Projection and presolve in MOSEK: exponential and power cones

ISMP 2018

Henrik A. Friberg

www.mosek.com

New cones in MOSEK

Friberg, Henrik A. (2017). *Power cones in second-order cone form and dual recovery*. SIAM Conference on Optimization. www.mosek.com/resources/presentations.

For rational numbers $\alpha_1, \ldots, \alpha_k \geq 0$:

$$K_{\mathrm{pow}(\alpha)} = \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T\alpha},\ x_1, \ldots x_k \geq 0\}$$
$$=$$
$$P\left(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots\right).$$

New cones in MOSEK

Friberg, Henrik A. (2017). *Power cones in second-order cone form and dual recovery.* SIAM Conference on Optimization. www.mosek.com/resources/presentations.

For rational numbers $\alpha_1, \ldots, \alpha_k \geq 0$:

$$K_{\mathrm{pow}(\alpha)} = \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T\alpha},\ x_1, \ldots x_k \geq 0\}$$

$$=$$

$$P\left(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots\right). \qquad ☹$$

New cones in MOSEK

Friberg, Henrik A. (2017). *Power cones in second-order cone form and dual recovery.* SIAM Conference on Optimization. www.mosek.com/resources/presentations.

For rational numbers $\alpha_1, \ldots, \alpha_k \geq 0$:

$$K_{\mathrm{pow}(\alpha)} = \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T\alpha},\ x_1, \ldots x_k \geq 0\}$$
$$=$$
$$P\left(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots\right). \qquad ☹$$

<u>In MOSEK 9:</u>

- $K_{\mathrm{pow}(\alpha,1-\alpha)} = \{x_1^{\alpha} x_2^{1-\alpha} \geq \|z\|_2,\ x_1, x_2 \geq 0\}$, parametrized by a real number $0 < \alpha < 1$.
- $K_{\mathrm{exp}} = \mathrm{cl}\{t \geq s\exp(r/s),\ s > 0\}$.
- The corresponding dual cones $K^*_{\mathrm{pow}(\alpha,1-\alpha)}$ and K^*_{exp}.

Exponential cone examples

1. Exponential.

$$t \geq \exp(x) \iff (t, 1, x) \in K_{\exp}.$$

Exponential cone examples

1. Exponential.

$$t \geq a^x \iff (t, 1, x\log(a)) \in K_{\exp}.$$

Exponential cone examples

1. Exponential.

$$t \geq a_1^{x_1} a_2^{x_2} \cdots a_n^{x_n} \iff \left(t, 1, \sum_{j=1}^{n} x_j \log(a_j) \right) \in K_{\exp}.$$

Exponential cone examples

❷ $\{t \leq \log(x)\} = \{(x,1,t) \in K_{\exp}\}$.

❸ $\{t \geq x\log(x/y)\} = \{(y,x,-t) \in K_{\exp}\}$.

❹ $\left\{t \geq (\log x)^2,\ 0 < x \leq 1\right\} = \left\{(\frac{1}{2},t,u) \in \mathcal{Q}_r^3,\ (x,1,u) \in K_{\exp},\ x \leq 1\right\}$.

❺ $\{t \leq \log\log x,\ x > 1\} = \{(u,1,t) \in K_{\exp},\ (x,1,u) \in K_{\exp}\}$.

❻ $\left\{t \geq (\log x)^{-1},\ x > 1\right\} = \left\{(u,t,\sqrt{2}) \in \mathcal{Q}_r^3,\ (x,1,u) \in K_{\exp}\right\}$.

❼ $\left\{t \leq \sqrt{\log x},\ x > 1\right\} = \left\{(\frac{1}{2},u,t) \in \mathcal{Q}_r^3,\ (x,1,u) \in K_{\exp}\right\}$.

❽ $\left\{t \leq \sqrt{x\log x},\ x > 1\right\} = \left\{(x,u,\sqrt{2}t) \in \mathcal{Q}_r^3,\ (x,1,u) \in K_{\exp}\right\}$.

❾ $\{t \geq x\exp(x),\ x \geq 0\} = \left\{(\frac{1}{2},u,x) \in \mathcal{Q}_r^3,\ (t,x,u) \in K_{\exp}\right\}$.

Exponential cone examples

10. Log-sum-exponential.

$$t \quad \geq \quad \log(e^{x_1} + \ldots + e^{x_n})$$

Exponential cone examples

10. Log-sum-exponential.

$$
\begin{array}{rcl}
t & \geq & \log(e^{x_1} + \ldots + e^{x_n}) \\
e^t & \geq & e^{x_1} + \ldots + e^{x_n} \qquad ☹
\end{array}
$$

Exponential cone examples

10. Log-sum-exponential.

$$\begin{array}{rcll} t & \geq & \log(e^{x_1} + \ldots + e^{x_n}) & \\ e^t & \geq & e^{x_1} + \ldots + e^{x_n} & ☹ \\ 1 & \geq & e^{x_1 - t} + \ldots + e^{x_n - t} & ☺ \end{array}$$

Exponential cone examples

10. Log-sum-exponential.

$$\begin{array}{rcll} t & \geq & \log(e^{x_1} + \ldots + e^{x_n}) & \\ e^t & \geq & e^{x_1} + \ldots + e^{x_n} & ☹ \\ 1 & \geq & e^{x_1 - t} + \ldots + e^{x_n - t} & ☺ \end{array}$$

Geometric programming in conic form:

$$\begin{array}{ll} \inf & x + y^2 z \\ \text{s.t.} & 0.1\sqrt{x} + 2y^{-1} \leq 1, \\ & z^{-1} + yx^{-2} \leq 1, \end{array} \quad \leftrightarrow \quad \begin{array}{ll} \inf & t \\ \text{s.t.} & \log(e^{u} + e^{2v+w}) \leq t, \\ & \log(e^{0.5u + \log(0.1)} + e^{-v+\log(2)}) \leq 0, \\ & \log(e^{-w} + e^{v-2u}) \leq 0, \end{array}$$

where $(x, y, z) = (e^u, e^v, e^w)$.

More information

Usage

- MOSEK Modeling Cookbook.
- Fri, 15:15. Michał Adamaszek: *Exponential cone in MOSEK: overview and applications*.

Implementation details

- Wed, 8:30. Joachim Dahl: *Extending MOSEK with exponential cones*.

Details for all of MOSEK 9

- Wed, 15:15. Erling Andersen: *MOSEK version 9*.

The curious case of error measuring

```
Interior-point solution summary
  Problem status  : PRIMAL_AND_DUAL_FEASIBLE
  Solution status : OPTIMAL

  Primal.  obj: 7.4390660847e-02
           nrm: 1e+00
           Viol.  con: 6e-09    var: 0e+00    cones: 4e-09

  Dual.    obj: 7.4390675795e-02
           nrm: 3e-01
           Viol.  con: 1e-19    var: 8e-09    cones: 0e+00
```

The curious case of error measuring

Error of $x = (0,\ 10^8,\ 1)$ in constraint $2x_1x_2 \geq |x_3|$?

The curious case of error measuring

Error of $x = (0,\ 10^8,\ 1)$ in constraint $2x_1x_2 \geq |x_3|$?

- $f(x) = |x_3| - 2x_1x_2 \leq 0$. Error $[f(x)]_+ = 1$.

The curious case of error measuring

Error of $x = (0,\ 10^8,\ 1)$ in constraint $2x_1x_2 \geq |x_3|$?

- $f(x) = |x_3| - 2x_1x_2 \leq 0$. Error $[f(x)]_+ = 1$.
- $f(x) = |x_3|/x_1 - 2x_2 \leq 0$. Error $[f(x)]_+ = \texttt{Inf}$.
- $f(x) = |x_3|/x_2 - 2x_1 \leq 0$. Error $[f(x)]_+ = 1e{-}8$.

The curious case of error measuring

Error of $x = (0,\ 10^8,\ 1)$ in constraint $2x_1x_2 \geq |x_3|$?

- $f(x) = |x_3| - 2x_1x_2 \leq 0$. Error $[f(x)]_+ = 1$.
- $f(x) = |x_3|/x_1 - 2x_2 \leq 0$. Error $[f(x)]_+ = \texttt{Inf}$.
- $f(x) = |x_3|/x_2 - 2x_1 \leq 0$. Error $[f(x)]_+ = 1e{-}8$.
- $\mathrm{dist}(x, \mathcal{Q}_r^3) = 5e{-}9$.

The curious case of error measuring

Error of $x = (0,\ 10^8,\ 1)$ in constraint $2x_1x_2 \geq |x_3|$?

- $f(x) = |x_3| - 2x_1x_2 \leq 0$. Error $[f(x)]_+ = 1$.
- $f(x) = |x_3|/x_1 - 2x_2 \leq 0$. Error $[f(x)]_+ = \texttt{Inf}$.
- $f(x) = |x_3|/x_2 - 2x_1 \leq 0$. Error $[f(x)]_+ = 1e{-}8$.
- $\mathrm{dist}(x, \mathcal{Q}_r^3) = 5e{-}9$.

The power and exponential cones are also representation sensitive:

$$x_1^{0.3333}x_2^{0.6666} \geq ||z||_2 \quad \Longleftrightarrow \quad x_1^1x_2^2 \geq ||z||_2^3$$

$$y \geq exp(x) \quad \Longleftrightarrow \quad x \leq log(y)$$

This sensitivity is a well-known caveat of forward error.
Projection is an example of backwards error.

The curious case of error measuring

```
Interior-point solution summary
  Problem status  : PRIMAL_AND_DUAL_FEASIBLE
  Solution status : OPTIMAL

  Primal.  obj: 7.4390660847e-02
           nrm: 1e+00
           Viol.  con: 6e-09    var: 0e+00    cones: 4e-09

  Dual.    obj: 7.4390675795e-02
           nrm: 3e-01
           Viol.  con: 1e-19    var: 8e-09    cones: 0e+00
```

Variable domains are measured with backwards error:
$\| \operatorname{dist}(x_1, \mathcal{K}_1),\ \operatorname{dist}(x_2, \mathcal{K}_2),\ \ldots \|_\infty$.

In need of projections!

$$
\boxed{\begin{aligned} \operatorname{dist}(\tilde{x}, \mathcal{K}) &= \min_{x \in \mathcal{K}} \|x - \tilde{x}\| \\ [\tilde{x}]_{\mathcal{K}} &= \arg\min_{x \in \mathcal{K}} \|x - \tilde{x}\| \end{aligned}}
$$

What is the hype about?

- Set membership conditions ($x \in \mathcal{K}$).
- Representation-free error measures.
- Maximal separating hyperplanes.
- First-order methods for feasible point searches (e.g., looking for specific properties).

...basically a useful low cost operation (time+memory).

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Trivial example: All scalars are uniquely decomposable as

$$v_0 \;=\; [v_0]_+ \;+\; [v_0]_-,$$

where $[\bullet]_+ = [\bullet]_{\mathbb{R}_+} = \max(0, \bullet)$,
and $[\bullet]_- = [\bullet]_{\mathbb{R}_-} = \min(0, \bullet)$.

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Dual cone projection:

$$[v_0]_{\mathcal{K}^*} = -\big[-v_0\big]_{-\mathcal{K}^*} = -\big[-v_0\big]_{\mathcal{K}^\circ}.$$

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Reflection (intrepid projection for obtuse cones):

$$\mathrm{Ref}_{\mathcal{K}}(v_0) = [v_0]_{\mathcal{K}} \;-\; [v_0]_{\mathcal{K}^\circ}.$$

Separation

For nonempty closed convex cones,

$$\mathcal{K} = \{x \mid a^T x \leq 0,\ \forall a \in \mathcal{K}^\circ\}.$$

- Separators of $\hat{x} \notin \mathcal{K}$ are points of $\{a \in \mathcal{K}^\circ \mid a^T \tilde{x} > 0\}$.

Separation

For nonempty closed convex cones,

$$\mathcal{K} = \{x \mid a^T x \leq 0, \forall a \in \mathcal{K}^\circ\}.$$

- Separators of $\hat{x} \notin \mathcal{K}$ are points of $\{a \in \mathcal{K}^\circ \mid a^T \tilde{x} > 0\}$.

Gradient separator

For positively homogeneous convex functions, the cone

$$\mathcal{K} = \{x \mid f(x) \leq 0\},$$

has separator $a = \nabla f(\hat{x})$ for $\hat{x} \notin \mathcal{K}$.

Lubin, Miles (2017). "Mixed-integer convex optimization: outer approximation algorithms and modeling power". PhD thesis. Massachusetts Institute of Technology.

Separation

For nonempty closed convex cones,

$$\mathcal{K} = \{x \mid a^T x \leq 0,\ \forall a \in \mathcal{K}^\circ\}.$$

- Separators of $\hat{x} \notin \mathcal{K}$ are points of $\{a \in \mathcal{K}^\circ \mid a^T \tilde{x} > 0\}$.
- The maximal separator solves $\max\limits_{a \in \mathcal{K}^\circ,\ \|a\|_2 \leq 1} a^T \hat{x}$.

Separation

For nonempty closed convex cones,

$$\mathcal{K} = \{x \mid a^T x \leq 0,\ \forall a \in \mathcal{K}^\circ\}.$$

- Separators of $\hat{x} \notin \mathcal{K}$ are points of $\{a \in \mathcal{K}^\circ \mid a^T \tilde{x} > 0\}$.
- The maximal separator solves $\max\limits_{a \in \mathcal{K}^\circ,\ \|a\|_2 \leq 1} a^T \hat{x}$.
- Its dual problem is $\min\limits_{x \in \mathcal{K}} \|x - \hat{x}\|_2$.
- Maximal separator is dual solution to projection problem.

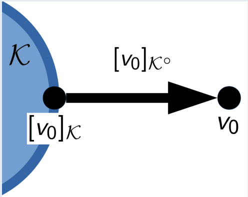
All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Maximal separation:

$$\mathrm{maxsep}_{\mathcal{K}}(v_0) = \frac{[v_0]_{\mathcal{K}^\circ}}{\|[v_0]_{\mathcal{K}^\circ}\|_2}$$



Projection theory

Moreau's decomposition theorem

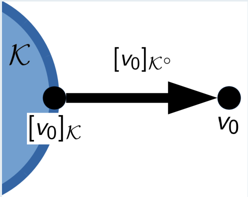
All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Maximal separation:

$$\operatorname{maxsep}_{\mathcal{K}}(v_0) = \frac{[v_0]_{\mathcal{K}^\circ}}{\|[v_0]_{\mathcal{K}^\circ}\|_2}$$



Differs from gradient separator when $\nabla f(v_0) \neq \nabla f\big([v_0]_{\mathcal{K}}\big)$.

Declaring infeasibility in presolve

Domain propagation

$$K_{\text{pow}(\alpha,1-\alpha)} = \{(x_1, x_2, z) \mid x_1^{\alpha} x_2^{1-\alpha} \geq \|z\|_2,\ x_1, x_2 \geq 0\}.$$

$$U = \sup_{l^x \leq x \leq u^x} x_1^{\alpha} x_2^{1-\alpha}, \qquad L = \inf_{l^z \leq z \leq u^z} \|z\|_2$$

If $U \geq L$, propagate bounds. Otherwise, declare infeasibility.

$$K_{\text{pow}(\alpha,1-\alpha)} = \{(x_1, x_2, z) \mid x_1^{\alpha} x_2^{1-\alpha} \geq \|z\|_2,\ x_1, x_2 \geq 0\}.$$

$$U = \sup_{l^x \leq x \leq u^x} x_1^{\alpha} x_2^{1-\alpha}, \qquad L = \inf_{l^z \leq z \leq u^z} \|z\|_2$$

If $U \geq L$, propagate bounds. Otherwise, declare infeasibility.

WHERE ARE MY DUAL VARIABLES TO CONSTRUCT THE INFEASIBILITY CERTIFICATE!?

Declaring infeasibility in presolve

Domain propagation

$$K_{\mathrm{pow}(\alpha,1-\alpha)} = \{(x_1, x_2, z) \mid x_1^{\alpha} x_2^{1-\alpha} \geq \|z\|_2,\ x_1, x_2 \geq 0\}.$$

$$U = \sup_{l^x \leq x \leq u^x} x_1^{\alpha} x_2^{1-\alpha}, \qquad L = \inf_{l^z \leq z \leq u^z} \|z\|_2$$

If $U \geq L$, propagate bounds. Otherwise, declare infeasibility.

Short answer: Maximal separator via projection.

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (in the 2-norm).

Symmetric cones

- $[v_0]_{\mathcal{K}}$ is well-known.
- $[v_0]_{\mathcal{K}^\circ} = v_0 - [v_0]_{\mathcal{K}}$ is simple, but not numerically optimal.

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (in the 2-norm).

Exponential and power cone

1. Mathematical foundation.
2. Pseudocode implementation (algorithmic idea + complexity).
3. Prototype implementation (behavior + edge cases).
4. Final implementation.

All matrices/vectors are uniquely decomposable as

$$v_0 \ = \ [v_0]_{\mathcal{K}} \ + \ [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Moreau conditions:

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \langle v_p, v_d \rangle = 0.$$

That is, conic KKT conditions for $\min\limits_{v_p \in \mathcal{K}} \frac{1}{2} \| v_p - v_0 \|_2^2$.

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Moreau conditions:

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \langle v_p, v_d \rangle = 0.$$

$$v_d \in \mathcal{K}^\circ \cap v_p^\perp.$$

All matrices/vectors are uniquely decomposable as

$$v_0 \;=\; [v_0]_{\mathcal{K}} \;+\; [v_0]_{\mathcal{K}^\circ},$$

for all nonempty, closed, convex cones $\mathcal{K}$ (and in any norm).

Moreau conditions:

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \langle v_p, v_d \rangle = 0.$$

$$v_d \in \mathcal{K}^\circ \cap v_p^\perp.$$

$$v_d \in N_{\mathcal{K}}(v_p).$$

Two strong results for normal cones:

1. Let $S_1 \cap S_2$ be constraint qualified; e.g., Slater. Then

$$N_{S_1 \cap S_2}(x) = N_{S_1}(x) + N_{S_2}(x)$$

2. Let $S = \{x \in \mathbb{R}^n : g(x) \le 0\}$ (proper convex function). Then

$$N_S(x) = \begin{cases} \mathbb{R}_+ \partial g(x) & \text{if } g(x) = 0, \\ \{0\} & \text{if } g(x) < 0, \\ \emptyset & \text{if } g(x) > 0. \end{cases}$$

Exponential cone projection

Primal cone

$$K_{\exp} = \mathrm{cl}([K_{\exp}]_{++}) = [K_{\exp}]_{++} \cup [K_{\exp}]_0,$$

in terms of

$$\begin{aligned}
[K_{\exp}]_{++} &= \left\{ \begin{pmatrix} t \\ s \\ r \end{pmatrix} \in \mathbb{R}^3 \mid s > 0, \qquad t \geq s \exp\left(\frac{r}{s}\right) \right\}, \\
[K_{\exp}]_0 &= \left\{ \begin{pmatrix} t \\ s \\ r \end{pmatrix} \in \mathbb{R}^3 \mid s = 0, \qquad t \geq 0,\ r \leq 0 \right\}.
\end{aligned}$$

Exponential cone projection

Polar cone

$$K_{\exp}^{\circ} = \mathrm{cl}([K_{\exp}^{\circ}]_{++}) = [K_{\exp}^{\circ}]_{++} \cup [K_{\exp}^{\circ}]_{0},$$

in terms of

$$\begin{aligned}
[K_{\exp}^{\circ}]_{++} &= \left\{ \begin{pmatrix} t \\ s \\ r \end{pmatrix} \in \mathbb{R}^3 \mid r > 0,\ (-\mathbf{e})t \geq r \exp\left(\frac{s}{r}\right) \right\}, \\
[K_{\exp}^{\circ}]_{0} &= \left\{ \begin{pmatrix} t \\ s \\ r \end{pmatrix} \in \mathbb{R}^3 \mid r = 0,\ (-\mathbf{e})t \geq 0,\ s \leq 0 \right\}.
\end{aligned}$$

Exponential cone projection

Step 1. Presolving edge cases away

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \boxed{v_p^T v_d = 0}.$$

What if $t_p t_d = s_p s_d = r_p r_d = 0$?

- Case $t_d = 0$ (implies $r_d = 0$) and $s_p = 0$. The Moreau system reduces to $t_0 = t_p \geq 0$, $s_0 = s_d \leq 0$ and $r_0 = r_p \leq 0$.
- ...

Exponential cone projection

Step 1. Presolving edge cases away

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \boxed{v_p^T v_d = 0}.$$

1. If $v_0 \in K_{\exp}$, then $v_p = v_0$ and $v_d = 0$.
2. If $v_0 \in K_{\exp}^\circ$, then $v_p = 0$ and $v_d = v_0$.
3. If $r_0, s_0 \leq 0$, then $v_p = \big([t_0]_+,\ 0,\ r_0\big)$ and $v_d = \big([t_0]_-,\ s_0,\ 0\big)$.

Literature

Parikh, Neal and Stephen Boyd (2013). “Proximal Algorithms”. In: *Foundations and Trends in Optimization* 1.3, pp. 123–231.

Exponential cone projection

Step 1. Presolving edge cases away

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \boxed{v_p^T v_d = 0}.$$

1. If $v_0 \in K_{\exp}$, then $v_p = v_0$ and $v_d = 0$.
2. If $v_0 \in K_{\exp}^\circ$, then $v_p = 0$ and $v_d = v_0$.
3. If $r_0, s_0 \leq 0$, then $v_p = \big([t_0]_+,\, 0,\, r_0\big)$ and $v_d = \big([t_0]_-,\, s_0,\, 0\big)$.

Case covered

1. $[v_0 \in K_{\exp}]$ or $[v_0 \in K_{\exp}^\circ]$ or $[r_0, s_0 \leq 0]$.
2. $[v_p^T v_d = 0] \Rightarrow [t_p t_d = s_p s_d = r_p r_d = 0]$.

   Sufficient conditions: $[t_p = 0]$ or $[t_d = 0]$ or $[s_p = 0]$or $[r_d = 0]$
   or $[s_d \leq 0$ and $r_p \leq 0]$.

Exponential cone projection

Step 1. Presolving edge cases away

$$v_0 = v_p + v_d, \quad v_p \in \mathcal{K}, \quad v_d \in \mathcal{K}^\circ, \quad \boxed{v_p^T v_d = 0}.$$

1. If $v_0 \in K_{\exp}$, then $v_p = v_0$ and $v_d = 0$.
2. If $v_0 \in K_{\exp}^\circ$, then $v_p = 0$ and $v_d = v_0$.
3. If $r_0, s_0 \leq 0$, then $v_p = \big([t_0]_+, 0, r_0\big)$ and $v_d = \big([t_0]_-, s_0, 0\big)$.

Case remaining

1. $[v_0 \notin K_{\exp}]$ and $[v_0 \notin K_{\exp}^\circ]$ and $[r_0 > 0$ or $s_0 > 0]$.
2. $[v_p^T v_d = 0] \not\Rightarrow [t_p t_d = s_p s_d = r_p r_d = 0]$.

   Necessary: $[t_p > 0]$ and $[t_d < 0]$ and $[s_p > 0]$ and $[r_d > 0]$ and $[s_d > 0$ or $r_p > 0]$.

Exponential cone projection

Step 2. Simplify Moreau for remaining case

The Moreau conditions are equivalent to the system

$$t_0 = t_p + t_d, \quad s_p > 0, \quad r_d > 0,$$

given definitions

$$v_p = (t_p, s_p, r_p) = (\exp(\rho), 1, \rho)\, s_p, \qquad s_p = \frac{(\rho - 1) r_0 + s_0}{\rho^2 - \rho + 1},$$

$$v_d = (t_d, s_d, r_d) = (-\exp(-\rho), 1 - \rho, 1)\, r_d, \quad r_d = \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1},$$

depending solely on the primal ratio, $\rho = \dfrac{r_p}{s_p} = 1 - \dfrac{s_d}{r_d}$.

Exponential cone projection

Step 2. Simplify Moreau for remaining case

The Moreau conditions are equivalent to the system

$$\boxed{t_0 = t_p + t_d, \quad s_p > 0, \quad r_d > 0,}$$

given definitions

$$v_p = (t_p, s_p, r_p) = (\exp(\rho), 1, \rho)\, s_p, \qquad s_p = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1},$$

$$v_d = (t_d, s_d, r_d) = (-\exp(-\rho), 1 - \rho, 1)\, r_d, \quad r_d = \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1},$$

depending solely on the primal ratio, $\rho = \dfrac{r_p}{s_p} = 1 - \dfrac{s_d}{r_d}$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1}\exp(\rho) - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1}\exp(-\rho) - t_0,$$

on the nonempty strict domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1} \exp(\rho) - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1} \exp(-\rho) - t_0,$$

on the nonempty strict domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign - $\boxed{\text{not locally convex}}$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1}\exp(\rho) - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1}\exp(-\rho) - t_0,$$

on the nonempty strict domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ \boxed{-\infty} & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \boxed{\infty} & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign - $\boxed{\text{not locally convex}}$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1}\exp(\rho) - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1}\exp(-\rho) - t_0,$$

on the nonempty strict domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign - not locally convex.

Take $(t_0, s_0, r_0) = (8, -8, 0.1)$, then $l = 801$ and $u = \infty$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1}\boxed{\exp(\rho)} - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1}\boxed{\exp(-\rho)} - t_0,$$

on the nonempty strict domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign - not locally convex.

Take $(t_0, s_0, r_0) = (8, -8, 0.1)$, then $\boxed{l = 801}$ and $u = \infty$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1) r_0 + s_0}{\rho^2 - \rho + 1} \boxed{\exp(\rho)} - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1} \boxed{\exp(-\rho)} - t_0,$$

on the nonempty $\boxed{\text{strict}}$ domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, if v_0 passes presolve, the function $h(\rho)$ is smooth, strictly increasing and changes sign - not locally convex.

Take $(t_0, s_0, r_0) = (8, -8, 0.1)$, then $\boxed{l = 801}$ and $u = \infty$.

Find a root of the function

$$h(\rho) = \frac{(\rho - 1)r_0 + s_0}{\rho^2 - \rho + 1}\boxed{\exp(\rho)} - \frac{r_0 - \rho s_0}{\rho^2 - \rho + 1}\boxed{\exp(-\rho)} - t_0,$$

on the nonempty $\boxed{\text{strict}}$ domain, $l < \rho < u$, given by

$$l = \begin{cases} 1 - s_0/r_0 & \text{if } r_0 > 0, \\ -\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad u = \begin{cases} r_0/s_0 & \text{if } s_0 > 0, \\ \infty & \text{otherwise.} \end{cases}$$

On this domain, $\boxed{\text{if } v_0 \text{ passes presolve}}$, the function $h(\rho)$ is smooth, strictly increasing and changes sign - not locally convex.

Take $(t_0, s_0, r_0) = (8, -8, 0.1)$, then $l = 801$ and $u = \infty$.

Exponential cone projection

In need of heuristics

Change v_0 such that the presolve rules apply.

1. If $s_0 > 0$, try increasing t_0 until the first rule applies:

$$\begin{aligned} \tilde{v}_p &= \big(s_0 \exp(r_0/s_0), s_0, r_0\big) \in K_{\exp} \\ \tilde{v}_d &= 0 \in K_{\exp}^{\circ} \end{aligned}$$

2. If $r_0 > 0$, try decreasing t_0 until the second rule applies:

$$\begin{aligned} \tilde{v}_d &= (-r_0 \exp(s_0/r_0 - 1), s_0, r_0) \in K_{\exp}^{\circ} \\ \tilde{v}_p &= 0 \in K_{\exp} \end{aligned}$$

3. One may always decrease s_0 and r_0 until the third rule applies:

$$\begin{aligned} \tilde{v}_p &= ([t_0]_+, 0, [r_0]_-) \in K_{\exp} \\ \tilde{v}_d &= ([t_0]_-, [s_0]_-, 0) \in K_{\exp}^{\circ} \end{aligned}$$

Exponential cone projection

Friberg, Henrik Alsing (2018). *Projection onto the exponential cone: a univariate root-finding problem*. To be published.

- These heuristics are all you need!

```
v0=[8,-8,0.01]
primal=[8,0,0]
polar =[-0,-8,0.01]
Moreau system errors: comp=1e-350. orth=1e-349 pexp=0. polar=0.
```

- Root can be lower and upper bounded (finite bracket).

Exponential cone projection

Prototype implementation (Julia language)

```
function proj_pexpcone(v0::Array{real,1})
    const t0,s0,r0 = v0

    vp,pdist    = projheu_pexpcone(v0)
    negvd,ddist = projheu_negdexpcone(v0)

    if !( (s0<=0 && r0<=0) || min(pdist,ddist)<=1e-12 )
        xl,xh = bracket(v0,pdist,ddist)
        rho   = rootsearch_ntinc(hfun,v0,xl,xh,0.5*(xl+xh))

        vp1,pdist1 = projsol_pexpcone(v0,rho)
        if (pdist1 <= pdist)
            vp,pdist=vp1,pdist1
        end

        negvd1,ddist1 = projsol_negdexpcone(v0,rho)
        if (ddist1 <= ddist)
            negvd,ddist = negvd1,ddist1
        end
    end
    return [vp,negvd]
end
```

Power cone projection

$$
\begin{aligned}
K_{\mathrm{pow}(\alpha)} &= \left\{ \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^k_+ \times \mathbb{R}^{n-k} \mid x^\alpha \geq \|z\|_2 \right\}, \\
K^\circ_{\mathrm{pow}(\alpha)} &= \left\{ \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^k_- \times \mathbb{R}^{n-k} \mid \alpha^{-\alpha}(-x)^\alpha \geq \|z\|_2 \right\},
\end{aligned}
$$

for a parameter vector $\alpha \in \mathbb{R}^k_{++}$ summing to $e^T\alpha = 1$.

Hien, Le Thi Khanh (2015). "Differential properties of Euclidean projection onto power cone". In: *Mathematical Methods of Operations Research* 82.3, pp. 265–284.

Microbenchmark

```
POW(0.45,0.55)
  AVGTIM 3.86704e-06  Viol  comp=2.85684e-08  orth=8.0448e-12
                            pfeas=3.0247e-16  dfeas=4.53284e-16

PPOW(0.1,0.9)
  AVGTIM 3.69943e-06  Viol  comp=4.57066e-07  orth=1.78308e-11
                            pfeas=3.4512e-16  dfeas=4.19057e-16

PPOW(0.01,0.99)
  AVGTIM 3.87547e-06  Viol  comp=2.85684e-08  orth=8.0448e-12
                            pfeas=3.0247e-16  dfeas=4.53284e-16

PEXP
  AVGTIM 9.92736e-07  Viol  comp=2.02018e-07  orth=2.06617e-10
                            pfeas=3.8096e-15  dfeas=1.4412e-14

QUAD
  AVGTIM 6.03383e-08  Viol  comp=2.61777e-16  orth=3.05808e-12
                            pfeas=2.5593e-16  dfeas=2.5593e-16
```

Translates into 270K, 1M and 1.7M projections per second.