Exercises for Lecture Course on Numerical Optimization (NUMOPT)
Albert-Ludwigs-Universität Freiburg – Summer Semester 2025

Exercise 5: Exam Type questions and last chapters

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 July 1st. To receive feedback on your solutions, please hand it in during the exercise session on July 1st, or by e-mail to leo.simpson@imtek.uni-freiburg.de before the same date.

I Hanging chain, the last episode: the Interior Point Method

Consider once again the hanging chain problem, with zero rest length, and ground constraints:

$$\begin{aligned} \underset{y \in \mathbb{R}^{N-1},\, z \in \mathbb{R}^{N-1}}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=0}^{N-1} D\left((y_i - y_{i+1})^2 + (z_i - z_{i+1})^2\right) + g\sum_{i=0}^{N} m\, z_i \\ \text{subject to} \quad & z_i \geq 0.5 \qquad \text{for } i = 1, \dots, N-1, \\ & z_i \geq 0.5 + 0.1 y_i \quad \text{for } i = 1, \dots, N-1, \\ & z_i \geq -1 - y_i \quad \text{for } i = 1, \dots, N-1 \end{aligned} \tag{1}$$

with the same parameters as before. In this exercise, we will treat the problem in its standard form:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) = \frac{1}{2} x^\top Q x + c^\top x \\ \text{subject to} \quad & a_j^\top x + b_j \geq 0 \quad \text{for } j = 1, \dots, m \end{aligned} \tag{2}$$

where $x = \begin{bmatrix} y & z \end{bmatrix}^\top$ is the decision variable. You do not need to compute Q, c, a_j and b_j explicitly, this is already implemented for you in the file `hanging_chain_ip_matrices.py`.
Here, we will implement an interior point method to solve this problem.

1. What type of problem is (2)?

   Solution: The problem is a convex quadratic program with linear constraints. The objective function is convex, and the constraints are linear.

2. A popular approach to solve this problem is *the interior point method*. In this algorithm, we choose a decreasing sequence of barrier parameters $\tau^{[k]}$, solve iteratively the following problems:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f^{[k]}(x) := f(x) - \tau^{[k]} \sum_{j=1}^{m} \log\left(a_j^\top x + b_j\right) \tag{3}$$

   Are these problems convex?

   Solution: Yes, these problems are convex. Indeed, the first term is quadratic in x and positive definite, the second term is linear in x, and the last term is the negative of a concave function.

3. Compute the gradient and the Hessian of the function $f^{[k]}(x)$.

Solution: The gradient is given by:

$$\nabla f^{[k]}(x) = Qx + c - \tau^{[k]} \sum_{j=1}^{m} \frac{a_j}{a_j^\top x + b_j}$$

The Hessian is given by:

$$\nabla^2 f^{[k]}(x) = Q + \tau^{[k]} \sum_{j=1}^{m} \frac{a_j a_j^\top}{(a_j^\top x + b_j)^2}$$

Note that a more compact way to write this is to define $\tilde{a}^{[k]} := \frac{a_j}{a_j^\top x + b_j}$, then write:

$$\nabla f^{[k]}(x) = Qx + c - \tau^{[k]} \sum_{j=1}^{m} \tilde{a}_j^{[k]}$$

$$\nabla^2 f^{[k]}(x) = Q + \tau^{[k]} \sum_{j=1}^{m} \tilde{a}_j^{[k]} \tilde{a}_j^{[k]\top}$$

4. A popular variant of this algorithm is to perform only one Newton step at each iteration to find the next iterate $x^{[k+1]}$. Write the update rule for $x^{[k+1]}$.

   Remark: *In practice, we use backtracking line-search as a globalization strategy, but you do not have to write the rules of the line-search here.*

Solution: The update rule can be written as:

$$x^{[k+1]} = x^{[k]} - t^{[k]} \left(\nabla^2 f^{[k]}(x^{[k]})\right)^{-1} \nabla f^{[k]}(x^{[k]})$$

where $x^{[k]}$ is the current iterate.

5. Complete the code in the file `hanging_chain_ip.py` to implement the interior point method.

   Like in exercise 3, a visualization of the iterates is provided.

   Comment what you see.

Solution: See Python file `hanging_chain_ip_sol.py`
We see that the ground constraint pushes the chain upwards, but as the iterations progres, this ”push” is less and less active, until the chain can almost touch the ground.

II Sequential Quadratic Programming

In this exercise, we study the Sequential Quadratic Programming (SQP) method for a general nonlinear programming (NLP) problem:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g(x) = 0, \\ & h(x) \geq 0 \end{aligned} \tag{4}$$

This method consists of solving an approximate problem at each iteration, where the objective is approximated by a quadratic function and the constraints are approximated by linear functions. These approximations are based on the Taylors expansion at the current solution points.

1. Write down the KKT conditions of the generic NLP (4).

> **Solution:** KKT conditions of the generic NLP (4) are:
>
> $$\begin{aligned} \nabla f(x) - \nabla g(x)\lambda - \nabla h(x)\mu &= 0 \\ g(x) &= 0 \\ h(x) &\geq 0 \\ \mu &\geq 0 \\ \mu^\top h(x) &= 0 \end{aligned}$$
>
> where $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$ are the Lagrange multipliers associated with the equality and inequality constraints, respectively.

2. Formulate the generic QP subproblem that would result if your current iterate is $x^{[k]}$.

   Hint: You can formulate it with the decision variable $p = x - x^{[k]}$.

> **Solution:**
>
> $$\begin{aligned} \underset{p \in \mathbb{R}^n}{\text{minimize}} \quad & f\left(x^{[k]}\right) + \nabla f\left(x^{[k]}\right)^\top p + \frac{1}{2} p^\top \nabla^2 f\left(x^{[k]}\right) p \\ \text{subject to} \quad & g\left(x^{[k]}\right) + \nabla g\left(x^{[k]}\right)^\top p = 0, \\ & h\left(x^{[k]}\right) + \nabla h\left(x^{[k]}\right)^\top p \geq 0 \end{aligned}$$

3. Prove that if $x^{[k]}$ is a KKT point (for some multipliers $\lambda^{[k]}$ and $\mu^{[k]}$) of the NLP problem, and that the Hessian is positive (i.e. $\nabla^2 f\left(x^{[k]}\right) \succ 0$), then it is a solution to the QP subproblem.

   Hint: Meaning that $p = 0$ is a solution if you choose $p = x - x^{[k]}$ as the decision variable.

Solution: The KKT conditions read:

$$\begin{aligned}
\nabla f\left(x^{[k]}\right) &= \underbrace{\nabla g\left(x^{[k]}\right)}_{=:A}\lambda^{[k]} + \underbrace{\nabla h\left(x^{[k]}\right)}_{=:B}\mu^{[k]} \\
g\left(x^{[k]}\right) &= 0 \\
h_i\left(x^{[k]}\right) &= 0 \quad \forall i \in \mathcal{A} \\
\mu_i^{[k]} &\geq 0 \quad \forall i \in \mathcal{A} \\
h_i\left(x^{[k]}\right) &> 0 \quad \forall i \notin \mathcal{A} \\
\mu_i^{[k]} &= 0 \quad \forall i \notin \mathcal{A}
\end{aligned}$$

where $\mathcal{A} \subset \{1, \ldots, m\}$ is the active set of constraints.
The goal is to show that the value of the QP subproblem is minimized at $p = 0$, i.e. that its value is greater or equal to $f\left(x^{[k]}\right)$. Its value is:

$$\begin{aligned}
\min_{p \in \mathbb{R}^n} \quad & f\left(x^{[k]}\right) + \underbrace{\nabla f\left(x^{[k]}\right)}_{=A\lambda^{[k]}+B\mu^{[k]}}{}^\top p + \frac{1}{2}p^\top \nabla^2 f\left(x^{[k]}\right) p \\
\text{s.t.} \quad & \underbrace{g\left(x^{[k]}\right)}_{=0} + \underbrace{\nabla g\left(x^{[k]}\right)}_{=A}{}^\top p = 0 \\
& h\left(x^{[k]}\right) + \underbrace{\nabla h\left(x^{[k]}\right)}_{=B}{}^\top p \geq 0
\end{aligned}$$

This simplifies to:

$$\begin{aligned}
\min_{p \in \mathbb{R}^n} \quad & f\left(x^{[k]}\right) + \underbrace{\lambda^{[k]\top} A^\top p}_{=0} + \sum_{i \in \mathcal{A}} \mu_i^{[k]} \underbrace{\left(B^\top p\right)_i}_{\geq 0} + \sum_{i \notin \mathcal{A}} \underbrace{\mu_i^{[k]}}_{=0} \left(B^\top p\right)_i + \underbrace{\frac{1}{2}p^\top \nabla^2 f\left(x^{[k]}\right) p}_{\geq 0} \\
\text{s.t.} \quad & A^\top p = 0 \\
& \left(B^\top p\right)_i \geq 0 \quad \forall i \in \mathcal{A} \\
& \left(B^\top p\right)_i \geq -h_i\left(x^{[k]}\right) \quad \forall i \notin \mathcal{A}
\end{aligned}$$

This allows to conclude that the problem is upper bounded by $f\left(x^{[k]}\right)$ which concludes the proof.

4. Show the converse, i.e. if $x = x^{[k]}$ is the solution of the QP subproblem, then $x^{[k]}$ is a KKT point.

   Hint: *You can use the fact that for two matrices A and B, if* $\operatorname{Ker}\left(A^\top\right) \subset \operatorname{Ker}\left(B^\top\right)$ *then* $\operatorname{Im}(B) \subset \operatorname{Im}(A)$

Solution: Since $x = x^{[k]}$ is a solution to the QP problem, then it is feasible. Define $\tilde{g}(x)$ as the stack of all the active constraints (including the equality constraints), and $\tilde{h}(x)$ as the stack of all of the inactive constraints.

Let $d \in \mathbb{R}^n$ be such that $\nabla \tilde{g}\left(x^{[k]}\right)^\top d = 0$, i.e. $d \in \operatorname{Ker}\left(\nabla \tilde{g}\left(x^{[k]}\right)^\top\right) = \operatorname{Im}\left(\nabla \tilde{g}\left(x^{[k]}\right)\right)^\perp$. It is clear that for ε small enough, $x_\varepsilon = x^{[k]} \pm \varepsilon d$ is feasible for the QP subproblem. Because $x = x^{[k]}$ is a solution, we have:

$$
\begin{aligned}
& f\left(x^{[k]}\right) + \nabla f\left(x^{[k]}\right)^\top \left(x_\varepsilon - x^{[k]}\right) + \frac{1}{2}\left(x_\varepsilon - x^{[k]}\right)^\top \nabla^2 f\left(x^{[k]}\right)\left(x_\varepsilon - x^{[k]}\right) \geq f\left(x^{[k]}\right) \\
\implies & \pm \varepsilon \nabla f\left(x^{[k]}\right)^\top d + \varepsilon^2 \frac{1}{2} d^\top \nabla^2 f\left(x^{[k]}\right) d \geq 0 \\
\implies & \pm \nabla f\left(x^{[k]}\right)^\top d + \underbrace{\varepsilon \frac{1}{2} d^\top \nabla^2 f\left(x^{[k]}\right) d}_{\to 0 \text{ when } \varepsilon \to 0} \geq 0 \\
\implies & \pm \nabla f\left(x^{[k]}\right)^\top d \geq 0 \\
\implies & \nabla f\left(x^{[k]}\right)^\top d = 0
\end{aligned}
$$

Hence, for all $d \in \operatorname{Ker}\left(\nabla \tilde{g}\left(x^{[k]}\right)^\top\right)$, we also have: $d \in \operatorname{Ker}\left(\nabla f\left(x^{[k]}\right)^\top\right)$. This implies that $\operatorname{Ker}\left(\nabla \tilde{g}\left(x^{[k]}\right)^\top\right) \subset \operatorname{Ker}\left(\nabla f\left(x^{[k]}\right)^\top\right)$, which itself implies: $\operatorname{Im}\left(\nabla f\left(x^{[k]}\right)\right) \subset \operatorname{Im}\left(\nabla \tilde{g}\left(x^{[k]}\right)\right)$, i.e. there exists $\tilde{\lambda}$ such that $\nabla f\left(x^{[k]}\right) = \nabla \tilde{g}\left(x^{[k]}\right)\tilde{\lambda}$. Putting everything together, we find that the KKT conditions are verified.

III A sample exam question

Regard the following minimization problem:

$$\begin{aligned} \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad & x_2^4 + (x_1 + 2)^4 \\ \text{subject to} \quad & x_1 - x_2 = 0, \\ & x_1^2 + x_2^2 \leq 8 \end{aligned}$$

1. How many scalar decision variables, how many equality, and how many inequality constraints does this problem have?

2	

Solution: two scalar decision variables, 1 equality constraint, 1 inequality constraint

2. Sketch the feasible set $\Omega \in \mathbb{R}^2$ of this problem.

3	

Solution: This is basically the segment from $(-2, -2)$ to $(2, 2)$.

3. Bring this problem into the NLP standard form

$$\begin{aligned} \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g(x) = 0, \\ & h(x) \geq 0 \end{aligned}$$

by defining the functions f, g, h appropriately.

3	

Solution:

$$f(x) = x_2^4 + (x_1+2)^4$$
$$g(x) = x_1 - x_2$$
$$h(x) = 8 - x_1^2 - x_2^2$$

FROM NOW ON UNTIL THE END TREAT THE PROBLEM IN THIS STANDARD FORM.

4. Is this optimization problem convex? Justify your answer.

2	

Solution: $f(x)$ is convex, $g(x)$ is affine, $h(x)$ is concave $\implies$ the problem is convex

5. Write down the Lagrangian function of this optimization problem.

2	

Solution:

$$\begin{aligned}\mathcal{L}(x,\lambda,\mu) &= f(x) - \lambda^\top g(x) - \mu^\top h(x)\\ &= x_2^4 + (x_1+2)^4 - \lambda(x_1 - x_2) - \mu(8 - x_1^2 - x_2^2)\end{aligned}$$

where $\lambda \in \mathbb{R}$ and $\mu \in \mathbb{R}^2$ are the Lagrange multipliers associated with the equality and inequality constraints respectively.

6. A feasible solution of the problem is $\bar{x} = \begin{bmatrix}2\\2\end{bmatrix}$. What is the active set $\mathcal{A}(\bar{x})$ at this point?

2	

Solution: $h(\bar{x}) = 8 - 2^2 - 2^2 = 0 \implies$ the constraint is active, $\mathcal{A}(\bar{x}) = \{1\}$ (This notation interprets $h(x)$ as vector valued function with only one dimension, i.e. a "scalar vector")

7. Is the *linear independence constraint qualification (LICQ)* satisfied at $\bar{x}$? Justify.

3	

Solution: Check linear independence of $\nabla g(\bar{x})$ and $\nabla h_i(\bar{x})$, $i \in \mathcal{A}$ or whether $\begin{bmatrix}\nabla g(\bar{x}) & \nabla h_1(\bar{x})\end{bmatrix}$ is full rank.

$$\nabla g(x) = \begin{bmatrix}1\\-1\end{bmatrix} = \nabla g(\bar{x}) \qquad \nabla h_1(x) = \begin{bmatrix}-2x_1\\-2x_2\end{bmatrix},\ \nabla h_1(\bar{x}) = \begin{bmatrix}-4\\-4\end{bmatrix},$$

$$\det\begin{bmatrix}\nabla g(\bar{x}) & \nabla h_1(\bar{x})\end{bmatrix} = \det\begin{bmatrix}1 & -4\\-1 & -4\end{bmatrix} = 8 > 0 \implies \text{full rank} \implies \text{LICQ satisfied}$$

8. An optimal solution of the problem is $x^\star = \begin{bmatrix}-1\\-1\end{bmatrix}$. What is the active set $\mathcal{A}(x^\star)$ at this point?

1	

Solution: $h(x^\star) = 8 - (-1)^2 - (-1)^2 = 6 > 0 \implies$ the constraint is inactive, $\mathcal{A}(x^\star) = \emptyset$

9. Is the *linear independence constraint qualification (LICQ)* satisfied at $x^\star$? Justify.

2	

Solution:

$$\nabla g(x^\star) = \begin{bmatrix}1\\-1\end{bmatrix} \implies \text{full rank} \implies \text{LICQ satisfied}$$

10. Describe the tangent cone $T_\Omega(x^\star)$ (the set of feasible directions) to the feasible set at this point $x^\star$, by a set definition formula with explicitly computed numbers.

2	

Solution: LICQ holds at $x^\star$, so the tangent cone and the linearized feasible cone coincide, i.e. the set of directions p such that $\nabla g(x^\star)^\top p = 0$ and $\nabla h_i(x^\star)^\top p = 0$ for $i \in \mathcal{A}(x^\star)$.
In the present case:

$$T_\Omega(x^\star) = \mathcal{F}(x^\star) = \{p \in \mathbb{R}^2 \mid \nabla g(x^\star)^\top p = 0\} = \{p \in \mathbb{R}^2 \mid \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = 0\} = \{p \in \mathbb{R}^2 \mid p_1 = p_2\}$$

11. Write down the KKT conditions for the point $x^\star$ and solve them to find the multipliers $\lambda^\star$ and $\mu^\star$.

4	

Solution: KKT conditions:

$$\begin{aligned} \nabla f(x^\star) - \nabla g(x^\star)\lambda^\star - \nabla h(x^\star)\mu &= 0 \\ g(x^\star) &= 0 \\ h(x^\star) &\geq 0 \\ \mu^\star &\geq 0 \\ h(x^\star)^\top \mu^\star &= 0 \end{aligned}$$

Here, $\mu^\star = 0$ because the constraint is active (i.e. $h(x^\star) = 0$). Furthermore, the following can be computed:

$$\nabla f(x^\star) = \begin{bmatrix} 4(x_1+2)^3 \\ 4x_2^3 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \end{bmatrix} \qquad \nabla g(x^\star) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Hence, we find $\lambda^\star = 4$.

12. Formulate the second order necessary conditions for optimality (SONC) for this problem and test if they are satisfied at $(x^\star, \lambda^\star, \mu^\star)$. Can you prove whether $x^\star$ is a local or even global minimizer?

Solution: SONC (when LICQ):

(a) $\exists \lambda^\star, \mu^\star$ such that KKT conditions hold

(b) $\forall p \in T_\Omega(x^\star)$ holds $p^\top \nabla_x^2 \mathcal{L}(x^\star, \lambda^\star, \mu^\star) p \geq 0$

Here:

$$\nabla_x^2 \mathcal{L}(x^\star, \lambda^\star, \mu^\star) = \begin{bmatrix} 12(x_1^\star + 2)^2 + 2\mu^\star & 0 \\ 0 & 12 x_2^{\star 2} + 2\mu^\star \end{bmatrix} = \begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix},$$
$$T_\Omega(x^\star) = \{p \in \mathbb{R}^2 \mid p_1 = p_2\}$$

It is clear that SONC are satisfied.
Since the problem is convex and LICQ is verified, KKT conditions are even equivalent to being a solution. Hence, $x^\star$ is a local and global minimizer.