Solving Mixed Integer Quadratic and Conic Optimization problems

Pietro Belotti

Xpress Optimizer Team
Fair Isaac

MINO/COST Spring School, CNAM, Paris
8 April 2016

MINLPs: lessons learned

MINLP is hard because it combines

- integrality constraints (non-convex!)
- nonlinear constraints (sometimes non-convex)

There are several ways to deal with this deadly combination.

- Most of them really aim at tightening the continuous relaxation ($\rightarrow$ better lower bound)
- We'll look at two such techniques
- (Not) coincidentally, they both use cones

Outline

- Nonlinearities and disjunctions
- Perspective reformulations
- Disjunctive cones

MINLP: integrality and nonlinearity

Instead of integrality, consider a disjunction:

- binary variable: $z = 0 \vee z = 1$
- variable disjunction: $z \leq \gamma \vee z \geq \gamma + 1$, with $\gamma \in \mathbb{Z}$
- general disjunction: $\boldsymbol{a}^\top \boldsymbol{x} \leq \alpha \vee \boldsymbol{b}^\top \boldsymbol{x} \leq \beta$

So for solving MINLPs (even the convex ones), a natural question is:

The $100 question

How do we combine a disjunction with a convex set?

MINLPs and (convex) cones

Cones are a useful object in MINLP.

For a cone $\mathcal{C}$ pointed at vertex $\boldsymbol{x}_0$,

- $\boldsymbol{x} \in \mathcal{C} \Rightarrow \boldsymbol{x}_0 + \alpha(\boldsymbol{x} - \boldsymbol{x}_0) \in \mathcal{C}$ for any $\alpha \geq 0$.

So their boundary is basically made of rays.

Perspective reformulation[1]

Consider the two sets

$$S_1 = \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \leq 0\} \qquad S_0 = \{\bar{\boldsymbol{x}}\}$$

With $f(\boldsymbol{x})$ a non-homogeneous polynomial of degree k.

Consider a binary variable z such that

$$z = 0 \Rightarrow \boldsymbol{x} = \bar{\boldsymbol{x}}$$
$$z = 1 \Rightarrow \boldsymbol{x} \in S_1$$

Then

- Multiply all terms in $f(\boldsymbol{x})$ of degree $d < k$ by z^{k-d}

⇒ Obtain a new (homogeneous) polynomial $g(\boldsymbol{x})$

☺ The constraint $g(\boldsymbol{x}) \leq 0$ is valid.

[1] Günlük and Linderoth, "Perspective reformulations of mixed integer nonlinear programs with indicator variables", *Mathematical programming* 124 (2010): 183-205.

Example 1: `uflquad` instances

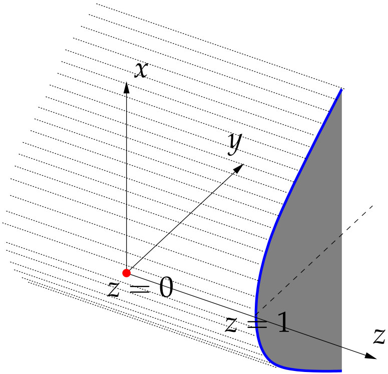
- Uncapacitated facility location problems
- Quadratic constraint: $y \geq x^2$
- x is turned off by z (binary): $x \leq z$

For $z = 0$, the constraint becomes redundant and the only solution is $x = y = 0$.

Then multiplying the constraint by z at an opportune degree yields

$$x^2 \leq yz$$

Perspective reformulation (cont'd)



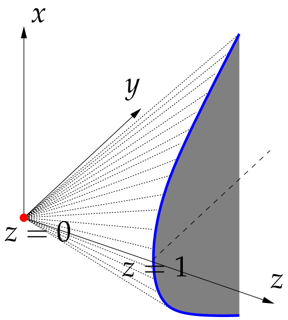
The red dot is the solution for $z = 0$, i.e., $x = y = 0$.

The grey area is the set of solutions for $z = 1$, i.e., $y \geq x^2$

The original constraint $x^2 \leq y$ does not depend on z, so it is a *cylinder* along the direction of z

The tightened constraint $x^2 \leq yz$ is a (rotated) second order cone

Perspective reformulation (cont'd)



The red dot is the solution for $z = 0$, i.e., $x = y = 0$.

The grey area is the set of solutions for $z = 1$, i.e., $y \geq x^2$

The original constraint $x^2 \leq y$ does not depend on z, so it is a *cylinder* along the direction of z

The tightened constraint $x^2 \leq yz$ is a (rotated) second order cone

Example 2: `sssd` instances[2]

Consider the following constraint:

$$y \leq \frac{x}{1+x}$$

with $x, y \geq 0$. Not exactly a nice convex quadratic constraint, but let's massage it a little:

$$\begin{aligned}
&y(1+x) - x \leq 0 \\
&y + xy - x \leq 0 \\
&y + xy - x + x^2 - x^2 \leq 0 \\
&x^2 \leq -y - xy + x + x^2 \\
&x^2 \leq (1+x)(x-y)
\end{aligned}$$

Which is a (rotated!) second order cone.

[2] S.Elhedhli, "Service system design with immobile servers, stochastic demand, and congestion", *Manufacturing & Service Operations Management* 8.1 (2006): 92-97.

Example 2: `sssd` instances[3]

So we have the conic constraint

$$x^2 \leq (1 + x)(x - y)$$

In addition, the binary variable $z = 0$ implies $x = 0$, which turns the whole constraint into $y \leq 0$, trivial.

Hence the constraint can be rewritten as

$$x^2 \leq (z + x)(x - y)$$

---

[3] S.Elhedhli, "Service system design with immobile servers, stochastic demand, and congestion", *Manufacturing & Service Operations Management* 8.1 (2006): 92-97.

Performance difference

Instance	Orig	Reformulated
$x^2 \le yz$		
uflquad-nopsc-20-100	46.2	3.4
uflquad-nopsc-20-150	253.9	4.5
uflquad-nopsc-30-100	212.3	5.9
uflquad-nopsc-30-150	568.4	1.0
uflquad-nopsc-30-200	5289.5	10.6
$x^2 \le (z + x)(x - y)$		
sssd-weak-15-8	11.5	3.2
sssd-weak-20-8	69.4	4.6
sssd-weak-25-8	1310.2	100.9
sssd-weak-30-8	211.8	38.3

To recap

- Any non-homogeneous quadratic constraint might be reformulated
- (if it falls into that S_0/S_1 class)
- Substantial computational gains
- Cones save the day!

This is not the only case where cones help tightening the relaxation.

Mixed integer convex optimization

$$\begin{array}{ll} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{E} \\ & \boldsymbol{x} \in \mathbb{Z}^d \times \mathbb{R}^{n-d} \end{array}$$

- $\mathcal{E}$ is a closed convex set
- Knowing $\text{conv}(\mathcal{E} \cap (\mathbb{Z}^d \times \mathbb{R}^{n-d}))$ is useful
- More realistic:
  - Take disjunction $\mathcal{A} \cup \mathcal{B}$
  - Find the convex hull of $\mathcal{E} \cap (\mathcal{A} \cup \mathcal{B})$

Main result

Consider a closed convex set $\mathcal{E}$ and a disjunction $\mathcal{A} \cup \mathcal{B}$, where

$$\mathcal{A} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^\top \boldsymbol{x} \leq \alpha\}; \qquad \mathcal{B} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{b}^\top \boldsymbol{x} \leq \beta\}.$$

Assume

- $\dim(\mathcal{E}) \geq 2$
- $\mathcal{A} \cap \mathcal{B} \cap \mathcal{E} = \emptyset$
- $\mathcal{E} \cap \partial\mathcal{A}$ and $\mathcal{E} \cap \partial\mathcal{B}$ are **bounded** and nonempty

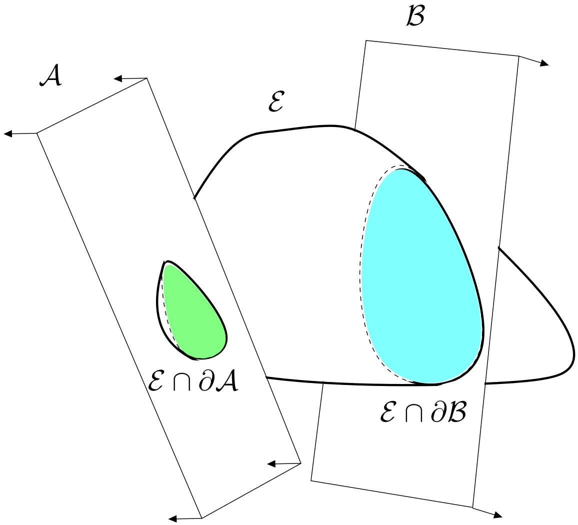
Theorem (B., Góez, Pólik, Ralphs, Terlaky 2012)

If there exists a pointed convex cone $\mathcal{K}$ such that

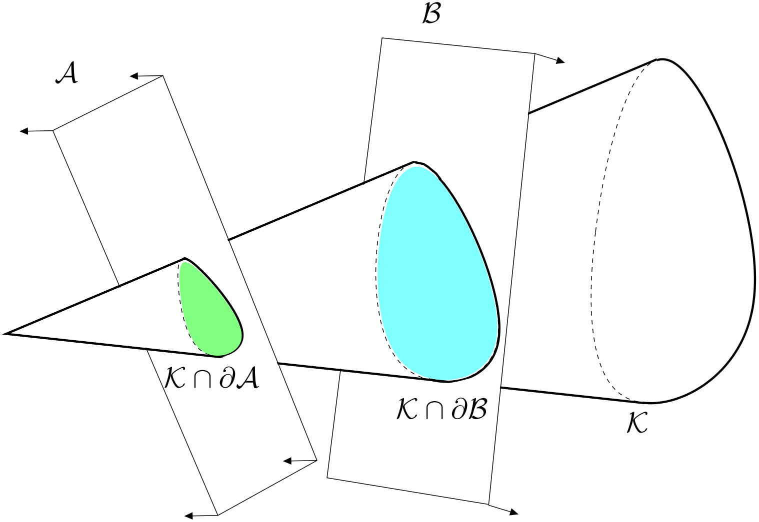
$$\mathcal{K} \cap \partial\mathcal{A} = \mathcal{E} \cap \partial\mathcal{A} \qquad \text{and} \qquad \mathcal{K} \cap \partial\mathcal{B} = \mathcal{E} \cap \partial\mathcal{B},$$

then $\mathcal{K}$ is unique and $\boxed{\operatorname{conv}(\mathcal{E} \cap (\mathcal{A} \cup \mathcal{B})) = \mathcal{K} \cap \mathcal{E}.}$

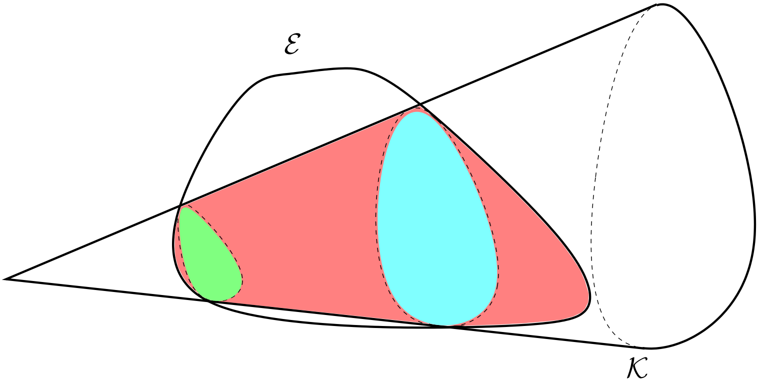
A geometrical explanation



A geometrical explanation



A geometrical explanation



Mixed integer **conic** optimization

Often, $\mathcal{E}$ = affine subspace $\cap$ convex cone:

$$\begin{array}{ll} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{g} \\ & \boldsymbol{x} \in \mathcal{C} \qquad \longleftarrow \text{convex cone} \\ & \boldsymbol{x} \in \mathbb{Z}^d \times \mathbb{R}^{n-d}, \end{array}$$

where $A \in \mathbb{R}^{m \times n}$, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{g} \in \mathbb{R}^m$.

MILP: $\mathcal{C} = \mathbb{R}^n_+$

MISDP: $\mathcal{C}$ is the cone of positive semidefinite matrices

MISOCP: $\mathcal{C} = \mathbb{L}_1^{n_1} \times \mathbb{L}_2^{n_2} \times \cdots \times \mathbb{L}_k^{n_k}$

$\mathbb{L}^h = \{\boldsymbol{x} \in \mathbb{R}^h : \; x_1 \geq \|\boldsymbol{x}_{2:h}\|_2\}$ is a **second order cone**

Related work

- The split closure of a strictly convex body is defined by a **finite** number of split disjunctions[4].
- The Chvátal-Gomory closure of an ellipsoid is a polyhedron[5].
- If $\mathcal{E}$ is bounded, conic quadratic representable, and strictly convex, its split closure is conic quadratic representable.
- Conic MIR cuts[6]
- Generalization: n-step conic MIR inequalities[7]

---

[4] D. Dadush, S.S. Dey, J.P. Vielma, OR *Letters* 39(2):121-126, 2011.

[5] D. Dadush, S.S. Dey, J.P. Vielma, *Math of OR* 36(2):227-239, 2011.

[6] A. Atamtürk, V. Narayanan, "Conic mixed-integer rounding cuts", *Math. Prog.* 122(1):1-20, 2010.

[7] S. Masihabadi, S. Sanjeevi, K. Kianfar, "n-step Conic Mixed Integer Rounding Inequalities", Texas A&M University, Apr. 2011.

More related work

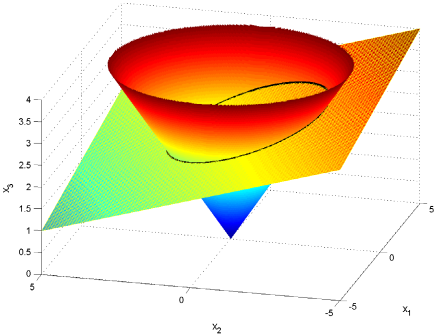
- Drewes (2009): nonlinear cuts for 0-1 MISOCO problems.
- Krokhmal and Soberanis (2010), Vielma et al. (2008): branch and bound algorithm based on linear outer approximations for second order cones.
- Cezik and Iyengar (2005): cuts for mixed 0-1 conic programming.
- Kılınç-Karzan (2015), Minimal inequalities for MICP
- Burer and Kılınç-Karzan (2014): convexifying the intersection of a second-order cone and a nonconvex quadratic constraint

Simple case: $\mathcal{C}$ is **one** second order cone

$$\begin{array}{ll} \min & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{g} \\ & \boldsymbol{x} \in \mathbb{L}^n \\ & \boldsymbol{x} \in \mathbb{Z}^d \times \mathbb{R}^{n-d} \end{array}$$

Example: $\mathcal{C}$ is **one** second order cone

$$\begin{array}{rl} \min & x_1 - 2x_2 + x_3 \\ \text{s.t.} & x_1 - 0.1x_2 + 0.2x_3 = 2.5 \\ & x_3 \geq \|(x_1, x_2)\| \end{array}$$



Example: $\mathcal{C}$ is **one** second order cone

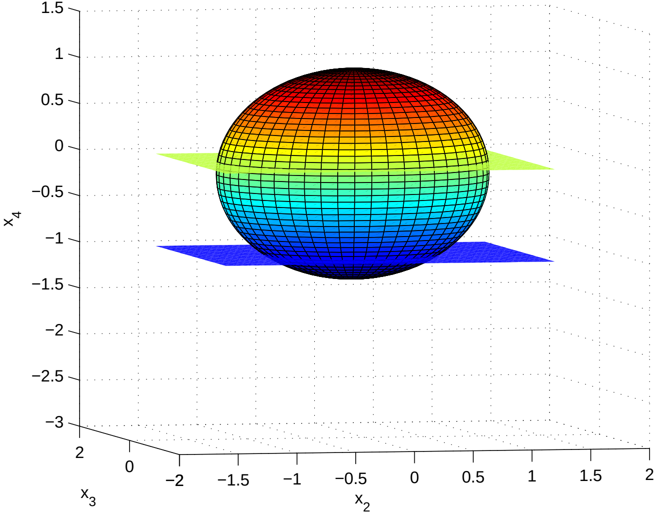
$$
\begin{array}{llllll}
\min & 3x_1 & +2x_2 & +2x_3 & +x_4 & & \\
\text{s.t.} & 9x_1 & +x_2 & +x_3 & +x_4 & = & 10 \\
 & & & & (x_1, x_2, x_3, x_4) & \in & \mathbb{L}^4 \\
 & & & & x_4 & \in & \mathbb{Z}.
\end{array}
$$

Optimum of the continuous relaxation:

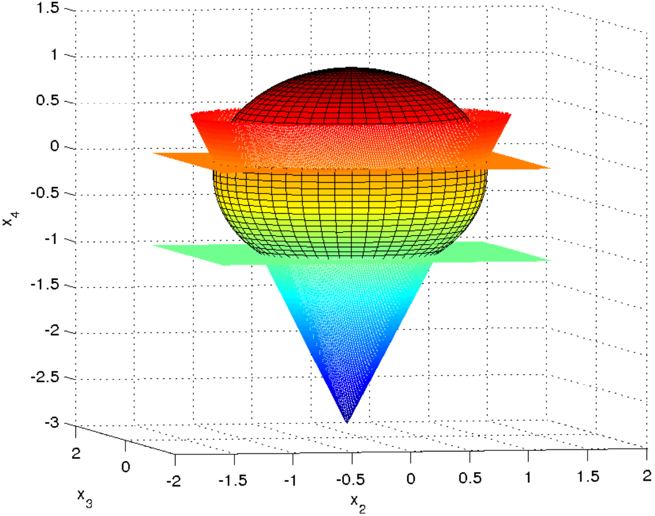
$$\boldsymbol{x}^* = (1.36, -0.91, -0.91, -0.45),$$

Objective function: $z^* = 0$.

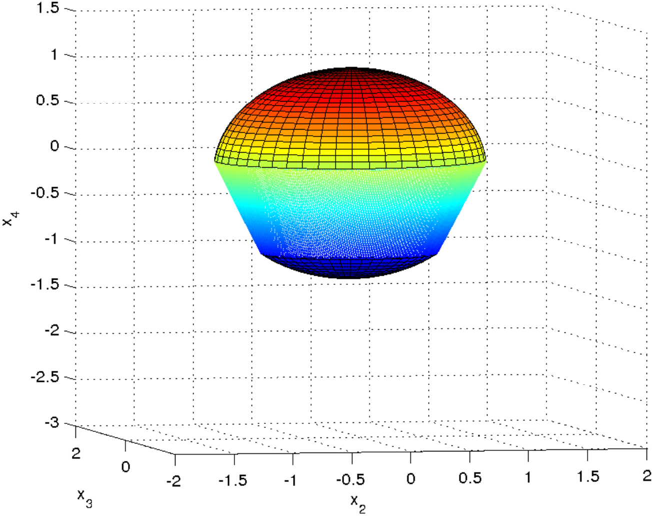
Disjunction violated by $\boldsymbol{x}^*$: $(x_4 \leq -1) \vee (x_4 \geq 0)$



Disjunctive conic cut



Strengthened relaxation



Relaxation + disjunctive cut

$$
\begin{aligned}
\min \quad & 3x_1 + 2x_2 + 2x_3 + x_4 \\
\text{s.t.} \quad & 9x_1 + x_2 + x_3 + x_4 = 10 \\
& 10.14 + 0.04x_2 + 0.04x_3 + 3.56x_4 \geq \\
& \quad \sqrt{(1.65 + 6.28x_2 + 6.28x_3 + 0.14x_4)^2 + (-6.36x_2 + 6.36x_3)^2} \\
& (x_1, x_2, x_3, x_4) \in \mathbb{L}^4 \\
& x_4 \in \mathbb{Z}.
\end{aligned}
$$

Optimal solution of the new continuous relaxation:

$$\boldsymbol{x}^* = (1.32,\ -0.93,\ -0.93,\ 0)$$

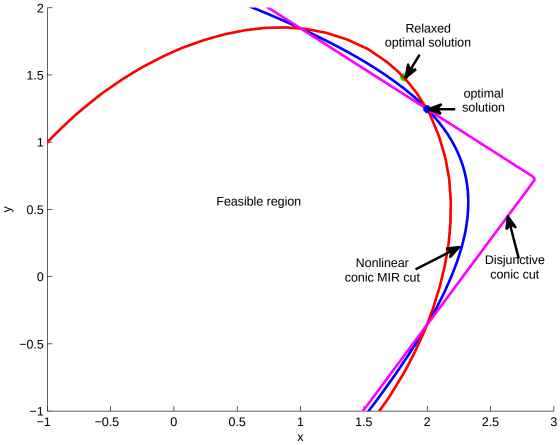
Objective value: $z^* = 0.23$.

Compare with conic MIR cut[8]

$$
\begin{aligned}
\min \quad & -x - y \\
\text{s.t.} \quad & x + y + 2t = \eta \\
& \sqrt{(x - \tfrac{4}{3})^2 + (y - 1)^2} \le t \\
& x \in \mathbb{Z}, y \in \mathbb{R}
\end{aligned}
$$

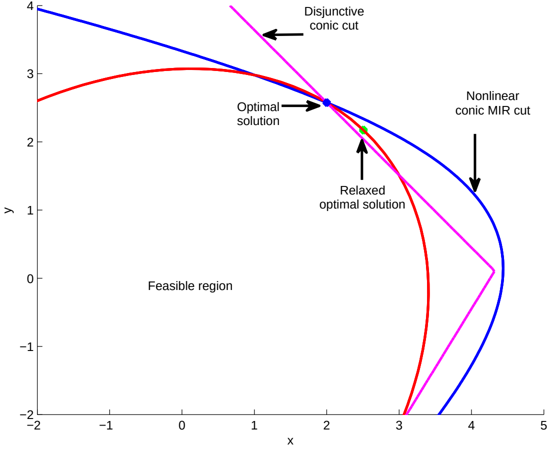
[8] A. Atamtürk, V. Narayanan, "Conic mixed-integer rounding cuts", *Math. Prog.* 122(1):1-20, 2010

Compare with conic MIR cut



$$\eta = 4.66$$

Compare with conic MIR cut



$$\eta = 8$$

Finding $\mathcal{K}$ in MISOCO

We are looking for a **second order cone** $\mathcal{K}$ such that

$$\begin{aligned} \mathcal{K} \cap \partial\mathcal{A} &= \mathcal{E} \cap \partial\mathcal{A} \\ \mathcal{K} \cap \partial\mathcal{B} &= \mathcal{E} \cap \partial\mathcal{B} \end{aligned}$$

$\Rightarrow$ Search the generic quadric $\mathcal{Q}$ such that

$$\begin{aligned} \mathcal{Q} \cap \partial\mathcal{A} &= \mathcal{E} \cap \partial\mathcal{A} \\ \mathcal{Q} \cap \partial\mathcal{B} &= \mathcal{E} \cap \partial\mathcal{B} \end{aligned}$$

Intersection of an affine space and a second order cone

Consider an affine subspace $\mathcal{H} = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} = \boldsymbol{g}\}$ and $\boldsymbol{x}_0 \in \mathcal{H}$.

$$\Rightarrow \mathcal{H} \cap \mathbb{L}^n = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{x}_0 + H\boldsymbol{z} \text{ with } \boldsymbol{z}^\top Q\boldsymbol{z} + 2\boldsymbol{q}^\top \boldsymbol{z} + \rho \leq 0\}.$$

Q has at most one negative eigenvalue.

Define the quadric $\mathcal{Q} = \{\boldsymbol{z} | \boldsymbol{z}^\top Q\boldsymbol{z} + 2\boldsymbol{q}^\top \boldsymbol{z} + \rho \leq 0\} = (Q, \boldsymbol{q}, \rho)$.

$\Rightarrow$ Their intersection $\mathcal{E} = \mathcal{H} \cap \mathbb{L}^n$ is a quadric

- (not necessarily finite)

$\Rightarrow$ $\mathcal{E} \cap \partial\mathcal{A}$ and $\mathcal{E} \cap \partial\mathcal{B}$ are quadrics

- (both ellipsoids for the boundedness assumption)

Uni-parametric family of quadrics $\mathcal{Q}(\tau)$

Assume now $Q \succeq 0$.

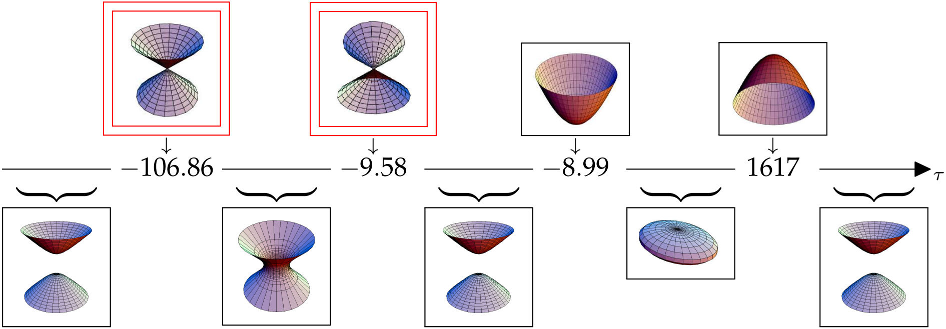
- Recall: $\mathcal{A} = \{\boldsymbol{z} | \boldsymbol{a}^\top \boldsymbol{z} \leq \alpha\}$ and $\mathcal{B} = \{\boldsymbol{z} | \boldsymbol{b}^\top \boldsymbol{z} \leq \beta\}$.

Q.: What is the **family of quadrics** having the same intersection with $\partial\mathcal{A}$ and $\partial\mathcal{B}$ as $\mathcal{Q}$?

A.: It is the family $\mathcal{Q}(\tau) = (Q(\tau), \boldsymbol{q}(\tau), \rho(\tau))$, where

$$
\begin{aligned}
Q(\tau) &= Q + \tau(\boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top) \\
\boldsymbol{q}(\tau) &= \boldsymbol{q} - \tau(\beta\boldsymbol{a} + \alpha\boldsymbol{b}) \\
\rho(\tau) &= \rho + 2\tau\alpha\beta,
\end{aligned}
$$

Family of quadrics $\mathcal{Q}(\tau)$



How do we find the right cone?

The disjunctive conic cut is the quadric generated by

$$\mathcal{Q}(\hat{\tau}) = (Q(\hat{\tau}), \boldsymbol{q}(\hat{\tau}), \rho(\hat{\tau}))$$

where $\hat{\tau}$ is the larger root of the **quadratic** equation

$$\boldsymbol{q}(\tau)^\top Q(\tau)^{-1}\boldsymbol{q}(\tau) - \rho(\tau) = 0.$$

Remember that we only use one “side” of the nonconvex cone.