



Projection and presolve in MOSEK: exponential and power cones


ISMP 2018

Henrik A. Friberg

www.mosek.com






-  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 $\alpha_1, \dots, \alpha_k \geq 0$:

$$\begin{aligned} K_{\text{pow}(\alpha)} &= \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T \alpha}, x_1, \dots, x_k \geq 0\} \\ &= \\ &P(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots). \end{aligned}$$






-  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 $\alpha_1, \dots, \alpha_k \geq 0$:

$$\begin{aligned} K_{\text{pow}(\alpha)} &= \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T \alpha}, x_1, \dots, x_k \geq 0\} \\ &= \\ &P(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots). \quad \text{☹} \end{aligned}$$





-  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 $\alpha_1, \dots, \alpha_k \geq 0$:

$$\begin{aligned} K_{\text{pow}(\alpha)} &= \{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k} \geq \|z\|_2^{e^T \alpha}, x_1, \dots, x_k \geq 0\} \\ &= \\ &P(\mathcal{L} \cap \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots). \quad \text{☺} \end{aligned}$$

In MOSEK 9:

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





① Exponential.

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





① Exponential.

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





① Exponential.

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





- ② $\{t \leq \log(x)\} = \{(x, 1, t) \in K_{\text{exp}}\}.$
- ③ $\{t \geq x \log(x/y)\} = \{(y, x, -t) \in K_{\text{exp}}\}.$
- ④ $\{t \geq (\log x)^2, 0 < x \leq 1\} = \left\{ \left(\frac{1}{2}, t, u\right) \in \mathcal{Q}_r^3, (x, 1, u) \in K_{\text{exp}}, x \leq 1 \right\}.$
- ⑤ $\{t \leq \log \log x, x > 1\} = \{(u, 1, t) \in K_{\text{exp}}, (x, 1, u) \in K_{\text{exp}}\}.$
- ⑥ $\{t \geq (\log x)^{-1}, x > 1\} = \{(u, t, \sqrt{2}) \in \mathcal{Q}_r^3, (x, 1, u) \in K_{\text{exp}}\}.$
- ⑦ $\{t \leq \sqrt{\log x}, x > 1\} = \left\{ \left(\frac{1}{2}, u, t\right) \in \mathcal{Q}_r^3, (x, 1, u) \in K_{\text{exp}} \right\}.$
- ⑧ $\{t \leq \sqrt{x \log x}, x > 1\} = \{(x, u, \sqrt{2}t) \in \mathcal{Q}_r^3, (x, 1, u) \in K_{\text{exp}}\}.$
- ⑨ $\{t \geq x \exp(x), x \geq 0\} = \left\{ \left(\frac{1}{2}, u, x\right) \in \mathcal{Q}_r^3, (t, x, u) \in K_{\text{exp}} \right\}.$



⑩ Log-sum-exponential.

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





⑩ Log-sum-exponential.

$$\begin{aligned} t &\geq \log(e^{x_1} + \dots + e^{x_n}) \\ e^t &\geq e^{x_1} + \dots + e^{x_n} \quad \text{☹} \end{aligned}$$





⑩ Log-sum-exponential.

$$\begin{aligned} t &\geq \log(e^{x_1} + \dots + e^{x_n}) \\ e^t &\geq e^{x_1} + \dots + e^{x_n} && \text{☹} \\ 1 &\geq e^{x_1-t} + \dots + e^{x_n-t} && \text{☺} \end{aligned}$$





⑩ Log-sum-exponential.

$$\begin{aligned}
 t &\geq \log(e^{x_1} + \dots + e^{x_n}) \\
 e^t &\geq e^{x_1} + \dots + e^{x_n} && \text{☹} \\
 1 &\geq e^{x_1-t} + \dots + e^{x_n-t} && \text{☺}
 \end{aligned}$$

Geometric programming in conic form:

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

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





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





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





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 $[f(x)]_+ = 1$.





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





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





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

The power and exponential cones are also representation sensitive:

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

$$y \geq \exp(x) \iff x \leq \log(y)$$

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





```
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:

$$\| \text{dist}(x_1, \mathcal{K}_1), \text{dist}(x_2, \mathcal{K}_2), \dots \|_\infty.$$





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

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:

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





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):

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





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 $\{a \in \mathcal{K}^\circ \mid a^T \hat{x} > 0\}$.





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 $\{a \in \mathcal{K}^\circ \mid a^T \hat{x} > 0\}$.

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.](#)



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 $\{a \in \mathcal{K}^\circ \mid a^T \hat{x} > 0\}$.
- The maximal separator solves $\max_{a \in \mathcal{K}^\circ, \|a\|_2 \leq 1} a^T \hat{x}$.





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 $\{a \in \mathcal{K}^\circ \mid a^T \hat{x} > 0\}$.
- The maximal separator solves $\max_{a \in \mathcal{K}^\circ, \|a\|_2 \leq 1} a^T \hat{x}$.
- Its dual problem is $\min_{x \in \mathcal{K}} \|x - \hat{x}\|_2$.
- Maximal separator is dual solution to projection problem.





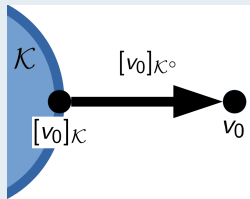
All matrices/vectors are uniquely decomposable as

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

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

Maximal separation:

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





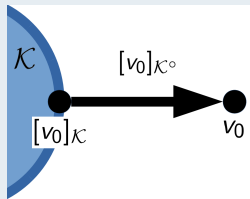
All matrices/vectors are uniquely decomposable as

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

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

Maximal separation:

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



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



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

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

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





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

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

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

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



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

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

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 $\min_{v_p \in \mathcal{K}} \frac{1}{2} \|v_p - v_0\|_2^2$.





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

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





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

in terms of

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





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

in terms of

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



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

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

Case covered

- 1 $[v_0 \in K_{\text{exp}}]$ or $[v_0 \in K_{\text{exp}}^\circ]$ 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 $[s_d \leq 0 \text{ and } r_p \leq 0]$.

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

Case remaining

- 1 $[v_0 \notin K_{\text{exp}}]$ and $[v_0 \notin K_{\text{exp}}^\circ]$ and $[r_0 > 0 \text{ or } s_0 > 0]$.
- 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 $[s_d > 0 \text{ or } r_p > 0]$.

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

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

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

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



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

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

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

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



Exponential cone projection



Step 2. Simplify Moreau for remaining case

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 - 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 - 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 $u = \infty$.

Exponential cone projection

Algorithmic ideas and concerns



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 $u = \infty$.



Change v_0 such that the presolve rules apply.

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

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

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


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

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

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





 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
```



$$K_{\text{pow}(\alpha)} = \left\{ \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}_+^k \times \mathbb{R}^{n-k} \mid x^\alpha \geq \|z\|_2 \right\},$$
$$K_{\text{pow}(\alpha)}^\circ = \left\{ \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}_-^k \times \mathbb{R}^{n-k} \mid \alpha^{-\alpha} (-x)^\alpha \geq \|z\|_2 \right\},$$

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.





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.