# Numerical Optimization (Numerische Optimierung) – Exam

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg, March 21st, 2018, 9:00-11:15, Freiburg,

Georges-Koehler-Allee 101 Room 00-010/014

| Page | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Sum |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Points on page (max) | 3 | 9 | 11 | 6 | 9 | 10 | 12 | 12 | 8 | 0 | 80 |
| Points obtained | | | | | | | | | | | |
| Intermediate sum | | | | | | | | | | | |
| Intermediate sum (max) | 3 | 12 | 23 | 29 | 38 | 48 | 60 | 72 | 80 | 80 | |

Mark: Exam inspected on: Signature of examiner:

Last Name: First Name: Matriculation number:

Subject: Programme: Bachelor ☐ Master ☐ Lehramt ☐ others ☐ Signature:

Please fill in your name above. For the multiple choice questions, which give exactly one point, tick exactly one box for the right answer. For the text questions, give a short formula or text answer just below the question in the space provided and, if necessary, write on the **backpage of the same sheet** where the question appears, and add a comment “see backpage”. You can request additional empty pages for drafting. They have to be handed in at the end, but their content will not be graded. Only answers on the original exam sheets count. The exam is a closed book exam, i.e., no books or other material are allowed besides 2 sheets (with 4 pages) of hand-written notes and a non-programmable calculator. Some legal comments are found in a footnote[1].

1. Given two convex functions $f(x)$ and $g(x)$, which of the following operations is NOT necessarily convex?

| | |
|---|---|
| (a) ☐ affine input transformation: $h(x) = f(Ax + b)$ | (b) ☐ sum: $h(x) = f(x) + g(x)$ |
| (c) ☐ composition: $h(x) = f(g(x))$ | (d) ☐ point-wise maximum: $h(x) = \max\{f(x), g(x)\}$ |

| 1 | |
|---|---|

2. Which of the following functions is strictly convex?

| | | | |
|---|---|---|---|
| (a) ☐ $f(x) = -5x^2$ | (b) ☐ $f(x) = \log(x)$ | (c) ☐ $f(x) = 3x + 1$ | (d) ☐ $f(x) = e^x$ |

| 1 | |
|---|---|

3. A point in the feasible set of an NLP that satisfies the KKT optimality conditions is

| | |
|---|---|
| (a) ☐ the global minimum | (b) ☐ a candidate for local minimum |
| (c) ☐ a boundary point | (d) ☐ a local minimum |

| 1 | |
|---|---|

[1]WITHDRAWING FROM AN EXAMINATION: In case of illness, you must supply proof of your illness by submitting a medical report to the Examinations Office. Please note that the medical examination must be done at the latest on the same day of the missed exam. In case of illness while writing the exam please contact the supervisory staff, inform them about your illness and immediately see your doctor. The medical certificate must be submitted latest 3 days after the medical examination. More information’s: http://www.tf.uni-freiburg.de/studies/exams/withdrawing_exam.html

CHEATING/DISTURBING IN EXAMINATIONS: A student who disrupts the orderly proceedings of an examination will be excluded from the remainder of the exam by the respective examiners or invigilators. In such a case, the written exam of the student in question will be graded as ’nicht bestanden’ (5.0, fail) on the grounds of cheating. In severe cases, the Board of Examiners will exclude the student from further examinations.



| Points on page (max. 3) | |
|---|---|

4. Why is “globalization” used in the context of optimization?

| | |
|---|---|
| (a) ☐ to make the iterations cheaper | (b) ☐ to accelerate convergence |
| (c) ☐ to ensure convergence to a local minimum | (d) ☐ to ensure convergence to the global minimum |

| 1 | |
|---|---|

5. Which of the following is NOT true for the `fmincon` solver that you used in the exercises

| | |
|---|---|
| (a) ☐ It can be used through `casadi` | (b) ☐ It can solve nonlinear programs |
| (c) ☐ It can be used through `YALMIP` | (d) ☐ It can solve quadratic programs |

| 1 | |
|---|---|

6. The operation $X = A\backslash B$ in `MATLAB`, where $A$ and $B$ are matrices of appropriate dimensions, returns

| | |
|---|---|
| (a) ☐ The solution to the equation $BX = A$ | (b) ☐ The solution to the equation $AX = B$ |
| (c) ☐ The result of $X = AB^{-1}$, if B is invertible | (d) ☐ The point-wise division of the two matrices |

| 1 | |
|---|---|

7. How does one define a vector $x \in \mathbb{R}^n$ of optimization variables in YALMIP?

| | |
|---|---|
| (a) ☐ `x = optvar(n,1);` | (b) ☐ `x = sdpvar(n);` |
| (c) ☐ `x = sdpvar(n,1);` | (d) ☐ `x = optvar(n);` |

| 1 | |
|---|---|

8. Define mathematically what is a strict local minimizer $x^*$ of the problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad x \in \Omega$$

| 2 | |
|---|---|

9. Compute gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ of the function $f : \mathbb{R}^3 \to \mathbb{R},\ (x_1, x_2, x_3) \mapsto f(x_1, x_2, x_3) := x_1^2 + x_2^3 + x_1 x_3$.

| 3 | |
|---|---|



| Points on page (max. 9) | |
|---|---|

10. Duality: Regard the following equality constrained NLP (the *primal problem*)

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \qquad g(x) \;=\; 0$$

with global optimal solution $x^*$.

(a) Define the Lagrangian function $\mathcal{L}(x, \lambda)$ of this NLP.

| 2 | |
|---|---|

(b) Define the Lagrangian dual function $q(\lambda)$ of this problem. Is it convex or concave or none of the two?

| 2 | |
|---|---|

(c) Regard now $f(x) = \frac{1}{2}x^T Q x + c^T x$ and $g(x) = 1 - \frac{1}{2}\|x\|_2^2$ with an arbitrary symmetric matrix $Q \in \mathbb{R}^{n \times n}$ and vector $c \in \mathbb{R}^n$. Define and explicitly compute, if possible, the Lagrange dual function. Depending on the eigenvalues of $Q$, for which values of $\lambda$ is $q(\lambda)$ finite?

| 3 | |
|---|---|

(d) Write down the dual problem and use weak duality to formulate a bound on the optimal objective value $f(x^*)$. Is it a lower or upper bound?

| 2 | |
|---|---|

(e) For the case $n = 1, c = 1, Q = -2$, sketch the dual function $q(\lambda)$ and compute the bound of the previous question explicitly.

| 2 | |
|---|---|



| Points on page (max. 11) | |
|---|---|

11. Line search: regard an iterative descent algorithm for unconstrained minimization of a differentiable function $\phi : \mathbb{R}^n \to \mathbb{R}$, at the current iterate $x_k$ and with a search direction $\Delta x_k$.

    (a) When does a trial point $x_{k+1} = x_k + \alpha_k \Delta x_k$ with step length $\alpha_k$ satisfy the "Armijo-Condition" ?

    | 2 | |
    |---|---|

    (b) For the same differentiable function $\phi$, assume that there exists a global minimizer $x^*$ with $f(x^*) > -\infty$. Regard now an algorithm that generates an (infinite) sequence of points $x_0, x_1, x_2, x_3, \ldots$ such that each step satisfies the Armijo condition. Will the sequence necessarily converge to a stationary point of $\phi$? Justify your answer.

    | 2 | |
    |---|---|

12. Convergence rates: You observe an iterative optimization algorithm while it converges towards a solution. In each iteration, it gives you the norm of the current step.

    (a) You see the sequence

    ```
    iter |step|
    1   1.067E-2
    2   2.163E-3
    3   4.275E-4
    4   7.917E-5
    5   1.678E-5
    6   3.250E-6
    7   6.436E-7
    **** convergence achieved ****
    ```

    What local convergence rate does the algorithm have?

    | 1 | |
    |---|---|

    (b) With another algorithm, you see the sequence

    ```
    iter |step|
    1  1.657E+2
    2  2.123E+1
    3  2.275E+0
    4  7.917E-1
    5  1.678E-2
    6  3.250E-4
    7  6.436E-8
    **** convergence achieved ****
    ```

    What local convergence rate does the algorithm have?

    | 1 | |
    |---|---|



| Points on page (max. 6) | |
|---|---|

13. Regard vectors $s, y \in \mathbb{R}^n$ and a symmetric matrix $B \in \mathbb{R}^{n \times n}$ that satisfy $Bs = y$ and $s^T y < 0$. Is $B$ positive definite? Justify your answer.

| 2 | |
|---|---|

14. Automatic differentiation: regard the task to compute the Jacobian of a function $f : \mathbb{R}^n \to \mathbb{R}^2, x \mapsto (f_1(x), f_2(x))^T$. If evaluating $f(x)$ uses one millisecond of CPU time and $n = 1000$, how much time do you need to compute $\frac{\partial f}{\partial x}(x) = \nabla f(x)^T$ using (a) the forward and (b) the backward mode mode of automatic differentiation? Name one disadvantage of the backward mode compared to the forward one.

| 3 | |
|---|---|

15. Automatic Differentiation: regard the following algorithm to evaluate the function $f : \mathbb{R}^3 \to \mathbb{R}^2$.

```
function [f1,f2]=myfunction(x1,x2,x3)
v1=x1*x2;
v2=log(v1);
f1=v2/v1;
f2=v1*x3;
```

Write an algorithm (by hand) that computes the directional derivative $\nabla f(x)^T \dot{x} = \dot{f}$ in the direction $\dot{x} \in \mathbb{R}^3$ using the forward mode of automatic differentiation. You can do this by adding extra lines to the following template function. Use the variables `v1dot, v2dot, f1dot, f2dot` in the intermediate lines.

```
function [f1,f2,f1dot,f2dot] = my_dir_der_func(x1,x2,x3,x1dot,x2dot,x3dot)

v1=x1*x2;


v2=log(v1);


f1=v2/v1;


f2=v1*x3;


end
```

| 4 | |
|---|---|



| Points on page (max. 9) | |
|---|---|

16. A three-page question: Regard the following optimization problem.

$$\boxed{\min_{x \in \mathbb{R}^3} x_3 \quad \text{s.t.} \quad \begin{cases} x_1^2 + x_2^2 & \leq & 1 \\ x_3 & \geq & (x_1 - 1)^4 \end{cases}}$$

(a) How many variables, how many equality, and how many inequality constraints does this problem have?

| 1 | |
|---|---|

(b) Sketch the feasible set $\Omega \in \mathbb{R}^3$ of this problem. (*Hint: you may draw it from different sides.*)

| 3 | |
|---|---|

(c) Bring this problem into the inequality constrained NLP standard form:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \qquad h(x) \quad \geq \quad 0$$

by defining the functions $f, h$ appropriately.

| 2 | |
|---|---|

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

(d) Is this optimization problem convex? Justify.

| 2 | |
|---|---|

(e) An optimal solution of the problem is $x^* = (1, 0, 0)^T$. What is the active set $\mathcal{A}(\bar{x})$ at this point?

| 2 | |
|---|---|



| Points on page (max. 10) | |
|---|---|

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

| 3 | |
|---|---|

(g) Write down the Lagrangian function of this optimization problem.

| 2 | |
|---|---|

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

| 3 | |
|---|---|

(i) Formulate the necessary optimality conditions of first order (also called Karush-Kuhn-Tucker (KKT) conditions) that a local minimizer $x^* \in \mathbb{R}^3$ of this problem must satisfy, both generally and specifically.

| 4 | |
|---|---|



| Points on page (max. 12) | |
|---|---|

(j) Find a multiplier vector $\mu^*$ so that the above point $x^*$ satisfies the KKT conditions

| 3 | |
|---|---|

(k) Describe the critical cone $C(x^*, \mu^*)$ at the point $(x^*, \mu^*)$ in a set definition using explicitly computed numbers.

| 3 | |
|---|---|

(l) Formulate the second order necessary conditions for optimality (SONC) for this problem and test if they are satisfied at $(x^*, \mu^*)$.

| 2 | |
|---|---|

(m) Also formulate the second order sufficient conditions for optimality (SOSC) and test if they are satisfied at $(x^*, \mu^*)$.

| 2 | |
|---|---|

(n) Can you prove that the point $x^*$ is a local (or even global) minimizer?

| 2 | |
|---|---|



| Points on page (max. 12) | |
|---|---|

17. Constrained Gauss-Newton: assume that you want to solve the following NLP

$$\min_{x \in \mathbb{R}^2} \frac{1}{2}\|x\|_2^2 \quad \text{s.t.} \quad (x_1 - 4)^2 + x_2^2 - 9 = 0$$

and that the current iterate $x^{[k]}$ is given by $x^{[k]} = (1,1)^T$. Using the Gauss-Newton Hessian approximation, formulate the quadratic program (QP) that delivers the solution $p^{[k]}$, needed to compute the next iterate $x^{[k+1]} = x^{[k]} + p^{[k]}$. Compute all numbers in the matrices of the QP explicitly and use $p = (p_1, p_2)^T \in \mathbb{R}^2$ as a variable.

| 4 | |
|---|---|

18. Exact Hessian SQP: now we want to solve the same NLP as in Question 17, but with an exact Hessian SQP method. Assume that the current primal iterate $x^{[k]}$ is again given by $x^{[k]} = (1,1)^T$ and the current multiplier guess by $\lambda^{[k]} = -1$. Formulate the QP that delivers the next step $p^{[k]}$ (and multiplier guess $\lambda^{[k+1]}$). What is different in the exact Hessian QP compared to the Gauss-Newton QP from Question 17?

| 4 | |
|---|---|



| Points on page (max. 8) | |
|---|---|

**Empty page for calculations**



| | |
|---|---|
| Points on page (max. 0) | |