Exercises on Chapter 2: Duality

Mathematics of Data Science, M1 IDD

Fall 2025 (V2, October 20)

Dauphine | PSL
UNIVERSITÉ PARIS

Formulations of optimization problems

Exercise 2.1: Chebyshev approximation

a) Write the epigraph reformulation of the problem

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

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 of a linear objective function and linear constraints.

c) More generally, consider the Chebyshev approximation problem

$$\operatorname*{minimize}_{\boldsymbol{y} \in \mathbb{R}^m} \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|_\infty \quad \text{where } \boldsymbol{A} \in \mathbb{R}^{n \times m} \text{ and } \boldsymbol{b} \in \mathbb{R}^n. \tag{1}$$

Find a reformulation of problem (1) as a linear program with the same optimal value.

d) Consider now a function $\boldsymbol{\phi} : \mathbb{R}^m \to \mathbb{R}^n$ 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

$$\operatorname*{minimize}_{\boldsymbol{y} \in \mathbb{R}^m} \|\boldsymbol{\phi}(\boldsymbol{y})\|_\infty$$

as a convex optimization problem in standard form.

e) Show that the problem

$$\operatorname*{minimize}_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{x}\|_1 = \sum_{i=1}^n |x_i| \tag{2}$$

can be reformulated as a linear program with $2n$ nonnegative variables.
Hint: use that $|t| = t^+ + t^-$ *with* $t^+ = \max(t, 0)$ *and* $t^- = \max(-t, 0)$.

f) Finally, consider the robust linear regression problem

$$\operatorname*{minimize}_{\boldsymbol{y} \in \mathbb{R}^m} \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|_1 \quad \text{where } \boldsymbol{A} \in \mathbb{R}^{n \times m} \text{ and } \boldsymbol{b} \in \mathbb{R}^n. \tag{3}$$

Find a reformulation of problem (3) as a linear program.

Exercise 2.2: Equivalent reformulation

Given any $\boldsymbol{x} \in \mathbb{R}^n$, we let $\|\boldsymbol{x}\|_0$ denote the number of nonzero components of $\boldsymbol{x}$, and we denote the ℓ_4 norm of $\boldsymbol{x}$ by $\|\boldsymbol{x}\|_4 := \left(\sum_{i=1}^n x_i^4\right)^{1/4}$. Note that by equivalence of norms, we have $\|\boldsymbol{x}\|_4 \leq \|\boldsymbol{x}\|_2 \leq n^{\frac{1}{4}}\|\boldsymbol{x}\|_4$, where $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ denotes the Euclidean norm.

a) Find the solutions and optimal value of

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}} \, \|\boldsymbol{x}\|_0 \text{ subject to } \ \|\boldsymbol{x}\|_2 = 1 \tag{4}$$

b) Find the solutions and optimal value of

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{maximize}} \, \|\boldsymbol{x}\|_4 \text{ subject to } \ \|\boldsymbol{x}\|_2 = 1. \tag{5}$$

c) Explain in which sense problems (4) and (5) are equivalent reformulations.

Exercise 2.3: 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 2.4: 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. *Hint: Use that f is strictly convex.*
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 $\text{minimize}_{\boldsymbol{x} \in \mathbb{R}^3} f(\boldsymbol{x})$) ?

Duality

Exercise 2.5: 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

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

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 2.6: Weak duality

Consider the problem

$$\begin{array}{ll} \mathrm{minimize}_{\boldsymbol{x} \in \mathbb{R}^2} & f(\boldsymbol{x}) \\ \text{subject to} & \frac{x_1^2}{x_2} \leq 0, \end{array}, \quad \text{where} \quad f(\boldsymbol{x}) = \begin{cases} \exp(-x_1) & \text{if } x_2 > 0 \\ \infty & \text{otherwise.} \end{cases} \tag{6}$$

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 (6), give a formula for its solution λ^* and its optimal value d^*. What is then the duality gap?

Exercise 2.7: Trust-region subproblem

Let $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ be a symmetric matrix, $\boldsymbol{g} \in \mathbb{R}^n$ and $\delta > 0$. We consider the so-called *trust-region subproblem* defined by

$$\begin{cases} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & q(\boldsymbol{x}) := \boldsymbol{g}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \le \delta, \end{cases} \tag{7}$$

where $\mathrm{dom}(q) = \mathbb{R}^n$.

a) Find a diagonal matrix $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ and a vector $\boldsymbol{c} \in \mathbb{R}^n$ such that problem (7) is equivalent to

$$\begin{cases} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & f(\boldsymbol{x}) := \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \le 1, \end{cases} \tag{8}$$

where $\mathrm{dom}(f) = \mathbb{R}^n$.

b) Write the primal and dual problems corresponding to (8), and justify that the primal problem is equivalent to (8).

c) Suppose now that $\boldsymbol{D}$ is not positive semidefinite, i.e. that there exists $i \in \{1, \dots, n\}$ such that $[\boldsymbol{D}]_{ii} < 0$. In that case, it can be shown that problem (8) is equivalent to

$$\begin{cases} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1, \end{cases} \tag{9}$$

i) Justify that problem (9) has a solution.

ii) Show that for any $\lambda > 0$,

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ f(\boldsymbol{x}) \middle| \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1 \right\} = -\lambda + \inf_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}(\boldsymbol{D} + \lambda\boldsymbol{I})\boldsymbol{x} \middle| \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1 \right\}.$$

iii) For any $\lambda \ge \max_{1\le i\le n} |[\boldsymbol{D}]_{ii}|$, show that strong duality holds for problem

$$\begin{array}{ll} \text{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} & f_\lambda(\boldsymbol{x}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}(\boldsymbol{D} + \lambda\boldsymbol{I})\boldsymbol{x} - \lambda \\ \text{s.t.} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1. \end{array} \tag{10}$$

How is strong duality expressed using the Lagrangian of problem (10)?

iv) Use the result from the previous question to conclude that strong duality holds between problem (9) and its dual.