DATA130026.01 Optimization
Solution of Assignment 6

1. Consider the primal optimization problem

$$\begin{aligned} \min \quad & x_1^4 - 2x_2^2 - x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 + x_2 \leq 0. \end{aligned}$$

(i) Is the problem convex?

(ii) Does there exist an optimal solution to the problem?

(iii) Write a dual problem. Solve it.

(iv) Is the optimal value of the dual problem equal to the optimal value of primal problem? Find the optimal solution of the primal problem.

Solution.

(i) Since $\nabla^2 f_0(x) = \begin{pmatrix} 12x_1^2 & 0 \\ 0 & -4 \end{pmatrix}$ is not semidefinite, the problem isn't convex.

(ii) We know that $\{(x_1, x_2) |\, x_1^2+x_2^2+x_2 \leq 0\}$ is a compact set, and $f_0(x) = x_1^4-2x_2^2-x_2$ is a continuous function. Since any continuous function on a compact set attains a maximum and minimum, there exists an optimal solution to the problem.

(iii) The Lagrangian function is

$$\begin{aligned} L(x_1, x_2, \lambda) &= x_1^4 - 2x_2^2 - x_2 + \lambda(x_1^2 + x_2^2 + x_2) \\ &= x_1^4 + \lambda x_1^2 + (\lambda - 2)x_2^2 + (\lambda - 1)x_2. \end{aligned}$$

- If $0 \leq \lambda < 2$, then the coefficient of x_2^2 is less than 0, so $\inf L(x_1, x_2, \lambda) = -\infty$;
- If $\lambda = 2$, then $L(x_1, x_2, \lambda) = x_1^4 + 2x_1^2 + x_2$. Since x_2 isn't bounded below, we have $\inf L(x_1, x_2, 2) = -\infty$;
- If $\lambda > 2$, since

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 4x_1^3 + 2\lambda x_1 = 0 \\[2ex] \dfrac{\partial L}{\partial x_2} = -4x_2 - 1 + 2\lambda x_2 + \lambda = 0, \end{cases}$$

we have

$$\begin{cases} x_1 = 0 \\ x_2 = \dfrac{1-\lambda}{2\lambda - 4}. \end{cases}$$

So $\inf L(x_1, x_2, \lambda) = L(0, \dfrac{1-\lambda}{2\lambda-4}, \lambda) = -\dfrac{(\lambda-1)^2}{4(\lambda-2)}$.

Therefore, the dual function is

$$g(\lambda) = \inf_{x_1, x_2} L(x_1, x_2, \lambda) = \begin{cases} -\dfrac{(\lambda-1)^2}{4(\lambda-2)}, & \lambda > 2 \\ -\infty, & 0 \le \lambda \le 2. \end{cases}$$

And the dual problem is

$$\begin{aligned} \max \quad & -\frac{(\lambda-1)^2}{4(\lambda-2)} \\ \text{s.t.} \quad & \lambda > 2. \end{aligned}$$

Since

$$\begin{aligned} g'(\lambda) &= -\frac{2(\lambda-1)4(\lambda-2) - 4(\lambda-1)^2}{16(\lambda-2)^2} \\ &= -\frac{(\lambda-1)(\lambda-3)}{4(\lambda-2)^2}, \end{aligned}$$

we have $\lambda^* = 3$ and $v_d^* = \max g(\lambda) = g(\lambda^*) = -1$.

(iv) Since Slater's condition holds for this problem, the necessity of the KKT conditions also holds. The KKT conditions are

$$\begin{aligned} 4x_1^3 + 2\lambda x_1 &= 0, & (1) \\ -4x_2 - 1 + 2\lambda x_2 + \lambda &= 0, & (2) \\ x_1^2 + x_2^2 + x_2 &\le 0, & (3) \\ \lambda &\ge 0, & (4) \\ \lambda(x_1^2 + x_2^2 + x_2) &= 0. & (5) \end{aligned}$$

We abtain that

$$(x_1, x_2, \lambda) = (0, -\frac{1}{4}, 0),\ (0, 0, 1),\ (0, -1, 3)$$

satisfy the KKT system. Hence,

$$(x_1, x_2) = (0, -\frac{1}{4}),\ (0, 0),\ (0, -1)$$

are KKT points.

- When $(x_1, x_2) = (0, -\frac{1}{4})$, $f_0(x_1, x_2) = \frac{1}{8}$;
- When $(x_1, x_2) = (0, 0)$, $f_0(x_1, x_2) = 0$;
- When $(x_1, x_2) = (0, -1)$, $f_0(x_1, x_2) = -1$.

So the optimal solution of the problem is $x^* = (0, -1)$, and the optimal value is $v_p^* = -1$. Obviously, the optimal value of the dual problem v_d^* is equal to the optimal value of prime problem v_p^*.

2. Find a dual problem to the following convex minimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{n}(a_i x_i^2 + 2b_i x_i + e^{\alpha_i x_i}) \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i = 1, \end{aligned}$$

where $\mathbf{a}, \alpha \in \mathbb{R}^{\mathbf{n}}_{++}, \mathbf{b} \in \mathbb{R}^{\mathbf{n}}$.

Solution. The primal problem can be converted to

$$\begin{aligned} \min \quad & \sum_{i=1}^{n}(a_i x_i^2 + 2b_i x_i + e^{\alpha_i z_i}) \\ \text{s.t.} \quad & z_i = x_i, \quad i = 1, ..., n \\ & \sum_{i=1}^{n} x_i = 1. \end{aligned}$$

where $\mathbf{a}, \alpha \in \mathbb{R}^{\mathbf{n}}_{++}, \mathbf{b} \in \mathbb{R}^{\mathbf{n}}$.
The Lagrangian function is

$$\begin{aligned} L(\mathbf{x}, \mathbf{z}, \boldsymbol{\nu}, \mu) &= \sum_{i=1}^{n} a_i x_i^2 + 2b_i x_i + e^{\alpha_i z_i} + \sum_{i=1}^{n} \nu_i (z_i - x_i) + \mu(\sum_{i=1}^{n} x_i - 1) \\ &= \sum_{i=1}^{n} L_i - \mu, \end{aligned}$$

where

$$L_i = a_i x_i^2 + (2b_i - \nu_i + \mu) x_i + e^{\alpha_i z_i} + \nu_i z_i.$$

Since $a_i > 0$, $a_i x_i^2 + (2b_i - \nu_i + \mu)x_i$ is always bounded below, and the lower boundary is

$$\inf_{x_i} a_i x_i^2 + (2b_i - \nu_i + \mu) x_i = -\frac{(\nu_i - \mu - 2b_i)^2}{4a_i}.$$

Then we need to discuss the lower boundary of $e^{\alpha_i z_i} + \nu_i z_i$.

- If $\nu_i > 0$, then $\inf e^{\alpha_i z_i} + \nu_i z_i = -\infty$ when $z_i \to -\infty$.
- If $\nu_i = 0$, then $\inf e^{\alpha_i z_i} + 0 = 0$ when $z_i \to -\infty$.
- If $\nu_i < 0$, then $\inf e^{\alpha_i z_i} + \nu_i z_i = -\frac{\nu_i}{\alpha_i} + \frac{\nu_i}{\alpha_i} \ln(-\frac{\nu_i}{\alpha_i})$ by $\frac{\partial L_i}{\partial z_i} = \alpha_i e^{\alpha_i z_i} + \nu_i = 0$.

Therefore, the dual function is

$$g(\boldsymbol{\nu}, \mu) = \inf_{\mathbf{x}, \mathbf{z}} L(\mathbf{x}, \mathbf{z}, \boldsymbol{\nu}, \mu) = \begin{cases} \sum_{i=1}^{n} g_i(\nu_i, \mu) - \mu, & \boldsymbol{\nu} \le 0 \\ -\infty, & \text{otherwise}, \end{cases}$$

where

$$g_i(\nu_i, \mu) = \begin{cases} -\dfrac{(\nu_i - \mu - 2b_i)^2 + 4\nu_i}{4a_i} + \dfrac{\nu_i}{\alpha_i}\ln(-\dfrac{\nu_i}{\alpha_i}), & \nu_i < 0 \\ -\dfrac{(\mu + 2b_i)^2}{4a_i}, & \nu_i = 0. \end{cases}$$

And the dual problem is

$$\begin{array}{ll} \max & \displaystyle\sum_{i=1}^{n} g_i(\nu_i, \mu) - \mu \\ \text{s.t.} & \boldsymbol{\nu} \leq \mathbf{0}, \end{array}$$

with $g_i(\nu_i, \mu)$ defined above.

3. Consider the following optimization problem in the variables $\alpha \in \mathbb{R}$ and $q \in \mathbb{R}^n$:

$$(\text{P}) \qquad \begin{array}{ll} \min & \alpha \\ \text{s.t.} & Aq = \alpha f \\ & \|q\|_2^2 \leq \epsilon, \end{array}$$

where $A \in \mathbb{R}^{m \times n}$, $f \in \mathbb{R}^m$, $\epsilon \in \mathbb{R}_{++}$. Assume in addition that the rows of A are linearly independent.

(a) Explain why strong duality holds for problem.

(b) Find a dual problem to problem (P). (Do not assign a Lagrange multiplier to the quadratic constraint.)

(c) Solve the dual problem obtained in part (ii) and find the optimal solution of problem (P).

Solution.

(a) Obviously it's a convex problem. Assume $(\alpha_0, \boldsymbol{q}_0) = (0, \mathbf{0})$, then

$$\begin{cases} \mathbf{A0} = 0 \cdot \boldsymbol{f} \\ \|\mathbf{0}\|_2^2 < \epsilon, \end{cases}$$

so the Slater's Condition holds. Thus the strong duality holds for the problem.

(b) The Lagrangian function is

$$\begin{aligned} L(\alpha, \boldsymbol{q}, \boldsymbol{\nu}) &= \alpha + \boldsymbol{\nu}^T(\mathbf{A}\boldsymbol{q} - \alpha \boldsymbol{f}) \\ &= (1 - \boldsymbol{\nu}^T \boldsymbol{f})\alpha + \boldsymbol{\nu}^T \mathbf{A}\boldsymbol{q} \quad (\|\boldsymbol{q}\|_2^2 \leq \epsilon). \end{aligned}$$

- For $(1 - \boldsymbol{\nu}^T \boldsymbol{f})\alpha$, it's bounded below only when $\boldsymbol{\nu}^T \boldsymbol{f} = 1$.

- For $\boldsymbol{\nu}^T \mathbf{A}\boldsymbol{q} \quad (\|\boldsymbol{q}\|_2^2 \le \epsilon)$,

$$\boldsymbol{\nu}^T \mathbf{A}\boldsymbol{q} = \|\mathbf{A}^T\boldsymbol{\nu}\|_2 \ \|\boldsymbol{q}\|_2 \cdot \cos\langle \mathbf{A}^T\boldsymbol{\nu}, \boldsymbol{q}\rangle.$$

  When $\boldsymbol{q} = -\sqrt{\epsilon}\ \dfrac{\mathbf{A}^T\boldsymbol{\nu}}{\|\mathbf{A}^T\boldsymbol{\nu}\|_2}$ $(\mathbf{A}^T\boldsymbol{\nu} \neq \mathbf{0})$, we have

$$\inf_{\boldsymbol{q}} \boldsymbol{\nu}^T \mathbf{A}\boldsymbol{q} = -\sqrt{\epsilon}\ \|\mathbf{A}^T\boldsymbol{\nu}\|_2.$$

  It also holds when $\mathbf{A}^T\boldsymbol{\nu} = \mathbf{0}$.

So the dual function is

$$g(\boldsymbol{\nu}) = \inf_{\alpha, \|\boldsymbol{q}\|_2^2 \le \epsilon} L(\alpha, \boldsymbol{q}, \boldsymbol{\nu}) = \begin{cases} -\sqrt{\epsilon}\ \|\mathbf{A}^T\boldsymbol{\nu}\|_2, & \boldsymbol{\nu}^T\boldsymbol{f} = 1 \\ -\infty, & \text{otherwise.} \end{cases}$$

And the dual problem is

$$\begin{array}{ll} \max & -\sqrt{\epsilon}\ \|\mathbf{A}^T\boldsymbol{\nu}\|_2 \\ \text{s.t.} & \boldsymbol{f}^T\boldsymbol{\nu} = 1. \end{array}$$

(c) The dual problem can be converted to

$$\begin{array}{ll} \min & \dfrac{1}{2}\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} \\ \text{s.t.} & \boldsymbol{f}^T\boldsymbol{\nu} = 1. \end{array}$$

It's a convex problem with an equality constraint, thus the KKT conditions are necessary and sufficiency.
The Lagrangian function of the dual problem is

$$L_d(\boldsymbol{\nu}, \mu) = \frac{1}{2}\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mu(\boldsymbol{f}^T\boldsymbol{\nu} - 1).$$

The KKT conditions are

$$\frac{\partial L_d}{\partial \boldsymbol{\nu}} = \mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mu\boldsymbol{f} = \mathbf{0} \tag{6}$$

$$\boldsymbol{f}^T\boldsymbol{\nu} = 1 \tag{7}$$

Since the rows of $\mathbf{A}$ are linerly independent, $\mathbf{A}\mathbf{A}^T$ is reversible and $\mathbf{A}\mathbf{A}^T \succ 0$. Then, we obtain that

$$(\boldsymbol{\nu}, \mu) = \left(\frac{(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}}{f^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}}, \frac{1}{f^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}} \right)$$

satisfies the KKT system. Hence, the optimal solution of the dual problem is

$$\boldsymbol{\nu}^* = \frac{(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}}{f^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}}.$$

And the optimal value of the dual problem is

$$v_d^* = -\sqrt{\epsilon}\,\|\mathbf{A}^T\boldsymbol{\nu}^*\|_2 = -\frac{\sqrt{\epsilon}}{\|\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}\|_2}.$$

By the strong duality, we have

$$v_p^* = \alpha^* = v_d^* = -\frac{\sqrt{\epsilon}}{\|\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}\|_2}.$$

And the optimal solution of the prime problem is

$$(\alpha^*, \boldsymbol{q}^*) = \left(-\frac{\sqrt{\epsilon}}{\|\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}\|_2}, -\frac{\sqrt{\epsilon}}{\|\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}\|_2}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\boldsymbol{f}\right).$$

4. Consider the convex optimization problem

$$\begin{array}{ll} \min & \sum_{j=1}^n x_j \ln \frac{x_j}{c_j} \\ \text{s.t.} & Ax \geq b \\ & \sum_{j=1}^n x_j = 1, \end{array}$$

where $c > 0$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Find a dual problem. Simplify it as much as possible.

Solution. The Lagrangian function is

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}, \nu) &= \sum_{j=1}^n x_j \ln \frac{x_j}{c_j} + \boldsymbol{\lambda}^T(-\mathbf{A}\mathbf{x} + \mathbf{b}) + \nu(\sum_{j=1}^n x_j - 1) \\ &= \sum_{j=1}^n \left(x_j \ln \frac{x_j}{c_j} - \boldsymbol{\lambda}^T\boldsymbol{a_j}x_j + \nu x_j\right) + \boldsymbol{\lambda}^T\mathbf{b} - \nu, \end{aligned}$$

where $\boldsymbol{a_j} \in \mathbb{R}^m$ is the i^{th} column of $\mathbf{A}$. Thus

$$\frac{\partial L}{\partial x_j} = \ln \frac{x_j}{c_j} + 1 - \boldsymbol{\lambda}^T\boldsymbol{a_j} + \nu.$$

Obviously, $\frac{\partial L}{\partial x_j}$ is strictly increasing, and $\inf \frac{\partial L}{\partial x_j} = -\infty$, $\sup \frac{\partial L}{\partial x_j} = \infty$. Then $\frac{\partial L}{\partial x_j} = 0$ has a unique solution $x_j = c_j e^{\boldsymbol{\lambda}^T\boldsymbol{a_j} - \nu - 1}$.
Therefore, the dual function is

$$\begin{aligned} g(\boldsymbol{\lambda}, \nu) &= \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \nu) \\ &= \inf_{\mathbf{x}} \sum_{j=1}^n \left(x_j(\ln \frac{x_j}{c_j} - \boldsymbol{\lambda}^T\boldsymbol{a_j} + \nu)\right) + \boldsymbol{\lambda}^T\mathbf{b} - \nu \\ &= -\sum_{j=1}^n c_j e^{\boldsymbol{\lambda}^T\boldsymbol{a_j} - \nu - 1} + \boldsymbol{\lambda}^T\mathbf{b} - \nu. \end{aligned}$$

And the dual problem is

$$
\begin{aligned}
\max \quad & -\sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j} - \nu - 1} + \boldsymbol{\lambda}^T \mathbf{b} - \nu \\
\text{s.t.} \quad & \boldsymbol{\lambda} \geq \mathbf{0},
\end{aligned}
$$

where $\boldsymbol{a_j} \in \mathbb{R}^m$ is the i^{th} column of $\mathbf{A}$.
Now let's simplify it. Because ν has no constraints, we have

$$
g(\boldsymbol{\lambda}, \nu) = -e^{-\nu-1} \sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j}} + \boldsymbol{\lambda}^T \mathbf{b} - \nu,
$$

$$
\frac{\partial g}{\partial \nu} = \sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j}} \cdot e^{-\nu-1} - 1
$$

$$
\frac{\partial^2 g}{\partial \nu^2} = -\sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j}} \cdot e^{-\nu-1} < 0.
$$

Thus, $\frac{\partial g}{\partial \nu}$ is strictly decreasing. Since $\inf \frac{\partial g}{\partial \nu} = -1$ and $\sup \frac{\partial g}{\partial \nu} = +\infty$, we know that $\frac{\partial g}{\partial \nu} = 0$ has a unique solution $\nu^* = \ln \left(\sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j}} \right) - 1$. Plugging it into the dual problem, we have

$$
\begin{aligned}
\max \quad & -\ln \left(\sum_{j=1}^{n} c_j e^{\boldsymbol{\lambda}^T \boldsymbol{a_j}} \right) + \boldsymbol{\lambda}^T \mathbf{b} \\
\text{s.t.} \quad & \boldsymbol{\lambda} \geq \mathbf{0},
\end{aligned}
$$

where $\boldsymbol{a_j} \in \mathbb{R}^m$ is the i^{th} column of $\mathbf{A}$.