Topics in Convex Analysis

Tim Hoheisel (McGill University, Montreal)

Spring School on Variational Analysis

Paseky, May 19–25, 2019

Outline

1. Fundamentals from Convex Analysis

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle\cdot,\cdot\rangle : \mathbb{E}\times\mathbb{E}\to\mathbb{R}$ of dimension $\kappa<\infty$.

Examples

- $\mathbb{E}=\mathbb{R}^n, \quad \langle x, y\rangle := x^T y, \quad \kappa = n$
- $\mathbb{E}=\mathbb{R}^{m\times n}, \quad \langle A, B\rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A\subset\mathbb{E}$

- $A+B := \{a+b \mid a\in A,\ b\in B\} \quad (B\subset\mathbb{E})$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A \subset \mathbb{E}$

- $A + B := \{a + b \mid a \in A,\ b \in B\} \quad (B \subset \mathbb{E})$
- $A + x := A + \{x\} \quad (x \in \mathbb{E})$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A \subset \mathbb{E}$

- $A + B := \{a + b \mid a \in A,\ b \in B\} \quad (B \subset \mathbb{E})$
- $A + x := A + \{x\} \quad (x \in \mathbb{E})$
- $\Lambda \cdot A := \{\lambda a \mid a \in A,\ \lambda \in \Lambda\} \quad (\Lambda \subset \mathbb{R})$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A \subset \mathbb{E}$

- $A + B := \{a + b \mid a \in A,\ b \in B\} \quad (B \subset \mathbb{E})$
- $A + x := A + \{x\} \quad (x \in \mathbb{E})$
- $\Lambda \cdot A := \{\lambda a \mid a \in A,\ \lambda \in \Lambda\} \quad (\Lambda \subset \mathbb{R})$
- $\lambda A := \{\lambda\} \cdot A \ (\lambda \in \mathbb{R})$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A \subset \mathbb{E}$

- $A + B := \{a + b \mid a \in A,\ b \in B\} \quad (B \subset \mathbb{E})$
- $A + x := A + \{x\} \quad (x \in \mathbb{E})$
- $\Lambda \cdot A := \{\lambda a \mid a \in A,\ \lambda \in \Lambda\} \quad (\Lambda \subset \mathbb{R})$
- $\lambda A := \{\lambda\} \cdot A \ (\lambda \in \mathbb{R})$

Examples:

- $U, V \subset \mathbb{E}$ subspaces. Then $U + V = \mathrm{span}\,(U \cup V)$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle\cdot,\cdot\rangle:\mathbb{E}\times\mathbb{E}\to\mathbb{R}$ of dimension $\kappa<\infty$.

Examples

- $\mathbb{E}=\mathbb{R}^n, \quad \langle x,y\rangle := x^Ty, \quad \kappa=n$
- $\mathbb{E}=\mathbb{R}^{m\times n}, \quad \langle A,B\rangle := \mathrm{tr}\,(A^TB), \quad \kappa=mn$

Minkowski addition/multiplication: Let $A\subset\mathbb{E}$

- $A+B := \{a+b \mid a\in A,\ b\in B\} \quad (B\subset\mathbb{E})$
- $A+x := A+\{x\} \quad (x\in\mathbb{E})$
- $\Lambda\cdot A := \{\lambda a \mid a\in A,\ \lambda\in\Lambda\} \quad (\Lambda\subset\mathbb{R})$
- $\lambda A := \{\lambda\}\cdot A \ (\lambda\in\mathbb{R})$

Examples:

- $U,V\subset\mathbb{E}$ subspaces. Then $U+V=\mathrm{span}\,(U\cup V)$
- $B_\varepsilon(x)=x+\varepsilon\mathbb{B}$

The Euclidean setting and Minkowski notation

In what follows $\mathbb{E}$ will be a **Euclidean space**, i.e. a *real-vector space* equipped with an *inner product* $\langle \cdot, \cdot \rangle : \mathbb{E} \times \mathbb{E} \to \mathbb{R}$ of dimension $\kappa < \infty$.

Examples

- $\mathbb{E} = \mathbb{R}^n, \quad \langle x, y \rangle := x^T y, \quad \kappa = n$
- $\mathbb{E} = \mathbb{R}^{m \times n}, \quad \langle A, B \rangle := \mathrm{tr}\,(A^T B), \quad \kappa = mn$

Minkowski addition/multiplication: Let $A \subset \mathbb{E}$

- $A + B := \{a + b \mid a \in A,\ b \in B\} \quad (B \subset \mathbb{E})$
- $A + x := A + \{x\} \quad (x \in \mathbb{E})$
- $\Lambda \cdot A := \{\lambda a \mid a \in A,\ \lambda \in \Lambda\} \quad (\Lambda \subset \mathbb{R})$
- $\lambda A := \{\lambda\} \cdot A \ (\lambda \in \mathbb{R})$

Examples:

- $U, V \subset \mathbb{E}$ subspaces. Then $U + V = \mathrm{span}\,(U \cup V)$
- $B_\varepsilon(x) = x + \varepsilon \mathbb{B}$
- $\mathrm{pos}\ S := \mathbb{R}_+ S \quad$ (conical hull)

Convex sets and cones

"The great watershed in optimization is not between linearity and nonlinearity, but convexity and nonconvexity." (R.T. Rockafellar, *1935)

Convex sets and cones

"The great watershed in optimization is not between linearity and nonlinearity, but convexity and nonconvexity." (R.T. Rockafellar, *1935)

$S \subset \mathbb{E}$ is said to be

- *convex* if $\lambda S + (1-\lambda)S \subset S \quad (\lambda \in (0,1))$;
- a *cone* if $\lambda S \subset S \quad (\lambda \geq 0)$.

Note that $K \subset \mathbb{E}$ is a *convex cone* iff $K + K \subset K$.

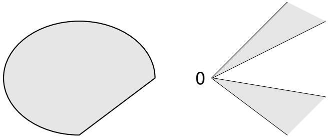


Figure: Convex set/non-convex cone

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\operatorname{conv} S := \bigcap \{ C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex} \}.$$

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\operatorname{conv} S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex}\}.$$

The *closed convex hull* of S is the smallest closed, convex set containing S, i.e.

$$\overline{\operatorname{conv}}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ closed and convex}\}.$$

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\mathrm{conv}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex}\}.$$

The *closed convex hull* of S is the smallest closed, convex set containing S, i.e.

$$\overline{\mathrm{conv}}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ closed and convex}\}.$$

- $\overline{\mathrm{conv}}\, S = \mathrm{cl}\,(\mathrm{conv}\, S)$

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\mathrm{conv}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex}\}.$$

The *closed convex hull* of S is the smallest closed, convex set containing S, i.e.

$$\overline{\mathrm{conv}}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ closed and convex}\}.$$

- $\overline{\mathrm{conv}}\, S = \mathrm{cl}\,(\mathrm{conv}\, S)$
- $\mathrm{conv}\, S = \left\{\sum_{i=1}^{\kappa+1} \lambda_i x_i \;\middle|\; x_i \in S,\ \lambda_i \geq 0\ (i = 1, \dots, \kappa + 1),\ \sum_{i=1}^{\kappa+1} \lambda_i = 1\right\}$ (Carathéodory's Theorem)

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\mathrm{conv}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex}\}.$$

The *closed convex hull* of S is the smallest closed, convex set containing S, i.e.

$$\overline{\mathrm{conv}}\, S := \bigcap \{C \subset \mathbb{E} \mid S \subset C,\ C \text{ closed and convex}\}.$$

- $\overline{\mathrm{conv}}\, S = \mathrm{cl}\,(\mathrm{conv}\, S)$
- $\mathrm{conv}\, S = \left\{\sum_{i=1}^{\kappa+1} \lambda_i x_i \;\middle|\; x_i \in S,\ \lambda_i \geq 0\ (i = 1, \dots, \kappa+1),\ \sum_{i=1}^{\kappa+1} \lambda_i = 1\right\}$ (Carathéodory's Theorem)
- conv preserves compactness and boundedness, not necessarily closedness

The convex hull and the closed convex hull

Definition 1 (Convex hull/closed convex hull).

Let $S \subset \mathbb{E}$ nonempty. Then the *convex hull* of S is the smallest convex set containing S, i.e.

$$\operatorname{conv} S := \bigcap \{ C \subset \mathbb{E} \mid S \subset C,\ C \text{ convex} \}.$$

The *closed convex hull* of S is the smallest closed, convex set containing S, i.e.

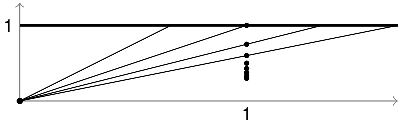
$$\overline{\operatorname{conv}}\, S := \bigcap \{ C \subset \mathbb{E} \mid S \subset C,\ C \text{ closed and convex} \}.$$

- $\overline{\operatorname{conv}}\, S = \operatorname{cl}(\operatorname{conv} S)$
- $\operatorname{conv} S = \left\{ \sum_{i=1}^{\kappa+1} \lambda_i x_i \;\middle|\; x_i \in S,\ \lambda_i \geq 0\ (i = 1, \dots, \kappa+1),\ \sum_{i=1}^{\kappa+1} \lambda_i = 1 \right\}$ (Carathéodory's Theorem)
- conv preserves compactness and boundedness, not necessarily closedness

Example: $S := \{\binom{0}{0}\} \cup \{\binom{a}{1} \mid a \geq 0\}$,

$\binom{1}{1/k} = \frac{1}{k}\binom{k}{1} + \left(1 - \frac{1}{k}\right)\binom{0}{0} \in \operatorname{conv} S.$

But: $\binom{1}{1/k} \to \binom{1}{0} \notin \operatorname{conv} S.$



The topology relative to the affine hull

Affine set: A set $S = U + x$ with $x \in \mathbb{E}$ and a subspace $U \subset$ is called *affine*. This is characterized by

$$\alpha S + (1 - \alpha) S \subset S \quad (\alpha \in \mathbb{R}).$$

The topology relative to the affine hull

Affine set: A set $S = U + x$ with $x \in \mathbb{E}$ and a subspace $U \subset$ is called *affine*. This is characterized by

$$\alpha S + (1 - \alpha)S \subset S \quad (\alpha \in \mathbb{R}).$$

Affine hull: $\operatorname{aff} M := \bigcap \{S \in \mathbb{E} \mid M \subset S,\ S \text{ affine}\}$.

The topology relative to the affine hull

Affine set: A set $S = U + x$ with $x \in \mathbb{E}$ and a subspace $U \subset$ is called *affine*. This is characterized by

$$\alpha S + (1-\alpha)S \subset S \quad (\alpha \in \mathbb{R}).$$

Affine hull: $\operatorname{aff} M := \bigcap \{S \in \mathbb{E} \mid M \subset S,\ S \text{ affine}\}$.

Relative interior/boundary: $C \subset \mathbb{E}$ convex.

$$\operatorname{ri} C \quad := \quad \left\{x \in C \,\middle|\, \exists \varepsilon > 0 : \ B_\varepsilon(x) \cap \operatorname{aff} C \subset C\right\} \qquad \text{(relative interior)}$$

The topology relative to the affine hull

Affine set: A set $S = U + x$ with $x \in \mathbb{E}$ and a subspace $U \subset$ is called *affine*. This is characterized by

$$\alpha S + (1-\alpha)S \subset S \quad (\alpha \in \mathbb{R}).$$

Affine hull: $\mathrm{aff}\, M := \bigcap \{S \in \mathbb{E} \mid M \subset S,\ S \text{ affine}\}$.

Relative interior/boundary: $C \subset \mathbb{E}$ convex.

$$\begin{array}{rcll} \mathrm{ri}\, C & := & \left\{x \in C \,\middle|\, \exists \varepsilon > 0 :\ B_\varepsilon(x) \cap \mathrm{aff}\, C \subset C\right\} & \text{(relative interior)} \\ \mathrm{rbd}\, C & := & \mathrm{cl}\, C \setminus \mathrm{ri}\, C & \text{(relative boundary)} \\ x \in \mathrm{ri}\, C & \Leftrightarrow & \mathrm{span}\, C = \mathbb{R}_+(C - x) & \end{array}$$

The topology relative to the affine hull

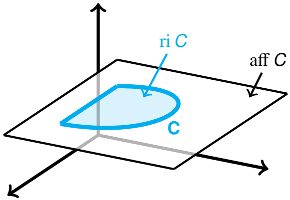
Affine set: A set $S = U + x$ with $x \in \mathbb{E}$ and a subspace $U \subset$ is called *affine*. This is characterized by

$$\alpha S + (1-\alpha)S \subset S \quad (\alpha \in \mathbb{R}).$$

Affine hull: $\operatorname{aff} M := \bigcap \{S \in \mathbb{E} \mid M \subset S,\ S \text{ affine}\}$.

Relative interior/boundary: $C \subset \mathbb{E}$ convex.

$$\begin{array}{rcll} \operatorname{ri} C & := & \left\{x \in C \,\middle|\, \exists \varepsilon > 0 : \ B_\varepsilon(x) \cap \operatorname{aff} C \subset C \right\} & \text{(relative interior)} \\ \operatorname{rbd} C & := & \operatorname{cl} C \setminus \operatorname{ri} C & \text{(relative boundary)} \\ x \in \operatorname{ri} C & \Leftrightarrow & \operatorname{span} C = \mathbb{R}_+(C - x) & \end{array}$$



C	$\operatorname{aff} C$	$\operatorname{ri} C$
$\{x\}$	$\{x\}$	$\{x\}$
$[x, x']$	$\{\lambda x + (1-\lambda)x' \mid \lambda \in \mathbb{R}\}$	(x, x')
$\overline{B}_\varepsilon(x)$	$\mathbb{E}$	$B_\varepsilon(x)$

Table: Examples for relative interiors

The horizon cone

Definition 2 (Horizon cone).

For a nonempty set $S \subset \mathbb{E}$ the set

$$S^{\infty} := \{v \in \mathbb{E} \mid \exists \{x_k \in S\}, \{t_k\} \downarrow 0 : \; t_k x_k \to v\}$$

is called the *horizon cone* of S. We put $\emptyset^{\infty} := \{0\}$.

The horizon cone

Definition 2 (Horizon cone).

For a nonempty set $S \subset \mathbb{E}$ the set

$$S^\infty := \{v \in \mathbb{E} \mid \exists \{x_k \in S\}, \{t_k\} \downarrow 0 : \ t_k x_k \to v\}$$

is called the *horizon cone* of S. We put $\emptyset^\infty := \{0\}$.

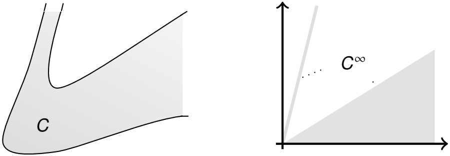


Figure: The horizon cone of an unbounded, nonconvex set

The horizon cone

Definition 2 (Horizon cone).

For a nonempty set $S \subset \mathbb{E}$ the set

$$S^\infty := \{v \in \mathbb{E} \mid \exists \{x_k \in S\}, \{t_k\} \downarrow 0 : \ t_k x_k \to v\}$$

is called the *horizon cone* of S. We put $\emptyset^\infty := \{0\}$.

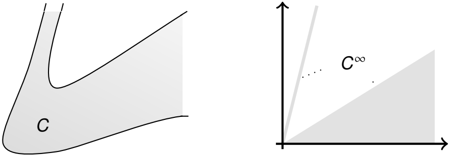


Figure: The horizon cone of an unbounded, nonconvex set

Proposition 3 (The convex case).

Let $C \subset \mathbb{E}$ be nonempty and convex. Then $C^\infty = \{v \mid \forall x \in \operatorname{cl} C, \lambda \geq 0 : \ x + \lambda v \in \operatorname{cl} C\}$. In particular, C^∞ is (a closed and) convex (cone) if C is convex.

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\mathrm{epi}\, f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \leq \alpha\right\}$ (epigraph)

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\operatorname{epi} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \le \alpha\right\}$ (epigraph)
- $\operatorname{epi}_< f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) < \alpha\right\}$ (strict epigraph)

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\operatorname{epi} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \leq \alpha \right\}$ (epigraph)
- $\operatorname{epi}_{<} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) < \alpha \right\}$ (strict epigraph)
- $\operatorname{dom} f := \left\{x \in \mathbb{E} \;\middle|\; f(x) < \infty \right\}$ (domain).

$\rightarrow$ f is uniquely determined through $\operatorname{epi} f$!

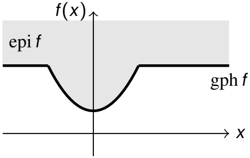


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

[1]Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\operatorname{epi} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \,\middle|\, f(x) \le \alpha\right\}$ (epigraph)
- $\operatorname{epi}_< f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \,\middle|\, f(x) < \alpha\right\}$ (strict epigraph)
- $\operatorname{dom} f := \left\{x \in \mathbb{E} \,\middle|\, f(x) < \infty\right\}$ (domain).

$\rightarrow$ f is uniquely determined through $\operatorname{epi} f$!

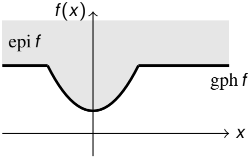


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

f proper $:\Leftrightarrow$ $-\infty < f \not\equiv +\infty$ $\Leftrightarrow$[1] $\operatorname{dom} f \neq \emptyset$

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\operatorname{epi} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \,\middle|\, f(x) \le \alpha\right\}$ (epigraph)
- $\operatorname{epi}_< f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \,\middle|\, f(x) < \alpha\right\}$ (strict epigraph)
- $\operatorname{dom} f := \left\{x \in \mathbb{E} \,\middle|\, f(x) < \infty\right\}$ (domain).

$\rightarrow$ f is uniquely determined through $\operatorname{epi} f$!

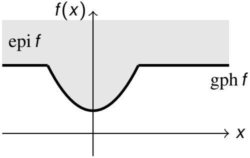


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

f proper $:\Leftrightarrow$ $-\infty < f \not\equiv +\infty$ $\Leftrightarrow$[1] $\operatorname{dom} f \neq \emptyset$

f convex $:\Leftrightarrow$ $\operatorname{epi} f/\operatorname{epi}_< f$ convex $\Leftrightarrow$[1] $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$ $(x, y \in \mathbb{E},\ \lambda \in [0,1])$

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\operatorname{epi} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \leq \alpha\right\}$ (epigraph)
- $\operatorname{epi}_< f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) < \alpha\right\}$ (strict epigraph)
- $\operatorname{dom} f := \left\{x \in \mathbb{E} \;\middle|\; f(x) < \infty\right\}$ (domain).

$\to$ f is uniquely determined through $\operatorname{epi} f$!

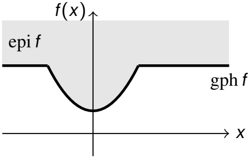


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

f proper	$:\Leftrightarrow$	$-\infty < f \not\equiv +\infty$	$\Leftrightarrow$[1]	$\operatorname{dom} f \neq \emptyset$
f convex	$:\Leftrightarrow$	$\operatorname{epi} f / \operatorname{epi}_< f$ convex	$\Leftrightarrow$[1]	$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \; (x, y \in \mathbb{E},\ \lambda \in [0,1])$
f pos. hom.	$:\Leftrightarrow$[1]	$\operatorname{epi} f$ cone	$\Leftrightarrow$	$\alpha f(x) = f(\alpha x) \quad (x \in \mathbb{E},\ \alpha \geq 0)$

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\mathrm{epi}\, f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \le \alpha\right\}$ (epigraph)
- $\mathrm{epi}_{<} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) < \alpha\right\}$ (strict epigraph)
- $\mathrm{dom}\, f := \left\{x \in \mathbb{E} \;\middle|\; f(x) < \infty\right\}$ (domain).

$\rightarrow$ f is uniquely determined through $\mathrm{epi}\, f$!

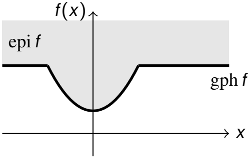


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

f proper	$:\Leftrightarrow$	$-\infty < f \not\equiv +\infty$	$\Leftrightarrow$ [1]	$\mathrm{dom}\, f \neq \emptyset$
f convex	$:\Leftrightarrow$	$\mathrm{epi}\, f / \mathrm{epi}_{<} f$ convex	$\Leftrightarrow$ [1]	$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$ $(x, y \in \mathbb{E},\ \lambda \in [0,1])$
f pos. hom.	$:\Leftrightarrow$ [1]	$\mathrm{epi}\, f$ cone	$\Leftrightarrow$	$\alpha f(x) = f(\alpha x)$ $(x \in \mathbb{E},\ \alpha \ge 0)$
f sublinear	$:\Leftrightarrow$ [1]	$\mathrm{epi}\, f$ cvx. cone	$\Leftrightarrow$	$f(\lambda x + \mu y) \le \lambda f(x) + \mu f(y)$ $(x, y \in \mathbb{E},\ \lambda, \mu \ge 0)$.

[1] Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Extended real-valued functions: An epigraphical perspective

Let $f : \mathbb{E} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.

- $\mathrm{epi}\, f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) \le \alpha \right\}$ (epigraph)
- $\mathrm{epi}_{<} f := \left\{(x, \alpha) \in \mathbb{E} \times \mathbb{R} \;\middle|\; f(x) < \alpha \right\}$ (strict epigraph)
- $\mathrm{dom}\, f := \left\{x \in \mathbb{E} \;\middle|\; f(x) < \infty \right\}$ (domain).

$\to$ f is uniquely determined through $\mathrm{epi}\, f$!

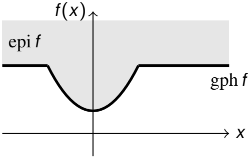


Figure: Epigraph of $f : \mathbb{R} \to \mathbb{R}$

f proper	$:\Leftrightarrow$	$-\infty < f \not\equiv +\infty$	$\Leftrightarrow$ [1]	$\mathrm{dom}\, f \neq \emptyset$
f convex	$:\Leftrightarrow$	$\mathrm{epi}\, f / \mathrm{epi}_{<} f$ convex	$\Leftrightarrow$ [1]	$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$ $(x, y \in \mathbb{E},\ \lambda \in [0,1])$
f pos. hom.	$:\Leftrightarrow$ [1]	$\mathrm{epi}\, f$ cone	$\Leftrightarrow$	$\alpha f(x) = f(\alpha x)$ $(x \in \mathbb{E},\ \alpha \ge 0)$
f sublinear	$:\Leftrightarrow$ [1]	$\mathrm{epi}\, f$ cvx. cone	$\Leftrightarrow$	$f(\lambda x + \mu y) \le \lambda f(x) + \mu f(y)$ $(x, y \in \mathbb{E},\ \lambda, \mu \ge 0)$.
			$\Leftrightarrow$	f convex $+$ positively homogeneous

[1]Only for $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x \to \bar{x}} f(x) := \inf \left\{ \alpha \,\middle|\, \exists\, x_k \to \bar{x} : \ f(x_k) \to \alpha \right\}$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x \to \bar{x}} f(x) := \inf \left\{ \alpha \;\middle|\; \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha \right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x}).$$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x\to\bar{x}} f(x) := \inf\left\{\alpha \,\middle|\, \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha\right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x\to\bar{x}} f(x) \geq f(\bar{x}).$$

Closure: $\mathrm{cl}\, f : \mathbb{E} \to \overline{\mathbb{R}}, \quad (\mathrm{cl}\, f)(\bar{x}) := \liminf_{x\to\bar{x}} f(x).$

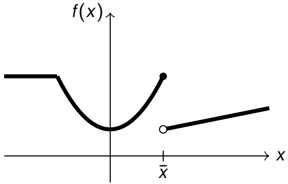


Figure: f *not* lsc at $\bar{x}$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x \to \bar{x}} f(x) := \inf \left\{ \alpha \;\middle|\; \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha \right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x}).$$

Closure: $\mathrm{cl}\, f : \mathbb{E} \to \overline{\mathbb{R}}, \quad (\mathrm{cl}\, f)(\bar{x}) := \liminf_{x \to \bar{x}} f(x).$

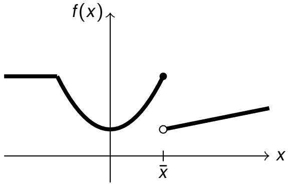


Figure: f *not* lsc at $\bar{x}$

Facts:

- f lsc $\iff$ $\mathrm{epi}\, f$ closed $\iff$ $f = \mathrm{cl}\, f$ $\iff$ $\mathrm{lev}_r f$ closed $(r \in \mathbb{R})$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x \to \bar{x}} f(x) := \inf \left\{ \alpha \;\middle|\; \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha \right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x}).$$

Closure: $\mathrm{cl}\, f : \mathbb{E} \to \overline{\mathbb{R}}, \quad (\mathrm{cl}\, f)(\bar{x}) := \liminf_{x \to \bar{x}} f(x).$

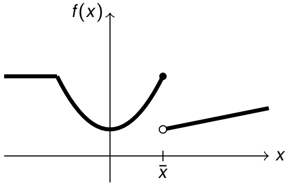


Figure: f *not* lsc at $\bar{x}$

Facts:

- f lsc $\iff$ $\mathrm{epi}\, f$ closed $\iff$ $f = \mathrm{cl}\, f$ $\iff$ $\mathrm{lev}_r f$ closed $(r \in \mathbb{R})$
- $\mathrm{cl}\, f \leq f$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x\to\bar{x}} f(x) := \inf \left\{ \alpha \,\middle|\, \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha \right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x\to\bar{x}} f(x) \geq f(\bar{x}).$$

Closure: $\mathrm{cl}\, f : \mathbb{E} \to \overline{\mathbb{R}}, \quad (\mathrm{cl}\, f)(\bar{x}) := \liminf_{x\to\bar{x}} f(x).$

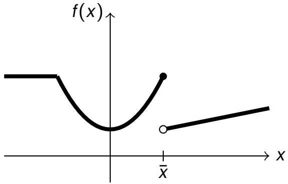


Figure: f *not* lsc at $\bar{x}$

Facts:

- f lsc $\iff$ $\mathrm{epi}\, f$ closed $\iff$ $f = \mathrm{cl}\, f$ $\iff$ $\mathrm{lev}_r f$ closed $(r \in \mathbb{R})$
- $\mathrm{cl}\, f \leq f$
- f proper, lsc and coercive (i.e. $\lim_{\|x\|\to\infty} f(x) = \infty$) then:

$$\operatorname*{argmin}_{\mathbb{E}} f \neq \emptyset \quad \text{and} \quad \inf_{\mathbb{E}} f \in \mathbb{R}$$

Lower semicontinuity

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathbb{E}$.

Lower limit:
$\liminf_{x \to \bar{x}} f(x) := \inf \left\{ \alpha \;\middle|\; \exists\, x_k \to \bar{x} : \; f(x_k) \to \alpha \right\}$

Lower semicontinuity: f is said to be *lsc* (or *closed*) at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x}).$$

Closure: $\mathrm{cl}\, f : \mathbb{E} \to \overline{\mathbb{R}}, \quad (\mathrm{cl}\, f)(\bar{x}) := \liminf_{x \to \bar{x}} f(x).$

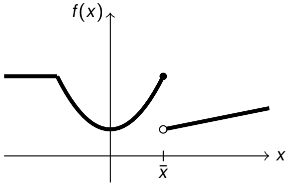


Figure: f *not* lsc at $\bar{x}$

Facts:

- f lsc $\iff$ $\mathrm{epi}\, f$ closed $\iff$ $f = \mathrm{cl}\, f$ $\iff$ $\mathrm{lev}_r f$ closed $(r \in \mathbb{R})$
- $\mathrm{cl}\, f \leq f$
- f proper, lsc and coercive (i.e. $\lim_{\|x\| \to \infty} f(x) = \infty$) then:

$$\underset{\mathbb{E}}{\operatorname{argmin}}\, f \neq \emptyset \quad \text{and} \quad \inf_{\mathbb{E}} f \in \mathbb{R}$$

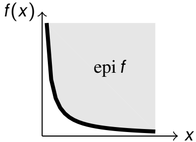


Figure: $f : x \mapsto \begin{cases} \frac{1}{x}, & x > 0, \\ +\infty, & \text{else.} \end{cases}$

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i \ (i \in I) \subset \mathbb{E}, \ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i \ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

- $F(C)$ (affine image)

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

- $F(C)$ (affine image)
- $F^{-1}(D)$ (affine pre-image)

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

- $F(C)$ (affine image)
- $F^{-1}(D)$ (affine pre-image)
- $C \times D$ (Cartesian product)

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

- $F(C)$ (affine image)
- $F^{-1}(D)$ (affine pre-image)
- $C \times D$ (Cartesian product)
- $C_1 + C_2$ (Minkowski sum)

Convexity preserving operations - new from old

1. Set Operations

For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

- $F(C)$ (affine image)
- $F^{-1}(D)$ (affine pre-image)
- $C \times D$ (Cartesian product)
- $C_1 + C_2$ (Minkowski sum)
- $\bigcap_{i \in I} C_i$ (Intersection)

Convexity preserving operations - new from old

1. Set Operations

   For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

   - $F(C)$ (affine image)
   - $F^{-1}(D)$ (affine pre-image)
   - $C \times D$ (Cartesian product)
   - $C_1 + C_2$ (Minkowski sum)
   - $\bigcap_{i \in I} C_i$ (Intersection)

2. Functional operations

   For $f_i, g : \mathbb{E} \to \overline{\mathbb{R}}$ convex and $F : \mathbb{E}' \to \mathbb{E}$ affine the following functions are convex:

Convexity preserving operations - new from old

1. Set Operations

   For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

   - $F(C)$ (affine image)
   - $F^{-1}(D)$ (affine pre-image)
   - $C \times D$ (Cartesian product)
   - $C_1 + C_2$ (Minkowski sum)
   - $\bigcap_{i \in I} C_i$ (Intersection)

2. Functional operations

   For $f_i, g : \mathbb{E} \to \overline{\mathbb{R}}$ convex and $F : \mathbb{E}' \to \mathbb{E}$ affine the following functions are convex:

   - (Affine pre-composition) $f := g \circ F$: $\mathrm{epi}\, f = T^{-1}(\mathrm{epi}\, g),\ T : (x, \alpha) \mapsto (T(x), \alpha)$

Convexity preserving operations - new from old

1. Set Operations

   For $C, C_i\ (i \in I) \subset \mathbb{E},\ D \subset \mathbb{E}'$ convex, $F : \mathbb{E} \to \mathbb{E}'$ affine the following sets are convex:

   - $F(C)$ (affine image)
   - $F^{-1}(D)$ (affine pre-image)
   - $C \times D$ (Cartesian product)
   - $C_1 + C_2$ (Minkowski sum)
   - $\bigcap_{i \in I} C_i$ (Intersection)

2. Functional operations

   For $f_i, g : \mathbb{E} \to \overline{\mathbb{R}}$ convex and $F : \mathbb{E}' \to \mathbb{E}$ affine the following functions are convex:

   - (Affine pre-composition) $f := g \circ F$: $\operatorname{epi} f = T^{-1}(\operatorname{epi} g),\ T : (x, \alpha) \mapsto (T(x), \alpha)$
   - (Pointwise supremum) $f := \sup_{i \in I} f_i$: $\operatorname{epi} f = \bigcap_{i \in I} \operatorname{epi} f_i$

Convexity preserving operations - new from old

1. Set Operations

   For $C, C_i\ (i\in I)\subset \mathbb{E},\ D\subset \mathbb{E}'$ convex, $F:\mathbb{E}\to\mathbb{E}'$ affine the following sets are convex:

   - $F(C)$ (affine image)
   - $F^{-1}(D)$ (affine pre-image)
   - $C\times D$ (Cartesian product)
   - C_1+C_2 (Minkowski sum)
   - $\bigcap_{i\in I} C_i$ (Intersection)

2. Functional operations

   For $f_i, g:\mathbb{E}\to\overline{\mathbb{R}}$ convex and $F:\mathbb{E}'\to\mathbb{E}$ affine the following functions are convex:

   - (Affine pre-composition) $f:=g\circ F$: $\mathrm{epi}\, f = T^{-1}(\mathrm{epi}\, g),\ T:(x,\alpha)\mapsto(T(x),\alpha)$
   - (Pointwise supremum) $f:=\sup_{i\in I} f_i$: $\mathrm{epi}\, f=\bigcap_{i\in I}\mathrm{epi}\, f_i$
   - (Moreau envelope) $f:=e_\lambda g: x\mapsto \inf_u\left\{g(u)+\frac{1}{2\lambda}\|x-u\|^2\right\}$: $\mathrm{epi}\, f=\mathrm{epi}\, g+\mathrm{epi}\,\frac{1}{2}\|\cdot\|^2$.

The convex subdifferential

Definition 4.

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$. A vector $v \in \mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad (x \in \mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

The convex subdifferential

Definition 4.

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$. A vector $v \in \mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad (x \in \mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

Slogan: *"The subgradients of f at $\bar{x}$ are the slopes of affine minorants of f that coincide with f at $\bar{x}$".*

The convex subdifferential

Definition 4.

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$. A vector $v \in \mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad (x \in \mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

Slogan: *"The subgradients of f at $\bar{x}$ are the slopes of affine minorants of f that coincide with f at $\bar{x}$".*

The subdifferential operator is a set-valued mapping $\partial f : \mathbb{E} \rightrightarrows \mathbb{E}$. Set

$$\mathrm{dom}\,\partial f := \left\{x \in \mathbb{E} \;\middle|\; \partial f(x) \neq \emptyset\right\}.$$

- $0 \in \partial f(x) \quad \Longleftrightarrow \quad x \in \mathrm{argmin}_{\mathbb{E}} f \quad$ (Fermat's rule)

The convex subdifferential

Definition 4.

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$. A vector $v \in \mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad (x \in \mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

Slogan: *"The subgradients of f at $\bar{x}$ are the slopes of affine minorants of f that coincide with f at $\bar{x}$".*

The subdifferential operator is a set-valued mapping $\partial f : \mathbb{E} \rightrightarrows \mathbb{E}$. Set

$$\operatorname{dom} \partial f := \left\{ x \in \mathbb{E} \,\middle|\, \partial f(x) \neq \emptyset \right\}.$$

- $0 \in \partial f(x) \iff x \in \operatorname{argmin}_{\mathbb{E}} f$ (Fermat's rule)
- $\partial f(x)$ closed and convex ($x \in \mathbb{E}$)

The convex subdifferential

Definition 4.

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$. A vector $v \in \mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad (x \in \mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

Slogan: *"The subgradients of f at $\bar{x}$ are the slopes of affine minorants of f that coincide with f at $\bar{x}$".*

The subdifferential operator is a set-valued mapping $\partial f : \mathbb{E} \rightrightarrows \mathbb{E}$. Set

$$\mathrm{dom}\, \partial f := \left\{ x \in \mathbb{E} \,\middle|\, \partial f(x) \neq \emptyset \right\}.$$

- $0 \in \partial f(x) \quad \Longleftrightarrow \quad x \in \mathrm{argmin}_{\mathbb{E}} f$ (Fermat's rule)
- $\partial f(x)$ closed and convex ($x \in \mathbb{E}$)
- $\partial f(x)$ is a singleton $\quad \Longleftrightarrow \quad f$ differentiable at $x \quad \Longleftrightarrow \quad f$ continuously differentiable at x

The convex subdifferential

Definition 4.

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$. A vector $v\in\mathbb{E}$ is called a *subgradient* of v at $\bar{x}$ if

$$f(x)\geq f(\bar{x})+\langle v,\, x-\bar{x}\rangle \quad (x\in\mathbb{E}). \tag{1}$$

We denote by $\partial f(\bar{x})$ the set of all subgradients of f at $\bar{x}$ and call it the *(convex) subdifferential* of f at $\bar{x}$.

The inequality (1) is referred to as *subgradient inequality*.

Slogan: *"The subgradients of f at $\bar{x}$ are the slopes of affine minorants of f that coincide with f at $\bar{x}$".*

The subdifferential operator is a set-valued mapping $\partial f:\mathbb{E}\rightrightarrows\mathbb{E}$. Set

$$\mathrm{dom}\,\partial f:=\left\{x\in\mathbb{E}\;\middle|\;\partial f(x)\neq\emptyset\right\}.$$

- $0\in\partial f(x) \iff x\in\mathrm{argmin}_{\mathbb{E}}\, f$ (Fermat's rule)
- $\partial f(x)$ closed and convex ($x\in\mathbb{E}$)
- $\partial f(x)$ is a singleton $\iff$ f differentiable at x $\iff$ f continuously differentiable at x
- $\mathrm{ri}\,(\mathrm{dom}\, f)\subset\mathrm{dom}\,\partial f\subset\mathrm{dom}\, f$ (f convex).

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

$$\partial \delta_S(\bar{x}) \quad = \quad \left\{ v \;\middle|\; \delta_C(x) \geq \delta_C(\bar{x}) + \langle v, x - \bar{x} \rangle \;\; (x \in \mathbb{E}) \right\}$$

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

$$\begin{aligned} \partial\delta_S(\bar{x}) &= \left\{ v \;\middle|\; \delta_C(x) \geq \delta_C(\bar{x}) + \langle v, x - \bar{x} \rangle \ (x \in \mathbb{E}) \right\} \\ &= \left\{ v \in \mathbb{E} \;\middle|\; \langle v, x - \bar{x} \rangle \leq 0 \ (x \in S) \right\} \end{aligned}$$

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

$$\begin{aligned} \partial \delta_S(\bar{x}) &= \left\{ v \;\middle|\; \delta_C(x) \geq \delta_C(\bar{x}) + \langle v, x - \bar{x} \rangle \;\; (x \in \mathbb{E}) \right\} \\ &= \left\{ v \in \mathbb{E} \;\middle|\; \langle v, x - \bar{x} \rangle \leq 0 \; (x \in S) \right\} \\ &=: N_S(\bar{x}) \quad (\bar{x} \in S) \end{aligned}$$

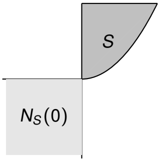


Figure: Normal cone

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

$$\begin{aligned} \partial \delta_S(\bar{x}) &= \left\{ v \,\middle|\, \delta_C(x) \geq \delta_C(\bar{x}) + \langle v,\, x - \bar{x} \rangle \ (x \in \mathbb{E}) \right\} \\ &= \left\{ v \in \mathbb{E} \,\middle|\, \langle v,\, x - \bar{x} \rangle \leq 0 \ (x \in S) \right\} \\ &=: N_S(\bar{x}) \quad (\bar{x} \in S) \end{aligned}$$

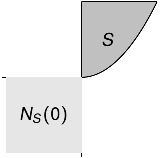


Figure: Normal cone

- (Euclidean norm) $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$. Then

$$\partial \|\cdot\|(\bar{x}) = \begin{cases} \left\{ \frac{\bar{x}}{\|\bar{x}\|} \right\} & \text{if} \quad \bar{x} \neq 0, \\ \mathbb{B} & \text{if} \quad \bar{x} = 0. \end{cases}$$

Examples of subdifferentiation

- (Indicator function/Normal cone) Let $S \subset \mathbb{E}$.

Indicator function of S:

$$\delta_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \delta_S(x) := \begin{cases} 0, & x \in S, \\ +\infty, & \text{else.} \end{cases}$$

$$\begin{aligned} \partial \delta_S(\bar{x}) &= \left\{ v \;\middle|\; \delta_C(x) \geq \delta_C(\bar{x}) + \langle v, x - \bar{x} \rangle \;\; (x \in \mathbb{E}) \right\} \\ &= \left\{ v \in \mathbb{E} \;\middle|\; \langle v, x - \bar{x} \rangle \leq 0 \;(x \in S) \right\} \\ &=: N_S(\bar{x}) \quad (\bar{x} \in S) \end{aligned}$$

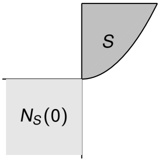


Figure: Normal cone

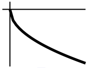
- (Euclidean norm) $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$. Then

$$\partial \|\cdot\|(\bar{x}) = \begin{cases} \left\{ \frac{\bar{x}}{\|\bar{x}\|} \right\} & \text{if} \quad \bar{x} \neq 0, \\ \mathbb{B} & \text{if} \quad \bar{x} = 0. \end{cases}$$

- (Empty subdifferential)

$$f : x \in \mathbb{R} \mapsto \begin{cases} -\sqrt{x} & \text{if} \quad x \geq 0, \\ +\infty & \text{else.} \end{cases}$$

$$\partial f(x) = \begin{cases} \left\{ -\frac{1}{2\sqrt{x}} \right\}, & x > 0, \\ \emptyset, & \text{else.} \end{cases}$$



The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\mathrm{epi}\, f^* \overset{!}{=} \Big\{(v,\beta)\;\Big|\;\langle v,\, x\rangle-\beta\le f(x)\quad (x\in\mathbb{E})\Big\}$$

The Fenchel conjugate

For $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$ let $f^* : \mathbb{E} \to \overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$
\begin{aligned}
& \mathrm{epi}\, f^* \overset{!}{=} \left\{ (v, \beta) \;\middle|\; \langle v,\, x \rangle - \beta \le f(x) \quad (x \in \mathbb{E}) \right\} \\
\Longrightarrow \quad & f^*(v) \le \beta \quad \Longleftrightarrow \quad \sup_{x \in \mathbb{E}} \{ \langle v,\, x \rangle - f(x) \} \le \beta \quad ((v, \beta) \in \mathbb{E} \times \mathbb{R})
\end{aligned}
$$

The Fenchel conjugate

For $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$ let $f^* : \mathbb{E} \to \overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \overset{!}{=} \left\{(v, \beta) \;\middle|\; \langle v,\, x\rangle - \beta \leq f(x) \quad (x \in \mathbb{E})\right\} \\ \Longrightarrow \quad & f^*(v) \leq \beta \quad \Longleftrightarrow \quad \sup_{x \in \mathbb{E}} \{\langle v,\, x\rangle - f(x)\} \leq \beta \quad ((v, \beta) \in \mathbb{E} \times \mathbb{R}) \\ \Longrightarrow \quad & f^*(v) = \sup_{x \in \mathbb{E}} \{\langle v,\, x\rangle - f(x)\} \quad (v \in \mathbb{E}). \end{aligned} \tag{2}$$

The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \overset{!}{=} \left\{(v,\beta)\;\middle|\;\langle v,\,x\rangle-\beta\le f(x)\quad(x\in\mathbb{E})\right\}\\ \Longrightarrow\quad & f^*(v)\le\beta \quad\Longleftrightarrow\quad \sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\le\beta \quad ((v,\beta)\in\mathbb{E}\times\mathbb{R})\\ \Longrightarrow\quad & f^*(v)=\sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\quad(v\in\mathbb{E}). \end{aligned} \tag{2}$$

Definition 5 (Fenchel conjugate).

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$ proper. The function $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ defined through (2) is called the <u>(Fenchel) conjugate of f</u>. The function $(f^{**}):=(f^*)^*$ is called the <u>biconjugate of f</u>.

The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \overset{!}{=} \left\{(v,\beta)\;\middle|\;\langle v,\,x\rangle-\beta\le f(x)\quad (x\in\mathbb{E})\right\}\\ \Longrightarrow\quad & f^*(v)\le\beta \quad\Longleftrightarrow\quad \sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\le\beta\quad ((v,\beta)\in\mathbb{E}\times\mathbb{R})\\ \Longrightarrow\quad & f^*(v)=\sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\quad (v\in\mathbb{E}). \end{aligned} \tag{2}$$

Definition 5 (Fenchel conjugate).

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$ proper. The function $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ defined through (2) is called the (Fenchel) conjugate of f. The function $(f^{**}):=(f^*)^*$ is called the biconjugate of f.

Define $\Gamma:=\left\{f:\mathbb{E}\to\overline{\mathbb{R}}\mid f \text{ convex and proper}\right\}$ and $\Gamma_0:=\{f\in\Gamma\mid f\text{ closed}\}$.

The Fenchel conjugate

For $f : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$ let $f^* : \mathbb{E} \to \overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$
\begin{aligned}
& \mathrm{epi}\, f^* \overset{!}{=} \left\{(v,\beta) \;\middle|\; \langle v,\, x\rangle - \beta \le f(x) \quad (x \in \mathbb{E})\right\} \\
\Longrightarrow \quad & f^*(v) \le \beta \quad \Longleftrightarrow \quad \sup_{x\in\mathbb{E}} \{\langle v,\, x\rangle - f(x)\} \le \beta \quad ((v,\beta) \in \mathbb{E}\times\mathbb{R}) \\
\Longrightarrow \quad & f^*(v) = \sup_{x\in\mathbb{E}} \{\langle v,\, x\rangle - f(x)\} \quad (v \in \mathbb{E}). \qquad (2)
\end{aligned}
$$

Definition 5 (Fenchel conjugate).

Let $f : \mathbb{E} \to \overline{\mathbb{R}}$ proper. The function $f^* : \mathbb{E} \to \overline{\mathbb{R}}$ defined through (2) is called the (Fenchel) conjugate of f. The function $(f^{**}) := (f^*)^*$ is called the biconjugate of f.

Define $\Gamma := \left\{f : \mathbb{E} \to \overline{\mathbb{R}} \mid f \text{ convex and proper}\right\}$ and $\Gamma_0 := \{f \in \Gamma \mid f \text{ closed}\}$.

- f^* closed and convex - proper if $f \not\equiv +\infty$ with an affine minorant

The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \overset{!}{=} \left\{(v,\beta)\;\middle|\;\langle v,\,x\rangle-\beta\leq f(x)\quad (x\in\mathbb{E})\right\} \\ \Longrightarrow\quad & f^*(v)\leq\beta \quad\Longleftrightarrow\quad \sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\leq\beta\quad ((v,\beta)\in\mathbb{E}\times\mathbb{R}) \\ \Longrightarrow\quad & f^*(v)=\sup_{x\in\mathbb{E}}\{\langle v,\,x\rangle-f(x)\}\quad (v\in\mathbb{E}). \end{aligned} \tag{2}$$

Definition 5 (Fenchel conjugate).

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$ proper. The function $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ defined through (2) is called the (Fenchel) conjugate of f. The function $(f^{**}):=(f^*)^*$ is called the biconjugate of f.

Define $\Gamma:=\left\{f:\mathbb{E}\to\overline{\mathbb{R}}\mid f \text{ convex and proper}\right\}$ and $\Gamma_0:=\{f\in\Gamma\mid f \text{ closed}\}$.

- f^* closed and convex - proper if $f\not\equiv+\infty$ with an affine minorant
- $f=f^{**}$ *proper* $\Longleftrightarrow$ $f\in\Gamma_0$ (Fenchel-Moreau)

The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \overset{!}{=} \left\{(v,\beta)\;\middle|\;\langle v,\, x\rangle-\beta\le f(x)\quad (x\in\mathbb{E})\right\} \\ \Longrightarrow\quad & f^*(v)\le\beta \quad\Longleftrightarrow\quad \sup_{x\in\mathbb{E}}\{\langle v,\, x\rangle - f(x)\}\le\beta \quad ((v,\beta)\in\mathbb{E}\times\mathbb{R}) \\ \Longrightarrow\quad & f^*(v)=\sup_{x\in\mathbb{E}}\{\langle v,\, x\rangle - f(x)\}\quad (v\in\mathbb{E}). \end{aligned} \tag{2}$$

Definition 5 (Fenchel conjugate).

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$ proper. The function $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ defined through (2) is called the (Fenchel) conjugate of f. The function $(f^{**}):=(f^*)^*$ is called the biconjugate of f.

Define $\Gamma:=\left\{f:\mathbb{E}\to\overline{\mathbb{R}}\mid f \text{ convex and proper}\right\}$ and $\Gamma_0:=\{f\in\Gamma\mid f \text{ closed}\}$.

- f^* closed and convex - proper if $f\not\equiv+\infty$ with an affine minorant
- $f=f^{**}$ *proper* $\iff$ $f\in\Gamma_0$ (Fenchel-Moreau)
- $f^*=(\mathrm{cl}\, f)^*\quad (f\in\Gamma)$

The Fenchel conjugate

For $f:\mathbb{E}\to\mathbb{R}\cup\{+\infty\}$ let $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ be the function whose epigraph encodes the affine minorants of $\mathrm{epi}\, f$:

$$\begin{aligned} & \mathrm{epi}\, f^* \stackrel{!}{=} \left\{(v,\beta)\;\middle|\;\langle v,\, x\rangle - \beta \le f(x) \quad (x\in\mathbb{E})\right\} \\ \Longrightarrow \quad & f^*(v)\le\beta \quad\Longleftrightarrow\quad \sup_{x\in\mathbb{E}}\{\langle v,\, x\rangle - f(x)\}\le\beta \quad ((v,\beta)\in\mathbb{E}\times\mathbb{R}) \\ \Longrightarrow \quad & f^*(v) = \sup_{x\in\mathbb{E}}\{\langle v,\, x\rangle - f(x)\} \quad (v\in\mathbb{E}). \end{aligned} \tag{2}$$

Definition 5 (Fenchel conjugate).

Let $f:\mathbb{E}\to\overline{\mathbb{R}}$ proper. The function $f^*:\mathbb{E}\to\overline{\mathbb{R}}$ defined through (2) is called the (Fenchel) conjugate of f. The function $(f^{**}) := (f^*)^*$ is called the biconjugate of f.

Define $\Gamma := \left\{f:\mathbb{E}\to\overline{\mathbb{R}} \mid f \text{ convex and proper}\right\}$ and $\Gamma_0 := \{f\in\Gamma \mid f \text{ closed}\}$.

- f^* closed and convex - proper if $f\not\equiv+\infty$ with an affine minorant
- $f = f^{**}$ *proper* $\iff$ $f\in\Gamma_0$ (Fenchel-Moreau)
- $f^* = (\mathrm{cl}\, f)^* \quad (f\in\Gamma)$
- $f(x)+f^*(y)\ge\langle x,\, y\rangle \quad (x,y\in\mathbb{E})$ (Fenchel-Young Inequality)

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y \rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y \rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Proof.

Notice that

$$y \in \partial f(x) \qquad \Longleftrightarrow \qquad f(z) \geq f(x) + \langle y, z - x \rangle \quad (z \in \mathbb{E})$$

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y \rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Proof.

Notice that

$$\begin{aligned} y \in \partial f(x) \quad &\iff \quad f(z) \geq f(x) + \langle y, z - x \rangle \quad (z \in \mathbb{E}) \\ &\iff \quad \langle y, x \rangle - f(x) \geq \sup_z \{\langle y, z \rangle - f(z)\} \end{aligned}$$

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y\rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Proof.

Notice that

$$\begin{aligned} y \in \partial f(x) \quad &\Longleftrightarrow \quad f(z) \geq f(x) + \langle y, z - x\rangle \quad (z \in \mathbb{E}) \\ &\Longleftrightarrow \quad \langle y, x\rangle - f(x) \geq \sup_z\{\langle y, z\rangle - f(z)\} \\ &\Longleftrightarrow \quad f(x) + f^*(y) \leq \langle x, y\rangle \end{aligned}$$

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y\rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Proof.

Notice that

$$\begin{aligned} y \in \partial f(x) \quad &\Longleftrightarrow \quad f(z) \geq f(x) + \langle y, z - x\rangle \quad (z \in \mathbb{E}) \\ &\Longleftrightarrow \quad \langle y, x\rangle - f(x) \geq \sup_z \{\langle y, z\rangle - f(z)\} \\ &\Longleftrightarrow \quad f(x) + f^*(y) \leq \langle x, y\rangle \\ &\overset{\text{Fenchel–Young}}{\Longleftrightarrow} \quad f(x) + f^*(y) = \langle x, y\rangle, \end{aligned}$$

Interplay of conjugation and subdifferentiation

Theorem 6 (Subdifferential and conjugate function).

Let $f \in \Gamma_0$. TFAE:

i) $y \in \partial f(x)$;

ii) $f(x) + f^*(y) = \langle x, y \rangle$;

iii) $x \in \partial f^*(y)$.

In particular, $\partial f^* = (\partial f)^{-1}$.

Proof.

Notice that

$$\begin{aligned} y \in \partial f(x) \quad & \iff \quad f(z) \geq f(x) + \langle y, z - x \rangle \quad (z \in \mathbb{E}) \\ & \iff \quad \langle y, x \rangle - f(x) \geq \sup_z \{\langle y, z \rangle - f(z)\} \\ & \iff \quad f(x) + f^*(y) \leq \langle x, y \rangle \\ & \overset{\text{Fenchel–Young}}{\iff} \quad f(x) + f^*(y) = \langle x, y \rangle, \end{aligned}$$

Applying the same reasoning to f^* and noticing that $f^{**} = f$ if $f \in \Gamma_0$, gives the missing equivalence. $\square$

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$
- $\sigma_S^* = \delta_{\overline{\mathrm{conv}}\, S}$

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$
- $\sigma_S^* = \delta_{\overline{\mathrm{conv}}\, S}$
- $\partial \sigma_S(x) = \left\{ z \in \overline{\mathrm{conv}}\, S \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, S}(z) \right\}$

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$
- $\sigma_S^* = \delta_{\overline{\mathrm{conv}}\, S}$
- $\partial \sigma_S(x) = \left\{ z \in \overline{\mathrm{conv}}\, S \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, S}(z) \right\}$
- $\mathrm{epi}\, \sigma_S = \bigcap_{s \in S} \mathrm{epi}\, \langle s, \cdot \rangle$ is a nonempty, closed, convex cone, i.e. σ_S is proper, closed and sublinear.

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$
- $\sigma_S^* = \delta_{\overline{\mathrm{conv}}\, S}$
- $\partial \sigma_S(x) = \left\{ z \in \overline{\mathrm{conv}}\, S \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, S}(z) \right\}$
- $\mathrm{epi}\, \sigma_S = \bigcap_{s \in S} \mathrm{epi}\, \langle s, \cdot \rangle$ is a nonempty, closed, convex cone, i.e. σ_S is proper, closed and sublinear.

Support functions: A special case of conjugacy

The *support function* σ_S of $S \subset \mathbb{E}$ (nonempty) is defined by

$$\sigma_S : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}, \quad \sigma_S(z) := \delta_S^*(z) = \sup_{x \in S} \langle x, z \rangle .$$

- σ_S is finite-valued if and only if S is bounded (and nonempty)
- $\sigma_S = \sigma_{\mathrm{conv}\, S} = \sigma_{\overline{\mathrm{conv}}\, S} = \sigma_{\mathrm{cl}\, S}$
- $\sigma_S^* = \delta_{\overline{\mathrm{conv}}\, S}$
- $\partial \sigma_S(x) = \left\{ z \in \overline{\mathrm{conv}}\, S \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, S}(z) \right\}$
- $\mathrm{epi}\, \sigma_S = \bigcap_{s \in S} \mathrm{epi}\, \langle s, \cdot \rangle$ is a nonempty, closed, convex cone, i.e. σ_S is proper, closed and sublinear.

Here's the complete picture:

Theorem 7 (Hörmander).

A function $f : \mathbb{E} \to \overline{\mathbb{R}}$ is proper, closed and sublinear if and only if it is a support function.

Proof.

Blackboard/Notes. $\square$

Gauges and polar sets

Definition 8 (Gauge function).

Let $C \subset \mathbb{E}$. The *gauge (function)* of C is defined by $\gamma_C : x \in \mathbb{E} \mapsto \inf \{\lambda \geq 0 \mid x \in \lambda C\}$.

Gauges and polar sets

Definition 8 (Gauge function).

Let $C \subset \mathbb{E}$. The *gauge (function)* of C is defined by $\gamma_C : x \in \mathbb{E} \mapsto \inf \{\lambda \geq 0 \mid x \in \lambda C\}$.

- If $C \subset \mathbb{E}$ be nonempty, closed and convex with $0 \in C$, then γ_C is proper, lsc and sublinear.

Gauges and polar sets

Definition 8 (Gauge function).

Let $C \subset \mathbb{E}$. The *gauge (function)* of C is defined by $\gamma_C : x \in \mathbb{E} \mapsto \inf \{\lambda \geq 0 \mid x \in \lambda C\}$.

- If $C \subset \mathbb{E}$ be nonempty, closed and convex with $0 \in C$, then γ_C is proper, lsc and sublinear.

Definition 9 (Polar sets).

Let $C \subset \mathbb{E}$. Then its *polar set* is defined by

$$C^\circ := \left\{ v \in \mathbb{E} \,\middle|\, \langle v, x \rangle \leq 1 \ (x \in C) \right\}.$$

Moreover, we put $C^{\circ\circ} := (C^\circ)^\circ$ and call it the *bipolar set* of C.

- If K is a cone then $K^\circ = \left\{ v \in \mathbb{E} \,\middle|\, \langle v, x \rangle \leq 0 \ (x \in K) \right\}$.
- For $C \subset \mathbb{E}$ we have $C^{\circ\circ} = \overline{\mathrm{conv}}\,(C \cup \{0\})$. (bipolar theorem)

Gauges and polar sets

Definition 8 (Gauge function).

Let $C \subset \mathbb{E}$. The *gauge (function)* of C is defined by $\gamma_C : x \in \mathbb{E} \mapsto \inf \{\lambda \geq 0 \mid x \in \lambda C\}$.

- If $C \subset \mathbb{E}$ be nonempty, closed and convex with $0 \in C$, then γ_C is proper, lsc and sublinear.

Definition 9 (Polar sets).

Let $C \subset \mathbb{E}$. Then its *polar set* is defined by

$$C^\circ := \left\{ v \in \mathbb{E} \,\middle|\, \langle v, x \rangle \leq 1 \ (x \in C) \right\}.$$

Moreover, we put $C^{\circ\circ} := (C^\circ)^\circ$ and call it the *bipolar set* of C.

- If K is a cone then $K^\circ = \left\{ v \in \mathbb{E} \,\middle|\, \langle v, x \rangle \leq 0 \ (x \in K) \right\}$.
- For $C \subset \mathbb{E}$ we have $C^{\circ\circ} = \overline{\mathrm{conv}}\,(C \cup \{0\})$. (bipolar theorem)

Proposition 10.

Let $C \subset \mathbb{E}$ be closed and convex with $0 \in C$. Then

$$\gamma_C = \sigma_{C^\circ} \overset{*}{\longleftrightarrow} \delta_{C^\circ} \quad \textit{and} \quad \gamma_{C^\circ} = \sigma_C \overset{*}{\longleftrightarrow} \delta_C.$$

Infimal projection

Theorem 11 (Infimal projection).

Let $\psi : \mathbb{E}_1 \times \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ *be convex. Then the* optimal value function

$$p : \mathbb{E}_1 \to \overline{\mathbb{R}}, \; p(x) := \inf_{y \in \mathbb{E}_2} \psi(x, y)$$

is convex.

Infimal projection

Theorem 11 (Infimal projection).

Let $\psi : \mathbb{E}_1 \times \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ *be convex. Then the* optimal value function

$$p : \mathbb{E}_1 \to \overline{\mathbb{R}}, \; p(x) := \inf_{y \in \mathbb{E}_2} \psi(x, y)$$

is convex.

Proof.

Let $L : (x, y, \alpha) \mapsto (x, \alpha)$ and observe that

$$\mathrm{epi}_{<} p \quad = \quad \left\{ (x, \alpha) \,\middle|\, \inf_y \psi(x, y) < \alpha \right\}$$

Infimal projection

Theorem 11 (Infimal projection).

Let $\psi : \mathbb{E}_1 \times \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ *be convex. Then the* optimal value function

$$p : \mathbb{E}_1 \to \overline{\mathbb{R}}, \; p(x) := \inf_{y \in \mathbb{E}_2} \psi(x, y)$$

is convex.

Proof.

Let $L : (x, y, \alpha) \mapsto (x, \alpha)$ and observe that

$$\begin{aligned} \operatorname{epi}_{<} p &= \left\{ (x, \alpha) \;\middle|\; \inf_y \psi(x, y) < \alpha \right\} \\ &= \left\{ (x, \alpha) \;\middle|\; \exists y : \; \psi(x, y) < \alpha \right\} \end{aligned}$$

Infimal projection

Theorem 11 (Infimal projection).

Let $\psi : \mathbb{E}_1 \times \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ *be convex. Then the* optimal value function

$$p : \mathbb{E}_1 \to \overline{\mathbb{R}},\ p(x) := \inf_{y \in \mathbb{E}_2} \psi(x, y)$$

is convex.

Proof.

Let $L : (x, y, \alpha) \mapsto (x, \alpha)$ and observe that

$$\begin{aligned} \mathrm{epi}_{<} p &= \left\{ (x, \alpha) \,\middle|\, \inf_y \psi(x, y) < \alpha \right\} \\ &= \left\{ (x, \alpha) \,\middle|\, \exists y :\ \psi(x, y) < \alpha \right\} \\ &= L(\mathrm{epi}_{<} \psi). \end{aligned}$$

Infimal projection

Theorem 11 (Infimal projection).

Let $\psi : \mathbb{E}_1 \times \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ *be convex. Then the* optimal value function

$$p : \mathbb{E}_1 \to \overline{\mathbb{R}}, \; p(x) := \inf_{y \in \mathbb{E}_2} \psi(x, y)$$

is convex.

Proof.

Let $L : (x, y, \alpha) \mapsto (x, \alpha)$ and observe that

$$\begin{aligned} \mathrm{epi}_{<} p &= \left\{ (x, \alpha) \;\middle|\; \inf_y \psi(x, y) < \alpha \right\} \\ &= \left\{ (x, \alpha) \mid \exists y : \; \psi(x, y) < \alpha \right\} \\ &= L(\mathrm{epi}_{<} \psi). \end{aligned}$$

Hence $\mathrm{epi}_{<} p$ is a convex set, and thus p is convex. $\square$

Infimal convolution - a special case of infimal projection

Definition 12 (Infimal convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the function

$$f \# g : \mathbb{E} \to \overline{\mathbb{R}}, \quad (f \# g)(x) := \inf_{u \in \mathbb{E}} \{f(u) + g(x - u)\}$$

is called the *infimal convolution* of f and g. We call the infimal convolution $f \# g$ *exact at* $x \in \mathbb{E}$ if

$$\operatorname*{argmin}_{u \in \mathbb{E}} \{f(u) + g(x - u)\} \neq \emptyset.$$

We simply call $f \# g$ exact if it is exact at every $x \in \operatorname{dom} f \# g$.

Infimal convolution - a special case of infimal projection

Definition 12 (Infimal convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the function

$$f\#g : \mathbb{E} \to \overline{\mathbb{R}}, \quad (f\#g)(x) := \inf_{u \in \mathbb{E}} \{f(u) + g(x-u)\}$$

is called the *infimal convolution* of f and g. We call the infimal convolution $f\#g$ *exact at* $x \in \mathbb{E}$ if

$$\operatorname*{argmin}_{u \in \mathbb{E}} \{f(u) + g(x-u)\} \neq \emptyset.$$

We simply call $f\#g$ exact if it is exact at every $x \in \operatorname{dom} f\#g$.

We always have:

- $\operatorname{dom} f\#g = \operatorname{dom} f + \operatorname{dom} g$;
- $f\#g = g\#f$;
- f, g convex, then $f\#g$ convex (as $(f\#g)(x) = \inf_y h(x,y)$ with $h : (x,y) \mapsto f(y) + g(x-y)$ convex).

Infimal convolution - a special case of infimal projection

Definition 12 (Infimal convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the function

$$f \# g : \mathbb{E} \to \overline{\mathbb{R}}, \quad (f \# g)(x) := \inf_{u \in \mathbb{E}} \{f(u) + g(x-u)\}$$

is called the *infimal convolution* of f and g. We call the infimal convolution $f \# g$ *exact at* $x \in \mathbb{E}$ if

$$\operatorname*{argmin}_{u \in \mathbb{E}} \{f(u) + g(x-u)\} \neq \emptyset.$$

We simply call $f \# g$ exact if it is exact at every $x \in \operatorname{dom} f \# g$.

We always have:

- $\operatorname{dom} f \# g = \operatorname{dom} f + \operatorname{dom} g$;
- $f \# g = g \# f$;
- f, g convex, then $f \# g$ convex (as $(f \# g)(x) = \inf_y h(x, y)$ with $h : (x, y) \mapsto f(y) + g(x-y)$ convex).

Example 13 (Distance functions).

Let $C \subset \mathbb{E}$. Then $d_C := \delta_C \# \|\cdot\|$, i.e.

$$d_C(x) = \inf_{u \in C} \|x - u\|$$

is the distance function of C, which is hence convex if C is a convex.

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. *Then the following hold:*

a) $(f \# g)^* = f^* + g^*$;

b) *If* $f, g \in \Gamma_0$ *such that* $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, *then* $(f+g)^* = \mathrm{cl}\,(f^* \# g^*)$.

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f + g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$(f \# g)^*(y) \quad = \quad \sup_x \left\{ \langle x, y \rangle - \inf_u \{f(u) + g(x - u)\} \right\}$$

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f+g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$\begin{aligned} (f\#g)^*(y) &= \sup_x \left\{\langle x, y\rangle - \inf_u \{f(u) + g(x-u)\}\right\} \\ &= \sup_{x,u} \{\langle x, y\rangle - f(u) - g(x-u)\} \end{aligned}$$

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f+g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$\begin{aligned} (f\#g)^*(y) &= \sup_x \left\{\langle x, y\rangle - \inf_u \{f(u) + g(x-u)\}\right\} \\ &= \sup_{x,u} \{\langle x, y\rangle - f(u) - g(x-u)\} \\ &= \sup_{x,u} \{(\langle u, y\rangle - f(u)) + (\langle x-u, y\rangle - g(x-u))\} \end{aligned}$$

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f+g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$\begin{aligned} (f\#g)^*(y) &= \sup_x \left\{\langle x, y\rangle - \inf_u \{f(u) + g(x-u)\}\right\} \\ &= \sup_{x,u} \{\langle x, y\rangle - f(u) - g(x-u)\} \\ &= \sup_{x,u} \{(\langle u, y\rangle - f(u)) + (\langle x-u, y\rangle - g(x-u))\} \\ &= f^*(y) + g^*(y). \end{aligned}$$

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f + g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$\begin{aligned}
(f \# g)^*(y) &= \sup_x \left\{ \langle x, y \rangle - \inf_u \{ f(u) + g(x-u) \} \right\} \\
&= \sup_{x,u} \{ \langle x, y \rangle - f(u) - g(x-u) \} \\
&= \sup_{x,u} \{ (\langle u, y \rangle - f(u)) + (\langle x-u, y \rangle - g(x-u)) \} \\
&= f^*(y) + g^*(y).
\end{aligned}$$

b) $(f^* \# g^*)^* \overset{a)}{=} f^{**} + g^{**} \overset{f,g \in \Gamma_0}{=} f + g \overset{clear?}{\in} \Gamma$

Conjugacy of infimal convolution

Proposition 14 (Conjugacy of inf-convolution).

Let $f, g : \mathbb{E} \to \mathbb{R} \cup \{+\infty\}$. Then the following hold:

a) $(f \# g)^* = f^* + g^*$;

b) *If $f, g \in \Gamma_0$ such that $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$, then $(f+g)^* = \mathrm{cl}\,(f^* \# g^*)$.*

Proof.

a) For all $y \in \mathbb{E}$, we have

$$\begin{aligned}
(f\#g)^*(y) &= \sup_x \left\{\langle x, y\rangle - \inf_u \{f(u) + g(x-u)\}\right\} \\
&= \sup_{x,u} \{\langle x, y\rangle - f(u) - g(x-u)\} \\
&= \sup_{x,u} \{(\langle u, y\rangle - f(u)) + (\langle x-u, y\rangle - g(x-u))\} \\
&= f^*(y) + g^*(y).
\end{aligned}$$

b) $(f^* \# g^*)^* \overset{a)}{=} f^{**} + g^{**} \overset{f,g\in\Gamma_0}{=} f + g \overset{\text{clear?}}{\in} \Gamma$

$$\Longrightarrow \quad \mathrm{cl}\,(f^* \# g^*) = (f^* \# g^*)^{**} = (f+g)^*.$$

$\square$

Attouch-Brézis - drop the closure!

Theorem 15 (Attouch-Brézis).

Let $f, g \in \Gamma_0$ such that

$$\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g) \neq 0 \quad (CQ).$$

Then $(f+g)^* = f^* \# g^*$, and the infimal convolution is exact, i.e. the infimum in the infimal convolution is attained on $\mathrm{dom}\, f^* \# g^*$.

Proof.

On blackboard. □

Attouch-Brézis - drop the closure!

Theorem 15 (Attouch-Brézis).

Let $f, g \in \Gamma_0$ such that

$$\mathrm{ri}\,(\mathrm{dom}\, f) \cap \mathrm{ri}\,(\mathrm{dom}\, g) \neq 0 \quad (CQ).$$

Then $(f+g)^* = f^* \# g^*$, and the infimal convolution is exact, i.e. the infimum in the infimal convolution is attained on $\mathrm{dom}\, f^* \# g^*$.

Proof.

On blackboard. □

We note that (CQ) is always satisfied under any of the following:

- $\mathrm{int}\,(\mathrm{dom}\, f) \cap \mathrm{dom}\, g \neq \emptyset$,
- $\mathrm{dom}\, f = \mathbb{E}$,

and is equivalent to saying that

$$0 \in \mathrm{ri}\,(\mathrm{dom}\, f - \mathrm{dom}\, g).$$

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f.

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda}\|x-u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda}\|x-u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)
- $e_\lambda f \in C^{1,1} \cap \Gamma_0$

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda}\|x-u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)
- $e_\lambda f \in C^{1,1} \cap \Gamma_0$
- $\nabla e_\lambda f = \frac{1}{\lambda}(\mathrm{id} - P_\lambda f)$

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda}\|x-u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)
- $e_\lambda f \in C^{1,1} \cap \Gamma_0$
- $\nabla e_\lambda f = \frac{1}{\lambda}(\mathrm{id} - P_\lambda f)$
- $x \in \operatorname{argmin} f \iff x \in \operatorname{argmin} e_\lambda f \iff x = P_\lambda f(x)$ ($\to$ proximal point/gradient method)

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda}\|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda}\|x - u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)
- $e_\lambda f \in C^{1,1} \cap \Gamma_0$
- $\nabla e_\lambda f = \frac{1}{\lambda}(\mathrm{id} - P_\lambda f)$
- $x \in \operatorname{argmin} f \iff x \in \operatorname{argmin} e_\lambda f \iff x = P_\lambda f(x)$ ($\to$ proximal point/gradient method)
- $e_\lambda f \uparrow f \quad (\lambda \downarrow 0)$ (monotone pointwise convergence)

[1] Not in lecture notes!

Excursion: Moreau envelope and proximal operator [1]

Let $f \in \Gamma_0$ and $\lambda > 0$. Then

$$e_\lambda f := f \# \frac{1}{2\lambda} \|\cdot\|^2$$

is called the *Moreau envelope* of f. The map $P_\lambda f : \mathbb{E} \to \mathbb{E}$ given by

$$P_\lambda f(x) := \operatorname*{argmin}_u \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\}.$$

We have

- $P_\lambda f$ is 1-Lipschitz (in fact, firmly non-expansive)
- $e_\lambda f \in C^{1,1} \cap \Gamma_0$
- $\nabla e_\lambda f = \frac{1}{\lambda}(\mathrm{id} - P_\lambda f)$
- $x \in \operatorname{argmin} f \iff x \in \operatorname{argmin} e_\lambda f \iff x = P_\lambda f(x)$ ($\to$ proximal point/gradient method)
- $e_\lambda f \uparrow f \quad (\lambda \downarrow 0)$ (monotone pointwise convergence)
- $\operatorname{epi} e_\lambda f \to \operatorname{epi} f \quad (\lambda \downarrow 0)$ (epi-convergence)

[1] Not in lecture notes!

Conjugacy for convex-linear composites

Let $f \in \Gamma$ and $L \in \mathcal{L}(\mathbb{E}, \mathbb{E}')$. Then

$$Lf : \mathbb{E}' \to \overline{\mathbb{R}}, \quad (Lf)(y) := \inf \left\{ f(x) \;\middle|\; L(x) = y \right\}$$

is convex[2].

[2] Show that $\operatorname{epi}_{<} Lf = T(\operatorname{epi}_{<} f)$ for $T : (x, y) \mapsto (Tx, y)$.

Conjugacy for convex-linear composites

Let $f \in \Gamma$ and $L \in \mathcal{L}(\mathbb{E}, \mathbb{E}')$. Then

$$Lf : \mathbb{E}' \to \overline{\mathbb{R}}, \quad (Lf)(y) := \inf\left\{f(x) \,\middle|\, L(x) = y\right\}$$

is convex[2].

Proposition 16.

Let $g : \mathbb{E} \to \overline{\mathbb{R}}$ be proper and $L \in \mathcal{L}(\mathbb{E}, \mathbb{E}')$ and $T \in \mathcal{L}(\mathbb{E}', \mathbb{E})$. Then the following hold:

a) $(Lg)^* = g^* \circ L^*$.

b) $(g \circ T)^* = \mathrm{cl}\,(T^* g^*)$ *if* $g \in \Gamma$.

c) *The closure in b) can be dropped and the infimum is attained when finite if* $g \in \Gamma_0$ *and*

$$\mathrm{ri}\,(\mathrm{rge}\, T) \cap \mathrm{ri}\,(\mathrm{dom}\, g) \neq \emptyset. \tag{3}$$

Proof.

Notes and Part 2. □

[2] Show that $\mathrm{epi}_{<} Lf = T(\mathrm{epi}_{<} f)$ for $T : (x, y) \mapsto (Tx, y)$.

Infimal projection revisited

Theorem 17 (Infimal projection II).

Let $\psi \in \Gamma_0(\mathbb{E}_1 \times \mathbb{E}_2)$ and define $p : \mathbb{E}_1 \to \overline{\mathbb{R}}$ by

$$p(x) := \inf_v \psi(x, v). \tag{4}$$

Then the following hold:

a) *p is convex.*

b) *$p^* = \psi^*(\cdot, 0)$ which is closed and convex.*

c) *The condition*

$$\mathrm{dom}\, \psi^*(\cdot, 0) \neq 0 \tag{5}$$

is equivalent to having $p^* \in \Gamma_0$.

d) *If (5) holds then $p \in \Gamma_0$ and the infimum in its definition is attained when finite.*

Proof.

Blackboard/Notes. □

2. Conjugacy of composite functions via K-convexity and inf-convolution

Cone-induced ordering

Given a cone $K \subset \mathbb{E}$, the relation

$$x \leq_K y \quad :\Longleftrightarrow \quad y - x \in K \quad (x, y \in \mathbb{E})$$

induces an *ordering* on $\mathbb{E}$ which is a *partial ordering* if K is convex and pointed[3].

[3] i.e. $K \cap (-K) = \{0\}$

Cone-induced ordering

Given a cone $K \subset \mathbb{E}$, the relation

$$x \leq_K y \quad :\Longleftrightarrow \quad y - x \in K \quad (x, y \in \mathbb{E})$$

induces an *ordering* on $\mathbb{E}$ which is a *partial ordering* if K is convex and pointed[3].

- Attach to $\mathbb{E}$ a *largest element* $+\infty_\bullet$ w.r.t. $\leq_K$ which satisfies $x \leq_K +\infty_\bullet \quad (x \in \mathbb{E})$.

[3] i.e. $K \cap (-K) = \{0\}$

Cone-induced ordering

Given a cone $K \subset \mathbb{E}$, the relation

$$x \leq_K y \quad :\Longleftrightarrow \quad y - x \in K \quad (x, y \in \mathbb{E})$$

induces an *ordering* on $\mathbb{E}$ which is a *partial ordering* if K is convex and pointed[3].

- Attach to $\mathbb{E}$ a *largest element* $+\infty_\bullet$ w.r.t. $\leq_K$ which satisfies $x \leq_K +\infty_\bullet \quad (x \in \mathbb{E})$.
- Set $\mathbb{E}^\bullet := \mathbb{E} \cup \{+\infty_\bullet\}$.
- For $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ define

$$\begin{aligned} \mathrm{dom}\, F &:= \left\{x \in \mathbb{E}_1 \;\middle|\; F(x) \in \mathbb{E}_2\right\} \quad \text{(domain)},\\ \mathrm{gph}\, F &:= \{(x, F(x)) \in \mathbb{E}_1 \times \mathbb{E}_2 \mid x \in \mathrm{dom}\, F\} \quad \text{(graph)},\\ \mathrm{rge}\, F &:= \{F(x) \in \mathbb{E}_2 \mid x \in \mathrm{dom}\, F\} \quad (\textit{range}). \end{aligned}$$

[3] i.e. $K \cap (-K) = \{0\}$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \leq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \leq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \leq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0, 1])$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \preceq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \preceq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0, 1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\ F(x) \preceq_{\mathrm{ri}\,(K)} v\right\}$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \leq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff F(\lambda x + (1-\lambda)y) \leq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0,1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\; F(x) \leq_{\mathrm{ri}\,(K)} v\right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \preceq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \preceq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0,1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\ F(x) \preceq_{\mathrm{ri}\,(K)} v\right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

Examples:

- $K = \mathbb{R}^m_+$ and $F : \mathbb{R}^n \to (\mathbb{R}^m)^\bullet$ with $F_i \in \Gamma$ $(i = 1, \dots, m)$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \leq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \leq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0, 1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\; F(x) \leq_{\mathrm{ri}\,(K)} v\right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

Examples:

- $K = \mathbb{R}^m_+$ and $F : \mathbb{R}^n \to (\mathbb{R}^m)^\bullet$ with $F_i \in \Gamma$ $(i = 1, \dots, m)$
- $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \leq t\}$ and $F : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}$, $F(x) = (x, \|x\|)$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \preceq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \preceq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0,1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\; F(x) \preceq_{\mathrm{ri}\,(K)} v\right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

Examples:

- $K = \mathbb{R}^m_+$ and $F : \mathbb{R}^n \to (\mathbb{R}^m)^\bullet$ with $F_i \in \Gamma$ $(i = 1, \dots, m)$
- $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \le t\}$ and $F : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R},\; F(x) = (x, \|x\|)$
- $K = \mathbb{S}^n_+$ and $F : \mathbb{S}^n \to (\mathbb{S}^n)^\bullet, F(X) = \begin{cases} X^{-1}, & X \succ 0, \\ +\infty_\bullet, & \text{else} \end{cases}$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{(x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \preceq_K v\right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \preceq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0,1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{(x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\; F(x) \preceq_{\mathrm{ri}\,(K)} v\right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

Examples:

- $K = \mathbb{R}^m_+$ and $F : \mathbb{R}^n \to (\mathbb{R}^m)^\bullet$ with $F_i \in \Gamma$ $(i = 1, \dots, m)$
- $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \le t\}$ and $F : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R},\; F(x) = (x, \|x\|)$
- $K = \mathbb{S}^n_+$ and $F : \mathbb{S}^n \to (\mathbb{S}^n)^\bullet, F(X) = \begin{cases} X^{-1}, & X \succ 0, \\ +\infty_\bullet, & \text{else} \end{cases}$
- $K = \mathbb{S}^n_+$ and $F : \mathbb{R}^{m\times n} \to \mathbb{S}^n,\; F(X) = XX^T$

K-convexity

Definition 18 (K-convexity).

Let $K \subset \mathbb{E}_2$ be a cone and $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$. Then we call F *K-convex* if

$$K\text{-epi}\, F := \left\{ (x, v) \in \mathbb{E}_1 \times \mathbb{E}_2 \;\middle|\; F(x) \preceq_K v \right\} \quad (K\text{-epigraph})$$

is convex (in $\mathbb{E}_1 \times \mathbb{E}_2$).

- F is K-convex $\iff$ $F(\lambda x + (1-\lambda)y) \preceq_K \lambda F(x) + (1-\lambda)F(y) \quad (x, y \in \mathbb{E}_1, \lambda \in [0,1])$
- F K-convex, then $\mathrm{ri}\,(K\text{-epi}\, F) = \left\{ (x, v) \;\middle|\; x \in \mathrm{ri}\,(\mathrm{dom}\, F),\; F(x) \preceq_{\mathrm{ri}\,(K)} v \right\}$
- $K \subset L$ cones: F K-convex $\Rightarrow$ L-convex

Examples:

- $K = \mathbb{R}^m_+$ and $F : \mathbb{R}^n \to (\mathbb{R}^m)^\bullet$ with $F_i \in \Gamma$ $(i = 1, \dots, m)$
- $K = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \le t\}$ and $F : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}$, $F(x) = (x, \|x\|)$
- $K = \mathbb{S}^n_+$ and $F : \mathbb{S}^n \to (\mathbb{S}^n)^\bullet, F(X) = \begin{cases} X^{-1}, & X \succ 0, \\ +\infty_\bullet, & \text{else} \end{cases}$
- $K = \mathbb{S}^n_+$ and $F : \mathbb{R}^{m \times n} \to \mathbb{S}^n$, $F(X) = XX^T$
- K arbitrary, F affine.

Convexity of composite functions

For $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ and $g : \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ we define

$$(g \circ F)(x) := \begin{cases} g(F(x)) & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

Convexity of composite functions

For $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ and $g : \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ we define

$$(g \circ F)(x) := \begin{cases} g(F(x)) & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

Proposition 19.

Let $K \subset \mathbb{E}_2$ be a convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex and $g \in \Gamma(\mathbb{E}_2)$ such that $\operatorname{rge} F \cap \operatorname{dom} g \neq \emptyset$. If

$$g(F(x)) \leq g(y) \quad ((x, y) \in K\text{-}\operatorname{epi} F) \tag{6}$$

then the following hold:

a) *$g \circ F$ is convex and proper.*

b) *If g is lsc and F is continuous then $g \circ F$ is lower semicontinuous.*

Convexity of composite functions

For $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ and $g : \mathbb{E}_2 \to \mathbb{R} \cup \{+\infty\}$ we define

$$(g \circ F)(x) := \begin{cases} g(F(x)) & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

Proposition 19.

Let $K \subset \mathbb{E}_2$ be a convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex and $g \in \Gamma(\mathbb{E}_2)$ such that $\operatorname{rge} F \cap \operatorname{dom} g \neq \emptyset$. If

$$g(F(x)) \leq g(y) \quad ((x, y) \in K\text{-}\operatorname{epi} F) \tag{6}$$

then the following hold:

a) *$g \circ F$ is convex and proper.*

b) *If g is lsc and F is continuous then $g \circ F$ is lower semicontinuous.*

Condition (6) holds if g is *K-increasing*, i.e.

$$x \leq_K y \quad \Longrightarrow \quad g(x) \leq g(y).$$

Scalarization

Given $v \in \mathbb{E}_2$ and the linear form $\langle v, \cdot\rangle : \mathbb{E}_2 \to \mathbb{R}$, we set $\langle v, F\rangle := \langle v, \cdot\rangle \circ F$, i.e.

$$\langle v, F\rangle(x) = \begin{cases} \langle v, F(x)\rangle & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

Scalarization

Given $v \in \mathbb{E}_2$ and the linear form $\langle v, \cdot\rangle : \mathbb{E}_2 \to \mathbb{R}$, we set $\langle v, F\rangle := \langle v, \cdot\rangle \circ F$, i.e.

$$\langle v, F\rangle (x) = \begin{cases} \langle v, F(x)\rangle & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

For K a closed, convex cone we have:

- F is K-convex $\iff$ $\langle v, F\rangle$ is convex ($v \in -K^\circ$)
- $\sigma_{\operatorname{gph} F}(u, -v) = \langle v, F\rangle^* (u)$.
- $\sigma_{K\text{-}\operatorname{epi} F}(u, v) = \sigma_{\operatorname{gph} F}(u, v) + \delta_{K^\circ}(v)$

Scalarization

Given $v \in \mathbb{E}_2$ and the linear form $\langle v, \cdot\rangle : \mathbb{E}_2 \to \mathbb{R}$, we set $\langle v, F\rangle := \langle v, \cdot\rangle \circ F$, i.e.

$$\langle v, F\rangle (x) = \begin{cases} \langle v, F(x)\rangle & \text{if} \quad x \in \operatorname{dom} F, \\ +\infty & \text{else.} \end{cases}$$

For K a closed, convex cone we have:

- F is K-convex $\iff$ $\langle v, F\rangle$ is convex ($v \in -K^\circ$)
- $\sigma_{\operatorname{gph} F}(u, -v) = \langle v, F\rangle^* (u)$.
- $\sigma_{K\text{-epi}\, F}(u, v) = \sigma_{\operatorname{gph} F}(u, v) + \delta_{K^\circ}(v)$

Lemma 20 (Pennanen, JCA 1999).

Let $f \colon \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ with a convex domain and let $K \subset \mathbb{E}_2$ be the smallest closed convex cone with respect to which F is convex. Then

$$(-K)^\circ = \{v \in \mathbb{E}_2 \mid \langle v, F\rangle \text{ is convex}\}.$$

The main result

Theorem 21 (Conjugacy for composite function, H./Nguyen '19, Bot et. al '11).

Let $K \subset \mathbb{E}_2$ be a closed convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex such that K-$\mathrm{epi}\, F$ is closed and $g_0 \in \Gamma(\mathbb{E}_2)$ such that (6) *is satisfied, i.e.*

$$x \leq_K y \quad \Longrightarrow \quad g(x) \leq g(y).$$

Under the CQ

$$F(\mathrm{ri}\,(\mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset \tag{7}$$

we have

$$(g \circ F)^*(p) = \min_{v \in -K^\circ} g^*(v) + \langle v, F\rangle^*(p)$$

with $\mathrm{dom}\,(g \circ F)^* = \left\{p \in \mathbb{E}_1 \;\middle|\; \exists v \in \mathrm{dom}\, g^* \cap (-K^\circ) : \langle v, F\rangle^*(p) < +\infty\right\}.$

Proof.

Blackboard/Notes. □

The main result

Theorem 21 (Conjugacy for composite function, H./Nguyen '19, Bot et. al '11).

Let $K \subset \mathbb{E}_2$ be a closed convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex such that $K\text{-}\mathrm{epi}\, F$ is closed and $g_0 \in \Gamma(\mathbb{E}_2)$ such that (6) is satisfied, i.e.

$$x \leq_K y \quad \Longrightarrow \quad g(x) \leq g(y).$$

Under the CQ

$$F(\mathrm{ri}\,(\mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset \tag{7}$$

we have

$$(g \circ F)^*(p) = \min_{v \in -K^\circ} g^*(v) + \langle v, F\rangle^*(p)$$

with $\mathrm{dom}\,(g \circ F)^* = \left\{p \in \mathbb{E}_1 \,\middle|\, \exists v \in \mathrm{dom}\, g^* \cap (-K^\circ) : \langle v, F\rangle^*(p) < +\infty\right\}.$

Proof.

Blackboard/Notes. □

Remark:

- The CQ (7) is trivially satisfied if g is finite-valued.

The main result

Theorem 21 (Conjugacy for composite function, H./Nguyen '19, Bot et. al '11).

Let $K \subset \mathbb{E}_2$ be a closed convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex such that K-$\mathrm{epi}\, F$ is closed and $g_0 \in \Gamma(\mathbb{E}_2)$ such that (6) *is satisfied, i.e.*

$$x \leq_K y \quad \Longrightarrow \quad g(x) \leq g(y).$$

Under the CQ

$$F(\mathrm{ri}\,(\mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset \tag{7}$$

we have

$$(g \circ F)^*(p) = \min_{v \in -K^\circ} g^*(v) + \langle v, F\rangle^*(p)$$

with $\mathrm{dom}\,(g \circ F)^* = \left\{p \in \mathbb{E}_1 \;\middle|\; \exists v \in \mathrm{dom}\, g^* \cap (-K^\circ) : \langle v, F\rangle^*(p) < +\infty\right\}$.

Proof.

Blackboard/Notes. □

Remark:

- The CQ (7) is trivially satisfied if g is finite-valued.
- Condition (6) can be replaced by the stronger condition that g be K-increasing.

The main result

Theorem 21 (Conjugacy for composite function, H./Nguyen '19, Bot et. al '11).

Let $K \subset \mathbb{E}_2$ be a closed convex cone, $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ K-convex such that $K\text{-epi}\, F$ is closed and $g_0 \in \Gamma(\mathbb{E}_2)$ such that (6) is satisfied, i.e.

$$x \leq_K y \quad \Longrightarrow \quad g(x) \leq g(y).$$

Under the CQ

$$F(\mathrm{ri}\,(\mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset \tag{7}$$

we have

$$(g \circ F)^*(p) = \min_{v \in -K^\circ} g^*(v) + \langle v, F \rangle^*(p)$$

with $\mathrm{dom}\,(g \circ F)^* = \left\{ p \in \mathbb{E}_1 \;\middle|\; \exists v \in \mathrm{dom}\, g^* \cap (-K^\circ) : \langle v, F \rangle^*(p) < +\infty \right\}$.

Proof.

Blackboard/Notes. □

Remark:

- The CQ (7) is trivially satisfied if g is finite-valued.
- Condition (6) can be replaced by the stronger condition that g be K-increasing.
- $K\text{-epi}\, F$ is closed if F is continuous.

Extension to the additive composite setting

Corollary 22 (Conjugate of additive composite functions, H./Nguyen '19).

Under the assumptions of Theorem 21 let $f \in \Gamma_0$ *such that*

$$F(\mathrm{ri}\,(\mathrm{dom}\, f \cap \mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset. \tag{8}$$

Then

$$(f + g \circ F)^*(p) = \min_{\substack{v \in -K^\circ, \\ y \in \mathbb{E}_1}} g^*(v) + f^*(y) + \langle v, F\rangle^*(p - y).$$

Extension to the additive composite setting

Corollary 22 (Conjugate of additive composite functions, H./Nguyen '19).

Under the assumptions of Theorem 21 let $f \in \Gamma_0$ such that

$$F(\mathrm{ri}\,(\mathrm{dom}\, f \cap \mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g - K) \neq \emptyset. \tag{8}$$

Then

$$(f + g \circ F)^*(p) = \min_{\substack{v \in -K^\circ, \\ y \in \mathbb{E}_1}} g^*(v) + f^*(y) + \langle v, F \rangle^* (p - y).$$

Proof.

(Sketch) Apply Theorem 21 to $\tilde{g} : (s, y) \in \mathbb{R} \times \mathbb{E}_2 \mapsto s + g(y)$, $\tilde{F} : x \in \mathbb{E}_1 \to (f(x), x)$ and $\tilde{K} := \mathbb{R}_+ \times K$. $\square$

The case $K = -\operatorname{hzn} g$

For $g \in \Gamma_0$ its *horizon function* g^∞ is given via

$$\operatorname{epi} g^\infty = (\operatorname{epi} g)^\infty.$$

The *horizon cone of* g is

$$\operatorname{hzn} g := \left\{x \mid g^\infty(x) \le 0\right\}^\infty.$$

The case $K = -\operatorname{hzn} g$

For $g \in \Gamma_0$ its *horizon function* g^∞ is given via

$$\operatorname{epi} g^\infty = (\operatorname{epi} g)^\infty.$$

The *horizon cone of* g is

$$\operatorname{hzn} g := \left\{x \mid g^\infty(x) \le 0\right\}^\infty.$$

- $\operatorname{hzn} g = (\operatorname{cone}(\operatorname{dom} g^*))^\circ$

The case $K = -\operatorname{hzn} g$

For $g \in \Gamma_0$ its *horizon function* g^∞ is given via

$$\operatorname{epi} g^\infty = (\operatorname{epi} g)^\infty.$$

The *horizon cone of* g is

$$\operatorname{hzn} g := \left\{x \mid g^\infty(x) \leq 0\right\}^\infty.$$

- $\operatorname{hzn} g = (\operatorname{cone}(\operatorname{dom} g^*))^\circ$
- g is K-increasing for $K = -\operatorname{hzn} g$:

The case $K = -\operatorname{hzn} g$

For $g \in \Gamma_0$ its *horizon function* g^∞ is given via

$$\operatorname{epi} g^\infty = (\operatorname{epi} g)^\infty.$$

The *horizon cone of* g is

$$\operatorname{hzn} g := \left\{x \mid g^\infty(x) \le 0\right\}^\infty.$$

- $\operatorname{hzn} g = (\operatorname{cone}(\operatorname{dom} g^*))^\circ$
- g is K-increasing for $K = -\operatorname{hzn} g$: Let $x \le_K y$, i.e. $y = x + b$ for some $b \in K$. Then

$$g(x) = \sup_{z \in \operatorname{dom} g^*} \{\langle x, z\rangle - g^*(z)\} = \sup_{z \in \operatorname{dom} g^*} \{\langle y, z\rangle - \langle b, z\rangle - g^*(z)\} \le \sup_{z \in \operatorname{dom} g^*} \{\langle y, z\rangle - g^*(z)\} = g(y),$$

The case $K = -\mathrm{hzn}\, g$

For $g \in \Gamma_0$ its *horizon function* g^∞ is given via

$$\mathrm{epi}\, g^\infty = (\mathrm{epi}\, g)^\infty.$$

The *horizon cone of* g is

$$\mathrm{hzn}\, g := \left\{x \;\middle|\; g^\infty(x) \le 0\right\}^\infty.$$

- $\mathrm{hzn}\, g = (\mathrm{cone}\,(\mathrm{dom}\, g^*))^\circ$
- g is K-increasing for $K = -\mathrm{hzn}\, g$: Let $x \le_K y$, i.e. $y = x + b$ for some $b \in K$. Then

$$g(x) = \sup_{z \in \mathrm{dom}\, g^*} \{\langle x, z\rangle - g^*(z)\} = \sup_{z \in \mathrm{dom}\, g^*} \{\langle y, z\rangle - \langle b, z\rangle - g^*(z)\} \le \sup_{z \in \mathrm{dom}\, g^*} \{\langle y, z\rangle - g^*(z)\} = g(y),$$

Corollary 23 (Burke ’91, H./Nguyen ’19).

Let $g \in \Gamma_0(\mathbb{E}_2)$ and let $F : \mathbb{E}_1 \to \mathbb{E}_2^\bullet$ be $(-\mathrm{hzn}\, g)$-convex with $-\mathrm{hzn}\, g$-epi F closed such that

$$F(\mathrm{ri}\,(\mathrm{dom}\, F)) \cap \mathrm{ri}\,(\mathrm{dom}\, g + \mathrm{hzn}\, g) \ne \emptyset.$$

Then

$$(g \circ F)^*(p) = \min_{v \in \mathbb{E}_2} g^*(v) + \langle v, F\rangle^*(p).$$

The linear case

Corollary 24 (The linear case).

Let $g \in \Gamma(\mathbb{E}_2)$ *and* $F : \mathbb{E}_1 \to \mathbb{E}_2$ *linear such that*

$$\mathrm{rge}\, F \cap \mathrm{ri}\,(\mathrm{dom}\, g) \neq \emptyset.$$

Then

$$(g \circ F)^*(p) = \min_{v \in \mathbb{E}_2} \left\{ g^*(v) \;\middle|\; F^*(v) = p \right\}$$

with $\mathrm{dom}\,(g \circ F) = (F^*)^{-1}(\mathrm{dom}\, g^*)$.

The linear case

Corollary 24 (The linear case).

Let $g \in \Gamma(\mathbb{E}_2)$ and $F : \mathbb{E}_1 \to \mathbb{E}_2$ linear such that

$$\operatorname{rge} F \cap \operatorname{ri}(\operatorname{dom} g) \neq \emptyset.$$

Then

$$(g \circ F)^*(p) = \min_{v \in \mathbb{E}_2} \left\{ g^*(v) \mid F^*(v) = p \right\}$$

with $\operatorname{dom}(g \circ F) = (F^*)^{-1}(\operatorname{dom} g^*)$.

Proof.

We notice that F is $\{0\}$-convex. Hence we can apply Theorem 21 with $K = \{0\}$. Condition (7) then reads $\operatorname{rge} F \cap \operatorname{ri}(\operatorname{dom} g) \neq \emptyset$, which is our assumption. Hence we obtain

$$(g \circ F)^*(p) = \min_{v \in -K^\circ} g^*(v) + \langle v, F \rangle^*(p) = \min_{v \in \mathbb{E}_2} g^*(v) + \delta_{\{F^*(v)\}}(p).$$

$\square$

Conic programming duality

Consider the general conic program

$$\min f(x) \quad \text{s.t.} \quad F(x) \in -K \tag{9}$$

or equivalently

Conic programming duality

Consider the general conic program

$$\min f(x) \quad \text{s.t.} \quad F(x) \in -K \tag{9}$$

or equivalently

$$\min_{x \in \mathbb{E}_1} f(x) + (\delta_{-K} \circ F)(x) \tag{10}$$

where $f : \mathbb{E}_1 \to \mathbb{R}$ is convex, $F : \mathbb{E}_1 \to \mathbb{E}_2$ is K-convex and $K \subset \mathbb{E}_2$ is a closed, convex cone.

Conic programming duality

Consider the general conic program

$$\min f(x) \quad \text{s.t.} \quad F(x) \in -K \tag{9}$$

or equivalently

$$\min_{x \in \mathbb{E}_1} f(x) + (\delta_{-K} \circ F)(x) \tag{10}$$

where $f : \mathbb{E}_1 \to \mathbb{R}$ is convex, $F : \mathbb{E}_1 \to \mathbb{E}_2$ is K-convex and $K \subset \mathbb{E}_2$ is a closed, convex cone. The qualification condition (7) turns into a *generalized Slater condition*

$$\mathrm{rge}\, F \cap \mathrm{ri}\,(-K) \neq \emptyset. \tag{11}$$

Conic programming duality

Consider the general conic program

$$\min f(x) \quad \text{s.t.} \quad F(x) \in -K \tag{9}$$

or equivalently

$$\min_{x \in \mathbb{E}_1} f(x) + (\delta_{-K} \circ F)(x) \tag{10}$$

where $f : \mathbb{E}_1 \to \mathbb{R}$ is convex, $F : \mathbb{E}_1 \to \mathbb{E}_2$ is K-convex and $K \subset \mathbb{E}_2$ is a closed, convex cone. The qualification condition (7) turns into a *generalized Slater condition*

$$\mathrm{rge}\, F \cap \mathrm{ri}\,(-K) \neq \emptyset. \tag{11}$$

Conic programming duality

Consider the general conic program

$$\min f(x) \quad \text{s.t.} \quad F(x) \in -K \tag{9}$$

or equivalently

$$\min_{x\in\mathbb{E}_1} f(x) + (\delta_{-K} \circ F)(x) \tag{10}$$

where $f : \mathbb{E}_1 \to \mathbb{R}$ is convex, $F : \mathbb{E}_1 \to \mathbb{E}_2$ is K-convex and $K \subset \mathbb{E}_2$ is a closed, convex cone. The qualification condition (7) turns into a *generalized Slater condition*

$$\mathrm{rge}\, F \cap \mathrm{ri}\,(-K) \neq \emptyset. \tag{11}$$

Theorem 25 (Strong duality and dual attainment for conic programming).

Let $f : \mathbb{E}_1 \to \mathbb{R}$ is convex, $K \subset \mathbb{E}_2$ a closed, convex cone, and let $F : \mathbb{E}_1 \to \mathbb{E}_2$ be K-convex with closed K-epigraph. If (11) *holds then*

$$\inf_{x\in\mathbb{E}_1} f(x) + (\delta_{-K} \circ F)(x) = \max_{v\in -K^\circ} -f^*(y) - (\delta_{-K} \circ F)^*(-y) = \max_{v\in -K^\circ} \inf_{x\in\mathbb{E}_1} f(x) + \langle v, F(x)\rangle .$$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \dots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\dots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^{m} v_i f_i \right)^* (x).$$

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \dots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \ldots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\ldots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \ldots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\ldots,m} x_i.$$

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \ldots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \dots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\dots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \dots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\dots,m} x_i.$$

Then F is $\mathbb{R}^m_+$-convex and g is $\mathbb{R}^m_+$-increasing with $\operatorname{dom} g = \mathbb{R}^m$, and $g^* = \delta_{\Delta_m}$[4].

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \dots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \dots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\dots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \dots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\dots,m} x_i.$$

Then F is $\mathbb{R}^m_+$-convex and g is $\mathbb{R}^m_+$-increasing with $\operatorname{dom} g = \mathbb{R}^m$, and $g^* = \delta_{\Delta_m}$[4]. Hence

$$(g \circ F)^*(x)$$

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \dots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \dots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\dots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \dots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\dots,m} x_i.$$

Then F is $\mathbb{R}^m_+$-convex and g is $\mathbb{R}^m_+$-increasing with $\operatorname{dom} g = \mathbb{R}^m$, and $g^* = \delta_{\Delta_m}$[4]. Hence

$$(g \circ F)^*(x) = \min_{v \in \mathbb{R}^m_+} g^*(v) + \langle v, F \rangle^* (x)$$

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \dots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \dots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\dots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \dots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\dots,m} x_i.$$

Then F is $\mathbb{R}^m_+$-convex and g is $\mathbb{R}^m_+$-increasing with $\operatorname{dom} g = \mathbb{R}^m$, and $g^* = \delta_{\Delta_m}$[4]. Hence

$$(g \circ F)^*(x) = \min_{v \in \mathbb{R}^m_+} g^*(v) + \langle v, F \rangle^* (x) = \min_{v \in \mathbb{R}^m_+} \delta_{\Delta_m}(v) + \langle v, F \rangle^* (x)$$

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \ (i1, \dots, m) \right\}$

Conjugate of pointwise maximum of convex functions

Proposition 26.

For $f_1, \ldots, f_m \in \Gamma_0(\mathbb{E})$ define $f := \max_{i=1,\ldots,m} f_i$. Then $f \in \Gamma_0(\mathbb{E})$ with

$$f^*(x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x).$$

Proof.

We have $f = g \circ F$ for

$$F : x \mapsto \begin{cases} (f_1(x), \ldots, f_m(x)) & \text{if } x \in \bigcap_{i=1}^m \operatorname{dom} f_i, \\ +\infty_\bullet & \text{otherwise,} \end{cases} \quad \text{and} \quad g : y \mapsto \max_{i=1,\ldots,m} x_i.$$

Then F is $\mathbb{R}^m_+$-convex and g is $\mathbb{R}^m_+$-increasing with $\operatorname{dom} g = \mathbb{R}^m$, and $g^* = \delta_{\Delta_m}$[4]. Hence

$$(g \circ F)^*(x) = \min_{v \in \mathbb{R}^m_+} g^*(v) + \langle v, F \rangle^* (x) = \min_{v \in \mathbb{R}^m_+} \delta_{\Delta_m}(v) + \langle v, F \rangle^* (x) = \min_{v \in \Delta_m} \left(\sum_{i=1}^m v_i f_i \right)^* (x)$$

□

[4] $\Delta_m = \left\{ \lambda \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \; (i1, \ldots, m) \right\}$

3. A new class of matrix support functionals

Motivation I: Nuclear norm minimization/smoothing

Rank minimization (→ Netflix recommender problem)

$$\min_{X \in \mathbb{R}^{n\times m}} \operatorname{rank} X \quad \text{s.t.} \quad MX = B \quad (M \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}) \tag{12}$$

[5] $\sigma(X) = (\sigma_1, \dots, \sigma_n)$ is the vector of positive singular values of X.

Motivation I: Nuclear norm minimization/smoothing

Rank minimization (→ Netflix recommender problem)

$$\min_{X\in\mathbb{R}^{n\times m}} \operatorname{rank} X \quad \text{s.t.} \quad MX = B \quad (M \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}) \tag{12}$$

- Approximating the rank function (→ combinatorial)

$$\operatorname{rank} X = \|\sigma(X)\|_0 \overset{\text{Convex approx.}}{\sim} \|\sigma(X)\|_1 =: \|X\|_* \quad (\textit{nuclear norm})^5$$

[5] $\sigma(X) = (\sigma_1, \ldots, \sigma_n)$ is the vector of positive singular values of X.

Motivation I: Nuclear norm minimization/smoothing

Rank minimization (→ Netflix recommender problem)

$$\min_{X\in\mathbb{R}^{n\times m}} \operatorname{rank} X \quad \text{s.t.} \quad MX = B \quad (M \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}) \tag{12}$$

- Approximating the rank function (→ combinatorial)

$$\operatorname{rank} X = \|\sigma(X)\|_0 \overset{\text{Convex approx.}}{\sim} \|\sigma(X)\|_1 =: \|X\|_* \quad (\textit{nuclear norm})^5$$

- Convex approximation of (12)

$$\min_{X\in\mathbb{R}^{n\times m}} \|X\|_* \quad \text{s.t.} \quad MX = B$$

[5] $\sigma(X) = (\sigma_1, \ldots, \sigma_n)$ is the vector of positive singular values of X.

Motivation I: Nuclear norm minimization/smoothing

Rank minimization (→ Netflix recommender problem)

$$\min_{X\in\mathbb{R}^{n\times m}} \operatorname{rank} X \quad \text{s.t.} \quad MX = B \quad (M \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}) \tag{12}$$

- Approximating the rank function (→ combinatorial)

$$\operatorname{rank} X = \|\sigma(X)\|_0 \overset{\text{Convex approx.}}{\sim} \|\sigma(X)\|_1 =: \|X\|_* \quad (\textit{nuclear norm})^5$$

- Convex approximation of (12)

$$\min_{X\in\mathbb{R}^{n\times m}} \|X\|_* \quad \text{s.t.} \quad MX = B$$

- Hsieh/Olsen '14: $\|X\|_* = \min_{V\in\mathbb{S}^n_{++}} \frac{1}{2}\operatorname{tr}(V) + \frac{1}{2}\operatorname{tr}(X^TV^{-1}X) \quad (X \in \mathbb{R}^{n\times m})$

[5] $\sigma(X) = (\sigma_1, \dots, \sigma_n)$ is the vector of positive singular values of X.

Motivation I: Nuclear norm minimization/smoothing

Rank minimization (→ Netflix recommender problem)

$$\min_{X\in\mathbb{R}^{n\times m}} \operatorname{rank} X \quad \text{s.t.} \quad MX = B \quad (M \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}) \tag{12}$$

- Approximating the rank function (→ combinatorial)

$$\operatorname{rank} X = \|\sigma(X)\|_0 \overset{\text{Convex approx.}}{\sim} \|\sigma(X)\|_1 =: \|X\|_* \quad (\textit{nuclear norm})^5$$

- Convex approximation of (12)

$$\min_{X\in\mathbb{R}^{n\times m}} \|X\|_* \quad \text{s.t.} \quad MX = B$$

- Hsieh/Olsen '14: $\|X\|_* = \min_{V\in\mathbb{S}^n_{++}} \frac{1}{2}\operatorname{tr}(V) + \frac{1}{2}\operatorname{tr}(X^TV^{-1}X) \quad (X \in \mathbb{R}^{n\times m})$
- Smooth approximation of (12)

$$\min_{(X,V)\in\mathbb{R}^{n\times n}\times\mathbb{S}^n_{++}} \frac{1}{2}\operatorname{tr}(V) + \frac{1}{2}\operatorname{tr}(X^TV^{-1}X) \quad \text{s.t.} \quad MX = B$$

[5] $\sigma(X) = (\sigma_1, \dots, \sigma_n)$ is the vector of positive singular values of X.

Motivation II: Maximum likelihood estimation

Let $y_i \in \mathbb{R}^n$ $(i = 1, \ldots, N)$ be measurements of

$$y \sim N(\mu, \Sigma) \quad (\mu \in \mathbb{R}^n, \Sigma \in \mathbb{S}^n_{++} \to \text{unknown})$$

Motivation II: Maximum likelihood estimation

Let $y_i \in \mathbb{R}^n$ $(i = 1, \dots, N)$ be measurements of

$$y \sim N(\mu, \Sigma) \quad (\mu \in \mathbb{R}^n, \Sigma \in \mathbb{S}^n_{++} \to \text{unknown})$$

- Likelihood function:

$$\ell(\mu, \Sigma) := \frac{1}{(2\pi)^{n/2}} \prod_{i=1}^{N} \frac{1}{(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(y_i - \mu)^T \Sigma^{-1} (y_i - \mu)\right)$$

Motivation II: Maximum likelihood estimation

Let $y_i \in \mathbb{R}^n$ $(i = 1, \dots, N)$ be measurements of

$$y \sim N(\mu, \Sigma) \quad (\mu \in \mathbb{R}^n, \Sigma \in \mathbb{S}^n_{++} \to \text{unknown})$$

- Likelihood function:

$$\ell(\mu, \Sigma) := \frac{1}{(2\pi)^{n/2}} \prod_{i=1}^{N} \frac{1}{(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(y_i - \mu)^T \Sigma^{-1}(y_i - \mu)\right)$$

- log-likelihood function

$$\log \ell(\mu, \Sigma) = -\frac{N}{2} \log(\det \Sigma) - \frac{1}{2} \sum_{i=1}^{N} (y_i - \mu)^T \Sigma^{-1} (y_i - \mu) - \frac{n}{2} \log(2\pi)$$

Motivation II: Maximum likelihood estimation

Let $y_i \in \mathbb{R}^n$ $(i = 1, \dots, N)$ be measurements of

$$y \sim N(\mu, \Sigma) \quad (\mu \in \mathbb{R}^n, \Sigma \in \mathbb{S}^n_{++} \to \text{unknown})$$

- Likelihood function:

$$\ell(\mu, \Sigma) := \frac{1}{(2\pi)^{n/2}} \prod_{i=1}^{N} \frac{1}{(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(y_i - \mu)^T \Sigma^{-1} (y_i - \mu)\right)$$

- log-likelihood function

$$\log \ell(\mu, \Sigma) = -\frac{N}{2} \log(\det \Sigma) - \frac{1}{2} \sum_{i=1}^{N} (y_i - \mu)^T \Sigma^{-1} (y_i - \mu) - \frac{n}{2} \log(2\pi)$$

- Maximum likelihood estimation

$$\max_{(\mu,\Sigma)} \ell(\mu, \Sigma) \quad \Leftrightarrow \quad \min_{(\mu,\Sigma)} -\log \ell(\mu, \Sigma)$$

$$\overset{x_i := y_i - \mu}{\Leftrightarrow} \quad \min_{(X,\Sigma) \in \mathbb{R}^{n\times N} \times \mathbb{S}^n_{++}} \frac{1}{2} \operatorname{tr}\left(X^T \Sigma^{-1} X\right) + \frac{N}{2} \log(\det \Sigma)$$

The Moore-Penrose pseudoinverse

Theorem 27 (Moore-Penrose pseudoinverse).

Let $A \in R^{m\times n}$ with $\mathrm{rank}\, A = r$ and the singular value decomposition

$$A = U\Sigma V^T \quad \text{with} \quad \Sigma = \mathrm{diag}\,(\sigma_i), \quad U, V \text{ orthogonal.}$$

The Moore-Penrose pseudoinverse

Theorem 27 (Moore-Penrose pseudoinverse).

Let $A \in \mathbb{R}^{m\times n}$ with $\mathrm{rank}\, A = r$ and the singular value decomposition

$$A = U\Sigma V^T \quad \text{with} \quad \Sigma = \mathrm{diag}\,(\sigma_i), \quad U, V \text{ orthogonal.}$$

The matrix

$$A^\dagger := V\Sigma^\dagger U^T \quad \text{with} \quad \Sigma^\dagger := \begin{pmatrix} \sigma_1^{-1} & & & & & \\ & \ddots & & & & \\ & & \sigma_r^{-1} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

called the (Moore-Penrose) pseudoinverse *of A is the unique matrix with the following properties.*

The Moore-Penrose pseudoinverse

Theorem 27 (Moore-Penrose pseudoinverse).

Let $A \in R^{m\times n}$ with $\mathrm{rank}\, A = r$ and the singular value decomposition

$$A = U\Sigma V^T \quad \text{with} \quad \Sigma = \mathrm{diag}\,(\sigma_i), \quad U, V \text{ orthogonal.}$$

The matrix

$$A^\dagger := V\Sigma^\dagger U^T \quad \text{with} \quad \Sigma^\dagger := \begin{pmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

called the (Moore-Penrose) pseudoinverse *of A is the unique matrix with the following properties.*

a) $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$
b) $(AA^\dagger)^T = AA^\dagger$ and $(A^\dagger A)^T = A^\dagger As$

The Moore-Penrose pseudoinverse

Theorem 27 (Moore-Penrose pseudoinverse).

Let $A \in R^{m\times n}$ with $\mathrm{rank}\, A = r$ and the singular value decomposition

$$A = U\Sigma V^T \quad \text{with} \quad \Sigma = \mathrm{diag}\,(\sigma_i), \quad U, V \text{ orthogonal.}$$

The matrix

$$A^\dagger := V\Sigma^\dagger U^T \quad \text{with} \quad \Sigma^\dagger := \begin{pmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

called the (Moore-Penrose) pseudoinverse *of A is the unique matrix with the following properties.*

a) $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$

b) $(AA^\dagger)^T = AA^\dagger$ and $(A^\dagger A)^T = A^\dagger As$

Moreover:

c) A invertible $\quad\Rightarrow\quad A^\dagger = A^{-1}$

d) $A \succ 0 \quad\Rightarrow\quad A^\dagger \succ 0$

The closure of the matrix-fractional function

Put $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$.

$$\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{-1} X) & \text{if} \quad V \succ 0, \\ +\infty & \text{else.} \end{cases} \qquad \text{(matrix-fractional function)}$$

The closure of the matrix-fractional function

Put $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$.

$$\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{-1} X) & \text{if} \quad V \succ 0, \\ +\infty & \text{else.} \end{cases} \qquad \text{(matrix-fractional function)}$$

$$\overset{\text{Schur}}{\Rightarrow} \qquad \mathrm{epi}\,\phi = \left\{(X, V, \alpha) \mid \exists\, Y \in \mathbb{S}^m : \begin{pmatrix} V & X \\ X^T & Y \end{pmatrix} \succeq 0,\ V \succ 0,\ \frac{1}{2}\mathrm{tr}\,(Y) \le \alpha \right\}$$

The closure of the matrix-fractional function

Put $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$.

$$\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{-1} X) & \text{if} \quad V \succ 0, \\ +\infty & \text{else.} \end{cases} \qquad \text{(matrix-fractional function)}$$

$$\overset{\text{Schur}}{\Rightarrow} \qquad \mathrm{epi}\,\phi = \left\{ (X, V, \alpha) \mid \exists\, Y \in \mathbb{S}^m : \begin{pmatrix} V & X \\ X^T & Y \end{pmatrix} \succeq 0,\ V \succ 0,\ \frac{1}{2}\mathrm{tr}\,(Y) \le \alpha \right\}$$

$\Rightarrow$ $\qquad \phi$ proper, sublinear and <u>not</u> lsc.

The closure of the matrix-fractional function

Put $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$.

$$\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{-1} X) & \text{if} \quad V \succ 0, \\ +\infty & \text{else.} \end{cases} \qquad \text{(matrix-fractional function)}$$

$$\overset{\text{Schur}}{\Rightarrow} \qquad \mathrm{epi}\,\phi = \left\{(X, V, \alpha) \mid \exists\, Y \in \mathbb{S}^m : \begin{pmatrix} V & X \\ X^T & Y \end{pmatrix} \succeq 0,\ V \succ 0,\ \frac{1}{2}\mathrm{tr}\,(Y) \le \alpha \right\}$$

$\Rightarrow$ ϕ proper, sublinear and <u>not</u> lsc.

$$\Rightarrow \qquad \mathrm{cl}\,\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{\dagger} X) & \text{if} \quad V \succeq 0, \mathrm{rge}\, X \in \mathrm{rge}\, V, \\ +\infty & \text{else} \end{cases}$$

is proper, lsc and sublinear

The closure of the matrix-fractional function

Put $\mathbb{E} := \mathbb{R}^{n\times m} \times \mathbb{S}^n$.

$$\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{-1} X) & \text{if} \quad V \succ 0, \\ +\infty & \text{else.} \end{cases} \qquad \text{(matrix-fractional function)}$$

$$\overset{\text{Schur}}{\Rightarrow} \qquad \mathrm{epi}\,\phi = \left\{(X, V, \alpha) \,|\, \exists\, Y \in \mathbb{S}^m : \begin{pmatrix} V & X \\ X^T & Y \end{pmatrix} \succeq 0,\ V \succ 0,\ \tfrac{1}{2}\mathrm{tr}\,(Y) \leq \alpha\right\}$$

$\Rightarrow$ ϕ proper, sublinear and <u>not</u> lsc.

$$\Rightarrow \qquad \mathrm{cl}\,\phi : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^{\dagger} X) & \text{if} \quad V \succeq 0, \mathrm{rge}\, X \in \mathrm{rge}\, V, \\ +\infty & \text{else} \end{cases}$$

is proper, lsc and sublinear

$\overset{\text{Hörmander's Theorem}}{\Rightarrow}$ $\mathrm{cl}\,\phi$ is a support function

Motivation III: Quadratic programming

For $A \in \mathbb{R}^{p\times n}$ and $V \in \mathbb{S}^n$ put

$$M(V) := \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{K}_A := \left\{ V \in \mathbb{S}^n \;\middle|\; u^T V u \geq 0 \ (u \in \ker A) \right\}.$$

Motivation III: Quadratic programming

For $A \in \mathbb{R}^{p\times n}$ and $V \in \mathbb{S}^n$ put

$$M(V) := \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{K}_A := \left\{ V \in \mathbb{S}^n \;\middle|\; u^T V u \geq 0 \ (u \in \ker A) \right\}.$$

Theorem 28 (Burke, H. '15).

For $b \in \operatorname{rge} A$, we have

$$\inf_{u\in\mathbb{R}^n} \left\{ \frac{1}{2} u^T V u - x^T u \mid Au = b \right\} = \begin{cases} -\frac{1}{2}\begin{pmatrix} x \\ b \end{pmatrix}^T M(V)^\dagger \begin{pmatrix} x \\ b \end{pmatrix} & \text{if} \quad x \in \operatorname{rge}[V\ A^T],\ V \in \mathcal{K}_A, \\ -\infty & \text{else.} \end{cases}$$

Motivation III: Quadratic programming

For $A \in \mathbb{R}^{p\times n}$ and $V \in \mathbb{S}^n$ put

$$M(V) := \begin{pmatrix} V & A^T \\ A & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{K}_A := \left\{ V \in \mathbb{S}^n \mid u^T V u \geq 0 \ (u \in \ker A) \right\}.$$

Theorem 28 (Burke, H. '15).

For $b \in \mathrm{rge}\, A$, we have

$$\inf_{u\in\mathbb{R}^n} \left\{ \frac{1}{2} u^T V u - x^T u \mid Au = b \right\} = \begin{cases} -\frac{1}{2} \begin{pmatrix} x \\ b \end{pmatrix}^T M(V)^\dagger \begin{pmatrix} x \\ b \end{pmatrix} & \text{if} \quad x \in \mathrm{rge}\,[V\ A^T],\ V \in \mathcal{K}_A, \\ -\infty & \text{else.} \end{cases}$$

Question: For $A \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}$, is

$$\varphi_{A,B} : (X, V) \in \mathbb{E} \mapsto \begin{cases} \frac{1}{2} \mathrm{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B} \right) & \text{if} \ \ \mathrm{rge}\,\binom{X}{B} \subset \mathrm{rge}\, M(V),\ V \in \mathcal{K}_A, \\ +\infty & \text{else} \end{cases}$$

a support function?

A new class of matrix support functions

Define

$$\mathcal{D}(A,B) := \left\{ \left(Y, -\frac{1}{2} YY^T \right) \in \mathbb{E} \;\middle|\; Y \in \mathbb{R}^{n \times m} : \; AY = B \right\} \quad (A \in \mathbb{R}^{p \times n}, B \in \mathbb{R}^{p \times m}).$$

A new class of matrix support functions

Define

$$\mathcal{D}(A,B) := \left\{ \left(Y, -\frac{1}{2} YY^T \right) \in \mathbb{E} \;\middle|\; Y \in \mathbb{R}^{n\times m} : \; AY = B \right\} \quad (A \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}).$$

Theorem 29 (Burke, H. '15).

For $\mathrm{rge}\, B \subset \mathrm{rge}\, A$

$$\sigma_{\mathcal{D}(A,B)}(X,V) = \begin{cases} \frac{1}{2}\mathrm{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B} \right) & \text{if } \; \mathrm{rge}\, \binom{X}{B} \subset \mathrm{rge}\, M(V), \;\; V \in \mathcal{K}_A, \\ +\infty & \text{else} \end{cases} \qquad ((X,V) \in \mathbb{E})$$

with

$$\mathrm{int}\,(\mathrm{dom}\, \sigma_{D(A,B)}) = \{ (X,V) \in \mathbb{E} \mid V \in \mathrm{int}\, \mathcal{K}_A \}.$$

A new class of matrix support functions

Define

$$\mathcal{D}(A,B) := \left\{ \left(Y, -\frac{1}{2}YY^T \right) \in \mathbb{E} \;\middle|\; Y \in \mathbb{R}^{n\times m} : \; AY = B \right\} \quad (A \in \mathbb{R}^{p\times n}, B \in \mathbb{R}^{p\times m}).$$

Theorem 29 (Burke, H. '15).

For $\mathrm{rge}\, B \subset \mathrm{rge}\, A$

$$\sigma_{\mathcal{D}(A,B)}(X,V) = \begin{cases} \frac{1}{2}\mathrm{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B} \right) & \text{if } \; \mathrm{rge}\, \binom{X}{B} \subset \mathrm{rge}\, M(V), \; V \in \mathcal{K}_A, \\ +\infty & \text{else} \end{cases} \quad ((X,V) \in \mathbb{E})$$

with

$$\mathrm{int}\,(\mathrm{dom}\, \sigma_{D(A,B)}) = \{(X,V) \in \mathbb{E} \mid V \in \mathrm{int}\, \mathcal{K}_A\}.$$

In particular,

$$\sigma_{\mathcal{D}(0,0)}(X,V) = \begin{cases} \frac{1}{2}\mathrm{tr}\,(X^T V^\dagger X) & \text{if} \quad V \succeq 0, \mathrm{rge}\, X \subset \mathrm{rge}\, V, \\ +\infty & \text{else} \end{cases} \quad = \mathrm{cl}\, \phi(X,V).$$

Proof.

Blackboard/Notes. □

Closed convex hull of $\mathcal{D}(A, B)$: Carathéodory-based description

Recall

$$\partial\sigma_{\mathcal{D}(A,B)}(X, V) = \left\{(Y, W) \in \overline{\mathrm{conv}}\,\mathcal{D}(A, B) \;\middle|\; (X, V) \in N_{\overline{\mathrm{conv}}\,\mathcal{D}(A,B)}(Y, W)\right\} \quad \text{and} \quad \sigma_{\mathcal{D}(A,B)} = \delta^*_{\overline{\mathrm{conv}}\,\mathcal{D}(A,B)}$$

where

$$\mathcal{D}(A, B) := \left\{\left(Y, -\frac{1}{2}YY^T\right) \in \mathbb{E} \;\middle|\; Y \in \mathbb{R}^{n\times m} : \; AY = B\right\}.$$

[6] $d \otimes I_m = (d_i I_m) \in \mathbb{R}^{m(\kappa+1)}$
[7] $\kappa := \dim \mathbb{E}$

Closed convex hull of $\mathcal{D}(A, B)$: Carathéodory-based description

Recall

$$\partial\sigma_{\mathcal{D}(A,B)}(X, V) = \left\{(Y, W) \in \overline{\mathrm{conv}}\, \mathcal{D}(A, B) \;\middle|\; (X, V) \in N_{\overline{\mathrm{conv}}\, \mathcal{D}(A,B)}(Y, W)\right\} \quad \text{and} \quad \sigma_{\mathcal{D}(A,B)} = \delta^*_{\overline{\mathrm{conv}}\, \mathcal{D}(A,B)}$$

where

$$\mathcal{D}(A, B) := \left\{\left(Y, -\frac{1}{2}YY^T\right) \in \mathbb{E} \;\middle|\; Y \in \mathbb{R}^{n\times m} : \; AY = B\right\}.$$

Proposition 30 (Burke, H. '15).

$$\overline{\mathrm{conv}}\, \mathcal{D}(A, B) = \left\{(Z(d \otimes I_m), -\frac{1}{2}ZZ^T) \;\middle|\; (d, Z) \in \mathcal{F}(A, B)\right\}.^6$$

where

$$\mathcal{F}(A, B) := \left\{(d, Z) \in \mathbb{R}^{\kappa+1} \times \mathbb{R}^{n\times m(\kappa+1)} \;\middle|\; \begin{array}{l} d \geq 0, \; \|d\| = 1, \\ AZ_i = d_i B \; (i = 1, \dots, \kappa + 1) \end{array}\right\}.^7$$

---

[6] $d \otimes I_m = (d_i I_m) \in \mathbb{R}^{m(\kappa+1)}$

[7] $\kappa := \dim \mathbb{E}$

Closed convex hull of $\mathcal{D}(A, B)$: A new description

Define

$$\Omega(A, B) := \left\{ (Y, W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ \right\}, \tag{13}$$

and observe that

$$\mathcal{K}_A^\circ = \mathbb{R}_+ \mathrm{conv}\left\{ -vv^T \mid v \in \ker A \right\}.$$

Closed convex hull of $\mathcal{D}(A, B)$: A new description

Define

$$\Omega(A,B) := \left\{ (Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ \right\}, \tag{13}$$

and observe that

$$\mathcal{K}_A^\circ = \mathbb{R}_+ \mathrm{conv}\left\{ -vv^T \mid v \in \ker A \right\}.$$

Theorem 31 (Burke, Gao, H. '17).

We have

$$\overline{\mathrm{conv}}\, \mathcal{D}(A,B) = \Omega(A,B).$$

In particular,

$$\overline{\mathrm{conv}}\, \mathcal{D}(0,0) = \left\{ (Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\; \textit{and} \;\; \frac{1}{2}YY^T + W \preceq 0 \right\}.$$

Proof.

Notes. □

Closed convex hull of $\mathcal{D}(A,B)$: A new description

Define

$$\Omega(A,B) := \left\{ (Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ \right\}, \tag{13}$$

and observe that

$$\mathcal{K}_A^\circ = \mathbb{R}_+ \mathrm{conv}\left\{ -vv^T \mid v \in \ker A \right\}.$$

Theorem 31 (Burke, Gao, H. '17).

We have

$$\overline{\mathrm{conv}}\,\mathcal{D}(A,B) = \Omega(A,B).$$

In particular,

$$\overline{\mathrm{conv}}\,\mathcal{D}(0,0) = \left\{ (Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\; \textit{and} \;\; \frac{1}{2}YY^T + W \preceq 0 \right\}.$$

Proof.

Notes. □

Corollary 32 (Conjugate of GMF).

We have

$$\sigma^*_{\mathcal{D}(A,B)} = \delta_{\Omega(A,B)}.$$

Convex geometry of $\Omega(A, B)$

Recall that $\Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}$

Convex geometry of $\Omega(A, B)$

Recall that $\Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}$

Proposition 33 (Burke, Gao, H. ’17).

Let $\Omega(A, B)$ be given as above. Then:

a) $\mathrm{ri}\,\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathrm{ri}\,(\mathcal{K}_A^\circ)\right\}.$

Convex geometry of $\Omega(A, B)$

Recall that $\Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}$

Proposition 33 (Burke, Gao, H. '17).

Let $\Omega(A,B)$ be given as above. Then:

a) $\operatorname{ri}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \operatorname{ri}(\mathcal{K}_A^\circ)\right\}.$

b) $\operatorname{aff}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \operatorname{span}\mathcal{K}_A^\circ\right\}.$

Convex geometry of $\Omega(A,B)$

Recall that $\Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ \right\}$

Proposition 33 (Burke, Gao, H. '17).

Let $\Omega(A,B)$ be given as above. Then:

a) $\operatorname{ri}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\; \textit{and} \;\; \frac{1}{2}YY^T + W \in \operatorname{ri}(\mathcal{K}_A^\circ)\right\}$.

b) $\operatorname{aff}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \;\; \textit{and} \;\; \frac{1}{2}YY^T + W \in \operatorname{span}\mathcal{K}_A^\circ\right\}$.

c) $\Omega(A,B)^\circ = \left\{(X,V) \;\middle|\; \operatorname{rge}\binom{X}{B} \subset \operatorname{rge} M(V),\; V \in \mathcal{K}_A,\; \frac{1}{2}\operatorname{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B}\right) \le 1\right\}$.

Convex geometry of $\Omega(A, B)$

Recall that $\Omega(A, B) := \left\{(Y, W) \in \mathbb{E} \;\middle|\; AY = B \;\text{ and }\; \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}$

Proposition 33 (Burke, Gao, H. '17).

Let $\Omega(A, B)$ be given as above. Then:

a) $\operatorname{ri}\Omega(A, B) = \left\{(Y, W) \in \mathbb{E} \;\middle|\; AY = B \;\textit{ and }\; \frac{1}{2}YY^T + W \in \operatorname{ri}(\mathcal{K}_A^\circ)\right\}$.

b) $\operatorname{aff}\Omega(A, B) = \left\{(Y, W) \in \mathbb{E} \;\middle|\; AY = B \;\textit{ and }\; \frac{1}{2}YY^T + W \in \operatorname{span}\mathcal{K}_A^\circ\right\}$.

c) $\Omega(A, B)^\circ = \left\{(X, V) \;\middle|\; \operatorname{rge}\binom{X}{B} \subset \operatorname{rge} M(V),\; V \in \mathcal{K}_A,\; \frac{1}{2}\operatorname{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B}\right) \le 1\right\}$.

d) $\Omega(A, B)^\infty = \{0_{n\times m}\} \times \mathcal{K}_A^\circ$.

Convex geometry of $\Omega(A, B)$

Recall that $\Omega(A,B) := \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \mathcal{K}_A^\circ\right\}$

Proposition 33 (Burke, Gao, H. '17).

Let $\Omega(A,B)$ be given as above. Then:

a) $\operatorname{ri}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \operatorname{ri}(\mathcal{K}_A^\circ)\right\}$.

b) $\operatorname{aff}\Omega(A,B) = \left\{(Y,W) \in \mathbb{E} \;\middle|\; AY = B \text{ and } \frac{1}{2}YY^T + W \in \operatorname{span}\mathcal{K}_A^\circ\right\}$.

c) $\Omega(A,B)^\circ = \left\{(X,V) \;\middle|\; \operatorname{rge}\binom{X}{B} \subset \operatorname{rge} M(V),\ V \in \mathcal{K}_A,\ \frac{1}{2}\operatorname{tr}\left(\binom{X}{B}^T M(V)^\dagger \binom{X}{B}\right) \le 1\right\}$.

d) $\Omega(A,B)^\infty = \{0_{n\times m}\} \times \mathcal{K}_A^\circ$.

Proposition 34 (Burke, Gao, H. '17).

Let $\Omega(A,B)$ be given as above and let $(Y,W) \in \Omega(A,B)$. Then

$$N_{\Omega(A,B)}(Y,W) = \left\{(X,V) \in \mathbb{E} \;\middle|\; \begin{array}{l} V \in \mathcal{K}_A,\ \left\langle V, \dfrac{1}{2}YY^T + W\right\rangle = 0 \\ \text{and } \operatorname{rge}(X - VY) \subset (\ker A)^\perp \end{array}\right\}.$$

Subdifferentiation of the GMF

For any set C recall that

$$\partial\sigma_C(x) = \left\{z \in \overline{\mathrm{conv}}\, C \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, C}(z)\right\} \tag{14}$$

Subdifferentiation of the GMF

For any set C recall that

$$\partial\sigma_C(x) = \left\{ z \in \overline{\mathrm{conv}}\, C \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, C}(z) \right\} \tag{14}$$

Corollary 35 (The subdifferential of $\sigma_{\mathcal{D}(A,B)}$).

For all $(X, V) \in \mathrm{dom}\, \sigma_{\mathcal{D}(A,B)}$, *we have*

$$\partial\sigma_{\mathcal{D}(A,B)} = \left\{ (Y, W) \in \Omega(A, B) \;\middle|\; \begin{array}{l} \exists Z \in \mathbb{R}^{p\times m} : \; X = VY + A^T Z, \\ \left\langle V, \, \frac{1}{2} YY^T + W \right\rangle = 0 \end{array} \right\}.$$

Subdifferentiation of the GMF

For any set C recall that

$$\partial\sigma_C(x) = \left\{ z \in \overline{\mathrm{conv}}\, C \;\middle|\; x \in N_{\overline{\mathrm{conv}}\, C}(z) \right\} \tag{14}$$

Corollary 35 (The subdifferential of $\sigma_{\mathcal{D}(A,B)}$).

For all $(X, V) \in \mathrm{dom}\, \sigma_{\mathcal{D}(A,B)}$, *we have*

$$\partial\sigma_{\mathcal{D}(A,B)} = \left\{ (Y, W) \in \Omega(A, B) \;\middle|\; \begin{array}{l} \exists Z \in \mathbb{R}^{p \times m} :\; X = VY + A^T Z, \\ \left\langle V, \frac{1}{2} YY^T + W \right\rangle = 0 \end{array} \right\}.$$

Corollary 36.

The GMF $\sigma_{\mathcal{D}(A,B)}$ *is (continuously) differentiable on the interior of its domain with*

$$\nabla\sigma_{\mathcal{D}(A,B)}(X, V) = \left(Y, -\frac{1}{2} YY^T \right) \quad ((X, V) \in \mathrm{int}\,(\mathrm{dom}\, \sigma_{\mathcal{D}(A,B)}))$$

where $Y := A^{\dagger} B + (P(P^T VP)^{-1} P^T)(X - A^{\dagger} X)$, $P \in \mathbb{R}^{n \times (n-p)}$ *is any matrix whose columns form an orthonormal basis of* $\ker A$ *and* $p := \mathrm{rank}\, A$.

Conjugate of variational Gram functions

For $M \subset \mathbb{S}^n_+$ (w.lo.g.) closed, convex, the associated *variational Gram function (VGF)* is given by

$$\Omega_M : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{+\infty\}, \quad \Omega_M(X) = \frac{1}{2}\sigma_M(XX^T).$$

With

$$F : \mathbb{R}^{n\times m} \to \mathbb{S}^n, \quad F(X) = \frac{1}{2}XX^T. \tag{15}$$

$\Omega_M = \sigma_M \circ F$ fits the composite scheme studied in Section 2.

Conjugate of variational Gram functions

For $M \subset \mathbb{S}^n_+$ (w.lo.g.) closed, convex, the associated *variational Gram function (VGF)* is given by

$$\Omega_M : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{+\infty\}, \quad \Omega_M(X) = \frac{1}{2}\sigma_M(XX^T).$$

With

$$F : \mathbb{R}^{n\times m} \to \mathbb{S}^n, \quad F(X) = \frac{1}{2}XX^T. \tag{15}$$

$\Omega_M = \sigma_M \circ F$ fits the composite scheme studied in Section 2.

- $\mathbb{S}^n_+$ is the smallest closed convex cone in $\mathbb{S}^n$ with respect to which F is convex;
- $-\mathrm{hzn}\,\sigma_M \supset \mathbb{S}^n_+$. In particular, F is $(-\mathrm{hzn}\,\sigma_M)$-convex.

Conjugate of variational Gram functions

For $M \subset \mathbb{S}^n_+$ (w.lo.g.) closed, convex, the associated *variational Gram function (VGF)* is given by

$$\Omega_M : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{+\infty\}, \quad \Omega_M(X) = \frac{1}{2}\sigma_M(XX^T).$$

With

$$F : \mathbb{R}^{n\times m} \to \mathbb{S}^n, \quad F(X) = \frac{1}{2}XX^T. \tag{15}$$

$\Omega_M = \sigma_M \circ F$ fits the composite scheme studied in Section 2.

- $\mathbb{S}^n_+$ is the smallest closed convex cone in $\mathbb{S}^n$ with respect to which F is convex;
- $-\mathrm{hzn}\,\sigma_M \supset \mathbb{S}^n_+$. In particular, F is $(-\mathrm{hzn}\,\sigma_M)$-convex.

Theorem 37 (Jalali et al. '17/ Burke, Gao, H. '19).

Let $M \subset \mathbb{S}^n_+$ be nonempty, convex and compact. Then Ω^*_M is finite-valued and given by

$$\Omega^*(X) \quad = \quad \frac{1}{2}\min_{V\in M}\left\{\mathrm{tr}\,(X^TV^\dagger X) \;\middle|\; \mathrm{rge}\,X \subset \mathrm{rge}\,V\right\}.$$

Proof.

Blackboard/Notes. □

Nuclear norm smoothing

For $A \in \mathbb{R}^{p\times n}$ set

$$\mathrm{Ker}\, A := \{V \in \mathbb{R}^{n\times n} \mid AV = 0\} \quad \text{and} \quad \mathrm{Rge}\, A := \left\{W \in \mathbb{R}^{n\times n} \;\middle|\; \mathrm{rge}\, W \subset \mathrm{rge}\, A\right\}.$$

Nuclear norm smoothing

For $A \in \mathbb{R}^{p \times n}$ set

$$\mathrm{Ker}\, A := \{V \in \mathbb{R}^{n \times n} \mid AV = 0\} \quad \text{and} \quad \mathrm{Rge}\, A := \left\{W \in \mathbb{R}^{n \times n} \;\middle|\; \mathrm{rge}\, W \subset \mathrm{rge}\, A\right\}.$$

Theorem 38.

Let $p : \mathbb{R}^{n \times m} \to \overline{\mathbb{R}}$ be defined by

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,0)}(X, V) + \left\langle \bar{U},\, V \right\rangle$$

for some $\bar{U} \in \mathbb{S}^n_+ \cap \mathrm{Ker}\, A$ and set $C(\bar{U}) := \left\{Y \;\middle|\; \frac{1}{2} YY^T \preceq \bar{U}\right\}$. Then we have:

a) *$p^* = \delta_{C(\bar{U}) \cap \mathrm{Ker}\, A}$ is closed, proper, convex.*

b) *$p = \sigma_{C(\bar{U}) \cap \mathrm{Ker}\, A} = \gamma_{C(\bar{U})^\circ + \mathrm{Rge}\, A^T}$ is sublinear, finite-valued, nonnegative and symmetric (i.e. a seminorm).*

c) *If $\bar{U} \succ 0$ with $2\bar{U} = LL^T$ ($L \in \mathbb{R}^{n \times n}$) and $A = 0$ then $p = \sigma_{C(\bar{U})} = \|L^T(\cdot)\|_*$, i.e. p is a norm with $C(\bar{U})^\circ$ as its unit ball and $\gamma_{C(\bar{U})}$ as its dual norm.*

Proof.

Blackboard/Notes. □

Current and future directions

- K-convexity
  - When is $\overline{\mathrm{conv}}\,(\mathrm{gph}\,F) = K\text{-}\mathrm{epi}\,F$ for F K-convex ?

Current and future directions

- *K-convexity*
  - When is $\overline{\mathrm{conv}}\,(\mathrm{gph}\, F) = K\text{-}\mathrm{epi}\, F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)

Current and future directions

- *K-convexity*
  - When is $\overline{\text{conv}}\,(\text{gph}\, F) = K\text{-epi}\, F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)
  - Learn more about existing literature!
  -

Current and future directions

- <u>K-convexity</u>
  - When is $\overline{\text{conv}}\,(\text{gph}\, F) = K\text{-epi}\, F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)
  - Learn more about existing literature!
  -
- <u>Generalized matrix-fractional function</u>
  - Systematic study of (partial) infimal projections

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,B)}(X, V) + h(V).$$

for $h \in \Gamma_0(\mathbb{S}^n)$. → SIOPT article to appear.

Current and future directions

- K-convexity
  - When is $\overline{\text{conv}}\,(\text{gph}\,F) = K\text{-epi}\,F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)
  - Learn more about existing literature!
  -
- Generalized matrix-fractional function
  - Systematic study of (partial) infimal projections

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,B)}(X, V) + h(V).$$

    for $h \in \Gamma_0(\mathbb{S}^n)$. $\rightarrow$ SIOPT article to appear.
  - Numerical methods based on GMF.

Current and future directions

- K-convexity
  - When is $\overline{\mathrm{conv}}\,(\mathrm{gph}\, F) = K\text{-epi}\, F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)
  - Learn more about existing literature!
  -
- Generalized matrix-fractional function
  - Systematic study of (partial) infimal projections

$$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,B)}(X, V) + h(V).$$

    for $h \in \Gamma_0(\mathbb{S}^n)$. $\to$ SIOPT article to appear.
  - Numerical methods based on GMF.
  - Compute (analytically/numerically) the projection onto $\Omega(A, B)$ ($\to$ projection/proximal-based algorithms).

Current and future directions

- K-convexity
  - When is $\overline{\mathrm{conv}}\,(\mathrm{gph}\, F) = K\text{-epi}\, F$ for F K-convex ?
  - Subdifferential analysis for convex convex-composites, unification with the nonconvex convex-composite case (BCQ etc.)
  - Learn more about existing literature!
  -
- Generalized matrix-fractional function
  - Systematic study of (partial) infimal projections

    $$p(X) = \inf_{V \in \mathbb{S}^n} \sigma_{\Omega(A,B)}(X, V) + h(V).$$

    for $h \in \Gamma_0(\mathbb{S}^n)$. $\rightarrow$ SIOPT article to appear.
  - Numerical methods based on GMF.
  - Compute (analytically/numerically) the projection onto $\Omega(A, B)$ ($\rightarrow$ projection/proximal-based algorithms).

References

H. Attouch and H. Brézis: *Duality for the Sum of Convex Functions in General Banach Spaces.* In Aspects of Mathematics and its Applications, Eds. I.A. Barroso, North Holland, Amsterdam, 1986, pp.125–133.

H.H. Bauschke and P.L. Combettes: *Convex analysis and Monotone Operator Theory in Hilbert Spaces.* CMS Books in Mathematics, Springer-Verlag, New York, 2011.

J.V. Burke and T. Hoheisel: *Matrix support functionals for inverse problems, regularization, and learning.* SIAM Journal on Optimization 25(2), 2015, pp. 1135–1159.

J.V. Burke, T. Hoheisel, and Y. Gao: *Convex geometry of the generalized matrix-fractional function.* SIAM Journal on Optimization 28(3), 2018, pp. 2189–2200.

J.V. Burke, T. Hoheisel, and Y. Gao: *Variational properties of matrix functions via the generalized matrix-fractional function.* SIAM Journal on Optimization, to appear.

J.V. Burke, T. Hoheisel, and Q.V. Nguyen: *A note on the Fenchel conjugate of composite functions via infimal convolution.* Working paper, McGill University, 2019.

M. Fazel, H. Hindi, and S. Boyd: *A Rank Minimization Heuristic with Application to Minimum Order System Approximation.* Proceedings of the American Control Conference, Arlington, Virginia, June 2001.

L. Hörmander: *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe.* Arkiv för Matematik 21(3), 1955, pp. 181–186.

C.-J. Hsieh and P. Olsen: *Nuclear norm minimization via active subspace selection.* Proceedings of the 31st International Conference on Machine Learning (ICML-14), 2014, pp. 575-583.

A. Jalali, M. Fazel, and L. Xiao: *Variational Gram functions: Convex analysis and optimization.* SIAM Journal on Optimimization 27(4), 2017, pp. 2634–2661.

Teemu Pennanen: *Graph-Convex Mappings and K-Convex Functions.* Journal of Convex Analysis 6(2), 1999, pp. 235–266.

R.T. Rockafellar: *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

R.T. Rockafellar and R.J.-B. Wets: *Variational Analysis.* A Series of Comprehensive Studies in Mathematics, Vol. 317, Springer, Berlin, Heidelberg, 1998.

J.F. Toland: *A duality principle for non-convex optimisation and the calculus of variations.* Archive for Rational Mechanics and Analysis 71(1), 1979, pp. 41-61.