Discussion 11

We might not cover all the problems on the worksheet, as discussion worksheets are not always designed to be finished within an hour. However, this is totally fine since they are deliberately made slightly longer so they can serve as resources you can use to practice, reinforce, and build upon concepts covered in lectures and homework.

1. A Linear Program

Let $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$ and $\mu > 0$. First, consider the following problem:

$$p^\star = \min_{\vec{x}} \|A\vec{x} - \vec{y}\|_1 . \tag{1}$$

For $i \in \{1, \ldots, m\}$, we denote by $\vec{a}_i^\top$ the i-th row of A, so that $A = \begin{bmatrix} \vec{a}_1^\top \\ \vdots \\ \vec{a}_m^\top \end{bmatrix}$.

(a) Express the problem as an LP.

Solution: Using epigraph reformulation and introducing slack variable $\vec{z}$, this problem can be written as

$$\begin{aligned} \min_{\vec{x}, \vec{z}} \quad & \vec{z}^\top \vec{1} \\ \text{s.t.} \quad & z_i \geq \left| \vec{a}_i^\top \vec{x} - y_i \right|, \quad i = 1, \ldots, m. \end{aligned}$$

Equivalently, we can express it as the following LP:

$$\begin{aligned} \min_{\vec{x}, \vec{z}} \quad & \vec{z}^\top \vec{1} \\ \text{s.t.} \quad & z_i \geq \vec{a}_i^\top \vec{x} - y_i, \quad z_i \geq -(\vec{a}_i^\top \vec{x} - y_i), \quad i = 1, \ldots, m. \end{aligned}$$

(b) Show that a dual to the problem can be written as

$$d^\star = \max_{\vec{u}} \; -\vec{u}^\top \vec{y} \; : \; A^\top \vec{u} = 0, \quad \|\vec{u}\|_\infty \leq 1. \tag{2}$$

HINT: Use the fact that, for any vector z:

$$\max_{\vec{u} \,:\, \|\vec{u}\|_1 \leq 1} \vec{u}^\top \vec{z} = \|\vec{z}\|_\infty , \quad \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} \vec{u}^\top \vec{z} = \|\vec{z}\|_1 . \tag{3}$$

Additionally, you may use the following fact: If the primal problem is expressed as $p^\star = \min_{\vec{x}} \max_{\vec{y}} L(\vec{x}, \vec{y})$ (which is $p^\star = \min_{\vec{x}} f_0(x)$ for $f_0(x) = \max_{\vec{y}} L(\vec{x}, \vec{y})$), then the dual problem can be obtained by swapping the max and min: $d^\star = \max_{\vec{y}} \min_{\vec{x}} L(\vec{x}, \vec{y})$.

Solution: Based on the hint, we substitute the objective function with the variational characterization:

$$\|A\vec{x} - \vec{y}\|_1 = \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} \vec{u}^\top (A\vec{x} - \vec{y}), \tag{4}$$

The optimization problem thus becomes:

$$p^\star = \min_{\vec{x}} \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} \mathcal{L}(\vec{x}, \vec{u}), \tag{5}$$

where

$$\mathcal{L}(\vec{x}, \vec{u}) = \vec{u}^\top (A\vec{x} - \vec{y}). \tag{6}$$

Note that the primal is already expressed in min-max form. Recall exchanging min and max directly leads to the dual problem:

$$\begin{aligned} p^\star \geq d^\star &= \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} \; \min_{\vec{x}} \; \mathcal{L}(\vec{x}, \vec{u}) & (7) \\ &= \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} g(\vec{u}), & (8) \end{aligned}$$

where g is the dual function:

$$g(\vec{u}) = \min_{\vec{x}} \; \mathcal{L}(\vec{x}, \vec{u}) = \begin{cases} -\vec{u}^\top \vec{y} & \text{if } A^\top \vec{u} = 0, \\ -\infty & \text{otherwise.} \end{cases} \tag{9}$$

The dual problem can therefore be written as

$$\begin{aligned} d^\star = \max_{\vec{u}} & -\vec{u}^\top \vec{y} \\ \text{s.t.} \quad & A^\top \vec{u} = 0, \;\; \|\vec{u}\|_\infty \leq 1. \end{aligned}$$

Now, consider the following more complicated problem involving both the ℓ_1 and ℓ_∞ norms:

$$p^\star = \min_{\vec{x}} \; \|A\vec{x} - \vec{y}\|_1 + \mu \|\vec{x}\|_\infty . \tag{10}$$

(c) Express the problem as an LP.

Solution: Similar to part (a), we can introduce slack variable $\vec{z}$ and write

$$\min_{\vec{x}, \vec{z}, t} \; \vec{z}^\top \vec{1} + \mu t \;:\; t \geq \|\vec{x}\|_\infty, \;\; z_i \geq |\vec{a}_i^\top \vec{x} - y_i|, \;\; i = 1, \ldots, m. \tag{11}$$

which is an LP:

$$\begin{aligned} \min_{\vec{x}, \vec{z}, t} \; \vec{z}^\top \vec{1} + \mu t \;:\; \quad & t \geq x_j, \;\; t \geq -x_j, j = 1, \ldots, n \\ & z_i \geq \vec{a}_i^\top \vec{x} - y_i, \;\; z_i \geq -(\vec{a}_i^\top \vec{x} - y_i), \;\; i = 1, \ldots, m. \end{aligned} \tag{12}$$

(d) Show that a dual to the problem can be written as

$$\begin{aligned} d^\star = \max_{\vec{u}} & \; -\vec{u}^\top \vec{y} \\ \text{s.t.} \quad & \|\vec{u}\|_\infty \leq 1, \;\; \left\|A^\top \vec{u}\right\|_1 \leq \mu. \end{aligned}$$

HINT: Use the fact that, for any vector z:

$$\max_{\vec{u} \,:\, \|\vec{u}\|_1 \leq 1} \; \vec{u}^\top \vec{z} = \|\vec{z}\|_\infty , \quad \max_{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1} \; \vec{u}^\top \vec{z} = \|\vec{z}\|_1 . \tag{13}$$

Additionally, you may use the following fact: If the primal problem is expressed as $p^\star = \min_{\vec{x}} \max_{\vec{y}} L(\vec{x}, \vec{y})$ (which is $p^\star = \min_{\vec{x}} f_0(x)$ for $f_0(x) = \max_{\vec{y}} L(\vec{x}, \vec{y})$), then the dual problem can be obtained by swapping the max and min: $d^\star = \max_{\vec{y}} \min_{\vec{x}} L(\vec{x}, \vec{y})$.

Solution: Similar to part (b), we can express the objective function as

$$\|A\vec{x} - \vec{y}\|_1 + \mu \|\vec{x}\|_\infty = \max_{\substack{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1 \\ \vec{v} \,:\, \|\vec{v}\|_1 \leq \mu}} \vec{u}^\top (A\vec{x} - \vec{y}) + \vec{v}^\top \vec{x}, \tag{14}$$

The optimization problem thus becomes:

$$p^\star = \min_{\vec{x}} \max_{\substack{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1 \\ \vec{v} \,:\, \|\vec{v}\|_1 \leq \mu}} \mathcal{L}(\vec{x}, \vec{u}, \vec{v}), \tag{15}$$

where

$$\mathcal{L}(\vec{x}, \vec{u}, \vec{v}) = u^\top (A\vec{x} - \vec{y}) + \vec{v}^\top \vec{x}. \tag{16}$$

Again, recall exchanging min and max leads to the dual:

$$p^\star \geq d^\star = \max_{\substack{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1 \\ \vec{v} \,:\, \|\vec{v}\|_1 \leq \mu}} \min_{\vec{x}} \mathcal{L}(\vec{x}, \vec{u}, \vec{v}) \tag{17}$$

$$= \max_{\substack{\vec{u} \,:\, \|\vec{u}\|_\infty \leq 1 \\ \vec{v} \,:\, \|\vec{v}\|_1 \leq \mu}} g(\vec{u}, \vec{v}), \tag{18}$$

where g is the dual function:

$$g(\vec{u}, \vec{v}) = \min_{\vec{x}} \mathcal{L}(\vec{x}, \vec{u}, \vec{v}) = \begin{cases} -\vec{u}^\top \vec{y} & \text{if } A^\top \vec{u} + \vec{v} = \vec{0}, \\ -\infty & \text{otherwise.} \end{cases} \tag{19}$$

The dual problem writes

$$\begin{aligned} d^\star = \max_{\vec{u}\,,\,\vec{v}} \; & -\vec{u}^\top \vec{y} \\ \text{s.t.} \quad & A^\top \vec{u} + \vec{v} = \vec{0}, \;\; \|\vec{u}\|_\infty \leq 1, \;\; \|\vec{v}\|_1 \leq \mu. \end{aligned}$$

Note we can eliminate $\vec{v}$:

$$\begin{aligned} d^\star = \max_{\vec{u}} \; & -\vec{u}^\top \vec{y} \\ \text{s.t.} \quad & \|\vec{u}\|_\infty \leq 1, \;\; \left\|A^\top \vec{u}\right\|_1 \leq \mu. \end{aligned}$$

2. A review of standard problem formulations

In this question, we review conceptually the standard forms of various problems and the assertions we can (and cannot!) make about each.

(a) ***Linear programming (LP).***

i. Write the most general form of a linear program (LP) and list its defining attributes.

Solution: A general LP can be written as

$$p^* = \min_{\vec{x} \in \mathbb{R}^n} \quad \vec{c}^\top \vec{x} + d \tag{20}$$

$$\text{s.t.} \quad A_{\text{eq}} \vec{x} = \vec{b}_{\text{eq}} \tag{21}$$

$$A\vec{x} \leq \vec{b}, \tag{22}$$

or equivalently,

$$p^* = \min_{\vec{x}\in\mathbb{R}^n} \quad \vec{c}^\top \vec{x} + d \tag{23}$$
$$\text{s.t.} \quad A_{\text{eq}}\vec{x} = \vec{b}_{\text{eq}} \tag{24}$$
$$\vec{x} \geq \vec{0}. \tag{25}$$

The first LP formulation is known as the *inequality form*; the second is known as the *conic form*. A full treatment of the equivalence of these forms and how to convert between them can be found in section 9.3 of Calafiore & El Ghaoui.

ii. Under what conditions is an LP convex?

Solution: An LP is **always convex**, as the objective function and all constraints are convex (affine) and all equality constraints are affine.

(b) ***Quadratic programming (QP).***

i. Write the most general form of a quadratic program (QP) and list its defining attributes.

Solution: A general QP can be written as

$$p^* = \min_{\vec{x}\in\mathbb{R}^n} \quad \frac{1}{2}\vec{x}^\top H\vec{x} + \vec{c}^\top \vec{x} + d \tag{26}$$
$$\text{s.t.} \quad A_{\text{eq}}\vec{x} = \vec{b}_{\text{eq}} \tag{27}$$
$$A\vec{x} \leq \vec{b}. \tag{28}$$

ii. Under what conditions is a QP convex?

Solution: A QP is convex if and only if $H \succeq 0$ (i.e., PSD).

(c) ***Quadratically-constrained quadratic programming (QCQP).***

i. Write the most general form of a quadratically-constrained quadratic program (QCQP) and list its defining attributes.

Solution: A general QCQP can be written as

$$p^* = \min_{\vec{x}\in\mathbb{R}^n} \quad \vec{x}^\top H_0\vec{x} + 2\vec{c}^\top \vec{x} + d \tag{29}$$
$$\text{s.t.} \quad \vec{x}^\top H_i\vec{x} + 2\vec{c}_i^\top \vec{x} + d_i \leq 0, \quad i = 1, \ldots, m \tag{30}$$
$$\vec{x}^\top H_j\vec{x} + 2\vec{c}_j^\top \vec{x} + d_j = 0, \quad j = 1, \ldots, q. \tag{31}$$

ii. Under what conditions is a QCQP convex?

Solution: A QCQP is convex if and only if all matrices H_0 and $H_i,\ i = 1, \ldots, m$ are PSD, and $H_j = 0$ for all $j = 1, \ldots, q$ (i.e., when the objective and all inequality constraints are convex quadratic, and all the equality constraints are actually affine.

(d) ***Second-order cone programming (SOCP).***

i. Write the most general form of a second-order cone program (SOCP) and list its defining attributes.

Solution: A general SOCP can be written as

$$p^* = \min_{\vec{x}\in\mathbb{R}^n} \quad \vec{c}^\top \vec{x} \tag{32}$$

$$\text{s.t.} \quad \left\| A_i\vec{x} + \vec{b}_i \right\|_2 \leq \vec{c}_i^\top \vec{x} + d_i, \quad i = 1, \ldots, m, \tag{33}$$

or equivalently,

$$p^* = \min_{\vec{x} \in \mathbb{R}^n} \quad \vec{c}^\top \vec{x} \tag{34}$$

$$\text{s.t.} \quad (A_i\vec{x} + \vec{b}_i, \vec{c}_i^\top \vec{x} + d_i) \in \mathcal{K}_{m_i} \quad i = 1, \ldots, m, \tag{35}$$

where second-order cone (SOC) $\mathcal{K}_n \doteq \{(\vec{x}, t), \vec{x} \in \mathbb{R}^n, t \in \mathbb{R} \mid \|\vec{x}\|_2 \leq t\}$. The first SOCP formulation is known as the *standard inequality form*; the second is known as the *conic standard form*.

ii. Under what conditions is an SOCP convex?

Solution: An SOCP is **always convex**, as the objective function is convex (linear) and the constraint stipulates that points lie within a convex set. A thorough discussion of the convexity of SOC constraints can be found in Calafiore & El Ghaoui chapter 10.

(e) ***Relationships.*** Recall that

$$\text{LP} \subset \text{QP}_{\text{convex}} \subset \text{QCQP}_{\text{convex}} \subset \text{SOCP} \subset \{\text{all convex programs}\}, \tag{36}$$

where LP denotes the set of all linear programs, $\text{QP}_{\text{convex}}$ denotes the set of all convex quadratic programs, etc. Which of these problems can be solved most efficiently? Why are these categorizations useful?

Solution: In general, problems to the left side of the subset sequence above can be solved more efficiently than those on the right, for problems of comparable size: LPs are arguably easiest (optima are always achieved at a critical point, so we just need to check those analytically, e.g. using the simplex algorithm), while general convex problems must use numerical algorithms like gradient descent and Newton's method, whose efficiency depends on the particular geometric characteristics of the problem. Intermediate forms (convex QPs/QCQPs, SOCPs) fall somewhere in between in terms of difficulty: there's lots of research on how to solve them efficiently (one popular set of approaches is called "interior point methods"), and there are a number of off-the-shelf solvers available (e.g., CVX).

Though we don't get a chance to explore it much in this class, knowing these classes of problems is useful when encountering optimization problems in the wild — if you can write your problem in one of the forms above, you know what kinds of solutions and convergence guarantees you can expect, and often employ existing software to help you solve it. For further discussion of these problem classes — as well as an even more general class known as semidefinite programming (SDP), of which SOCPs are a subset — we encourage you to take EECS 227B and 227C.