Exercises on Chapter 2: Convex optimization

Mathematics of Data Science, M1 IDD

November 2023

Exercise 2.1: Reformulations

a) Write the epigraph reformulation of the problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}}\, \|\boldsymbol{x}\|_\infty,$$

where $\|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i|$.

b) Using that $\|\boldsymbol{x}\|_\infty \ge 0$, reformulate the epigraph problem so that it consists in a linear of objective function and n linear constraints.

c) More generally, consider the Chebyshev approximation problem

$$\underset{\boldsymbol{y}\in\mathbb{R}^m}{\text{minimize}}\, \|\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b}\|_\infty \quad \text{where } \boldsymbol{A}\in\mathbb{R}^{n\times m} \text{ and } \boldsymbol{b}\in\mathbb{R}^n.$$

Find a reformulation of this problem as a problem with linear constraints and linear objective. Justify that the two problems have the same optimal value.

d) Consider now a function $\boldsymbol{\phi}:\mathbb{R}^m\to\mathbb{R}^n$ be defined by $\boldsymbol{\phi}(\boldsymbol{y}) = [\phi_i(\boldsymbol{y})]_{i=1}^n$, where every $\phi_i:\mathbb{R}^m\to\mathbb{R}$ is a nonnegative convex function.
Using the result of the previous question, reformulate the problem

$$\underset{\boldsymbol{y}\in\mathbb{R}^m}{\text{minimize}}\, \|\boldsymbol{\phi}(\boldsymbol{y})\|_\infty$$

as a convex optimization problem in standard form.

Exercise 2.2: Nonlinear equality constraints

Consider the optimization problem

$$\begin{cases} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & f(\boldsymbol{x}) := \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{a}\|_2^2 \\ \text{subject to} & \sum_{i=1}^{n-1} x_i^2 = 0. \end{cases}$$

where $\boldsymbol{a}\in\mathbb{R}^n$.

a) Show that the feasible set of this problem is convex, and give a description of this set that only involves convex inequalities and/or linear equalities.

b) Justify that the problem is convex

c) Find the unique solution to this problem by hand.

d) Use the first-order optimality condition to compute the solution to this problem. *NB: The gradient of the objective is* $\nabla f(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{a}$.

Exercise 2.3: Existence and optimality

We consider the following optimization problem:

$$\begin{cases} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^3} & f(\boldsymbol{x}) := x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 + 2x_2 - x_3 = 4 \\ & x_1 - x_2 + x_3 = -2. \end{cases}$$

a) Justify that the problem is convex, and give a standard form reformulation.

b) Show that there is a unique solution to the problem..

c) Using the first-order optimality condition, check that the point $\boldsymbol{x}^* = \left[\frac{2}{7}, \frac{10}{7}, -\frac{6}{7}\right]^{\mathrm{T}}$ is the solution.

d) What happens to the solution set of this problem in absence of constraints (i.e. the problem becomes $\text{minimize}_{\boldsymbol{x}\in\mathbb{R}^3} f(\boldsymbol{x})$) ?

Exercise 2.4: Convex conjugate

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. The convex conjugate of f, denoted by $\overline{f}(\cdot)$, is given by

$$\forall \boldsymbol{y} \in \mathbb{R}^n, \quad \overline{f}(\boldsymbol{y}) = \sup_{\boldsymbol{x}\in\mathrm{dom}(f)} (\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x})).$$

a) Using that the supremum of convex functions is always convex, justify that the function $\overline{f}$ is convex.

b) Suppose that f is convex and differentiable on $\mathrm{dom}(f) = \mathbb{R}^n$. Using the result of question a), show that for $\boldsymbol{z} \in \mathbb{R}^n$, we have

$$\overline{f}(\nabla f(\boldsymbol{z})) = \boldsymbol{z}^{\mathrm{T}}\nabla f(\boldsymbol{z}) - f(\boldsymbol{z}).$$

c) Give a closed-form expression for $\overline{f}$ when f is linear on $\mathbb{R}^n$, i.e. $f(\boldsymbol{x}) = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} + b$ for some $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$;

d) Show that if $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}\|^2 = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$, then $\overline{f} = f$.

e) Consider the problem

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

where $\boldsymbol{A} \in \mathbb{R}^{\ell\times n}$ et $\boldsymbol{b} \in \mathbb{R}^{\ell}$. Show that the dual function of this problem, denoted by $d(\boldsymbol{\mu})$, is given by

$$d(\boldsymbol{\mu}) = -\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu} - \overline{f}(-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\mu}).$$

Exercise 2.5: Weak duality

Consider the problem

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^2} & f(\boldsymbol{x}) \\ \text{subject to} & \frac{x_1^2}{x_2} \leq 0, \end{array}, \quad \text{where} \quad f(\boldsymbol{x}) = \begin{cases} \exp(-x_1) & \text{if } x_2 > 0 \\ \infty & \text{otherwise.} \end{cases} \tag{1}$$

The domain of the problem is $\mathcal{X} = \{(x_1, x_2) | x_2 > 0\}$.

a) Justify that this problem is convex. What is its optimal value?

b) Write the dual problem of (1), give a formula for its solution λ^* and its optimal value d^*. What is then the duality gap?

c) Show that Slater's constraint qualification does not hold for this problem. *Hint: Use the property* $\mathrm{ri}(\mathcal{X}) = \mathcal{X}$.

Exercise 2.6: KKT and optimality conditions

Let $f : \boldsymbol{x} \mapsto 2x_1^2 + 4x_2^2 - x_1 - x_2$ be a function from $\mathbb{R}^2$ to $\mathbb{R}$ and let

$$\mathcal{C} = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0 \text{ et } x_2 \geq 0\right\}.$$

a) Justify that f is strongly convex.

b) Justify that $\mathcal{C}$ is a convex cone and express then the first-order optimality condition for the problem $\text{minimize}_{\boldsymbol{x}\in\mathcal{C}} f(\boldsymbol{x})$. Give then the solution of this problem.

c) Write the KKT conditions for this problem, and justify that they can be used to compute a solution.

Solutions

Solution for Exercise 2.1: Reformulations

Question a) By definition, the epigraph reformulation of the problem is

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ t\in\mathbb{R}}} & t \\ \text{subject to} & \|\boldsymbol{x}\|_\infty \leq t. \end{array}$$

Question b) Using the definition of $\|\boldsymbol{x}\|_\infty$, we have that

$$\begin{aligned} \|\boldsymbol{x}\|_\infty \leq t \quad &\Leftrightarrow \quad \max_{1\leq i\leq n} |x_i| \leq t \\ &\Leftrightarrow \quad |x_i| \leq t \quad \forall i = 1,\dots,n \\ &\Leftrightarrow \quad t \geq 0 \text{ and } -t \leq x_i \leq t \quad \forall i = 1,\dots,n \end{aligned}$$

As a result, the problem can be equivalently reformulated as the linear program

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ t\in\mathbb{R}}} & t \\ \text{subject to} & -t - x_i \leq 0 \quad \forall i = 1,\dots,n \\ & -t + x_i \leq 0 \quad \forall i = 1,\dots,n \\ & t \leq 0. \end{array}$$

Question c) Let $[\boldsymbol{a}_i^{\mathrm{T}}]_{i=1}^n$ denote the rows of the matrix $\boldsymbol{A}$, and let $\phi_i(\boldsymbol{y}) = \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y} - b_i$, where b_i is the corresponding coefficient of the vector $\boldsymbol{b}$. Applying the same reasoning as in question b), we obtain the following reformulation:

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{y}\in\mathbb{R}^m \\ t\in\mathbb{R}}} & t \\ \text{subject to} & -t + \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y} - b_i \leq 0 \quad \forall i = 1,\dots,n \\ & -t - \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{y} + b_i \leq 0 \quad \forall i = 1,\dots,n \\ & -t \leq 0. \end{array}$$

This problem has a linear objective as well as linear constraints.

The solutions of the problem are that of the epigraph formulation of the original problem. As a result, if $(\boldsymbol{y}^*, t^*)$ is a solution of the problem, then $\boldsymbol{y}^*$ is a solution of the original problem, and we must have $t^* = \|\boldsymbol{A}\boldsymbol{y}^* - \boldsymbol{b}\|_\infty$, otherwise this point would not be an optimal solution.

Question d) By using the epigraph form of the problem and applying the same reasoning as in question b), we obtain the formulation

$$\begin{array}{ll} \text{minimize}_{\substack{\boldsymbol{y}\in\mathbb{R}^m \\ t\in\mathbb{R}}} & t \\ \text{subject to} & -t - \phi_i(\boldsymbol{y}) \leq 0 \quad \forall i = 1,\dots,n \\ & t - \phi_i(\boldsymbol{y}) \leq 0 \quad \forall i = 1,\dots,n \\ & -t \leq 0. \end{array}$$

Since $\phi_i(\boldsymbol{y}) \geq 0$ for any $\boldsymbol{y}$, the constraint $-t - \phi_i(\boldsymbol{y}) \leq 0$ is always satisfied for $t \geq 0$. As a result, the problem can be equivalently reformulated as

$$\begin{array}{lll} \text{minimize}_{\substack{\boldsymbol{y}\in\mathbb{R}^m \\ t\in\mathbb{R}}} & t & \\ \text{subject to} & -t + \phi_i(\boldsymbol{y}) \leq 0 & \forall i = 1, \dots, n \\ & -t \leq 0. & \end{array}$$

Since a sum of convex functions is convex, all functions of the form $(\boldsymbol{y}, t) \mapsto -t + \phi_i(\boldsymbol{y})$ involved in the inequality constraints are convex, hence the problem is a convex problem in standard form.

Solution for Exercise 2.2: Nonlinear equality constraints

Question a) Let $\mathcal{F}$ denote the feasible set. For any $\boldsymbol{x} \in \mathcal{F}$, we have $x_1 = x_2 = \cdots = x_{n-1}$, and this property is preserved by convex combinations. As a result,

$$\forall (\boldsymbol{x}, \boldsymbol{y}) \in (\mathbb{R}^n)^2, \forall \alpha \in [0,1], \quad \sum_{i=1}^{n-1} (\alpha x_i + (1-\alpha) y_i)^2 = 0,$$

hence $\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \mathcal{F}$, proving that the set is convex. Our reasoning provides the following description of $\mathcal{F}$:

$$\mathcal{F} = \{\boldsymbol{x} \in \mathbb{R}^n \mid x_1 = 0, \dots, x_{n-1} = 0\}.$$

Question b) The objective function is a sum of quadratic functions of the entries of $\boldsymbol{x}$, that are all convex functions. As a result, the objective function is convex, hence the problem is a convex optimization problem.

Question c) For any $\boldsymbol{x} \in \mathcal{F}$, $f(\boldsymbol{x}) = \frac{1}{2}(x_n - a_n)^2 + \frac{1}{2}\sum_{i=1}^n a_i^2$. As a result, for any $\boldsymbol{x} \in \mathcal{F}$, we have $f(\boldsymbol{x}) \geq \frac{1}{2}\sum_{i=1}^n a_i^2$, and this value is only attained for $\boldsymbol{x} = [0 \cdots 0 a_n]^{\mathrm{T}}$. As a result, we have

$$\inf \{f(\boldsymbol{x}) | \boldsymbol{x} \in \mathcal{F}\} = \frac{1}{2}\sum_{i=1}^n a_i^2 \quad \text{and} \quad \operatorname{argmin} \{f(\boldsymbol{x}) | \boldsymbol{x} \in \mathcal{F}\} = \{[0 \cdots 0 a_n]^{\mathrm{T}}\}.$$

Question d) Since the problem is convex and the objective is differentiable, we know that a point $\bar{\boldsymbol{x}}$ is a solution if and only if

$$\bar{\boldsymbol{x}} \in \mathcal{F} \quad \text{and} \quad \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{z} - \bar{\boldsymbol{x}}) = 0 \forall \boldsymbol{z} \in \mathcal{F}.$$

Thanks to the description of $\mathcal{F}$, we know that any vector in $\mathcal{F}$ has its first $n-1$ components equal to 0. Thus, the conditions above translate to

$$\begin{array}{lll} x_1 & = & 0 \\ \vdots & & \\ x_{n-1} & = & 0 \\ [\nabla f(\bar{\boldsymbol{x}})]_n (z_n - \bar{x}_n) = 0 \forall z_n \in \mathbb{R}. & & \end{array}$$

Using $\nabla f(\bar{\boldsymbol{x}}) = \bar{\boldsymbol{x}} - \boldsymbol{a}$, the last condition becomes $(\bar{x}_n - a_n)(z_n - \bar{x}_n) = 0$, which only holds for all $z_n \in \mathbb{R}$ if $\bar{x}_n - a_n = 0$. As a result, we obtain the system

$$\begin{array}{lcl} x_1 & = & 0 \\ \vdots & & \\ x_{n-1} & = & 0 \\ x_n - a_n & = & 0, \end{array}$$

confirming the result of Question c).

Solution for Exercise 2.3: Existence and optimality

Question a) The objective function is convex on $\mathbb{R}^3$ as the sum of three convex functions. Moreover, the feasible set is an affine set in $\mathbb{R}^3$, hence it is convex. As a result, the problem is a convex optimization problem. Up to the right-hand sides of the linear equality constraints, it is already in standard form.

Question b) The objective function is strongly convex since it is a quadratic function based on a positive definite matrix (the identity). Moreover, the feasible set is not empty as it contains the vector $[0 \ \ -2 \ 0]^{\mathrm{T}}$. As a result, we are guaranteed that there exists a unique solution.

Question c) Given that the feasible set of the problem is an affine set, we know that the solution $\boldsymbol{x}^*$ of the problem is the unique vector such that

$$\begin{array}{lcl} x_1^* + 2x_2^* - x_3^* & = & 4 \\ x_1^* - x_2^* + x_3^* & = & -2 \\ \nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}^*) & = & 0 \quad \forall \boldsymbol{y} \in \mathcal{F}, \end{array}$$

where $\mathcal{F}$ denotes the feasible set. It is clear that the choice $\boldsymbol{x}^* = \left[\frac{2}{7}, \frac{10}{7}, -\frac{6}{7}\right]^{\mathrm{T}}$ satisfies the first two conditions. Using $\nabla f(\boldsymbol{x}) = 2\boldsymbol{x}$, the latter condition can be written (after division by 2) as

$$x_1^*(y_1 - x_1^*) + x_2^*(y_2 - x_2^*) + x_3^*(y_3 - x_3^*) = 0 \Leftrightarrow 14y_1 + 70y_2 - 42y_3 = 140.$$

Since $\boldsymbol{y}$ is feasible, it satisfies the two constraints

$$\begin{array}{lcl} y_1 + 2y_2 - y_3 & = & 4 \\ y_1 - y_2 + y_3 & = & -2 \end{array} \Leftrightarrow \begin{array}{lcl} y_3 & = & -3y_1 \\ y_2 & = & 2 - 2y_1, \end{array}$$

implying that

$$14y_1 + 70y_2 - 42y_3 = 14y_1 + 140 - 140y_1 + 126y_1 = 140,$$

as desired. This shows that $\boldsymbol{x}^*$ is the solution to the problem.

Question d) In absence of constraints, the problem still has a unique solution, as the function is strongly convex over $\mathbb{R}^n$. This solution is immediately found to be the zero vector.

Solution for Exercise 2.4: Convex conjugate

Question a) For any $\boldsymbol{x} \in \mathrm{dom}(f)$, the function $\boldsymbol{y} \mapsto \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x})$ is a linear function of $\boldsymbol{y}$ hence it is a convex function of $\boldsymbol{y}$. Since the supremum of convex functions is convex, this shows that $\overline{f}$ is convex. *N.B. The proof of convexity of $\sup_{i\in\mathcal{I}} f_i$ with f_i convex follows from the definition of the supremum.*

Question b) Computing $\overline{f}(\nabla f(\boldsymbol{z})) = \sup_{\boldsymbol{x}\in\mathbb{R}^n} \left\{\nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x})\right\}$ amounts to computing the optimal value for the problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\mathrm{maximize}}\, \nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x}),$$

which is the opposite of that of

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\mathrm{minimize}}\, \phi(\boldsymbol{x}) = f(\boldsymbol{x}) - \nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x}.$$

This problem is convex and differentiable. As a result, we know that the solutions to this problem are characterized by the equation

$$\nabla\phi(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{z}) = \mathbf{0}.$$

The point $\boldsymbol{x} = \boldsymbol{z}$ is a solution of this equation, hence it gives the optimal value for the problem. As a result,

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} \phi(\boldsymbol{x}) = f(\boldsymbol{z}) - \nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{z},$$

and thus

$$\overline{f}(\nabla f(\boldsymbol{z})) = \sup_{\boldsymbol{x}\in\mathbb{R}^n} \left\{\nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x})\right\} = -\inf_{\boldsymbol{x}\in\mathbb{R}^n} \phi(\boldsymbol{x}) = \nabla f(\boldsymbol{z})^{\mathrm{T}}\boldsymbol{z} - f(\boldsymbol{z}).$$

Question c) If $f(\boldsymbol{x}) = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} - b$, then $\overline{f}(\boldsymbol{y}) = \sup_{\boldsymbol{x}\in\mathbb{R}^n} \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} - b$, hence $\overline{f}$ is defined as the supremum of linear functions. Given that a linear function is only bounded when it is constant (recall Exercise 1.6), we have $\overline{f}(\boldsymbol{a}) = -b$ and $\overline{f}(\boldsymbol{y}) = \infty$ for any $\boldsymbol{y} \neq \boldsymbol{a}$.

Question d) For any $\boldsymbol{x}, \boldsymbol{y} \in (\mathbb{R}^n)^2$, we have

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - \frac{1}{2}\|\boldsymbol{x}\|^2 \leq \|\boldsymbol{y}\|\|\boldsymbol{x}\| - \frac{1}{2}\|\boldsymbol{x}\|^2$$

by Cauchy-Schwarz inequality. The function $t \mapsto \|\boldsymbol{y}\|t - \frac{t^2}{2}$ is maximized at $t = \|\boldsymbol{y}\|$, hence

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x}) \leq \frac{1}{2}\|\boldsymbol{y}\|^2 \;\Rightarrow\; \overline{f}(\boldsymbol{y}) \leq \frac{1}{2}\|\boldsymbol{y}\|^2.$$

It suffices then to notice that the first inequality is an equality for $\boldsymbol{x} = \boldsymbol{y}$ to obtain

$$\frac{1}{2}\|\boldsymbol{y}\|^2 = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - f(\boldsymbol{y}) \leq \sup_{\boldsymbol{x}\in\mathbb{R}^n} \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x} - f(\boldsymbol{x}) = f^*(\boldsymbol{y}) \leq \frac{1}{2}\|\boldsymbol{y}\|^2,$$

hence the result.

Question e) By definition of the dual function, we have for any $\boldsymbol{\mu} \in \mathbb{R}^{\ell}$:

$$\begin{aligned} d(\boldsymbol{\mu}) &= \inf_{\boldsymbol{x}\in\mathbb{R}^n} \left(f(\boldsymbol{x}) + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\right) \\ &= -\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu} + \inf_{\boldsymbol{x}\in\mathbb{R}^n} \left(\left[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\mu}\right]^{\mathrm{T}}\boldsymbol{x} + f(x)\right) \\ &= -\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu} - \sup_{\boldsymbol{x}\in\mathbb{R}^n} \left(\left[-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\mu}\right]^{\mathrm{T}}\boldsymbol{x} - f(x)\right) \\ &= -\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu} - f^*(-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\mu}). \end{aligned}$$

Solution for Exercise 2.5: Weak duality

Question a) The function f is convex (it satisfies the function inequality for every pair of points in the domain) on the feasible set of the problem $\{(0, x_2) \mid x_2 > 0\}$. The feasible set is a convex set, hence the problem is convex. Given that the objective function is constant on the problem domain, we have $p^* = f(0) = 1$.

Question b) The Lagrangian of the problem is given by

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \exp(-x_1) + \lambda \frac{x_1^2}{x_2}$$

where $\lambda \in \mathbb{R}$. As a result, the dual function of the problem is defined for every $\lambda \in \mathbb{R}$ as

$$d(\lambda) = \inf_{\boldsymbol{x}\in\mathcal{X}} \left\{\exp(-x_1) + \lambda \frac{x_1^2}{x_2}\right\}.$$

When $\lambda < 0$, we have $d(\lambda) = -\infty$ (taking x_2 to be constant and taking x_1 to $+\infty$). When $\lambda = 0$, we obtain $d(0) = \inf_{\boldsymbol{x}\in\mathcal{X}} \exp(-x_1) = 0$. Finally, when $\lambda > 0$, we get

$$\inf_{x_1\in\mathbb{R}, x_2>0} \exp(-x_1) + \lambda \frac{x_1^2}{x_2} = \inf_{x_1\in\mathbb{R}} \inf_{x_2>0} \exp(-x_1) + \lambda \frac{x_1^2}{x_2} = \inf_{x_1\in\mathbb{R}} \exp(-x_1) = 0$$

hence $d(\lambda) = 0$ in that case as well. Overall, we obtain that

$$d(\lambda) = \begin{cases} -\infty & \text{if } \lambda < 0 \\ 0 & \text{if } \lambda \geq 0, \end{cases}$$

hence $d^* = \sup_{\lambda\geq 0} d(\lambda) = 0$. As a result, the duality gap is $p^* - d^* = 1$, and strong duality does not hold.

Question c) Recall that Slater's condition applies to the relative interior of the problem domain, which is $\mathcal{X} = \{(x_1, x_2) | x_1 = 0, x_2 > 0\}$. Since $\mathrm{ri}(\mathcal{X}) = \mathcal{X}$ (which is clear from the fact that $\mathrm{aff}(\mathcal{X}) = \{(x_1, x_2) | x_1 = 0\}$), to check Slater's condition we would need a point from this domain that satisfies the constraint $\frac{x_1^2}{x_2} \leq 0$ with strict inequality. However, for any point in $\mathcal{X}$, we have $\frac{x_1^2}{x_2} = 0$, and thus Slater's condition does not hold. *NB: This makes sense since Slater's condition would imply strong duality, contradicting the result of the previous question.*

Solution for Exercise 2.6: KKT and optimality conditions

Question a) The function f is a quadratic defined by the positive definite matrix $\boldsymbol{A} = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$, hence it is convex (even strongly convex). Another possibility to prove convexity relies on computing the gradient

$$\forall \boldsymbol{x} \in \mathbb{R}^2, \nabla f(\boldsymbol{x}) = \begin{bmatrix} 4x_1 - 1 \\ 8x_2 - 1 \end{bmatrix}.$$

and checking the characterization of strong convexity for differentiable functions.

Question b) The set $\mathcal{C}$ is a cone since $t\boldsymbol{x} \geq \mathbf{0}$ for any $\boldsymbol{x} \in \mathcal{C}$ and $t > 0$, and it is also convex since $\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \geq \mathbf{0}$ for any $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{C}^2$ and $\alpha \in [0,1]$. As a result, a solution $\boldsymbol{x}^* = (x_1^*, x_2^*)$ of the problem satisfies $x_1^* \geq 0$, $x_2^* \geq 0$ as well as

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{x}^* = 0 \quad \text{and} \quad \forall \boldsymbol{y} \in \mathcal{C}, \nabla f(\boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{y} \geq 0.$$

Putting all conditions together, we obtain

$$\begin{cases} x_1^* \geq 0 \\ x_2^* \geq 0 \\ (4x_1^* - 1)x_1^* + (8x_2^* - 1)x_2^* = 0 \\ \forall y_1 \geq 0, \forall y_2 \geq 0, (4x_1^* - 1)y_1 + (8x_2^* - 1)y_2 \geq 0. \end{cases}$$

Letting $y_1 = 0$ and $y_2 = 0$ in the last condition gives $4x_1^* - 1 \geq 0$ and $8x_2^* - 1 \geq 0$. Combining with the previous two, we obtain

$$(4x_1^* - 1)x_1^* = (8x_2^* - 1)x_2^* = 0.$$

If $x_1^* = 0$, then $4x_1^* - 1 < 0$, contradicting what we had before. Similarly, $x_2^* = 0$ leads to the contradiction $8x_2^* - 1 < 0$. Overall, we must have

$$4x_1^* - 1 = 8x_2^* - 1 = 0,$$

hence $x_1^* = \frac{1}{4}$ and $x_2^* = \frac{1}{8}$.

Question c) Since the constraints defining $\mathcal{C}$ are linear, constraint qualification and strong duality hold for this problem. A solution of the problem is then also a solution of the KKT equations

$$\begin{cases} 4x_1^* - 1 - \lambda_1^* = 0 \\ 8x_2^* - 1 - \lambda_2^* = 0 \\ x_1^* \geq 0 \\ x_2^* \geq 0 \\ \lambda_1^* \geq 0 \\ \lambda_2^* \geq 0 \\ \lambda_1^* x_1^* = 0 \\ \lambda_2^* x_2^* = 0, \end{cases}$$

where $\boldsymbol{\lambda}^* = [\lambda_1^* \ \lambda_2^*]^{\mathrm{T}}$ is a vector of dual variables, and the first two conditions correspond to the gradient of the Lagrangian function

$$\nabla_x \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \nabla f(\boldsymbol{x}) - \boldsymbol{\lambda}.$$

Solving this system leads to $x_1^* = \frac{1}{4}$, $x_2^* = \frac{1}{8}$ en $\lambda_1^* = \lambda_2^* = 0$, recovering the solution computed in question b).