# A Conic Smörgåsbord

Bruno F. Lourenço
ISM

August 10th, 2023

# Today's contents

1. Cones, convexity and optimization
2. Duality and facial reduction
3. Bonus content

**Software**: CVXPY (there are also versions for Julia, R and others):
`https://www.cvxpy.org/`

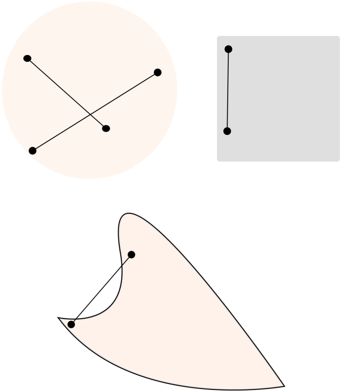
# Part 1 - Cones, convexity and optimization

# Convex sets

**Definition (Convex set)**

Let $C \subseteq \mathbb{R}^n$. $C$ is convex iff

$$x, y \in C \Rightarrow \alpha x + (1-\alpha) y \in C, \forall \alpha \in [0,1]$$



# Basic types of convex sets - affine sets

Affine set $\stackrel{\text{def}}{\Longleftrightarrow}$ the solution set of finitely many **equations**

- $\mathcal{C} \subseteq \mathbb{R}^n$ is affine $\Leftrightarrow$ exists $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ such that $\mathcal{C} = \{x \in \mathbb{R}^n \mid Ax = b\}$

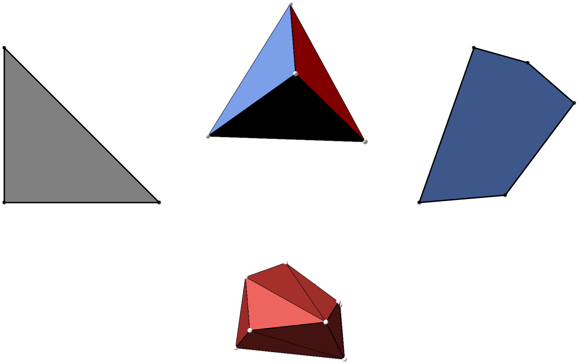
Examples

- A hyperplane $\{x \in \mathbb{R}^n \mid \langle x, v \rangle = \alpha\}$
- A vector subspace in $\mathbb{R}^n$
- Affine space = "translated subspace".

# Basic types of convex sets - polyhedral sets

Polyhedral sets $\overset{\text{def}}{\iff}$ the solution set of finitely many **equalities** and **inequalities**

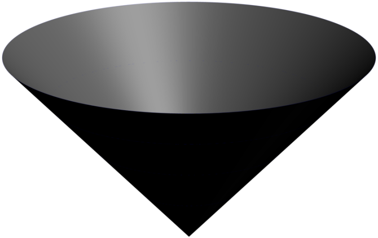
- $\mathcal{C} \subseteq \mathbb{R}^n$ is polyhedral $\Leftrightarrow$ exists $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ such that $\mathcal{C} = \{x \in \mathbb{R}^n \mid Ax \leq b\}$



# Basic types of convex sets - convex cones

$\mathcal{K}$ is a convex cone $\stackrel{\text{def}}{\iff}$ $\alpha x + \beta y \in \mathcal{K}$, whenever $x, y \in \mathcal{K}$ and $\alpha, \beta \geq 0$.

- $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall i\}$
- $n \times n$ symmetric positive semidefinite matrices $\mathcal{S}^n_+$.



# Convex functions

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

- $f$ is convex $\stackrel{\text{def}}{\iff}$ $f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$, $\forall x, y \in \mathbb{R}^n$, $\forall \alpha \in [0, 1]$.
- $f$ is convex $\Leftrightarrow$ the epigraph of $f$ given by $\operatorname{epi} f := \{(x, \mu) \mid f(x) \le \mu\}$ is a convex set.

**Examples**:

- $f(x) = x^2$
- $f(x) = ax$
- $f(x) = -\ln(x)$.

**Non-examples**:

- $f(x) = \ln(x)$
- $f(x) = x^3$

# Conic linear programming

$$
\begin{aligned}
\min_{x} \quad & \langle c, x\rangle \\
\text{subject to} \quad & \mathcal{A}x = b \\
& x \in \mathcal{K}
\end{aligned}
$$

- $\mathcal{K} \subseteq \mathcal{E}$: closed convex cone,
- $\mathcal{A} : \mathcal{E} \to \mathbb{R}^m$: linear map, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$,
- $\mathcal{E}$ is an Euclidean space equipped with an inner product $\langle \cdot, \cdot \rangle$ and a norm $\|\cdot\|$ induced by $\langle \cdot, \cdot \rangle$.

Feasible region $\{x \in \mathcal{K} \mid Ax = b\}$ = "a cone intersected by an affine set".

# Conic Linear Programming - Alternative forms

- "Minimize/Maximize a linear function, subject to equalities, inequalities and cone constraints"
- These are all CLPs:

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \\ & Ex - d \in \mathcal{K} \end{aligned}$$

$$\begin{aligned} \min_{x,y} \quad & c_1^T x + c_2^T y \\ \text{subject to} \quad & A_1 y \leq b_1, \\ & A_2 x = b_2 \\ & (x_1, x_2) \in \mathcal{K}_1 \times \mathcal{K}_2 \end{aligned}$$

# Linear Programming (LP)

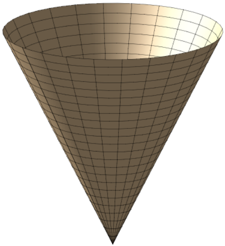
$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathbb{R}^n_+ \end{aligned}$$

- $\mathcal{A}$ is a $m \times n$ matrix, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

# The second-order cone (a.k.a ice-cream cone)

$$\mathcal{Q}^{n+1} := \{(x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^n \mid x_0 \geq \|\bar{x}\|_2\},$$

where $\|\bar{x}\|_2 = \sqrt{\bar{x}_1^2 + \cdots + \bar{x}_n^2}$



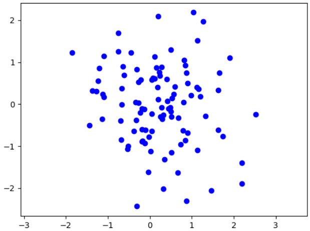
# Second-order cone programming (SOCP)

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{Q}^{n_1} \times \cdots \times \mathcal{Q}^{n_r} \end{aligned} \tag{P}$$

- $\mathcal{A}$ is a $m \times n$ matrix, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

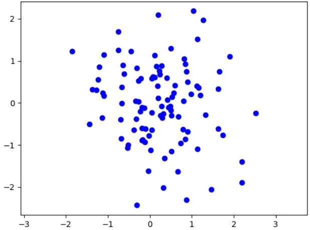
# Antenna placing problem

We want place an antenna that sends a signal that covers the whole region below.




Where should the antenna be placed and what is the minimum radius of the signal capable of covering all the points?

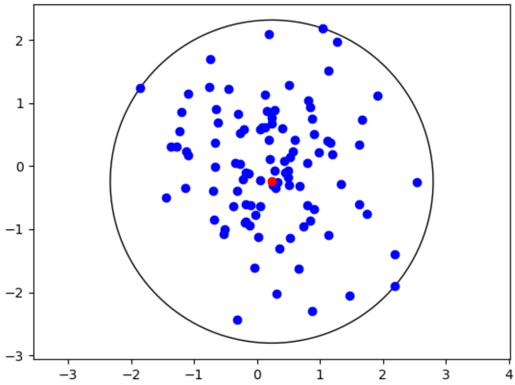
# Antenna placing problem- Formulation




$$\min_{x \in \mathbb{R}^3} \quad x_0 \qquad \text{(P)}$$
$$\text{subject to} \quad \|\bar{x} - p_i\|_2 \leq x_0, \forall i = 1, \ldots, m$$

$$\min_{x \in \mathbb{R}^3} \quad x_0 \qquad \text{(P)}$$
$$\text{subject to} \quad (x_0, \bar{x} - p_i) \in \mathcal{Q}^3, \qquad \forall i = 1, \ldots, m$$

# Antenna placing problem - solution



# Semidefinite Programming (SDP)

$$\begin{aligned} \min_{X \in \mathcal{S}^n} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X \succeq 0 \end{aligned} \tag{P}$$

- $\mathcal{S}^n$: $n \times n$ symmetric matrices.
- $X \succeq 0 \iff X \in \mathcal{S}^n_+ \stackrel{\text{def}}{\iff} v^T X v \geq 0, \forall v \in \mathbb{R}^n$.
- $\langle X, Y \rangle := \mathrm{trace}(X^\top Y) = \sum_{i,j} X_{ij} Y_{ij}$
- $\|X\|_F := \sqrt{\mathrm{trace}(X^\top X)} = \sqrt{\sum_{i,j} X_{ij}^2}$

"Linear programming for the 21st century"

# Linear Algebra Review

Let $X \in \mathcal{S}^n$ and $v \in \mathbb{R}^n$.

- $X \succeq 0 \iff$ all the eigenvalues of $X$ are nonnegative
- $X \succeq 0 \iff$ there exists a $n \times n$ symmetric matrix $V$ such that $X = V^2$.
- $\langle X, vv^T \rangle = v^T X v$
- If $X \succeq 0$, then $\langle X, vv^T \rangle = 0 \iff Xv = 0$.

# SDP Example: Nearest correlation matrix problem

- Suppose we are given a $H \in \mathcal{S}^n$ with diagonal entries equal to 1.
- **Problem:** We want to find the correlation matrix that is the nearest possible to $H$.

$$\begin{aligned} \min_{X \in \mathcal{S}^n} \quad & \|X - H\|_F \\ \text{subject to} \quad & X_{ii} = 1, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{aligned} \tag{Cor}$$

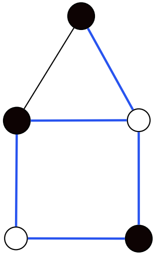
$\|\cdot\|_F$ is the Frobenius norm: $\|A\|_F = \sqrt{\mathrm{trace}(AA^\top)}$.

# SDP Example: Nearest correlation matrix problem (continued)

$$\begin{aligned} \min_{X\in\mathcal{S}^n, t\in\mathbb{R}} \quad & t \\ \text{subject to} \quad & \|X - H\|_F \leq t \\ & X_{ii} = 1, \quad i = 1, \ldots, n \\ & X \succeq 0 \end{aligned} \tag{Cor}$$

The constraint "$\|X - H\| \leq t$" can be written as a second order cone.

# MAX-CUT



**Goal**: Separate the vertices in two sets $S$, $S'$, such that the weight of the crossing edges is maximized. (**NP-Hard**)

- $a_{ij}$: weight of the edge between the $i$-th and $j$-th vertices.
- $x_i$: 1 if the $i$-th vertex is in $S$, $-1$ if in $S'$.

$$\begin{array}{lll} \max\limits_{x \in \mathbb{R}^n} & \displaystyle\sum_{i,j=1}^{n} \frac{a_{ij}}{4}(1 - x_i x_j) & \\ \text{subject to} & x_i^2 = 1, & i = 1, \ldots, n \end{array}$$

# The SDP relaxation - GW'95

- $X \in \mathcal{S}_+^n$ and $\operatorname{rank}(X) = 1 \Leftrightarrow X = xx^T$, for some $x \in \mathbb{R}^n$.
  - $X_{ij} = x_i x_j$ holds.

$$\begin{array}{lll} \max\limits_{x \in \mathcal{S}^n} & \sum\limits_{i,j=1}^{n} \frac{a_{ij}}{4}(1 - X_{ij}) & \\ \text{subject to} & X_{ii} = 1, & i = 1, \dots, n \\ & X \in \mathcal{S}_+^n, & \operatorname{rank}(X) = 1 \end{array}$$

**SDP relaxation**:

$$\begin{array}{lll} \max\limits_{x \in \mathcal{S}^n} & \sum\limits_{i,j=1}^{n} \frac{a_{ij}}{4}(1 - X_{ij}) & \\ \text{subject to} & X_{ii} = 1, & i = 1, \dots, n \\ & X \in \mathcal{S}_+^n, & \cancel{\operatorname{rank}(X) = 1} \end{array}$$

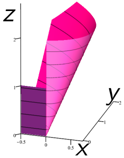
- Approximation ratio: $\frac{\text{MCUT}}{\text{SDP}} > 87\%$.
- Similar idea applies to many combinational optimization problems.

# Interlude - Some history

- **Symmetric Cone Programming**: LP + SOCP + SDP + $\alpha$.
  - SOCP and SDPs : researched intensively from the 90s on, partly because of the advent of interior point methods.
- **Non-symmetric cone optimization**: exponential cones, power cones, *p*-cones and many others.
  - More recent topic, with several new solvers developed in the past few years.

# The exponential cone

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




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

V. Chandrasekaran, P. Shah
Relative entropy optimization and its applications.
*Math. Program. 161, 1–32 (2017)*

# A geometric programming example

1. B-san wants to give a box-like present to a friend.
2. However, B-san wants to wrap it using a *special wrapping paper* and B-san only has $1m^2$ of it.
3. Because B-san is pretentious, B-san wants the ratio between height of the box and its width to be in $[1.5, \phi]$, where $\phi$ is the golden ratio $\phi = \frac{1+\sqrt{5}}{2}$
4. **Problem**: What is the biggest box (in volume) that can be wrapped with the special paper?

# A geometric programming example

$$\begin{aligned} \max_{w,h,d} \quad & whd \\ \text{subject to} \quad & 2(wh + wd + hd) \leq 1 \\ & 1.5 \leq \frac{h}{w} \leq \phi \\ & w > 0, h > 0, d > 0 \end{aligned}$$

**Not a convex problem** but if we make the substitutions $w = e^{\hat{w}}$, $d = e^{\hat{d}}$ and $h = e^{\hat{h}}$ we get

$$\begin{aligned} \max_{\hat{w},\hat{h},\hat{d}} \quad & e^{\hat{w}+\hat{h}+\hat{d}} \\ \text{subject to} \quad & (e^{\hat{w}+\hat{h}} + e^{\hat{w}+\hat{d}} + e^{\hat{h}+\hat{d}}) \leq 0.5 \\ & 1.5 \leq e^{\hat{h}-\hat{w}} \leq \phi \end{aligned}$$

# A geometric programming example

**Taking logs** linearizes the objective function and some of the constraints.

$$\begin{aligned} \max_{\hat{w},\hat{h},\hat{d}} \quad & \hat{w}+\hat{h}+\hat{d} \\ \text{subject to} \quad & e^{\hat{w}+\hat{h}} + e^{\hat{w}+\hat{d}} + e^{\hat{h}+\hat{d}} \le 0.5 \\ & \log(1.5) \le \hat{h}-\hat{w} \le \log(\phi) \end{aligned}$$

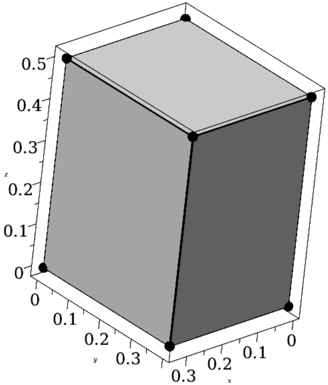
Noting that $e^x \le t$ holds if and only if $(x, 1, t) \in K_{\text{exp}}$, we have

$$\begin{aligned} \max_{\hat{w},\hat{h},\hat{d},t} \quad & \hat{w}+\hat{h}+\hat{d} \\ \text{subject to} \quad & t_1+t_2+t_3 \le 0.5 \\ & (\hat{w}+\hat{h}, 1, t_1) \in K_{\text{exp}}, (\hat{w}+\hat{d}, 1, t_2) \in K_{\text{exp}}, (\hat{h}+\hat{d}, 1, t_3) \in K_{\text{exp}} \\ & \log(1.5) \le \hat{h}-\hat{w} \le \log(\phi) \end{aligned}$$

Reminder:

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

# Solution



# Discrete distribution estimation

- We want to estimate a discrete distribution $p$ based on some prior information.
  - We might know some bounds on the moments
  - We might have some information on the $p_i$'s themselves.
- **Maximum entropy principle**: we try to find the "most random" $p$ that is consistent with the prior information $\mathcal{P}$.

$$
\begin{aligned}
\max_{p \in \mathbb{R}^n} \quad & \sum_{i=1}^{n} -p_i \ln p_i \\
\text{subject to} \quad & p \in \mathcal{P} \\
& \sum_{i=1}^{n} p_i = 1 \\
& p \in \mathbb{R}^n_+
\end{aligned}
$$

# Exponential cone formulation

Reminder:

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

$$\begin{aligned}
\max_{p,t\in\mathbb{R}^n} \quad & \sum_{i=1}^{n} t_i \\
\text{subject to} \quad & t_i \leq -p_i \ln p_i, \quad i = 1,\dots,n \\
& p \in \mathcal{P}, p \in \mathbb{R}^n_+ \\
& \sum_{i=1}^{n} p_i = 1
\end{aligned}$$

$$\begin{aligned}
\max_{p,t\in\mathbb{R}^n} \quad & \sum_{i=1}^{n} t_i \\
\text{subject to} \quad & (t_i, p_i, 1) \in K_{\exp}, \quad i = 1,\dots,n \\
& p \in \mathcal{P}, p \in \mathbb{R}^n_+ \\
& \sum_{i=1}^{n} p_i = 1
\end{aligned}$$

# Expressive power

# Convex optimization

**Convex optimization**:

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x \in C, \end{aligned}$$

$C$ is a convex set and $f$ is a convex function.
**Conic linear programming** (CLP)

$$\begin{aligned} \min_{x} \quad & \langle c, x \rangle \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{K} \end{aligned}$$

If we let $C := \{x \in \mathcal{K} \mid \mathcal{A}x = b\}$, then $C$ is convex.

- CLP is a particular case of convex optimization. However

# CLP $\cong$ Convex Optimization

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & x \in C \end{aligned} \qquad \begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & x \in C \\ & f(x) \leq t \end{aligned}$$

Let $C_2 := \{(x, t) \mid x \in C, f(x) \leq t\}$ and let $\mathcal{K}$ be the convex cone in $\mathcal{E} \times \mathbb{R}^2$ generated by $C_2 \times \{1\}$. That is

$$\mathcal{K} := \{\alpha(x, t, 1) \mid \alpha \geq 0, (x, t) \in C_2\}.$$

$$\begin{aligned} \min_{x,t,\alpha} \quad & t \\ \text{subject to} \quad & \alpha = 1 \\ & (x, t, \alpha) \in \mathcal{K} \end{aligned}$$

- Every convex optimization problem has an equivalent CLP formulation!
- CLP philosophy: concentrate the hard part of the problem inside the cone.
- CVXPY works by converting a convex problem into an equivalent CLP and calling a CLP solver.

# More on expressive power

- Some researchers believe a few cones are enough to model the vast majority of convex applications.

  > *The following chapters present modeling with four types of convex cones: quadratic cones, power cones, exponential cone, semidefinite cone. It is "well-known" in the convex optimization community that this family of cones is sufficient to express almost all convex optimization problems appearing in practice.* [MOSEK Modelling cookbook, 2023]

- That said, a cone may be "too general" for a certain application ⇒ a more specific cone may be better.
- Some new solvers (alfonso, DDS, Hypatia, etc) support multiple cones
  - User can select the cone that best fit the application.

# More specific vs more general cones

$$\mathcal{Q}^{n+1} := \{(x_0, \bar{x}) \in \mathbb{R}^n \times \mid x_0 \geq \|\bar{x}\|_2\},$$

where $\|\bar{x}\|_2 = \sqrt{\bar{x}_1^2 + \cdots + \bar{x}_n^2}$

$$\mathcal{S}_+^n := \{X \in \mathcal{S}^n \mid v^T X v \geq 0, \forall v \in \mathbb{R}^n\}$$

$$(\bar{x}, x_0) \in \mathcal{Q}^{n+1} \Leftrightarrow \begin{pmatrix} x_0 & \bar{x}_1 & \cdots & \bar{x}_n \\ \bar{x}_1 & x_0 & 0 & \cdots \\ \vdots & & \ddots & \cdots \\ \bar{x}_n & 0 & \cdots & x_0 \end{pmatrix} \in \mathcal{S}_+^{n+1}$$

- Everything that can be expressed using $\mathcal{Q}^{n+1}$ can also be expressed using $\mathcal{S}_+^{n+1}$
- However, $\mathcal{S}_+^{n+1}$ requires $(n+1) \times (n+1)$ matrices, while $\mathcal{Q}^{n+1}$ is a cone in $\mathbb{R}^{n+1}$.

# https://docs.mosek.com/cheatsheets/conic.pdf

mosek

## Conic Modeling Cheatsheet

### Cones

Quadratic cone $\mathcal{Q}^n$

$$x_1 \geq \sqrt{x_2^2 + \cdots + x_n^2}$$

Rotated quadratic cone $\mathcal{Q}_r^n$

$$2x_1x_2 \geq x_3^2 + \cdots + x_n^2,\ x_1, x_2 \geq 0$$

Power cone $\mathcal{P}_3^{\alpha,1-\alpha}$, $\alpha \in (0,1)$

$$x_1^\alpha x_2^{1-\alpha} \geq |x_3|,\ x_1, x_2 \geq 0$$

Exponential cone $K_{\exp}$

$$x_1 \geq x_2 e^{x_3/x_2},\ x_2 \geq 0$$

### Simple bounds

| | |
|---|---|
| $t \geq x^2$ | $(0.5, t, x) \in \mathcal{Q}_r^3$ |
| $\lvert t\rvert \leq \sqrt{x}$ | $(0.5, x, t) \in \mathcal{Q}_r^3$ |
| $t \geq \lvert x\rvert$ | $(t, x) \in \mathcal{Q}^2$ |
| $t \geq 1/x,\ x > 0$ | $(x, t, \sqrt{2}) \in \mathcal{Q}_r^3$ |
| $t \geq \lvert x\rvert^p,\ p > 1$ | $(t, 1, x) \in \mathcal{P}_3^{1/p,1-1/p}$ |
| $t \geq 1/x^p,\ x > 0,\ p > 0$ | $(t, x, 1) \in \mathcal{P}_3^{1/(1+p),p/(1+p)}$ |
| $\lvert t\rvert \leq x^p,\ x > 0,\ p \in (0,1)$ | $(x, 1, t) \in \mathcal{P}_3^{p,1-p}$ |
| $t \geq \lvert x\rvert^p/y^{p-1},\ y \geq 0$<br>$p > 1$ | $(t, y, x) \in \mathcal{P}_3^{1/p,1-1/p}$ |
| $t \geq x^T x/y,\ y \geq 0$ | $(0.5t, y, x) \in \mathcal{Q}_r^{n+2}$ |
| $t \geq e^x$ | $(t, 1, x) \in K_{\exp}$ |
| $t \leq \log x$ | $(x, 1, t) \in K_{\exp}$ |
| $t \geq 1/\log x,\ x > 1$ | $(u, t, \sqrt{2}) \in \mathcal{Q}_r^3$<br>$(x, 1, u) \in K_{\exp}$ |
| $t \geq a_1^{x_1} \cdots a_n^{x_n},\ a_i > 0$ | $(t, 1, \sum x_i \log a_i) \in K_{\exp}$ |
| $t \geq xe^x,\ x \geq 0$ | $(t, x, u) \in K_{\exp}$<br>$(0.5, u, x) \in \mathcal{Q}_r^3$ |
| $t \geq \log(1 + e^x)$ | $u + v \leq 1$<br>$(u, 1, x - t) \in K_{\exp}$<br>$(v, 1, -t) \in K_{\exp}$ |
| $t \geq \lvert x\rvert^{3/2}$ | $(t, 1, x) \in \mathcal{P}_3^{2/3,1/3}$ |
| $t \geq x^{3/2},\ x \geq 0$ | $(s, t, x), (x, 1/8, s) \in \mathcal{Q}_r^3$ |
| $t \geq 1/x^3,\ x > 0$ | $(t, x, 1) \in \mathcal{P}_3^{3/4,1/4}$ |
| $0 \leq t \leq x^{2/5},\ x \geq 0$ | $(x, 1, t) \in \mathcal{P}_3^{2/5,3/5},\ t \geq 0$ |

### Means and averaging

| | |
|---|---|
| Log-sum-exp<br>$t \geq \log(\sum e^{x_i})$ | $(z_i, 1, x_i - t) \in K_{\exp}$<br>$i = 1, \ldots, n$<br>$\sum z_i \leq 1$ |
| Harmonic mean<br>$0 \leq t \leq n(\sum x_i^{-1})^{-1}$<br>$x_i > 0$ | $(z_i, x_i, t) \in \mathcal{Q}_r^3$<br>$i = 1, \ldots, n$<br>$\sum z_i = nt/2$ |
| Geometric mean<br>$\lvert t\rvert \leq (x_1 \cdots x_n)^{1/n}$<br>$x_i > 0$ | $(z_i, x_i, z_{i+1}) \in \mathcal{P}_3^{1-1/i,1/i}$<br>$i = 2, \ldots, n$<br>$z_2 = x_1,\ z_{n+1} = t$ |
| $\lvert t\rvert \leq \sqrt{xy},\ x, y > 0$ | $(x, y, \sqrt{2}t) \in \mathcal{Q}_r^3$ |
| Weighted geom. mean<br>$\lvert t\rvert \leq x_1^{\alpha_1} \cdots x_n^{\alpha_n},\ x_i > 0$<br>$\alpha_i > 0, \sum \alpha_i = 1$ | $(z_i, x_i, z_{i+1}) \in \mathcal{P}_3^{1-\beta_i,\beta_i}$<br>$\beta_i = \alpha_i/(\alpha_1 + \cdots + \alpha_i)$<br>$i = 2, \ldots, n$<br>$z_2 = x_1,\ z_{n+1} = t$ |
| $\lvert t\rvert \leq x^{1/4}y^{5/12}z^{1/3}$<br>$x, y, z \geq 0$ | $(s, z, t) \in \mathcal{P}_3^{2/3,1/3}$<br>$(x, y, s) \in \mathcal{P}_3^{3/8,5/8}$ |

### Entropy

| | |
|---|---|
| $t \leq -x \log x$ | $(1, x, t) \in K_{\exp}$ |
| $t \geq x \log(x/y)$ | $(y, x, -t) \in K_{\exp}$ |
| $t \geq \log(1 + 1/x)$<br>$x > 0$ | $(x + 1, u, \sqrt{2}) \in \mathcal{Q}_r^3$<br>$(1 - u, 1, -t) \in K_{\exp}$ |
| $t \leq \log(1 - 1/x)$<br>$x > 1$ | $(x, u, \sqrt{2}) \in \mathcal{Q}_r^3$<br>$(1 - u, 1, t) \in K_{\exp}$ |
| $t \geq x \log(1 + x/y)$<br>$x, y > 0$ | $(y, x + y, u) \in K_{\exp}$<br>$(x + y, y, v) \in K_{\exp}$<br>$t + u + v = 0$ |

### Convex quadratic problems

Let $\Sigma \in \mathbb{R}^{n \times n}$, symmetric, p.s.d.
Find $\Sigma = LL^T$, $L \in \mathbb{R}^{n \times k}$ (Cholesky factor).
Then $x^T \Sigma x = \|L^T x\|_2^2$.

| | |
|---|---|
| $t \geq \frac{1}{2} x^T \Sigma x$ | $(1, t, L^T x) \in \mathcal{Q}_r^{k+2}$ |
| $t \geq \sqrt{x^T \Sigma x}$ | $(t, L^T x) \in \mathcal{Q}^{k+1}$ |
| $\frac{1}{2} x^T \Sigma x + p^T x + q \leq 0$ | $(1, -p^T x - q, L^T x) \in \mathcal{Q}_r^{k+2}$ |
| $\max_x c^T x - \frac{1}{2} x^T \Sigma x$ | $\max c^T x - r$<br>$(1, r, L^T x) \in \mathcal{Q}_r^{k+2}$ |
| $c^T x + d \geq \|Ax + b\|_2$ | $(c^T x + d, Ax + b) \in \mathcal{Q}^{m+1}$ |

### Norms, $x \in \mathbb{R}^n$

| | |
|---|---|
| $\|\cdot\|_1,\ t \geq \sum \lvert x_i\rvert$ | $(z_i, x_i) \in \mathcal{Q}^2,\ t = \sum z_i$ |
| $\|\cdot\|_2,\ t \geq (\sum x_i^2)^{1/2}$ | $(t, x) \in \mathcal{Q}^{n+1}$ |
| $\|\cdot\|_p,\ p > 1$<br>$t \geq (\sum \lvert x_i\rvert^p)^{1/p}$ | $(z_i, t, x_i) \in \mathcal{P}_3^{1/p,1-1/p}$<br>$i = 1, \ldots, n$<br>$\sum z_i = t$ |

### Geometry

| | |
|---|---|
| Bounding ball<br>$\min_x \max_i \|x - x_i\|_2$ | $\min r$<br>$(r, x - x_i) \in \mathcal{Q}^{n+1}$ |
| Geometric median<br>$\min_x \sum \|x - x_i\|_2$ | $\min \sum t_i$<br>$(t_i, x - x_i) \in \mathcal{Q}^{n+1}$ |
| Analytic center<br>$\max_x \sum \log(b_i - a_i^T x)$ | $\max \sum t_i$<br>$(b_i - a_i^T x, 1, t_i) \in K_{\exp}$ |

### Regression and fitting

| | |
|---|---|
| Regularized least squares<br>$\min_w \|Xw - y\|_2^2 + \lambda \|w\|_2^2$ | $\min t + \lambda r$<br>$(0.5, t, Xw - y) \in \mathcal{Q}_r^{m+2}$<br>$(0.5, r, w) \in \mathcal{Q}_r^{n+2}$ |
| Max likelihood<br>$\max_p p_1^{a_1} \cdots p_n^{a_n}$ | $\max \sum a_i t_i$<br>$(p_i, 1, t_i) \in K_{\exp}$ |
| Logistic cost function<br>$t \geq -\log(1/(1 + e^{-\theta^T x}))$ | $u + v \leq 1$<br>$(u, 1, -\theta^T x - t) \in K_{\exp}$<br>$(v, 1, -t) \in K_{\exp}$ |

### Risk-return

$\Sigma \in \mathbb{R}^{n \times n}$ – covariance, $\Sigma = LL^T,\ L \in \mathbb{R}^{n \times k}$

| | |
|---|---|
| $\max_x \alpha^T x$<br>s.t. $x^T \Sigma x \leq \gamma$ | $\max_x \alpha^T x$<br>$(\sqrt{\gamma}, L^T x) \in \mathcal{Q}^{k+1}$ |
| $\max_x \alpha^T x - \delta x^T \Sigma x$ | $\max_x \alpha^T x - \delta r$<br>$(0.5, r, L^T x) \in \mathcal{Q}_r^{k+2}$ |
| Risk plus $x^{1.5}$ impact cost<br>$t \geq \delta x^T \Sigma x + \beta \sum \lvert x_i\rvert^{3/2}$ | $t \geq \delta r + \beta \sum u_i$<br>$(0.5, r, L^T x) \in \mathcal{Q}_r^{k+2}$<br>$(u_i, 1, x_i) \in \mathcal{P}_3^{2/3,1/3}$ |
| Risk in factor model<br>$\gamma \geq x^T(D + FSF^T)x$<br>$D$ – specific risk (diag.)<br>$F \in \mathbb{R}^{n \times k}$ – factor loads<br>$S = UU^T$ – factor cov. | $\gamma \geq t + s$<br>$(0.5, t, \sqrt{D}x) \in \mathcal{Q}_r^{n+2}$<br>$(0.5, s, U^T F^T x) \in \mathcal{Q}_r^{k+2}$ |

More: https://docs.mosek.com/modeling-cookbook/index.html

# Part 2 - Duality and facial reduction

# More convex analysis

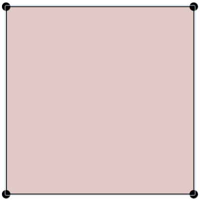
# Topological Interior

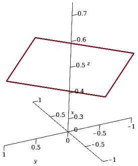
- $\mathcal{E}$: Euclidean space (i.e., $\mathbb{R}^n$) with inner product $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$
- $B(x,r)$ is the open ball centered in $x$ with radius $r$, i.e., $B(x,r) = \{y \mid \|y-x\| < r\}$.

Let $C \subseteq \mathcal{E}$

**Interior**

$\operatorname{int} C := \{x \in C \mid \exists r > 0, s.t., B(x,r) \subseteq C\}$.





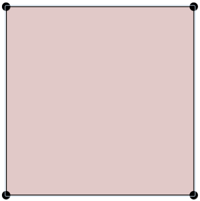
# Relative interior

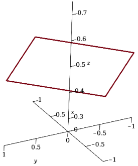
**Definition (Relative interior)**

$x$ is a relative interior point of $C$ (i.e., $x \in \operatorname{ri} C$) if for every $y \in C$, the line segment connecting $x$ and $y$ can be extended past $x$ while staying inside $C$.

$$x \in \operatorname{ri} C \overset{\text{def}}{\iff} \forall y \in C, \exists \mu > 1, \text{ s.t. } \mu x + (1-\mu) y \in C$$

- $\operatorname{ri} C = C \overset{\text{def}}{\iff} C$ is **relatively open**.





# Closure

$\mathcal{E}$: finite dimensional Euclidean space
$C \subseteq \mathcal{E}$: convex set

**Definition (Closure)**

The closure $\operatorname{cl} C$ of $C$ is the set of limit points of $C$ $\Leftrightarrow$ smallest closed set containing $C$.

- $\operatorname{cl} C = C \overset{\text{def}}{\iff} C$ is **closed**

# Properties of closures and relative interiors

$\mathcal{E}$: finite dimensional Euclidean space
$C \subseteq \mathcal{E}$: convex set

- $\mathrm{ri}\, C$ and $\mathrm{cl}\, C$ are convex.
- $\mathrm{ri}\, C \neq \emptyset$ if $C \neq \emptyset$.
- $\mathrm{ri}\,(\mathrm{cl}\, C) = \mathrm{ri}\, C$
- $\mathrm{ri}\,\mathrm{ri}\,(C) = \mathrm{ri}\, C$ "relative interiors are relatively open"
- $\mathrm{cl}\,(\mathrm{cl}\, C) = \mathrm{cl}\, C$ "closures are closed"

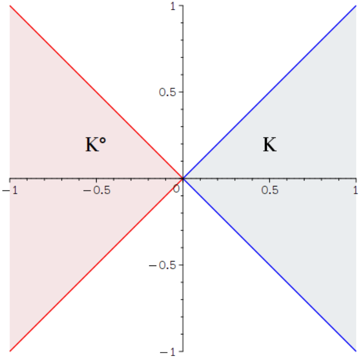
**Examples**

- $\mathrm{ri}\, \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i > 0, \forall i\}$
- $\mathrm{ri}\, \mathcal{S}^n_+ =$ symmetric positive definite matrices.

# Polars and duals of cones

$\mathcal{K} \subseteq \mathcal{E}$: convex cone.

$$\mathcal{K}^\circ = \{y \in \mathcal{E} \mid \langle x, y \rangle \leq 0, \forall x \in \mathcal{K}\}.$$



# Polars of cones - Examples and Properties

$\mathcal{K} \subseteq \mathcal{E}$: convex cone.

- **Bipolar Theorem**: $\mathcal{K}^{\circ\circ} = \mathrm{cl}\,(\mathcal{K})$.
- $(\mathbb{R}^n_+)^\circ = -\mathbb{R}^n_+$
- $(\mathcal{S}^n_+)^\circ = -\mathcal{S}^n_+$.
- $(\mathcal{Q}^n_p)^\circ = -\mathcal{Q}^n_q$, where $1/p + 1/q = 1$, $p \in (1, \infty)$,
  $\mathcal{Q}^n_p := \{(x_0, \bar{x}) \mid \|\bar{x}\|_p \leq x_0\}$.

**Dual cone**

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

- **Bipolar Theorem**: $\mathcal{K}^{**} = \mathrm{cl}\,(\mathcal{K})$.
- $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$.
- $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$.
- $(\mathcal{Q}^n_p)^* = \mathcal{Q}^n_q$, where $1/p + 1/q = 1$, $p \in (1, \infty)$,
  $\mathcal{Q}^n_p := \{(x_0, \bar{x}) \mid \|\bar{x}\|_p \leq x_0\}$.

# Recall our basic conic linear program

$$\begin{aligned} \cancel{\min_x}\ \inf \quad & \langle c, x \rangle \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{K} \end{aligned} \tag{P}$$

Suppose we wish to **relax the linear constraints**:

$$\begin{aligned} \mathcal{L}(y) &:= \inf_{x \in \mathcal{K}} \left[ \langle c, x \rangle + \langle y, b - \mathcal{A}x \rangle) \right] \\ &= \inf_{x \in \mathcal{K}} \left[ \langle c - \mathcal{A}^* y, x \rangle + \langle b, y \rangle \right] \\ &= \langle b, y \rangle + \inf_{x \in \mathcal{K}} \langle c - \mathcal{A}^* y, x \rangle \\ &= \begin{cases} \langle b, y \rangle & \text{if } c - \mathcal{A}^* y \in \mathcal{K}^* \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Denote the optimal value of (P) by $\theta_P$. Then:

$$\theta_P \geq \mathcal{L}(y), \qquad \forall y$$

# Relaxing the CLP

$$\begin{aligned} \inf_x \quad & \langle c, x \rangle \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{K} \end{aligned} \tag{P}$$

We have

$$\mathcal{L}(y) = \begin{cases} \langle b, y \rangle & \text{if } c - \mathcal{A}^* y \in \mathcal{K}^* \\ -\infty & \text{otherwise} \end{cases}$$

and

$$\theta_P \geq \mathcal{L}(y), \qquad \forall y$$

Which leads to

$$\theta_P \geq \sup_y \mathcal{L}(y)$$

- The dual problem is the task of finding the $y$ that provides the **tightest (Lagrangian) relaxation** to (P)

# Primal dual conic linear program (CLP)

$$\begin{aligned} \inf_x \quad & \langle c, x \rangle \qquad \text{(P)} \\ \text{subject to} \quad & \mathcal{A}x = b \\ & x \in \mathcal{K} \end{aligned}$$

$$\begin{aligned} \sup_y \quad & \langle b, y \rangle \qquad \text{(D)} \\ \text{subject to} \quad & c - \mathcal{A}^* y \in \mathcal{K}^*. \end{aligned}$$

- $\mathcal{K}^* := \{s \in \mathcal{E} \mid \langle s, x \rangle \geq 0, \forall x \in \mathcal{K}\}$. (dual cone)
- We denote the primal and dual optimal values by $\theta_P$ and $\theta_D$.

**Proposition (Weak duality)**

$$\theta_P \geq \theta_D$$

# Example - Eigenvalues via SDP duality

Suppose that $C \in \mathcal{S}^n$ is a fixed matrix and consider the SDP:

$$\begin{aligned} \sup_{y \in \mathbb{R}} \quad & y \\ \text{s.t.} \quad & C - yI_n \succeq 0, \end{aligned} \tag{D}$$

where $I_n$ is the $n \times n$ identity matrix. Then $\theta_D = \lambda_{\min}(C)$, where $\lambda_{\min}(C)$ is the minimum eigenvalue of $C$. The primal is:

$$\begin{aligned} \inf_{X \in \mathcal{S}^n} \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle I_n, X \rangle = \mathrm{trace}(X) = 1 \\ & X \succeq 0 \end{aligned} \tag{P}$$

If $v \in \mathbb{R}^n$ is an eigenvector of $C$ associated to $\lambda_{\min}(C)$ with $\|v\| = 1$, then $X^* := vv^\top$ is optimal to (P).

$$\theta_P = \theta_D = \lambda_{\min}(C).$$

# Strong duality in CLP

**Theorem (Strong duality Theorem - Primal version)**

*Suppose that*

- (P) *has a relative interior feasible solution, i.e., there exists $x$ such that $\mathcal{A}x = b$ and $x \in \operatorname{ri} \mathcal{K}$ (Primal Slater Condition)*

*Then:*

- $\theta_P = \theta_D$.
- (D) *has optimal solutions if $\theta_P$ is finite.*

**Theorem (Strong duality Theorem - Dual version)**

*Suppose that*

- (D) *has a relative interior feasible solution, i.e., there exists $y$ such that $c - \mathcal{A}^* y \in \operatorname{ri} \mathcal{K}^*$. (Dual Slater Condition)*

*Then:*

- $\theta_P = \theta_D$.
- (P) *has optimal solutions if $\theta_D$ is finite.*

# Optimality conditions

$$\begin{array}{rll} \inf_x & \langle c, x\rangle & \text{(P)} \\ \text{subject to} & \mathcal{A}x = b & \\ & x \in \mathcal{K} & \end{array} \qquad \begin{array}{rll} \sup_y & \langle b, y\rangle & \text{(D)} \\ \text{subject to} & c - \mathcal{A}^* y \in \mathcal{K}^*. & \end{array}$$

A **sufficient condition** for $(x^*, y^*)$ to be optimal is that the following are satisfied:

- Primal feasibility: $\mathcal{A}x^* = b$, $x^* \in \mathcal{K}$
- Dual feasibility: $s^* \in \mathcal{K}^*$, where $s^* := c - \mathcal{A}^* y^*$
- **Complementary slackness** (i.e., zero duality gap[1]): $\langle s^*, x^*\rangle = 0$.

If the primal and dual Slater conditions hold, the conditions above are **necessary** too.

[1] Note that $\langle s^*, x^*\rangle = \langle c, x^*\rangle - \langle b, y^*\rangle$

# Ex1 - MAXCUT-SDP

$$\begin{array}{lll} \inf\limits_{x \in \mathcal{S}^n} & \langle A, X \rangle & \text{(P)} \\ \text{s.t} & X_{ii} = 1, \quad i = 1, \ldots, n & \\ & X \in \mathcal{S}^n_+ & \end{array} \qquad \begin{array}{lll} \sup\limits_{y \in \mathbb{R}^n} & y_1 + \cdots + y_n & \text{(D)} \\ \text{s.t.} & A - \sum_{i=1}^n E_i y_i \in \mathcal{S}^n_+, & \end{array}$$

where $E_i$ is the matrix that has 1 in the $(i, i)$-entry and zero elsewhere.

- Both primal and dual Slater conditions are satisfied $\Rightarrow \quad \theta_P = \theta_D$ and both problems are attained.

# Ex2 - Dual Slater Condition holds, but no dual optimal solution

$$\begin{array}{rl} \sup_{t,s} & -s \\ \text{s.t.} & \begin{pmatrix} t & 1 \\ 1 & s \end{pmatrix} \succeq 0 \end{array} \qquad \text{(D)}$$

$$\begin{array}{rl} \inf_{X \in \mathcal{S}^2} & 2X_{12} \\ \text{s.t.} & -X_{11} = 0 \\ & -X_{22} = -1 \\ & X \succeq 0. \end{array} \qquad \text{(P)}$$

- The dual satisfies Slater condition, $\theta_D$ is finite but no dual optimal solutions exists. $\theta_D$ **is unattained**.
- The primal does not satisfy Slater conditions, but has an optimal solution.
- $\theta_P = \theta_D$ holds.

# Some clarification

Keep in mind the following:

$\inf(0,1) = 0$, but $0 \notin (0,1)$. "The infimum is finite but an optimal solution does not exist".

**Primal side**

- $\theta_P$ is **finite** $\Leftrightarrow$ $\theta_P$ is a **real number**.
- $\theta_P$ is **attained** $\Leftrightarrow$ there is a feasible $x^*$ such that $\theta_P = \langle c, x^* \rangle$.
- $\theta_P = -\infty$ ((P) is **unbounded**) $\Leftrightarrow$ there is a sequence $\{x^k\}$ of feasible solutions such that $\lim_{k\to\infty} \langle c, x^k \rangle \to -\infty$
- By convention $\theta_P = +\infty$ iff (P) is infeasible

**Dual side**

- $\theta_D$ is **finite** $\Leftrightarrow$ $\theta_D$ is a **real number**.
- $\theta_D$ is **attained** $\Leftrightarrow$ there is a feasible $y^*$ such that $\theta_D = \langle b, y^* \rangle$.
- $\theta_D = -\infty$ ((D) is **unbounded**) $\Leftrightarrow$ there is a sequence $\{y^k\}$ of feasible solutions such that $\lim_{k\to\infty} \langle b, y^k \rangle \to -\infty$
- By convention $\theta_D = +\infty$ iff (D) is infeasible

# Ex3 - Positive gap SDP

$$\sup_{t,s} \quad -s \qquad \text{(D)}$$
$$\text{s.t.} \quad \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \succeq 0$$

$$\inf_{X \in \mathcal{S}^3} \quad 2X_{12} - 2X_{13} \qquad \text{(P)}$$
$$\text{s.t.} \quad X_{11} = 0$$
$$-X_{22} - 2X_{13} = -1$$
$$X \succeq 0.$$

# Positive gap SDP

$$\sup_{t,s} \; -s \to -1 \qquad \text{(D)}$$
$$\text{s.t.} \quad \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \succeq 0$$

$$\inf_{X \in \mathcal{S}^3} \; 2X_{12} \to 0 \; - 2X_{13} \to 0 \qquad \text{(P)}$$
$$\text{s.t.} \quad X_{11} = 0$$
$$-X_{22} - 2X_{13} = -1$$
$$X \succeq 0.$$

$\theta_D = -1$ and $\theta_P = 0$. **Neither the primal nor the dual satisfy Slater**

- Ok, so what? How bad can this be?

To correct this we substitute $\mathcal{S}^3_+$ for

$$\mathcal{S}^2_+ \oplus 0 = \left\{ \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathcal{S}^2_+ \right\}.$$

# Example

$$\sup_{t,s} \quad \cancel{-s} \to -1 \qquad \text{(D')}$$

$$\text{s.t.} \quad \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \in \mathcal{S}^2_+ \oplus 0$$

Still, $\theta_{D'} = -1$. Let's take a look at the primal problem...

$$(\mathcal{S}^2_+ \oplus 0)^* = \left\{ \begin{pmatrix} a & b & * \\ b & c & * \\ * & * & * \end{pmatrix} \middle| \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathcal{S}^2_+ \right\}.$$

$$\inf_{x} \quad \cancel{2x_{12}} \to 0 \; \cancel{-2x_{13}} \to -1 \qquad \text{(P')}$$

$$\text{s.t.} \quad x_{11} = 0$$

$$-x_{22} - 2x_{13} = -1$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \in \mathcal{S}^2_+.$$

# What happened? 

- $\mathcal{S}^3_+ \Rightarrow \mathcal{S}^2_+ \oplus 0$: The feasible region of (D) stays the same $\Rightarrow$ $\theta_D = \theta_{D'} = -1$.
- $\mathcal{S}^3_+ \Rightarrow (\mathcal{S}^2_+ \oplus 0)^*$: The feasible region of (P) expands $\Rightarrow$ $-1 = \theta_{P'} \leq \theta_P = 0$.

$\mathcal{S}^2_+ \oplus 0$ a *face* of $\mathcal{S}^3_+$ with two key properties:

- it contains the feasible region of (D)
- Slater's condition is satisfied at (D').

This is an example of *Facial Reduction*

# Even more convex analysis

# Separating hyperplanes

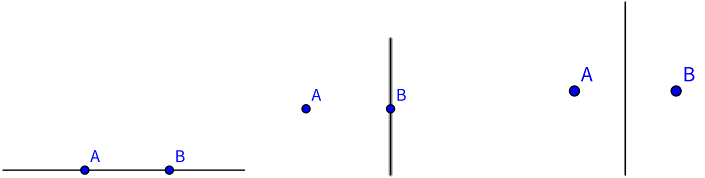
$H = \{x \in \mathcal{E} \mid \langle x, y \rangle = \theta\}$: hyperplane ($x \neq 0$)
$C_1, C_2$: convex sets
Define the closed half-spaces

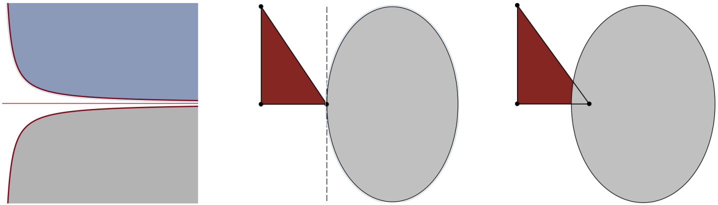
$$H^+ := \{x \in \mathcal{E} \mid \langle x, y \rangle \geq \theta\}, \qquad H^- := \{x \in \mathcal{E} \mid \langle x, y \rangle \leq \theta\}$$

$B_\epsilon$: unit ball of radius $\epsilon$

- $C_1$ and $C_2$ are **separated** by $H \overset{\text{def}}{\Longleftrightarrow}$ $C_1$ and $C_2$ belong to different closed half-spaces defined by $H$.
- $C_1$ and $C_2$ are **properly separated** by $H \overset{\text{def}}{\Longleftrightarrow}$ $C_1$ and $C_2$ belong to different closed half-spaces and **at least one of them is not contained in** $H$.
- $C_1$ and $C_2$ are **strongly separated** by $H \overset{\text{def}}{\Longleftrightarrow}$ $\exists \epsilon > 0$ such that $C_1 + B_\epsilon$ and $C_2 + B_\epsilon$ belong to different **open** half-spaces defined by $H$.

# Separating hyperplanes - Examples



# Some results

$C_1, C_2 \subseteq \mathcal{E}$: nonempty closed convex sets.

1. $C_1$ and $C_2$ can be **strongly separated** $\Leftrightarrow$ $\mathrm{dist}\,(C_1, C_2) = \inf_{x,y} \|x - y\| > 0 \Leftrightarrow 0 \notin \mathrm{cl}\,(C_1 - C_2)$
2. $C_1$ and $C_2$ can be **properly separated** $\Leftrightarrow (\mathrm{ri}\, C_1) \cap (\mathrm{ri}\, C_2) = \emptyset$.

# Faces of convex sets

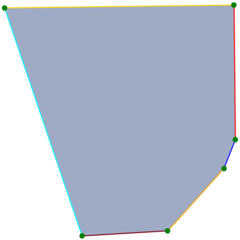
**Definition (Face)**

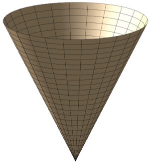
Let $C, \mathcal{F}$ be convex sets such that $\mathcal{F} \subseteq C$. $\mathcal{F}$ is a face of $C$ $\overset{\text{def}}{\Longleftrightarrow}$ for every $\alpha \in (0,1)$ and every $x, y \in C$

$$\alpha x + (1-\alpha) y \in \mathcal{F} \Rightarrow x, y \in \mathcal{F}$$

We write $\mathcal{F} \trianglelefteq C$.

- A face that is a singleton $\{x\}$ is called an **extreme point**
- A face $\mathcal{F}$ of dimension 1 of a cone $\mathcal{K}$ is called an **extreme ray**.



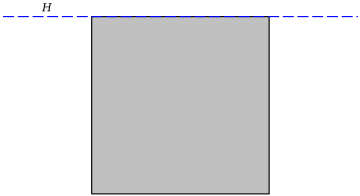


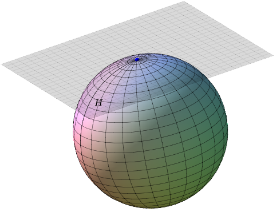
# Supporting hyperplanes

$H = \{x \in \mathcal{E} \mid \langle c, x \rangle = \theta\}$: hyperplane
$C \subseteq \mathcal{E}$: convex set
$H^+ := \{x \in \mathcal{E} \mid \langle c, x \rangle \geq \theta\}, \qquad H^- := \{x \in \mathcal{E} \mid \langle c, x \rangle \leq \theta\}$

$H$ is a supporting hyperplane of $C \overset{\text{def}}{\iff} H \cap C \neq \emptyset$ and $C$ is contained in one of the closed half-spaces defined by $H$.

# Examples of supporting hyperplanes



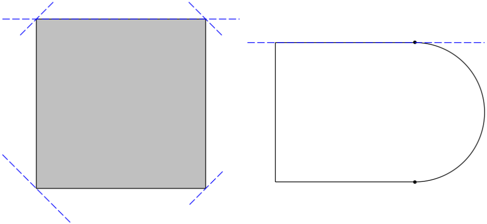



# Exposed faces

$\mathcal{F} \trianglelefteq C$ is **exposed** $\overset{\text{def}}{\iff}$ $\mathcal{F} = C \cap H$ holds for some supporting hyperplane $H$ of $C$

If all nonempty faces of $C$ are exposed we say that $C$ is facially exposed.

1. If $\mathcal{K}$ is a cone, $\mathcal{F} \trianglelefteq \mathcal{K}$ is exposed iff $\mathcal{F} = \mathcal{K} \cap \{s\}^\perp$ for some some $s \in \mathcal{K}^*$.



# Faces of $\mathcal{S}^n_+$

**Theorem**

*Let $\emptyset \neq \mathcal{F} \trianglelefteq \mathcal{S}^n_+$. Then, there exists a $n \times n$ orthogonal matrix $Q$ such that*

$$Q^\top \mathcal{F} Q = \left\{ \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \mid A \in \mathcal{S}^r_+ \right\}$$

Every nonempty face of $\mathcal{S}^n_+$ is exposed and is linearly isomorphic to a $\mathcal{S}^s_+$ for $s \leq n$.

# Minimal Face

$C, K \subseteq \mathcal{E}$: convex sets

**Definition (Minimal Face)**

Suppose $C \subseteq K$. The minimal face of $C$ with respect to $K$, is the *smallest* face of $K$ containing $C$. We write $\mathcal{F}_{\min}(C, K)$.

$$\mathcal{F}_{\min}(C, K) = \bigcap_{\substack{\mathcal{F} \trianglelefteq K \\ C \subseteq \mathcal{F}}} \mathcal{F}$$

**Key property**

Let $\emptyset \neq C \subseteq K$.

$$\mathcal{F}_{\min}(C, K) = \mathcal{F} \iff C \subseteq \mathcal{F} \text{ and } C \cap \operatorname{ri} \mathcal{F} \neq \emptyset.$$

# Facial Reduction - The basic idea

$$\sup_{y\in\mathbb{R}^m} \quad \langle b, y\rangle \qquad \text{(D)}$$
$$\text{s.t.} \quad c - \mathcal{A}^* y \in \mathcal{K}.$$

- Let $\mathcal{F}_{\text{D}} = \{c - \mathcal{A}^* y \mid c - \mathcal{A}^* y \in \mathcal{K}\} = (c + \operatorname{range} \mathcal{A}^*) \cap \mathcal{K}$, this are the feasible slacks of (D).
- We define the minimal face of (D) as $\mathcal{F}^D_{\text{min}} = \mathcal{F}_{\text{min}}(\mathcal{F}_{\text{D}}, \mathcal{K})$.
- Note: $\mathcal{F}^D_{\text{min}} = \mathcal{K} \iff$ (D) satisfies Slater's condition.

$$\inf_x \quad \langle c, x\rangle \qquad (\hat{\text{P}})$$
$$\text{subject to} \quad \mathcal{A}x = b$$
$$x \in (\mathcal{F}^D_{\text{min}})^*$$

$$\sup_y \quad \langle b, y\rangle \qquad (\hat{\text{D}})$$
$$\text{subject to} \quad c - \mathcal{A}^* y \in \mathcal{F}^D_{\text{min}}.$$

- Now, $(\hat{\text{D}})$ satisfies Slater's condition.

# Facial Reduction - Example

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

$$\mathcal{F}_{\mathrm{D}} = \left\{ \begin{pmatrix} t & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \begin{pmatrix} t & 1 \\ 1 & 1 \end{pmatrix} \succeq 0 \right\}$$

$$\mathcal{F}^D_{\min} = \left\{ \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \begin{pmatrix} a & b \\ b & c \end{pmatrix} \succeq 0 \right\} = \mathcal{S}^2_+ \oplus 0.$$

# Facial Reduction - Continued

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

$$\mathcal{F}^D_{\min} = \left\{ \begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \begin{pmatrix} a & b \\ b & c \end{pmatrix} \succeq 0 \right\} = \mathcal{S}^2_+ \oplus 0.$$

$$(\mathcal{F}^D_{\min})^* = \left\{ \begin{pmatrix} a & b & * \\ b & c & * \\ * & * & * \end{pmatrix} \mid \begin{pmatrix} a & b \\ b & c \end{pmatrix} \succeq 0 \right\} = (\mathcal{S}^2_+ \oplus 0)^*.$$

# The Facial Reduction Algorithm

- So... How do we compute $\mathcal{F}^D_{\min}$ in practice?

**Answer**: separating hyperplane theorem.

$\mathcal{V}$: Polyhedral set and $K$ a convex set

- $\mathcal{V} \cap (\operatorname{ri} K) = \emptyset \Leftrightarrow \mathcal{V}$ and $K$ can be properly separated by $H$ in such a way that $H$ does not contain $K$.

# The Facial Reduction Algorithm

$$\sup_{y \in \mathbb{R}^m} \quad \sum_{i=1}^m b_i y_i \qquad \text{(D)}$$
$$\text{s.t.} \quad C - \sum_{i=1}^m y_i A_i \succeq 0.$$

- Let $\mathcal{V} = \{C - \sum_{i=1}^m y_i A_i \mid y \in \mathbb{R}^m\}$ and $\mathcal{K} = \mathcal{S}_+^n$.
- Slater's condition is **not** satisfied $\iff \mathcal{V} \cap (\operatorname{ri} \mathcal{K}) = \emptyset$.

There exists $0 \neq X \in \mathcal{S}^n$ and $\alpha \in \mathbb{R}$ such that

$$\langle X, C - \sum_{i=1}^m y_i A_i \rangle \leq \alpha \leq \langle X, Z \rangle, \quad \forall y \in \mathbb{R}^m, \forall Z \in \mathcal{S}_+^n.$$

Therefore

- $\alpha \leq 0$, $X \in \mathcal{S}_+^n$
- $\langle X, C \rangle \leq 0$ and $\langle X, A_i \rangle = 0$, for every $i$.

# The Facial Reduction Algorithm

$$\sup_{y\in\mathbb{R}^m} \quad \sum_{i=1}^m b_i y_i \qquad \text{(D)}$$
$$\text{s.t.} \quad C - \sum_{i=1}^m y_i A_i \in \mathcal{S}_+^n.$$

Suppose Slater's condition is not satisfied, then there exists $0 \neq X_1 \in \mathcal{S}^n$ such that

- $X_1 \in \mathcal{F}_1 := \mathcal{S}_+^n$
- $\langle X_1, C\rangle \leq 0$ and $\langle X_1, A_i\rangle = 0$, for every $i$.

Two cases:

1. $\langle X_1, C\rangle < 0 \Rightarrow$ (D) is **infeasible**.
2. $\langle X_1, C\rangle = 0 \Rightarrow X_1 \notin (\mathcal{S}_+^n)^\perp$, so $\mathcal{F}_\mathrm{D} \subseteq \mathcal{S}_+^n \cap \{X_1\}^\perp \subsetneq \mathcal{S}_+^n$. $\mathcal{F}_2 := \mathcal{S}_+^n \cap \{X_1\}^\perp$ is a *face* of $\mathcal{S}_+^n$ that is smaller than $\mathcal{S}_+^n$.

# The Facial Reduction Algorithm

$$\sup_{y\in\mathbb{R}^m} \quad \sum_{i=1}^m b_i y_i \qquad (D_2)$$
$$\text{s.t.} \quad C - \sum_{i=1}^m y_i A_i \in \mathcal{F}_2.$$

If Slater's condition is not satisfied for $(D_2)$, then there exists $0 \neq X_2 \in \mathcal{S}^n$ such that

- $X_2 \in (\mathcal{F}_2)^*$.
- $\langle X_2, C\rangle \leq 0$ and $\langle X_2, A_i\rangle = 0$, for every $i$.

Two cases:

1. $\langle X_2, C\rangle < 0 \Rightarrow$ (D) is infeasible.
2. $\langle X_2, C\rangle = 0 \Rightarrow X_2 \notin (\mathcal{F}_2)^\perp$, so $\mathcal{F}_D \subseteq \mathcal{F}_2 \cap \{X_2\}^\perp \subsetneq \mathcal{F}_2$.
   $\mathcal{F}_3 = \mathcal{F}_2 \cap \{X_2\}^\perp$ is a *face* of $\mathcal{S}^n_+$ that is smaller than $\mathcal{F}_2$.

# The Facial Reduction Algorithm

$$\begin{aligned} \sup_{y \in \mathbb{R}^m} \quad & \sum_{i=1}^m b_i y_i \\ \text{s.t.} \quad & C - \sum_{i=1}^m y_i A_i \in \mathcal{F}_3. \end{aligned} \tag{$D_3$}$$

If Slater's condition is not satisfied for $(D_3)$, then there exists $0 \neq X_3 \in \mathcal{S}^n$ such that

- $X_3 \in (\mathcal{F}_3)^*$.
- $\langle X_3, C \rangle \leq 0$ and $\langle X_3, A_i \rangle = 0$, for every $i$.

Two cases:

1. $\langle X_3, C \rangle < 0 \Rightarrow$ (D) is infeasible.
2. $\langle X_3, C \rangle = 0 \Rightarrow X_3 \notin (\mathcal{F}_3)^\perp$, so $\mathcal{F}_D \subseteq \mathcal{F}_3 \cap \{X_3\}^\perp \subsetneq \mathcal{F}_3$.
   $\mathcal{F}_4 := \mathcal{F}_3 \cap \{X_3\}^\perp$ is a *face* of $\mathcal{S}_+^n$ that is smaller than $\mathcal{F}_3$.

# The Facial Reduction Algorithm - General form

**Assumptions**:

$$(c + \operatorname{range} \mathcal{A}^*) \cap \mathcal{K} \neq \emptyset.$$

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

# Facial Reduction - Continued

$$\begin{aligned} \sup_{t,s} \quad & -s \qquad \text{(D)} \\ \text{s.t.} \quad & \begin{pmatrix} t & 1 & s-1 \\ 1 & s & 0 \\ s-1 & 0 & 0 \end{pmatrix} \succeq 0 \end{aligned}$$

- We can take $X_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

# The Facial Reduction Algorithm

- If (D) is feasible, the algorithm construct a chain of faces:

$$\mathcal{F}^D_{\min} = \mathcal{F}_\ell \subsetneq \cdots \subseteq \mathcal{F}_1 = \mathcal{K}.$$

  Therefore, the Facial Reduction Algorithm always finds the minimal face $\mathcal{F}^D_{\min}$.

# Part 3 - Bonus contents

# Farkas Lemma' in LP

$$\nexists y \text{ s.t., } \; c - \mathcal{A}^* y \geq 0 \iff \exists x \geq 0, \text{s.t.} \quad \langle c, x \rangle = -1, \mathcal{A}x = 0$$

Let $e := (1, 1, \ldots, 1)$.

## Proof.

$$\begin{aligned} \inf_{x} \quad & \langle c, x \rangle & \text{(P)} \\ \text{subject to} \quad & \mathcal{A}x = 0 \\ & x_1 + \cdots + x_n = 1 \\ & x \geq 0 \end{aligned}$$

$$\begin{aligned} \sup_{t,y} \quad & t & \text{(D)} \\ \text{subject to} \quad & c - te - \mathcal{A}^* y \geq 0. \end{aligned}$$

First, (D) is always feasible.
$\theta_D < 0 \iff \nexists y \text{ s.t.}, c - \mathcal{A}^* y \geq 0$
By LP strong duality,
$\theta_D < 0 \iff \exists x^* \geq 0, \langle c, x^* \rangle = \theta_D, \mathcal{A}x^* = 0, x_1^* + \cdots + x_n^* = 1.$ (Divide $x^*$ by $-\theta_D$) □

# Find the problem in the "proof" below

$$\nexists y \text{ s.t., } C - \mathcal{A}^* y \succeq 0 \iff \exists X \succeq 0, \text{s.t. } \langle C, X \rangle = -1, \mathcal{A}X = 0$$

Let $I$ be the identity matrix.

**"Proof."**

$$\begin{aligned} \inf_{X} \quad & \langle C, X \rangle && \text{(P)} \\ \text{subject to} \quad & \mathcal{A}X = 0 \\ & \mathrm{trace}(X) = 1 \\ & X \succeq 0 \end{aligned}$$

$$\begin{aligned} \sup_{t,y} \quad & t && \text{(D)} \\ \text{subject to} \quad & C - tI - \mathcal{A}^* y \succeq 0. \end{aligned}$$

First, (D) is always feasible and satisfies Slater.
$\theta_D < 0 \iff \nexists y \text{ s.t.}, C - \mathcal{A}^* y \succeq 0$
By CLP strong duality under Slater,
$\theta_D < 0 \iff \exists X^* \geq 0, \langle C, X^* \rangle = \theta_D, \mathcal{A}X^* = 0, \mathrm{trace}(X) = 1.$ (Divide $X^*$ by $-\theta_D$)

# Farkas' Lemma in SDP?

$$\nexists y \text{ s.t. }, \ C - \mathcal{A}^* y \succeq 0 \overset{?}{\Longleftrightarrow} \exists X \succeq 0, \text{ s.t. } \langle C, X \rangle = -1, \mathcal{A}X = 0$$

$$\begin{aligned} \sup_t \quad & 0 \\ \text{s.t.} \quad & \begin{pmatrix} t & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - t \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \succeq 0 \end{aligned} \qquad \text{(D)}$$

However, $\langle C, X \rangle = -1, \mathcal{A}X = 0 \Rightarrow X_{12} = -0.5$, $X_{11} = 0$, $X$ cannot be positive semidefinite.

- (D) is infeasible but there is no $X \succeq 0$ with $\langle C, X \rangle = -1, \mathcal{A}X = 0$
- (D) is **weakly infeasible**, i.e., $(C + \operatorname{range} \mathcal{A}) \cap \mathcal{K} = \emptyset$ but $\operatorname{dist}(C + \operatorname{range} \mathcal{A}, \mathcal{K}) = 0$.

# The Facial Reduction Algorithm Again

**Assumptions**: ~~$(c + \operatorname{range} \mathcal{A}^*) \cap \mathcal{K} \neq \emptyset$~~.

1. Let $\mathcal{F}_1 = \mathcal{K}$ and $i \leftarrow 1$.
2. If $(c + \operatorname{range} \mathcal{A}^*) \cap \operatorname{ri} \mathcal{F}_i \neq \emptyset$, we are done, $\mathcal{F}_i$ is the minimal face.
3. If $(c + \operatorname{range} \mathcal{A}^*) \cap \operatorname{ri} \mathcal{F}_i = \emptyset$, then we invoke the (partial polyhedral) proper separation theorem.

   There exists $x_i \in \mathcal{E}$ and $\alpha \in \mathbb{R}$ such that

   $$\langle x_i, c - \mathcal{A}^* y \rangle \leq \alpha \leq \langle x_i, z \rangle, \quad \forall y \in \mathbb{R}^m, \forall z \in \mathcal{F}_i$$

   Therefore

   - $\alpha \leq 0$, $x_i \in \mathcal{F}_i^*$
   - $\langle x_i, c \rangle \leq 0$ and $\mathcal{A} x_i = 0$.

   Two cases:

   a. If $\alpha < 0$, then $(c + \operatorname{range} \mathcal{A}^*) \cap \mathcal{K} = \emptyset$ (Infeasibility detected)
   b. If $\alpha = 0$, then $x_i \notin \mathcal{F}_i^\perp$ holds and we let $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{x_i\}^\perp$ and $i \leftarrow i + 1$. Go to Step 2. ($c + \operatorname{range} \mathcal{A}^* \subseteq \{x_i\}^\perp$ holds)

# The FR Farkas' Lemma

"$c - \mathcal{A}^* y \in \mathcal{K}$" is infeasible **if and only if** there are $x_1, \ldots, x_\ell$ such that

- $x_i \in \mathcal{F}_i^* \cap \ker \mathcal{A} \cap \{c\}^\perp$, for $i = 1, \ldots, \ell - 1$, where
  - $\mathcal{F}_1 = \mathcal{K}$
  - $\mathcal{F}_i = \mathcal{F}_{i-1} \cap \{x_{i-1}\}^\perp$, for $i \geq 2$.
- $x_\ell \in \mathcal{F}_\ell^* \cap \ker \mathcal{A}$ and $\langle c, x_\ell \rangle = -1$.

**Theorem**

*An infeasible CLP has a finite certificate of infeasibility.*

# Example

$$\begin{aligned} \sup_t \quad & 0 \\ \text{s.t.} \quad & \begin{pmatrix} t & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - t \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \succeq 0 \end{aligned} \tag{D}$$

- $X_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{S}^2_+ \cap \ker \mathcal{A} \cap \{C\}^\perp$
- $X_2 = \begin{pmatrix} 0 & -0.5 \\ -0.5 & 0 \end{pmatrix} \in (\mathcal{S}^2_+ \cap \{X_1\}^\perp)^* = \left\{ \begin{pmatrix} a & * \\ * & * \end{pmatrix} \mid a \geq 0 \right\}$ and $\langle C, X_2 \rangle = -1$.

$X_1$ and $X_2$ form a certificate that (D) is infeasible.

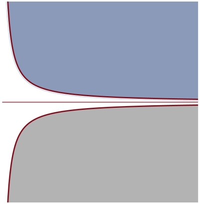
# Roots of bad behavior in CLP

$$\mathrm{dist}\,(U, V) := \inf_{x \in U, y \in V} \|x - y\|$$

Note that $\mathrm{dist}\,(U, V) = \mathrm{dist}\,(0, U - V)$.

- $\mathrm{dist}\,(U, V) = 0 \Leftrightarrow 0 \in \mathrm{cl}\,((U - V))$
- $U \cap V = \emptyset$ and $\mathrm{dist}\,(U, V) = 0 \Rightarrow U - V$ **is not closed.**

Many strange phenomena in CLP can be traced to the lack of closedness of certain maps or sums



# Example 1 - Failure of Farkas' Lemma

$$C - \mathcal{A}t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - t \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} t & 1 \\ 1 & 0 \end{pmatrix} \succeq 0$$

- $(C + \operatorname{range} \mathcal{A}) \cap \mathcal{S}^2_+ = \emptyset$ but $\operatorname{dist}(C + \operatorname{range} \mathcal{A}, \mathcal{S}^2_+) = 0$, so $\mathcal{S}^2_+ - \operatorname{range} \mathcal{A} - C$ **is not closed.**
  - In particular, $\mathcal{S}^2_+ + \operatorname{range} \mathcal{A}$ **is not closed.**

# Example 2 - Unattained optima

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

$\theta_D = 0$ but there is no feasible solution with $s = 0$.
Define

$$\hat{\mathcal{A}}(t,s) := \left(-s, \begin{pmatrix} t & 0 \\ 0 & s \end{pmatrix}\right)$$

and

$$\hat{C} := \left(-\theta_D, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)$$

Then, $\operatorname{dist}(\hat{C} + \operatorname{range}\hat{\mathcal{A}}, 0 \times \mathcal{S}^2_+) = \operatorname{dist}(\hat{C}, 0 \times \mathcal{S}^2_+ + \operatorname{range}\hat{\mathcal{A}}) = 0$

- $(\hat{C} + \operatorname{range}\hat{\mathcal{A}}) \cap (0 \times \mathcal{S}^2_+) = \emptyset$, so $(0 \times \mathcal{S}^2_+) + \operatorname{range}\hat{\mathcal{A}}$ **is not closed.**

# Fundamental questions

**A common pattern**:

- Strange thing happens $\Rightarrow \mathcal{K} + \mathcal{L}$ fails to be closed, for a certain closed convex cone $\mathcal{K}$ and subspace $\mathcal{L}$.

**Fundamental question**

Given convex cones $\mathcal{K}_1, \mathcal{K}_2$ when is $\mathcal{K}_1 + \mathcal{K}_2$ closed?

Let $S(x, y) := x + y$.

- $\mathcal{K}_1 + \mathcal{K}_2$ is closed $\iff$ $S(\mathcal{K}_1 \times \mathcal{K}_2)$ is closed.

**Fundamental question 2**

Let $\mathcal{K}$ be a convex cone and $M$ a linear map. When is $M\mathcal{K}$ closed?

# A classical result

If $\mathrm{ri}\,(\mathcal{K}_1^*) \cap \mathrm{ri}\,(\mathcal{K}_2^*) \neq \emptyset$ then $\mathcal{K}_1 + \mathcal{K}_2$ is closed.

Proof. See exercise list.

# Nice cones

For $\mathcal{F} \trianglelefteq \mathcal{K}, \mathcal{F} \neq \emptyset$ we have

$$\mathcal{F} = \mathcal{K} \cap \operatorname{span} \mathcal{F}.$$

Therefore

$$\mathcal{F}^* = \operatorname{cl}(\mathcal{K}^* + \mathcal{F}^\perp).$$

$\mathcal{K}$ is **nice** $\iff \mathcal{F}^* = \mathcal{K}^* + \mathcal{F}^\perp, \forall \mathcal{F} \trianglelefteq \mathcal{K} \iff$
$\mathcal{K}^* + \mathcal{F}^\perp$ is closed, $\forall \mathcal{F} \trianglelefteq \mathcal{K}$.

- $\mathbb{R}^n_+, \mathcal{Q}^n, \mathcal{S}^n_+$ (and all symmetric cones) are nice.
- Many applications we will not discuss here: extended duals, lifts of convex sets...

# Preliminary - Conjugate Faces

Let $\mathcal{F} \trianglelefteq \mathcal{K}$, $\mathcal{F} \neq \emptyset$.

The conjugate face of $\mathcal{F}$ is the face $\mathcal{F}^\Delta := \mathcal{K}^* \cap \mathcal{F}^\perp$

- (Exercise) $\mathcal{F}^\Delta = \mathcal{K}^* \cap \{x\}^\perp$ holds for $x \in \operatorname{ri} \mathcal{F}$. In particular $\mathcal{F}^\Delta$ is an exposed face of $\mathcal{K}^*$.

**Example** Let $\mathcal{F} \trianglelefteq \mathcal{S}^n_+$ be such that

$$\mathcal{F} = \left\{ \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \mid A \in \mathcal{S}^r_+ \right\}$$

Then

$$\mathcal{F}^\Delta = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix} \mid C \in \mathcal{S}^{n-r}_+ \right\}$$

# A closedness criterion by Pataki

$\mathcal{K}_1, \mathcal{K}_2$: closed convex **nice** cones.
Let $x \in \mathrm{ri}\,(\mathcal{K}_1 \cap \mathcal{K}_2)$, $\mathcal{F}_1 := \mathcal{F}_{\min}(x, \mathcal{K}_1)$ and $\mathcal{F}_2 := \mathcal{F}_{\min}(x, \mathcal{K}_2)$. Then

$\mathcal{K}_1^* + \mathcal{K}_2^*$ is closed if and only if $\mathcal{F}_1^\Delta + \mathcal{F}_2^\Delta = \mathcal{F}_1^\perp + \mathcal{F}_2^\perp$.

G. Pataki,
On the closedness of the linear image of a closed convex cone,
Math. Oper. Res., 32 (2007), pp. 395–412.

# Other exposedness properties

- Nices cones are nice, but niceness is hard to check.
- There are simpler **sufficient conditions**: *projectional exposedness*, *amenability*

J. M. Borwein and H. Wolkowicz.
Regularizing the abstract convex program.
*Journal of Mathematical Analysis and Applications*, 83(2):495 – 530, 1981.

B. F. Lourenço,
Amenable cones: error bounds without constraint qualifications,
Mathematical Programming, 186 (2021), pp. 1–48,

# Amenability

## Definition (Amenable cones)

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

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

Or, equivalently, if there is $\kappa > 0$ such that

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

Amenable cones are (particularly) nice


B. F. Lourenço,
Amenable cones: error bounds without constraint qualifications,
Mathematical Programming, 186 (2021), pp. 1–48,

B. F. Lourenço, V. Roshchina, and J. Saunderson.
Amenable cones are particularly nice.
SIAM J. Optim., 32(3):2347–2375, September 2022. arXiv:2011.07745.

# A comparison table

| | | Exposed | Nice | Amenable | Projectionally |
|---|---|---|---|---|---|
| Preserved under | finite intersections | ✓ | ✓ | ✓ | ? |
| | direct product | ✓ | ✓ | ✓ | ✓ |
| | injective linear image | ✓ | ✓ | ✓ | ✓ |
| Symmetric cones | | ✓ | ✓(CT'08) | ✓ | ✓L'21 |
| Homogeneous cones | | ✓ | ✓(CT'08) | ✓LRS'22 | ? |
| Hyperbolicity cones | | ✓(R'05) | ✓ | ✓LRS'23 | ? |

- Facially exposed $\overset{\text{P'13}}{\Leftarrow}$ Nice $\overset{\text{L'21}}{\Leftarrow}$ **Amenable** $\overset{\text{EPBR}}{\Leftarrow}$ Projectionally exposed.
- There exists a 4D cone that is facially exposed but not nice (Roschina, SIOPT'14).
- There exists a 4D cone that is nice but not amenable LRS'22
- In dimension 4 or less: Amenable $\Leftrightarrow$ Projectionally exposed. LRS'22

# Feasibility vs Optimization

**Optimization problem**:

$$\begin{array}{rl} \sup\limits_{y} & \langle b, y\rangle \\ \text{subject to} & c - \mathcal{A}^* y \in \mathcal{K} \end{array}$$

**Feasibility problem**:

$$\begin{array}{rl} \text{find} & y \\ \text{subject to} & c - \mathcal{A}^* y \in \mathcal{K} \end{array}$$

Are optimization problems harder than feasibility problems?

Depends, but in a very important sense **no**.

# In Linear Programming

Consider two oracles:

- **Feasibility Oracle**: Receives LP data and returns a feasible solution if one exists or **NO** if no solution exists.
- **Optimization Oracle**: Receives LP data and returns an optimal solution if one exists or **NO** if no solution exists.

1 call to **Feasibility Oracle** is enough to simulate the **Optimization Oracle**

**Proof.**

Ask the **Feasibility Oracle** for a solution to the KKT system $\{(x, y) \mid Ax = b, c - \mathcal{A}^T y \geq 0, x \geq 0, \langle c, x \rangle - \langle b, y \rangle = 0\}$. □

# In Linear Programming

Maybe you thought that was unfair. How about this?

- **Dangerous Feasibility Oracle**: Receives LP data and returns a feasible solution if one exists or **EXPLODES** if no solution exists.
- **Optimization oracle**: Receives LP data and returns an optimal solution if one exists or **NO** if no solution exists.

Can **Dangerous Feasibility Oracle** simulate **Optimization Oracle** in finite calls (without exploding)?

The short answer is **yes**.

# In Conic Linear Programming

Consider two oracles:

- **Dangerous Feasibility Oracle**: Receives CLP data and returns a feasible solution if one exists or **EXPLODES** if no solution exists.
- **Optimization Oracle**: Receives CLP data returns an optimal solution if one exists or **NO** if no solution exists.

KKT trick no longer works because (P) or (D) may be unattained and/or there may be a duality gap.

# A FR subproblem

The directions appearing in facial reduction can be found by solving the following subproblem

$$
\begin{aligned}
\inf_{x,t,w}\quad & t && && (P_{\mathcal{K}})\\
\text{subject to}\quad & -\langle c, x - te^*\rangle + t - w && = 0 && (1)\\
& \langle e, x\rangle + w && = 1 && (2)\\
& \mathcal{A}x - t\mathcal{A}e^* && = 0 && (3)\\
& (x, t, w) \in \mathcal{K}^* \times \mathbb{R}_+ \times \mathbb{R}_+
\end{aligned}
$$

$$
\begin{aligned}
\sup_{y_1,y_2,y_3}\quad & y_2 && (D_{\mathcal{K}})\\
\text{subject to}\quad & cy_1 - ey_2 - \mathcal{A}^\top y_3 \in \mathcal{K} && (4)\\
& 1 - y_1(1 + \langle c, e^*\rangle) + \langle e^*, \mathcal{A}^\top y_3\rangle \geq 0 && (5)\\
& y_1 - y_2 \geq 0 && (6)
\end{aligned}
$$

It has the following properties:

- Slater's condition is satisfied at both sides. Common optimal value is finite.
- **KKT trick** works and **Dangerous Feasibility Oracle** never explodes.

# Dangerously doing Facial Reduction

FR applied to $c - \mathcal{A}^* y \in \mathcal{K}$

1. Let $\mathcal{F}_1 = \mathcal{K}$ and $i \leftarrow 1$.
2. We invoke the **Dangerous Feasibility oracle with the KKT trick applied to the auxiliary problems** to get either $y$ such that $c - \mathcal{A}^* y \in \operatorname{ri} \mathcal{F}_i$ (in this case, we stop) or $x_i$, $\alpha$ such that

$$\langle x_i, c - \mathcal{A}^* y \rangle \leq \alpha \leq \langle x_i, z \rangle, \quad \forall y \in \mathbb{R}^m, \forall z \in \mathcal{F}_i$$

Therefore

- $\alpha \leq 0$, $x_i \in \mathcal{F}_i^*$
- $\langle x_i, c \rangle \leq 0$ and $\mathcal{A} x_i = 0$.

Two cases:

(a) If $\alpha < 0$, then $(c + \operatorname{range} \mathcal{A}^*) \cap \mathcal{K} = \emptyset$ (Infeasibility detected)

(b) If $\alpha = 0$, then $x \notin \mathcal{F}_i^*$ holds and we let $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{x_i\}^\perp$ and $i \leftarrow i + 1$. Go to Step 2. ($c + \operatorname{range} \mathcal{A}^* \subseteq \{x_i\}^\perp$ holds)

# Dangerous Optimization

Simulating the **Optimization Oracle**:

1. Do Facial Reduction **twice** to get a pair of problems $(\hat{D})$ $(\hat{P})$ satisfying Slater's condition. If FR declares infeasibility at some point return **NO**
2. Call the **Dangerous Feasibility Oracle** to solve the pair $(\hat{D})$ $(\hat{P})$ and obtain $\theta_D$.
3. Use FR to either compute a solution with value or $\theta_D$ or to check that none exists (return **NO** in this case).

**Dangerous Feasibility Oracle** can simulate **Optimization Oracle** with at most $O(\dim \mathcal{E})$ calls.


M. V. Ramana.
An exact duality theory for semidefinite programming and its complexity implications.
Math. Prog. 77, 1995

B. F. Lourenço, M. Muramatsu, and T. Tsuchiya.
Solving SDP completely with an interior point oracle.
Optimization Methods and Software 36(2-3), 425–471 (2021)

# Bibliography I

**Convex analysis**:


R. T. Rockafellar.
Convex Analysis.
Princeton University Press, 1997.

J.-B. Hiriart-Urruty and C. Lemaréchal.
Convex Analysis and Minimization Algorithms I: Fundamentals.
Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 1996.


**Conic programming**:


A. Ben-Tal and A. Nemirovski
Lectures on Modern Convex Optimization
MOS-SIAM Series on Optimization, 2001

# Bibliography II


MOSEK ApS.
MOSEK Modeling Cookbook Release 3.3.0, 2023.
URL:
https://docs.mosek.com/modeling-cookbook/index.html.


**Semidefinite programming**:


M. Todd
Semidefinite optimization.
Acta Numerica, 10:515 – 560, 2001.

M. X. Goemans and D. P. Williamson
Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming
Journal of the ACM, 42(6): 1115 –1145, 1995

# Bibliography III


M. V. Ramana.
An exact duality theory for semidefinite programming and its complexity implications.
Math. Prog. 77, 1995

L. Vandenberghe and S. Boyd
Semidefinite Programming
SIAM Review, 38:49–95, 1996


**Theory of cones**


G. P. Barker,
Theory of cones,
Linear Algebra and its Applications, 39 (1981), pp. 263 – 291.

C.-H. Sung and B.-S. Tam,
A study of projectionally exposed cones,
Linear Algebra and its Applications, 139 (1990), pp. 225 – 252.

# Bibliography IV


G. Pataki,
The geometry of semidefinite programming, in Handbook of semidefinite programming: theory, algorithms, and applications, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds.,
Kluwer Academic Publishers, online version at
http://www.unc.edu/~pataki/papers/chapter.pdf, 2000.

J. Renegar.
Hyperbolic programs, and their derivative relaxations.
*Foundations of Computational Mathematics*, 6(1):59–79, 2005.

G. Pataki,
On the closedness of the linear image of a closed convex cone,
Math. Oper. Res., 32 (2007), pp. 395–412.

C. B. Chua and L. Tunçel.
Invariance and efficiency of convex representations.
Math. Program., 111(1-2, Ser. B):113–140, 2008.

# Bibliography V

G. Pataki,
On the connection of facially exposed and nice cones,
J. Math. Anal. Appl., 400 (2013), pp. 211–221.

V. Roshchina.
Facially exposed cones are not always nice.
SIAM J. Optim., 24(1):257–268, 2014.

B. F. Lourenço,
Amenable cones: error bounds without constraint qualifications,
Math. Prog., 186 (2021), pp. 1–48,

B. F. Lourenço, V. Roshchina, and J. Saunderson.
Amenable cones are particularly nice.
SIAM J. Optim., 32(3):2347–2375, September 2022.
arXiv:2011.07745.

# Bibliography VI

B. F. Lourenço, V. Roshchina, and J. Saunderson.
Hyperbolicity cones are amenable.
Math. Prog., 2023. arXiv:2102.063595.

**Facial reduction**:

J. M. Borwein and H. Wolkowicz.
Regularizing the abstract convex program.
Journal of Mathematical Analysis and Applications, 83(2):495 – 530, 1981.

H. Waki and M. Muramatsu.
Facial reduction algorithms for conic optimization problems.
Journal of Optimization Theory and Applications, 158(1):188–215, 2013.

# Bibliography VII


G. Pataki.
Strong duality in conic linear programming: Facial reduction and extended duals.
In Computational and Analytical Mathematics, volume 50, pages 613–634. Springer New York, 2013.

M. Liu and G. Pataki,
Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming,
Math. Program., 167 (2018), pp. 435–480.

B. F. Lourenço, M. Muramatsu, and T. Tsuchiya.
Facial reduction and partial polyhedrality.
SIAM Journal on Optimization, 28(3):2304–2326, 2018.

B. F. Lourenço, M. Muramatsu, and T. Tsuchiya.
Solving SDP completely with an interior point oracle.
Optimization Methods and Software 36(2-3), 425–471 (2021)