Error bounds for conic problems and an application to the exponential cone

Bruno F. Lourenço
ISM
joint work with Scott Lindstrom (Curtin University) and
Ting Kei Pong (HK Poly. U)

July 22th, 2021
SIAM conference on optimization

Feasibility problems over convex cones

Consider the following *feasibility problem over a convex cone* $\mathcal{K}$.

$$\begin{array}{ll} \text{find} & x \\ \text{subject to} & x \in (\mathcal{L} + a) \cap \mathcal{K} \end{array}$$

- $\mathcal{K}$: closed convex cone contained in some space $\mathcal{E}$.
- $\mathcal{L}$: subspace contained in $\mathcal{E}$.
- $a \in \mathcal{E}$.

($\mathcal{L} + a$ is an affine space)

Scott B. Lindstrom; L and Ting Kei Pong
Error bounds, facial residual functions and applications to the exponential cone
arXiv:2010.16391

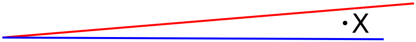
Motivation

Let $\|\cdot\|$ be the Euclidean norm and fix $x \in \mathcal{E}$.

$$\mathrm{dist}\,(x, \mathcal{L} + a) = \inf\{\|x - y\| \mid y \in \mathcal{L} + a\}$$
$$\mathrm{dist}\,(x, \mathcal{K}) = \inf\{\|x - y\| \mid y \in \mathcal{K}\}$$
$$\mathrm{dist}\,(x, (\mathcal{L} + a) \cap \mathcal{K}) = \inf\{\|x - y\| \mid y \in (\mathcal{L} + a) \cap \mathcal{K}\}$$

Fundamental question

Can we estimate $\mathrm{dist}\,(x, (\mathcal{L} + a) \cap \mathcal{K})$ using $\mathrm{dist}\,(x, \mathcal{L} + a)$ and $\mathrm{dist}\,(x, \mathcal{K})$?



- Convergence analysis often leads to this type of questions.

Hölderian error bounds

C_1, C_2: closed convex sets.
$C := C_1 \cap C_2$

Definition (Hölderian error bound)

C_1, C_2 satisfy a **Hölderian error bound** $\overset{\text{def}}{\iff}$ for every bounded set B there exist $\theta_B > 0$, $\gamma_B \in (0,1]$ such that

$$\mathrm{dist}\,(x,\ C) \le \theta_B(\mathrm{dist}\,(x, C_1) + \mathrm{dist}\,(x, C_2))^{\gamma_B} \quad \forall\ x \in B.$$

If $\gamma_B = \gamma \in (0,1]$ for all B, the bound is **uniform**. If the bound is uniform with $\gamma = 1$, we call it a **Lipschitzian** error bound.

Some known results:

- $\mathrm{ri}\, C_1 \cap \mathrm{ri}\, C_2 \neq \emptyset \Rightarrow$ Lipschitzian
- C_1, C_2 are polyhedral $\Rightarrow$ Lipschitzian (Hoffman's Lemma)

Conic results

- C_1: affine space, C_2: PSD cone $\Rightarrow$ Uniform Hölderian (Sturm's error bound)
- C_1: affine space, C_2: **amenable cone** $\Rightarrow$ generalized error bound holds (L'21)
- C_1: affine space, C_2: Symmetric cone $\Rightarrow$ Uniform Hölderian (L'21)

L.
Amenable cones: error bounds without constraint qualifications.
Mathematical Programming, 186:1–48, 2021.

J. F. Sturm.
Error bounds for linear matrix inequalities.
SIAM Journal on Optimization, 10(4):1228–1248, Jan. 2000.

Today's goals

- Prove error bounds for cones that might not be amenable
- Constraint qualifications are **forbidden**! 
- Apply our theory to the exponential cone.

Ingredient 1 - Error bounds under a constraint qualification

$$\begin{array}{ll} \text{find} & x \\ \text{subject to} & x \in (\mathcal{L} + a) \cap \mathcal{K} \end{array} \qquad \text{(CFP)}$$

Proposition (An error bound for when a face satisfying a CQ is known)

Let $\mathcal{F} \trianglelefteq \mathcal{K}$ be such that

- (a) $\mathcal{F}$ *contains* $\mathcal{K} \cap (\mathcal{L} + a)$
- (b) $(\operatorname{ri} \mathcal{F}) \cap (\mathcal{L} + a) \neq \emptyset$

Then, for every bounded set B, there exists $\kappa_B > 0$ such that

$$\operatorname{dist}(x, \mathcal{K} \cap (\mathcal{L} + a)) \leq \kappa_B (\operatorname{dist}(x, \mathcal{F}) + \operatorname{dist}(x, \mathcal{L} + a)), \qquad \forall x \in B.$$

It is not an error bound with respect to $\mathcal{L} + a$ and $\mathcal{K}$, but it is close.

Ingredient 2 - Facial Reduction

Theorem (The facial reduction theorem)

Suppose (CFP) *is feasible. There is a chain of faces*

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$$

and vectors $(z_1, \ldots, z_{\ell-1})$ *such that:*

(i) *For all* $i \in \{1, \ldots, \ell - 1\}$, *we have*

$$z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{a\}^\perp,$$
$$\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp.$$

(ii) $\mathcal{F}_\ell \cap (\mathcal{L} + a) = \mathcal{K} \cap (\mathcal{L} + a)$ *and* $(\mathrm{ri}\, \mathcal{F}_\ell) \cap (\mathcal{L} + a) \neq \emptyset$.

L, M. Muramatsu and T. Tsuchiya.
Facial reduction and partial polyhedrality.
SIAM Journal on Optimization, 28(3), 2018 (http://arxiv.org/abs/1512.02549).

J. M. Borwein and H. Wolkowicz.
Regularizing the abstract convex program.
Journal of Mathematical Analysis and Applications, 83(2):495 – 530, 1981.

Ingredient 3 - Facial Residual Functions

Let

- $\mathcal{K}$: closed convex pointed cone.
- $\mathcal{F}$: face of $\mathcal{K}$
- $z \in \mathcal{F}^*$

Fact:

$$\mathcal{F} \cap \{z\}^\perp = \mathcal{K} \cap \operatorname{span} \mathcal{F} \cap \{z\}^\perp.$$

Definition (Facial residual function for $\mathcal{F}$ and z with respect to $\mathcal{K}$)

If $\psi_{\mathcal{F},z} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

1. $\psi_{\mathcal{F},z}$ is nonnegative, monotone nondecreasing in each argument and $\psi(0,\alpha) = 0$ for every $\alpha \in \mathbb{R}_+$.
2. whenever $x \in \operatorname{span} \mathcal{K}$ satisfies the inequalities

$$\operatorname{dist}(x, \mathcal{K}) \leq \epsilon, \quad \langle x, z \rangle \leq \epsilon, \quad \operatorname{dist}(x, \operatorname{span} \mathcal{F}) \leq \epsilon$$

we have:

$$\operatorname{dist}(x, \mathcal{F} \cap \{z\}^\perp) \leq \psi_{\mathcal{F},z}(\epsilon, \|x\|).$$

Main result

Theorem (Error bound without amenable cones, Lindstrom, L., Pong)

Let $\mathcal{K}$ be a closed convex cone such that $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$. Let

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$$

be a chain of faces of $\mathcal{K}$ together with $z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$ such that

$$(\mathcal{L} + a) \cap \operatorname{ri} \mathcal{F}_\ell \neq \emptyset.$$

and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp$ for every i. Let ψ_i be a facial residual function for $\mathcal{F}_i$, z_i. Then, after positive rescaling the ψ_i, for every bounded set B there are constants $\kappa > 0$, $M > 0$ such that if $x \in \operatorname{span} \mathcal{K} \cap B$ satisfies the inequalities

$$\operatorname{dist}(x, \mathcal{K}) \leq \epsilon, \quad \operatorname{dist}(x, \mathcal{L} + a) \leq \epsilon,$$

we have

$$\operatorname{dist}\, (x, (\mathcal{L} + a) \cap \mathcal{K}) \leq \kappa(\epsilon + \varphi(\epsilon, M)),$$

where $\varphi = \psi_{\ell-1} \diamondsuit \cdots \diamondsuit \psi_1$, if $\ell \geq 2$. If $\ell = 1$, we let φ be the function satisfying $\varphi(\epsilon, \|x\|) = \epsilon$.

$$(f \diamondsuit g)(a, b) := f(a + g(a, b), b).$$

Main result

Theorem (Error bound without amenable cones, Lindstrom, L., Pong)

Let $\mathcal{K}$ be a closed convex cone such that $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$. Let

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$$

be a chain of faces of $\mathcal{K}$ together with $z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$ such that

$$(\mathcal{L} + a) \cap \mathrm{ri}\, \mathcal{F}_\ell \neq \emptyset.$$

and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp$ for every i. Let ψ_i be a facial residual function for $\mathcal{F}_i$, z_i. Then, after positive rescaling the ψ_i, for every bounded set B there are constants $\kappa > 0$, $M > 0$ such that if $x \in \mathrm{span}\, \mathcal{K} \cap B$ satisfies the inequalities

$$\mathrm{dist}\,(x, \mathcal{K}) \leq \epsilon, \quad \mathrm{dist}\,(x, \mathcal{L} + a) \leq \epsilon,$$

we have

$$\mathrm{dist}\,(x, (\mathcal{L} + a) \cap \mathcal{K}) \leq \kappa(\epsilon + \varphi(\epsilon, M)),$$

where $\varphi = \psi_{\ell-1} \diamondsuit \cdots \diamondsuit \psi_1$, if $\ell \geq 2$. If $\ell = 1$, we let φ be the function satisfying $\varphi(\epsilon, \|x\|) = \epsilon$.

$$(f \diamondsuit g)(a, b) := f(a + g(a, b), b).$$

Facial Residual Functions (FRFs) - Examples

- If $\mathcal{K}$ is a symmetric cone, then

$$\psi_{\mathcal{F},z}(\epsilon, t) = \kappa\epsilon + \kappa\sqrt{\epsilon t}$$

  is a FRF, for some $\kappa > 0$. (L'21)
  - Recovers Sturm's error bound when $\mathcal{K} = \mathcal{S}^n_+$.
- If $\mathcal{K}$ is polyhedral, then $\psi_{\mathcal{F},z}(\epsilon, \|x\|) = \kappa\epsilon$ is a FRF, for some $\kappa > 0$.

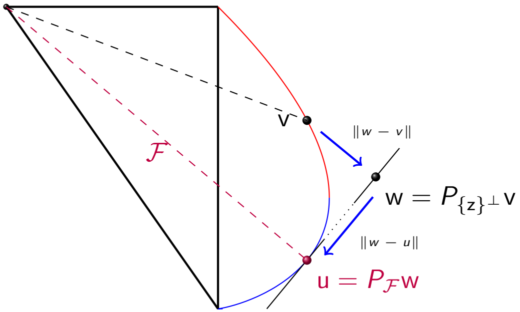
Reminder:

$$\mathrm{dist}\,(x, \mathcal{K}) \le \epsilon, \quad \langle x, z\rangle \le \epsilon, \quad \mathrm{dist}\,(x, \mathrm{span}\,\mathcal{F}) \le \epsilon$$

implies

$$\mathrm{dist}\,(x, \mathcal{F} \cap \{z\}^\perp) \le \psi_{\mathcal{F},z}(\epsilon, \|x\|).$$

FRFs without projection - LLP'21



$$\inf\left\{\frac{\|w-v\|^\alpha}{\|w-u\|}\right\} > 0 \quad \Rightarrow \quad \varphi(\epsilon,t) := \kappa_t\epsilon + \kappa_t\epsilon^\alpha \text{ is FRF}$$

$$\inf\left\{\frac{\mathfrak{g}(\|w-v\|)}{\|w-u\|}\right\} > 0 \quad \Rightarrow \quad \varphi(\epsilon,t) := \kappa_t\epsilon + \kappa_t\mathfrak{g}(2\epsilon) \text{ is FRF}$$

Facial residual functions and $\mathfrak{g}$-amenability

$\mathfrak{g} : \mathbb{R}_+ \to \mathbb{R}_+$: monotone nondecreasing function with $\mathfrak{g}(0) = 0$.

Definition ($\mathfrak{g}$-amenability)

$\mathcal{F} \trianglelefteq \mathcal{K}$ is $\mathfrak{g}$-*amenable* if for every bounded set B, there exists $\kappa > 0$ such that

$$\mathrm{dist}\,(x, \mathcal{F}) \leq \kappa \mathfrak{g}(\mathrm{dist}\,(x, \mathcal{K})), \quad \forall x \in (\mathrm{span}\, \mathcal{F}) \cap B.$$

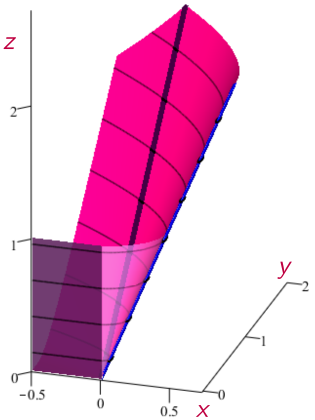
If all faces of $\mathcal{K}$ are $\mathfrak{g}$-amenable, then $\mathcal{K}$ is an $\mathfrak{g}$-*amenable cone*.

Suppose $\mathcal{K}^1$ and $\mathcal{K}^2$ are $\mathfrak{g}$-amenables

- There are calculus rules for the FRFs of $\mathcal{K}^1 \times \mathcal{K}^2$.
- A FRF of a **face** of $\mathcal{K}^1$ can be lifted to a FRF of the whole cone $\mathcal{K}^1$.
- Amenability is recovered when $\mathfrak{g} = |\cdot|$.
  - Amenable cones must be facially exposed.

The exponential cone

$$K_{\exp} := \left\{(x,y,z) \mid y > 0, z \geq ye^{x/y}\right\} \cup \{(x,y,z) \mid x \leq 0, z \geq 0, y = 0\}.$$



The exponential cone

$$K_{\exp} := \left\{(x,y,z) \mid y > 0, z \geq ye^{x/y}\right\} \cup \{(x,y,z) \mid x \leq 0, z \geq 0, y = 0\}\,.$$

$$K^*_{\exp} := \left\{(x,y,z) \mid x < 0, ez \geq -xe^{y/x}\right\} \cup \{(x,y,z) \mid x = 0, ez \geq 0, y \geq 0\}\,.$$

1. **Not exposed!** (So not amenable...)
2. Applications to entropy optimization, logistic regression, geometric programming and etc.
3. Available in Alfonso, Hypatia, Mosek. `https://docs.mosek.com/modeling-cookbook/expo.html`.

V. Chandrasekaran, P. Shah
Relative entropy optimization and its applications.
Math. Program. 161, 1–32 (2017)

The faces of the exponential cone

(a) exposed extreme rays (1D faces) parametrized by $\beta \in \mathbb{R}$:

$$\mathcal{F}_\beta := \left\{ \left(-\beta y + y, y, e^{1-\beta} y\right) \;\middle|\; y \in [0, \infty) \right\}. \qquad \text{(amenable)}$$

(b) an "exceptional" exposed extreme ray:

$$\mathcal{F}_\infty := \{(x, 0, 0) \mid x \leq 0\}. \qquad \text{(amenable)}$$

(c) a **non-exposed** extreme ray: $\mathcal{F}_{ne}$:

$$\mathcal{F}_{ne} := \{(0, 0, z) \mid z \geq 0\}. \qquad \text{($\mathfrak{g}$-amenable, not amenable)}$$

(d) a single 2D exposed face:

$$\mathcal{F}_{-\infty} := \{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\}, \qquad \text{(amenable)}$$

where $\mathcal{F}_\infty$ and $\mathcal{F}_{ne}$ are the extreme rays of $\mathcal{F}_{-\infty}$.

The facial residual functions - 1D exposed faces

1. $\mathcal{F}_\beta := \left\{ \left(-\beta y + y, y, e^{1-\beta} y\right) \;\middle|\; y \in [0, \infty) \right\}$:

$$\psi_{\mathcal{K},z}(\epsilon, t) := \kappa\epsilon + \rho(t)\sqrt{\epsilon}$$

2. $\mathcal{F}_\infty := \{(x, 0, 0) \mid x \leq 0\}$:
   1. if $z_y > 0$, then:

$$\psi_{\mathcal{K},z}(\epsilon, t) := \rho(t)\epsilon$$

   2. if $z_y = 0$, then:

$$\psi_{\mathcal{K},z}(\epsilon, t) := \kappa\epsilon + \rho(t)\mathfrak{g}_\infty(\epsilon),$$

$$\mathfrak{g}_\infty(t) := \begin{cases} 0 & \text{if } t = 0, \\ -\frac{1}{\ln(t)} & \text{if } 0 < t \leq \frac{1}{e^2}, \\ \frac{1}{4} + \frac{1}{4}e^2 t & \text{if } t > \frac{1}{e^2}. \end{cases}$$

The facial residual functions - The 2D face and the non-exposed face

$$\mathfrak{g}_{-\infty}(t) := \begin{cases} 0 & \text{if } \; t = 0, \\ -t\ln(t) & \text{if } \; t \in \left(0, 1/e^2\right], \\ t + \frac{1}{e^2} & \text{if } \; t > 1/e^2. \end{cases}$$

1. $\mathcal{F}_{-\infty} := \{(x, y, z) \mid x \le 0, z \ge 0, y = 0\}$:

$$\psi_{\mathcal{K},z}(\epsilon, t) := \kappa\epsilon + \rho(t)\mathfrak{g}_{-\infty}(\epsilon)$$

2. $\mathcal{F}_{ne} := \{(0, 0, z) \mid z \ge 0\}$:

$$\psi_{\mathcal{F}_{-\infty},z}(\epsilon, t) := \sigma(t)\epsilon + \sigma(t)\mathfrak{g}_{-\infty}(\epsilon)$$

Error bound for problems over the exponential cone

$$\begin{aligned} &\text{find} \quad x \\ &\text{subject to} \quad x \in (\mathcal{L} + a) \cap K_{\exp} \end{aligned} \tag{CFP}$$

Let $z \in (K_{\exp})^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$, $z \neq 0$. Let $\mathcal{F} = K_{\exp} \cap \{z\}^\perp$.

- $\mathcal{F} = \{0\}$: Lipschitzian error bound.
- $\mathcal{F} = \mathcal{F}_\beta$: a Hölderian error bound with exponent $1/2$.
- $\mathcal{F} = \mathcal{F}_\infty$, either a Lipschitzian or a log-type error bound holds depending on the exposing vector.
- $\mathcal{F} = \mathcal{F}_{-\infty}$, an **entropic error bound**: for every bounded set B, there exists $\kappa_B > 0$

$$\operatorname{dist}(x, (\mathcal{L} + a) \cap K_{\exp}) \leq \kappa_B \mathfrak{g}_{-\infty}(\max(\operatorname{dist}(x, \mathcal{L} + a), \operatorname{dist}(x, K_{\exp}))), \quad \forall x \in B.$$

The results above are **optimal**.

Strange error bounds

From the exponential cone we can:

- Obtain sets that **do not have** a Hölderian error bound, but have a logarithmic error bound:
  - Or, a function that does not have a KL exponent.

$$\mathcal{F}_{\infty} = K_{\exp} \cap \{z\}^{\perp},$$

  where $z = (0, 0, 1)$.
- Obtain sets that satisfy a Hölderian bound for all $\gamma \in (0, 1)$ but not $\gamma = 1$. Furthermore, the best error bound is an entropic one.
  - Or, a KL function whose exponent can be arbitrary close to $1/2$ but not $1/2$.

$$\mathcal{F}_{-\infty} = K_{\exp} \cap \{z\}^{\perp},$$

  where $z = (0, 1, 0)$.

Final remarks

- Much more stuff in the paper! Ex: direct products, techniques for obtaining FRFs and so on.

Scott B. Lindstrom; L and Ting Kei Pong
Error bounds, facial residual functions and applications to the exponential cone
arXiv:2010.16391

Other advertisement:

L; Vera Roshchina and James Saunderson
Amenable cones are particularly nice (**MS24**)
arXiv:2011.07745

T. Liu and L.
Convergence analysis under consistent error bounds (**MS69**)
arXiv:2008.12968

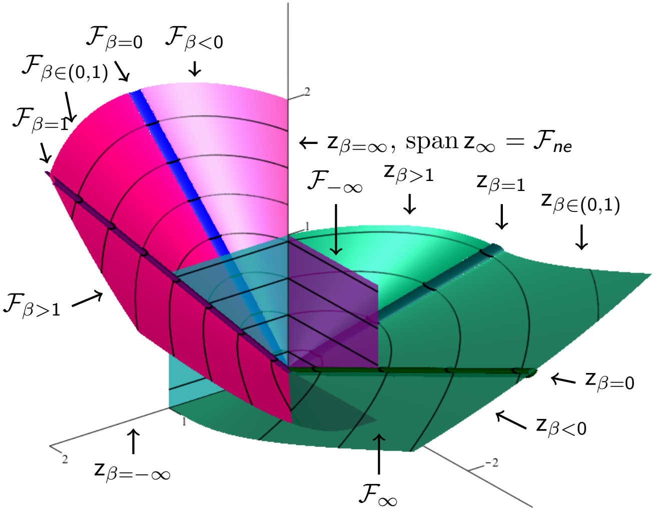


Figure: The exponential cone, its faces and exposing vectors