APPLICATIONS OF CONVEX ANALYSIS WITHIN MATHEMATICS

Victoria Martín-Márquez

Fran Aragón*, Jon Borwein†, Liangjin Yao†

Grupo de análisis no lineal

Dpto. de Análisis Matemático
Universidad de Sevilla

UNIVERSIDAD DE SEVILLA

* Universidad de Alicante
† University of Newcastle, Australia

XII Encuentro de Análisis Funcional y Aplicaciones, Cáceres

Introduction

This talk is based on the paper:

Aragón, Borwein, Martín-Márquez, Yao
Applications of convex analysis within mathematics,
Math. Program., Ser B, December 2014, Volume 148, Issue 1, pp 49-88.

in a special issue to celebrate the 50th birthday of **Modern Convex Analysis and convex optimization** that became a tribute to the memory of Jean Jacques Moreau who passed away (on January 9, 2014) as the edition was being completed.

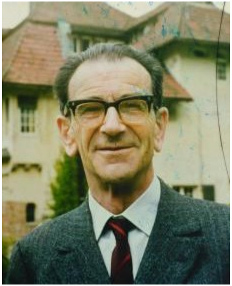
Introduction

The years 1962 – 1963 can be considered as birth date of modern convex analysis as the now familiar notions of **subdifferential, conjugate, proximal mappings, and infimal convolution** all date back to this period.

Introduction

The years 1962 – 1963 can be considered as birth date of modern convex analysis as the now familiar notions of **subdifferential, conjugate, proximal mappings, and infimal convolution** all date back to this period.

The development of convex analysis during the last fifty years owes much to



W. Fenchel (1905 – 1988)

J. J. Moreau (1923 – 2014)

R. T. Rockafellar (1935 –)

Introduction

Given a function $f : X \to (-\infty, +\infty]$, $x \in X$, various terms appeared in 1963 to name a vector s satisfying

$$f(x) + \langle y - x, s \rangle \leq f(y) \quad \forall y \in X.$$

Introduction

Given a function $f : X \to (-\infty, +\infty]$, $x \in X$, various terms appeared in 1963 to name a vector s satisfying

$$f(x) + \langle y - x, s \rangle \leq f(y) \quad \forall y \in X.$$

- In 1963 (Ph.D. thesis) Rockafellar called s “a **differential of f at x**”.

Introduction

Given a function $f : X \to (-\infty, +\infty]$, $x \in X$, various terms appeared in 1963 to name a vector s satisfying

$$f(x) + \langle y - x, s \rangle \leq f(y) \quad \forall y \in X.$$

- In 1963 (Ph.D. thesis) Rockafellar called s "a **differential of f at x**".
- At the same time Moreau coined the term "**sous-gradient**" which became "**subgradient**" in English, and investigated the properties of the associated set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x \rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

  term initially used by Moreau ("le sous-différentiel" in French)

Introduction

Given a function $f : X \to (-\infty, +\infty]$, $x \in X$, various terms appeared in 1963 to name a vector s satisfying

$$f(x) + \langle y - x, s \rangle \leq f(y) \quad \forall y \in X.$$

- In 1963 (Ph.D. thesis) Rockafellar called s "a **differential of f at x**".
- At the same time Moreau coined the term "**sous-gradient**" which became "**subgradient**" in English, and investigated the properties of the associated set-valued **subdifferential** operator ∂f:

  $$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x \rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

  term initially used by Moreau ("le sous-différentiel" in French)
- In the USSR, researchers were interested in similar concepts. For instance, in 1962, N. Z. Shor published the first instance of the use of a **subgradient method for minimizing a nonsmooth convex function**.

Introduction

The transformation $f \mapsto f^*$, where

$$f^* : X^* \to [-\infty, +\infty] : x^* \mapsto f^*(x^*) := \sup_{x \in X} \{\langle x^*, x\rangle - f(x)\}.$$

- has its origins in a publication of A. Legendre (1752 – 1833).

Introduction

The transformation $f \mapsto f^*$, where

$$f^* : X^* \to [-\infty, +\infty] : x^* \mapsto f^*(x^*) := \sup_{x \in X} \{\langle x^*, x \rangle - f(x)\}.$$

- has its origins in a publication of A. Legendre (1752 – 1833).
- Now is generally called Legendre-Fenchel transform or conjugate.

The transformation $f \mapsto f^*$, where

$$f^* : X^* \to [-\infty, +\infty] : x^* \mapsto f^*(x^*) := \sup_{x \in X} \{\langle x^*, x\rangle - f(x)\}.$$

- has its origins in a publication of A. Legendre (1752 – 1833).
- Now is generally called Legendre-Fenchel transform or conjugate.

The inf-convolution of two functions f and g is the function

$$f \Box g : X \to [-\infty, +\infty] : x \mapsto \inf_{y \in X} \{f(y) + g(x-y)\} = \inf_{u+v=x} \{f(u) + g(v)\}.$$

- key operation in modern convex analysis used by Fenchel.

Introduction

The transformation $f \mapsto f^*$, where

$$f^* : X^* \to [-\infty, +\infty] : x^* \mapsto f^*(x^*) := \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}.$$

- has its origins in a publication of A. Legendre (1752 – 1833).
- Now is generally called Legendre-Fenchel transform or conjugate.

The inf-convolution of two functions f and g is the function

$$f \Box g : X \to [-\infty, +\infty] : x \mapsto \inf_{y \in X}\{f(y) + g(x-y)\} = \inf_{u+v=x}\{f(u) + g(v)\}.$$

- key operation in modern convex analysis used by Fenchel.
- Moreau coined the term and use it in a more general setting.

Introduction

In a Hilbert space H, the proximal or proximity mapping is the operator

$$\operatorname{prox}_f : H \to H : x \mapsto \operatorname{prox}_f(x) := \operatorname*{argmin}_{y \in H} \left\{ f(y) + \frac{1}{2} \|x - y\|^2 \right\}.$$

Introduction

In a Hilbert space H, the proximal or proximity mapping is the operator

$$\operatorname{prox}_f : H \to H : x \mapsto \operatorname{prox}_f(x) := \underset{y \in H}{\operatorname{argmin}} \left\{ f(y) + \frac{1}{2}\|x - y\|^2 \right\}.$$

- These fundamental notions of **proximal mapping, subdifferential, conjugation, and inf- convolution** come together in Moreau's decomposition for a *proper lower semicontinuous convex function* f in a Hilbert space:

$$x = \operatorname{prox}_f(x) + \operatorname{prox}_{f^*}(x)$$

$$\frac{1}{2}\|\cdot\|^2 = f \Box \frac{1}{2}\|\cdot\|^2 + f^* \Box \frac{1}{2}\|\cdot\|^2$$

$$\operatorname{prox}_{f^*}(x) \in \partial(\operatorname{prox}_f(x)).$$

- Moreau's decomposition in terms of the proximal mapping is a powerful nonlinear analysis tool in the Hilbert setting that has been used in various areas of optimization and applied mathematics.

Context

X real Banach space

$f\colon X \to (-\infty, +\infty]$

- *proper* $(\operatorname{dom} f \neq \emptyset)$
- *convex* $(f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y),\ \forall x, y \in \operatorname{dom} f,\ \lambda \in [0,1])$
  $\Leftrightarrow \operatorname{epi} f$ is convex
- *lower-semicontinuous* (lsc) $(\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x})$ for all $\bar{x} \in X)$
  $\Leftrightarrow \operatorname{epi} f$ is closed.
- *Lipschitz* $(\exists\, M \geq 0$ so that $|f(x) - f(y)| \leq M\|x - y\|$ for all $x, y \in X)$

$\triangleright$ *epigraph* of f is $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} \mid f(x) \leq r\}$

Basic properties of convexity

1. (lsc) convex functions form a convex cone closed under pointwise suprema: f_γ convex (and lsc) $\forall \gamma \in \Gamma \Longrightarrow x \mapsto \sup_{\gamma \in \Gamma} f_\gamma(x)$ convex (and lsc).
2. Global minima and local minima in the domain coincide for proper convex functions.
3. f proper convex and $x \in \operatorname{dom} f$.
    - f locally Lipschitz at $x \Longleftrightarrow f$ continuous at $x \Longleftrightarrow f$ locally bounded at x.
    - f lower semicontinuous $\Longrightarrow f$ continuous at every point in $\operatorname{int} \operatorname{dom} f$.
4. A proper lower semicontinuous and convex function is bounded from below by a continuous affine function.
5. If C is a nonempty set, then $\mathrm{d}_C(\cdot)$ is non-expansive (Lipschitz function with constant one). Additionally, if C is convex, then $\mathrm{d}_C(\cdot)$ is convex.

Basic properties of subdifferential

Set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y-x\rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

- ∂f may be empty $\left(\text{example: } \partial f(0) = \varnothing \text{ for } f(x) = \begin{cases} -\sqrt{x} & x \geq 0 \\ +\infty & \text{otherwise} \end{cases}\right)$

Basic properties of subdifferential

Set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x\rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

Basic properties of subdifferential

Set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x\rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

- f proper, convex, lsc and **Gâteaux dif.** at $\bar{x} \in \operatorname{dom} f \implies \partial f(\bar{x}) = \nabla f$

Basic properties of subdifferential

Set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x \rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

- f proper, convex, lsc and **Gâteaux dif.** at $\bar{x} \in \operatorname{dom} f \Longrightarrow \partial f(\bar{x}) = \nabla f$
- Fundamental significance of **subgradients** in optimization:

Subdifferential at optimality

$f : X \to]-\infty, +\infty]$ proper convex

$$\bar{x} \in \operatorname{dom} f \text{ is a (global) minimizer of } f \Longleftrightarrow 0 \in \partial f(\bar{x}).$$

Basic properties of subdifferential

Set-valued **subdifferential** operator ∂f:

$$\partial f : X \rightrightarrows X^* : x \mapsto \{x^* \in X^* \mid \langle x^*, y - x\rangle \leq f(y) - f(x), \text{ for all } y \in X\}.$$

- f proper, convex, lsc and **Gâteaux dif.** at $\bar{x} \in \operatorname{dom} f \implies \partial f(\bar{x}) = \nabla f$
- Fundamental significance of **subgradients** in optimization:

Subdifferential at optimality

$f : X \to]-\infty, +\infty]$ proper convex

$$\bar{x} \in \operatorname{dom} f \text{ is a (global) minimizer of } f \iff 0 \in \partial f(\bar{x}).$$

- Relationship between **subgradients** and **directional derivatives**

Moreau's max formula

$f : X \to]-\infty, +\infty]$ convex and continuous at $\bar{x}$. $d \in X$. Then $\partial f(\bar{x}) \neq \emptyset$ and

$$f'(\bar{x}; d) := \lim_{t \to 0^+} \frac{f(\bar{x} + td) - f(\bar{x})}{t} = \max\{\langle x^*, d\rangle \mid x^* \in \partial f(\bar{x})\}.$$

Basic properties of conjugate

Fenchel conjugate (*Legendre-Fenchel transform or conjugate*)

$$f^* : X^* \to [-\infty, +\infty] : x \mapsto f^*(x^*) := \sup_{x\in X}\{\langle x^*, x\rangle - f(x)\}.$$

► By direct construction and Property 1 of convexity, for any function f, the conjugate function f^* is <u>always convex and lower semicontinuous.</u>

Basic properties of conjugate

Fenchel conjugate (*Legendre-Fenchel transform or conjugate*)

$$f^* : X^* \to [-\infty, +\infty] : x \mapsto f^*(x^*) := \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}.$$

► By direct construction and Property 1 of convexity, for any function f, the conjugate function f^* is always convex and lower semicontinuous.
► It plays a role in *convex analysis* in many ways analogous to the role played by the **Fourier transform** in *harmonic analysis* with infimal convolution replacing integral convolution

Basic properties of conjugate

Fenchel conjugate (*Legendre-Fenchel transform or conjugate*)

$$f^*: X^* \to [-\infty, +\infty] : x \mapsto f^*(x^*) := \sup_{x \in X}\{\langle x^*, x\rangle - f(x)\}.$$

► By direct construction and Property 1 of convexity, for any function f, the conjugate function f^* is always convex and lower semicontinuous.
► It plays a role in *convex analysis* in many ways analogous to the role played by the **Fourier transform** in *harmonic analysis* with infimal convolution replacing integral convolution

Fenchel–Young inequality

$f: X \to]-\infty, +\infty]$, $x^* \in X^*$ and $x \in \operatorname{dom} f$:

$$f(x) + f^*(x^*) \geq \langle x^*, x\rangle.$$

Equality holds if and only if $x^* \in \partial f(x)$.

Basic properties of conjugate

Example: $f(x) := \frac{\|x\|^p}{p} \quad (1 < p < \infty) \quad \Longrightarrow \quad f^*(x^*) = \frac{\|x^*\|_*^q}{q} \quad (\frac{1}{p} + \frac{1}{q} = 1).$

$$f^*(x^*) = \sup_{\lambda \in \mathbb{R}_+} \sup_{\|x\|=1} \left\{ \langle x^*, \lambda x \rangle - \frac{\|\lambda x\|^p}{p} \right\} = \sup_{\lambda \in \mathbb{R}_+} \left\{ \lambda \|x^*\|_* - \frac{\lambda^p}{p} \right\} = \frac{\|x^*\|_*^q}{q}.$$

Basic properties of conjugate

Example: $f(x) := \frac{\|x\|^p}{p} \quad (1 < p < \infty) \quad \Longrightarrow \quad f^*(x^*) = \frac{\|x^*\|_*^q}{q} \quad (\frac{1}{p} + \frac{1}{q} = 1).$

$$f^*(x^*) = \sup_{\lambda \in \mathbb{R}_+} \sup_{\|x\|=1} \left\{ \langle x^*, \lambda x \rangle - \frac{\|\lambda x\|^p}{p} \right\} = \sup_{\lambda \in \mathbb{R}_+} \left\{ \lambda \|x^*\|_* - \frac{\lambda^p}{p} \right\} = \frac{\|x^*\|_*^q}{q}.$$

From Fenchel–Young inequality: $\boxed{\frac{\|x\|^p}{p} + \frac{\|x^*\|_*^q}{q} \geq \langle x^*, x \rangle,}$

When $X = \mathbb{R}$ one recovers the original Young inequality.

Basic properties of conjugate

Example: $f(x) := \frac{\|x\|^p}{p} \quad (1 < p < \infty) \implies f^*(x^*) = \frac{\|x^*\|_*^q}{q} \quad (\frac{1}{p} + \frac{1}{q} = 1).$

$$f^*(x^*) = \sup_{\lambda \in \mathbb{R}_+} \sup_{\|x\|=1} \left\{ \langle x^*, \lambda x \rangle - \frac{\|\lambda x\|^p}{p} \right\} = \sup_{\lambda \in \mathbb{R}_+} \left\{ \lambda \|x^*\|_* - \frac{\lambda^p}{p} \right\} = \frac{\|x^*\|_*^q}{q}.$$

From Fenchel–Young inequality: $\boxed{\frac{\|x\|^p}{p} + \frac{\|x^*\|_*^q}{q} \geq \langle x^*, x \rangle,}$

When $X = \mathbb{R}$ one recovers the original Young inequality.

Biconjugate: $f^{**} = (f^*)^*$ defined on $X^{**} \implies f^{**}|_X \leq f$.

Basic properties of conjugate

Example: $f(x) := \frac{\|x\|^p}{p} \quad (1 < p < \infty) \quad \Longrightarrow \quad f^*(x^*) = \frac{\|x^*\|_*^q}{q} \quad (\frac{1}{p} + \frac{1}{q} = 1).$

$$f^*(x^*) = \sup_{\lambda \in \mathbb{R}_+} \sup_{\|x\|=1} \left\{ \langle x^*, \lambda x \rangle - \frac{\|\lambda x\|^p}{p} \right\} = \sup_{\lambda \in \mathbb{R}_+} \left\{ \lambda \|x^*\|_* - \frac{\lambda^p}{p} \right\} = \frac{\|x^*\|_*^q}{q}.$$

From Fenchel–Young inequality: $\boxed{\frac{\|x\|^p}{p} + \frac{\|x^*\|_*^q}{q} \geq \langle x^*, x \rangle,}$

When $X = \mathbb{R}$ one recovers the original Young inequality.

Biconjugate: $f^{**} = (f^*)^*$ defined on $X^{**} \Longrightarrow f^{**}|_X \leq f$.

Hormander (Legendre, Fenchel, Moreau)

$f : X \to]-\infty, +\infty]$ proper:

$$f \text{ convex and lsc} \quad \Longleftrightarrow \quad f = f^{**}|X$$

Basic properties of conjugate

Example: $f(x) := \frac{\|x\|^p}{p} \quad (1 < p < \infty) \implies f^*(x^*) = \frac{\|x^*\|_*^q}{q} \quad (\frac{1}{p} + \frac{1}{q} = 1).$

$$f^*(x^*) = \sup_{\lambda \in \mathbb{R}_+} \sup_{\|x\|=1} \left\{ \langle x^*, \lambda x \rangle - \frac{\|\lambda x\|^p}{p} \right\} = \sup_{\lambda \in \mathbb{R}_+} \left\{ \lambda \|x^*\|_* - \frac{\lambda^p}{p} \right\} = \frac{\|x^*\|_*^q}{q}.$$

From Fenchel–Young inequality: $\boxed{\frac{\|x\|^p}{p} + \frac{\|x^*\|_*^q}{q} \geq \langle x^*, x \rangle,}$

When $X = \mathbb{R}$ one recovers the original Young inequality.

Biconjugate: $f^{**} = (f^*)^*$ defined on $X^{**} \implies f^{**}|_X \leq f.$

Hormander (Legendre, Fenchel, Moreau)

$f : X \to]-\infty, +\infty]$ proper:

$$f \text{ convex and lsc} \iff f = f^{**}|X$$

Application in establishing convexity (to compute conjugates: SCAT Maple software)

Basic properties of infimal convolution

The inf-convolution of f and g:

$$f \Box g : X \to [-\infty, +\infty] : x \mapsto \inf_{y \in X} \left\{ f(y) + g(x-y) \right\} = \inf_{u+v=x} \left\{ f(u) + g(v) \right\}.$$

The largest extended real-valued function whose epigraph contains the sum of epigraphs of f and g $\implies$ $f \Box g$ is convex.

Basic properties of infimal convolution

The inf-convolution of f and g:

$$f\Box g : X \to [-\infty, +\infty] : x \mapsto \inf_{y\in X} \{f(y) + g(x-y)\} = \inf_{u+v=x} \{f(u) + g(v)\}.$$

The largest extended real-valued function whose epigraph contains the sum of epigraphs of f and g $\implies$ $f\Box g$ is convex.

- f, g proper $\implies$ $(f\Box g)^* = f^* + g^*$

Basic properties of infimal convolution

The inf-convolution of f and g:

$$f\Box g : X \to [-\infty,+\infty] : x \mapsto \inf_{y\in X}\{f(y)+g(x-y)\} = \inf_{u+v=x}\{f(u)+g(v)\}.$$

The largest extended real-valued function whose epigraph contains the sum of epigraphs of f and g $\implies$ $f\Box g$ is convex.

- f,g proper $\implies$ $(f\Box g)^* = f^* + g^*$

Example:

$$f(x) := \begin{cases} -\sqrt{1-x^2}, & \text{for } -1 \le x \le 1,\\ +\infty & \text{otherwise,}\end{cases}$$

$$g(x) := |x|$$

The inf-convolution of f and g:

$$f\Box g : X \to [-\infty,+\infty] : x \mapsto \inf_{y\in X}\{f(y)+g(x-y)\} = \inf_{u+v=x}\{f(u)+g(v)\}.$$

The largest extended real-valued function whose epigraph contains the sum of epigraphs of f and g $\Longrightarrow$ $f\Box g$ is convex.

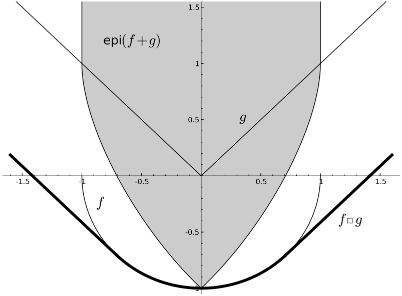
- f, g proper $\Longrightarrow$ $(f\Box g)^* = f^* + g^*$

Example:

$$f(x) := \begin{cases} -\sqrt{1-x^2}, & \text{for } -1 \le x \le 1, \\ +\infty & \text{otherwise,} \end{cases}$$

$$g(x) := |x|$$

$$(f\Box g)(x) = \begin{cases} -\sqrt{1-x^2}, & -\frac{\sqrt{2}}{2} \le x \le -\frac{\sqrt{2}}{2} \\ |x|-\sqrt{2}, & \text{otherwise} \end{cases}$$



Fenchel duality theorem

Fenchel duality theorem

X, Y Banach spaces, $f\colon X \to]-\infty, +\infty]$ and $g\colon Y \to]-\infty, +\infty]$ convex
$T\colon X \to Y$ bounded linear operator

$$p := \inf_{x \in X} \{f(x) + g(Tx)\} \qquad \text{primal problem}$$

$$d := \sup_{y^* \in Y^*} \{-f^*(T^* y^*) - g^*(-y^*)\} \qquad \text{dual problem}$$

Then $\boxed{p \geq d}$ weak duality inequality

Fenchel duality theorem

X, Y Banach spaces, $f\colon X \to\,]-\infty, +\infty]$ and $g\colon Y \to\,]-\infty, +\infty]$ convex
$T\colon X \to Y$ bounded linear operator

$$p := \inf_{x\in X}\{f(x) + g(Tx)\} \qquad \text{primal problem}$$

$$d := \sup_{y^*\in Y^*}\{-f^*(T^*y^*) - g^*(-y^*)\} \qquad \text{dual problem}$$

Then $\boxed{p \geq d}$ weak duality inequality

Suppose further that f, g and Tsatisfy either

$$\bigcup_{\lambda>0} \lambda\,[\operatorname{dom} g - T\operatorname{dom} f] = Y \text{ and both } f \text{ and } g \text{ lsc} \qquad \text{CQ1}$$

or the condition $\operatorname{cont} g \cap T\operatorname{dom} f \neq \emptyset$ CQ2

Then $\boxed{p = d}$ and the supremum in d is attained when finite.

Consequences of Fenchel duality

Consequences of Fenchel duality

Under the hypotheses of the Fenchel duality theorem

$$\boxed{(f+g)^*(x^*) = (f^* \Box g^*)(x^*)}$$

with attainment when finite

Consequences of Fenchel duality

Under the hypotheses of the Fenchel duality theorem

$$\boxed{(f+g)^*(x^*) = (f^* \Box g^*)(x^*)}$$

with attainment when finite

Obtaining primal solutions from dual ones

If the conditions for equality in the Fenchel duality Theorem hold, and $\bar{y}^* \in Y^*$ is an optimal dual solution:

$$\bar{x} \in X \text{ optimal for primal problem} \iff \begin{cases} T^*\bar{y}^* \in \partial f(\bar{x}) \\ -\bar{y}^* \in \partial g(T\bar{x}) \end{cases}$$

Consequences of Fenchel duality

Under the hypotheses of the Fenchel duality theorem

$$\boxed{(f+g)^*(x^*) = (f^* \Box g^*)(x^*)}$$

with attainment when finite

Obtaining primal solutions from dual ones

If the conditions for equality in the Fenchel duality Theorem hold, and $\bar{y}^* \in Y^*$ is an optimal dual solution:

$$\bar{x} \in X \text{ optimal for primal problem} \iff \begin{cases} T^*\bar{y}^* \in \partial f(\bar{x}) \\ -\bar{y}^* \in \partial g(T\bar{x}) \end{cases}$$

Extended sandwich theorem

f, g and T as in Fenchel duality theorem. If $f \geq -g \circ T$ then: $\exists\ \alpha : X \to \mathbb{R}$

$$\boxed{f \geq \alpha \geq -g \circ T} \qquad \left(\alpha(x) = \langle T^* y^*, x\rangle + r \text{ where } \bar{y}^* \in Y^* \text{ is an optimal dual solution}\right)$$

Moreover, for any $\bar{x}$ satisfying $f(\bar{x}) = (-g \circ T)(\bar{x})$, we have $-y^* \in \partial g(T\bar{x})$.

When constraint qualifications are not satisfied

When constraint qualifications are not satisfied

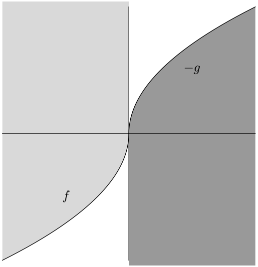
Examples:

$$f(x) := \begin{cases} -\sqrt{-x}, & \text{for } x \leq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

$$g(x) := \begin{cases} -\sqrt{x}, & \text{for } x \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

$\bigcup_{\lambda>0} \lambda\,[\operatorname{dom} g - \operatorname{dom} f] = [0, +\infty[\neq \mathbb{R}$

$\nexists\ \alpha$ separating f and $-g$



When constraint qualifications are not satisfied

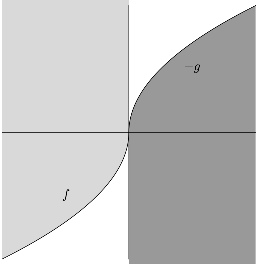
Examples:

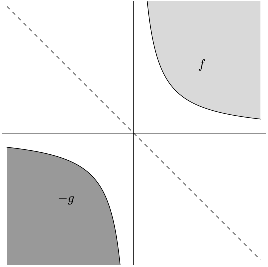
$$f(x) := \begin{cases} -\sqrt{-x}, & \text{for } x \leq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

$$g(x) := \begin{cases} -\sqrt{x}, & \text{for } x \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

$\bigcup_{\lambda>0} \lambda\, [\operatorname{dom} g - \operatorname{dom} f] = [0, +\infty[\, \neq \mathbb{R}$

$\nexists\, \alpha$ separating f and $-g$





$$f(x) := \begin{cases} \frac{1}{x}, & \text{for } x > 0, \\ +\infty & \text{otherwise,} \end{cases}$$

$$g(x) := \begin{cases} -\frac{1}{x}, & \text{for } x < 0, \\ +\infty & \text{otherwise.} \end{cases}$$

$\bigcup_{\lambda>0} \lambda\, [\operatorname{dom} g - \operatorname{dom} f] = \,]-\infty, 0[\, \neq \mathbb{R}$

$\alpha(x) := -x$ satisfies $f \geq \alpha \geq -g$

Consequences of Fenchel duality

Subdifferential Sum rule

f, g and T as in Fenchel duality theorem

- *without constraint qualifications*:

$$\partial(f + g \circ T)(x) \supseteq \partial f(x) + T^*(\partial g(Tx))$$

- *with a constraint qualification*:

$$\partial(f + g \circ T)(x) = \partial f(x) + T^*(\partial g(Tx))$$

Consequences of Fenchel duality

Subdifferential Sum rule

f, g and T as in Fenchel duality theorem

- *without constraint qualifications*:

$$\partial(f + g \circ T)(x) \supseteq \partial f(x) + T^*(\partial g(Tx))$$

- *with a constraint qualification*:

$$\partial(f + g \circ T)(x) = \partial f(x) + T^*(\partial g(Tx))$$

Hahn-Banach extension

$f : X \to \mathbb{R}$ continuous sublinear function with $\mathrm{dom} f = X$
L linear subspace of Banach space X
and $h : L \to \mathbb{R}$ linear and *dominated* by f ($f \geq h$) on L.

Then $\exists\, x^* \in X^*$ dominated by f on X such that

$$h(x) = \langle x^*, x \rangle, \text{ for all } x \in L.$$

Consequences of Fenchel duality

Remark:

$$\left.\begin{array}{l} non-emptiness\ of\ the\ subdifferential\ at\ a\ point\ of\ continuity \\ Moreau's\ max\ formula \\ Fenchel\ duality \\ Sandwich\ theorem \\ subdifferential\ sum\ rule \\ Hahn-Banach\ extension\ theorem \end{array}\right\} \textbf{equivalent}$$

in the sense that they are easily inter-derivable.

Consequences of Fenchel duality

Remark:

$$\left.\begin{array}{l} non-emptiness\ of\ the\ subdifferential\ at\ a\ point\ of\ continuity \\ Moreau's\ max\ formula \\ Fenchel\ duality \\ Sandwich\ theorem \\ subdifferential\ sum\ rule \\ Hahn-Banach\ extension\ theorem \end{array}\right\} \textbf{equivalent}$$

in the sense that they are easily inter-derivable.

More consequences of Fenchel duality:

- **Existence of Banach limits**
- **Chebyshev problem**:

  C weakly closed subset of a Hilbert space H

$$C \text{ convex} \iff C \text{ is a Chebyshev set.}$$

Monotone operator theory

$A : X \rightrightarrows X^*$ *set-valued operator* $\quad (\forall x \in X, Ax \subseteq X^*)$

graph of A: $\quad \operatorname{gra} A := \left\{(x,x^*) \in X \times X^* \mid x^* \in Ax\right\}$

domain of A: $\quad \operatorname{dom} A := \left\{x \in X \mid Ax \neq \varnothing\right\}$

range of A: $\quad \operatorname{ran} A := A(X)$

- A is *monotone* if $\langle x - y, x^* - y^* \rangle \geq 0, \quad$ for all $(x,x^*),(y,y^*) \in \operatorname{gra} A$
- A is *maximal monotone* if A is monotone and A has no proper monotone extension (in the sense of graph inclusion)

Monotone operator theory

$A : X \rightrightarrows X^*$ *set-valued operator* $\quad (\forall x \in X, Ax \subseteq X^*)$

graph of A: $\quad \operatorname{gra} A := \{(x,x^*) \in X \times X^* \mid x^* \in Ax\}$

domain of A: $\quad \operatorname{dom} A := \{x \in X \mid Ax \neq \varnothing\}$

range of A: $\quad \operatorname{ran} A := A(X)$

- A is *monotone* if $\langle x - y, x^* - y^* \rangle \geq 0, \quad$ for all $(x,x^*),(y,y^*) \in \operatorname{gra} A$
- A is *maximal monotone* if A is monotone and A has no proper monotone extension (in the sense of graph inclusion)

Minty 1962 (Extension to reflexive spaces by Rockafellar)

$A : H \rightrightarrows H$ monotone in a Hilbert space H

$$A \text{ maximal monotone} \iff \operatorname{ran}(A + \mathrm{Id}) = H$$

Monotone operator theory

$A : X \rightrightarrows X^*$ *set-valued operator* $\quad (\forall x \in X, Ax \subseteq X^*)$

graph of A: $\quad \operatorname{gra} A := \{(x,x^*) \in X \times X^* \mid x^* \in Ax\}$

domain of A: $\quad \operatorname{dom} A := \{x \in X \mid Ax \neq \varnothing\}$

range of A: $\quad \operatorname{ran} A := A(X)$

- A is *monotone* if $\langle x - y, x^* - y^* \rangle \geq 0, \quad$ for all $(x,x^*), (y,y^*) \in \operatorname{gra} A$
- A is *maximal monotone* if A is monotone and A has no proper monotone extension (in the sense of graph inclusion)

Minty 1962 (Extension to reflexive spaces by Rockafellar)

$A : H \rightrightarrows H$ monotone in a Hilbert space H

$$A \text{ maximal monotone} \quad \Longleftrightarrow \quad \operatorname{ran}(A + \operatorname{Id}) = H$$

Sum theorem (Rockafellar 1970, ...)

X reflexive Banach space.

$$\left.\begin{array}{l} A, B : X \rightrightarrows X \text{ maximal monotone} \\ \bigcup_{\lambda > 0} \lambda\, [\operatorname{dom} A - \operatorname{dom} B] \text{ closed subspace} \end{array}\right\} \Longrightarrow \begin{array}{c} A + B \\ \text{maximal monotone} \end{array}$$

Monotone operator theory

The *Fitzpatrick function* associated with A is $F_A : X \times X^* \to \,]-\infty, +\infty]$

$$F_A(x,x^*) := \sup_{(a,a^*)\in \operatorname{gra} A} \Big(\langle x, a^* \rangle + \langle a, x^* \rangle - \langle a, a^* \rangle \Big).$$

Monotone operator theory

The *Fitzpatrick function* associated with A is $F_A : X \times X^* \to \left]-\infty, +\infty\right]$

$$F_A(x,x^*) := \sup_{(a,a^*)\in \operatorname{gra} A} \Big(\langle x, a^* \rangle + \langle a, x^* \rangle - \langle a, a^* \rangle \Big).$$

$A : X \rightrightarrows X^*$ monotone with $\operatorname{dom} A \neq \varnothing$. Then:
F_A proper, convex, lsc in the norm $\times$ weak*-topology $\omega(X^*, X)$, and

$$\langle x, x^* \rangle = F_A(x,x^*) \ \forall (x,x^*) \in \operatorname{gra} A.$$

If A maximal monotone: $\langle x, x^* \rangle \leq F_A(x,x^*) \leq F_A^*(x^*,x)$, $\forall (x,x^*) \in X \times X^*$

Monotone operator theory

The *Fitzpatrick function* associated with A is $F_A : X \times X^* \to \,]-\infty, +\infty]$

$$F_A(x,x^*) := \sup_{(a,a^*)\in \operatorname{gra} A} \Big(\langle x, a^* \rangle + \langle a, x^* \rangle - \langle a, a^* \rangle \Big).$$

$A : X \rightrightarrows X^*$ monotone with $\operatorname{dom} A \neq \varnothing$. Then:
F_A proper, convex, lsc in the norm $\times$ weak*-topology $\omega(X^*, X)$, and

$$\langle x, x^* \rangle = F_A(x,x^*) \; \forall (x,x^*) \in \operatorname{gra} A.$$

If A maximal monotone: $\langle x, x^* \rangle \leq F_A(x,x^*) \leq F_A^*(x^*,x)$, $\forall (x,x^*) \in X \times X^*$

$F : X \times X^* \to \,]-\infty, +\infty]$

- *autoconjugate* if $F(x,x^*) = F^*(x^*,x)$, $\forall (x,x^*) \in X \times X^*$
- *representer* for A if $\operatorname{gra} A = \big\{ (x,x^*) \in X \times X^* \mid F(x,x^*) = \langle x, x^* \rangle \big\}$

Monotone operator theory

The *Fitzpatrick function* associated with A is $F_A : X \times X^* \to]-\infty, +\infty]$

$$F_A(x,x^*) := \sup_{(a,a^*)\in \operatorname{gra} A} \Big(\langle x, a^* \rangle + \langle a, x^* \rangle - \langle a, a^* \rangle \Big).$$

$A : X \rightrightarrows X^*$ monotone with $\operatorname{dom} A \neq \varnothing$. Then:
F_A proper, convex, lsc in the norm $\times$ weak*-topology $\omega(X^*, X)$, and

$$\langle x, x^* \rangle = F_A(x,x^*) \ \forall (x,x^*) \in \operatorname{gra} A.$$

If A maximal monotone: $\langle x,x^* \rangle \leq F_A(x,x^*) \leq F_A^*(x^*,x), \ \forall (x,x^*) \in X \times X^*$

$F : X \times X^* \to]-\infty, +\infty]$

- *autoconjugate* if $F(x,x^*) = F^*(x^*,x), \ \forall (x,x^*) \in X \times X^*$
- *representer* for A if $\operatorname{gra} A = \big\{ (x,x^*) \in X \times X^* \mid F(x,x^*) = \langle x,x^* \rangle \big\}$

If $A : X \rightrightarrows X^*$ is maximally monotone, does there necessarily exist an autoconjugate representer for A? Fitzpatrick 1988

Monotone operator theory

Bauschke, Wang (2009) gave an affirmative answer in reflexive spaces by construction of the function $\mathscr{B}_A : X \times X^* \to \left]-\infty, +\infty\right]$

$$\mathscr{B}_A(x,x^*) = \inf_{(y,y^*)\in X\times X^*} \left\{ \tfrac{1}{2}F_A(x+y,x^*+y^*) + \tfrac{1}{2}F_A^*(x^*-y^*,x-y) + \tfrac{1}{2}\|y\|^2 + \tfrac{1}{2}\|y^*\|^2 \right\}$$

Monotone operator theory

Bauschke, Wang (2009) gave an affirmative answer in reflexive spaces by construction of the function $\mathscr{B}_A : X \times X^* \to \,]-\infty, +\infty]$

$$\mathscr{B}_A(x,x^*) = \inf_{(y,y^*)\in X\times X^*} \left\{ \tfrac{1}{2}F_A(x+y,x^*+y^*) + \tfrac{1}{2}F_A^*(x^*-y^*,x-y) + \tfrac{1}{2}\|y\|^2 + \tfrac{1}{2}\|y^*\|^2 \right\}$$

Is $\mathscr{B}_A$ still an autoconjugate representer for a maximally monotone operator A in a general Banach space?

Monotone operator theory

Bauschke, Wang (2009) gave an affirmative answer in reflexive spaces by construction of the function $\mathscr{B}_A : X \times X^* \to \,]-\infty, +\infty]$

$$\mathscr{B}_A(x,x^*) = \inf_{(y,y^*)\in X\times X^*} \left\{ \tfrac{1}{2} F_A(x+y, x^*+y^*) + \tfrac{1}{2} F_A^*(x^*-y^*, x-y) + \tfrac{1}{2}\|y\|^2 + \tfrac{1}{2}\|y^*\|^2 \right\}$$

Is $\mathscr{B}_A$ still an autoconjugate representer for a maximally monotone operator A in a general Banach space?

We give a negative answer

Monotone operator theory

Examples: $\mathscr{B}_A$ **is not always autoconjugate**

$X := c_0$ with $\|\cdot\|_\infty$ so that $X^* = \ell^1$ with $\|\cdot\|_1$ and $X^{**} = \ell^\infty$ with $\|\cdot\|_*$.
Fix $\alpha := (\alpha_n)_{n\in\mathbb{N}} \in \ell^\infty$ with $\limsup \alpha_n \neq 0$ and $\|\alpha\|_* < \frac{1}{\sqrt{2}}$, and define $A_\alpha : \ell^1 \to \ell^\infty$:

$$(A_\alpha x^*)_n := \alpha_n^2 x_n^* + 2\sum_{i>n} \alpha_n \alpha_i x_i^*, \quad \forall x^* = (x_n^*)_{n\in\mathbb{N}} \in \ell^1.$$

Let $T_\alpha : c_0 \rightrightarrows X^*$ be defined by

$$\begin{aligned} \operatorname{gra} T_\alpha &:= \big\{(-A_\alpha x^*, x^*) \mid x^* \in X^*, \langle \alpha, x^* \rangle = 0\big\} \\ &= \Big\{\Big(\big(-\sum_{i>n} \alpha_n \alpha_i x_i^* + \sum_{i<n} \alpha_n \alpha_i x_i^*\big)_{n\in\mathbb{N}}, x^*\Big) \mid x^* \in X^*, \langle \alpha, x^* \rangle = 0\Big\}. \end{aligned}$$

Then

$$\mathscr{B}_{T_\alpha}(-Aa^*, a^*) > \mathscr{B}^*_{T_\alpha}(a^*, -Aa^*), \quad \forall a^* \notin \{e\}_\perp.$$

In consequence, $\mathscr{B}_{T_\alpha}$ is not autoconjugate.

More to read in the paper...

- Convex functions and maximal monotone operators.
- Symbolic convex analysis.
- Asplund averaging: existence of equivalent norms.
- Convexity and partial fractions

More to read in the paper...

- Convex functions and maximal monotone operators.
- Symbolic convex analysis.
- Asplund averaging: existence of equivalent norms.
- Convexity and partial fractions

THANKS YOU

Australia, December 2013