Optimization – Conjugate

Nelly Pustelnik

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

Conjugate

Adrien-Marie Legendre
(1752–1833)

Werner Fenchel
(1905–1988)

Conjugate

Adrien-Marie Legendre
(1752–1833)

Werner Fenchel
(1905–1988)

Conjugate: definition

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

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

Conjugate: definition

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

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

Examples :

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

  Proof : 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 at $x = u$.
  Consequently, $f^*(u) = \frac{1}{2}\|u\|^2$.

Conjugate: definition

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

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

Examples :

- $f = \frac{1}{2}\|\cdot\|^2 \Rightarrow f^* = \frac{1}{2}\|\cdot\|^2$.
- $(\forall x \in \mathbb{R}^N)\ f(x) = \frac{1}{q}\|x\|_q^q$ with $q \in]1, +\infty[$
  $\Rightarrow (\forall u \in \mathbb{R}^N)\ \boxed{f^*(u) = \frac{1}{q^*}\|u\|_{q^*}^{q^*} \text{ with } \frac{1}{q} + \frac{1}{q^*} = 1}$

Conjugate: definition

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

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

- If f is even, then f^* is even.

Conjugate: definition

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

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

- 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 $\alpha \in \mathbb{R}$,
  $\big(f(\cdot - y) + \langle \cdot \mid v \rangle + \alpha\big)^* = f^*(\cdot - v) + \langle y \mid \cdot - v \rangle - \alpha$.
- Let $\mathcal{G}$ be a Hilbert space and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ be an isomorphism.
  $(f \circ L)^* = f^* \circ (L^{-1})^*$.
- f^* is l.s.c. and convex.

Conjugate: definition

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

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

Moreau-Fenchel theorem

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

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

Conjugate: definition

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

The conjugate of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

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

Moreau-Fenchel theorem

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

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

Conjugate: definition

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

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

- Consequence: If $f \in \Gamma_0(\mathbb{R})$, then f^* is proper, hence $f^* \in \Gamma_0(\mathbb{R})$.

Conjugate: properties

Fenchel-Young inequality: If f is proper, then

1. $\big(\forall (x,u) \in \mathcal{H}^2\big) \qquad f(x) + f^*(u) \geq \langle x \mid u \rangle$

2. $\big(\forall (x,u) \in \mathcal{H}^2\big) \qquad u \in \partial f(x) \Leftrightarrow \quad f(x) + f^*(u) = \langle x \mid u \rangle.$

If $f \in \Gamma_0(\mathcal{H})$, then

$$\big(\forall (x,u) \in \mathcal{H}^2\big) \qquad u \in \partial f(x) \;\Leftrightarrow\; x \in \partial f^*(u).$$

Proximity operator: Moreau decomposition

Moreau decomposition formula

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0, +\infty[$.

$$(\forall x \in \mathcal{H}) \qquad \boxed{\operatorname{prox}_{\gamma f^*} x = x - \gamma \operatorname{prox}_{\gamma^{-1} f}(\gamma^{-1} x)}\,.$$

Proof:

$$\begin{aligned}
p = \operatorname{prox}_{\gamma f^*} x &\Leftrightarrow x - p \in \gamma \partial f^*(p) \\
&\Leftrightarrow p \in \partial f\left(\frac{x-p}{\gamma}\right) \\
&\Leftrightarrow \frac{x}{\gamma} - \frac{x-p}{\gamma} \in \frac{1}{\gamma} \partial f\left(\frac{x-p}{\gamma}\right) \\
&\Leftrightarrow \frac{x-p}{\gamma} = \operatorname{prox}_{\gamma^{-1} f}(\gamma^{-1} x) \\
&\Leftrightarrow p = x - \gamma \operatorname{prox}_{\gamma^{-1} f}(\gamma^{-1} x).
\end{aligned}$$

Proximity operator: Moreau decomposition

Moreau decomposition formula

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0, +\infty[$.

$$(\forall x \in \mathcal{H}) \qquad \boxed{\mathrm{prox}_{\gamma f^*} x = x - \gamma \mathrm{prox}_{\gamma^{-1} f}(\gamma^{-1} x)}\,.$$

Example: If $\mathcal{H} = \mathbb{R}^N$, $f = \frac{1}{q}\|\cdot\|_q^q$ with $q \in]1, +\infty[$, then $f^* = \frac{1}{q^*}\|\cdot\|_{q^*}^{q^*}$ with $1/q + 1/q^* = 1$, and

$$(\forall x \in \mathbb{R}^N) \qquad \mathrm{prox}_{\frac{\gamma}{q^*}\|\cdot\|_{q^*}^{q^*}} x = x - \gamma \mathrm{prox}_{\frac{1}{\gamma q}\|\cdot\|_q^q}(\gamma^{-1} x).$$

Conjugate: properties

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 $f_i \colon \mathcal{H}_i \to \,]-\infty, +\infty]$. Let

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

Then,

$$\big(\forall u = (u_i)_{i\in I} \in \mathcal{H}\big) \qquad f^*(u) = \sum_{i\in I} f_i^*(u_i).$$

Conjugate: properties

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 $f_i : \mathcal{H}_i \to \,]-\infty, +\infty]$. Let

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

Then,

$$(\forall u = (u_i)_{i\in I} \in \mathcal{H}) \qquad f^*(u) = \sum_{i\in I} f_i^*(u_i) .$$

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

$$\begin{aligned} f^*(u) &= \sup_{x\in\mathcal{H}} \langle x \mid u\rangle - f(x) \\ &= \sup_{x=(x_i)_{i\in I}\in\mathcal{H}} \sum_{i\in I} \langle x_i \mid u_i\rangle - f_i(x_i) \\ &= \sum_{i\in I} \sup_{x_i\in\mathcal{H}_i} \langle x_i \mid u_i\rangle - f_i(x_i) \\ &= \sum_{i\in I} f_i^*(u_i). \end{aligned}$$

Conjugate: example

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$.

σ_C is the support function of C if

$$(\forall u \in \mathcal{H}) \qquad \sigma_C(u) = \sup_{x \in C} \langle x \mid u \rangle$$
$$= \iota_C^*(u).$$

Proximity operator: support function

Support function :

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$ be nonempty closed convex.

$$(\forall x \in \mathcal{H}) \qquad \mathrm{prox}_{\sigma_C} = \mathrm{Id} - P_C.$$

Conjugate: example

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

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

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

Conjugate: example

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

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

$$C = \left\{ y \in \mathbb{R}^N \;\middle|\; \|y\|_{q^*} \leq 1 \right\} \qquad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

Conjugate: example

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

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

$$C = \left\{ y \in \mathbb{R}^N \mid \|y\|_{q^*} \leq 1 \right\} \qquad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

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

Conjugate: example

- Let $f : \mathbb{R} \to \left]-\infty, +\infty\right] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases}$
  with $-\infty \leq \delta_1 < \delta_2 \leq +\infty$.
  Then, $f = \sigma_C$ where C is the closed real interval such that $\inf C = \delta_1$ et $\sup C = \delta_2$.

Conjugate: example

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

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

$$C = \left\{ y \in \mathbb{R}^N \;\middle|\; \|y\|_{q^*} \leq 1 \right\} \qquad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

Conjugate: example

- Let $f\colon \mathbb{R} \to \,]-\infty, +\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases}$
  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
  $$C = \left\{ y \in \mathbb{R}^N \;\middle|\; \|y\|_{q^*} \le 1 \right\} \qquad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

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

Conjugate: example

- Let $f = \iota_C$ with $C =]-\infty, b]^N$ with $b \in \mathbb{R}$, then

$$(\forall u \in \mathbb{R}^N) \qquad f^*(u) = \langle b, u \rangle + \iota_{\geq 0}(u)$$

- Let $f = \iota_C$ with $C = [b, +\infty[^N$ with $b \in \mathbb{R}$, then

$$(\forall u \in \mathbb{R}^N) \qquad f^*(u) = \langle b, u \rangle + \iota_{\leq 0}(u)$$

- Many others in:
    - the work by Komodakis and Pesquet [link]
    - Beck's table [link]

Link with SVM

- Reminder about SVM

$$\min_{\mathrm{w}} \frac{1}{2}\|\mathrm{w}\|^2 \ \text{s.t.} \qquad (\forall \ell) \qquad \mathrm{c}_\ell\Big(\langle \mathrm{w}, \mathrm{u}_\ell\rangle + \mathrm{b}\Big) \geq 1$$

- Simplified formulation:

$$\begin{aligned}
& \min_{\theta} \frac{1}{2}\|\theta\|^2 \qquad \text{s.t.} \qquad X\theta \geq \mathbf{1} \\
\Leftrightarrow \quad & \min_{\theta} \max_{\mathrm{v}} \quad \frac{1}{2}\|\theta\|^2 + \langle \mathrm{v}, X\theta\rangle - f^*(\mathrm{v}) \quad \text{where} \quad f = \iota_{\geq \mathbf{1}} \\
\Leftrightarrow \quad & \max_{\mathrm{v}} \quad \frac{1}{2}\|X^\top \mathrm{v}\|^2 - \langle \mathrm{v}, XX^\top \mathrm{v}\rangle - f^*(\mathrm{v}) \quad \text{where} \quad f = \iota_{\geq \mathbf{1}} \\
\Leftrightarrow \quad & \max_{\mathrm{v}} \quad -\frac{1}{2}\|X^\top \mathrm{v}\|^2 - \langle \mathrm{v}, \mathbf{1}\rangle \quad \text{s.t.} \quad \mathrm{v} \leq 0 \\
\Leftrightarrow \quad & \max_{\mathrm{y}} \quad -\frac{1}{2}\|X^\top \mathrm{y}\|^2 + \langle \mathrm{y}, \mathbf{1}\rangle \quad \text{s.t.} \quad \mathrm{y} \geq 0 \\
\Leftrightarrow \quad & \max_{\mathrm{y}} \quad -\frac{1}{2}\langle \mathrm{y}, XX^\top \mathrm{y}\rangle + \langle \mathrm{y}, \mathbf{1}\rangle \quad \text{s.t.} \quad \mathrm{y} \geq 0
\end{aligned}$$