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

Winter term 2024/2025
Sheet 10, 46 Points

Numerical Optimization
Solution to 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 $g(x) := x^T Ax - a$, 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)

Solution: The explanations are just for sake of clarity and not needed in an exam.

(a) False, the parameter $\mu^{(k)}$ is given in the optimization problem

(b) True, because $g(x) \leq 0$ can be equivalently be rewritten as $\max\{g(x), 0\} = 0$.

(c) True, because $h(x) = 0$ is equivalent to $h(x) \leq 0$ and $-h(x) \leq 0$.

(d) False, because not every point of the line $\varphi(x_1, x_2) = \alpha \cdot (1,1)^T$, $\alpha > 0$ is included in the set, but $(0,0)^T$ and $(1,1)^T$ is.

(e) False, it should be the set $\{x \in \mathbb{R}^2 : x_2 = 0,\ x_1 \geq 0\}$.

(f) True, because from Theorem 3.1.17 it follows that x^* is a KKT point and the multiplier rule states that is is also a stationary point for the Lagrange function.

(g) False, because the function does not have to be convex.

(h) False, on sheet 3 we have seen an example (the McKinnon function) for which the Nelder-Mead method gets stuck.

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 $i = 1, \ldots, 5$ are determined by an optimization problem. The function $f : \mathbb{R}^{3 \cdot 5 + 5} \to \mathbb{R}$, $y := (x_1, \ldots, x_5, \rho_1, \ldots, \rho_5) \mapsto \|g(y) - q\|_2^2$ 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)

Solution:

a) - <u>Situation:</u> 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.
- <u>Problem:</u> $\min\limits_{x \in \mathbb{R}^{10^5}} E_{\text{tot}}(x)$
- <u>Requirements:</u> 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: Nonlinear cg-method after Fletcher-Reeves or Polak-Ribiére with strong wolfe line search
- Justification: It is an unconstrained minimization problem and the nonlinear cg-method requires only a few vectors of the size 10^5 to be stored, because it uses only the gradient information. The gradient can be calculated exactly as mentioned in the text. Moreover, the norm of the gradient can also serve as a termination criterion for the algorithm, such that we stop if we are close enough to a stationary point.

b) - <u>Situation:</u> 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.
- <u>Problem:</u> $\min\limits_{y \in \mathbb{R}^{20}} f(y)$
- <u>Requirements:</u> Stable algorithm which should not be stuck in saddle points.
- Numerical method: Levenberg-Marquardt method
- Justification: The problem is a nonlinear regression problem, so that either Gauß-Netwon or Levenberg-Marquardt can be chosen. The more stable algorithm is the Levenberg-Marquardt method due to the damping parameter μ for the Newton-step. Moreover, if the critical point x^* is a local maximum or a saddle point, the scheme is repelling.

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

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

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)

Solution:

a) From the inequality constraint we get

$$x_2^2 - 1 \leq 0 \quad \Leftrightarrow \quad x_2 \in [-1, 1]$$

and the equality constraint function is affine linear, so that we get a straight line in $\mathbb{R}^2$ which can be parameterized as $\varphi(t) = (t, t+1)^T$, $t \in \mathbb{R}$. From the condition $x_2 \in [-1, 1]$ we get $t \in I := [-2, 0]$. Therefore the feasible set is given by the points $\mathcal{F} = \{\varphi(t) : t \in I\}$.

b) We derive the KKT conditions. The Lagrange function is given by $\mathcal{L}(x, \lambda, \mu) := f(x) + \lambda h(x) + \mu g(x)$. Therefore, we get

$$\begin{aligned} \nabla_x \mathcal{L} = \begin{pmatrix} x_1 + 1 \\ \exp(x_2) - 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 2x_2 \end{pmatrix} &= 0 \\ h(x) = -x_1 + x_2 - 1 &= 0 \\ \mu &\geq 0 \\ g(x) = x_2^2 - 1 &\leq 0 \\ \mu(x_2^2 - 1) &= 0 \end{aligned}$$

c) To prove that this problem is convex, we need to show that f as well as g are convex and that h is affine linear. The latter statement is obviously true, because $h(x) = a^T x + b$ with $a := (-1, 1)^T$ and $b := -1$. To prove that f and g are convex we use the second order equivalent conditions, namely that the Hessian of the functions are semi-positive definite. The Hessian of f is given by

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ 0 & e^{x_2} \end{pmatrix}.$$

For g we get

$$\nabla^2 g(x_1, x_2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Therefore, we have $x^T \nabla^2 f(x) x \geq 0$ for all $x \in \mathbb{R}^2$ for both functions. This means, by the KKT Theorem for convex problems, that any local solution is a global solution.

d) We prove GCQ at $(-2, -1)^T$ by proving the Slater condition. The set $\mathring{\mathcal{F}} = \{x \in \mathbb{R}^n : h(x) = 0,\ g(x) < 0\}$ is non-empty, because $(-1, 0)^T \in \mathring{\mathcal{F}}$.

e) Yes the KKT conditions are fulfilled, because we showed in d) that the Slater condition holds. Therefore, a CQ holds and by the KKT-Theorem for convex problems we have that any local solution x^* fulfills the KKT conditions.

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 an 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)

Solution:

a) We set $x^{(0)} := 8$, $(p^{(k)})_{k\in\mathbb{N}} := (-1)_{k\in\mathbb{N}}$ and $(\alpha^{(k)})_{k\in\mathbb{N}} := (2^{-k})_{k\in\mathbb{N}}$. With this we have

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - 2^{-k} = \ldots = x^{(0)} - \sum_{i=0}^{k} 2^{-i} \\ &= 8 - 2 + \frac{1 - 2^{k+1}}{1/2} = 6 + 2^k. \end{aligned}$$

The sequence is monotonically decreasing, i.e. $x^{(k+1)} < x^{(k)}$ and therefore the sequence $(f^{(k)})_{k\in\mathbb{N}} := (f(x^{(k)}))_{k\in\mathbb{N}}$ is as well, due to the monotonicity of f for $x \geq 5$. We have (2) but the sequence converges to $\lim_{k\to\infty} x^{(k)} = 6 \neq 5 = x^*$.

b) Note that $\Phi(\alpha) := f(x^{(k)} + \alpha p^{(k)})$ is bounded from below for all $\alpha > 0$. Since $0 < c_1 < 1$, the line $l(\alpha) := f(x^{(k)}) + \alpha c_1 \nabla f(x^{(k)})^T p^{(k)}$ is unbounded below and must therefore intersect the graph of Φ at least once. This means there must be an $\tilde{\alpha} > 0$ with $l(\tilde{\alpha}) = \Phi(\tilde{\alpha})$ (and so $l(\tilde{\alpha}) \geq \Phi(\tilde{\alpha})$).

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

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

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)

Solution:

a) We prove the statement by showing that the minimization problem above is convex. The Hessian of f is given by $\nabla^2 f(x) = (A + A^T)$ and due to the positive semi-definitness of A we get $x^T(A+A^T)x = 2x^TAx \geq 0$, i.e. the Hessian is also positive semi-definite. Therefore f is convex. The equality constraint function h can equivalently be rewritten as $\tilde{h}(x) = (L,0)^T x - b$ which is obviously affine linear. The problem is convex and by the Theorem 3.1.20 we have that every KKT point is global solution.

b) To determine the kernel of B one determine a QR decomposition of $B^T = Q\tilde{R}$. Q is given by the identity matrix as

$$B^T = \begin{pmatrix} L^T \\ 0 \end{pmatrix} := \tilde{R}.$$

The kernel of B is then given as the image of the matrix $Z := (e_{m+1}, \ldots e_n)$, where $e_i \in \mathbb{R}^n$ is the unit vector. By assumption A is positive definite on the kernel of B, i.e. $Z^TAZ > 0$ which is equivalent to $x^TA_{22}x > 0$ for all $x \in \mathbb{R}^{n-m}$. From this it follows that A_{22} is positive definite, thus invertible.

c) The solution is given by the KKT system, which reads

$$\begin{pmatrix} A_{11} & 0 & L^T \\ 0 & A_{22} & 0 \\ L & 0 & 0 \end{pmatrix} \begin{pmatrix} x_Y \\ x_Z \\ \lambda \end{pmatrix} = \begin{pmatrix} -a_1 \\ -a_2 \\ b \end{pmatrix}.$$

From the second block equation we get $x_Z = -A_{22}^{-1}a_2$. From the last equation we deduce $x_Y = L^{-1}b$ and with that we have $\lambda = -L^{-T}(a_1 + A_{11}x_Y)$.
The code could look like:

```
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

%-- (2) determine x_Y
x_Y = L\b;

%-- (3) determine x_Z and set x
x_Z     = -A_22\a(m+1:end);
x       = [x_Y; x_Z];

%-- (4) determine lambda
lambda    = -L'\(a(1:m)+A_11*x_Y);
end
```