Optimization
Part VI: Conjugate

Nelly Pustelnik

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

(several slides in this part traced back Tutorial ICASSP 2014 written in collaboration with **Jean-Christophe Pesquet** from Centre de Vision Numérique, CentraleSupelec, University Paris-Saclay, Inria, 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 $f\colon \mathcal{H} \to \,]-\infty, +\infty]$.

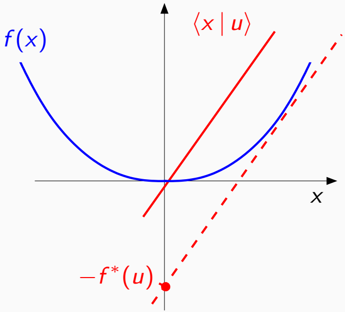
The conjugate of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

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

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

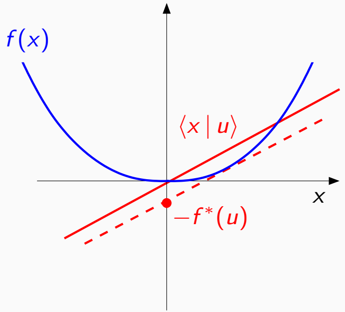
$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x \in \operatorname{dom} f} \big(\langle x \mid u \rangle - f(x) \big)$$



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

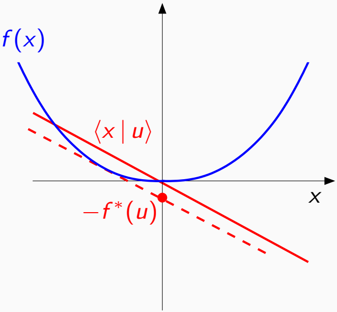
$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x \in \operatorname{dom} f} \big(\langle x \mid u \rangle - f(x) \big)$$



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

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



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

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

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

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

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

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

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

Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$.
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 $\alpha \in \left]0, +\infty\right[$, $(\alpha f)^* = \alpha f^*(\cdot/\alpha)$.
- 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 $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$.
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 $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ 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

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

Moreau-Fenchel theorem

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

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

- 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 $\gamma \in \,]0, +\infty[$.

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

<u>Proof</u>:

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

Proximity operator: Moreau decomposition

Moreau decomposition formula

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

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

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

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

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

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

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

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

Then,

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

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

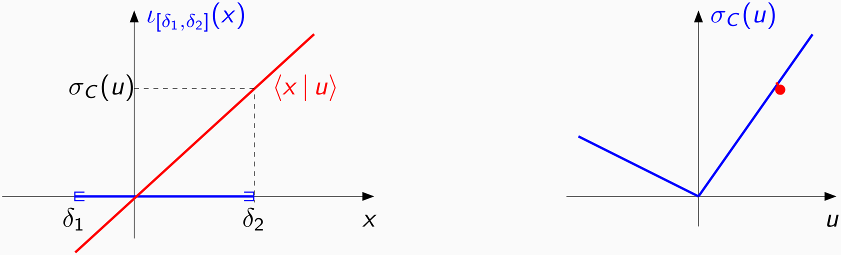
$$\begin{aligned} f^*(u) &= \sup_{x\in\mathcal{H}} \langle x \mid u\rangle - f(x) \\ &= \sup_{x=(x_i)_{i\in I}\in\mathcal{H}} \sum_{i\in I} \langle x_i \mid u_i\rangle - f_i(x_i) \\ &= \sum_{i\in I} \sup_{x_i\in\mathcal{H}_i} \langle x_i \mid u_i\rangle - f_i(x_i) \\ &= \sum_{i\in I} f_i^*(u_i). \end{aligned}$$

Conjugate: example

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

σ_C is the support function of C if

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

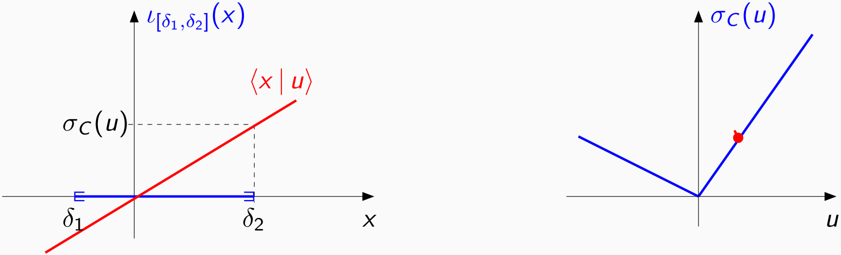


Conjugate: example

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

σ_C is the support function of C if

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

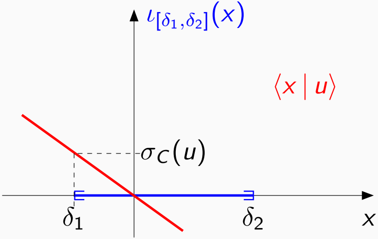


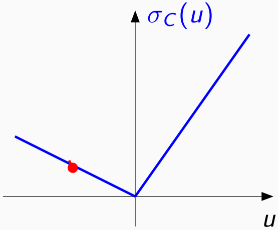
Conjugate: example

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.

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

Conjugate: example

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

  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 $\sup C = \delta_2$.

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

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

Conjugate: example

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

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

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

Proximity operator: existence and uniqueness

Let $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in\ ]0, +\infty[$.
For every $x \in \mathcal{H}$, there exists a unique vector $p \in \mathcal{H}$ such that

$$f(p) + \frac{1}{2\gamma}\|p - x\|^2 = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma}\|y - x\|^2 .$$

Proof: $f \in \Gamma_0(\mathcal{H}) \Leftrightarrow f^* \in \Gamma_0(\mathcal{H})$. Thus, there exists $u \in \mathcal{H}$ such that $f^*(u) \in \mathbb{R}$. According to Fenchel-Young inequality, we have

$$(\forall y \in \mathcal{H}) \qquad f(y) \geq \langle u \mid y \rangle - f^*(u).$$

Then, $f(y) + (2\gamma)^{-1}\|y - x\|^2 \to +\infty$ when $\|y\| \to +\infty$.
Furthermore $(2\gamma)^{-1}\|\cdot - x\|^2$ being strictly convex, $f + (2\gamma)^{-1}\|\cdot - x\|^2$ is a strictly convex coercive function.

α-averaged operator: example

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $\nu \in]0, 2[$.
If f is differentiable with a ν-lipschitzian gradient then, $\mathrm{Id} - \nabla f$ is a $\nu/2$-averaged operator.

Remark : $\mathrm{Id} - \nabla f$ denotes the gradient descent operator.

α-averaged operator: example

Proof : 1) Descent lemma

For every $(x, y) \in \mathcal{H}^2$ and $t \in \mathbb{R}$, let $\varphi(t) = f(x + t(y - x))$.
φ is differentiable and $\varphi'(t) = \langle y - x \mid \nabla f(x + t(y - x))\rangle$. We have then

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt$$

$$\Leftrightarrow \; f(y) - f(x) - \langle y - x \mid \nabla f(x)\rangle = \int_0^1 \langle y - x \mid \nabla f(x + t(y - x)) - \nabla f(x)\rangle \, dt.$$

But, according to the Cauchy-Schwarz inequality,

$$\begin{aligned} &\langle y - x \mid \nabla f(x + t(y - x)) - \nabla f(x)\rangle \\ &\le \|y - x\| \|\nabla f(x + t(y - x)) - \nabla f(x)\| \le t\nu \|y - x\|^2. \end{aligned}$$

This leads to

$$\boxed{\left(\forall (x, y) \in \mathcal{H}^2\right) \quad f(y) \le f(x) + \langle y - x \mid \nabla f(x)\rangle + \frac{\nu}{2}\|y - x\|^2.}$$

α-averaged operator: example

Proof : 2) $\mathrm{Id} - \nabla f$ is α-averaged

From the descent lemma, for every $(x, y, z) \in \mathcal{H}^3$,

$$\begin{aligned} f^*(\nabla f(y)) &\geq \langle z \mid \nabla f(y) \rangle - f(z) \\ &\geq \langle z \mid \nabla f(y) - \nabla f(x) \rangle + \langle x \mid \nabla f(x) \rangle - f(x) - \frac{\nu}{2}\|z - x\|^2. \end{aligned}$$

Moreover, according to the Fenchel-Young inequality,

$$\langle x \mid \nabla f(x) \rangle - f(x) = f^*(\nabla f(x)).$$

Thus,

$$f^*(\nabla f(y)) \geq \langle z \mid \nabla f(y) - \nabla f(x) \rangle + f^*(\nabla f(x)) - \frac{\nu}{2}\|z - x\|^2$$

α-averaged operator: example

Proof : 2) $\mathrm{Id} - \nabla f$ is α-averaged

Thus,

$$f^*(\nabla f(y)) \geq \langle z \mid \nabla f(y) - \nabla f(x)\rangle + f^*(\nabla f(x)) - \frac{\nu}{2}\|z - x\|^2$$
$$= f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \langle z - x \mid \nabla f(y) - \nabla f(x)\rangle - \frac{\nu}{2}\|z - x\|^2.$$

This yields

$$\begin{aligned} f^*(\nabla f(y)) \geq & f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle \\ & + \left(\nu\|\cdot\|^2/2\right)^*\left(\nabla f(y) - \nabla f(x)\right) \\ \geq & f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2. \end{aligned}$$

Consequently,

$$\boxed{f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2.}$$

α-averaged operator: example

Proof : 2) $\mathrm{Id} - \nabla f$ is α-averaged

For every $(x, y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x) \rangle + \frac{1}{2\nu} \|\nabla f(y) - \nabla f(x)\|^2$$

and symmetrically

$$f^*(\nabla f(x)) \geq f^*(\nabla f(y)) + \langle y \mid \nabla f(x) - \nabla f(y) \rangle + \frac{1}{2\nu} \|\nabla f(x) - \nabla f(y)\|^2.$$

By summing

$$-\langle y - x \mid \nabla f(y) - \nabla f(x) \rangle + \frac{1}{\nu} \|\nabla f(x) - \nabla f(y)\|^2 \leq 0.$$

It results that

$$\boxed{\|(\mathrm{Id} - \nabla f)x - (\mathrm{Id} - \nabla f)y\|^2 + \frac{1 - \nu/2}{\nu/2} \|\nabla f(x) - \nabla f(y)\|^2 \leq \|x - y\|^2.}$$

Optimization
Part VII: Duality

Nelly Pustelnik

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

Fenchel-Rockafellar duality

Primal problem

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

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$, $g\colon \mathcal{G} \to \,]-\infty, +\infty]$. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

We want to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + g(Lx).$$

Dual problem

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

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$, $g\colon \mathcal{G} \to \,]-\infty, +\infty]$. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

We want to

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; f^*(-L^*v) + g^*(v).$$

Fenchel-Rockafellar duality

Weak duality

Let $\mathcal{H}$ and $\mathcal{G}$ be two real Hilbert spaces.
Let f be a proper fonction from $\mathcal{H}$ to $]-\infty,+\infty]$, g be a proper function from $\mathcal{G}$ to $]-\infty,+\infty]$, and $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. Let

$$\mu = \inf_{x\in\mathcal{H}} f(x) + g(Lx) \quad \text{and} \quad \mu^* = \inf_{v\in\mathcal{G}} f^*(-L^*v) + g^*(v).$$

We have $\mu \geq -\mu^*$. If $\mu \in \mathbb{R}$, $\mu + \mu^*$ is called the **duality gap**.

Fenchel-Rockafellar duality

Weak duality

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

Let f be a proper fonction from $\mathcal{H}$ to $]-\infty,+\infty]$, g be a proper function from $\mathcal{G}$ to $]-\infty,+\infty]$, and $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. Let

$$\mu = \inf_{x\in\mathcal{H}} f(x) + g(Lx) \quad \text{and} \quad \mu^* = \inf_{v\in\mathcal{G}} f^*(-L^*v) + g^*(v).$$

We have $\mu \geq -\mu^*$. If $\mu \in \mathbb{R}$, $\mu + \mu^*$ is called the **duality gap**.

Proof: According to Fenchel-Young inequality, for every $x \in \mathcal{H}$ and $v \in \mathcal{G}$,

$$f(x) + g(Lx) + f^*(-L^*v) + g^*(v) \geq \langle x \mid -L^*v\rangle + \langle Lx \mid v\rangle = 0.$$

Fenchel-Rockafellar duality

Strong duality

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

$$\mu = \inf_{x \in \mathcal{H}} f(x) + g(Lx) = -\min_{v \in \mathcal{G}} f^*(-L^* v) + g^*(v) = -\mu^* .$$

Fenchel-Rockafellar duality

Duality theorem (1)

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 $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

$$\operatorname{zer}\left(\partial f + L \partial g L^*\right) \neq \varnothing \quad \Leftrightarrow \quad \operatorname{zer}\left((-L)\partial f^*(-L^*) + \partial g^*\right) \neq \varnothing.$$

Fenchel-Rockafellar duality

Duality theorem (1)

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 $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

$$\operatorname{zer}\left(\partial f + L \partial g L^*\right) \neq \varnothing \quad \Leftrightarrow \quad \operatorname{zer}\left((-L)\partial f^*(-L^*) + \partial g^*\right) \neq \varnothing.$$

Proof:

$$\begin{aligned}
(\exists x \in \mathcal{H}) \;\; 0 \in \partial f(x) + L^* \partial g(Lx) \quad &\Leftrightarrow \quad (\exists x \in \mathcal{H})(\exists v \in \mathcal{G}) \quad \begin{cases} -L^* v \in \partial f(x) \\ v \in \partial g(Lx) \end{cases} \\
&\Leftrightarrow \quad (\exists x \in \mathcal{H})(\exists v \in \mathcal{G}) \quad \begin{cases} x \in \partial f^*(-L^* v) \\ Lx \in \partial g^*(v) \end{cases} \\
&\Leftrightarrow \quad (\exists v \in \mathcal{G}) \qquad 0 \in -L\partial f^*(-L^* v) + \partial g^*(v).
\end{aligned}$$

Fenchel-Rockafellar duality

Duality theorem (2)

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 there exists $\widehat{x} \in \mathcal{H}$ such that $0 \in \partial f(\widehat{x}) + L^* \partial g(L\widehat{x})$, then $\widehat{x}$ is a solution to the primal problem. Moreover, there exists a solution $\widehat{v}$ to the dual problem such that $-L^*\widehat{v} \in \partial f(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v})$.
- If there exists $(\widehat{x}, \widehat{v}) \in \mathcal{H} \times \mathcal{G}$ such that $-L^*\widehat{v} \in \partial f(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v})$ then $\widehat{x}$ (resp. $\widehat{v}$) is a solution to the primal (resp. dual) problem.

If $(\widehat{x}, \widehat{v}) \in \mathcal{H} \times \mathcal{G}$ is such that $-L^*\widehat{v} \in \partial f(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v})$,
then $(\widehat{x}, \widehat{v})$ is called a **Kuhn-Tucker point**.

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
⇒ **Lagrangian interpretation**

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ f(x)+g(Lx) \quad\Leftrightarrow\quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\ f(x)+g(y)$$

Lagrange function:

$$\big(\forall(x,y,v)\in\mathcal{H}\times\mathcal{G}^2\big)\ \ \mathcal{L}(x,y,z)=f(x)+g(y)+\langle z\mid Lx-y\rangle$$

where $z\in\mathcal{G}$ denotes the Lagrange multiplier.

Alternating-direction method of multipliers

Idea: iterations for finding a saddle point $(\widehat{x}, \widehat{y}, \widehat{z})$:

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n \in \operatorname{Argmin} \mathcal{L}(\cdot, y_n, z_n) \\ y_{n+1} \in \operatorname{Argmin} \mathcal{L}(x_n, \cdot, z_n) \\ v_{n+1} \text{ such that } \mathcal{L}(x_n, y_{n+1}, v_{n+1}) \geq \mathcal{L}(x_n, y_{n+1}, v_n). \end{cases}$$

But the convergence is not guaranteed in general !

Alternating-direction method of multipliers

Idea: iterations for finding a saddle point $(\widehat{x}, \widehat{y}, \widehat{z})$:

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n \in \operatorname{Argmin} \mathcal{L}(\cdot, y_n, z_n) \\ y_{n+1} \in \operatorname{Argmin} \mathcal{L}(x_n, \cdot, z_n) \\ v_{n+1} \text{ such that } \mathcal{L}(x_n, y_{n+1}, v_{n+1}) \geq \mathcal{L}(x_n, y_{n+1}, v_n). \end{cases}$$

But the convergence is not guaranteed in general !

Solution: introduce an Augmented Lagrange function.

Let $\gamma \in \,]0, +\infty[$, we define

$$\left(\forall (x, y, z) \in \mathcal{H} \times \mathcal{G}^2\right) \; \widetilde{\mathcal{L}}(x, y, z) = f(x) + g(y) + \gamma \langle z \mid Lx - y \rangle + \frac{\gamma}{2}\|Lx - y\|^2$$

The Lagrange multiplier is $v = \gamma z$.

Alternating-direction method of multipliers

Algorithm for finding a saddle point:

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \;\; \widetilde{\mathcal{L}}(x, y_n, z_n) \\ y_{n+1} = \underset{y \in \mathcal{G}}{\operatorname{argmin}} \;\; \widetilde{\mathcal{L}}(x_n, y, z_n) \\ z_{n+1} \text{ such that } \widetilde{\mathcal{L}}(x_n, y_{n+1}, z_{n+1}) \geq \widetilde{\mathcal{L}}(x_n, y_{n+1}, z_n). \end{cases}$$

By performing a gradient ascent on the Lagrange multiplier,

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \;\; f(x) + \gamma \langle z_n \mid Lx - y_n \rangle + \frac{\gamma}{2}\|Lx - y_n\|^2 \\ y_{n+1} = \underset{y \in \mathcal{G}}{\operatorname{argmin}} \;\; g(y) + \gamma \langle z_n \mid Lx_n - y \rangle + \frac{\gamma}{2}\|Lx_n - y\|^2 \\ z_{n+1} = z_n + \frac{1}{\gamma} \nabla_z \widetilde{\mathcal{L}}(x_n, y_{n+1}, z_n) \end{cases}$$

$$\Leftrightarrow \quad (\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \;\; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}} (z_n + Lx_n) \\ z_{n+1} = z_n + Lx_n - y_{n+1}. \end{cases}$$

Augmented Lagrange method

ADMM algorithm (*Alternating-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ et $g \in \Gamma_0(\mathcal{G})$.

Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}} (z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

Augmented Lagrange method

ADMM algorithm (*Alternating-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ et $g \in \Gamma_0(\mathcal{G})$.

Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.

We assume that $\mathrm{int}\,(\mathrm{dom}\, g) \cap L(\mathrm{dom}\, f) \neq \varnothing$ or $\mathrm{dom}\, g \cap \mathrm{int}\,\big(L(\mathrm{dom}\, f)\big) \neq \varnothing$ and that $\mathrm{Argmin}(f + g \circ L) \neq \varnothing$. Let

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\mathrm{argmin}} \ \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \mathrm{prox}_{\frac{g}{\gamma}} (z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

We have:

- $x_n \rightharpoonup \widehat{x}$ where $\widehat{x} \in \mathrm{Argmin}(f + g \circ L)$
- $\gamma z_n \rightharpoonup \widehat{v}$ where $\widehat{v} \in \mathrm{Argmin}(f^* \circ (-L^*) + g^*)$.

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = \mathrm{prox}_{\gamma g^*} u_n \\ w_n = \mathrm{prox}_{\gamma f^* \circ (-L^*)}(2v_n - u_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n \in \gamma \partial\big(f^* \circ (-L^*)\big) w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \partial f^* \circ (-L^*) w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \underbrace{\partial f^* \circ (-L^*) w_n}_{x_n \in} \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \left]0, +\infty\right[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n = -\gamma L x_n \\ x_n \in \partial f^*(-L^* w_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n = -\gamma L x_n \\ -L^* w_n \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ L^*(u_n - 2v_n - \gamma L x_n) \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

using $y_n = \gamma^{-1}(u_n - v_n)$ and $z_n = \gamma^{-1} v_n$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ L^*(y_n - z_n - Lx_n) \in \frac{1}{\gamma}\partial f(x_n) \\ u_{n+1} = \gamma z_n + \gamma L x_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ x_n = \underset{x \in \mathcal{H}}{\mathrm{argmin}} \; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ u_{n+1} = \gamma z_n + \gamma L x_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)
$\equiv$ **Douglas-Rachford for the dual problem**

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \; \frac{1}{2} \| Lx - y_n + z_n \|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}} (z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm:

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:
  * $h \in C^{1,1}_{\zeta}(\mathcal{H})$
  * No operator inversion.
  * Allow the use of proximable or/and differentiable functions.
  * Convergence when $\frac{1}{\tau} - \sigma\|L\|^2 > \frac{\zeta}{2}$.
  * $x_n \rightharpoonup \widehat{x} \in \mathrm{Argmin}\big(f + h + g \circ L\big)$

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ **Chambolle-Pock algorithm**

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ **Chambolle-Pock algorithm**

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_{n+1} = \mathrm{prox}_{\tau f}\big(x_n - \tau L^* v_n\big) \\ y_n = 2x_{n+1} - x_n \\ v_{n+1} = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma L y_n\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ Forward-backward algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$ and $L = 0$, this yields the forward-backward algorithm.

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ Forward-backward algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau \nabla h(x_n)\big) \\ x_{n+1} = x_n + \lambda_n\big(p_n - x_n\big)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$ and $L = 0$, this yields the forward-backward algorithm.

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ Douglas-Rachford algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:
  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.
  * When $g = 0$ and $L = 0$, this yields the forward-backward algorithm.
  * In the limit case when $h = 0$, $\lambda_n \equiv 1$, $L = \mathrm{Id}$ and $\sigma = 1/\tau$, this yields the Douglas-Rachford algorithm.

Primal-dual optimization algorithm : $\min_x f(x) + h(x) + g(Lx)$

Condat-Vũ algorithm: $\Rightarrow$ Douglas-Rachford algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_{n+1} = \mathrm{prox}_{\tau f}(x_n - \tau v_n) \\ s_n = \mathrm{prox}_{\tau g}(2x_{n+1} - (x_n - \tau v_n)) \\ x_{n+1} - \tau v_{n+1} = (x_n - \tau v_n) + s_n - x_{n+1} \end{cases}$$

- Remark:

  * When $h = 0$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$ and $L = 0$, this yields the forward-backward algorithm.

  * In the limit case when $h = 0$, $\lambda_n \equiv 1$, $L = \mathrm{Id}$ and $\sigma = 1/\tau$, this yields the Douglas-Rachford algorithm.

Optimization algorithms

Forward-Backward	$f_1 + f_2$	f_1 grad. Lipschitz prox_{f_2}	[Combettes,Wajs,2005]
ISTA	$f_1 + f_2$	f_1 grad. Lipschitz $f_2 = \lambda\\| \cdot \\|_1$	[Daubechies et al, 2003]
Douglas-Rachford	$f_1 + f_2$	prox_{f_1} prox_{f_2}	[Combettes,Pesquet, 2007]
PPXA	$\sum_i f_i$	prox_{f_i}	[Combettes,Pesquet, 2008]
ADMM	$\sum_i f_i \circ L_i$	prox_{f_i} $(\sum_{i=1}^m L_i^* L_i)^{-1}$	[Eckstein, Yao, 2015]
Chambolle-Pock	$f_1 + f_2 \circ L_2$	prox_{f_1} prox_{f_2}	[Chambolle, Pock, 2011]
Condat-Vũ	$f_1 + f_2 \circ L_2 + f_3$	prox_{f_1} prox_{f_2} f_3 grad. Lipschitz	[Condat, 2013][Vũ, 2013]