$\mathcal{H}_0$-Conjugacies and the ℓ_0 Pseudonorm

Jean-Philippe Chancelier, *Michel De Lara*

Cermics
École nationale des ponts et chaussées
IP Paris, France

PGMO Days, 19-20 November 2024

Support and the ℓ_0 pseudonorm

Let $n \in \mathbb{N}^*$ be a natural number and

$$[\![0, n]\!] = \{0, 1, \ldots, n\} \ , \ \ [\![1, n]\!] = \{1, \ldots, n\}$$

For any vector $x \in \mathbb{R}^n$, we define

- its support by $\mathrm{supp}(x) = \{j \in [\![1, n]\!] \mid x_j \neq 0\}$

$$\mathrm{supp}((0, *, 0, *, *, 0)) = \{2, 4, 5\} \subset [\![1, 6]\!]$$

- its ℓ_0 pseudonorm by

$$\ell_0(x) = \overbrace{|\mathrm{supp}(x)|}^{\text{cardinality}} = \overbrace{\sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq 0\}}}^{\substack{\text{number of} \\ \text{nonzero entries}}}$$

$$\ell_0\big((0, *, 0, *, *, 0)\big) = |\{2, 4, 5\}| = 3 \in [\![0, 6]\!]$$

The ℓ_0 pseudonorm is ~~not~~ a ~~norm~~

The function ℓ_0 pseudonorm $: \mathbb{R}^n \to [\![0, n]\!]$
satisfies 3 out of 4 axioms of a norm

- $\ell_0(x) \geq 0$ ✓
- $\Big(\ell_0(x) = 0 \iff x = 0 \Big)$ ✓
- $\ell_0(x + x') \leq \ell_0(x) + \ell_0(x')$ ✓
- But... instead of absolute 1-homogeneity,
  it is absolute 0-homogeneity that holds true

$$\ell_0(\lambda x) = \ell_0(x) \ , \ \ \forall \lambda \neq 0$$

$$\mathrm{supp}(\lambda x) = \mathrm{supp}(x) \ , \ \ \forall \lambda \neq 0$$

The ℓ_0 pseudonorm is used in typical sparse optimization problems

- Spark of a matrix A

$$\mathrm{spark}(A) = \min \left\{ \ell_0(x) \,\middle|\, Ax = 0 \,, \;\; x \neq 0 \right\}$$

- Compressed sensing: recovery of a sparse signal $x \in \mathbb{R}^n$ from a measurement $b = Ax$

$$\min_{\substack{x \in \mathbb{R}^n \\ Ax = b}} \ell_0(x)$$

- Least squares sparse regression (best subset selection):

$$\text{for } k \in [\![1, n]\!] \qquad \min_{\substack{x \in \mathbb{R}^n \\ \ell_0(x) \leq k}} \|Ax - b\|^2$$

"explaining" the output b by at most k components of x

SNAPSHOTS OF OUR MAIN RESULTS

A menagerie of couplings (and two more)

Coupling c $c(u,v)$	c-convex functions $f^{cc'} = f$
$\langle u\,,\,v \rangle$	closed convex $\quad f^{\star\star'} = f$
$u(v)$, u continuous	lower semicontinuous
$\log\langle u\,,\,v \rangle_+$	$\log \circ$ sublinear
$-N\left\Vert u-v \right\Vert^{\alpha}$, $0 < \alpha \leq 1$	α-Hölder continuous with constant N
$\min_{i,v_i>0} u_i v_i$	increasing and convex-along-rays
Capra $\quad ¢(u,v) = \langle \frac{u}{\Vert u \Vert}\,,\,v \rangle$	$\ell_0^{¢¢'} = \ell_0$
$\mathcal{H}_0$	convex $\circ$ 0-homogeneous

Fenchel conjugacy ($\star$) *versus* E-Capra conjugacy ($¢$) for the ℓ_0 pseudonorm

- Fenchel conjugacy ($\star$)

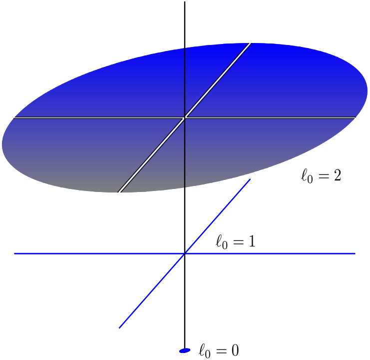
$$\ell_0^{\star\star'} = 0$$

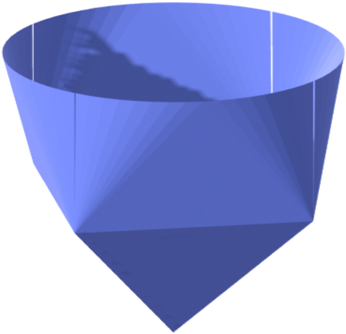
- E-Capra conjugacy ($¢$)

$$\boxed{\ell_0^{¢¢'} = \ell_0}$$

[Chancelier and De Lara, 2021]

The ℓ_0 pseudonorm coincides, on the unit sphere, with the proper convex lower semicontinuous ℓ_0-cup function $\mathcal{L}_0 = \ell_0^{¢\star\prime}$





Variational formulas for the ℓ_0 pseudonorm

Proposition

[Chancelier and De Lara, 2021]

$$\ell_0(x) = \frac{1}{\|x\|_2} \min_{\substack{x^{(1)}\in\mathbb{R}^n,\dots,x^{(d)}\in\mathbb{R}^n \\ \sum_{l=1}^d \left\|x^{(l)}\right\|_{(l)}^{\top\star} \leq \|x\|_2 \\ \sum_{l=1}^d x^{(l)} = x}} \sum_{l=1}^{d} l \left\|x^{(l)}\right\|_{(l)}^{\top\star}, \quad \forall x \in \mathbb{R}^n$$

$$\ell_0(x) = \sup_{y\in\mathbb{R}^n} \inf_{l=1,\dots,d} \left(\frac{\langle x \mid y \rangle}{\|x\|_2} - \left[\|y\|_{2,l}^{\top} - l \right]_+ \right), \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

Talk outline

$\mathcal{H}_0$-conjugacies [10 min]

- Couplings and $\mathcal{H}_0$-couplings
- Capra-couplings and radial $\mathcal{H}_0$-couplings
- Conjugacies and $\mathcal{H}_0$-conjugacies
- Biconjugates and duality, $\mathcal{H}_0$-biconjugates
- Subdifferentials and $\mathcal{H}_0$-subdifferentials

Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0 [10 min]

- Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0
- Orthant-strict monotonicity and Capra-convexity of ℓ_0
- Orthant-strict monotonicity

Conclusion

Outline of the presentation

$\mathcal{H}_0$-conjugacies [10 min]

Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0 [10 min]

Conclusion

Outline of the presentation

Couplings

Motivation: Legendre transform and Fenchel conjugacy in convex analysis

Definition

Two vector spaces $\mathcal{X}$ and $\mathcal{Y}$, paired by a bilinear form $\langle\ ,\ \rangle$, (in the sense of convex analysis [Rockafellar, 1974, p. 13])) give rise to the classic Fenchel conjugacy

$$f \in \overline{\mathbb{R}}^{\mathcal{X}} \mapsto f^{\star} \in \overline{\mathbb{R}}^{\mathcal{Y}}$$

given by the Legendre transform

$$f^{\star}(y) = \sup_{x \in \mathcal{X}} \Big(\underbrace{\langle x , y \rangle}_{\text{coupling}} + \big(-f(x)\big) \Big) , \ \forall y \in \mathcal{Y}$$

Coupling function between sets

- Let be given two sets $\mathcal{U}$ ("primal") and $\mathcal{V}$ ("dual")
  not necessarily paired vector spaces (nodes and arcs, etc.)
- We consider a coupling function

$$c : \mathcal{U} \times \mathcal{V} \to \overline{\mathbb{R}}$$

We also use the notation $\mathcal{U} \overset{c}{\leftrightarrow} \mathcal{V}$ for a coupling

[Moreau, 1966-1967, 1970]

$\mathcal{H}_0$-couplings

Background on homogeneous functions and mappings

Let $\mathcal{X}$ be a vector space
We denote $\mathbb{R}_+ = [0, +\infty[$, $\mathbb{R}_{++} =]0, +\infty[$
A set $C \subset \mathcal{X}$ is a cone if $\mathbb{R}_{++} C \subset C$

Definition

Let $\alpha \in \mathbb{R}$ and let $C \subset \mathcal{X}$ be a nonempty cone of $\mathcal{X}$
We say that a function $f \colon C \to \overline{\mathbb{R}}$ or a mapping $f \colon C \to \mathcal{X}$ is

1. (strictly positively) α-homogeneous (on the cone C) if

$$f(\lambda x) = \lambda^{\alpha} f(x) \; , \; \forall \lambda \in \mathbb{R}_{++} \; , \; \forall x \in C$$

2. absolutely α-homogeneous (on the cone C) if

$$f(\lambda x) = |\lambda|^{\alpha} f(x) \; , \; \forall \lambda \in \mathbb{R} \setminus \{0\} \; , \; \forall x \in C$$

We use the convention $\lambda^0 = 1 \; , \; \forall \lambda \in \mathbb{R}_{++}$

Examples of homogeneous functions

Examples

- The ℓ_0 pseudonorm is absolutely 0-homogeneous on $\mathbb{R}^n$
- Any norm $\|\cdot\|$ on $\mathcal{X}$ is absolutely 1-homogeneous on $\mathcal{X}$, and the radial projection $R_{\|\cdot\|}$, defined by

$$R_{\|\cdot\|} : \mathcal{X} \to \mathcal{X} \;, \quad R_{\|\cdot\|}(x) = \begin{cases} \dfrac{x}{\|x\|} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

  is a 0-homogeneous mapping

Definition of $\mathcal{H}_0$-couplings

Definition

Let two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ be paired by a bilinear form $\langle \ , \rangle$, and let $C \subset \mathcal{X}$ be a nonempty cone of $\mathcal{X}$

A $\mathcal{H}_0$-coupling[a] between C and $\mathcal{Y}$ is a function $c \colon C \times \mathcal{Y} \to \mathbb{R}$ which is

1. 0-homogeneous in the first variable (on the cone C)
2. linear continuous in the second variable (on the vector space $\mathcal{Y}$)

[a] We thank Prof. David L. Donoho for his suggestion to call such couplings $\mathcal{H}_0$, standing for homogeneous of degree 0

$\mathcal{H}_0$-couplings are left-sided 0-homogeneous and right-sided linear

The $\mathcal{H}_0$-couplings are a special case of a one-sided linear couplings [Chancelier and De Lara, 2021]

Proposition

Let two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ be paired by a bilinear form $\langle \ , \ \rangle$, and let $C \subset \mathcal{X}$ be a nonempty cone of $\mathcal{X}$

The function $c \colon C \times \mathcal{Y} \to \mathbb{R}$ is a $\mathcal{H}_0$-coupling
if and only if
there exists a 0-homogeneous mapping $\varrho \colon C \to \mathcal{X}$

such that
$$\boxed{c(x, y) = \langle \varrho(x), \, y \rangle} \ , \quad \forall x \in C \ , \quad \forall y \in \mathcal{Y}$$

and we denote
$$\boxed{c = \star_\varrho}$$

Outline of the presentation

Capra $\subsetneq$ radial $\mathcal{H}_0$ $\subsetneq$ $\mathcal{H}_0$

norms $\subsetneq$ 1-homogeneous functions

Examples of $\mathcal{H}_0$-couplings:
Capra-couplings Chancelier and De Lara [2022a]

Constant Along Primal RAys (Capra) couplings are $\mathcal{H}_0$-couplings

[Chancelier and De Lara, 2022a, Definition 8]

Definition

Let $\|\cdot\|$ be a (source) norm on $\mathcal{X}$

The Capra coupling (Capra) $\mathcal{X} \overset{¢}{\longleftrightarrow} \mathcal{Y}$ is given by

$$\forall y \in \mathcal{Y} \; , \; \begin{cases} ¢(x, y) & = \dfrac{\langle x \, , \, y \rangle}{\|x\|} \; , \; \forall x \in \mathcal{X} \backslash \{0\} \\ ¢(0, y) & = 0 \end{cases}$$

The coupling Capra is a $\mathcal{H}_0$-coupling between $\mathcal{X}$ and $\mathcal{Y}$ itself, as $¢ = \star_{R_{\|\cdot\|}}$ since

$$¢(x, y) = \langle \underbrace{R_{\|\cdot\|}(x)}_{\text{radial projection}} \, , \, y \rangle \; , \; \forall x \in \mathcal{X} \; , \; \forall y \in \mathcal{X}$$

Examples of $\mathcal{H}_0$-couplings:
radial $\mathcal{H}_0$-couplings

For any function $f : \mathcal{X} \to \overline{\mathbb{R}}$, we introduce

the level sets	$f^{\leq r} = \left\{ x \in \mathcal{X} \,\middle|\, f(x) \leq r \right\} , \ \forall r \in \overline{\mathbb{R}}$
the strict level sets	$f^{< r} = \left\{ x \in \mathcal{X} \,\middle|\, f(x) < r \right\} , \ \forall r \in \overline{\mathbb{R}}$
and the level curves	$f^{= r} = \left\{ x \in \mathcal{X} \,\middle|\, f(x) = r \right\} , \ \forall r \in \overline{\mathbb{R}}$

Radial $\mathcal{H}_0$-couplings

Definition

We call normalization function a 1-homogeneous nonnegative function $\nu \colon \mathcal{X} \to [0, +\infty]$ such that $\nu^{=1} \neq \emptyset$

With any normalization function ν, we associate the primal normalization mapping $\varrho_\nu \colon \mathcal{X} \setminus \nu^{=0} \to \mathcal{X}$

$$\text{defined by} \qquad \mathcal{X} \setminus \nu^{=0} \ni x \mapsto \boxed{\varrho_\nu(x) = \frac{1}{\nu(x)} x}$$

as well as the radial $\mathcal{H}_0$-coupling $\star_{\varrho_\nu}$, between $\mathcal{X} \setminus \nu^{=0}$ and $\mathcal{Y}$,

$$\text{as the function} \qquad \star_{\varrho_\nu} \colon (\mathcal{X} \setminus \nu^{=0}) \times \mathcal{Y} \to \mathbb{R}$$

$$\text{defined by} \qquad \star_{\varrho_\nu}(x, y) = \boxed{\frac{1}{\nu(x)} \langle x, y \rangle = \langle \varrho_\nu(x), y \rangle}$$

$$\forall x \in \mathcal{X} \setminus \nu^{=0}, \ \forall y \in \mathcal{Y}$$

Capra-couplings are radial $\mathcal{H}_0$-couplings

Proposition

Regarding classes of couplings, we have the following strict inclusions

$$\mathrm{Capra} \subsetneq \mathrm{radial}\ \mathcal{H}_0 \subsetneq \mathcal{H}_0$$

For the proof that $\mathrm{Capra} \subsetneq \mathrm{radial}\ \mathcal{H}_0$, we will need the following

Definition

The Minkowski functional associated with the subset $X \subset \mathcal{X}$ is the function $m_X : \mathcal{X} \to [0, +\infty]$ defined by
(with the convention that $\inf \emptyset = +\infty$)

$$m_X(x) = \inf \left\{ \lambda > 0 \,\middle|\, x \in \lambda X \right\} , \ \forall x \in \mathcal{X}$$

The Minkowski functional m_X is a
1-homogeneous (nonnegative) function

Proof that Capra-couplings are radial $\mathcal{H}_0$-couplings

Proof.

- Let S denote the unit sphere of the norm $\|\cdot\|$ and let $\nu = m_S$ be the Minkowski functional of S
- Thus defined, ν is a normalization function, as an easy calculation shows that

$$\nu = m_S = \|\cdot\| + \iota_{\mathcal{X}\setminus\{0\}} \ , \quad \nu^{=0} = \emptyset$$

$$\text{(Capra) } ¢ = \star_{\varrho_\nu} \text{ (radial } \mathcal{H}_0)$$

□

Notice that the norm $\|\cdot\|$ is also a normalization function, but that, as $\|\cdot\|^{=0} = \{0\}$, $\|\cdot\|$ leads to the $\mathcal{H}_0$-coupling between $\mathcal{X} \setminus \{0\}$ and $\mathcal{Y}$ defined by $\star_{\varrho_{\|\cdot\|}}(x, y) = \langle x, y \rangle / \|x\|$ when $x \neq 0$, which is not a Capra-coupling — as Capra-couplings are defined between $\mathcal{X}$ and $\mathcal{Y}$ — so that $¢ \neq \star_{\varrho_{\|\cdot\|}}$

Outline of the presentation

Fenchel-Moreau conjugate of a function

$$f \in \overline{\mathbb{R}}^{\mathcal{U}} \mapsto f^c \in \overline{\mathbb{R}}^{\mathcal{V}}$$

Definition

The c-Fenchel-Moreau conjugate $f^c : \mathcal{V} \to \overline{\mathbb{R}}$ of a function $f : \mathcal{U} \to \overline{\mathbb{R}}$ is defined by

$$f^c(v) = \sup_{u \in \mathcal{U}} \Big(c(u, v) \stackrel{.}{+} \big(-f(u) \big) \Big) \ , \quad \forall v \in \mathcal{V}$$

We use the Moreau lower and upper additions on $\overline{\mathbb{R}}$
that extend the usual addition with

$$(+\infty) \stackrel{.}{+} (-\infty) = (-\infty) \stackrel{.}{+} (+\infty) = -\infty$$
$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = +\infty$$

Blanket assumptions

1. Let two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ be paired by a bilinear form $\langle\, , \rangle$, and let $C \subset \mathcal{X}$ be a nonempty cone of $\mathcal{X}$
2. Let $\star_{\varrho}$ be a $\mathcal{H}_0$-coupling
   with associated 0-homogeneous mapping $\varrho \colon C \to \mathcal{X}$,
   satisfying
   $$\mathrm{im}\varrho = \varrho(C) \subset C$$
3. Let $f \colon C \to \overline{\mathbb{R}}$ be a function satisfying
   $$f \circ \varrho = f$$
   (hence, necessarily, $f \colon C \to \overline{\mathbb{R}}$ is 0-homogeneous)

$\mathcal{H}_0$-conjugates of 0-homogeneous functions

Proposition

Under the blanket assumptions, we have that

$$\underbrace{f^{\star_\varrho}}_{\star_\varrho-\text{conjugate}} = \underbrace{\left(f \dot{+} \iota_{\mathrm{im}\varrho}\right)^{\star}}_{\text{Fenchel conjugate}}$$

Capra-conjugate of the ℓ_0 pseudonorm

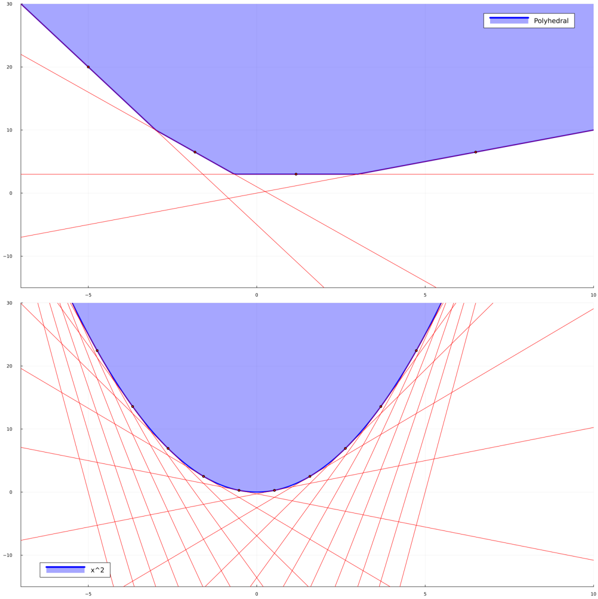
$S \subset \mathbb{R}^n$ unit sphere, $\nu = m_S$, $\cent = \star_{\varrho_\nu}$, $\varrho_\nu(x) = \frac{1}{\nu(x)} x$, $\forall x \in \mathbb{R}^n$,
$\mathrm{im}\varrho_\nu = S \cup \{0\}$ $(1/\nu(0) = 0)$, $C = \mathbb{R}^n$, $f = \ell_0$

$$\begin{aligned}
\ell_0^{\cent}(y) = \ell_0^{\star_{\varrho_\nu}}(y) &= \left(\ell_0 \dotplus \iota_{\mathrm{im}\varrho_\nu}\right)^\star \\
&= \left(\ell_0 \dotplus \iota_{S\cup\{0\}}\right)^\star \\
&= \sup\Big\{0, \sup_{s\in S}\big\{\langle s \mid y\rangle - \ell_0(s)\big\}\Big\} \\
&= \sup\Big\{0, \sup_{i\in[\![1,n]\!]}\big\{ \overbrace{\underbrace{\sup_{\substack{s\in S \\ \ell_0(s)=i}} \langle s \mid y\rangle}_{\|y\|^{\mathcal{R}}_{(i)}}}^{\text{coordinate-}i\text{ norm}} - i\big\}\Big\} \\
&= \sup_{i\in[\![1,n]\!]} \left[\|y\|^{\mathcal{R}}_{(i)} - i\right]_+
\end{aligned}$$

[Chancelier and De Lara, 2021, 2022a]

Outline of the presentation

Motivation: duality in convex analysis



Reverse coupling and Fenchel-Moreau biconjugate

With the coupling c, we associate the reverse coupling c'

$$c' : \mathcal{V} \times \mathcal{U} \to \overline{\mathbb{R}} \ , \ \ c'(v,u) = c(u,v) \ , \ \ \forall (v,u) \in \mathcal{V} \times \mathcal{U}$$

$$f \in \overline{\mathbb{R}}^{\mathcal{U}} \mapsto f^{c} \in \overline{\mathbb{R}}^{\mathcal{V}}$$

$$g \in \overline{\mathbb{R}}^{\mathcal{V}} \mapsto g^{c'} \in \overline{\mathbb{R}}^{\mathcal{U}}$$

Reverse coupling and Fenchel-Moreau biconjugate

With the coupling c, we associate the reverse coupling c'

$$c' : \mathcal{V} \times \mathcal{U} \to \overline{\mathbb{R}} \ , \ \ c'(v, u) = c(u, v) \ , \ \ \forall (v, u) \in \mathcal{V} \times \mathcal{U}$$

$$f \in \overline{\mathbb{R}}^{\mathcal{U}} \mapsto f^c \in \overline{\mathbb{R}}^{\mathcal{V}}$$

$$g \in \overline{\mathbb{R}}^{\mathcal{V}} \mapsto g^{c'} \in \overline{\mathbb{R}}^{\mathcal{U}}$$

$$g^{c'}(u) = \sup_{v \in \mathcal{V}} \Big(c(u, v) \dot{+} \big(-g(v)\big) \Big) \ , \ \ \forall u \in \mathcal{U}$$

$$f^{cc'}(u) = \big(f^c\big)^{c'}(u) = \sup_{v \in \mathcal{V}} \Big(c(u, v) \dot{+} \big(-f^c(v)\big) \Big) \ , \ \ \forall u \in \mathcal{U}$$

$\mathcal{H}_0$-biconjugates of 0-homogeneous functions

Proposition

Under the blanket assumptions, we have that

$$f^{\star_{\varrho}\star'_{\varrho}} = \overbrace{f^{\star_{\varrho}\star'}}^{\text{convex lsc}} \circ \varrho = \left(f \dotplus \iota_{\mathrm{im}\varrho}\right)^{\star\star'} \circ \varrho$$

In generalized convexity, one defines so-called c-convex functions

$$f \in \overline{\mathbb{R}}^{\mathcal{U}} \mapsto f^{c} \in \overline{\mathbb{R}}^{\mathcal{V}} \mapsto f^{cc'} \in \overline{\mathbb{R}}^{\mathcal{U}}$$

For any function $f : \mathcal{U} \to \overline{\mathbb{R}}$, one has that

$$f^{cc'} \leq f$$

Definition

The function $f : \mathcal{U} \to \overline{\mathbb{R}}$ is said to be c-convex if

$$f^{cc'} = f$$

c-convex functions have dual representations as suprema of elementary functions (abstract convexity)

If the function $f : \mathcal{U} \to \overline{\mathbb{R}}$ is c-convex, we have that

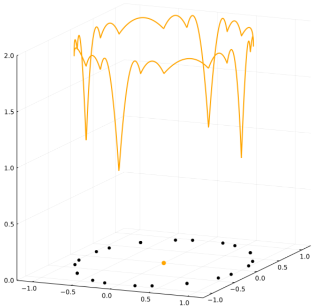
$$f(u) = \sup_{v \in \mathcal{V}} \underbrace{\Big(c(u,v) \dotplus \big(-f^{c}(v)\big) \Big)}_{\text{elementary function of } u}, \quad \forall u \in \mathcal{U}$$

Example: $\star$-convex functions
$=$ closed convex functions
$=$ proper convex lsc or $\equiv -\infty$ or $\equiv +\infty$
$=$ suprema of affine functions

$\star_\varrho$-convex functions

If the function $f : C \to \overline{\mathbb{R}}$ is $\star_\varrho$-convex, we have that

$$\boxed{f(x) = \sup_{y \in \mathcal{Y}} \underbrace{\Big(\langle \varrho(x), y \rangle \dot{+} \left(-f^{\star_\varrho}(y)\right) \Big)}_{\text{affine function of } \varrho(x)}} \; , \quad \forall x \in C$$



Outline of the presentation

$\mathcal{H}_0$-conjugacies [10 min]

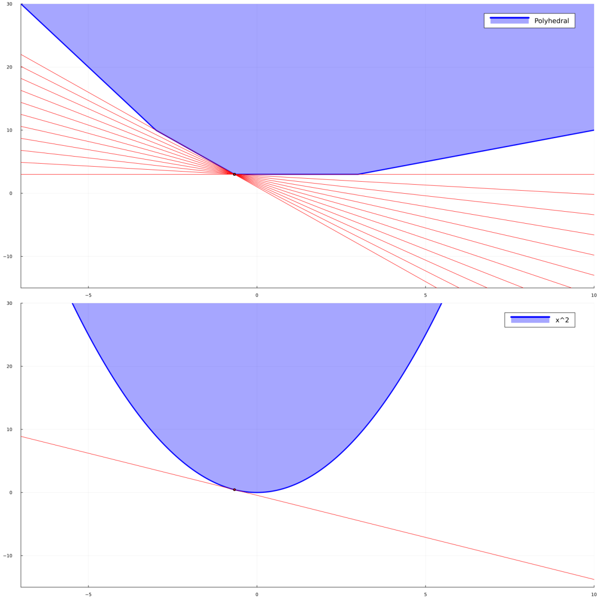
- Couplings and $\mathcal{H}_0$-couplings
- Capra-couplings and radial $\mathcal{H}_0$-couplings
- Conjugacies and $\mathcal{H}_0$-conjugacies
- Biconjugates and duality, $\mathcal{H}_0$-biconjugates
- Subdifferentials and $\mathcal{H}_0$-subdifferentials

Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0 [10 min]

- Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0
- Orthant-strict monotonicity and Capra-convexity of ℓ_0
- Orthant-strict monotonicity

Conclusion

Motivation: subgradients in convex analysis



Subdifferentials of a conjugacy

For any function $f : \mathcal{U} \to \overline{\mathbb{R}}$ and $u \in \mathcal{U}$, $v \in \mathcal{V}$

Definition

Upper subdifferential (following [Martinez-Legaz and Singer, 1995])

$$v \in \partial^c f(u) \iff f(u) = c(u, v) \dot{+} (-f^c(v))$$

The upper subdifferential $\partial^c f$ has the property that

$$\partial^c f(u) \neq \emptyset \implies \underbrace{f(u) = f^{cc'}(u)}_{\text{the function } f \text{ is } c\text{-convex at } u}$$

Definition

Lower subdifferential

$$v \in \partial_c f(u) \iff f^c(v) = c(u, v) \dot{+} (-f(u))$$

$\mathcal{H}_0$-subdifferential of 0-homogeneous functions

Proposition

Under the blanket assumptions, we have that

$$\underbrace{\partial_{\star_\varrho} f}_{\star_\varrho-\text{subdifferential}} = \underbrace{\partial(f \dotplus \iota_{\mathrm{im}\varrho})}_{\substack{\text{Moreau-Rockafellar}\\ \text{subdifferential}}} \circ \varrho$$

Roadmap: convex factorization of the ℓ_0 pseudonorm

Find

- a normalization function $\nu : \mathbb{R}^n \to [0, +\infty]$
- a nonempty cone $C \subset \mathbb{R}^n \setminus \nu^{=0}$

such that

$$\partial_{\star_{\varrho_\nu}}(\ell_0 + \iota_C)(x) = \partial(\ell_0 + \iota_{\mathrm{im}\varrho_\nu})\big(\varrho_\nu(x)\big) \neq \emptyset \ , \quad \forall x \in C$$

hence that $\ell_0 + \iota_C = \overbrace{(\ell_0 + \iota_C)^{\star_{\varrho_\nu}\star'}}^{\text{convex lsc}} \circ \varrho_\nu = \big(\ell_0 + \iota_{\mathrm{im}\varrho_\nu}\big)^{\star\star'} \circ \varrho_\nu$
and, in particular, when $C = \mathbb{R}^n$,

$$\boxed{\ell_0 = \underbrace{\ell_0^{\star_{\varrho_\nu}\star'}}_{\text{convex lsc } \mathcal{L}_0^\nu} \circ \varrho_\nu = \big(\ell_0 + \iota_{\mathrm{im}\varrho_\nu}\big)^{\star\star'} \circ \varrho_\nu}$$

Outline of the presentation

$\mathcal{H}_0$-conjugacies [10 min]

Orthant-strict monotonicity and $\mathcal{H}_0$-convexity of ℓ_0 [10 min]

Conclusion

Outline of the presentation

Polar transform of a function

The polar set $X^{\odot}$ of the subset $X \subset \mathbb{R}^n$ is the closed convex set

$$X^{\odot} = \big\{ y \in \mathbb{R}^n \,\big|\, \langle x \mid y \rangle \leq 1 \,, \;\; \forall x \in X \big\}$$

The bipolar theorem states that $X^{\odot\odot} = \overline{\mathrm{co}}(X \cup \{0\})$

Definition

For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}_+$, we introduce the polar transform $f^{\circ} : \mathbb{R}^n \to \overline{\mathbb{R}}_+$ defined by

$$f^{\circ}(y) = \sup_{x \in \mathbb{R}^n} \Big(\langle x, y \rangle_+ \,\underset{\cdot}{\times}\, \big(f(x) \big)^{-1} \Big) \,, \quad \forall y \in \mathbb{R}^n$$

where $0 \,\dot{\times}\, (+\infty) = +\infty$

Polar transform of Minkowski and support functions

If $\nu\colon \mathbb{R}^n \to [0,+\infty]$ is 1-homogeneous, we have that

$$
\begin{aligned}
\nu &= m_{\nu^{\leq 1}} \\
\nu^{\circ} &= m_{(\nu^{\leq 1})^{\odot}} = \sigma_{(\nu^{\leq 1})^{\odot\odot}} && \text{convex lsc} \\
\nu^{\circ\circ} &= m_{(\nu^{\leq 1})^{\odot\odot}} = \sigma_{(\nu^{\leq 1})^{\odot}} && \text{convex lsc}
\end{aligned}
$$

Example

When $\|\cdot\|$ is norm on $\mathbb{R}^n$, with unit sphere S,
unit ball B and dual norm $\|\cdot\|_\star$

$$
\begin{aligned}
\nu &= m_S = \|\cdot\| + \iota_{\mathcal{X}\setminus\{0\}} \\
\nu^{\leq 1} &= B \setminus \{0\}\ ,\ \ (\nu^{\leq 1})^{\odot} = B^{\odot}\ ,\ \ (\nu^{\leq 1})^{\odot\odot} = B \\
\nu^{\circ} &= m_{B^{\odot}} = \sigma_B = \|\cdot\|_\star \\
\nu^{\circ\circ} &= m_B = \sigma_{B^{\odot}} = \|\cdot\|
\end{aligned}
$$

Convex factorization of the ℓ_0 pseudonorm

Theorem

Let $Y \subset \mathbb{R}^n$ be a compact subset such that $Y^{\odot\odot} = Y$ and that the sets Y and $Y^{\odot}$ are orthant-strictly monotonic
Let us define

$$\begin{aligned} \text{the function} \quad & \nu = m_{Y^{\odot}\setminus\{0\}} \\ \text{the cone} \quad & C = \{0\} \cup \mathbb{R}^n \setminus Y^{\ominus} \end{aligned}$$

Then, ν is a normalization function, and

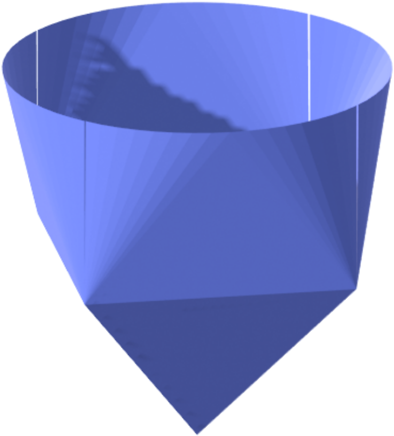
$$\partial_{\star_{\varrho\nu}}(\ell_0 + \iota_C)(x) \neq \emptyset \,, \quad \forall x \in C$$

hence

$$\boxed{\ell_0(x) = \overbrace{\mathcal{L}_0^{\nu}}^{\text{convex lsc}} \Big(\frac{x}{\nu(x)}\Big)} \,, \quad \forall x \in C$$

Outline of the presentation

Graph of the Euclidean ℓ_0-cup function $\mathcal{L}_0$



Orthant-strictly monotonic norms and hidden convexity in the ℓ_0 pseudonorm

[Chancelier and De Lara, 2022b]

Theorem

If both the norm $\|\cdot\|$ and the dual norm $\|\cdot\|_\star$ are orthant-strictly monotonic, there exists a proper convex lsc function $\mathcal{L}_0$, the ℓ_0-cup function, with domain the unit ball B, such that

$$\boxed{\ell_0(x) = \underbrace{\mathcal{L}_0}_{\text{convex lsc function}} \left(\frac{x}{\|x\|}\right)} \; , \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

and, as a consequence, the ℓ_0 pseudonorm coincides, on the unit sphere S, with the proper convex lsc function $\mathcal{L}_0$

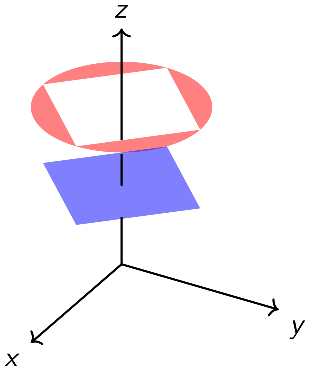
$$\ell_0(x) = \mathcal{L}_0(x) \; , \quad \forall x \in S$$

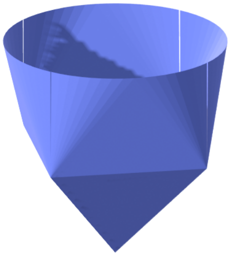
The ℓ_0-cup function as a convex envelope

Proposition

The proper convex lsc function $\mathcal{L}_0$ is the convex envelope of the following piecewise constant function

$$L_0(x) = \begin{cases} 0 & \text{if } x = 0, \\ \ell & \text{if } x \in B^{\top\star}_{(\ell)} \setminus B^{\top\star}_{(\ell-1)} , \quad \ell \in [\![1, n]\!] \\ +\infty & \text{if } x \notin B^{\top\star}_{(n)} = B \end{cases}$$





The ℓ_0-cup function as best proper convex lsc lower approximation of the ℓ_0 pseudonorm on the unit ball

Theorem

The ℓ_0-cup function $\mathcal{L}_0$ is
the best convex lsc lower approximation of the ℓ_0 pseudonorm on the unit ball B

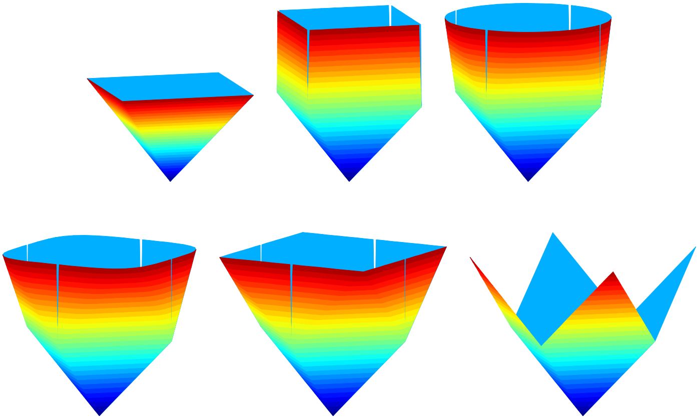
$$\text{best convex lsc function} \quad \mathcal{L}_0(x) \leq \ell_0(x) \ , \ \forall x \in B$$

and, as seen above, coincides with the ℓ_0 pseudonorm

on the unit sphere S

$$\ell_0(x) = \mathcal{L}_0(x) \ , \ \forall x \in S$$

Tightest closed convex function below the ℓ_0 pseudonorm on the ℓ_p-unit balls on $\mathbb{R}^2$ for $p \in \{1, 1.1, 2, 4, 300, \infty\}$



The Capra-convex functions are 0-homogeneous and coincide, on the unit sphere, with a closed convex function

Proposition

$$\text{¢-convexity of the function } h \colon \mathbb{R}^n \to \overline{\mathbb{R}}$$

$$\iff h = h^{¢¢'}$$

$$\iff h = \underbrace{\left(h^{¢}\right)^{\star'}}_{\text{convex lsc function}} \circ R_{\|\cdot\|}$$

$$\iff \text{hidden convexity in the function } h \colon \mathbb{R}^n \to \overline{\mathbb{R}}$$

there exists a closed convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$

such that $h = f \circ R_{\|\cdot\|}$, that is, $h(x) = f\left(\frac{x}{\|x\|}\right)$

[Chancelier and De Lara, 2022b]

Theorem

If both the norm $\|\cdot\|$ and the dual norm $\|\cdot\|_\star$ are orthant-strictly monotonic, we have that

$$\partial_{¢}\ell_0(x) \neq \emptyset \;, \quad \forall x \in \mathbb{R}^n \;,$$

and, as a consequence,

$$\boxed{\ell_0^{¢¢'} = \ell_0}$$

and thus

$$\ell_0 = \ell_0^{¢¢'} = \ell_0^{¢\star'} \circ R_{\|\cdot\|} = \underbrace{\left(\ell_0^{¢}\right)^{\star'}}_{\substack{\text{convex lsc}\\ \text{function } \mathcal{L}_0}} \circ \overbrace{R_{\|\cdot\|}}^{\substack{\text{radial}\\ \text{projection}}}$$

Variational formulas for the ℓ_0 pseudonorm

[Chancelier and De Lara, 2022b]

Proposition

If both the norm $\|\cdot\|$ and the dual norm $\|\cdot\|_\star$ are orthant-strictly monotonic, we have that

$$\ell_0(x) = \frac{1}{\|x\|} \min_{\substack{x^{(1)}\in\mathbb{R}^n,\dots,x^{(n)}\in\mathbb{R}^n \\ \sum_{\ell=1}^n \|x^{(\ell)}\|_{(\ell)}^{\top\star} \leq \|x\| \\ \sum_{\ell=1}^n x^{(\ell)} = x}} \sum_{\ell=1}^{n} \ell \|x^{(\ell)}\|_{(\ell)}^{\top\star} \;, \quad \forall x \in \mathbb{R}^n$$

$$\ell_0(x) = \sup_{y\in\mathbb{R}^n} \inf_{\ell\in[\![1,n]\!]} \left(\frac{\langle x \mid y \rangle}{\|x\|} - \Big[\|y\|_{(\ell)}^{\top} - \ell \Big]_+ \right) \;, \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

Outline of the presentation

We reformulate sparsity in terms of coordinate subspaces

$$y = (*, *, *, *, *, *) \to \pi_{\{2,4,5\}}(y) = (0, *, 0, *, *, 0) \in \mathcal{R}_{\{2,4,5\}}$$

- For any subset $K \subset [\![1, n]\!]$ of indices, we set

$$\mathcal{R}_K = \left\{ y \in \mathbb{R}^n \,\middle|\, y_j = 0 \,, \ \forall j \notin K \right\} \subset \mathbb{R}^n$$

- The connection with the level sets of the ℓ_0 pseudonorm is

$$\ell_0^{\leq k} = \underbrace{\left\{ x \in \mathbb{R}^n \,\middle|\, \ell_0(x) \leq k \right\}}_{k\text{-sparse vectors}} = \bigcup_{|K| \leq k} \mathcal{R}_K \ , \quad \forall k \in [\![0, n]\!]$$

- We denote by $\pi_K : \mathbb{R}^n \to \mathcal{R}_K$ the orthogonal projection

  For any vector $y \in \mathbb{R}^n$, $\pi_K(y) = y_K \in \mathcal{R}_K \subset \mathbb{R}^n$ is the vector whose entries coincide with those of y,
  except for those outside of K that vanish

Orthant-monotonic norms and sets

Orthant-monotonic norms

For any $x \in \mathbb{R}^n$, we denote by $|x|$
the vector of $\mathbb{R}^n$ with components $|x_i|$, $i \in [\![1, n]\!]$

Definition

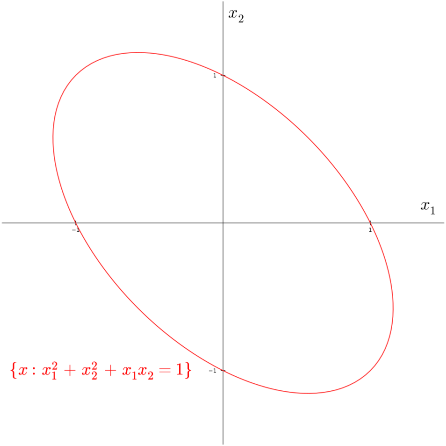
A norm $\|\cdot\|$ on the space $\mathbb{R}^n$ is called orthant-monotonic [Gries, 1967]
if, for all x, x' in $\mathbb{R}^n$, we have

$$|x| \leq |x'| \text{ and } x \circ x' \geq 0 \implies \|x\| \leq \|x'\|$$

where $x \circ x' = (x_1 x'_1, \ldots, x_n x'_n)$
is the Hadamard (entrywise) product

$$\left.\begin{array}{l} |x_1| \leq |x'_1| \,, \quad \ldots \,, \quad |x_n| \leq |x'_n| \\ \text{and} \\ x_1 x'_1 \geq 0 \,, \quad \ldots \,, \quad x_n x'_n \geq 0 \end{array}\right\} \implies \|x\| \leq \|x'\|$$

Example of unit sphere of a non orthant-monotonic norm



In the bottom right orthant, consider

$$|(0,-1)| \leq |(0.5,-1)|$$

and

$$(0,-1) \circ (0.5,-1) \geq (0,0)$$

but

$$1 = \|(0,-1)\| > \|(0.5,-1)\|$$

We define orthant-monotonic sets

Definition

The closed convex subset $X \subset \mathbb{R}^n$ is said to be orthant-monotonic if it satisfies any one of the equivalent conditions

1. $\sigma_X \circ \pi_K \leq \sigma_X$, for all $K \subset [\![1, n]\!]$
2. $\pi_K(X) \subset X$, for all $K \subset [\![1, n]\!]$
3. $\pi_K(X) \subset X \cap \mathcal{R}_K$, for all $K \subset [\![1, n]\!]$
4. $\pi_K(X) = X \cap \mathcal{R}_K$, for all $K \subset [\![1, n]\!]$

[Chancelier and De Lara, 2023]

Orthant-strictly monotonic (OSM) norms and sets

Orthant-strictly monotonic norms

[Chancelier and De Lara, 2023]

Definition

A norm $\|\cdot\|$ on the space $\mathbb{R}^n$ is called orthant-strictly monotonic if, for all x, x' in $\mathbb{R}^n$, we have

$$|x| < |x'| \text{ and } x \circ x' \geq 0 \implies \|x\| < \|x'\|$$

where $|x| < |x'|$ means that
there exists $j \in [\![1, n]\!]$ such that $|x_j| < |x_j'|$

Intuition: $\epsilon \neq 0 \implies \|(0, *, 0, *, *, 0)\| < \|(0, *, \epsilon, *, *, 0)\|$

An orthant-strictly monotonic norm is orthant-monotonic

Examples of orthant-strictly monotonic norms

$$\|x\|_\infty = \sup_{i \in [\![1,n]\!]} |x_i| \text{ and } \|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} \text{ for } p \in [1,\infty[$$

with unit ball B_p and unit sphere S_p

- All the ℓ_p-norms $\|\cdot\|_p$ on the space $\mathbb{R}^n$, for $p \in [1,\infty]$, are monotonic, hence orthant-monotonic

$$\ell_1, \ell_2, \ell_\infty$$

- All the ℓ_p-norms $\|\cdot\|_p$ on the space $\mathbb{R}^n$, for $p \in [1,\infty[$, are orthant-strictly monotonic

$$\ell_1, \ell_2, \cancel{\ell_\infty}$$

$$|\epsilon| < 1 \implies \|(1,0)\|_\infty = 1 = \|(1,\epsilon)\|_\infty$$

We define orthant-strictly monotonic sets

Definition

The closed convex subset $X \subset \mathbb{R}^n$
is said to be orthant-strictly monotonic (OSM)
if it satisfies

$$\sigma_X(y) < +\infty \text{ and } y \neq \pi_K y \implies \sigma_X(\pi_K y) < \sigma_X(y)$$

An orthant-strictly monotonic set is orthant-monotonic

Conclusion

- We have introduced $\mathcal{H}_0$-couplings, and indicated in what they are suitable tool for convex factorization of 0-homogeneous functions
- We have recalled Capra-couplings, induced by a norm, and how they reveal convex factorization of the ℓ_0 pseudonorm on the unit ball, when both the norm and the dual norm are orthant-strictly monotonic (OSM)
- We have generalized the notion of OSM and, using radial $\mathcal{H}_0$-couplings, we expect to display convex factorization of ℓ_0 on bipolar subsets (closed convex sets that contain 0) that are more general than unit balls (not necessarily symmetric, 0 is not necessarily in the interior)

Jean-Philippe Chancelier and Michel De Lara. Hidden convexity in the l_0 pseudonorm. *Journal of Convex Analysis*, 28(1):203–236, 2021.

Jean-Philippe Chancelier and Michel De Lara. Constant along primal rays conjugacies and the l_0 pseudonorm. *Optimization*, 71(2):355–386, 2022a. doi: 10.1080/02331934.2020.1822836.

Jean-Philippe Chancelier and Michel De Lara. Capra-convexity, convex factorization and variational formulations for the ℓ_0 pseudonorm. *Set-Valued and Variational Analysis*, 30:597–619, 2022b.

Jean-Philippe Chancelier and Michel De Lara. Orthant-strictly monotonic norms, generalized top-k and k-support norms and the ℓ_0 pseudonorm. *Journal of Convex Analysis*, 30(3):743–769, 2023.

D. Gries. Characterization of certain classes of norms. *Numerische Mathematik*, 10:30–41, 1967.

Juan-Enrique Martinez-Legaz and Ivan Singer. Subdifferentials with respect to dualities. *Mathematical Methods of Operations Research*, 42(1):109–125, February 1995.

J. J. Moreau. Fonctionnelles convexes. *Séminaire Jean Leray*, 2:1–108, 1966-1967.

Jean Jacques Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl. (9)*, 49:109–154, 1970.

R. Tyrrell Rockafellar. *Conjugate Duality and Optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1974.

Thank you :-)