Optimization 1

Lecture 12 - Conic Duality

Proper Cones

Reminder: a convex cone $K \subseteq \mathbb{R}^n$ is a set which satisfies:

- $\mathbf{x} \in K \Rightarrow \lambda \mathbf{x} \in K, \; \forall \lambda \geq 0$.
- $\mathbf{x}, \mathbf{y} \in K \Rightarrow \mathbf{x} + \mathbf{y} \in K$.

definition: A convex cone $K \subseteq \mathbb{R}^n$ is called proper if

- K is a closed set.
- K has a non-empty interior.
- If $\mathbf{x} \in K$ and $-\mathbf{x} \in K$ then $\mathbf{x} = 0$. (Pointed)

Examples:

- $\mathbb{R}^n_+$ - The nonnegative orthant.
- $L^n = \{(\mathbf{x}, t) \in \mathbb{R}^n : \|\mathbf{x}\| \leq t\}$ - The Lorenz ("ice-cream") cone.
- $\mathbb{S}^n_+ = \{\mathbf{X} \in \mathbb{S}^n : \mathbf{X} \succeq 0\}$ - The set of all symmetric PSD matrices.
- $K = K_1 \times K_2 \times \ldots \times K_m$ where K_j is a proper convex cone for all $j = 1, 2, \ldots, m$.

Why cones?

- Easier derivation of dual problem.
- Many solvers are conic solvers:
  - Use conic problem representation
  - Solution techniques using conic duality (interior point methods).

Conic constraints

For any K proper and convex cone we use the following notation

- $\mathbf{x} \in K \Leftrightarrow \mathbf{x} \succeq_K 0$
- $\mathbf{x} \in \text{int}(K) \Leftrightarrow \mathbf{x} \succ_K 0$

Thus we can write linear conic constraints of the form

- $\mathbf{x} \succeq_K \mathbf{y} \Leftrightarrow \mathbf{x} - \mathbf{y} \in K$
- $\mathbf{Ax} + \mathbf{b} \succeq_K 0 \Leftrightarrow \mathbf{Ax} + \mathbf{b} \in K$

Note that the set defined by these constraints are convex. Why?

Examples of linear conic constraints

- $\|\mathbf{x}\| \leq t$
- $\|\mathbf{x}\|^2 = yz,\ y, z \geq 0$
- $\mathbf{Y} - \mathbf{x}\mathbf{x}^\top \succeq 0$

In class.

Conic programming

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_j \succeq_{K_j} 0, \ j = 1, \ldots, m \end{aligned}$$

Properties:

- Linear objective.
- Linear equality constraints.
- K_j are proper convex cones.

Conic programming - Examples

Linear Programming:

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$

Adding slack variables $\mathbf{s} \geq 0$ we can rewrite the problem as

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{s} = \mathbf{b} \\ & \mathbf{s} \succeq_{\mathbb{R}^m_+} 0. \end{aligned}$$

Conic programming - Examples

Convex QCQP:

$$\begin{array}{ll} \min & \mathbf{x}^\top \mathbf{D}_0^\top \mathbf{D}_0 \mathbf{x} + 2\mathbf{b}_0^\top \mathbf{x} + c_0 \\ \text{s.t.} & \mathbf{x}^\top \mathbf{D}_j^\top \mathbf{D}_j \mathbf{x} + 2\mathbf{b}_j^\top \mathbf{x} + c_j \leq 0, \; j = 1, 2, \ldots, m \end{array}$$

For each $j = 0, 1, \ldots, m$ we add the variables $\mathbf{z}_j = \mathbf{D}_j \mathbf{x}$,

$\|\mathbf{z}_j\|^2 \leq y_j \Leftrightarrow \mathbf{w}_j = \begin{bmatrix} \mathbf{z}_j \\ (y_j - 1)/4 \\ (y_j + 1)/4 \end{bmatrix} \equiv \mathbf{A}_j \mathbf{z}_j + \mathbf{d}_j y_j + g_j \in L^{d_j+2}, y_j \geq 0$ we can rewrite the problem as

$$\begin{array}{lll} \min & y_0 + 2\mathbf{b}_0^\top \mathbf{x} + c_0 & \\ \text{s.t.} & y_j + 2\mathbf{b}_j^\top \mathbf{x} + c_j \leq 0, & j = 1, 2, \ldots, m, \\ & \mathbf{z}_j - \mathbf{D}_j \mathbf{x} = 0, & j = 0, 1, \ldots, m, \\ & \mathbf{w}_j - \mathbf{A}_j \mathbf{z}_j - \mathbf{d}_j y_j - g_j = 0, & j = 0, 1, \ldots, m, \\ & \mathbf{y} \succeq_{\mathbb{R}_+^{m+1}} 0, \; \mathbf{w}_j \succeq_{L^{d_j+2}} 0, & j = 0, 1, \ldots, m. \end{array}$$

Conic programming - Examples

Semidefinite programming (SDP):

$$
\begin{aligned}
\min \quad & \mathbf{A}_0 \cdot \mathbf{X} \\
\text{s.t.} \quad & \mathbf{A}_j \cdot \mathbf{X} = b_j, \; j = 1, \ldots, m, \\
& \mathbf{X} \succeq 0.
\end{aligned}
$$

- $\mathbf{A}_j$ are symmetric matrices.
- The inner product is equivalent to the vector inner product. $\mathbf{A}_j \cdot \mathbf{X} = \sum_{i=1}^n \sum_{k=1}^n [\mathbf{A}_j]_{ik} \mathbf{X}_{ik} = \mathrm{Tr}(\mathbf{A}_j^\top \mathbf{X})$.

Dual cones

definition: A dual cone of set $C \subseteq \mathbb{R}^n$ is

$$C^* = \{\mathbf{y} : \mathbf{y}^\top \mathbf{x} \geq 0, \forall \mathbf{x} \in C\}.$$

- C^* is always a convex cone. Why?
- The dual cone of $\mathbb{R}^n$ is $\{0\}$.
- If K is a proper and convex cone then so is K^*, and $(K^*)^* = K$.

Example:

$$K = \{(\mathbf{x}, t) : \|\mathbf{x}\|_p \leq t\}, \quad K^* = \{(\mathbf{y}, s) : \|\mathbf{y}\|_q \leq s\}$$

definition: A cone $K \subseteq \mathbb{R}^n$ is self dual if $K = K^*$.

Examples of self dual cones: $\mathbb{R}^n_+$, L^n, $\mathbb{S}^n_+$.

Conic duality

Consider the primal conic problem:

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_j \succeq_{K_j} 0, \ j = 1, \ldots, m \end{aligned}$$

- Its Lagrangian is given by

$$L(\mathbf{x}, \boldsymbol{\mu}) = \mathbf{c}^\top \mathbf{x} + \boldsymbol{\mu}^\top (\mathbf{b} - \mathbf{A}\mathbf{x}) = \sum_{j=1}^{m+1} (\mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu})^\top \mathbf{x}_j + \mathbf{b}^\top \boldsymbol{\mu},$$

  where we don not assign dual variables to the conic constraints.
- Minimizing over the primal variables

$$\begin{aligned} q(\mu) &= \min_{\mathbf{x}:\mathbf{x}_j \succeq_{K_j} 0, \, j=1,\ldots,m} \sum_{j=1}^{m+1} (\mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu})^\top \mathbf{x}_j + \mathbf{b}^\top \boldsymbol{\mu} \\ &= \sum_{j=1}^{m} \min_{\mathbf{x}_j \in K_j} (\mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu})^\top \mathbf{x}_j + \min_{\mathbf{x}_{m+1}} (\mathbf{c}_{m+1} - \mathbf{A}_{m+1}^\top \boldsymbol{\mu})^\top \mathbf{x}_{m+1} + \mathbf{b}^\top \boldsymbol{\mu} \end{aligned}$$

Conic duality

- We have that

$$\min_{\mathbf{x}_j \in K_j} (\mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu})^\top \mathbf{x}_j = \begin{cases} 0 & \mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu} \in K_j^* \\ -\inf & \text{otherwise} \end{cases}$$

$$\min_{\mathbf{x}_{m+1}} (\mathbf{c}_{m+1} - \mathbf{A}_{m+1}^\top \boldsymbol{\mu})^\top \mathbf{x}_j = \begin{cases} 0 & \mathbf{c}_{m+1} - \mathbf{A}_{m+1}^\top \boldsymbol{\mu} = 0 \\ -\inf & \text{otherwise} \end{cases}$$

Therefore, the dual problem is given by

$$\begin{aligned} \max \quad & q(\mu) \equiv \mathbf{b}^\top \boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{A}_j^\top \boldsymbol{\mu} \preceq_{K_j^*} \mathbf{c}_j, \ j = 1, \ldots, m \\ & \mathbf{A}_{m+1}^\top \boldsymbol{\mu} = \mathbf{c}_{m+1} \end{aligned}$$

Strong duality

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_j \succeq_{K_j} 0, \; j = 1, \ldots, m \end{aligned} \tag{P}$$

$$\begin{aligned} \max \quad & \mathbf{b}^\top \boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{A}_j^\top \boldsymbol{\mu} \preceq_{K_j^*} \mathbf{c}_j, \; j = 1, \ldots, m \\ & \mathbf{A}_{m+1}^\top \boldsymbol{\mu} = \mathbf{c}_{m+1} \end{aligned} \tag{D}$$

Theorem: If both primal and dual satisfy the generalized Slater condition then:

- The primal and dual values are equal and attained.
- $\mathbf{x}$ and $\boldsymbol{\mu}$ are optimal solutions to the primal and dual problem, respectively, if and only if, they are feasible and

$$\mathbf{x}_j^\top (\mathbf{c}_j - \mathbf{A}_j^\top \boldsymbol{\mu}) = 0$$

The proof is a straightforward extension of the proof for regular duality.

SDP Duality

$$\begin{aligned} \min \quad & \mathbf{A}_0 \cdot \mathbf{X} \\ \text{s.t.} \quad & \mathbf{A}_j \cdot \mathbf{X} = b_j,\ j = 1, \ldots, m, \\ & \mathbf{X} \succeq 0. \end{aligned}$$

The dual problem is given by

$$\begin{aligned} \max \quad & \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}_0 - \sum_{j=1}^{m} \mathbf{A}_j y_j \succeq 0 \end{aligned}$$

The constraints are called Linear Matrix Inequalities (LMI)

Using complementary-slackness to solve SDP

For the specific problem

$$\begin{aligned} \min \quad & \mathbf{A}_0 \cdot \mathbf{X} \\ \text{s.t.} \quad & \operatorname{trace}(\mathbf{X}) \equiv \mathbf{I} \cdot \mathbf{X} = 1, \\ & \mathbf{X} \succeq 0. \end{aligned}$$

The dual problem is given by

$$\begin{aligned} \max \quad & y \\ \text{s.t.} \quad & \mathbf{A}_0 - \mathbf{I} y \succeq 0. \end{aligned}$$

- The optimal solution of the dual problem is $y = \lambda_{min}(\mathbf{A}_0)$.
- The primal optimal solution exists (why?) and satisfies

$$\begin{aligned} & (\mathbf{A}_0 - \lambda_{min}(\mathbf{A}_0)\mathbf{I}) \cdot \mathbf{X} = 0 \\ & \mathbf{I} \cdot \mathbf{X} = 1, \\ & \mathbf{X} \succeq 0. \end{aligned}$$

- $\mathbf{A}_0 \cdot \mathbf{X} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top \cdot \mathbf{X} = \boldsymbol{\Lambda} \cdot \mathbf{U}^\top \mathbf{X} \mathbf{U} = \lambda_{min}(\mathbf{A}_0)$.
- Thus,

$$\mathbf{U}^\top \mathbf{X} \mathbf{U} = \mathbf{I}_n \Leftrightarrow \mathbf{X} = \mathbf{U} \mathbf{I}_n \mathbf{U}^\top = \mathbf{u}_n \mathbf{u}_n^\top$$

where $\mathbf{u}_n$ is an eigenvector associated with the minimal eigenvalue.

Aharon Ben-Tal and Arkadi Nemirovski. "Lectures on modern convex optimization: analysis, algorithms, and engineering applications". In: vol. 2. Siam, 2001. Chap. 1.4.