Lecture 7

Last time
- Fenchel conjugate
- Fenchel duality

Today
- Conic optimization
- Examples
- Duality

Conic optimization

To illustrate Fenchel duality, we consider a broad template. Consider the primal

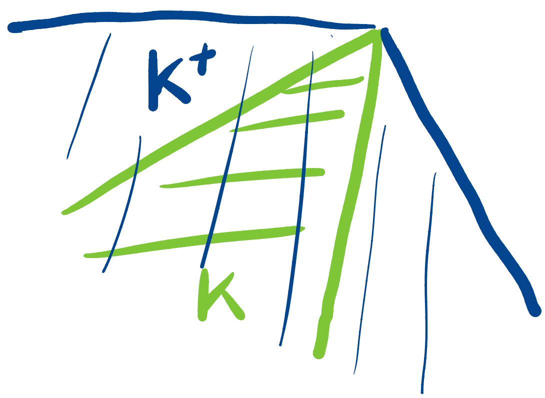
$$p^\star = \begin{cases} \inf\limits_{x} & \langle c, x\rangle \\ \text{s.t.} & Ax \in b + H \\ & x \in K. \end{cases}$$

$c \in E$ $\quad$ $b \in Y$

Convex cones $\left(\begin{matrix} K + K \subseteq K \\ tK \subseteq K \quad \forall t \geq 0 \end{matrix} \right)$

Linear map $A : E \to Y$.

Define the **dual cone** $K^+ = \{x \in E \mid \langle x, y\rangle \geq 0 \ \forall y \in K\}$.



Notice that $E \times Y$ is also an Euclidean space with inner product given by

$$\langle (x,y),(x',y')\rangle := \underbrace{\langle x, x'\rangle}_{\text{inner prod. in } E} + \underbrace{\langle y, y'\rangle}_{\text{inner prod. in } Y}.$$

Then, we can see that $K \times H$ is a cone in $E \times Y$ and, further, $(K \times H)^+ = K^+ \times H^+$.

Examples of cones

▷ Nonnegative orthant $\mathbb{R}^n_+$.

This cone models Linear programming (LP)

$$\min \langle c, x\rangle \quad \text{s.t.} \quad Ax \le b$$

Equivalent to $Ax \in b - \mathbb{R}^m_+$

We can always write an LP (after a a transformation of A, b, and c) as

$$\min \langle c, x\rangle \quad \text{s.t} \quad Ax = b, \quad x \ge 0.$$

(Why?)

We already saw an example of LPs in the syllabus. It has many more applications in logistics, economics, networks, among others.

▷ Second order cone (Ice cream cone)

Consider

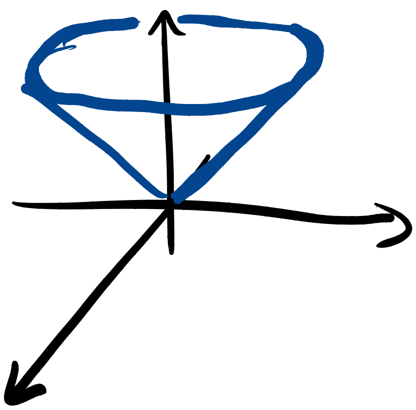
$$SO^n_+ = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\| \le r\}$$

This models second order cone programing (SOCP).

$$\min \; c^T x \quad \Longleftrightarrow \quad \begin{cases} y = A_i x \\ t = f_i^T x \\ (y,t) \in (b_i, -d_i) + SO_+^k \end{cases}$$

$$\text{s.t.} \quad \|A_i x - b_i\| \leq f_i^T x + d_i \quad \forall i \in [m]$$

Here $E = \mathbb{R}^n$, $A_i \in \mathbb{R}^{k_i \times n}$, $b_i \in \mathbb{R}^{k_i}$, $c \in \mathbb{R}^n$, $f_i \in \mathbb{R}^n$ and $d_i \in \mathbb{R}$.



One can encode any LP as an SOCP by setting $A_i = 0$.

Sub example: Group LASSO

Statisticians often encounter problems of the form

$$(\ddot{o}) \qquad \min_{\beta} \; \|\underbrace{X}_{\mathbb{R}^{n\times p}}\beta - \underbrace{y}_{\mathbb{R}^n}\|_2^2 + \lambda \sum_{g=1}^{G} \|\underbrace{\beta_{I_g}}_{\text{vector with indices in } I_g.}\|_2$$

where $I_g \subseteq [p]$ $\forall p \in [G]$ are subsets of indices. This is a funda mental problem for variable selection.

Exercise: show that (☺) can be written as an SOCP. ⊣

▷ Semidefinite cone

Consider the cone

$$S_+^n = \{ M \in S^n \mid \underbrace{x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n}_{\text{denoted } M \succeq 0} \}.$$

This cone models semidefinite programming (SDP):

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \mathcal{A}(X) = b \\ & X \succeq 0 \end{aligned}$$

trace inner product

Linear map

$\mathcal{A} : S_+^n \to \mathbb{R}^m$

Fact: The following equivalence holds:

$$\|x\|_2 \leq t \iff \begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \succeq 0. \qquad ⊣$$

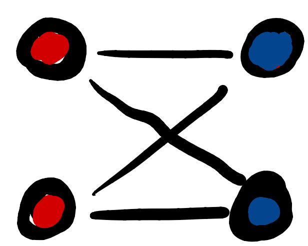
This fact follows easily from a Schur complement computation.
Thus any SOCP is also an SDP.

Subexample: Max cut

Suppose we had an undirected weighted graph on n nodes and we wanted to find a subset of the nodes $S \subseteq [n]$ that maximizes

cut-value(S) :

$$\max_{S} \sum_{\substack{i \in S \\ j \in S^c}} w_{ij} .$$



This can be modeled as an integer optimization problem

$$\max \sum w_{ij} \left(\frac{1 - x_i x_j}{2} \right)$$
$$\text{s.t.} \quad x \in \{\pm 1\}$$

or equivalently

matrix of all 1's.

$$\mathrm{OPT} = \begin{cases} \max \frac{1}{2} \langle W, J - xx^T \rangle \\ \text{s.t.} \quad x_i^2 = 1. \end{cases}$$

vector $\mathbb{R}^n$

We can relax this problem by considering a larger set

$$\mathrm{OPT} \le \left\{ \begin{array}{l} \max \frac{1}{2} \langle W, J - VV^T \rangle \\ \text{s.t. } \|V^{(i)}\|^2 = 1. \end{array} \right. = \left\{ \begin{array}{l} \max \frac{1}{2} \langle W, J - M \rangle \\ \text{s.t. } \quad M_{ii} = 1 \\ \qquad\quad M \in S_+^n. \end{array} \right.$$

ith row.

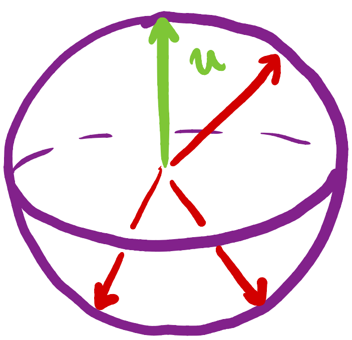
matrix $n \times n$

In turn, after solving this problem, one can get pretty good cuts via:

Goemans - Williamson

sphere

- Sample $u \in \mathrm{Unif}(\mathbb{S}^{n-1})$
- Return $\hat{S} = \{ i \mid \langle V^{(i)}, u \rangle \le 0 \}$.



Fact (Goemans-Williamson '94)

$\mathrm{OPT} \ge \mathbb{E}\, \text{cut-value}(\hat{S}) \ge 0.87856\, \mathrm{OPT}.$

Duality for conic optimization

We can use our template $\inf_x \{ f(x) + g(Ax) \}$ to write conic optimization problems

$A^*y \in c + N_K(\bar{x})$ $\langle A^*y - c, k - \bar{x}\rangle \le 0$

with $-y \in N_{b+H}(A\bar{x})$ $\langle -y, b+h - A\bar{x}\rangle \le 0$

$$f(x) = \langle c, x\rangle + \iota_K(x)$$

$$g(z) = \iota_{b+H}(z).$$

Recall that the dual was

$$\sup_y \; -f^*(A^*y) - g^*(-y)$$

Exercise: Show that the dual reduces to

$$d^* = \begin{cases} \sup_y & \langle b, y\rangle \\ \text{s.t.} & A^*y \in c - K^+ \\ & y \in H^+. \end{cases}$$

Theorem: For conic problems $p^* \ge d^*$. If H and K are convex cones and either

1) $\exists \hat{x} \in K$ such that $A\hat{x} - b \in \operatorname{int} H$.
2) $\exists \hat{x} \in \operatorname{int} K$ such that $A\hat{x} - b \in H$, and A is surjective.

Then, $p^* = d^*$ and if d^* is finite it is attained. If $(\bar{x}, \bar{y})$ are feasi blem, then, they are optimal iff

$$\langle \bar{x}, A^*\bar{y} - c\rangle = 0 \quad \text{and} \quad \langle Ax - b, \bar{y}\rangle = 0.$$

These properties doesn't always hold

Example Consider

$$\begin{aligned} \inf \quad & x_1 \\ \text{s.t.} \quad & x_2 - t = 0 \\ & (x_1, x_2, t) \in SO^2_+. \end{aligned}$$

$$\begin{aligned} 0 = \inf \quad & x_1 \\ \text{s.t.} \quad & x_2 \geq \sqrt{x_1^2 + x_2^2} \end{aligned}$$

This can be cast with $c = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $A = (0, 1, -1)$, $b = 0$, $H = \{0\}$, and $K = SO^2_+$. The dual is

$$\begin{aligned} \sup \quad & 0 \\ \text{s.t.} \quad & \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} y \in \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - SO^2_+. \end{aligned}$$

The orange formulation makes it clear that $p^* = 0$. On the other hand the dual constrained yields

$$\begin{pmatrix} 1 \\ -y \\ y \end{pmatrix} \in SO^2_+ \iff y \geq \sqrt{1 + y^2}$$

infeasible.

Thus, $d^* = -\infty$. What goes wrong?
Consider the value function

$$\nu(z) = \begin{cases} \inf_x & x_1 \\ \text{s.t.} & x_2 - t + z = 0 \\ & (x_1, x_2, t) \in SO_+^2 \end{cases} = \begin{cases} \inf & x_1 \\ \text{s.t.} & x_2 + z \geq \|x\|. \end{cases}$$

Let's consider two cases

Case 1: $z < 0$, then

$$x_2 > x_2 + z \geq \|x\| \geq |x_2|.$$

Thus, the problem is infeasible and $\nu(z) = \infty$.

Case 2: $z > 0$, then

$$x_2^2 + 2x_2 z + z^2 \geq x_1^2 + x_2^2$$

$$\Updownarrow$$

$$2x_2 z + z^2 \geq x_1^2$$

If we let $x_2 \uparrow \infty$, the upper bound is arbitrary large and $\nu(z) = -\infty$.

Thus, $\nu(z) = \begin{cases} -\infty & z > 0 \\ 0 & z = 0 \\ \infty & z < 0 \end{cases}$ and $0 \notin \operatorname{int} \operatorname{dom} \nu$.