

Exercise 2: Duality and Fitting Problems

Prof. Dr. Moritz Diehl, Dimitris Kouzoupis, Florian Messerer, Yizhen Wang

1. Lagrange duality and dual problems:

(a) Consider the following *logarithmic barrier* problem,

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x - \sum_{j=1}^n \log x_j \\ \text{s.t.} \quad & a^T x = b, \end{aligned}$$

where $a, c \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

Remark 1: Problems using a logarithmic barrier as the one above will be at the core of interior point methods that we will analyze later in this course.

Remark 2: $\log x_j$ is only defined for $x_j \in \mathbb{R}_{++}$. For simplicity, and without discussing this further here, we will assume that $-\log x_j$ takes the value $+\infty$ whenever $x_j \in \mathbb{R}_-$.

- i. Derive the explicit form of the dual of this problem.
- ii. Does strong duality hold?

(b) Consider the following *mixed-integer quadratic program* (MIQP):

$$\begin{aligned} \min_{x \in \{0,1\}^n} \quad & x^T Q x + q^T x \\ \text{s.t.} \quad & A x \geq b, \end{aligned}$$

where $Q \in \mathbb{R}^{n \times n}$, $q \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. where the optimization variables x_i are restricted to take values in $\{0, 1\}$. Solving mixed-integer problems is in general a challenging task, thus it is common practice to exploit continuous reformulations as the following:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T Q x + q^T x \\ \text{s.t.} \quad & A x \geq b \\ & x_i(1 - x_i) = 0 \quad i = 1, \dots, n. \end{aligned}$$

- i. Is this reformulation convex?
- ii. A lower bound to the optimal solution can be computed by solving the (convex) dual problem (not required here). Derive the explicit form of the dual of the continuous reformulation.

2. **Regularized linear least squares:** Given a matrix $J \in \mathbb{R}^{m \times n}$, a symmetric positive definite matrix $Q \succ 0$, a vector of measurements $\eta \in \mathbb{R}^m$ and a point $\bar{x} \in \mathbb{R}^n$, compute the limit:

$$\lim_{\substack{\alpha \rightarrow 0 \\ \alpha > 0}} \arg \min_x \frac{1}{2} \|\eta - Jx\|_2^2 + \frac{\alpha}{2} (x - \bar{x})^\top Q (x - \bar{x}). \quad (1)$$

Hint: Use matrix square root and Lemma 6.1 from the lecture notes.

3. **Linear L_2 fitting:** Assume we have modeled the dependency of some output $y \in \mathbb{R}$ on some input $x \in \mathbb{R}$ as the linear model $y = ax + b$ with parameters $a, b \in \mathbb{R}$. The value of these parameters is unknown, but we have a data set of N noisy measurements (x_i, \tilde{y}_i) , $i = 1, \dots, N$. These measurements are obtained as $\tilde{y}_i = ax_i + b + \eta_i$, where η_i is noise drawn from a normal distribution with zero mean and variance one, $\eta_i \sim \mathcal{N}(0, 1)$.

One way of finding an estimate of the parameter values is to minimize a least-squares loss of the residuals $ax_i + b - \tilde{y}_i$, which can be formulated as the optimization problem

$$\min_{a, b \in \mathbb{R}} \sum_{i=1}^N \frac{1}{2} (ax_i + b - \tilde{y}_i)^2 = \min_{a, b} \frac{1}{2} \left\| J \begin{bmatrix} a \\ b \end{bmatrix} - \tilde{y} \right\|_2^2, \quad (2)$$

where $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_N)$ and it will be part of the exercise to define J . As discussed in the lecture, the optimal solution of (2) can be calculated explicitly by solving the linear system

$$J^\top J \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = J^\top \tilde{y}, \quad (3)$$

where \hat{a}, \hat{b} are the resulting estimates of the parameter values.

- Define J by writing it down on paper.
- Generate the problem data. Take $N = 30$ and generate $x = (x_1, \dots, x_N)$ as N equally spaced points in the interval $[0, 5]$ and, for $i = 1, \dots, N$, generate the measurements as $\tilde{y}_i = 3x_i + 4 + \eta_i$, where η_i is sampled from the normal distribution $\mathcal{N}(0, 1)$. Plot the results.
Hint: look up the `linspace` and `randn` commands, e.g., via NumPy documentation (Python) / using `help` or `doc` command (MATLAB). If you want a reproducible 'random' sequence, you can use `rng`.
- Calculate the estimates \hat{a}, \hat{b} in MATLAB using Equation (3) and plot the obtained line in the same graph as the measurements.
- Introduce 3 outliers in \tilde{y} by replacing arbitrary measurements and plot the new fitted line in your plot.

You will need the measurements \tilde{y} (both with and without outliers) and the matrix J for the next task.

4. **Linear L_1 fitting:** In this task we want to fit a line to the same set of measurements, but we use a different cost function:

$$\min_{a,b \in \mathbb{R}} \sum_{i=1}^N |(ax_i + b - \tilde{y}_i)|. \quad (4)$$

- (a) Problem (4) is not differentiable. Find an (equivalent) smooth reformulation.
Hint 1: Introduce slack variables $s_1, \dots, s_N \in \mathbb{R}$ as additional decision variables.
Hint 2: The resulting problem will be a Linear Program (LP).
- (b) The result of the previous task is a LP. In order to solve it with `linprog`, the native LP solver of MATLAB, we need to bring it to the form:

$$\min_{z \in \mathbb{R}^n} f^T z \quad (5a)$$

$$\text{s.t. } Az \leq b \quad (5b)$$

$$Cz = d \quad (5c)$$

$$l_z \leq z \leq u_z, \quad (5d)$$

Define matrices A, C and vectors f, b, d, l_z, u_z by writing them down on paper. You may not need all of these. In this case you can define them as 'empty'. Order your variables as $z = (a, b, s_1, \dots, s_N)$. Use matrix J from the previous exercise to define A .

- (c) Solve the problem with `linprog` (SciPy / MATLAB). Use the measurements \tilde{y} from the previous exercise (both with and without outliers) and plot the results against those of the L2 fitting. Which norm performs better?
- (d) Solve the problem resulting from task 4a with CasADi and compare the results.