Perspectives on Duality in Convex Optimization

James V Burke
Mathematics, University of Washington

Joint work with
Aleksandr Aravkin (UW), Michael Friedlander (UBC),
Dmitriy Drusvyatskiy (UW) and Kellie MacPhee (UW)

McGill University
Discrete Mathematics and Optimization Seminar
February 20, 2017

Dual pairs of problems

A prototype problem: $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$

$$\mathcal{P} \qquad \begin{array}{rl} \min & \|x\|_1 \\ \text{s.t.} & \|b - Ax\|_2 \le \tau \end{array}$$

Dual pairs of problems

A prototype problem: $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$

$$\mathcal{P} \qquad \begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & \|b - Ax\|_2 \le \tau \end{array}$$

The (Fenchel-Rockafellar) dual problem:

$$\mathcal{D}_{\mathrm{L}} \qquad \begin{array}{ll} \sup & \langle b, z\rangle - \tau\|z\|_2 \\ \text{s.t.} & \left\|A^T z\right\|_\infty \le 1. \end{array}$$

Dual pairs of problems

A prototype problem: $A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m$

$$\mathcal{P} \qquad \begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & \|b - Ax\|_2 \le \tau \end{array}$$

The (Fenchel-Rockafellar) dual problem:

$$\mathcal{D}_{\mathrm{L}} \qquad \begin{array}{ll} \sup & \langle b, z\rangle - \tau\|z\|_2 \\ \text{s.t.} & \left\|A^T z\right\|_\infty \le 1. \end{array}$$

The gauge dual problem:

$$\mathcal{D}_{\mathrm{G}} \qquad \begin{array}{ll} \inf & \left\|A^T z\right\|_\infty \\ \text{s.t.} & \langle b, z\rangle - \tau\|z\|_2 \ge 1. \end{array}$$

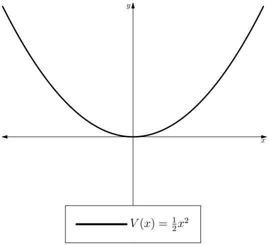
Piecewise Linear-Quadratic Penalties

$$\phi(x) := \sup_{u \in U} [\langle x, u \rangle - \frac{1}{2} u^T B u]$$

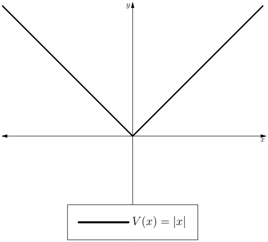
$U \subset \mathbb{R}^n$ is nonempty, closed and convex with $0 \in U$.
$B \in \mathbb{R}^{n \times n}$ is symmetric positive semi-definite.

Examples:
Norms, gauges, support functions, least-squares, Huber density

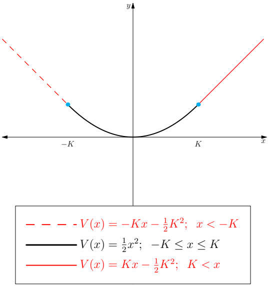
PLQ Densities: Gauss, Laplace, Huber, Vapnik



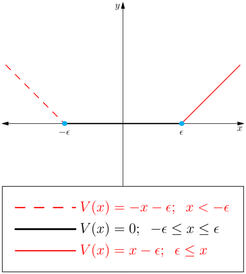
Gauss



ℓ_1



Huber



Vapnik

Convex Sets

Convex Sets

A subset C of $\mathbb{R}^n$ is convex if

$$[x, y] \subset C \quad \forall\, x, y \in C,$$

where

$$[x, y] := \{(1 - \lambda x) + \lambda y \,|\, 0 \leq \lambda \leq 1\}$$

is the line segment connecting x and y.

Convex Sets

<u>Convex Sets</u>

A subset C of $\mathbb{R}^n$ is convex if

$$[x, y] \subset C \quad \forall\, x, y \in C,$$

where

$$[x, y] := \{(1 - \lambda x) + \lambda y \mid 0 \leq \lambda \leq 1\}$$

is the line segment connecting x and y.

<u>Convex Cones</u>

A subset K of $\mathbb{R}^n$ is convex if

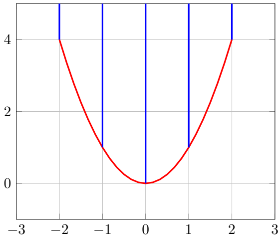
$$\lambda K \subset K \;\; \forall\, \lambda > 0 \;\; \text{and } K + K \subset K.$$

Convex functions and the epigraphical perspective

A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be convex if

$$\operatorname{epi} f := \{(x, \mu) \mid f(x) \leq \mu\},$$

is convex.

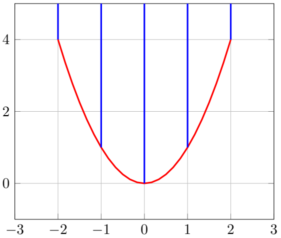


Convex functions and the epigraphical perspective

A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be convex if

$$\operatorname{epi} f := \{(x, \mu) \mid f(x) \leq \mu\},$$

is convex.



f is lower semi-continuous (lsc) $\iff$ $\operatorname{epi}(f)$ is closed

Linear transformations and their pre-images

If $A, B^T \in \mathbb{R}^{m \times n}$, then both

$$\begin{aligned} AC &:= \{Ax \mid x \in C\} \subset \mathbb{R}^m \text{ and} \\ B^{-1}C &:= \{y \mid By \in C\} \subset \mathbb{R}^m \end{aligned}$$

are convex.

Linear transformations and their pre-images

If $A, B^T \in \mathbb{R}^{m\times n}$, then both

$$\begin{aligned} AC &:= \{Ax \mid x \in C\} \subset \mathbb{R}^m \text{ and} \\ B^{-1}C &:= \{y \mid By \in C\} \subset \mathbb{R}^m \end{aligned}$$

are convex.

For example, the projection of a convex set onto an affine set is convex, where the affine sets are translates of subspaces.

Linear transformations and their pre-images

If $A, B^T \in \mathbb{R}^{m\times n}$, then both

$$AC := \{Ax \mid x \in C\} \subset \mathbb{R}^m \text{ and}$$
$$B^{-1}C := \{y \mid By \in C\} \subset \mathbb{R}^m$$

are convex.

For example, the projection of a convex set onto an affine set is convex, where the affine sets are translates of subspaces.

Caution:
The linear image of a closed convex set may not be closed.

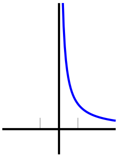


Figure : The graph of $1/x$

Linear transformations and their pre-images

If $A, B^T \in \mathbb{R}^{m\times n}$, then both

$$AC := \{Ax \mid x \in C\} \subset \mathbb{R}^m \text{ and}$$
$$B^{-1}C := \{y \mid By \in C\} \subset \mathbb{R}^m$$

are convex.

For example, the projection of a convex set onto an affine set is convex, where the affine sets are translates of subspaces.

Caution:
The linear image of a closed convex set may not be closed.

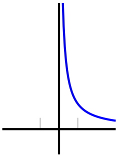


Figure : The graph of $1/x$

Convex hull: The convex hull of $S \subset \mathbb{R}^n$ is the intersection of all convex sets in $\mathbb{R}^n$ containing S, denoted $\operatorname{conv}(S)$.

Affine sets and relative interior

Affine sets: Any set of the form $x^0 + S$ where $x^0 \in \mathbb{R}^n$ and $S \subset \mathbb{R}^n$ is a subspace.

Affine sets and relative interior

Affine sets: Any set of the form $x^0 + S$ where $x^0 \in \mathbb{R}^n$ and $S \subset \mathbb{R}^n$ is a subspace.

Affine hull: The affine hull of a set $S \subset \mathbb{R}^n$ is the intersection of all affine sets that contain S, denoted $\operatorname{aff}(S)$.

Affine sets and relative interior

Affine sets: Any set of the form $x^0 + S$ where $x^0 \in \mathbb{R}^n$ and $S \subset \mathbb{R}^n$ is a subspace.

Affine hull: The affine hull of a set $S \subset \mathbb{R}^n$ is the intersection of all affine sets that contain S, denoted aff (S).

Relative interior: The relative interior of a convex set is the interior relative to its affine hull:

$$\operatorname{ri} C := \{x \in C \mid \exists \epsilon > 0 \text{ s.t. } (x + \epsilon \mathbb{B}) \cap \operatorname{aff}(C) \subset C\}$$

Affine sets and relative interior

Affine sets: Any set of the form $x^0 + S$ where $x^0 \in \mathbb{R}^n$ and $S \subset \mathbb{R}^n$ is a subspace.

Affine hull: The affine hull of a set $S \subset \mathbb{R}^n$ is the intersection of all affine sets that contain S, denoted $\operatorname{aff}(S)$.

Relative interior: The relative interior of a convex set is the interior relative to its affine hull:

$$\operatorname{ri} C := \{x \in C \mid \exists \epsilon > 0 \text{ s.t. } (x + \epsilon \mathbb{B}) \cap \operatorname{aff}(C) \subset C\}$$

Properties: Let $C \subset \mathbb{R}^n$ be convex and $A, B^T \in \mathbb{R}^{m \times n}$, then

$$\begin{aligned} A \operatorname{ri}(C) &= \operatorname{ri}(AC) \quad \text{and} \\ B^{-1} \operatorname{ri}(C) &= \operatorname{ri}(B^{-1}C), \quad \text{whenever } B^{-1}\operatorname{ri}(C) \neq \emptyset. \end{aligned}$$

The Hahn-Banach Theorem

Hyperplanes: Affine sets of co-dimension 1, or equivalently, any set of the form

$$\{x \mid \langle z, x\rangle = \beta\}$$

for some $\beta \in \mathbb{R}$ and non-zero $z \in \mathcal{L}$.

The Hahn-Banach Theorem: Let M be a nonempty affine set such that

$$M \cap \operatorname{ri} C = \emptyset.$$

Then there is a hyperplane H such that

$$M \subset H \quad \text{and} \quad H \cap \operatorname{ri} C = \emptyset.$$

Supporting hyperplanes

If $\overline{x} \in \operatorname{rbdry}(C) := \operatorname{cl} C \setminus \operatorname{ri} C$, then there is a hyperplane H containing $\overline{x}$ that does not meet the relative interior of C, or equivalently,

$$\exists\, z \text{ s.t. } \langle z, x\rangle < \langle z, \overline{x}\rangle \quad \forall\, x \in \operatorname{ri} C. \tag{1}$$

Supporting hyperplanes

If $\overline{x} \in \operatorname{rbdry}(C) := \operatorname{cl} C \setminus \operatorname{ri} C$, then there is a hyperplane H containing $\overline{x}$ that does not meet the relative interior of C, or equivalently,

$$\exists\, z \text{ s.t. } \langle z, x\rangle < \langle z, \overline{x}\rangle \quad \forall\, x \in \operatorname{ri} C. \tag{1}$$

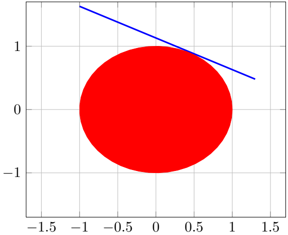
In this case, H is said to be a supporting hyperplane to C at $\overline{x}$.

Supporting hyperplanes

If $\overline{x} \in \operatorname{rbdry}(C) := \operatorname{cl} C \setminus \operatorname{ri} C$, then there is a hyperplane H containing $\overline{x}$ that does not meet the relative interior of C, or equivalently,

$$\exists\, z \text{ s.t. } \langle z, x\rangle < \langle z, \overline{x}\rangle \quad \forall\, x \in \operatorname{ri} C. \tag{1}$$

In this case, H is said to be a supporting hyperplane to C at $\overline{x}$.

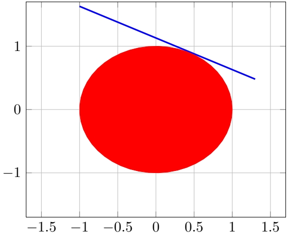


Supporting hyperplanes

If $\overline{x} \in \operatorname{rbdry}(C) := \operatorname{cl} C \setminus \operatorname{ri} C$, then there is a hyperplane H containing $\overline{x}$ that does not meet the relative interior of C, or equivalently,

$$\exists\, z \text{ s.t. } \langle z, x \rangle < \langle z, \overline{x} \rangle \quad \forall\, x \in \operatorname{ri} C. \tag{1}$$

In this case, H is said to be a supporting hyperplane to C at $\overline{x}$.



Question: Given $z \in \mathbb{R}^n$, does it define a supporting hyperplane to C and what are the associated support points.

Support functions

The support function for a set $S \subset \mathbb{R}^n$ is given by

$$\sigma_S(z) := \sup_{x \in S} \langle z, x \rangle.$$

It is straightforward to show that

$$\sigma_S(z) = \sigma_C(z), \quad \text{where} \quad C := \overline{\mathrm{conv}}(S).$$

Support functions

The support function for a set $S \subset \mathbb{R}^n$ is given by

$$\sigma_S(z) := \sup_{x \in S} \langle z, x \rangle.$$

It is straightforward to show that

$$\sigma_S(z) = \sigma_C(z), \quad \text{where} \quad C := \overline{\operatorname{conv}}(S).$$

Fact: Support functions are sublinear:

Support functions

The support function for a set $S \subset \mathbb{R}^n$ is given by

$$\sigma_S(z) := \sup_{x \in S} \langle z, x \rangle.$$

It is straightforward to show that

$$\sigma_S(z) = \sigma_C(z), \quad \text{where} \quad C := \overline{\text{conv}}(S).$$

Fact: Support functions are sublinear:

1. (positively homogeneous) $\sigma(\lambda x) = \lambda \sigma(x) \; \forall \, \lambda \geq 0$,

Support functions

The support function for a set $S \subset \mathbb{R}^n$ is given by

$$\sigma_S(z) := \sup_{x \in S} \langle z, x \rangle.$$

It is straightforward to show that

$$\sigma_S(z) = \sigma_C(z), \quad \text{where} \quad C := \overline{\mathrm{conv}}(S).$$

Fact: Support functions are sublinear:

1. (positively homogeneous) $\sigma(\lambda x) = \lambda \sigma(x) \; \forall \, \lambda \geq 0$,
2. (subadditive) $\sigma(x + y) \leq \sigma(x) + \sigma(y)$.

Support functions

The support function for a set $S \subset \mathbb{R}^n$ is given by

$$\sigma_S(z) := \sup_{x \in S} \langle z, x \rangle .$$

It is straightforward to show that

$$\sigma_S(z) = \sigma_C(z), \quad \text{where} \quad C := \overline{\text{conv}}(S).$$

Fact: Support functions are sublinear:

1. (positively homogeneous) $\sigma(\lambda x) = \lambda \sigma(x) \; \forall \, \lambda \geq 0$,
2. (subadditive) $\sigma(x+y) \leq \sigma(x) + \sigma(y)$.

Hörmander's Theorem: $\sigma : \mathbb{R}^n \to \bar{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\}$ lsc.

σ is subadditive $\iff$ epi(σ) is a closed cvx cone $\iff$ $\sigma = \sigma_C$,

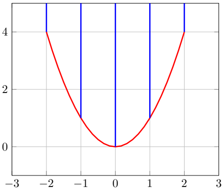
where $C := \{ z \mid \langle z, x \rangle \leq f(x) \; \forall \, x \} = \{ z \mid \langle z, x \rangle \leq 1 \; \forall \, f(x) \leq 1 \}$.

Convex functions and the epigraphical perspective

A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be convex if

$$\operatorname{epi} f := \{(x, \mu) \mid f(x) \le \mu\},$$

is convex.



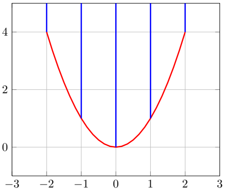
$\operatorname{epi}(f)$

Convex functions and the epigraphical perspective

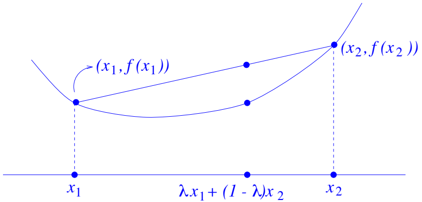
A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be convex if

$$\text{epi}\, f := \{(x, \mu) \mid f(x) \leq \mu\},$$

is convex.



epi (f)



$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda) f(x_1) + \lambda f(x_2)$$

$$\forall\ x_1, x_2 \in \text{dom}\, f \text{ and } \lambda \in [0, 1]$$

$$\text{dom}\, f := \{x \mid f(x) < \infty\}$$

Coordinate inf-projection of a convex set

Let $C \subset \mathbb{R}^{m+1}$ be a convex set such that the projection of C onto its last coordinate is bounded below. Define $f : \mathbb{R}^m \to \bar{\mathbb{R}}\}$ by

$$f(x) := \inf \{\overline{x}_{m+1} \mid \exists\, \overline{x} \in C \text{ s.t. } \overline{x} = (x, \overline{x}_{m+1})\},$$

where, again, the infimum over the empty set is $+\infty$.

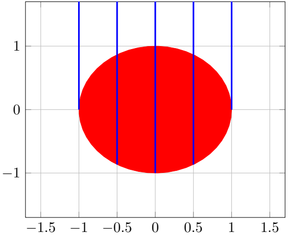
Coordinate inf-projection of a convex set

Let $C \subset \mathbb{R}^{m+1}$ be a convex set such that the projection of C onto its last coordinate is bounded below. Define $f : \mathbb{R}^m \to \bar{\mathbb{R}}\}$ by

$$f(x) := \inf \{\overline{x}_{m+1} \mid \exists\, \overline{x} \in C \text{ s.t. } \overline{x} = (x, \overline{x}_{m+1})\},$$

where, again, the infimum over the empty set is $+\infty$.

$$\mathrm{epi}\,(f) = C + (\{0\}^m \times \mathbb{R}_+)$$



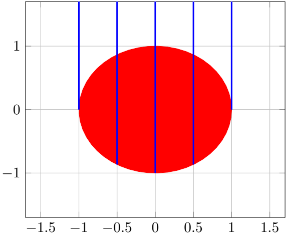
Coordinate inf-projection of a convex set

Let $C \subset \mathbb{R}^{m+1}$ be a convex set such that the projection of C onto its last coordinate is bounded below. Define $f : \mathbb{R}^m \to \bar{\mathbb{R}}\}$ by

$$f(x) := \inf\{\overline{x}_{m+1} \mid \exists\ \overline{x} \in C \text{ s.t. } \overline{x} = (x, \overline{x}_{m+1})\},$$

where, again, the infimum over the empty set is $+\infty$.

$$\text{epi}\,(f) = C + (\{0\}^m \times \mathbb{R}_+)$$



Example: $f(x) := \inf\limits_{(x,\mu)\in \text{epi}\,(f)} \mu$

Inf-projection: $h(x) := \inf_y f(y, x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y, x, \mu) = (x, \mu).$$

Inf-projection: $h(x) := \inf_y f(y, x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y, x, \mu) = (x, \mu).$$

Since the set $P\mathrm{epi}\,(f)$ is convex, the function $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by $\mathrm{epi}\,(h) := P\mathrm{epi}\,(f)$ is also convex:

Inf-projection: $h(x) := \inf_y f(y,x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y, x, \mu) = (x, \mu).$$

Since the set $P\operatorname{epi}(f)$ is convex, the function $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by $\operatorname{epi}(h) := P\operatorname{epi}(f)$ is also convex:

$$h(x) := \inf_{(x,\mu) \in P\operatorname{epi}(f)} \mu$$

Inf-projection: $h(x) := \inf_y f(y,x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y, x, \mu) = (x, \mu).$$

Since the set $P\text{epi}\,(f)$ is convex, the function $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by $\text{epi}\,(h) := P\text{epi}\,(f)$ is also convex:

$$\begin{aligned} h(x) &:= \inf_{(x,\mu) \in P\text{epi}\,(f)} \mu \\ &= \inf \{\mu \mid \exists\, y \in \mathbb{R}^m \text{ s.t. } \ f(y,x) \le \mu\} \end{aligned}$$

Inf-projection: $h(x) := \inf_y f(y,x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y,x,\mu) = (x,\mu).$$

Since the set $P\mathrm{epi}\,(f)$ is convex, the function $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by $\mathrm{epi}\,(h) := P\mathrm{epi}\,(f)$ is also convex:

$$\begin{aligned} h(x) &:= \inf_{(x,\mu)\in P\mathrm{epi}\,(f)} \mu \\ &= \inf\,\{\mu \mid \exists\, y \in \mathbb{R}^m \text{ s.t. } \; f(y,x) \le \mu\} \\ &= \inf\left\{\mu \,\middle|\, \inf_y f(y,x) < \mu\right\} \end{aligned}$$

Inf-projection: $h(x) := \inf_y f(y,x)$

Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex and consider the projection

$$P(y,x,\mu) = (x,\mu).$$

Since the set $P\operatorname{epi}(f)$ is convex, the function $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ given by $\operatorname{epi}(h) := P\operatorname{epi}(f)$ is also convex:

$$\begin{aligned} h(x) &:= \inf_{(x,\mu)\in P\operatorname{epi}(f)} \mu \\ &= \inf\{\mu \mid \exists\, y \in \mathbb{R}^m \text{ s.t. } f(y,x) \le \mu\} \\ &= \inf\left\{\mu \,\middle|\, \inf_y f(y,x) < \mu\right\} \\ &= \inf_y f(y,x). \end{aligned}$$

The support function for epi (f): the convex conjugate

$$\begin{aligned}
\sigma_{\operatorname{epi} f}((z,-1)) &= \sup_{f(x)\le\mu} \langle (z,-1),(x,\mu)\rangle \\
&= \sup_{f(x)\le\mu} [\langle z,x\rangle - \mu] \\
&= \sup_{x}[\langle z,x\rangle - f(x)] \\
&=: f^*(z)
\end{aligned}$$

The support function for epi (f): the convex conjugate

$$
\begin{aligned}
\sigma_{\mathrm{epi}\, f}\left((z,-1)\right) &= \sup_{f(x)\le\mu} \langle (z,-1),(x,\mu)\rangle \\
&= \sup_{f(x)\le\mu} \left[\langle z,x\rangle - \mu\right] \\
&= \sup_{x}[\langle z,x\rangle - f(x)] \\
&=: f^*(z)
\end{aligned}
$$

Subgradients:

$$
\overline{x} \in \operatorname*{argmax}_{x}[\langle z,x\rangle - f(x)] \iff (z,-1) \text{ supports epi}\,(f) \text{ at } (\overline{x}, f(\overline{x})).
$$

The support function for epi (f): the convex conjugate

$$\begin{aligned}\sigma_{\mathrm{epi}\, f}\left((z,-1)\right) &= \sup_{f(x)\le\mu} \langle (z,-1),(x,\mu)\rangle \\ &= \sup_{f(x)\le\mu} \left[\langle z,x\rangle - \mu\right] \\ &= \sup_{x}[\langle z,x\rangle - f(x)] \\ &=: f^*(z)\end{aligned}$$

Subgradients:

$$\overline{x} \in \operatorname*{argmax}_x[\langle z,x\rangle - f(x)] \quad \Longleftrightarrow \quad (z,-1) \text{ supports epi}\,(f) \text{ at } (\overline{x}, f(\overline{x})).$$

$$\Longleftrightarrow \quad \langle (z,-1),(\overline{x}, f(\overline{x})\rangle \ge \langle (z,-1),(x,f(x))\rangle\ \forall\, x\in \operatorname{dom} f,$$

The support function for epi (f): the convex conjugate

$$\begin{aligned}
\sigma_{\mathrm{epi}\,f}((z,-1)) &= \sup_{f(x)\leq\mu} \langle (z,-1),(x,\mu)\rangle \\
&= \sup_{f(x)\leq\mu} [\langle z,x\rangle - \mu] \\
&= \sup_{x}[\langle z,x\rangle - f(x)] \\
&=: f^*(z)
\end{aligned}$$

Subgradients:

$$\overline{x} \in \operatorname*{argmax}_{x}[\langle z,x\rangle - f(x)] \iff (z,-1) \text{ supports epi}\,(f) \text{ at } (\overline{x}, f(\overline{x})).$$

$$\iff \langle (z,-1),(\overline{x}, f(\overline{x})\rangle \geq \langle (z,-1),(x,f(x))\rangle \;\forall\, x \in \operatorname{dom} f,$$

$$\iff f(x) \geq f(\overline{x}) + \langle z, x-\overline{x}\rangle \quad \forall\, x \in \mathbb{R}^n$$

The support function for epi (f): the convex conjugate

$$\begin{aligned}
\sigma_{\mathrm{epi}\, f}\left((z,-1)\right) &= \sup_{f(x)\leq\mu} \langle (z,-1),(x,\mu)\rangle \\
&= \sup_{f(x)\leq\mu} \left[\langle z,x\rangle - \mu\right] \\
&= \sup_{x}\left[\langle z,x\rangle - f(x)\right] \\
&=: f^*(z)
\end{aligned}$$

Subgradients:

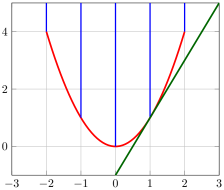
$$\overline{x} \in \operatorname*{argmax}_x\left[\langle z,x\rangle - f(x)\right] \iff (z,-1) \text{ supports epi}\,(f) \text{ at } (\overline{x}, f(\overline{x})).$$

$$\iff \langle (z,-1),(\overline{x}, f(\overline{x})\rangle \geq \langle (z,-1),(x,f(x))\rangle \;\forall\, x \in \operatorname{dom} f,$$

$$\iff f(x) \geq f(\overline{x}) + \langle z, x-\overline{x}\rangle \quad \forall\, x \in \mathbb{R}^n$$

$z \in \partial f(\overline{x})$, the subdifferential of f at $\overline{x}$.

The conjugate and subgradients



$$\partial f(\overline{x}) \text{ is a singleton} \iff \partial f(\overline{x}) = \{\nabla f(\overline{x})\}.$$

The conjugate and subgradients

$$f^*(z) \geq \langle z, x \rangle - f(x) \quad \forall \;\; x \in \operatorname{dom}(f) \text{ and } z \in \mathbb{R}^n$$

$$\iff \quad f(x) \geq \langle z, x \rangle - f^*(z) \qquad \forall \;\; z \in \operatorname{dom}(f^*) \text{ and } x \in \mathbb{R}^n$$

$$\implies \quad f(x) \geq f^{**}(x) \qquad \forall\, x \in \mathbb{R}^n$$

The conjugate and subgradients

$$f^*(z) \geq \langle z, x\rangle - f(x) \quad \forall \;\; x \in \mathrm{dom}\,(f) \text{ and } z \in \mathbb{R}^n$$

$$\iff \quad f(x) \geq \langle z, x\rangle - f^*(z) \qquad \forall \;\; z \in \mathrm{dom}\,(f^*) \text{ and } x \in \mathbb{R}^n$$

$$\implies \quad f(x) \geq f^{**}(x) \qquad \forall\, x \in \mathbb{R}^n$$

But

$$z \in \partial f(x) \quad \iff \quad \langle z, x\rangle \geq f(x) + f^*(z),$$

so $\forall\, x \in \mathrm{dom}\,(\partial f) := \{x \,|\, \partial f(x) \neq \emptyset\}$ and $z \in \partial f(x)$,

$$f(x) \leq \langle z, x\rangle - f^*(z) \leq \sup_w [\langle w, x\rangle - f^*(w)] = f^{**}(x) \leq f(x).$$

The conjugate and subgradients

$$f^*(z) \geq \langle z, x\rangle - f(x) \quad \forall \;\; x \in \operatorname{dom}(f) \text{ and } z \in \mathbb{R}^n$$

$$\iff \quad f(x) \geq \langle z, x\rangle - f^*(z) \qquad \forall \;\; z \in \operatorname{dom}(f^*) \text{ and } x \in \mathbb{R}^n$$

$$\implies \quad f(x) \geq f^{**}(x) \qquad \forall\, x \in \mathbb{R}^n$$

But

$$z \in \partial f(x) \quad \iff \quad \langle z, x\rangle \geq f(x) + f^*(z),$$

so $\forall\, x \in \operatorname{dom}(\partial f) := \{x \mid \partial f(x) \neq \emptyset\}$ and $z \in \partial f(x)$,

$$f(x) \leq \langle z, x\rangle - f^*(z) \leq \sup_w [\langle w, x\rangle - f^*(w)] = f^{**}(x) \leq f(x).$$

So $f(x) = f^{**}(x)$ on $\operatorname{dom}(\partial f)$, where $\operatorname{ri}\operatorname{dom}(f) \subset \operatorname{dom}(\partial f)$.
Consequently $f^{**} = \operatorname{cl} f$, so if $f = \operatorname{cl} f$, $\partial f^* = (\partial f)^{-1}$.

The convex indicator function

$C \subset \mathbb{R}^n$ non-empty closed convex

$$\delta_C\,(x) := \begin{cases} 0 & ,\ x \in C, \\ +\infty & ,\ x \notin C \end{cases}$$

$$\delta_C^*(z) = \sigma_C\,(z)$$

$$\begin{aligned} \partial\delta_C\,(x) &= \{z \mid \langle z, y - x\rangle \leq 0 \ \ \forall\, y \in C\} \quad (x \in C) \\ &=: N\,(x \,|\, C) \qquad \text{the normal cone to } C \text{ at } x \\ &= \text{set of supporting vectors to } C \text{ at } x \end{aligned}$$

The conjugate under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection:

$$p(y) := \inf_x F(x, y).$$

Then

$$\begin{aligned} p^*(z) &= \sup_y [\langle z, y\rangle - p(y)] \\ &= \sup_y [\langle z, y\rangle - \inf_x F(x, y)] \\ &= \sup_y \sup_x [\langle z, y\rangle - F(x, y)] \\ &= \sup_{(x,y)} [\langle (0, z), (x, y)\rangle - F(x, y)] \\ &= F^*(0, z) \end{aligned}$$

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x, y)$.

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x, y)$. Then

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x, y).$$

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x, y)$. Then

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x, y).$$

$$\Longleftrightarrow$$

$$F(\overline{x}, y) + F^*(0, z) = p(y) + p^*(z) \leq \langle z, y \rangle = \langle (0, z), (\overline{x}, y) \rangle$$

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x, y)$. Then

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x, y).$$

$$\Longleftrightarrow$$

$$F(\overline{x}, y) + F^*(0, z) = p(y) + p^*(z) \leq \langle z, y \rangle = \langle (0, z), (\overline{x}, y) \rangle$$

$$\Longrightarrow$$

$$(0, z) \in \partial F(\overline{x}, y) \quad \text{or equivalently} \quad (\overline{x}, y) \in \partial F^*(0, z)$$

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x,y)$. Then

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x,y).$$

$$\Longleftrightarrow$$

$$F(\overline{x},y) + F^*(0,z) = p(y) + p^*(z) \leq \langle z, y \rangle = \langle (0,z), (\overline{x}, y) \rangle$$

$$\Longrightarrow$$

$$(0,z) \in \partial F(\overline{x},y) \quad \text{or equivalently} \quad (\overline{x},y) \in \partial F^*(0,z)$$

$$\Longleftrightarrow$$

$$p(y) + p^*(z) \leq F(\overline{x},y) + F^*(0,z) \leq \langle (0,z), (\overline{x},y) \rangle = \langle z, y \rangle$$

The subdifferential under inf-projection

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \bar{\mathbb{R}}$ and define the following optimal value function by inf-projection: $p(y) := \inf_x F(x, y)$. Then

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x, y).$$

$$\Longleftrightarrow$$

$$F(\overline{x}, y) + F^*(0, z) = p(y) + p^*(z) \le \langle z, y \rangle = \langle (0, z), (\overline{x}, y) \rangle$$

$$\Longrightarrow$$

$$(0, z) \in \partial F(\overline{x}, y) \quad \text{or equivalently} \quad (\overline{x}, y) \in \partial F^*(0, z)$$

$$\Longleftrightarrow$$

$$p(y) + p^*(z) \le F(\overline{x}, y) + F^*(0, z) \le \langle (0, z), (\overline{x}, y) \rangle = \langle z, y \rangle$$

$$\Longleftrightarrow$$

$$z \in \partial p(y) \quad \text{and} \quad \overline{x} \in \operatorname*{argmin}_x F(x, y).$$

Duality Theory

Let $F\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and define the following optimal value functions by inf-projection:

$$p(y) := \inf_x \, F(x, y) \quad \text{and} \quad q(w) := \inf_z \, F^*(w, z).$$

Duality Theory

Let $F\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and define the following optimal value functions by inf-projection:

$$p(y) := \inf_x \; F(x,y) \quad \text{and} \quad q(w) := \inf_z \; F^*(w,z).$$

This set-up yields the primal-dual pair

$$p(0) = \inf_x \; F(x,0) \quad \text{and} \quad p^{**}(0) = \sup_z \; -F^*(0,z) \;\; = \;\; -q(0).$$

Duality Theory

Let $F\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and define the following optimal value functions by inf-projection:

$$p(y) := \inf_x \, F(x,y) \quad \text{and} \quad q(w) := \inf_z \, F^*(w,z).$$

This set-up yields the primal-dual pair

$$p(0) = \inf_x \, F(x,0) \quad \text{and} \quad p^{**}(0) = \sup_z \, -F^*(0,z) \; = \; -q(0).$$

$$p(0) \;\geq\; p^{**}(0) \;=\; -q(0) \text{ always holds}$$

Duality Theory

$$p(y) := \inf_x \; F(x,y) \quad \text{and} \quad q(w) := \inf_z \; F^*(w,z).$$

Duality Theory

$$p(y) := \inf_x F(x,y) \quad \text{and} \quad q(w) := \inf_z F^*(w,z).$$

1. If $0 \in \operatorname{ri}(\operatorname{dom} p)$, then $p(0) = -q(0)$ and the infimum $q(0)$ is attained, if finite.

   Similarly, if $0 \in \operatorname{ri}(\operatorname{dom} q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite.

Duality Theory

$$p(y) := \inf_x \; F(x,y) \quad \text{and} \quad q(w) := \inf_z \; F^*(w,z).$$

1. If $0 \in \operatorname{ri}(\operatorname{dom} p)$, then $p(0) = -q(0)$ and the infimum $q(0)$ is attained, if finite.

   Similarly, if $0 \in \operatorname{ri}(\operatorname{dom} q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite.

2. The set $\operatorname{argmax}_z -F^*(0,z)$ is nonempty and bounded if and only if $0 \in \operatorname{int}(\operatorname{dom} p)$ and $p(0)$ is finite, in which case $\partial p(0) = \operatorname{argmax}_z -F^*(0,z)$.

Duality Theory

$$p(y) := \inf_x \; F(x,y) \quad \text{and} \quad q(w) := \inf_z \; F^*(w,z).$$

1. If $0 \in \mathrm{ri}\,(\mathrm{dom}\, p)$, then $p(0) = -q(0)$ and the infimum $q(0)$ is attained, if finite.

   Similarly, if $0 \in \mathrm{ri}\,(\mathrm{dom}\, q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite.

2. The set $\mathrm{argmax}_z \, -F^*(0,z)$ is nonempty and bounded if and only if $0 \in \mathrm{int}\,(\mathrm{dom}\, p)$ and $p(0)$ is finite, in which case $\partial p(0) = \mathrm{argmax}_z \, -F^*(0,z)$.

3. $\mathrm{argmin}_x \, F(x,0)$ is nonempty and bounded if and only if $0 \in \mathrm{int}\,(\mathrm{dom}\, q)$ and $q(0)$ is finite, in which case $\partial q(0) = \mathrm{argmin}_x \, F(x,0)$.

Duality Theory

$$p(y) := \inf_x \; F(x,y) \quad \text{and} \quad q(w) := \inf_z \; F^*(w,z).$$

1. If $0 \in \mathrm{ri}\,(\mathrm{dom}\, p)$, then $p(0) = -q(0)$ and the infimum $q(0)$ is attained, if finite.

   Similarly, if $0 \in \mathrm{ri}\,(\mathrm{dom}\, q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite.

2. The set $\mathrm{argmax}_z \; -F^*(0,z)$ is nonempty and bounded if and only if $0 \in \mathrm{int}\,(\mathrm{dom}\, p)$ and $p(0)$ is finite, in which case $\partial p(0) = \mathrm{argmax}_z \; -F^*(0,z)$.

3. $\mathrm{argmin}_x \, F(x,0)$ is nonempty and bounded if and only if $0 \in \mathrm{int}\,(\mathrm{dom}\, q)$ and $q(0)$ is finite, in which case $\partial q(0) = \mathrm{argmin}_x \, F(x,0)$.

4. Optimal solutions are characterized by

$$\left.\begin{array}{l} \bar{x} \in \mathrm{argmin}_x \;\; F(x,0) \\ \bar{y} \in \mathrm{argmax}_z \;\; -F^*(0,z) \\ F(\bar{x},0) = -F^*(0,\bar{z}) \end{array}\right\} \iff (0,\bar{z}) \in \partial F(\bar{x},0) \iff (\bar{x},0) \in \partial F^*(0,\bar{z}).$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z,y\rangle] \;=\; -\sup_y [\langle z,y\rangle - F(x,y)]$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v)$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$L(x,(z,w)) = \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned} L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\ &= -\{\ \sup_u [\langle z,u\rangle - h(Ax+u)] \;+\; \sup_v [\langle w,v\rangle - g(x+v)]\ \} \end{aligned}$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned} L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\ &= -\{ \sup_u [\langle z, u\rangle - h(Ax+u)] \;+\; \sup_v [\langle w, v\rangle - g(x+v)] \} \end{aligned}$$

$(r := Ax + u)$ $\qquad$ $(s := x + v)$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z,y\rangle] \;=\; -\sup_y [\langle z,y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned} L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\ &= -\{ \sup_u [\langle z,u\rangle - h(Ax+u)] \quad + \quad \sup_v [\langle w,v\rangle - g(x+v)] \;\} \\ &\qquad\qquad (r := Ax+u) \qquad\qquad\qquad\qquad (s := x+v) \\ &= -\{ \sup_r [\langle z, r - Ax\rangle - h(r)] \quad + \quad \sup_s [\langle w, s-x\rangle - g(s)] \;\} \end{aligned}$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned}
L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\
&= -\{ \sup_u [\langle z,u\rangle - h(Ax+u)] \;+\; \sup_v [\langle w,v\rangle - g(x+v)] \} \\
&\qquad\qquad (r := Ax+u) \qquad\qquad\qquad (s := x+v) \\
&= -\{ \sup_r [\langle z, r - Ax\rangle - h(r)] \;+\; \sup_s [\langle w, s-x\rangle - g(s)] \} \\
&= [\langle z, Ax\rangle - h^*(z)] \;+\; [\langle w, x\rangle - g^*(w)]
\end{aligned}$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z,y\rangle] \;=\; -\sup_y [\langle z,y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned}
L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\
&= -\{\sup_u [\langle z,u\rangle - h(Ax+u)] \quad + \quad \sup_v [\langle w,v\rangle - g(x+v)]\ \} \\
&\qquad\qquad (r := Ax+u) \qquad\qquad\qquad (s := x+v) \\
&= -\{\sup_r [\langle z, r - Ax\rangle - h(r)] \quad + \quad \sup_s [\langle w, s-x\rangle - g(s)]\ \} \\
&= [\langle z, Ax\rangle - h^*(z)] \; + \; [\langle w,x\rangle - g^*(w)] \\
&= \langle A^T z + w, x\rangle - [h^*(z) + g^*(w)]
\end{aligned}$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned}
L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\
&= -\{\, \sup_u [\langle z, u\rangle - h(Ax+u)] \;+\; \sup_v [\langle w, v\rangle - g(x+v)] \,\} \\
&\qquad\qquad (r := Ax+u) \qquad\qquad (s := x+v) \\
&= -\{\, \sup_r [\langle z, r - Ax\rangle - h(r)] \;+\; \sup_s [\langle w, s - x\rangle - g(s)] \,\} \\
&= [\langle z, Ax\rangle - h^*(z)] \;+\; [\langle w, x\rangle - g^*(w)] \\
&= \langle A^T z + w, x\rangle - [h^*(z) + g^*(w)]
\end{aligned}$$

$$p(0) = \inf_x \sup_{(z,w)} L(x,z,w) = \inf_x [h(Ax) + g(x)]$$

Lagrangian Duality

The Lagrangian function:

$$L(x,z) := \inf_y [F(x,y) - \langle z, y\rangle] \;=\; -\sup_y [\langle z, y\rangle - F(x,y)]$$

Example: $F(x,(u,v)) := h(Ax+u) + g(x+v) + \frac{1}{2}\|u\|_2^2 + \|v\|_2^2$

$$\begin{aligned}
L(x,(z,w)) &= \inf_{(u,v)} h(Ax+u) + g(x+v) - \langle (z,w),(u,v)\rangle \\
&= -\{\, \sup_u [\langle z,u\rangle - h(Ax+u)] \;+\; \sup_v [\langle w,v\rangle - g(x+v)] \,\} \\
&\qquad\qquad (r := Ax+u) \qquad\qquad\qquad (s := x+v) \\
&= -\{\, \sup_r [\langle z, r-Ax\rangle - h(r)] \;+\; \sup_s [\langle w, s-x\rangle - g(s)] \,\} \\
&= [\langle z, Ax\rangle - h^*(z)] \;+\; [\langle w, x\rangle - g^*(w)] \\
&= \langle A^T z + w, x\rangle - [h^*(z) + g^*(w)]
\end{aligned}$$

$$p(0) = \inf_x \sup_{(z,w)} L(x,z,w) = \inf_x [h(Ax) + g(x)]$$

$$p^{**}(0) = \sup_{(z,w)} \inf_x L(x,z,w) = \sup_z -[h^*(z) + g^*(-A^T z)]$$

Example

Fenchel-Rockafellar Duality: $F(x, y) = h(Ax + y) + g(x)$

$$p(0) = \inf_x \{ h(Ax) + g(x) \} \quad \text{and} \quad p^{\star\star}(0) = \sup_z \{ -h^*(z) - g^*(-A^*z) \}$$

Example

Fenchel-Rockafellar Duality: $F(x,y) = h(Ax+y) + g(x)$

$$p(0) = \inf_x \{\, h(Ax) + g(x) \,\} \quad \text{and} \quad p^{\star\star}(0) = \sup_z \{\, -h^*(z) - g^*(-A^*z) \,\}$$

A prototype problem:

$$\mathcal{P} \qquad \begin{array}{rl} \min & \|x\|_1 \\ \text{s.t.} & \|Ax-b\|_2 \leq \tau \end{array}$$

Example

Fenchel-Rockafellar Duality: $F(x,y) = h(Ax+y) + g(x)$

$$p(0) = \inf_x \{\, h(Ax) + g(x) \,\} \quad \text{and} \quad p^{\star\star}(0) = \sup_z \{\, -h^*(z) - g^*(-A^*z) \,\}$$

A prototype problem:

$$\mathcal{P} \qquad \begin{array}{l} \min \ \|x\|_1 \\ \text{s.t. } \|Ax - b\|_2 \le \tau \end{array}$$

$$g(x) = \|x\|_1 = \delta^*\left(x \mid \mathbb{B}_\infty\right) \qquad g^*(w) = \delta\left(w \mid \mathbb{B}_\infty\right)$$

$$h(y) = \delta\left(y - b \mid \tau\mathbb{B}_2\right) \qquad h^*(z) = -\langle z, b\rangle + \delta^*\left(z \mid \tau\mathbb{B}_2\right) = -\langle z, b\rangle + \tau\|z\|_2$$

Example

Fenchel-Rockafellar Duality: $F(x,y) = h(Ax+y) + g(x)$

$$p(0) = \inf_x \{\, h(Ax) + g(x) \,\} \quad \text{and} \quad p^{\star\star}(0) = \sup_z \{\, -h^*(z) - g^*(-A^*z) \,\}$$

A prototype problem:

$$\mathcal{P} \qquad \begin{array}{l} \min \ \|x\|_1 \\ \text{s.t. } \|Ax - b\|_2 \le \tau \end{array}$$

$$g(x) = \|x\|_1 = \delta^*\left(x \mid \mathbb{B}_\infty\right) \qquad g^*(w) = \delta\left(w \mid \mathbb{B}_\infty\right)$$

$$h(y) = \delta\left(y - b \mid \tau\mathbb{B}_2\right) \qquad h^*(z) = -\langle z, b\rangle + \delta^*\left(z \mid \tau\mathbb{B}_2\right) = -\langle z, b\rangle + \tau\|z\|_2$$

$$\mathcal{D}_{\mathrm{L}} \qquad \begin{array}{l} \sup \ \langle b, z\rangle - \tau\|z\|_2 \\ \text{s.t. } \left\|A^T z\right\|_\infty \le 1. \end{array}$$

Gauge Duality

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \leq 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_{\circ} := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \leq 1\}.$$

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \leq 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_\circ := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \leq 1\}.$$

Hence,

$$\langle z, x\rangle \leq \|z\|_\circ \|x\| \quad \forall\, z, x \in \mathbb{R}^n.$$

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \le 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_\circ := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \le 1\}.$$

Hence,

$$\langle z, x\rangle \le \|z\|_\circ \|x\| \quad \forall\, z, x \in \mathbb{R}^n.$$

For $S \subset \mathbb{R}^n$, the *polar* of S is

$$S^\circ := \{z \mid \langle z, x\rangle \le 1 \ \forall\, x \in S\}.$$

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \le 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_\circ := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \le 1\}.$$

Hence,

$$\langle z, x\rangle \le \|z\|_\circ \|x\| \quad \forall\, z, x \in \mathbb{R}^n.$$

For $S \subset \mathbb{R}^n$, the *polar* of S is

$$S^\circ := \{z \mid \langle z, x\rangle \le 1 \;\forall\, x \in S\}.$$

Since $\mathbb{B}^\circ = \{z \mid \langle z, x\rangle \le 1 \;\forall\, x \in \mathbb{B}\} = \{z \mid \|z\|_\circ \le 1\}$, we have

$$\|x\| = \sigma_{\mathbb{B}^\circ}(x).$$

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \leq 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_\circ := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \leq 1\}.$$

Hence,

$$\langle z, x\rangle \leq \|z\|_\circ \|x\| \quad \forall\, z, x \in \mathbb{R}^n.$$

For $S \subset \mathbb{R}^n$, the *polar* of S is

$$S^\circ := \{z \mid \langle z, x\rangle \leq 1 \;\forall\, x \in S\}.$$

Since $\mathbb{B}^\circ = \{z \mid \langle z, x\rangle \leq 1 \;\forall\, x \in \mathbb{B}\} = \{z \mid \|z\|_\circ \leq 1\}$, we have

$$\|x\| = \sigma_{\mathbb{B}^\circ}(x).$$

Properties:

1. $(S^\circ)^\circ = \overline{\mathrm{conv}}\,(S \cup \{0\})$

Dual Norms and Polars

$\mathbb{B} := \{x \mid \|x\| \le 1\}$ is the closed unit ball of norm $\|\cdot\|$.

The norm dual to $\|\cdot\|$ is defined to be

$$\|z\|_\circ := \sigma_{\mathbb{B}}(z) = \sup\{\langle z, x\rangle \mid \|x\| \le 1\}.$$

Hence,

$$\langle z, x\rangle \le \|z\|_\circ \|x\| \quad \forall\, z, x \in \mathbb{R}^n.$$

For $S \subset \mathbb{R}^n$, the *polar* of S is

$$S^\circ := \{z \mid \langle z, x\rangle \le 1 \;\forall\, x \in S\}.$$

Since $\mathbb{B}^\circ = \{z \mid \langle z, x\rangle \le 1 \;\forall\, x \in \mathbb{B}\} = \{z \mid \|z\|_\circ \le 1\}$, we have

$$\|x\| = \sigma_{\mathbb{B}^\circ}(x)\,.$$

Properties:

1. $(S^\circ)^\circ = \overline{\mathrm{conv}}\,(S \cup \{0\})$
2. K is a close convex cone ($\lambda K \subset K \;\forall\, \lambda > 0,\; K + K \subset K$), then

$$K^\circ = \{z \mid \langle z, x\rangle \le 0 \;\forall\, x \in K\}.$$

Minkowski (gauge) functionals and polarity

$0 \in C \subset \mathbb{R}^n$ nonempty closed convex

$$\gamma_C(x) := \inf\{t \mid 0 \le t,\ x \in tC\},$$

where the infimum over the empty set is $+\infty$.

Minkowski (gauge) functionals and polarity

$0 \in C \subset \mathbb{R}^n$ nonempty closed convex

$$\gamma_C(x) := \inf\{t \mid 0 \le t,\ x \in tC\},$$

where the infimum over the empty set is $+\infty$.

Example: $\|x\| = \gamma_{\mathbb{B}}(x)$ for any norm with unit ball $\mathbb{B}$.

Minkowski (gauge) functionals and polarity

$0 \in C \subset \mathbb{R}^n$ nonempty closed convex

$$\gamma_C(x) := \inf\{t \mid 0 \leq t,\ x \in tC\},$$

where the infimum over the empty set is $+\infty$.

Example: $\|x\| = \gamma_{\mathbb{B}}(x)$ for any norm with unit ball $\mathbb{B}$.

Gauge functions are sublinear, and so by Hörmander,

$$\gamma_C(x) = \sigma_D(x),$$

where

$$D = \{z \mid \langle z, x\rangle \leq 1\ \forall\ x \in C\} = C^\circ.$$

Polar Gauges

$$\kappa^\circ(y) = \sup\{\, \langle y, x\rangle \mid \kappa(x) \le 1 \,\} = \sigma_{\mathbb{U}_\kappa}(y)\,,$$

where $\mathbb{U}_\kappa := \{x \mid \kappa(x) \le 1\}$.

Polar Gauges

$$\kappa^\circ(y) = \sup\left\{\, \langle y, x\rangle \mid \kappa(x) \le 1 \,\right\} = \sigma_{\mathbb{U}_\kappa}(y)\,,$$

where $\mathbb{U}_\kappa := \{x \mid \kappa(x) \le 1\,\}$.

If κ is a norm then κ° is the corresponding dual norm.

Polar Gauges

$$\kappa^\circ(y) = \sup\{\, \langle y, x\rangle \mid \kappa(x) \le 1 \,\} = \sigma_{\mathbb{U}_\kappa}(y)\,,$$

where $\mathbb{U}_\kappa := \{x \mid \kappa(x) \le 1\,\}$.

If κ is a norm then κ° is the corresponding dual norm.

$$\operatorname{epi}\kappa^\circ = \{(y, -\lambda) : (y, \lambda) \in (\operatorname{epi}\kappa)^\circ\}.$$

Polar Gauges

$$\kappa^\circ(y) = \sup \{ \langle y, x \rangle \mid \kappa(x) \le 1 \} = \sigma_{\mathbb{U}_\kappa}(y),$$

where $\mathbb{U}_\kappa := \{x \mid \kappa(x) \le 1\}$.

If κ is a norm then κ° is the corresponding dual norm.

$$\operatorname{epi} \kappa^\circ = \{(y, -\lambda) : (y, \lambda) \in (\operatorname{epi} \kappa)^\circ\}.$$

The generalized Hölder inequality

$$\langle x, y \rangle \le \kappa(x) \cdot \kappa^\circ(y) \quad \forall x \in \operatorname{dom} \kappa, \ \forall y \in \operatorname{dom} \kappa^\circ,$$

is known as the *polar-gauge inequality.*

Gauge Duality

κ and ρ are gauges.

$$\min_x \quad \kappa(x) \qquad \text{s.t.} \quad \rho(b - Ax) \leq \tau, \tag{G$_p$}$$

$$\max_y \quad \langle b, y \rangle - \tau\rho^\circ(y) \qquad \text{s.t.} \quad \kappa^\circ(A^T y) \leq 1, \tag{L$_d$}$$

$$\min_y \quad \kappa^\circ(A^T y) \qquad \text{s.t.} \quad \langle b, y \rangle - \tau\rho^\circ(y) \geq 1. \tag{G$_d$}$$

Example: In $\mathcal{P}$ we set $\kappa(x) = \|x\|_1$ and $\rho(y) = \|y\|_2$

Gauge Duality

κ and ρ are gauges.

$$v_p := \min_x \quad \kappa(x) \qquad \text{s.t.} \quad \rho(b - Ax) \le \tau, \tag{G_p}$$

$$\max_y \quad \langle b, y\rangle - \tau\rho^\circ(y) \qquad \text{s.t.} \quad \kappa^\circ(A^T y) \le 1, \tag{L_d}$$

$$v_d := \min_y \quad \kappa^\circ(A^T y) \qquad \text{s.t.} \quad \langle b, y\rangle - \tau\rho^\circ(y) \ge 1. \tag{G_d}$$

Example: In $\mathcal{P}$ we set $\kappa(x) = \|x\|_1$ and $\rho(y) = \|y\|_2$

Feasibility

Primal, Dual Domains:

$$\mathcal{F}_p := \{\, x \mid \rho(b - Ax) \le \tau \,\} \qquad \text{and} \qquad \mathcal{F}_d := \{\, y \mid \langle b, y\rangle - \tau\rho^\circ(y) \ge 1 \,\}\,.$$

$$\textbf{Feasibilty}: \quad \begin{array}{ll} \text{Primal} & \mathcal{F}_p \cap (\operatorname{dom}\kappa) \\ \text{Dual} & A^T\,\mathcal{F}_d \cap (\operatorname{dom}\kappa^\circ) \end{array}$$

$$\textbf{Relative Strict Feasibilty}: \quad \begin{array}{ll} \text{Primal} & \operatorname{ri}\mathcal{F}_p \cap (\operatorname{ri}\operatorname{dom}\kappa) \\ \text{Dual} & A^T\operatorname{ri}\mathcal{F}_d \cap (\operatorname{ri}\operatorname{dom}\kappa^\circ) \end{array}$$

$$\textbf{Strict Feasibilty}: \quad \begin{array}{ll} \text{Primal} & \operatorname{int}(\mathcal{F})_p \cap (\operatorname{ri}\operatorname{dom}\kappa) \\ \text{Dual} & A^T\operatorname{int}(\mathcal{F})_d \cap (\operatorname{ri}\operatorname{dom}\kappa^\circ) \end{array}$$

Gauge Duality

$$v_p \;=\; \min_{\rho(b-AX)\leq\tau} \kappa(x) \qquad\qquad v_d = \min_{\langle b,y\rangle - \tau\rho^\circ(y)\geq 1} \kappa^\circ(A^T y)$$

Theorem: (2014)

1. (Weak duality)
   If x and y are P-D feasible, then
   $$1 \leq v_p v_d \leq \kappa(x)\cdot\kappa^\circ(A^T y).$$

2. (Strong duality)
   If the dual (resp. primal) is feasible and the primal (resp. dual) is relatively strictly feasible, then $\nu_p\nu_d = 1$ and the gauge dual (resp. primal) attains its optimal value.

Freund (1987), Friedlander-Macedo-Pong (2014)

Gauge Duality and Sensitivity

$$v_p(y) := \inf_{\mu>0,\,x} \{\mu \mid \rho\,(b - Ax + \mu y) \le \tau,\ \kappa(x) \le \mu\}$$

Gauge Duality and Sensitivity

$$v_p(y) := \inf_{\mu>0,\, x} \{\mu \mid \rho\,(b - Ax + \mu y) \le \tau,\ \kappa(x) \le \mu\}$$

$$\lambda := 1/\mu \text{ and } w := x/\mu$$

Gauge Duality and Sensitivity

$$v_p(y) := \inf_{\mu>0,\, x} \{\mu \mid \rho\,(b - Ax + \mu y) \le \tau,\ \kappa(x) \le \mu\}$$

$\lambda := 1/\mu$ and $w := x/\mu$

$$= \inf_{\lambda>0,\, w} \{1/\lambda \mid \rho(\lambda b - Aw + y) \le \tau\lambda,\ w \in \mathbb{U}_\kappa\},$$

Gauge Duality and Sensitivity

$$v_p(y) := \inf_{\mu>0,\, x} \{\mu \mid \rho\,(b - Ax + \mu y) \le \tau,\ \kappa(x) \le \mu\}$$

$$\lambda := 1/\mu \text{ and } w := x/\mu$$

$$= \inf_{\lambda>0,\, w} \{1/\lambda \mid \rho(\lambda b - Aw + y) \le \tau\lambda,\ w \in \mathbb{U}_\kappa\},$$

Variational framework:

$$F(w,\lambda,y) := -\lambda + \delta_{(\operatorname{epi}\rho)\times\mathbb{U}_\kappa}\left(W\begin{pmatrix} w \\ \lambda \\ y \end{pmatrix}\right), \quad W := \begin{pmatrix} -A & b & I \\ 0 & \tau & 0 \\ I & 0 & 0 \end{pmatrix}$$

$$F^*(w,\lambda,y) = \delta_{\operatorname{epi}\rho^\circ}\begin{pmatrix} y \\ -\sigma^{-1}(1+\lambda-\langle b,y\rangle) \end{pmatrix} + \kappa^\circ(w + A^T y)$$

Gauge Duality and Sensitivity

$$p(y) := \inf_{w,\lambda} F(w, \lambda, y)$$

$$\begin{aligned} v_p(y) := \inf_x \;& \mu \\ \text{s.t.}\;& \rho(b - Ax + \mu y) \le \tau \\ & \kappa(x) \le \mu \end{aligned}$$

$$\begin{aligned} v_d(w, \lambda) := \inf_z \;& \kappa^\circ(A^T z + w) \\ \text{s.t.}\;& \langle b, y\rangle - \tau\rho^\circ(z) \ge 1 + \lambda \end{aligned}$$

Gauge Duality and Sensitivity

$$p(y) := \inf_{w,\lambda} F(w,\lambda,y)$$

$$\begin{aligned} v_p(y) := \inf_x \ & \mu \\ \text{s.t. } \ & \rho(b - Ax + \mu y) \le \tau \\ & \kappa(x) \le \mu \end{aligned} \qquad \begin{aligned} v_d(w,\lambda) := \inf_z \ & \kappa^\circ(A^T z + w) \\ \text{s.t. } \ & \langle b, y\rangle - \tau\rho^\circ(z) \ge 1 + \lambda \end{aligned}$$

1. If the primal is strictly feasible and the dual is feasible, then the set of optimal solutions for the dual is nonempty and bounded, and coincides with

$$\partial p(0) = \partial(-1/v_p)(0) = \frac{1}{v_p^2(0)} \partial v_p(0).$$

Gauge Duality and Sensitivity

$$p(y) := \inf_{w,\lambda} F(w,\lambda,y)$$

$$\begin{aligned} v_p(y) := \inf_x \ & \mu \\ \text{s.t.} \ & \rho(b - Ax + \mu y) \le \tau \\ & \kappa(x) \le \mu \end{aligned} \qquad \begin{aligned} v_d(w,\lambda) := \inf_z \ & \kappa^\circ(A^T z + w) \\ \text{s.t.} \ & \langle b, y\rangle - \tau\rho^\circ(z) \ge 1 + \lambda \end{aligned}$$

1. If the primal is strictly feasible and the dual is feasible, then the set of optimal solutions for the dual is nonempty and bounded, and coincides with

$$\partial p(0) = \partial(-1/v_p)(0) = \frac{1}{v_p^2(0)} \partial v_p(0).$$

2. If the dual is strictly feasible and the primal is feasible, then the set of optimal solutions for the primal is nonempty and bounded with solutions given by $x^* = w^*/\lambda^*$, where

$$(w^*, \lambda^*) \in \partial v_d(0,0) \text{ and } \lambda^* > 0.$$

Gauge Duality and Optimality Conditions

Suppose both the gauge primal and dual are strictly feasible. Then the pair (x^*, y^*) is primal-dual optimal if and only

$$\sigma = \rho(b - Ax^*) \qquad \text{(primal activity)}$$

$$1 = \langle b, y^* \rangle - \sigma \rho^\circ(y^*) \qquad \text{(dual activity)}$$

$$\langle x^*, A^T y^* \rangle = \kappa(x^*) \cdot \kappa^\circ(A^T y^*) \qquad \text{(objective alignment)}$$

$$\langle b - Ax^*, y^* \rangle = \rho(b - Ax^*) \cdot \rho^\circ(y^*). \qquad \text{(constraint alignment)}$$

Gauge primal-dual recovery

Suppose that the gauge primal and dual are strictly feasible. If y is optimal for G_d, then for any $x \in \mathbb{R}^n$ the following conditions are equivalent:

(a) x is optimal for G_p;

(b) $\langle x, A^T y\rangle = \kappa(x) \cdot \kappa^\circ(A^T y)$ and $b - Ax \in \sigma\partial\rho^\circ(y)$;

(c) $A^T y \in \kappa^\circ(A^T y) \cdot \partial\kappa(x)$ and $b - Ax \in \sigma\partial\rho^\circ(y)$.

Gauge primal-dual recovery from the Lagrange dual

Suppose that the gauge dual G_d is strictly feasible and the primal G_p is feasible.

Let $\mathcal{L}G_d$ denote the Lagrange dual of G_d, and let ν_L denote its optimal value.

Then

$$z^* \text{ is optimal for } \mathcal{L}G_d \iff z^*/\nu_L \text{ is optimal for } G_p.$$

Perspective Duality

The Perspective-Polar Transform

$$
\begin{aligned}
f^\sharp(x,\xi) &:= (f^\pi)^\circ(x,\xi) \\
&= \sigma_{\mathrm{epi}\,(f^*)^\circ}\,(x,-\xi) \\
&= \gamma_{\mathrm{epi}\,(f^*)}\,(x,-\xi) \\
&= \inf\{\mu > 0 \mid \xi + \langle z, x\rangle \le \mu f(z),\ \forall z\}
\end{aligned}
$$

The Perspective-Polar Transform

$$\begin{aligned} f^\sharp(x,\xi) &:= (f^\pi)^\circ(x,\xi) \\ &= \sigma_{\mathrm{epi}\,(f^*)^\circ}(x,-\xi) \\ &= \gamma_{\mathrm{epi}\,(f^*)}(x,-\xi) \\ &= \inf\{\mu > 0 \mid \xi + \langle z,x\rangle \le \mu f(z),\ \forall z\} \end{aligned}$$

$f^\sharp$ is a gauge.

The Perspective-Polar Transform

$$
\begin{aligned}
f^\sharp(x,\xi) &:= (f^\pi)^\circ(x,\xi) \\
&= \sigma_{\operatorname{epi}(f^*)^\circ}(x,-\xi) \\
&= \gamma_{\operatorname{epi}(f^*)}(x,-\xi) \\
&= \inf\{\mu > 0 \mid \xi + \langle z, x\rangle \le \mu f(z),\ \forall z\}
\end{aligned}
$$

$f^\sharp$ is a gauge.

If f is a gauge, then $f^\sharp(x,\xi) = f^\circ(x) + \delta_{\mathbb{R}_-}(\xi)$.

The Perspective-Polar of a PLQ

Piecewise linear-quadratic (PLQ) functons:

$$g(y) := \sup_{u\in U} \left\{ \langle u, y\rangle - \tfrac{1}{2}\|Lu\|_2^2 \right\}, \quad U := \left\{ u \in \mathbb{R}^l \mid Wu \le w \right\},$$

$$\begin{aligned} g^{\sharp}(y,\mu) &= \delta_{\mathbb{R}_-}(\mu) + \max\left\{\gamma_U(y),\, -(1/2\mu)\|Ly\|^2\right\} \\ &= \delta_{\mathbb{R}_-}(\mu) + \max\left\{-(1/2\mu)\|Ly\|^2,\ \max_{i=1,\dots,k}\left\{ {W_i}^T y/w_i \right\}\right\}, \end{aligned}$$

where $W_1^T, \dots, W_k^T$ are the rows of W.

Perspective duality

Suppose $f : \mathbb{R}^n \to_+$ and $g : \mathbb{R}^m \to_+$ are closed, convex and nonnegative over their domains.

$$N_p \quad \min_{x} \quad f(x) \qquad \text{s.t.} \quad g(b - Ax) \leq \sigma,$$

$$N_d \quad \min_{y,\ \alpha,\ \mu} \quad f^\sharp(A^T y, \alpha) \qquad \text{s.t.} \quad \langle b, y \rangle - \sigma \cdot g^\sharp(y, \mu) \geq 1 - (\alpha + \mu)$$

The Perspective Duality for PLQ

Assume f is a gauge and g is PLQ, then

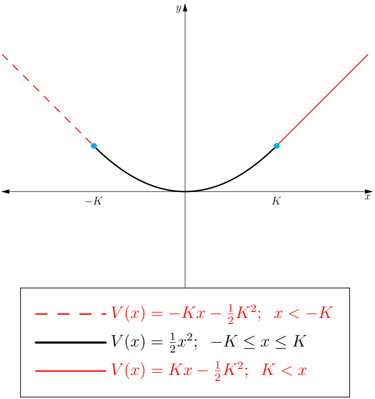
$$\begin{array}{ll} \min_{(y,\mu,\xi)} & f^\circ\big(A^T y\big) \\ \text{s.t.} & \langle b, y\rangle + \mu - \sigma\xi = 1 \\ & \quad Wy \le \xi w,\ \left\|\begin{bmatrix} 2Ly \\ \xi + 2\mu \end{bmatrix}\right\|_2 \le \xi - 2\mu \end{array}$$

Perspective Duality Numerics

$$\min_x \quad \|x\|_1$$

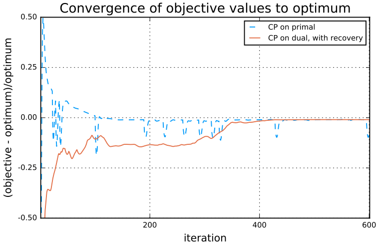
$$\text{s.t.} \quad \sum_{i=1}^{m} V((Ax-b)_i) \le \sigma,$$

where V is the Huber function

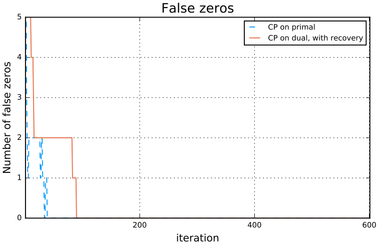


Experiment:
$m = 120$, $n = 512$, $\sigma = 0.2$, $\eta = 1$, and A is a Gaussian matrix. The true solution $x_{\text{true}} \in \{-1, 0, 1\}$ is a spike train which has been constructed to have 20 nonzero entries, and the true noise $b - Ax_{\text{true}}$ has been constructed to have 5 outliers.

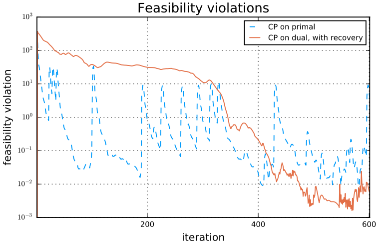
Perspective Duality Numerics



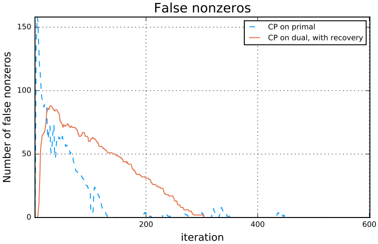
(a)



(c)



(b)



(d)

Chambolle- Pock (CP) algorithm

The Perspective Transform

$$f^\pi(x,\mu) := \operatorname{cl} \inf \ \tau \quad \text{s.t. } (x,\tau,\mu) \in \mathbb{R}_+[\operatorname{epi}(f) \times \{1\}]$$

$$= \begin{cases} \mu f(\mu^{-1}x), & \mu > 0 \\ f^\infty(x), & \mu = 0 \\ +\infty, & \mu < 0 \end{cases}$$

where

$$f^\infty(x) := \sup_{z \in \operatorname{dom}(f)} [f(x+z) - f(x)]$$

is the horizon function of f.

The Subdifferential of the Perspective

$$(f^\pi)^*(y,\xi) = \delta_{\mathrm{epi}\, f^*}\left((y,-\xi)\right)$$

$$\partial f^\pi(x,\mu) = \begin{cases} \{(z,-f^*(z)) \mid z \in \partial f(x/\mu)\} & \text{if } \mu > 0 \\ \\ \{(z,-\gamma) \mid (z,\gamma) \in \mathrm{epi}\, f^*,\;\; z \in \partial f^\infty(x)\} & \text{if } \mu = 0. \end{cases}$$

Properties of the Perspective

$$\sigma_{\mathrm{epi}\,h}\left((y,\mu)\right) = (h^*)^\pi(y,-\mu)$$

$$\sigma_{\mathrm{lev}_h(\tau)}(y) = \mathrm{cl}\inf_{\mu\geq 0}\left[\tau\mu + (h^*)^\pi(y,\mu)\right]$$