DATA620008 Optimization
Solution of Midterm

1. Please answer true of false. No explanation is needed. A correct answer is worth 2 points, no answer 0 points, a wrong answer -1.

   (a) The feasible set of a linear program (LP) in standard form is always bounded.

   (b) Consider an LP. If the dual problem is infeasible, then so is the primal problem.

   (c) The problem $\max \sum_{j=1}^n c_j|x_j|$ subject to $\sum_{j=1}^n a_j|x_j| \leq b$, with $c_j, a_j \geq 0$, can be modeled as a linear optimization problem.

   (d) Given a local optimum $\bar{x}$ for a nonlinear optimization problem, it always satisfies the KKT conditions when the gradients of the active constraints and the gradients of the equality constraints at the point $\bar{x}$ are linearly independent.

   (e) For a quadratic function $f(x) = x^T Ax + b^T x + c$ with $A \succ 0$, the convergence rate of Newton's method depends on the condition number of the matrix A.

   Solution.

   (a) False. The feasible set can be $\mathbb{R}_+^n$.

   (b) False. e.g. min x, s.t. $x \leq 0$. It's dual problem is infeasible, but the primal problem is feasible (although doesn't have optimal solution).

   (c) True. Define $u_j = \frac{x_j + |x_j|}{2}$, $v_j = \frac{|x_j| - x_j}{2}$, then the problem can be represented as a linear optimization problem:

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j(u_j + v_j) \\ \text{s.t.} \quad & \sum_{j=1}^n a_j(u_j + v_j) \leq b, \\ & u \geq 0, \quad v \geq 0. \end{aligned}$$

   (d) True. See Theorem 11.6 in
   Beck, Amir. *Introduction to nonlinear optimization: theory, algorithms, and applications with MATLAB.*

   (e) False. When the objective function is quadratic function, then Newton's method only need one step to converge.

2. A function f is called log-convex if $f(x) > 0$, for all $x \in \text{dom}(f)$ and $\log(f(x))$ is convex.

   (a) If f is log-convex, then f is also convex.

   (b) f is log-convex if and only if $\forall \lambda \in [0, 1], \forall x, y \in \text{dom}(f)$, we have $f(\lambda x + (1-\lambda)y) \leq f(x)^\lambda f(y)^{1-\lambda}$.

Solution.

(a) Let $g(x) = \log(f(x)), h(x) = e^x$, then $f(x) = h(g(x))$. Since $g(x)$ is convex, $h(x)$ is convex and nondecreasing, $f(x)$ is also convex.

(b) We have

$$\begin{aligned}
&f \text{ is log convex} \\
\Leftrightarrow\ &\log(f(x)) \text{ is convex} \\
\Leftrightarrow\ &\forall\lambda \in [0,1], \forall x, y \in \mathrm{dom}(f) : \log f(\lambda x + (1-\lambda)y) \leq \lambda \log f(x) + (1-\lambda)\log f(y) \\
\Leftrightarrow\ &\forall\lambda \in [0,1], \forall x, y \in \mathrm{dom}(f) : f(\lambda x + (1-\lambda)y) \leq f(x)^{\lambda} f(y)^{1-\lambda}.
\end{aligned}$$

The final equivalence follows from the fact that e^x is increasing.

3. Write down the conjugate function of $f(z) = \log(1 + e^{-z})$.

Solution. From the conjugate function, we obtain $f^*(y) = \sup_z[yz - \log(1 + e^{-z})]$. Denote $g(z) := yz - \log(1 + e^{-z})$, then

$$g'(z) = y - \frac{-e^{-z}}{1 + e^{-z}} = y + \frac{1}{e^z + 1}.$$

Notice that $g'(z)$ is strictly decreasing, and $\frac{1}{e^z+1} \in (0, 1)$.

(a) When $-1 < y < 0$, then $g'(z) = 0 \Rightarrow z = \log(-\frac{y+1}{y})$. Thus $f^*(y) = (y+1)\log(y+1) - y\log(-y)$.

(b) When $y \leq -1$, then $g'(z) < 0$, i.e. $g(z)$ is strictly decreasing. Thus $\sup_z g(z) = \lim_{z\to-\infty} g(z)$. By L'Hospital, we have

$$\begin{aligned}
\lim_{z\to-\infty} \frac{z}{\log(1 + e^{-z})} &= \lim_{z\to-\infty} \frac{1}{\frac{-e^{-z}}{1+e^{-z}}} \\
&= -1.
\end{aligned}$$

Therefore,

$$f^*(y) = \sup_z g(z) = \begin{cases} 0, & \text{if } y = -1, \\ +\infty, & \text{if } y < -1. \end{cases}$$

(c) When $y \geq 0$, then $g'(z) > 0$, i.e. $g(z)$ is strictly increasing. Thus $\sup_z g(z) = \lim_{z\to+\infty} g(z)$. We have

$$f^*(y) = \sup_z g(z) = \begin{cases} 0, & \text{if } y = 0, \\ +\infty, & \text{if } y > 0. \end{cases}$$

To summarize,

$$f^*(y) = \begin{cases} +\infty, & \text{if } y < -1, \\ 0, & \text{if } y = -1, \\ (y+1)\log(y+1) - y\log(-y), & \text{if } -1 < y < 0, \\ 0, & \text{if } y = 0, \\ +\infty, & \text{if } y > 0. \end{cases}$$

4. Let $a \in \mathbb{R}^n_+$ with $a \neq 0$ and $b > 0$ be given. Consider the following problem:

$$\begin{aligned} \text{(P)} \quad \min \quad & \frac{1}{2}\|x\|_2^2 \\ \text{s.t.} \quad & a^T x = b, \\ & x \geq 0. \end{aligned}$$

(a) Show that the above problem is feasible, i.e., there exists at least one feasible solution.

(b) Show that an optimal solution always exists, and that it is unique.

(c) Write down the KKT conditions associated with problem (P).

(d) Suppose that you know that the above KKT conditions are necessary and sufficient for optimality. Use the KKT conditions, or otherwise, express the optimal solution to problem (P) in terms of a and b.

Solution.

(a) Since $a \in \mathbb{R}^n_+$ and $a \neq 0$, there exist an i such that $a_i > 0$. Thus, define $\tilde{x}$ as

$$\tilde{x_j} = \begin{cases} \frac{b}{a_i}, & \text{if } j = i, \\ 0, & \text{if } j \neq i. \end{cases}$$

It is easy to check that $a^T \tilde{x} = b$ and $\tilde{x} \geq 0$.

(b) i. Prove that an optimal solution always exists.
Method 1. Denote $S = \{x \mid a^T x = b, x \geq 0\}$, $f(x) = \frac{1}{2}\|x\|_2^2$. Obviously, $\lim_{\|x\| \to \infty} \|x\|_2^2 = \infty$, i.e., f is a coercive function defined over S. Therefore f has a global minimum point over S.
Method 2. Recall that we defined $\tilde{x}$ in (a). It is trivial that $\|x^*\| \leq \|\tilde{x}\|$ if x^* is an optimal solution for (P). Denote $\tilde{S} = \{x \mid a^T x = b, x \geq 0, \|x\| \leq \|\tilde{x}\|\}$, which is a compact set. f is continuous over $\tilde{S}$. By Weierstrass's Theorem, an optimal solution always exists.

ii. Prove that the optimal solution is unique.
To prove it is unique, we just need to consider the Hessian matrix. Since $H = \nabla^2 f(x) = I_n \succ 0$, the objective function is strictly convex, then the optimal solution is unique.

(c) The Lagrangian function is $\mathcal{L}(x, \lambda, \mu) = \frac{1}{2}x^T x - \lambda^T x + \mu(a^T x - b)$, and the KKT conditions are:

$$\begin{aligned} x - \lambda + \mu a &= 0, \\ \lambda^T x &= 0, \\ \lambda &\geq 0, \\ a^T x &= b, \\ x &\geq 0. \end{aligned}$$

Or equivalently:

$$\begin{aligned} x_i - \lambda_i + \mu a_i = 0, &\quad i = 1, 2, \ldots, n \\ \lambda_i x_i = 0, &\quad i = 1, 2, \ldots, n \\ \lambda_i \geq 0, &\quad i = 1, 2, \ldots, n \\ \sum_{i=1}^{n} a_i x_i = b, & \\ x_i \geq 0, &\quad i = 1, 2, \ldots, n. \end{aligned}$$

(d) Since $\lambda^T x = 0,\ \lambda \geq 0, x \geq 0$, we have

$$\lambda_i x_i = 0, \quad i = 1, \ldots, n.$$

From $x - \lambda + \mu a = 0$ and $a^T x = b$, we know

$$x_i = \lambda_i - \mu a_i, \tag{1}$$
$$a^T(\lambda - \mu a) = b \tag{2}$$

From (1) and $\lambda_i x_i = 0$, we have $\lambda_i(\lambda_i - \mu a_i) = 0$, which, together with $\lambda_i \geq 0$, we obtain

$$\lambda_i = \begin{cases} \text{either 0 or } \mu a_i, & \text{if } \mu > 0 \quad (\text{since } a \geq 0), \\ 0, & \text{if } \mu = 0, \\ 0, & \text{if } \mu < 0 \quad (\text{since } a \geq 0). \end{cases}$$

From (2) we see

$$\sum_{i=1}^{n} a_i(\lambda_i - \mu a_i) = b. \tag{3}$$

Thus if $\mu > 0$, we have $\lambda_i - \mu a_i \leq 0$ and then $a_i(\lambda_i - \mu a_i) \leq 0$, which contradicts (3). Hence $\mu \leq 0$ and thus $\lambda = 0$.
Then $x = -\mu a$. Substituting this into $a^T x = b$, we have $\mu = -\frac{b}{a^T a}$ and thus $x = \frac{ba}{a^T a}$.

5. Consider the Quadratically Constrained Quadratic Program (QCQP):

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & x^T Q_0 x + 2b_0^T x \\ \text{s.t.} & x^T Q_i x + 2b_i^T x + c_i \leq 0,\ i = 1, \ldots, m \\ & Ax = b, \end{array}$$

where $Q_i \in \mathbb{S}^n$ (i.e., $n \times n$ symmetric matrix), $b_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R},\ i = 0, \ldots, m,\ A \in \mathbb{R}^{p \times n}$ and $b \in \mathbb{R}^p$ are given.

(a) What constraints must each Q_i satisfy for the problem to be convex?

(b) Derive the dual of the problem.

(c) Derive the SDP relaxation of the problem.

Solution.

(a) $Q_i \succeq 0,\ i = 0, 1, \ldots, m.$

(b) The Lagrangian function is

$$\begin{aligned}\mathcal{L}(x, \lambda, \mu) &= x^T Q_0 x + 2b_0^T x + \sum_{i=1}^m \lambda_i (x^T Q_i x + 2b_i^T x + c_i) + \mu^T (Ax - b) \\ &= x^T (Q_0 + \sum_{i=1}^m \lambda_i Q_i) x + 2(b_0 + \sum_{i=1}^m \lambda_i b_i + \frac{1}{2} A^T \mu)^T x + (\sum_{i=1}^m \lambda_i c_i - \mu^T b).\end{aligned}$$

Then,

$$\begin{aligned}g(\lambda, \mu) &= \inf_x \mathcal{L}(x, \lambda, \mu) \\ &= \sup_t \{t \mid \mathcal{L}(x, \lambda, \mu) \geq t, \forall x \in \mathbb{R}^n\}.\end{aligned}$$

Claim $\mathcal{L}(x, \lambda, \mu) \geq t$ for all $x \in \mathbb{R}^n$ is equivalent to

$$\begin{pmatrix} Q_0 + \sum\limits_{i=1}^m \lambda_i Q_i & b_0 + \sum\limits_{i=1}^m \lambda_i b_i + \frac{1}{2} A^T \mu \\ (b_0 + \sum\limits_{i=1}^m \lambda_i b_i + \frac{1}{2} A^T \mu)^T & \sum\limits_{i=1}^m \lambda_i c_i - \mu^T b - t \end{pmatrix} \succeq 0.$$

Therefore, the dual problem is

$$\begin{aligned}\max_{\lambda, t, \mu} \quad & t \\ \text{s.t.} \quad & \begin{pmatrix} Q_0 + \sum\limits_{i=1}^m \lambda_i Q_i & b_0 + \sum\limits_{i=1}^m \lambda_i b_i + \frac{1}{2} A^T \mu \\ (b_0 + \sum\limits_{i=1}^m \lambda_i b_i + \frac{1}{2} A^T \mu)^T & \sum\limits_{i=1}^m \lambda_i c_i - \mu^T b - t \end{pmatrix} \succeq 0. \\ & \lambda_i \geq 0, \quad i = 1, \ldots, m.\end{aligned}$$

(c)

$$\begin{aligned}\text{(SDR)} \quad \min \quad & Q_0 \bullet X + 2b_0^T x \\ \text{s.t.} \quad & Q_i \bullet X + 2b_i^T x + c_i \leq 0,\ i = 1, \ldots, m \\ & Ax = b, \\ & \begin{pmatrix} X & x \\ x & 1 \end{pmatrix} \succeq 0.\end{aligned}$$

6. Let $L : \mathbb{R}^l \to \mathbb{R}$ be a convex function, $A \in \mathbb{R}^{n \times l}$ and $v \in \mathbb{R}^l$. Denote $L^* : \mathbb{R}^l \to \mathbb{R}$ the conjugate function of L. Show that the dual problem of

$$\min_{w \in \mathbb{R}^n, b \in \mathbb{R}} L(A^T w - bv) + \frac{1}{2} \|w\|_2^2,$$

can be written as

$$\begin{aligned} \min_{\alpha \in \mathbb{R}^l} \quad & L^*(-\alpha) + \frac{1}{2}\|A\alpha\|_2^2 \\ \text{s.t.} \quad & v^T \alpha = 0. \end{aligned}$$

Solution. The problem is equivalent to

$$\begin{aligned} \min_{w,b,x} \quad & L(x) + \frac{1}{2}\|w\|_2^2 \\ \text{s.t.} \quad & x = A^T w - bv. \end{aligned}$$

The Lagrangian function is

$$\begin{aligned} \mathcal{L}(w,b,x,\alpha) &= L(x) + \frac{1}{2}\|w\|_2^2 + \alpha^T(x - A^T w + bv) \\ &= \alpha^T x + L(x) + \frac{1}{2}\|w\|_2^2 - \alpha^T A^T w + b\alpha^T v. \end{aligned}$$

- For x,
$$\inf_x \left(\alpha^T x + L(x)\right) = -\sup_x \left(-\alpha^T x - L(x)\right) = -L^*(-\alpha).$$
- For w, $\frac{1}{2}\|w\|_2^2 - \alpha^T A^T w$ is quadratic. Therefore, when $w = A\alpha$,
$$\inf_w \left(\frac{1}{2}\|w\|_2^2 - \alpha^T A^T w\right) = -\frac{1}{2}\|A\alpha\|_2^2$$
.
- For b,
$$\inf_b b\alpha^T v = \begin{cases} 0, & \alpha^T v = 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

Therefore,

$$g(\alpha) = \inf_{w,b,x} \mathcal{L}(w,b,x,\alpha) = \begin{cases} -L^*(-\alpha) - \frac{1}{2}\|A\alpha\|_2^2, & \alpha^T v = 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{aligned} \max_{\alpha} \quad & -L^*(-\alpha) - \frac{1}{2}\|A\alpha\|_2^2 \\ \text{s.t.} \quad & \alpha^T v = 0, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \min_{\alpha} \quad & L^*(-\alpha) + \frac{1}{2}\|A\alpha\|_2^2 \\ \text{s.t.} \quad & v^T \alpha = 0. \end{aligned}$$