Convexity and Duality

April 30, 2018

Dual pairs of problems

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

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

Dual pairs of problems

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

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

The (Fenchel-Rockafellar) dual problem:

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

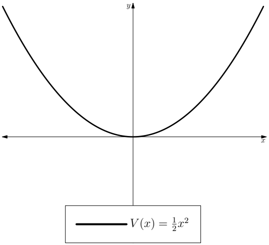
Piecewise Linear-Quadratic Penalties

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

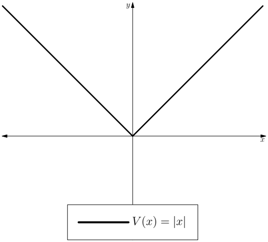
$U \subset \mathbb{R}^n$ is nonempty, closed and convex with $0 \in U$.
$B \in \mathbb{R}^{n \times n}$ is symmetric positive semi-definite.

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

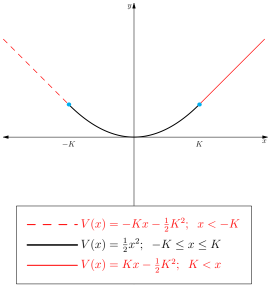
PLQ Densities: Gauss, Laplace, Huber, Vapnik



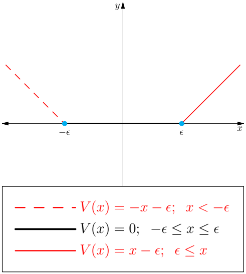
Gauss



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

Convex Sets

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

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

where

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

is the line segment connecting x and y.

Convex Cones

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

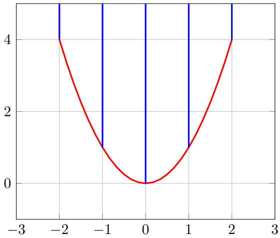
$$\lambda K \subset K \;\; \forall\, \lambda > 0 \;\; \text{and} \; K + K \subset K.$$

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.

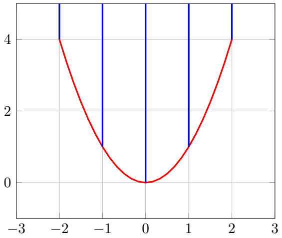


Convex functions and the epigraphical perspective

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

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

are convex.

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.

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.

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.

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

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 $\mathrm{conv}\,(S)$.

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 $\operatorname{aff}(S)$.

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

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

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 $\mathrm{aff}\,(S)$.

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

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

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

$$A\,\mathrm{ri}\,(C) = \mathrm{ri}\,(AC) \quad \text{and}$$
$$B^{-1}\mathrm{ri}\,(C) = \mathrm{ri}\,(B^{-1}C), \quad \text{whenever } B^{-1}\mathrm{ri}\,(C) \neq \emptyset.$$

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 C be a non-empty convex set in the Euclidean space $\mathbb{E}$, and let $M \subset \mathbb{E}$ be a nonempty affine set such that

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

Then there is a hyperplane H in $\mathbb{E}$ 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,

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

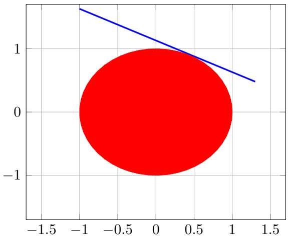
In this case, H is said to be a supporting hyperplane to C at $\overline{x}$.

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 $\overline{x}$.

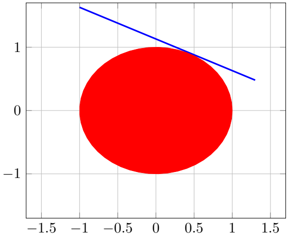


Supporting hyperplanes

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

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

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

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

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

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

Fact: Support functions are sublinear:

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

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

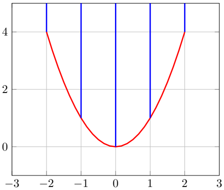
where $C := \{z \mid \langle z, x \rangle \leq \sigma(x) \; \forall \, x\} = \{z \mid \langle z, x \rangle \leq 1 \; \forall \, \sigma(x) \leq 1\}$.

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.



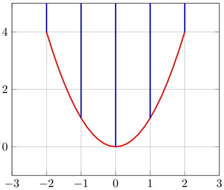
$\operatorname{epi}(f)$

Convex functions and the epigraphical perspective

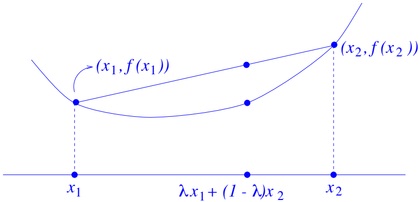
A function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be convex if

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

is convex.



epi (f)



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

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

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

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 $+\infty$.

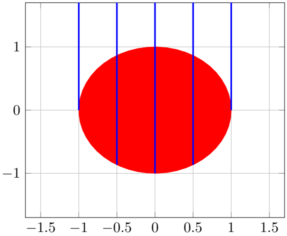
Coordinate inf-projection of a convex set

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 $+\infty$.

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



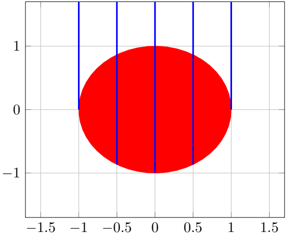
Coordinate inf-projection of a convex set

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 $+\infty$.

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



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

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

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

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:

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

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

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

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

Subgradients:

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

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

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

Subgradients:

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

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

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

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

Subgradients:

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

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

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

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

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

Subgradients:

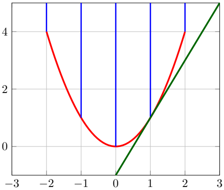
$$\overline{x} \in \operatorname*{argmax}_x [\langle z,x\rangle - f(x)] \iff (z,-1) \text{ supports epi}\,(f) \text{ at } (\overline{x}, f(\overline{x})).$$

$$\iff \langle (z,-1),(\overline{x}, f(\overline{x})\rangle \ge \langle (z,-1),(x,f(x))\rangle \ \forall\, x \in \mathrm{dom}\, f,$$

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

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

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

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

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

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

But

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

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

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

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

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

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

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

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

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

$$\iff$$

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

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

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

$$\Longleftrightarrow$$

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

$$\Longrightarrow$$

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

$$\Longleftrightarrow$$

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

$$\Longleftrightarrow$$

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

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 $0 \in \mathrm{ri}\,(\mathrm{dom}\, p)$, then $p(0) = -q(0)$ and the infimum $q(0)$ is attained, if finite, in which case $\mathrm{argmax}_z\, -F^*(0,z)$ is nonempty and $\partial p(0) = \mathrm{argmax}_z\, -F^*(0,z)$. If, in fact, $0 \in \mathrm{int}\,(\mathrm{dom}\, p)$ and $p(0)$ is finite, then $\partial p(0)$ is bounded.

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, in which case $\operatorname{argmax}_z -F^*(0,z)$ is nonempty and $\partial p(0) = \operatorname{argmax}_z -F^*(0,z)$. If, in fact, $0 \in \operatorname{int}(\operatorname{dom} p)$ and $p(0)$ is finite, then $\partial p(0)$ is bounded.

2. If $0 \in \operatorname{ri}(\operatorname{dom} q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite, in which case $\operatorname{argmin}_x F(x,0)$ is nonempty and $\partial q(0) = \operatorname{argmin}_x F(x,0)$. If, in fact, $0 \in \operatorname{int}(\operatorname{dom} q)$ and $q(0)$ is finite, then $\partial q(0)$ is bounded.

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, in which case $\operatorname{argmax}_z -F^*(0,z)$ is nonempty and $\partial p(0) = \operatorname{argmax}_z -F^*(0,z)$. If, in fact, $0 \in \operatorname{int}(\operatorname{dom} p)$ and $p(0)$ is finite, then $\partial p(0)$ is bounded.
2. If $0 \in \operatorname{ri}(\operatorname{dom} q)$, then $p(0) = -q(0)$ and the infimum $p(0)$ is attained, if finite, in which case $\operatorname{argmin}_x F(x,0)$ is nonempty and $\partial q(0) = \operatorname{argmin}_x F(x,0)$. If, in fact, $0 \in \operatorname{int}(\operatorname{dom} q)$ and $q(0)$ is finite, then $\partial q(0)$ is bounded.
3. Optimal solutions are characterized by

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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:

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

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:

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

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

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

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