Lecture Series on "Convex analysis with applications in inverse problems"

Lecture 1: Convex analysis: basics, conjugation and duality

Radu Ioan Boţ

Chemnitz University of Technology
Department of Mathematics
09107 Chemnitz
www.tu-chemnitz.de/∼rabot

Institute for Numerical and Applied Mathematics
University of Goettingen
June 11, 2012

TECHNISCHE UNIVERSITÄT CHEMNITZ

Deutsche Forschungsgemeinschaft DFG

Lecture Series
on
"Convex analysis with applications in inverse problems"

- Lecture 1: Convex analysis: basics, conjugation and duality (Monday, June 11, 2012)
- Lecture 2: Proximal methods in convex optimization (Wednesday, June 13, 2012)
- Lecture 3: Convex regularization techniques for linear inverse problems (Thursday, June 14, 2012)

Contents

Convex functions

Algebraic properties of convex functions

Let $(X, \|\cdot\|)$ be a normed space, $(X^*, \|\cdot\|_*)$ its topological dual space and the duality pairing on $X^* \times X$, $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{R}, \langle x^*, x \rangle = x^*(x)$.

Convex function

A function $f : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is said to be convex, if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \ \forall x, y \in X \ \forall \lambda \in [0,1].$$

- Conventions: $(+\infty) + (-\infty) = +\infty, 0(+\infty) = +\infty, 0(-\infty) = 0$.
- The effective domain of the function $f : X \to \overline{\mathbb{R}}$ is the set $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$. If f is convex, then $\operatorname{dom} f$ is a convex set.
- A function $f : X \to \overline{\mathbb{R}}$ is said to be proper if $f(x) > -\infty \ \forall x \in X$ and $\operatorname{dom} f \neq \emptyset$.

Some examples of convex functions

- The norm $\|\cdot\| : X \to \mathbb{R}$ is a convex function.
- The indicator function of a set $S \subseteq X$ is defined as

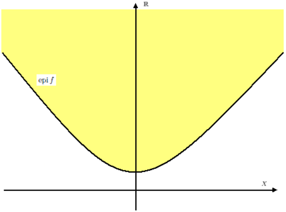
$$\delta_S : X \to \overline{\mathbb{R}}, \delta_S(x) = \begin{cases} 0, & \text{if } x \in S, \\ +\infty, & \text{otherwise.} \end{cases}$$

The function δ_S is convex if and only if S is a convex set.

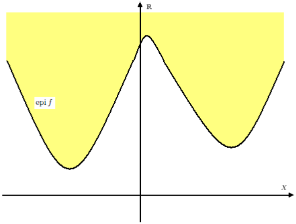
- When $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix, then $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T A x$, is convex if and only if A is positive semidefinite.

- The epigraph of a function $f : X \to \overline{\mathbb{R}}$ is the set

$$\operatorname{epi} f = \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}.$$



- The function f is convex if and only if the set $\operatorname{epi} f$ is convex.



Level set

If $f : X \to \overline{\mathbb{R}}$ is a convex function, then for each $\lambda \in \mathbb{R}$ its upper level set

$$\{x \in X : f(x) \leq \lambda\}$$

is convex. However, the opposite statement is not true. A counterexample in this sense is provided by the function $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$.

Sublinear function

A function $f : X \to \overline{\mathbb{R}}$ is said to be sublinear, if it is:

- positively homogeneous: $f(0) = 0$ and $f(\lambda x) = \lambda f(x)\ \forall \lambda > 0\ \forall x \in X$;
- subadditive: $f(x + y) \leq f(x) + f(y)\ \forall x, y \in X$.

- A function is sublinear if and only if it is positively homogeneous and convex.
- A function $f : X \to \overline{\mathbb{R}}$ is sublinear if and only if $\operatorname{epi} f$ is a convex cone with $(0, -1) \notin \operatorname{epi} f$.

Composition with an affine mapping

When $(Y, \|\cdot\|)$ is another normed space, the operator $T : X \to Y$ is said to be affine, if

$$T(\lambda x + (1 - \lambda)y) = \lambda T(x) + (1 - \lambda)T(y)\ \forall x, y \in X\ \forall \lambda \in \mathbb{R}.$$

When $f : Y \to \overline{\mathbb{R}}$ is convex and $T : X \to Y$ is affine, then $f \circ T : X \to \overline{\mathbb{R}}$ is convex.

Pointwise supremum

The pointwise supremum of a family of convex functions $f_i : X \to \overline{\mathbb{R}}$,

$$\sup_{i\in I} f_i : X \to \overline{\mathbb{R}}, \sup_{i\in I} f_i(x) = \sup\{f_i(x) : i \in I\},$$

is convex. Notice that $\operatorname{epi}\left(\sup_{i\in I} f_i\right) = \bigcap_{i\in I} \operatorname{epi} f_i$.

Infimal value function

When $\Phi : X \times Y \to \overline{\mathbb{R}}$ is convex, then its infimal value function

$$h : Y \to \overline{\mathbb{R}}, h(y) = \inf\{\Phi(x,y) : x \in X\},$$

is convex, too.

Infimal convolution

The infimal convolution of two functions $f, g : X \to \overline{\mathbb{R}}$ is defined as

$$f\Box g : X \to \overline{\mathbb{R}}, (f\Box g)(x) = \inf\{f(x-y) + g(y) : y \in X\}.$$

One has $\operatorname{epi}(f\Box g) = \operatorname{epi} f + \operatorname{epi} g$. When f and g are convex, then $f\Box g$ is convex, too.

Example (distance function)

When $S \subseteq X$ is a convex set, then its distance function $d_S : X \to \overline{\mathbb{R}}$ fulfills

$$d_S(x) = \inf\{\|x-y\| : y \in S\} = (\|\cdot\|\Box\delta_S)(x) \ \forall x \in X,$$

thus it is convex.

Topological properties of convex functions

Lower semicontinuous function

A function $f : X \to \overline{\mathbb{R}}$ is said to be

- lower semicontinuous at $x \in X$, if $\liminf_{y \to x} f(y) := \sup_{\delta > 0} \inf_{y \in B(x,\delta)} f(y) \geq f(x)$;
- lower semicontinuous, if it is lower semicontinuous at every $x \in X$.

For a given function $f : X \to \overline{\mathbb{R}}$ the following statements are equivalent:

- f is lower semicontinuous;
- $\operatorname{epi} f$ is closed;
- every upper level set $\{x \in X : f(x) \leq \lambda\}$, $\lambda \in \mathbb{R}$, is closed.

Example (indicator function)

For the indicator function δ_S of a set $S \subseteq X$ one has $\operatorname{epi} \delta_S = S \times \mathbb{R}_+$. Thus δ_S is lower semicontinuous if and only if S is closed.

Pointwise supremum

The pointwise supremum of a family of lower semicontinuous functions $f_i : X \to \overline{\mathbb{R}}$,

$$\sup_{i \in I} f_i : X \to \overline{\mathbb{R}}, \sup_{i \in I} f_i(x) = \sup\{f_i(x) : i \in I\},$$

is lower semicontinuous.

Lower semicontinuous hull

The lower semicontinuous hull of a function $f : X \to \overline{\mathbb{R}}$ is defined as

$$\bar{f} : X \to \overline{\mathbb{R}}, \bar{f}(x) = \inf\{r : (x,r) \in \mathrm{cl}(\mathrm{epi}\, f)\}.$$

The following statements are true:

- $\liminf_{y \to x} f(y) = \bar{f}(x) \ \forall x \in X$;
- $\mathrm{epi}\, \bar{f} = \mathrm{cl}(\mathrm{epi}\, f)$;
- $\bar{f} = \sup\{h : X \to \overline{\mathbb{R}} : h \leq f$ and h is lower semicontinuous$\}$.

Affine minorant

One says that $x \mapsto \langle x^*, x\rangle + \alpha$, where $(x^*, \alpha) \in X^* \times \mathbb{R}$, is an affine minorant of $f : X \to \overline{\mathbb{R}}$, if

$$\langle x^*, y\rangle + \alpha \leq f(y) \ \forall y \in X.$$

Fundamental result

A function $f : X \to \overline{\mathbb{R}}$ is convex, lower semicontinuous and it fulfills $f > -\infty$ if and only if there exists $(x^*, \alpha) \in X^* \times \mathbb{R}$ such that $\langle x^*, y\rangle + \alpha \leq f(y)$ for all $y \in X$ and

$$f(x) = \sup\{\langle x^*, x\rangle + \alpha : (x^*, \alpha) \in X^* \times \mathbb{R}, \langle x^*, y\rangle + \alpha \leq f(y) \ \forall y \in X\} \ \forall x \in X.$$

Weak lower semicontinuity

- A function $f : X \to \overline{\mathbb{R}}$ is said to be weakly lower semicontinuous, if $\operatorname{epi} f$ is weakly closed.
- Since

$$\operatorname{epi} f \subseteq \operatorname{cl}(\operatorname{epi} f) \subseteq \operatorname{cl}_{\omega(X,X^*)\times\mathbb{R}}(\operatorname{epi} f),$$

every weakly lower semicontinuous function is lower semicontinuous, too.

- If $f : X \to \overline{\mathbb{R}}$ is convex, then f is weakly lower semicontinuous if and only if f is lower semicontinuous.

Continuity via convexity

If a convex function $f : X \to \overline{\mathbb{R}}$ is bounded above on a neighborhood of a point of its domain, then f is continuous on $\operatorname{int}(\operatorname{dom} f)$.

Local Lipschitz continuity via convexity

If a proper and convex function $f : X \to \overline{\mathbb{R}}$ is bounded above on a neighborhood of a point of its domain, then f is locally Lipschitz continuous on $\operatorname{int}(\operatorname{dom} f)$, i.e. for all $x \in \operatorname{int}(\operatorname{dom} f)$ there exist $\varepsilon > 0$ and $L \geq 0$ such that

$$|f(y) - f(z)| \leq L\|y - z\| \ \forall y, z \in B(x, \varepsilon).$$

An intermezzo: the algebraic interior of a convex set

The algebraic interior of a convex set $S \subseteq X$ is

$$\operatorname{core}(S) := \{s \in S : \operatorname{cone}(S - s) = \bigcup_{\lambda > 0} \lambda(S - s) = X\}.$$

- One always has $\operatorname{int}(S) \subseteq \operatorname{core}(S)$.
- If $\operatorname{int}(S) \neq \emptyset$ or X is finite-dimensional, then $\operatorname{int}(S) = \operatorname{core}(S)$.

Example

Let $x^\sharp : X \to \mathbb{R}$ be a discontinuous linear functional and $S := \{x \in X : |\langle x^\sharp, x\rangle| \leq 1\}$. Then $\operatorname{int}(S) = \emptyset$, while $0 \in \operatorname{core}(S) \neq \emptyset$.

From lower semicontinuity to continuity

If X is a Banach space and $f : X \to \overline{\mathbb{R}}$ is a convex and lower semicontinuous function, then $\operatorname{int}(\operatorname{dom} f) = \operatorname{core}(\operatorname{dom} f)$ and f is continuous on $\operatorname{int}(\operatorname{dom} f)$.

Example

If X is a Banach space and $S \subseteq X$ is a convex and closed set, then $\operatorname{int}(S) = \operatorname{int}(\operatorname{dom} \delta_S) = \operatorname{core}(\operatorname{dom} \delta_S) = \operatorname{core}(S)$. However, these sets can be also empty. This is, for instance, the case when

$$p \in [1, +\infty), X = \ell^p \text{ and } S = \ell^p_+ := \{(x_k)_{k \geq 1} \in \ell_p : x_k \geq 0 \ \forall k \geq 1\}.$$

Conjugacy and subdifferentiability

Conjugate functions

(Fenchel-Legendre-) Conjugate function of a function $f : X \to \overline{\mathbb{R}}$:

$$f^* : X^* \to \overline{\mathbb{R}}, f^*(x^*) = \sup_{x \in X} \{\langle x^*, x\rangle - f(x)\}.$$

Properties of the conjugate function (I)

For a given function $f : X \to \overline{\mathbb{R}}$ we have:

- f^* is convex and weak* lower semicontinuous;
- Young-Fenchel-inequality:

$$f(x) + f^*(x^*) \geq \langle x^*, x\rangle \ \forall (x, x^*) \in X \times X^*;$$

- when, for $g : X \to \overline{\mathbb{R}}$, $f \leq g$, then $g^* \leq f^*$;
- $f^* = (\bar{f})^*$.

Examples

- The conjugate function of the indicator function of a set $S \subseteq X$ is the so-called support function of S,

$$\sigma_S : X^* \to \overline{\mathbb{R}}, \sigma_S(x^*) = \delta_S^*(x^*) = \sup_{x \in S} \langle x^*, x \rangle.$$

- For $f = \| \cdot \|$, one has $f^*(x^*) = \begin{cases} 0, & \text{if } \|x^*\|_* \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$
- For $f = \frac{1}{2}\| \cdot \|^2$, one has $f^* = \frac{1}{2}\| \cdot \|_*^2$.

Properties of the conjugate function (II)

For a given function $f : X \to \overline{\mathbb{R}}$ we have:

- $-f^*(0) = \inf_{x \in X} f(x)$;
- $(\lambda f)^*(x^*) = \lambda f^*\left(\frac{1}{\lambda} x^*\right) \ \forall \lambda > 0 \ \forall x^* \in X^*$;
- for $\bar{x} \in X$:
$$(f(\cdot + \bar{x}))^*(x^*) = f^*(x^*) - \langle x^*, \bar{x} \rangle \ \forall x^* \in X^*;$$
- for $\bar{x}^* \in X^*$:
$$(f + \langle \bar{x}^*, \cdot \rangle)^*(x^*) = f^*(x^* - \bar{x}^*) \ \forall x^* \in X^*.$$

Properties of the conjugate function (III)

Let be $\Phi : X \times Y \to \overline{\mathbb{R}}$.

- If
$$h : Y \to \overline{\mathbb{R}}, h(y) = \inf\{\Phi(x, y) : x \in X\},$$
then
$$h^*(y^*) = \Phi^*(0, y^*) \ \forall y^* \in Y^*.$$
- If
$$\Phi(x, y) = f(x) + g(y),$$
where $f : X \to \overline{\mathbb{R}}$ and $g : Y \to \overline{\mathbb{R}}$, then
$$\Phi^*(x^*, y^*) = f^*(x^*) + g^*(y^*) \ \forall (x^*, y^*) \in X^* \times Y^*.$$

The conjugate of the infimal convolution

For $f, g : X \to \overline{\mathbb{R}}$ proper functions one has
$$(f \Box g)^* = f^* + g^*.$$

Biconjugate function of a function $f : X \to \overline{\mathbb{R}}$

$$f^{**} : X \to \overline{\mathbb{R}}, f^{**}(x) = \sup_{x^* \in X^*} \{\langle x^*, x\rangle - f^*(x^*)\}.$$

- When X^* is endowed with the weak* topology, then $f^{**} = (f^*)^*$.
- One always has: $f^{**} \leq \bar{f} \leq f$.

Theorem of Fenchel-Moreau

If $f : X \to \overline{\mathbb{R}}$ is a proper, convex and lower semicontinuous function, then f^* is proper and it holds $f^{**} = f$.

Conjugate of the biconjugate

For $f : X \to \overline{\mathbb{R}}$ a given function it holds

$$f^{***} = (f^{**})^* = (f^*)^{**} = f^*.$$

The conjugate of the sum

For $f, g : X \to \overline{\mathbb{R}}$ proper, convex and lower semicontinuous functions with $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$ it holds

$$(f+g)^* = (f^{**} + g^{**})^* = (f^* \Box g^*)^{**} = \left(\overline{f^* \Box g^*}\right)^{**} = \overline{f^* \Box g^*}.$$

The convex subdifferential

The convex subdifferential of f at $x \in X$:

$$\partial f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x \rangle \ \forall y \in X\},$$

for $f(x) \in \mathbb{R}$. Otherwise, $\partial f(x) := \emptyset$.

Properties of the convex subdifferential (I)

For a given function $f : X \to \overline{\mathbb{R}}$ and $x \in X$ we have:

- the set $\partial f(x)$ is convex and weak* closed and it can be empty, even if $f(x) \in \mathbb{R}$;
- $x^* \in \partial f(x) \Leftrightarrow f(x) + f^*(x^*) = \langle x^*, x \rangle$;
- if $\partial f(x) \neq \emptyset$, then $\bar{f}(x) = f(x)$ and $\partial \bar{f}(x) = \partial f(x)$;
- when f proper:

$$x \text{ is a global minimum of } f \Leftrightarrow 0 \in \partial f(x).$$

Examples

- The convex subdifferential of the indicator function of a set $S \subseteq X$ at $x \in X$ is the so-called normal cone of S at X,

$$N_S(x) := \partial(\delta_S)(x) = \begin{cases} \{x^* \in X^* : \langle x^*, y - x\rangle \leq 0 \ \forall y \in S\}, & \text{if } x \in S, \\ \emptyset, & \text{otherwise.} \end{cases}$$

- One has

$$\partial\| \cdot \|(x) = \begin{cases} \{x^* \in X^* : \|x^*\|_* \leq 1\}, & \text{if } x = 0, \\ \{x^* \in X^* : \|x^*\|_* = 1, \|x\| = \langle x^*, x\rangle\}, & \text{otherwise.} \end{cases}$$

- One has $\partial\left(\frac{1}{2}\|\cdot\|^2\right)(x) = \{x^* \in X^* : \|x^*\|_* = \|x\|, \|x^*\|_*\|x\| = \langle x^*, x\rangle\}$.

Properties of the convex subdifferential (II)

For a given function $f : X \to \overline{\mathbb{R}}$ and $x \in X$ we have:

- $\partial(\lambda f)(x) = \lambda \partial f(x) \ \ \forall \lambda > 0$;
- for $\bar{x} \in X$:

$$\partial f(\cdot + \bar{x})(x) = \partial f(x + \bar{x});$$

- for $\bar{x}^* \in X^*$:

$$\partial(f + \langle \bar{x}^*, \cdot\rangle)(x) = \partial f(x) + x^*.$$

Properties of the convex subdifferential (III)

For a proper function $f : X \to \overline{\mathbb{R}}$ and $x \in \operatorname{dom} f$ we have:

- $x^* \in \partial f(x) \Rightarrow x \in \partial f^*(x^*)$, where

$$\partial f^*(x^*) := \{z \in X : f^*(y^*) - f^*(x^*) \geq \langle y^* - x^*, z\rangle \ \forall y^* \in X^*\};$$

- if f is convex and lower semicontinuous at x, then

$$x^* \in \partial f(x) \Leftrightarrow x \in \partial f^*(x^*).$$

The convex subdifferential of the sum of two functions

For $f : X \to \overline{\mathbb{R}}$, $g : Y \to \overline{\mathbb{R}}$ given functions and $A : X \to Y$ a linear continuous operator it holds

$$\partial f(x) + A^*(\partial g(Ax)) \subseteq \partial(f + g \circ A)(x) \ \forall x \in X,$$

where $A^* : Y^* \to X^*$,

$$\langle A^* y^*, x\rangle = \langle y^*, Ax\rangle \ \forall (x, y^*) \in X \times Y^*,$$

denotes the adjoint operator of A.
Thus, when $X = Y$ and A is the identity on X, it holds

$$\partial f(x) + \partial g(x) \subseteq \partial(f + g)(x) \ \forall x \in X.$$

Convex subdifferential and directional derivatives

Let $f : X \to \overline{\mathbb{R}}$ be a proper and convex function and $x \in \operatorname{dom} f$. The following statements are true:

- the directional derivative of f at x fulfills for every direction $d \in X$:

$$f'(x;d) = \lim_{t \downarrow 0} \frac{f(x+td)-f(x)}{t} = \inf_{t>0} \frac{f(x+td)-f(x)}{t} \in \overline{\mathbb{R}};$$

- it holds:

$$\partial f(x) = \{x^* \in X^* : f'(x;d) \geq \langle x^*, d\rangle \ \forall d \in X\};$$

- if f is Gâteaux differentiable at x, i.e

$$\exists \nabla f(x) \in X^* \text{ such that } f'(x;d) = \langle \nabla f(x), d\rangle \ \forall d \in X,$$

then

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

Examples

When $(X, \|\cdot\|)$ is a Hilbert space one has

- $\partial \|\cdot\|(x) = \begin{cases} \{x^* \in X : \|x^*\| \leq 1\}, & \text{if } x = 0, \\ \left\{\frac{1}{\|x\|}x\right\}, & \text{otherwise.} \end{cases}$
- $\partial \left(\frac{1}{2}\|\cdot\|^2\right)(x) = \{x\}$ for all $x \in X$.

Subdifferentiability via continuity

Let $f : X \to \overline{\mathbb{R}}$ be proper, convex and continuous at $x \in \operatorname{dom} f$. The following statements are true:

- $\partial f(x) \neq \emptyset$;
- $\partial f(x)$ is weak* compact and, consequently, norm-bounded;
- $f'(x;\cdot)$ is continuous and it holds

$$f'(x;d) = \max\{\langle x^*, d\rangle : x^* \in \partial f(x)\} \ \forall d \in X;$$

- if $\partial f(x)$ is a singleton, then f is Gâteaux differentiable at x.

Example

When $f : X \to \overline{\mathbb{R}}$ is a proper, convex and lower semicontinuous function at $x \in \operatorname{dom} f$, which fails to be continuous at $x \in \operatorname{dom} f$, $\partial f(x)$ may be empty. For

$$f : \mathbb{R} \to \overline{\mathbb{R}}, f(x) = \begin{cases} -\sqrt{1-x^2}, & \text{if } |x| \leq 1, \\ +\infty, & \text{otherwise,} \end{cases}$$

one has $\partial f(1) = \emptyset$.
Moreover,

$$\emptyset = 0\partial f(1) \neq \partial(0f)(1) = \mathbb{R}_-.$$

Convex duality

Fenchel duality

For $f : X \to \overline{\mathbb{R}}$ and $g : Y \to \overline{\mathbb{R}}$ proper and convex functions fulfilling $A(\operatorname{dom} f) \cap \operatorname{dom} g \neq \emptyset$, we consider the unconstrained optimization problem

$$(P) \quad \inf_{x \in X} \{f(x) + g(Ax)\}.$$

Particular case included (I)

For $X = Y$, A the identity operator on X and $f, g : X \to \overline{\mathbb{R}}$ proper and convex functions fulfilling $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, problem (P) reads

$$\inf_{x \in X} \{f(x) + g(x)\}.$$

Particular case included (II)

Let $f_i : X \to \overline{\mathbb{R}}$, $i = 1, ..., k$, be proper and convex functions fulfilling $\cap_{i=1}^k \operatorname{dom} f_i \neq \emptyset$. By taking $Y := \prod_{i=1}^k X$, $A : X \to Y$, $Ax = (x, ..., x)$, $f(x) = 0$ for all $x \in X$ and $g : Y \to \overline{\mathbb{R}}$, $g(x_1, ..., x_k) = \sum_{i=1}^k f_i(x_i)$, problem (P) becomes

$$\inf_{x \in X} \left\{ \sum_{i=1}^k f_i(x) \right\}.$$

Fenchel dual problem to (P):

$$(D) \quad \sup_{y^* \in Y^*} \left\{ -f^*(-A^*y^*) - g^*(y^*) \right\}.$$

Weak duality (is always fulfilled):

$$\inf_{x \in X} \{ f(x) + g(Ax) \} \geq \sup_{y^* \in Y^*} \left\{ -f^*(-A^*y^*) - g^*(y^*) \right\}.$$

Strong duality holds, if:

$$\inf_{x \in X} \{ f(x) + g(Ax) \} = \max_{y^* \in Y^*} \left\{ -f^*(-A^*y^*) - g^*(y^*) \right\}.$$

Example (nonzero duality gap)

Let $A : \mathbb{R}^2 \to \mathbb{R}^2$, $A(x_1, x_2) = (x_1, x_2)$,

$$f : \mathbb{R}^2 \to \overline{\mathbb{R}},\ f(x_1, x_2) = \max\{-1, -\sqrt{x_1 x_2}\} + \delta_{\mathbb{R}^2_+}(x_1, x_2)$$

and

$$g : \mathbb{R}^2 \to \overline{\mathbb{R}},\ g(x_1, x_2) = \delta_{\{0\}\times\mathbb{R}}(x_1, x_2).$$

The optimal objective value of (P) is equal to 0, while the optimal objective value of (D) is equal to -1.

Example (zero duality gap, but no strong duality)

Let $A : \mathbb{R} \to \mathbb{R}$, $Ax = x$,

$$f : \mathbb{R} \to \overline{\mathbb{R}},\ f(x) = \begin{cases} x(\ln x - 1), & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ +\infty, & \text{otherwise,} \end{cases}$$

and

$$g : \mathbb{R} \to \overline{\mathbb{R}},\ g(x) = \frac{1}{2}x^2 + \delta_{\mathbb{R}_-}(x).$$

The optimal objective values of (P) and (D) are both equal to 0, however the dual problem has no optimal solution.

An intermezzo: the strong-quasi relative interior

The strong-quasi relative interior of a convex set $S \subseteq X$ is

$$\operatorname{sqri}(S) := \{s \in S : \operatorname{cone}(S - s) \text{ is a closed linear subspace}\}.$$

- Recall: $\operatorname{core}(S) = \{s \in S : \operatorname{cone}(S - s) = X\}$.
- One always has $\operatorname{int}(S) \subseteq \operatorname{core}(S) \subseteq \operatorname{sqri}(S)$.
- If $\operatorname{int}(S) \neq \emptyset$, then $\operatorname{int}(S) = \operatorname{core}(S) = \operatorname{sqri}(S)$.
- If X is finite-dimensional, then

$$\operatorname{int}(S) = \operatorname{core}(S) \text{ and } \operatorname{sqri}(S) = \operatorname{ri}(S) = \operatorname{int}_{\operatorname{aff}(S)}(S).$$

Interiority-type qualification conditions for Fenchel duality:

- (F): $\exists x' \in \operatorname{dom} f \cap A^{-1}(\operatorname{dom} g)$ such that g is continuous at Ax';
- (MR) (Moreau-Rockafellar, 1966): $0 \in \operatorname{core}(A(\operatorname{dom} f) - \operatorname{dom} g)$;
- (AB) (Attouch-Brezis, 1986): $0 \in \operatorname{sqri}(A(\operatorname{dom} f) - \operatorname{dom} g)$.

Strong duality statements:

- $(F) \Rightarrow$ strong duality for $(P) - (D)$;
- When X and Y are Banach spaces and f, g are lower semicontinuous, then $(F) \Rightarrow (MR) \Rightarrow (AB) \Rightarrow$ strong duality for $(P) - (D)$.

The finite-dimensional case

If $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, then $(AB) \Leftrightarrow A(\operatorname{ri}(\operatorname{dom} f)) \cap \operatorname{ri}(\operatorname{dom} g) \neq \emptyset \Rightarrow$ strong duality for $(P) - (D)$.

Closedness-type qualification condition for Fenchel duality:

- (B): $(A^* \times \mathrm{id}_{\mathbb{R}})(\operatorname{epi} f^*) + \operatorname{epi} g^*$ is closed in $(X^*, \omega(X^*, X)) \times \mathbb{R}$.

- If f, g are lower semicontinuous, then $(B) \Rightarrow$ strong duality for $(P) - (D)$.
- If X, Y are Banach spaces and f, g are lower semicontinuous, then $(F) \Rightarrow (MR) \Rightarrow (AB) \Rightarrow (B)$.

Example

Let $A : \mathbb{R} \to \mathbb{R}$, $Ax = x$,

$$f : \mathbb{R} \to \overline{\mathbb{R}},\ f(x) = \frac{1}{2}x^2 + \delta_{\mathbb{R}_+}(x) \text{ and } g : \mathbb{R} \to \overline{\mathbb{R}},\ g(x) = \delta_{\mathbb{R}_-}(x).$$

The functions f and g are proper, convex and lower semicontinuous and none of the interiority-type qualification conditions is fulfilled. On the other hand,

$$(A^* \times \mathrm{id}_{\mathbb{R}})(\operatorname{epi} f^*) + \operatorname{epi} g^* = \mathbb{R} \times \mathbb{R}_+$$

and (B) is valid, i.e. for (P) and (D) one has strong duality.

Subdifferential formulae

- Recall:

$$\partial(f + g \circ A)(x) \supseteq \partial f(x) + A^*(\partial g(Ax)) \ \forall x \in X.$$

- Each of the qualification conditions $(F), (MR), (AB)$ and (B) guarantees (under corresponding topological assumptions) that

$$\partial(f + g \circ A)(x) = \partial f(x) + A^*(\partial g(Ax)) \ \forall x \in X.$$

Optimality conditions for (P)

Assume that one of the qualification conditions $(F), (MR), (AB)$ and (B) (under corresponding topological assumptions) is fulfilled. Then $\bar{x} \in X$ is an optimal solution to (P) if and only if

$$0 \in \partial f(\bar{x}) + A^*(\partial g(A\bar{x})).$$

Lagrange duality

Consider the geometric and cone-constrained optimization problem

$$
\begin{array}{lll}
(P) & \inf & f(x), \\
 & \text{s.t.} & g(x) \in -K, \\
 & & x \in S
\end{array}
$$

where

- X, Z are two normed spaces;
- $K \subseteq Z$ is a nonempty convex cone, i.e., $\forall \lambda \geq 0\ \forall k \in K \Rightarrow \lambda k \in K$. By $\leq_K$ we denote the partial order induced by K on Z, i.e.,

  $$\text{for } u, v \in Z \text{ it holds } u \leq_K v \Leftrightarrow v - u \in K$$

  and by

  $$K^* := \{\lambda \in Z^* : \langle \lambda, k \rangle \geq 0\ \forall k \in K\}$$

  the dual cone of K;
- $S \subseteq X$ is a convex set;
- $f : X \to \overline{\mathbb{R}}$ is a proper and convex function;
- $g : X \to Z$ is a K-convex function, i.e.,

  $$\text{the } K\text{-epigraph of } g,\ \operatorname{epi}_K g = \{(x, z) \in X \times Z : g(x) \leq_K z\}, \text{ is convex}$$

  or, equivalently,

  $$g(\lambda x + (1-\lambda)y) \leq_K \lambda g(x) + (1-\lambda) g(y)\ \forall x, y \in X\ \forall \lambda \in [0,1];$$

- the feasiblity condition $\operatorname{dom} f \cap \mathcal{A} \neq \emptyset$ is fulfilled, with

  $$\mathcal{A} := \{x \in S : g(x) \in -K\}.$$

Particular case included (I)

For $Z = \mathbb{R}^m$, $K = \mathbb{R}^m_+$ and $g = (g_1, ..., g_m)^T : X \to \mathbb{R}^m$, problem (P) reads

$$\begin{array}{ll} \inf & f(x). \\ \text{s.t.} & g_i(x) \leq 0, i = 1, ..., m, \\ & x \in S \end{array}$$

The function g is K-convex $\Leftrightarrow g_i, i = 1, ..., m$, is convex.

Particular case included (II)

For $Z = \mathbb{R}^{m+p}$, $K = \mathbb{R}^m_+ \times \{0_{\mathbb{R}^p}\}$ and $g = (g_1, ..., g_m, h_1, ..., h_p)^T : X \to \mathbb{R}^{m+p}$, problem (P) reads

$$\begin{array}{ll} \inf & f(x). \\ \text{s.t.} & g_i(x) \leq 0, i = 1, ..., m, \\ & h_j(x) = 0, j = 1, ..., p, \\ & x \in S \end{array}$$

The function g is K-convex $\Leftrightarrow g_i, i = 1, ..., m$, is convex and $h_j, j = 1, ..., p$, is affine.

Particular case included (III)

For an arbitrary index set I, $Z = \mathbb{R}^I := \{z | z : I \to \mathbb{R}\}$, $K = (\mathbb{R}^I)_+ := \{z \in \mathbb{R}^I | z(i) \geq 0 \ \forall i \in I\}$ and $g = (g_i)_{i \in I} : \mathcal{X} \to \mathbb{R}^I$, problem (P) reads

$$\begin{array}{ll} \inf & f(x). \\ \text{s.t.} & g_i(x) \leq 0, i \in I, \\ & x \in S \end{array}$$

The function g is K-convex $\Leftrightarrow$ g_i is convex for every $i \in I$.

Lagrange dual problem to (P):

$$(D) \quad \sup_{\lambda \in K^*} \inf_{x \in S} \{f(x) + \langle \lambda, g(x) \rangle\}.$$

Weak duality (is always fulfilled):

$$\inf_{x \in \mathcal{A}} f(x) \geq \sup_{\lambda \in K^*} \inf_{x \in S} \{f(x) + \langle \lambda, g(x) \rangle\}.$$

Strong duality holds, if:

$$\inf_{x \in \mathcal{A}} f(x) = \max_{\lambda \in K^*} \inf_{x \in S} \{f(x) + \langle \lambda, g(x) \rangle\}.$$

Example (nonzero duality gap)

Let $X = \mathbb{R}^2$, $Z = \mathbb{R}$, $K = \mathbb{R}_+$, $S = \{0\} \times [3,4] \cup (0,2] \times (1,4] \subseteq \mathbb{R}^2$,

$$f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1, x_2) = x_2$$

and

$$g : \mathbb{R}^2 \to \mathbb{R},\ g(x_1, x_2) = x_1.$$

Then $\mathcal{A} = 0 \times [3,4]$ and the optimal objective value of (P) is equal to 3, while the optimal objective value of (D) is equal to 1.

Interiority-type qualification conditions for Lagrange duality:

- (S) (Slater qualification condition): $\exists x' \in \operatorname{dom} f \cap S$ such that $g(x') \in -\operatorname{int}(K)$;
- (R) (Rockafellar, 1974): $0 \in \operatorname{core}(g(\operatorname{dom} f \cap S) + K)$;
- (JW) (Jeyakumar-Wolkowicz, 1992): $0 \in \operatorname{sqri}(g(\operatorname{dom} f \cap S) + K)$.

Strong duality statements:

- $(S) \Rightarrow$ strong duality for $(P) - (D)$;
- If X and Z are Banach spaces, S is closed, f is lower semicontinuous and g is K-epi closed (i.e. $\operatorname{epi}_K g$ is closed), then $(S) \Rightarrow (R) \Rightarrow (JW) \Rightarrow$ strong duality for $(P) - (D)$.

The finite-dimensional case

If $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $K = \mathbb{R}^m_+$ and $g = (g_1, ..., g_m)^T : \mathbb{R}^n \to \mathbb{R}^m$, then the three conditions become

$$\exists x' \in \operatorname{dom} f \cap S \text{ such that } g_i(x') < 0, i = 1, ..., m.$$

Recall also the following weak Slater qualification condition

- (WS) (Rockafellar, 1970): $\exists x' \in \operatorname{ri}(\operatorname{dom} f \cap S)$ such that $g_i(x') \leq 0, i \in L$, and $g_i(x') < 0, i \in N$,

where $L = \{i \in \{1, ..., m\} : g_i \text{ is affine}\}$ and $N = \{1, ..., m\} \setminus L$.

Closedness-type qualification condition for Lagrange duality:

- (B): $\bigcup_{\lambda \in K^*} \operatorname{epi}(f + \langle \lambda, g \rangle + \delta_S)^*$ is closed in $(X^*, \omega(X^*, X)) \times \mathbb{R}$.

- If S is closed, f is lower semicontinuous and g is K-epi closed, then $(B) \Rightarrow$ strong duality for $(P) - (D)$.
- If X and Z are Banach spaces, S is closed, f is lower semicontinuous and g is K-epi closed, then $(S) \Rightarrow (R) \Rightarrow (JW) \Rightarrow (B)$.

Example

Let $X = Z = \mathbb{R}^2$, $K = \mathbb{R}^2_+$, $S = \mathbb{R}^2_+$,

$$f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1, x_2) = \frac{1}{2}x_1^2 + x_2 \text{ and } g : \mathbb{R}^2 \to \mathbb{R}^2,\ g(x_1, x_2) = (x_1, x_2 - x_1).$$

The set S is convex and closed, the function f is proper, convex and lower semicontinuous, the function g is $\mathbb{R}^2_+$-convex and $\mathbb{R}^2_+$-epi closed and none of the interiority-type qualification conditions is fulfilled. On the other hand,

$$\bigcup_{\lambda \in \mathbb{R}^2_+} \operatorname{epi}(f + \langle \lambda, g(\cdot) \rangle + \delta_{\mathbb{R}^2_+})^* = \mathbb{R}^2 \times \mathbb{R}_+$$

and (B) is valid, i.e. for (P) and (D) one has strong duality.

Subdifferential formulae

- One always has:

$$\partial\left(f+\delta_{\mathcal{A}}\right)(x) \supseteq \bigcup_{\substack{\lambda \in K^*, \\ \langle \lambda, g(x) \rangle = 0}} \partial(f+\langle \lambda, g\rangle+\delta_S)(x) \ \forall x \in \operatorname{dom} f \cap \mathcal{A}.$$

- Each of the qualification conditions $(S), (R), (JW)$, (WS) and (B) guarantees (under corresponding topological assumptions) that

$$\partial\left(f+\delta_{\mathcal{A}}\right)(x) = \bigcup_{\substack{\lambda \in K^*, \\ \langle \lambda, g(x) \rangle = 0}} \partial(f+\langle \lambda, g\rangle+\delta_S)(x) \ \forall x \in \operatorname{dom} f \cap \mathcal{A}.$$

Generalized KKT optimality conditions for (P)

Assume that one of the qualification conditions $(S), (R), (JW)$, (WS) and (B) is (under corresponding topological assumptions) fulfilled. Then $\bar{x} \in X$ is an optimal solution to (P) if and only if there exists $\bar{\lambda} \in K^*$ such that

$$0 \in \partial(f+\langle \bar{\lambda}, g\rangle+\delta_S)(\bar{x})$$

and

$$\langle \bar{\lambda}, g(\bar{x}) \rangle = 0.$$

References

- H.H. Bauschke, P.-L. Combettes (2011): *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer-Verlag, New York
- J.M. Borwein, A.S. Lewis (2006): *Convex Analysis and Nonlinear Optimization*, Springer-Verlag, New York
- J.M. Borwein, J.D. Vanderwerff (2010): *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, New York
- R.I. Boţ (2010): *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer-Verlag, Berlin Heidelberg
- R.I. Boţ, S.-M. Grad, G. Wanka (2009): *Duality in Vector Optimization*, Springer-Verlag, Berlin Heidelberg
- I. Ekeland, R. Temam (1976): *Convex Analysis and Variational Problems*, North-Holland Publishing Company, Amsterdam
- R.T. Rockafellar (1970): *Convex Analysis*, Princeton University Press, Princeton
- C. Zălinescu (2002): *Convex Analysis in General Vector Spaces*, World Scientific, River Edge