The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381K: Convex Optimization — Fall 2019

LECTURE 20

Aryan Mokhtari Monday, November 11, 2019.

Goal: In this lecture, we talk about using dual problem for solving problem, tricks for writing different versions (possibly simpler) of the dual of a problem, and the dual of SOCP and SDP.

1 Solving the primal problem via the dual problem

In cases that strong duality holds, if we find optimal dual solutions $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, then we know that the minimizer of $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is primal optimal. Using this connection we can first write $\mathbf{x}^*$ as a function of $\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*$ and then solve the dual problem and find $\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*$ and use the expression for $\mathbf{x}^*$ to compute the optimal solution of the primal problem.

Least-norm solution of linear equations: Consider the following problem

$$\min : \mathbf{x}^\top \mathbf{x} \qquad \text{s.t.} : \mathbf{A}\mathbf{x} = \mathbf{b}.$$

We know that for this problem Slater's condition holds (if the problem is feasible) and we have strong duality. The Lagrangian in this case is defined as

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^\top \mathbf{x} + \boldsymbol{\nu}^\top (\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{x}^\top \mathbf{x} + \mathbf{x}^\top \mathbf{A}^\top \boldsymbol{\nu} - \boldsymbol{\nu}^\top \mathbf{b}$$

According to the KKT conditions, we know that optimal $\mathbf{x}^*$ and $\boldsymbol{\nu}^*$ satisfy

$$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\nu}^*) = \mathbf{0} \quad \Leftrightarrow \quad 2\mathbf{x}^* + \mathbf{A}^\top \boldsymbol{\nu}^* = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{x}^* = -\frac{1}{2}\mathbf{A}^\top \boldsymbol{\nu}^*$$

Further note that in this case, the dual problem is

$$\max : -\frac{1}{4}\boldsymbol{\nu}^\top \mathbf{A}\mathbf{A}^\top \boldsymbol{\nu} - \mathbf{b}^\top \boldsymbol{\nu}$$

Hence, if we solve the dual problem and find $\boldsymbol{\nu}^*$, we can use the relation $\mathbf{x}^* = -\frac{1}{2}\mathbf{A}^\top \boldsymbol{\nu}^*$ to find an optimal solution for the primal problem. For instance, when $\mathbf{A}\mathbf{A}^\top$ is invertible, we know that the optimal solution for the dual problem is $\boldsymbol{\nu}^* = -2(\mathbf{A}\mathbf{A}^\top)^{-1}\mathbf{b}$ and as a result $\mathbf{x}^* = \mathbf{A}^\top (\mathbf{A}\mathbf{A}^\top)^{-1}\mathbf{b}$ is an optimal solution for the primal problem.

2 Introducing new variables and equality constraints

Equivalent formulations of a problem can lead to very different duals. Therefore, reformulating the primal problem can be useful when the dual is difficult to derive, or uninteresting.

Common reformulations techniques:
- Introduce new variables and equality constraints (will use this trick for the dual of an SOCP)
- Make explicit constraints implicit or vice-versa

2.1 Example

Consider the following LP with box constraints:

$$\begin{aligned} \min: \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.}: \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1} \end{aligned}$$

Indeed, the dual problem can be written as an LP too and has the following form:

$$\begin{aligned} \max: \quad & -\mathbf{b}^\top \boldsymbol{\nu} - \mathbf{1}^\top \boldsymbol{\lambda}_1 - \mathbf{1}^\top \boldsymbol{\lambda}_2 \\ \text{s.t.}: \quad & \mathbf{c} + \mathbf{A}^\top \boldsymbol{\nu} + \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 = \mathbf{0} \\ & \boldsymbol{\lambda}_1 \geq \mathbf{0}, \quad \boldsymbol{\lambda}_2 \geq \mathbf{0}. \end{aligned}$$

But, we can also write the dual problem in a different way. Basically, we do not dualize the box constraint. To do so we rewrite the primal objective function in a way that is well defined if the box-constrained is satisfied.

$$\begin{aligned} \min: \quad & f_0(\mathbf{x}) = \begin{cases} \mathbf{c}^\top \mathbf{x} & \text{if } -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{s.t.}: \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

Considering this new formulation the Lagrangian can be written as

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\nu}) &= \inf_{-\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}} \left(\mathbf{c}^\top \mathbf{x} + \boldsymbol{\nu}^\top (\mathbf{A}\mathbf{x} - \mathbf{b}) \right) \\ &= -\boldsymbol{\nu}^\top \mathbf{b} + \inf_{-\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}} \left((\mathbf{c} + \mathbf{A}^\top \boldsymbol{\nu})^\top \mathbf{x} \right) \\ &= -\boldsymbol{\nu}^\top \mathbf{b} - \|\mathbf{c} + \mathbf{A}^\top \boldsymbol{\nu}\|_1 \end{aligned}$$

Hence, the dual problem is unconstrained and can be written as

$$\max: \quad -\boldsymbol{\nu}^\top \mathbf{b} - \|\mathbf{c} + \mathbf{A}^\top \boldsymbol{\nu}\|_1$$

3 Dual of SOCP

Recall that an SOCP in general can be written as

$$\begin{aligned} \min: \quad & \mathbf{q}^\top \mathbf{x} \\ \text{s.t.}: \quad & \|\mathbf{A}_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^\top \mathbf{x} + d_i, \qquad i = 1, \ldots, m. \end{aligned}$$

with primal variable $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$.

In this subsection, we study the dual of this problem. To do so, we introduce new variables $\mathbf{y}_i \in \mathbb{R}^{n_i}$ defined as $\mathbf{y}_i := \mathbf{A}_i \mathbf{x} + \mathbf{b}_i$ and $t_i \in \mathbb{R}$ defined as $t_i := \mathbf{c}_i^\top \mathbf{x} + d_i$. Then the original SOCP problem can be written as

$$\begin{aligned} \min: \quad & \mathbf{q}^\top \mathbf{x} \\ \text{s.t.}: \quad & \|\mathbf{y}_i\|_2 \leq t_i, \qquad i = 1, \ldots, m. \\ & \mathbf{y}_i = \mathbf{A}_i \mathbf{x} + \mathbf{b}_i, \qquad i = 1, \ldots, m. \\ & t_i = \mathbf{c}_i^\top \mathbf{x} + d_i, \qquad i = 1, \ldots, m. \end{aligned}$$

Note that the primal variables for the new problem are $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y}_1 \in \mathbb{R}^{n_1}$, $\ldots$, $\mathbf{y}_m \in \mathbb{R}^{n_m}$ and $\mathbf{t} \in \mathbb{R}^m$. The Lagrangian of the new problem can be written as

$$
\begin{aligned}
& L(\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_m, \mathbf{t}, \boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\mu}) \\
& = \mathbf{q}^\top \mathbf{x} + \sum_{i=1}^m \lambda_i (\|\mathbf{y}_i\|_2 - t_i) + \sum_{i=1}^m \boldsymbol{\nu}_i^\top (\mathbf{y}_i - \mathbf{A}_i \mathbf{x} - \mathbf{b}_i) + \sum_{i=1}^m \mu_i (t_i - \mathbf{c}_i^\top \mathbf{x} - d_i) \\
& = \left(\mathbf{q} - \sum_{i=1}^m \mathbf{A}_i^\top \boldsymbol{\nu}_i - \sum_{i=1}^m \mu_i \mathbf{c}_i \right)^\top \mathbf{x} + \sum_{i=1}^m \left(\lambda_i \|\mathbf{y}_i\|_2 + \boldsymbol{\nu}_i^\top \mathbf{y}_i \right) + \sum_{i=1}^m (-\lambda_i + \mu_i)\, t_i - \sum_{i=1}^m \left(\boldsymbol{\nu}_i^\top \mathbf{b}_i + \mu_i d_i \right)
\end{aligned}
$$

Now to find the dual function we need to minimize the Lagrangian with respect to $\mathbf{x}$, $\mathbf{y}_1$, $\ldots$, $\mathbf{y}_m$, t_1, $\ldots$, t_m.

The minimum over $\mathbf{x}$ is bounded below if and only if

$$
\mathbf{q} - \sum_{i=1}^m \mathbf{A}_i^\top \boldsymbol{\nu}_i - \sum_{i=1}^m \mu_i \mathbf{c}_i = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{q} = \sum_{i=1}^m (\mathbf{A}_i^\top \boldsymbol{\nu}_i + \mu_i \mathbf{c}_i)
$$

To minimize over $\mathbf{y}_i$, we note that for any given λ_i and $\boldsymbol{\nu}_i$ we have that

$$
\begin{aligned}
\inf_{\mathbf{y}_i} \left(\lambda_i \|\mathbf{y}_i\|_2 + \boldsymbol{\nu}_i^\top \mathbf{y}_i \right) &= \inf_{r \geq 0, \|\mathbf{q}_i\| = 1} \left(\lambda_i \|r \mathbf{q}_i\|_2 + r \boldsymbol{\nu}_i^\top \mathbf{q}_i \right) \\
&= \inf_{r \geq 0, \|\mathbf{q}_i\| = 1} \left(\lambda_i r + r \boldsymbol{\nu}_i^\top \mathbf{q}_i \right) \\
&= \inf_{r \geq 0, \|\mathbf{q}_i\| = 1} r \left(\lambda_i + \boldsymbol{\nu}_i^\top \mathbf{q}_i \right) \\
&= \inf_{r \geq 0} r \left(\lambda_i + \boldsymbol{\nu}_i^\top \frac{-\boldsymbol{\nu}_i}{\|\boldsymbol{\nu}_i\|} \right) \\
&= \inf_{r \geq 0} r \left(\lambda_i - \|\boldsymbol{\nu}_i\| \right) \\
&= \begin{cases} 0 & \text{if } \|\boldsymbol{\nu}_i\| \leq \lambda_i \\ -\infty & \text{otherwise} \end{cases}
\end{aligned}
$$

The minimum over t_i is bounded below if and only if

$$
\lambda_i = \mu_i
$$

Considering these results we obtain that the dual function $g(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\mu})$ is given by

$$
g(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\mu}) = \begin{cases} -\sum_{i=1}^m \left(\boldsymbol{\nu}_i^\top \mathbf{b}_i + \mu_i d_i \right) & \text{if } \mathbf{q} = \sum_{i=1}^m (\mathbf{A}_i^\top \boldsymbol{\nu}_i + \mu_i \mathbf{c}_i),\ \|\boldsymbol{\nu}_i\| \leq \lambda_i,\ \lambda_i = \mu_i, \\ -\infty & \text{otherwise} \end{cases}
$$

As a result, the dual problem of SCOP can be written as

$$
\begin{aligned}
\max: \quad & -\sum_{i=1}^m \left(\boldsymbol{\nu}_i^\top \mathbf{b}_i + \mu_i d_i \right) \\
\text{s.t.}: \quad & \sum_{i=1}^m (\mathbf{A}_i^\top \boldsymbol{\nu}_i + \mu_i \mathbf{c}_i) = \mathbf{q} \\
& \|\boldsymbol{\nu}_i\| \leq \lambda_i, \qquad i = 1, \ldots, m. \\
& \lambda_i = \mu_i, \qquad i = 1, \ldots, m.
\end{aligned}
$$

which can be simplified as

$$\begin{aligned}
\max: \quad & -\sum_{i=1}^{m}\left(\boldsymbol{\nu}_i^\top \mathbf{b}_i + \lambda_i d_i\right) \\
\text{s.t.}: \quad & \sum_{i=1}^{m}(\mathbf{A}_i^\top \boldsymbol{\nu}_i + \lambda_i \mathbf{c}_i) = \mathbf{q} \\
& \|\boldsymbol{\nu}_i\| \leq \lambda_i, \qquad i = 1, \ldots, m.
\end{aligned}$$

4 Dual of SDP

Recall the general formulation of SDP in inequality form

$$\begin{aligned}
\min: \quad & \mathbf{c}^\top \mathbf{x} \\
\text{s.t.}: \quad & x_1\mathbf{F}_1 + x_2\mathbf{F}_2 + \cdots + x_n\mathbf{F}_n \preceq \mathbf{G},
\end{aligned}$$

where $\mathbf{G}, \mathbf{F}_1, \ldots, \mathbf{F}_n \in \mathbf{S}^k$ and $\mathbf{A} \in \mathbb{R}^{p\times n}$.
In this case, the Lagrange multiplier is matrix $\mathbf{Z} \in \mathbf{S}^k$ and the Lagrangian is

$$\begin{aligned}
L(\mathbf{x}, \mathbf{Z}) &= \mathbf{c}^\top \mathbf{x} + \mathbf{tr}(\mathbf{Z}(x_1\mathbf{F}_1 + x_2\mathbf{F}_2 + \cdots + x_n\mathbf{F}_n - \mathbf{G})) \\
&= (c_1 + \mathbf{tr}(\mathbf{Z}\mathbf{F}_1))x_1 + \cdots + (c_n + \mathbf{tr}(\mathbf{Z}\mathbf{F}_n))x_n - \mathbf{tr}(\mathbf{Z}\mathbf{G})
\end{aligned}$$

Hence, the dual function is

$$g(\mathbf{Z}) = \begin{cases} -\mathbf{tr}(\mathbf{Z}\mathbf{G}) & \text{if } \; c_i + \mathbf{tr}(\mathbf{Z}\mathbf{F}_i) = 0 \text{ for } i = 1, \ldots, n \\ -\infty & \text{otherwise} \end{cases}$$

and therefore the dual problem is

$$\begin{aligned}
\max: \quad & -\mathbf{tr}(\mathbf{Z}\mathbf{G}) \\
\text{s.t.}: \quad & \mathbf{Z} \succeq \mathbf{0}, \qquad \mathbf{tr}(\mathbf{Z}\mathbf{F}_i) + c_i = 0, \qquad i = 1, \ldots, n.
\end{aligned}$$

which is an SDP in standard form. Similar as LP, when we start with a problem in inequality form the dual is in standard form.

Note: Strong duality holds if primal SDP is strictly feasible ($\exists \mathbf{x} : x_1\mathbf{F}_1 + x_2\mathbf{F}_2 + \cdots + x_n\mathbf{F}_n \prec \mathbf{G}$).

If we write the primal problem in standard form as

$$\begin{aligned}
\min: \quad & \mathbf{tr}(\mathbf{C}\mathbf{X}) \\
\text{s.t.}: \quad & \mathbf{tr}(\mathbf{A}_i\mathbf{X}) = b_i, \qquad \text{for } i = 1, \ldots, p, \\
& \mathbf{X} \succeq \mathbf{0},
\end{aligned}$$

then the Lagrange multipliers are $\boldsymbol{\Lambda} \in \S^{n\times n}$ and $\boldsymbol{\nu} \in \mathbb{R}^p$, and the Lagrangian is defined as

$$\begin{aligned}
L(\mathbf{X}, \boldsymbol{\Lambda}, \boldsymbol{\nu}) &= \mathbf{tr}(\mathbf{C}\mathbf{X}) - \mathbf{tr}(\boldsymbol{\Lambda}\mathbf{X}) + \sum_{i=1}^{p} \nu_i(\mathbf{tr}(\mathbf{A}_i\mathbf{X}) - b_i) \\
&= \mathbf{tr}\left(\left(\mathbf{C} - \boldsymbol{\Lambda} + \sum_{i=1}^{p} \nu_i \mathbf{A}_i\right)\mathbf{X}\right) - \sum_{i=1}^{p} \nu_i b_i \\
&= \mathbf{tr}\left(\left(\mathbf{C} - \boldsymbol{\Lambda} + \sum_{i=1}^{p} \nu_i \mathbf{A}_i\right)\mathbf{X}\right) - \boldsymbol{\nu}^\top \mathbf{b}
\end{aligned}$$

Therefore, the dual function is

$$g(\mathbf{Z}) = \begin{cases} -\boldsymbol{\nu}^\top \mathbf{b} & \text{if } \ \mathbf{C} - \boldsymbol{\Lambda} + \sum_{i=1}^{p} \nu_i \mathbf{A}_i = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

and the dual problem is

$$\begin{aligned} \max: \quad & -\boldsymbol{\nu}^\top \mathbf{b} \\ \text{s.t.}: \quad & \mathbf{C} - \boldsymbol{\Lambda} + \sum_{i=1}^{p} \nu_i \mathbf{A}_i = \mathbf{0}, \qquad \boldsymbol{\Lambda} \succeq \mathbf{0} \end{aligned}$$

which can be simplified as

$$\begin{aligned} \max: \quad & -\boldsymbol{\nu}^\top \mathbf{b} \\ \text{s.t.}: \quad & \sum_{i=1}^{p} \nu_i \mathbf{A}_i + \mathbf{C} \succeq \mathbf{0} \end{aligned}$$

which is an SDP in inequality form.