Prof. Dr. Karsten Urban
Institute of Numerical Mathematics

Winter Term 2021/22
Ulm University

Exam 1 - Numerical Optimization

March 8, 10:15 am

Exercise 1 **(8+3+3=14 Points)**

a) In numerical optimization, we first check which type of optimization problem we are dealing with and then we pick an appropriate numerical method to solve it.

$$(P1)\ \begin{cases} \max & f(x) = x^\top Ax + b^\top x; \\ \text{S.t.} & Cx \leq r; \\ \multicolumn{2}{l}{A \in \mathbb{R}^{n\times n} \text{ symmetric}, b \in \mathbb{R}^n, C \in \mathbb{R}^{p\times n}, r \in \mathbb{R}^p.} \end{cases}$$

$$(P2)\ \begin{cases} \min & f(x); \\ \text{S.t.} & Bx = g; \\ B \in \mathbb{R}^{p\times n}, g \in \mathbb{R}^p. \end{cases}$$

$$(P3)\ \begin{cases} \min & f(x) = (Ax-d)^\top (Ax-d); \\ A \in \mathbb{R}^{n\times n}, d \in \mathbb{R}^n. \end{cases}$$

$$(P4)\ \begin{cases} \min & f(x); \\ \text{S.t.} & x \in \mathcal{F}; \\ f : \mathcal{F} \to \mathbb{R} \text{ convex}, \mathcal{F} \subset \mathbb{R}^n \text{ open and convex.} \end{cases}$$

For each for the 4 problems, suggest a method from the lecture to solve it. If necessary, adjust the problem accordingly.

b) Figure 1 shows the contour plot of the Rosenbrock function as well as iterations of the classical gradient scheme and the nonlinear conjugate gradient method after Fletcher-Reeves. Fill the legend with the fitting method and shortly explain your choice.

c) In Figure 2, we see the simplexes computed by the Nelder-Mead method applied to the function

$$g(x, y) = \begin{cases} 360x^2 + y + y^2, & x < 0, \\ 6x^2 + y + y^2, & x \geq 0, \end{cases}$$

using the initial simplex

$$x^{(0)} = \{(1, 1)^\top,\ (0.8, -0.6)^\top,\ (0, 0)^\top\}.$$

We used parameters $\gamma = 0.5$, $\beta = 2$ and $\alpha = 1$ as well as $\alpha = 0.8$. Decide which plot is for $\alpha = 1$ and which one for $\alpha = 0.8$. Shortly explain your choice.

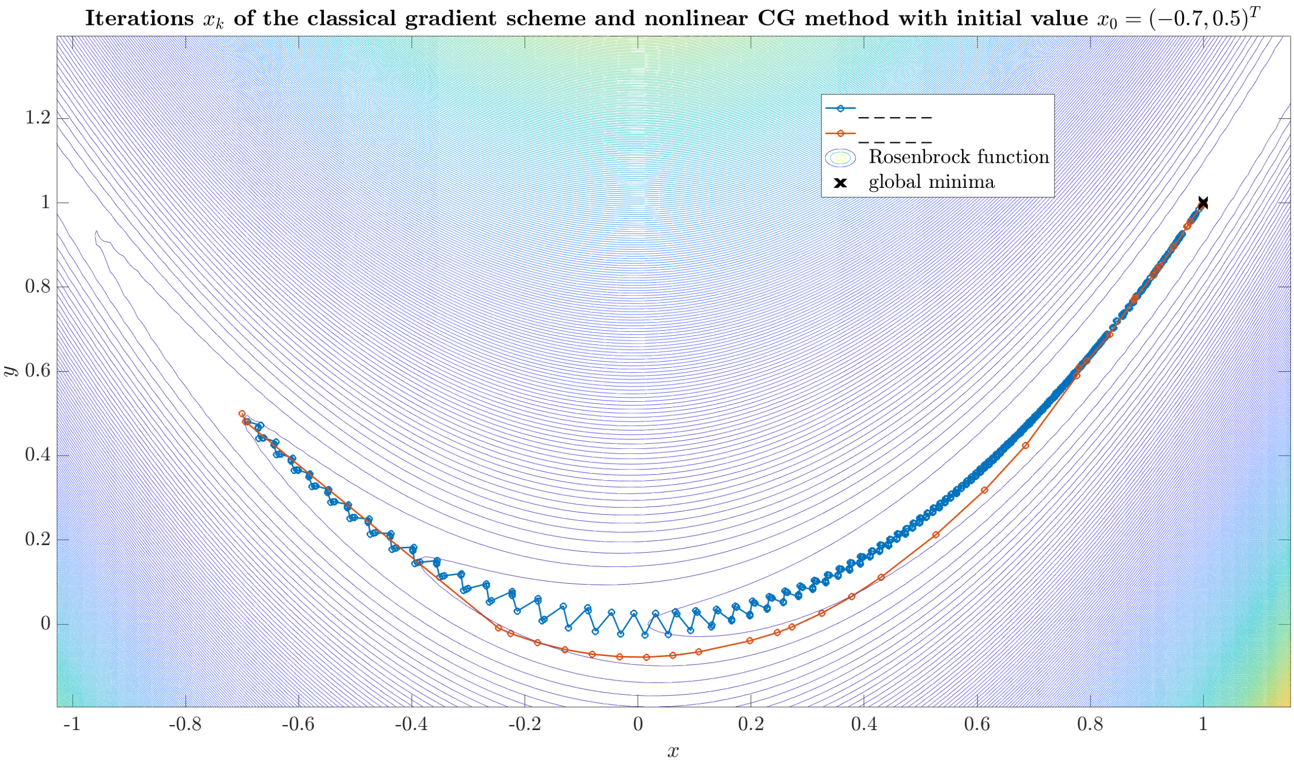

Figure 1: Exercise 1b) - Classical gradient scheme versus nonlinear CG.

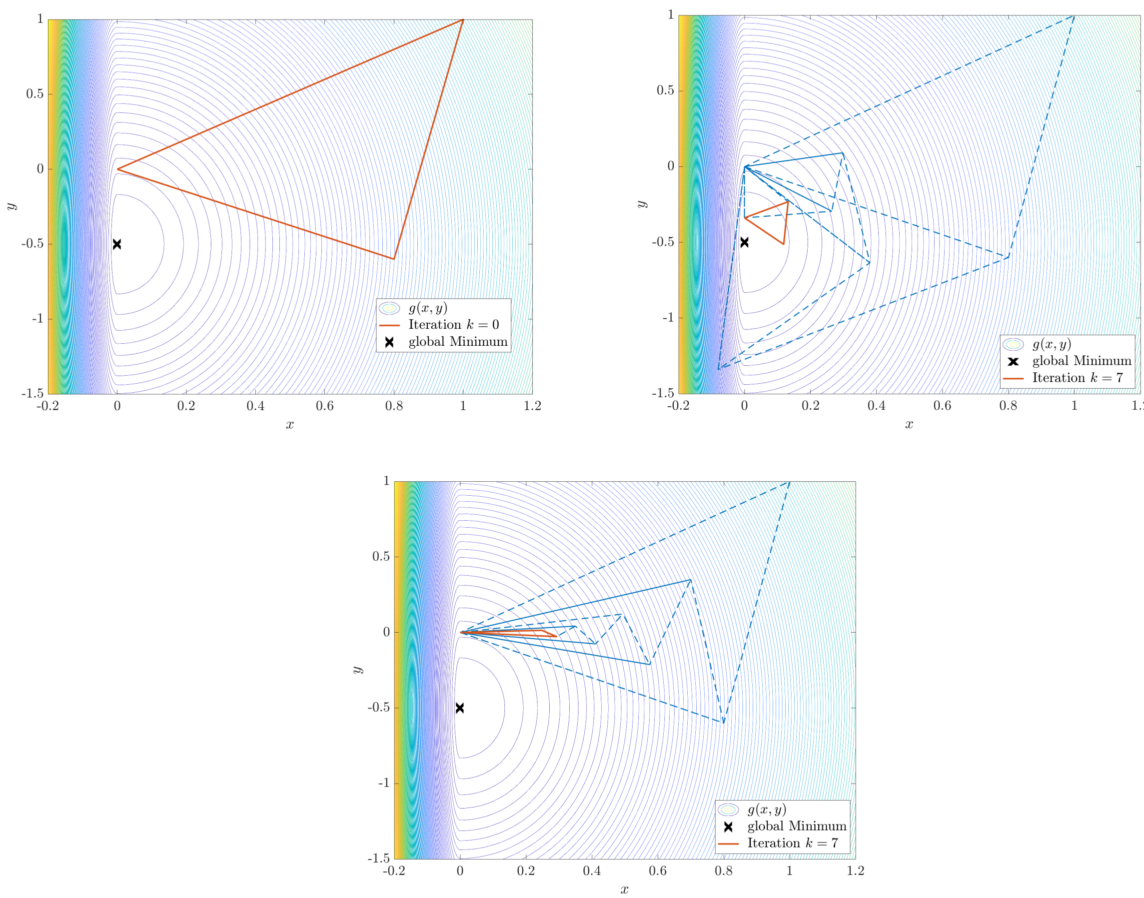

Figure 2: Exercise 1c) - Nelder-Mead simplexes with two different α.

Exercise 2 (Multiple Choice) **(10 Points)**

Decide whether the following statements are right or wrong. Correct the wrong statements.

No	Statement	Wrong	Correct
1:	In general, the Gauß-Newton method is locally second order convergent.		
2:	In the Active Set method, after performing the inactivation step one has $\tilde{d}^{(k)} \neq 0$. Then it holds that $\tilde{d}^{(k)}$ is a feasible direction in the point $x^{(k)}$.		
3:	If for $f \in C^1(\mathcal{F})$ it holds $\nabla f(x)^\top (x-y) \leq f(x) - f(y)$, for all $x, y \in \mathcal{F}$, then f is convex.		
4:	Let $q^{(k)}(p)$ be the quadratic model of the Trust-Region Method. The predicted reduction is defined as $\mathrm{pr}^{(k)} := f(x^{(k)}) - q^{(k)}(p^{(k)})$.		
5:	In the case of a convex optimization problem, if a local solution x^* satisfies the KKT conditions, then x^* is a global solution of the problem.		
6:	Let $f : \mathcal{F} \to \mathbb{R}$ be convex on the convex and open set $\mathcal{F} \subset \mathbb{R}^n$. Then, the negative subgradient of maximal norm is a descent direction.		
7:	Let x^* be a local solution of a constrained optimization problem with objective function f. Then it holds $\nabla f(x^*)^\top d < 0$ for all $d \in T(\mathcal{F}, x^*)$, where $T(\mathcal{F}, x^*)$ is the tangential cone.		
8:	To apply the semismooth Newton method, it is a precondition that the objective function f is convex.		

Exercise 3 (Nonlinear least squares problems) (4 + 5 + 5 = 14 Points)

To solve nonlinear least squares problems, we introduced the Gauß-Newton as well as the Levenberg-Marquardt method in the lecture.

a) Finish the MATLAB routine `GaussNewton.m` that realizes the Gauß-Newton method, by completing lines 5,6,7,8.

```
1  function xk = GaussNewton(F,DF,x0,maxIt,tol)
2      k=0;
3      xk=x0;
4      sk=-(DF(x0)'*DF(x0))\(DF(x0)'*F(x0));
5      while %TODO
6          %TODO
7          %TODO
8          %TODO
9      end
10 end
```

b) Find and correct the errors in the Levenberg-Marquardt MATLAB routine `LevenbergMarquardt.m`. You can correct the code on the exercise sheet.

```
function xk = LevenbergMarquardt(F,DF,x0,mu0,beta0,beta1,maxIt,tol)
    n = length(x0);
    k=0;
    mu=mu0;
    xk=x0;
    sk=-(DF(x0)'*DF(x0)+mu^2*eye(n))-(DF(x0)'*F(x0));
    while norm(sk)<tol && k<maxIt
        [sk,mu] = correction(F,DF,xk,mu,beta0,beta1);
        xk=xk+sk;
        k=k+1;
    end
end

function [s,mu] = correction(F,DF,x,mu,beta0,beta1)
    n = length(x0);
    s=-(DF(x)'*DF(x)+mu^2*eye(n))-(DF(x)'*F(x));
    eps_mu=(F(x)'*F(x)-F(x+s)*F(x+s)')/...
        (F(x)'*F(x)-(F(x)+DF(x)*s)'*(F(x)+DF(x)*s));
    if eps_mu <= beta0
        [s,mu]= correction(F,DF,x,2*mu,beta0,beta1);
    elseif eps_mu >= beta0
        mu=mu/2;
    end
end
```

c) After correctly implementing both methods, we use them to solve the nonlinear least squares problem

$$||F(x_1,x_2)||_2^2 \to \min!, \quad F(x_1,x_2) := \begin{pmatrix} x_1 \exp(0.1 \cdot x_2) \cdot \cos(0.1 \cdot 2\pi) - 0.395 \\ x_1 \exp(0.2 \cdot x_2) \cdot \cos(0.2 \cdot 2\pi) - 0.134 \\ x_1 \exp(0.3 \cdot x_2) \cdot \cos(0.3 \cdot 2\pi) + 0.119 \end{pmatrix},$$

and get the following errors.

Iteration	Error `GaussNewton.m`	Error `LevenbergMarquardt.m`
0	2.2522	2.2522
1	791.2193	2.1945
2	1.2659e+34	2.1798
3		2.1647
4		1.9487
$\vdots$		$\vdots$
20		1.6340e-04
$\vdots$		$\vdots$
30		5.0520e-08
$\vdots$		$\vdots$
38		2.1264e-10

The Gauß-Newton-Method breaks in iteration 3 with the output

```
Warning:  A is rank deficient to within machine precision.
```

where the Jacobi matrix in iteration 3 is

```
dF(x(:,k+1)) =
```

$$\begin{array}{cc} 0.0000 & 0.0395 \\ 0.0000 & 0.0000 \\ -0.0000 & -0.0000 \end{array}$$

Use this as a motivation to explain the difference of both methods for computing the Newton correction.

Exercise 4 (Armijo Condition) (5+7=12 Points)

a) When using gradient descent methods, the condition

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

is in general not enough to guarantee that a resulting sequence $(x^{(k)})_{k\in\mathbb{N}}$ converges to the minimum x^*.
Find a decreasing sequence $(x^{(k)})_{k\in\mathbb{N}}$ for $f(x) := (x-5)^2 - 2$ that does not converge to the unique minimum $x^* = 5$, $f(x^*) = -2$.

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 $p^{(k)} \in \mathbb{R}^n$ 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 $\sigma \in (0,1)$. Prove that there exists an $\tilde{\alpha} \in \mathbb{R}_+$, s.t.

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

Exercise 5 (KKT optimality conditions) (6+3+4=13 Points)

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) Set up the KKT conditions of this problem and solve the equations.

b) Can we use the KKT theorem for convex problems to find the global solution x^* of this problem?

c) Use the local solution x^* you computed in a). Prove that it is a global solution by using a Constraint Qualification on x^*.
This might be ACQ, GCQ or some alternative Constraint Qualification from the exercise session.

Exercise 6 (Linear Interior Point Method) **(8+5=13 Points)**

In this exercise, we aim to solve the primal problem

$$(P) \quad \begin{cases} c^\top x \to \min, \\ \text{subject to} \quad Ax = b,\ x \geq 0, \end{cases}$$

with it's dual

$$(D) \quad \begin{cases} b^\top y \to \max, \\ \text{subject to} \quad A^\top y \leq c. \end{cases}$$

The matrix $A \in \mathbb{R}^{m \times n}, m \leq n$, has full rank.

a) We assume that the system

$$F_\mu(x, y, s) := \begin{pmatrix} A^\top y + s - c \\ Ax - b \\ Xs - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad x \geq 0,\ s \geq 0, \tag{1}$$

with $X = \operatorname{diag}(x), e = (1, \ldots, 1)^\top$, has a unique solution for every $\mu > 0$. Derive equation (2) in Algorithm 1 using the Newton's Method for the system (1) and prove that system (2) can be solved uniquely in every step, where you can assume that $x, s > 0$ in every iteration.

Algorithm 1 Interior-Point Algorithm

Require: Starting vector $(x^{(0)}, y^{(0)}, s^{(0)}) \in \mathcal{F}^0 := \mathcal{F} \cap \{x > 0, s > 0\}$ and a tolerance $\epsilon > 0$, as well as centralizing parameter $\sigma \in [0, 1]$ is given;
  Initialize $(x, y, s) := (x^{(0)}, y^{(0)}, s^{(0)})$ and $\tau := \frac{1}{n} x^\top s$;
  while $\tau > \epsilon$ **do**
    Set $\mu = \sigma\tau$;
    Solve the system

$$DF_\mu(x, y, s) \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = F_\mu(x, y, s); \tag{2}$$

    Set $(x, y, s) := (x, y, s) - \alpha(\Delta x, \Delta y, \Delta s)$, where $\alpha \in (0, 1]$ denotes a suitable step size which we choose such that $x^\top s > 0$;
    $\tau := \frac{1}{n} x^\top s$;
  end while

---

b) A toy store owner sells three types of Lego boxes in his store, we name them A, B, C. The store owner pays for each box A \$5, for a box of type B \$11, and \$15 for C. One unit of toys A yields a profit of \$1, a unit of B yields a profit of \$2 and one unit C yields a profit of \$3. The store owner estimates that no more than 100 lego boxes will be sold every month and he can't order more than 40 boxes of one kind in a month. Further, since he also sells different toys in his store, he does not plan to invest more than \$1500 in inventory of Lego boxes. How many units of each type should be stocked in order to maximize his monthly total profit?

   Formulate this problem into a constrained optimization problem such that we can apply Algorithmn 1. If done correctly, the computer gives us the optimal solution $(A, B, C)^\top = (20, 40, 40)^\top$.