An (hopefully gentle) introduction to error bounds for conic problems

Bruno F. Lourenço
ISM

July 30th, 2021
SOMA

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & h(x) = 0 \end{aligned}$$

- Suppose I use my favourite solver and obtain x^*.
- The solver tells me that the KKT conditions are satisfied to $\epsilon = 10^{-6}$.
- It also tells me that $\|h(x^*)\| \leq 10^{-7}$.

Question 1

Is x^* close to the set of **optimal** solutions?

Question 2

Is x^* close to the set of **feasible** solutions?

Distance to a set C: $\operatorname{dist}(x, C) := \inf_{y \in C} \|x - y\|$.

An example by Sturm

$$
\begin{aligned}
\min_{x} \quad & x_{22} \\
\text{subject to} \quad & x_{22} = 0 \\
& x_{12} = x_{33} \\
& x \in \mathcal{S}^3_+
\end{aligned}
$$

- $\mathcal{S}^3_+$: 3×3 positive semidefinite matrices.

An example by Sturm

$$\begin{aligned} \min_x \quad & 0 \\ \text{subject to} \quad & \begin{pmatrix} x_{11} & x_{33} & x_{13} \\ x_{33} & 0 & 0 \\ x_{13} & 0 & x_{33} \end{pmatrix} \succeq 0. \end{aligned}$$

- Feasible set: matrices $\begin{pmatrix} x_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ with $x_{11} \geq 0$.

An example by Sturm

Let $\epsilon > 0$

$$x_\epsilon := \begin{pmatrix} 3 & \sqrt{\epsilon} & \sqrt[4]{\epsilon} \\ \sqrt{\epsilon} & \epsilon & 0 \\ \sqrt[4]{\epsilon} & 0 & \sqrt{\epsilon} \end{pmatrix}$$

- The constraints are "$x_{22} = 0$", "$x_{12} = x_{33}$" and "$x \in \mathcal{S}^3_+$".
- Suppose we measure the violation of constraints by x using

$$\mathrm{Res}(x) := [x_{22}^2 + (x_{12} - x_{33})^2 + \max\{-\lambda_{\min}(x), 0\}^2]^{1/2}$$

($\mathrm{Res}(x) = 0 \Leftrightarrow x$ is feasible.) x_ϵ does not seem a bad point:

$$\mathrm{Res}(x_\epsilon) = \epsilon$$

But...

$$\mathrm{dist}\,(x_\epsilon, \mathrm{Feas}) \geq \sqrt[4]{\epsilon}.$$

If $\epsilon = 10^{-5}$, we have $\mathrm{Res}(x_\epsilon) = 10^{-5}$, but $\mathrm{dist}\,(x_\epsilon, \mathrm{Feas}) \geq 0.1$.

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ \text{subject to} & h(x) = 0 \end{array}$$

- Suppose I use my favourite solver and obtain x^*.
- The solver tells me that the KKT conditions are satisfied to $\epsilon = 10^{-6}$.
- It also tells me that $\|h(x^*)\| \leq 10^{-7}$.

Question 1

Is x^* close to the set of **optimal** solutions?

Question 2

Is x^* close to the set of **feasible** solutions?

Answer: **Not necessarily!** Also $\mathrm{Res}(x_\epsilon) \to 0$ does not imply $\mathrm{dist}\,(x_\epsilon, \mathrm{Feas}) \to 0$...

Conclusions

- Using solvers, we input the constraints one by one: $h_1(x) = 0, \ldots, h_n(x) = 0,\ g_1(x) \le 0, g_2(x) \le 0, \ldots, g_m(x) \le 0$.
- Solvers can only compute the residuals with respect the g_i and h_j. (**Backward error**)
  - Some measure of error using $|h_j(x)|$, $\max\{g_i(x), 0\}$, or similar quantities are used
- The **true** distance to the feasible region is almost never computable. (**Forward error**)

Backward Error: $\mathrm{Res}(x) := [x_{22}^2 + (x_{12} - x_{33})^2 + \max\{-\lambda_{\min}(x), 0\}^2]^{1/2}$
Forward Error: $\mathrm{dist}\,(x, \mathrm{Feas})$.

Key point

Forward error $\neq O($**Backward Error**$)$

- The same phenomenon happens for optimal sets: small KKT residual $\not\Rightarrow$ the point is close to the optimal set.

What next?

Error bounds provide relations between **Forward error** and **Backward error**.

Feasibility problems over convex cones

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

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

- $\mathcal{K}$: 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)

Motivation

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

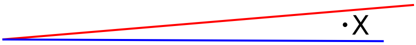
$$\operatorname{dist}(x, \mathcal{L} + a) = \inf\{\|x - y\| \mid y \in \mathcal{L} + a\}$$

$$\operatorname{dist}(x, \mathcal{K}) = \inf\{\|x - y\| \mid y \in \mathcal{K}\}$$

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

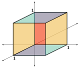
Fundamental question

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



- **Backward error**: $\operatorname{dist}(x, \mathcal{L} + a) + \operatorname{dist}(x, \mathcal{K})$
- **Forward error**: $\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{K})$

Hoffman's Lemma

Polyhedral set: a set that can be writen as the set of solutions of a *finite* number of linear inequalities. 

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

Theorem (Hoffman's Lemma '52)

If $\mathcal{K}$ is polyhedral, there is a constant $\kappa > 0$ such that

$$\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{K}) \leq \kappa \operatorname{dist}(x, \mathcal{L} + a) + \kappa \operatorname{dist}(x, \mathcal{K}), \quad \forall x \in \mathcal{E}.$$

Application to Linear Programming

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

- $\mathbb{R}^n_+$: nonnegative orthant.
- $\mathrm{Feas} = \{x \mid Ax = b, x \in \mathbb{R}^n_+\}$.

$$\mathrm{Res}(x) := \|Ax - b\| + \sum_{i=1}^{n} \max(-x_i, 0).$$

Because of Hoffman's Lemma:

$$\mathrm{dist}\,(x, \mathrm{Feas}) \leq \kappa \mathrm{Res}(x).$$

LPs are nice!

In LP, **Forward error** $= O($**Backward error**$)$

Application to Linear Programming - Optimal sets

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

- θ: optimal value
- $\mathrm{Opt} = \{x \mid c^T x = \theta, Ax = b, x \in \mathbb{R}^n_+\}$.

$$\mathrm{Res}_{\mathrm{opt}}(x) := \|c^T x - \theta\| + \|Ax - b\| + \sum_{i=1}^{n} \max(-x_i, 0).$$

Because of Hoffman's Lemma:

$$\mathrm{dist}\,(x, \mathrm{Opt}) \leq \kappa(\mathrm{Res}_{\mathrm{opt}}(x)).$$

LPs are nice!

Even for optimal sets we have **Forward error** $= O($**Backward Error**$)$

Lipschitzian error bound

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

Definition (Lipschitzian error bound)

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

$$\operatorname{dist}(x, C) \leq \theta_B(\operatorname{dist}(x, C_1) + \operatorname{dist}(x, C_2)) \quad \forall\, x \in B.$$

If θ_B is the same for all B, the bound is **global**.

Some known results:

- $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset \Rightarrow$ local Lipschitzian
- C_1, C_2 are polyhedral $\Rightarrow$ global Lipschitzian (Hoffman's Lemma)
- C_1 is polyhedral and $C_1 \cap (\operatorname{ri} C_2) \neq \emptyset \Rightarrow$ local Lipschitzian

Consequences to conic programming

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \in \mathcal{K} \end{aligned}$$

- $\mathcal{K}$: closed convex cone.
- $\mathrm{Feas} = \{x \mid Ax = b, x \in \mathcal{K}\}$.
- Slater's condition: $\mathrm{Feas} \cap \mathrm{ri}\, \mathcal{K} \neq \emptyset$

Define

$$\mathrm{Res}(x) := \|Ax - b\| + \mathrm{dist}\,(x, \mathcal{K})$$

If Slater's condition holds, for every bounded set B, $\exists \kappa_B$

$$\mathrm{dist}\,(x, \mathrm{Feas}) \leq \kappa_B \mathrm{Res}(x).$$

Under Slater's

Forward error $= O($**Backward error**$)$ over every fixed bounded set

Consequences to conic programming - optimal sets

$$\begin{array}{rl} \min_x & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{S}^n_+ \end{array}$$

- θ: optimal value
- $\mathrm{Opt} = \{x \mid c^T x = \theta, Ax = b, x \in \mathcal{S}^n_+\}$.
- Suppose **Slater's condition holds.**

$$\text{In general, } \mathrm{Opt} \cap \mathrm{ri}\, \mathcal{S}^n_+ = \emptyset$$

If x is primal optimal and s is dual optimal then

$$xs = 0$$

so $s \neq 0$ implies x is **not positive definite**.

Optimal sets are hard

Even under Slater, we may have **Forward error** $\neq O($**Backward Error**$)$

In conic linear programming...

- **For feasible regions**: Slater's condition holds $\Rightarrow$ **Forward error** $=$ O(**Backward error**) over every fixed bounded set
- **For optimal sets**: even under Slater's, **Forward error** and **Backward error** might be quite different.

Key point

We need error bounds that hold when Slater fails!

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

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

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.

Sturm's bound

S^n: $n \times n$ symmetric matrices.
$\mathcal{S}^n_+$: $n \times n$ positive semidefinite matrices.

Theorem (Sturm's Error Bound)

Suppose $(\mathcal{L} + a) \cap \mathcal{S}^n_+ \neq \emptyset$. *There exists* $\gamma \geq 0$ *such that for every bounded set* B*, there exists* κ_B *such that*

$$\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{S}^n_+) \leq \kappa_B (\operatorname{dist}(x, \mathcal{L} + a) + \operatorname{dist}(x, \mathcal{S}^n_+))^{(2^{-\gamma})}, \quad \forall\, x \in B$$

where $\gamma \leq \min\{n-1, \dim(\mathcal{L}^\perp \cap \{a\}^\perp), \dim \operatorname{span}(\mathcal{L} + a)\}$.

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

Consequence for optimal sets: if **strict complementarity holds**, over a fixed bounded set we have

$$\textbf{Forward error} = O(\sqrt{\textbf{Backward Error}})$$

Beyond Sturm's error bound

Today's goals

- Prove error bounds for general cones beyond $\mathcal{S}^n_+$
- Constraint qualifications are **forbidden**! 

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

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

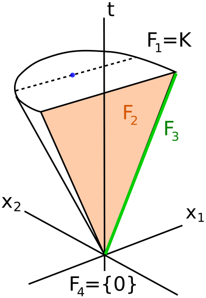
Review of faces

- $\mathcal{K}$: closed convex cone
- $\mathcal{F} \subseteq \mathcal{K}$: closed convex cone

Definition (Face of a cone)

$\mathcal{F}$ is a face of $\mathcal{K}$ $\Leftrightarrow$ if $x + y \in \mathcal{F}$, with $x, y \in \mathcal{K}$, then $x, y \in \mathcal{F}$.

If $\mathcal{F} \subseteq \mathcal{K}$ is a face, we write $\mathcal{F} \trianglelefteq \mathcal{K}$.



Ingredient 1 - Error bounds under a constraint qualification

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

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

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

- (a) $\mathcal{F} \cap (\mathcal{L} + a) = \mathcal{K} \cap (\mathcal{L} + a)$
- (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.

General strategy

Goal: We want to bound $\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{K})$ using $\operatorname{dist}(x, \mathcal{L} + a)$ and $\operatorname{dist}(x, \mathcal{K})$.

1. Find $\mathcal{F}$ such that
   a. $\mathcal{F} \cap (\mathcal{L} + a) = \mathcal{K} \cap (\mathcal{L} + a)$
   b. $(\operatorname{ri} \mathcal{F}) \cap (\mathcal{L} + a) \neq \emptyset$

   Therefore,

$$\operatorname{dist}(x, \mathcal{K} \cap (\mathcal{L} + a)) \leq \kappa_B(\operatorname{dist}(x, \mathcal{F}) + \operatorname{dist}(x, \mathcal{L} + a)), \quad \forall x \in B. \tag{1}$$

2. Upper bound $\operatorname{dist}(x, \mathcal{F})$ using $\operatorname{dist}(x, \mathcal{K})$ and $\operatorname{dist}(x, \mathcal{L} + a)$.
3. Plug the upper bound in (1).

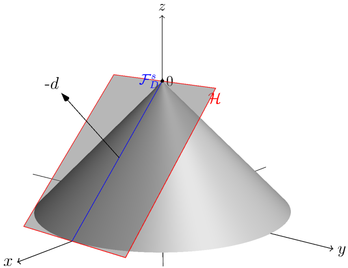
How to find $\mathcal{F}$?

We want $\mathcal{F}$ such that

- (a) $\mathcal{F} \cap (\mathcal{L} + a) = \mathcal{K} \cap (\mathcal{L} + a)$
- (b) $(\mathrm{ri}\,\mathcal{F}) \cap (\mathcal{L} + a) \neq \emptyset$

Idea:

1. Let $\mathcal{F}_1 = \mathcal{K}$ and $i \leftarrow 1$.
2. If $(\mathcal{L} + a) \cap \mathrm{ri}\,\mathcal{F}_i \neq \emptyset$, we are done.
3. If $(\mathcal{L} + a) \cap \mathrm{ri}\,\mathcal{F}_i = \emptyset$, we invoke a separation theorem.
    - There exists $z_i \in \mathcal{F}_i^* \setminus \mathcal{F}_i^\perp$ and $z_i \in \mathcal{L}^\perp \cap \{a\}^\perp$.
    - Let $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{z_i\}^\perp$ and $i \leftarrow i + 1$. Go to Step 2.



How to find $\mathcal{F}$? - 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* $(\operatorname{ri} \mathcal{F}_\ell) \cap (\mathcal{L} + a) \neq \emptyset$.

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.

Facial Reduction - Example

$$
\begin{array}{ll}
\sup\limits_{t,s} & -s \\
\text{s.t.} & \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \succeq 0
\end{array}
\tag{D}
$$

$$
\begin{aligned}
\mathcal{K} &= \mathcal{S}^3_+, \\
\mathcal{L} + a &= \left\{ \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \mid t, s \in \mathbb{R} \right\} \\
\mathcal{S}^3_+ \cap (\mathcal{L} + a) &= \left\{ \begin{pmatrix} t & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \begin{pmatrix} t & 1 \\ 1 & 1 \end{pmatrix} \succeq 0 \right\}.
\end{aligned}
$$

Facial Reduction - Continued

$$\sup_{t,s} \quad -s \qquad \text{(D)}$$

$$\text{s.t.} \quad \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \succeq 0$$

Let

$$z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\mathcal{S}^3_+ \cap (\mathcal{L} + a) \subseteq \{z\}^\perp.$$

So, the feasible region is contained in

$$\mathcal{S}^3_+ \cap \{z\}^\perp = \left\{ \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \begin{pmatrix} a & b \\ b & c \end{pmatrix} \succeq 0 \right\}$$

$\mathcal{F} = \mathcal{S}^3_+ \cap \{z\}^\perp$ is the face we want, since $(\mathcal{L} + a) \cap \operatorname{ri} \mathcal{F} \neq \emptyset$.

General strategy

Goal: We want to bound $\mathrm{dist}\,(x, (\mathcal{L}+a)\cap\mathcal{K})$ using $\mathrm{dist}\,(x, \mathcal{L}+a)$ and $\mathrm{dist}\,(x, \mathcal{K})$.

1. Find $\mathcal{F}$ such that
   a. $\mathcal{F}\cap(\mathcal{L}+a) = \mathcal{K}\cap(\mathcal{L}+a)$
   b. $(\mathrm{ri}\,\mathcal{F})\cap(\mathcal{L}+a) \neq \emptyset$

   Therefore,

   $$\mathrm{dist}\,(x, \mathcal{K}\cap(\mathcal{L}+a)) \leq \kappa_B(\mathrm{dist}\,(x, \mathcal{F}) + \mathrm{dist}\,(x, \mathcal{L}+a)), \quad \forall x \in B. \tag{1}$$

2. Upper bound $\mathrm{dist}\,(x, \mathcal{F})$ using $\mathrm{dist}\,(x, \mathcal{K})$ and $\mathrm{dist}\,(x, \mathcal{L}+a)$.
3. Plug the upper bound in (1).

Step 1 done!

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

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

we have:

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

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

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

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 - simplified

Suppose $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
Let

$$d(x) := \operatorname{dist}(x, \mathcal{L} + a) + \operatorname{dist}(x, \mathcal{K}).$$

Then, for every B, we have

$$\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{K}) \leq \kappa_B(d(x) + \varphi(d(x), M_B)), \quad \forall x \in B$$

where φ is a composition of **facial residual functions**.

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)
- If $\mathcal{K}$ is polyhedral, then $\psi_{\mathcal{F},z}(\epsilon, \|x\|) = \kappa\epsilon$ is a FRF, for some $\kappa > 0$.

Reminder:

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

implies

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

The case of symmetric cones - L'21

- $\mathcal{K}$: symmetric cone (psd matrices, second order cone and etc)
- Facial residual function (FRFs): $\psi_{\mathcal{F},z}(\epsilon, t) = \kappa\epsilon + \kappa\sqrt{\epsilon t}$

Suppose $(\mathcal{L} + a) \cap \mathcal{K} \neq \emptyset$. There exists $\gamma \geq 0$ such that for every bounded set B, there exists κ_B such that

$$\operatorname{dist}(x, (\mathcal{L} + a) \cap \mathcal{K}) \leq \kappa_B (\operatorname{dist}(x, \mathcal{L} + a) + \operatorname{dist}(x, \mathcal{K}))^{(2^{-\gamma})}, \quad \forall\, x \in B$$

where γ is the number of facial reduction steps.

Consequences for symmetric cone programming

$$\begin{array}{rl} \min\limits_{x} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \end{array}$$

For the feasible set:

- Under Slater: **Forward error** $= O($**Backward Error**$)$.
- Without Slater: **Forward error** $= O(($**Backward Error**$)^{2^{-\gamma}})$

For the optimal set:

- Strict complementarity holds: $x^* + s^* \in \operatorname{ri} \mathcal{K} \Leftrightarrow x^* \in \operatorname{ri}(\mathcal{K} \cap \{s^*\}^\perp)$
  - $\mathrm{Opt} = \{x \mid c^T x = \theta, Ax = b, x \in \mathcal{K}\}$ intersects $\operatorname{ri}(\mathcal{K} \cap \{s^*\}^\perp)$
  - Facial reduction finishes in 1 step.
- Under Strict complementarity:
  Forward error $= O(\sqrt{\textbf{Backward Error}})$

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

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

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 $\mathcal{K}^1$.

Amenable cones

Definition (Amenable cones)

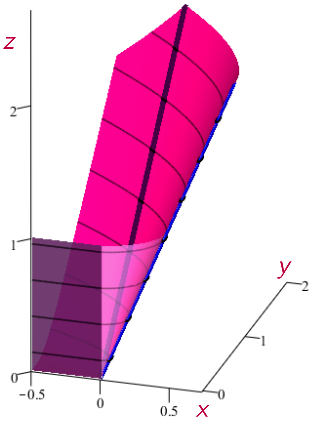
$\mathcal{K}$ is **amenable** if for every face $\mathcal{F}$ of $\mathcal{K}$ there is $\kappa > 0$ such that

$$\operatorname{dist}(x, \mathcal{F}) \leq \kappa \operatorname{dist}(x, \mathcal{K}), \qquad \forall x \in \operatorname{span} \mathcal{F}.$$

- Symmetric cones (e.g., PSD cone) are amenable ($\kappa = 1$)
- Polyhedral cones are amenable
- Strictly convex cones are amenable. (p-cones, second order cones and so on)
- $\mathcal{K}_1, \mathcal{K}_2 \Rightarrow$ FRFs of $\mathcal{K}_1 \times \mathcal{K}_2$ are sums of FRFs of $\mathcal{K}_1$ and $\mathcal{K}_2$.

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

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

1. Applications to entropy optimization, logistic regression, geometric programming and etc.
2. 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)

Error bounds for the exponential cone - LLP'20

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

Four types of error bounds are possible:

- Lipschitzian error bound
- Hölderian error bound with exponent $1/2$
- **Entropic error bound**: for every bounded set B, there exists $\kappa_B > 0$

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

- **Logarithmic 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**.

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

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:

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

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.

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

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:

T. Liu and L.
Convergence analysis under consistent error bounds
arXiv:2008.12968

L; Vera Roshchina and James Saunderson
Amenable cones are particularly nice
arXiv:2011.07745

Amenable cones

Definition (Amenable cones)

$\mathcal{K}$ is **amenable** if for every face $\mathcal{F}$ of $\mathcal{K}$ there is $\kappa > 0$ such that

$$\mathrm{dist}\,(x, \mathcal{F}) \leq \kappa \mathrm{dist}\,(x, \mathcal{K}), \qquad \forall x \in \mathrm{span}\,\mathcal{F}.$$

- Symmetric cones (e.g., PSD cone) are amenable ($\kappa = 1$)
- Polyhedral cones are amenable
- Strictly convex cones are amenable. (p-cones, second order cones and so on)
- Amenability is preserved under linear isomorphism and direct products

Facial exposedness

$$\mathcal{F} \text{ is a face of } \mathcal{K} \overset{\text{def}}{\iff} \mathcal{F} \trianglelefteq \mathcal{K}$$

$$\mathcal{K}^* := \{y \mid \langle y, x\rangle \geq 0, \forall x \in \mathcal{K}\}$$

1. Projectionally exposed cone $\overset{\text{def}}{\iff}$ $\forall \mathcal{F} \trianglelefteq \mathcal{K}$ there exists a projection such that $P\mathcal{K} = \mathcal{F}$.
2. Amenable cones $\overset{\text{def}}{\iff}$ for every face $\mathcal{F}$ of $\mathcal{K}$ there is $\kappa > 0$ such that

$$\operatorname{dist}(x, \mathcal{F}) \leq \kappa \operatorname{dist}(x, \mathcal{K}), \qquad \forall x \in \operatorname{span} \mathcal{F}.$$

3. Nice cone $\overset{\text{def}}{\iff}$ $\forall \mathcal{F} \trianglelefteq \mathcal{K}, \quad \mathcal{F}^* = \mathcal{K}^* + \mathcal{F}^\perp$.
4. Facially exposed cone $\overset{\text{def}}{\iff}$
$\forall \mathcal{F} \trianglelefteq \mathcal{K}, \quad \exists z \in \mathcal{K}$, s.t. $\mathcal{F} = \mathcal{K} \cap \{z\}^\perp$.

Comparison of exposedness properties

Known results:

- Facially exposed $\Leftarrow$ Nice $\Leftarrow$ **Amenable** $\overset{\text{EPBR}}{\Leftarrow}$ Projectionally exposed.
- $\dim \mathcal{K} \leq 3$: Facially exposed $\Leftrightarrow$ Projectionally exposed (Barker and Poole, SIADM'87)
- There exists a 4D cone that is facially exposed but not nice (Vera, SIOPT'14).

New results (see LRS'20):

- There exists a 4D cone that is nice but not amenable
- In dimension 4 or less: Amenable $\Leftrightarrow$ Projectionally exposed.

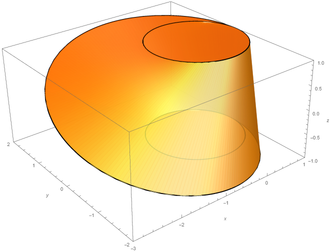


Figure: A 3D slice of a 4D convex cone that is nice but not amenable

Hyperbolicity cone

Let

- $p : \mathbb{R}^n \to \mathbb{R}$: homogenous polynomial
- $e \in \mathbb{R}^n$, with $p(e) > 0$

Hyperbolic polynomial

if for every $x \in \mathbb{R}^n$

$$t \mapsto p(te - x)$$

has only real roots, then p is **hyperbolic** along e.

For $x \in \mathbb{R}^n$, denotes the roots of

$$t \mapsto p(te - x)$$

by $\lambda_1(x), \ldots, \lambda_r(x)$.

Hyperbolicity cones

$\Lambda_+(p, e) := \{x \in \mathbb{R}^n \mid \lambda_i(x) \geq 0, i = 1, \ldots, r\}$.

Example

Let

- $p(X) : \mathcal{S}^n \to \mathbb{R}$, $p(X) = \det X$.
- $e = I_n$.

The roots of

$$t \mapsto p(tI_n - X) = \det(tI_n - X)$$

are the eigenvalues of X.

$$\Lambda_+(p, e) = \mathcal{S}^n_+.$$

Some history

- Studied in the 50's by Gårding in the context of partial differential equations.
- Güler brought them to attention of optimizers in 97.
    - $-\log p$ is a self-concordant barrier for the interior of $\Lambda_+(p, e)$.
- Renegar proved key results on the structure of $\Lambda_+(p, e)$ in 2005.

Some classes of cones

More general	Hyperbolicity cone
	Homogeneous cone
	Symmetric cone
	PSD cone
	Second order cone
Less general	$\mathbb{R}^n_+$

- Example of cone that is not a hyperbolicity cone: exponential cone
- Renegar proved that hyperbolicity cones are facially exposed.

Some classes of cones

	Hyperbolicity cone
Slice of a PSD cone (**spectrahedral**)	Homogeneous cone Symmetric cone PSD cone Second order cone $\mathbb{R}^n_+$

Spectrahedral cone

$\mathcal{K}$ is spectrahedral $\overset{\text{def}}{\Longleftrightarrow}$ $A(\mathcal{K}) = \mathcal{S}^n_+ \cap V$ holds for some injective linear map A, subspace V and n.

Lax conjecture

Spectrahedral cone

$\mathcal{K}$ is spectrahedral $\overset{\text{def}}{\iff}$ $A(\mathcal{K}) = \mathcal{S}^n_+ \cap V$ holds for some injective linear map A, subspace V and n.

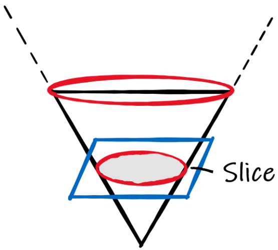
Generalized Lax Conjecture

Is every hyperbolicity cone spectrahedral?

Recent results on amenability

A few results (L, Roshchina and Saunderson)

- Hyperbolicity cones and spectrahedral cones are amenable.
- Amenability is preserved by intersections and taking slices.
- A cone constructed from an amenable compact convex set is amenable.



Open questions

- Is there an amenable cone that is not projectionally exposed? (dim $\mathcal{K} \geq 5$ must hold!)
- Which cones are projectionally exposed?

L, V. Roshchina and J. Saunderson
Amenable cones are particularly nice.
arxiv:2011.07745

L, V. Roshchina and J. Saunderson
Hyperbolicity cones are amenable.
arxiv:2102.06359

Thank you!

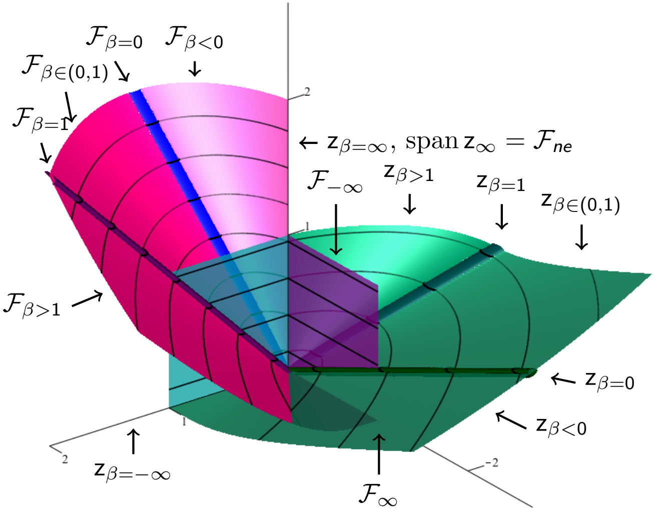
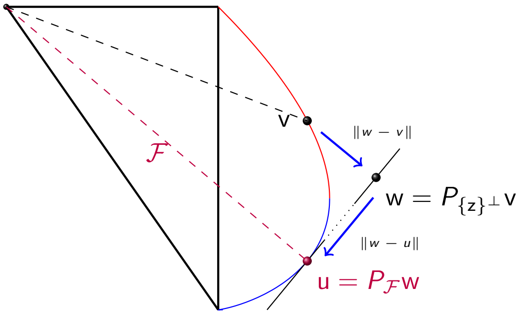


Figure: The exponential cone, its faces and exposing vectors

FRFs without projection - LLP'21



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

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