Convex Optimization: Fall 2021, Homework #2 (Due: 12/17)

1. (20 points) Show that the following the epigraph problem form is equivalent to the standard problem by using KKT conditions.

   (Standard problem) $\underset{x}{\text{minimize}}\ \ f_0(x)$

   (Epigraph problem) $\underset{x,t}{\text{minimize}}\ \ t$ subject to $f_0(x) \le t$

(sol)

The Lagrangian of epigraph problem is $L(x,t) = t + \lambda(f_0(x) - t)$. By KKT condition,

$$\frac{\partial L}{\partial t} = 1 - \lambda = 0$$
$$\lambda(f_0(x) - t) = 0$$

Then, $\lambda = 1$ and $f_0(x) = t$. Hence, the optimal solutions of these two problems are the same, i.e., they are equivalent problems.

2. (40 points) We have the following norm approximation problem: $\underset{x}{\text{minimize}}\ \ \|Ax - b\|$

   (a) For $\|Ax - b\| = \|Ax - b\|_\infty$, show that the following LP is an equivalent problem.

   $$\underset{x,t}{\text{minimize}}\ \ t \qquad \text{subject to } -t\mathbf{1} \le Ax - b \le t\mathbf{1}$$

   where $x \in R^n$, $t \in R, A \in R^{m\times n}$, and $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

(sol)

From the solution of Problem 1, we know that $\|Ax - b\|_\infty = \max_i\{|a_i^T x - b_i|\} \le t$ where a_i^T is the i-th row of A. Since $\max_i\{|a_i^T x - b_i|\} \le t$ is equivalent to $-t\mathbf{1} \le Ax - b \le t\mathbf{1}$, the LP problem is equivalent to the ℓ_∞-norm approximation problem.

   (b) For $\|Ax - b\| = \|Ax - b\|_1$, show that the following LP is an equivalent problem.

   $$\underset{x,y}{\text{minimize}}\ \ \mathbf{1}^T y \qquad \text{subject to } -y \le Ax - b \le y$$

   where $x \in R^n$, $y \in R^m$, $A \in R^{m\times n}$, and $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

(sol)

The Lagrangian is given by

$$L(x,y,\lambda,\nu) = \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} \lambda_i\left(a_i^T x - b_i + y_i\right) + \sum_{i=1}^{m} \nu_i\left(a_i^T x - b_i - y_i\right)$$

From KKT conditions,

$$\frac{\partial L}{\partial y_i} = 1 - \lambda_i - \nu_i = 0,$$
$$\lambda_i\left(a_i^T x - b_i + y_i\right) = 0,$$
$$\nu_i\left(a_i^T x - b_i - y_i\right) = 0.$$

First, we will show that $a_i^T x - b_i = -y_i$ or $a_i^T x - b_i = y_i$ (i.e., $|a_i^T x - b_i| = y_i$). If $|a_i^T x - b_i| \neq y_i$, then $\lambda_i = \nu_i = 0$ because of complementary slackness. However, $\lambda_i = \nu_i = 0$ cannot satisfy $1 - \lambda_i - \nu_i = 0$. Hence, $|a_i^T x - b_i| = y_i$ for any i.

The objective function $\mathbf{1}^T y = \sum_{i=1}^m |a_i^T x - b_i| = \|Ax - b\|_1$, the LP problem is equivalent to the ℓ_1-norm approximation problem.

3. (20 points) Suppose that $f(x)$ is *concave* of $x = (x_1, \dots, x_k)$ when x is a probability vector (probability mass function). Consider the following optimization problem.

$$\underset{x}{\text{maximize}} \;\; f(x)$$

$$\text{subject to } \sum_{i=1}^k x_i = 1,\; x_i \geq 0 \text{ for } 1 \leq i \leq k$$

Then, derive the following condition for the optimal x^* by using KKT conditions:

$$\frac{\partial f(x)}{\partial x_i} = \nu \text{ for } i \text{ such that } x_i > 0;$$

$$\frac{\partial f(x)}{\partial x_i} \leq \nu \text{ for } i \text{ such that } x_i = 0;$$

(sol)

The Lagrangian is $L(x, \lambda, \nu) = -f(x) - \sum_{i=1}^k \lambda_i x_i + \nu(\sum_{i=1}^k x_i - 1)$. KKT conditions are as follows:

$$\frac{\partial L}{\partial x_i} = -\frac{\partial f}{\partial x_i} - \lambda_i + \nu = 0$$
$$\lambda_i \geq 0,\;\; \lambda_i x_i = 0.$$

If $x_i > 0$, then $\lambda_i = 0$, which results in $\frac{\partial f}{\partial x_i} = \nu$. If $x_i = 0$, then $\lambda_i \geq 0$. Hence, $\frac{\partial f(x)}{\partial x_i} \leq \nu$.

4. (20 points) For an underdetermined linear equation: $Ax = b$ where $A \in R^{p\times n}$ where $p < n$ and $\text{rank}(A) = p$. In order to transform an equality-constrained optimization problem into an unconstrained optimization problem, we parametrize the affine feasible set: $\{x | Ax = b\} = \{Fz + \hat{x} | z \in R^{n-p}\}$.

Suppose that the first p columns of A are independent, i.e., $A = [A_1\;\; A_2]$ where $A_1 \in R^{p\times p}$ is nonsingular. Then, show that $F = \begin{bmatrix} -A_1^{-1}A_2 \\ I \end{bmatrix}$ and $\hat{x} = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix}$.

(sol)

$Ax = [A_1\;\; A_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A_1x_1 + A_2x_2 = b$. Then, $x_1 = A_1^{-1}b - A_1^{-1}A_2x_2$ since A_1 is nonsingular. For a particular solution $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$, we can set $\hat{x}_2 = 0$. Then $\hat{x}_1 = A_1^{-1}b$, i.e., $\hat{x} = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix}$.

$Az = [A_1\;\; A_2]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = A_1z_1 + A_2z_2 = 0$. Then, $z_1 = -A_1^{-1}A_2z_2$. Hence, $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -A_1^{-1}A_2 \\ I \end{bmatrix} z_2$ where $z_2 \in R^{n-p}$.

Then, $x = Fz + \hat{x} = \begin{bmatrix} -A_1^{-1}A_2 \\ I \end{bmatrix} z + \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix}$, where $z = z_2$.

5. (Optional) Boyd 10.1 (a) (p. 557) [Nonsingularity of the KKT matrix]

(sol)

The second and third are clearly equivalent. To see this, if $Ax = 0$, $x \neq 0$, then x must have the form $x = Fz$, where $z \neq 0$. Then we have $x^T P x = z^T F^T P F z$.

Similarly, the first and second are equivalent. To see this, if $x \in \mathcal{N}(A) \cap \mathcal{N}(P)$, $x \neq 0$, then $Ax = 0$, $x \neq 0$, but $x^T P x = 0$, contradicting the second statement. Conversely, suppose the second statement fails to hold, *i.e.*, there is an x with $Ax = 0$, $x \neq 0$, but $x^T P x = 0$. Since $P \succeq 0$, we conclude $Px = 0$, *i.e.*, $x \in \mathcal{N}(P)$, which contradicts the first statement.

Finally, the second and fourth statements are equivalent. If the second holds then the last statement holds with $Q = I$. If the last statement holds for some $Q \succeq 0$ then it holds for all $Q \succ 0$, and therefore the second statement holds.

Now let's show that the four statements are equivalent to nonsingularity of the KKT matrix. First suppose that x satisfies $Ax = 0$, $Px = 0$, and $x \neq 0$. Then

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 0,$$

which shows that the KKT matrix is singular.

Now suppose the KKT matrix is singular, *i.e.*, there are x, z, not both zero, such that

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = 0.$$

This means that $Px + A^T z = 0$ and $Ax = 0$, so multiplying the first equation on the left by x^T, we find $x^T P x + x^T A^T z = 0$. Using $Ax = 0$, this reduces to $x^T P x = 0$, so we have $Px = 0$ (using $P \succeq 0$). This contradicts (a), unless $x = 0$. In this case, we must have $z \neq 0$. But then $A^T z = 0$ contradicts $\operatorname{rank} A = p$.