Exercises for Lecture Course on Numerical Optimization (NUMOPT)
Albert-Ludwigs-Universität Freiburg – Summer Semester 2026

Exercise 2: Convexity, Duality and Fitting problems

Léo Simpson, Prof. Dr. Moritz Diehl, with contributions from previous teaching assistants

The solutions for these exercises will be given and discussed during the exercise session on May the 12th, 2026.
To receive feedback on your solutions, please hand it in during the exercise session on May the 12th, 2026, or by e-mail to leo.simpson@imtek.uni-freiburg.de before the same date.

I Convex optimization

In this part we learn how to recognize convex sets and functions. Moreover we revisit the hanging chain problem from the previous exercise sheet and add different types of constraints.

I.1 Convex sets

Determine whether the following sets are convex or not. Justify your answers.

1. A halfspace, i.e., a set of the form:
$$\Omega = \{x \in \mathbb{R}^n | a^\top x \leq b\}$$
for some $a \in \mathbb{R}^n$, and $b \in \mathbb{R}$.
2. A wedge, i.e., a set of the form
$$\Omega = \{x \in \mathbb{R}^n \mid a_1^\top x \leq b_1, a_2^\top x \leq b_2\}$$
for some $a_1, a_2 \in \mathbb{R}^n$, and $b_1, b_2 \in \mathbb{R}$.
3. The set of points closer to a given point than to a given set:
$$\Omega = \{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq \|x - y\|_2 \ \forall y \in \mathcal{S}\}$$
where $x_0 \in \mathbb{R}^n$ is a vector and $\mathcal{S} \subseteq \mathbb{R}^n$ is a set of points.
4. The set of points closer to one set than to another:
$$\Omega = \{x \in \mathbb{R}^n \mid \|x - z\|_2 \leq \|x - y\|_2 \ \forall y \in \mathcal{S}, \forall z \in \mathcal{T}\}$$
where $x_0 \in \mathbb{R}^n$ is a vector and $\mathcal{S}, \mathcal{T} \subseteq \mathbb{R}^n$ are sets of points.

I.2 Convex functions

Determine whether the following functions are convex or not. Justify your answers.

1. The function $f(x) = -\log(x)$ on $\mathbb{R}_{++}$.
2. The function $f(x) = x^3$ on $\mathbb{R}$.
3. The function $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbb{R}^2_{++}$.
4. The function $f(x_1, x_2) = x_1/x_2$ on $\mathbb{R}^2_{++}$.

I.3 Hanging chain, revisited

Recall the hanging chain problem from the previous exercise sheet.

1. What would happen if you add, instead of the linear ground constraints, the nonlinear ground constraints $z_i \geq -0.2 + 0.1y_i^2$, for $i = 2, \ldots, N-1$ to your problem? What type of optimization problem is the resulting problem? Is it convex?
2. What would happen if you add instead the nonlinear ground constraints $z_i \geq -y_i^2$, $i = 2, \ldots, N-1$? Do you expect this optimization problem to be convex?
3. Do you expect these two problems to have several local minima? Why?
4. Now solve both variations using CasADi and plot the results (both chain and constraints).

   <u>Remark:</u> *For readibility of your code, we suggest to introduce a variable* `TYPE_OF_CONTRAINTS` *to easily switch between the different constraints that are proposed in the serie of hanging chain exercises.*
5. Find numerically two different local minimum for the second variation, by initializing the solver with different initial values.

   <u>Hint:</u> *In CasADi, you can provide an initial value* `x0` *for variable* `x` *via*

   ```
   opti.set_initial(x,x0)
   ```

II Minimum of a coercive function

We say that a function $f : \mathbb{R}^n \to \mathbb{R}$ is coercive when $\lim_{\|x\| \to \infty} f(x) = +\infty$, i.e. $\forall\, M,\ \exists\, R$ such that for all $\|x\| > R$, $f(x) > M$.

Prove that for a function $f : \mathbb{R}^n \to \mathbb{R}$ continuous and coercive, the unconstrained optimization problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

always has (at least) one global minimum point.
<u>Hint:</u> *Use the Weierstrass Theorem.*

III Lagrange duality and dual problems

III.1 Logarithmic barrier

Consider the following *logarithmic barrier* problem,

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\top x - \sum_{j=1}^{n} \log x_j \\ \text{subject to} & a^\top x = b \end{array} \tag{1}$$

where $a, c \in \mathbb{R}^n$ and $b \in \mathbb{R}$.
Remark: *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: *Even though* $\log x$ *is only defined for* $x > 0$*, for some reasons, it is not really a problem here. Consider that we extend its definition with* $\log x = -\infty$ *for* $x \leq 0$.

1. Derive the explicit form of the dual of this problem.

2. What connection is there between the optimal value of the dual problem, and the optimal value of the initial problem (1)?

   Hint: *You are asked to find a "strong" connection.*

III.2 Linear programming

Consider the following *integer linear program* (ILP):

$$\begin{array}{ll} \underset{x \in \{0,1\}^n}{\text{minimize}} & -c^\top x \\ \text{subject to} & Ax \geq b \end{array} \tag{2}$$

where $c \in \mathbb{R}^n$, $q \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The optimization variables x_i are restricted to take values in $\{0, 1\}$. Solving such problems is in general a challenging task. A common practice is to reformulate the binary constraint.

1. Reformulate the mixed-integer program (2) into a nonlinear program in the standard form:

   $$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & g(x) = 0, \\ & h(x) \geq 0 \end{array} \tag{3}$$

   where the functions $f(\cdot), g(\cdot)$, and $h(\cdot)$ are continuous functions that you have to provide.

   Hint: *What are the* $x_i \in \mathbb{R}$ *such that* $x_i(1 - x_i) = 0$*?*

2. Is this reformulation convex?

3. Write the Lagrangian of the reformulation.

4. Derive the explicit form of the dual of your reformulation.

5. Is the dual problem convex?

6. Using what you have found so far, propose a procedure to compute a lower bound to the ILP (2).

   Hint: *we do not ask you to compute this lower bound! only explain how to compute it numerically...*

IV Fitting problems

IV.1 Affine L_2 fitting

We consider the linear regression task: we want to predict some output $y \in \mathbb{R}$ using some input $x \in \mathbb{R}$. Some input-output data $(x_i, y_i) \in \mathbb{R}^2$ are available for $i = 1, \dots, N$, and we want to use them to estimate an affine model, i.e.

$$y_i \approx a x_i + b, \tag{4}$$

where the parameters a and b have to be estimated. For this purpose, we solve the following least-square problem:

$$\underset{a,b \in \mathbb{R}}{\text{minimize}} \quad \frac{1}{2}\sum_{i=1}^{N}(a x_i + b - y_i)^2 \tag{5}$$

1. Rewrite the least square problem (5) in matrix form:

$$\underset{a,b \in \mathbb{R}}{\text{minimize}} \quad \frac{1}{2}\left\| J \begin{bmatrix} a \\ b \end{bmatrix} - y \right\|_2^2 \tag{6}$$

   (you have to the matrix J and the vector y).

2. Assume that J is full column-rank. Give an explicit form of the solution $(\hat{a}, \hat{b}) \in \mathbb{R}^2$ to (5).

   Hint: *J is full column-rank if and only if $J^\top J$ is invertible.*

3. Create a Python script that generates the data points (x_i, y_i). For this, you will choose:

   - $N = 30$
   - $x_1, \dots, x_N$ to be N equally spaced points in the interval $[0, 5]$,
   - y_i is an actually affine function of the input, but corrupted with random noise. More precisely: $y_i = 3x_i + 4 + \eta_i$, where η_i is sampled from the normal distribution $\mathcal{N}(0, 1)$.

   Plot the data points.

   Hint: *look up the `numpy.linspace` command and the function `numpy.random.Generator.normal` in NumPy documentation*

4. Calculate the estimates $\hat{a}, \hat{b}$ in Python using what you have found in question 2. Plot the corresponding line in the same plot as the data points.

5. Introduce 3 outliers in y by replacing arbitrary measurements with some nonsense data and plot the new fitted line in your plot. Comment the result.

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

IV.2 Affine L_1 fitting

In this exercise, like in the previous one, we will learn an affine model from an input-output data set. However, instead of using the least squares formulation, we will try another cost function, *the least absolute deviations*:

$$\underset{a, b \in \mathbb{R}}{\text{minimize}} \quad \sum_{i=1}^{N} |ax_i + b - y_i| \tag{7}$$

1. Problem (7) is not differentiable. Find an (equivalent) smooth reformulation.

   Hint: *Introduce slack variables $s_1, \ldots, s_N \in \mathbb{R}$ as additional decision variables.*

   Hint: *Note that $s \geq |x|$ if and only if $s \pm x \geq 0$.*

   Hint: *The resulting problem will be a Linear Program (LP).*

2. Solve the optimization problem you have just formulated using CasADi for the data generated in the previous exercise (both with and without outliers).

   Plot the results against those of the L2 fitting problem.

3. Which norm performs better? Interpret the results.

IV.3 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 $y \in \mathbb{R}^m$ and a point, and a penalization parameter $\alpha \geq 0$, we consider the following regularized least squares problem:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} \|y - Jx\|_2^2 + \frac{\alpha}{2} x^\top Q x \tag{8}$$

1. Does the problem (8) always have a unique solution? Justify.

   Hint: *Distinguish the casew where $\alpha = 0$ from the case where $\alpha > 0$. Also distinguish the case where J is full column rank and from the case where it is not.*

2. We denote by $x^\star(\alpha)$ the solution of the initial problem (8) for a given $\alpha > 0$. Write $x^\star(\alpha)$ explicitly.

3. Transform the problem into the standard regularized least-squares form:

   $$\underset{\tilde{x} \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} \left\| y - \tilde{J}\tilde{x} \right\|_2^2 + \frac{\alpha}{2} \|\tilde{x}\|^2 \tag{9}$$

   where $\tilde{J}$ should be written explicitly, and the connection between $\tilde{x}$ and x should be given.

   Hint: *Use matrix square-root*

4. Prove that the solution $x^\star(\alpha)$ converges to some vector $x^\star \in \mathbb{R}^n$ when $\alpha \to 0$ (even when J is not full column-rank). Express $x^\star$ explicitly using $\tilde{J}$.

   Hint: *Do not use the explicit formula for $x^\star(\alpha)$. Instead, use the result from last question and Lemma 6.1 from the lecture notes about matrix pseudo-inverse.*

5. Prove that $x^\star$ is a solution of the non-regularized least squares problem

   $$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} \|y - Jx\|_2^2 \tag{10}$$