Gauge optimization: Duality and polar envelope

Ting Kei Pong
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong

ICCOPT 2019
August 2019

(Joint work with Michael Friedlander and Ives Macêdo)

Motivating example

Minimum norm solutions:

- In sparse optimization:

$$\begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & \|b - Ax\| \le \sigma, \end{array}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $\sigma < \|b\|$.

Motivating example

Minimum norm solutions:

- In sparse optimization:

$$\begin{array}{ll} \min & \|x\|_1 \\ \text{s.t.} & \|b - Ax\| \leq \sigma, \end{array}$$

  where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $\sigma < \|b\|$.

- More generally, minimization of atomic norm (Chandrasekaran et al. '12)

$$\|x\|_{\mathcal{A}} = \inf\{\lambda \geq 0 : \; x \in \lambda \operatorname{conv} \mathcal{A}\},$$

  where $\mathcal{A}$ is a set of "atoms" characterizing the notion of sparsity:
  - ⋆ $\mathcal{A} = \{\pm e_i : \; i = 1, ..., n\} \;\; \Rightarrow \;\; \|x\|_{\mathcal{A}} = \sum_{i=1}^n |x_i|$.
  - ⋆ $\mathcal{A} =$ unit norm rank 1 matrices $\;\; \Rightarrow \;\; \|X\|_{\mathcal{A}} = \sum_{i=1}^n \sigma_i(X)$.

Gauges

- Gauges are generalizations of norms: nonnegative convex positively homogeneous functions that are zero at the origin.
- $\kappa(x) = \inf\{\lambda \geq 0 : \; x \in \lambda U\}$ for some convex set U.

Gauges

- Gauges are generalizations of norms: nonnegative convex positively homogeneous functions that are zero at the origin.
- $\kappa(x) = \inf\{\lambda \geq 0 : \; x \in \lambda U\}$ for some convex set U.
- Polar gauge generalizes dual norm:

$$\begin{aligned}\kappa^\circ(y) &= \inf\{\lambda > 0 : \; \langle x, y\rangle \leq \lambda\kappa(x) \; \forall x\} \\ &= \sup\{\langle x, y\rangle : \; \kappa(x) \leq 1\}.\end{aligned}$$

- Generalized Cauchy inequality: for all $x \in \operatorname{dom}\kappa$ and $y \in \operatorname{dom}\kappa^\circ$,

$$\langle x, y\rangle \leq \kappa(x)\kappa^\circ(y).$$

Further example of gauge

Conic gauge optimization:

- In conic optimization:

$$\begin{array}{ll} \min & \langle c, x \rangle \\ \text{s.t.} & Ax = b,\ x \in \mathcal{K}. \end{array}$$

  If $c \in \mathcal{K}^*$, then $\langle c, \cdot \rangle + \delta_{\mathcal{K}}(\cdot)$ is a gauge.

- Examples: SDP relaxation of max-cut, phase retrieval...

Further example of gauge

Conic gauge optimization:

- In conic optimization:

$$\begin{array}{rl} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b,\ x \in \mathcal{K}. \end{array}$$

  If $c \in \mathcal{K}^*$, then $\langle c, \cdot\rangle + \delta_{\mathcal{K}}(\cdot)$ is a gauge.

- Examples: SDP relaxation of max-cut, phase retrieval...

- More generally, if there exists $\hat{y}$ so that $c - A^*\hat{y} \in \mathcal{K}^*$, then $\hat{c} := c - A^*\hat{y} \in \mathcal{K}^*$ and $\langle \hat{c}, \cdot\rangle + \delta_{\mathcal{K}}(\cdot)$ is a gauge.

Gauge optimization

$$v_p := \min \quad \kappa(x) \quad \text{s.t.} \quad \rho(b - Ax) \le \sigma. \tag{P_ρ}$$

- κ is a gauge.
- ρ is a closed gauge with $\rho^{-1}(0) = \{0\}$, $0 \le \sigma < \rho(b)$.

Gauge optimization

$$v_p := \min \quad \kappa(x) \quad \text{s.t.} \quad \rho(b - Ax) \le \sigma. \tag{P_ρ}$$

- κ is a gauge.
- ρ is a closed gauge with $\rho^{-1}(0) = \{0\}$, $0 \le \sigma < \rho(b)$.
- Lagrange and gauge dual problems:

$$v_\ell := \max \quad \langle b, y\rangle - \sigma\rho^\circ(y) \quad \text{s.t.} \quad \kappa^\circ(A^*y) \le 1. \qquad v_g := \min \quad \kappa^\circ(A^*y) \quad \text{s.t.} \quad \langle b, y\rangle - \sigma\rho^\circ(y) \ge 1.$$

- The role of objective and constraint is reversed in the gauge dual.

Outline

- Gauge duality: general framework.
- Gauge duality: structured problem.
- Smoothing technique: polar envelope.

Gauge duality framework

Let $\mathcal{C}$ be a nonempty closed convex set not containing the origin, and define its anti-polar

$$\mathcal{C}' = \{u : \ \langle u, x \rangle \geq 1 \ \ \forall x \in \mathcal{C}\}.$$

Gauge duality framework

Let $\mathcal{C}$ be a nonempty closed convex set not containing the origin, and define its anti-polar

$$\mathcal{C}' = \{u : \langle u, x \rangle \geq 1 \;\; \forall x \in \mathcal{C}\}.$$

Freund ('87) defined the following primal-dual gauge pairs:

$$\begin{aligned} v_p := \min \quad & \kappa(x) \\ \text{s.t.} \quad & x \in \mathcal{C}, \end{aligned} \tag{P}$$

$$\begin{aligned} v_g := \min \quad & \kappa^\circ(u) \\ \text{s.t.} \quad & u \in \mathcal{C}'. \end{aligned} \tag{D}$$

Gauge duality framework

Let $\mathcal{C}$ be a nonempty closed convex set not containing the origin, and define its anti-polar

$$\mathcal{C}' = \{u : \langle u, x\rangle \geq 1 \ \ \forall x \in \mathcal{C}\}.$$

Freund ('87) defined the following primal-dual gauge pairs:

$$\begin{aligned} v_p := \min \quad & \kappa(x) \\ \text{s.t.} \quad & x \in \mathcal{C}, \end{aligned} \tag{P}$$

$$\begin{aligned} v_g := \min \quad & \kappa^\circ(u) \\ \text{s.t.} \quad & u \in \mathcal{C}'. \end{aligned} \tag{D}$$

Theorem 1. [Strong duality] (Freund '87)
Suppose that κ is closed, $\operatorname{ri}\operatorname{dom}\kappa^\circ \cap \operatorname{ri}\mathcal{C}' \neq \emptyset$ and $\operatorname{ri}\operatorname{dom}\kappa \cap \operatorname{ri}\mathcal{C} \neq \emptyset$. Then $v_p v_g = 1$ and both values are attained.

Anti-polar calculus

Let $\mathcal{D} := \{u : \ \rho(b-u) \le \sigma\}$. Then

$$\mathcal{C} = \{x : \ \rho(b-Ax) \le \sigma\} = A^{-1}\mathcal{D}.$$

How do we compute $\mathcal{C}'$?

Fact 2.

$$\mathcal{D}' = \{y : \ \langle b, y\rangle - \sigma\rho^\circ(y) \ge 1\}.$$

Proposition 1. (Friedlander, Macêdo, P. '14)

$$(A^{-1}\mathcal{D})' = \mathrm{cl}(A^*\mathcal{D}').$$

If, in addition, $\mathrm{ri}\,\mathcal{D} \cap \mathrm{Range}\,A \neq \emptyset$, then

$$(A^{-1}\mathcal{D})' = A^*\mathcal{D}' = \{A^*y : \ \langle b, y\rangle - \sigma\rho^\circ(y) \ge 1\}.$$

Strong duality

Consider the following primal-dual gauge pairs:

$$
\begin{aligned}
v_p := \min \quad & \kappa(x) \\
\text{s.t.} \quad & \rho(b - Ax) \le \sigma,
\end{aligned}
\tag{P_ρ}
$$

$$
\begin{aligned}
v_g := \min \quad & \kappa^\circ(A^* y) \\
\text{s.t.} \quad & \langle b, y \rangle - \sigma \rho^\circ(y) \ge 1.
\end{aligned}
\tag{D_ρ}
$$

Strong duality

Consider the following primal-dual gauge pairs:

$$v_p := \min \quad \kappa(x) \quad \text{s.t.} \quad \rho(b - Ax) \le \sigma, \tag{P$_\rho$}$$

$$v_g := \min \quad \kappa^\circ(A^* y) \quad \text{s.t.} \quad \langle b, y \rangle - \sigma \rho^\circ(y) \ge 1. \tag{D$_\rho$}$$

Let $\mathcal{D} := \{u : \ \rho(b - u) \le \sigma\}$ so that primal feasible set is $A^{-1}\mathcal{D}$.

Theorem 2. (Friedlander, Macêdo, P. '14)
Suppose that κ is closed, $\operatorname{ri}\operatorname{dom}\kappa^\circ \cap \operatorname{ri} A^*\mathcal{D}' \neq \emptyset$ and $\operatorname{ri}\operatorname{dom}\kappa \cap A^{-1}\operatorname{ri}\mathcal{D} \neq \emptyset$. Then $v_p v_g = 1$ and both values are attained.

Strong duality

Consider the following primal-dual gauge pairs:

$$v_p := \min \quad \kappa(x) \quad \text{s.t.} \quad \rho(b - Ax) \le \sigma, \tag{P_ρ}$$

$$v_g := \min \quad \kappa^\circ(A^* y) \quad \text{s.t.} \quad \langle b, y\rangle - \sigma\rho^\circ(y) \ge 1. \tag{D_ρ}$$

Let $\mathcal{D} := \{u : \ \rho(b - u) \le \sigma\}$ so that primal feasible set is $A^{-1}\mathcal{D}$.

Theorem 2. (Friedlander, Macêdo, P. '14)
Suppose that κ is closed, $\operatorname{ri}\operatorname{dom}\kappa^\circ \cap \operatorname{ri} A^*\mathcal{D}' \neq \emptyset$ and $\operatorname{ri}\operatorname{dom}\kappa \cap A^{-1}\operatorname{ri}\mathcal{D} \neq \emptyset$. Then $v_p v_g = 1$ and both values are attained.

Unlike the Lagrange dual, the gauge dual (D_ρ) has a complicated objective and simple constraint.

Solution method: Smoothing

For proper closed convex functions f_1 and f_2, with suitable CQ:

$$\begin{array}{ccc}
\underset{x}{\text{minimize}}\ f_1(x) + f_2(x) & \xrightarrow{\text{addition}} & \underset{x}{\text{minimize}}\ (f_1 + \frac{\alpha}{2}\|\cdot\|_2^2)(x) + f_2(x) \\
\text{Fenchel duality} \Big\updownarrow\ x \in \partial f_1^*(y) \cap \partial f_2^*(-y) & & \text{Fenchel duality} \Big\updownarrow\ x = \nabla(f_1^* \,\square\, \frac{1}{2\alpha}\|\cdot\|_2^2)(y) \\
\underset{y}{\text{minimize}}\ f_1^*(y) + f_2^*(-y) & \xrightarrow{\text{sum convolution}} & \underset{y}{\text{minimize}}\ (f_1^* \,\square\, \frac{1}{2\alpha}\|\cdot\|_2^2)(y) + f_2^*(-y)
\end{array}$$

Here, $f_1^* \square \frac{1}{2\alpha}\|\cdot\|^2$ is the Moreau envelope of αf_1^*:

$$(f_1^* \square \frac{1}{2\alpha}\|\cdot\|^2)(x) := \inf_{y \in \mathbb{R}^n} \left\{ f_1^*(y) + \frac{1}{2\alpha}\|x - y\|^2 \right\}$$

which is smooth.

Solution method: Smoothing

For proper closed convex functions f_1 and f_2, with suitable CQ:

$$\begin{array}{ccc}
\underset{x}{\text{minimize}}\ f_1(x) + f_2(x) & \xrightarrow{\text{addition}} & \underset{x}{\text{minimize}}\ (f_1 + \tfrac{\alpha}{2}\|\cdot\|_2^2)(x) + f_2(x) \\
\text{Fenchel duality} \Big\updownarrow\ x \in \partial f_1^*(y) \cap \partial f_2^*(-y) & & \text{Fenchel duality} \Big\updownarrow\ x = \nabla(f_1^* \,\Box\, \tfrac{1}{2\alpha}\|\cdot\|_2^2)(y) \\
\underset{y}{\text{minimize}}\ f_1^*(y) + f_2^*(-y) & \xrightarrow{\text{sum convolution}} & \underset{y}{\text{minimize}}\ (f_1^* \,\Box\, \tfrac{1}{2\alpha}\|\cdot\|_2^2)(y) + f_2^*(-y)
\end{array}$$

Here, $f_1^* \Box \frac{1}{2\alpha}\|\cdot\|^2$ is the Moreau envelope of αf_1^*:

$$(f_1^* \Box \frac{1}{2\alpha}\|\cdot\|^2)(x) := \inf_{y \in \mathbb{R}^n} \left\{ f_1^*(y) + \frac{1}{2\alpha}\|x - y\|^2 \right\}$$

which is smooth. Does not preserve gauge structure!

Polar envelope

Definition. (Friedlander, Macêdo, P. '18)
Let κ be a gauge, $\alpha > 0$. The polar envelope and polar proximal mapping are

$$\kappa_\alpha(x) := \inf_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\},$$

$$\text{pprox}_{\alpha\kappa}(x) := \operatorname*{Arg\,min}_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\}.$$

Polar envelope

Definition. (Friedlander, Macêdo, P. '18)
Let κ be a gauge, $\alpha > 0$. The polar envelope and polar proximal mapping are

$$\kappa_\alpha(x) := \inf_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\},$$

$$\mathrm{pprox}_{\alpha\kappa}(x) := \operatorname{Arg\,min}_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\}.$$

Note:

- For proper convex functions f_1 and f_2, their max-convolution is

$$(f_1 \diamond f_2)(x) := \inf_z \max\{f_1(z), f_2(x - z)\}$$

Polar envelope

Definition. (Friedlander, Macêdo, P. '18)
Let κ be a gauge, $\alpha > 0$. The polar envelope and polar proximal mapping are

$$\kappa_\alpha(x) := \inf_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\},$$

$$\text{pprox}_{\alpha\kappa}(x) := \operatorname*{Arg\,min}_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\}.$$

Note:

- For proper convex functions f_1 and f_2, their max-convolution is

$$(f_1 \diamond f_2)(x) := \inf_z \max\{f_1(z), f_2(x - z)\}$$

- For gauges κ_1 and κ_2: (Friedlander, Macêdo, P. '18)

$$(\kappa_1 \diamond \kappa_2)^\circ = \kappa_1^\circ + \kappa_2^\circ.$$

Polar envelope

Definition. (Friedlander, Macêdo, P. '18)
Let κ be a gauge, $\alpha > 0$. The polar envelope and polar proximal mapping are

$$\kappa_\alpha(x) := \inf_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\},$$

$$\mathrm{pprox}_{\alpha\kappa}(x) := \operatorname*{Arg\,min}_z \max\left\{\kappa(z), \frac{1}{\alpha}\|x - z\|\right\}.$$

Note:

- For proper convex functions f_1 and f_2, their max-convolution is

$$(f_1 \diamond f_2)(x) := \inf_z \max\{f_1(z), f_2(x - z)\}$$

- For gauges κ_1 and κ_2: (Friedlander, Macêdo, P. '18)

$$(\kappa_1 \diamond \kappa_2)^\circ = \kappa_1^\circ + \kappa_2^\circ.$$

Recall that for proper convex functions f_1 and f_2, $(f_1 \square f_2)^* = f_1^* + f_2^*$.

Polar envelope: Differential properties

Theorem 3. (Friedlander, Macêdo, P. '18)
Let κ be a gauge and $\alpha > 0$.

(i) If $\bar{x} \in \mathrm{pprox}_{\alpha\kappa}(x)$, then $\|x - \bar{x}\| \geq \alpha\kappa(\bar{x})$. If in addition κ is continuous, then $\|x - \bar{x}\| = \alpha\kappa(\bar{x})$.

(ii) If κ is closed, then $\mathrm{pprox}_{\alpha\kappa}(x)$ is a singleton for all x, and $\mathrm{pprox}_{\alpha\kappa}$ is continuous and positively homogeneous.

(iii) Suppose κ is closed. Then κ_α is differentiable at all x such that $\kappa_\alpha(x) > 0$. Moreover, at these x, it holds that $\langle x, x - \bar{x}\rangle > 0$ and

$$\nabla\kappa_\alpha(x) = \frac{\|x - \bar{x}\|}{\alpha\langle x, x - \bar{x}\rangle}(x - \bar{x}),$$

where $\bar{x} = \mathrm{pprox}_{\alpha\kappa}(x)$.

Polar envelope: Explicit example

Proposition 2. (Friedlander, Macêdo, P. '18)
Let κ be a continuous gauge. Then for any x satisfying $\kappa_\alpha(x) > 0$, it holds that

$$\kappa_\alpha(x) = \bar{r} \;\text{ and }\; \mathrm{pprox}_{\alpha\kappa}(x) = \mathrm{Proj}_{[\kappa \le \bar{r}]}(x),$$

where $\bar{r}$ is the unique root satisfying

$$\alpha^2 \bar{r}^2 = \|x - \mathrm{Proj}_{[\kappa \le \bar{r}]}(x)\|^2. \qquad (\clubsuit)$$

Polar envelope: Explicit example

Proposition 2. (Friedlander, Macêdo, P. '18)
Let κ be a continuous gauge. Then for any x satisfying $\kappa_\alpha(x) > 0$, it holds that

$$\kappa_\alpha(x) = \bar{r} \ \text{ and } \ \mathrm{pprox}_{\alpha\kappa}(x) = \mathrm{Proj}_{[\kappa \le \bar{r}]}(x),$$

where $\bar{r}$ is the unique root satisfying

$$\alpha^2 \bar{r}^2 = \|x - \mathrm{Proj}_{[\kappa \le \bar{r}]}(x)\|^2. \qquad (\clubsuit)$$

Example: Consider $\kappa = \|\cdot\|_\infty$. For $x \neq 0$, ($\clubsuit$) becomes

$$\alpha^2 \bar{r}^2 = \sum_{i=1}^{n} (|x_i| - \bar{r})_+^2 \quad \overset{\bar{r} = \bar{\gamma}^{-1}}{\Longleftrightarrow} \quad \alpha^2 = \sum_{i=1}^{n} (\bar{\gamma}|x_i| - 1)_+^2.$$

Solved by simple root-finding procedure.

Solution method: Smoothing revisited

For closed gauges κ, ρ satisfying $\kappa^{-1}(0) = \{0\}$ and $\rho^{-1}(0) = \{0\}$, $\mathcal{C} = \{x : \ \rho(b - Ax) \le \sigma\}$ and $\sigma \in [0, \rho(b))$, with suitable CQ:

$$
\begin{array}{ccc}
\displaystyle\operatorname*{minimize}_{x\in\mathcal{C}} \ \kappa(x) & \xrightarrow{\quad\text{addition}\quad} & \displaystyle\operatorname*{minimize}_{x\in\mathcal{C}} \ (\kappa + \alpha\|\cdot\|_2)(x) \\
\text{gauge duality} \updownarrow \ x \in [\mathrm{cl\,cone}\,\partial\kappa^\circ(y)] \cap \partial\delta^*_{\mathcal{C}}(-y) & & \text{gauge duality} \updownarrow \ x = (\text{Theorem 6.2})(y) \\
\displaystyle\operatorname*{minimize}_{y\in\mathcal{C}'} \ \kappa^\circ(y) & \xrightarrow{\quad\text{max convolution}\quad} & \displaystyle\operatorname*{minimize}_{y\in\mathcal{C}'} \ (\kappa^\circ \diamond \tfrac{1}{\alpha}\|\cdot\|_2)(y)
\end{array}
$$

Here, "$x = (\text{Theorem 6.2})(y)$" refers to (Friedlander, Macêdo, P. '18)

$$
\bar{x} = \frac{\bar{r}^{-1}}{\kappa(\mathrm{prox}_{\bar{r}\kappa}(A^*\bar{u})) + \alpha\|\mathrm{prox}_{\bar{r}\kappa}(A^*\bar{u}))\|}\mathrm{prox}_{\bar{r}\kappa}(A^*\bar{u})),
$$

where $A^*\bar{u}$ solves the smoothed gauge dual with optimal value $\bar{r}$.

Conclusion

- Gauge optimization framework captures many applications.
- Gauge strong duality holds under conditions similar to standard CQ in Lagrange duality theory.
- Polar envelope and polar proximal mapping appear naturally in dual smoothing and primal solution recovery.

Reference:

- M. Friedlander, I. Macêdo and T. K. Pong.
  Gauge Optimization and Duality.
  SIAM J. Optim. 24, 2014, pp. 1999–2022.
- M. Friedlander, I. Macêdo and T. K. Pong.
  Polar convolution.
  SIAM J. Optim. 29, 2019, pp. 1366–1391.

Thanks for coming! $\ddot\smile$