Technische Universität München
Department of Mathematics
Chair of Mathematical Optimization, M1
Prof. Dr. Michael Ulbrich
Dr. Andre Milzarek

Winter Semester 2016/17

Nonlinear Optimization: Advanced

Exercise Sheet 7

Exercise 7.1 (Multiple Choice Test):
The fourth multiple choice test is available on `www.moodle.tum.de`. Take the test and answer the different multiple choice questions online on moodle.

Exercise 7.2 (A Local Lagrange-Newton Type Method):
We consider the constrained problem

$$\min_{x \in \mathbb{R}^n} \; f(x) \qquad \text{s.t.} \qquad h(x) = 0, \tag{P$_1$}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are twice continuously differentiable functions. Let $\bar{x}$ be a local solution of problem (P$_1$) and suppose that the LICQ is satisfied at $\bar{x}$. Moreover, let $\bar{\mu}$ be a corresponding Lagrange multiplier and assume that the second-order sufficient conditions are satisfied at $(\bar{x}, \bar{\mu})$.

In this exercise, we consider a local Lagrange-Newton type method which uses a symmetric matrix $H_k \in \mathbb{R}^{n \times n}$ as a substitute for the Hessian $\nabla^2_{xx} L(x^k, \mu^k)$ of the standard Lagrange-Newton approach. The different steps of this method are shown and summarized in Algorithm 1.

Algorithm 1: A Local Lagrange-Newton Type Method

S0 Initialization: Choose an initial point $(x^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^p$.

for $k = 0, 1, 2, \ldots$ **do**

S1 STOP, if $F(x^k, \mu^k) := (\nabla_x L(x^k, \mu^k)^\top, h(x^k)^\top)^\top = 0$, i.e., if (x^k, μ^k) is a KKT pair of (P$_1$).

S2 Compute the direction $d^k = (d^k_x, d^k_\mu)$ as solution of the linear system of equations

$$M_k d^k = -F(x^k, \mu^k), \quad \text{where} \quad M_k = \begin{pmatrix} H_k & \nabla h(x^k) \\ \nabla h(x^k)^\top & 0 \end{pmatrix}.$$

S3 Set $x^{k+1} = x^k + d^k_x$ and $\mu^{k+1} = \mu^k + d^k_\mu$.

a) Suppose that each step of Algorithm 1 is well-defined. Let the sequence (x^k, μ^k) be generated by Algorithm 1 and assume that (x^k, μ^k) converges to $(\bar{x}, \bar{\mu})$. Show that the following two statements are equivalent:

   i) (x^k, μ^k) converges q-superlinearly to $(\bar{x}, \bar{\mu})$ and it holds $F(\bar{x}, \bar{\mu}) = 0$.

   ii) $\|(H_k - \nabla^2_{xx} L(x^k, \mu^k))(x^{k+1} - x^k)\| = o(\|(x^{k+1} - x^k, \mu^{k+1} - \mu^k)\|)$.

Hint: use the Dennis-Moré condition and Lemma 5.2 (more details on the Dennis-Moré condition can be found in chapter 11 of the book „M. Ulbrich, S. Ulbrich, *Nichtlineare Optimierung*, Birkhäuser Verlag, 2012").

In the following, we consider the special choice $H_k = \nabla^2_{xx} L(x^k, \mu^k) + \beta_k I$ and $\beta_k \geq 0$.

b) Let x^k be an arbitrary iterate and suppose that it holds rank $\nabla h(x^k) = p$. Show that, in this case, there exists $\beta_k^* \geq 0$ such that the equation in step **S2** has a unique solution for all $\beta_k \geq \beta_k^*$. Investigate whether the matrix M_k is positive definite for those β_k.

   Hint: you can use without proof that Lemma 5.2 remains valid if $\nabla^2_{xx} L(x, \mu)$ is replaced by an arbitrary matrix of the same size.

c) Let $(\beta_k) \subset \mathbb{R}_+$ be given and let (x^k, μ^k) be a sequence of iterates that is generated by Algorithm 1 and that converges to $(\bar{x}, \bar{\mu})$. Formulate a condition on the sequence (β_k) that guarantees superlinear convergence of (x^k, μ^k) to $(\bar{x}, \bar{\mu})$. This condition should also (but not exclusively) include the case $(\beta_k) \equiv 0$.

Exercise 7.3 (An Infeasible SQP Subproblem):
Let us consider the optimization problem

$$\min_{x \in \mathbb{R}^2} \; f(x) := -x_1 - x_2 \qquad \text{s.t.} \qquad g(x) := -x \leq 0, \quad h(x) := x_1^2 + x_2^2 - 1 = 0. \tag{P$_2$}$$

a) Compute a local solution of problem (P$_2$) and specify the corresponding Lagrange multipliers.

b) Is problem (P$_2$) convex?

c) Show that the local solution, which was calculated in part a), is the unique and global solution of problem (P$_2$).

d) Let the initial points $x^0 = (-0.5, -0.5)^\top$, $\lambda^0 = (1, 0)^\top$, and $\mu^0 = 1$ be given. For $H_0 = \nabla^2_{xx} L(x^0, \lambda^0, \mu^0)$ derive the objective function and the feasible set of the corresponding SQP subproblem explicitly.

e) Draw the constraints of the SQP subproblem of part d) and verify computationally that the feasible set of this subproblem is empty.

Exercise 7.4 (A Little Bit of Everything):
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function and let $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^p$ be twice continuously differentiable. We consider the optimization problem:

$$\min_{x \in \mathbb{R}^n} \; f(x) \qquad \text{s.t.} \qquad g(x) \leq 0, \quad h(x) = 0. \tag{P$_3$}$$

Let X denote the corresponding feasible set.

a) Let $x \in X$ be given. Does this imply $T(X, x) \cap T(X, x)^\circ = \{0\}$? (Explain your answer!)

b) Suppose that $m = 2$, $p = 1$, $g(x) := \left(e^{x_1} + \frac{1}{2}x_2^2 - 10, \|x\|^2 - 2\right)^\top$, and $h(x) := x_1 - x_2$ are given. Moreover, let x be a feasible point of this problem. Show that a CQ is satisfied at x.

c) Here, we set $n = 2$, $m = 2$, $p = 0$, $f(x) := |x_1| + x_2$, and $g(x) := (x_1, -x_2)^\top$. Apparently, $\bar{x} := (0, 0)^\top$ is the unique, global solution of this problem and the optimal objective function value satisfies $f(\bar{x}) = 0$. Investigate whether the corresponding dual problem has the same objective function value and if it possesses a unique, global solution.

d) Let f be twice continuously differentiable. Prove or disprove the following statement: *in contrast to the ℓ_1-penalty function, the quadratic penalty function of problem* (P_3) *is twice continuously differentiable for all* $\alpha > 0$.

e) We apply the quadratic penalty method (Algorithm 4.1) to problem (P_3). Suppose that the method generates a sequence (x^k, λ^k, μ^k) with $\lambda^k = \alpha_k (g(x^k))_+$ and $\lambda^k = \alpha_k h(x^k)$ such that x^k converges to some $\bar{x} \in \mathbb{R}^n$. Furthermore, assume that $\bar{x}$ is not a KKT point of problem (P_3). Show that this implies $\lim_{k\to\infty} \|(\lambda^k, \mu^k)\| \to \infty$.

f) Suppose that the constraint function h is affine linear and g is concave. Show that the SQP method produces only feasible iterates $x^k \in X$ for $k \geq 1$.

The exercises will be discussed from **Feb, 1st** to **Feb, 9th**.