Convex Optimization – Final Exam, Thursday June 28, 2018.

Answer prepared by Tzu-Yu Jeng.

Textbook: Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge U. Press, 2004.

Exam policy: Open book. You can bring any books, handouts, and any kinds of paper-based notes with you, but use of any electronic devices (including cellphones, laptops, tablets, etc.) is strictly prohibited.

Note: the total score of all problems is 112 points.

1. (45%) For each of the following optimization problems, find (i) the lagrangian $L(x, \lambda, \nu)$, (ii) dual function $g(\lambda, \nu)$, and (iii) the dual problem.

   (a) (5% + 5% + 5%)

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$.

$$\begin{aligned} & L(x, \lambda, \nu) \\ =& \|x\|_2^2 + \langle \lambda,\ Ax - b \rangle \end{aligned}$$

Note that ν is absent. By straightforward arrangement,

$$\begin{aligned} & L(x, \lambda, \nu) \\ =& \left\|x - A^T \lambda\right\|_2^2 - \frac{1}{4}\|A^T \lambda\|_2^2 - \langle \lambda,\ b \rangle \end{aligned}$$

Thus

$$\begin{aligned} & g(\lambda, \nu) \\ =& \inf_x L(x, \lambda, \nu) \\ =& -\frac{1}{4}\|A^T \lambda\|_2^2 - \langle \lambda,\ b \rangle \end{aligned}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -\frac{1}{4}\|A^T \lambda\|_2^2 - \langle \lambda,\ b \rangle \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

   (b) (5% + 5% + 5%)

$$\begin{array}{ll} \text{minimize} & x^T P x \\ \text{subject to} & x^T x \leq 1 \\ & Ax = b \end{array}$$

where $P \in \mathbf{S}^n_{++}$.

$$
\begin{aligned}
&L(x, \lambda, \nu) \\
=&\langle x,\ Px\rangle + \lambda \cdot (\|x\|_2^2 - 1) + \langle \nu,\ Ax - b\rangle
\end{aligned}
$$

Here the λ is a scalar, and the inner product reduces to a multiplication.

$$
\begin{aligned}
&g(\lambda, \nu) \\
=&\inf_x L(x, \lambda, \nu) \\
=&\inf_x \left[\langle x,\ (P + \lambda I)x\rangle + \langle \nu,\ Ax\rangle - \langle \nu,\ b\rangle - \lambda\right] \\
=&\inf_x \left[\langle x,\ (P + \lambda I)x\rangle + \left\langle A^T\nu,\ x\right\rangle - (\langle \nu,\ b\rangle + \lambda)\right]
\end{aligned}
$$

It is well known that the infimum of the quadratic form $\langle x,\ A'x\rangle + \langle b',\ x\rangle$ is $\|\sqrt{A}^{-1}b'\|_2^2/4$, for self-adjoint (symmetric) positive definite linear operator A (which is assumed here). As a reminder, diagonalize A' (why is that possible?) to be UDU^T where D has all-positive entries as assumed, then $\sqrt{A'} = U\sqrt{D}U^T$ and the generalization of "completion of square" works.

$$
g(\lambda, \nu) = -\frac{1}{4}\|\sqrt{P + \lambda I}^{-1}(A^T\nu)\|_2^2 - (\langle \nu,\ b\rangle + \lambda)
$$

The dual problem is

$$
\begin{array}{ll}
\text{maximize} & -\frac{1}{4}\|\sqrt{P + \lambda I}^{-1}(A^T\nu)\|_2^2 - (\langle \nu,\ b\rangle + \lambda) \\
\text{subject to} & \lambda \succeq 0
\end{array}
$$

(c) (5% + 5% + 5%)

$$
\begin{array}{ll}
\text{minimize} & 2x_1 + 3x_2 + 4x_3 \\
\text{subject to} & x_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + x_3 \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \preceq_{\mathbf{S}_+^2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\end{array}
$$

Hint: the Lagrange multiplier for this problem is in the form of a symmetric matrix. You can use the notation Z, as in, e.g., $L(x, Z)$ and $g(Z)$, etc.

Relevant material is Lecture 12, p.8-9. Let

$$
\begin{aligned}
c &= \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \\
F_1 &= \begin{bmatrix} 1, & 0 \\ 0, & 1 \end{bmatrix} \\
F_2 &= \begin{bmatrix} 1, & 1 \\ 1, & 1 \end{bmatrix} \\
F_3 &= \begin{bmatrix} 1, & -1 \\ -1, & 2 \end{bmatrix} \\
G &= \begin{bmatrix} 0, & -1 \\ -1, & 0 \end{bmatrix}
\end{aligned}
$$

Then (if you have defined them clearly, you are allowed not to expand them)

$$
L(x, Z) = \langle c,\ x\rangle + \operatorname{tr}((x_1F_1 + x_2F_2 + x_3F_3 + G)Z)
$$

$$
g(Z) = \begin{cases} \operatorname{tr}(GZ), & \operatorname{tr}(F_iZ) + c_i = 0,\ i = 1, 2, 3 \\ -\infty, & \text{otherwise} \end{cases}
$$

And finally the dual problem is

$$\begin{array}{ll} \text{maximize} & \operatorname{tr}(GZ) \\ \text{subject to} & \begin{cases} \operatorname{tr}(F_i Z) + c_i = 0, \ i = 1, 2, 3 \\ Z \succeq O \end{cases} \end{array}$$

2. (40%) Consider the equality constrained problem

$$\begin{array}{ll} \text{minimize} & f_0(x) = c^T x + \sum_{i=1}^{n} x_i^3 \\ \text{subject to} & Ax = b \end{array}$$

where $x \in \textbf{dom}\ f_0 = \mathbf{R}_+^n$, $A \in \mathbf{R}^{p \times n}$, rank $A = p$, $b \in \mathbf{R}^p$, and $c \in \mathbf{R}^n$.

(a) (10%) Derive the Lagrange dual function $g(\nu)$ of the problem (2a). Find also **dom** g.

$$L(x, \lambda, \nu) = \langle c,\ x \rangle + \sum_{i=1}^{n} x_i^3 + \langle \nu,\ Ax - b \rangle$$

This function is a cubic function of each of x_i, and always attains $-\infty$ if we just set $x = (\infty, \ldots, \infty)$, and thus $g(\lambda, \nu) = \infty$.

(b) (5%) Formulate the dual problem of the problem (2a).
We have the trivial dual problem

$$\begin{array}{ll} \text{maximize} & -\infty \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

The dual problem is never dual feasible.

(c) (10%) Derive $\nabla f_0(x)$ and $\nabla^2 f_0(x)$.
I think this should be straightforward and only give the result.

$$\nabla f_0(x) = c + 3\langle x_1^2, \ldots, x_n^2 \rangle$$
$$\nabla^2 f_0(x) = 6 \operatorname{diag}(x)$$

(d) (5%) Find the KKT conditions for the problem (2a).
(1) Primal feasibility for inequality constraint: No.
(2) Primal feasibility for inequality constraint: $Ax - b = 0$.
(3) Dual feasibility: No (since g is trivial).
(4) Complementary slackeness: No (since there is no inequality constraint).
(5) Vanishing of gradient of Lagrangian:

$$\nabla L = c + \langle 3x_1^2, \ldots, 3x_n^2 \rangle + A^T \nu = 0$$

Or

$$3\langle x_1^2, \ldots, x_n^2 \rangle + c + A^T \nu = 0.$$

(e) (10%) Given a feasible point x (i.e., $Ax = b$), derive the Newton's step Δx_{nt} by writing down a linear system in which $(\Delta x_{nt}, \nu^+)$ is the variable (ν^+ is the updated version of dual variable ν):

$$\begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ \nu^+ \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Find M_1, M_2, M_3, M_4, v_1, and v_2 explicitly.

Quote the result from textbook p.526, eqn. (10.19).

$$\begin{bmatrix} 6 \text{ diag}(x), & A^T \\ A, & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ \nu^+ \end{bmatrix} = \begin{bmatrix} -c - 3\langle x_1^2, \ldots, x_n^2 \rangle \\ 0 \end{bmatrix}$$

This result need not be memorized. The first equation comes from the second KKT equation. Notice that $(\nabla^2 f_0(x))\Delta x_{nt}$ the variation of $\nabla f_0(x)$, and $A^T\nu^+$ is simply plugged in the new ν value. And likewise the second equation comes from the first KKT equation $Ax - b = 0$ by varying Δx_{nt} a little.

3. (15%) Consider the simple problem

$$\begin{array}{ll} \text{minimize} & x^2 - 1 \\ \text{subject to} & 1 \le x \le 3 \end{array}$$

whose feasible set is $[1, 3]$. Suppose you are going to apply the barrier method to this problem (whose objective function is denoted f_0 and constraint functions f_1 and f_2).

(a) (6%) Derive $tf_0 + \phi$ as a function of x, where t is any given positive number and ϕ denotes the logarithmic barrier for the original problem.

$f_0 = x^2 - 1$, and for convenience set $f_1 = 1 - x$, $f_2 = x - 3$. Then

$$\begin{aligned} tf_0 + \phi =& tf_0 - \log(-f_1) - \log(-f_2) \\ =& t \cdot (x^2 - 1) - \log(x - 1) - \log(3 - x) \end{aligned}$$

(b) (7%) Find the optimal point for the "approximated" problem

$$\underset{x}{\text{minimize}} \qquad tf_0 + \phi$$

for any given $t > 0$. In other words, find the central point $x^*(t)$ for any given $t > 0$.

Relevant material: textbook p.597.

Since $tf_0 + \phi$ is convex in $x \in (1, 3)$,

$$\frac{\partial}{\partial x}(tf_0 + \phi) = 0$$

Or

$$2xt - \frac{1}{x-1} - \frac{1}{x-3} = 0$$

When this condition is met, $x \leftarrow x^*(t)$ is plugged in, after some direct arrangement ($x \neq 1, 3$ where the function is singular and thus arrangement is valid),

$$tx^*(t)^3 - 4tx^*(t)^2 + (3t - 1)x^*(t) + 2 = 0$$

(c) (2%) What is $\lim_{t\to\infty} x^*(t)$?

Cast as

$$x^*(t)^3 - 4x^*(t)^2 + (3 - 1/t)x^*(t) + 2/t = 0$$

When $t \to \infty$,

$$x^*(t)^3 - 4x^*(t)^2 + 3x^*(t) = 0$$

Or $x = 0, 1, 3$. Of course, $x = 0$ is not considered, $x = 1$ gives the minimum, and $x = 3$ the maximum, agreeing the (originaly) obvious result. Unless we calculate the second derivative w.r.t. x, we have to plug in the values to know which is minimum.

4. (12%) For the following pairs of proper cones $K \subseteq \mathbf{R}^q$ and functions $\psi : \mathbf{R}^q \to \mathbf{R}$, determine whether ψ is a **generalized logarithm** for K. Briefly explain why.

(a) (4%) $K = \mathbf{R}^3_+$, $\psi(x) = \log x_1 + \log x_2 + \log x_3$.

(b) (4%) $K = \mathbf{R}^3_+$, $\psi(x) = \log x_1 + 2\log x_2 + 3\log x_3$.

(c) (4%) $K = \mathbf{R}^2_+$, $\psi(x) = \log x_1 - \log x_2$.

(a) Yes (b) Yes (c) No

We quote Lecture 15, p.62: "We say that $\mathbf{R}^q \mapsto \mathbf{R}$ is a generalized logarithm for K if (i) ψ is concave, (ii) closed, (iii) twice continuously differentiable, (iv) dom $= \mathrm{int}(K)$, (v) $\nabla^2\psi(y) \prec \mathbf{0}$ for $y \in \mathrm{int}(K)$, (vi) that there is a constant $\theta > 0$ such that for all $y \succ_K 0$, and all $s > 0$, we have $\psi(sy) = \psi(y) + \theta \log s$".

(i), (ii), (iii), and (iv) are entirely obvious. For (ii), see textbook p.457 for the definition of a closed function: a function is closed if it always has closed sublevel sets. For (iv), note that the definition of K ensures $x_1, x_2, x_3 > 0$ (and no more so) in each case, and the generalized inequality here reduces to an usual inequality.

The explanation for (b) is extremely similar to that of (a) in just above.

For (c), the (2,2)-entry of Hessian matrix fails to be negative, as required by (v):

$$\begin{aligned}\frac{\partial^2\psi}{\partial x_2^2} &= x_2^2 \\ &> 0\end{aligned}$$

always.