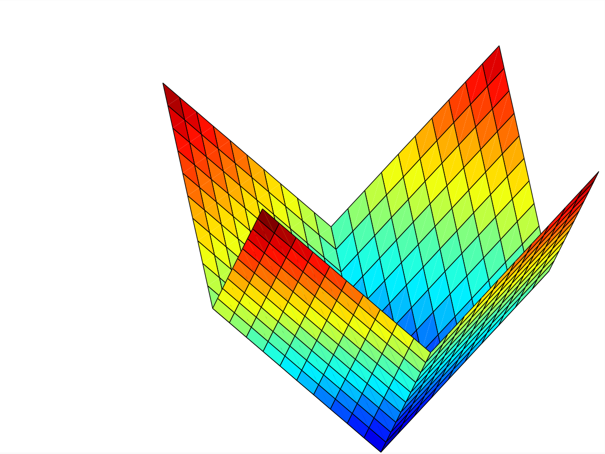
Optimization
– Subdifferential

Nelly Pustelnik

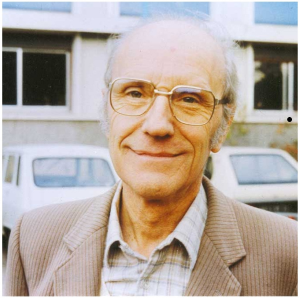
CNRS, Laboratoire de Physique de l'ENS de Lyon, France

(several slides in this part traced back to Tutorial ICASSP 2014 written in collaboration with **Jean-Christophe Pesquet** from Centre de Vision Numérique, CentraleSupelec, University Paris-Saclay, Inria, France.)

Non-smooth convex optimization



A pioneer



Jean-Jacques Moreau
(1923–2014)

Subdifferential of function: definition

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

Subdifferential of function: definition

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

Subdifferential of function: definition

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

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leqslant f(y)\} \end{aligned}$$

Subdifferential of a function: properties

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

$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leqslant f(y)\}$$

Fermat's rule: $0 \in \partial f(\hat{x}) \Leftrightarrow \hat{x} \in \underset{x}{\mathrm{Argmin}}\ f(x)$

Subdifferential of a function: properties

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

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leqslant f(y)\} \end{aligned}$$

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

Subdifferential of a convex function: properties

If $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ is convex and it is Gâteaux differentiable at x, then

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

Subdifferential of a convex function: example

Let C be a nonempty subset of $\mathcal{H}$.

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

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

Subdifferential calculus

Let $\mathcal{H}$ and $\mathcal{G}$ be two 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 $f : \mathcal{H} \to]-\infty, +\infty]$, $g : \mathcal{G} \to]-\infty, +\infty]$, and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
  If $dom\, g \cap L(dom\, f) \neq \varnothing$, then

$$(\forall x \in \mathcal{H}) \qquad \partial f(x) + L^* \partial g(Lx) \subset \partial(f + g \circ L)(x).$$

Subdifferential calculus

Let $\mathcal{H}$ and $\mathcal{G}$ be two real Hilbert spaces.
Let $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
If $int\,(dom\,g) \cap L(dom\,f) \neq \varnothing$ or $dom\,g \cap int\,\big(L(dom\,f)\big) \neq \varnothing$, then

$$\partial f + L^* \partial g\, L = \partial(f + g \circ L).$$

Notations:

- $int\, C$ denotes the interior of a subset C of $\mathcal{H}$.
- $ran\, A$ denotes the range of a set-valued operator from $\mathcal{H}$ to $\mathcal{G}$ such that $ran\, A = A(\mathcal{H})$.

Particular case:

- If $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{H})$, and $dom\, g = \mathcal{H}$ (or $dom\, f = \mathcal{H}$), then $\partial f + \partial g = \partial(f + g)$.
- If $g \in \Gamma_0(\mathcal{G})$, $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$, and $int\,(dom\, g) \cap ran\, L \neq \varnothing$, then $L^* \partial g\, L = \partial(g \circ L)$.

Subdifferential calculus

Let $(\mathcal{H})_{i\in I}$ where $I \subset \mathbb{N}$ be Hilbert spaces and let $\mathcal{H} = \bigoplus_{i\in I} \mathcal{H}_i$.
For every $i \in I$, let $f_i \colon \mathcal{H}_i \to \left]-\infty, +\infty\right]$ be a proper function. Let

$$f \colon \mathcal{H} \to \left]-\infty, +\infty\right] : x = (x_i)_{i\in I} \mapsto \sum_{i\in I} f_i(x_i)$$

Then,

$$\big(\forall x = (x_i)_{i\in I} \in \mathcal{H}\big) \qquad \partial f(x) = \bigtimes_{i\in I} \partial f_i(x_i).$$

Subdifferential calculus

Let $(\mathcal{H})_{i\in I}$ where $I \subset \mathbb{N}$ be Hilbert spaces and let $\mathcal{H} = \bigoplus_{i\in I} \mathcal{H}_i$.
For every $i \in I$, let $f_i \colon \mathcal{H}_i \to \,]-\infty, +\infty]$ be a proper function. Let

$$f \colon \mathcal{H} \to \,]-\infty, +\infty] : x = (x_i)_{i\in I} \mapsto \sum_{i\in I} f_i(x_i)$$

Then,

$$\big(\forall x = (x_i)_{i\in I} \in \mathcal{H}\big) \qquad \partial f(x) = \bigtimes_{i\in I} \partial f_i(x_i).$$

Proof: Let $x = (x_i)_{i\in I} \in \mathcal{H}$. We have

$$\begin{aligned}
& t = (t_i)_{i\in I} \in \bigtimes_{i\in I} \partial f_i(x_i) \\
\Leftrightarrow \quad & (\forall i \in I)(\forall y_i \in \mathcal{H}_i) \;\; f_i(y_i) \geqslant f_i(x_i) + \langle t_i \mid y_i - x_i \rangle \\
\Rightarrow \quad & \big(\forall y = (y_i)_{i\in I} \in \mathcal{H}\big) \;\; \sum_{i\in I} f_i(y_i) \geqslant \sum_{i\in I} f_i(x_i) + \sum_{i\in I} \langle t_i \mid y_i - x_i \rangle \\
\Leftrightarrow \quad & \big(\forall y \in \mathcal{H}\big) \;\; f(y) \geqslant f(x) + \langle t \mid y - x \rangle .
\end{aligned}$$

Subdifferential calculus

Let $(\mathcal{H})_{i\in I}$ where $I \subset \mathbb{N}$ be Hilbert spaces and let $\mathcal{H} = \bigoplus_{i\in I} \mathcal{H}_i$.
For every $i \in I$, let $f_i \colon \mathcal{H}_i \to \,]-\infty, +\infty]$ be a proper function. Let

$$f \colon \mathcal{H} \to \,]-\infty, +\infty] : x = (x_i)_{i\in I} \mapsto \sum_{i\in I} f_i(x_i)$$

Then,

$$\big(\forall x = (x_i)_{i\in I} \in \mathcal{H}\big) \qquad \partial f(x) = \bigtimes_{i\in I} \partial f_i(x_i).$$

Proof: Conversely,

$$\begin{aligned} & t = (t_i)_{i\in I} \in \partial f(x) \\ \Leftrightarrow \quad & \big(\forall y = (y_i)_{i\in I} \in \mathcal{H}\big) \ \ \sum_{i\in I} f_i(y_i) \geqslant \sum_{i\in I} f_i(x_i) + \sum_{i\in I} \langle t_i \mid y_i - x_i \rangle . \end{aligned}$$

Let $j \in I$. By setting $(\forall i \in I \backslash \{j\})\ y_i = x_i \in dom\, f_i$, we get

$$(\forall y_j \in \mathcal{H}_j) \ \ f_j(y_j) \geqslant f_j(x_j) + \langle t_j \mid y_j - x_j \rangle .$$

L1 norm

→ **ℓ_1−norm**

$$f : \mathbb{R}^N \to \mathbb{R} : (x_i)_{1 \leqslant i \leqslant N} \mapsto \sum_{i=1}^{N} |x_i|$$

Then

$$\partial|\cdot| : \zeta \mapsto \begin{cases} -1 & \text{if } \zeta < 0; \\ [-1, 1] & \text{if } \zeta = 0, \\ 1 & \text{if } \zeta > 0; \end{cases}$$

Huber function

→ **Smooth approximation of the ℓ_1−norm** parametrized by $\mu > 0$. [**Combettes-Glaudin,2019**]

$$f : \mathbb{R}^N \to \mathbb{R} : (x_i)_{1 \leqslant i \leqslant N} \mapsto \sum_{i=1}^{N} f_i(x_i)$$

and

$$f_i : \zeta \mapsto \begin{cases} |\zeta| - \frac{\mu}{2}, & \text{if } |\zeta| > \mu; \\ \frac{|\zeta|^2}{2\mu}, & \text{if } |\zeta| \leqslant \mu. \end{cases}$$

Note that, since

$$\partial f_i = \nabla f_i : \zeta \mapsto \begin{cases} \frac{\zeta}{|\zeta|}, & \text{if } |\zeta| > \mu; \\ \frac{\zeta}{\mu}, & \text{if } |\zeta| \leqslant \mu, \end{cases}$$