Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization I Spring 2016

Exercise set 6

The first three of the following exercises are concerned with the exact solution of constrained optimisation problems by means of the first and second order optimality conditions. Note that the constraints in problems 2 and 3 have already been discussed in exercise sheet 5. Exercise 1 describes the "typical" case, there the KKT-conditions should work perfectly fine for finding the solution of the problem. In contrast, both exercise 2 and 3 deal with slightly problematic situations.

Exercises 4–6 deal with approximations of constrained optimisation problems using penalty methods. Apart from possibly 4d) all the computations should be straight-forward.

Exercises 7 and 8 can somehow be seen as a counterpoint to exercise 1 on sheet 3. There, we have discussed the solution of over-determined linear systems by means of the least-squares method, which replaces the actual solution of the equation by the best possible approximation. Here we are studying under-determined linear systems, where the problem lies not in actually finding a solution, but rather in selecting a good solution from the infinitely many one usually expects. One of the possible approaches to this problem consists in selecting the solution with the smallest Euclidean norm, which leads to the problem discussed in exercise 7. Additionally, if it is known that the right hand side b of the data includes some measurement errors, it can make sense to try to solve the equation not exactly, but only up to some accuracy depending on the size of the errors. In exercise 7b, the relation between this problem and the quadratic penalty method for the equality constraint problem is discussed. Note also that the matrix $A^T(AA^T)^{-1}$ that appears in exercise 7 actually coincides (in this particular case) with the *Moore–Penrose pseudoinverse* $A^\dagger$ of A (cf. the lecture *Linear methods*). In addition, the matrix in 7b is an approximation of the Moore–Penrose pseudoinverse; this allows another definition of $A^\dagger$ as $A^\dagger = \lim_{\mu\to\infty} A^T(\frac{1}{\mu}\operatorname{Id} + AA^T)^{-1}$, which avoids the usage of a singular value decomposition. Finally, exercise 8 deals with some duality theory for the problem in exercise 7. In particular it shows that in this case we have strong duality; that is, the value of the dual problem coincides with the value of the primal problem.

Exercises 9–12 are concerned with linear and quadratic optimisation problems and, again, deal a bit with duality.

1 Consider the constrained optimization problem

$$x^2 + y^2 \to \min \qquad \text{such that} \qquad \begin{cases} x + y \geq 1, \\ y \leq 2, \\ y^2 \geq x. \end{cases}$$

a) Formulate the KKT-conditions for this optimization problem.

b) Find all KKT points for this optimization problem.

c) Find all local and global minima for this optimization problem.

(Part b can be very tedious. One strategy is to consider all possible active sets and determine for each active set whether KKT-points exist. It can also be extremely helpful to sketch the feasible set and the function.)

2 Consider the constrained optimization problem

$$x \to \min \qquad \text{such that} \qquad \begin{cases} y \geq x^4, \\ y \leq x^3. \end{cases}$$

Find all KKT points and local minima for this optimization problem.

3 Consider the constrained optimization problem

$$xy \to \min \qquad \text{such that} \qquad \begin{cases} y \geq x, \\ y^4 \leq x^3. \end{cases}$$

a) Find all KKT points and local minima for this optimization problem.

b) Compute the critical cone at $(0,0)$ as defined in the lecture and Nocedal & Wright, and show that there exist directions d contained in the critical cone for which $d^T \nabla^2 \mathcal{L}((0,0);\lambda^*)d < 0$.

c) Show that $d^T \nabla^2 \mathcal{L}((0,0);\lambda^*)d \geq 0$ for all vectors d contained in the tangent cone to the feasible set at $(0,0)$.

4 Consider the constrained optimisation problem

$$\frac{1}{2}(x^2 + y^2) \to \min \qquad \text{subject to } xy = 1.$$

a) Find (by whatever means) the solutions of this problem. In addition, find the values of the corresponding Lagrange multipliers.

b) Formulate the unconstrained optimisation problem that results from the application of the quadratic penalty method with parameter $\mu > 0$. Solve these problems for all possible parameters μ and verify that the solutions converge to the solutions of the constrained optimization problem as $\mu \to \infty$.

c) Formulate the augmented Lagrangian for this constrained optimization problem and find (for all possible parameters $\lambda \in \mathbb{R}$ and $\mu > 0$) the global solutions of this (unconstrained) optimization problem. For which parameters does one recover the solution of the original constrained problem?

d) The ℓ^1-penalty function for this optimisation problem is defined, for some parameter $\mu > 1$, as

$$\Phi_1(x, y; \mu) := \frac{1}{2}(x^2 + y^2) + \mu|xy - 1|.$$

Find for each parameter $\mu > 0$ the global minimisers of this function. For which parameters $\mu > 0$ do they coincide with the solutions of the original problem?

5 Consider the constrained optimisation problem

$$x + y \to \min \qquad \text{subject to } x^2 + y^2 \leq 1.$$

Formulate a logarithmic barrier method for the solution of this constrained optimisation problem and compute its solution for each parameter $\mu > 0$ in the barrier functional.

6 Find an equality constrained optimisation problem for which the augmented Lagrangian is unbounded for all Lagrange parameters and all $\mu > 0$.

7 Assume that $A \in \mathbb{R}^{m \times n}$ with $m < n$ is a matrix of full rank and that $b \in \mathbb{R}^m \setminus \{0\}$. Consider the optimization problem

$$\frac{1}{2}\|x\|^2 \to \min \qquad \text{subject to} \qquad Ax = b. \tag{1}$$

a) Formulate the KKT-conditions for this problem and show that the unique solution is given by

$$x^* = A^T(AA^T)^{-1}b.$$

b) Formulate the quadratic penalty method for this constrained optimization problem, and show that the unique minimizer with parameter $\mu > 0$ is given by

$$x_\mu := A^T\Big(\frac{1}{\mu}\operatorname{Id} + AA^T\Big)^{-1}b$$

with $\operatorname{Id} \in \mathbb{R}^{m \times m}$ denoting the identity matrix.

c) Now consider the optimization problem

$$\frac{1}{2}\|x\|^2 \to \min \qquad \text{subject to} \qquad \frac{1}{2}\|Ax - b\|^2 \leq \varepsilon$$

for some $\varepsilon > 0$, and denote its solution by $\hat{x}_\varepsilon$. Show that either $\frac{1}{2}\|b\|^2 \leq \varepsilon$ (in which case $\hat{x}_\varepsilon = 0$), or there exists $\mu > 0$ such that $\hat{x}_\varepsilon = x_\mu$.

8 The Lagrangian of the problem (1) is the function $\mathcal{L}\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) = \frac{1}{2}\|x\|^2 - \lambda^T(Ax - b).$$

Now define the function $g\colon \mathbb{R}^m \to \mathbb{R}$,

$$g(\lambda) := \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda).$$

Then we can define the dual of the optimisation problem (1) as the maximisation problem

$$\max_{\lambda \in \mathbb{R}^m} g(\lambda). \tag{2}$$

a) Derive an explicit formula for the function g.

b) Show that $\lambda^* \in \mathbb{R}^m$ solves (2), if and only if

$$AA^T\lambda^* = b.$$

c) Verify that in this situation

$$\min_{x \in \mathbb{R}^n} \max_{\lambda \in \mathbb{R}^m} \mathcal{L}(x, \lambda) = \max_{\lambda \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda).$$

9 Exercise 13.1 in Nocedal & Wright.

10 Exercise 13.5 in Nocedal & Wright.

11 Find the dual of the linear optimisation problem

$$5x_1 + 3x_2 + 4x_3 \to \min \qquad \text{subject to} \quad \begin{cases} x_1 + x_2 + x_3 = 1, & \\ \qquad\qquad x_i \geq 0, & i = 1, 2, 3, \end{cases}$$

and compute its (i.e., the *dual's*) solution.

12 (*See exercise 16.1 in Nocedal & Wright.*) Consider the quadratic programme

$$f(x, y) := 2x + 3y + 4x^2 + 2xy + y^2 \to \min$$

subject to

$$x - y \geq 0, \qquad x + y \leq 4, \qquad x \leq 3.$$

a) Solve the quadratic programme and sketch its geometry (that is, the domain of the problem and the level lines of the function f).

b) What happens if one replaces the function f by $-f$? Does the problem still have solutions or local solutions?