STAT 593
Duality / Conjugacy

Joseph Salmon
http://josephsalmon.eu

Télécom Paristech, Institut Mines-Télécom
&
University of Washington, Department of Statistics
(Visiting Assistant Professor)

Outline

Table of Contents

Convex function

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **convex** if for all $x, y \in \mathbb{R}^d$ and $\alpha \in [0, 1]$

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

Proposition

The function f is convex iff $\mathrm{epi}(f)$ is a convex set.

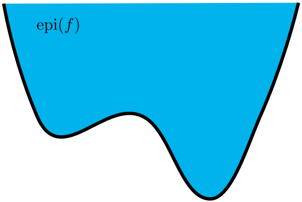
Convex function

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **convex** if for all $x, y \in \mathbb{R}^d$ and $\alpha \in [0, 1]$

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

Proposition

The function f is convex iff $\mathrm{epi}(f)$ is a convex set.



Non-convex

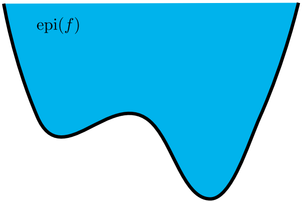
Convex function

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **convex** if for all $x, y \in \mathbb{R}^d$ and $\alpha \in [0, 1]$

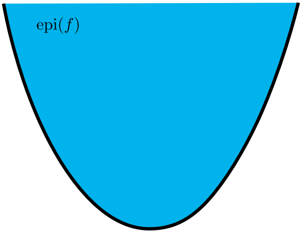
$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Proposition

The function f is convex iff $\mathrm{epi}(f)$ is a convex set.



Non-convex



Convex

Convex hull of a function

Definition

The **convex hull** of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is the function $\text{conv}(f)$ whose epigraph is the convex hull of the epigraph of f :

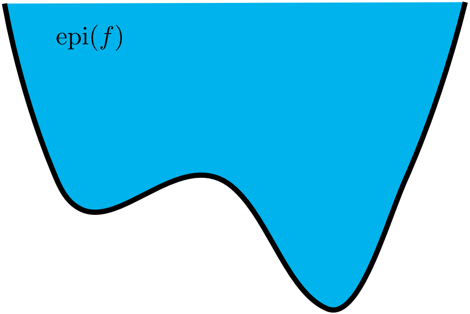
$$\text{epi}(\text{conv}(f)) = \text{conv}(\text{epi}(f))$$

Convex hull of a function

Definition

The **convex hull** of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is the function $\mathrm{conv}(f)$ whose epigraph is the convex hull of the epigraph of f :

$$\mathrm{epi}(\mathrm{conv}(f)) = \mathrm{conv}(\mathrm{epi}(f))$$

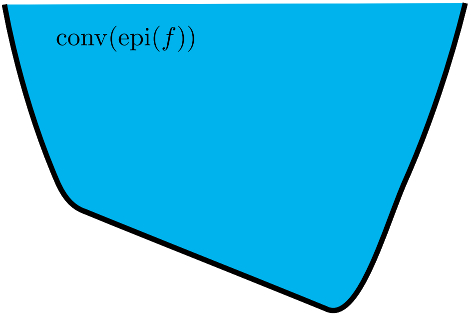


Convex hull of a function

Definition

The **convex hull** of a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is the function $\mathrm{conv}(f)$ whose epigraph is the convex hull of the epigraph of f :

$$\mathrm{epi}(\mathrm{conv}(f)) = \mathrm{conv}(\mathrm{epi}(f))$$



Closed function

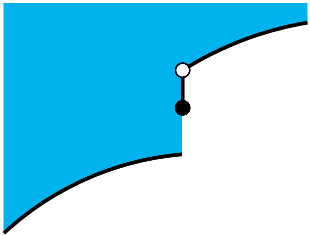
Definition

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **closed** if its epigraph $\mathrm{epi}(f)$ is a (topologically) closed set.

Closed function

Definition

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **closed** if its epigraph $\mathrm{epi}(f)$ is a (topologically) closed set.

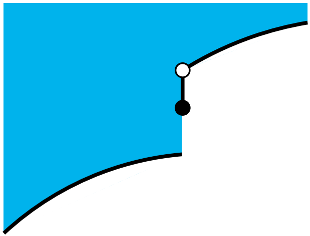


Non-closed function

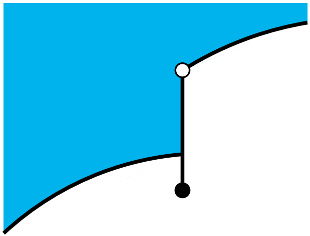
Closed function

Definition

A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is **closed** if its epigraph $\mathrm{epi}(f)$ is a (topologically) closed set.



Non-closed function



Closed function

Closure of a function

Definition

The **closure** $\mathrm{cl}(f)$ of the function f is defined as the function having for epigraph the closure of the epigraph of f :

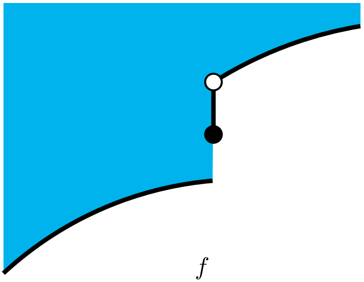
$$\mathrm{epi}(\mathrm{cl}(f)) := \mathrm{cl}(\mathrm{epi}(f))$$

Closure of a function

Definition

The **closure** $\mathrm{cl}(f)$ of the function f is defined as the function having for epigraph the closure of the epigraph of f :

$$\mathrm{epi}(\mathrm{cl}(f)) := \mathrm{cl}(\mathrm{epi}(f))$$

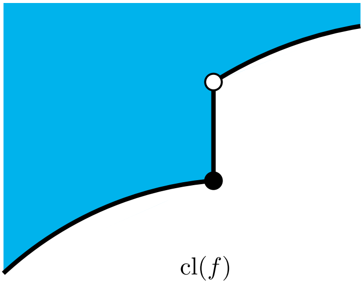


Closure of a function

Definition

The **closure** $\mathrm{cl}(f)$ of the function f is defined as the function having for epigraph the closure of the epigraph of f :

$$\mathrm{epi}(\mathrm{cl}(f)) := \mathrm{cl}(\mathrm{epi}(f))$$



Why closed functions ?

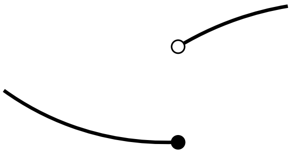
Motivation : close function convenient for minimization

- a closed function, defined on a nonempty closed and bounded set, is bounded below and attains its infimum

Why closed functions ?

Motivation : close function convenient for minimization

- a closed function, defined on a nonempty closed and bounded set, is bounded below and attains its infimum

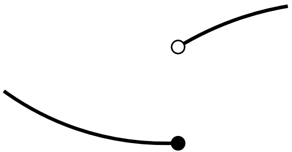


Closed function : inf is reached

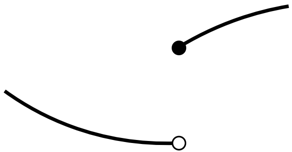
Why closed functions ?

Motivation : close function convenient for minimization

- a closed function, defined on a nonempty closed and bounded set, is bounded below and attains its infimum



Closed function : inf is reached



Non closed function : inf is not reached

Table of Contents

Convex conjugate

Assumption : $f : \mathbb{R}^d \to \overline{\mathbb{R}}$, satisfies $f \not\equiv +\infty$ and there exists an affine function function minorizing f on $\mathbb{R}^d$

Definition

The convex **conjugate** (a.k.a. the **Legendre-Fenchel transform**) of f is the function f^* defined by

$$
\begin{aligned}
f^* : \mathbb{R}^d &\to \overline{\mathbb{R}} \\
s &\mapsto \sup \{ \langle s , x \rangle - f(x) \ : \ x \in \operatorname{dom} f \}
\end{aligned}
$$

Rem: equivalently $\mapsto f^*(s) = -\inf \{ f(x) - \langle s , x \rangle \ : \ x \in \operatorname{dom} f \}$

Rem: f^* is convex as a $\sup$ of convex (affine) functions

https://people.ok.ubc.ca/bauschke/Research/68.pdf

Caracterisation

Legendre caracterisation https://annals.math.princeton.edu/wp-content/uploads/annals-v169-n2-p08.pdf

Intuition

$$\boxed{f^*(s) = \sup\{\langle\, s\,,\, x\,\rangle - f(x) \;:\; x \in \operatorname{dom} f\}}$$

Interpretation : in the smooth case and when ∇f is one-one and $\operatorname{dom} f = \mathbb{R}^d$, let us define :

$$x(s) = \arg\max_{x \in \mathbb{R}^d} \langle\, s\,,\, x\,\rangle - f(x)$$

Then,

$$\begin{aligned} \nabla f(x(s)) = s \iff & (\nabla f)^{-1}(s) = x(s) \\ \iff & \nabla(f^*)(s) = x(s) \end{aligned}$$

All in all :

$$\nabla(f^*) = (\nabla f)^{-1}$$

(1)

(1). J.-B. Hiriart-Urruty et C. Lemaréchal. *Convex analysis and minimization algorithms. II*. T. 306. Berlin : Springer-Verlag, 1993, Chapter X.

Intermission

Movies on conjugacy insights (see associated notebook)

First properties

Remind : $f^*(s) = \sup \{ \langle s , x \rangle - f(x) \ : \ x \in \operatorname{dom} f \}$

Fenchel's inequality :

$$\forall (x, s) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(x) + f^*(s) \geq \langle x , s \rangle$$

Fenchel's equality : equality in the former is equivalent to

$s \in \partial f(x) \iff x \in \operatorname*{arg\,max}_{y \in \mathbb{R}^d} \langle s, y \rangle - f(y)$ (Fermat's rule w.r.t. x)

First properties

Remind : $\quad f^*(s) = \sup\{\langle\, s\,,\, x\,\rangle - f(x)\ :\ x \in \operatorname{dom} f\}$

Fenchel's inequality :

$$\forall (x,s) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(x) + f^*(s) \geq \langle\, x\,,\, s\,\rangle$$

Fenchel's equality : equality in the former is equivalent to

$$\begin{aligned} s \in \partial f(x) &\iff x \in \underset{y \in \mathbb{R}^d}{\arg\max} \langle s, y\rangle - f(y) && \text{(Fermat's rule w.r.t. } x\text{)} \\ &\iff f(x) + f^*(s) = \langle\, x\,,\, s\,\rangle && \text{(maximum achieved)} \end{aligned}$$

First properties

Remind : $f^*(s) = \sup \{ \langle s , x \rangle - f(x) \ : \ x \in \operatorname{dom} f \}$

Fenchel's inequality :

$$\forall (x, s) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(x) + f^*(s) \geq \langle x , s \rangle$$

Fenchel's equality : equality in the former is equivalent to

$$\begin{aligned} s \in \partial f(x) &\iff x \in \arg\max_{y \in \mathbb{R}^d} \langle s, y \rangle - f(y) && \text{(Fermat's rule w.r.t. } x\text{)} \\ &\iff f(x) + f^*(s) = \langle x , s \rangle && \text{(maximum achieved)} \\ &\iff s \in \arg\max_{t \in \mathbb{R}^d} \langle t, x \rangle - f^*(t) && \text{(maximum achieved)} \end{aligned}$$

First properties

Remind : $\quad f^*(s) = \sup\{\langle\, s\,,\, x\,\rangle - f(x) \ :\ x \in \operatorname{dom} f\}$

Fenchel's inequality :

$$\forall (x,s) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(x) + f^*(s) \geq \langle\, x\,,\, s\,\rangle$$

Fenchel's equality : equality in the former is equivalent to

$$\begin{aligned}
s \in \partial f(x) &\iff x \in \operatorname*{arg\,max}_{y \in \mathbb{R}^d} \langle s, y\rangle - f(y) && \text{(Fermat's rule w.r.t. } x\text{)}\\
&\iff f(x) + f^*(s) = \langle\, x\,,\, s\,\rangle && \text{(maximum achieved)}\\
&\iff s \in \operatorname*{arg\,max}_{t \in \mathbb{R}^d} \langle t, x\rangle - f^*(t) && \text{(maximum achieved)}\\
s \in \partial f(x) &\iff x \in \partial f^*(s) && \text{(Fermat's rule w.r.t. } s\text{)}
\end{aligned}$$

Bi-conjugate / closed convex hull

$$
\begin{aligned}
f^{**}(x) &= \sup_s \{ \langle\, s \,,\, x \,\rangle - f^*(s) \} \\
&= \sup_s \Big\{ \langle\, s \,,\, x \,\rangle - \sup_z \{ \langle\, s \,,\, z \,\rangle - f(z) \} \Big\}
\end{aligned}
$$

Bi-conjugate / closed convex hull

$$
\begin{aligned}
f^{**}(x) &= \sup_s \{ \langle\, s \,,\, x \,\rangle - f^*(s) \} \\
&= \sup_s \left\{ \langle\, s \,,\, x \,\rangle - \sup_z \{ \langle\, s \,,\, z \,\rangle - f(z) \} \right\} \\
&= \sup_s \left\{ \langle\, s \,,\, x \,\rangle - r | r = \sup_z \{ \langle\, s \,,\, z \,\rangle - f(z) \} \right\}
\end{aligned}
$$

Bi-conjugate / closed convex hull

$$\begin{aligned}
f^{**}(x) &= \sup_s \{\langle\, s\,,\, x\,\rangle - f^*(s)\} \\
&= \sup_s \Big\{\, \langle\, s\,,\, x\,\rangle - \sup_z \{\langle\, s\,,\, z\,\rangle - f(z)\}\Big\} \\
&= \sup_s \Big\{\, \langle\, s\,,\, x\,\rangle - r | r = \sup_z \{\langle\, s\,,\, z\,\rangle - f(z)\}\Big\} \\
&= \sup_{s,r} \Big\{\, \langle\, s\,,\, x\,\rangle - r | r \geq \sup_z \{\langle\, s\,,\, z\,\rangle - f(z)\}\Big\}
\end{aligned}$$

Bi-conjugate / closed convex hull

$$
\begin{aligned}
f^{**}(x) &= \sup_s \{\langle\, s \,,\, x \,\rangle - f^*(s)\} \\
&= \sup_s \left\{ \langle\, s \,,\, x \,\rangle - \sup_z \{\langle\, s \,,\, z \,\rangle - f(z)\} \right\} \\
&= \sup_s \left\{ \langle\, s \,,\, x \,\rangle - r | r = \sup_z \{\langle\, s \,,\, z \,\rangle - f(z)\} \right\} \\
&= \sup_{s,r} \left\{ \langle\, s \,,\, x \,\rangle - r | r \geq \sup_z \{\langle\, s \,,\, z \,\rangle - f(z)\} \right\} \\
&= \sup_{s,r} \left\{ \langle\, s \,,\, x \,\rangle - r | r \geq \langle\, s \,,\, z \,\rangle - f(z), \forall z \right\}
\end{aligned}
$$

Bi-conjugate / closed convex hull

$$
\begin{aligned}
f^{**}(x) &= \sup_s \{\langle s\,,\, x \rangle - f^*(s)\} \\
&= \sup_s \Big\{ \langle s\,,\, x \rangle - \sup_z \{\langle s\,,\, z \rangle - f(z)\} \Big\} \\
&= \sup_s \Big\{ \langle s\,,\, x \rangle - r | r = \sup_z \{\langle s\,,\, z \rangle - f(z)\} \Big\} \\
&= \sup_{s,r} \Big\{ \langle s\,,\, x \rangle - r | r \geq \sup_z \{\langle s\,,\, z \rangle - f(z)\} \Big\} \\
&= \sup_{s,r} \Big\{ \langle s\,,\, x \rangle - r | r \geq \langle s\,,\, z \rangle - f(z), \forall z \Big\} \\
&= \sup_{s,r} \Big\{ \langle s\,,\, x \rangle - r | f(z) \geq \langle s\,,\, z \rangle - r, \forall z \Big\}
\end{aligned}
$$

Bi-conjugate / closed convex hull

$$
\begin{aligned}
f^{**}(x) &= \sup_s \{\langle s , x \rangle - f^*(s)\} \\
&= \sup_s \Big\{ \langle s , x \rangle - \sup_z \{\langle s , z \rangle - f(z)\} \Big\} \\
&= \sup_s \Big\{ \langle s , x \rangle - r | r = \sup_z \{\langle s , z \rangle - f(z)\} \Big\} \\
&= \sup_{s,r} \Big\{ \langle s , x \rangle - r | r \geq \sup_z \{\langle s , z \rangle - f(z)\} \Big\} \\
&= \sup_{s,r} \Big\{ \langle s , x \rangle - r | r \geq \langle s , z \rangle - f(z), \forall z \Big\} \\
&= \sup_{s,r} \Big\{ \langle s , x \rangle - r | f(z) \geq \langle s , z \rangle - r, \forall z \Big\}
\end{aligned}
$$

Hence : f^{**} is the **closed convex hull** of f (whose epigraph is obtained as the intersection of all supporting half spaces)

Standard functions

- $\boxed{f = \frac{1}{2} \|\cdot\|_2^2 \Rightarrow f^* = \frac{1}{2} \|\cdot\|_2^2}$; it is the only [(2)] f s.t. $f^* = f$

(2). J.-J. Moreau. "Proximité et dualité dans un espace hilbertien". In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.

Standard functions

- $\boxed{f = \frac{1}{2}\left\|\cdot\right\|_2^2 \Rightarrow f^* = \frac{1}{2}\left\|\cdot\right\|_2^2}$; it is the only [(2)] f s.t. $f^* = f$
- Let $p, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then $\boxed{f = \frac{1}{p}\left\|\cdot\right\|_p^p \Rightarrow f^* = \frac{1}{q}\left\|\cdot\right\|_q^q}$

(2). J.-J. Moreau. "Proximité et dualité dans un espace hilbertien". In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.

Standard functions

- $\boxed{f = \frac{1}{2}\|\cdot\|_2^2 \Rightarrow f^* = \frac{1}{2}\|\cdot\|_2^2}$; it is the only [(2)] f s.t. $f^* = f$

- Let $p, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then $\boxed{f = \frac{1}{p}\|\cdot\|_p^p \Rightarrow f^* = \frac{1}{q}\|\cdot\|_q^q}$

- Let $\iota_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C} \\ +\infty & \text{if } x \notin \mathcal{C} \end{cases}$

  Constant case : $\forall c \in \mathbb{R}^d$, $\boxed{f \equiv c \Rightarrow f^* = \iota_{\{-c\}}}$

(2). J.-J. Moreau. "Proximité et dualité dans un espace hilbertien". In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.

Standard functions

- $\boxed{f = \frac{1}{2}\|\cdot\|_2^2 \Rightarrow f^* = \frac{1}{2}\|\cdot\|_2^2}$; it is the only [(2)] f s.t. $f^* = f$
- Let $p, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then $\boxed{f = \frac{1}{p}\|\cdot\|_p^p \Rightarrow f^* = \frac{1}{q}\|\cdot\|_q^q}$
- Let $\iota_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C} \\ +\infty & \text{if } x \notin \mathcal{C} \end{cases}$

  Constant case : $\forall c \in \mathbb{R}^d$, $\boxed{f \equiv c \Rightarrow f^* = \iota_{\{-c\}}}$
- Let $\mathcal{B}_{\|\cdot\|}(0,1)$ be a unit ball associated to a norm $\|\cdot\|$ and $\|\cdot\|_*$ is the dual norm associated to it where $\|x\|_* = \sup_{y:\|y\|\leq 1} \langle x , y \rangle$.

  Then, $\boxed{f = \iota_{\mathcal{B}_{\|\cdot\|}(0,1)} \Rightarrow f^* = \|\cdot\|_*}$

(2). J.-J. Moreau. "Proximité et dualité dans un espace hilbertien". In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.

Simple properties

$$
\begin{array}{rcr}
\forall \alpha > 0, & (\alpha f)^* = \alpha f^*\left(\frac{\cdot}{\alpha}\right) & \text{(Scaling)} \\
\forall \alpha > 0, & \left(\alpha f\left(\frac{\cdot}{\alpha}\right)\right)^* = \alpha f^* & \text{(Scaling)} \\
\forall x \in \mathbb{R}^d, & \Big(f(\cdot - x)\Big)^* = f^* + \langle\, x \,,\, \cdot\,\rangle & \text{(Shifting)} \\
\forall \tau \in \mathbb{R}, & (f + \tau)^* = f^* - \tau & \text{(Shifting)} \\
\forall x \in \mathbb{R}^d, & \big(f + \langle\, x \,,\, \cdot\,\rangle\,\big)^* = f^*(\cdot - x) & \text{(Shifting)} \\
\forall L \in \mathbb{R}^{d\times d}, & (f \circ L)^* = f^* \circ (L^{-1})^\top & \text{(Linear composition)} \\
 & & \text{(when } L \text{ invertible)}
\end{array}
$$

Table of Contents

Fenchel duality [3]

Optimization : $\quad \arg\min_{x \in \mathbb{R}^d} \mathcal{P}(x)$ where $\mathcal{P}(x) = f(Lx) + g(x)$

- $f : \mathbb{R}^n \to \mathbb{R}$ closed convex
- $g : \mathbb{R}^d \to \mathbb{R}$ closed convex
- $L : n \times d$ matrix

Theorem

$$\sup_s \{-f^*(s) - g^*(-L^\top s)\} \leq \inf_x \{f(Lx) + g(x)\}$$

Moreover, under mild assumptions, equality holds (**strong duality**)

(3). H. H. Bauschke et P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. New York : Springer, 2011, p. xvi+468.

Duality gap / vocabulary

Primal function : $\mathcal{P}(x) = f(Lx) + g(x)$ (to **minimize**)

Primal solution : $x^\star \in \arg\min_x \mathcal{P}(x)$

Dual function : $\mathcal{D}(s) = -f^*(s) - g^*(-L^\top s)$ (to **maximize**)

Dual solution : $s^\star \in \arg\max_s \mathcal{D}(s)$

Duality gap : $\Delta(x, s) = \mathcal{P}(x) - \mathcal{D}(s)$

Theorem

$$\forall (x, s), \quad 0 \leq \mathcal{P}(x) - \mathcal{P}(x^\star) \leq \Delta(x, s)$$

Also $\Delta(x, s) = 0$ implies that $\mathcal{P}(x) = \mathcal{P}(x^\star)$ and $\mathcal{P}(s) - \mathcal{P}(s^\star)$

Duality gap for stopping algorithms

$$\Delta(x, s) = \mathcal{P}(x) - \mathcal{D}(s)$$

Stopping criterion for iterative solver :
Fix $\varepsilon > 0$ (small). Whenever one has point x, s then

$$\boxed{\Delta(x, s) \leq \varepsilon \Rightarrow \mathcal{P}(x) - \mathcal{P}(x^\star) \leq \varepsilon}$$

Hence, stopping with duality gap criterion leads to a precise statement on the sub-optimality of the solution obtained.

Rem: this is a more precise criterion that choices like

$$\frac{\mathcal{P}(x^{t+1}) - \mathcal{P}(x_t)}{\mathcal{P}(x^t)} \leq \varepsilon \quad \text{or} \quad |\nabla \mathcal{P}(x^t)| \leq \varepsilon$$

Example : duality gap for the standard median

$$\mathrm{Med}_n(\mathbf{x}) \in \underset{\theta \in \mathbb{R}}{\arg\min} \left\| \theta \mathbf{1}_n - \mathbf{x} \right\|_1 = \mathcal{P}(\theta)$$

where $\mathbf{1}_n = (1, \dots, 1)^\top \in \mathbb{R}^n$

$$f = \left\| \cdot - \mathbf{x} \right\|_1, \qquad g = 0$$

$$f^* = \iota_{\mathcal{B}_{\|\cdot\|_\infty}(0,1)} + \langle \mathbf{x}, \cdot \rangle, \qquad g^* = \iota_{\{0\}}$$

Hence, $\mathcal{D}(s) = \langle s, \mathbf{x} \rangle + \iota_{\{s \in \mathbb{R}^n : \|s\|_\infty \leq 1\}} + \iota_{\{s \in \mathbb{R}^n : \mathbf{1}_n^\top s = 0\}}$

Rem: with iterate θ^t aiming at solving the primal problem, following would create dual feasible points :

$$s^t = \frac{\theta^t \mathbf{1}_n - \mathbf{x} - \mathrm{Ave}(\theta^t \mathbf{1}_n - \mathbf{x})}{\left\| \theta^t \mathbf{1}_n - \mathbf{x} - \mathrm{Ave}(\theta^t \mathbf{1}_n - \mathbf{x}) \right\|_\infty}$$

Example : duality gap for Lasso

Lasso objective : $\boxed{\mathcal{P}(\theta) = \frac{1}{2}\|X\theta - y\|_2^2 + \lambda\|\theta\|_1}$

- $f(z) = \frac{1}{2}\|z - y\|_2^2$; $f^*(s) = \frac{1}{2}\|s\|_2^2 + \langle s, y\rangle$ (use gradient)
- $g(\theta) = \lambda\|\theta\|_1$; $g^*(s) = \iota_{\{s, \|s\|_\infty \leq \lambda\}}$ (ℓ_∞ ball indicator)

- Duality gap :

$$
\begin{aligned}
\Delta(\theta, s) =& \mathcal{P}(\theta) + f^*(s) + g^*(-X^\top s) \\
=& \mathcal{P}(\theta) + \frac{1}{2}\|s\|_2^2 + \langle s, y\rangle
\end{aligned}
$$

as soon as $\|X^\top s\|_\infty \leq \lambda$, o.w. the bound is $+\infty$ (useless)

More references

- Material mostly inspired by the lecture notes by Pontus Giselsson : `http://www.control.lth.se/ls-convex-2015/`
- Further reading on an intrinsic property of Fenchel duality : `https://annals.math.princeton.edu/wp-content/uploads/annals-v169-n2-p08.pdf`

References I

- Bauschke, H. H. et P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. New York : Springer, 2011, p. xvi+468.
- Hiriart-Urruty, J.-B. et C. Lemaréchal. *Convex analysis and minimization algorithms. II*. T. 306. Berlin : Springer-Verlag, 1993.
- Moreau, J.-J. "Proximité et dualité dans un espace hilbertien". In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.