Chalmers/Gothenburg University
Mathematical Sciences

**EXAM SOLUTION**

# TMA947/MMG621
# NONLINEAR OPTIMISATION

**Date:** 18–01–09
**Examiner:** Michael Patriksson

Note: the solutions presented here are brief in relation to the requirements on your answers, in particular regarding your motivations.

## Question 1

(the simplex method)

**(2p)** a) Rewrite the problem into standard form by adding/subtracting slack variables $s_1$ and $s_2$ to the left-hand side in the first and second constraint, respectively. Moreover, let $z := -z$ to get the problem on minimization form. Thus, we get the following linear program:

$$\begin{array}{llcccccl} \text{minimize} & z = & -x_1 & - & 2x_2, & & & \\ \text{subject to} & & -x_1 & - & x_2 & + s_1 & & = 1, \\ & & x_1 & - & x_2 & & - s_2 & = 1, \\ & & x_1, & & x_2, & s_1, & s_2 & \geq 0. \end{array}$$

Introducing the artificial variable $a$, phase I gives the problem

$$\begin{array}{llcccccl} \text{minimize} & w = & a, & & & & & \\ \text{subject to} & & -x_1 & - x_2 & + s_1 & & & = 1, \\ & & x_1 & - x_2 & & - s_2 & + a & = 1, \\ & & x_1, & x_2, & s_1, & s_2, & a & \geq 0. \end{array}$$

Using the starting basis $(s_1, a)^T$ gives

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

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

Thus, we move on to phase II using the basis $(s_1, x_1)^T$, and

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

The new reduced costs are $\bar{\boldsymbol{c}}_N^T = \begin{pmatrix} -3 & -1 \end{pmatrix}$

which means that $x_2$ enters the basis. From the minimum ratio test we get $\boldsymbol{B}^{-1}\boldsymbol{N}_1 = \begin{pmatrix} -2 & -1 \end{pmatrix}^T < \boldsymbol{0}$, meaning that the problem is unbounded.

**(1p)** b) A direction of unboundedness is $\boldsymbol{l}(\mu) = \begin{pmatrix}1 & 0 & 2 & 0\end{pmatrix}^T + \mu \begin{pmatrix}1 & 1 & 2 & 0\end{pmatrix}^T$, $\mu \geq 0$.

---

## (3p) Question 2

(gradient projection)

The gradient of $f$ at the point $\boldsymbol{x}_0 = (0,2)^T$ is $\nabla f(\boldsymbol{x}_0) = (2,8)^T$. Taking a step in the negative gradient direction with $\alpha = 1/8$ gives the new point $\boldsymbol{x}_0 - (1/8)(2,8)^T = (-1/4, 1)$.

Projecting this point to the feasible set yields the new iterate $\boldsymbol{x}_1 = (0,1)$.

This point is clearly neither a local nor a global minimum. To check this, perform another iteration and see that the new iterate is not the same as $\boldsymbol{x}_1$.

---

## (3p) Question 3

(optimality conditions for special feasible sets)

Thanks to the linearity of the constraints, the problem satisfies the Abadie constraint qualification and the Karush–Kuhn–Tucker conditions are necessary for the local optimality of $\boldsymbol{x}$. Introducing the multiplier $\mu$ for the equality constraint and $\lambda_j$ for the sign constraints on $x_j$ we obtain the Lagrangian function $L(\boldsymbol{x}, \mu, \boldsymbol{\lambda}) := b\mu + \sum_{j=1}^n (f_j(x_j) + [\mu - \lambda_j]x_j)$. Assume that $(\boldsymbol{x}^*, \mu^*, \boldsymbol{\lambda}^*)$ is a KKT point. Setting the partial derivatives of $L$ to zero yields

$$\phi_j'(x_j^*) = \lambda_j^* - \mu^*, \qquad j = 1, \dots, n, \tag{1}$$

and further, from complementarity, that

$$\lambda_j^* x_j^* = 0, \qquad j = 1, \dots, n.$$

For a $j$ with $x_j^* > 0$ it must then hold that $\phi_j'(x_j^*) = -\mu^*$. Suppose instead that $x_j^* = 0$. Then since $\lambda_j^* \geq 0$ must hold, we find, from the characterization (1), that $\phi_j'(x_j^*) = \lambda_j^* - \mu^* \geq -\mu^*$, and we are done.

---

## Question 4

(Karush–Kuhn–Tucker)

**(2p)** a) Let $g_1(\boldsymbol{x}) := -x_1^2 - x_2^2 + 25$, $g_2(\boldsymbol{x}) := x_1 - 4$ , $g_3(\boldsymbol{x}) := x_2 - 4$, $g_4(\boldsymbol{x}) := -x_1$ and $g_5(\boldsymbol{x}) := -x_2$with respective gradients $\nabla g_1 = (-2x_1, -2x_2)^T$, $\nabla g_2 = (1,0)^T$ ,$\nabla g_3 = (0,1)^T$ , $\nabla g_4 = (-1,0)^T$ and $\nabla g_5 = (0,-1)^T$.
Moreover, $\nabla f = (-2x_1 + 2, 0)^T$. The KKT-conditions are as follows:

$$\nabla f(x^*) + \sum_{i=1}^{5} \mu_i \nabla g_i(x^*) = 0,$$

$$\mu_i g_i(x^*) = 0, i = 1, ..., 5,$$

$$\mu_i \geq 0, i = 1, ..., 5.$$

Since the objective function $f$ is not convex, the KKT conditions are not sufficient.

To prove KKT conditions are necessary, we use LICQ. For the interior points, there is no active constraints, and for the points on the boundary but not extreme points, there is only one active constraint, so the gradient of the active constraint must be linearly independent. So we only need to check three extreme points: $(4,3)^T, (3,4)^T, (4,4)^T$. For point $(4,3)^T$, the gradient of the active constraints are $(-8,-6)^T$ and $(1,0)^T$, obviously they are linearly independent. For point $(3,4)^T$, the gradient of the active constraints are $(-6,-8)^T$ and $(1,0)^T$, obviously they are linearly independent. For point $(4,4)^T$, the gradient of the active constraints are $(0,1)^T$ and $(1,0)^T$, obviously they are linearly independent. So LICQ holds at every feasible point. Thus, the KKT-conditions are necessary.

**(1p)** b) By letting different combinations of constraints be active, we can see when only $g_2$ active, we get $(4,a), 3 < a < 4$ are KKT points. When $g_1$ and $g_2$ are active, we get $(4,3)$ is a KKT point. When $g_2$ and $g_3$ are active, we get $(4,4)$ is a KKT point. So $(4,a), 3 \leq a \leq 4$ are KKT points. Since KKT conditions are necessary, so the optimal solution must be KKT points. Since all KKT points give the same objective function value $-8$, so all the KKT points are optimal.

## (3p) Question 5

(modelling)

Let $S_1, S_2, \ldots, S_m$ be the sets, and let $U = \{1, \ldots n\}$ be the universe to cover. Now let the binary parameters $s_{ij} = 1$ if the element $j$ is in the set $S_i$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, and $s_{ij} = 0$ otherwise. Let $w_i$ be the weight of set $S_i$.

Let $x_i$ be a binary variable where $x_i = 1$ if set $S_i$ is included in the sub-collection, where $i \in \{1, \ldots, m\}$, and $x_i = 0$ otherwise. The weighted set covering problem can now be formulated as:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{m} w_i s_i, \\
\text{subject to} \quad & \sum_{i=1}^{m} s_{ij} x_i \geq 1, \quad j \in \{1, \ldots, n\}, \\
& x_{ij} \in \{0, 1\}
\end{aligned}
$$

---

## Question 6

(true or false)

**(1p)** a) False. Consider $f(x) = x^3$ at $x = 0$; a negative direction from 0 clearly reduces the value of $f$, while $f'(0) = 0$.

**(1p)** b) True. The claim is a characterization of the line search being exact in the direction of the vector $\boldsymbol{x}^{t+1} - \boldsymbol{x}^t$.

**(1p)** c) False. The solution set of the two linear inequalities $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \geq b$ and $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leq b$, defines, by definition, a polyhedron, as it is the solution set of a collection of linear inequalities. On the other hand, the solution set also is a line segment in $\mathbb{R}^n$.

---

## (3p) Question 7

(Farkas' lemma)

See the course book.