FOUNDATIONS OF OPTIMIZATION: IE6001

Lagrangian Duality

Napat Rujeerapaiboon

Semester I, AY2022/2023

Lagrangian Duality

Optimization problem in standard form (possibly nonconvex):

$$
\begin{aligned}
\mathrm{P}: \quad & \text{minimize} && f_0(x) \\
& \text{subject to} && f_i(x) \leq 0 \quad \forall i = 1, \dots, m \\
& && h_i(x) = 0 \quad \forall i = 1, \dots, p
\end{aligned}
$$

$x \in \mathbb{R}^n$	decision variable
$f_0 : \mathbb{R}^n \to \mathbb{R}$	objective function
$f_i : \mathbb{R}^n \to \mathbb{R}$	inequality constraint functions
$h_i : \mathbb{R}^n \to \mathbb{R}$	equality constraint functions

Assume for simplicity that value cannot be ∞.

Lagrangian

The *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ for problem P is defined as

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

- ~ weighted sum of objective and constraint functions
- λ_i is *Lagrange multiplier* corresponding to $f_i(x) \leq 0$
- μ_i is *Lagrange multiplier* corresponding to $h_i(x) = 0$

The Lagrangian is *concave* (affine) in (λ, μ) for any fixed x.

If P is a *convex* optimization problem, then the Lagrangian is *convex* in x for any fixed (λ, μ).

Lagrangian

The Lagrangian enables us to reexpress P as a *min-max problem*.
To see this, we define

$$f(x) = \sup_{\lambda \in \mathbb{R}^m_+,\ \mu \in \mathbb{R}^p} L(x, \lambda, \mu)$$

and note that

$$f(x) = \begin{cases} f_0(x) & \text{if } \begin{cases} f_i(x) \leq 0 \ \ \forall i = 1, \dots, m, \\ h_i(x) = 0 \, \forall i = 1, \dots, p, \end{cases} \\ \infty & \text{else.} \end{cases}$$

Thus, we obtain

$$\inf \mathsf{P} = \inf_{x \in \mathbb{R}^n} f(x) = \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^m_+,\ \mu \in \mathbb{R}^p} L(x, \lambda, \mu).$$

The Dual Problem

Below we refer to P as the *primal problem*. Using the Lagrangian we can introduce a *dual problem* with objective function

$$g(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

The dual *maximization* problem is then defined as

$$\text{D:} \quad \underset{\lambda \in \mathbb{R}^m_+,\, \mu \in \mathbb{R}^p}{\text{maximize}} \quad g(\lambda, \mu).$$

By construction, D is equivalent to a *max-min problem*.

$$\sup \mathrm{D} = \sup_{\lambda \in \mathbb{R}^m_+,\, \mu \in \mathbb{R}^p} g(\lambda, \mu) = \sup_{\lambda \in \mathbb{R}^m_+,\, \mu \in \mathbb{R}^p} \; \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

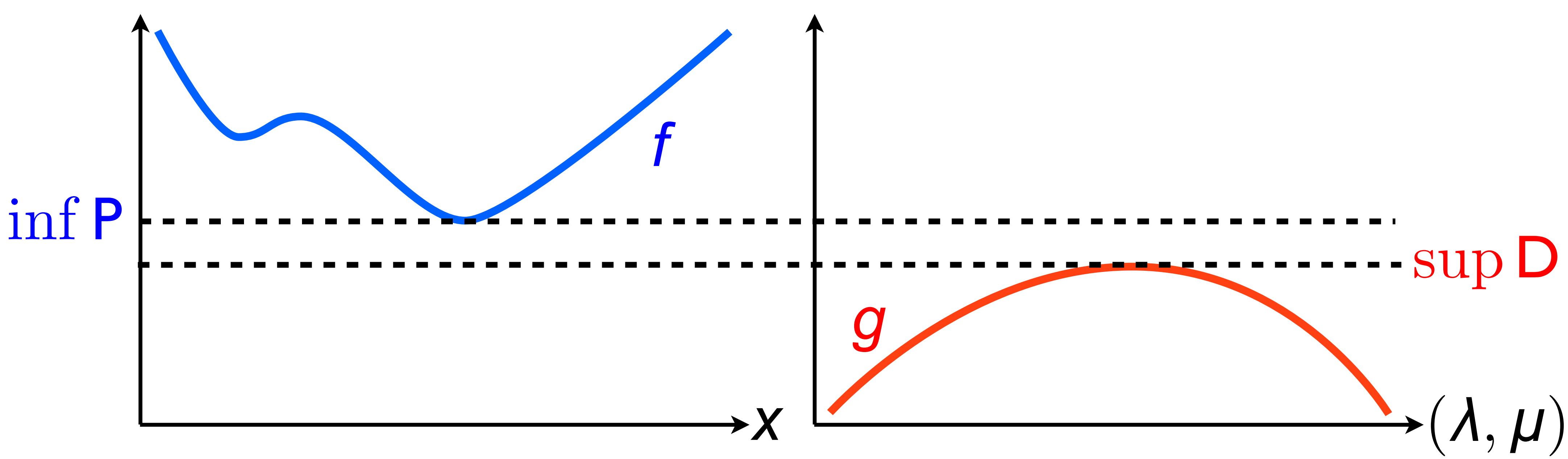
Weak Duality

Proposition: $g(\lambda, \mu) \le f(x) \quad \forall x \in \mathbb{R}^n,\ \lambda \in \mathbb{R}^m_+,\ \mu \in \mathbb{R}^p.$

Proof: By definition, we have

$$g(\lambda, \mu) = \inf_{\hat{x} \in \mathbb{R}^n} L(\hat{x}, \lambda, \mu) \le L(x, \lambda, \mu) \le \sup_{\hat{\lambda} \in \mathbb{R}^m_+,\ \hat{\mu} \in \mathbb{R}^p} L(x, \hat{\lambda}, \hat{\mu}) = f(x).$$

Corollary (Weak Duality): $\sup \mathsf{D} \le \inf \mathsf{P}$



Thus, if D is unbounded ($\sup \mathsf{D} = \infty$), P must be infeasible. If P is unbounded ($\inf \mathsf{P} = -\infty$), D must be infeasible.

Significance of Dual Solutions

Note: Every *feasible solution of* P (D) provides an *upper* (*lower*) *bound* on both $\inf$ P and $\sup$ D.

Assume $\hat{x}$ is a (feasible) candidate solution for P. Its quality is quantified by $f_0(\hat{x}) - \inf \mathsf{P}$. However, $\inf$ P is unknown!

A dual feasible solution (λ, μ) provides a *proof* or *certificate* that

$$\inf \mathsf{P} \geq g(\lambda, \mu) \quad \Longrightarrow \quad f_0(\hat{x}) - \inf \mathsf{P} \leq f_0(\hat{x}) - g(\lambda, \mu).$$

Thus, if $f_0(\hat{x}) - g(\lambda, \mu) \leq \varepsilon$, then $\hat{x}$ is an ε-optimal solution.

Strong Duality

Definition: $\Delta = \inf \mathrm{P} - \sup \mathrm{D}$ is called the *duality gap*. By weak duality, we have $\Delta \geq 0$. If $\Delta = 0$, we say that *strong duality* holds.

$$
\begin{array}{llll}
\mathrm{P}: & \text{minimize} & f_0(x) & \\
& \text{subject to} & f_i(x) \leq 0 & \forall i = 1, \ldots, m \\
& & h_i(x) = 0 & \forall i = 1, \ldots, p
\end{array}
$$

Strong duality

- does not hold in general
- but always holds if P is a *convex* problem satisfying a *constraint qualification* (proof later)

Slater's constraint qualification holds if there exists x_{S} with

- $f_i(x_{\mathrm{S}}) < 0 \quad \forall i = 1, \ldots, m,$
- $h_i(x_{\mathrm{S}}) = 0 \quad \forall i = 1, \ldots, p.$

Least Squares Problem

Primal problem: minimize $x^\top x$
subject to $Ax = b$

Lagrangian: $L(x, \mu) = x^\top x + \mu^\top (Ax - b)$

Dual objective:
$$0 = \nabla_x L(x, \mu) = 2x + A^\top \mu$$
$$\implies \quad x = -\tfrac{1}{2} A^\top \mu$$
$$\implies \quad g(\mu) = -\tfrac{1}{4} \mu^\top A A^\top \mu - b^\top \mu$$

Dual problem: maximize $-\frac{1}{4} \mu^\top A A^\top \mu - b^\top \mu$

Standard Form Linear Program

Primal problem:

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Lagrangian:

$$L(x, \lambda, \mu) = c^\top x - \lambda^\top x + \mu^\top (Ax - b)$$

Dual objective:

$$g(\lambda, \mu) = \begin{cases} -b^\top \mu & \text{if } c - \lambda + A^\top \mu = 0 \\ -\infty & \text{else} \end{cases}$$

Dual problem:

$$\left.\begin{array}{ll} \text{maximize} & -b^\top \mu \\ \text{subject to} & c - \lambda + A^\top \mu = 0 \\ & \lambda \geq 0 \end{array}\right\} \iff \left\{\begin{array}{ll} \text{maximize} & b^\top \mu \\ \text{subject to} & A^\top \mu \leq c \end{array}\right.$$

Quadratic Program

Primal problem:

$$\begin{array}{ll} \text{minimize} & x^\top P x \\ \text{subject to} & Ax \leq b \end{array} \qquad (P \succ 0)$$

Lagrangian:

$$L(x, \lambda) = x^\top P x + \lambda^\top (Ax - b)$$

Dual objective:

$$g(\lambda) = -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda$$

Dual problem:

$$\begin{array}{ll} \text{maximize} & -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda \\ \text{subject to} & \lambda \geq 0 \end{array}$$

Second-Order Cone Program

Primal problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f^\top x \\ \text{subject to} \quad & \|A_i x + b_i\|_2 \leq c_i^\top x + d_i \quad i = 1, \ldots, m \end{aligned}$$

Second-Order Cone Program

Primal problem:

$$\begin{array}{ll} \underset{x,y,t}{\text{minimize}} & f^\top x \\ \text{subject to} & \|y_i\|_2 \le t_i \quad i = 1, \ldots, m \\ & y_i = A_i x + b_i, \;\; t_i = c_i^\top x + d_i \quad i = 1, \ldots, m \end{array}$$

Lagrangian:

$$\begin{aligned} L(x, y, t, \lambda, v, \mu) \\ &= f^\top x + \sum_{i=1}^m \lambda_i(\|y_i\|_2 - t_i) + \sum_{i=1}^m v_i^\top (y_i - A_i x - b_i) + \sum_{i=1}^m \mu_i (t_i - c_i^\top x - d_i) \\ &= \boxed{\left(f - \sum_{i=1}^m A_i^\top v_i - \sum_{i=1}^m \mu_i c_i\right)^\top x} + \boxed{\sum_{i=1}^m (\lambda_i \|y_i\|_2 + v_i^\top y_i)} + \boxed{\sum_{i=1}^m (-\lambda_i + \mu_i) t_i} \\ &\quad - \sum_{i=1}^m (b_i^\top v_i + d_i \mu_i) \end{aligned}$$

Second-Order Cone Program

To evaluate the dual objective, we note that

- the infimum over x is finite iff

$$\sum_{i=1}^{m} A_i^\top v_i + \sum_{i=1}^{m} \mu_i c_i = f$$

- the infimum over y_i satisfies

$$\inf_{y_i} \lambda_i \|y_i\|_2 + v_i^\top y_i = \begin{cases} 0 & \text{if } \|v_i\|_2 \leq \lambda_i \\ -\infty & \text{othewise} \end{cases}$$

- the infimum over t_i is finite iff $\lambda_i = \mu_i$

Dual problem:

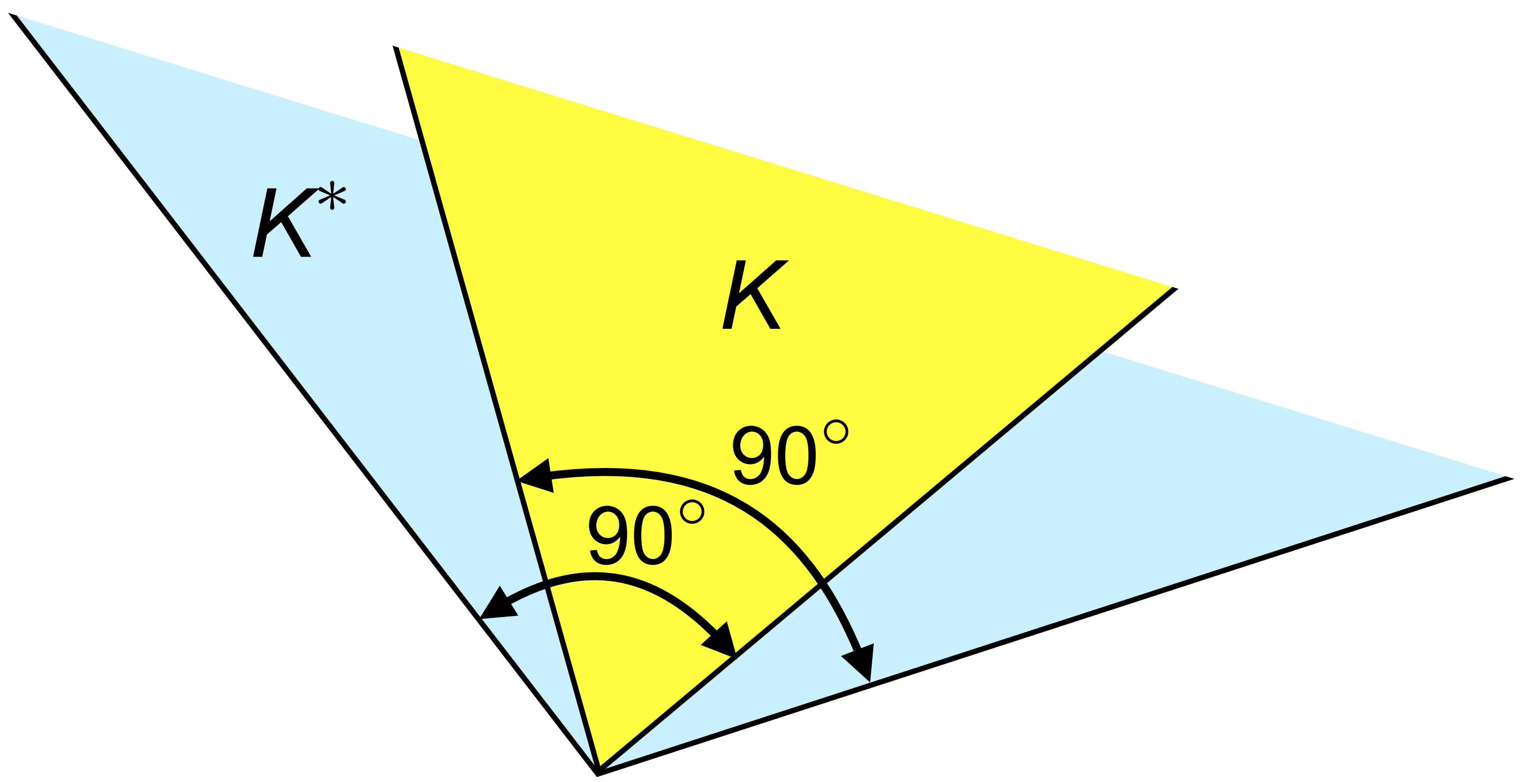
$$\begin{array}{ll} \underset{v,\,\mu}{\text{maximize}} & -\sum_{i=1}^{m}(b_i^\top v_i + d_i \mu_i) \\ \text{subject to} & \sum_{i=1}^{m}(A_i^\top v_i + c_i \mu_i) = f \\ & \|v_i\|_2 \leq \mu_i \quad \forall i = 1, \ldots, m \end{array}$$

Dual Cones

Definition: If K is a cone, then the set

$$K^* = \{y \in \mathbb{R}^n : x^\top y \geq 0 \ \forall x \in K\}$$

is the *dual cone* of K.



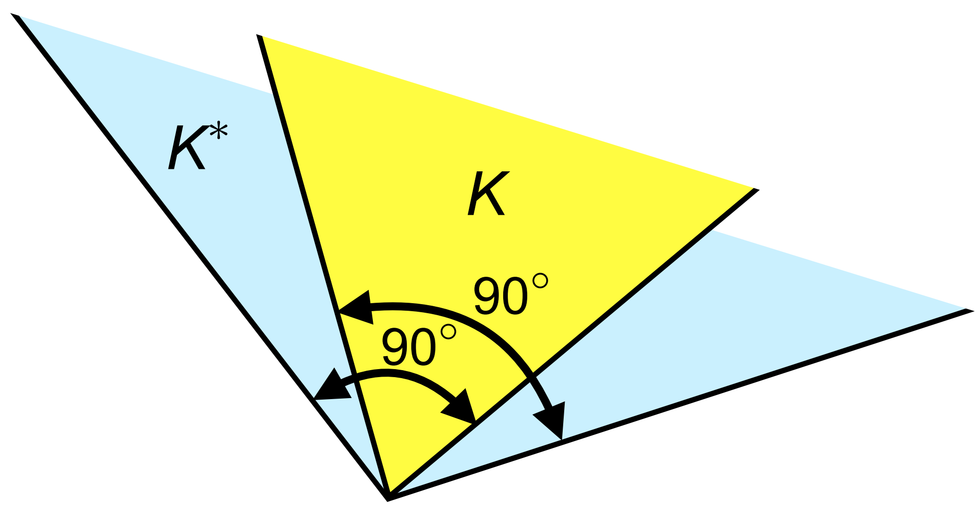
Note that K^* is a cone by construction.

Example: $(\mathbb{S}^n_+)^* = \{Y \in \mathbb{S}^n : \mathrm{tr}(X^\top Y) \geq 0 \ \forall X \in \mathbb{S}^n_+\}$

Dual Cones

Properties of the dual cone:

- K^* is closed and convex
- $K_2 \subseteq K_1 \implies K_1^* \subseteq K_2^*$ (the smaller K, the larger K^*)
- $K^{**} = \mathrm{cl}(\mathrm{conv}(K))$, the smallest convex closed superset of K
- If a convex cone K is proper, then K^* is proper and $K^{**} = K$



Definition: A cone K is called *self-dual*, if $K^* = K$.

Examples of self-dual cones: $\mathbb{R}^n_+$, the second-order cone, $\mathbb{S}^n_+$

Problems with Generalized Inequalities

$$\begin{array}{lll} \mathrm{P}: & \text{minimize} & f_0(x) \\ & \text{subject to} & f_i(x) \preceq_{K_i} 0 \quad \forall i = 1, \ldots, m \\ & & h_i(x) = 0 \quad \forall i = 1, \ldots, p \end{array}$$

where $K_i \subseteq \mathbb{R}^{r_i}$ is a proper convex cone

- assign a *Lagrange multiplier* $\lambda_i \in K_i^*$ to $f_i(x) \preceq_{K_i} 0$
- assign a *Lagrange multiplier* $\mu_i \in \mathbb{R}$ to $h_i(x) = 0$

The *Lagrangian* $L : \mathbb{R}^n \times K_1^* \times \cdots \times K_m^* \times \mathbb{R}^p \to \mathbb{R}$ is defined as:

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i^\top f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

Problems with Generalized Inequalities

$$\text{P}: \quad \inf_{x\in\mathbb{R}^n} f(x) \qquad\qquad f(x) = \sup_{\lambda\in K^*,\, \mu\in\mathbb{R}^p} L(x,\lambda,\mu)$$

$$\text{D}: \quad \sup_{\lambda\in K^*,\, \mu\in\mathbb{R}^p} g(\lambda,\mu) \qquad g(\lambda,\mu) = \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu)$$

- *Weak duality* holds.
- *Strong duality* holds for convex problems satisfying a constraint qualification (proof later).

Slater's constraint qualification holds if there exists x_{S} with

- $f_i(x_{\mathrm{S}}) \prec_{K_i} 0 \quad \forall i = 1, \dots, m,$
- $h_i(x_{\mathrm{S}}) = 0 \quad \forall i = 1, \dots, p.$

SDP Duality

Primal SDP ($F_i, G \in \mathbb{S}^m$):

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & F_1 x_1 + \cdots + F_n x_n \preceq G \end{array}$$

Lagrangian ($\Lambda \in \mathbb{S}^m_+$):

$$L(x, \Lambda) = c^\top x + \mathrm{tr}(\Lambda^\top [\textstyle\sum_{i=1}^n x_i F_i - G])$$

Dual objective:

$$g(\Lambda) = \inf_x L(x, \Lambda) = \begin{cases} -\mathrm{tr}(\Lambda^\top G) & \text{if } \mathrm{tr}(\Lambda^\top F_i) = -c_i \ \forall i = 1, \ldots, n \\ -\infty & \text{else} \end{cases}$$

Dual SDP:

$$\begin{array}{ll} \text{maximize} & -\mathrm{tr}(\Lambda^\top G) \\ \text{subject to} & \mathrm{tr}(\Lambda^\top F_i) = -c_i \ \forall i = 1, \ldots, n \\ & \Lambda \succeq 0 \end{array}$$

Main Take-Away Points

- **Lagrangian:** weighted sum of objective and constraints; saddle function for convex problems; used to construct primal objective (partial maximum w.r.t. Lagrange multipliers) and dual objective (partial minimum w.r.t. decision variables)
- **Duality:** sup D never exceeds inf P (weak duality); sup D equals inf P for convex problems satisfying a constraint qualification (strong duality); all primal (dual) feasible solutions offer upper (lower) bounds on both inf P and sup D
- **Explicit dual problems:** the dual of an LP/QP/SOCP/SDP is also an LP/QP/SOCP/SDP, respectively
- **Dual cones:** a proper convex cone coincides with the dual of its dual (bidual) cone; the nonnegative orthant, the second-order cone and the positive semidefinite cone are self-dual