A course on Robust Optimization

Aharon Ben-Tal

Minerva Optimization center Technion – Israel Institute Of Technology

Lecture 1

Conic Optimization

From Linear to Conic Programming

When passing from a generic LP problem

$$\min_{x} \left\{ c^{T}x : Ax - b \ge 0 \right\} \quad [A : m \times n]$$
 (LP)

to its nonlinear extensions, some components of the problem become nonlinear.

• The traditional way is to allow nonlinearity of the objective and the constraints:

$$c^T x \mapsto c(x); a_i^T x - b_i \mapsto a_i(x)$$

and to preserve the "coordinate-wise" interpretation of the vector inequality $A(x) \geq 0$:

$$A(x) \equiv \begin{pmatrix} a_1(x) \\ \vdots \\ a_m(x) \end{pmatrix} \ge 0 \Leftrightarrow a_i(x) \ge 0, \ i = 1, ..., m.$$

• An alternative is to preserve the linearity of the objective and the constraint functions and to modify the interpretation of the vector inequality \geq .

In Convex Programming, both approaches are equivalent.

♣ In our course, we prefer the second option, due to its strong "unifying abilities": it turns out that an extremely wide variety of nonlinear optimization problems can be covered by just 3 "standard" types of vector inequalities.

• Example: The problem with nonlinear objective and constraints

$\mathbf{minimize} \ \ \underset{\ell=1}{\overset{n}{\sum}} x_{\ell}^2$							
(a)	$x \ge 0;$						
(b)	$a_{\ell}^T x \le b_{\ell}, \ \ell = 1,, n;$						
(c)	$ Px - p _2 \le c^T x + d;$						
(d)	$x_{\ell}^{\frac{\ell+1}{\ell}} \le e_{\ell}^T x + f_{\ell}, \ \ell = 1,, n;$						
(e)	$x_{\ell}^{\frac{l}{l+3}} x_{l+1}^{\frac{1}{l+3}} \ge g_{\ell}^T x + h_{\ell}, \ \ell = 1,, n-1;$						
			x_2 x_1	x_3 x_2		x_n x_{n-1}	
(f)	Det	x_3	x_2	x_1		x_{n-2}	
		$\langle x_n \rangle$	x_{n-1}	x_{n-2}		x_1	
(g)	$1 \le \sum_{\ell=1}^{n}$	x_{ℓ} co	$s(\ell\omega) \leq$	$\leq 1 + \sin \theta$	$n^2(5\omega)$	$\forall \omega \in$	$\left[-\frac{\pi}{7}, 1.3\right]$

can be converted, in a systematic way, into an equivalent problem

$$\min_{x} \left\{ c^T x : Ax - b \succeq 0 \right\},\,$$

" \succeq " being one of our 3 standard vector inequalities, so that seemingly highly diverse constraints of the original problem allow for unified treatment.

• A significant part of nice mathematical properties of an LP program

$$\min_{x} \left\{ c^T x : Ax - b \ge 0 \right\}$$

comes from the fact that the underlying coordinate-wise vector inequality

$$a \ge b \Leftrightarrow a_i \ge b_i, \ i = 1, ..., m$$

$$[a, b \in \mathbf{R}^m]$$

satisfies a number of quite general axioms, namely:

- I. It defines a partial ordering of \mathbb{R}^m , i.e., it is
- I.a) Reflexive: $a \ge a \ \forall a \in \mathbf{R}^m$;
- I.b) Anti-symmetric:

$$\forall a, b \in \mathbf{R}^m : a \ge b \text{ and } b \ge a \Rightarrow a = b;$$

• I.c) <u>Transitive</u>:

$$\forall a, b, c \in \mathbf{R}^m : a \ge b \text{ and } b \ge c \Rightarrow a \ge c.$$

- II. It is compatible with linear structure of \mathbb{R}^m , i.e., is
- II.a) Additive: one can add vector inequalities of the same sign:

$$\forall a, b, c, d \in \mathbf{R}^m : \left\{ \begin{array}{l} a \geq b \\ c \geq d \end{array} \right\} \Rightarrow a + c \geq b + d;$$

• II.b) <u>Homogeneous:</u> one can multiply vector inequality by a nonnegative real:

$$\forall a, b \in \mathbf{R}^m, \lambda \in \mathbf{R} : \left\{ \begin{array}{l} a \geq b \\ \lambda \geq 0 \end{array} \right\} \Rightarrow \lambda a \geq \lambda b.$$

"Good" vector inequalities

• A vector inequality \succeq on \mathbb{R}^m is a binary relation – a set of ordered pairs (a,b) with $a,b\in\mathbb{R}^m$. The fact that a pair (a,b) belongs to this set is written down as $a\succeq b$ (" $a\succeq$ -dominates b").

Let us call a vector inequality $\succeq \underline{good}$, if it satisfies the outlined axioms, namely, is

- I.a) Reflexive: $a \succeq a \ \forall a \in \mathbf{R}^m$;
- I.b) Anti-symmetric:

$$\forall a, b \in \mathbf{R}^m : a \succeq b \text{ and } b \succeq a \Rightarrow a = b;$$

• I.c) <u>Transitive:</u>

$$\forall a, b, c \in \mathbf{R}^m : a \succeq b \text{ and } b \succeq c \Rightarrow a \succeq c.$$

• II.a) Additive:

$$\forall a, b, c, d \in \mathbf{R}^m : \begin{cases} a \succeq b \\ c \succeq d \end{cases} \Rightarrow a + c \succeq b + d;$$

• II.b) Homogeneous:

$$\forall a, b \in \mathbf{R}^m, \lambda \in \mathbf{R} : \left\{ \begin{array}{l} a \succeq b \\ \lambda \geq 0 \end{array} \right\} \Rightarrow \lambda a \succeq \lambda b.$$

- Claims:
 - The coordinate-wise vector inequality

$$a \ge b \Leftrightarrow a_i \ge b_i, i = 1, ..., m$$

is neither the only possible, nor the only interesting good vector inequality on \mathbb{R}^m .

• The generic optimization program

$$\min_{x} \left\{ c^T x : Ax - b \succeq 0 \right\}$$

obtained from LP by replacing the coordinate-wise vector inequality by another good vector inequality inherits a significant part of nice properties of LP programs.

At the same time, specifying " \succeq " properly, one come to generic optimization problems covering a lot of important applications which cannot be treated by the standard LP.

Geometry of good vector inequalities

• Observation I: A good vector inequality \succeq on \mathbb{R}^m is completely characterized by the set $\mathbb{K} = \{a \in \mathbb{R}^m : a \succeq 0\}$ of " \succeq "-nonnegative vectors:

$$a \succeq b \Leftrightarrow a - b \succeq 0 \quad [\Leftrightarrow a - b \in \mathbf{K}]$$

Indeed,

"
$$a \succeq b \Rightarrow a - b \succeq 0$$
":
$$\begin{array}{c} a \succeq b \\ -b \succeq -b & \text{[reflexivity]} \\ \downarrow & \text{[additivity]} \end{array}$$

$$a - b \succeq 0 \Rightarrow a \succeq b$$
":

$$\begin{array}{ccc} a-b \succeq 0 \\ b \succeq b & \text{[reflexivity]} \\ & & & & \\ & & & & \\ \end{array}$$

 $a \succ b$

• Observation II: The necessary and sufficient condition for a set $K \subset \mathbb{R}^m$ to define a good vector inequality \geq_K via the rule

$$a \ge_{\mathbf{K}} b \Leftrightarrow a - b \in \mathbf{K}$$

is to be a nonempty pointed cone:

A: nonemptiness: $K \neq \emptyset$;

B: closedness w.r.t. addition: $a, b \in \mathbf{K} \Rightarrow a + b \in \mathbf{K}$;

C: closedness w.r.t. multiplication by nonnegative reals: $a \in \mathbf{K}, \lambda \geq 0 \Rightarrow \lambda a \in \mathbf{K}$;

D: "pointedness": $a, -a \in \mathbf{K} \Rightarrow a = 0$.

• The cone underlying the usual coordinate-wise vector inequality

$$a \ge b \Leftrightarrow a_i - b_i \ge 0, \ i = 1, ..., m$$

is the nonnegative orthant

$$\mathbf{R}_{+}^{m} = \{x = (x_{1}, ..., x_{m})^{T} \in \mathbf{R}^{m} : x_{i} \geq 0 \quad \forall i \}.$$

This particular cone, in addition to the outlined algebraic properties, possesses two important topological properties: it

E: <u>is closed</u>: if a sequence $\{x_i\}$ of points from the cone converges to a point x, then x belongs to the cone as well;

F: <u>possesses a nonempty interior</u>, i.e., contains a Euclidean ball of positive radius.

• From now on, we call a vector inequality $\geq_{\mathbf{K}}$ on \mathbf{R}^m good, if the underlying cone K possesses, along with properties A-D, also the properties E,F.

- Working with a good vector inequality \geq_{K} , we may
 - use the standard "arithmetics of inequalities", i.e., add vector inequalities of the same sign and multiply vector inequalities by nonnegative reals;
 - pass to limits in both sides of a vector inequality:

$$a_i \ge_{\mathbf{K}} b_i, \lim_i a_i = a, \lim_i b_i = b \Rightarrow a \ge_{\mathbf{K}} b$$
 [by E]

• define the <u>strict</u> version $>_{K}$ of the vector inequality:

$$\underbrace{a >_{\mathbf{K}} b \Leftrightarrow a - b \in \operatorname{int} \mathbf{K}}_{\mathbf{definition of} >_{\mathbf{K}}}$$

$$\underbrace{\exists \epsilon > 0 : a + x' \geq_{\mathbf{K}} b + x'' \quad \forall x', x'' : \|x'\|, \|x''\| \leq \epsilon}$$

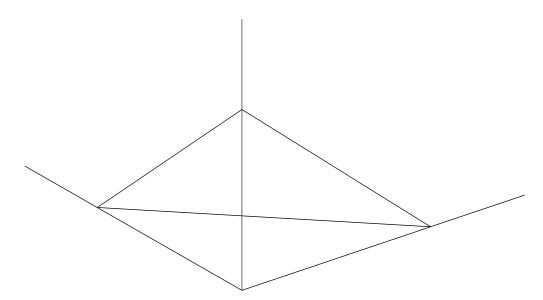
stability w.r.t. small perturbations of both sides

• Note that the arithmetics of strict and non-strict inequalities is governed by the usual rules like

$$a \ge_{\mathbf{K}} b, \ c >_{\mathbf{K}} d \Rightarrow a + c >_{\mathbf{K}} b + d;$$

 $a >_{\mathbf{K}} b, \ \lambda > 0 \Rightarrow \lambda a >_{\mathbf{K}} \lambda b.$

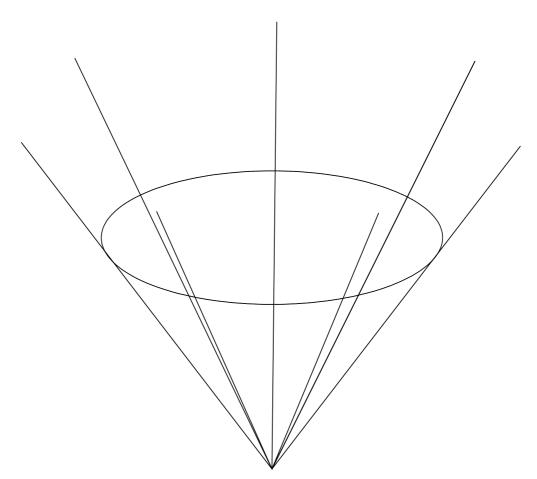
- In our course, we will be especially interested in the vector inequalities coming from the following 3 standard cones:
 - 1. The nonnegative orthant \mathbb{R}_{+}^{m} :



The nonnegative orthant \mathbb{R}^3

2. The $\underline{\text{Lorentz}}$ (or the second-order, or the ice-cream) cone

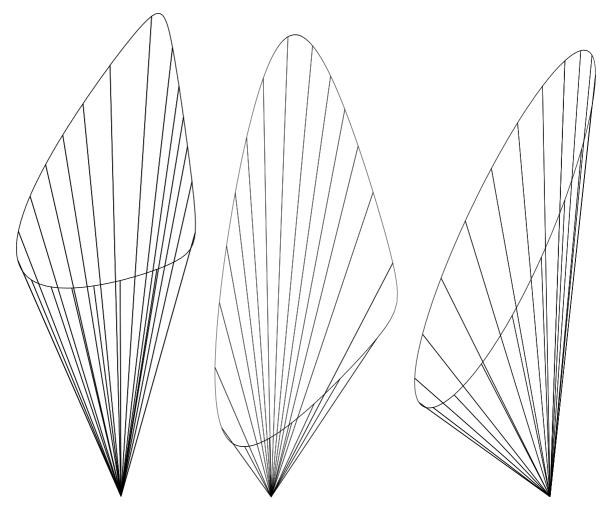
$$\mathbf{L}^{m} = \left\{ x = (x_{1}, ..., x_{m-1}, x_{m})^{T} \in \mathbf{R}^{m} : x_{m} \ge \sqrt{\sum_{i=1}^{m-1} x_{i}^{2}} \right\}$$



The ice-cream cone L^3

3. The <u>positive semidefinite cone</u> S_+^m . This cone belongs to the space S^m of $m \times m$ symmetric matrices and is comprised of all $m \times m$ symmetric positive semidefinite matrices, i.e., of $m \times m$ matrices A such that

$$A = A^T, \quad x^T A x \ge 0 \quad \forall x \in \mathbf{R}^m.$$



3 random 3D cross-sections of \mathbf{S}^3_+

Conic optimization program

• Let $K \subset \mathbb{R}^m$ be a cone defining a good vector inequality \geq_K (i.e., K is a closed pointed cone with a nonempty interior).

A generic conic problem associated with K is an optimization program of the form

$$\min_{x} \left\{ c^{T} x : Ax - b \ge_{\mathbf{K}} 0 \right\}. \tag{CP}$$

$$\min_{x} \left\{ c^{T} x : Ax - b \ge_{\mathbf{K}} 0 \right\}. \tag{CP}$$

Examples:

• Linear Programming

$$\min_{x} \left\{ c^T x : Ax - b \ge 0 \right\} \tag{LP}$$

(K is a nonnegative orthant)

• Conic Quadratic Programming:

$$\min_{x} \left\{ c^{T}x : \|D_{\ell}x + d_{\ell}\|_{2} \leq e_{\ell}^{T}x + f_{\ell}, \ \ell = 1, ..., k \right\}$$

$$\lim_{x} \left\{ c^{T}x : Ax - b \equiv \begin{bmatrix} D_{1}x + d_{1} \\ e_{1}^{T}x + f_{1} \end{bmatrix} \right\} \geq_{\mathbf{K}} 0$$

$$\left[D_{k}x + d_{k} \\ e_{k}^{T}x + f_{k} \end{bmatrix} \right]$$

$$\mathbf{K} = \mathbf{L}^{m_{1}} \times ... \times \mathbf{L}^{m_{k}}$$

$$(CQP)$$

is a direct product of Lorentz cones

• Semidefinite Programming:

$$\min_{x} \left\{ c^{T}x : \mathcal{A}x - B \equiv x_{1}A_{1} + \dots + x_{n}A_{n} - B \succeq 0 \right\}$$

$$\left[P \succeq Q \Leftrightarrow P \geq_{\mathbf{S}_{+}^{m}} Q \right]$$

$$(SDP)$$

Conic Duality

• Let us look at the origin of the problem dual to an LP program

$$\min_{x} \left\{ c^{T} x : Ax - b \ge 0 \right\}. \tag{LPr}$$

Observing that any <u>nonnegative</u> "weight vector" $y \in \mathbb{R}^m$ is "admissible" for the constraint-wise vector inequality on \mathbb{R}^m :

$$\forall a, b, y \in \mathbf{R}^m : a \ge b \text{ and } y \ge 0 \Rightarrow \underbrace{y^T a \ge y^T b}_{\text{usual scalar inequality}}$$

we conclude that all scalar linear inequalities of the type

$$\left[A^T y\right]^T x \ge b^T y \qquad [y \in \mathbf{R}_+^m]$$

with variables x are consequences of the constraints of (LPr).

In particular,

- (*) If $y \ge 0$ is such that $A^T y = c$, then $b^T y$ is a lower bound on the optimal value in (LPr).
- The LP dual to (LPr) is exactly the problem

$$\min_{y} \left\{ b^{T} y : A^{T} y = c, \ y \ge 0 \right\}$$
 (LDI)

of finding the best (the largest) lower bound on the optimal value of (LPr) among those given by (*).

• Conic Duality (same as the LP one) is inspired by the desire to bound from below the optimal value in a conic program

$$\min_{x} \left\{ c^{T} x : Ax - b \ge_{\mathbf{K}} 0 \right\} \tag{CP}$$

and follows the just outlined scheme.

• The scheme is based on "linear aggregation" of vector inequalities into the scalar ones:

$$a \ge_{\mathbf{K}} b \Rightarrow y^T a \ge y^T b,$$
 (*)

and the crucial issue is:

(?) Given a vector inequality $\geq_{\mathbf{K}}$, what are the associated <u>admissible</u> weight vectors y – those for which the implication (*) is true, so that the scalar inequality

$$[A^T y]^T x \ge b^T y$$

is a consequence of the constraints of (CP)?

In the case when \geq_K is the usual coordinate-wise inequality, the admissible weight vectors are those ≥ 0 , but this is not the case for more general vector inequalities:

$$\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \ge_{\mathbf{L}^3} 0$$

$$\begin{pmatrix} 1 \\ 1 \\ 0.1 \end{pmatrix} \ge 0$$

$$\Rightarrow \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0.1 \end{pmatrix}^T \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}}_{-1.8} \ge 0$$

$$- \mathbf{No}$$

• Question: Given a vector inequality $\geq_{\mathbf{K}}$, what are vectors $y \in \mathbf{R}^m$ such that the implication

$$\forall a, b \in \mathbf{R}^m : \quad a \ge_{\mathbf{K}} b \Rightarrow y^T a \ge y^T b$$

is true?

• Answer: These are exactly the vectors y satisfying the requirement

$$y^{T}a \ge 0 \quad \forall \ \underline{a} \ge_{\mathbf{K}} \underline{0}$$

$$\updownarrow$$

$$a \in \mathbf{K}$$

i.e., the vectors from the set

$$\mathbf{K}_* = \left\{ y \in \mathbf{R}^n : y^T x \ge 0 \quad \forall x \in \mathbf{K} \right\}. \tag{*}$$

<u>Theorem.</u> If $K \subset \mathbb{R}^m$ is a closed pointed cone with a nonempty interior, then so is the <u>dual cone</u> K_* given by (*).

For closed pointed cones with nonempty interior, the duality is symmetric:

$$\left(\mathbf{K}_{*}\right)_{*}=\mathbf{K}.$$

• Now we can define the problem dual to a conic problem

$$\min_{x} \left\{ c^{T} x : Ax - b \ge_{\mathbf{K}} 0 \right\}. \tag{CPr}$$

For every $y \geq_{\mathbf{K}_*} 0$, the inequality

$$0 \le y^T A x - y^T b \Leftrightarrow [A^T y]^T x \ge b^T y$$

is a consequence of the constraints in (CPr).

Consequently, whenever

$$y \geq_{\mathbf{K}_*} 0$$
 and $A^T y = c$,

the quantity $b^T y$ is a lower bound on the optimal value of (CPr).

• The problem dual to (CPr) is just the problem

$$\max \left\{ b^T y : A^T y = c, \ y \ge_{\mathbf{K}_*} 0 \right\}$$
 (CDI)

of finding the best (the largest) of the bounds yielded by the outlined construction.

$$\min_{x} \left\{ c^{T}x : Ax - b \ge_{\mathbf{K}} 0 \right\}$$
 (CPr)
$$\max_{x} \left\{ b^{T}y : A^{T}y = c, \ y \ge_{\mathbf{K}_{*}} 0 \right\}$$
 (CDl)

• The origin of the dual problem yields the Weak Duality Theorem: The value of the primal objective $\overline{c^T}x$ at every primal feasible solution (one feasible for (CPr)) is \geq the value of the dual objective b^Ty at every dual feasible solution y (one feasible for (CDl)).

Equivalently: For every primal-dual feasible pair (x, y), the associated duality gap

DualityGap
$$(x, y) = c^T x - b^T y$$
,

is nonnegative.

Conic Duality Theorem

• A conic problem

$$\min_{x} \left\{ e^{T} s : s \in [\mathcal{L} + f] \cap \mathbf{K} \right\}$$

is called <u>strictly feasible</u>, if its feasible plane intersects the <u>interior</u> of the cone K.

(P):
$$\min_{x} \{c^{T}x : Ax - b \ge_{\mathbf{K}} 0\}$$

(D): $\max_{y} \{b^{T}y : A^{T}y = c, y \ge_{\mathbf{K}_{*}} 0\}$

- Conic Duality Theorem. Consider a conic problem (P) along with its dual (D).
- 1. <u>Symmetry:</u> The duality is symmetric: the problem dual to dual is (equivalent to) the primal;
- 2. Weak duality: The value of the dual objective at any dual feasible solution is \leq the value of the primal objective at any primal feasible solution;
- 3. Strong duality in strictly feasible case: If one of the problems (P), (D) is strictly feasible and bounded, then the other problem is solvable, and the optimal values in (P) and (D) are equal to each other.

If both (P), (D) are strictly feasible, then both problems are solvable with equal optimal values.

(P):
$$\min_{x} \{c^{T}x : Ax - b \ge_{\mathbf{K}} 0\}$$

(D): $\max_{y} \{b^{T}y : A^{T}y = c, y \ge_{\mathbf{K}_{*}} 0\}$

- Corollary. Let one of the problems (P), (D) be strictly feasible and bounded. Then a pair (x,y) of primal and dual feasible solutions is comprised of optimal solutions to the respective problems
 - If and only if the duality gap at the pair is zero:

DualityGap
$$(x, y) \equiv c^T x - b^T y = 0$$

Indeed, in the case in question Opt(P) = Opt(D), hence

$$c^{T}x - b^{T}y = \left[\underbrace{c^{T}x - \operatorname{Opt}(P)}_{\geq 0}\right] + \left[\underbrace{\operatorname{Opt}(D) - b^{T}y}_{\geq 0}\right]$$

as well as

• If and only if the complementary slackness holds:

$$[\underbrace{y}_{\in \mathbf{K}_*}]^T [\underbrace{Ax - b}] = 0$$

Indeed, for primal-dual feasible (x, y) we have

$$c^{T}x - b^{T}y = [A^{T}y]^{T}x - b^{T}y = y^{T}[Ax - b].$$

Example: Dual to the Steiner sum problem

Steiner sum problem:

$$\min_{x \in \mathbf{R}^n} \sum_{i=1}^m \|x - a_i\|_2. \quad [m > 1, a_1, ..., a_m \text{ are distinct points in } \mathbf{R}^n]$$

"Cover story" (n = 2): There are m oil wells located at points $a_1, ..., a_m \in \mathbb{R}^2$. Where should one place an oil collector in order to minimize the total length of pipelines connecting the wells to the collector?

A The problem can be reformulated as conic:

$$\min_{t_1,...,t_m,x} \left\{ \sum_{i=1}^m t_i : \|x - a_i\|_2 \le t_i, \ i = 1,...,m \right\}$$
(S)

Note: $||x - a_i||_2 \le t_i$ if and only if $\begin{bmatrix} x - a_i \\ t_i \end{bmatrix} \in \mathbf{L}^{n+1}$ (Lorentz cone).

• (S) is a conic problem of the generic form

$$\min_{z} \left\{ c^{T}z : \begin{bmatrix} A_{1}z - b_{1} \\ \vdots \\ A_{m}z - b_{m} \end{bmatrix} \in \mathbf{K} \right\}, \tag{1}$$

where K is a direct product of closed pointed cones K_i with nonempty interiors:

$$\mathbf{K} = \mathbf{K}_1 \times ... \times \mathbf{K}_m = \left\{ \xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_m \end{bmatrix} : \xi_i \in \mathbf{K}_i, \ i = 1, ..., m \right\}$$
(2)

• Conic duality "recipe" as applied to problem

$$\min_{x} \left\{ c^{T} x : Ax \ge_{\mathbf{K}} b \right\} \tag{P}$$

is as follows:

- 1. Specify "admissible aggregation weights" y, those for which vector inequality $a \ge_{\mathbf{K}} b$ always implies scalar inequality $y^T a \ge y^T b$ (these weights are exactly vectors from \mathbf{K}_*);
- 2. Every admissible aggregation weight y combines with the vector inequality constraint in (P) to produce the scalar inequality

$$[(A^T y)^T x \equiv] y^T A x \ge y^T b$$

which is valid on the entire feasible set of (P). Select those admissible aggregation weights y for which the left hand side in this inequality, as a function of x, is exactly the objective c^Tx (i.e., $A^Ty = c$). Every selected y produces the lower bound b^Ty on the optimal value in (P), and the dual of (P) is exactly the problem

$$\max_{y} \left\{ b^T y : y \in \mathbf{K}_*, A^T y = c \right\}$$

of maximizing this bound over the selected aggregation weights.

Remark: In the case when the vector inequality in (P) has a "direct product structure", i.e.,

$$Ax \geq_{\mathbf{K}} b \Leftrightarrow A_i x \geq_{\mathbf{K}_i} b_i, i = 1, ..., m,$$

(so that $K = K_1 \times ... \times K_m$), we have

$$\mathbf{K}_* = (\mathbf{K}_1)_* \times ... \times (\mathbf{K}_m)_*,$$

that is, admissible aggregation weights y for $\geq_{\mathbf{K}}$ are collections of admissible aggregation weights y_i for the respective vector inequalities $\geq_{\mathbf{K}_i}$.

♣ Let us apply the outlined scheme to the Steiner sum problem

$$\min_{t,x} \left\{ \sum_{i} t_{i} : \underbrace{\|x - a_{i}\|_{2} \leq t_{i}}_{t_{i}}, i = 1, ..., m \right\}$$

$$\begin{bmatrix} x \\ t_{i} \end{bmatrix} \geq_{\mathbf{L}^{n+1}} \begin{bmatrix} a_{i} \\ 0 \end{bmatrix}$$
(S)

1. Lorentz cones are self-dual; thus, admissible aggregation weights y_i for i-th vector inequality constraint in (S) are of the form $\begin{bmatrix} \xi_i \\ \tau_i \end{bmatrix}$ with $\|\xi_i\|_2 \leq \tau_i$. Aggregating i-th vector inequality constraint in (S) with such a weight, we get

$$\xi_i^T x + \tau_i t_i \ge \xi_i^T a_i. \tag{*_i}$$

2. Admissible aggregation weights for the system of vector inequality constraints in (S) (this system is a vector inequality of the direct product structure) are collections $\{(\xi_i, \tau_i), \|\xi_i\|_2 \leq \tau_i\}_{i=1}^m$ of admissible aggregation weights for particular constraints. Aggregation yields

$$\sum_{i} \left[\xi_i^T x + \tau_i t_i \right] \ge \sum_{i} \xi_i^T a_i. \tag{*}$$

3. Now we should select those admissible aggregation weights $\{(\xi_i, \tau_i), \|\xi_i\|_2 \leq \tau_i\}_{i=1}^m$ for which the left hand side in (*) is, as a function of (x, t), the objective of (S), i.e., is $\equiv \sum_i t_i$, and should maximize the right hand side of (*) over the selected collections. Thus, the conic dual of (S) is the problem

$$\max_{\xi_i, \tau_i} \left\{ \sum_i a_i^T \xi_i : \sum_i \xi_i = 0, \ \|\xi_i\|_2 \le 1 = \tau_i, \ i = 1, ..., m \right\}$$

Conic Duality Theorem: Sketch of the proof

(P):
$$\min_{x} \left\{ c^{T}x : Ax - b \ge_{\mathbf{K}} 0 \right\}$$

(D): $\max_{y} \left\{ b^{T}y : A^{T}y = c, y \ge_{\mathbf{K}_{*}} 0 \right\}$

- ♣ In view of Weak Conic Duality and primal-dual symmetry (which follows from geometric interpretation of the primal-dual pair), all we need to prove is the following fact:
 - (!) Let (P) be strictly feasible and below bounded. Then (D) admits a feasible solution y such that

$$b^T y \ge c^* \equiv \text{Opt}(P).$$
 (*)

- \bigvee Step 0: (!) is trivial when c=0: in this case, $c^*=0$, hence one can take 0 as the required y. Thus, assume that $c\neq 0$.
- ♡ Step 1: Let

$$M = \{ y = Ax - b : c^T x \le c^* \}$$

Claim I: M is a nonempty convex set which does not intersect int **K**.

Nonemptiness and convexity of M are evident. Assuming $M \cap \operatorname{int} \mathbf{K} \neq \emptyset$, there exists \bar{x} such that $A\bar{x} - b \in \operatorname{int} K$ and $c^T\bar{x} \leq c^*$. By the first of these relations, all points x close enough to \bar{x} are feasible for (P), and by the second relation and due to $c \neq 0$, among these points are such that $c^Tx < c^*$, which is impossible. Thus, M does not intersect $\operatorname{int} \mathbf{K}$.

Situation:

$$\underbrace{\{y = Ax - b : c^T x \le c^*\}}_{M \text{ (nonempty and convex)}} \cap \operatorname{int} \mathbf{K} = \emptyset$$
 (1)

 \heartsuit Step 2: Since K is convex and int $K \neq \emptyset$, the set int K is convex. Applying to the pair of nonempty convex non-intersecting sets M, int K the Separation Theorem:

Whenever S, T are nonempty and non-intersecting convex sets, there exists $e \neq 0$ such that

$$\sup_{x \in S} e^T x \le \inf_{y \in T} e^T y$$

we conclude that

$$\exists e \neq 0 : \sup_{x:c^T x \leq c^*} e^T (Ax - b) \leq \inf_{y \in \text{int } \mathbf{K}} e^T y.$$
 (2)

 \bigvee Step 3: Since int K contains, along with every one of its points y, the entire open ray $\{ty: t>0\}$, the right hand side in (2), being finite, must be 0, i.e., $e \in K_*$.

The left hand side in (2) is $\sup_{x:c^Tx\leq c^*}(A^Te)^Tx-e^Tb$. Thus quantity can be finite only when $A^Te=\mu c$ with certain $\mu\geq 0$, and in this case $\sup_{x:c^Tx\leq c^*}(A^Te)^Tx=\mu c^*$. Thus, (2) implies that

$$e \in \mathbf{K}_* \& A^T e = \mu c \& \mu \ge 0 \& \mu c^* - e^T b \le 0.$$

Situation:

$$\exists (e \neq 0, \mu \geq 0) : e \in \mathbf{K}_* \& A^T e = \mu c \& \mu c^* \leq e^T b.$$
 (3)

 \heartsuit Step 4: We claim that $\mu > 0$.

Indeed, the only other option is $\mu = 0$. In this case (3) implies that $0 \le e^T b$.

Since (P) is strictly feasible, there exists \bar{x} such that $\bar{y} = A\bar{x} - b \in \text{int } \mathbf{K}$. Since $0 \neq e \in \mathbf{K}_*$ and $\bar{y} \in \text{int } \mathbf{K}$, we have

$$0 < e^T \bar{y} = (A^T e)^T \bar{x} - e^T b = \mu c^T \bar{x} - e^T b = -e^T b.$$

Thus, e^Tb should be both nonnegative and strictly negative, which is impossible.

 \bigvee <u>Last Step</u>: Since $\mu > 0$, we can set $y = e/\mu$. By (3), we have $y \in \mathbf{K}^* \ \& \ A^T y = c \ \& \ b^T y > c^*$,

as required in the conclusion of (!).

Whether strict feasibility is needed?

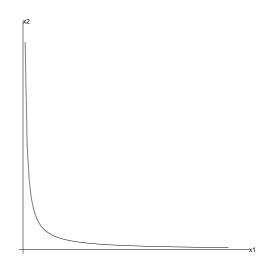
• Example 1:

$$\min_{x_1, x_2, x_3} \left\{ x_1 : Ax - b \equiv \begin{bmatrix} x_1 - x_2 \\ 1 \\ x_1 + x_2 \end{bmatrix} \ge_{\mathbf{L}^3} 0 \right\}$$

$$\lim_{x_1, x_2, x_3} \left\{ x_1 : \sqrt{(x_1 - x_2)^2 + 1} \le x_1 + x_2 \right\}$$

$$\lim_{x_1, x_2, x_3} \left\{ x_1 : 4x_1 x_2 \ge 1, x_1 + x_2 > 0 \right\}.$$

Geometrically: We are minimizing x_1 over the intersection of the 3D ice-cream cone with a 2D plane; the projection of the intersection onto the (x_1, x_2) -plane is the part of the 2D nonnegative orthant bounded by the hyperbola $x_1x_2 \ge 1/4$:



• <u>Conclusion</u>: A conic problem can be strictly feasible and bounded and at the same time – unsolvable!

• Example 2:

$$\min_{x_1, x_2} \left\{ x_2 : Ax - b = \begin{bmatrix} x_1 \\ x_2 \\ x_1 \end{bmatrix} \ge_{\mathbf{L}^3} 0 \right\}.$$

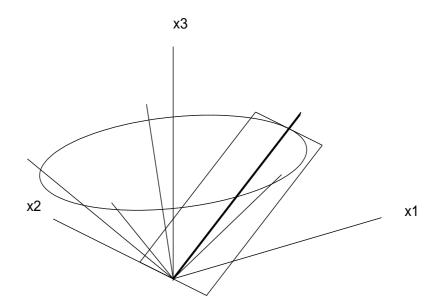
$$\lim_{x_1, x_2} \left\{ x_2 : \sqrt{x_1^2 + x_2^2} \le x_1 \right\}$$

$$\lim_{x_1, x_2} \left\{ x_2 \to \min : x_2 = 0, x_1 \ge 0 \right\}$$

$$\downarrow \downarrow$$

The optimal set: the ray $\{(x_1,0)^T: x_1 \geq 0\}$

Geometrically: We are minimizing the linear objective x_2 over the intersection of the 3D ice-cream cone and the plane tangent to the cone along the ray $\{(x_1, 0, x_1)^T : x_1 \ge 0\}$. The intersection is the ray itself.



• The ice-cream cone is self-dual:

$$(\mathbf{L}^m)_* = \mathbf{L}^m,$$

so that the dual problem is

$$\max_{\lambda} \left\{ b^{T} \lambda \equiv 0 : A^{T} \lambda \equiv \begin{bmatrix} \lambda_{1} + \lambda_{3} \\ \lambda_{2} \end{bmatrix} = c \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \lambda \geq_{\mathbf{L}^{3}} 0 \right\}.$$

$$\lambda \geq_{\mathbf{L}^{3}} 0 \Leftrightarrow \sqrt{\lambda_{1}^{2} + \lambda_{2}^{2}} \leq \lambda_{3}$$

$$\lambda_{1} + \lambda_{3} = 0$$

$$\lambda_{2} = 1$$
infeasible!

• Conclusion: A conic problem can be solvable and possess an infeasible dual!

• Example 3: The problem

$$\min_{x_1, x_2, x_3} \left\{ x_2 : Ax - b \equiv \begin{pmatrix} 1 + x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{pmatrix} \ge_{\mathbf{S}_+^3} 0 \right\}$$

$$\min_{x_1, x_2, x_3} \left\{ x_2 : x_1 \ge 0, x_2 = 0 \right\}$$
(P)

is solvable with optimal value Opt(P) = 0.

The semidefinite cone S_+^m is self-dual, and the problem dual to (P) is

$$\max_{y} \left\{ -y_{11} : \begin{cases} \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{12} & y_{22} & y_{23} \\ y_{13} & y_{23} & y_{33} \end{pmatrix} \geq_{\mathbf{S}_{+}^{3}} 0 \\ y_{22} & = 0, \\ y_{11} + 2y_{23} & = 1 \end{cases} \right.$$

$$\left\{ -y_{11} : \left\{ \begin{pmatrix} y_{11} & 0 & y_{13} \\ 0 & 0 & 0 \\ y_{13} & 0 & y_{33} \end{pmatrix} \geq_{\mathbf{S}_{+}^{3}} 0, \\ y_{13} & 0 & y_{33} \end{pmatrix} \geq_{\mathbf{S}_{+}^{3}} 0, \right\}$$

$$y_{11} & = 1$$
(D)

(D) is solvable with optimal value $\mathrm{Opt}(D) = -1$. Conclusion: It may happen that both primal and dual conic problems are solvable and have different optimal values!

$$Ax \ge_{\mathbf{K}} b$$
 (I)

$$\lambda \ge_{\mathbf{K}_*} 0; \quad A^T \lambda = 0; \quad \lambda^T b > 0.$$
 (II)

Conic Theorem on Alternative:

- 1. If (II) has a solution, then (I) has no solutions.
- 2. If (II) has no solutions, then (I) is "almost solvable":

For every $\epsilon > 0$, you may perturb b by no more than ϵ to get a solvable system (I):

$$\forall \epsilon > 0 \ \exists b' : ||b - b'|| < \epsilon \ \text{and} \ Ax \ge_{\mathbf{K}} b' \ \text{is solvable}$$

3. (II) has no solutions if and only if (I) is almost solvable.

Conic Quadratic Problems

♣ The *m*-dimensional Lorentz cone is

$$\mathbf{L}^m = \{ x = (x_1, ..., x_m) \in \mathbf{R}^m : x_m \ge \sqrt{x_1^2 + ... + x_{m-1}^2} \};$$

A conic quadratic problem is a conic problem

$$\min_{x} \left\{ c^{T} x : Ax - b \ge_{\mathbf{K}} 0 \right\} \tag{CP}$$

for which the cone K is a direct product of Lorentz cones:

$$\mathbf{K} = \mathbf{L}^{m_1} \times \mathbf{L}^{m_2} \times ... \times \mathbf{L}^{m_k}$$

$$= \left\{ y = \begin{pmatrix} y[1] \\ y[2] \\ ... \\ y[k] \end{pmatrix} : y[i] \in \mathbf{L}^{m_i}, \ i = 1, ..., k \right\}.$$

• Thus, a conic quadratic problem is an optimization problem with linear objective and finitely many "conic quadratic constraints":

$$\min_{x} \left\{ c^{T} x : A_{i} x - b_{i} \ge_{\mathbf{L}^{m_{i}}} 0, \ i = 1, ..., k \right\}.$$
 (*)

$$\min_{x} \left\{ c^{T} x : A_{i} x - b_{i} \ge_{\mathbf{L}^{m_{i}}} 0, \ i = 1, ..., k \right\}.$$
 (*)

Representing

$$[A_i; b_i] = \begin{pmatrix} D_i & d_i \\ p_i^T & q_i \end{pmatrix}$$

 $(q^i \text{ is a real}), \text{ we may rewrite (*) "explicitly" as}$

$$\min_{x} \left\{ c^{T}x : \underbrace{\|D_{i}x - d_{i}\|_{2} \leq p_{i}^{T}x - q_{i}, i = 1, ..., k}_{\mathbf{A}_{i}x - b_{i} \geq_{\mathbf{L}^{m}} 0} \right\}.$$
(CQ_s)

• A scalar linear inequality $a^Tx \ge b$ is the same as the conic quadratic inequality

$$\begin{bmatrix} 0 \\ a^T x - b \end{bmatrix} \ge_{\mathbf{L}^2} 0,$$

so that adding to (CQ) finitely many scalar linear inequalities, we do not vary the structure of the problem.

Problem dual to a Conic Quadratic Problem

A Consider a conic quadratic program:

$$\begin{aligned}
\min_{x} c^{T} x \\
Rx &= r, \\
\|D_{i}x - d_{i}\|_{2} &\leq p_{i}^{T} x - q_{i} \Leftrightarrow \begin{cases}
A_{i}x \geq_{\mathbf{L}^{m_{i}}} b_{i}, \\
A_{i} &= \begin{bmatrix} D_{i} \\ p_{i}^{T} \end{bmatrix}, \\
b_{i} &= \begin{bmatrix} d_{i} \\ q_{i} \end{bmatrix}
\end{aligned} (CQp)$$

The Lorentz cone is self-dual, so that the dual to (CQp) is again a conic quadratic program:

$$\begin{cases} \max_{\nu,\eta} \left[r^T \nu + \sum_{i=1}^k b_i^T \eta_i \right] \\ R^T \nu + \sum_{i=1}^k A_i^T \eta_i &= c \\ \eta_i \geq_{\mathbf{L}^{m_i}} 0, \ i = 1, ..., k \end{cases}$$

$$\begin{cases} \max_{\nu,\eta} \left[p^T \nu + \sum_{i=1}^k \left[d_i^T \xi_i + q_i \sigma_i \right] \right] \\ P^T \nu + \sum_{i=1}^k \left[D_i^T \xi_i + \sigma_i p_i \right] = c \\ \|\xi_i\|_2 \leq \sigma_i, \\ i = 1, ..., m \end{cases}$$

$$(\mathbf{CQd})$$

with design variables

$$\nu \in \mathbf{R}^{\dim r}, \left\{ \eta_i = \begin{pmatrix} \xi_i \}_{m_i - 1} \\ \sigma_i \}_1 \end{pmatrix} \right\}_{i=1}^k$$

What can be expressed via conic quadratic constraints?

A Normally, an initial form of an optimization model is

$$\min\{f(x) : x \in X\}$$

$$X = \bigcap_{i=1}^{m} X_{i}$$
[often $X_{i} = \{x : g_{i}(x) \leq 0\}$]

We always may make the objective linear:

$$\min\{f(x) : x \in X\} \qquad \left[X = \bigcap_{i=1}^{m} X_i\right]$$

$$\lim_{t,x} \left\{t : (t,x) \in X^+\right\}, \qquad X^+ = \bigcap_{i=0}^{m} X_i^+$$

$$[X_0^+ = \{(t,x) : t \ge f(x)\}; X_i^+ = \{(t,x) : x \in X_i\}]$$

From now on, assume that the objective is linear, so that the original problem is

$$\min_{x} \left\{ c^{T}x : x \in X \right\} \qquad \left[X = \bigcap_{i=1}^{m} X_{i} \right]$$
 (Ini)

♣ Question: When (Ini) can be reformulated as a conic quadratic problem?

$$\min_{x} \left\{ c^{T}x : x \in X \right\} \qquad \left[X = \bigcap_{i=1}^{m} X_{i} \right]$$
 (Ini)

- ♣ Question: When (Ini) can be reformulated as a conic quadratic problem?
- Answer: This is the case when X is a $\underline{CQr \ set}$ it is "Conic Quadratic representable".

<u>Definition</u>. Let $X \subset \mathbb{R}^n$. We say that X is CQr, if there exist

- a vector $u \in \mathbf{R}^{\ell}$ of additional variables and
 - an affine mapping

$$A(x,u) = \begin{bmatrix} A_1(x,u) \\ A_2(x,u) \\ \dots \\ A_k(x,u) \end{bmatrix} : \mathbf{R}_x^n \times \mathbf{R}_u^\ell \to \mathbf{R}^{m_1} \times \dots \times \mathbf{R}^{m_k}$$

such that

$$X = \{x \in \mathbf{R}^n \mid \exists u : A_j(x, u) \ge_{\mathbf{L}^{m_j}} 0, j = 1, ..., k\}.$$

The collection $(\ell, k, A(\cdot, \cdot), m_1, ..., m_k)$ is called a CQR ("Conic Quadratic Representation") of X.

 \clubsuit Given a CQR of X, we can pose (Ini) as a conic quadratic program, namely, as

$$\min_{x,u} \left\{ c^T x : A_j(x,u) \ge_{\mathbf{L}^{m_j}} 0 \right\}.$$

\$ Example: Consider the program

$$\min_{x} \left\{ x : x \in X = \{ x^2 + 2x^4 \le 1 \} \right\}$$
 (Ini)

A CQR for X can be obtained as follows:

$$x^{2} + 2x^{4} \le 1 \Leftrightarrow \exists t_{1}, t_{2} : \begin{cases} x^{2} \le t_{1} \\ t_{1}^{2} \le t_{2} \\ t_{1} + 2t_{2} \le 1 \end{cases}$$

and

$$s^2 \le r \Leftrightarrow 4s^2 + (r-1)^2 \le (r+1)^2 \Leftrightarrow \begin{pmatrix} 2s \\ r-1 \\ r+1 \end{pmatrix} \ge_{\mathbf{L}^3} 0,$$

so that

$$X = \left\{ x \middle| \exists t_{1}, t_{2} : \left\{ \begin{array}{c} \underbrace{\begin{pmatrix} 2x \\ t_{1} - 1 \\ t_{1} + 1 \end{pmatrix}}_{\text{"says" that } x^{2} \leq t_{1}} \\ \underbrace{\begin{pmatrix} 2t_{1} \\ t_{2} - 1 \\ t_{2} + 1 \end{pmatrix}}_{\text{"says" that } t_{1}^{2} \leq t_{2}} \right\},$$

and (Ini) is the conic quadratic program

$$\min_{x,t_1,t_2} \left\{ x : \begin{pmatrix} 2x \\ t_1 - 1 \\ t_1 + 1 \end{pmatrix} \ge_{\mathbf{L}^3} 0, \begin{pmatrix} 2t_1 \\ t_2 - 1 \\ t_2 + 1 \end{pmatrix} \ge_{\mathbf{L}^3} 0, t_1 + 2t_2 \le 1 \right\}.$$

<u>Definition</u>. Let $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a function. We say that f is CQr ("Conic Quadratic representable"), if its epigraph

$$\mathrm{Epi}\{f\} = \{(t, x) \in \mathbf{R} \times \mathbf{R}^n : f(x) \le t\}$$

is a CQr set.

Every CQR of $\text{Epi}\{f\}$ is called a CQR ("Conic Quadratic Representation") of f.

Example: The function $f(x) = x^2 + 2x^4 : \mathbf{R} \to \mathbf{R}$ is CQr:

$$t \geq x^{2} + 2x^{4} \Leftrightarrow \exists t_{1}, t_{2} : \begin{cases} \begin{pmatrix} 2x \\ t_{1} - 1 \\ t_{1} + 1 \end{pmatrix} \geq_{\mathbf{L}^{3}} 0 \quad [\Leftrightarrow x^{2} \leq t_{1}] \\ \begin{pmatrix} 2t_{1} \\ t_{2} - 1 \\ t_{2} + 1 \end{pmatrix} \geq_{\mathbf{L}^{3}} 0 \quad [\Leftrightarrow t_{1}^{2} \leq t_{2}],$$

$$t_{1} + 2t_{2} \leq t$$

<u>Theorem.</u> If $f: \mathbb{R}^n \to \mathbb{R}$ is a CQr function, then its level sets

$$X_a = \{x : f(x) \le a\}$$

are CQr, and CQR for X_a is given by a CQR for f:

$$\underbrace{t \geq f(x) \Leftrightarrow \exists u : A_j(t, x, u) \geq_{\mathbf{L}^{m_j}} 0, \ j = 1, ..., k}_{\Downarrow}$$
$$\{x : f(x) \leq a\} = \{x \mid \exists u : A_j(a, x, u) \geq_{\mathbf{L}^{m_j}} 0, \ j = 1, ..., k\}$$

"Calculus" of CQr functions and sets

Elementary CQr functions/sets. The following functions/sets are CQr with explicit CQr's:

1. Closed half-space and affine function. Closed half-space $\Pi = \{x : a^T x \leq b\}$ is CQr:

$$\Pi = \{x : \begin{bmatrix} 0 \\ b - a^T x \end{bmatrix} \ge_{\mathbf{L}^2} 0\}.$$

Affine function $f(x) = a^T x + b$ is CQr, since its epigraph is a closed half-space:

$$\operatorname{Epi}\{f\} = \{(t, x) : t \ge a^T x + b\} = \{(t, x) : \begin{bmatrix} 0 \\ t - a^T x - b \end{bmatrix} \ge_{\mathbf{L}^2} 0\}$$

2. The Euclidean norm $f(x) = ||x||_2 : \mathbf{R}^n \to \mathbf{R}$:

$$\operatorname{Epi}\{f\} = \{(t, x) : \left[\begin{array}{c} x \\ t \end{array}\right] \ge_{\mathbf{L}^{n+1}} 0\}$$

3. The squared Euclidean norm $f(x) = x^T x : \mathbf{R}^n \to \mathbf{R}$:

$$t \ge x^T x \Leftrightarrow (t+1)^2 \ge (t-1)^2 + 4x^T x \Leftrightarrow \begin{bmatrix} 2x \\ t-1 \\ t+1 \end{bmatrix} \ge_{\mathbf{L}^{n+2}} 0$$

4. The fractional-quadratic function

$$f(x,s) = \begin{cases} \frac{x^T x}{s}, & s > 0\\ 0, & x = 0, s = 0\\ +\infty, & \text{all remaining cases} \end{cases}$$
 :
$$[x \in \mathbf{R}^n, s \in \mathbf{R}]$$

$$\operatorname{Epi}\{f\} = \{(t, x, s) : \begin{bmatrix} 2x \\ t - s \\ t + s \end{bmatrix} \ge_{\mathbf{L}^{n+2}} 0\}$$

5. The (branch of) hyperbola

$$\{(t,s) \in \mathbf{R}^2 : ts \ge 1, t, s \ge 0\}$$
:

$$\{(t,s): ts \ge 1, t, s \ge 0\} = \{(t,s): \begin{bmatrix} 2 \\ t-s \\ t+s \end{bmatrix} \ge_{\mathbf{L}^3} 0\}$$

Operations preserving CQ-representability of sets

- ♣ The following operations with sets preserve the CQr property, and, moreover, allow, given CQRs for the operands, to point out explicitly a CQR for the result:
- S.A. Intersection: If sets $X_i \subset \mathbf{R}^n$, i = 1, ..., N, are CQr, then so is their intersection:

$$X_{i} = \{x \in \mathbf{R}^{n} : A_{ij}(x, u^{i}) \geq_{\mathbf{L}^{m_{ij}}} 0, j = 1, ..., k_{i} \}$$

$$\downarrow \downarrow$$

$$\bigcap_{1}^{N} X_{i} = \{x \in \mathbf{R}^{n} : \alpha_{ij}(x, u) \geq_{\mathbf{L}^{m_{ij}}} 0, j = 1, ..., k_{i} \}$$

$$[u = (u^{1}, ..., u^{N}), \alpha_{ij}(x, u) = A_{ij}(x, u^{i})]$$

Corollary: A polyhedral set

$$\{x: Ax - b \ge 0\}$$

is CQr (as the intersection of closed half-spaces, which are CQr).

S.B. <u>Direct product</u>: If sets $X_i \subset \mathbb{R}^{n_i}$, i = 1, ..., N, are CQr, so is their direct product

$$X = X_1 \times X_2 \times ... \times X_N \subset \mathbf{R}^{n_1} \times ... \times \mathbf{R}^{n_N}$$
.

Indeed,

$$\begin{cases}
X_{i} = \{x^{i} \in R^{n_{i}} : \exists u^{i} : \mathcal{A}_{i}(x^{i}, y^{i}) \geq_{\mathbf{K}_{i}} 0\} \\
[\mathbf{K}_{i} : \mathbf{products of Lorentz cones}]
\end{cases}$$

$$X = \{x = (x^{1}, ..., x^{N}) : x^{i} \in X_{i}\}$$

$$= \{x = (x^{1}, ..., x^{N}) : \exists u = (u^{1}, ..., u^{N}) : \exists u = (u^{1},$$

 $[\mathbf{K} = \mathbf{K}_1 \times ... \times \mathbf{K}_N \text{ product of Lorentz cones}]$

S.C. Affine image ("Projection"): If $X \subset \mathbb{R}^n$ is CQr and $x \mapsto Ax + \overline{b : \mathbb{R}^n \to \mathbb{R}^k}$ is an affine mapping, then the set

$$X^{\rightarrow} = \{ y = Ax + b : x \in X \}$$

is CQr:

$$\begin{cases}
X = \{x : \exists u : \mathcal{A}(x, u) \geq_{\mathbf{K}} 0\} \\
X^{\rightarrow} = \{y : \exists x \in X : y = Ax + b\}
\end{cases}$$

$$\downarrow X^{\rightarrow} = \{y : \exists (x, u) : \mathcal{A}(x, u) \geq_{\mathbf{K}} 0, y = Ax + b\}$$

$$= \{y : \exists (x, u) : \mathcal{B}(y, (x, u)) \equiv \begin{cases}
\frac{A(x, u)}{0} \\
\frac{(y - Ax - b)_{1}}{0} \\
\frac{(Ax + b - y)_{1}}{0}
\end{cases}$$

$$\frac{(y - Ax - b)_{k}}{0} \\
\frac{(Ax + b - y)_{k}}{0}$$

$$[\mathbf{L} = \mathbf{K} \times \mathbf{L}^{2} \times ... \times \mathbf{L}^{2}]$$

Corollary: Let S be a finite system of linear and conic quadratic inequalities in variables (x, u). Then the set

$$X = \{x : \exists u : (x, u) \text{ solves } \mathcal{S}\}$$

is CQr.

Indeed, the set \mathcal{Y}_i of all solutions (x, u) to *i*-th inequality of \mathcal{S} clearly is CQr, whence the set $\bigcap_i Y_i$ also is CQr. It remains to note that X is the projection of $\bigcap_i Y_i$ onto the x-plane.

S.D. Inverse affine image ("Cross-section"): If $X \subset \mathbb{R}^n$ is CQr and $y \mapsto Ax + b : \mathbb{R}^k \to \mathbb{R}^n$ is an affine mapping, then the set

$$X^{\leftarrow} = \{ y : Ay + b \in X \}$$

is CQr:

$$\left\{ \begin{array}{l} X = \left\{ x : \exists u : \mathcal{A}(x, u) \geq_{\mathbf{K}} 0 \right\} \\ X^{\leftarrow} = \left\{ y : Ay + b \in X \right\} \end{array} \right\}$$

$$\qquad \qquad \qquad \downarrow$$

$$X^{\leftarrow} = \left\{ y : \exists u : \mathcal{A}(Ay + b, u) \geq_{\mathbf{K}} 0 \right\}$$

S.E. Arithmetic summation: If sets $X_i \subset \mathbb{R}^n$, i = 1, ..., N, are CQr, so is their arithmetic sum

$$X \equiv X_1 + ... + X_N = \{x = x_1 + ... + x_N : x_i \in X_i, i = 1, ..., N\}$$

Indeed, X is the image of the direct product $Y = X_1 \times ... \times X_N$ under the linear mapping

$$y \equiv (x_1, ..., x_N) \mapsto x_1 + ... + x_N,$$

and both operations preserve the CQr property.

S.F*. Taking polar: If a set $X \subset \mathbb{R}^n$, $0 \in X$, is CQr with a strictly feasible CQR:

$$X = \{x : \exists u : Ax + Bu + b \ge_{\mathbf{K}} 0\}, \quad \exists (\bar{x}, \bar{u}) : A\bar{x} + B\bar{u} + b >_{\mathbf{K}} 0\},$$

then the polar of X, that is, the set

$$Polar(X) = \{ \xi : \xi^T x \le 1 \quad \forall x \in X \}$$

is CQr:

$$\begin{aligned} \operatorname{Polar}(X) &= \left\{ \xi : \xi^T x \leq 1 \quad \forall x \in X \right\} \\ &= \left\{ \xi : \inf_{\underline{x}, u} \left\{ -\xi^T x : Ax + Bu + b \geq_{\mathbf{K}} 0 \right\} \geq -1 \right\} \\ &= \max_{\underline{y}} \left\{ -b^T y : A^T y + \xi = 0, B^T y = 0, y \geq_{\mathbf{K}} 0 \right\} \\ &= \left\{ \xi : \exists y : b^T y \leq 1, A^T y + \xi = 0, B^T y = 0, y \geq_{\mathbf{K}} 0 \right\} \end{aligned}$$

Thus, $\operatorname{Polar}(X)$ is the projection onto the ξ -plane of the CQr set

$$\underbrace{\{(\xi,y): y \ge_{\mathbf{K}} 0\}}_{\mathbf{CQr}} \cap \underbrace{\{(\xi,y): b^T y \le 1, A^T y + \xi = 0, B^T y = 0\}}_{\mathbf{polyhedral \ and \ thus \ CQr}$$

and as such is CQr itself.

S.G*. Taking conic hull: If a set $X \subset \mathbb{R}^n$ is CQr, then the conic hull of X – the set

$$X^+ \equiv \{(x,t) : t > 0, x/t \in X\}$$

is CQr:

$$X = \{x : \exists u : Ax + Bu + b \geq_{\mathbf{K}} 0\}$$

$$\downarrow \downarrow$$

$$X^{+} \equiv \{(x,t) : t > 0, x/t \in X\}$$

$$= \{((x,t) : \exists u : A(x/t) + Bu + b \geq_{\mathbf{K}} \mathbf{K}, t > 0\}$$

$$= \{(x,t) : \exists v : Ax + Bv + tb \geq_{\mathbf{K}} 0, t > 0\}$$

$$= \{(x,t) : \exists (v,s) : Ax + Bv + tb \geq_{\mathbf{K}} 0, t \geq 0, st \geq 1\}$$

$$= \{(x,t) : \exists (v,s) : \left[\frac{Ax + Bv + tb}{2}\right] \geq_{\mathbf{L}} 0,$$

$$[\mathbf{L} = \mathbf{K} \times \mathbf{L}^{3}]$$

Note: If X is closed, then the CQr set

$$\widetilde{X}^+ = \{(x,t) : \exists v : Ax + Bv + tb \ge_{\mathbf{K}} 0\} \cap \{(x,t) : t \ge 0\}$$

is "in-between" the conic hull of X and its closure:

$$X^+ \subset \widetilde{X}^+ \subset \operatorname{cl} X^+.$$

If X is closed and bounded, then $\widetilde{X}^+ = \operatorname{cl} X^+$, so that the closure of the convex hull of X is CQr.

S.H*. Taking convex hull of union: If $X_i \subset \mathbb{R}^n$, i = 1, ..., N, are closed CQr sets:

$$X_i = \{x^i : \exists u^i : A_i x^i + B_i u^i + b_i \ge_{\mathbf{K}_i} 0\},\$$

then the convex hull of their union

$$\widehat{X} = \operatorname{Conv}(\bigcup_{i=1}^{N} X_i) = \{ y = \sum_{i} \lambda_i x^i : x^i \in X_i, \lambda_i \ge 0, \sum_{i} \lambda_i = 1 \}$$

is "nearly CQr": given a CQR for X, one can explicitly point out a CQR for a set \widehat{X}' which is in-between \widehat{X} and $\operatorname{cl} \widehat{X}$. In particular, if \widehat{X} is closed, it is CQr.

Indeed,

$$\operatorname{Conv}(\bigcup_{i=1}^{N} X_{i})$$

$$= \left\{ y = \sum_{i} \underline{\lambda_{i} x^{i}} : x^{i} \in X_{i}, \lambda_{i} \geq 0, \sum_{i} \lambda_{i} = 1 \right\}$$

$$= \left\{ y : \exists \{y^{i}, \lambda_{i}\}_{i=1}^{N} : \begin{cases} y = \sum_{i} y^{i} \\ \lambda \geq 0 \\ \sum_{i} \lambda_{i} = 1 \\ \lambda_{i} > 0 \Rightarrow y^{i} / \lambda_{i} \in X_{i}, \\ \lambda_{i} = 0 \Rightarrow y^{i} = 0 \end{cases} \right\}$$

$$\subset \left\{ y : \exists \{y^{i}, \lambda_{i}, v^{i}\}_{i=1}^{N} : A_{i} y^{i} + B_{i} v^{i} + \lambda_{i} b_{i} \geq_{\mathbf{K}_{i}} 0, i = 1, ..., N \right\}$$

$$\lambda_{i} \geq 0, i = 1, ..., N \sum_{i} \lambda_{i} = 1$$

$$\widehat{X}'$$

 \widehat{X}' is the required CQr set.

Operations preserving CQ-representability of functions

- ♣ The following operations with functions preserve the CQr property, and, moreover, allow, given CQRs for the operands, to point out explicitly a CQR for the result:
- F.A. Taking maximum: If $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, i = 1, ..., N, are CQr functions, then so is their maximum

$$f(x) = \max_{i} f_i(x).$$

Indeed, $\operatorname{Epi}\{f\} = \bigcap_{i} \operatorname{Epi}\{f_i\}$.

F.B. Summation with nonnegative weights: If functions f_i : $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, i = 1, ..., N, are CQr and $\alpha_i \geq 0$, then the function

$$f(x) = \sum_{i=1}^{n} \alpha_i f_i(x)$$

is CQr:

F.C. <u>Direct summation</u>: If $f_i : \mathbf{R}^{n_i} \to \mathbf{R} \cup \{\infty\}$, i = 1, ..., N, are CQr, so is the function

$$f(x^1, ..., x^N) = \sum_{i=1}^N f_i(x^i) : \mathbf{R}_{x^1}^{n_1} \times ... \times \mathbf{R}_{x^N}^{n_N} \to \mathbf{R} \cup \{\infty\}.$$

F.D. Affine substitution of argument: If

$$f: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$$

is CQr and $y\mapsto Ay+b:\mathbf{R}^k\to\mathbf{R}^n$ is an affine mapping, then the superposition

$$f^{\rightarrow}(y) = f(Ay + b)$$

is CQr.

Indeed,

$$\operatorname{Epi}\{f^{\to}\} = \{(t,y) : t \ge f(Ay+b)\}$$
$$= \left\{ (t,y) : \mathcal{A}(t,y) \equiv \begin{bmatrix} t \\ Ay+b \end{bmatrix} \in \operatorname{Epi}\{f\} \right\}$$

F.E. Partial minimization: If $f(\xi, \eta) : \mathbf{R}_{\xi}^{n_1} \times \mathbf{R}_{\eta}^{n_2} \to \mathbf{R} \cup \{+\infty\}$ is CQr and the minimum in the parametric problem

$$\min_{\eta} f(\xi, \eta),$$

is achieved for every ξ , then the function

$$\widehat{f}(\xi) = \min_{\eta} f(\xi, \eta)$$

is CQr. Indeed, under our assumption

$$\operatorname{Epi}\{\widehat{f}\} = \{(t,\xi) : \exists \eta : t \ge f(\xi,\eta)\}\$$

that is, $\operatorname{Epi}\{\widehat{f}\}\$ is a projection of $\operatorname{Epi}\{f\}.$

F.G*. Taking Legendre transformation: If

$$f: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$$

is CQr with a strictly feasible CQR

$$\{(t,x): t \ge f(x)\} = \{(t,x): \exists u: at + Ax + Bu + b \ge_{\mathbf{K}} 0\}$$

then the Legendre transformation of f

$$f_*(\xi) = \sup_{x} \left[\xi^T x - f(x) \right]$$

is CQr:

$$\{(\tau, \xi) : \tau \geq f_*(\xi)\}$$

$$= \{(\tau, \xi) : \tau \geq \xi^T x - t \forall (t, x) \in \operatorname{Epi}\{f\}\}$$

$$= \{(\tau, \xi) : \tau \geq \sup_{(t, x) \in \operatorname{Epi}\{f\}} [\xi^T x - t]\}$$

$$= \left\{ (\tau, \xi) : \tau \geq \sup_{\underbrace{x, t, u}} \{\xi^T x - t : ta + Ax + Bu + b \geq_{\mathbf{K}} 0\} \right\}$$

$$= \left\{ (\tau, \xi) : \exists y : \begin{cases} a^T y = 1 \\ A^T y + \xi = 0 \\ B^T y = 0 \\ \tau - b^T y \geq 0 \\ y \geq_{\mathbf{K}} 0 \end{cases} \right\}$$

F.H. Theorem on superposition: If functions $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, i = 1, ..., N, are CQr:

$$\{(t,x): t \ge f_i(x)\} = \{(t,x): \exists u^i: ta_i + A_i x + B_i u^i + b_i \ge_{\mathbf{K}_i} 0\}$$

and a function $g: \mathbf{R}^N \to \mathbf{R} \cup \{\infty\}$ is \geq -nonincreasing and CQr:

$$\{(t,y): t \ge g(y)\} = \{(t,y): \exists v: ta + Ay + Bv + b \ge_{\mathbf{K}} 0\},\$$

then the superposition

$$h(x) = g(f_1(x), ..., f_N(x)) : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$$

is CQr:

$$\{(t,x): t \ge h(x)\}$$

$$= \left\{ (t,x): \exists \{t_i, u^i\}_{i=1}^N, v: \begin{cases} \underbrace{t_i a_i + A_i x + B_i u^i + b_i \ge_{\mathbf{K}_i} 0}_{\text{"says" that } t_i \ge f_i(x)} \\ t_1 \\ \vdots \\ t_N \end{bmatrix} + Bv + b \ge_{\mathbf{K}} 0 \end{cases} \right\}$$

$$= \left\{ (t,x): \exists \{t_i, u^i\}_{i=1}^N, v: \begin{cases} \underbrace{t_i a_i + A_i x + B_i u^i + b_i \ge_{\mathbf{K}_i} 0}_{\text{"says" that } t \ge g(t_1, \dots, t_N)} \end{cases} \right\}$$

More examples of CQr functions/sets

6. Convex quadratic form $f(x) = x^T Q^T Q x + q^T x + r$ is CQr, since it can be obtained from the squared Euclidean norm and affine function (both are CQr) by affine substitution of argument and addition. Here is an explicit CQR for f:

$$\operatorname{Epi}\{f\} = \{(x,t) : \begin{bmatrix} 2Qx \\ t - q^Tx - r - 1 \\ t - q^Tx - r + 1 \end{bmatrix} \ge_{\mathbf{L}^{m+2}} 0\} \qquad [Q : m \times n]$$

7. The cone

$$K = \{(x, \sigma_1, \sigma_2) \in \mathbf{R}_x^n \times \mathbf{R}_{\sigma_1, \sigma_2}^2 : \sigma_1, \sigma_2 \ge 0, x^T x \le \sigma_1 \sigma_2\}$$

is CQr (this is just affine image of the Lorentz cone): Indeed,

$$K = \{(x, \sigma_1, \sigma_2) \in \mathbf{R}_x^n \times \mathbf{R}_{\sigma_1, \sigma_2}^2 : \begin{bmatrix} 2x \\ \sigma_1 - \sigma_2 \\ \sigma_1 + \sigma_2 \end{bmatrix} \ge_{\mathbf{L}^{n+2}} 0\}.$$

Observation: Let m be nonnegative integer and let $M=2^m$. The set

$$\mathcal{X}_m = \{(t, x_1, x_2, ..., x_M) \in \mathbf{R}_+^{M+1} : t^M \le x_1 ... x_M \}$$

is CQr. Indeed,

$$\mathcal{X}_m = \{(t, x_1, ..., x_M) \ge 0 : \exists y_{ij} \ge 0 :$$

$$y_{1,1}^{2} \leq x_{1}x_{2}, y_{1,2}^{2} \leq x_{3}x_{4}, \dots, y_{1,M/2}^{2} \leq x_{M-1}x_{M}$$

$$y_{2,1}^{2} \leq y_{1,1}y_{1,2}, \dots, y_{2,M/4}^{2} \leq y_{1,M/2-1}y_{1,M/2}$$

$$\vdots$$

$$t^{2} \leq y_{m-1,1}y_{m-1,2}$$

8. Convex increasing power function $f(x) = x_+^{\pi}$ with rational degree $\pi = \frac{p}{q} \ge 1$ is CQr.

Indeed, let m be such that $M \equiv 2^m \ge p + q$. We have

- 9. Convex even power function $f(x) = |x|^{\pi}$ with rational degree $\pi \ge 1$ is CQr.
- 10. Decreasing power function

$$f(x) = \begin{cases} x^{-\pi} & , x > 0 \\ +\infty & , x \le 0 \end{cases}$$

of rational degree $-\pi < 0$ is CQr.

11. The hypograph of a concave monomial

$$f(x) = x_1^{\pi_1} ... x_n^{\pi_n} : \mathbf{R}_+^n \to \mathbf{R}$$

with positive rational $\pi_1, ..., \pi_n, \sum_i \pi_i \leq 1$, is CQr.

12. The epigraph of a convex monomial

$$f(x) = x_1^{-\pi_1} ... x_n^{-\pi_n} : \mathbf{R}_+^n \to \mathbf{R} \cup \{+\infty\}$$

with positive rational $\pi_1, ..., \pi_n$ is CQr.

13. The epigraph of the $\|\cdot\|_p$ -norm

$$||x||_p = \begin{cases} \left(\sum_i |x_i|^p\right)^{1/p}, & 1 \le p < \infty \\ \max_i |x_i|, & p = \infty \end{cases}$$

is CQr, provided $p \in [1, \infty]$ is rational (or $p = \infty$).

Semidefinite Programming

Preliminaries: Spaces of Matrices and Semidefinite cone

• The space $\mathbf{M}^{m,n}$ of $m \times n$ matrices can be identified with \mathbf{R}^{mn}

$$A = [a_{ij}]_{\substack{i=1,...,m\\j=1,...,n}}$$

$$\updownarrow$$

$$Vec(A) = (a_{11},...,a_{1n},a_{21},...,a_{2n},...,a_{m1},...,a_{mn})^T$$

The inner product of matrices induced by this representation is

$$\langle A, B \rangle \equiv \sum_{i,j} A_{ij} B_{ij} = \text{Tr}(A^T B) \quad [A, B \in \mathbf{M}^{m,n}]$$

$$\left[\text{Tr}(C) = \sum_{i=1}^n C_{ii}, C \in \mathbf{M}^{n,n}, \text{ is the trace of } C \right]$$

• In particular, the space S^m of $m \times m$ symmetric matrices equipped with the inner product inherited from $M^{m,m}$:

$$\langle A, B \rangle \equiv \sum_{i,j} A_{ij} B_{ij} = \text{Tr}(A^T B) = \text{Tr}(AB)$$

is a Euclidean space (dim $S^m = \frac{m(m+1)}{2}$).

• The <u>positive semidefinite</u> symmetric $m \times m$ matrices form a cone (closed, convex, pointed and with a nonempty interior) in S^m :

$$\mathbf{S}_{+}^{m} = \left\{ A \in \mathbf{S}^{m} \mid \xi^{T} A \xi \ge 0 \ \forall \xi \in \mathbf{R}^{m} \right\}$$

$$\mathbf{S}_{+}^{m} = \left\{ A \in \mathbf{S}^{m} \mid \xi^{T} A \xi \ge 0 \ \forall \xi \in \mathbf{R}^{m} \right\}$$

- Equivalent descriptions of S_{+}^{m} : an $m \times m$ matrix A is positive semidefinite
- * if and only if A is symmetric $(A = A^T)$ and all its eigenvalues are nonnegative;
 - * if and only if A can be decomposed as $A = D^T D$
- \ast if and only A can be represented as a sum of symmetric dyadic matrices:

$$A = \sum_{j} d_{j} d_{j}^{T};$$

- * if and only if $A = U^T \Lambda U$ with orthogonal U and diagonal Λ , the diagonal entries of Λ being nonnegative;
- * if and only if A is symmetric $(A = A^T)$ and all principal minors of A are nonnegative.
- As every cone, S_{+}^{m} defines a "good" partial ordering on S^{m} :

$$A \succeq B \Leftrightarrow A - B \succeq 0 \Leftrightarrow \xi^T A \xi \ge \xi^T B \xi \qquad \forall \xi$$
$$\left[A = A^T, B = B^T \text{ are of the same size} \right]$$

• <u>Useful observation</u>: It is possible to multiply both sides of a \succeq -inequality from the right and from the left by a matrix and its transpose:

$$A \succeq B \Rightarrow Q^T A Q \succeq Q^T B Q$$
 $\left[A, B \in \mathbf{S}^m, Q \in \mathbf{M}^{m,k} \right]$

Indeed,

$$\left\{ \xi^T A \xi \ge \xi^T B \xi \ \forall \xi \right\} \Rightarrow \left\{ \eta^T Q^T A \underbrace{Q \eta}_{\xi} \ge \eta^T Q^T B Q \eta \ \forall \eta \right\}$$

• Observation: The semidefinite cone is self-dual:

$$\left(\mathbf{S}_{+}^{m}\right)_{*} \equiv \left\{A \in \mathbf{S}^{m} \mid \operatorname{Tr}(AB) \geq 0 \ \forall B \in \mathbf{S}_{+}^{m}\right\} = \mathbf{S}_{+}^{m}.$$

Indeed,

$$\xi^T A \xi = \operatorname{Tr}(\xi^T A \xi) = \operatorname{Tr}(A \xi \xi^T)$$

$$PQ \text{ makes sense and is square } \downarrow \downarrow$$

$$\operatorname{Tr}(PQ) = \operatorname{Tr}(QP)$$

It follows that if $A \in \mathbf{S}^m$ is such that $\text{Tr}(AB) \geq 0$ for all $B \succeq 0$, then $A \succeq 0$:

$$\xi \in \mathbf{R}^m \Rightarrow B = \xi \xi^T \succeq 0 \Rightarrow \operatorname{Tr}(AB) = \xi^T A \xi \ge 0$$

Vice versa, if $A \in \mathbf{S}_{+}^{m}$, then $\operatorname{Tr}(AB) \geq 0$ for all $B \succeq 0$:

$$B \succeq 0$$

$$\downarrow \\ B = \sum_{j} d_{j} d_{j}^{T}$$

$$\downarrow \\ Tr(AB) = \sum_{j} Tr(Ad_{j} d_{j}^{T}) = \sum_{j} d_{j}^{T} Ad_{j} \geq 0.$$

Semidefinite program

• A <u>semidefinite program</u> is a conic program associated with the <u>semidefinite</u> cone:

$$\min_{x} \left\{ c^{T}x \mid \mathcal{A}x - B \succeq 0 \quad \left[\Leftrightarrow \mathcal{A}x - B \geq_{\mathbf{S}_{+}^{m}} 0 \right] \right\}$$
$$\left[\mathcal{A}x = \sum_{i=1}^{\infty} x_{i} A_{i}, A_{i} \in \mathbf{S}^{m} \right]$$

A constraint of the type

$$x_1A_1 + \ldots + x_nA_n \succeq B$$

with variables $x_1, ..., x_n$ is called an <u>LMI</u> – Linear Matrix Inequality. Thus, a semidefinite program is to minimize a linear objective under an LMI constraint.

• Observation: A system of LMI constraints

$$\begin{cases} A_1 x - B_1 &\equiv x_1 A_1^1 + \dots + x_n A_1^n - B_1 \succeq 0 \\ A_2 x - B_2 &\equiv x_1 A_2^1 + \dots + x_n A_2^n - B_2 \succeq 0 \\ \dots \\ A_k x - B_k &\equiv x_1 A_k^1 + \dots + x_n A_k^n - B_k \succeq 0 \end{cases}$$

is equivalent to the single LMI constraint

Diag
$$(A_1x - B_1, ..., A_kx - B_k)$$

$$= \begin{pmatrix} A_1x - B_1 & & & \\ \hline & A_2x - B_2 & & \\ \hline & & & \ddots & \\ \hline & & & & A_kx - B_k \end{pmatrix} \succeq 0.$$

Program dual to an SDP program

$$\min_{x} \left\{ c^{T}x \mid \mathcal{A}x - B \equiv \sum_{j=1}^{n} x_{j}A_{j} - B \succeq 0 \right\}$$
 (SDPr)

According to our general scheme, the problem dual to (SDPr) is

$$\max_{Y} \{ \langle B, Y \rangle \mid \mathcal{A}^*Y = c, Y \succeq 0 \}$$
 (SDDI)

(recall that S^m_+ is self-dual!).

It is easily seen that the operator \mathcal{A}^* conjugate to \mathcal{A} is given by

$$\mathcal{A}^*Y = (\operatorname{Tr}(YA_1), ..., \operatorname{Tr}(YA_n))^T : \mathbf{S}^m \to \mathbf{R}^n.$$

Consequently, the dual problem is

$$\max_{Y} \{ \text{Tr}(BY) \mid \text{Tr}(YA_i) = c_i, i = 1, ..., n, Y \succeq 0 \}$$
 (SDDI)

SDP optimality conditions

$$\min_{x} \left\{ c^{T}x \mid \mathcal{A}x - B \equiv \sum_{j=1}^{n} x_{j}A_{j} - B \succeq 0 \right\}$$
 (SDPr)

$$\max_{Y} \{ \text{Tr}(BY) \mid \text{Tr}(A_{j}Y) = c_{j}, \ j = 1, ..., n; \ Y \succeq 0 \}$$
 (SDDI)

- Assume that
- (!) both (SDPr) and (SDDl) are strictly feasible, so that by Conic Duality Theorem both problems are solvable with equal optimal values.

By Conic Duality, the necessary and sufficient condition for a primal-dual feasible pair (x,Y) to be primal-dual optimal is that

$$\operatorname{Tr}(\underbrace{[\mathcal{A}x - B]}_{\text{"primal slack"}X}Y) = 0$$

• For a pair of symmetric positive semidefinite matrices X and Y,

$$Tr(XY) = 0 \Leftrightarrow XY = YX = 0.$$

Thus, under assumption (!) a primal-dual feasible pair (x, Y) is primal-dual optimal if and only if

$$[\mathcal{A}x - B]Y = Y[\mathcal{A}x - B] = 0$$

$$\min_{x} \left\{ c^{T}x \mid \mathcal{A}x - B \equiv \sum_{j=1}^{n} x_{j}A_{j} - B \succeq 0 \right\}$$
 (SDPr)

$$\max_{Y} \left\{ \text{Tr}(BY) \mid \text{Tr}(A_{j}Y) = c_{j}, \ j = 1, ..., n; \ Y \succeq 0 \right\}$$
 (SDDI)

- Assume that
 - (!) both (SDPr) and (SDDl) are strictly feasible.

Then a primal-dual feasible pair (x,Y) is primal-dual optimal if and only if

$$[\mathcal{A}x - B]Y = Y[\mathcal{A}x - B] = 0$$

Cf. Linear Programming:

(P):
$$\min_{x} \{c^{T}x \mid Ax - b \ge 0\}$$

(D): $\max_{y} \{b^{T}y \mid A^{T}y = c, y \ge 0\}$

(x,y) primal-dual optimal



(x,y) primal-dual feasible and $y_j[Ax-b]_j=0$ $\forall j$

What can be expressed via SDP?

$$\min_{x} \left\{ c^{T} x \mid x \in X \right\} \tag{Ini}$$

• A sufficient condition for (Ini) to be equivalent to an SD program is that X is a SDr ("Semidefinite-representable") set:

<u>Definition.</u> A set $X \subset \mathbb{R}^n$ is called SDr, if it admits SDR ("Semidefinite representation")

$$X = \{x \mid \exists u : \mathcal{A}(x, u) \succeq 0\}$$
$$\left[\mathcal{A}(x, u) = \sum_{j} x_{j} A_{j} + \sum_{\ell} u_{\ell} B_{\ell} + C : \mathbf{R}_{x}^{n} \times \mathbf{R}_{u}^{k} \to \mathbf{S}^{m}\right]$$

• Given a SDR of X, we can write down (Ini) equivalently as the semidefinite program

$$\min_{x,u} \left\{ c^T x \mid \mathcal{A}(x,u) \succeq 0 \right\}.$$

- Same as in the case of Conic Quadratic Programming, we can
 - 1. Define the notion of a SDr function

$$f: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$$

as a function with SDr epigraph:

$$\{(t,x) \mid t \ge f(x)\} = \left\{ (t,x) \mid \exists u : \underbrace{\mathcal{A}(t,x,u) \succeq 0}_{\mathbf{LMI}} \right\}$$

and verify that if f is a SDr function, then all its level sets

$$\{x \mid f(x) \le a\}$$

are SDr;

2. Develop a "calculus" of SDr functions/sets with <u>exactly</u> the same combination rules as for CQ-representability.

When a function/set is SDr?

Proposition. Every CQr set/function is SDr as well.

Proof.

- 1⁰. Lemma. Every direct product of Lorentz cones is SDr.
- **2**⁰ Lemma \Rightarrow Proposition: Let $X \subset \mathbb{R}^n$ be CQr:

$$X = \{x \mid \exists u : A(x, u) \in \mathbf{K}\},\$$

K being a direct product of Lorentz cones and A(x,u) being affine.

By Lemma,

$$\mathbf{K} = \{ y \mid \exists v : \mathcal{B}(y, v) \succeq 0 \}$$

with affine $\mathcal{B}(\cdot,\cdot)$. It follows that

$$X = \left\{ x \mid \exists u, v : \underbrace{\mathcal{B}(A(x, u), v) \succeq 0}_{\mathbf{LMI}} \right\},\,$$

which is a SDR for X.

Lemma. Every direct product of Lorentz cones is SDr.

<u>Proof.</u> It suffices to prove that a Lorentz cone L^m is a SDr set (since SD-representability is preserved when taking direct products).

To prove that L^m is SDr, let us make use of the following Lemma on Schur Complement. A symmetric block matrix

$$A = \left(\frac{P \mid Q^T}{Q \mid R}\right)$$

with positive definite R is positive (semi)definite if and only if the matrix

$$P - Q^T R^{-1} Q$$

is positive (semi)definite.

LSC⇒Lemma: Consider the linear mapping

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} \mapsto \mathcal{A}x = \begin{pmatrix} x_m & x_1 & x_2 & x_3 & \dots & x_{m-1} \\ \hline x_1 & x_m & & & & \\ x_2 & & x_m & & & \\ x_3 & & & x_m & & \\ \vdots & & & & \ddots & \\ x_{m-1} & & & & x_m \end{pmatrix}$$

We claim that

$$\mathbf{L}^m = \{ x \mid \mathcal{A}(x) \succeq 0 \}.$$

Indeed,

$$\mathbf{L}^{m} = \left\{ x \in \mathbf{R}^{m} \mid x_{m} \ge \sqrt{x_{1}^{2} + \dots + x_{m-1}^{2}} \right\}$$

and therefore

• if $x \in \mathbf{L}^m$ is nonzero, then $x_m > 0$ and

$$x_m - \frac{x_1^2 + x_2^2 + \dots + x_{m-1}^2}{x_m} \ge 0$$

so that $A(x) \succeq 0$ by LSC. If x = 0, then $A(x) = 0 \succeq 0$.

• if $A(x) \succeq 0$ and $A(x) \neq 0$, then $x_m > 0$ and, by LSC,

$$x_m - \frac{x_1^2 + x_2^2 + \dots + x_{m-1}^2}{x_m} \ge 0 \Rightarrow x \in \mathbf{L}^m.$$

And if A(x) = 0, then $x = 0 \in \mathbf{L}^m$.

• Lemma on Schur Complement. A symmetric block matrix

$$A = \begin{pmatrix} P & Q^T \\ Q & R \end{pmatrix}$$

with positive definite R is positive (semi)definite if and only if the matrix

$$P - Q^T R^{-1} Q$$

is positive (semi)definite.

<u>Proof.</u> A is $\succeq 0$ if and only if

$$\inf_{v} \begin{pmatrix} u \\ v \end{pmatrix}^{T} \begin{pmatrix} P & Q^{T} \\ Q & R \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \ge 0 \quad \forall u. \tag{*}$$

When $R \succ 0$, the left hand side inf can be easily computed and turns to be

$$u^T(P - Q^T R^{-1}Q)u.$$

Thus, (*) is valid if and only if

$$u^T(P - Q^T R^{-1}Q)u \ge 0 \quad \forall u,$$

i.e., if and only if

$$P - Q^T R^{-1} Q \succeq 0.$$

More examples of SD-representable functions/sets

• The largest eigenvalue $\lambda_{\max}(X)$ regarded as a function of $m \times m$ symmetric matrix X is SDr:

$$\lambda_{\max}(X) \le t \quad \Leftrightarrow \quad tI_m - X \succeq 0,$$

 I_k being the unit $k \times k$ matrix.

• The largest eigenvalue of a matrix pencil. Let $M, A \in \mathbf{S}^m$ be such that $M \succ 0$.

The eigenvalues of the pencil [M,A] are reals λ such that the matrix $\lambda M - A$ is singular, or, equivalently, such that

$$\exists e \neq 0 : Ae = \lambda Me.$$

The eigenvalues of the pencil [M,A] are the usual eigenvalues of the symmetric matrix $D^{-1}AD^{-T}$, where D is such that $M=DD^T$.

The largest eigenvalue $\lambda_{\max}(X:M)$ of a pencil [M,X] with $M\succ 0$, regarded as a function of X, is SDr:

$$\lambda_{\max}(X:M) \le t \quad \Leftrightarrow \quad tM - X \succeq 0.$$

• Sum of k largest eigenvalues. For a symmetric $m \times m$ matrix X, let $\lambda(X)$ be the vector of eigenvalues of X taken with their multiplicities in the non-ascending order:

$$\lambda_1(X) \ge \lambda_2(X) \ge \dots \ge \lambda_m(X),$$

and let $S_k(X)$ be the sum of k largest eigenvalues of X:

$$S_k(X) = \sum_{i=1}^k \lambda_i(X) \qquad [1 \le k \le m]$$
$$[S_1(X) = \lambda_{\max}(X); \ S_m(X) = \operatorname{Tr}(X)]$$

The functions $S_k(X)$ are SDr:

$$S_k(X) \le t \Leftrightarrow \exists s, Z : \begin{cases} (a) & ks + \text{Tr}(Z) \le t \\ (b) & Z \succeq 0 \\ (c) & X \preceq Z + sI_m \end{cases}$$

- We should prove that
- (i) If a pair X, t can be extended, by properly chosen s, Z, to a solution of (a) (c), then $S_k(X) \leq t$;
- (ii) If $S_k(X) \leq t$, then the pair X, t can be extended by properly chosen s, Z, to a solution of (a) (c).
- (i): We use the following

Basic Fact: The vector $\lambda(X)$ is a \succeq -monotone function of $X \in \mathbf{S}^m \colon X \succeq X' \Rightarrow \lambda(X) \geq \lambda(X')$.

Let (X, t, s, Z) solve (a) - (c). Then

$$X \leq Z + sI_m$$
 [by (c)]
 $\Rightarrow \lambda(X) \leq \lambda(Z + sI_m) = \lambda(Z) + s \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ [by Basic Fact]
 $\Rightarrow S_k(X) \leq S_k(Z) + sk$
 $\Rightarrow S_k(X) \leq \text{Tr}(Z) + sk$ [since $S_k(Z) \leq \text{Tr}(Z)$]
 $\Rightarrow S_k(X) \leq t$ [by (a)]

(ii): Let $S_k(X) \leq t$, and let $X = U \operatorname{Diag}\{\lambda\}U^T$, $\lambda = \lambda(X)$, be the eigenvalue decomposition of X.

$$s = \lambda_k, \quad Z = U \begin{bmatrix} \lambda_1 - \lambda_k & & & \\ & \ddots & & \\ & & \lambda_{k-1} - \lambda_k & \\ & & 0 & \\ & & & 0 \end{bmatrix} U^T,$$

$$\underbrace{\text{Diag}\{\lambda(Z)\}}$$

we have

$$Z \succeq 0,$$

$$\operatorname{Diag}\{\lambda(X)\} \leq \operatorname{Diag}\{\lambda(Z)\} + s \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \Rightarrow X \preceq Z + sI_m,$$

$$t \geq S_k(X) = ks + \operatorname{Tr}(Z),$$

so that (t, X, s, Z) solves the system of LMIs

$$(a) ks + Tr(Z) \le t$$

$$(b)$$
 $Z \succ 0$

$$\begin{array}{ccc} (a) & ks + \mathrm{Tr}(Z) & \leq & t \\ (b) & Z & \succeq & 0 \\ (c) & X & \preceq & Z + sI_m \end{array}$$

• Norm of rectangular matrix. Let X be a $m \times n$ matrix. Its operator norm

$$||X|| = \max_{\|\xi\|_2 \le 1} ||X\xi||_2$$

is SDr:

$$t \ge ||X|| \quad \Leftrightarrow \quad \begin{pmatrix} tI_n & X^T \\ X & tI_m \end{pmatrix} \succeq 0.$$

More generally, let

$$\sigma_i(X) = \sqrt{\lambda_i(X^T X)}$$

be the singular values of a rectangular matrix X. Then

- The sum of k largest singular values $\Sigma_k(X) = \sum_{i=1}^k \sigma_i(X)$ is a SDr function of $X \in \mathbf{M}^{m,n}$.
- SDr of symmetric monotone functions of singular values. Let $f(\lambda): \mathbf{R}^n_+ \to \mathbf{R} \cup \{\infty\}$ be a symmetric w.r.t. permutations of coordinates and \geq -nondecreasing SDr function. Then the function

$$F(X) = f(\sigma(X)) : \mathbf{M}^{m,n} \to \mathbf{R} \cup \{\infty\}$$

is SDr.

In particular, the functions

$$|X|_{\pi} = ||\sigma(X)||_{\pi}$$

with rational $\pi \in [1, \infty)$ are SDr with explicit SDR's.

Nonnegative polynomials

- \clubsuit For every positive integer k, the following sets are SDr:
- The set $P_{2k}^+(\mathbf{R})$ of coefficients of algebraic polynomials of degree $\leq 2k$ which are nonnegative on the entire axis:

$$P_{2k}^{+} = \left\{ p = (p_0, ..., p_{2k})^T : \exists Q = [Q_{ij}]_{i,j=0}^k \in \mathbf{S}_{+}^{k+1} : p_{\ell} = \sum_{i+j=\ell} Q_{ij}, \ \ell = 0, 1, ..., 2k \right\}$$

- The set $P_k^+(\mathbf{R}_+)$ of coefficients of algebraic polynomials of degree $\leq k$ which are nonnegative on the nonnegative ray \mathbf{R}_+
- The set $P_k^+([0,1])$ of coefficients of algebraic polynomials of degree $\leq k$ which are nonnegative on the segment [0,1]
- The set $T_k^+(\Delta)$ of coefficients of trigonometric polynomials of degree $\leq k$ which are nonnegative on a given segment $\Delta \in [-\pi,\pi]$.
- \clubsuit As a corollary, for every segment $\Delta \subset \mathbb{R}$ and every positive integer k, the function

$$f(p) = \min_{t \in \Delta} p(t)$$

of the vector p of coefficients of an algebraic (or a trigonometric) polynomial $p(\cdot)$ of degree $\leq k$ is SDr.

Basic Theorem [Nesterov '97]: The set $P_{2k}^+(R)$ of (coefficients of) polynomials of degree $\leq 2k$ nonnegative on the entire axis is SDr: it is the image of the cone S_+^{k+1} under the linear mapping \mathcal{A}

$$X \mapsto p_X(t) \equiv \sum_{i,j=0}^k X_{ij} t^{i+j}$$

$$\updownarrow \qquad [X \in S^{k+1}]$$

$$X \mapsto x = \operatorname{Coef}(p_X) : x_{\ell} = \sum_{i+j=\ell} X_{ij}$$

Proof. We should prove that

- (i) If $X \succeq 0$, then the polynomial $p_X(t)$ is nonnegative on R. $\mathcal{A}S_+^{k+1} \subset P_{2k}^+(R)$
- (ii) Vice versa, if a polynomial p of degree $\leq 2k$ is nonnegative on R, then there exists $X \succeq 0$ such that $p = p_X$. $P_{2k}^+(R) \subset \mathcal{A}S_+^{k+1}$.

Proof of (i): If $X \succeq 0$, then $X = \sum_{\ell=1}^{k+1} b_{\ell} b_{\ell}^{T}$ with $b_{\ell} \in \mathbb{R}^{k+1}$, so that

$$p_{X}(t) = \sum_{\ell=1}^{k+1} p_{b_{\ell}b_{\ell}^{T}}(t)$$

$$= \sum_{\ell=1}^{k+1} \left(\sum_{i,j=0}^{k} (b_{\ell})_{i} (b_{\ell})_{j} t^{i+j} \right)$$

$$= \sum_{\ell=1}^{k+1} \left[\sum_{i=0}^{k} (b_{\ell})_{i} t^{i} \right]^{2}$$

$$> 0 \quad \forall t.$$

$$\mathcal{A}: S^{k+1} \Rightarrow \mathbb{R}^{2k+1}$$

$$X \in S^{k+1}: \mathcal{A}(X) = \operatorname{Coef}\left(\sum_{j=0}^{k} \sum_{\ell=0}^{k} X_{ij} t^{i+j}\right)$$

Theorem

$$P_{2k}^+(R) = \mathcal{A}S_+^{k+1} =: \{\mathcal{A}X : X \in S_+^{k+1}\}$$

<u>Proof of (ii)</u>: We should prove that if a polynomial p of degree $\leq 2k$ is nonnegative on R, then there exists $X \succeq 0$ such that $p = p_X$. (Coef. of $t^{\ell} = \sum_{i+j=\ell} X_{ij}$)

We have seen that the polynomials which are sums of squares:

$$p(t) = \sum_{\ell=1}^{N} \left[\sum_{i=0}^{k} (b_{\ell})_{i} t^{i} \right]^{2}$$

can be obtained as p_X with $X \succeq 0$ (one can take $X = \sum_{\ell} b_{\ell} b_{\ell}^T$).

It remains to note that

(!) An algebraic polynomial on the axis is nonnegative if and *only if* it is a sum of squares (in fact, just *two* squares!). "spectral factorization"

Indeed, let p be a nontrivial polynomial nonnegative on the entire axis. Then the degree of p is even, its leading coefficient π is positive and we can factorize p as

$$p(t) = \pi \left[(t - \lambda_1)^{m_1} \dots (t - \lambda_r)^{m_r} \right]_1 \left[\prod_{\ell=1}^s \left[(t - \mu_\ell)(t - \bar{\mu}_\ell) \right] \right]_2,$$

where $\lambda_1, \ldots, \lambda_r$ are real roots of p, and $(\mu_\ell, \bar{\mu}_\ell)$ are conjugate pairs of complex roots of p.

Every factor $(t - \mu_{\ell})(t - \bar{\mu}_{\ell}) = (t - \Re \mu_{\ell})^2 + (\Im \mu_{\ell})^2$ is a sum of squares. Since p is nonnegative, its real roots are of even multiplicity each, so that $[\cdot]_1$ is a product of squares. Thus, p is a product of sums of squares, i.e., is itself a sum of squares. [This is due to "Liouville identity":

$$(a^{2} + b^{2})(c^{2} + d^{2}) = (ad + bc)^{2} + (ac - bd)^{2}$$

Complexity bounds for \mathcal{LP}_b

 \clubsuit A program from \mathcal{LP}_b :

$$(p) : \min_{x} \{c^{T}x \mid Ax \ge b, ||x||_{\infty} \le R\} \quad [A \in \mathbf{M}^{m,n}]$$

can be solved within accuracy ϵ in

$$\mathcal{N}_{\mathcal{LP}} = O(1)\sqrt{m}\ln\left(\frac{\|\mathrm{Data}(p)\|_1 + \epsilon^2}{\epsilon}\right)$$

iterations.

The computational effort per iteration is dominated by the necessity, given a positive definite diagonal matrix Δ and a vector r, to assemble the matrix

$$H = [A; I; -I] \Delta \begin{pmatrix} A^T \\ I \\ -I \end{pmatrix}$$

and to solve the linear system

$$Hu = r$$
.

• In the case m = O(n), the overall complexity of solving (p) within accuracy ϵ is cubic in n:

$$O(1)mn^2 \ln \left(\frac{\|\mathrm{Data}(p)\|_1 + \epsilon^2}{\epsilon} \right)$$

Complexity bounds for \mathcal{CQP}_b

\clubsuit A program from \mathcal{CQP}_b :

(p):
$$\{c^T x \mid ||D_i x - d_i||_2 \le e_i^T x - c_i, i = 1, ..., k; ||x||_2 \le R\}$$

can be solved within accuracy ϵ in

$$\mathcal{N}_{\mathcal{CQP}} = O(1)\sqrt{k} \ln \left(\frac{\|\mathrm{Data}(p)\|_1 + \epsilon^2}{\epsilon} \right)$$

iterations.

The computational effort per iteration is dominated by the necessity, given vectors δ_i , i = 1, ..., k and a vector r, to assemble the matrices

$$H_i = D_i^T (I - \delta_i \delta_i^T) D_i, \ i = 1, ..., k$$

and to solve a $\dim x \times \dim x$ linear system

$$Hu = r$$

with positive definite matrix H "readily given" by $H_1, ..., H_k$.

Complexity bounds for \mathcal{SDP}_b

\clubsuit A program from \mathcal{SDP}_b :

(p):
$$\min_{x} \left\{ c^T x \mid \mathcal{A}(x) = \sum_{i=1}^{n} x_i A_i - B \succeq 0, \|x\|_2 \le R \right\}$$

can be solved within accuracy ϵ in

$$\mathcal{N}_{\mathcal{SDP}} = O(1)\sqrt{\mu} \ln \left(\frac{\|\mathrm{Data}(p)\|_1 + \epsilon^2}{\epsilon} \right)$$

iterations, where μ is the row size of matrices $A_1,...,A_n$.

The computational effort per iteration is dominated by the necessity, given a positive definite matrix X of the same size and block-diagonal structure as those of A_i and a vector r,

ullet to compute $n \times n$ symmetric matrix \widehat{H} with entries

$$\widehat{H}_{ij} = \text{Tr}(X^{-1}A_iX^{-1}A_j), \ i, j = 1, ..., n;$$

• to solve $n \times n$ linear system

$$Hu = r$$

with positive definite matrix H "readily given" by \widehat{H} .