Chalmers/GU
Mathematics

EXAM

TMA947/MMG621
NONLINEAR OPTIMISATION

Date:	17–04–12
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:	Olof Elias, tel. 5325
Result announced:	17–05–03 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

(linear programming)

The following linear optimization problem is given:

$$\begin{array}{llrrrrl} \text{minimize} & z = & 5x_1 & +\,40x_2 & +\,4x_3 & -\,x_4, & \\ \text{subject to} & & -\frac{1}{2}x_1 & +\;\;3x_2 & +\;\;x_3 & -\;\;x_4 & \geq 6, \\ & & x_1 & +\;\;4x_2 & -\;\;x_3 & -\;\;x_4 & \geq 7, \\ & & x_1, & x_2, & x_3, & x_4 & \geq 0. \end{array}$$

(2p) a) Show that x_2, x_3 are optimal basic variables. (*Note:* The simplex method need not be used.)

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) Assume that the right-hand side of both linear constraints decrease by $\delta \geq 0$. For what range of values of δ are x_2 and x_3 still optimal basic variables?

(3p) Question 2

(the Separation Theorem)

Given a closed and convex set $C \subset \mathbb{R}^n$ and a vector $\boldsymbol{y} \in \mathbb{R}^n$ which does not belong to C, the Separation Theorem states a result on the existence of a separating hyperplane. State the Separation Theorem precisely, and establish its correctness with a complete proof.

(3p) Question 3

(Lagrangian duality)

Given is the following problem over $\boldsymbol{x} \in \mathbb{R}^n$:

$$\begin{aligned}
\text{minimize} \quad & f(\boldsymbol{x}), \\
\text{subject to} \quad & g_1(\boldsymbol{x}) \leq 0, && (1)\\
& g_2(\boldsymbol{x}) \leq 0, && (2)\\
& \boldsymbol{x} \in S,
\end{aligned}$$

where $f,\ g_1,\ g_2 : \mathbb{R}^n \to \mathbb{R}$, and $S \subseteq \mathbb{R}^n$.

Constraints (1) and (2) are Lagrangian relaxed using multipliers $\mu_1 \geq 0$ and $\mu_2 \geq 0$, respectively.

A heuristic was used to try and find feasible solutions of $\boldsymbol{x}$ by making suitable adjustments of the multipliers $\boldsymbol{\mu}$. The table below shows numerical results for a number of different values of the multipliers, given in the order they were examined. The point $\boldsymbol{x}^k$ is the optimal value from the Lagrangian problem in $\boldsymbol{x}$ using the multipliers $(\mu_1, \mu_2)^k$.

k	$(\mu_1, \mu_2)^k$	$\boldsymbol{x}^k$	$f(\boldsymbol{x}^k)$	$g_1(\boldsymbol{x}^k)$	$g_2(\boldsymbol{x}^k)$
1	(0, 0)	$\boldsymbol{x}^1$	−3.0	8.0	12.0
2	(3, 3)	$\boldsymbol{x}^2$	1.0	−3.0	5.0
3	(1.5, 6)	$\boldsymbol{x}^3$	9.0	2.0	−1.0
4	(2.25, 4.5)	$\boldsymbol{x}^4$	12.0	−1.0	−0.5
5	(2, 3.75)	$\boldsymbol{x}^5$	8.0	0.0	1.0
6	(2.16, 4)	$\boldsymbol{x}^6$	12.25	−0.25	−0.25

Use the information in the table to give the best possible estimate of the optimal objective function value of the given problem, i.e. the smallest interval for f^*.

(3p) Question 4

(modelling)

In the labyrinth puzzle game there are usually several possible routes, and people need to pick the route that gives the highest probability to reach the required end point. Let's simplify this into a mathematical modelling problem. Figure 1 shows a grid of dimension $I \times J$. The starting point is $(1, 1)$, and the end point is (I, J). At each point, it is possible to move to one of the (at most four) adjacent points. Thus, a move is defined as rolling the ball from one point to another adjacent point. Let p be the probability of failure during a move. Each move then has three possible scenarios:

1. The passage contains a hole, where the ball could fall down with the probability $p \in [0, 1]$. Falling down indicates failure at the game.
2. The passage is blocked by a wall, which makes it impossible to pass. Trying to pass through this means failure at the game with a probability $p = 1$.
3. The passage contains no obstacle, which gives the probability of failing $p = 0$.

Thus, when a move is done from a point (i_1, j_1) to an adjacent point (i_2, j_2), there is a probability $p_{(i_1,j_1)(i_2,j_2)}$ of failing which will end the game. Moreover, performing moves is tiresome for the player and there is therefore an upper limit of performing S moves. Furthermore assume that a move can be made (at most) once, where a move is defined as moving from one point to another adjacent point.

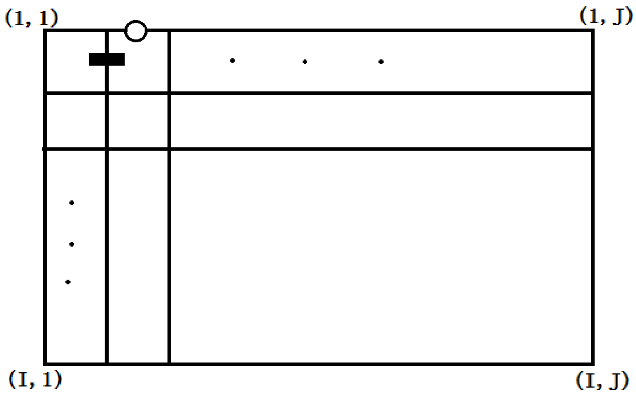

Figure 1: A simplified model of the labyrinth puzzle game

Formulate an integer program to determine the route which maximizes the probability of succeeding at reaching the end point.

Question 5

(necessary local and sufficient global optimality conditions)

Consider an optimization problem of the following general form:

$$\text{minimize } f(\boldsymbol{x}), \tag{1a}$$
$$\text{subject to } \boldsymbol{x} \in S, \tag{1b}$$

where $S \subseteq \mathbb{R}^n$ is nonempty, closed and convex, and $f : \mathbb{R}^n \to \mathbb{R}$ is in C^1 on S.

(1p) a) Establish the following result on the local optimality of a vector $\boldsymbol{x}^* \in S$ in this problem.

Proposition 1 (necessary optimality conditions, C^1 case) *If $\boldsymbol{x}^* \in S$ is a local minimum of f over S then*

$$\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0, \qquad \boldsymbol{x} \in S \tag{2}$$

holds.

(2p) b) Establish the following result on the global optimality of a vector $\boldsymbol{x}^* \in S$ in this problem.

Theorem 2 (necessary and sufficient global optimality conditions, C^1 case) *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ further is convex on S. Then,*

$$\boldsymbol{x}^* \text{ is a global minimum of } f \text{ over } S \iff (2) \text{ holds.}$$

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) Suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a vector $\boldsymbol{x} \in \mathbb{R}^n$.

Claim: for the vector $\boldsymbol{p} \in \mathbb{R}^n$ to be a descent direction with respect to f at $\boldsymbol{x}$ it is necessary that $\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{p} < 0$.

(1p) b) Consider solving a linear program (call it "P") through the process of utilizing "phase I" and "phase II" of the Simplex method. Suppose that the optimal value in the phase I-problem is zero.

Claim: There exists an optimal solution to the linear program P.

(1p) c) *Claim:* If the function $g : \mathbb{R}^n \to \mathbb{R}$ is *concave* on $\mathbb{R}^n$ and $c \in \mathbb{R}$, then the set $\{\, \boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) \leq c \,\}$ is convex.

(3p) Question 7

(the Karush–Kuhn–Tucker conditions)

Consider the problem to:

$$
\begin{aligned}
\text{maximize} \quad & f(\boldsymbol{x}) := x_1 - x_1^2, \\
\text{subject to} \quad & x_1 \geq 2, \\
& x_2 - (x_1 - 3)^2 \geq -2, \\
& x_1 - x_2 \geq 1.
\end{aligned}
$$

(2p) a) Express the KKT conditions, and find all KKT points.

(1p) b) Are the KKT points optimal? Motivate!

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 17–04–12
Examiner: Michael Patriksson

Question 1

(linear programming)

(2p) a) Rewrite the problem into standard form by subtracting slack variables x_5 and x_6 from the left-hand side in the first and second constraint, respectively. If x_2 and x_3 are basic variables, the basic solution is

$$\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 6 \\ 7 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 13 \\ 3 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and thus the basic solution is *feasible*.

Now we can check the reduced costs $\bar{\boldsymbol{c}}^T = \boldsymbol{c}_N^T - \boldsymbol{y}^T \boldsymbol{N}$, where $\boldsymbol{y} = \boldsymbol{c}_B^T \boldsymbol{B}^{-1} = \begin{pmatrix} 40 & 4 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 4 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, for the non-basic variables:

$$\bar{c}_1 = 5 - \begin{pmatrix} 8 & 4 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix} = 5 \geq 0,$$

$$\bar{c}_4 = -1 - \begin{pmatrix} 8 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = 11 \geq 0,$$

$$\bar{c}_5 = 0 - \begin{pmatrix} 8 & 4 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} = 8 \geq 0,$$

$$\bar{c}_6 = 0 - \begin{pmatrix} 8 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = 4 \geq 0.$$

All reduced costs are non-negative, and thus the basis is optimal.

(It is also possible to show this using LP duality and complementary slackness conditions.)

(1p) b) The dual solution and the reduced costs are not affected by a small enough perturbation in the right-hand side, and it is therefore enough to study how feasibility is affected.

Basic solution as a function of δ:

$$\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \boldsymbol{B}^{-1}(\boldsymbol{b} - \delta) = \begin{pmatrix} 3 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 6 - \delta \\ 7 - \delta \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 13 \\ 3 \end{pmatrix} - \frac{1}{7} \begin{pmatrix} 2\delta \\ \delta \end{pmatrix}.$$

Constraints on $\delta \geq 0$ for feasibility:

$13 - 2\delta \geq 0 \implies \delta \leq \frac{13}{2}$,
$3 - \delta \geq 0 \implies \delta \leq 3$.

Thus, x_2 and x_3 are optimal basic variables if $0 \leq \delta \leq 3$.

(3p) Question 2

(the Separation Theorem)

See Theorem 4.29 in the course book.

(3p) Question 3

(Lagrangian duality)

Dual problem:

$$\begin{aligned} q^* = \quad & \max_{\boldsymbol{\mu} \geq \mathbf{0}} \quad q(\boldsymbol{\mu}), \\ & \text{where} \quad q(\boldsymbol{\mu}) = \min_{\boldsymbol{x} \in S} f(\boldsymbol{x}) + \mu_1 g_1(\boldsymbol{x}) + \mu_2 g_2(\boldsymbol{x}). \end{aligned}$$

Since the optimal solution to the dual problem is given in the table, it is easy to calculate the dual function $q(\boldsymbol{\mu}^k) = f(\boldsymbol{x}^k) + \mu_1^k g_1(\boldsymbol{x}^k) + \mu_2^k g_2(\boldsymbol{x}^k)$.

Thus, the following calculations can be done:

$q(\boldsymbol{\mu}^1) = -3.0 + 0 \cdot 8.0 + 0 \cdot 12.0 = -3.0,$
$q(\boldsymbol{\mu}^2) = 1.0 - 3 \cdot 3.0 + 3 \cdot 5.0 = 7.0,$
$q(\boldsymbol{\mu}^2) = 9.0 + 1.5 \cdot 2.0 - 6 \cdot 1.0 = 6.0,$
$q(\boldsymbol{\mu}^4) = 12.0 - 2.25 \cdot 1.0 - 4.5 \cdot 0.5 = 7.5,$
$q(\boldsymbol{\mu}^5) = 8.0 + 2 \cdot 0.0 + 3.75 \cdot 1.0 = 11.75,$
$q(\boldsymbol{\mu}^6) = 12.25 - 2.16 \cdot 0.25 - 4 \cdot 0.25 = 10.71.$

Each $q(\boldsymbol{\mu}^k)$ gives an optimistic estimation of the optimal objective function value, f^*. Thus, the best optimistic estimation is $f^* \geq 11.75$.

Every *feasible solution* gives a pessimistic estimation of f^*:

$\boldsymbol{x}^4$ feasible $\implies f^* \leq 12,$
$\boldsymbol{x}^6$ feasible $\implies f^* \leq 12.25.$

Thus, $f^* \leq 12$.

Therefore, the best possible estimation is $11.75 \leq f^* \leq 12$.

(3p) Question 4

(modelling)

To simplify the notations, we change the two dimensions notations into one dimension. So change point (i, j) to $(i-1)\cdot J + j$, and $p_{(i_1,j_1)(i_2,j_2)}$ changes to $p_{(i_1-1)\cdot J+j_1,(i_2-1)\cdot J+j_2}$.

Sets:
$\mathcal{M} := \{i | i \in \{1, ..., I\cdot J\}\}$, the set of possible points,
$\mathcal{N} := \{(i, j)|$ all pairs of points (i, j) where $i \in \mathcal{M}$ is an adjacent point of $j \in \mathcal{M}\}$.

The decision variables are:

$$x_{i,j} = \begin{cases} 1 & \text{part of the optimal route goes from } i \text{ to } j, \\ 0 & \text{otherwise,} \end{cases}$$

where $\{i, j\} \in \mathcal{N}$.

Model:

$$\begin{aligned}
\text{maximize} \quad & \prod_{(i,j)\in N} (1 - p_{i,j}x_{i,j}), \\
\text{subject to} \quad & \sum_{j|(i,j)\in N} x_{i,j} = \sum_{k|(k,i)\in N} x_{k,i} && i \in \mathcal{M}\backslash\{1, I\cdot J\}, \\
& \sum_{j|(1,j)\in N} x_{1,j} = \sum_{k|(k,1)\in N} x_{k,1} + 1, \\
& \sum_{j|(I\cdot J,j)\in N} x_{I\cdot J,j} = \sum_{k|(k,I\cdot J)\in N} x_{k,I\cdot J} - 1, \\
& \sum_{(i,j)\in N} x_{i,j} \le S, \\
& x_{i,j} \in \{0, 1\} && (i, j) \in \mathcal{N}.
\end{aligned}$$

Question 5

(necessary local and sufficient global optimality conditions)

(1p) a) See Proposition 4.22 in course book.

(2p) b) See Theorem 4.23 in the course book.

Question 6

(true or false)

(1p) a) False. Let $f(x) = -x^2$. At the point $\bar{x} = 0$, all feasible directions $p \neq 0$ are descent directions. However, $f'(\bar{x}) = 0$ and thus $f'(\bar{x})p = 0$. Therefore, the claim is false.

(It is however sufficient, i.e. if $\nabla f(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{p} < 0$, then $\boldsymbol{p}$ is a descent direction with respect to f at $\boldsymbol{x}$.)

(1p) b) False. The problem is feasible but may have an unbounded solution.

(1p) c) False. Consider the function g where $g(x) = 4 - x^2$ and the two points $x^1 = -2$ and $x^2 = 3$ which belong to the set $S = \{x \in \mathbb{R} \mid g(x) \leq 0\}$. By Definitions 3.39 and 3.40, g is concave. However, the point $\frac{1}{2}x^1 + \frac{1}{2}x^2 = \frac{1}{2} \notin S$. Hence, by Definition 3.1, the set S is not convex.

Question 7

(the Karush–Kuhn–Tucker conditions)

(2p) a) First, rewrite the problem to the following form:

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) := x_1^2 - x_1, \\ \text{subject to} \quad & 2 - x_1 \leq 0, \\ & (x_1-3)^2 - x_2 - 2 \leq 0, \\ & 1 - x_1 + x_2 \leq 0. \end{aligned}$$

Let:

$g_1(\boldsymbol{x}) = 2 - x_1,$
$g_2(\boldsymbol{x}) = (x_1-3)^2 - x_2,$
$g_3(\boldsymbol{x}) = 1 - x_1 + x_2.$

The KKT conditions are:

$$\nabla f(\boldsymbol{x}) + \sum_{i=1}^{3} \mu_i \nabla g_i(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 - 6 \\ -1 \end{pmatrix} + \mu_3 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$\mu_1, \mu_2, \mu_3 \geq 0,$

$\mu_i g_i(\boldsymbol{x}) = 0, \quad i = 1,2,3,$

$g_i(\boldsymbol{x}) \leq 0, \quad i = 1,2,3.$

The following cases of active constraints are possible:

- Let g_1 be active. Solving the KKT conditions gives $x_1 = 2$, $-1 < x_2 < 1$, $\mu_1 = 3$, $\mu_2 = 0$, and $\mu_3 = 0$.
- Let g_1 and g_2 be active. Solving the KKT conditions gives $x_1 = 2$, $x_2 = -1$, $\mu_1 = 3$, $\mu_2 = 0$, $\mu_3 = 0$.
- Let g_2 be active. The KKT conditions do not give any points.
- Let g_2 and g_3 be active. The KKT conditions do not give any points.
- Let g_3 be active. The KKT conditions do not give any points.
- Let $g_!$ and g_3 be active. Solving the KKT conditions gives $x_1 = 2$, $x_2 = 1$, $\mu_1 = 3$, $\mu_2 = 0$, $\mu_3 = 0$.
- Let no constraints be active. The KKT conditions do not give any points.

Thus, the feasible points fulfilling the KKT conditions are $\boldsymbol{x} = \begin{pmatrix} 2 \\ a \end{pmatrix}$, where $-1 \leq a \leq 1$.

(1p) b) The objective function f and the constraint functions g_i are convex. Therefore the KKT conditions are sufficient for global optimality, and thus all KKT points are optimal.