Prof. Dr. Karsten Urban
M. Sc. Lewin Ernst

Winter term 2024/2025
Sheet 10, 46 Points

Numerical Optimization
exercise sheet a.k.a. test exam

review on 08.01.2025 during the exercise class

Throughout we use the following notation for a constraint optimization problem. For $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ we consider

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ \text{s.t. } &h(x) = 0, \\ &g(x) \leq 0 \end{aligned} \tag{1}$$

The feasible set is defined as $\mathcal{F} := \{x \in \mathbb{R}^n : h(x) = 0,\ g(x) \leq 0\}$.

1. *(Exercise 1, Numerical Optimization: General Knowledge)*
   Decide whether the following statements are true or false. Mark your answers **clearly**. Each correct answer gives one point.

   (a) In each step of the Levenberg-Marquardt algorithm, one tries to find a step $s^{(k)}$ and a parameter $\mu^{(k)}$ by solving $\|F'(x^{(k)})s^{(k)} + F(x^{(k)})\|_2^2 + (\mu^{(k)})^2\|s^{(k)}\|_2^2 \to \min$. true ◯ false ◯

   (b) A minimization problem with equality and inequality constraints (1) can always be rewritten as a minimization problem with only equality constraints. true ◯ false ◯

   (c) A minimization problem with equality and inequality constraints (1) can always be rewritten as a minimization problem with only inequality constraints. true ◯ false ◯

   (d) The set $\{x \in \mathbb{R}^2 : 0 \leq x_2 \leq x_1^2,\ x_1 \geq 0\}$ is a cone. true ◯ false ◯

   (e) The tangential cone of the set $\{x \in \mathbb{R}^2 : 0 \leq x_2 \leq (x_1 - 1)^2,\ x_1 \geq 1\}$ in $(1, 0)^T \in \mathbb{R}^2$ is given by $\{x \in \mathbb{R}^2 : x_2 = 0,\ x_1 > 1\}$. true ◯ false ◯

   (f) Let $x^* \in \mathbb{R}^n$ be a local solution of the minimization problem (1) such that a constraint qualification holds. Then, x^* is a stationary point of the function $F(x) = f(x) + \lambda^T h(x) + \mu^T g(x)$, for some parameters $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p$. true ◯ false ◯

   (g) The minimization problem (1) with $h(x) := Bx - Cx + v$ and $g(x) := x^T A x - a$, where $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite is convex. true ◯ false ◯

   (h) Let $f \in C^0(\mathbb{R}^n)$ be given. Then, the Nelder-Mead method applied to the problem $\min_{x \in \mathbb{R}^n} f(x)$ will converge to a local minimum. true ◯ false ◯

   (8 Points)

2. *(Exercise 2, Application of Numerical Methods)*
In this exercise you are given situations (already mathematically modeled) with different requirements for which you should decide which numerical method you would use and why.
Remark: A maximum of *one point* is awarded for a method that can be applied but does not meet the requirements. However, the applicability must be justified.
Moreover, if a method reduces a minimization problem to a sequence of minimization problems, then the numerical method for the subproblems also has to be given.

a)
- Situation: The steady state $x \in \mathbb{R}^{10^5}$ of a chemical system with 10^5 species should be determined by minimizing the total energy function $E_{\text{tot}} \in C^\infty(\mathbb{R}^{10^5})$. Derivatives can be calculated exactly by automatic differentiation algorithms. One expects many local minima.
- Problem: $\min\limits_{x \in \mathbb{R}^{10^5}} E_{\text{tot}}(x)$
- Requirements: The available computer memory is limited and can only store at most 10^7 values. The accuracy of the solution $x^* \in \mathbb{R}^{10^5}$ is important for the application.
- Numerical method:
- Justification:

b)
- Situation: A manufacturer of photo cameras of mobile phones develops a new camera system. The position $x_i \in \mathbb{R}^3$ and the curvature $\rho_i \in \mathbb{R}$ of the lenses $i = 1, \dots, 5$ are determined by an optimization problem. The function $f : \mathbb{R}^{3 \cdot 5 + 5} \to \mathbb{R}$, $y := (x_1, \dots, x_5, \rho_1, \dots, \rho_5) \mapsto \|g(y) - q\|_2^2$ to be minimized measures the sharpness of the picture $g(y) : \mathbb{R}^{20} \to \mathbb{R}^{100}$ of the optical system from the desired sharpness $q \in \mathbb{R}^{100}$. The derivatives are computable.
- Problem: $\min\limits_{y \in \mathbb{R}^{20}} f(y)$
- Requirements: Stable algorithm which should not be stuck in saddle points.
- Numerical method:
- Justification:

(3 + 3 = 6 Points)

3. *(Exercise 3, KKT optimality conditions)*
Consider the problem

$$\begin{cases} \min\limits_{x\in\mathbb{R}^2} & f(x_1,x_2) := \exp(x_2) + \dfrac{1}{2}(x_1+1)^2 - x_2; \\ \text{s.t.} & g(x_1,x_2) := {x_2}^2 - 1 \leq 0; \\ & h(x_1,x_2) := -x_1 + x_2 - 1 = 0. \end{cases}$$

a) Determine the feasible set $\mathcal{F}$ explicitly by finding a function $\varphi : I \subset \mathbb{R} \to \mathbb{R}^2$ and a set $I \subset \mathbb{R}$, such that $\mathcal{F} = \{\varphi(t) : t \in I\}$.

b) Set up the KKT conditions for this problem.

c) Prove that this problem is convex. What does that mean for a local solution x^* of the problem?

d) Prove that the Guignard Constraint Qualification (GCQ) is fulfilled at $(-2,-1)^T \in \mathbb{R}^2$.

e) Are the KKT conditions fulfilled at the local solution $x^* = (-1,0)^T \in \mathbb{R}^2$?

$(3+3+3+2+1=12$ Points)

4. *(Step size control)*

a) When using gradient descent methods, the condition

$$f(x^{(k+1)}) = f(x^{(k)} + \alpha^{(k)} p^{(k)}) < f(x^{(k)}) \tag{2}$$

is in general not enough to guarantee that a resulting sequence $(x^{(k)})_{k\in\mathbb{N}}$ converges to the minimum x^*.
Find sequences $(\alpha^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}_+$, $(p^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}$ and a initial value $x^{(0)} \in \mathbb{R}$, such that the sequence defined by $x^{(k+1)} = x^{(k)} + \alpha^{(k)} p^{(k)}$ satisfies (2) for $f(x) := (x-5)^2 - 2$ and does not converge to the unique minimum $x^* = 5$.

b) To ensure some decay of the objective function, we need step-size conditions like the so-called Armijo condition.
For $f \in C^1(\mathbb{R}^n)$, let $x^{(k)} \in \mathbb{R}^n$ and $p^{(k)} \in \mathbb{R}^n$ be vectors such that

$$\nabla f(x^{(k)})^\top p^{(k)} < 0$$

is satisfied. Furthermore, let f be bounded from below on the set $\{x^{(k)} + \alpha\, p^{(k)} : \alpha \in \mathbb{R}_+\}$.
Let $c_1 \in (0,1)$. Prove that there exists an $\tilde{\alpha} \in \mathbb{R}_+$, s.t.

$$f(x^{(k)} + \tilde{\alpha} p^{(k)}) \leq f(x^{(k)}) + c_1 \tilde{\alpha}\, \nabla f(x^{(k)})^\top p^{(k)}.$$

(3 + 3 = 6 Points)

5. *(Quadratic problems)*
Consider the problem

$$\begin{cases} \min\limits_{x \in \mathbb{R}^n} & f(x) := x^T A x + a^T x \\ \text{s.t.} & h(x) := L x_1 - b = 0, \end{cases}$$

where $A \in \mathbb{R}^{n \times n}$ is positive semi-definite. A has the form

$$A = \begin{pmatrix} A_{11} \in \mathbb{R}^{m \times m} & 0 \\ 0 & A_{22} \in \mathbb{R}^{(n-m) \times (n-m)} \end{pmatrix}$$

and

$$a = \begin{pmatrix} a_1 & \in \mathbb{R}^m \\ a_2 & \in \mathbb{R}^{n-m} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 & \in \mathbb{R}^m \\ x_2 & \in \mathbb{R}^{n-m} \end{pmatrix}, \quad L = (\ell_{ij})_{1 \le i,j \le m} \text{ with } \ell_{ij} \neq 0, \ \forall\, i \ge j$$

a) Prove that a KKT-point $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ of the problem above is a global solution.

b) Prove that, if A is positive definite on the kernel of $B := (L, 0) \in \mathbb{R}^{m \times n}$, then A_{22} is invertible.

c) Assume that the statement in b) holds. Determine the solution $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ of the problem and implement a specialized MATLAB-solver for this problem in the following code-snipped.

```
function [x,lambda] = nullspace_method_special(A_11, A_22, a, L, b)
% Specialized nullspace method.
% Input:
%   A_11: matrix in \R^{m x m}
%   A_22: matrix in \R^{(n-m) x (n-m)}
%   a   : vector in \R^n
%   L   : matrix in \R^{m x m}
%   b   : vector in \R^m
% Output:
%   x     : solution vector in \R^n
%   lambda: Lagrange multiplier solution in \R^m


% TODO















end
```

(4 + 4 + 6 = 14 Points)