Optimization
Reminder and exercises

Barbara Pascal, **Nelly Pustelnik**

CNRS, Laboratoire de Physique de l'ENS de Lyon, Univ. Lyon 1
nelly.pustelnik@ens-lyon.fr

barbara.pascal@ens-lyon.fr

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x\in\mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,
- lower semi-continuous: closed *epigraph*

$$\operatorname{epi} f = \{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\},$$

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,
- lower semi-continuous: closed *epigraph*

$$\operatorname{epi} f = \{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\},$$

- convex: convex epigraph.

Reminder of the context

Questions:

- Existence and uniqueness of $\widehat{x}$
  $\longrightarrow$ coercivity or compactness *(existence)*
  $\longrightarrow$ strict convexity *(uniqueness)*

- Characterization of $\widehat{x}$

  $\longrightarrow$ $\nabla f(\widehat{x}) = 0$ if f Gâteaux-differentiable
  $\longrightarrow$ $0 \in \partial f(\widehat{x})$ is f non-smooth

$$\partial f : \begin{cases} \mathcal{H} & \rightarrow \quad 2^{\mathcal{H}} \\ x & \mapsto \quad \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}), \quad \langle y - x \mid u \rangle + f(x) \leq f(y)\} \end{cases}$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_{n+1})$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$x_{n+1} + \gamma \partial f(x_{n+1}) = x_n$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Proximal operator

$$\mathrm{prox}_{\gamma f}(x) = \operatorname*{arg\,min}_{y \in \mathcal{H}} \frac{1}{2}\|y - x\|^2 + \gamma f(y) = (\mathrm{Id} + \gamma \partial f)^{-1}(x)$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Proximal operator

$$\mathrm{prox}_{\gamma f}(x) = \operatorname*{arg\,min}_{y \in \mathcal{H}} \frac{1}{2}\|y - x\|^2 + \gamma f(y) = (\mathrm{Id} + \gamma \partial f)^{-1}(x)$$

Another way to say that $p = \mathrm{prox}_f(x) \quad \Leftrightarrow \quad x - p \in \partial f(p)$.

Exercise 1: Huber function

Let $\rho > 0$ and set

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

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

1. $(\forall x \in \mathbb{R}) \quad f(x) < \infty$, thus $\operatorname{dom} f = \mathbb{R}$.
2. f is differentiable on $\mathbb{R} \setminus \{\pm\rho\}$. Further

$$\lim_{x \to \rho^-} \underbrace{f'(x)}_{=x} = \rho = \lim_{x \to \rho^+} \underbrace{f'(x)}_{=\rho}$$

Thus f is differentiable at $x = \rho$ and by symmetry, f is also differentiable at $-\rho$. Finally f is differentiable on $\mathbb{R}$.

Exercise 1: Huber function

Let $\rho > 0$ and set

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

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

2. f is twice-differentiable on $\mathbb{R} \setminus \{\pm\rho\}$ and

$$f''(x) = \begin{cases} 1 & \text{if } |x| < \rho \\ 0 & \text{if } |x| > \rho \end{cases}$$

thus it is not twice-differentiable.

Exercise 1: Huber function

Let $\rho > 0$ and set

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

1. What is the domain of f ?
2. Is f differentiable ? twice-differentiable ?
3. Prove that f is convex.
3. f is differentiable on $\mathbb{R}$ and

$$f'(x) = \begin{cases} -\rho & \text{if } x < -\rho \\ x & \text{if } |x| < \rho \\ \rho & \text{otherwise} \end{cases}$$

which is increasing. Thus f is convex.

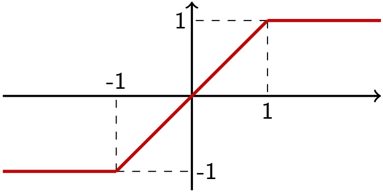
Exercise 1: Huber function

Let $\rho > 0$ and set

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

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

2. (See the computation of $f'(x)$ done above.) For $\rho = 1$



Exercise 2

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

- Give a sufficient condition for $x \in \mathcal{H}$ to be a global minimizer of $f + \iota_C$.
- *Assume that $f \in \Gamma_0(\mathcal{H})$ and that C is a closed convex set.*

  Then, from the properties of C, $\iota_C \in \Gamma_0(\mathcal{H})$.
  From Fermat's rule,

  $$\widehat{x} \text{ is a minimizer of } f + \iota_C \text{ iff } 0 \in \partial(f + \iota_C)(\widehat{x}).$$

  Since $\operatorname{dom} f \cap C \neq \varnothing$, then $\partial(f + \iota_C) = \partial f + \partial \iota_C$.
  Moreover $(\forall x \in \mathcal{H}) \quad \partial \iota_C(x) = N_C(x)$, the normal cone of C at x.

  $$\text{Thus, } \widehat{x} \text{ is a minimizer of } f + \iota_C \text{ iff } 0 \in \partial f(\widehat{x}) + N_C(\widehat{x}).$$

  That is *if the normal cone of C at $\widehat{x}$ contains a subgradient of f at $\widehat{x}$*.