Chalmers/GU
Mathematics

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	17–08–24
Time:	8^{30}–13^{30}
Aids:	Text memory-less calculator, English–Swedish dictionary
Number of questions:	7; passed on one question requires 2 points of 3.
	Questions are *not* numbered by difficulty.
	To pass requires 10 points and three passed questions.
Examiner:	Michael Patriksson
Teacher on duty:	Caroline Granfeldt, tel. 5325
Result announced:	17–09–14
	Short answers are also given at the end of
	the exam on the notice board for optimization
	in the MV building.

Exam instructions

When you answer the questions

Use generally valid theory and methods.
State your methodology carefully.

Only write on one page of each sheet. Do not use a red pen.
Do not answer more than one question per page.

At the end of the exam

Sort your solutions by the order of the questions.
Mark on the cover the questions you have answered.
Count the number of sheets you hand in and fill in the number on the cover.

Question 1

(simplex method)

The following linear optimization problem is given:

$$\begin{array}{lrcrcrl} \text{maximize} & z = 4x_1 & + & 2x_2 & + & 2x_3, & \\ \text{subject to} & x_1 & - & x_2 & + & 2x_3 & \leq 2, \\ & 2x_1 & + & x_2 & + & x_3 & \leq 8, \\ & x_1, & & x_2\,, & & x_3 & \geq 0. \end{array}$$

(2p) a) Solve the problem using phase I and phase II of the simplex method.

Aid: You may utilize the identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(1p) b) State the LP dual to the above problem and solve it graphically. Does it have an optimal solution?

(3p)

Question 2

(Lagrangian duality)

For a symmetric real matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ consider the problem to find

$$\text{minimum}_x\, f(\boldsymbol{x}) := -\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}, \tag{1}$$

subject to the constraint that $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} = 1$.

Derive the KKT conditions to this problem, and interpret the solution.

(3p) Question 3

(Newton's method)

An engineer has decided to verify numerically that the exponential function $x \to \exp(x) = e^x$ grows faster than any polynomial. In order to do so, he/she studies the optimization problem to

$$\underset{x\in\mathbb{R}}{\text{minimize}}\ f(x) = x^\alpha - e^x, \tag{1}$$

where α is the highest power of the polynomial (we assume it is an even, positive integer number). The engineer uses a Newton method (with unit steps!) to solve the problem. He/she argues that if the exponential function grows faster than any polynomial, then the sequence $\{x^k\}$ generated by the method should diverge to infinity, because the objective function f can be decreased indefinitely by increasing the value of x.

(1p) a) State the Newton iteration explicitly for the given problem (1).

(1p) b) Construct a numerical example (that is, choose a value of $\alpha \in \{2, 4, ...\}$ and a starting point of the Newton algorithm) illustrating the engineers error in reasoning.

(1p) c) Find the error in the engineer's reasoning and formally explain it.

(3p) Question 4

(unconstrained optimization)

Suppose that you have attacked the problem of minimizing a differentiable function f over $\mathbb{R}^n$. Explain as well as you can how you can measure, and motivate, whether or not a vector $\boldsymbol{x}$ is near-stationary.

(3p) Question 5

(modelling)

Suppose that people from two groups, G_1 and G_2, wish to pair up with each other. Each group contains $n \in N$ people, and thus the total number of pairings will be n as well. A pair has to consist of one person from each group.

All persons have been asked to rank people from the other group with numbers, where the higher number means a higher preferability to be paired up. For person $i \in G_1$, the ranking of person $j \in G_2$ is a_{ij}, and for person $j \in G_2$, the ranking of person $i \in G_1$ is b_{ji}. Thus, if person $i \in G_1$ and person $j \in G_2$ pair up, the sum of the rankings of these two persons is $a_{ij} + b_{ji}$. If, however, the individual ranking is a value below p, it means a person really don't like the other one and therefore refuses to be paired up with that person. In other words, these two persons can't pair up.

Formulate an integer linear model which pairs up people from both groups while maximizing the sum of the total rankings.

Question 6

(true or false)

The below three individual claims should be assessed individually. Are they true or false, or is it impossible to say? For each of the three statements, provide an answer, together with a short—but complete— motivation.

(1p) a) *Claim:* A convex quadratic function always has a minimum over $\mathbb{R}^n$.

(1p) b) *Claim:* A linear program always has a non-empty polyhedral set of optimal solutions.

(1p) c) *Claim:* The Lagrangian dual problem to any problem is a convex one.

Question 7

(a simple optimization problem)

In a recent optimization exam at a Swedish technical university, the following optimization problem was addressed:

$$\begin{aligned} \text{maximize} \quad & f(\boldsymbol{x}) := \sum_{j=1}^{n} a_j/x_j, \\ \text{subject to} \quad & \sum_{j=1}^{n} \log x_j \leq b, \\ & x_j > 0, \qquad j = 1, ..., n, \end{aligned}$$

where $a_j > 0$ for all j, and $b > 0$.

The students where asked to derive the optimal solution to this problem through a Lagrangian relaxation of the first constraint, and by then solving the resulting dual problem. Explain what is wrong with this exam question. In other words, prove that there does not exist an optimal solution to this problem.

[Hint: Utilize the KKT conditions.]

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 17–08–24
Examiner: Michael Patriksson

Question 1

(simplex method)

(2p) a) Rewrite the problem into standard form by adding slack variables s_1 and s_1 to the left-hand side in the first and second constraint, respectively. Thus, we get the following linear program:

$$\begin{array}{lrcrcrcrcrcl}
\text{maximize} & z = 4x_1 & + & 2x_2 & + & 2x_3, & & & & & & \\
\text{subject to} & x_1 & - & x_2 & + & 2x_3 & + & s_1 & & & = & 2, \\
 & 2x_1 & + & x_2 & + & x_3 & & & + & s_2 & = & 8, \\
 & x_1, & & x_2\,, & & x_3, & & s_1, & & s_2 & \geq & 0.
\end{array}$$

We start directly with phase II at the origin, using the starting basis $(s_1, s_2)^T$. This iteration,

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 2 \\ 8 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}.$$

The reduced costs, $\bar{\boldsymbol{c}}^T = \boldsymbol{c}_N^T - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{N}$, for this basis is $\bar{\boldsymbol{c}}^T = \begin{pmatrix} 4 & 2 & 2 \end{pmatrix}$, which means that x_1 enters the basis. The minimum ratio test implies that s_1 leaves.

Updating the basis, we now have $(x_1, s_2)^T$ in the basis and

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}^T = \begin{pmatrix} -4 & 6 & -6 \end{pmatrix}$ which means that x_2 enters the basis. The minimum ratio test implies that s_2 leaves.

Once again updating the basis, now with $(x_1, x_2)^T$, gives

$$\boldsymbol{B} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}, \boldsymbol{N} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \boldsymbol{x}_B = \begin{pmatrix} 3\frac{1}{3} \\ 1\frac{1}{3} \end{pmatrix}, \boldsymbol{c}_B = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \boldsymbol{c}_N = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

The new reduced costs are $\bar{\boldsymbol{c}}^T = \begin{pmatrix} 0 & -2 & 0 \end{pmatrix}$ which means that the current basis is optimal. The optimal solution is thus
$\boldsymbol{x}^* = \begin{pmatrix} x_1 & x_2 & x_3 & s_1 & s_2 \end{pmatrix}^T = \begin{pmatrix} 3\frac{1}{3} & 1\frac{1}{3} & 0 & 0 & 0 \end{pmatrix}^T$
with optimal objective function value $z^* = 16$.

(1p) b) The LP dual is

$$\begin{array}{lrrrrl} \text{minimize} & w = & 2y_1 & + & 8y_2, & \\ \text{subject to} & & y_1 & + & 2y_2 & \geq 4, \\ & & -y_1 & + & y_2 & \geq 2, \\ & & 2y_1 & + & y_2 & \geq 2, \\ & & y_1, & & y_2 & \geq 0. \end{array}$$

Drawing this problem, it easy to see that the optimal solution is $\boldsymbol{y}^* = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ with optimal objective function value $w^* = 16$.

(Note: Since there exists an optimal solution to the primal problem, strong duality actually implies that the dual problem also has an optimal solution.)

(3p) Question 2

(Lagrangian duality)

Introducing the scalar $\lambda \in \mathbb{R}$, the KKT conditions state that

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \qquad \text{and} \qquad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = 1.$$

The interpretation is that the (primal) solutions in $\boldsymbol{x}$ are eigenvectors of the matrix $\boldsymbol{A}$, and that the (dual) variable λ is an eigenvalue.

See also Example 5.54 from the course book.

(3p) Question 3

(Newton's method)

(1p) a) Newton's equation:

$$x_{k+1} = x_k - \frac{\alpha {x_k}^{\alpha-1} - e^{x_k}}{\alpha(\alpha-1){x_k}^{\alpha-2} - e^{x_k}}$$

(1p) b) For example, take $x_0 = 1$, $\alpha = 4$. Then we can get $x_1 = 0.8619$, $x_2 = 0.8323$, $x_3 = 0.8310$, $x_4 = 0.8310$, ... These initial values cause the Newtons method to generate a sequence stuck at $x = 0.8310$ which is the local minimum.

(1p) c) The objective function of the problem is not convex in general [may be verified by analyzing the sign of the Hessian $\alpha(\alpha-1)x^{\alpha-2} - e^x$]. Since the convergence of the Newton method is local in nature, the method is most likely to converge to the nearest local minimum (or maximum if the Hessian is negative definite). The engineer thus wrongly assumes the global convergence of the Newton method on non convex functions.

Question 4

(unconstrained optimization)

The main problem in this question lies in the fact that we need to cope with the fact that the value of $\nabla f(\boldsymbol{x})$ needs to be exactly zero in order to conclude that $\boldsymbol{x}$ is stationary. Hence the exact measure

$$\nabla f(\boldsymbol{x}) = \mathbf{0}^n$$

needs to be replaced by a sensible tolerance. The course book include, in Section 11.5, a list of three combinations of criteria, based on a small norm of the gradient of f, a small decrease in the value of f between two iterates in relation to the size of problem data, and a small shift in the vector $\boldsymbol{x}$ between iterations.

(3p) Question 5

(modelling)

The decision variables are:

$$x_{i,j} = \begin{cases} 1 & \text{if person } i \in G_1 \text{ and person } j \in G_2 \text{ pair up,} \\ 0 & \text{otherwise.} \end{cases}$$

Model:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i \in G_1} \sum_{j \in G_2} (a_{ij} + b_{ji}) x_{ij}, \\
\text{subject to} \quad & \sum_{i \in G_1} x_{ij} = 1 && j \in G_2, \\
& \sum_{j \in G_2} x_{ij} = 1 && i \in G_1, \\
& (a_{ij} - p) x_{ij} \geq 0 && i \in G_1, j \in G_2, \\
& (b_{ji} - p) x_{ij} \geq 0 && i \in G_1, j \in G_2, \\
& x_{i,j} \in \{0, 1\} && i \in G_1, j \in G_2, .
\end{aligned}$$

Question 6

(true or false)

(1p) a) The claim is false in general. (However, if f is lower bounded, then it has a minimum.)

Example: Let $\boldsymbol{A} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

The quadratic form $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{x}$ has no minimum.

(1p) b) The claim is false, as a linear program may have no feasible solutions.

(1p) c) The claim is true, and is established in Theorem 6.4 in the textbook.

Question 7

(a simple optimization problem)

The KKT conditions for this problem amount, apart from complementary and primal feasibility, to finding a solution in the pair $(\boldsymbol{x}, \mu)^T \in R^n \times R_+$ to the nonlinear equations formed by the stationarity conditions for the Lagrangian with respect to $\boldsymbol{x}$, that is, for all $j = 1, ..., n$,

$$\frac{a_j}{x_j^2} + \frac{\mu}{x_j} = 0$$

This is clearly impossible, as $x_j > 0$ must be fulfilled, and $a_j > 0$ holds. We therefore conclude that there are no KKT points for this problem.

Can there be optimal solutions that are not KKT points? No, because the linear independence CQ (LICQ) is fulfilled for this problem, so the KKT conditions are necessary conditions for both local and global optimal solutions.