f, denoted by ∂f ,

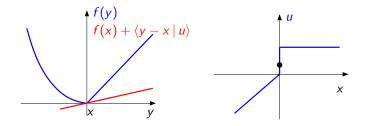


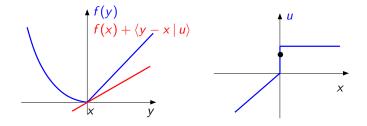


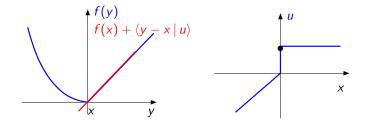


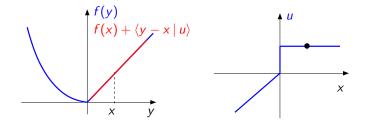




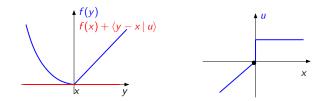








Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function. The (Moreau) subdifferential of f, denoted by ∂f , is such that $\partial f : \mathcal{H} \to 2^{\mathcal{H}}$ $x \to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x | u \rangle + f(x) \leq f(y)\}$



Fermat rule : $0 \in \partial f(x) \Leftrightarrow x \in \operatorname{Argmin} f$

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function. The (Moreau) subdifferential of f, denoted ∂f , is such that $\partial f : \mathcal{H} \to 2^{\mathcal{H}}$ $x \to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x | u \rangle + f(x) \leq f(y)\}$

• $u \in \partial f(x)$ is a subgradient of f at x.

- $u \in \partial f(x)$ is a subgradient of f at x.
- If $x \notin \operatorname{dom} f$, then $\partial f(x) = \emptyset$.

- $u \in \partial f(x)$ is a subgradient of f at x.
- If $x \notin \operatorname{dom} f$, then $\partial f(x) = \emptyset$.
- For every $x \in \text{dom } f$, $\partial f(x)$ is a closed and convex set.

If $f: \mathcal{H} \to]-\infty, +\infty]$ is convex and it is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$

If $f: \mathcal{H} \to]-\infty, +\infty]$ is convex and it is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$

$$(\forall y \in \mathcal{H}) \qquad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x + \alpha y) - f(x)}{\alpha}.$$

Proof:

For every $\alpha \in [0,1]$ and $y \in \mathcal{H}$,

$$f(x + \alpha(y - x)) \le (1 - \alpha)f(x) + \alpha f(y)$$

$$\Rightarrow \quad \langle \nabla f(x) \mid y - x \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x + \alpha(y - x)) - f(x)}{\alpha} \le f(y) - f(x)$$

Then $\nabla f(x) \in \partial f(x)$.

If $f: \mathcal{H} \to]-\infty, +\infty]$ is convex and it is Gâteaux differentiable at x, then $\partial f(x) = \{ \nabla f(x) \}$

$$(\forall y \in \mathcal{H}) \qquad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x + \alpha y) - f(x)}{\alpha}.$$

Proof:

Inversely, if $u \in \partial f(x)$, then, for every $\alpha \in [0, +\infty[$ and $y \in \mathcal{H}$,

$$f(x + \alpha y) \ge f(x) + \langle u \mid x + \alpha y - x \rangle$$

$$\Rightarrow \quad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x + \alpha y) - f(x)}{\alpha} \ge \langle u \mid y \rangle$$

By selecting $y = u - \nabla f(x)$, it results that $||u - \nabla f(x)||^2 \le 0$ and then $u = \nabla f(x)$.

Let $f: \mathcal{H} \to]-\infty, +\infty]$ be Gâteaux differentiable on dom f. Then, f is convex if and only if

 $(\forall (x,y) \in (\operatorname{dom} f)^2) \quad f(y) \ge f(x) + \langle \nabla f(x) \mid y - x \rangle.$

Let $f: \mathcal{H} \to]-\infty, +\infty]$ be Gâteaux differentiable on dom f. Then, f is convex if and only if

 $(\forall (x,y) \in (\operatorname{dom} f)^2) \quad f(y) \geq f(x) + \langle \nabla f(x) \mid y - x \rangle.$

Proof:

We have already seen that the gradient inequality holds when f is convex and differentiable at $x \in \mathcal{H}$.

Let $f: \mathcal{H} \to]-\infty, +\infty]$ be Gâteaux differentiable on dom f. Then, f is convex if and only if

 $(\forall (x,y) \in (\operatorname{dom} f)^2) \quad f(y) \geq f(x) + \langle \nabla f(x) \mid y - x \rangle.$

Proof:

Conversely, if the gradient inequality is satisfied, we have, for every $(x, y) \in (\text{dom } f)^2$ and $\alpha \in [0, 1]$, $\alpha x + (1 - \alpha)y \in \text{dom } f$, and

$$\begin{aligned} f(x) &\geq f(\alpha x + (1 - \alpha)y) + (1 - \alpha) \left\langle \nabla f(\alpha x + (1 - \alpha)y) \mid x - y \right\rangle \\ f(y) &\geq f(\alpha x + (1 - \alpha)y) + \alpha \left\langle \nabla f(\alpha x + (1 - \alpha)y) \mid y - x \right\rangle. \end{aligned}$$

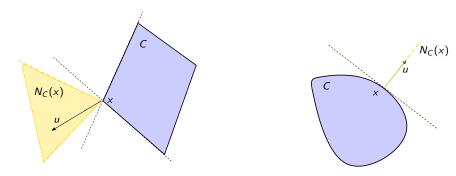
By multiplying the first inequality by α and the second one by $1-\alpha$ and summing them, we get

$$\alpha f(x) + (1-\alpha)f(y) \ge f(\alpha x + (1-\alpha)y).$$

Subdifferential of a convex function: example

For every
$$x \in \mathcal{H}$$
, $\partial \iota_C(x)$ is the normal cone to C at x defined by

$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \ \langle u \mid y - x \rangle \leq 0 \} & \text{if } x \in C \\ \varnothing & \text{otherwise.} \end{cases}$$



Subdifferential of a convex function: example

For every
$$x \in \mathcal{H}$$
, $\partial \iota_C(x)$ is the normal cone to C at x defined by
$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \ \langle u \mid y - x \rangle \leq 0 \} & \text{if } x \in C \\ \varnothing & \text{otherwise.} \end{cases}$$

Si
$$x \in \text{int } C$$
 alors $N_C(x) = \{0\}$.

- ▶ If C is a vector space, then for every $x \in C$, $N_C(x) = C^{\perp}$.
- ▶ Let $c \in \mathcal{H}$, $\rho \in]0, +\infty[$ and $C = \overline{\mathcal{B}}(c, \rho) = \{y \in \mathcal{H} \mid ||y c|| \le \rho\}.$ For every $x \in C$,

$$N_C(x) = \begin{cases} \{\alpha(x-c) \mid \alpha \in [0,+\infty[\} & \text{if } \|x-c\| = \rho \\ \{0\} & \text{si } \|x-c\| < \rho. \end{cases}$$

Subdifferential calculus

Let \mathcal{H} and \mathcal{G} be real Hilbert spaces.

- ▶ Let $f: \mathcal{H} \to]-\infty, +\infty]$ be proper, then for every $\lambda \in]0, +\infty[$ $\partial(\lambda f) = \lambda \partial f$.
- ▶ Let $f: \mathcal{H} \to]-\infty, +\infty]$ and $g: \mathcal{G} \to]-\infty, +\infty]$ be two proper functions, and let $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$. Then, $\partial f + L^* \partial gL \subset \partial (f + g \circ L)$. In addition, if $f \in \Gamma_0(\mathcal{H})$, $f \in \Gamma_0(\mathcal{G})$, and

 $\operatorname{dom} g \cap \operatorname{int} (L(\operatorname{dom} f)) \neq \varnothing \text{ or } \operatorname{int} (\operatorname{dom} g) \cap L(\operatorname{dom} f) \neq \varnothing,$

then $\partial f + L^* \partial gL = \partial (f + g \circ L)$. <u>Particular case</u>: If $f \in \Gamma_0(\mathcal{H}), f \in \Gamma_0(\mathcal{G}), L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ and f is finite valued, then $\partial f + \partial g = \partial (f + g)$.

Exercise 1: Huber function

Let $\rho > 0$ and set

$$f: \mathbb{R} \to \mathbb{R}: \mapsto \begin{cases} rac{x^2}{2}, & \text{if } |x| \le \rho \\ \rho |x| - rac{
ho^2}{2}, & \text{otherwise} \end{cases}$$

- 1. What is the domain of f?
- 2. Is f differentiable ? twice-differentiable ?
- 3. Prove that f is convex.

Exercise 2

Let \mathcal{H} be a Hilbert space. Let $f: \mathcal{H} \to]-\infty, +\infty]$ and let $C \subset \mathcal{H}$ such that dom $f \cap C \neq \emptyset$.

- 1. Give a sufficient condition for $x \in \mathcal{H}$ to be a global minimizer of $f + \iota_C$.
- When f ∈ Γ₀(H), C is a nonempty closed convex set, and f is Gâteaux-differentiable on H, give a necessary and sufficient condition for x ∈ H to be a minimizer of f + ι_C.

Conjugate



Adrien-Marie Legendre (1752–1833)



Werner Fenchel (1905–1988)

Conjugate



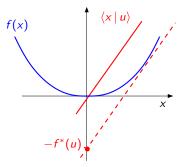
Adrien-Marie Legendre (1752–1833)



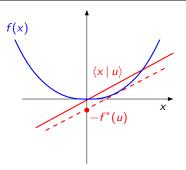
Werner Fenchel (1905–1988)

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \mathcal{H}} (\langle x \mid u \rangle - f(x))$

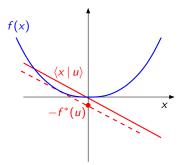
$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$



$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$



$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$



Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

Examples :

$$\blacktriangleright f = \frac{1}{2} \| \cdot \|^2 \Rightarrow f^* = \frac{1}{2} \| \cdot \|^2$$

<u>Proof</u>: For every $(x, u) \in \mathcal{H}^2$, $\langle x \mid u \rangle - \frac{1}{2} ||x||^2 = \frac{1}{2} ||u||^2 - \frac{1}{2} ||u - x||^2$ is maximum in x = u. Consequently, $f^*(u) = \frac{1}{2} ||u||^2$.

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

Examples :

 $\blacktriangleright f = \frac{1}{2} \| \cdot \|^2 \Rightarrow f^* = \frac{1}{2} \| \cdot \|^2 \ .$

▶ Let ϕ : $\mathbb{R} \to]-\infty, +\infty]$ be a even function. $(\phi \circ \| \cdot \|)^* = \phi^* \circ \| \cdot \|$.

$$(\forall x \in \mathbb{R}^N) f(x) = \frac{1}{q} ||x||_q^q \text{ with } q \in]1, +\infty[\Rightarrow (\forall u \in \mathbb{R}^N) f^*(u) = \frac{1}{q^*} ||u||_{q^*}^q \text{ with } \frac{1}{q} + \frac{1}{q^*} = 1$$

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

• If f is even then f^* is even.

$$\frac{\text{Proof}}{\text{Proof}}: \qquad (\forall u \in \mathcal{H}) \quad f^*(-u) = \sup_{x \in \mathcal{H}} \left(\langle x \mid -u \rangle - f(x) \right)$$
$$= \sup_{x \in \mathcal{H}} \left(\langle x \mid u \rangle - f(-x) \right)$$
$$= \sup_{x \in \mathcal{H}} \left(\langle x \mid u \rangle - f(x) \right)$$
$$= f^*(u)$$

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

- ▶ If *f* is even then *f*^{*} is even.
- For every $\alpha \in]0, +\infty[$, $(\alpha f)^* = \alpha f^*(\cdot/\alpha)$.
- ► For every $(y, v) \in \mathcal{H}^2$ et $\alpha \in \mathbb{R}$, $(f(\cdot - y) + \langle \cdot | v \rangle + \alpha)^* = f^*(\cdot - v) + \langle y | \cdot - v \rangle - \alpha$.
- Let G be a Hilbert space and L ∈ B(G, H) be an isomorphism. (f ∘ L)* = f* ∘ (L⁻¹)*.
- ► f* is l.s.c. and convex.

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

Moreau-Fenchel theorem

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$ be a proper function.

f is l.s.c. and convex $\Leftrightarrow f^{**} = f$.

14/19

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$. The conjugate of f is the function $f^*: \mathcal{H} \to [-\infty, +\infty]$ such that

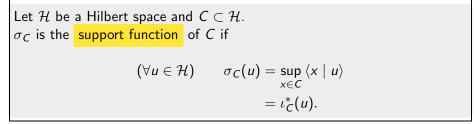
$$(\forall u \in \mathcal{H})$$
 $f^*(u) = \sup_{x \in \text{dom } f} (\langle x \mid u \rangle - f(x))$

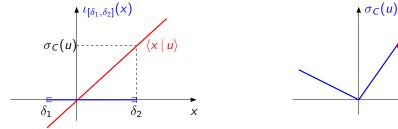
Moreau-Fenchel theorem

Let \mathcal{H} be a Hilbert space and $f: \mathcal{H} \to]-\infty, +\infty]$ be a proper function.

f is l.s.c. and convex $\Leftrightarrow f^{**} = f$.

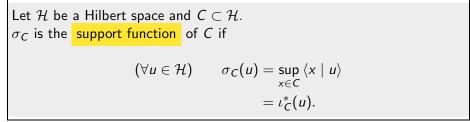
► Consquence: Si $f \in \Gamma_0(\mathbb{R})$ alors f^* est propre et donc $f^* \in \Gamma_0(\mathbb{R})$.

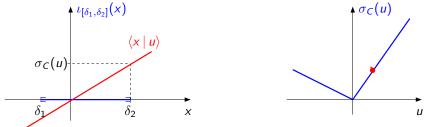


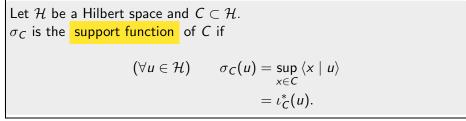


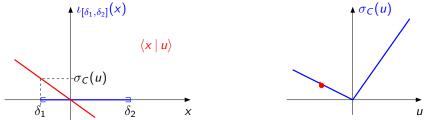
15/19

11









Let \mathcal{H} be a Hilbert space. $f: \mathcal{H} \to]-\infty, +\infty]$ is positively homogeneous if $(\forall x \in \mathcal{H})(\forall \alpha \in]0, +\infty[) \qquad f(\alpha x) = \alpha f(x).$

Let \mathcal{H} be a Hilbert space. $f: \mathcal{H} \to]-\infty, +\infty]$ is positively homogeneous if

 $(\forall x \in \mathcal{H})(\forall \alpha \in]0, +\infty[)$ $f(\alpha x) = \alpha f(x).$

Proof: (
$$\Leftarrow$$
)
 $f = \iota_{\mathcal{C}}^*$ and $\iota_{\mathcal{C}} \in \Gamma_0(\mathcal{H})$. Consequently $\sigma_{\mathcal{C}} \in \Gamma_0(\mathcal{H})$.
Moreover, $(\forall x \in \mathcal{H}) \ (\forall \alpha \in]0, +\infty[) \ \sigma_{\mathcal{C}}(\alpha x) = \alpha \sigma_{\mathcal{C}}(x)$.

Let \mathcal{H} be a Hilbert space. $f: \mathcal{H} \to]-\infty, +\infty]$ is positively homogeneous if

 $(\forall x \in \mathcal{H})(\forall \alpha \in]0, +\infty[)$ $f(\alpha x) = \alpha f(x).$

$$\begin{array}{l} \underline{\operatorname{Proof}}: (\Rightarrow)\\ \operatorname{Soit} y \in \operatorname{dom} f.\\ f(0) = \lim_{\alpha \to 0, \alpha \geq 0} f\left((1-\alpha)0 + \alpha y\right) = \lim_{\alpha \to 0} \alpha f(y) = 0.\\ \operatorname{Let} C = \left\{ u \in \mathcal{H} \mid (\forall x \in \mathcal{H}) \langle x \mid u \rangle \leq f(x) \right\}.\\ \operatorname{We have, for every} u \in C,\\ f^*(u) = \sup_{x \in \mathcal{H}} \langle x \mid u \rangle - f(x) \leq 0 = \langle 0 \mid u \rangle - f(0) \leq f^*(u).\\ \operatorname{Consequently,} f^*(u) = 0. \end{array}$$

Let \mathcal{H} be a Hilbert space. $f: \mathcal{H} \to]-\infty, +\infty]$ is positively homogeneous if $(\forall x \in \mathcal{H})(\forall \alpha \in]0, +\infty[) \qquad f(\alpha x) = \alpha f(x).$

<u>Proof</u>: (\Rightarrow) Moreover, for every $u \notin C$, there exists $x \in \mathcal{H}$ such that $\langle x \mid u \rangle > f(x)$. We have then, for every $\alpha \in]0, +\infty[$, $f^*(u) \ge \langle \alpha x \mid u \rangle - f(\alpha x) = \alpha (\langle x \mid u \rangle - f(x))$. By taking α to $+\infty$, we obtain $f^*(u) = +\infty$. To conclude, $f^* = \iota_C \in \Gamma_0(\mathcal{H}) \Rightarrow f = \sigma_C$ and C is a nonempty closed convex set.

► Let
$$f : \mathbb{R} \to]-\infty, +\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0\\ 0 & \text{if } x = 0\\ \delta_2 x & \text{if } x > 0 \end{cases}$$

with $-\infty \le \delta_1 < \delta_2 \le +\infty$.

Then, $f = \sigma_C$ where C is the closed real interval such that $inf C = \delta_1$ et sup $C = \delta_2$.

► Let
$$f : \mathbb{R} \to]-\infty, +\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0\\ 0 & \text{if } x = 0\\ \delta_2 x & \text{if } x > 0 \end{cases}$$

with $-\infty \le \delta_1 < \delta_2 \le +\infty$.

Then, $f = \sigma_C$ where C is the closed real interval such that $\inf C = \delta_1$ et sup $C = \delta_2$.

► Let f be a ℓ^q norm of \mathbb{R}^N with $q \in [1, +\infty]$. We have $f = \sigma_C$ where

$$\mathcal{C} = \left\{ y \in \mathbb{R}^{\mathsf{N}} \mid \|y\|_{q^*} \leq 1
ight\} \qquad ext{with } rac{1}{q} + rac{1}{q^*} = 1.$$

► Let
$$f : \mathbb{R} \to]-\infty, +\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases}$$

with $-\infty \le \delta_1 < \delta_2 \le +\infty$.

Then, $f = \sigma_C$ where C is the closed real interval such that $\inf C = \delta_1$ et sup $C = \delta_2$.

► Let f be a ℓ^q norm of \mathbb{R}^N with $q \in [1, +\infty]$. We have $f = \sigma_C$ where

$$\mathcal{C} = \left\{ y \in \mathbb{R}^{\mathcal{N}} \mid \|y\|_{q^*} \leq 1
ight\} \qquad ext{with } rac{1}{q} + rac{1}{q^*} = 1.$$

<u>Particular case</u>: ℓ^1 norm of \mathbb{R}^N : $C = [-1, 1]^N$.

Conjugate: properties

Fenchel-Young inequality : If f is proper, then 1. $(\forall (x, u) \in \mathcal{H}^2)$ $f(x) + f^*(u) \ge \langle x \mid u \rangle$ 2. $(\forall (x, u) \in \mathcal{H}^2)$ $u \in \partial f(x) \Leftrightarrow f(x) + f^*(u) = \langle x \mid u \rangle.$ If $f \in \Gamma_0(\mathcal{H})$, then $(\forall (x, u) \in \mathcal{H}^2)$ $u \in \partial f(x) \Leftrightarrow x \in \partial f^*(u).$

Exercise

▶ For every $c \in \mathbb{R}^N$, compute the conjugate of

$$f: \mathbb{R}^N \to]-\infty, +\infty]$$
$$x \to c^\top x + \iota_{[0, +\infty[}^N(x))$$

Same question for

$$g: \mathbb{R}^N \to]-\infty, +\infty]$$

 $x \to \iota_{[0,+\infty[}N(x-c))$