Homework accompanying the lecture "Basics in Applied Mathematics"

Homework 11

Hand in: Tuesday, 14.01.2025, after the lecture in the mailbox at the Math Institut
(Don't forget to put your name on your homework.
Please hand in your solutions in groups of two.)

Note: We recall that the notation "$\underset{x\in\mathcal{X}}{\text{minimize}}\ f(x)$" refers to an *optimization problem*, whereas the notation "$\min_{x\in\mathcal{X}} f(x)$" refers to the *optimal value* of the corresponding optimization problem.

Exercise 1 (1D optimization problems; 5 points)

In this exercise, we study the solutions of some optimization problems in the following form:

$$\underset{x\in\mathbb{R}}{\text{minimize}} \quad f(x) \tag{1}$$

For each of the following functions $f(x)$, find *all* local minima, and specify which ones are global minima (when they exist) for the problem (1)

a) $f(x) := x^2 + 4x + 10$

b) $f(x) = x^3 - 3x^2 + 1$

c) $f(x) = x^2 + \frac{1}{x^2+1}$

d) $f(x) = |x| + 2|x-1|$

e) $f(x) = e^{-x^2}$

Hint: For smooth functions, use first- and second-order optimality conditions to find and classify stationary points.
More precisely, when f is smooth, look at the points where $f'(x) = 0$ holds, and check if $f''(x) > 0$ also holds for these points.

Exercise 2 (2D optimization ; 2 points)

Let $\mathcal{X}$ and $\mathcal{Y}$ be two compact sets. Let $f : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ be a continuous function.
Show the following properties (and prove the existence of each minimum):

a)

$$\min_{x\in\mathcal{X},y\in\mathcal{Y}} f(x,y) = \min_{x\in\mathcal{X}} \left(\min_{y\in\mathcal{Y}} f(x,y)\right)$$

Note that on the right-hand side of the equation, $x \to \min_{y\in\mathcal{Y}} f(x,y)$ is viewed as a function of x. Note: You can use without proof that the function $x \to \min_{y\in\mathcal{Y}} f(x,y)$ is also continuous with respect to x.

Hint: Use the compactness of the sets $\mathcal{X}$ and $\mathcal{Y}$ to prove the existence of each minimum.

b)

$$\max_{y\in\mathcal{Y}} \left(\min_{x\in\mathcal{X}} f(x,y)\right) \leq \min_{x\in\mathcal{X}} \left(\max_{y\in\mathcal{Y}} f(x,y)\right)$$

Note: You can use without proof that the functions $y \to \min_{x \in \mathcal{X}} f(x,y)$ and $x \to \max_{y \in \mathcal{Y}} f(x,y)$ are continuous with respect to y and x, respectively.

Hint 1: For maximization problems, you can use the same existence theorem as in the first part.

Hint 2: Start with the following inequality for all $\bar{x} \in \mathcal{X}$ and $y \in \mathcal{Y}$:

$$\min_{x \in \mathcal{X}} f(x,y) \leq f(\bar{x},y)$$

then maximize or minimize twice on both sides of the inequality with respect to some variables...

Exercise 3 (A function approximation problem ; 7 points)

In this exercise, we will study a function approximation problem. More precisely, a function $g : \mathbb{R} \to \mathbb{R}$ is unknown, but we have access to some of its values: $y_j = g(t_j)$ for some points $t_1 < t_2 < \cdots < t_m$.

The goal is to approximate $g(\cdot)$ by a function $\Phi(\cdot; x)$ of the following form:

$$\Phi(t; x) = \sum_{i=1}^{n} \varphi_i(t) x_i \tag{2}$$

where $\varphi_1(\cdot), \dots, \varphi_n(\cdot)$ are known functions and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ are the parameters to be determined.

We formulate the approximation problem as a least-square problem:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \sum_{j=1}^{m} \big(y_j - \Phi(t_j; x)\big)^2 \tag{3}$$

a) Show that if the function $g(\cdot)$ is already in the form: $g(t) = \Phi(t; x^\star)$ for some $x^\star \in \mathbb{R}^n$, then the point $x^\star$ is a solution of the problem (3).

b) Show that the problem (3) can be written in the following general form:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \|y - Ax\|^2 \tag{4}$$

where $y = (y_1, \dots, y_m) \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ is a matrix that you need to specify.

c) Show that the problem (4) can be reformulated in the general QP form:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} x^\top Q x - c^\top x + r \tag{5}$$

where $Q \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ and $r \in \mathbb{R}$ are matrices/vectors/values that you need to specify.

d) Assume that the matrix $A^\top A$ is invertible. Then show that the problem (4) has a unique solution, and specify it as a function of A and y.

e) Discuss a necessary condition on the number of samples m, and the number of parameters n for $A^\top A$ to be invertible. Also provide an intuitive explanation of this condition.

f) Consider the following example of basis functions:

$$\varphi_1(t) = 1, \quad \varphi_2(t) = t, \quad \varphi_3(t) = 2t - 3$$

Is it possible that the problem (3) has a unique solution in that case? Explain why.

g) Write a Python script to find a solution $\hat{x}$ of the problem (3) for the following data:

$$t_1 = 0.1, \quad t_2 = 0.2, \quad \dots \quad t_{10} = 1.0$$
$$g(t) := \begin{cases} t & \text{if } t \leq 0.5 \\ 1 - t & \text{if } t > 0.5 \end{cases}$$
$$\varphi_1(t) = 1, \quad \varphi_2(t) = \sin(\pi t), \quad \varphi_3(t) = \sin(2\pi t)$$

Note: Do not forget to describe the computations that you implement with some comments to facilitate the understanding of your code.

h) (Optional) Visualize the function $g(\cdot)$, data points (t_j, y_j), and the approximation function $\Phi(\cdot, \hat{x})$.

Exercise 4 (A variance estimation problem ; 4 points)

In this exercise, we will study a famous estimation method called the *Maximum Likelihood Estimation* problem, in the particular case of a variance estimation problem.
Here, the measurements $y_1, \dots, y_m$ are measurements, and are generated through the stochastic model $y_j \sim \mathcal{N}(0, \sigma^2)$, i.e. the probability density function of y_j is given by:

$$p(y_j) = f(y_j; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y_j^2}{2\sigma^2}\right)$$

In order to to estimate the parameter σ from the measurements $y_1, \dots, y_m$, one classical approcah is to solve the following problem:

$$\underset{\sigma \in \mathbb{R}_{>0}}{\text{maximize}} \quad \prod_{j=1}^{m} f(y_j; \sigma^2) \tag{6}$$

In this exercise, we will assume that at least one measurement is such that $y_j \neq 0$.

a) Show that the problem (6) can be reformulated as the following *minimization* problem:

$$\underset{\sigma \in \mathbb{R}_{>0}}{\text{minimize}} \quad \sum_{j=1}^{m} \frac{y_j^2}{\sigma^2} + \log(\sigma^2) \tag{7}$$

Hint: Maximizing a function $f(x)$ is equivalent to minimizing the funcion $-2\log(f(x)) + \text{cst}$ for some constant cst.

b) Let $\bar{v} > 0$ be a positive scalar value. Find the solution(s) of the following optimization problem:

$$\underset{v \in \mathbb{R}_{>0}}{\text{minimize}} \quad \frac{\bar{v}}{v} + \log(v)$$

Hint: Sketch a graph of the function $v \to \frac{\bar{v}}{v} + \log(v)$ by looking at its derivative.

c) Use the two previous questions to find the solution $\hat{\sigma}$ of the problem (6).

d) Write a Python script that does the following things:

- Generate the data $y_1, \dots, y_{100}$ according to the random distribution $y_j \sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 2$.

- Compute the estimate $\hat{\sigma}$ using the formula derived in the previous questions. Compare it to the true value $\sigma = 2$.
- Plot the graph of the objective function of the problem (6) as a function of σ, and highlight the maximum.

Note: Do not forget to describe the computations that you implement with some comments to facilitate the understanding of your code.