Assignment 2

Nonlinear Optimization WS2024/25

Alexander Falk, falk@tugraz.at
Clemens Krenn, clemens.krenn@student.tugraz.at
Stefanie Stoppacher, stoppacher@student.tugraz.at

November 19, 2024

Deadline: Dec 10^{th}, 2024 at 16:00

Submission: Upload your report and your implementation to the TeachCenter. Please use the provided framework-file for your implementation. Make sure that the total size of your submission does not exceed 50MB. Include **all** of the following files in your submission:

- `report.pdf`: This file includes your notes for the individual tasks. Keep in mind that we must be able to follow your *calculation process*. Thus, it is not sufficient to only present the final results. Please submit your findings in a compiled LaTeX document.
- `main.py`: This file includes your `python` code to solve the different tasks. Please only change the marked code sections. Also please follow the instructions defined in `main.py`.
- `figures.pdf`: This file is generated by running `main.py`. It includes a plot of all mandatory figures on separate pdf pages. Hence, you do not have to embed the plots in your report.

1 Lagrange Multiplier Problem (15 P.)

Given the following constrained optimization problems over $\mathbf{x} = (x_1, x_2)^\top$, compute the optimal solution using Lagrangian formulations.

a)
$$\min_{\mathbf{x}} \; x_2 - x_1 \quad \text{s.t.} \quad x_1 \geq 4x_2, \quad x_2 = \frac{1}{10}x_1^2 - 3$$

b)
$$\min_{\mathbf{x}} \; \|\mathbf{x}\|_2^2 \quad \text{s.t.} \quad x_2 \geq 3 - x_1, \quad x_2 \geq 2$$

c)
$$\min_{\mathbf{x}} \; (x_1 - 1)^2 + x_1 x_2^2 - 2 \quad \text{s.t.} \quad x_1^2 + x_2^2 \leq 4$$

For each example - **In Python**:

1. Plot the level sets of the above functions using a (filled) contour plot[1].
2. Draw the equality-/inequality-constraints on top[2].
3. Add markers to all points which fulfill the KKT-conditions (computed with Pen & Paper). Highlight the overall optimal solution with a point in a different color.

For each example - **With Pen & Paper**:

4. Formulate the Lagrangian equation. As shown in the exercise.

[1] https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.contourf.html
[2] https://matplotlib.org/stable/gallery/images_contours_and_fields/contours_in_optimization_demo.html

5. Formulate the KKT-conditions.
6. Discuss every case with respect to the given constraints. Compute all possible solutions and identify valid ones.
7. Show which solution is the optimal one.

2 Augmented Lagrangian (10 P.)

Consider an optimization problem with linear constraints of the form

$$\min_{\mathbf{x}} \; f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \tag{2.1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and the associated Lagrangian function is given as

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top(\mathbf{A}\mathbf{x} - \mathbf{b}). \tag{2.2}$$

Recall from the lecture that the Lagrangian function provides a lower bound to the original constrained optimization problem which can be written in the form of a min-max problem

$$\min_{\mathbf{x}} \big\{ \max_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \big\}. \tag{2.3}$$

A naive approach to arrive at a saddle point of the Lagrangian is to start with some initial $(\mathbf{x}^0, \boldsymbol{\lambda}^0)$ and perform gradient descent in $\mathbf{x}$ and gradient ascent in $\boldsymbol{\lambda}$ as

$$\begin{cases} \mathbf{x}^{k+1} = \mathbf{x}^k - \tau \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^k, \boldsymbol{\lambda}^k), \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \tau \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{x}^{k+1}, \boldsymbol{\lambda}^k), \end{cases} \qquad \text{with } \tau > 0, \; k = 0, 1, 2, \ldots \tag{2.4}$$

Another way to tackle Eq. (2.1) is to introduce the augmented Lagrangian function with parameter $\mu > 0$ as

$$\mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) + \frac{\mu}{2} \left\| \mathbf{A}\mathbf{x} - \mathbf{b} \right\|_2^2. \tag{2.5}$$

This modification of the original Lagrangian is valid as at optimality we have that $\mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0}$. The advantage of the augmented Lagrangian is that we can come up with a better update rule for $\mathbf{x}$ and $\boldsymbol{\lambda}$ as

$$\begin{cases} \mathbf{x}^{k+1} \in \operatorname{argmin}_{\mathbf{x}} \mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\lambda}^k), \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \mu(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{b}). \end{cases} \tag{2.6}$$

Tasks: Given the optimization problem over $\mathbf{x} = (x_1, x_2)^\top$

$$\min_{\mathbf{x}} f(\mathbf{x}) := (x_1 - 1)^2 - x_1 x_2 \quad \text{s.t.} \quad -x_1 + 4 = x_2$$

With Pen & Paper:

1. Identify $\mathbf{A}$ and $\mathbf{b}$ and state the Lagrangian $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ and the augmented Lagrangian $\mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\lambda})$ for the given problem.
2. State the KKT-conditions with $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$. Solve for $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$.
3. Come up with the update rule for $\mathbf{x}$ in procedure Eq. (2.6).
4. Also state the KKT-conditions for $\mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\lambda})$ and justify the update rule for $\boldsymbol{\lambda}$ in Eq. (2.6) by comparing with the original KKT-conditions.

In Python:

4. Plot the level sets of the above function using a filled contour plot[3] and draw the constraint on top.
5. Add a marker to the optimal solution which fulfills the original KKT-conditions (computed with Pen & Paper).
6. Implement both procedures in Eqs. (2.4) and (2.6) and compute the optimal solution iteratively. Test procedure Eq. (2.4) also with the augmented Lagrangian. Choose the same $(\mathbf{x}^0, \boldsymbol{\lambda}^0)$ for each method. Pick suitable values for τ and μ. Also choose an adequate number of iterations K.

[3] `https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.contourf.html`

7. Plot the points $\mathbf{x}_k$ over the iterations.
8. Plot the value of the used Lagrangian over the iterations for the three methods.
9. Discuss the convergence behavior of the three methods by plotting $\|\mathbf{x}^k - \mathbf{x}^*\|$ over the iterations k. Do you observe any problems?
10. Discuss the effects of picking different values for τ and μ. For the discussion on μ, you may also consider the definitness of $\nabla^2_{\mathbf{x}} \mathcal{L}_\mu(\mathbf{x}, \boldsymbol{\lambda})$.